Unsupervised Network Discovery for **Brain Imaging Data**

Zilong **Bai** (University of California, Davis)
LCDR Peter Walker, Anna Tschiffely (Naval Medical Research Center)
Professor Fei Wang (Cornell University)
Professor Ian Davidson (University of California, Davis)

KDD2017

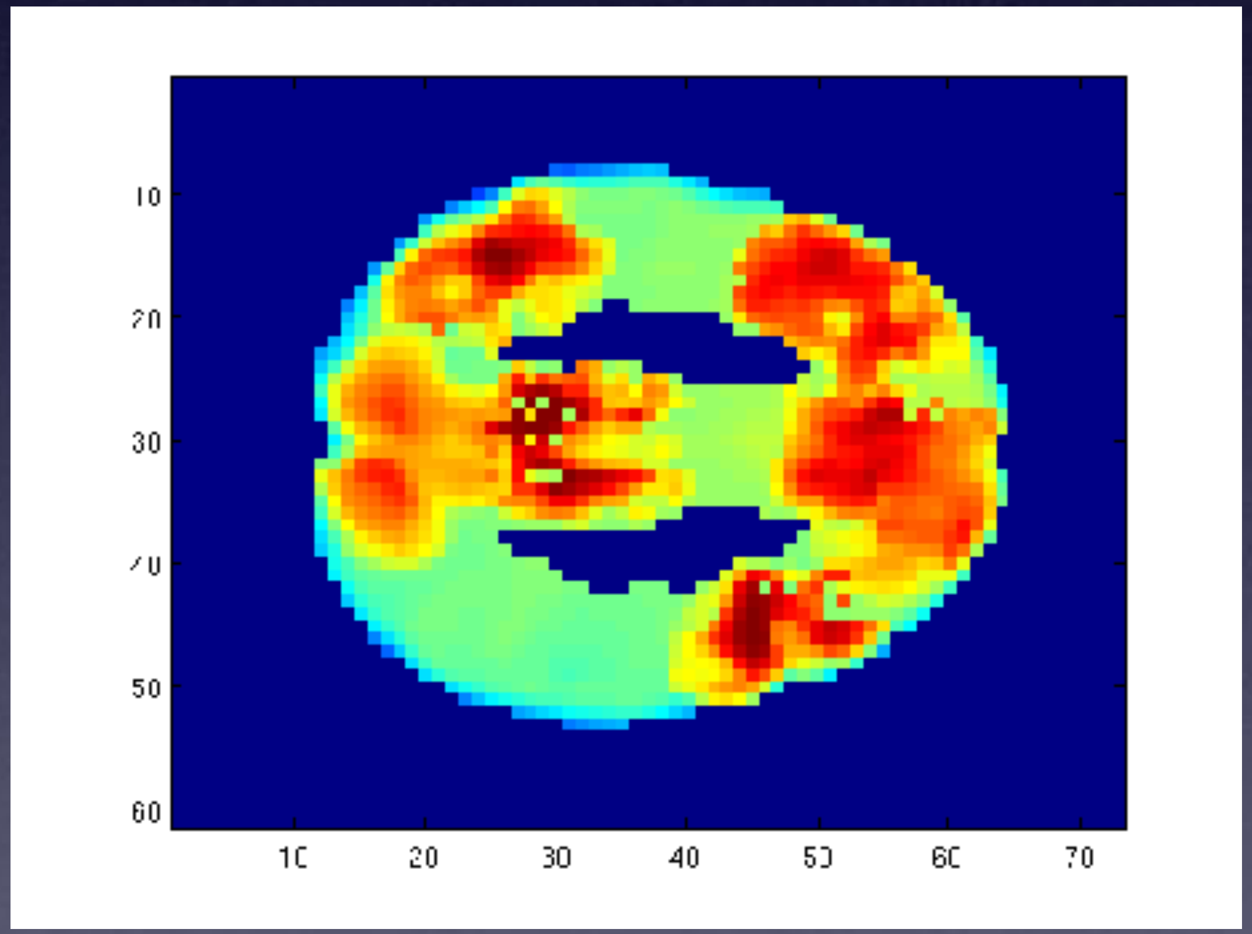
Outline

- Introduction to fMRI
- Define the problem setting
- Previous work & Our Method
- Experiments: Synthetic & Real-world Data
- Future Work

KDD2017

Brain Imaging Data

fMRI



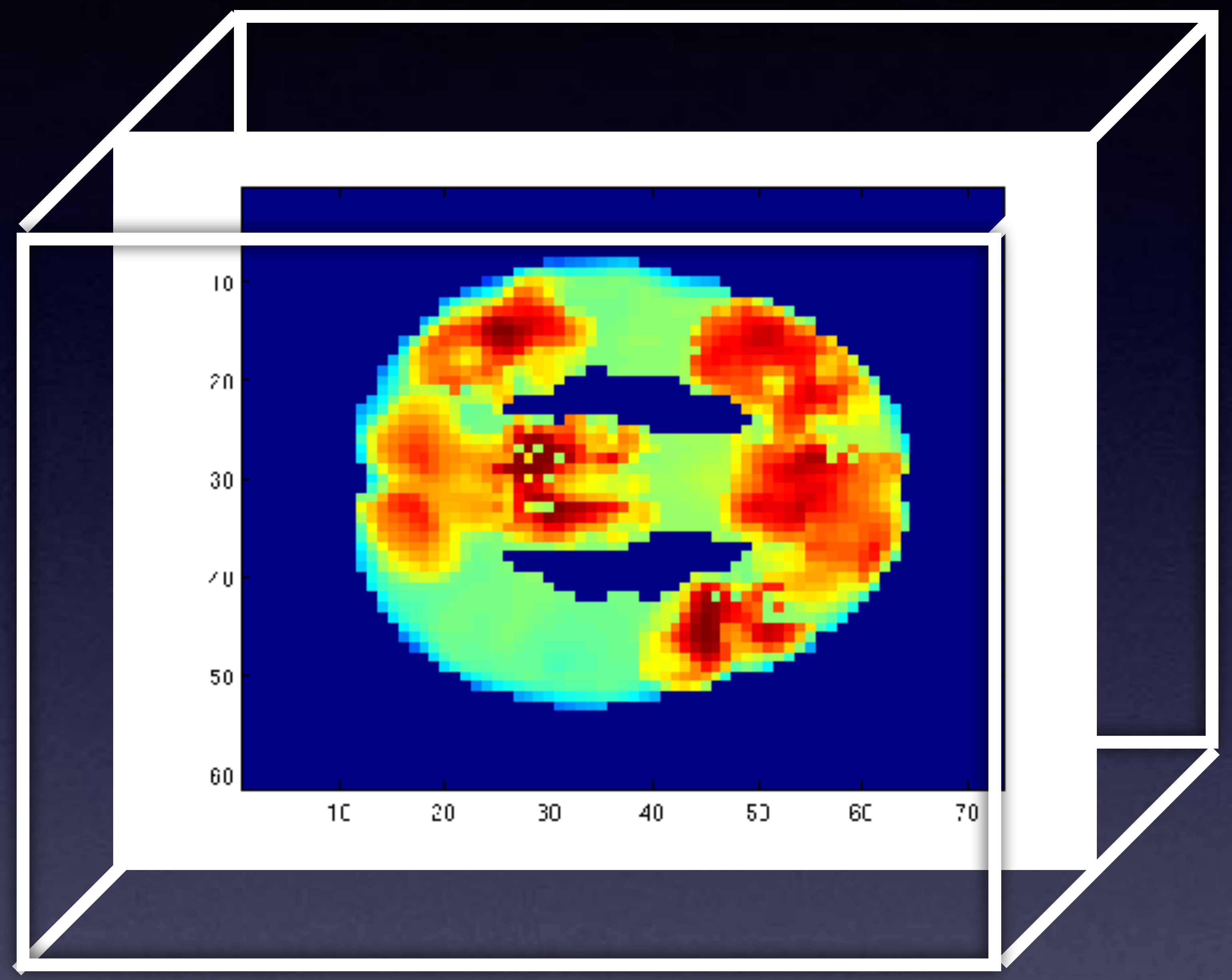
KDD2017

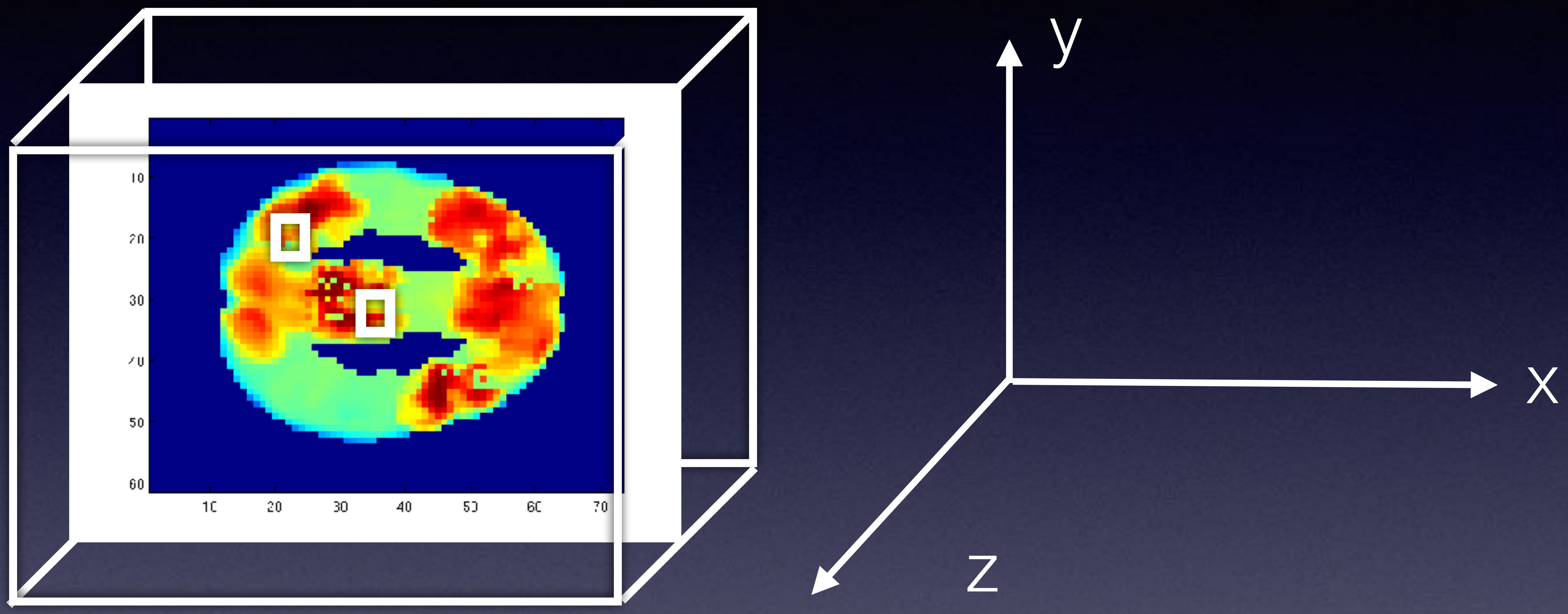
Brain Imaging Data

fMRI

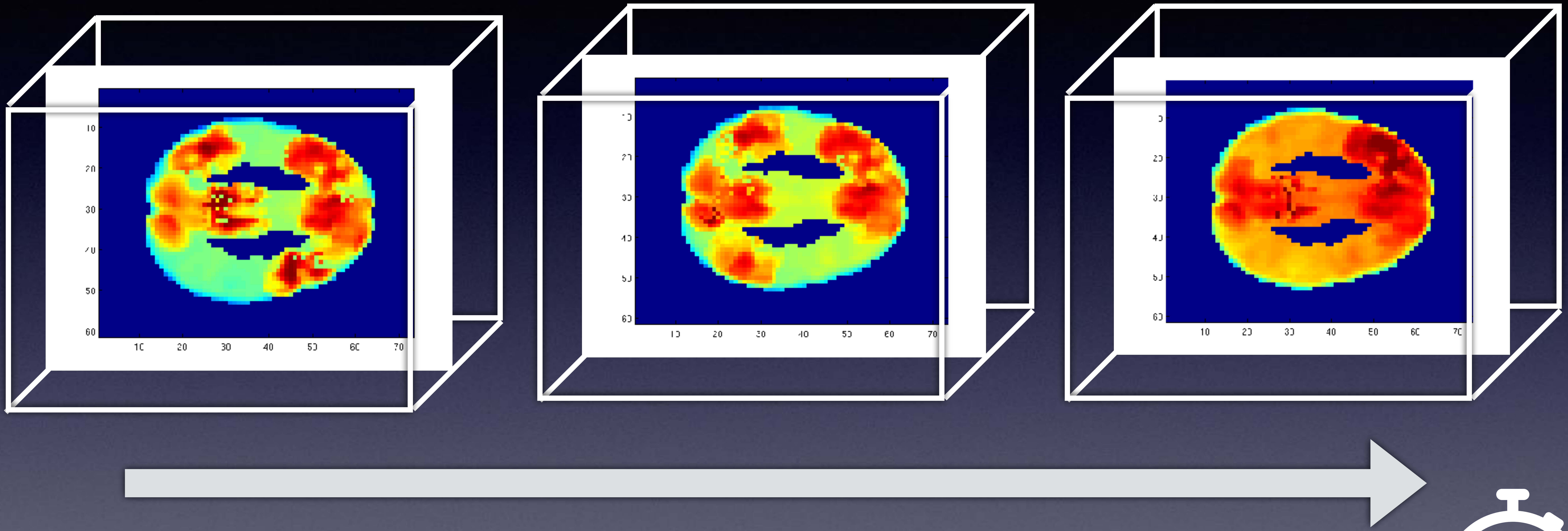
Functional Magnetic Resonance Imaging

- Changes associated with blood flow: “**BOLD**”
- **Non-invasive**




10
20
30
40
50
60
10
20
30
40
50
60
70
y
x
z

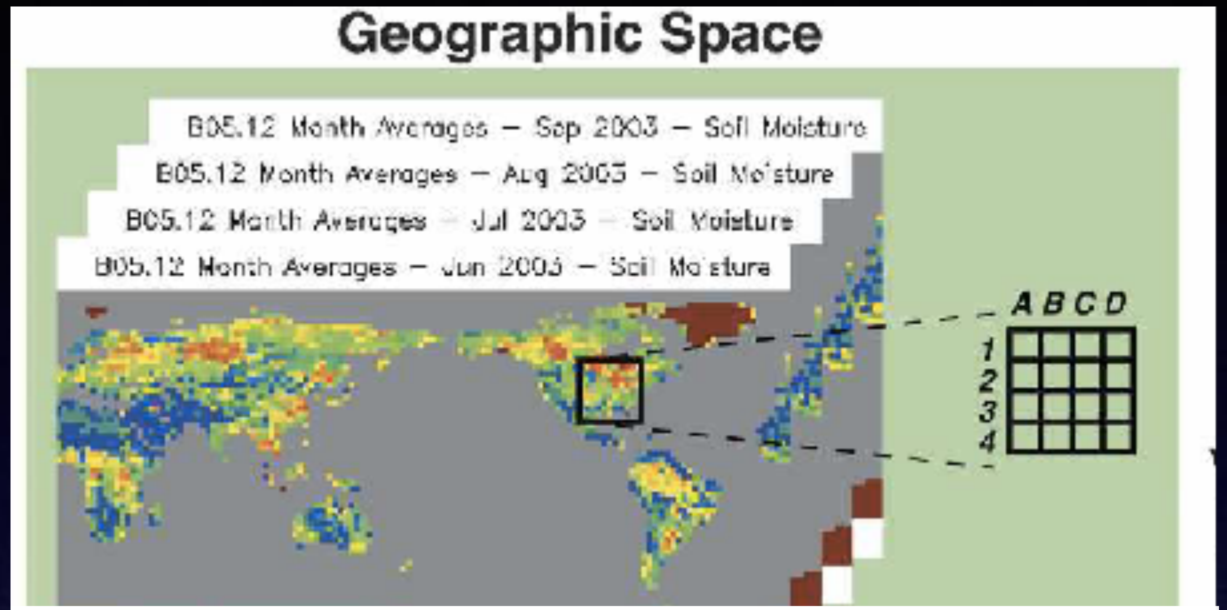
“Raw” scan: 4th-order tensor

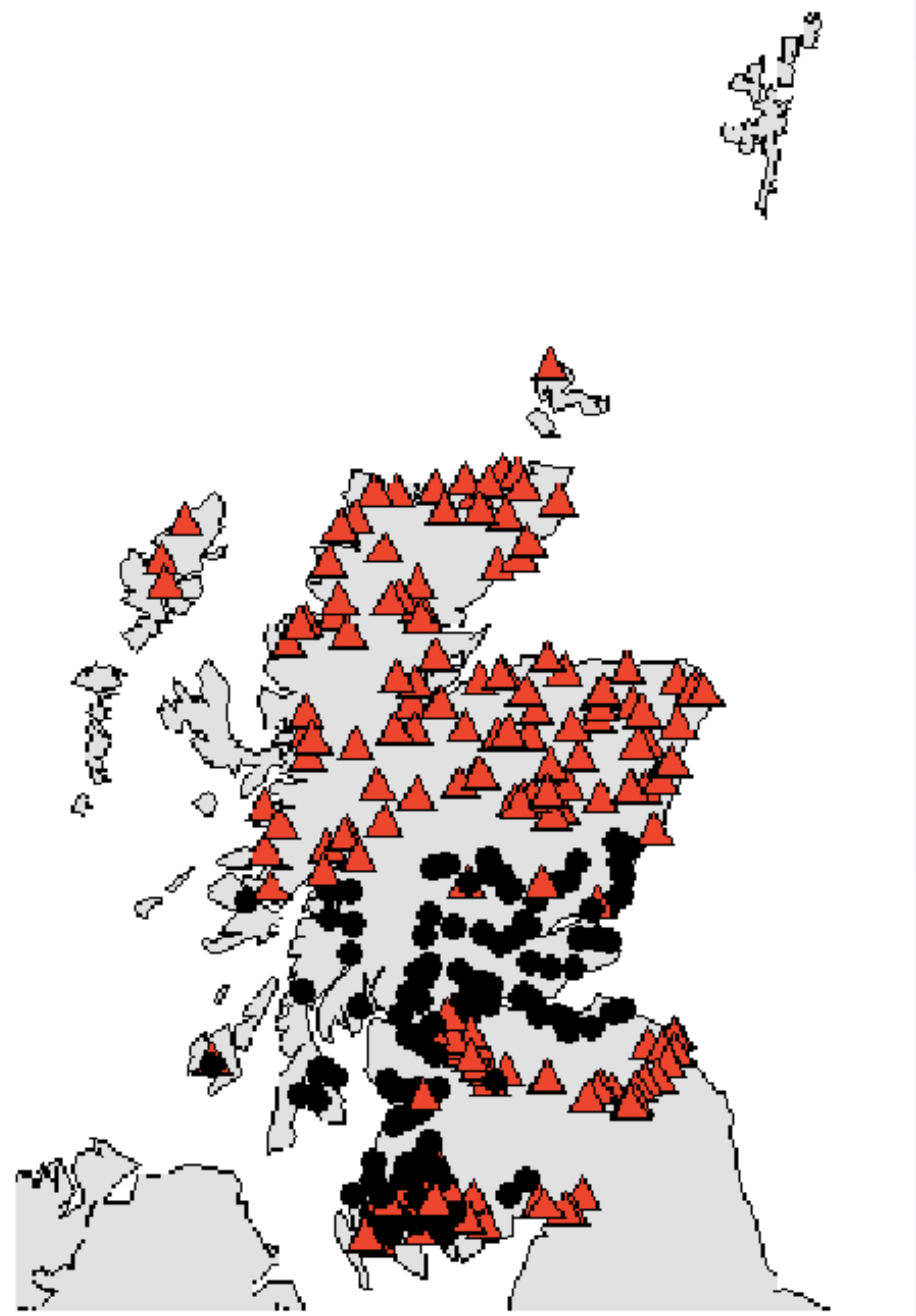


Spatial + Temporal

Spatial + Temporal

Forrest M Hoffman, William W Hargrove Jr, David J Erickson III, and Robert J Oglesby. 2005. Using clustered climate regimes to analyze and compare predictions from fully coupled general circulation models. Earth Interactions 9, 10 (2005), 1–27.



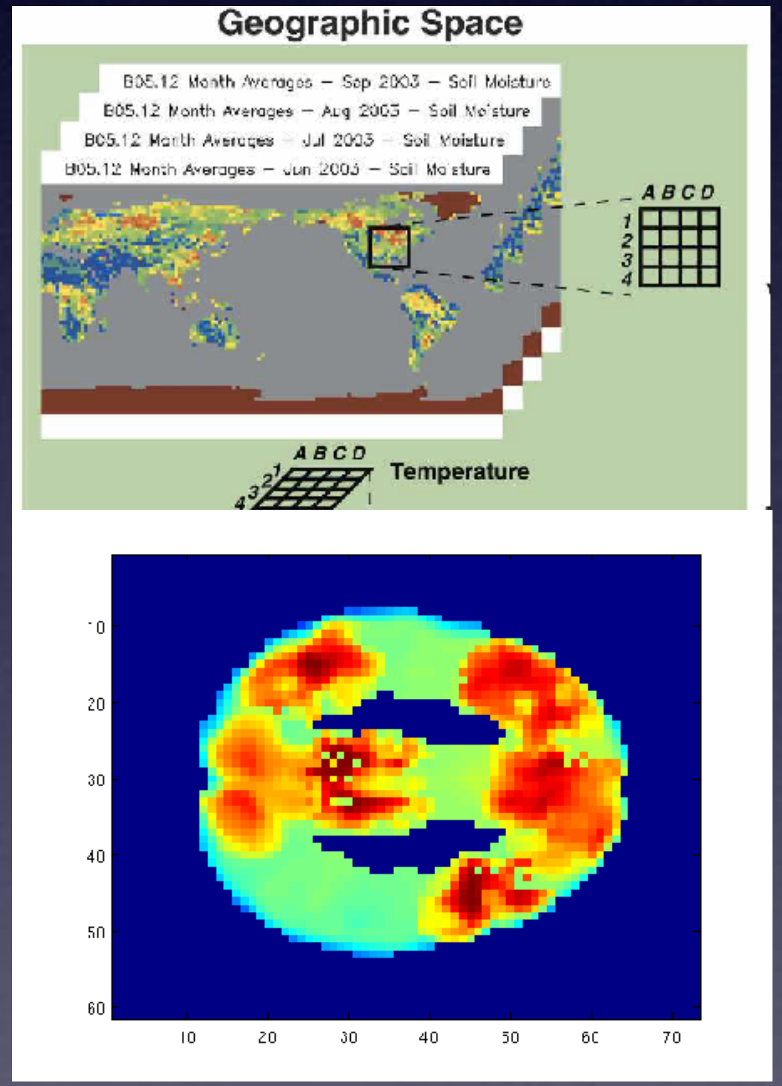


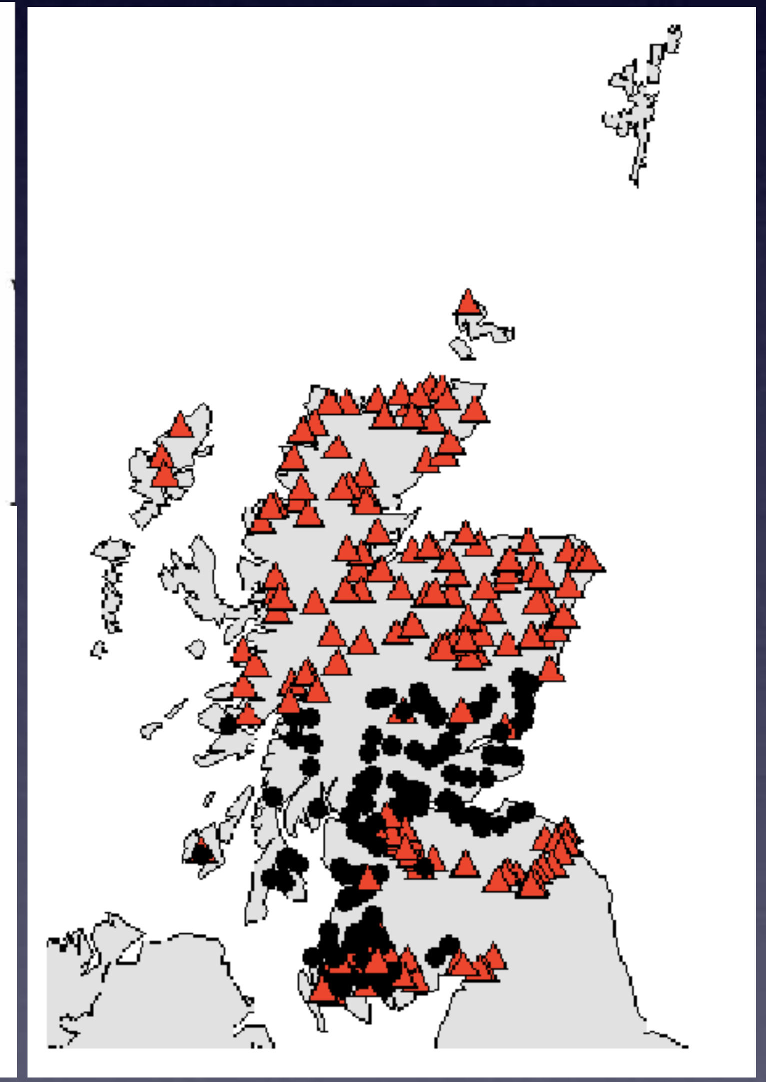
Marian Scott, Claire Miller, Francesco Finazzi, and Ruth Haggarty. 2013. Coherency in space of lake and river temperature and water quality records. (2013).

Marc Postman and P Murdin. 2001. Distribution of Galaxies, Clusters, and Superclusters. Encyclopedia of Astronomy and Astrophysics (2001).

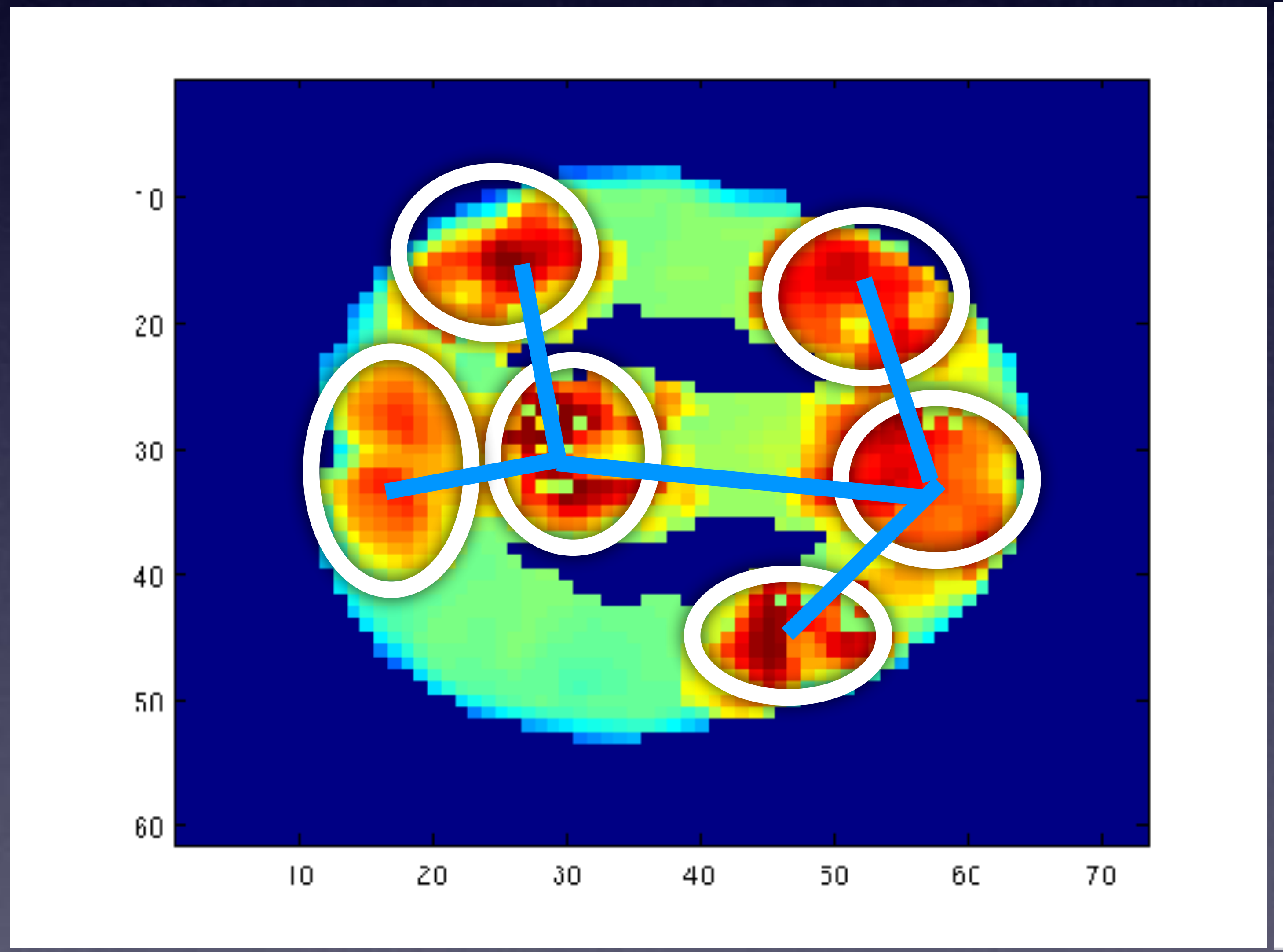
KDD2017

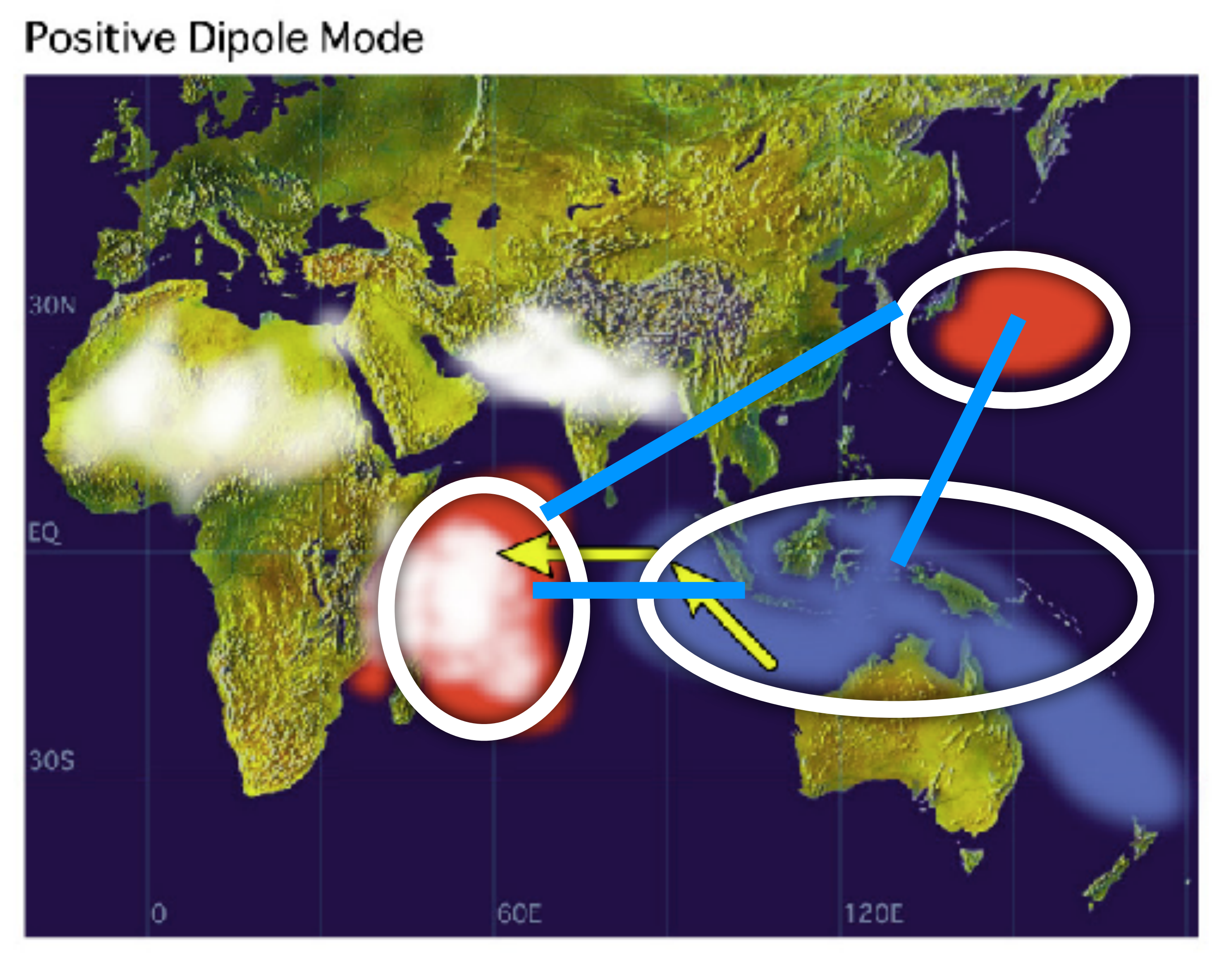
Spatial + Temporal





Network Discovery





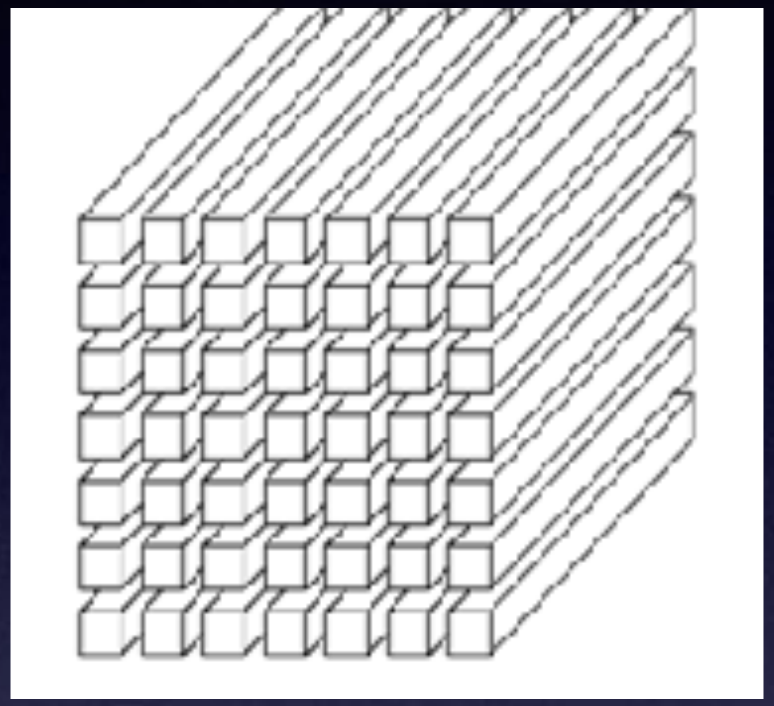
http://www.oceansatlas.org/subtopic/en/c/656/

Outline

- Introduction to fMRI
- **Define the problem setting**
- Previous work & Our Method
- Experiments: Synthetic & Real-world Data
- Future Work

KDD2017

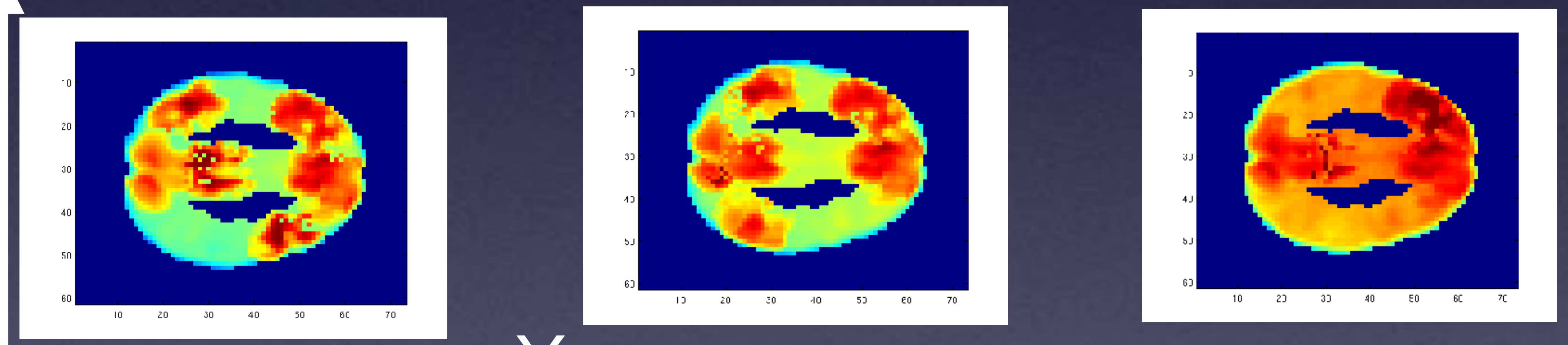
Network Discovery



Spatial + Temporal

36-th Slice

y



x

Graph $\mathbf{X}$

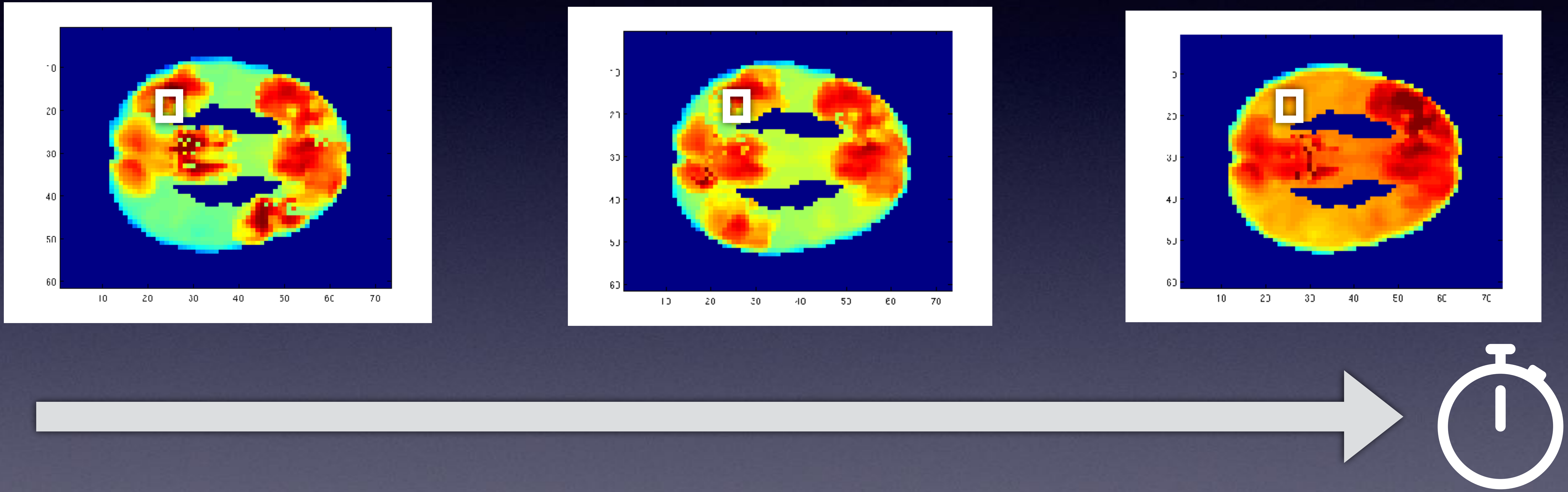
<u>Vertices</u>

Voxels in Spatial Domain

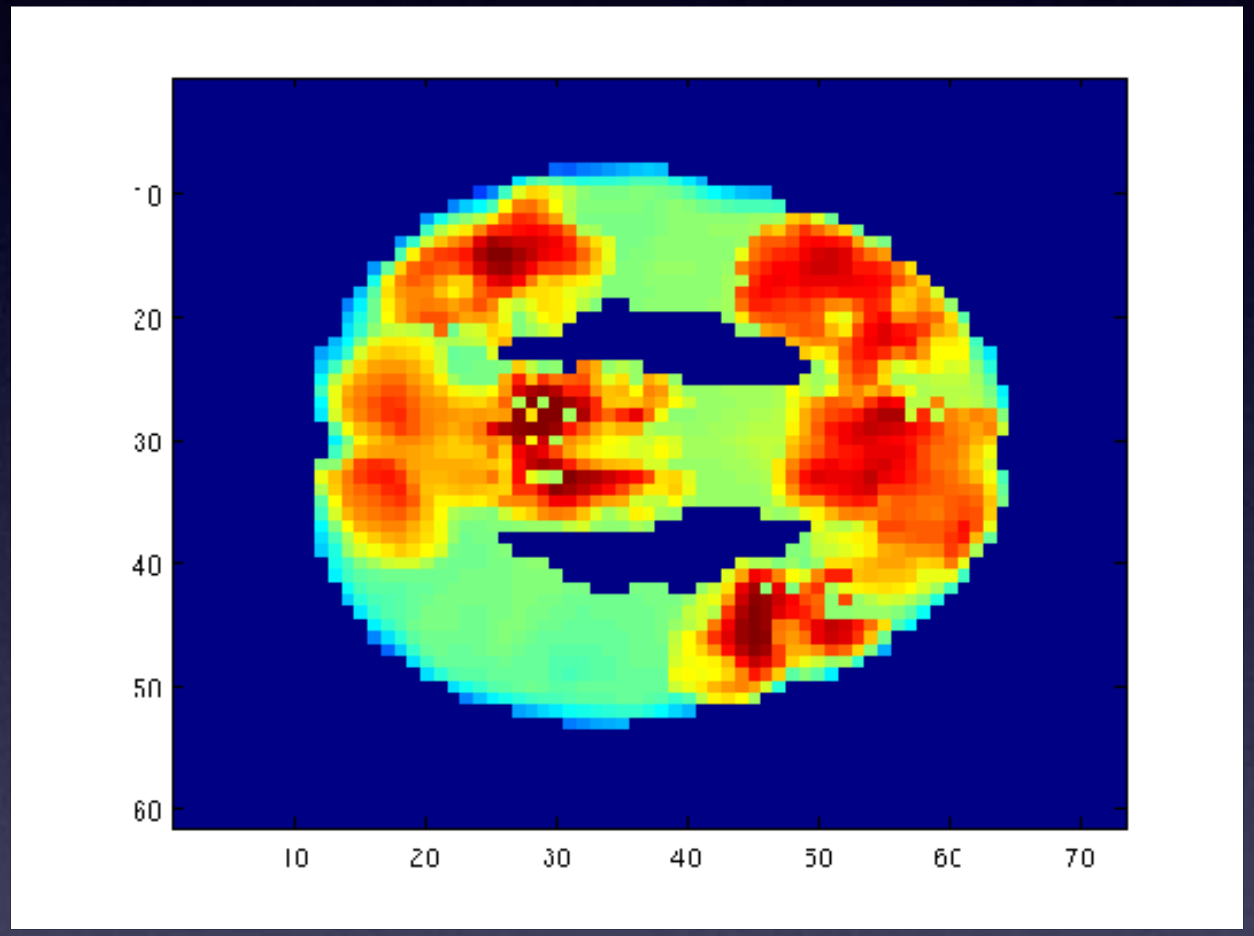
<u>Links</u>

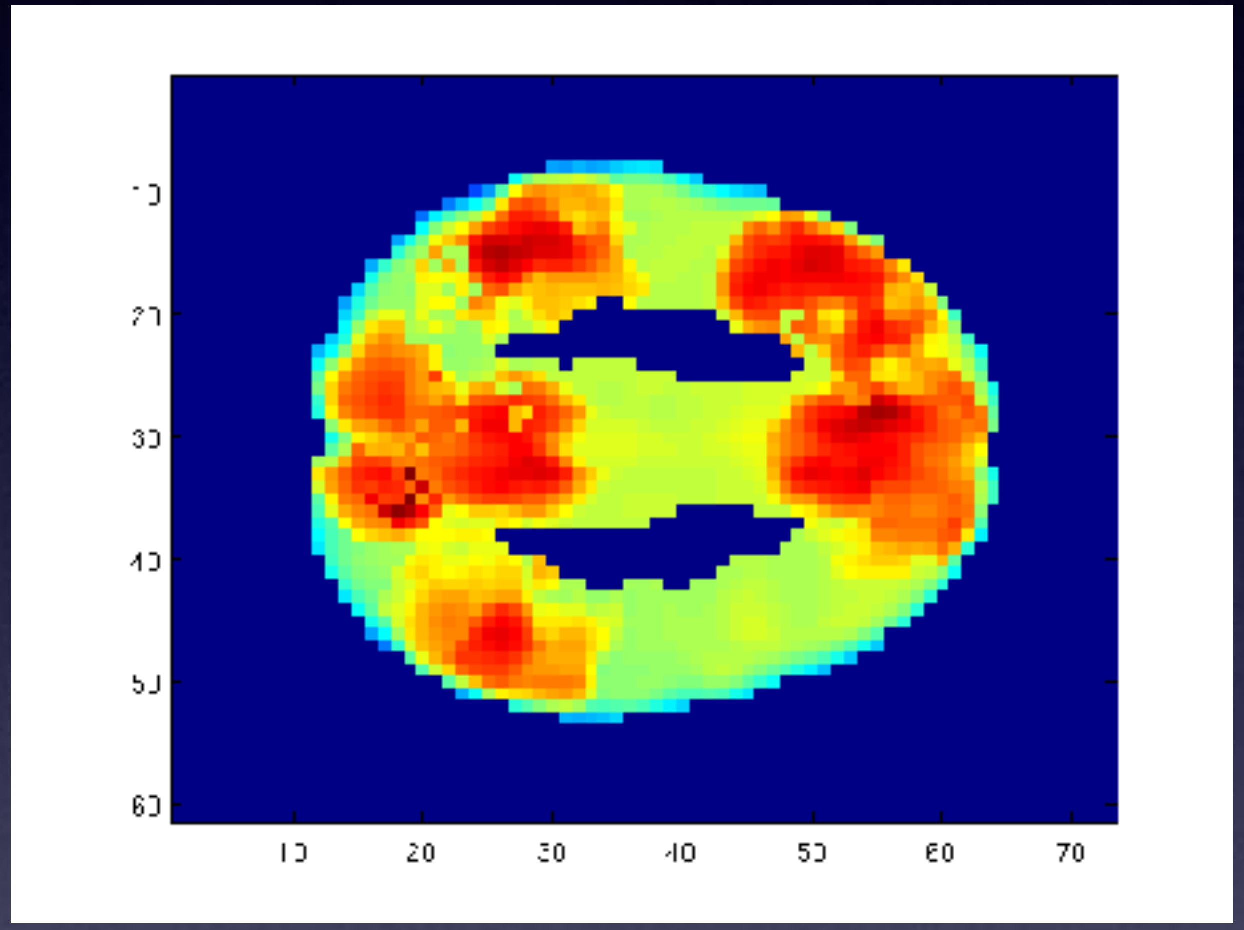
Temporal Pattern Correlations

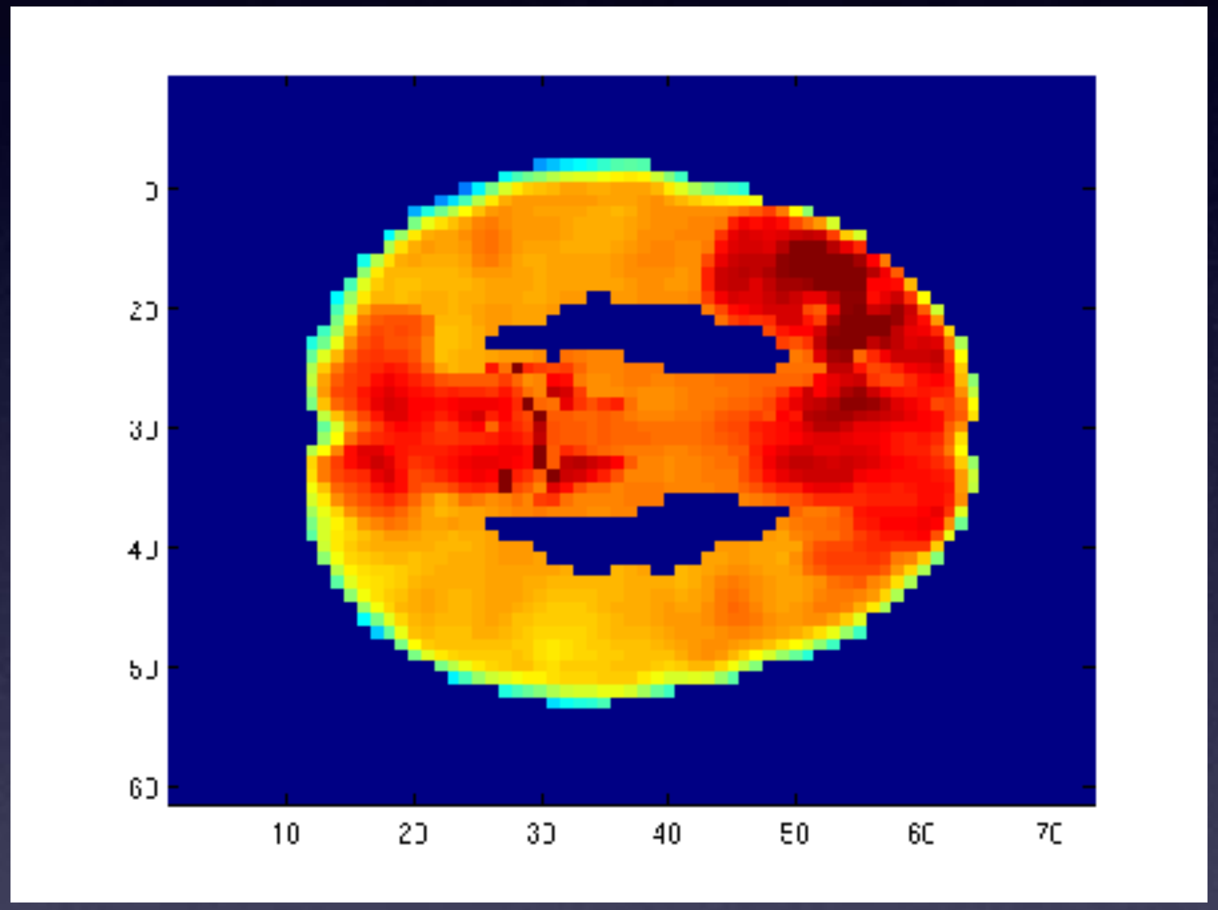
Network Discovery



Network Discovery

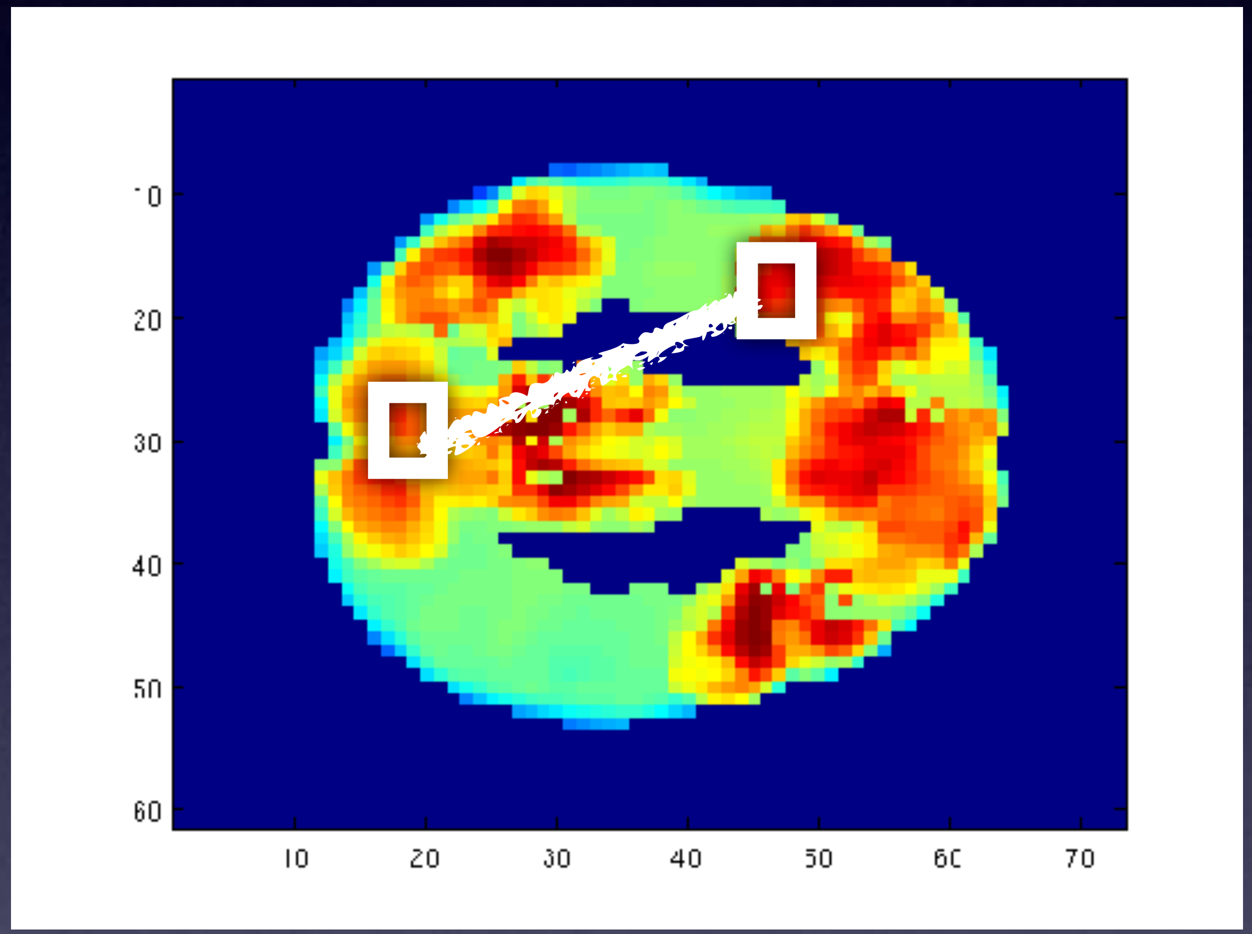


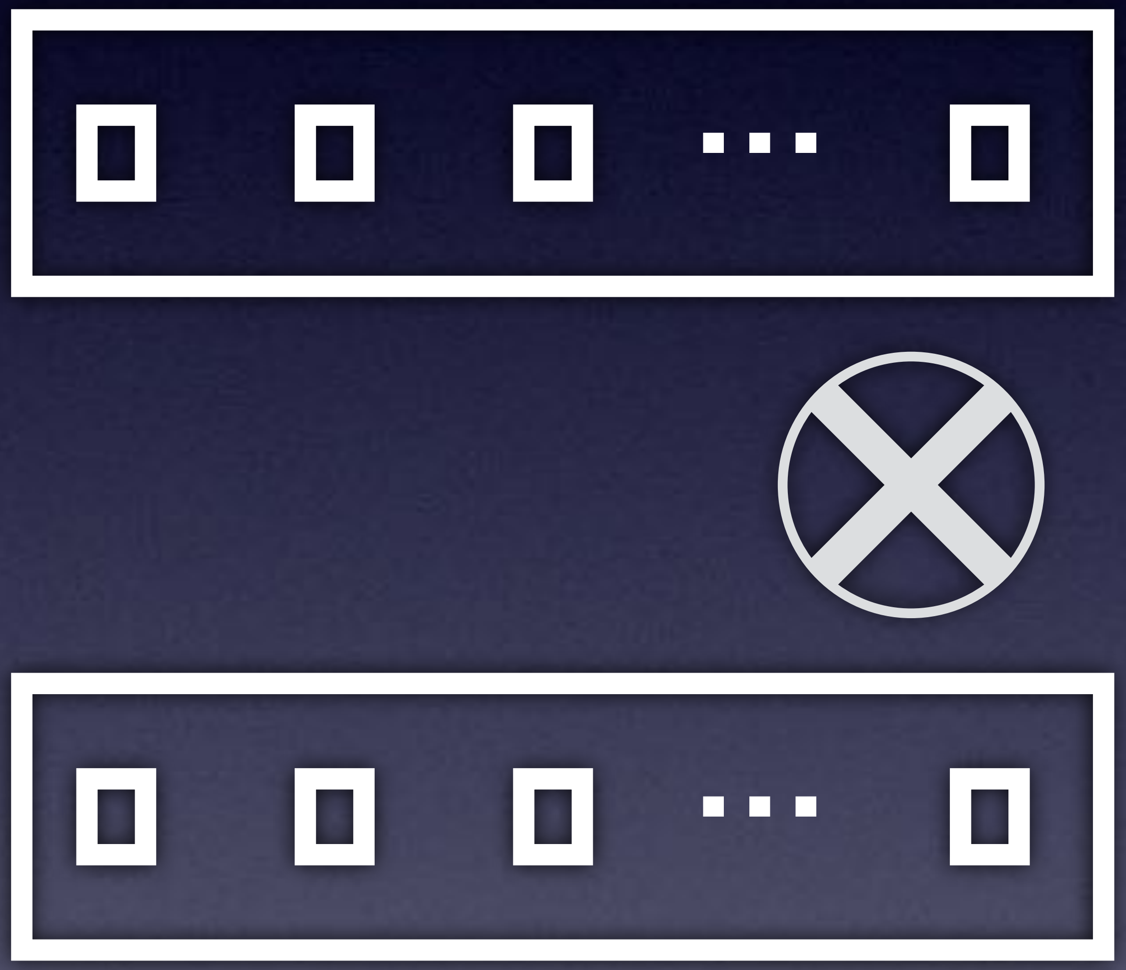




Temporal pattern of one spatial voxel

Network Discovery

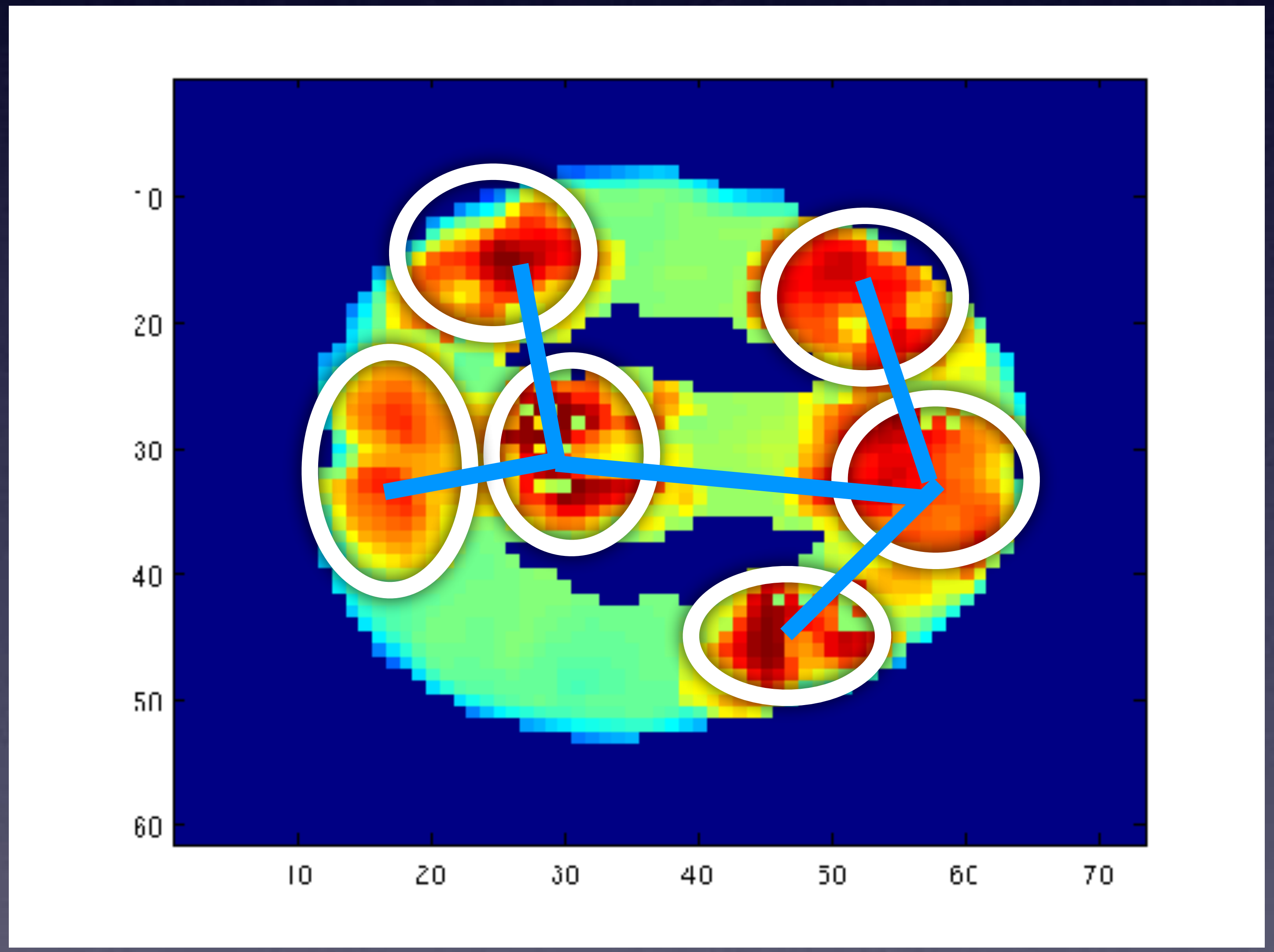




Absolute Pearson Correlation

$\mathbf{X}_{i,j}$

Network Discovery



Co-activation Network

From $\mathbf{X}$

Constrained & Regularized

Node: cluster of voxels that are co-active

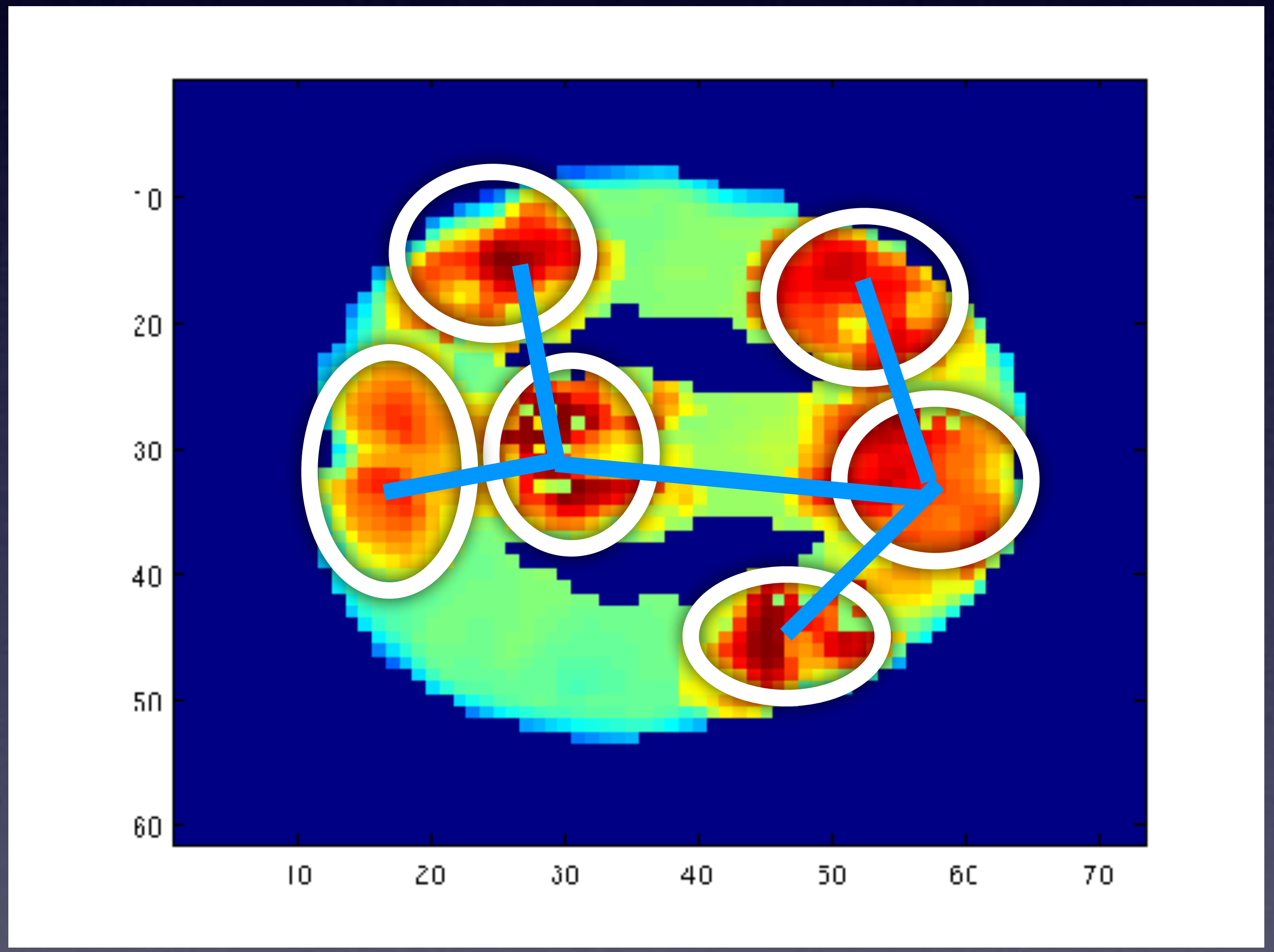
Edge: association between a pair of Nodes

KDD2017

Outline

- Introduction to fMRI
- Define the problem setting
- **Previous work & Our Method**
- Experiments: Synthetic & Real-world Data
- Future Work

Previous Work



Related Work

Voxel Clustering

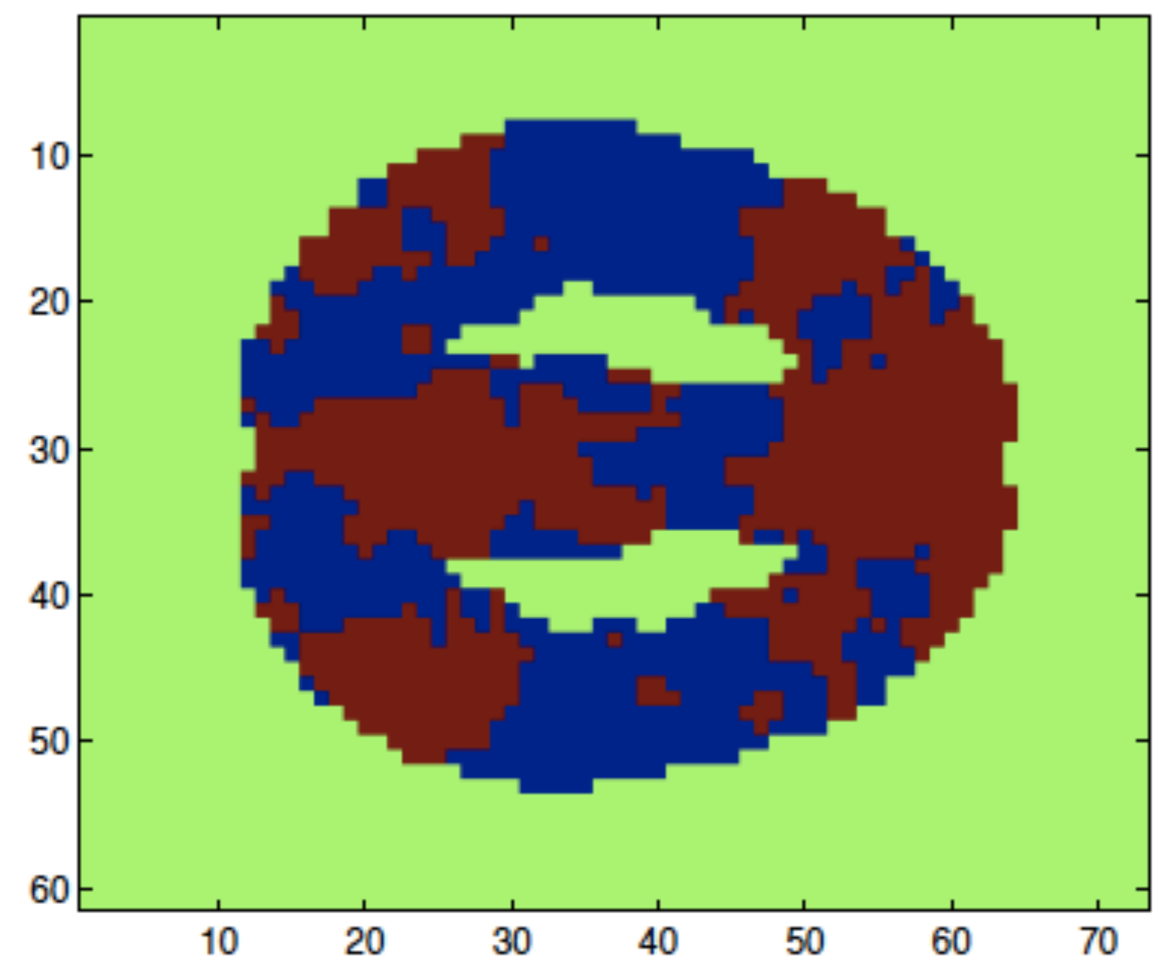
- Neuron' 10 - Andrews-Hanna et al.
- Human Brain Mapping'04 - Ven et al.

OR

Edge Learning

- Human Brain Mapping'09 - Burge et al.
- KDD'09 - Sun et al.

Previous Work



(c) Constrained cut by transferring Scan 1 to 2

Previously

- **KDD' 10, DMKD'14 - Wang & Davidson**
- **Foreground vs Background**

KDD'15 - Kuo and Davidson

Stacks of Graphs

KDD'13 - Gilpin & Davidson

Network discovery with strong supervision

KDD2017

Previous Work

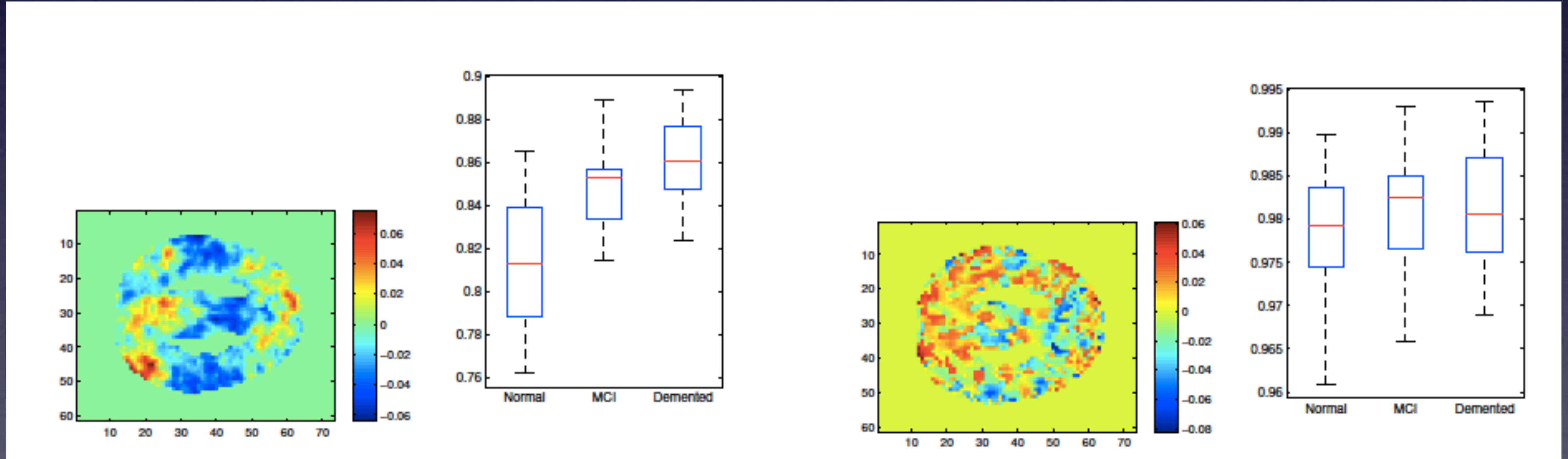
Previously

KDD'10 - Wang & Davidson

Foreground vs Background

- **KDD'15 - Kuo and Davidson**

Stacks of Graphs



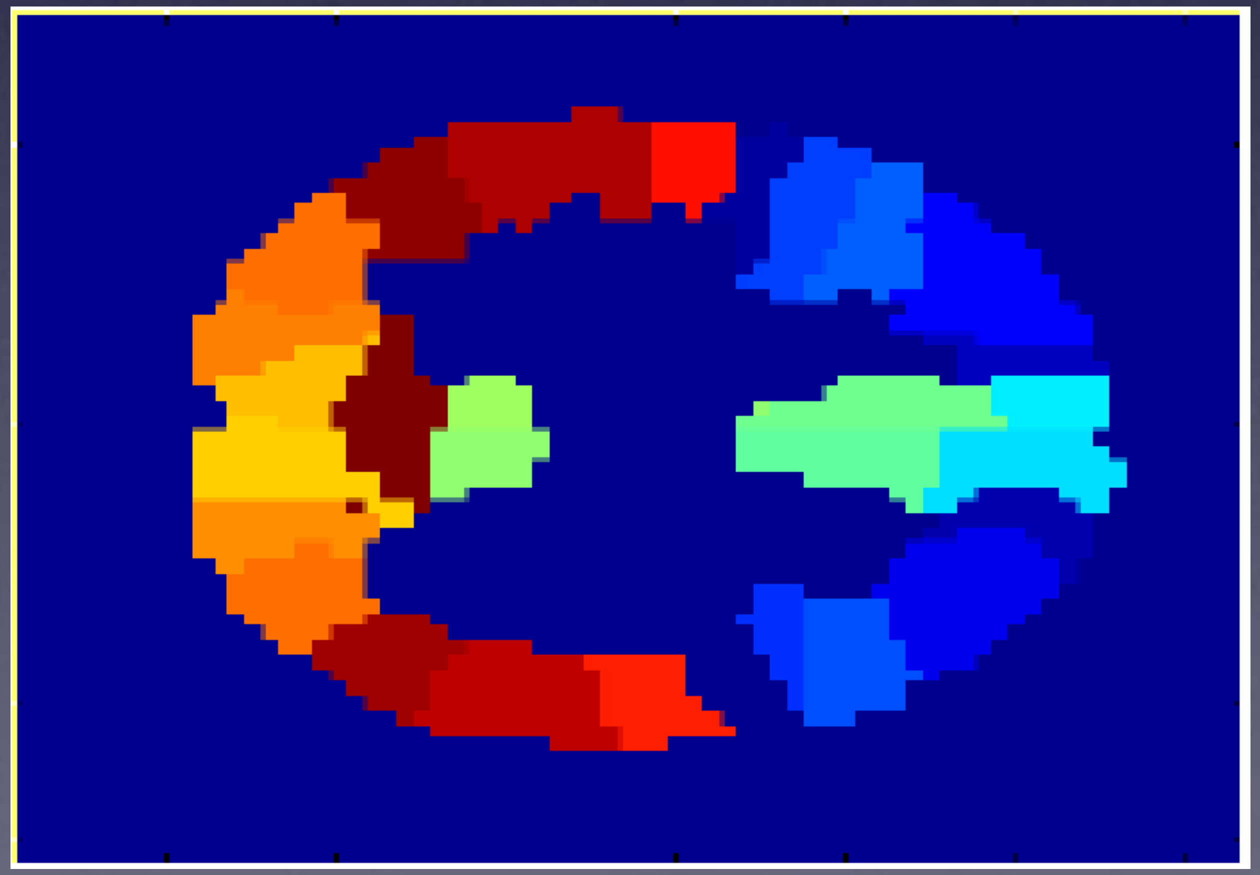
(a) Contrast cut (b) Grouped cut costs (a) Unified cut (b) Grouped cut costs

KDD'13 - Gilpin & Davidson

Network discovery with strong supervision

Previous Work

Neurologists like Professor Carmichael Know the Brain contains 116 anatomical regions



Previously

KDD’10 - Wang & Davidson

Foreground vs Background

KDD’15 - Kuo and Davidson

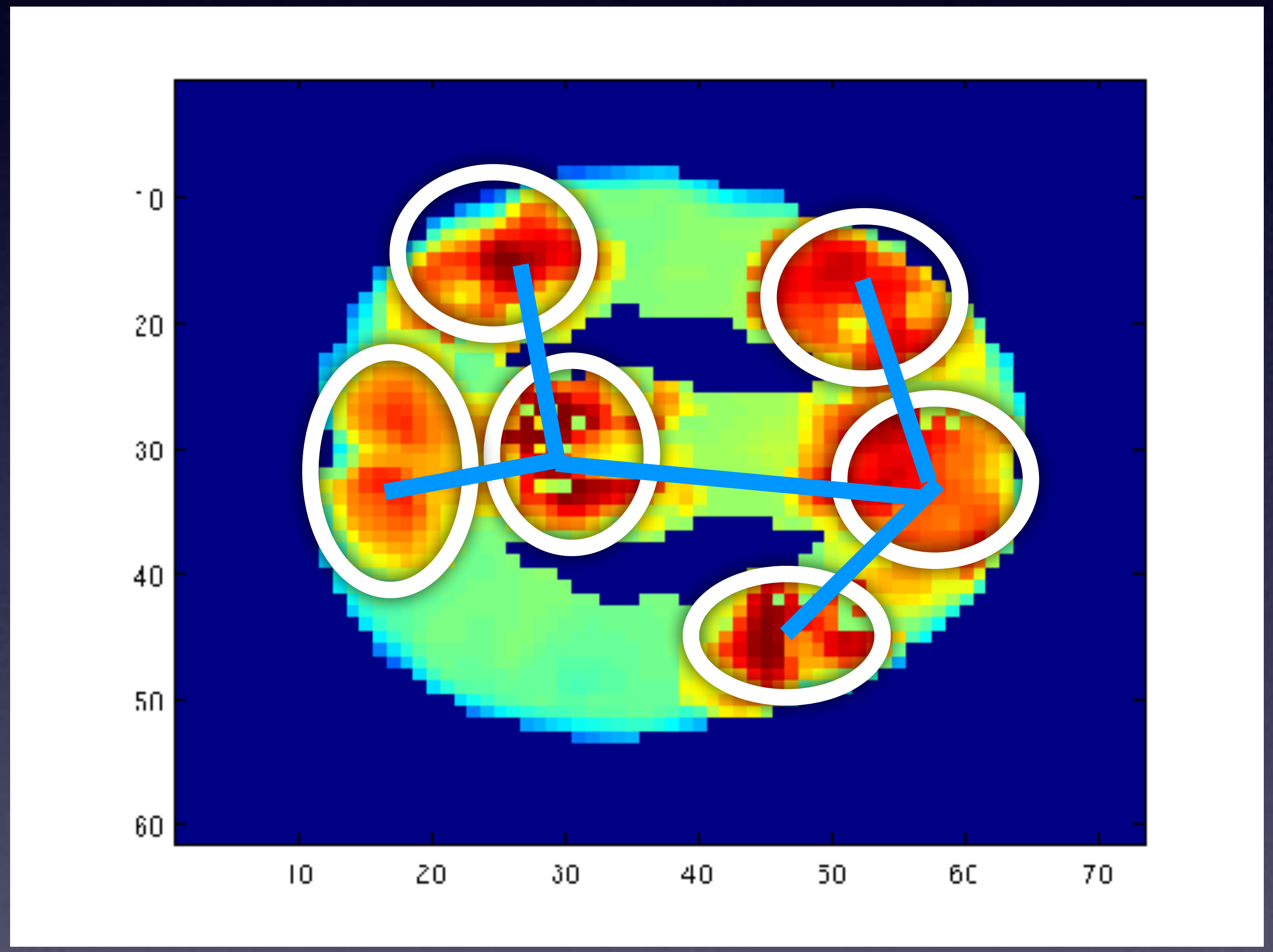
Stacks of Graphs

- **KDD’13 - Gilpin & Davidson**

  Network discovery with
- **STRONG SUPERVISION**

LIMITED

Our Method



KDD’17 Bai and Davidson

Complete Network Discovery

Nodes & Edges + **NO** Supervision

Previously

KDD’10 - Wang & Davidson

Foreground vs Background

KDD’15 - Kuo and Davidson

Stacks of Graphs

KDD’13 - Gilpin & Davidson

Network discovery with strong supervision

KDD2017

Our Method

$$\underset{\mathbf{F}\geq 0,\mathbf{M}\geq 0}{Minimize} \|\mathbf{X}-\mathbf{F}\mathbf{M}\mathbf{F}^T\|_F^2+\beta tr(\mathbf{F}^T\boldsymbol{\Theta}\mathbf{F})$$

$$s.t. \quad \mathbf{F}^T\mathbf{F} = \mathbf{I}$$

- Block Modeling:
  - Input symmetric affinity matrix
- Nonnegative Matrix tri-Factorization
- Regularize cluster indicator for Spatial Continuity

KDD2017

Our Method

- "Affinity" matrix
- Absolute Correlations
- Graph: ***N*** by ***N***

Spatial Continuity Regularization

$$\underset{\mathbf{F}\geq 0, \mathbf{M}\geq 0}{Minimize} \|\mathbf{X}-\mathbf{F}\mathbf{M}\mathbf{F}^T\|_F^2 + \beta tr(\mathbf{F}^T\mathbf{\Theta}\mathbf{F})$$
$$s.t. \quad \mathbf{F}^T\mathbf{F} = \mathbf{I}$$

- Cluster indicator matrix **(Nodes)**
- ***N*** by ***k***
- [0,1]
- Column-wise orthogonal

- Mixing matrix **(Edges)**
- ***k*** by ***k***
- Nonnegative
- Associations between clusters

Our Method

Spatial Continuity Regularization

$$\underset{\mathbf{F}\geq 0, \mathbf{M}\geq 0}{Minimize} \|\mathbf{X}-\mathbf{F}\mathbf{M}\mathbf{F}^T\|_F^2 + \beta tr(\mathbf{F}^T \mathbf{\Theta} \mathbf{F})$$

$$s.t. \quad \mathbf{F}^T\mathbf{F} = \mathbf{I}$$

KDD2017

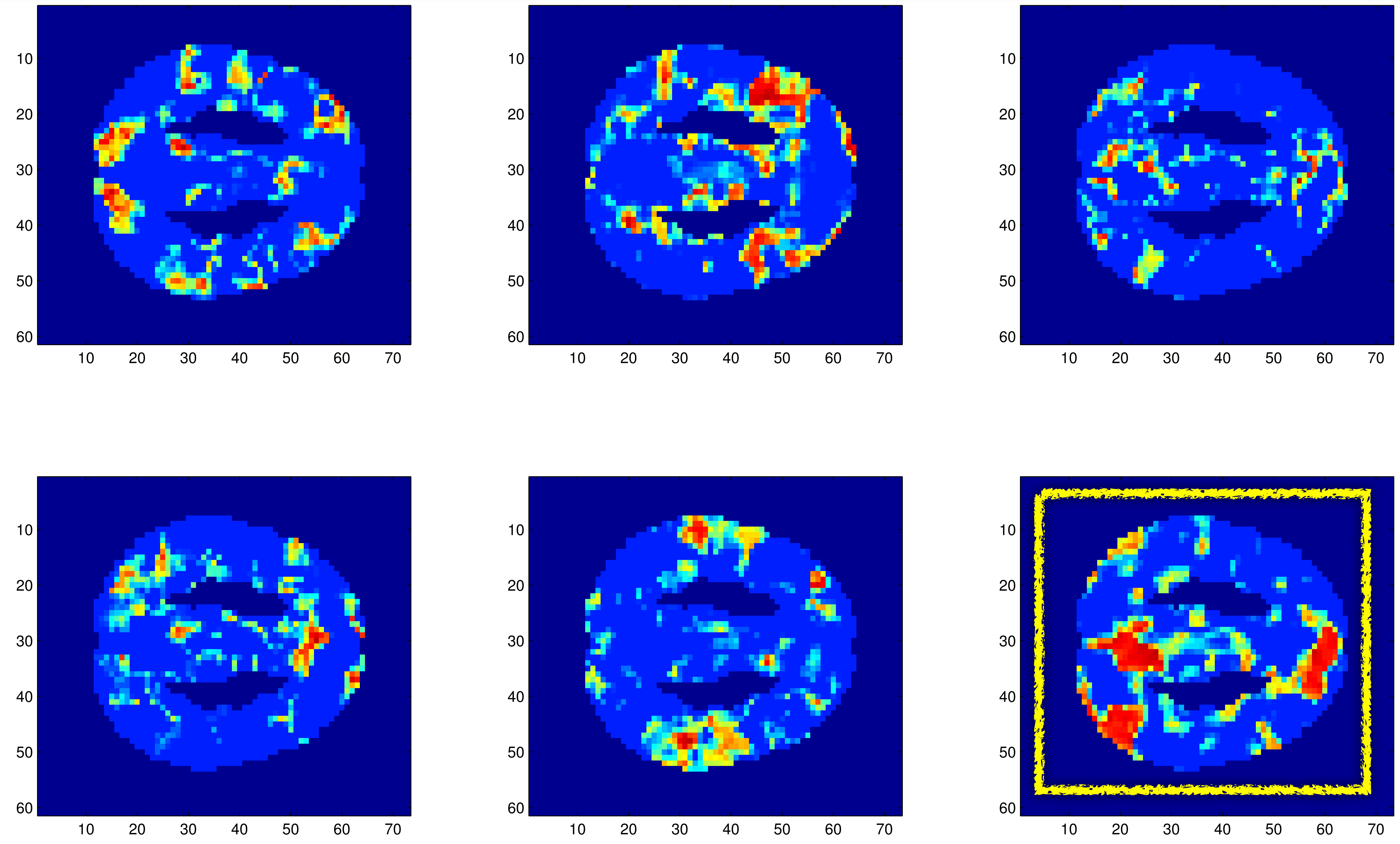
Baseline Method

$$\underset{\mathbf{F}\geq 0,\mathbf{M}\geq 0}{Minimize} \|\mathbf{X}-\mathbf{F}\mathbf{M}\mathbf{F}^T\|_F^2 + \cancel{\beta tr(\mathbf{F}^T\Theta\mathbf{F})}$$

$$s.t. \quad \mathbf{F}^T\mathbf{F} = \mathbf{I}$$

~~Spatial Continuity Regularization~~

- Why do we need it?

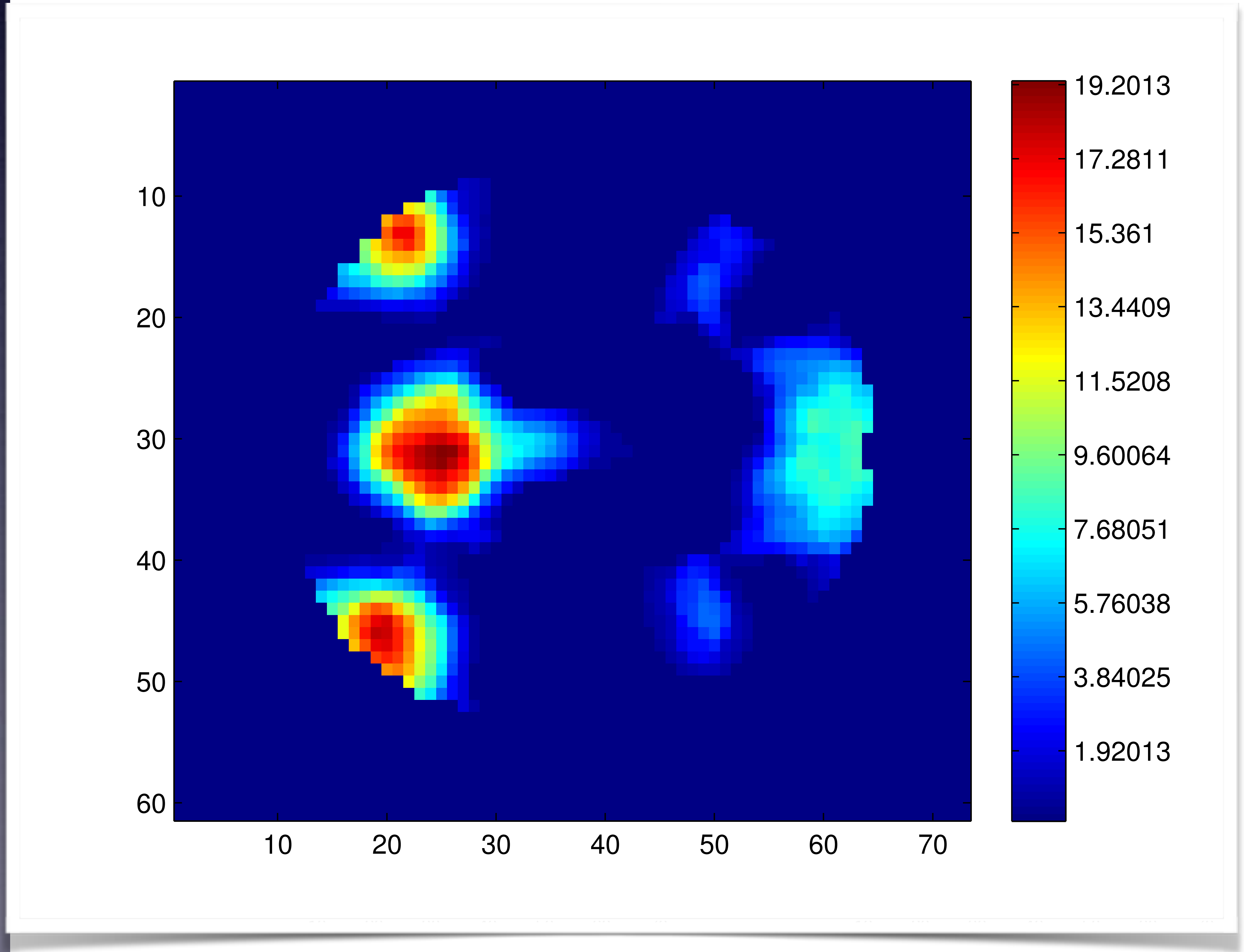


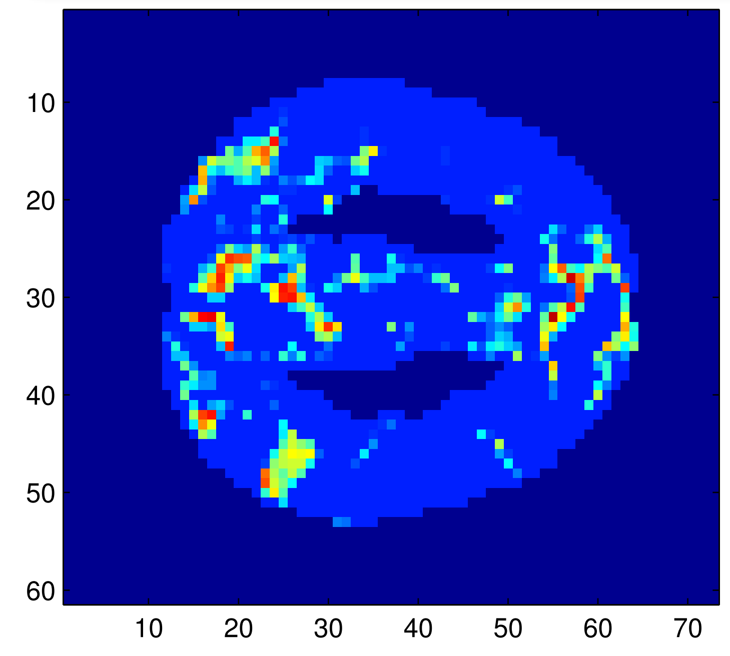
Baseline Method

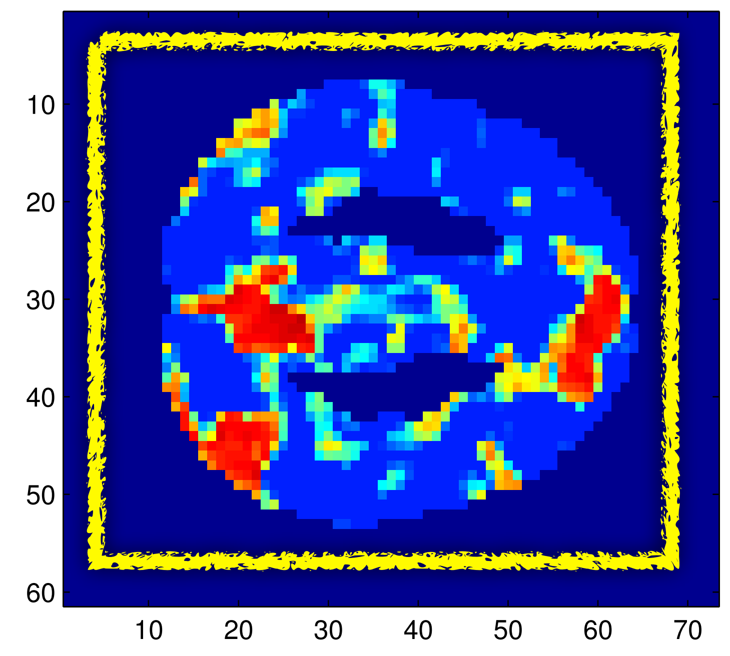
$$\underset{\mathbf{F}\geq 0,\mathbf{M}\geq 0}{Minimize} \|\mathbf{X}-\mathbf{F}\mathbf{M}\mathbf{F}^T\|_F^2 + \cancel{\beta tr(\mathbf{F}^T\Theta\mathbf{F})}$$

$$s.t. \quad \mathbf{F}^T\mathbf{F} = \mathbf{I}$$

~~Spatial Continuity Regularization~~

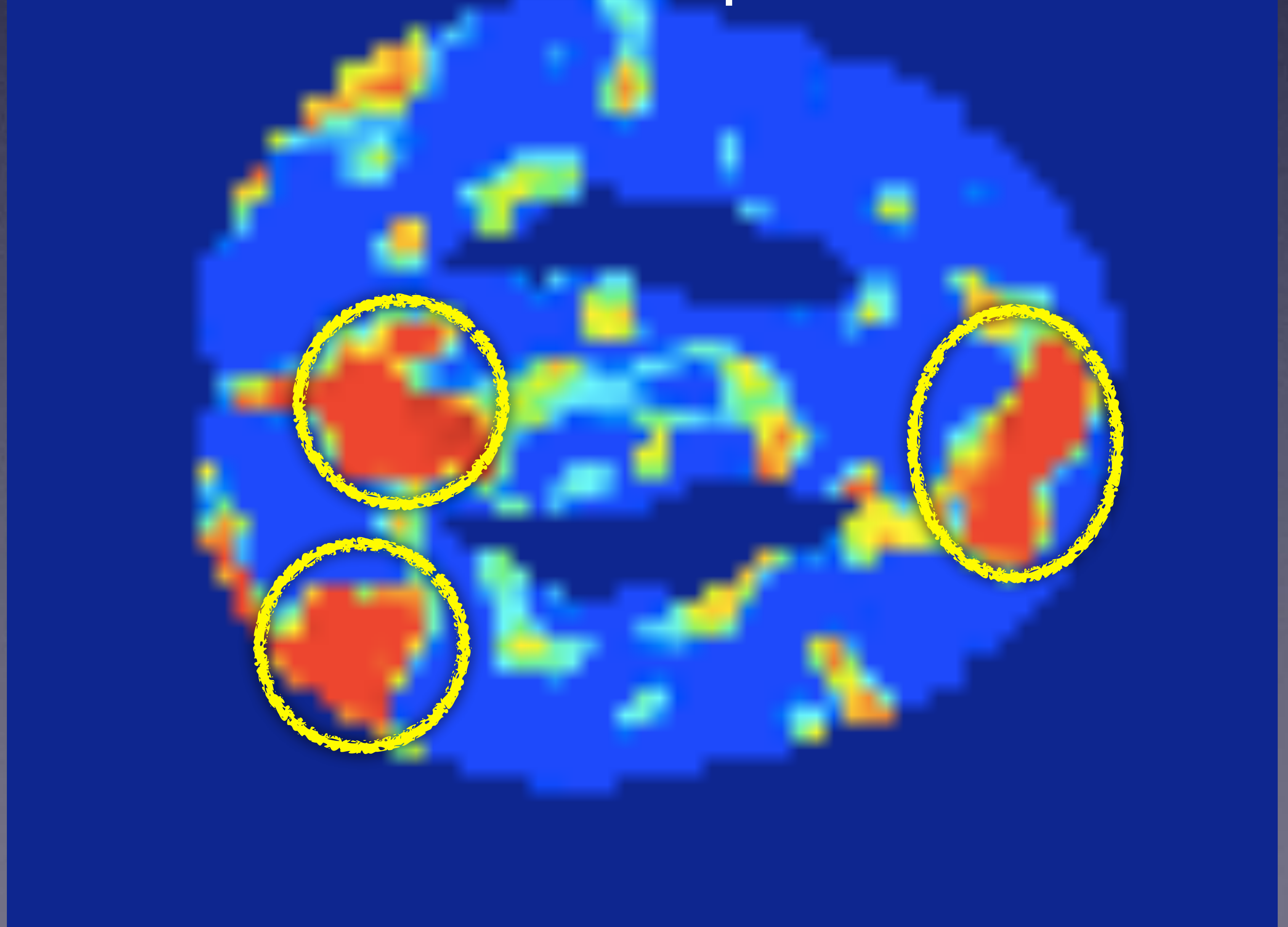






One Cluster

Contains Multiple Nodes



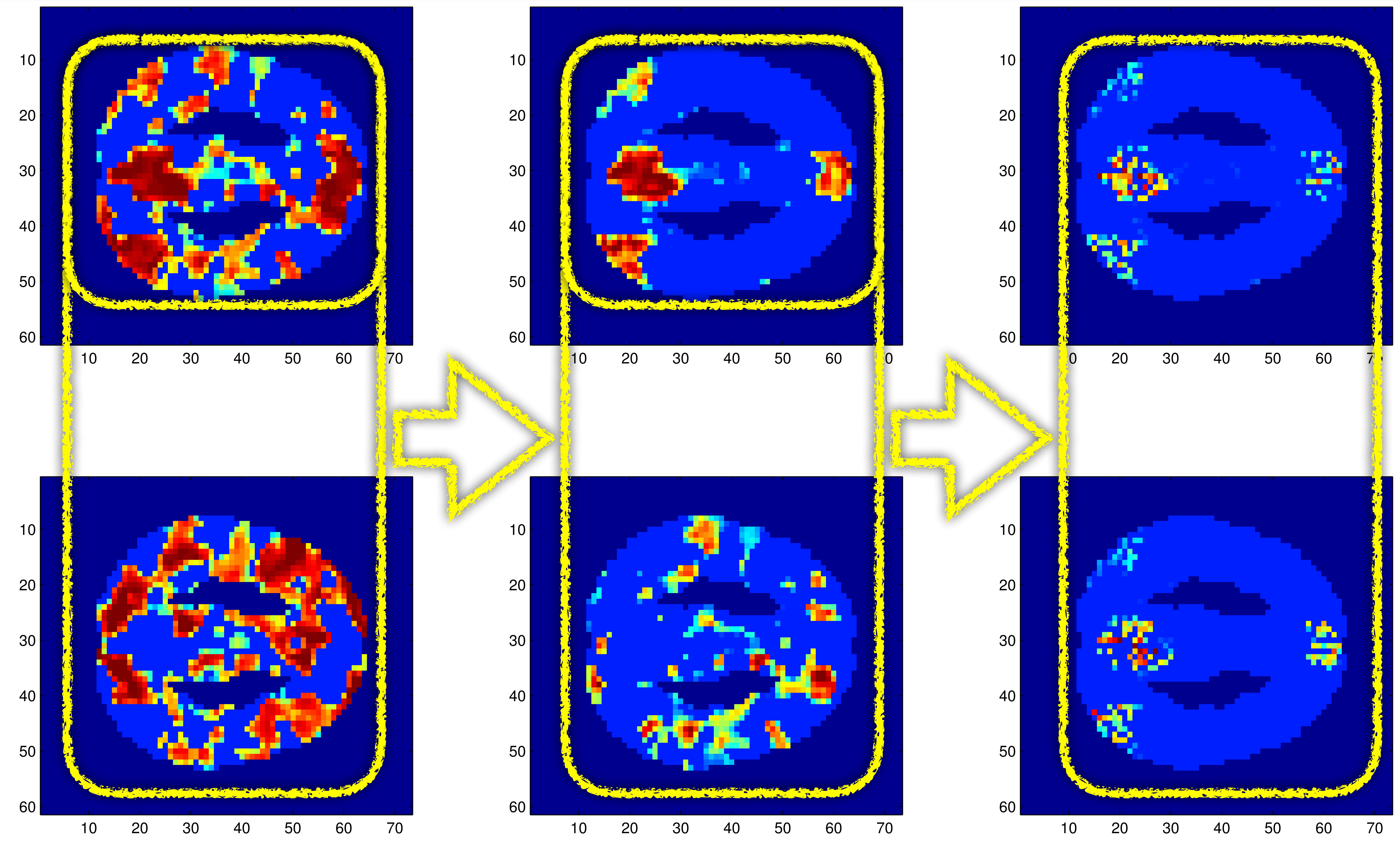
Baseline Method

$$\underset{\mathbf{F}\geq 0,\mathbf{M}\geq 0}{Minimize} \|\mathbf{X}-\mathbf{F}\mathbf{M}\mathbf{F}^T\|_F^2+\cancel{\beta tr(\mathbf{F}^T\mathbf{\Theta}\mathbf{F})}$$

$$s.t. \quad \mathbf{F}^T\mathbf{F}=\mathbf{I}$$

~~Spatial Continuity Regularization~~

- Why do we need it?
  - How about recursive decomposition?



KDD2017

Our Method

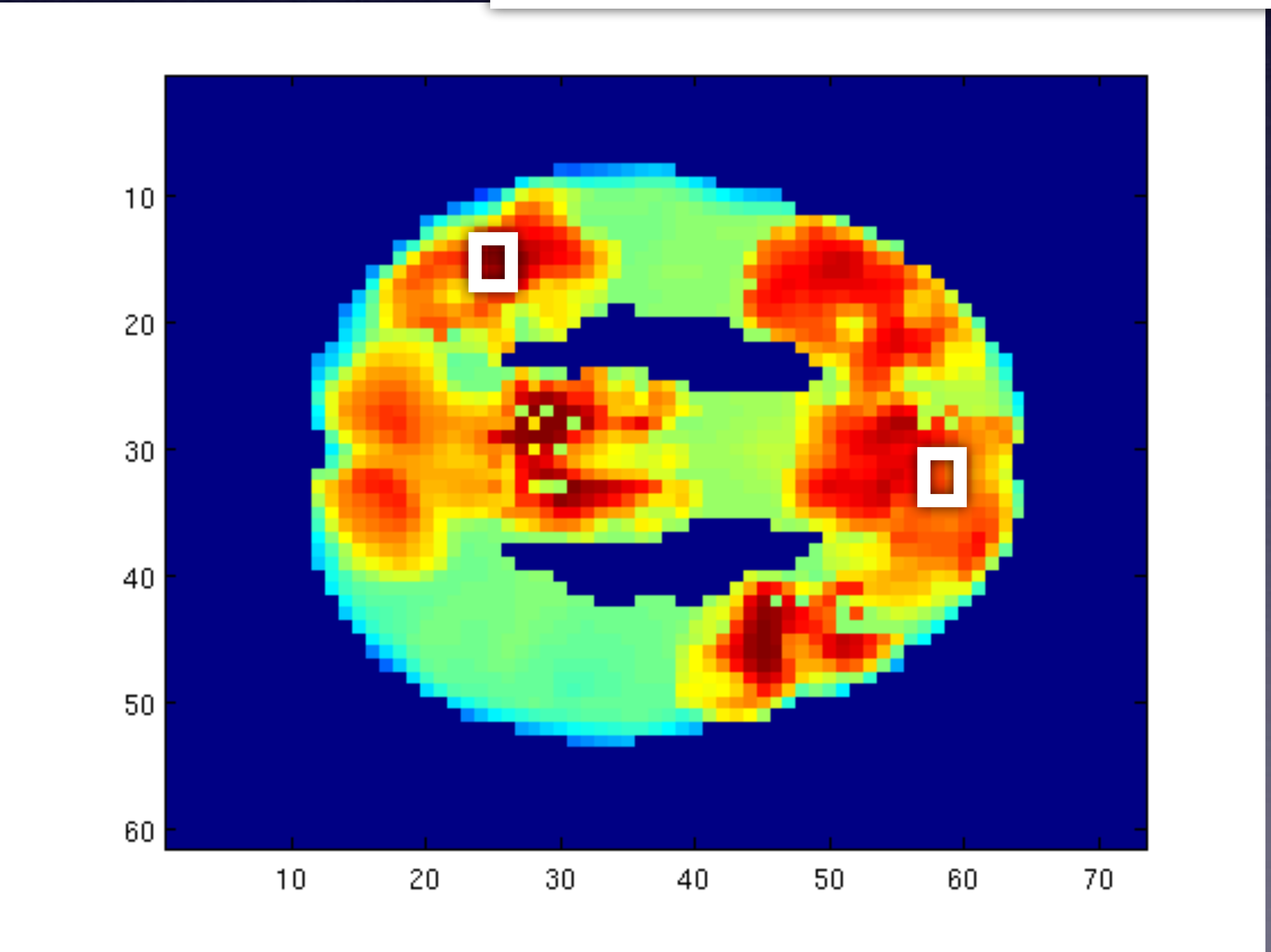
$$\underset{\mathbf{F}\geq 0,\mathbf{M}\geq 0}{Minimize}\|\mathbf{X}-\mathbf{F}\mathbf{M}\mathbf{F}^T\|_F^2+\beta tr(\mathbf{F}^T\boldsymbol{\Theta}\mathbf{F})$$

$$s.t.\quad \mathbf{F}^T\mathbf{F}=\mathbf{I}$$

Spatial Continuity Regularization

- How does it work?

$$(\boldsymbol{\Theta})_{i,j}=e^{\|v_i-v_j\|_2^2}$$



KDD2017

Our Method

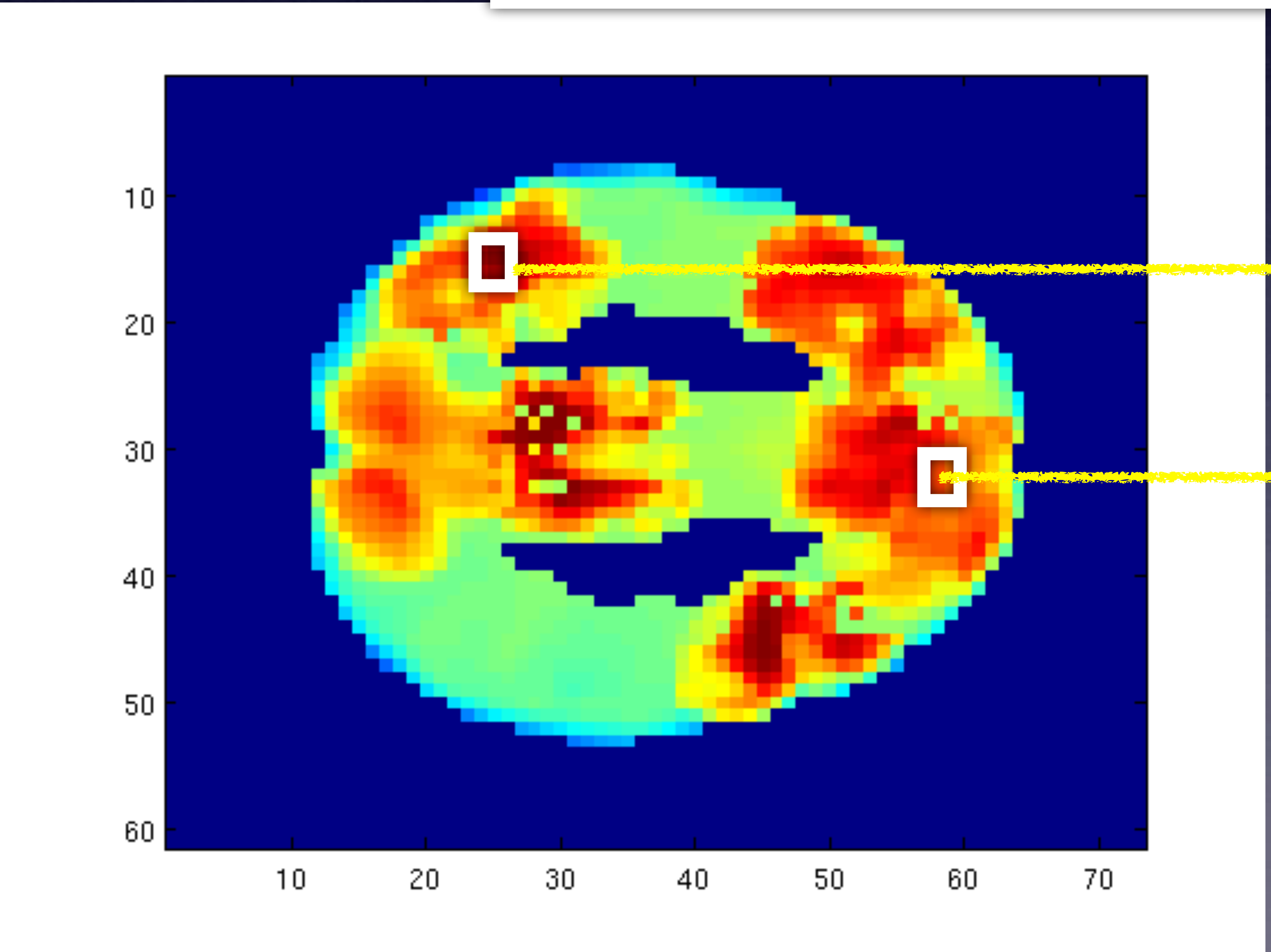
$$\underset{\mathbf{F}\geq 0,\mathbf{M}\geq 0}{Minimize}\|\mathbf{X}-\mathbf{F}\mathbf{M}\mathbf{F}^T\|_F^2+\beta tr(\mathbf{F}^T\mathbf{\Theta}\mathbf{F})$$

$$s.t.\quad \mathbf{F}^T\mathbf{F}=\mathbf{I}$$

Spatial Continuity Regularization

- How does it work?

$$(\mathbf{\Theta})_{i,j}=e^{\|v_i-v_j\|_2^2}$$



v_i

v_j

- Spatial Coordinates
- Reciprocal Gaussian Kernel

KDD2017

Our Method

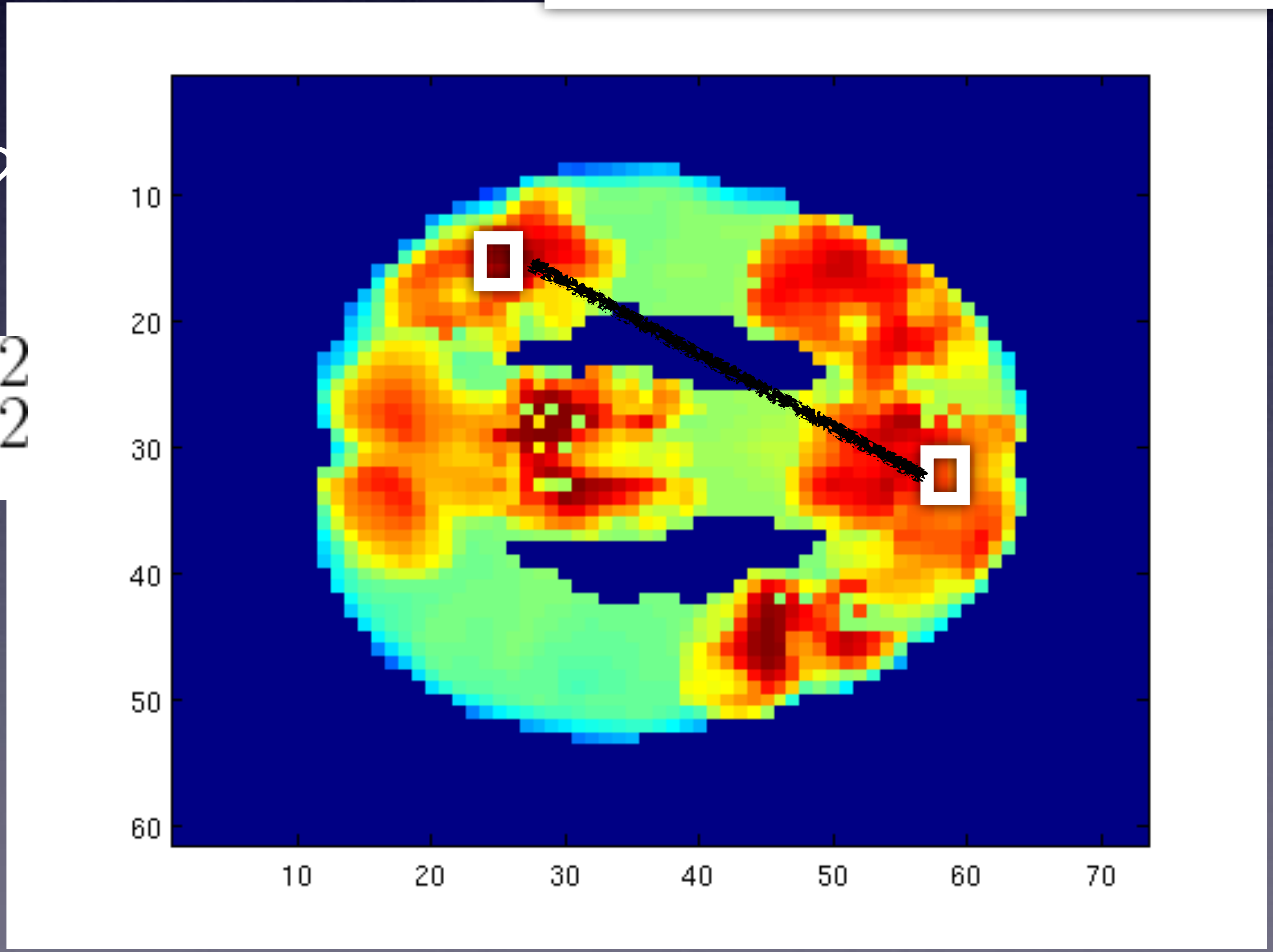
$$\underset{\mathbf{F}\geq 0, \mathbf{M}\geq 0}{Minimize} \|\mathbf{X}-\mathbf{F}\mathbf{M}\mathbf{F}^T\|_F^2 + \beta tr(\mathbf{F}^T \mathbf{\Theta} \mathbf{F})$$

$$s.t. \quad \mathbf{F}^T\mathbf{F} = \mathbf{I}$$

Spatial Continuity Regularization

- How does it work?

$$(\mathbf{\Theta})_{i,j} = e^{\|v_i - v_j\|_2^2}$$



- Further apart voxels suffer greater penalty to be in the same cluster.

KDD2017

Outline

- Introduction to fMRI
- Define the problem setting
- Previous work & Our Method
- **Experiments: Synthetic & Real-world Data**
- Future Work

KDD2017

Experiments **Purpose**

Synthetic

Know ground truth. Rediscover it.

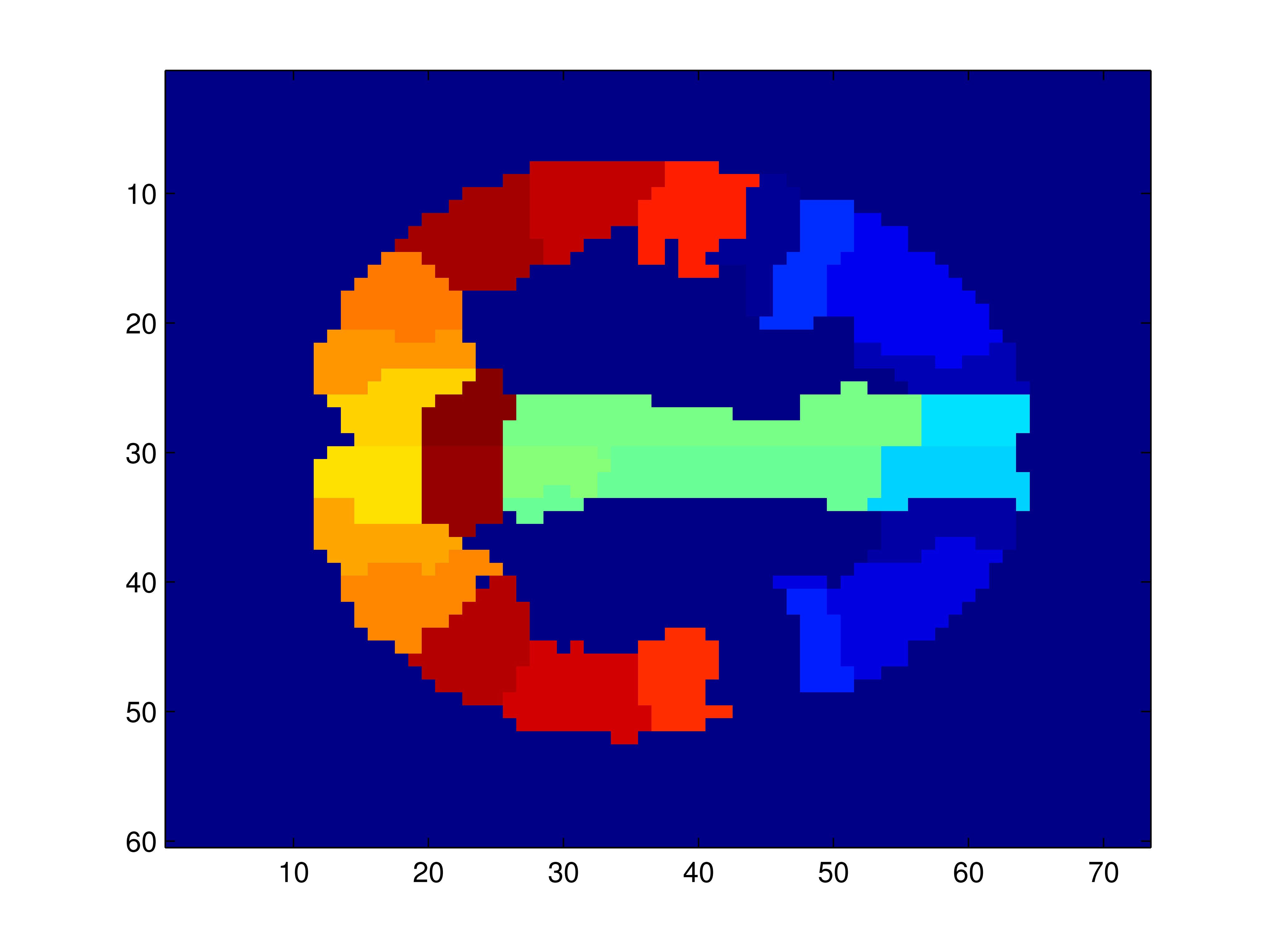
Real fMRI

Find insights consistent with domain scientists.

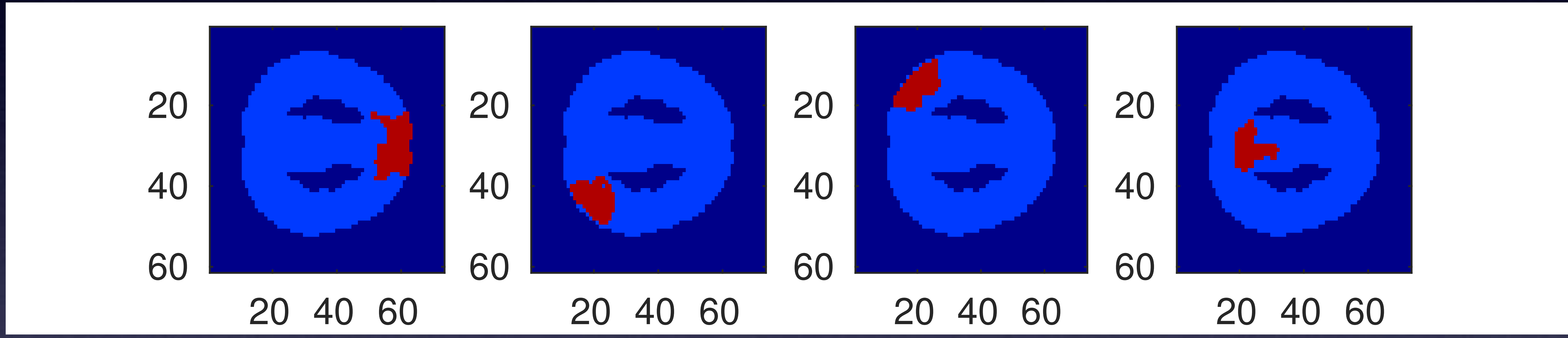
KDD2017

Experiments: Synthetic Data

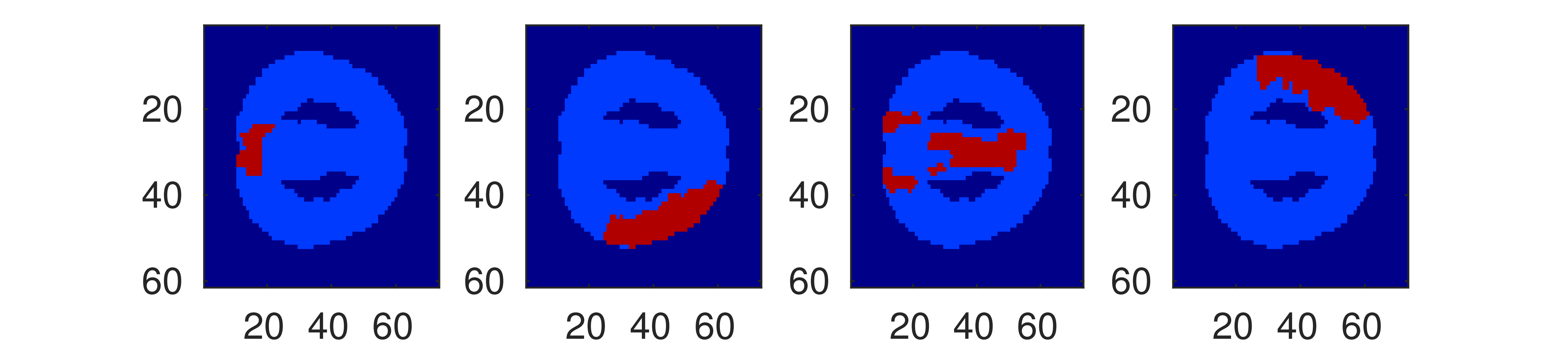
An Atlas on the 36-th Slice



Artificial but Realistic Functional Nodes



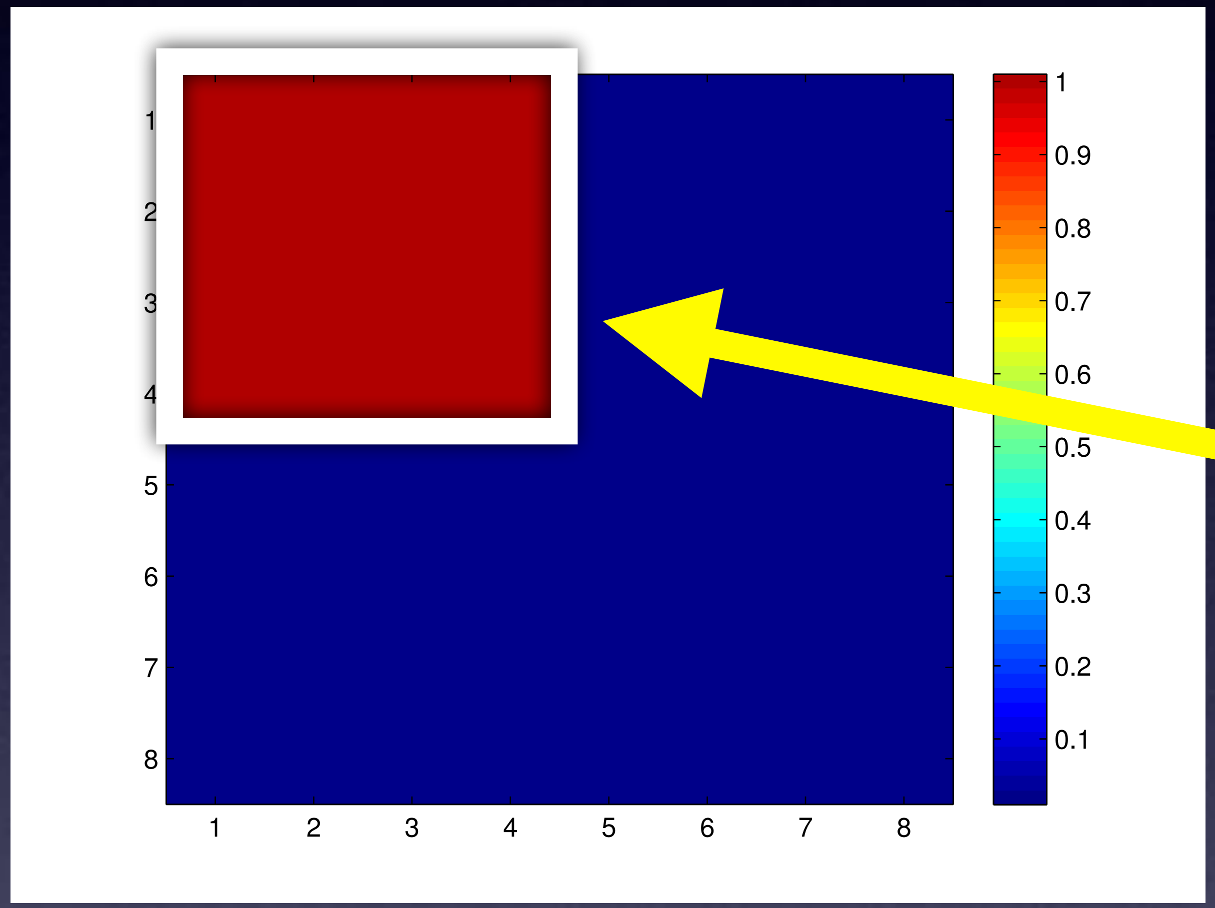
Background Nodes/Regions



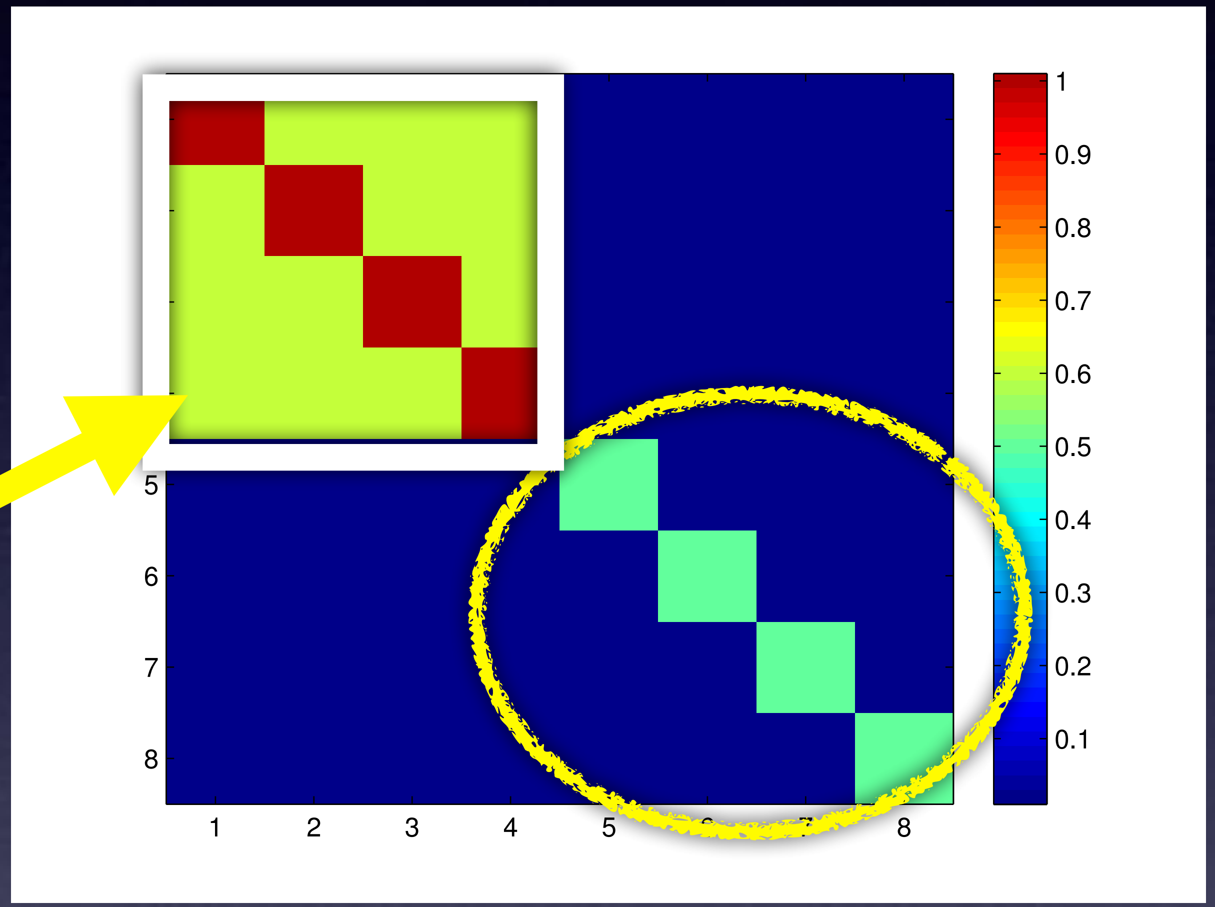
Experiments: Synthetic Ground-truth

Artificial correlations: different noisy settings

Simple

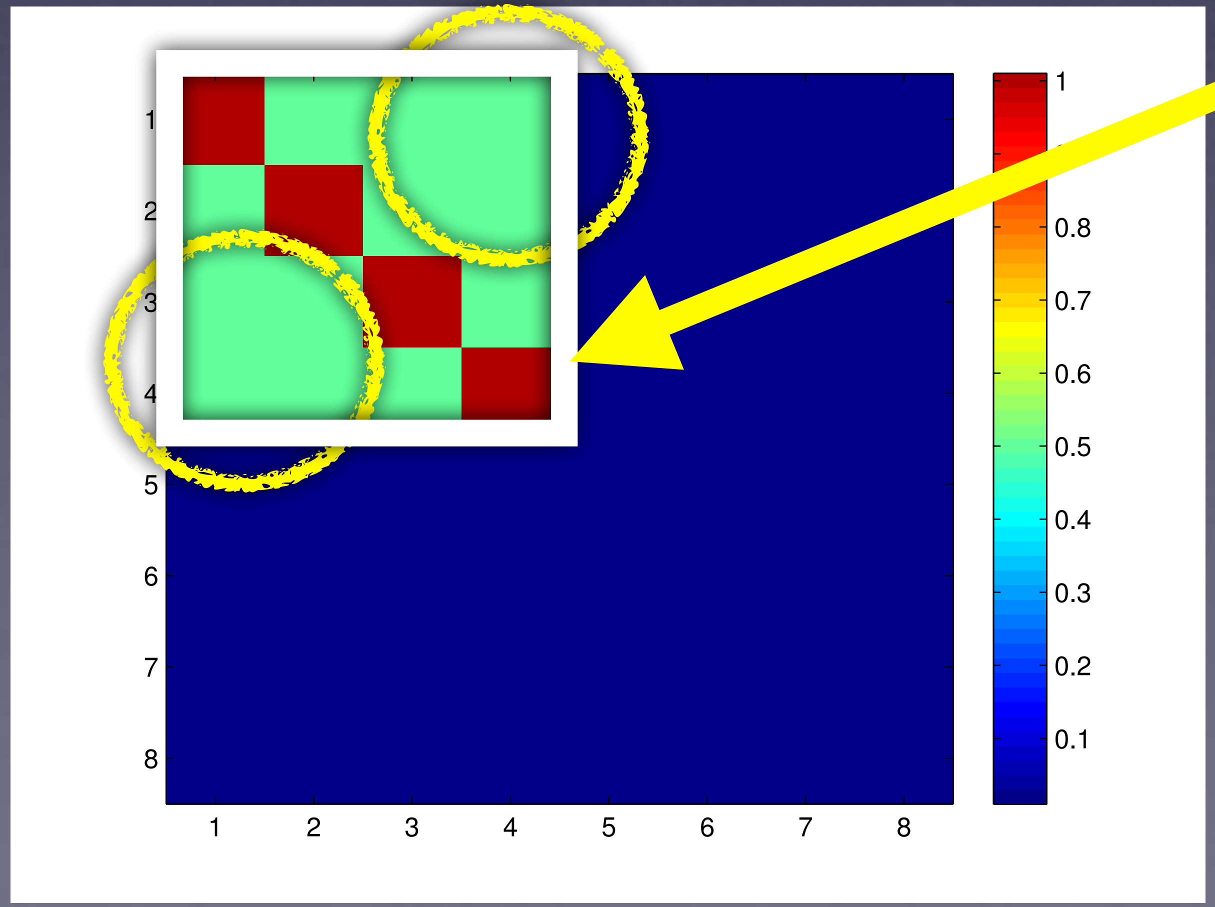


Local

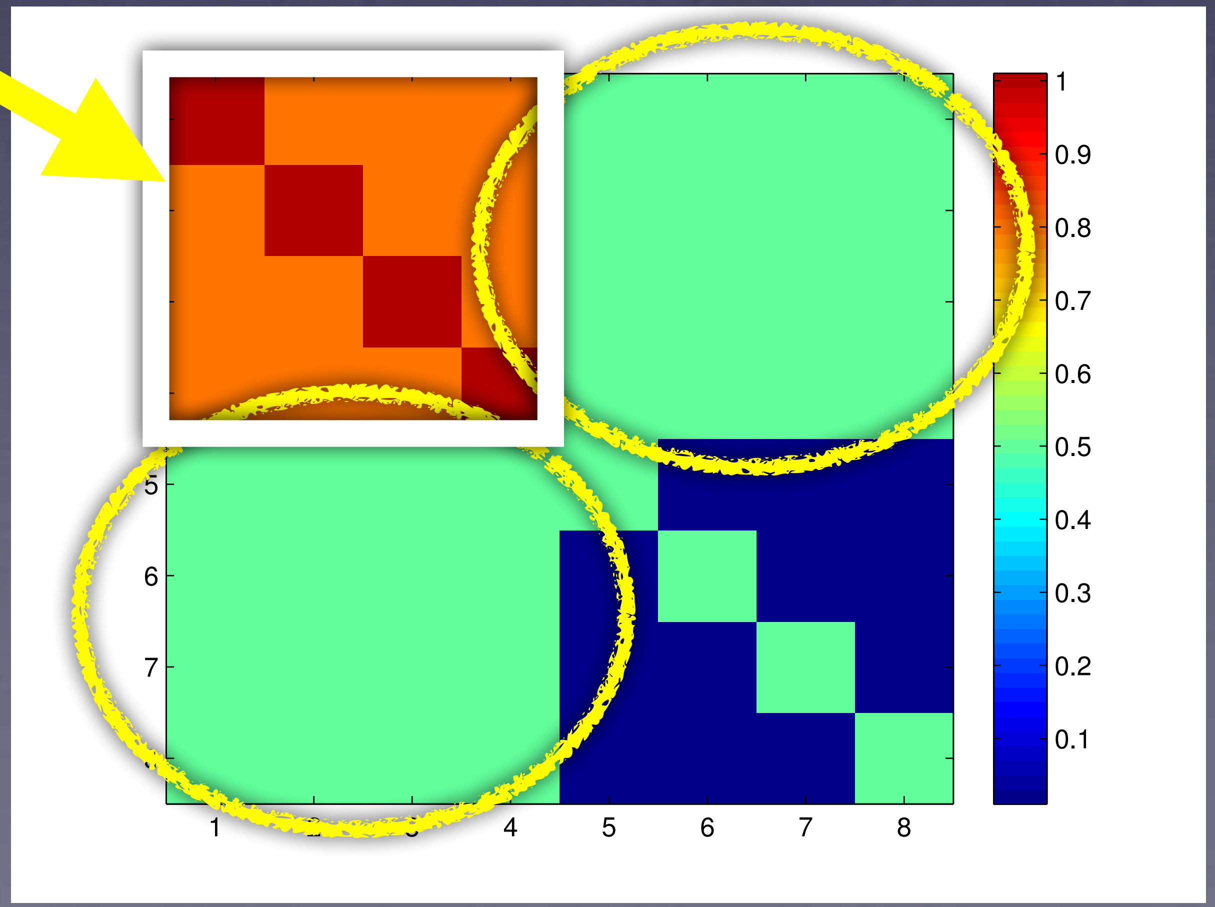


Edges between
Functional Nodes
To be reported on the next slide

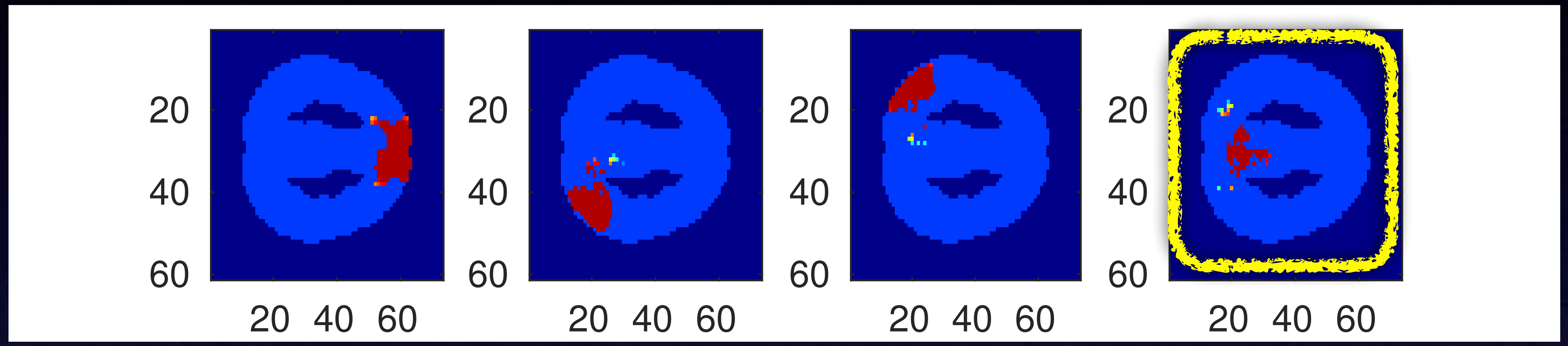
Degrading



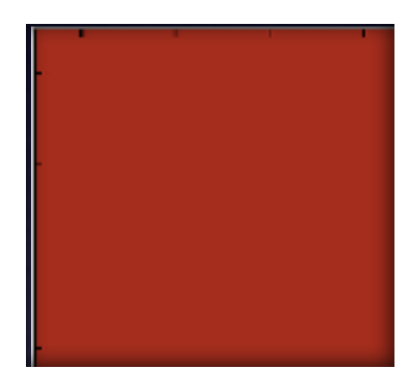
Global



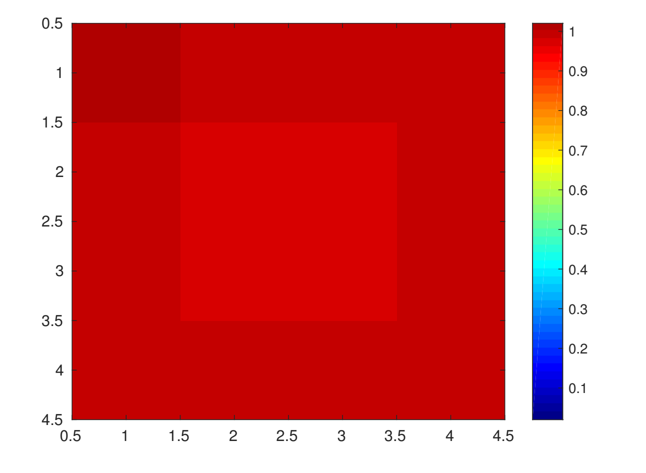
From Synthetic Data: Rediscovered Nodes & Edges

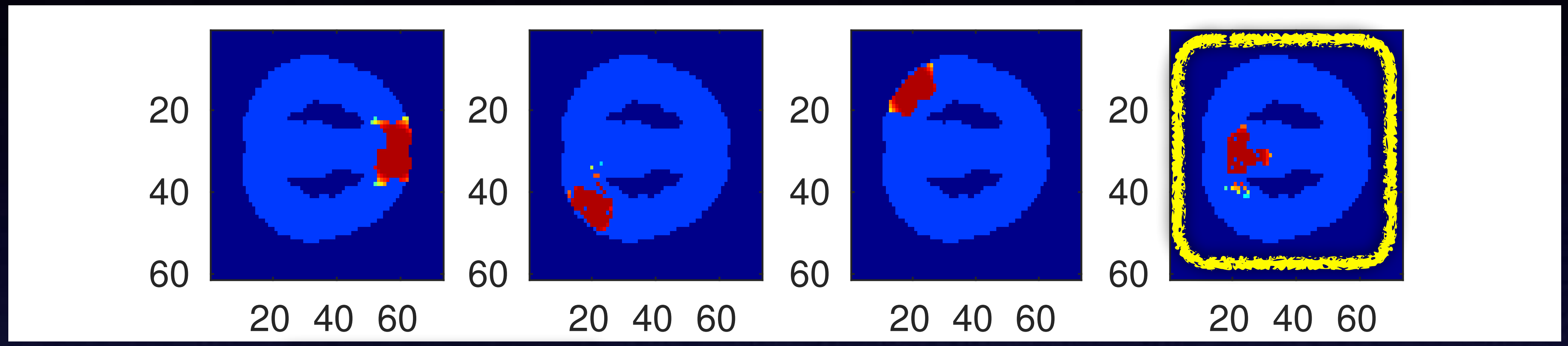


Simple

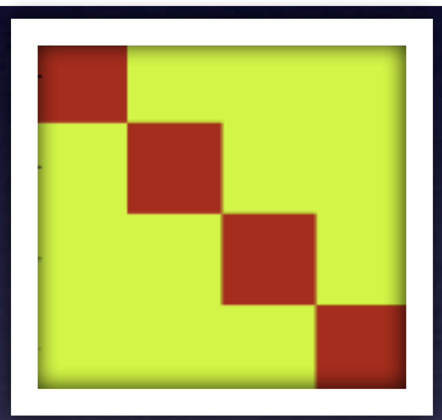


Ground Truth

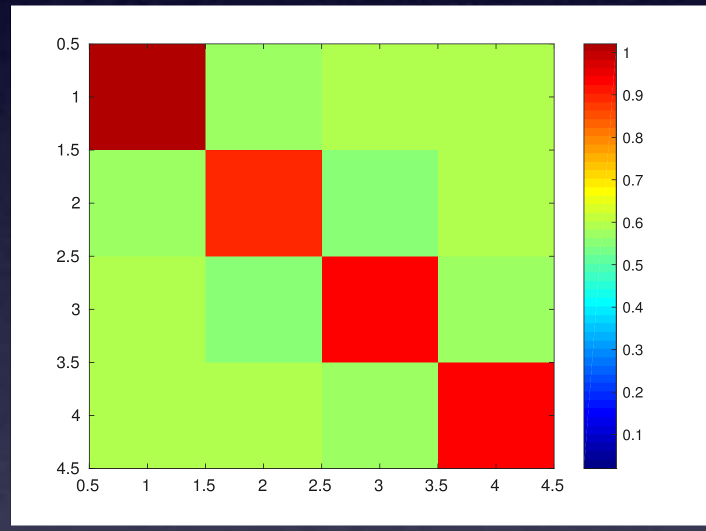


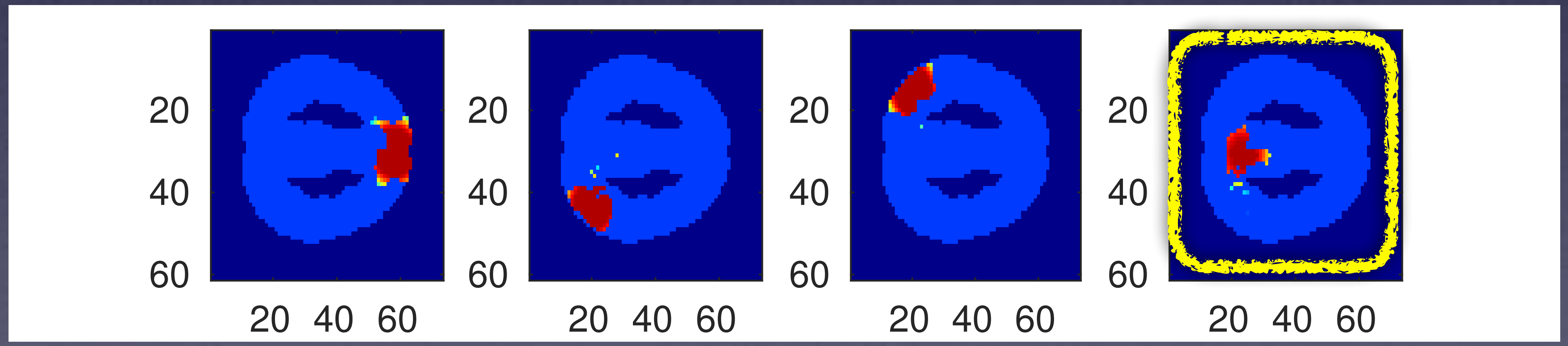


Local Noise

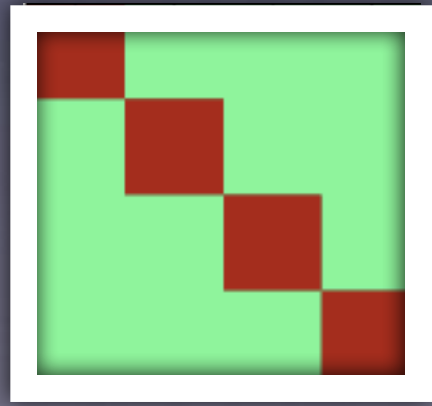


Ground Truth

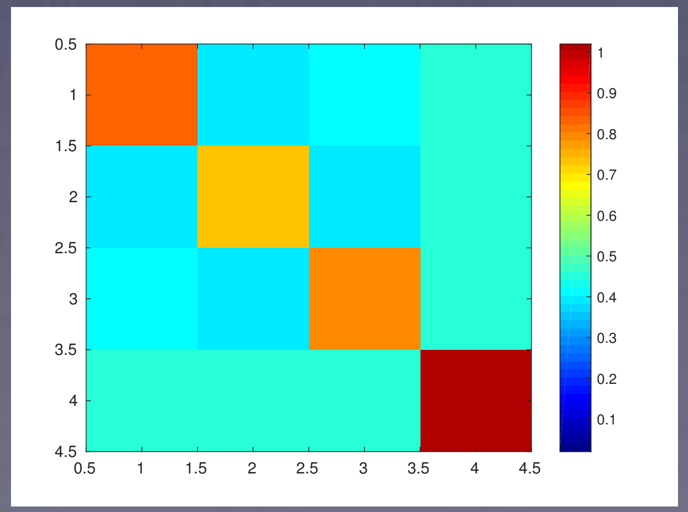


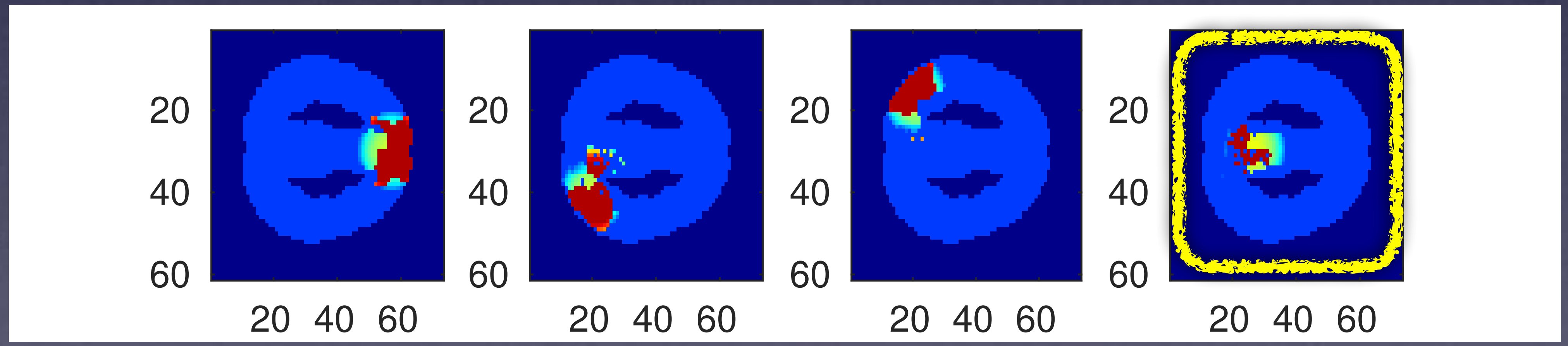


Degrading

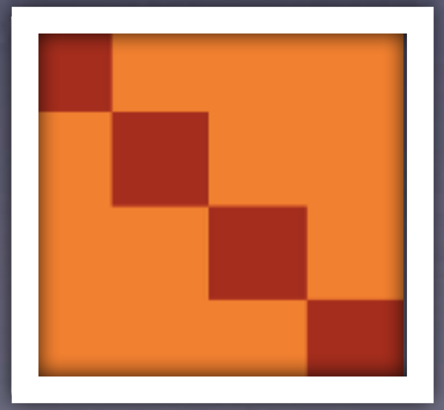


Ground Truth

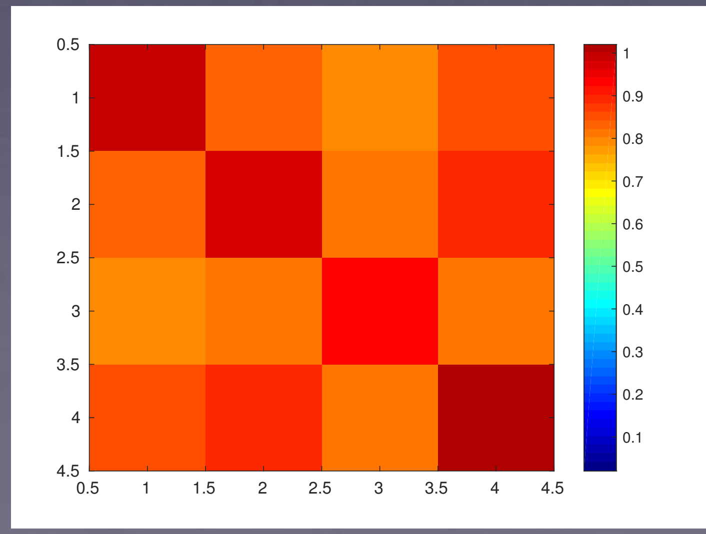


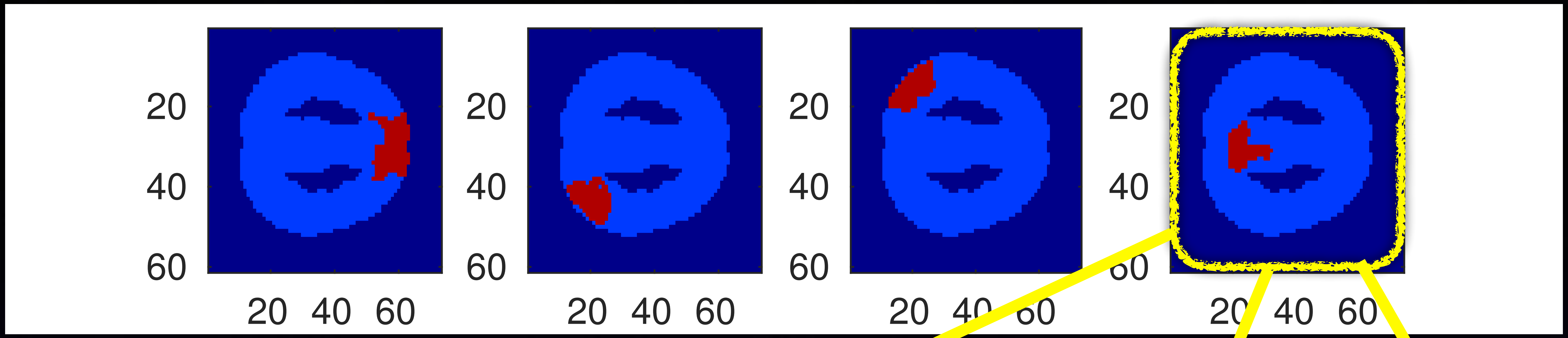


Global Noise

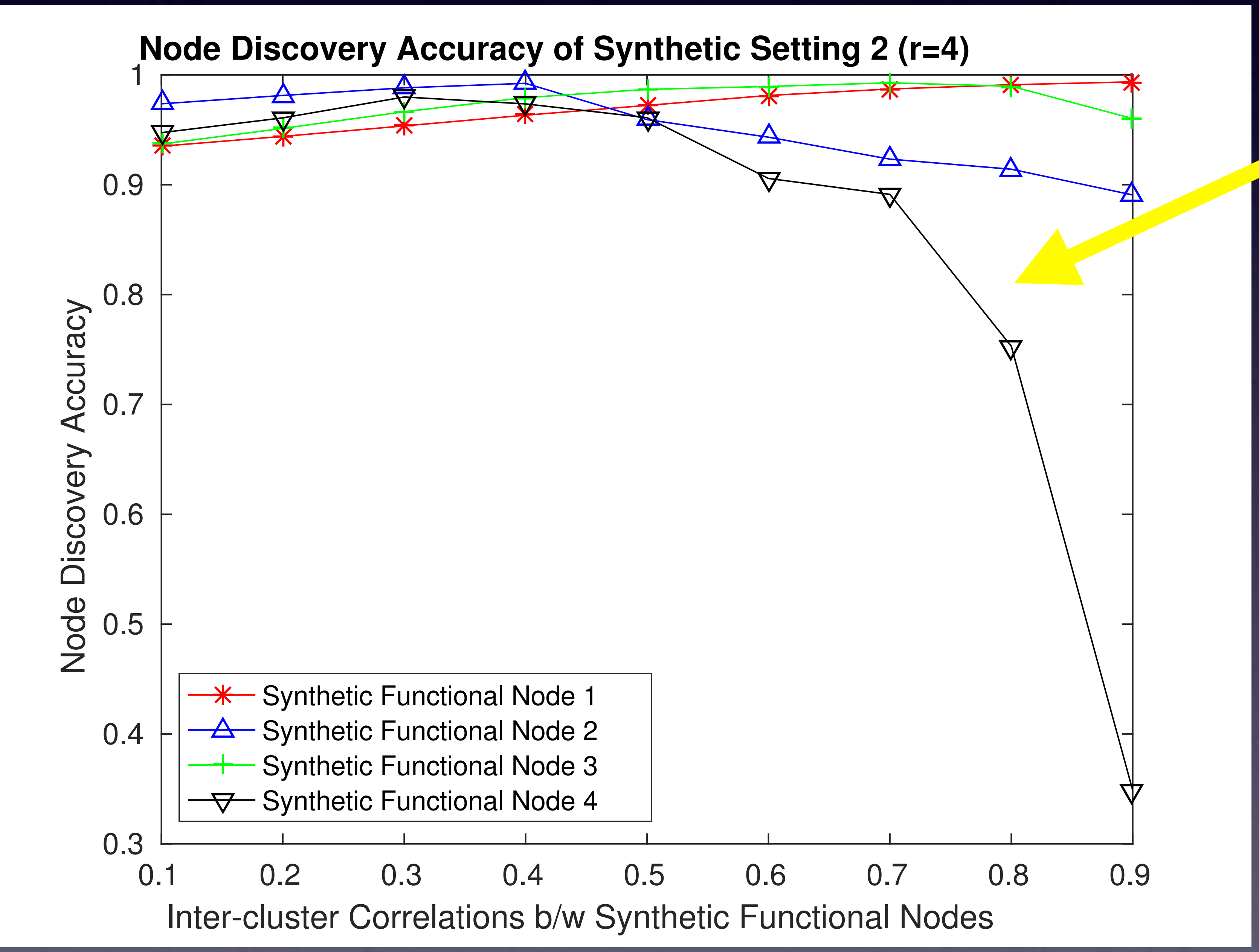


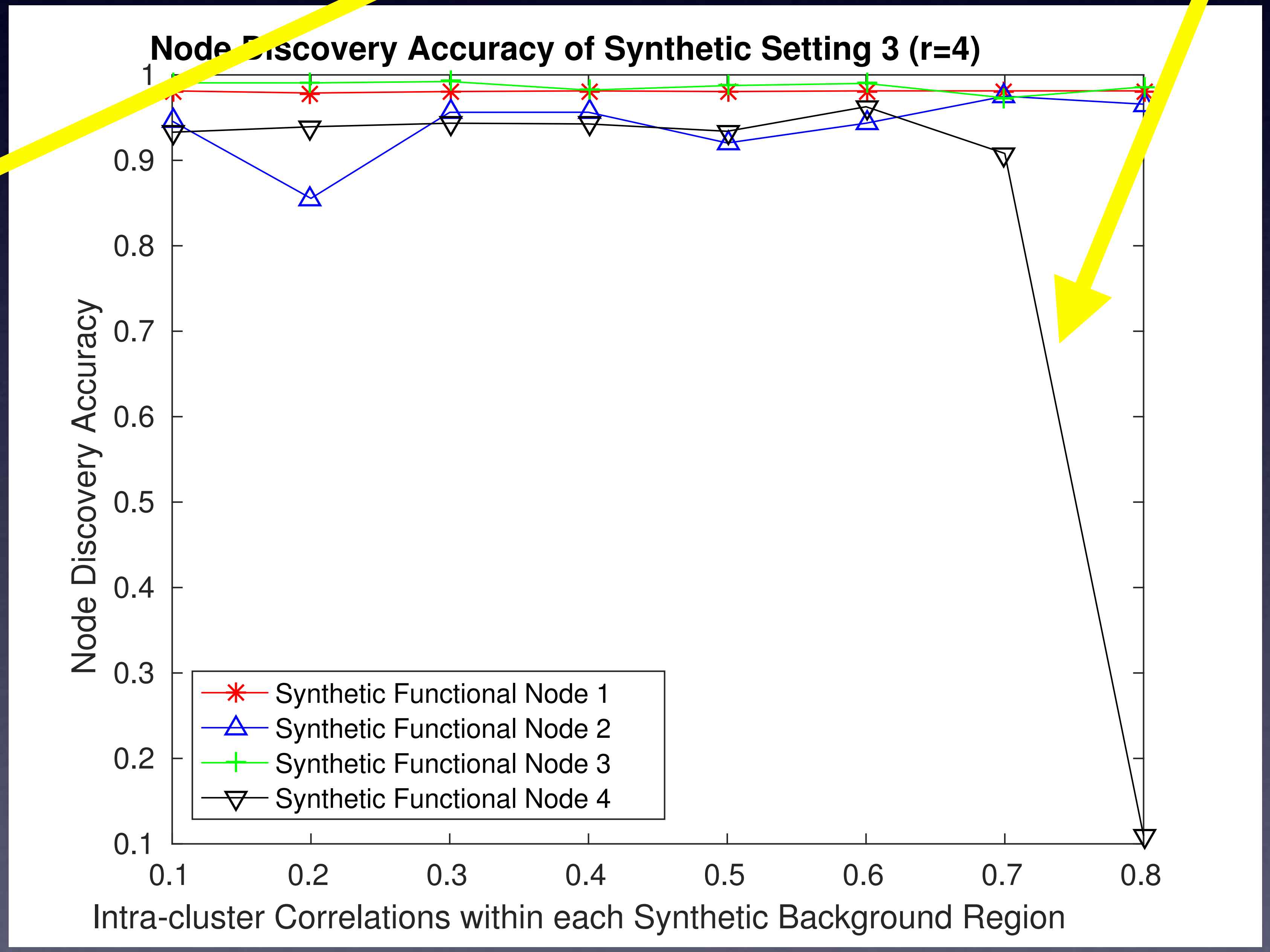
Ground Truth

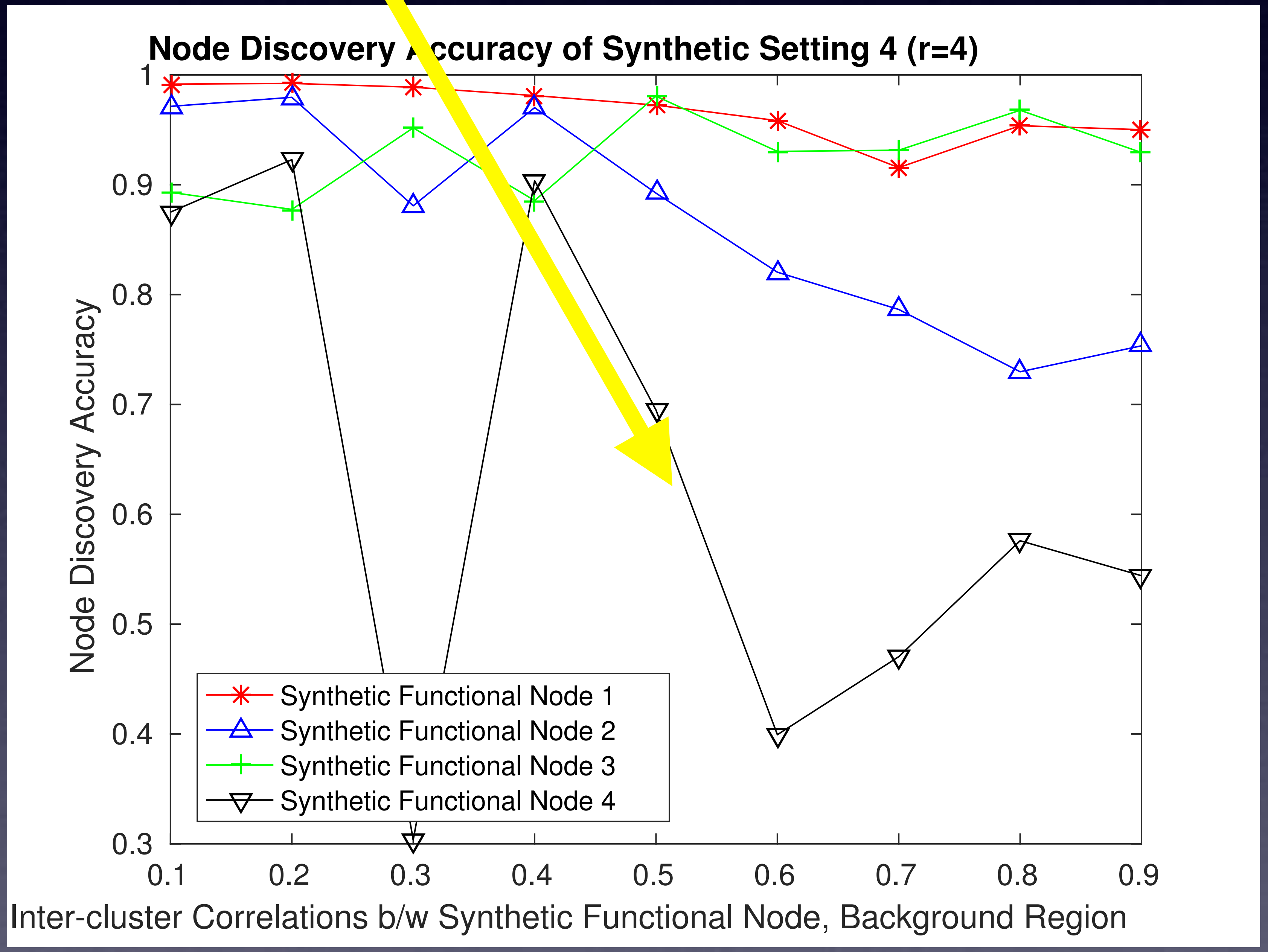




Degrading | Local Noise | Global Noise







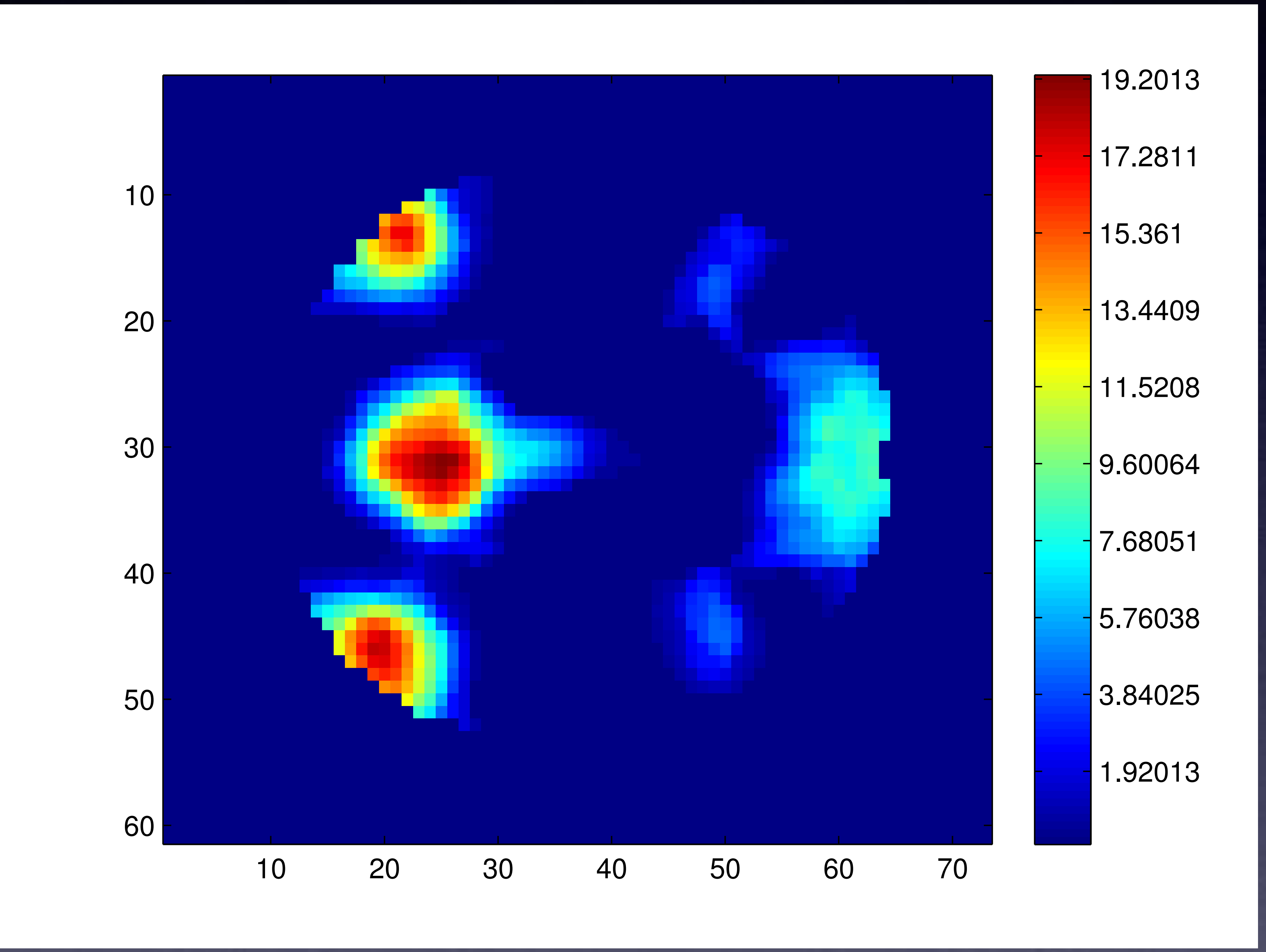
Why is Node No. 4's difficult to discover:
Irregular shape + surrounded by background nodes/regions

Experiments: ADNI fMRI

Young

Elderly Normal

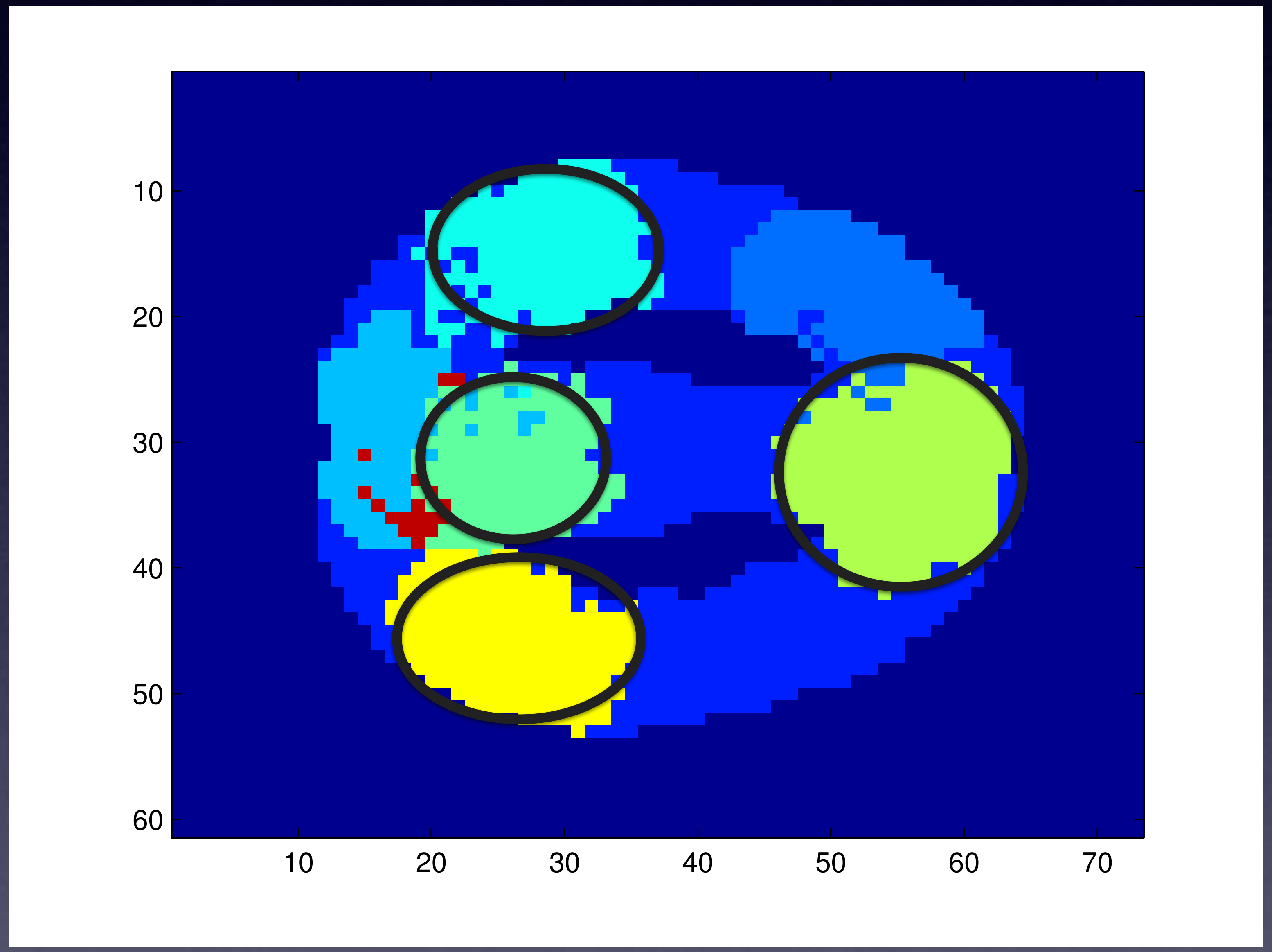
Demented



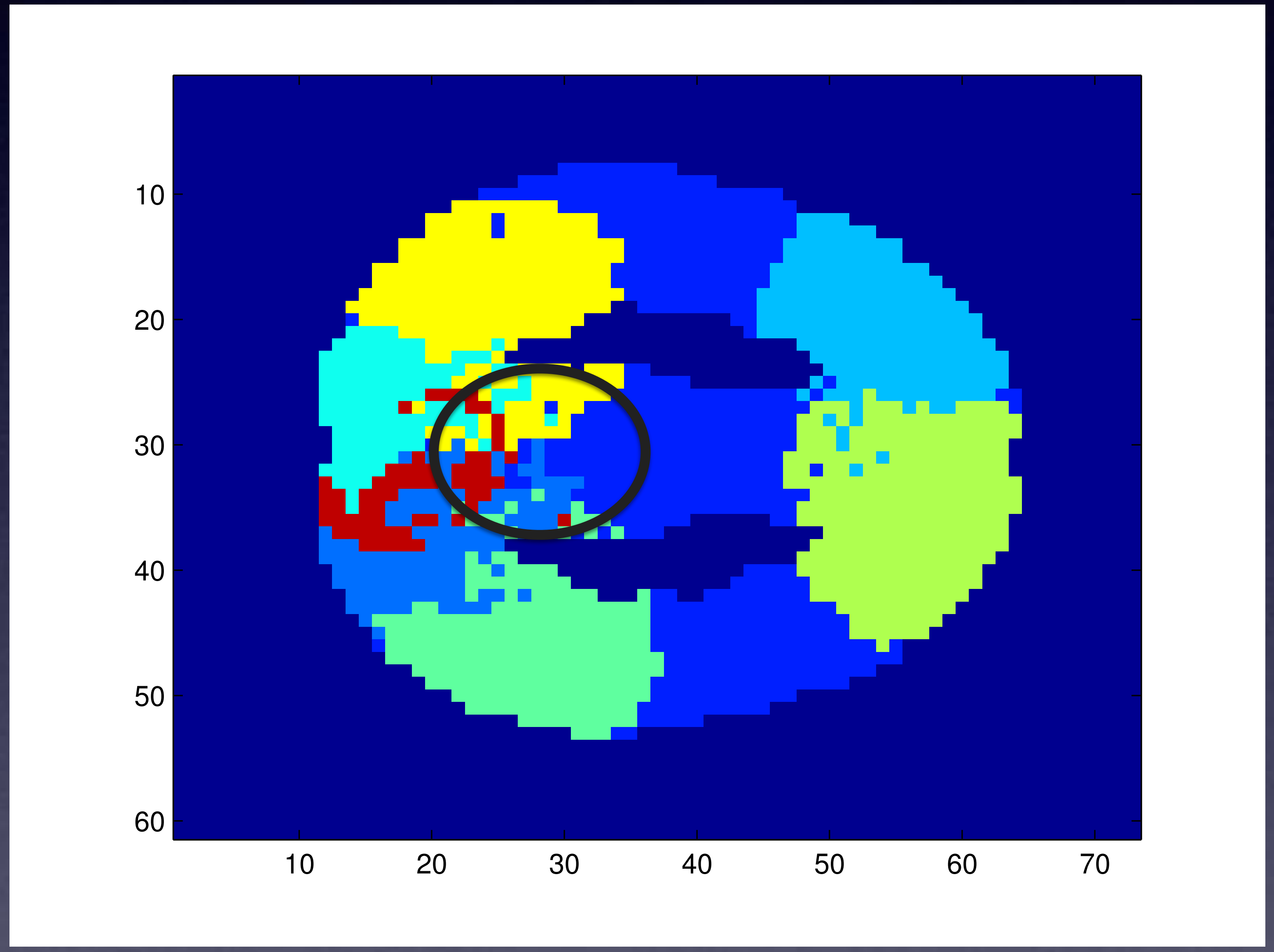
Foreground activation network:
Resting-mode Default Mode Network

The clinical interpretations by LCDR Walker and Doctor Tschiffely

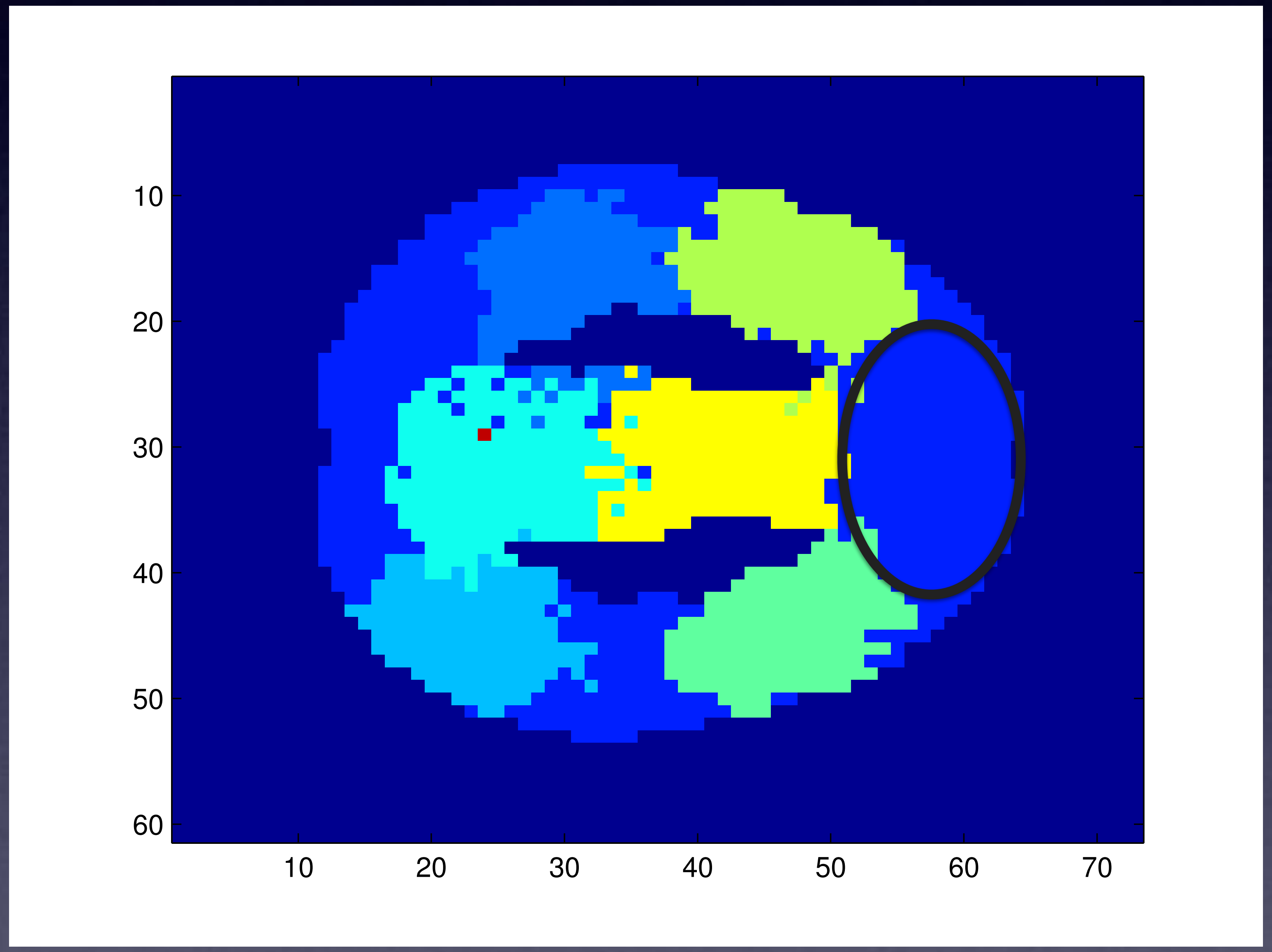
Experiments: ADNI fMRI



Young



Elderly Normal

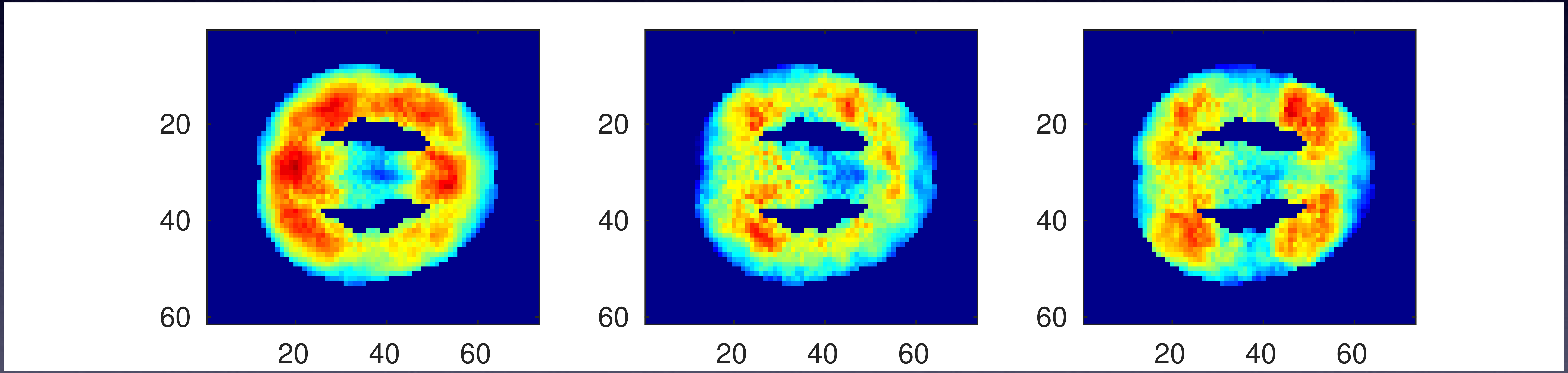


Demented
(Alzheimer Affected)

KDD2017

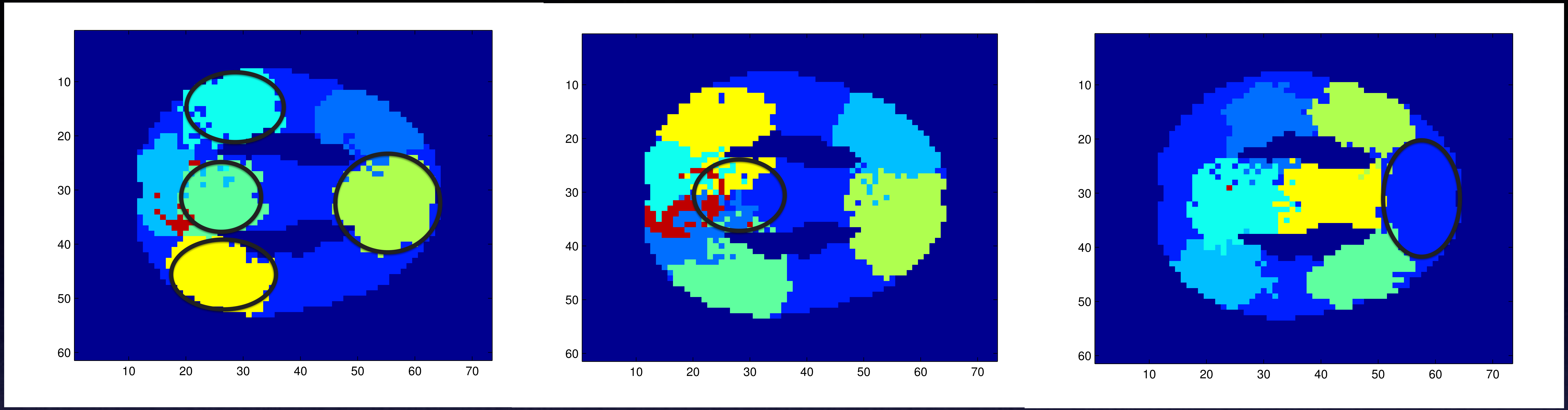
Experiments: ADNI fMRI

Average

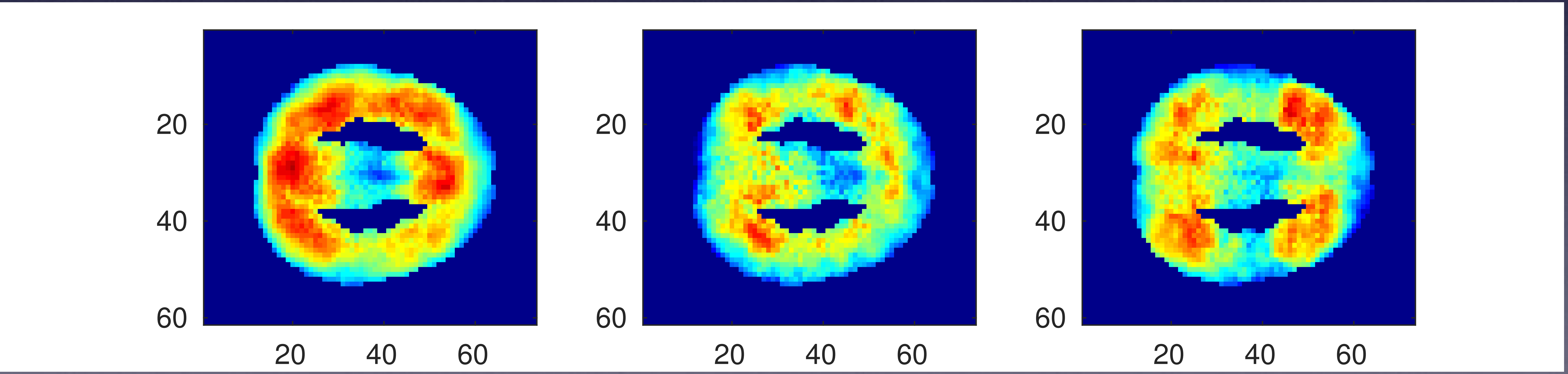


Young Elderly normal Demented

KDD2017



Young Elderly normal Demented



Future Work

- More than just resting-mode
  - Task-driven fMRI scans
  - Explore other underlying networks
- Other domains
  - Astronomical, climatic, geographical, etc
- New variations of our formulation
  - Stacks of graphs (RESCAL style model)
  - Constrain **M** to be binary

KDD2017

Acknowledgement

NSF Grant IIS-1422218

‘Functional Network Discovery for Brain Connectivity’

The Henry M. Jackson Foundation Grant

‘Small World and Other Graph Properties in Brains’

THANK YOU!

Questions?

KDD2017