Synchronization: Geometry Meets Learning

Tingran Gao

Committee on Computational and Applied Mathematics (CCAM)
Department of Statistics
The University of Chicago

Dynamics, Topology, and Computations
Będlewo, Poland

Friday June 22, 2018

Background
- Evolutionary Biology
- Collection Shape Analysis
- Cryo-Electron Microscopy

Graph Synchronization Problems: Principal Bundle Perspective
- Representation Variety
- Discrete Twisted Hodge Theory

Learning Aspects
- Horizontal Diffusion Maps
- Learning Group Actions

Simultaneous Matching and Clustering (SMAC)

Doug Boyer
Duke

Ingrid Daubechies
Duke

Sayan Mukherjee
Duke

Jacek Brodzki
Southamption

Qixing Huang
UT Austin

Chandrajit Bajaj
UT Austin

Outline

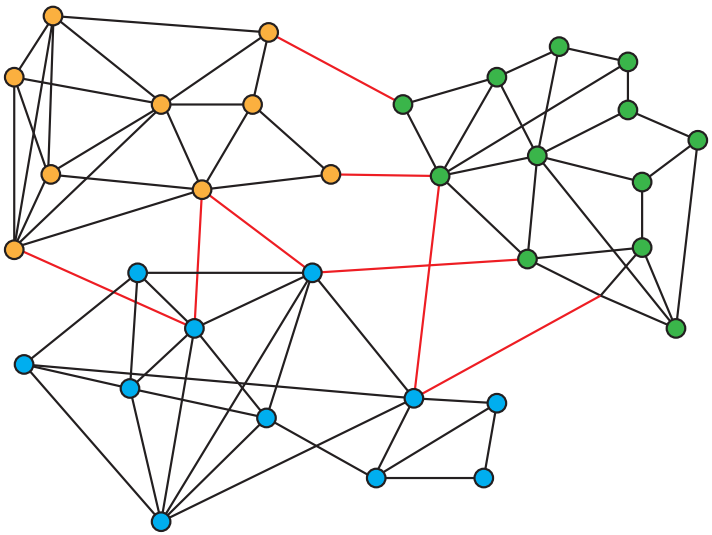
Background
- Evolutionary Biology
- Collection Shape Analysis
- Cryo-Electron Microscopy

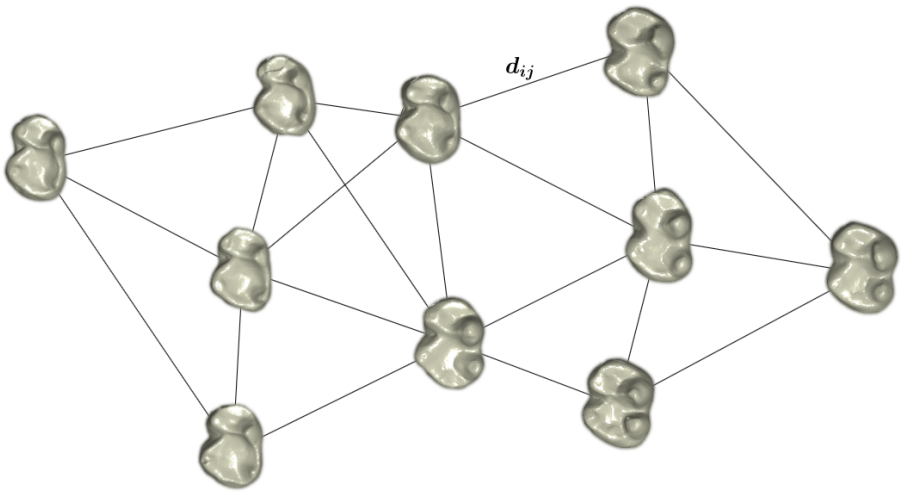
Graph Synchronization Problems: Principal Bundle Perspective
- Representation Variety
- Discrete Twisted Hodge Theory

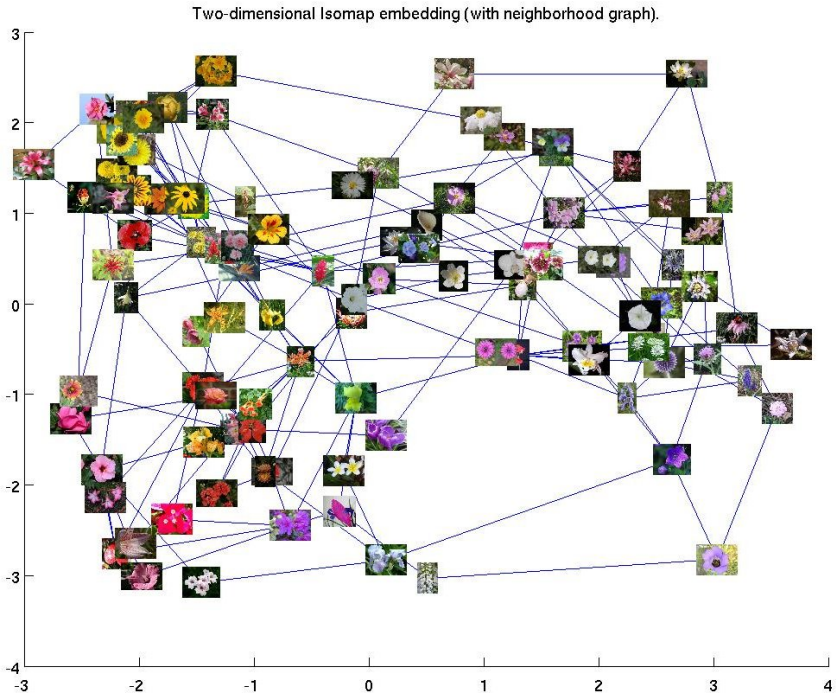
Learning Aspects
- Horizontal Diffusion Maps
- Learning Group Actions

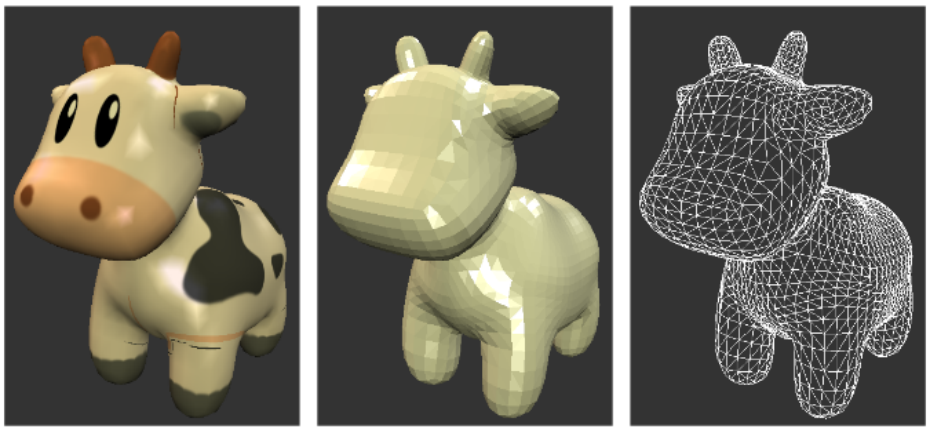
Simultaneous Matching and Clustering (SMAC)

(Weighted) Graph-Based Data Analysis





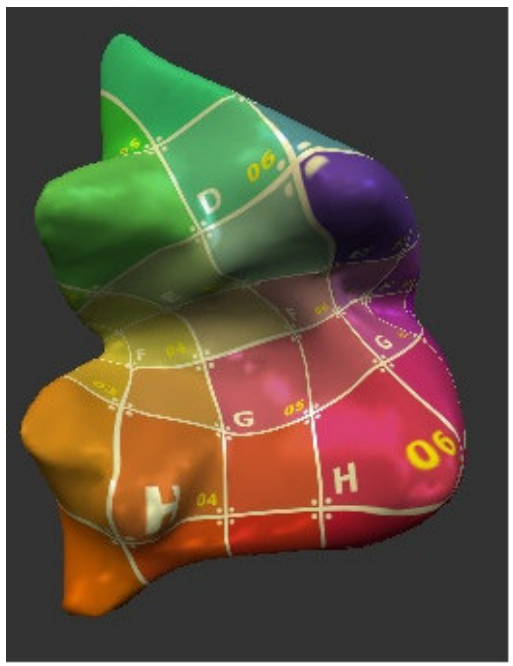




Evolutionary Biology

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

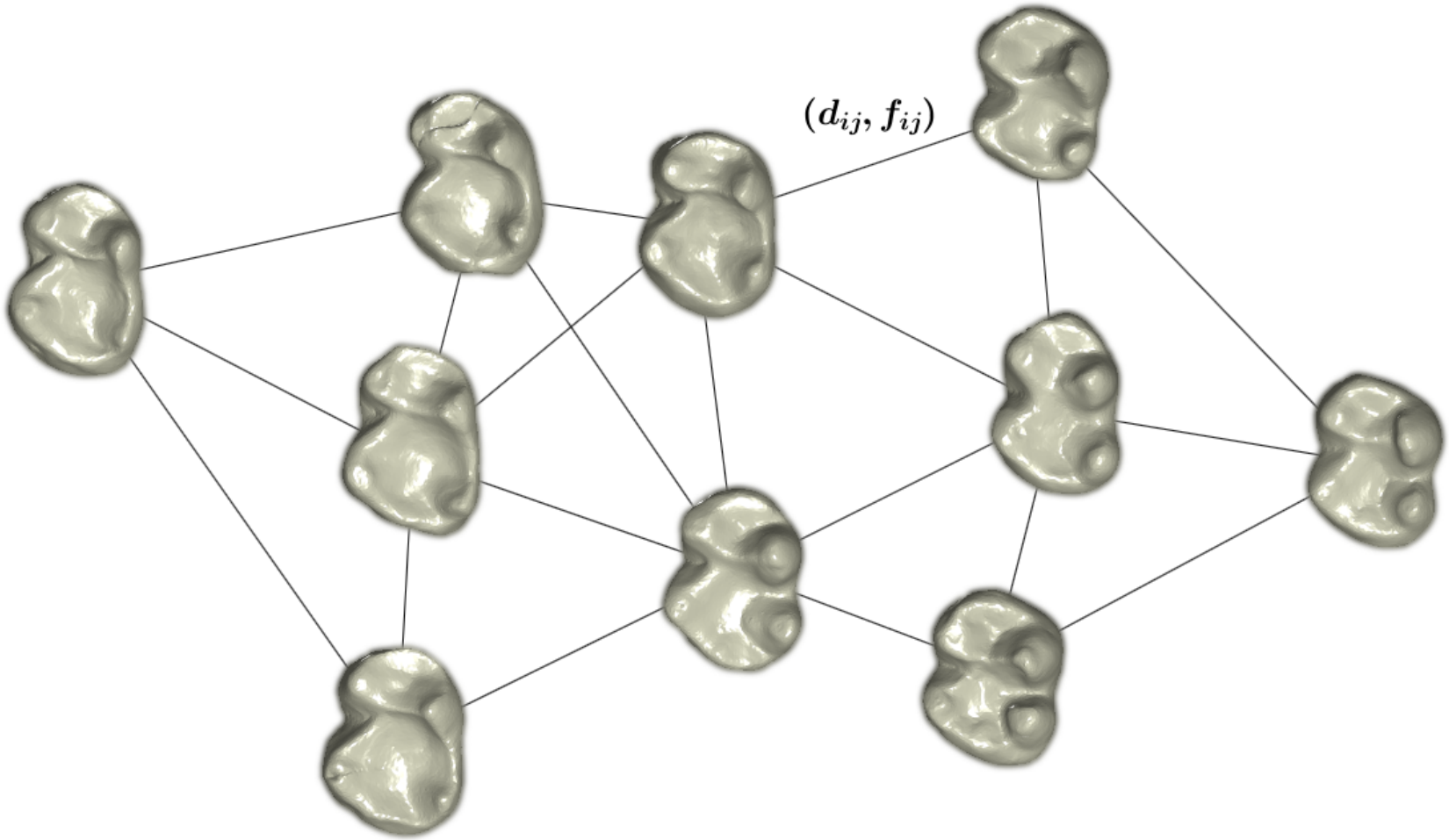
S_i



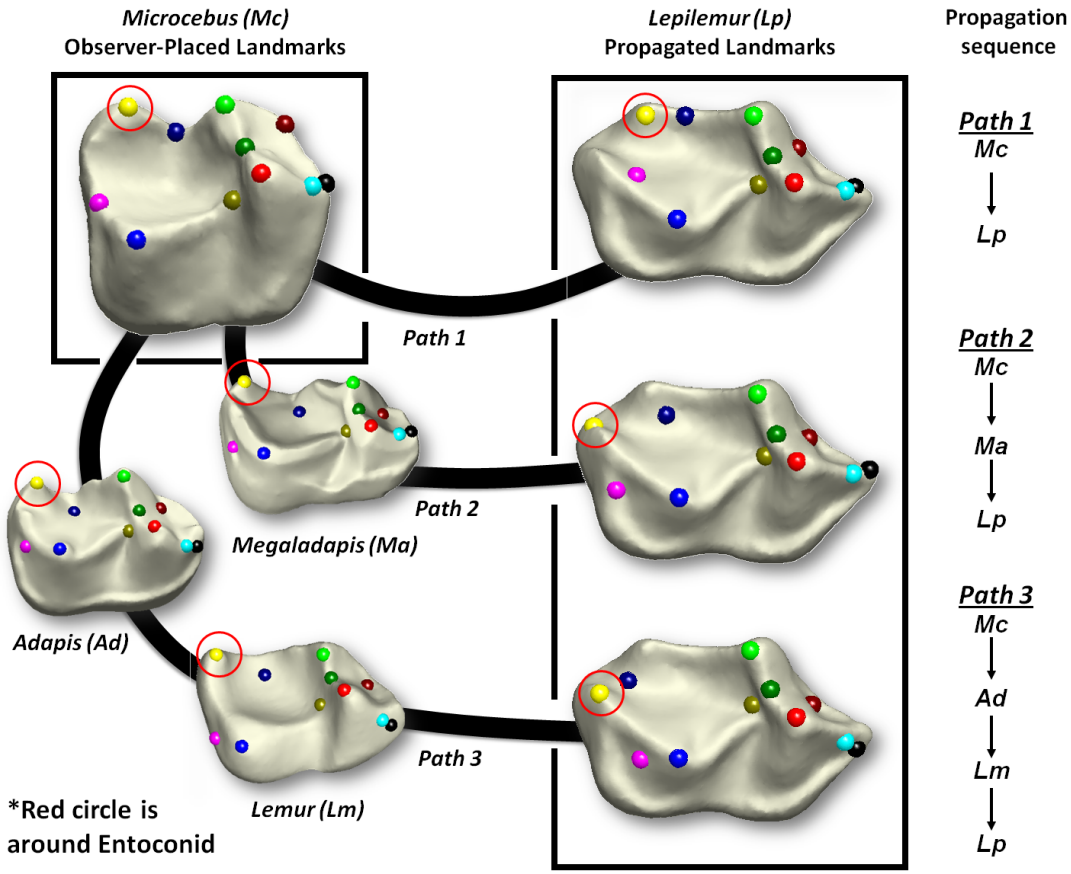
d_{ij}

f_{ij}

S_j

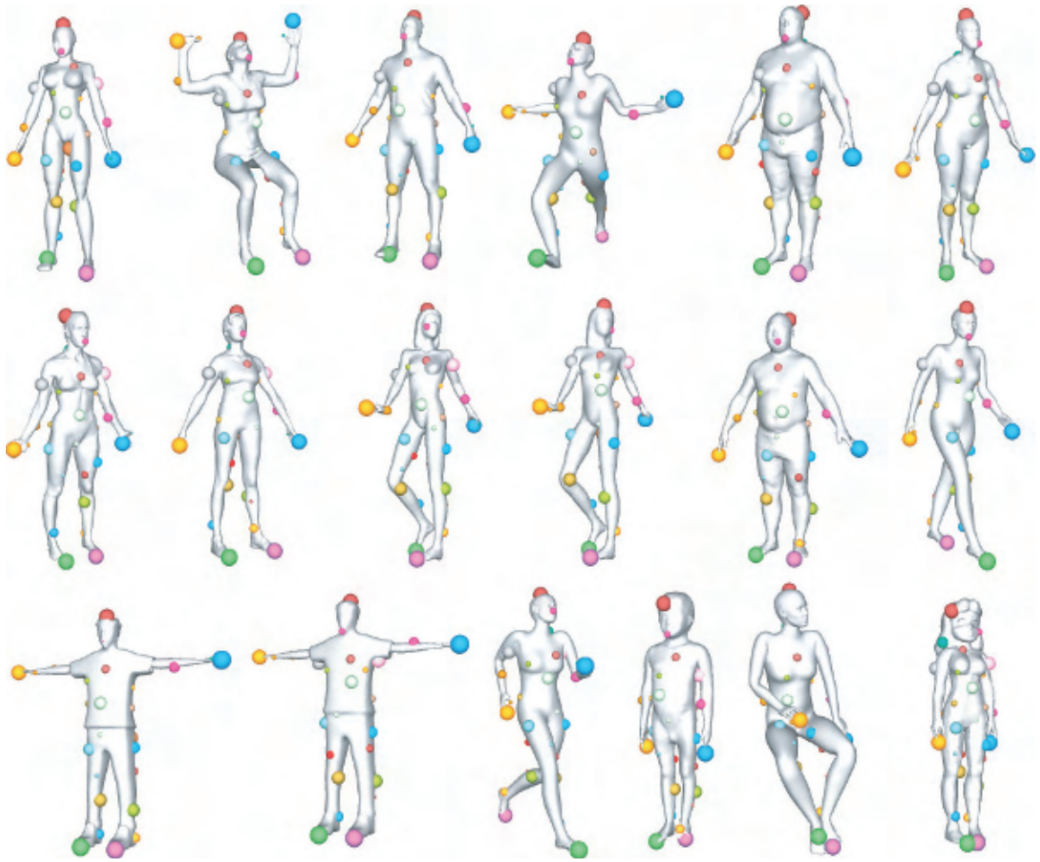

(d_{ij}, f_{ij})

Interpretability Issue



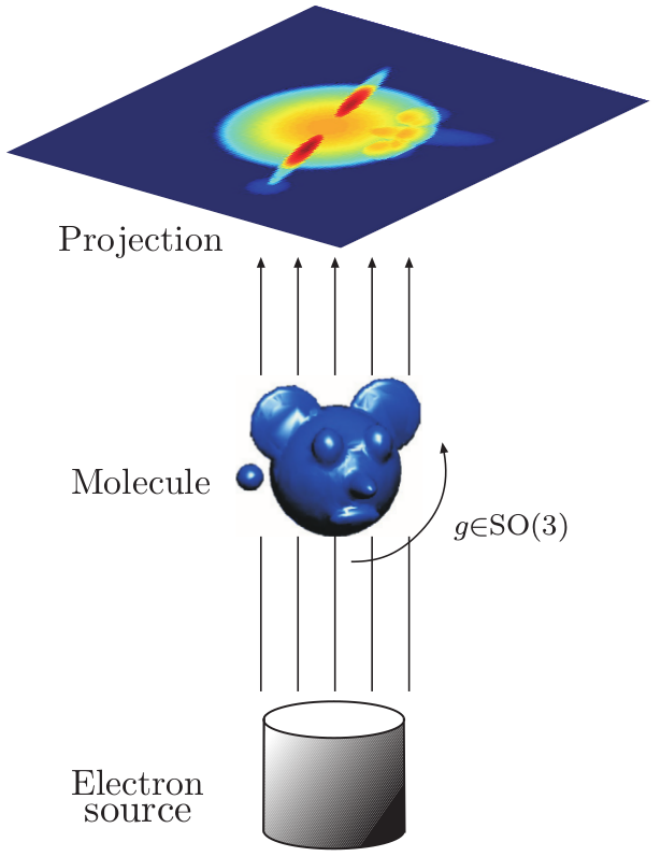
Boyer et al. Ąlgorithms to Automatically Quantify the Geometric Similarity of Anatomical Surfaces." *Proceedings of the National Academy of Sciences* 108.45 (2011): 18221-18226.

Collection Shape Analysis



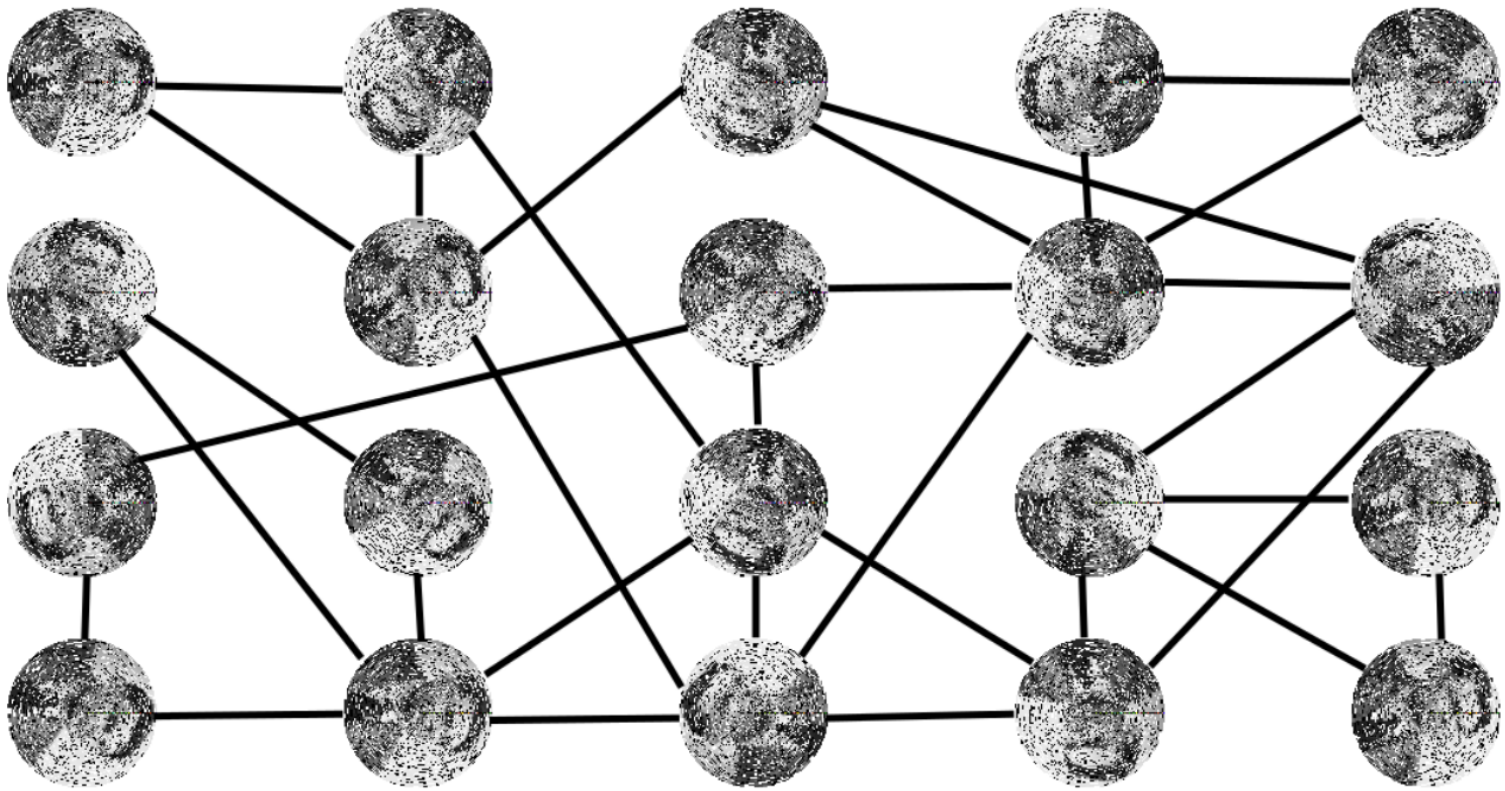
Q. Huang, L. Guibas. "Consistent Shape Maps via Semidefinite Programming", *Computer Graphics Forum*, vol. 32, no. 5, pp. 177-186 (2013).

Cryo-Electron Microscopy



A. Singer, Z. Zhao, Y. Shkolnisky, R. Hadani. "Viewing Angle Classification of Cryo-Electron Microscopy Images using Eigenvectors", *SIAM Journal on Imaging Sciences*, 4 (2), pp. 543-572 (2011).

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



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

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

$$\frac{1}{2} \sum_{1 \leq i,j \leq N} w_{ij} \left\| R_{ij} - R_i^\top R_j \right\|_{\mathrm{F}}^2$$

$$
\begin{aligned}
&\frac{1}{2}\sum_{1\le i,j\le N} w_{ij}\left\|R_{ij}-R_i^\top R_j\right\|_{\mathrm{F}}^2 \\
&\quad=\frac{1}{2}\sum_{1\le i,j\le N} w_{ij}\mathrm{Tr}\left[\left(R_i^\top-R_{ij}R_j^\top\right)\left(R_i-R_jR_{ij}^\top\right)\right] \\
&\quad=\sum_{1\le i,j\le N} w_{ij}\mathrm{Tr}\left(I_{d\times d}-R_iR_{ij}R_j^\top\right)
\end{aligned}
$$

$$
\begin{aligned}
&\frac{1}{2}\sum_{1\le i,j\le N} w_{ij}\left\|R_{ij}-R_i^\top R_j\right\|_{\mathrm{F}}^2 \\
&\quad=\frac{1}{2}\sum_{1\le i,j\le N} w_{ij}\mathrm{Tr}\left[\left(R_i^\top-R_{ij}R_j^\top\right)\left(R_i-R_jR_{ij}^\top\right)\right] \\
&\quad=\sum_{1\le i,j\le N} w_{ij}\mathrm{Tr}\left(I_{d\times d}-R_iR_{ij}R_j^\top\right) \\
&\quad=\mathrm{Tr}\left(\left[R_1,\cdots,R_N\right]\left(\mathscr{D}-\mathscr{W}\right)\begin{bmatrix}R_1^\top\\ \vdots\\ R_N^\top\end{bmatrix}\right)
\end{aligned}
$$

where

$$
\mathscr{D}=\begin{bmatrix}d_1I_{d\times d} & & \\ & \ddots & \\ & & d_NI_{d\times d}\end{bmatrix},\mathscr{W}=\begin{bmatrix}w_{11}R_{11} & \cdots & w_{1N}R_{1N}\\ \vdots & & \vdots\\ w_{N1}R_{N1} & \cdots & w_{NN}R_{NN}\end{bmatrix}
$$

QO-OC

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

$$\Leftrightarrow \min_{R_1,\cdots,R_N\in\mathbb{R}^{d\times d}} \mathrm{Tr}\left(\left[R_1,\cdots,R_N\right] \left(\mathscr{D}-\mathscr{W}\right) \begin{bmatrix} R_1^\top \\ \vdots \\ R_N^\top \end{bmatrix} \right)$$

$$\text{s.t.} \qquad R_1^\top R_1 = \cdots = R_N^\top R_N = I_{d\times d}$$

A. Nemirovski, "Sums of Random Symmetric Matrices and Quadratic Optimization Under Orthogonality Constraints", *Math. Programming*, 109 (2007), pp. 283-317.

Convex Relaxations (I)

$$\min_{R_1,\cdots,R_N \in \mathbb{R}^{d\times d}} \operatorname{Tr}\left(\begin{bmatrix} R_1, \cdots, R_N \end{bmatrix} (\mathscr{D} - \mathscr{W}) \begin{bmatrix} R_1^\top \\ \vdots \\ R_N^\top \end{bmatrix} \right)$$
$$\text{s.t.} \qquad R_1^\top R_1 = \cdots = R_N^\top R_N = I_{d\times d}$$

- Spectral Relaxation

$$\boxed{\begin{aligned} &\min_{\mathcal{R}\in\mathbb{R}^{d\times Nd}} \operatorname{Tr}\left(\mathcal{R}\left(\mathscr{D} - \mathscr{W}\right)\mathcal{R}^\top\right) \\ &\text{s.t.} \qquad \mathcal{R}\mathcal{R}^\top = N \cdot I_{d\times d} \end{aligned}}$$

Kunal N. Chaudhury et al. "Global Registration of Multiple Point Clouds using Semidefinite Programming", *SIAM Journal on Optimization*, 25(1):468-501, 2015.

Convex Relaxations (II)

$$\min_{R_1,\cdots,R_N \in \mathbb{R}^{d\times d}} \operatorname{Tr}\left((\mathscr{D} - \mathscr{W}) \begin{bmatrix} R_1^\top \\ \vdots \\ R_N^\top \end{bmatrix} \begin{bmatrix} R_1, \cdots, R_N \end{bmatrix} \right)$$
$$\text{s.t.} \qquad R_1^\top R_1 = \cdots = R_N^\top R_N = I_{d\times d}$$

-

$$\begin{aligned} \min_{\mathcal{G}\in\mathbb{R}^{Nd\times Nd}} \quad & \operatorname{Tr}\left((\mathscr{D} - \mathscr{W})\,\mathcal{G}\right) \\ \text{s.t.} \quad & \mathcal{G} \succeq 0,\ \mathcal{G} = \mathcal{G}^\top \\ & \mathcal{G}_{ii} = I_{d\times d},\ 1 \le i \le N \\ & \operatorname{rank}(\mathcal{G}) = d \end{aligned}$$

Kunal N. Chaudhury et al. "Global Registration of Multiple Point Clouds using Semidefinite Programming", *SIAM Journal on Optimization*, 25(1):468-501, 2015.

Convex Relaxations (II)

$$\min_{R_1,\cdots,R_N \in \mathbb{R}^{d\times d}} \operatorname{Tr}\left((\mathscr{D} - \mathscr{W}) \begin{bmatrix} R_1^\top \\ \vdots \\ R_N^\top \end{bmatrix} \begin{bmatrix} R_1, \cdots, R_N \end{bmatrix} \right)$$
$$\text{s.t.} \qquad R_1^\top R_1 = \cdots = R_N^\top R_N = I_{d\times d}$$

- SDP Relaxation

$$\min_{\mathcal{G} \in \mathbb{R}^{Nd\times Nd}} \operatorname{Tr}\left((\mathscr{D} - \mathscr{W}) \mathcal{G} \right)$$
$$\text{s.t.} \qquad \mathcal{G} \succeq 0, \ \mathcal{G} = \mathcal{G}^\top$$
$$\mathcal{G}_{ii} = I_{d\times d}, \ 1 \leq i \leq N$$
$$\cancel{\operatorname{rank}(\mathcal{G}) = d}$$

Kunal N. Chaudhury et al. "Global Registration of Multiple Point Clouds using Semidefinite Programming", *SIAM Journal on Optimization*, 25(1):468-501, 2015.

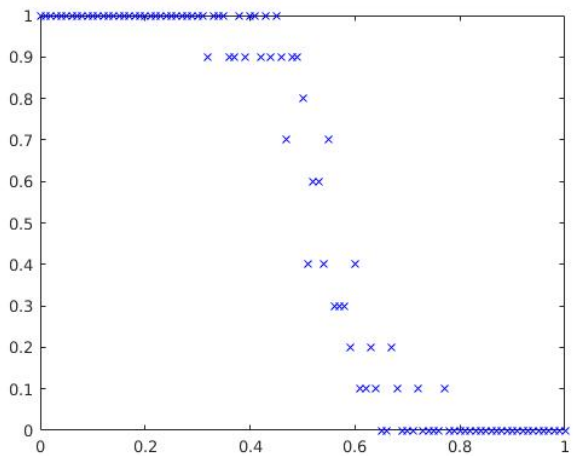
Tightness of SDP Relaxation?

$$A_i = R_i A + \sigma \eta_i, \quad \eta \sim \text{i.i.d. Standard Gaussian}$$

$$C \in \mathbb{R}^{Nd \times Nd}, \quad C_{ij} = A_i A_j^\top$$

$$\begin{aligned} \min_{\mathcal{G} \in \mathbb{R}^{Nd \times Nd}} \quad & \operatorname{Tr}(C\mathcal{G}) \\ \text{s.t.} \quad & \mathcal{G} \succeq 0, \ \mathcal{G} = \mathcal{G}^\top \\ & \mathcal{G}_{ii} = I_{d \times d}, \ 1 \leq i \leq N \end{aligned}$$

- Rank Recovery: Empirically, solution has rank d with high probability for noise level below a certain threshold!



Outline

Background
- Evolutionary Biology
- Collection Shape Analysis
- Cryo-Electron Microscopy

Graph Synchronization Problems: Principal Bundle Perspective
- Representation Variety
- Discrete Twisted Hodge Theory

Learning Aspects
- Horizontal Diffusion Maps
- Learning Group Actions

Simultaneous Matching and Clustering (SMAC)

Synchronization Problems

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

Synchronization Problems

- Data:
  - graph $\Gamma = (V, E)$
  - topological 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

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

Synchronization Problems

- Data:
  - graph $\Gamma = (V, E)$
  - topological 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 $\rho_{ij} = f_i f_j^{-1}$

Synchronization Problems

- Data:
  - graph $\Gamma = (V, E)$
  - topological 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!

Synchronization Problems

- Data:
  - graph $\Gamma = (V, E)$
  - topological 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

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

Synchronization Problems

- 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 variant: G acts on a normed vector space $(F,\|\cdot\|)$,

$$\eta(f)=\frac{1}{2}\frac{\sum_{i,j\in V}\|f_i-\rho_{ij}f_j\|^2}{\sum_{i\in V}\|f_i\|^2}=\frac{\langle f,L_1f\rangle}{\langle f,f\rangle}$$

where L_1 is the *Graph Connection Laplacian* (GCL)

$$(L_1f)_i=\frac{1}{\deg(i)}\sum_{j:(i,j)\in E}(f_i-\rho_{ij}f_j)$$

Synchronization Problems: Examples

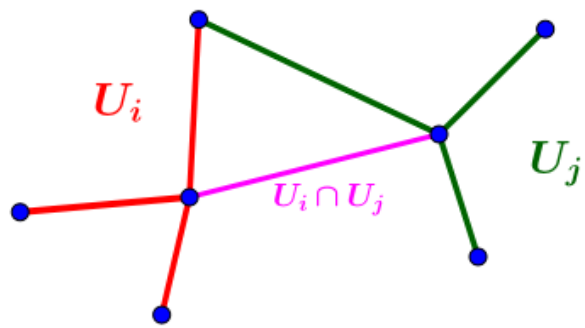
- Manifold Orientability [Singer, Wu (2011)]: $G = O(1)$
- Angular Synchronization [Singer (2011)]: $G = U(1)$
- Vector Diffusion Maps [Singer, Wu (2012)]: $G = O(d)$
- Multireference Alignment [Bandeira et al. (2014)]: $G = \{\text{cyclic shifts}\}$
- Global Registration of Point Clouds [Chaudhury (2015)]: $G = \mathbb{E}_d$
- Collection Shape Matching [Nguyen et al. (2011)], [Huang, Guibas (2013)], [Chen et al. (2014)], [Maron et al. (2016)]: $G = \text{permutation group}$
- Cryo-EM Structural Reconstruction [Singer et al. (2011)], [Shkolnisky, Singer (2012)], [Zhao, Singer (2014)], [Bandeira et al. (2015)]: $G = SO(3)$

The Geometry of Synchronization Problems

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

The Geometry of Synchronization Problems

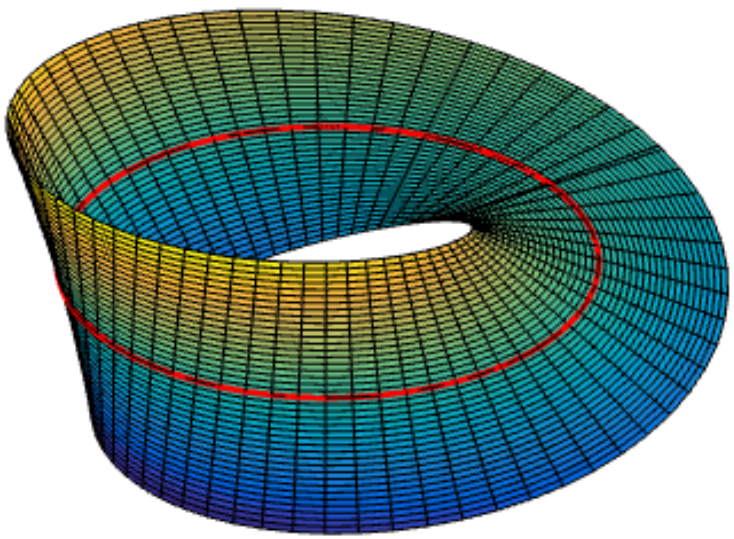
- Data:
  - graph $\Gamma = (V, E)$
  - topological group G, equipped with a norm $\|\cdot\|$
  - **edge potential** $\rho : E \to G$ satisfying $\rho_{ij} = \rho_{ji}^{-1}, \ \forall (i,j) \in E$
- Sheaf:
  - Let $\mathfrak{U} = \{U_i \mid 1 \leq i \leq |V|\}$ be an open cover of Γ (viewed as a 1-dimensional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

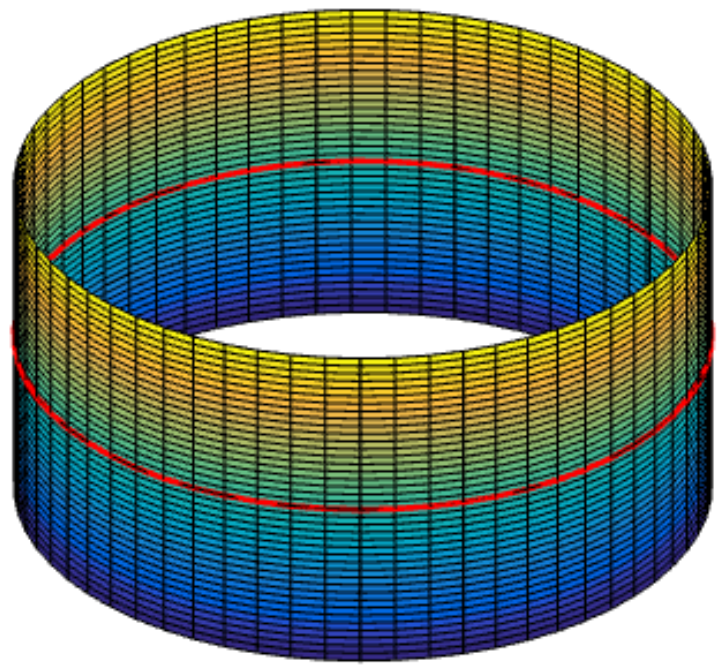


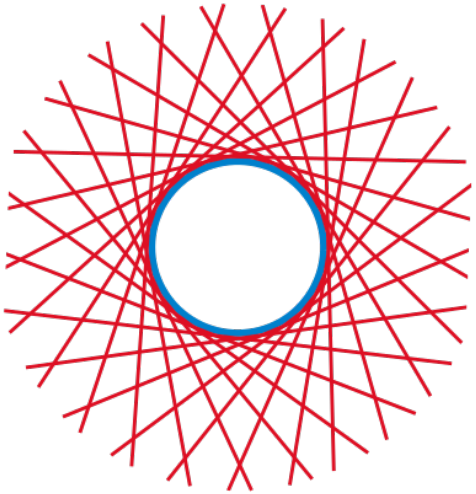
  - Triplet (ρ, G, Γ) defines a *flat principal G-bundle* $\mathscr{B}_\rho$ over Γ

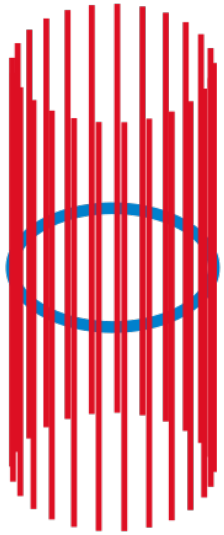
Fibre Bundle $\mathscr{E} = (E, M, F, \pi)$

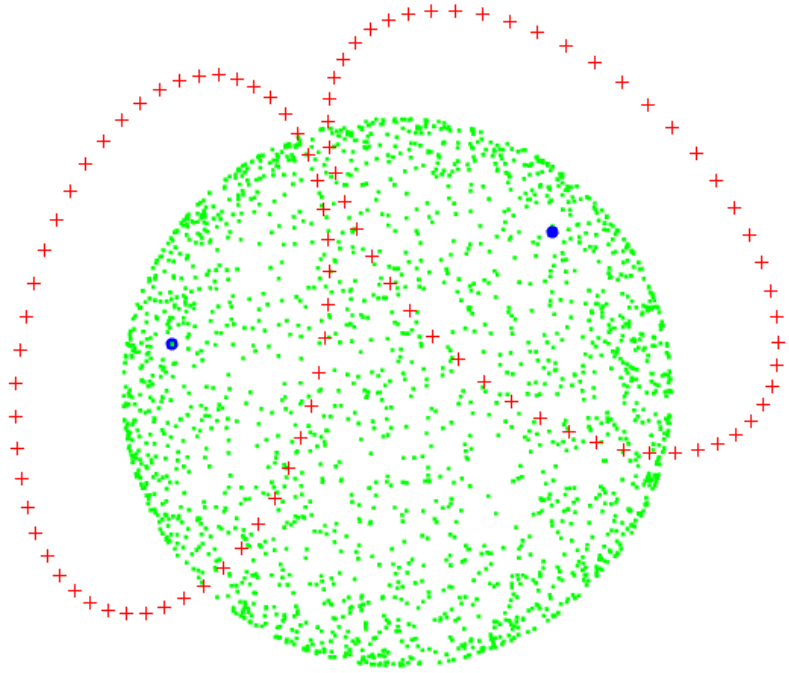
- M: *base* manifold
- F: *fibre* manifold
- E: *total* manifold
- $\pi : E \to M$: smooth surjective map (*bundle projection*)
- *local triviality*: for "small" open set $U \subset M$, $\pi^{-1}(U)$ is diffeomorphic to $U \times F$

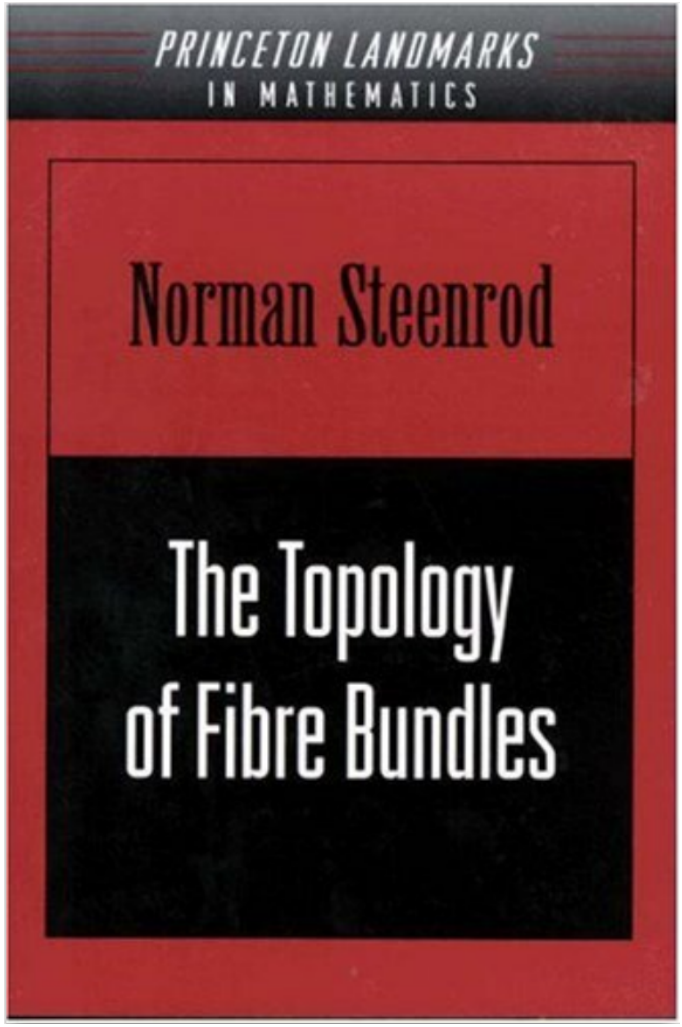








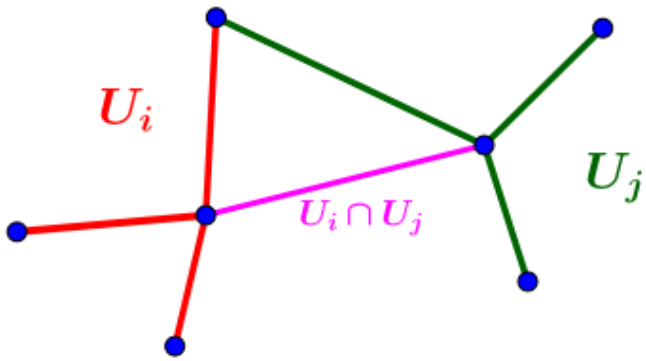




Theorem (Steenrod 1951, §2). If topological group G acts on Y and $\{U_i\}$, $\{g_{ij}\}$ is a system of coordinate transformations in the space X such that

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

then there exists a fibre bundle $\mathscr{B}$ with base space X, fibre Y, group G, and coordinate transforms $\{g_{ij}\}$.



No triple intersections!

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

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

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

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

then there exists a fibre bundle $\mathscr{B}$ with base space X, fibre Y, group G, and coordinate transforms $\{g_{ij}\}$.

Geometric Observations

- Denote

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

- 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 $(\tau_f \rho)_{ij} := f_i^{-1} \rho_{ij} f_j, \quad \forall (i,j) \in E$. Denote the orbit space of τ_f by $C^1(\Gamma; G)/C^0(\Gamma; G)$

- ρ synchronizable $\Leftrightarrow$ $\tau_f \rho$ synchronizable for all $f \in C^0(\Gamma; G)$, i.e. synchronizability is defined up to orbits of τ_f

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:

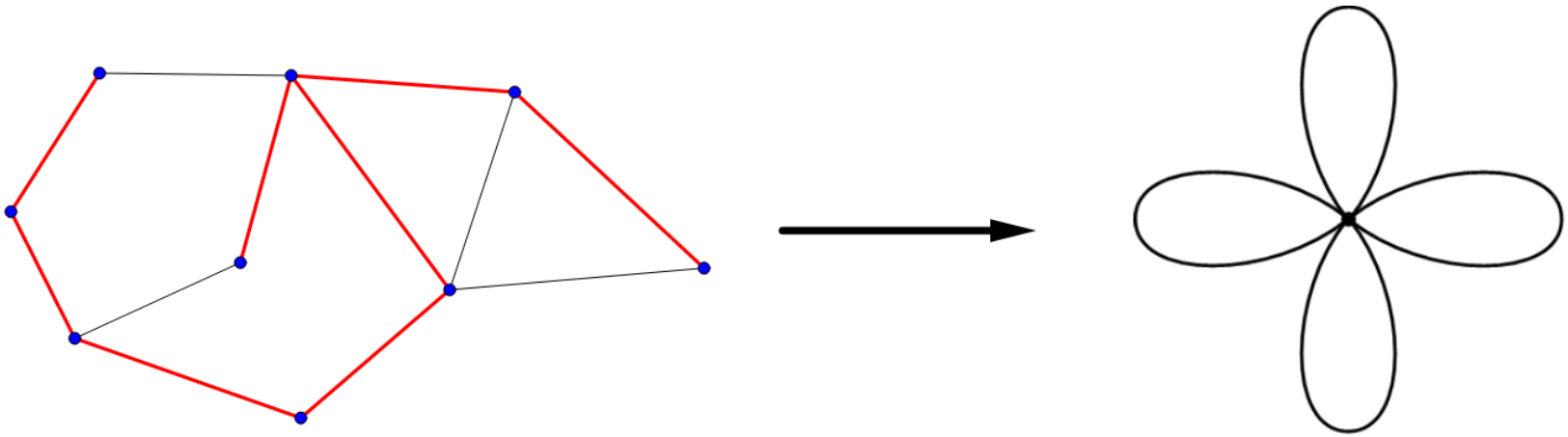
$$\begin{aligned} \mathrm{Hom}(\pi_1(\Gamma), G) \times G &\longrightarrow \mathrm{Hom}(\pi_1(\Gamma), G) \\ (\phi, g) &\longmapsto g^{-1}\phi g \end{aligned}$$

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

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

- **LHS**: The *first cohomology set* $\check{H}^1((\Gamma, \mathfrak{U}), \underline{G})$
- **RHS**: The *representation variety* of $\pi_1(\Gamma)$ with value in G
- Proof builds upon construction of a *holonomy homomorphism*
- In particular, $\rho \in C^1(\Gamma; G)$ synchronizable
  $\Leftrightarrow [\rho] = [e]$ as equivalence classes in $C^1(\Gamma; G)/C^0(\Gamma; G)$
  $\Leftrightarrow$ the principal G-bundle $\mathscr{B}_\rho$ is trivial
- On-going work: study synchronization problems through the geometry of the moduli space/character variety

Fundamental Group of a Graph?



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

$$= \text{free group of rank } (|E| - |V| + 1)$$

Twisted De Rham Cohomology

- From now on, let F be a vector space such that $G \subset GL(F)$, and consider only $f \in C^0(\Gamma; F)$

Twisted De Rham Cohomology

- From now on, let F be a vector space such that $G \subset GL(F)$, and consider only $f \in C^0(\Gamma; F)$
- **Motivation**: Write

$$\eta(f) = \frac{1}{2} \frac{\sum_{i,j \in V} \|f_i - \rho_{ij} f_j\|^2}{\sum_{i \in V} \|f_i\|^2}$$

into a Rayleigh quotient:

$$\eta(f) = \frac{\langle f, L_1 f \rangle}{\langle f, f \rangle}$$

where L_1 is the *Graph Connection Laplacian* (GCL)

$$(L_1 f)_i = \frac{1}{\deg(i)} \sum_{j:(i,j) \in E} (f_i - \rho_{ij} f_j)$$

Graph Connection Laplacian

Hau-tieng Wu, M.D., Ph.D. (Duke)

- A. Singer, H.-T. Wu, "Vector Diffusion Maps and the Connection Laplacian," *Communications in Pure and Applied Mathematics*, 65 (8): 1067-1144, 2012
- N. El Karoui, H.-T. Wu, "Graph Connection Laplacian and Random Matrices with Random Blocks," *Information and Inference*, 4 (1): 1-44, 2015
- N. El Karoui, H.-T. Wu, "Graph Connection Laplacian Methods can Be Made Robust to Noise," *Annals of Statistics*, 44 (1), 346-372, 2016
- Afonso S. Bandeira, Amit Singer, and Daniel A. Spielman, "A Cheeger Inequality for the Graph Connection Laplacian," *SIAM. J. Matrix Anal. & Appl.*, 34(4), 1611âĂŞ1630.
- A. Singer, H.-T. Wu, "Spectral Convergence of the Connection Laplacian From Random Samples," *Information and Inference*, in press, 2016.

Question: Can GCL be realized as a Hodge Laplacian in a De Rham-Hodge theory?

Why De Rham-Hodge for GCL?

- **Motivation:** Combinatorial Hodge theory does realize the graph Laplacian as a Hodge Laplacian in a de Rham-Hodge theory associated with any weighted graph $\Gamma = (V, E)$:

$$\Omega^0(\Gamma) := \{f : V \to \mathbb{K}\},$$
$$\Omega^1(\Gamma) := \{\omega : E \to \mathbb{K} \mid \omega_{ij} = -\omega_{ji} \ \forall (i,j) \in E\},$$

Define cochain complex

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

where

$$(df)_{ij} = f_i - f_j, \quad \forall f \in \Omega^0(\Gamma),$$
$$(\delta\omega)_i = \frac{1}{\deg(i)} \sum_{j:(i,j)\in E} \omega_{ij}, \quad \forall \omega \in \Omega^1(\Gamma).$$

Why De Rham-Hodge for GCL?

Define cochain complex

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

where

$$(df)_{ij} = f_i - f_j, \quad \forall f \in \Omega^0(\Gamma),$$
$$(\delta\omega)_i = \frac{1}{\deg(i)} \sum_{j:(i,j)\in E} \omega_{ij}, \quad \forall \omega \in \Omega^1(\Gamma).$$

Then

$$(L_0 f)_i := (\delta d f)_i$$
$$= \frac{1}{\deg(i)} \sum_{j:(i,j)\in E} (f_i - f_j) \quad \forall i \in V, \ \forall f \in \Omega^0(\Gamma).$$

Twisted De Rham-Hodge Theory

$$(L_0 f)_i = \frac{1}{\deg(i)} \sum_{j:(i,j)\in E} (f_i - f_j), \quad \forall f : V \to \mathbb{K}$$

$$(L_1 f)_i = \frac{1}{\deg(i)} \sum_{j:(i,j)\in E} (f_i - \rho_{ij} f_j) \quad \forall f : V \to F$$

Twisted De Rham-Hodge Theory

$$(L_0 f)_i = \frac{1}{\deg(i)} \sum_{j:(i,j)\in E} (f_i - f_j), \quad \forall f : V \to \mathbb{K}$$

$$(L_1 f)_i = \frac{1}{\deg(i)} \sum_{j:(i,j)\in E} (f_i - \rho_{ij} f_j) \quad \forall f : V \to F$$

Naïvely:

$$(d_\rho f)_{ij} \sim f_i - \rho_{ij} f_j, \quad \forall f \in C^0(\Gamma; F)$$

$$(\delta_\rho \omega)_i \sim \frac{1}{\deg(i)} \sum_{j:(i,j)\in E} \omega_{ij}, \quad \forall \omega \in C^1(\Gamma; F)$$

then $L_1 = \delta_\rho d_\rho$. **What's wrong with this?**

$$(d_\rho f)_{ij} \sim f_i - \rho_{ij} f_j, \quad \forall f \in C^0(\Gamma; F)$$

$$(\delta_\rho \omega)_i \sim \frac{1}{\deg(i)} \sum_{j:(i,j)\in E} \omega_{ij}, \quad \forall \omega \in C^1(\Gamma; F)$$

Issue: d_ρ does not map into $C^1(\Gamma; F)$! (no skew-symmetry)

$$f_j - \rho_{ji} f_j = -\rho_{ji}(f_i - \rho_{ij} f_j) \neq -(f_i - \rho_{ij} f_j).$$

$$(d_\rho f)_{ij} \sim f_i - \rho_{ij} f_j, \quad \forall f \in C^0(\Gamma; F)$$

$$(\delta_\rho \omega)_i \sim \frac{1}{\deg(i)} \sum_{j:(i,j)\in E} \omega_{ij}, \quad \forall \omega \in C^1(\Gamma; F)$$

Issue: d_ρ does not map into $C^1(\Gamma; F)$! (no skew-symmetry)

$$f_j - \rho_{ji} f_j = -\rho_{ji}(f_i - \rho_{ij} f_j) \neq -(f_i - \rho_{ij} f_j).$$

Fix: Interpret $f_i - \rho_{ij} f_j$ as the "local expression" of $(d_\rho f)_{ij}$ in a local trivialization over $\mathfrak{U} = \{U_i \mid 1 \leq i \leq |V|\}$ of the associated F-bundle of $\mathscr{B}_\rho$, denoted as $\mathscr{B}_\rho[F]$, such that the extra ρ_{ji} factor encodes a bundle transformation from U_i to U_j.

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:

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

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:

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

Twisted De Rham-Hodge Theory

- **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:

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

- Future work:
  - Generalize to higher-dimensional simplicial complexes
  - Use this Hodge structure to design synchronization algorithms
  - Synchronization beyond the regime of linear algebraic groups (not unseen in real applications)
  - Towards a spectral sheaf theory?

Outline

Background
- Evolutionary Biology
- Collection Shape Analysis
- Cryo-Electron Microscopy

Graph Synchronization Problems: Principal Bundle Perspective
- Representation Variety
- Discrete Twisted Hodge Theory

Learning Aspects
- Horizontal Diffusion Maps
- Learning Group Actions

Simultaneous Matching and Clustering (SMAC)

Learning Aspect I: Horizontal Diffusion Maps (HDM)

Doug Boyer

Jukka Jernvall

More Precisely: biological morphologists

↓

Study Teeth & Bones of

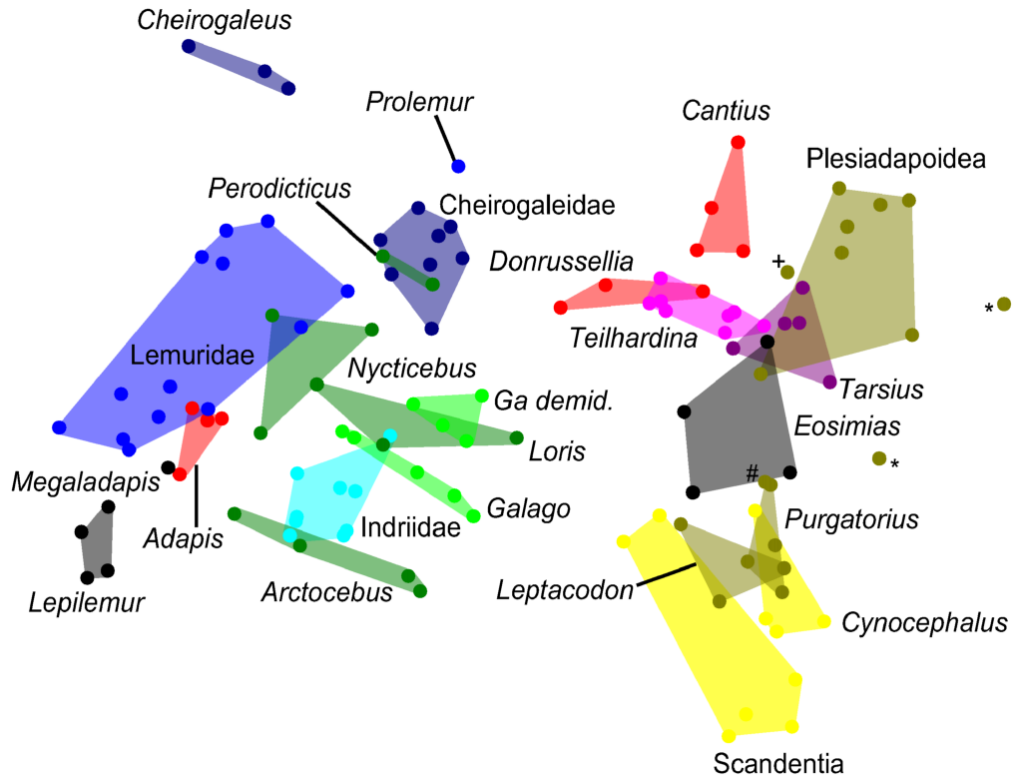
extant & extinct animals

still live today — extant

fossils — extinct

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.

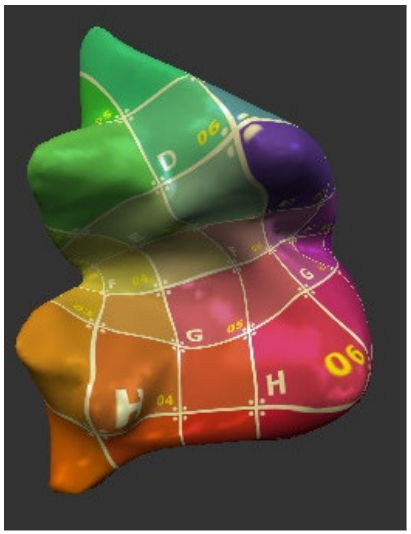
Shape Distances

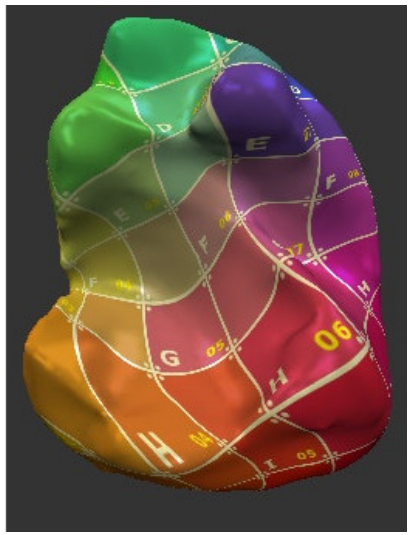
$d_{\mathrm{cWn}}(S_1, S_2)$: Conformal Wasserstein Distance (CWD)

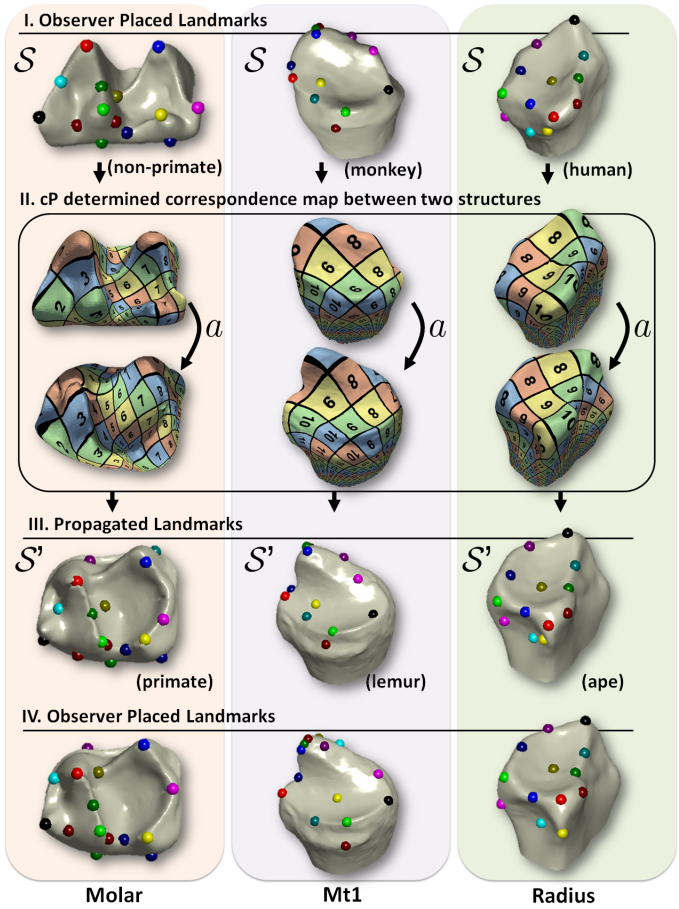
$d_{\mathrm{cP}}(S_1, S_2)$: Continuous Procrustes Distance (CPD)

$d_{\mathrm{cKP}}(S_1, S_2)$: Continuous Kantorovich-Procrustes Distance (CKPD)

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

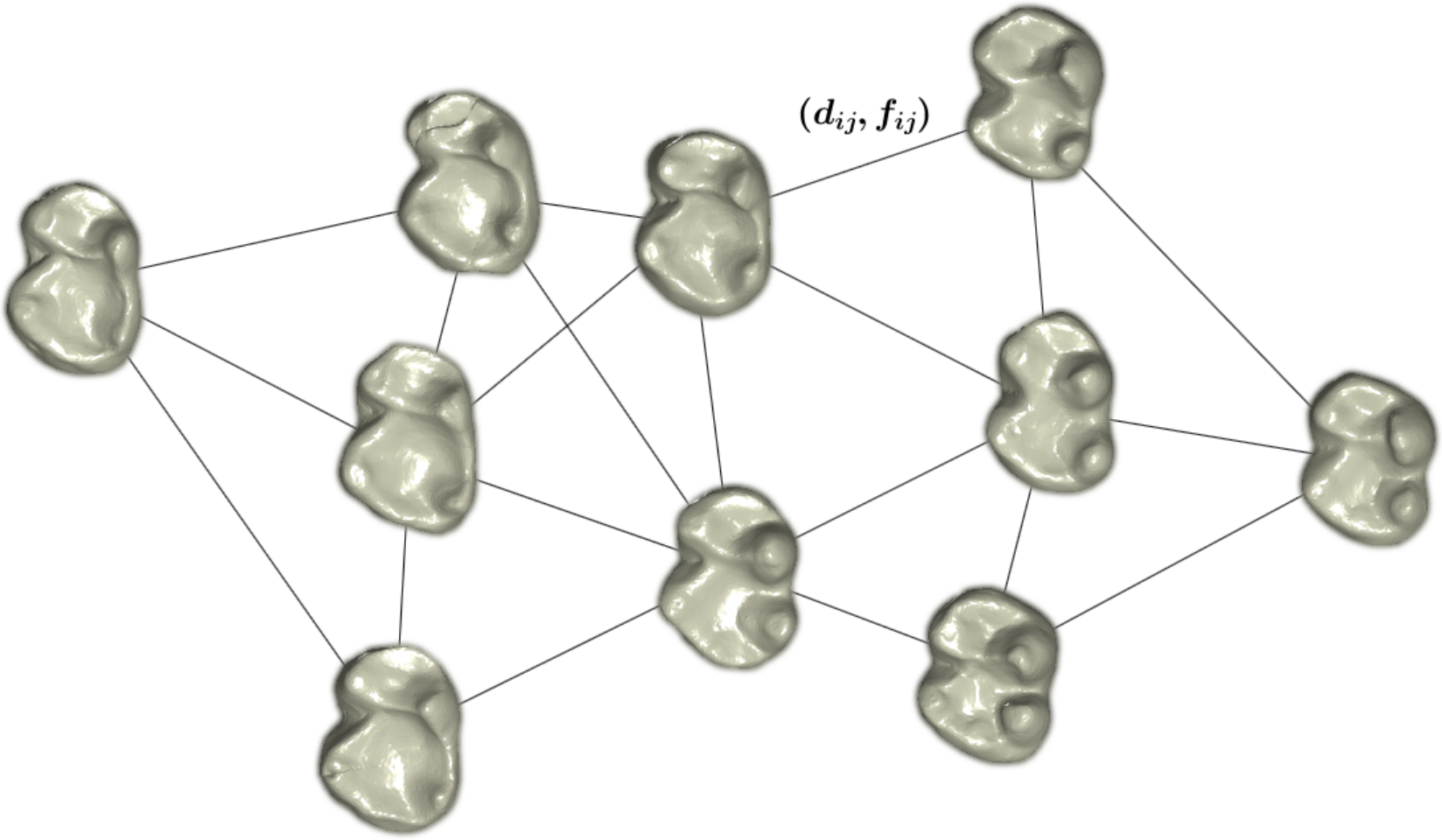




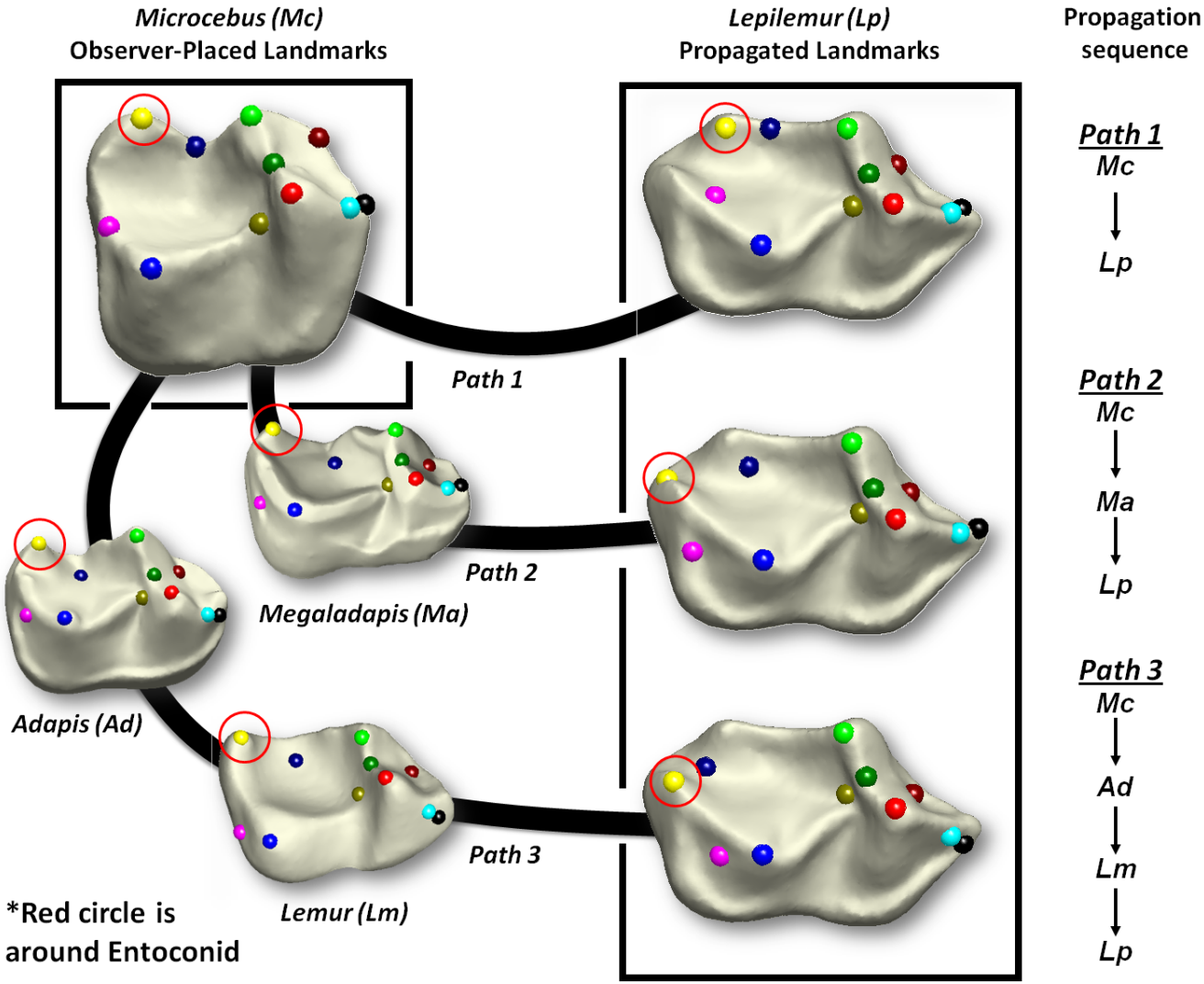


Shape Distances

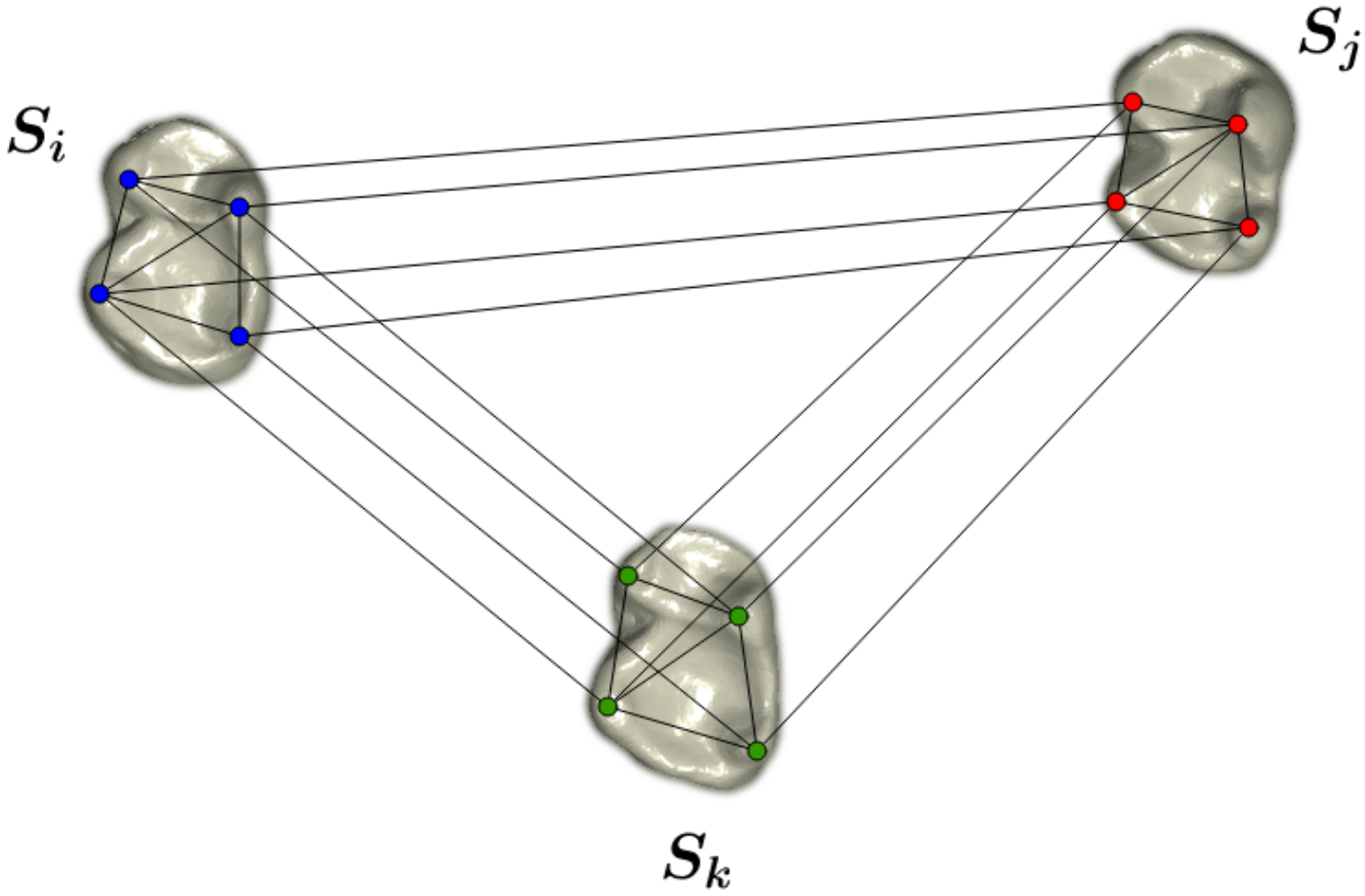
- John Clifford Gower, John C. Gower, and Garmt B. Dijksterhuis. "**Procrustes Problems.**" vol. 3 of Oxford Statistical Science Series, Oxford University Press Oxford, 2004.
- Rima Alaifari, Ingrid Daubechies, and Yaron Lipman. "**Continuous Procrustes distance between two surfaces.**" *Communications on Pure and Applied Mathematics 66*, no. 6 (2013): 934-964.
- Doug M. Boyer, Yaron Lipman, Elizabeth St Clair, Jesus Puente, Biren A. Patel, Thomas Funkhouser, Jukka Jernvall, and Ingrid Daubechies. "**Algorithms to automatically quantify the geometric similarity of anatomical surfaces.**" *Proceedings of the National Academy of Sciences* 108, no. 45 (2011): 18221-18226.
- Facundo Mémoli. "**Gromov-Wasserstein distances and the metric approach to object matching.**" *Foundations of Computational Mathematics* 11, no. 4 (2011): 417-487.
- Rongjie Lai, Hongkai Zhao. "**Multi-scale Non-Rigid Point Cloud Registration Using Robust Sliced-Wasserstein Distance via Laplace-Beltrami Eigenmap**", *SIAM Journal on Imaging Sciences* 10(2), pp. 449âĂŤ483, 2017.
- Patrice Koehl, Joel Hass, "**Landmark-Free Geometric Methods in Biological Shape Analysis**", *Journal of The Royal Society Interface*, 12 no. 113 (2015): 20150795
- .....


(d_{ij}, f_{ij})

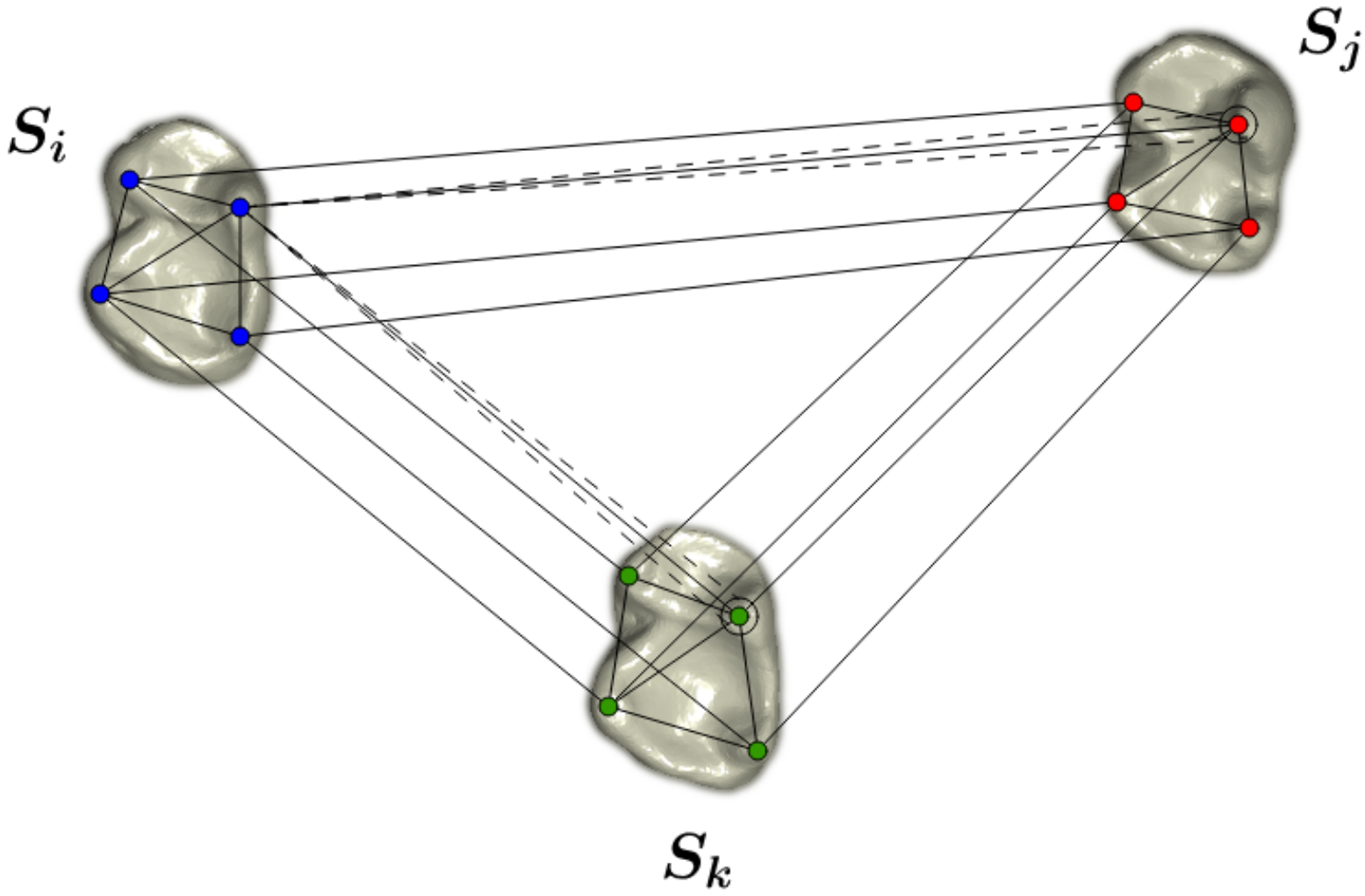
Shape Space is NOT a Trivial Fibre Bundle (due to noise?)



"Synchronize" or "Denoise" Maps: Ideally



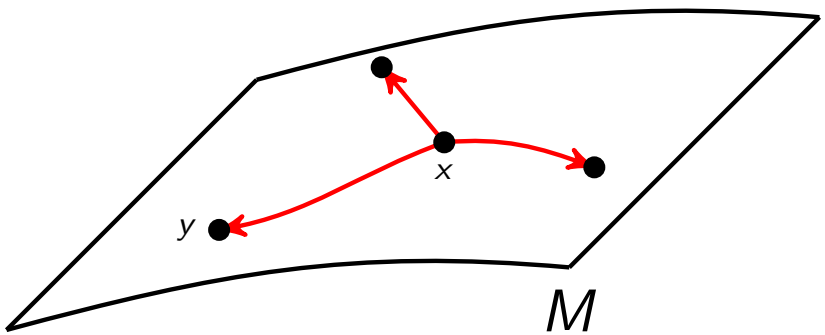
"Synchronize" or "Denoise" Maps: Reality



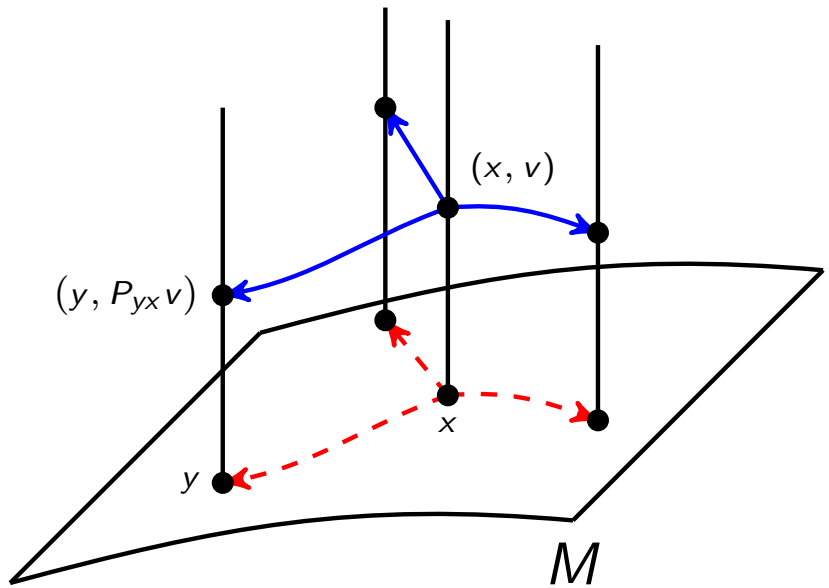
Horizontal Random Walk on a *Fibre Bundle*

$$H_\epsilon^{(\alpha)} f(x,v) = \frac{\int_M K_\epsilon^{(\alpha)}(x,y)\, f(y, P_{yx} v)\, p(y)\, d\mathrm{vol}_M(y)}{\int_M K_\epsilon^{(\alpha)}(x,y)\, p(y)\, d\mathrm{vol}_M(y)}$$

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

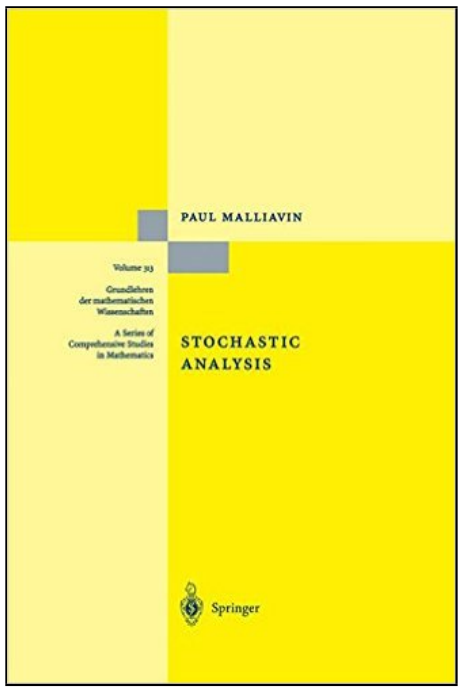


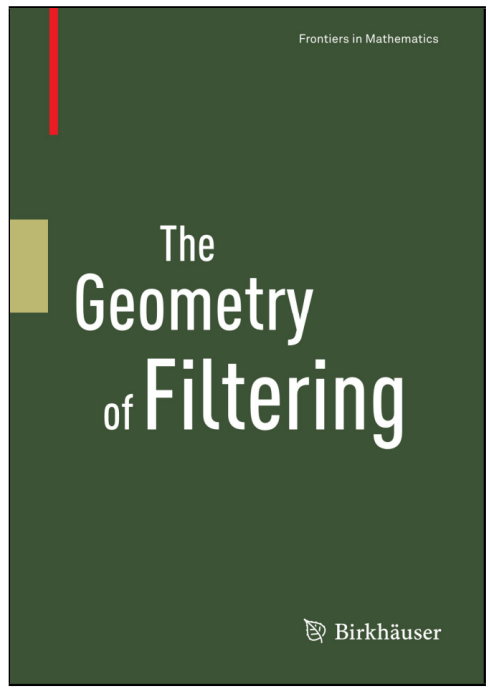
$H_{\epsilon,\delta}^{(\alpha)}$

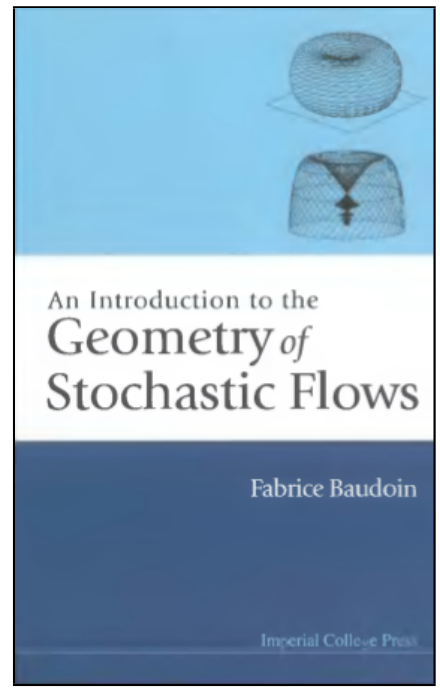


Horizontal Diffusion in Stochastic Differential Geometry

- 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 ProbabilitÃl's 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.

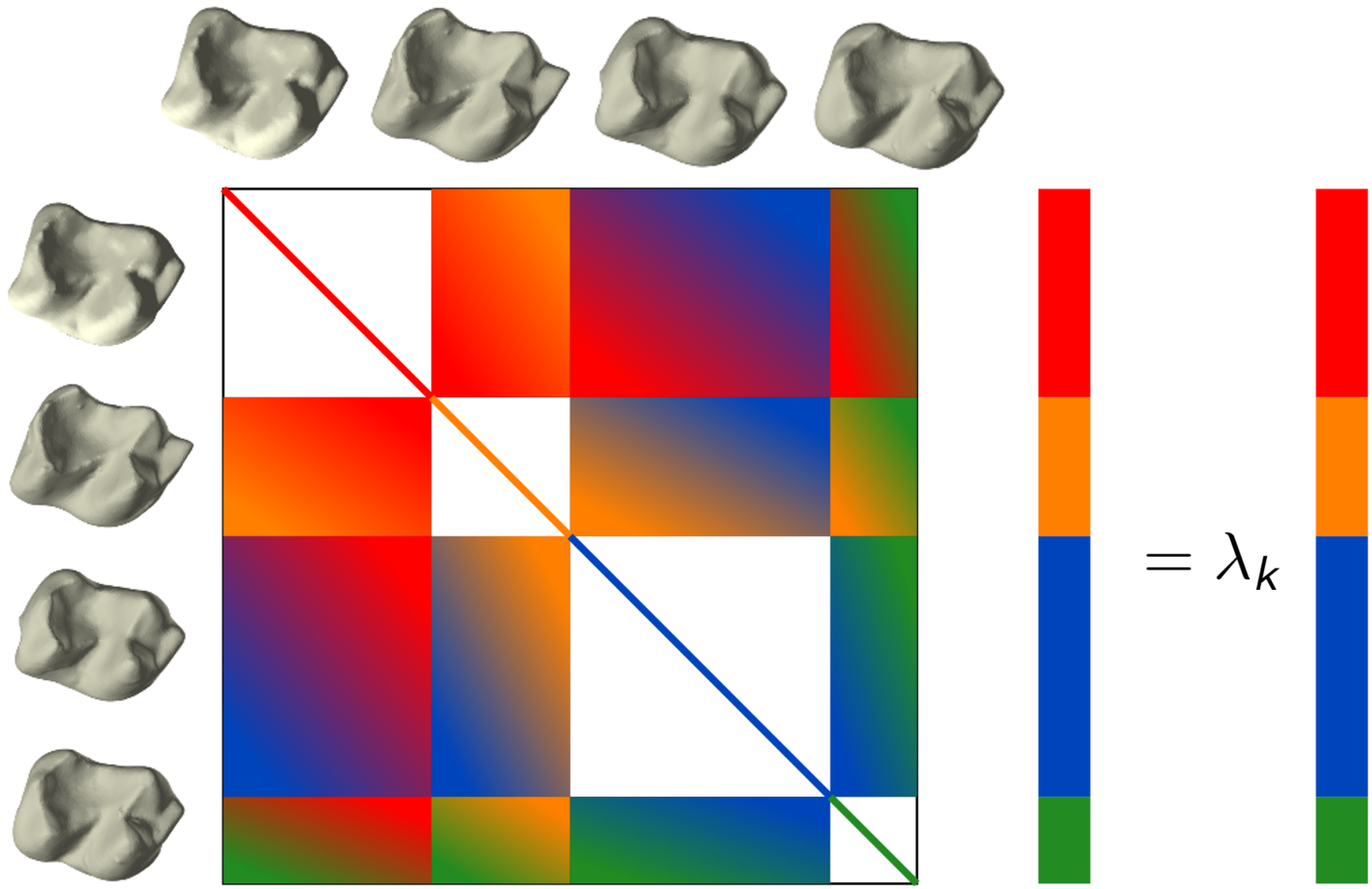




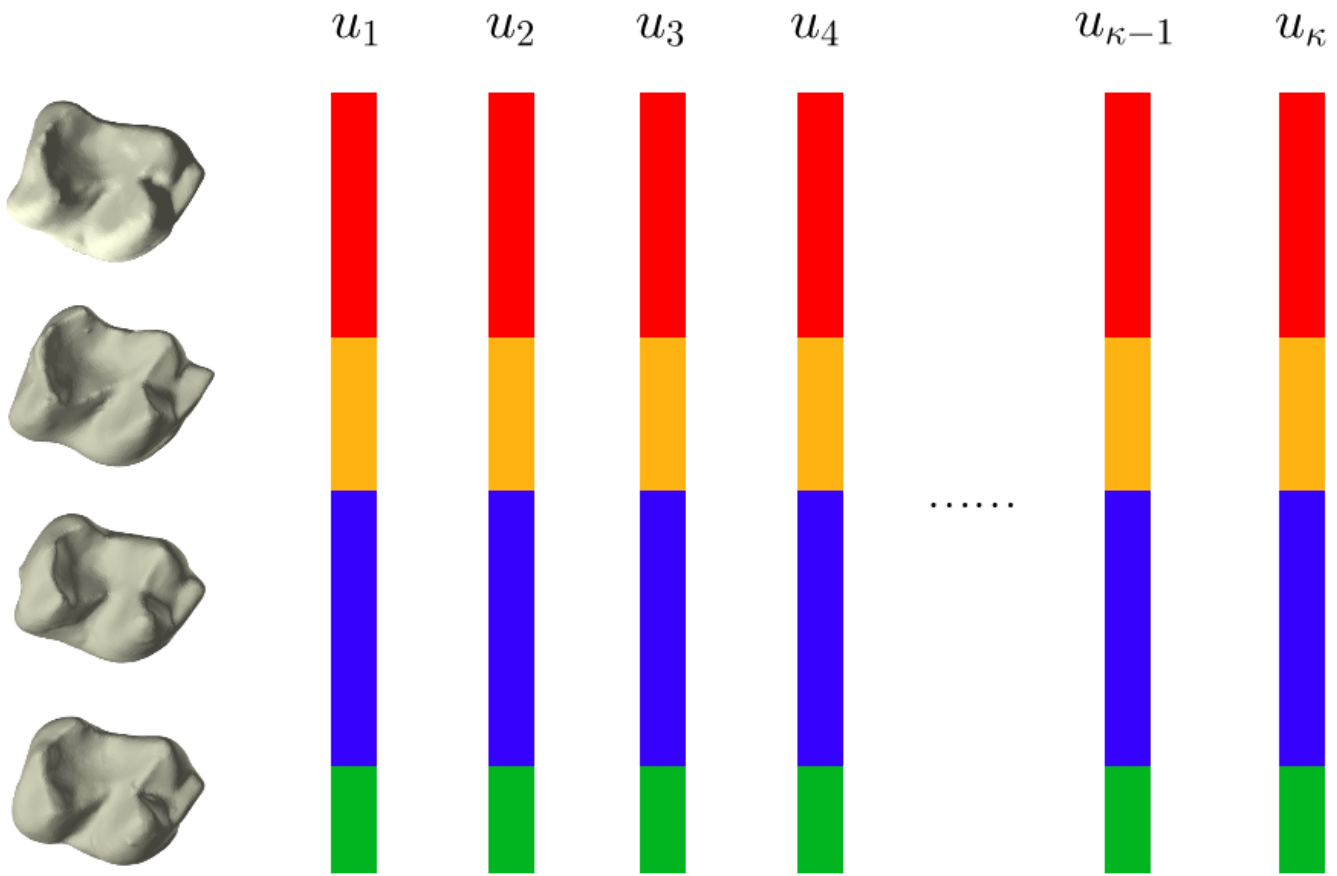


Horizontal Diffusion Maps

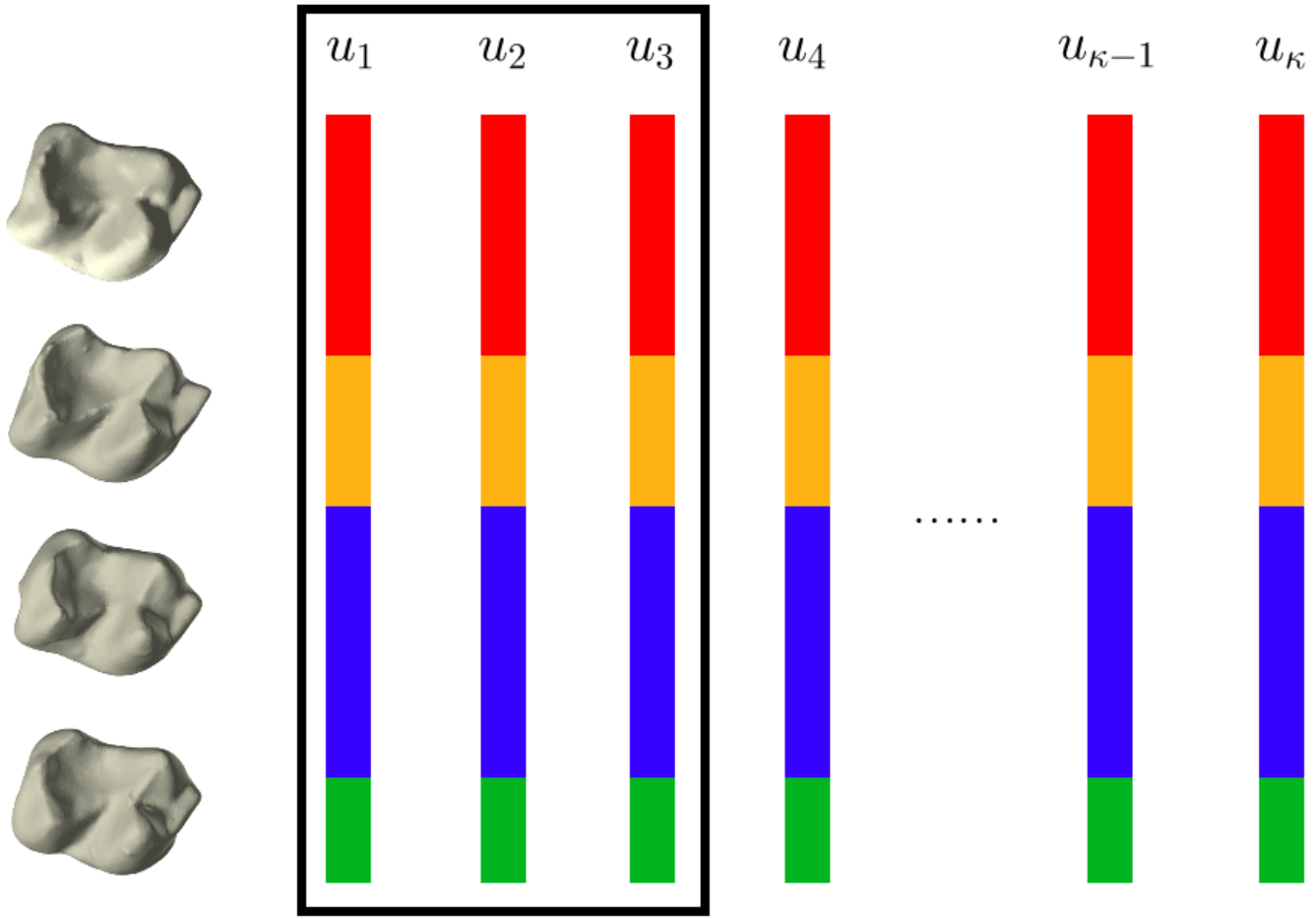
$$\mathcal{D}^{-1}\mathcal{W}u_k = \lambda_k u_k, \quad 1 \leq k \leq \kappa$$



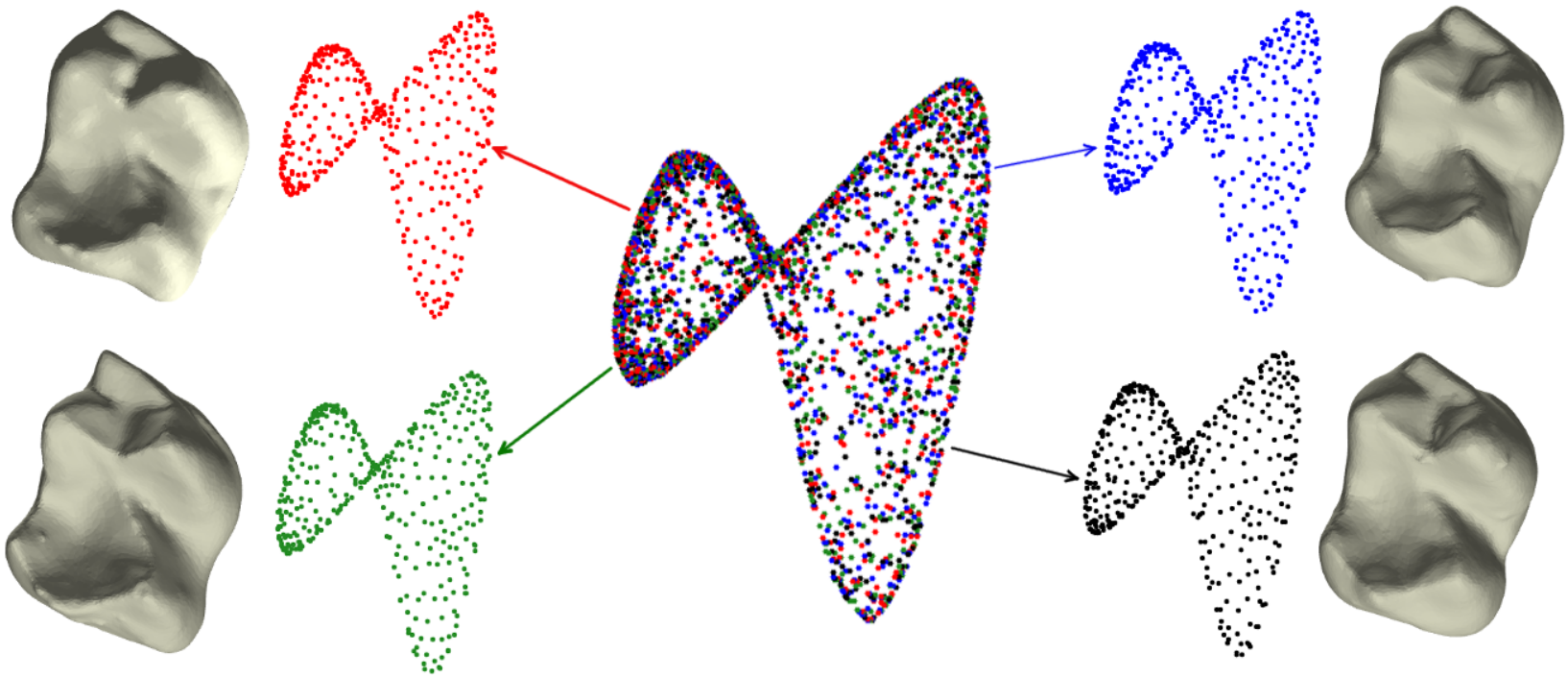
Horizontal Diffusion Maps: Embedding the Entire Bundle



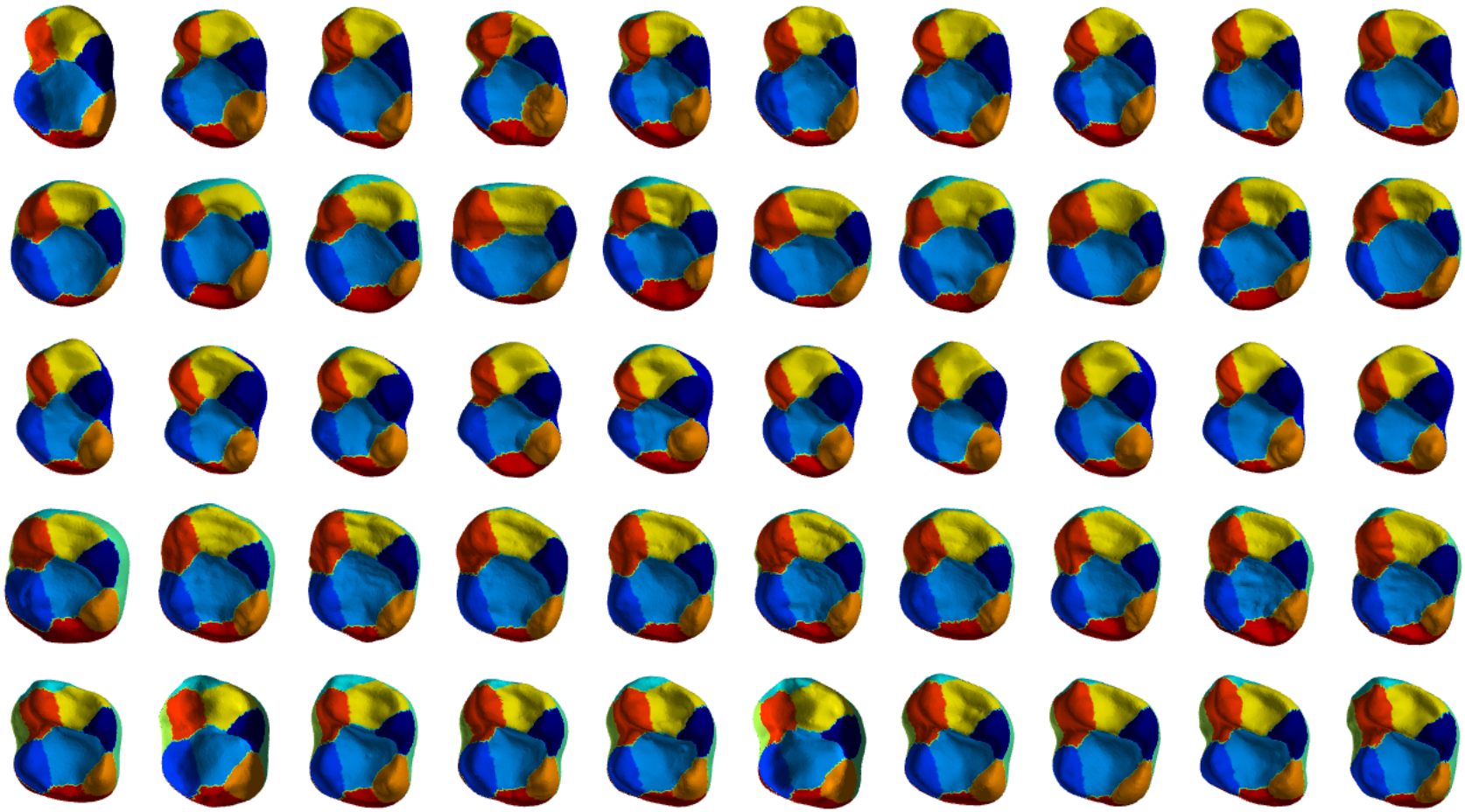
Horizontal Diffusion Maps: Embedding the Entire Bundle



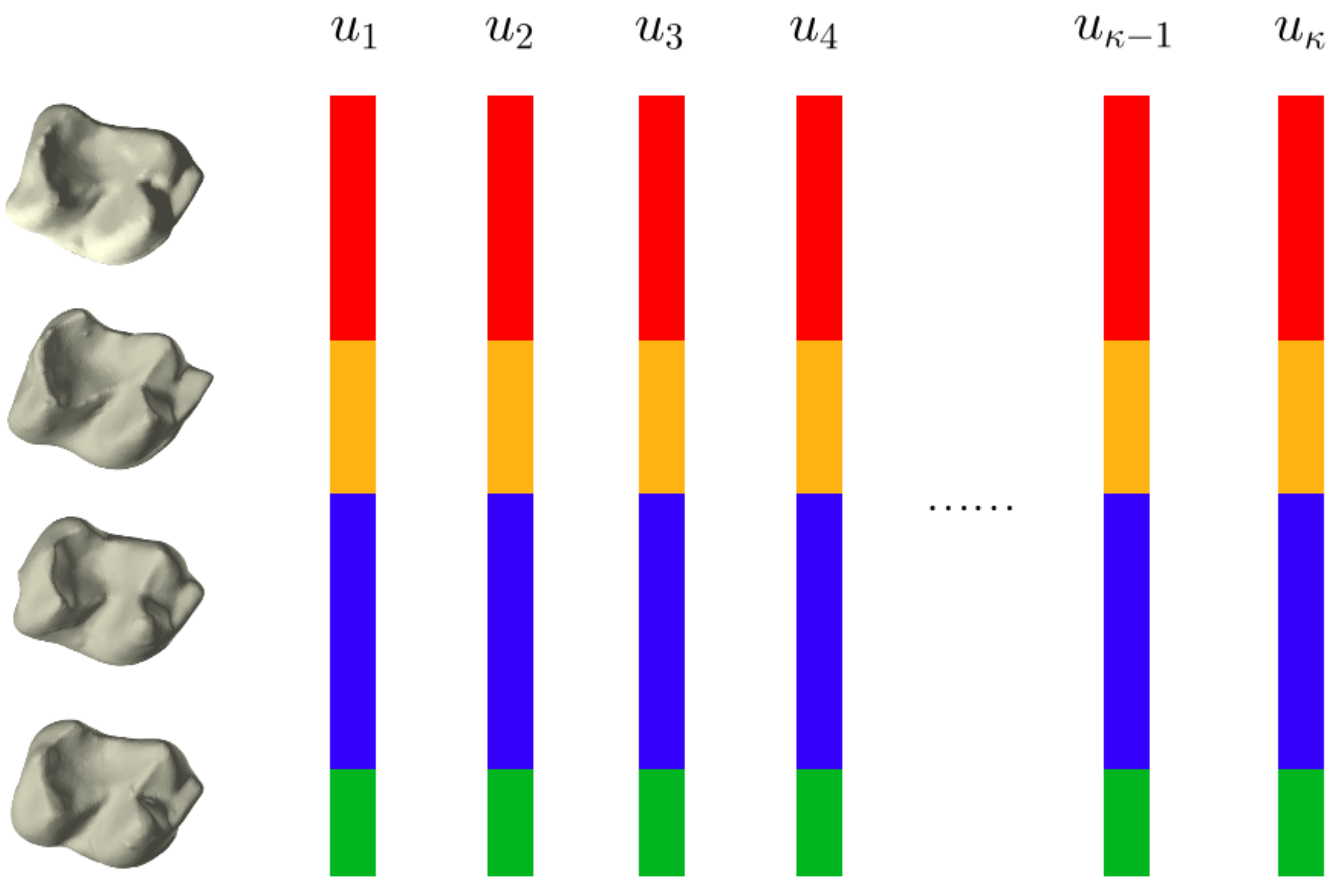
Horizontal Diffusion Maps



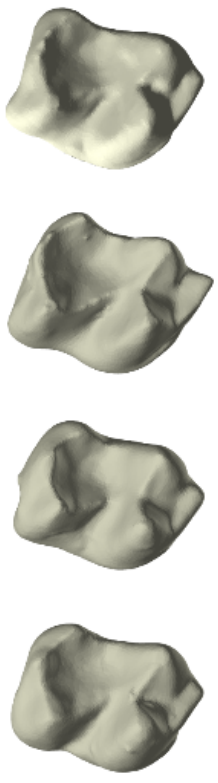
Automatic Landmarking — Interpretability

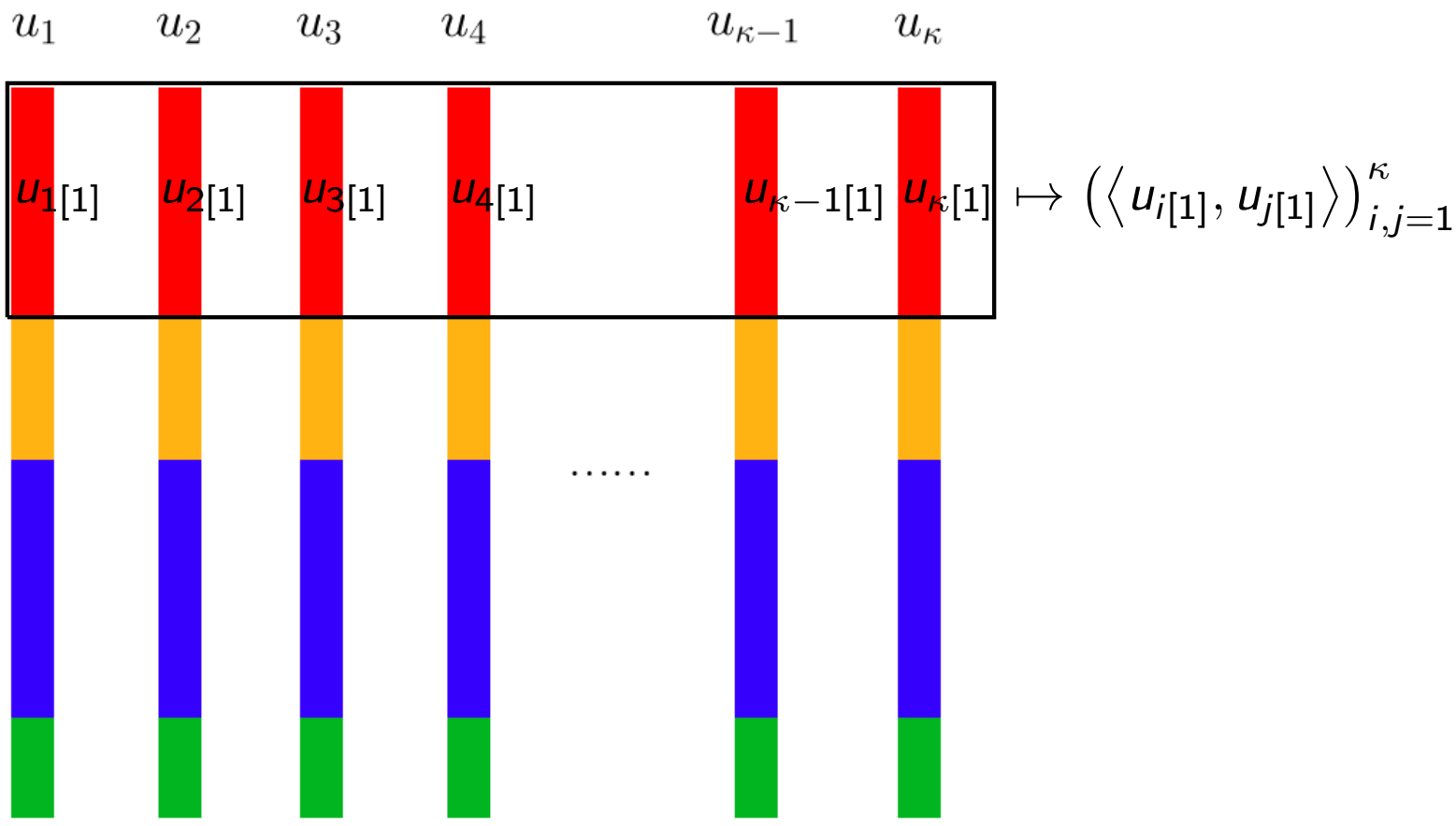


Horizontal Diffusion Maps: Embedding the Base Manifold

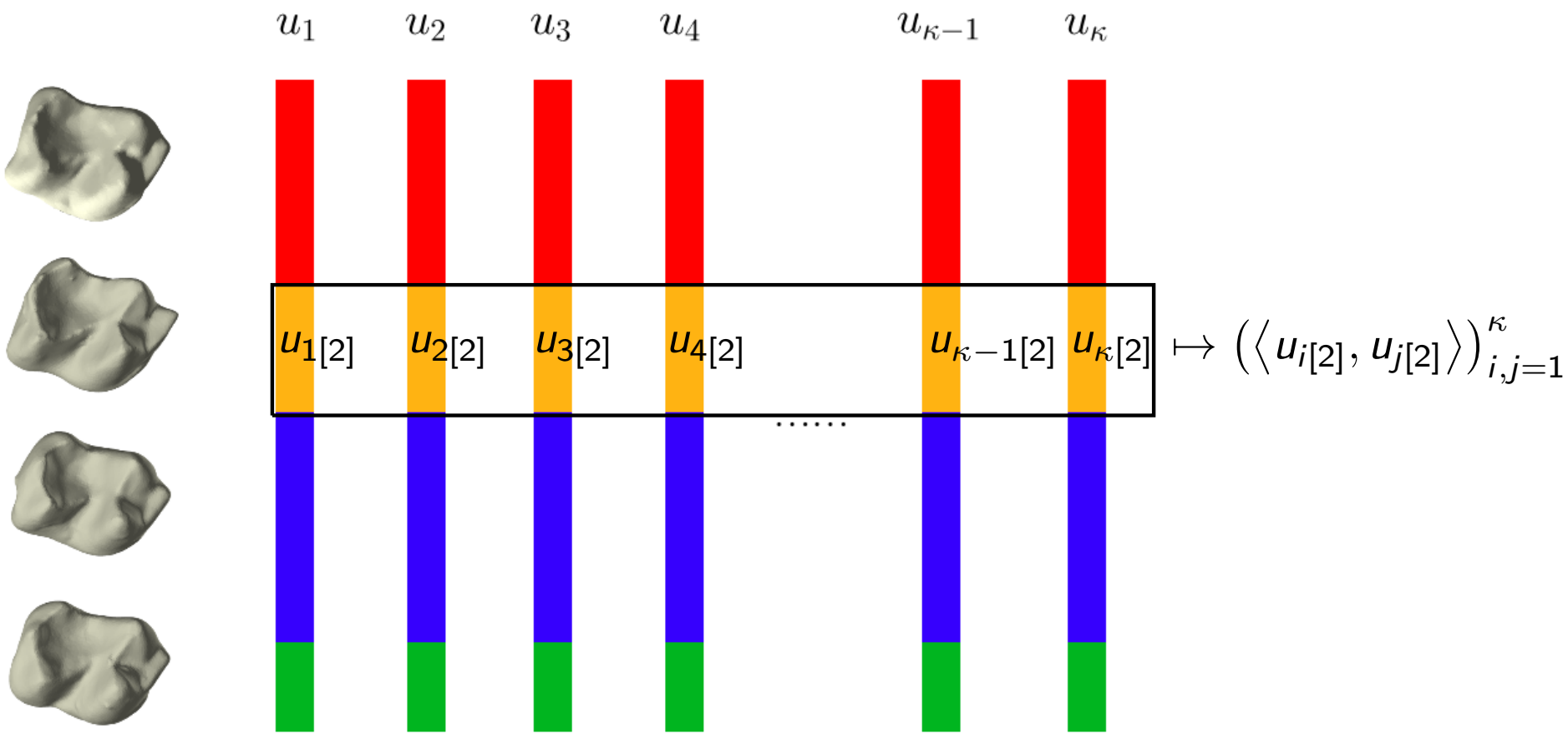


Embedding the Base Manifold with Horizontal Heat Kernel

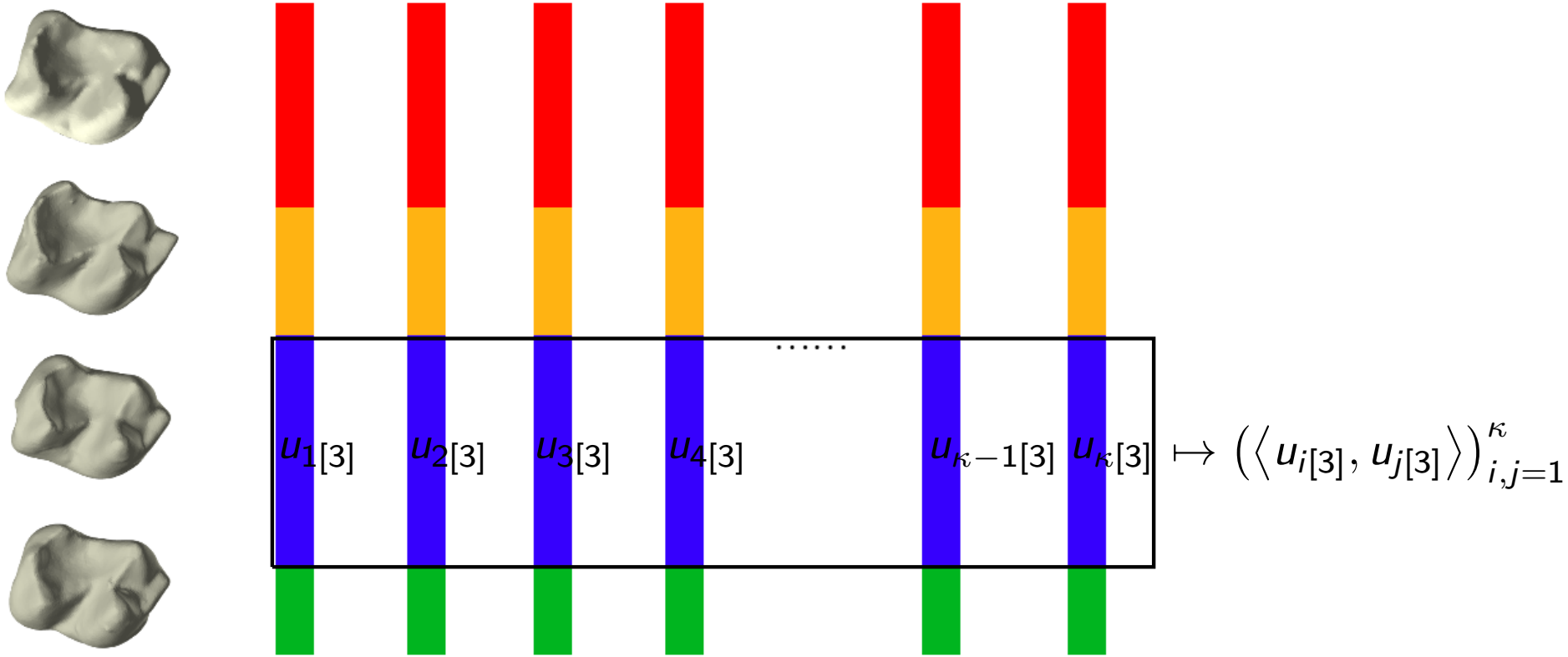




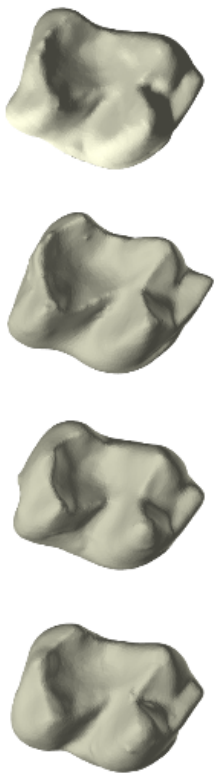
Embedding the Base Manifold with Horizontal Heat Kernel

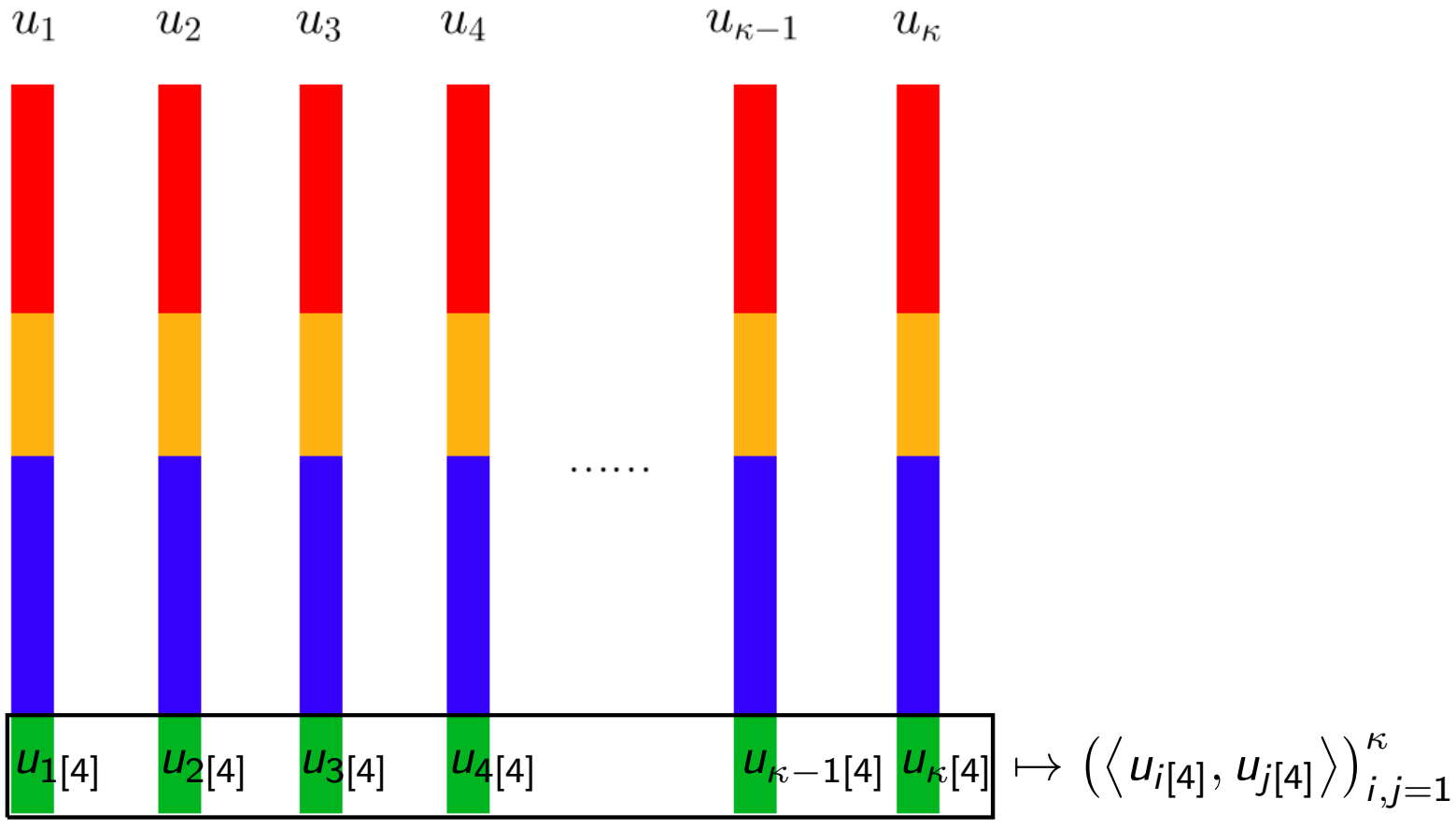


Embedding the Base Manifold with Horizontal Heat Kernel

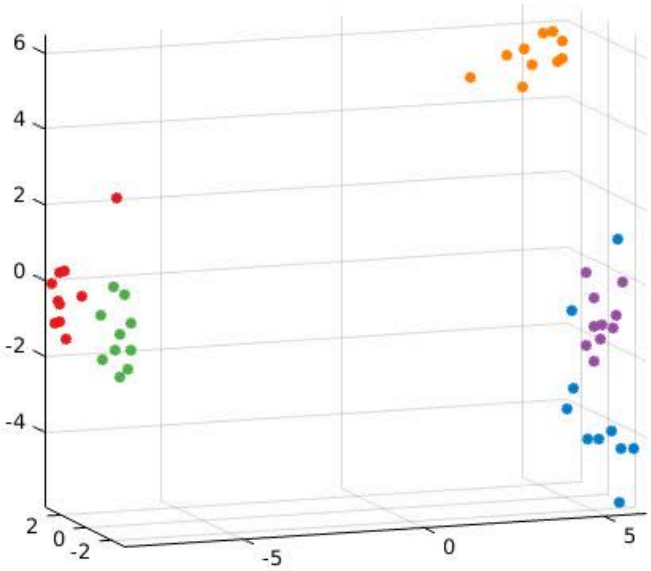


Embedding the Base Manifold with Horizontal Heat Kernel

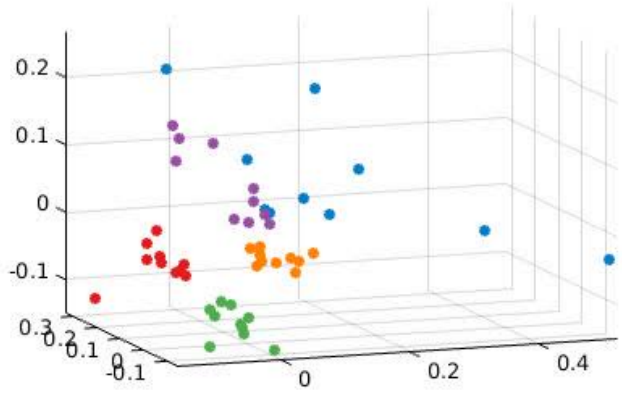




Species Clustering

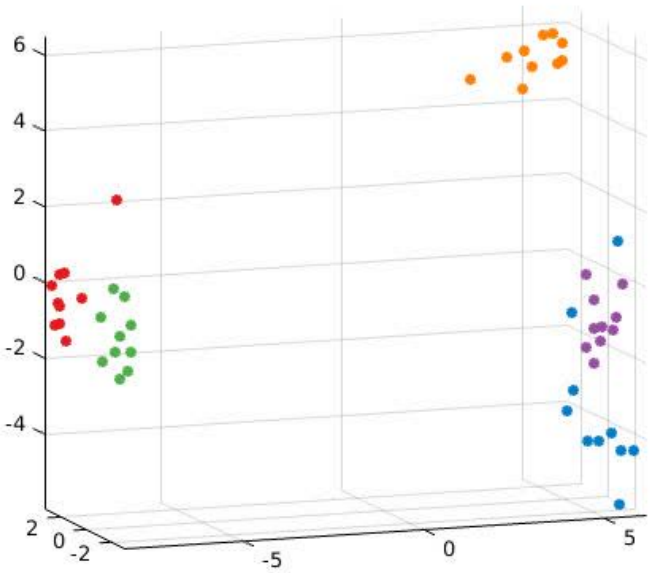


Horizontal Base Diffusion Distance (with Maps)

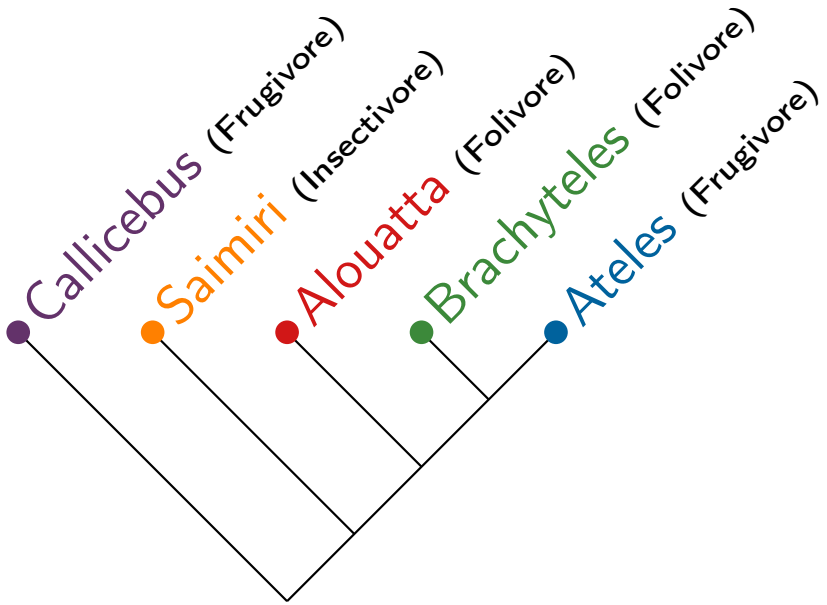


Diffusion Distance (without Maps)

Species Clustering



Horizontal Base Diffusion Distance (with Maps)

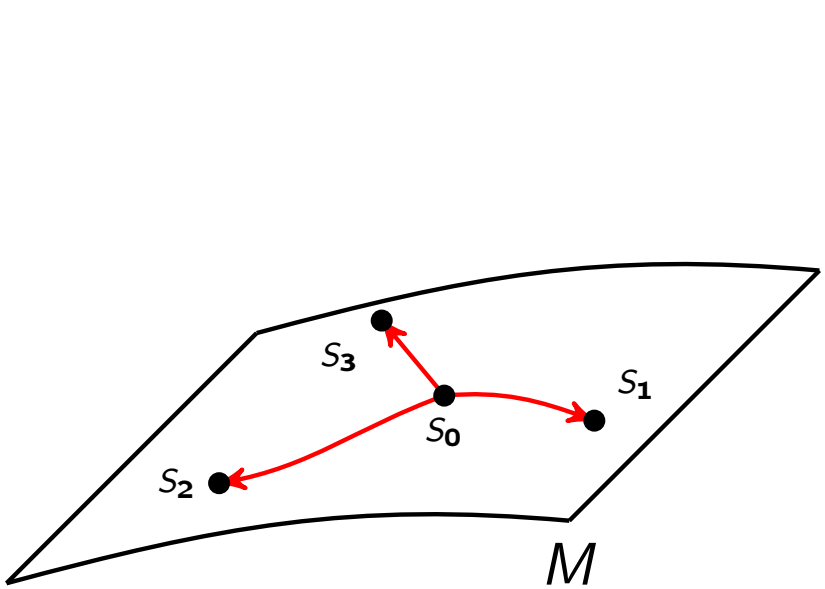


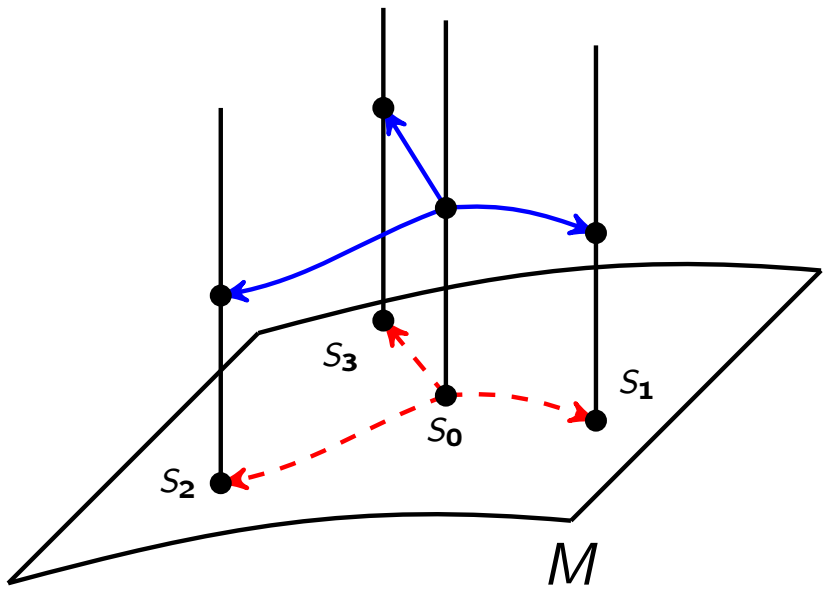
HDM: Mathematical Theory

ϵ : diffusion bandwidth on the base

δ : diffusion bandwidth along each fibre

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





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.

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

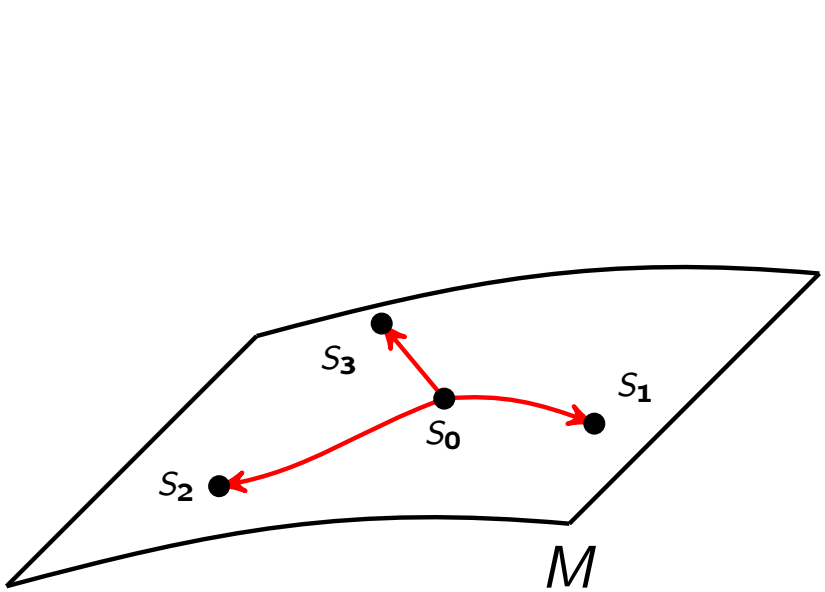
HDM: Horizontal Random Walk on a *Fibre Bundle*

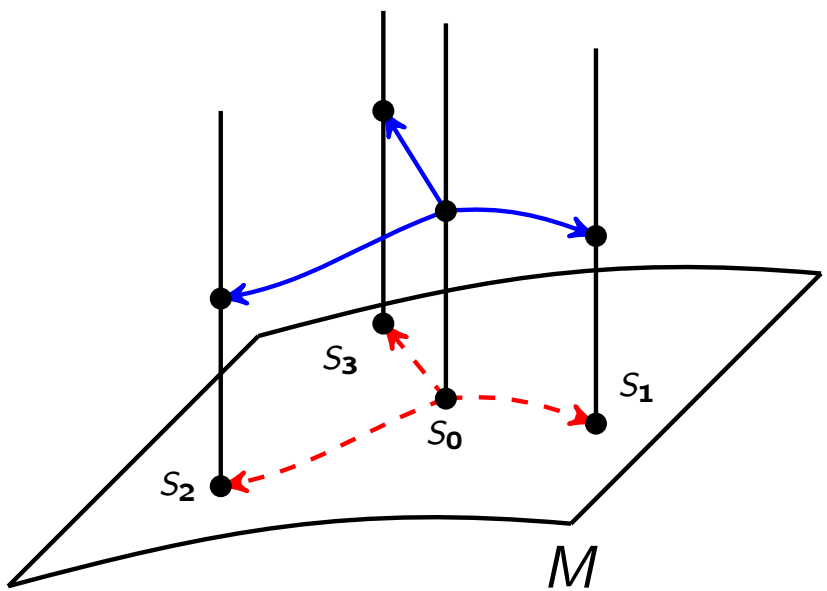
ϵ : diffusion bandwidth on the base

δ : diffusion bandwidth along each fibre

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

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





HDM: Infinitesimal Generator

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 $\epsilon \to 0$,

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

- 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*
- $\Rightarrow$ HDM is not applying diffusion maps to the total space!

HDM: Convergence Rate 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

Learning Aspect II: Learning Group Actions

Problem (Learning Group Actions)

Given a group G acting on a set X, simultaneously learn a new action of G on X and a partition of X into disjoint subsets $X_1, \cdots, X_K$, such that the new action is as close as possible to the given action and cycle-consistent on each X_i $(1 \leq i \leq K)$.

Learning Group Actions by Synchronization

Problem (Learning Group Actions by Synchronization)

Denote $\mathscr{X}_K$ for all partitions of Γ into K nonempty connected subgroups ($K \leq n$) and

$$\nu(S_i) = \inf_{f \in C^0(\Gamma; G)} \sum_{j,k \in S_i} w_{jk} \mathrm{Cost}_G(f_j, \rho_{jk} f_k), \ \mathrm{vol}(S_i) = \sum_{j \in S_i} d_j,$$

Solve the optimization problem

$$\min_{\{S_1, \cdots, S_K\} \in \mathscr{X}_K} \frac{\max\limits_{1 \leq i \leq K} \nu(S_i)}{\min\limits_{1 \leq i \leq K} \mathrm{vol}(S_i)} \tag{1}$$

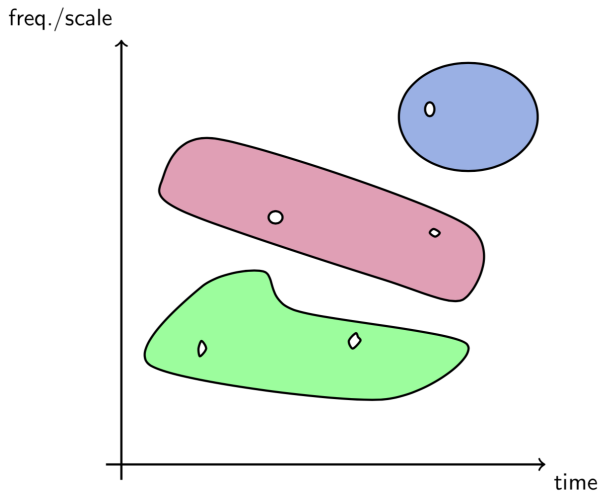
and output an optimal partition $\{S_1, \cdots, S_K\}$ together with the minimizing $f \in C^0(\Gamma; G)$.

Analogy with the New Phase Retrieval Paradigm

R Alaifari, I Daubechies, P Grohs, R Yin. Stable Phase Retrieval in Infinite Dimensions, arXiv:1609.00034, 2017

Allow different phase factors α_1, α_2, α_3 on the different regions:

$$V_\varphi f \sim e^{i\alpha_1} V_\varphi f_1 + e^{i\alpha_2} V_\varphi f_2 + e^{i\alpha_3} V_\varphi f_3$$



Stability for holomorphic measurements

Theorem: [A., Daubechies, Grohs, Yin '17]

Suppose that $F_1 \in \mathcal{H}(D, (D_0^i)_{i=1}^l, \delta, \Delta)$ and $F_2 \in C^1(D)$ such that there exists a continuous function $\eta : D \to \mathbb{C}$ for which

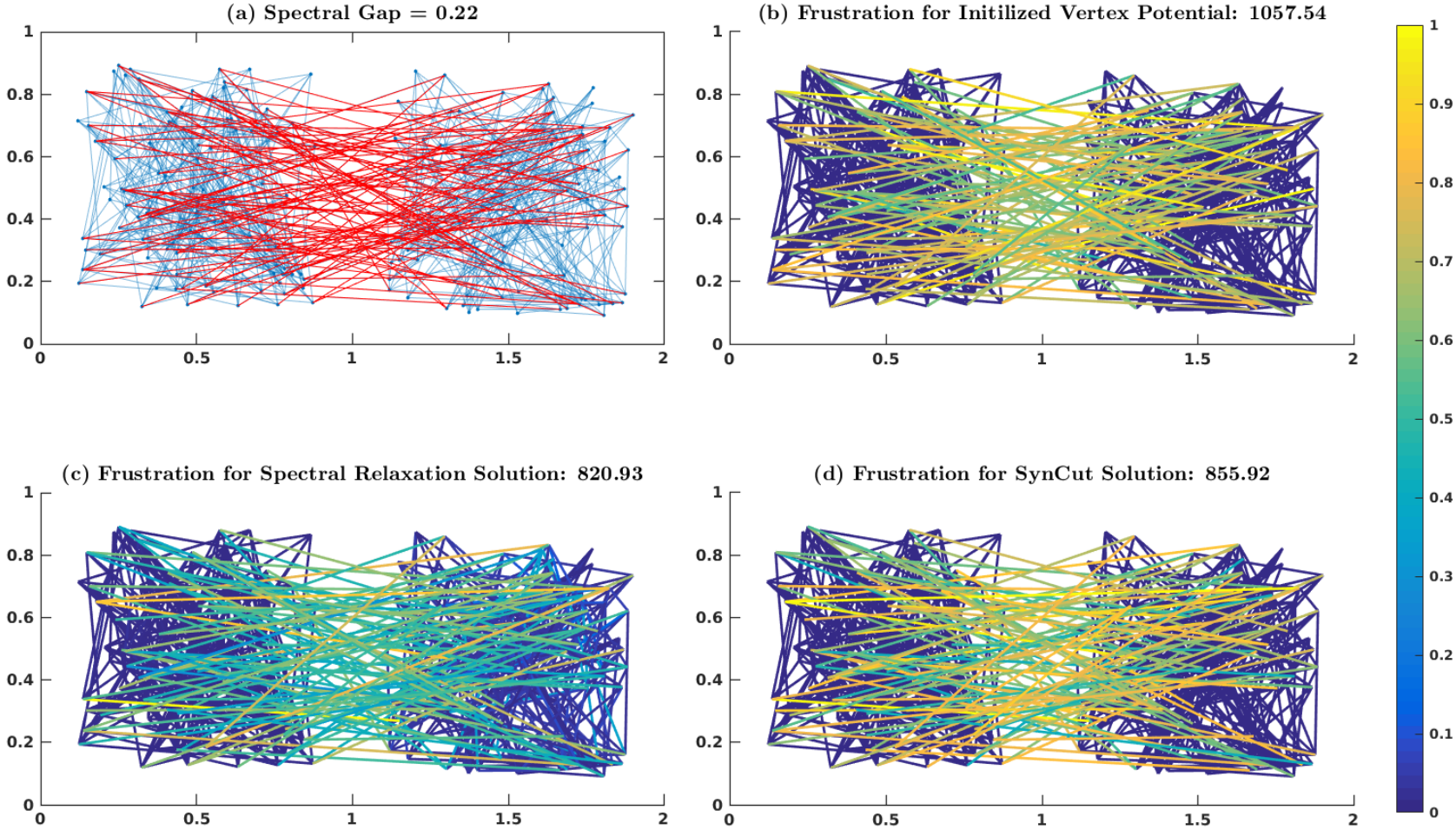
$$\eta \cdot F_1, \ \eta \cdot F_2 \in \mathcal{O}(D).$$

Suppose that $1 \leq p \leq \infty$. Pick $z_0 \in D_+$.
Then the following estimate holds:

$$\inf_{\alpha \in \mathbb{R}} \left\| F_1 - e^{i\alpha} F_2 \right\|_{L^p(D)} \leq$$

$$\leq C(z_0, p, D_+, (D_0^i)_{i=1}^l) \frac{\Delta^2}{\delta^2} \left\| |F_1| - |F_2| \right\|_{W^{1,p}(D_+)}.$$

Clustering based on Synchronization Residuals



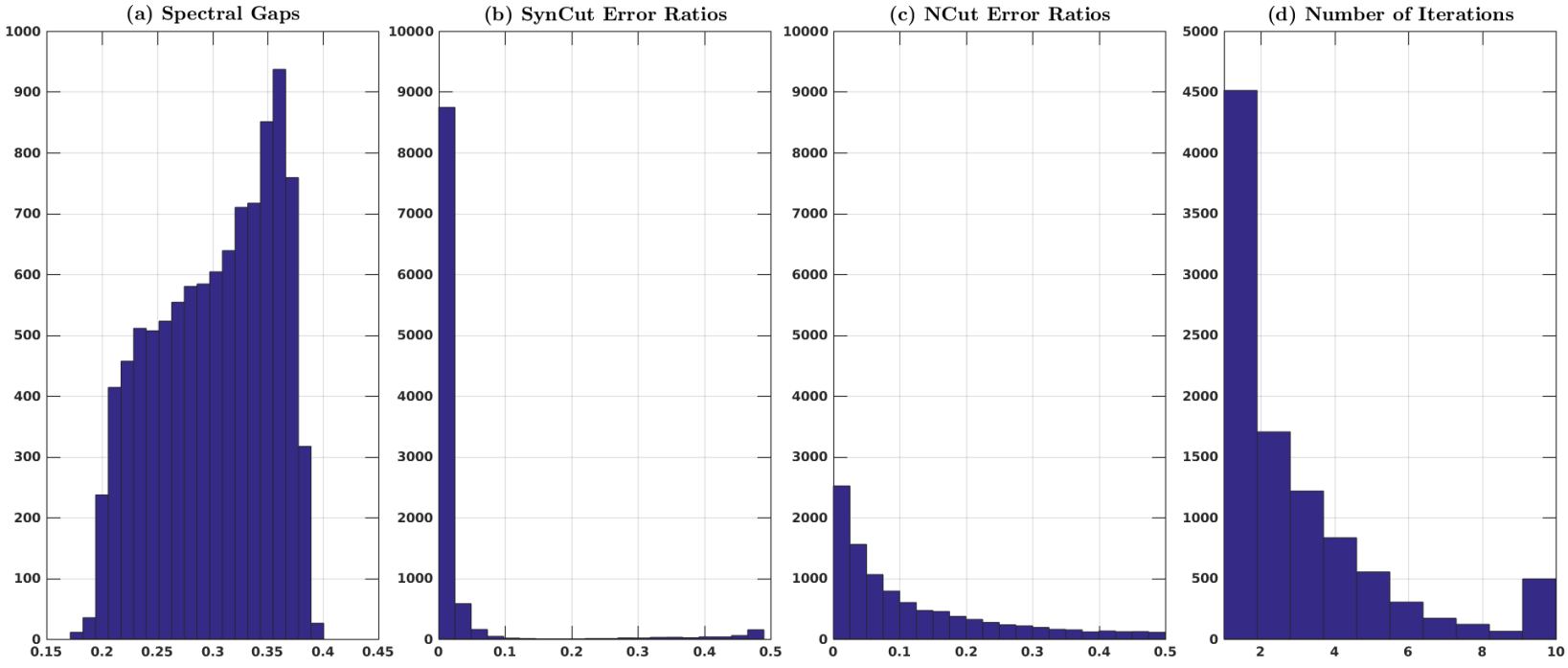
Algorithm: SynCut

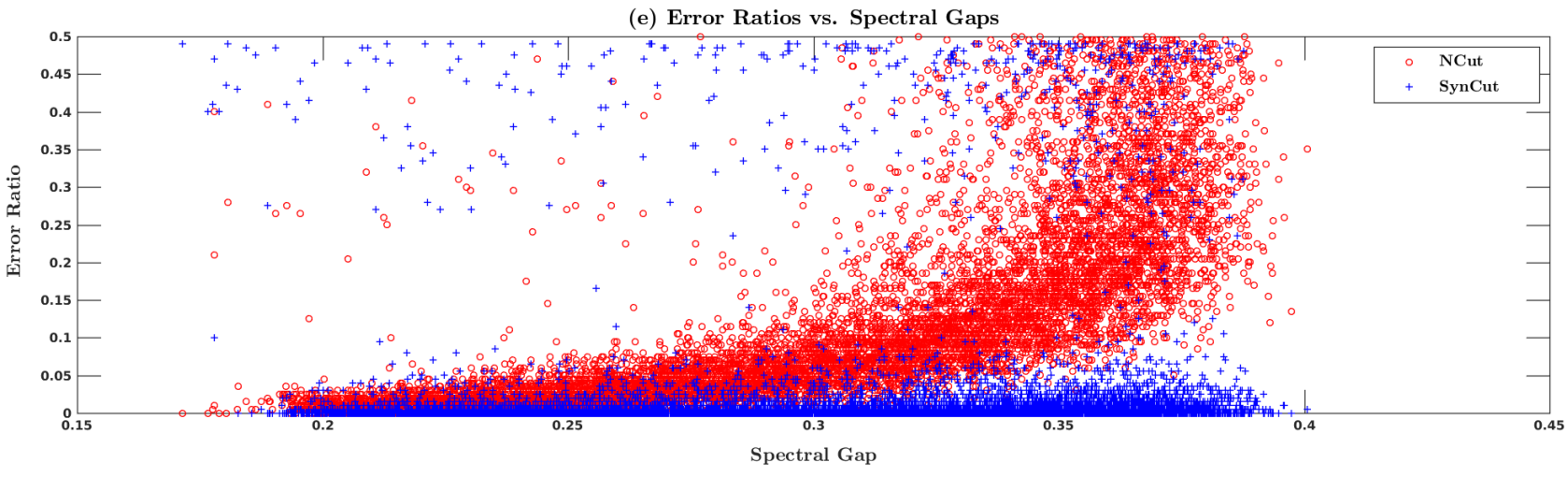
Input: $\Gamma = (V, E, w)$, $\rho \in C^1(\Gamma; G)$, number of partitions K
Output: Partitions $\{S_1, \cdots, S_K\}$

1. Solve synchronization problem over Γ for ρ, obtain $f \in C^0(\Gamma; G)$
2. Compute $d_{ij} = \exp(-w_{ij} \|f_i - \rho_{ij} f_j\|)$ on all edges $(i, j) \in E$
3. Spectral clustering on weighted graph (V, E, d) to get $\{S_1, \cdots, S_k\}$
4. Solve synchronization problem within each partition S_j, "glue up" the local solutions to obtain $f_* \in C^0(\Gamma; G)$
5. $f \leftarrow f_*$, repeat from Step 2

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

Compare with Graph-based Spectral Clustering





Outline

Background
- Evolutionary Biology
- Collection Shape Analysis
- Cryo-Electron Microscopy

Graph Synchronization Problems: Principal Bundle Perspective
- Representation Variety
- Discrete Twisted Hodge Theory

Learning Aspects
- Horizontal Diffusion Maps
- Learning Group Actions

Simultaneous Matching and Clustering (SMAC)

Learning Group Action: SBM+Random Corruption Model

- Consider graph $\Gamma = (V, E)$, where $V = (v_1, \cdots, v_n)$ are separated into 2 communities
- Fix orthogonal matrices $R_1, \cdots, R_n \in O(d)$, one at each corresponding vertex
- Fix $0 \leq q < p \leq 1$. Consider the following $\rho \in C^1(\Gamma; O(d))$:
  - If v_i, v_j belong to the same community,

$$\rho_{ij} = \begin{cases} R_i^\top R_j & \text{with probability } p \\ \mathrm{Unif}(O(d)) & \text{with probability } 1-p \end{cases}$$

  - If v_i, v_j belong to different communities

$$\rho_{ij} = \begin{cases} R_i^\top R_j & \text{with probability } q \\ \mathrm{Unif}(O(d)) & \text{with probability } 1-q \end{cases}$$

Learning Group Action: SBM+Random Corruption Model

- Fix $0 \leq q < p \leq 1$. Consider the following $\rho \in C^1(\Gamma; O(d))$:
  - If v_i, v_j belong to the same community,

$$\rho_{ij} = \begin{cases} R_i^\top R_j & \text{with probability } p \\ \mathrm{Unif}(O(d)) & \text{with probability } 1-p \end{cases}$$

  - If v_i, v_j belong to different communities

$$\rho_{ij} = \begin{cases} R_i^\top R_j & \text{with probability } q \\ \mathrm{Unif}(O(d)) & \text{with probability } 1-q \end{cases}$$

- "Clean" case ($p = q = 0$): synchronization solution

$$\tilde{\rho}_{ij} = R_i^\top R_j$$

Learning Group Action: SBM+Random Corruption Model

- Fix $0 \leq q < p \leq 1$. Consider the following $\rho \in C^1(\Gamma; O(d))$:
  - If v_i, v_j belong to the same community,

$$\rho_{ij} = \begin{cases} R_i^\top R_j & \text{with probability } p \\ \mathrm{Unif}(O(d)) & \text{with probability } 1-p \end{cases}$$

  - If v_i, v_j belong to different communities

$$\rho_{ij} = \begin{cases} R_i^\top R_j & \text{with probability } q \\ \mathrm{Unif}(O(d)) & \text{with probability } 1-q \end{cases}$$

- "Perturbed case ($q > p > 0$)": synchronization solution

$$\tilde{\rho}_{ij} \approx R_i^\top R_j$$

Learning Group Action: SBM+Random Corruption Model

- Fix $0 \leq q < p \leq 1$. Consider the following $\rho \in C^1(\Gamma; O(d))$:
  - If v_i, v_j belong to the same community,

$$\rho_{ij} = \begin{cases} R_i^\top R_j & \text{with probability } p \\ \mathrm{Unif}(O(d)) & \text{with probability } 1-p \end{cases}$$

  - If v_i, v_j belong to different communities

$$\rho_{ij} = \begin{cases} R_i^\top R_j & \text{with probability } q \\ \mathrm{Unif}(O(d)) & \text{with probability } 1-q \end{cases}$$

- "Perturbed case $(q > p > 0)$": synchronization solution

$$\tilde{\rho}_{ij} \approx R_i^\top R_j$$

- But $\|\tilde{\rho}_{ij} - \rho_{ij}\|$ deviates from 0 more frequently on edges connecting different communities!

Learning Group Action: SBM+Random Corruption Model

- Fix $0 \leq q < p \leq 1$. Consider the following $\rho \in C^1(\Gamma; O(d))$:
  - If v_i, v_j belong to the same community,

$$\rho_{ij} = \begin{cases} R_i^\top R_j & \text{with probability } p \\ \mathrm{Unif}(O(d)) & \text{with probability } 1-p \end{cases}$$

  - If v_i, v_j belong to different communities

$$\rho_{ij} = \begin{cases} R_i^\top R_j & \text{with probability } q \\ \mathrm{Unif}(O(d)) & \text{with probability } 1-q \end{cases}$$

- "Perturbed case $(q > p > 0)$": synchronization solution

$$\tilde{\rho}_{ij} \approx R_i^\top R_j$$

- But $\|\tilde{\rho}_{ij} - \rho_{ij}\|$ deviates from 0 more frequently on edges connecting different communities!
- suggesting normalized graph cut with new weights $\|\tilde{\rho}_{ij} - \rho_{ij}\|$

Simultaneous Matching and Clustering (SMAC)

- A first attempt at implementing the program of SBM & LGA
- Established results for permutation groups in the context of collection shape matching
- SBM+Random Corruption for the maps, Erdős-Rényi for the observation graph, $K \geq 2$ balanced clusters, synchronization with spectral methods
- The recovery rate for the underlying clusters matches the information theoretic lower bound in [Chen et al. 2017]

• Yuxin Chen, Changho Suh, and Andrea J. Goldsmith. "Information Recovery from Pairwise Measurements." *IEEE Transactions on Information Theory* 62, no. 10 (2016): 5881-5905.
• Chandrajit Bajaj, **Tingran Gao**, Zihang He, Qixing Huang, and Zhenxiao Liang. "SMAC: Simultaneous Mapping and Clustering Using Spectral Decompositions." *2018 International Conference on Machine Learning.* (ICML 2018)

Algorithm 1 PermSMAC: Simultaneously mapping and clustering

Input: Observation graph $\mathcal{G} = (\mathcal{S}, \mathcal{E})$ and initial pairwise maps $X_{ij}^{in}, (i,j) \in \mathcal{E}$

Output: Underlying clusters $\mathcal{S} = c_1 \cup \cdots \cup c_k$ and optimized pairwise maps $X_{ij}, 1 \leq i, j \leq n$

1: {**Step 1**} Simultaneously compute the intra-cluster maps and extract the underlying clusters:

2: {**Step 1.1**} Form data matrix based on (1).

3: {**Step 1.2**} Compute the critical value $r = \underset{2 \leq i \leq nm}{\text{argmax}} \frac{\lambda_i - \lambda_{i+1}}{\lambda_i + \lambda_{i+1}}$.

4: {**Step 1.3**} Let $U \in \mathbb{R}^{nm \times r}$ store the leading r eigenvectors of X. Compute pair-wise maps $X_{ij}^{\star}$ by solving (2)

5: {**Step 1.4**} Use $f_{ij}(X_{ij}^{\star})$ as the affinity score and apply single-linkage clustering to obtain the underlying clusters

6: {**Step 2**} compute the inter-cluster maps by solving (6)

SMAC: Exact Recovery Conditions

Theorem (Bajaj et al. 2018). Under mild assumptions, `PermSMAC` recovers the underlying clusters and intra-cluster maps with high probability provided

$$p - q \geq ck\sqrt{\frac{\log n}{nt}}$$

for some constant $c > 0$, where $t = \Omega(\log(n)/n)$. The inter-cluster maps are recovered with high probability provided

$$q \geq c'k\sqrt{\frac{\log n}{n^2 t}}$$

for some constant $c' > 0$.

• Chandrajit Bajaj, **Tingran Gao**, Zihang He, Qixing Huang, and Zhenxiao Liang. "SMAC: Simultaneous Mapping and Clustering Using Spectral Decompositions." *2018 International Conference on Machine Learning.* (ICML 2018)

Summary: Towards a Learning Theory on Sheaves?

- Identify a sheaf in the wild world of applied harmonic analysis
- Sheaves encode abundant structure information
- Open questions from a learning perspective
  - Statistical paradigm for **sheaf learning**?
  - Fundamental (information-theoretic) limits?
  - Cheeger's inequality for learning group actions?
  - Spectral sheaf theory?

Thank You!

• **T. Gao**. "The Diffusion Geometry of Fibre Bundles." *submitted.* arXiv:1602.02330, (2016)

• **T. Gao**, J. Brodzki, S. Mukherjee. "The Geometry of Synchronization Problems and Learning Group Actions." *submitted.* arXiv:1610.09051, (2016)

• **T. Gao**, G.S. Yapuncich, I. Daubechies, S. Mukherjee, and D.M. Boyer, "Development and Assessment of Fully Automated and Globally Transitive Geometric Morphometric Methods, with Application to a Biological Comparative Dataset with High Interspecific Variation." *The Anatomical Record.* Volume 301, Issue 4, Pages 636-658. (2018)

• C. Bajaj, **T. Gao**, Z. He, Q. Huang, and Z. Liang. "SMAC: Simultaneous Mapping and Clustering Using Spectral Decompositions." *2018 International Conference on Machine Learning.* (2018)