Topological Dimensionality Reduction

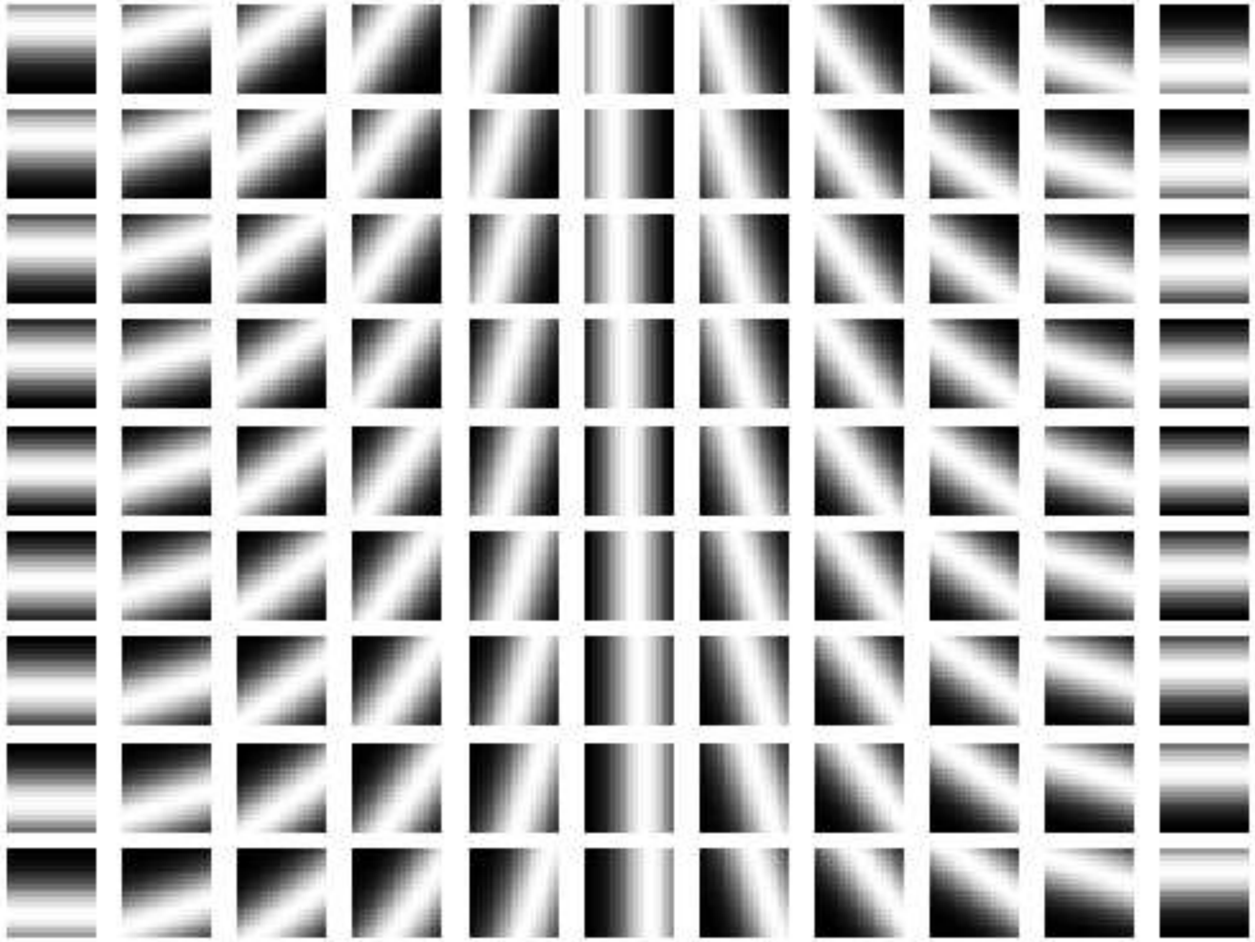
Jose Perea

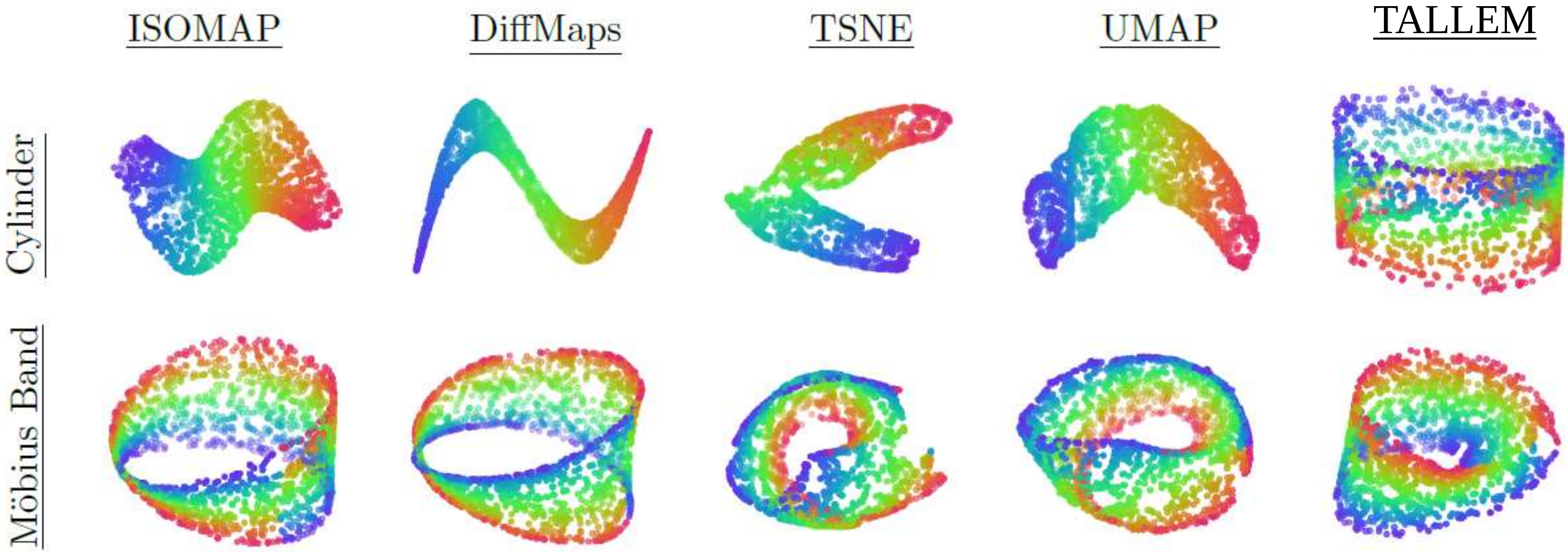
Dpt. of Computational Mathematics, Science & Engineering (CMSE)

Dpt. of Mathematics

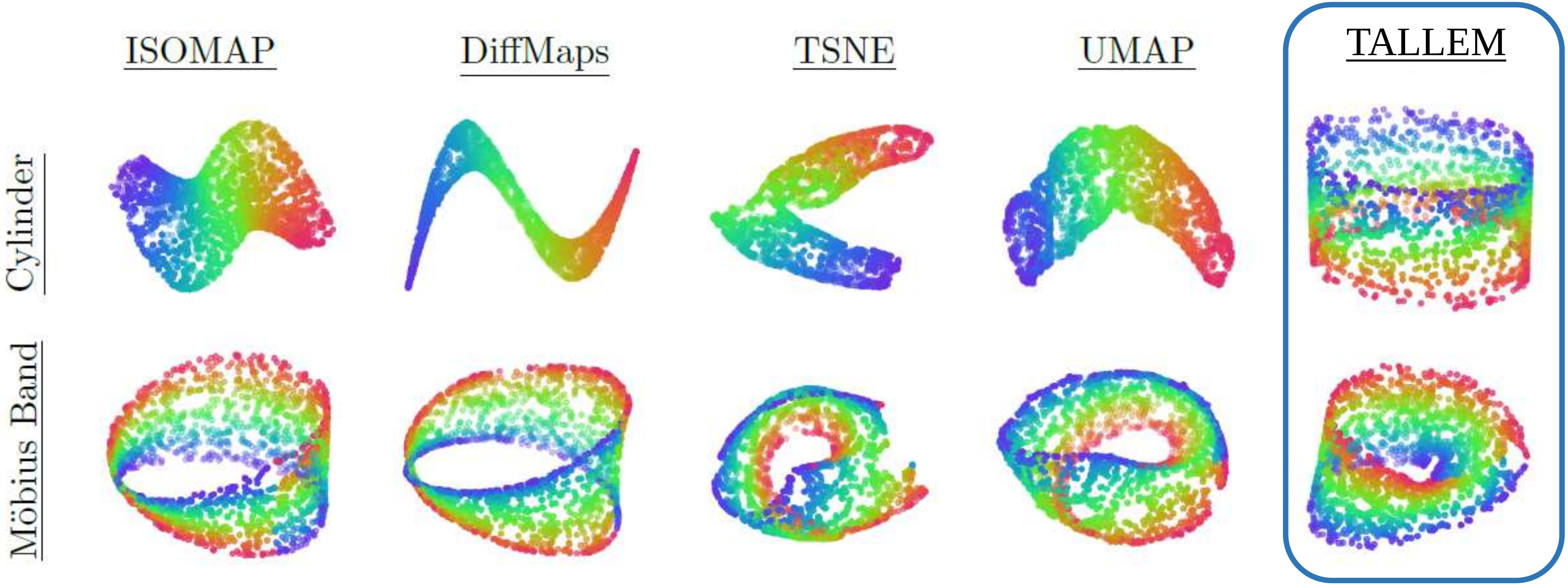
Twisted Cylinder $\subset \mathbb{R}^3$

Mobius Band $\subset \mathbb{R}^{n^2}$




ISOMAP
DiffMaps
TSNE
UMAP
TALLEM
Cylinder
Möbius Band

Topological Assembly of Locally Euclidean Models

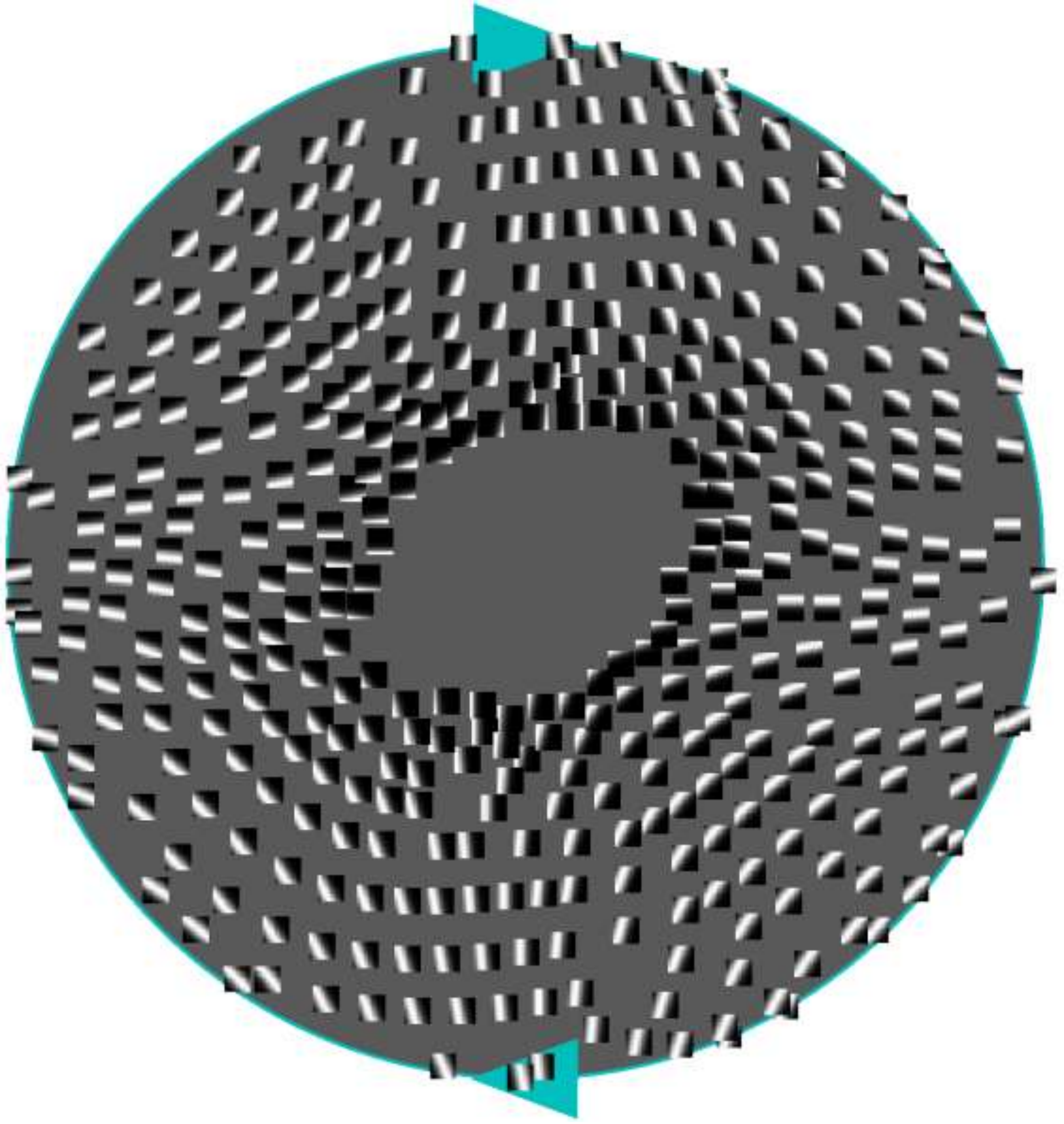


Metric ↔ Topological

TALLEM: Topological Alignment of Locally Euclidean Models, J. Mike & J. Perea, **Preprint**, 2019

$\mathbb{RP}^2$

Projective coordinates



Multiscale Projective Coordinates via Persistent Cohomology, J. Perea, **DC&G**, 2018

Eilenberg-MacLane Coordinates

Theorem (Brown representability):

CW-complex

Homotopy classes of maps

$$H^n(B;G) \cong [B, K(G,n)]$$

Abelian group

Eilenberg-MacLane Space

Theorem (Brown representability):

$$H^n(B;G) \cong \left[B, K(G,n)\right]$$

Where

$$\pi_j\left(K(G,n)\right) = \begin{cases} G & , \ j = n \\ \{0\} & , \ j \neq n \end{cases}$$

Theorem (Brown representability):

$$H^n(B;G) \cong [B, K(G,n)]$$

$$f^*(1_G) \qquad\qquad f : B \longrightarrow K(G,n)$$

$$\nwarrow\!\!\!\!\dashv \qquad\qquad\qquad \downarrow$$

$$f^* : H^n(K(G,n);G) \longrightarrow H^n(B;G)$$

$$\wr\|$$

$$1_G \in \mathsf{Hom}(G,G)$$

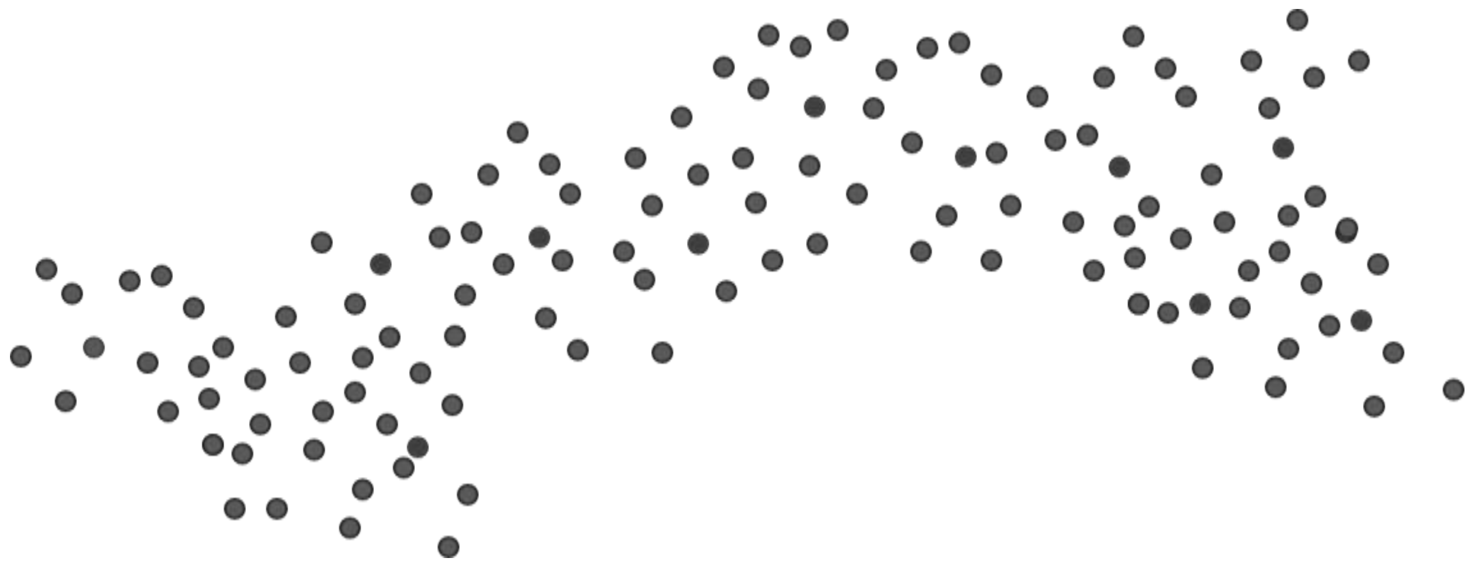
Eilenberg-MacLane Coordinates

$$f^*(1_G) \qquad\qquad f : B \longrightarrow K(G,n)$$

$$H^n(B;G) \cong [B, K(G,n)]$$

n	1	1	1	2
G	$\mathbb{Z}$	$\mathbb{Z}_2$	$\mathbb{Z}_q$	$\mathbb{Z}$
$K(G,n)$	S^1	$\mathbb{R}\mathbf{P}^\infty$	L_q^∞	$\mathbb{C}\mathbf{P}^\infty$

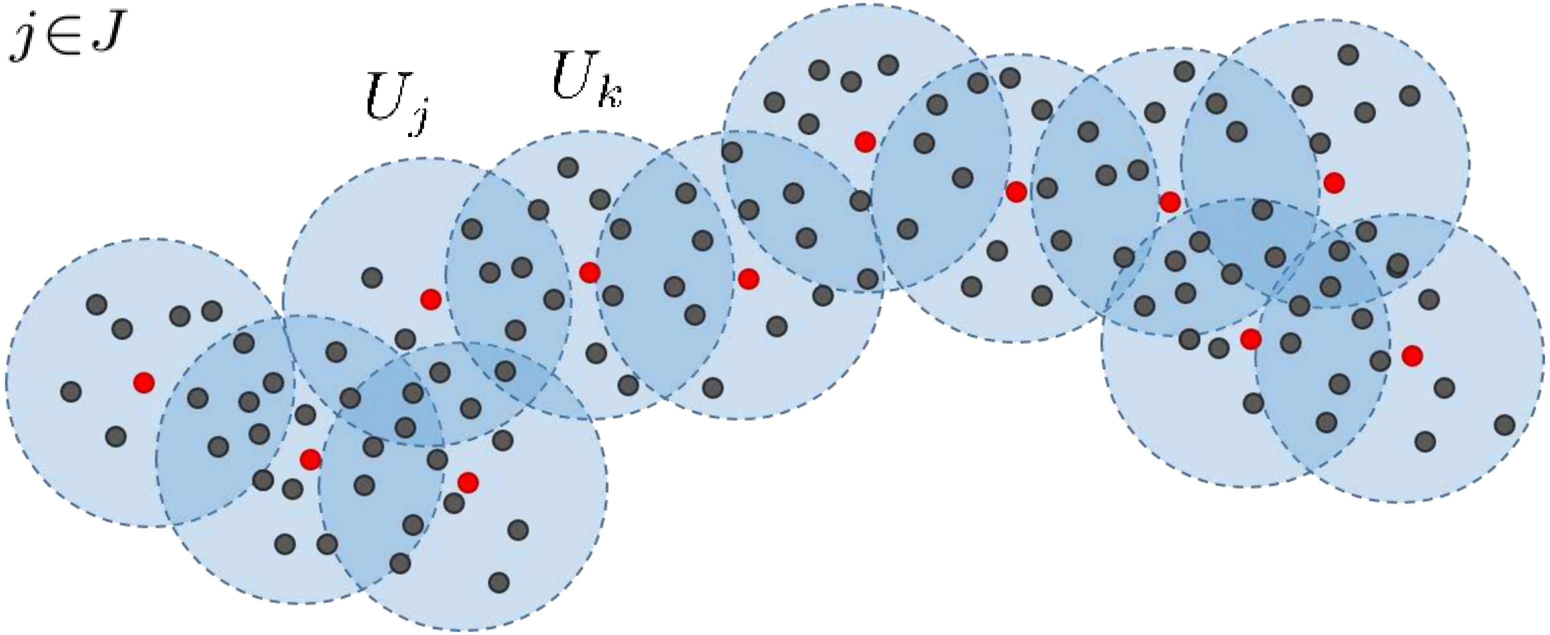
$X \subset (\mathbb{M}, \mathbf{d})$



$$L \subset X \subset (\mathbb{M}, \mathbf{d})$$

$$\mathcal{U} = \{U_j\}_{j \in J}$$

$$B = \bigcup \mathcal{U}$$



Partition of unity:

$$\varphi_j : B \longrightarrow [0,1] \quad j \in J$$

$$\sum_{j \in J} \varphi_j(b) = 1$$

$$\mathsf{supp}(\varphi_j) \subset \mathsf{clos}(U_j)$$

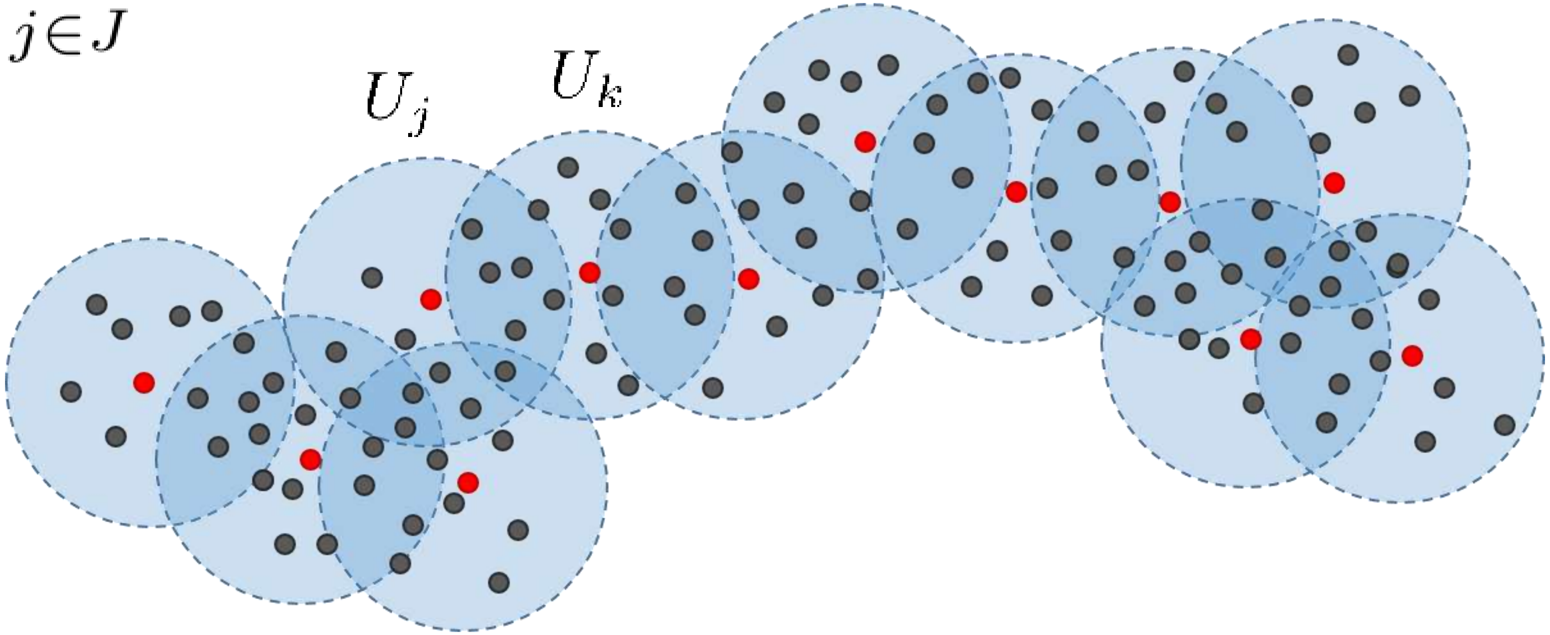
$$L \subset X \subset (\mathbb{M}, \mathbf{d})$$

$$\mathcal{U} = \{U_j\}_{j \in J}$$

$$B = \bigcup \mathcal{U}$$

$$\{\varphi_j\}_{j \in J}$$

Partition of 1



Example:

$$\varphi_j(b) = \frac{|\alpha - \mathbf{d}(b, x_j)|_+}{\sum_{k \in J} |\alpha - \mathbf{d}(b, x_k)|_+}$$

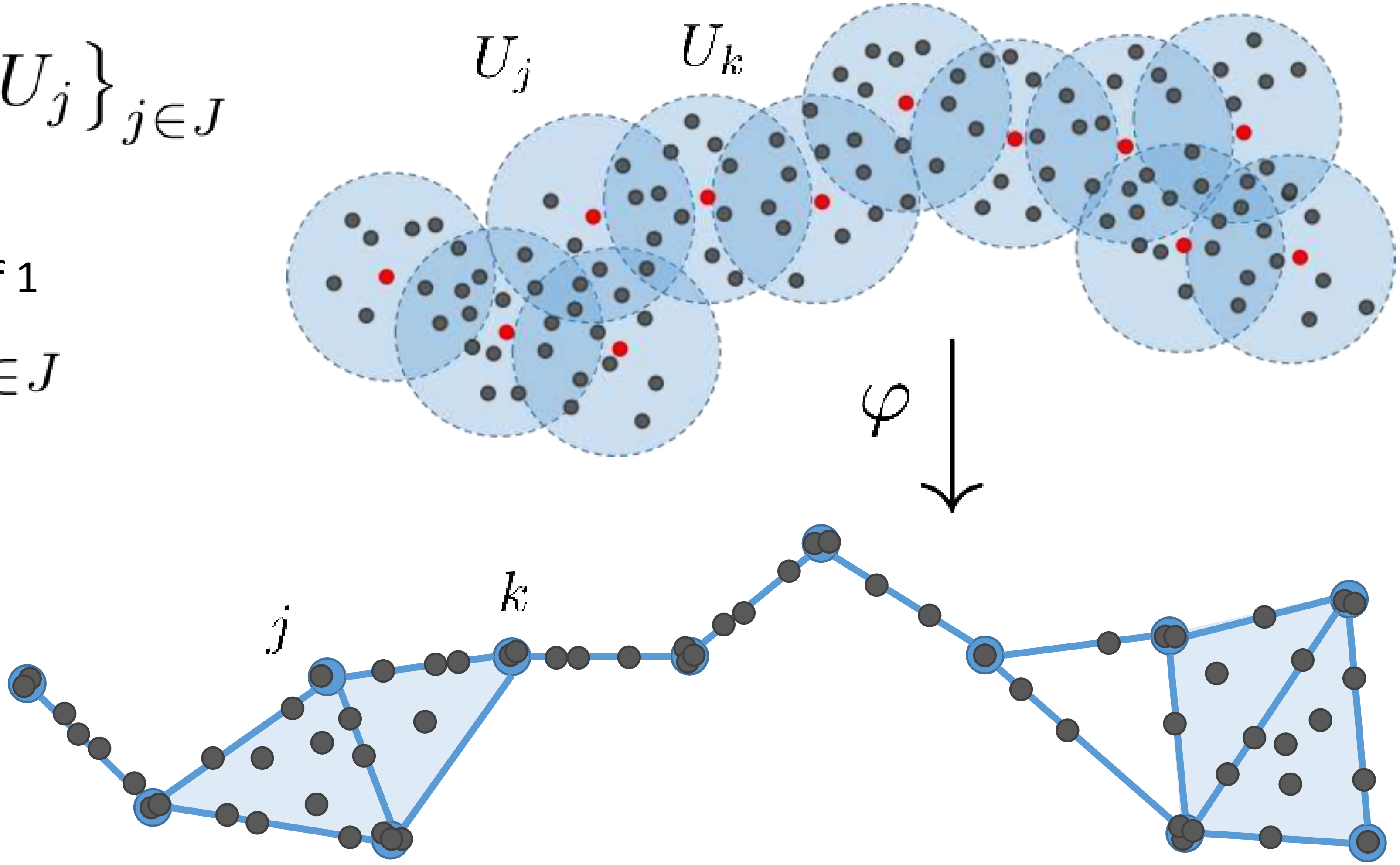
$$U_j = \{b \in B : \mathbf{d}(b, x_j) < \alpha\}$$
$$= B_\alpha(x_j)$$

$$|\lambda|_+ = \max\{\lambda, 0\} \quad , \quad \lambda \in \mathbb{R}$$

$$\mathcal{U} = \{U_j\}_{j \in J}$$

Partition of 1

$$\{\varphi_j\}_{j \in J}$$



$$B = \bigcup \mathcal{U}$$

$$\Downarrow$$

$$b$$

$$\downarrow$$

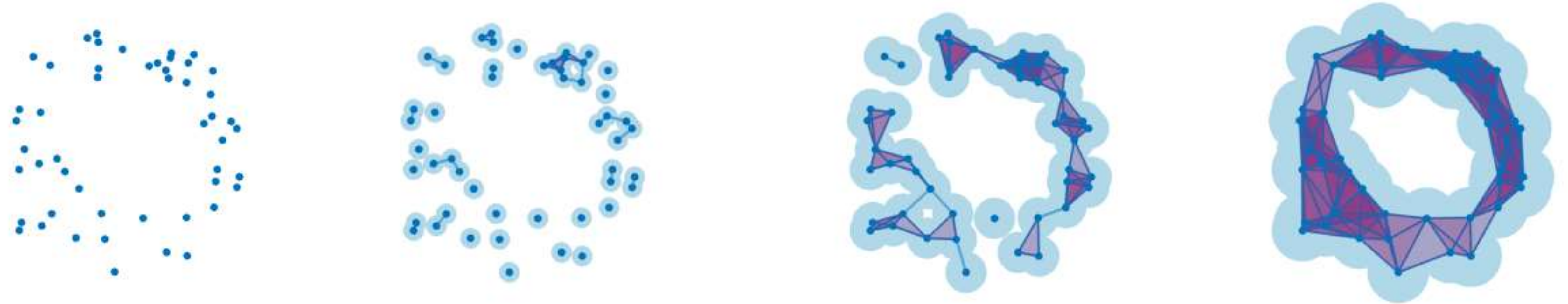
$$[\varphi_1(b), \ldots, \varphi_N(b)]$$

$$\cap$$

$$|\mathcal{N}(\mathcal{U})|$$

$$\#(J) = N$$

$$H^n(\mathcal{N}(\mathcal{U}); G) \xrightarrow{\varphi^*} H^n(B; G) \cong [B, K(G, n)]$$



$$(\mathbb{M}, \mathbf{d}) \supset L$$

$$\mathcal{N}(\mathcal{U}^{\alpha}) \quad \hookrightarrow \quad \mathcal{N}(\mathcal{U}^{\beta}) \quad \hookrightarrow \quad \mathcal{N}(\mathcal{U}^{\gamma})$$

$$\begin{array}{ccccccc}
0 \leftarrow & H^n(\mathcal{N}(\mathcal{U}^{\alpha}); G) & \leftarrow & H^n(\mathcal{N}(\mathcal{U}^{\beta}); G) & \leftarrow & H^n(\mathcal{N}(\mathcal{U}^{\gamma}); G) \\
& \downarrow & & \downarrow & & \downarrow \\
& [\cup\mathcal{U}^{\alpha}, K(G,n)] & \leftarrow & [\cup\mathcal{U}^{\beta}, K(G,n)] & \leftarrow & [\cup\mathcal{U}^{\gamma}, K(G,n)]
\end{array}$$

Circular Coordinates:

$$[\eta] \in H^1(\mathcal{N}(\mathcal{U}); \mathbb{Z}) \longrightarrow H^1(\mathcal{N}(\mathcal{U}); \mathbb{R})$$

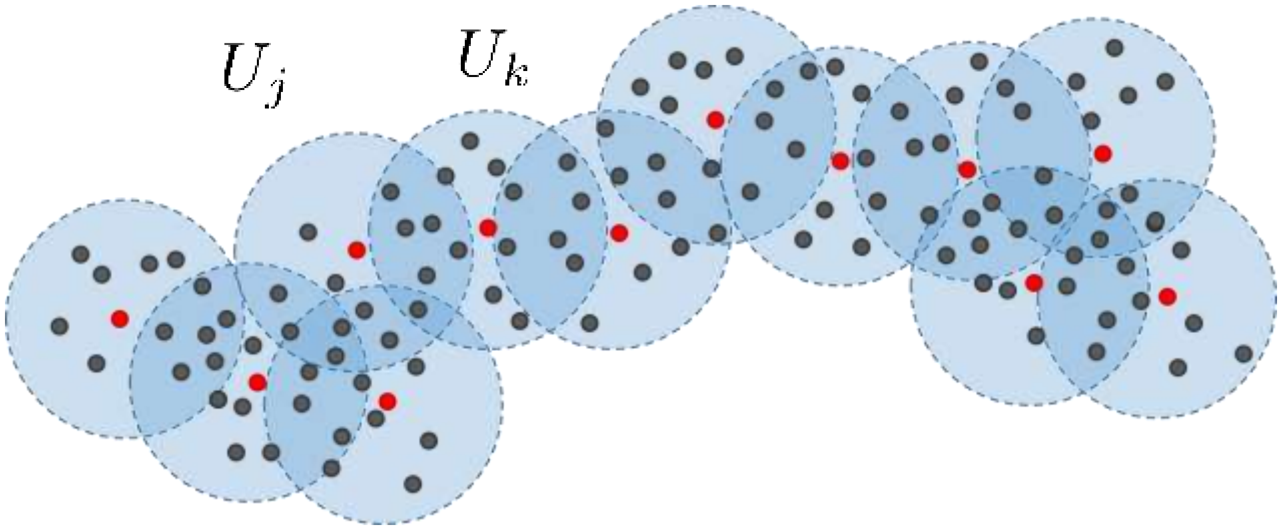
$$\theta = \eta + \delta^0 \tau$$

$$S^1 = K(\mathbb{Z}, 1)$$

$$f_\eta : \bigcup_{j=J} U_j \longrightarrow S^1$$

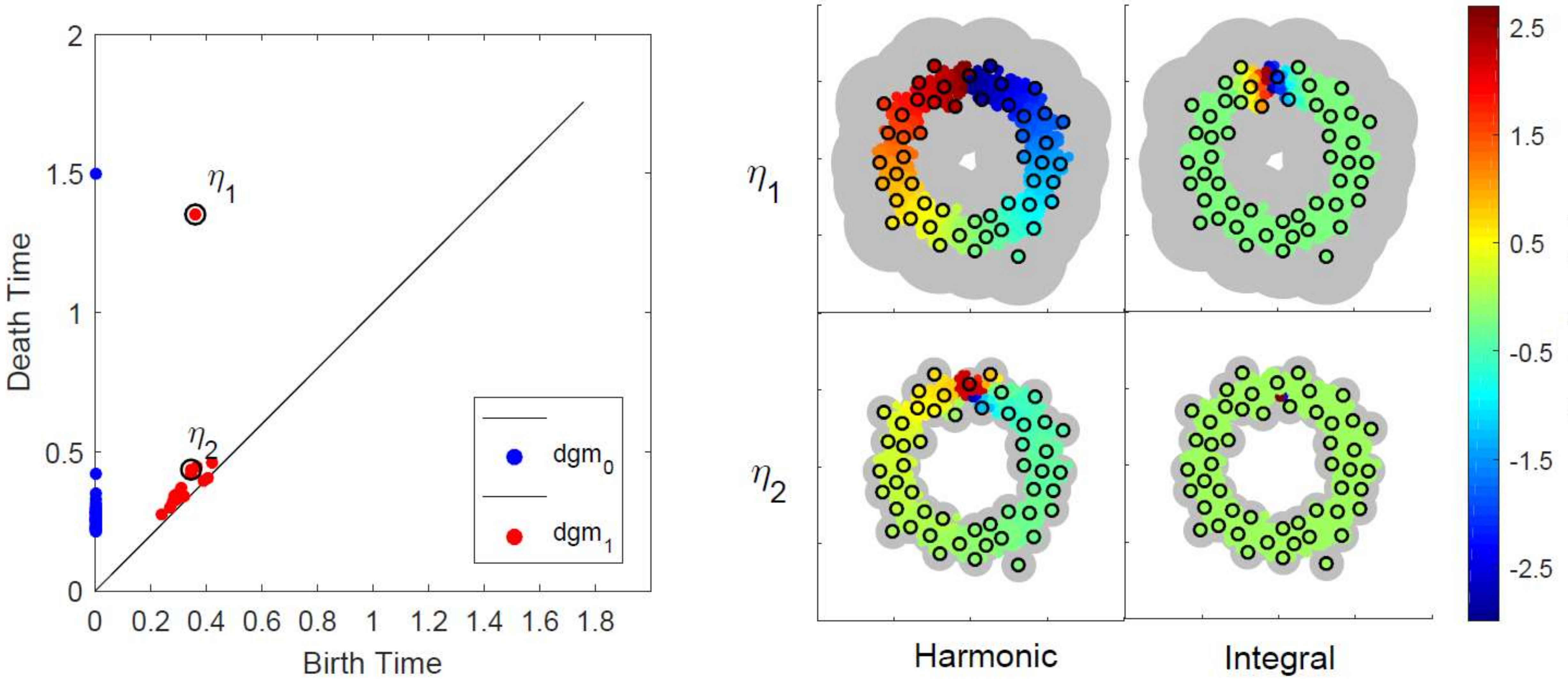
$$U_j \ni b \mapsto \exp\left\{2\pi i\left(\tau_j + \sum_k \varphi_k(b)\theta_{jk}\right)\right\}$$

$$\mathcal{U} = \{U_j\}_{j\in J}$$



Partition of 1

$$\{\varphi_j\}_{j\in J}$$



2
1.5
1
0.5
0
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
Death Time
Birth Time
η_1
η_2
dgm_0
dgm_1
Harmonic
Integral
2.5
1.5
0.5
-0.5
-1.5
-2.5
Circular Coordinate

Mumford Data

J. van Hateren natural (4,000) images dataset

Mumford Data

- Consider the set of high-contrast intensity-centered and contrast-normalized 3×3 patches from natural images.

$$\mathbf{v} \sim (\mathbf{v} + \lambda)$$

Intensity-centered

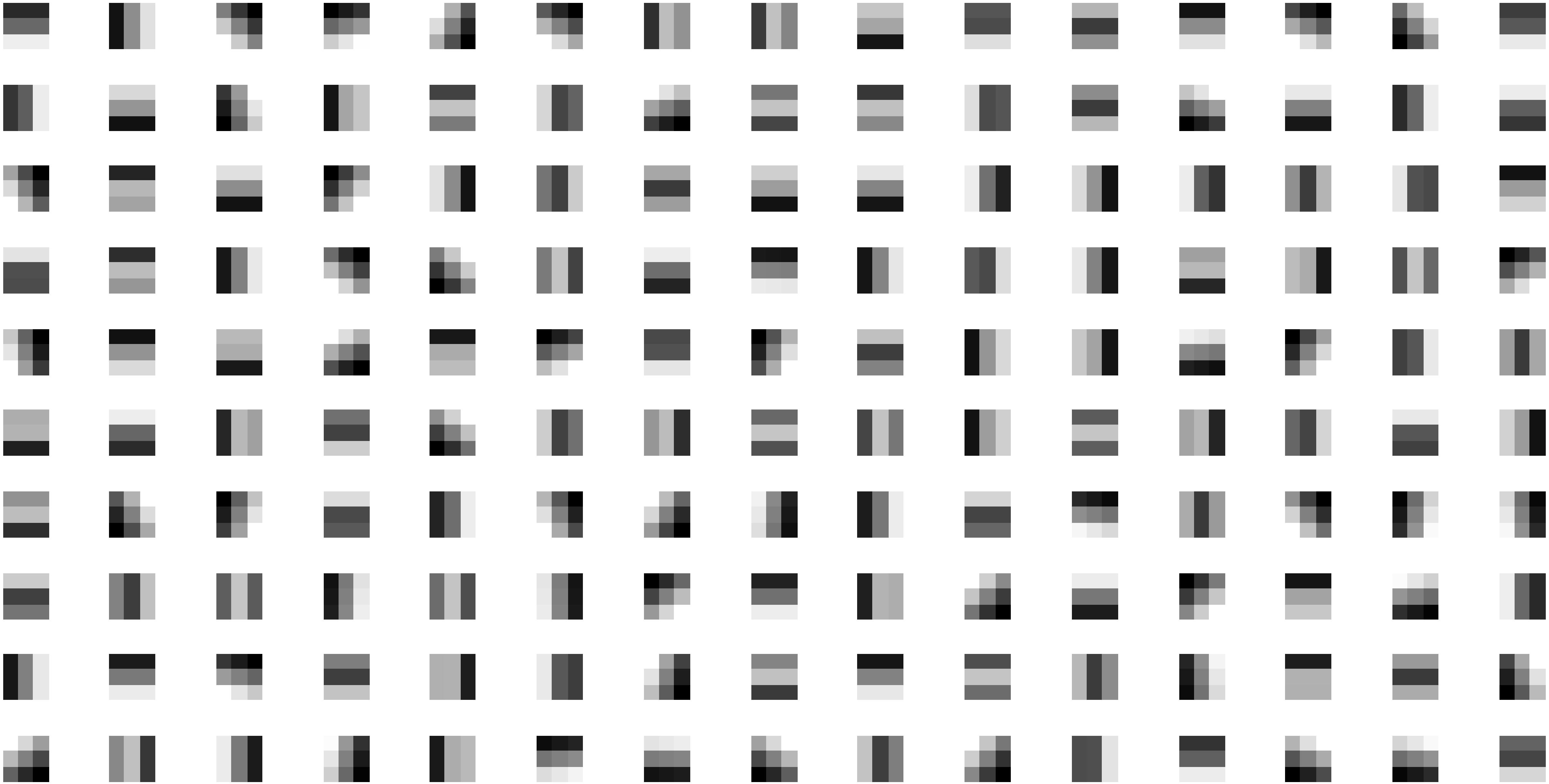
$$\mathbf{v} \sim \beta \mathbf{v}$$

Contrast-normalized

$$\mathbf{v} \in \mathbb{R}^9$$

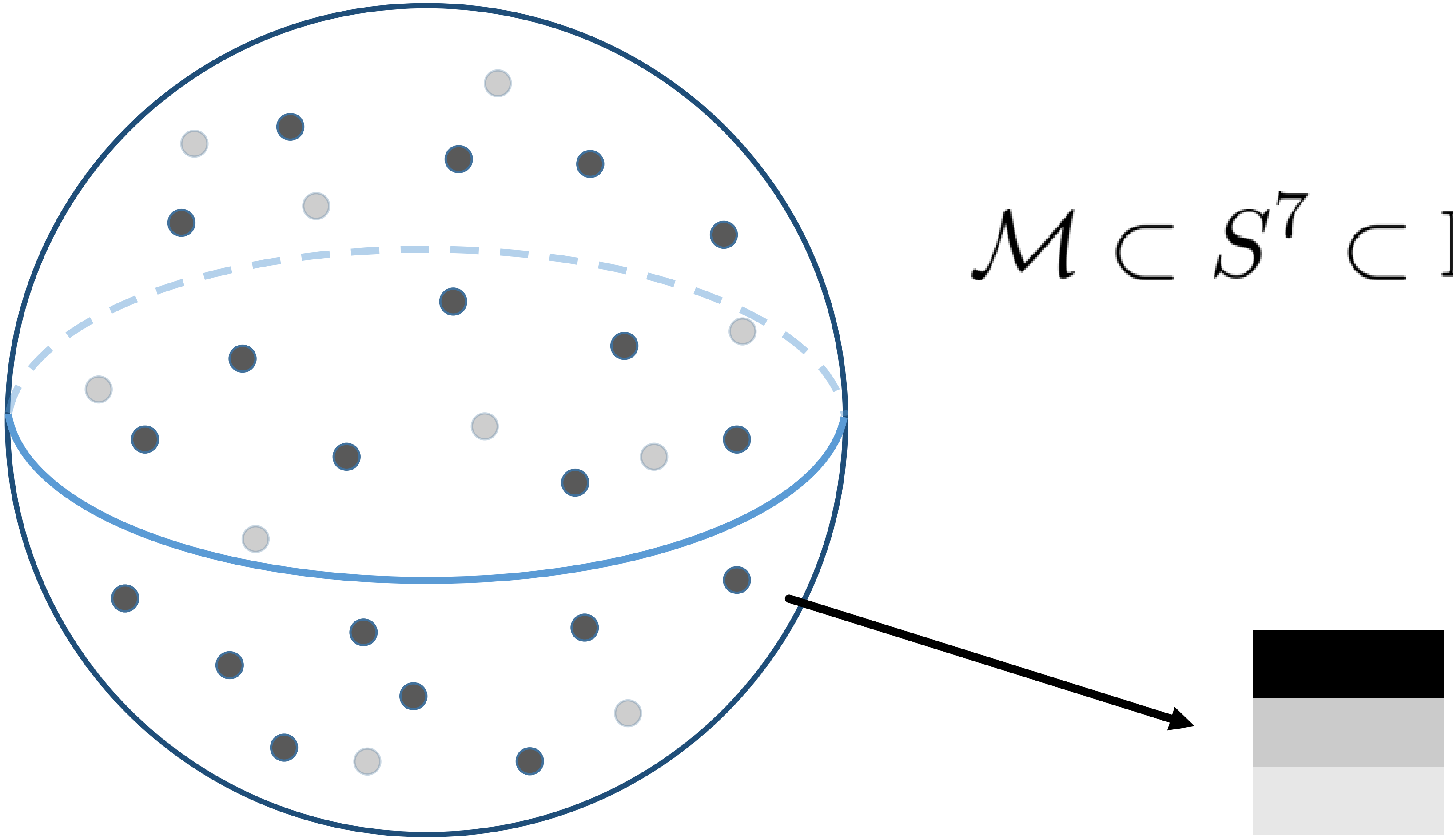
$$\lambda, \beta \in \mathbb{R}$$

The non-linear statistics of high-contrast patches in natural images, A. Lee et. al., 2002



$$\mathbf{v} \sim (\mathbf{v} + \lambda)$$

$$\mathbf{v} \sim \beta \mathbf{v}$$



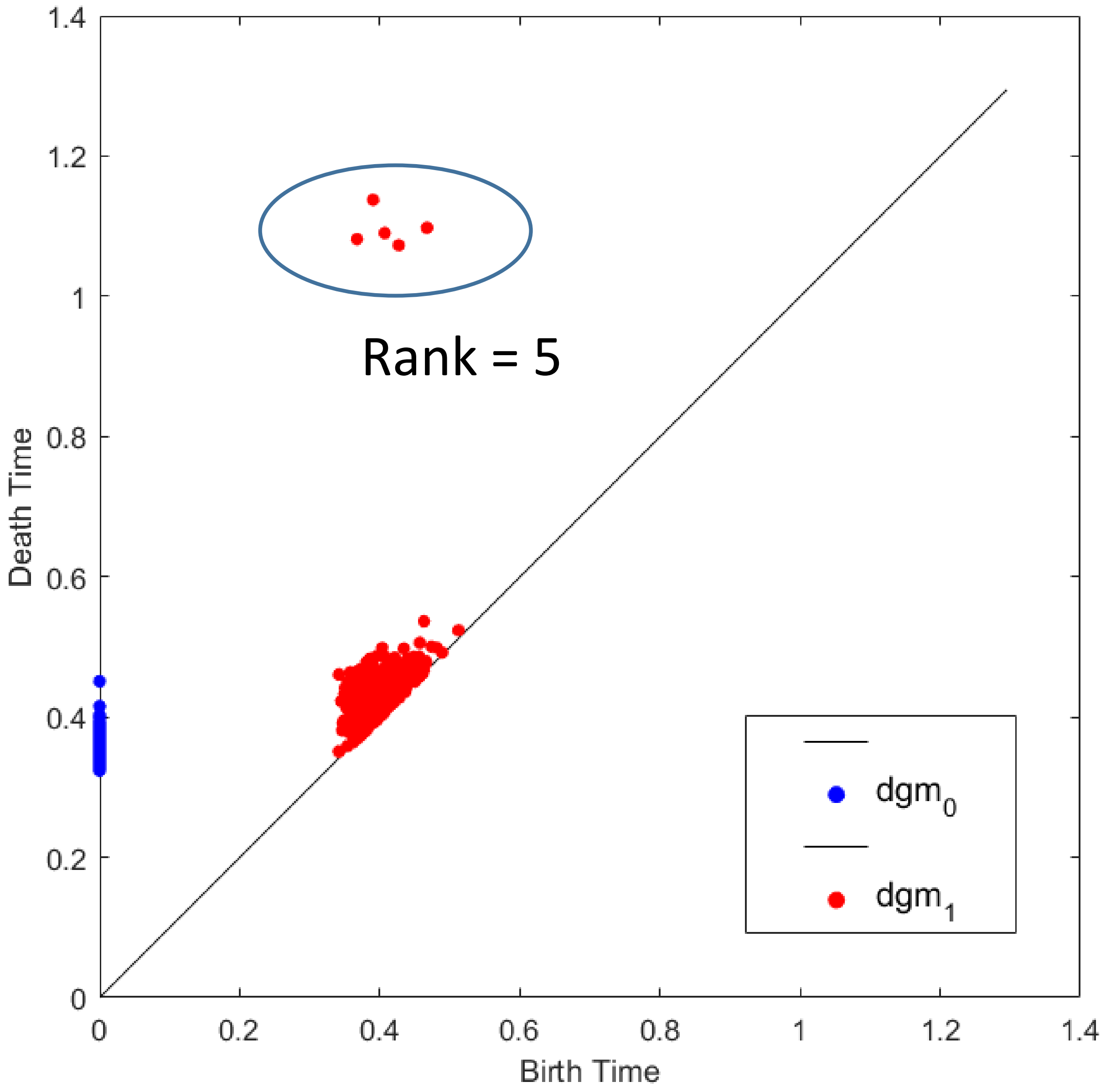
$$\mathcal{M} \subset S^7 \subset \mathbb{R}^9$$

Facts:

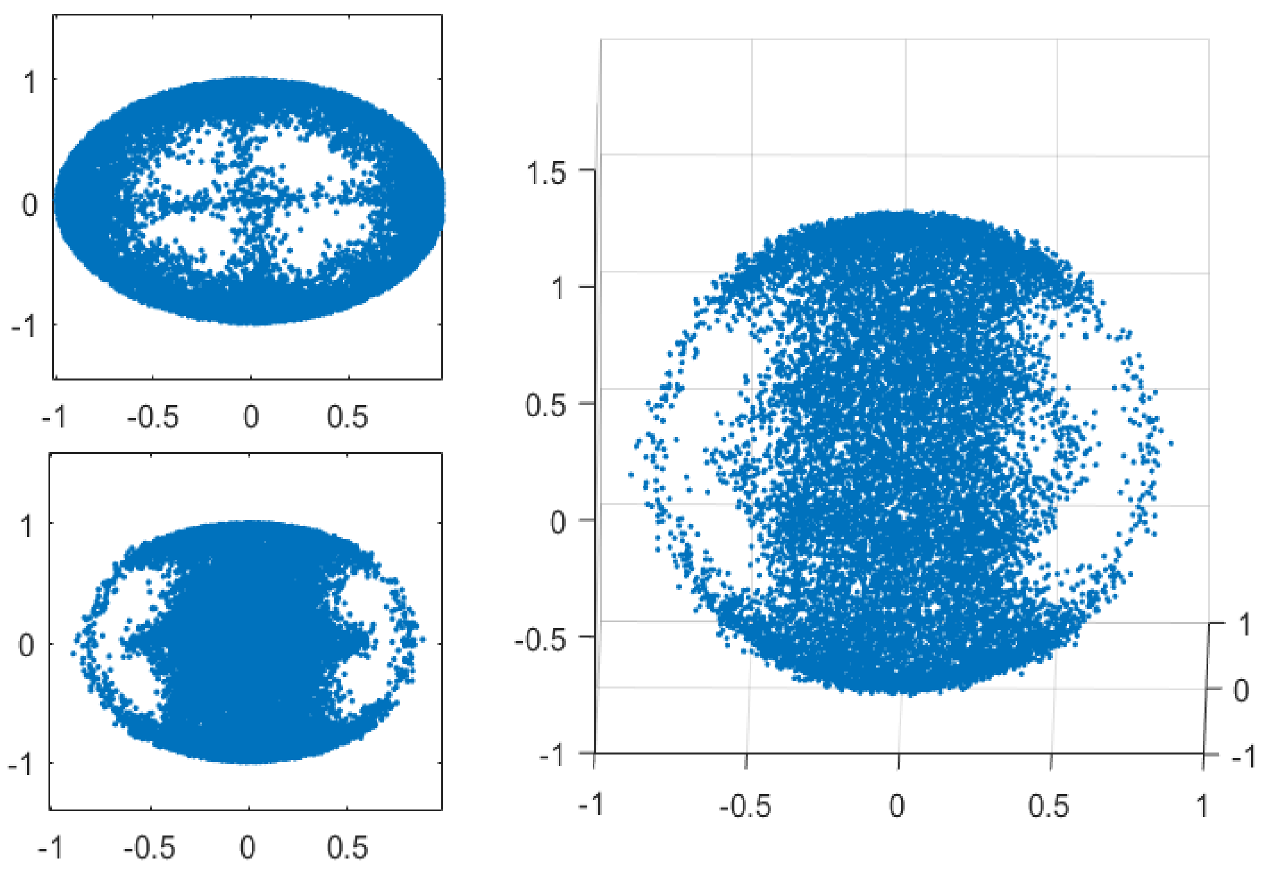
- $\mathcal{M}$ fills out S^7
- Not all regions of S^7 are equally densely populated

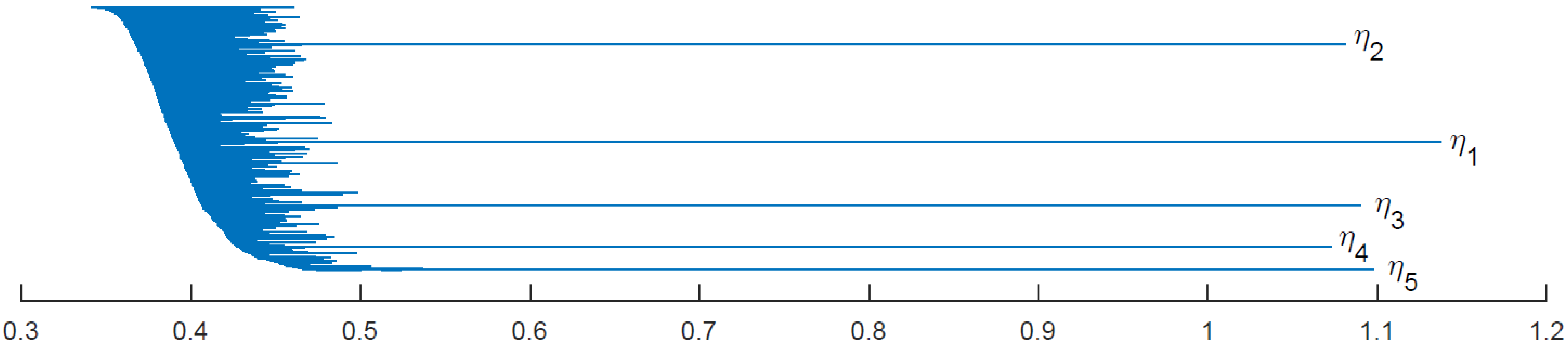
The non-linear statistics of high-contrast patches in natural images, A. Lee et. al., 2002

$X(15, 30)$ Top 30% densest points (w.r.t. distance to 15th nearest neighbor)

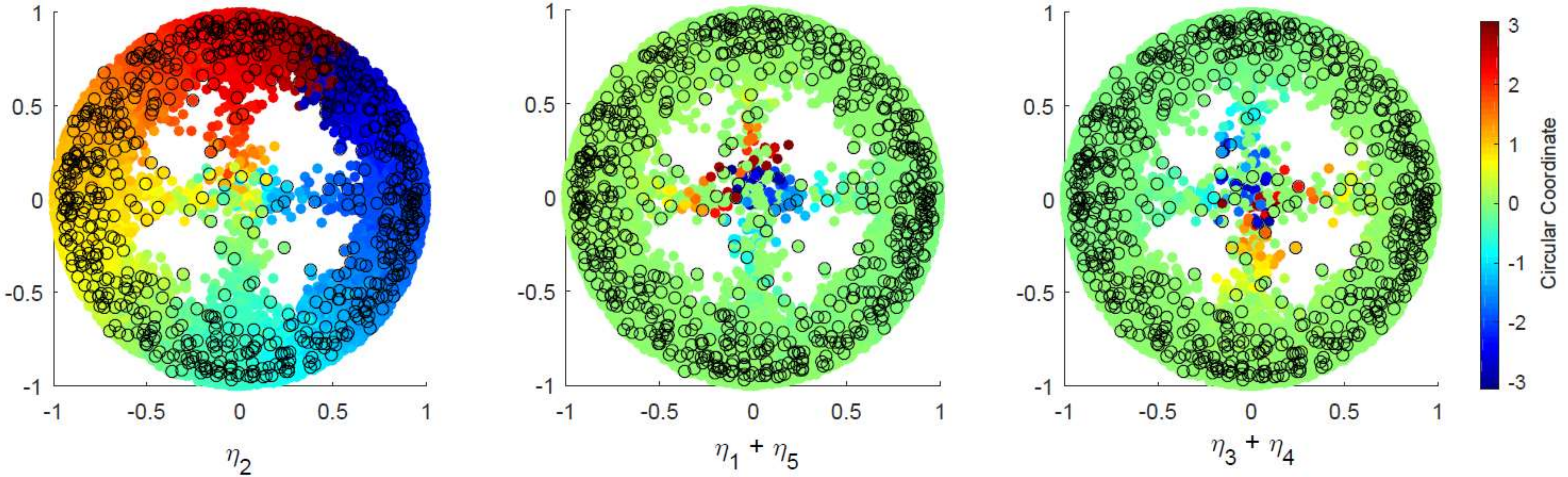


3D - PCA

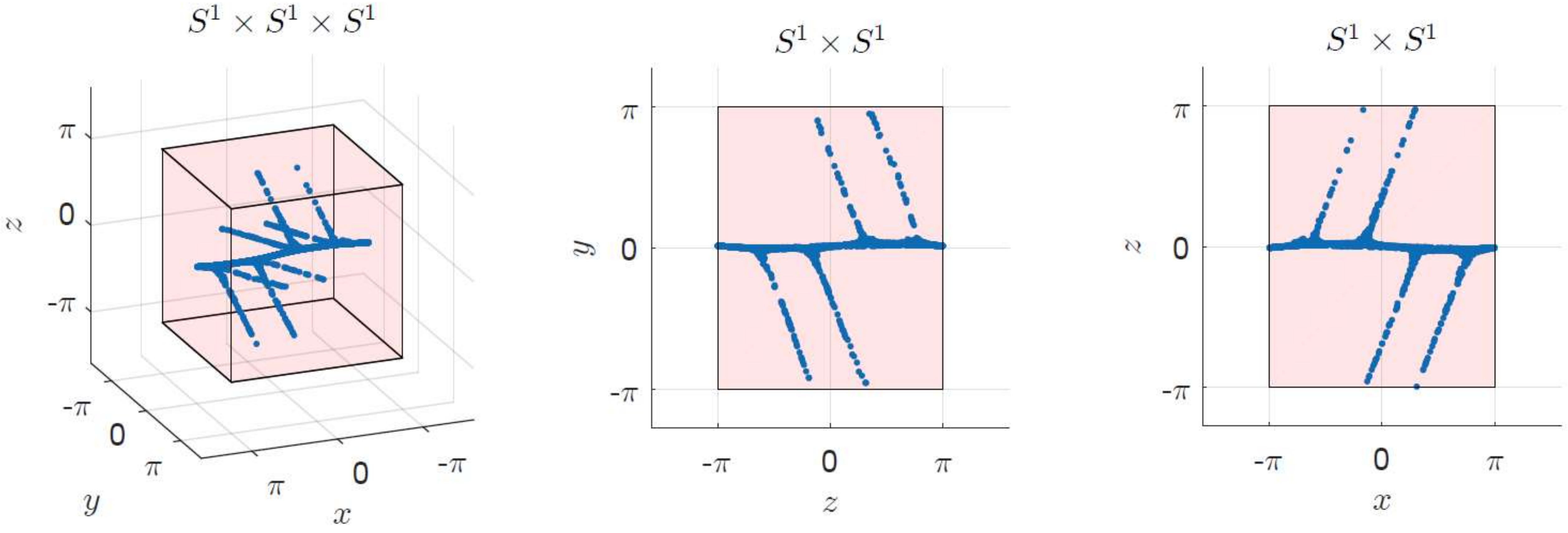


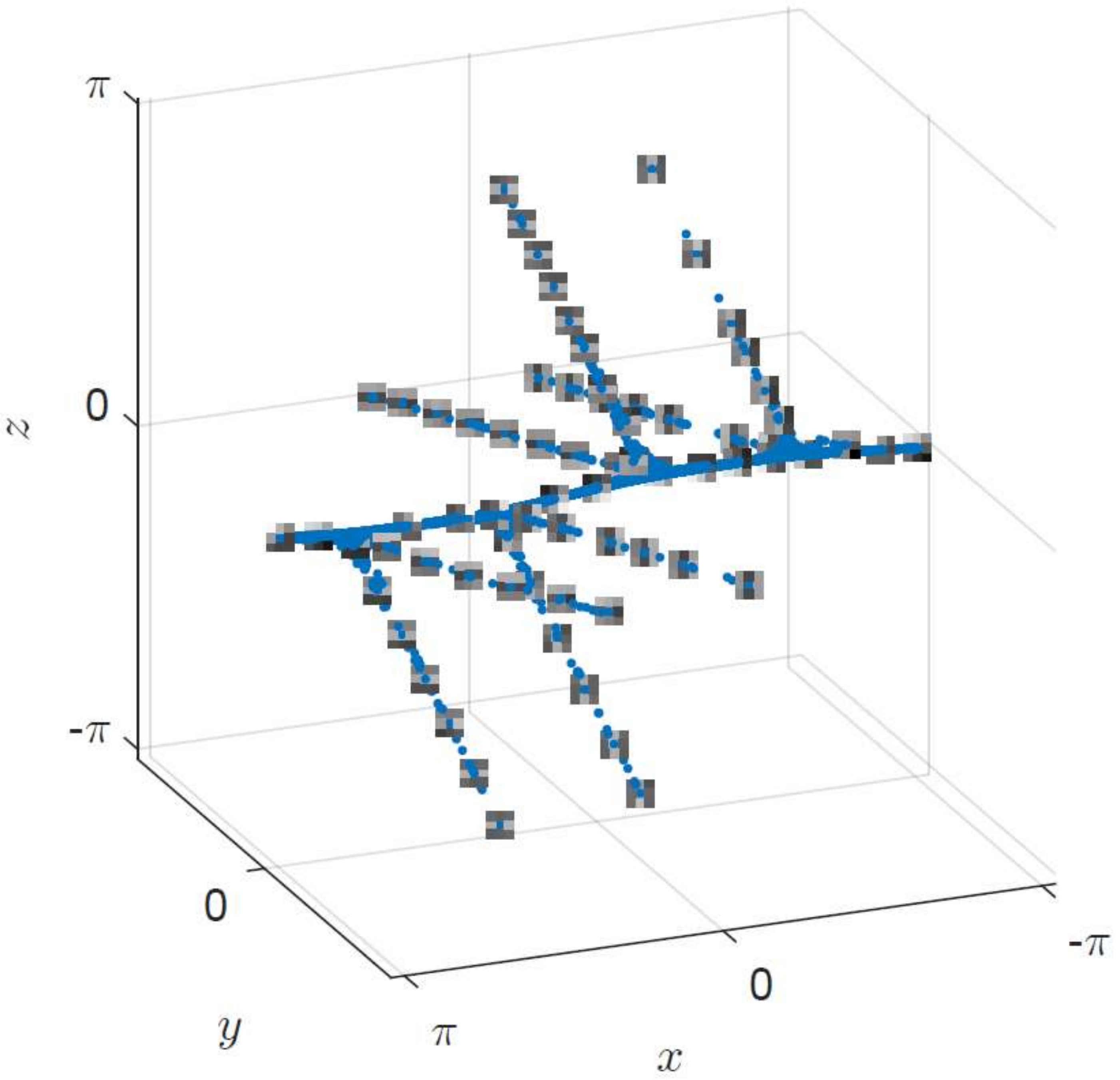


η_2
η_1
η_3
η_4
η_5
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
1.1
1.2

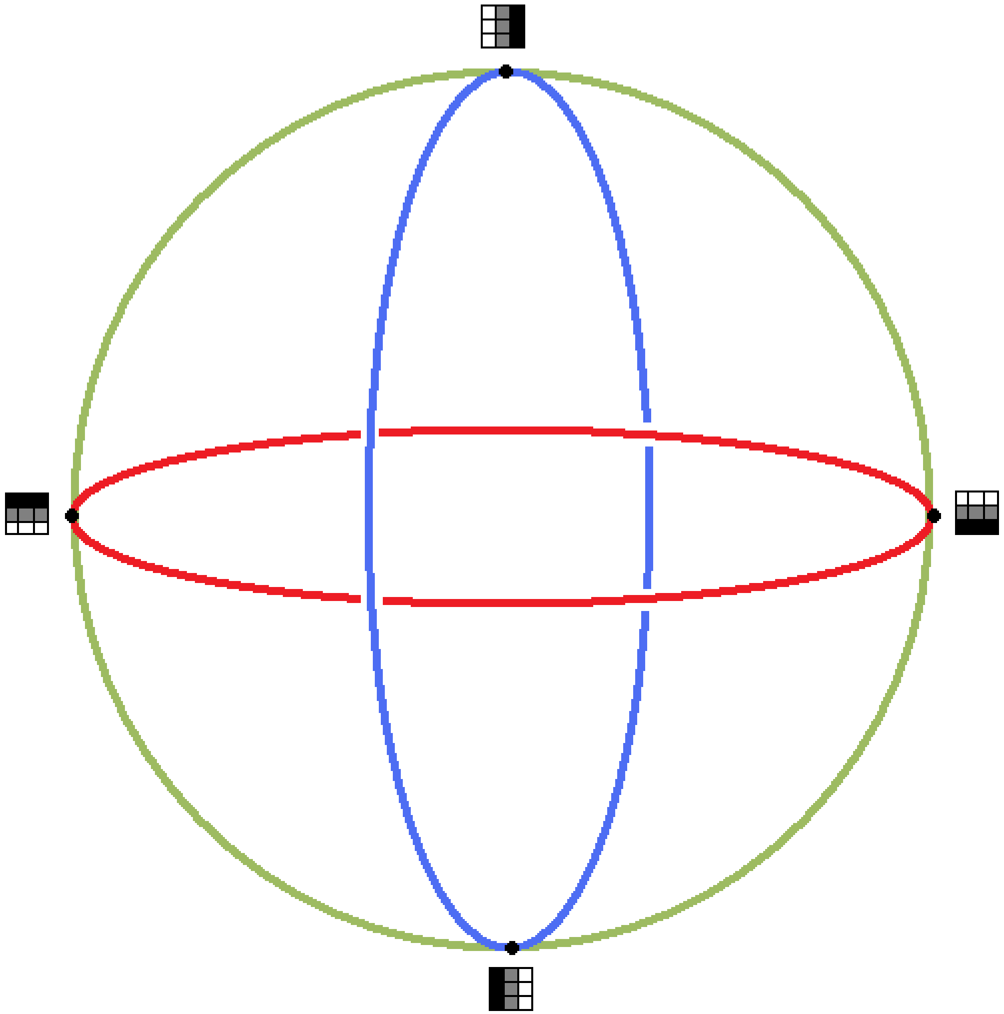


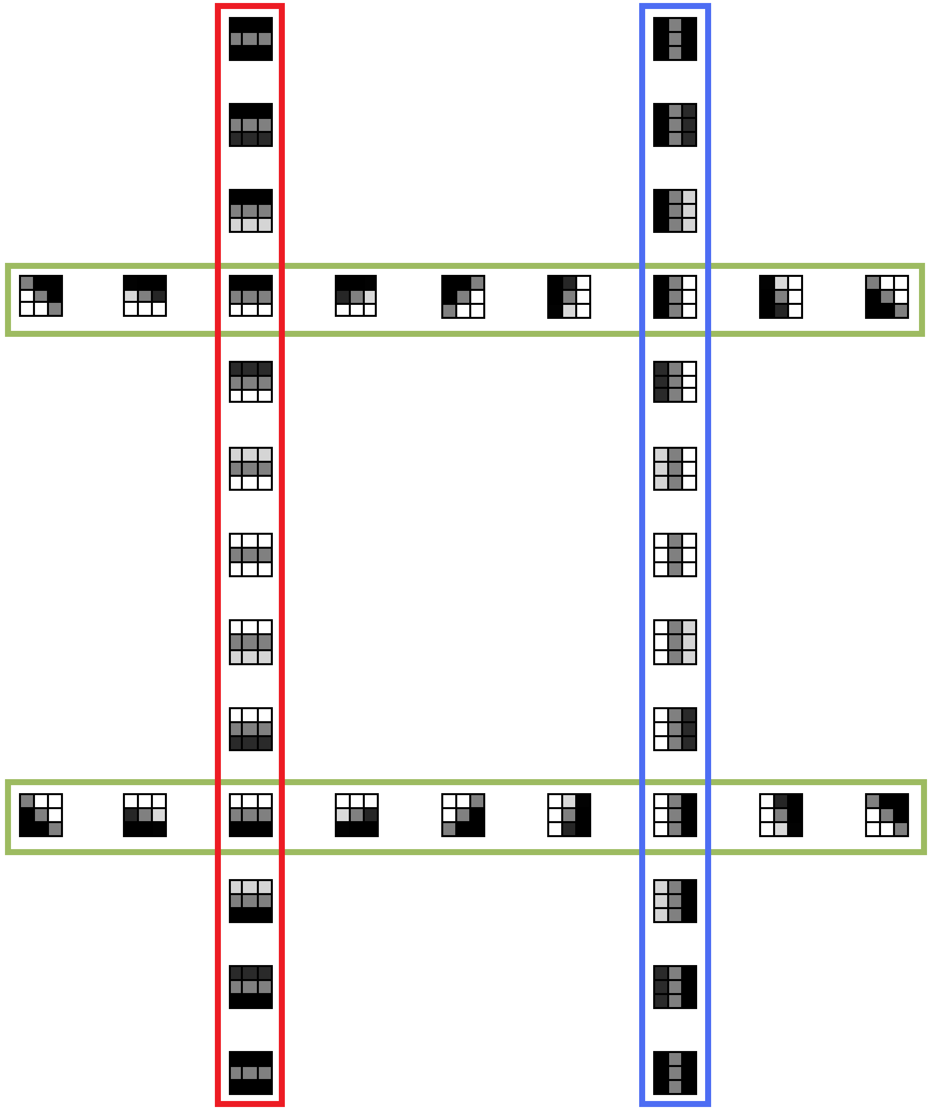
1
0.5
0
-0.5
-1
-1
-0.5
0
0.5
1
η_2
$\eta_1 + \eta_5$
$\eta_3 + \eta_4$
3
2
1
0
-1
-2
-3
Circular Coordinate


$S^1 \times S^1 \times S^1$
π
0
-π
z
-π
0
π
y
π
0
-π
x
$S^1 \times S^1$
π
0
-π
y
-π
0
π
z
$S^1 \times S^1$
π
0
-π
z
-π
0
π
x


π
0
-π
z
y
0
π
x
0
-π

The three circle model





On the Local Behavior of Spaces of Natural Images, Carlsson et al, 2008

Projective Coordinates ($\mathbb{R}$):

$$\mathbb{R}\mathbf{P}^\infty = K(\mathbb{Z}_2, 1)$$

$$[\theta] \in H^1(\mathcal{N}(\mathcal{U}); \mathbb{Z}_2)$$

$$\mathcal{U} = \{U_j\}_{j=0}^n$$

$$\{\varphi_j\}_{j=0}^n$$

Partition of 1

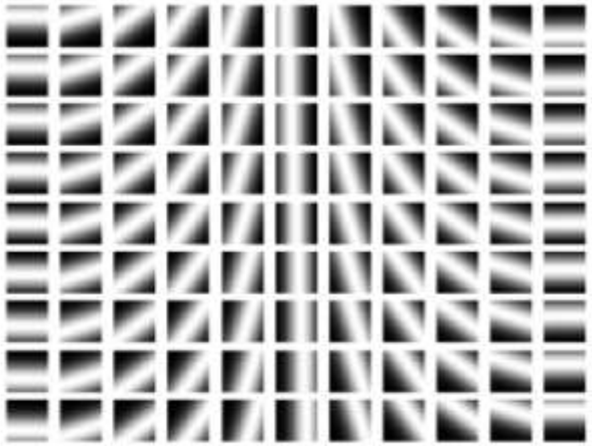
$$f_\theta : \bigcup \mathcal{U} \longrightarrow \mathbb{R}\mathbf{P}^n$$

$$U_j \ni b \mapsto \left[\theta_{j0}\sqrt{\varphi_0(b)} : \cdots : \theta_{jn}\sqrt{\varphi_n(b)}\right]$$

$\mathbb{R}\mathbf{P}^n$ Supports a version of Principal Component Analysis (Projective PCA)

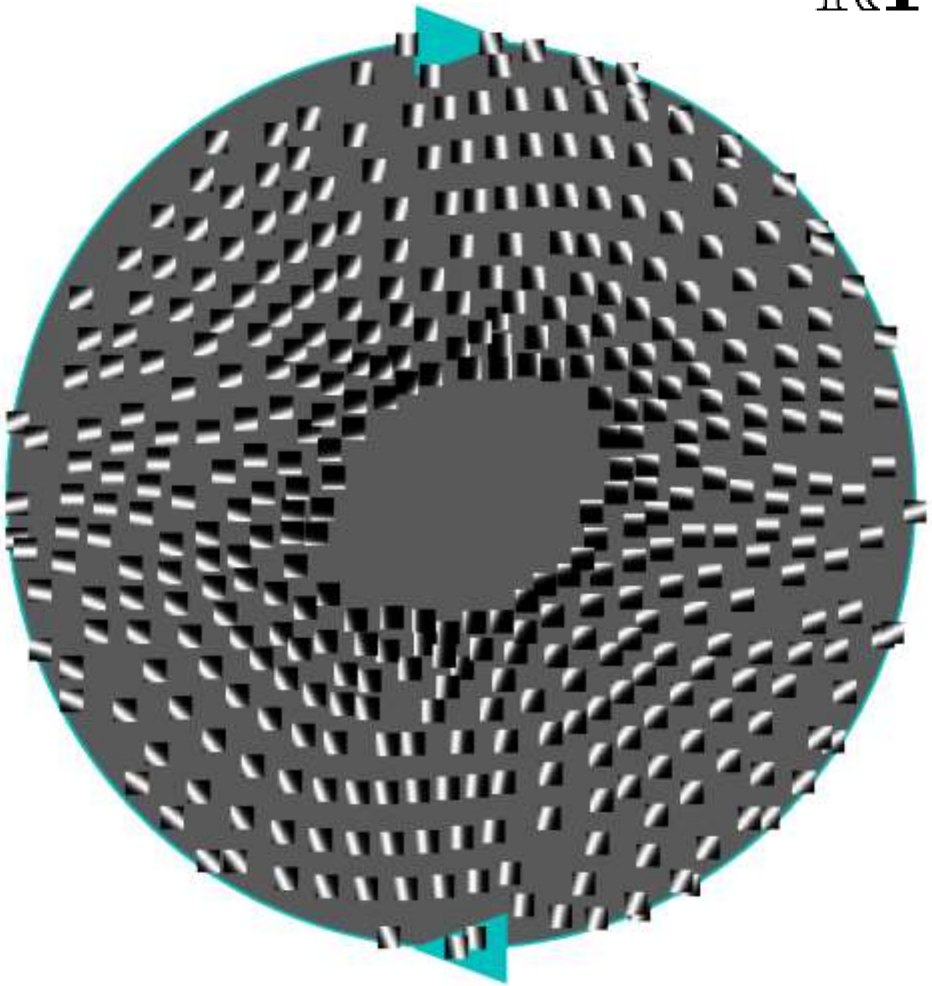
Multiscale Projective Coordinates via Persistent Cohomology, J. Perea, **DC&G**, 2018

Projective coordinates



Data $\subset \mathbb{R}^{n^2}$

$\mathbb{RP}^2$



Multiscale Projective Coordinates via Persistent Cohomology, J. Perea, **DC&G**, 2018

Lens Space Coordinates:

$$L_q^\infty = K(\mathbb{Z}_q, 1)$$

$$[\nu] \in H^1(\mathcal{N}(\mathcal{U}); \mathbb{Z}_q)$$

$$f_\nu : \bigcup_{j=1}^{N} U_j \longrightarrow L_q^N = S^{2N-1}/\mathbb{Z}_q$$

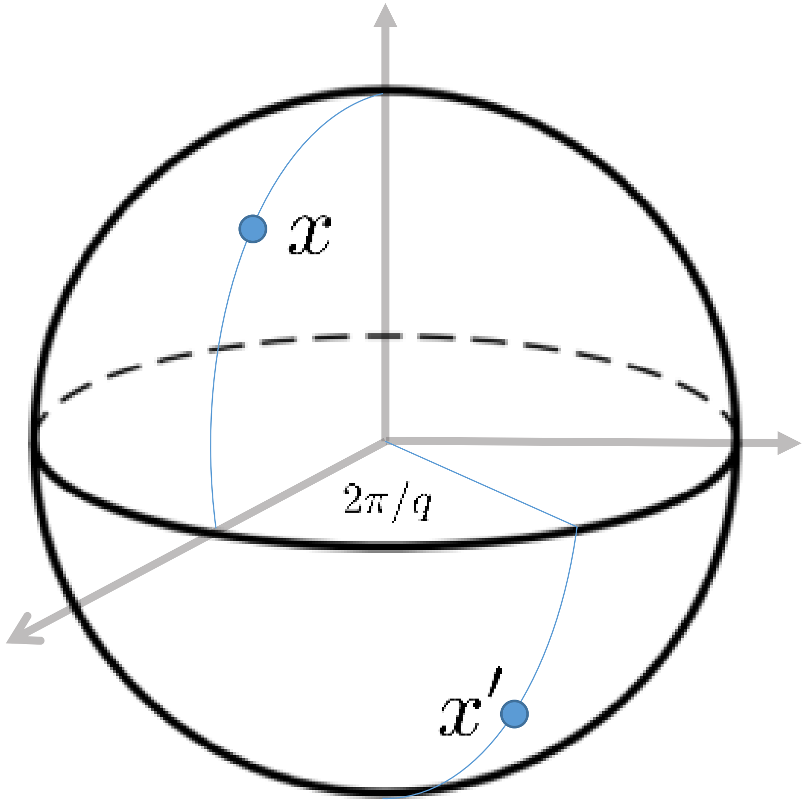
L_q^N Supports a version of Principal Component Analysis (Lens PCA)

Coordinatizing Data with Lens Spaces and Persistent Cohomology, L. Polanco & J. Perea, **CCCG**, 2019

Lens Spaces

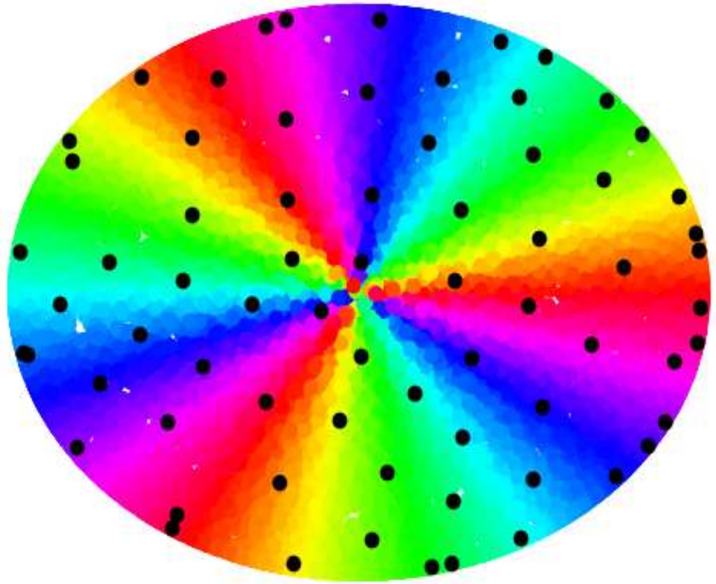
$$B^3 = \{x \in \mathbb{R}^3 : \|x\| \leq 1\}$$

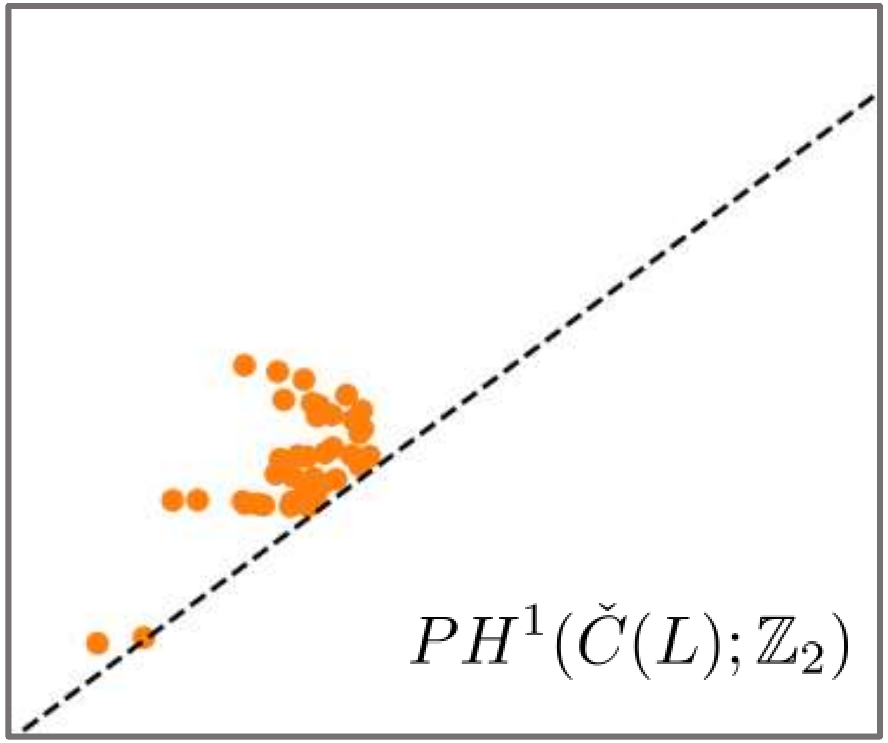
$$x \sim x'$$

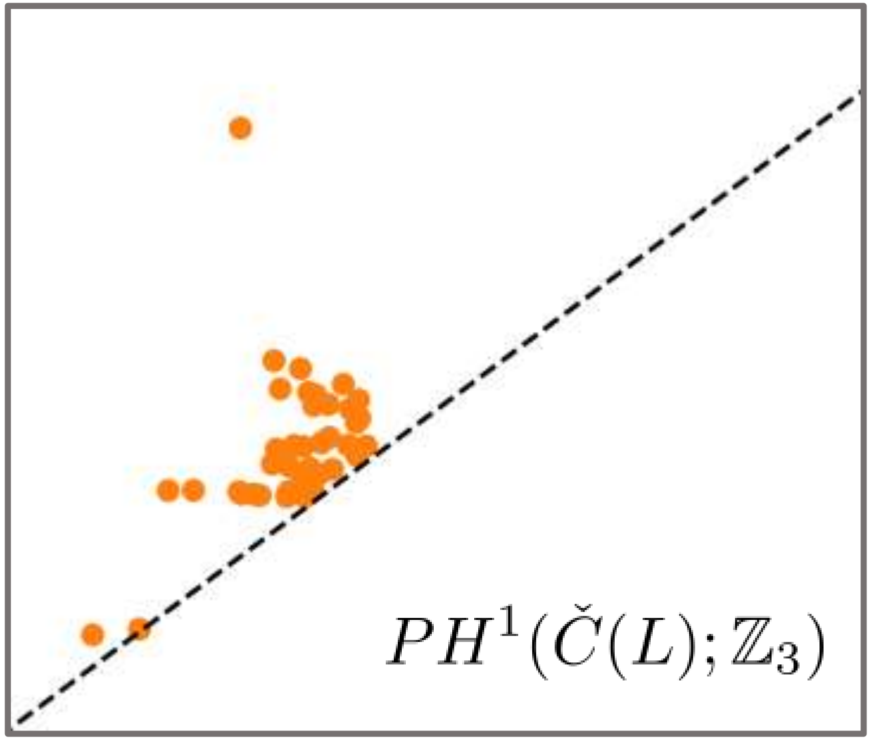


$$L_q^2 \cong B^3 / x \sim x'$$

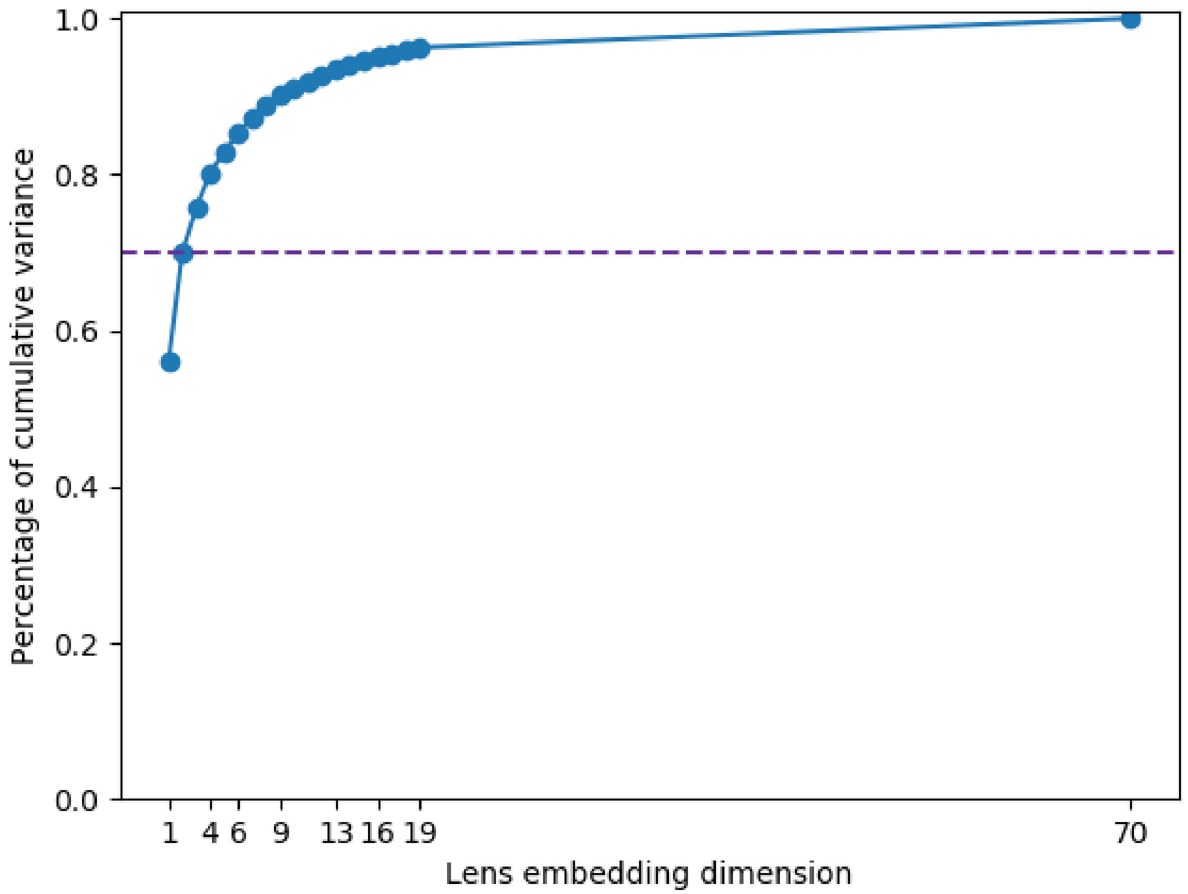
The Moore space $M(\mathbb{Z}_3, 1)$





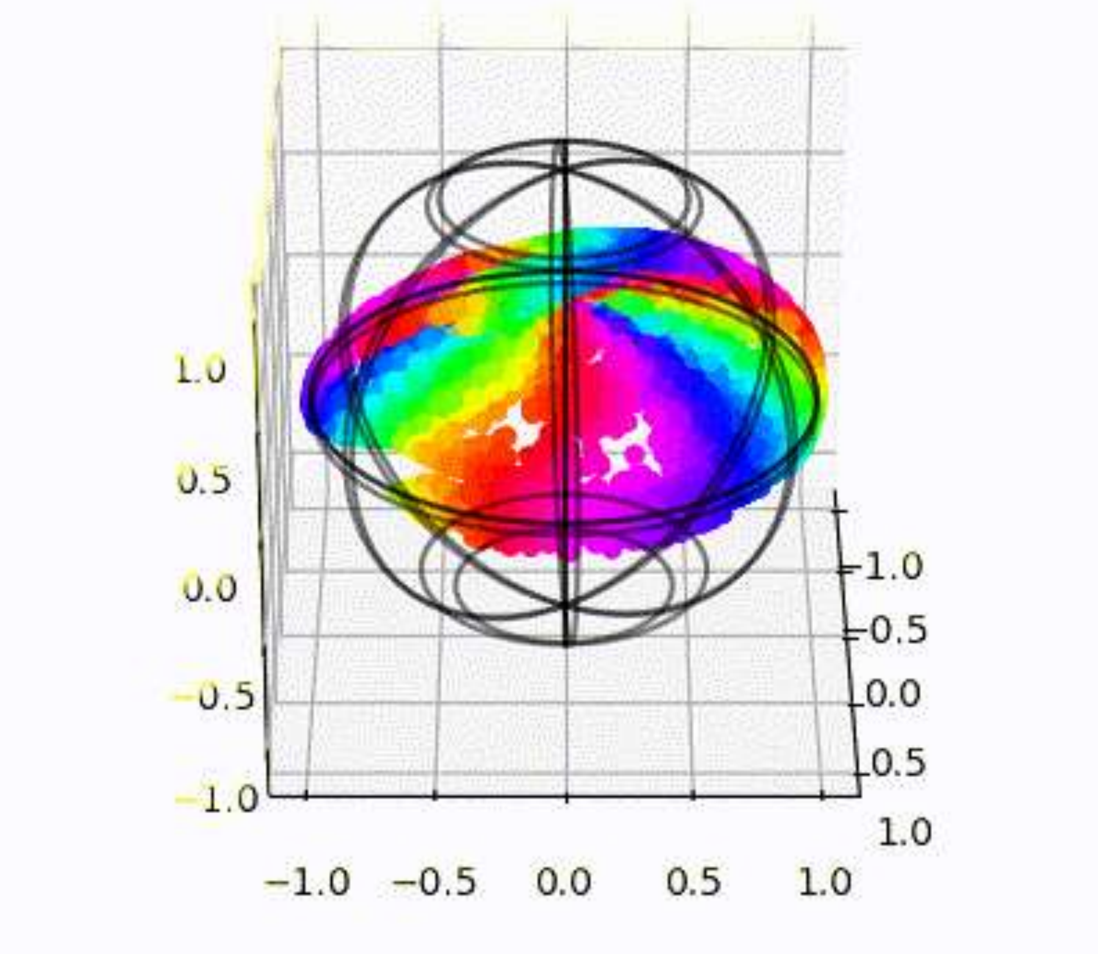


Lens PCA (recovered variance)



Visualization in

$$L_3^2 \cong B^3/x \sim x'$$



https://youtu.be/_Ic730_xFkw

The state of the art:

$$B = \bigcup \mathcal{U}$$

- Sparse, stable and transductive circular coordinates:

$$[\eta] \in H^1(\mathcal{N}(\mathcal{U}); \mathbb{Z}) \Rightarrow f_\eta : B \longrightarrow S^1$$

- Real Projective coordinates :

$$[\theta] \in H^1(\mathcal{N}(\mathcal{U}); \mathbb{Z}/2) \Rightarrow f_\theta : B \longrightarrow \mathbb{R}\mathbf{P}^n$$

- Complex Projective coordinates:

$$[\nu] \in H^2(\mathcal{N}(\mathcal{U}); \mathbb{Z}) \Rightarrow f_\nu : B \longrightarrow \mathbb{C}\mathbf{P}^n$$

- Lens Space Coordinates:

$$[\mu] \in H^1(\mathcal{N}(\mathcal{U}); \mathbb{Z}_q) \Rightarrow f_\mu : B \longrightarrow S^{2n-1}/(\mathbb{Z}_q)$$

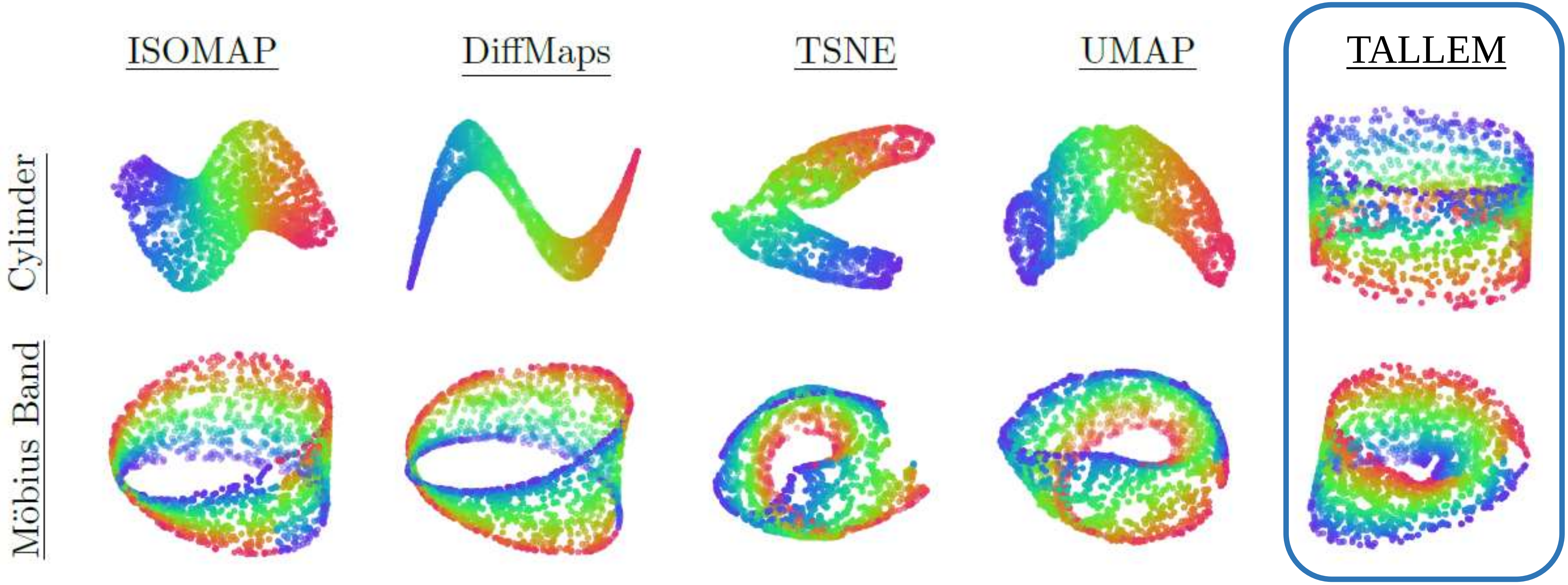
Code

- Python Library (w/ Chris Tralie):

DREiMac: Dimension Reduction with Eilenberg-MacLane Coordinates

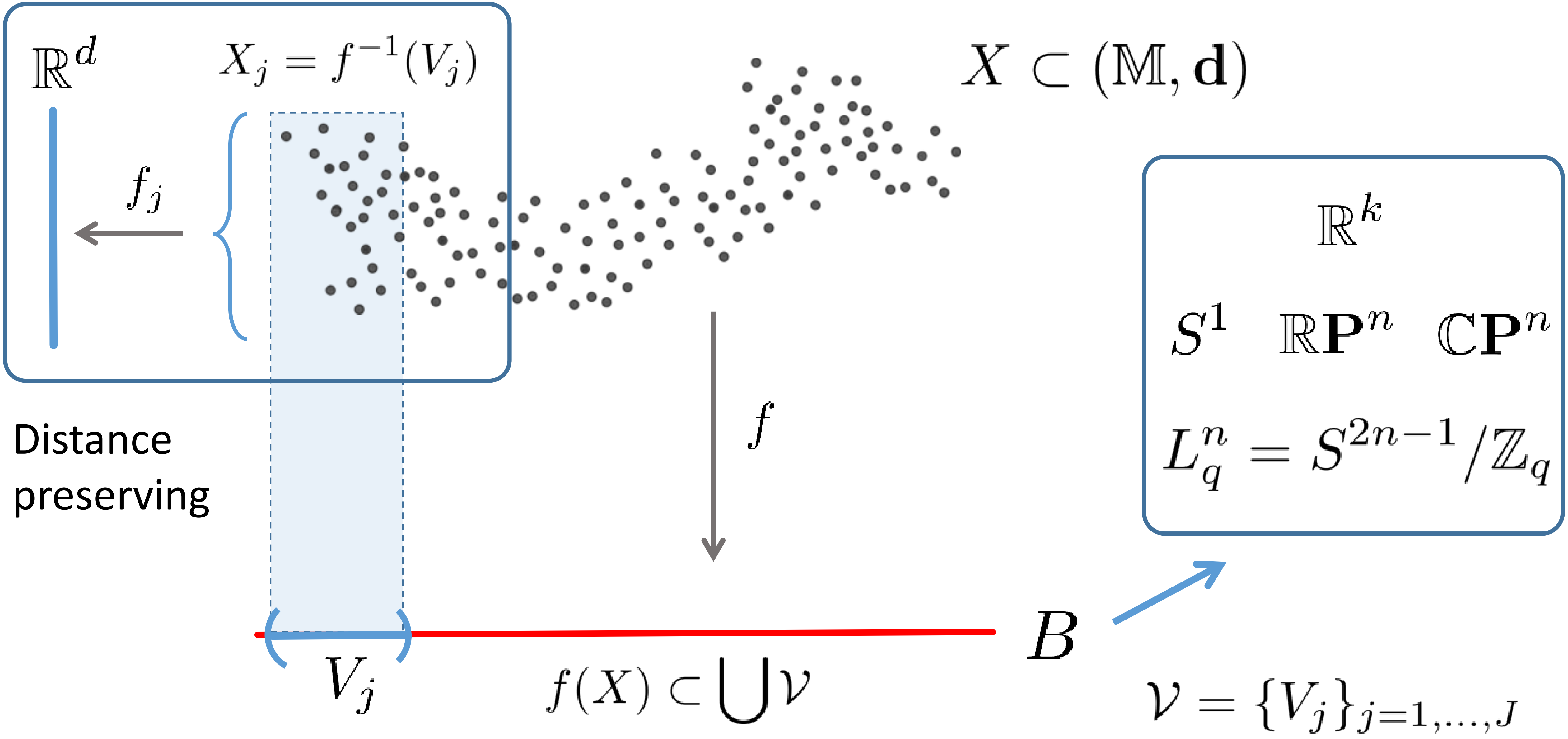
https://github.com/ctralie/DREiMac

Topological Assembly of Locally Euclidean Models



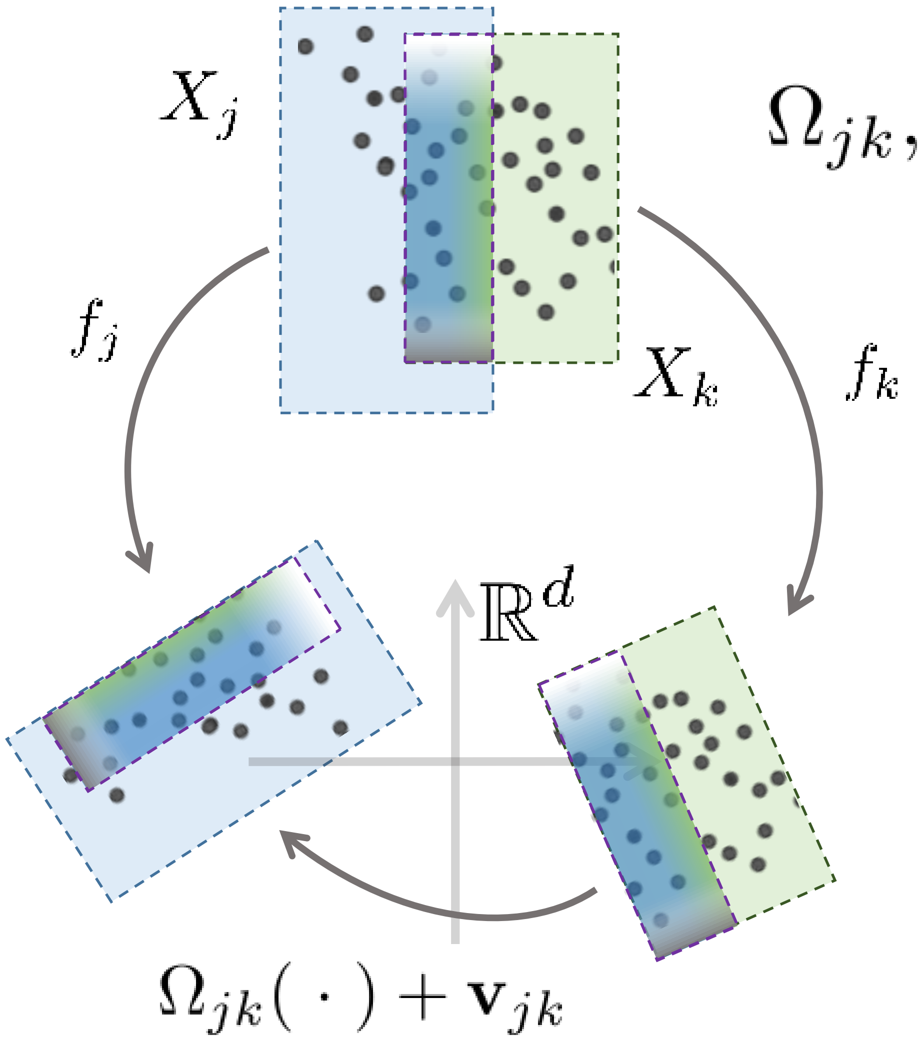
TALLEM: Topological Alignment of Locally Euclidean Models, J. Mike & J. Perea, **Preprint**, 2019

Locally Euclidean Models



Assembly of Locally Euclidean Models

Orthogonal Procrustes Problem

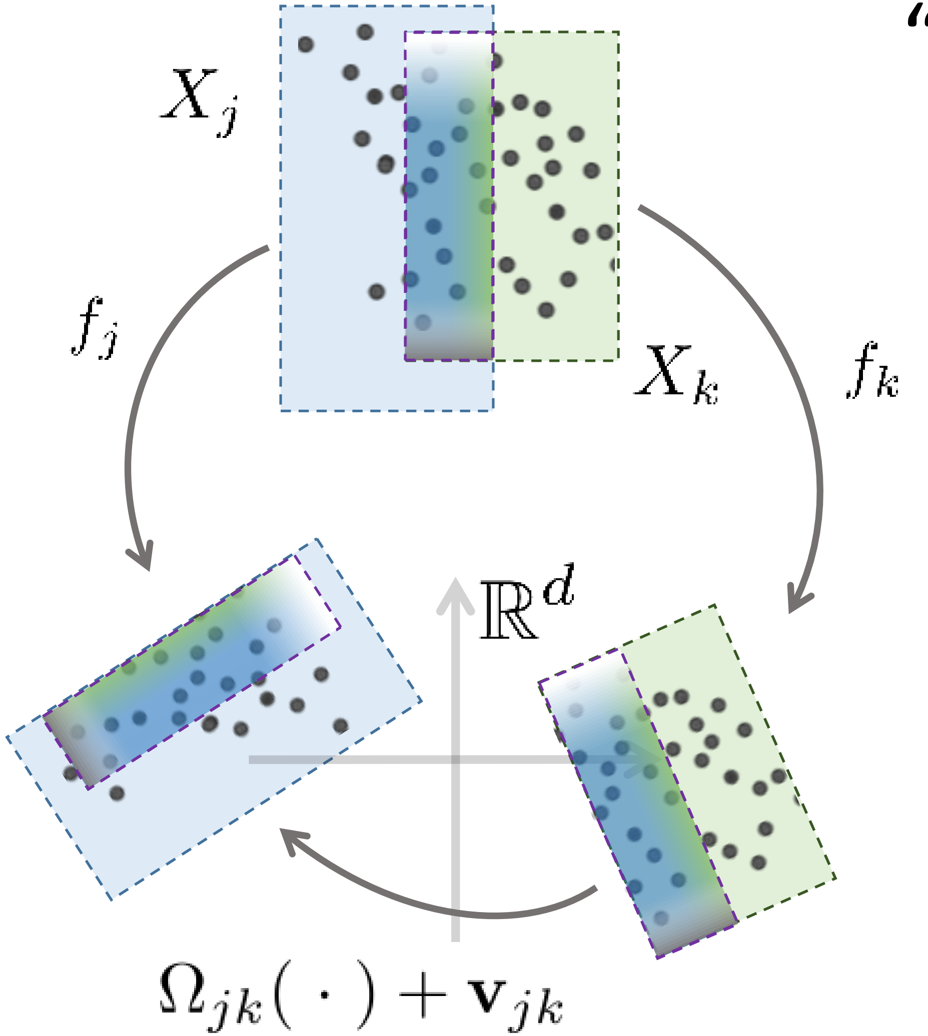


$$\Omega_{jk}, \mathbf{v}_{jk} = \underset{\substack{\Omega \in O(d) \\ \mathbf{v} \in \mathbb{R}^d}}{\text{argmin}} \sum_{x \in X_{jk}} \|\Omega f_k(x) + \mathbf{v} - f_j(x)\|^2$$

$$\Omega_{jj} = id_{\mathbb{R}^d} \qquad \in O(d)$$

$$\Omega_{kj} = \Omega_{jk}^{-1}$$

Assembly of Locally Euclidean Models



"Theorem":

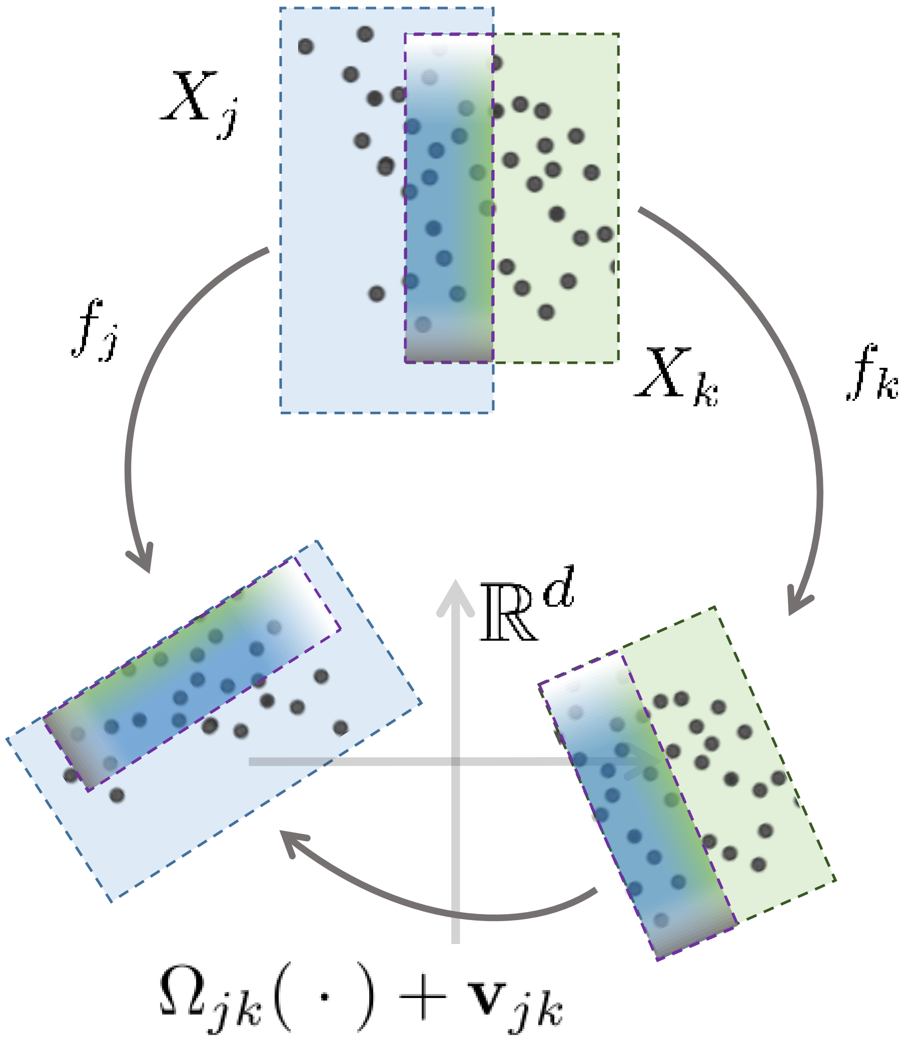
If conditions on X_{jkl} and f_j, f_k, f_l ,

then $\Omega_{jk}\Omega_{kl} \approx \Omega_{jl}$.

$$\Omega_{jj} = id_{\mathbb{R}^d} \qquad \in O(d)$$

$$\Omega_{kj} = \Omega_{jk}^{-1}$$

Topological Assembly of Locally Euclidean Models



$$\mathcal{X} = \{X_j\}_{1 \le j \le J}$$

$$X = \bigcup \mathcal{X}$$

Approximate rank d Vector Bundle Over $|\mathcal{N}(\mathcal{X})|$

$$\left.\begin{aligned} \Omega_{jk}\Omega_{kl} &\approx \Omega_{jl} \\ \Omega_{jj} &= id_{\mathbb{R}^d} \\ \Omega_{kj} &= \Omega_{jk}^{-1} \end{aligned}\right\} \in O(d)$$

Approximate Cocycle Condition

Topological Assembly of Locally Euclidean Models

$$\mathcal{X} = \{X_j\}_{1 \leq j \leq J}$$

Frames:

$$\Phi_j : X_j \longrightarrow V_d(\mathbb{R}^{Jd})$$ Stiefel manifold

$$x \mapsto \left[\varphi_1(x)\Omega_{1j}, \ldots, \varphi_J(x)\Omega_{Jj}\right]^*$$

$$\Phi_k(x) \approx \Phi_j(x) \cdot \Omega_{jk}$$

$$x \in X_{jk}$$

$V_d(\mathbb{R}^{Jd})$ Supports $O(d)$ - Equivariant PCA!!

$$\phi_j : X_j \longrightarrow V_d(\mathbb{R}^n) \qquad \phi_k(x) \approx \phi_j(x) \cdot \Omega_{jk} \qquad d \leq n < Jd$$

Topological Assembly of Locally Euclidean Models

$$F : X \longrightarrow \mathbb{R}^n$$

Assemblage

$$x \mapsto \sum_{j=1}^{J} \varphi_j(x)\phi_j(x)(f_j(x) + \theta_j)$$

$$\exists\, \theta_j \in \mathbb{R}^d \;,\; 1 \leq j \leq J$$

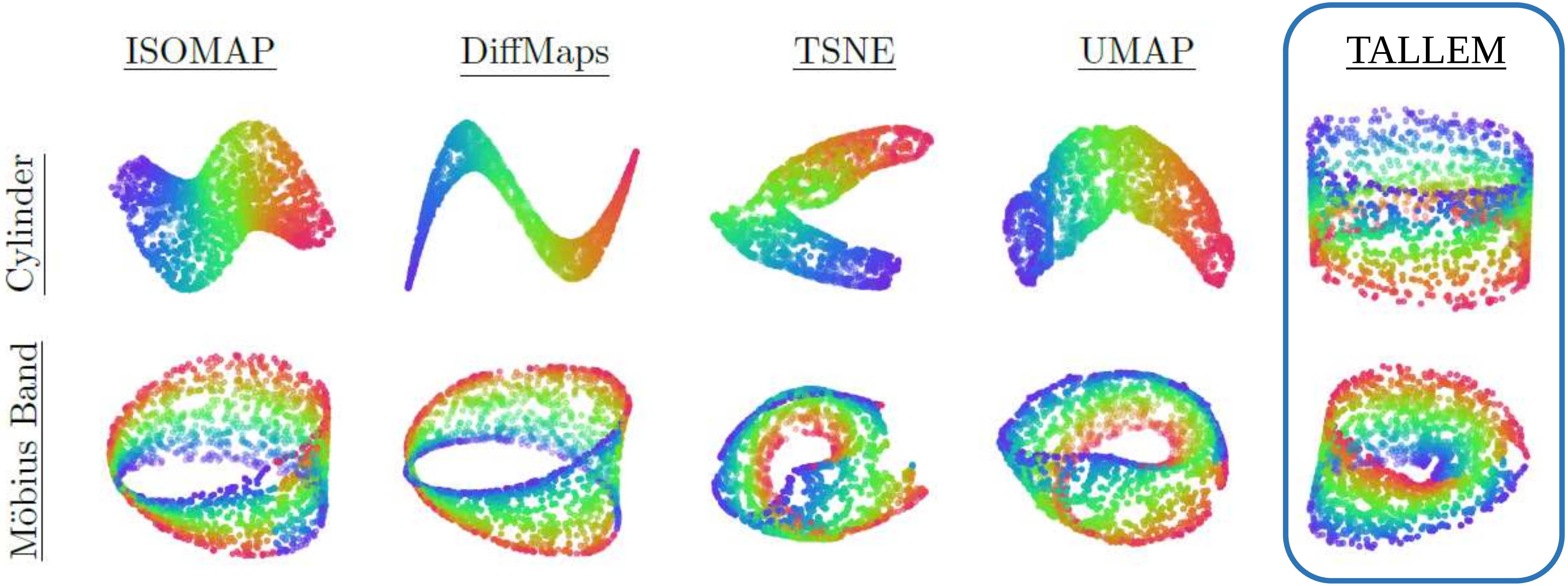
$$\mathbf{v}_{jk} \approx \Omega_{jk}\theta_k - \theta_j$$

$$\varphi_j : X \longrightarrow [0,1]$$

Partition of unity

$$\phi_j : X_j \longrightarrow V_d(\mathbb{R}^n) \qquad \phi_k(x) \approx \phi_j(x) \cdot \Omega_{jk} \qquad d \leq n < Jd$$

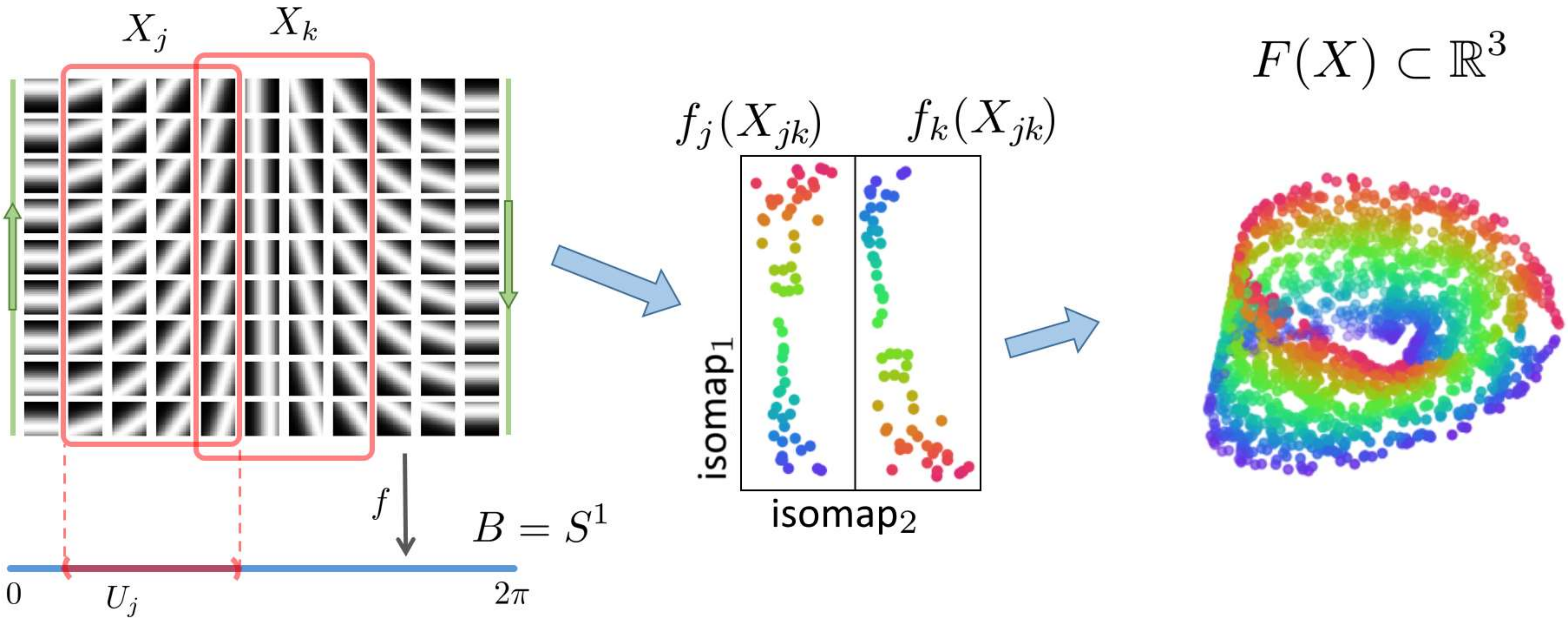
Topological Assembly of Locally Euclidean Models



TALLEM: Topological Alignment of Locally Euclidean Models, J. Mike & J. Perea, **Preprint**, 2019

Persistence → circular coordinates

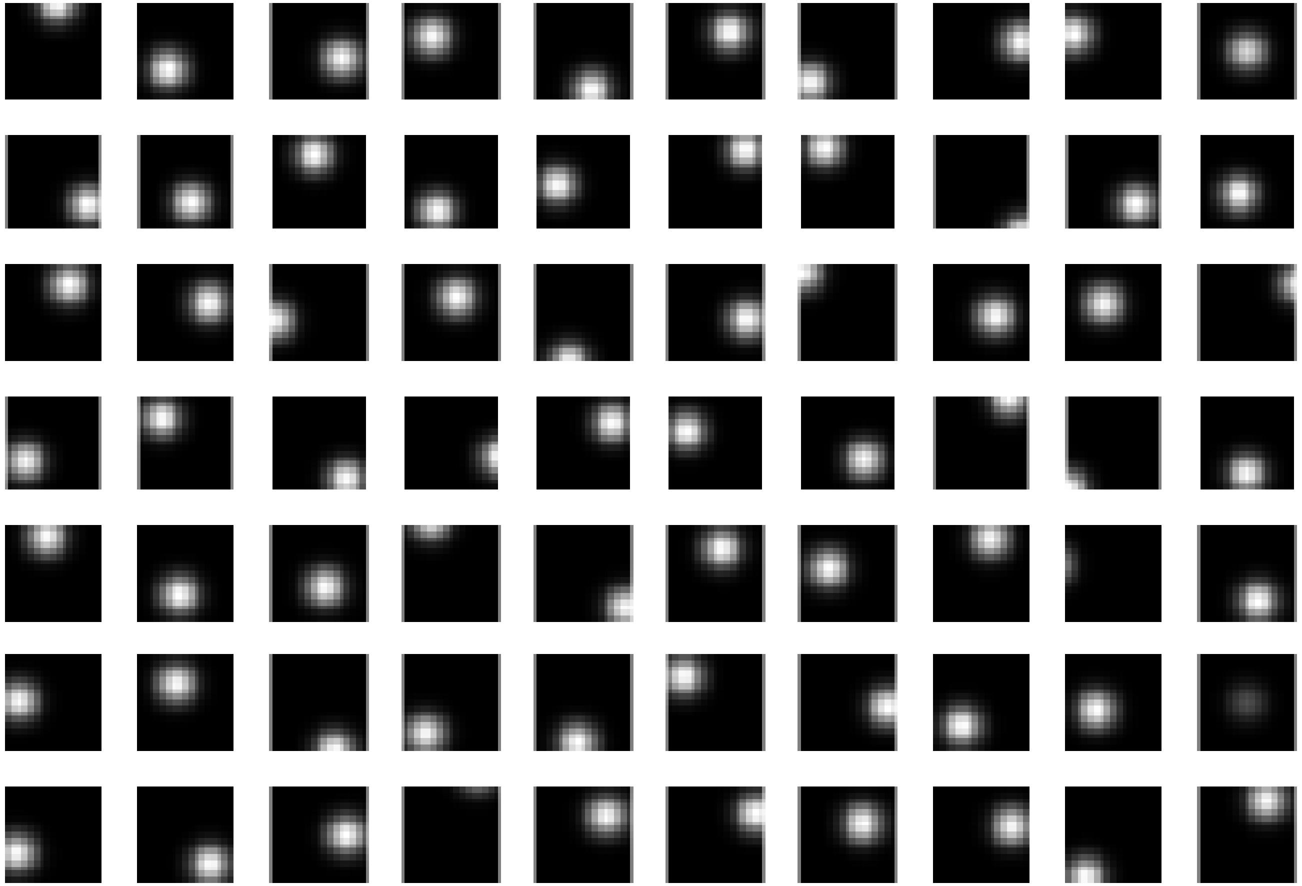
TALLEM assemblage



TALLEM: Topological Alignment of Locally Euclidean Models, J. Mike & J. Perea, **Preprint**, 2019

TALLEM vs Charting Methods

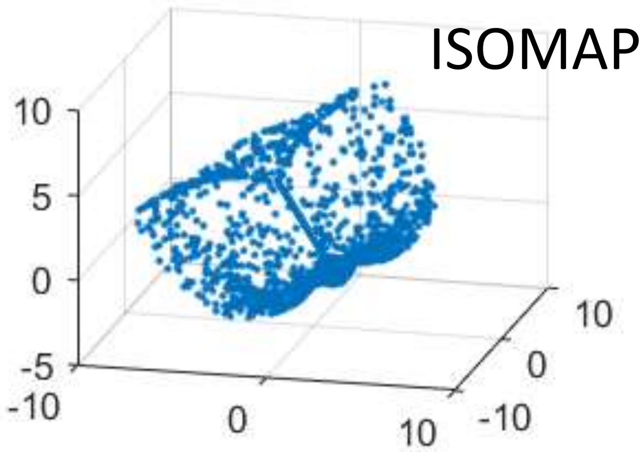
The data: a 2-sphere whose north and south pole are connected by a line segment.



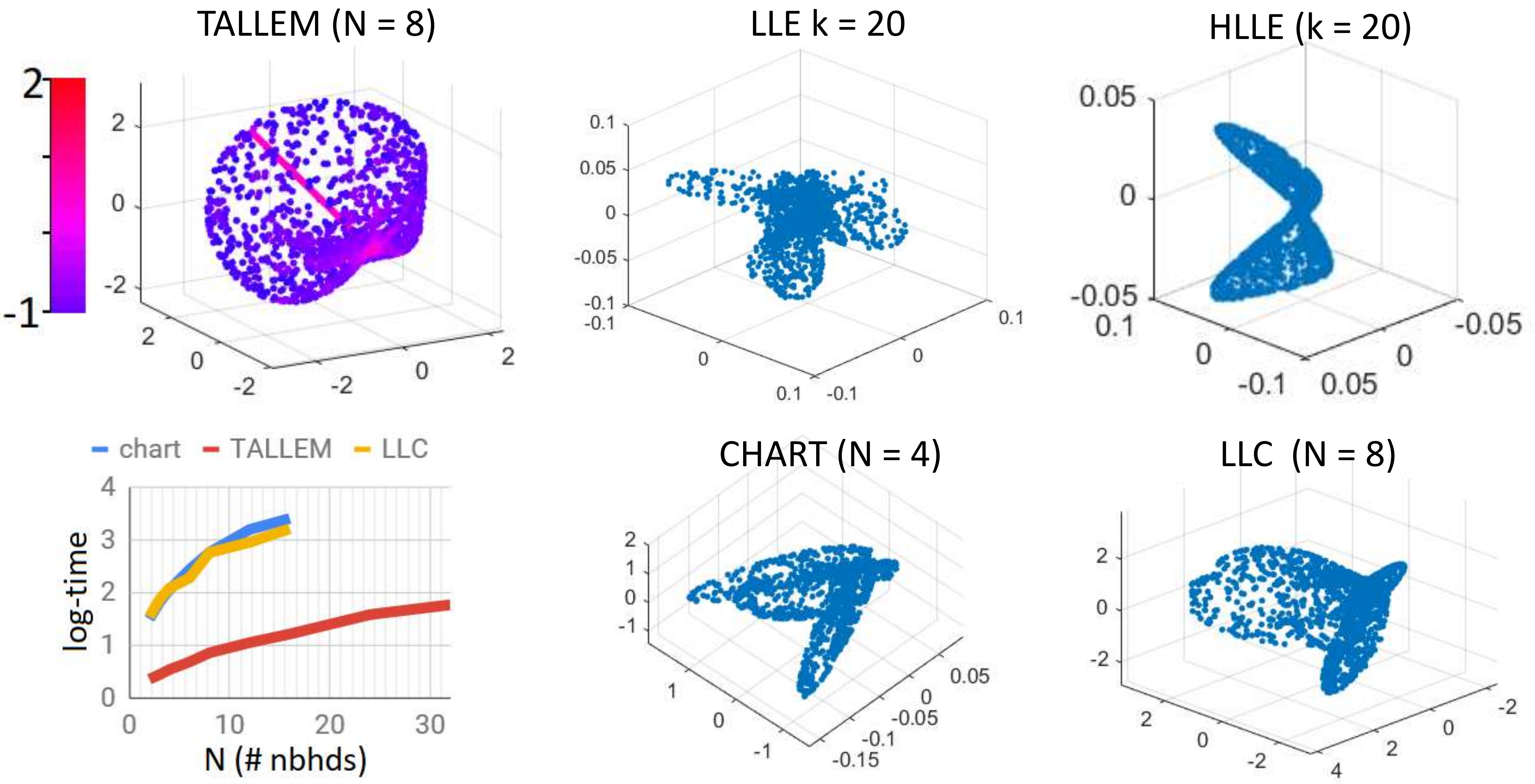
17 × 17 images of a white dot of fixed intensity and varying location
(1,900 of these — a 2-sphere)

17 × 17 images of a white dot of variable intensity at the center of the image patch
(100 of these — a line segment).

ISOMAP (k = 20)



TALLEM: Topological Alignment of Locally Euclidean Models, J. Mike & J. Perea, **Preprint**, 2019



TALLEM: Topological Alignment of Locally Euclidean Models, J. Mike & J. Perea, **Preprint**, 2019

References

J. A. Perea, *Multiscale Projective Coordinates via Persistent Cohomology of Sparse Filtrations*
Discrete & Computational Geometry, 2018

J. A. Perea, *Sparse Circular Coordinates via Principal $\mathbb{Z}$-Bundles*
Proceedings of the Abel Symposium, 2019

L. Polanco and J. A. Perea, *Coordinatizing Data with Lens Spaces and Persistent Cohomology*
Proceedings of CCCG, 2019

J. Mike and J. A. Perea, *TALLEM: Topological Assembly of Locally Euclidean Models*
Preprint, 2019