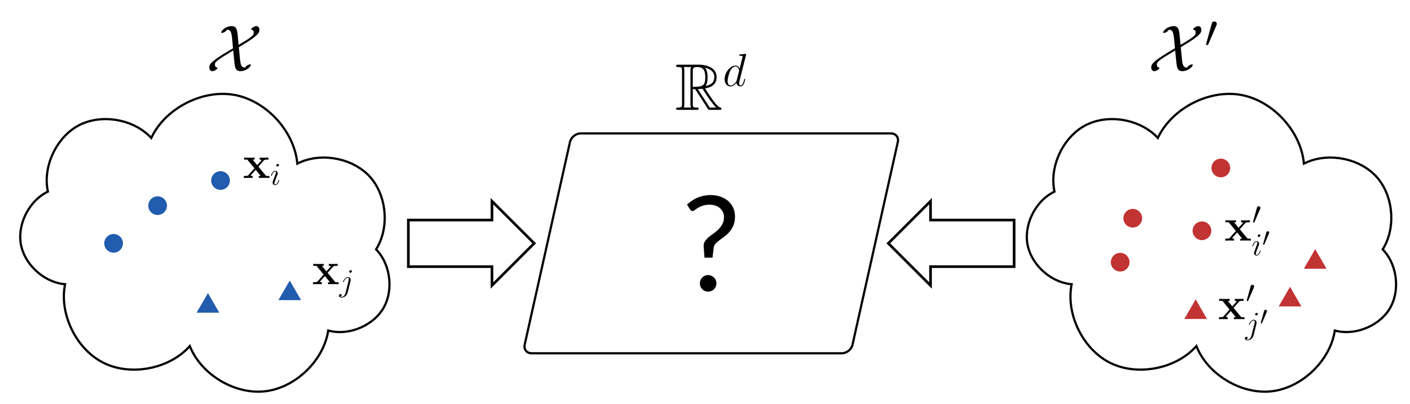
Unsupervised Manifold Alignment with Joint Multidimensional Scaling

Dexiong Chen †,1,2 Bowen Fan †,1,2 Carlos Oliver 1,2 Karsten Borgwardt 1,2

†Equal contribution [1]Department of Biosystems Science and Engineering, ETH Zürich, Switzerland [2]SIB Swiss Institute of Bioinformatics, Switzerland

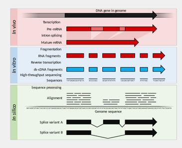
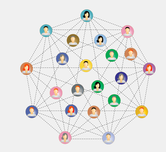
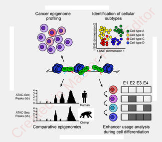
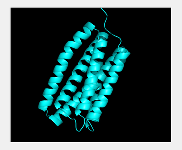
ETH zürich · MLCB

Problem and contribution



- Input: $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathcal{X}$ and $\mathbf{x}'_1, \dots, \mathbf{x}'_{n'} \in \mathcal{X}'$.
- Output: correspondence $\mathbf{P} \in \mathbb{R}^{n \times n'}$ and embeddings $\mathbf{z}_i, \mathbf{z}'_{i'} \in \mathbb{R}^d$.

Examples

	RNA-seq	AlphaFold	Network 1
Domain 1			
	ATAC-seq	Experimental	Network 2
Domain 2			

Contribution

- We formulate the above problem as an optimization problem, jointly optimizing the correspondence matrix and the embeddings.
- We propose an alternating optimization strategy, by alternatingly solving a multidimensional scaling (MDS) and a Wasserstein Procrustes problem.
- Our algorithm, named Joint MDS, can effectively benefit from the optimization techniques for solving each individual sub-problem.
- We demonstrate the effectiveness of joint MDS in several machine learning applications.

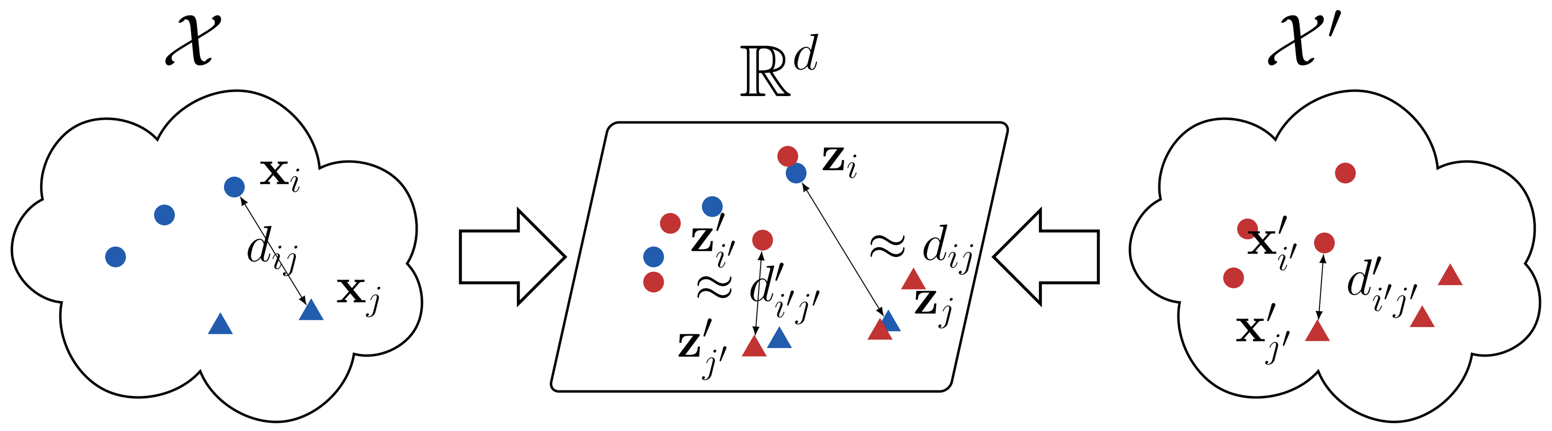
Existing methods

Method	Input	Output	Limitations
Multidimensional scaling	$[d_{ij}] \in \mathbb{R}^{n \times n}$	$\mathbf{z}_i, \mathbf{z}_j \in \mathbb{R}^d$	only 1 dataset
Discrete OT	$\mathcal{X} = \mathcal{X}'$	$\mathbf{P} \in \mathbb{R}^{n \times n'}$	same metric space
Gromov-Wasserstein	$[d_{ij}], [d'_{i'j'}]$	$\mathbf{P} \in \mathbb{R}^{n \times n'}$	only correspondence

Joint MDS: overall algorithm

- **Initialization:** Solve $\mathbf{Z} = \mathrm{MDS}(\mathbf{D}, \mathbf{W})$ and $\mathbf{Z}' = \mathrm{MDS}(\mathbf{D}', \mathbf{W}')$ using SMACOF.
- **Alignment:** Solve Wasserstein Procrustes $\mathbf{P}, \mathbf{O} = \mathrm{WP}(\mathbf{Z}, \mathbf{Z}')$.
- Update $\tilde{\mathbf{Z}}$, $\tilde{\mathbf{D}}$, and $\tilde{\mathbf{W}}$ using $\mathbf{Z}, \mathbf{Z}', \mathbf{P}$, and $\mathbf{O}$.
- **Embedding:** Solve $\mathbf{Z}, \mathbf{Z}' = \mathrm{MDS}(\tilde{\mathbf{D}}, \tilde{\mathbf{W}})$ using SMACOF with $\tilde{Z}$ as intialization.
- Repeat step **Alignment** and **Embedding** until convergence.

Overview of the Joint MDS problem



- Input: intra-dataset pairwise dissimilarities $\mathbf{D} \in \mathbb{R}^{n \times n}$ and $\mathbf{D}' \in \mathbb{R}^{n' \times n'}$.
- Output: correspondence $\mathbf{P} \in \mathbb{R}^{n \times n'}$ and embeddings $\mathbf{z}_i, \mathbf{z}'_{i'} \in \mathbb{R}^d$.

Optimization problem and solution

$$\min_{\substack{\mathbf{Z} \in \mathbb{R}^{n \times d}, \mathbf{Z}' \in \mathbb{R}^{n' \times d} \\ \mathbf{P} \in \Pi(\mathbf{a}, \mathbf{b}), \mathbf{O} \in \mathcal{O}_d}} \mathrm{stress}(\mathbf{Z}, \mathbf{D}, \mathbf{W}) + \mathrm{stress}(\mathbf{Z}', \mathbf{D}', \mathbf{W}') + 2\lambda \langle \mathbf{P}, d^2(\mathbf{Z}\mathbf{O}, \mathbf{Z}') \rangle_F.$$

- When $\mathbf{P}, \mathbf{O}$ fixed, it amounts to minimizing $\mathrm{stress}(\tilde{\mathbf{Z}}, \tilde{\mathbf{D}}, \tilde{\mathbf{W}})$ where

$$\tilde{\mathbf{Z}} := \begin{bmatrix} \mathbf{Z} \\ \mathbf{Z}' \end{bmatrix}, \quad \tilde{\mathbf{D}} := \begin{bmatrix} \mathbf{D} & 0 \\ 0 & \mathbf{D}' \end{bmatrix}, \quad \tilde{\mathbf{W}} := \begin{bmatrix} \mathbf{W} & \lambda \mathbf{P} \\ \lambda \mathbf{P}^\top & \mathbf{W}' \end{bmatrix}.$$

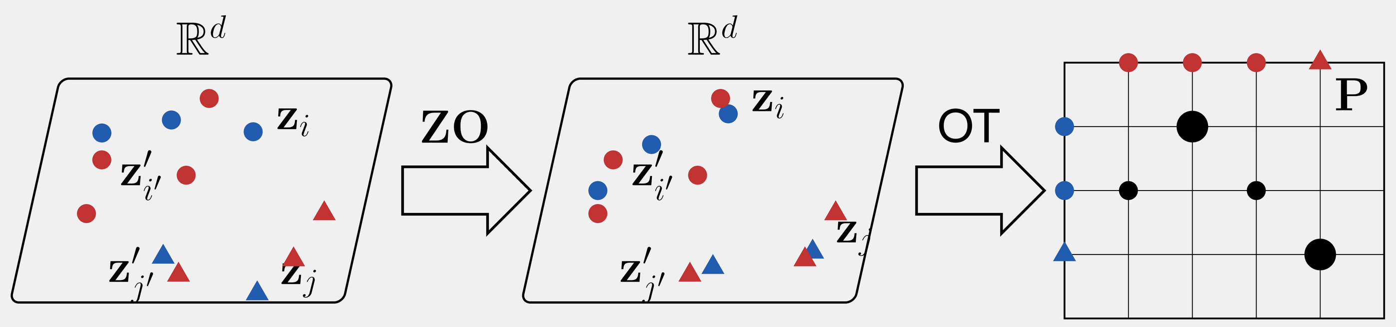
- When $\mathbf{Z}, \mathbf{Z}'$ fixed, one recovers the Wasserstein Procrustes problem.

Weighted MDS [2]

$$\mathrm{MDS}(\mathbf{D}, \mathbf{W}) := \min_{\mathbf{Z} \in \mathbb{R}^{n \times d}} \mathrm{stress}(\mathbf{Z}, \mathbf{D}, \mathbf{W}) := \sum_{i,j=1}^{n} w_{ij} (d_{ij} - d(\mathbf{z}_i, \mathbf{z}_j))^2,$$

- Input: pairwise distance $\mathbf{D}$ and weight matrix $\mathbf{W}$.
- Output: embeddings $\mathbf{z}_1, \dots, \mathbf{z}_n \in \mathbb{R}^d$.
- Algorithm: iterative majorization method SMACOF.

Wasserstein Procrustes [1]



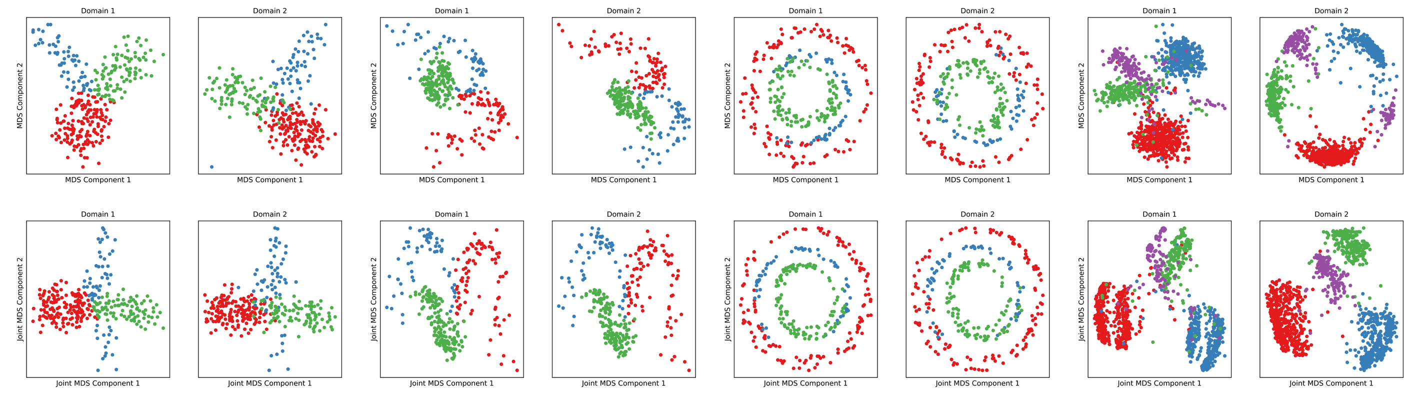
$$\mathrm{WP}(\mathbf{Z}, \mathbf{Z}') := \min_{\mathbf{P} \in \Pi(\mathbf{a}, \mathbf{b})} \min_{f \in \mathcal{F}} \langle \mathbf{P}, d^2(f(\mathbf{Z}), \mathbf{Z}') \rangle_F$$

- $\mathcal{F}$ is a pre-defined invariance class.
- In particular if $\mathcal{F} = \mathcal{O}_d$, one recovers the Wasserstein Procrustes problem.
- Input: $\mathbf{Z} \in \mathbb{R}^{n \times d}, \mathbf{Z}' \in \mathbb{R}^{n' \times d}$. Output: $\mathbf{P} \in \mathbb{R}^{n \times n'}$, $\mathbf{O} \in \mathcal{O}_d$.
- Algorithm: Sinkhorn-Knopp (for OT) + SVD (for orthogonal Procrustes).

Joint visualization of two datasets

MDS

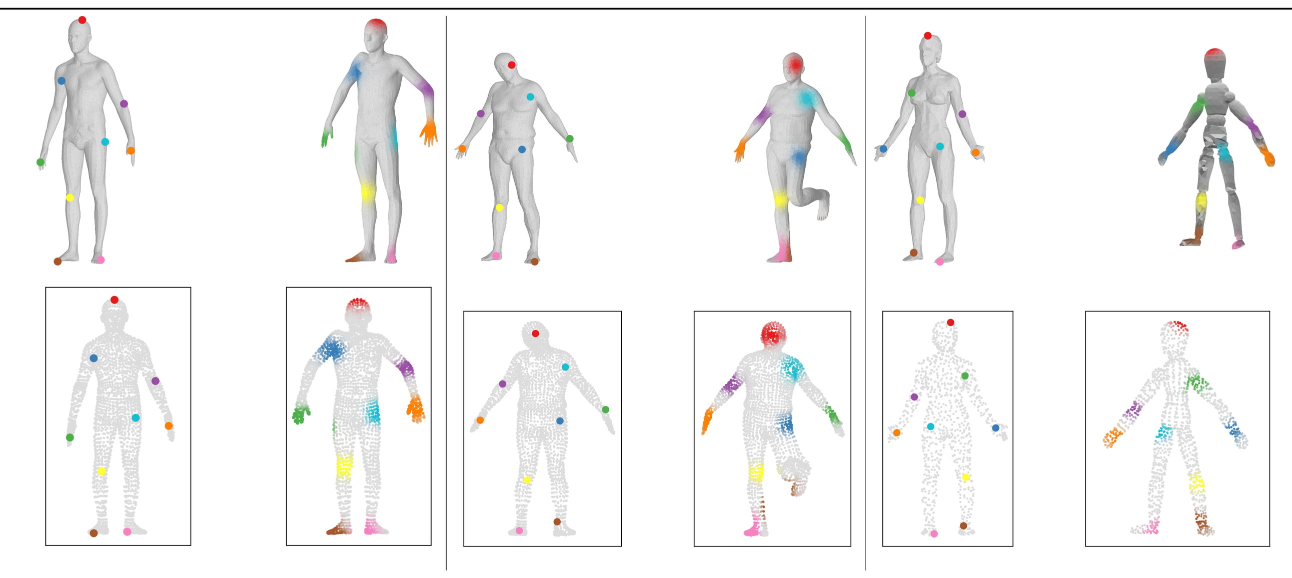
Joint MDS



Unsupervised heterogeneous domain adaption

Method	Bifurcation	Swiss roll	Circular frustum	SNAREseq	scGEM	MNIST-USPS
SCOT	93.7	97.7	**95.7**	**98.2**	57.6	26.7
EGW	95.7	**99.3**	94.7	93.8	62.7	43.1
Joint MDS (d=2)	96.0	**99.3**	94.0	85.5	64.4	15.0
Joint MDS (d=16)	**96.7**	**99.3**	94.7	94.7	**72.9**	**60.2**

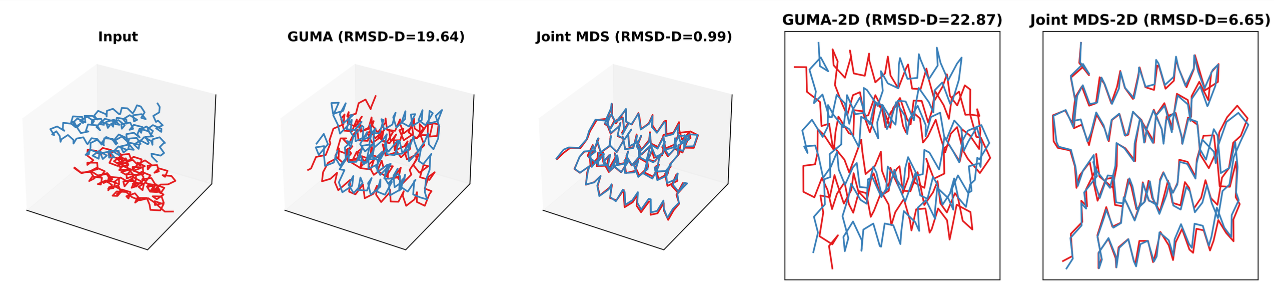
Human body pose alignment



Graph matching

Method	PPI 5%	PPI 15%	PPI 25%	MIMIC top 3	MIMIC top 5
MAGNA++	50.00	35.16	12.85	–	–
HubAlign	46.06	32.47	27.39	–	–
GWL	84.31	**74.35**	**67.42**	27.98	42.14
Joint MDS	**86.44±0.33**	72.31±0.62	55.3±0.78	**30.24±1.66**	**46.28±1.51**

Protein structure alignment



References

[1] David Alvarez-Melis, Stefanie Jegelka, and Tommi S Jaakkola. Towards optimal transport with global invariances. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2019.

[2] Warren S Torgerson. Multidimensional scaling of similarity. *Psychometrika*, 30(4):379–393, 1965.