Bottom-up manifold learning with low distortion

Gal Mishne

Halıcıoğlu Data Science Institute
UC San Diego

Joint work with Dhruv Kohli, Alex Cloninger, Bas Nieuwenhuis and Devika Narain

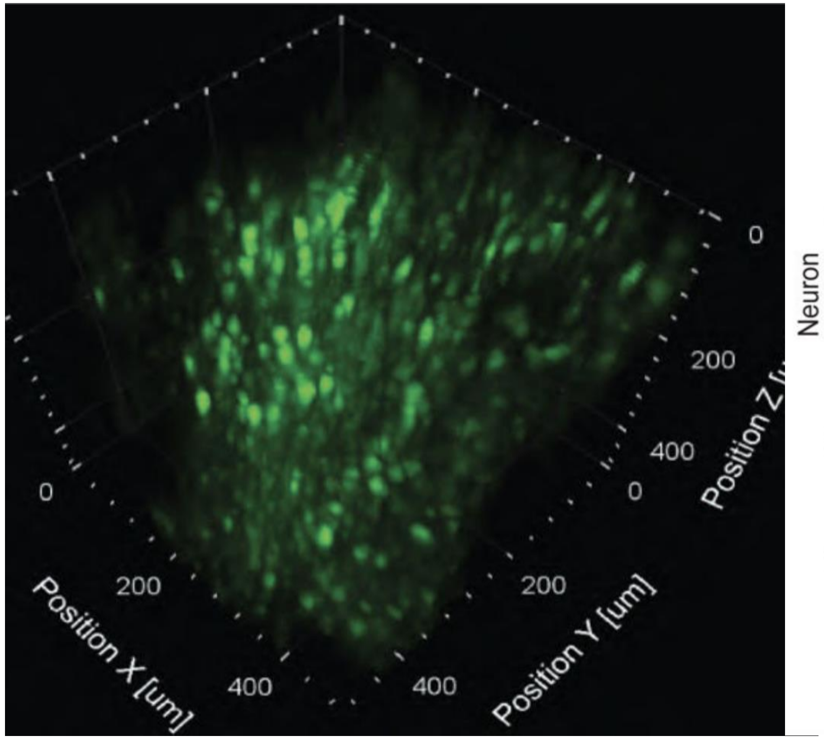
UC San Diego

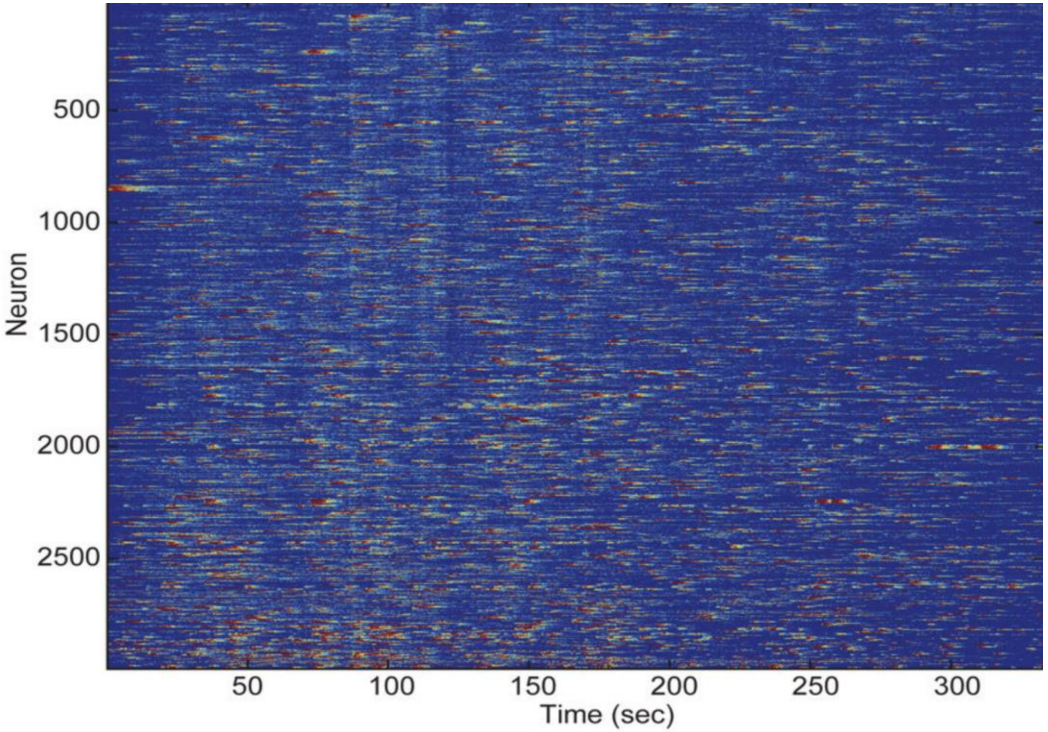
ENCORE

Neuro-data-science

Neuro-data-science

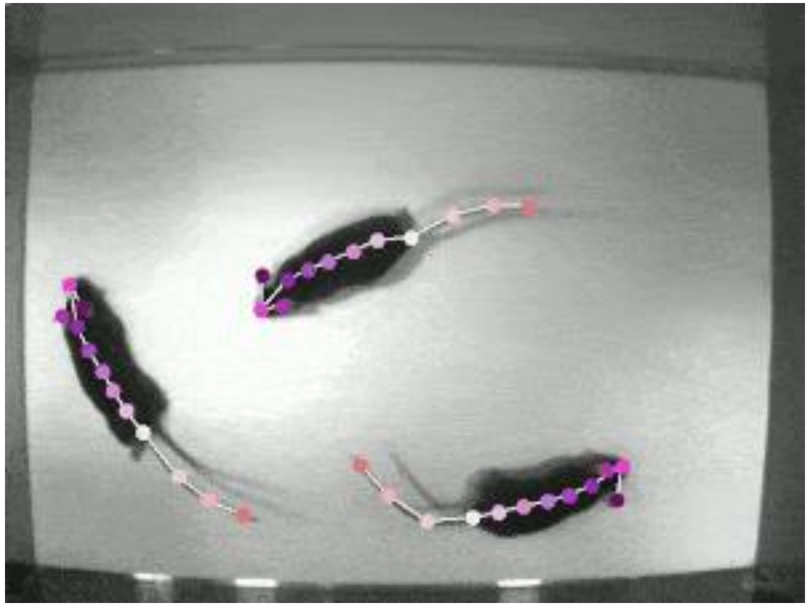
- Data explosion in neuroscience

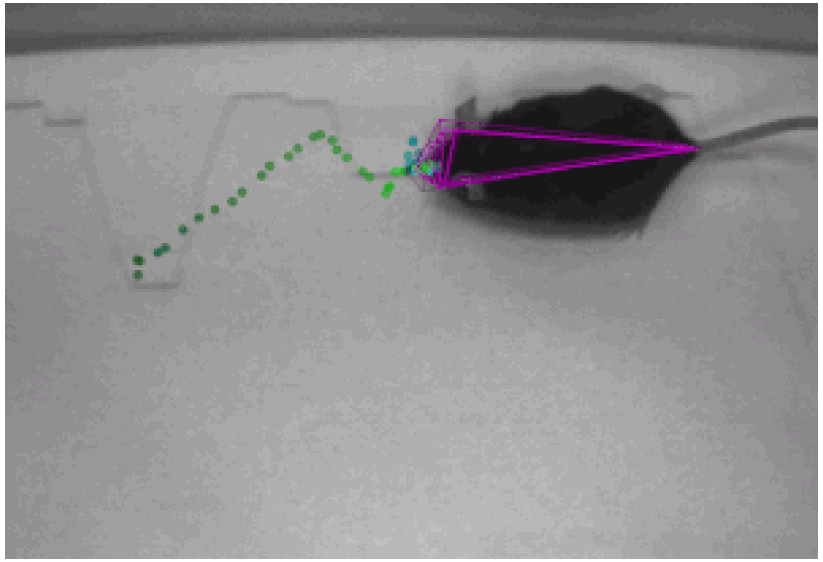


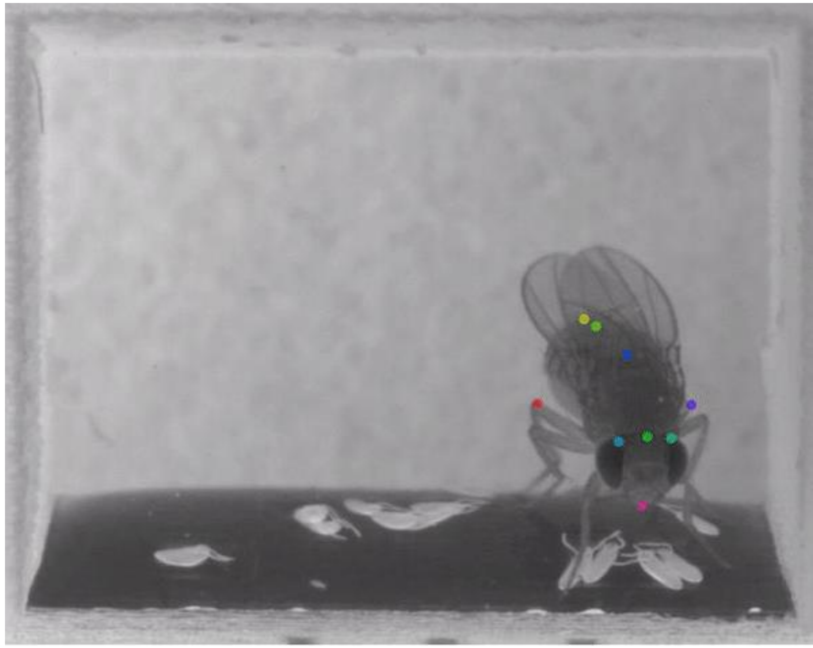


Volumetric calcium imaging
thousands of neurons
across multiple cortical layers
Source: [Prevedel et al. 2016]

Neuro-data-science

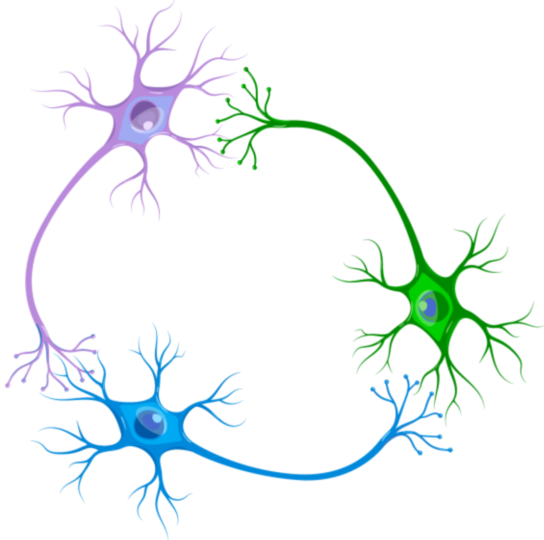


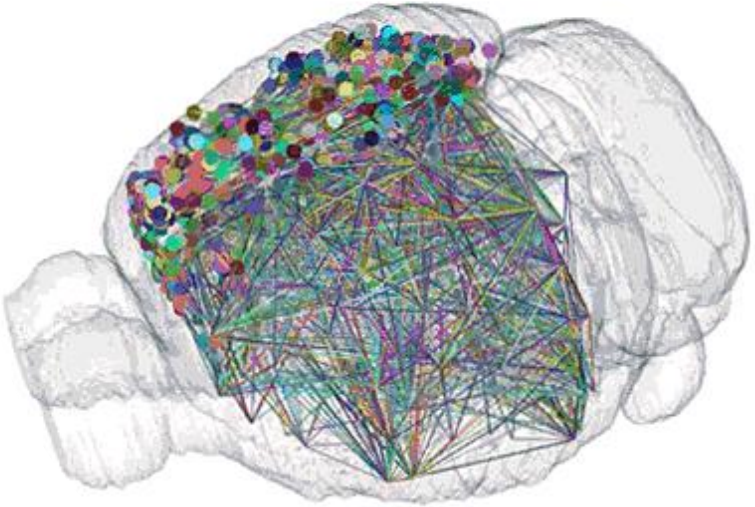




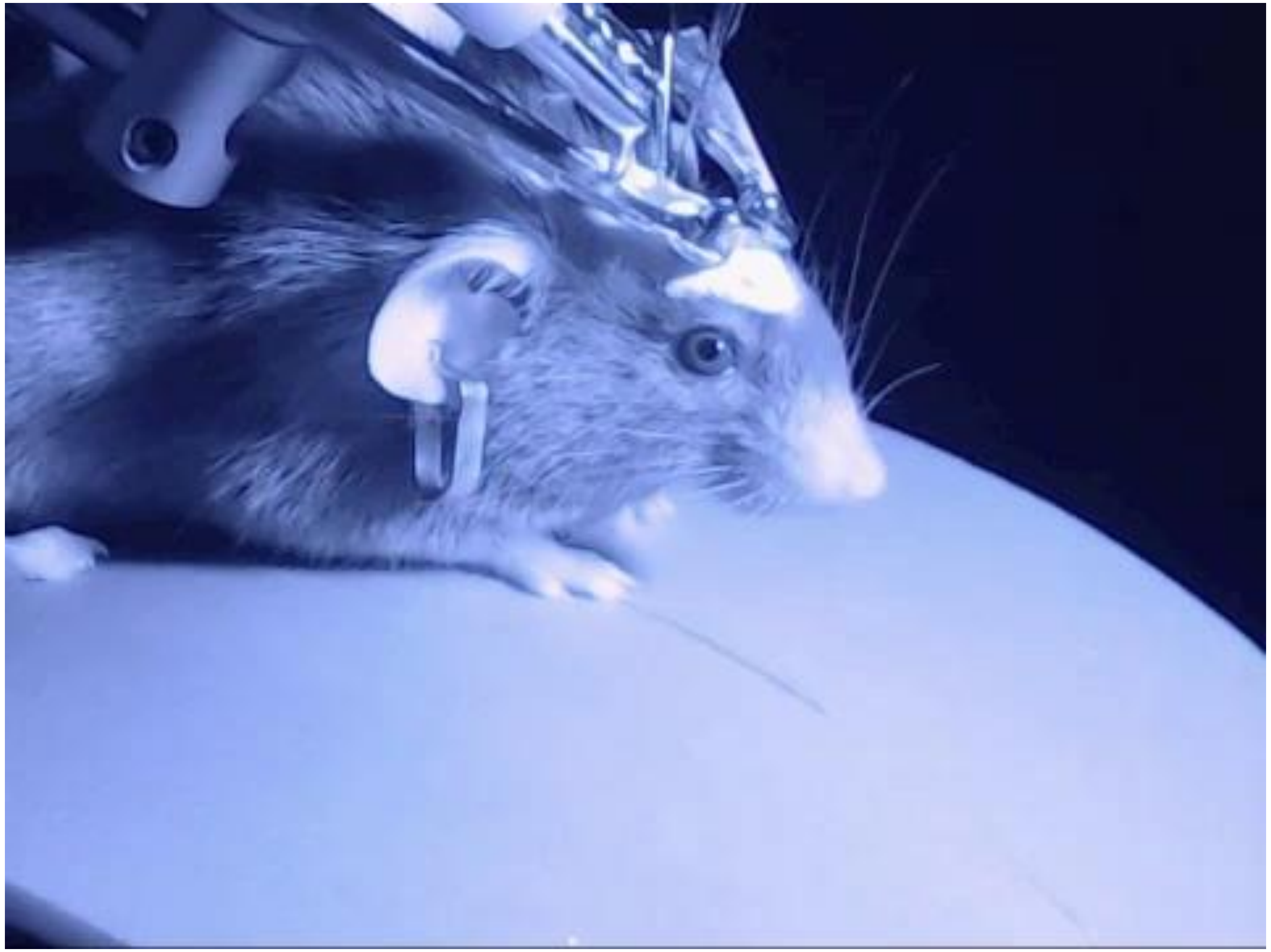
Behavioral videos of animal models
[Deeplabcut, 2019]

How does the brain generate behavior?



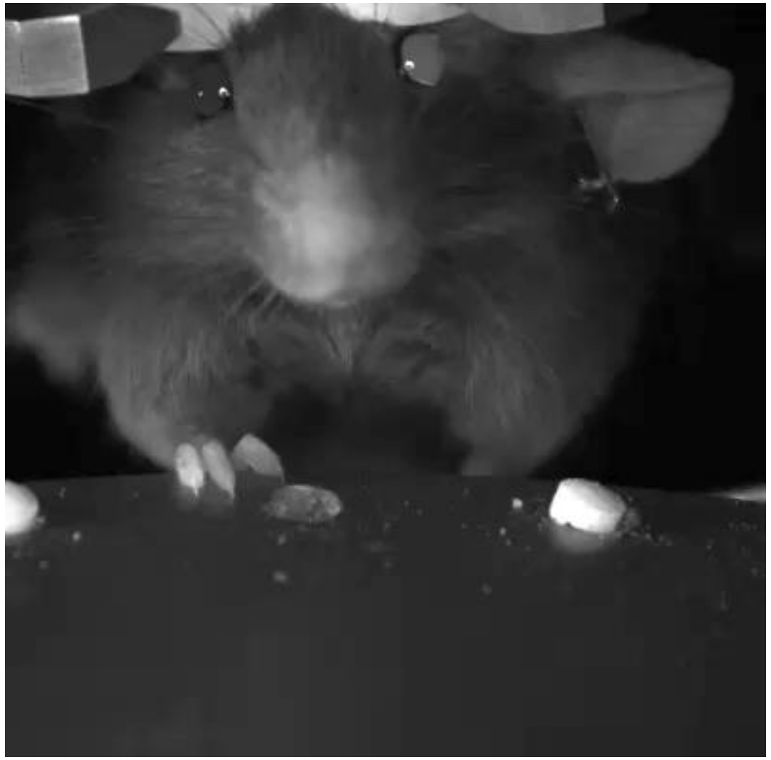


Spontaneous Behavior

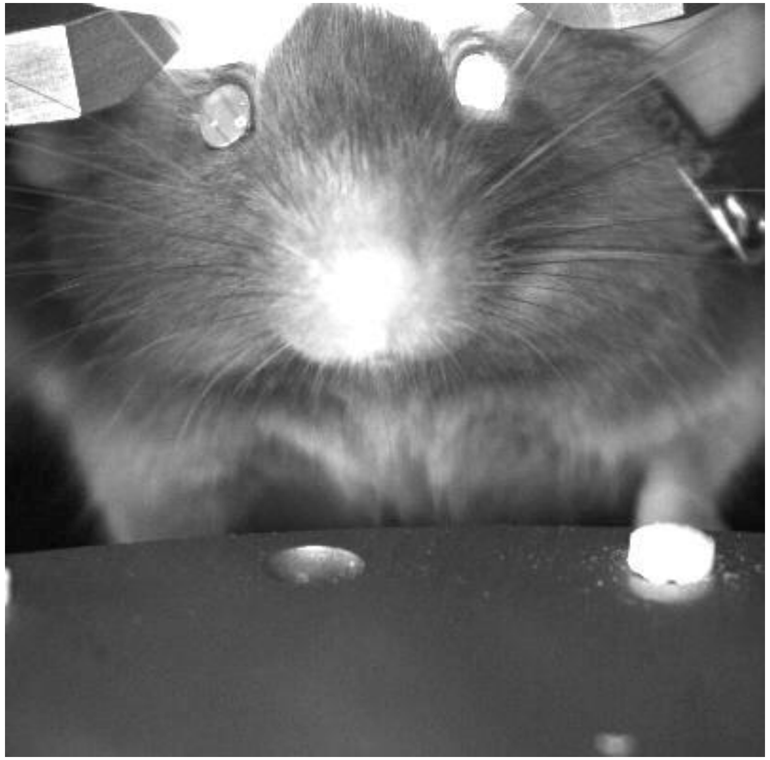


[Higley lab]

Learned motor task



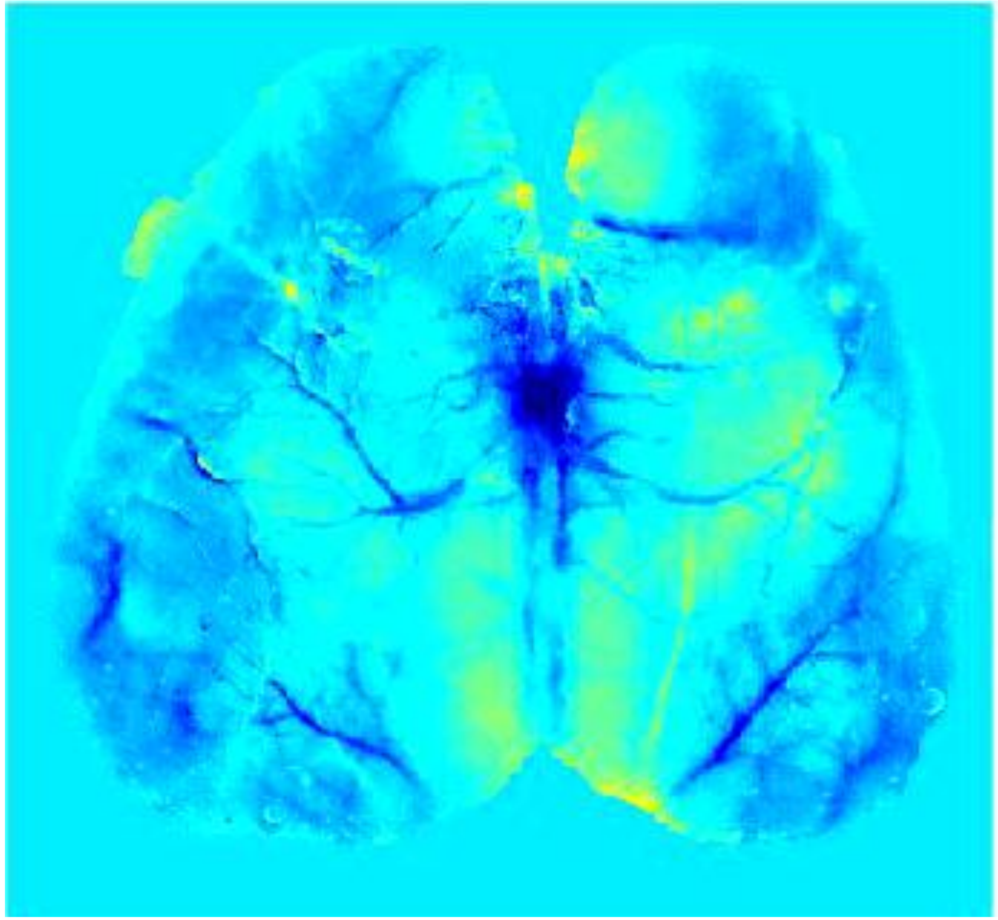
Novice animal

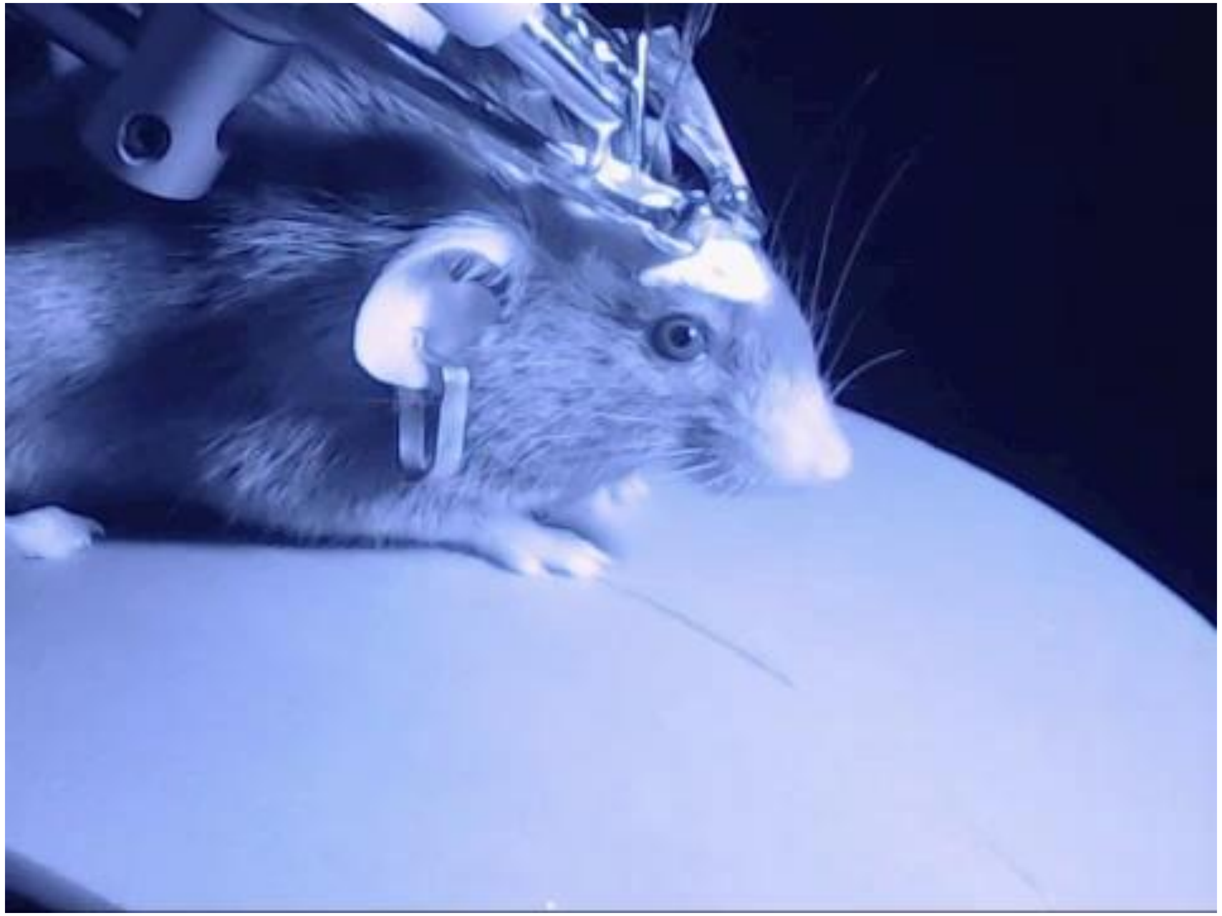


Expert animal

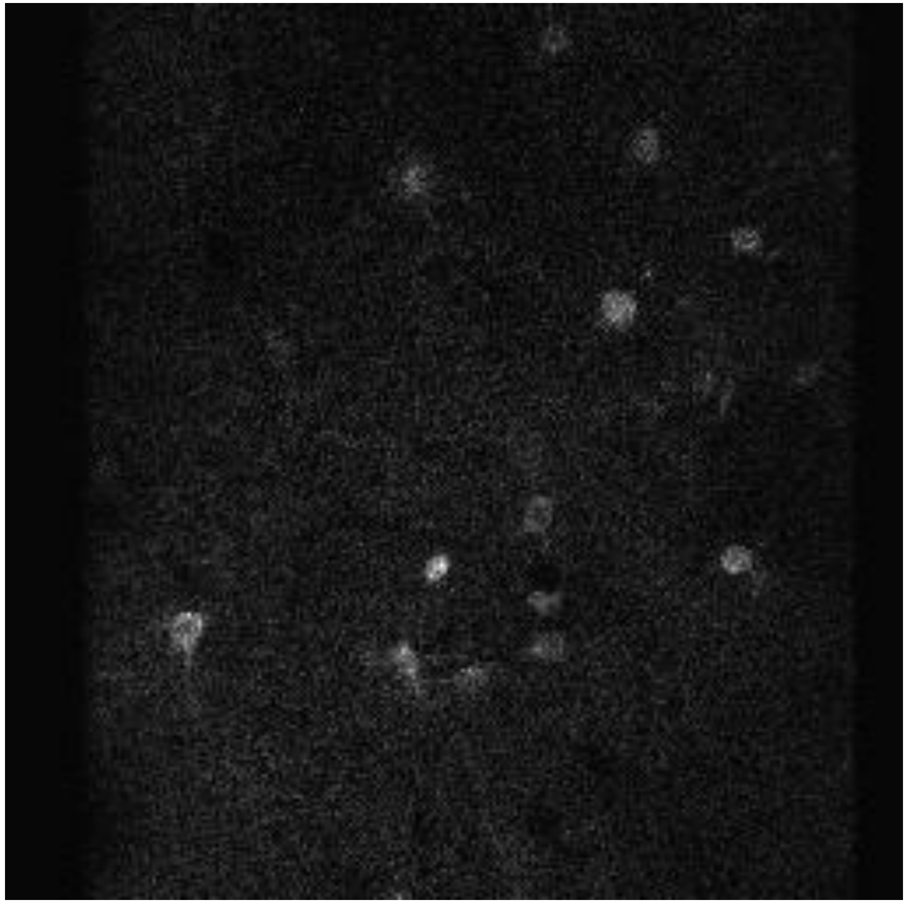
[Schiller lab]

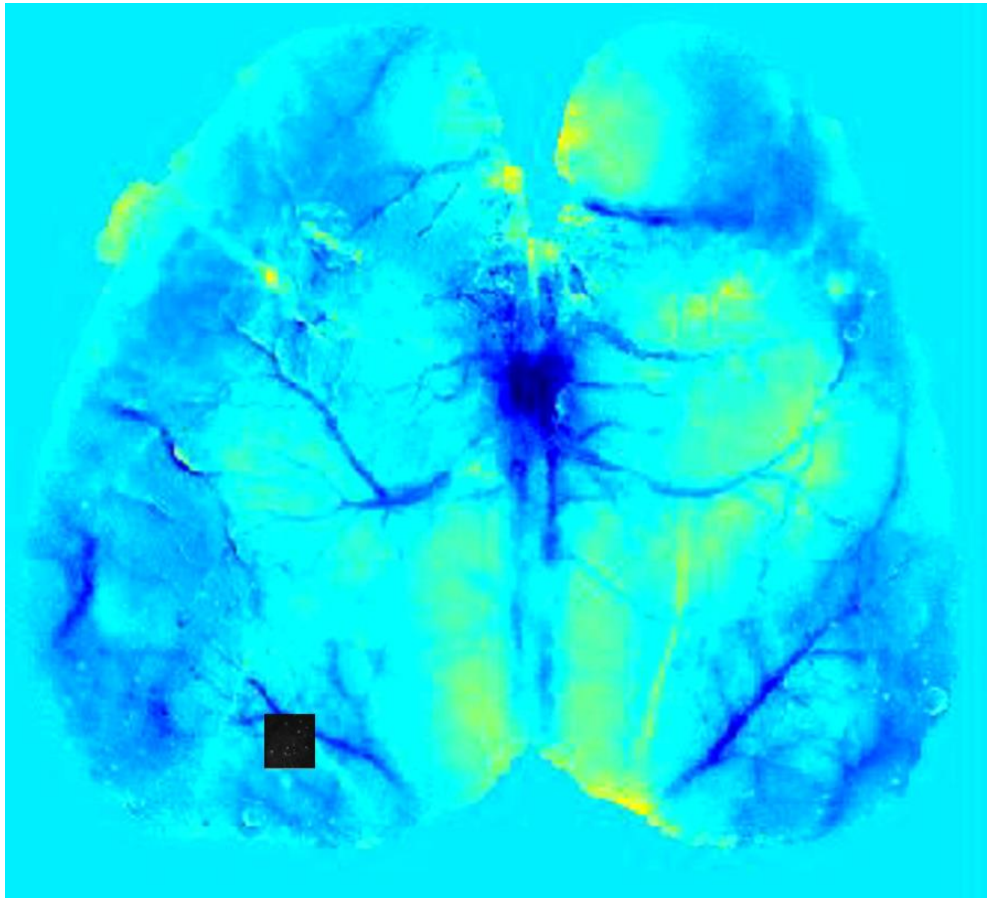
Imaging of large-scale networks





Imaging of large-scale networks

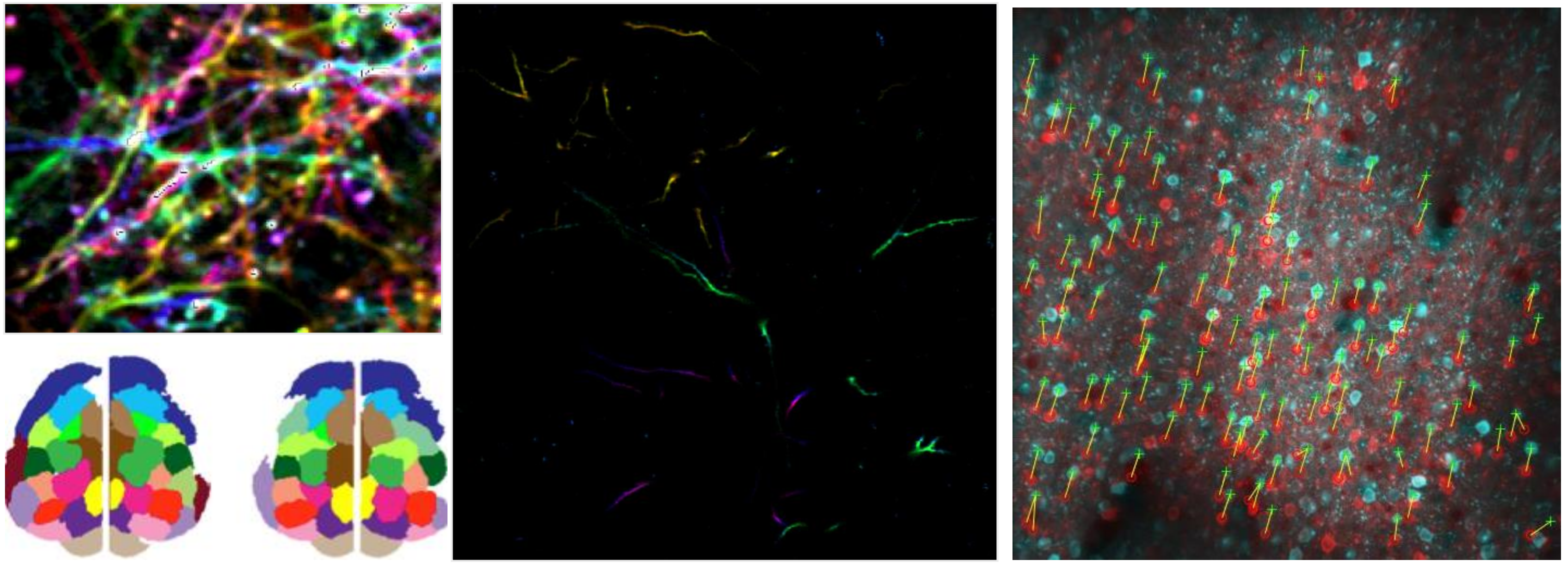




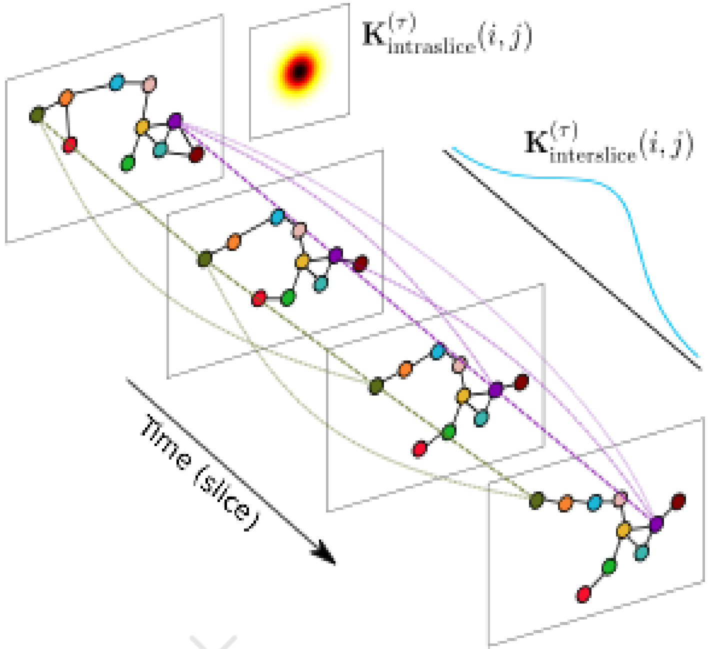
[Barson et al., 2020]

What does my lab do? Neuro-data-science

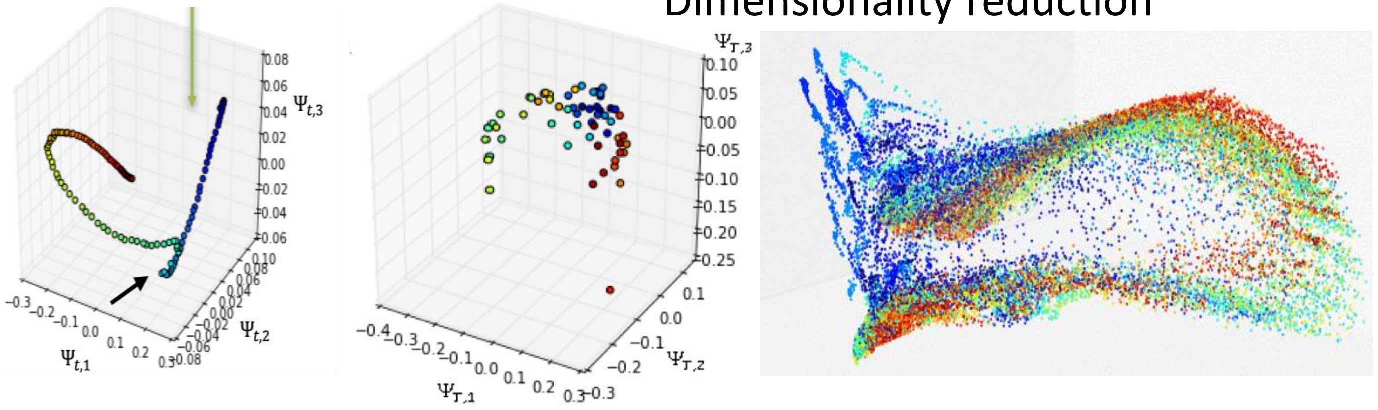
Imaging analysis



Learning connectivity



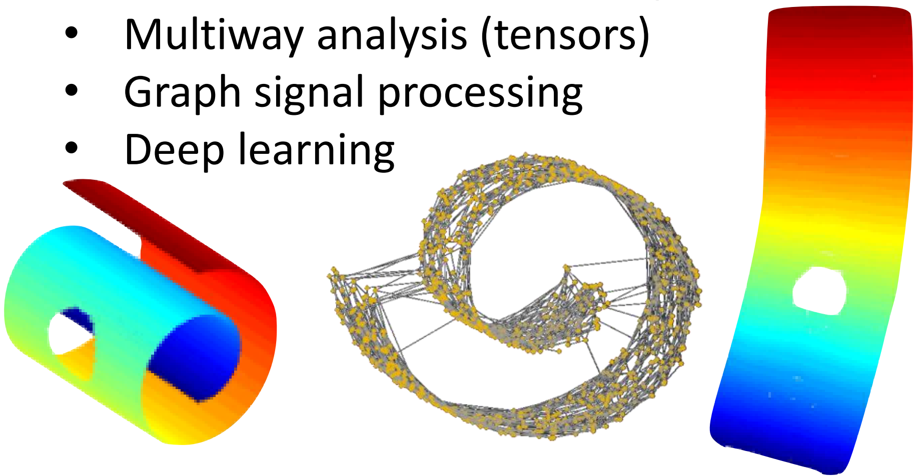
Dimensionality reduction



Time   Trials   learning over days

Methods

- Dimensionality reduction / Nonlinear manifold learning
- Multiway analysis (tensors)
- Graph signal processing
- Deep learning



Bottom-up manifold learning with low distortion

Gal Mishne
Halıcıoğlu Data Science Institute
UC San Diego

Joint work with Dhruv Kohli, Alex Cloninger, Bas Nieuwenhuis and Devika Narain

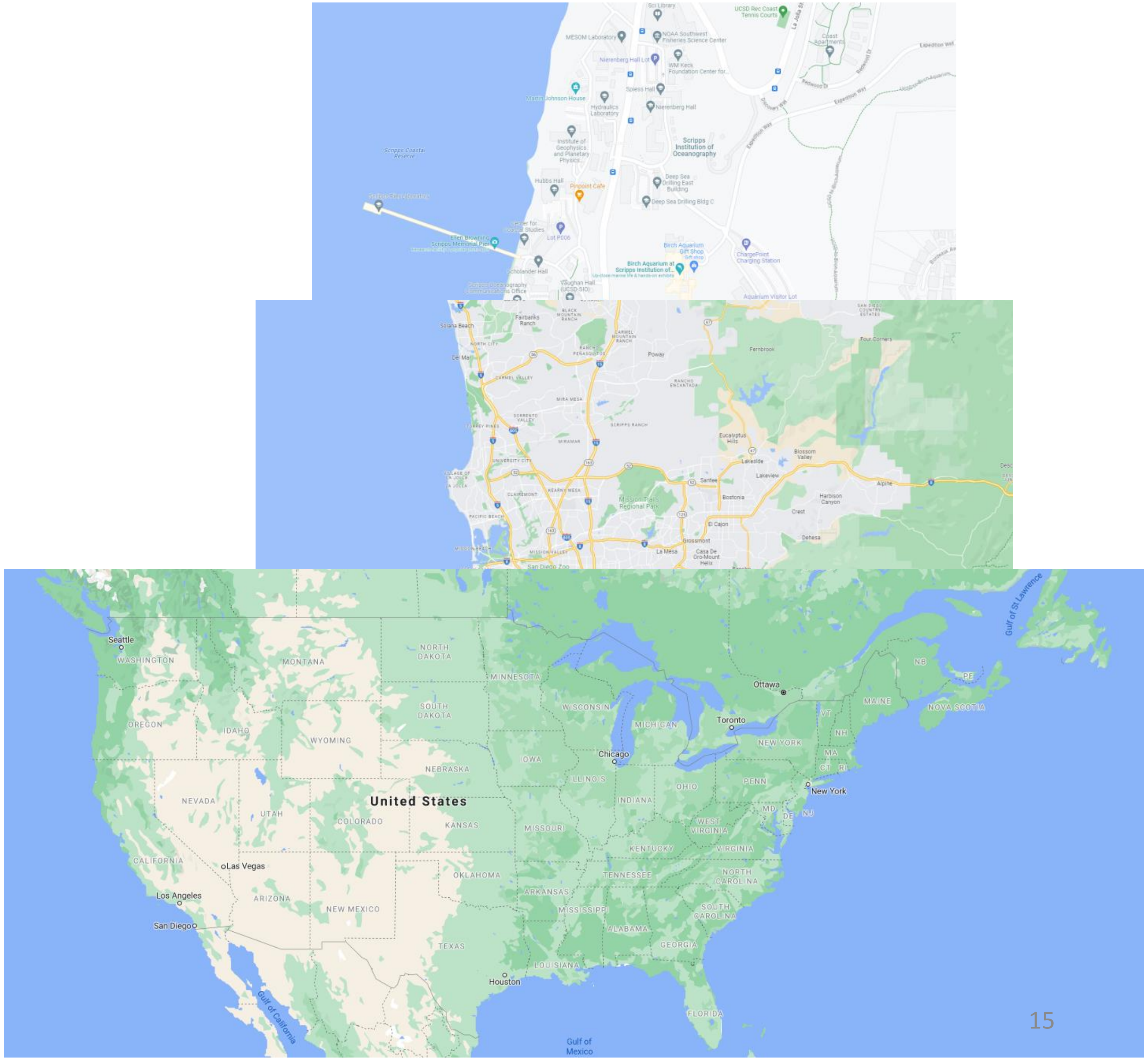
UC San Diego

EnCORE

Outline

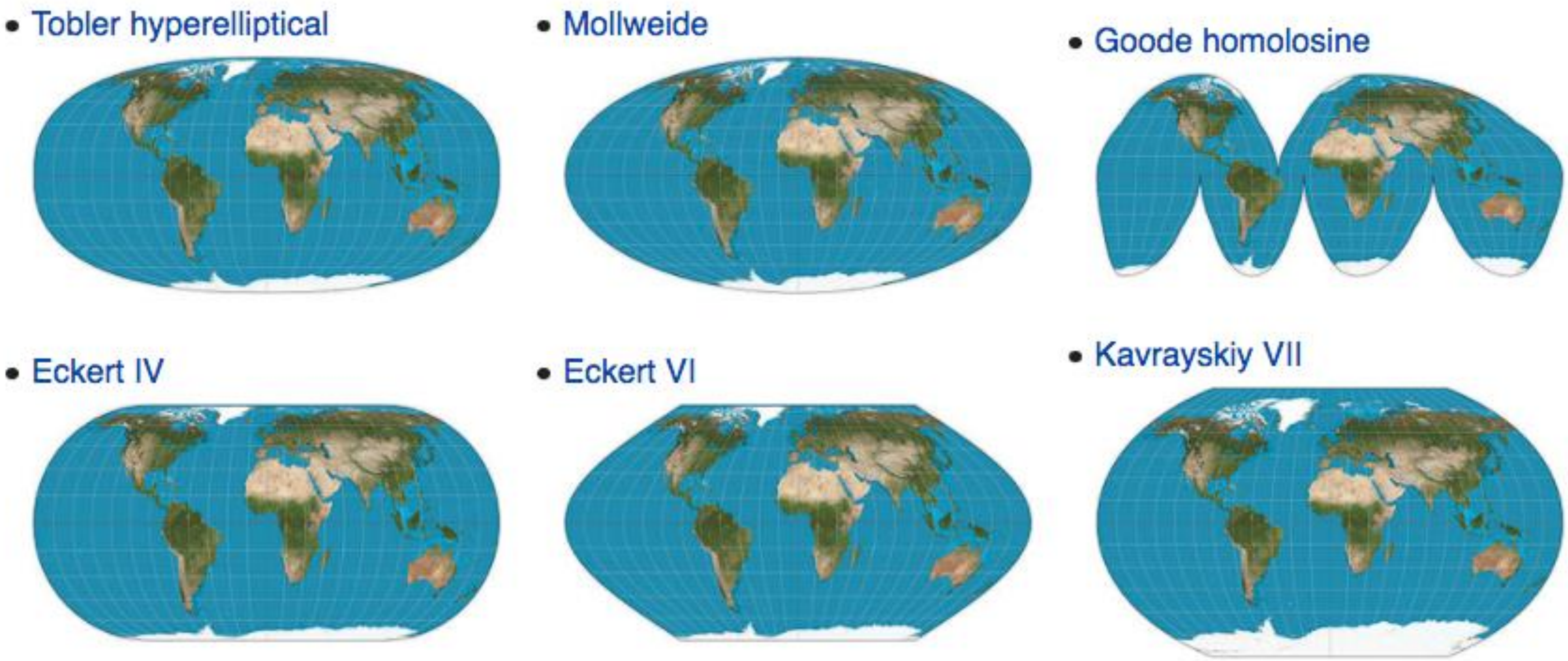
- **Introduction**
- Distortion
- Bottom-up manifold learning
- Results

Introduction



[Wikipedia, google maps]

Introduction



[Wikipedia]

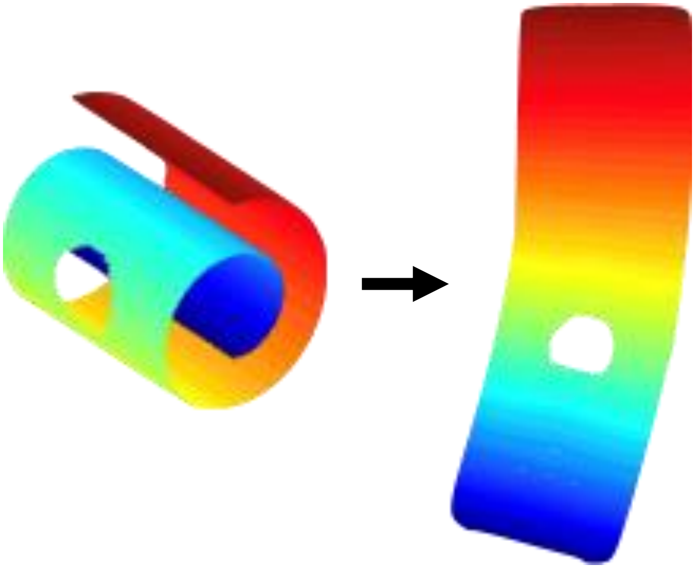
Introduction

What is a manifold?

- **Definition**: a topological space that locally resembles Euclidean space near each point.
- 1D:

- 2D:

Manifold learning



- Learn manifold from data
- Non-linear **representation** of low-dimensional manifold.
- Preserve **geometric** properties
- Embedding with top eigenvectors of the
    - Covariance matrix (PCA)
    - Normalized graph Laplacian (Laplacian Eigenmaps)
    - Random-walk graph Laplacian (Diffusion maps)
    - …

*[Tenenbaum et al., 2000,
Roweis and Saul, 2000,
Belkin and Niyogi, 2001,
Donoho and Grimes, 2002,
Coifman and Lafon, 2004,
van der Maaten and Hinton 2008,
McInnes et al. 2018, ...]*

Manifold learning frameworks

Manifold learning frameworks

Top-down

- Initial global embedding of the data
- Optional: refine it iteratively by minimizing a measure of local distortion.
- ISOMAP, Laplacian Eigenmaps, t-SNE, UMAP…

Manifold learning frameworks

Top-down

- Initial global embedding of the data
- Optional: refine it iteratively by minimizing a measure of local distortion.
- ISOMAP, Laplacian Eigenmaps, t-SNE, UMAP…

Manifold learning frameworks

Top-down

- Initial global embedding of the data
- Optional: refine it iteratively by minimizing a measure of local distortion.
- ISOMAP, Laplacian Eigenmaps, t-SNE, UMAP…

Bottom-up

- Calculate local views with low distortion
- Solve alignment problem to register views to global embedding
- LTSA, LDLE, RATS

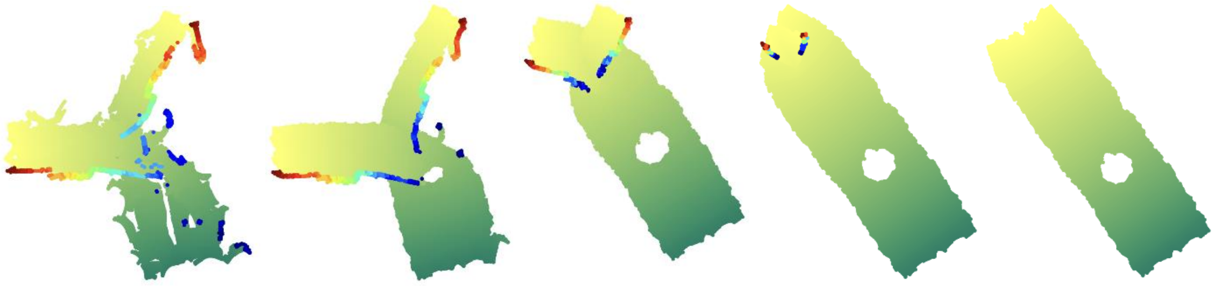
Manifold learning frameworks

Top-down

- Initial global embedding of the data
- Optional: refine it iteratively by minimizing a measure of local distortion.
- ISOMAP, Laplacian Eigenmaps, t-SNE, UMAP…

Bottom-up

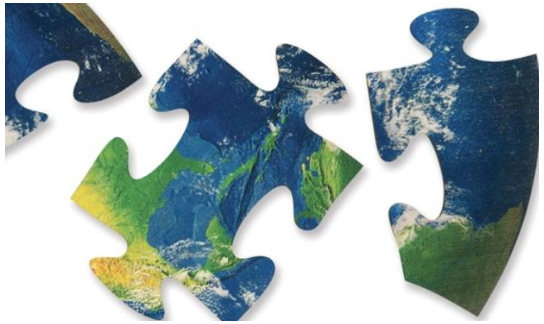
- Calculate local views with low distortion
- Solve alignment problem to register views to global embedding
- LTSA, LDLE, RATS



Bottom-up manifold learning

Bottom-up manifold learning:

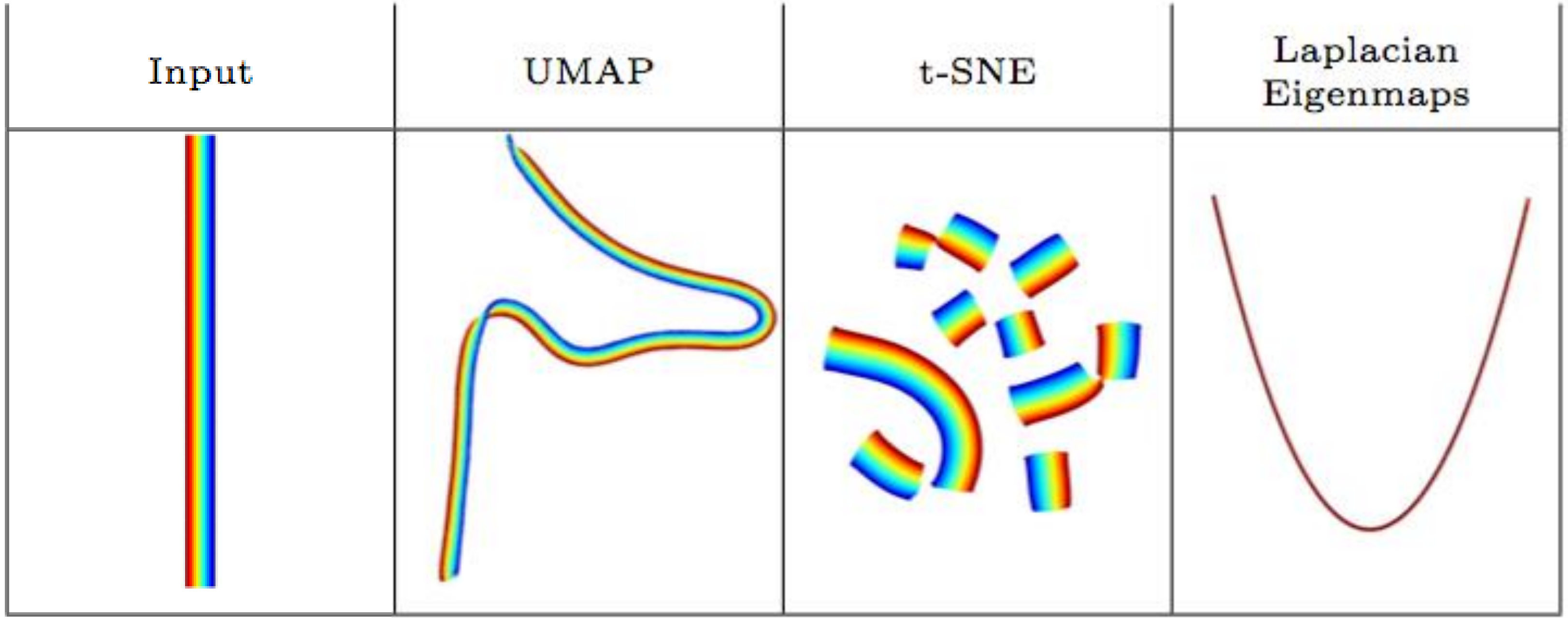
- Local neighborhoods in the data have their own parameterization (local views) with low distortion
- Local views are aligned to obtain a global embedding



[Zhang & Zha, 2004; Kohli, Cloninger & Mishne, JMLR 2021]

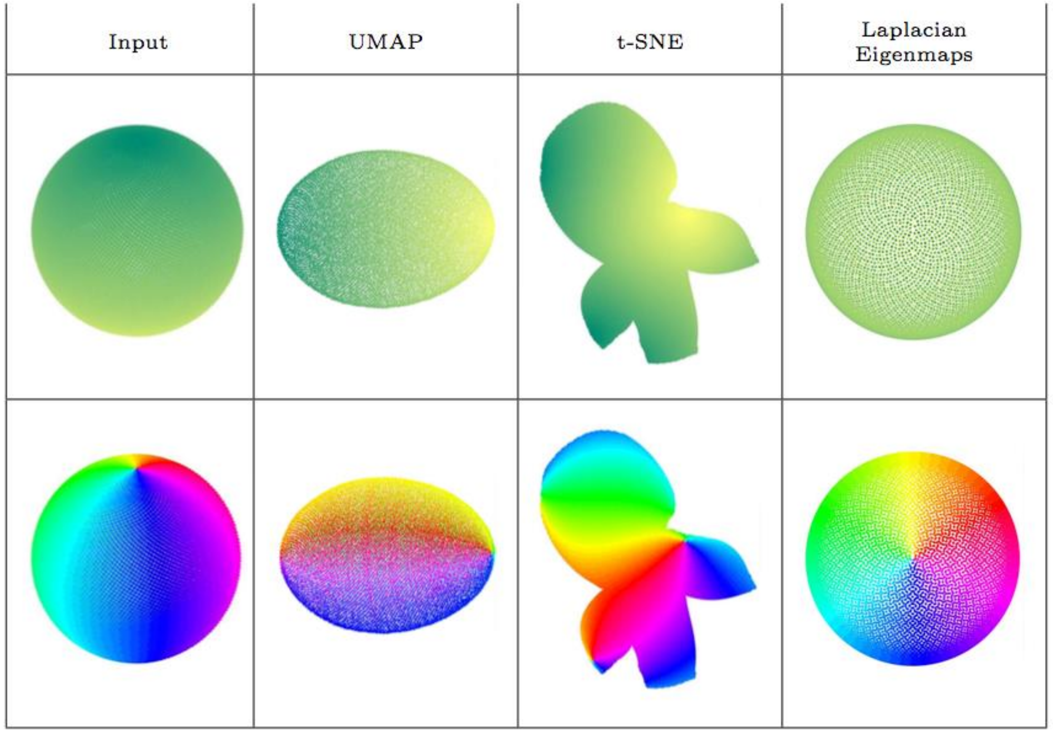
Gaps in existing manifold learning approaches

- Some methods rely on a **fixed set of global eigenvectors** of the graph Laplacian
- Embedding may not have low **distortion** everywhere.



Gaps in existing manifold learning approaches

- Cannot embed **closed and non-orientable manifolds** into their intrinsic dimension.

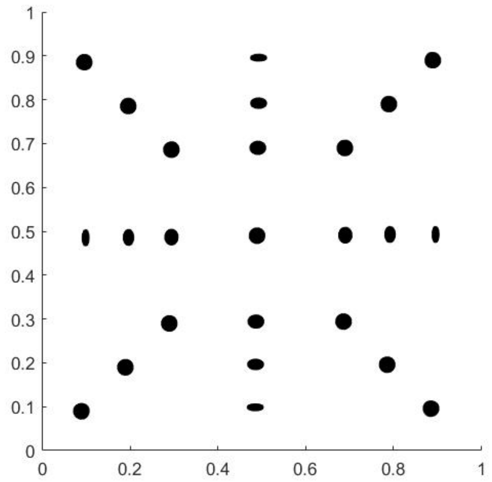


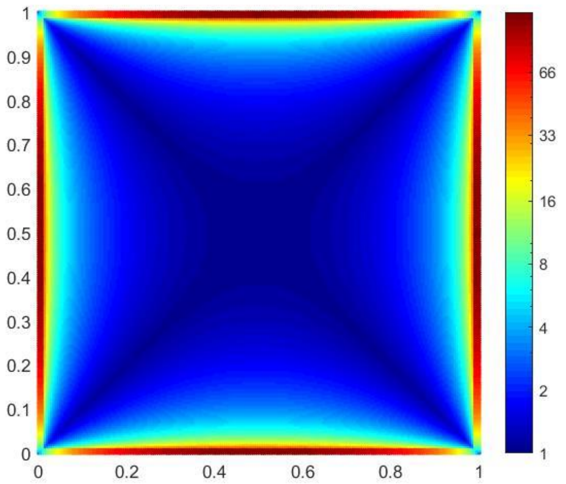
Outline

- Introduction
- **Distortion**
- Bottom-up manifold learning
- Results

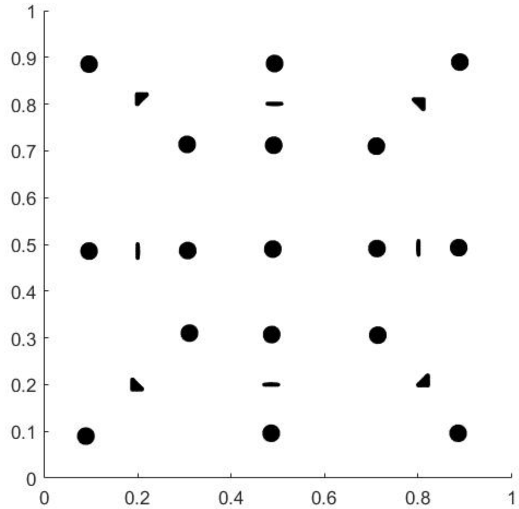
Distortion on the unit square

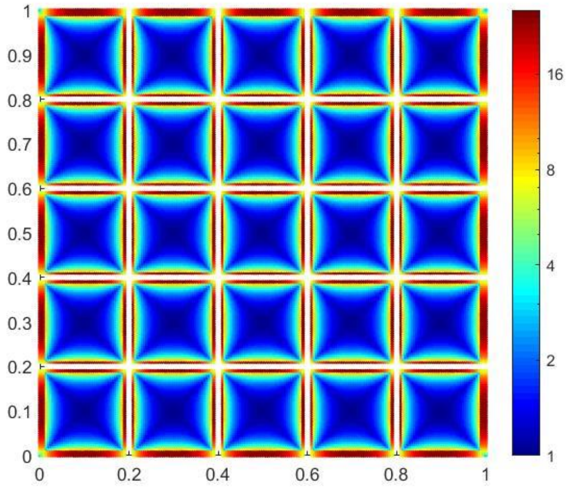
$$\Phi_1^*(x,y) = (\cos(\pi x), \cos(\pi y))$$





$$\Phi_2^*(x,y) = (\cos(5\pi x), \cos(5\pi y))$$



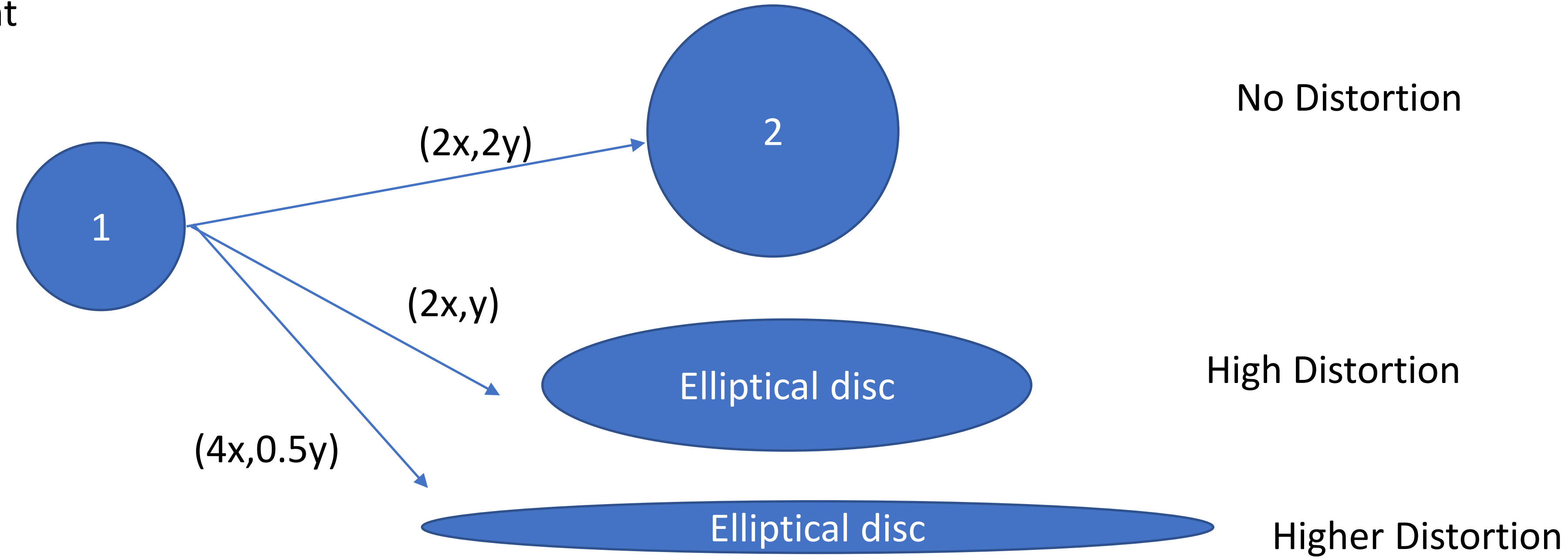


Distortion

$$\text{expansion}(\Psi) = \sup_{x,y \in X} \frac{\|\Psi(x) - \Psi(y)\|}{d(x,y)} \qquad \text{contraction}(\Psi) = \sup_{x,y \in X} \frac{d(x,y)}{\|\Psi(x) - \Psi(y)\|}$$

$$\text{distortion}(\Psi) = \text{expansion}(\Psi)\text{contraction}(\Psi)$$

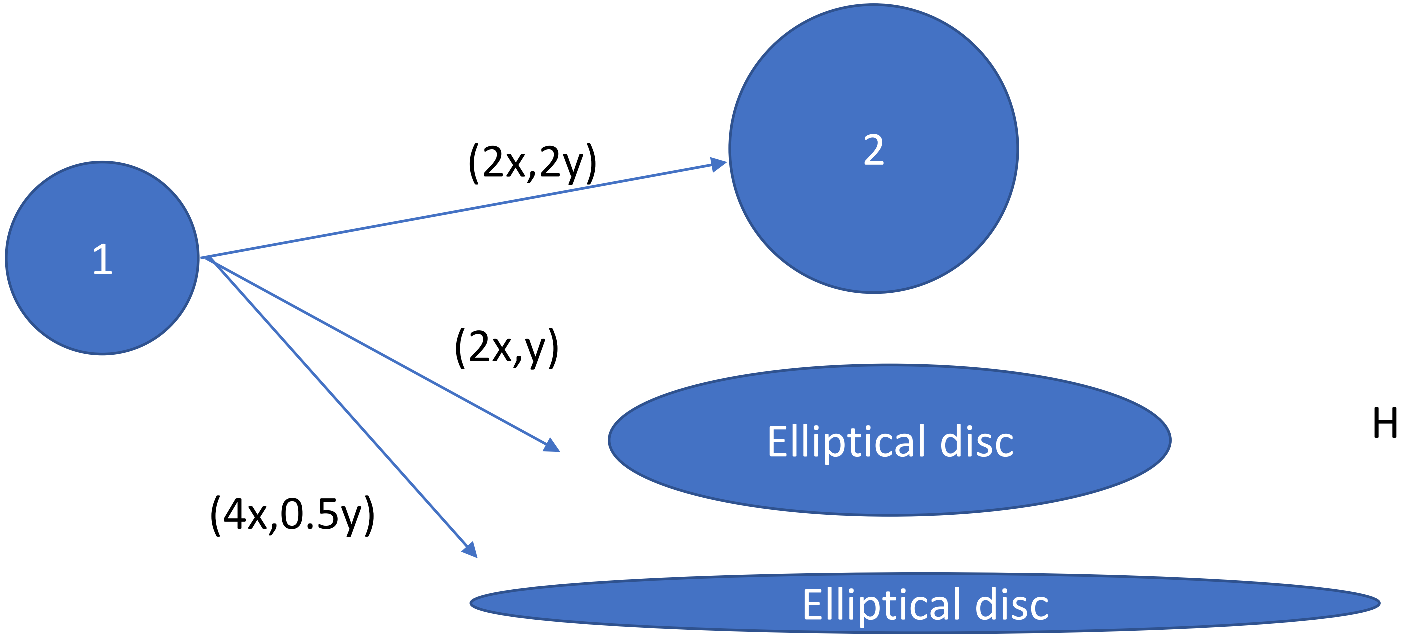
- Translation invariant
- Scale invariant



Distortion

$$\text{distortion}(\Psi) = \sup_{x,y \in X} \frac{\|\Psi(x) - \Psi(y)\|}{d(x,y)} \sup_{x,y \in X} \frac{d(x,y)}{\|\Psi(x) - \Psi(y)\|}$$

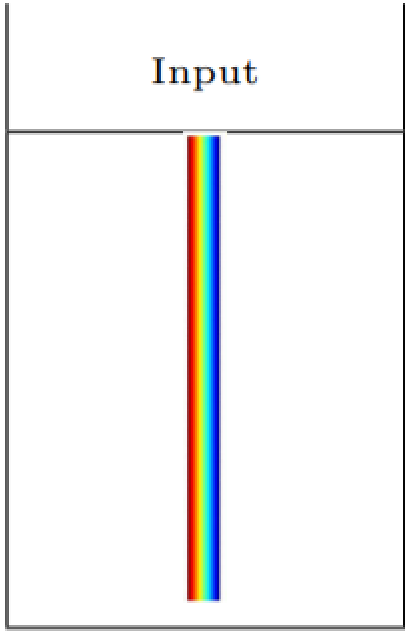
- distortion = 1 if and only if $\|\Psi_k(x) - \Psi_k(y)\|_2 = cd(x,y)$ for all $x, y \in X$ and constant c

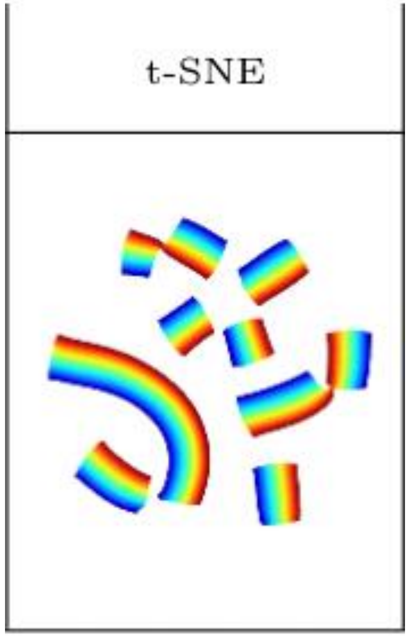


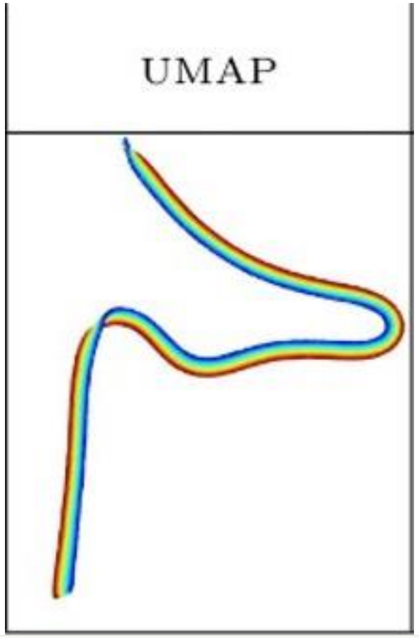
No Distortion

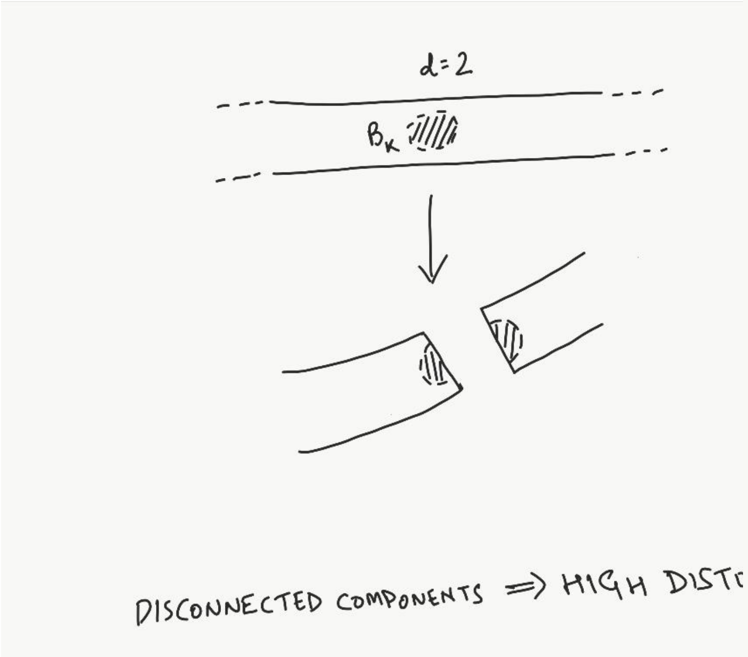
High Distortion

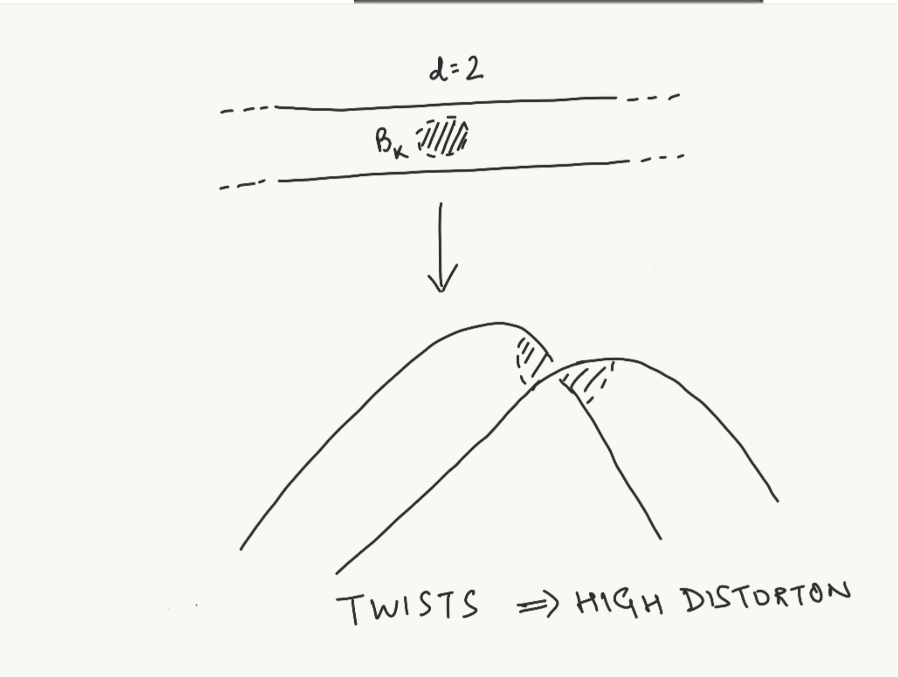
Higher Distortion


Input

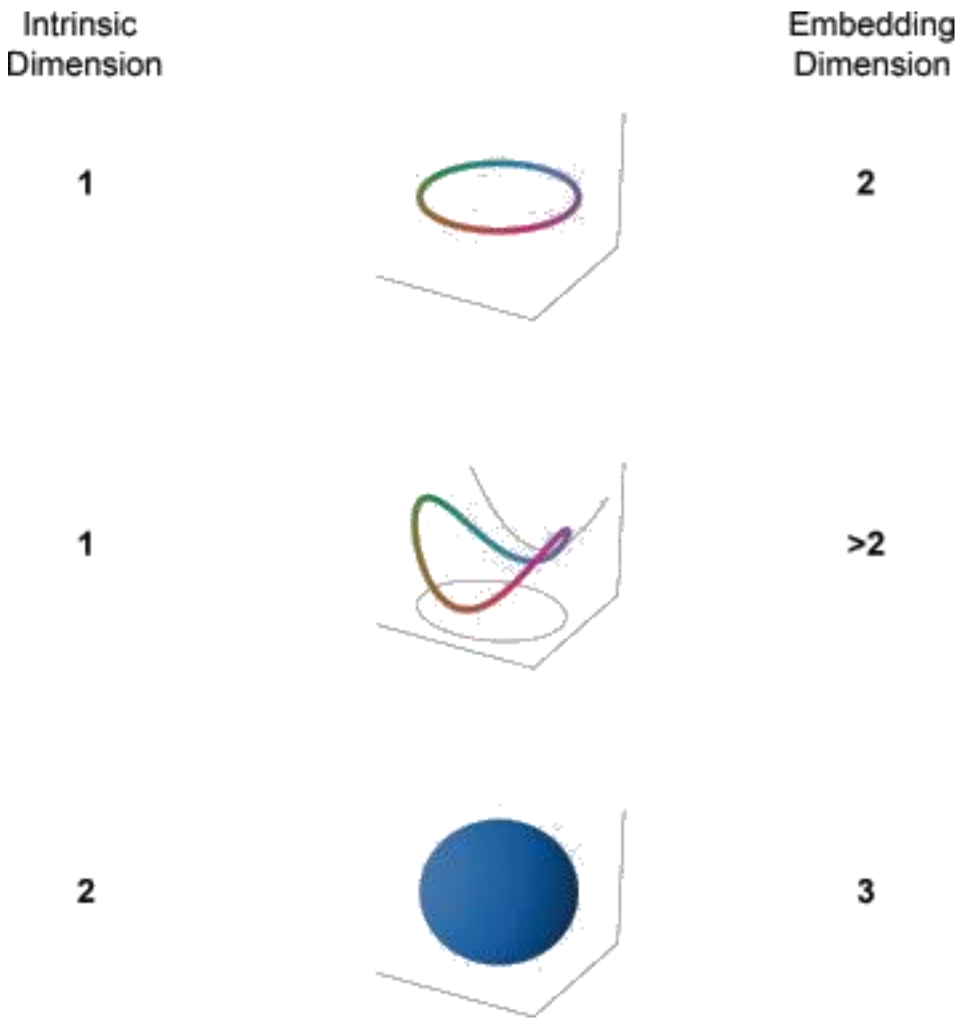

t-SNE


UMAP


d=2
B_K
DISCONNECTED COMPONENTS => HIGH DIST

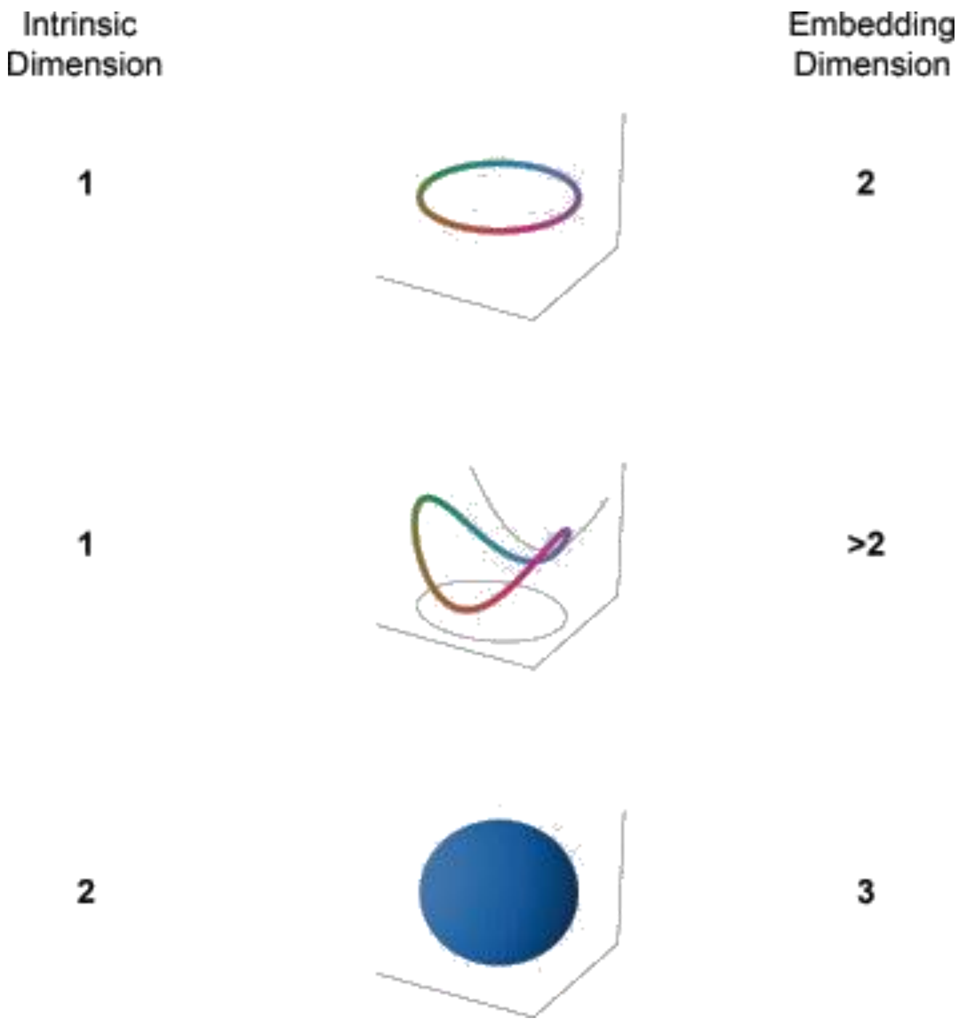

d=2
B_K
TWISTS => HIGH DISTORTON

Embedding in intrinsic dimension



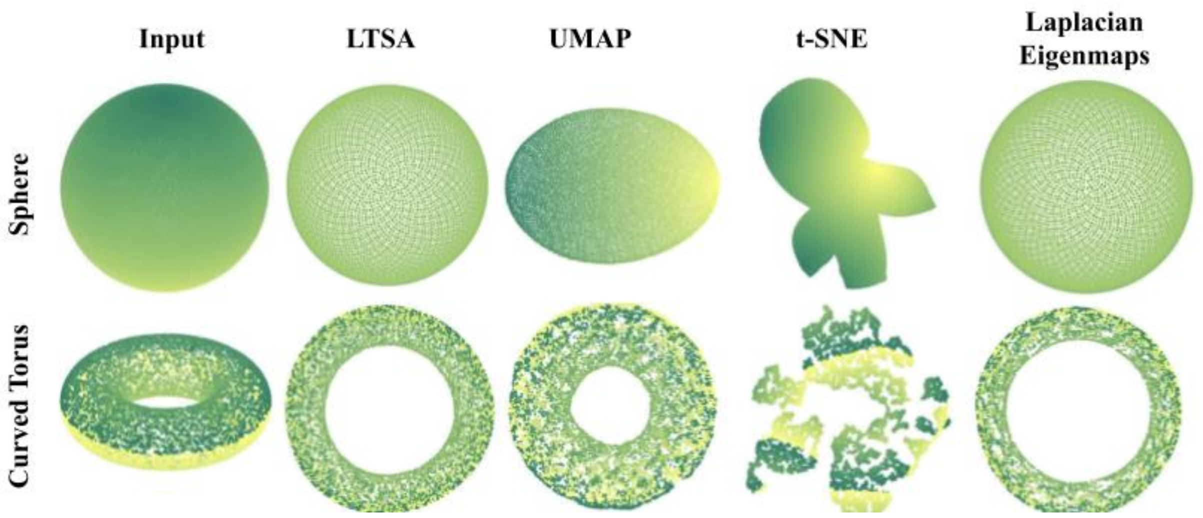
[Jazayeri & Ostojic, 2021]

Embedding in intrinsic dimension

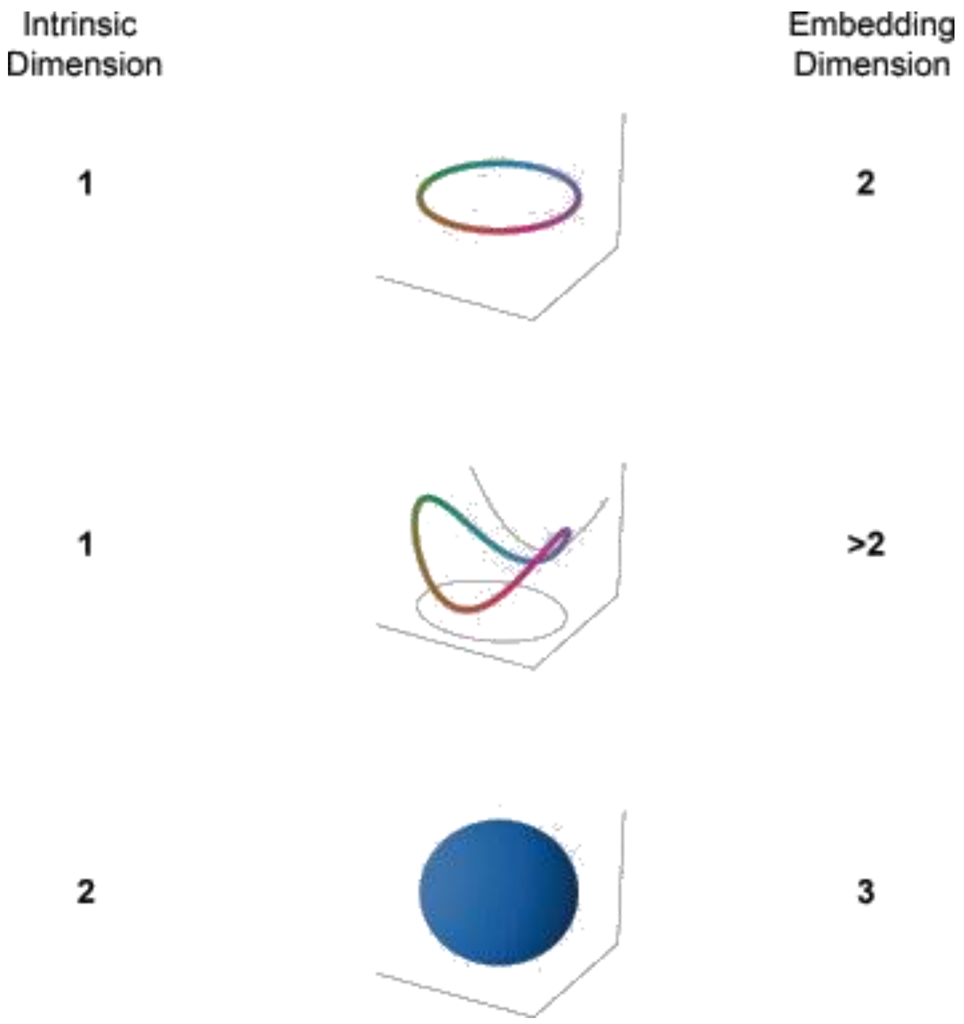


[Jazayeri & Ostojic, 2021]

What if we embed in intrinsic dimension?

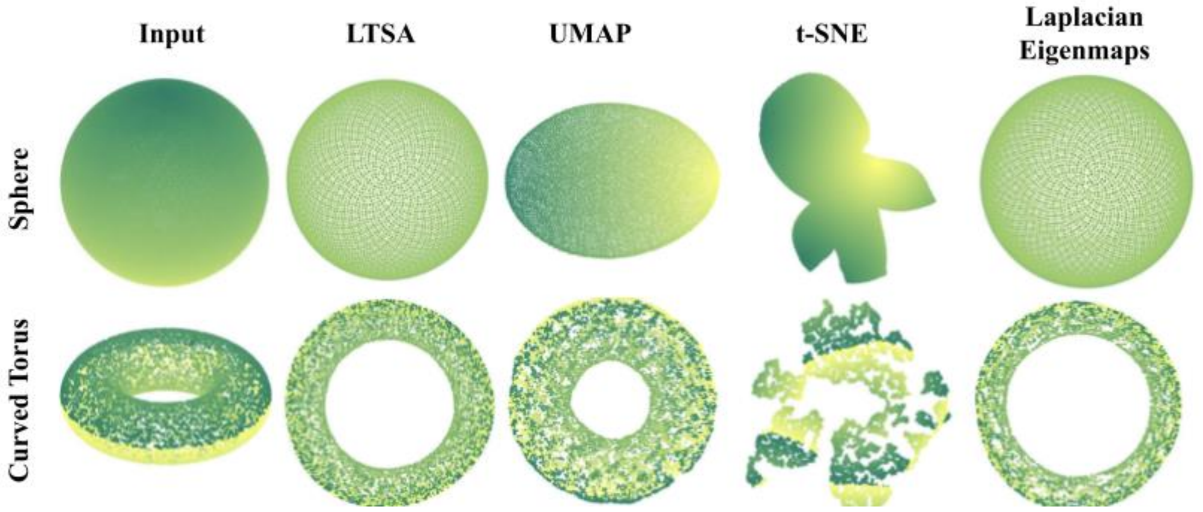


Embedding in intrinsic dimension

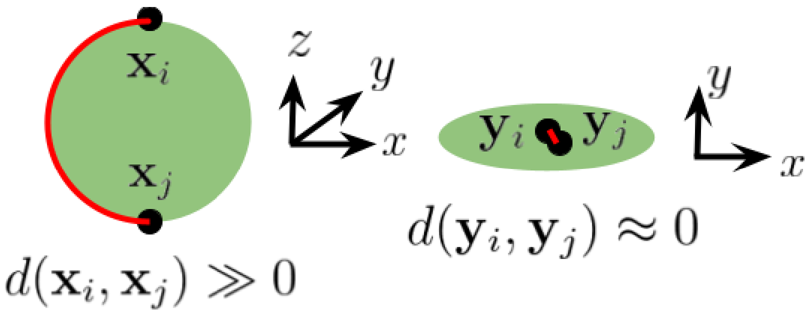


[Jazayeri & Ostojic, 2021]

What if we embed in intrinsic dimension?



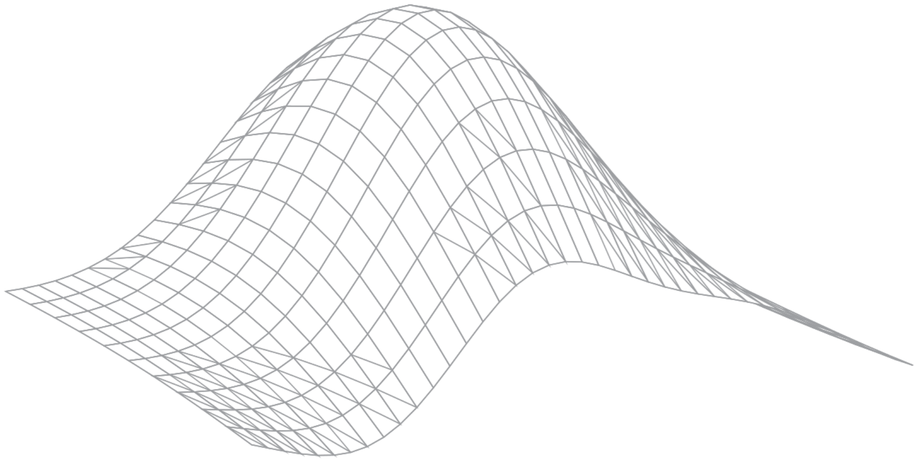
Distance distortion



Outline

- Introduction
- Distortion
- **Bottom-up manifold learning**
- Results

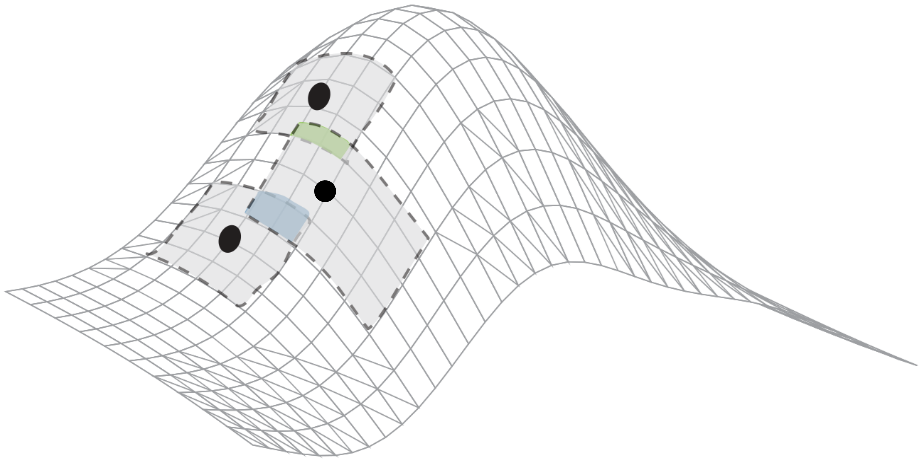
Bottom-up manifold learning



Manifold

[Nieuwenhuis*, Kohli*, Cloninger, Mishne# & Narain #, in prep]

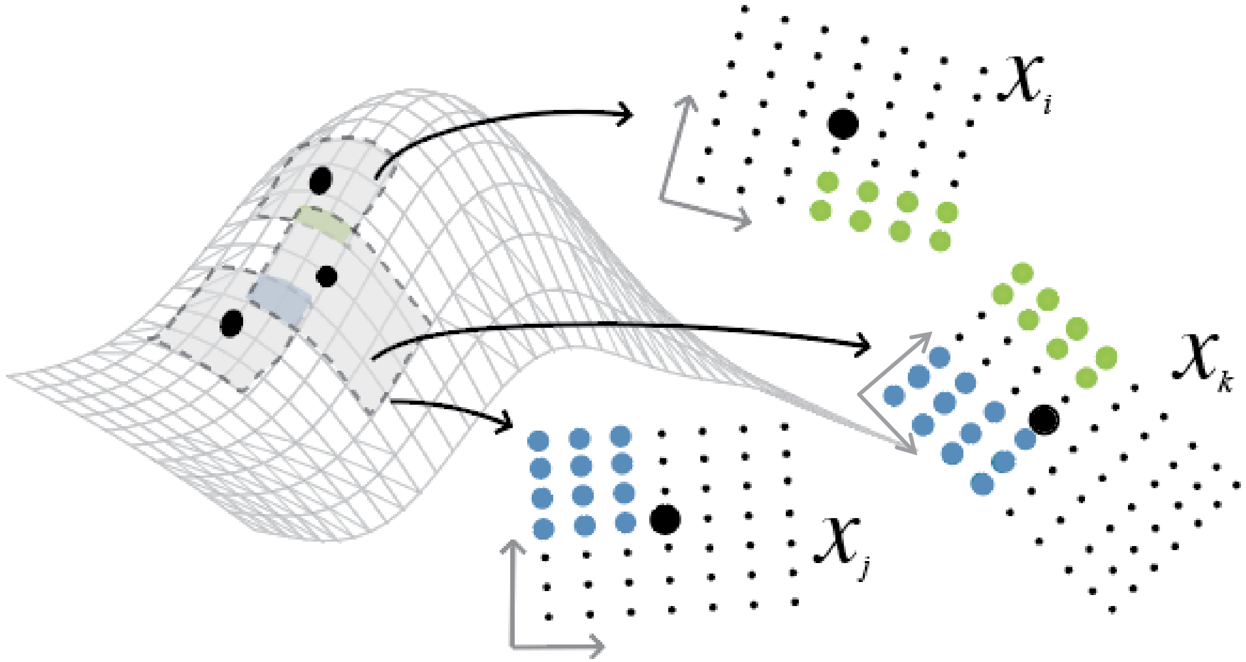
Bottom-up manifold learning



Local neighborhoods

[Nieuwenhuis*, Kohli*, Cloninger, Mishne# & Narain #, in prep]

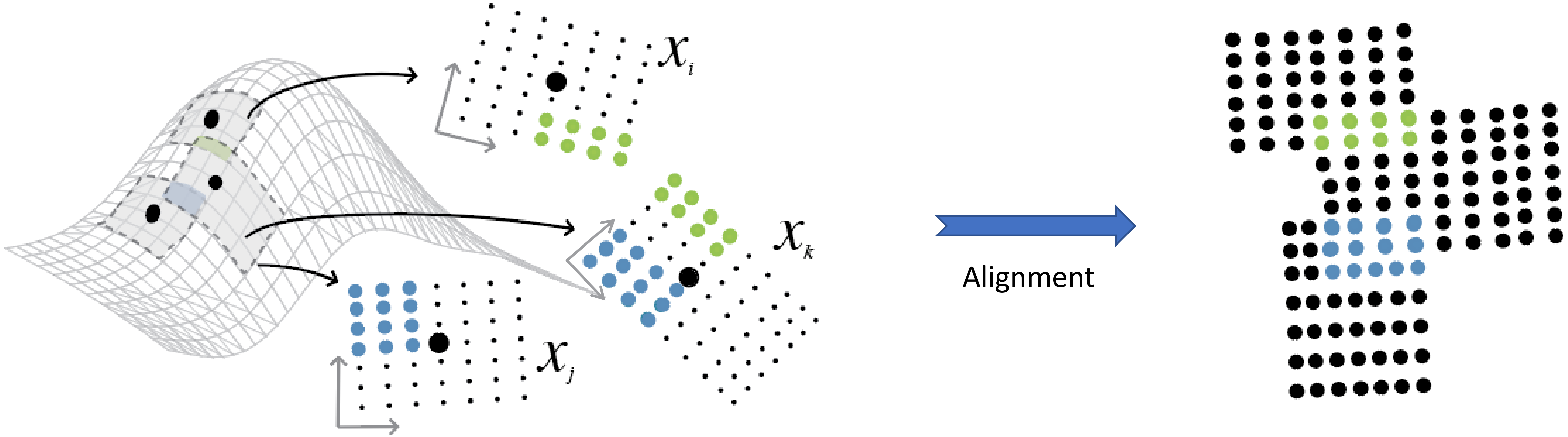
Bottom-up manifold learning



Local views

[Nieuwenhuis*, Kohli*, Cloninger, Mishne# & Narain #, in prep]

Bottom-up manifold learning



Local views

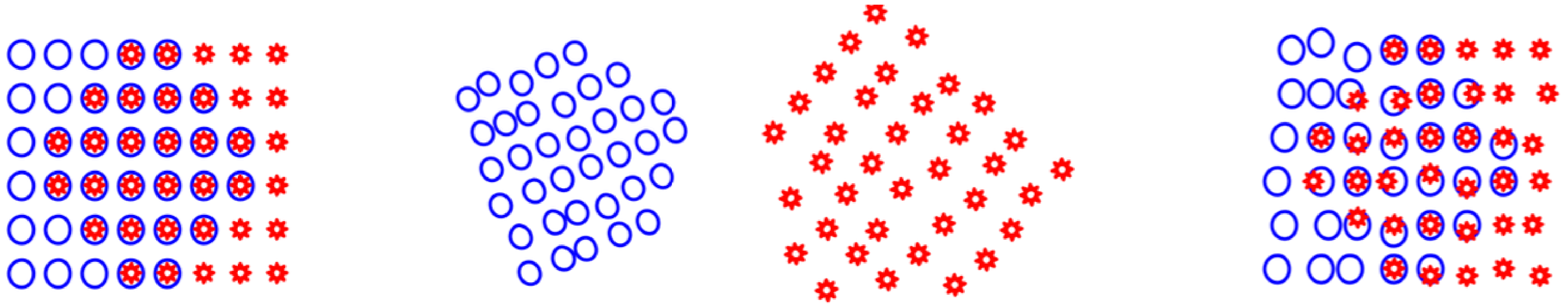
Global embedding

[Nieuwenhuis*, Kohli*, Cloninger, Mishne# & Narain #, in prep]

Global Alignment

- Goal: find a rigid transformation for each view to obtain global embedding

$$\min_{\substack{(S_i)_{i=1}^{m} \subseteq \mathbb{O}(d) \\ (t_i)_{i=1}^{m} \subseteq \mathbb{R}^d}} \sum_{\tilde{x}_k \in \widetilde{U}_i \cap \widetilde{U}_j} \left\| S_i^T (x_{k,i} + t_i) - S_j^T (x_{k,j} + t_j) \right\|_2^2$$



Local views in ambient dimension | Local views in embedding | Aligned local views in global embedding

Tearing manifolds

- Main Idea: Start with an over torn embedding laid down in intrinsic dimension and adequately stitch it back together in an iterative manner.
- In each iteration only align views who are neighbors in both the ambient space and embedding

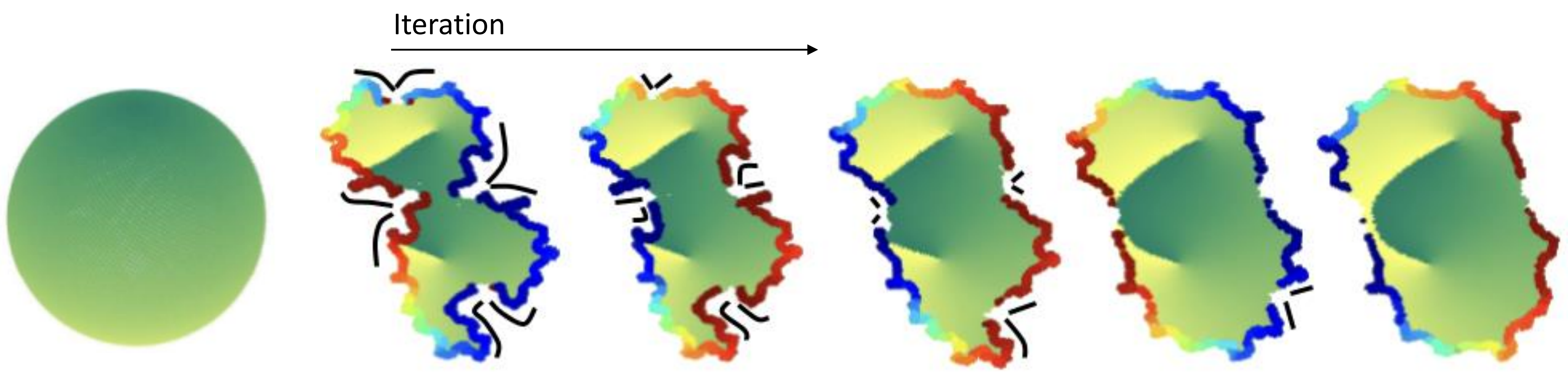
$$\min_{\substack{(S_i)_{i=1}^{m} \subseteq \mathbb{O}(d) \\ (t_i)_{i=1}^{m} \subseteq \mathbb{R}^d}} \sum_{\substack{\tilde{x}_k \in \tilde{U}_i \cap \tilde{U}_j \\ x_k \in U_i \cap U_j}} \left\| S_i^T (x_{k,i} + t_i) - S_j^T (x_{k,j} + t_j) \right\|_2^2$$

$\tilde{x}_k \in \tilde{U}_i \cap \tilde{U}_j$ - Neighbors in ambient space

$x_k \in U_i \cap U_j$ - Neighbors in embedding space

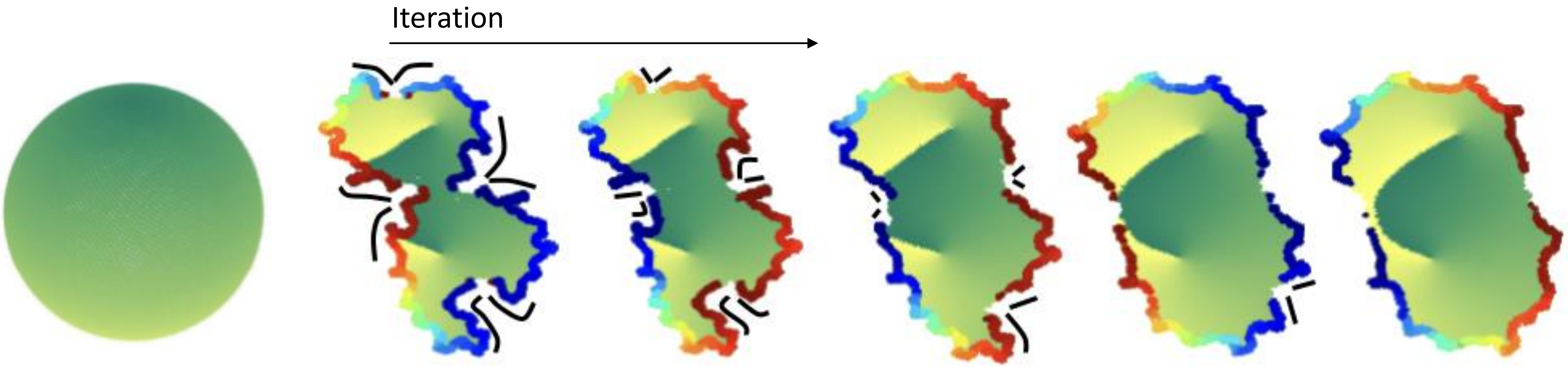
- End goal: for closed and non-orientable manifolds, a tear must be retained while for other manifolds the tear will vanish automatically.

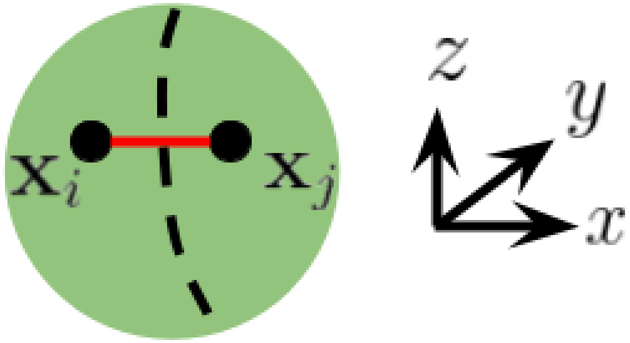
Tear-aware Riemannian Alignment



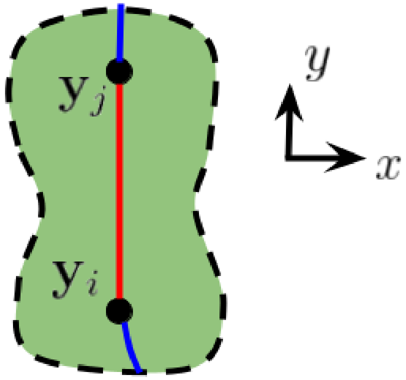
t-SNE

Tear-aware Riemannian Alignment





Tear-aware “teleportation” distance



Global distortion

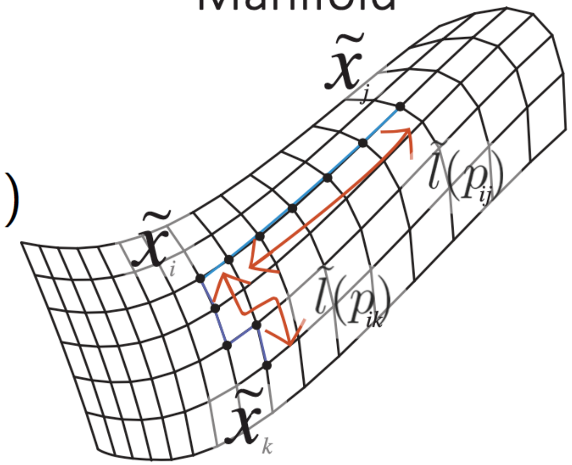
Shortest path in the embedding space

Shortest path in the ambient space

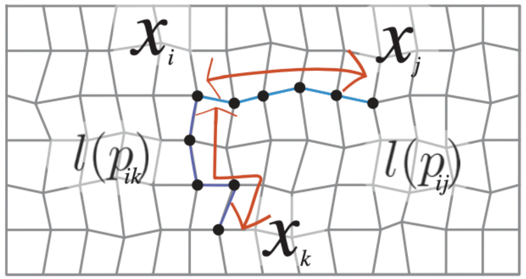
$$\mathcal{G} = \max_{j \neq k} \frac{l(\tilde{p}_{kj})}{\tilde{l}(\tilde{p}_{kj})} \max_{j \neq k} \frac{\tilde{l}(\tilde{p}_{kj})}{l(\tilde{p}_{kj})}$$

$$\tilde{l}(p) = \sum_{j=1}^{s-1} d_g(\tilde{\mathbf{x}}_{i_j}, \tilde{\mathbf{x}}_{i_{j+1}})$$

Manifold



Low-Dim Embedding



$$l(p) = \sum_{j=1}^{s-1} \left\| \mathbf{x}_{i_j} - \mathbf{x}_{i_{j+1}} \right\|_2$$

[Kohli, Cloninger & Mishne, in prep]

Bounds on global distortion

- Intuitively, if the local distortions are low and the alignment error is low then the global distortion should be low too

Theorem

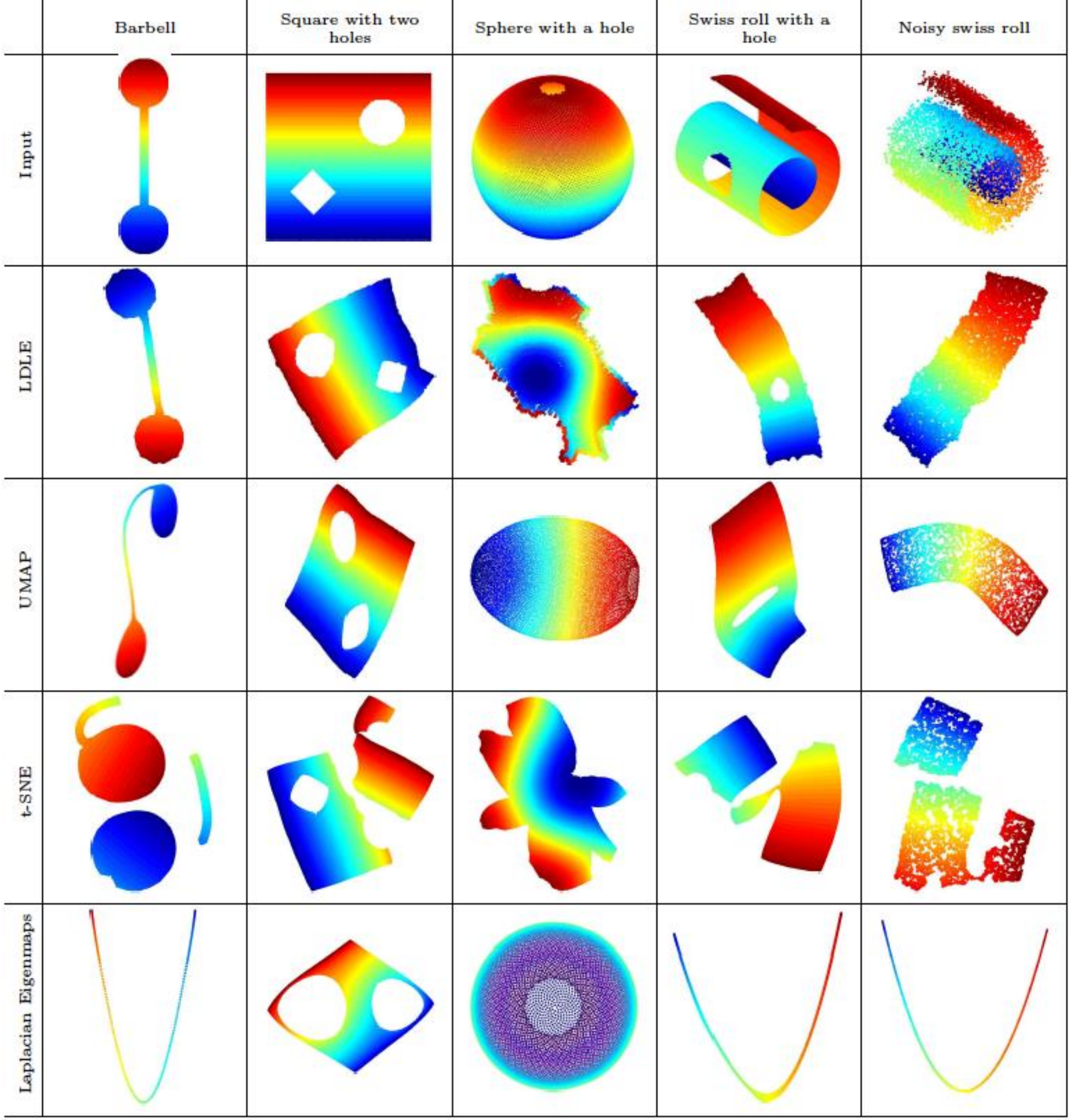
Let κ_i be the distortion of the ith view. Suppose that the alignment error $F_\Gamma(\mathbf{S})$ is small enough then $\mathcal{G}^w \leq \frac{1}{2}\kappa[\kappa + O(\sqrt{F_\Gamma(\mathbf{S})})]$, where $\kappa = \max_1^m \kappa_i$.

[Kohli, Cloninger & Mishne, in prep]

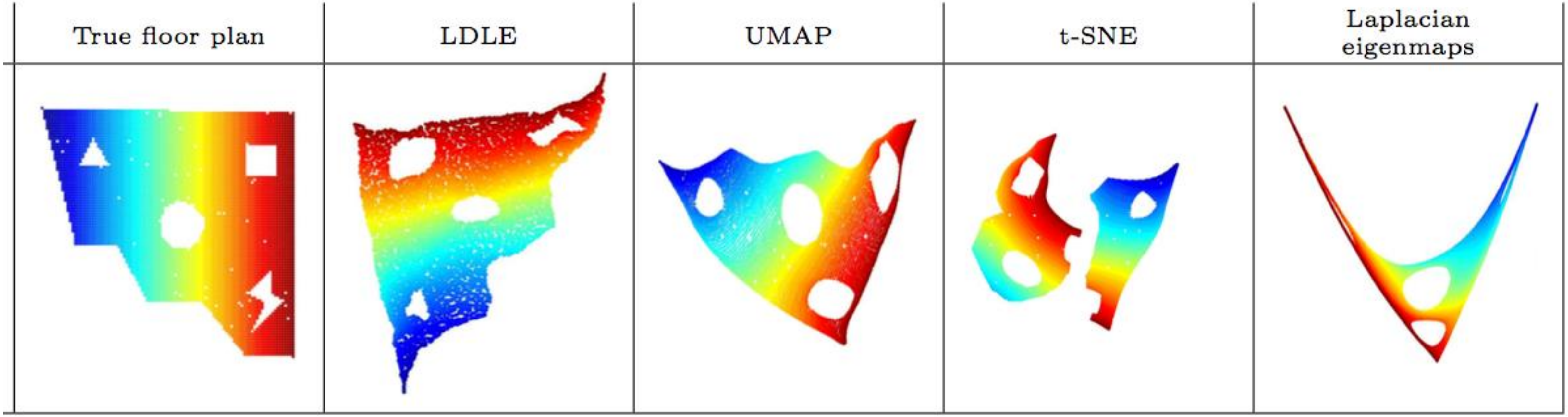
Outline

- Introduction
- Distortion
- Bottom-up manifold learning
- **Results**

Manifolds with boundary

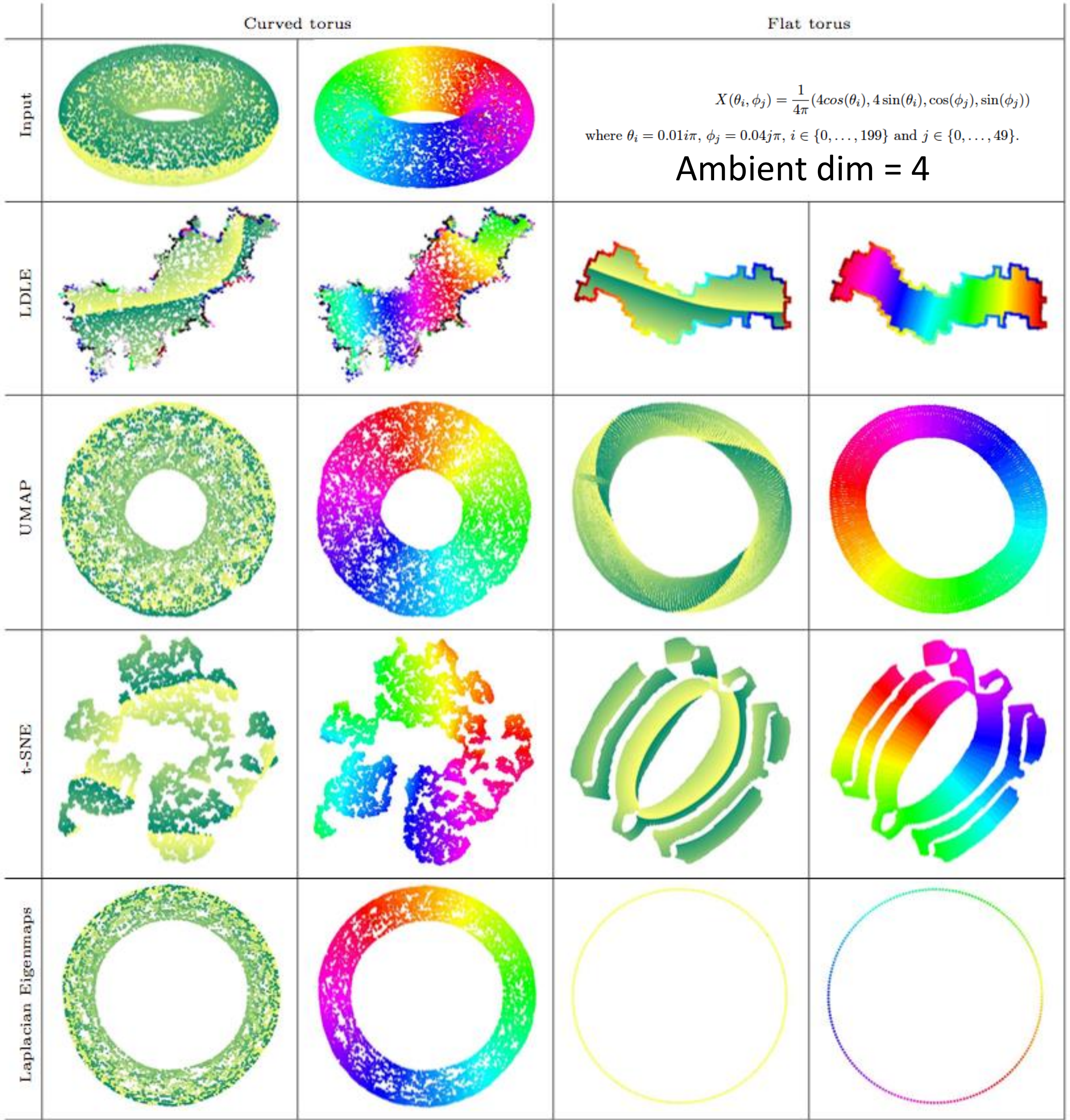


Manifolds with boundary – high-dim example

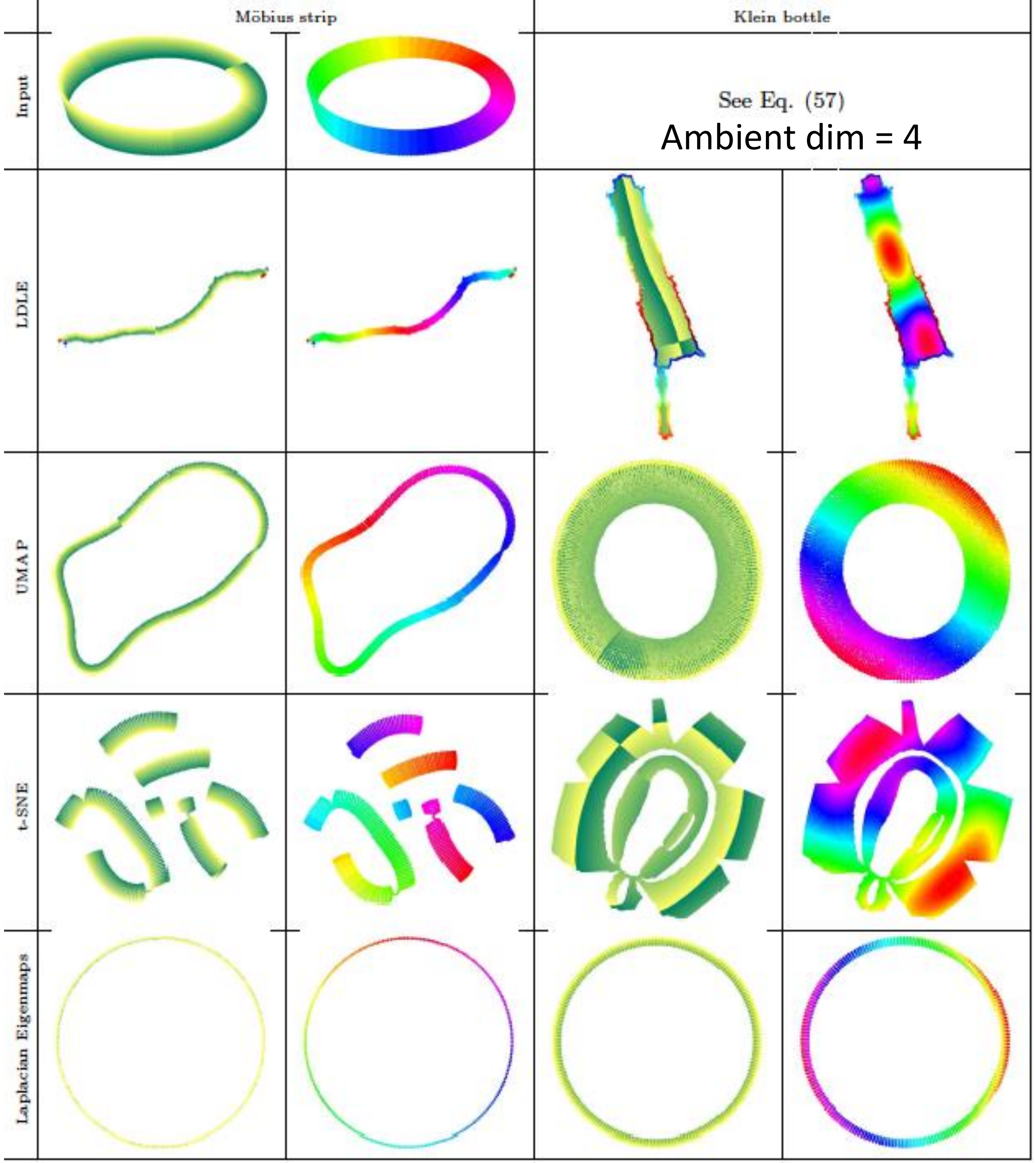


Ambient dim = 42

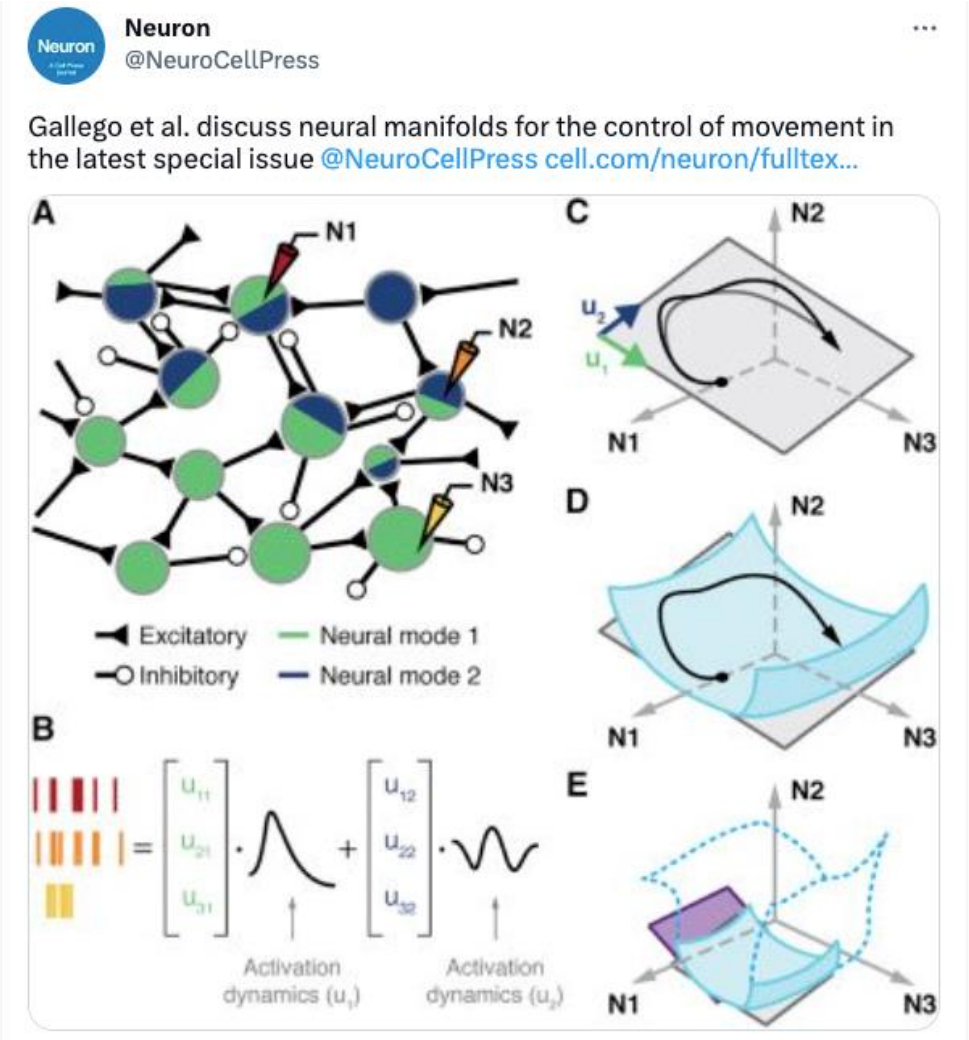
Manifolds without boundary

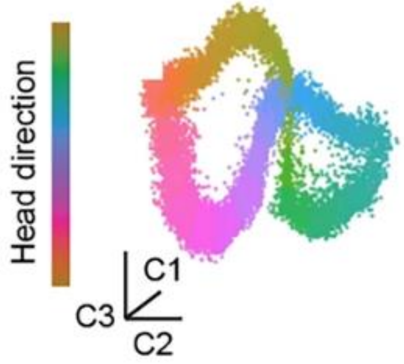


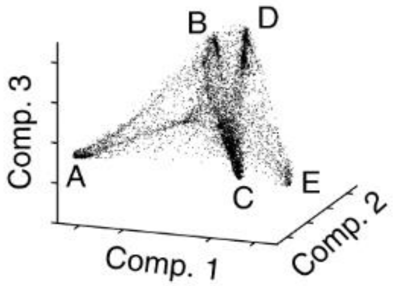
Non-orientable manifolds

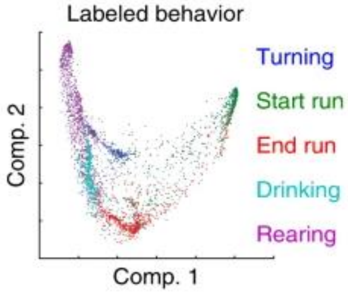


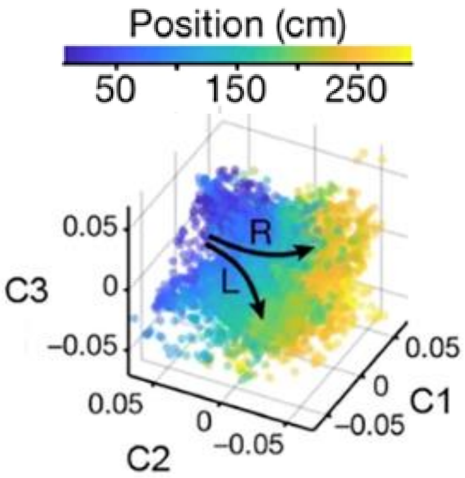
Neural Manifolds





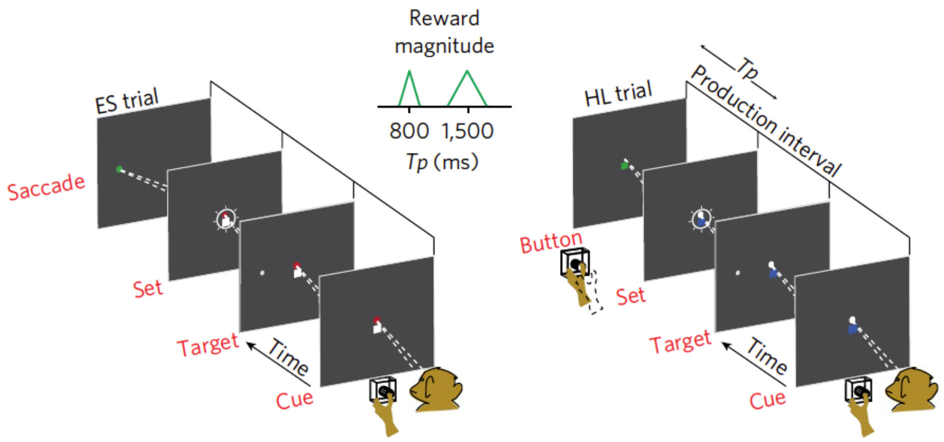




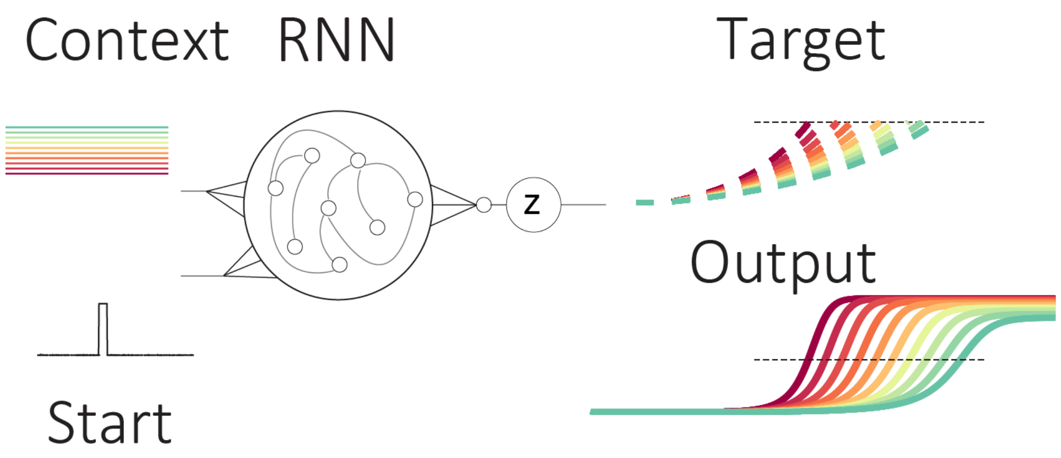


[Cunningham and Yu 2014,
Gao and Ganguli 2015,
Remington et al. 2018,
Chaudhuri et al. 2019,
Rubin et al. 2019,
Nieh at al. 2021,
…]

Motor timing

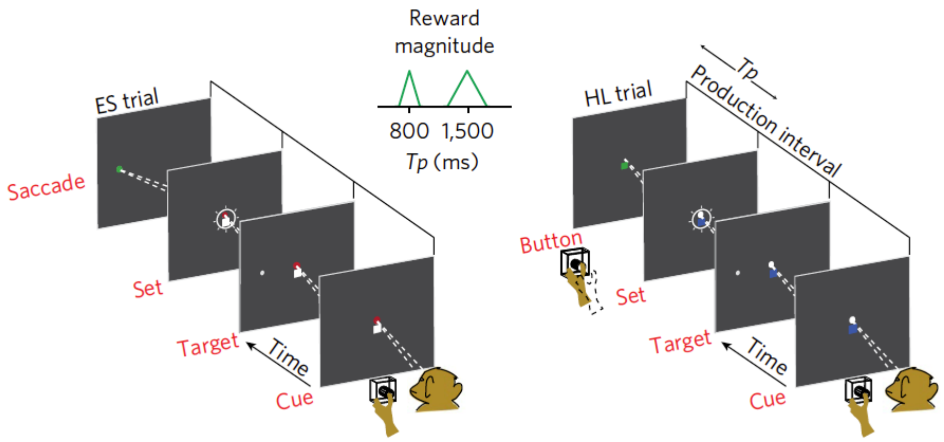


[Wang*, Narain*, Hosseini & Jazayeri, 2018]

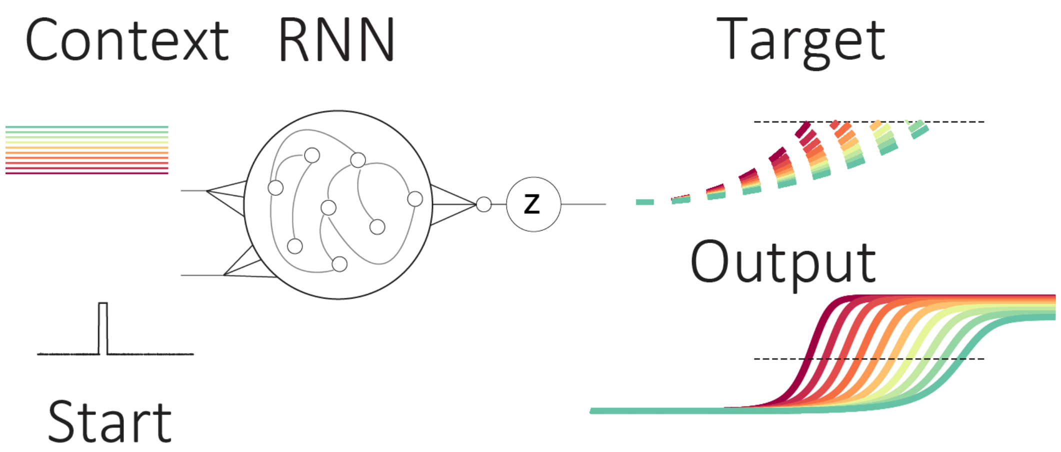


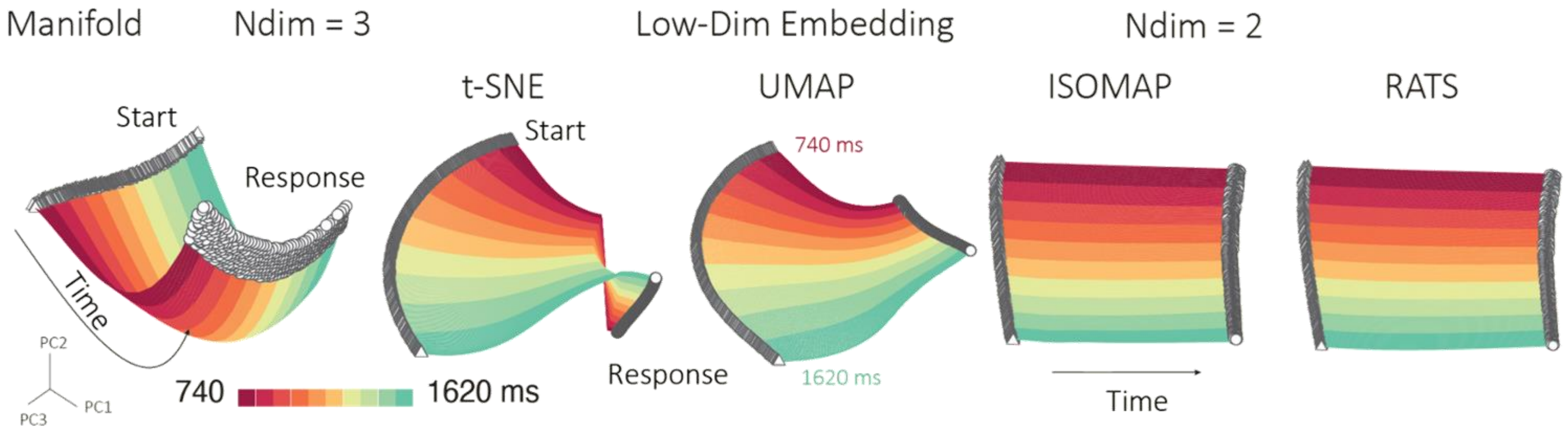
[Nieuwenhuis*, Kohli*, Cloninger, Mishne# & Narain #, in prep]

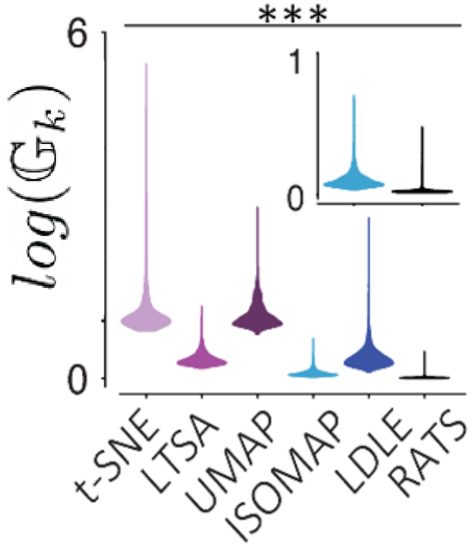
Motor timing



[Wang*, Narain*, Hosseini & Jazayeri, 2018]



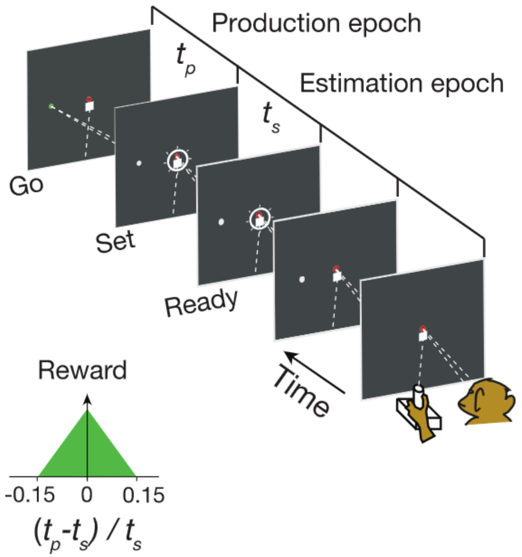


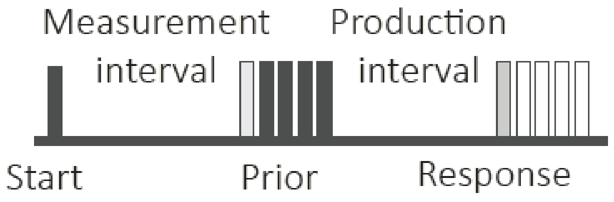


[Nieuwenhuis*, Kohli*, Cloninger, Mishne# & Narain #, in prep]

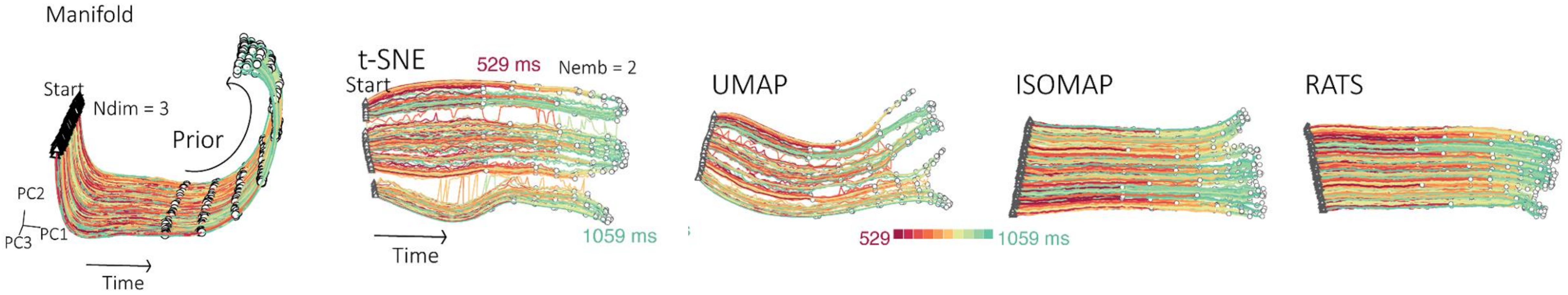
Interval reproduction

Ready-Set-Go task: time reproduction
[Sohn*, Narain*, Meirhaeghe* & Jazayeri (2019)]



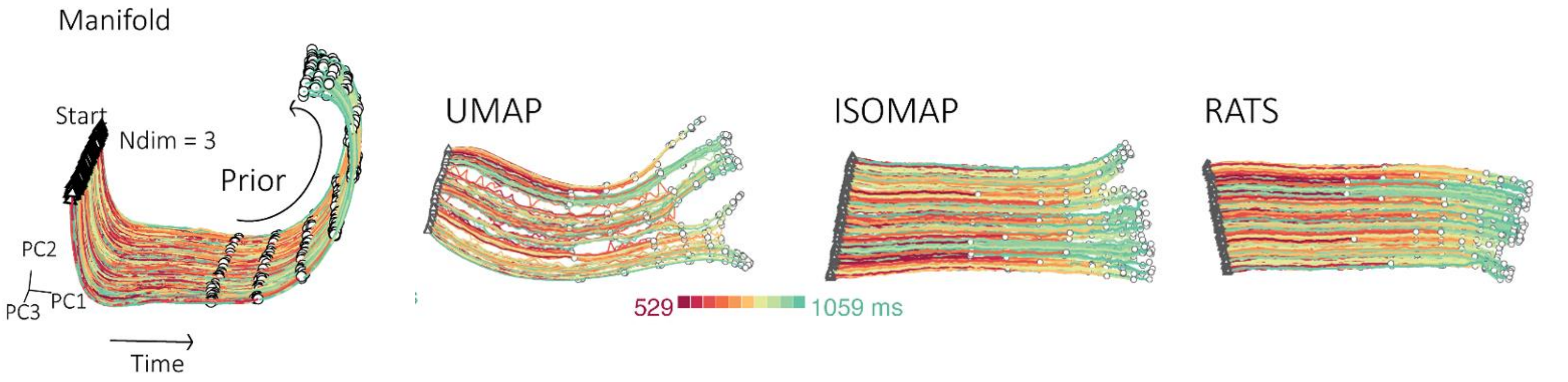


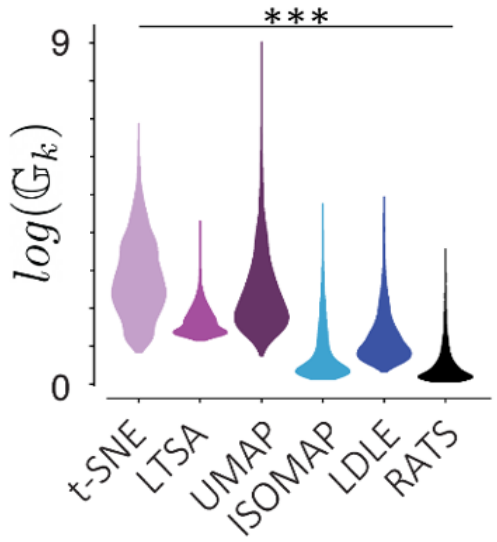
We trained an RNN that recapitulates dMFC population dynamics

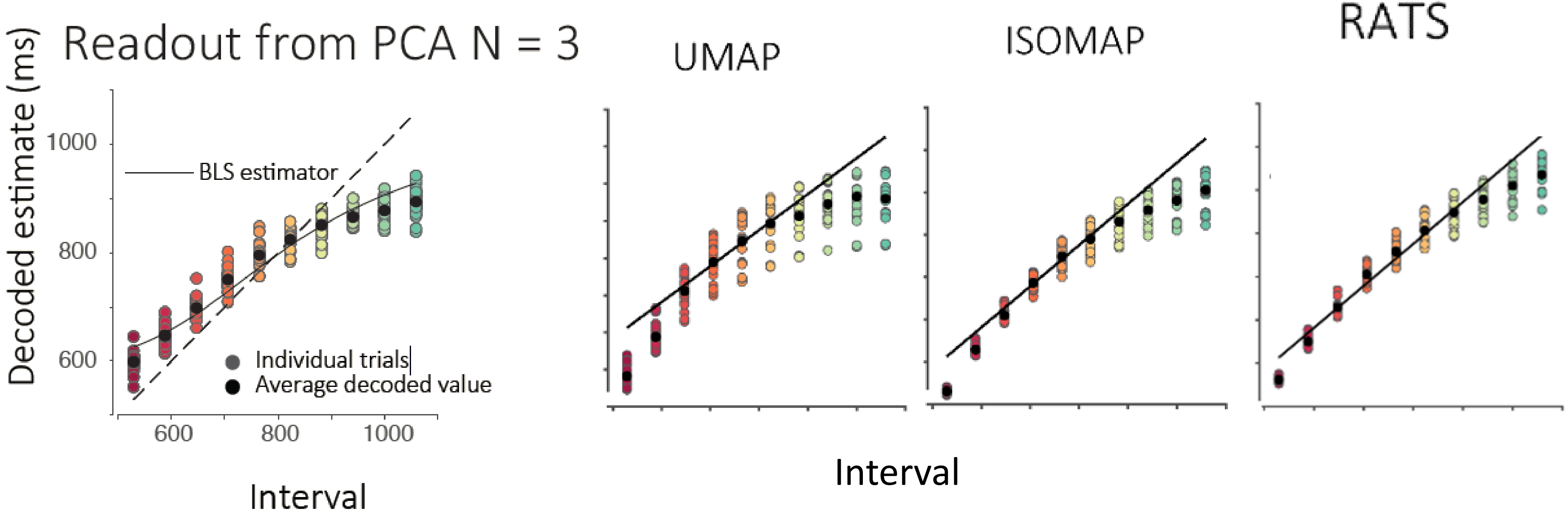


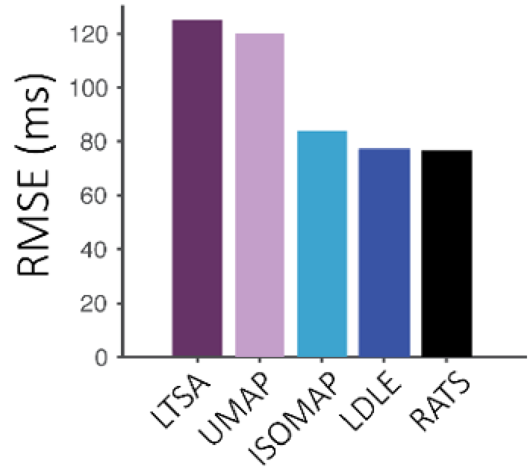
[Nieuwenhuis*, Kohli*, Cloninger, Mishne# & Narain #, in prep]

Interval reproduction





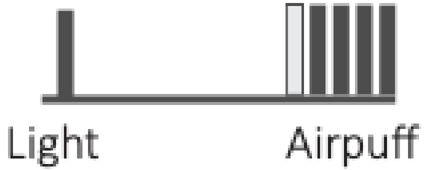


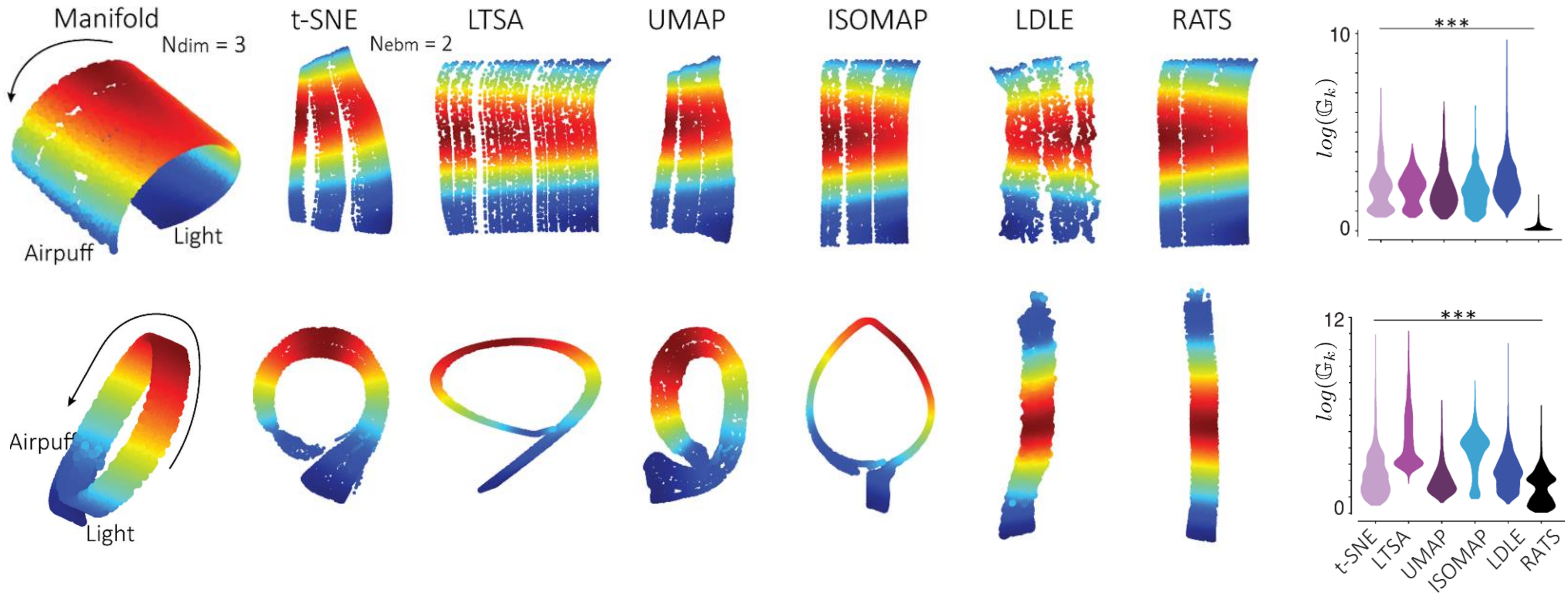


[Nieuwenhuis*, Kohli*, Cloninger, Mishne# & Narain #, in prep]

Trace conditioning

Associative learning of time elapsed between two brief sensory cues

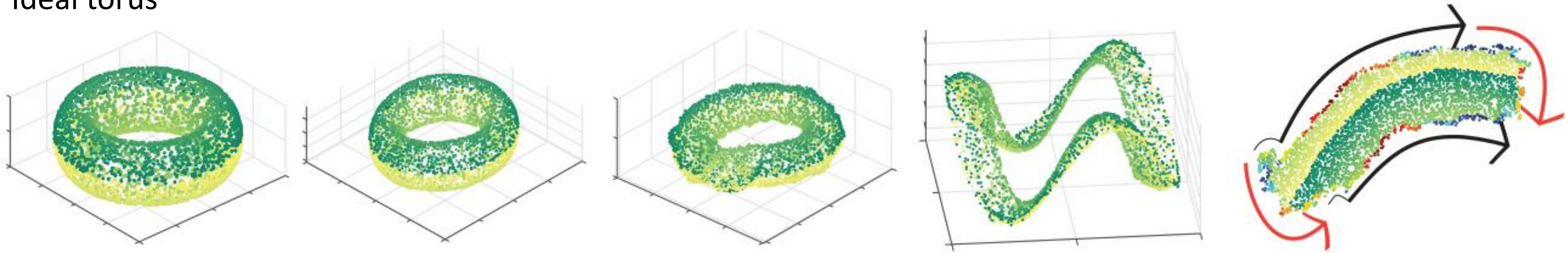




[Nieuwenhuis*, Kohli*, Cloninger, Mishne# & Narain #, in prep]

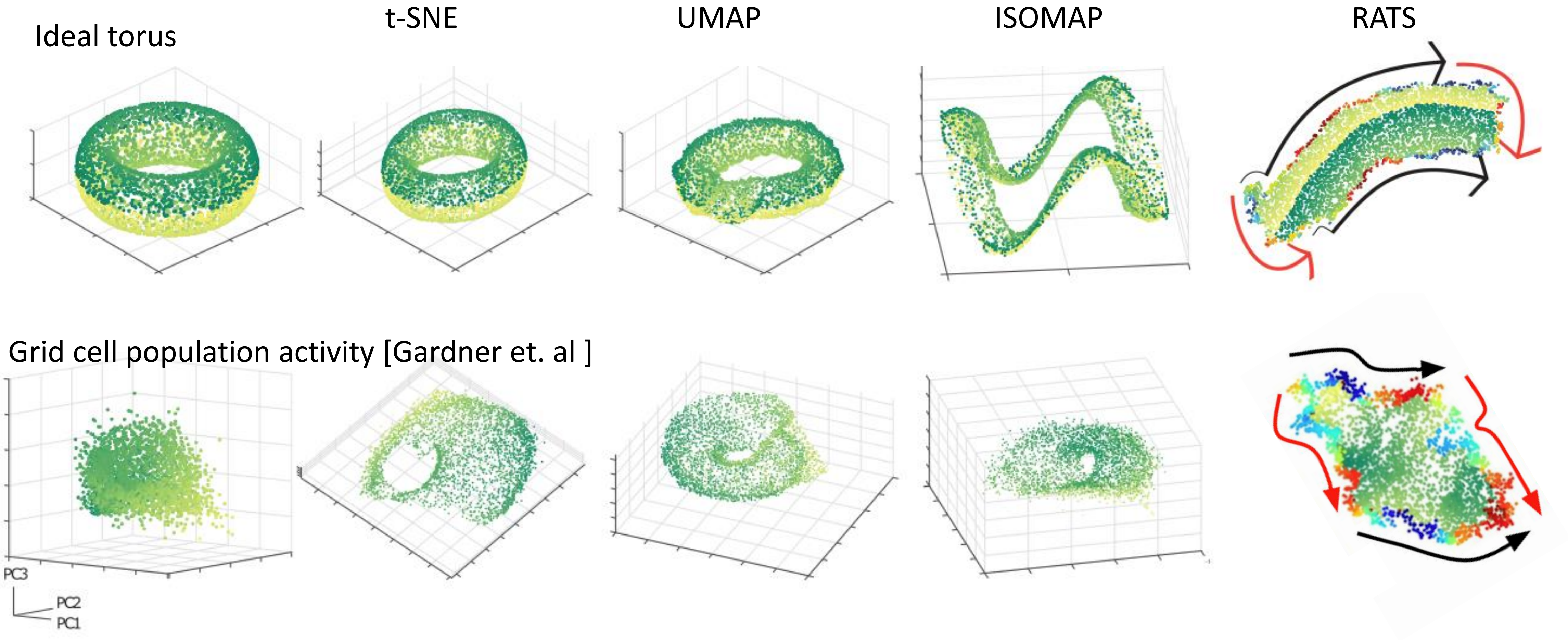
Toroidal manifolds

Ideal torus | t-SNE | UMAP | ISOMAP | RATS



[Nieuwenhuis*, Kohli*, Cloninger, Mishne# & Narain #, in prep]

Toroidal manifolds

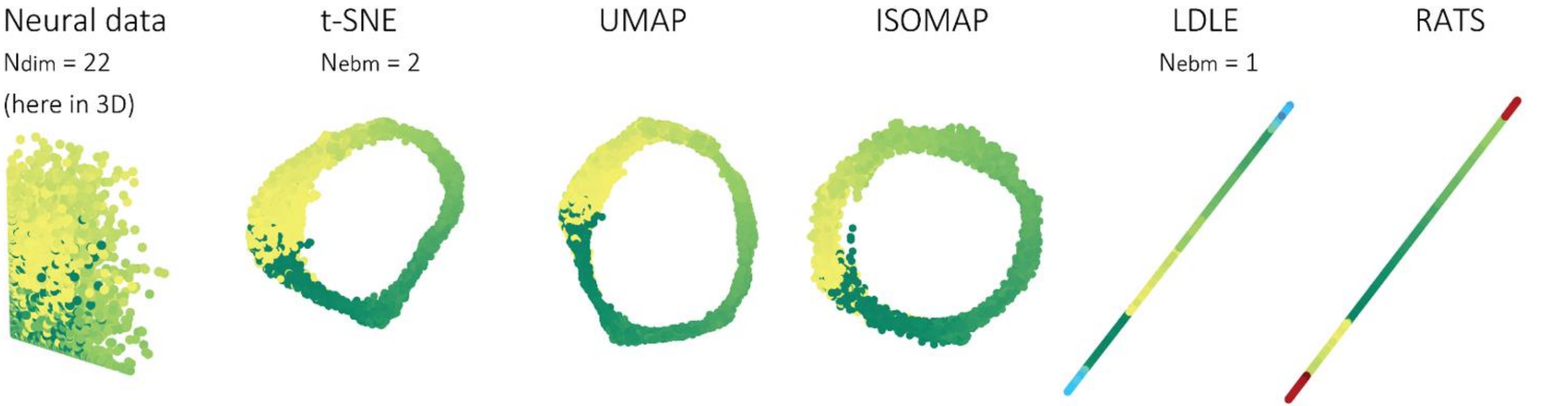


[Nieuwenhuis*, Kohli*, Cloninger, Mishne# & Narain #, in prep]

Head direction

Circuit that uses external and internal cues to estimate the direction the animal is heading with respect to the external world

[Chaudhuri et al. 2019]



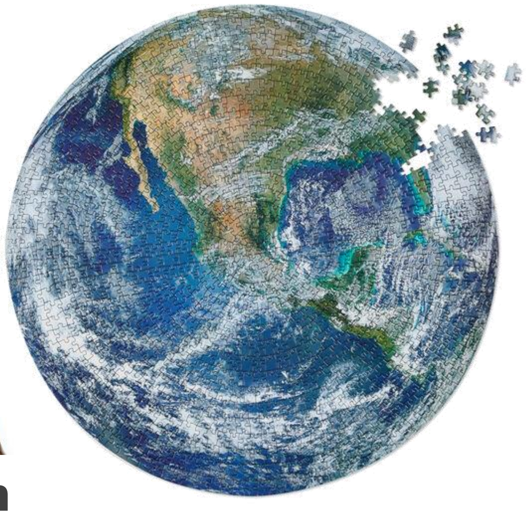
[Nieuwenhuis*, Kohli*, Cloninger, Mishne# & Narain #, in prep]

Bas Nieuwenhuis

Dhruv Kohli

Alex Cloninger

Devika Narain



W.M. KECK FOUNDATION

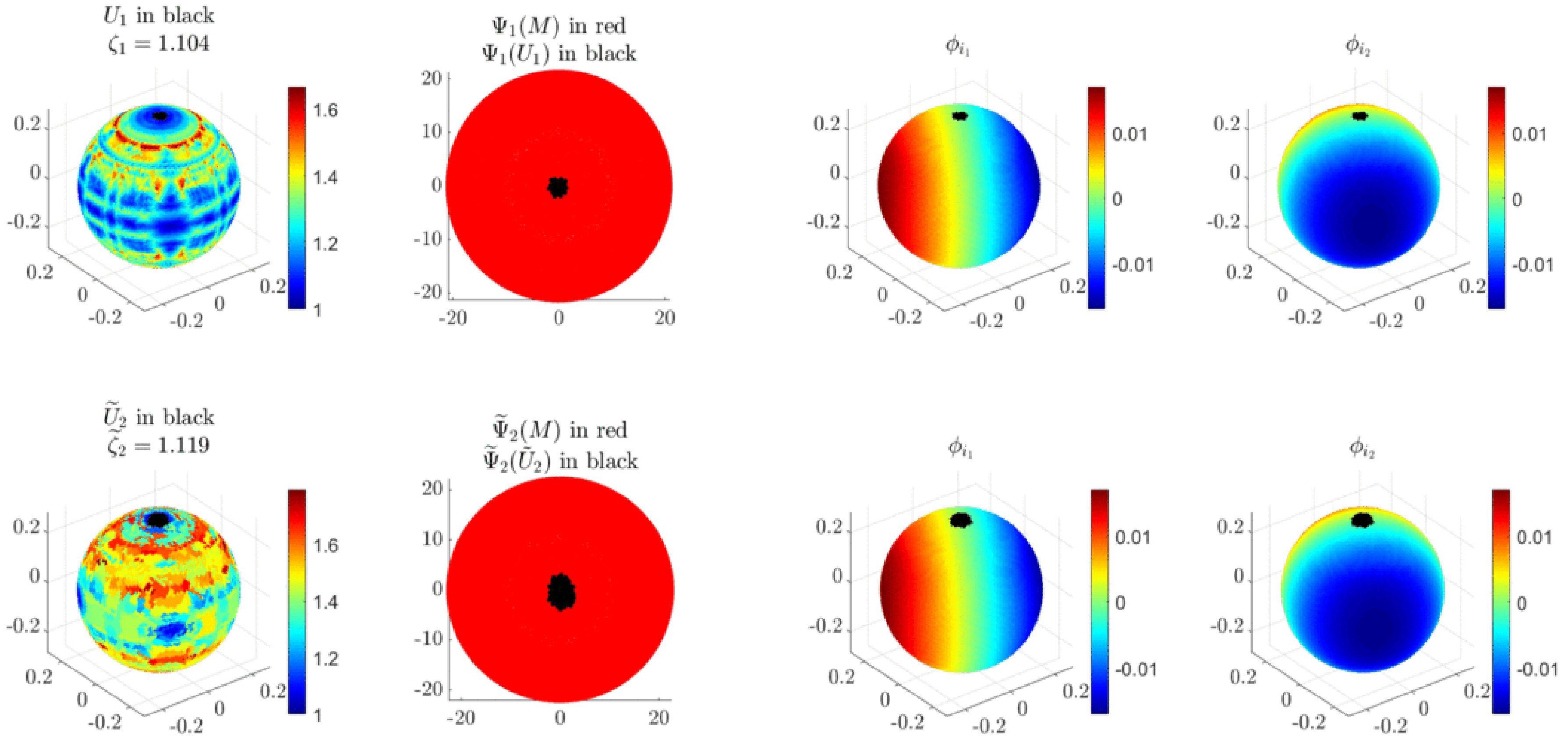
NSF

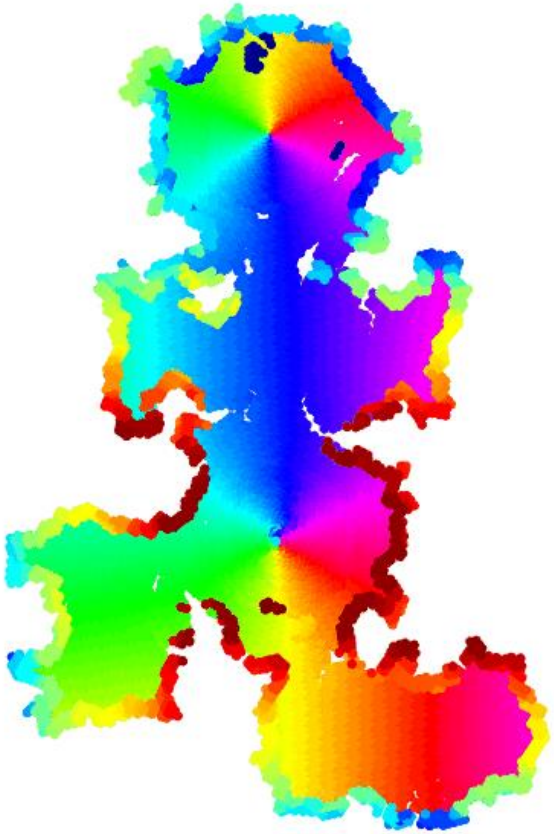
NIH

SIMONS FOUNDATION

gmishne@ucsd.edu

@gmishne





Iter 0

Thank You!

arXiv:2101.11055

https://pyldle2.readthedocs.io/en/latest/

gmishne@ucsd.edu

@gmishne