Distance Preserving Embeddings of Riemannian Manifolds from Samples

Nakul Verma
Janelia, HHMI

Sample-based manifold embeddings

Let's fix a desirable property: **preserving geodesic distances**.

We are interested in the following question:

Given: a sample *X* from *n*-dimensional manifold $M \subset \mathbf{R}^D$, and

an embedding procedure $\mathcal{A} : M \to \mathbf{R}^d$

Define: the **quality** of embedding as $(1 \pm \epsilon)$-isometric, if for all distinct p, q

$$(1-\epsilon) \leq \frac{D(\mathcal{A}(p),\mathcal{A}(q))}{D(p,q)} \leq (1+\epsilon)$$

Questions:

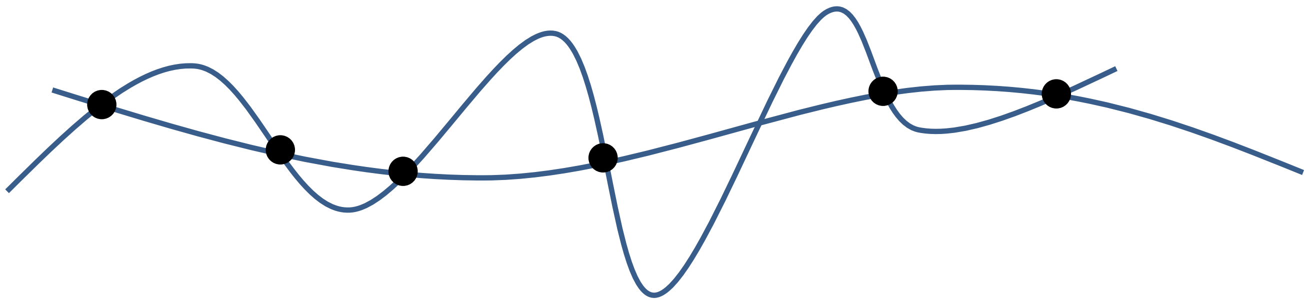
I. Can one come up an $\mathcal{A}$ that achieves $(1 \pm \epsilon)$-isometry?

II. How much can one reduce *d* and still have $(1 \pm \epsilon)$-isometry?

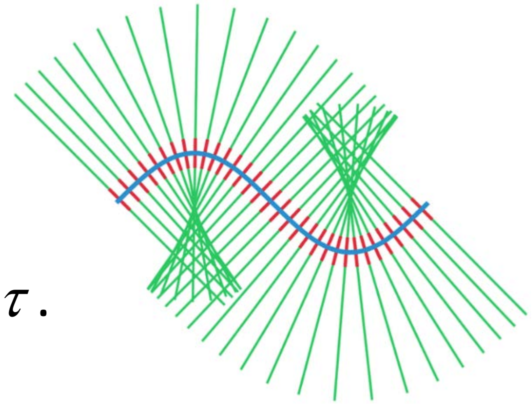
III. Do we need any restriction on *X* or *M*?

Preliminaries

We only have samples



Manifold condition number [Niyogi, Smale, Weinberger '06]



A submanifold $M \subset \mathbf{R}^D$ has condition number ($1/\tau$),
if τ is the largest number such that:
normals of M of length r are nonintersecting for all $r < \tau$.

Sample-based manifold embeddings

Let's fix a desirable property: **preserving geodesic distances**.

We are interested in the following question:

Given: a sample *X* from *n*-dimensional manifold $M \subset \mathbf{R}^D$, and
an embedding procedure $\mathcal{A} : M \rightarrow \mathbf{R}^d$

Define: the **quality** of embedding as $(1 \pm \epsilon)$-isometric, if for all distinct p, q

$$(1-\epsilon) \leq \frac{D(\mathcal{A}(p),\mathcal{A}(q))}{D(p,q)} \leq (1+\epsilon)$$

Questions:

I. Can one come up an $\mathcal{A}$ that achieves $(1 \pm \epsilon)$-isometry?

II. How much can one reduce *d* and still have $(1 \pm \epsilon)$-isometry?

III. Do we need any restriction on *X* or *M*?

Recent progress in this area

An interesting result:

Let $M \subset \mathbf{R}^D$ be a compact *n*-dimensional manifold with volume *V* and curvature τ.

Then projecting it to a random **linear** subspace of dimension

$$O\left(\frac{n}{\epsilon^2} \log \frac{V}{\tau}\right)$$

achieves $(1 \pm \epsilon)$-isometry with high probability.

Does not need any samples from the underlying manifold!

[Baraniuk & Wakin ’08, Clarkson ’08, V. ’11]

Recent progress in this area

An interesting result:

Let $M \subset \mathbf{R}^D$ be a compact *n*-dimensional manifold with volume *V* and curvature τ.

Then projecting it to a random **linear** subspace of dimension

$$O\left(\frac{n}{\epsilon^2} \log \frac{V}{\tau}\right)$$

Highly undesirable!
To have all distances within factor of 99% requires projection dimension > 10,000!

achieves $(1 \pm \epsilon)$-isometry with high probability.

Does not need any samples from the underlying manifold!

[Baraniuk & Wakin '08, Clarkson '08, V. '11]

What we show

For **any** compact *n*-dimensional manifold $M \subset \mathbf{R}^D$ we present an algorithm that can embed *M* in

$$O(n + \ln(V/\tau^n))$$

dimensions that achieves $(1 \pm \epsilon)$-isometry (using only samples from *M*).

Embedding dimension is independent of ϵ !

Sample size is a function of ϵ

The Algorithm

Embedding Stage: Find a representation of *M* in lower dimensional space without worrying about maintaining any distances.

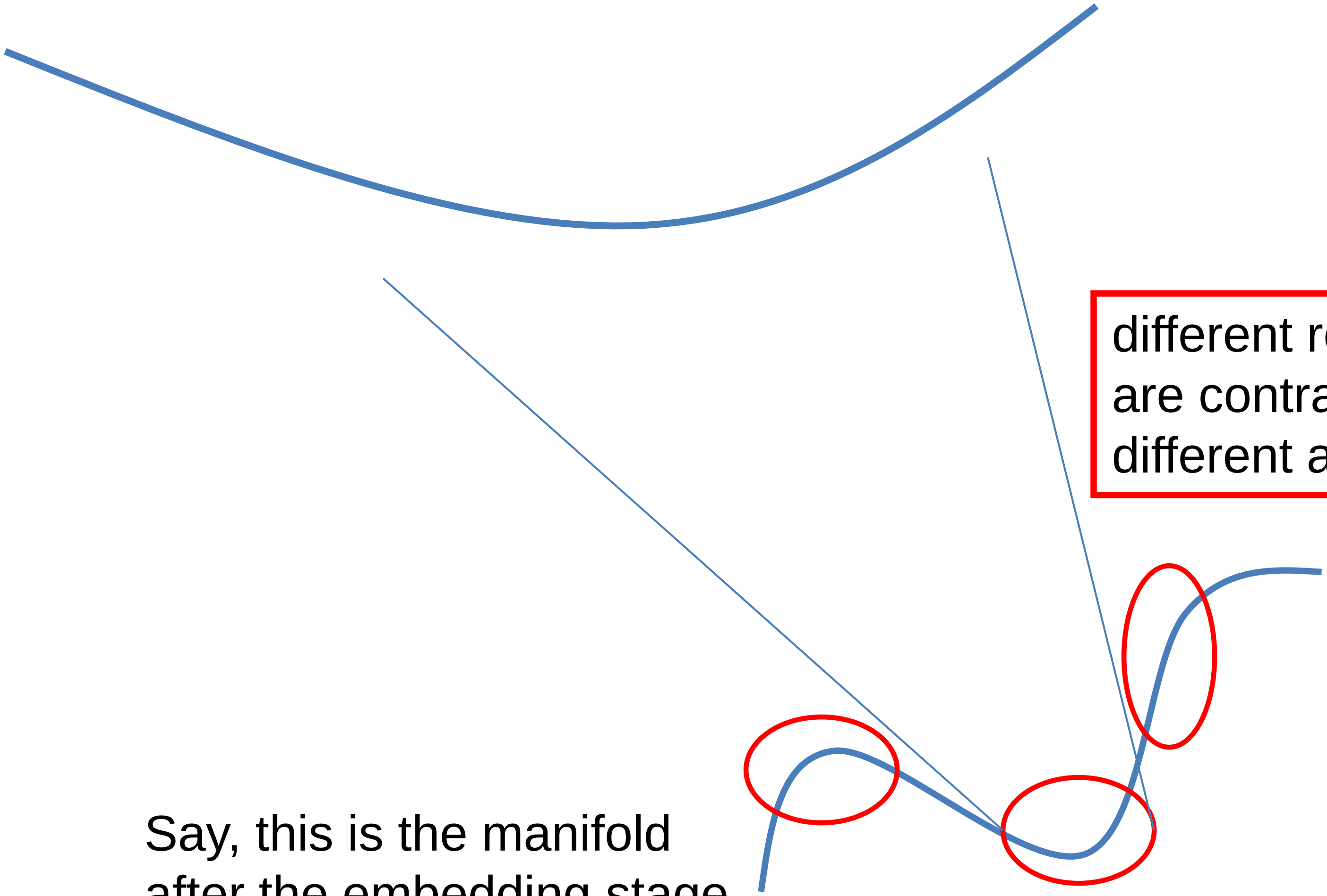
We can use a random linear projection without $1/\epsilon^2$ penalty

Correction Stage: Apply a series of corrections, each corresponding to a different region of the manifold, to restore back the distances.

This requires a bit of thinking…

Corrections

Zoomed in a local region



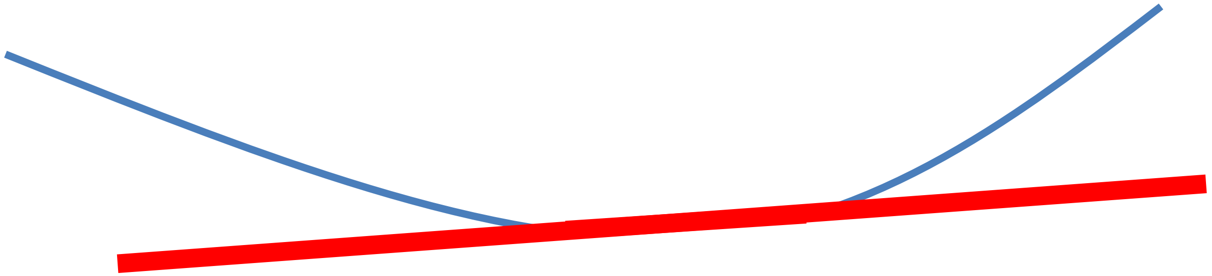
different regions are contracted by different amounts

Say, this is the manifold after the embedding stage

$\sigma \leq 1$ shrinkage amount

Corrections

Zoomed in a local region



Suppose we linearly stretch this local region

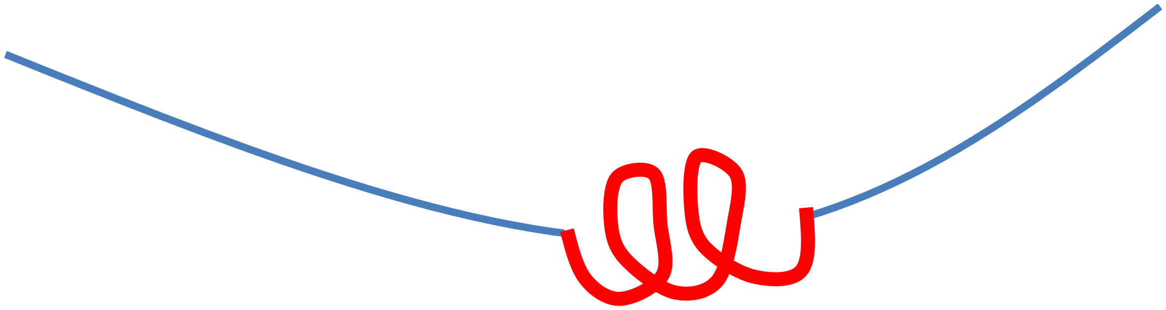
cannot systematically attach the boundary of the stretched region back to the manifold…

$t \mapsto \sigma^{-1} t$

$\sigma \leq 1$ shrinkage amount

Corrections

Zoomed in a local region



Instead we locally twist the space!

This creates the necessary stretch to restore back the local distances!

$C = \sqrt{\sigma^{-2} - 1}$ correction

$t \mapsto (t, \sin(Ct), \cos(Ct))$

$\sigma \leq 1$ shrinkage amount

Technical challenges

Since working with samples, each step of the algorithm results in additive amount of approximation error!

- Need to **estimate the contraction** at every local region.
- Find sufficient amount of ambient space to **create the local twist**.
- Care needs to be taken so as not to have **sharp** (non-differentiable) **edges on the boundary** while locally twisting the space.
- **Interference between different corrections** at overlapping localities need to be reconciled.

The algorithm

Input: Sample X from M, local neighborhood size ρ.

Let Φ denote the initial random projection in $O(n)$ dim.

Preprocess:

- For each $x \in X$, let F_x be the local tangent space approximation using neighborhood size ρ.
- Let $U_x \Sigma_x V_x^{\mathsf{T}}$ be the SVD of ΦF_x.
- Estimate local correction around x as:

$$C_x := (\Sigma_x^{-2} - I)^{1/2} U_x^{\mathsf{T}}$$

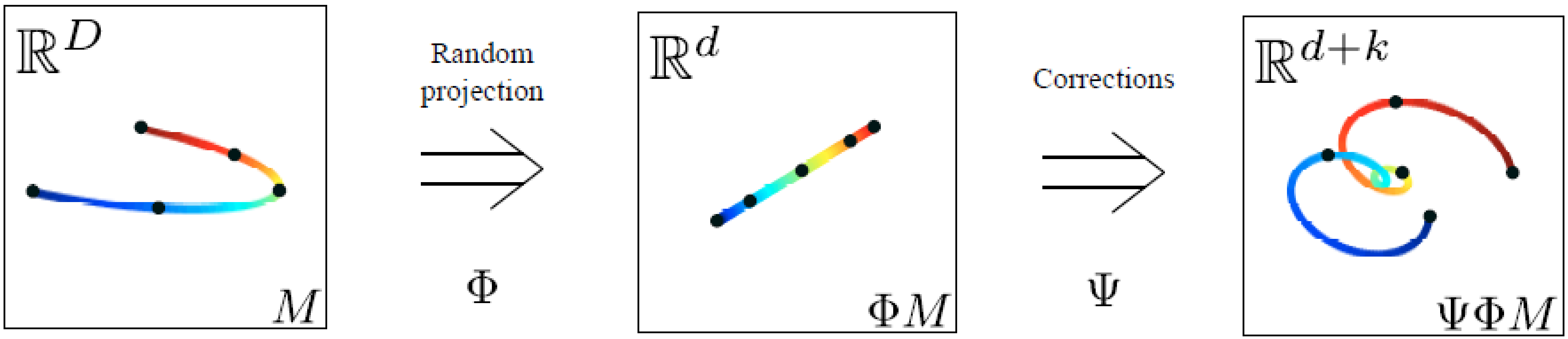
Embedding: For any $p \in M$

- $t = \Phi p$.
- **for** every $x \in X$:
  - let $\Psi_{i\text{-}1}(t)$ be the embedding from previous iteration.
  - let η and ν be vectors normal to $\Psi_{i\text{-}1}(t)$.
  - let Λ_x be a localizing kernel.
  - apply correction

$$\Psi_i = \Psi_{i-1} + \eta\sqrt{\Lambda_x}\sin(C_x t) + \nu\sqrt{\Lambda_x}\cos(C_x t)$$

- **return** $\Psi_{|X|}(t)$

The algorithm at work



Theoretical guarantee

Theorem:

Let $M \subset \mathbf{R}^D$ be compact *n*-dimensional manifold with volume *V* and curvature τ. For any ϵ > 0, let *X* be $(D/\epsilon\tau)^n$-dense sample from *M*.

Then with high probability, our algorithm (given access to *X*) embeds any point from *M* in dimension

$$O(n + \ln(V/\tau^n))$$

with $(1 \pm \epsilon)$-isometry.

Our embedding is C^∞

A quick proof overview

The goal is to **prove that the geodesic distances** between all pairs of points *p* and *q* in *M* are approximately preserved.

Recall that **length of any curve** γ is given by the expression:

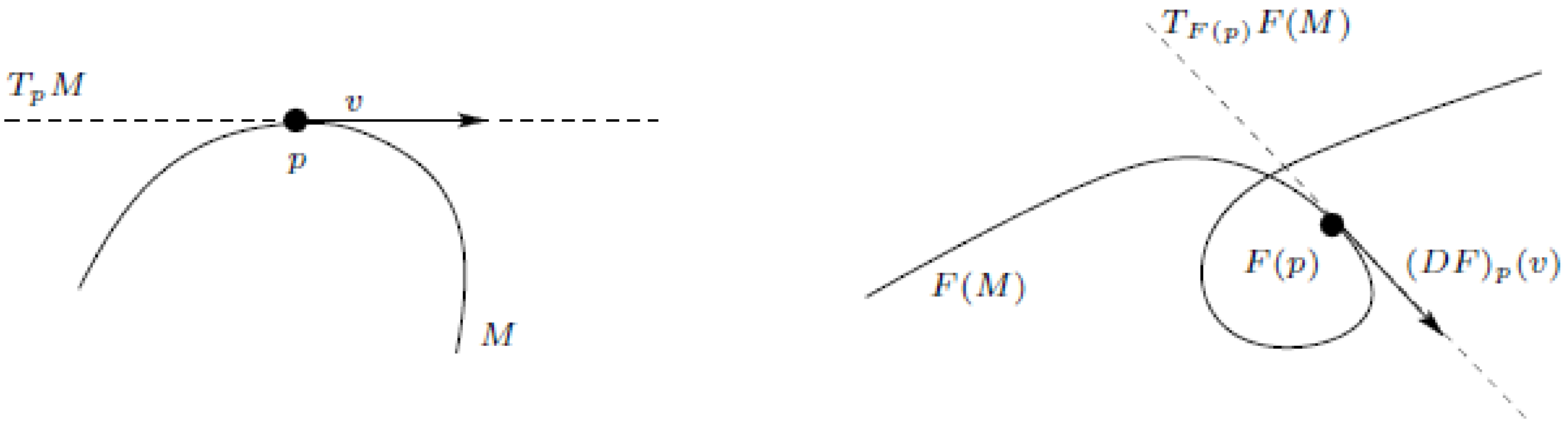
$$\int_a^b \|\gamma'(s)\| \, ds$$

length of a curve is the infinitesimal sum of the length of the tangent vectors along its path

Therefore, suffices to show that our algorithm preserves lengths of all vectors tangent to at all points in *M*.

A quick proof overview

From differential geometry, we know that for any smooth map *F*



the **exterior derivative** or the **pushforward** map *DF* acts on the tangent vectors.

We carefully analyze how each correction step of the algorithm changes the corresponding pushforward map.

Conclusion and implications

- Gave the first **sample complexity result** for approximately isometric embedding for a manifold learning algorithms.
- **Novel algorithmic and analysis techniques** are of independent interest.
- One can use an existing manifold learning algorithm as the ‘embedding’ step. The corrections in second step enhance the embedding to make it isometric, making this as a **universal procedure**.

Summary of known embedding results

Riemannian Geometry	Manifold Learning
Whitney's result (medium form) 2n+1 Differential structure preserved	Random projection 2n+1 Differential structure preserved a.s.
	Random projection $O(\varepsilon^{-2} n \log(V / \tau))$ Euclidean and Geodesic $(1 \pm \varepsilon)$ w.h.p.
Nash / Kuiper 2n+1 All paths preserved	Our result $O(n + \log(V / \tau^n))$ All paths preserved upto $(1 \pm \varepsilon)$ w.h.p.

Open problems

- How can we determine the curvature bound τ, or other geometric properties?
- Is it possible to embed a manifold with constant distortion that only depends on n?
- Is it possible to reduce the sampling requirement?

Questions / Discussion

Thank You!