Spatially-Encouraged Spectral Clustering

A Critical Revision of Spatially-Constrained Spectral Clustering

Levi John Wolf

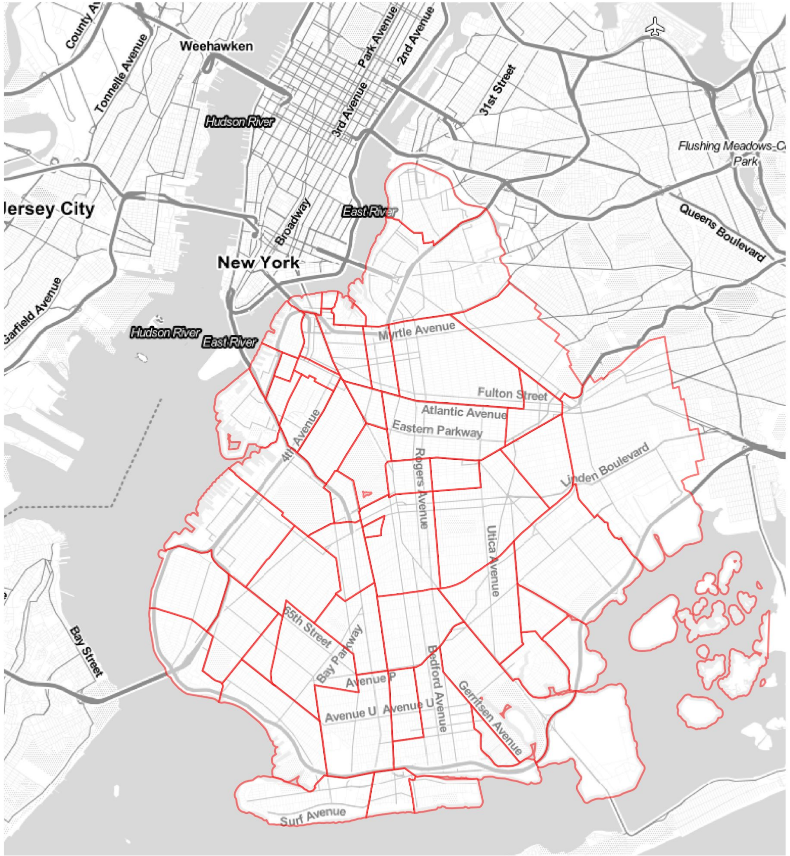
ljwolf.org/post/2018-gisruk

MOTIVATION

MECHANICS

EXAMPLE

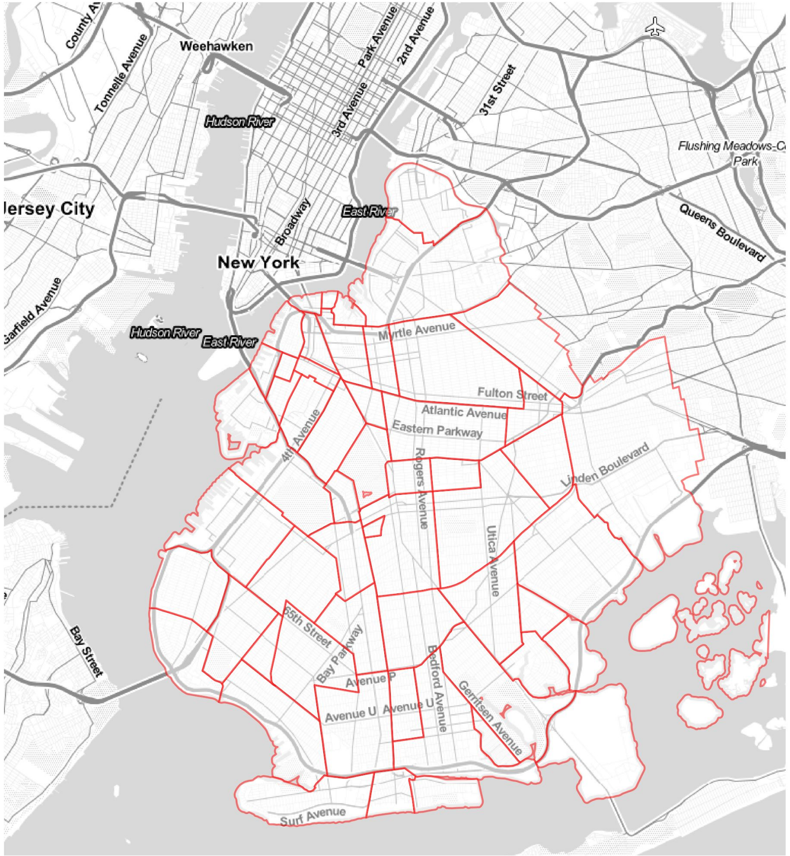
EXTENSION



GOAL:

Identify coherent sets of areas with internally-similar attributes or values.

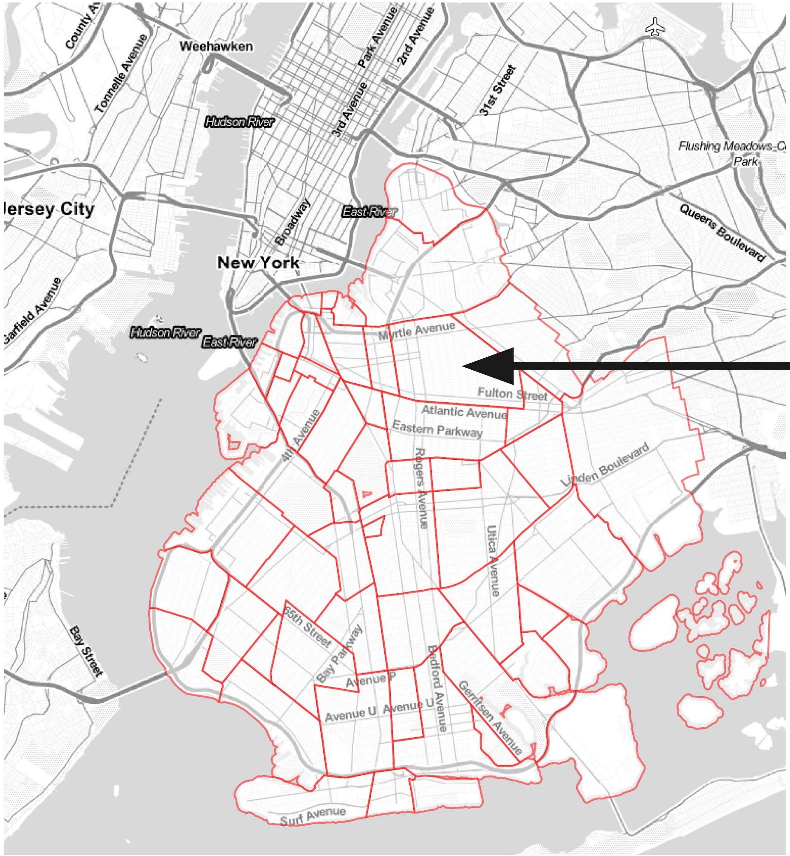
Basics of Affinity Clustering



GOAL:

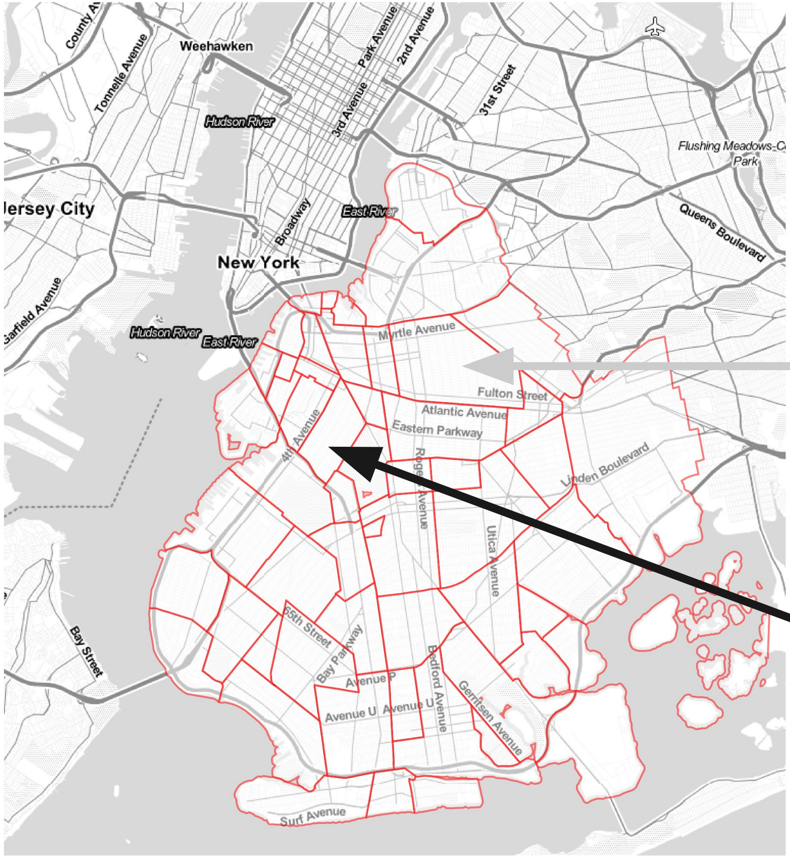
Identify coherent
neighbourhoods
regions
clusters
market areas
communities

Basics of Affinity Clustering



Income	45
% Young	22
% Uni	40
# Coffeeshops	200

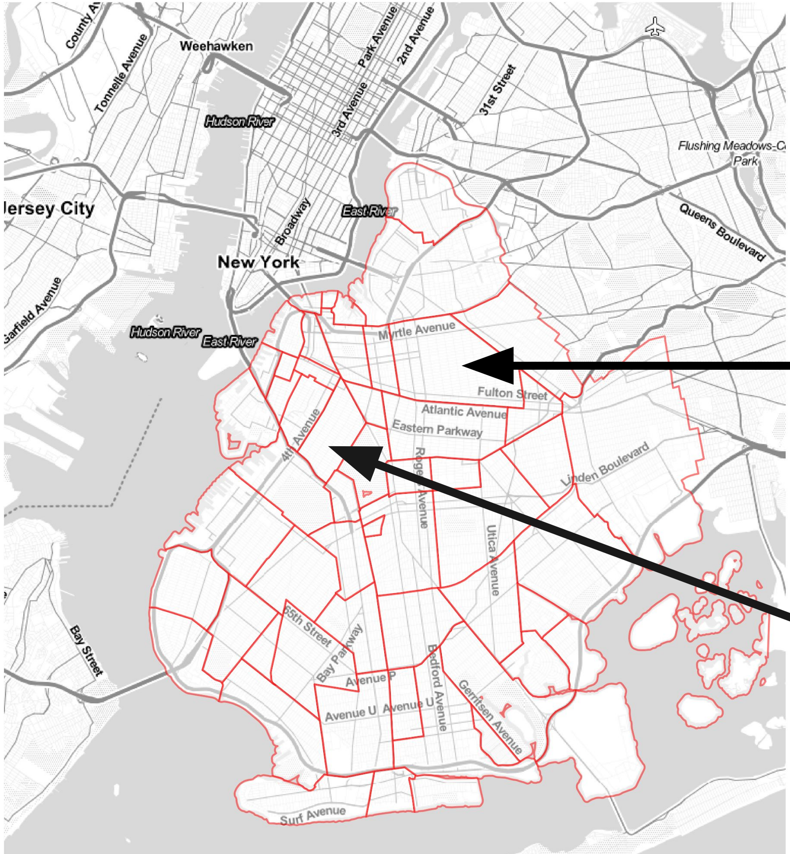
Basics of Affinity Clustering



Income	45
% Young	22
% Uni	40
# Coffeeshops	200

Income	63
% Young	23
% Uni	66
# Coffeeshops	400

Basics of Affinity Clustering



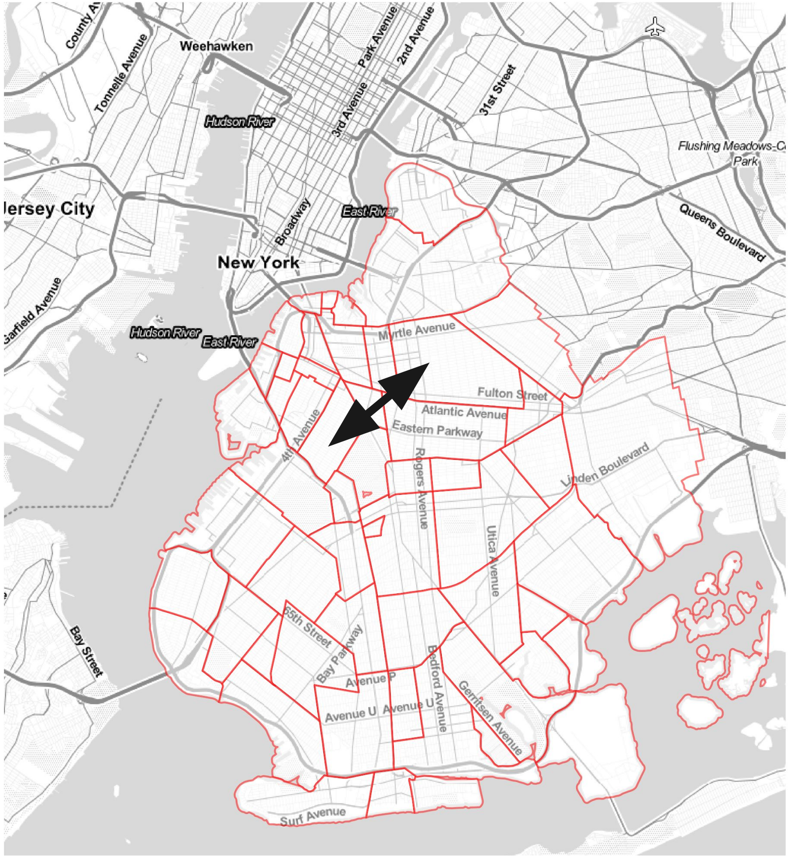
X_i

Income	45
% Young	22
% Uni	40
# Coffeeshops	200

X_j

Income	63
% Young	23
% Uni	66
# Coffeeshops	400

Basics of Affinity Clustering

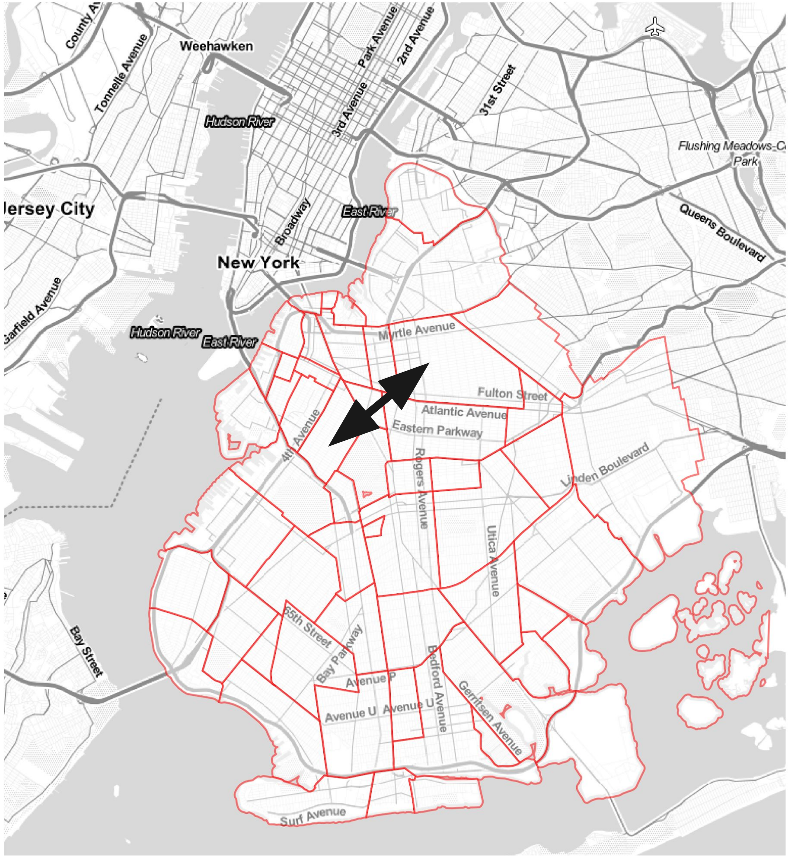


$$\varrho(X_i,X_j) \propto e^{-d(X_i,X_j)^{-1}}$$

Income	45
% Young	22
% Uni	40
# Coffeeshops	200

Income	63
% Young	23
% Uni	66
# Coffeeshops	400

Basics of Affinity Clustering

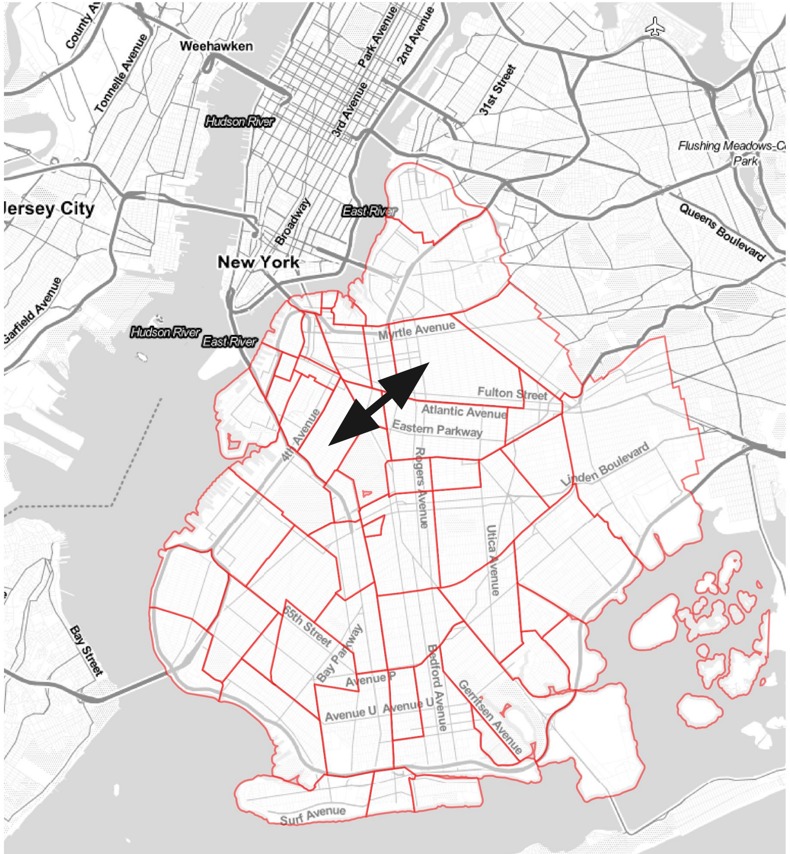


$$\varrho(X_i, X_j) \propto e^{-d(X_i, X_j)^{-1}}$$

AFFINITY METRIC

Score between zero and one reflecting similarity of two sites in the problem.

Basics of Affinity Clustering

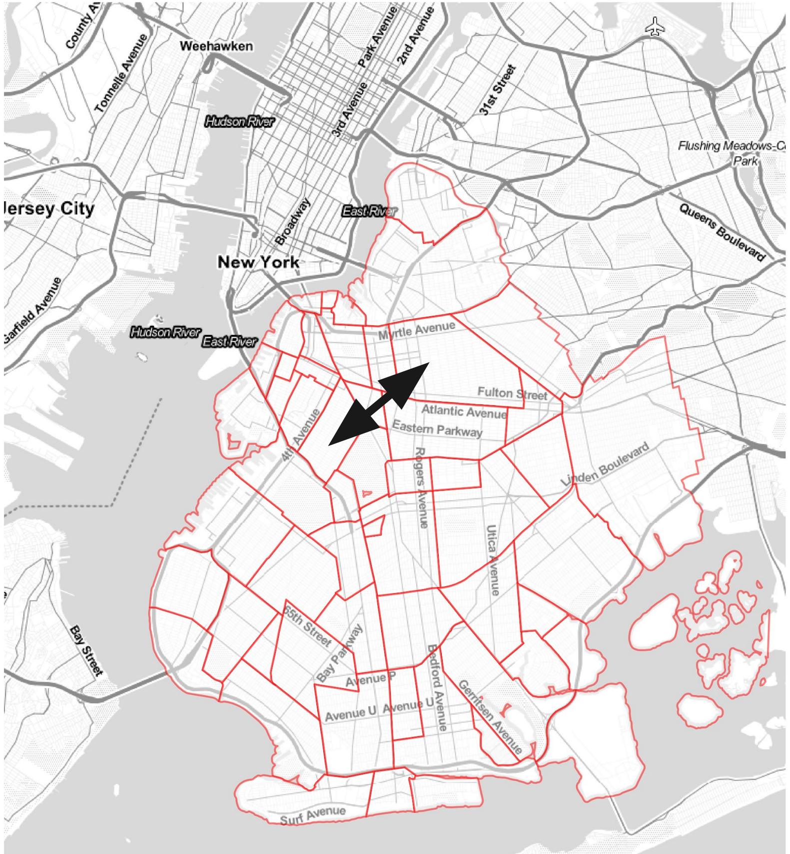


Income	45
% Young	22
% Uni	40
# Coffeeshops	200

$$\varrho(X_i, X_j) \propto e^{-d(X_i, X_j)^{-1}} \propto -d(X_i, X_j)^{-1} \propto X_i^t X_j$$

Income	63
% Young	23
% Uni	66
# Coffeeshops	400

Basics of Affinity Clustering



Income	45
% Young	22
% Uni	40
# Coffeeshops	200

$$\varrho(X_i,X_j) \propto e^{-d(X_i,X_j)^{-1}}$$

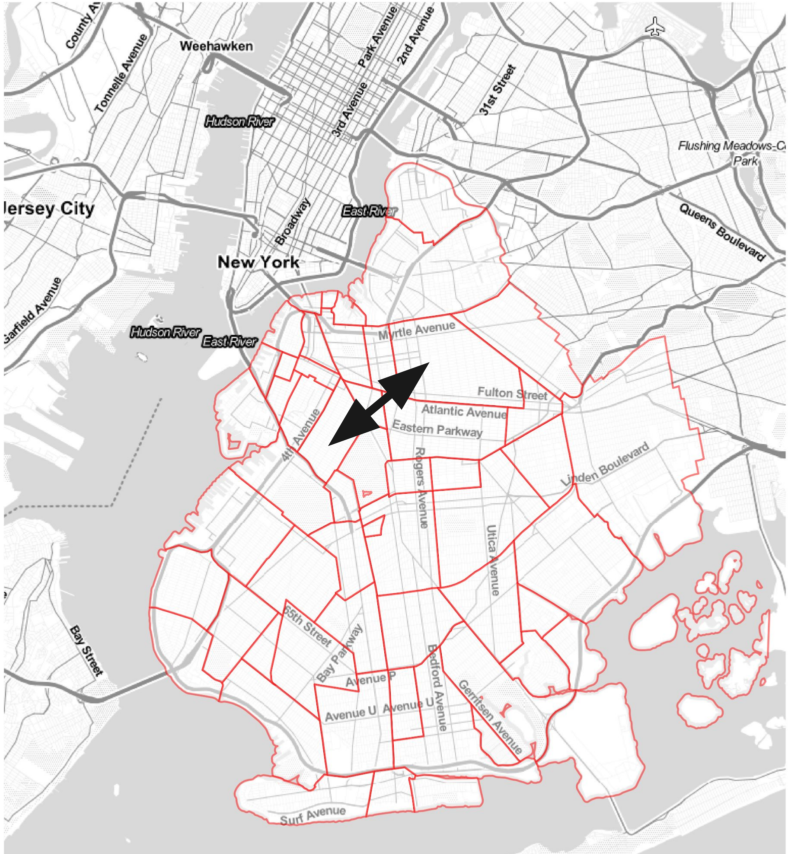
$$\propto -d(X_i,X_j)^{-1}$$

Income	63
% Young	23
% Uni	66
# Coffeeshops	400

$$\propto X_i^t X_j$$

$$\varrho(X_i,X_j) \propto \mathbf{K}(X_i,X_j \mid \theta)$$

Basics of Affinity Clustering



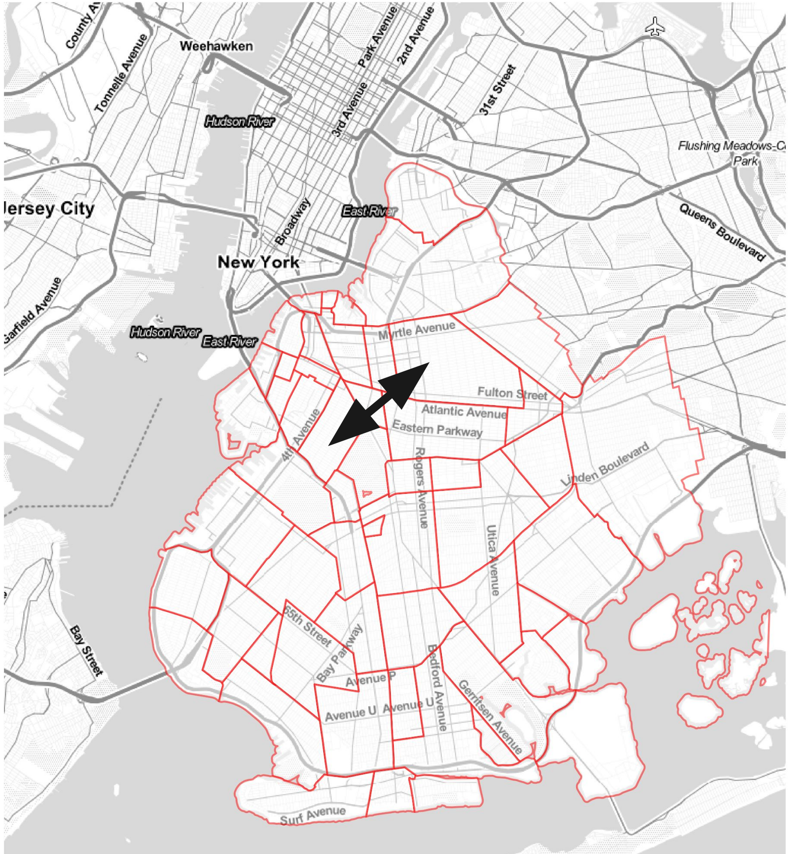
Income	45

```
[[1.  , 0.71, 0.59, 1.  , 0.02, 0.18],
 [0.71, 1.  , 0.98, 0.69, 0.13, 0.58],
 [0.59, 0.98, 1.  , 0.57, 0.19, 0.7 ],
 [1.  , 0.69, 0.57, 1.  , 0.02, 0.16],
 [0.02, 0.13, 0.19, 0.02, 1.  , 0.62],
 [0.18, 0.58, 0.7 , 0.16, 0.62, 1.  ]]
```

Income	63
% Young	23
% Uni	[illegible]
# Coffeeshops	400

$$\varrho(X_i, X_j) \propto \mathbf{K}(X_., X_. \mid \theta)$$

Basics of Affinity Clustering



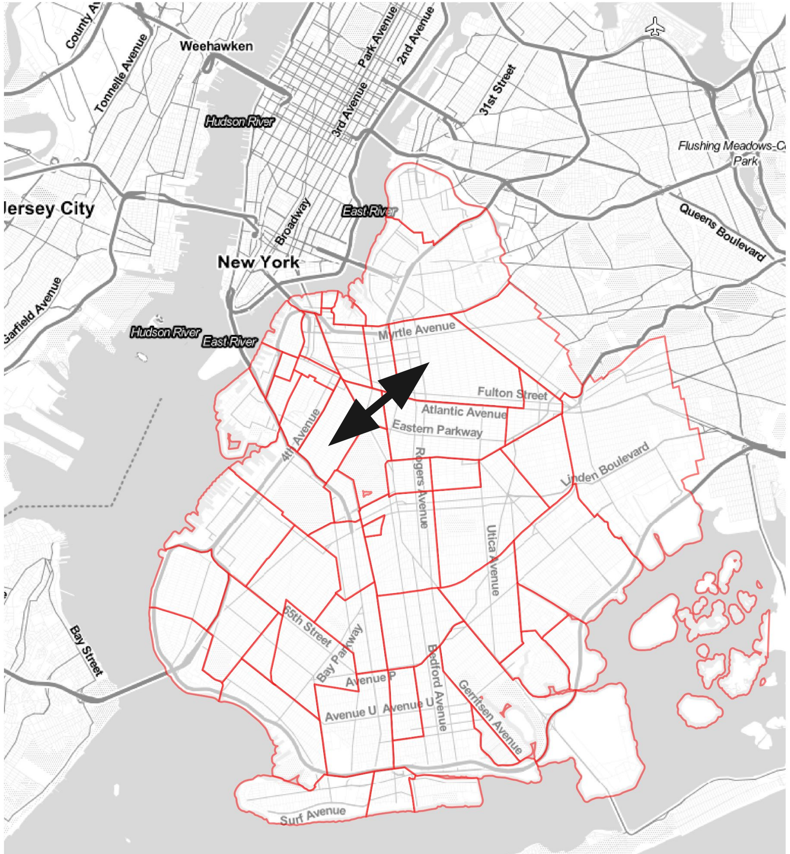
Income	45

```
[[1.  , 0.71, 0.59, 1.  , 0.02, 0.18],
 [0.71, 1.  , 0.98, 0.69, 0.13, 0.58],
 [0.59, 0.98, 1.  , 0.57, 0.19, 0.7 ],
 [1.  , 0.69, 0.57, 1.  , 0.02, 0.16],
 [0.02, 0.13, 0.19, 0.02, 1.  , 0.62],
 [0.18, 0.58, 0.7 , 0.16, 0.62, 1.  ]]
```

Income	63
% Young	23
% Uni	[illegible]
# Coffeeshops	400

$$\varrho(X_i, X_j) \propto \mathbf{K}(X_i, X_\cdot \mid \theta)$$

Basics of Affinity Clustering



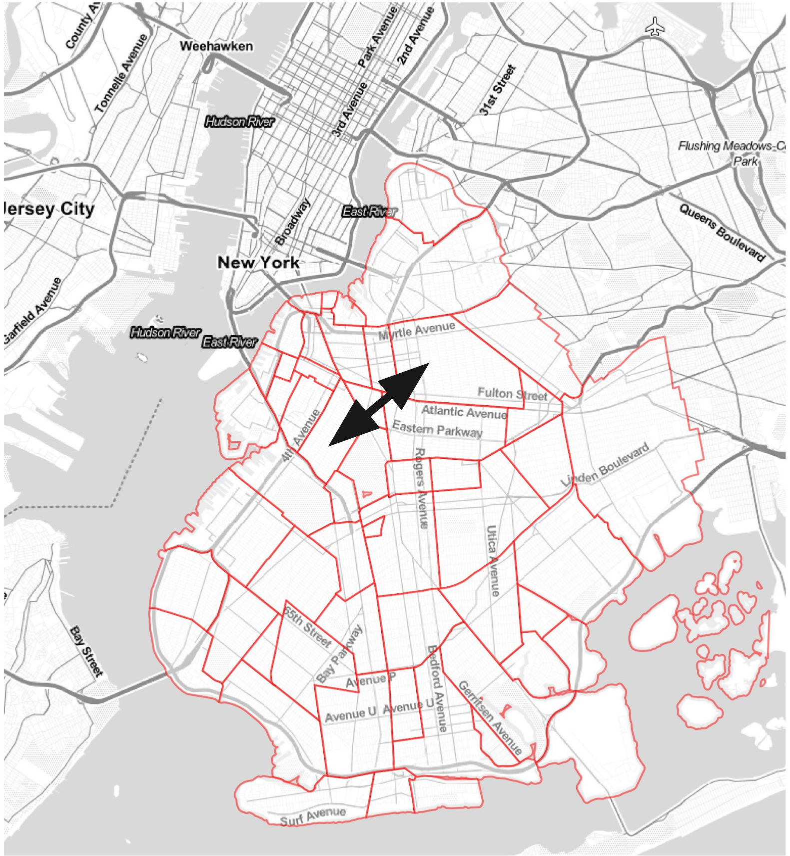
Income	45

```
[[1.  , 0.71, 0.59, 1.  , 0.02, 0.18],
 [0.71, 1.  , 0.98, 0.69, 0.13, 0.58],
 [0.59, 0.98, 1.  , 0.57, 0.19, 0.7 ],
 [1.  , 0.69, 0.57, 1.  , 0.02, 0.16],
 [0.02, 0.13, 0.19, 0.02, 1.  , 0.62],
 [0.18, 0.58, 0.7 , 0.16, 0.62, 1.  ]]
```

Income	63
% Young	23
% Uni	56
# Coffeeshops	400

$$\varrho(X_i, X_j) \propto \mathbf{K}(X_i, X_j \mid \theta)$$

Basics of Affinity Clustering




```
[[1.  , 0.03, 0.01, 0.99, 0.  , 0.  ],
 [0.03, 1.  , 0.82, 0.02, 0.  , 0.  ],
 [0.01, 0.82, 1.  , 0.  , 0.  , 0.03],
 [0.99, 0.02, 0.  , 1.  , 0.  , 0.  ],
 [0.  , 0.  , 0.  , 0.  , 1.  , 0.01],
 [0.  , 0.  , 0.03, 0.  , 0.01, 1.  ]]
```

Income	63
% Young	23
% Uni	65
# Coffeeshops	400

$$\varrho(X_i, X_j) \propto \mathbf{K}(X_i, X_j \mid \theta')$$

Basics of Affinity Clustering

$$\mathbf{L} = \mathbf{D} - \mathbf{A}_f(\theta)$$

The Graph Laplacian

$$\mathbf{L} = \mathbf{D} - \mathbf{A}_f(\theta)$$

Affinity Matrix

N x N set of relationships between observations, varying in $[0,1]$

Alternative spelling of $\mathbf{K}(X_i, X_j \mid \theta)$

$$\mathbf{L} = \mathbf{D} - \mathbf{A}_f(\theta)$$

Degree Matrix

Total "affinity mass" associated with each observation on a diagonal, zero elsewhere

$$\mathbf{L} = \mathbf{D} - \mathbf{A}_f(\theta)$$

Laplacian Matrix

Sufficient representation of affinity graph of $\mathbf{A}_f(\theta)$

$$\mathbf{L} = \mathbf{D} - \mathbf{A}_f(\theta)$$

Laplacian Matrix

Sufficient representation of affinity graph of $\mathbf{A}_f(\theta)$

$$\mathbf{L}^* = I - \mathbf{D}^{-\frac{1}{2}}\mathbf{A}_f(\theta)\mathbf{D}^{-\frac{1}{2}}$$

Symmetric Normalized Laplacian

Sufficient representation of affinity graph of $\mathbf{A}_f(\theta)$ with simpler eigensystem

$$\mathbf{L} = \mathbf{D} - \mathbf{A}_f(\theta)$$

```
[[1.03, 0.  , 0.  , 0.  , 0.  , 0.  ],        [[0.  , 0.03, 0.01, 0.99, 0.  , 0.  ],
 [0.  , 0.87, 0.  , 0.  , 0.  , 0.  ],         [0.03, 0.  , 0.82, 0.02, 0.  , 0.  ],
 [0.  , 0.  , 0.86, 0.  , 0.  , 0.  ],         [0.01, 0.82, 0.  , 0.  , 0.  , 0.03],
 [0.  , 0.  , 0.  , 1.01, 0.  , 0.  ],    -    [0.99, 0.02, 0.  , 0.  , 0.  , 0.  ],
 [0.  , 0.  , 0.  , 0.  , 0.01, 0.  ],         [0.  , 0.  , 0.  , 0.  , 0.  , 0.01],
 [0.  , 0.  , 0.  , 0.  , 0.  , 0.04]]         [0.  , 0.  , 0.03, 0.  , 0.01, 0.  ]]
```

The Graph Laplacian, computationally

$$\mathbf{L} = \mathbf{D} - \mathbf{A}_f(\theta)$$

```
[[ 1.03, -0.03, -0.01, -0.99,  0.  ,  0.  ],
 [-0.03,  0.87, -0.82, -0.02,  0.  ,  0.  ],
 [-0.01, -0.82,  0.86,  0.  ,  0.  , -0.03],
 [-0.99, -0.02,  0.  ,  1.01,  0.  ,  0.  ],
 [ 0.  ,  0.  ,  0.  ,  0.  ,  0.01, -0.01],
 [ 0.  ,  0.  , -0.03,  0.  , -0.01,  0.04]]
```

The Graph Laplacian

$$\mathbf{L}^* = I - \mathbf{D}^{-\frac{1}{2}} \mathbf{A}_f(\theta) \mathbf{D}^{-\frac{1}{2}}$$

```
[[ 1.  , -0.03, -0.01, -0.97,  0.  ,  0.  ],
 [-0.03,  1.  , -0.95, -0.02,  0.  ,  0.  ],
 [-0.01, -0.95,  1.  ,  0.  ,  0.  , -0.16],
 [-0.97, -0.02,  0.  ,  1.  ,  0.  ,  0.  ],
 [ 0.  ,  0.  ,  0.  ,  0.  ,  1.  , -0.5 ],
 [ 0.  ,  0.  , -0.16,  0.  , -0.5 ,  1.  ]]
```

The Graph Laplacian*

IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 22, NO. 8, AUGUST 2000

Normalized Cuts and Image Segmentation

Jianbo Shi and Jitendra Malik, *Member, IEEE*

Abstract—We propose a novel approach for solving the perceptual grouping problem in vision. Rather than focusing on local features and their consistencies in the image data, our approach aims at extracting the global impression of an image. We treat image segmentation as a graph partitioning problem and propose a novel global criterion, the *normalized cut*, for segmenting the graph. The *normalized cut* criterion measures both the total dissimilarity between the different groups as well as the total similarity within the groups. We show that an efficient computational technique based on a generalized eigenvalue problem can be used to optimize this criterion. We have applied this approach to segmenting static images, as well as motion sequences, and found the results to be very encouraging.

Index Terms—Grouping, image segmentation, graph partitioning.

Shi & Malik's Realization

IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 22, NO. 8, AUGUST 2000

Normalized Cuts and Image Segmentation

Jianbo Shi and Jitendra Malik, *Member, IEEE*

Point - Eigenspectrum of $\mathbf{L}$ yields the approximate solution for minimum weighted cut of $A_f(\theta)$.

Shi & Malik's Realization: MWC is a Rayleigh Quot.

$$\mathbf{L} = \mathbf{D} - \mathbf{A}_f(\theta)$$

1. **Affinity matrix computed** *for all sites using parameter* θ *to structure the attribute kernel*
2. **Obtain top** k **eigenvectors of** $\mathbf{L}$ (*approx. OK!*)
3. **Compute** k**-means clustering** *or recursively partition based on unique values in eigenspace*
4. **Label** *original data using the eigenvector labels*

Spectral Clustering for the MWC

$$\mathbf{L} = \mathbf{D} - \mathbf{A}_f(\theta) \circ \mathbf{A}_s(\eta)$$

Spatially-Encouraged Spectral Clustering™

(a constrained laplace clustering brand)

Mixing Spatial & Attribute Kernels

$$\mathbf{L} = \mathbf{D} - \mathbf{A}_f(\theta) \circ \mathbf{A}_s(\eta)$$

Spatial Affinity Matrix

Model of the spatial proximity of relationships in the problem

$$\mathbf{L} = \mathbf{D} - \mathbf{A}_f(\theta) \circ \mathbf{A}_s(\eta)$$

Spatially-Encouraged Laplacian

Sufficient representation of attribute affinity graph of $\mathbf{A}_f(\theta)$ *restricted/informed by spatial relationships* $\mathbf{A}_s(\eta)$

Mixing Spatial & Attribute Kernels

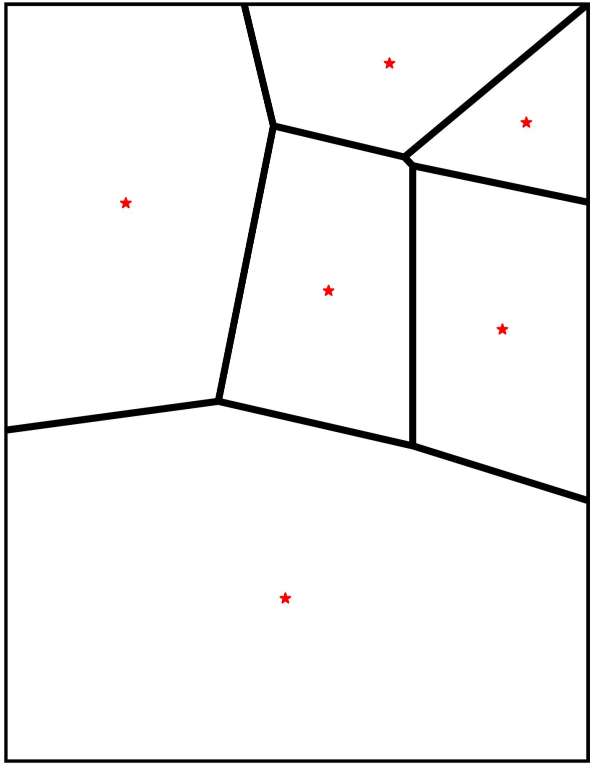
$$\mathbf{L} = \mathbf{D} - \mathbf{A}_f(\theta) \circ \mathbf{A}_s(\eta)$$

Spatially-Encouraged Laplacian

If $\mathbf{A}_s(\eta)_{ij} = -1$, *then i,j never connect!*

If $\mathbf{A}_s(\eta)_{ij} = 0$, *then i,j may only connect iff*

$\mathbf{A}_s(\eta)_{ik} \neq 0$ **AND** $\mathbf{A}_s(\eta)_{jk} \neq 0$ *for some k*

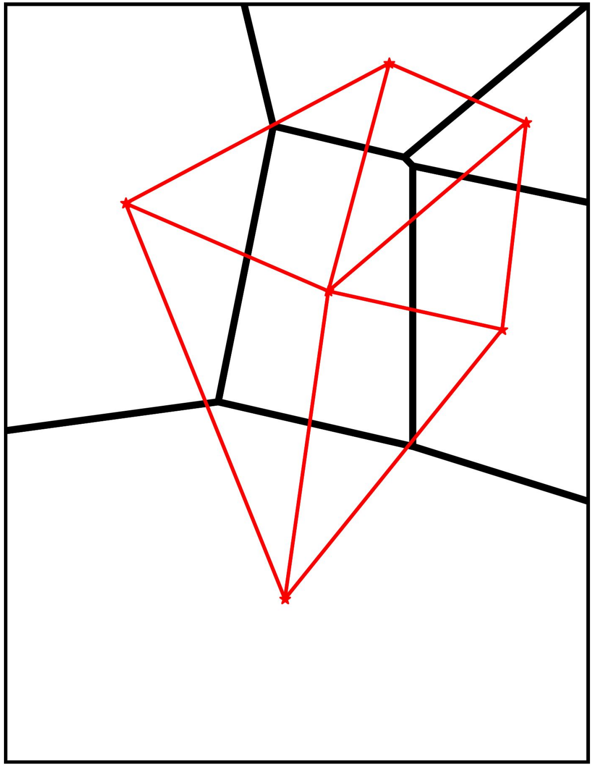


$$\mathbf{A}_f(\theta) =$$

```
[[0.  , 0.03, 0.01, 0.99, 0.  , 0.  ],
 [0.03, 0.  , 0.82, 0.02, 0.  , 0.  ],
 [0.01, 0.82, 0.  , 0.  , 0.  , 0.03],
 [0.99, 0.02, 0.  , 0.  , 0.  , 0.  ],
 [0.  , 0.  , 0.  , 0.  , 0.  , 0.01],
 [0.  , 0.  , 0.03, 0.  , 0.01, 0.  ]]
```

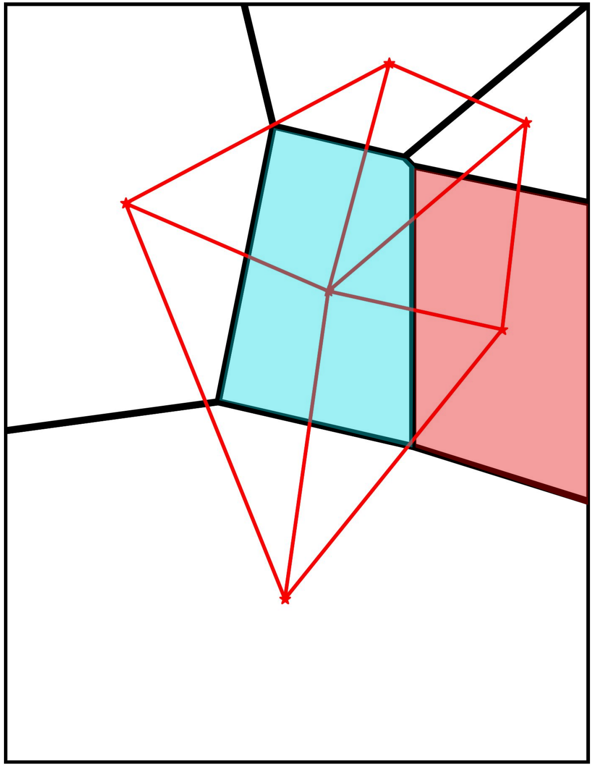
Example: Sixset

$$\mathbf{A}_f(\theta) = \begin{bmatrix} 0. & 0.03 & 0.01 & 0.99 & 0. & 0. \\ 0.03 & 0. & 0.82 & 0.02 & 0. & 0. \\ 0.01 & 0.82 & 0. & 0. & 0. & 0.03 \\ 0.99 & 0.02 & 0. & 0. & 0. & 0. \\ 0. & 0. & 0. & 0. & 0. & 0.01 \\ 0. & 0. & 0.03 & 0. & 0.01 & 0. \end{bmatrix}$$



$$\mathbf{A}_s(\eta) = \begin{bmatrix} 0. & 1. & 1. & 0. & 1. & 0. \\ 1. & 0. & 1. & 1. & 0. & 0. \\ 1. & 1. & 0. & 1. & 1. & 1. \\ 0. & 1. & 1. & 0. & 0. & 1. \\ 1. & 0. & 1. & 0. & 0. & 1. \\ 0. & 0. & 1. & 1. & 1. & 0. \end{bmatrix}$$

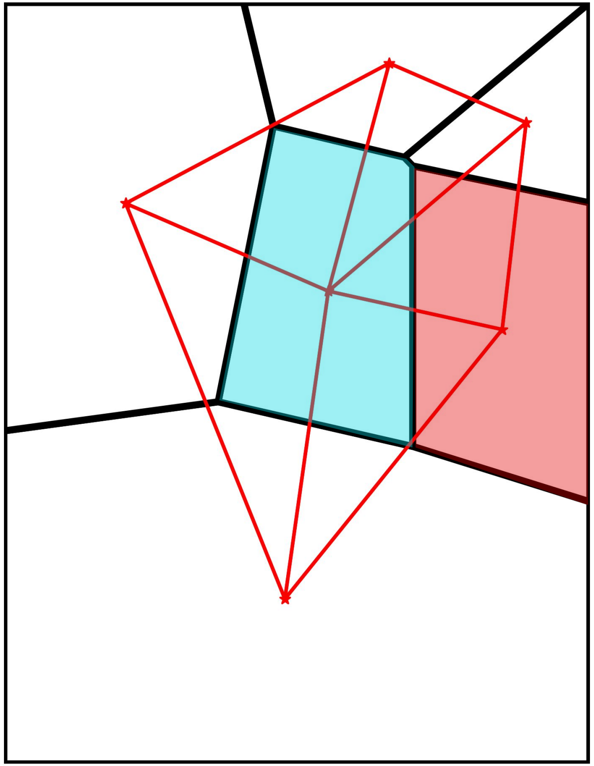
Example: Sixset



$$\mathbf{A}_f(\theta) = \begin{bmatrix} 0. & 0.03 & 0.01 & 0.99 & 0. & 0. \\ 0.03 & 0. & 0.82 & 0.02 & 0. & 0. \\ 0.01 & 0.82 & 0. & 0. & 0. & 0.03 \\ 0.99 & 0.02 & 0. & 0. & 0. & 0. \\ 0. & 0. & 0. & 0. & 0. & 0.01 \\ 0. & 0. & 0.03 & 0. & 0.01 & 0. \end{bmatrix}$$

$$\mathbf{A}_s(\eta) = \begin{bmatrix} 0. & 1. & 1. & 0. & 1. & 0. \\ 1. & 0. & 1. & 1. & 0. & 0. \\ 1. & 1. & 0. & 1. & 1. & 1. \\ 0. & 1. & 1. & 0. & 0. & 1. \\ 1. & 0. & 1. & 0. & 0. & 1. \\ 0. & 0. & 1. & 1. & 1. & 0. \end{bmatrix}$$

Example: Sixset



```
[[0.  , 0.03, 0.01, 0.99, 0.  , 0.  ],
 [0.03, 0.  , 0.82, 0.02, 0.  , 0.  ],
 [0.01, 0.82, 0.  , 0.  , 0.  , 0.03],
 [0.99, 0.02, 0.  , 0.  , 0.  , 0.  ],
 [0.  , 0.  , 0.  , 0.  , 0.  , 0.01],
 [0.  , 0.  , 0.03, 0.  , 0.01, 0.  ]]
```

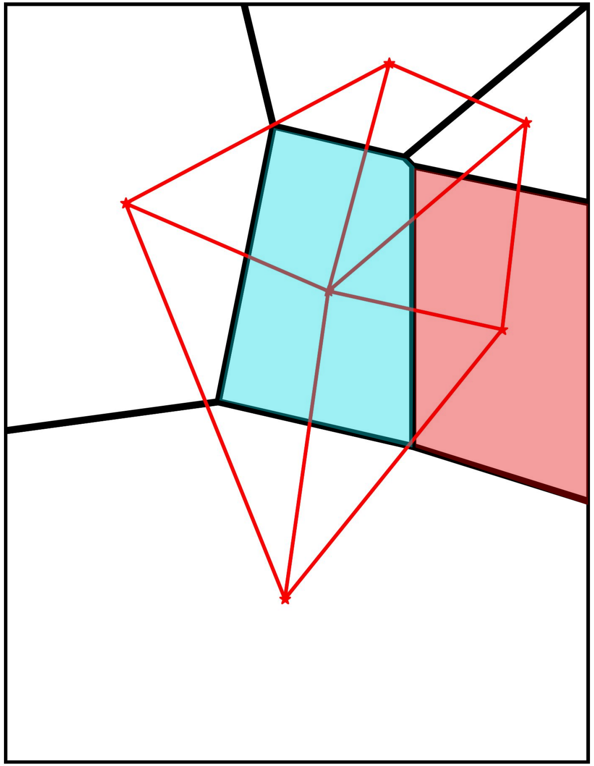
$\circ$

```
[[0., 1., 1., 0., 1., 0.],
 [1., 0., 1., 1., 0., 0.],
 [1., 1., 0., 1., 1., 1.],
 [0., 1., 1., 0., 0., 1.],
 [1., 0., 1., 0., 0., 1.],
 [0., 0., 1., 1., 1., 0.]]
```

```
[[ 0.04, -0.03, -0.01,  0.  ,  0.  ,  0.  ],
 [-0.03,  0.87, -0.82, -0.02,  0.  ,  0.  ],
 [-0.01, -0.82,  0.86,  0.  ,  0.  , -0.03],
 [ 0.  , -0.02,  0.  ,  0.02,  0.  ,  0.  ],
 [ 0.  ,  0.  ,  0.  ,  0.  ,  0.01, -0.01],
 [ 0.  ,  0.  , -0.03,  0.  , -0.01,  0.04]]
```

$$\mathbf{L} = \mathbf{D} - \mathbf{A}_f(\theta) \circ \mathbf{A}_s(\eta)$$

Example: Sixset



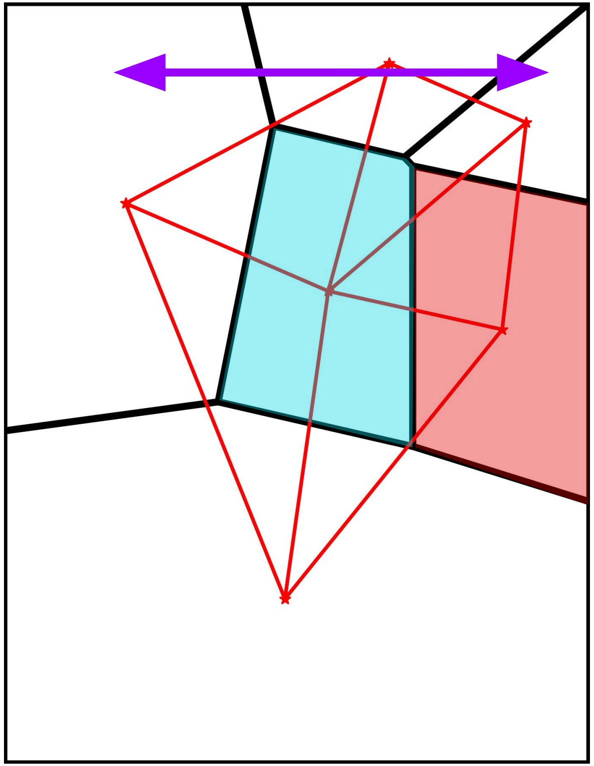
```
[[0.  , 0.03, 0.01, 0.99, 0.  , 0.  ],
 [0.03, 0.  , 0.82, 0.02, 0.  , 0.  ],
 [0.01, 0.82, 0.  , 0.  , 0.  , 0.03],
 [0.99, 0.02, 0.  , 0.  , 0.  , 0.  ],
 [0.  , 0.  , 0.  , 0.  , 0.  , 0.01],
 [0.  , 0.  , 0.03, 0.  , 0.01, 0.  ]]
```

```
[[0., 1., 1., 0., 1., 0.],
 [1., 0., 1., 1., 0., 0.],
 [1., 1., 0., 1., 1., 1.],
 [0., 1., 1., 0., 0., 1.],
 [1., 0., 1., 0., 0., 1.],
 [0., 0., 1., 1., 1., 0.]]
```

```
[[ 1.03, -0.03, -0.01, -0.99,  0.  ,  0.  ],
 [-0.03,  0.87, -0.82, -0.02,  0.  ,  0.  ],
 [-0.01, -0.82,  0.86,  0.  ,  0.  , -0.03],
 [-0.99, -0.02,  0.  ,  1.01,  0.  ,  0.  ],
 [ 0.  ,  0.  ,  0.  ,  0.  ,  0.01, -0.01],
 [ 0.  ,  0.  , -0.03,  0.  , -0.01,  0.04]]
```

$$\mathbf{L} = \mathbf{D} - \mathbf{A}_f(\theta)$$

Example: Sixset



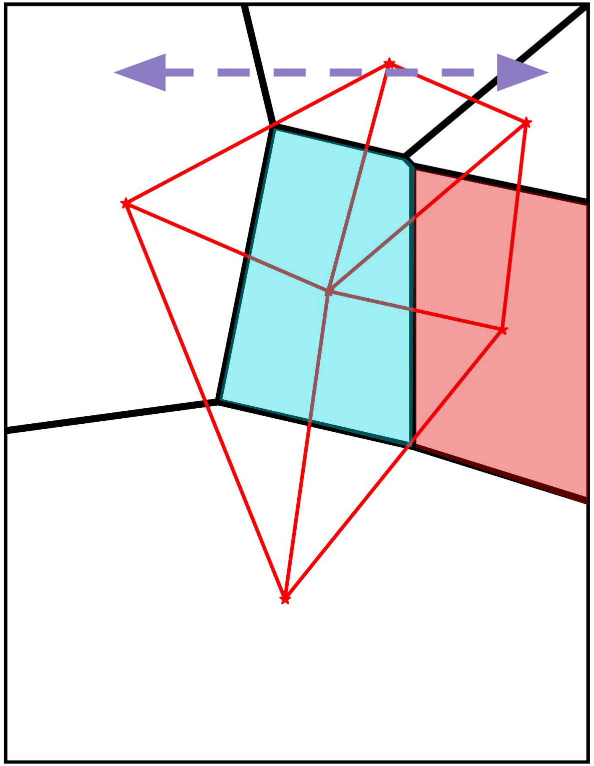
```
[[0.  , 0.03, 0.01, 0.99, 0.  , 0.  ],
 [0.03, 0.  , 0.82, 0.02, 0.  , 0.  ],
 [0.01, 0.82, 0.  , 0.  , 0.  , 0.03],
 [0.99, 0.02, 0.  , 0.  , 0.  , 0.  ],
 [0.  , 0.  , 0.  , 0.  , 0.  , 0.01],
 [0.  , 0.  , 0.03, 0.  , 0.01, 0.  ]]
```

```
[[0., 1., 1., 0., 1., 0.],
 [1., 0., 1., 1., 0., 0.],
 [1., 1., 0., 1., 1., 1.],
 [0., 1., 1., 0., 0., 1.],
 [1., 0., 1., 0., 0., 1.],
 [0., 0., 1., 1., 1., 0.]]
```

```
[[ 1.03, -0.03, -0.01, -0.99,  0.  ,  0.  ],
 [-0.03,  0.87, -0.82, -0.02,  0.  ,  0.  ],
 [-0.01, -0.82,  0.86,  0.  ,  0.  , -0.03],
 [-0.99, -0.02,  0.  ,  1.01,  0.  ,  0.  ],
 [ 0.  ,  0.  ,  0.  ,  0.  ,  0.01, -0.01],
 [ 0.  ,  0.  , -0.03,  0.  , -0.01,  0.04]]
```

$$\mathbf{L} = \mathbf{D} - \mathbf{A}_f(\theta)$$

Example: Sixset



```
[[0.  , 0.03, 0.01, 0.99, 0.  , 0.  ],
 [0.03, 0.  , 0.82, 0.02, 0.  , 0.  ],
 [0.01, 0.82, 0.  , 0.  , 0.  , 0.03],
 [0.99, 0.02, 0.  , 0.  , 0.  , 0.  ],
 [0.  , 0.  , 0.  , 0.  , 0.  , 0.01],
 [0.  , 0.  , 0.03, 0.  , 0.01, 0.  ]]
```

$\circ$

```
[[0., 1., 1., 0., 1., 0.],
 [1., 0., 1., 1., 0., 0.],
 [1., 1., 0., 1., 1., 1.],
 [0., 1., 1., 0., 0., 1.],
 [1., 0., 1., 0., 0., 1.],
 [0., 0., 1., 1., 1., 0.]]
```

```
[[ 0.04, -0.03, -0.01,  0.  ,  0.  ,  0.  ],
 [-0.03,  0.87, -0.82, -0.02,  0.  ,  0.  ],
 [-0.01, -0.82,  0.86,  0.  ,  0.  , -0.03],
 [ 0.  , -0.02,  0.  ,  0.02,  0.  ,  0.  ],
 [ 0.  ,  0.  ,  0.  ,  0.  ,  0.01, -0.01],
 [ 0.  ,  0.  , -0.03,  0.  , -0.01,  0.04]]
```

$$\mathbf{L} = \mathbf{D} - \mathbf{A}_f(\theta) \circ \mathbf{A}_s(\eta)$$

Example: Sixset

$$\mathbf{L} = \mathbf{D} - \mathbf{A}_f(\theta)$$

1. **Affinity matrix computed** *for all sites using parameter θ to structure the attribute kernel.*
2. **Obtain top k eigenvectors of $\mathbf{L}$** (*approx. OK!*)
3. **Compute k-means clustering** *or recursively partition based on unique values in eigenspace.*
4. **Label** *original data using the eigenvector labels.*

Spectral Clustering for the MWC

$$\mathbf{L} = \mathbf{D} - \mathbf{A}_f(\theta) \circ \mathbf{A}_s(\eta)$$

1. **Affinity matrix computed** *for all* ***near*** *sites using parameters* θ,η *to structure the kernels.*
2. **Obtain top** k **eigenvectors of** $\mathbf{L}$ (*approx. OK!*)
3. **Compute** k**-means clustering** *or recursively partition based on unique values in eigenspace.*
4. **Label** *original data using the eigenvector labels.*

Spectral Clustering for the Spatial MWC

Yuan et al. (2015)
(kind of)

$$\mathbf{L} = \mathbf{D} - \mathbf{A}_f(\theta) \circ \mathbf{A}_s(\eta)$$

1. **Affinity matrix computed** *for all* ***near*** *sites using parameters* θ, η *to structure the kernels.*
2. **Obtain top** k **eigenvectors of** $\mathbf{L}$ (*approx. OK!*)
3. **Compute** k**-means clustering** *or recursively partition based on unique values in eigenspace.*
4. **Label** *original data using the eigenvector labels.*

Spectral Clustering for the Spatial MWC

Yuan et al. (2015)
(kind of)

$$\mathbf{L} = \mathbf{D} - \mathbf{A}_f(\theta) \circ \mathbf{A}_s(\eta)$$

Constrained Spectral Clustering for Regionalization: Exploring the Trade-off between Spatial Contiguity and Landscape Homogeneity

Shuai Yuan, Pang-Ning Tan
Department of Computer Science and Engineering
Michigan State University
East Lansing, MI-48824
Email: {yuanshu2,ptan}@msu.edu

Kendra Spence Cheruvelil
Lyman Briggs College and Department of Fisheries and Wildlife
Michigan State University
East Lansing, MI-48824
Email: ksc@msu.edu

Sarah M. Collins, Patricia A. Soranno
Department of Fisheries and Wildlife
Michigan State University
East Lansing, MI-48824
Email: {colli636,soranno}@msu.edu

Mixing Spatial & Attribute Kernels

$$\mathbf{L} = \mathbf{D} - \mathbf{A}_f(\cancel{\theta}) \circ \mathbf{A}_s(\eta)$$

Constrained Spectral Clustering for Regionalization:

CLAIM (Yuan et al, 2015):

For a contiguity $\mathbf{A}_s(\eta)$, η controls solution.

θ never mentioned.

η is like convex weight, w in $w\mathbf{A} + (1\text{-}w)\mathbf{B}$

Mixing Spatial & Attribute Kernels

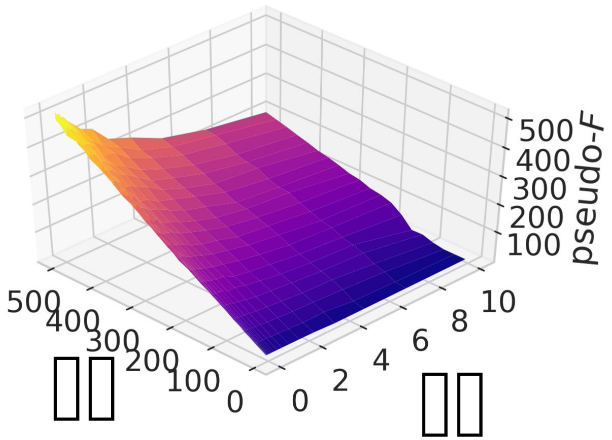
$$\mathbf{L} = \mathbf{D} - \mathbf{A}_f(\cancel{\theta}) \circ \mathbf{A}_s(\eta)$$

CLAIM (Yuan et al, 2015):

For a contiguity $\mathbf{A}_s(\eta)$, η controls solution.

θ never mentioned.

η is like convex weight, w in $w\mathbf{A} + (1\text{-}w)\mathbf{B}$



Mixing Spatial & Attribute Kernels

$$\mathbf{L} = \mathbf{D} - \mathbf{A}_f(\cancel{\theta}) \circ \mathbf{A}_s(\eta)$$

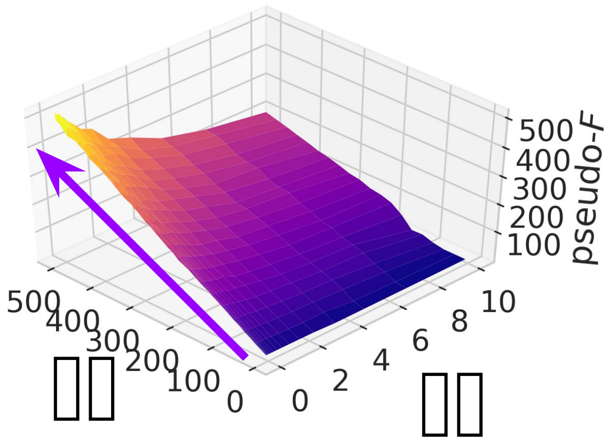
CLAIM (Yuan et al, 2015):

For a contiguity $\mathbf{A}_s(\eta)$, η controls solution.

θ never mentioned.

η is like convex weight, w in $w\mathbf{A} + (1\text{-}w)\mathbf{B}$

Change η while θ held constant



Mixing Spatial & Attribute Kernels

$$\mathbf{L} = \mathbf{D} - \mathbf{A}_f(\cancel{\theta}) \circ \mathbf{A}_s(\eta)$$

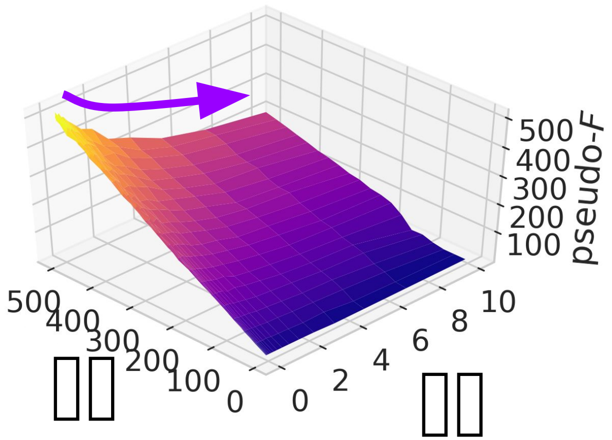
CLAIM (Yuan et al, 2015):

For a contiguity $\mathbf{A}_s(\eta)$, η controls solution.

θ never mentioned.

η is like convex weight, w in $w\mathbf{A} + (1\text{-}w)\mathbf{B}$

Hold η constant while θ changes



Mixing Spatial & Attribute Kernels

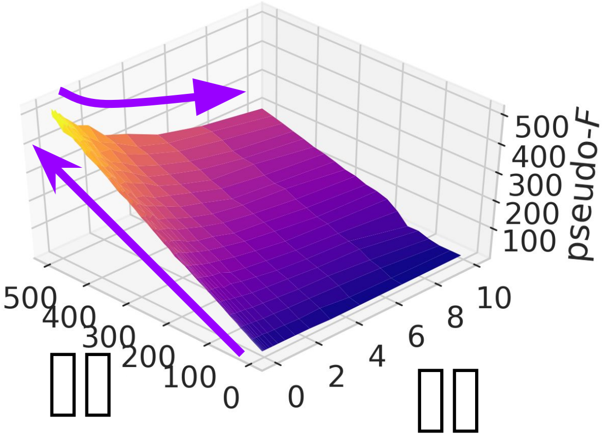
$$\mathbf{L} = \mathbf{D} - \mathbf{A}_f(\theta) \circ \mathbf{A}_s(\eta)$$

~~**CLAIM (Yuan et al, 2015):**~~

~~*For a contiguity* $\mathbf{A}_s(\eta)$*,* η *controls solution.*~~

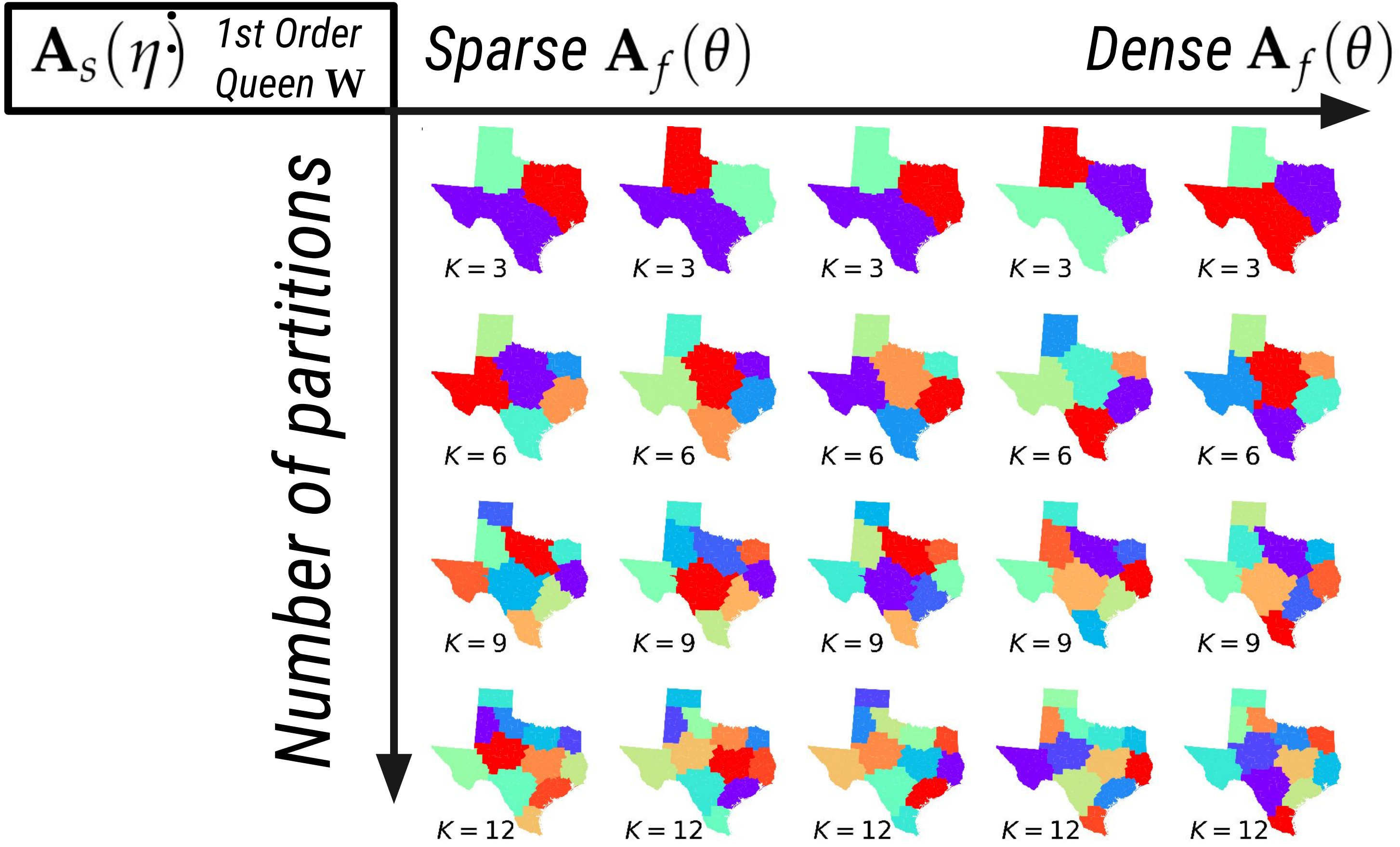
~~θ *never mentioned.*~~

~~η *is like convex weight,* w *in* $w\mathbf{A} + (1\text{-}w)\mathbf{B}$~~

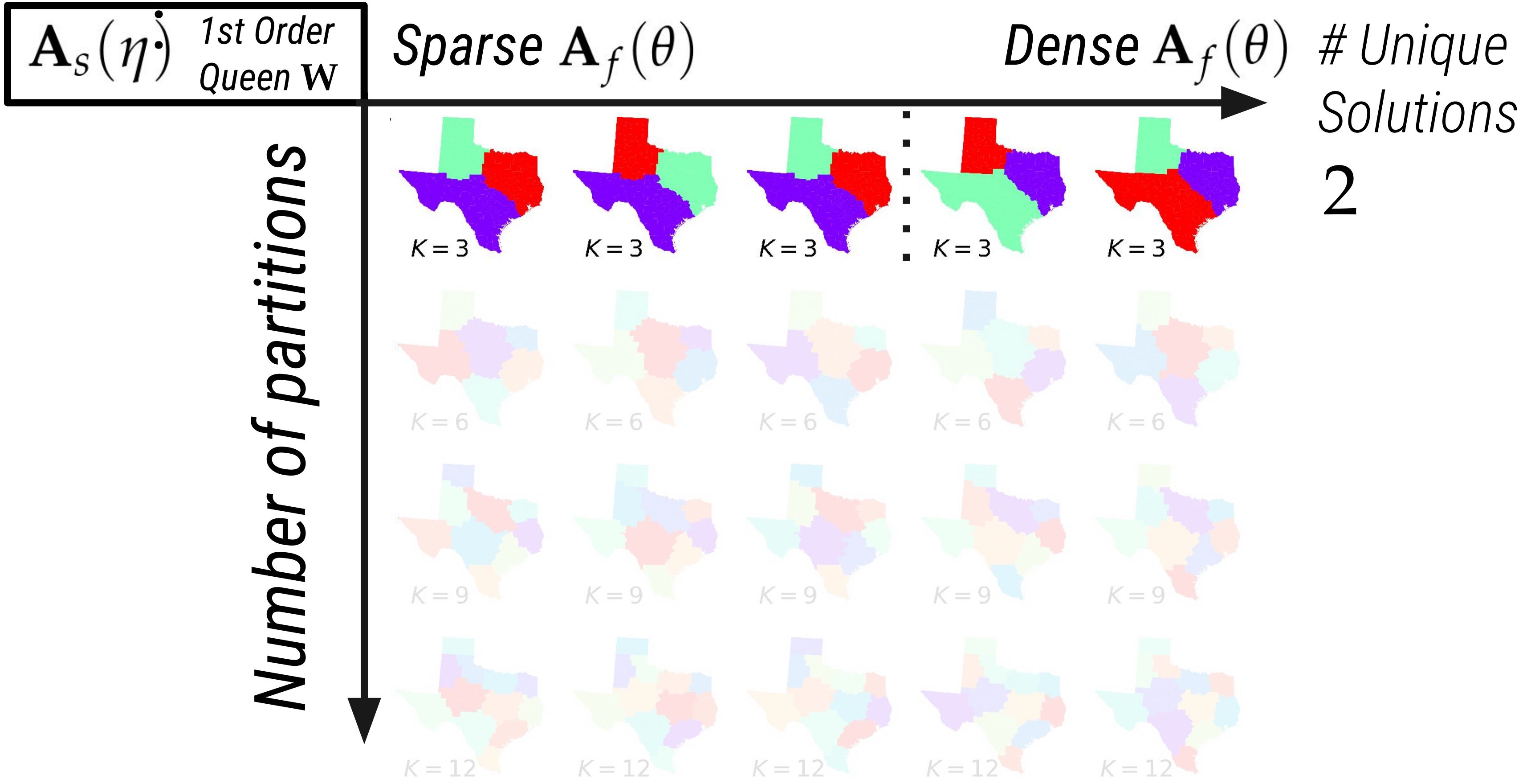


BOTH MATTER

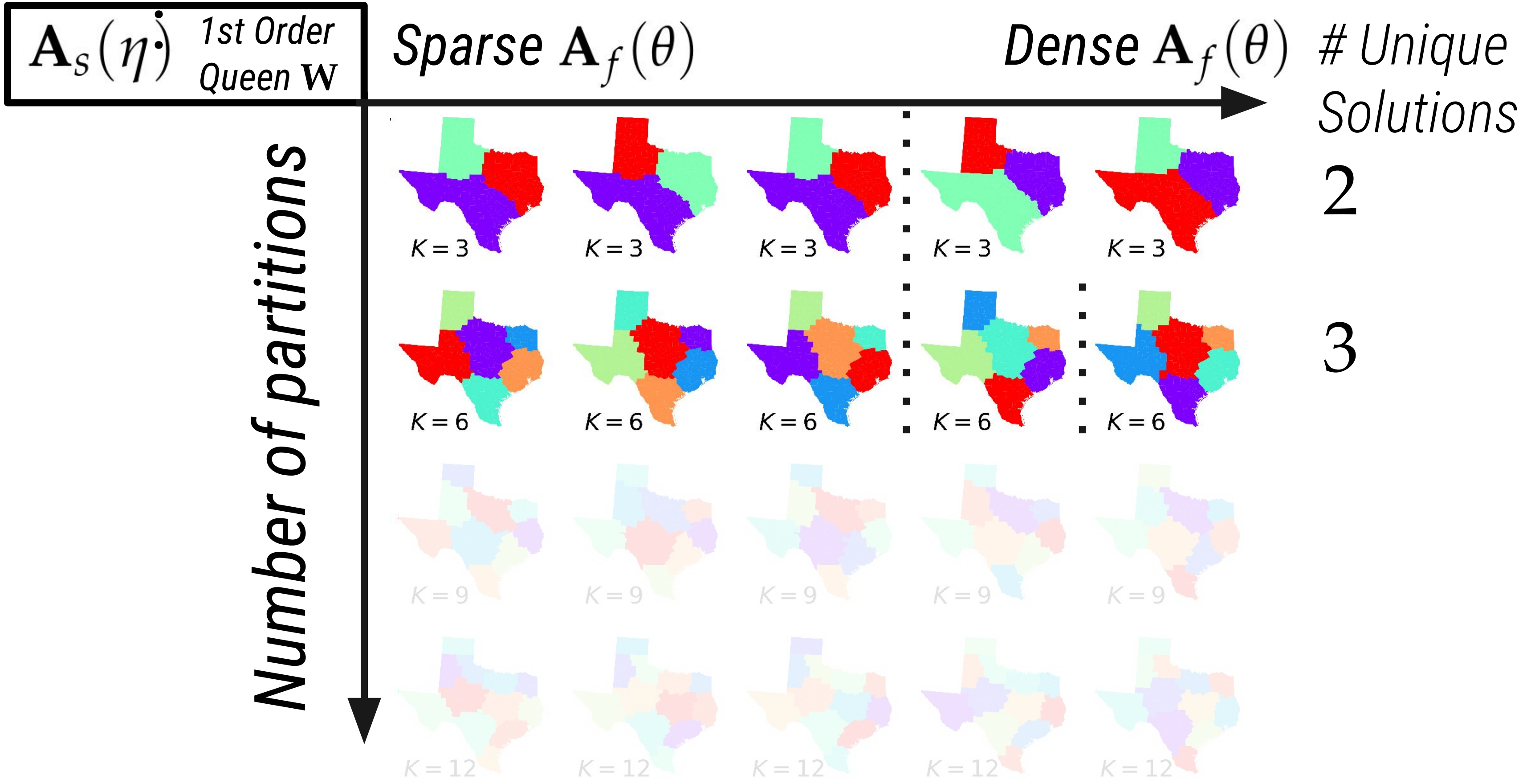
Mixing Spatial & Attribute Kernels



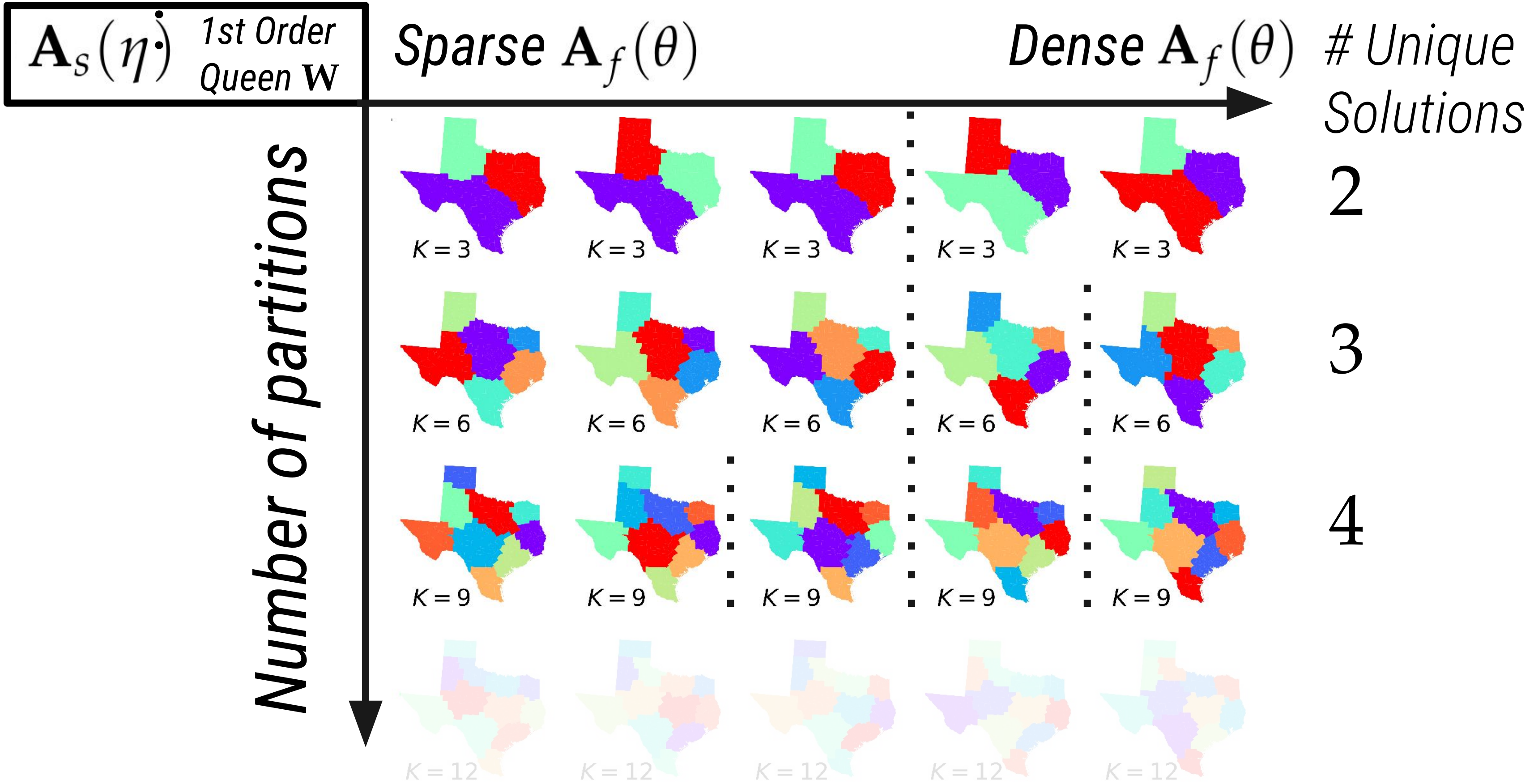
Mixing Spatial & Attribute Kernels



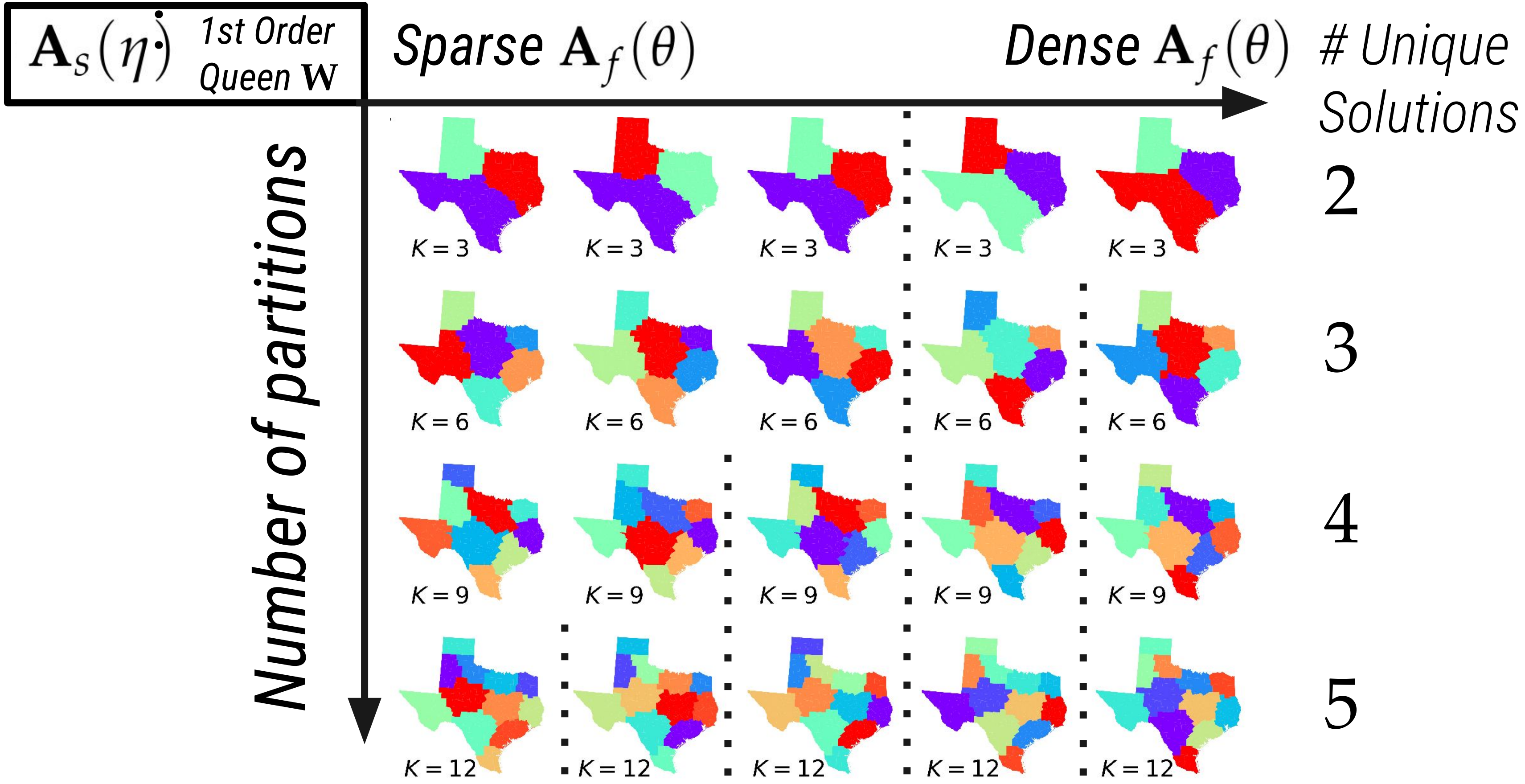
Mixing Spatial & Attribute Kernels



Mixing Spatial & Attribute Kernels



Mixing Spatial & Attribute Kernels



Mixing Spatial & Attribute Kernels

$$\mathbf{L} = \mathbf{D} - \mathbf{A}_f(\theta) \circ \mathbf{A}_s(\eta)$$

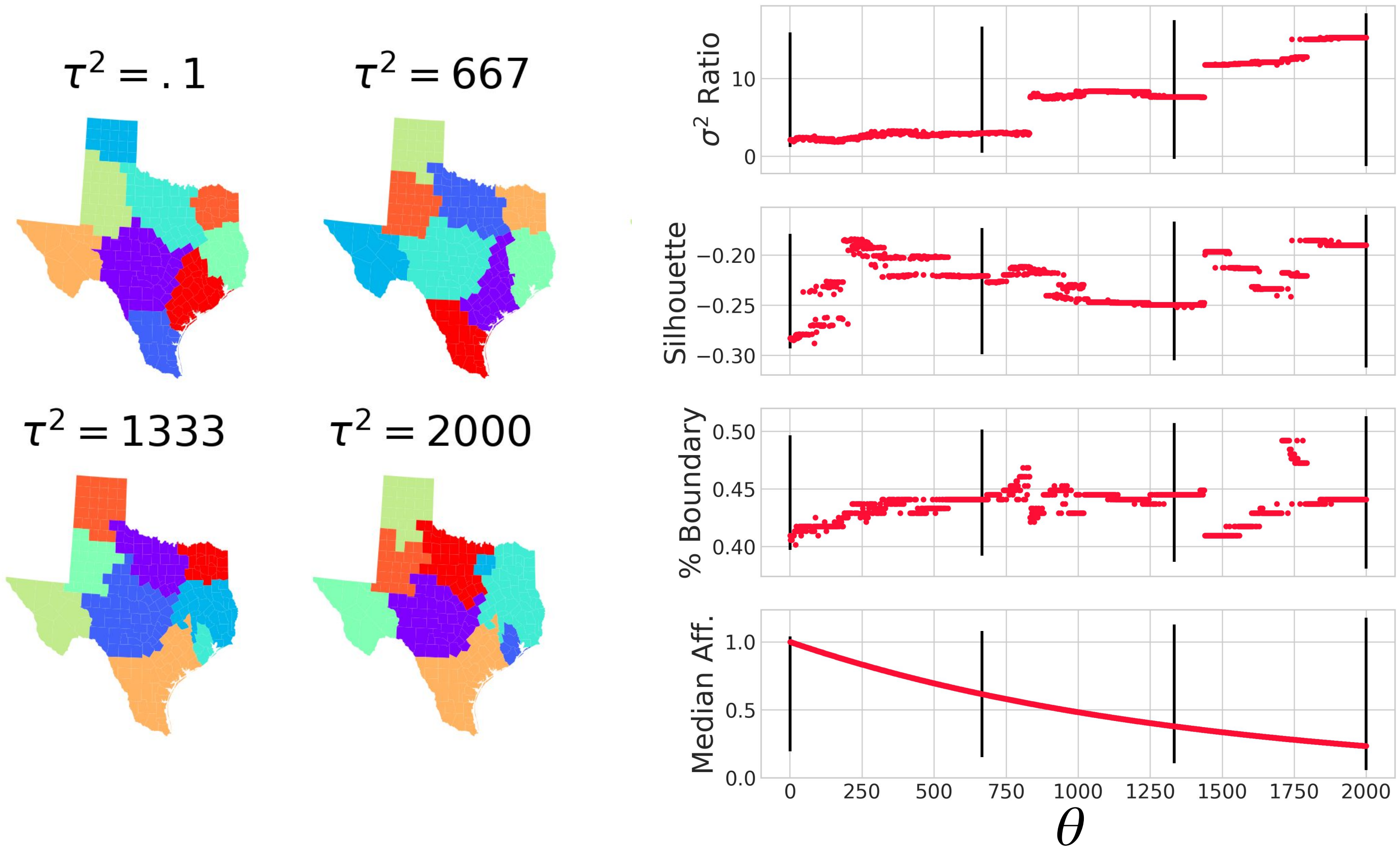
1. **Affinity matrix computed** *for all sites using attribute bandwidth θ to structure the attribute affinity.*
2. **Convolve** *attribute affinity using the spatial kernel filter with width* η
3. **Obtain top k eigenvectors of $\mathbf{L}$** *(approx. OK!)*
4. **Compute k-means clustering** *or recursively partition based on unique values in eigenspace.*
5. **Label** *original data using the eigenvector labels.*

Spatially-Encouraged Spectral Clustering

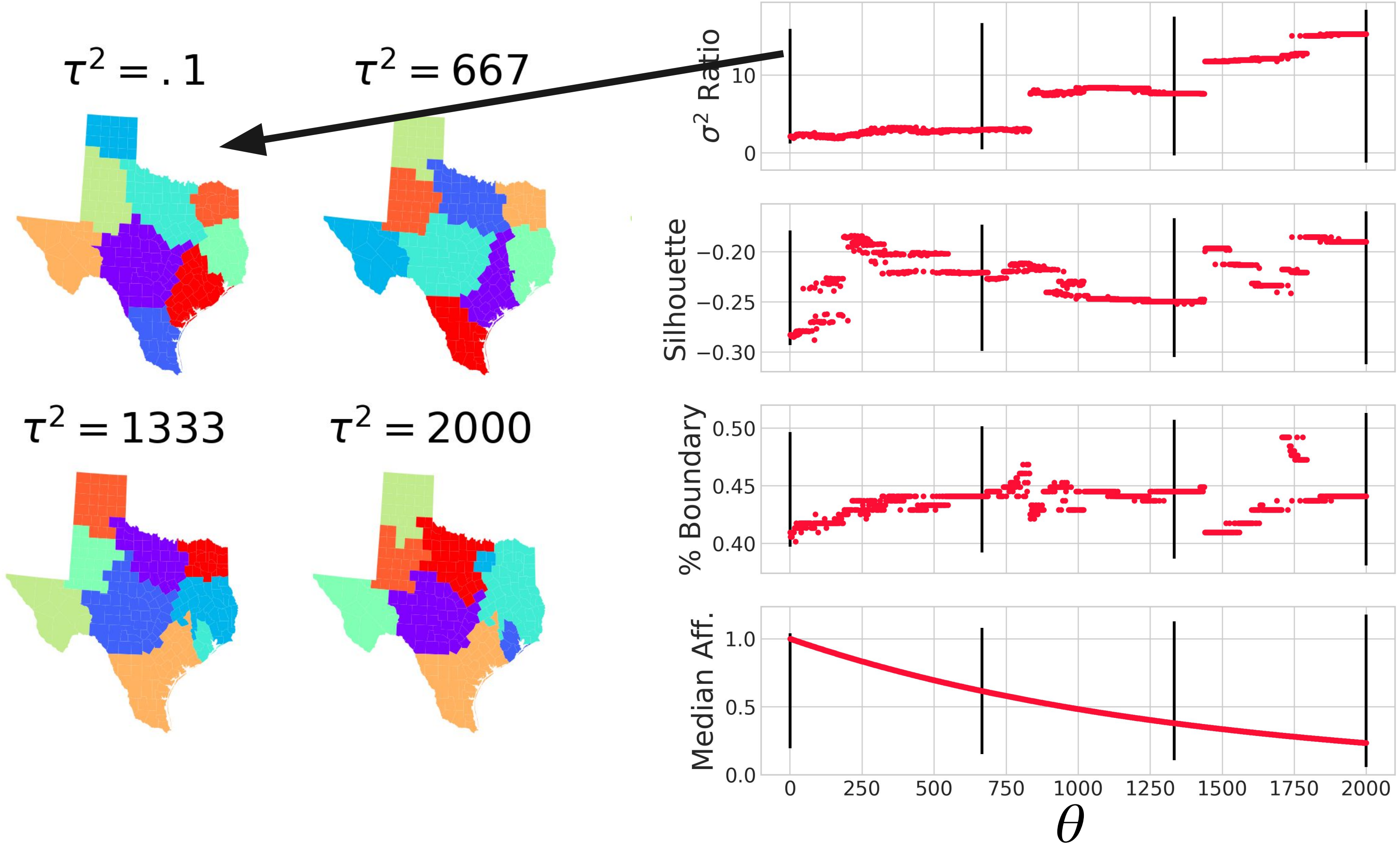
$$\mathbf{L} = \mathbf{D} - \mathbf{A}_f(\theta) \circ \boxed{\mathbf{A}_s(\eta)}$$

1. **Affinity matrix computed** *for all sites using attribute bandwidth θ to structure the attribute affinity.*
2. **Convolve** *attribute affinity using the spatial kernel filter with width η*
3. **Obtain top k eigenvectors of $\mathbf{L}$** *(approx. OK!)*
4. **Compute k-means clustering** *or recursively partition based on unique values in eigenspace.*
5. **Label** *original data using the eigenvector labels.*

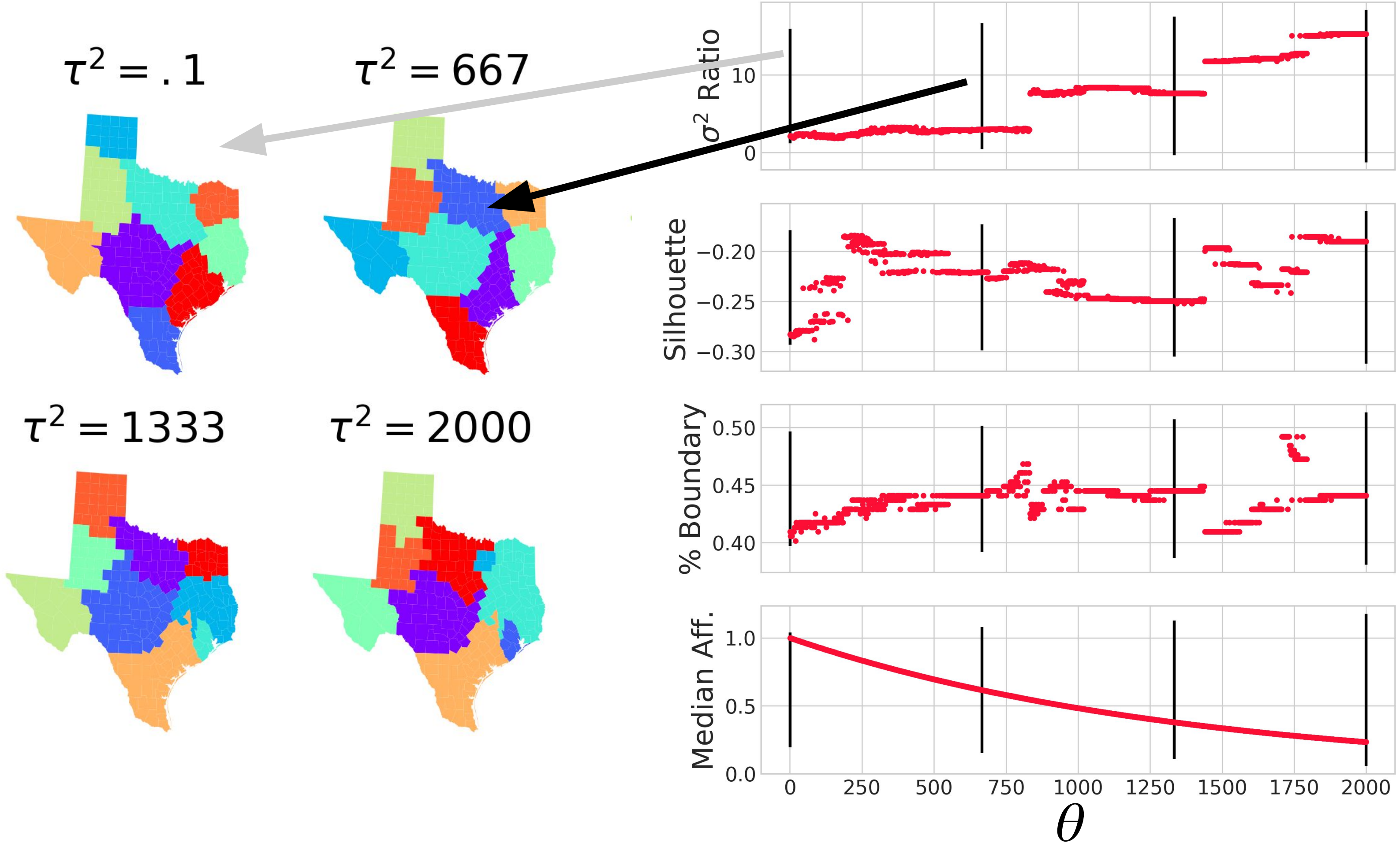
Spatially-Encouraged Spectral Clustering



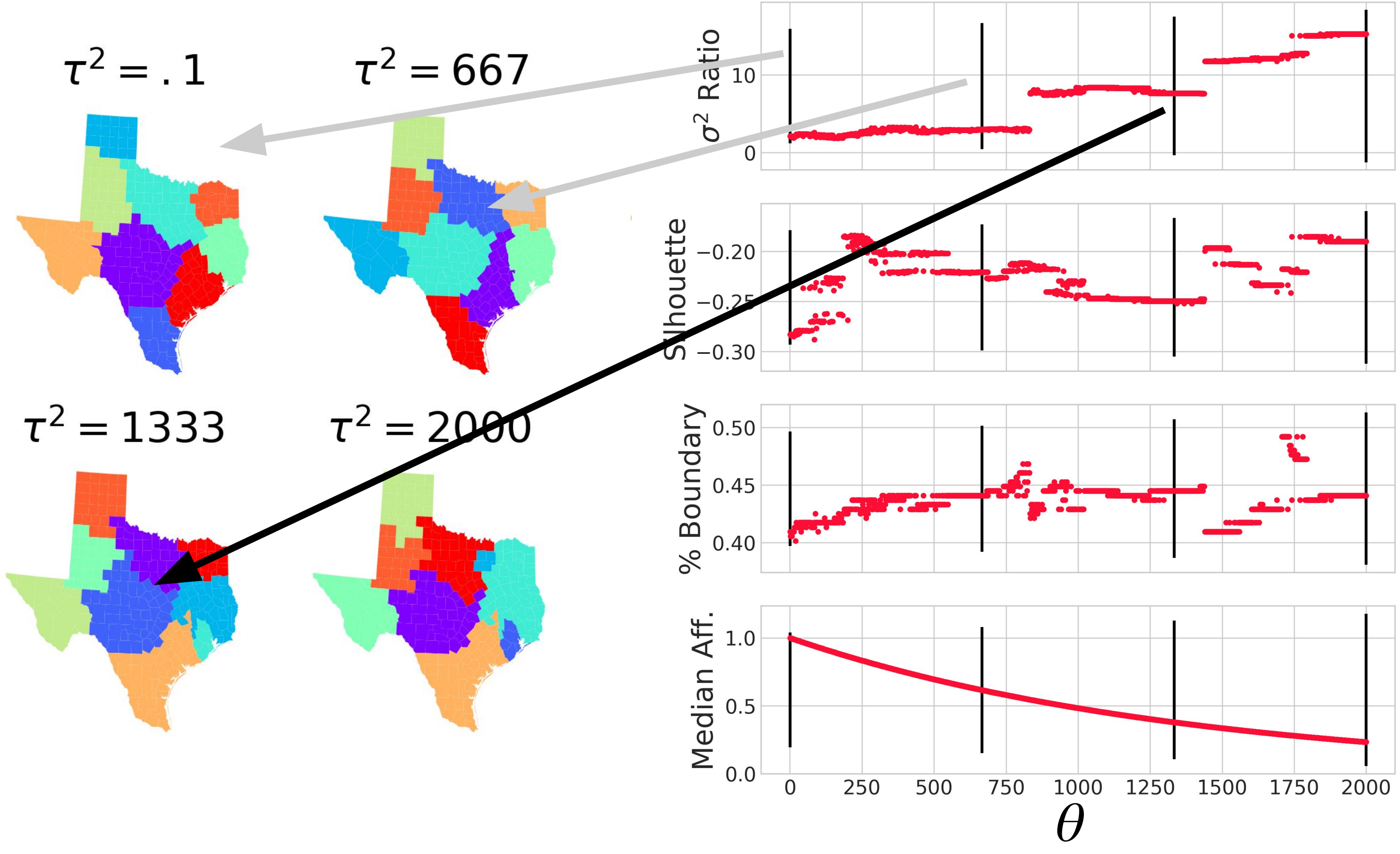
$\tau^2 = .1$
$\tau^2 = 667$
$\tau^2 = 1333$
$\tau^2 = 2000$
σ^2 Ratio
Silhouette
% Boundary
Median Aff.
θ



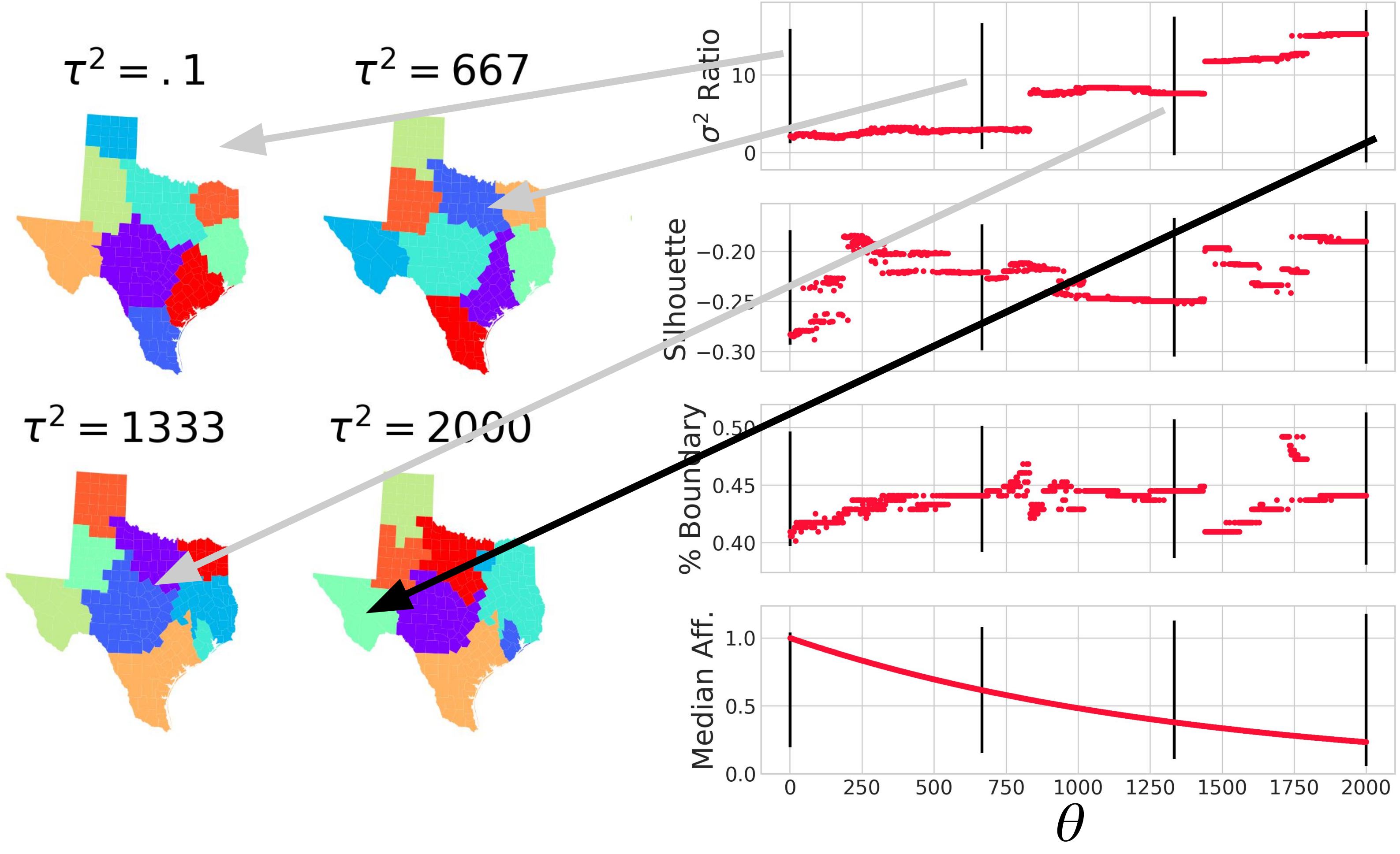
$\tau^2 = .1$
$\tau^2 = 667$
$\tau^2 = 1333$
$\tau^2 = 2000$
σ^2 Ratio
Silhouette
% Boundary
Median Aff.
θ



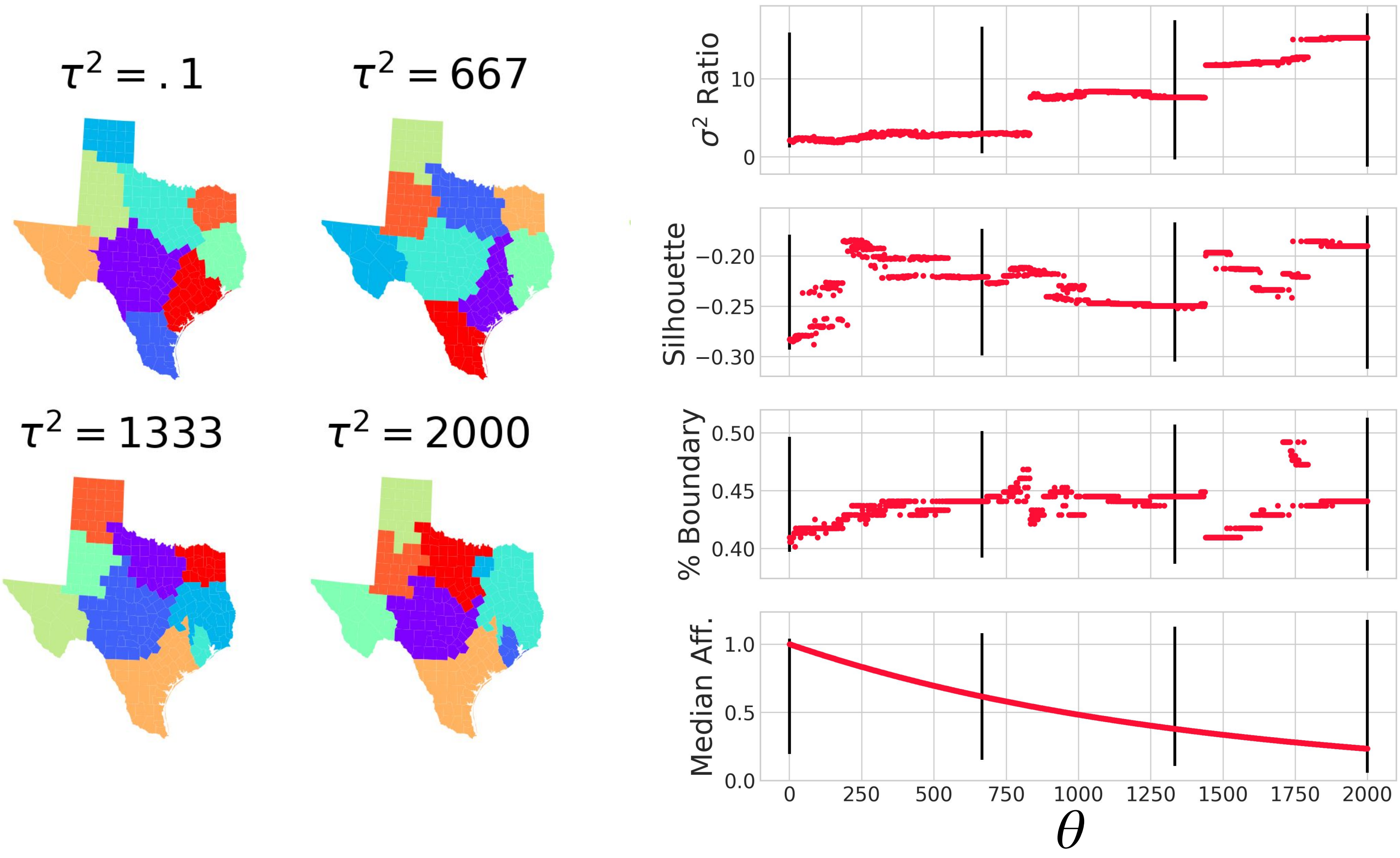
$\tau^2 = .1$
$\tau^2 = 667$
$\tau^2 = 1333$
$\tau^2 = 2000$
σ^2 Ratio
Silhouette
% Boundary
Median Aff.
0
10
−0.20
−0.25
−0.30
0.50
0.45
0.40
1.0
0.5
0.0
0
250
500
750
1000
1250
1500
1750
2000
θ



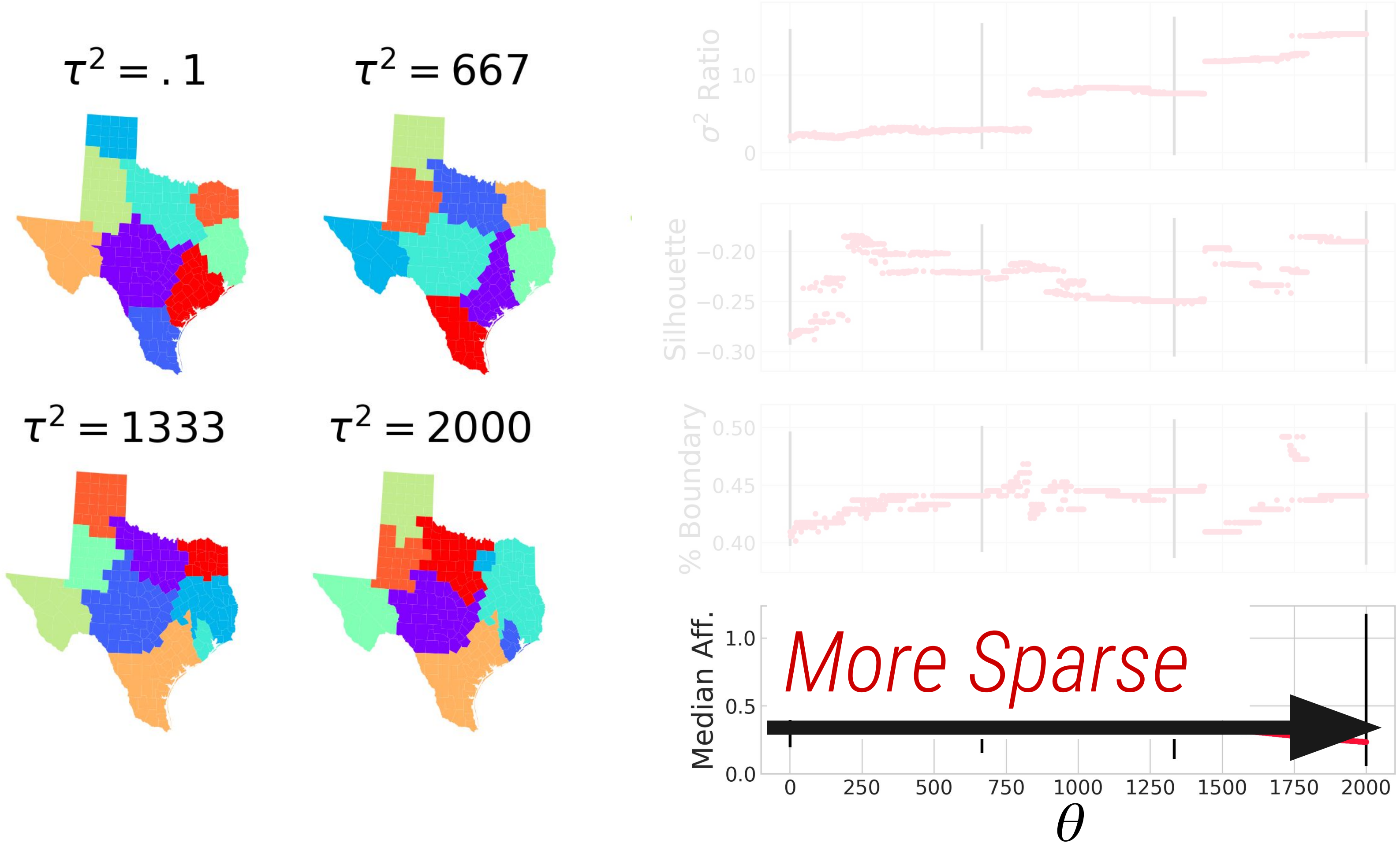
$\tau^2 = .1$
$\tau^2 = 667$
$\tau^2 = 1333$
$\tau^2 = 2000$
σ^2 Ratio
10
0
Silhouette
−0.20
−0.25
−0.30
% Boundary
0.50
0.45
0.40
Median Aff.
1.0
0.5
0.0
0
250
500
750
1000
1250
1500
1750
2000
θ



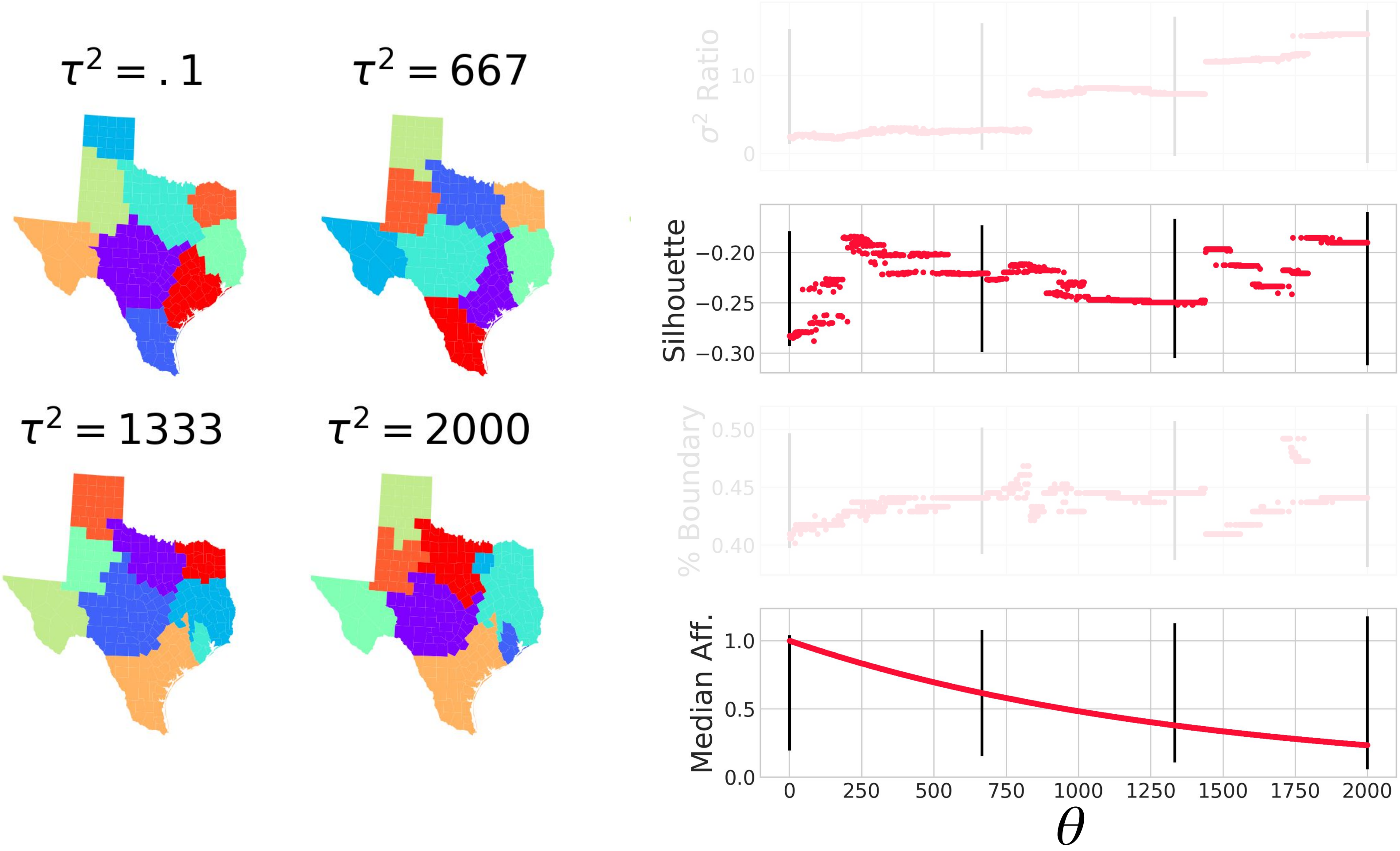
$\tau^2 = .1$
$\tau^2 = 667$
$\tau^2 = 1333$
$\tau^2 = 2000$
σ^2 Ratio
Silhouette
% Boundary
Median Aff.
θ
0
10
−0.20
−0.25
−0.30
0.40
0.45
0.50
0.0
0.5
1.0
0
250
500
750
1000
1250
1500
1750
2000



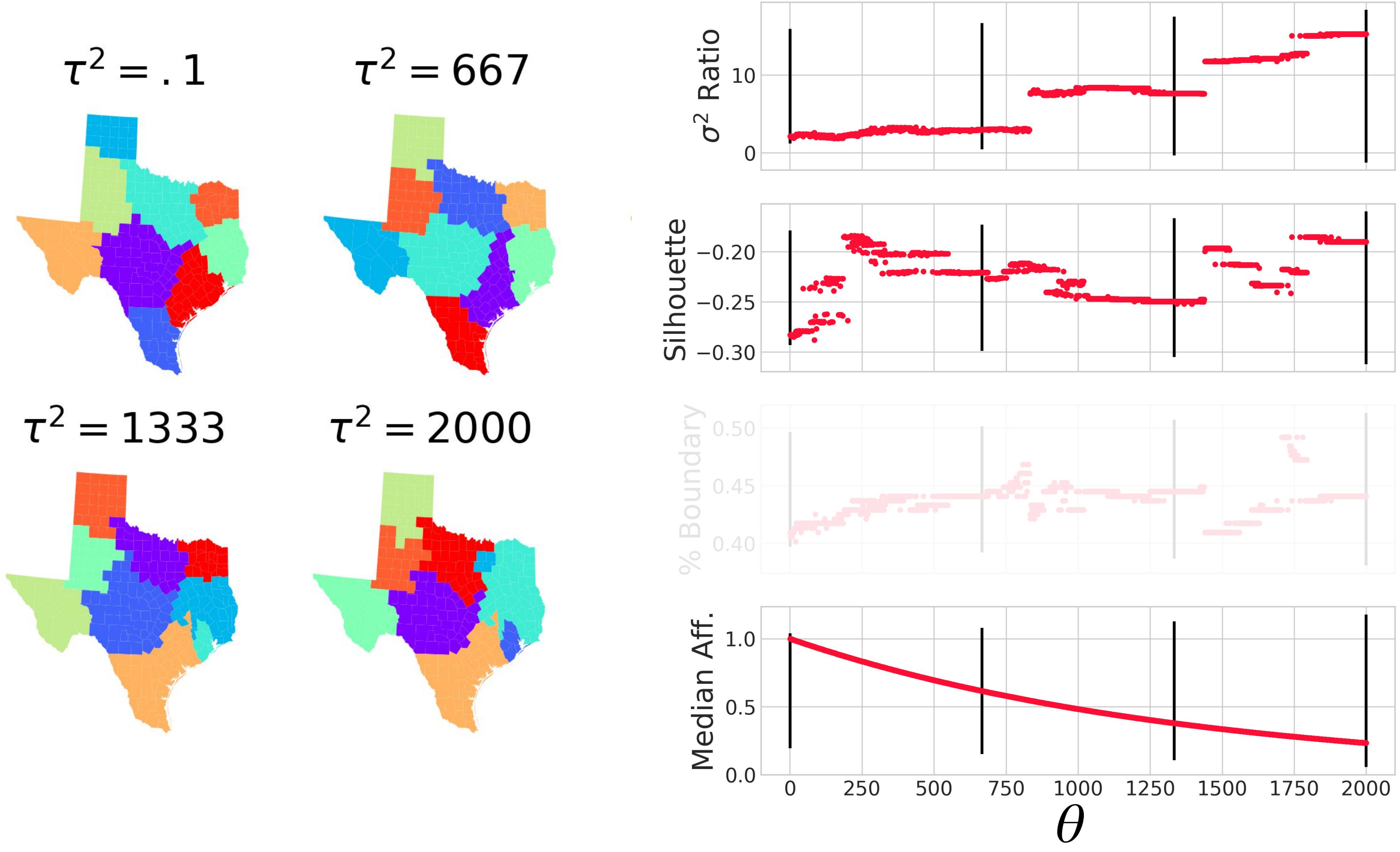
$\tau^2 = .1$
$\tau^2 = 667$
$\tau^2 = 1333$
$\tau^2 = 2000$
σ^2 Ratio
Silhouette
% Boundary
Median Aff.
0
10
−0.20
−0.25
−0.30
0.50
0.45
0.40
1.0
0.5
0.0
0
250
500
750
1000
1250
1500
1750
2000
θ



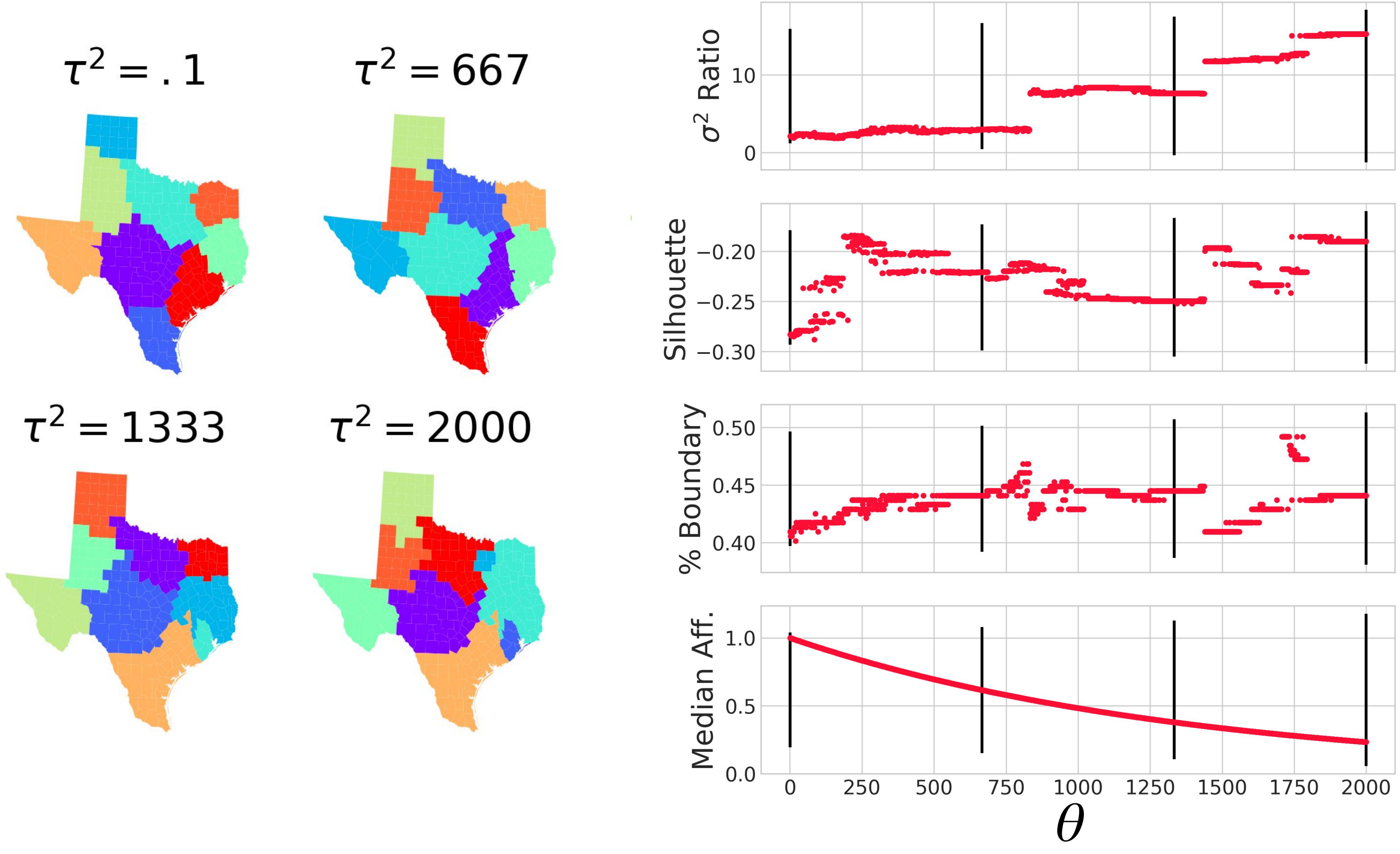
$\tau^2 = .1$
$\tau^2 = 667$
$\tau^2 = 1333$
$\tau^2 = 2000$
σ^2 Ratio
Silhouette
% Boundary
Median Aff.
More Sparse
θ



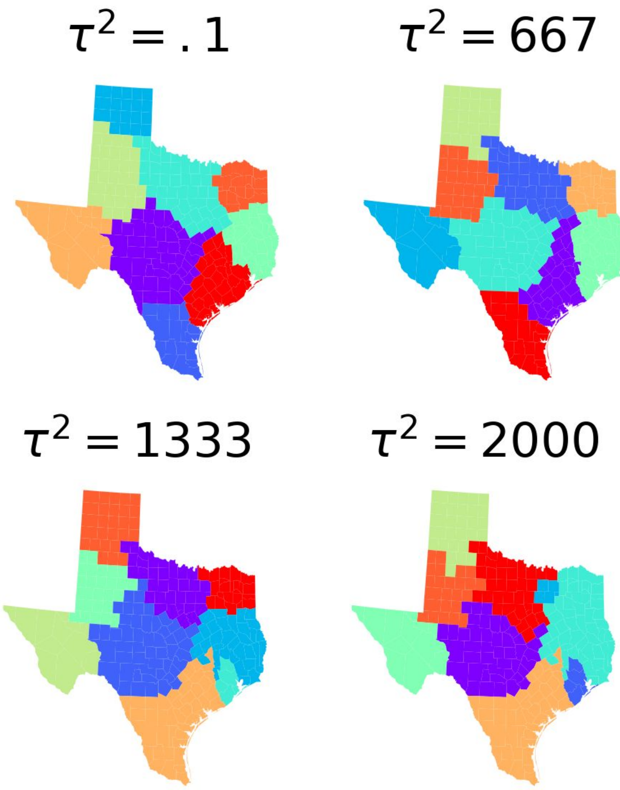
$\tau^2 = .1$
$\tau^2 = 667$
$\tau^2 = 1333$
$\tau^2 = 2000$
σ^2 Ratio
0
10
Silhouette
−0.20
−0.25
−0.30
% Boundary
0.50
0.45
0.40
Median Aff.
1.0
0.5
0.0
0
250
500
750
1000
1250
1500
1750
2000
θ

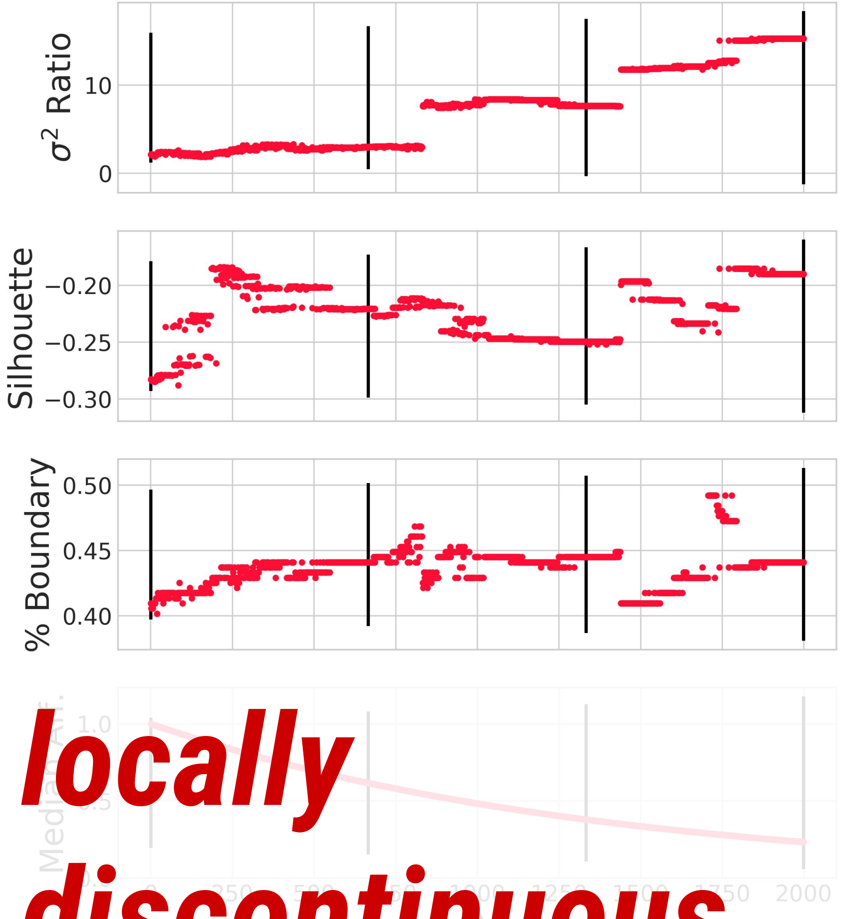


$\tau^2 = .1$
$\tau^2 = 667$
$\tau^2 = 1333$
$\tau^2 = 2000$
σ^2 Ratio
0
10
Silhouette
−0.30
−0.25
−0.20
% Boundary
0.40
0.45
0.50
Median Aff.
0.0
0.5
1.0
0
250
500
750
1000
1250
1500
1750
2000
θ

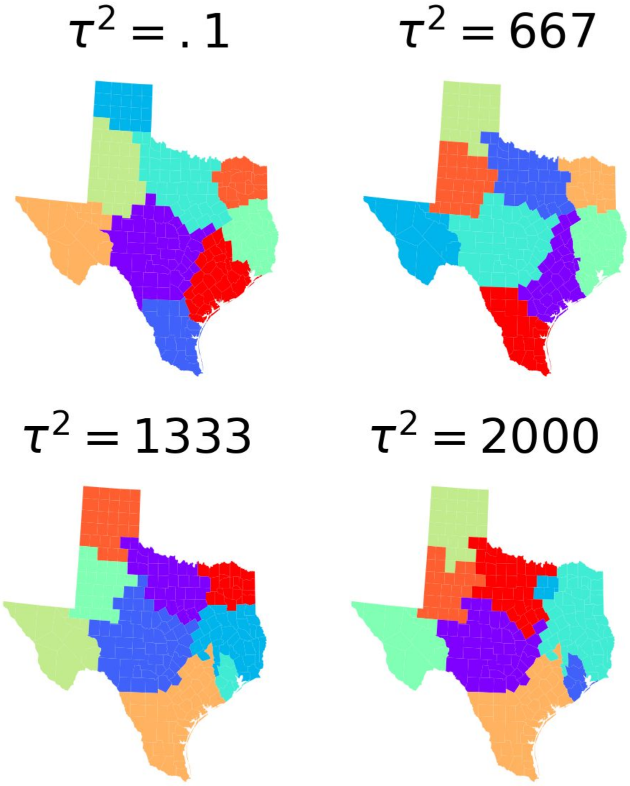


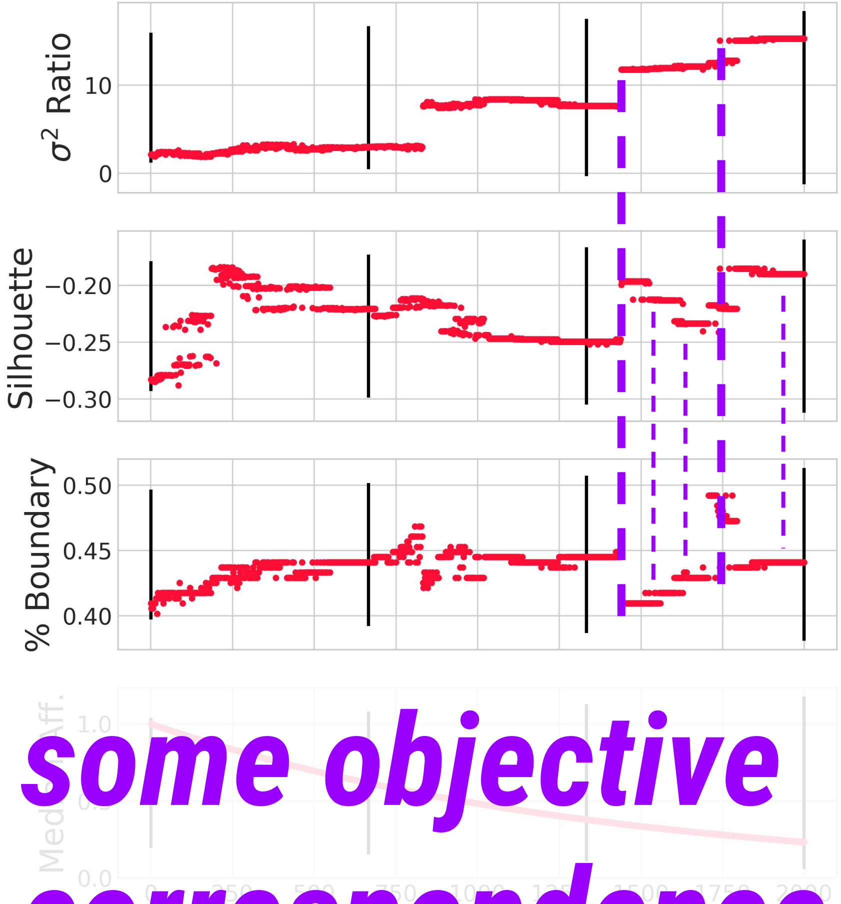
$\tau^2 = .1$
$\tau^2 = 667$
$\tau^2 = 1333$
$\tau^2 = 2000$
σ^2 Ratio
Silhouette
% Boundary
Median Aff.
0
10
−0.20
−0.25
−0.30
0.50
0.45
0.40
1.0
0.5
0.0
0
250
500
750
1000
1250
1500
1750
2000
θ



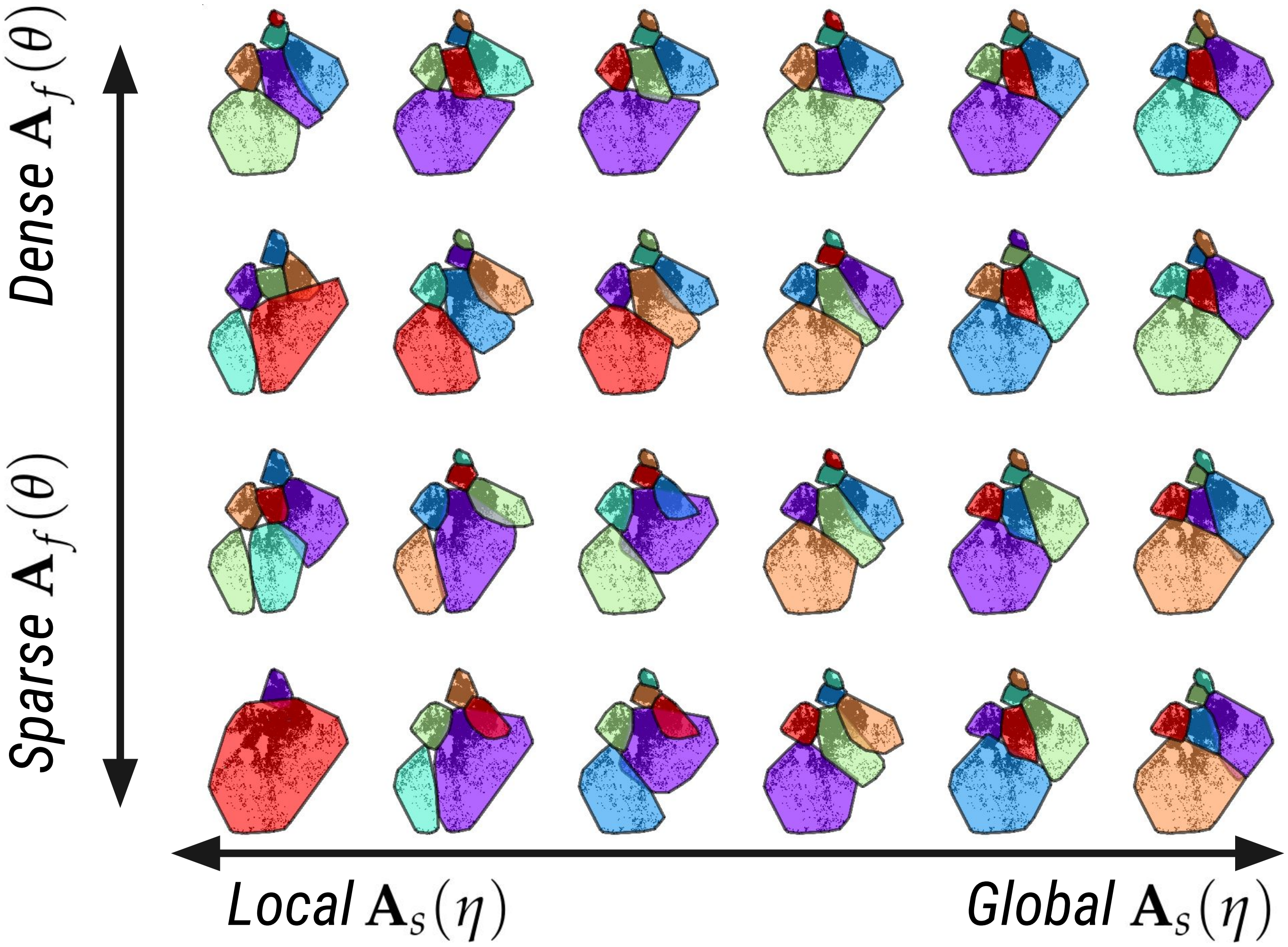


locally discontinuous



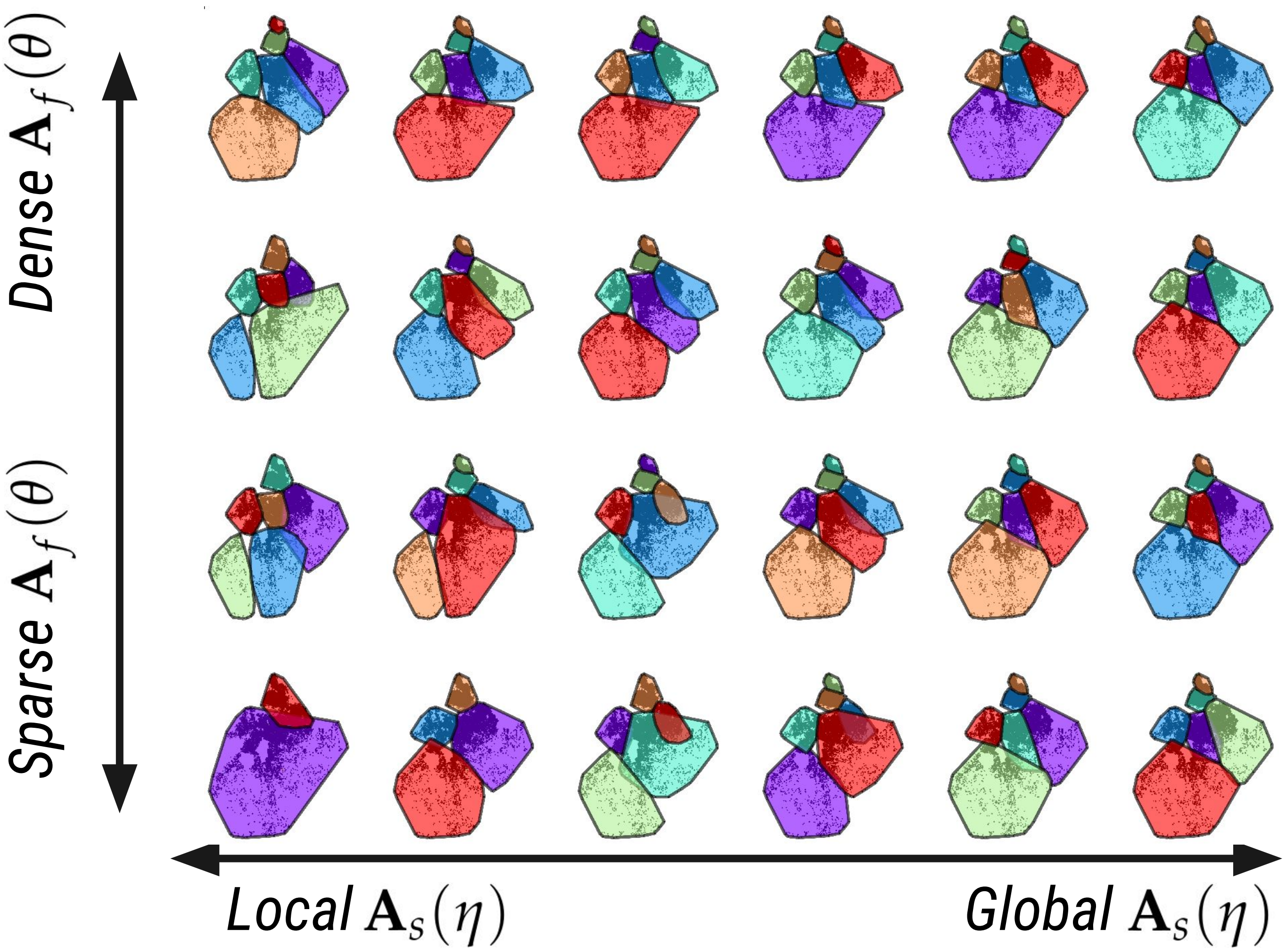


some objective correspondence



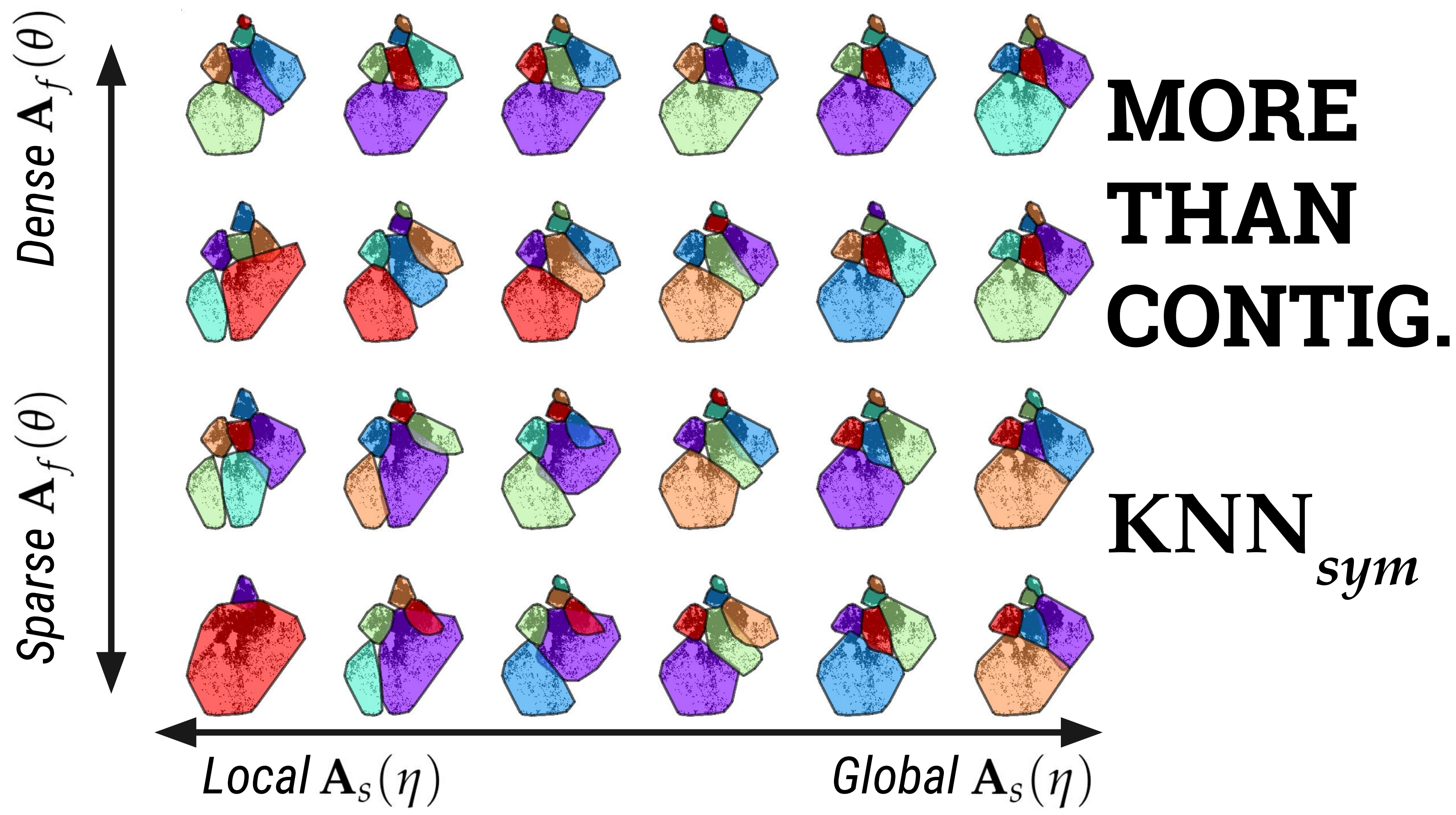
MORE THAN CONTIG.

Prices, KNN_{sym}



MORE THAN CONTIG.

Prices,
N-Kernel


Dense $\mathbf{A}_f(\theta)$
Sparse $\mathbf{A}_f(\theta)$
Local $\mathbf{A}_s(\eta)$
Global $\mathbf{A}_s(\eta)$
MORE THAN CONTIG.
KNN_{sym}

***Sp**atially-**En**couraged Spectral **C**lustering is a spectral clustering method using convolutions of spatial & attribute kernels.*

*Solutions from **SPENC** balance using both spatial & attribute kernel bandwidths.*

*Spatial kernels can be fully general when convolved in **SPENC**.*

Closing Thoughts

QUESTIONS ABOUT

University of BRISTOL

Spatially-Encouraged Spectral Clustering

A Critical Revision of Spatially-Constrained Spectral Clustering

Levi John Wolf

ljwolf.org/post/2018-gisruk