Stable Manifold Embeddings with Operators Satisfying the Restricted Isometry Property

Han Lun Yap,* Michael B. Wakin, and Christopher J. Rozell

***Abstract*—Signals of interests can often be thought to come from a low dimensional signal model. The exploitation of this fact has led to many recent interesting advances in signal processing, one notable example being in the field of compressive sensing (CS). The literature on CS has established that many matrices satisfy the Restricted Isometry Property (RIP), which guarantees a stable (i.e., distance-preserving) embedding of a sparse signal model from an undersampled linear measurement system. In this work, we study the stable embedding of manifold signal models using matrices that satisfy the RIP. We show that by paying reasonable additional factors in the number of measurements, all matrices that satisfy the RIP can also be used (in conjunction with a random sign sequence) to obtain a stable embedding of a manifold.**

***Index Terms*—Manifold Embedding, Dimensionality Reduction, Restricted Isometry Property**

I. Introduction

A significant amount of modern work in signal processing rests on the observation that many high-dimensional signals in fact have an intrinsic low-dimensional structure. Models for characterizing such structure can often be interpreted geometrically; sparse signals, for example, are constrained to live near a union of low-dimensional subspaces within the ambient high-dimensional signal space [1], while parametric signals and certain non-parametric signal collections are constrained to live near low-dimensional manifolds [2], [3]. Such low-dimensional geometric structures can be useful for reducing the burden of acquiring, storing, and processing high-dimensional signals thanks to the fact that these structures have been shown to support stable, distance-preserving embeddings via linear mappings into a lower-dimensional space.

In the field of compressive sensing (CS), for example, it is known that random matrices populated with independent and identically distributed (i.i.d.) Gaussian or subgaussian entries will, with high probability, satisfy a condition known as the Restricted Isometry Property (RIP) [4]. The RIP guarantees that a matrix will approximately preserve distances between all pairs of sparse signals. These results have been previously extended to show that an undersampled random orthoprojector can also provide a stable embedding of signals living along a low-dimensional manifold [5], [6]. Such stable embeddings are valuable because they ensure that key properties of the signal model are retained in the low-dimensional measurement space, and this often leads to guarantees on our ability to recover the original signal or to perform processing in the measurement space. In the particular case of a manifold signal model, a stable embedding guarantees that manifold learning algorithms (e.g., Isomap) can be run in the low-dimensional measurement space nearly as accurately as (and much more efficiently than) in the original signal space [7].

Recently the CS community has turned to investigating structured measurement systems because unstructured systems (e.g., those corresponding to i.i.d. random matrices or random orthoprojectors) may be impractical due to memory constraints, computational costs, or prohibitions in the data acquisition architecture. Several structured CS systems, such as random convolution systems (described by partial Toeplitz [8] and circulant matrices [9]) and deterministic matrix constructions [10], have emerged that satisfy the RIP while requiring only a small increase in the number of measurements beyond what is needed for an unstructured random matrix. At first glance, it may appear that these structured embedding results are limited to sparsity-based signal models and cannot be generalized to models such as manifolds.

In this paper, we show that by paying reasonable factors in the number of measurements, all measurement systems that satisfy the RIP for sparse signals can also be used (in conjunction with a random sign sequence) to obtain a stable embedding of a manifold. Therefore, the advantages afforded by structured random matrices over unstructured ones can be carried over to measurement systems for manifold modeled signals. Our work rests

*Corresponding Author. HLY and CJR are with the School of Electrical and Computer Engineering at the Georgia Institute of Technology. MBW is with the Division of Engineering at the Colorado School of Mines. This work was partially supported by NSF grants CCF-0830456 and CCF-0830320 and by DSO National Laboratories of Singapore. The authors are grateful to Armin Eftekhari for valuable discussions about this work.

on a recent result [11] showing that when the columns of an RIP matrix are modulated by a random sign sequence, the matrix will obey a form of the Johnson-Lindenstrauss (JL) lemma and can therefore provide a stable embedding of an arbitrary finite point cloud. Following similar arguments to [5], we extend the finite JL result to all points living along a manifold. After covering the necessary background in Section II, we state our main result and its implications in Section III. Section IV contains the supporting proofs, while Section V concludes.

II. Background

A. RIP and Stable Embeddings

We formally state the RIP in the following definition.

Definition II.1. *A linear operator $\Phi : \mathbb{R}^N \to \mathbb{R}^M$ satisfies the* Restricted Isometry Property *of* order S *and* conditioning δ *(or RIP-(S,δ) in short) if for all $x \in \mathbb{R}^N$ with at most S nonzero entries, we have*

$$(1-\delta)\|x\|_2^2 \le \|\Phi x\|_2^2 \le (1+\delta)\|x\|_2^2.$$

For more general signal families, we make the following definition.

Definition II.2. *A linear operator $\widehat{\Phi}$ provides a* stable embedding *of a subset $\mathcal{M} \subset \mathbb{R}^N$ with conditioning ϵ if for all pairs $x_1, x_2 \in \mathcal{M}$, we have*

$$(1-\epsilon) \le \frac{\|\widehat{\Phi}x_1 - \widehat{\Phi}x_2\|_2}{\|x_1 - x_2\|_2} \le (1+\epsilon). \tag{1}$$

Alternatively, one may define

$$U(\mathcal{M}) = \left\{ \frac{x-y}{\|x-y\|_2} \mid x, y \in \mathcal{M},\ x \ne y \right\}, \tag{2}$$

which is the set of all directions (normalized difference vectors) between pairs of points in $\mathcal{M}$. Then $\widehat{\Phi}$ is a stable embedding of $\mathcal{M}$ with conditioning ϵ if and only if

$$\sup_{x \in U(\mathcal{M})} \left| \|\widehat{\Phi}x\|_2 - 1 \right| \le \epsilon.$$

A well known interpretation of the RIP is as a stable embedding of sparse vectors. In particular, an operator satisfying RIP-$(2S,\delta)$ provides a stable embedding with conditioning δ of the union of all S-sparse subspaces.

B. Manifold Characteristics

We present here some relevant terminology used for discussing manifolds. For a Riemannian submanifold $\mathcal{M} \subset \mathbb{R}^N$, we let $d_{\mathcal{M}}(x,y)$ denote the geodesic distance between two points $x, y \in \mathcal{M}$; in other words, $d_{\mathcal{M}}(x,y)$ denotes the length of the shortest path between x and y along on the manifold. Also, for a given point $x \in \mathcal{M}$, we let Tan_x denote the tangent space of $\mathcal{M}$ at x. For a manifold of dimension K, Tan_x can be thought of as a K-dimensional linear subspace of $\mathbb{R}^N$.

We also describe two quantities that are used for characterizing the properties of manifolds. The first is the condition number, $\frac{1}{\tau}$, first described in [12] and later used in [5]. The condition number provides a bound on the worse case curvature of any unit speed geodesic path along the manifold. More explicitly, for larger values of τ, it is guaranteed that for two nearby points $x, y \in \mathcal{M}$ the angle between Tan_x and Tan_y will be small. A large value for τ also ensures that two points $x, y \in \mathcal{M}$ with large geodesic distance $d_{\mathcal{M}}(x,y)$ cannot be arbitrarily close in Euclidean distance $\|x-y\|_2$. The interested reader is referred to [12], [5] for an explicit definition of the condition number and for formal descriptions of its implications.

A second useful quantity concerns the covering regularity of a manifold. Before describing this quantity, we require the following definition.

Definition II.3. *Let $\mathcal{M}$ be a K-dimensional Riemannian submanifold of $\mathbb{R}^N$. The* geodesic covering number *of $\mathcal{M}$ at resolution $T > 0$, denoted by $G(T)$, is the smallest number such that there exists a finite set $A \subset \mathcal{M}$ with $|A| = G(T)$ such that for all $x \in \mathcal{M}$,*

$$\min_{a \in A} d_{\mathcal{M}}(x,a) \le T.$$

Any set A that satisfies the above condition is called a geodesic covering set of resolution T *and in particular, any such set A with $|A| = G(T)$ is called a* minimal geodesic covering set of resolution T.

Now, the following definition of covering regularity allows us to enumerate minimal geodesic covering sets of a manifold. This will be useful in our analysis when we approximate a manifold with its covering set.

Definition II.4. *Let $\mathcal{M}$ be a K-dimensional Riemannian submanifold of $\mathbb{R}^N$ having volume V. We say that $\mathcal{M}$ has* geodesic covering regularity R *for resolutions $T \le T_0$ if*

$$G(T) \le \frac{R^K V K^{K/2}}{T^K}$$

holds for all $T \le T_0$.

As in [5], we shall neglect the minor dependence of the geodesic covering regularity R on the maximum resolution T_0 and assume that the formula for $G(T)$ works for the range of T considered in this paper.

C. Related Work

The work in this paper is most closely related to [5] and [6], which both showed that a random orthogonal projection from $\mathbb{R}^N$ into $\mathbb{R}^M$ will, with high probability, provide a stable embedding of a K-dimensional manifold $\mathcal{M} \subset \mathbb{R}^N$. In each of these works, the requisite number of measurements scales only linearly in the manifold dimension K and logarithmically in certain other parameters of the manifold. In [5] as well as in this work, there is an additional logarithmic dependence on the ambient dimension N.

The proof of each of these results requires consideration of a finite covering of points carefully chosen from the manifold. Using the JL lemma [13], it then can be argued that, with high probability, a random orthogonal projection will provide a stable embedding of these points. Following this, various geometric arguments allow one to conclude that the same orthogonal projection will provide a (slightly weaker) stable embedding of the entire manifold $\mathcal{M}$.

Formally, the JL lemma guarantees that for any finite set $E \subset \mathbb{R}^N$, a random orthogonal projector $\widehat{\Phi} \in \mathbb{R}^{M \times N}$ will provide a stable embedding of E with conditioning ϵ with high probability if $M = O(\log(|E|) \cdot \epsilon^{-2})$. Recent research has uncovered several other techniques for constructing $\widehat{\Phi}$ that also satisfy the JL lemma. In particular, it has recently been shown [11] that any $M \times N$ matrix Φ that satisfies the RIP can also be used in the JL lemma if its columns are randomized by a length-N sequence of ± 1's. To state this result formally, we make the following definition.

Definition II.5. *A* diagonal Rademacher matrix *is an $N \times N$ diagonal matrix populated with a length-N sequence of independent ± 1-Bernoulli random variables on the main diagonal and with zeros elsewhere.*

Multiplying any fixed matrix on the right by a diagonal Rademacher matrix will therefore randomize its column signs. The following theorem is a slightly modified version of Theorem 3.1 from [11].

Theorem II.1. [11] *Let $0 < \rho < 1$ and $0 < \epsilon' < 1$, and let E be a finite set of points in $\mathbb{R}^N$. Suppose $\Phi \in \mathbb{R}^{M \times N}$ satisfies RIP-($S, \frac{\epsilon'}{4}$) with*

$$S \geq 40 \log \left(\frac{4|E|}{\rho} \right),$$

and let $\widehat{\Phi} = \Phi D$ where D is a diagonal Rademacher matrix. Then with probability exceeding $1 - \rho$, $\widehat{\Phi}$ will provide a stable embedding of E with conditioning ϵ'.

It is this result that we leverage to expand the range of compressive linear operators that can provide a stable manifold embedding.

This work is also related to current research on extending the JL lemma to data sets that have low "intrinsic dimension" [14], [15], [16]. In this line of research, the open question is whether a finite data set $B \in \mathbb{R}^N$ that has low *doubling dimension*[1] $D < \log(|B|)$ can be stably embedded into a space whose dimension scales with D but not $\log(|B|)$ as in the JL lemma. Our work can be considered a small step in this direction as it allows consideration of the special case where B is a Riemannian submanifold.

III. Main Result

The main result of this paper is captured in the following theorem.

Theorem III.1. *Let $\mathcal{M}$ be a compact K-dimensional Riemannian submanifold of $\mathbb{R}^N$ having condition number $\frac{1}{\tau}$, volume V, and geodesic covering regularity R. Fix a failure probability $0 < \rho < 1$ and a conditioning $0 < \epsilon < 1$. Suppose $\Phi \in \mathbb{R}^{M \times N}$ satisfies RIP-($S, \frac{\epsilon}{64}$) with*

$$S \geq O \left(K \log \left(\frac{RVN}{\tau \epsilon} \right) + \log \left(\frac{1}{\rho} \right) \right).$$

Let D be a diagonal Rademacher matrix, and define $\widehat{\Phi} = \Phi D$. Then with probability greater than $1 - \rho$, $\widehat{\Phi}$ provides a stable embedding of $\mathcal{M}$ with conditioning ϵ.

The proof of this theorem can be found in Section IV. The theorem essentially says that stable manifold embedding can be obtained, with high probability, from any RIP matrix Φ, provided that the RIP order is proportional to the dimension of the manifold, and provided that the RIP conditioning is sufficiently tight. Therefore, from any of the known constructions of RIP matrices with a number of rows proportional to the RIP order, we can obtain a stable manifold embedding with a number of measurements proportional to the manifold dimension. We illustrate the implications of our result with a few notable examples below.

In the corollaries that follow, we assume that $\mathcal{M}$ is a compact K-dimensional Riemannian submanifold of $\mathbb{R}^N$ having condition number $\frac{1}{\tau}$, volume V, and geodesic covering regularity R. We also assume a fixed failure probability $0 < \rho < 1$ and conditioning $0 < \epsilon < 1$.

Corollary III.1 (Subgaussian matrices)**.** *Suppose $\Phi \in \mathbb{R}^{M \times N}$ is a random matrix populated with i.i.d. subgaus-*

[1]The *doubling constant* λ_B of a set B is the smallest number $\lambda \geq 1$ such that every ball in B can be covered by at most λ balls of half its radius. The *doubling dimension* of B is $D = \log_2(\lambda_B)$.

sian random variables having zero mean and variance $\frac{1}{M}$. If

$$M \geq O\left(\frac{1}{\epsilon^2}\left(K\log\left(\frac{RVN}{\tau\epsilon}\right)+\log\left(\frac{1}{\rho}\right)\right)\log\left(\frac{N}{K}\right)\right)$$

then with probability greater than $1-(e^{-O(M)}+\rho)$, Φ provides a stable embedding of $\mathcal{M}$ with conditioning ϵ.

The proof of this corollary follows from the fact [1] that subgaussian random matrices satisfy RIP-(S,δ) with high probability whenever $M \geq O\left(\frac{S}{\delta^2}\log\left(\frac{N}{S}\right)\right)$ and the fact that, for a diagonal Rademacher matrix D and i.i.d. subgaussian matrix Φ, both Φ and ΦD have the same distribution. This corollary formally proves a remark made briefly in [5]—that stable manifold embeddings can arise not only from random orthogonal matrices but also from random subgaussian matrices.

Corollary III.2 (Partial circulant matrices)**.** *Suppose $\Phi \in \mathbb{R}^{M\times N}$ is a partial circulant matrix with Rademacher entries and arbitrarily selected rows (see [9] for a detailed construction). If*

$$M \geq O\left(\frac{\log\frac{1}{\rho}}{\epsilon}\left(K\log\left(\frac{RVN}{\tau\epsilon}\right)+\log\left(\frac{1}{\rho}\right)\right)^{\frac{3}{2}}\log^{\frac{3}{2}}(N)\right)$$

and D is a diagonal Rademacher matrix, then with probability greater than $1-O(\rho)$, $\widehat{\Phi}=\Phi D$ is a stable embedding of $\mathcal{M}$ with conditioning ϵ.

Here, the proof follows from the fact [9] that partial circulant matrices satisfy RIP-(S,δ) with probability greater than $1-\rho$ whenever $M \geq O\left(\frac{S^{3/2}}{\delta}\log^{3/2}(N)\log\left(\frac{1}{\rho}\right)\right)$. This result makes possible one efficient implementation of a dimensionality reduction scheme for manifold-modeled data. One would first preprocess the ambient data in $\mathbb{R}^N$ by multiplying the entries with a Rademacher sequence (that is fixed beforehand). Then one would simply convolve the processed data with a random sequence (made up of Gaussian random variables or a separate Rademacher sequence) and arbitrarily select M samples of the convolution output. One can also imagine using other efficient dimensionality reduction schemes that require only $M = O\left(K\log^p(N)\right)$ measurements for some integer p, for example by taking Φ to be a subsampled Fourier matrix [17].

Corollary III.3 (Deterministic binary matrices)**.** *Suppose $\Phi \in \{0,1\}^{M\times N}$ is a deterministic matrix following the construction given in [10]. If*

$$M \geq O\left(\frac{1}{\epsilon^2}\left(K\log\left(\frac{RVN}{\tau\epsilon}\right)+\log\left(\frac{1}{\rho}\right)\right)^2\log^2(N)\right)$$

and D is a diagonal Rademacher matrix, then with probability greater than $1-\rho$, $\widehat{\Phi}=\Phi D$ is a stable embedding of $\mathcal{M}$ with conditioning ϵ.

Again, this corollary follows from the fact [10] that such matrices satisfy RIP-(S,δ) whenever $M \geq O\left(\frac{S^2}{\delta^2}\log^2(N)\right)$. Despite the additional number of required measurements, deterministic matrices are of interest to the CS community because there currently do not exist any tractable algorithms for verifying whether a randomly constructed matrix satisfies the RIP.

IV. Proof of Main Result

Our proof of Theorem III.1 follows the basic structure of the proof of Theorem 3.1 in [5]. We also borrow a bit of terminology from [18], which helps delineate the geometric and probabilistic aspects of the problem.

To be specific, in order to prove Theorem III.1, our goal is to show that $\sup_{x\in U(\mathcal{M})}\left|\|\widehat{\Phi}x\|_2-1\right| \leq \epsilon$ holds with probability at least $1-\rho$. To reduce the complexity of bounding the the random process $\left|\|\widehat{\Phi}x\|_2-1\right|$ over the infinite set of points $U(\mathcal{M})$, we explain that it is in some sense sufficient to consider $\left|\|\widehat{\Phi}x\|_2-1\right|$ only over some finite sampling $U(B)\subset U(\mathcal{M})$ which we refer to as a *covering set*. Finally, we apply the RIP/JL result from [11] to guarantee that $\max_{x\in U(B)}\left|\|\widehat{\Phi}x\|_2-1\right|$ is small with high probability.

A. Covering Set

The covering set $U(B)\subset U(\mathcal{M})$ is defined in terms of a set of points $B\subset\mathcal{M}$, which we refer to as *anchor points*. We define the set B according to the following steps. Let $T>0$, $\gamma>0$ denote some variables whose values will be fixed later in Lemma IV.1. Then let A be any minimal geodesic covering set of $\mathcal{M}$ with resolution T. According to Definitions II.3 and II.4, this means that for any point $x\in\mathcal{M}$, $\min_{a\in A} d_{\mathcal{M}}(x,a)\leq T$ and that

$$|A| = G(T) \leq \frac{R^K V K^{K/2}}{T^K}.$$

Next, for every $a\in A$, define $Q_1(a)$ to be a minimal γ-net of the unit sphere of Tan_a in the sense that for every $u\in\mathrm{Tan}_a$ with $\|u\|_2=1$, we have

$$\min_{q\in Q_1(a)}\|u-q\|_2\leq\gamma.$$

From standard volumetric estimates (see for example [18]), this can be accomplished with

$$|Q_1(a)| \leq \left(1+\frac{2}{\gamma}\right)^K \leq \left(\frac{3}{\gamma}\right)^K.$$

Note that since every $q \in Q_1(a)$ belongs in the unit sphere, $\|q\|_2 = 1$. We also define a renormalized set

$$Q_2(a) = \{Tq \ : \ q \in Q_1(a)\},$$

so that $\|q\|_2 = T$ for all $q \in Q_2(a)$ and so that for every $u \in \mathrm{Tan}_a$ such that $\|u\|_2 = T$, we have

$$\min_{q \in Q_2(a)} \|u - q\|_2 \le T\gamma.$$

Since $Q_2(a)$ is merely a renormalization of $Q_1(a)$, $|Q_2(a)| = |Q_1(a)|$. Finally, the set of anchor points of $\mathcal{M}$ is defined as

$$B = \bigcup_{a \in A} \{a\} \cup (a + Q_2(a)),$$

where $a + Q_2(a)$ is the set of tangents anchored at the point a (instead of the origin).

B. Embedding Guarantee

Returning to the goal of our proof, we wish to show that for some $0 < \epsilon < 1$,

$$\sup_{x \in U(\mathcal{M})} \left| \|\widehat{\Phi} x\|_2 - 1 \right| \le \epsilon. \tag{3}$$

The following lemma, which collects and modifies several results from [5], relates the supremum of our random process over $U(\mathcal{M})$ to its maximum over $U(B)$.

Lemma IV.1. *Let $\mathcal{M}$ and ϵ be as in the statement of Theorem III.1. Suppose $\Phi \in \mathbb{R}^{M \times N}$ satisfies RIP-($S, \frac{\epsilon}{4}$) for any $S \le N$ and let $\widehat{\Phi} = \Phi D$ where D is a diagonal Rademacher matrix. Let $B \subset \mathcal{M}$ be a set of anchor points as defined above, specifically where $T = \frac{C_1 \tau \epsilon^2}{N}$, $\gamma = \frac{C_2 \epsilon}{\sqrt{N}}$, $C_1 \le \frac{1}{308}$, and $C_2 \le \frac{1}{3123}$. Then*

$$\sup_{x \in U(\mathcal{M})} \left| \|\widehat{\Phi} x\|_2 - 1 \right| \le 16 \max_{x \in U(B)} \left| \|\widehat{\Phi} x\|_2 - 1 \right|. \tag{4}$$

Proof: See the Appendix for a proof sketch, which closely follows the results in [5]. ∎

Therefore, if the assumptions of Lemma IV.1 are met, then (3) will be attained if we can prove that

$$\max_{x \in U(B)} \left| \|\widehat{\Phi} x\|_2 - 1 \right| \le \frac{\epsilon}{16}. \tag{5}$$

To ensure that (5) is true with high probability, we apply Theorem II.1 with $E = U(B)$ and with conditioning $\epsilon' = \frac{\epsilon}{16}$ to conclude that, if Φ satisfies RIP-($S, \frac{\epsilon}{64}$) with

$$S \ge 40 \log \left(\frac{4|U(B)|}{\rho} \right), \tag{6}$$

then with probability exceeding $1 - \rho$, $\widehat{\Phi}$ will provide a stable embedding of $U(B)$ with conditioning $\frac{\epsilon}{16}$. We note that if Φ satisfies RIP-($S, \frac{\epsilon}{64}$) as required here, it also satisfies RIP-($S, \frac{\epsilon}{4}$) as required in Lemma IV.1.

For the cardinality of $|U(B)|$, we note that $|U(B)| \le |B|^2$ and that

$$\begin{aligned} |B| &\le \sum_{a \in A} (1 + |Q_2(a)|) = \sum_{a \in A} (1 + |Q_1(a)|) \\ &\le \left(\frac{R^K V K^{K/2}}{T^K} \right) \left(\frac{3}{\gamma} \right)^K \\ &\le \left(\frac{3 \cdot 2^{1/K} R V^{1/K} \sqrt{K} N^{3/2}}{C_1 C_2 \tau (\epsilon')^3} \right)^K. \end{aligned}$$

Therefore,

$$\begin{aligned} \log \left(\frac{4|B|^2}{\rho} \right) &= 2 \log(|B|) + \log(4/\rho) \\ &\le 2K \log \left(\frac{3 \cdot 2^{1/K} R V^{1/K} \sqrt{K} N^{3/2}}{C_1 C_2 \tau (\epsilon')^3} \right) \\ &\quad + \log(4/\rho). \end{aligned} \tag{7}$$

Plugging (7) into (6), we conclude that if Φ satisfies the RIP of the order S stated in Theorem III.1 with conditioning $\frac{\epsilon}{64}$, then

$$\begin{aligned} &\mathbb{P} \left\{ \sup_{x \in U(\mathcal{M})} \left| \|\widehat{\Phi} x\|_2 - 1 \right| > \epsilon \right\} \\ &\le \mathbb{P} \left\{ \max_{x \in U(B)} \left| \|\widehat{\Phi} x\|_2 - 1 \right| > \frac{\epsilon}{16} \right\} \le \rho. \end{aligned}$$

V. Conclusions and Future Work

In this paper, we showed that all measurement operators Φ satisfying the RIP can also be used (in conjunction with a random sign sequence) to obtain a stable embedding of a manifold. Moreover, we gave specific examples of such operators, namely subgaussian random matrices, random convolution matrices and deterministically constructed matrices, together with the requisite number of measurements needed for each operator to ensure a stable embedding of the manifold with high probability and for a pre-determined conditioning. This result represents a combination of two directions of recent interest in the CS community: structured measurement matrices and the development of structured signal models beyond sparsity.

While this result is encouraging, there are several areas of improvement that would be valuable. For example, we note that our Lemma IV.1 is an extension of Lemma 3 in [18]. There, the author showed that we can replace the unit sphere in $\mathbb{R}^N$ with a finite covering of the sphere and we extend this idea to discretizing the set of all directions of a manifold $U(\mathcal{M})$. However, compared to [18], our lemma is *not* operator agnostic in the sense that the operator $\widehat{\Phi}$ involved in the random process needs to be derived from an operator satisfying the RIP. There are advantages to being operator agnostic, including the result that a simple concentration of measure inequality for the operator $\widehat{\Phi}$ will suffice to ensure a stable embedding of the manifold (instead of specifically requiring Theorem II.1). Second, our result includes a logarithmic dependency on the ambient dimension N which we suspect is suboptimal because [6] shows a stable manifold embedding without this dependence. We are currently pursuing ideas related to different constructions of the covering set to try and tackle these issues.

Finally, another open question is whether the introduction of the random Rademacher sign sequence was absolutely necessary to prove our desired result. Note that this sign sequence introduces a notion of universality into Φ that allows it to be used with manifold signal models. However, just as the RIP can be shown without recourse to universality (e.g., subsampled Fourier matrices measuring only sparse signals but with degradation when working with signals that are sparse in other bases), it is perhaps possible that a less stringent condition can be imposed on Φ and still achieve stable embeddings of a class of manifolds.

APPENDIX
PROOF SKETCH FOR LEMMA IV.1

The proof of this lemma essentially follows from Lemmas 3.1, 3.2, 3.3 and Section 3.2.5 in [5]. In words, we express $U(\mathcal{M})$ as the union of two sets:

$$U^n_\alpha(\mathcal{M}) := \left\{ \frac{x_1 - x_2}{\|x_1 - x_2\|_2} \in U(\mathcal{M}) \mid d_{\mathcal{M}}(x_1, x_2) \le \alpha \right\}$$
$$U^f_\alpha(\mathcal{M}) := \left\{ \frac{x_1 - x_2}{\|x_1 - x_2\|_2} \in U(\mathcal{M}) \mid d_{\mathcal{M}}(x_1, x_2) > \alpha \right\}.$$

The first of these sets contains normalized difference vectors between all nearby points on $\mathcal{M}$, while the second contains normalized difference vectors between all faraway points on $\mathcal{M}$; the parameter α delineates nearness from farness. Denoting the random process by $Z(x) := \left| \|\widehat{\Phi}x\|_2 - 1 \right|$, we then see that

$$\sup_{x \in U(\mathcal{M})} Z(x) = \max \left\{ \sup_{x \in U^n_\alpha(\mathcal{M})} Z(x), \sup_{x \in U^f_\alpha(\mathcal{M})} Z(x) \right\}.$$

From here, we first observe that the each of the vectors in $U^n_\alpha(\mathcal{M})$ can be approximated by some tangent vector in some tangent plane to the manifold from the set $\bigcup_{a \in A}(a + Q_2(a))$. This fact is the topic of Lemmas 3.1 and 3.2 in [5]. Second, we observe that each vector in $U^f_\alpha(\mathcal{M})$ can be approximated by some pair of points from the geodesic covering set A. This is the topic of Lemma 3.3 in [5]. Additionally, we require the ℓ_2 norm of our operator $\widehat{\Phi}$ to be bounded, and an adequate bound follows from the fact that Φ satisfies RIP-$(S, \frac{\epsilon}{4})$. In particular, using elementary geometry we have: $\|\widehat{\Phi}\|_2 = \|\Phi\|_2 \le \left(1 + \frac{\epsilon}{4}\right)\sqrt{N}$. Finally, Section 3.2.5 in [5] determines the specific values for T and γ that we require to define the covering set B.

Another way to view Lemma IV.1 is as follows. Suppose we fix Φ with RIP conditioning δ and fix parameters T and γ of the covering set B. Then these parameters limit the possible range of values of the conditioning parameter ϵ that can be achieved for a stable embedding of $\mathcal{M}$. More specifically, $\epsilon \ge \max\left\{4\delta, \sqrt{\frac{NT}{C_1 \tau}}, \frac{\sqrt{N}\gamma}{C_2}\right\}$.

REFERENCES

[1] R. G. Baraniuk, M. A. Davenport, R. A. DeVore, and M. B. Wakin, "A simple proof of the restricted isometry property for random matrices," *Const. Approx.*, vol. 28, no. 3, pp. 253–263, Dec. 2008.

[2] D. L. Donoho and C. Grimes, "Image manifolds which are isometric to Euclidean space," *J. Math. Imaging Comp. Vision*, vol. 23, no. 1, pp. 5–24, July 2005.

[3] M. Turk and A. Pentland, "Eigenfaces for recognition," *J. Cogn. Neurosci.*, vol. 3, no. 1, pp. 71–83, 1991.

[4] E. J. Candès, "Compressive sampling," in *Proc. Int. Congr. Math.*, vol. 3, Madrid, Spain, August 2006, pp. 1433–1452.

[5] R. G. Baraniuk and M. B. Wakin, "Random projections of smooth manifolds," *Found. Comp. Math.*, vol. 9, no. 1, pp. 51–77, 2009.

[6] K. L. Clarkson, "Tighter bounds for random projections of manifolds," in *Proc. 24th Annual Symp. on Computational Geometry*. ACM, 2008, pp. 39–48.

[7] C. Hegde, M. Wakin, and R. Baraniuk, "Random Projections for Manifold Learning: Proofs and Analysis," *Technical Report TREE 0710, Rice University*, 2007.

[8] J. Haupt, W. U. Bajwa, G. Raz, and R. Nowak, "Toeplitz compressed sensing matrices with applications to sparse channel estimation," *IEEE Trans. Inform. Theory*, 2008.

[9] H. Rauhut, J. Romberg, and J. Tropp, "Restricted isometries for partial random circulant matrices," *Arxiv preprint arXiv:1010.1847*, 2010.

[10] R. A. DeVore, "Deterministic constructions of compressed sensing matrices," *J. Complexity*, vol. 23, no. 4-6, pp. 918–925, 2007.

[11] F. Krahmer and R. Ward, "New and improved Johnson-Lindenstrauss embeddings via the Restricted Isometry Property," *Arxiv preprint arXiv:1009.0744*, 2010.

[12] P. Niyogi, S. Smale, and S. Weinberger, "Finding the homology of submanifolds with high confidence from random samples," *J. Discrete and Computational Geometry*, March 2008.

[13] S. Dasgupta and A. Gupta, "An elementary proof of the Johnson-Lindenstrauss lemma," *Random Struct. Algor.*, vol. 22, no. 1, pp. 60–65, 2002.

[14] L. A. Gottlieb and R. Krauthgamer, "A nonlinear approach to dimension reduction," *Arxiv preprint arXiv:0907.5477*, 2009.

[15] A. Gupta, R. Krauthgamer, and J. Lee, "Bounded geometries, fractals, and low-distortion embeddings," 2003.

[16] P. Indyk and A. Naor, "Nearest-neighbor-preserving embeddings," *ACM Trans. Algorithms (TALG)*, vol. 3, no. 3, p. 31, 2007.

[17] M. Rudelson and R. Vershynin, "On sparse reconstruction from Fourier and Gaussian measurements," *Communications on Pure and Applied Mathematics*, vol. 61, no. 8, pp. 1025–1045, 2008.

[18] R. Vershynin, "Introduction to the non-asymptotic analysis of random matrices," *Compressed sensing: theory and applications. Cambridge University Press. Submitted*.