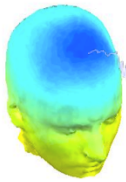
Clustering of ICA components

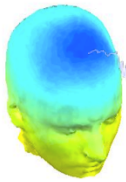
Arnaud Delorme

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



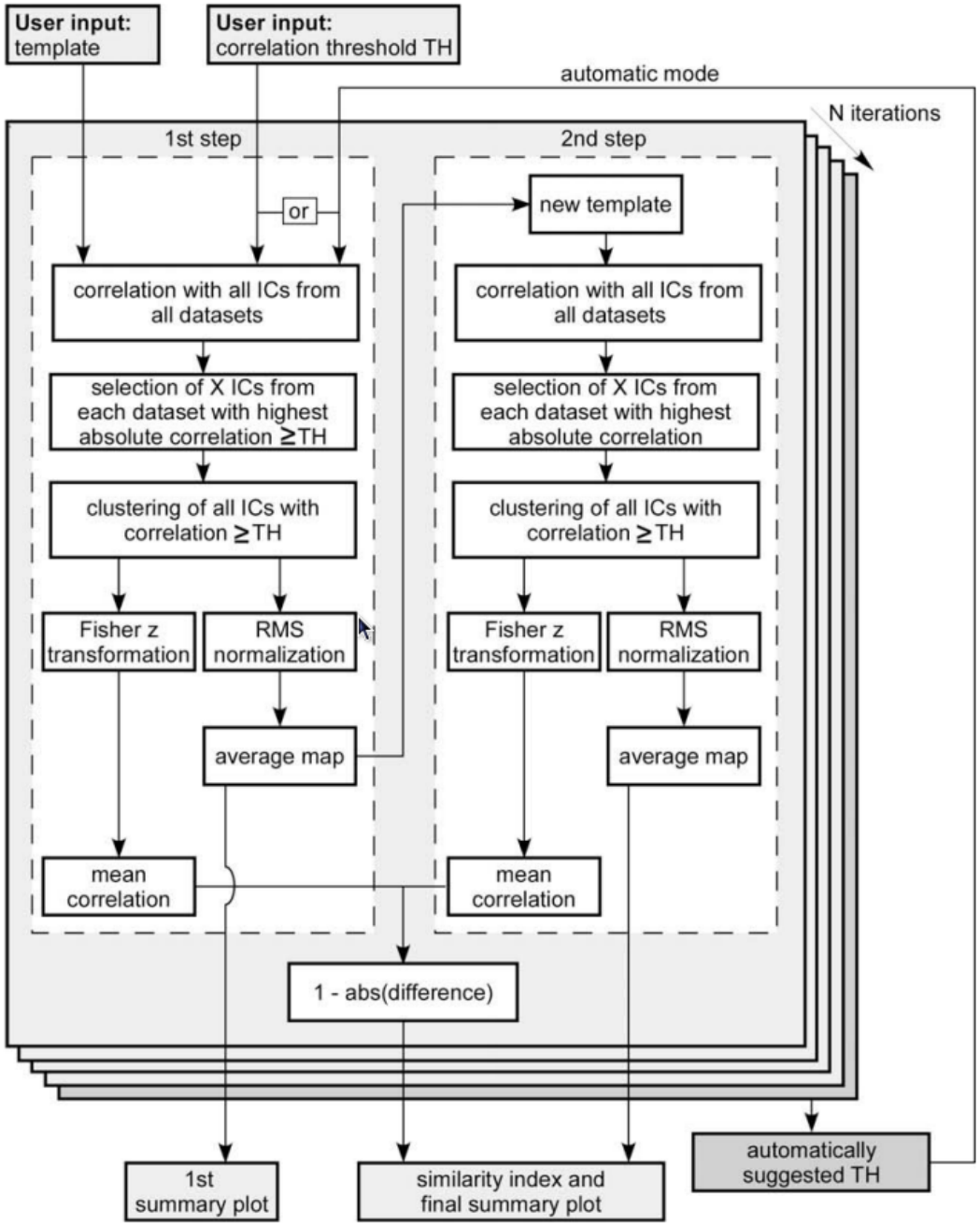
Outline

- ICA clusters and reliability within subjects
- ICA clusters and reliability across subjects
- Clustering in EEGLAB theory & Practice



ICA decomposition of multiple data sets from the same individuals

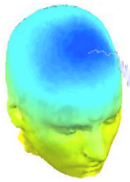
- Experimental protocol
  - Mind wandering experiment
  - 2 subjects
  - 11 x 30 min. sessions
  - 2 sessions per week
  - EEG from Biosemi 64 channels
  - Fs=1024 Hz



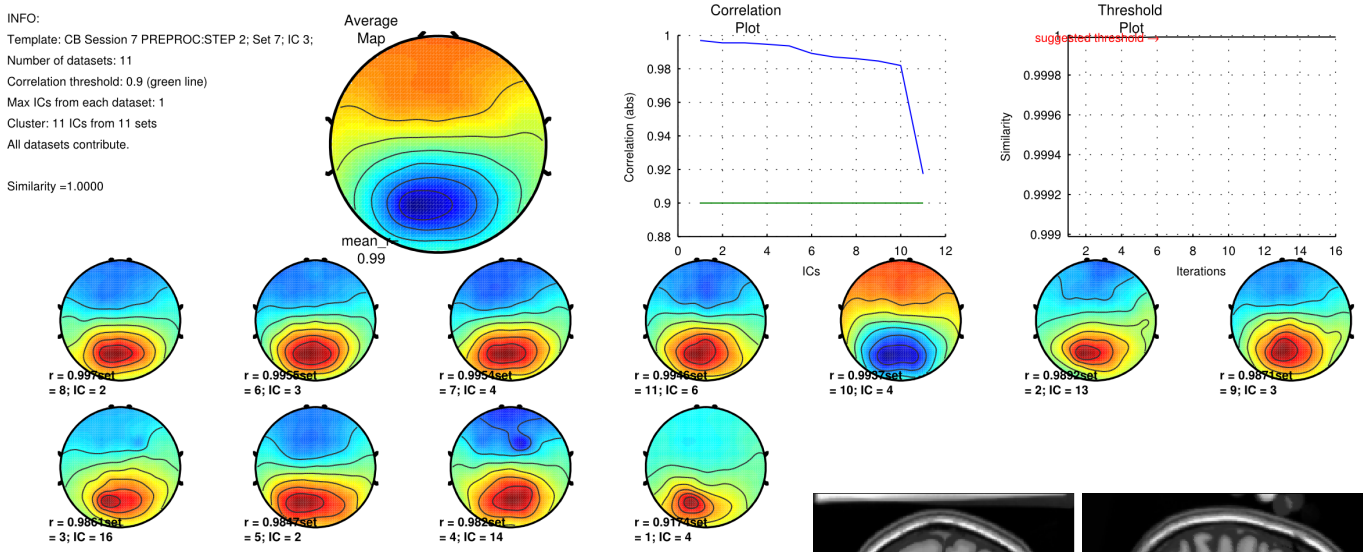
F. Campos Viola et al., “Semi-automatic identification of independent components representing EEG artifact,” Clinical Neurophysiology 120, no. 5 (2009): 868–877.

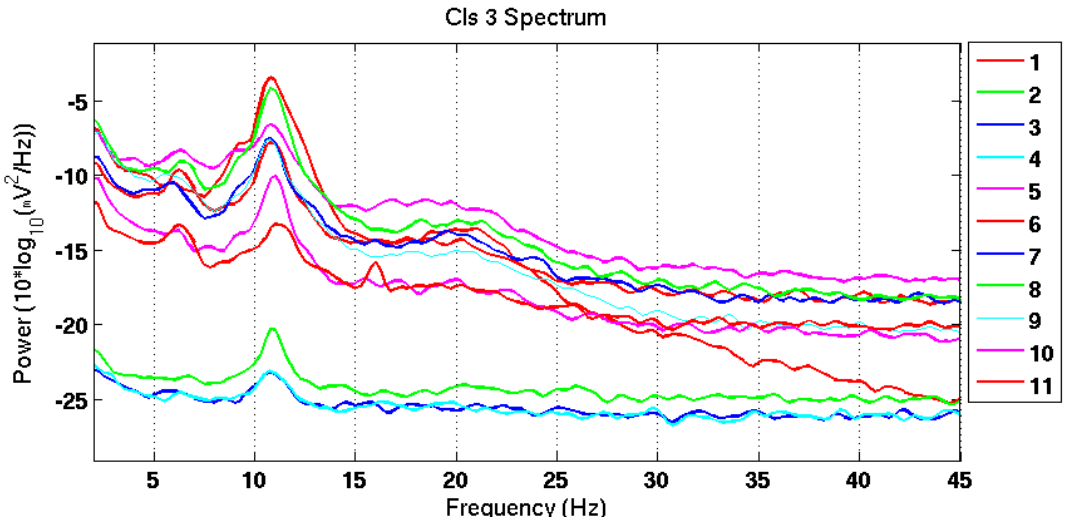
suggested as the automatic correlation threshold.

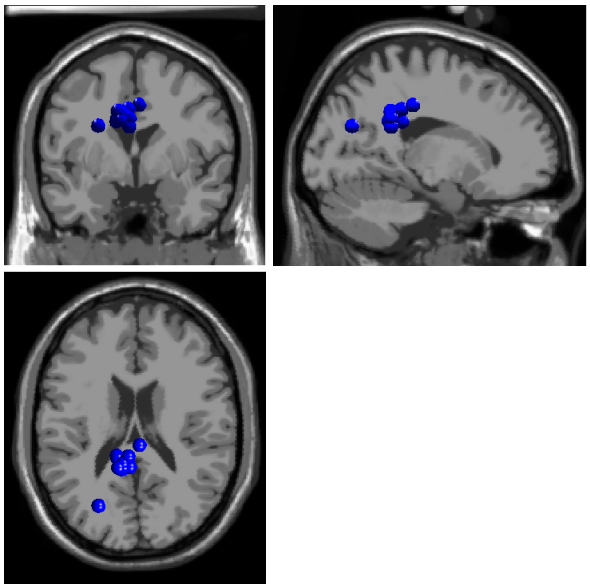
Results (Cluster 1)



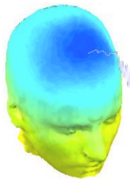
100 % Sessions contribute



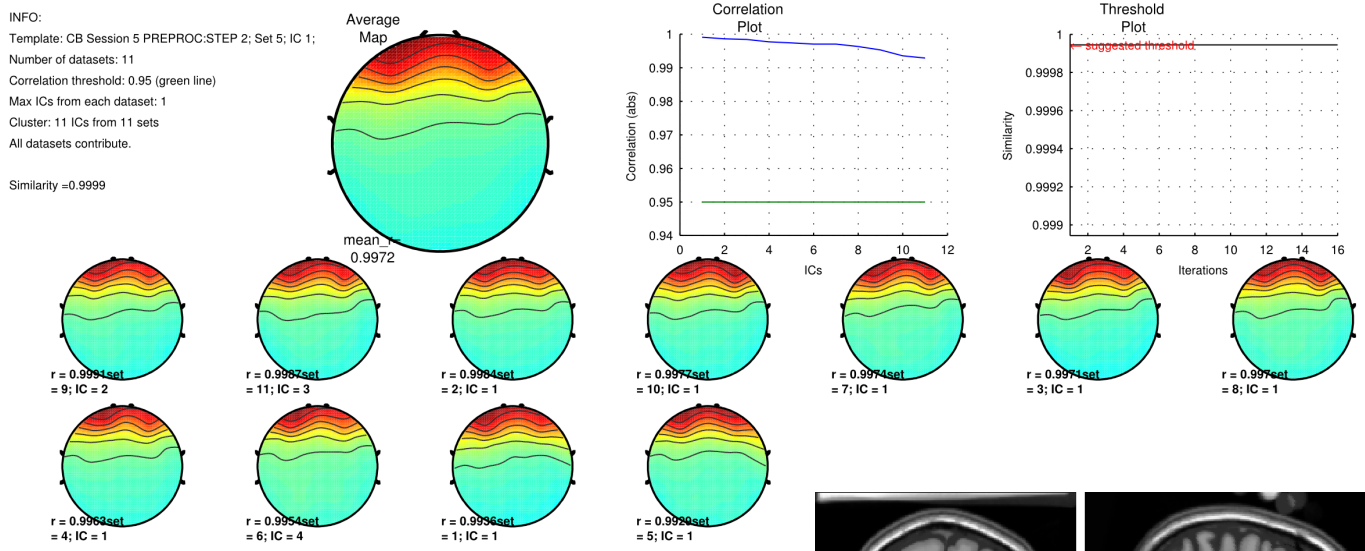


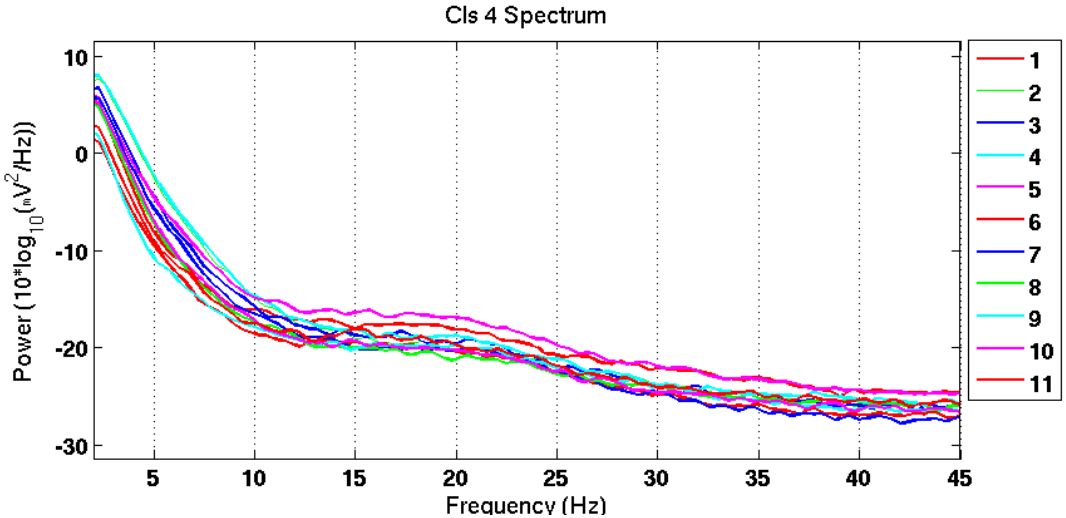


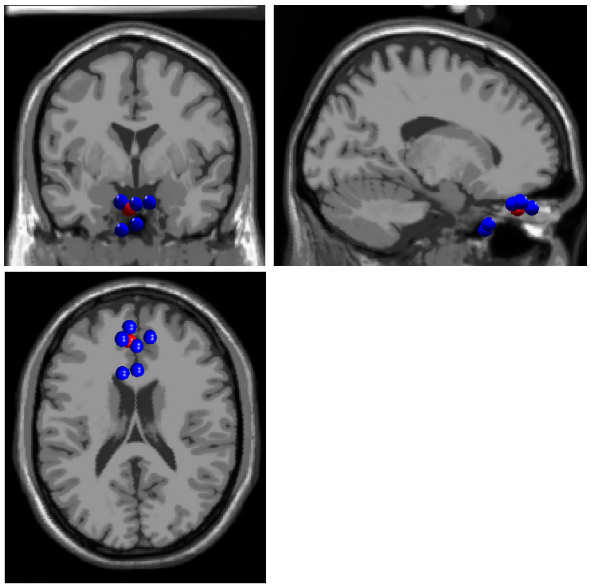
Results (Cluster 2)



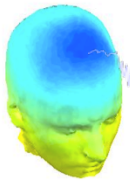
100 % Sessions contribute



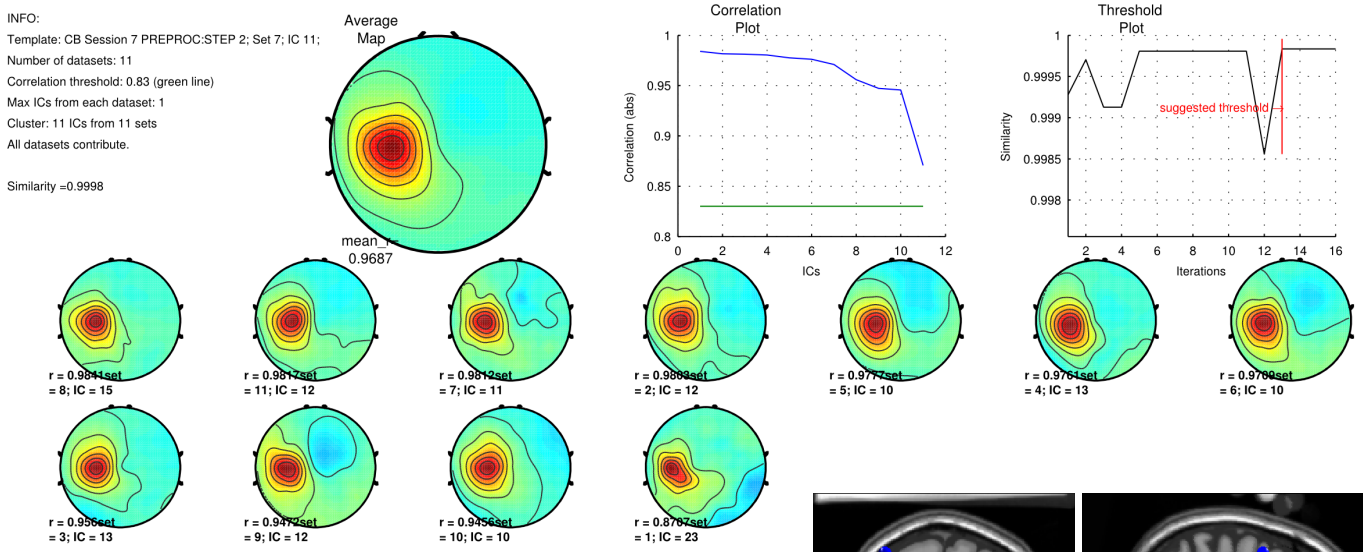


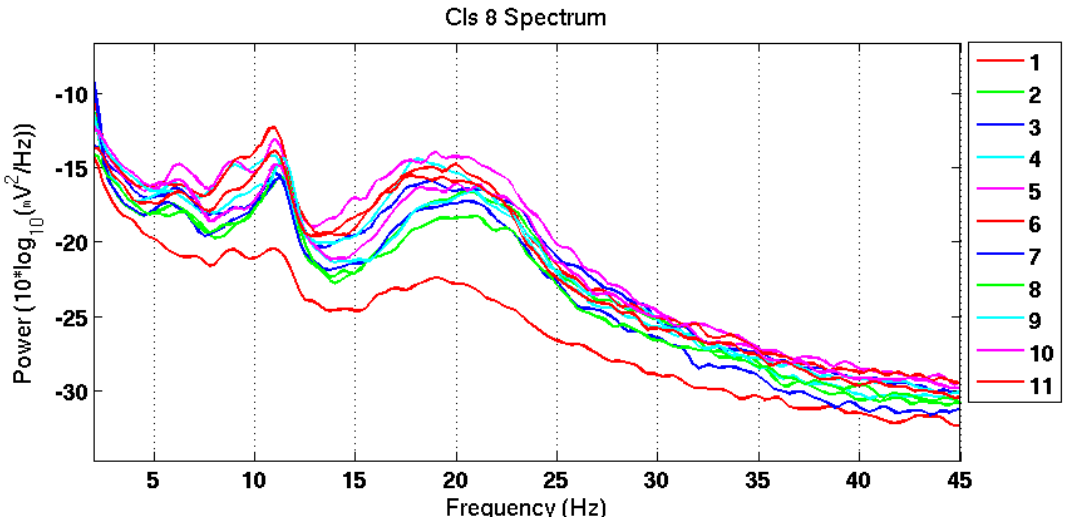


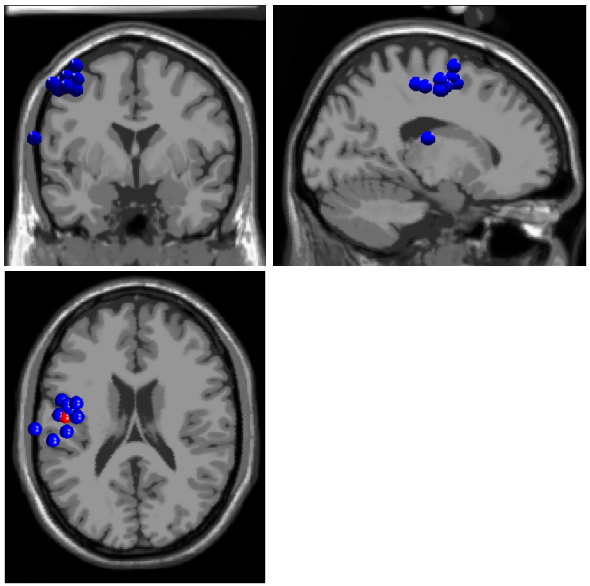
Results (Cluster 8)



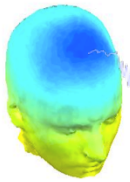
100 % Sessions contribute



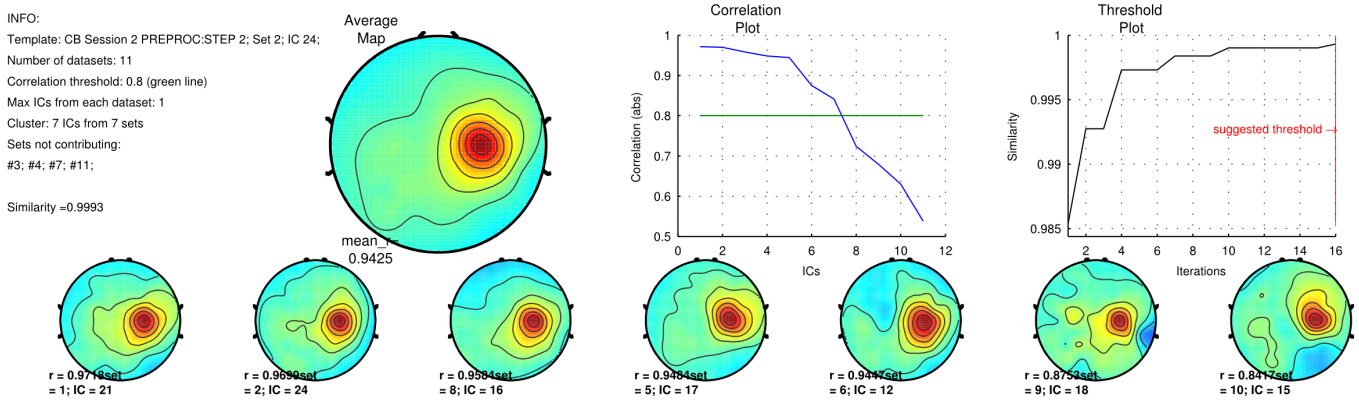


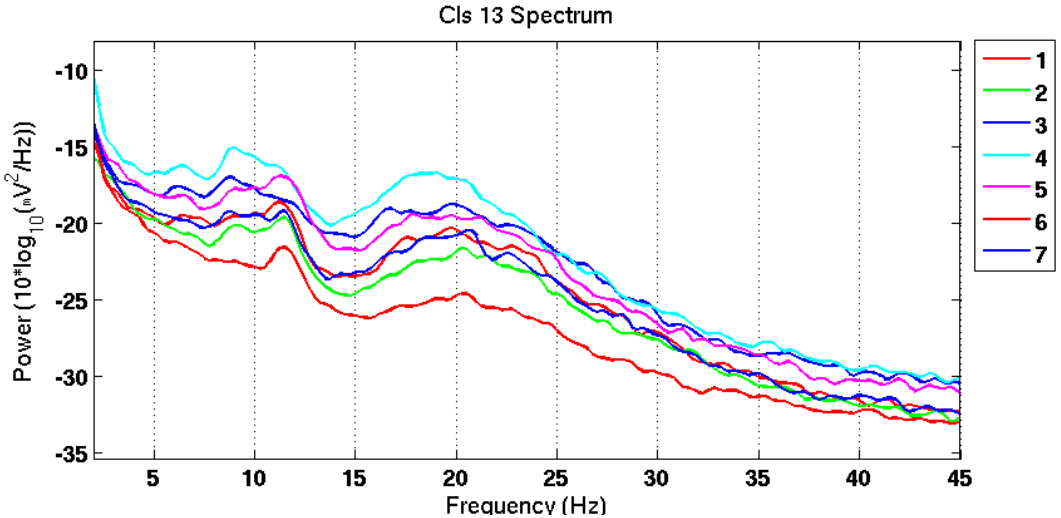


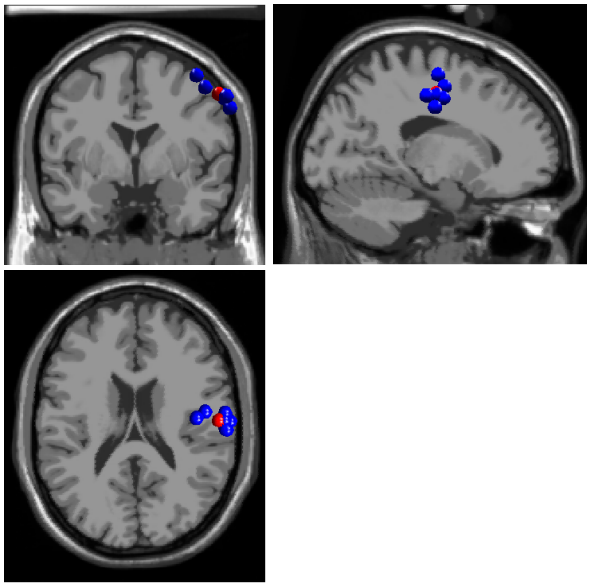
Results (Cluster 13)



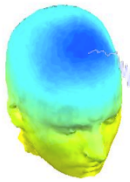
63.64% Sessions contribute



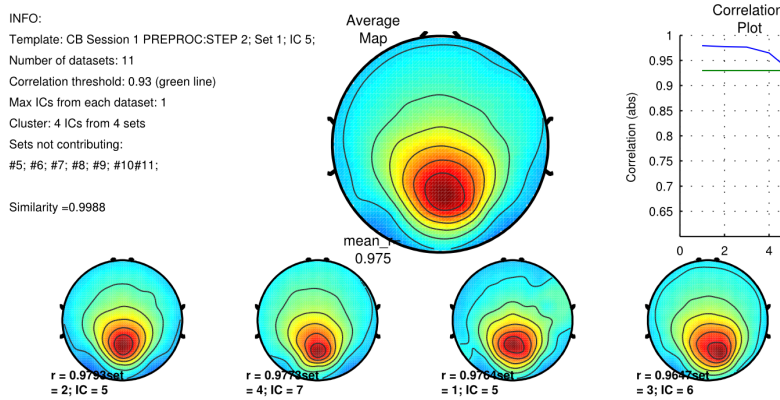


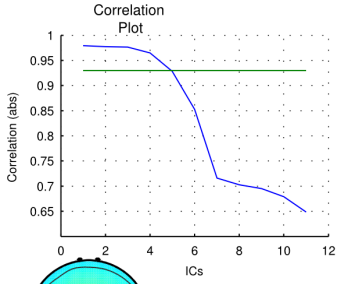


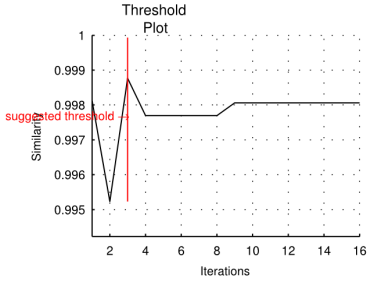
Results (Cluster 14)

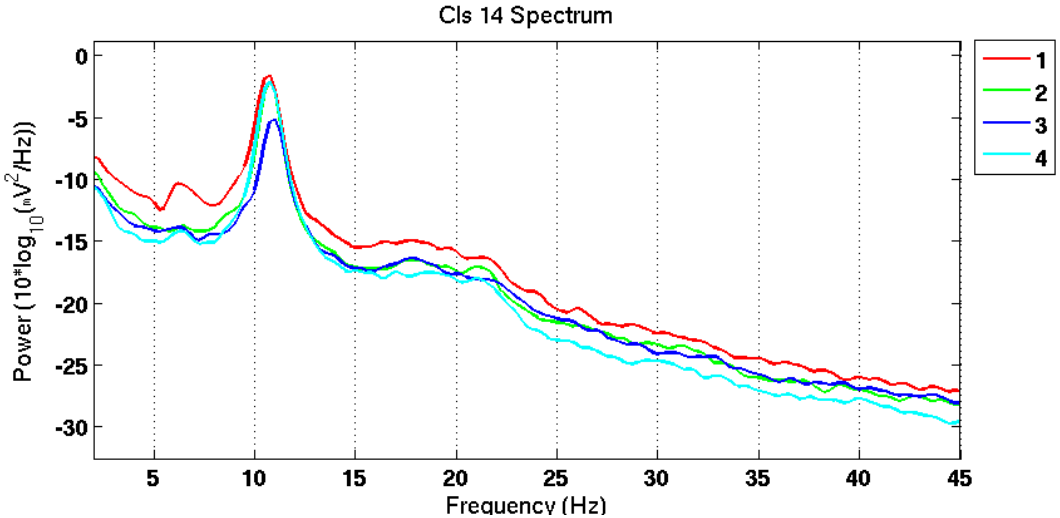


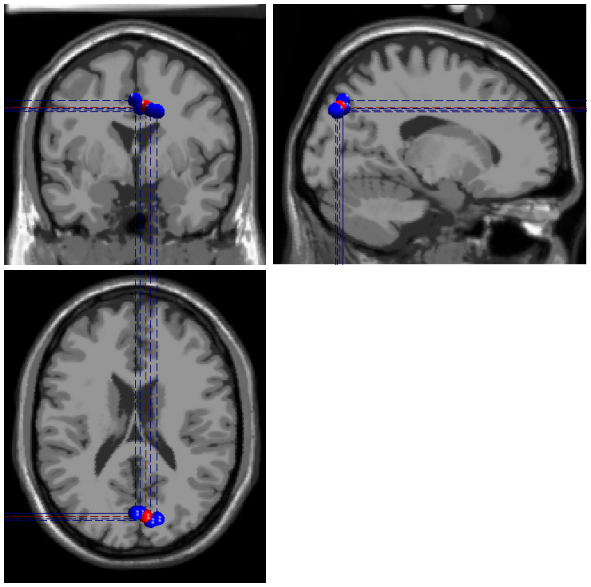
36.36% Sessions contribute



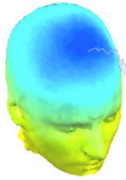








Inter iteration Cluster Consistency

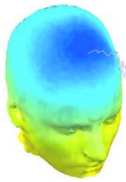


Swartz Center for Computational Neuroscience

Iterations

Clusters

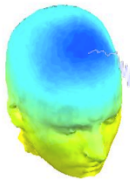
	1	2	3	4	5	6	7	8	9	10	Mean
3	100	100	100	100	100	100	100	100	100	100	**100**
4	100	100	100	100	100	100	90	100	100	100	**99**
5	90	40	10	90	90	60	100	10	60	90	**64**
6	60	0	100	60	100	90	60	60	90	60	**68**
7	90	100	90	90	60	90	90	100	90	90	**89**
8	80	80	60	80	40	80	80	80	80	100	**76**
9	60	90	50	60	80	60	0	10	60	50	**52**
10	40	90	10	40	0	50	50	0	50	60	**39**
11	60	20	0	0	10	60	10	90	60	60	**37**
12	100	50	50	100	50	100	100	50	100	50	**75**
13	50	10	20	50	90	50	50	10	50	20	**40**
14	20	10	10	20	20	30	20	20	30	30	**21**



Outline

- ICA clusters and reliability within subjects
- ICA clusters and reliability across subjects
- Clustering in EEGLAB theory & Practice

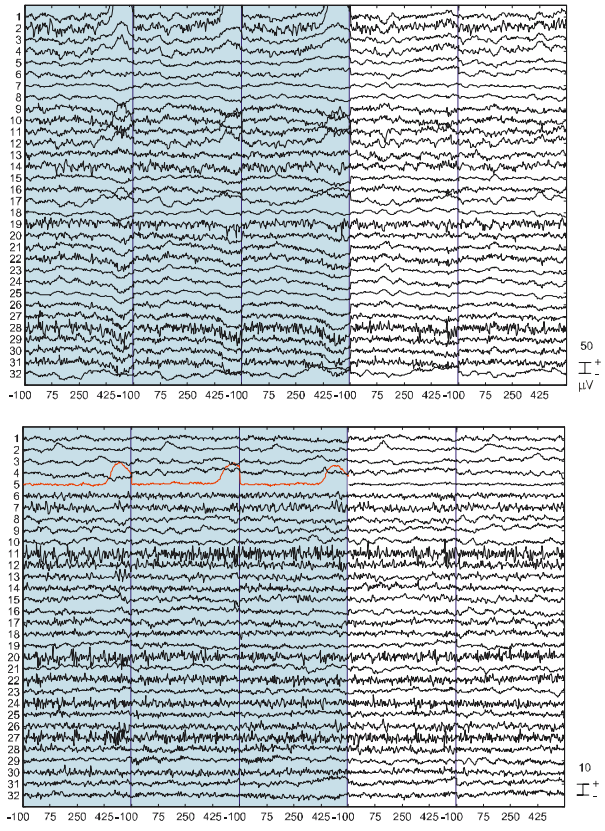
Localization



ICA

Electrodes

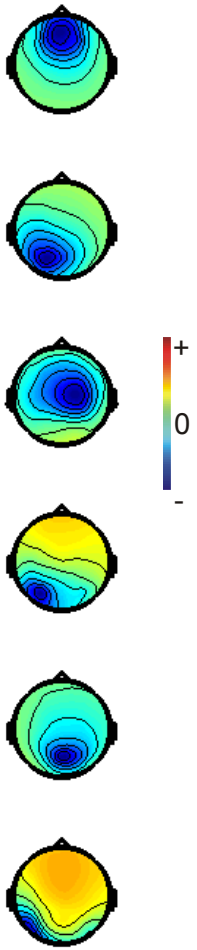
Components



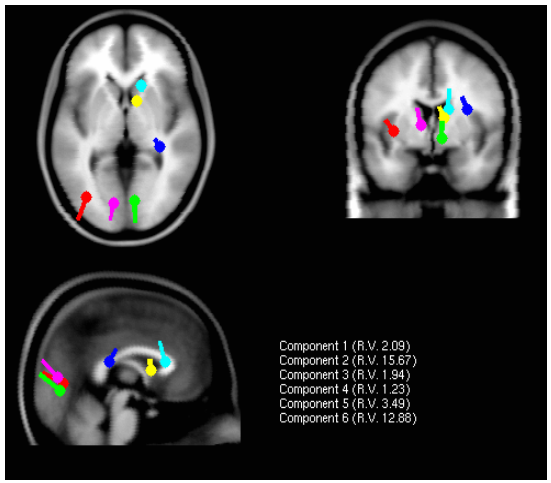
Time (ms)

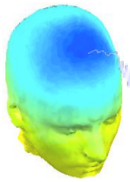
ICA component

scalp maps



Localization

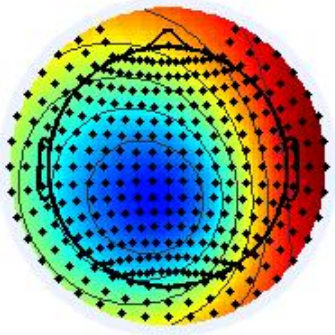




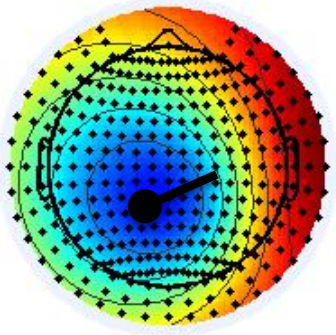
Computing residual variance (%)

Swartz Center for Computational Neuroscience

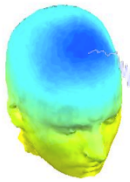
Actual



Dipole projection



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



Validation of the ICA algorithm for EEG

Swartz Center for Computational Neuroscience

Data

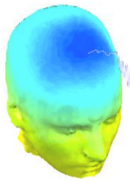
- 13 subjects performing a memory task
- 71 electrodes including EOGs
- more than 300,000 data points/subject

Decomposition

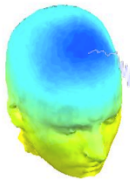
- 23 ICA algorithms plus PCA and Promax

Analysis

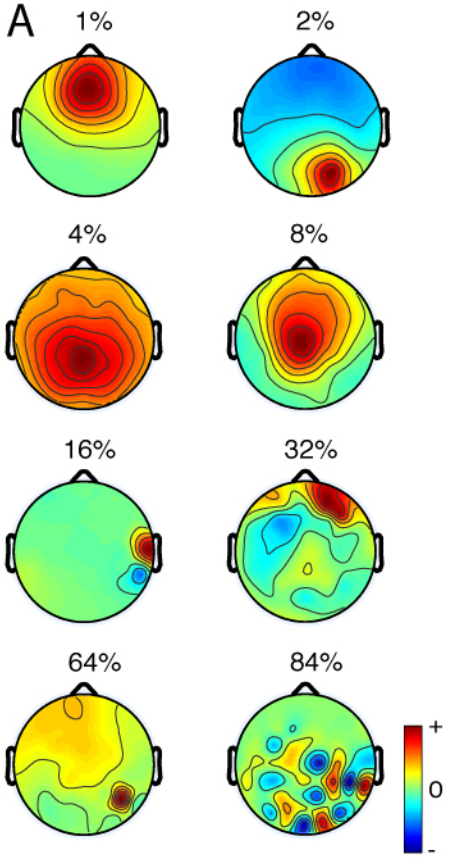
- Localization of all components with a single dipole (4-shell spherical model)



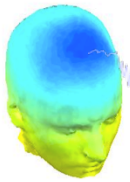
Algorithm (Matlab func.)	D%	LL	Origin
Extended Infomax (runica)	29.9	178	EEGLAB 4.515
Pearson	29.1	169	ICAcentral (6)
Infomax (runica)	28.2	160	EEGLAB 4.515
ERICA	26.9	184	ICALAB 1.5.2
SONS	25.4	183	ICALAB 1.5.2
SHIBBS	23.7	169	ICAcentral (5)
FastICA*	23.5	169	ICAcentral (2)
JADE (jader)	23.4	169	EEGLAB 4.515
TICA	23.4	169	ICALAB 1.5.2
JADE optimized (jade_op)	21.4	169	ICALAB 1.5.2
JADE w/ time delay (jade_td)	20.2	169	ICALAB 1.5.2
eeA	19.0	305	ICAcentral (8)
Infomax (icaML) †	18.8	212	ICA DTU Tbox
FOBI	18.6	169	ICALAB 1.5.2
SOBIRO (acsobiro)	17.9	167	EEGLAB 4.515
EVD 24	17.7	169	ICALAB 1.5.2
EVD	17.0	169	ICALAB 1.5.2
SOBI	16.1	583	EEGLAB 4.515
icaMS†	10.6	169	ICA DTU Tbox
AMUSE	8.5	169	ICALAB 1.5.2
PCA	3.1	583	EEGLAB 4.515
Promax	33.7	467	EEGLAB 4.515
Whitening/Sphering	57.6	164	EEGLAB 4.515

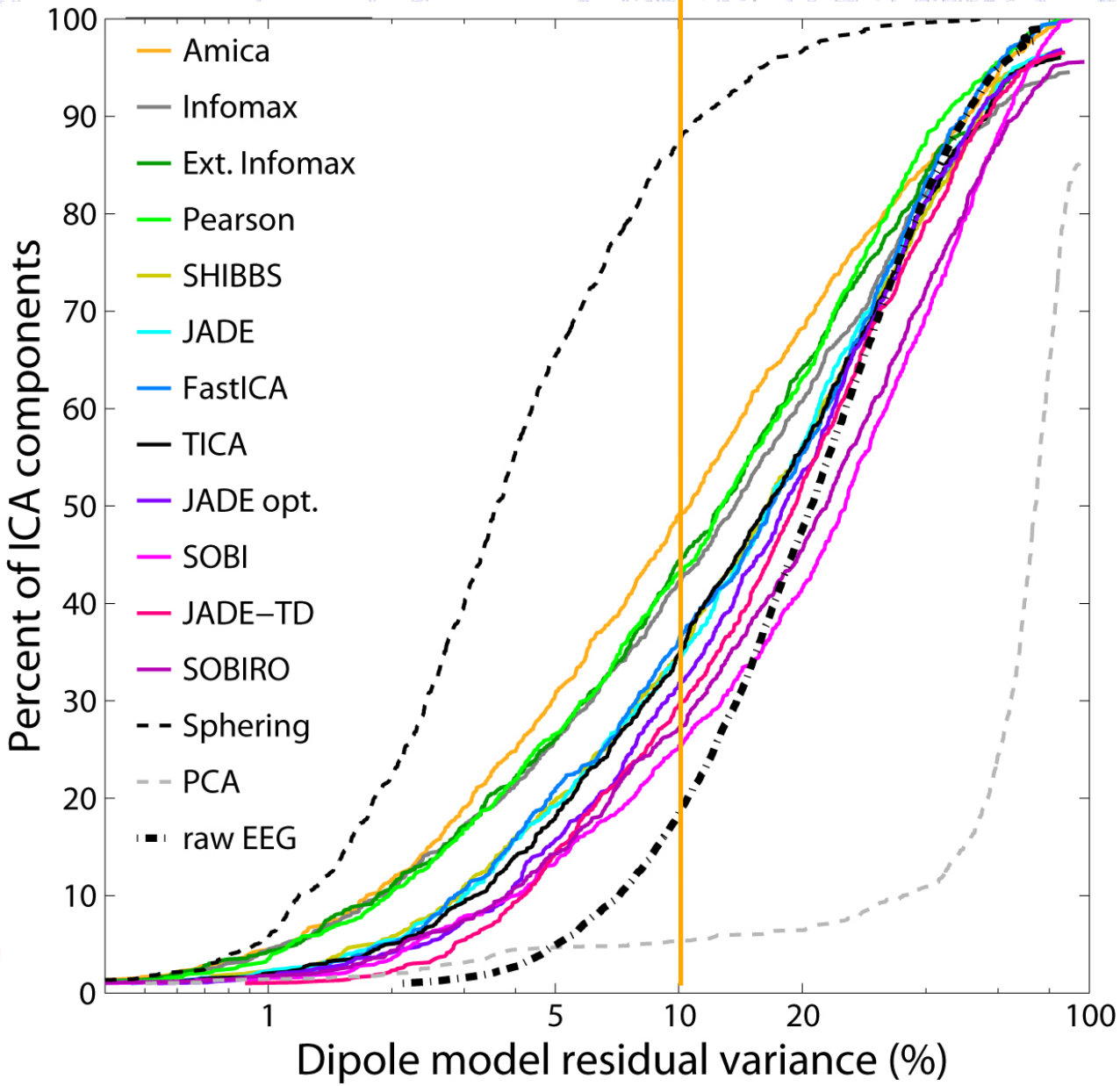


Component examples

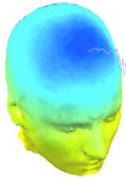


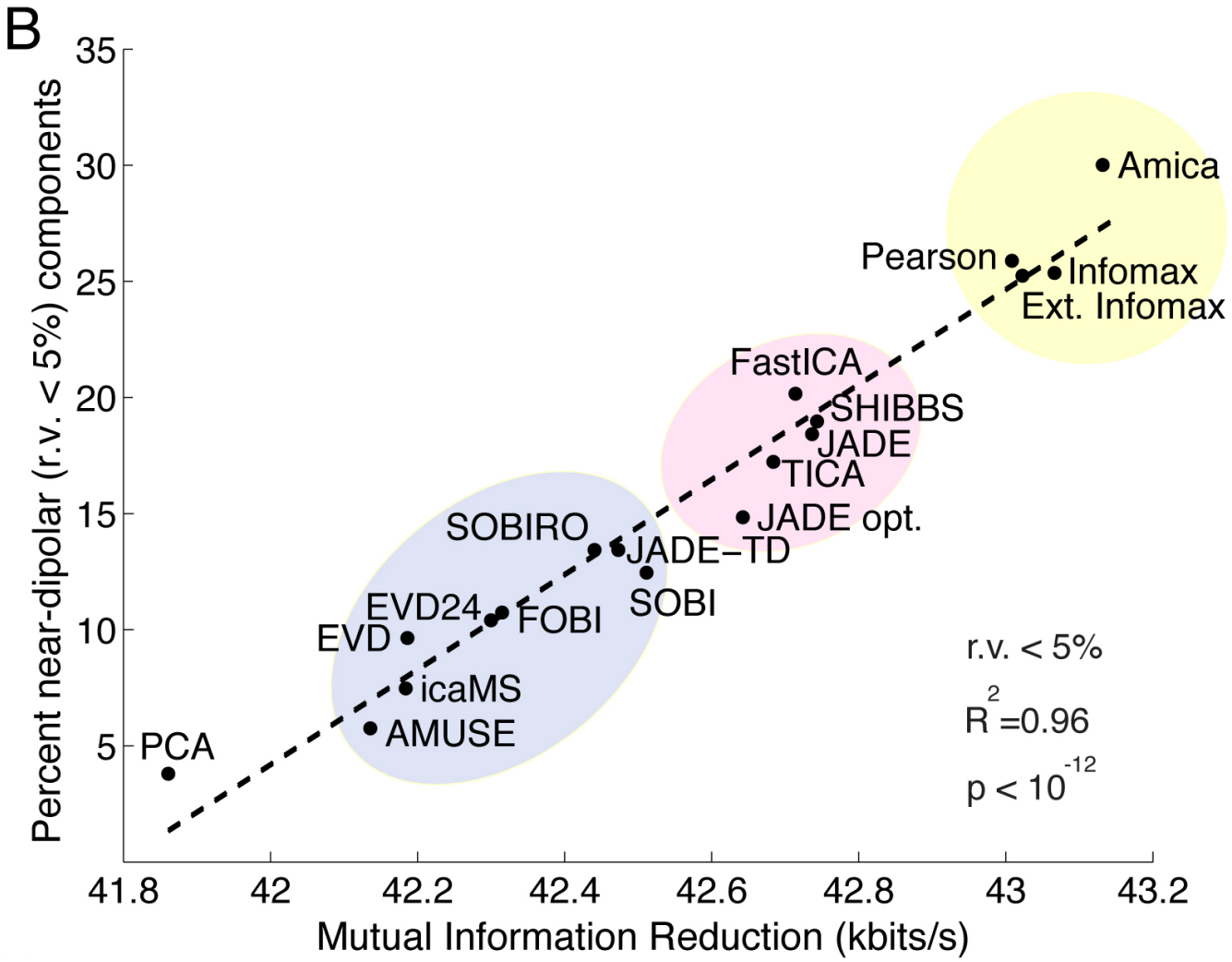
Cumulative number of component below residual variance threshold

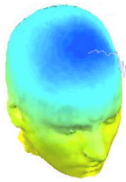




More independence -> more biological components







C

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

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

Amica
Ext. Infomax
Pearson
Infomax
FastICA
SHIBBS
JADE
TICA
JADE opt.
SOBIRO
JADE–TD
SOBI
FOBI
EVD
EVD24
icaMS
AMUSE
PCA

r.v. < 5%

R^2=0.71

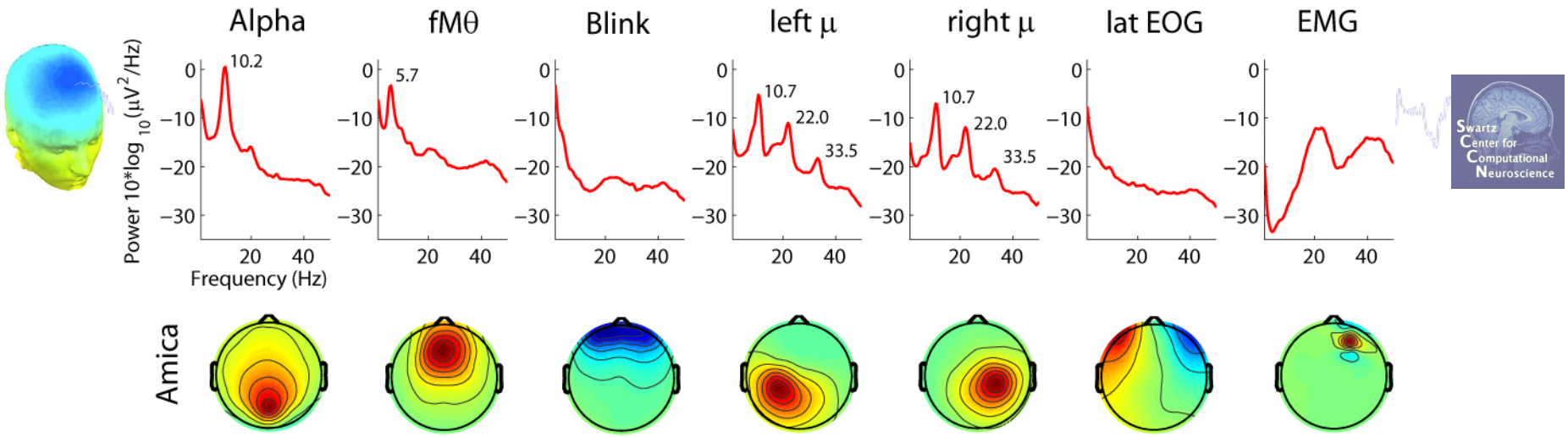
$p = 10^{-5}$

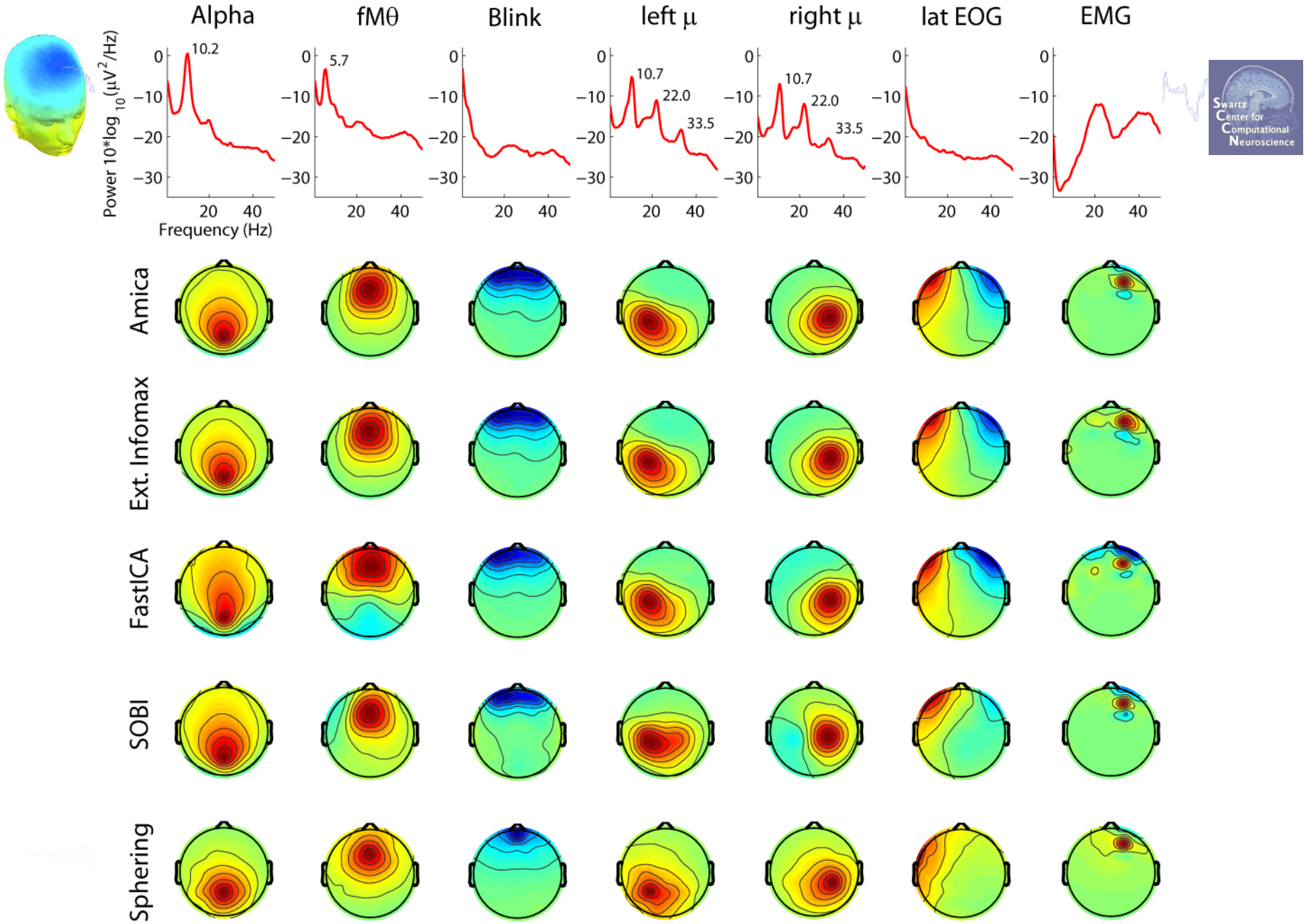
R^2 value

Probability (10^x)

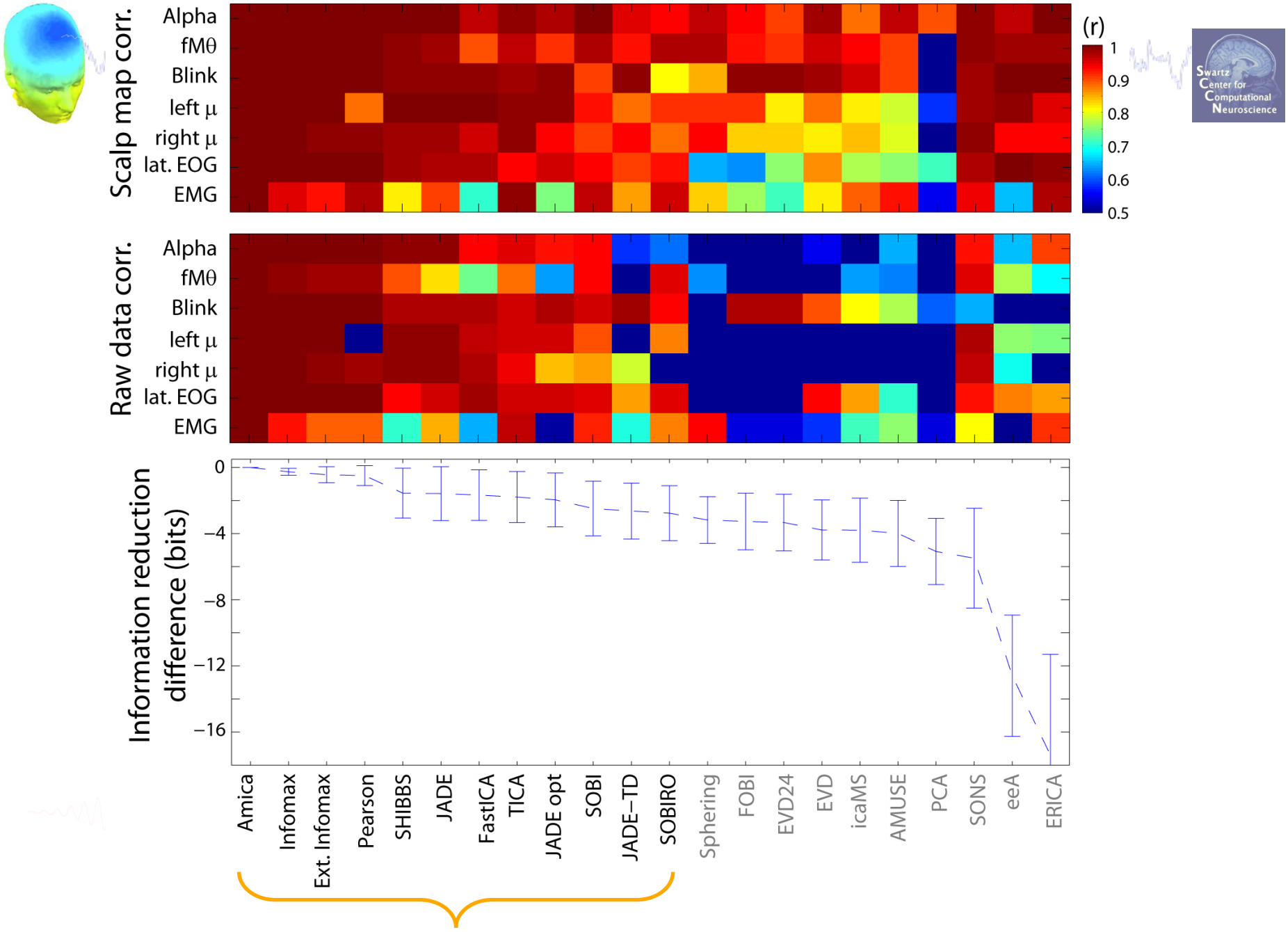
R^2
probability

Near dipolar cutoff (% r.v.)

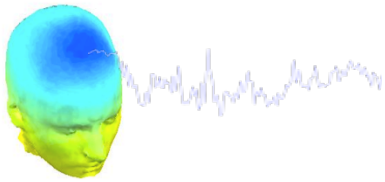

Alpha
fMθ
Blink
left μ
right μ
lat EOG
EMG
Power 10*log 10(μV2/Hz)
10.2
5.7
10.7
22.0
33.5
10.7
22.0
33.5
0
−10
−20
−30
20
40
Frequency (Hz)
Amica
Swartz Center for Computational Neuroscience

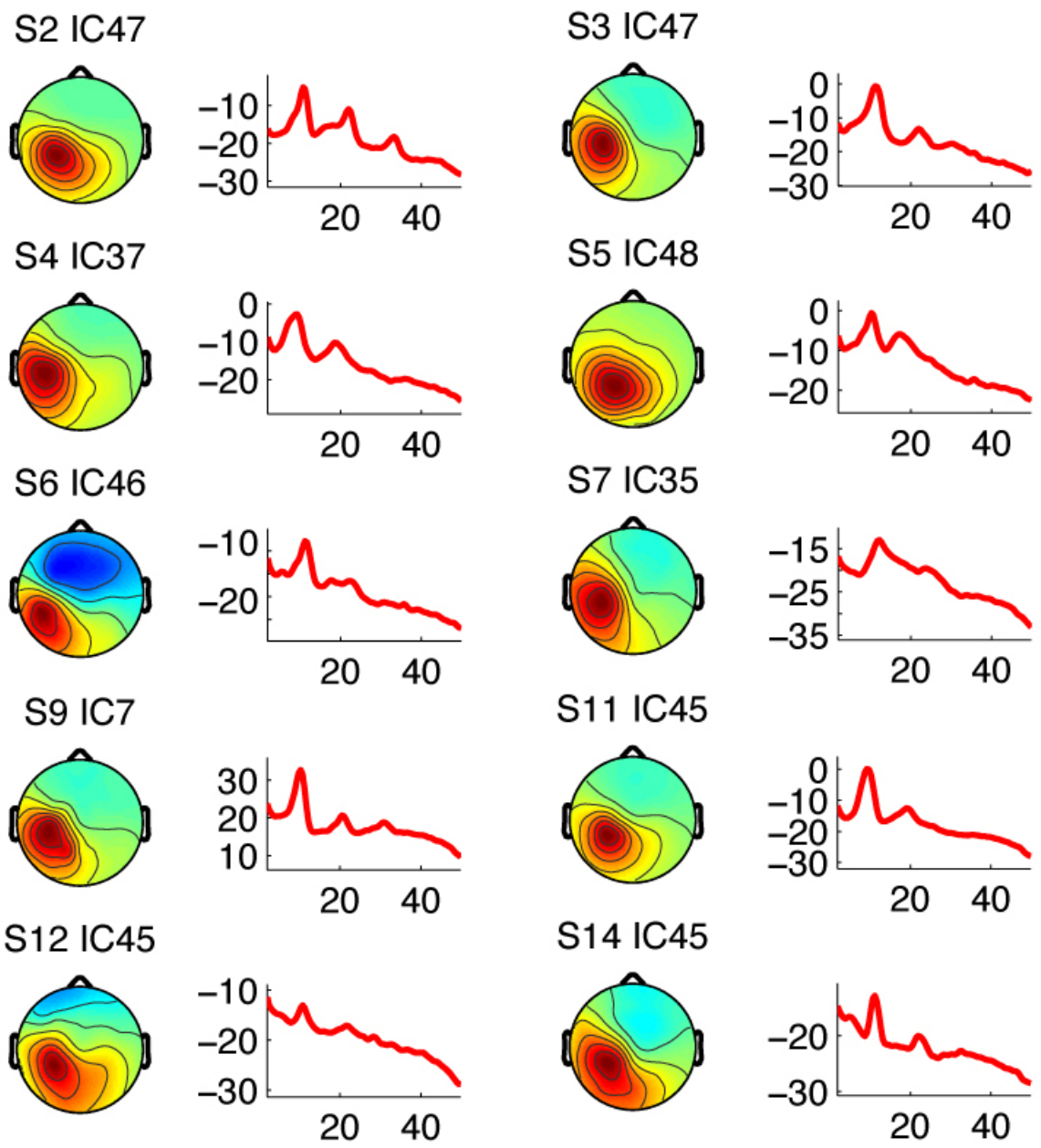

Alpha
fMθ
Blink
left μ
right μ
lat EOG
EMG
10.2
5.7
10.7
22.0
33.5
10.7
22.0
33.5
Power 10*log 10 (μV2/Hz)
Frequency (Hz)
Amica
Ext. Infomax
FastICA
SOBI
Sphering
Swartz Center for Computational Neuroscience

Correlations between decompositions

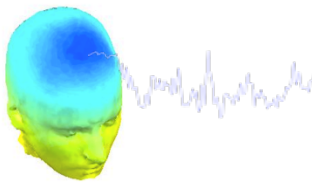


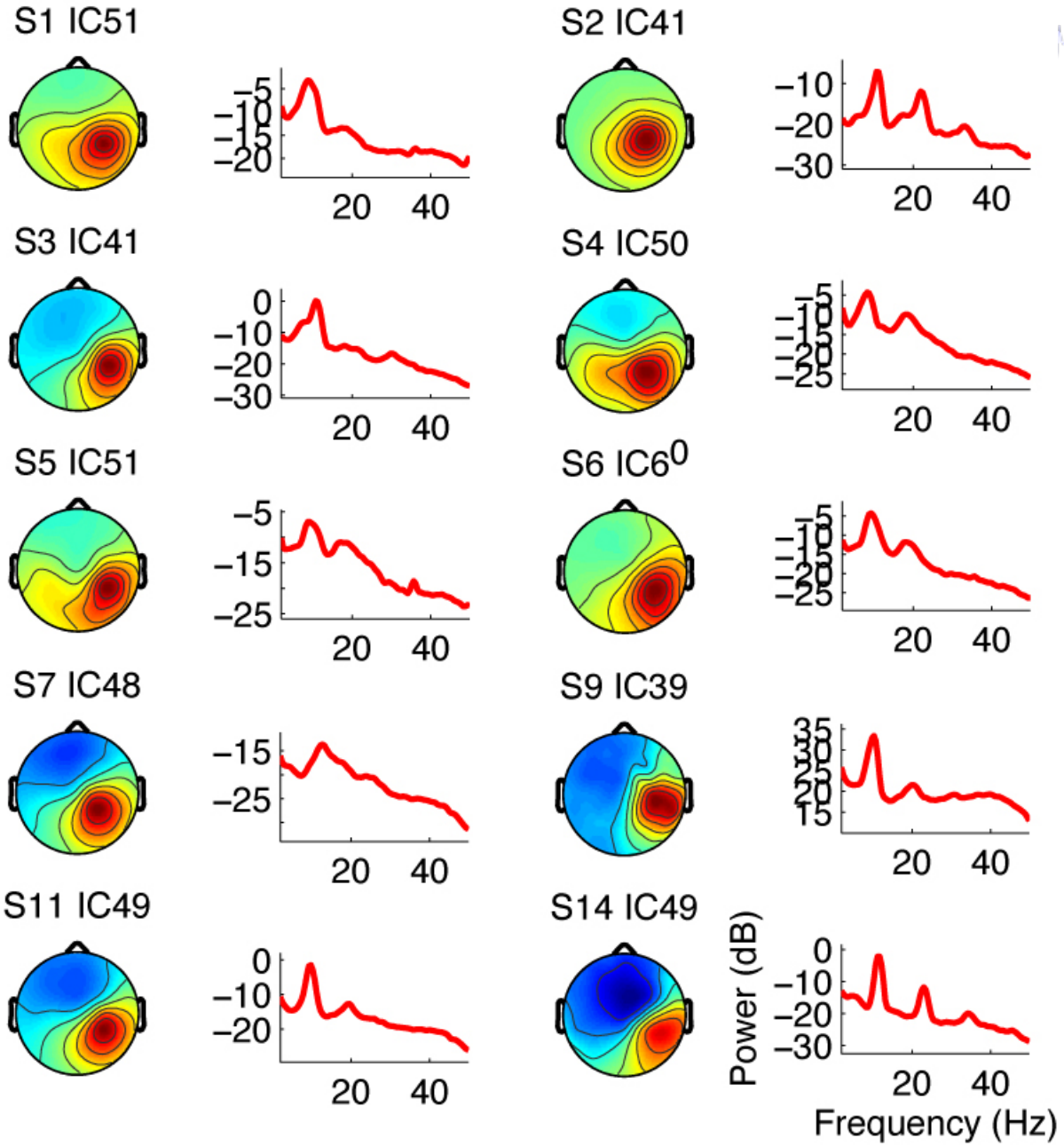
Left μ cluster

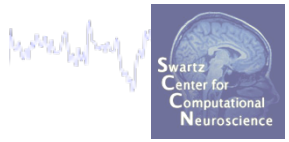




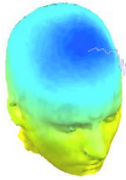
Right μ cluster

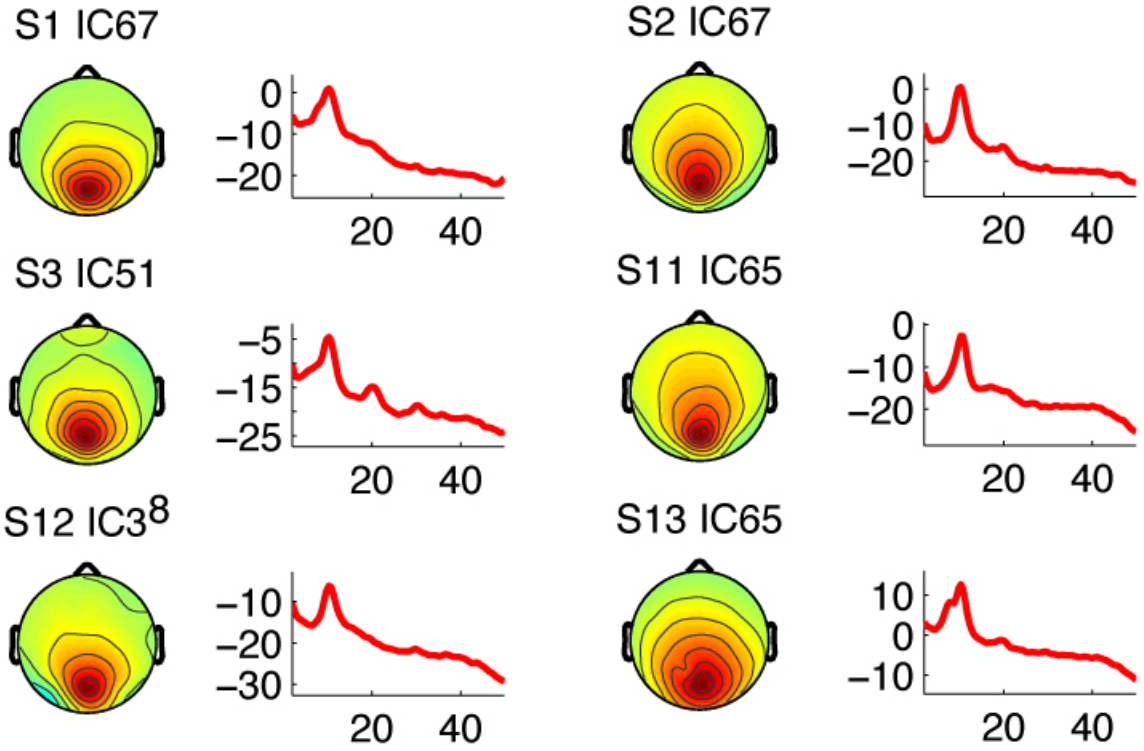




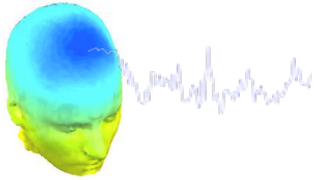


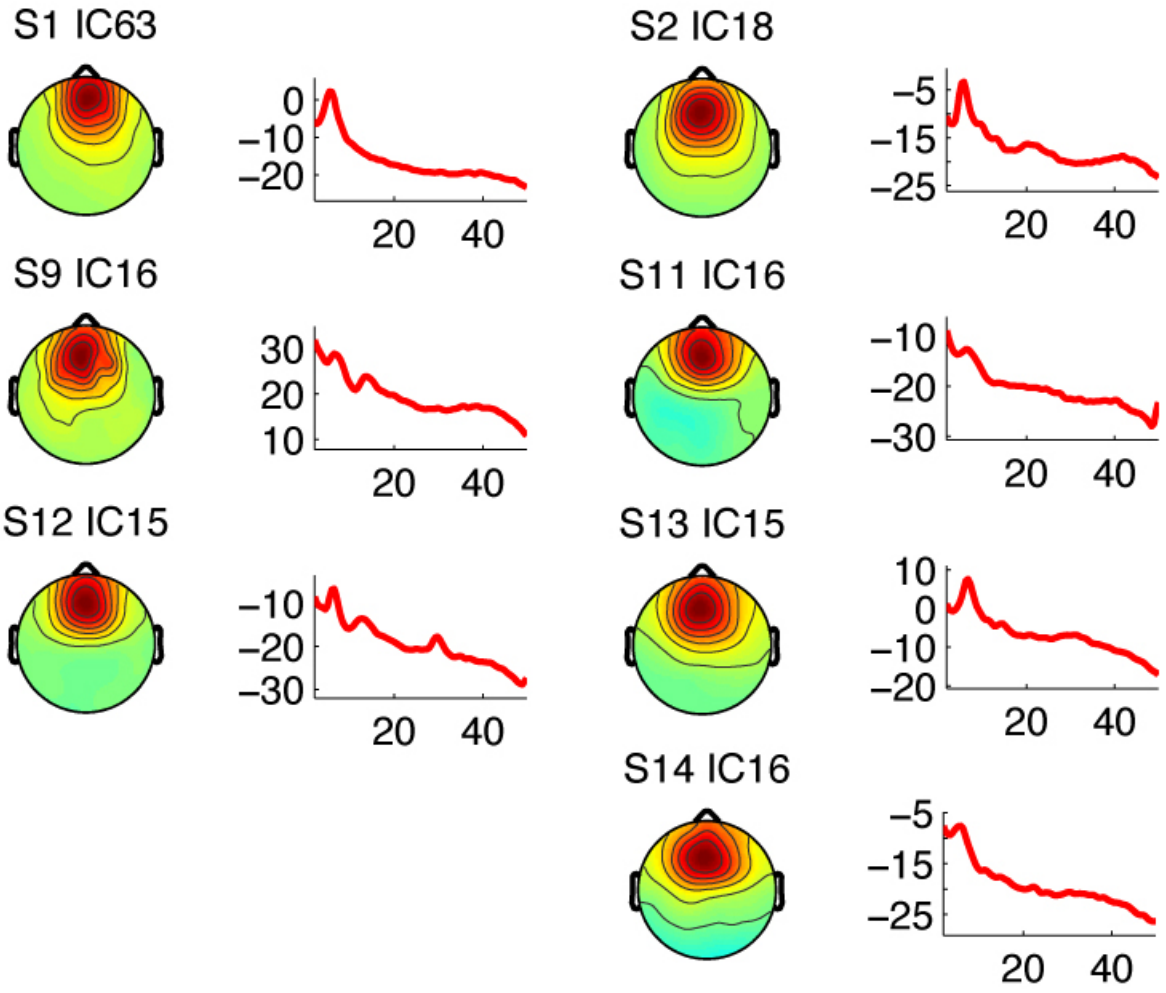
Occipital α cluster

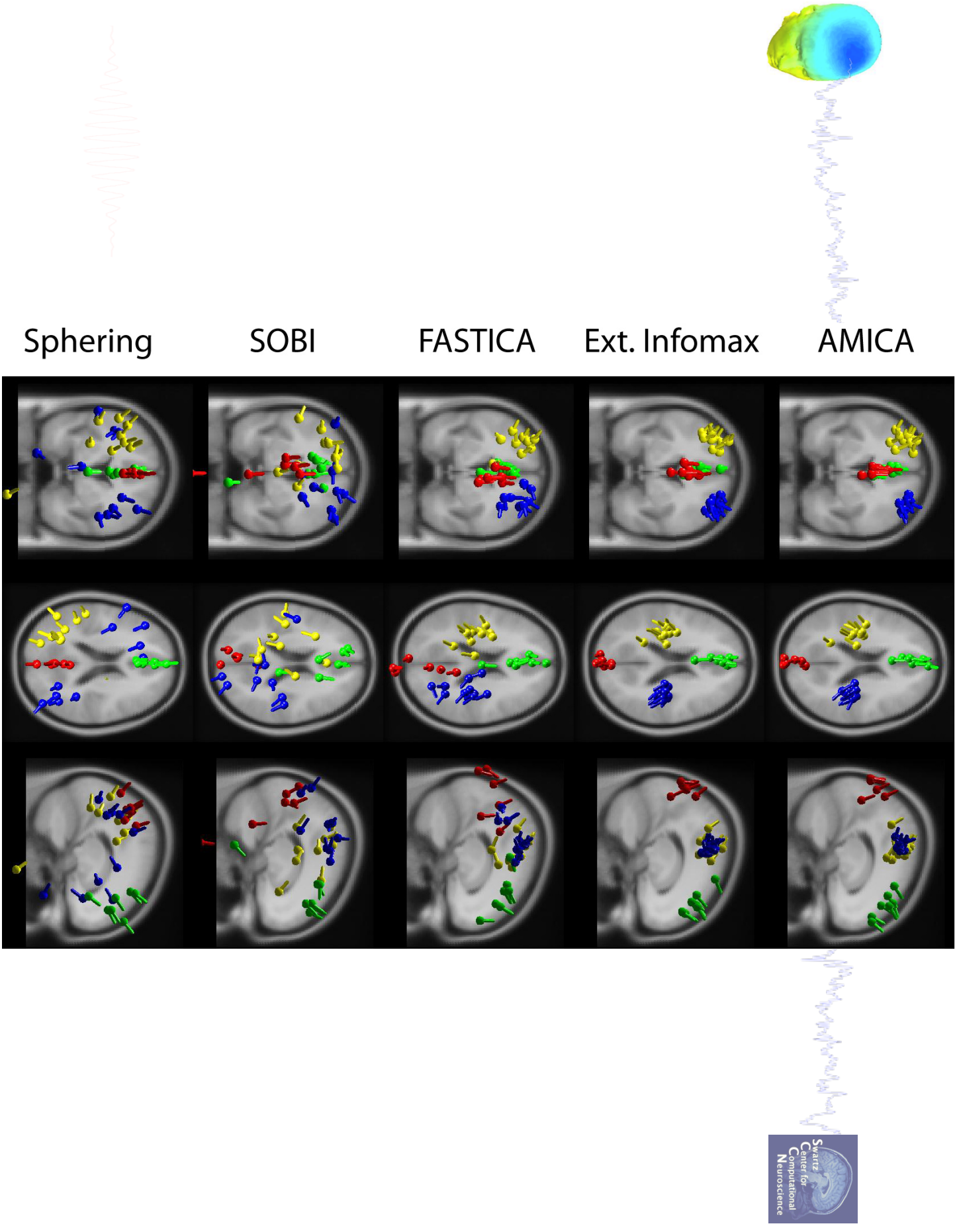


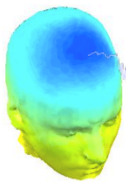


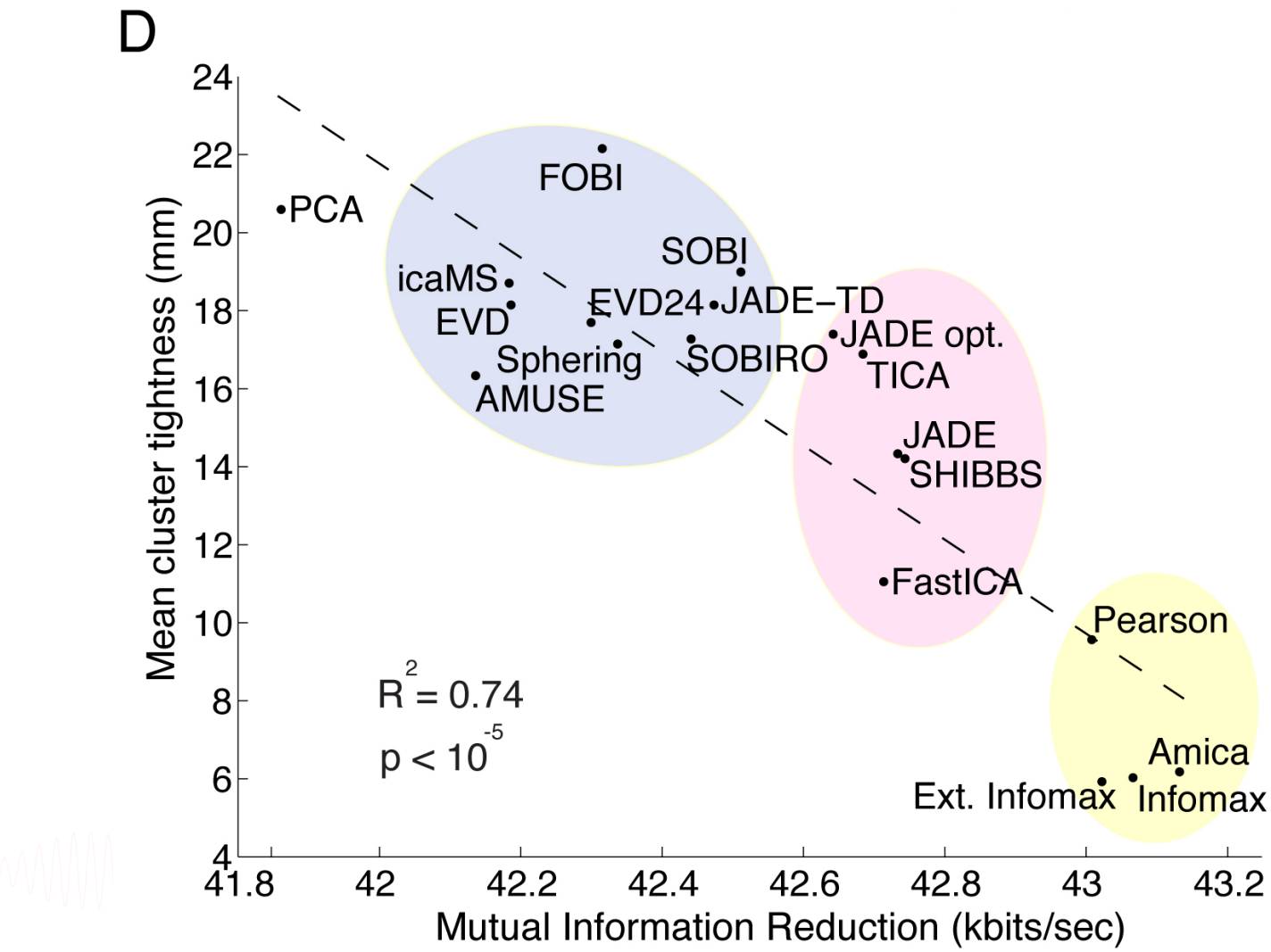
Frontal Midline θ cluster

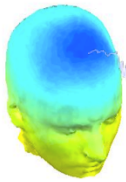





Sphering
SOBI
FASTICA
Ext. Infomax
AMICA
Swartz Center for Computational Neuroscience



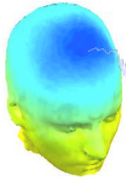

D
Mean cluster tightness (mm)
Mutual Information Reduction (kbits/sec)
24
22
20
18
16
14
12
10
8
6
4
41.8
42
42.2
42.4
42.6
42.8
43
43.2
PCA
FOBI
SOBI
icaMS
EVD
EVD24
JADE–TD
JADE opt.
TICA
Sphering
SOBIRO
AMUSE
JADE
SHIBBS
FastICA
Pearson
Amica
Ext. Infomax
Infomax
R2 = 0.74
p < 10-5

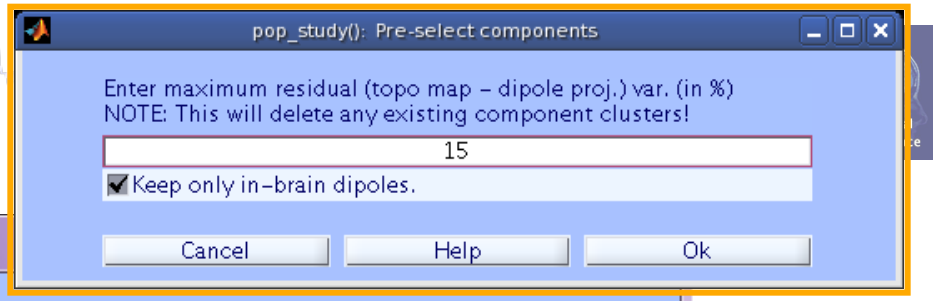


Outline

- ICA clusters and reliability within subjects
- ICA clusters and reliability across subjects
- Clustering in EEGLAB theory & Practice

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\Wor	...	S01		memorize		Comp.: 3 5 ...	Clear
2	C:\Users\julie\Documents\Wor	...	S01		ignore		Comp.: 3 5 ...	Clear
3	C:\Users\julie\Documents\Wor	...	S01		probe		Comp.: 3 5 ...	Clear
4	C:\Users\julie\Documents\Wor	...	S02		memorize		Comp.: 5 6 ...	Clear
5	C:\Users\julie\Documents\Wor	...	S02		ignore		Comp.: 5 6 ...	Clear
6	C:\Users\julie\Documents\Wor	...	S02		probe		Comp.: 5 6 ...	Clear
7	C:\Users\julie\Documents\Wor	...	S03		memorize		Comp.: 6 7 ...	Clear
8	C:\Users\julie\Documents\Wor	...	S03		ignore		Comp.: 6 7 ...	Clear
9	C:\Users\julie\Documents\Wor	...	S03		probe		Comp.: 6 7 ...	Clear
10	C:\Users\julie\Documents\Wor	...	S04		memorize		Comp.: 1 2 ...	Clear

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

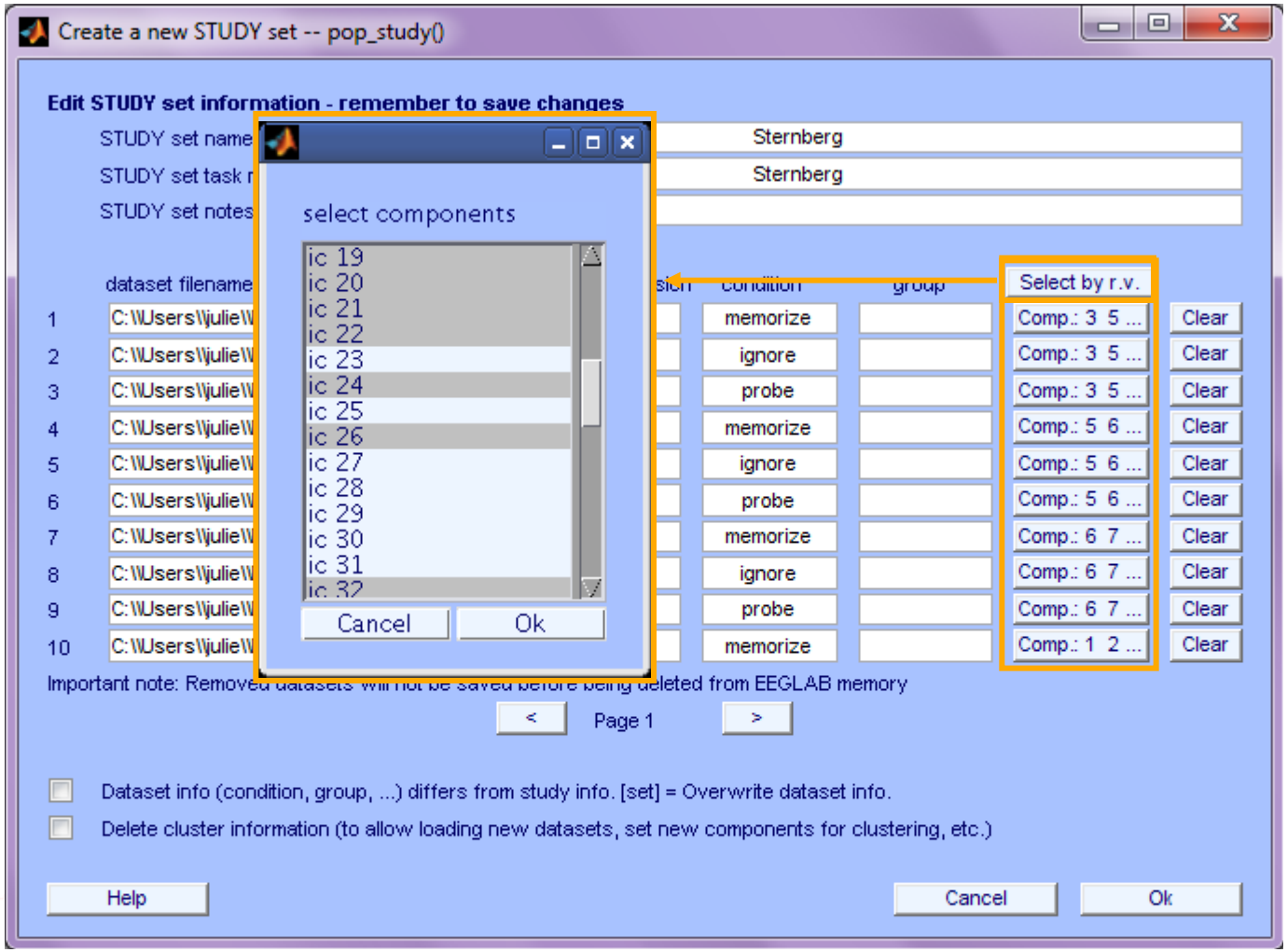
< Page 1 >

Dataset info (condition, group, ...) differs from study info. [set] = Overwrite dataset info.

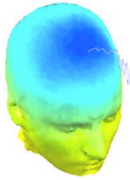
Delete cluster information (to allow loading new datasets, set new components for clustering, etc.)

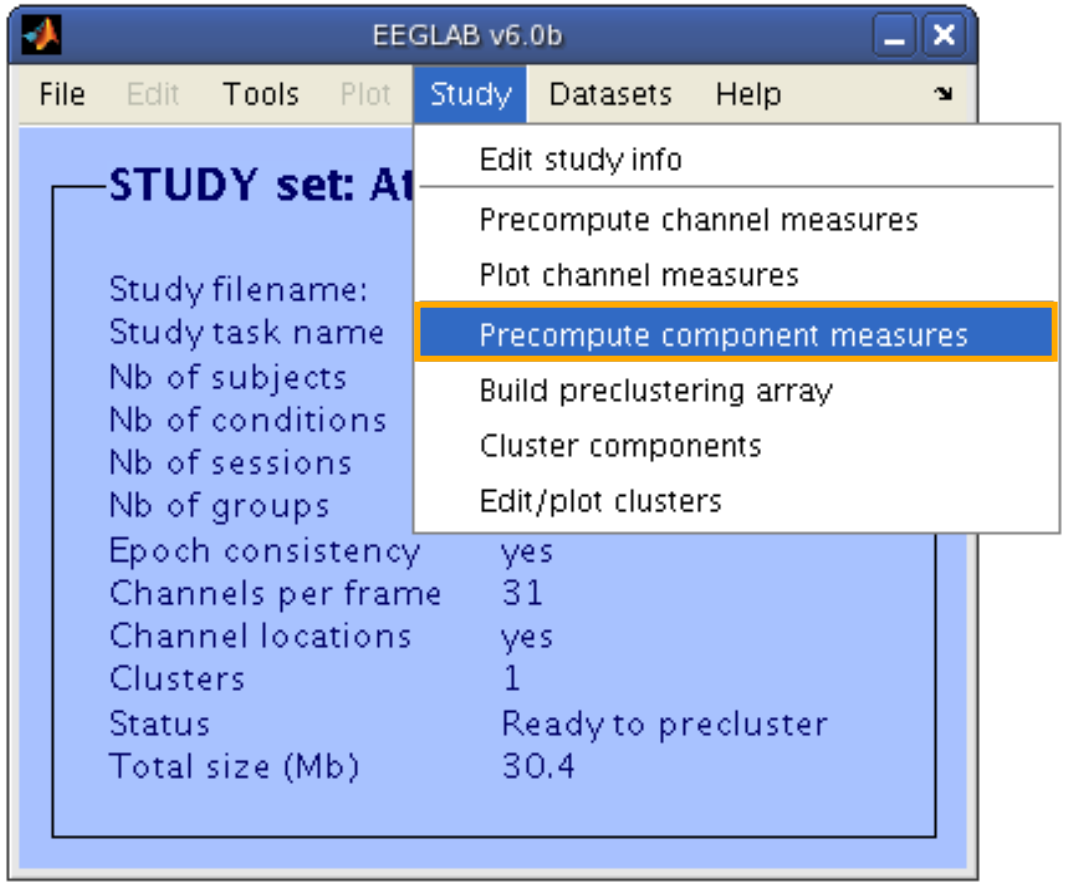
Help Cancel Ok

ICs to cluster

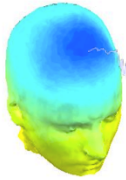


Precompute data measures

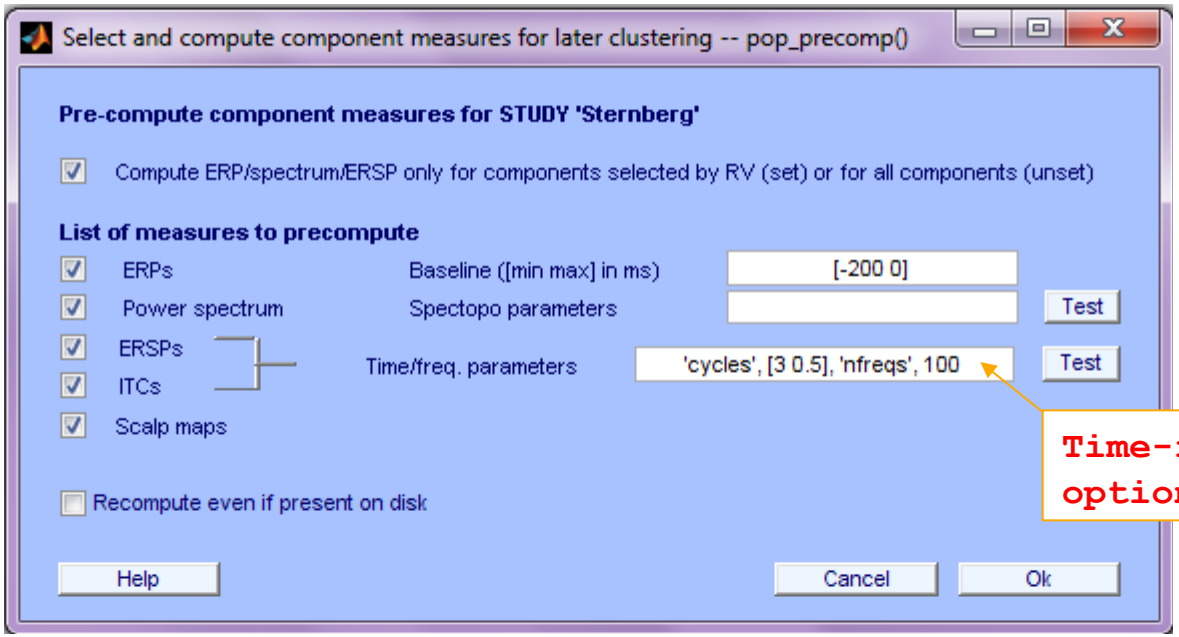




Precompute data measures

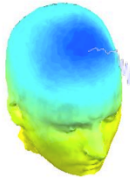


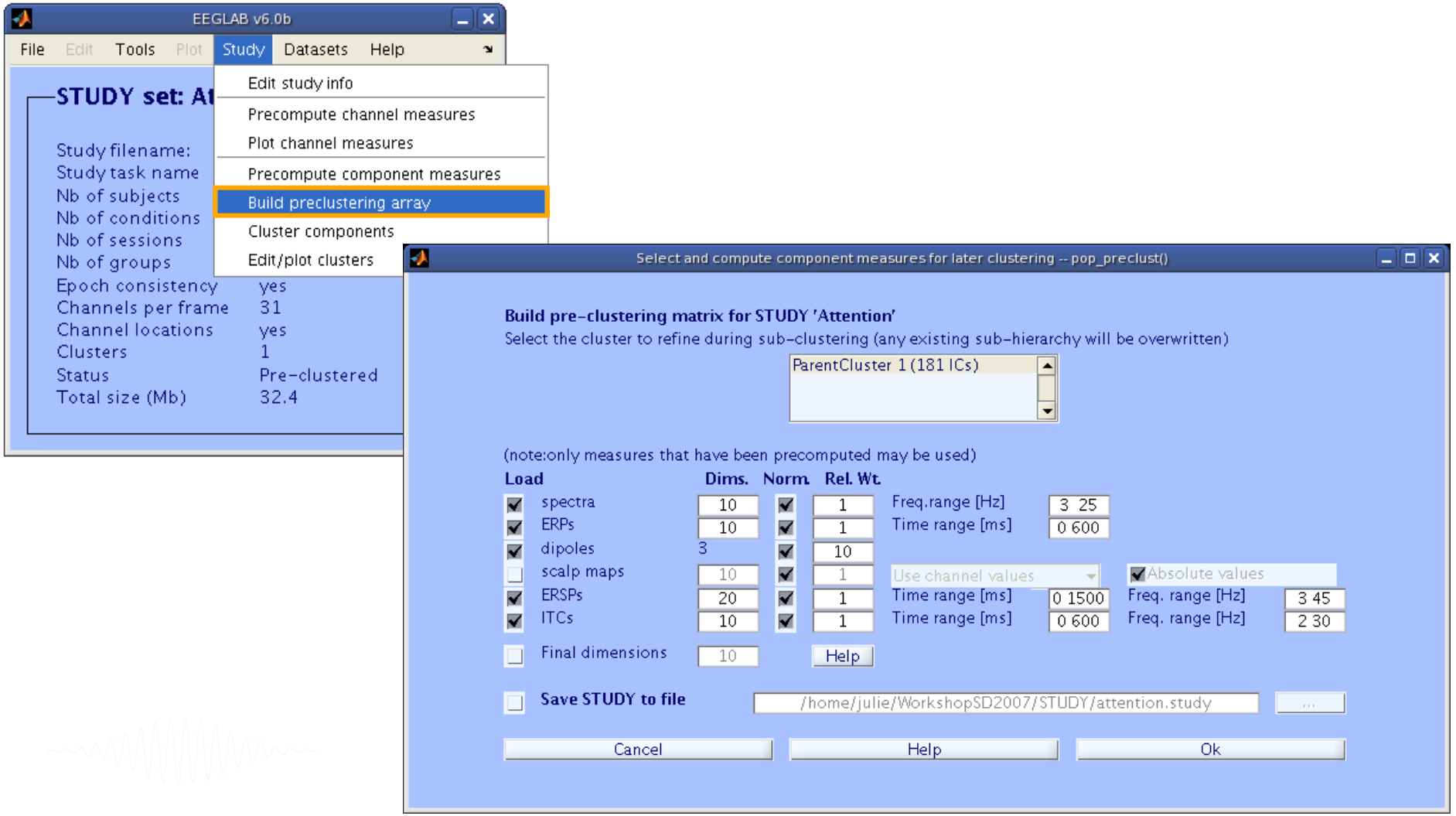
TIP: Compute all measures so you can test different combinations for clustering



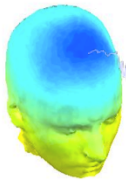
Time-frequency options

3. Cluster components





Precluster schematic



ERSP (time/freq)

Mean ERSPs

ICs (all subj)

PCA

ICs (all subj)

PC templates

PCs

* PC weights

Power spectrum

Mean spectra

ICs (all subj)

PCA

ICs (all subj)

PC templates

PCs

* PC weights

3D dipole position

ICs (all subj)

PC templates

x3

* PC weights

concatenate

concatenate

ICs (all subj)

ERSP

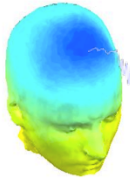
Spectrum

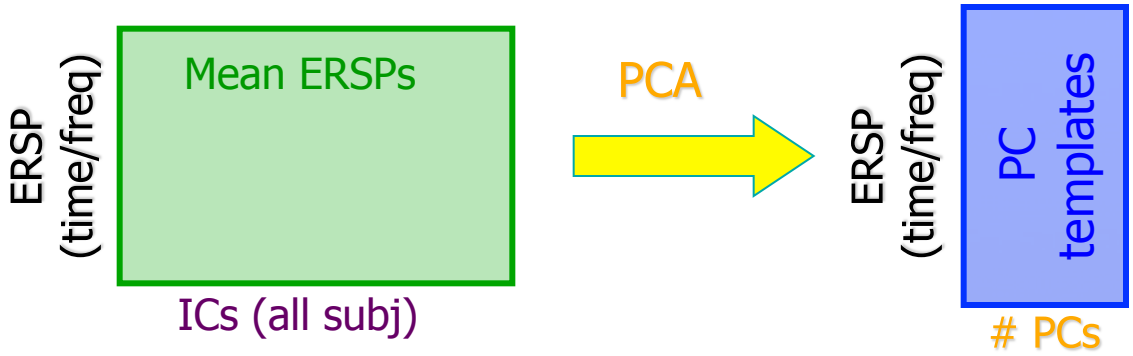
Dipoles

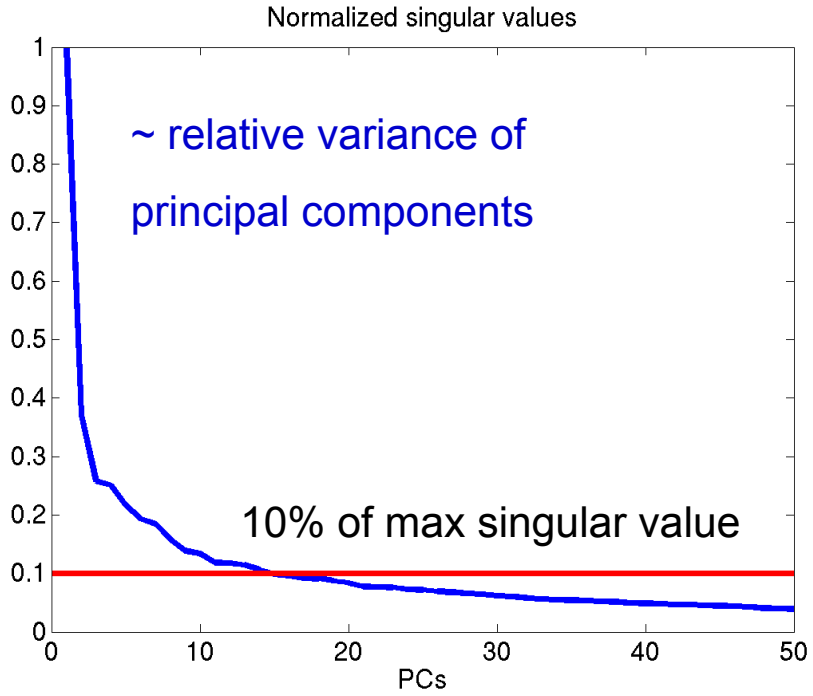
PCA

ICs (all subj)

Precluster: Use singular values from PCA

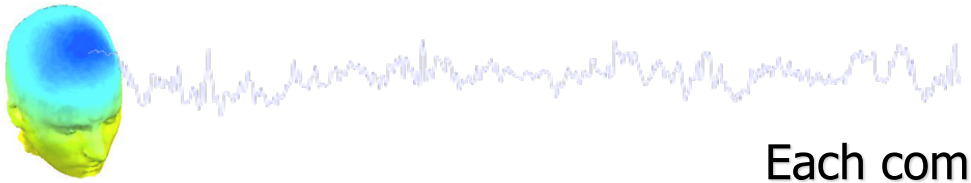






Credit: Julie Onton

Precluster schematic



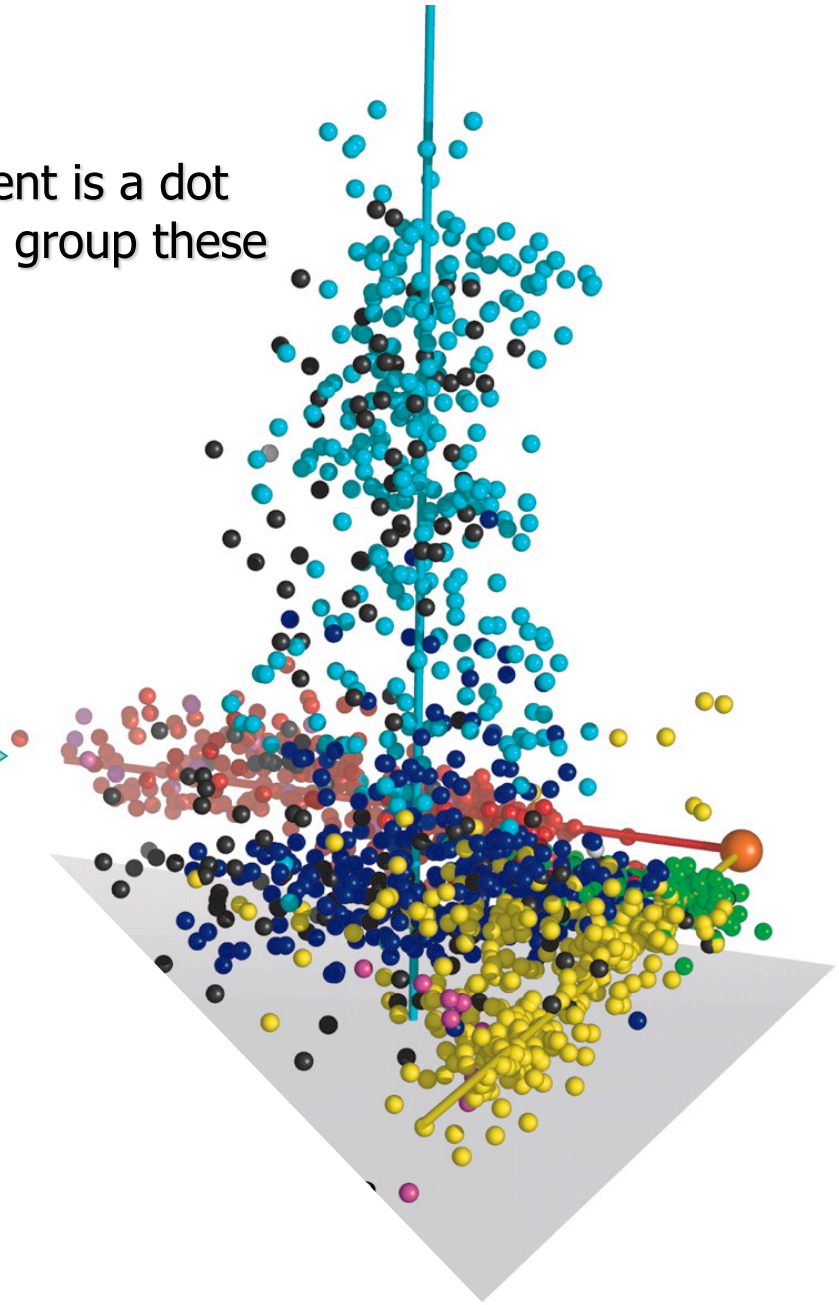
Each component is a dot
Clustering will group these dots

ICs (all subj)

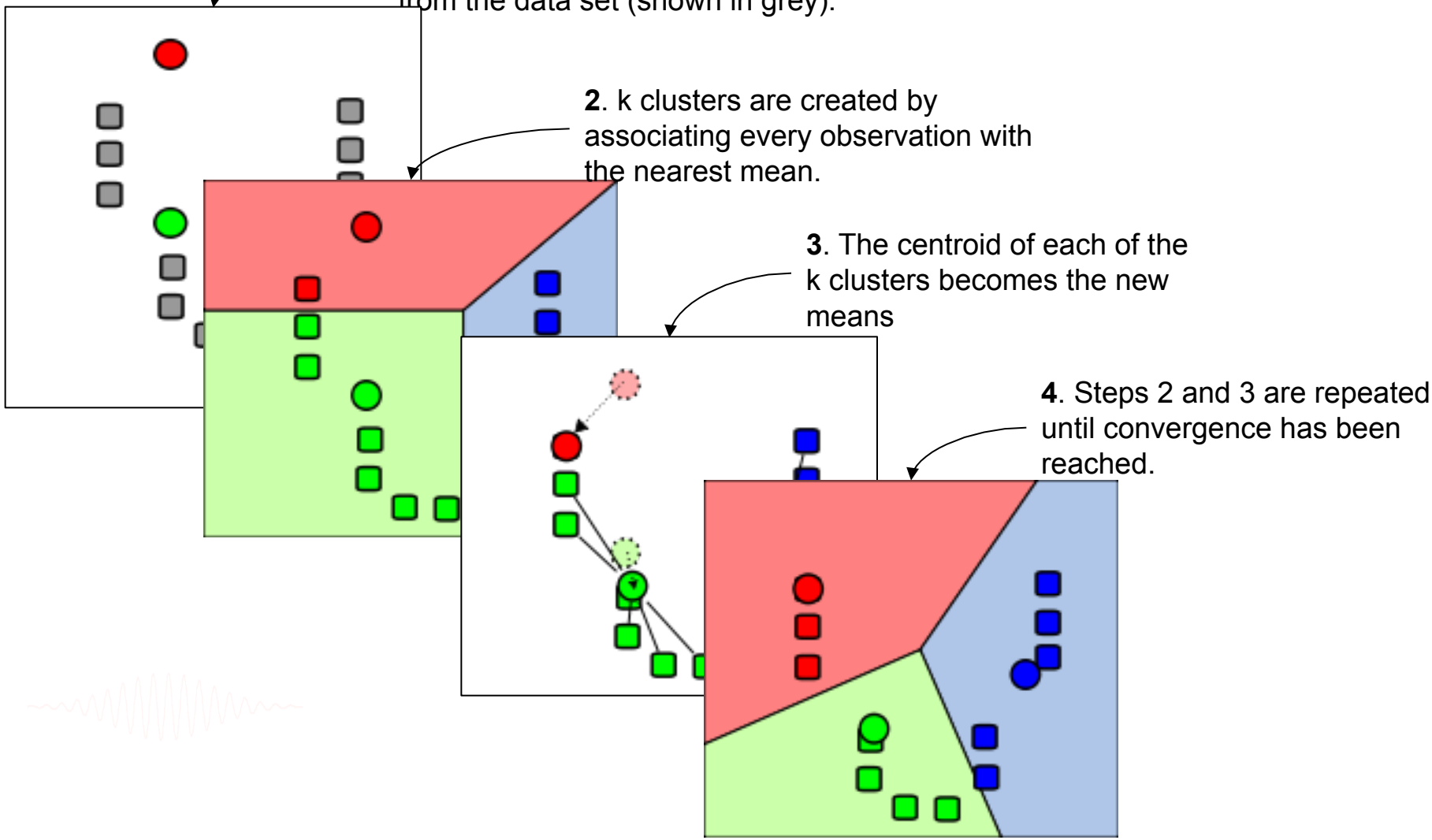
ERSP | Spectrum | Dipoles

OR

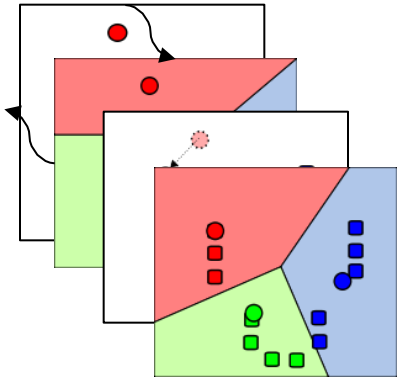
ICs (all subj)



Classical KMean



Customized KMean
(no more than 1 session per cluster)



1. A first KMean solution is computed for N clusters

2. Select the cluster with minimum residual distance to centroid

3. Keeps at most one component per session

(min dist. to centroid)

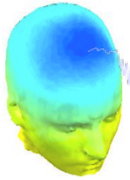
4. Store the resulting cluster

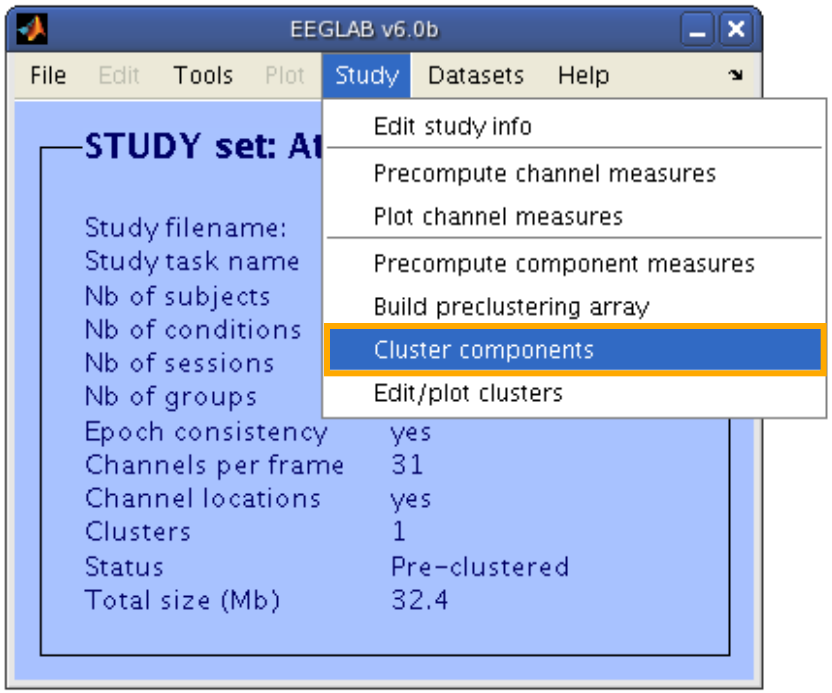
5. Remove the cluster's ICs from the pool of all ICs

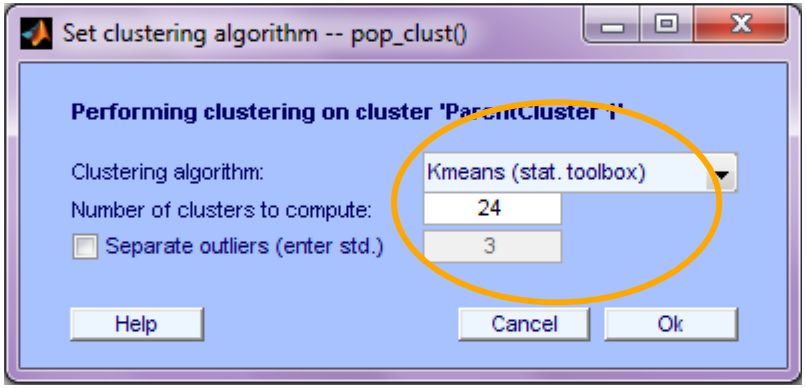
6. Compute a new KMean solution for N-1 clusters on the new pool

7. Loop until the desired number of selected clusters is reached

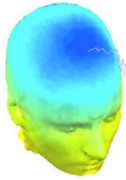
Cluster components







Choosing data measures

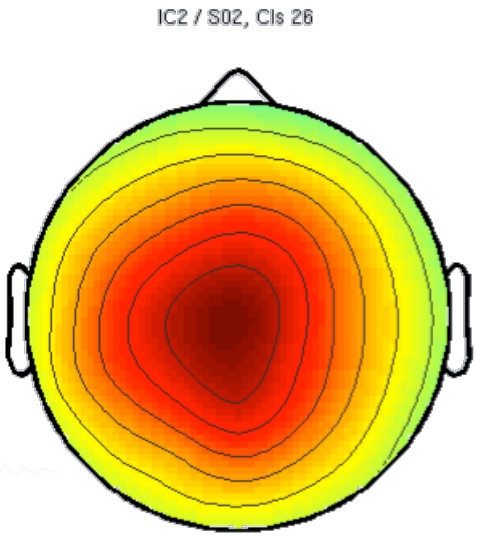


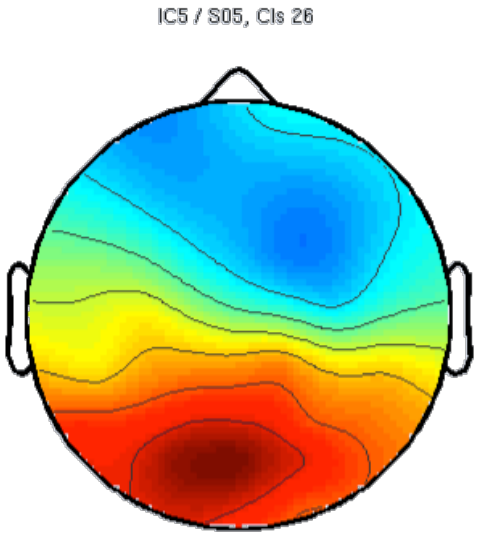
What measure(s) should you use?

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

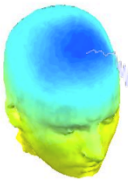
But consider:

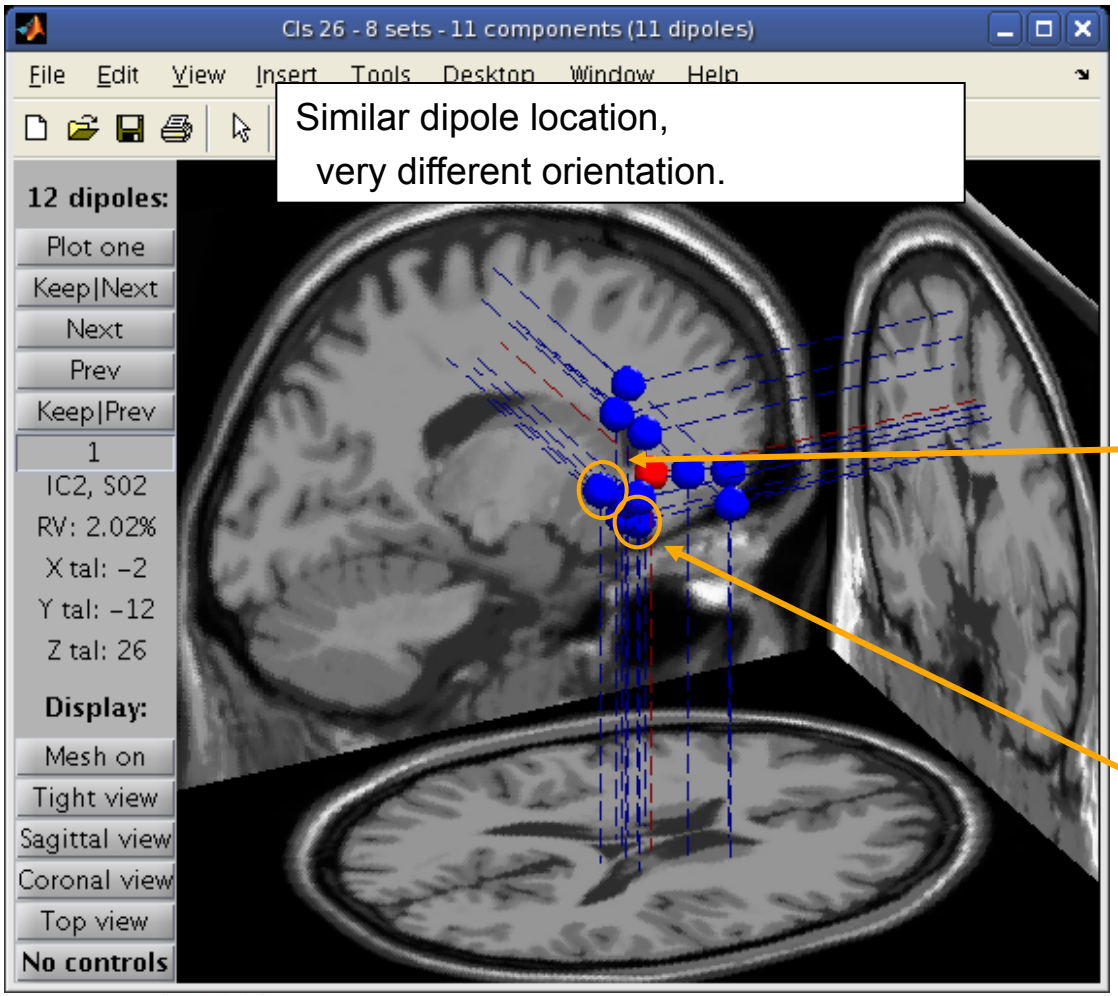
- What is the difference between these two components?





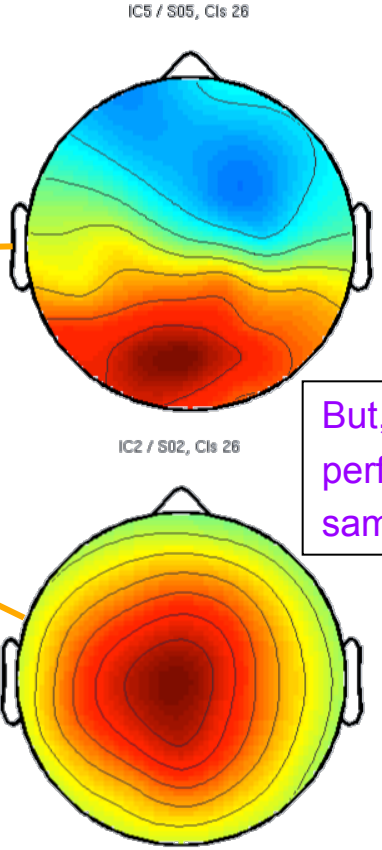
Choosing data measures





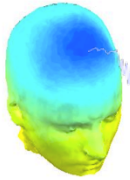
Similar dipole location,
very different orientation.

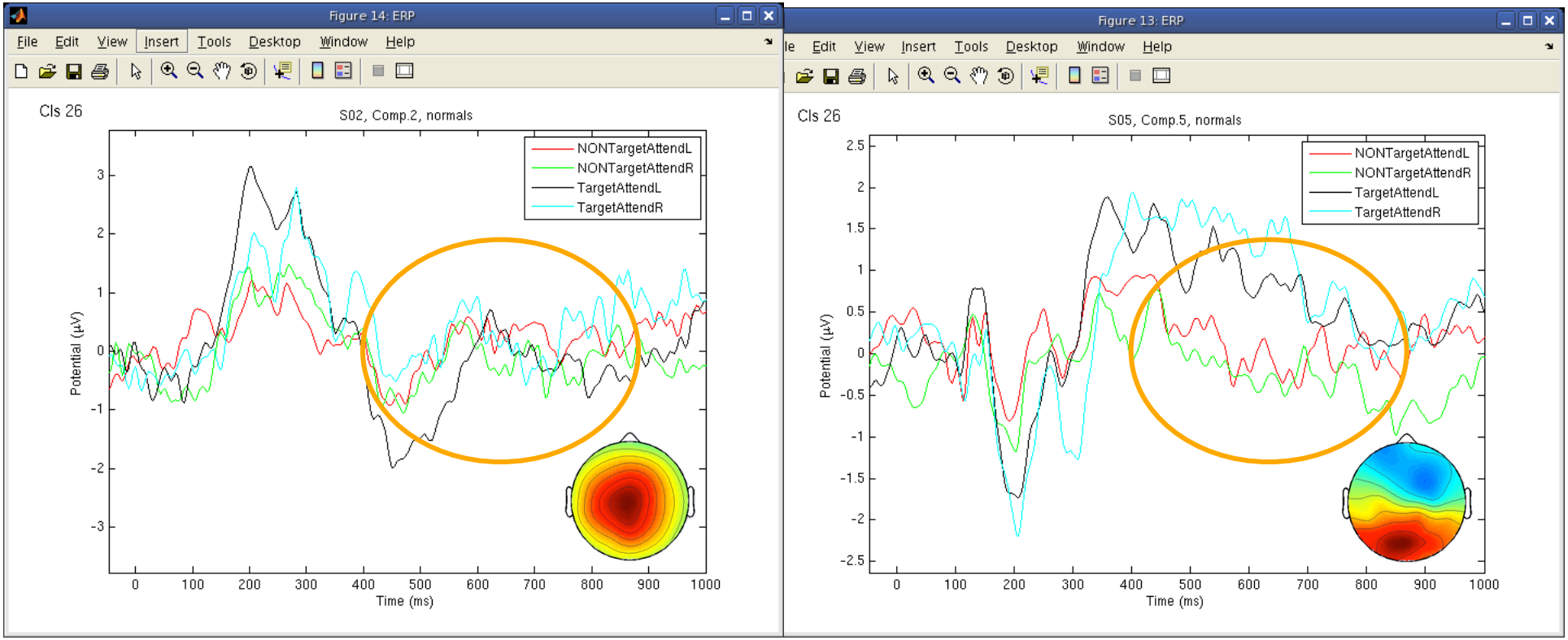
Obvious dramatic effect on scalp map topography:



But, do they perform the same functions?

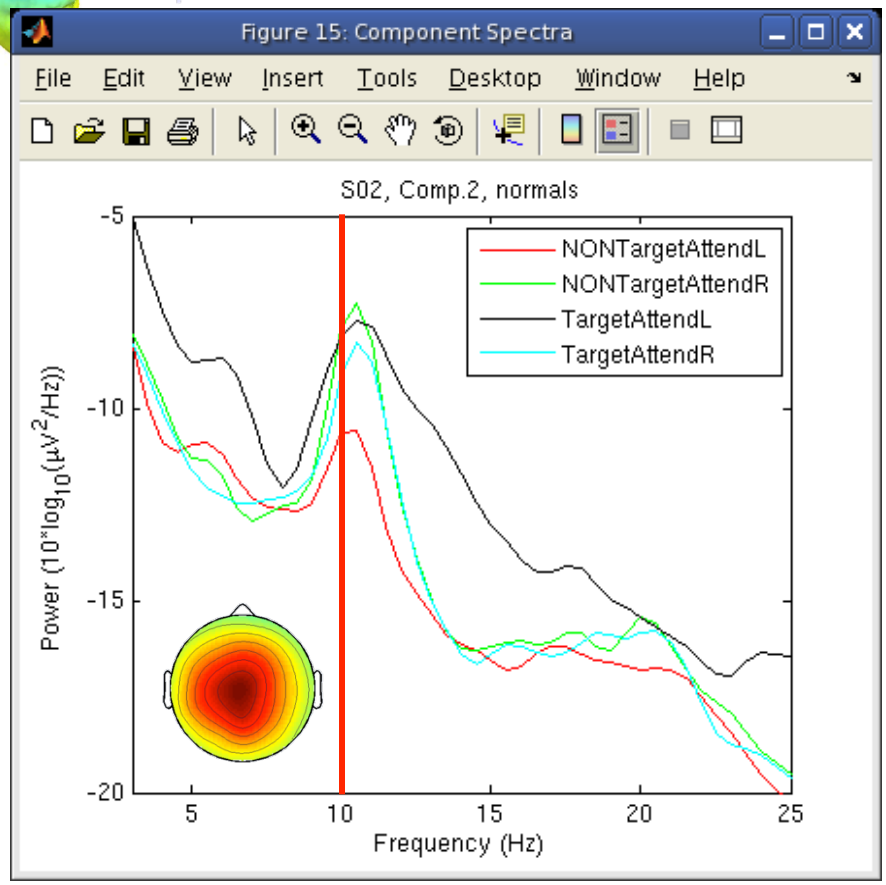
Choosing data measures

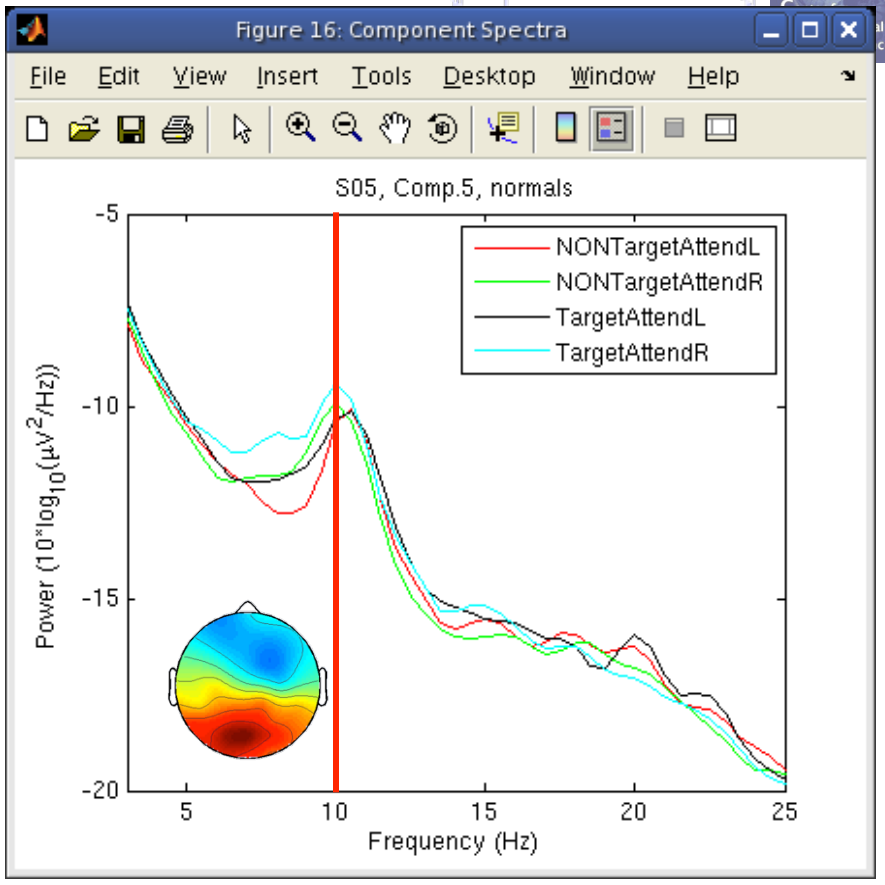




ERPs seem different…

Choosing data measures

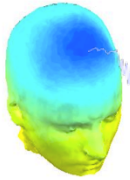


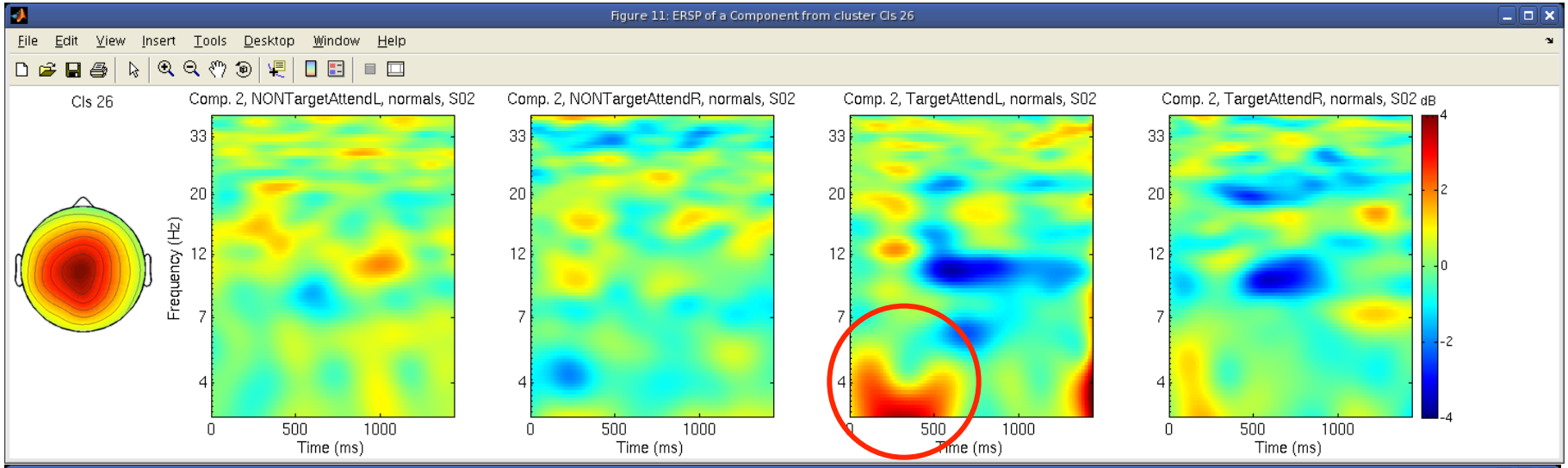


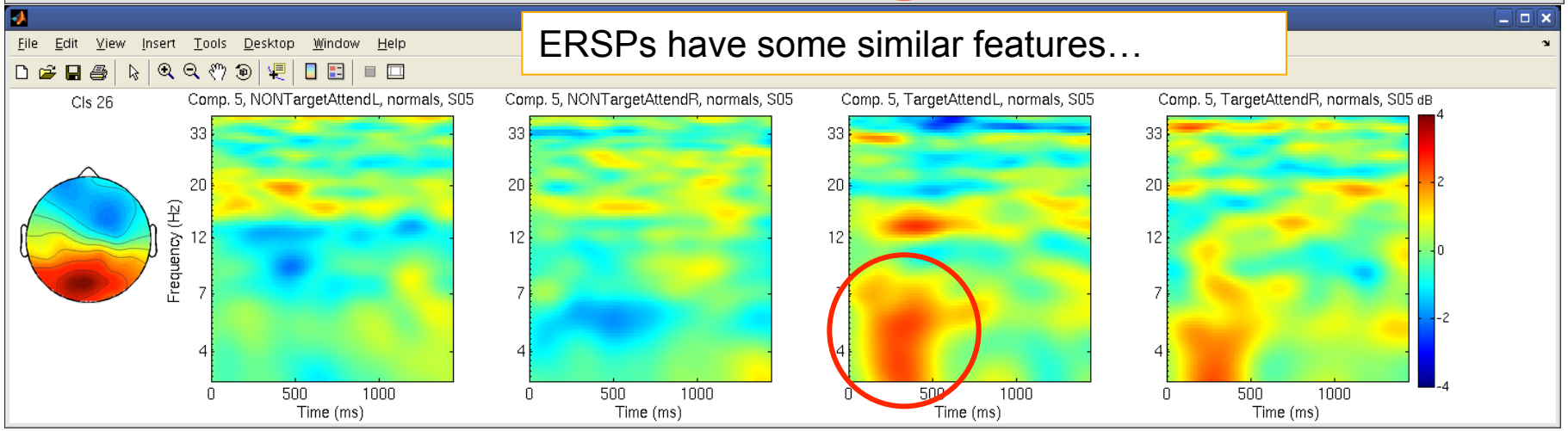
Spectra are similar, but they have

variable responses to different conditions…

Choosing data measures

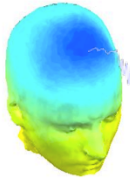


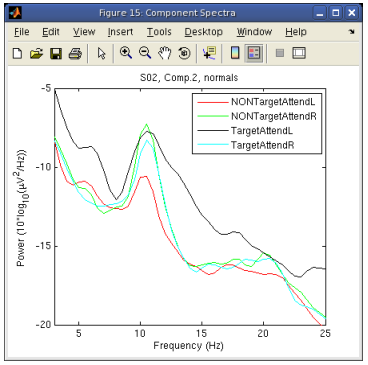


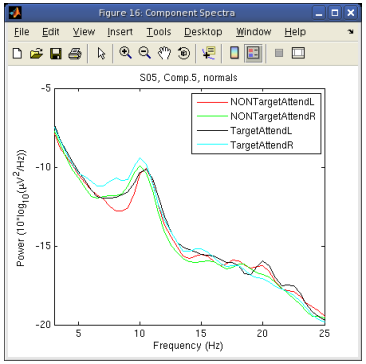


ERSPs have some similar features…

Choosing data measures



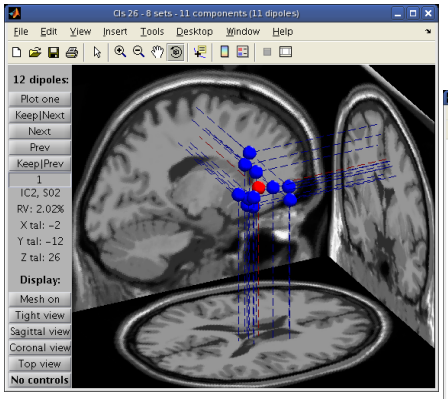


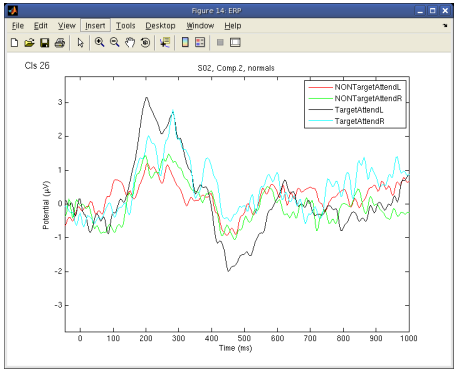


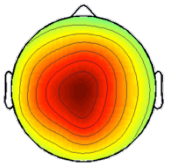
What data measures should you use?

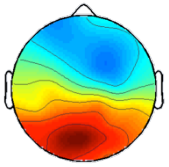
It depends…

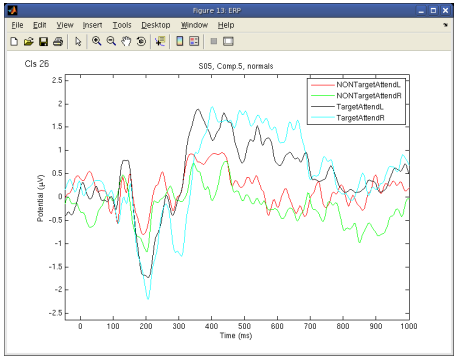
- broadly-matched ICs: use many/all of the measures.
- specifically-matched ICs: use one/few of the measures.

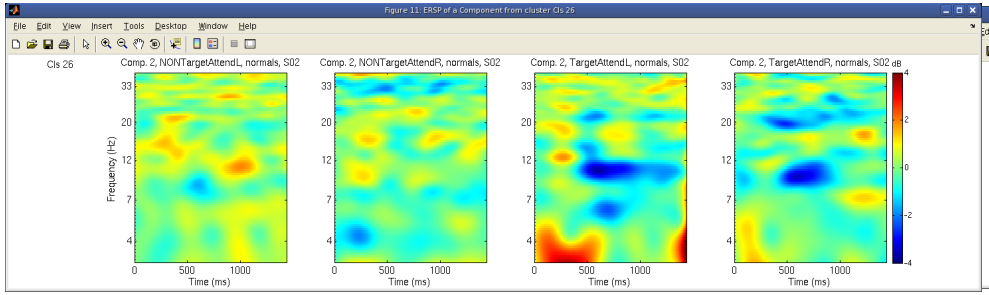


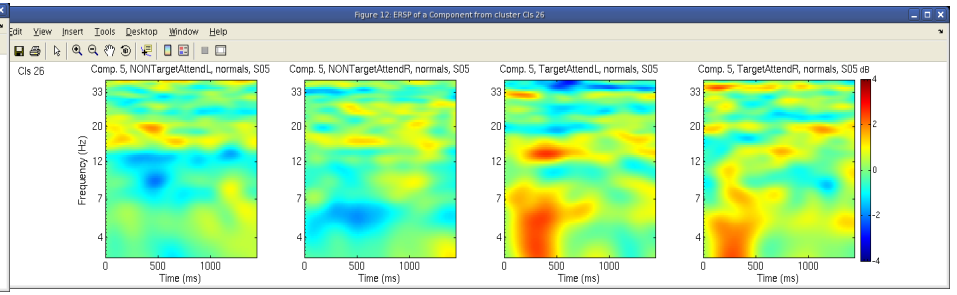




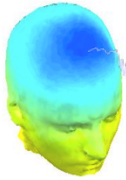


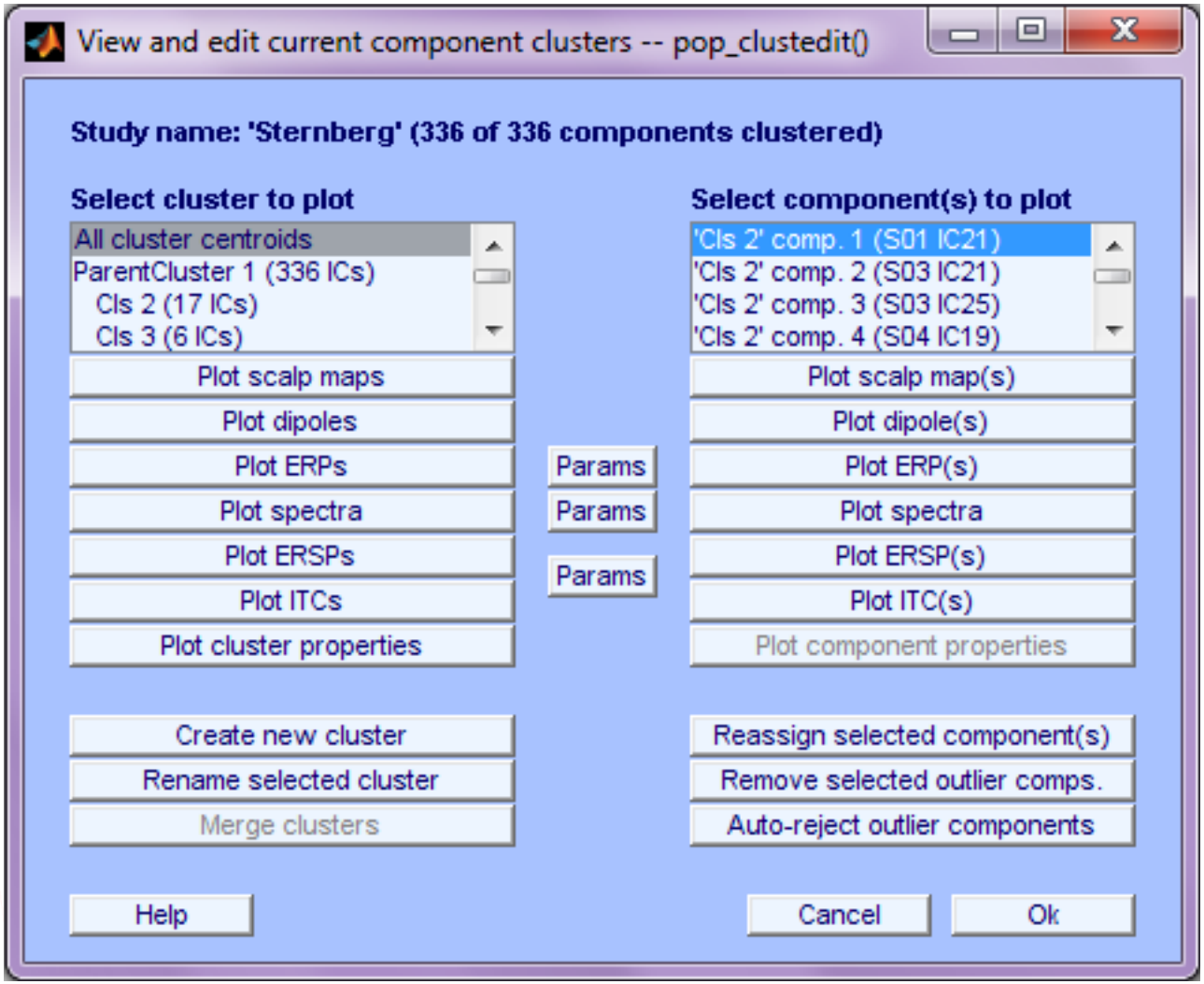




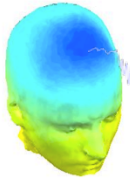


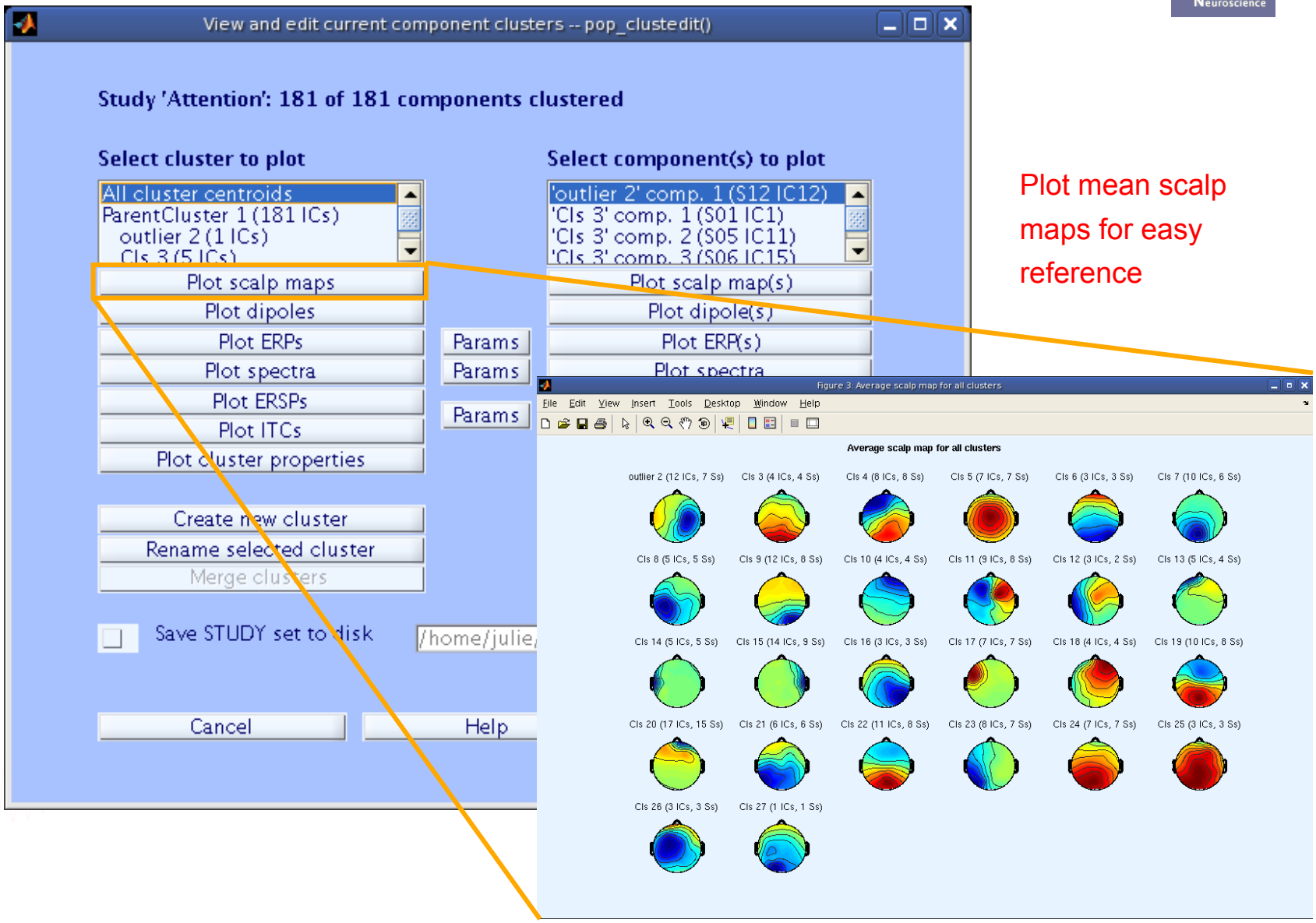
Plot/edit clusters





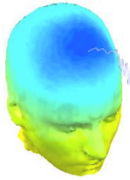
Plot cluster data





Plot mean scalp maps for easy reference

Plot cluster data



Swartz Center for Computational Neuroscience

Choose which cluster

Choose which components

View and edit current component clusters -- pop_clustedit()

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)

Select component(s) to plot

All components
S01 IC6
S05 IC9
S06 IC12

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

Params
Params
Params

Plot scalp map(s)

Create new cluster
Rename selected cluster
Merge clusters

Save STUDY set to disk

/home/julie/WorkshopSD2

Cancel
Help

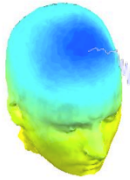
Figure 4

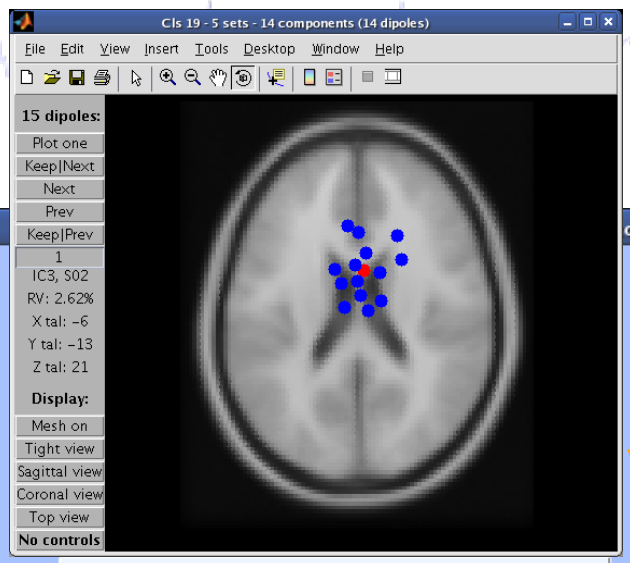
File Edit View Insert Tools Desktop Window Help

Cls 21 (7 ICs, 7 Ss)

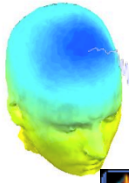
ic4/S01
ic11/S02
ic1/S06
ic1/S08
ic2/S10
ic6/S12
ic3/S15

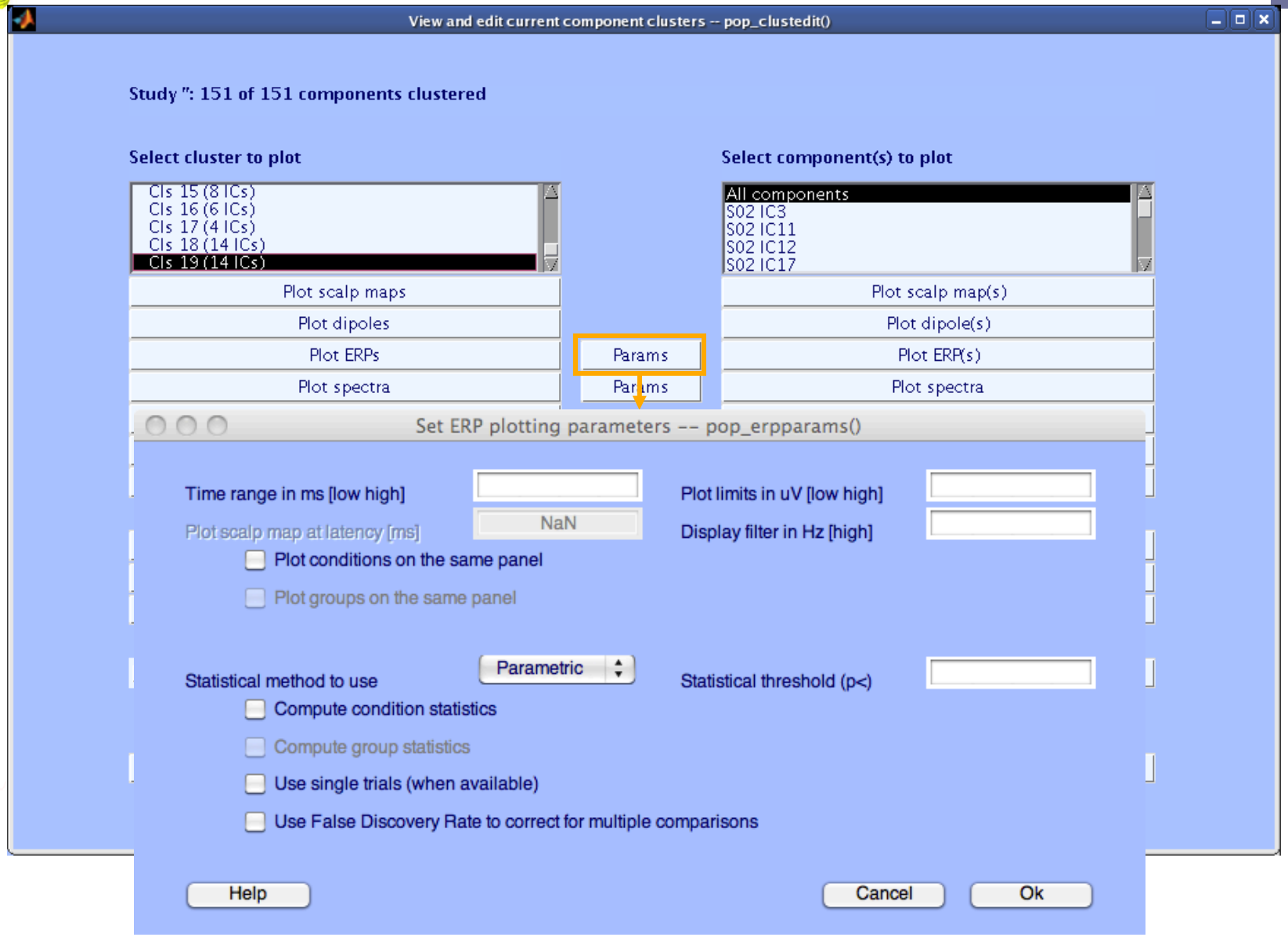
Plot cluster data



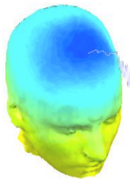


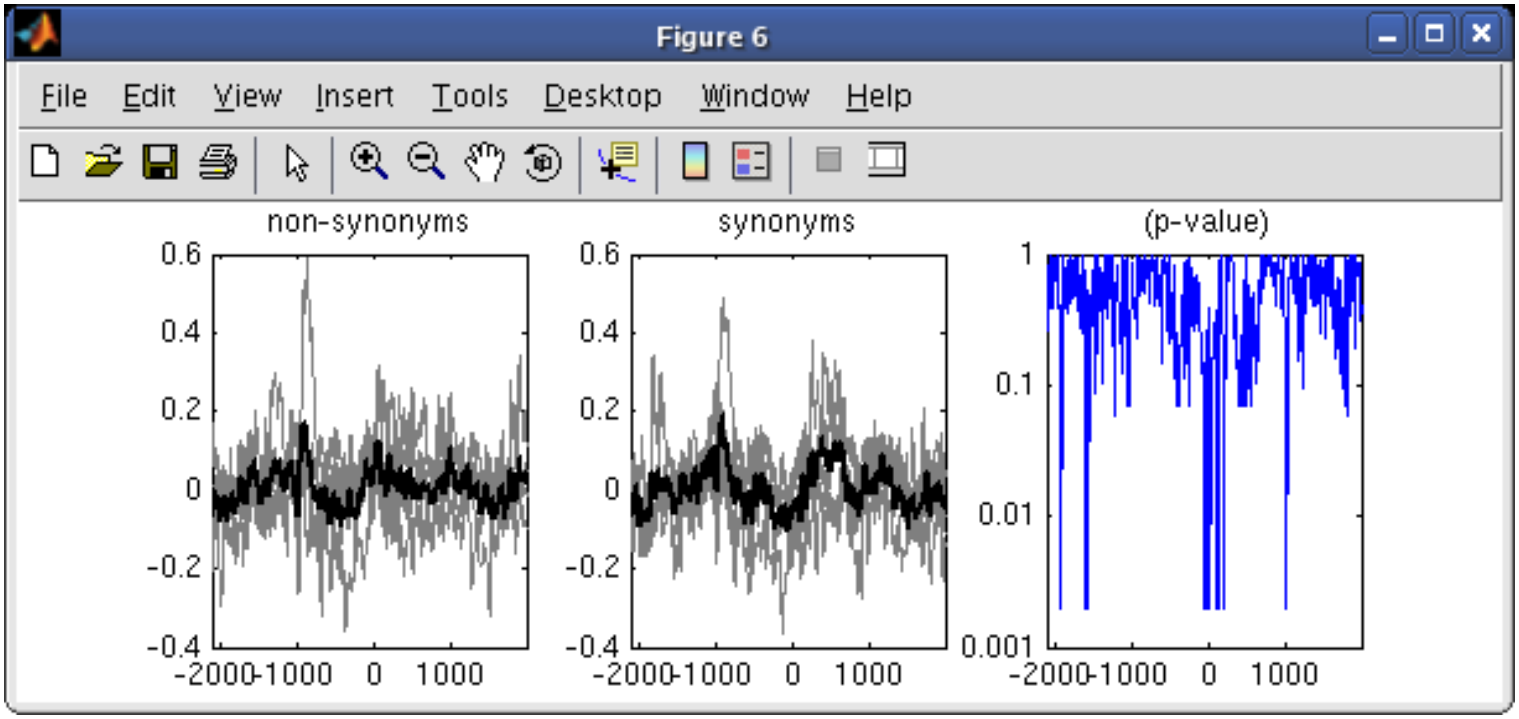
Plot cluster data



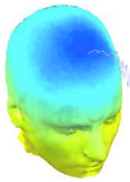


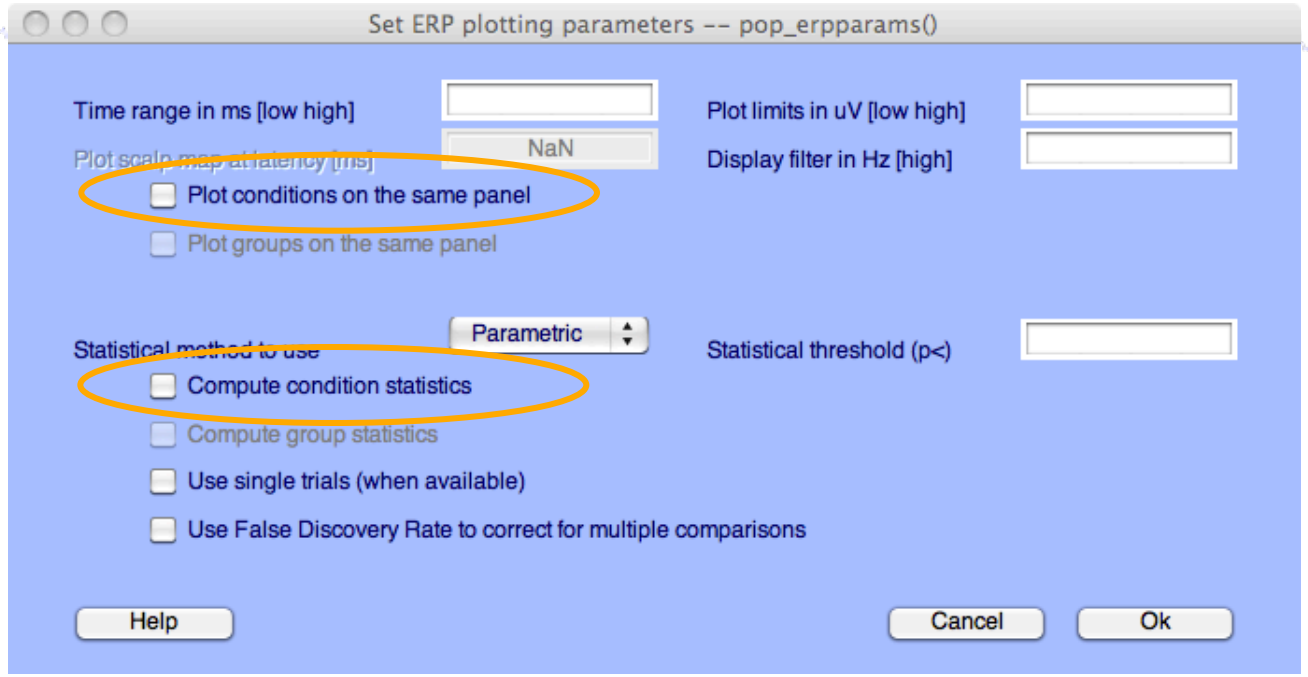
Plot cluster ERP

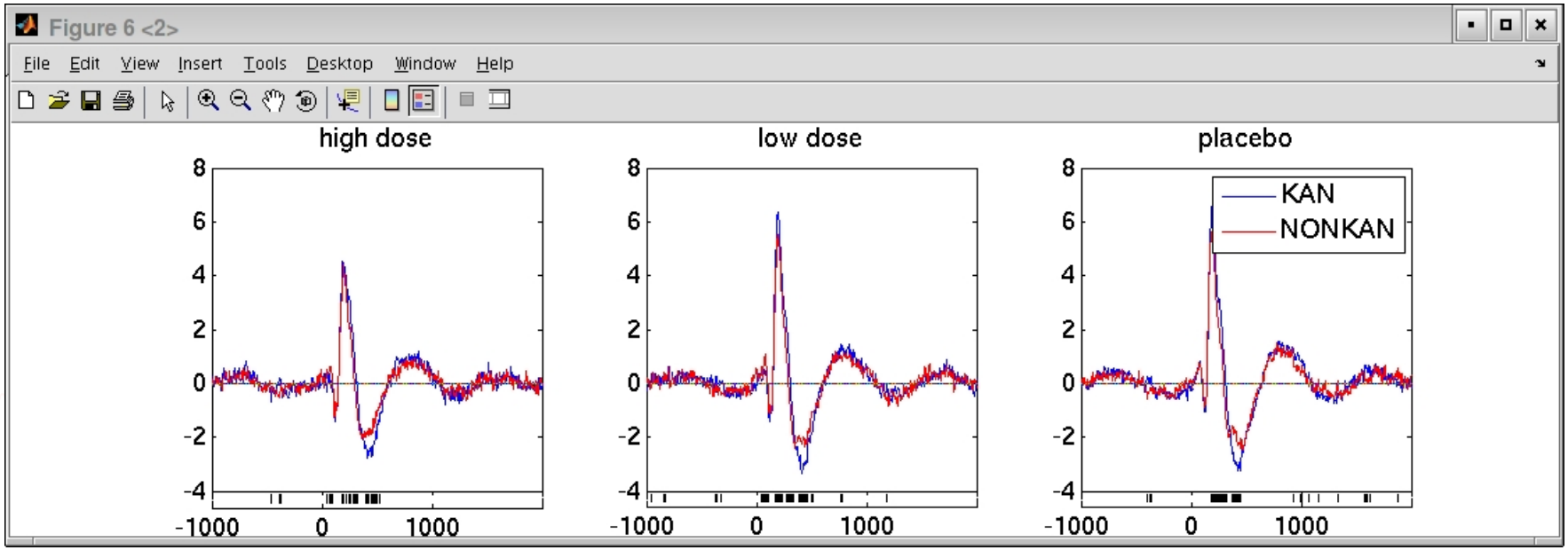




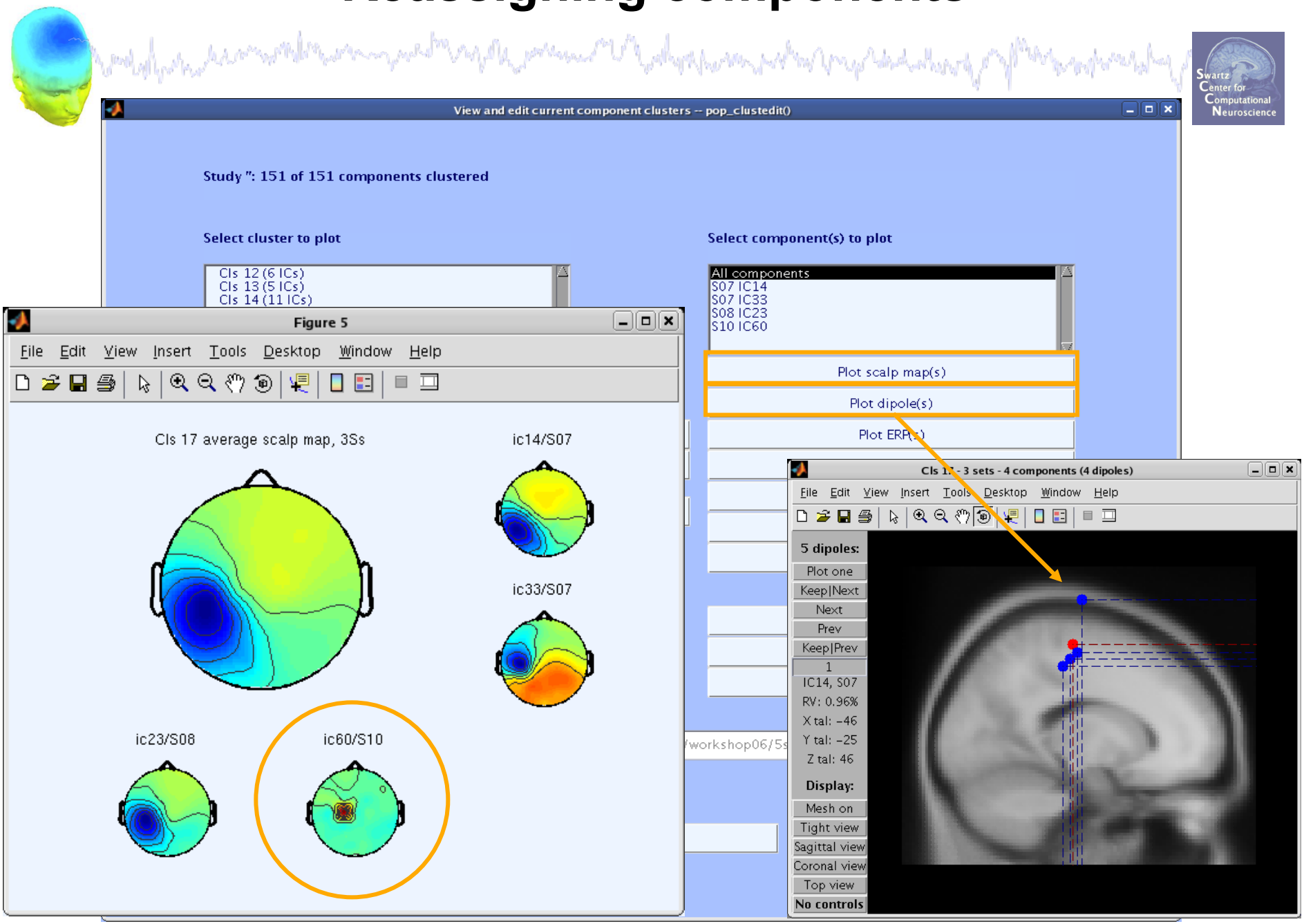
Other plotting options…



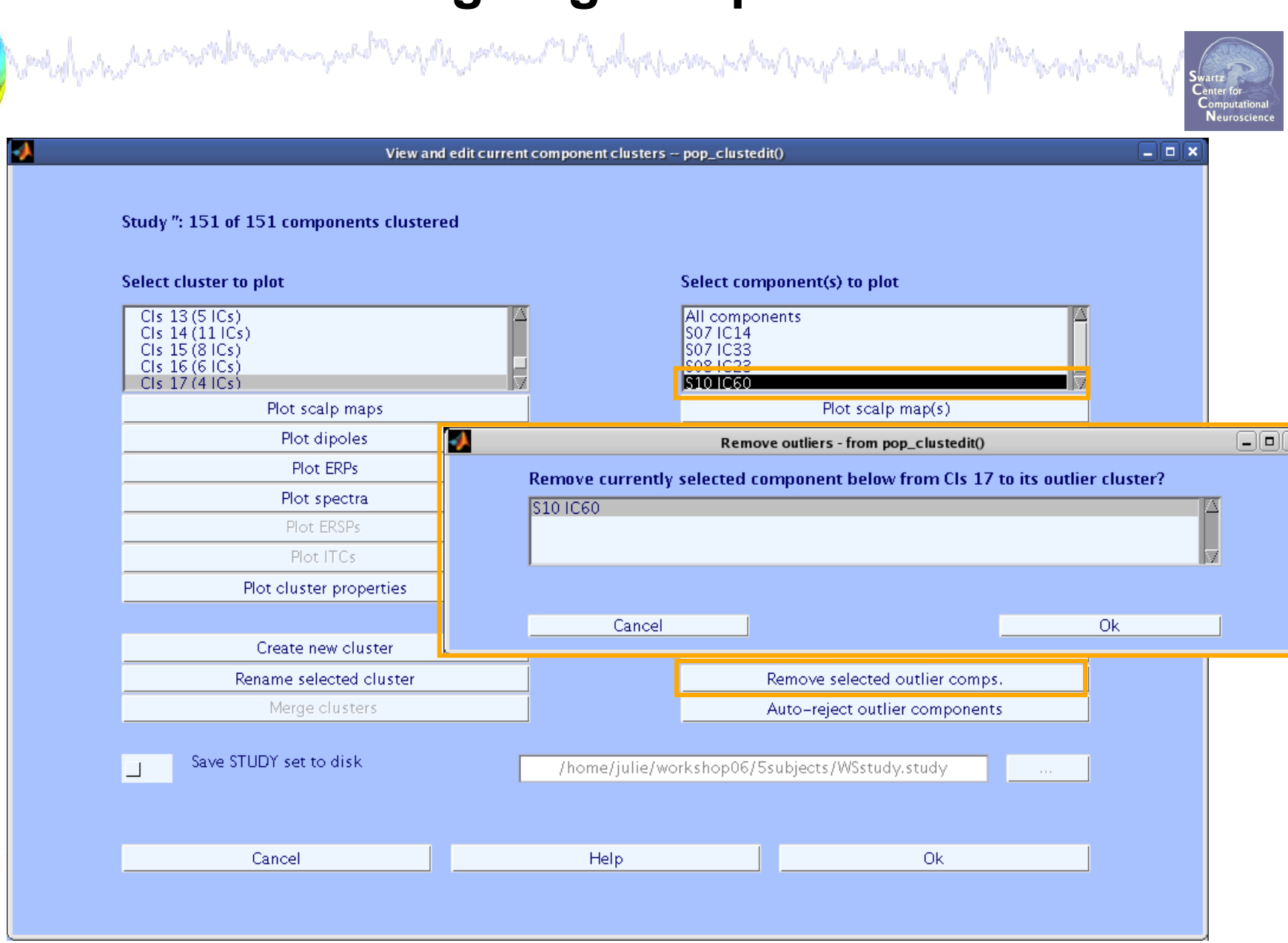




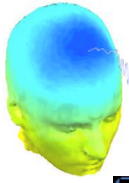
Reassigning components

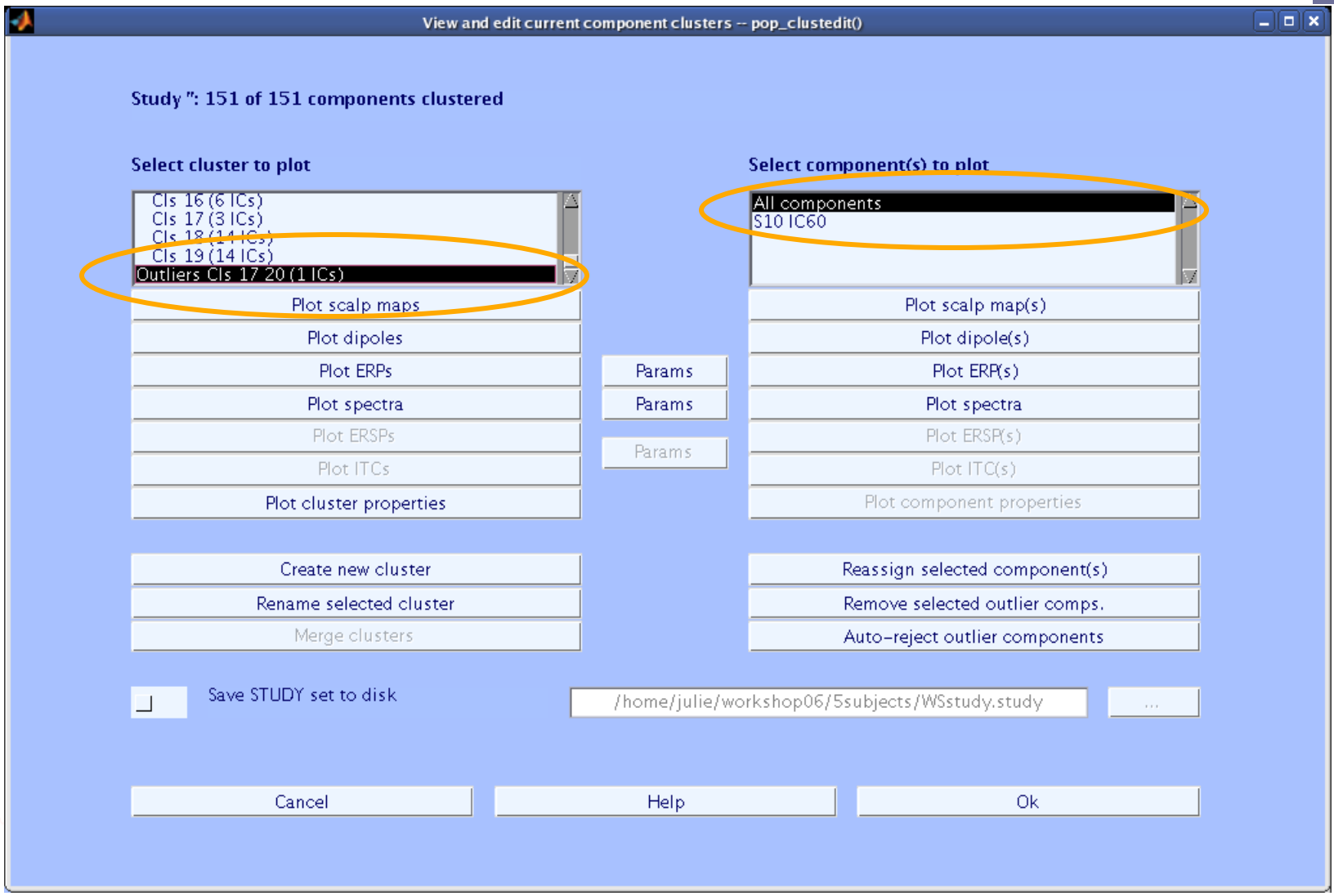


Reassigning components

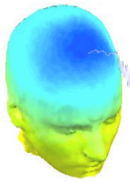


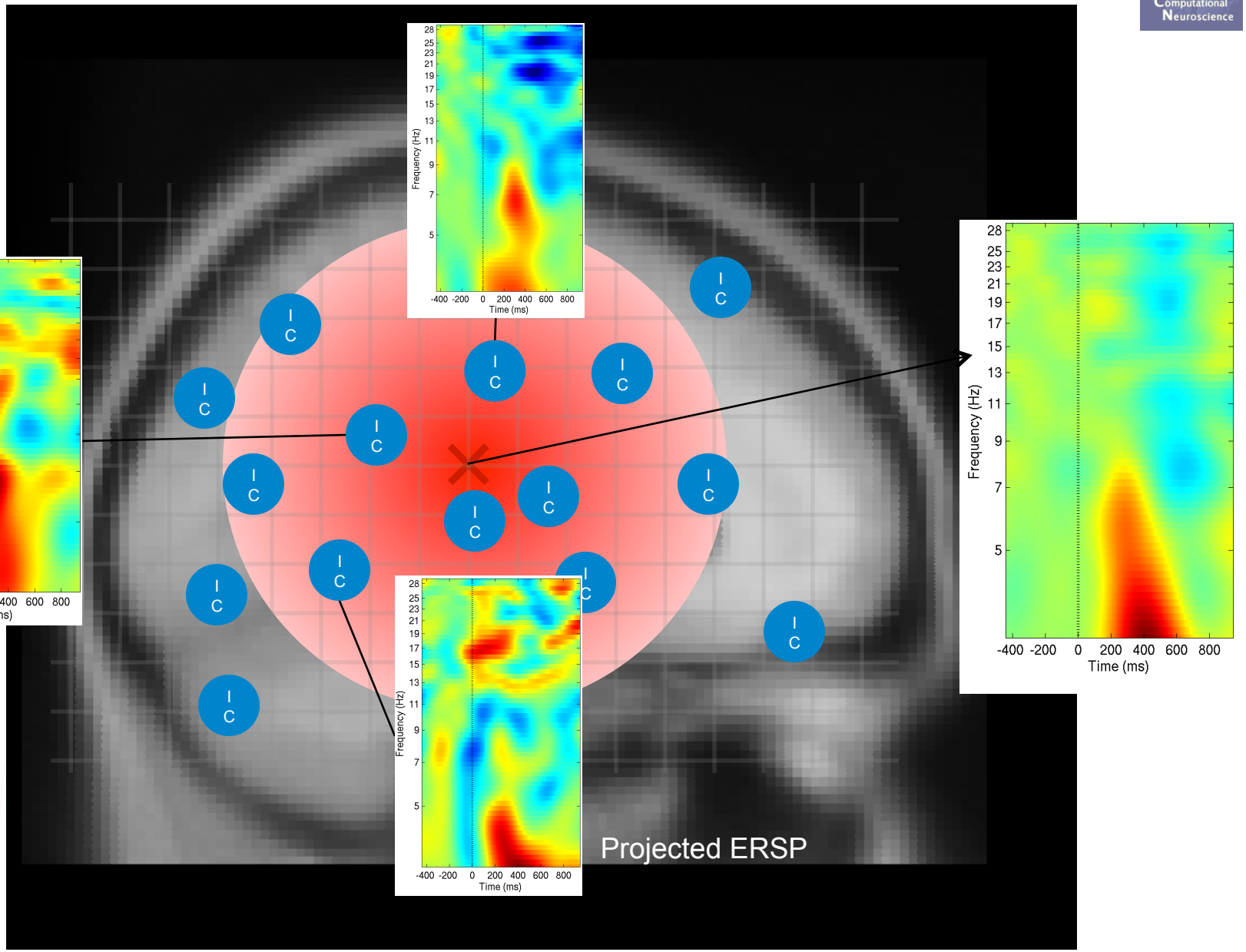
Outlier cluster reassignment



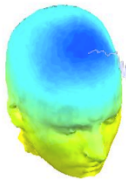


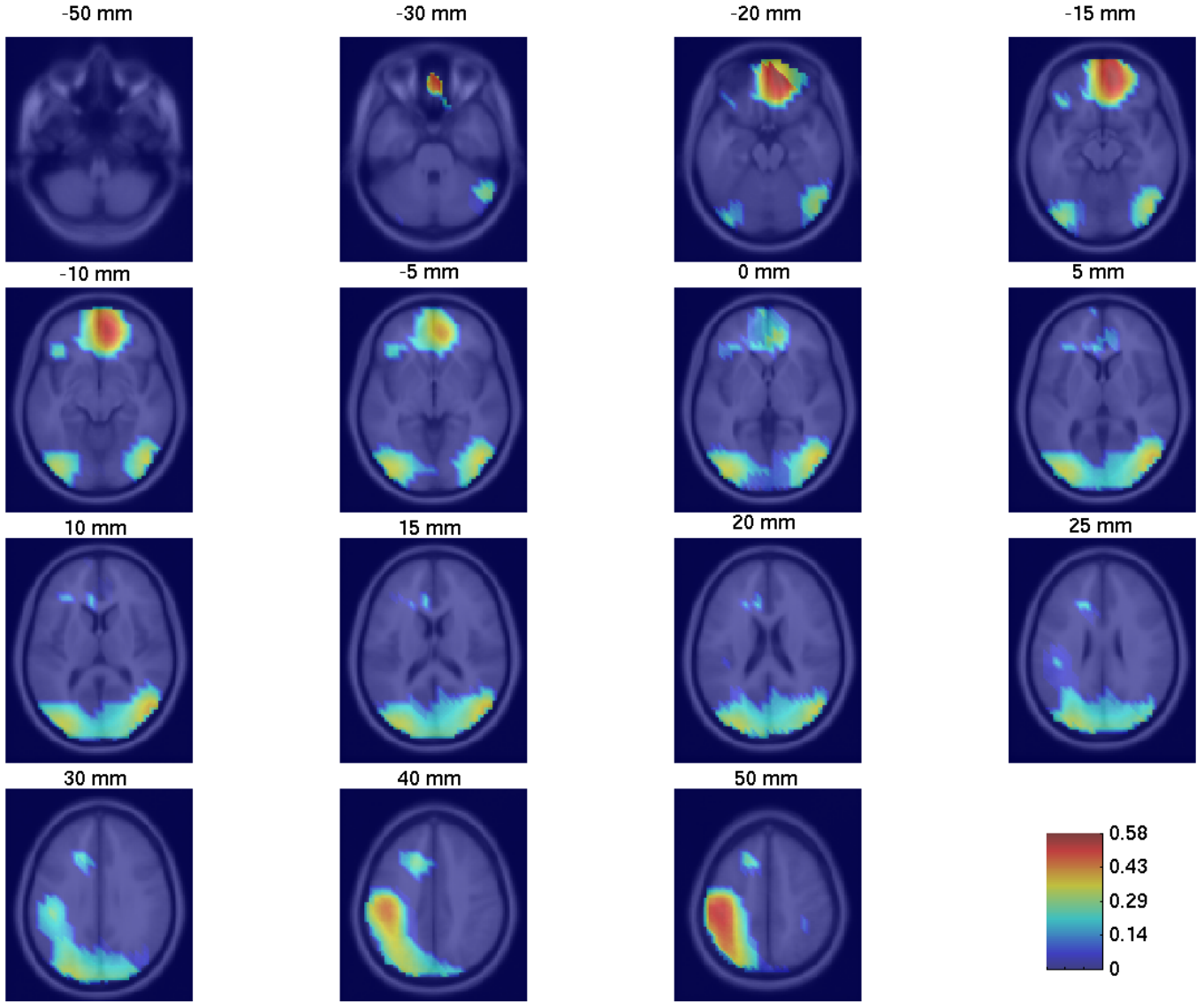
Measure Projection Toolbox



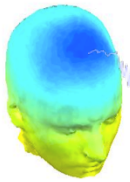


Measure Projection Toolbox





Swartz Center for Computational Neuroscience



Exercise

Precluster (pre-computation already done) and cluster components using measures of your choice. Experiment with different measures.