Minimax Rates for Estimating the Dimension of a Manifold

Jisu Kim, Alessandro Rinaldo and Larry Wasserman
Carnegie Mellon University
January 4 2016

Many algorithms in machine learning and computational geometry require, as input, the intrinsic dimension of the manifold that supports the probability distribution of the data. This parameter is rarely known and therefore has to be estimated from the data. We characterize the statistical difficulty of this problem. Specifically, we derive upper and lower bounds on the minimax rate for estimating the dimension. First we consider the problem of testing the hypothesis that the support of the data-generating probability distribution is a well-behaved manifold of intrinsic dimension d_1 versus the alternative that it is of different dimension d_2. With an i.i.d. sample of size n, we provide an upper bound on the sum of typeI and II errors of order $O\left(n^{-(d_2/d_1-1-\epsilon)n}\right)$ based on the travelling salesman path through the data points, where ϵ is an arbitrarily small positive number. We also demonstrate a lower bound of $\Omega\left(n^{-(2d_2-2d_1+\epsilon)n}\right)$, for any $\epsilon > 0$, by applying Le Cam's lemma with a specific set of d_1-dimensional probability distributions. We then extend these results to get minimax rates for estimating the dimension of a well-behaved manifolds. We obtain an upper bound of order $O\left(n^{-(\frac{1}{m-1}-\epsilon)n}\right)$ and a lower bound of order $\Omega\left(n^{-(2+\epsilon)n}\right)$, where m is the embedding dimension.

1. Introduction

Suppose that $X_1, \ldots, X_n$ is an i.i.d. sample from a distribution P whose support is an unknown manifold M of dimension d in $\mathbb{R}^m$, where $1 \leq d \leq m$. Manifold learning refers broadly to a suite of techniques from statistics and machine learning aimed at estimating M or some of its features based on the sample.

Manifold learning methods are widely used in high dimensional data analysis mainly to alleviate the curse of dimensionality. Indeed, manifold learning algorithms typically map the data to a new, lower dimensional coordinate system [Bellman, 1961, Lee and Verleysen, 2007a, Hastie et al., 2009]. By using such a mapping, manifold learning can greatly reduce the dimensionality of the data with little loss in accuracy..

Most manifold learning algorithms require, as input, the intrinsic dimension of the manifold. However, such quantity is almost never known in advance and therefore has to be estimated.

Various intrinsic dimension estimators have been proposed and analyzed; see, e.g., Lee and Verleysen [2007b], Koltchinskii [2000], Kégl [2003], Levina and Bickel [2004], Hein and Audibert [2005], Raginsky and Lazebnik [2005], Little et al. [2009, 2011], Sricharan et al. [2010], Rozza et al. [2012], Camastra and Staiano [2015]. However, characterizing the intrinsic statistical hardness of the dimension estimation problem remains an open problem. The traditional way of measuring the difficulty of a statistical problem is to bound its *minimax risk,* which is in the present setting is defined as the worst possible statistical performance of a best dimension estimator. Formally, given a class of probability distribution $\mathcal{P}$, the minimax risk $R_n = R_n(\mathcal{P})$ is defined as

$$R_n = \inf_{\widehat{d}} \sup_{P \in \mathcal{P}} \mathbb{E}_P[I(\widehat{d} \neq d(P))]. \tag{1.1}$$

Here $d(P)$ is the dimension of the support of P, $\mathbb{E}_P$ denotes the expectation with respect to the distribution P, $I(\cdot)$ is the indicator function and the infimum is over all estimators $\widehat{d} = \widehat{d}(X_1, \ldots, X_n)$ (measurable functions of the data) of the dimension. Notice in particular that the statistical performance or risk $\mathbb{E}_P[I(\widehat{d} \neq d(P))]$ of a dimension estimator $\widehat{d}$ is the probability that $\widehat{d}$ differs from the true dimension $d(P)$ of the support of the data generating distribution. The minimax risk, a function of both the sample size n and the class $\mathcal{P}$, quantifies the intrinsic hardness of the dimension estimation problem in the sense that *any dimension estimator* cannot have a risk smaller than R_n uniformly over all $P \in \mathcal{P}$. The purpose of this paper is to obtain upper and lower bounds on the minimax risk.

We start by assuming that the manifold supporting the data generating distribution P has two possible dimensions, d_1 and d_2. This assumption is then relaxed to any dimension $d(P)$ between 1 and the embedding dimension m in Section 5. We will impose several regularity conditions on the supporting manifold in order to make the problem analytically tractable and also to avoid intractable or trivial cases, such as space filling manifolds. Our main result may be summarized as follows. Let $X_1, \ldots, X_n \sim P$ be i.i.d., where P belongs to a class $\mathcal{P}$ of probability distributions supported on well-behaved manifolds in $\mathbb{R}^m$, as defined in Section 2.

Theorem 1. *The minimax risk R_n satisfies, $a_n \leq R_n \leq b_n$ where*

$$a_n = (C_{K_I}^{(5,2)})^n \kappa_\ell^{-n} \min\{\kappa_\ell^3 n^{-2}, 1\}^n$$
$$b_n = (C_{K_I,K_p,K_v,K_m}^{(5,1)})^n (1 + \kappa_g^{(m^2-m)n}) n^{-\frac{n}{m-1}}$$

where the constants κ_ℓ, κ_g, $C_{K_I}^{(5,2)}$ and and $C_{K_I,K_p,K_v,K_m}^{(5,1)}$ depends on $\mathcal{P}$ and are defined in Section 5.

We now make a few remarks about the result.

- First, as the dimension is a discrete quantity, the rates are exponential in sample size, a finding consistent with the results obtained by Koltchinskii [2000].
- The key constants that appear in the bounds depend on the local curvature κ_ℓ and the global curvature κ_g of the manifold, which are defined in Section 2. These curvature parameters affect the performance of any dimension estimator: a manifold with high curvature may appear more space filling than a manifold of the same dimension but with low curvature, thus making the task of resolving the dimension harder. Indeed, our analysis shows formally that the minimax risk is increasing in the values of the curvatures. Given their crucial role, we have made the dependence of the minimax risk on the curvatures as explicit as possible.

- Finally, there is a gap between the lower and upper bound, as the two rates do not match. Nonetheless, as far as we are aware, these are the most precise bounds on R_n that are available.

This paper is organized as follows. In Section 2, regularity conditions on distributions and their supporting manifolds are discussed. In Section 3, we give an upper bound on the minimax rate, by considering TSP path. In Section 4, we give a lower bound on the minimax rate by applying Le Cam's lemma with a specific set of d_1-dimensional probability distribution. In Section 5, we extended our upper bound and lower bound for case where possible intrinsic dimension varies from 1 to m.

2. Definitions and Regularity Conditions

In this section we define the model which consists of the set of distributions that live on manifolds whose dimension d is between $1 \leq d \leq m$. The manifolds are required to have an upper bound on their curvature. The resulting class of distributions is denoted by

$$\mathcal{P} = \bigcup_{d=1}^{m} \mathcal{P}^d_{\kappa_l,\kappa_g,K_p,K_v}.$$

The rest of this section makes the definition precise. Readers who are not interested in the precise details may skip the rest of the section.

2.1. **Notation and Basic Definitions.** Throughout the paper, we will use the following notation. For positive integers n_1, n_2, d such that $1 \leq n_1 \leq n_2 \leq d$, the coordinate projection map $\Pi_{n_1:n_2} : \mathbb{R}^d \to \mathbb{R}^{n_2-n_1+1}$ is defined by $\Pi_{n_1:n_2}(x_1, \cdots, x_d) = (x_{n_1}, x_{n_1+1}, \cdots, x_{n_2})$. We let S_n denote the permutation group on $\{1, \ldots, n\}$. For any product set $J^n \subset \mathbb{R}^n$, S_n acts on J^n and its subsets by applying a coordinate change, i.e. for $\sigma \in S_n$ and $x \in J^n$, $\sigma x := (x_{\sigma(1)}, \cdots, x_{\sigma(n)})$, and, for any $A \subset J^n$, $S_n A := \{\sigma x \in J^n : \ \sigma \in S_n, \ x \in A\}$. Finally, for a metric space (X, d_X) and $x \in X$, we let $B_X(x, r) = \{y \in X : \ d_X(y, x) < r\}$ be the ball with center x and radius r. We will set ω_d to be the volume of the unit ball in the d-dimensional Euclidean space $\mathbb{R}^d$, which can be computed exactly as $\omega_d = \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2}+1)}$.

We next briefly review some key, basic concepts in differential geometry. For more detailed treatment, we refer the reader to standard textbooks on this topic [see, e.g., **??**]. A topological manifold of dimension d is a topological space M and a family of homeomorphisms $\mathbf{x}_\alpha : U_\alpha \subset \mathbb{R}^d \to V_\alpha \subset M$ from open subset of $\mathbb{R}^d$ to open subset of M such that $\bigcup_\alpha \mathbf{x}_\alpha(U_\alpha) = M$. If M is a d-dimensional manifold, such d is unique and is called the dimension of manifold. If, for any pair α, β, with $\mathbf{x}_\alpha(U_\alpha) \cap \mathbf{x}_\beta(U_\beta) \neq \emptyset$, $\mathbf{x}_\beta^{-1} \circ \mathbf{x}_\alpha : U_\alpha \cap U_\beta \to U_\alpha \cap U_\beta$ is C^k, then M is C^k-manifold.

Let T_pM denote the tangent space to M at p. Given $p \in M$, there exist a set $0 \in \mathcal{E} \subset T_p(M)$ and a mapping $\exp_p : \mathcal{E} \subset T_pM \to M$ such that $t \to \exp_p(tv)$, $t \in (-1, 1)$, is the unique geodesic of M which, at $t = 0$, passes through p with velocity v, for all $v \in \mathcal{E}$. The map $\exp_p : \mathcal{E} \subset T_pM \to M$ is called exponential map on p.

2.2. **Minimax Theory.** The minimax rate is the risk of an estimator that performs best in the worst case, as a function of the sample size [**?**, see, e.g.]. Let $\mathcal{P}$ be a collection of probability distributions over the same sample space $\mathcal{X}$ and $\theta : \mathcal{P} \to \Theta$ be a functional over $\mathcal{P}$ taking value in some space Θ, the parameter space. We can think of $\theta(P)$ as the feature of interest of the probability distribution P, such as its mean, or, like in our case, the dimension of its support. For fixed sample size n, suppose $X = (X_1, \cdots, X_n)$ is an i.i.d. drawn from a fixed probability distribution $P \in \mathcal{P}$. Thus X takes values in the n-fold product space $\mathcal{X}^n = \mathcal{X} \times \ldots \times \mathcal{X}$ and is distributed as $P^{(n)}$, the n-fold product measure. An estimator $\widehat{\theta}_n : \mathbb{R}^n \to \Theta$ is any measurable function that maps the observations X into parameter space Θ. Let $\ell : \Theta \times \Theta \to \mathbb{R}$ be a loss function, a non-negative, non-decreasing bounded function that measures how different two parameters are. Then for a fixed estimator $\widehat{\theta}_n$ and a fixed distribution P, risk of $\widehat{\theta}_n$ is defined as

$$\mathbb{E}_{P^{(n)}}\left[\ell\left(\widehat{\theta}_n(X), \theta(P)\right)\right].$$

Then for fixed estimator $\widehat{\theta}_n$, its maximum risk is the supremum of its risk over all distribution $P \in \mathcal{P}$, that is,

$$\sup_{P \in \mathcal{P}} \mathbb{E}_{P^{(n)}}\left[\ell\left(\widehat{\theta}_n(X), \theta(P)\right)\right].$$

The minimax risk associated to $\mathcal{P}$, θ, ℓ and n is the maximal risk of any estimator that performs best under the worst possible choice of P. Formally, the minimax risk is

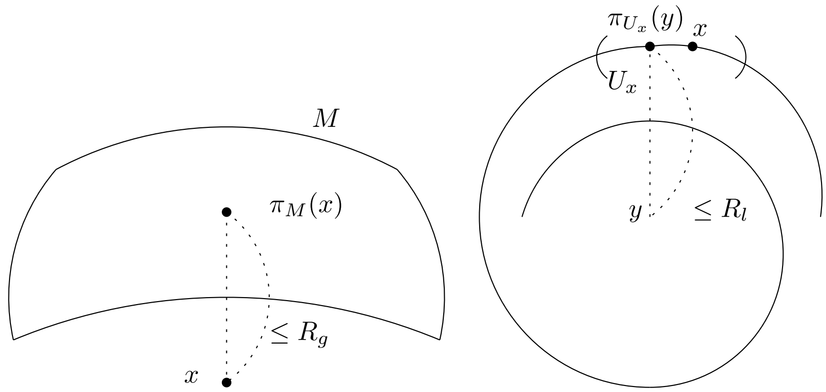
$$R_n = \inf_{\widehat{\theta}_n} \sup_{P \in \mathcal{P}} \mathbb{E}_{P^{(n)}}\left[\ell\left(\widehat{\theta}_n(X), \theta(P)\right)\right]. \tag{2.1}$$

The minimax risk R_n is often viewed as a function of sample size n, in which case any positive sequence ψ_n such that $\lim_{n\to\infty} R_n/\psi_n$ remains bounded away from 0 and ∞ is called a *minimax rate.* Notice that minimax rates are unique up to constants and lower order terms.

The determination of a minimax rate for a given problem requires two separate calculations: that of an upper bound on R_n and that of a lower bound. In order to derive an upper bound, one analyzes the asymptotic risk a specific estimator $\widehat{\theta}_n$. Lower bounds are instead usually computed by measuring the difficulty of a multiple hypothesis testing problems that entails identifying finitely many distributions in $\mathcal{P}$ that are maximally difficult to discriminate and yet their parameter values are well-separated under the loss ℓ [see, e.g. **?**, Section 2.2].

For the dimension estimation problem, we obtain an upper bound on the minimax risk by analyzing the performance of an estimator based on the length of the traveling salesman problem, as described in Section 3. On the other hand, the determination of the lower bound presents non-trivial technical difficulties, due to the fact that that probability distributions supported on manifolds of different dimensions are singular to each other, and therefore trivially discriminable. In order to overcome such an issue, we resort to constructing mixtures of mutually singular distributions. We detail such construction in Section 4.

More detail on how we get the final lower bound?



(A) global curvature (B) local curvature

FIGURE 2.1. A manifold M with (a) *global curvature* less than κ_g, or (b) *local curvature* less than κ_l.

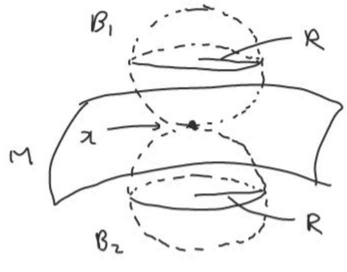
As anticipated, the lower and upper bounds we derive on the minimax risk are not asymptotic equivalent. While we do not know which one, if any, is sharp, nonetheless the derivation of such bounds is of use in understanding the difficulty of the dimension estimation problem.

2.3. **Regularity conditions on Distributions and Supporting Manifolds.** To derive bounds on minimax risk, we will impose some regularity conditions. First, the supporting manifold M is assumed to be bounded, that is, $M \subset I := [-K_I, K_I]^m \subset \mathbb{R}^m$, where $K_I \in [1, \infty)$. Second, the curvature is assumed to be bounded to avoid an arbitrarily complicated manifold. In fact there are several types of curvature so we will need the following definitions.

Definition 1. Fix $0 \leq \kappa_l \leq \kappa_g < \infty$ and let $R_l := \frac{1}{\kappa_l}$, $R_g := \frac{1}{\kappa_g} \in (0, \infty]$. A compact d-dimensional topological manifold $M \subset I$ (with boundary) is of *global curvature* less than κ_g, if for all $x \in \mathbb{R}^m$ with $d_{\mathbb{R}^m}(x, M) < R_g$ has unique projection $\pi_M(x)$ to M, i.e. there uniquely exists $\pi_M(x) \in M$ such that $d(x, \pi_M(x)) = \inf_{y \in M} d(x, y)$. M has *local curvature* less than κ_l if for all $x \in M$, there exists neighborhood $U_x \subset M$ of x such that U_x is of global curvature less than κ_l. See Figure 2.1.

Definition 2. We define $\mathcal{M}^d_{\kappa_l, \kappa_g}$ to be the set of all d-dimensional topological manifolds in I with local and global curvature bounded by κ_l and κ_g, respectively.

Remark 1. The above definition is equivalent to the following: M is of global curvature $\leq \kappa_g$ if $\forall x \in M$, $\forall y \in M$ s.t. $y - x \perp T_x M$ and $\|y - x\|_2 = R_g$, $B_{\mathbb{R}^m}(y, R_g) \cap M = \emptyset$, where $T_x M$ is



tangent space of M at x.

We assume that the data are generated from a distribution supported on a manifold M of dimension d, and with density with respect to the Hausdorff measure vol_M on M bounded away from ∞.

Definition 3. Let $\mathcal{B}(I)$ be the Borel subsets of I and $\mathcal{P}$ be a set of probability measures on $(I, \mathcal{B}(I))$. Fix $K_p \geq (2K_I)^m$. Let $\mathcal{P}^d_{\kappa_l,\kappa_g,K_p}$ be the set of probability distributions P supported on a d-dimensional manifold $M \in \mathcal{M}^d_{\kappa_l,\kappa_g}$, absolutely continuous with respect to the restriction vol_M of the d-dimensional Hausdorff measure on M and such that $\sup_{x\in M} \frac{dP}{dvol_M}(x) \leq K_p$. For all $P \in \mathcal{P}^d_{\kappa_l,\kappa_g,K_p}$, define $\dim(P) := d$.

To deal with manifolds with boundary, we further need to assume local geodesic completeness and essential dimension.

Definition 4. For a manifold $M \in \mathcal{M}^d_{\kappa_l,\kappa_g}$, the interior $\text{int}M := M \backslash \partial M$. $intM$ is said to be locally (geodesically) complete if for all $p \in \text{int}M$ and for all $q_1, q_2 \in B_M(p, 2\sqrt{3}R_g)$, there exists a geodesic γ joining q_1 and q_2 whose image lies on $intM$. Fix $K_v \leq 2^{-m}$, then M is said to be of essential dimension d (with respect to bound K_v) if for all $p \in M$ and for all $r \leq \sqrt{3}R_g$, $vol_M(B_M(p,r)) \geq K_v r^d \omega_d$. Let $\mathcal{M}^d_{\kappa_l,\kappa_g,K_v} := \{M \subset \mathcal{M}^d_{\kappa_l,\kappa_g} : M$ is locally complete and of essential dimension $d\}$. Corrrespondingly define $\mathcal{P}^d_{\kappa_l,\kappa_g,K_p,K_v} := \{P \in \mathcal{P} :$ there exists $M \in \mathcal{M}^d_{\kappa_l,\kappa_g,K_v}$ such that $P \ll vol_M$ and $\frac{dP}{dvol_M} \leq K_p\}$.

Remark 2. For manifolds without boundary, the local completeness condition and the essential dimension condition always hold. The Hopf Rinow Theorem (Theorem 16 in [Petersen, 2006]) implies that any compact closed manifold without boundary is geodesic complete, which implies it is locally complete. Also, Lemma 5.3 in [Niyogi et al., 2008] implies that when $M \in \mathcal{M}^d_{\kappa_l,\kappa_g}$ and $r \leq 2R_g$, then $\forall p \in M$,

$$vol_M(B_M(p,r)) \geq r^d \left(1 - \left(\frac{\kappa_g r}{2}\right)^2\right)^{\frac{d}{2}} \omega_d. \tag{2.2}$$

Hence by setting $r = \sqrt{3}R_g$, $vol_M(B_M(p,r)) \geq 2^{-d} r^d \omega_d$, so the essential dimension condition is satisfied.

The preceding regularity conditions imply additional conditions on both the distribution and the supporting manifold. These additional conditions, given now in Lemma 2, 3, and 4 will be used later. The proofs for Lemma 2 and Lemma 4 are in the appendix.

Lemma 2. *Let $M \in \mathcal{M}^d_{\kappa_l,\kappa_g,K_v}$ satisfies $M \subset A$, and let $r \leq R_g$. Let $A_r := \{x \in \mathbb{R}^m : d_{\mathbb{R}^m}(x, A) < r\}$ be r-neighborhood of A in $\mathbb{R}^m$. Then, volume of M is bounded by*

$$vol_M(M) \leq C^{(2,1)}_{d,m} r^{d-m} vol_{\mathbb{R}^m}(A_r), \tag{2.3}$$

where $C^{(2,1)}_{d,m}$ depends only on d and m. In particular, considering the case $A = I$ and $r = \min\left\{R_g, \frac{m-d}{d}K_I\right\}$,

$$vol_M(M) \leq C^{(2,2)}_{K_I,d,m}\left(1 + \kappa_g^{m-d}\right), \tag{2.4}$$

where $C^{(2,2)}_{K_I,d,m}$ *depends only on* K_I, d *and* m.

Lemma 3. *Let* $M \in \mathcal{M}^d_{\kappa_l,\kappa_g,K_v}$ *and* $r \leq 4R_g$. *Then* M *can be covered by* N *radius* r *balls* $B_M(p_1, r)$, $\cdots$, $B_M(p_N, r)$, *with*

$$N = \left\lfloor \frac{2^d vol(M)}{K_v r^d \omega_d} \right\rfloor . \tag{2.5}$$

Proof. See 4.3.1. Lemma 3 in [Ma and Fu, 2012]. □

Lemma 4. *Let* $M \in \mathcal{M}^d_{\kappa_l,\kappa_g,K_v}$ *and let* $exp_{p_k} : \mathcal{E}_k \subset \mathbb{R}^m \to \mathcal{M}$ *be an exponential map, where* $T_{p_k}M$ *is identified with* $\mathbb{R}^m$. *Then for all* $v, w \in \mathcal{E}_k \cap B_{\mathbb{R}^d}(0, R_k)$,

$$\| \exp_{p_k}(v) - \exp_{p_k}(w)\|_{\mathbb{R}^m} \leq \frac{e^{\kappa_l R_k} \sinh \kappa_l R_k}{\kappa_l R_k} \|v - w\|_{\mathbb{R}^d}. \tag{2.6}$$

Under these regularity conditions, and given $d_1 < d_2$, the minimax rate R_n is defined as

$$R_n = \inf_{\widehat{\dim}_n} \sup_{P \in \mathcal{P}} \mathbb{E}_{P^{(n)}} \left[\ell \left(\widehat{\dim}_n(X), \dim(P) \right) \right] \tag{2.7}$$

where

$$\mathcal{P} = \mathcal{P}^{d_1}_{\kappa_l,\kappa_g,K_p,K_v} \bigcup \mathcal{P}^{d_2}_{\kappa_l,\kappa_g,K_p,K_v} \tag{2.8}$$

Here $\widehat{\dim}_n$ is any dimension estimator based on data $X = (X_1, \cdots , X_n)$, and the loss function $\ell(\cdot, \cdot)$ is $0 - 1$ loss, so for all $x, y \in \mathbb{R}$, $\ell(x, y) = I(x = y)$.

3. Upper Bound

In this section, we give an upper bound on the minimax rate. Our strategy to accomplish this task is to focus on a particular estimator $\widehat{\dim}_n$and demonstrate an upper bound on its risk uniformly over the class $\mathcal{P}$. This will in turn provides an upper bound on the minimax risk since since

$$R_n = \inf_{\widehat{\dim}^*_n} \sup_{P \in \mathcal{P}} \mathbb{E}_{P^{(n)}} \left[\ell \left(\widehat{\dim}^*_n(X), \dim(P) \right) \right] \leq \sup_{P \in \mathcal{P}} \mathbb{E}_{P^{(n)}} \left[\ell \left(\widehat{\dim}^*_n(X), \dim(P) \right) \right] .$$

Choosing an appropriate estimator is critical to get a good bound.

For now, the intrinsic dimension of data is assumed to be either d_1 or d_2.

Our estimator is based on the d_1-squared length of TSP (Traveling Salesman Path) generated by the data, and estimating dimension to be d_1 if and only if the length is below a certain threshold:

$$\widehat{\dim}_n(X) = d_1 \iff \exists \sigma \in S_n \ s.t \sum_{i=1}^{n-1} \|X_{\sigma(i+1)} - X_{\sigma(i)}\|^{d_1}_{\mathbb{R}^m} \leq C^{(3,2)}_{K_I,K_v,d_1,m} \left(1 + \kappa_g^{m-d_1}\right),$$

where $C^{(3,2)}_{K_I,K_v,d_1,m}$ is a constant to be defined later. Then in Proposition 7, it is shown that this estimator $\widehat{\dim}_n$ is always correct when the intrinsic dimension is d_1, and makes error with probability at most $O\left(n^{-\left(\frac{d_2}{d_1}-1\right)n}\right)$ if intrinsic dimension is d_2.

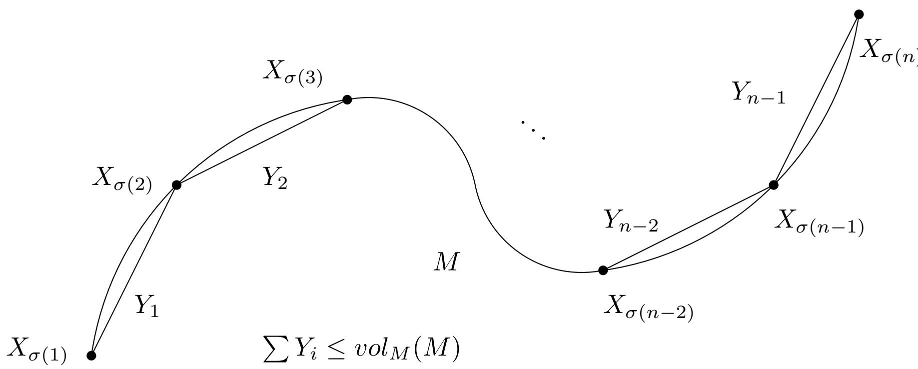


FIGURE 3.1. When the manifold is a curve, length of TSP path $\sum_{i=1}^{n-1} \|X_{\sigma(i+1)} - X_{\sigma(i)}\|_{\mathbb{R}^m}$ is bounded by the length of the curve $vol_M(M)$.

Lemma 5 shows that the estimator $\widehat{\dim}_n$ makes error with probability at most $O\left(n^{-\left(\frac{d_2}{d_1}-1\right)n}\right)$ if the intrinsic dimension is d_2. It states that d_1-squared length of a piecewise linear path from X_1 to X_n, $\sum_{i=1}^{n-1} \|X_{i+1} - X_i\|_{\mathbb{R}^m}^{d_1}$, is bounded with probability $O\left(n^{-\left(\frac{d_2}{d_1}-1\right)n}\right)$, and hence not likely to be bounded by any threshold L.

Lemma 5. *Let $X_1, \cdots, X_n \sim P \in \mathcal{P}^{d_2}_{\kappa_l,\kappa_g,K_p,K_v}$, then*

$$P^{(n)}\left[\sum_{i=1}^{n-1} \|X_{i+1} - X_i\|^{d_1} \le L\right] \le \frac{\left(C^{(3,1)}_{K_I,K_p,d_1,d_2,m}\right)^{n-1} L^{\frac{d_2}{d_1}(n-1)} \left(1 + \kappa_g^{(m-d_2)(n-1)}\right)}{(n-1)^{\left(\frac{d_2}{d_1}-1\right)(n-1)}(n-1)!}, \tag{3.1}$$

where $C^{(3,1)}_{K_I,K_p,d_1,d_2,m}$ depends only on K_I, K_p, d_1, d_2, m.

Proof. in Appendix B. ☐

Lemma 6 show that the estimator $\widehat{\dim}_n$ is always correct when intrinsic dimension is d_1. Lemma 6 states that d_1-squared length of TSP path generated from data is bounded by $C^{(3,2)}_{K_I,K_v,d_1,m}\left(1 + \kappa_g^{m-d_1}\right)$, i.e. there exists $\sigma \in S_n$ such that $\sum_{i=1}^{n-1} \|X_{\sigma(i+1)} - X_{\sigma(i)}\|_{\mathbb{R}^m}^{d_1} \le C^{(3,2)}_{K_I,K_v,d_1,m}\left(1 + \kappa_g^{m-d_1}\right)$. When $d_1 = 1$, this lemma is straightforward: length of TSP path is bounded by length of curve $vol_M(M)$ as in Figure 3.1, and from Lemma 2 we have $vol_M(M) \le C^{(2,2)}_{K_I,d,m}\left(1 + \kappa_g^{m-1}\right)$, hence $C^{(3,2)}_{K_I,K_v,d_1,m}$ can be set as $C^{(3,2)}_{K_I,K_v,d_1,m} = C^{(2,2)}_{K_I,d,m}$.

When $d_1 > 1$, Lemma 2, 3, and 4, combined with Hölder continuity of d_1-dimensional space-filling curve [Steele, 1997, Buchin, 2007], is used to show Lemma 6.

Lemma 6. *Let $M \in \mathcal{M}^{d_1}_{\kappa_l,\kappa_g,K_p}$ and $X_1, \cdots, X_n \in M$. Then there exists $C^{(3,2)}_{K_I,K_v,d_1,m}$ which depends only on m, d_1, K_v, and K_I, and there exists $\sigma \in S_n$ such that*

$$\sum_{i=1}^{n-1} \|X_{\sigma(i+1)} - X_{\sigma(i)}\|_{\mathbb{R}^m}^{d_1} \le C^{(3,2)}_{K_I,K_v,d_1,m}\left(1 + \kappa_g^{m-d_1}\right) \tag{3.2}$$

Proof. in Appendix B. ☐

Proposition 7 is a combination of Lemma 5 and Lemma 6.

Proposition 7. *Let $1 \leq d_1 < d_2 \leq m$. Then*

$$\inf_{\widehat{\dim}_n} \sup_{P \in \mathcal{P}_1 \cup \mathcal{P}_2} \mathbb{E}_{P^{(n)}} \left[l \left(\widehat{\dim}_n, \dim(P) \right) \right] \tag{3.3}$$

$$\leq \left(C^{(3,3)}_{K_I, K_p, K_v, d_1, d_2, m} \right)^n \left(1 + \kappa_g^{\left(\frac{d_2}{d_1} m + m - 2d_2 \right) n} \right) n^{-\left(\frac{d_2}{d_1} - 1 \right) n}. \tag{3.4}$$

for some $C^{(3,3)}_{K_I, K_p, K_v, d_1, d_2, m}$ that depends only on $K_I, K_p, K_v, d_1, d_2, m$ where

$$\mathcal{P}_1 = \mathcal{P}^{d_1}_{\kappa_l, \kappa_g, K_p, K_v}, \quad \mathcal{P}_2 = \mathcal{P}^{d_2}_{\kappa_l, \kappa_g, K_p, K_v}.$$

Proof. in Appendix B. □

4. Lower Bound

In this section, a lower bound for the minimax rate is derived. The lower bound measures how hard it is to tell whether the data come from a d_1 or d_2-dimensional manifold. More precisely, a subset $T \subset I^n$ and set of distributions $\mathcal{P}_1^{d_1}$, $\mathcal{P}_2^{d_2}$ are found so that, whenever $X = (X_1, \cdots, X_n) \in T$, we cannot distinguish the models. T, $\mathcal{P}_1^{d_1}$ and $\mathcal{P}_2^{d_2}$ are linked to the lower bound by using Le Cam's lemma[Yu, 1997] which provides lower bounds based on the minimum of two densities $q_1 \wedge q_2$, where q_1, q_2 are constructed from $\mathcal{P}_1^{d_1}$ and $\mathcal{P}_2^{d_2}$, respectively.

Lemma 8. *(Le Cam's Lemma) Let $\mathcal{P}$ be a set of probability measures on $(\Omega, \mathcal{F})$, and $\mathcal{P}_1, \mathcal{P}_2 \subset \mathcal{P}$ be such that for all $P \in \mathcal{P}_i$, $\theta(P) = \theta_i$ for $i = 1, 2$. For any $Q_i \in co(\mathcal{P}_i)$, let q_i be density of Q_i with respect to measure ν. Then*

$$\inf_{\widehat{\theta}} \sup_{P \in \mathcal{P}} \mathbb{E}_P [d(\widehat{\theta}, \theta(P))] \geq \frac{\Delta}{4} \int [q_1(x) \wedge q_2(x)] d\nu(x), \tag{4.1}$$

where $\Delta = d(\theta_1, \theta_2)$.

Proof. See Chapter 29.2, Lemma 1 in [Yu, 1997]. □

Lemmas 9 and Lemma 10 below are ingredients for constructing subset $T \subset I^n$ and the sets of distributions $\mathcal{P}_1$, $\mathcal{P}_2$ of d_1- and d_2- dimension.

Lemma 9. *Let $M \in \mathcal{M}^d_{\kappa_l, \kappa_g}$ be d-dimensional manifold of global curvature $\leq \kappa_g$, local curvature $\leq \kappa_l$, which is imbedded in $\mathbb{R}^{m - \Delta d}$. Then*

$$M \times [-K_I, K_I]^{\Delta d} \in \mathcal{M}^{d + \Delta d}_{\kappa_l, \kappa_g}, \tag{4.2}$$

which is imbedded in $\mathbb{R}^m$.

Proof. in Appendix C. □

Lemma 10. *Let $X : [-K_\delta, K_\delta] \to I$ be parametrized curve which is C^1 and piecewise C^2. Suppose that, for all $t \in [-K_\delta, K_\delta]$,*

$$\|X''(t)\| < \|X'(t)\|_2^2 \kappa_l. \tag{4.3}$$

Then image(X) is of local curvature $\leq \kappa_l$.

Proof. in Appendix C. ☐

Corollary 11. *Let X be C^1 curve which is piecewise line or arc of circle of radius$\geq R_l$. Then image(X) is of local curvature $\leq \kappa_l$.*

Proof. Follows from Lemma 10. ☐

Lemma 12 below is for constructing the subset $T \subset I^n$ and the sets of distributions $\mathcal{P}_1$, $\mathcal{P}_2$. Claim 13 is for showing that whenever $X = (X_1, \cdots, X_n) \in T$, it is both likely that X is sampled from some distribution P that is either in $\mathcal{P}_1$ or $\mathcal{P}_2$. From Le Cam's lemma, a lower bound is given by $\int q_1(x) \wedge q_2(x)$. Hence if $q_1(x) \geq C q_2(x)$ for every $x \in T$ with $C < 1$, then $q_1(x) \wedge q_2(x) \geq C q_2(x)$, hence $C \int q_2(x)$ can serve as lower bound of minimax rate. This inequality $q_1(x) \geq C q_2(x)$ is shown in Claim 13. This intuitively means that if $X \in T$, it is hard to determine whether X is sampled from distribution P in either $\mathcal{P}_1$ or $\mathcal{P}_2$.

In Lemma 12, as in Figure C.2, we construct T_i's that are cylinder sets aligned along boundary of $[-K_I, K_I]^{d_2}$, and then T is constructed as $T = S_n \prod_{i=1}^{n} T_i$, where the permutation group S_n acts on $\prod_{i=1}^{n} T_i$ as a coordinate change. And it is also shown in Lemma 12 that for any $x \in \prod T_i$, there exists a manifold $M \in \mathcal{M}^{d_2}_{\kappa_l, \kappa_g, \infty}$ that passes through $x_1, \cdots, x_n$.

Then in Proposition 14, $\mathcal{P}_1$ is constructed as set of distributions that are supported on such a manifold, and $\mathcal{P}_2$ is a singleton set consisting of the uniform distribution on $[-K_I, K_I]^{d_2}$.

Lemma 12. *Suppose $R_l \leq K_I$. There exists $T_1, \cdots, T_n \subset [-K_I, K_I]^{d_2}$ such that*

(1) each T_i's are distinct

(2) For each T_i, there exists isometry Φ_i such that

$$T_i = \Phi_i \left([-K_I, K_I]^{d_1 - 1} \times [0, a] \times B_{\mathbb{R}^{d_2 - d_1}}(0, w) \right), \tag{4.4}$$

where $c = \left\lceil \frac{K_I + R_l}{2R_l} \right\rceil$, $a = \frac{K_I - R_l}{\left(d + \frac{1}{2}\right) \left\lceil \frac{n}{c^{d_2 - d_1}} \right\rceil}$, *and* $w = \min \left\{ R_l,\ \frac{d^2 (K_I - R_l)^2}{2R_l \left(d + \frac{1}{2}\right)^2 \left(\left\lceil \frac{n}{c^{d_2 - d_1}} \right\rceil + 1 \right)^2} \right\}$.

(3) $\exists \mathscr{M} : (B_{\mathbb{R}^{d_2 - d_1}}(0, w))^n \to \mathcal{M}^{d_1}_{\kappa_l, \kappa_g, K_v}$ one-to-one such that for each $Y_i \in B_{\mathbb{R}^{d_2 - d_1}}(0, w)$, $1 \leq i \leq n$, $\mathscr{M}(Y_1, \cdots, Y_n) \cap T_i = \Phi_i([-K_I, K_I]^{d_1 - 1} \times [0, a] \times \{Y_i\})$. Hence for any $X_1 \in T_1, \cdots, X_n \in T_n$, $\mathscr{M}(\{\Pi^{-1}_{(d_1+1):d_2} \Phi_i^{-1}(X_i)\}_{1 \leq i \leq n})$ passes through $X_1, \cdots, X_n$.

Proof. in Appendix C. ☐

Claim 13. Let $T = S_n \prod_{i=1}^{n} T_i$. Then for all $x \in \mathrm{int} T$, there exists $r_x > 0$ such that for all $r < r_x$,

$$Q_1 \left(\prod_{i=1}^{n} B_{\|\|_{\mathbb{R}^{d_2}, \infty}}(x_i, r) \right) \geq \frac{2^{(1 - d_1) n}}{\kappa_l^{(d_2 - d_1) n} K_I^{(d_2 - d_1) n}} Q_2 \left(\prod_{i=1}^{n} B_{\|\|_{\mathbb{R}^{d_2}, \infty}}(x_i, r) \right). \tag{4.5}$$

Proof. in Appendix C. ☐

Proposition 14 is a combination of Le Cam's lemma, Lemma 12, and Claim 13.

Proposition 14. *Suppose $R^l < K_I$, then*

$$\inf_{\widehat{\dim}} \sup_{P \in \mathcal{P}^{d_1}_{\kappa_l,\kappa_g,K_p,K_v} \cup \mathcal{P}^{d_2}_{\kappa_l,\kappa_g,K_p,K_v}} \mathbb{E}_{P^{(n)}}[l(\widehat{\dim}_n, \dim(P))] \tag{4.6}$$

$$\geq \left(K^{(4,1)}_{d_1,d_2,K_I}\right)^n \kappa_l^{-(d_2-d_1)n} \min\left\{\kappa_l^{2(d_2-d_1)+1} n^{-2}, 1\right\}^{(d_2-d_1)n}, \tag{4.7}$$

for some constant $K^{(4,1)}_{d_1,d_2,K_I}$ that depends only on d_1, d_2, and K_I.

Proof. in Appendix C. ☐

5. Upper bound and Lower bound for General case

Now we generalize our results to allow the intrinsic dimension to be any integer between 1 and m. Thus the model is $\mathcal{P} = \bigcup_{d=1}^{m} \mathcal{P}^d_{\kappa_l,\kappa_g,K_p,K_v}$. The estimator we consider to derive an upper bound is the simply the smallest integer d between 1 and m such that (3.2) holds. As for the lower bound, we simply use the lower bound derived in Section 4 with $d_1 = 1$ and $d_2 = 2$.

Proposition 15.

$$\inf_{\widehat{\dim}_n} \sup_{P \in \mathcal{P}} \mathbb{E}_{P^{(n)}}\left[l\left(\widehat{\dim}_n, \dim(P)\right)\right] \leq \left(C_{K_I,K_p,K_v,K_m}\right)^n \left(1 + \kappa_g^{(m^2-m)n}\right) n^{-\frac{1}{m-1}n} \tag{5.1}$$

for some $C^{(5,1)}_{K_I,K_p,K_v,K_m}$ that depends only on K_I, K_p, K_v, m.

Proof. in Appendix D. ☐

Proposition 16. *Suppose $R_l < K_I$, then*

$$\inf_{\widehat{\dim}} \sup_{P \in \mathcal{P}} \mathbb{E}_{P^{(n)}}[l(\widehat{\dim}_n, \dim(P))] \geq \left(C^{(5,2)}_{K_I}\right)^n \kappa_l^{-n} \min\left\{\kappa_l^3 n^{-2}, 1\right\}^n \tag{5.2}$$

for some $C^{(5,2)}_{K_I}$ that depends only on K_I.

Proof. in Appendix D. ☐

6. Conclusion

On a logarithmic scale, the leading terms of the lower and upper bounds have the form

$$-nc \log \kappa$$

for some constant c, where κ is the global curvature for the upper bound and the lower curvature for the lower bound. This shows that the difficulty of the problem of estimating the dimension goes to 0 rapidly with sample size, in a way that depends on the curvature of the manifold. It is an open question whether oen can obtain tighter bounds on R_n.

References

Richard. Bellman. *Adaptive Control Processes: A Guided Tour*. Princeton University Press, 1961.

M.R. Bridson and A. Häfliger. *Metric Spaces of Non-Positive Curvature*. Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen. Springer, 1999. ISBN 9783540643241.

Kevin. Buchin. *2. Space-Filling Curves*, chapter 2, pages 5–29. Freien Universität Berlin, 2007.

Francesco Camastra and Antonio Staiano. Intrinsic dimension estimation: Advances and open problem. *Information Sciences*, (38):26–41, 2015.

Trevor. Hastie, Robert. Tibshirani, and Jerome Friedman. *14. Unsupervised Learning*, chapter 14, pages 485–586. Springer-Verlag, 2009.

Matthias. Hein and Jean-Yves Audibert. Intrinsic dimensionality estimation of submanifolds in $\mathbb{R}^d$. In *22nd International Conference on Machine Learning*, 2005.

Balázs Kégl. Intrinsic dimension estimation using packing numbers, 2003.

V. I. Koltchinskii. Empirical geometry of multivariate data: a deconvolution approach. *Ann. Statist.*, 28(2):591–629, 04 2000. doi: 10.1214/aos/1016218232. URL `http://dx.doi.org/10.1214/aos/1016218232`.

John A. Lee and Michel. Verleysen. *1. High-Dimensional Data*, chapter 1, pages 1–16. Springer New York, 2007a.

John A. Lee and Michel. Verleysen. *3. Estimation of the Intrinsic Dimension*, chapter 3, pages 47–68. Springer New York, 2007b.

Elizaveta Levina and Peter J Bickel. Maximum likelihood estimation of intrinsic dimension. In *Advances in neural information processing systems*, pages 777–784, 2004.

Anna V Little, Yoon-Mo Jung, and Mauro Maggioni. Multiscale estimation of intrinsic dimensionality of data sets. In *AAAI Fall Symposium: Manifold Learning and Its Applications*, 2009.

Anna V Little, Mauro Maggioni, and Lorenzo Rosasco. Multiscale geometric methods for estimating intrinsic dimension. *Proc. SampTA*, 2011.

Yunqian Ma and Yun Fu. *Manifold Learning Theory and Applications*. CRC Press, 2012.

Partha Niyogi, Stephen Smale, and Shmuel Weinberger. Finding the homology of submanifolds with high confidence from random samples. *Discrete & Computational Geometry*, 39(1-3):419–441, 2008. ISSN 0179-5376. doi: 10.1007/s00454-008-9053-2. URL `http://dx.doi.org/10.1007/s00454-008-9053-2`.

Peter Petersen. *Riemannian Geometry*. Springer New York, 2006. doi: 10.1007/978-0-387-29403-2.

Maxim Raginsky and Svetlana Lazebnik. Estimation of intrinsic dimensionality using high-rate vector quantization. In *Advances in neural information processing systems*, pages 1105–1112, 2005.

Alessandro Rozza, Gabriele Lombardi, Claudio Ceruti, Elena Casiraghi, and Paola Campadelli. Novel high intrinsic dimensionality estimators. *Machine learning*, 89(1-2):37–65, 2012.

Kumar Sricharan, Raviv Raich, and AO Hero. Optimized intrinsic dimension estimator using nearest neighbor graphs. In *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, pages 5418–5421. IEEE, 2010.

J. Michael Steele. *2. Concentration of Measure and the Classical Theorems*, chapter 2, pages 27–51. Society for Industrial and Applied Mathematics, 1997. doi: 10.1137/1.9781611970029.ch2. URL http://epubs.siam.org/doi/abs/10.1137/1.9781611970029.ch2.

Bin Yu. Assouad, fano, and le cam. In David Pollard, Erik Torgersen, and GraceL. Yang, editors, *Festschrift for Lucien Le Cam*. Springer New York, 1997.

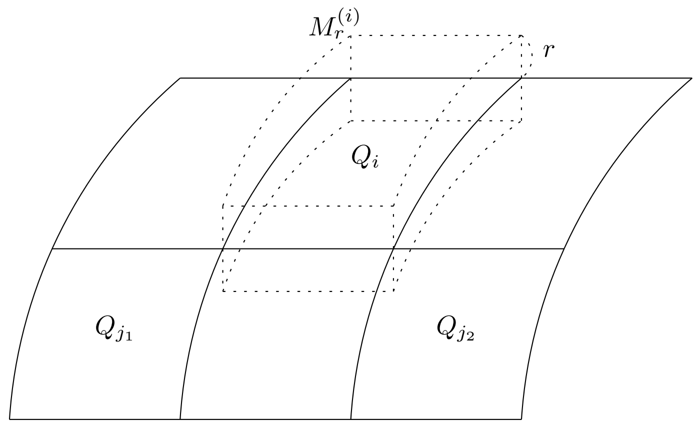


FIGURE A.1. $\{Q_1, \cdots, Q_l\}$ is a disjoint cover of M, and each Q_i is a projection of $M_r^{(i)}$ on M.

APPENDIX A. PROOFS FOR SECTION 2

Lemma. *2. Let $M \in \mathcal{M}^d_{\kappa_l,\kappa_g,K_v}$ satisfies $M \subset A$, and let $r \leq R_g$. Let $A_r := \{x \in \mathbb{R}^m : d_{\mathbb{R}^m}(x, A) < r\}$ be r-neighborhood of A in $\mathbb{R}^m$. Then, volume of M is bounded by*

$$vol_M(M) \leq C^{(2,1)}_{d,m} r^{d-m} vol_{\mathbb{R}^m}(A_r), \tag{A.1}$$

where $C^{(2,1)}_{d,m}$ depends only on d and m. In particular, considering the case $A = I$ and $r = \min\left\{R_g, \frac{m-d}{d} K_I\right\}$,

$$vol_M(M) \leq C^{(2,2)}_{K_I,d,m}\left(1 + \kappa_g^{m-d}\right), \tag{A.2}$$

where $C^{(2,2)}_{K_I,d,m}$ depends only on K_I, d and m.

Proof. Let $M_r := \{x \in \mathbb{R}^m : \ d_{\|\cdot\|_1}(x, M) < r\}$ be r-neighborhood of M in $\mathbb{R}^m$, then trivially

$$vol_{\mathbb{R}^m}(M_r) \leq vol_{\mathbb{R}^m}(A_r). \tag{A.3}$$

Suppose $\{Q_1, \cdots, Q_l\}$ is a disjoint cover of M, i.e. $\{Q_1, \cdots, Q_l\}$ are measurable subset of M that $Q_i \cap Q_j = \emptyset$, $\bigcup\limits_{i=1}^{l} Q_i = M$, and each Q_i is equipped with chart maps $X^{(i)} : U_i \to Q_i$. Such a triangulation is always possible. For each Q_i, define $M_r^{(i)} := \{x \in \mathbb{R}^m : \ \pi_M(x) \in Q_i, \ d_{\|\cdot\|_{\mathbb{R}^m,1}}(x, M) \leq r\}$ so that each Q_i is a projection of $M_r^{(i)}$ on M, as in Figure A.1. Then,

$$vol_{\mathbb{R}^m}(M_r) = \sum_{i=1}^{l} vol_{\mathbb{R}^m}(M_r^{(i)}). \tag{A.4}$$

Fix $i \in \{1, \cdots, l\}$. $M_r^{(i)}$can be parametrized as $Y^{(i)} : U_i \times B_{\|\cdot\|_{\mathbb{R}^{m-d},1}}(0, r) \to M_r^{(i)}$with

$$Y^{(i)}(u, t) = X^{(i)}(u) + R^{(i)}(u)t = X^{(i)}(u) + \sum_{j=1}^{m-d} t_j R^{(i,j)}(u), \tag{A.5}$$

where $R^{(i)}(u):\mathbb{R}^{m-d}\to(T_{X^{(i)}(u)}M)^{\perp}$ is a linear isometry for each $u\in U$, and can be identified as an $m\times(m-d)$ matrix, and $R^{(i,j)}(u)$ is j^{th} column of $R^{(i)}(u)$. Then, because of the isometry,

$$R^{(i)}(u)^T R^{(i)}(u) = I_{m-d}. \tag{A.6}$$

Let $Y_u^{(i)} = \frac{\partial Y^{(i)}}{\partial u} = \left(\frac{\partial Y^{(i)}}{\partial u_1},\cdots,\frac{\partial Y^{(i)}}{\partial u_d}\right)\in\mathbb{R}^{m\times d}$ be partial derivatives with respect to u and $Y_t^{(i)} = \frac{\partial Y^{(i)}}{\partial t}$ as well. define $X_u^{(i)}$, $X_t^{(i)}$, $R_u^{(i,j)}$, $R_t^{(i,j)}$ similarily. Then, since $image(R^{(i)}(u))\subset(T_{X^{(i)}(u)}M)^{\perp}$,

$$R^{(i)}(u)^T X_u^{(i)}(u) = 0. \tag{A.7}$$

Also from $R^{(i)}(u)^T R^{(i)}(u) = I_{m-d}$,

$$\forall j,\ R^{(i)}(u)^T R_u^{(i,j)}(u) = 0. \tag{A.8}$$

Then from the fact that

$$Y_u^{(i)}(u,t) = X_u^{(i)}(u) + \sum_{j=1}^{m-d} t_j R_u^{(i,j)}(u) \tag{A.9}$$

and

$$Y_t^{(i)}(u,t) = R^{(i)}(u), \tag{A.10}$$

we have

$$Y_t^{(i)}(u,t)^T Y_u^{(i)}(u,t) = R^{(i)}(u)^T X_u^{(i)}(u) + R^{(i)}(u)^T R_u^{(i)}(u)t = 0. \tag{A.11}$$

Also,

$$Y_t^{(i)}(u,t)^T Y_t^{(i)}(u,t) = R^{(i)}(u)^T R^{(i)}(u) = I_{m-d}. \tag{A.12}$$

Now let's consider $Y_u^{(i)}(u,t)^T Y_u^{(i)}(u,t)$. Since $R_u^{(i,j)}(u)^T R^{(i)}(u) = 0$, column space generated by $R_u^{(i,j)}(u)$ is $< R_u^{(i,j)}(u) > \subset T_{X^{(i)}(u)}(M) = span(X_u^{(i)}(u))$. Therefore, there exists $\Lambda^{(i,j)}(u) : d\times d$ matrix such that

$$R_u^{(i,j)}(u) = X_u^{(i)}(u)\Lambda^{(i,j)}(u). \tag{A.13}$$

Then,

$$Y_u^{(i)}(u,t) = X_u^{(i)}\left(I + \sum_{j=1}^{m-d} t_j \Lambda^{(i,j)}(u)\right). \tag{A.14}$$

M being of global curvature $\le \kappa_g$ implies $Y_u^{(i)}(u,t)$ is of full rank for all $t\in B_{\mathbb{R}^{m-d}}(0,R_g)$. Hence this implies $I + \sum_{j=1}^{m-d} t_j\Lambda^{(i,j)}(u)$ is invertible for all $t\in B_{\mathbb{R}^{m-d}}(0,R_g)$, and this implies all singular values of $\Lambda^{(i,j)}(u)$ is bounded by κ_g. Hence $\forall v\in\mathbb{R}^d$,

$$\left|v^T\Lambda^{(i,j)}(u)v\right| \le \kappa_g\|v\|_2^2. \tag{A.15}$$

From this, $\forall v \in \mathbb{R}^d$,

$$\left| v^T \left(I + \sum_{j=1}^{m-d} t_j \Lambda^{(i,j)}(u) \right) v \right| \geq \|v\|_2^2 - \sum_{j=1}^{m-d} |t_j| \left| v^T \Lambda^{(i,j)}(u) v \right| \tag{A.16}$$

$$\geq (1 - \|t\|_1 \kappa_g) \|v\|_2^2, \tag{A.17}$$

Hence for any singular values σ of $I + \sum_{j=1}^{m-d} t_j \Lambda^{(i,j)}(u)$, $|\sigma| \geq 1 - \|t\|_1 \kappa_g$. And since $\|t\|_1 \leq R_g$,

$$\left| I + \sum_{j=1}^{m-d} t_j \Lambda^{(i,j)}(u) \right| \geq (1 - \|t\|_1 \kappa_g)^d. \tag{A.18}$$

Then determinant of Riemannian metric tensor is lower bounded by

$$|\det(g_{ij})|^2 = \left| \left(Y_u^{(i)}(u,t) \; Y_t^{(i)}(u,t) \right)^T \left(Y_u^{(i)}(u,t) \; Y_t^{(i)}(u,t) \right) \right| \tag{A.19}$$

$$= \left| \begin{array}{cc} Y_u^{(i)}(u,t)^T Y_u^{(i)}(u,t) & Y_u^{(i)}(u,t)^T Y_t^{(i)}(u,t) \\ Y_u^{(i)}(u,t)^T Y_t^{(i)}(u,t) & Y_t^{(i)}(u,t)^T Y_t^{(i)}(u,t) \end{array} \right| \tag{A.20}$$

$$= \left| \; Y_u^{(i)}(u,t)^T Y_u^{(i)}(u,t) \; \right| \tag{A.21}$$

$$\geq (1 - \|t\|_1 \kappa_g)^d. \tag{A.22}$$

Hence

$$vol_{\mathbb{R}^m}(M_r^{(i)}) = \int_{U_i \times B_{\|\cdot\|_{\mathbb{R}^{m-d},1}}(0,r)} |\det(g_{ij})| du dt \tag{A.23}$$

$$\geq \int_{U_i} \int_{B_{\|\cdot\|_{\mathbb{R}^{m-d},1}}(0,r)} (1 - \|t\|_1 \kappa_g)^{\frac{d}{2}} dt du \tag{A.24}$$

$$= vol(U_i) \int_0^r \int_{t_1 + \cdots + t_{m-d-1} \leq s} (1 - s\kappa_g)^{\frac{d}{2}} dt_1 \cdots dt_{m-d-1} ds \tag{A.25}$$

$$= \frac{1}{(m-d-1)!} vol(U_i) \int_0^r s^{m-d-1} (1 - s\kappa_g)^{\frac{d}{2}} ds \tag{A.26}$$

$$\geq \frac{1}{C_{d,m}^{(2)}} r^{m-d} vol(U_i), \tag{A.27}$$

where $C_{d,m}^{(2)}$ depends only on d and m. Therefore,

$$vol_{\mathbb{R}^m}(M_r) \geq \frac{1}{C_{d,m}^{(2)}} r^{m-d} \sum_{i=1}^{l} vol(U_i) \tag{A.28}$$

$$= \frac{1}{C_{d,m}^{(2)}} r^{m-d} vol_M(M) \tag{A.29}$$

and hence we have following result:

$$vol_M(M) \leq C_{d,m}^{(2)} r^{d-m} vol_{\mathbb{R}^m}(M_r) \leq C_{d,m}^{(2)} r^{d-m} vol_{\mathbb{R}^m}(A_r). \tag{A.30}$$

□

Jisu: remove next lemm unless it is needed

Lemma. *(Now this seems useless) Let* $M \in \mathcal{M}^d_{\kappa_l,\kappa_g,\infty}$ *and* $x, y \in M$*. Let* M_1 *be geodesic path connecting* $x, y \in M$*. Then* $M_1 \in \mathcal{M}^1_{\kappa_l,4\kappa_g,\infty}$*.*

Proof. Since M_1 is geodesic on M and the local curvature is $\leq \kappa_l$, we have that M_1 is also of local curvature $\leq \kappa_l$. Hence we only need to show that global curvature is $\leq 4\kappa_g$.

Assume to the contrary, so that there exists $x \in \mathbb{R}^m$ such that $d(x, M_1) = r < \frac{R_g}{4}$ with at least two projection on M_1, i.e. there exists $p \neq q \in M_1$ such that $\|x - p\|_{\mathbb{R}^m} = \|x - q\|_{\mathbb{R}^m} = r$. Define $X_1 : [-1, 1] \to M$ as

$$X_1(\lambda) = \begin{cases} \pi_M(-\lambda p + (1+\lambda)x) & \lambda \in [-1, 0] \\ \pi_M(\lambda x + (1-\lambda)q) & \lambda \in [0, 1] \end{cases}. \tag{A.31}$$

Such projection is well defined and X_1 is continuous since M is of global curvature $\leq \kappa_g$. Let $d_\lambda = \|p - X_1(\lambda)\|_{\mathbb{R}^m}$. Then for $\lambda \in [-1, 0]$,

$$\|(-\lambda p + (1+\lambda)x) - \pi_M(-\lambda p + (1+\lambda)x)\|_{\mathbb{R}^m} < \|(-\lambda p + (1+\lambda)x) - p\|_{\mathbb{R}^m} \tag{A.32}$$

$$= |1+\lambda|\|x - p\|_{\mathbb{R}^m} \leq |1+\lambda| r, \tag{A.33}$$

we have

$$\|p - X_1(\lambda)\|_{\mathbb{R}^m} \leq \|p - (-\lambda p + (1+\lambda)x)\|_{\mathbb{R}^m} + \|(-\lambda p + (1+\lambda)x) - \pi_M(-\lambda p + (1+\lambda)x)\|_{\mathbb{R}^m} \tag{A.34}$$

$$< 2|1+\lambda|\|x - p\|_{\mathbb{R}^m} \leq 2r. \tag{A.35}$$

For $\lambda \in [0, 1]$, we have similarly $\|(\lambda x + (1-\lambda)q) - \pi_M(\lambda x + (1-\lambda)q)\|_{\mathbb{R}^m} \leq |\lambda| r$, so

$$\|p - X_1(\lambda)\|_{\mathbb{R}^m} \tag{A.36}$$

$$\leq \|p - x\|_{\mathbb{R}^m} + \|x - (\lambda x + (1-\lambda)q)\|_{\mathbb{R}^m} + \|(\lambda x + (1-\lambda)q) - \pi_M(\lambda x + (1-\lambda)q)\|_{\mathbb{R}^m} \tag{A.37}$$

$$< \|p - x\|_{\mathbb{R}^m} + (1-\lambda)\|p - x\|_{\mathbb{R}^m} + \lambda\|p - x\|_{\mathbb{R}^m} \tag{A.38}$$

$$= 2\|x - p\|_{\mathbb{R}^m} = 2r. \tag{A.39}$$

Hence we have for all $\lambda \in [-1, 1]$,

$$d_\lambda \leq 2r < 2 \cdot \frac{R_g}{4} = \frac{1}{2} R_g. \tag{A.40}$$

Let $c_\lambda : [0, s_\lambda] \to M$ be (arc length parametrized) geodesic connecting p and $X_1(\lambda)$, so that $c_\lambda(0) = p$, $c_\lambda(s) = X_1(\lambda)$, and $d_M(p, X_1(\lambda)) = s_\lambda$. Then from M being of global curvature$\leq \kappa_g$,

$\|c''(t)\| \leq \kappa_g$. Then from

$$c_\lambda(s_\lambda) - c_\lambda(0) = \int_0^{s_\lambda} c'_\lambda(t)dt \tag{A.41}$$

$$= \int_0^{s} c'_\lambda(0)dt + \int_0^{s_\lambda}\int_0^t c''_\lambda(u)dudt, \tag{A.42}$$

we have

$$d_\lambda = \|c_\lambda(s_\lambda) - c_\lambda(0)\|_{\mathbb{R}^m} \tag{A.43}$$

$$\geq \|\int_0^{s_\lambda} c'_\lambda(0)dt\| - \int_0^{s_\lambda}\int_0^t \|c''_\lambda(u)\|dudt \tag{A.44}$$

$$\geq s_\lambda - \int_0^{s_\lambda}\int_0^t \kappa_g dudt = s_\lambda - \frac{1}{2}\kappa_g s_\lambda^2 \tag{A.45}$$

hence we have

$$s_\lambda^2 - 2R_g s_\lambda + 2R_g d_\lambda \geq 0. \tag{A.46}$$

This is satisfied if and only if $s_\lambda \leq R_g - R_g\sqrt{1-2\kappa_g d_\lambda}$ or $s_\lambda \geq R_g + R_g\sqrt{1-2\kappa_g d_\lambda}$ (Note that $1-2\kappa_g d_\lambda > 1 - \kappa_g \cdot R_g = 0$). Then since $s_\lambda$varies continuously by λ and $s_0 = 0$, so for all $\lambda \in [0,1]$,

$$s_\lambda \leq R_g - R_g\sqrt{1-2\kappa_g d_\lambda} = \frac{2d_\lambda}{1+\sqrt{1-2\kappa_g d_\lambda}} \tag{A.47}$$

Then in particular,

$$d_M(p,q) = s_1 \leq \frac{2d_1}{1+\sqrt{1-2\kappa_g d_1}} \tag{A.48}$$

$$\leq 2d_1. \tag{A.49}$$

On the other hand, since $c_1(t)$ is geodesic path joining p and q, $c_1([0,s_1]) \subset M_1$. Then since $p = \underset{p' \in M_1}{\arg\min} d(x,p')$, $\|c_1(t)-x\|^2 = (c_1(t)-x)^T(c_1(t)-x)$ is minimized at $t=0$. Hence

$$c'_1(0)^T(p-x) = \frac{d}{dt}(c_1(t)-x)^T(c_1(t)-x)|_{t=0} = 0. \tag{A.50}$$

Let θ be angle between p, x, and q, i.e. $\theta = \angle pxq$. Then $\angle xpq = \frac{\pi}{2} - \frac{\theta}{2}$ and angle between $\overline{pq}$ and T_pM_1 is $\frac{\theta}{2}$, as in figure ?. Now, decompose $q-p$ by components parallel to $p-x$ and orthogonal to $p-x$, i.e.

$$q - p = \Pi_{p-x}(q-p) + (I - \Pi_{p-x})(q-p), \tag{A.51}$$

where $\Pi_{p-x} = \frac{1}{\|p-x\|_{\mathbb{R}^m}^2}(p-x)(p-x)^T$ is projection matrix. Then from $\angle xpq = \frac{\pi}{2} - \frac{\theta}{2}$,

$$\|(I-\Pi_{p-x})(q-p)\|_{\mathbb{R}^m} = \sin\frac{\theta}{2}\|q-p\|_{\mathbb{R}^m}. \tag{A.52}$$

Since $c_1'(0)$ is orthogonal to $p-x$, $c_1'(0)^T\left(\Pi_{p-x}(q-p)\right)=0$, hence

$$|c_1'(0)^T(q-p)| = |c_1'(0)^T(I-\Pi_{p-x})(q-p)| \tag{A.53}$$

$$\leq \|c_1'(0)\|_{\mathbb{R}^m}\|(I-\Pi_{p-x})(q-p)\|_{\mathbb{R}^m} \tag{A.54}$$

$$= \sin\frac{\theta}{2}\|q-p\|_{\mathbb{R}^m}. \tag{A.55}$$

Now, let $v := -\frac{(I-\Pi_{p-x})c_1'(0)}{\|(I-\Pi_{p-x})c_1'(0)\|}$, then

$$v^T c_1'(0) \leq -\sin\frac{\theta}{2}, \tag{A.56}$$

then from

$$(c_1(s_1)-c_1(0))^T v = 0, \tag{A.57}$$

we have

$$0 = \int_0^{s_1} c_1'(0)^T v dt + \int_0^{s_1}\int_0^t c_1''(u)^T v du dt \tag{A.58}$$

$$\leq -s_1\sin\frac{\theta}{2} + \int_0^{s_1}\int_0^t \|c_1''(u)\|_{\mathbb{R}^m} du dt \tag{A.59}$$

$$\leq -s_1\sin\frac{\theta}{2} + \frac{1}{2}\kappa_g s_1^2 \tag{A.60}$$

Therefore we have

$$s_1 \geq 2R_g\sin\frac{\theta}{2} = \frac{R_g}{r}d_1 > 4d_1. \tag{A.61}$$

Hence

$$4d_1 < s_1 \leq 2d_1, \tag{A.62}$$

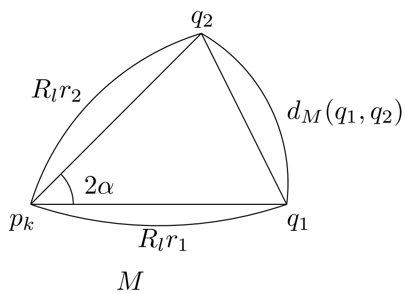
which is a contradiction. ☐

Lemma 17. *(Toponogov comparison theorem, 1959) Let (M,g) be a complete Riemannian manifold with sectional curvature$\geq k$, and let S_k be a surface of constant Gaussian curvature k. Given any geodesic triangle with vertices $p,q,r\in M$ forming an angle α at q, consider a (comparison) triangle with vertices $\bar{p},\bar{q},\bar{r}\in S_k$ such that $d(\bar{p},\bar{q})=d(p,q)$, $d(\bar{r},\bar{q})=d(r,q)$, and $\angle pqr=\angle\bar{p}\bar{q}\bar{r}$. Then $d(p,r)\leq d(q,r)$.*

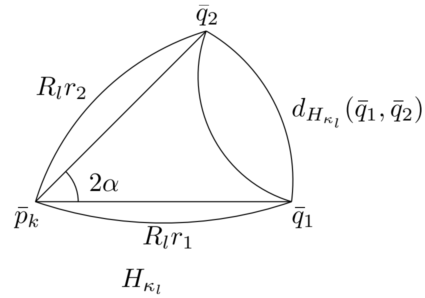
Proof. See Theorem 79 in [Petersen, 2006], p.339. Note that for a manifold with boundary, the complete Riemannian manifold condition can be relaxed to requiring the existence of a geodesic path joining p and q whose image lies on $intM$. ☐

Lemma 18. *(Hyperbolic law of cosines) Let H_κ be a hyperbolic plane whose Gaussian curvature is $-\kappa^2$. Then given a hyperbolic triangle ABC with angles α, β, γ, and side lengths $BC=a$, $CA=b$, and $AB=c$, the following holds:*

$$\cosh(\kappa a) = \cosh(\kappa b)\cosh(\kappa c) - \sinh(\kappa b)\sinh(\kappa c)\cos\alpha. \tag{A.63}$$



(A) triangle in M



(B) comparison triangle in H_{κ_l}

FIGURE A.2. (a) triangle in M formed by p_k, q_1, q_2, and (b) its comparison triangle in H_{κ_l}.

Proof. See 2.13 The Law of Cosines in M^n_κ in [Bridson and Häfliger, 1999], p.24. $\square$

Lemma. *4. Let $M \in \mathcal{M}^d_{\kappa_l,\kappa_g,K_v}$ and let $exp_{p_k} : \mathcal{E}_k \subset \mathbb{R}^m \to \mathcal{M}$ be an exponential map, where $T_{p_k}M$ is identified with $\mathbb{R}^m$. Then for all $v, w \in \mathcal{E}_k \cap B_{\mathbb{R}^d}(0, R_k)$,*

$$\|\exp_{p_k}(v) - \exp_{p_k}(w)\|_{\mathbb{R}^m} \le \frac{e^{\kappa_l R_k} \sinh \kappa_l R_k}{\kappa_l R_k} \|v - w\|_{\mathbb{R}^d}. \tag{A.64}$$

Proof. Let $q_1 = \exp_{p_k}(v)$ and $q_2 = \exp_{p_k}(w)$. Let $d_M(p_k, q_1) = R_l r_1$, $d_M(p_k, q_2) = R_l r_2$, and $\angle q_1 p_k q_2 = 2\alpha$ with $0 \le \alpha \le \pi$. Then

$$\|v - w\|_{\mathbb{R}^{d_1}} = R_l \sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos 2\alpha} \tag{A.65}$$

$$= R_l \sqrt{\sin^2 \alpha (r_1 + r_2)^2 + \cos^2 \alpha (r_1 - r_2)^2}. \tag{A.66}$$

Let H_{κ_l} be a surface of constant sectional curvature $-\kappa_l^2$, and let $\bar{p}_k, \bar{q}_1, \bar{q}_2 \in H_{\kappa_l}$ be such that $d_{H_{\kappa_l}}(\bar{p}_k, \bar{q}_1) = d_M(p_k, q_1)$, $d_{H_{\kappa_l}}(\bar{p}_k, \bar{q}_2) = d_M(p_k, q_2)$, and $\angle \bar{q}_1 \bar{p}_k \bar{q}_2 = \angle q_1 p_k q_2$, so that $\triangle \bar{p}_k \bar{q}_1 \bar{q}_2$ becomes a comparison triangle of $p_k q_1 q_2$, as in Figure A.2. Then since (sectional curvature of $M) \ge -\kappa_l^2$, from the Toponogov comparison theorem in Lemma 17,

$$d_M(q_1, q_2) \le d_{H_{\kappa_l}}(\bar{q}_1, \bar{q}_2). \tag{A.67}$$

From the hyperbolic law of cosines in Lemma 18,

$$\cosh \kappa_l d_{M_{\kappa_l}}(\bar{q}_1, \bar{q}_2) = \cosh r_1 \cosh r_2 - \sinh r_1 \sinh r_2 \cos 2\alpha \tag{A.68}$$

$$= (\sin^2 \alpha) \cosh(r_1 + r_2) + (\cos^2 \alpha) \cosh(r_1 - r_2). \tag{A.69}$$

Therefore,

$$\frac{d_{H_{\kappa_l}}(\bar{q}_1, \bar{q}_2)}{\|v - w\|_{\mathbb{R}^d}} = \frac{\cosh^{-1}\left(\sin^2 \alpha \cosh(r_1 + r_2) + \cos^2 \alpha \cosh(r_1 - r_2)\right)}{\sqrt{(\sin^2 \alpha)(r_1 + r_2)^2 + (\cos^2 \alpha)(r_1 - r_2)^2}}. \tag{A.70}$$

Let $F(a, b, \lambda) = f^{-1}(\lambda f(a) + (1-\lambda) f(b))$ and $G(a, b, \lambda) = g^{-1}(\lambda g(a) + (1-\lambda) g(b))$, with $0 \le a < b$, $\lambda \in [0, 1]$, $f(t) = \cosh t$ and $g(t) = t^2$. Toponogov's theorem implies $F(a, b, \lambda) \ge G(a, b, \lambda)$, and f

and g being strictly increasing function implies $a < G(a,b,\lambda) \le F(a,b,\lambda) < b$. Also,

$$\frac{\partial}{\partial a}\log\frac{F(a,b,\lambda)}{G(a,b,\lambda)} \tag{A.71}$$

$$= \frac{\lambda f'(a)}{f'(F(a,b,\lambda))F(a,b,\lambda)} - \frac{\lambda g'(a)}{g'(G(a,b,\lambda))G(a,b,\lambda)} \tag{A.72}$$

$$= \frac{\lambda}{F(a,b,\lambda)}\exp\left(-\int_a^{F(a,b,\lambda)}(\log f')'(t)dt\right) - \frac{\lambda}{G(a,b,\lambda)}\exp\left(-\int_a^{G(a,b,\lambda)}(\log g')'(t)dt\right). \tag{A.73}$$

Then $(\log f')'(t) = \coth t > (\log g')'(t) = \frac{1}{t}$ and $F(a,b,\lambda) \ge G(a,b,\lambda)$ implies

$$0 < \forall a < b,\ \frac{\partial}{\partial a}\log\frac{F(a,b,\lambda)}{G(a,b,\lambda)} < 0. \tag{A.74}$$

Hence

$$\frac{F(a,b,\lambda)}{G(a,b,\lambda)} \le \frac{F(0,b,\lambda)}{G(0,b,\lambda)} \tag{A.75}$$

and by plugging in $a = |r_1 - r_2|$, $b = r_1 + r_2$, $\lambda = \cos^2\alpha$ implies

$$\frac{d_{H_{\kappa_l}}(\bar{q}_1,\bar{q}_2)}{\|v-w\|_{\mathbb{R}^{d_1}}} \tag{A.76}$$

$$\le \frac{\cosh^{-1}\left(\sin^2\alpha\cosh(r_1+r_2)+\cos^2\alpha\right)}{\sqrt{(r_1+r_2)^2\sin^2\alpha}} \tag{A.77}$$

$$= \frac{\cosh^{-1}\left(1+2\sinh^2\left(\frac{r_1+r_2}{2}\right)\sin^2\alpha\right)}{(r_1+r_2)\sin\alpha} \tag{A.78}$$

$$= \frac{\log\left(1+2\sinh^2\left(\frac{r_1+r_2}{2}\right)\sin^2\alpha + 2\sinh\left(\frac{r_1+r_2}{2}\right)\sin\alpha\sqrt{1+\sinh^2\left(\frac{r_1+r_2}{2}\right)\sin^2\alpha}\right)}{(r_1+r_2)\sin\alpha} \tag{A.79}$$

$$\le \frac{2\sinh\left(\frac{r_1+r_2}{2}\right)\sin\alpha\left[\sinh\left(\frac{r_1+r_2}{2}\right)\sin\alpha+\sqrt{1+\sinh^2\left(\frac{r_1+r_2}{2}\right)\sin^2\alpha}\right]}{(r_1+r_2)\sin\alpha}\ \text{(using }\log(1+x)\le x\text{)} \tag{A.80}$$

$$\le \frac{e^r\sinh r}{r},\ \text{with } r = \frac{r_1+r_2}{2}. \tag{A.81}$$

Then since $t \mapsto \frac{e^t\sinh t}{t}$ is increasing function of t and $r = \frac{r_1+r_2}{2} \le \kappa_l R_k$, so

$$\frac{d_{H_{\kappa_l}}(\bar{q}_1,\bar{q}_2)}{\|v-w\|_{\mathbb{R}^{d_1}}} \le \frac{e^{\kappa_l R_k}\sinh\kappa_l R_k}{\kappa_l R_k}. \tag{A.82}$$

Therefore,

$$\|\exp_{p_k}(v)-\exp_{p_k}(w)\|_{\mathbb{R}^m}\le d_M(q_1,q_2)\le d_{H_{\kappa_l}}(\bar q_1,\bar q_2)\le \frac{e^{\kappa_l R_k}\sinh\kappa_l R_k}{\kappa_l R_k}\|v-w\|_{\mathbb{R}^d}. \tag{A.83}$$

□

Appendix B. Proofs for Section 3

Lemma. *5. Let $X_1,\cdots,X_n\sim P\in\mathcal{P}^{d_2}_{\kappa_l,\kappa_g,K_p,K_v}$, then*

$$P^{(n)}\left[\sum_{i=1}^{n-1}\|X_{i+1}-X_i\|^{d_1}\le L\right]\le \frac{\left(C^{(3,1)}_{K_I,K_p,d_1,d_2,m}\right)^{n-1}L^{\frac{d_2}{d_1}(n-1)}\left(1+\kappa_g^{(m-d_2)(n-1)}\right)}{(n-1)^{\left(\frac{d_2}{d_1}-1\right)(n-1)}(n-1)!}, \tag{B.1}$$

where $C^{(3,1)}_{K_I,K_p,d_1,d_2,m}$ depends only on K_I,K_p,d_1,d_2,m.

Proof. Let $Y_i:=\|X_{i+1}-X_i\|^{d_1}_{\mathbb{R}^m}$, $i=1,\cdots,n-1$. Then

$$P^{(n)}\left(Y_{n-1}\le y|X_1,\cdots,X_{n-1}\right) \tag{B.2}$$

$$=P^{(n)}\left(X_n\in B_{\mathbb{R}^m}\left(X_{n-1},y^{\frac{1}{d_1}}\right)\mid X_1,\cdots,X_{n-1}\right) \tag{B.3}$$

$$=\int_{M\cap\left(B_{\mathbb{R}^m}\left(X_{n-1},y^{\frac{1}{d_1}}\right)\right)}f_{X_n}\left(x_n\right)dvol_M(x_n) \tag{B.4}$$

$$\le K_p vol_M\left(M\cap B\left(X_{n-1},y^{\frac{1}{d_1}}\right)\right) \tag{B.5}$$

$$\le K_pC^{(2)}_{d_2,m}\min\left\{y^{\frac{1}{d_1}},R_g\right\}^{d_2-m}vol_{\mathbb{R}^m}\left(B\left(X_{n-1},y^{\frac{1}{d_1}}+\min\left\{y^{\frac{1}{d_1}},R_g\right\}\right)\right)\ \text{(Lemma 1)} \tag{B.6}$$

$$=K_pC^{(2)}_{d_2,m}\omega_m\left(y^{\frac{d_2}{d_1}}2^mI(y^{\frac{1}{d_1}}\le R_g)+y^{\frac{d_2}{d_1}}\left(\frac{R_g}{y^{\frac{1}{d_1}}}\right)^{d_2-m}\left(1+\left(\frac{R_g}{y^{\frac{1}{d_1}}}\right)\right)^mI(y^{\frac{1}{d_1}}>R_g)\right) \tag{B.7}$$

$$\le K_pC^{(2)}_{d_2,m}\omega_m2^m\left(y^{\frac{d_2}{d_1}}I(y^{\frac{1}{d_1}}\le R_g)+y^{\frac{d_2}{d_1}}\left(\frac{R_g}{2K_I\sqrt{m}}\right)^{d_2-m}I(y^{\frac{1}{d_1}}>R_g)\right) \tag{B.8}$$

$$\le C^{(3,1,1)}_{K_I,K_p,d_2,m}\left(1+\kappa_g^{m-d_2}\right)y^{\frac{d_2}{d_1}} \tag{B.9}$$

where $C^{(3,1,1)}_{K_I,K_p,d_2,m}=K_pC^{(2)}_{d_2,m}\omega_m2^m\left(2K_I\sqrt{m}\right)^{m-d_2}$. Then since $\sum_{i=2}^{n-1}Y_i$ is function of $X_1,\cdots,X_{n-1}$, so

$$P^{(n)}\left(Y_{n-1}\le y|\sum_{i=2}^{n-1}Y_i\right)\le C^{(3,1,1)}_{K_I,K_p,d_2,m}\left(1+\kappa_g^{m-d_2}\right)y^{\frac{d_2}{d_1}}. \tag{B.10}$$

Hence

(B.11)

$$P^{(n)}\left(\sum_{i=1}^{n-1}|X_{i+1}-X_i|^{d_1}\leq L\right)$$

(B.12)

$$=P^{(n)}\left(\sum_{i=1}^{n-1}Y_i\leq L\right)$$

(B.13)

$$=\int_0^L P^{(n)}\left(Y_{n-1}\leq y_{n-1}|\sum_{i=1}^{n-2}Y_i=L-y_{n-1}\right)dF_{\sum\limits_{i=1}^{n-2}Y_i}(L-y_{n-1})$$

(B.14)

$$\leq C^{(3,1,1)}_{K_I,K_p,d_2,m}\left(1+\kappa_g^{m-d_2}\right)\int_0^L y_{n-1}^{\frac{d_2}{d_1}}dF_{\sum\limits_{i=1}^{n-2}Y_i}(L-y_{n-1})$$

(B.15)

$$=C^{(3,1,1)}_{K_I,K_p,d_2,m}\left(1+\kappa_g^{m-d_2}\right)\left(\left[-y_{n-1}^{\frac{d_2}{d_1}}P\left(\sum_{i=1}^{n-2}Y_i\leq L-y_{n-1}\right)\right]_0^L+\int_0^L P\left(\sum_{i=1}^{n-2}Y_i\leq L-y_{n-1}\right)d\left(y_{n-1}^{\frac{d_2}{d_1}}\right)\right)$$

(B.16)

$$=C^{(3,1,1)}_{K_I,K_p,d_2,m}\left(1+\kappa_g^{m-d_2}\right)\int_0^L P\left(\sum_{i=1}^{n-2}Y_i\leq L-y_{n-1}\right)\frac{d_2}{d_1}y_{n-1}^{\frac{d_2-d_1}{d_1}}dy_{n-1}$$

By repeating this argument, we get

(B.17)

$$P^{(n)}\left(\sum_{i=1}^{n-1}|X_{i+1}-X_i|^{d_1}\leq L\right)$$

(B.18)

$$\leq\left(\frac{d_2}{d_1}C^{(3,1,1)}_{K_I,K_p,d_2,m}\left(1+\kappa_g^{m-d_2}\right)\right)^{n-1}\int_{\sum\limits_{i=1}^{n-1}y_i\leq L}\prod_{i=1}^{n-1}y_i^{\frac{d_2-d_1}{d_1}}dy$$

(B.19)

$$\leq\left(\frac{2d_2}{d_1}C^{(3,1,1)}_{K_I,K_p,d_2,m}\right)^{n-1}L^{\frac{d_2}{d_1}(n-1)}\left(1+\kappa_g^{(m-d_2)(n-1)}\right)\int_{\sum\limits_{i=1}^{n-1}y_i\leq 1}\left(\frac{1}{n-1}\sum_{i=1}^{n-1}y_i\right)^{\frac{(d_2-d_1)(n-1)}{d_1}}dy_{n-1}\cdots dy_1$$

(B.20)

$$=\frac{\left(C^{(3,1)}_{K_I,K_p,d_1,d_2,m}\right)^{n-1}L^{\frac{d_2}{d_1}(n-1)}\left(1+\kappa_g^{(m-d_2)(n-1)}\right)}{(n-1)^{\left(\frac{d_2}{d_1}-1\right)(n-1)}}\int_0^1\int_{\sum\limits_{i=1}^{n-2}y_i\leq z}z^{\frac{(d_2-d_1)(n-1)}{d_1}}dy_{n-2}\cdots dy_1dz$$

(B.21)

$$=\frac{\left(C^{(3,1)}_{K_I,K_p,d_1,d_2,m}\right)^{n-1}L^{\frac{d_2}{d_1}(n-1)}\left(1+\kappa_g^{(m-d_2)(n-1)}\right)}{(n-1)^{\left(\frac{d_2}{d_1}-1\right)(n-1)}(n-2)!}\int_0^1 z^{\frac{d_2(n-1)}{d_1}-1}dz$$

(B.22)

$$\leq\frac{\left(C^{(3,1)}_{K_I,K_p,d_1,d_2,m}\right)^{n-1}L^{\frac{d_2}{d_1}(n-1)}\left(1+\kappa_g^{(m-d_2)(n-1)}\right)}{(n-1)^{\left(\frac{d_2}{d_1}-1\right)(n-1)}(n-1)!},$$

where $C^{(3,1)}_{K_I,K_p,d_1,d_2,m}=\frac{2d_2}{d_1}C^{(3,1,1)}_{K_I,K_p,d_2,m}$. □

Lemma 19. *(Space-filling curve) There exists a surjective map $\psi_d:\ \mathbb{R}\to\mathbb{R}^d$ which is Hölder continuous of order $1/d$, i.e.*

$$0\leq\forall s,t\leq 1,\ \|\psi_d(s)-\psi_d(t)\|_{\mathbb{R}^d}\leq 2\sqrt{d+3}|s-t|^{1/d}. \tag{B.23}$$

Such a map is called a space-filling curve.

Proof. See chapter 2.1.6 in [Buchin, 2007]. □

Lemma. *6. Let $M\in\mathcal{M}^{d_1}_{\kappa_l,\kappa_g,K_p}$ and $X_1,\cdots,X_n\in M$. Then there exists $C^{(3,2)}_{K_I,K_v,d_1,m}$ which depends only on m, d_1, K_v, and K_I, and there exists $\sigma\in S_n$ such that*

$$\sum_{i=1}^{n-1}\|X_{\sigma(i+1)}-X_{\sigma(i)}\|^{d_1}_{\mathbb{R}^m}\leq C^{(3,2)}_{K_I,K_v,d_1,m}\left(1+\kappa_g^{m-d_1}\right). \tag{B.24}$$

Proof. When $d_1 = 1$, length of TSP path is bounded by length of curve $vol_M(M)$ as in Figure 3.1, and from Lemma 2 we have $vol_M(M) \leq C^{(2,2)}_{K_I,d,m}\left(1+\kappa_g^{m-1}\right)$, hence $C^{(3,2)}_{K_I,K_v,d_1,m}$ can be set as $C^{(3,2)}_{K_I,K_v,d_1,m} = C^{(2,2)}_{K_I,d,m}$, as described before.

Consider $d_1 > 1$. By scaling the space-filling curve in Lemma 19, there exists a surjective map $\psi_{d_1} : [0,1] \to [-r,r]^{d_1}$ and $\psi_m : [0,1] \to [-K_I, K_I]^m$ that satisfies

$$0 \leq \forall s,t \leq 1,\ \|\psi_{d_1}(s) - \psi_{d_1}(t)\|_{R^{d_1}} \leq 4r\sqrt{d_1+3}|s-t|^{1/d_1} \tag{B.25}$$

$$0 \leq \forall s,t \leq 1,\ \|\psi_m(s) - \psi_m(t)\|_{\mathbb{R}^m} \leq 4K_I\sqrt{m+3}|s-t|^{1/m} \tag{B.26}$$

Let $r := 2\sqrt{3}R_g$. From Lemma 3, M can be covered by N balls of radius r, denoted by $B_M(p_1, r)$, $\cdots$, $B_{\mathcal{M}}(p_N, r)$, with $N = \left\lfloor \frac{2^{d_1} vol_M(M)}{K_v r^{d_1} \omega_{d_1}} \right\rfloor$. Since $\psi_m : [0,1] \to [-K_I, K_I]^m$ is surjective, we can find a right inverse $\Psi_m : [-K_I, K_I]^m \to [0,1]$ that satisfies $\psi_m(\Psi_m(p)) = p$, i.e.

$$\begin{tikzcd} {[0,1]} \arrow[r, bend left, "\psi_m"] & {[-K_I, K_I]^m}. \arrow[l, bend left, "\Psi_m"] \end{tikzcd} \tag{B.27}$$

Reindex p_k so that

$$\Psi_m(p_1) < \cdots < \Psi_m(p_N). \tag{B.28}$$

Now fix k. Then for all $p \in B_M(p_k, r)$, since $d_M(p_k, p) < r$, we can find $\varphi_k(p) \in B_{\mathbb{R}^{d_1}}(0, r)$ such that $\exp_{p_k}(\varphi_k(p)) = p$. So this shows

$$B_M(p_k, r) \subset \exp_{p_k}\left(B_{\mathbb{R}^{d_1}}(0,r)\right). \tag{B.29}$$

Now consider the map $\exp_{p_k} \circ \psi_{d_1} : [0,1] \to M$. Then

$$B_M(p_k, r) \subset \exp_{p_k}\left(B_{\mathbb{R}^{d_1}}(0,r)\right) \subset \exp_{p_k}\left([-r,r]^{d_1}\right) = \exp_{p_k} \circ \psi_{d_1}\left([0,1]\right). \tag{B.30}$$

So $\exp_{p_k} \circ \psi_{d_1} : [0,1] \to M$ is surjective on $B_M(p, r)$, so we can find right inverse $\Psi_k : B_M(p_k, r) \to [0,1]$ that satisfies $(\exp_{p_k} \circ \psi_{d_1})(\Psi_k(p)) = p$, i.e.

$$\begin{tikzcd} {[0,1]} \arrow[r, bend left, "\psi_{d_1}"] & {[-r,r]} \arrow[r, bend left, "\exp_{p_k}"] & M \supset B_M(p_k, r). \arrow[ll, bend left, "\Psi_k"] \end{tikzcd} \tag{B.31}$$

Then, reindex $X_1, \cdots, X_n$ as $\{X_{k,j}\}_{1\le k\le l,\ 1\le j\le n_k}$, where $X_{k,1}, \cdots, X_{k,n_k} \in B_M(p_k, r)$ and $\Psi_k(X_{k,1}) < \cdots < \Psi_k(X_{k,n_i})$. Then for all $1 \le k \le l$,

$$\sum_{j=1}^{n_l-1} \|X_{k,j+1} - X_{k,j}\|_{\mathbb{R}^m}^{d_1} \le \sum_{j=1}^{n_l-1} \|(\exp_{p_k} \circ\psi_{d_1})(\Psi_k(X_{k,j+1})) - (\exp_{p_k} \circ\psi_{d_1})(\Psi_k(X_{k,j}))\|_{\mathbb{R}^m}^{d_1} \tag{B.32}$$

$$\le \left(\frac{e^{\kappa_l r}\sinh \kappa_l r}{\kappa_l r}\right)^{d_1} \sum_{j=1}^{n_k-1} \|\psi_{d_1}(\Psi_k(X_{k,j+1})) - \psi_{d_1}(\Psi_k(X_{k,j}))\|_{\mathbb{R}^{d_1}}^{d_1} \tag{B.33}$$

$$\le \left(\frac{4\sqrt{d_1+3}e^{\kappa_l r}\sinh \kappa_l r}{\kappa_l r}\right)^{d_1} r^{d_1} \sum_{j=1}^{n_k-1} |\Psi_k(X_{k,j+1}) - \Psi_k(X_{k,j})| \tag{B.34}$$

$$\le \left(\frac{4\sqrt{d_1+3}e^{\kappa_l r}\sinh \kappa_l r}{\kappa_l r}\right)^{d_1} r^{d_1}. \tag{B.35}$$

And

$$\sum_{k=1}^{N-1} \|X_{k+1,1} - X_{k,n_l}\|_{\mathbb{R}^m}^{d_1} \tag{B.36}$$

$$\le \sum_{k=1}^{N-1} \left(\|X_{k+1,1} - p_{k+1}\|_{\mathbb{R}^m}^{d_1} + \|p_{k+1} - p_k\|_{\mathbb{R}^m}^{d_1} + \|p_k - X_{k,n_l}\|_{\mathbb{R}^m}^{d_1}\right) \tag{B.37}$$

$$\le 2(N-1)r^{d_1} + \sum_{k=1}^{N-1} \|\psi_m(\Psi_m(p_{k+1})) - \psi_m(\Psi_m(p_k))\|_{\mathbb{R}^{d_1}}^{d_1} \tag{B.38}$$

$$\le 2(N-1)r^{d_1} + 4\sqrt{m+3}K_I \sum_{k=1}^{N-1} |\Psi_m(p_{k+1}) - \Psi_m(p_k)|^{\frac{d_1}{m}} \tag{B.39}$$

$$\le 2(N-1)r^{d_1} + 4\sqrt{m+3}K_I \left(\sum_{k=1}^{N-1} |\Psi_m(p_{k+1}) - \Psi_m(p_k)|^{\frac{d_1}{m}\times\frac{m}{d_1}}\right)^{\frac{d_1}{m}} \left(\sum_{k=1}^{N-1} 1^{\frac{m}{m-d_1}}\right)^{\frac{m-d_1}{m}} \tag{B.40}$$

$$\le 2(N-1)r^{d_1} + 4\sqrt{m+3}K_I (N-1)^{1-\frac{d_1}{m}}, \tag{B.41}$$

where the second from the last inequality comes from Hölder's inequality. Hence, by ordering as $X_{1,1}, \cdots, X_{1,n_1}, X_{2,1}, \cdots, X_{2,n_2}, \cdots, X_{l,1}, \cdots, X_{l,n_2}$,

$$\sum_{i=1}^{n-1} \|X_{\sigma(i+1)} - X_{\sigma(i)}\|_{\mathbb{R}^m}^{d_1} \tag{B.42}$$

$$\le \sum_{k=1}^{N} \sum_{j=1}^{n_l-1} \|X_{k,j+1} - X_{k,j}\|_{\mathbb{R}^m}^{d_1} + \sum_{k=1}^{N-1} \|X_{k+1,1} - X_{k,n_l}\|_{\mathbb{R}^m}^{d_1} \tag{B.43}$$

$$\le N\left(\frac{4\sqrt{d_1+3}e^{\kappa_l r}\sinh \kappa_l r}{\kappa_l r}\right)^{d_1} r^{d_1} + 2(N-1)r^{d_1} + 4\sqrt{m+3}K_I(N-1)^{1-\frac{d_1}{m}}. \tag{B.44}$$

Then, since $\kappa_l r \leq 2\sqrt{3}$ and the fact that $t \mapsto \frac{e^t \sinh t}{t}$ is increasing function on $t \geq 0$, we have

(B.45)
$$\sum_{i=1}^{n-1} \|X_{\sigma(i+1)} - X_{\sigma(i)}\|_{\mathbb{R}^m}^{d_1}$$

(B.46)
$$< \left(\left(\frac{4\sqrt{d_1+3}e^{2\sqrt{3}} \sinh 2\sqrt{3}}{2\sqrt{3}} \right)^{d_1} + 2 \right) r^{d_1} N + 4\sqrt{m+3} K_I N^{1-\frac{d_1}{m}}$$

(B.47)
$$< \frac{2^{d_1} \left(4\sqrt{d_1+3}e^{2\sqrt{3}} \sinh 2\sqrt{3}\right)^{d_1} + 2^{d_1+1}}{K_v \omega_{d_1}} vol_M(M) + \frac{4\sqrt{m+3} K_I 2^{d_1\left(1-\frac{d_1}{m}\right)}}{\left(K_v \left(2\sqrt{3}\right)^{d_1} \omega_{d_1}\right)^{1-\frac{d_1}{m}}} \kappa_g^{d_1\left(1-\frac{d_1}{m}\right)} \left(vol_M(M)\right)^{1-\frac{d_1}{m}}$$

(B.48)
$$\leq C_{K_I,K_v,d_1,m}^{(3,2)} \left(1 + \kappa_g^{m-d_1}\right)$$

by some $C_{K_I,K_v,d_1,m}^{(3,2)}$ which depends only on m, d_1, K_v, and K_I, where the last line comes from inequality in Lemma 2. □

Proposition. *7. Let $1 \leq d_1 < d_2 \leq m$. Then*

$$\inf_{\widehat{\dim}_n} \sup_{P \in \mathcal{P}_{\kappa_l,\kappa_g,K_p,K_v}^{d_1} \cup \mathcal{P}_{\kappa_l,\kappa_g,K_p,K_v}^{d_2}} \mathbb{E}_{P^{(n)}} \left[l\left(\widehat{\dim}_n, \dim(P)\right) \right] \tag{B.49}$$

$$\leq \left(C_{K_I,K_p,K_v,d_1,d_2,m}^{(3,3)}\right)^n \left(1 + \kappa_g^{\left(\frac{d_2}{d_1}m+m-2d_2\right)n}\right) n^{-\left(\frac{d_2}{d_1}-1\right)n}. \tag{B.50}$$

for some $C_{K_I,K_p,K_v,d_1,d_2,m}^{(3,3)}$ that depends only on $K_I, K_p, K_v, d_1, d_2, m$.

Proof. Suppose $X = (X_1, \cdots, X_n) \in I^n$ is observed, then define $\widehat{\dim}(X)$ as

$$\widehat{\dim}_n(X) := \begin{cases} d_1 & if\ \exists \sigma \in S_n\ s.t \sum_{i=1}^{n-1} \|X_{\sigma(i+1)} - X_{\sigma(i)}\|_{\mathbb{R}^m}^{d_1} \leq C_{K_I,K_v,d_1,m}^{(3,2)} \left(1 + \kappa_g^{m-d_1}\right) \\ d_2 & \text{otherwise} \end{cases} \tag{B.51}$$

Then for all $P \in \mathcal{P}_{\kappa_l,\kappa_g,K_p,K_v}^{d_1}$ and $X_1, \cdots, X_n \sim P$, by Lemma 6,

$$\sum_{i=1}^{n-1} \|X_{\sigma(i+1)} - X_{\sigma(i)}\|_{\mathbb{R}^m}^{d_1} \leq C_{K_I,K_v,d_1,m}^{(3,2)} \left(1 + \kappa_g^{m-d_1}\right) \tag{B.52}$$

holds for some $\sigma \in S_n$, hence $\widehat{\dim}_n(X) = d_1 = \dim(P)$, i.e. $P^{(n)} \left[\widehat{\dim}_n(X_1, \cdots, X_n) = d_2\right] = 0$.

On the other hand, for all $P \in \mathcal{P}^{d_2}_{\kappa_l,\kappa_g,K_p,K_v}$,

$$P^{(n)}\left[\widehat{\dim}_n(X_1,\cdots,X_n)=d_1\right] \tag{B.53}$$

$$= P\left[\bigcup_{\sigma\in S_n}\sum_{i=1}^{n-1}|X_{\sigma(i+1)}-X_{\sigma(i)}| \leq C^{(3,2)}_{K_I,K_v,d_1,m}\left(1+\kappa_g^{m-d_1}\right)\right] \tag{B.54}$$

$$\leq \sum_{\sigma\in S_n} P\left[\sum_{i=1}^{n-1}|X_{\sigma(i+1)}-X_{\sigma(i)}| \leq C^{(3,2)}_{K_I,K_v,d_1,m}\left(1+\kappa_g^{m-d_1}\right)\right] \tag{B.55}$$

$$= n!P\left[\sum_{i=1}^{n-1}|X_{i+1}-X_i| \leq C^{(3,2)}_{K_I,K_v,d_1,m}\left(1+\kappa_g^{m-d_1}\right)\right] \tag{B.56}$$

$$= \frac{n\left(C^{(2,2)}_{K_p,d_1,d_2,m}\right)^{n-1}\left(C^{(3,2)}_{K_I,K_v,d_1,m}\left(1+\kappa_g^{m-d_1}\right)\right)^{\frac{d_2}{d_1}(n-1)}\left(1+\kappa_g^{(m-d_2)(n-1)}\right)}{(n-1)^{\left(\frac{d_2}{d_1}-1\right)(n-1)}} \text{ (by Lemma 5).} \tag{B.57}$$

Therefore,

$$\inf_{\widehat{\dim}_n}\sup_{P\in\mathcal{P}^{d_1}_{\kappa_l,\kappa_g,\infty}\cup\mathcal{P}^{d_2}_{\kappa_l,\kappa_g,K_p,K_v}}\mathbb{E}_{P^{(n)}}\left[l\left(\widehat{\dim}_n,\dim(P)\right)\right] \tag{B.58}$$

$$\frac{n\left(C^{(2,2)}_{K_p,d_1,d_2,m}\left(C^{(3,2)}_{K_I,K_v,d_1,m}\right)^{\frac{d_2}{d_1}}\right)^{n-1}\left(1+\kappa_g^{\left(\frac{d_2}{d_1}m+m-2d_2\right)(n-1)}\right)}{(n-1)^{\left(\frac{d_2}{d_1}-1\right)(n-1)}} \tag{B.59}$$

$$\leq \left(C^{(3,3)}_{K_I,K_p,K_v,d_1,d_2,m}\right)^n\left(1+\kappa_g^{\left(\frac{d_2}{d_1}m+m-2d_2\right)n}\right)n^{-\left(\frac{d_2}{d_1}-1\right)n} \tag{B.60}$$

for some $C^{(3,3)}_{K_I,K_p,K_v,d_1,d_2,m}$ that depends only on K_I,K_p,K_v,d_1,d_2,m. $\square$

Appendix C. Proofs for Section 4

Lemma. *9. Let $M \in \mathcal{M}^d_{\kappa_l,\kappa_g}$ be d-dimensional manifold of global curvature $\leq \kappa_g$, local curvature $\leq \kappa_l$, which is imbedded in $\mathbb{R}^{m-\Delta d}$. Then*

$$M \times [-K_I,K_I]^{\Delta d} \in \mathcal{M}^{d+\Delta d}_{\kappa_l,\kappa_g}, \tag{C.1}$$

which is imbedded in $\mathbb{R}^m$.

Proof. Let $x \in \mathbb{R}^m$ be with $d_{\mathbb{R}^m}(x, M\times[-K_I,K_I]^{\Delta d}) < R_g$, and let $\pi_{M\times[-K_I,K_I]^{\Delta d}}(x)$ be its closest point in $M\times[-K_I,K_I]$. We will show that $\pi_{M\times[-K_I,K_I]^{\Delta d}}(x)$ is unique. Then as in Figure C.1,

$$d_{\mathbb{R}^m}\left(x,\pi_{M\times[-K_I,K_I]^{\Delta d}}(x)\right) \geq d_{\mathbb{R}^{m-\Delta d}}\left(\Pi_{1:m-\Delta d}(x),\Pi_{1:m-\Delta d}\left(\pi_{M\times[-K_I,K_I]^{\Delta d}}(x)\right)\right) \tag{C.2}$$

where equality holds iff $\Pi_{(m-\Delta d+1):m}\left(\pi_{M\times[-K_I,K_I]^{\Delta d}}(x)\right) = \Pi_{(m-\Delta d+1):m}(x)$. Also,

$$d_{\mathbb{R}^{m-\Delta d}}\left(\Pi_{1:m-\Delta d}(x),M\right) = d_{\mathbb{R}^m}(x,M\times[-K_I,K_I]^{\Delta d}) < R_g. \tag{C.3}$$

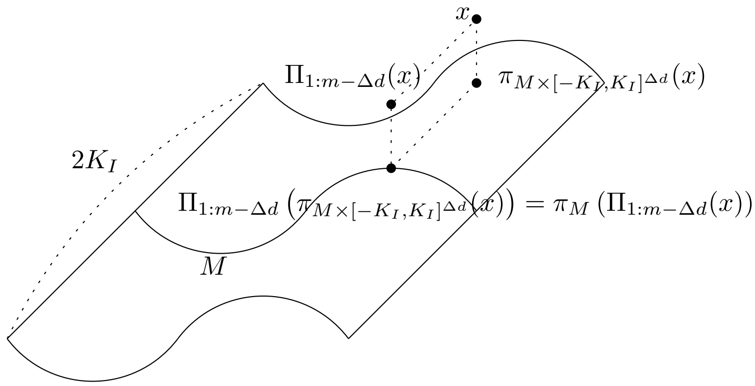


FIGURE C.1. $\pi_{M\times[-K_I,K_I]^{\Delta d}}$ satisfies $\Pi_{1:m-\Delta d}\left(\pi_{M\times[-K_I,K_I]^{\Delta d}}(x)\right) = \pi_M\left(\Pi_{1:m-\Delta d}(x)\right)$.

Hence there uniquely exists $\pi_M\left(\Pi_{1:m-\Delta d}(x)\right) \in M$. And from $\Pi_{1:m-\Delta d}\left(\pi_{M\times[-K_I,K_I]^{\Delta d}}(x)\right) \in M$,

$$d_{\mathbb{R}^{m-\Delta d}}\left(\Pi_{1:m-\Delta d}(x), \Pi_{1:m-\Delta d}\left(\pi_{M\times[-K_I,K_I]^{\Delta d}}(x)\right)\right) \geq d_{\mathbb{R}^{m-\Delta d}}\left(\Pi_{1:m-\Delta d}(x), \pi_M\left(\Pi_{1:m-\Delta d}(x)\right)\right) \tag{C.4}$$

where equality holds iff $\Pi_{1:m-\Delta d}\left(\pi_{M\times[-K_I,K_I]^{\Delta d}}(x)\right) = \pi_M\left(\Pi_{1:m-\Delta d}(x)\right)$, as in Figure C.1. Hence $\pi_{M\times[-K_I,K_I]^{\Delta d}}$ is uniquely determined as $\pi_{M\times[-K_I,K_I]^{\Delta d}}(x) = \left(\pi_M\left(\Pi_{1:m-\Delta d}(x)\right),\ \Pi_{(m-\Delta d+1):m}(x)\right)$. □

Lemma. *10. Let $X : [-K_\delta, K_\delta] \to I$ be a paramatrized curve which is C^1 and piecewise C^2. Suppose that, for all $t \in [-K_\delta, K_\delta]$,*

$$\|X''(t)\| < \|X'(t)\|_2^2 \kappa_l. \tag{C.5}$$

Then image(X) is of local curvature $\leq \kappa_l$.

Proof. $\forall p \in image(X)$, let $\epsilon > 0$ be sufficiently small and $U_p = B(p,\epsilon) \cap image(X)$ be an ϵ-neighborhood of p. Let $U_p = X(a,b)$, and $x \in \mathbb{R}^m$ be such that $d(x, U_p) < R_l - \epsilon$. Then $\forall t \in (a,b)$, if $X''(t)$ exists,

$$\frac{d}{dt}(X(t)-x)^T(X(t)-x) = X'(t)^T(X(t)-x) \tag{C.6}$$

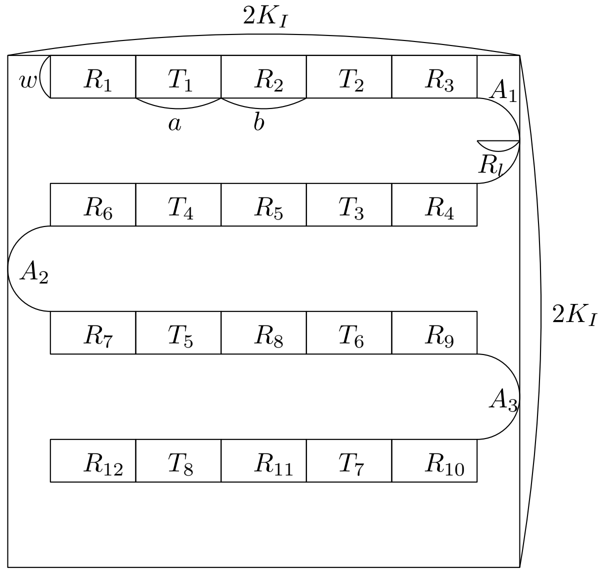
$$\frac{d^2}{dt^2}(X(t)-x)^T(X(t)-x)|_{t=t_0} = X''(t_0)^T(X(t_0)-x) + \|X'(t_0)\|_2^2 \tag{C.7}$$

$$> -\|X''(t_0)\|R^l + \|X'(t_0)\|_2^2 > 0 \tag{C.8}$$

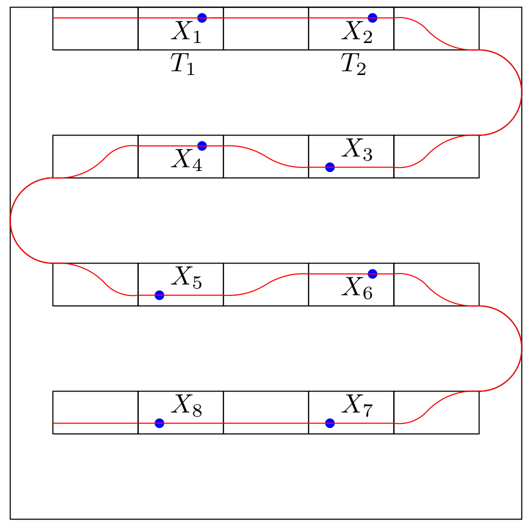
Since X is piecewise C^2, $\|X(t)-x\|_2^2$ is strictly convex function of $t \in (a,b)$. Hence a unique global minimizer t_0 exists, $X(t_0) = \pi_{U_p}(x)$, which is the unique projection of x to U_p. Therefore, *image*(X) is of local curvature $\leq \kappa_l + \epsilon$, for all $\epsilon > 0$. And this asserts that *image*(X) is of local curvature $\leq \kappa_l$. □

Lemma. *12. Suppose $R_l \leq K_I$. There exists $T_1, \cdots, T_n \subset [-K_I, K_I]^{d_2}$ such that*

(1) each T_i's are distinct



(A) alignment of T_i, R_i, and A_i



(B) manifold passing through X_i's

FIGURE C.2. This figure illustrates the case where $d_1 = 1$ and $d_2 = 2$. (A) shows how T_i, R_i, and A_i's are aligned in a zigzag. (B) shows for given $X_1 \in T_1, \cdots, X_n \in T_n$(represented as blue points), how $\mathscr{M}(\{\Pi^{-1}_{(d_1+1):d_2}\Phi_i^{-1}(X_i)\}_{1\leq i\leq n})$(represented as a red curve) passes through $X_1, \cdots, X_n$.

(2) For each T_i, there exists isometry Φ_i such that

$$T_i = \Phi_i\left([-K_I, K_I]^{d_1-1} \times [0,a] \times B_{\mathbb{R}^{d_2-d_1}}(0,w)\right), \tag{C.9}$$

where $c = \left\lceil \frac{K_I+R_l}{2R_l}\right\rceil$, $a = \frac{K_I-R_l}{\left(d+\frac{1}{2}\right)\left\lceil\frac{n}{c^{d_2-d_1}}\right\rceil}$, *and* $w = \min\left\{R_l,\ \frac{d^2(K_I-R_l)^2}{2R_l\left(d+\frac{1}{2}\right)^2\left(\left\lceil\frac{n}{c^{d_2-d_1}}\right\rceil+1\right)^2}\right\}$.

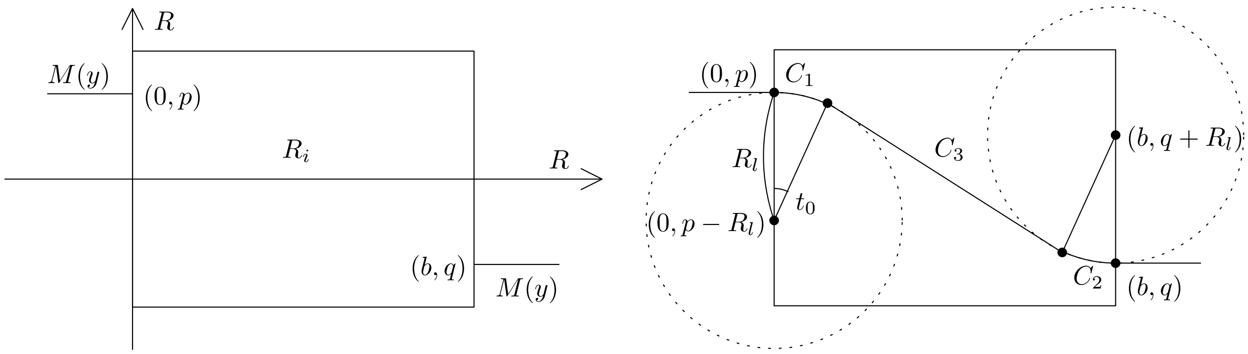
(3) $\exists \mathscr{M} : (B_{\mathbb{R}^{d_2-d_1}}(0,w))^n \to \mathcal{M}^{d_1}_{\kappa_l,\kappa_g,K_v}$ one-to-one such that for each $Y_i \in B_{\mathbb{R}^{d_2-d_1}}(0,w)$, $1 \leq i \leq n$, $\mathscr{M}(Y_1,\cdots,Y_n) \cap T_i = \Phi_i([-K_I,K_I]^{d_1-1} \times [0,a] \times \{Y_i\})$. Hence for any $X_1 \in T_1, \cdots, X_n \in T_n$, $\mathscr{M}(\{\Pi^{-1}_{(d_1+1):d_2}\Phi_i^{-1}(X_i)\}_{1\leq i\leq n})$ passes through $X_1, \cdots, X_n$.

Proof. By Lemma 9, we only need to show the case for $d_1 = 1$. Let $b = \frac{2d(K_I-R_l)}{\left(d+\frac{1}{2}\right)\left(\left\lfloor\frac{n}{c^{d_2-d_1}}\right\rfloor+1\right)}$, so that $b \geq 2\sqrt{2wR_l}$ and $2R_l + \left\lfloor\frac{n}{c^{d_2-d_1}}\right\rfloor a + \left(\left\lfloor\frac{n}{c^{d_2-d_1}}\right\rfloor + 1\right) b = 2K_I$ holds.

With such values of a, b, and w, align T_i, R_i, and A_i in a zigzag; see Figure C.2.

Then from the definition of T_i, it is apparent that (1) the T_i's are distinct and (2) for each T_i, there exists an isometry Φ_i such that $T_i = \Phi_i\left([-K_I,K_I]^{d_1-1} \times [0,a] \times B_{\mathbb{R}^{d_2-d_1}}(0,w)\right)$. There exists isometry Ψ_i such that $R_i = \Psi_i\left([-K_I,K_I]^{d_1-1} \times [0,b] \times B_{\mathbb{R}^{d_2-d_1}}(0,w)\right)$ as well.

Now define $\mathscr{M} : (B_{\mathbb{R}^{d_2-d_1}}(0,w))^n \to \mathcal{M}^{d_1}_{\kappa_l,\kappa_g}$ as follows. For each $Y_i \in B_{\mathbb{R}^{d_2-d_1}}(0,w)$, $1 \leq i \leq n$, $\bigcup_{i=1}^{4} A_i \subset \mathscr{M}(Y_1,\cdots,Y_n) \subset \left(\bigcup_{i=1}^{4} A_i\right) \cup \left(\bigcup_{i=1} T_i\right) \cup \left(\bigcup_{i=1} R_i\right)$. The intersection of $\mathscr{M}(Y_1,\cdots,Y_n)$ and T_i is a line segment $\Phi_i([-K_I,K_I]^{d_1-1} \times [0,a] \times \{Y_i\})$. Our goal is to make $\mathscr{M}(Y_1,\cdots,Y_n)$ be C^1 and piecewise C^2.



(A) A (B) B

FIGURE C.3. (A) We need to find C^2 curve with curvature $\leq R^l$ that starts from $(0,p) \in \mathbb{R}^2$, ends at (b,q), and velocity at each end points are both parallel to $(1,0)$. (B) C_1 and C_2 are arcs of circles of radius R_l, and C_3 is the cotangent segment of two circles.

Given that $\mathscr{M}(Y_1,\cdots,Y_n)\cap\left(\left(\bigcup_{i=1}^{4}A_i\right)\bigcup\left(\bigcup_{i=1}T_i\right)\right)$ is determined, two points on $\mathscr{M}(Y_1,\cdots,Y_n)\cap\partial R_i$ is already determined. By translation and rotation if necessary, $\forall p,q$ with $-w\leq q\leq p\leq w$, we need to find C^2 curve with curvature $\leq R_l$ that starts from $(0,p)\in\mathbb{R}^2$, ends at $(b,q)\in\mathbb{R}^2$, and velocity at each end points are both parallel to $(1,0)\in\mathbb{R}^2$, as in Figure C.3a.

Let

$$t_0 = \cos^{-1}\left(\frac{2R_l\,(2R_l-(p-q))+b\sqrt{b^2-(p-q)\,(4R_l-(p-q))}}{b^2+(2R_l-(p-q))^2}\right), \tag{C.10}$$

and

$$C_1 = \left\{(0,p-R_l)+R_l\,(\sin t,\cos t)\ \mid\ 0\leq t\leq t_0\right\}. \tag{C.11}$$

Then C_1 is an arc of circle of which center is $(0,p-R_l)$, and starts at $(0,p)$ when $t=0$ and ends at $(R_l\sin t_0,\ p-R_l(1-\cos t_0))$ when $t=t_0$. Also, velocity of C_1 at $(0,p)$ is $(1,0)$. Similarily, let

$$C_2 = \left\{(b,q+R_l)-R_l\,(\sin t,\cos t)\ \mid\ 0\leq t\leq t_0\right\}. \tag{C.12}$$

Then C_2 is an arc of a circle of whose center is $(b,q+R_l)$, and starts at (b,q) when $t=0$ and ends at $(b-R_l\sin t_0,\ q+R_l\,(1-\cos t_0))$ when $t=t_0$. Also, the velocity of C_2 at (b,q) is $(-1,0)$. Let

$$C_3 = \left\{(1-s)\,(R_l\sin t_0,\ p-R_l(1-\cos t_0))+s\,(b-R_l\sin t_0,\ q+R_l\,(1-\cos t_0))\ \mid\ 0\leq s\leq 1\right\}, \tag{C.13}$$

so that C_3 is a segment joining $(R_l\sin t_0,\ p-R_l(1-\cos t_0))$ and $(b-R_l\sin t_0,\ q+R_l\,(1-\cos t_0))$. Then,

$$\cos t_0\,(q-p+2R_l\,(1-\cos t_0))+\sin t_0\,(b-2R_l\sin t_0)=0 \tag{C.14}$$

implies that $(b - 2R_l \sin t_0,\ q - p + 2R_l(1 - \cos t_0))$ is parallel to $(\cos t_0, -\sin t_0)$, and hence C_3 is cotangent to both C_1 and C_2. Therefore from Corollary 11, $C_1 \cup C_2 \cup C_3$ is of local curvature $\leq \kappa_l$. Refer to Figure C.3b.

Hence by defining $\mathscr{M}(Y_1, \cdots, Y_n) \cap R_i$ as appropriate translation and rotation of $C_1 \cup C_2$, $\mathscr{M}(Y_1, \cdots, Y_n)$ is of local curvature $\leq \kappa_l$. $\square$

Claim. 13. Let $T = S_n \prod_{i=1}^{n} T_i$. Then for all $x \in intT$, there exists $r_x > 0$ such that for all $r < r_x$,

$$Q_1\left(\prod_{i=1}^{n} B_{\|\|_{\mathbb{R}^{d_2},\infty}}(x_i, r)\right) \geq \frac{2^{(1-d_1)n}}{\kappa_l^{(d_2-d_1)n} K_I^{(d_2-d_1)n}} Q_2\left(\prod_{i=1}^{n} B_{\|\|_{\mathbb{R}^{d_2},\infty}}(x_i, r)\right). \tag{C.15}$$

Proof. By symmetry, we can assume that $x \in \prod_{i=1}^{n} T_i$, i.e. $x_1 \in T_1, \cdots, x_n \in T_n$. Choose r_x small enough so that $B(x, r_x) \subset intU_n$. Then $\forall r < r_x$,

$$Q_1\left(\prod_{i=1}^{n} B_{\|\|_{\mathbb{R}^{d_2},\infty}}(x_i, r)\right) = \int_{\mathcal{P}_1} P^{(n)}\left(\prod_{i=1}^{n} B_{\|\|_{\mathbb{R}^{d_2},\infty}}(x_i, r)\right) d\mu_1(P) \tag{C.16}$$

$$= \int_{C^n} \Phi(y)^{(n)}\left(\prod_{i=1}^{n} B_{\|\|_{\mathbb{R}^{d_2},\infty}}(x_i, r)\right) \lambda_{C^n}(y) \tag{C.17}$$

$$= \int_{C^n} \prod_{i=1}^{n} \lambda_{\mathscr{M}(y)}\left(B_{\|\|_{\mathbb{R}^{d_2},\infty}}(x_i, r)\right) \lambda_{C^n}(y) \tag{C.18}$$

Then since $\mathscr{M}(y) \cap T_i = \Phi_i\left([-K_I, K_I]^{d_1-1} \times [0, a] \times \{y_i\}\right)$,

$$\mathscr{M}(y) \cap B_{\|\|_{\mathbb{R}^{d_2},\infty}}(x_i, r) \tag{C.19}$$

$$= \begin{cases} \Phi_i\left(B_{\|\cdot\|_{\mathbb{R}^{d_2},\infty}}\left(\Pi_{1:d_1}(\Phi_i^{-1}(x_i)),\ r\right) \times \{y_i\}\right) & \text{if } \left\|y_i - \Pi_{(d_1+1):d_2}(\Phi_i^{-1}(x_i))\right\|_{\mathbb{R}^{d_2-d_1}} < r \\ \emptyset & \text{otherwise.} \end{cases} \tag{C.20}$$

Hence

$$\lambda_{\mathscr{M}(y)}\left(B_{\|\|_{\mathbb{R}^{d_2},\infty}}(x_i, r)\right) = \frac{r^{d_1}}{(2K_I)^{d_1-1} an} I\left(\left\|y_i - \Pi_{(d_1+1):d_2}(\Phi_i^{-1}(x_i))\right\|_{\mathbb{R}^{d_2-d_1},\infty} < r\right) \tag{C.21}$$

and

$$Q_0\left(\prod_{i=1}^{n}B_{\|\|_{\mathbb{R}^{d_2},\infty}}(x_i,r)\right) \tag{C.22}$$

$$= \int_{C^n}\prod_{i=1}^{n}\frac{r^{d_1}}{(2K_I)^{d_1-1}an}I\left(\left\|y_i-\Pi_{(d_1+1):d_2}(\Phi_i^{-1}(x_i))\right\|_{\mathbb{R}^{d_2-d_1},\infty}<r\right)\lambda_{C^n}(y) \tag{C.23}$$

$$= \frac{r^{d_1 n}}{(2K_I)^{(d_1-1)n}(an)^n}\prod_{i=1}^{n}\int_C I\left(\left\|y_i-\Pi_{(d_1+1):d_2}(\Phi_i^{-1}(x_i))\right\|_{\mathbb{R}^{d_2-d_1},\infty}<r\right)\lambda_C(y_i) \tag{C.24}$$

$$= \frac{r^{d_1 n}}{(2K_I)^{(d_1-1)n}(an)^n}\left(\frac{(2r)^{d_2-d_1}}{w^{d_2-d_1}\omega_{d_2-d_1}}\right)^n \tag{C.25}$$

$$= \frac{2^{(d_2-2d_1+1)n}r^{d_2 n}}{K_I^{(d_1-1)n}w^{(d_2-d_1)n}(an)^n\omega_{d_2-d_1}^n} \tag{C.26}$$

$$\geq \frac{2^{(d_2-2d_1+1)n}r^{d_2 n}}{\kappa_l^{(d_2-d_1)n}K_I^{(2d_2-d_1)n}\omega_{d_2-d_1}^n}, \tag{C.27}$$

where the last inequality uses $an \leq c^d K_I \leq \frac{K_I^{d_2-d_1+1}}{R_l^{d_2-d_1}}$ and $w \leq K_I$.

On the other hand, $Q_2\left(\prod_{i=1}^{n}B_{\|\|_{\mathbb{R}^{d_2},\infty}}(x_i,r)\right) = \left(\frac{2r}{2K_I}\right)^{d_2 n} = \frac{r^{d_2 n}}{K_I^{d_2 n}}$, so

$$Q_1\left(\prod_{i=1}^{n}B_{\|\|_{\mathbb{R}^{d_2},\infty}}(x_i,r)\right) \geq \frac{2^{(d_2-2d_1+1)n}}{\kappa_l^{(d_2-d_1)n}K_I^{(d_2-d_1)n}\omega_{d_2-d_1}^n}Q_2\left(\prod_{i=1}^{n}B_{\|\|_{\mathbb{R}^{d_2},\infty}}(x_i,r)\right) \tag{C.28}$$

$$\geq \frac{2^{(1-d_1)n}}{\kappa_l^{(d_2-d_1)n}K_I^{(d_2-d_1)n}}Q_2\left(\prod_{i=1}^{n}B_{\|\|_{\mathbb{R}^{d_2},\infty}}(x_i,r)\right) \tag{C.29}$$

holds. ∎

Proposition. *14. Suppose $R^l < K_I$. Then*

$$\inf_{\widehat{\dim}}\sup_{P\in\mathcal{P}^{d_1}_{\kappa_l,\kappa_g,K_p,K_v}\cup\mathcal{P}^{d_2}_{\kappa_l,\kappa_g,K_p,K_v}}\mathbb{E}_{P^{(n)}}[l(\widehat{\dim}_n,\dim(P))] \tag{C.30}$$

$$\geq \left(C^{(4,1)}_{d_1,d_2,K_I}\right)^n\kappa_l^{-(d_2-d_1)n}\min\left\{\kappa_l^{2(d_2-d_1)+1}n^{-2},1\right\}^{(d_2-d_1)n}, \tag{C.31}$$

for some constant $C^{(4,1)}_{d_1,d_2,K_I}$ that depends only on d_1, d_2, and K_I.

Proof. Let $J=[-K_I,K_I]^{d_2}$. Let S_n be the permutation group, and $S_n \curvearrowright J^n$ by coordinate change, i.e. $\sigma\in S_n$, $x\in J^n$, $\sigma x := (x_{\sigma(1)},\cdots,x_{\sigma(n)})$. For any set $A\subset J^n$, let $S_nA := \{\sigma x\in J^n:\ \sigma\in S_n,\ x\in A\}$.

Let T_i be T_i's from Lemma 12. Let $T := S_n\prod_{i=1}^{n}T_i$, and $V := \bigcup_{i=1}^{n}T_i = \Pi_{1:d_2}(T)$. (Intuitively, T is the set of points $x=(x_1,\cdots,x_n)$ where x_i lies on one of the T_j.)

Let $C = B_{\mathbb{R}^{d_2-d_1}}(0,w)$, and let $\mathcal{P}_1 = \{P\in\mathcal{P}^{d_1}_{\kappa_l,\kappa_g}$: there $\exists M\in\mathscr{M}(C^n)$ such that P is uniform on $M\}$, and let $\mathcal{P}_2 = \{\lambda_J\}\subset\mathcal{P}^{d_2}_{\kappa_l,\kappa_g}$.

Define $\Phi : C^n \to \mathcal{P}_1$ by $\Phi(y_1, \cdots, y_n) = \lambda_{\mathscr{M}(y_1,\cdots,y_n)}$, i.e. the uniform measure on $\mathscr{M}(y_1, \cdots, y_n)$. Impose a topology and probability measure structure on $\mathcal{P}_1$ by the pushforward topology and the uniform measure on C^n, i.e. $\mathcal{P}' \subset \mathcal{P}_1$ is open iff $\Phi^{-1}(\mathcal{P}')$ is open in C^n, $\mathcal{P}' \subset \mathcal{P}_1$ is measurable iff $\Phi^{-1}(\mathcal{P}') \in \mathcal{B}(C^n)$, and $\mu_1(\mathcal{P}') = \lambda_{C^n}(\Phi^{-1}(\mathcal{P}'))$.

Define a probability measure Q_1, Q_2 on $(J^n, \mathcal{B}(J^n))$ by $Q_1(A) := \int_{\mathcal{P}_1} P^{(n)}(A)d\mu_1(P)$ and $Q_2 = \lambda_{J^n}$. Fix $P \in \mathcal{P}_1$, let $x = \Phi^{-1}(P)$. Then $P^{(n)}(A) = \lambda^{(n)}_{\mathscr{M}(x)}(A)$ is a measurable function of x and Φ is a homeomorphism. Hence, $p^{(n)}(A)$ is measurable function and $Q_1(A)$ is well defined. Define $\nu = Q_1 + \lambda_J$. Then $Q_1,\ Q_2 \ll \nu$, so there exist densities $q_1,\ q_2$ with respect to ν.

Then from Claim 13, $\forall x \in intT$, $\exists r_x > 0$ s.t. $\forall r < r_x$,

$$Q_1\left(\prod_{i=1}^{n} B_{\|\|_{\mathbb{R}^{d_2},\infty}}(x_i, r)\right) \geq \frac{2^{(1-d_1)n}}{\kappa_l^{(d_2-d_1)n} K_I^{(d_2-d_1)n}} Q_2\left(\prod_{i=1}^{n} B_{\|\|_{\mathbb{R}^{d_2},\infty}}(x_i, r)\right). \tag{C.32}$$

Hence q_1 satisfies $q_1(x) \geq \frac{2^{(d_2-2d_1+1)n}}{\kappa_l^{(d_2-d_1)n} K_I^{(d_2-d_1)n} \omega_{d_2-d_1}} q_2(x)$ if $x \in T$ (and $q_1(x) = 0$ if $x \notin T$). Then,

$$\frac{Q_1(\psi \neq 0) + Q_2(\psi \neq 1)}{2} \geq \frac{1}{4}\int_T \min\left\{\frac{2^{(1-d_1)n}}{\kappa_l^{(d_2-d_1)n} K_I^{(d_2-d_1)n}}, 1\right\} q_2(x)d\nu(x) \text{ (by lemma)} \tag{C.33}$$

$$= \frac{2^{(1-d_1)n-2}}{\kappa_l^{(d_2-d_1)n} K_I^{(d_2-d_1)n}} \lambda_{J^n}(T) \tag{C.34}$$

Then from $a = \frac{K_I - R_l}{\left(d+\frac{1}{2}\right)\left\lceil \frac{n}{c^{d_2-d_1}} \right\rceil}$ and $w = \min\left\{R_l,\ \frac{d^2(K_I-R_l)^2}{2R_l\left(d+\frac{1}{2}\right)^2\left(\left\lceil \frac{n}{c^{d_2-d_1}} \right\rceil + 1\right)^2}\right\}$

$$\lambda_{J^n}\left(S_n \prod_{i=1}^{n} T_i\right) = n!\lambda_{J^1}(T_1)^n = n!\left(\frac{(2K_I)^{d_1-1}\omega_{d_2-d_1} a w^{d_2-d_1}}{(2K_I)^{d_2}}\right)^n \tag{C.35}$$

$$\geq \left(C^{(4,1,1)}_{d_1,d_2,K_I}\right)^n \min\left\{\kappa_l^{2(d_2-d_1)+1} n^{-2}, 1\right\}^{(d_2-d_1)n}, \tag{C.36}$$

for some constant $C^{(4,1,1)}_{d_1,d_2,K_I}$ that depends only on d_1, d_2, and K_I. Hence

$$\inf_{\widehat{\dim}} \sup_{P \in \mathcal{P}_1 \cup \mathcal{P}_2} \mathbb{E}_P[l(\widehat{\dim}_n, \dim(P))] \geq \left(C^{(4,1)}_{d_1,d_2,K_I}\right)^n \kappa_l^{-(d_2-d_1)n} \min\left\{\kappa_l^{2(d_2-d_1)+1} n^{-2}, 1\right\}^{(d_2-d_1)n}, \tag{C.37}$$

for some constant $C^{(4,1)}_{d_1,d_2,K_I}$ that depends only on d_1, d_2, and K_I. Then

$$\inf_{\widehat{\dim}} \sup_{P \in \mathcal{P}^{d_1}_{\kappa_l,\kappa_g,K_p,K_v} \cup \mathcal{P}^{d_2}_{\kappa_l,\kappa_g,K_p,K_v}} \mathbb{E}_P[l(\widehat{\dim}_n, \dim(P))] \geq \inf_{\widehat{\dim}} \sup_{P \in \mathcal{P}_1 \cup \mathcal{P}_2} \mathbb{E}_P[l(\widehat{\dim}_n, \dim(P))], \tag{C.38}$$

which completes the proof. □

Appendix D. Proofs For Secton 5

Proposition. *15.*

$$\inf_{\widehat{\dim}_n} \sup_{P \in \mathcal{P}} \mathbb{E}_{P^{(n)}}\left[l\left(\widehat{\dim}_n, \dim(P)\right)\right] \leq \left(C_{K_I,K_p,K_v,K_m}\right)^n \left(1 + \kappa_g^{(m^2-m)n}\right) n^{-\frac{1}{m-1}n} \tag{D.1}$$

for some $C^{(5,1)}_{K_I,K_p,K_v,K_m}$ that depends only on K_I, K_p, K_v, m.

Proof. define $\widehat{\dim}(X)$ as

(D.2)
$$\widehat{\dim}_n(X) := \min\left\{ d \in [1,m] : \ \exists \sigma \in S_n \ s.t \sum_{i=1}^{n-1} \|X_{\sigma(i+1)} - X_{\sigma(i)}\|^d_{\mathbb{R}^m} \le C^{(3,2)}_{K_I,K_v,d,m}\left(1+\kappa_g^{m-d}\right)\right\}.$$

Then for all $P \in \mathcal{P}^d_{\kappa_l,\kappa_g,K_p,K_v}$ and $X_1, \cdots, X_n \sim P$, by Lemma 6,

$$\sum_{i=1}^{n-1} \|X_{\sigma(i+1)} - X_{\sigma(i)}\|^d_{\mathbb{R}^m} \le C^{(3,2)}_{K_I,K_v,d,m}\left(1+\kappa_g^{m-d}\right) \tag{D.3}$$

holds for some $\sigma \in S_n$, hence

$$\widehat{\dim}_n(X) \le d = \dim(P). \tag{D.4}$$

Therefore,

(D.5)
$$P^{(n)}\left[\widehat{\dim}_n(X_1, \cdots, X_n) \neq d\right]$$
(D.6)
$$= P^{(n)}\left[\max\left\{ d \in [1,m] : \ \exists \sigma \in S_n \ s.t \sum_{i=1}^{n-1} \|X_{\sigma(i+1)} - X_{\sigma(i)}\|^d_{\mathbb{R}^m} \le C^{(3,2)}_{K_I,K_v,d,m}\left(1+\kappa_g^{m-d}\right)\right\} < d\right]$$
(D.7)
$$\le \sum_{k=1}^{d-1} P^{(n)}\left[\exists \sigma \in S_n \ s.t \sum_{i=1}^{n-1} \|X_{\sigma(i+1)} - X_{\sigma(i)}\|^k_{\mathbb{R}^m} \le C^{(3,2)}_{K_I,K_v,k,m}\left(1+\kappa_g^{m-k}\right)\right]$$
(D.8)
$$\le \sum_{k=1}^{d-1} \left(C^{(3,3)}_{K_I,K_p,K_v,k,d,m}\right)^n \left(1+\kappa_g^{\left(\frac{d}{k}m+m-2d\right)n}\right) n^{-\left(\frac{d}{k}-1\right)n}$$
(D.9)
$$\le \left(C^{(5,1)}_{K_I,K_p,K_v,m}\right)^n \left(1+\kappa_g^{(m^2-m)n}\right) n^{-\frac{1}{m-1}n},$$

for some $C^{(5,1)}_{K_I,K_p,K_v,m}$ that depends only on K_I, K_p, K_v, m. Therefore,

$$\inf_{\widehat{\dim}_n} \sup_{P\in\mathcal{P}} \mathbb{E}_{P^{(n)}}\left[l\left(\widehat{\dim}_n, \dim(P)\right)\right] \le \left(C^{(5,1)}_{K_I,K_p,K_v,m}\right)^n \left(1+\kappa_g^{(m^2-m)n}\right) n^{-\frac{1}{m-1}n}. \tag{D.10}$$

□

Proposition. *16. Suppose $R_l < K_I$, then*

$$\inf_{\widehat{\dim}} \sup_{P\in\mathcal{P}} \mathbb{E}_{P^{(n)}}[l(\widehat{\dim}_n, \dim(P))] \ge \left(C^{(5,2)}_{K_I}\right)^n \kappa_l^{-n} \min\left\{\kappa_l^3 n^{-2}, 1\right\}^n \tag{D.11}$$

for some $C^{(5,2)}_{K_I}$ that depends only on K_I.

Proof. For any d_1 and d_2, from Proposition 14,

$$\inf_{\widehat{\dim}} \sup_{P\in\mathcal{P}} \mathbb{E}_{P^{(n)}}[l(\widehat{\dim}_n, \dim(P))] \geq \inf_{\widehat{\dim}} \sup_{P\in\mathcal{P}^{d_1}_{\kappa_l,\kappa_g,K_p,K_v}\cup\mathcal{P}^{d_2}_{\kappa_l,\kappa_g,K_p,K_v}} \mathbb{E}_{P^{(n)}}[l(\widehat{\dim}_n, \dim(P))] \tag{D.12}$$

$$\geq \left(C^{(4,1)}_{d_1,d_2,K_I}\right)^n \kappa_l^{-(d_2-d_1)n} \min\left\{\kappa_l^{2(d_2-d_1)+1} n^{-2}, 1\right\}^{(d_2-d_1)n} \tag{D.13}$$

Hence by plugging in $d_1 = 1$ and $d_2 = 2$, we have

$$\inf_{\widehat{\dim}} \sup_{P\in\mathcal{P}} \mathbb{E}_{P^{(n)}}[l(\widehat{\dim}_n, \dim(P))] \geq \left(C^{(5,2)}_{K_I}\right)^n \kappa_l^{-n} \min\left\{\kappa_l^{3} n^{-2}, 1\right\}^{n} \tag{D.14}$$

with $C^{(5,2)}_{K_I} = C^{(4,1)}_{d_1=1,d_2=2,K_I}$.

$\square$