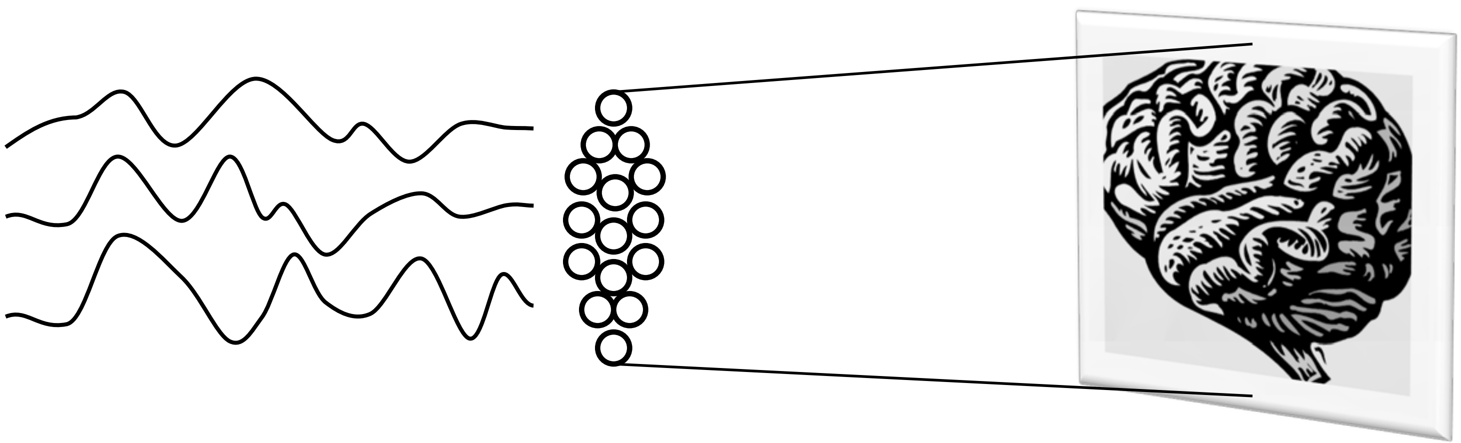
Measure Projection Analysis

or: how I learned to stop clustering and love the grid.



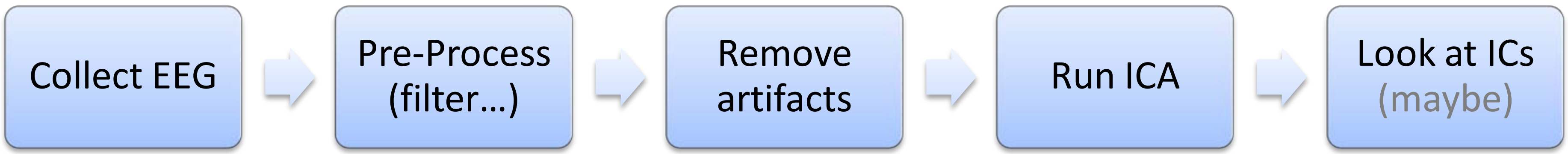
Nima Bigdely-Shamlo, Tim Mullen, Ozgur Yigit Balkan
Swartz Center for Computational Neuroscience
INC, UCSD, 2010

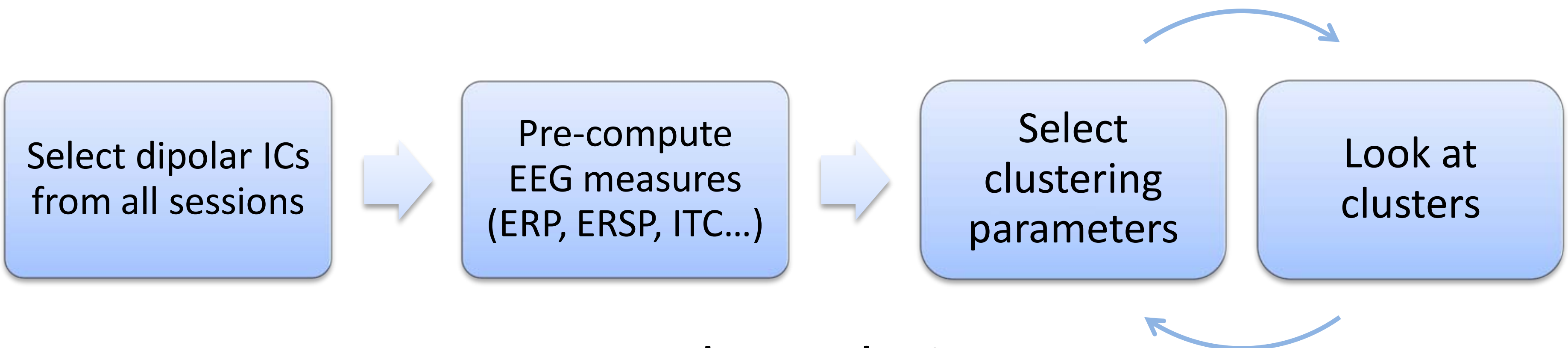
Outline

1. Current EEGLAB Workflow
2. STUDY IC Clustering
3. Conceptual Problems of IC Clustering
4. Practical issues of IC Clustering
5. Introduce Measure Projection method
6. Show its application on a study
7. Introduce MPA Beta Demo EEGLAB GUI

Current EEGLAB Workflow

Single EEG Session Analysis





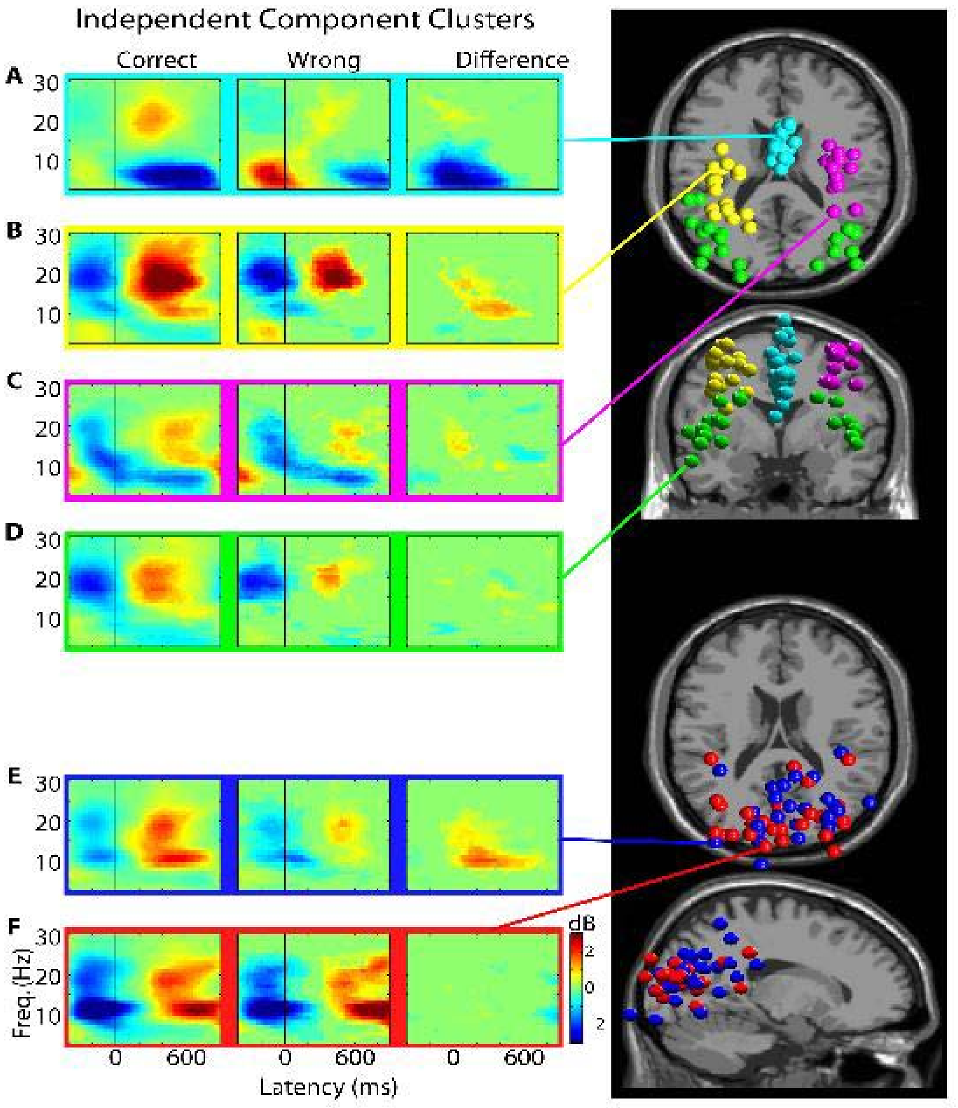
Study Analysis

Until **expected** results are produced

Study IC Clustering

- Assume there are *Functionally Equivalent* ICs across most subjects.
- Assumes these ICs have *Similar Responses* to experimental conditions across all measures (ERP, ERSP, ITC…)
- Creates *Non-Overlapping partitions*: each IC belongs only to one cluster

Study IC Clustering



Sometime clusters are spatially separate AND have distinct responses.

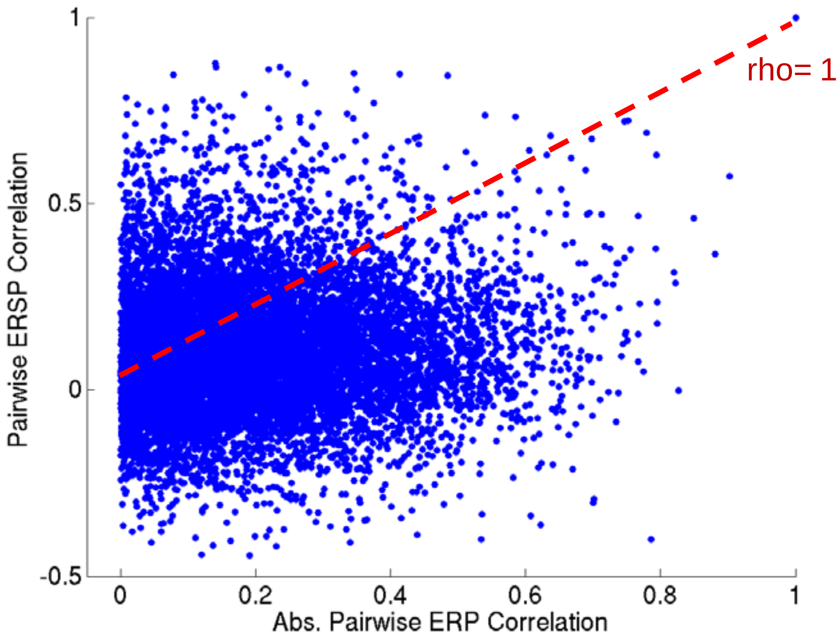
In other cases, they have similar responses or they overlap spatially.

Conceptual Problems with Study IC Clustering

1. Condition responses in the same brain area may significantly differ across subject groups. (often the goal of the study is finding these differences)
2. Components may have similar responses for one measure (e.g. ERSP) but not for the other (e.g. ERP). This is one of the worst issues.
3. *Boosts* evidence by rejecting ICs that are in the same brain area but show different responses. This makes calculating significance values extremely hard.
   How can we be sure that we are not *imagining* clusters?

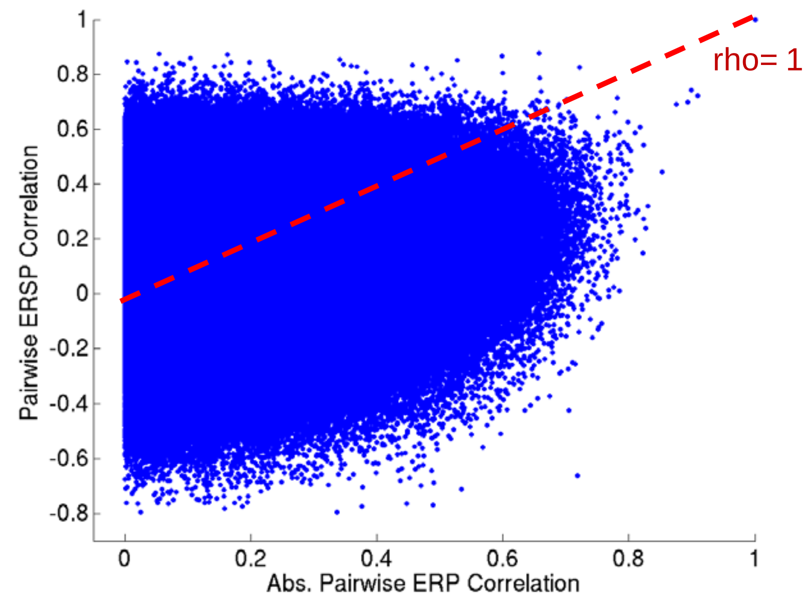
ERP and ERSP have different similarity structures

EEGLAB sample (N400) study with 5 subjects (151 ICs)



Correlation between correlations = **0.21**

ADHD study with 132 subjects (3477 ICs)

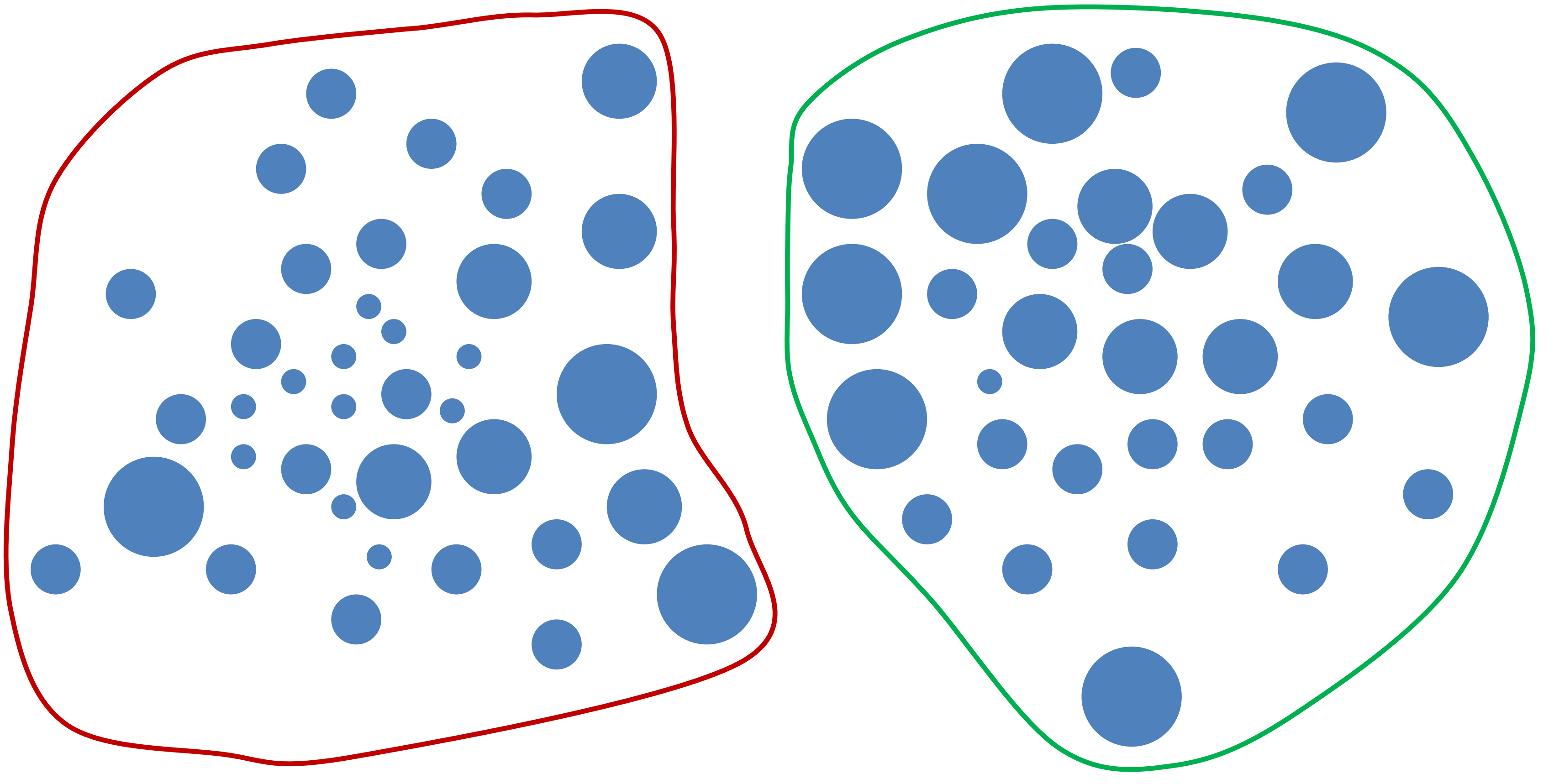


Correlation between correlations = **0.11**

This shows that ERSP and ERP generally have quite different similarity structures and it is probably not a good idea to combine them for clustering.

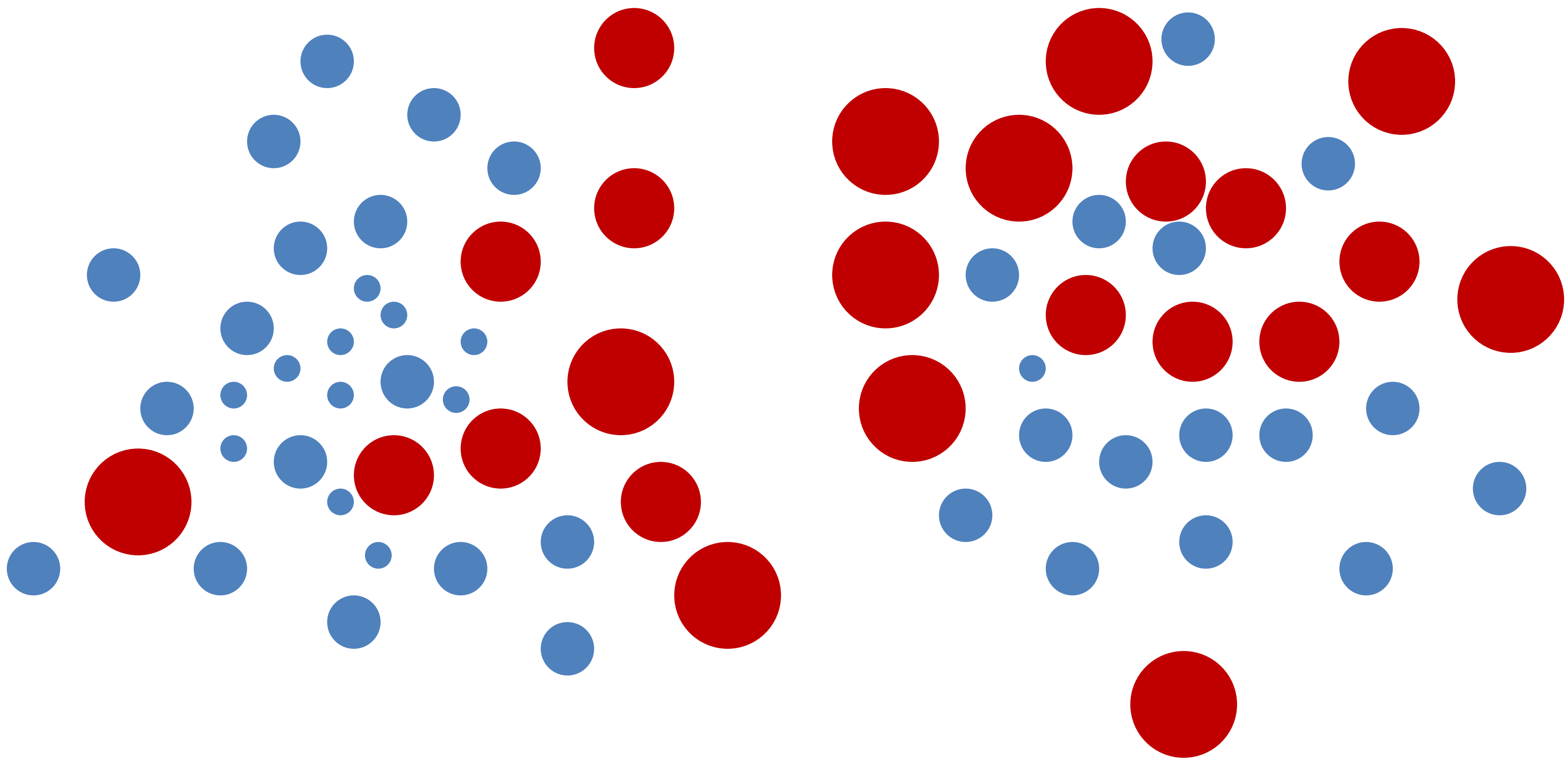
Problems with multi-measure clustering

What are the clusters according to location?



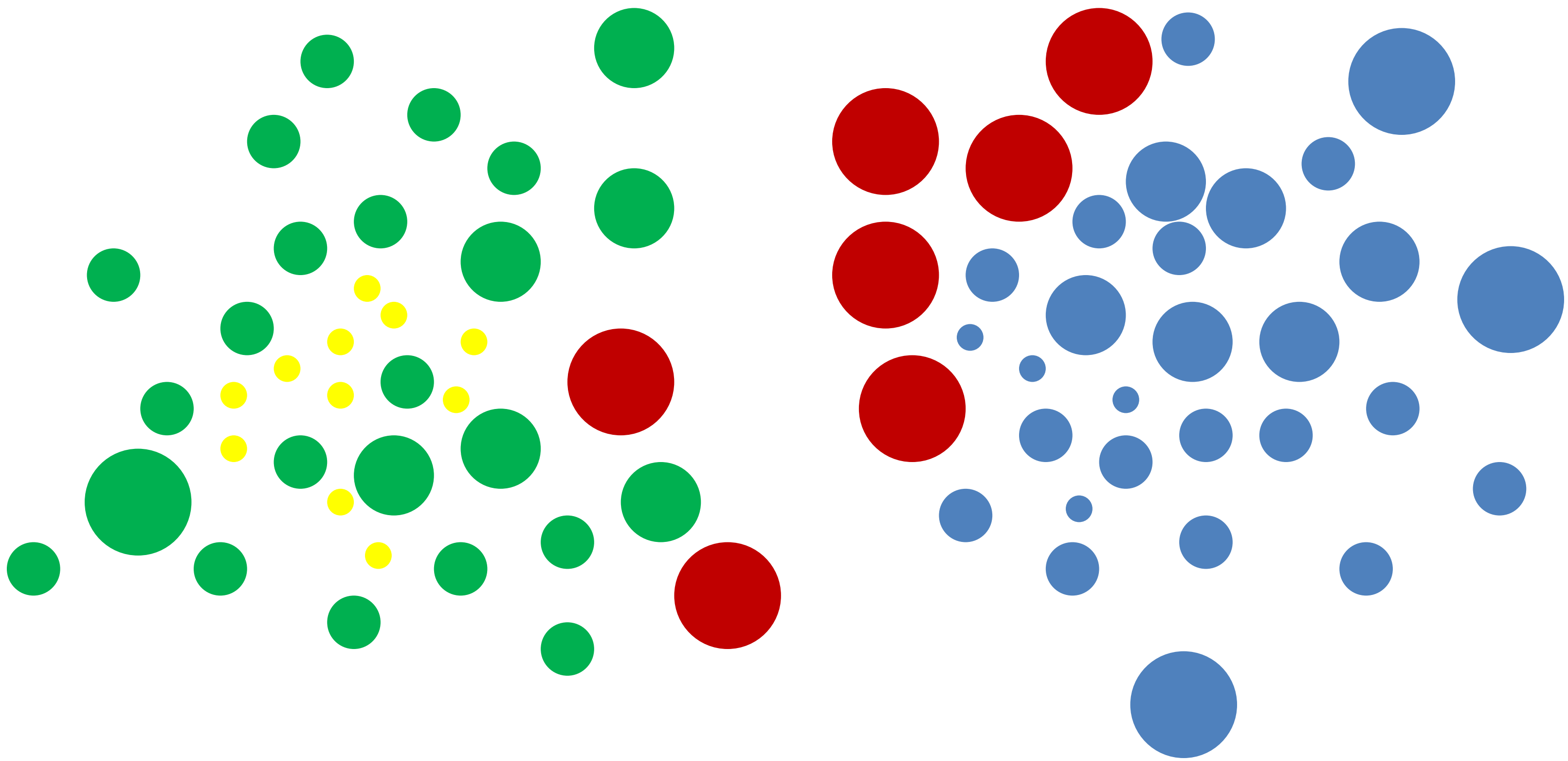
Problems with multi-measure clustering

What are the clusters according to size ?



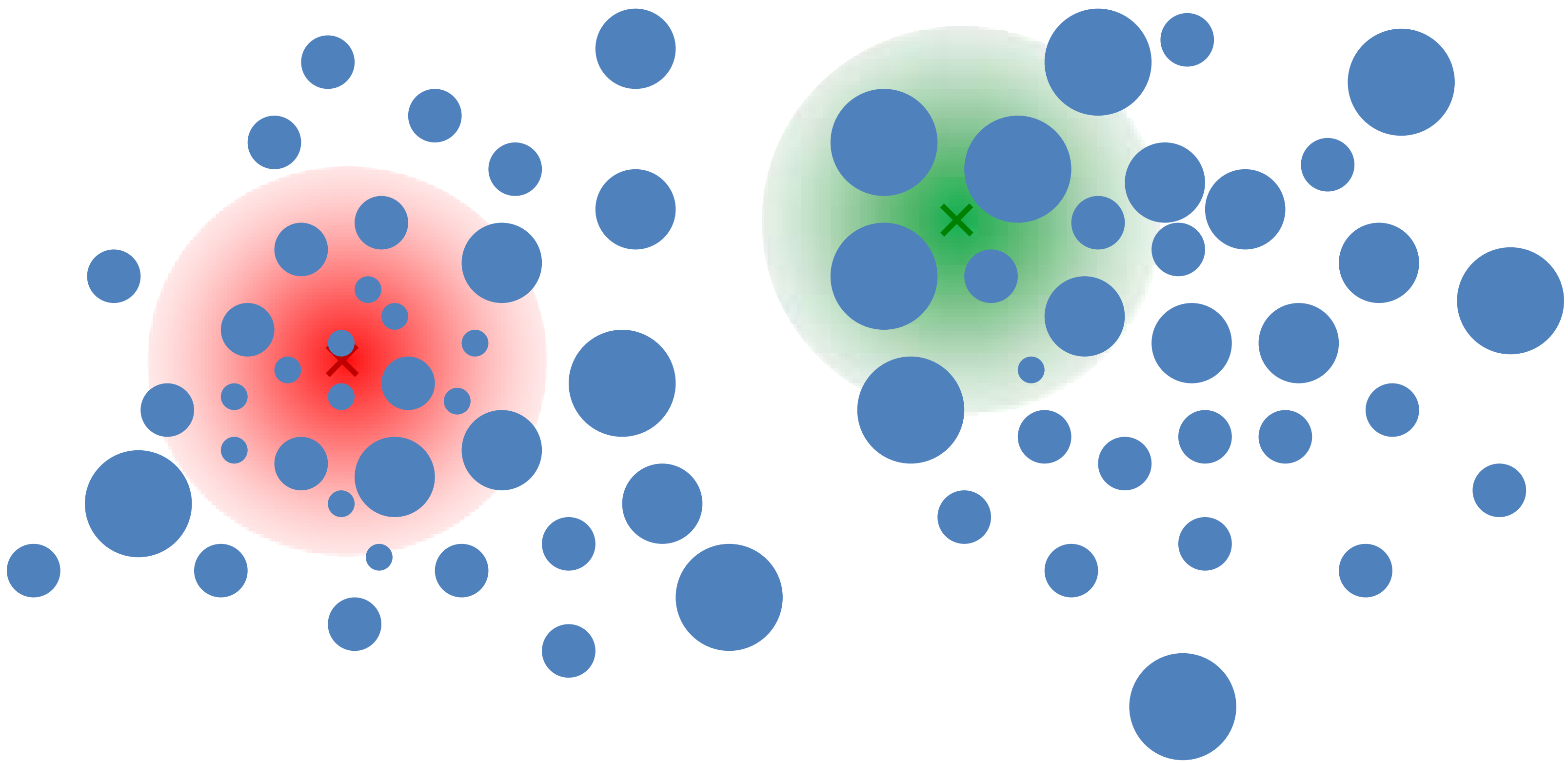
Problems with multi-measure clustering

What are the clusters according to location and size?
Well, it highly depends on how much weight we give to each *measure*.



Problems with multi-measure clustering

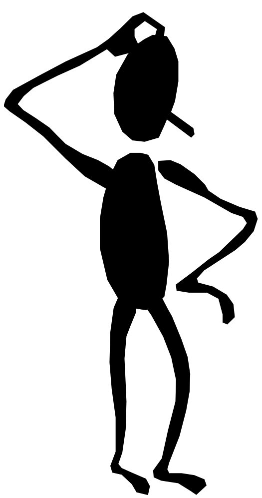
It might be more useful to find local neighborhoods with significantly (more than chance) similarity in size.

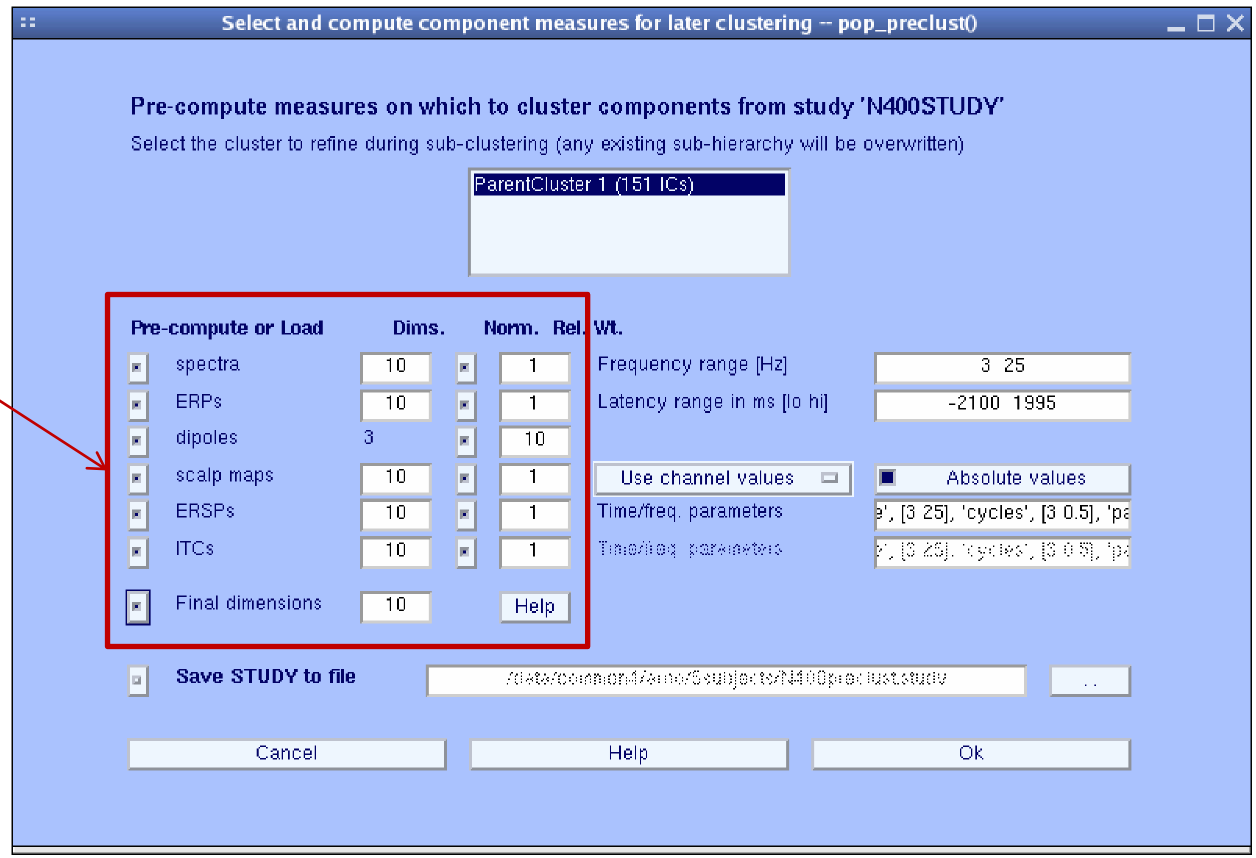


Practical problems with current methods of Study IC Clustering

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

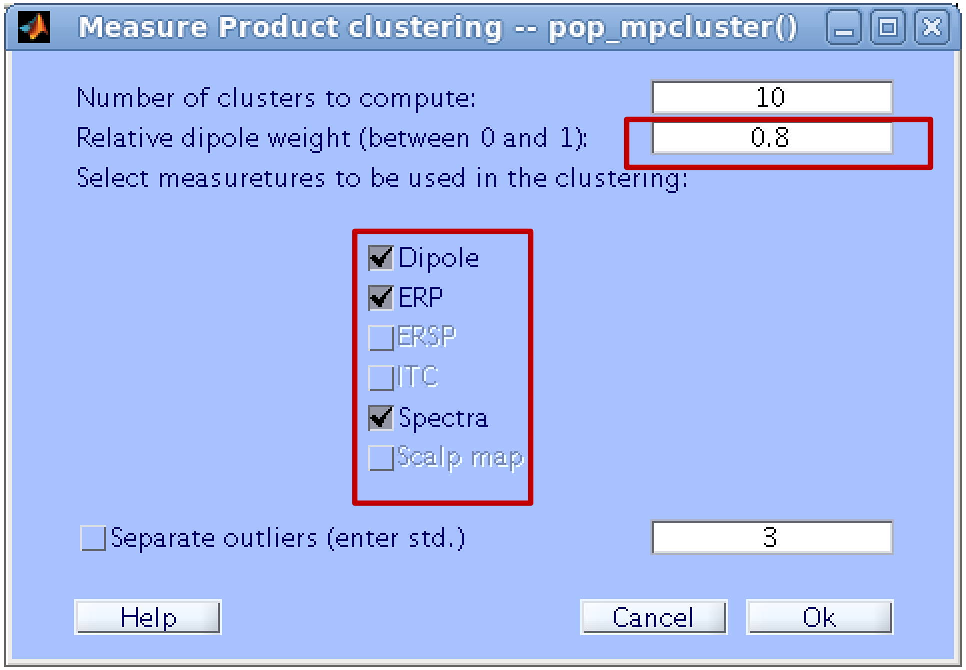
EEGLAB original clustering has ~12 parameters





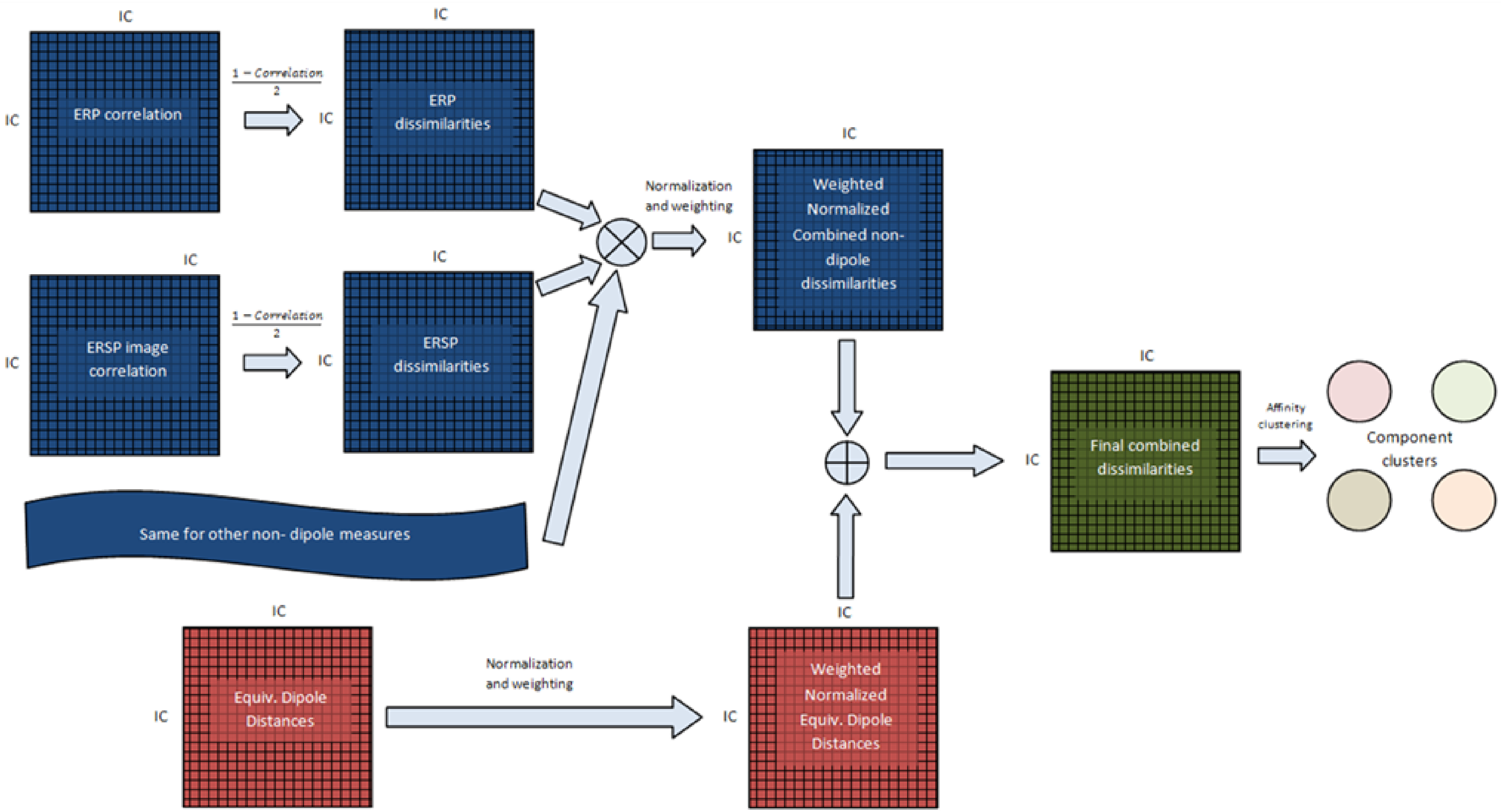
Affinity Product Clustering

Affinity product method (formerly know as Measure Product) only has one pre-clustering parameter.



Of course we still have to select the number of clusters…. but it's better than 12 parameters required in PCA clustering.

Affinity Product Clustering



Practical problems with current methods of Study IC Clustering

- We still have to select the number of clusters.
- With both these clustering methods, we basically mix (either add or multiply) distances for a subset of EEG measures (ERP, ERSP, ITC, Spectrum, Dipole location) together. This makes clustering parameters less meaningful.

Study IC Clustering

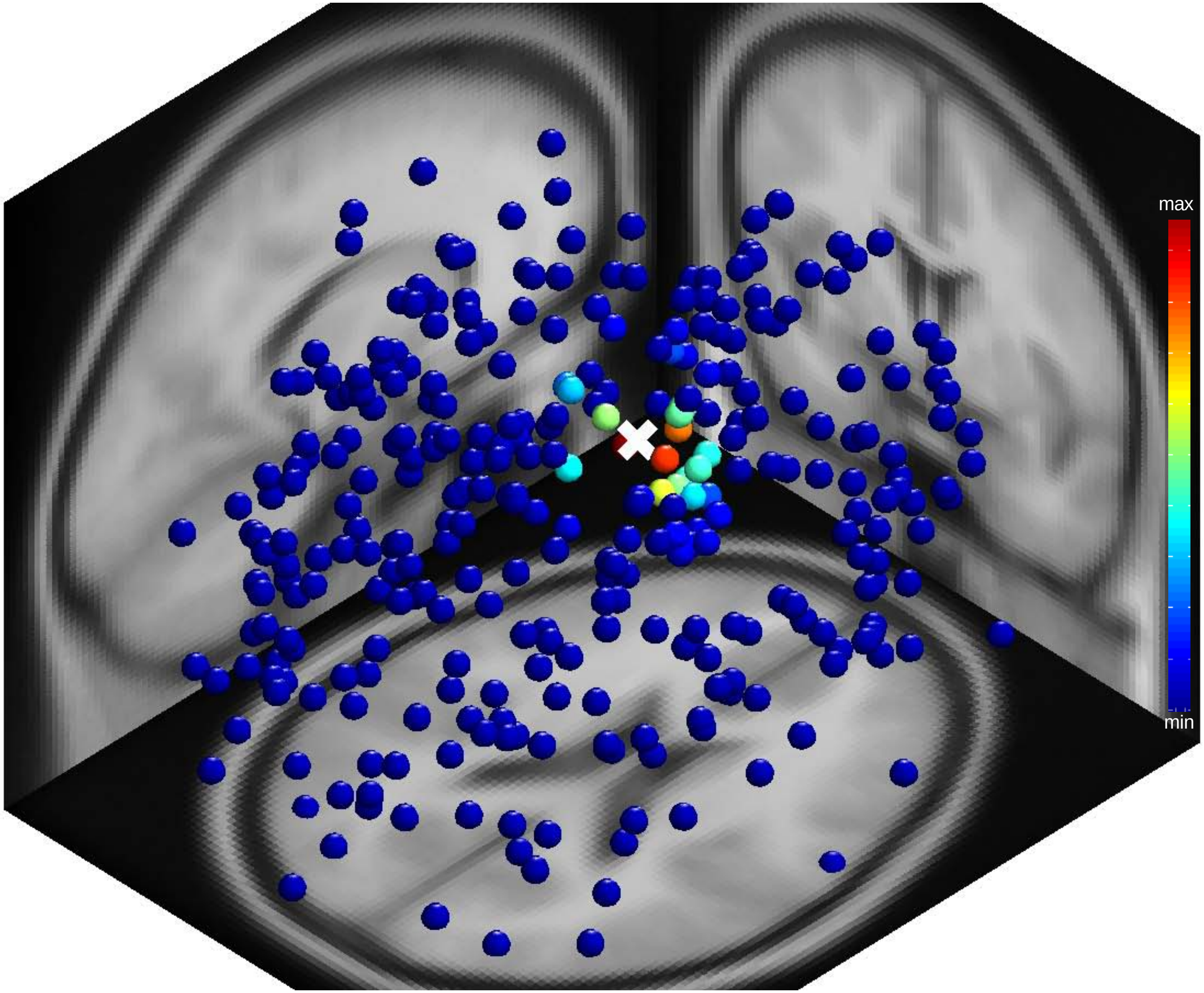
IC Clusters are not necessarily *incorrect*, or *imaginary*, but it is quite hard to provide a good argument against such objections.

Measure Projection

- Instead of clustering, we assign to each location in the brain a unique EEG response.
- The response at each location is calculated as the weighted sum of IC responses in its neighborhood.
- Weights are assigned by passing the distance between the location and IC dipole through a Gaussian function.
- The std. of this function represent expected error in dipole localization and inter-subject variability.

Measure Projection

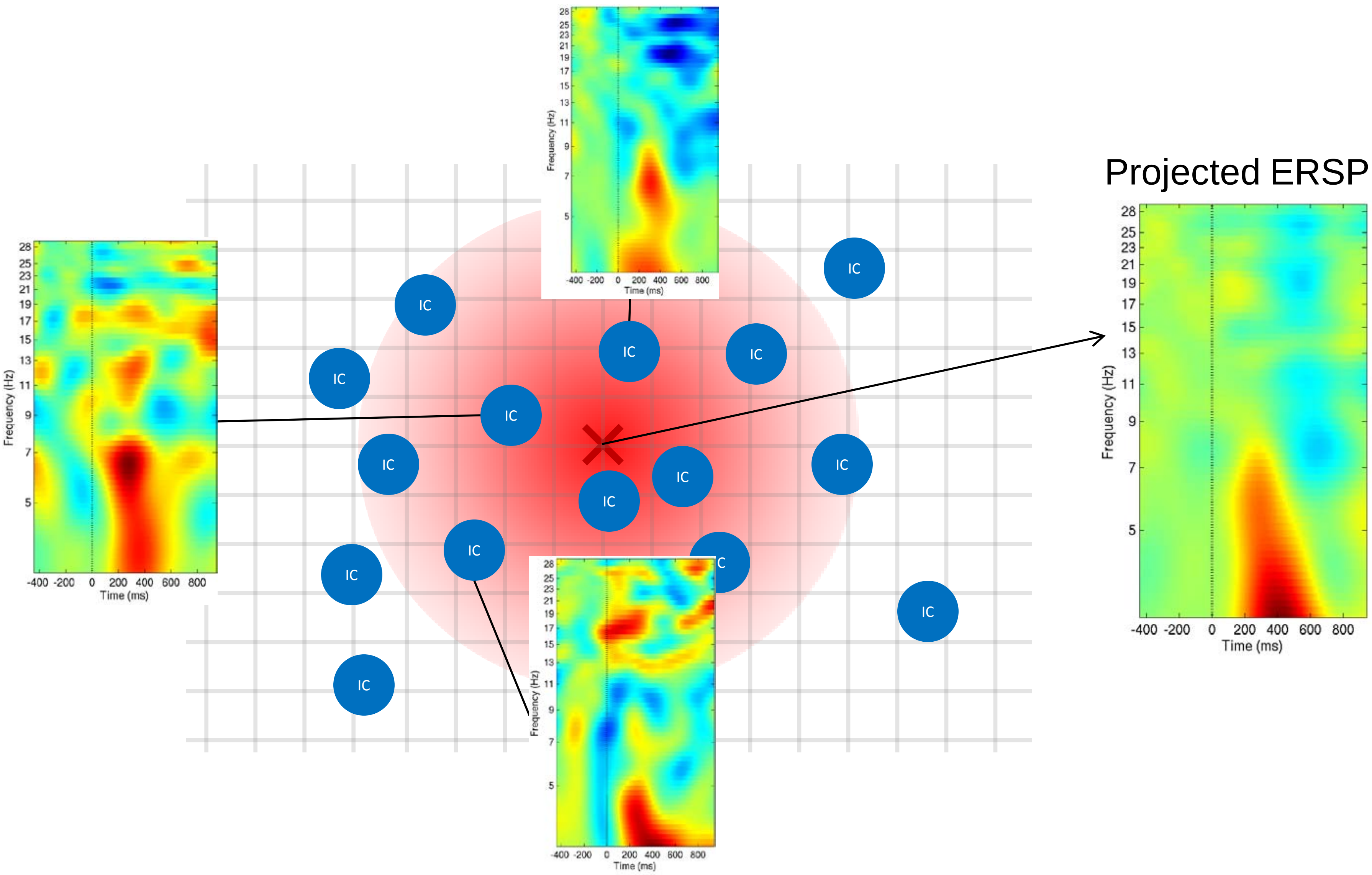
Gaussian weights
(12 mm std.)



Measure Projection

- Each EEG measure (ERP, ERSP..) is projected separately.
- Only has one (1) parameter: std. of Gaussian (which has a biological meaning).
- Bootstrap (permutation) statistics can be easily and quickly performed for each point in the brain.
- A regular grid is placed in the brain to investigate every area (with ~10 mm or less spacing).

Measure Projection



Measure Projection

- Not all projected values are significant.
- Some are weighted means of ICs with very dissimilar responses.
- Only projected values in neighborhoods with convergent responses are significant.
- Convergence can be expressed as the mean of pairwise similarities in a spatial neighborhood.
- The significance of convergence at each location can be calculated with bootstrapping (permutation).

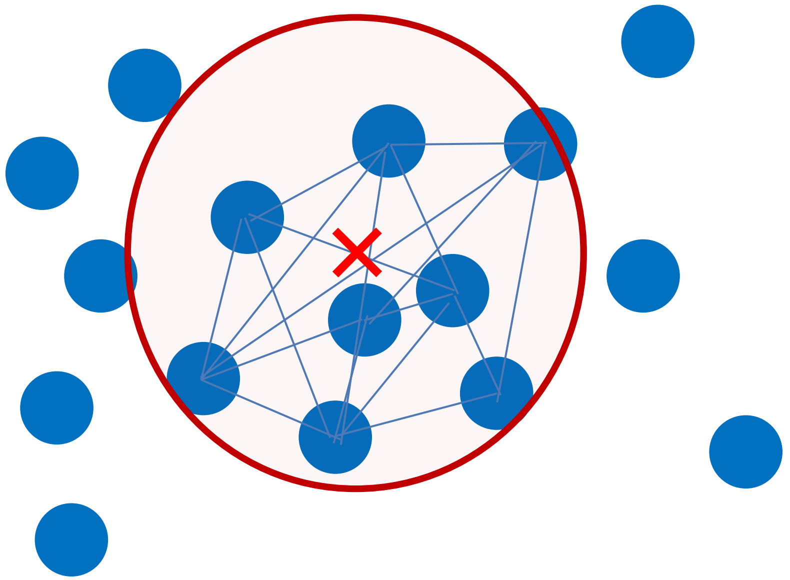
Measure Projection

For a neighborhood with a ‘fixed’ boundary, for each IC pair we can define a membership function:

$$M(IC_i, IC_j) = M(IC_i)M(IC_j)$$

Where M(IC) is one (1) if IC is in the neighborhood and zero (0) otherwise. Convergence can then be defined:

$$convergence = \frac{\sum_{i=1}^{n} \sum_{j=1, j<>i}^{n} S(i,j)M(i,j)}{\sum_{i=1}^{n} \sum_{j=1, j<>i}^{n} M(i,j)}$$



Where M is the neighborhood membership matrix and S is the pairwise similarity matrix. This is basically the mean of pairwise IC similarities around a location in the brain.

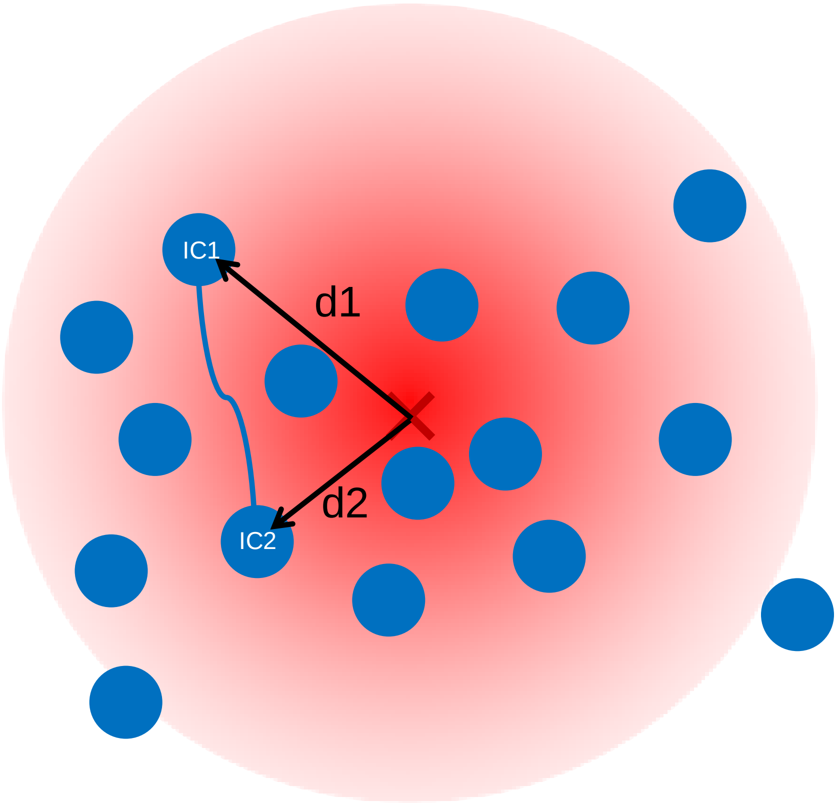
Measure Projection

Now we can extend this concept of convergence to neighborhoods with 'soft' Gaussian boundaries, for each IC pair we modify the membership function:

$$MG(IC_i, IC2_j) = MG(IC_i)MG(IC_j)$$

Where $MG(IC) = e^{-\frac{d^2}{2\sigma^2}}$ (d is distance from IC equiv. dipole to neighborhood center). Convergence can now be defined as:

$$convergence = \frac{\sum_{i=1}^{n} \sum_{j=1, j<>i}^{n} S(i,j) e^{-\frac{d_1^2 + d_2^2}{2\sigma^2}}}{\sum_{i=1}^{n} \sum_{j=1, j<>i}^{n} e^{-\frac{d_1^2 + d_2^2}{2\sigma^2}}}$$

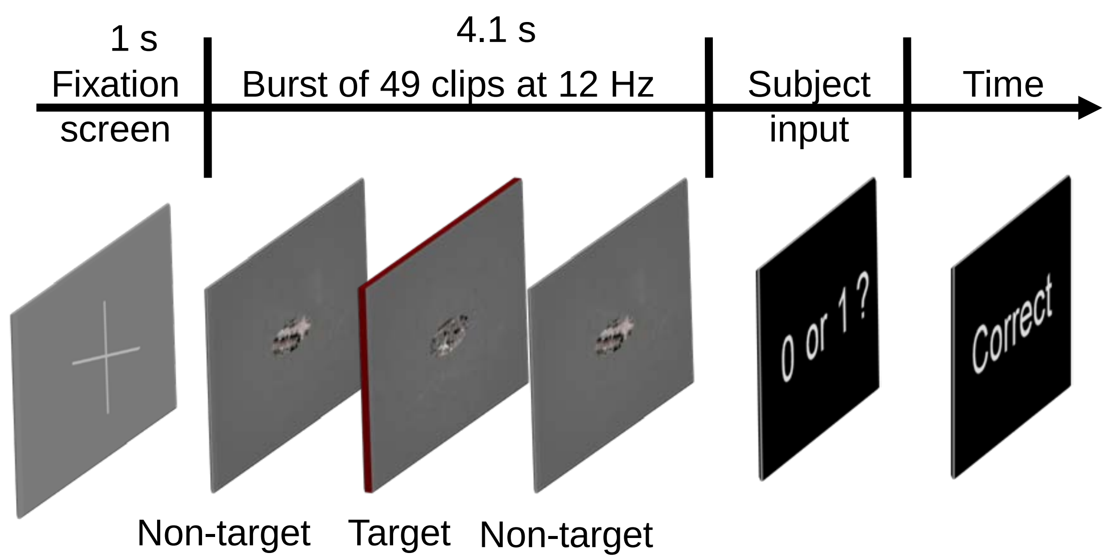


Where S is the pairwise similarity matrix. This is basically the weighted mean of IC similarities around a location in the brain.

Measure Projection: RSVP Example

Rapid Serial Visual Presentation Experiment

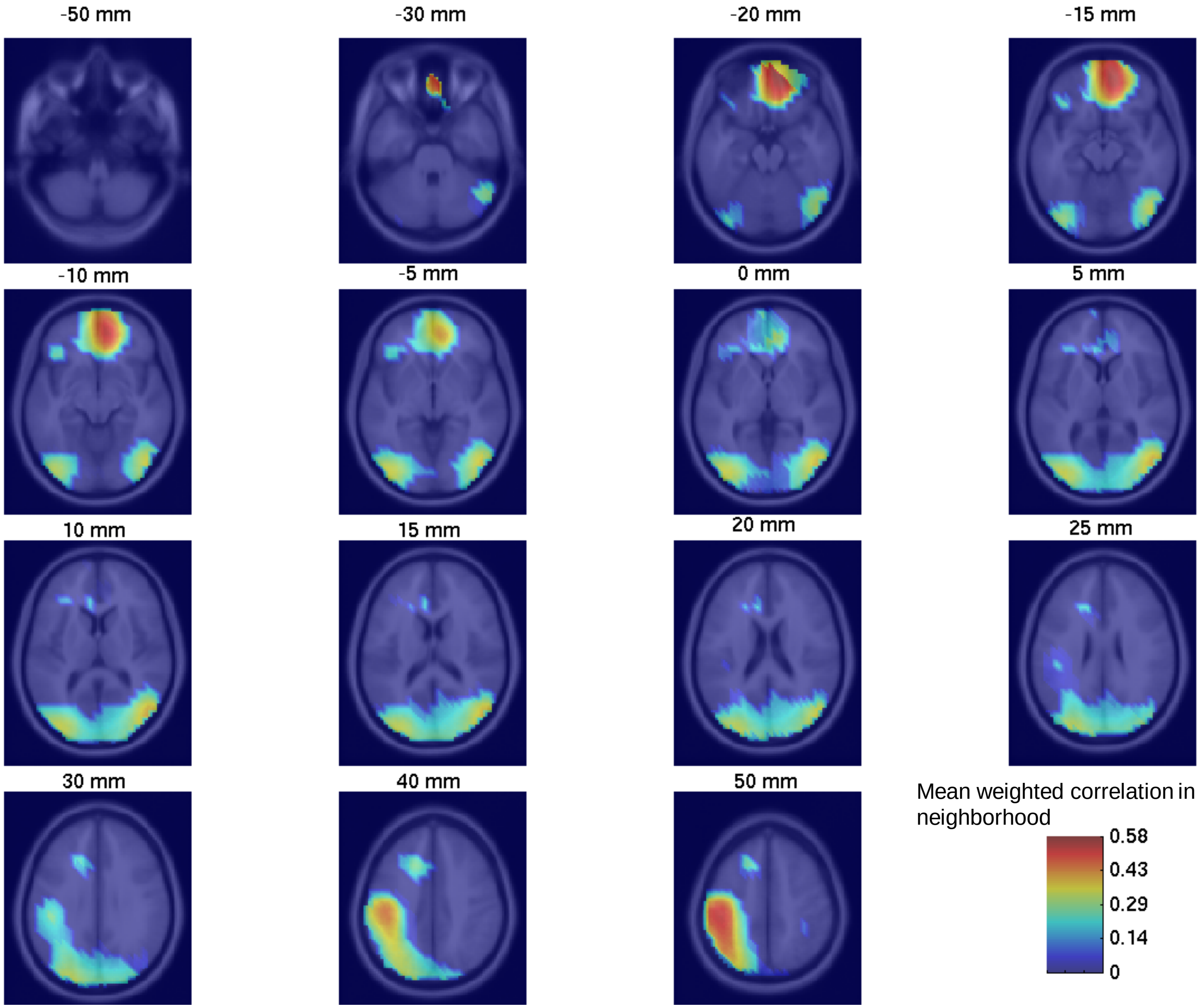
- 8 subjects
- 15 Sessions
- Visual target detection
- 257 components with equiv. dipoles inside the brain



Measure Projection: RSVP Example

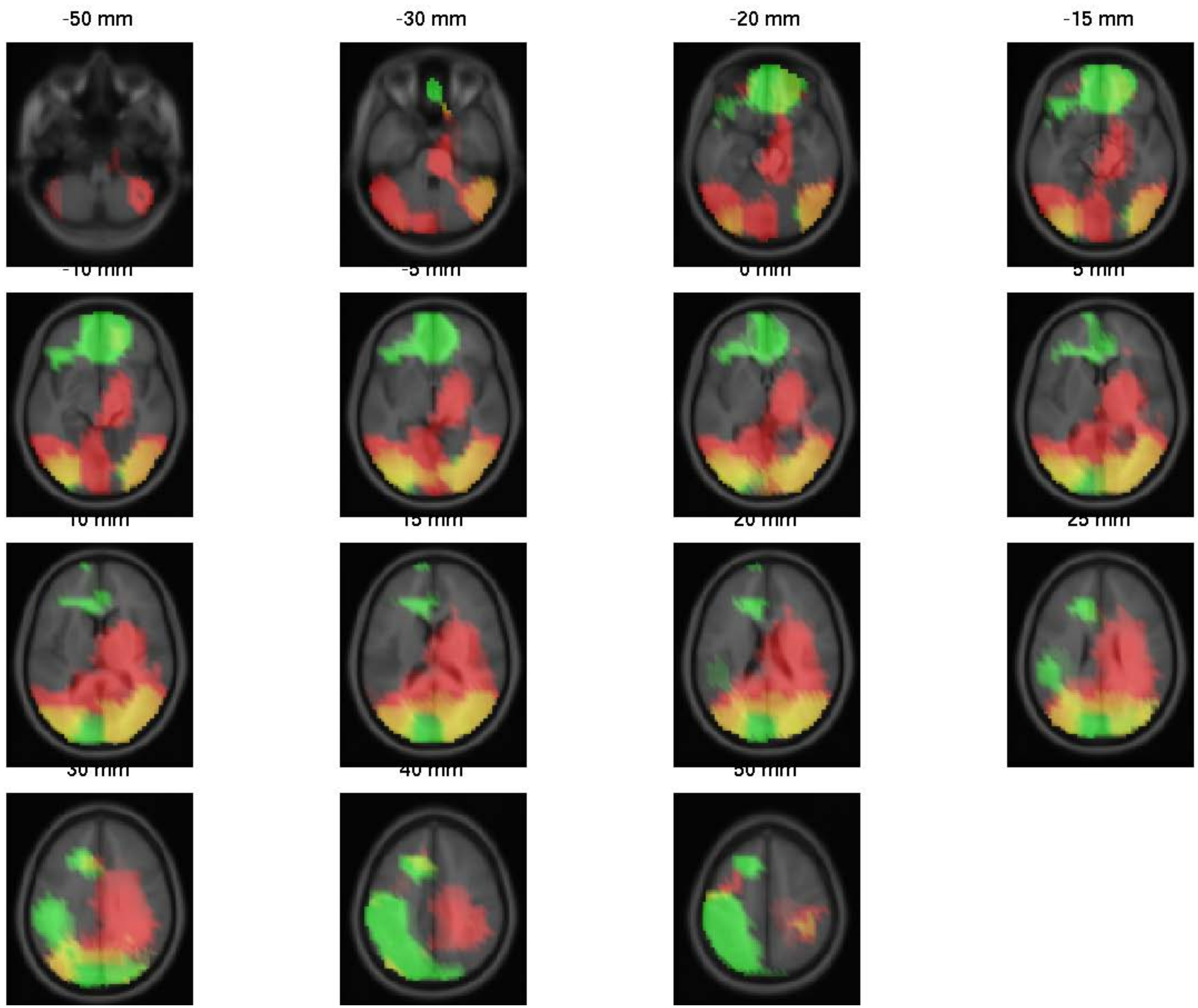
Areas in which convergence is significant ($p<0.01$).

Gaussian std. = 12 mm



Measure Projection: RSVP Example

ERP and ERSP locations with significant convergence ($p<0.01$)



ERSP

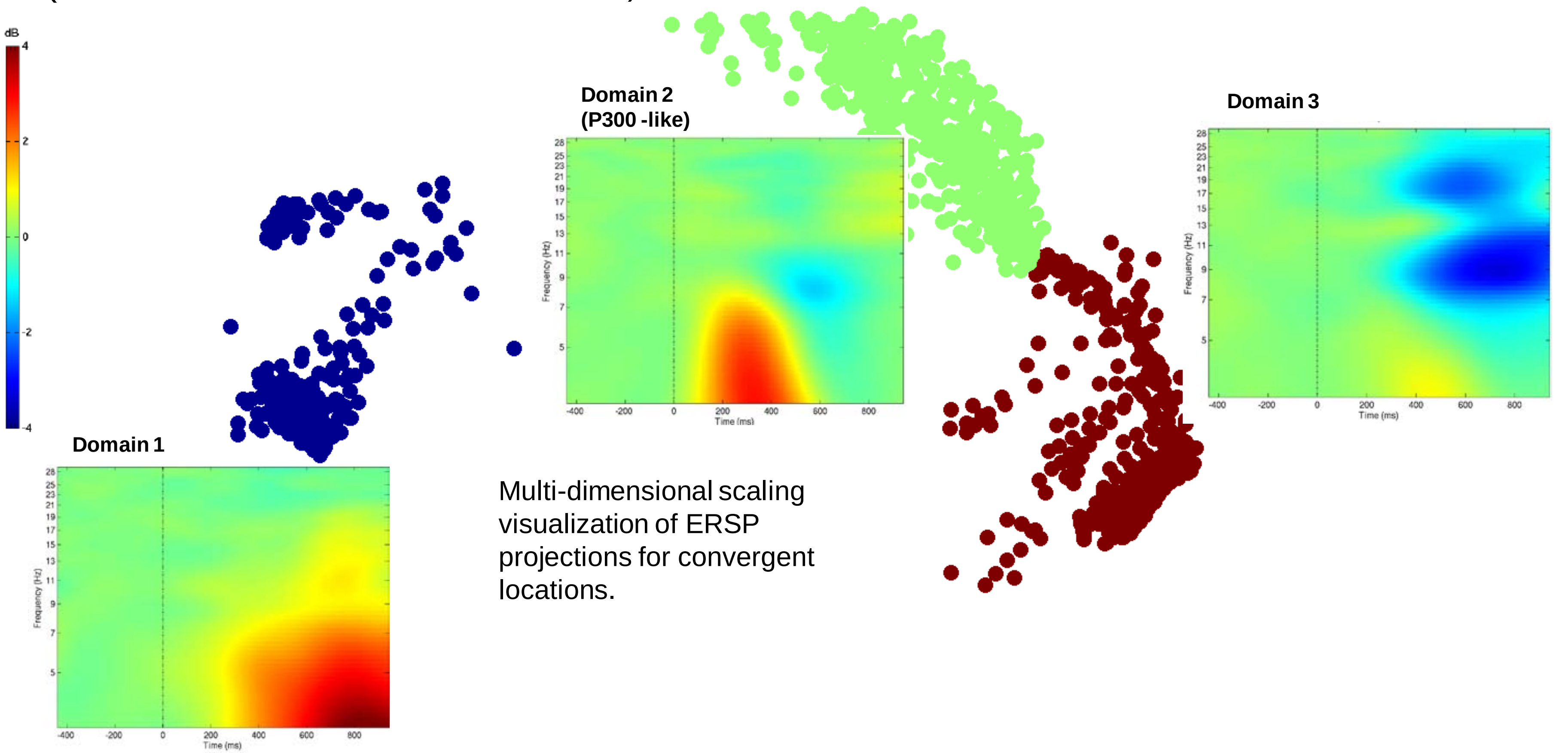
ERP

ERP and ERSP

Measure Projection: RSVP Example

To better visualize measure responses in areas with significant convergence, they can be summarized into different *domains*. The exact number of these domains depends on how similar their exemplars are allowed to be.

Below you can see ERSP responses in the RSVP experiment form three (3) domains (with the correlation threshold of 0.8).



Multi-dimensional scaling visualization of ERSP projections for convergent locations.

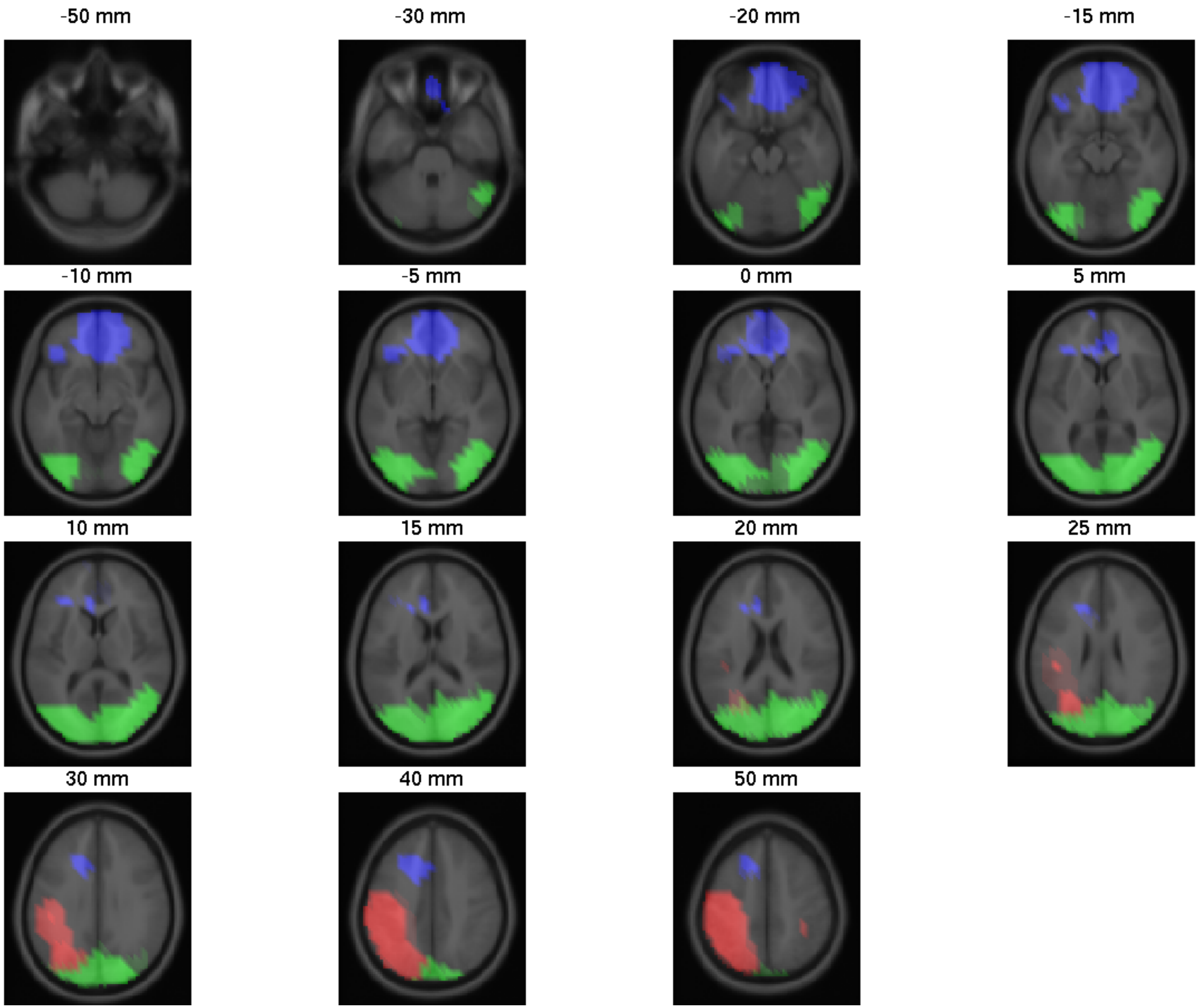
Measure Projection: RSVP Example

ERSP domains (exemplar similarity <0.8)

Domain 1

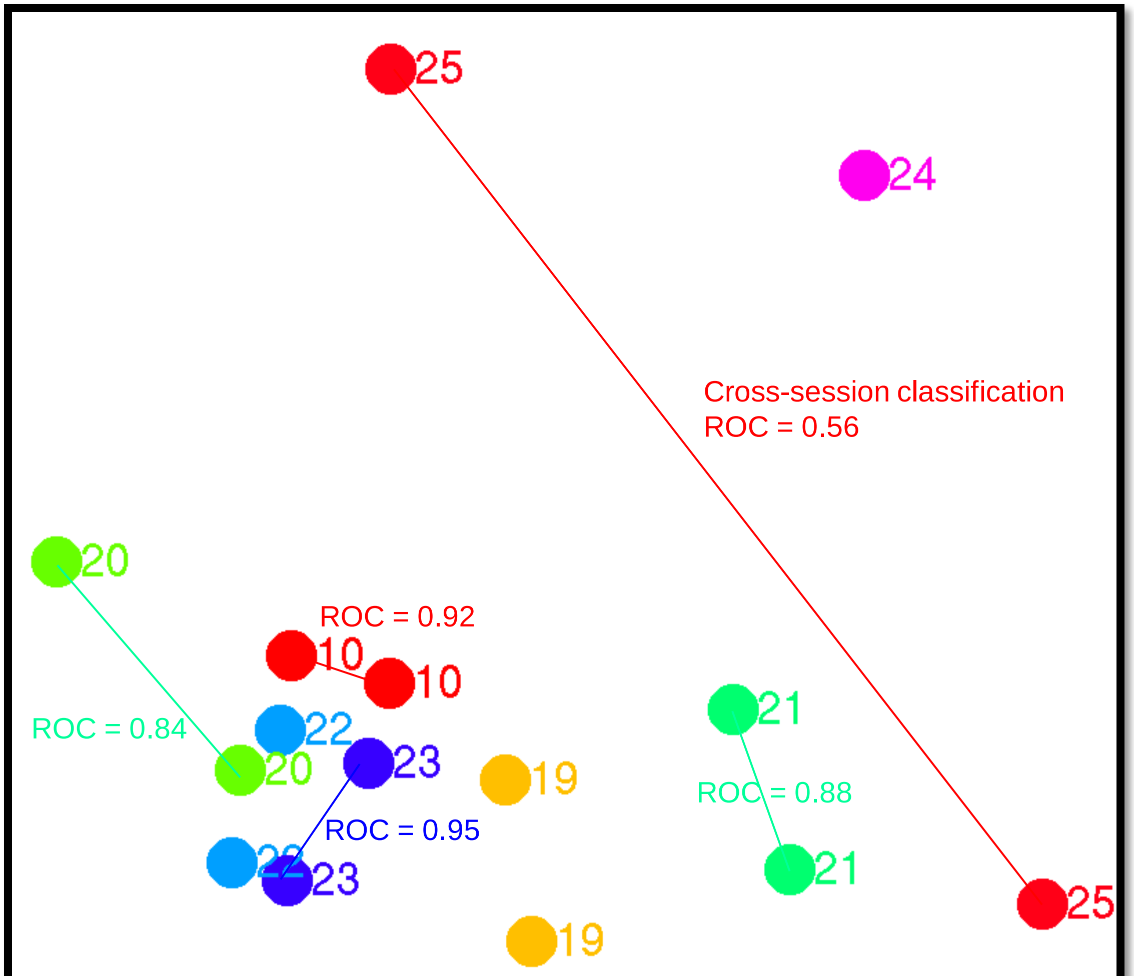
Domain 2

Domain 3

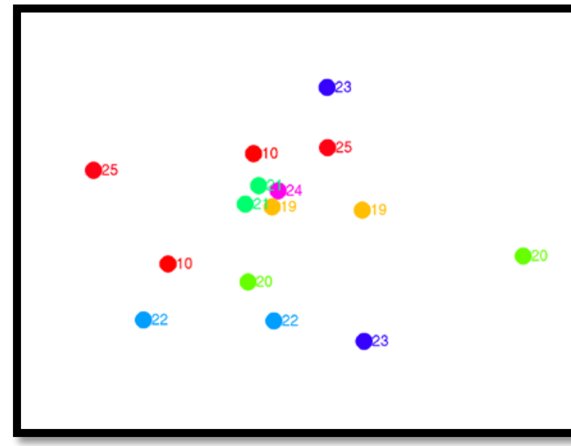


Measure Projection: RSVP Example

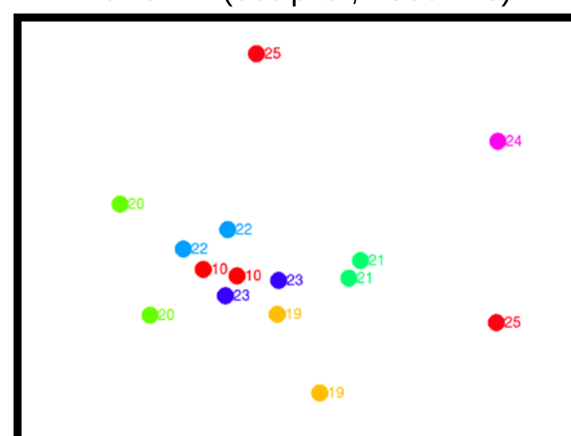
Subject-Session Similarity Space (S^4), All domains



Domain 1 (frontal)



Domain 2 (occipital, P300-like)



Measure Projection: Summary

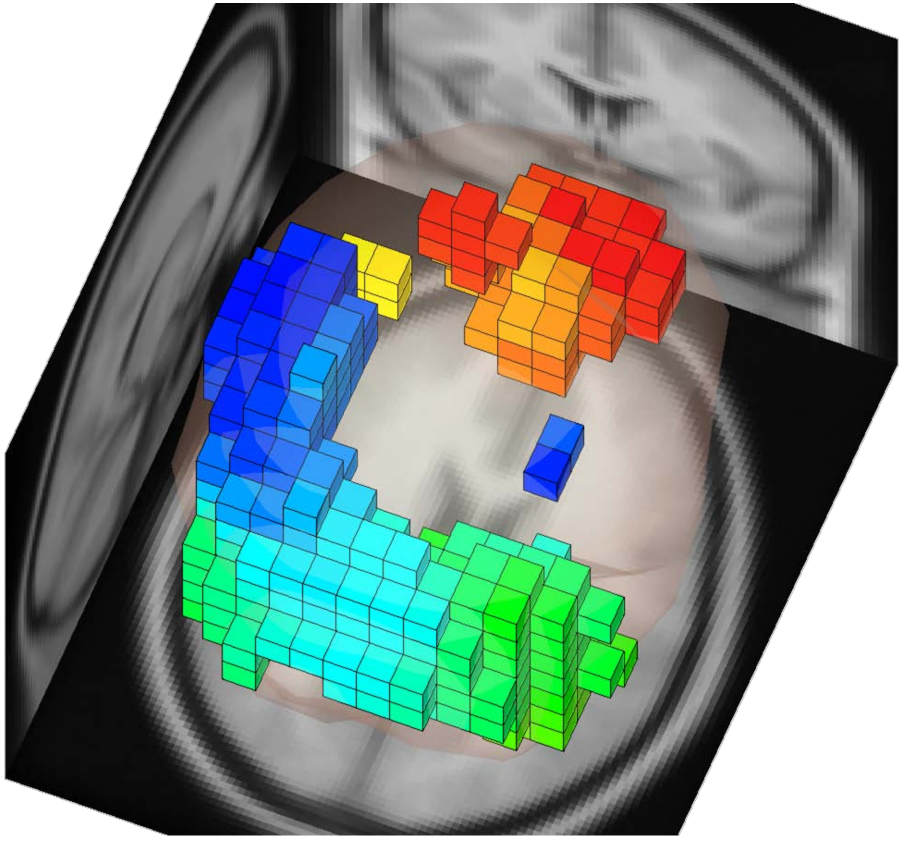
- Enables us to compare subjects, groups and conditions at every brain location.
- Enables us to calculate significance on every step.
- Enables us to perform new types of analysis that we could not do with IC clusters (e.g. subject similarity space)
- All types of analysis that can be done on IC clusters, can also be performed in Measure Projection framework.

Measure Projection: Summary

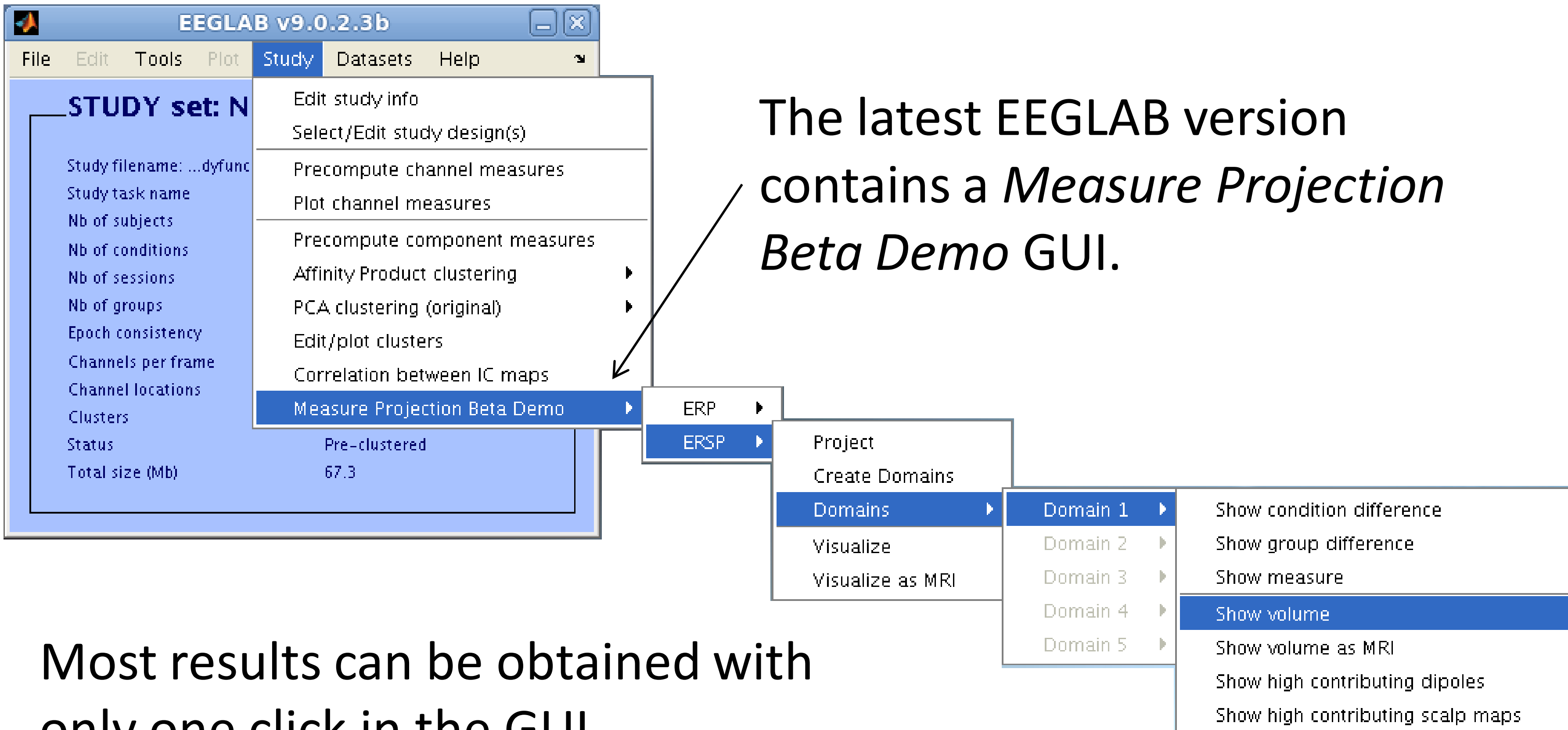
- Assigns measures directly to brain location. This simplifies comparing results across different studies and modalities (e.g. fMRI).
- Provides a framework for analyzing EEG data that has several advantages over clustering and does not suffer from its conceptual and practical problems.
- The closest concept in MPA to a cluster is a Domain, which identifies a region of brain volume with highly similar projected measure.

Measure Projection toolbox

- 95% of capabilities alternative to clustering are currently implemented in the toolbox.
- Pair-wise comparison of Conditions or Groups is implemented .
- MRI-like visualization and analysis.
- The image on the right shows areas with significant ($p<0.03$) high IC ERSP similarity in their neighborhood in an RSVP experiment.
- Voxels with similar ERSP are colored similarly (with multi-dimensional scaling of projected measure to Hue)



Measure Projection toolbox



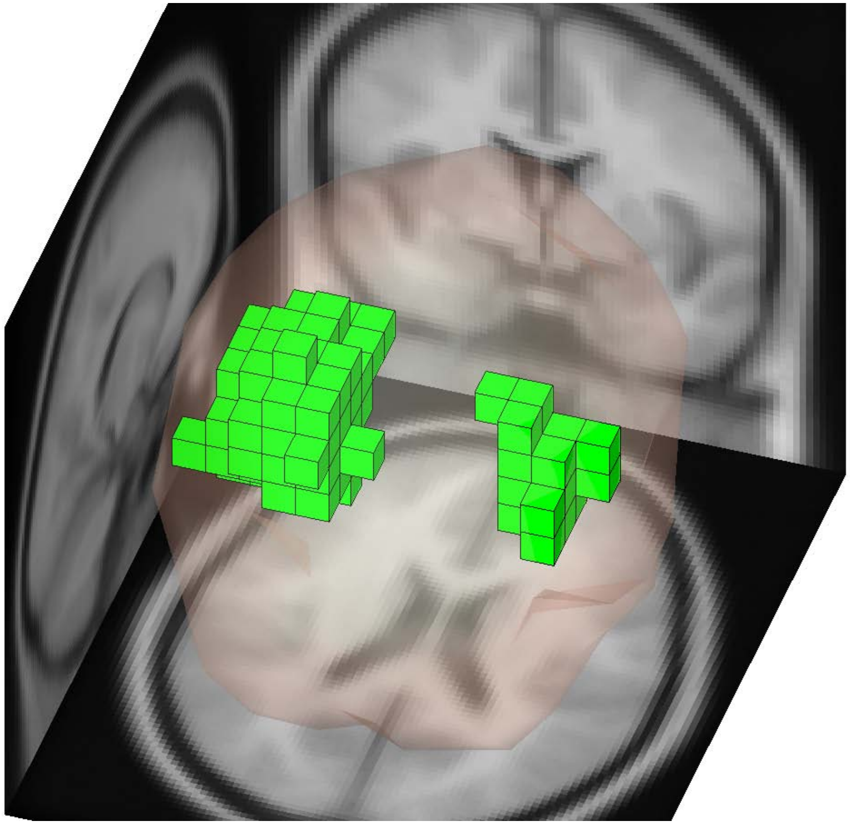
The latest EEGLAB version contains a *Measure Projection Beta Demo* GUI.

Most results can be obtained with only one click in the GUI.
All results are available within two clicks.

Measure Projection toolbox

Roadmap:

- Full support of STUDY Design
- Single trial projection (possibly useful for building classifier from multiple subjects)
- Multiple ICA models for each session.
- Subject-Session comparison on regions of interest (ROIs).



An ERP domain from Sternberg task

Exercise: Affinity Product clustering

1. Install the latest version of EEGLAB or download, extract and copy plug-in files from the wiki to your /eeglab/plugins/ folder.
2. Type >> eeglab rebuild
3. Load 5 subject N 400 study.
4. Click on *Study->Affinity Product clustering-> Build Pre-clustering matrices.*
5. Click on *Study->Affinity Product clustering-> Cluster Components.*
6. Select only *Dipole* and *ERSP* measure with 10 clusters, then click OK.

Exercise: Measure Projection clustering

1. Go to *Study->Measure Projection Beta Demo -> ERSP -> Visualize*
2. Try Create Domains under the ERP menu.
3. Click on *Domains -> Domain 1 (or n) -> Show Volume.*
4. Click on *Domains -> Domain 1 (or n) Show Measure*
5. Try other *Domain* menu options.