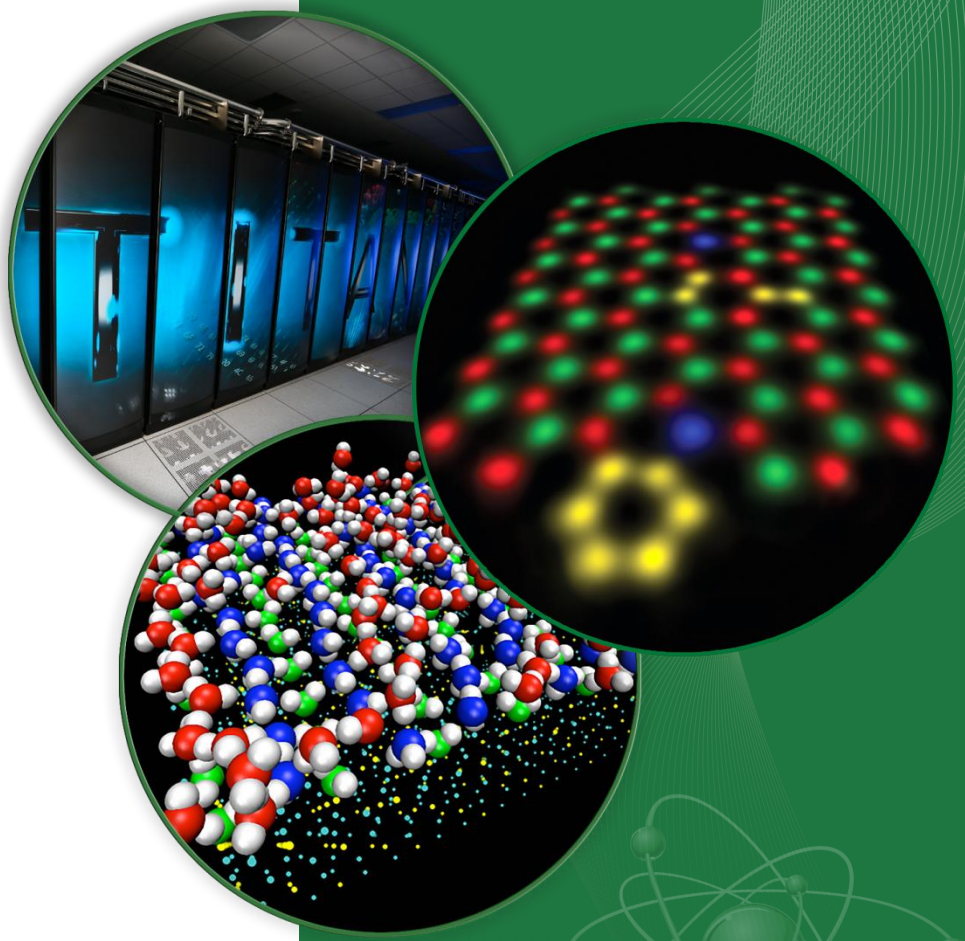
Multivariate Methods Applied to 4D STEM data

Stephen Jesse

OAK RIDGE National Laboratory | CENTER FOR NANOPHASE MATERIALS SCIENCES

INSTITUTE FOR FUNCTIONAL IMAGING OF MATERIALS
OAK RIDGE NATIONAL LABORATORY

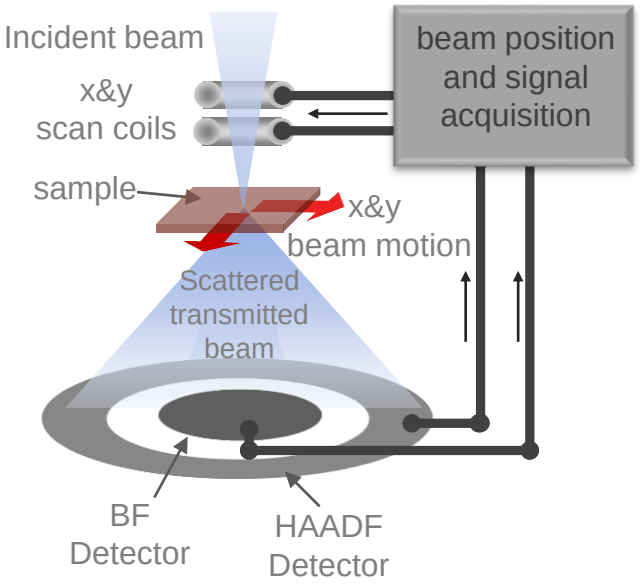


ORNL is managed by UT-Battelle
for the US Department of Energy

OAK RIDGE National Laboratory

Introduction

- The standard STEM configuration is exquisitely tuned to capture material structure (in projection)
- However, the transmitted electrons carry far more information than is captured by monolithic detectors



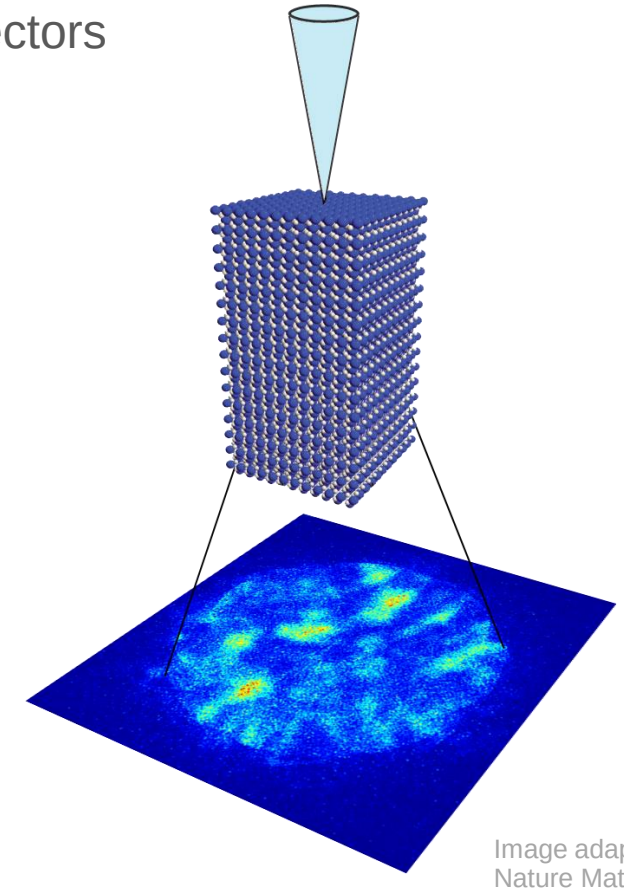
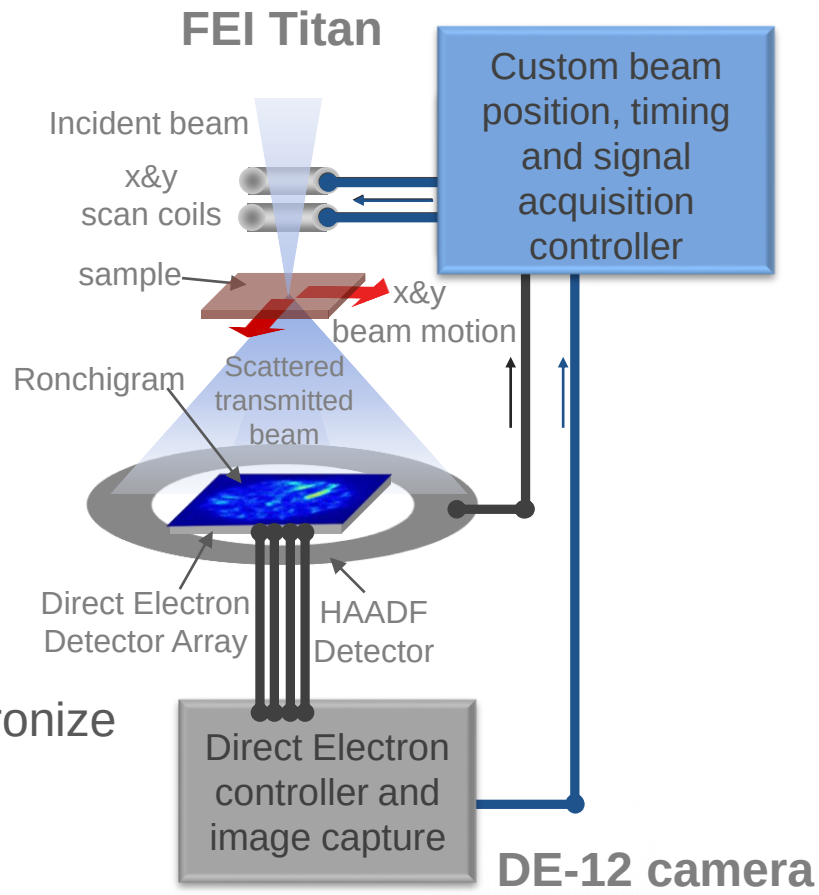


Image adapted from:
Nature Materials 10, 270–271 (2011)

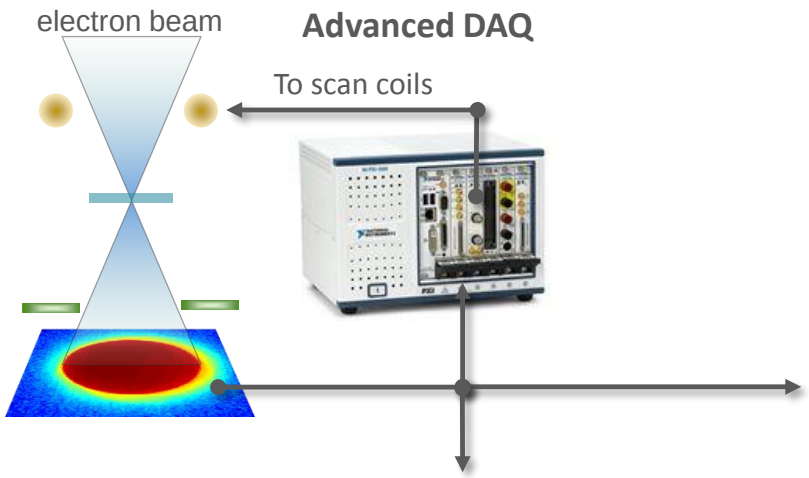
Introduction

- The standard STEM configuration is exquisitely tuned to capture material structure (in projection)
- However, the transmitted electrons carry far more information than is captured by monolithic detectors
- Recent advances in detectors has enabled rapid and sensitive capture of Ronchigrams
- Developed FPGA based controller to synchronize beam positioning with Ronchigram capture enabling:
  - Ronchigram capture at ~400 frames/sec
  - 192X192 pixels per image with 384X384 pixels per Ronchigram
  - 4-dimensional data set (4 GB)



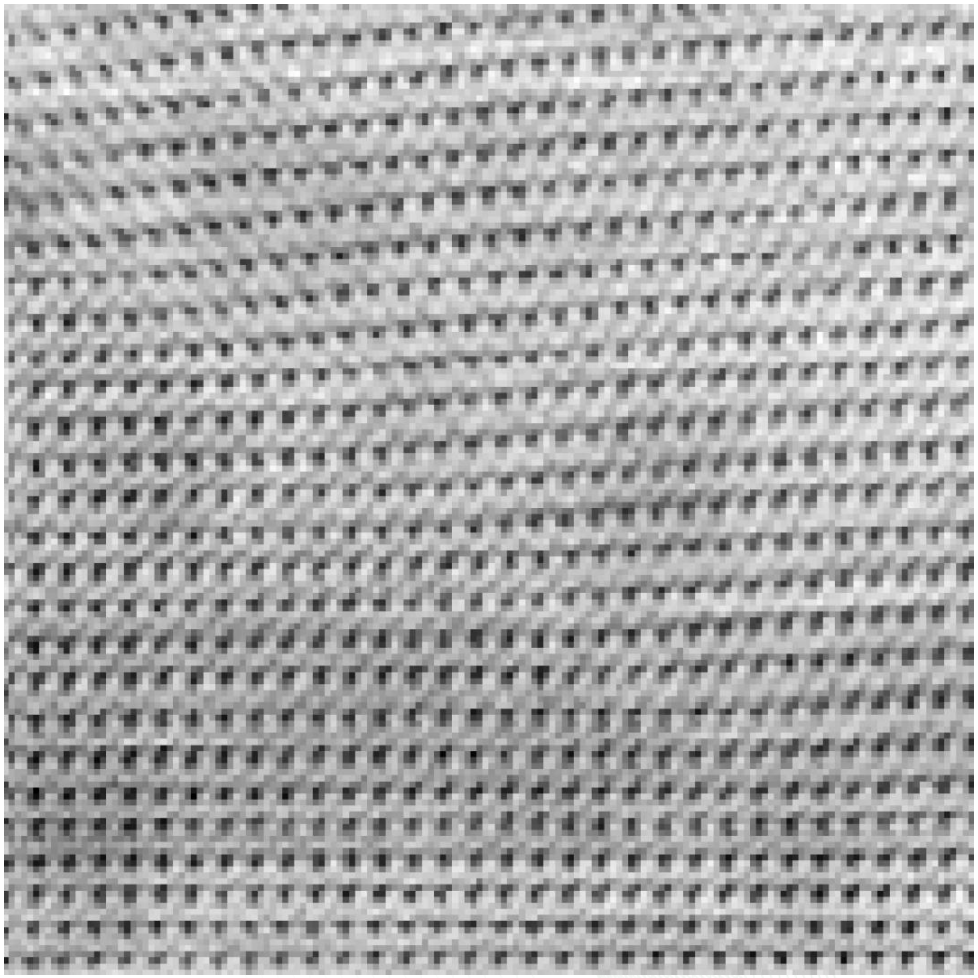
How can we explore this data to yield new and useful information?

Ptychography Data Capture



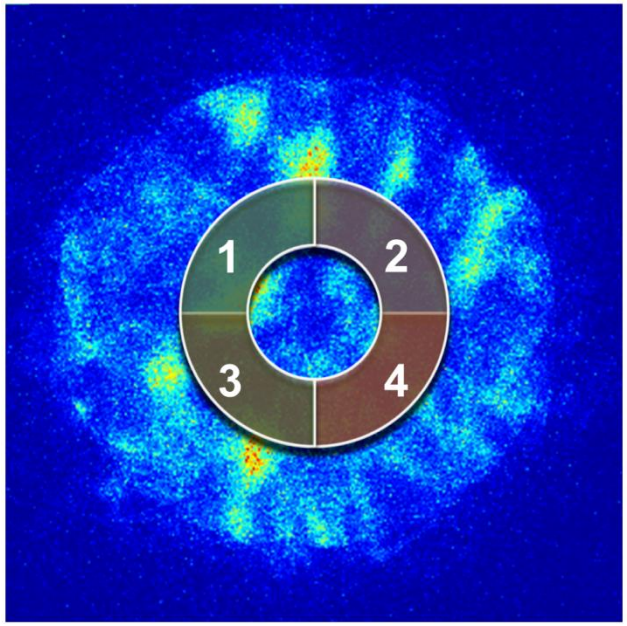
Bismuth Ferrite (Ferroelectric) Domain Boundary

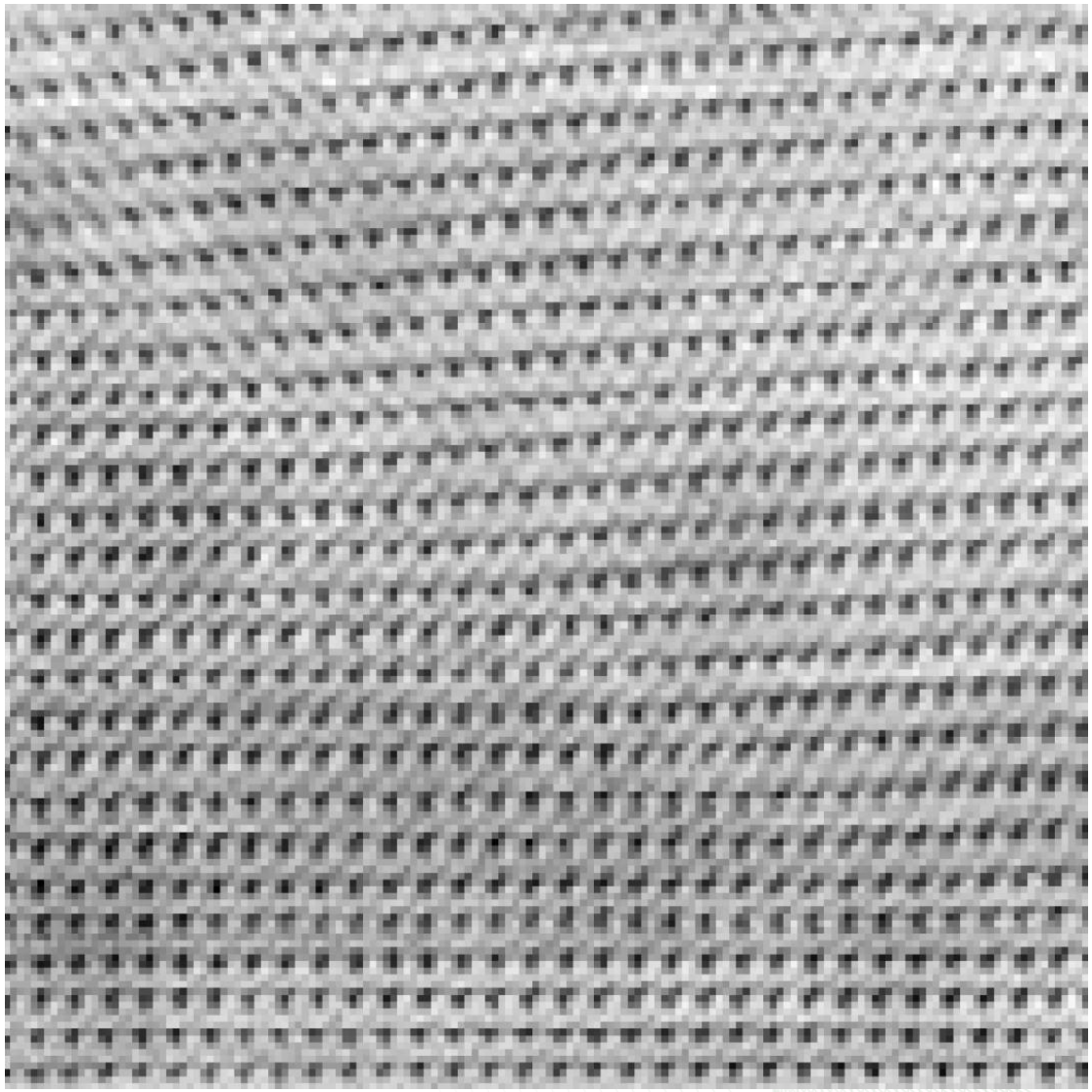
Average intensity of each Ronchigram ~ BF image



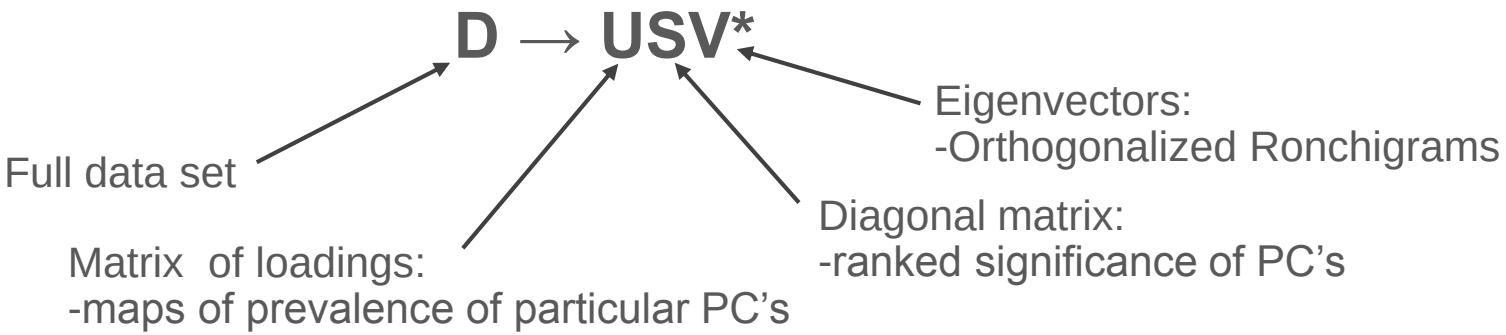
Arbitrary Detector Shape

We can simulate arbitrary and optimal detectors after the data has been captured





Principal Component Analysis (PCA)

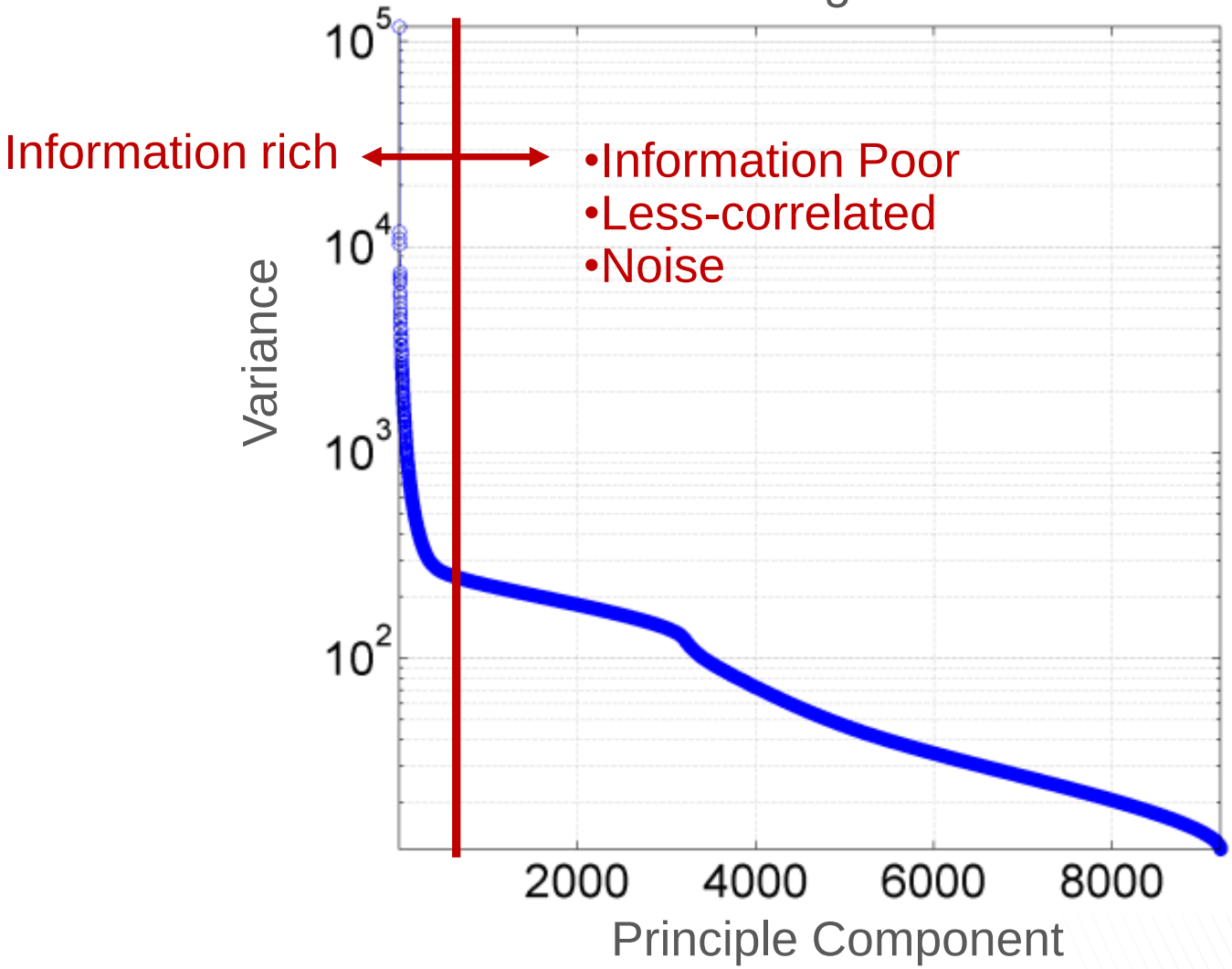


- Compares every individual observable with every other observable
- Find the most common behavior
- Find the next most information rich (orthogonal) variation on the above behavior
- Find the next most information rich (orthogonal) variation on the above behavior
- … loop

Principal Component Analysis (PCA)

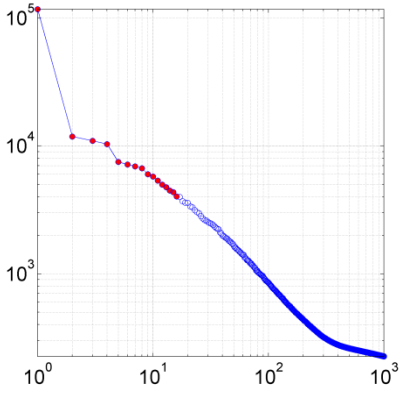
$D \rightarrow USV^*$

Scree Plot -ranked significance of PC's

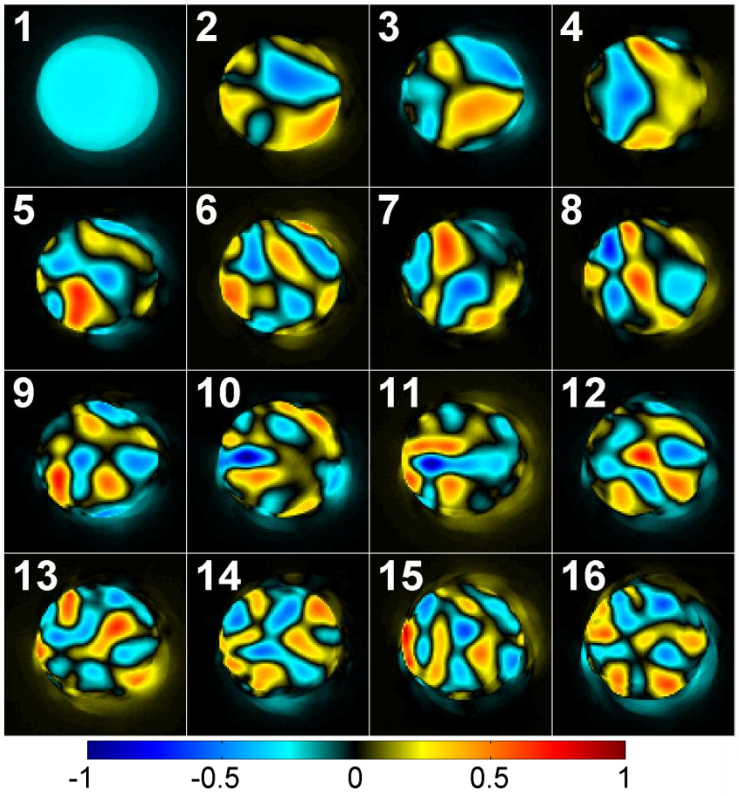


Principal Component Analysis (PCA)

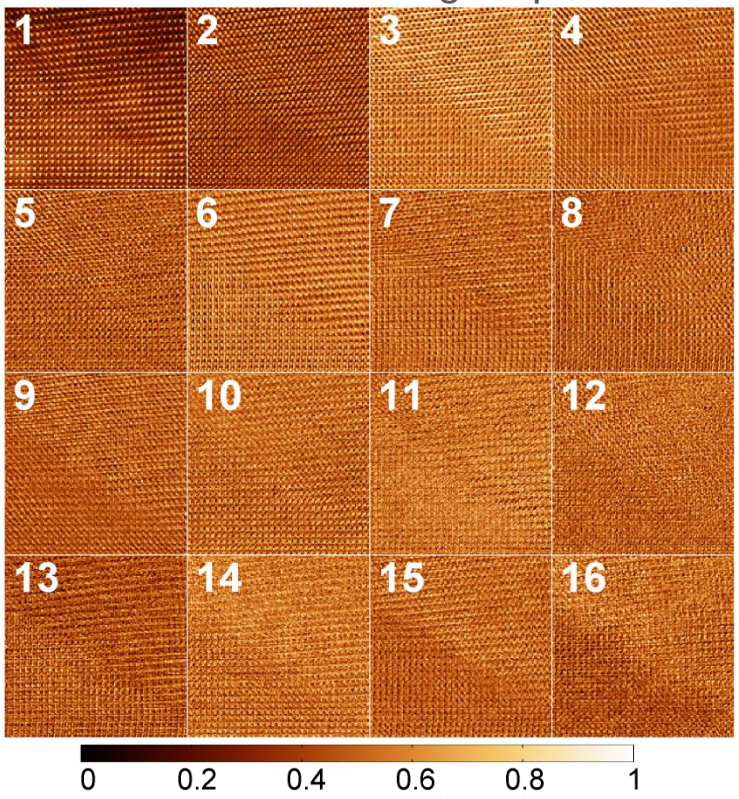
Scree Plot



First 16 eigenvectors

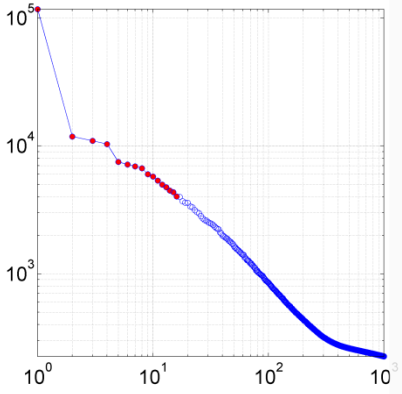


First 16 loading maps

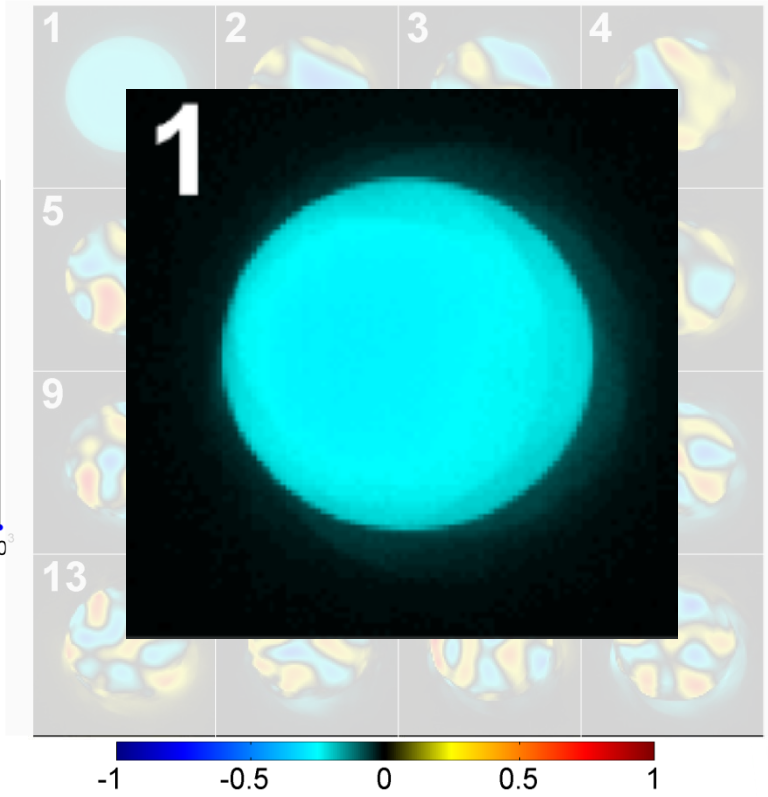


Principal Component Analysis (PCA)

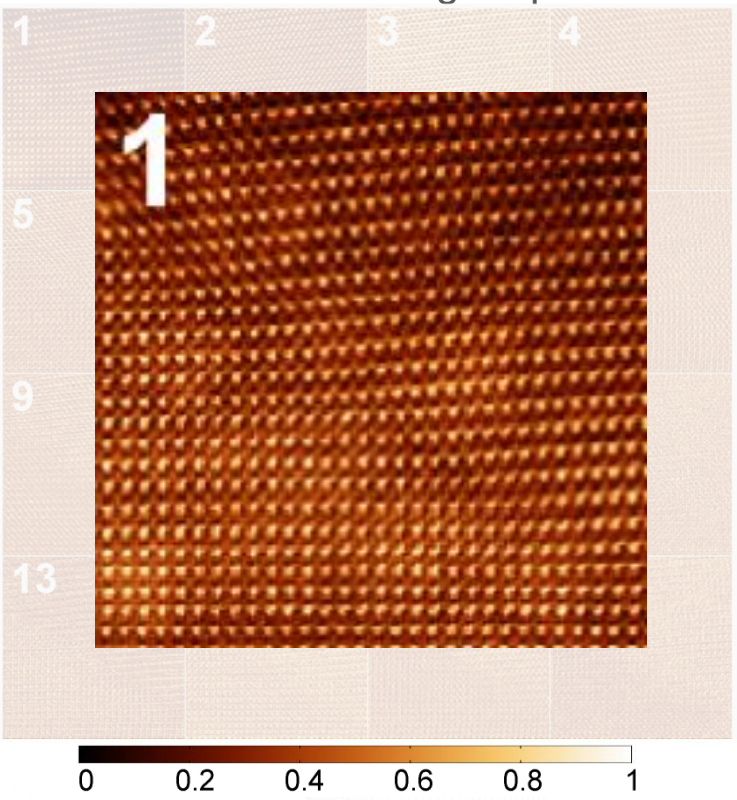
Scree Plot



First 16 eigenvectors

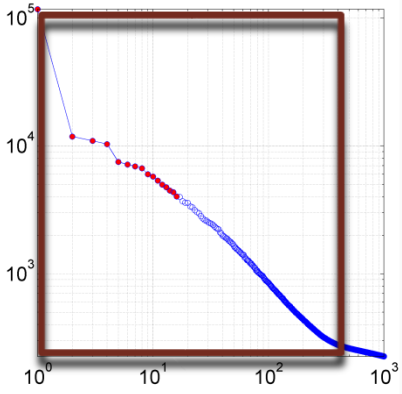


First 16 loading maps

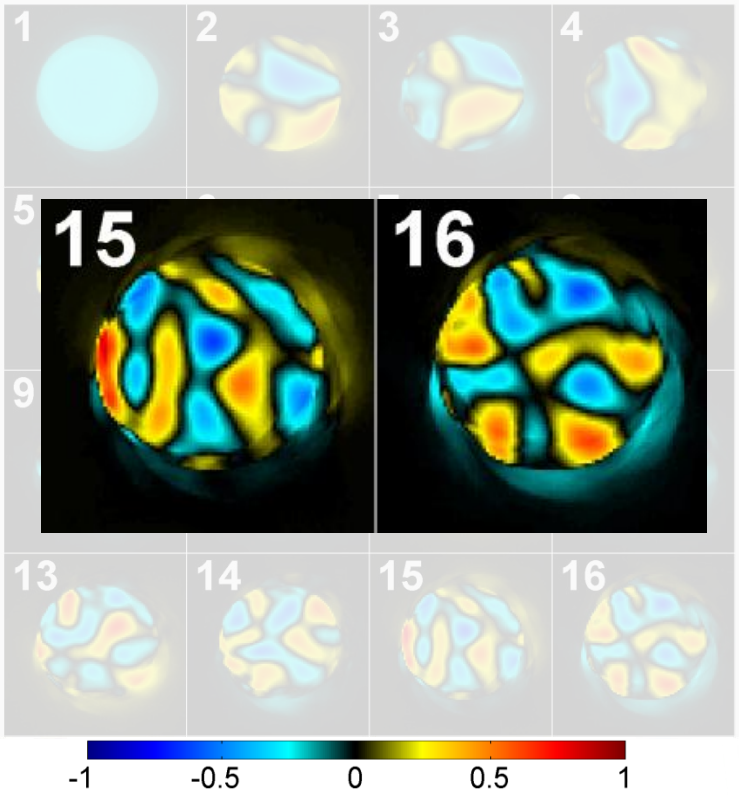


Principal Component Analysis (PCA)

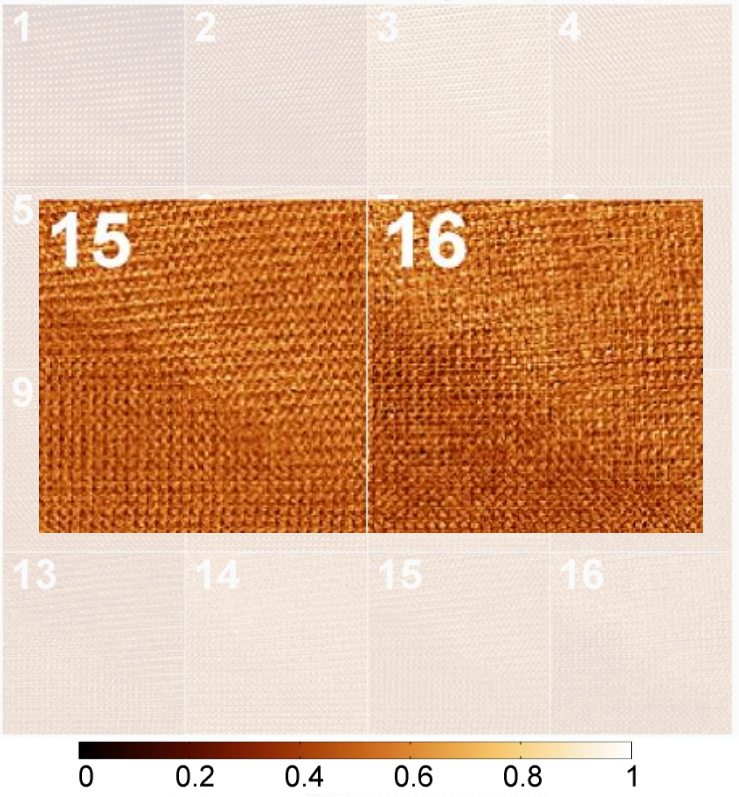
Scree Plot



First 16 eigenvectors



First 16 loading maps

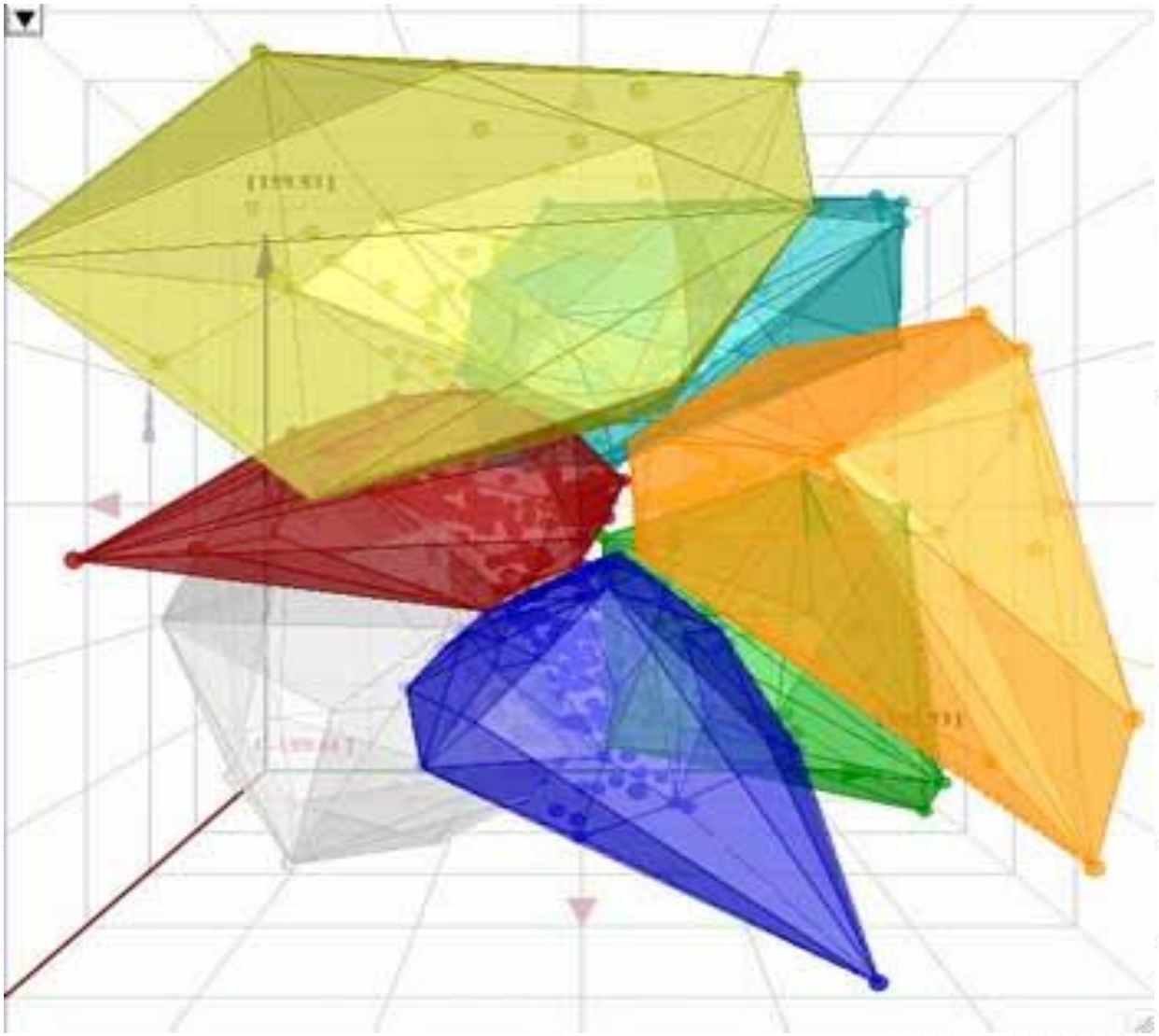


PCA typically fails at: elucidating underlying physics

PCA typically succeeds at:
- being fast and efficient
- de-noising
- data compression

Cluster analysis

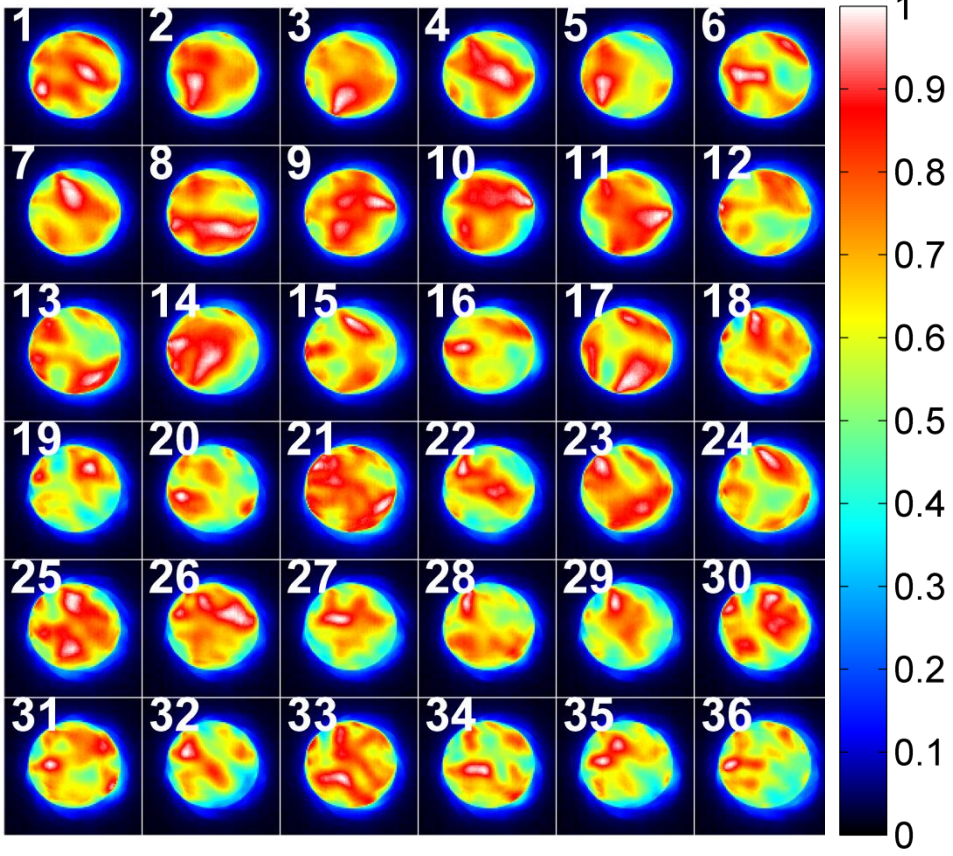
Group observables by “distance” to one another



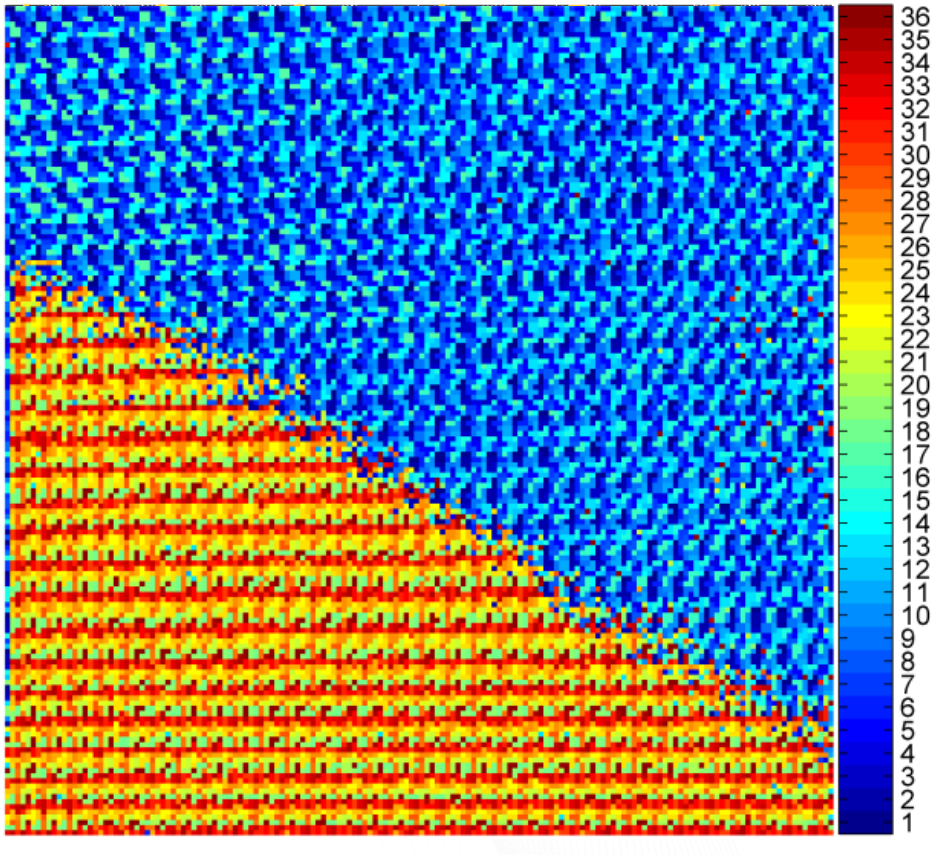
Cluster analysis

K-means clustering: 36 Clusters

Ronchigram corresponding to each cluster

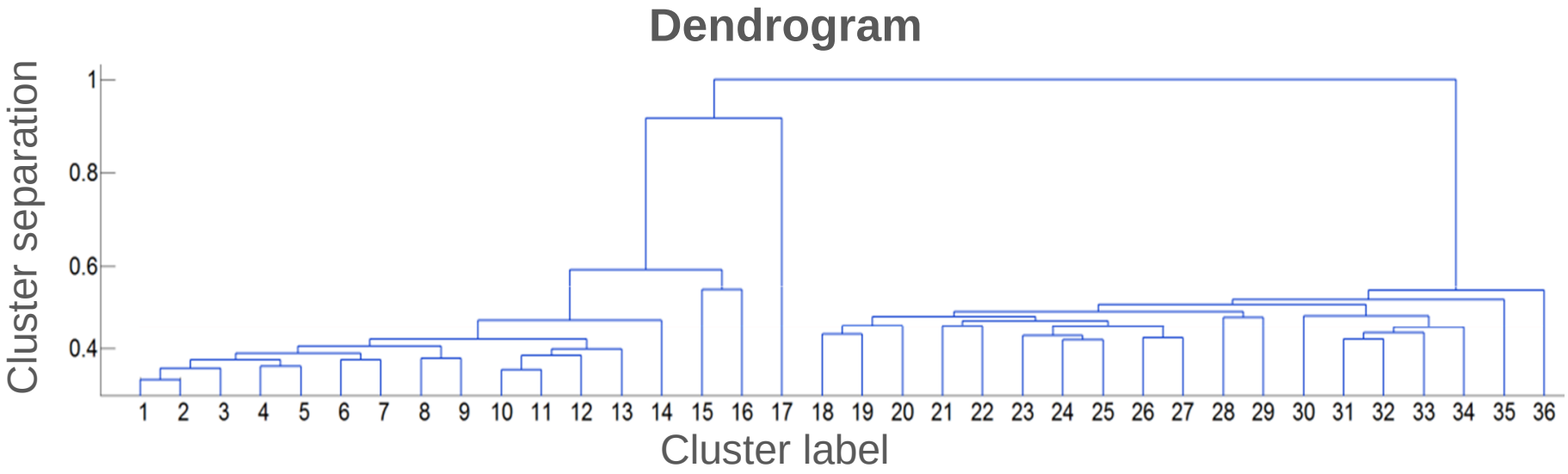


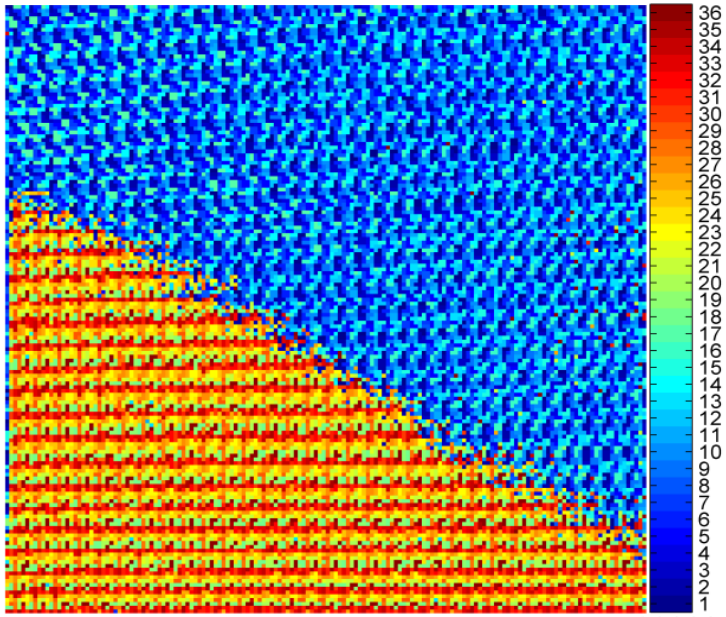
Cluster Locations



Cluster analysis

How many clusters is the right number of clusters?





Cluster analysis

How many clusters is the right number of clusters?