STUDY ICA component clustering
### Edit STUDY set information - remember to save changes

<table>
<thead>
<tr>
<th>dataset filename</th>
<th>browse</th>
<th>subject</th>
<th>session</th>
<th>condition</th>
<th>group</th>
</tr>
</thead>
<tbody>
<tr>
<td>C:.../data1.txt</td>
<td>...</td>
<td>S01</td>
<td></td>
<td>memorize</td>
<td></td>
</tr>
<tr>
<td>C:.../data2.txt</td>
<td>...</td>
<td>S01</td>
<td></td>
<td>ignore</td>
<td></td>
</tr>
<tr>
<td>C:.../data3.txt</td>
<td>...</td>
<td>S02</td>
<td></td>
<td>probe</td>
<td></td>
</tr>
<tr>
<td>C:.../data4.txt</td>
<td>...</td>
<td>S02</td>
<td></td>
<td>memorize</td>
<td></td>
</tr>
<tr>
<td>C:.../data5.txt</td>
<td>...</td>
<td>S03</td>
<td></td>
<td>ignore</td>
<td></td>
</tr>
<tr>
<td>C:.../data6.txt</td>
<td>...</td>
<td>S03</td>
<td></td>
<td>probe</td>
<td></td>
</tr>
<tr>
<td>C:.../data7.txt</td>
<td>...</td>
<td>S04</td>
<td></td>
<td>memorize</td>
<td></td>
</tr>
<tr>
<td>C:.../data8.txt</td>
<td>...</td>
<td>S04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C:.../data9.txt</td>
<td>...</td>
<td>S04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C:.../data10.txt</td>
<td>...</td>
<td>S04</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Important note: Removed datasets will not be saved before being deleted from EEGLAB memory.

- Dataset info (condition, group, ...) differs from study info. [set] = Overwrite dataset info.
- Delete cluster information (to allow loading new datasets, set new components for clustering, etc.)

[Help] [Cancel] [Ok]
ICs to cluster
Precompute data measures
Precompute data measures

TIP: Compute all measures so you can test different combinations for clustering

Time-frequency options
Cluster components
Precluster schematic

Each component is a dot. Clustering will group these dots.
Classical KMean

1. k initial "means" (in this case k=3, shown in color) are randomly selected from the data set (shown in grey).

2. k clusters are created by associating every observation with the nearest mean.

3. The centroid of each of the k clusters becomes the new means

4. Steps 2 and 3 are repeated until convergence has been reached.
Cluster components
Choosing data measures

What measure(s) should you use?

It depends on your final cluster criteria…
  - If for example, your priority is dipole location,
    then cluster only based on dipole location…

But consider:
  - What is the difference between these two components?
Choosing data measures

Similar dipole location, very different orientation.

Obvious dramatic effect on scalp map topography:
Plot/edit clusters

View and edit current component clusters -- pop_clustedit()

Study name: 'Sternberg' (336 of 336 components clustered)

Select cluster to plot
- All cluster centroids
  - ParentCluster 1 (336 ICs)
    - Cls 2 (17 ICs)
    - Cls 3 (6 ICs)
- Plot scalp maps
- Plot dipoles
- Plot ERPs
- Plot spectra
- Plot ERSPs
- Plot ITCs
- Plot cluster properties

Select component(s) to plot
- 'Cls 2' comp. 1 (S01 IC21)
- 'Cls 2' comp. 2 (S03 IC25)
- 'Cls 2' comp. 3 (S03 IC25)
- 'Cls 2' comp. 4 (S04 IC19)
- Plot scalp map(s)
- Plot dipole(s)
- Plot ERP(s)
- Plot spectra
- Plot ERSP(s)
- Plot ITC(s)

- Params
- Params
- Params

Create new cluster
 Rename selected cluster
 Merge clusters

Reassign selected component(s)
 Remove selected outlier comps.
 Auto-reject outlier components

Help
 Cancel
 Ok
Plot cluster data

Study ‘Attention’: 181 of 181 components clustered

Select cluster to plot

- All cluster centroids
- ParentCluster 1 (181 ICs)
- outlier 2 (1 ICs)
- Cls 3 (5 ICs)

- Plot scalp maps
- Plot dipoles
- Plot ERP
- Plot spectra
- Plot ERPs
- Plot ITCS
- Plot cluster properties

Select component(s) to plot

- outlier 2’ comp. 1 (S12 IC12)
- Cls 3’ comp. 1 (S01 IC1)
- Cls 3’ comp. 2 (S05 IC11)
- Cls 3’ comp. 3 (S06 IC15)

- Plot scalp map(s)
- Plot dipoles(s)
- Plot ERP(s)
- Plot spectra

Create new cluster
Rename selected cluster
Merge clusters

Save STUDY set to disk /home/julie

Plot mean scalp maps for easy reference
Plot cluster data

Choose which cluster

Choose which components
Plot cluster data
Issue with standard clustering

Large parameter space problem: many different clustering solutions can be produced by changing parameters and measure subsets. Which one should we choose?

EEGLAB clustering has ~12 parameters
Measure projection

(EEGLAB plug-in by Nima Bigdely Shamlo) only has one pre-clustering parameter.
Measure projection

- **Instead**, we can directly work on pair-wise similarity matrices and prevent ICs with similarities less than certain threshold (e.g., ERSP corr. < 0.5) to be clustered together.

- The most important measure is **equivalent dipole location**
Measure Projection Toolbox

Projected ERSP

ERP
Exercise

• Load the Stern STUDY (STUDY folder)
• Precluster (pre-computation already done) and cluster components using measures of your choice. Experiment with different measures.