



Clustering of ICA components

EEGLAB

Arnaud Delorme
Julie Onton
Romain Grandchamp
Nima Bigdely Shamlo
Scott Makeig

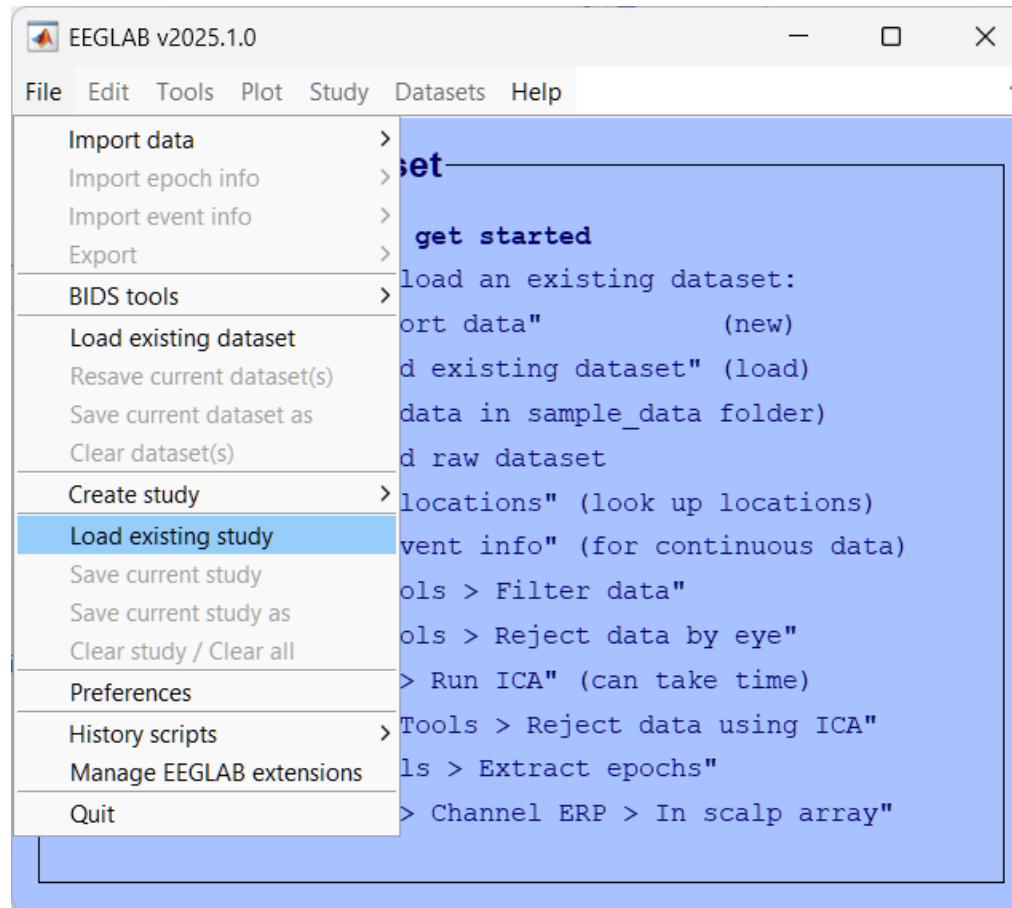
Steps of clustering

1. Select ICA components for clustering
2. Precompute measures of interest
3. Cluster measures
4. Plot clusters and edit them if necessary



Load an existing STUDY

ds002718_processed/Face_detection.study



EEGLAB v2025.1.0

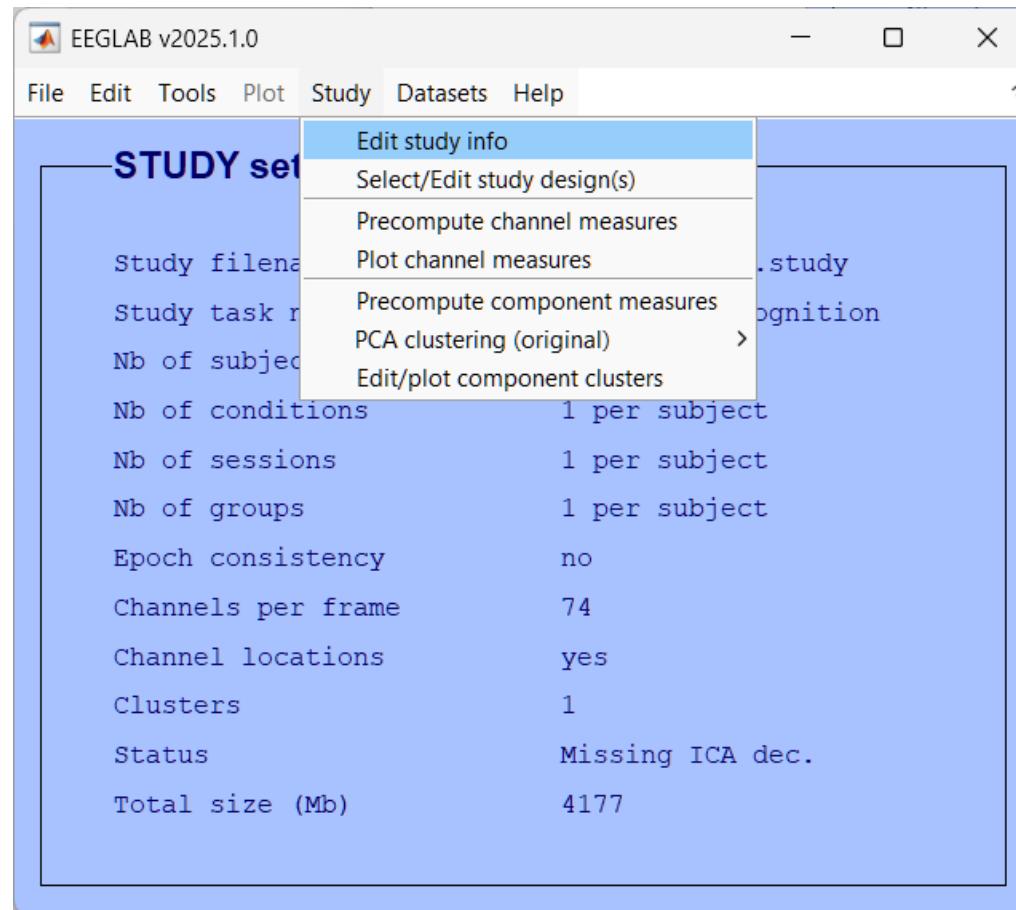
File Edit Tools Plot Study Datasets Help

STUDY set:

Study filename:	...cessed\Face_detection.study
Study task name	
Nb of subjects	18
Nb of conditions	1 per subject
Nb of sessions	1 per subject
Nb of groups	1 per subject
Epoch consistency	yes
Channels per frame	60, 61, 62, 64, 65, 66, 67, 68
Channel locations	yes
Clusters	19
Status	Pre-clustered
Total size (Mb)	2571.8



Study/Edit study info



Edit dataset info

Create a new STUDY set -- pop_study0

Edit STUDY set information - remember to save changes

STUDY set name:

STUDY set task name:

STUDY set notes:

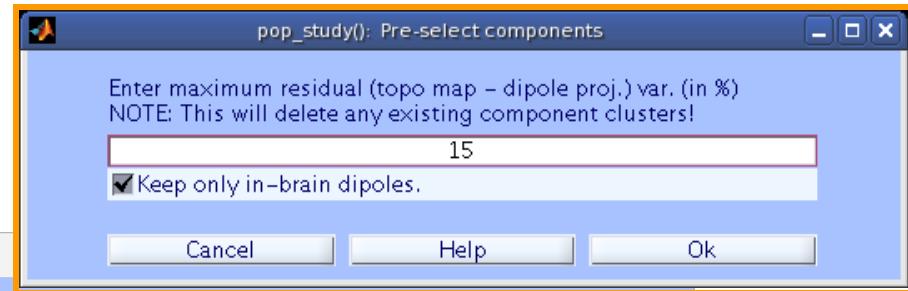
	dataset filename	browse	subject	session	run	condition	group	Select by r.v.	
1	C:\Users\Romain\Workspace\	...	sub-002	1	1			All comp.	Clear
2	C:\Users\Romain\Workspace\	...	sub-003	1	1			All comp.	Clear
3	C:\Users\Romain\Workspace\	...	sub-004	1	1			All comp.	Clear
4	C:\Users\Romain\Workspace\	...	sub-005	1	1			All comp.	Clear
5	C:\Users\Romain\Workspace\	...	sub-006	1	1			All comp.	Clear
6	C:\Users\Romain\Workspace\	...	sub-007	1	1			All comp.	Clear
7	C:\Users\Romain\Workspace\	...	sub-008	1	1			All comp.	Clear
8	C:\Users\Romain\Workspace\	...	sub-009	1	1			All comp.	Clear
9	C:\Users\Romain\Workspace\	...	sub-010	1	1			All comp.	Clear
10	C:\Users\Romain\Workspace\	...	sub-011	1	1			All comp.	Clear

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

< Page 1 >

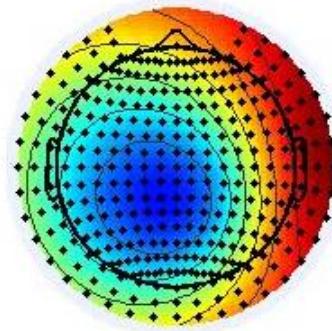
Delete cluster information (to allow loading new datasets, set new components for clustering, etc.)

Help Cancel Ok

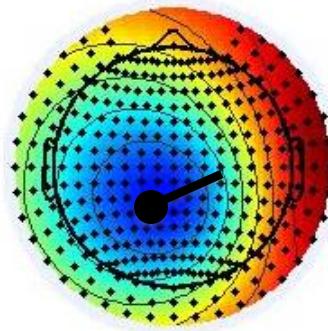


Computing residual variance (%)

Actual IC map (\mathcal{X}_i)



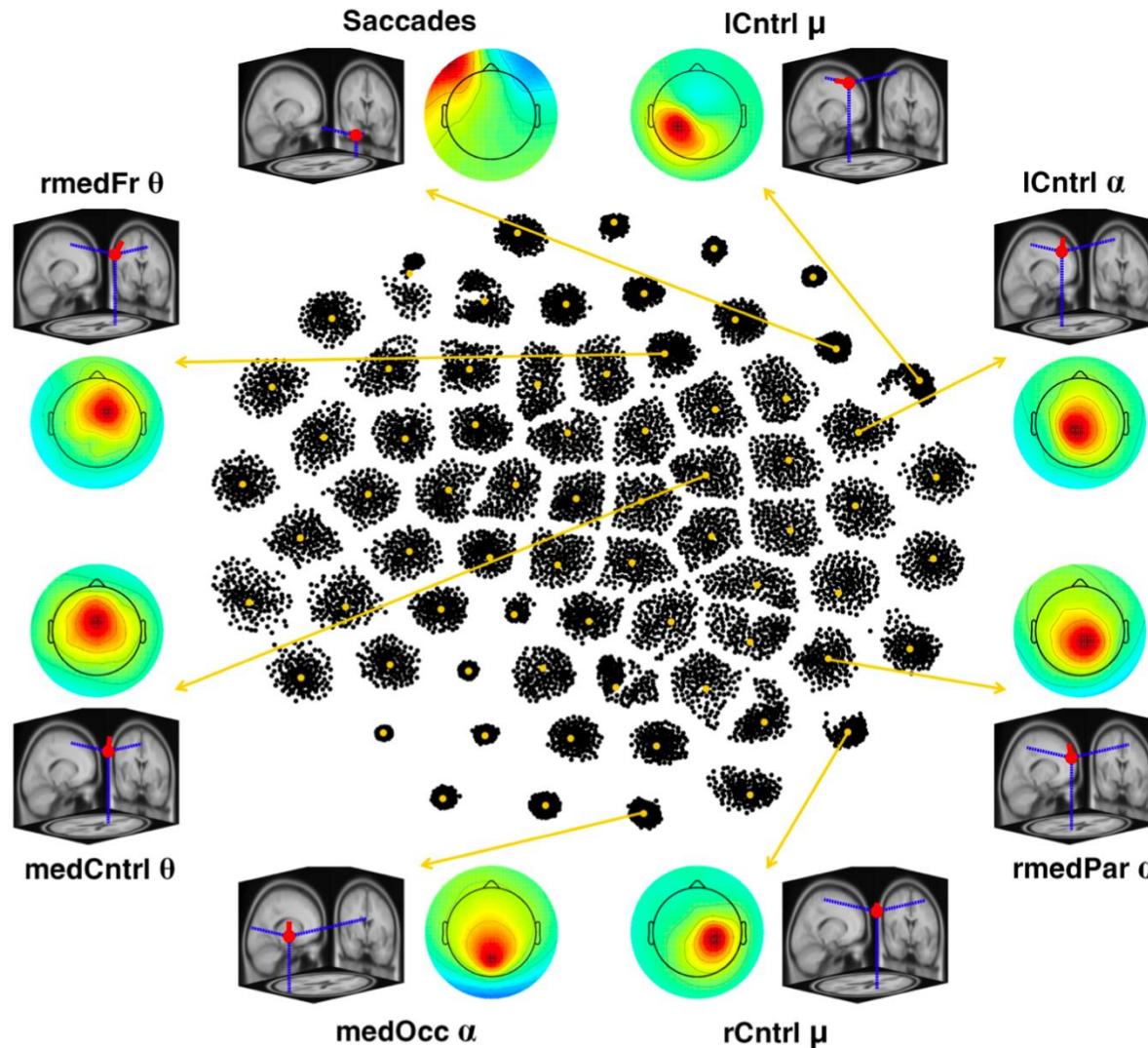
Dipole projection ($\tilde{\mathcal{X}}_i$)



$$rv = \frac{\sum_i (x_i - \tilde{x}_i)^2}{\sum_i x_i^2}$$



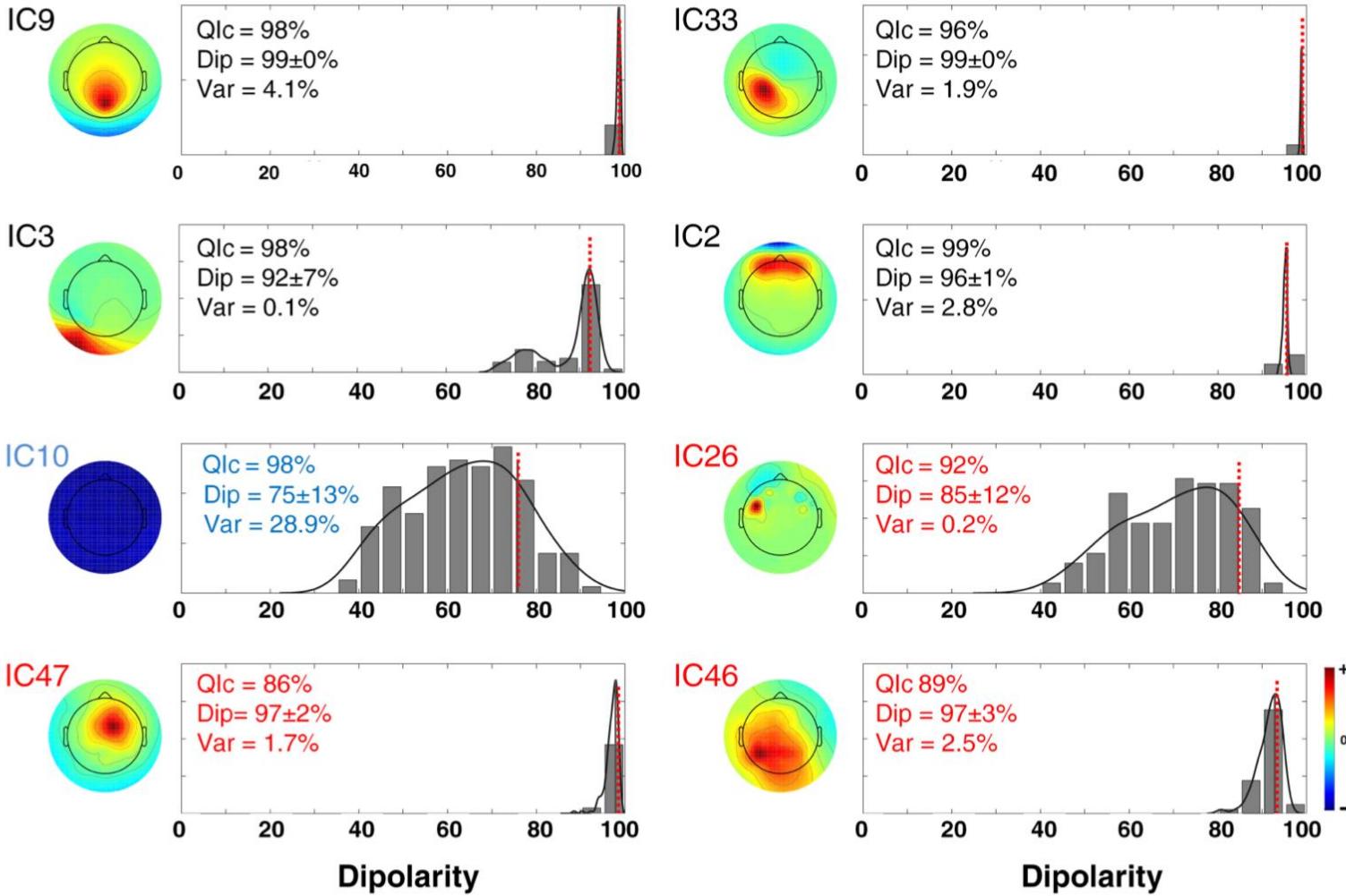
Clustering results example



RELICA: A method for estimating the reliability of independent components

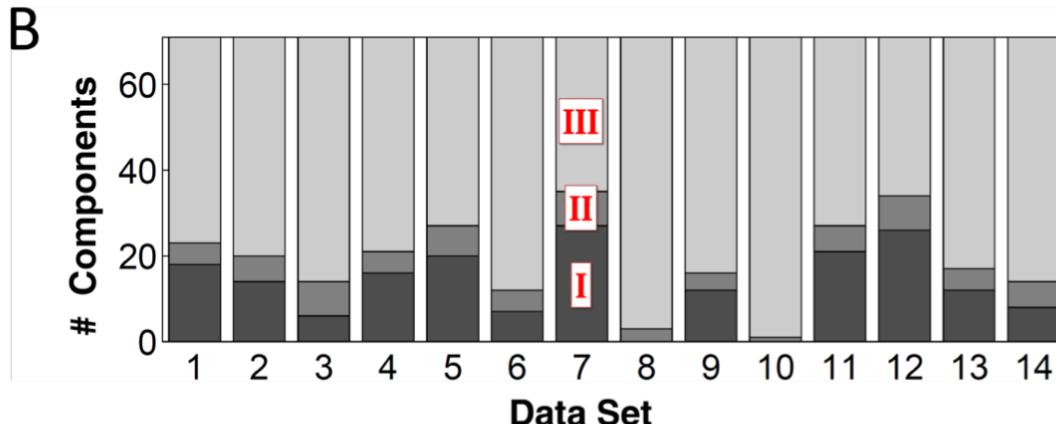
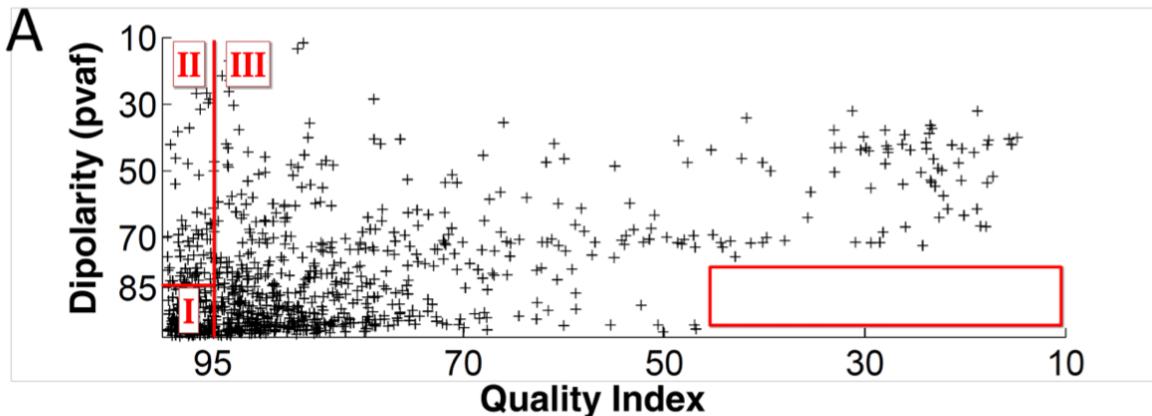
Within-cluster reliability

The distribution of dipolarity within the cluster helps assessing the **quality** and characteristics of Independent Components



Reliability criteria and the rv<15%

First justification why we should select an $r.v < 15\%$ for components to include in further analyses: there is a forbidden region underlined in red, that indicates the absence of



CLASS I

Quality Index and Dipolarity above Retention threshold: **Good**

CLASS II

Quality Index above threshold, dipolarity below: **artifact** or mixing of multiple processes

CLASS III

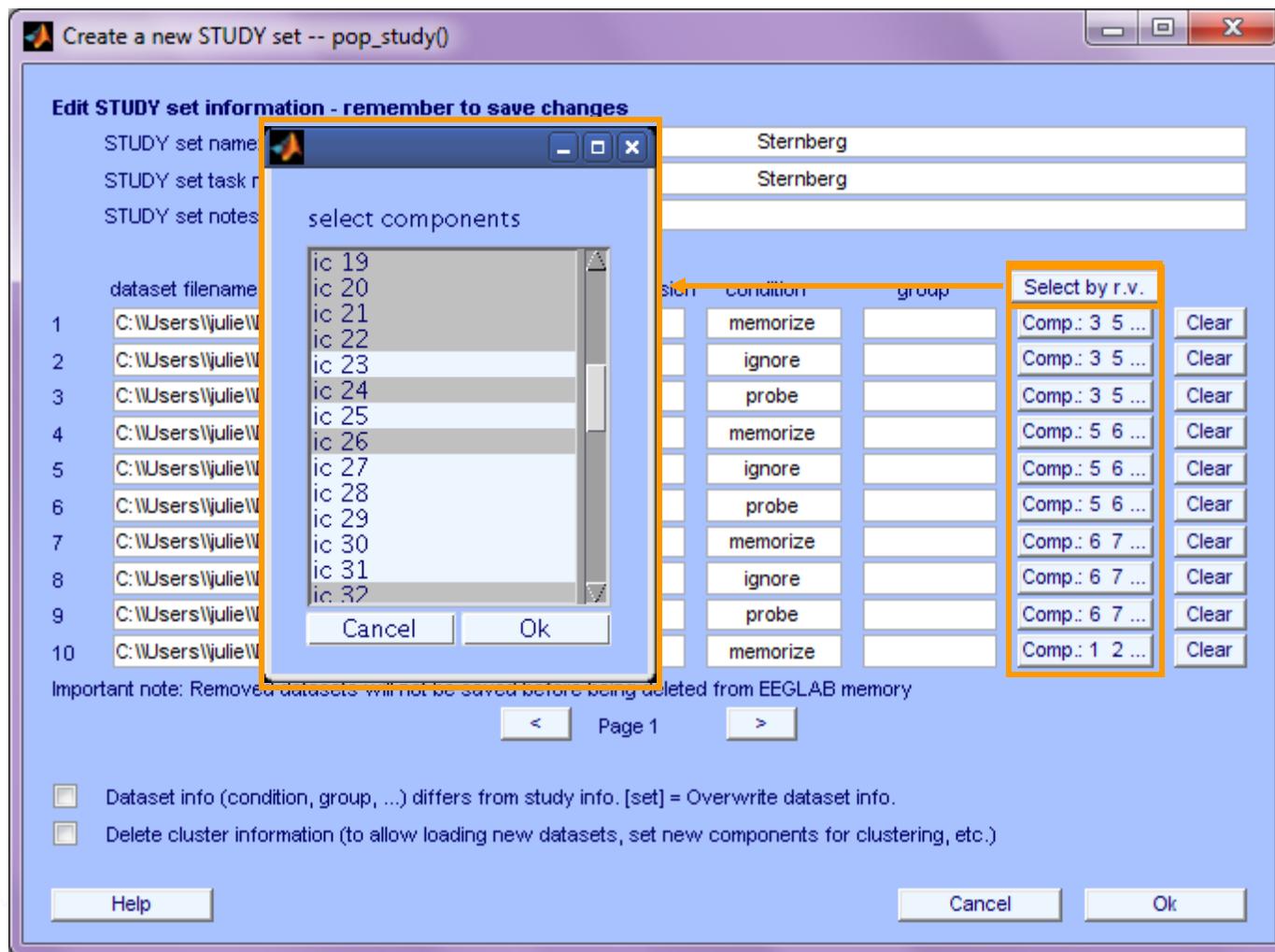
Quality Index below retention threshold

$dip \pm std > th$

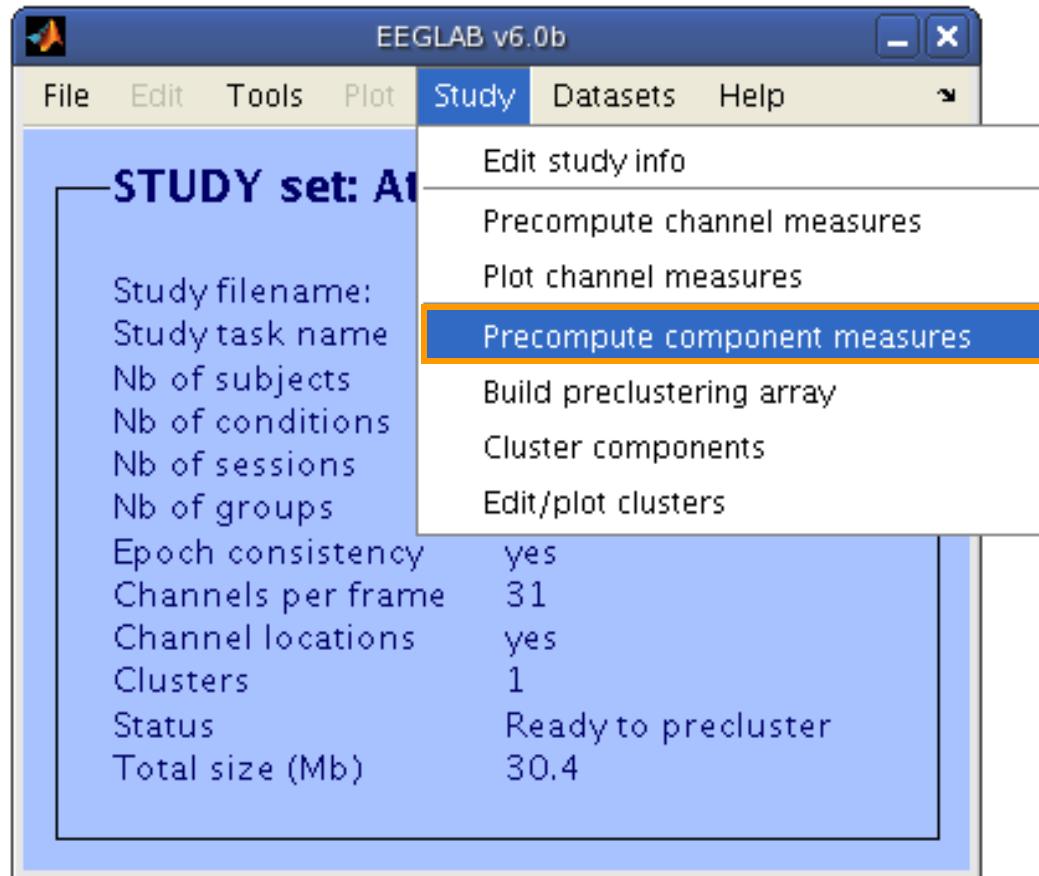
$dip \pm std < th$

Probable Inseparable noise: variance explained useful or **multiple subject classification**

ICs to cluster



Precompute data measures



Pre-compute measures

EEGLAB v9.0.0.0b

File Edit Tools Plot Study Datasets Help

STUDY set:

- Study filename: ...s/data
- Study task name
- Nb of subjects
- Nb of conditions
- Nb of sessions
- Nb of groups
- Epoch consistency
- Channels per frame
- Channel locations
- Clusters
- Status
- Total size (Mb)

61
yes
1
Pre-clustered
8.2

Precompute component measures

- Measure Product clustering
- PCA clustering (original)
- Edit/plot clusters

61
yes
1
Pre-clustered
8.2

Select and compute component measures for later clustering -- pop_precomp()

Pre-compute component measures for STUDY 'Sternberg' - 'STUDY.design 1'

Compute ERP/spectrum/ERSP only for components selected by RV (set) or for all components (unset)

List of measures to precompute

- ERPs Baseline ([min max] in ms)
- Power spectrum Spectopo parameters 'specmode', 'fft'
- ERSPs Time/freq. parameters 'cycles', [3 0.5], 'nfreqs', 100
- ITCs
- Scalp maps

Save single-trial measures for single-trial statistics - requires disk space

Recompute even if present on disk

Help Cancel Ok

EEGLAB v9.0.0.0b

File Edit Tools Plot Study Datasets Help

STUDY set:

- Study filename: ...s/data
- Study task name
- Nb of subjects
- Nb of conditions
- Nb of sessions
- Nb of groups
- Epoch consistency
- Channels per frame
- Channel locations
- Clusters
- Status
- Total size (Mb)

61
yes
1
Pre-clustered
8.2

Precompute channel measures

- Plot channel measures
- Precompute component measures
- Measure Product clustering
- PCA clustering (original)
- Edit/plot clusters

Select and compute component measures for later clustering -- pop_precomp()

Pre-compute channel measures for STUDY 'Sternberg' - 'STUDY.design 1'

Channel list (default:all)

- Spherical interpolation of missing channels (performed after optional ICA removal below)
- Remove ICA artifactual components pre-tagged in each dataset
- Remove artifactual ICA cluster or clusters (hold shift key)

ParentCluster 1
Cls 2
Cls 3
Cls 4

List of measures to precompute

- ERPs Baseline ([min max] in ms)
- Power spectrum Spectopo parameters 'specmode', 'fft'
- ERSPs Time/freq. parameters 'cycles', [3 0.5], 'nfreqs', 100
- ITCs

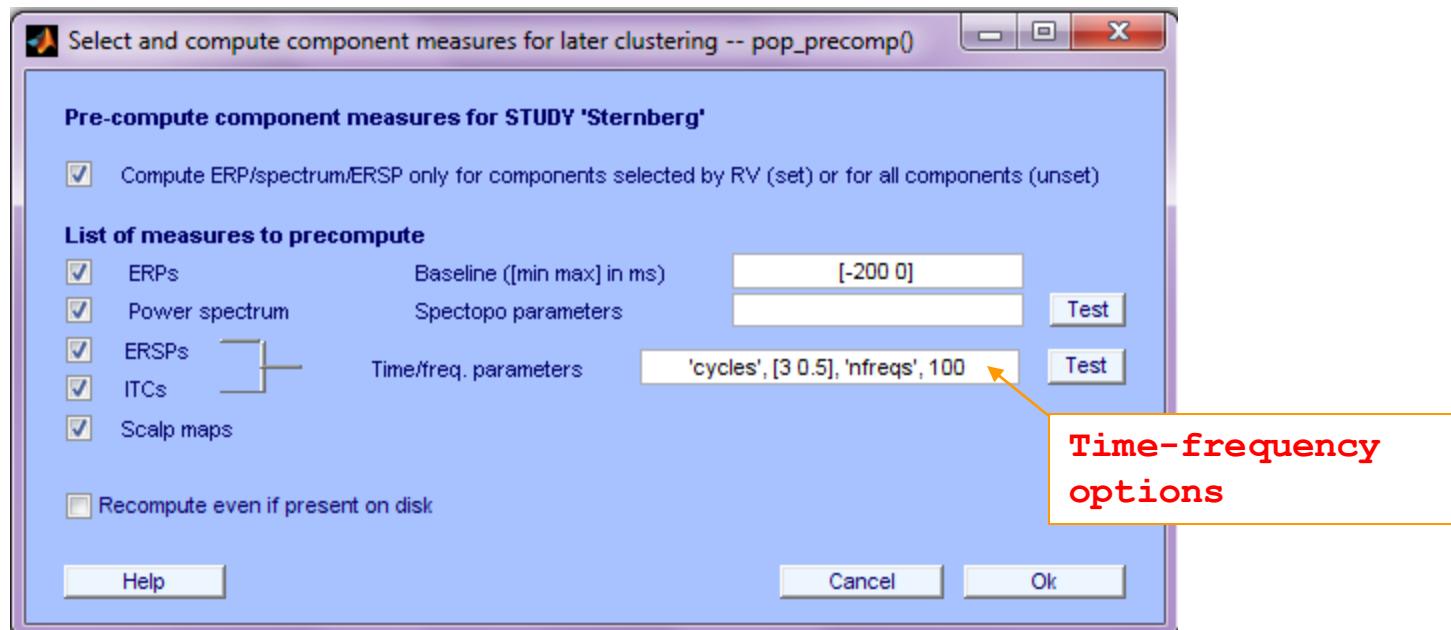
Save single-trial measures for single-trial statistics - requires disk space

Recompute even if present on disk

Help Cancel Ok

Precompute data measures

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



Cluster components

EEGLAB v15.x (dev)

File Edit Tools Plot Study Datasets Help

STUDY set: Sternber

Study filename:
Study task name
Nb of subjects
Nb of conditions
Nb of sessions
Nb of groups
Epoch consistency
Channels per frame
Channel locations
Clusters
Status
Total size (Mb)

Study info
Select/Edit study design(s)
Precompute channel measures
Plot channel measures
Precompute component measures
PCA clustering (original) > Build preclustering array
Edit/plot clusters
Cluster components

Select and compute component measures for later clustering -- pop_preclust()

Build pre-clustering matrix for STUDY set: Sternberg

Only measures that have been precomputed may be used for clustering

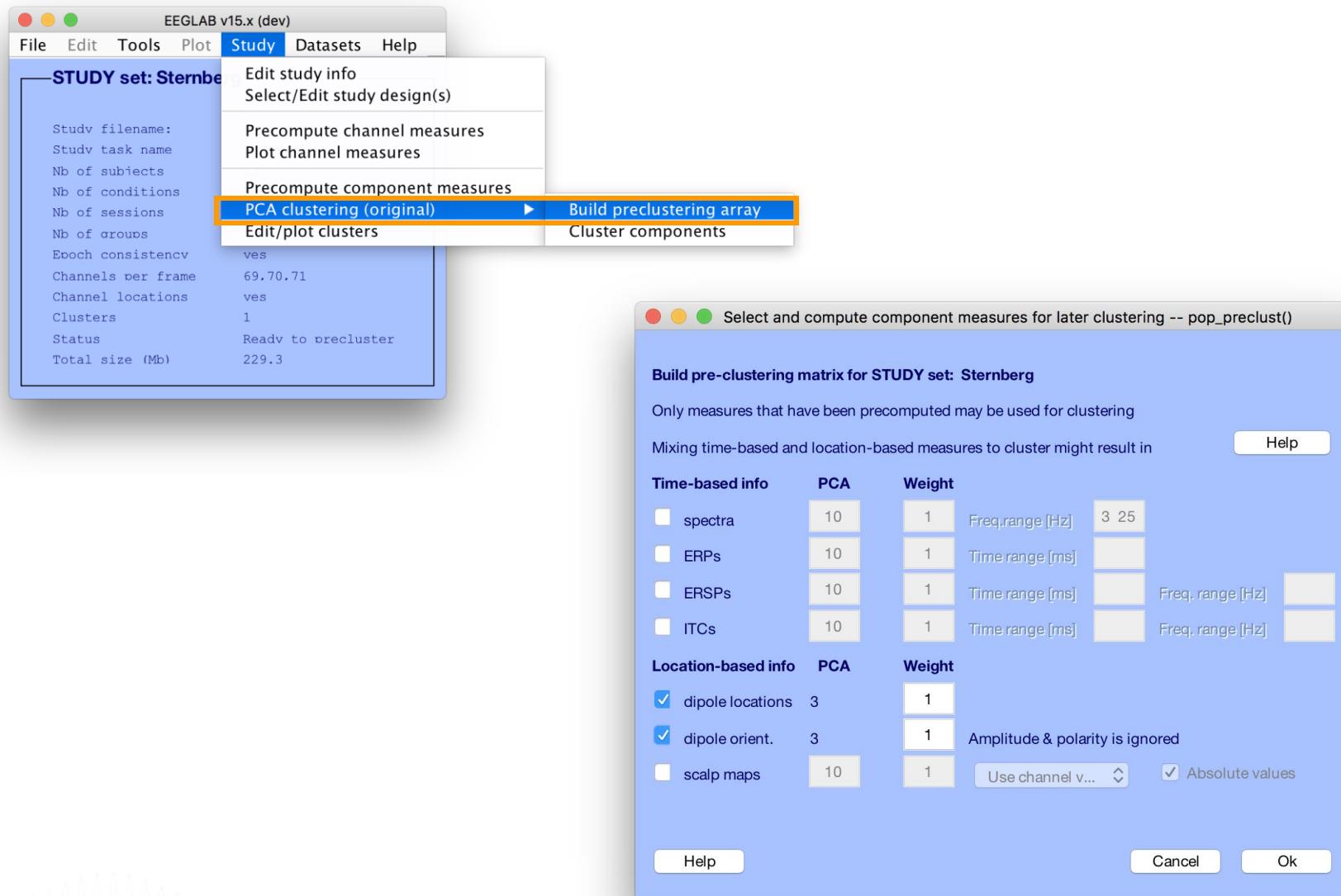
Mixing time-based and location-based measures to cluster might result in

Help

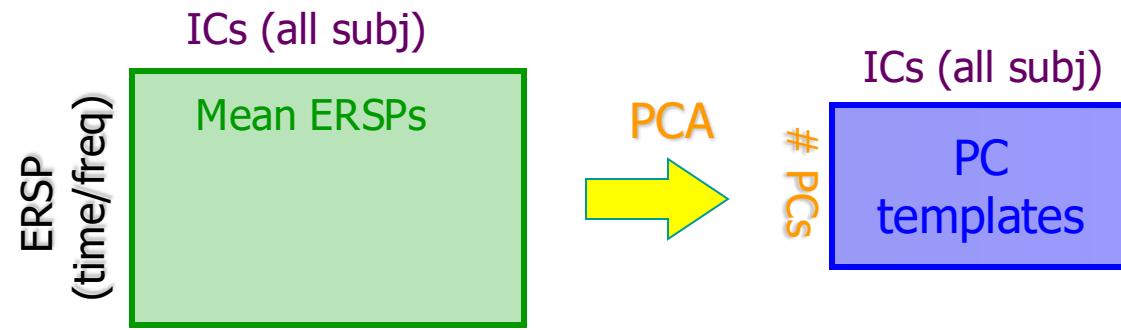
Time-based info	PCA	Weight	
<input type="checkbox"/> spectra	10	1	Freq.range [Hz] 3.25
<input type="checkbox"/> ERPs	10	1	Time range [ms]
<input type="checkbox"/> ERSPs	10	1	Time range [ms]
<input type="checkbox"/> ITCs	10	1	Freq. range [Hz]
			Freq. range [Hz]

Location-based info	PCA	Weight	
<input checked="" type="checkbox"/> dipole locations	3	1	
<input checked="" type="checkbox"/> dipole orient.	3	1	Amplitude & polarity is ignored
<input type="checkbox"/> scalp maps	10	1	<input type="checkbox"/> Use channel v... <input type="checkbox"/> Absolute values

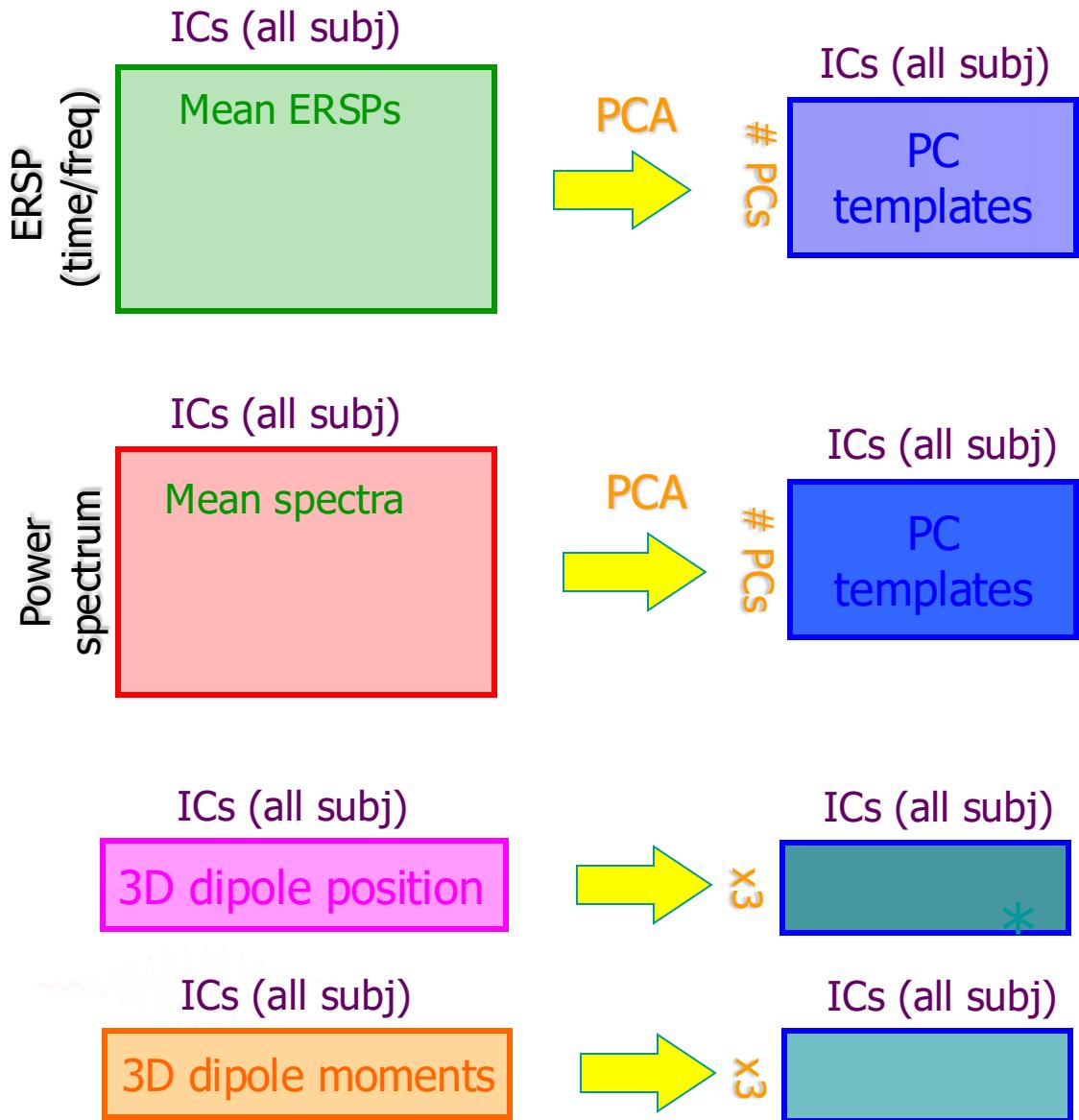
Help Cancel Ok



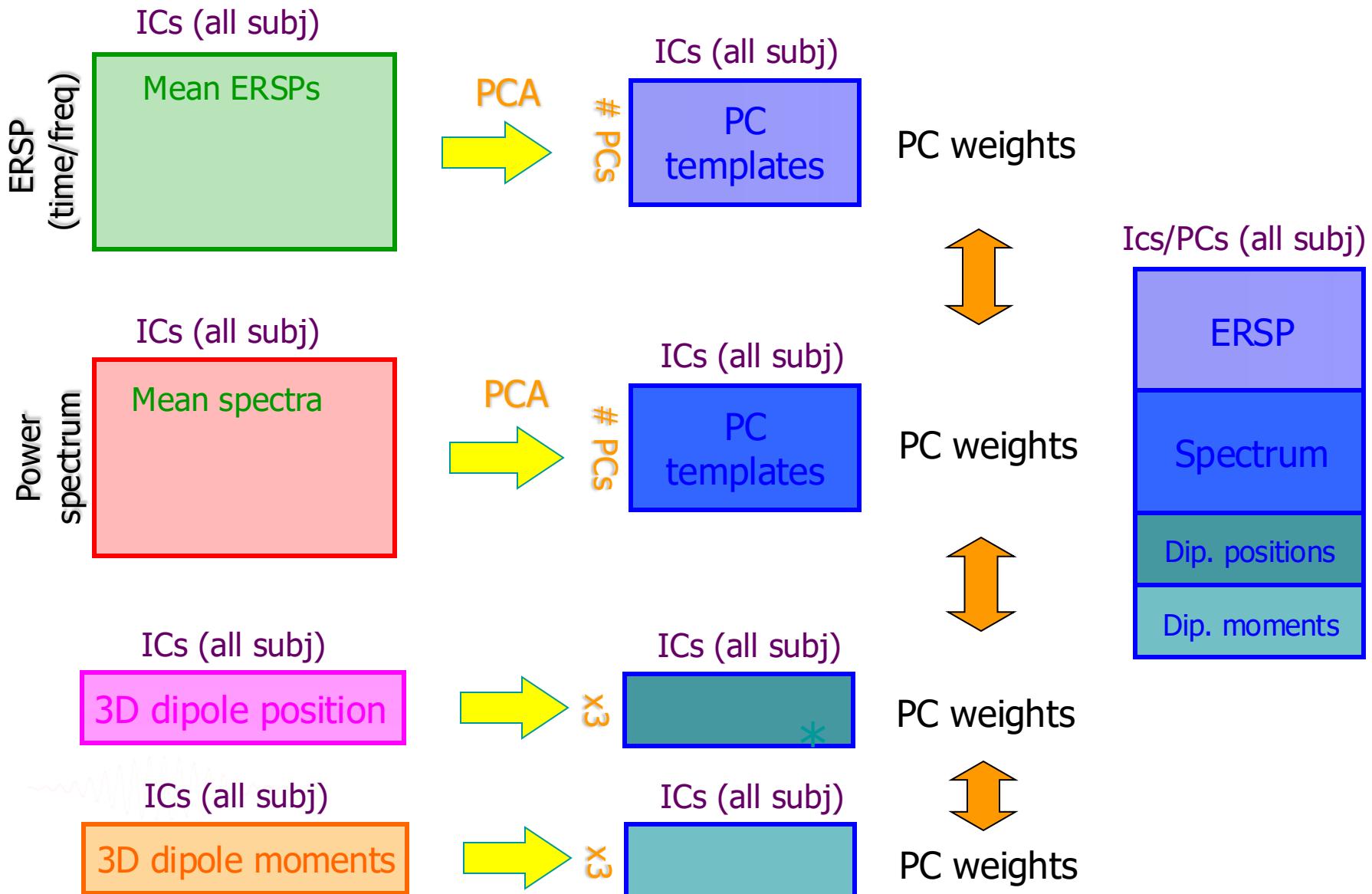
Precluster schematic



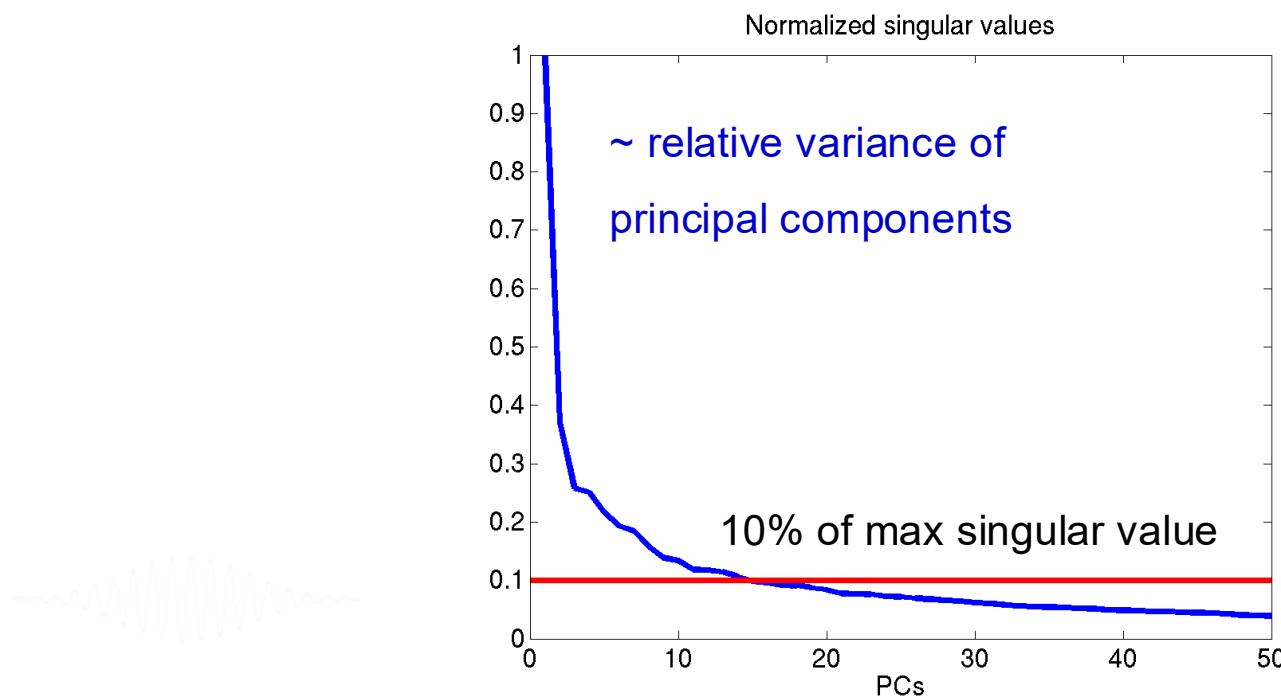
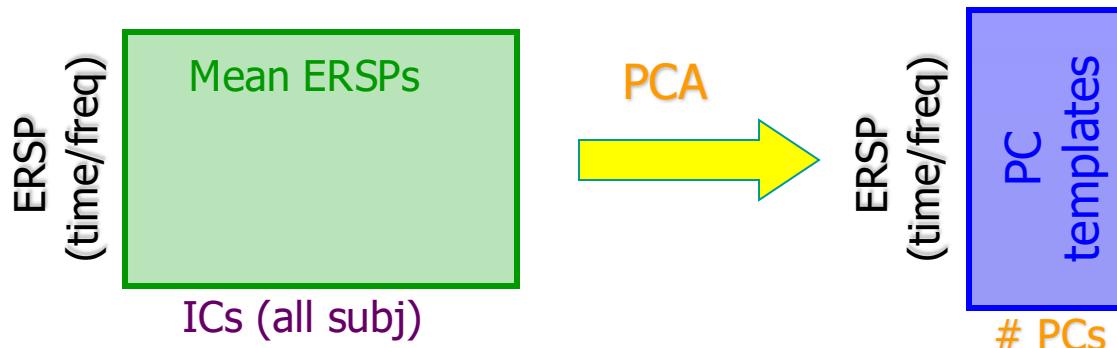
Precluster schematic



Precluster schematic

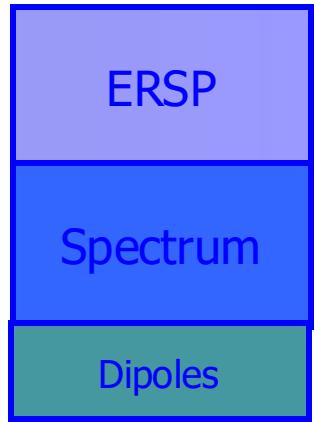


Precluster: Use singular values from PCA

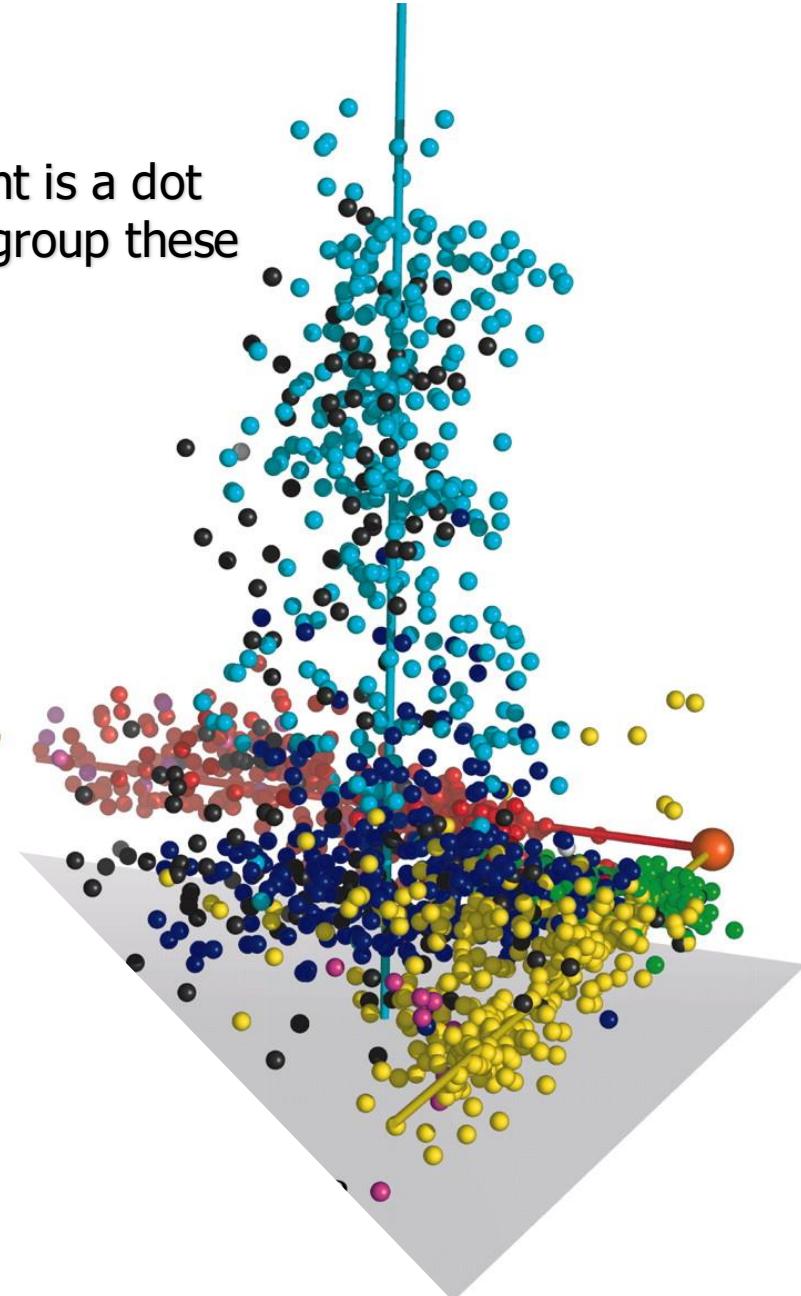


Precluster schematic

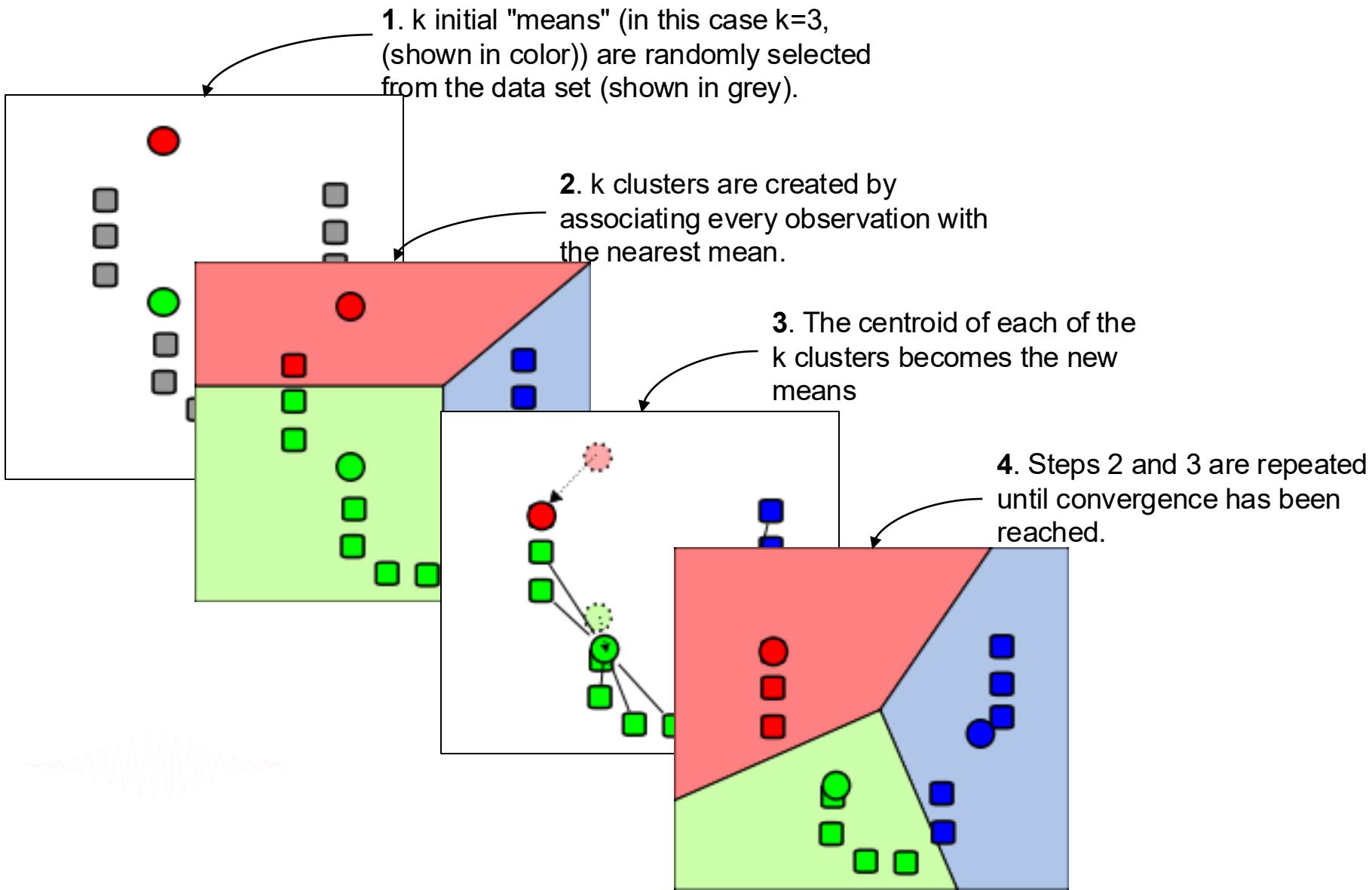
ICs (all subj)



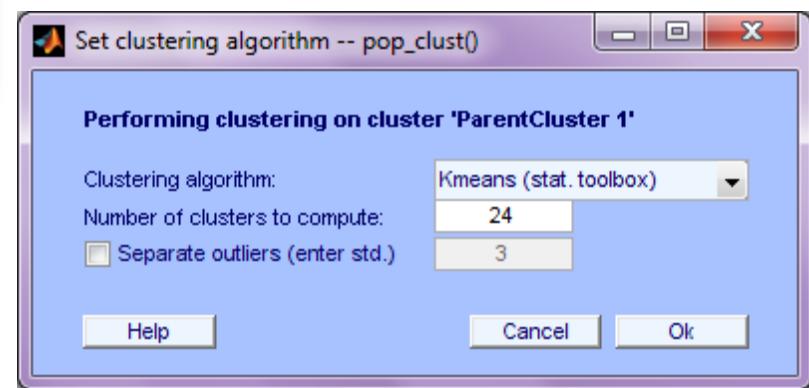
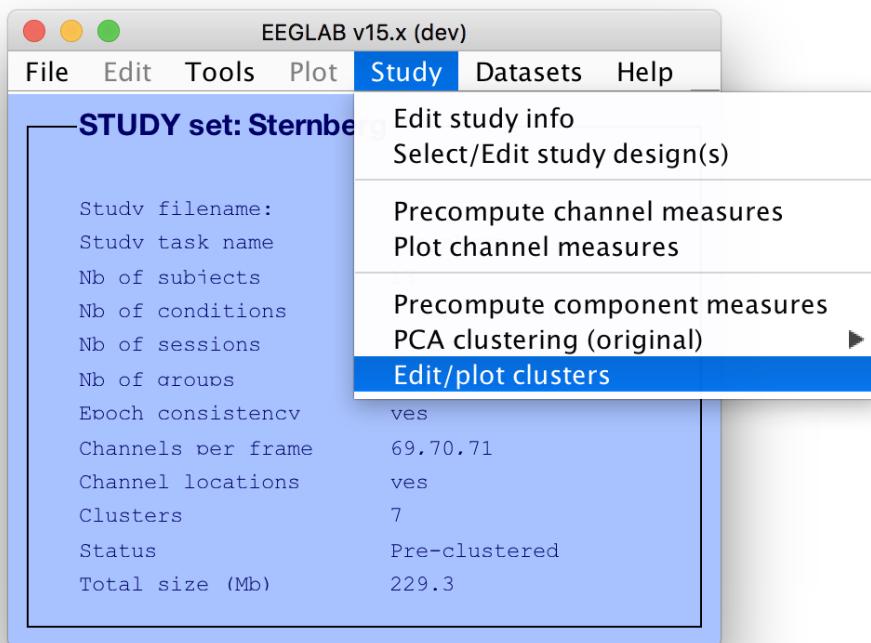
Each component is a dot
Clustering will group these dots



Standard Kmean Clustering



Cluster components



Choosing data measures

What measure(s) should you use?

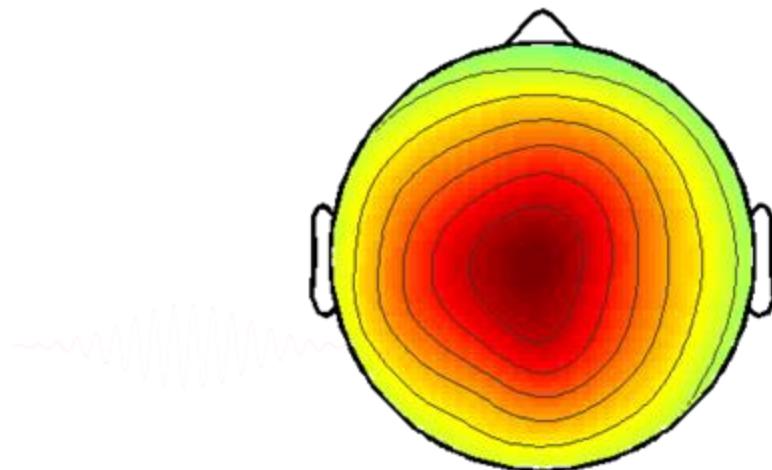
It depends on your final cluster criteria...

- If for example, your priority is dipole location,
then cluster only based on dipole location...

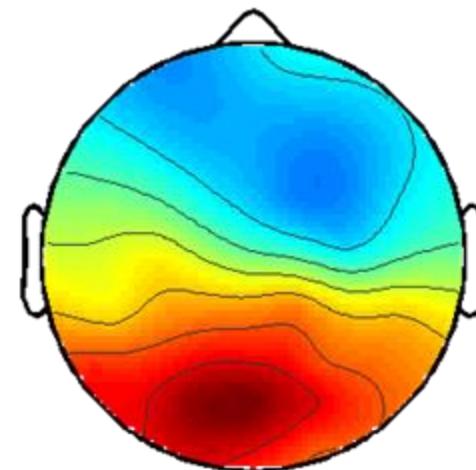
But consider:

- What is the difference between these two components?

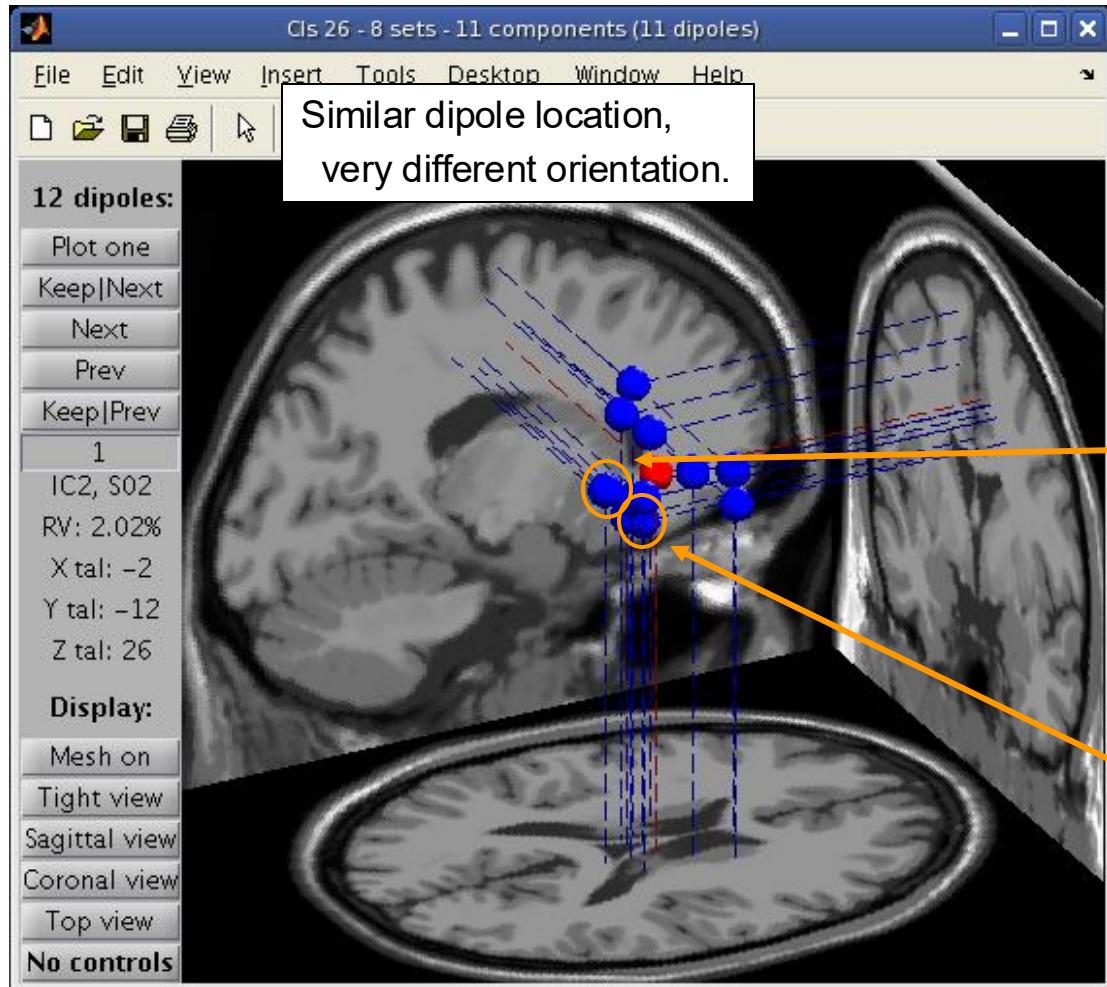
IC2 / S02, Cls 26



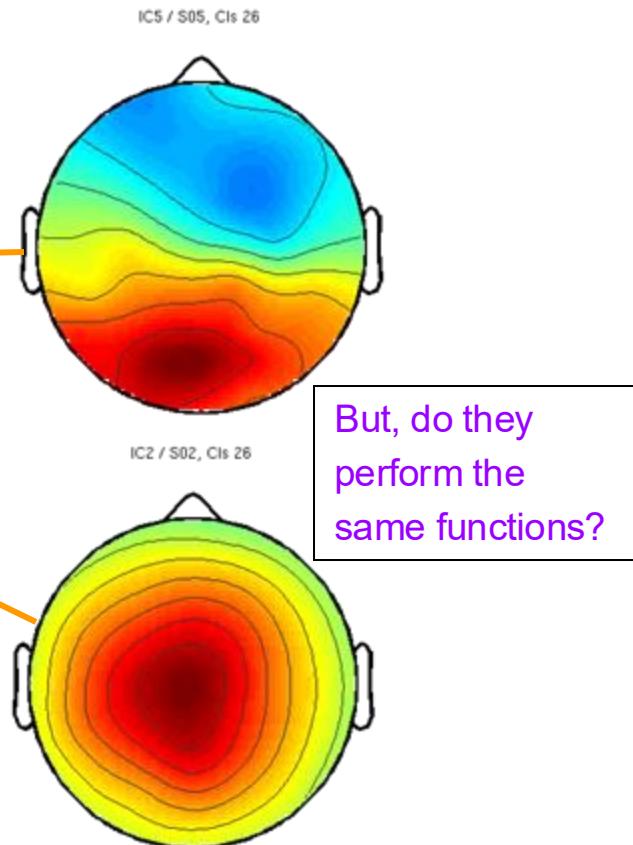
IC5 / S05, Cls 26



Choosing data measures



Obvious dramatic effect on
scalp map topography:

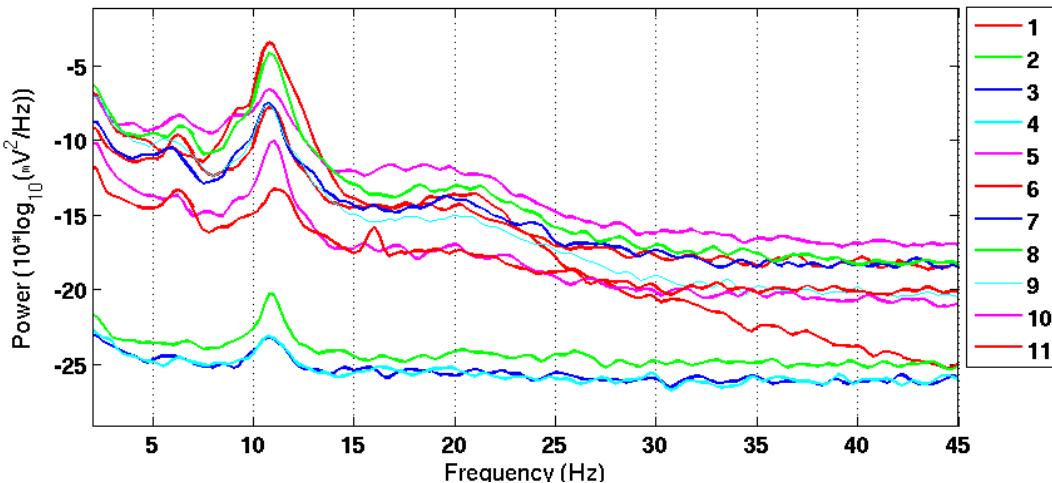
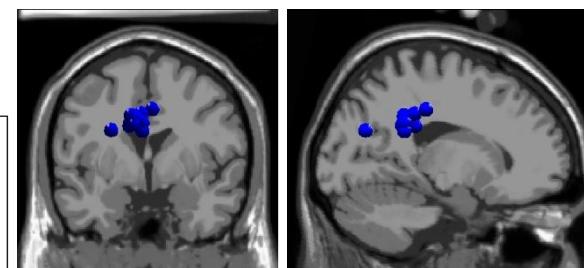
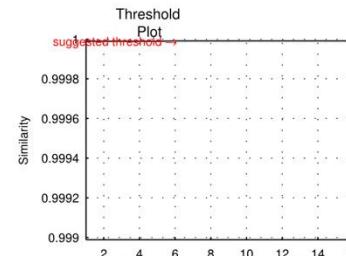
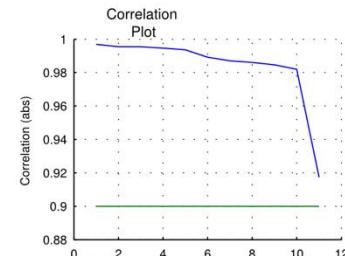
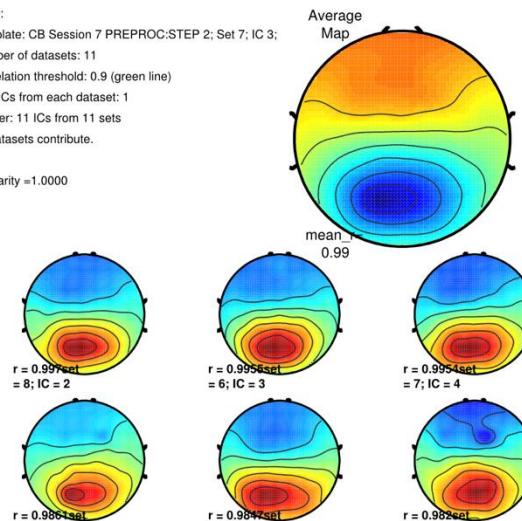


Results (Cluster 1 within subject)

100 % Sessions contribute

INFO:
Template: CB Session 7 PREPROC:STEP 2; Set 7; IC 3;
Number of datasets: 11
Correlation threshold: 0.9 (green line)
Max ICs from each dataset: 1
Cluster: 11 ICs from 11 sets
All datasets contribute.

Similarity = 1.0000

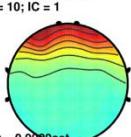
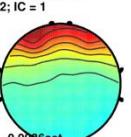
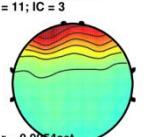
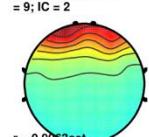
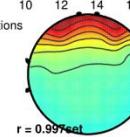
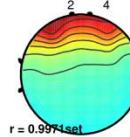
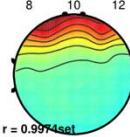
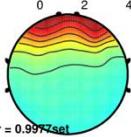
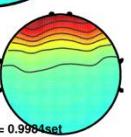
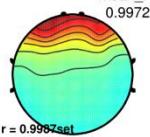
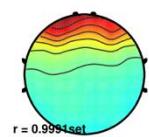
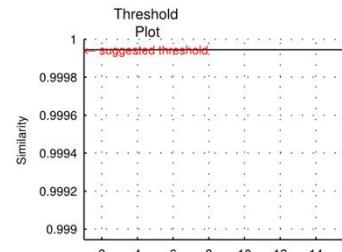
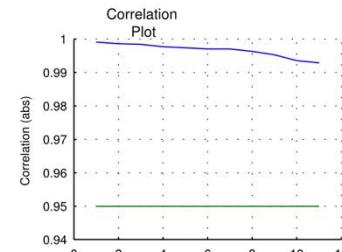
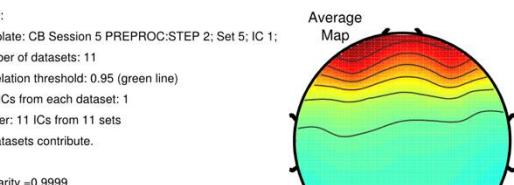


Results (Cluster 2 within subject)

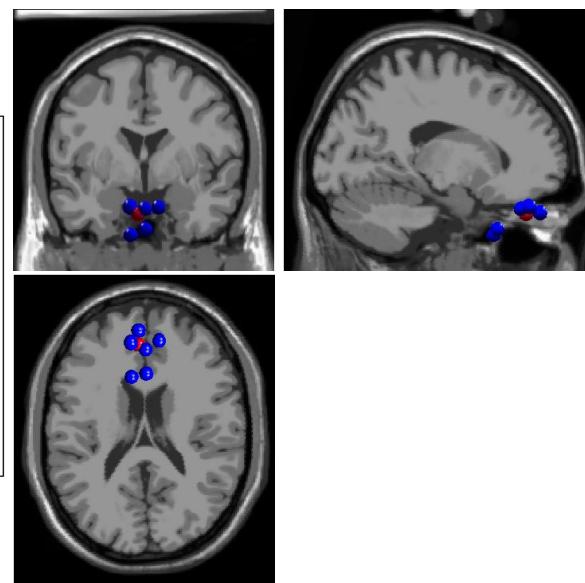
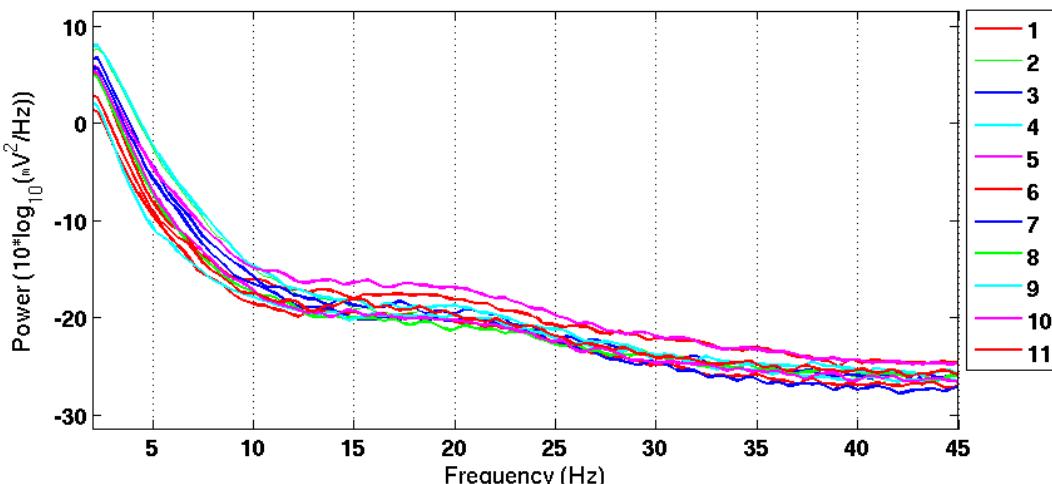
100 % Sessions contribute

INFO:
Template: CB Session 5 PREPROC:STEP 2; Set 5; IC 1;
Number of datasets: 11
Correlation threshold: 0.95 (green line)
Max ICs from each dataset: 1
Cluster: 11 ICs from 11 sets
All datasets contribute.

Similarity = 0.9999



Cl 4 Spectrum

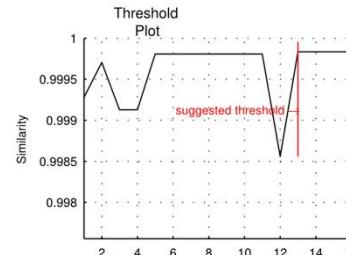
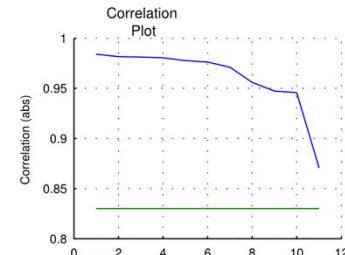
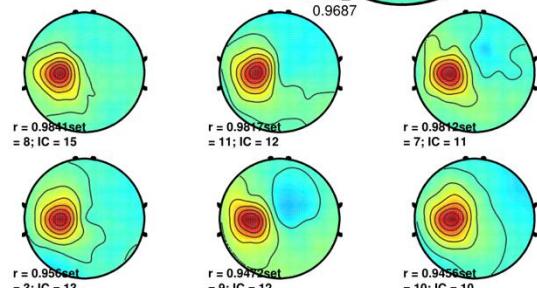
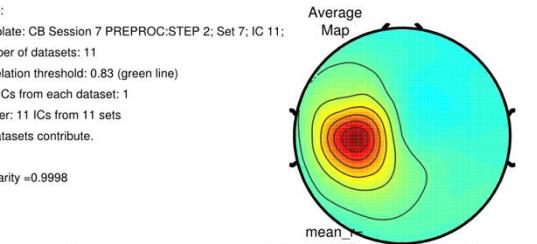


Results (Cluster 8 within subject)

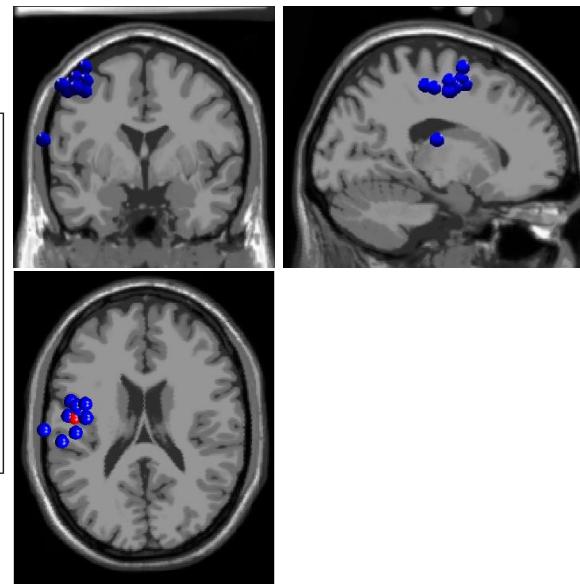
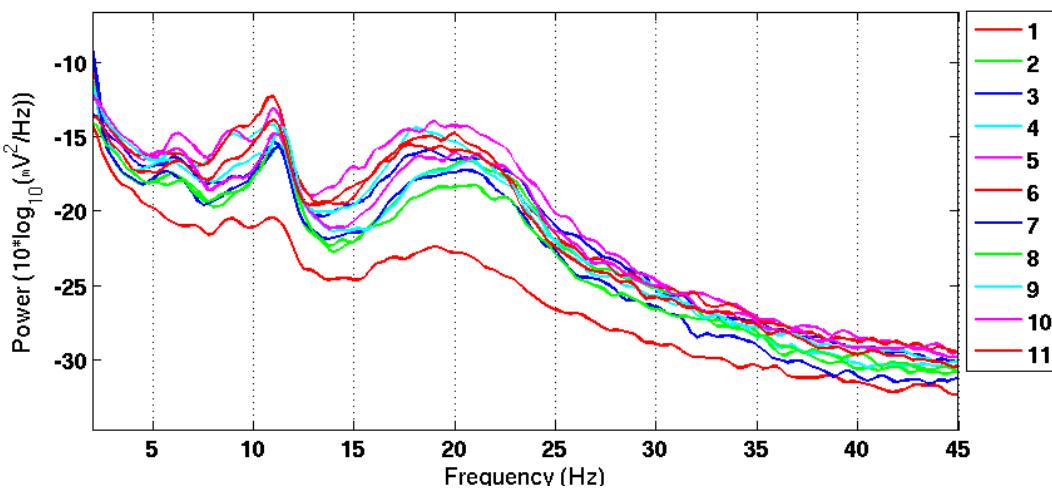
100 % Sessions contribute

INFO:
 Template: CB Session 7 PREPROC:STEP 2; Set 7; IC 11;
 Number of datasets: 11
 Correlation threshold: 0.83 (green line)
 Max ICs from each dataset: 1
 Cluster: 11 ICs from 11 sets
 All datasets contribute.

Similarity = 0.9998

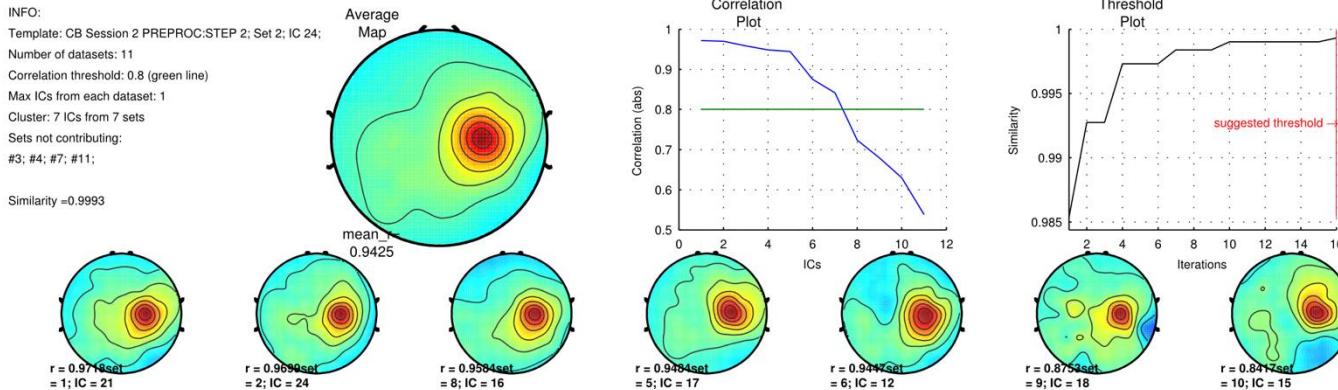


Cl 8 Spectrum

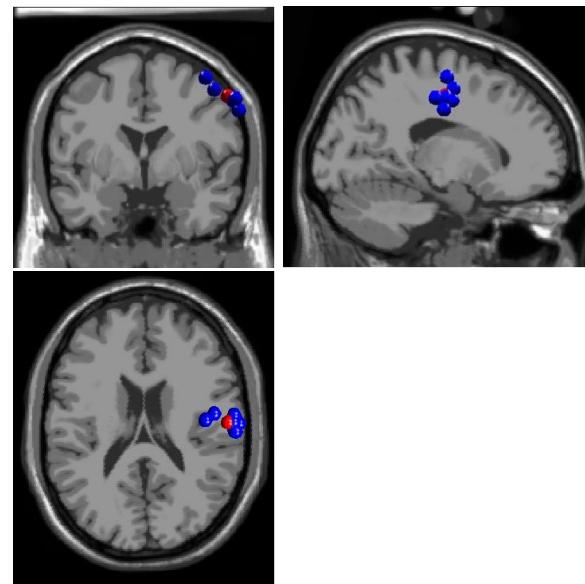
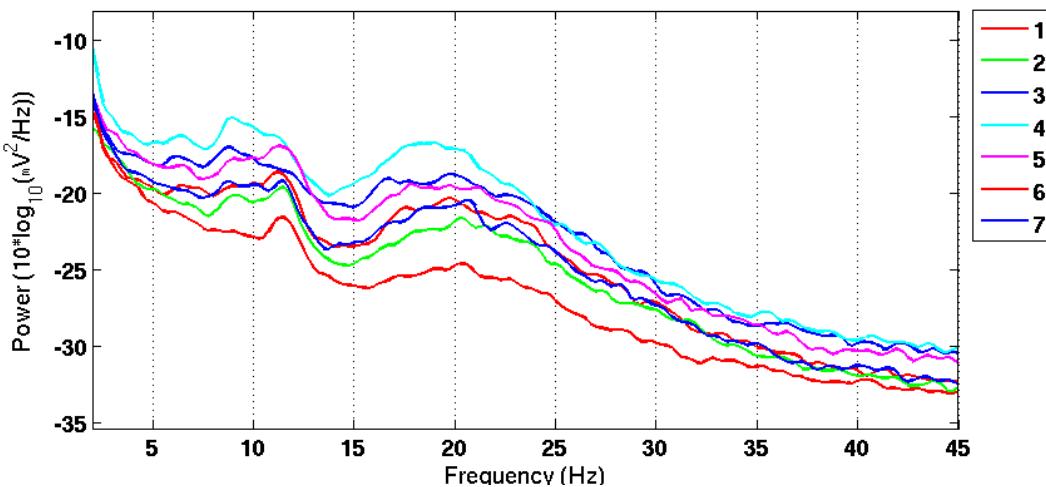


Results (Cluster 13 within subject)

63.64% Sessions contribute

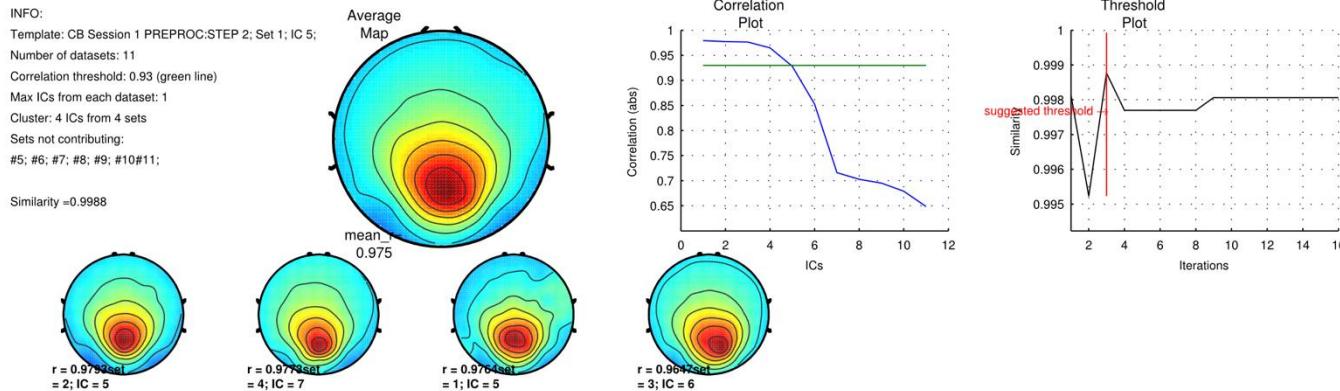


Cls 13 Spectrum

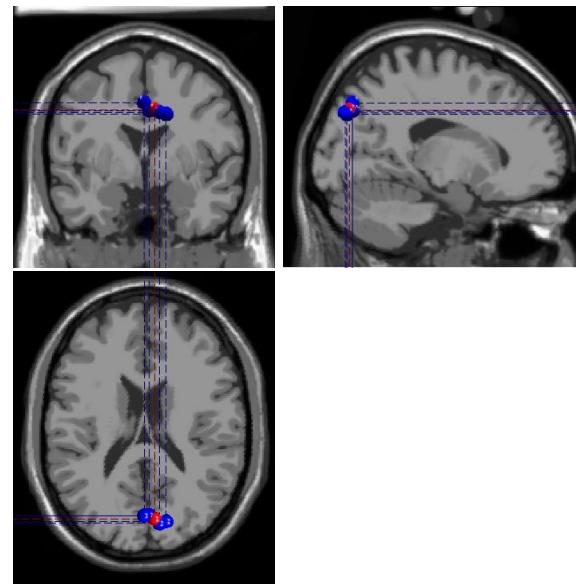
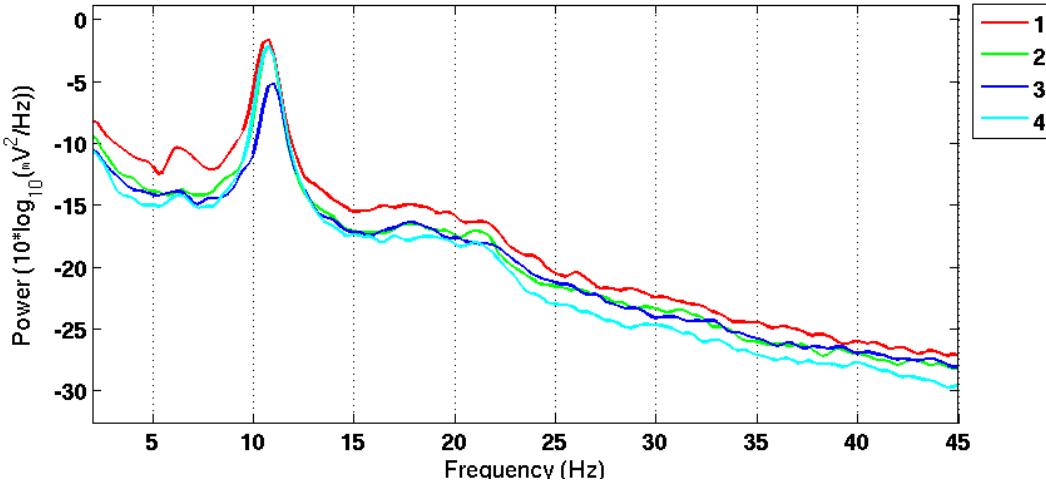


Results (Cluster 14 within subject)

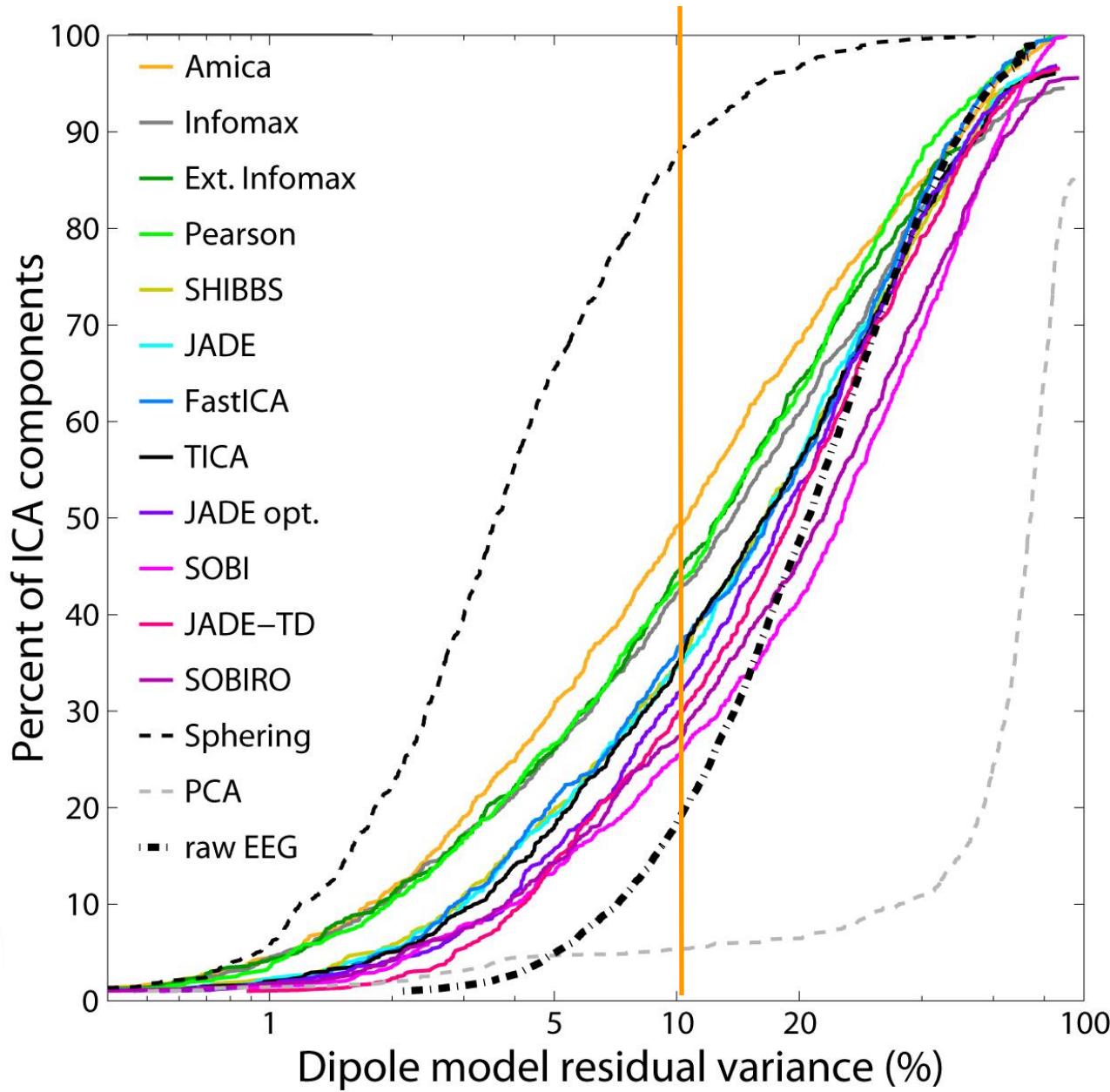
36.36% Sessions contribute



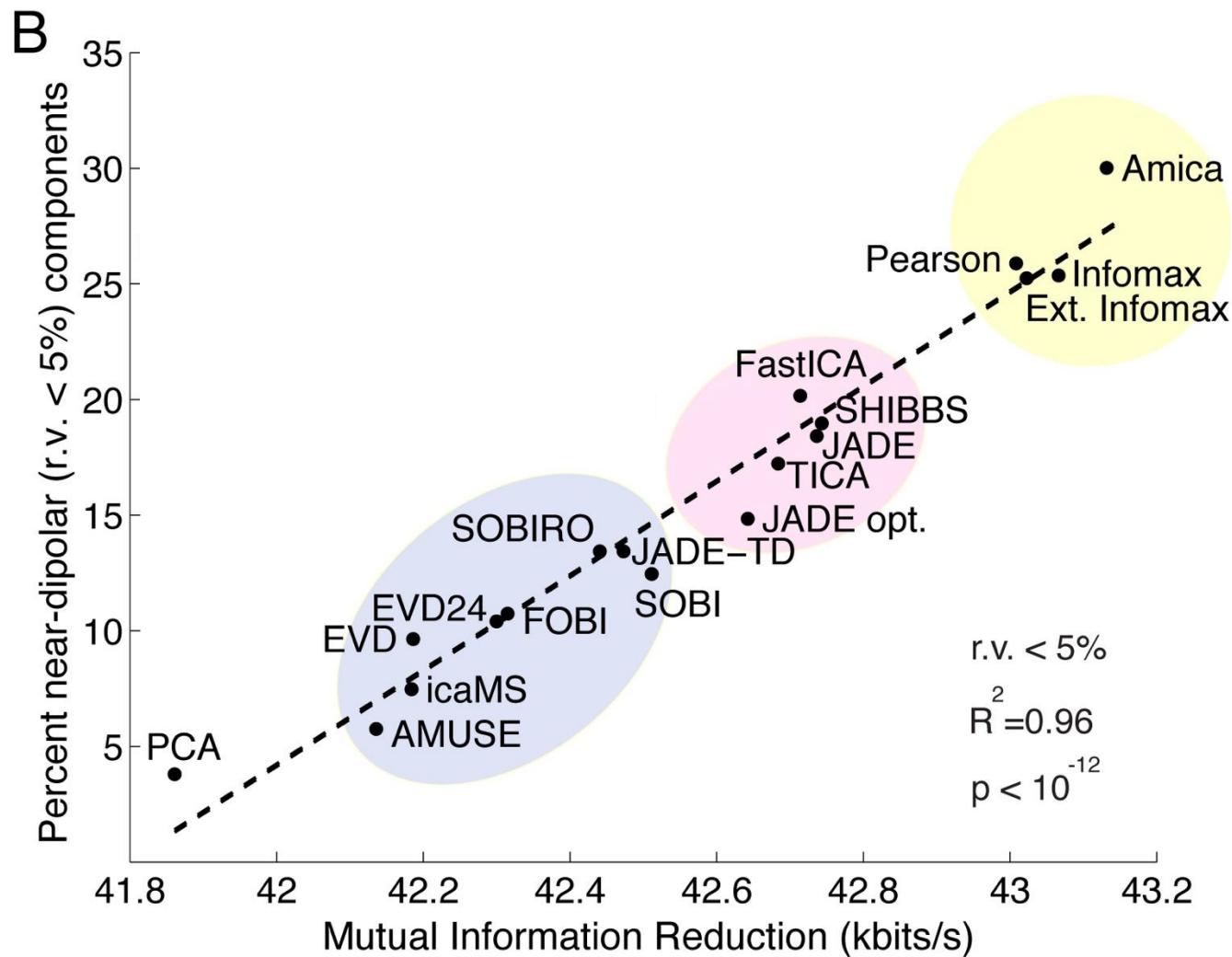
Cl 14 Spectrum

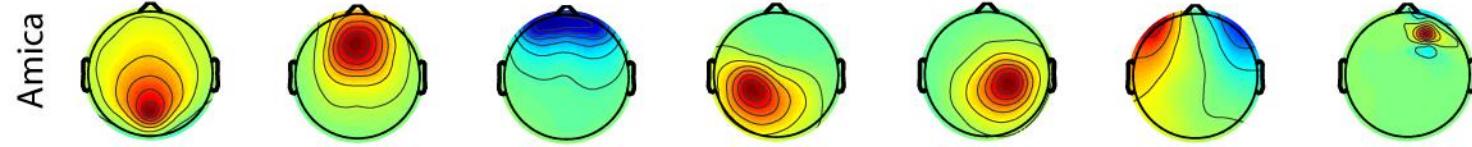
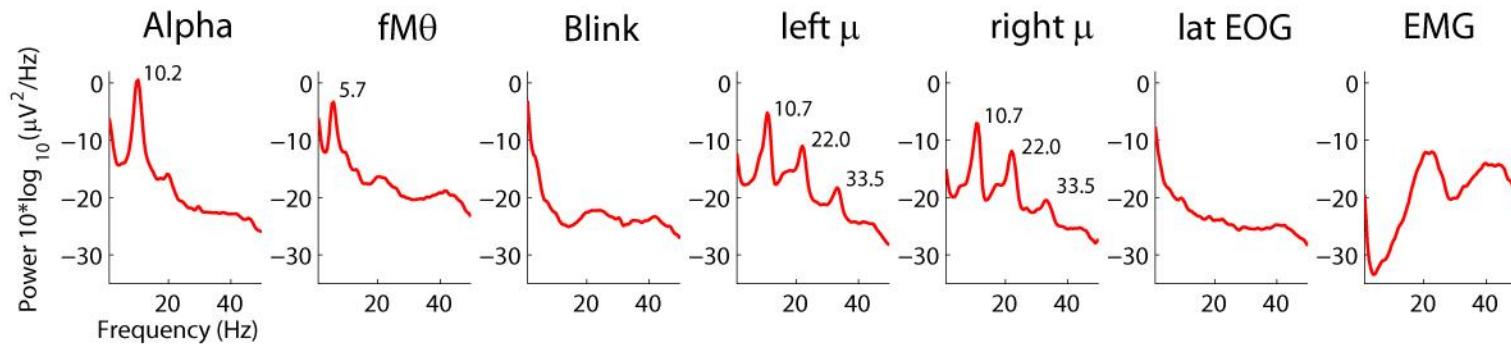


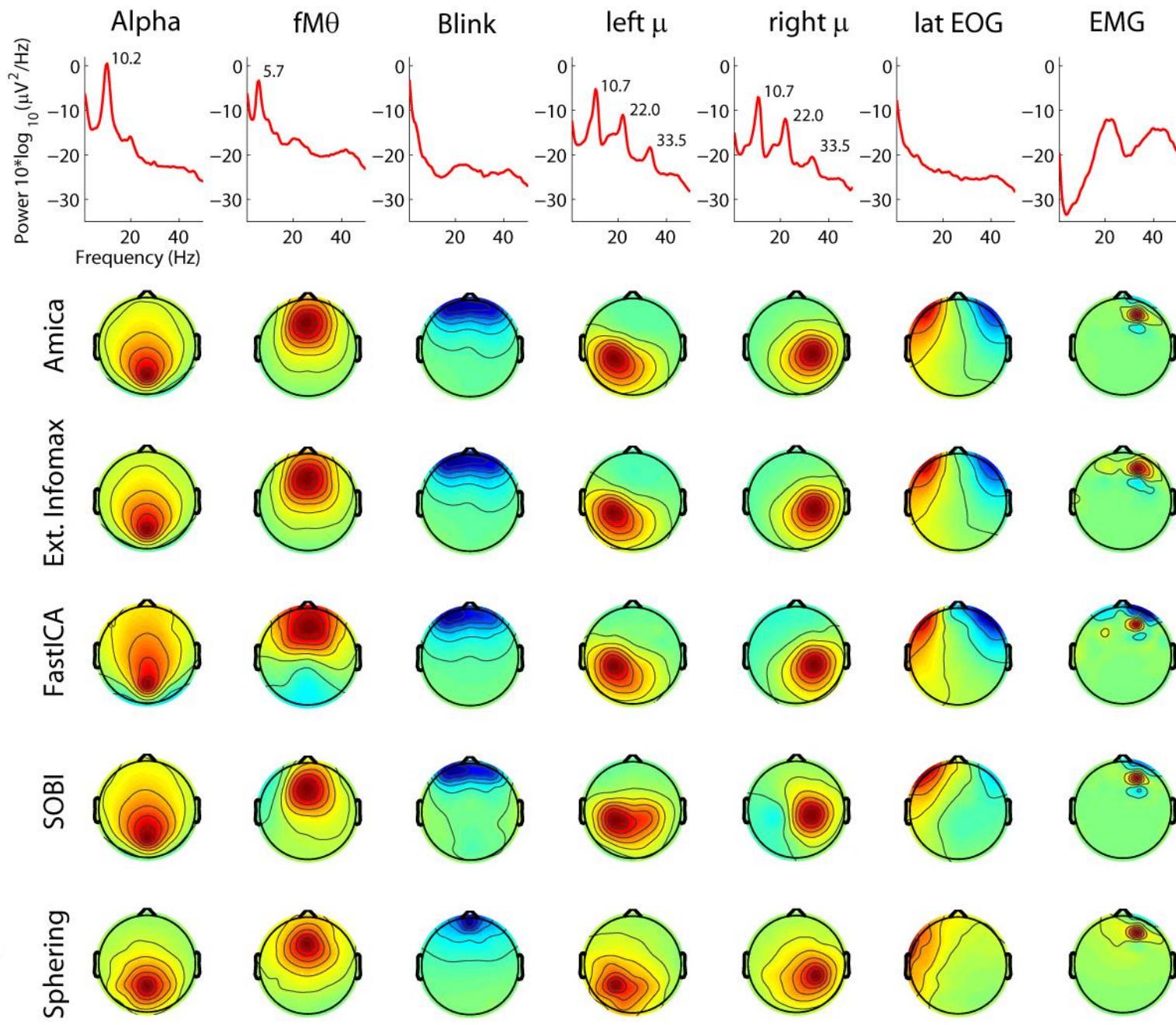
Number of components with residual variance lower than a specific threshold



More independence -> more biological components

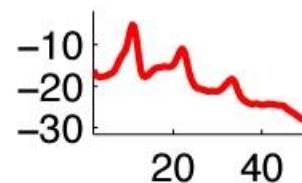




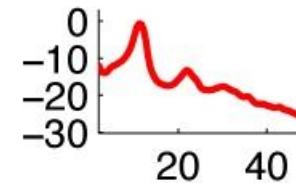
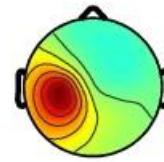


Left μ cluster (across subjects)

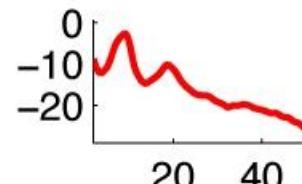
S2 IC47



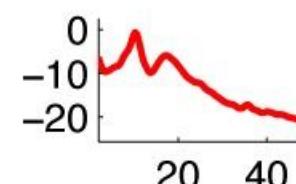
S3 IC47



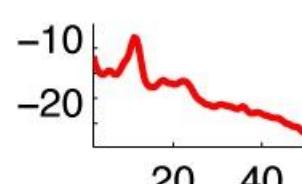
S4 IC37



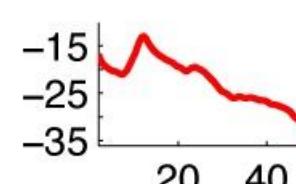
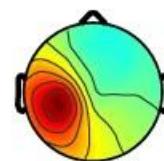
S5 IC48



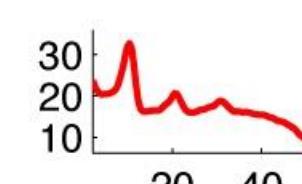
S6 IC46



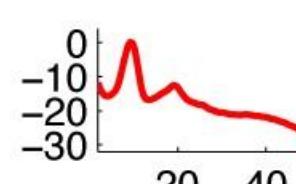
S7 IC35



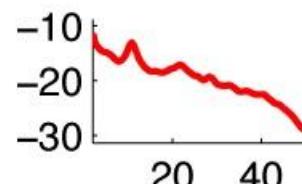
S9 IC7



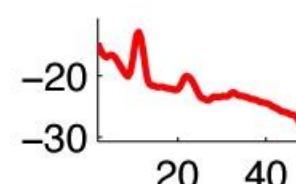
S11 IC45



S12 IC45

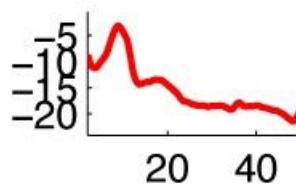


S14 IC45

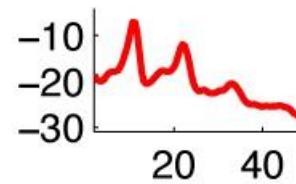


Right μ cluster

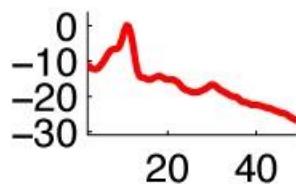
S1 IC51



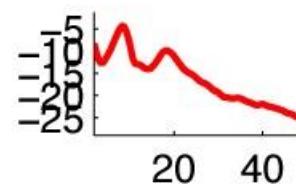
S2 IC41



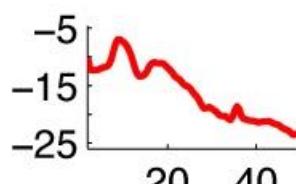
S3 IC41



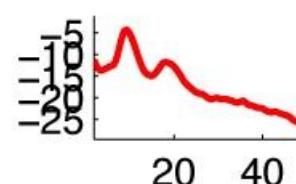
S4 IC50



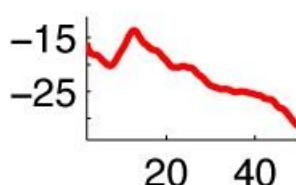
S5 IC51



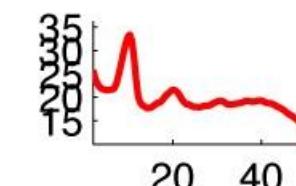
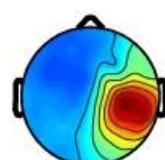
S6 IC 6^0



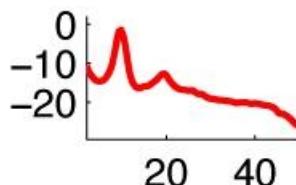
S7 IC48



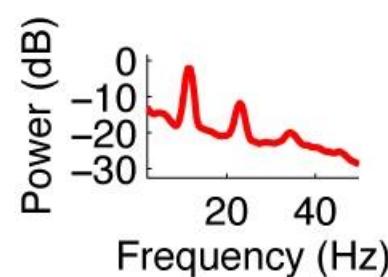
S9 IC39



S11 IC49

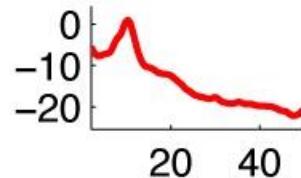


S14 IC49

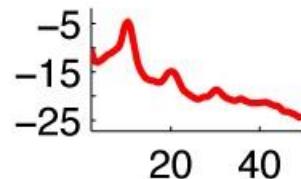


Occipital α cluster

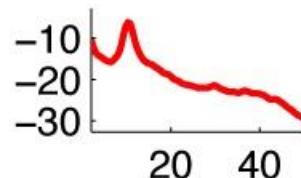
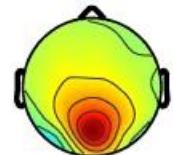
S1 IC67



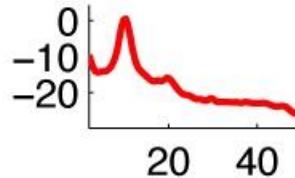
S3 IC51



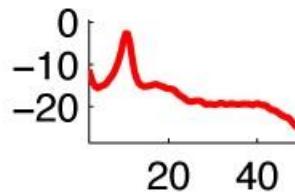
S12 IC3⁸



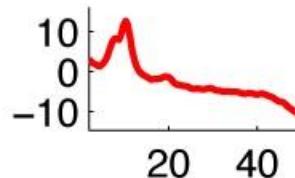
S2 IC67



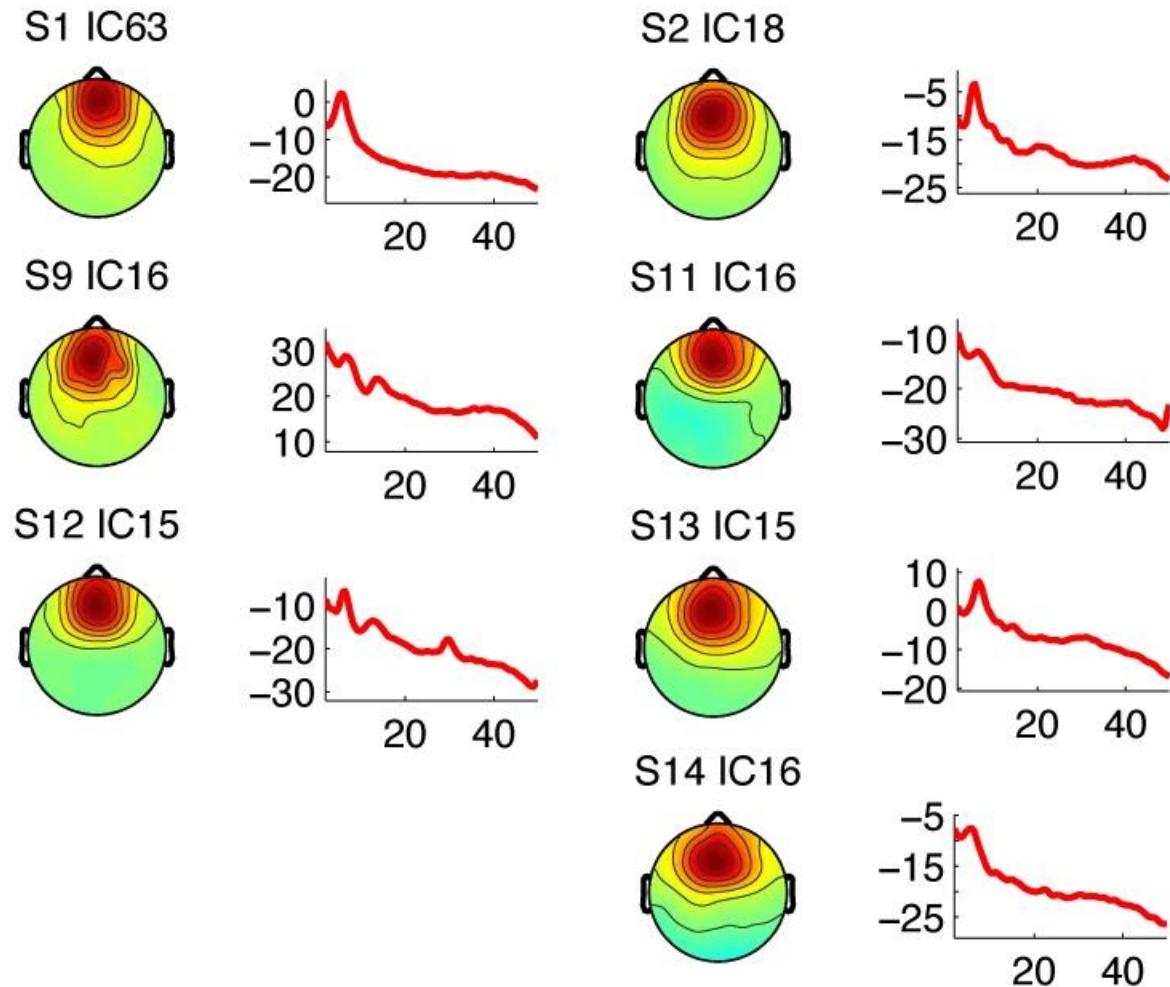
S11 IC65



S13 IC65



Frontal Midline θ cluster

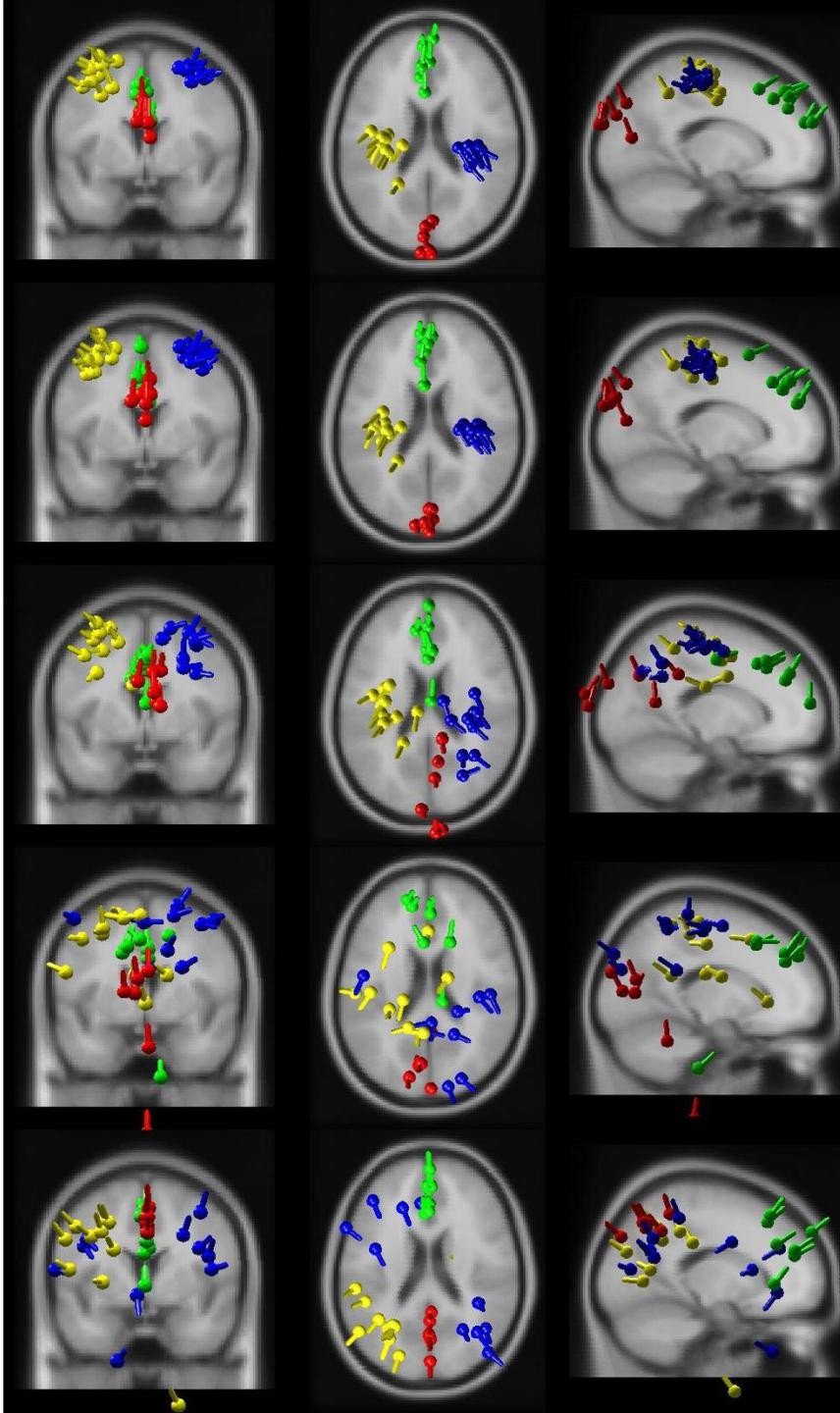


Delorme, A., Palmer, J., Onton, J., Oostenveld, R., & Makeig, S. (2012). Independent EEG sources are dipolar. *PLOS ONE*, 7(2), e30135. <https://doi.org/10.1371/journal.pone.0030135>



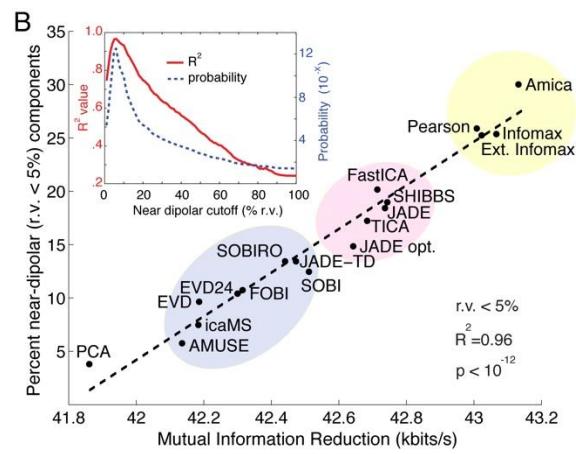
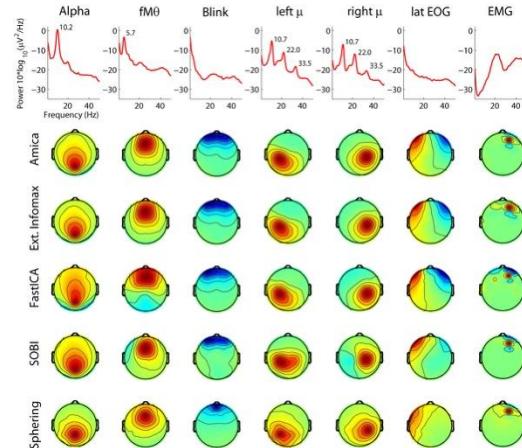
Sphering

SOBI FASTICA Ext. Infomax AMICA

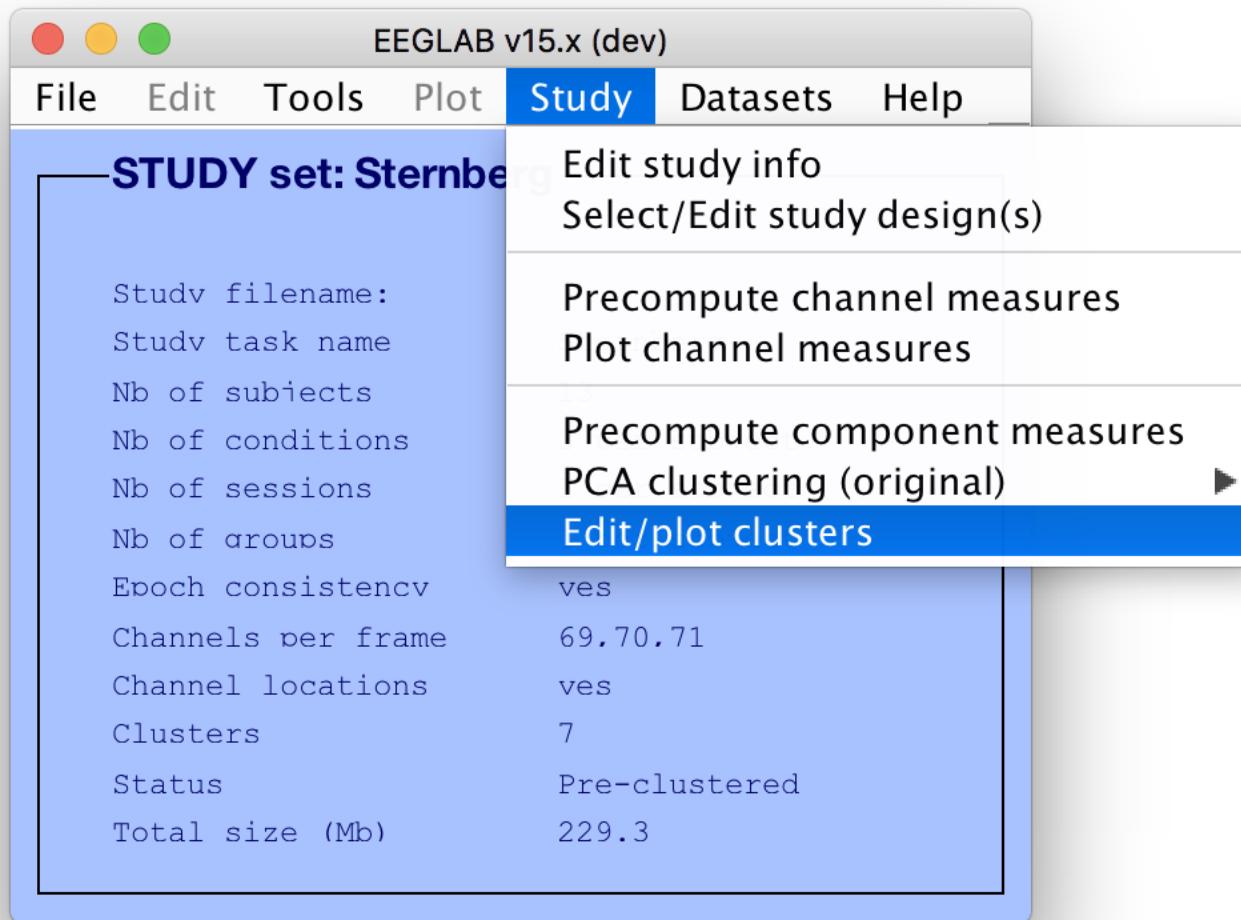


ICA reliability across subjects

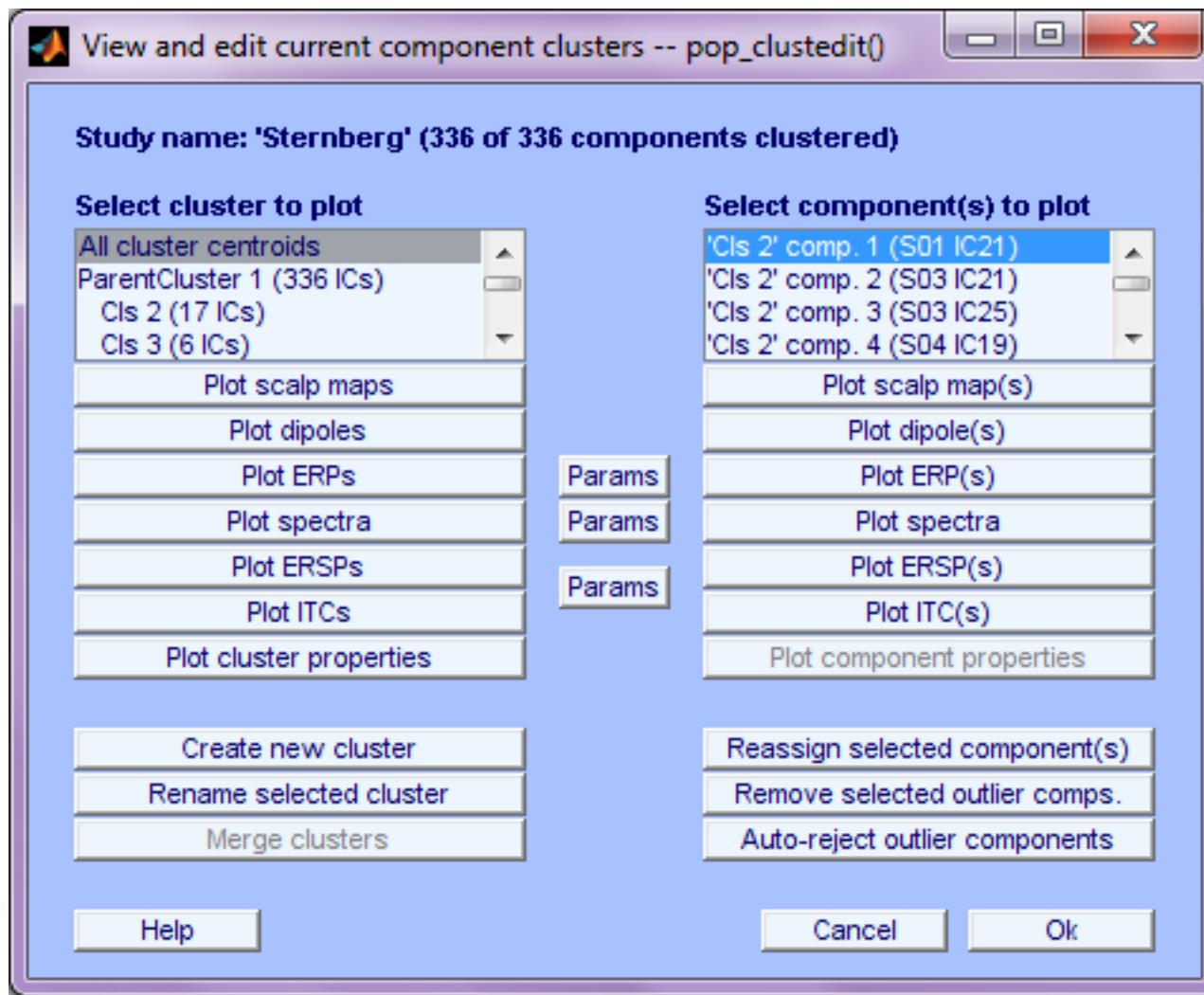
- Different ICA algorithms return similar solutions
- The ICA algorithms that return the most biologically plausible solutions are also the one that return the most independent decompositions



View and edit clusters



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 ERSPs

Plot ITCs

Plot cluster properties

Create new cluster

Rename selected cluster

Merge clusters

Save STUDY set to disk /home/julie/

Cancel Help

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 dipole(s)

Plot ERP(s)

Plot spectra

Figure 3: Average scalp map for all clusters

Plot mean scalp maps for easy reference

Average scalp map for all clusters

Cluster Name	ICs	Ss
outlier 2	12	7
Cls 3	4	4
Cls 4	8	8
Cls 5	7	7
Cls 6	3	3
Cls 7	10	6
Cls 8	5	5
Cls 9	12	8
Cls 10	4	4
Cls 11	9	8
Cls 12	3	2
Cls 13	5	4
Cls 14	5	5
Cls 15	14	9
Cls 16	3	3
Cls 17	7	7
Cls 18	4	4
Cls 19	10	8
Cls 20	17	15
Cls 21	6	6
Cls 22	11	8
Cls 23	6	7
Cls 24	7	7
Cls 25	3	3
Cls 26	3	3
Cls 27	1	1

Plot cluster data

Choose which cluster

Choose which components

Study 'Attention': 181 of 181 components clustered

Select cluster to plot

- Cls 6 (3 ICs)
- Cls 7 (10 ICs)**
- Cls 8 (5 ICs)
- Cls 9 (12 ICs)

Plot scalp maps
Plot dipoles
Plot ERPs
Plot spectra
Plot ERSPs
Plot ITCs
Plot cluster properties

Params
Params
Params

Create new cluster
Rename selected cluster
Merge clusters

Save STUDY set to disk /home/julie/WorkshopSD2

Cancel Help

View and edit current component clusters -- pop_clustedit()

Select component(s) to plot

- All components
- S01 IC6
- S05 IC9
- S06 IC12

Plot scalp map(s)
Plot dipole(s)
Plot ERPs
Plot spectra
Plot ITCs
Plot cluster properties

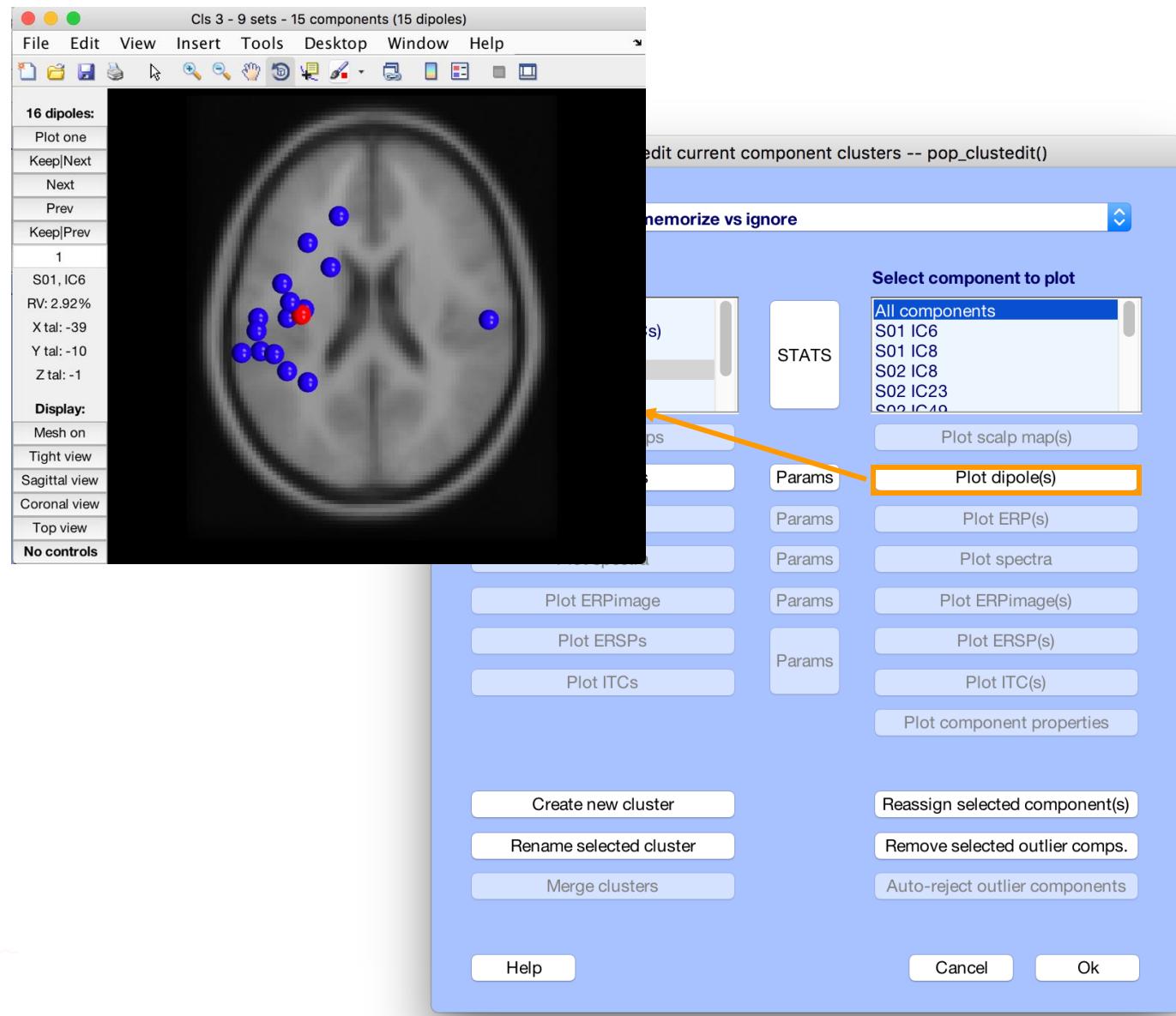
Params
Params
Params

Reassign selected component
Remove selected component
Auto-reject component

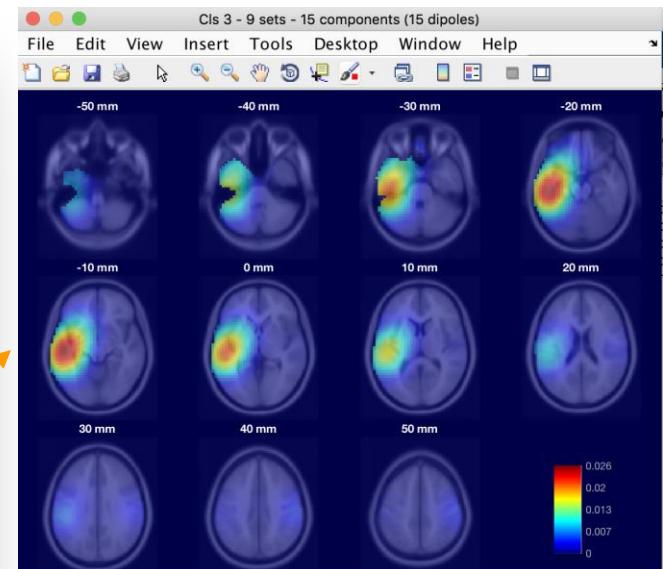
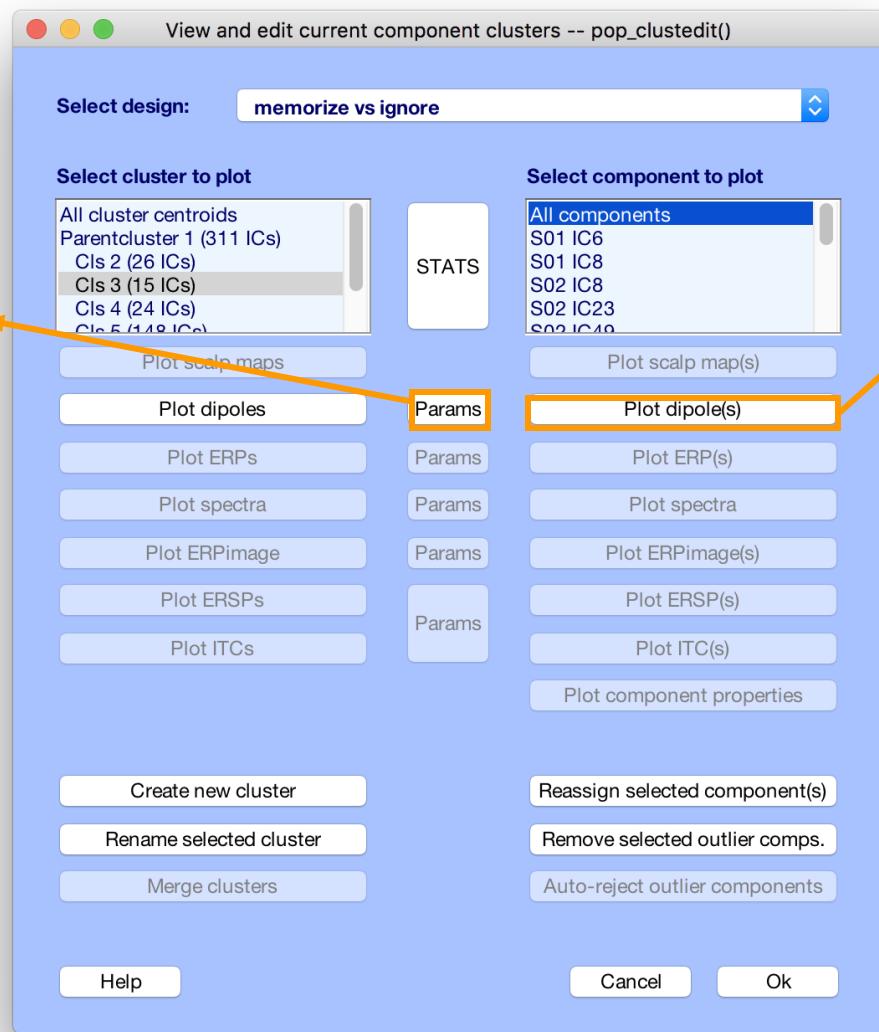
Figure 4

The figure window contains nine scalp maps arranged in a grid. The top row shows 'Cls 21 (7 ICs, 7 Ss)' and three maps labeled 'ic4/S01', 'ic11/S02', and 'ic1/S06'. The middle row shows 'ic2/S10' and 'ic6/S12'. The bottom row shows 'ic3/S15'. Each scalp map displays a topographic distribution of activity across electrode sites.

Plot cluster data

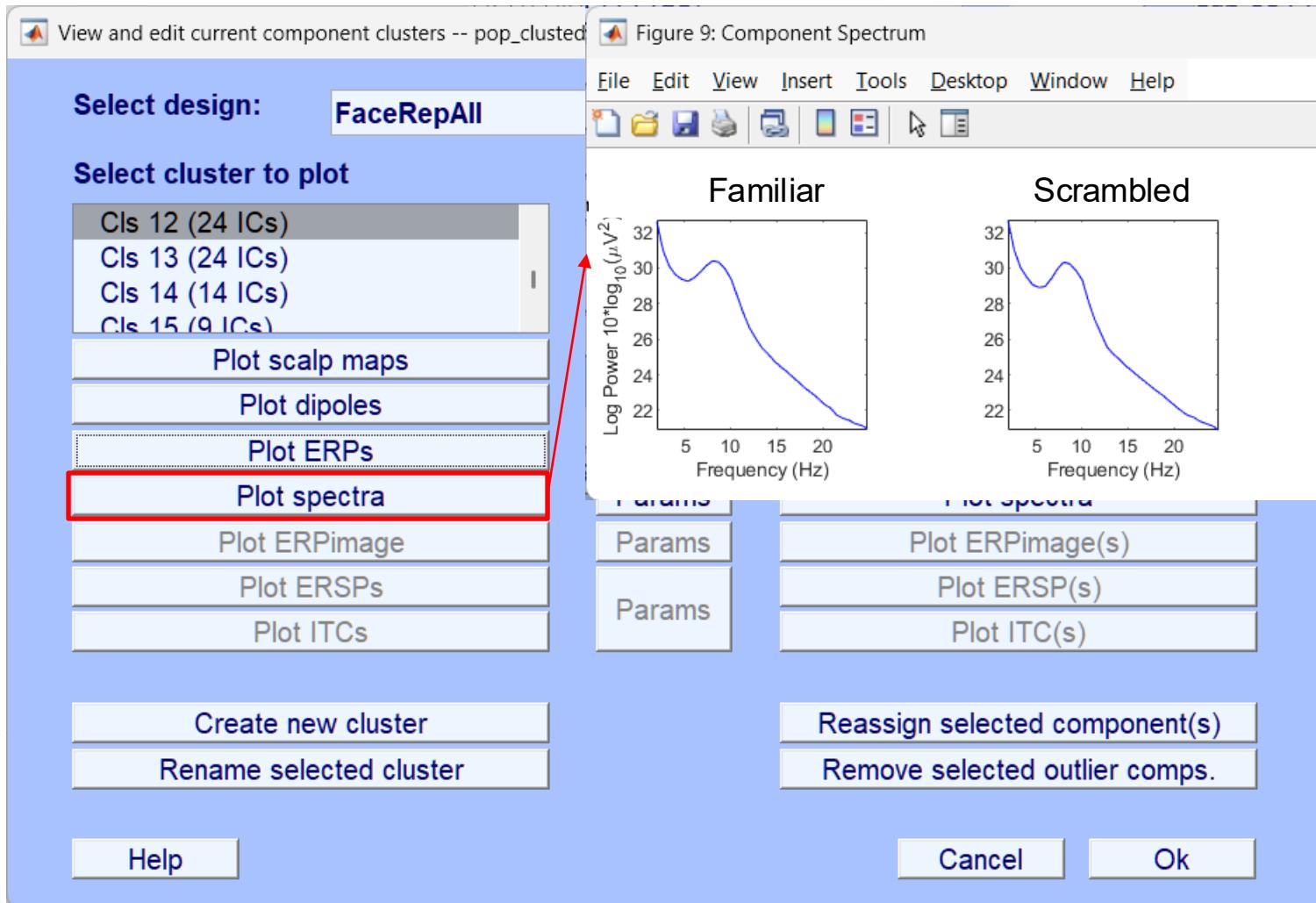


Plot cluster data

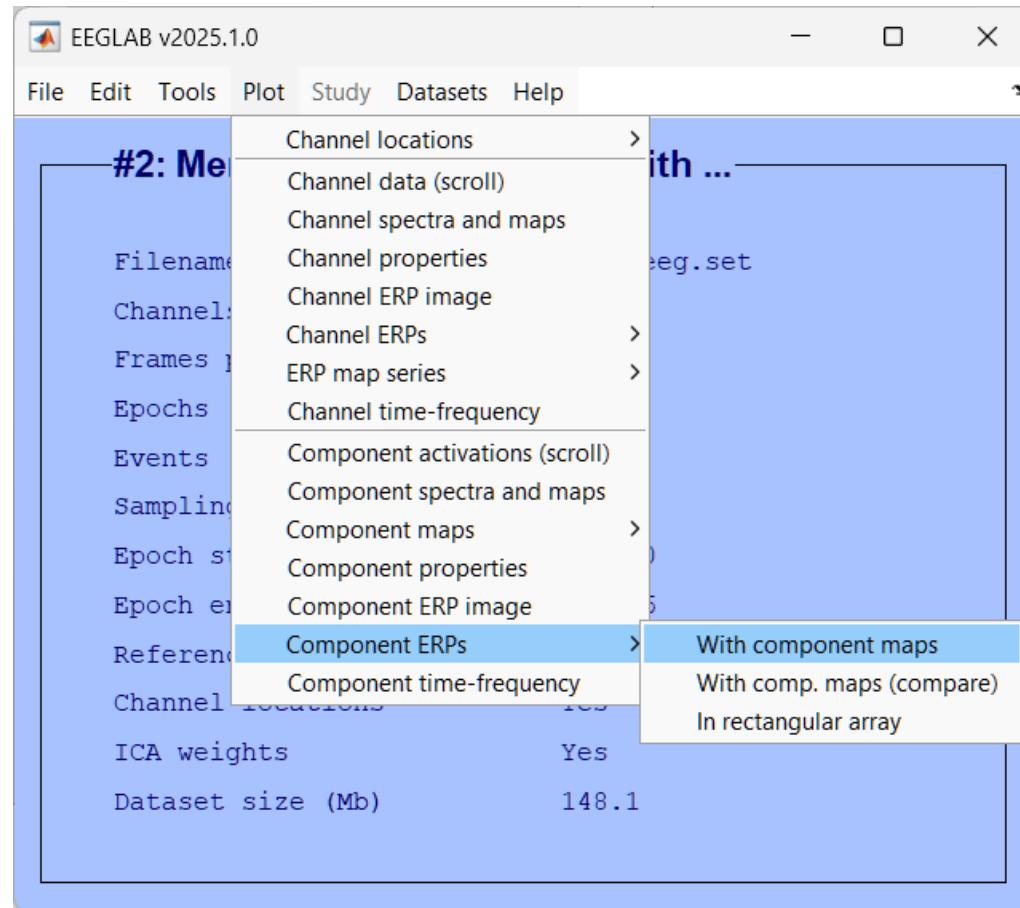


Plot cluster data

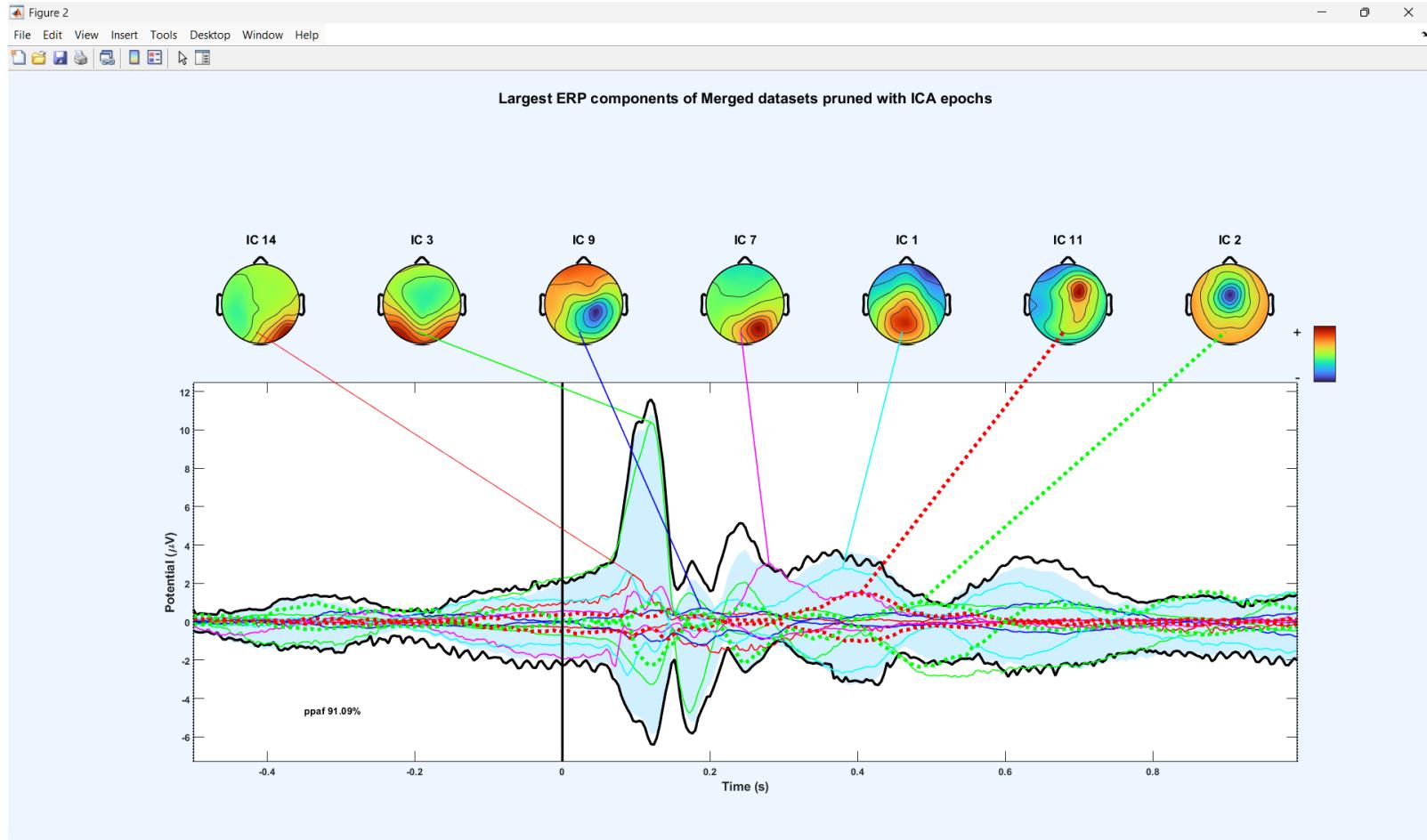




IC contributions to ERP envelope



IC contributions to ERP envelope



Hands on

- Load the STUDY ds002718/Face_detection.study
(download ds002718 from NEMAR and run the script Session_5_ds002718_preproc.m if needed)
 - Precompute **spectrum** for components, scalp topo.
 - Precluster and cluster components using **dipole locations** and **dipole moments** (KMean)
 - Look at your clusters. Identify frontal midline theta cluster(s) and occipital alpha cluster(s)
 - Look for a cluster generating a N170 and check dipoles location
 - Remove outliers if any