Clustering of ICA components

Arnaud Delorme

(with Julie Onton, Romain Grandchamp, Nima Bigdely Shamlo, Scott Makeig)
Rational of clustering

• ICA transforms the data from a channel basis (activity recorded at each channel) to a component basis (activity computed for each IC)

• Normally, EEG researchers assume that an electrode, say channel F7, represent the same activity in each subject F7 == F7 == F7 ... and then ‘cluster‘ their data channel by channel ...

• But this is only *roughly* correct!
Subject 2
Contribution of brain sources to electrode

Onton et al., *Neuroimage* 2005
ICA clustering

Sometime clusters are spatially separate AND have distinct responses.

In other cases, they have similar responses or they overlap spatially.
Clustering independent components

- Across Subjects, components don’t even have “the same” scalp maps! → Are “the same” components found across subjects?
- What should define “the same” (i.e., “component equivalence”)?
- Similar scalp maps?
- Similar cortical or 3-D equivalent dipole locations?
- Similar activity power spectra?
- Similar ERPs?
- Similar ERSPs?
- Similar ITCs?
- OR ..., Similar combinations of the above?? ...
Outline

- ICA clusters and reliability within subjects
  REF: Grandchamp, Makeig, Delorme, IEEE, 2012

- ICA clustering across subjects
  REF: Delorme et al., PLOS One, 2012

- Clustering in EEGLAB theory & Practice
ICA decomposition of multiple data sets from the same individuals

- **Experimental protocol**
  - Mind wandering experiment
  - 2 subjects
  - 11 x 30 min. sessions
  - 2 sessions per week
  - EEG from Biosemi 64 channels
  - Fs=1024 Hz
Results (Cluster 1)

100 % Sessions contribute
Results (Cluster 2)

100 % Sessions contribute
Results (Cluster 8)

100 % Sessions contribute
Results (Cluster 13)

63.64% Sessions contribute
Results (Cluster 14)

36.36% Sessions contribute

INFO:
Template: CB Session 1 PREPROC STEP 2: Set 1; IC 5.
Number of datasets: 11
Correlation threshold: 0.93 (green line)
Max ICs from each dataset: 1
Cluster: 4 ICs from 4 sets
Sets not contributing:
#5, #6, #7, #8, #9, #10/11;

Similarity = 0.9988

Average Map:

Correlation Plot:

Threshold Plot:

Cls 14 Spectrum:

Power (10^10/log(V^2/Hz))

Frequency (Hz):
## Inter iteration Cluster Consistency

<table>
<thead>
<tr>
<th>Clusters</th>
<th>Iterations</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td><strong>100</strong></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>90</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td><strong>99</strong></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>90</td>
<td>40</td>
<td>10</td>
<td>90</td>
<td>90</td>
<td>60</td>
<td>100</td>
<td>10</td>
<td>60</td>
<td>90</td>
<td><strong>64</strong></td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>60</td>
<td>0</td>
<td>100</td>
<td>60</td>
<td>100</td>
<td>90</td>
<td>60</td>
<td>60</td>
<td>90</td>
<td>60</td>
<td><strong>68</strong></td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>90</td>
<td>100</td>
<td>90</td>
<td>90</td>
<td>60</td>
<td>90</td>
<td>90</td>
<td>100</td>
<td>90</td>
<td>90</td>
<td><strong>89</strong></td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>80</td>
<td>80</td>
<td>60</td>
<td>80</td>
<td>40</td>
<td>80</td>
<td>80</td>
<td>80</td>
<td>80</td>
<td>100</td>
<td><strong>76</strong></td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>60</td>
<td>90</td>
<td>50</td>
<td>60</td>
<td>80</td>
<td>60</td>
<td>0</td>
<td>10</td>
<td>60</td>
<td>50</td>
<td><strong>52</strong></td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>40</td>
<td>90</td>
<td>10</td>
<td>40</td>
<td>0</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>50</td>
<td>60</td>
<td><strong>39</strong></td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>60</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>60</td>
<td>10</td>
<td>90</td>
<td>60</td>
<td>60</td>
<td><strong>37</strong></td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>100</td>
<td>50</td>
<td>100</td>
<td>50</td>
<td>100</td>
<td>50</td>
<td>75</td>
<td><strong>50</strong></td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>50</td>
<td>10</td>
<td>20</td>
<td>50</td>
<td>90</td>
<td>50</td>
<td>50</td>
<td>10</td>
<td>50</td>
<td>20</td>
<td><strong>40</strong></td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>20</td>
<td>10</td>
<td>10</td>
<td>20</td>
<td>20</td>
<td>30</td>
<td>20</td>
<td>20</td>
<td>30</td>
<td>30</td>
<td><strong>21</strong></td>
</tr>
</tbody>
</table>
Installing CORRMAP
### Plutings available for install on the internet

<table>
<thead>
<tr>
<th>Plugin</th>
<th>Version</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>brainmovie</td>
<td>0.1</td>
<td>Brainmovies (command line only)</td>
</tr>
<tr>
<td>cormap</td>
<td>2.00</td>
<td>Cluster ICA components using correlation of scalp maps</td>
</tr>
<tr>
<td>eeg_toolbox</td>
<td>1.0</td>
<td>Interface EEG toolbox functions for ERP peak detection</td>
</tr>
<tr>
<td>ERPLABfilters</td>
<td>1.00</td>
<td>Interface ERPLAB filters (requires separate ERPLAB installation)</td>
</tr>
<tr>
<td>fMRIb</td>
<td>1.21</td>
<td>Remove fMRI artifacts from EEG</td>
</tr>
<tr>
<td>MP_clustering</td>
<td>1.00</td>
<td>Measure projection clustering of ICA components</td>
</tr>
<tr>
<td>Mutual_Info_Clustering</td>
<td>1.00</td>
<td>Mutual information clustering</td>
</tr>
<tr>
<td>VisEd</td>
<td>1.05</td>
<td>Add/Edit dataset events</td>
</tr>
<tr>
<td>iirfilt</td>
<td>1.02</td>
<td>Non linear filtering</td>
</tr>
<tr>
<td>loreta</td>
<td>1.1</td>
<td>Export and import data to/from LORETA software</td>
</tr>
<tr>
<td>std_envtopo</td>
<td>2.13</td>
<td>Plot STUDY cluster envtopo</td>
</tr>
<tr>
<td>std_dipplotWithDensity</td>
<td>0.10</td>
<td><em>BETA</em> Dipole density plot for STUDY clusters</td>
</tr>
<tr>
<td>std_backproj</td>
<td>0.10</td>
<td><em>BETA</em> Select and backproject STUDY cluster ICs</td>
</tr>
<tr>
<td>std_ErpCalc</td>
<td>0.10</td>
<td><em>BETA</em> Flexible calculation and plotting for STUDY cluster ERPs</td>
</tr>
<tr>
<td>PACT</td>
<td>0.10</td>
<td><em>BETA</em> Computes phase-amplitude coupling for continuous data</td>
</tr>
<tr>
<td>SIFT</td>
<td>1.3</td>
<td><em>BETA</em> Analysis and visualization of multivariate connectivity</td>
</tr>
<tr>
<td>clean_rawdata</td>
<td>0.12</td>
<td><em>BETA</em> Cleans continuous data using Artifact Subspace Reconstructions</td>
</tr>
<tr>
<td>firstPassOutlierTrimmer</td>
<td>0.10</td>
<td><em>BETA</em> Rapid browsing and simple interactive cleaning for co...</td>
</tr>
</tbody>
</table>

### Installed plutings

<table>
<thead>
<tr>
<th>Plugin</th>
<th>Version</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cleanline</td>
<td>1.0</td>
<td>Removes sinusoidal artifacts (line noise)</td>
</tr>
</tbody>
</table>
Outline

- ICA clusters and reliability within subjects
  REF: Grandchamp, Makeig, Delorme, IEEE, 2012

- ICA clustering across subjects
  REF: Delorme et al., PLOS One, 2012

- Clustering in EEGLAB theory & Practice
Localization

Electrodes

Components

Time (ms)

ICA component scalp maps

Localization
Computing residual variance (%) 

\[ r = \frac{\sum (x_i - \tilde{x}_i)^2}{\sum x_i^2} \]
Validation of the ICA algorithm for EEG

Data
• 13 subjects performing a memory task
• 71 electrodes including EOGs
• more than 300,000 data points/subject

Decomposition
• 23 ICA algorithms plus PCA and Promax

Analysis
• Localization of all components with a single dipole
  (4-shell spherical model)
<table>
<thead>
<tr>
<th>Algorithm (Matlab func.)</th>
<th>D%</th>
<th>LL</th>
<th>Origin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extended Infomax (runica)</td>
<td>29.9</td>
<td>178</td>
<td>EEGLAB 4.515</td>
</tr>
<tr>
<td>Pearson</td>
<td>29.1</td>
<td>169</td>
<td>ICAcentral (6)</td>
</tr>
<tr>
<td>Infomax (runica)</td>
<td>28.2</td>
<td>160</td>
<td>EEGLAB 4.515</td>
</tr>
<tr>
<td>ERICA</td>
<td>26.9</td>
<td>184</td>
<td>ICALAB 1.5.2</td>
</tr>
<tr>
<td>SONS</td>
<td>25.4</td>
<td>183</td>
<td>ICALAB 1.5.2</td>
</tr>
<tr>
<td>SHIBBS</td>
<td>23.7</td>
<td>169</td>
<td>ICAcentral (5)</td>
</tr>
<tr>
<td>FastICA*</td>
<td>23.5</td>
<td>169</td>
<td>ICAcentral (2)</td>
</tr>
<tr>
<td>JADE (jader)</td>
<td>23.4</td>
<td>169</td>
<td>EEGLAB 4.515</td>
</tr>
<tr>
<td>TICA</td>
<td>23.4</td>
<td>169</td>
<td>ICALAB 1.5.2</td>
</tr>
<tr>
<td>JADE optimized (jade_op)</td>
<td>21.4</td>
<td>169</td>
<td>ICALAB 1.5.2</td>
</tr>
<tr>
<td>JADE w/ time delay (jade_td)</td>
<td>20.2</td>
<td>169</td>
<td>ICALAB 1.5.2</td>
</tr>
<tr>
<td>eeA</td>
<td>19.0</td>
<td>305</td>
<td>ICAcentral (8)</td>
</tr>
<tr>
<td>Infomax (icaML) †</td>
<td>18.8</td>
<td>212</td>
<td>ICA DTU Tbox</td>
</tr>
<tr>
<td>FOBI</td>
<td>18.6</td>
<td>169</td>
<td>ICALAB 1.5.2</td>
</tr>
<tr>
<td>SOBIRO (acsohiro)</td>
<td>17.9</td>
<td>167</td>
<td>EEGLAB 4.515</td>
</tr>
<tr>
<td>EVD 24</td>
<td>17.7</td>
<td>169</td>
<td>ICALAB 1.5.2</td>
</tr>
<tr>
<td>EVD</td>
<td>17.0</td>
<td>169</td>
<td>ICALAB 1.5.2</td>
</tr>
<tr>
<td>SOBI</td>
<td>16.1</td>
<td>583</td>
<td>EEGLAB 4.515</td>
</tr>
<tr>
<td>icaMST†</td>
<td>10.6</td>
<td>169</td>
<td>ICA DTU Tbox</td>
</tr>
<tr>
<td>AMUSE</td>
<td>8.5</td>
<td>169</td>
<td>ICALAB 1.5.2</td>
</tr>
<tr>
<td>PCA</td>
<td>3.1</td>
<td>583</td>
<td>EEGLAB 4.515</td>
</tr>
<tr>
<td>Promax</td>
<td>33.7</td>
<td>467</td>
<td>EEGLAB 4.515</td>
</tr>
<tr>
<td>Whitening/Sphering</td>
<td>57.6</td>
<td>164</td>
<td>EEGLAB 4.515</td>
</tr>
</tbody>
</table>
Component examples
Cumulative number of component below residual variance threshold
More independence -> more biological components
C

Percent near-dipolar (r.v. < 5%) components

Exten, Infomax, Pearson, Infomax

FastICA, SHIBBS, TICA, JADE opt.

r.v. < 5%

$R^2 = 0.71$

$p = 10^{-5}$

Remaining Pairwise Mutual Information (10^{-3} %)

$R^2$ value

Probability ($10^x$)

Near dipolar cutoff (% r.v.)
Correlations between decompositions

Scalp map corr.
- Alpha
- fMθ
- Blink
- left µ
- right µ
- lat. EOG
- EMG

Raw data corr.
- Alpha
- fMθ
- Blink
- left µ
- right µ
- lat. EOG
- EMG

Information reduction difference (bits)

Methods:
- Amica
- Infomax
- Ext. Infomax
- Pearson
- SHIBBS
- JADE
- FastICA
- TICA
- JADE opt
- SOBI
- JADE–TD
- SOBRO
- Sphering
- FOBI
- EVD24
- EVD
- icaMS
- AMUSE
- PCA
- SONS
- eeA
- ERICA
Left $\mu$ cluster

S2 IC47

S4 IC37

S6 IC36

S9 IC7

S12 IC45

S3 IC47

S5 IC48

S7 IC35

S11 IC45

S14 IC45
Right $\mu$ cluster

S1 IC51
S3 IC41
S5 IC51
S7 IC48
S11 IC49
S2 IC41
S4 IC50
S6 IC60
S9 IC39
S14 IC49

Power (dB)
Frequency (Hz)
Occipital $\alpha$ cluster
Frontal Midline $\theta$ cluster

S1 IC63

S9 IC16

S12 IC15

S2 IC18

S11 IC16

S13 IC15

S14 IC16
$R^2 = 0.74$
$p < 10^{-5}$
Outline

- ICA clusters and reliability within subjects
  REF: Grandchamp, Makeig, Delorme, IEEE, 2012

- ICA clustering across subjects
  REF: Delorme et al., PLOS One, 2012

- Clustering in EEGLAB theory & Practice
## Edit dataset info

### Create a new STUDY set -- pop_study()

**Edit STUDY set information - remember to save changes**

<table>
<thead>
<tr>
<th>STUDY set name:</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Sternberg</th>
</tr>
</thead>
<tbody>
<tr>
<td>STUDY set task name:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sternberg</td>
</tr>
<tr>
<td>STUDY set notes:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Dataset Information

<table>
<thead>
<tr>
<th>Dataset filename</th>
<th>Browse</th>
<th>Subject</th>
<th>Session</th>
<th>Condition</th>
<th>Group</th>
<th>Select by r.v.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C:Users\julie\Documents\Work</td>
<td>...</td>
<td>SC1</td>
<td></td>
<td>memorize</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C:Users\julie\Documents\Work</td>
<td>...</td>
<td>SC1</td>
<td></td>
<td>ignore</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C:Users\julie\Documents\Work</td>
<td>...</td>
<td>SC1</td>
<td></td>
<td>probe</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C:Users\julie\Documents\Work</td>
<td>...</td>
<td>SC2</td>
<td></td>
<td>memorize</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C:Users\julie\Documents\Work</td>
<td>...</td>
<td>SC2</td>
<td></td>
<td>ignore</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C:Users\julie\Documents\Work</td>
<td>...</td>
<td>SC2</td>
<td></td>
<td>probe</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C:Users\julie\Documents\Work</td>
<td>...</td>
<td>SC3</td>
<td></td>
<td>memorize</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C:Users\julie\Documents\Work</td>
<td>...</td>
<td>SC3</td>
<td></td>
<td>ignore</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C:Users\julie\Documents\Work</td>
<td>...</td>
<td>SC3</td>
<td></td>
<td>probe</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C:Users\julie\Documents\Work</td>
<td>...</td>
<td>SC4</td>
<td></td>
<td>memorize</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C:Users\julie\Documents\Work</td>
<td>...</td>
<td>SC4</td>
<td></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.)
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

[Image of EEGlab software interface with options for studying and clustering components]
Precluster schematic

ERSP (time/freq)
Mean ERSPs
ICs (all subj)

Power spectrum
Mean spectra
ICs (all subj)

3D dipole position

PCA

ICs (all subj)
PC templates
# PCs

PC weights
concatenate

ICs (all subj)
ERSP
Spectrum
Dipoles

PCA

ICs (all subj)
PC templates
# PCs

PC weights
concatenate

ICs (all subj)
PC templates
x3

PC weights

Precluster: Use singular values from PCA

Mean ERSPs

ICs (all subj)

ERSP (time/freq)

PCA

ERSP (time/freq)

PC templates

# PCs

Normalized singular values

~ relative variance of principal components

10% of max singular value

0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
0 10 20 30 40 50 PCs

Normalized singular values
Precluster schematic

Each component is a dot. Clustering will group these dots.
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:

But, do they perform the same functions?
Choosing data measures

ERPs seem different…
Choosing data measures

Spectra are similar, but they have variable responses to different conditions…
Choosing data measures

What data measures should you use?

It depends…

• broadly-matched ICs: use many/all of the measures.

• specifically-matched ICs: use one/few of the measures.
Plot/edit clusters
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 ERPs
- Plot spectra
- Plot ERPs
- Plot ITCs
- Plot cluster properties

Create new cluster
Rename selected cluster
Merge clusters

Save STUDY set to disk /home/julie

Plot mean scalp maps for easy reference

Average scalp map for all clusters
- outlier 2 (12 ICs, 7 Sc)
- Chs 3 (4 ICs, 4 Sc)
- Chs 4 (8 ICs, 8 Sc)
- Chs 5 (7 ICs, 7 Sc)
- Chs 6 (3 ICs, 3 Sc)
- Chs 7 (9 ICs, 8 Sc)
- Chs 8 (5 ICs, 5 Sc)
- Chs 9 (2 ICs, 0 Sc)
- Chs 11 (4 ICs, 6 Sc)
- Chs 12 (3 ICs, 2 Sc)
- Chs 13 (3 ICs, 3 Sc)
- Chs 14 (4 ICs, 4 Sc)
- Chs 15 (4 ICs, 3 Sc)
- Chs 16 (7 ICs, 7 Sc)
- Chs 17 (7 ICs, 7 Sc)
- Chs 18 (5 ICs, 5 Sc)
- Chs 19 (9 ICs, 8 Sc)
- Chs 20 (17 ICs, 15 Sc)
- Chs 21 (8 ICs, 8 Sc)
- Chs 22 (11 ICs, 8 Sc)
- Chs 23 (8 ICs, 7 Sc)
- Chs 24 (4 ICs, 4 Sc)
- Chs 25 (3 ICs, 3 Sc)
- Chs 26 (1 ICs, 1 Sc)
- Chs 27 (5 ICs, 3 Sc)
Plot cluster data

Choose which cluster

Choose which components
Plot cluster data
Plot cluster ERP
Reassigning components
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**

- Assuming a certain variability estimate for dipole location (due to error in localization and subject variability), one can also estimate an optimum number of clusters.
Measure Projection Toolbox