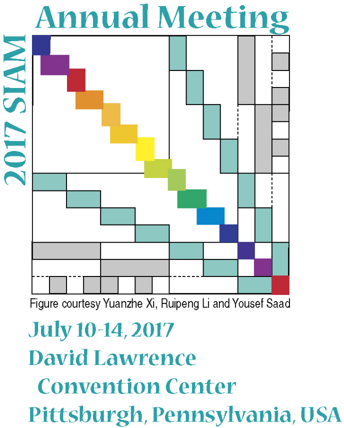
Geometry and Computational Challenges in Data Science (GCCDS)



Diffusion Geometry and Manifold Learning on Fibre Bundles

Tingran Gao

Visiting Assistant Professor
Department of Mathematics
Duke University

2017 SIAM Annual Meeting
Minisymposium
Geometry and Computational Challenges in Data Science (GCCDS)
PIttsburgh, PA

Tuesday July 11, 2017

Outline

Background & Motivations

- Graph Synchronization Problems

Manifold Learning on Fibre Bundles

- Diffusion Geometry
- Fibre Bundles
- Horizontal Diffusion Maps

Applications

- Evolutionary Anthropology

Graph Synchronization Problems

- Data:
  - graph $\Gamma = (V, E)$
  - matrix group G, equipped with a norm $\|\cdot\|$
  - **edge potential** $\rho : E \to G$ satisfying $\rho_{ij} = \rho_{ji}^{-1}, \; \forall (i,j) \in E$

Graph Synchronization Problems

- Data:
  - graph $\Gamma = (V, E)$
  - matrix group G, equipped with a norm $\|\cdot\|$
  - **edge potential** $\rho : E \to G$ satisfying $\rho_{ij} = \rho_{ji}^{-1}, \ \forall (i,j) \in E$
- Goal:
  - find a **vertex potential** $f : V \to G$ such that

$$f_i = \rho_{ij} f_j, \quad \forall (i,j) \in E$$

Graph Synchronization Problems

- Data:
  - graph $\Gamma = (V, E)$
  - matrix group G, equipped with a norm $\|\cdot\|$
  - **edge potential** $\rho : E \to G$ satisfying $\rho_{ij} = \rho_{ji}^{-1}, \ \forall (i,j) \in E$
- Goal:
  - find a **vertex potential** $f : V \to G$ such that

$$f_i = \rho_{ij} f_j, \quad \forall (i,j) \in E$$

- The goal can be achieved if and only if $\rho_{ij} = f_i f_j^{-1}$

Graph Synchronization Problems

- Data:
  - graph $\Gamma = (V, E)$
  - matrix group G, equipped with a norm $\|\cdot\|$
  - **edge potential** $\rho : E \to G$ satisfying $\rho_{ij} = \rho_{ji}^{-1}, \ \forall (i,j) \in E$
- Goal:
  - find a **vertex potential** $f : V \to G$ such that

$$f_i = \rho_{ij} f_j, \quad \forall (i,j) \in E$$

- The goal can be achieved if and only if $\rho_{ij} = f_i f_j^{-1}$
- **Not** always feasible!

Graph Synchronization Problems

- Data:
  - graph $\Gamma = (V, E)$
  - matrix group G, equipped with a norm $\|\cdot\|$
  - **edge potential** $\rho : E \to G$ satisfying $\rho_{ij} = \rho_{ji}^{-1}, \ \forall (i,j) \in E$
- Goal:
  - find a **vertex potential** $f : V \to G$ such that

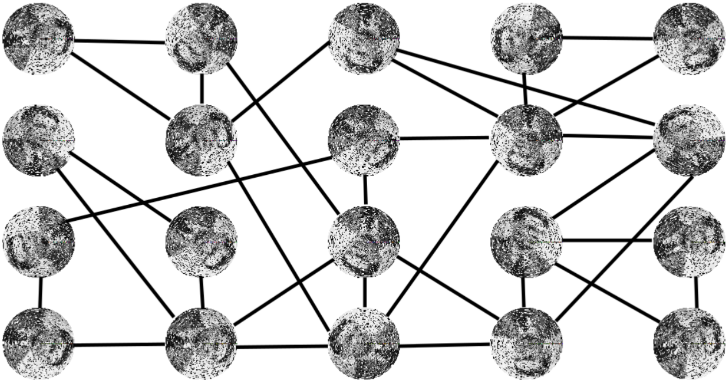
$$f_i = \rho_{ij} f_j, \quad \forall (i,j) \in E$$

- The goal can be achieved if and only if $\rho_{ij} = f_i f_j^{-1}$
- **Not** always feasible!
- If infeasible, find the "closest solution" in the sense of

$$\min_{\substack{f:V\to G \\ \|f\|\neq 0}} \frac{1}{2} \frac{\sum_{i,j\in V} \|f_i - \rho_{ij} f_j\|^2}{\sum_{i\in V} \|f_i\|^2} (=: \eta(f))$$

A Toy Example

$$y_i = R_i x + \xi_i$$
$$R_i \in O(d), \quad \xi_i \sim \text{i.i.d. noise}$$



Afonso S. Bandeira. "Ten Lectures and Forty-Two Open Problems in the Mathematics of Data Science." (2015).

$$\boxed{\begin{array}{c} y_i = R_i x + \xi_i \\ R_i \in O(d), \quad \xi_i \sim \text{i.i.d. noise} \end{array}}$$

$$\begin{aligned} \text{Measurement:} \quad & R_{ij} \approx R_i^\top R_j \\ \text{Recover:} \quad & R_1, R_2, \cdots \end{aligned}$$

$$\boxed{\begin{array}{c} y_i = R_i x + \xi_i \\ R_i \in O(d), \quad \xi_i \sim \text{i.i.d. noise} \end{array}}$$

$$\begin{aligned} \text{Measurement:} \quad & R_{ij} \approx R_i^\top R_j \\ \text{Recover:} \quad & R_1, R_2, \cdots \end{aligned}$$

$\Rightarrow$ Solve the minimization problem

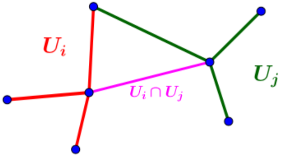
$$\min_{R_1, R_2, \cdots \in O(d)} \sum_{(i,j) \in E} w_{ij} \left\| R_{ij} - R_i^\top R_j \right\|_{\mathrm{F}}^2$$

Synchronization Problems: Examples

- Manifold Orientability [Singer, Wu (2011)]: $G = O(1)$
- Angular Synchronization [Singer (2011)]: $G = U(1)$
- Vector Diffusion Maps [Singer, Wu (2012)]: $G = O(d)$
- Multireference Alignment [Bandeira et al. (2014)]: $G = \{\text{cyclic shifts}\}$
- Global Registration of Point Clouds [Chaudhury (2015)]: $G = \mathbb{E}_d$
- Collection Shape Matching [Nguyen et al. (2011)], [Huang, Guibas (2013)], [Chen et al. (2014)], [Maron et al. (2016)]: $G = S_n$ (symm. group of n elements)
- Cryo-EM Structural Reconstruction [Singer et al. (2011)], [Shkolnisky, Singer (2012)], [Zhao, Singer (2014)], [Bandeira et al. (2015)]: $G = SO(3)$
- Cartan Motion Groups [Ozyesil et al. (2016)]: $G = K \ltimes V$ (more about this soon — in Nir Sharon's talk)

The Geometry of Synchronization Problems

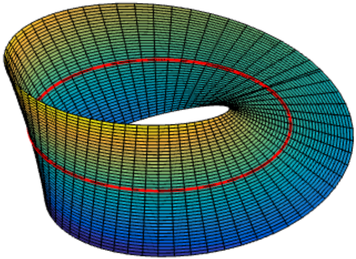
- Data:
  - graph $\Gamma = (V, E)$
  - linear algebraic group G, equipped with a norm $\|\cdot\|$
  - **edge potential** $\rho : E \to G$ satisfying $\rho_{ij} = \rho_{ji}^{-1}, \ \forall (i,j) \in E$
- Observation:
  - Let $\mathfrak{U} = \{U_i \mid 1 \leq i \leq |V|\}$ be an open cover of Γ (viewed as a 1-dimensional simplicial complex), where U_i is the *(open) star neighborhood* of vertex i.

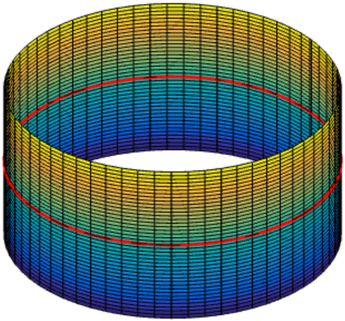


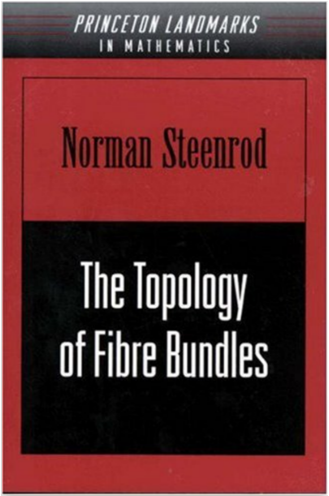
  - The ρ defines a *flat principal G-bundle* over Γ (denoted as $\mathscr{B}_\rho$).

Fibre Bundle $\mathscr{E} = (E, M, F, \pi)$

- E: total manifold
- M: base manifold
- F: fibre
- E is "locally equivalent" to $M \times F$, but not necessarily so globally!



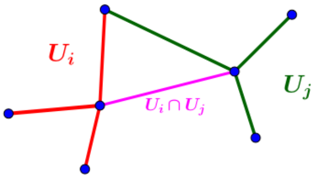




Theorem (Steenrod 1951, §2).
If topological group G acts on F and $\{U_i\}$, $\{\rho_{ij}\}$ is a system of coordinate transformations in the space M such that

$$\rho_{ii} = e \in G \quad \text{for all } U_i$$
$$\rho_{ij} = \rho_{ji}^{-1} \quad \text{if } U_i \cap U_j \neq \emptyset$$
$$\rho_{ij}\rho_{jk} = \rho_{ik} \quad \text{if } U_i \cap U_j \cap U_k \neq \emptyset$$

then there exists a fibre bundle $\mathscr{B}$ with base space M, fibre F, group G, and coordinate transforms $\{\rho_{ij}\}$.



No triple intersections!

Theorem (Steenrod 1951, §2). If topological group G acts on F and $\{U_i\}$, $\{\rho_{ij}\}$ is a system of coordinate transformations in the space M such that

$$\rho_{ii} = e \in G \quad \text{for all } U_i$$

$$\rho_{ij} = \rho_{ji}^{-1} \quad \text{if } U_i \cap U_j \neq \emptyset$$

~~$\rho_{ij}\rho_{jk} = \rho_{ik} \quad \text{if } U_i \cap U_j \cap U_k \neq \emptyset$~~

then there exists a fibre bundle $\mathscr{B}$ with base space M, fibre F, group G, and coordinate transforms $\{\rho_{ij}\}$.

Geometric Observations

- Denote

$$C^0(\Gamma; G) := \{f : V \to G\} \text{ vertex potentials}$$
$$C^1(\Gamma; G) := \left\{\rho : E \to G \mid \rho_{ij} = \rho_{ji}^{-1}, \forall (i,j) \in E\right\} \text{ edge potentials}$$

Geometric Observations

- Denote

$$C^0(\Gamma; G) := \{f : V \to G\} \text{ vertex potentials}$$
$$C^1(\Gamma; G) := \left\{\rho : E \to G \mid \rho_{ij} = \rho_{ji}^{-1}, \forall (i,j) \in E\right\} \text{ edge potentials}$$

- Consider the right action of $C^0(\Gamma; G)$ on $C^1(\Gamma; G)$:

$$C^1(\Gamma; G) \times C^0(\Gamma; G) \to C^1(\Gamma; G)$$
$$(\rho, f) \longmapsto \tau_\rho f$$

defined as $(\tau_f \rho)_{ij} := f_i^{-1} \rho_{ij} f_j, \quad \forall (i,j) \in E$.

Geometric Observations

- Denote

$$C^0(\Gamma;G) := \{f : V \to G\} \text{ vertex potentials}$$
$$C^1(\Gamma;G) := \left\{\rho : E \to G \mid \rho_{ij} = \rho_{ji}^{-1}, \forall (i,j) \in E\right\} \text{ edge potentials}$$

- Consider the right action of $C^0(\Gamma;G)$ on $C^1(\Gamma;G)$:

$$C^1(\Gamma;G) \times C^0(\Gamma;G) \to C^1(\Gamma;G)$$
$$(\rho, f) \longmapsto \tau_\rho f$$

  defined as $(\tau_f \rho)_{ij} := f_i^{-1} \rho_{ij} f_j, \quad \forall (i,j) \in E$.

- ρ synchronizable $\Leftrightarrow$ $\tau_f \rho$ synchronizable for all $f \in C^0(\Gamma;G)$, i.e. synchronizability is defined at the level of equivalence classes $C^1(\Gamma;G) / C^0(\Gamma;G)$

Moduli Space of Synchronization Data

Theorem (G., Brodzki, Muhkerjee (2016)). There exists a one-to-one correspondence (between *sets*)

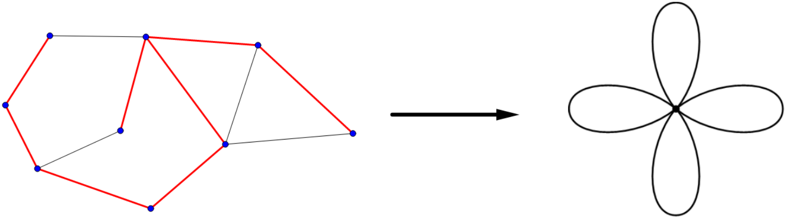
$$C^1(\Gamma; G) / C^0(\Gamma; G) \cong \mathrm{Hom}(\pi_1(\Gamma), G) / G$$

where G acts on $\mathrm{Hom}(\pi_1(\Gamma), G)$ by conjugations:

$$\begin{aligned} \mathrm{Hom}(\pi_1(\Gamma), G) \times G &\longrightarrow \mathrm{Hom}(\pi_1(\Gamma), G) \\ (\phi, g) &\longmapsto g^{-1}\phi g \end{aligned}$$

Tingran Gao, Jacek Brodzki, Sayan Mukherjee. "The Geometry of Synchronization Problems and Learning Group Actions." *submitted.* arXiv:1610.09051, 2016

Fundamental Group of a Graph?



$$\pi_1(\Gamma) = \bigvee_{k=1}^{|E|-|V|+1} S^1$$

$$\mathrm{Hom}(\pi_1(\Gamma), G) = \underbrace{G * G * G * \cdots * G * G * G}_{(|E|-|V|+1)\text{-copy free product}}$$

$$\boxed{C^0(\Gamma; G)/C^1(\Gamma; G) \cong \mathrm{Hom}(\pi_1(\Gamma), G)/G}$$

- Proof builds upon construction of a *holonomy homomorphism*
- The orbit space $C^0(\Gamma; G)/C^1(\Gamma; G)$ is exactly the *first cohomology set* $\check{H}^1((\Gamma, \mathfrak{U}), \underline{G})$
- "Synchronizability" is a property at the level of equivalence classes $\left[f_i^{-1}\rho_{ij}f_j\right]_{(i,j)\in E}$
- $\rho \in C^1(\Gamma; G)$ synchronizable
  $\Leftrightarrow [\rho] = [e]$ as equivalence classes in $C^0(\Gamma; G)/C^1(\Gamma; G)$
  $\Leftrightarrow$ the principal G-bundle $\mathscr{B}_\rho$ is trivial
- Future work: study synchronization problems through the geometry of the moduli space/character variety

Quick Aside: A Twisted De Rham-Hodge Theory

- Combinatorial Hodge Theory:

$$0 \rightleftarrows \Omega^0(\Gamma) \underset{\delta}{\overset{d}{\rightleftarrows}} \Omega^1(\Gamma) \rightleftarrows 0,$$

- *Twisted* Combinatorial Hodge Theory:

$$0 \rightleftarrows C^0(\Gamma; F) \underset{\delta_\rho}{\overset{d_\rho}{\rightleftarrows}} \Omega^1(\Gamma; \mathscr{B}_\rho[F]) \rightleftarrows 0.$$

- **Theorem (G., Brodzki, Muhkerjee (2016)).** Define

$$\Delta_\rho^{(0)} := \delta_\rho d_\rho, \quad \Delta_\rho^{(1)} := d_\rho \delta_\rho$$

then the following Hodge-type decomposition holds:

$$C^0(\Gamma; F) = \ker \Delta_\rho^{(0)} \oplus \operatorname{im} \delta_\rho = \ker d_\rho \oplus \operatorname{im} \delta_\rho,$$
$$\Omega^1(\Gamma; \mathscr{B}_\rho[F]) = \operatorname{im} d_\rho \oplus \ker \Delta_\rho^{(1)} = \operatorname{im} d_\rho \oplus \ker \delta_\rho.$$

Application: Evolutionary Anthropology

Doug Boyer

Jukka Jernvall

More Precisely: biological morphologists

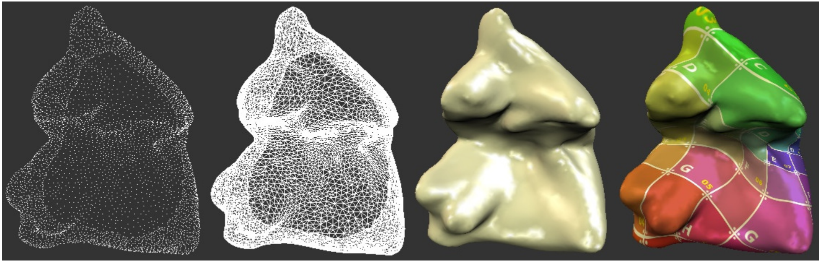
↓

Study Teeth & Bones of

extant & extinct animals

still live today   fossils

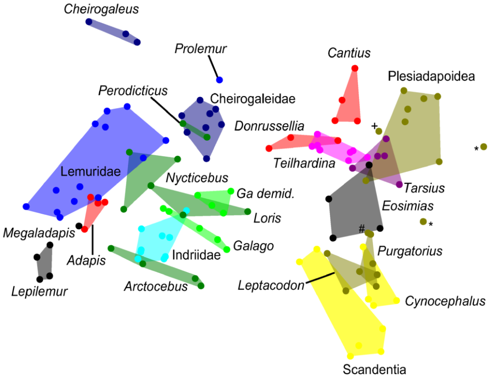
Data Acquisition: microCT (High Resolution X-ray CT)



Surface reconstructed from μCT-scanned voxel data

Landmarked Teeth $\longrightarrow$

$$d^2_{Procrustes}(S_1, S_2) = \min_{R \text{ rigid motion}} \frac{1}{k} \sum_{j=1}^{k} \| R(x_j) - y_j \|^2$$

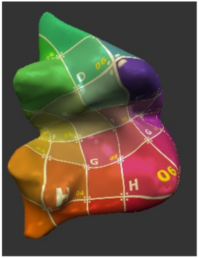


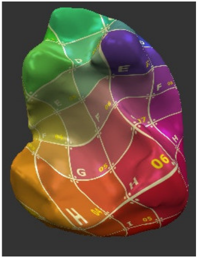
Boyer et al. "Algorithms to Automatically Quantify the Geometric Similarity of Anatomical Surfaces." *Proceedings of the National Academy of Sciences* 108.45 (2011): 18221-18226.

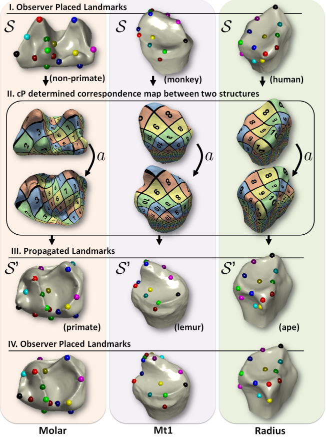
A Zoo of Shape Distances...

$d_{\mathrm{cWn}}(S_1, S_2)$: Conformal Wasserstein Distance (CWD)
$d_{\mathrm{cP}}(S_1, S_2)$: Continuous Procrustes Distance (CPD)
$d_{\mathrm{cKP}}(S_1, S_2)$: Continuous Kantorovich-Procrustes Distance (CKPD)

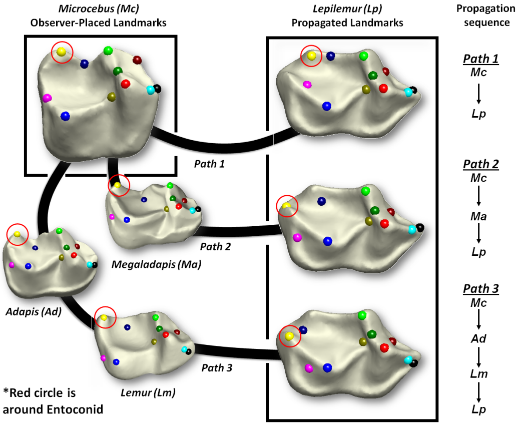
$$d_{\mathrm{cP}}(S_1, S_2) = \inf_{\mathcal{C} \in \mathcal{A}(S_1, S_2)} \inf_{R \in \mathbb{E}(3)} \left(\int_{S_1} \| R(x) - \mathcal{C}(x) \|^2 \, d\mathrm{vol}_{S_1}(x) \right)^{\frac{1}{2}}$$





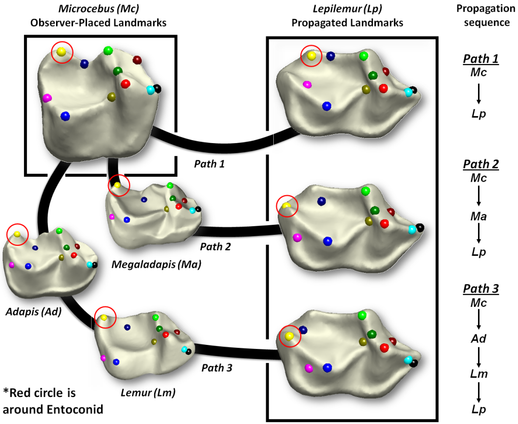


Interpretability Issue



Even mistakes made by CPD were similar to biologists' mistakes!

Resolving Interpretability Issue #1: Trust Small Distances



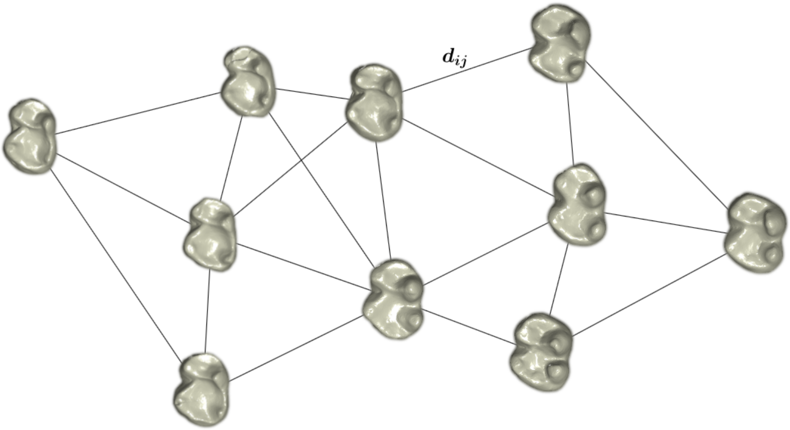
"Correct" like a biologist, but *automatically*?

small distances between $S_1, S_2 \longrightarrow$ OK maps

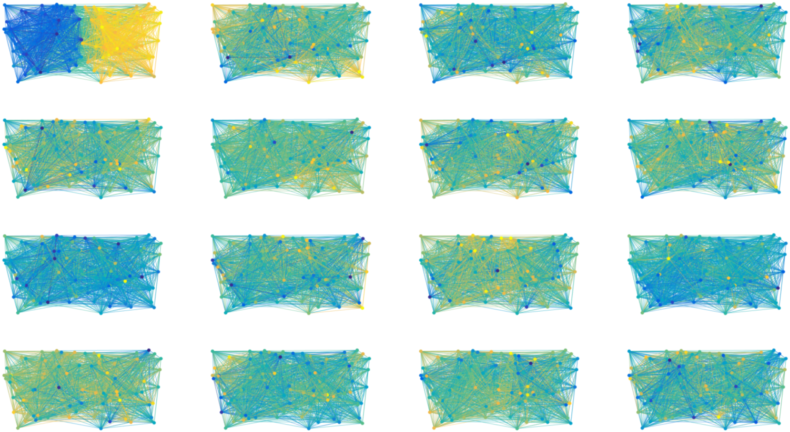
larger distances $\longrightarrow$ not OK

Gao et al. (2016) "Development and Assessment of Fully Automated and Globally Transitive Geometric Morphometric Methods." *submitted.* DOI: `http://dx.doi.org/10.1101/086280`

Trust Only *Small* Distances: *Geodesics* in Shape Space



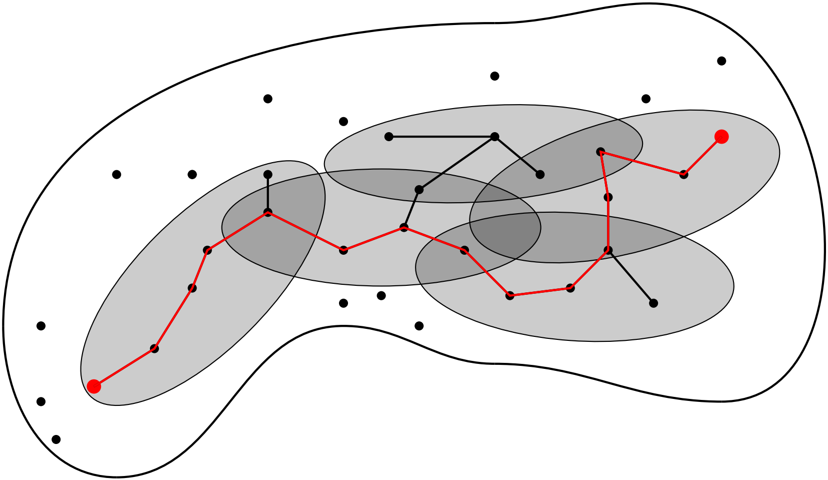
Diffusion Maps and Diffusion Distances



Diffusion Maps: Embedding Graphs into ℓ_2 using Eigenfunctions and the Heat Kernel of the Graph Laplacian

Coifman, R. R., and Lafon, S. "Diffusion Maps." *Appl. & Comput. Harmonic Analysis* 21, no. 1 (2006): 5-30.

Diffusion Maps: “Knit Together” Local Geometry

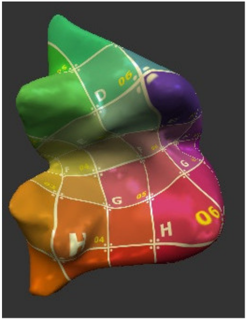


Small distances are much more reliable!

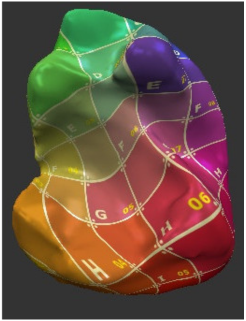
Resolving Interpretability Issue #2: Use Maps!

$$d_{\mathrm{cP}}(S_i, S_j) = \inf_{\mathcal{C}\in\mathcal{A}(S_i,S_j)} \inf_{R\in\mathbb{E}(3)} \left(\int_{S_i} \| R(x) - \mathcal{C}(x) \|^2 \, d\mathrm{vol}_{S_i}(x) \right)^{\frac{1}{2}}$$

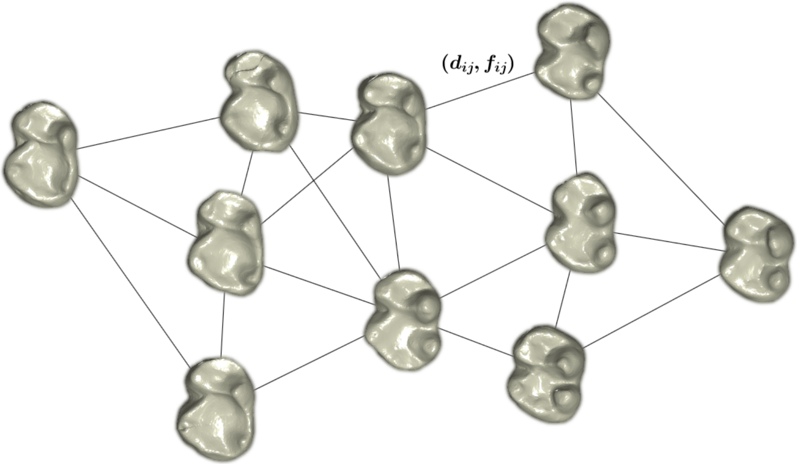
S_i

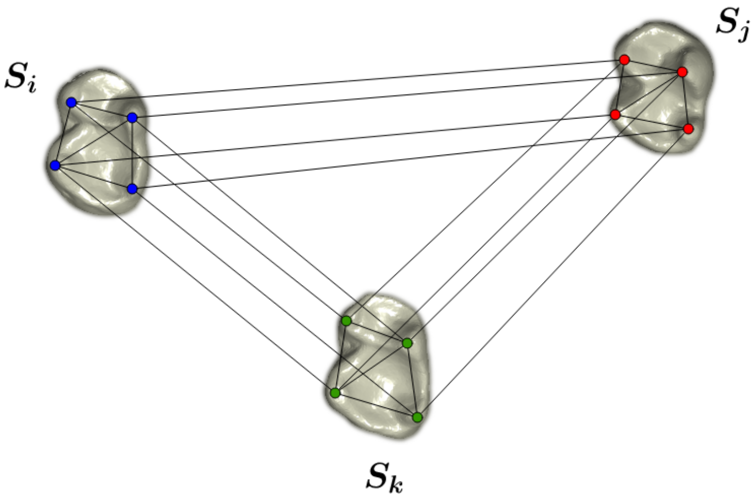


$\xrightarrow[f_{ij}]{d_{ij}}$



S_j


(d_{ij}, f_{ij})

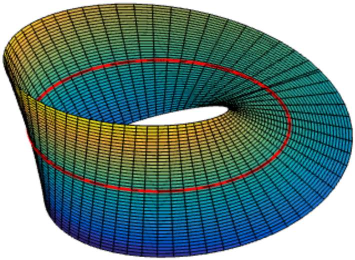


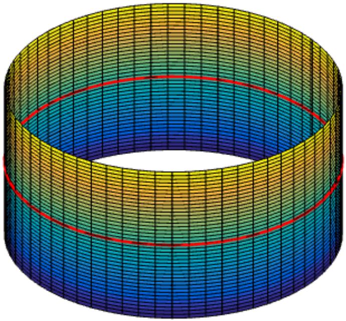
S_i
S_j
S_k

Geometric Model —— *Fibre Bundles*

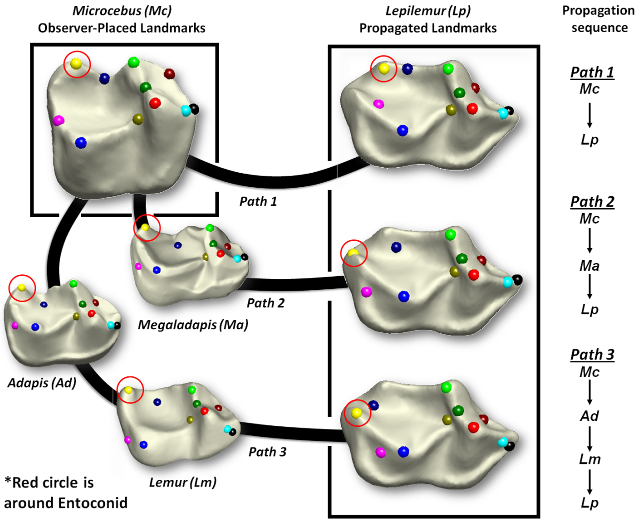
Fibre Bundle $\mathscr{E} = (E, M, F, \pi)$

- E: total manifold
- M: base manifold
- F: fibre
- E is "locally equivalent" to $M \times F$, but not necessarily so globally!

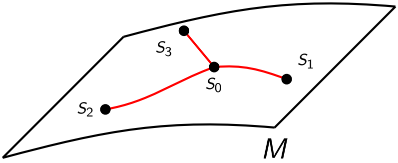


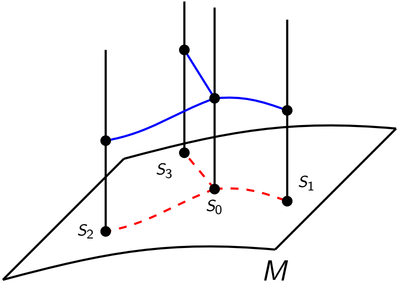


Shape Space is NOT a Trivial Fibre Bundle



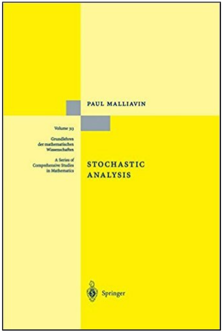
Horizontal Random Walk on a *Fibre Bundle*

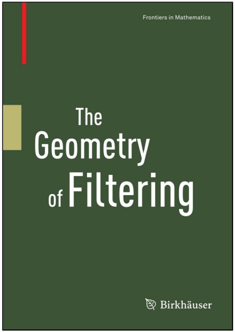


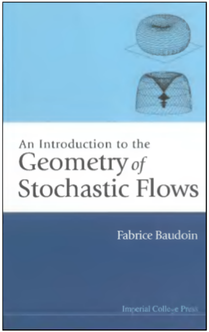


Horizontal Diffusion Process in Stochastic Geometry

- K.D. Elworthy, W.S. Kendall. "Factorization of Harmonic Maps and Brownian Motions." University of Warwick, 1985.
- M. Liao, "Factorization of Diffusions on Fibre Bundles." *Transactions of the American Mathematical Society.* 311.2 (1989): 813-827.
- M. Arnaudon, A. Thalmaier. "Horizontal Martingales in Vector Bundles." *Séminaire de Probabilits de Strasbourg.* 36 (2002): 419-456.
- K.D. Elworthy, Y. Le Jan, and X. Li. "The Geometry of Filtering." Springer Basel, 2010. 33-59.
- F. Baudoin. "An Introduction to the Geometry of Stochastic Flows." London: Imperial College Press, 2004.



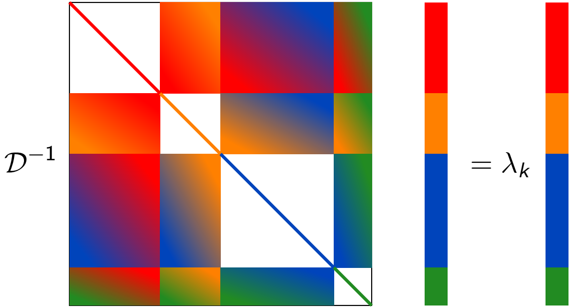




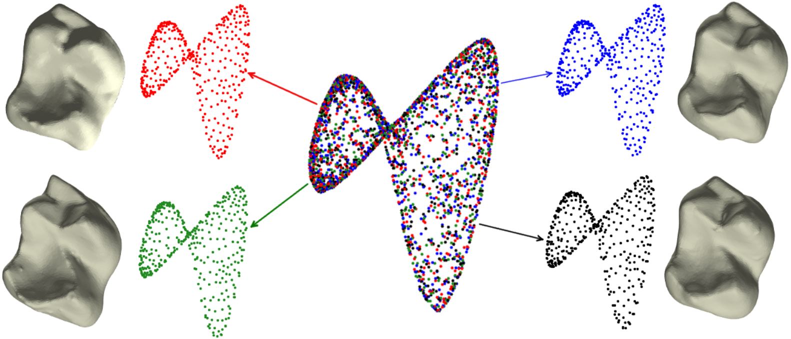
Horizontal Diffusion Maps

Horizontal Diffusion Maps

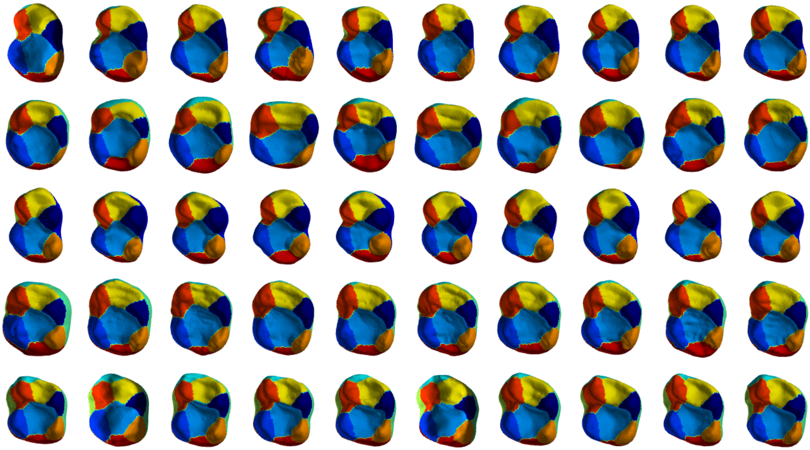
$$\mathcal{D}^{-1}\mathcal{W}u_k = \lambda_k u_k, \quad 1 \leq k \leq \kappa$$



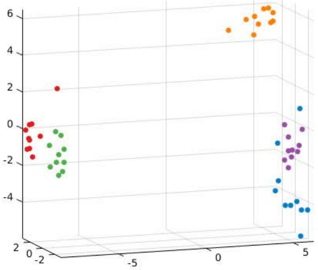
Horizontal Diffusion Maps



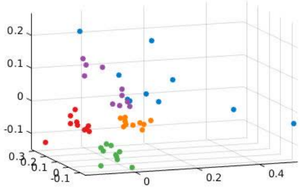
Automatic Landmarking — Interpretability



Species Clustering

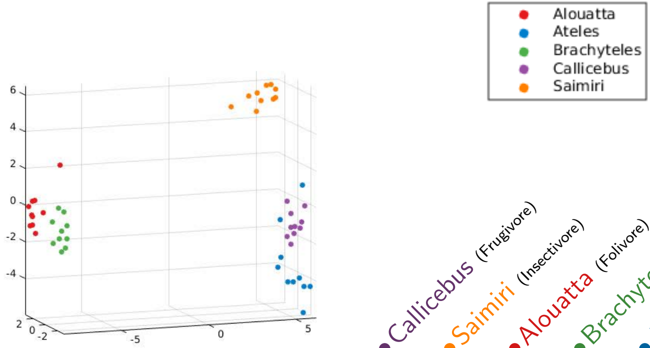


Horizontal Base Diffusion Distance (with Maps)

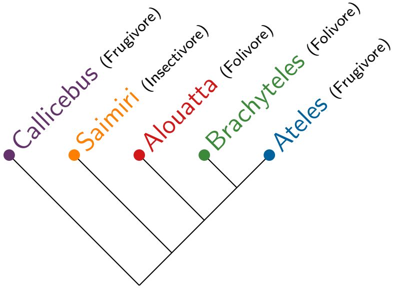


Diffusion Distance (without Maps)

Species Clustering



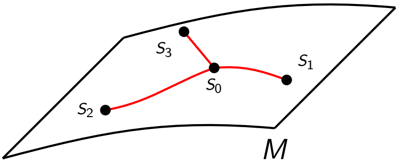
Horizontal Base Diffusion Distance (with Maps)

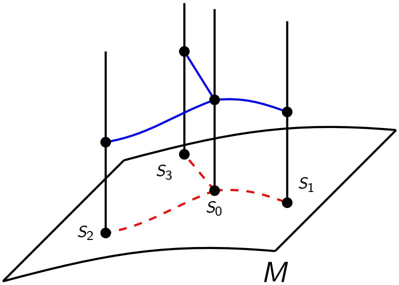


HDM: Mathematical Theory

$$P_\epsilon^{(\alpha)} = \left(D_\epsilon^{(\alpha)}\right)^{-1} W_\epsilon^{(\alpha)}$$

$$H_{\epsilon,\delta}^{(\alpha)} = \left(\mathscr{D}_{\epsilon,\delta}^{(\alpha)}\right)^{-1} \mathscr{W}_{\epsilon,\delta}^{(\alpha)}$$





Asymptotic Theory for Diffusion Maps

Theorem (Belkin-Niyogi 2005). Let data points $x_1, \cdots, x_n$ be sampled from a **uniform** distribution on M. Under mild technical assumptions, there exist a sequence of real numbers $t_n \to 0$ and a constant C such that for any $f \in C^\infty(M)$

$$\lim_{n\to\infty} C \frac{(4\pi t_n)^{-\frac{k+2}{2}}}{n} \frac{P_{t_n} - I}{t_n} f(x) = \Delta_M f(x), \quad \forall x \in M.$$

Theorem (Coifman-Lafon 2006). As $\epsilon \to 0$, for any $f \in C^\infty(M)$ and $x \in M$, if $\{x_i\}_{i=1}^n \sim p(x)\mathrm{dvol}_M(x)$, then w.h.p.

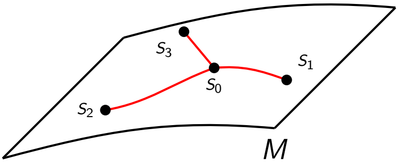
$$P_\epsilon^{(\alpha)} f(x)$$

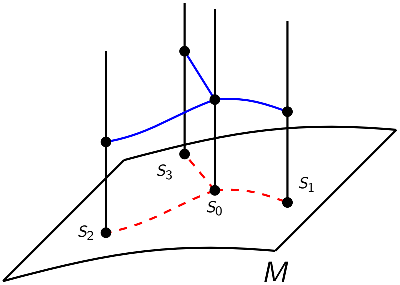
$$= f(x) + \epsilon \frac{m_2}{2m_0} \left[\frac{\Delta_M \left[f p^{1-\alpha} \right](x)}{p^{1-\alpha}(x)} - f(x) \frac{\Delta_M p^{1-\alpha}(x)}{p^{1-\alpha}(x)} \right] + O\left(\epsilon^2\right).$$

HDM: Horizontal Random Walk on a *Fibre Bundle*

$$P_{\epsilon}^{(\alpha)} = \left(D_{\epsilon}^{(\alpha)}\right)^{-1} W_{\epsilon}^{(\alpha)}$$

$$H_{\epsilon,\delta}^{(\alpha)} = \left(\mathscr{D}_{\epsilon,\delta}^{(\alpha)}\right)^{-1} \mathscr{W}_{\epsilon,\delta}^{(\alpha)}$$





Asymptotic Theory for HDM on (E, M, F, π)

Theorem (G. 2016). If $\delta = O(\epsilon)$ as $\epsilon \to 0$, then for any $f \in C^\infty(E)$ and $(x, v) \in E$, as $\epsilon \to 0$,

$$
\begin{aligned}
&H^{(\alpha)}_{\epsilon,\delta} f(x, v) \\
&= f(x, v) + \epsilon \frac{m_{21}}{2m_0} \left[\frac{\Delta_H \left(f p^{1-\alpha} \right)(x, v)}{p^{1-\alpha}(x, v)} - f(x, v) \frac{\Delta_H p^{1-\alpha}(x, v)}{p^{1-\alpha}(x, v)} \right] \\
&\quad + \delta \frac{m_{22}}{2m_0} \left[\frac{\Delta_E^V \left(f p^{1-\alpha} \right)(x, v)}{p^{1-\alpha}(x, v)} - f(x, v) \frac{\Delta_E^V p^{1-\alpha}(x, v)}{p^{1-\alpha}(x, v)} \right] \\
&\quad + O\left(\epsilon^2 + \epsilon\delta + \delta^2 \right).
\end{aligned}
$$

Tingran Gao. *The Diffusion Geometry of Fibre Bundles.* arXiv:1602.02330, 2016

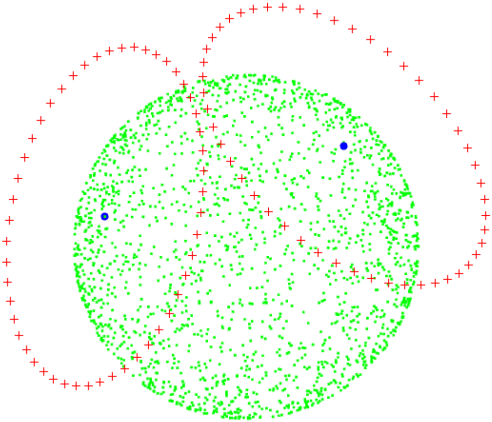
Asymptotic Theory for HDM on (E, M, F, π)

Theorem (G. 2016). If $\delta = O(\epsilon)$ as $\epsilon \to 0$, then for any $f \in C^\infty(E)$ and $(x, v) \in E$, as $\epsilon \to 0$,

$$
\begin{aligned}
&H^{(\alpha)}_{\epsilon,\delta} f(x, v) \\
&= f(x, v) + \epsilon \frac{m_{21}}{2m_0} \left[\frac{\Delta_H \left(fp^{1-\alpha}\right)(x, v)}{p^{1-\alpha}(x, v)} - f(x, v) \frac{\Delta_H p^{1-\alpha}(x, v)}{p^{1-\alpha}(x, v)} \right] \\
&\quad + \delta \frac{m_{22}}{2m_0} \left[\frac{\Delta^V_E \left(fp^{1-\alpha}\right)(x, v)}{p^{1-\alpha}(x, v)} - f(x, v) \frac{\Delta^V_E p^{1-\alpha}(x, v)}{p^{1-\alpha}(x, v)} \right] \\
&\quad + O\left(\epsilon^2 + \epsilon\delta + \delta^2\right).
\end{aligned}
$$

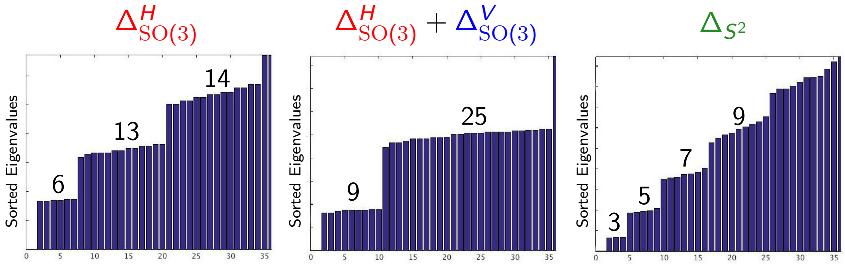
- Δ^V_E is the vertical Laplacian on E
- Δ_H is the Bochner horizontal Laplacian on E
- In general $\Delta_H + \Delta^V_E \neq \Delta_E$, true if and only if π is *harmonic*

HDM on Unit Tangent Bundles: Validation on $\mathrm{SO}(3)$



$\mathrm{SO}(3)$ as the unit tangent bundle of $S^2 \subset \mathbb{R}^3$

HDM on Unit Tangent Bundles: Validation on $\mathrm{SO}(3)$



(a) $\delta = 0.002$ (b) $\delta = 0.015$ (c) $\delta = 20$

Bar plots of the smallest 36 eigenvalues of *horizontal*, *total*, and *base* Laplacians on $\mathrm{SO}(3)$, with fixed $\epsilon = 0.2$ and varying δ

Tingran Gao. *The Diffusion Geometry of Fibre Bundles.* arXiv:1602.02330, 2016

The Convergence Rate: Diffusion Maps

Theorem (Singer 2006). Suppose N points are i.i.d. uniformly sampled from a d-dimensional Riemannian manifold M. The graph diffusion operator $P_{\epsilon,\alpha}$ converges to its smooth limit at rate

$$O\left(N^{-\frac{1}{2}}\epsilon^{\frac{1}{2}-\frac{d}{4}}\right).$$

Corollary. Under the same assumption, non-uniform sampling has convergence rate

$$O\left(N^{-\frac{1}{2}}\epsilon^{-\frac{d}{4}}\right).$$

The Convergence Rate: HDM on Unit Tangent Bundles

Theorem (G. 2016). Suppose N_B points are i.i.d. sampled from a d-dimensional Riemannian manifold M, and N_F unit tangent vectors are i.i.d. sampled at each of the N_B samples. The graph horizontal diffusion operator $H^{\alpha}_{\epsilon,\delta}$ converges to its smooth limit at rate

$$O\left(\theta_*^{-1} N_B^{-\frac{1}{2}} \epsilon^{-\frac{d}{4}}\right),$$

where

$$\theta_* = 1 - \frac{1}{1 + \epsilon^{\frac{d}{4}} \delta^{\frac{d-1}{4}} \sqrt{\frac{N_F}{N_B}}}.$$

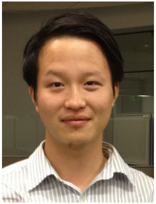
Tingran Gao. "The Diffusion Geometry of Fibre Bundles." *submitted.* arXiv:1602.02330, 2016

Collaborators

Rima Alaifari
ETH Zürich

Doug Boyer
Duke

Ingrid Daubechies
Duke



Tingran Gao
Duke

Yaron Lipman
Weizmann

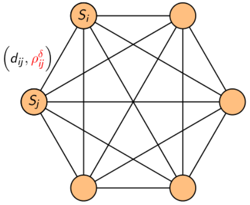
Roi Poranne
ETH Zürich

Jesús Puente
J.P. Morgan

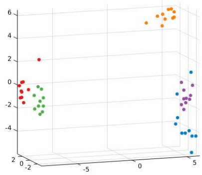
Robert Ravier
Duke

Thank You!

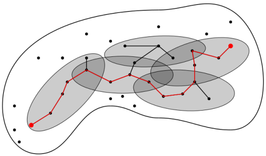
Augmented Distance Graph



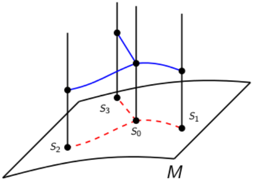
Horizontal Base Diffusion Distance



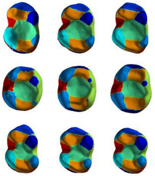
"Knit Together" Local Geometry



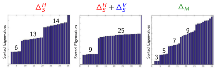
Random Walk on a Fibre Bundle



Biological Application



Mathematical Theory



$$O\left(\theta_*^{-1} N_B^{-\frac{1}{2}} \epsilon^{-\frac{d}{4}}\right),$$

$$\theta_* = 1 - \frac{1}{1+\epsilon^{\frac{d}{4}}\delta^{\frac{d-1}{4}}\sqrt{\frac{N_F}{N_B}}}.$$

Tingran Gao. "The Diffusion Geometry of Fibre Bundles." *submitted.* arXiv:1602.02330, 2016

Tingran Gao, Jacek Brodzki, Sayan Mukherjee. "The Geometry of Synchronization Problems and Learning Group Actions." *submitted.* arXiv:1610.09051, 2016