THREE VARIATIONS ON JOHNSON-LINDENSTRAUSS MAPS FOR SUBMANIFOLDS OF EUCLIDEAN SPACE VIA REACH

By

Arman Tavakoli

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Mathematics – Doctor of Philosophy

2021

ABSTRACT

THREE VARIATIONS ON JOHNSON-LINDENSTRAUSS MAPS FOR SUBMANIFOLDS OF EUCLIDEAN SPACE VIA REACH

By

Arman Tavakoli

In this thesis we investigate 3 variations of the classical Johnson-Lindenstrauss (JL) maps. In one direction we build on the earlier work of Wakin and Eftekhari (2015), by considering generalizations to manifolds with boundary. In a second direction we extend the work of Noga Alon (2003) for lower bounds for the final embedding dimension in JL maps. In the third direction, we consider matrices with fast matrix-vector multiply and improve the run-time in the earlier work of Oymak, Recht and Soltanolkotabi (2018), and Ailon and Liberty (2009).

This thesis is organized into 6 chapters. The three variations are discussed in chapters 4, 5 and 6. The variation for manifolds with boundary is presented in chapter 4. The lower bound problem is discussed in chapter 5, and chapter 6 is regarding the run-time improvements. The first chapter is an introduction to Johnson-Lindenstrauss maps. The second chapter is about a regularity parameter called reach and geometrical estimates for manifolds. The third chapter is regarding two geometry questions about reach that arise from the discussions in chapter 2.

ACKNOWLEDGEMENTS

I would like to thank my advisers Profs. Mark Iwen and Ben Schmidt for their kind support and generous spirit. They listened to me express my own ideas, and helped me make progress with my research. They shared their experience and knowledge with me and I am grateful for that. I thank Profs. Rongrong Wang and Willie Wong for being on my dissertation committee and providing feedback about my work.

I am indebted to many people at MSU for providing a supportive community that improved my experience. Sami Mehri generously hosted me in my first visit at MSU; Emmet Harrington and Gorapada Bera were my amazing office mates who shared many mathematical discussions with me. Keshav Sutrave was there for me on numerous occasions. I am thankful to Tingting Jian, Lijie Han, Xiangyu (Orange) Wang, Xiaoying Yuan for many cakes, pastor John Bell, the Greek squad (Michalis Paparizos, Georgios Psaromiligkos, Christos Grigoriadis, Ioannis Zachos, Dimitris Vardakis), Zhe Zhang, Dan Normand, Timothy Szocinski, Minh Hoang Le, Yesfun Yeh, Anand Chandrasekhar, Henry Manoj D'Souza, Nopphachai Sirinakbumrung, Liwei Deng, Farzan Masrour, Hossein Mazaheri, Vojtěch Kejzlar, and Nayeli Rodriguez Briones for hosting me on multiple occasions.

My family always supported me and helped me recharge during the holidays. To my parents Fara and Mohammad Hossein, my sisters Aylin and Iren, and little (at the time of writing) Ethan, I love you very much.

During the first year of the Covid-19 pandemic, I could stay home and work remotely. Yet many front line workers specially in health care services, provided service to public everyday while risking their lives. This was at a time where limited information was available about Covid-19 and there were no vaccines. I am grateful to the numerous people in the front lines who made life safer for others.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

In this thesis we study different variations of Johnson-Lindenstrauss (JL) maps. A great source for learning about such maps is the textbook of Roman Vershynin, *High Dimensional Probability*, [35, Chapter 5]. The variations we consider are based on the theme of improving existing bounds for JL maps, but they come in different flavors such as sufficient conditions, necessary conditions or runtime time complexity bounds.

In this chapter we give an introduction to JL maps and discuss some of the tools that are used in their construction.

1.1 Johnson-Lindenstrauss Maps

JL maps belong to the general field of dimension reduction (also known as dimensionality reduction). Informally, given T as a subset of $\mathbb{R}^N$, a JL map associated to T is a map that reduces N while preserving the "geometry of T", i.e. $L : \mathbb{R}^N \to \mathbb{R}^m$ with $m < N$ such that T and $L(T)$ have the "same geometry". There are many possible interpretations for a rigorous replacement for preserving the geometry. Perhaps the most strict requirement would be to ask for preservation of all pairwise Euclidean distances; that would mean for $x, y \in T$,

$$||L(x-y)||_2 = ||x-y||_2 \tag{1.1}$$

where $||.||_2$ is the Euclidean norm. For an example one can consider L to be the projection onto the span of T. If T is a finite set, then its span is at most a $|T|$-dimensional subspace. If $|T| < N$, L would satisfy the above requirements and it would be the simplest example of a map that "preserves the geometry" of T while reducing its ambient dimension. In this example the final embedding dimension is at most the number of points. With condition (1.1), this number is also necessary; consider n points with all the pairwise distances equal to 1. Such an arrangement can be embedded into $\mathbb{R}^N$ for N large enough, but it can not be found when the dimension is strictly less than $n-1$. For

$n = 3, 4$ this is the familiar fact that an equilateral triangle doesn't fit in $\mathbb{R}^1$ or a regular tetrahedron doesn't fit in $\mathbb{R}^2$. Relaxing (1.1) allows us to reduce the ambient dimension more.

One could be less strict and allow the map to distort the pairwise distances, but in a controlled manner. In this way, given $0 < \epsilon < 1$ and $x, y \in T$ we require

$$(1-\epsilon)||x-y|| \leq ||L(x-y)|| \leq (1+\epsilon)||x-y||. \tag{1.2}$$

Inequality (1.2) is known as the JL condition. In this sense, if we consider T as a metric space with the chordal metric, the JL map is a metric space embedding. With condition (1.2), we have a celebrated theorem known as the Johnson-Lindenstrauss lemma.

Theorem 1.1.1. *[35, Theorem 5.3.1] Let T be a set of n points in $\mathbb{R}^N$ and $\epsilon > 0$. Let P be an orthogonal projection onto a random m-dimensional subspace in $\mathbb{R}^N$ selected uniformly in the Grassmanian $G_{N,m}$. Then there are universal constants c_1, c_2 such that if*

$$m \geq c_1 \frac{\log(n)}{\epsilon^2}. \tag{1.3}$$

then with probability at least $1 - 2\exp(-c_2\epsilon^2 m)$, the scaled projection

$$Q = \sqrt{\frac{N}{m}} P$$

satisfies

$$(1-\epsilon)||x-y|| \leq ||Q(x-y)|| \leq (1+\epsilon)||x-y||. \tag{1.4}$$

The main tool for proving this theorem is the concentration of measure. In this proof technique one estimates the expected length of a vector under all projections. One calculates the averaged length of that vector when it is projected onto all possible directions. Then by the concentration of measure, most projections do almost the same as the average, with a small probability for deviation. Having obtained a failure probability for a single vector, we apply it $\binom{n}{2}$ times to get an estimate for all pairwise secants. We use the union bound to combine the the failure probabilities over the different vectors and we get the claimed result.

The bound (1.3) is logarithmic in the number of points. This is a improvement compared to the discussion for (1.1) where we saw a linear bound. The price to pay for this logarithmic improvement is to tolerate the $\frac{1}{\epsilon^2}$ factor coming from the distortion ϵ. The break even point would be $\epsilon = \sqrt{\frac{c_1 \log(n)}{n}}$. As long as we allow ϵ to be larger that this threshold, the JL lemma offers a great improvement.

The bound (1.3) is also independent of N. This property makes the JL maps robust for applications in dimension reduction since it allows them to handle models that encode their data in spaces with large N.

A remarkable property of the above theorem is that it is possible to design the JL map *before* knowing the data (the n points). This property is known as being *oblivious*, and is essentially the result of introducing probability and allowing for a controllable chance of failure. Since the expressions for the probability only depend on the number of points, and not where they are, one can pick a map from all the possible projections before we know the data, and with high probability (1.2) would still be satisfied. The obliviousness makes this technique very useful and robust in applications since the dimension reduction technique becomes decoupled from the intricacies of the data.

The next step is to construct a JL map for infinite points. The theorem mentioned above depends explicitly on the number of points. However it is possible to give a different variation, known as the matrix deviation inequality [35, theorem 9.1.1], where the embedding dimension depends on how the vectors are spread out in different directions, through a quantity called Gaussian width (we will describe it below), or a variation of it called Gaussian complexity. The deviation inequality also differs from the JL lemma discussed above because it works with Gaussian random matrices instead of random projections. Informally given $T \subset \mathbb{R}^N$, the matrix deviation inequality says that with high probability

$$||Ax|| = \mathbb{E}||Ax|| + O(\gamma(T)), \qquad \text{for all } x \in T \tag{1.5}$$

where A is a Gaussian random matrix, $\mathbb{E}$ is the expectation (average) operator and $\gamma(T)$ is the Gaussian complexity of T. Conceptually (1.5) tells us that the length of a vector in a set T after its is mapped by a random a matrix A is, with a high chance, closer than $O(\gamma(T))$ to the average length. Compared to the earlier discussion about pairwise distances, this perspective is different because we only consider individual vectors. To transition from one to the other, we form a set of all pairwise secants and work with that set of vectors.

A quantitative version of the matrix deviation inequality, with tail bounds, is as follows.

Theorem 1.1.2. *[35, theorem 9.1.1 and 9.1.8] Let A be a $m \times N$ matrix whose rows A_i are independent isotropic and sub-Gaussian random vectors in $\mathbb{R}^N$. Let $K = \max_i ||A_i||_{\psi_2}$ where ψ_2 is the sub-Gaussian norm. Let $T \subset \mathbb{R}^N$ with Gaussian width $\omega(T)$. Then there is a universal constant C such that for $u \geq 0$, the probability of event*

$$\sup_{x\in T} \left| ||Ax|| - \sqrt{m}||x|| \right| \leq CK^2[\omega(T) + u\ rad(T)] \tag{1.6}$$

is at least $1 - 2\exp(-u^2)$.

Here $\mathrm{rad}(T)$ is the radius of the smallest ball that can contain T. Again in this theorem we transition from earlier pairwise distances $x - y$ in theorem 1.1.1 to individual vectors with x only. We note that all vectors of the form $\frac{x-y}{||x-y||}$ are put into a set T and theorem 1.1.2 is applied to this set. The proof of this theorem is based on the generic chaining technique [35, chapter 9]. The term $\sqrt{m}||x||$ acts as the average $\mathbb{E}||Ax||$. The right hand sides is the deviation from the mean. To achieve a higher success probability, we must allow for a larger deviation.

If we have a choice over m, by dividing both sides of (1.6) by $\sqrt{m}$ and choosing m large enough we can bound the deviation to an error ϵ of our choosing. In this sense, a large enough m allows us to construct an oblivious JL map for any set in terms of its Gaussian width.

Now we discuss the Gaussian width. We present two ways to think about it. One way is using probability and Gaussian random variables. The second way is geometrical, and doesn't use probability. Having the different view points allows us to switch from one to the other depending on the application. The probabilistic definition is as follows.

Definition 1.1.3. *[35, Definiiton 7.5.1] Let g be a standard Gaussian random vector,* $g \sim N(0, I_N)$. *The Gaussian width of a subset* $T \subset \mathbb{R}^N$ *is defined as*

$$\omega(T) = \mathbb{E} \sup_{x \in T} \langle g, x \rangle. \tag{1.7}$$

With an absolute value $\mathbb{E} \sup_{x \in T} |\langle g, x \rangle|$, one gets the Gaussian complexity of T, denoted by $\gamma(T)$. In general $\omega(T) \leq \gamma(T)$ but if T is origin-symmetric then $\omega(T) = \gamma(T)$.

To get an interpretation of how Gaussian width measures the width of a set, first we need the concept of spherical width. Then we show how it relates to Gaussian width.

Definition 1.1.4. *[35, Definition 7.5.5] The spherical width of a subset* $T \subset \mathbb{R}^N$ *is defined as follows where* θ *is a direction uniformly distributed in* $\mathbb{S}^{N-1}$, $\theta \sim$ *Unif*(S^{N-1}).

$$\omega_s(T) = \mathbb{E} \sup_{x \in T} \langle \theta, x \rangle \tag{1.8}$$

Figure 1.1 shows how width of a set in a particular direction is measured. When this directional width is averaged over all possible directions we get the spherical width.

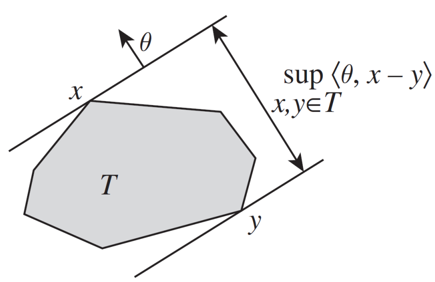


Figure 1.1: Two parallel hyper-planes containing a set determine its width in that direction. From [35].

For a Gaussian random vector, one can consider its direction θ and length r separately. Direction θ is uniformly distributed and decouples from r. Next we recall that the expected length of Gaussian random variable differs from $\sqrt{N}$ by a universal constant, $|\mathbb{E}||g|| - \sqrt{N}| \leq C$. This lets us relate the Gaussian width to spherical width as follows.

$$(\sqrt{N} - C)\omega_s(T) \leq \omega(T) \leq (\sqrt{N} + C)\omega_s(T) \tag{1.9}$$

In this sense Gaussian width is the spherical width of a set scaled by $\sqrt{N}$, and spherical width has a clear geometrical interpretation.

We give 4 examples of Gaussian widths that are often encountered in high dimensional probability, see [35, 7.5.8-7.5.10]. They are the L^1 ball B_1^N, the L^2 ball B_2^N, the L^∞ ball B_∞^N, and a set T with n points. We use C for any universal constant.

$$C_1\sqrt{\log(N)} \leq \omega(B_1^n) \leq C_2\sqrt{\log(N)} \tag{1.10}$$

$$|\omega(B_2^N) - \sqrt{N}| \leq C \tag{1.11}$$

$$\omega(B_\infty^n) = \sqrt{\frac{2}{\pi}}N \tag{1.12}$$

$$\omega(T) \leq C\sqrt{\log(n)}\,\mathrm{diam}(T) \tag{1.13}$$

Dividing the first three above expressions by $\sqrt{N}$ gives us spherical width estimates for the corresponding L^p balls. One sees that the width of L^2 ball is a constant as expected, while the width of the L^1 ball decreases to zero as $\sqrt{\frac{\log(N)}{N}}$ while the width of the L^∞ ball grows as $\sqrt{N}$.

With Gaussian width, theorem 1.1.2 lets us get JL maps for infinite sets. However one still needs techniques for estimating the Gaussian width. In this thesis, we restrict our discussion to compact smooth manifolds. To get a JL map for a compact manifold we need estimates for the Gaussian width of its set of unit secants. The regularity provided by being a manifold allows us to find explicit estimates and thus close the argument for JL maps of this class of objects. More about this estimate is in section 2.6.

The key inequality that allows us to relate geometrical properties of a manifold to its Gaussian width is called Dudley's inequality [18, page 226]. For a universal constant c, one can relate covering number of set T to its Gaussian width as follows.

$$\omega(T) \leq c \int_0^\infty \sqrt{\log(C(T,\epsilon))} d\epsilon. \tag{1.14}$$

Motivated by this inequality we calculate various covering numbers in chapter 2.

1.2 Fast Matrices

Having discussed random projections and Gaussian matrices for dimension reduction, the next step would be to speed up the computation. For a $m \times N$ matrix, the general matrix-vector multiplication takes $O(mN)$ operations. Since Gaussian matrices are densely populated with random variables this is difficult to improve.

If instead we specialize to discrete Fourier matrices, one can use the Fast Fourier Transform (FFT) algorithm. FFT allows for a $O(N \log(N))$ matrix-vector multiplication, for square matrices. Therefore the Discrete Fourier Transform (DFT) matrices are ideal candidates for fast JL maps. Other examples in this class are also the Welsh-Hadamard matrices. However for JL maps we need short and wide matrices with $m \leq N$. This will come from sub-sampling the Fourier matrix. Also since Fourier matrices have a fast matrix-vector multiplication, one can do the full multiplication and throw away the unwanted rows. Constructing JL maps from Fourier matrices requires a process that we briefly describe below. Such an approach has been used before in earlier works. One example is the work in [29] where the authors give a JL construction via Fourier matrices for an arbitrary object in terms of its Gaussian width.

We use an indirect approach from the field of the Compressive Sensing (CS). This topic was first discussed by Terence Tao and Emmanuel Candes in their seminal papers [11], [12] and [13]. ([11] is the highest cited paper of Tao according to Google scholar at the time of writing). CS is a large field; for our purposes we need three tools from it. First is the concept of Restricted Isometry Property (RIP), second is an RIP estimate for discrete Fourier transform matrices (one can think of RIP estimates as the fundamental theorems of CS), and third is a theorem called Krahmer-Ward that transforms an RIP estimate to a JL theorem. Below we describe each step separately. If F is a DFT matrix, then the construction scheme is as follows.

$$\underbrace{\sqrt{\frac{N}{m}}\ (\text{Sample m rows from } F_{N\times N})}_{\text{Restricted Isometry Property}}\ \underbrace{\begin{pmatrix} \pm 1 & & \\ & \ddots & \\ & & \pm 1 \end{pmatrix}}_{\text{Random Diagonal Matrix}}$$

The random diagonal matrix is used in the Krahmer-Ward theorem. We also note that there are multiple conventions for normalizing a DFT matrix. Here we adopt the convention that the matrix is unitary. If ω is the nth root of unity, $\omega^N = 1$, then the matrix F would be

$$F = \frac{1}{\sqrt{N}} \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \omega & \omega^2 & \dots & \omega^{N-1} \\ \vdots & & & & \\ 1 & \omega^{N-1} & \omega^{2(N-1)} & \dots & \omega^{(N-1)^2}. \end{pmatrix} \tag{1.15}$$

The first step is the Restricted Isometry Property. RIP is about how much a matrix distorts sparse vectors. It is similar to the JL condition but it is restricted to sparse vectors. This property is reported through pairs of numbers. In each pair, there is a sparsity level, s, and a deviation value δ. In short sometimes it is written as δ_s. As a reminder, sparsity of a vector is the number of non-zero components of that vectors, denoted by $||x||_0$ where the zero comes from the $\lim_{p\to 0} ||x||_p$. So if a matrix A possess the δ_s RIP then for all vectors x that are s or less sparse we simultaneously have

$$(1-\delta_s)||x||_2 \leq ||Ax||_2 \leq (1+\delta_s)||x||_2. \tag{1.16}$$

Some authors put squares in the RIP condition as in $(1-\delta_s)||x||^2 \leq ||Ax||^2 \leq (1+\delta_s)||x||^2$. These definitions are equivalent up to scaling the δ_s by an absolute constant.

The second step is an RIP estimate for the discrete Fourier transform matrices. The estimate we provide below is for the general family of orthonormal matrices with m rows sampled uniformly in an i.i.d fashion.

Lemma 1.2.1. *[29, Lemma 4.2], [18, theorem 12.32] Let $F \in \mathbb{R}^{N\times N}$ be an orthonormal matrix obeying*

$$F^*F = I, \qquad max_{i,j}|F_{ij}| \leq \frac{K}{\sqrt{N}}. \tag{1.17}$$

Define the random subsampled matrix $H \in \mathbb{R}^{m\times N}$ with i.i.d rows chosen uniformly at random from the rows of F. Then the RIP(s, δ) holds for $\sqrt{\frac{N}{m}}H$ with probability at least $1 - e^{-\eta}$ for all $\delta > 0$ as long as

$$m \geq CK^2(1+\eta)\frac{s\log^4(N)}{\delta^2}. \tag{1.18}$$

Some remarkable characteristics for the bound in (1.18) are that it is linear in s and logarithmic (to 4th power) in N. It has an explicit dependence on the ambient dimension N; since N could be prohibitively large, improving the depence on N in this bound has been an active area of research in CS. For a recent improvement see [10].

The third step is a theorem that allows us to get a JL theorem for finite points from an RIP estimate. This theorem is known as the Krahmer-Ward theorem; it takes a matrix with a RIP estimate, n points as data and applies a sequence of random reflections through the coordinate axes. The result is a probabilistic theorem for a JL map with a controlled failure probability.

Theorem 1.2.2. ***(Krahmer-Ward)*** *[18, theorem 9.36] Let x_i be n points in $\mathbb{R}^n$. Let $\rho, \epsilon \in (0, 1)$. Let $A \in \mathbb{R}^{m\times N}$ be a matrix where its restricted isometry constant for sparsity $2s$, i.e. δ_{2s}, satisfies $\delta_{2s} \leq \frac{\epsilon}{4}$ for $s \geq 16\log(\frac{4n}{\rho})$. Let D be a diagonal matrix with Rademacher random variables, i.e. with uniform ± 1 random variables, on the diagonal. Then with probability exceeding $1 - \rho$ the following holds simultaneously for all x_i.*

$$(1-\epsilon)||x_i||^2 \leq ||ADx_i||^2 \leq (1+\epsilon)||x_i||^2. \tag{1.19}$$

Combining the above 3 steps we get a JL map for finite points using DFT matrices. The benefit of this approach is the computational speed. The cost to pay is a higher final dimension compared to Gaussian or random projection matrices. In chapter 6, we show how one can combine the two approaches and get the best of both worlds.

CHAPTER 2

GEOMETRIC PROPERTIES OF REACH AND COVERING ESTIMATES

In this chapter we discuss geometrical estimates based on a regularity parameter known as reach. The reader interested in only the JL applications can skip to further chapters and refer back as necessary.

We provide 4 main theorems. These theorems provide us with the necessary geometrical tools to construct our desired JL maps for manifolds. In theorem 2.1.7 we give a comprehensive list of properties of reach. In theorem 2.2.6 we give an upper bound for covering numbers of a compact manifold with boundary. We do so by first covering the boundary as an independent manifold. This covers a collar of the boundary, after which we cover the interior. The method of proof is based on Gunther's volume comparison theorem and is restricted by the injectivity radius of the exponential map in the interior of M away from the boundary. In theorem 2.4.1 we provide a lower bound for the covering numbers of a compact manifold. In theorem 2.5.2 we give a covering number bound from above for the unit secants of a submanifold.

2.1 Reach and its Properties

Here we review the concept of reach for a submanifold of the Euclidean space and in theorem 2.1.7 provide a comprehensive list of its properties. Theorem 2.1.7 provides the main properties of reach in a convenient and ready-to-use manner for application to geometric problems, and we hope it is helpful for other researchers. We also discuss the case when the submanifold possibly possesses a boundary, as is expected to be the case for manifold models in applications.

Reach is an extrinsic parameter that is defined based on how far one can move away from an embedded submanifold while maintaining a unique closest point property. This parameter also controls the extrinsic acceleration of unit speed geodesics. It has been used extensively as a regularity parameter since 1959 when it was defined by Federer in [16]. A historical viewpoint of its development can be found in [34]. Some earlier results regarding reach can be found in [1], [8] and [15]. In [1], the authors offer a probabilistic method for estimating the reach of a manifold. In [8] the authors consider the intersection of a ball of small radius with a set of positive reach, and for manifolds they estimate the angle between tangent spaces at different points in terms of reach. In [15, theorem 2], the authors give a JL theorem for a closed manifold via reach.

Here, we begin by recalling the definition of reach and then we summarize its properties.

Definition 2.1.1. *(**Reach** [16, definition 4.1]) For a closed subset of Euclidean space $A \subset \mathbb{R}^n$, the reach τ is defined as*

$$\tau(A) = \sup\{t \geq 0 \,\big|\, \forall x \in \mathbb{R}^n \text{ such that } d(x, A) < t,\ x \text{ has a unique closest point in } A\}.$$

The above definition is for closed subsets of $\mathbb{R}^n$. When restricted to submanifolds, reach depends in part on the second fundamental form of the embedding. However control of the second fundamental form does not fully control reach as two sheets of the submanifold may come close to each other. One can obtain an equivalent characterization of reach in terms of the injectivity radius of the normal exponential map, defined below following [22, Section 8.1]. First we need the concept of cut points for the normal exponential map.

Definition 2.1.2. *(**Cut point**) Let ξ be a line segment in $\mathbb{R}^n$ meeting M orthogonally at $m \in M$. We say $x \in \xi$ is cut point (cut-focal point) along ξ provided distance from x to M is no longer minimized along ξ past x.*

Definition 2.1.3. *(**Cut distance**) Define the function $e_c : \{(p, u) \mid p \in M, u \in N_pM, \|u\| = 1\} \to \mathbb{R}$ such that*

$$e_c(p, u) = \sup\{t > 0 \mid distance(p + tu, M) = t\}.$$

In words, $e_c(p, u)$ is the distance along direction u starting orthogonal at p, past which the line segment along u stops being distance minimizing to M.

Definition 2.1.4. *([22, Section 8.1])* ***The injectivity radius of the normal exponential map*** $r_{inj}(\exp_N)$ *(also known as minimal focal-cut distance) is* $\inf\{e_c(p,u) \mid p \in M, u \in N_pM, \|u\| = 1\}$

When boundary is present, one must specify the criteria for normal vectors at the boundary. They are defined as the ones that make an obtuse angle with inward pointing tangent vectors, following [16, definition 4.4]. These vectors are normal to the boundary, when boundary is viewed as an independent manifold, but only make obtuse angles with the manifold itself. Let $p \in \partial(M) \subset \mathbb{R}^n$ and $v \in T_pM$ that is inward pointing. Then $u \in \mathbb{R}^n$ is normal to M at p if $u \bullet v \leq 0$ for all such v.

We also need the concept of focal points; this helps us describe the behavior where the distance to the submanifold is equal to reach.

Definition 2.1.5. *(**Focal point**, [22, section 8.1]) Let M be a submanifold of $\mathbb{R}^n$. A focal point of M is a point $x \in \mathbb{R}^n$ such that the exponential map exp_N of the normal bundle of M, is singular somewhere on $exp_N^{-1}(x)$.*

Definition 2.1.6. *[22, section 8.1] Let $e_f(p,u)$ be the distance from $p \in M$ to its first focal point along the geodesic $t \to exp_N(p,tu)$. In other words, let $e_f := \{(p,u) \mid p \in M, u \in M_p^{\perp}, ||u|| = 1\} \to \mathbb{R}$ be the function defined by*

$$e_f(p,u) = \inf\{t > 0 \mid kernel(((exp_N)_*)_{(p,tu)}) \neq \{0\}\}.$$

We are now ready to state the key properties of reach. In this approach we are in part following the work of [1, proposition A.1]. We provide a survey of results that are scattered in the literature and provide proofs where they can not be easily located elsewhere.

Theorem 2.1.7 (Properties of Reach)**.** *Let M be a compact, smooth submanifold of $\mathbb{R}^n$ possibly with boundary. Let τ be the reach of M. Let $p, q \in M$ and $x \in \mathbb{R}^N$. Then the following properties hold.*

1. *All geodesics in M are C^1. In the interior of M, they are smooth. At all points unit speed geodesics have one-sided second derivatives. When switching from interior to boundary, geodesics may bifurcate.*

2. *The one sided second derivatives for unit speed geodesics are bounded above in norm by $\frac{1}{\tau}$. In particular in the interior of M the norm of the second fundamental form of M is bounded by $\frac{1}{\tau}$.*

3. *In the interior of M, any sectional curvature k satisfies $\frac{-2}{\tau^2} \leq k \leq \frac{1}{\tau^2}$.*

4. *(Federer Tubular Neighborhood Theorem) If the line segment px is normal to M at p, and $\|p - x\| < \tau$ then p is the closest point to x in M.*

5. *If line segment pq is normal to M at p then $\|p - q\| \geq 2\tau$.*

6. *At distance exactly τ away from M, either there is a focal point or there are two points $p_1, p_2 \in M$ and $y \in \mathbb{R}^n$ such that $\|p_1 - y\| = \|p_2 - y\| = \tau$ and the line segments $p_1 y$ and $p_2 y$ meet at y at an angle of π; or both.*

7. $\tau = r_{inj}(\exp_N)$.

8. $\tau > 0$.

9. *For $p \in M$, the injectivity radius at p is at least $\min\{d_M(p, \partial M), \pi\tau\}$ where $d_M(p, \partial M)$ is the geodesic distance of p to the boundary of M. If there is no boundary, injectivity radius is at least $\pi\tau$.*

For the next two properties, let d and l be the Euclidean and geodesic distances between p and q respectively.

10. $l - \frac{l^2}{2\tau} \leq d$.

11. *When restricted to $d \leq \frac{\tau}{2}$, we further have $l \leq d + \frac{2d^2}{\tau}$.*

For the last two properties, let $\gamma(t)$ be a unit speed geodesic connecting p to q. Let $v \in T_pM$ be $\dot{\gamma}(0)$. Let ϕ be the angle between v and the secant line connecting p to q. Let $w \in T_pM$ be any unit vector and $w^* \in T_qM$ its parallel transport via the connection on M along γ. Let θ be the angle between w and w^* after they are parallel transported to the origin in $\mathbb{R}^n$.

12. $\theta \leq \frac{l}{\tau}$.

13. When restricted to $d \leq \frac{\tau}{2}$, $\sin(\phi) \leq \frac{d}{2\tau}(1 + \frac{2d}{\tau})^2$.

Proof. 1. See [3, theorem 1].

2. See [8, lemma 4] or [27, proposition 6.1]. In the case where a unit speed geodesic $\gamma(t)$ is part of the boundary, the only modification is that $\ddot{\gamma}(t)$ is normal to the boundary and the same argument applies.

3. This follows from item 2 above and Gauss's equation [28, page 100, theorem 5].

4. See [16, theorem 4.8(12)] or [8, theorem 2].

5. By Federer's tubular neighborhood theorem [8, section 3.1], for points on line segment pq with distance to p less than τ, p is the closest point to them. Therefore q is at least τ away from the point that is τ away from p. This gives the claimed $\|p - q\| \geq 2\tau$.

6. See [33].

7. Let $r = r_{\mathrm{inj}}(\exp_N)$. First we show $\tau \leq r$. For the sake of contradiction assume $r < \tau$. Then by definition of r, there is a normal line segment to M that stops being distance minimizing pass distance r away from M. But this contradicts Federer's tubular neighborhood theorem, as such normal line segments must remain distance minimizing up to distance τ.

Next we show $\tau < r$ is impossible and that is enough for obtaining $\tau = r$. For contradiction, assume $\tau < r$. By item 6, at distance exactly τ away from M there must be either a focal point or two equal length line segments normal to M and intersecting each other at angle of π. If there is a focal point, it is known that past a focal point, geodesics stop being distance minimizing, see [31, lemma 2.11]. Hence past a distance τ, there is a normal line segment to M that stops being distance minimizing; but this contradicts the definition of r because it guarantees they remain distance minimizing up to distance r away from M.

In the case of two equal line segments, call them p_1y and p_2y with $p_1, p_2 \in M$. Since they meet at an angle of π, p_1, y and p_2 are collinear; and since $\|p_1 - y\| = \|p_2 - y\|$, p_1y stops being distance minimizing past y. Again this contradicts the guarantee of being distance minimizing up to distance r, showing $\tau < r$ is impossible.

8. See [7, page 151, theorem 4].

9. By [4, theorem 3], in a manifold with boundary and reach τ, two geodesics with the same starting point must travel a distance of $\pi\tau$ before they meet. Separately all sectional curvatures are bounded above by $\frac{1}{\tau^2}$. Then by Klingenberg's theorem [6, theorem 89], we get that the injectivity radius is at least $\pi\tau$. Since we take a minimum with the distance to the boundary, we are guaranteed not to intersect it.

10. See [27, lemma 6.3].

11. For partial and complementary proofs see [27, lemma 6.3] and [15, lemma 7]. Here we streamline the above arguments. Reversing $l - \frac{l^2}{2\tau} \leq d$, one of

$$l \leq \tau - \tau\sqrt{1 - \frac{2d}{\tau}} \text{ or } l \geq \tau + \tau\sqrt{1 - \frac{2d}{\tau}}$$

must hold. If the strict inequality $d < \frac{\tau}{2}$ holds then $l = \tau$ is impossible. We use this fact to show when $d \leq \frac{\tau}{2}$ we must have $l \leq \tau$. From $l \leq \tau$, only $l \leq \tau - \tau\sqrt{1 - \frac{2d}{\tau}}$ remains, and using $1 - \sqrt{1-x} \leq \frac{x+x^2}{2}$ we get the claimed $l \leq d + \frac{2d^2}{\tau}$.

For the sake of contradiction assume $d \leq \frac{\tau}{2}$ and $\tau < l$. Then $S_1 = \{a \in M \mid d_M(p, a) = \tau\}$ is not empty. Let q^* be a minimizer for the Euclidean distance to p in S_1:

$$q^* = \arg\inf_{a \in S_1} \|p - a\|.$$

Then q^* is also a minimizer for Euclidean distance to p over $S_2 = \{a \in M \mid d_M(p, a) \geq \tau\}$. Because if there is a minimizer in the interior of S_2 such as a, then line segment pa must be normal to M at a. Then $\|p - a\| \geq 2\tau$. But since a is a minimizer, it holds that $\|p - a\| \leq \|p - q\| \leq \frac{\tau}{2}$ and that contradicts $\|p - a\| \geq 2\tau$.

Similarly q can't be a minimizer because it is in the interior of S_2. So there is a strict inequality $\|p - q^*\| < \|p - q\|$. But for $\|p - q^*\| < \frac{\tau}{2}$, it is impossible to have $d_M(p, q^*) = \tau$ and this contradicts how q^* was constructed.

12. See [8, lemma 6].

13. See figure 2.1. We have $\sin(\phi) = \frac{h}{d}$ where h is the distance of q to the line through p with direction v. Since we are restricted to $d \leq \frac{\tau}{2}$, using item 11, we have that $l \leq \tau$. Now we bound h. Let $\theta(s)$ be the angle between v and the tangent vector to the geodesic at distance s along the geodesic, when viewed as vectors in $\mathbb{R}^N$. From item 12 we get that $\theta(s) \leq \frac{s}{\tau}$. The incremental height gain in the h direction is at most $\sin(\theta(s))ds$. Therefore we obtain $h \leq \int_0^l \sin(\theta(s))ds \leq \int_0^l \sin\left(\frac{s}{\tau}\right) ds = \tau - \tau\cos\left(\frac{l}{\tau}\right) \leq \frac{l^2}{2\tau} \leq \frac{d^2}{2\tau}\left(1 + \frac{2d}{\tau}\right)^2$. Dividing by d gives the claimed result. ■

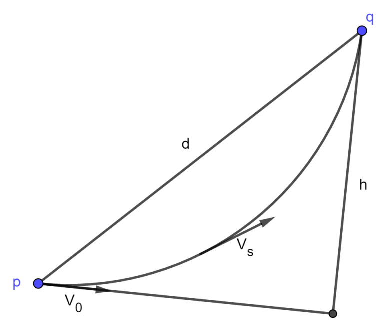


Figure 2.1: Two points with the geodesic and secant line between them. We bound the angle between d and V_0 using τ.

2.2 Covering Estimate from Above via Gunther's Theorem

We first present a covering bound for closed manifolds. Since the boundary of a manifold is itself a closed manifold, this argument is useful in the case the manifold has boundary.

First we need a lemma relating covering and packing numbers. For definitions of packing and covering see [35, lemma 4.2.8].

Lemma 2.2.1. *Let X be a metric space with A a subset of it. Then for any $0 < \epsilon$, the ϵ-covering number of A is not greater than the $\frac{\epsilon}{2}$-packing number of A.*

$$C(\epsilon) \leq P(\frac{\epsilon}{2}) \tag{2.1}$$

Proof. Start with a packing with balls of radius $\frac{\epsilon}{2}$. Then enlargen the balls to radius ϵ. We claim this is a covering. If not there is a point $y \in A$ such that it is more than ϵ away from all the centers of balls. Then a ball of radius $\frac{\epsilon}{2}$ centered at y would be disjoint with the starting packing balls and that contradicts the assumption that any packing is maximal.

Theorem 2.2.2. *(Covering a Closed Manifold) Let $M \hookrightarrow \mathbb{R}^n$ be a closed (compact, no boundary) submanifold of $\mathbb{R}^n$ of dimension $d \geq 0$. Let $\tau > 0$ be the reach of M, and V denote the volume of M, as induced by the embedding. When $d = 0$, we use the counting measure for volume.*
Then if $d = 0$, V balls are sufficient to cover M.
If $d > 0$, let ω_d denote the volume of the d-dimensional Euclidean ball of radius 1. Then for $0 < \epsilon < 2\sqrt{6}\tau$, one can cover M with at most $\frac{V}{\omega_d(1-\frac{\epsilon^2}{24\tau^2})^{d-1}(\frac{\epsilon}{2})^d}$ n-dimensional Euclidean balls of radius ϵ with centers on M.

Proof. The case of $d = 0$ corresponds to isolated points and 1 ball per point is sufficient. So assume $d > 0$. The covering number with balls of radius ϵ is bounded above by the packing number with radius $\frac{\epsilon}{2}$, [35, lemma 4.2.8]. We claim each such packing ball contains a geodesic ball of radius $\frac{\epsilon}{2}$ and the volume of the geodesic ball is at least $\omega_d(1 - \frac{\epsilon^2}{24\tau^2})^{d-1}(\frac{\epsilon}{2})^d$. Since the packing balls are disjoint, the claim implies the theorem immediately. For $x \in M$, $B_{\mathbb{R}^n}(x, \frac{\epsilon}{2})$ contains the geodesic ball $B_M(x, \frac{\epsilon}{2})$ because the geodesic distance between two points of M is not smaller than the Euclidean distance between them. We will work with this geodesic ball.

To bound the volume of the geodesic ball from below, we use Günther's volume comparison theorem, see [19, page 169, theorem 3.101, part ii]. To use it, we need to bound the sectional curvatures of M from above, and ensure that the geodesic balls don't touch the cut locus of their center. From properties of reach τ, theorem 2.1.7, all sectional curvatures of M are bounded above by $\frac{1}{\tau^2}$ and the injectivity radius of M is at least $\pi\tau$. To avoid intersecting the cut locus, we need $\frac{\epsilon}{2} < \pi\tau$; this is satisfied by the assumption that $\epsilon < 2\sqrt{6}\tau$ (one could allow for a larger range of ϵ, see remark 2.2.3). By Günther's theorem, the volume of the geodesic ball is at least the volume of a geodesic ball of radius $\frac{\epsilon}{2}$ in the simply connected model space of constant sectional curvature $\frac{1}{\tau^2}$, which is a sphere with radius τ. Let $V_{\frac{1}{\tau^2}}(\frac{\epsilon}{2})$ denote this volume. It is explicitly given by the formula (see [25]):

$$V_{\frac{1}{\tau^2}}(\frac{\epsilon}{2}) = \frac{2\pi^{d/2}}{\Gamma(\frac{d}{2})}\int_0^{\frac{\epsilon}{2}} S_{\frac{1}{\tau^2}}(x)^{d-1}dx \tag{2.2}$$

where for $k > 0, S_k(x) = \frac{1}{\sqrt{k}}\sin(\sqrt{k}x)$. It follows that

$$\frac{2\pi^{d/2}}{\Gamma(\frac{d}{2})}\int_0^{\frac{\epsilon}{2}} S_{\frac{1}{\tau^2}}(x)^{d-1}dx = \frac{2\pi^{d/2}}{\Gamma(\frac{d}{2})}\int_0^{\frac{\epsilon}{2}} (\tau\sin(\frac{x}{\tau}))^{d-1}dx. \tag{2.3}$$

Since $\frac{\sin(x)}{x}$ is decreasing on $[0, \pi]$, and $\epsilon < 2\sqrt{6}\tau < 2\pi\tau$,

$$V_{\frac{1}{\tau^2}}(\frac{\epsilon}{2}) \geq \frac{2\pi^{d/2}}{\Gamma(\frac{d}{2})}(\frac{\sin(\frac{\epsilon}{2\tau})}{\frac{\epsilon}{2\tau}})^{d-1}\frac{1}{d}(\frac{\epsilon}{2})^d. \tag{2.4}$$

Using $0 < 1 - \frac{x^2}{6} < \frac{\sin(x)}{x}$ for $0 < x < \sqrt{6}$, $\omega_d = \frac{\pi^{d/2}}{\Gamma(\frac{d}{2}+1)}$ and $\Gamma(\frac{d}{2}+1) = \Gamma(\frac{d}{2})\frac{d}{2}$ we get

$$V_{\frac{1}{\tau^2}}(\frac{\epsilon}{2}) \geq \omega_d(1 - \frac{\epsilon^2}{24\tau^2})^{d-1}(\frac{\epsilon}{2})^d. \tag{2.5}$$

■

Remark 2.2.3. In the proof of the theorem 2.2.2, we restricted the range of ϵ as $0 < \epsilon < 2\sqrt{6}\tau$. The reason is that $1 - \frac{x^2}{6}$, which is the second order Taylor expansion of $\sin(x)/x$, remains positive in this range. One could work with the $\text{sinc}(x)$ function to have the larger range $0 < \epsilon < 2\pi\tau$, or one could work with the 6th order Taylor expansion of $\sin(x)/x$, which remains positive slightly beyond 3 and we have $0 \leq T_6(x) \leq \text{sinc}(x)$ on $[0, 3]$. However for the purposes of JL applications such improvements only offer marginal benefits and hence for simplicity we state our bounds for the second order expansion.

Next as an example we apply our theorem to $\mathbb{S}^d$. A standard estimate for covering $\mathbb{S}^d$ with balls centered on the sphere is $(\frac{3}{\epsilon})^d$, see [35, corollary 4.2.13]. Here we get $(3.4\sqrt{d})\frac{2.1^d}{\epsilon^d}$. In the special case of $\mathbb{S}^d$, there are more economical bounds on the order of $O(\frac{d^{1.5}\log(d)}{\epsilon^d})$ ([9, theorem 6.8.1]), but they apply only to $\mathbb{S}^d$, and not arbitrary manifolds.

Corollary 2.2.4. *For* $0 < \epsilon < 1$ *the standard* $\mathbb{S}^d \hookrightarrow \mathbb{R}^n$ *can be covered with at most* $(3.4\sqrt{d})\frac{2.1^d}{\epsilon^d}$ *Euclidean n-dimensional balls of radius* ϵ*.*

Proof. By theorem 2.2.6, we need at most $\frac{V_M}{\omega_d(1-\frac{\epsilon^2}{24\tau_M^2})^{d-1}(\frac{\epsilon}{2})^d}$. We have $\tau = 1$, $V_M = \frac{2\pi^{\frac{d+1}{2}}}{\Gamma(\frac{d+1}{2})}$, $\omega_d = \frac{\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2}+1)}$, and $\frac{V_M}{\omega_d} = 2\sqrt{\pi}\frac{\Gamma(\frac{d}{2}+1)}{\Gamma(\frac{d+1}{2})} < 2\sqrt{\pi d}$. In the last equation we used the Gautschi's inequality [20]. This leads to an upper bound of $2\sqrt{\pi d}(\frac{24}{23})^{d-1}\frac{2^d}{\epsilon^d} \leq 3.4\sqrt{d}(\frac{2.1}{\epsilon})^d$.

■

Next we focus on a covering estimate for a compact manifold with boundary. First we present a lemma showing that the number of connected components of the boundary must be finite.

Lemma 2.2.5. *Let* M *be a compact connected embedded submanifold of* $\mathbb{R}^n$ *with boundary* ∂M*. Then* ∂M *has a finite number of connected components.*

Proof. Assume not. Let $p_i \in M$ be a sequence of points where each p_i belongs to a different component of ∂M. Since M is compact, p_i converge to a point $p \in M$, after passing to a subsequence. Either p is in interior of M or it belongs to the boundary of M. If p is in the interior there is an open neighborhood of p in $\mathbb{R}^n$ intersecting M only in interior points. Since p_i are boundary points, this contradicts their convergence to p. If p is in the ∂M, then it is in one of the connected components such as j. Again there is an open neighborhood of p in $\mathbb{R}^n$ intersecting M only in points in component j of boundary or interior points; this contradicts p_i converging to p. ■

Using theorem 2.2.2 and lemma 2.2.5 we present a covering estimate for a manifold with boundary.

Theorem 2.2.6. *(Covering a Manifold with Boundary) Let* $M \hookrightarrow \mathbb{R}^n$ *be a compact d-dimensional,* $d \geq 1$*, submanifold of* $\mathbb{R}^n$ *with boundary. Let* $\tau_M > 0$ *and* V_M *be the reach and volume of* M*. Let* ∂M *denote the boundary of* M *as an independent submanifold of* $\mathbb{R}^n$ *with* $V_{\partial M}$ *as the volume of* ∂M*. Let* $\tau_{\partial_i M} > 0$ *be the reach of the i-th component of boundary of* M*. Let* $\tau_{\partial M} = \inf_i\{\tau_{\partial_i M}\}$*. Fix* $0 < \epsilon \leq \min\{4\sqrt{6}\tau_{\partial M}, 2\sqrt{6}\tau_M\}$*. Then for* $d = 1$*,* M *can be covered with* $\frac{V_M}{\epsilon} + V_{\partial M}$ *n-dimensional*

Euclidean balls of radius ϵ, where $V_{\partial M}$ corresponds to the counting measure. For $d \geq 2$, M can be covered with $\frac{V_M}{\omega_d(1-\frac{\epsilon^2}{24\tau_M^2})^{d-1}(\frac{\epsilon}{2})^d} + \frac{V_{\partial M}}{\omega_{d-1}(1-\frac{\epsilon^2}{96\tau_{\partial M}^2})^{d-2}(\frac{\epsilon}{4})^{d-1}}$ many balls.

Proof. Consider two subsets of M, $S_1 = \{x \in M | d_M(x, \partial M) < \frac{\epsilon}{2}\}$ and $S_2 = M \backslash S_1$. We will cover each region separately.

To cover S_1 we cover ∂M with balls of radius $\frac{\epsilon}{2}$, and then enlarge them to balls of radius ϵ. We claim such balls cover S_1. For each point $x \in S_1$, there is a point $y \in \partial M$ such that $d_M(x, y) < \frac{\epsilon}{2}$. Also $d_{\mathbb{R}^n}(x, y) \leq d_M(x, y) < \frac{\epsilon}{2}$. Since we have a $\frac{\epsilon}{2}$ covering of ∂M, center of one of the balls, o_i, satisfies $d_{\mathbb{R}^n}(o_i, y) < \frac{\epsilon}{2}$. Therefore $d_{\mathbb{R}^n}(o_i, x) \leq d_{\mathbb{R}^n}(o_i, y) + d_{\mathbb{R}^n}(y, x) < \epsilon$ as claimed.

When $d = 1$, the boundary is a collections of points, so $V_{\partial M}$ is enough to cover the region S_1. For region S_2, we pack it with balls of radius $\frac{\epsilon}{2}$. Each Euclidean ball will overlap with the geodesic ball of the same radius; for $d = 1$ that is a curve of length ϵ. Since volume of S_2 is at most V_M and the packing balls are disjoint we can pack at most $\frac{V}{\epsilon}$ balls. This gives our covering claim of $\frac{V_M}{\epsilon} + V_{\partial M}$.

Now assume $d \geq 2$. We claim $\frac{V_{\partial M}}{\omega_{d-1}(1-\frac{\epsilon^2}{96\tau_{\partial M}^2})^{d-2}(\frac{\epsilon}{4})^{d-1}}$ balls of radius $\frac{\epsilon}{2}$ are enough to cover ∂M. To show this, we use theorem 2.2.2 which requires ∂M to be compact and without boundary. Boundary of M is compact since it is a closed subset of M; furthermore boundary of boundary of a manifold is empty. Therefore by theorem 2.2.2 for $0 < \frac{\epsilon}{2} \leq 2\sqrt{6}\tau_{\partial M}$, ∂M can be covered with $\frac{V_{\partial M}}{\omega_{d-1}(1-\frac{\epsilon^2}{96\tau_{\partial M}^2})^{d-2}(\frac{\epsilon}{4})^{d-1}}$ n-dimensional balls of radius $\frac{\epsilon}{2}$.

To cover region S_2, we instead pack it with balls of radius $\frac{\epsilon}{2}$ and centers in S_2 to get an upper bound. Since the centers are in S_2, each ball will contain the geodesic ball of the same radius, $B_M(o_i, \frac{\epsilon}{2})$. The open geodesic ball $B_M(o_i, \frac{\epsilon}{2})$ can not touch the boundary ∂M since its center is at least $\frac{\epsilon}{2}$ geodesic distance away. The cut-locus of o_i is either at least $\pi\tau_M$ away in the geodesic distance or at least the distance of o_i to the boundary in geodesic distance; and for $0 < \frac{\epsilon}{2} \leq \sqrt{6}\tau_M < \pi\tau_M$ the geodesic ball $B_M(o_i, \frac{\epsilon}{2})$ does not touch the cut locus of o_i. By the volume comparison in the (2.5), we have $0 < \epsilon \leq 2\sqrt{6}\tau_M$ the volume of the the geodesic ball $B_M(o_i, \frac{\epsilon}{2})$ is at least $\omega_d(1 - \frac{\epsilon^2}{24\tau^2})^{d-1}(\frac{\epsilon}{2})^d$. Since in a packing such balls are disjoint, the total volume of such balls can not exceed V_M and their number is at most $\frac{V_M}{\omega_d(1-\frac{\epsilon^2}{24\tau_M^2})^{d-1}(\frac{\epsilon}{2})^d}$ as claimed. ■

We apply our estimate to the standard d-dimensional closed ball B^d as a manifold with boundary.

Corollary 2.2.7. *Let $B^d \hookrightarrow \mathbb{R}^n$ for $n \geq 2$ be the standard d dimensional closed ball of radius* 1. *Then to cover it with n-dimensional balls of radius ϵ for $0 < \epsilon < 1$, $(\frac{2}{\epsilon})^d + 2\pi(\frac{4.05}{\epsilon})^{d-1}$ balls are sufficient.*

Proof. From theorem 2.2.6, we need at most $\frac{V_M}{\omega_d(1-\frac{\epsilon^2}{24\tau_M^2})^{d-1}(\frac{\epsilon}{2})^d} + \frac{V_{\partial M}}{\omega_{d-1}(1-\frac{\epsilon^2}{96\tau_{\partial M}^2})^{d-2}(\frac{\epsilon}{4})^{d-1}}$ balls. The parameters are

M	V_{B^d}	τ_{B^d}	∂B^d	$V_{S^{d-1}}$	$\tau_{S^{d-1}}$	ω_d	ω_{d-1}
B^d	$\frac{\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2}+1)}$	∞	S^{d-1}	$2\frac{\pi^{\frac{d+1}{2}}}{\Gamma(\frac{d+1}{2})}$	1	V_{B^d}	$V_{B^{d-1}}$

Using $0 < \epsilon < 1$ we get, $(\frac{2}{\epsilon})^d + 2\pi\frac{95}{96}(\frac{96}{95}\frac{4}{\epsilon})^{d-1} < (\frac{2}{\epsilon})^d + 2\pi(\frac{4.05}{\epsilon})^{d-1}$.

■

2.3 Covering Estimate from Above Without Radius Restriction via Tubes

In this section we show it is possible to remove the restriction imposed on the covering numbers by the injectivity radius in theorem 2.2.2. This is at the cost of an explicit dependence on the ambient dimension. This dependence could be prohibitive when one is in spaces with extremely high dimensions. For a covering radius larger than the diameter of a manifold, one expects the covering number to be 1. Therefore one encounters 3 regimes for the covering radius: less than the injectivity radius, larger than the injectivity radius but up to the diameter, and larger than the diameter. Interpolating between these regimes is an interesting direction for further research. After we derive our formula in this section, we will discuss this issue again in remark 2.3.3.

Proposition 2.3.1. *Let $M \hookrightarrow \mathbb{R}^n$ be a closed (compact, no boundary) submanifold of $\mathbb{R}^n$ of dimension d. Let $\tau > 0$ be the reach of M, and V denote the volume of M. Let ω_d denote the* *volume of the d-dimensional Euclidean ball of radius* 1*. Then for any* $0 < \epsilon$*, one can cover* M *with* $\frac{V\omega_{n-d}(1+\frac{\epsilon}{2\tau})^d}{\omega_n(\frac{\epsilon}{2})^d}$ *n-dimensional Euclidean balls of radius ϵ with centers on M.*

Remark. Compared to theorem 2.2.2, there is no upper bound on ϵ.

Proof. Instead of covering M, we pack it with n-Euclidean ball of radius $\frac{\epsilon}{2}$ with centers on M. All such balls are within the tube of radius $\frac{\epsilon}{2}$ around M. Since the balls are disjoint the number of balls is bounded above by

$$\frac{\text{Volume of Tube of radius } \frac{\epsilon}{2} \text{ around } M}{\omega_n(\frac{\epsilon}{2})^n} \tag{2.6}$$

To establish the proposition, we need to show that

$$\text{Volume of Tube of radius } \frac{\epsilon}{2} \text{ around } M \leq V_M\omega_{n-d}(\frac{\epsilon}{2})^{n-d}(1+\frac{\epsilon}{2\tau})^d \tag{2.7}$$

$$\leq V_M \text{Vol}(B^{n-d}(\frac{\epsilon}{2}))(1+\frac{\epsilon}{2\tau})^d. \tag{2.8}$$

For $\epsilon < \tau$, the exact volume of this tube was found by Herman Weyl in [36] (see the textbook of Gray [22] for a more detailed explanation). Lotz [26, theorem 3.1] gives an overestimate for the volume of the tube for all values of ϵ as follows.

Let S be the shape operator of M defined via a local orthonormal frame $E_1, .., E_d$ on M and v a unit normal vector to M

$$S(v)_{ij} := S_v(E_i, E_j) = \langle \nabla_{E_i} E_j, v\rangle \tag{2.9}$$

Let $\psi_i(v)$ be defined using the characteristic polynomials of the shape operator S as

$$\det(\text{Id} - tS(v)) = \sum_{i=0}^{d} t^i \psi_i(v) \tag{2.10}$$

Let $S(NM)$ be the unit sphere bundle of the normal bundle of M and define the volume form $\omega_{S(NM)}$ via the Sasaki metric [32]. Define K_i as

$$K_i := \int_{S(NM)} |\psi_i(v)| \omega_{S(NM)} \tag{2.11}$$

Then Lotz's overestimate is give by

$$(\text{Volume of Tube of radius } \frac{\epsilon}{2} \text{ around } M) \leq (\frac{\epsilon}{2})^{n-d} \sum_{i=0}^{d} \frac{K_i}{n-d+i} (\frac{\epsilon}{2})^i \tag{2.12}$$

We have that $\frac{1}{\tau}$ is an upper bound for all eigenvalues of the shape operator. Since $\psi_i(v)$ are the coefficients of the characteristic polynomial of S(v), we have

$$|\psi_i(v)| \leq \binom{d}{i} (\frac{1}{\tau})^i \tag{2.13}$$

$$K_i := \int_{S(NM)} |\psi_i(v)| \omega_{S(NM)} \tag{2.14}$$

$$\leq \text{vol}(S(NM)) \binom{d}{i} (\frac{1}{\tau})^i. \tag{2.15}$$

where $\text{vol}(S(NM)) = \text{vol}(M)\text{vol}(S^{n-d-1})$. Therefore using (2.12) we get

$$(\text{Volume of Tube of radius } \frac{\epsilon}{2} \text{ around } M) \leq \text{vol}(M)\text{vol}(S^{n-d-1})(\frac{\epsilon}{2})^{n-d} \sum_{i=0}^{d} \frac{\binom{d}{i}(\frac{1}{\tau})^i}{n-d+i} (\frac{\epsilon}{2})^i \tag{2.16}$$

Using $\text{vol}(S^{n-d-1}) = (n-d)\text{vol}(B^{n-d})$

$$(\text{Volume of Tube of radius } \frac{\epsilon}{2} \text{ around } M) \leq V_M \text{Vol}(B^{n-d}(\frac{\epsilon}{2})) \sum_{i=0}^{d} \frac{n-d}{n-d+i} \binom{d}{i} (\frac{\epsilon}{2\tau})^i \tag{2.17}$$

$$\leq V_M \text{Vol}(B^{n-d}(\frac{\epsilon}{2}))(1 + \frac{\epsilon}{2\tau})^d \tag{2.18}$$

as claimed. ■

Now we present a similar argument for when a boundary is present.

Proposition 2.3.2. *(Via Tube Formula) Let $M \hookrightarrow \mathbb{R}^n$ be a compact d-dimensional submanifold of $\mathbb{R}^n$ with boundary. Let $\tau_M > 0$ be the reach of $M \backslash \partial M$, and V_M denote the volume of M. Let ∂M denote the boundary of M as an independent submanifold of $\mathbb{R}^n$. Let $\tau_{\partial M} > 0$ be the injectivity radius of the normal exponential map of ∂M and let $V_{\partial M}$ be the volume of ∂M. Let ω_n be volume of the unit Euclidean n-dimensional ball. Then for all $0 < \epsilon$, M can be covered with $V_M \frac{\omega_{n-d}}{\omega_n}(\frac{1}{\tau_M} + \frac{2}{\epsilon})^d + V_{\partial M} \frac{\omega_{n-d+1}}{\omega_n}(\frac{1}{\tau_{\partial M}} + \frac{2}{\epsilon})^{d-1}$ n-dimensional Euclidean balls of radius ϵ.*

Proof. Instead of covering M with Euclidean balls of radius ϵ, we pack it with balls of radius $\frac{\epsilon}{2}$. We restrict the centers to be on M. Any such ball is contained in the set of points of distance $\frac{\epsilon}{2}$

from M. We will bound this volume from above. Let A_1 be the image of normal exponential map of $M \backslash \partial M$ with radius $\frac{\epsilon}{2}$. Let A_2 be the image of normal exponential map of ∂M with radius $\frac{\epsilon}{2}$. Volume of A is bounded above by $V_M \text{Vol}(B^{n-d}(\frac{\epsilon}{2}))(1+\frac{\epsilon}{\tau_M})^d$. Volume of B is bounded above by $V_{\partial M}\text{Vol}(B^{n-(d-1)}(\frac{\epsilon}{2}))(1+\frac{\epsilon}{\tau_{\partial M}})^{d-1}$. Therefore the number of disjoint packing balls can be bounded by dividing by the volume of a single ball.

$$\text{num of balls} \leq \frac{V_M \text{Vol}(B^{n-d}(\frac{\epsilon}{2}))(1+\frac{\epsilon}{\tau_M})^d + V_{\partial M}\text{Vol}(B^{n-(d-1)}(\frac{\epsilon}{2}))(1+\frac{\epsilon}{\tau_{\partial M}})^{d-1}}{\omega_n(\frac{\epsilon}{2})^n} \tag{2.19}$$

$$\leq V_M \frac{\omega_{n-d}}{\omega_n}(1+\frac{\epsilon}{2\tau_M})^d \frac{1}{(\frac{\epsilon}{2})^d} + V_{\partial M}\frac{\omega_{n-d+1}}{\omega_n}(1+\frac{\epsilon}{2\tau_{\partial M}})^{d-1}\frac{1}{(\frac{\epsilon}{2})^{d-1}} \tag{2.20}$$

$$= V_M \frac{\omega_{n-d}}{\omega_n}(\frac{1}{\tau_M}+\frac{2}{\epsilon})^d + V_{\partial M}\frac{\omega_{n-d+1}}{\omega_n}(\frac{1}{\tau_{\partial M}}+\frac{2}{\epsilon})^{d-1} \tag{2.21}$$

Remark 2.3.3. The bound in proposition 2.3.2, which is based on overestimating the volume of a tube around the manifold, is valid for all radii of covering balls. This is in contrast to the approach using Gunther's theorem where one is restricted by the injectivity radius of the manifold. However the tube bounds depend on the ambient dimension n while Gunther's bound only depends on the dimension of the manifold. This is because for small enough radii, one is close to the surface of the manifold. However for large radii, one must account for the dimension of the ambient space. This dependence on the ambient dimension is prohibitive specially when one is in spaces with extremely large dimension.

We show the dependence on the ambient dimension with a calculation. Using the formula for the volume of a Euclidean ball in dimension d, $\omega_d = \frac{\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2}+1)}$. One gets $\frac{\omega_{n-d}}{\omega_n} = \frac{1}{\pi^{\frac{d}{2}}}\frac{\Gamma(\frac{n}{2}+1)}{\Gamma(\frac{n-d}{2}+1)}$. In particular if both n, d are even and $d < \frac{n}{2}$,

$$\frac{\omega_{n-d}}{\omega_n} = \frac{1}{\pi^{\frac{d}{2}}}(\frac{n-d}{2}+1)...(\frac{n}{2}) \geq \frac{1}{\pi^{\frac{d}{2}}}(\frac{n-d}{2}+1)^{\frac{d}{2}} \geq (\frac{n}{4\pi})^{\frac{d}{2}}.$$

One can consider the limit $\epsilon \to \infty$ in the formula $\frac{V\omega_{n-d}(1+\frac{\epsilon}{2\tau})^d}{\omega_n(\frac{\epsilon}{2})^d}$ from proposition 2.3.1. The limit scales as $(\frac{\omega_{n-d}}{\omega_n})\frac{V}{\tau^d}$. For a compact manifold, the covering number is 1 once $\epsilon >$ diameter(M). This difference is in part because in the overestimate we ignore the overlaps as the tube gets large. Incorporating the overlaps in the covering bound is an interesting direction for future research.

2.4 Covering Estimate from Below via Bishop's Theorem

In this section we give lower bounds for covering numbers of a manifold. These bounds are used in chapter 5 for discussing necessary conditions for JL maps.

Theorem 2.4.1. *Let M be a d-dimensional smooth submanifold of $\mathbb{R}^n$ possibly with boundary, with volume V and reach τ. Let $0 < \epsilon \leq \frac{\tau}{2}$, and $r = \epsilon(1 + \frac{2\epsilon}{\tau})$. Let $\omega_d = \frac{\pi^{d/2}}{\Gamma(\frac{d}{2}+1)}$ be the volume of the standard unit d-ball. Then the covering number of M with balls centered on M and radius ϵ satisfies*

$$\frac{V}{\omega_d(1 + \frac{\sqrt{2}r}{3\tau})^{d-1} r^d} \leq C(M, \epsilon) \ .$$

Proof. We claim a single ball can cover at most a subset of M with volume $\omega_d(1 + \frac{\sqrt{2}r}{3\tau})^{d-1} r^d$. This claim immediately implies the theorem.

For two points $p, q \in M$ with Euclidean distance δ, geodesic distance l, and $\delta \leq \frac{\tau}{2}$, one has $l \leq \delta(1 + \frac{2\delta}{\tau})$. Hence the intersection of M and a Euclidean ball of radius ϵ centered on M is contained in a geodesic balls of radius $r = \epsilon(1 + \frac{2\epsilon}{\tau})$ with the same center provided $\epsilon \leq \frac{\tau}{2}$. Therefore it is sufficient to bound the volume of such a geodesic ball from above with $\omega_d(1 + \frac{\sqrt{2}r}{3\tau})^{d-1} r^d$.

We use Bishop's volume comparison theorem [See [19], page 169, theorem 3.101, part i]. By properties of reach, theorem 2.1.7, the injectivity radius of each point of M is at least $\pi\tau$, and all sectional curves of M are bounded below by $\frac{-2}{\tau^2}$. Hence by Bishop's theorem the volume of a geodesic ball of radius r in M is bounded above by the volume of the geodesic ball of radius r in the simply connected space with constant sectional curvature $\frac{-2}{\tau^2}$. Denote this volume by $V_{\frac{-2}{\tau^2}}(r)$; it is given by the formula

$$V_{\frac{-2}{\tau^2}}(r) = \frac{2\pi^{d/2}}{\Gamma(\frac{d}{2})} \int_0^r S_{\frac{-2}{\tau^2}}(x)^{d-1} dx$$

where $S_k(x) = \frac{1}{\sqrt{-k}} \sinh(\sqrt{-k}x)$. Since $\frac{\sinh(x)}{x}$ is increasing on $0 < x$,

$$V_{\frac{-2}{\tau^2}}(r) = \frac{2\pi^{d/2}}{\Gamma(\frac{d}{2})} \int_0^r S_{\frac{-2}{\tau^2}}(x)^{d-1} dx$$

$$\leq \frac{2\pi^{d/2}}{\Gamma(\frac{d}{2})} \int_0^r \left(\frac{\sinh(\frac{\sqrt{2}x}{\tau})}{\frac{\sqrt{2}x}{\tau}}\right)^{d-1} x^{d-1} dx$$

$$\leq \frac{2\pi^{d/2}}{\Gamma(\frac{d}{2})} \left(\frac{\sinh(\frac{\sqrt{2}r}{\tau})}{\frac{\sqrt{2}r}{\tau}}\right)^{d-1} \frac{r^d}{d}.$$

Here we use an approximation for $\frac{\sinh(x)}{x}$. Since $0 < \epsilon \leq \frac{\tau}{2}$, then $r = \epsilon(1 + \frac{2\epsilon}{\tau}) \leq \tau$, and $\frac{\sqrt{2}r}{\tau} \leq \sqrt{2}$. Using $\frac{\sinh(x)}{x} \leq 1 + \frac{x}{3}$ for $0 < x \leq \sqrt{2}$,

$$\leq \frac{2\pi^{d/2}}{\Gamma(\frac{d}{2})} (1 + \frac{\sqrt{2}r}{3\tau})^{d-1} \frac{r^d}{d}$$

$$= \omega_d (1 + \frac{\sqrt{2}r}{3\tau})^{d-1} r^d$$

■

Corollary 2.4.2. *Let M be a d-dimensional smooth submanifold of $\mathbb{R}^n$ possibly with boundary, with volume V and reach τ. Let $0 < \epsilon \leq \frac{\tau}{2}$. Let $\omega_d = \frac{\pi^{d/2}}{\Gamma(\frac{d}{2}+1)}$ be the volume of the standard d-ball. Then the covering number of M with balls centered on M and radius ϵ satisfies*

$$\frac{V}{\omega_d (3\epsilon)^d} \leq C(M, \epsilon) \ .$$

Proof. We use the bound $\frac{V}{\omega_d (1+\frac{\sqrt{2}r}{3\tau})^{d-1} r^d} \leq C(M, \epsilon)$. From $0 < \epsilon \leq \frac{\tau}{2}$, we simplify $r = \epsilon(1+\frac{2\epsilon}{\tau})$ to $r \leq 2\epsilon$ and $r \leq \tau$. We get the factor of 3 from $(1 + \frac{\sqrt{2}}{3})2 < 3$. ■

2.5 Covering Estimate from Above for the Unit Secants of a Submanifold

In this section we calculate a covering number for the unit secants generated by a compact submanifold of the Euclidean space with boundary in terms of the covering numbers for the manifold. The secants of a manifold have been studied previously in [24, section 3], [30, page 1323], [37, section 1] and [38, section 3]. In our approach here, we use the covering numbers to estimate the Gaussian width of the set of unit secants.

We give a list of secants that form an ϵ-net for the set of unit secants. The distance between two unit secants is calculated by their distance on the unit sphere. The secants of M are separated into long and short secants; the long secants are compared to the secants generated by a net on the manifold and the short secants are compared to tangent vectors anchored at the net points. The technique of separating the cords into long and short have been previously used in [14, page 5], [15, lemma 9], and [23, section 3].

For long secants, we have the following lemma.

Lemma 2.5.1. *[14, lemma 4.1] Let p, p^*, q and q^* be 4 points in $\mathbb{R}^n$. Let $0 < l = ||p - q||$ and $||p - p^*||, ||q - q^*|| < d$. Let $0 < \epsilon < 1$ and assume $\frac{4d}{l} \leq \epsilon$. Then*

$$||\frac{p - q}{||p - q||} - \frac{p^* - q^*}{||p^* - q^*||}|| \leq \epsilon. \tag{2.22}$$

We now present the covering argument for the unit secants. In the proof, we develop a lemma for short secants similar to long ones.

Theorem 2.5.2. ***(Covering the unit secants)*** *Let $M \hookrightarrow \mathbb{R}^n$ be a compact d-dimensional submanifold of $\mathbb{R}^n$ with boundary ∂M. Let τ_M be the reach of M. Let τ_i be the reach of the i-th connected component of ∂M as a submanifold of $\mathbb{R}^n$. Let $\tau = \inf_{M,i}\{\tau_M, \tau_i\}$. Let V_M be the volume of M and $V_{\partial M}$ be the volume of ∂M. Let $U(M - M) = \{\frac{p-q}{||p-q||} \mid p \neq q,\ p, q \in M\}$ be the set of unit secants of M, and let $\overline{U(M - M)}$ be its closure. Fix $0 < \epsilon < 1$. Define $\alpha = \frac{V_M}{\omega_d}(\frac{41}{\tau})^d + \frac{V_{\partial M}}{\omega_{d-1}}(\frac{81}{\tau})^{d-1}$. Then $\overline{U(M - M)}$ can be covered with $\left(\frac{\alpha^2}{2} + \alpha(2d3^{d-1})\right)\frac{1}{\epsilon^{4d}}$ n-dimensional Euclidean balls of radius ϵ.*

Proof. We give a list of unit vectors in $\mathbb{R}^n$, count its elements, and prove that every element of $\overline{U(M-M)}$ is within ϵ distance of one of them. In this approach we follow [15]. The list is as follows.

1. Consider a $(\frac{\tau\epsilon^2}{20})$-net of points in M. Include all the unit secant lines generated by pair of these points in the list. If $\tau = \infty$ then 1 point is sufficient for the net. (That is because infinite reach corresponds to affine spaces and all secants can be compared to tangent vectors at one point).

2. For each point in the above net, consider a $(\frac{\epsilon}{3})$-net for its unit tangent sphere. Include all such unit vectors in the list. If an anchor point is in the boundary, there is a half unit sphere for inward pointing vectors and we consider the unit S^{d-1} as its extension.

We must show that the list above is an ϵ-net for $\overline{U(M-M)}$. To do so divide the secant lines into a long and short set based on their length. The cut-off between long and short secants is $\frac{\tau\epsilon}{5}$. If $\tau = \infty$ then all secants count as short.

Long Secants: For a long secant between $p, q \in M$, let p^* and q^* be the closest point to them in the net on M. Since $2(\frac{\tau\epsilon^2}{20}) < \frac{\tau\epsilon}{5} \leq ||p-q||$, we are guaranteed that p^* and q^* are distinct points. By lemma 2.5.1 the distance between the unit secants of $p-q$ and p^*-q^* is bounded by $\frac{4(\frac{\tau\epsilon^2}{20})}{||p-q||} \leq \frac{(\frac{\tau\epsilon^2}{5})}{(\frac{\tau\epsilon}{5})} \leq \epsilon$. Therefore all long secants are covered by the secants in item 1 above.

Short Secants: For a short secant, we first exchange it for the starting tangent direction for the distance minimizing geodesic connecting its base points. Second, we exchange that tangent vector for its parallel transport along the distance minimizing geodesic to the tangent space of the closest point to the base point in the $(\frac{\tau\epsilon^2}{20})$-net of M. Third, we exchange the parallel transported vector for one of the vectors in the item 2 above. In this process, we will incur 3 errors which we denote by ϵ_1, ϵ_2 and ϵ_3. We must show that $\epsilon_1 + \epsilon_2 + \epsilon_3 \leq \epsilon$.

Exchanging a short secant for a tangent vector. Here we bound ϵ_1 by $\frac{\epsilon}{3}$. Using theorem 2.1.7, the angle between the initial tangent vector and the secant satisfies $\sin(\phi) \leq \frac{d}{2\tau}(1+\frac{2d}{\tau})^2$.

Because the cutoff for short secants is $\frac{\tau\epsilon}{5}$, substituting for d, $\sin(\phi) \leq \frac{49\epsilon}{250}$. Using $2\sin(\frac{\phi}{2})$ as the distance on a unit sphere for 2 vectors with angle ϕ,

$$\epsilon_1 = 2\sin(\frac{\phi}{2}) < \frac{3}{2}\sin(\phi) \leq \frac{147\epsilon}{500} < \frac{\epsilon}{3}.$$

Parallel transport to an anchor point. Here we bound ϵ_2 by $\frac{\epsilon}{3}$. Here we parallel transport an initial geodesic tangent vector from the base point of the secant to an anchor point. Using theorem 2.1.7, $\theta \leq \frac{l}{\tau}$. The Euclidean distance to the closest net point is at most $d \leq (\frac{\tau\epsilon^2}{20})$. We translate to the geodesic distance using theorem 2.1.7, $l \leq d + \frac{2d^2}{\tau}$. It follows that using $\epsilon \leq 1$

$$\epsilon_2 = 2\sin(\frac{\theta}{2}) \leq \theta \leq \frac{l}{\tau} \leq \frac{d}{\tau}(1 + \frac{2d}{\tau})^2 \leq \frac{\epsilon^2}{20}(1 + \frac{\epsilon^2}{10})^2 < \frac{\epsilon}{3}.$$

Here we had room to spare for obtaining the bound $\frac{\epsilon}{3}$. However the cut-off of $\frac{\tau\epsilon}{3}$ and $\frac{\tau\epsilon^2}{20}$ are needed for the exchanging a short secant with a tangent vector and for handling the long secants.

Exchange with a preselected tangent vector: At the anchor point, we exchange the parallel transported tangent vector with a vector in a net for the unit sphere at the anchor point. By choice a $(\frac{\epsilon}{3})$-net is considered at the anchor point hence, $\epsilon_3 \leq \frac{\epsilon}{3}$. Therefore together $\epsilon_1 + \epsilon_2 + \epsilon_3 \leq \epsilon$.

It remains to count the number of unit vectors in the above list. For item 1, by theorem 2.2.6, one can create a $(\frac{\tau\epsilon^2}{20})$-net for M with $C(\frac{\tau\epsilon^2}{20})$ when $C(\epsilon)$ is defined as

$$C(\epsilon) = \frac{V_M}{\omega_d(1 - \frac{\epsilon^2}{24\tau_M^2})^{d-1}(\frac{\epsilon}{2})^d} + \frac{V_{\partial M}}{\omega_{d-1}(1 - \frac{\epsilon^2}{96\tau_{\partial M}^2})^{d-2}(\frac{\epsilon}{4})^{d-1}}.$$

The number of unit secants generated by such points is bounded above by $\frac{1}{2}C(\frac{\tau\epsilon^2}{20})^2$. For directions in the list 2, the anchor points could be in the interior or in the boundary of M. We consider the tangent S^{d-1} for the unit vectors for all of them; if an anchor point is in the boundary, there is a half unit sphere for inward pointing vectors and we consider the unit S^{d-1} as its extension. By [35, corollary 4.2.13], each S^{d-1} can be covered by $(\frac{3}{\epsilon})^d$ unit vectors. Since there is a unit sphere for each point in the net from item 1, there are $C(\frac{\tau\epsilon^2}{20})(\frac{3}{\epsilon})^d$. Adding the two estimates together there are at most $\frac{1}{2}C(\frac{\tau\epsilon^2}{20})^2 + C(\frac{\tau\epsilon^2}{20})(\frac{3}{\epsilon})^d$ unit vectors. This number is rather inconvenient and we will simplify it. One has $C(\frac{\tau\epsilon^2}{20})$

$$C(\frac{\tau\epsilon^2}{20}) = \frac{V_M}{\omega_d(1-\frac{(\frac{\tau\epsilon^2}{20})^2}{24\tau_M^2})^{d-1}(\frac{(\frac{\tau\epsilon^2}{20})}{2})^d} + \frac{V_{\partial M}}{\omega_{d-1}(1-\frac{(\frac{\tau\epsilon^2}{20})^2}{96\tau_{\partial M}^2})^{d-2}(\frac{(\frac{\tau\epsilon^2}{20})}{4})^{d-1}}.$$

We note $0 < \epsilon \leq 1$, $\tau = \min\{\tau_M, \tau_{\partial M}\}$, and

$$\frac{(2\times 20)^d}{(1-\frac{\epsilon^4}{(24)(20^2)})^{d-1}} \leq 40.005^d < 41^d$$
$$\frac{(4\times 20)^d}{(1-\frac{\epsilon^4}{(96)(20^2)})^{d-1}} \leq 80.003^d < 81^d.$$

Define $\alpha = \frac{V_M}{\omega_d}(\frac{41}{\tau})^d + \frac{V_{\partial M}}{\omega_{d-1}}(\frac{81}{\tau})^{d-1}$. Then

$$C(\frac{\tau\epsilon^2}{20}) \leq \frac{\alpha}{\epsilon^{2d}}$$

Therefore we get the overall upper bound as

$$\frac{\alpha^2}{2}\frac{1}{\epsilon^{4d}} + (\frac{\alpha}{\epsilon^{2d}})(\frac{3}{\epsilon})^d \leq \left(\frac{\alpha^2}{2} + \alpha 3^d\right)\frac{1}{\epsilon^{4d}}.$$

■

For an example we apply our covering estimate to the standard S^d as a corollary. We note that the set of the unit secants of S^d is equal to S^d itself, hence there is redundancy as many pairs of points achieve the same unit secant. It is an interesting geometry question to find submanifolds that their secants avoid being parallel. In this direction there is the work on totally skew embeddings [21]. Such submanifolds would be great candidates for benchmarking JL maps as their unit secants is expected to be large.

Corollary 2.5.3. *For covering the unit secants of $S^d \hookrightarrow \mathbb{R}^n$, $\frac{14d\sqrt{d}41^{2d}}{\epsilon^{4d}}$ balls of radius ϵ are sufficient.*

Proof. We calculate using theorem 2.5.2. We need $\left(\frac{\alpha^2}{2} + \alpha 3^d\right)\frac{1}{\epsilon^{4d}}$ balls for $\alpha = \frac{V_M}{\omega_d}(\frac{41}{\tau})^d + \frac{V_{\partial M}}{\omega_{d-1}}(\frac{81}{\tau})^{d-1}$. We drop the boundary term since S^d has no boundary. We have $\tau = 1$, $V_{S^d} = 2\frac{\pi^{\frac{d+1}{2}}}{\Gamma(\frac{d+1}{2})}$ and $V_{B^d} = \omega_d = \frac{\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2}+1)}$. Then $\frac{V_{S^d}}{\omega_d} = 2\sqrt{\pi}\frac{\Gamma(\frac{d}{2}+1)}{\Gamma(\frac{d+1}{2})} \leq 2\sqrt{\pi d}$. Therefore $\alpha \leq 2\sqrt{\pi d}41^d$, and $\frac{\alpha^2}{2} + \alpha 3^d \leq 14d\sqrt{d}41^{2d}$. ∎

2.6 Gaussian Width Estimate for Unit Secants From Above

Definition 2.6.1. *Let g be a standard Gaussian random variable in $\mathbb{R}^n$. Define the Gaussian Width of $S \subset \mathbb{R}^n$, $\omega(S)$, as follows.*

$$\omega(S) = \mathbb{E} \sup_{x \in S} \langle g, x \rangle \tag{2.23}$$

Theorem 2.6.2. *(Via Gunther) Let $M \hookrightarrow \mathbb{R}^n$ be a compact d-dimensional submanifold of $\mathbb{R}^n$ with boundary ∂M. Let τ_M be the reach of M. Let τ_i be the reach of the i-th connected component of ∂M as a submanifold of $\mathbb{R}^n$. Let $\tau = \inf_{M,i}\{\tau_M, \tau_i\}$. Let V_M be the volume of M and $V_{\partial M}$ be the volume of ∂M. Let $U(M - M) = \{\frac{p-q}{||p-q||} \mid p \neq q,\ p, q \in M\}$ be the set of unit secants of M, and let $\overline{U(M - M)}$ be its closure. Let Gaussian Width of $\overline{U(M - M)}$ be $\omega(\overline{U(M - M)})$.*

Define

$$\alpha = \frac{V_M}{\omega_d}(\frac{41}{\tau})^d + \frac{V_{\partial M}}{\omega_{d-1}}(\frac{81}{\tau})^{d-1} \tag{2.24}$$

$$c = (\frac{\alpha^2}{2} + 3^d \alpha) \tag{2.25}$$

Then

$$\omega(\overline{U(M - M)}) \leq 4\sqrt{2}\sqrt{\log(c) + 4d} \tag{2.26}$$

Proof. We use the covering number bounds in theorem 2.5.2 and Dudley's inequality [18, page 226]. Let $C(\epsilon)$ denote the covering number of $\overline{U(M - M)}$ with n-dimensional Euclidean balls of radius ϵ. We have that $C(\epsilon) \leq (\frac{\alpha^2}{2} + 2d3^d\alpha)\frac{1}{\epsilon^{4d}}$ for $0 < \epsilon \leq 1$ and $C(\epsilon) = 1$ for $\epsilon > 1$ because $\overline{U(M - M)}$ is a subset of the S^{n-1}. Dudley's inequality gives us

$$\omega(\overline{U(M - M)}) \leq 4\sqrt{2} \int_0^\infty \sqrt{\log(C(\epsilon))} d\epsilon. \tag{2.27}$$

Therefore

$$\omega(\overline{U(M - M)}) \leq 4\sqrt{2} \int_0^1 \sqrt{\log((\frac{\alpha^2}{2} + 2d3^d\alpha)\frac{1}{\epsilon^{4d}})} d\epsilon. \tag{2.28}$$

Let $c = (\frac{\alpha^2}{2} + 2d3^d\alpha)$, then by Cauchy-Schwartz

$$\omega(\overline{U(M-M)}) \leq 4\sqrt{2}\int_0^1 \sqrt{\log(c) + 4d\log(\frac{1}{\epsilon})} \tag{2.29}$$

$$\leq 4\sqrt{2}\sqrt{\log(c)}\int_0^1 \sqrt{1 + \frac{4d}{\log(c)}\log(\frac{1}{\epsilon})} \tag{2.30}$$

$$\leq 4\sqrt{2}\sqrt{\log(c)}\sqrt{\int_0^1 1 + \frac{4d}{\log(c)}\log(\frac{1}{\epsilon})}. \tag{2.31}$$

With the identity

$$\int_0^1 1 + k\log(\frac{1}{x})dx = 1 - k(x\log(x) - x)\big|_0^1 = 1 + k. \tag{2.32}$$

We continue

$$\omega(\overline{U(M-M)}) \leq 4\sqrt{2}\sqrt{\log(c)}\sqrt{1 + \frac{4d}{\log(c)}} \tag{2.33}$$

$$\leq 4\sqrt{2}\sqrt{\log(c) + 4d} \tag{2.34}$$

as required. ∎

CHAPTER 3

MORE ESTIMATES WITH REACH

3.1 Closed Submanifolds with Reach 1 and Minimum Volume

In the earlier covering estimates such as theorem 2.6.2 we encounter the ratio $\frac{V}{\tau^d}$ for d dimensional manifolds. This ratio is scaling invariant and hence it is natural to ask if there is a minimum value for it over closed manifolds. We will also encounter this ratio again in theorem 5.0.5 for a necessary condition for JL maps. Therefore here we investigate if it has a lower bound.

In particular if we consider the case when $d = 1$, and the manifold is topologically equivalent to $\mathbb{S}^1$, and $\tau = 1$, then the minimum of the ratio is just the minimum length of a closed curve with reach one. Here we can show that for a closed d dimensional manifold the minimum value of this ratio is given by the volume of the standard $\mathbb{S}^d$. In section 3.2 we also consider the minimum value of this ratio for a torus embedded in 3 dimensions, which has a topological constraint.

Proposition 3.1.1. *Let $M \subset \mathbb{R}^N$ be a closed smooth d dimensional manifold of reach 1. Then the volume of M is at least $Vol(\mathbb{S}^d)$ which is the volume of the standard unit $\mathbb{S}^d$.*

Proof. We present the proof in 4 steps.

1. The injectivity radius of a closed manifold of reach τ is at least $\pi\tau$. So for $\tau = 1$, the injectivy radius is at least π. See theorem 2.1.7

2. Sectional curvatures of M are all abounded above by $\frac{1}{\tau^2}$. For $\tau = 1$, they are all bounded by 1.

3. Consider the geodesic ball of radius π around an arbitrary point. Because the injectivity radius is at least π, then we can apply Gunther's volume comparison theorem. Then the volume of the geodesic ball is at least the volume of the geodesic ball of radius π in the space of uniform

curvature one. The general formula is

$$V_{\frac{1}{\tau^2}}(r) = \frac{2\pi^{d/2}}{\Gamma(\frac{d}{2})}\int_0^r S_{\frac{1}{\tau^2}}(x)^{d-1}dx$$

where $S_k(x) = \frac{1}{\sqrt{k}}\sin(\sqrt{k}x)$ for $k > 0$. So we can rewrite

$$\frac{2\pi^{d/2}}{\Gamma(\frac{d}{2})}\int_0^r S_{\frac{1}{\tau^2}}(x)^{d-1}dx = \frac{2\pi^{d/2}}{\Gamma(\frac{d}{2})}\int_0^r \left(\tau\sin\left(\frac{x}{\tau}\right)\right)^{d-1} dx.$$

Here we want $r = \pi$ and $\tau = 1$ so we get

$$V = \frac{2\pi^{d/2}}{\Gamma(\frac{d}{2})}\int_0^\pi (\sin(x))^{d-1}\, dx.$$

This is the formula for the volume of the d sphere, as

$$\mathrm{Vol}(\mathbb{S}^d) = \frac{2\pi^{(d+1)/2}}{\Gamma(\frac{d+1}{2})}. \tag{3.1}$$

4. The volume of M is at least the volume of the geodesic ball we found in the previous item. ■

By scaling we get the following more general bound.

Corollary 3.1.2. *Let $M \subset \mathbb{R}^N$ be a closed smooth d dimensional manifold of reach τ. Then the we have*

$$\frac{V}{\tau^d} \geq Vol(\mathbb{S}^d). \tag{3.2}$$

Proof. We scale by a factor of $\frac{1}{\tau}$ so that τ becomes 1. However under the scaling the ratio $\frac{V}{\tau^d}$ doesn't change. After the scaling we have $\frac{V_{\text{new}}}{1} \geq \mathrm{Vol}(\mathbb{S}^d)$. Therefore the same is true before scaling. ■

3.1.1 Codimension 1

In this subsection we give an alternative proof with the isoperimetric inequality for the lower bound on the ratio $\frac{V}{\tau^d}$ in the special case where $d = n - 1$ (codimension 1) and the manifold is closed. In the case of equality in our bound, we get a rigidity statement coming from the isoperimetric inequality. The difference with the previous discuss is in having a different method of proof and a rigidity result in the case of equality.

Theorem 3.1.3. *Let $M \subset \mathbb{R}^n$ be a codimension 1 closed submanifold of $\mathbb{R}^n$. Assume the reach of M is 1. Then the $n-1$-dimensional volume of M is at least vol($\mathbb{S}^{n-1}$) where $\mathbb{S}^{n-1}$ is the unit $n-1$ dimensional sphere. In the case of volume equality, M is a rigid transformation of $\mathbb{S}^{n-1}$.*

Proof. Since M is codimension 1 and is closed, its complement will have two connected components. One of them is the inward component. Call it A. We give a lower bound for the volume of A.

Since the reach of M is 1, the principle curvatures of M in all normal directions are bounded above by 1, [27, Proposition 6.1]. Consider the parallel surface of distance $0 < t < 1$ in the inwards normal direction. Let the shape operator of M be S, then the change of volume formula is given by $\det(I - tS)$ and

$$\det(I - tS) \geq (1 - t)^{n-1}. \tag{3.3}$$

Therefore if V is the (n-1)-dimensional volume of M, the n-dimensional volume of A is at least

$$V \int_{t=0}^{1} (1 - t)^{n-1} dt = \frac{V}{n} \tag{3.4}$$

Now we apply the isoperimetric inequality to A with $\partial A = M$. Let ω_n be the volume of the unit n-ball.

$$n(\omega_n)^{\frac{1}{n}} \mathrm{vol}_n(A)^{\frac{n-1}{n}} \leq \mathrm{vol}_{n-1}(M) \tag{3.5}$$

$$n(\omega_n)^{\frac{1}{n}} (\frac{V}{n})^{\frac{n-1}{n}} \leq V \tag{3.6}$$

$$\mathrm{vol}(\mathbb{S}^{n-1}) = n\omega_n \leq V. \tag{3.7}$$

In the last line we used the fact that $\mathrm{vol}(\mathbb{S}^{n-1}) = \frac{d}{dr}|_{r=1}\mathrm{vol}(rB_n)$. The equality case of the isoperimetric inequality gives that M is a rigid motion of $\mathbb{S}^{n-1}$. ■

3.2 Torus in $\mathbb{R}^3$ with Reach 1 and Minimum Area

Here we consider the problem of minimizing the ratio $\frac{V}{\tau^d}$ when $d = 2$, $\tau = 1$ for a smooth torus in $\mathbb{R}^3$. In this special case there is a topological constraint of having genus 1 and therefore we expect the minimum to be more than the value $4\pi \approx 13$ given in section 3.1.

Our candidate torus for minimizing the ratio is the torus of revolution with larger radius 2 and smaller radius 1. The reach of this torus is 1 and its surface area is $8\pi^2 = (2\pi(2))(2\pi(1)) \approx 79$.

Here we offer two proofs for smaller lower bounds. The first proof is based on Loewner's systolic inequality [6, page 325] and gives a lower bound of $2\sqrt{3}\pi^2 \approx 34$. The second proof is based on the isoperimetric inequality and gives a lower bound of $2^{7/3}3^{2/3}\pi^{5/3} \approx 70$.

Theorem 3.2.1. *(Via systolic inequality) Let $T \subset \mathbb{R}^3$ be a smooth embedded torus with reach* 1. *Then the surface area of T is at least* $2\sqrt{3}\pi^2$.

Proof. Consider the systole of T (shortest non-contractible loop). It is necessarily a periodic geodesic [6, page 326]. Since the reach of T is 1, the external acceleration of the systole as a unit speed geodesic is at most 1 in norm. Any closed curve with acceleration bounded by 1 has length at least 2π [17]. Therefore our theorem follows from Loewner's systolic inequality [6, page 325]

$$\text{sys}^2 \frac{\sqrt{3}}{2} \leq A \tag{3.8}$$

$$(2\pi)^2 \frac{\sqrt{3}}{2} \leq A \tag{3.9}$$

$$2\sqrt{3}\pi^2 \leq A. \tag{3.10}$$

■

Theorem 3.2.2. *(Via isoperimetric inequality) Let $T \subset \mathbb{R}^3$ be a smooth embedded torus with reach* 1. *Then the surface area of T is at least* $2^{7/3}3^{2/3}\pi^{5/3} \approx 70$.

Proof. Since torus is an orientable 2 dimensional surface there is a well-defined outward and inward normal direction.

Consider the parallel surface T_{out} defined by moving in the outward direction by $1 - \epsilon$. T_{out} is another smooth torus because reach of T is 1. We get our bound by applying the isoperimetric inequality to this surface.

First we show the volume contained inside T_{out} (i.e. the volume of the bounded component of the complement of T_{out}) is at least $16\pi^2$. Second we show that the surface area of T_{out} is at most twice the surface area of T. Then we apply the isoperimetric inequality.

First we discuss the issue of volume of T_{out}. Let $\epsilon \in (0, 1)$. Consider the surface Σ_ϵ, that is obtained from T by moving in the inward normal direction by a distance of $1 - \epsilon$. Then Σ_ϵ is a smooth surface because the reach of T is 1.

Consider the solid torus given by the filling the bounded component of the complement of Σ_ϵ. This solid torus is a 3-manifold with boundary and its systole is a periodic geodesic [6, page 326]. Call this curve l_ϵ. We claim the reach of l_ϵ is at least $2 - \epsilon$. Consider $l_\epsilon(t)$ as a unit speed geodesic; for a t_0 if $l_\epsilon(t_0)$ is in the interior of solid bounded by Σ_ϵ, then the acceleration of $l_\epsilon(t_0)$ is zero because $\mathbb{R}^3$ is flat. If $l_\epsilon(t_0)$ is on Σ_ϵ, the boundary of the solid torus, then the acceleration of $l_\epsilon(t_0)$ is outward pointing, and since reach of T is 1 and we moved another $1 - \epsilon$ inside, we get that reach of l_ϵ is at least $2 - \epsilon$. By [17, page 50], the length of l_ϵ is at least $2\pi(2 - \epsilon)$.

The tube of radius $2 - 2\epsilon$ around l_ϵ is contained in the inside of T_{out}. The volume of this tube by [22, page 7] is at least $2(2 - \epsilon)(2 - 2\epsilon)^2\pi^2$. Therefore the volume inside of T_{out} is at least $2(2 - 2\epsilon)^3\pi^2$.

Now we discuss the surface area of T_{out}. Let H and K denote the mean and Gaussian curvature for T respectively. By the parallel surface area formula [22, page 8], we have

$$\text{Area}(T_{out}) = \text{Area(T)} + (1 - \epsilon) \int_T 2H + (1 - \epsilon)^2 \int_T K \tag{3.11}$$

By the Gauss–Bonnet theorem $\int_T K = 0$.

We claim $\int_T 2H \leq \text{Area}(T)$. This is because the surface area of the parallel surface in the inward direction, obtained by moving from T in the inward direction by distance t, is given by $\text{Area(T)} - t \int_T 2H$. Since the reach of T is 1, this area must remain positive up to $t = 1$ which implies $\int_T 2H \leq \text{Area}(T)$. Returning to (3.11), we get that

$$\text{Area}(T_{out}) \leq (2 - \epsilon)\text{Area}(T) \tag{3.12}$$

Now we apply the isoperimetric inequality to T_{out}.

$$3\text{Vol}(T_{out})^{2/3}(\frac{4}{3}\pi)^{1/3} \leq \text{Area}(T_{out}) \tag{3.13}$$

$$3(2(2-2\epsilon)^3\pi^2)^{2/3}(\frac{4}{3}\pi)^{1/3} \leq (2-\epsilon)\text{Area}(T) \tag{3.14}$$

Letting $\epsilon \to 0$ and simplifying we get

$$70 \approx 2^{7/3}3^{2/3}\pi^{5/3} \leq \text{Area}(T). \tag{3.15}$$

■

3.3 Doing Better than Reach in Covering Estimates

Our earlier covering estimate for a compact manifold was based on reach. We used Gunther's theorem for volume comparison and the main ingredient in that theorem is an upper bound on sectional curvatures. For a compact manifold of reach τ, $\frac{1}{\tau^2}$ is a global bound on all sectional curvatures. Since this bound is based on reach, it is sensitive to perturbations. A small bump on the manifold could change the value of τ substantially; see figure 3.1. However for a covering argument, one should be able to find a global estimate which is based on averaging a local estimate.

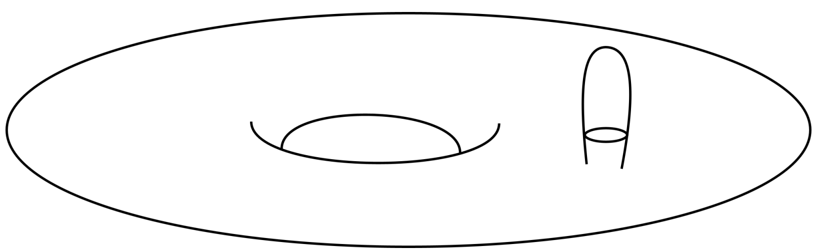


Figure 3.1: Reach is a global parameter. A bump on the surface of a manifold can change its reach, and therefore reach is sensitive to local perturbations.

Consider the function $k : M \to \mathbb{R}$

$$k(p) = \text{max sectional curvature of } M \text{ at } p \tag{3.16}$$

Since sectional curvatures at a point are defined in the space of all planes in T_pM, Grassmanian $Gr(2,d)$, and this space is compact then the max exists. Let $\tilde{k}, \underset{\sim}{k}, \fint k$ be the max, min and the average of k over M. In this notation $\fint$ takes precedence, hence "$\fint k + \tilde{k}$" means "$(\fint k) + \tilde{k}$". Again since M is compact, all three exist. With respect to the covering estimates, the approach with reach is based on $\tilde{k} \le \frac{1}{\tau^2}$. Here we can show that in the asymptotic regime as the radius of the covering goes to zero, one can use $\frac{2^d \tilde{k} \fint k - (\fint k)\underset{\sim}{k} - (2^d - 1)\tilde{k}\underset{\sim}{k}}{\tilde{k} + (2^d - 1)\fint k - 2^d \underset{\sim}{k}}$. When k is not a constant function, this quantity is strictly bounded by $\fint k$ and $\tilde{k}$ and therefore it is an improvement over $\tilde{k}$ and $\frac{1}{\tau^2}$. We show this with a calculation below:

$$0 < (\tilde{k} - \fint k)(\tilde{k} - \underset{\sim}{k}) \tag{3.17}$$

$$2^d \tilde{k} \fint k - \underset{\sim}{k} \fint k - (2^d - 1)\underset{\sim}{k}\tilde{k} < (2^d - 1)\tilde{k} \fint k + \tilde{k}^2 - 2^d \tilde{k}\underset{\sim}{k} \tag{3.18}$$

$$\frac{2^d \tilde{k} \fint k - \underset{\sim}{k} \fint k - (2^d - 1)\tilde{k}\underset{\sim}{k}}{\tilde{k} + (2^d - 1)\fint k - 2^d \underset{\sim}{k}} < \tilde{k}, \tag{3.19}$$

and

$$(\fint k - \tilde{k})(\fint k - \underset{\sim}{k}) < 0 \tag{3.20}$$

$$(2^d - 1)\fint^2 k + \tilde{k} \fint k - 2^d \underset{\sim}{k} \fint k < 2^d \tilde{k} \fint k - \underset{\sim}{k} \fint k - (2^d - 1)\underset{\sim}{k}\tilde{k} \tag{3.21}$$

$$\fint k < \frac{2^d \tilde{k} \fint k - \underset{\sim}{k} \fint k - (2^d - 1)\tilde{k}\underset{\sim}{k}}{\tilde{k} + (2^d - 1)\fint k - 2^d \underset{\sim}{k}}, \tag{3.22}$$

and therefore we can conclude

$$\fint k < \frac{2^d \tilde{k} \fint k - \underset{\sim}{k} \fint k - (2^d - 1)\tilde{k}\underset{\sim}{k}}{\tilde{k} + (2^d - 1)\fint k - 2^d \underset{\sim}{k}} < \tilde{k}. \tag{3.23}$$

The averaged quantity comes from treating each packing ball separately. As usual we use the packing-covering relationship to switch to a packing estimate. For each packing ball, we showed it isolates a geodesic ball of some radius with a minimum volume. This minimum volume gives the desired bound. Consider the supremum of all sectional curvatures only in that geodesic ball, then the Gunther's theorem still holds with this sectional curvature. We can average the values of max sectional curvatures with one value per packing ball to get an averaged value. Now we show this calculation. First we need a volume estimate lemma.

Lemma 3.3.1. *Let M be a simply connected d-dimensional manifold of constant sectional curvature $k \in \mathbb{R}$. Consider a geodesic ball of radius r and volume $V_k(r)$. Then for r small enough we have*

$$\frac{2\pi^{d/2}}{\Gamma(\frac{d}{2})}(\frac{r^d}{d} - \frac{k(d-1)}{6(d+2)}r^{d+2}) \leq V_k(r). \tag{3.24}$$

Proof. The volume of a ball of radius r in a space form is

$$V_k(r) = \frac{2\pi^{d/2}}{\Gamma(\frac{d}{2})}\int_0^r S_k(x)^{d-1}dx$$

where for $k \in \mathbb{R}$

$$S_k(r) = \begin{cases} \frac{1}{\sqrt{k}}\sin(\sqrt{k}r) & k > 0 \\ r & k = 0 \\ \frac{1}{\sqrt{-k}}\sinh(\sqrt{-k}r) & k < 0 \end{cases} \tag{3.25}$$

For $k > 0$

$$\frac{1}{\sqrt{k}}\sin(\sqrt{k}x) = x - \frac{kx^3}{6} + \frac{k^2x^5}{5!} + \ldots \tag{3.26}$$

and $k < 0$

$$\frac{1}{\sqrt{-k}}\sinh(\sqrt{-k}x) = x + \frac{(-k)x^3}{6} + \frac{k^2x^5}{5!} + \ldots \tag{3.27}$$

Therefore

$$\int_0^r S_k(x)^{d-1}dx = \int_0^r (x - \frac{kx^3}{6} + \ldots)^{d-1}dx \tag{3.28}$$

$$= \int_0^r x^{d-1} + \binom{(d-1)}{1}x^{d-2}(\frac{-kx^3}{6}) + \ldots dx \tag{3.29}$$

$$= \frac{r^d}{d} - \frac{k(d-1)}{6(d+2)}r^{d+2} + \ldots \tag{3.30}$$

The next term in the expansion is always positive so for a small enough r we have

$$\frac{r^d}{d} - \frac{k(d-1)}{6(d+2)}r^{d+2} \leq \int_0^r S_k(x)^{d-1}dx \tag{3.31}$$

■

Next we give a lemma about averaging.

Lemma 3.3.2. *Let M be a closed orientable smooth Riemannian manifold of dimension d with volume V. Let $k : M \to \mathbb{R}$ be the map that gives the maximum sectional curvature at each point. Let $B_i(x_i, r), 1 \leq i \leq N$ be disjoint closed geodesic balls of radius r and centers x_i. Let k_i be the supremum of all sectional curvatures in B_i. Let $k^* = \frac{\sum k_i}{N}$ be the average of k_i. Then for r small enough we have*

$$N \leq V \Big/ \left[\frac{2\pi^{d/2}}{\Gamma(\frac{d}{2})} \left(\frac{r^d}{d} - \frac{k^*(d-1)}{6(d+2)} r^{d+2} \right) \right] \tag{3.32}$$

Proof. Let $V_i(r)$ be the volume of ball i. Then $\sum_i V_i(r) \leq V$. By the Gunther's theorem we have

$$\frac{2\pi^{d/2}}{\Gamma(\frac{d}{2})} \left(\frac{r^d}{d} - \frac{k_i(d-1)}{6(d+2)} r^{d+2} \right) \leq V_i(r) \tag{3.33}$$

Summing over the balls we get

$$\sum_{i=1}^{N} \frac{2\pi^{d/2}}{\Gamma(\frac{d}{2})} \left(\frac{r^d}{d} - \frac{k_i(d-1)}{6(d+2)} r^{d+2} \right) \leq V \tag{3.34}$$

$$N \left[\frac{2\pi^{d/2}}{\Gamma(\frac{d}{2})} \left(\frac{r^d}{d} - \frac{k^*(d-1)}{6(d+2)} r^{d+2} \right) \right] \leq V \tag{3.35}$$

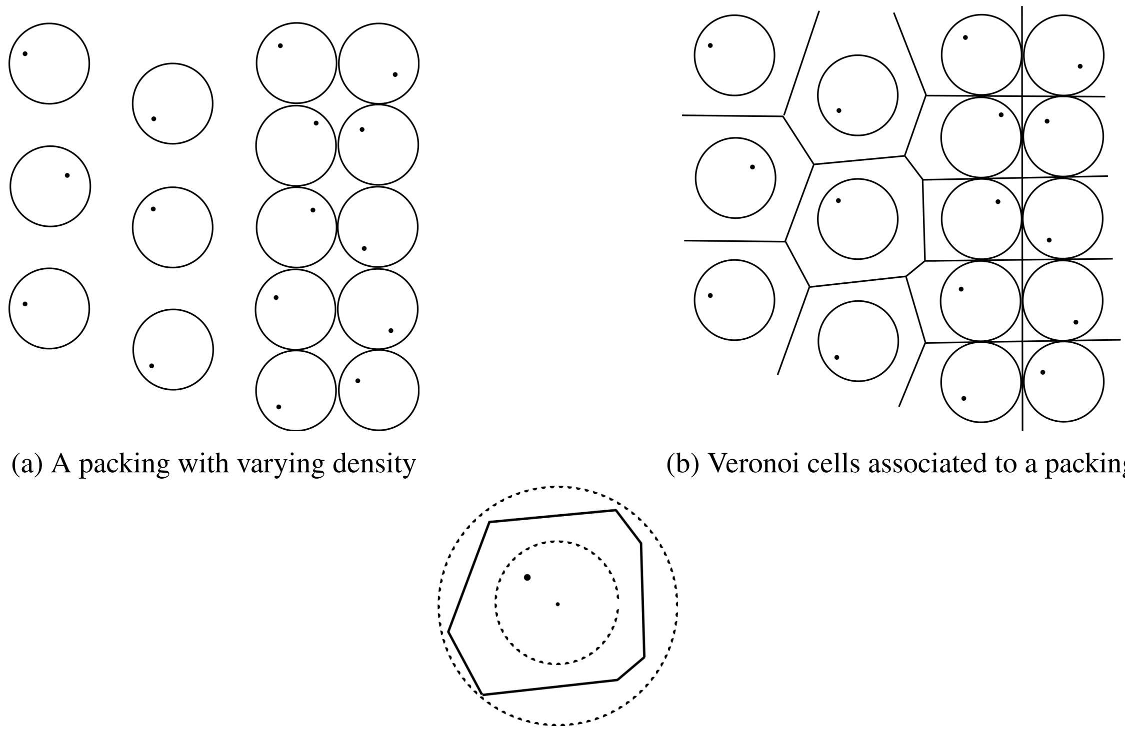
which is our claim. ■

So based on the above lemma if we find an estimate for $k^* = \sum_i \frac{k_i}{N}$ we can use sharper bound in the packing estimates compared to the ones given by reach. In the asymptotic regime, as the radius of packing balls goes to zero, a natural candidates for k^* is $\frac{1}{V} \int_M k$. But because of varying packing density this is not necessary the case.

Because of varying density as in figure 3.2, it is possible that $\frac{1}{V} \int_M k$ deviates from $k^* = \sum \frac{k_i}{N}$. However, if $B_i(x_i, r)$ is a packing ball with center x_i and v_i is the Voronoi cell corresponding to the B_i, then

$$B_i(x_i, r) \subset v_i \subset B_i(x_i, 2r) \tag{3.36}$$

This property restricts the deviation of our average from the continuous average over the manifold. We use this relation to obtain our bound $\frac{2^d \tilde{k} \fint k - (\fint k) \underset{\sim}{k} - (2^d - 1) \tilde{k} \underset{\sim}{k}}{\tilde{k} + (2^d - 1) \fint k - 2^d \underset{\sim}{k}}$.



(a) A packing with varying density

(b) Veronoi cells associated to a packing

(c) Voronoi cell contained in a doubled ball

Figure 3.2: Figure-a shows a packing with varying density. Figure-b shows the Veronoi cells associated with a packing. Figure-c shows a Veronoi cell contained by a doubled packing ball.

Theorem 3.3.3. *Let M be a closed smooth Riemannian manifold with volume V. Let $B^i_j(x^i_j, r), 1 \leq j \leq N_i$ be a packing of M with geodesic balls of radius r_i and centers x_i. Let $p^i_j \in B^j_i$ be points that are arbitrarily selected from each ball. Let σ_i be a probability measure obtained by placing a mass of $1/N_i$ weight at p^i_j. Assume $r_i \to 0$ and σ_i converges weakly to σ.*

Let $k : M \to [\underset{\sim}{k}, \tilde{k}]$ be the map that gives the maximum sectional curvature at each point. Assume k is a Lipschitz function. Let $\tilde{k}, \underset{\sim}{k}$ and $\fint k$ be the max, min and the average of k over M. Let $k^i_j := k(p^i_j)$.

Let μ and μ_σ be the push forward of measure induced by the volume form and σ on $[\underset{\sim}{k}, \tilde{k}]$ via k. Assume μ and μ_σ are absolutely continuous with respect to the Lebesgue measure.

If k is not a constant function, then $\underset{\sim}{k} < \fint k < \tilde{k}$ and

$$\limsup_{r_i \to 0} \frac{\sum_{j=1}^{N_i} k^i_j}{N_i} \leq \frac{2^d \tilde{k} \fint k - (\fint k)\underset{\sim}{k} - (2^d - 1)\tilde{k}\underset{\sim}{k}}{\tilde{k} + (2^d - 1) \fint k - 2^d \underset{\sim}{k}}. \tag{3.37}$$

Proof. We first fix i and work with a single packing. Later we consider $i \to \infty$. We partition M into Voronoi cells based on their distance to the center of the balls x^i_j. Define

$$v^i_j = \{p \in M | d_g(p, x^i_j) < d_g(p, x^i_k)\ j \neq k\} \tag{3.38}$$

where d_g is the geodesic distance. We assume the measure of the points that are equ-distant to two of x^i_j is zero and we have

$$V = \sum_{j=1}^{N_i} |v^i_j| \tag{3.39}$$

Since balls B^i_j are a packing of M we have

$$B^i_j(x^i_j, r_i) \subset v^i_j \subset B^i_j(x^i_j, 2r_i) \tag{3.40}$$

For $a, b \in [\underset{\sim}{k}, \tilde{k}]$, we count the number $k^i_j \in [a, b]$ from above and below using the v^i_j and μ. Let c be the Lipschitz constant of k.

For an upper bound, if $k^i_j \in [a, b]$, then $\forall x \in v^i_j$, $a - 4r_i c \leq k(x) \leq b + 4r_i c$. Then

$$\bigcup_{i: k^i_j \in [a,b]} v^i_j \subset \{p | a - 4r_i c \leq k(p) \leq b + 4r_i c\} \tag{3.41}$$

Let $V_\alpha(r)$ be the volume of geodesic ball of radius r in simply connected space of constant sectional curvature α. By Gunther's volume comparison theorem and relation ((3.40)), for r small enough we have

$$V_{b+4rc}(r_i) \leq |v^i_j| \tag{3.42}$$

Since the Voronoi cells are disjoint, from equation ((3.41))

$$(\# k^i_j \in [a, b]) \leq \frac{\mu(a - 4r_i c, b + 4r_i c)}{V_{b+4r_i c}(r)} \tag{3.43}$$

where $\mu(a - 4r_i c, b + 4r_i c) = |\{p : a - 4r_i c \leq k(p) \leq b + 4r_i c\}|$.

For a lower bound, consider all Voronoi cells that intersect the set $\{a + 4r_i c \leq k(p) \leq b - 4r_i c\}$. If $r_i < \frac{b-a}{8c}$ this set is non-empty. For such Voronoi cells we have $\forall x \in v^i_j, a \leq k(x) \leq b$ since v^i_j is contained in a ball of radius $2r_i$ and the k^i_j from v^i_j satisfies $k^i_j \in [a, b]$. Thus

$$\{p | a + 4r_i c \leq k(p) \leq b - 4r_i c\} \subset \bigcup_{j: k^i_j \in [a,b]} v^i_j \tag{3.44}$$

Again let $V_\alpha(r)$ be the volume of geodesic ball of radius r in simply connected space of constant sectional curvature α. Using the Bishop-Gromov's volume comparison theorem

$$|v_j^i| \leq V_a(2r_i) \tag{3.45}$$

Since the Voronoi cells are disjoint, from equation ((3.44))

$$\frac{\mu(a+4r_ic, b-4r_ic)}{V_a(2r_i)} \leq (\#k_j^i \in [a,b]) \tag{3.46}$$

Combining ((3.43)) and ((3.46)), we have

$$\frac{\mu(a+4r_ic, b-4r_ic)}{V_a(2r_i)} \leq (\#k_j^i \in [a,b]) \leq \frac{\mu(a-4r_ic, b+4r_ic)}{V_{b+4r_ic}(r_i)} \tag{3.47}$$

To get the proportion of $k_i \in [a,b]$ we must divide by N_i. We divide the ((3.47)) for $[a,b]$ and $[d,e] \subset [\underset{\sim}{k}, \tilde{k}]$ to cancel the N_i and get

$$\frac{V_{e+4r_ic}(r_i)}{V_a(2r_i)} \frac{\mu(a+4r_ic, b-4r_ic)}{\mu(d-4r_ic, e+4r_ic)} \leq \frac{(\#k_j^i \in [a,b])}{(\#k_j^i \in [d,e])} \leq \frac{V_d(2r_i)}{V_{b+4r_ic}(r_i)} \frac{\mu(a-4r_ic, b+4r_ic)}{\mu(d+4r_ic, e-4r_ic)} \tag{3.48}$$

Note that since curvature terms are second order in the volume expansion

$$\lim_{r \to 0} \frac{V_\alpha(2r)}{V_\beta(r)} = 2^d \tag{3.49}$$

Letting $r_i \to 0$ and using ((3.49)) in ((3.48)) we get

$$\frac{1}{2^d} \frac{\mu(a,b)}{\mu(d,e)} \leq \frac{(\#k_j^i \in [a,b])}{(\#k_j^i \in [d,e])} \leq 2^d \frac{\mu(a,b)}{\mu(d,e)} \tag{3.50}$$

Let $f(a) = \sigma'(a)$ and $g(k) = \frac{\mu'(k)}{V}$. Note that $g(k)$ is a probability density on $[\underset{\sim}{k}, \tilde{k}]$. Letting $b \to a$ and $e \to d$ and using convergence of $\sigma_i \to \sigma$, absolute continuity of σ and μ,

$$\frac{1}{2^d} \frac{g(a)}{g(d)} \leq \frac{f(a)}{f(d)} \leq 2^d \frac{g(a)}{g(d)} \tag{3.51}$$

By assumption both $f(k)$ and $g(k)$ are positive and continuous on $[\underset{\sim}{k}, \tilde{k}]$. Therefore there are constants $c_1 < 1 < c_2$ such that

$$c_1 g(k) \leq f(k) \leq c_2 g(k) \tag{3.52}$$

Suppose $f(k_1) = c_1 g(k_1)$ and $f(k_2) = c_2 g(k_2)$ for some k_1, k_2. Using (3.51) with $a = k_2$ and $d = k_1$we get

$$\frac{1}{2^d}\frac{g(k_2)}{g(k_1)} \leq \frac{c_2}{c_1}\frac{g(k_2)}{g(k_1)} \leq 2^d \frac{g(k_2)}{g(k_1)} \tag{3.53}$$

which gives

$$\frac{c_1}{2^d} \leq c_2 \leq 2^d c_1 \tag{3.54}$$

Since $c_1 < 1 < c_2$

$$\frac{c_2}{2^d} g(k) \leq f(k) \leq c_2 g(k) \tag{3.55}$$

Now consider that both $f(k)$ and $g(k)$ are probability densities on $[\underset{\sim}{k}, \tilde{k}]$. Given the linear constraints on $f(k)$ given by condition (3.55), and $\int f(k) = \int g(k) = 1$, from the simplex algorithm, we can conclude that $\int_k kf(k)$ is maximized when the inequalities in (3.55) are saturated except possibility at one point such as k_0 to transition between the two sides of (3.55). Therefore consider

$$f(k) = \begin{cases} \frac{c_2}{2^d} g(k) & \text{if } k \leq k_0 \\ c_2 g(k) & \text{if } k_0 < k \end{cases} \tag{3.56}$$

Let $\int_{\underset{\sim}{k}}^{k_0} g(k) = a$ and $\int_{k_0}^{\tilde{k}} g(k) = b$ with $a, b > 0, a + b = 1$. Then from $\int_{\underset{\sim}{k}}^{\tilde{k}} f(k) = 1$ we get $\frac{c_2}{2^d} a + c_2 b = 1$, and $c_2 = \frac{2^d}{a + 2^d b}$. Now we bound $\int_k kf(k)$ from above in two ways.

$$\int_k kf(k) = \frac{c_2}{2^d}\int_{\underset{\sim}{k}}^{k_0} kg(k) + c_2 \int_{k_0}^{\tilde{k}} kg(k) \tag{3.57}$$

$$= \frac{c_2}{2^d}\int_{\underset{\sim}{k}}^{\tilde{k}} kg(k) + c_2(1 - \frac{1}{2^d})\int_{k_0}^{\tilde{k}} kg(k) \tag{3.58}$$

$$\leq \frac{c_2}{2^d} \fint k + c_2(1 - \tfrac{1}{2^d}) b\tilde{k} \tag{3.59}$$

$$= \frac{c_2}{2^d} \fint k + (1 - \tfrac{c_2}{2^d})\tilde{k} \qquad \text{Ⓐ} \tag{3.60}$$

where we use $(1-\frac{1}{2^d})c_2 b = (1-\frac{c_2}{2^d})$. In another way,

$$\int_k kf(k) = \frac{c_2}{2^d}\int_{\underset{\sim}{k}}^{k_0} kg(k) + c_2\int_{k_0}^{\tilde{k}} kg(k) \tag{3.61}$$

$$= -c_2(1-\frac{1}{2^d})\int_{\underset{\sim}{k}}^{k_0} kg(k) + c_2\int_{\underset{\sim}{k}}^{\tilde{k}} kg(k) \tag{3.62}$$

$$\leq -c_2(1-\frac{1}{2^d})a\underset{\sim}{k} + c_2 ⨍ k \tag{3.63}$$

$$= -(c_2-1)\underset{\sim}{k} + c_2 ⨍ k \qquad \text{Ⓑ} \tag{3.64}$$

where we used $c_2(1-\frac{1}{2^d})a = c_2 - 1$. Hence $\int_k kf(k) \leq \min\{A, B\}$. If we allow c_2 to be a variable then A and B are lines with positive and negative slopes respectively. Hence their intersection is an upper bound for $\int_k kf(k)$.

$$\frac{c}{2^d} ⨍ k + (1-\tfrac{c}{2^d})\tilde{k} = -(c-1)\underset{\sim}{k} + c ⨍ k \tag{3.65}$$

$$c = \frac{\tilde{k} - \underset{\sim}{k}}{(1-\frac{1}{2^d}) ⨍ k + \frac{1}{2^d}\tilde{k} - \underset{\sim}{k}} \tag{3.66}$$

Substituting back into A or B, we get

$$\int_k kf(k) \leq \frac{2^d\tilde{k} ⨍ k - (⨍ k)\underset{\sim}{k} - (2^d-1)\tilde{k}\underset{\sim}{k}}{\tilde{k} + (2^d-1) ⨍ k - 2^d\underset{\sim}{k}} \tag{3.67}$$

■

Corollary 3.3.4. *Let $T \subset \mathbb{R}^3$ be a smooth torus. Let B_i be packing balls of radius r and let p_i be arbitrary points selected from B_i, and $k(p_i)$ be the sectional curvature at p_i. Then*

$$\limsup_{r\to 0} \frac{\sum_i k(p_i)}{\#\{k_i\}} \leq \frac{-3\tilde{k}\underset{\sim}{k}}{\tilde{k} - 4\underset{\sim}{k}} \tag{3.68}$$

Proof. We use the previous theorem with $d = 2$. By Gauss-Bonnet theorem $⨍ k = 2 - 2g = 0$, and the claim follows. Note that in a torus in $\mathbb{R}^3$, the max and min of sectional curvatures are respectively positive and negative. Hence our bound is a positive quantity. ■

CHAPTER 4

SUFFICIENT CONDITIONS FOR JL MAPS OF MANIFOLDS

4.1 With sub-Gaussian Matrices (Variation 1)

In this section we improve on the work in [15, theorem 2]. In that work, a JL map is constructed for a compact manifold via Gaussian matrices. The method is based on first finding an overestimate for covering for the unit secants of a compact manifold based on its reach and volume, and second using a chaining argument to get a JL map. The chaining method is based on using a sequence of progressively denser ϵ-nets and representing each point in the manifold as a branch in the tree of ϵ-nets. [35, Chapter 8] provides a detailed explanation of this technique.

Here we improve on the previous work in two ways. First we allow the manifold to have boundary. This generalization allows the JL map to apply to data models where the data terminates at a boundary before it can wrap around and form a close manifold. Second we use the covering numbers to calculate the Gaussian width of the unit secants of the manifold via Dudley's inequality, and then we apply the matrix deviation inequality to get JL maps for sub-Gaussian matrices. This calculation streamlines and simplifies the generic chaining calculation in the earlier work. Our main theorem in this section is theorem 4.1.7.

First we state the result in [15, theorem 2].

Theorem 4.1.1. *[15, theorem 2] Let M be a compact d-dimensional Riemannian submanifold of $\mathbb{R}^N$ with reach τ [1] and volume V_M. Conveniently assume that*

$$\frac{V_M}{\tau^d} \geq \left(\frac{21}{2\sqrt{d}}\right)^d \tag{4.1}$$

For $0 < \epsilon \leq \frac{1}{3}$ and $0 < \rho < 1$. Let Φ be a random $m \times N$ matrix populated with i.i.d zero mean

[1] In [15] the concept of condition number is used which is just the reciprocal of reach.

Gaussian random variables with variance of $1/m$ *with*

$$m \geq 18\epsilon^{-2} \max(24d + 2d\log(\frac{\sqrt{d}}{\tau\epsilon^2}) + \log(2V_M^2), \log(\frac{8}{\rho})). \tag{4.2}$$

Then with probability at least $1-\rho$*, for every pair of points* $x_1, x_2 \in M$

$$(1-\epsilon)||x_1 - x_2|| \leq ||\Phi x_1 - \Phi x_2\,|| \leq (1+\epsilon)||x_1 - x_2||. \tag{4.3}$$

Next we state the definitions for sub-Gaussian random variables, vectors and matrices, and then we state the matrix deviation inequality.

Definition 4.1.2. *[35, section 2.5](sub-Gaussian Random Variable) A random variable X is sub-Gaussian if it satisfies*

$$\mathbb{E}\exp(X^2/a^2) \leq 2 \tag{4.4}$$

for some positive constant a. The sub-Gaussian norm of X denoted by $||X||_{\psi_2}$ *is defined as*

$$||X||_{\psi_2} = \inf\{t > 0 : \mathbb{E}\exp(X^2/t^2) \leq 2\} \tag{4.5}$$

Alternatively, a random variable is called sub-Gaussian if it satisfies

$$\mathbb{P}\{|X| \geq t\} \leq 2\exp(-ct^2/||X||_{\psi_2}^2) \tag{4.6}$$

for all $t \geq 0$ *where* $c > 0$ *is an absolute constant.*

Definition 4.1.3. *[35, definition 3.4.1](sub-Gaussian Random Vector) A random vector* $X \in \mathbb{R}^n$ *is called sub-Gaussian if the one-dimensional marginals* $\langle X, x\rangle$ *are sub-Gaussian random variables for all* $x \in \mathbb{R}^n$*. The sub-Gaussian norm of X is defined as*

$$||X||_{\psi_2} = \sup_{x \in S^{n-1}} ||\langle X, x\rangle||_{\psi_2}. \tag{4.7}$$

Definition 4.1.4. *[35, definition 3.4.1](sub-Gaussian Random Matrix) A random* $m \times n$ *matrix A such that its rows* A_i *are independent, isotropic and sub-Gaussian random vectors in* $\mathbb{R}^n$ *is called a sub-Gaussian random matrix.*

Theorem 4.1.5. *[35, theorem 9.1.1 and 9.1.8] Let A be $m \times n$ sub-Gaussian random matrix. Let A_i be the rows of A with $K = \max ||A_i||_{\psi_2}$. Let $T \subset \mathbb{R}^n$ with Gaussian width $\omega(T)$. Then there exists an absolute constant C such that for any $u \geq 0$ the following statement holds with probability at least $1 - 2\exp(-u^2)$.*

$$\sup_{x \in T} \left| ||Ax||_2 - \sqrt{m}||x||_2 \right| \leq CK^2[\omega(T) + u \cdot rad(T)].$$

■

We rewrite the matrix deviation inequality as lower bound for the number of rows of a sub-Gaussian matrix.

Theorem 4.1.6. *Let $T \subset \mathbb{S}^{n-1} \subset \mathbb{R}^n$ with Gaussian width $\omega(T)$. Let $0 < \epsilon, \rho < 1$. Let A be a $m \times n$ sub-Gaussian random matrix with $K = \max ||A_i||_{\psi_2}$. Then there exists a universal constant C such that if*

$$\frac{C^2K^4\left(\omega(T) + \sqrt{\log(\frac{2}{1-\rho})}\right)^2}{\epsilon^2} \leq m \tag{4.8}$$

then with probably at least ρ the following expressions

$$(1-\epsilon)||x|| \leq ||\frac{1}{\sqrt{m}}Ax|| \leq (1+\epsilon)||x||$$

holds simultaneously for all $x \in T$.

Proof. We use theorem 4.1.5. Since $T \subset \mathbb{S}^{n-1}$, $\mathrm{rad}(T) = 1$. Inverting the probability expression $\rho = 1 - 2\exp(-u^2)$, we obtain $u = \sqrt{\log(\frac{2}{1-\rho})}$. To control the error, we need

$$\frac{CK^2}{\sqrt{m}}[\omega(T) + u] \leq \epsilon$$

$$\frac{C^2K^4\left(\omega(T) + \sqrt{\log(\frac{2}{1-\rho})}\right)^2}{\epsilon^2} \leq m.$$

This completes the proof. ■

Now we put our matrix deviation inequality and the Gaussian width upper bound calculation together to a JL map for a compact manifold with boundary.

Theorem 4.1.7. *(JL Map of a Compact Submanifold of $\mathbb{R}^n$ with Boundary via sub-Gaussian Random Matrices) Let $M \hookrightarrow \mathbb{R}^n$ be a compact d-dimensional submanifold of $\mathbb{R}^n$ with boundary ∂M. Let τ_M be the reach of M. Let τ_i be the reach of the i-th connected component of ∂M as a submanifold of $\mathbb{R}^n$. Let $\tau = \inf_{M,i}\{\tau_M, \tau_i\}$. Let V_M be the volume of M and $V_{\partial M}$ be the volume of ∂M. Define*

$$\alpha = \frac{V_M}{\omega_d}(\frac{41}{\tau})^d + \frac{V_{\partial M}}{\omega_{d-1}}(\frac{81}{\tau})^{d-1} \tag{4.9}$$

$$c = (\frac{\alpha^2}{2} + d3^d\alpha) \tag{4.10}$$

$$\omega^* = 4\sqrt{2}\sqrt{\log(c) + 4d}. \tag{4.11}$$

Let A be a $m \times n$ sub-Gaussian random matrix such that $K = \max_i ||A_i||_{\psi_2}$.

Let $0 < \epsilon, \rho < 1$. There exists a universal constant C such that if m satisfies

$$\frac{C^2K^4\left(\omega^* + \sqrt{\log(\frac{2}{1-\rho})}\right)^2}{\epsilon^2} \leq m \tag{4.12}$$

then with probably at least ρ, the following bounds hold simultaneously for all $x_1, x_2 \in M$,

$$(1-\epsilon)||x_1 - x_2|| \leq ||A(x_1 - x_2)|| \leq (1+\epsilon)||x_1 - x_2||. \tag{4.13}$$

Proof. We use theorem 4.1.6. We know that ω^* is an upper bound for the Gaussian width of $U(M)$. Since $U(M)$ is a subset of the unit sphere, we can apply theorem 4.1.6 and this immediately gives the result. ∎

Remark 4.1.8. We analyze the dependence of m on d while keeping the other variables fixed. If one puts $m = \frac{C^2K^4\left(\omega(T)+\sqrt{\log(\frac{2}{1-\rho})}\right)^2}{\epsilon^2}$ as the least sufficient value of m, then m depends on d with order of $O(d\log(d))$. We have $\frac{1}{\omega_d} = \frac{\Gamma(\frac{d}{2}+1)}{\pi^{\frac{d}{2}}} = O(d^d)$, $\alpha = \frac{V_M}{\omega_d}(\frac{41}{\tau})^d + \frac{V_{\partial M}}{\omega_{d-1}}(\frac{81}{\tau})^{d-1} = O(d^d)$, $c = (\frac{\alpha^2}{2} + 2d3^{d-1}\alpha) = O(d^{2d})$, and $\omega(M) = 4\sqrt{2}\sqrt{\log(c) + 4d} = \sqrt{\log(O(d^{2d}) + 4d} = O(\sqrt{d\log(d)})$. Since $m = \frac{C^2K^4\left(\omega(T)+\sqrt{\log(\frac{2}{1-\rho})}\right)^2}{\epsilon^2}$, then $m = O(d\log(d))$.

Remark 4.1.9. Comparing our result to the work in [15, theorem 2], we have removed the mild geometric condition on reach $\frac{V}{\tau^d} \geq (\frac{21}{2\sqrt{d}})^d$ and can accommodate the presence of a boundary.

However the dependence of the final dimension, m, on the dimension of the manifold, d, remains $O(d\log(d))$.

4.2 With SORS Matrices

In this section we improve the work in [29], where a sufficient condition is given for finding JL maps via SORS matrices in terms of Gaussian width of the target set. The method is base on the Multi-Resolution Restricted Isometry Property (MRIP). Separately in the work of [10], an improvement in RIP estimates for SORS matrices is presented. Here we combine the improved RIP estimate with the MRIP framework and obtain an improved JL map for compact manifolds based on the Gaussian width of their unit secants. Theorem 4.2.16 is our main theorem in this section.

We start with the definition of SORS matrices and some properties.

Definition 4.2.1. *(SORS Matrix) Let M denote a $\mathbb{R}^{n\times n}$ orthogonal matrix (or $\mathbb{C}^{n\times n}$ where some authors use unitary matrices) obeying*

$$M^*M = I \qquad \text{and} \qquad \max_{i,j}|M_{i,j}| \leq \frac{K}{\sqrt{n}} \tag{4.14}$$

Let $H \in \mathbb{R}^{m\times n}$ be a random matrix created from selecting rows of M in an i.i.d fashion. Let $D \in \mathbb{R}^{n\times n}$ be a random diagonal matrix with diagonal entries of ± 1 of equal probability. Then $A = \sqrt{n}HD$ is a Subsampled Orthogonal with Random Sign (SORS) matrix with constant K. Note that $K \geq 1$ since $1 = \sum_i M_{ij}^2 \leq n(\frac{K}{\sqrt{n}})^2 = K^2$.

Definition 4.2.2. *(RIP)[18, Prop 6.1] The sth restricted isometry constant ϵ_s of a matrix $A \in \mathbb{R}^{m\times N}$ is the smallest $\epsilon \geq 0$ such that for all s-sparse $x \in \mathbb{R}^N$*

$$(1-\epsilon)||x||^2 \leq ||Ax||^2 \leq (1+\epsilon)||x||^2.$$

Proposition 4.2.3. *[18, Prop 6.6] For matrix A, let ϵ_s be the sth restricted isometry constant of A. Then for integers $1 \leq s \leq t$*

$$\epsilon_t \leq \frac{t-d}{s}\epsilon_{2s} + \frac{d}{s}\epsilon_s, \qquad d = gcd(s,t).$$

In particular since $\epsilon_s \leq \epsilon_{2s}$

$$\epsilon_t \leq \frac{t}{s}\epsilon_{2s}.$$

Notation 4.2.4. *For a matrix A, if* $\epsilon_s \leq \epsilon$*, then we say A satisfies the RIP of order* (s, ϵ)*.*

Proposition 4.2.5. *Let* $s \in \mathbb{N}$*, and* $1 \leq k$ *be a real number. Then*

$$\epsilon_s \leq k\epsilon_{(2\lceil \frac{s}{k} \rceil)}.$$

Proof. From proposition 4.2.3, or [18, Prop 6.6], for $1 \leq s \leq t$, we have $\epsilon_t \leq \frac{t}{s}\epsilon_{2s}$. Since $1 \leq k$ and s is an integer, $1 \leq \lceil \frac{s}{k} \rceil \leq s$, and hence $\epsilon_s \leq \frac{s}{\lceil \frac{s}{k} \rceil}\epsilon_{(2\lceil \frac{s}{k} \rceil)} \leq k\epsilon_{(2\lceil \frac{s}{k} \rceil)}$.

Definition 4.2.6. *[29, definition 2.1](Extended Restricted Isometry Property (ERIP)) A matrix* $A \in \mathbb{R}^{m\times N}$ *satisfies the ERIP of order* (s, ϵ) *if every* $x \in \mathbb{R}^N$ *with* $||x||_0 = s$ *satisfies*

$$\left| ||Ax||^2 - ||x||^2 \right| \leq \max\{\epsilon, \epsilon^2\}||x||^2$$

Remark 4.2.7. The above definition differs from the RIP when $1 < \epsilon$.

Definition 4.2.8. *[29, definition 2.2] (Multiresolution Restricted Isometry Property (MRIP)) A matrix* $A \in \mathbb{R}^{m\times N}$ *satisfies the MRIP with* (s, ϵ) *if it possesses the extended RIP for* $(2^l s, 2^{l/2}\epsilon)$ *for* $0 \leq l \leq \lfloor \log_2(\frac{N}{s}) \rfloor$.

Next we state an existing RIP estimate for SORS matrices from [18], followed by an improvement from [10]. The improvement has one fewer log factor.

Theorem 4.2.9. ***(RIP for SORS matrices)*** *[18, theorem 12.31]Let* $A \in \mathbb{C}^{m\times N}$ *be a SORS matrix with constant* $K \geq 1$*. For* $\epsilon \in (0, 1)$ *if*

$$m \geq CK^2\epsilon^{-2}s\log^4(N) \tag{4.15}$$

then with probability at least $1 - N^{-\log^3(N)}$ *the restricted isometry constant of* $\frac{1}{\sqrt{m}}A$ *for sparsity s, i.e* ϵ_s*, satisfies* $\epsilon_s \leq \epsilon$*. The constant* $C > 0$ *is universal.*

Now we present the improvement from [10].

Theorem 4.2.10. *[10, Thm 1.1] There exist absolute constants $\kappa, c_0, c_1 > 0$ such that the following holds. Let $X_1, ..., X_m$ be independent copies of a random vector $X \in \mathbb{C}^N$ with bounded coordinates, i.e. $\max_{1\le i\le N} |\langle X, \boldsymbol{e}_i\rangle| \le K$ for some $K > 0$, where $\{e_i\}_{i=1}^N$ is the standard basis of $\mathbb{C}^N$. Let $T \subseteq \{x \in \mathbb{C}^N : \|x\|_1 \le \sqrt{s}\}$, $\epsilon \in (0, \kappa)$, and assume that*

$$m \ge c_0 K^2 \epsilon^{-2} s \log(eN) \log^2 \left(sK^2/\epsilon\right).$$

Then with probability exceeding $1 - 2\exp\left(-\epsilon^2 m/(sK^2)\right)$,

$$\sup_{y\in T} \left| \frac{1}{m} \sum_{i=1}^{m} |\langle y, X_i\rangle|^2 - \mathbb{E}|\langle y, X\rangle|^2 \right| \le c_1 \epsilon \left(1 + \sup_{y\in T} \mathbb{E}|\langle y, X\rangle|^2\right).$$

Approximate values of constants are $\kappa \approx 0.306$, $c_0 \approx 316792$ and for c_1 we have $c_1 = 492$.

Corollary 4.2.11. *(Adapted from [10, Thm 1.1]) There exists absolute constants, $a_0, a_1, a_2 > 0$ such that the following holds for $\epsilon \in (0, a_2]$. Assume A is a $m \times N$ SORS matrix with constant K such that*

$$m \ge a_0 K^2 \frac{s}{\epsilon^2} \log(eN) \log^2 \left(\frac{a_1 s K^2}{\epsilon}\right),$$

then, with probability at least $1 - 2\exp(-\epsilon^2 m/(a_1^2 s K^2))$, for all s-sparse vectors $x \in \mathbb{R}^n$ we have

$$(1 - \epsilon)\|x\|^2 \le \left\| \frac{1}{\sqrt{m}} A x \right\|^2 \le (1 + \epsilon)\|x\|^2.$$

Proof. Using theorem 4.2.10, we consider unit length vectors $T := \{x \in \mathbb{C}^N$ with $\|x\|_2 = 1$ and $\|x\|_1 \le \sqrt{s}\}$. Since A is a SORS matrix, there is a unitary $\mathbb{C}^{N\times N}$ matrix U where A is sampled from $\sqrt{N}U$. Let X be a random vector uniformly distributed in the rows of $\sqrt{N}U$. Then we can apply theorem 4.2.10 to X. We have $\mathbb{E}|\langle x, X\rangle|^2 = \|x\|_2^2 = 1$ and $\left\|\frac{1}{\sqrt{m}} A x\right\|_2^2 = \frac{1}{m}\sum_{i=1}^m |\langle x, X_i\rangle|^2$. Thus

$$\sup_{x\in T} \left| \left\| \frac{1}{\sqrt{m}} A x \right\|_2^2 - \|x\|_2^2 \right| \le 2c_1\epsilon.$$

Changing constants to account for the extra $2c_1$ factor accompanying the ϵ above gives the stated bounds on the probability and m. The new constants and their approximate values are $a_0 = c_0(2c_1)^2 \approx 3 \times 10^{11}$, $a_1 = 2c_1 = 984$, $a_2 = 2c_1\kappa \approx 301$. ■

In the next corollary, we remove the restriction on the size of ϵ in corollary 4.2.11. We also introduce a parameter ρ for more explicit control on the probability of success. Using ρ we relate the number of rows, m, to the success probability.

Corollary 4.2.12. *Let A be a $m \times N$ SORS matrix with constant K. Let $0 < s, \epsilon, \rho$, with s an integer and ϵ, ρ real numbers. There exists absolute constants, $a_0, a_1, a_2 > 0$ such that the following holds.*

If $\epsilon \le a_2$, and

$$m \ge a_0 K^2 \frac{s}{\epsilon^2} \log(eN) \log^2(\frac{a_1 s K^2}{\epsilon})(1+\rho) \tag{4.16}$$

then with probability at least $1 - 2\exp\left(\frac{-a_0}{a_1^2} \log(eN) \log^2(\frac{a_1 s K^2}{\epsilon})(1+\rho)\right)$, for all s-sparse vectors $x \in \mathbb{R}^n$ we have

$$(1-\epsilon)||x||^2 \le ||\frac{1}{\sqrt{m}} A x||^2 \le (1+\epsilon)||x||^2. \tag{4.17}$$

If $\epsilon \ge a_2$ and

$$m \ge a_0 K^2 \frac{2\lceil \frac{a_2 s}{\epsilon} \rceil}{a_2^2} \log(eN) \log^2(\frac{a_1 K^2 2\lceil \frac{a_2 s}{\epsilon} \rceil}{a_2})(1+\rho) \tag{4.18}$$

then with probability at least $1 - 2\exp\left(\frac{-a_0}{a_1^2} \log(eN) \log^2(\frac{a_1 2\lceil \frac{a_2 s}{\epsilon} \rceil K^2}{a_2})(1+\rho)\right)$, for all s-sparse vectors $x \in \mathbb{R}^n$ we have

$$(1-\epsilon)||x||^2 \le ||\frac{1}{\sqrt{m}} A x||^2 \le (1+\epsilon)||x||^2.$$

Proof. The case of $\epsilon \le a_2$ is the same as before, with a simplified expression for probability. In the simplification we plug in the expression for m into the probability bound and simplify. For the case of $\epsilon > a_2$, in the formula for $\epsilon \le a_2$ we put $\epsilon = a_2$ and $s = 2\lceil \frac{a_2 s}{\epsilon} \rceil$. By proposition 4.2.5, from RIP of order $(2\lceil \frac{a_2 s}{\epsilon} \rceil, a_2)$, putting $k = \frac{\epsilon}{a_2}$, we can get RIP of order (s, ϵ). In this process we need

$$m \ge a_0 K^2 \frac{2\lceil \frac{a_2 s}{\epsilon} \rceil}{a_2^2} \log(eN) \log^2(\frac{a_1 K^2 2\lceil \frac{a_2 s}{\epsilon} \rceil}{a_2})(1+\rho).$$

Substituting the value of m in the probability bound, we get the claimed expression. ■

In the next corollary we combine the two regimes of ϵ into one expression.

Corollary 4.2.13. *There exist constants a_0, a_1, a_2 such that for a SORS matrix A of size $m \times N$ with constant K, and positive integer s and positive reals ϵ and ρ the following holds. If*

$$m \geq a_0 K^2 (\frac{1}{\epsilon^2} + \frac{2}{a_2^2}) s \log(eN) \log^2 \left((\frac{1}{\epsilon} + \frac{2}{a_2}) a_1 s K^2 \right) (1 + \rho)$$

then the matrix $\frac{1}{\sqrt{m}} A$ satisfies the RIP condition for (s, ϵ) with probability at least

$$1 - 2 \exp \left(\frac{-a_0}{a_1^2} \log(eN) \log^2 (\frac{a_1}{a_2}) (1 + \rho) \right).$$

Proof. We combine the two bounds (4.16) and (4.18) for m from the cases of $\epsilon \leq a_2$ and $\epsilon > a_2$ into a single bound. Similarly we combine the probability bounds into one expressions. This final expression is independent of s and ϵ and it relates to m through the variable ρ. For the m bound, when $\epsilon \geq a_2$ we used $\frac{2\lceil \frac{a_2 S}{\epsilon} \rceil}{a_2^2} \leq \frac{2S}{a_2^2}$. For the probability bounds first note that $1 - \exp(-x)$ is an increasing function. Then when $\epsilon \leq a_2$ we have $\frac{a_1 s K^2}{\epsilon} \geq \frac{a_1}{a_2}$, since $s, K \geq 1$. When $\epsilon \geq a_2$, then $\frac{a_1 K^2 2 \lceil \frac{a_2 s}{\epsilon} \rceil}{a_2} \geq \frac{a_1}{a_2}$. ■

The focus of the next lemma is transitioning from RIP bounds to their multi-resolution (MRIP) counterparts.

Proposition 4.2.14. *Let A be a $m \times N$ SORS matrix with constant K. Let s be a positive integer, $\epsilon < 1$ and ρ be positive reals. There exists absolute constants, $a_0, a_1, a_2 > 0$ such that if*

$$m \geq 2 a_0 K^2 (\frac{1}{\epsilon^2} + \frac{1}{a_2^2} + \frac{1}{a_2^3}) s \log(eN) \log^2 \left(2 a_1 s K^2 (\frac{1}{\epsilon^2} + \frac{1}{a_2} + \frac{1}{a_2^2}) \right) (1 + \rho) \tag{4.19}$$

then $\frac{1}{\sqrt{m}} A$ satisfies the MRIP (s, ϵ) criteria with probability at least

$$1 - 2 \exp \left(\log(\log(N)) - \frac{a_0}{a_1^2} \log(eN) \log^2 (\frac{a_1}{a_2}) (1 + \rho) \right).$$

Proof. The probability expression is simpler to obtain so we start with it. We take the expression of probability from corollary 4.2.13, and apply it $\log N$ times. In the definition of MRIP, given a sparsity level s and distortion ϵ, we need to get ERIP conditions of the form $(2^i s, 2^{\frac{i}{2}} \epsilon)$ or $(2^i s, 2^i \epsilon^2)$ for i starting from 0 and continuing until $2^i s$ is larger than N. Therefore there are at most $\log N$ steps. Since the probability expression in corollary 4.2.13 is independent of s and ϵ, we can use the same expression for different steps. Combining the failure probabilities by the union bound we arrive at the expression

$$1 - 2\exp\left(\log(\log(N)) - \frac{a_0}{a_1^2}\log(eN)\log^2(\frac{a_1}{a_2})(1+\rho)\right). \tag{4.20}$$

Now with respect to m, we need to show that the given bound for m, (4.19), is large enough for the different ERIP conditions required in the MRIP. Here we find it more convenient to return to corollary 4.2.12. First we check the condition for the regime $\epsilon \le a_2$, over the different i values $(s, \epsilon) \rightarrow (2^i s, 2^{\frac{i}{2}}\epsilon)$ when $2^{\frac{i}{2}}\epsilon \le 1$.

$$m \ge a_0 K^2 \frac{s}{\epsilon^2}\log(eN)\log^2(\frac{a_1 s K^2}{\epsilon})(1+\rho) \rightarrow \tag{4.21}$$

$$m \ge a_0 K^2 \frac{2^i s}{(2^{\frac{i}{2}}\epsilon)^2}\log(eN)\log^2(\frac{a_1 2^i s K^2}{(2^{\frac{i}{2}}\epsilon)})(1+\rho) \tag{4.22}$$

$$m \ge a_0 K^2 \frac{s}{\epsilon^2}\log(eN)\log^2(2^{\frac{i}{2}}\frac{a_1 s K^2}{\epsilon})(1+\rho) \tag{4.23}$$

The transition from $(2^i s, 2^{\frac{i}{2}}\epsilon)$ to $(2^i s, 2^i \epsilon^2)$ happens when for $\epsilon < 1$,

$$2^{\frac{i}{2}}\epsilon = 1 \rightarrow 2^{\frac{i}{2}} = \frac{1}{\epsilon} \tag{4.24}$$

So we get the bound for m as

$$m \ge a_0 K^2 \frac{s}{\epsilon^2}\log(eN)\log^2(\frac{a_1 s K^2}{\epsilon^2})(1+\rho). \tag{4.25}$$

Now we plug-in $(2^i s, 2^i \epsilon^2)$ for the regime when $2^i \epsilon^2 > 1$ to get

$$m \ge a_0 K^2 \frac{2^i s}{(2^i \epsilon^2)^2}\log(eN)\log^2(\frac{a_1 2^i s K^2}{(2^i \epsilon)})(1+\rho) \tag{4.26}$$

$$\ge a_0 K^2 \frac{s}{2^i \epsilon^4}\log(eN)\log^2(\frac{a_1 s K^2}{\epsilon})(1+\rho) \tag{4.27}$$

and this bound for m get smaller as i grows so we can just work with the earlier one in (4.25). Now we work with the case for $\epsilon > a_2$. Again we plug in $(2^i s, 2^{\frac{i}{2}}\epsilon)$ or $(2^i s, 2^i \epsilon^2)$, and look for a combined upper bound. First substitute $(2^i s, 2^{\frac{i}{2}}\epsilon)$ into (4.18)

$$m \geq a_0 K^2 \frac{2\lceil \frac{a_2 s}{\epsilon} \rceil}{a_2^2} \log(eN) \log^2(\frac{a_1 K^2 2\lceil \frac{a_2 s}{\epsilon} \rceil}{a_2})(1+\rho) \rightarrow \tag{4.28}$$

$$m \geq a_0 K^2 \frac{2\lceil \frac{a_2 2^i s}{2^{\frac{i}{2}}\epsilon} \rceil}{a_2^2} \log(eN) \log^2(\frac{a_1 K^2 2\lceil \frac{a_2 2^i s}{2^{\frac{i}{2}}\epsilon} \rceil}{a_2})(1+\rho) \tag{4.29}$$

$$m \geq a_0 K^2 \frac{2\lceil \frac{a_2 2^{\frac{i}{2}} s}{\epsilon} \rceil}{a_2^2} \log(eN) \log^2(\frac{a_1 K^2 2\lceil \frac{a_2 2^{\frac{i}{2}} s}{\epsilon} \rceil}{a_2})(1+\rho) \tag{4.30}$$

Similar to above $2^{\frac{i}{2}}$ is at most $\frac{1}{\epsilon}$ when $2^{\frac{i}{2}}\epsilon \leq 1$ so we arrive at

$$m \geq a_0 K^2 \frac{2\lceil \frac{a_2 s}{\epsilon^2} \rceil}{a_2^2} \log(eN) \log^2(\frac{a_1 K^2 2\lceil \frac{a_2 s}{\epsilon^2} \rceil}{a_2})(1+\rho)$$

Next we try plugging in $(2^i s, 2^i \epsilon^2)$ and we arrive at

$$m \geq a_0 K^2 \frac{2\lceil \frac{a_2 2^i s}{(2^i \epsilon^2)^2} \rceil}{a_2^2} \log(eN) \log^2(\frac{a_1 K^2 2\lceil \frac{a_2 2^i s}{(2^i \epsilon^2)^2} \rceil}{a_2})(1+\rho) \tag{4.31}$$

$$m \geq a_0 K^2 \frac{2\lceil \frac{a_2 s}{\epsilon^2 (2^i \epsilon^2)} \rceil}{a_2^2} \log(eN) \log^2(\frac{a_1 K^2 2\lceil \frac{a_2 s}{\epsilon^2 (2^i \epsilon^2)} \rceil}{a_2})(1+\rho). \tag{4.32}$$

We note that in this case by assumption $2^i \epsilon^2 \geq 1$ so we can drop it and get a stricter lower bound for m as

$$m \geq a_0 K^2 \frac{2\lceil \frac{a_2 s}{\epsilon^2} \rceil}{a_2^2} \log(eN) \log^2(\frac{a_1 K^2 2\lceil \frac{a_2 s}{\epsilon^2} \rceil}{a_2})(1+\rho). \tag{4.33}$$

So overall for $\epsilon \leq a_2$ and $\epsilon > a_2$ we respectively need

$$m \geq a_0 K^2 \frac{s}{\epsilon^2} \log(eN) \log^2(\frac{a_1 s K^2}{\epsilon^2})(1+\rho)$$

and

$$m \geq a_0 K^2 \frac{2\lceil \frac{a_2 s}{\epsilon^2} \rceil}{a_2^2} \log(eN) \log^2(\frac{a_1 K^2 2\lceil \frac{a_2 s}{\epsilon^2} \rceil}{a_2})(1+\rho). \tag{4.34}$$

We can combine them both into one expression with as claimed in the proposition.

$$m \geq 2a_0 K^2 (\frac{1}{\epsilon^2} + \frac{1}{a_2^2} + \frac{1}{a_2^3}) s \log(eN) \log^2\left(2a_1 s K^2 (\frac{1}{\epsilon^2} + \frac{1}{a_2} + \frac{1}{a_2^2})\right)(1+\rho). \tag{4.35}$$

where for $\epsilon > a_2$, we used $(\frac{1}{\epsilon^2} + \frac{1}{a_2^2} + \frac{1}{a_2^3})s \geq \frac{s}{a_2^3} + \frac{1}{a_2^2} \geq \frac{\lceil \frac{a_2 s}{\epsilon^2} \rceil}{a_2^2}$. ■

Having established the MRIP estimate for SORS matrices, we use the results of [29] to get a JL map for compact manifolds. First we state the result in [29].

Theorem 4.2.15. *Let $T \subset \mathbb{R}^N$ and suppose the matrix $H \in \mathbb{R}^{m \times N}$ obeys the multiresolution RIP with sparsity and distortion levels*

$$s = 150(1+\eta), \qquad \delta = \frac{\epsilon \, rad(T)}{C \max(rad(T), \omega(T))} \tag{4.36}$$

with $C > 0$ an absolute constant. Then, for a diagonal matrix D with an i.i.d random sign pattern on the diagonal, the matrix $A = HD$ obeys

$$\sup_{x \in T} \left| ||Ax||^2 - ||x||^2 \right| \leq \max(\epsilon, \epsilon^2).rad(T)^2 \tag{4.37}$$

with probability at least $1 - \exp(-\eta)$. Here $rad(T) = \sup_{v \in T} ||v||$ is the maximum of the Euclidean norm of points inside T. ■

Now we use our proposition 4.2.14 with theorem 4.2.15.

Theorem 4.2.16. *Let $M \hookrightarrow \mathbb{R}^N$ be a compact d-dimensional submanifold of $\mathbb{R}^N$ with boundary ∂M. Let τ_M be the reach of M. Let τ_i be the reach of the i-th connected component of ∂M as a submanifold of $\mathbb{R}^N$. Let $\tau = \inf_{M,i}\{\tau_M, \tau_i\}$. Let V_M be the volume of M and $V_{\partial M}$ be the volume of ∂M. Let $U = U(M - M) = \{\frac{p-q}{||p-q||} \mid p \neq q,\ p, q \in M\}$ be the set of unit secants of M, and let $\overline{U}$ be its closure. Let $\omega(\overline{U})$ be the Gaussian width of U. Define*

$$\alpha = \frac{V_M}{\omega_d}(\frac{41}{\tau})^d + \frac{V_{\partial M}}{\omega_{d-1}}(\frac{81}{\tau})^{d-1} \tag{4.38}$$

$$c = (\frac{\alpha^2}{2} + 3^d \alpha) \tag{4.39}$$

$$w^* = 4\sqrt{2}\sqrt{\log(c) + 4d} \tag{4.40}$$

Let ρ, η and ϵ be positive reals. Let $0 < \epsilon < 1$. Let A be a SORS matrix of size $m \times N$ with constant K.

There exist absolute constants a_0, a_1, a_2 and C such that if

$$m \geq a_0 K^2 \frac{(Cw^{*2})}{\epsilon^2} + \frac{1}{a_2^2} + \frac{1}{a_2^3})(1+\eta)\log(eN)\log^2\left((1+\eta)a_1 K^2 \frac{(Cw^{*2})}{\epsilon^2} + \frac{1}{a_2} + \frac{1}{a_2^2})\right)(1+\rho) \tag{4.41}$$

then for all $x, y \in M$ we have

$$\left|\|A(x-y)\|^2 - \|x-y\|^2\right| \leq \epsilon||x-y||^2 \tag{4.42}$$

with probability

$$1 - \exp(-\eta) - 2\exp\left(\log(\log(N)) - \frac{a_0}{a_1^2}\log(eN)\log^2(\frac{a_1}{a_2})(1+\rho)\right). \tag{4.43}$$

Proof. We use proposition 4.2.14 with theorem 4.2.15. We consider $U(M-M) \subset S^{N-1}$. We have w^* as an upper bound for the Gaussian width of $U(M-M)$. Therefore to get an ϵ JL map for $U(M-M)$ we need to establish $\frac{1}{\sqrt{m}}A$ has the MRIP property for (s, δ) where s determines the probability and $\delta = \frac{\epsilon \operatorname{rad}(U(M-M))}{C\max(\operatorname{rad}(U(M-M)), \omega(U(M-M)))}$. We have $\operatorname{rad}(U(M-M)) = 1$, and $w^* > 1$, and we get $\delta = \frac{\epsilon}{Cw^*}$.

To get such an MRIP we need

$$m \geq 2a_0 K^2(\frac{(Cw*)^2}{\epsilon^2} + \frac{1}{a_2^2} + \frac{1}{a_2^3})150(1+\eta)\log(eN)\log^2\left(2a_1(150)(1+\eta)K^2(\frac{(Cw*)^2}{\epsilon^2} + \frac{1}{a_2} + \frac{1}{a_2^2})\right)(1+\rho). \tag{4.44}$$

and the success probability is

$$1 - 2\exp\left(\log(\log(N)) - \frac{a_0}{a_1^2}\log(eN)\log^2(\frac{a_1}{a_2})(1+\rho)\right). \tag{4.45}$$

Going from MRIP to an ϵ-JL map has a failure probability of its own as $\exp(-\eta)$. Combining the failure probabilities we arrive at

$$1 - \exp(-\eta) - 2\exp\left(\log(\log(N)) - \frac{a_0}{a_1^2}\log(eN)\log^2(\frac{a_1}{a_2})(1+\rho)\right). \tag{4.46}$$

■

CHAPTER 5

NECESSARY CONDITIONS FOR JL MAPS OF MANIFOLDS (VARIATION 2)

In this section we give a necessary condition for existence of JL maps for compact manifolds. Earlier work of Alon in [5, Theorem 9.3.] provides a necessary condition for a JL map of finite points. Our work is a partial generalization of that bound to manifolds. Our method is based on the Sudokov inequality [35, Theorem 7.4.1] which bounds the Gaussian width from below using covering numbers. Complementing with Sudakov-Fernique's comparison theorem [35, Def 7.2.11], we control the reduction in the Gaussian width under the JL map. The image of the manifold under the map must fit in the target space and comparison of the Gaussian widths of the unit sphere of the target space with the Gaussian width of the manifold gives us the desired necessary condition. Our main theorem in this section is 5.0.5. After that we consider a reduction from manifolds to finite points, and compare our work to the lower bound in [5, Theorem 9.3.].

We start by reviewing the definitions and give estimates of the Gaussian width.

Definition 5.0.1. *[35, Def 7.5.1] Let g be a standard Gaussian random variable in $\mathbb{R}^n$. Define the Gaussian Width of $T \subset \mathbb{R}^n$, $\omega(T)$, as follows.*

$$\omega(T) = \mathbb{E} \sup_{x \in T} \langle g, x \rangle$$

Proposition 5.0.2. *Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map, $T \subset \mathbb{R}^n$ with a and b positive reals such that for $x, y \in T$, $a||x-y|| \leq ||L(x-y)|| \leq b||x-y||$. Then $a\,\omega(T) \leq \omega(L(T)) \leq b\,\omega(T)$.*

Proof. We use the Sudakov-Fernique's comparison theorem [35, Def 7.2.11]. For $t \in T$, define the Gaussian processes $A_t = a\,\langle g_n, t\rangle, B_t = \langle g_m, Lt\rangle, C_t = b\,\langle g_n, t\rangle$. For their increments we have the estimates

$$\mathbb{E}(A_{t_1} - A_{t_2})^2 = a^2\mathbb{E}(\langle g_n, t_1 - t_2\rangle)^2 = a^2||t_1 - t_2||^2 \leq$$

$$\mathbb{E}(B_{t_1} - B_{t_2})^2 = \mathbb{E}(\langle g_m, L(t_1 - t_2)\rangle)^2 = ||L(t_1 - t_2)||^2 \leq$$

$$\mathbb{E}(C_{t_1} - C_{t_2})^2 = b^2\mathbb{E}(\langle g_n, t_1 - t_2\rangle)^2 = b^2||t_1 - t_2||^2.$$

By the comparison theorem $\mathbb{E}\sup_t A_t \leq \mathbb{E}\sup_t B_t \leq \mathbb{E}\sup_t C_t$. This implies the result as $\mathbb{E}\sup_t A_t = a\,\mathbb{E}\sup_t\langle g_n, t\rangle = a\omega(T)$ with similar expressions for B_t and C_t. ■

The next proposition gives the necessary condition on the target dimension given the requirement of the Gaussian widths.

Proposition 5.0.3. *Let $M \subset \mathbb{R}^n$ have Gaussian width $\omega(M)$, and L be a linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$, and $0 < \epsilon < 1$. Assume for all $x, y \in M$*

$$(1-\epsilon)||x-y|| \leq ||L(x-y)|| \leq (1+\epsilon)||x-y||.$$

Then

$$\left(\frac{1}{2}(\frac{1-\epsilon}{1+\epsilon})\frac{\omega(M)}{\operatorname{diam}(M)}\right)^2 \leq m.$$

Proof. Consider the set $M - M$ (defined as $M + (-M)$). By proposition 5.0.2,

$$(1-\epsilon)\omega(M-M) \leq \omega(L(M-M))$$

We have $\omega(M-M) = \frac{1}{2}\omega(M)$ [35, proposition 7.5.2]. Image of $L(M-M)$ is contained in a ball in $\mathbb{R}^m$ with radius $(1+\epsilon)(\operatorname{diam}(M))$. Since $\omega(B^m(1)) = \sqrt{m}$ [35, 7.5.7], then by monotonicity of Gaussian width and scaling $\omega(L(M-M)) \leq (1+\epsilon)\operatorname{diam}(M)\sqrt{m}$. Then

$$(1-\epsilon)\frac{1}{2}\omega(M) = (1-\epsilon)\omega(M-M) \leq \omega(L(M-M)) \leq (1+\epsilon)\operatorname{diam}(M)\sqrt{m}.$$

■

In the next proposition we combine the Sudakov inequality with the lower covering estimates in corollary 2.4.2. When we optimize the Sudakov inequality over the choice of covering radii, we arrive at two possible regimes.

Proposition 5.0.4. *Let M be a d-dimensional smooth submanifold of $\mathbb{R}^n$ possibly with boundary, with volume V and reach τ. Let $0 < \epsilon \leq \frac{\tau}{2}$. Let $\omega_d = \frac{\pi^{d/2}}{\Gamma(\frac{d}{2}+1)}$ be the volume of the unit d-ball. There there is a universal constant c such that, if $\frac{2}{3\sqrt{e}}(\frac{V}{\omega_d})^{\frac{1}{d}} \leq \tau$ then*

$$\frac{c}{3}\sqrt{\frac{d}{2e}}(\frac{V}{\omega_d})^{\frac{1}{d}} \leq \omega(M).$$

If $\tau \leq \frac{2}{3\sqrt{e}}(\frac{V}{\omega_d})^{\frac{1}{d}}$ *then*

$$c\frac{\tau}{2}\sqrt{\log(\frac{V}{\omega_d(\frac{3}{2}\tau)^d})} \leq \omega(M).$$

Proof. We use Sudakov's inequality [35, theorem 7.4.1].

$$c\epsilon\sqrt{\log(C(M,\epsilon))} \leq \omega(M)$$

Sudakov's inequality is valid for any radius ϵ. We further restrict to $\epsilon \leq \frac{\tau}{2}$ so we can use $\frac{V}{\omega_d(3\epsilon)^d} \leq C(M,\epsilon)$, corollary 2.4.2. We obtain

$$c\epsilon\sqrt{\log(\frac{V}{\omega_d(3\epsilon)^d})} \leq \omega(M) \tag{5.1}$$

Maximizing the left hand side with respect to ϵ, for the optimal $\epsilon = \epsilon^*$, we obtain

$$\sqrt{\log(\frac{V}{\omega_d(3\epsilon)^d})} - \frac{d}{2\sqrt{\log(\frac{V}{\omega_d(3\epsilon)^d})}} = 0$$

$$\log(3\epsilon^*) = \frac{1}{d}\log(\frac{V}{\omega_d}) - \frac{1}{2}$$

$$\epsilon^* = \frac{1}{3\sqrt{e}}(\frac{V}{\omega_d})^{\frac{1}{d}}$$

Using this optimal choice in equation (5.1),we get

$$\frac{c}{3}\sqrt{\frac{d}{2e}}(\frac{V}{\omega_d})^{\frac{1}{d}} \leq \omega(M).$$

If the optimal ϵ^* is forbidden because of $\epsilon \leq \frac{\tau}{2}$, then we just use $\epsilon = \frac{\tau}{2}$. It gives the inequality

$$c\frac{\tau}{2}\sqrt{\log(\frac{V}{\omega_d(\frac{3}{2}\tau)^d})} \leq \omega(M).$$

For a valid formula we must have

$$0 \leq \log(\frac{V}{\omega_d(\frac{3}{2}\tau)^d})$$

$$1 \leq \frac{V}{\omega_d(\frac{3}{2}\tau)^d}$$

$$\tau \leq \frac{2}{3}(\frac{V}{\omega})^{\frac{1}{d}} \tag{5.2}$$

Because we are considering the case where the optimal ϵ^* is forbidden, we have

$$\frac{\tau}{2} < \epsilon^* = \frac{1}{3\sqrt{e}}(\frac{V}{\omega_d})^{\frac{1}{d}}$$
$$\tau < \frac{2}{3\sqrt{e}}(\frac{V}{\omega_d})^{\frac{1}{d}}$$

This is enough for satisfying (5.2). ■

Now we put together the requirement on the target dimension of a JL map, with the geometric properties of the manifold and its Gaussian width.

Theorem 5.0.5. *Let $M \subset \mathbb{R}^n$ with Gaussian width $\omega(M)$, L a linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$, and $0 < \epsilon < 1$ and let c represent multiple universal constants. Assume for all $x, y \in M$*

$$(1-\epsilon)||x-y|| \le ||L(x-y)|| \le (1+\epsilon)||x-y||.$$

If $\frac{2}{3\sqrt{e}}(\frac{V}{\omega_d})^{\frac{1}{d}} \le \tau$ then

$$c(\frac{1-\epsilon}{1+\epsilon})^2\frac{d}{\text{diam}^2(M)}(\frac{V}{\omega_d})^{\frac{2}{d}} \le m.$$

If $\tau \le \frac{2}{3\sqrt{e}}(\frac{V}{\omega_d})^{\frac{1}{d}}$ then

$$c\tau^2(\frac{1-\epsilon}{1+\epsilon})^2\frac{1}{\text{diam}^2(M)}\log(\frac{V}{\omega_d(\frac{3}{2}\tau)^d}) \le m.$$

Proof. We use proposition 5.0.3 and proposition 5.0.4. We have a lower bound for the embedding dimension m using the distortion ϵ and the Gaussian width of M, $\omega(M)$. Then we have a lower bound for the $\omega(M)$ using the covering numbers of M. Putting the two together we get the following.

If $\frac{2}{3\sqrt{e}}(\frac{V}{\omega_d})^{\frac{1}{d}} \le \tau$

$$\frac{2c}{3}\sqrt{\frac{d}{2e}}(\frac{V}{\omega_d})^{\frac{1}{d}} \le \omega(M)$$
$$\left(\frac{1}{2}(\frac{1-\epsilon}{1+\epsilon})\frac{\omega(M)}{\text{diam}(M)}\right)^2 \le m$$
$$c(\frac{1-\epsilon}{1+\epsilon})^2\frac{d}{\text{diam}^2(M)}(\frac{V}{\omega_d})^{\frac{2}{d}} \le m.$$

If $\tau \leq \frac{2}{3\sqrt{e}}(\frac{V}{\omega_d})^{\frac{1}{d}}$ then

$$c\frac{\tau}{2}\sqrt{\log(\frac{V}{\omega_d(\frac{3}{2}\tau)^d})} \leq \omega(M)$$
$$\left(\frac{1}{2}(\frac{1-\epsilon}{1+\epsilon})\frac{\omega(M)}{\text{diam}(M)}\right)^2 \leq m$$
$$c\tau^2(\frac{1-\epsilon}{1+\epsilon})^2\frac{1}{\text{diam}^2(M)}\log(\frac{V}{\omega_d(\frac{3}{2}\tau)^d}) \leq m.$$

■

Standard Examples: We apply our theorem 5.0.5 to the standard examples of unit sphere S^d and unit ball B^d as submanifold of $\mathbb{R}^N$.

Corollary 5.0.6.

For B^d, the data is as follows.

M	V_{B^d}	τ_{B^d}	ω_d	diam(M)
B^d	$\frac{\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2}+1)}$	∞	V_{B^d}	2

Since τ is infinite, we have

$$c(\frac{1-\epsilon}{1+\epsilon})^2\frac{d}{\text{diam}^2(M)}(\frac{V}{\omega_d})^{\frac{2}{d}} \leq m$$
$$c(\frac{1-\epsilon}{1+\epsilon})^2 d \leq m$$

which agrees with our intuition that for a perfect, no distortion embedding of unit B^d, target dimension grows at least linearly in d.

■

Corollary 5.0.7.

For the unit $\mathbb{S}^d$ the data is as follows.

M	V_{S^d}	τ_{S^d}	ω_d	diam(M)
S^d	$2\frac{\pi^{\frac{d+1}{2}}}{\Gamma(\frac{d+1}{2})}$	1	$\frac{\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2}+1)}$	2

We check the relation between reach and the other parameters to find that τ is big enough.

$$\frac{2}{3\sqrt{e}}(\frac{V}{\omega_d})^{\frac{1}{d}} = \frac{2}{3\sqrt{e}}\left(2\sqrt{\pi}\frac{\Gamma(\frac{d}{2}+1)}{\Gamma(\frac{d+1}{2})}\right)^{\frac{1}{d}}$$

Using $\lim_{d\to\infty}\left(2\sqrt{\pi}\frac{\Gamma(\frac{d}{2}+1)}{\Gamma(\frac{d+1}{2})}\right)^{\frac{1}{d}} = 1$, for large enough d we have

$$\frac{2}{3\sqrt{e}}(\frac{V}{\omega_d})^{\frac{1}{d}} < \tau = 1.$$

So we get

$$c(\frac{1-\epsilon}{1+\epsilon})^2\frac{d}{\operatorname{diam}^2(M)}(\frac{V}{\omega_d})^{\frac{2}{d}} \leq m$$
$$c(\frac{1-\epsilon}{1+\epsilon})^2 d \leq m.$$

This gives a linear dependence on d which agrees which our intuition that to embed S^d without distortion at least d dimensions are needed. The reason is standard S^d circumscribes $d+2$ points that are equidistant from each other and such points need at least $\mathbb{R}^{d+1}$ for an embedding.

∎

Reduction to the case of finite points: Now we show that when we reduce to the zero-dimensional case we recover the finite point case. Consider n points in $\mathbb{R}^N$. There are at most $\binom{n}{2}$ pair-wise secants. Normalizing them we get some points on the unit S^{N-1}. Call this set M. We apply Sudakov's inequality to the set M. The covering number with balls of radius ϵ for $\epsilon < \frac{\tau}{2}$ is exactly $|M|$. Using the counting measure we put $V = |M|$. Then by Sudakov's inequality for $\epsilon = \frac{\tau}{2}$ we get

$$c\frac{\tau}{2}\sqrt{\log(V)} \leq \omega(M)$$
$$c\frac{\tau}{2}\sqrt{\log(\frac{n(n-1)}{2})} \leq \omega(M)$$
$$c\frac{\tau}{4}\sqrt{\log(n)} \leq \omega(M).$$

Using proposition 5.0.3

$$\left(\frac{1}{2}(\frac{1-\epsilon}{1+\epsilon})\frac{c\frac{\tau}{4}\sqrt{\log(n)}}{\operatorname{diam}(M)}\right)^2 \leq \left(\frac{1}{2}(\frac{1-\epsilon}{1+\epsilon})\frac{\omega(M)}{\operatorname{diam}(M)}\right)^2 \leq m. \tag{5.3}$$

For finite points the reach τ is equal to half of the smallest distance between the points and diameter is the largest distance hence $\frac{\tau}{\text{diam}} \leq \frac{1}{2}$. Therefore this term won't push the lower bound up enough to contradict the known upper bounds. Simplifying equation (5.3) we arrive at

$$c(\frac{\tau}{\text{diam}})^2(\frac{1-\epsilon}{1+\epsilon})^2 \log(n) \leq m.$$

Comparing to the well known upper JL bound $O(\frac{1}{\epsilon^2})\log(n)$, [35, Theorem 5.3.1], we see that we recover the $\log(n)$. A known lower bound for finite points from [5, Theorem 9.3.] is $O(\frac{1}{\epsilon^2 \log(\frac{1}{\epsilon})})\log(n)$. Comparing to our bound, we have the same dependence on $\log(n)$, however our ϵ dependence of $\frac{1-\epsilon}{1+\epsilon}$ is weaker than $\frac{1}{\epsilon^2 \log(\frac{1}{\epsilon})}$. ■

CHAPTER 6

A JL ALGORITHM WITH $N\log(\log(N))$ RUN-TIME (VARIATION 3)

In this chapter we give an algorithm for computing a JL map with a $N\log(\log(N))$ run-time. This approach improves on the work of Ailon and Liberty in [2] for fast JL maps. First we present the algorithm for finite point sets, and after we upgrade it to the case of subspaces. Since we restrict the discussion to finite points and subspaces, our bounds don't directly involve the reach as a parameter.

Our algorithm is based on the divide and conquer approach, where we split a vector into pieces, apply a common fast JL map to each piece, combine the outputs and apply a sub-Gaussian JL map at the end. An schematic diagram for the algorithm in presented in figure 6.1. In this way we maintain the best of two worlds: the run-time speed of fast JL maps and the optimal embedding dimension of sub-Gaussian matrices. In lemma 6.0.1, we show the general framework for combining fast and sub-Gaussian JL maps work together. Our main theorems are 6.1.5 for finite points and 6.2.5 for the case of subspaces.

Although our theorems are stated for general matrices with fast matrix-vector multiply, Discrete Fourier Transform (DFT) and Welsh-Hadamard matrices are prime examples of matrices that have the fast runtimes.

Lemma 6.0.1. *Let $\epsilon \in \left(0, \frac{1}{3}\right)$, $S \subset \mathbb{C}^N$ be a set with n points. Assume $N > m_1^2$ for integers $m_1 \geq m_2$ where m_1^2 divides N. Furthermore*

1. *Split every vector in S into N/m_1^2 pieces each in $\mathbb{C}^{m_1^2}$ and obtain the set $S' \subset \mathbb{C}^{m_1^2}$ of cardinality nN/m_1^2.*
2. *Let $\mathbf{A} \in \mathbb{C}^{m_1 \times m_1^2}$ be an ϵ-JL map of S' into $\mathbb{C}^{m_1}$.*
3. *Suppose $\mathbf{A}$ has a fast matrix vector multiply so that $\mathbf{Ay}$ can be computed in $m_1^2 \cdot f(m_1) = o(m_1^3)$-time for all $\mathbf{y} \in \mathbb{C}^{m_1^2}$.*

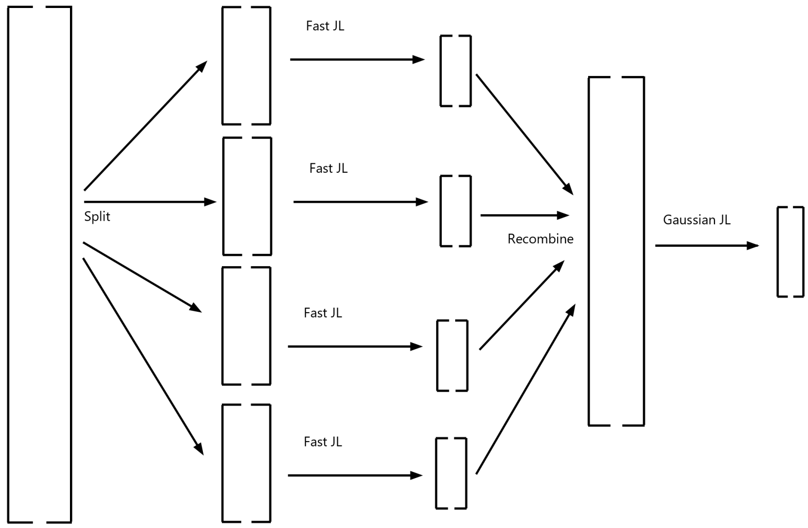


Figure 6.1: A vector is split into pieces, each part is processed with a fast JL map, outputs are combined and fed into a Gaussian JL map. The Gaussian maps have optimal dimensional reduction while fast JL maps have better run-time. In this mixed approach, we combine the benefits of both methods.

4. *Let $S'' \subset \mathbb{C}^{N/m_1}$ be a set of n points obtained by reshaping the N/m_1^2 pieces of each vector in S back into a single vector after they've been mapped into $\mathbb{C}^{m_1}$ by* $\mathbf{A}$*, and*

5. *Let $\mathbf{B} \in \mathbb{C}^{m_2 \times N/m_1}$ be any ϵ-JL map of S'' into $\mathbb{C}^{m_2}$.*

Then, $L(\mathbf{x}) := \mathbf{B}\begin{pmatrix} A & & \\ & \ddots & \\ & & A \end{pmatrix}$ *will be a 3ϵ-JL map of $S \subset \mathbb{C}^N$ into $\mathbb{C}^{m_2}$. Furthermore, L can be applied to any vector in just $N \cdot f(m_1)$-time.*

Proof. For $x \in S$, let x_i, $1 \le i \le \frac{N}{m_1^2}$ be its pieces in S'. Let $x' \in S''$ be the concatenated vector $\begin{bmatrix} Ax_1 \\ Ax_2 \\ \vdots \end{bmatrix}$. Then $\|x\|^2 = \sum_i \|x_i\|^2$, $\|x'\|^2 = \sum_i \|Ax_i\|^2$, and

$$(1-\epsilon)\|x_i\|^2 \le \|Ax_i\|^2 \le (1+\epsilon)\|x_i\|^2$$

$$(1-\epsilon)\|x'\|^2 \le \|Bx'\|^2 \le (1+\epsilon)\|x'\|^2$$

Therefore we get the claimed 3ϵ-JL map as follows

$$(1-3\epsilon)\|x\|^2 \leq (1-\epsilon)^2\|x\|^2 \leq \|Bx'\|^2 \leq (1+\epsilon)^2\|x\|^2 \leq (1+3\epsilon)\|x\|^2.$$

For the time complexity, the number of required operations is

$$\frac{N}{m_1^2}(m_1^2 f(m_1)) + m_2\frac{N}{m_1} = O(Nf(m_1)).$$

The first term is $\frac{N}{m_1^2}$ applications of matrix A, and the second term is a single application of matrix B. ■

6.1 Case of Finite Points

In this section we review the the necessary theorems for JL maps for finite points via sub-Gaussian matrices and SORS matrices. For sub-Gaussian matrices, the main tools are the matrix deviation inequality and an upper bound on the Gaussian width of n points. For the SORS matrices, we don't have an equivalent matrix deviation inequality, hence we go through an indirect approach. First an RIP estimate is derived for the SORS matrices, and then one considers a sequence of random reflections by the coordinate axes, a theorem known as the Krahmer-Ward, to get a JL map.

6.1.1 JL Maps for Finite Points via sub-Gaussian Matrices

We specialize the matrix deviation inequality to a finite point set and use the Gaussian width bound for finite points.

Lemma 6.1.1. *Let x_i be n points in $\mathbb{R}^N$. Let $0 < \epsilon, \rho < 1$. Let A be a $m \times N$ sub-Gaussian random matrix with $K = \max ||A_i||_{\psi_2}$. Then there exists a universal constant C such that if*

$$m \geq C^2K^4\epsilon^{-2}\log(n)\left(1+\log(\frac{2}{\rho})\right) \tag{6.1}$$

then with probably at least $1-\rho$ the following expressions holds simultaneously for all $x \in T$.

$$(1-\epsilon)||x||^2 \leq ||\frac{1}{\sqrt{m}}Ax||^2 \leq (1+\epsilon)||x||^2. \tag{6.2}$$

Proof. By normalizing the x_i we can assume they are on the unit sphere. Next we use theorem 4.1.6 and we note that (6.2) is scaling invariant. If we let T be the set of normalize x_i, then by [35, lemma 7.5.10], we have

$$\omega(T) \leq C\sqrt{\log(n)}. \tag{6.3}$$

Then using (6.1) and adjusting the constants we get the desired bound. To get the squares in (6.2), we note that for $0 < \epsilon < 1$, $1 - 2\epsilon \leq (1-\epsilon)^2 \leq (1+\epsilon)^2 \leq 1 + 3\epsilon$.

6.1.2 JL Maps for Finite Points via SORS Matrices

We first review the Krahmer-Ward theorem that allows one to get a JL map for finite point via an RIP estimate.

Theorem 6.1.2. ***(Krahmer-Ward)*** *[18, theorem 9.36] Let x_i be n points in $\mathbb{R}^n$. Let $\rho, \epsilon \in (0,1)$. Let $A \in \mathbb{R}^{m\times N}$ be a matrix where its restricted isometry constant for sparsity 2s, i.e. δ_{2s}, satisfies $\delta_{2s} \leq \frac{\epsilon}{4}$ for $s \geq 16\log(\frac{4n}{\rho})$. Let D be a diagonal matrix with ± 1 variables on the diagonal. Then with probability exceeding $1-\rho$ the following holds simultaneously for all x_i.*

$$(1-\epsilon)||x_i||^2 \leq ||ADx_i||^2 \leq (1+\epsilon)||x_i||^2. \tag{6.4}$$

Here we present an RIP estimate for SORS matrices that we can pair with the Krahmer-Ward theorem.

Theorem 6.1.3. ***(RIP for SORS Matrices)*** *[18, theorem 12.31]Let $A \in \mathbb{C}^{m\times N}$ be a SORS matrix with constant $K \geq 1$. Let $\epsilon \in (0,1)$. There is a universal constant such that if*

$$m \geq CK^2\epsilon^{-2}s\log^4(N) \tag{6.5}$$

then with probability at least $1 - N^{-\log^3(N)}$ the restricted isometry constant of $\frac{1}{\sqrt{m}}A$ for sparsity s, δ_s, satisfies $\delta_s \leq \epsilon$.

Finally we present a finite point JL map for the SORS matrices.

Proposition 6.1.4. *Let x_i be n points in $\mathbb{C}^N$. Let $A \in \mathbb{C}^{m\times N}$ be a SORS matrix with constant K. There is a universal constant C such that if*

$$m \geq CK^2\epsilon^{-2}\log(\frac{n}{\rho})\log^4(N), \tag{6.6}$$

then with probability at least $1-\rho-N^{-\log^3(N)}$

$$(1-\epsilon)||x_i||^2 \leq ||\frac{1}{\sqrt{m}}ADx_i||^2 \leq (1+\epsilon)||x_i||^2, \tag{6.7}$$

where D is a random diagonal matrix with ±1 on its diagonal with uniform distribution.

Proof. We combine theorems 6.1.2 and 6.1.3; using the union bound, we get the claimed success probability. ■

6.1.3 Combining the Two Matrix Types

We combine the SORS and sub-Gaussian JL maps as in the scheme in lemma 6.0.1 to get our main theorem for finite points.

Theorem 6.1.5. *Let x_i be n points in $\mathbb{R}^n$. Let $\epsilon, \rho \in (0,1)$. Then with probability at least $1-\rho_1-\rho_2-N^{-\log^3(N)}$, one can find a linear map $L : \mathbb{R}^N \to \mathbb{R}^m$ such that*

$$(1-\epsilon)||x_i||^2 \leq ||Lx_i||^2 \leq (1+\epsilon)||x_i||^2 \tag{6.8}$$

$$m = O\left(\frac{\log(n)}{\epsilon^2}(1+\log(2/\rho_2))\right) \tag{6.9}$$

and a single Lx_i can be computed with time complexity $O\left(N\log\left(\epsilon^{-2}\log(\frac{nN}{p_1})\log^4(N)\right)\right)$.

Proof. We follow the in the scheme in lemma 6.0.1. Hence we have A and B such that $A \in \mathbb{C}^{m_1\times m_1^2}$ and $B \in \mathbb{C}^{m_2\times N/m_1}$. The task is to find sufficient requirements for m_1 and m_2, where m_1 will determine the runtime and m_2 will determine the final embedding dimension.

First we determine m_1. We need a $\frac{\epsilon}{3}$-JL map for $\frac{nN}{m_1^2}$ points from the SORS matrix A. For this by proposition 6.1.4, we need

$$m_1 \geq CK^2\frac{9}{\epsilon^2}\log(\frac{nN}{m_1^2\rho_1})\log^4(N). \tag{6.10}$$

Removing the reference of m_1 on the right hand side makes the required condition on m_1 stricter, hence it is sufficient to have

$$m_1 \geq CK^2 \frac{9}{\epsilon^2} \log(\frac{nN}{\rho_1}) \log^4(N). \tag{6.11}$$

The success probability is $1 - \rho_1 - N^{-\log^3(N)}$. For m_2, we need B to be a $\frac{\epsilon}{3}$-JL map of n points . Hence by lemma 6.1.1, we need

$$m_2 \geq C^2 K^4 \frac{9}{\epsilon^2} \log(n)(1 + \log(\frac{2}{\rho_2})) \tag{6.12}$$

with success probability at least $1 - \rho_2$. So combining the two steps together and specializing to DFT matrix that has a $d\log(d)$ runtime, we see that our algorithm has

$$\text{Final embedding dimension: } m_2 = O\left(\frac{\log(n)}{\epsilon^2}(1 + \log(2/\rho_2))\right) \tag{6.13}$$

$$\text{Runtime : } O(N\log(m_1)) = O\left(N\log\left(\epsilon^{-2}\log(\frac{nN}{p_1})\log^4(N)\right)\right) \tag{6.14}$$

$$\text{Success Probability: } 1 - \rho_1 - \rho_2 - N^{-\log^3(N)}. \tag{6.15}$$

This finishes the proof. ■

6.2 Upgrade for Subspaces

In this section, we repeat the process as in the previous case, but here we change our theorems so that we can get a JL embedding of a subspace. By the scaling invariance of the JL condition $(1 - \epsilon)||x|| \leq ||Lx|| \leq (1 + \epsilon)||x_i||$, we can reduce the subspace problem to its corresponding unit sphere.

We have a theorem from [29, theorem 3.3] that provides a JL map for an arbitrary set in terms of its Gaussian width. With this theorem we can get a JL embedding of a unit sphere in a subspace. This theorem uses SORS matrices.

6.2.1 SORS Matrices for Subspaces

Theorem 6.2.1. *[29, theorem 3.3] Let $T \subset \mathbb{R}^N$ and suppose $A \in \mathbb{R}^{m\times N}$ is selected from a SORS distribution with a constant K. Then*

$$\sup_{x\in T} \left| ||Ax||^2 - ||x||^2 \right| \leq \max\{\epsilon, \epsilon^2\} rad(T)^2 \tag{6.16}$$

holds with probability at least $1-\rho$ as long as

$$m \geq CK^2(1+\log(\frac{2}{\rho}))^2 \log^4(N) \frac{\max(1, \frac{\omega^2(T)}{rad(T)^2})}{\epsilon^2}. \tag{6.17}$$

■

Now we specialize to a unit sphere of a subspace using an estimate for the Gaussian width of a unit sphere.

Corollary 6.2.2. *Let T be a d dimensional affine subspace of $\mathbb{R}^N$. Assume $A \in \mathbb{R}^{m\times N}$ is selected from a SORS distribution with a constant K. Suppose $0 < \epsilon, \rho < 1$ and*

$$m \geq CK^2(1+\log(\frac{2}{\rho}))^2 \log^4(N) \frac{d}{\epsilon^2} \tag{6.18}$$

Then with probability at least $1-\rho$, the following holds for all $x \in T$

$$(1-\epsilon)||x||^2 \leq ||Ax||^2 \leq (1+\epsilon)||x||^2. \tag{6.19}$$

Proof. We use theorem 6.2.1. Since the equation (6.19) is scaling invariant, we can scale T to be a unit S^{d-1}. We have that the $\omega(S^{d-1})$, the Gaussian width of S^{d-1}, is bounded above by $\sqrt{d}+C$ for a universal constant [35, Example 7.5.7]. Using this fact and adjusting the constants the result follows. ■

6.2.2 Sub-Gaussian Matrices for Subspaces

Now we state a JL theorem for subspace embeddings using sub-Gaussian matrices. The main tool is the matrix deviation inequality [35, Chapter 9].

Corollary 6.2.3. *Let T be a d-dimensional affine subspace of $\mathbb{R}^N$. Let A be a $m \times N$ sub-Gaussian random matrix with $K = \max ||A_i||_{\psi_2}$. Suppose*

$$m \geq CK^4(1 + \log(\frac{2}{\rho}))^2 \frac{d}{\epsilon^2}. \tag{6.20}$$

Then with probability at least $1 - \rho$, the following holds for all $x \in T$

$$(1 - \epsilon)||x||^2 \leq ||Ax||^2 \leq (1 + \epsilon)||x||^2. \tag{6.21}$$

Proof. We use theorem 4.1.6. Again since (6.21) is scaling invariant we can reduce to the unit sphere. Since $\omega(S^{d-1}) \leq \sqrt{d} + C$ for a universal constant, we get

$$m \geq \frac{C^2K^4\left(\sqrt{d} + C + \sqrt{\log(\frac{2}{\rho})}\right)^2}{\epsilon^2} \tag{6.22}$$

adjusting the constants and using $3(a^2 + b^2 + c^2) \geq (a + b + c)^2$ gives the desired result. ∎

6.2.3 Mixing the Matrix Types

We use the scheme in lemma 6.0.1 to combine the SORS and Gaussian matrices. We break the vectors into smaller pieces, apply a fast JL map, recombine the pieces and then apply a sub-Gaussian JL map.

To the prove the desired properties of our algorithm, we also need the to review the proposition 5.0.2 that lets us control the Gaussian width of a set after a JL map is applied to it.

Proposition 6.2.4. *Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map, $T \subset \mathbb{R}^n$ with a and b positive reals such that for $x, y \in T$, $a||x - y|| \leq ||L(x - y)|| \leq b||x - y||$. Then $a\,\omega(T) \leq \omega(L(T)) \leq b\,\omega(T)$.*

Now we are ready to present the proof of algorithm in the case of subspaces. The proof strategy first uses the properties of SORS matrices for the pieces of an input vector, then we use proposition 6.2.4 to control the Gaussian width of the image of the set after the SORS matrix has been applied. This in turn allows us to apply the matrix deviation inequality as the final step.

Theorem 6.2.5. *Let T be a d-dimensional affine subspace of $\mathbb{R}^N$. Let $0 < \epsilon, \rho_1, \rho_2 < 1$ such that $\rho_1 + \rho_2 \leq 1$ and $K \geq 1$. Let m_1 and m_2 be integers such that*

$$m_1 \geq CK^2(1 + \log(\frac{2N}{\rho_1 m_1^2}))^2 \log^4(N)\frac{d}{\epsilon^2} \tag{6.23}$$

$$m_2 \geq \frac{CK^4(1+\epsilon)^2 d(1 + \log(\frac{2}{p_2}))}{\epsilon^2} \tag{6.24}$$

Then with probability at least $1 - \rho_1 - \rho_2$ one can find a linear map $L : \mathbb{R}^N \to \mathbb{R}^{m_2}$ such that for all $x, y \in T$

$$(1 - \epsilon)||x - y||^2 \leq ||L(x - y)||^2 \leq (1 + \epsilon)||x - y||^2 \tag{6.25}$$

Furthermore, for $x \in T$, the Lx can be computed with a $O\left(N \log\left(K^2(1 + \log(\frac{2N}{\rho_1 m_1^2}))^2 \log^4(N)\frac{d}{\epsilon^2}\right)\right)$ run-time.

Proof. Following the scheme in lemma 6.0.1, we have matrices A and B such that $\mathbf{A} \in \mathbb{C}^{m_1 \times m_1^2}$ and $\mathbf{B} \in \mathbb{C}^{m_2 \times N/m_1}$. Our goal is to find sufficient conditions on m_1 and m_2 to satisfy the required JL properties.

The projection of T onto a m_1^2 dimensional subspace is another subspace of dimension at most d. The Gaussian width of the unit sphere in this subspace is bounded above by $\sqrt{d} + C$ for a universal constant C. Now we apply corollary 6.2.2, for $\frac{\epsilon}{3}$, $\frac{N}{m_1^2}$ different times since there are that many different projections. Applying each with a success probability $1 - \frac{\rho_1 m_1^2}{N}$, and combining the failure probabilities using the union bound we get the success probability of $1 - \rho_1$. Therefore we need

$$m_1 \geq CK^2(1 + \log(\frac{2N}{\rho_1 m_1^2}))^2 \log^4(N)\frac{d}{\epsilon^2}. \tag{6.26}$$

For a universal constant C, where K is the constant associated to the SORS matrix A.

Now we consider the block diagonal matrix $C = \text{diag}(A)$ where $C \in \mathbb{R}^{\frac{N}{m_1} \times N}$. For any $x \in T$, since matrix A distorts each individual projection of x at must by $1 \pm \epsilon$, then same is true for x. Now we apply proposition 6.2.4 to the unit sphere of T, call it $U(T)$. Hence we can estimate the Gaussian width of the image of $U(T)$ from above as $\omega(C(U(T))) \leq (1 + \epsilon)(\sqrt{d} + C')$.

Now we can apply the matrix deviation inequality, theorem 4.1.5. With probability at least $1 - p_2$, we have

$$\sup_{x \in C(U(T))} \left| \|Bx\|_2 - \sqrt{m_2}\|x\|_2 \right| \leq C'K^2 \left[w(C(U(T))) + \sqrt{\log(2/p_2)} \cdot \sup_{x \in C(U(S))} \|x\|_2 \right], \tag{6.27}$$

$$\sup_{x \in C(U(T))} \left| \|Bx\|_2 - \sqrt{m_2}\|x\|_2 \right| \leq C'K^2 \left[(1+\epsilon)(\sqrt{d} + C') + \sqrt{\log(2/p_2)} \cdot (1+\epsilon) \right]. \tag{6.28}$$

Rewriting, we get that for $x \in C(U(T))$

$$\begin{aligned} &\|x\|_2 - \frac{C'K^2(1+\epsilon)\left[(\sqrt{d}+C') + \sqrt{\log(2/p_2)}\right]}{\sqrt{m_2}} \\ &\leq \|\frac{1}{\sqrt{m_2}}Bx\|_2 \\ &\leq \|x\|_2 + \frac{C'K^2(1+\epsilon)\left[(\sqrt{d}+C') + \sqrt{\log(2/p_2)}\right]}{\sqrt{m_2}}. \end{aligned}$$

Now Since C is an ϵ-JL map of a $U(T)$ we have that

$$1 - \epsilon \leq \|x\|_2 \leq 1 + \epsilon.$$

Therefore we get

$$\begin{aligned} 1 - \epsilon - \frac{C'K^2(1+\epsilon)\left[(\sqrt{d}+C') + \sqrt{\log(2/p_2)}\right]}{\sqrt{m_2}} &\leq \|\frac{1}{\sqrt{m_2}}Bx\|_2 \\ &\leq 1 + \epsilon + \frac{C'K^2(1+\epsilon)\left[(\sqrt{d}+C') + \sqrt{\log(2/p_2)}\right]}{\sqrt{m_2}}. \end{aligned}$$

Finally as long as

$$m_2 \geq \frac{C'K^4(1+\epsilon)^2 d(1 + \log(\frac{2}{p_2}))}{\epsilon^2} \tag{6.29}$$

we get

$$1 - 2\epsilon \leq \|\frac{1}{\sqrt{m_2}}Bx\|_2 \leq 1 + 2\epsilon$$

Adjusting the constants we get the desired bound. To summarize we have 3 estimates.

$$\text{Final embedding dimension: } m_2 = \mathcal{O}\left(\frac{K^4(1+\epsilon)^2 d(1+\log(\frac{2}{p_2}))}{\epsilon^2}\right). \tag{6.30}$$

$$\text{Runtime : } \mathcal{O}(N\log(m_1)) = \mathcal{O}\left(N\log\left(K^2(1+\log(\frac{2N}{\rho m_1^2}))^2\log^4(N)\frac{d}{\epsilon^2}\right)\right). \tag{6.31}$$

$$\text{Success Probability: } 1-\rho_1-\rho_2. \tag{6.32}$$

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] Eddie Aamari, Jisu Kim, Frédéric Chazal, Bertrand Michel, Alessandro Rinaldo, and Larry Wasserman. Estimating the reach of a manifold. *Electronic Journal of Statistics*, 13(1):1359–1399, 2019. 12, 13

[2] Nir Ailon and Edo Liberty. Fast Dimension Reduction Using Rademacher Series on Dual BCH Codes. *Discrete & Computational Geometry*, 42(4):615–630, December 2009. 69

[3] Ralph Alexander and Stephanie Alexander. Geodesics in Riemannian manifolds-with-boundary. *Indiana University Mathematics Journal*, 30(4):481–488, 1981. Publisher: JSTOR. 15

[4] Stephanie B. Alexander, I. David Berg, and Richard L. Bishop. The Riemannian obstacle problem. *Illinois Journal of Mathematics*, 31(1):167–184, March 1987. 16

[5] Noga Alon. Problems and results in extremal combinatorics—i. *Discrete Mathematics*, 273(1-3):31–53, 2003. 62, 68

[6] Marcel Berger. *A Panoramic View of Riemannian Geometry*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2003. 16, 38, 39

[7] R.L. Bishop and R.J. Crittenden. *Geometry of manifolds*. ISSN. Elsevier Science, 2011. 16

[8] Jean-Daniel Boissonnat, André Lieutier, and Mathijs Wintraecken. The reach, metric distortion, geodesic convexity and the variation of tangent spaces. *Journal of Applied and Computational Topology*, 3(1-2):29–58, June 2019. 12, 15, 17

[9] Károly Böröczky, Jr. *Finite Packing and Covering*. Cambridge Tracts in Mathematics. Cambridge University Press, 2004. 19

[10] Simone Brugiapaglia, Sjoerd Dirksen, Hans Christian Jung, and Holger Rauhut. Sparse recovery in bounded riesz systems with applications to numerical methods for pdes. *arXiv preprint arXiv:2005.06994*, 2020. 9, 53, 54, 55

[11] E.J. Candes, J. Romberg, and T. Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(2):489–509, 2006. 7

[12] E.J. Candes and T. Tao. Decoding by linear programming. *IEEE Transactions on Information Theory*, 51(12):4203–4215, 2005. 7

[13] Emmanuel J. Candes and Terence Tao. Near-optimal signal recovery from random projections: Universal encoding strategies? *IEEE Transactions on Information Theory*, 52(12):5406–5425, 2006. 7

[14] Kenneth L. Clarkson. Tighter bounds for random projections of manifolds. In *Proceedings of the twenty-fourth annual symposium on Computational geometry - SCG '08*, page 39, College Park, MD, USA, 2008. ACM Press. 28

[15] Armin Eftekhari and Michael B. Wakin. New analysis of manifold embeddings and signal recovery from compressive measurements. *Applied and Computational Harmonic Analysis*, 39(1):67–109, July 2015. 12, 16, 28, 29, 49, 52

[16] Herbert Federer. Curvature measures. *Transactions of the American Mathematical Society*, 93(3):418–418, March 1959. 11, 12, 13, 15

[17] Werner Fenchel. On the differential geometry of closed space curves. *Bulletin of the American Mathematical Society*, 57(1. P1):44–54, 1951. 38, 39

[18] Simon Foucart and Holger Rauhut. *A Mathematical Introduction to Compressive Sensing*. Applied and Numerical Harmonic Analysis. Springer New York, New York, NY, 2013. 6, 9, 33, 53, 54, 72

[19] Sylvestre Gallot, Dominique Hulin, and Jacques Lafontaine. *Riemannian geometry*, volume 2. Springer, 1990. 18, 26

[20] Walter Gautschi. Some elementary inequalities relating to the gamma and incomplete gamma function. *J. Math. Phys*, 38(1):77–81, 1959. 20

[21] Mohammad Ghomi and Serge Tabachnikov. Totally skew embeddings of manifolds. *Mathematische Zeitschrift*, 258(3):499–512, 2008. 31

[22] Alfred Gray. *Tubes*, volume 221. Birkhäuser, 2012. 12, 13, 23, 39

[23] Subhaneil Lahiri, Peiran Gao, and Surya Ganguli. Random projections of random manifolds. *arXiv preprint arXiv:1607.04331*, 2016. 28

[24] R. Lashof and S. Smale. On the Immersion of Manifolds in Euclidean Space. *The Annals of Mathematics*, 68(3):562, November 1958. 28

[25] S. Li. Concise Formulas for the Area and Volume of a Hyperspherical Cap. *Asian Journal of Mathematics & Statistics*, 4(1):66–70, December 2010. 19

[26] Martin Lotz. On the volume of tubular neighborhoods of real algebraic varieties. *Proceedings of the American Mathematical Society*, 143(5):1875–1889, December 2014. 23

[27] 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, March 2008. 15, 16, 37

[28] Barrett O'Neill. *Semi-Riemannian geometry: with applications to relativity*. Number 103 in Pure and applied mathematics. Academic Press, New York, 1983. 15

[29] Samet Oymak, Benjamin Recht, and Mahdi Soltanolkotabi. Isometric sketching of any set via the Restricted Isometry Property. *Information and Inference: A Journal of the IMA*, 7(4):707–726, December 2018. 7, 9, 53, 54, 60, 74, 75

[30] William F. Pohl. Some Integral Formulas for Space Curves and Their Generalization. *American Journal of Mathematics*, 90(4):1321, October 1968. 28

[31] Takashi Sakai. Riemannian geometry. translations of mathematical monographs, vol. 149. *American Mathematical Society, Providence, RI*, pages 262–272, 1996. 15

[32] Shigeo Sasaki. On the differential geometry of tangent bundles of riemannian manifolds. *Tohoku Mathematical Journal, Second Series*, 10(3):338–354, 1958. 23

[33] Hukum Singh. On the cut locus and the focal locus of a submanifold in a riemannian manifold. ii. *Publ. Inst. Math.(Beograd)(NS)*, 41(55):119–124, 1987. 15

[34] Christoph Thäle. 50 years sets with positive reach–a survey. *Surveys in Mathematics and its Applications*, 3:123–165, 2008. Publisher: University Constantin Brancusi. 12

[35] Roman Vershynin. *High-dimensional probability: an introduction with applications in data science*. Number 47 in Cambridge series in statistical and probabilistic mathematics. Cambridge University Press, Cambridge ; New York, NY, 2018. v, 1, 2, 3, 4, 5, 6, 17, 18, 19, 30, 49, 50, 51, 62, 63, 64, 68, 72, 75

[36] Hermann Weyl. On the Volume of Tubes. *American Journal of Mathematics*, 61(2):461, April 1939. 23

[37] James H. White. Self-Linking and the Gauss Integral in Higher Dimensions. *American Journal of Mathematics*, 91(3):693, July 1969. 28

[38] James H. White. Self-linking and the directed secant span of a differentiable manifold. *Journal of Differential Geometry*, 5(3-4):357–369, 1971. 28