Finding Neural Codes Using Random Projections

Brendan Mumey[*]

Aditi Sarkar[#]
Tomas Gedeon[+]
Alex Dimitrov[#]
John Miller[#]

Departments of [*]*Computer Science,*
[#]*Cell Biology and Neuroscience,* and
[+]*Mathematical Sciences,*
Montana State University - Bozeman

Support: NSF BITS grant

Outline

- **Coding Problem:** find the neural code that maximizes mutual information between input and output

- **Previous Work:** local search methods for nonlinear temporal codes, single channel data

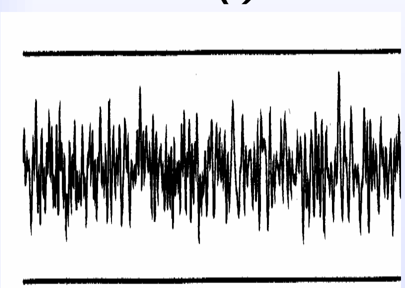
- **A New Approach:** Random Projections

- **Preliminary Experimental Results**

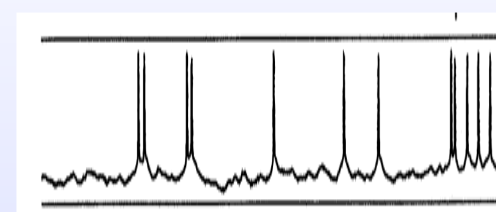
Background: The Coding Problem

- How does neural activity represent information about environmental stimuli?

stimulus
X(t)



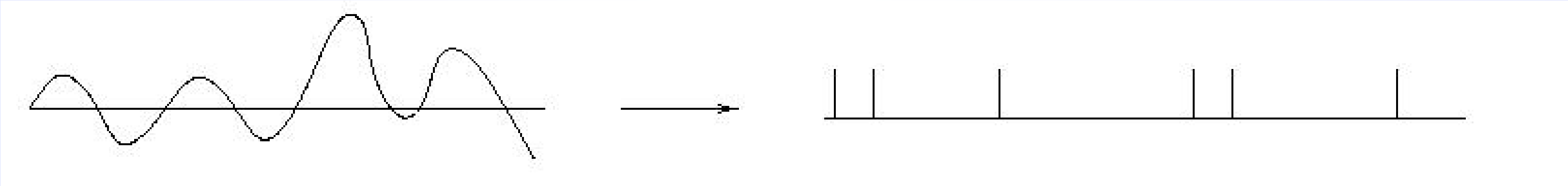
response
Y(t)



Stochastic Framework

Random Variables:

- X stimulus (waveform)
- Y neural response (single channel for now)



stimulus *X=x*

neural response *Y=y*

… (X,Y) forms a joint distribution

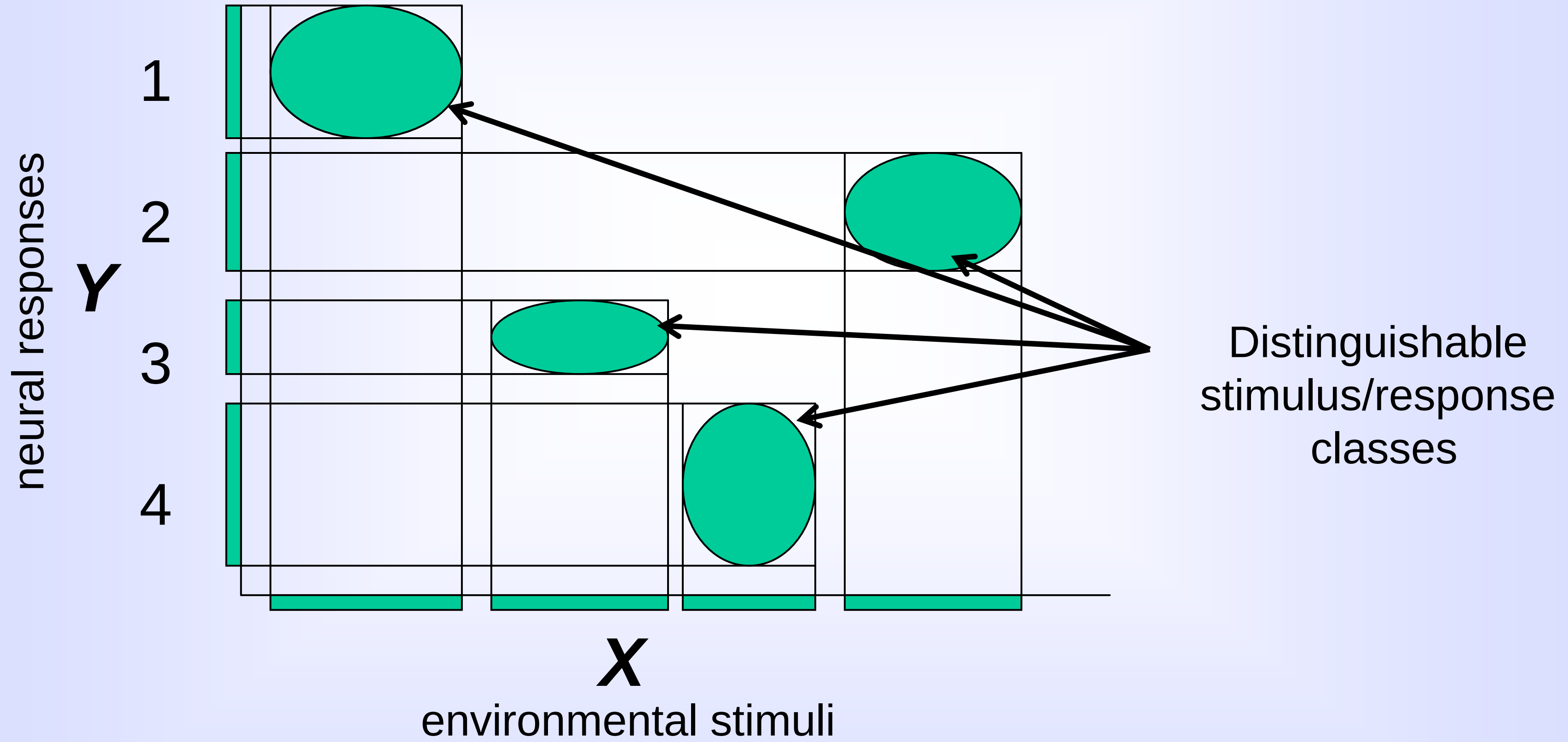
Partitioning the Joint Space

- **Problem**: partition (X,Y) into discrete classes (X_N, Y_M) so as to maximally preserve *mutual information:*

$$I(X_N, Y_M) = \sum_{i,j} \Pr(X_i, Y_j) \log \frac{\Pr(X_i, Y_j)}{\Pr(X_i)\Pr(Y_j)}$$

- <u>NB</u>: this is an *NP-complete* problem
- *Assumption*: there are a small number of distinguishable clusters in the joint space

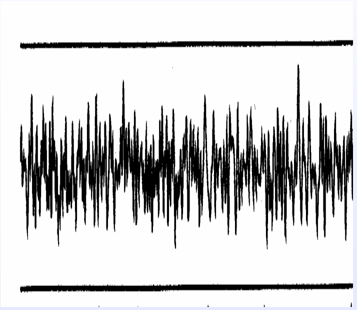
Example Partition



The Curse of Dimensionality

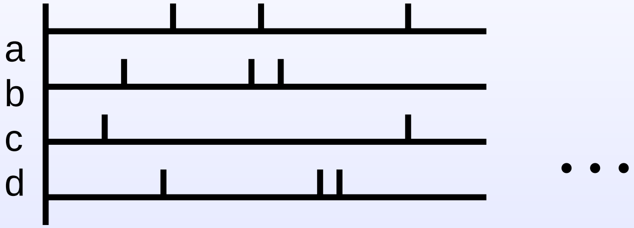
- Both X (input stimuli) and Y (neural response) can be very high dimensional…

X



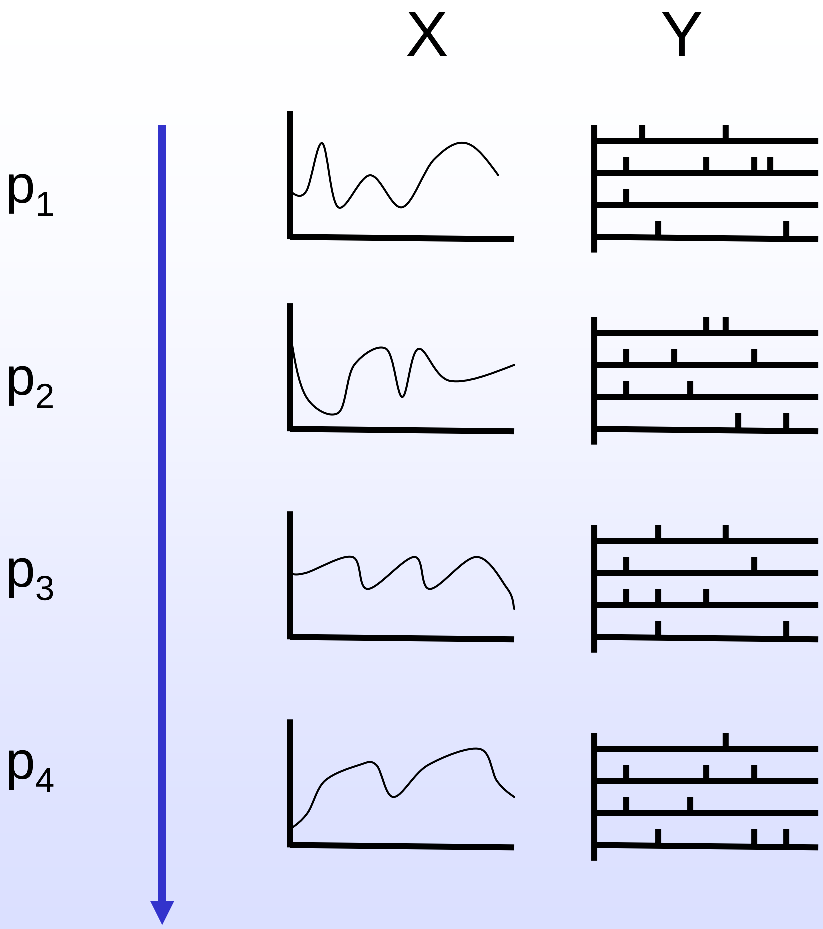
stimulus waveform represented as a vector in $R^{>100}$

Y



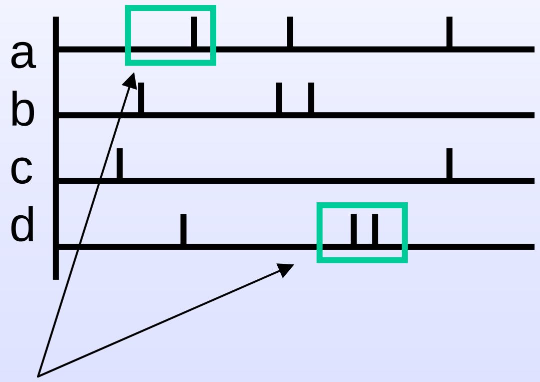
response may be multiplexed across several channels… may require several hundred bits if encoded in binary

Observed Data



Hidden Codes

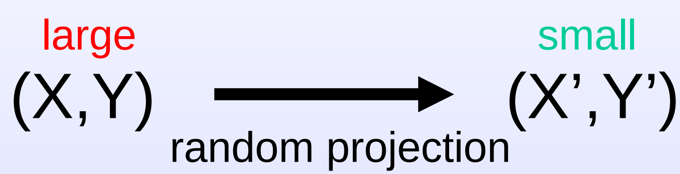
- The optimal partitioning may be based on subtle signals in the data.
- It is difficult to know what features of the data are relevant beforehand…



e.g. perhaps this pattern represents
an important response codeword

Random Projections

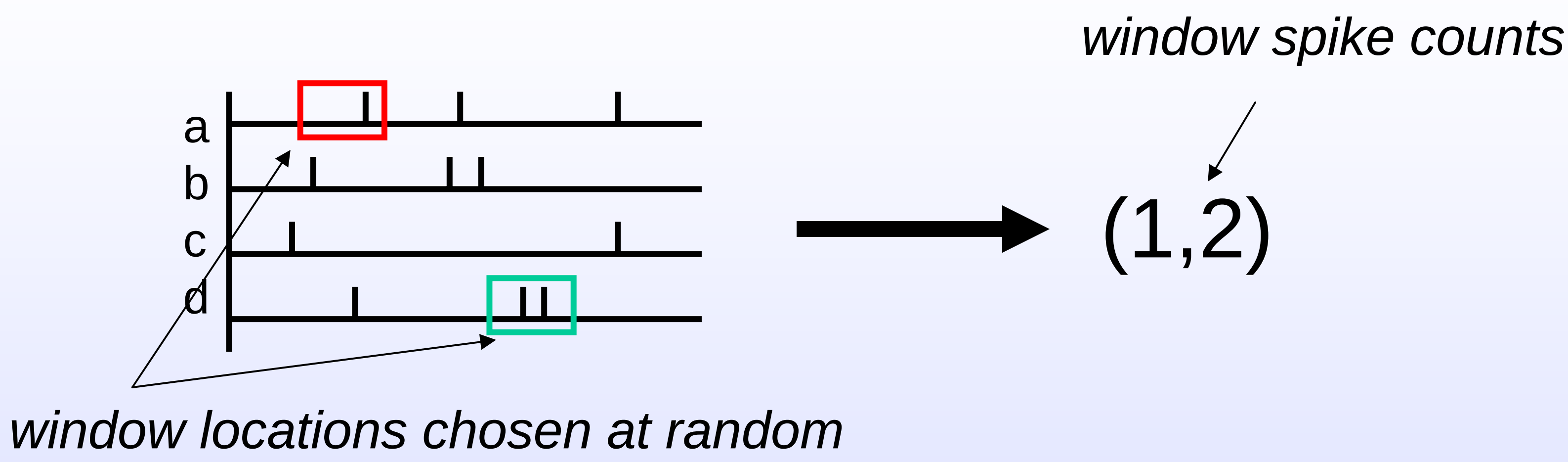
- **Idea**: use random projections to measure “similarity” between observed data points…
  (originally proposed by Indyk et al. for computational geometry problems)
- **Outline**: project to a much smaller space; check for patterns in the projected space…



- …repeat many times

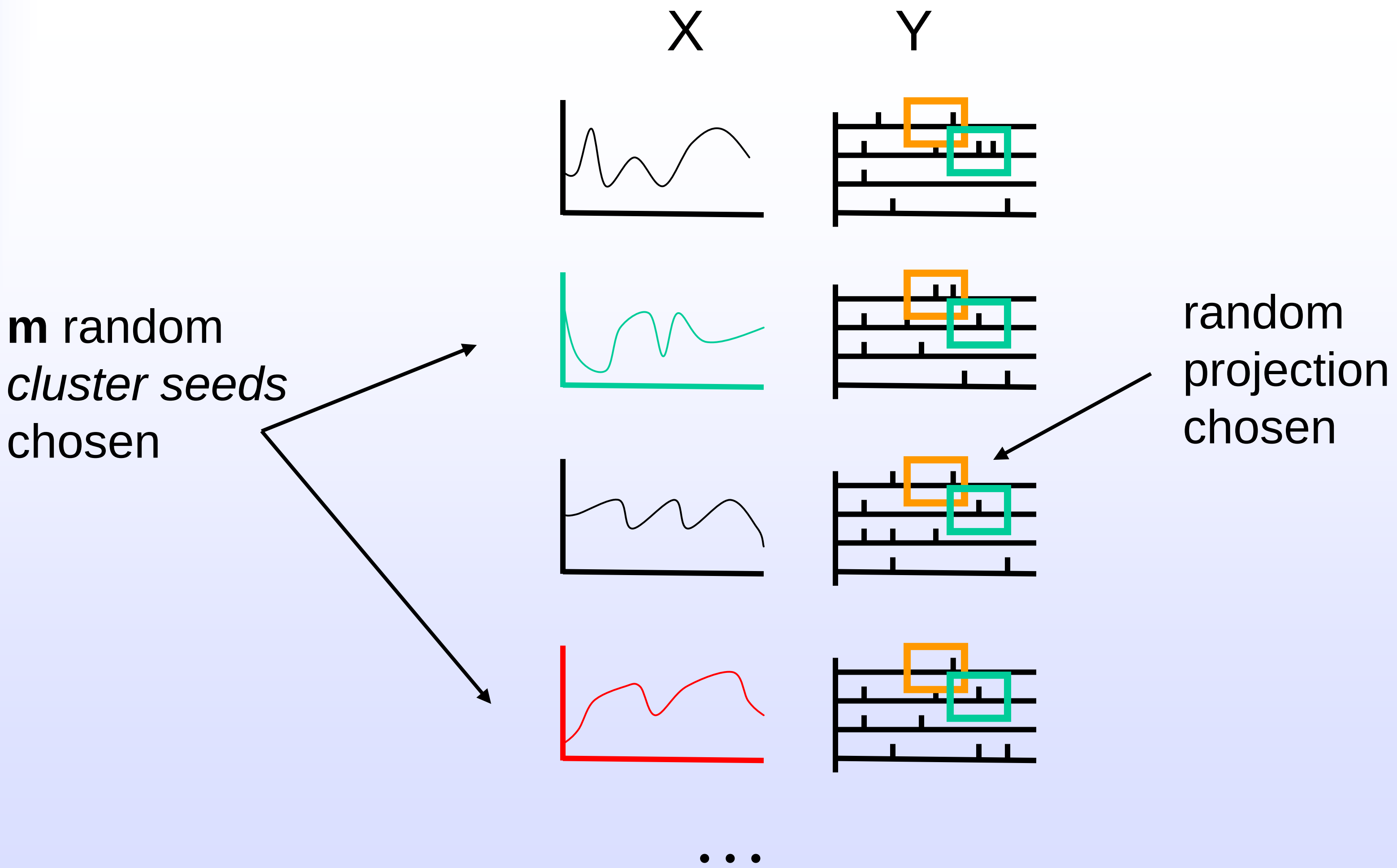
Choosing Projections

- Projections are drawn at random from projection function families:

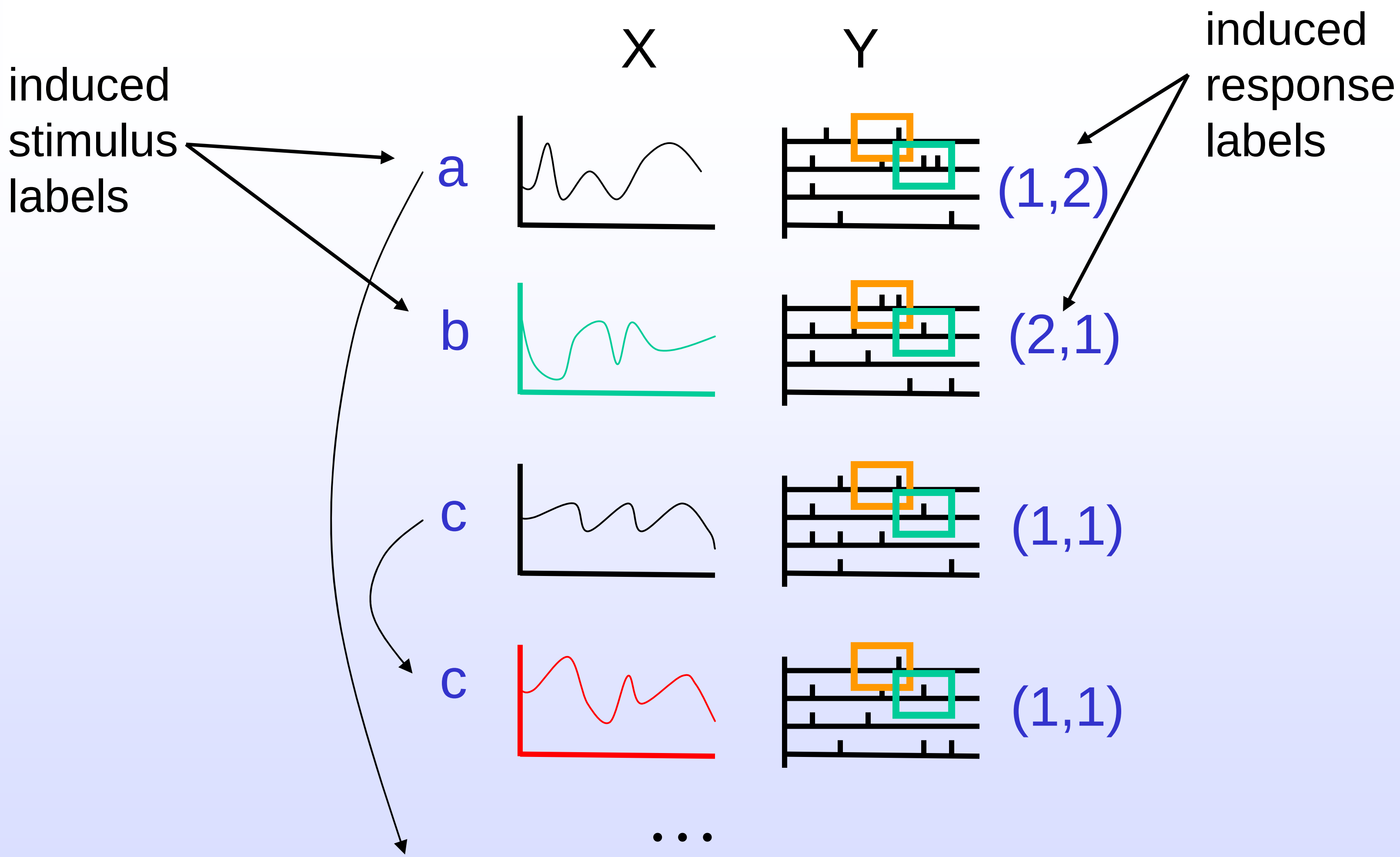


- Less restrictive than pre-imposed metric based methods

Example



Label Assignment



Incorporate Mutual Information

Labeled joint points:

p_1 a (1,2)
p_2 b (2,1)
p_3 c (1,1)
p_4 c (1,1)
…

→ Calculate mutual information, decide how *interesting* this random projection was…

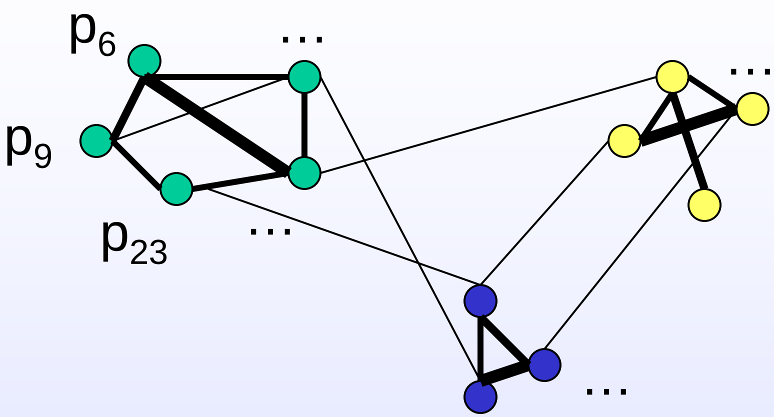
↓

If *interesting*, this provides some evidence that $\mathbf{p_3}$ and $\mathbf{p_4}$ are in the same joint cluster (distinct from $\mathbf{p_1}$ and $\mathbf{p_2}$)

Idea: repeat this many times…

Joint Similarity Graph

- Construct a weighted graph summarizing the similarity evidence found between each pair of data points over random trials:



- Look for “near-cliques” in this graph - these represent the clusters of the underlying joint distribution and so define the neural code… (we use the *CLICK* software for this)

Some Theoretical Analysis

- **Assumptions:**
  - joint space is "clustered" into distinct clusters $C_1, \ldots, C_k$
  - cluster C_i sampled with probability p_i, $\Sigma p_i = c <= 1$.
  - "Good projection family":

$$x \in C_i, y \in C_j$$

$$i = j \Rightarrow \Pr[\mathrm{proj}(x) = \mathrm{proj}(y)] \geq q$$

$$i \neq j \Rightarrow \Pr[\mathrm{proj}(x) = \mathrm{proj}(y)] \leq r$$

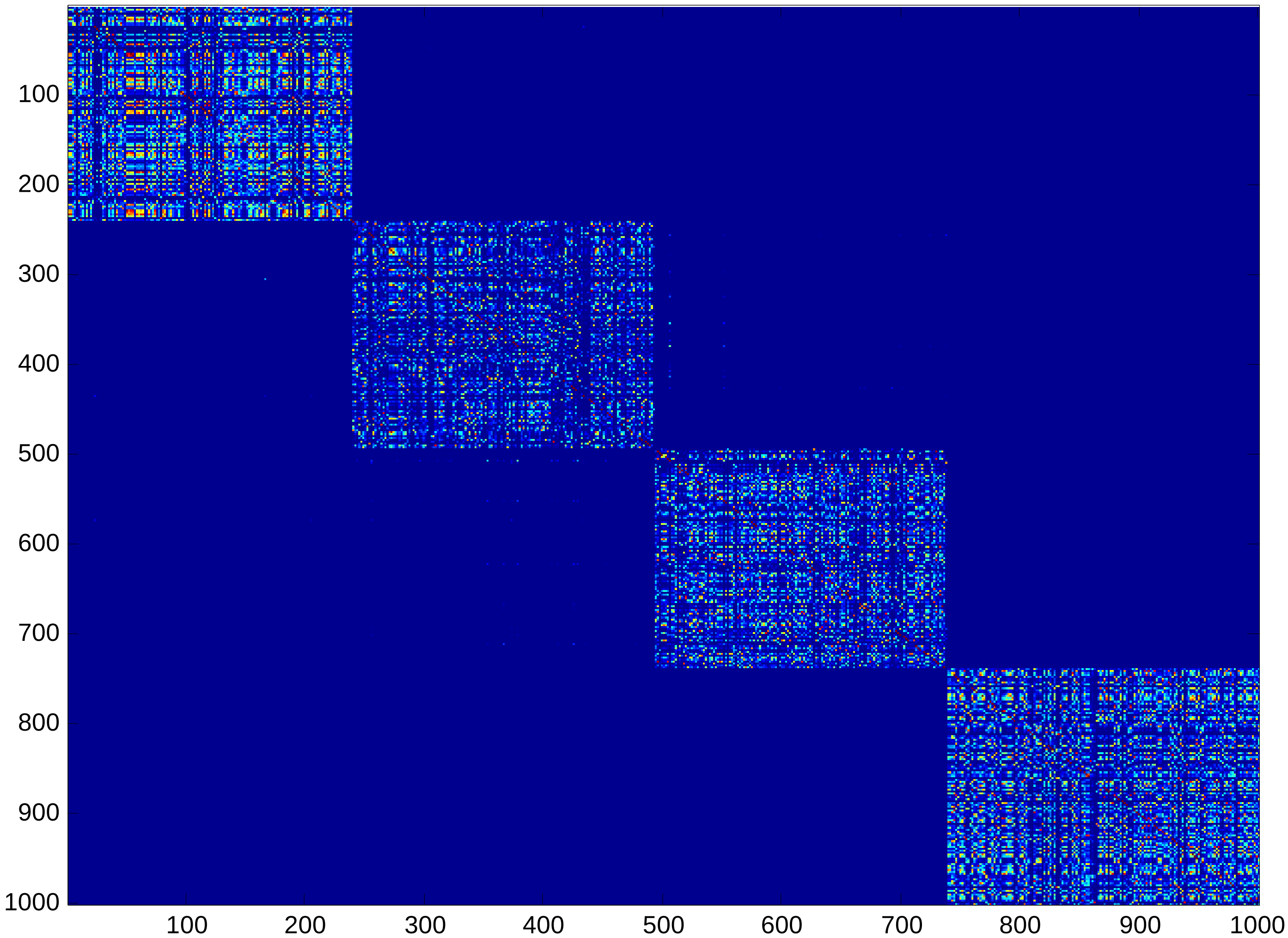
$$q > r$$

- **One Result**: If $q-r = \Omega(1/n)$, can recover true clusters with high probability with $O(n)$ random projections (using some recent random graph clustering results)

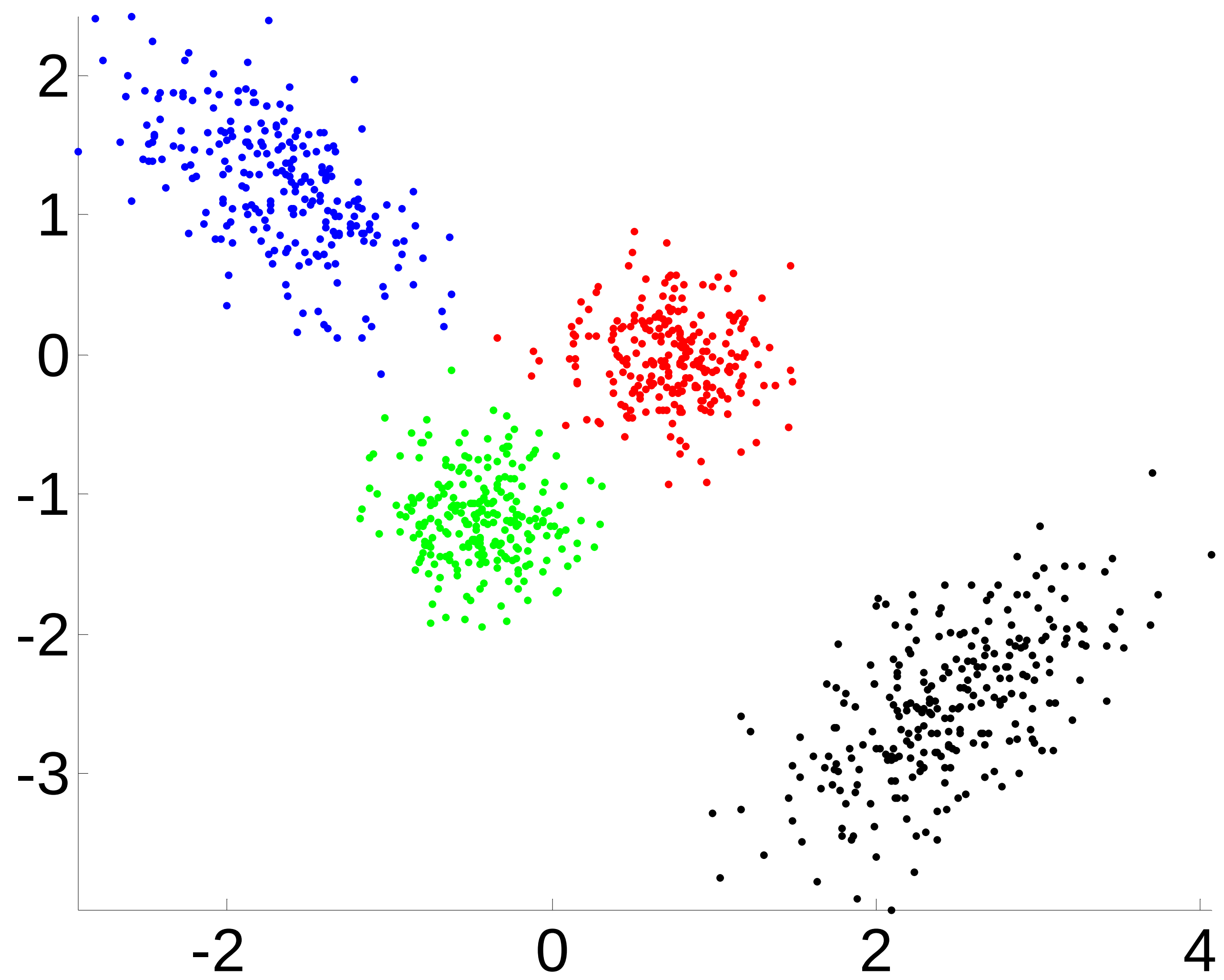
Simple Test Case

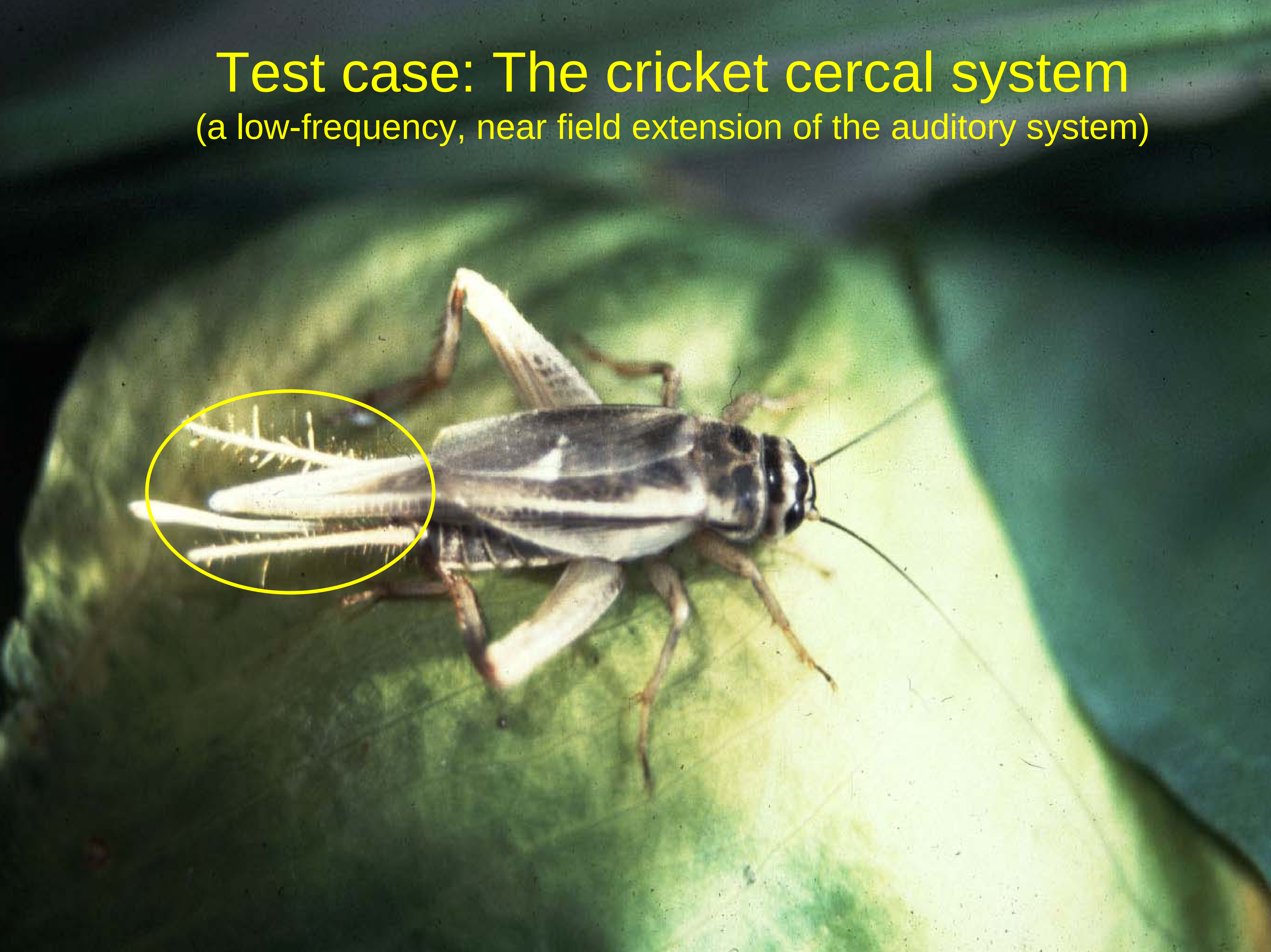
- Four Gaussian joint clusters (mixture)
- Recovers clusters perfectly…

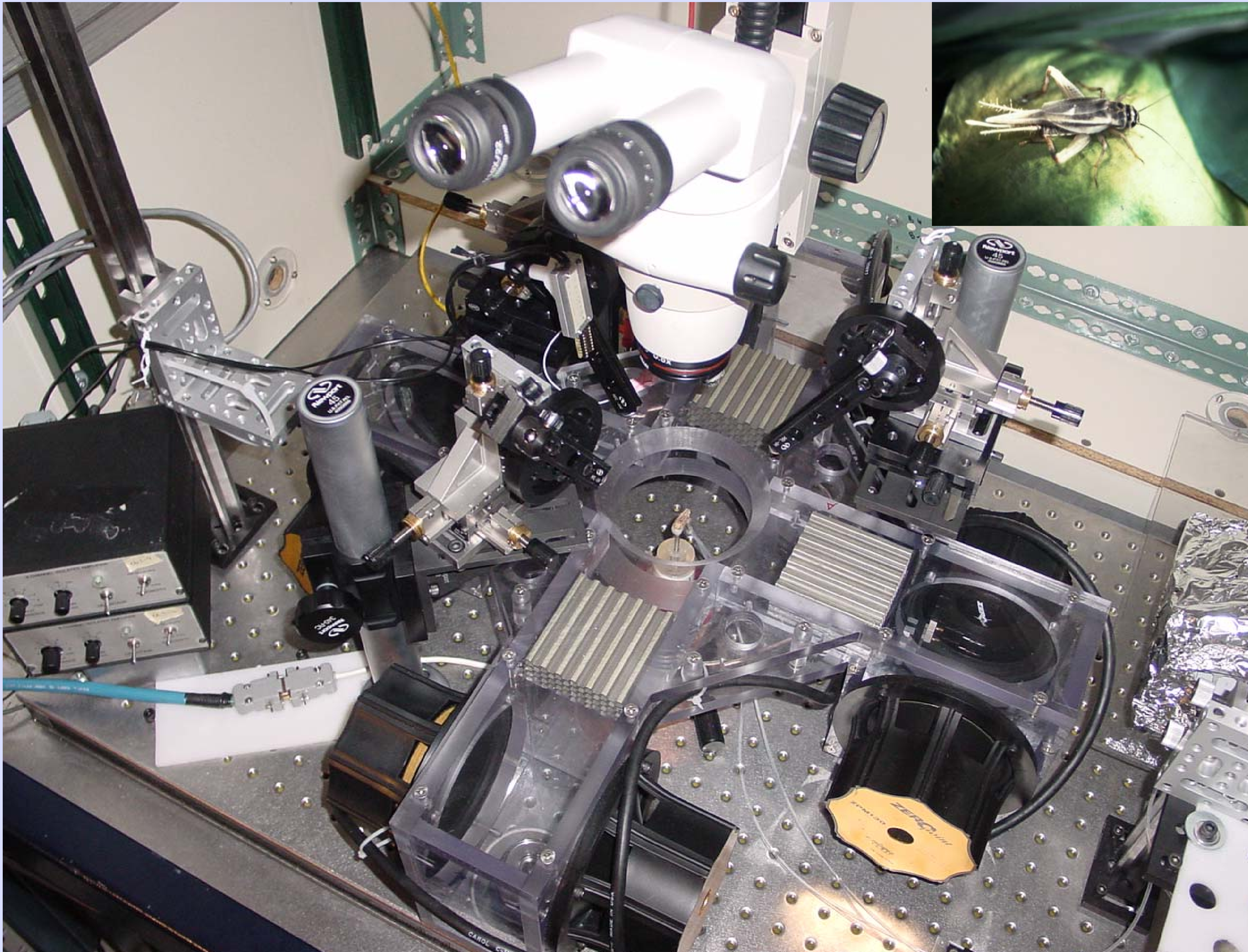
Permuted Similarity Matrix



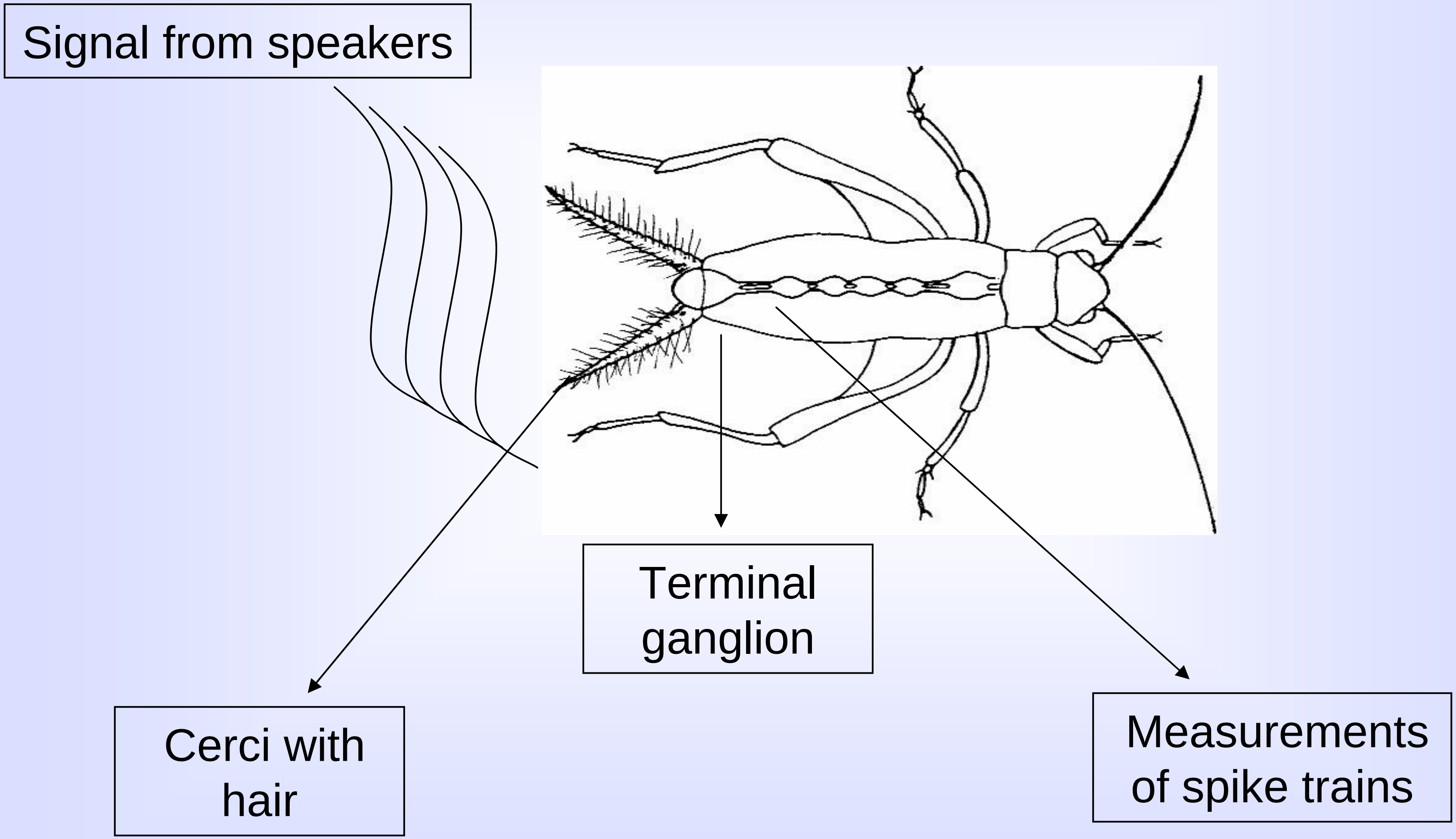
Cluster Assignments




Test case: The cricket cercal system
(a low-frequency, near field extension of the auditory system)



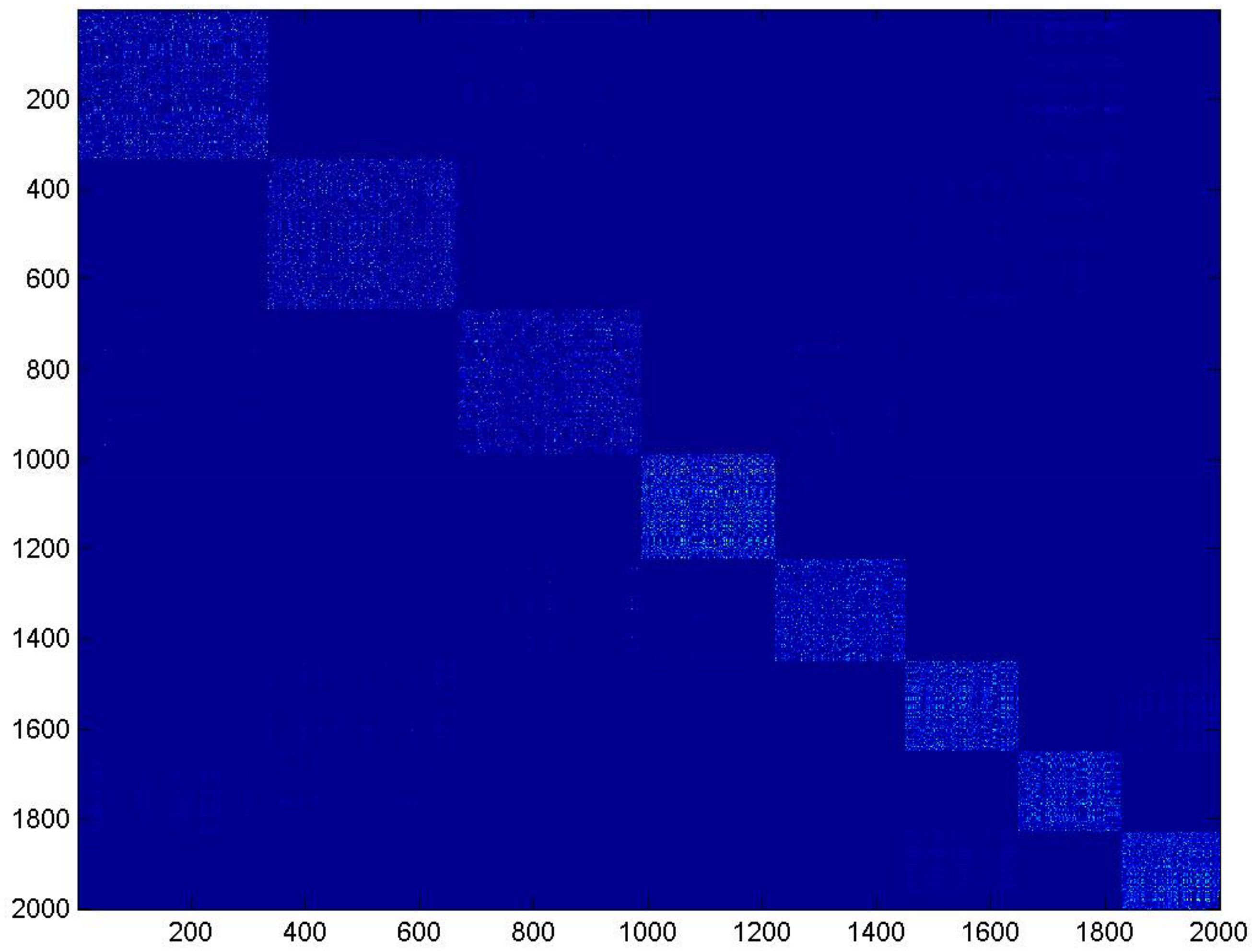
Cricket cerci sensory system

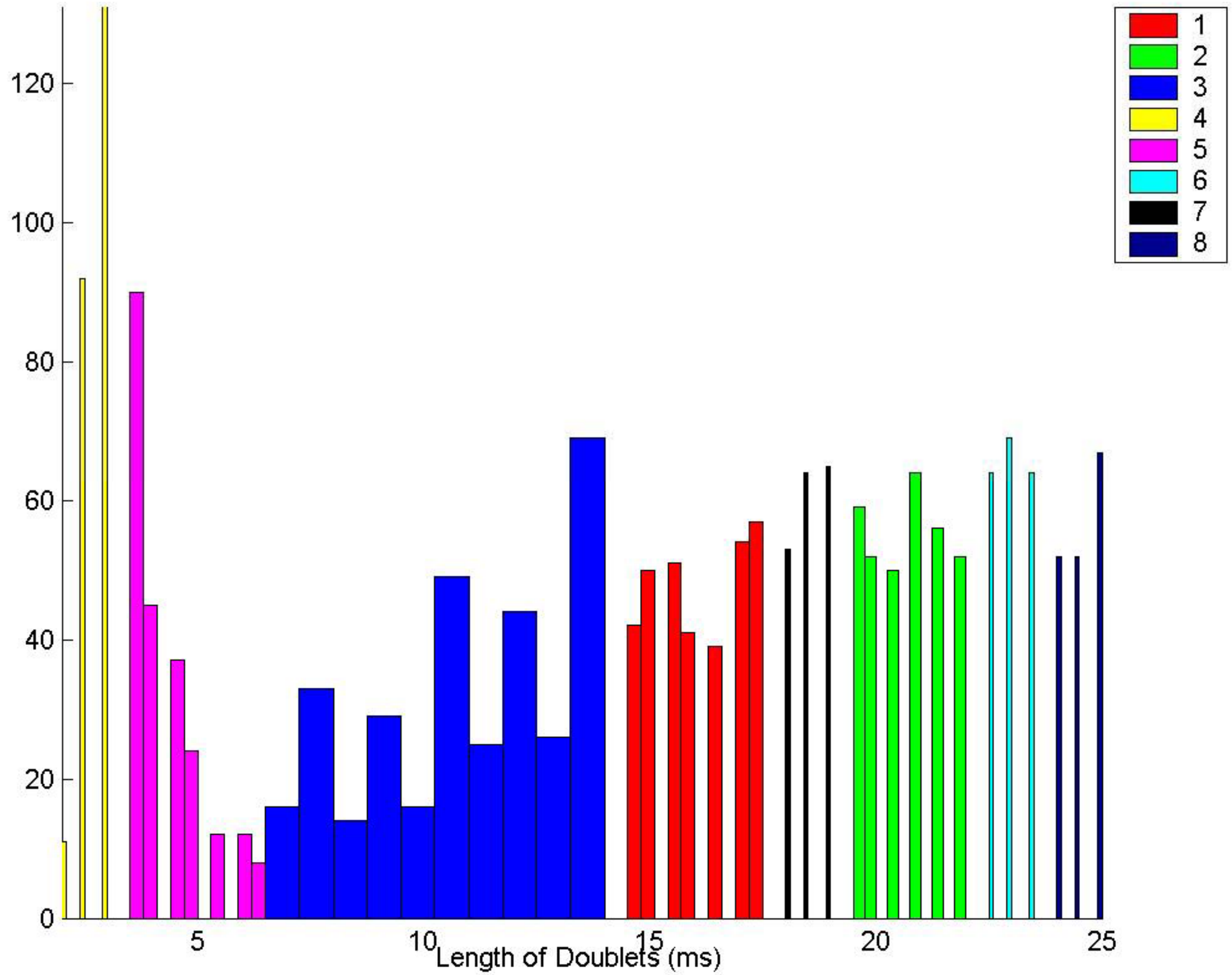


Sample Neural Data Set

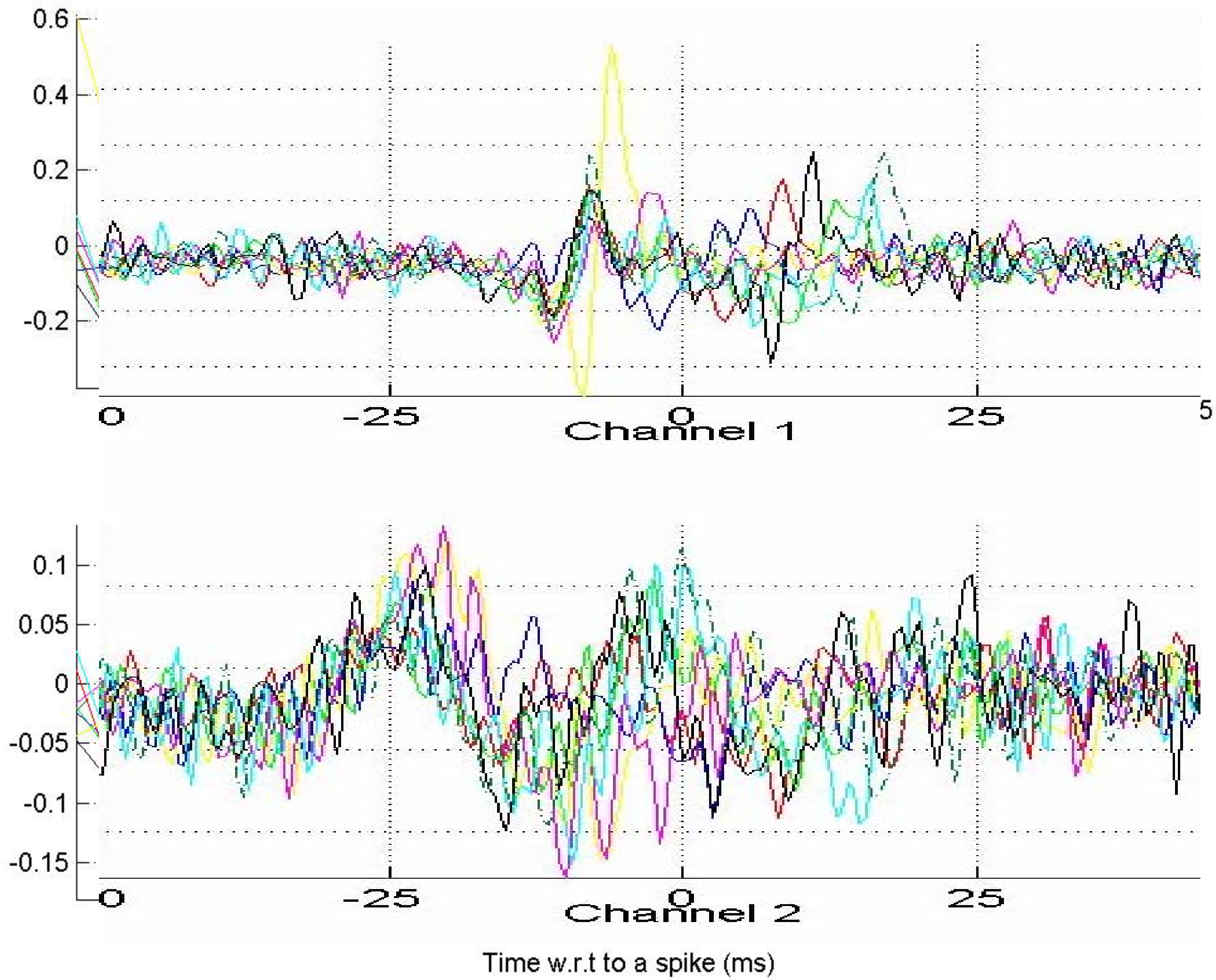
- 1000 input/output pairs
- two channel input
- single channel output
- 8 joint clusters found

Data Point Similarity Matrix





120
100
80
60
40
20
0
5
10
15
20
25
Length of Doublets (ms)
1
2
3
4
5
6
7
8


0.6
0.4
0.2
0
-0.2
0
-25
0
25
5
Channel 1
0.1
0.05
0
-0.05
-0.1
-0.15
0
-25
0
25
Channel 2
Time w.r.t to a spike (ms)

Current Work

- Optimizing the class of projection functions used
- Improving bounds on the number of random trials required
- More experiments…
  - multiple channel data
  - synthetic test data