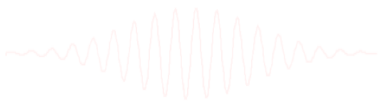


Clustering of ICA components

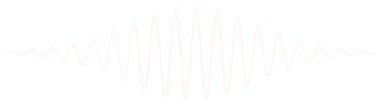
Arnaud Delorme

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



Steps of clustering

- Select ICA components for clustering
- Precompute measures of interest
- Cluster measures
- Plot clusters and edit them if necessary



Edit dataset info

Create a new STUDY set -- pop_study()

Edit STUDY set information - remember to save changes

STUDY set name: Sternberg

STUDY set task name: Sternberg

STUDY set notes:

	dataset filename	browse	subject	session	condition	group	Select by r.v.	
1	C:\Users\julie\Documents\Wvor	...	S01		memorize		Comp.: 3 5 ...	Clear
2	C:\Users\julie\Documents\Wvor	...	S01		ignore		Comp.: 3 5 ...	Clear
3	C:\Users\julie\Documents\Wvor	...	S01		probe		Comp.: 3 5 ...	Clear
4	C:\Users\julie\Documents\Wvor	...	S02		memorize		Comp.: 5 6 ...	Clear
5	C:\Users\julie\Documents\Wvor	...	S02		ignore		Comp.: 5 6 ...	Clear
6	C:\Users\julie\Documents\Wvor	...	S02		probe		Comp.: 5 6 ...	Clear
7	C:\Users\julie\Documents\Wvor	...	S03		memorize		Comp.: 6 7 ...	Clear
8	C:\Users\julie\Documents\Wvor	...	S03		ignore		Comp.: 6 7 ...	Clear
9	C:\Users\julie\Documents\Wvor	...	S03		probe		Comp.: 6 7 ...	Clear
10	C:\Users\julie\Documents\Wvor	...	S04		memorize		Comp.: 1 2 ...	Clear

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

< Page 1 >

☐ 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

pop_study(): Pre-select components

Enter maximum residual (topo map - dipole proj.) var. (in %)
NOTE: This will delete any existing component clusters!

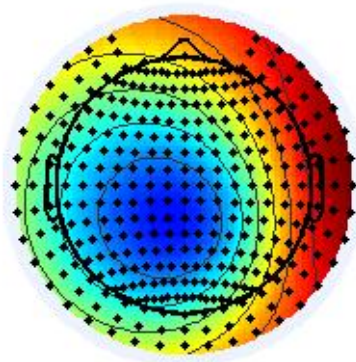
15

☒ Keep only in-brain dipoles.

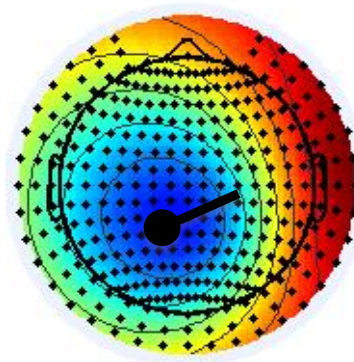
Cancel Help Ok

Computing residual variance (%)

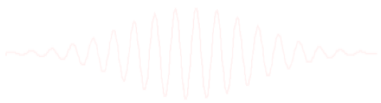
Actual



Dipole projection

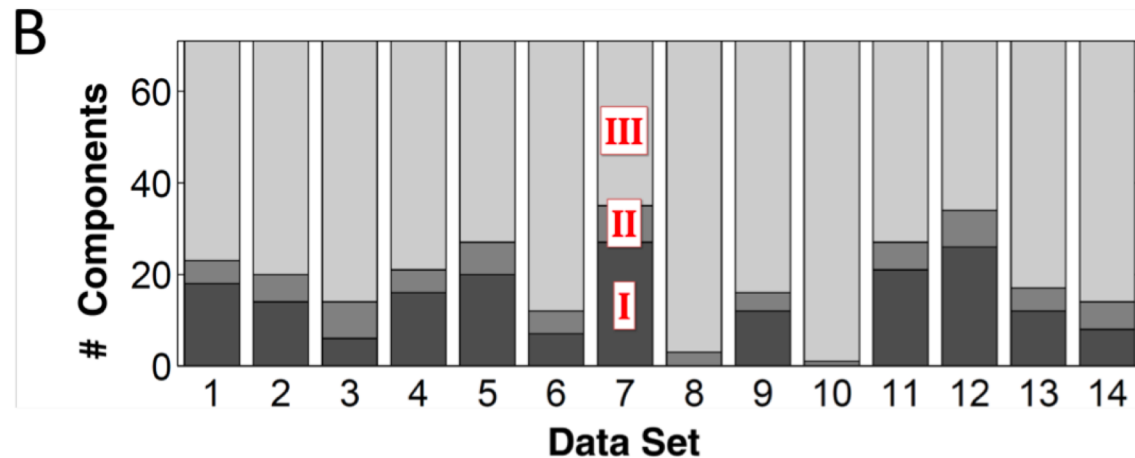
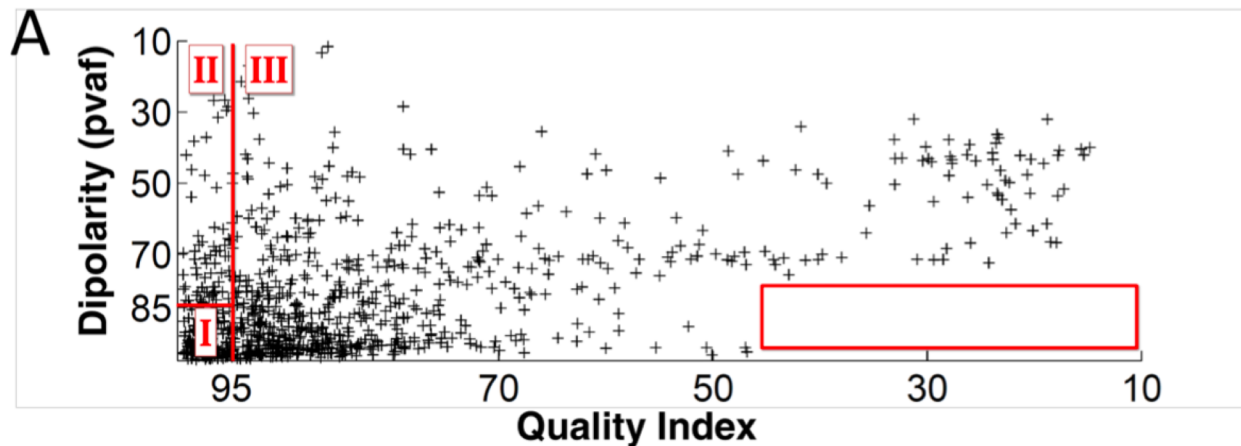


$$r = \Sigma(x_i - \tilde{x}_i)^2 / \Sigma x_i^2$$



Reliability criteria and the $rv < 15\%$

First justification why we should select an $rv < 15\%$ 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

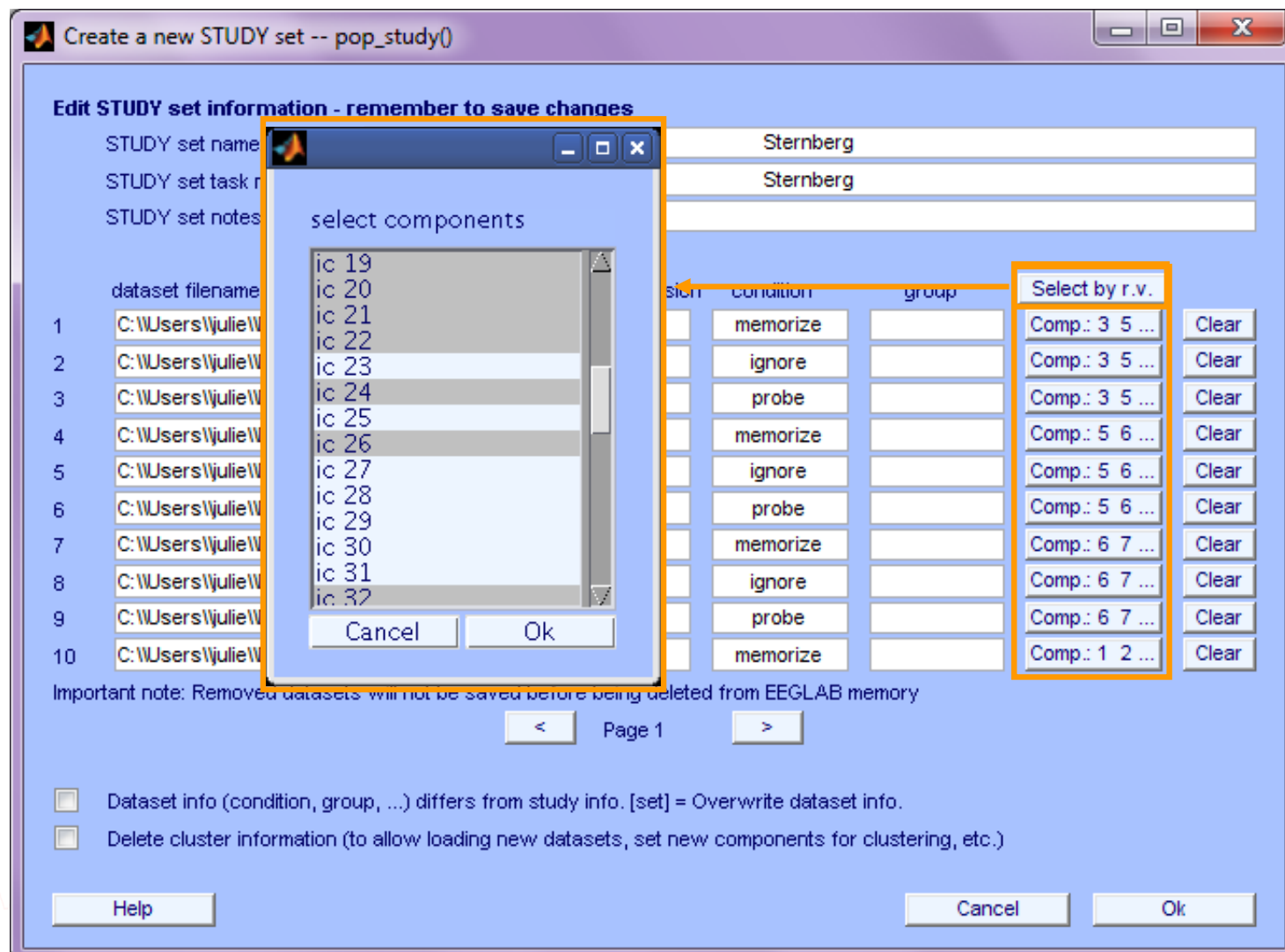
$$dip \pm std > th$$

$$dip \pm std < th$$

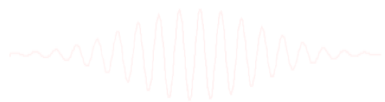
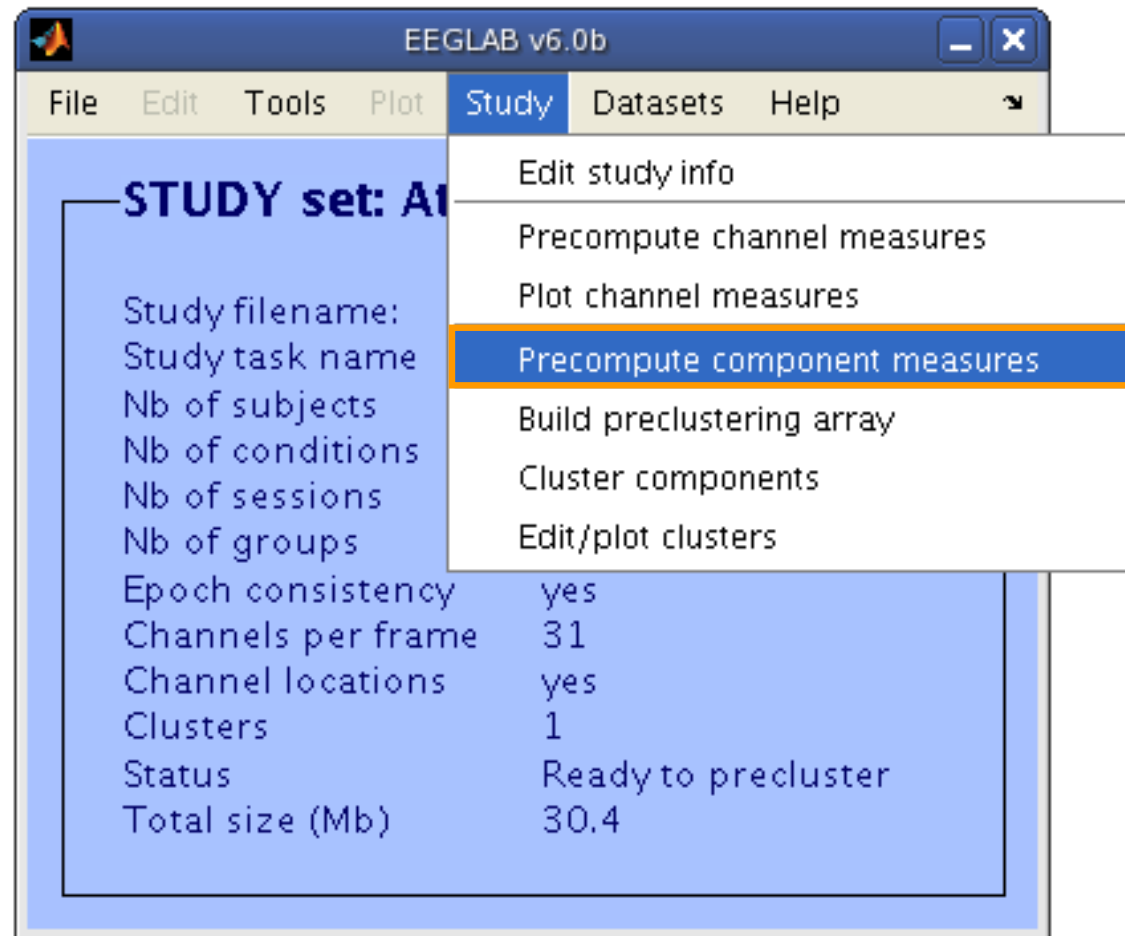
Probable Inseparable noise: variance explained useful or **multiple subject confirmation**

Discard

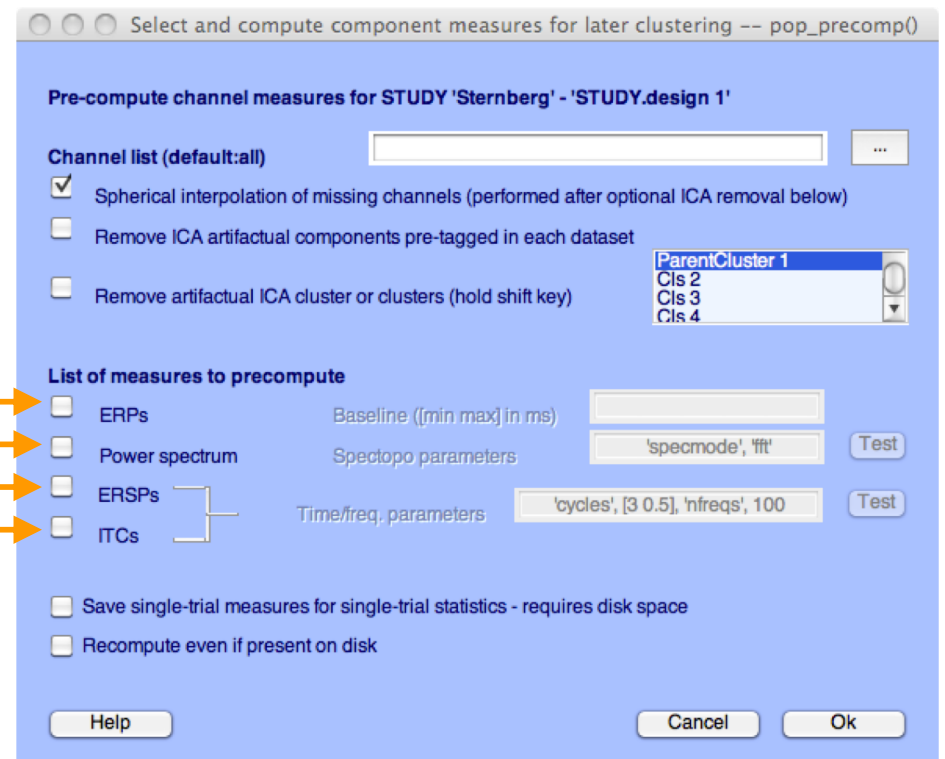
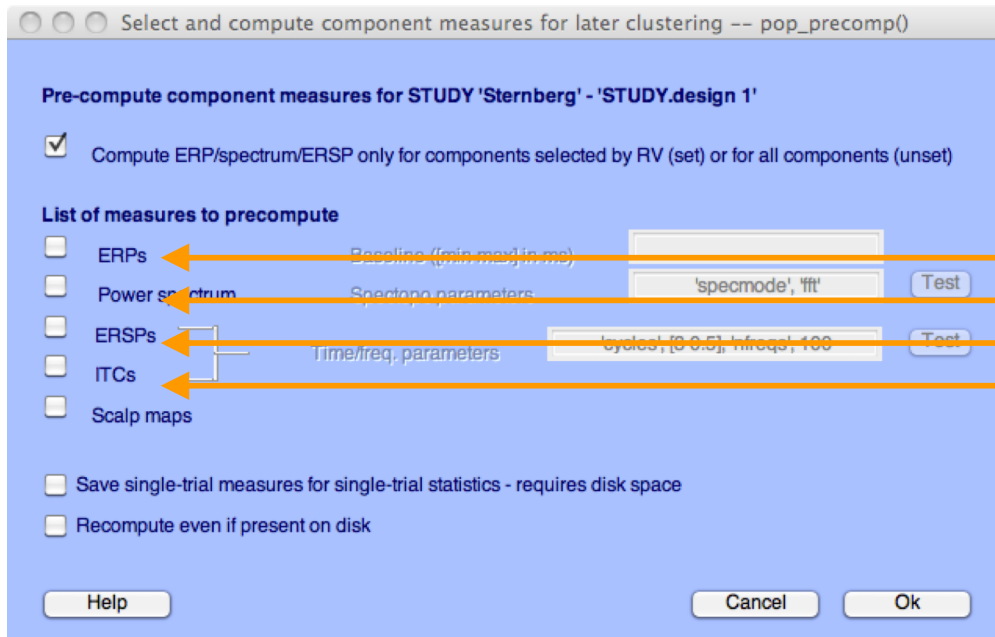
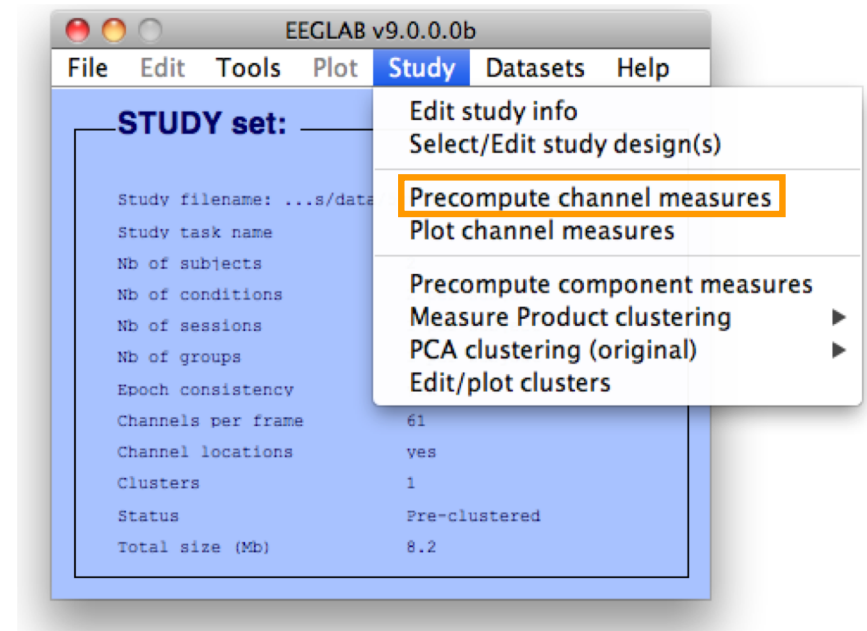
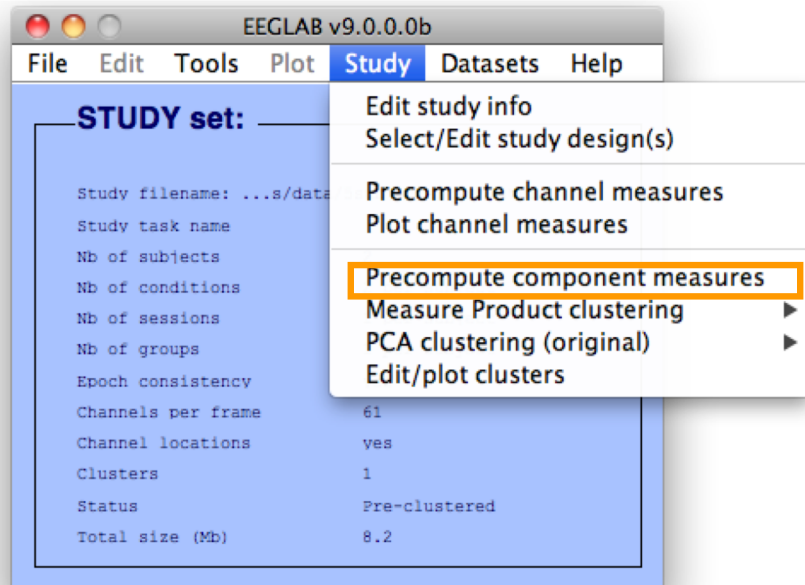
ICs to cluster



Precompute data measures



Pre-compute measures



Precompute data measures

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

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

Pre-compute component measures for STUDY 'Sternberg'

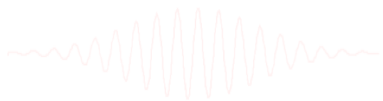
☒ Compute ERP/spectrum/ERSP only for components selected by RV (set) or for all components (unset)

List of measures to precompute

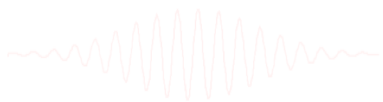
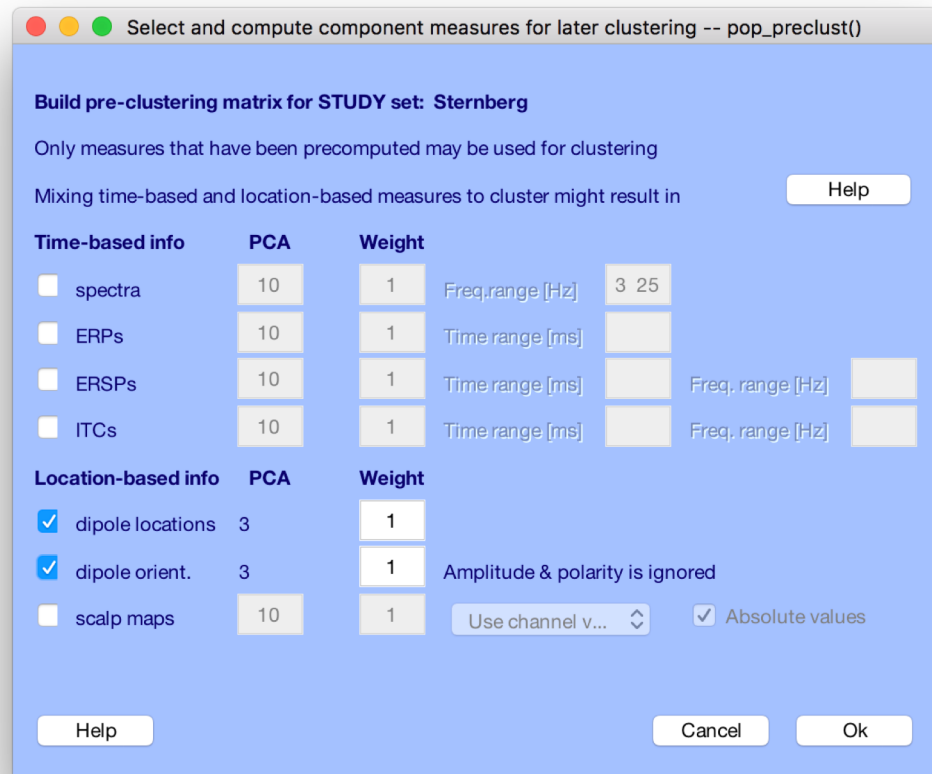
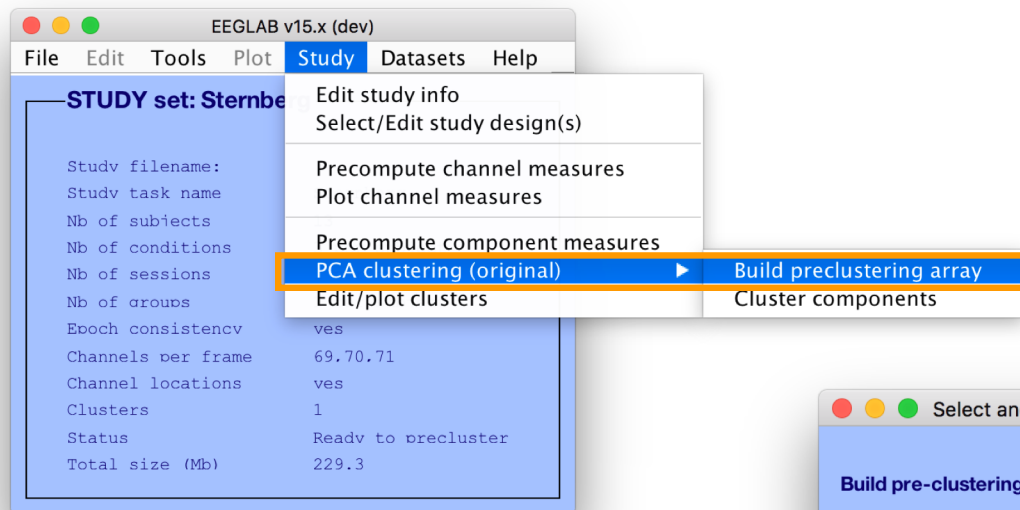
<input checked="" type="checkbox"/> ERPs	Baseline ([min max] in ms)	<input type="text" value="[-200 0]"/>	
<input checked="" type="checkbox"/> Power spectrum	Spectopo parameters	<input type="text"/>	<input type="button" value="Test"/>
<input checked="" type="checkbox"/> ERSPs	Time/freq. parameters	<input type="text" value="'cycles', [3 0.5], 'nfreqs', 100"/>	<input type="button" value="Test"/>
<input checked="" type="checkbox"/> ITCs			
<input checked="" type="checkbox"/> Scalp maps			

☐ Recompute even if present on disk

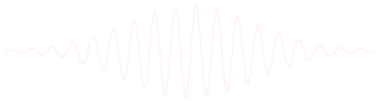
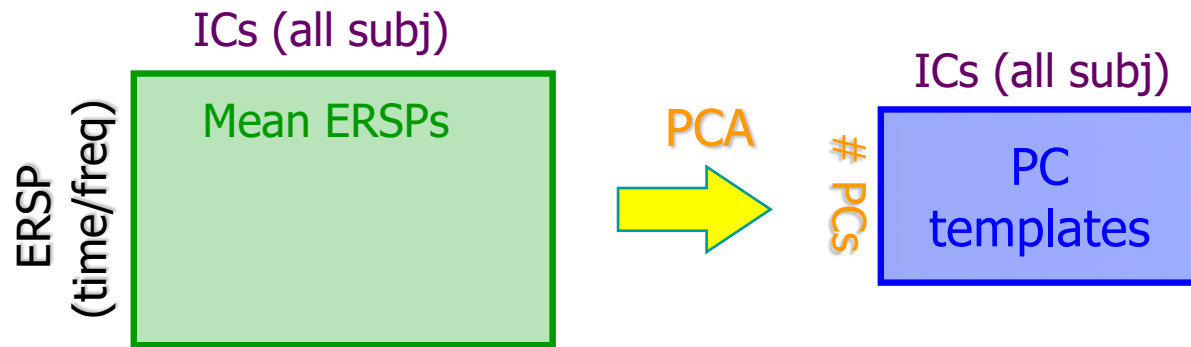
Time-frequency options



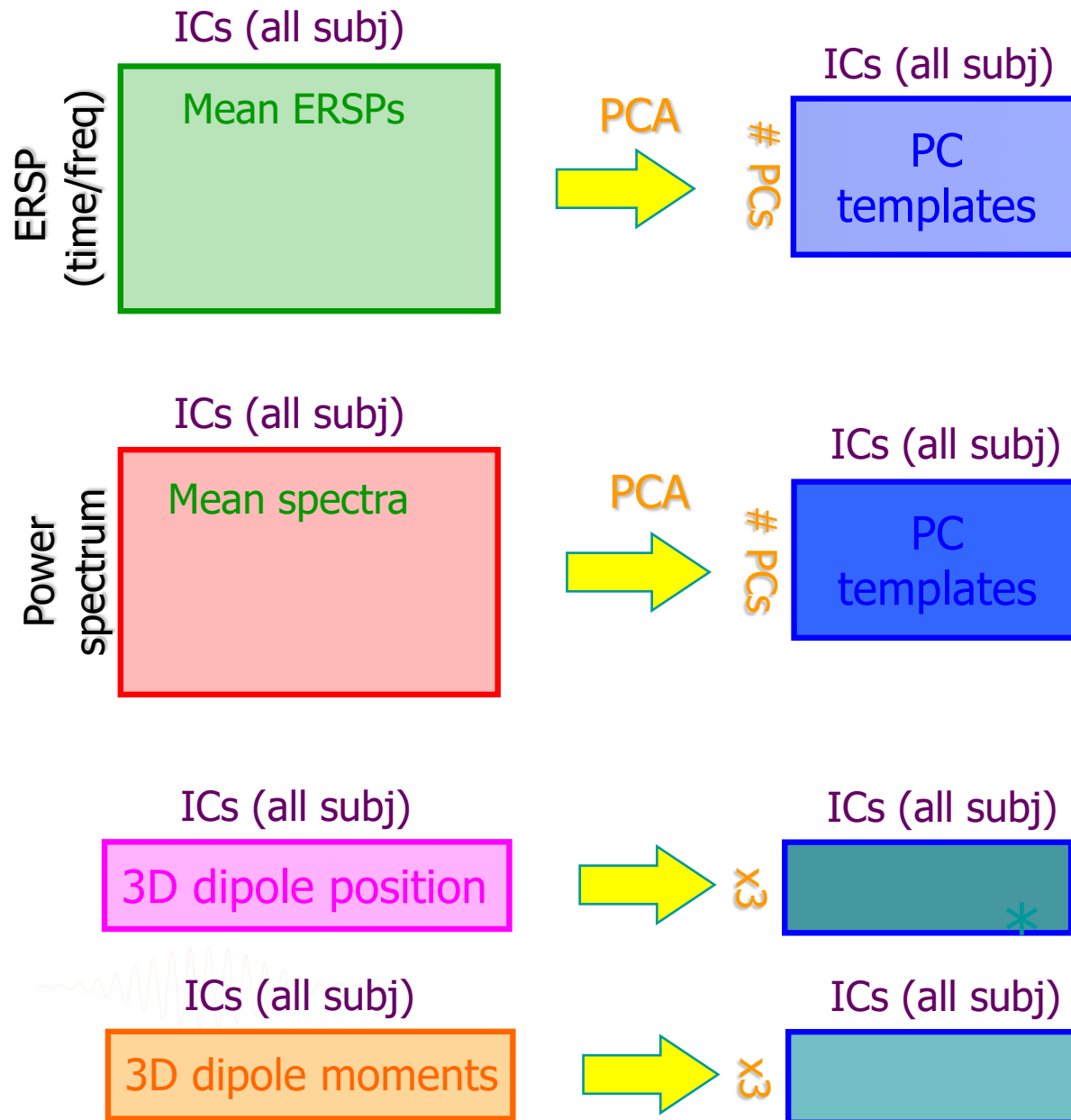
Cluster components



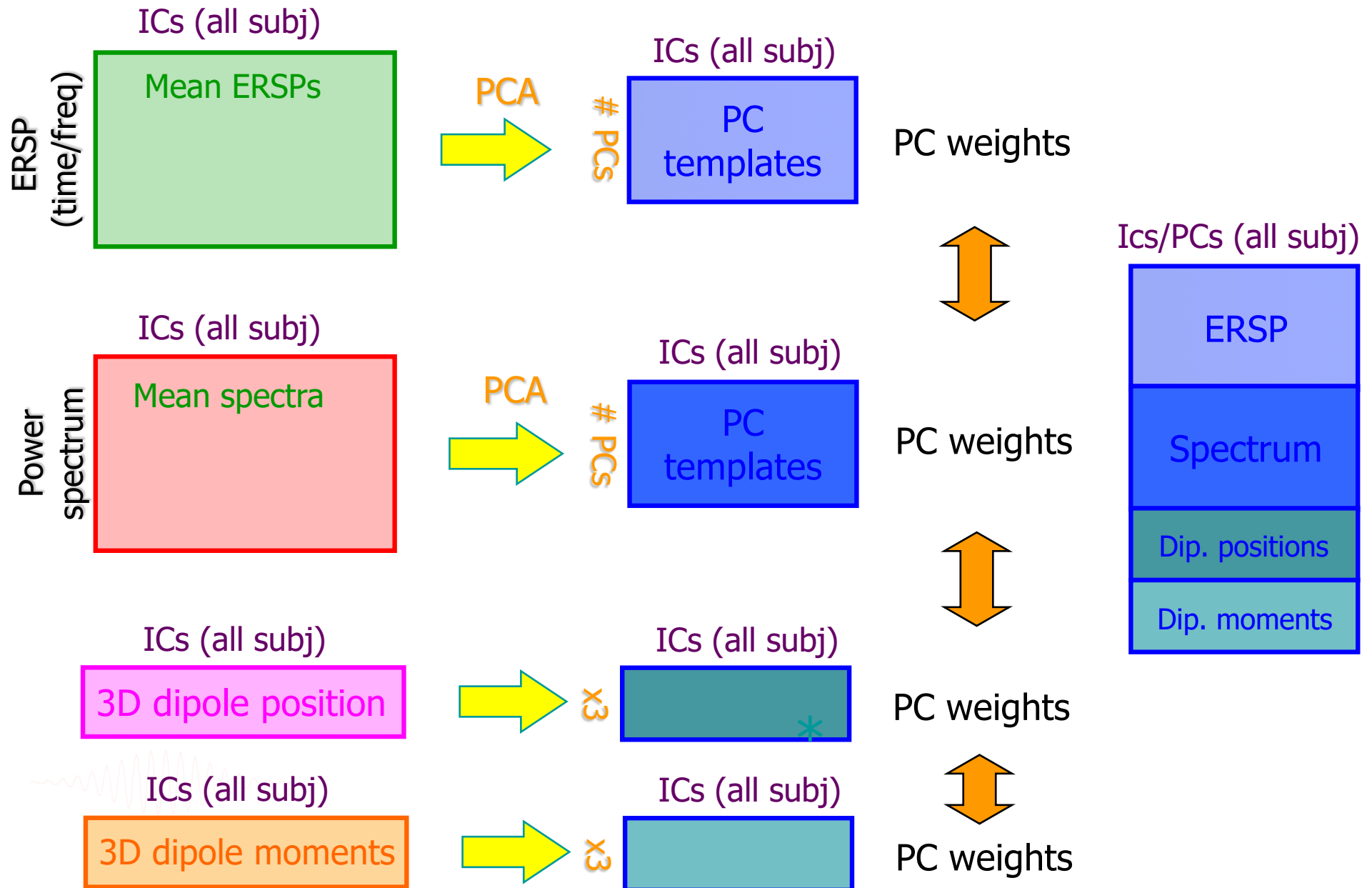
Precluster schematic



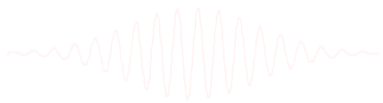
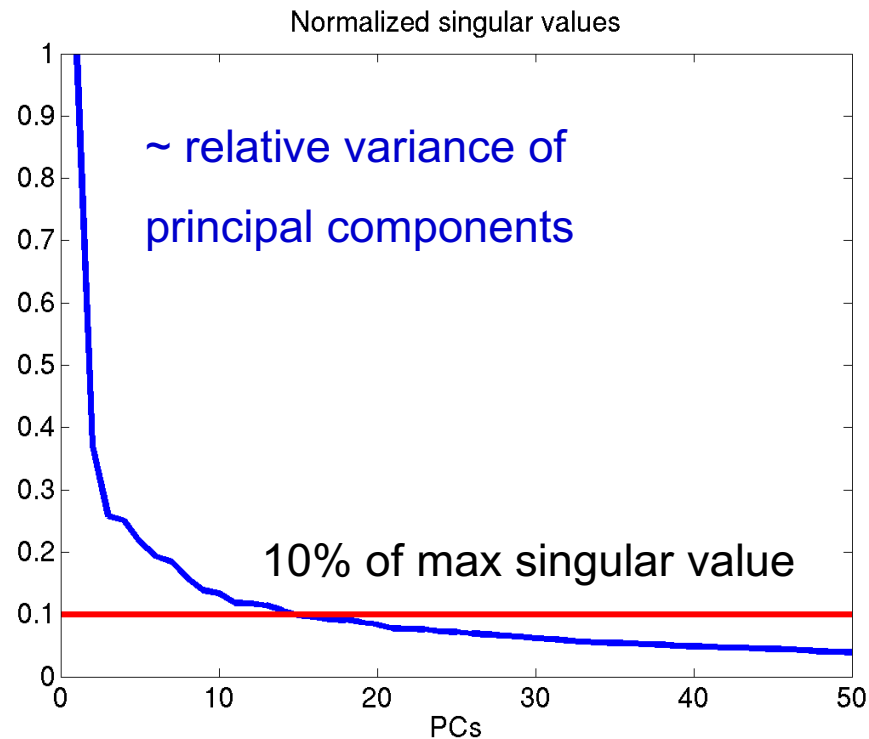
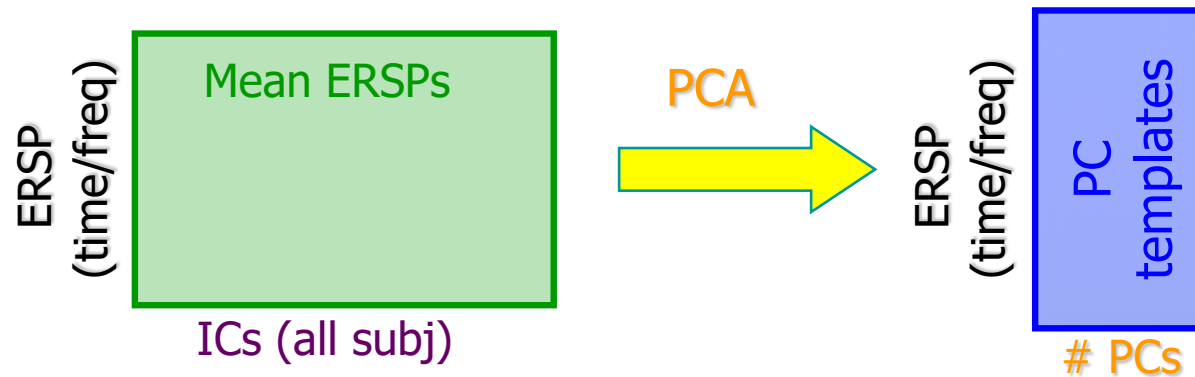
Precluster schematic



Precluster schematic

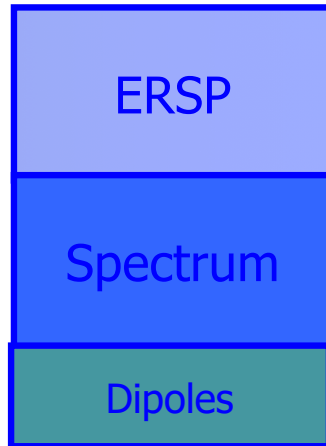


Precluster: Use singular values from PCA

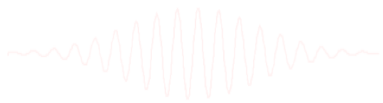
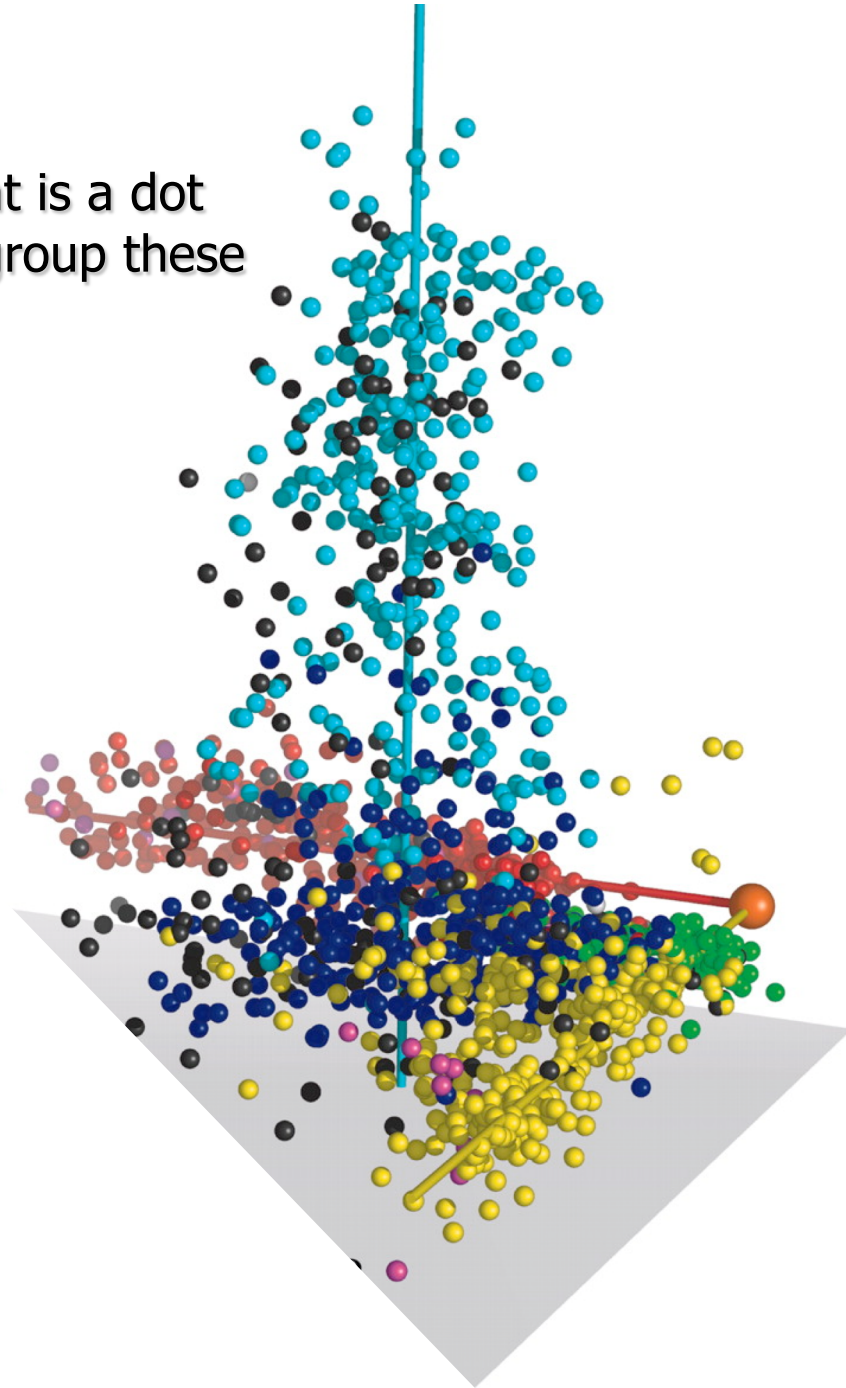


Precluster schematic

ICs (all subj)



Each component is a dot
Clustering will group these
dots



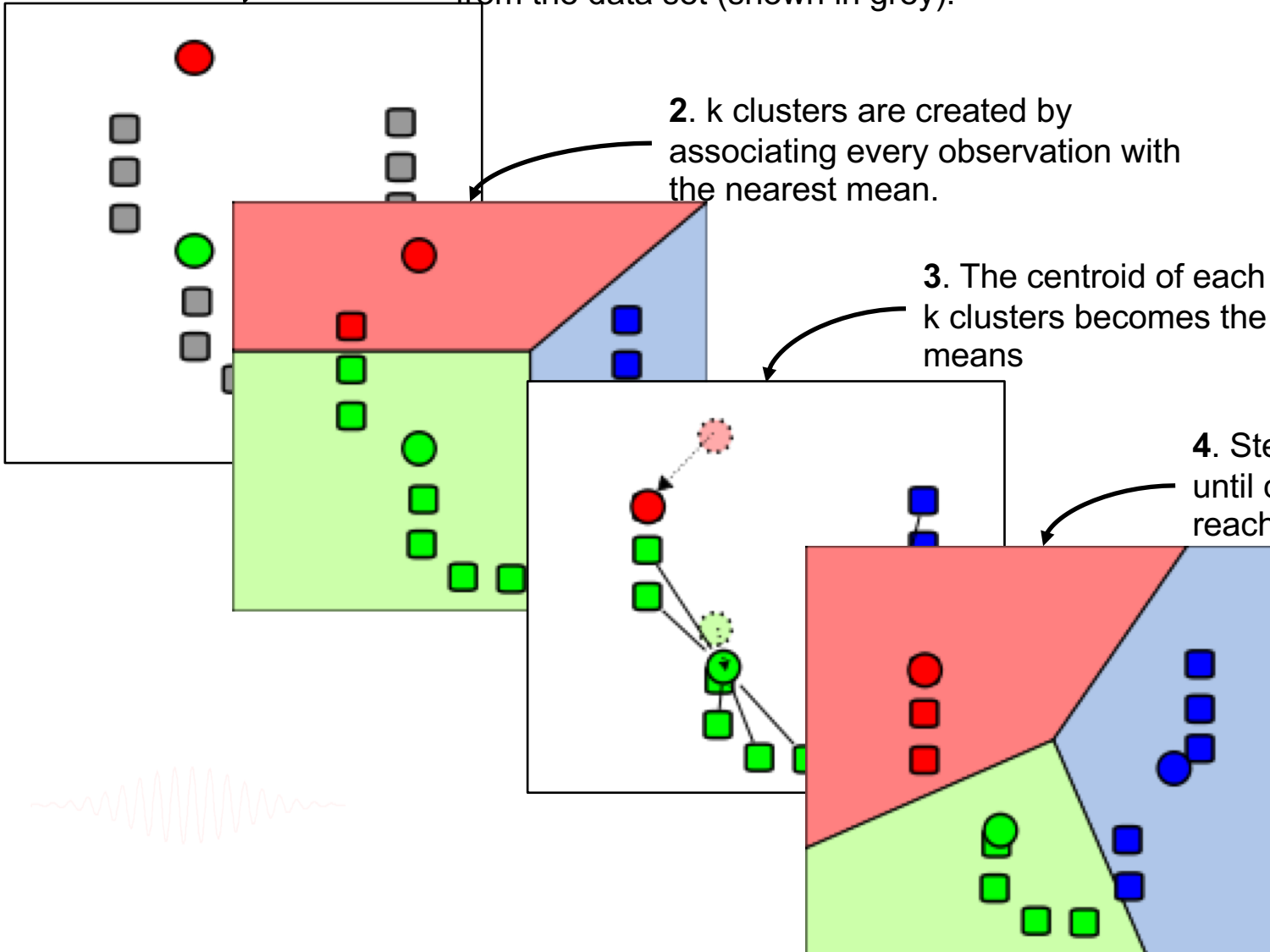
Classical KMean

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

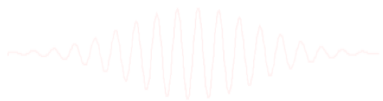
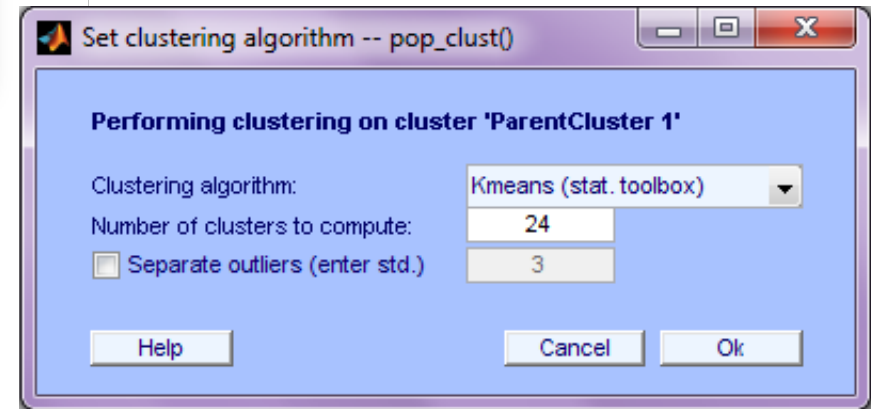
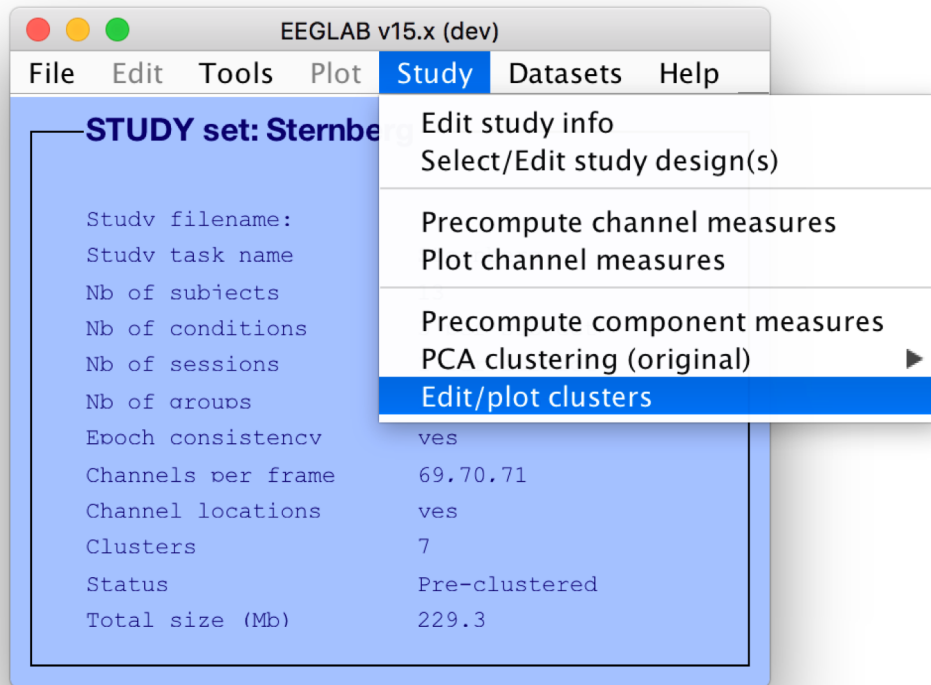
2. k clusters are created by associating every observation with the nearest mean.

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

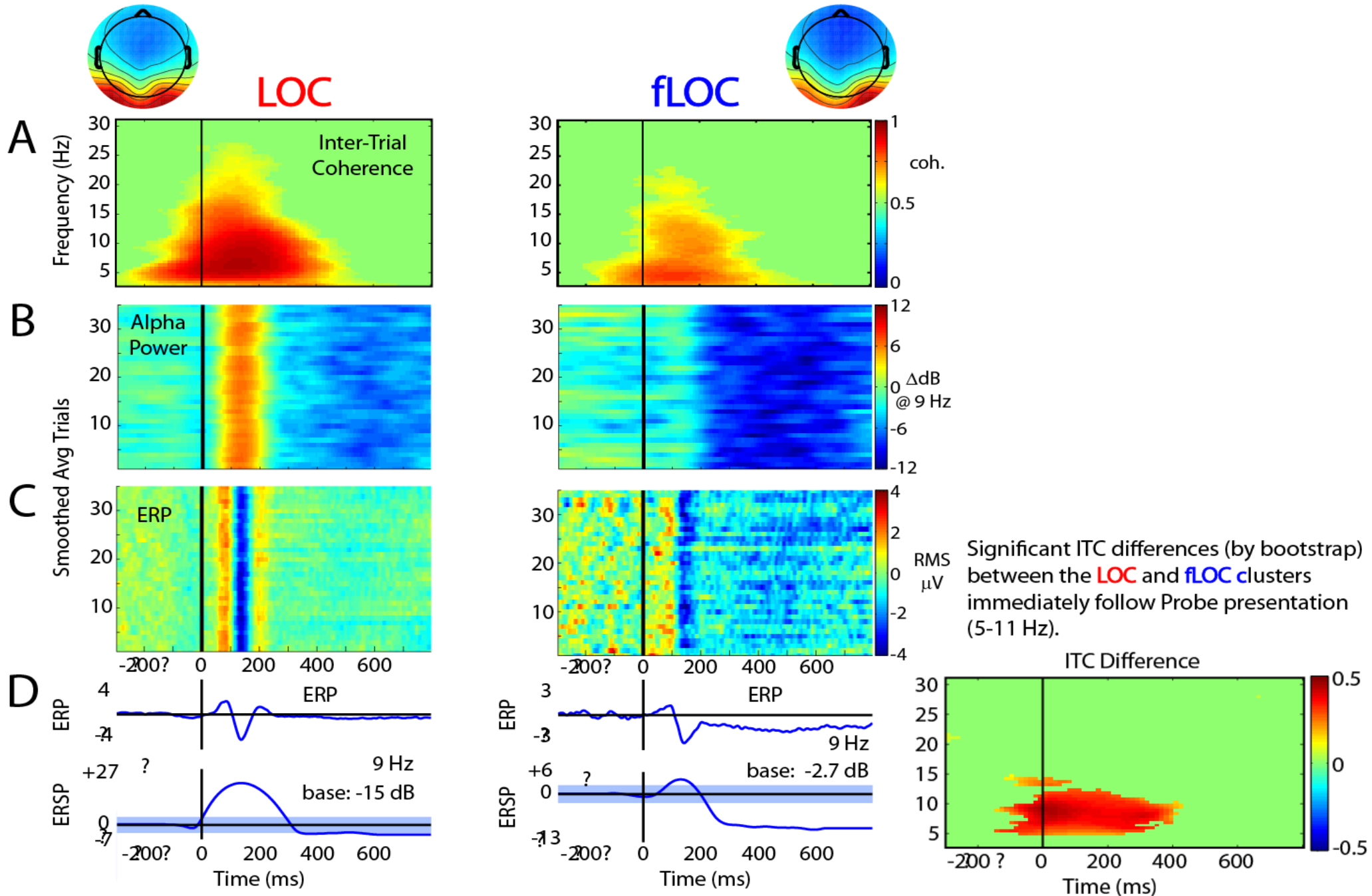
4. Steps 2 and 3 are repeated until convergence has been reached.



Cluster components



Subject differences?

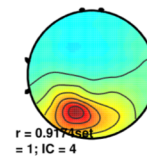
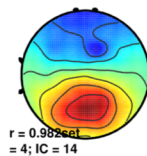
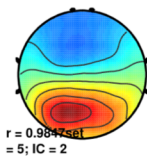
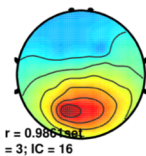
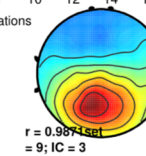
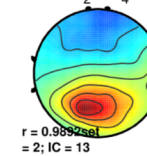
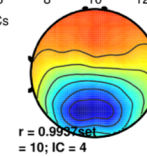
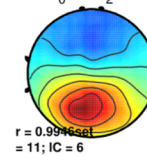
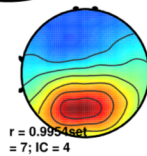
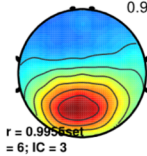
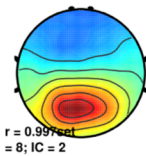
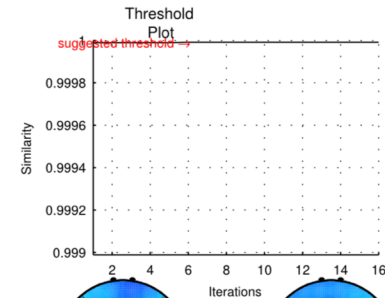
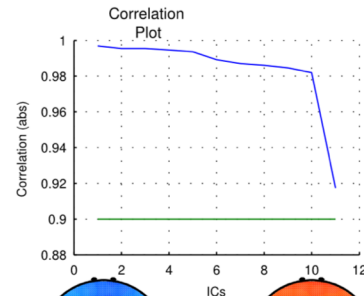
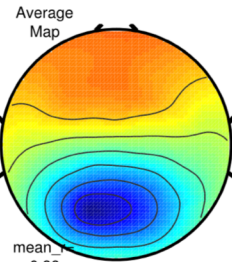


Results (Cluster 1 within subject)

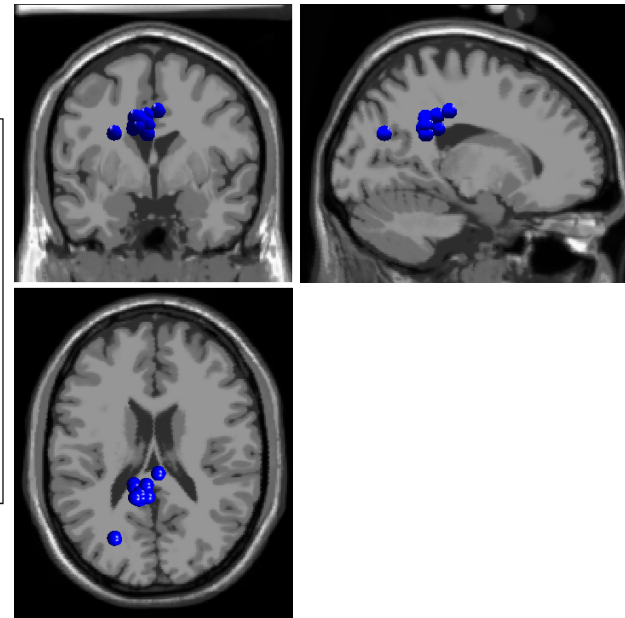
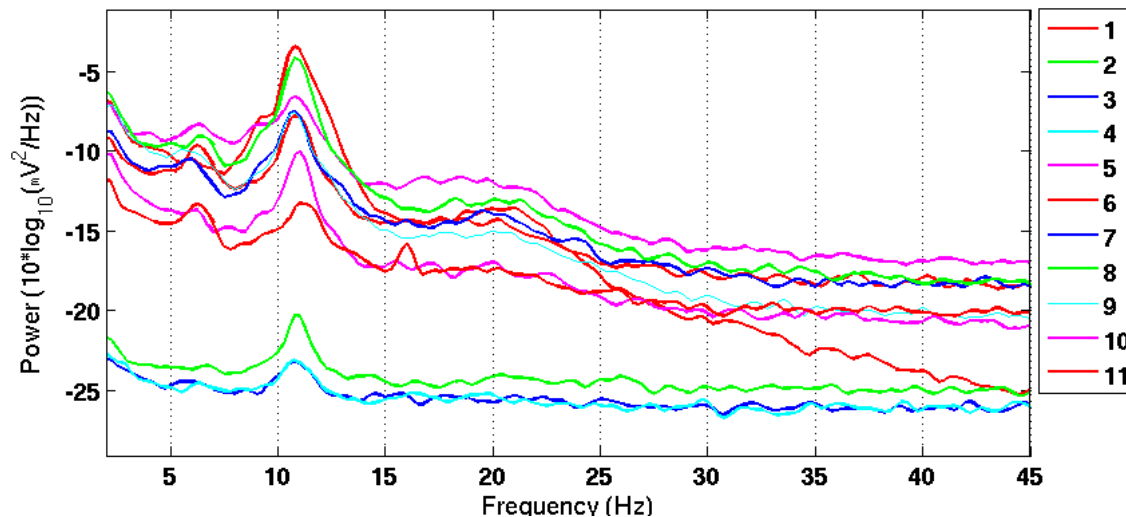
100 % Sessions contribute

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



Cls 3 Spectrum

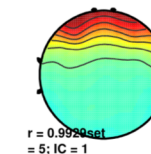
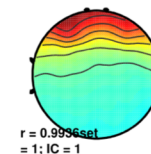
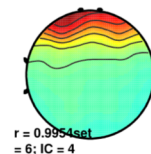
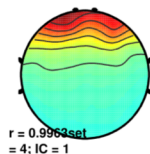
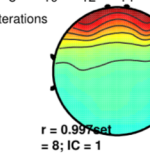
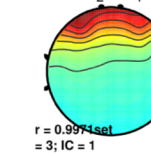
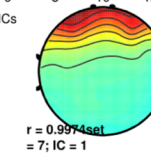
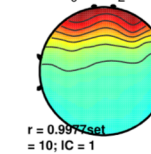
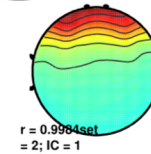
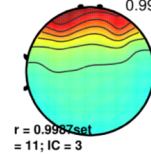
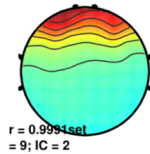
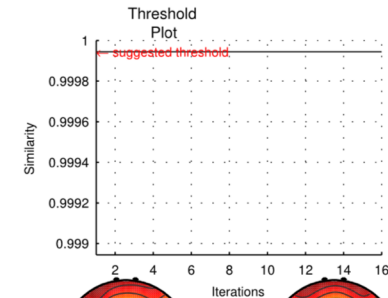
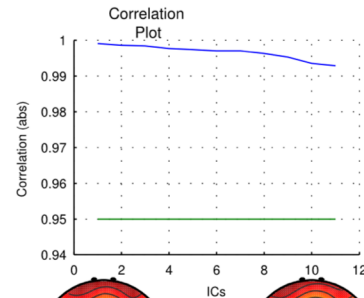
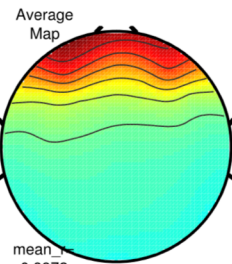


Results (Cluster 2 within subject)

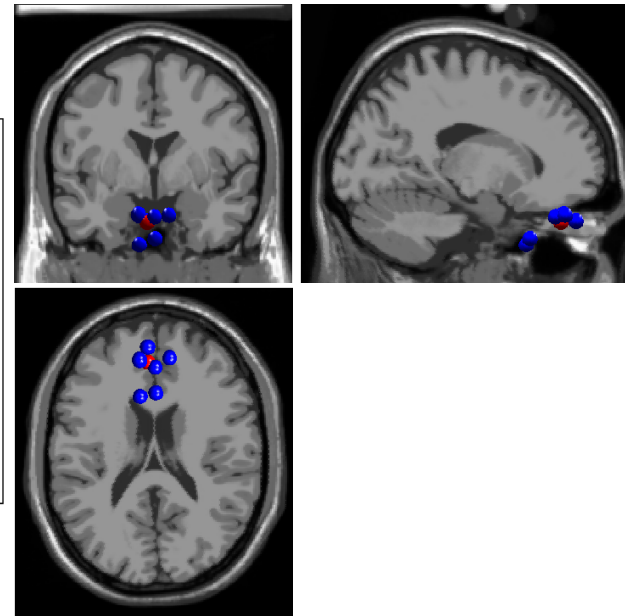
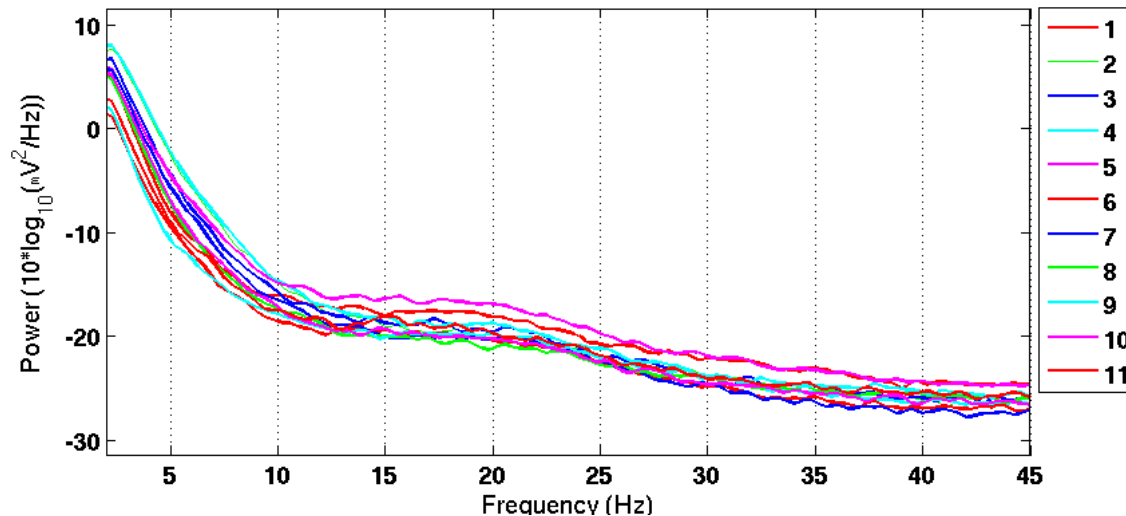
100 % Sessions contribute

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



Cls 4 Spectrum



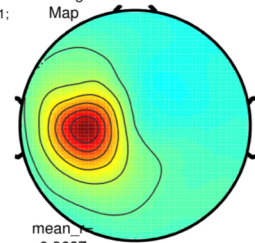
Results (Cluster 8 within subject)

100 % Sessions contribute

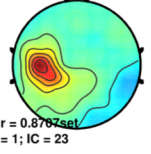
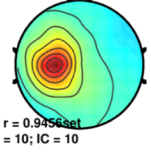
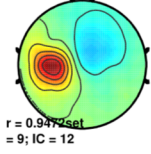
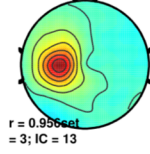
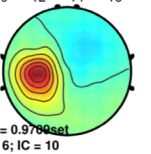
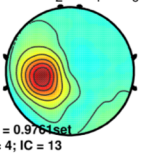
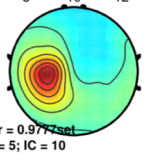
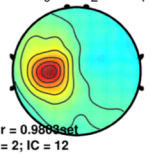
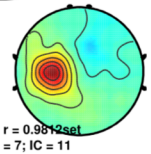
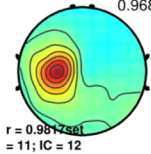
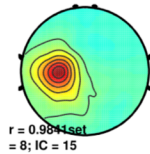
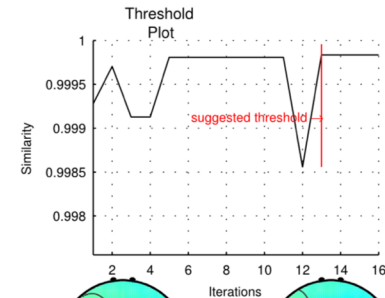
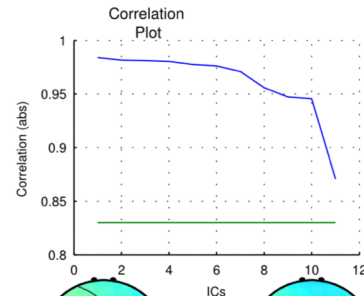
INFO:
Template: CB Session 7 PREPROC:STEP 2; Set 7; IC 11;
Number of datasets: 11
Correlation threshold: 0.83 (green line)
Max ICs from each dataset: 1
Cluster: 11 ICs from 11 sets
All datasets contribute.

Similarity = 0.9998

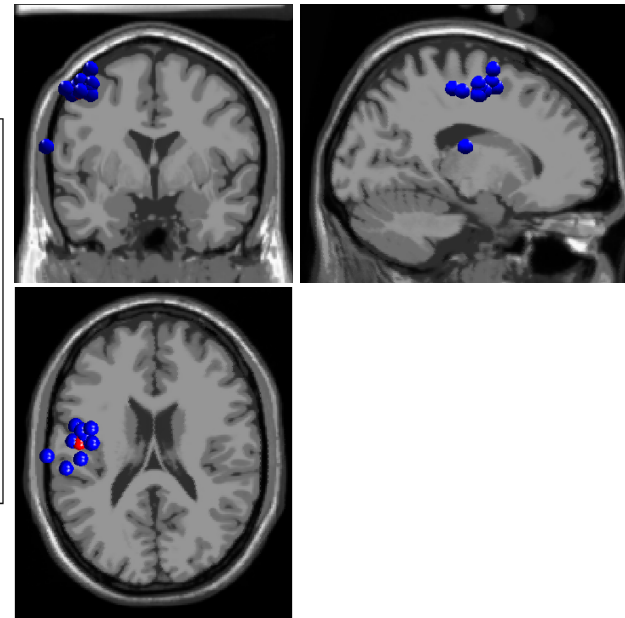
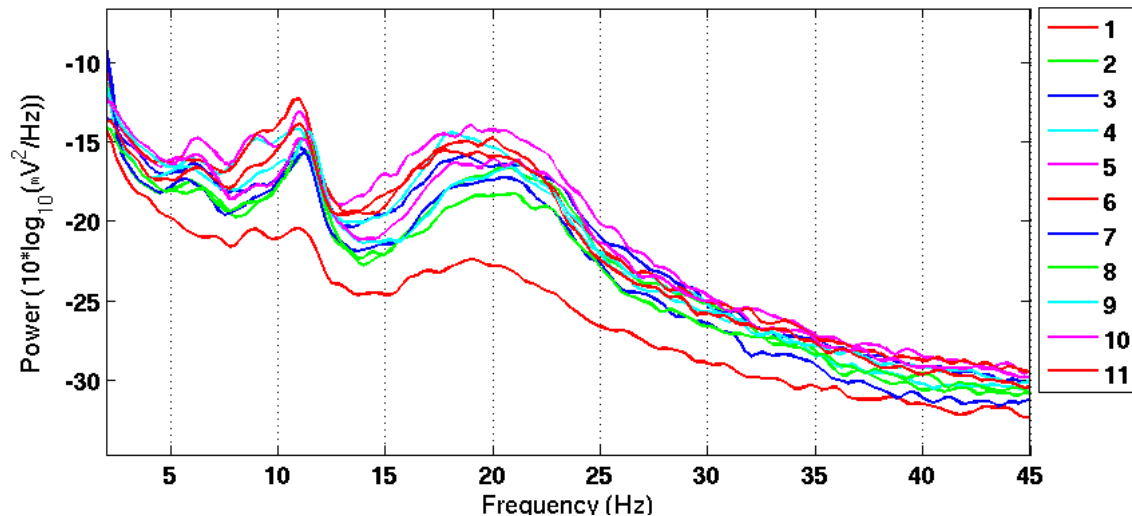
Average
Map



mean r
0.9687

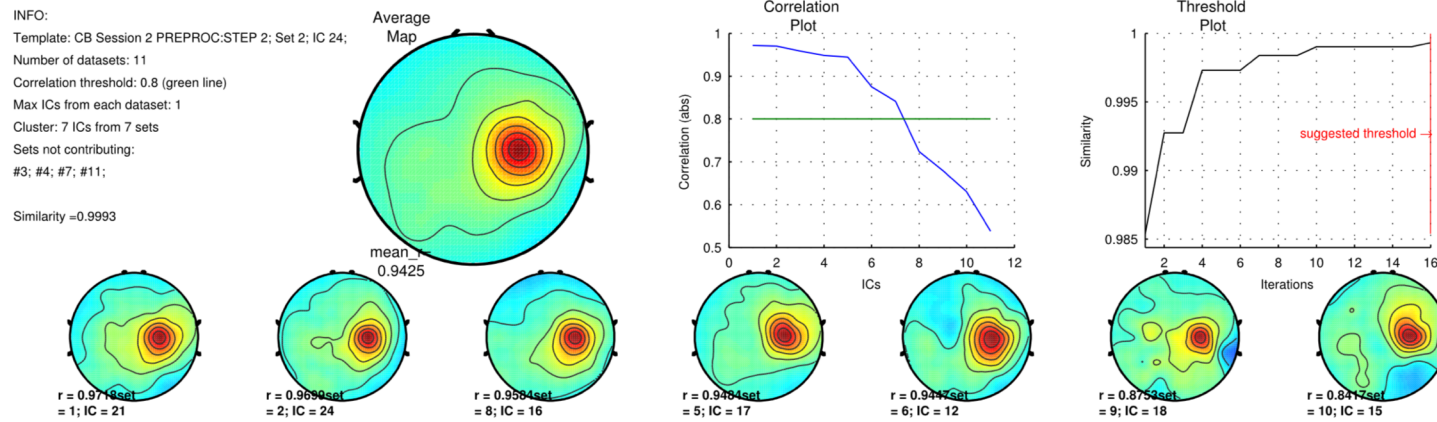


Cls 8 Spectrum

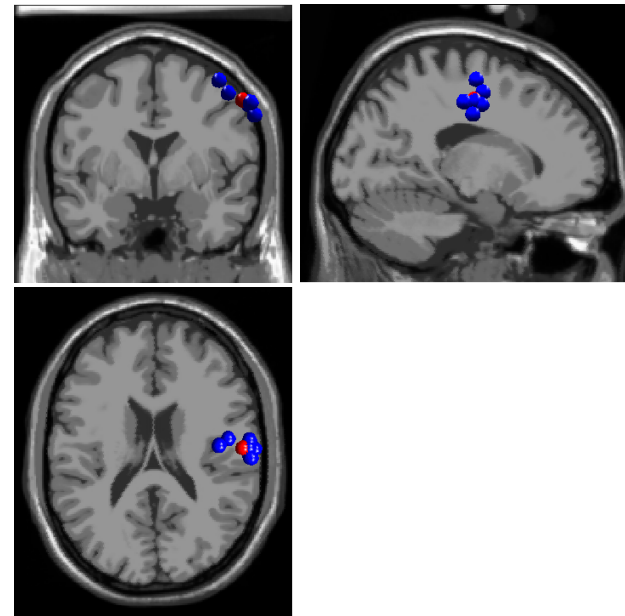
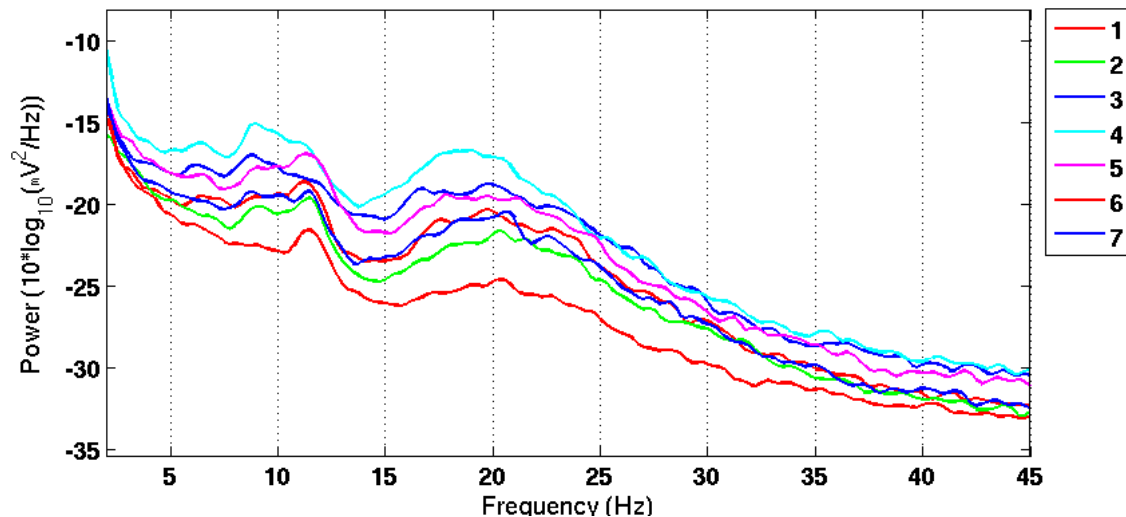


Results (Cluster 13 within subject)

63.64% Sessions contribute



Cls 13 Spectrum



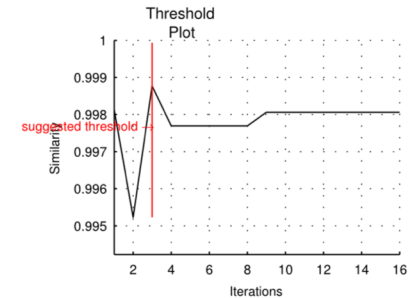
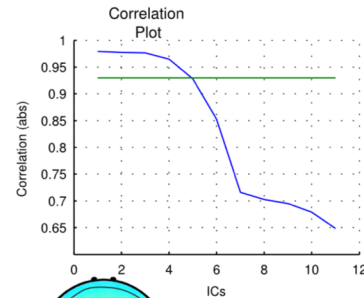
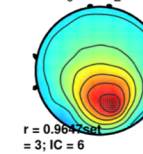
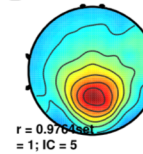
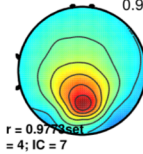
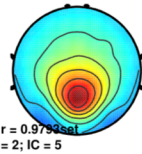
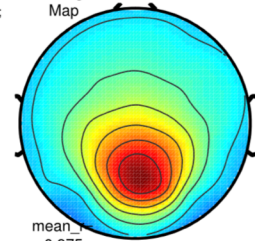
Results (Cluster 14 within subject)

36.36% Sessions contribute

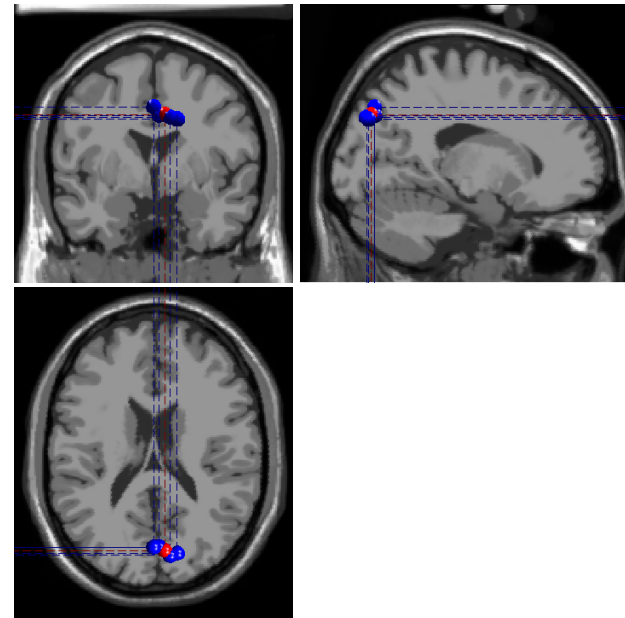
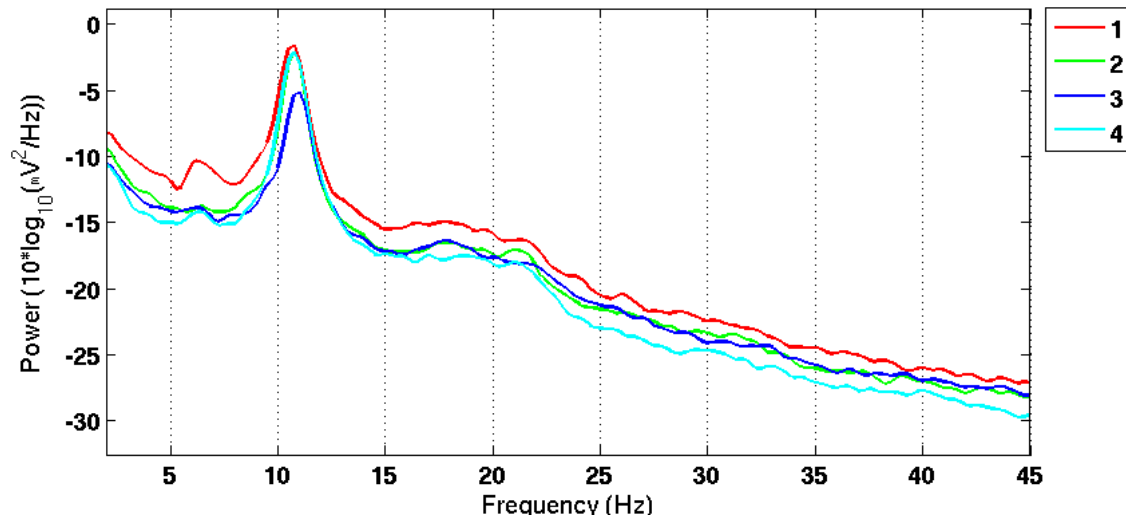
INFO:
Template: CB Session 1 PREPROC:STEP 2; Set 1; IC 5;
Number of datasets: 11
Correlation threshold: 0.93 (green line)
Max ICs from each dataset: 1
Cluster: 4 ICs from 4 sets
Sets not contributing:
#5; #6; #7; #8; #9; #10#11;

Similarity = 0.9988

Average Map

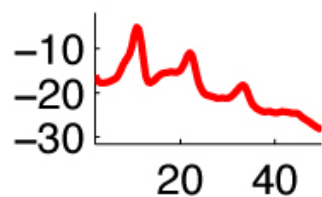
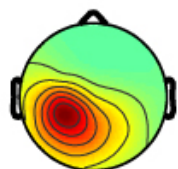


Cls 14 Spectrum

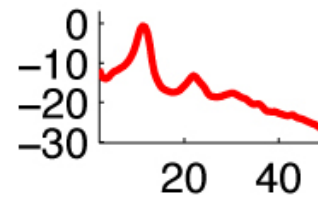
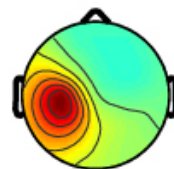


Left μ cluster (across subjects)

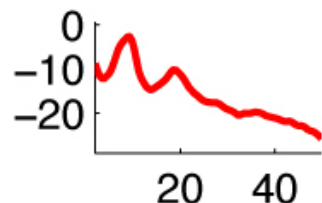
S2 IC47



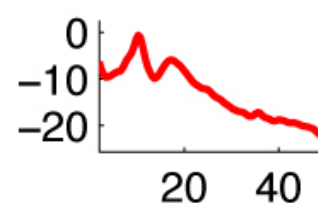
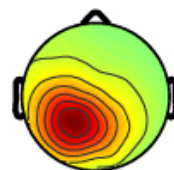
S3 IC47



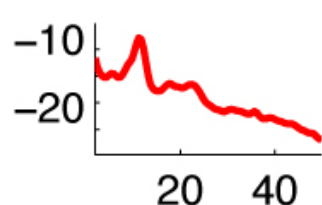
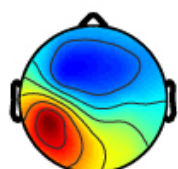
S4 IC37



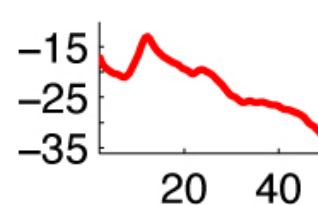
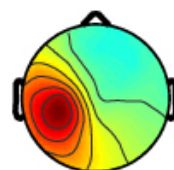
S5 IC48



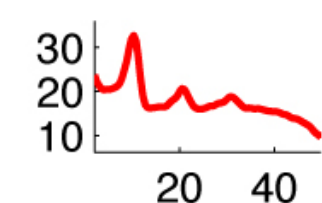
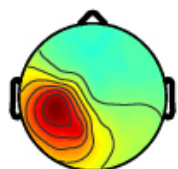
S6 IC46



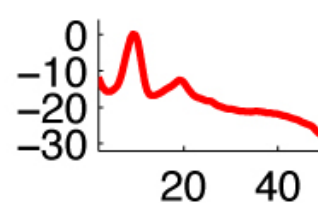
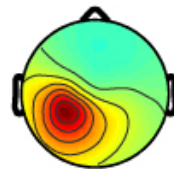
S7 IC35



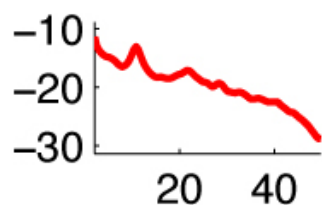
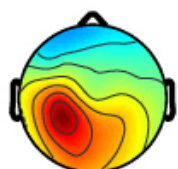
S9 IC7



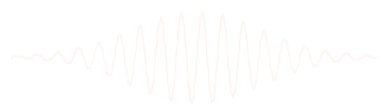
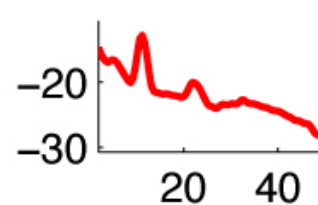
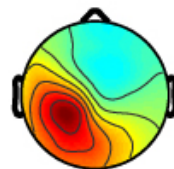
S11 IC45



S12 IC45

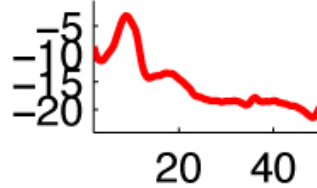
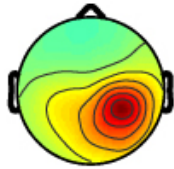


S14 IC45

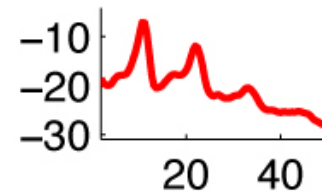
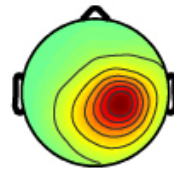


Right μ cluster

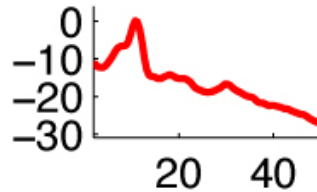
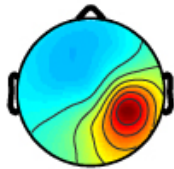
S1 IC51



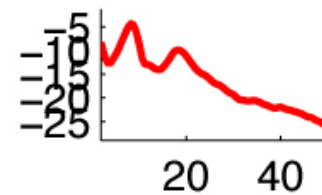
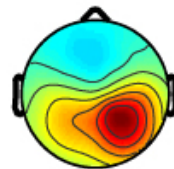
S2 IC41



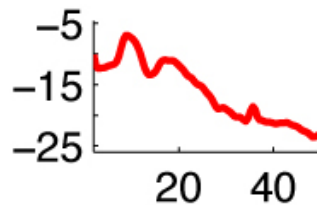
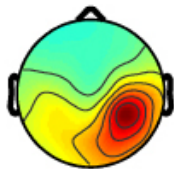
S3 IC41



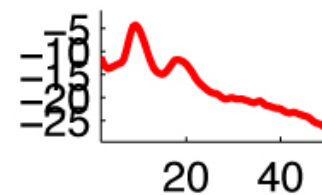
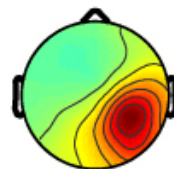
S4 IC50



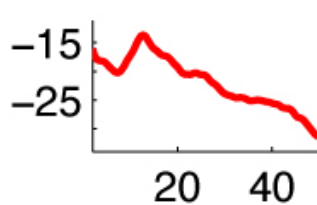
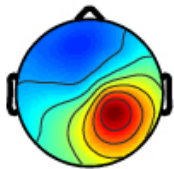
S5 IC51



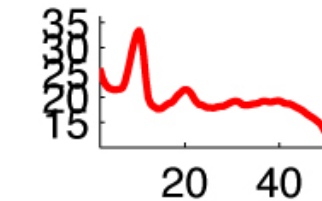
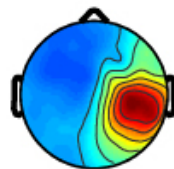
S6 IC6⁰



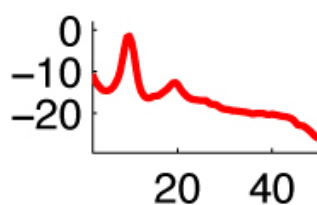
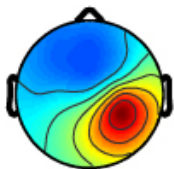
S7 IC48



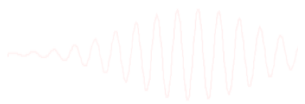
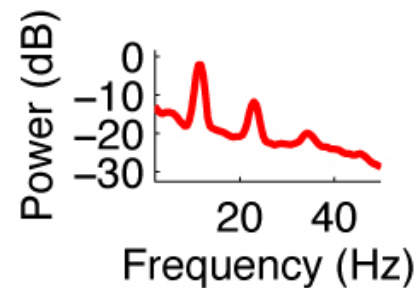
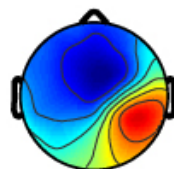
S9 IC39



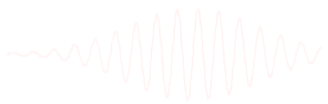
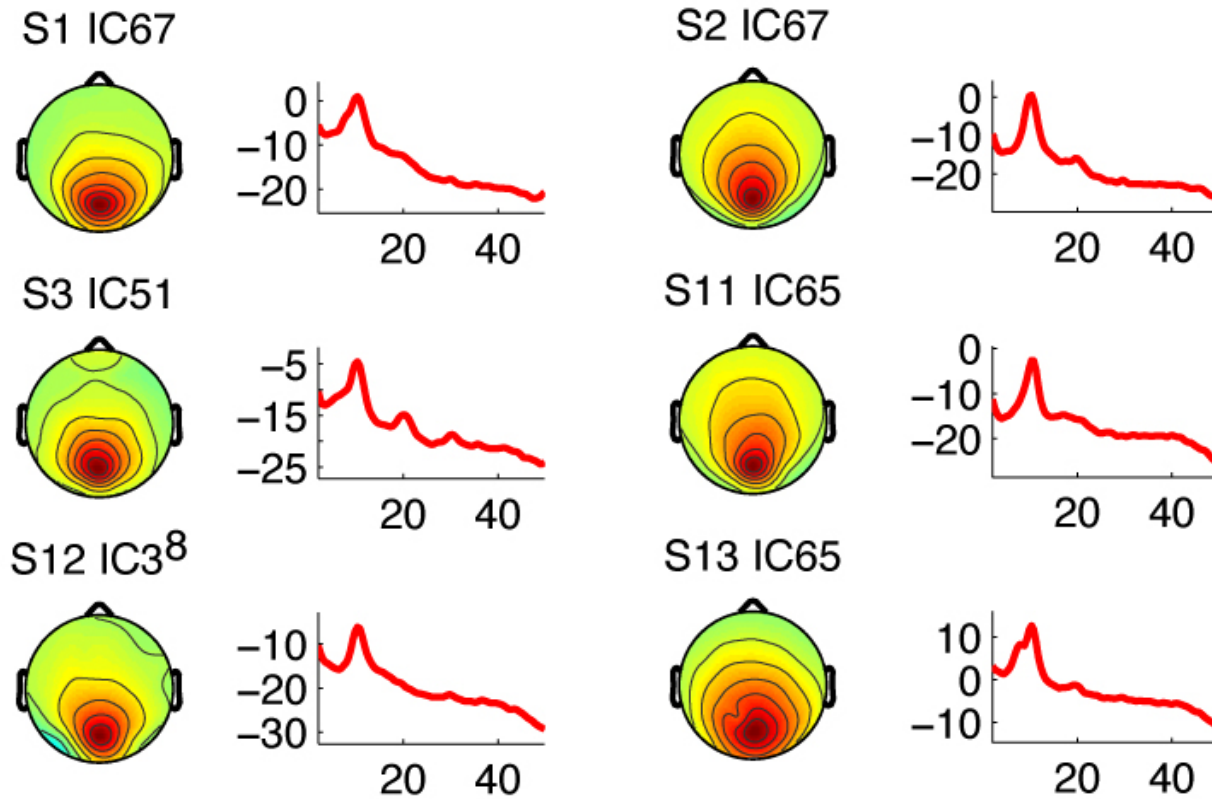
S11 IC49



S14 IC49

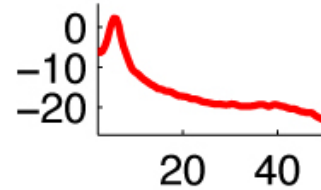
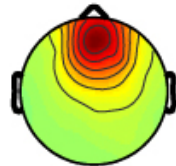


Occipital α cluster

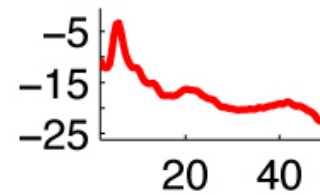
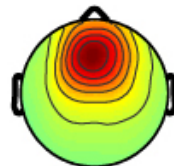


Frontal Midline θ cluster

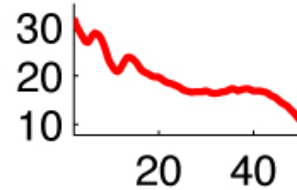
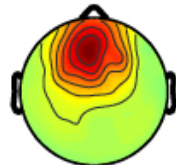
S1 IC63



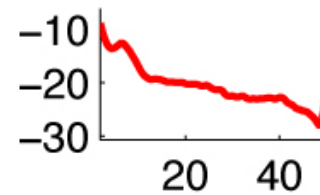
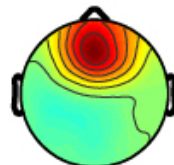
S2 IC18



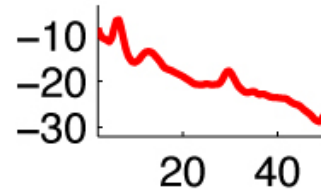
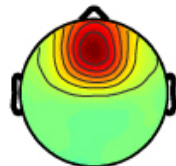
S9 IC16



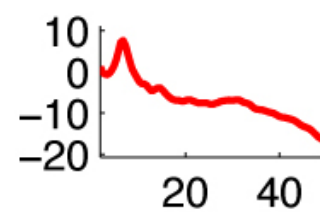
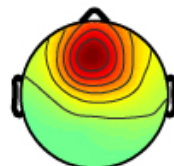
S11 IC16



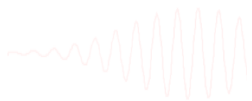
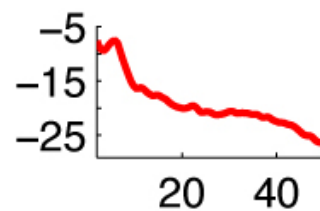
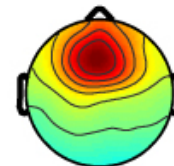
S12 IC15



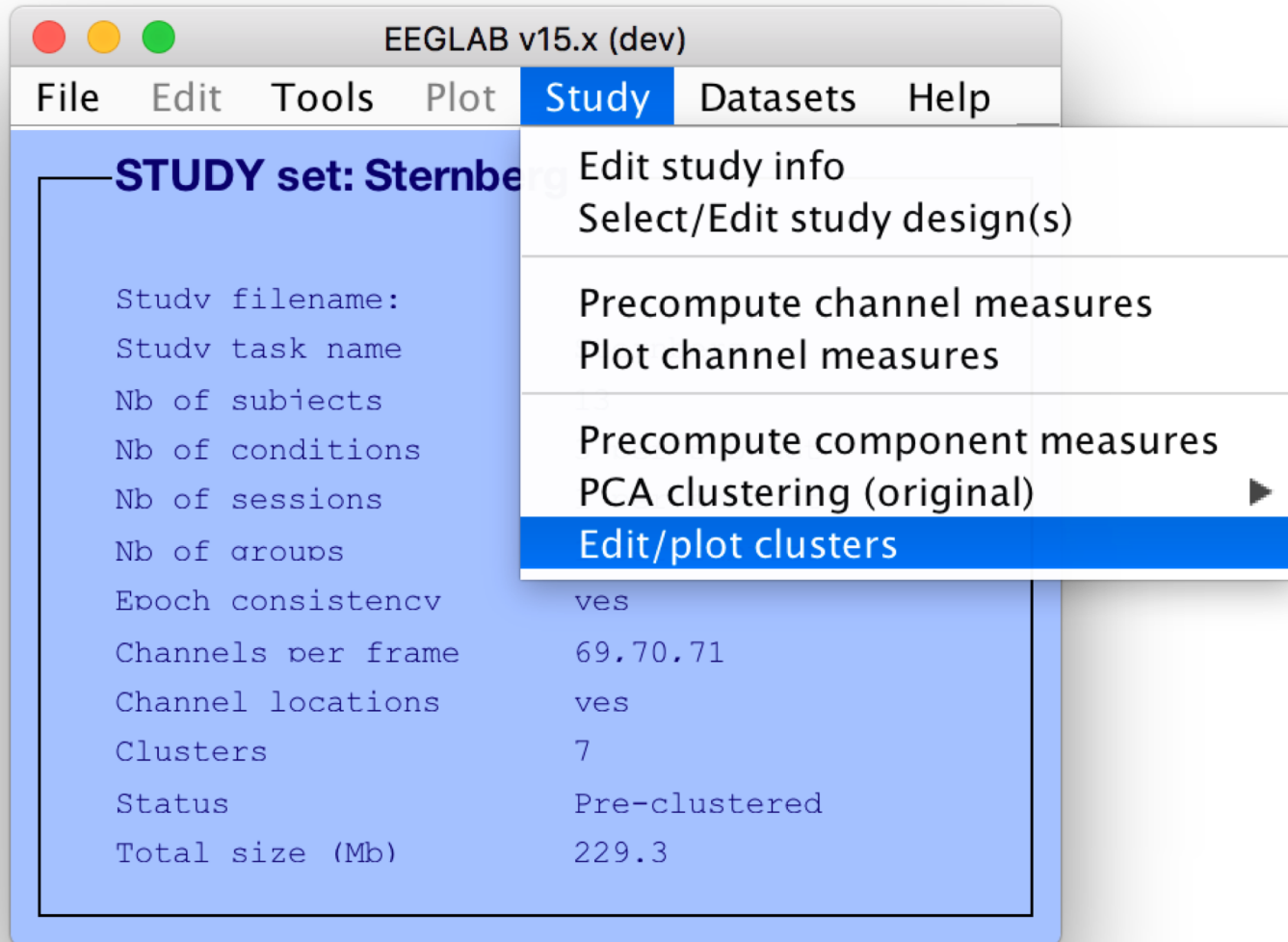
S13 IC15



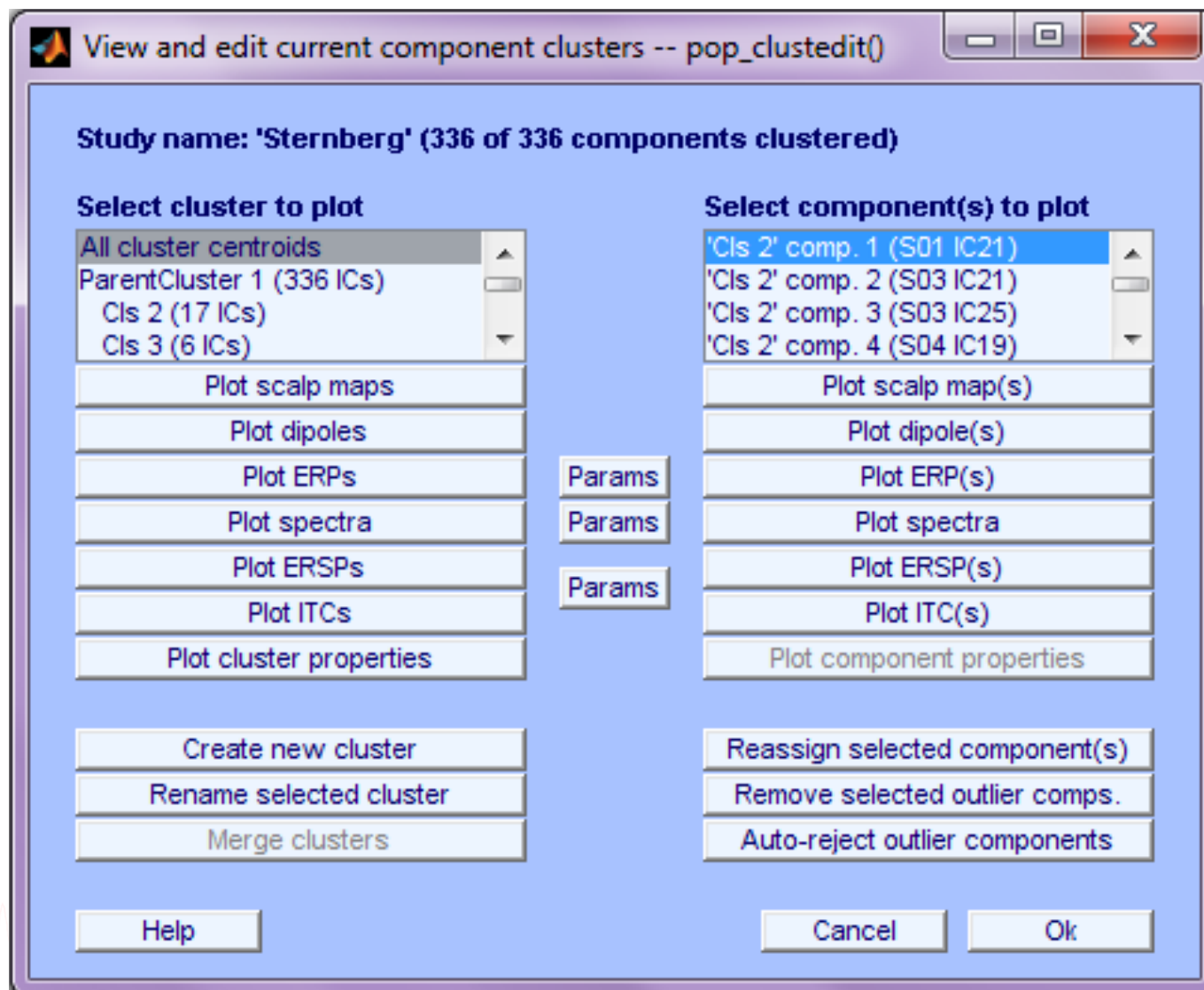
S14 IC16



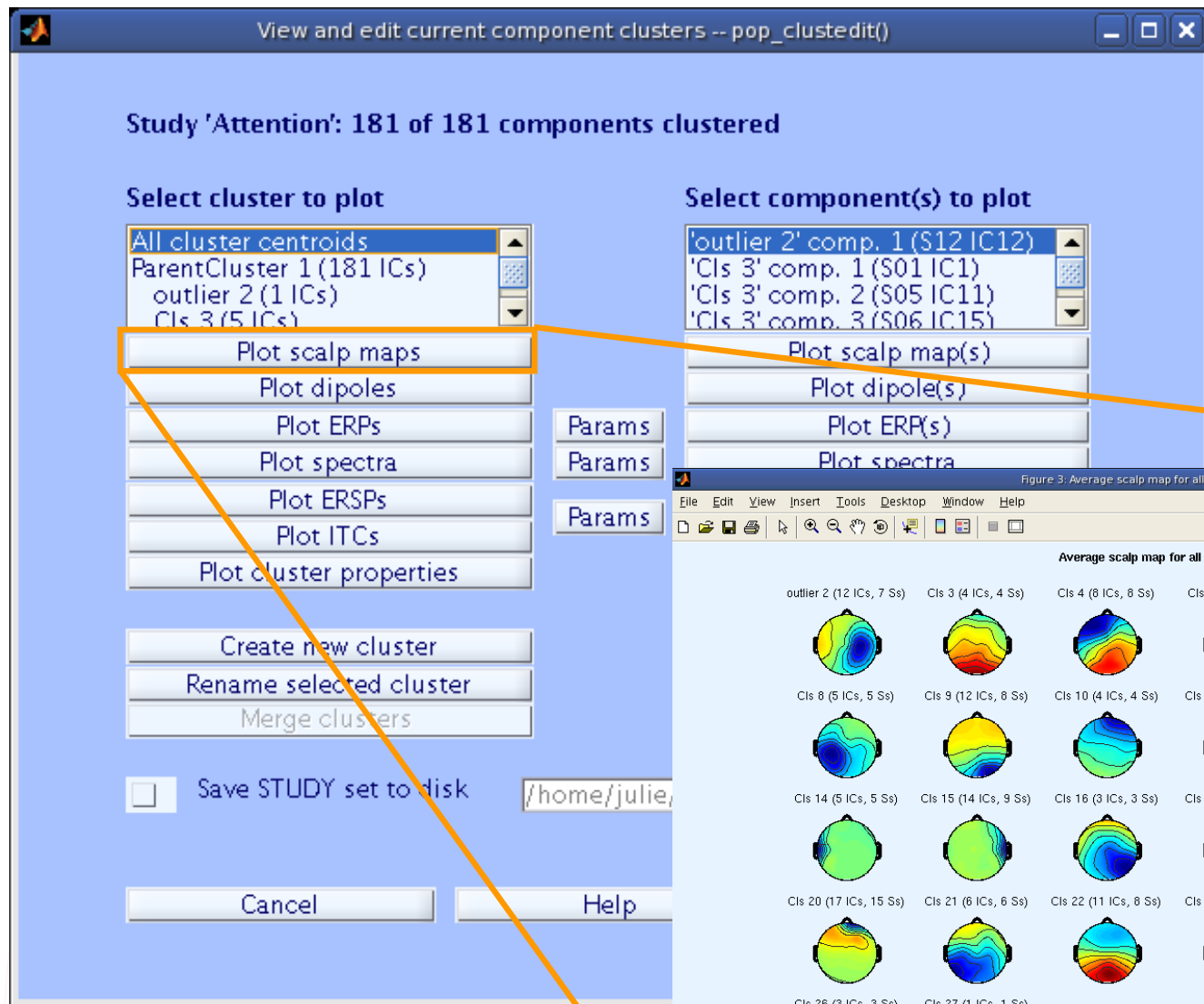
View and edit clusters



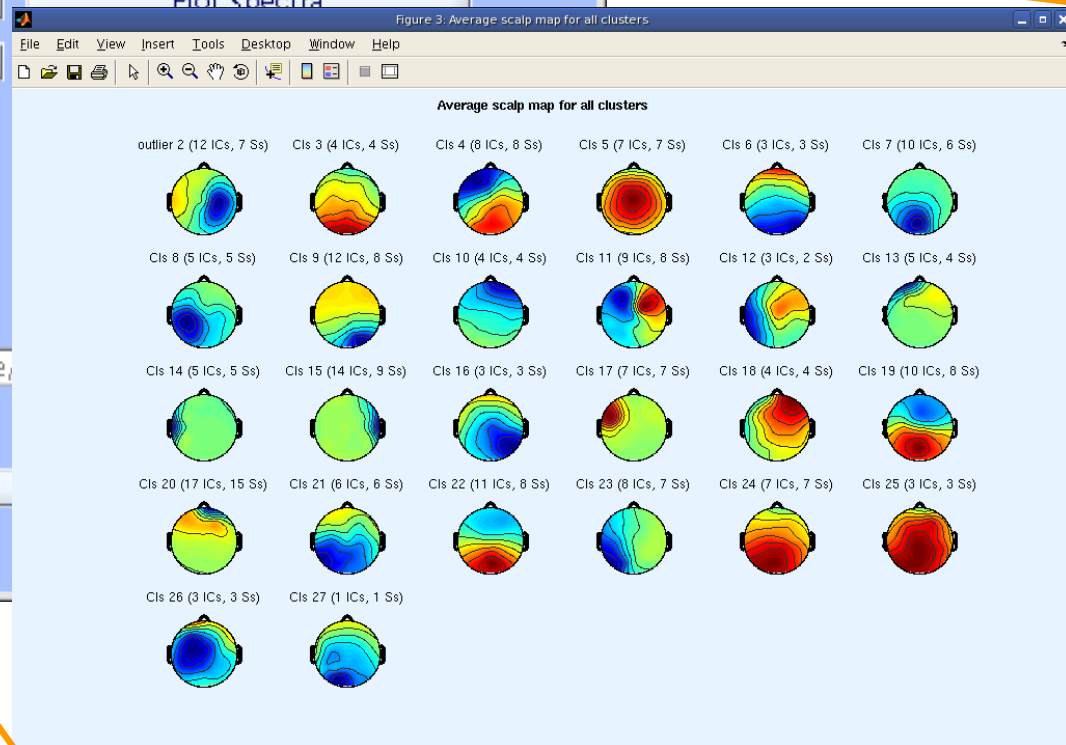
Plot/edit clusters



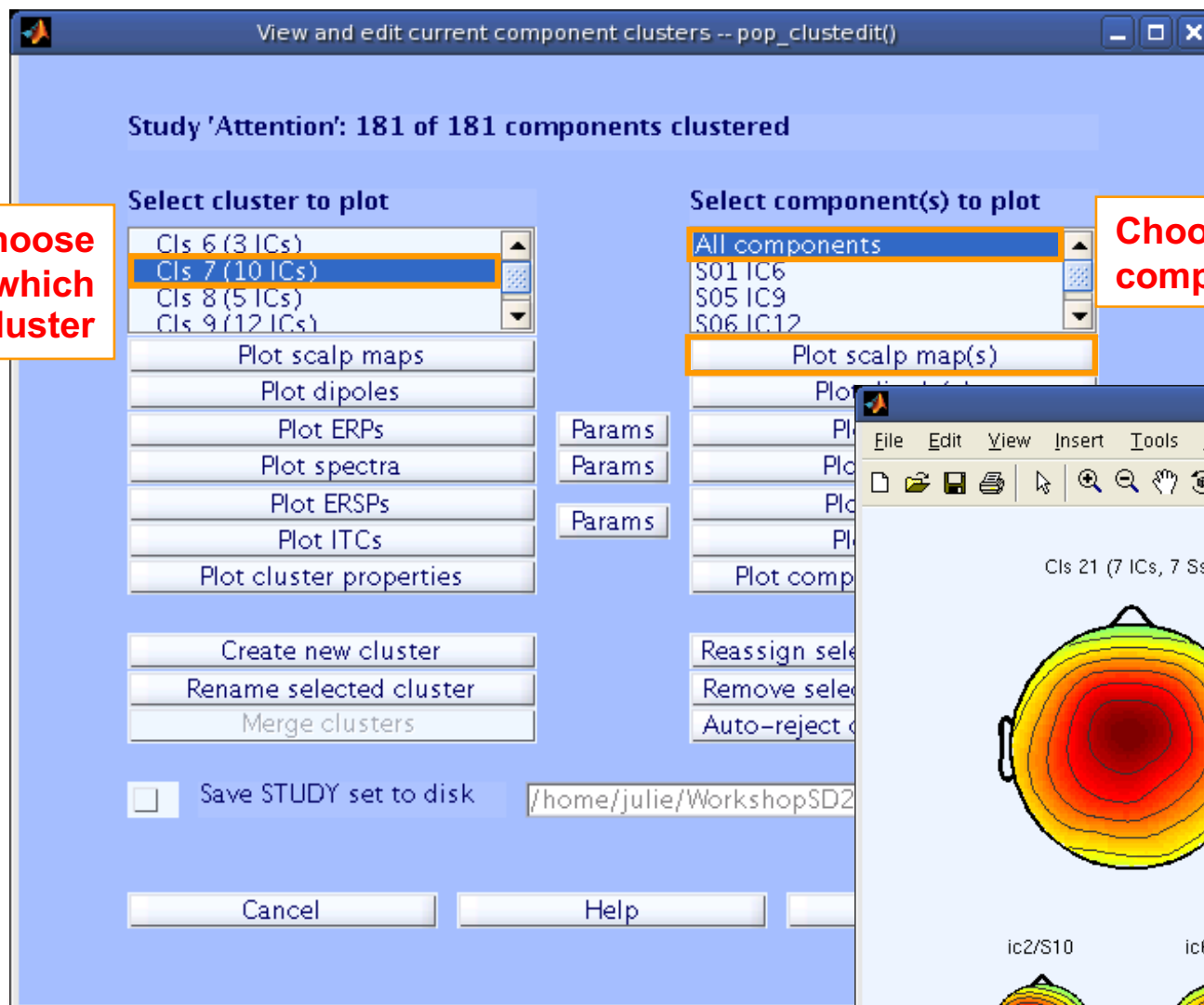
Plot cluster data



Plot mean scalp maps for easy reference

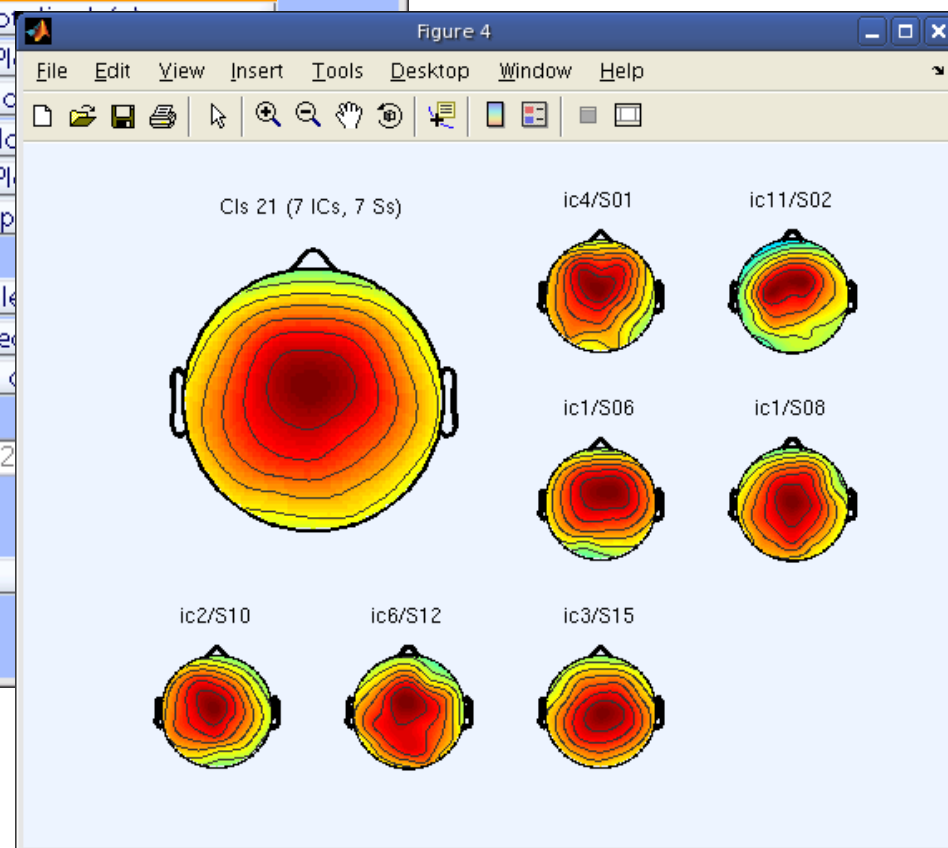


Plot cluster data

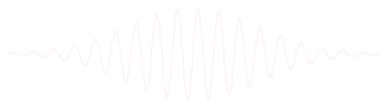
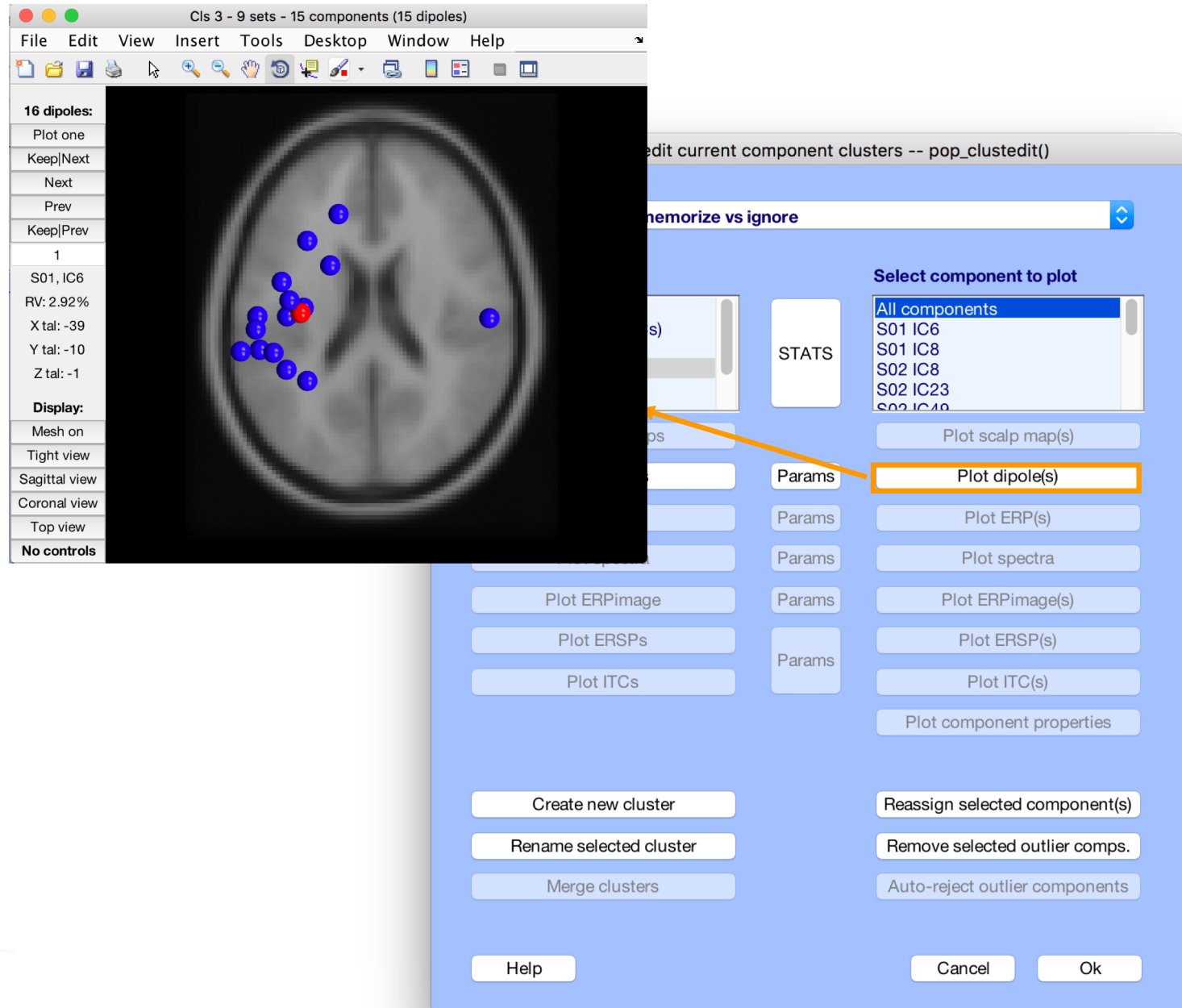


Choose which cluster

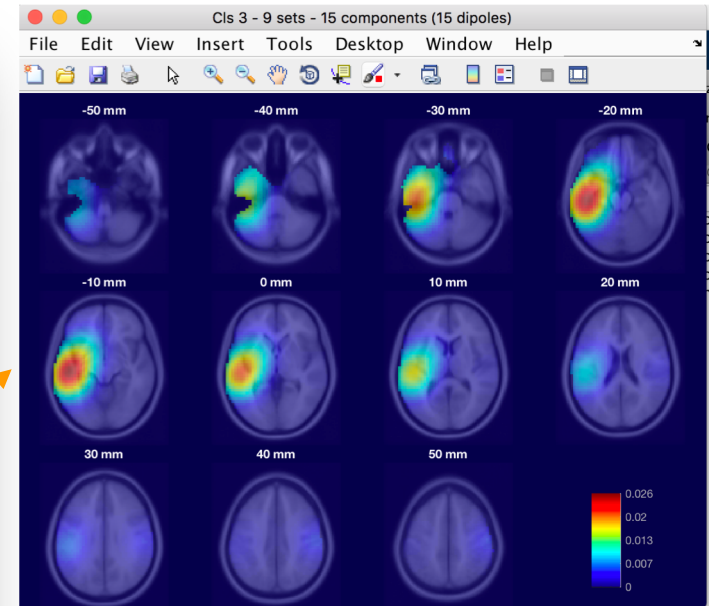
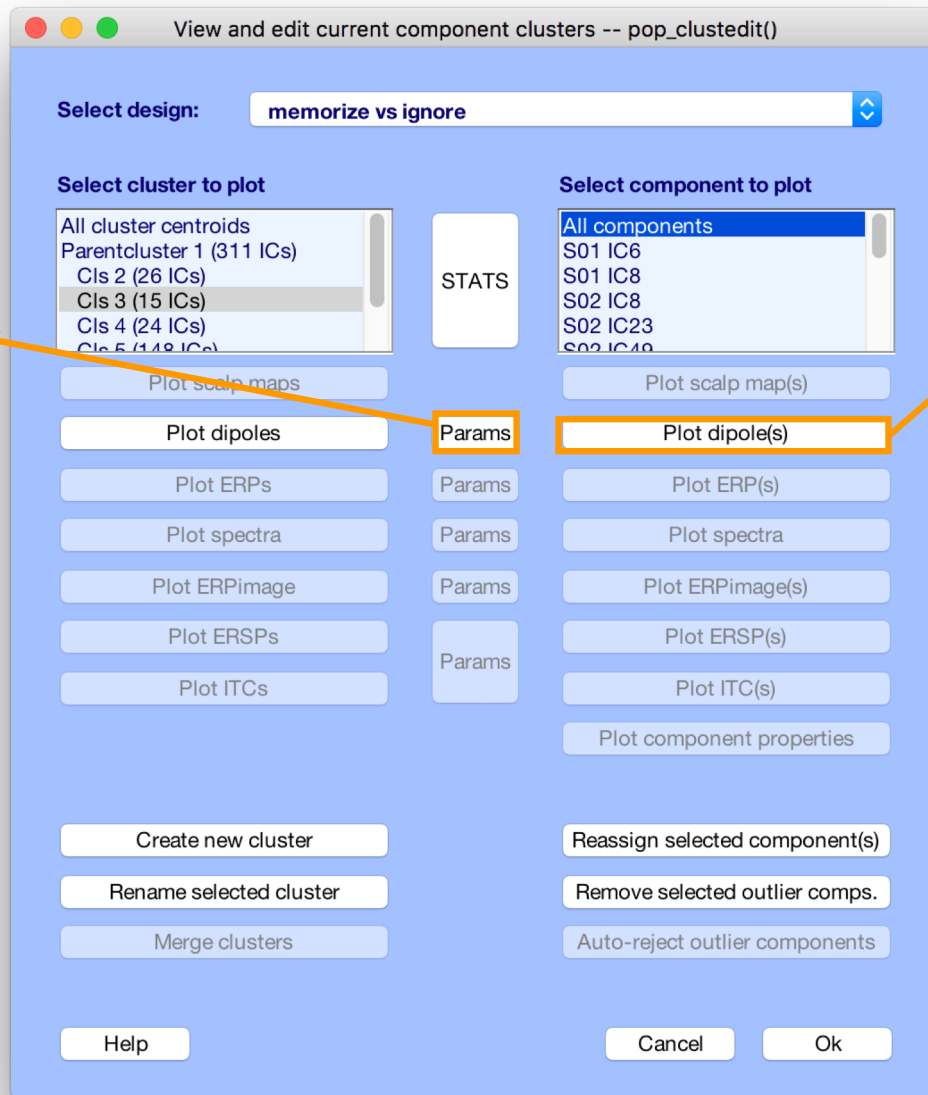
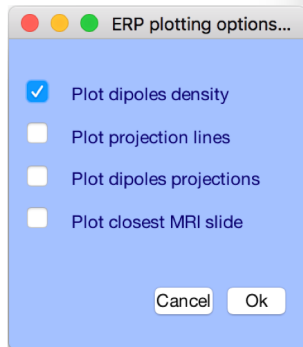
Choose which components



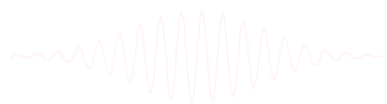
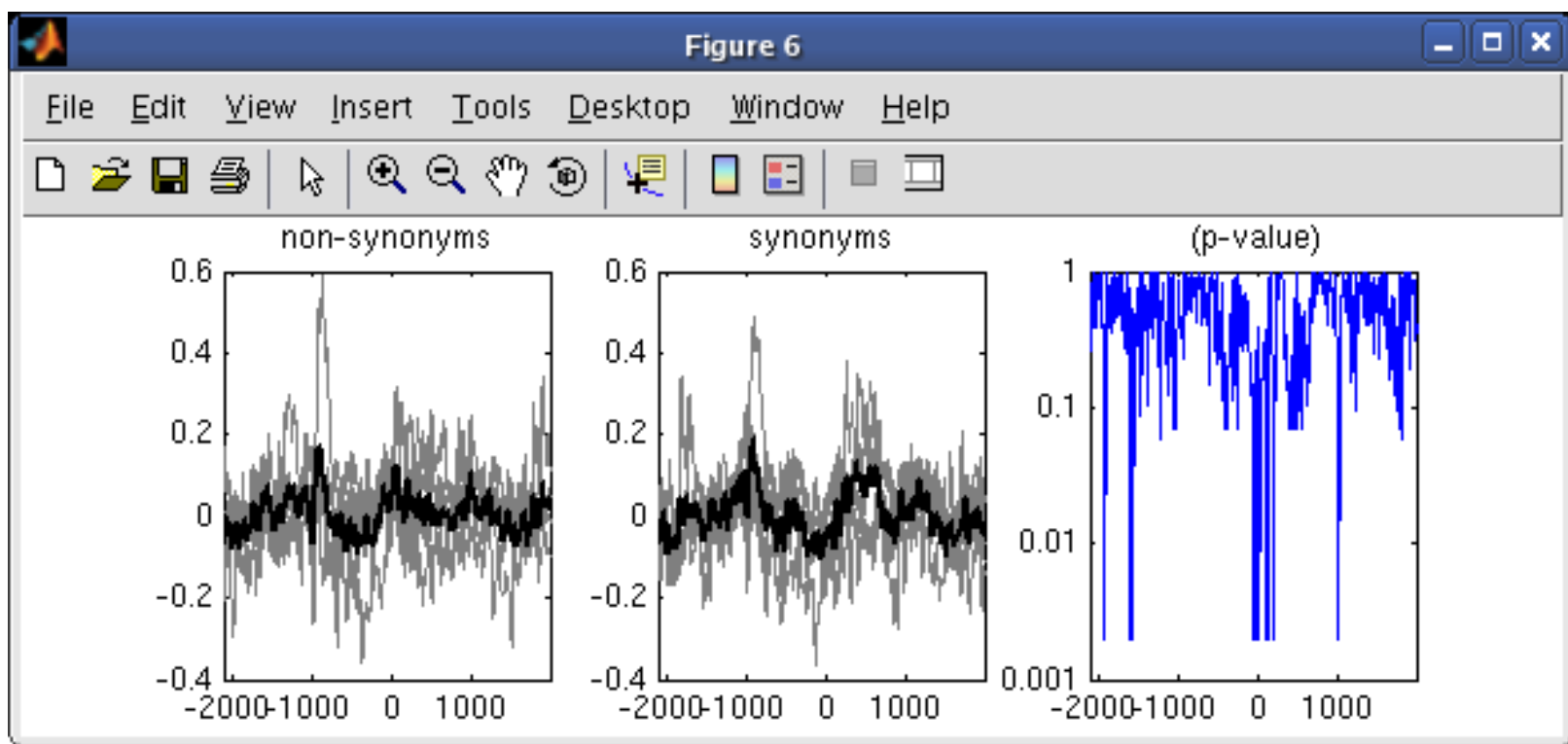
Plot cluster data



Plot cluster data



Plot cluster ERP



Exercise

- Load the STUDY stern.study
- Precompute **spectrum** and **scalp maps** for components
- Precluster and cluster components using **dipole locations** and **dipole moments** (affinity clustering)
- Look at your cluster. Identify frontal midline theta cluster(s) and occipital alpha cluster(s)
- Remove outliers if any
- Plot significant difference (parametric statistics) for one component cluster spectrum between the two conditions ignore vs memorize

