

Consensus Dimension Reduction via Multi-View Learning

Tiffany Tang

University of Notre Dame

Applied and Computational Mathematics and Statistics (ACMS)

July 8, 2026

Joint work with Bingxue An (University of Notre Dame)



Motivation: Dimension Reduction

- In today's age of “**big data**,” modern datasets in genetics, neuroscience, finance, and beyond contain a large number of features (or dimensions).

Motivation: Dimension Reduction

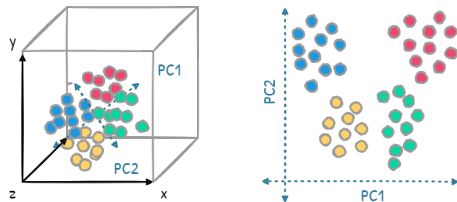
- In today's age of “**big data**,” modern datasets in genetics, neuroscience, finance, and beyond contain a large number of features (or dimensions).
- A key first step in any machine learning or data analysis pipeline is exploratory data analysis, where we want to **visualize** the data.

Motivation: Dimension Reduction

- In today's age of “**big data**,” modern datasets in genetics, neuroscience, finance, and beyond contain a large number of features (or dimensions).
- A key first step in any machine learning or data analysis pipeline is exploratory data analysis, where we want to **visualize** the data.
- **Dimension reduction:** aims to find a lower-dimensional representation of the data which preserves as much information as possible

Motivation: Dimension Reduction

- In today's age of “**big data**,” modern datasets in genetics, neuroscience, finance, and beyond contain a large number of features (or dimensions).
- A key first step in any machine learning or data analysis pipeline is exploratory data analysis, where we want to **visualize** the data.
- **Dimension reduction**: aims to find a lower-dimensional representation of the data which preserves as much information as possible



Motivation: Dimension Reduction

- In today's age of “**big data**,” modern datasets in genetics, neuroscience, finance, and beyond contain a large number of features (or dimensions).
- A key first step in any machine learning or data analysis pipeline is exploratory data analysis, where we want to **visualize** the data.
- **Dimension reduction:** aims to find a lower-dimensional representation of the data which preserves as much information as possible
- **Use cases:**
 - Exploratory Data Analysis/Visualization
 - Feature Engineering
 - Clustering
 - Unsupervised Discovery

Existing Dimension Reduction Methods

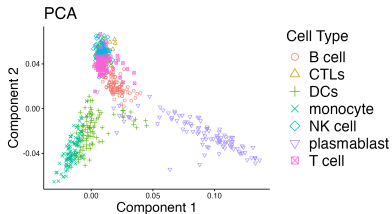
There are numerous existing dimension reduction methods
(10k-50k citations each):

- Principal components analysis (PCA)
[Pearson, 1901, Hotelling, 1933]
- Multidimensional scaling (MDS)
[Kruskal, 1964]
- Kernel principal components analysis (kPCA) [Schölkopf et al., 1998]
- Independent components analysis (ICA) [Hyvärinen and Oja, 2000]
- Isomap [Tenenbaum et al., 2000]
- Locally linear embedding (LLE)
[Roweis and Saul, 2000]
- t-distributed stochastic neighbor embedding (tSNE)
[Van der Maaten and Hinton, 2008]
- Uniform manifold approximation and projection (UMAP) [McInnes et al., 2018]
- ...

Which method is best?

Data: Single-cell RNA sequencing from HIV-infected individuals ($n = 700, p = 70$)

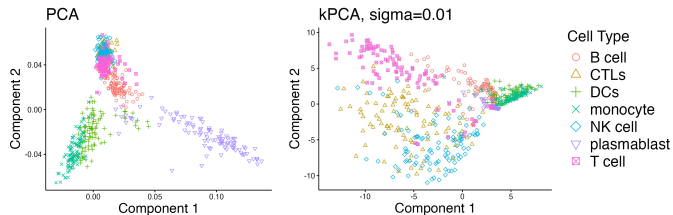
[Huang et al., 2022]



Which method is best?

Data: Single-cell RNA sequencing from HIV-infected individuals ($n = 700, p = 70$)

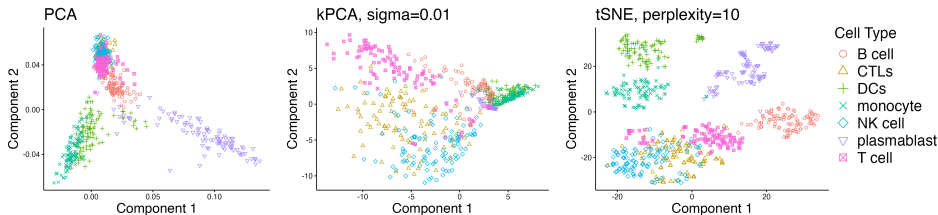
[Huang et al., 2022]



Which method is best?

Data: Single-cell RNA sequencing from HIV-infected individuals ($n = 700, p = 70$)

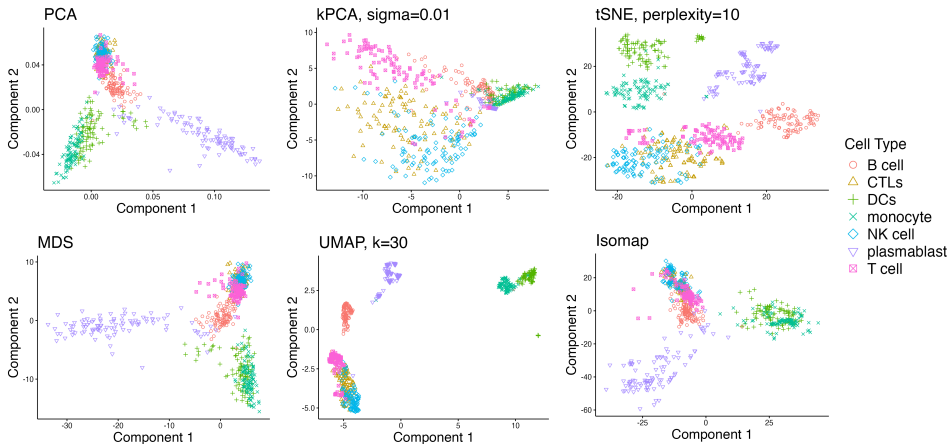
[Huang et al., 2022]



Which method is best?

Data: Single-cell RNA sequencing from HIV-infected individuals ($n = 700, p = 70$)

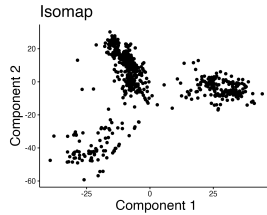
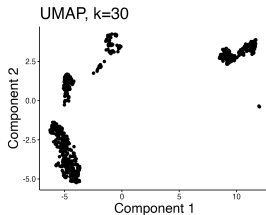
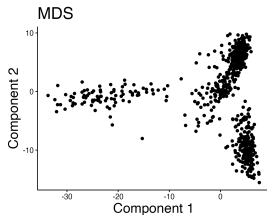
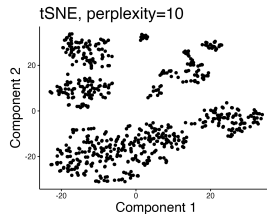
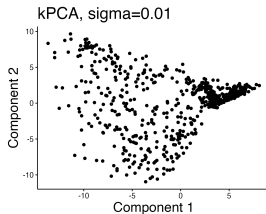
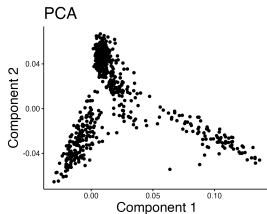
[Huang et al., 2022]



Which method is best?

Data: Single-cell RNA sequencing from HIV-infected individuals ($n = 700, p = 70$)

[Huang et al., 2022]



The Problem + Our Approach

- **Problem:** In unsupervised settings, validation of dimension reduction findings is challenging, if not impossible

The Problem + Our Approach

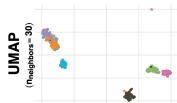
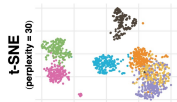
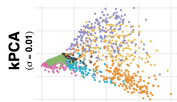
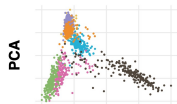
- **Problem:** In unsupervised settings, validation of dimension reduction findings is challenging, if not impossible
- **Our Approach:** Can we create a **consensus visualization** which captures the **shared structure** across many different dimension reduction visualizations?

The Problem + Our Approach

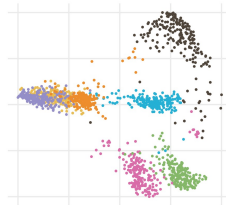
- **Problem:** In unsupervised settings, validation of dimension reduction findings is challenging, if not impossible
- **Our Approach:** Can we create a **consensus visualization** which captures the **shared structure** across many different dimension reduction visualizations?
 - **Stability Principle:** patterns that appear across multiple dimension reduction methods are less likely to be spurious noise compared to patterns found in only one dimension reduction method [Yu and Kumbier, 2020]

Overview: Consensus Dimension Reduction

*Candidate Dimension
Reduction Methods*

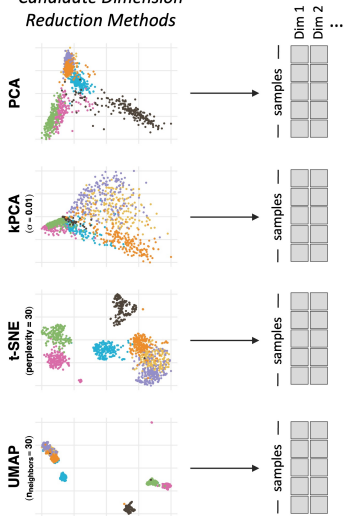


*Consensus
Dimension Reduction*

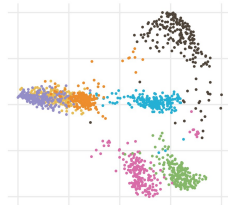


Overview: Consensus Dimension Reduction

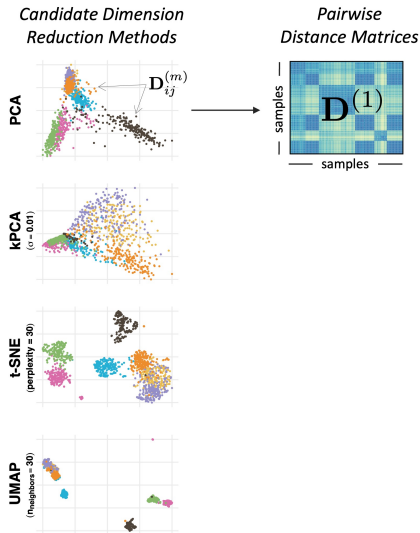
*Candidate Dimension
Reduction Methods*



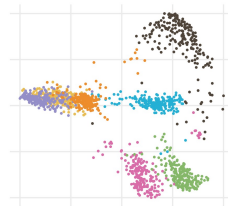
*Consensus
Dimension Reduction*



Overview: Consensus Dimension Reduction

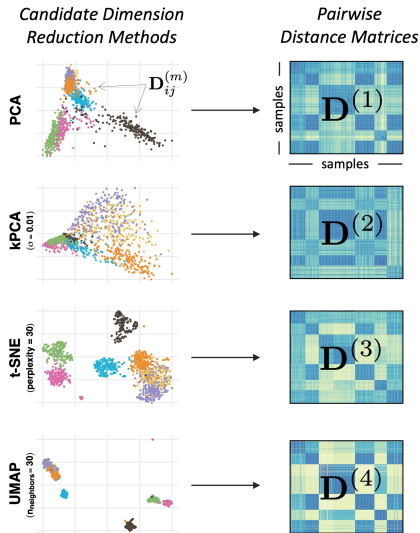


*Consensus
Dimension Reduction*

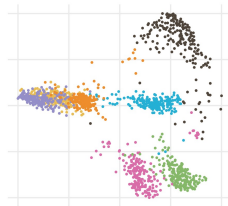


* [Viswanath and Madabhushi, 2012],
Meta-spec [Ma et al., 2023],
MCE [Tomo and Yoneoka, 2025]

Overview: Consensus Dimension Reduction

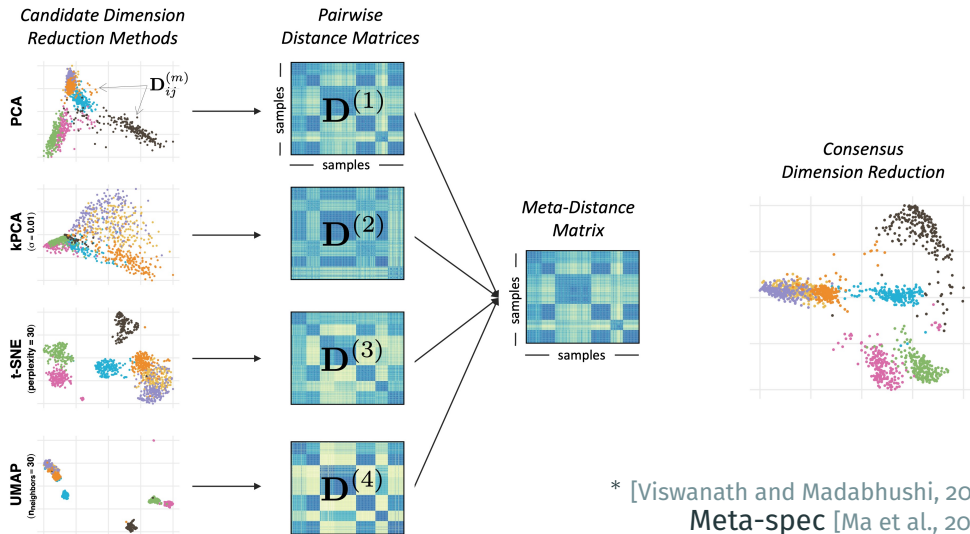


Consensus Dimension Reduction



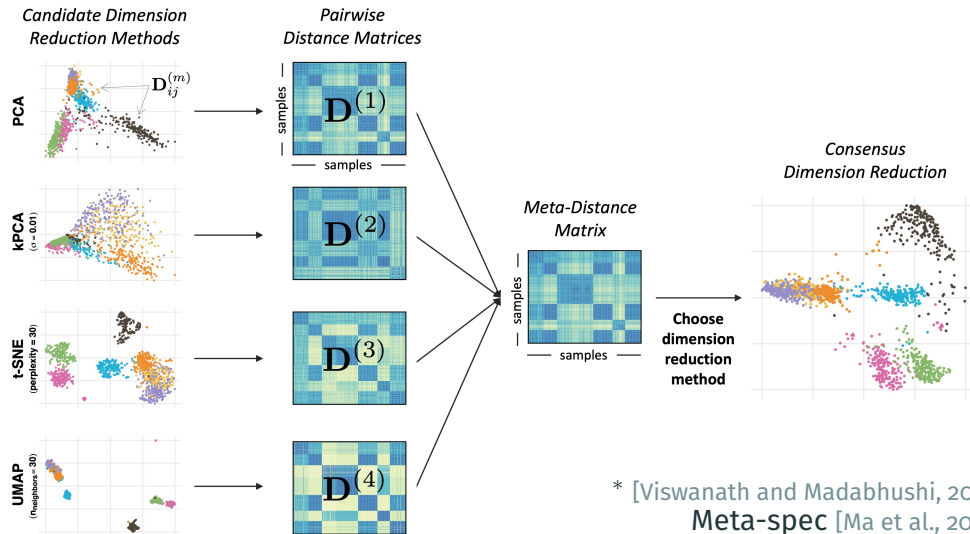
* [Viswanath and Madabhushi, 2012],
Meta-spec [Ma et al., 2023],
MCE [Tomo and Yoneoka, 2025]

Overview: Consensus Dimension Reduction



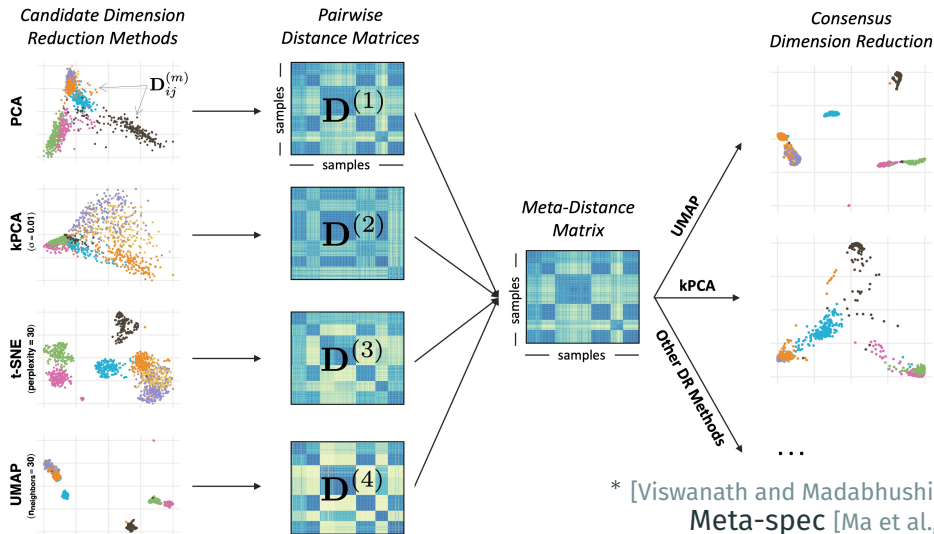
* [Viswanath and Madabhushi, 2012],
Meta-spec [Ma et al., 2023],
MCE [Tomo and Yoneoka, 2025]

Overview: Consensus Dimension Reduction



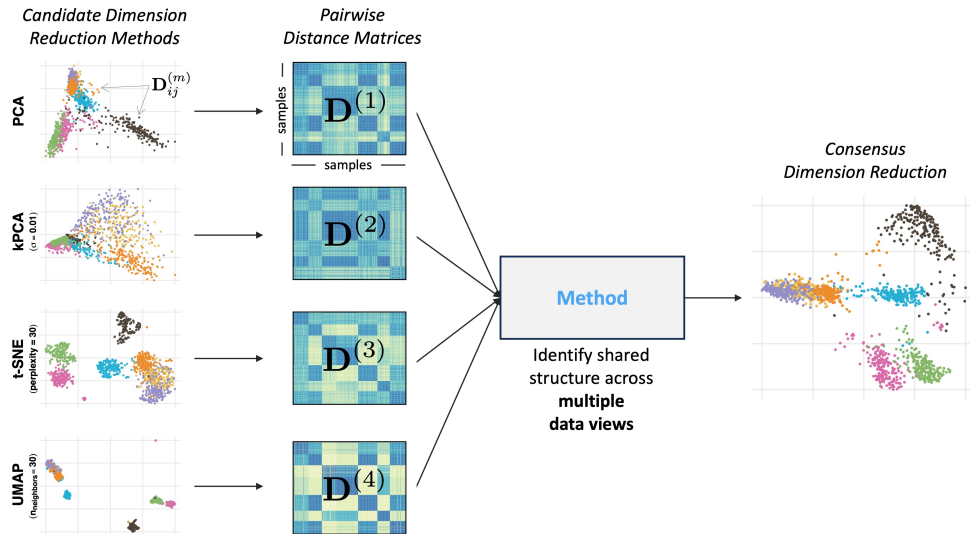
* [Viswanath and Madabhushi, 2012],
Meta-spec [Ma et al., 2023],
MCE [Tomo and Yoneoka, 2025]

Overview: Consensus Dimension Reduction

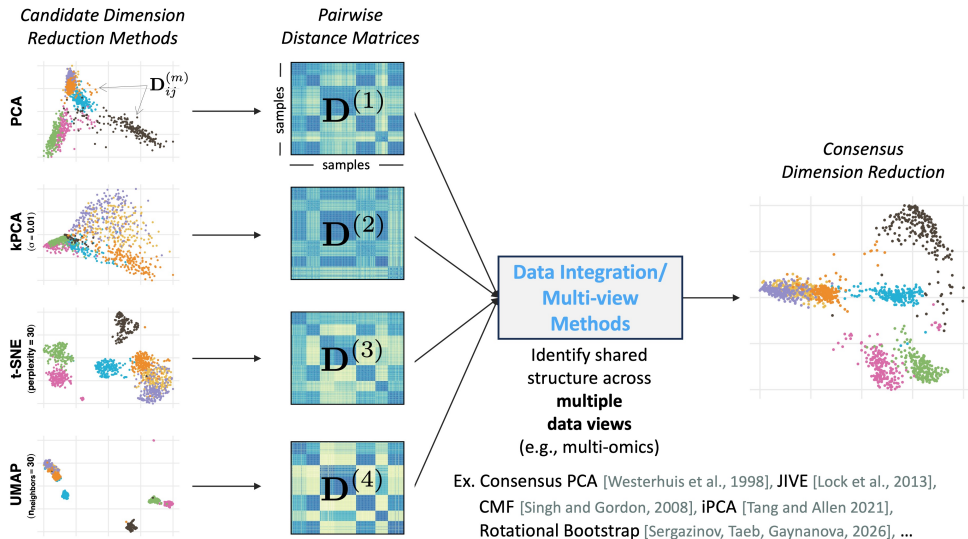


* [Viswanath and Madabhushi, 2012],
Meta-spec [Ma et al., 2023],
MCE [Tomo and Yoneoka, 2025]

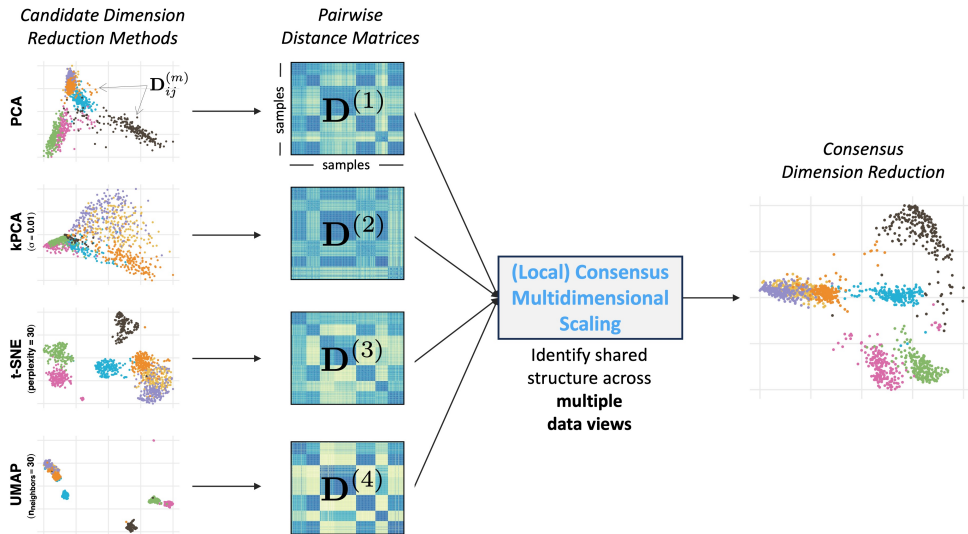
Overview: Consensus Dimension Reduction



Overview: Consensus Dimension Reduction



Overview: Consensus Dimension Reduction



Consensus Multidimensional Scaling (CoMDS)

Multidimensional Scaling (MDS) [Kruskal, 1964]

Given a distance matrix $\mathbf{D} \in \mathbb{R}^{n \times n}$, MDS finds a low-dimensional representation $\hat{\mathbf{Z}} \in \mathbb{R}^{n \times r}$ ($r \ll n$) such that

$$\hat{\mathbf{Z}} = \arg \min_{\mathbf{Z} \in \mathbb{R}^{n \times r}} \sum_{1 \leq i < j \leq n} \left(\underbrace{\mathbf{D}_{ij}}_{\text{original distance}} - \underbrace{d(\mathbf{Z}_i, \mathbf{Z}_j)}_{\text{distance in embedding space}} \right)^2$$

Multidimensional Scaling (MDS) [Kruskal, 1964]

Given a distance matrix $\mathbf{D} \in \mathbb{R}^{n \times n}$, MDS finds a low-dimensional representation $\hat{\mathbf{Z}} \in \mathbb{R}^{n \times r}$ ($r \ll n$) such that

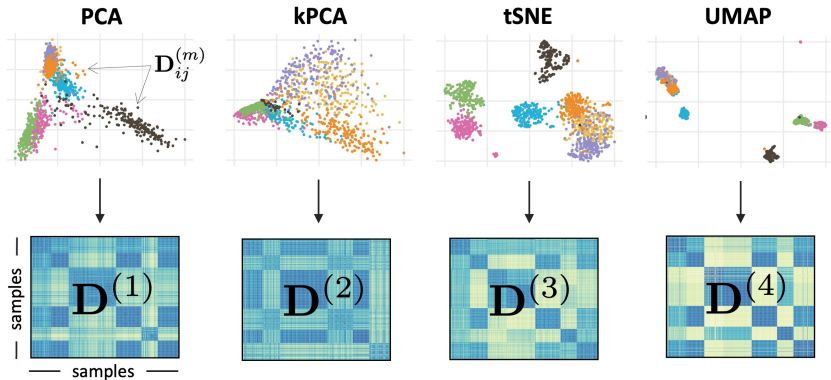
$$\hat{\mathbf{Z}} = \arg \min_{\mathbf{Z} \in \mathbb{R}^{n \times r}} \sum_{1 \leq i < j \leq n} \left(\underbrace{\mathbf{D}_{ij}}_{\text{original distance}} - \underbrace{d(\mathbf{Z}_i, \mathbf{Z}_j)}_{\text{distance in embedding space}} \right)^2$$

In other words, MDS finds a low-dimensional representation $\hat{\mathbf{Z}}$, which best preserves the original pairwise distances (measured in the original domain).

Consensus Multidimensional Scaling (CoMDS)

Given M different dimension reduction (DR) embeddings $\mathbf{Z}^{(1)}, \dots, \mathbf{Z}^{(M)}$,

1. Encode each DR embedding as a pairwise distance matrix $\rightarrow \mathbf{D}^{(1)}, \dots, \mathbf{D}^{(M)}$



Consensus Multidimensional Scaling (CoMDS)

Given M different dimension reduction (DR) embeddings $\mathbf{Z}^{(1)}, \dots, \mathbf{Z}^{(M)}$,

1. Encode each DR embedding as a pairwise distance matrix $\rightarrow \mathbf{D}^{(1)}, \dots, \mathbf{D}^{(M)}$
2. Find consensus low-dimension representation $\hat{\mathbf{Z}}^* \in \mathbb{R}^{n \times r}$ ($r \ll n$) such that

$$\hat{\mathbf{Z}}^*, \hat{\mathbf{W}}^{(1)}, \dots, \hat{\mathbf{W}}^{(M)} = \arg \min_{\substack{\mathbf{Z} \in \mathbb{R}^{n \times r} \\ \mathbf{W}^{(m)} \in \text{Diag}(r)}} \sum_{m=1}^M \sum_{i < j} \left\{ \underbrace{\mathbf{D}_{ij}^{(m)}}_{\text{distance in } m^{\text{th}} \text{ DR space}} - \underbrace{d(\mathbf{W}^{(m)} \mathbf{Z}_i, \mathbf{W}^{(m)} \mathbf{Z}_j)}_{\text{distance in consensus space}} \right\}^2,$$

Consensus Multidimensional Scaling (CoMDS)

Given M different dimension reduction (DR) embeddings $\mathbf{Z}^{(1)}, \dots, \mathbf{Z}^{(M)}$,

1. Encode each DR embedding as a pairwise distance matrix $\rightarrow \mathbf{D}^{(1)}, \dots, \mathbf{D}^{(M)}$
2. Find consensus low-dimension representation $\hat{\mathbf{Z}}^* \in \mathbb{R}^{n \times r}$ ($r \ll n$) such that

$$\hat{\mathbf{Z}}^*, \hat{\mathbf{W}}^{(1)}, \dots, \hat{\mathbf{W}}^{(M)} = \arg \min_{\substack{\mathbf{Z} \in \mathbb{R}^{n \times r} \\ \mathbf{W}^{(m)} \in \text{Diag}(r)}} \sum_{m=1}^M \sum_{i < j} \left\{ \underbrace{\mathbf{D}_{ij}^{(m)}}_{\text{distance in } m^{\text{th}} \text{ DR space}} - \underbrace{d(\mathbf{W}^{(m)} \mathbf{Z}_i, \mathbf{W}^{(m)} \mathbf{Z}_j)}_{\text{distance in consensus space}} \right\}^2,$$

where $\hat{\mathbf{W}}^{(1)}, \dots, \hat{\mathbf{W}}^{(M)}$ are $r \times r$ diagonal matrices to allow for different scalings among the input dimension reduction embeddings

Consensus Multidimensional Scaling (CoMDS)

Given M different dimension reduction (DR) embeddings $\mathbf{Z}^{(1)}, \dots, \mathbf{Z}^{(M)}$,

1. Encode each DR embedding as a pairwise distance matrix $\rightarrow \mathbf{D}^{(1)}, \dots, \mathbf{D}^{(M)}$
2. Find consensus low-dimension representation $\hat{\mathbf{Z}}^* \in \mathbb{R}^{n \times r}$ ($r \ll n$) such that

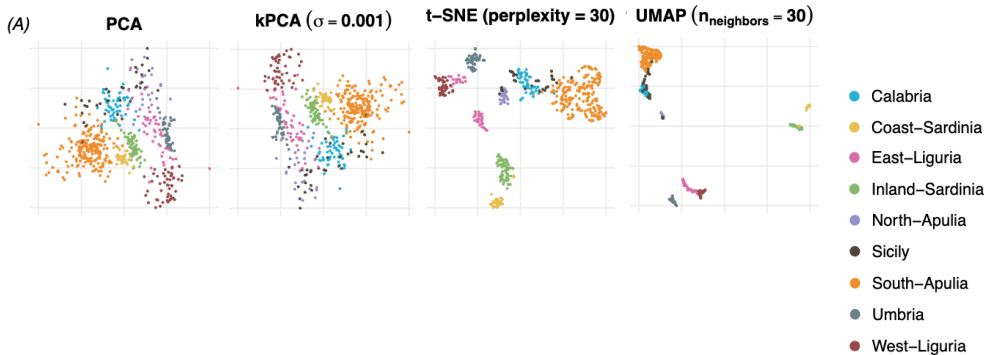
$$\hat{\mathbf{Z}}^*, \hat{\mathbf{W}}^{(1)}, \dots, \hat{\mathbf{W}}^{(M)} = \arg \min_{\substack{\mathbf{Z} \in \mathbb{R}^{n \times r} \\ \mathbf{W}^{(m)} \in \text{Diag}(r)}} \sum_{m=1}^M \sum_{i < j} \left\{ \underbrace{\mathbf{D}_{ij}^{(m)}}_{\substack{\text{distance in} \\ m^{\text{th}} \text{ DR space}}} - \underbrace{d(\mathbf{W}^{(m)} \mathbf{Z}_i, \mathbf{W}^{(m)} \mathbf{Z}_j)}_{\substack{\text{distance in} \\ \text{consensus space}}} \right\}^2,$$

where $\hat{\mathbf{W}}^{(1)}, \dots, \hat{\mathbf{W}}^{(M)}$ are $r \times r$ diagonal matrices to allow for different scalings among the input dimension reduction embeddings

- In other words, CoMDS finds a low-dimensional representation $\hat{\mathbf{Z}}^*$, which best preserves the pairwise distances (measured in the dimension-reduced space), averaged across all dimension reduction visualizations.

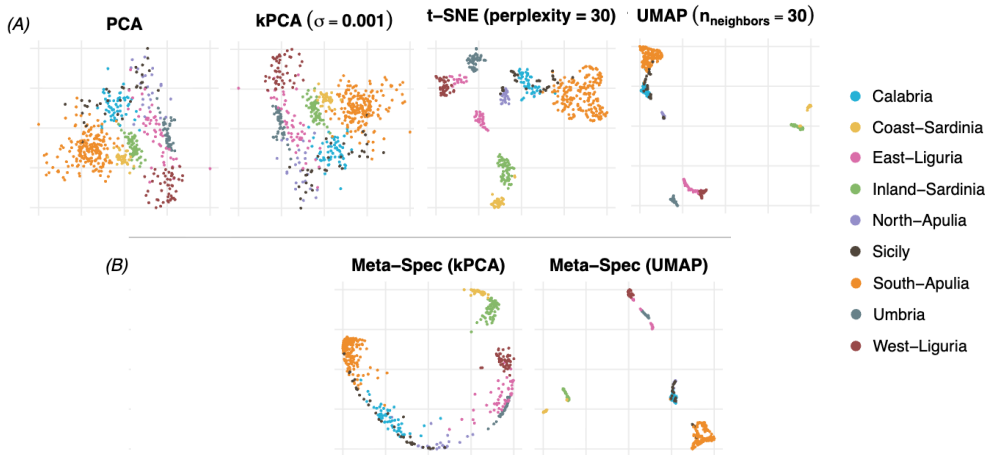
Case Study: Olive Oil Samples [Forina et al., 1983]

Data: Chemical measurements of **olive oils** from 9 regions in Italy ($n = 572, p = 8$)



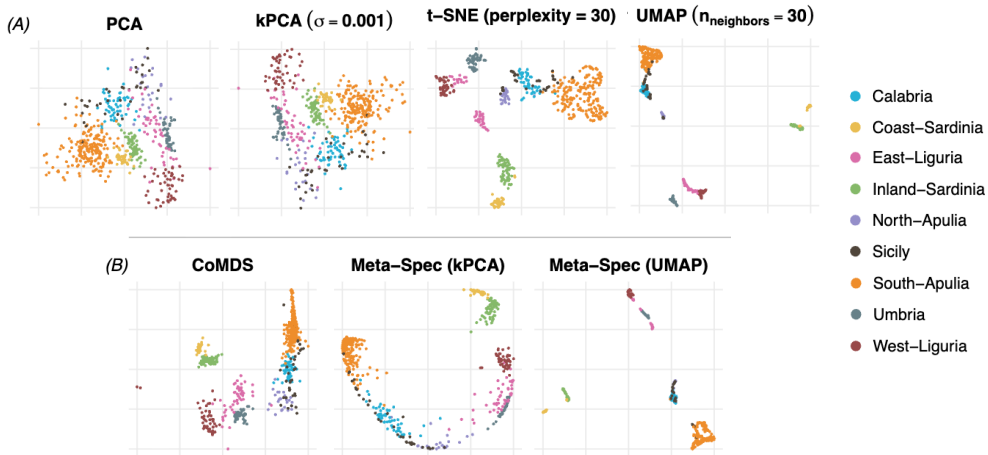
Case Study: Olive Oil Samples [Forina et al., 1983]

Data: Chemical measurements of **olive oils** from 9 regions in Italy ($n = 572, p = 8$)



Case Study: Olive Oil Samples [Forina et al., 1983]

Data: Chemical measurements of **olive oils** from 9 regions in Italy ($n = 572, p = 8$)



Local Consensus Multidimensional Scaling (LoCoMDS)

Local Consensus Multidimensional Scaling (LoCoMDS)

LoCoMDS provides a **localized** alternative to CoMDS, which is useful in the presence of outliers and “local” dimension reduction methods (e.g., t-SNE, UMAP)

Local Consensus Multidimensional Scaling (LoCoMDS)

LoCoMDS provides a **localized** alternative to CoMDS, which is useful in the presence of outliers and “local” dimension reduction methods (e.g., t-SNE, UMAP)

Idea: Focus on pairs of samples that are close to each other and de-emphasize pairs of samples that are far away from each other [Chen and Buja, 2009]

Local Consensus Multidimensional Scaling (LoCoMDS)

LoCoMDS provides a **localized** alternative to CoMDS, which is useful in the presence of outliers and “local” dimension reduction methods (e.g., t-SNE, UMAP)

Idea: Focus on pairs of samples that are close to each other and de-emphasize pairs of samples that are far away from each other [Chen and Buja, 2009]

$$\arg \min_{\substack{\mathbf{Z} \in \mathbb{R}^{n \times r} \\ \mathbf{W}^{(m)} \in \text{Diag}(r)}} \underbrace{\sum_{m=1}^M \sum_{(i,j) \in \mathcal{N}_{\pi}^{(m)}} \left\{ \mathbf{D}_{ij}^{(m)} - d(\mathbf{W}^{(m)} \mathbf{z}_i, \mathbf{W}^{(m)} \mathbf{z}_j) \right\}^2}_{\text{Local Stress (close points)}} - t \underbrace{\sum_{m=1}^M \sum_{(i,j) \notin \mathcal{N}_{\pi}^{(m)}} d(\mathbf{W}^{(m)} \mathbf{z}_i, \mathbf{W}^{(m)} \mathbf{z}_j)}_{\text{Repulsion (distant points)}}$$

Local Consensus Multidimensional Scaling (LoCoMDS)

LoCoMDS provides a **localized** alternative to CoMDS, which is useful in the presence of outliers and “local” dimension reduction methods (e.g., t-SNE, UMAP)

Idea: Focus on pairs of samples that are close to each other and de-emphasize pairs of samples that are far away from each other [Chen and Buja, 2009]

$$\arg \min_{\substack{\mathbf{Z} \in \mathbb{R}^{n \times r} \\ \mathbf{W}^{(m)} \in \text{Diag}(r)}} \underbrace{\sum_{m=1}^M \sum_{(i,j) \in \mathcal{N}_{\pi}^{(m)}} \left\{ \mathbf{D}_{ij}^{(m)} - d(\mathbf{W}^{(m)} \mathbf{z}_i, \mathbf{W}^{(m)} \mathbf{z}_j) \right\}^2}_{\text{Local Stress (close points)}} - t \underbrace{\sum_{m=1}^M \sum_{(i,j) \notin \mathcal{N}_{\pi}^{(m)}} d(\mathbf{W}^{(m)} \mathbf{z}_i, \mathbf{W}^{(m)} \mathbf{z}_j)}_{\text{Repulsion (distant points)}}$$

where

- $t > 0$ indicates the strength of penalty applied to large distances
- $\mathcal{N}_{\pi}^{(m)}$ is the set of “neighboring” sample pairs (i, j) such that the distance between samples i and j is less than the π^{th} percentile of all pairwise distances

Hyperparameter Tuning for LoCoMDS

Following [Chen and Buja \(2009\)](#):

Hyperparameter Tuning for LoCoMDS

Following Chen and Buja (2009):

- For each choice of π (neighborhood size) and t (penalty strength),

Hyperparameter Tuning for LoCoMDS

Following Chen and Buja (2009):

- For each choice of π (neighborhood size) and t (penalty strength),
 - (i) Compute proportion of overlap between k -nearest neighbors in the original (high-dimensional) space and in the embedded (low-dimensional) space

Hyperparameter Tuning for LoCoMDS

Following Chen and Buja (2009):

- For each choice of π (neighborhood size) and t (penalty strength),
 - (i) Compute proportion of overlap between k -nearest neighbors in the original (high-dimensional) space and in the embedded (low-dimensional) space
 - Also sometimes called *Local-Continuity Meta-Criterion* (LCMC)

Hyperparameter Tuning for LoCoMDS

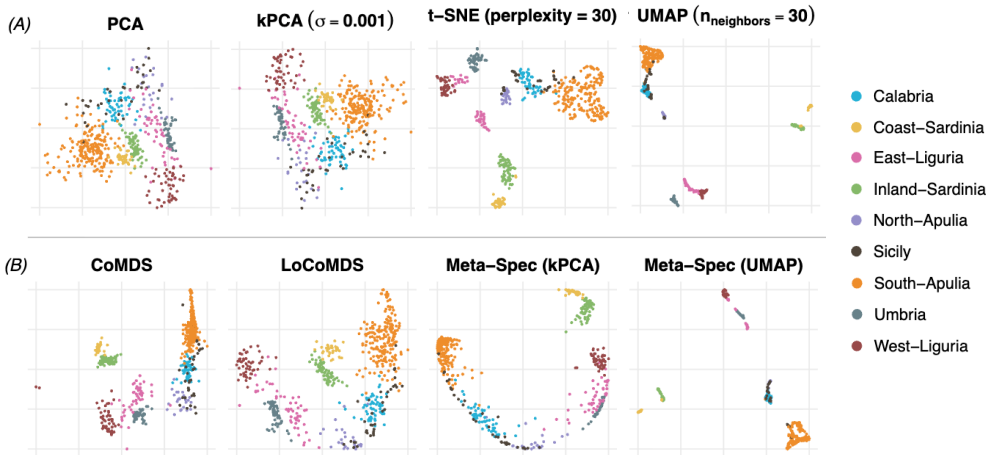
Following Chen and Buja (2009):

- For each choice of π (neighborhood size) and t (penalty strength),
 - (i) Compute proportion of overlap between k -nearest neighbors in the original (high-dimensional) space and in the embedded (low-dimensional) space
 - Also sometimes called *Local-Continuity Meta-Criterion* (LCMC)
 - (ii) Choose π and t which maximizes the LCMC (i.e., preserves the neighborhood structure)

Empirical Results

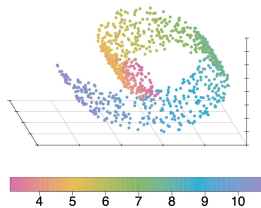
Case Study: Olive Oil Samples [Forina et al., 1983]

Data: Chemical measurements of **olive oils** from 9 regions in Italy ($n = 572, p = 8$)



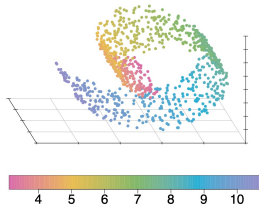
Case Study: Swiss Roll

(A)



Case Study: Swiss Roll

(A)



(B)

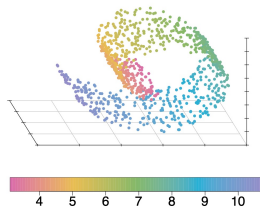
PCA

kPCA
($\sigma = 0.01$)



Case Study: Swiss Roll

(A)

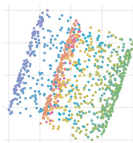


(B)

PCA



kPCA
($\sigma = 0.01$)



HLLE



t-SNE
(perplexity = 30)

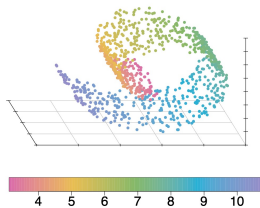


UMAP
($n_{\text{neighbors}} = 30$)



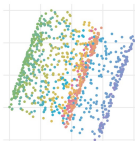
Case Study: Swiss Roll

(A)



(B)

PCA



kPCA
($\sigma = 0.01$)



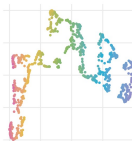
HLLE



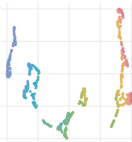
t-SNE
(perplexity = 30)



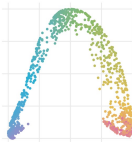
UMAP
($n_{\text{neighbors}} = 30$)



Meta-Spec (UMAP)



Meta-Spec (kPCA)

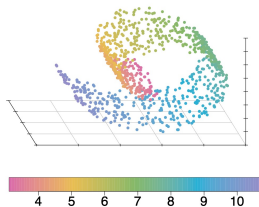


Multi-SNE

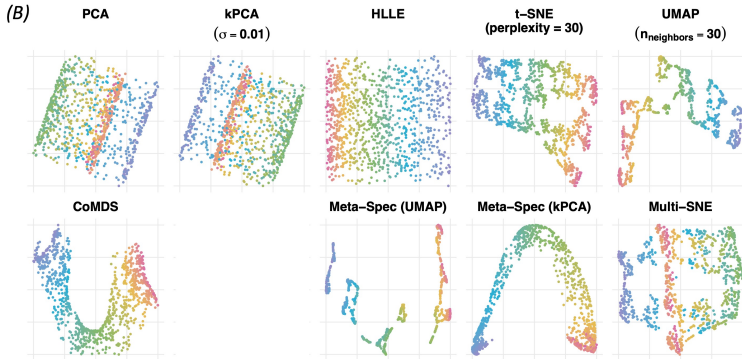


Case Study: Swiss Roll

(A)

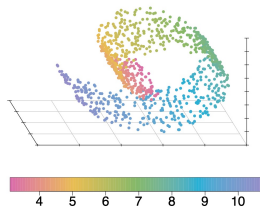


(B)

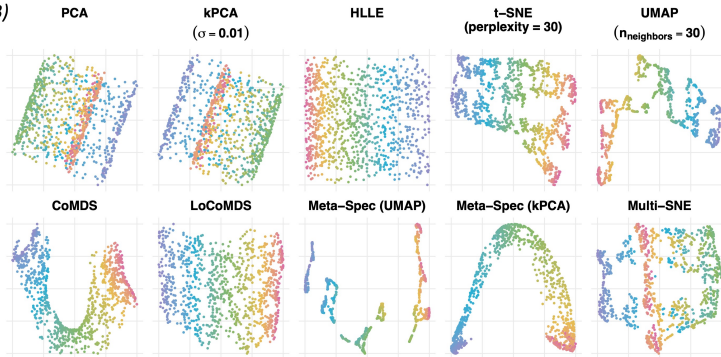


Case Study: Swiss Roll

(A)



(B)



Summary of Empirical Evaluations

Across 9 datasets spanning a variety of domains (e.g., agriculture, biology, and astrophysics), we show that CoMDS and LoCoMDS:

Summary of Empirical Evaluations

Across 9 datasets spanning a variety of domains (e.g., agriculture, biology, and astrophysics), we show that CoMDS and LoCoMDS:

- Strike an effective balance between preserving both **global** (i.e., far-away) and **local** (i.e., nearby) **structure** in the data

Summary of Empirical Evaluations

Across 9 datasets spanning a variety of domains (e.g., agriculture, biology, and astrophysics), we show that CoMDS and LoCoMDS:

- Strike an effective balance between preserving both **global** (i.e., far-away) and **local** (i.e., nearby) **structure** in the data
- Maintain **high predictive accuracy** and good separation of known relevant clusters, **without being over-confident** in spurious clusters

Summary of Empirical Evaluations

Across 9 datasets spanning a variety of domains (e.g., agriculture, biology, and astrophysics), we show that CoMDS and LoCoMDS:

- Strike an effective balance between preserving both **global** (i.e., far-away) and **local** (i.e., nearby) **structure** in the data
- Maintain **high predictive accuracy** and good separation of known relevant clusters, **without being over-confident** in spurious clusters
- Yield stable visualizations that are relatively **robust** to the choice of input dimension reduction methods

Summary of Empirical Evaluations

Across 9 datasets spanning a variety of domains (e.g., agriculture, biology, and astrophysics), we show that CoMDS and LoCoMDS:

- Strike an effective balance between preserving both **global** (i.e., far-away) and **local** (i.e., nearby) **structure** in the data
- Maintain **high predictive accuracy** and good separation of known relevant clusters, **without being over-confident** in spurious clusters
- Yield stable visualizations that are relatively **robust** to the choice of input dimension reduction methods
- Can reveal and mitigate **hyperparameter instability** in existing dimension reduction methods

Summary of Empirical Evaluations

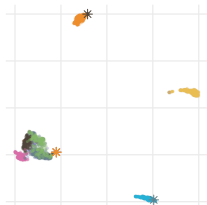
Across 9 datasets spanning a variety of domains (e.g., agriculture, biology, and astrophysics), we show that CoMDS and LoCoMDS:

- Strike an effective balance between preserving both **global** (i.e., far-away) and **local** (i.e., nearby) **structure** in the data
- Maintain **high predictive accuracy** and good separation of known relevant clusters, **without being over-confident** in spurious clusters
- Yield stable visualizations that are relatively **robust** to the choice of input dimension reduction methods
- **Can reveal and mitigate hyperparameter instability in existing dimension reduction methods**

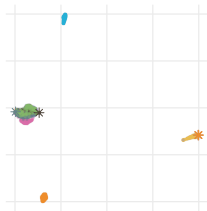
Reconciling Differences Across UMAP Hyperparameters

Data: scRNA-seq data ($n = 1600, p = 50$) [Duò et al., 2018]

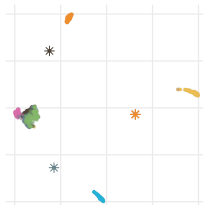
(A) UMAP ($n_{\text{neighbors}} = 10$)



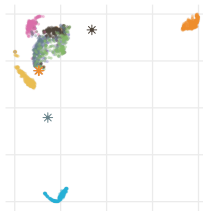
UMAP ($n_{\text{neighbors}} = 100$)



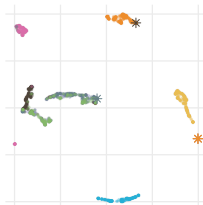
(B) CoMDS



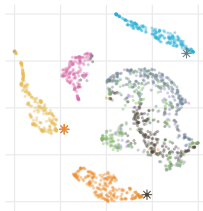
LoCoMDS



Meta-Spec (UMAP)



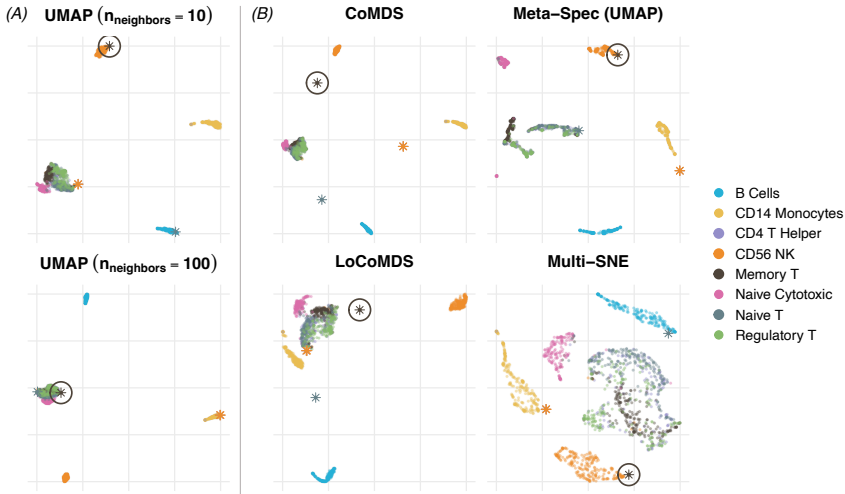
Multi-SNE



- B Cells
- CD14 Monocytes
- CD4 T Helper
- CD56 NK
- Memory T
- Naive Cytotoxic
- Naive T
- Regulatory T

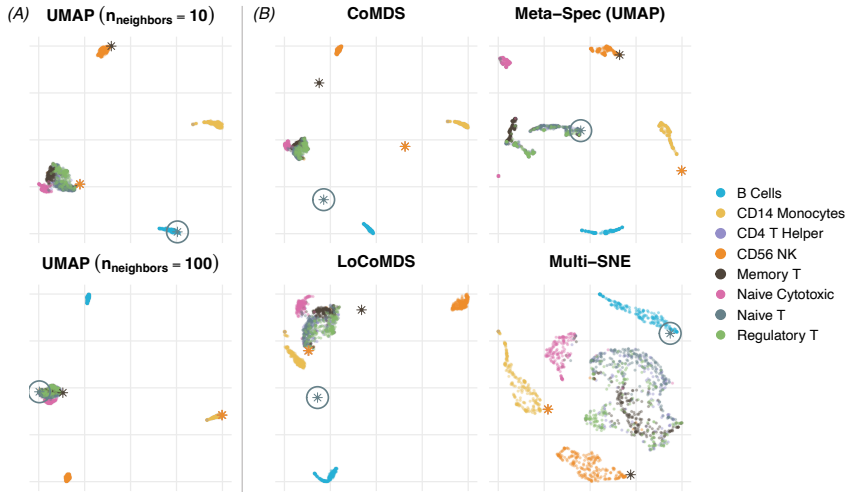
Reconciling Differences Across UMAP Hyperparameters

Data: scRNA-seq data ($n = 1600, p = 50$) [Duò et al., 2018]



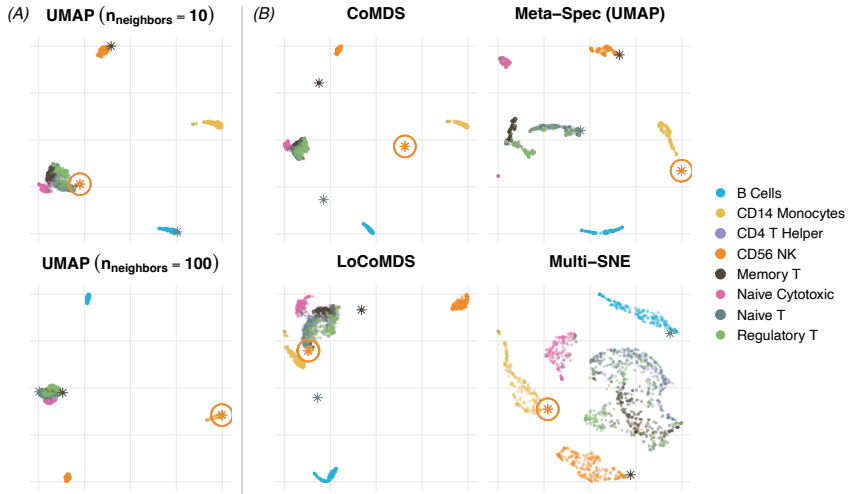
Reconciling Differences Across UMAP Hyperparameters

Data: scRNA-seq data ($n = 1600, p = 50$) [Duò et al., 2018]



Reconciling Differences Across UMAP Hyperparameters

Data: scRNA-seq data ($n = 1600, p = 50$) [Duò et al., 2018]



Discussion

- We develop a novel framework to generate a **consensus visualization** that captures the **shared structure** across different dimension reduction methods
 - Can be used to aggregate across different methods and/or hyperparameters

Discussion

- We develop a novel framework to generate a **consensus visualization** that captures the **shared structure** across different dimension reduction methods
 - Can be used to aggregate across different methods and/or hyperparameters
- Driving idea behind CoMDS/LoCoMDS is the **stability principle**
 - Patterns shared across multiple DR methods are less likely to be spurious noise compared to patterns that are idiosyncratic to a single DR method

Discussion

- We develop a novel framework to generate a **consensus visualization** that captures the **shared structure** across different dimension reduction methods
 - Can be used to aggregate across different methods and/or hyperparameters
- Driving idea behind CoMDS/LoCoMDS is the **stability principle**
 - Patterns shared across multiple DR methods are less likely to be spurious noise compared to patterns that are idiosyncratic to a single DR method
- Most useful in unsupervised contexts where validation is challenging and avoiding false discoveries is crucial

Discussion

- We develop a novel framework to generate a **consensus visualization** that captures the **shared structure** across different dimension reduction methods
 - Can be used to aggregate across different methods and/or hyperparameters
- Driving idea behind CoMDS/LoCoMDS is the **stability principle**
 - Patterns shared across multiple DR methods are less likely to be spurious noise compared to patterns that are idiosyncratic to a single DR method
- Most useful in unsupervised contexts where validation is challenging and avoiding false discoveries is crucial

Link: <https://arxiv.org/abs/2512.15802>

R Software: coMDS (<https://github.com/tiffanymtang/coMDS>)

Email: ttang4@nd.edu

Website: tiffanymtang.github.io/

References i

-  Chen, L. and Buja, A. (2009).
Local multidimensional scaling for nonlinear dimension reduction, graph drawing, and proximity analysis.
Journal of the American Statistical Association, 104(485):209–219.
-  Duò, A., Robinson, M. D., and Soneson, C. (2018).
A systematic performance evaluation of clustering methods for single-cell RNA-seq data.
F1000Res, 7:1141.
-  Forina, M., Armanino, C., Lanteri, S., and Tiscornia, E. (1983).
Classification of olive oils from their fatty acid composition.
Food research and data analysis, 1983:189–214.
-  Hotelling, H. (1933).
Analysis of a complex of statistical variables into principal components.
Journal of educational psychology, 24(6):417.

References ii

-  Huang, H., Wang, Y., Rudin, C., and Browne, E. P. (2022).
Towards a comprehensive evaluation of dimension reduction methods for transcriptomic data visualization.
Communications Biology, 5:719.
-  Hyvärinen, A. and Oja, E. (2000).
Independent component analysis: algorithms and applications.
Neural networks, 13(4-5):411-430.
-  Kruskal, J. B. (1964).
Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis.
Psychometrika, 29(1):1-27.
-  Ma, R., Sun, E. D., and Zou, J. (2023).
A spectral method for assessing and combining multiple data visualizations.
Nature Communications, 14:780.

References iii

-  McInnes, L., Healy, J., and Melville, J. (2018).
Umap: Uniform manifold approximation and projection for dimension reduction.
arXiv preprint arXiv:1802.03426.
-  Pearson, K. (1901).
Liii. on lines and planes of closest fit to systems of points in space.
The London, Edinburgh, and Dublin philosophical magazine and journal of science, 2(11):559–572.
-  Roweis, S. T. and Saul, L. K. (2000).
Nonlinear dimensionality reduction by locally linear embedding.
science, 290(5500):2323–2326.
-  Schölkopf, B., Smola, A., and Müller, K.-R. (1998).
Nonlinear component analysis as a kernel eigenvalue problem.
Neural computation, 10(5):1299–1319.

References iv

-  Tenenbaum, J. B., Silva, V. d., and Langford, J. C. (2000).
A global geometric framework for nonlinear dimensionality reduction.
science, 290(5500):2319–2323.
-  Tomo, Y. and Yoneoka, D. (2025).
Median consensus embedding for dimensionality reduction.
arXiv preprint arXiv:2503.08103.
-  Van der Maaten, L. and Hinton, G. (2008).
Visualizing data using t-sne.
Journal of machine learning research, 9(11).
-  Viswanath, S. and Madabhushi, A. (2012).
Consensus embedding: theory, algorithms and application to segmentation and classification of biomedical data.
BMC Bioinformatics, 13(1):26.

References v



Yu, B. and Kumbier, K. (2020).

Veridical data science.

Proceedings of the National Academy of Sciences, 117(8):3920–3929.