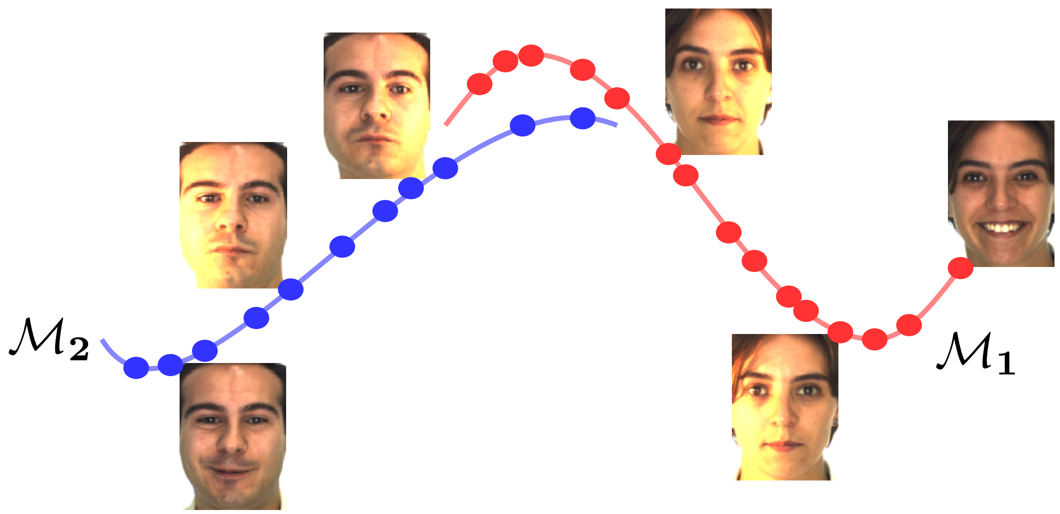
Learning All by Selecting A Few

Ehsan Elhamifar & Rene Vidal

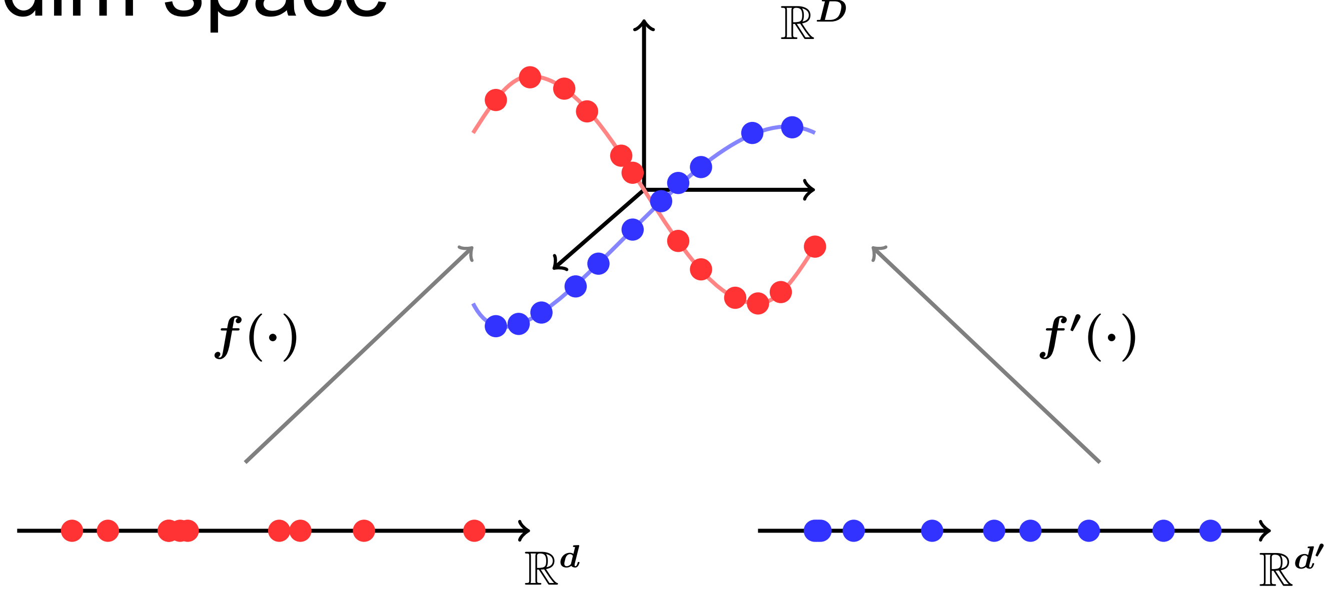
Center for Imaging Science
Johns Hopkins University

Intrinsic low-dimensionality

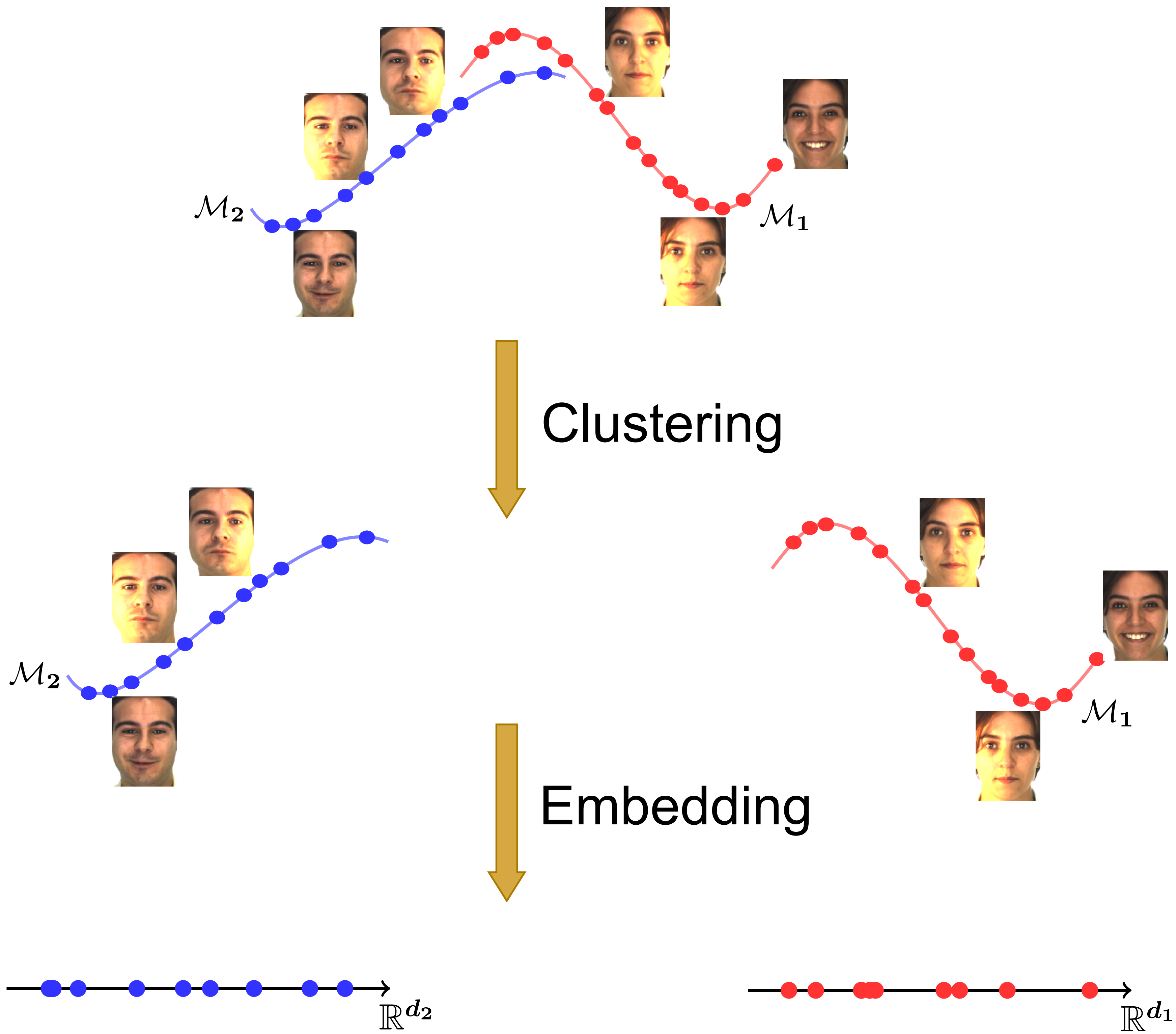
- Data concentrate around low-dimensional structures



- Mapping from low-dim to high-dim space
  - linear / nonlinear
  - one / multiple

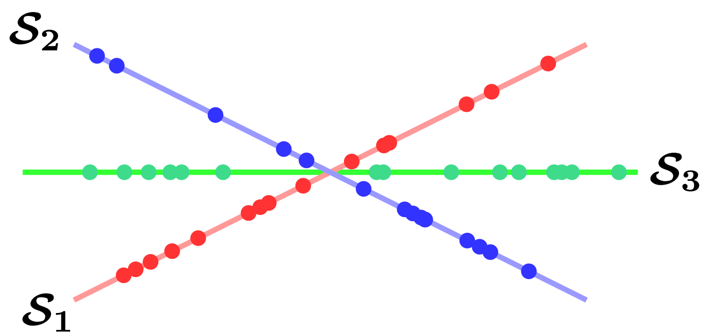


Two fundamental tasks

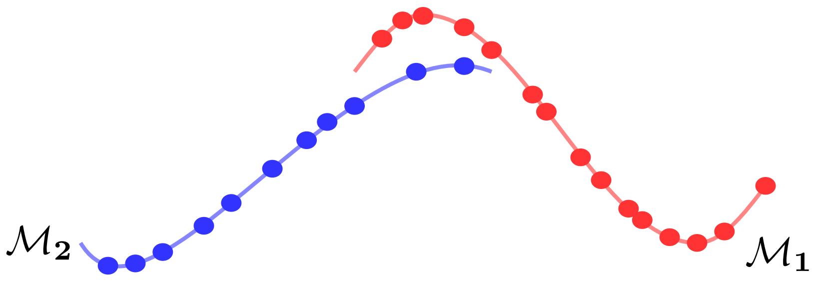


This talk

- Clustering and embedding on
  - multiple subspaces



  - multiple nonlinear manifolds



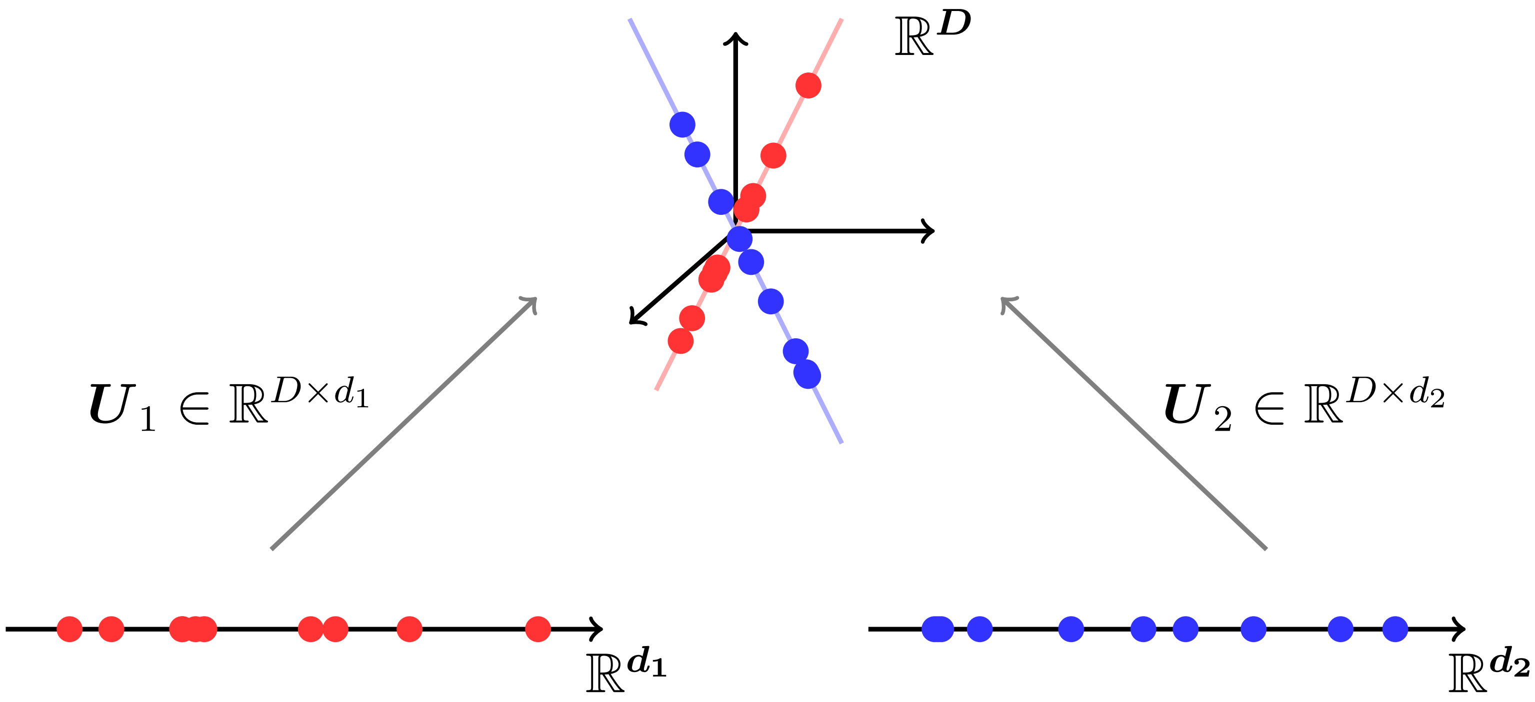
- Techniques from sparse representation theory
  - well suited for high-dim data (blessing of dimensionality, Donoho`00)

Sparse Subspace Clustering

Ehsan Elhamifar & Rene Vidal

Center for Imaging Science
Johns Hopkins University

Subspace clustering

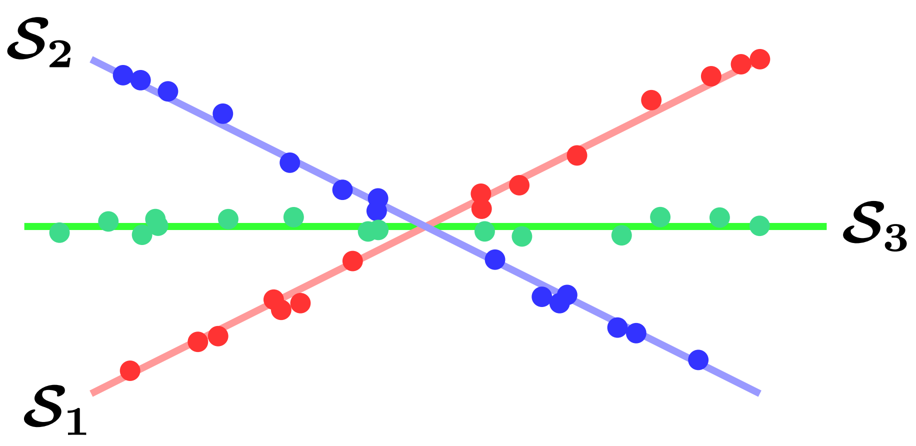


- Linear mappings: flat manifolds
  - faces, digits, motions, text corpus, gene expression
- Tasks:
  - separate data into subspaces
  - find low-dimensional representations

Subspace clustering

- Prior work:
  - algebraic: Ksubspaces (Tseng`00), GPCA (Vidal-Ma`03), median Kflats (Zhang`09)
  - statistical: RANSAC (Fischler`81), MPPCA (Tipping`99), ALC (Rao-Ma`09)
  - spectral clustering: LSA (Yan`06), SCC (Chen-Lerman`09), LRR (Liu`09), SLBF (Zhang-Lerman`10)

- Challenges:
  - intersecting subspaces
  - noise, outliers, missing entries
  - computational complexity
  - knowledge of dimension/number of subspaces



Our approach

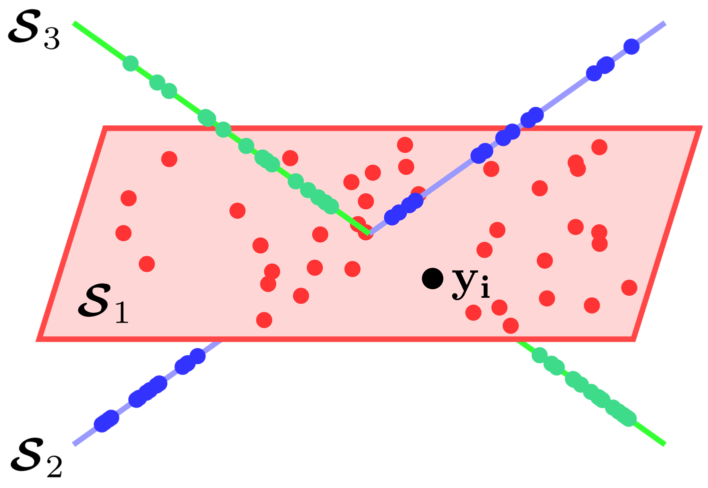
- Global sparse optimization
  - deal with intersection
  - deal with noise, outlying / missing entries
  - do not require dimension / number of subspaces
- Theoretical guarantees
- Achieves/outperforms state-of-the-art results in
  - segmentation of rigid-body motions
  - clustering of face images
  - temporal segmentation of videos

E. Elhamifar and R. Vidal, CVPR09, ICASSP10, TPAMI.

Sparse subspace clustering: idea

- Self-expressiveness property
  - $\boldsymbol{y}_i = \boldsymbol{Y}\boldsymbol{c}_i$ ⟶ many solutions
  - $\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{Y}_1 & \boldsymbol{Y}_2 & \cdots & \boldsymbol{Y}_n \end{bmatrix} \boldsymbol{\Gamma}$

    low column-rank



- In $\mathcal{S}_i$ of dim d_i, each point can be reconstructed by d_i other points
  - sparse representation comes from same subspace

$$\min \|\boldsymbol{c}_i\|_0 \quad \mathrm{s.t.} \quad \boldsymbol{y}_i = \boldsymbol{Y}\boldsymbol{c}_i, \;\; c_{ii} = 0$$

⇓

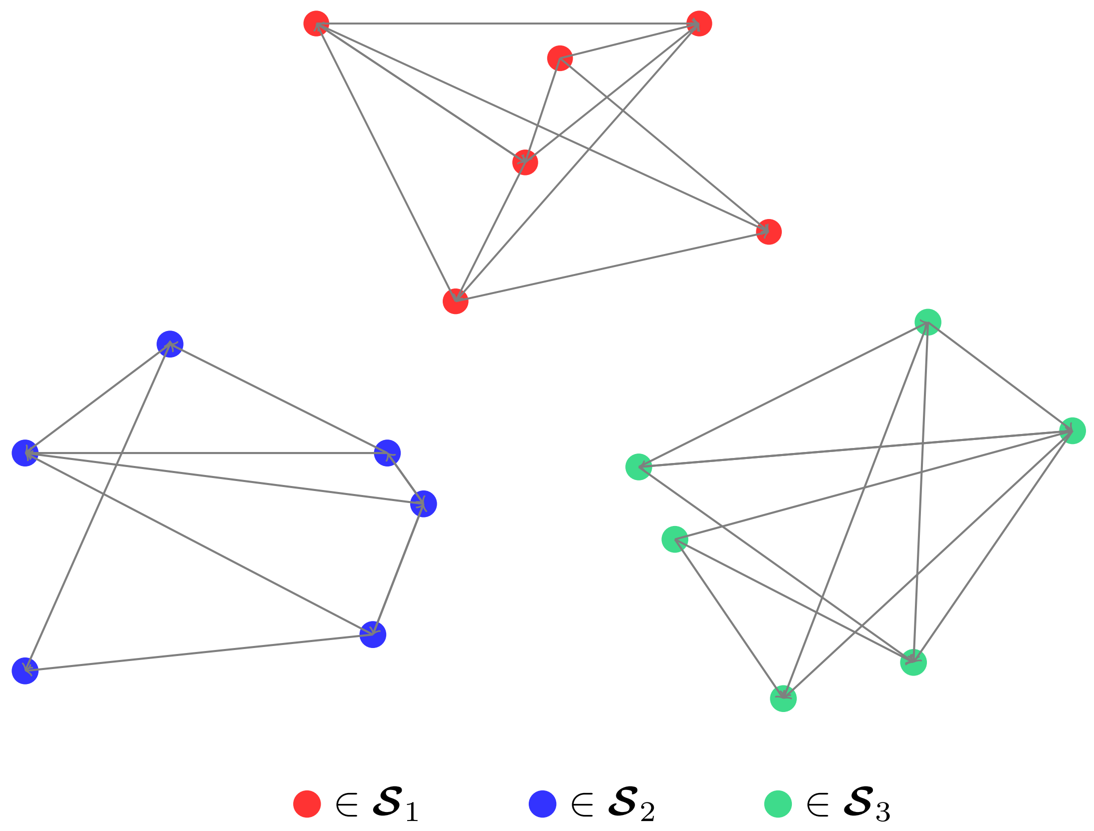
$$\min \|\boldsymbol{c}_i\|_1 \quad \mathrm{s.t.} \quad \boldsymbol{y}_i = \boldsymbol{Y}\boldsymbol{c}_i, \;\; c_{ii} = 0$$

Sparse subspace clustering

- SSC algorithm
  - 1: solve the sparse optimization

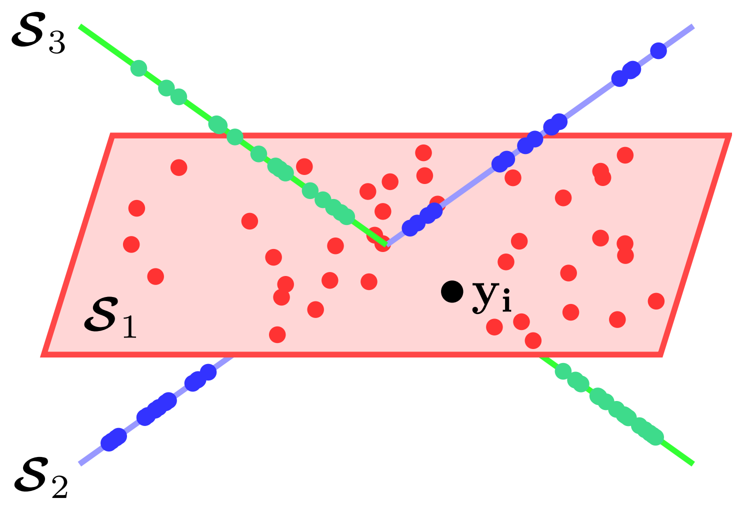
$$\min \|\boldsymbol{c}_i\|_1 \quad \text{s. t.} \quad \boldsymbol{y}_i = \boldsymbol{Y}\boldsymbol{c}_i,\; c_{ii} = 0 \qquad \boldsymbol{c}_i = \begin{bmatrix} c_{i1} \\ c_{i2} \\ \vdots \\ c_{iN} \end{bmatrix}$$

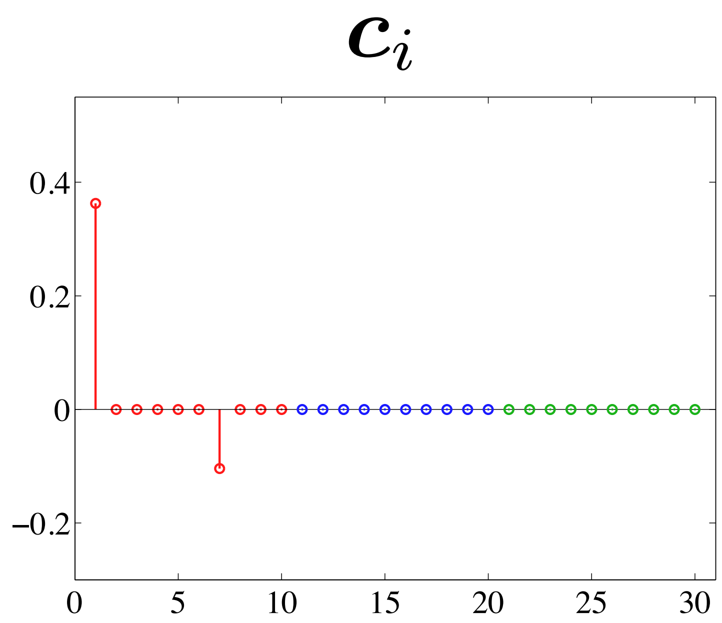
  - 2: infer clustering from similarity graph
    - connect points using sparse weights
    - symmetrize the weights $w_{ij} = c_{ij} + c_{ji}$
    - apply spectral clustering



SSC theory

- When does the algorithm succeed?
  - sparse representation from the correct subspace
  - subspace-sparse representation

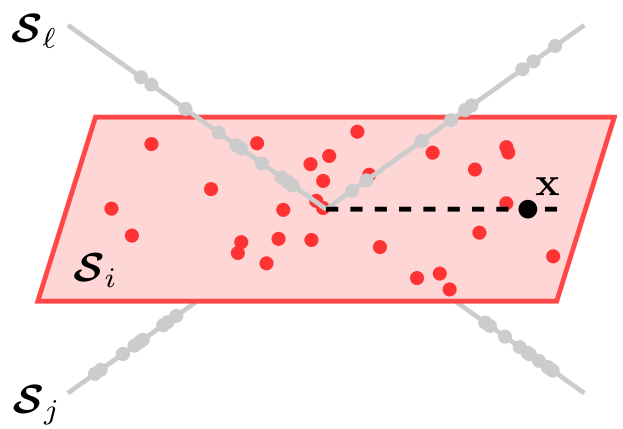




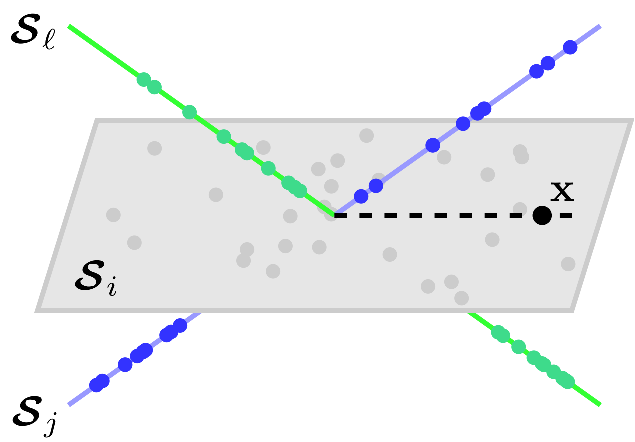
SSC theory

- Take any nonzero $\boldsymbol{x}$ in the intersection of $\mathcal{S}_i$ and $\oplus_{j \neq i} \mathcal{S}_j$

$$\boldsymbol{a}_i = \operatorname{argmin} \|\boldsymbol{a}\|_1 \quad \text{s.t.} \quad \boldsymbol{x} = \boldsymbol{Y}_i \, \boldsymbol{a}$$



$$\boldsymbol{a}_{-i} = \operatorname{argmin} \|\boldsymbol{a}\|_1 \quad \text{s.t.} \quad \boldsymbol{x} = \boldsymbol{Y}_{-i} \, \boldsymbol{a}$$

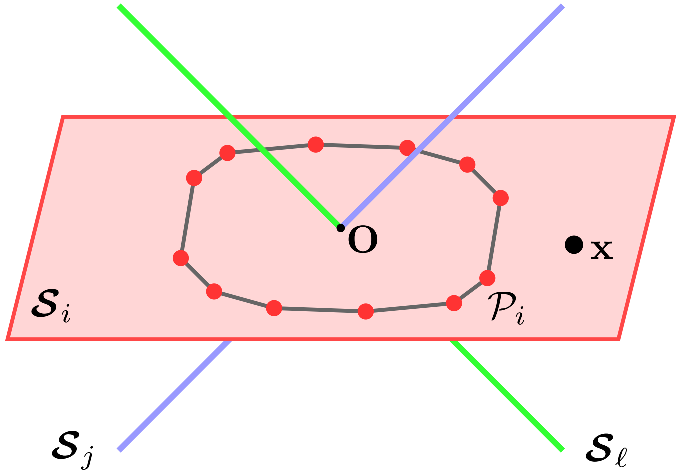


- **Theorem:** SSC recovers a subspace-sparse representation for any $\boldsymbol{y}$ in $\mathcal{S}_i$ iff

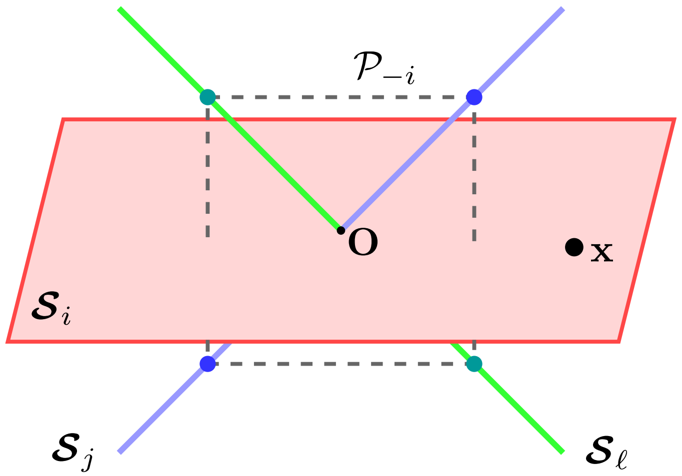
$$\forall \boldsymbol{x} \in \mathcal{S}_i \cap (\oplus_{j \neq i} \mathcal{S}_j), \boldsymbol{x} \neq \boldsymbol{0} \implies \|\boldsymbol{a}_i\|_1 < \|\boldsymbol{a}_{-i}\|_1$$

Geometric interpretation

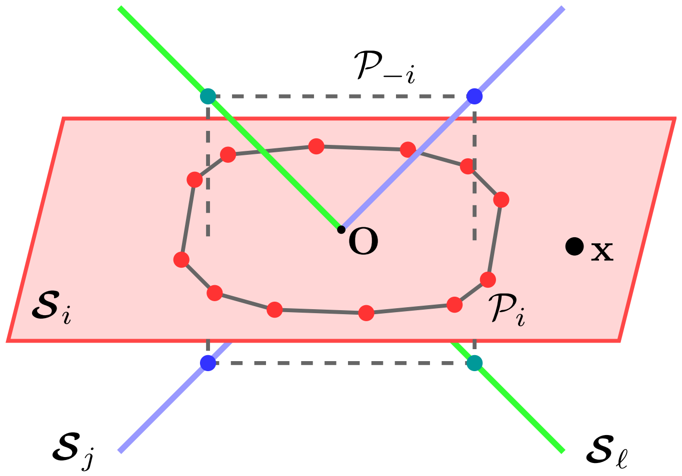
- $\min \|\boldsymbol{a}\|_1 \quad \text{s.t.} \quad \boldsymbol{x} = \boldsymbol{Y}_i \boldsymbol{a} \;\longrightarrow\; \boldsymbol{a}_i$



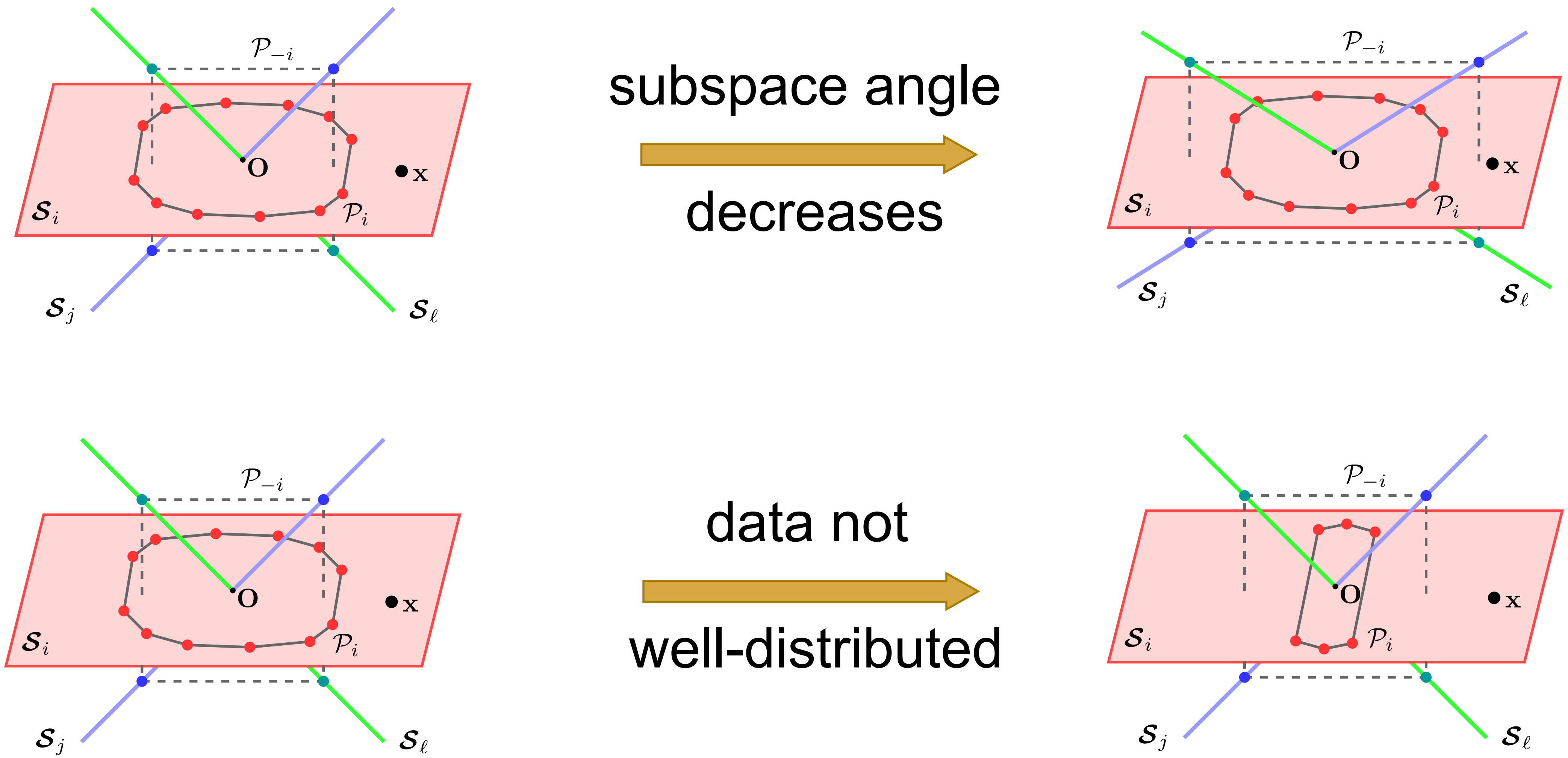
- $\min \|\boldsymbol{a}\|_1 \quad \text{s.t.} \quad \boldsymbol{x} = \boldsymbol{Y}_{-i} \boldsymbol{a} \;\longrightarrow\; \boldsymbol{a}_{-i}$



- $\|\boldsymbol{a}_i\|_1 < \|\boldsymbol{a}_{-i}\|_1$



Geometric interpretation

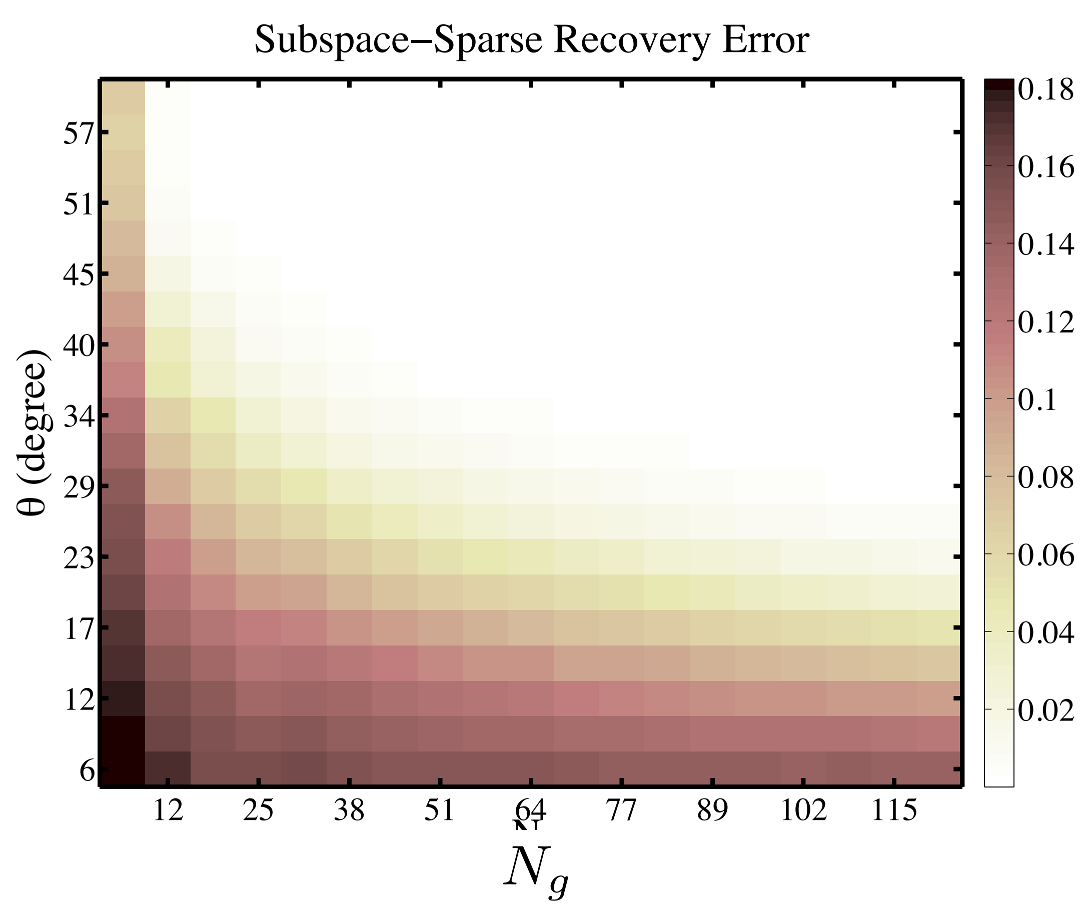


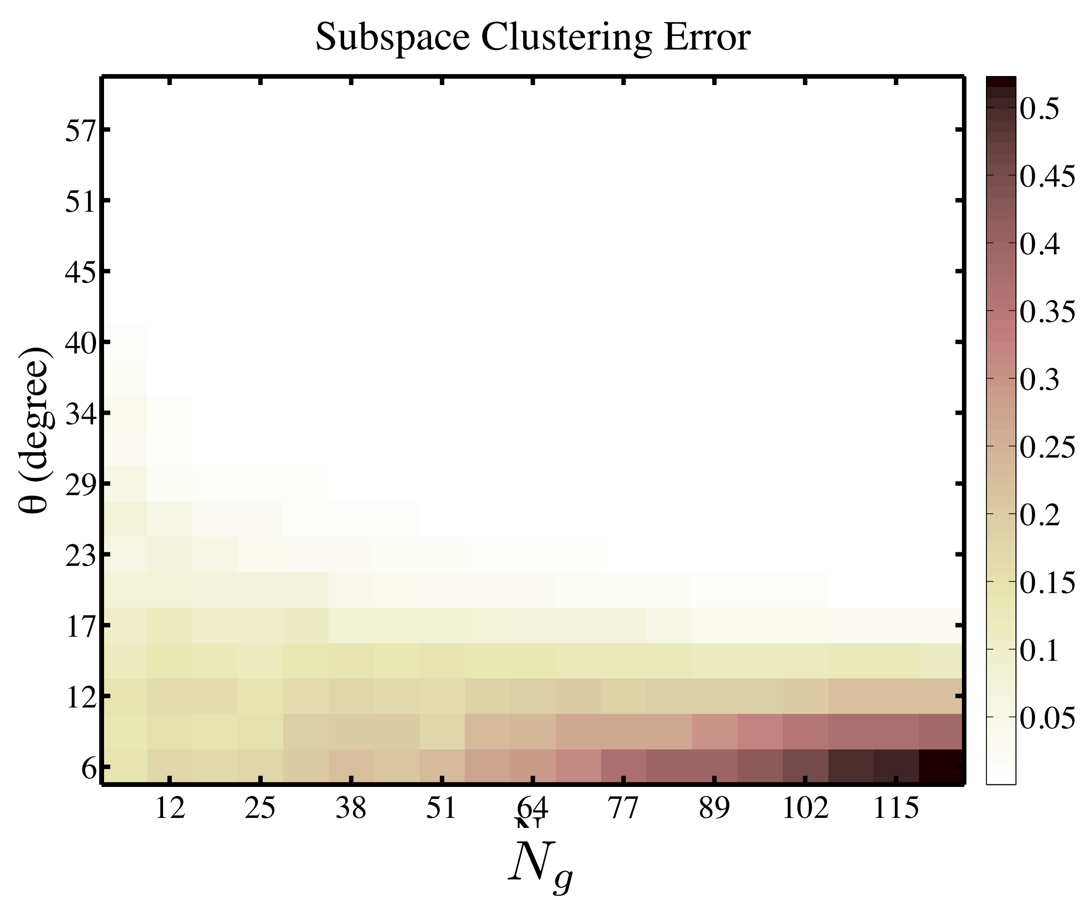
- **Theorem:** ℓ_1 optimization is successful if

$$\max_{\tilde{\boldsymbol{Y}}_i \in \mathbb{W}_i} \sigma_{d_i}(\tilde{\boldsymbol{Y}}_i) > \sqrt{d_i} \max_{j \neq i} \cos(\theta_{ij})$$

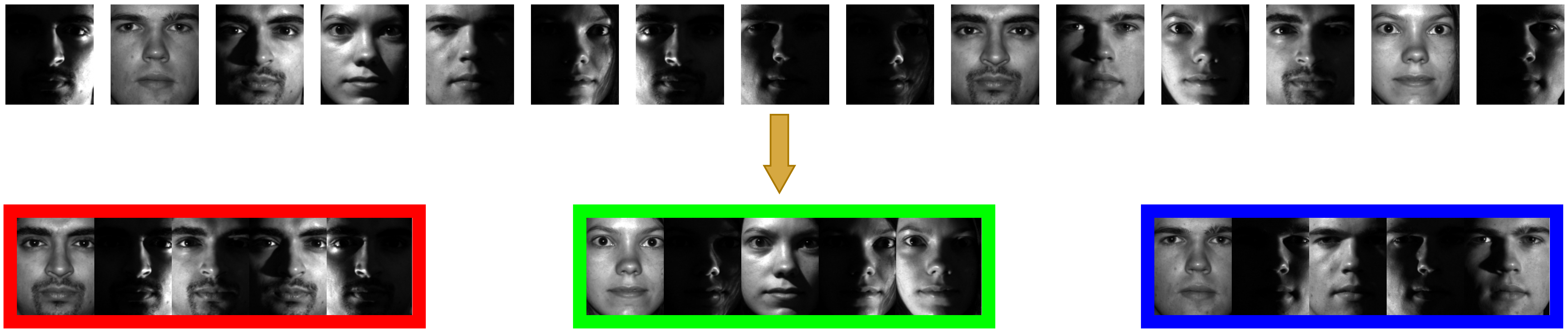
Experiments: synthetic

- Three subspaces of dimension $d = 4$
  - θ : subspace angle
  - N_g : # points in each subspace

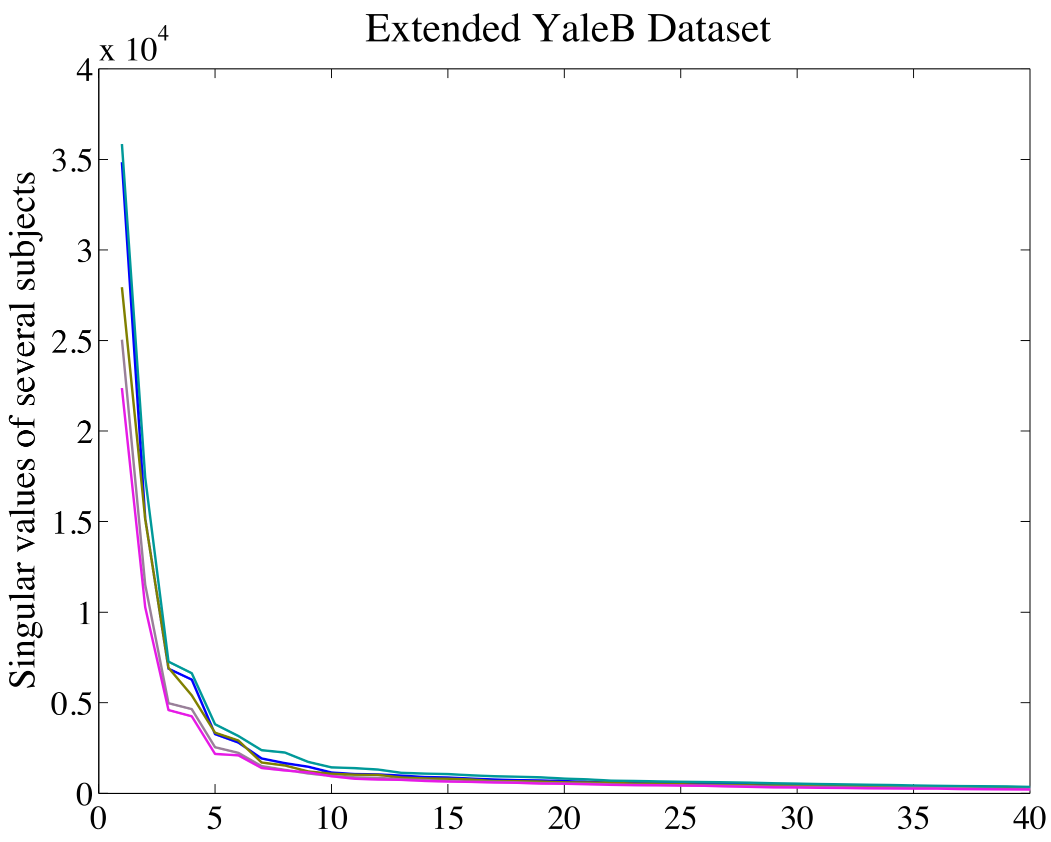




Experiments: face clustering

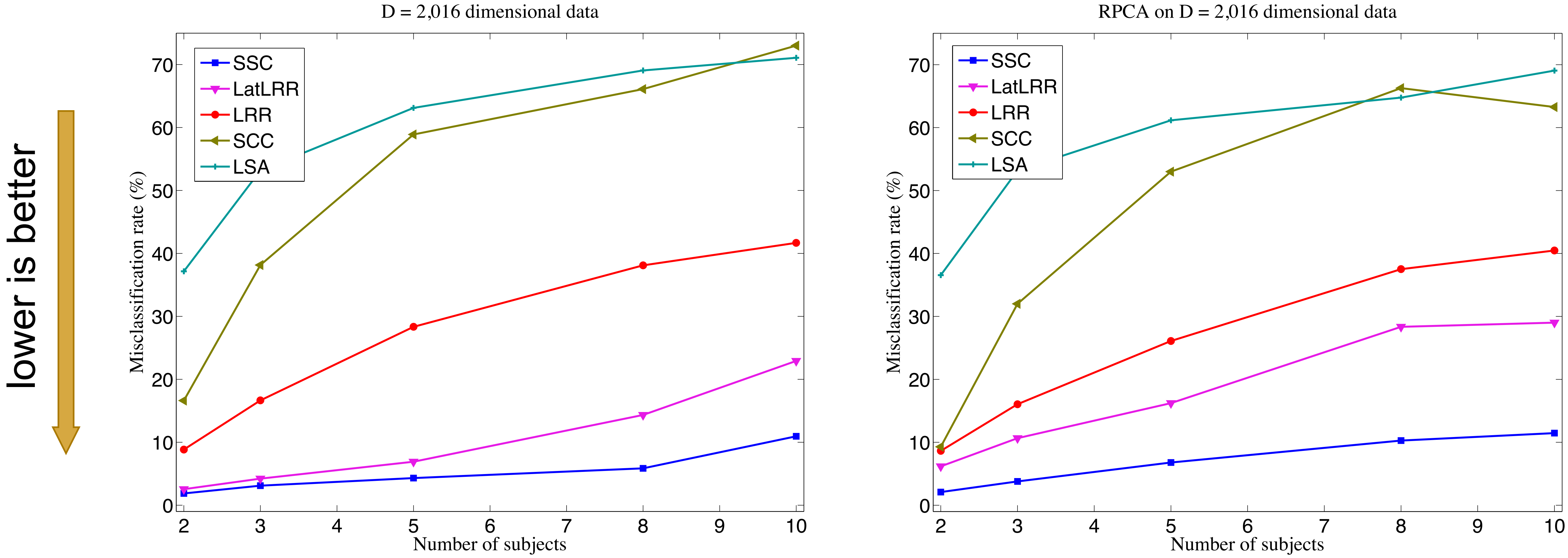


- Faces under varying illumination
  - lie in a 9-dim subspace
  - Extended Yale B dataset



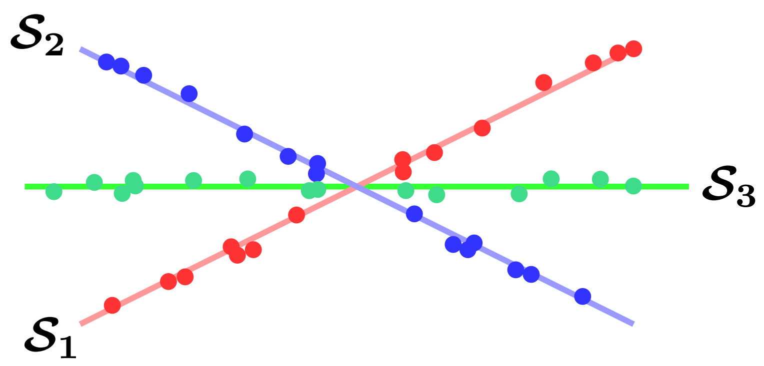
Experiments: face clustering

- SSC < 2.0% error for 2 subjects
- SSC < 11.0% error for 10 subjects



Other problems

- Noisy data $\boldsymbol{y}_i = \underbrace{\boldsymbol{y}_i^0}_{\text{noise-free}} + \underbrace{\boldsymbol{z}_i^0}_{\text{noise}}$



- Outlying, missing entries

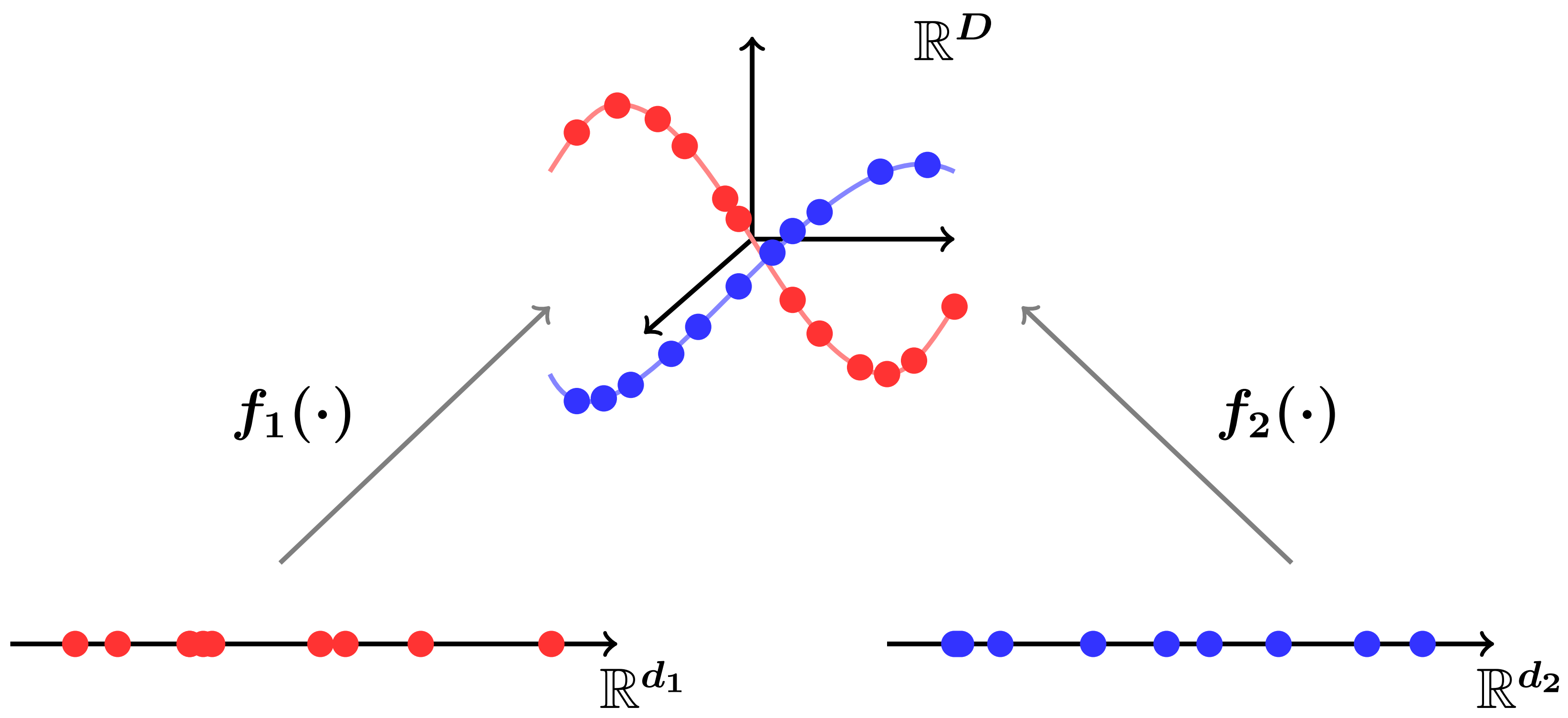
- Motion segmentation

Sparse Manifold Clustering & Embedding

Ehsan Elhamifar & Rene Vidal

Center for Imaging Science
Johns Hopkins University

Nonlinear manifolds

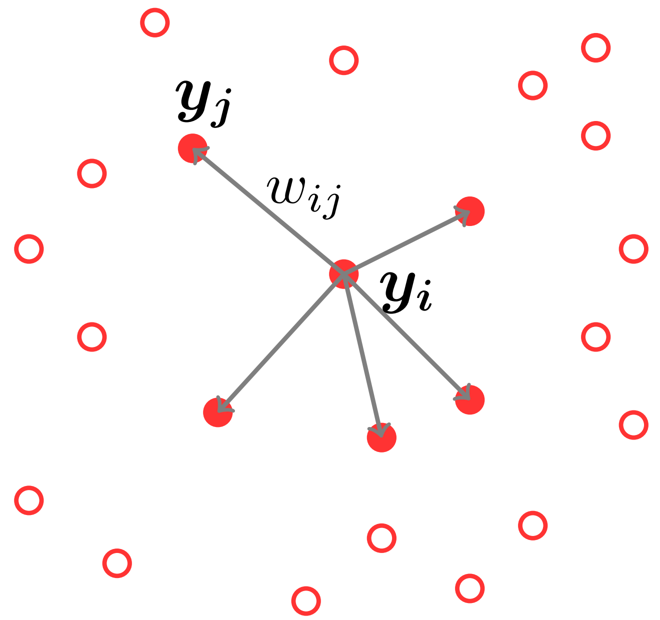


- Mappings are nonlinear
- Tasks:
  - cluster data into manifolds
  - find low-dimensional representations

Nonlinear dimensionality reduction

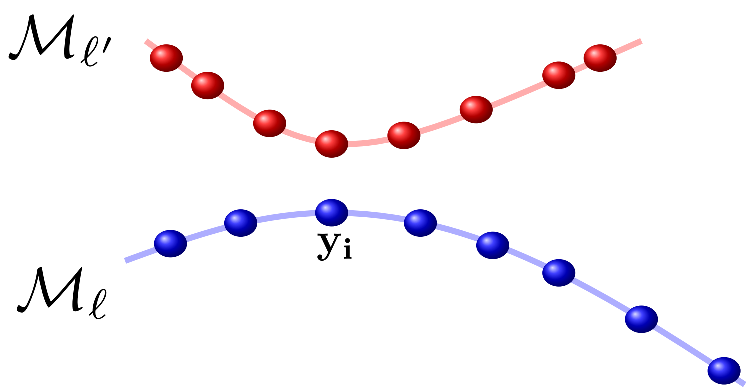
- LLE (Roweis`00), ISOMap (Tenenbaum`00), HLLE (Grimms`03), LEM (Belkin`02), MVU (Weinberger`04), MVE (Shaw`07), SPE (Jebara`09), ...
  - same in the first step
  - different in the second step

- Nonlinear dimension reduction
  - 1: build nearest neighbor graph
  - 2: learn weights
  - 3: find embedding from weights



Sparse manifold clustering and embedding

- Our method (SMCE)
  - 1: learn the neighborhood graph and its weights
  - 2: find embedding from weights

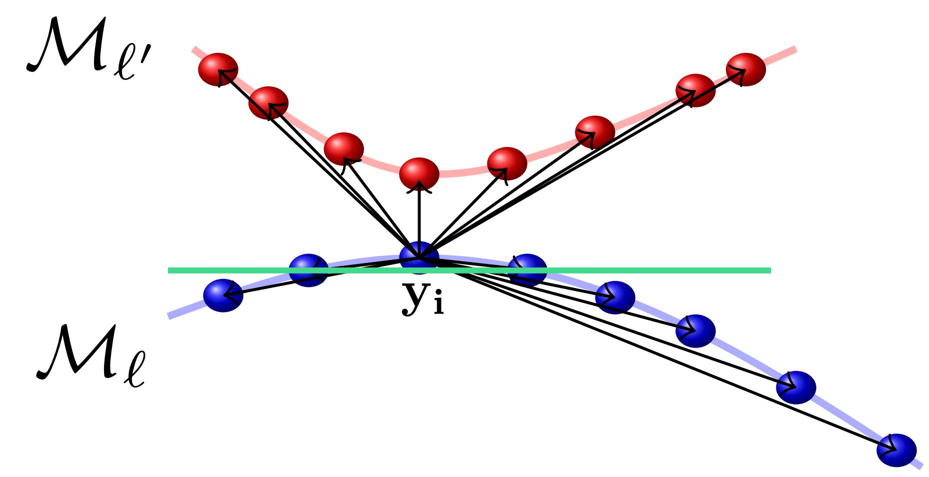


- Weights encode information for both clustering and embedding
  - deal with manifolds close to each other
  - deal with manifolds of different dimensions
    - automatically pick the right number of neighbors

E. Elhamifar and R. Vidal, Sparse Manifold Clustering and Embedding, NIPS11

Sparse manifold clustering and embedding

- $\mathcal{M}_\ell$ of intrinsic dimension d_ℓ
- Affine span of $d_\ell + 1$ points from $\mathcal{M}_\ell$ is close to $\boldsymbol{y}_i$



- Optimization program

$$\min \|\boldsymbol{q}_i \odot \boldsymbol{c}_i\|_1 \quad \text{s.t.} \quad \left[\frac{\boldsymbol{y}_1 - \boldsymbol{y}_i}{\|\boldsymbol{y}_1 - \boldsymbol{y}_i\|_2} \quad \cdots \quad \frac{\boldsymbol{y}_N - \boldsymbol{y}_i}{\|\boldsymbol{y}_N - \boldsymbol{y}_i\|_2}\right] \boldsymbol{c}_i \approx \mathbf{0}, \quad \mathbf{1}^\top \boldsymbol{c}_i = 1$$

few close points — span affine subspace

- proximity inducing vector: $\boldsymbol{q}_i \triangleq \left[\frac{\|\boldsymbol{y}_1 - \boldsymbol{y}_i\|_2}{\sum_{t \neq i} \|\boldsymbol{y}_t - \boldsymbol{y}_i\|_2} \quad \cdots \quad \frac{\|\boldsymbol{y}_N - \boldsymbol{y}_i\|_2}{\sum_{t \neq i} \|\boldsymbol{y}_t - \boldsymbol{y}_i\|_2}\right]^\top$

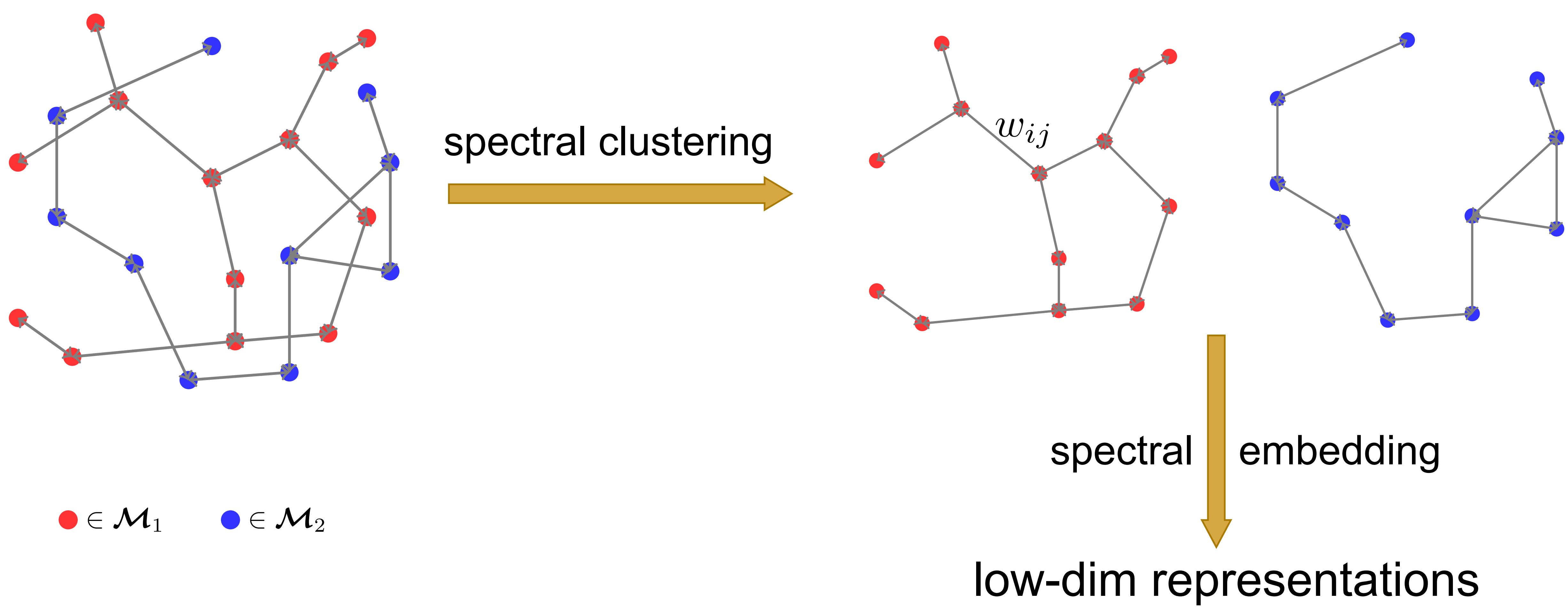
$\odot$: Hadamard (entry-wise) product

SMCE: algorithm

- For each data point solve

$$\min \ \lambda \|\boldsymbol{q}_i \odot \boldsymbol{c}_i\|_1 + \frac{1}{2} \left\| \begin{bmatrix} \frac{\boldsymbol{y}_1 - \boldsymbol{y}_i}{\|\boldsymbol{y}_1 - \boldsymbol{y}_i\|_2} & \cdots & \frac{\boldsymbol{y}_N - \boldsymbol{y}_i}{\|\boldsymbol{y}_N - \boldsymbol{y}_i\|_2} \end{bmatrix} \boldsymbol{c}_i \right\|_2^2 \quad \text{s. t.} \quad \mathbf{1}^\top \boldsymbol{c}_i = 1$$

- Build a similarity graph

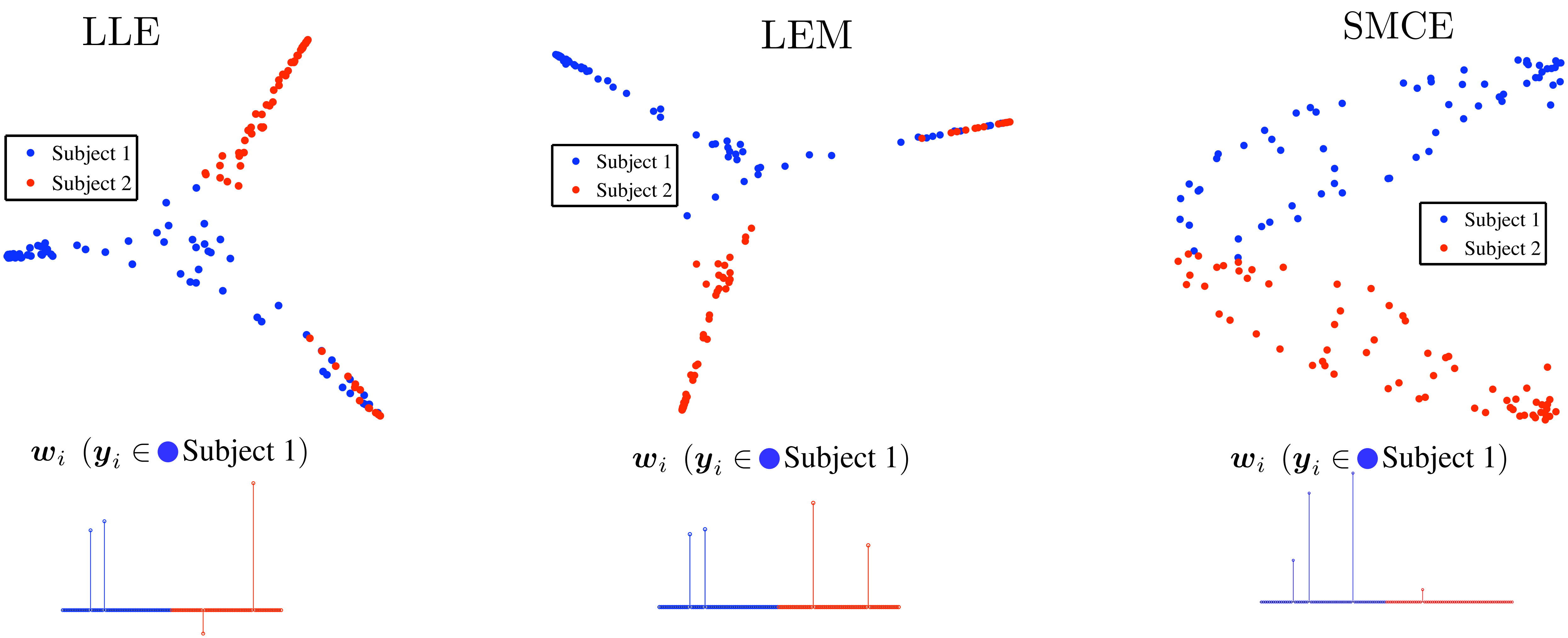


Experiments: real data

- Clustering and DR: faces

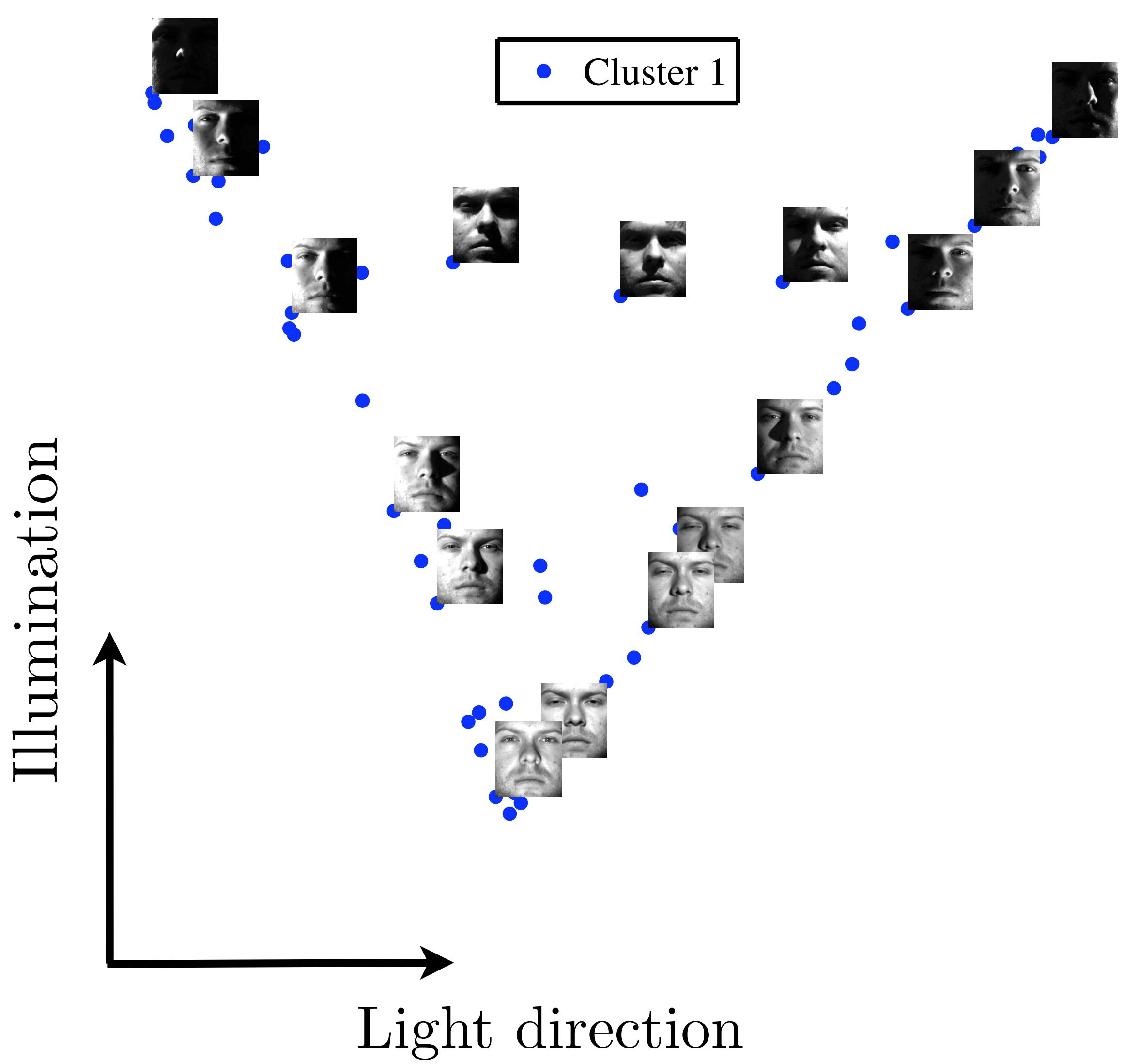
Table 1: Percentage of points whose K-NNs contain points from the other manifold.

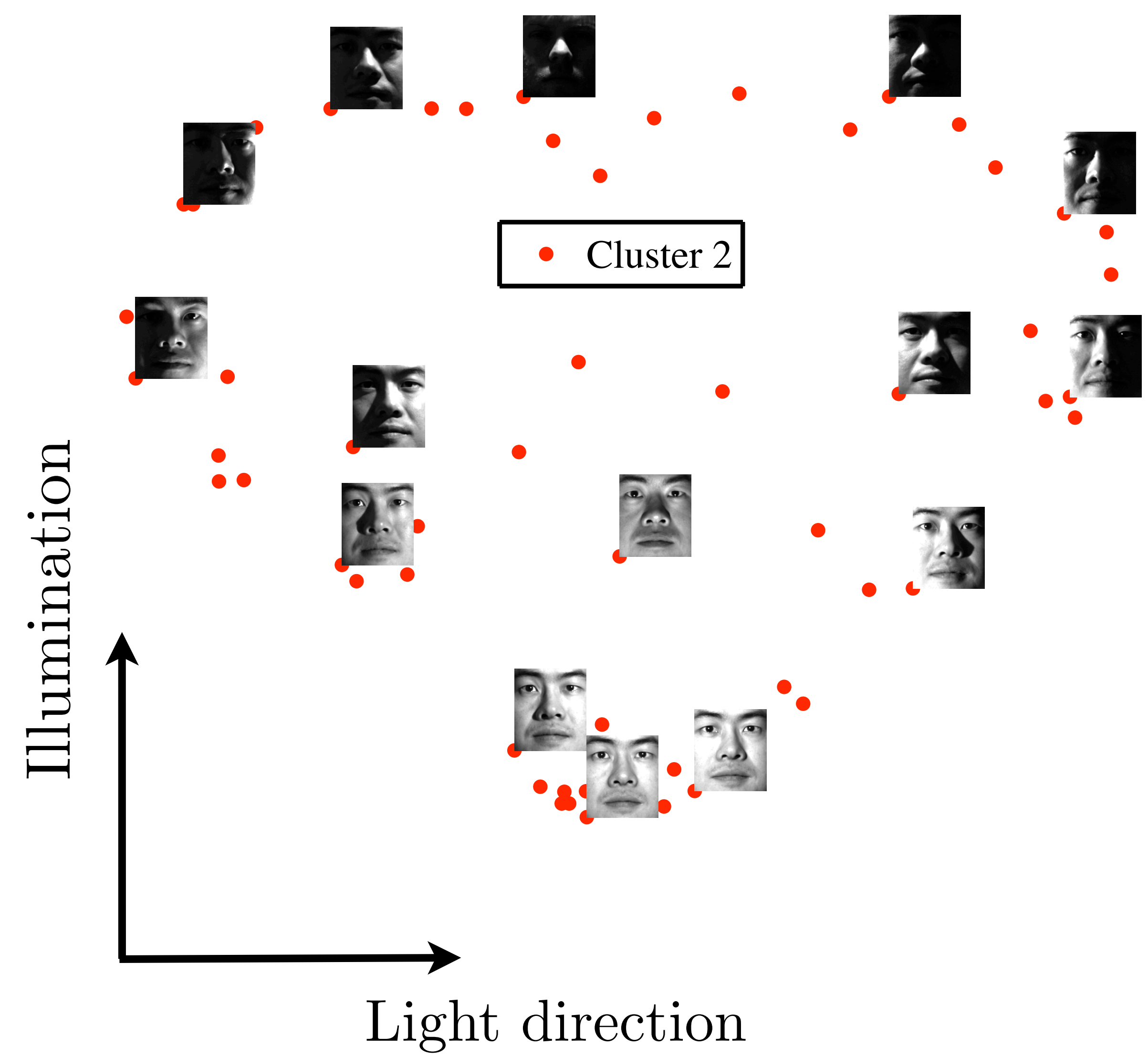
K	1	2	3	4	7	10
	3.9%	10.2%	23.4%	35.2%	57.0%	64.8%



Experiments: real data

- Clustering and DR: faces





Conclusions

- Exploited the self-expressiveness property of the data for
  - clustering subspaces
  - clustering and embedding of nonlinear manifolds
- Used sparse representation techniques
- Developed theoretical guarantees

Thanks!

References:

E. Elhamifar and R. Vidal, Sparse Subspace Clustering: Algorithm, Theory, and Applications, TPAMI.

E. Elhamifar and R. Vidal, Sparse Manifold Clustering and Embedding, NIPS 2011.

E. Elhamifar and R. Vidal, Sparse Subspace Clustering, CVPR 2009.