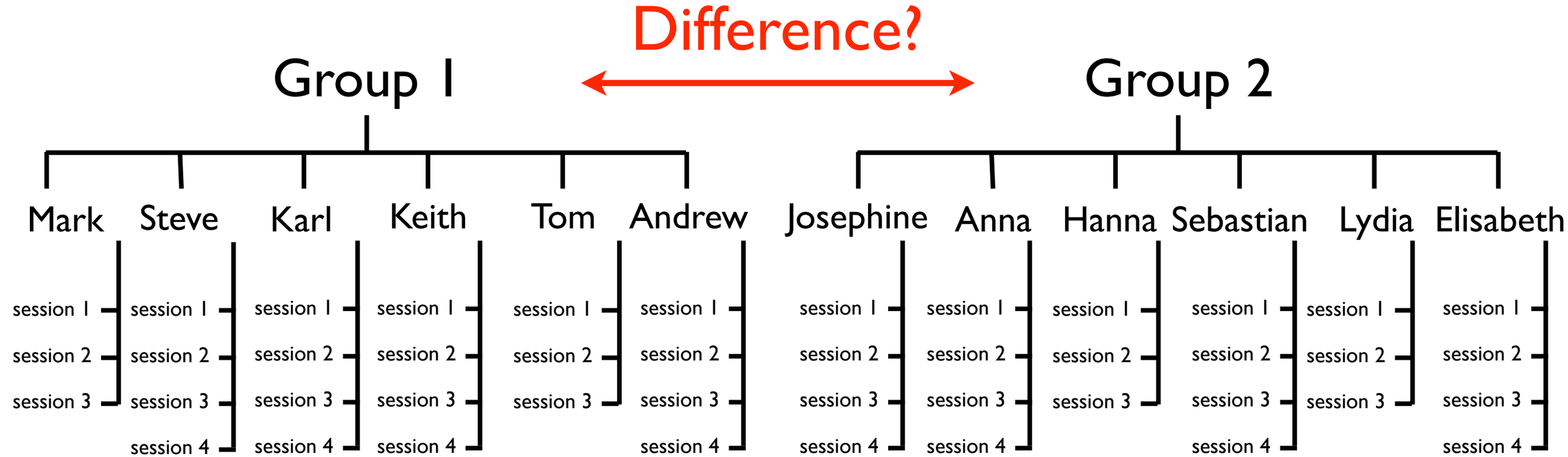
Group-level Analysis in OSL

OHBA Analysis Workshop

Group Analysis

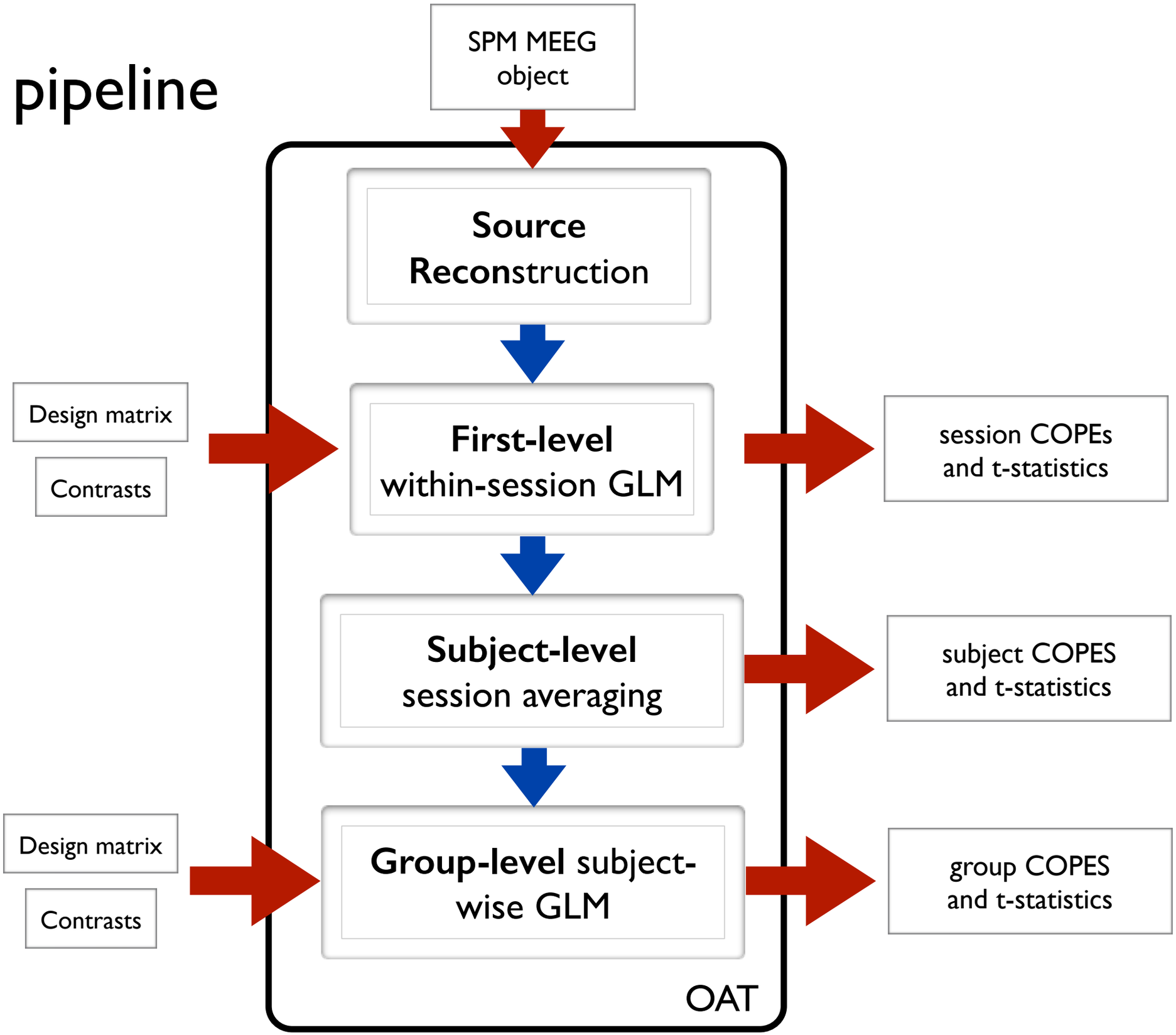
- typically need to infer across multiple subjects, sometimes multiple groups and/or multiple sessions



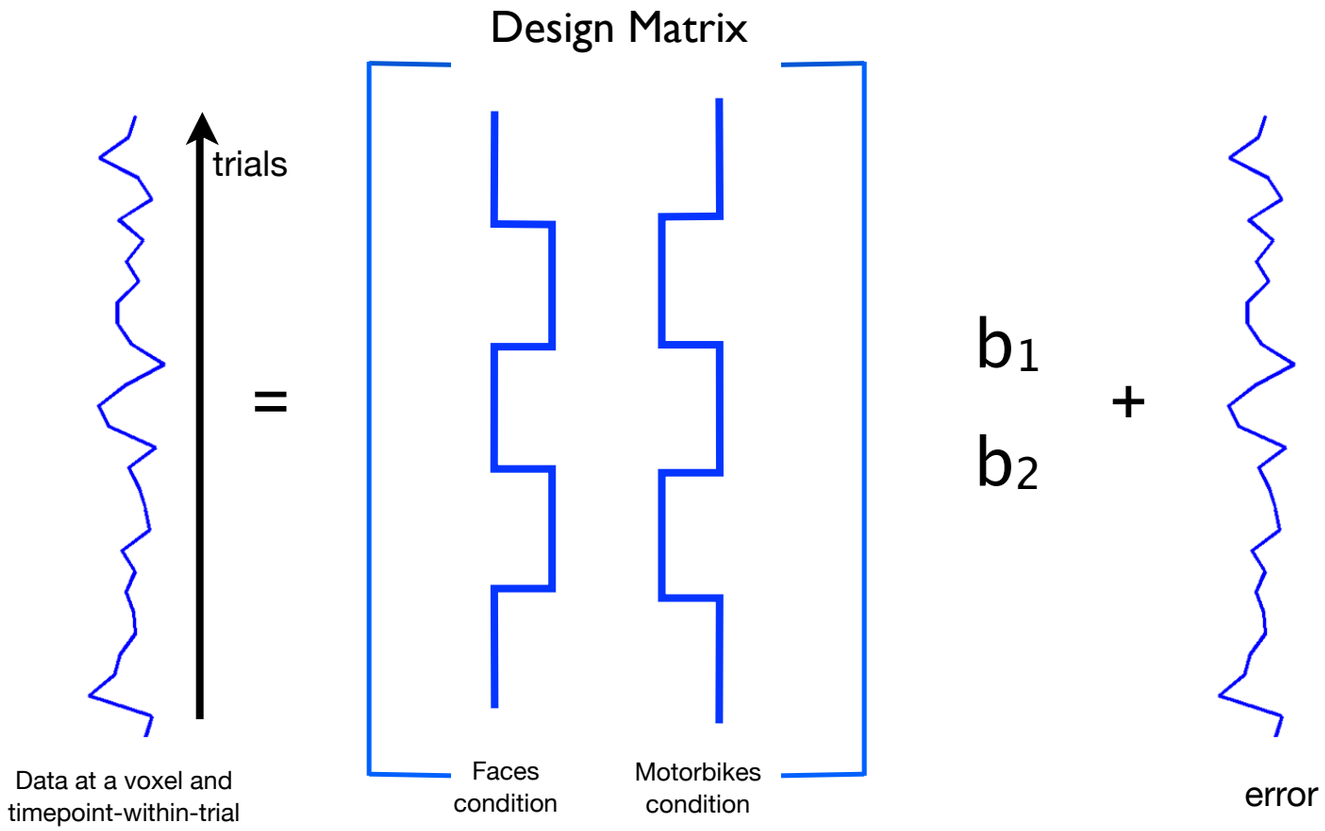
- questions of interest involve tests/comparisons at the group level

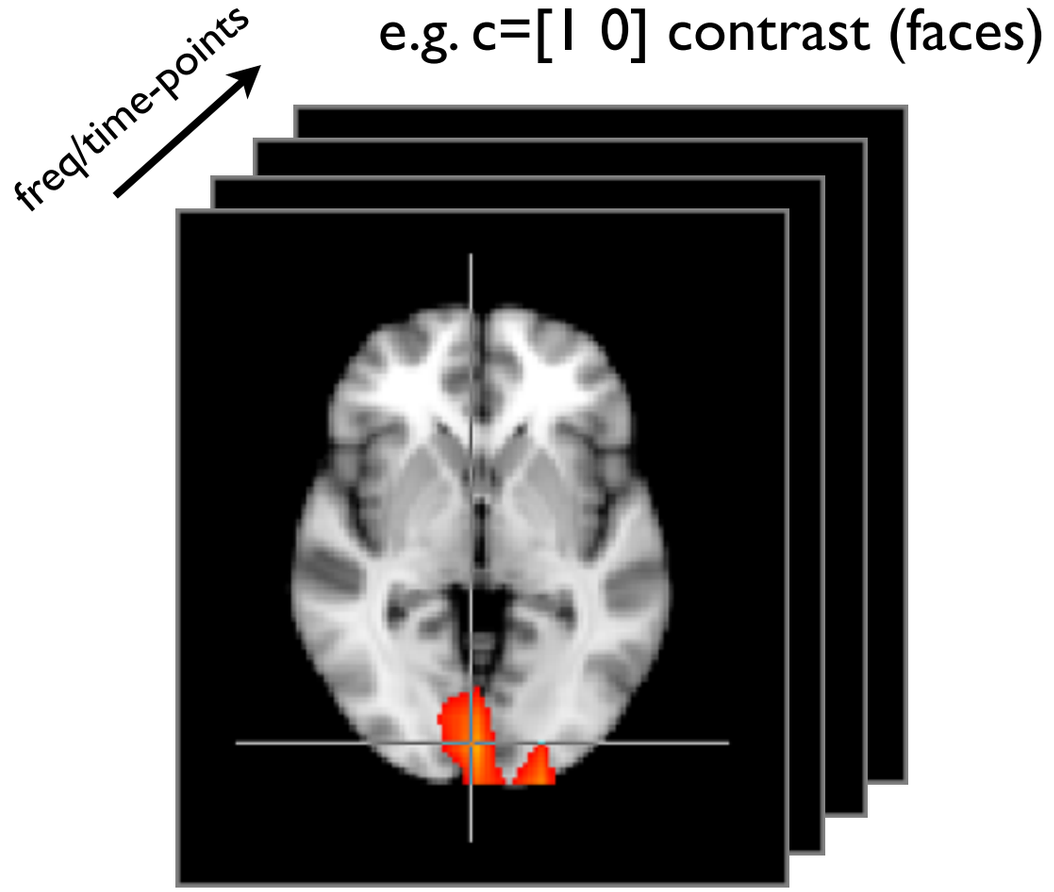
OAT Pipeline Stages

- 4 distinct pipeline stages:



First-Level (Trial-wise) GLM



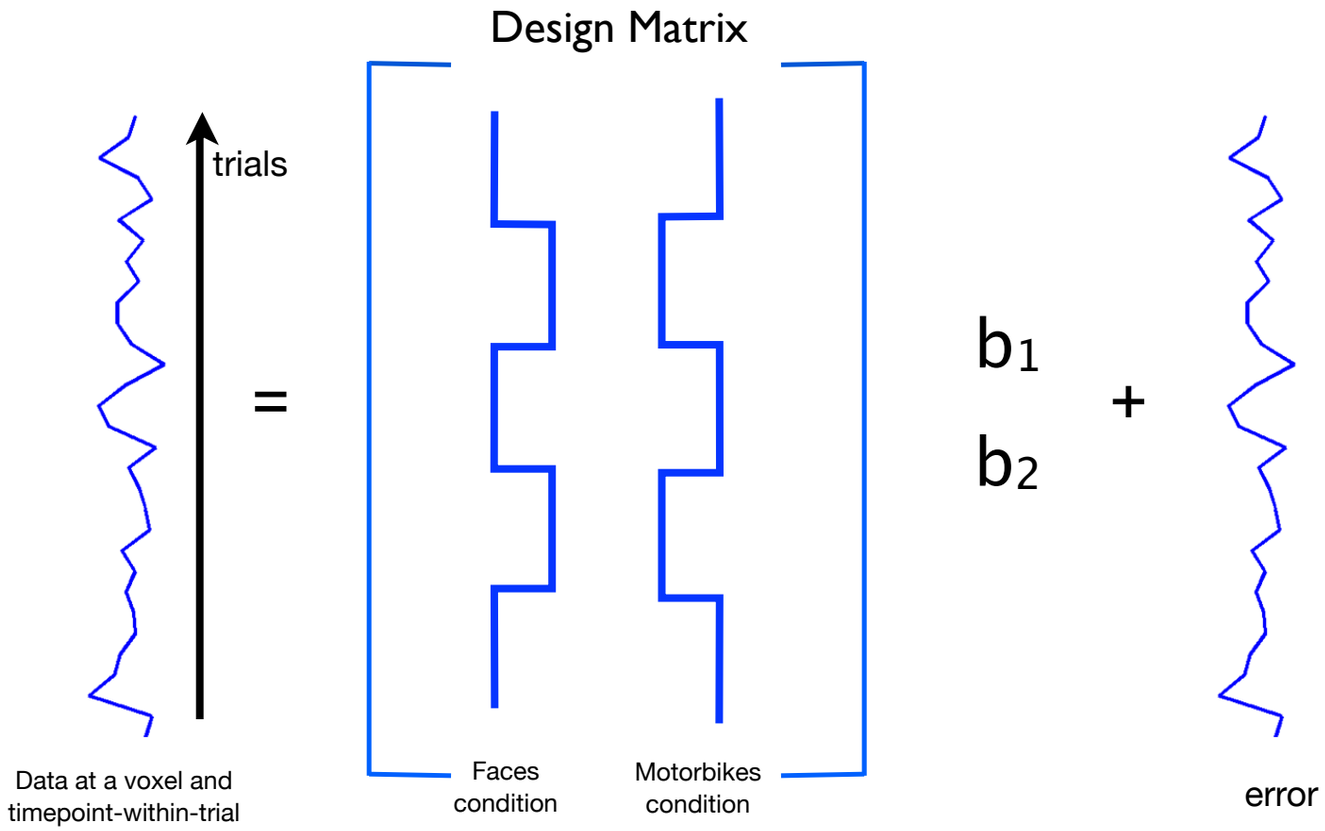


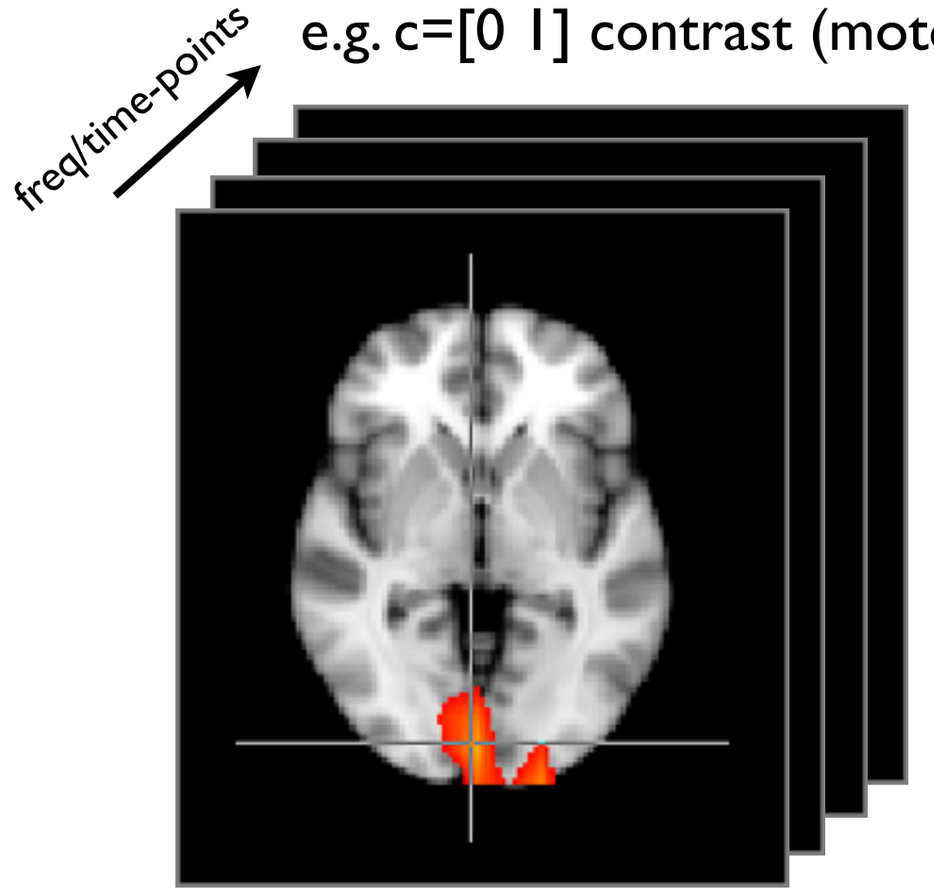
Repeat for all voxels/ freq/time-points, then compute contrasts: COPE=c*B

$$Y=Xb+e \rightarrow B=pinv(X)Y$$

B is the Parameter Estimates (PEs) of b

First-Level (Trial-wise) GLM



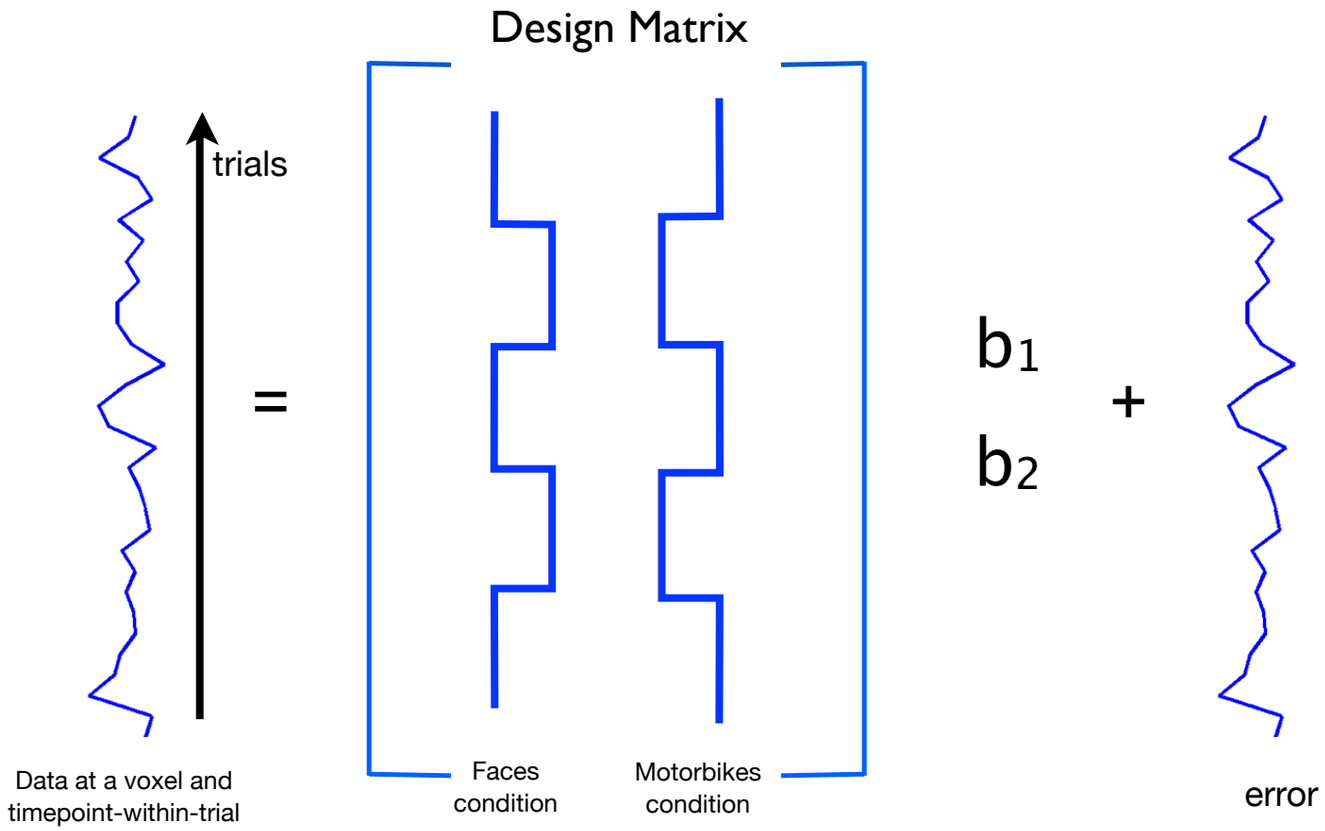


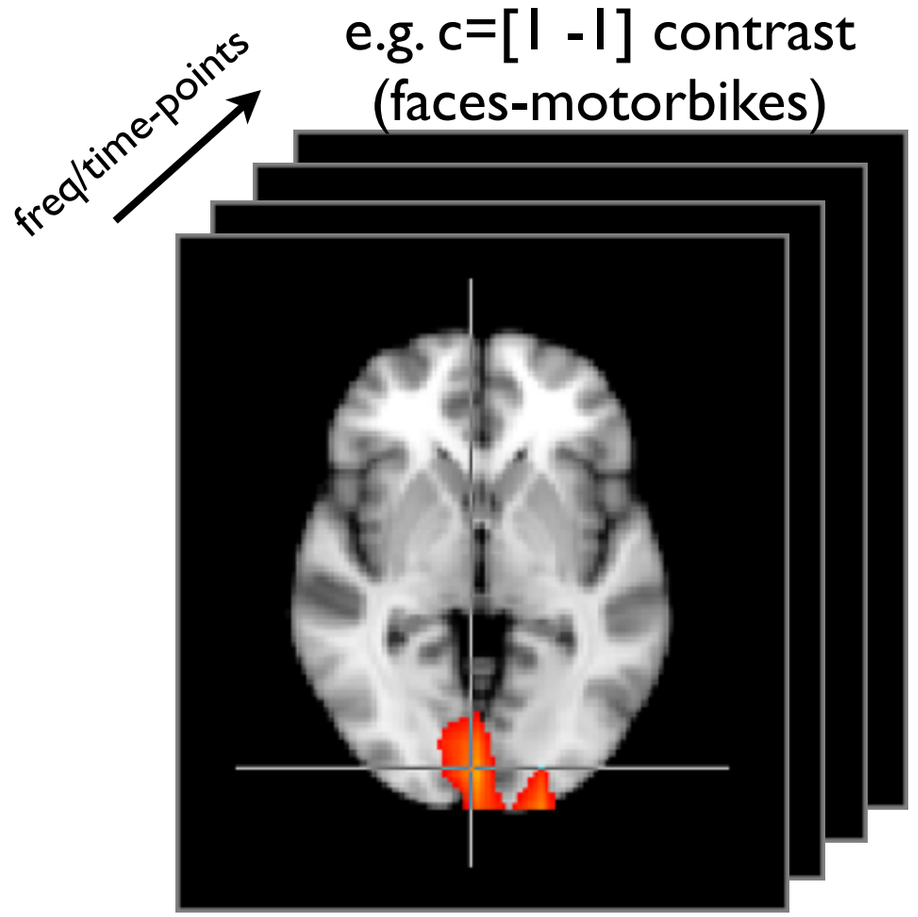
Repeat for all voxels/ freq/time-points, then compute contrasts: COPE=c*B

$Y=Xb+e \rightarrow B=pinv(X)Y$

B is the Parameter Estimates (PEs) of b

First-Level (Trial-wise) GLM





Repeat for all voxels/ freq/time-points, then compute contrasts: COPE=c*B

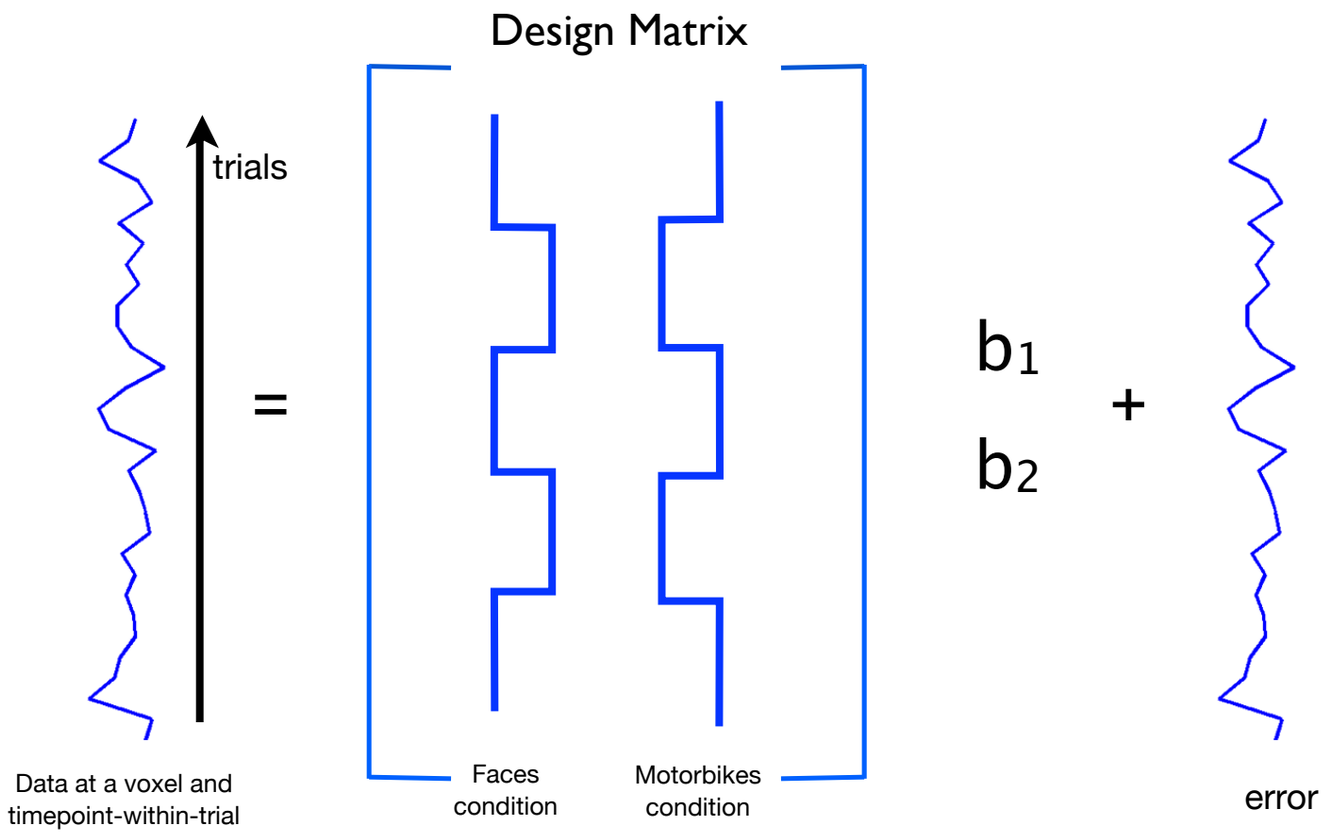
$$Y=Xb+e \rightarrow B=pinv(X)Y$$

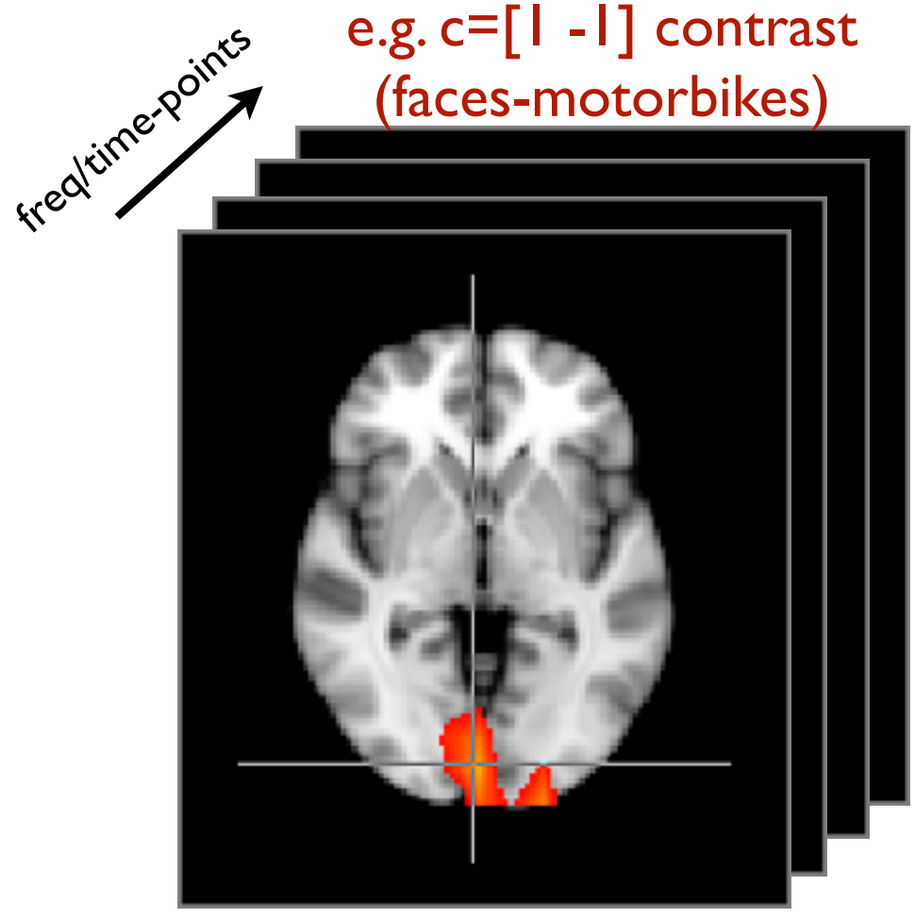
B is the Parameter Estimates (PEs) of b

First-Level (Trial-wise) GLM

first-level COPEs are the INPUTS (data) in the Group Analysis:

e.g. c=[1 -1] contrast
(faces-motorbikes)





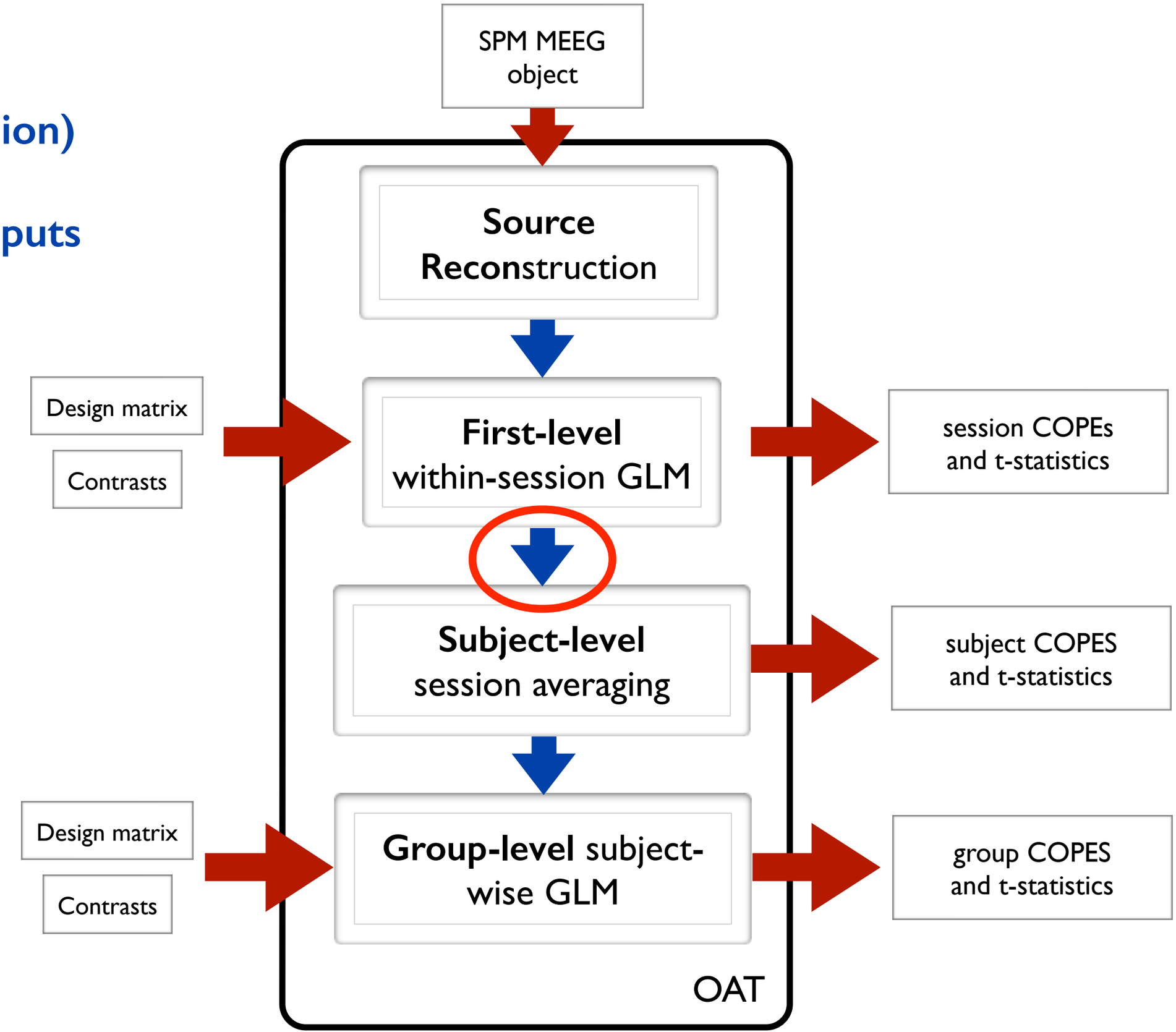
Repeat for all voxels/
freq/time-points, then
compute contrasts:
COPE=c*B

$$Y=Xb+e \rightarrow B=pinv(X)Y$$

B is the Parameter Estimates (PEs) of b

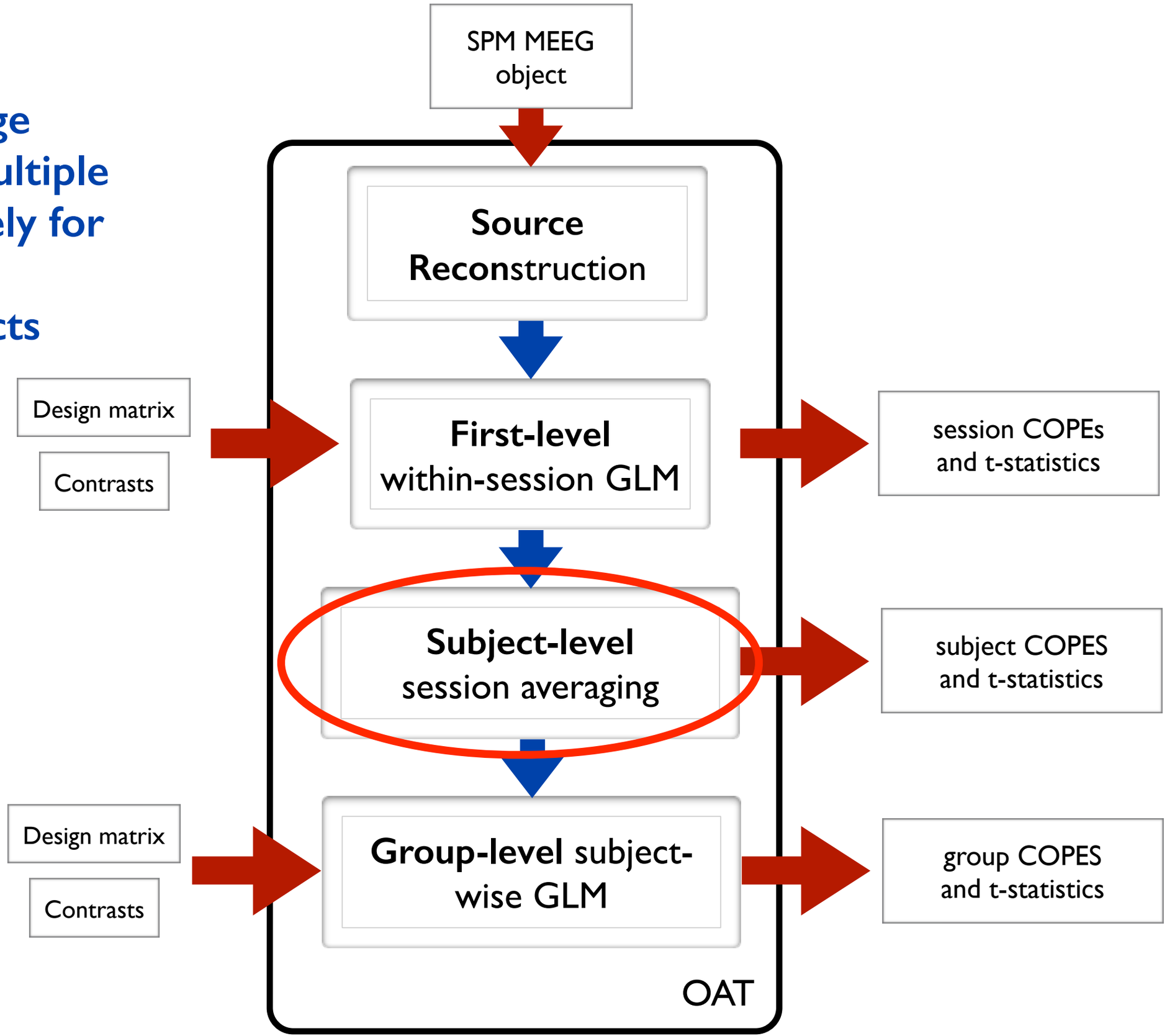
Multiple Session Analysis

First-level (session) COPEs are the Subject-level Inputs



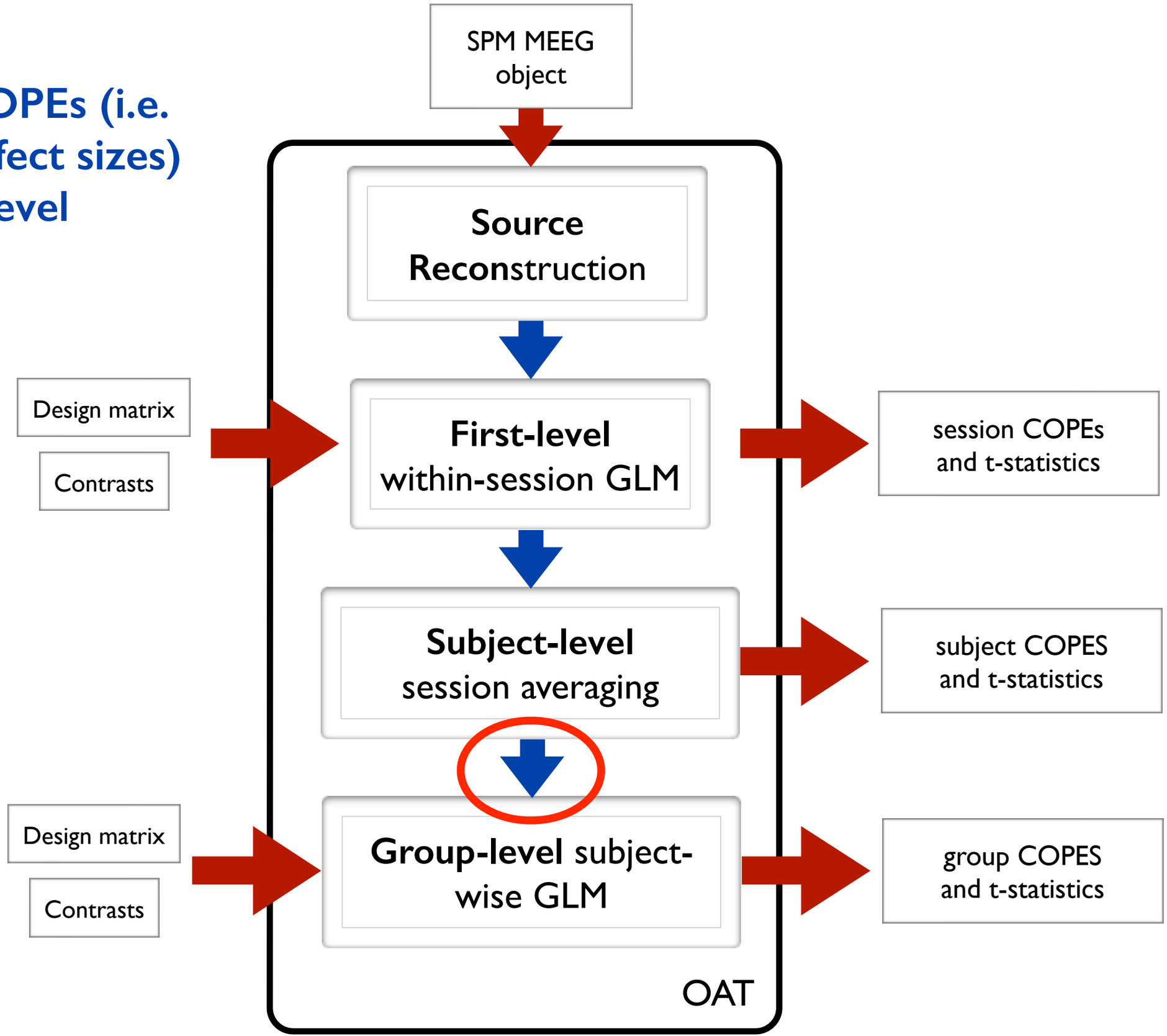
Multiple Session Analysis

Subject-level stage averages over multiple sessions separately for each subject
- using fixed effects



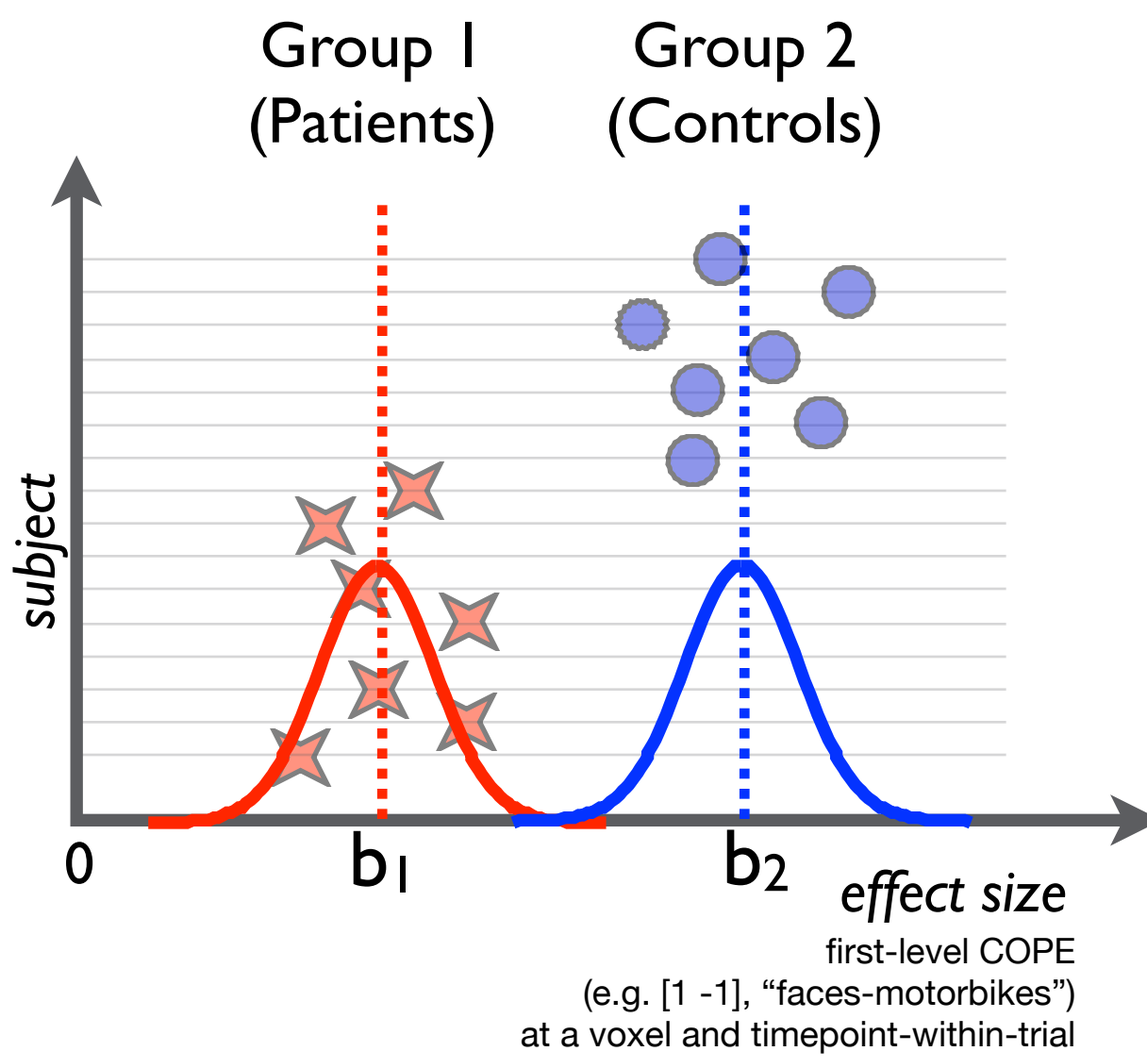
Multiple Subject Analysis

Subject-level COPEs (i.e. mean subject effect sizes) are the Group-level Inputs



Subject-wise GLM (Multiple Regression)

- We have two groups (e.g. 7 patients, 7 controls)

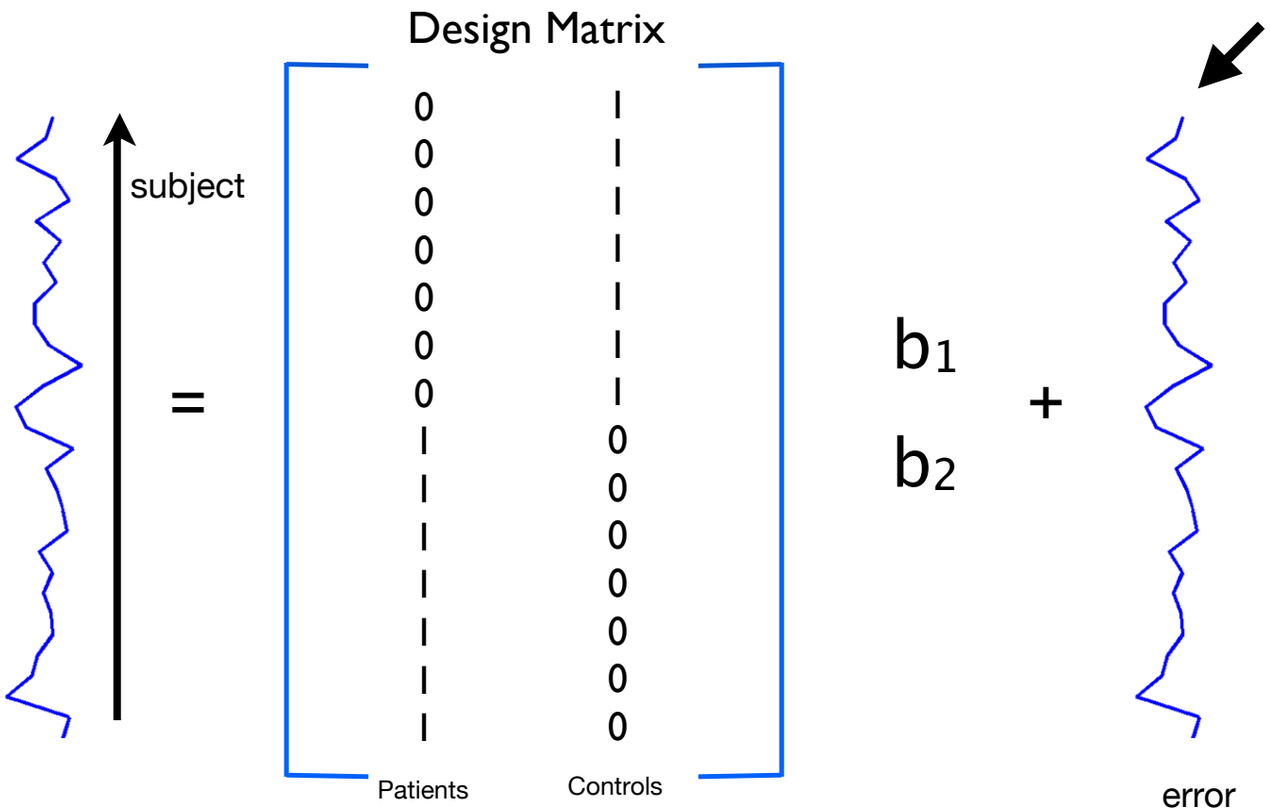


first-level COPE
(e.g. [1 -1], “faces-motorbikes”)
at a voxel and timepoint-within-trial

Subject-wise GLM (Multiple Regression)

- We have two groups (e.g. 7 patients, 7 controls)

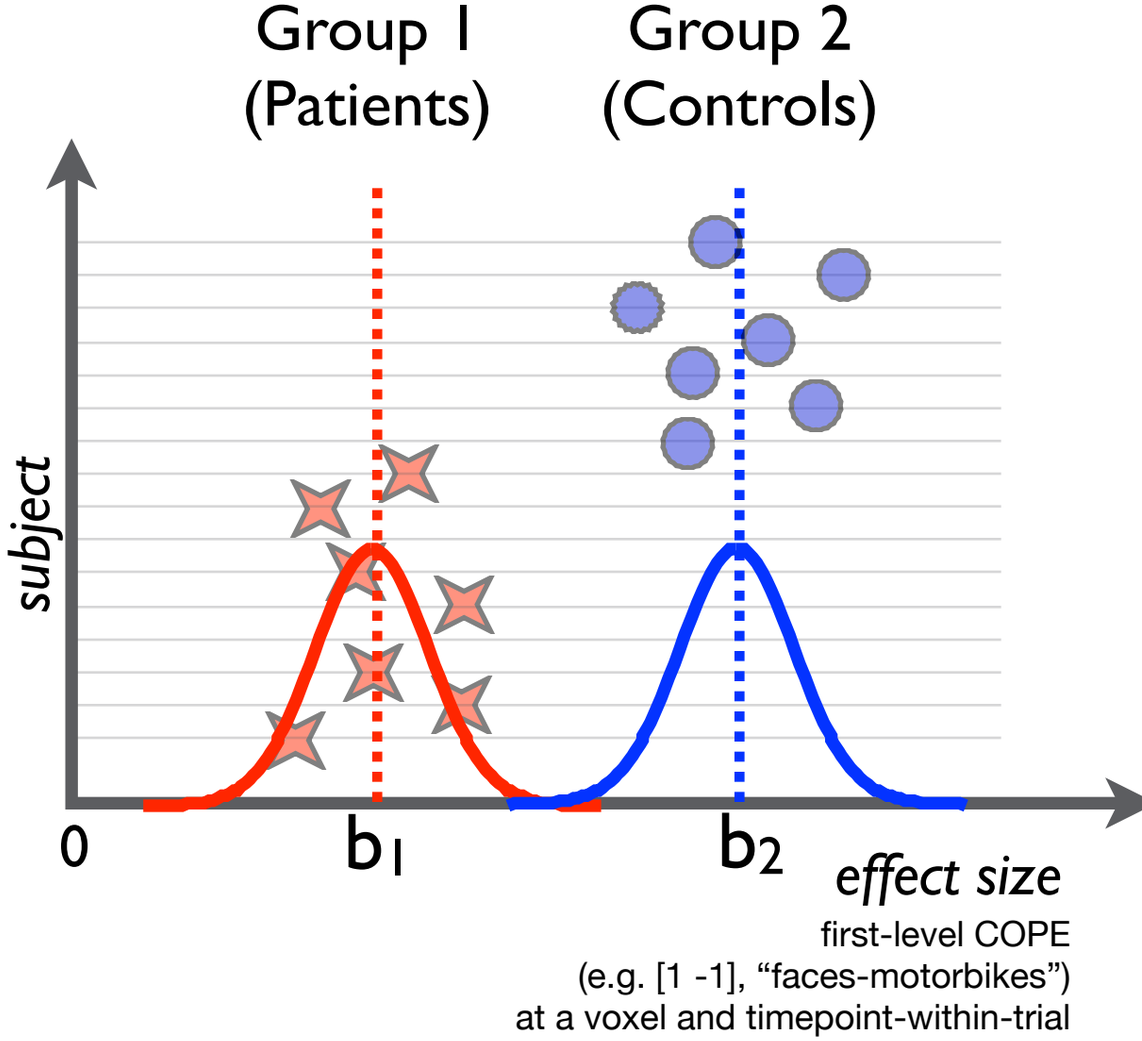
We model the between-subject variance giving inference on the populations (i.e. mixed-effects)



$Y = X b + e$

first-level COPE (e.g. [1 -1], "faces-motorbikes") at a voxel and timepoint-within-trial

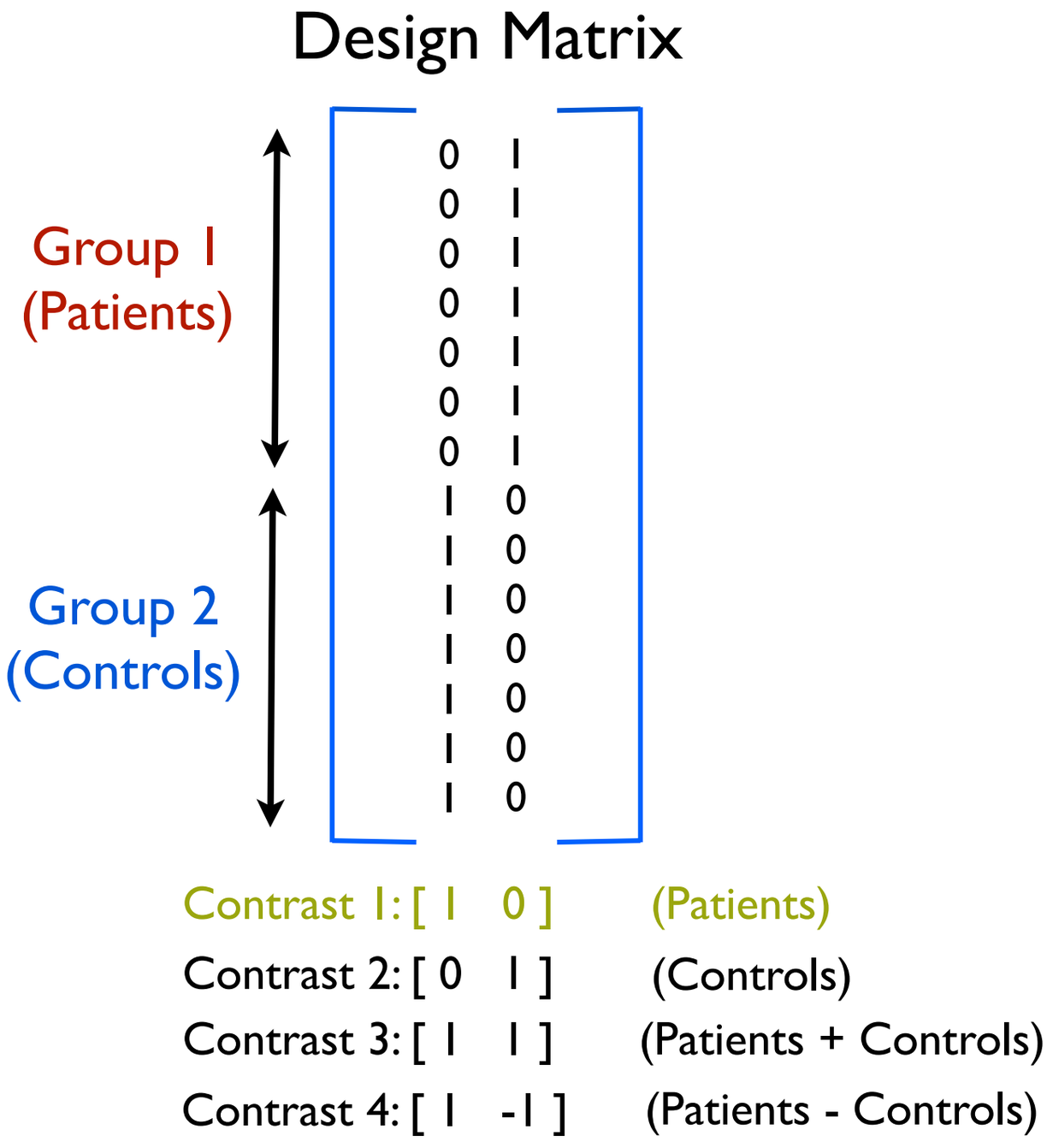
$Y=Xb+e$

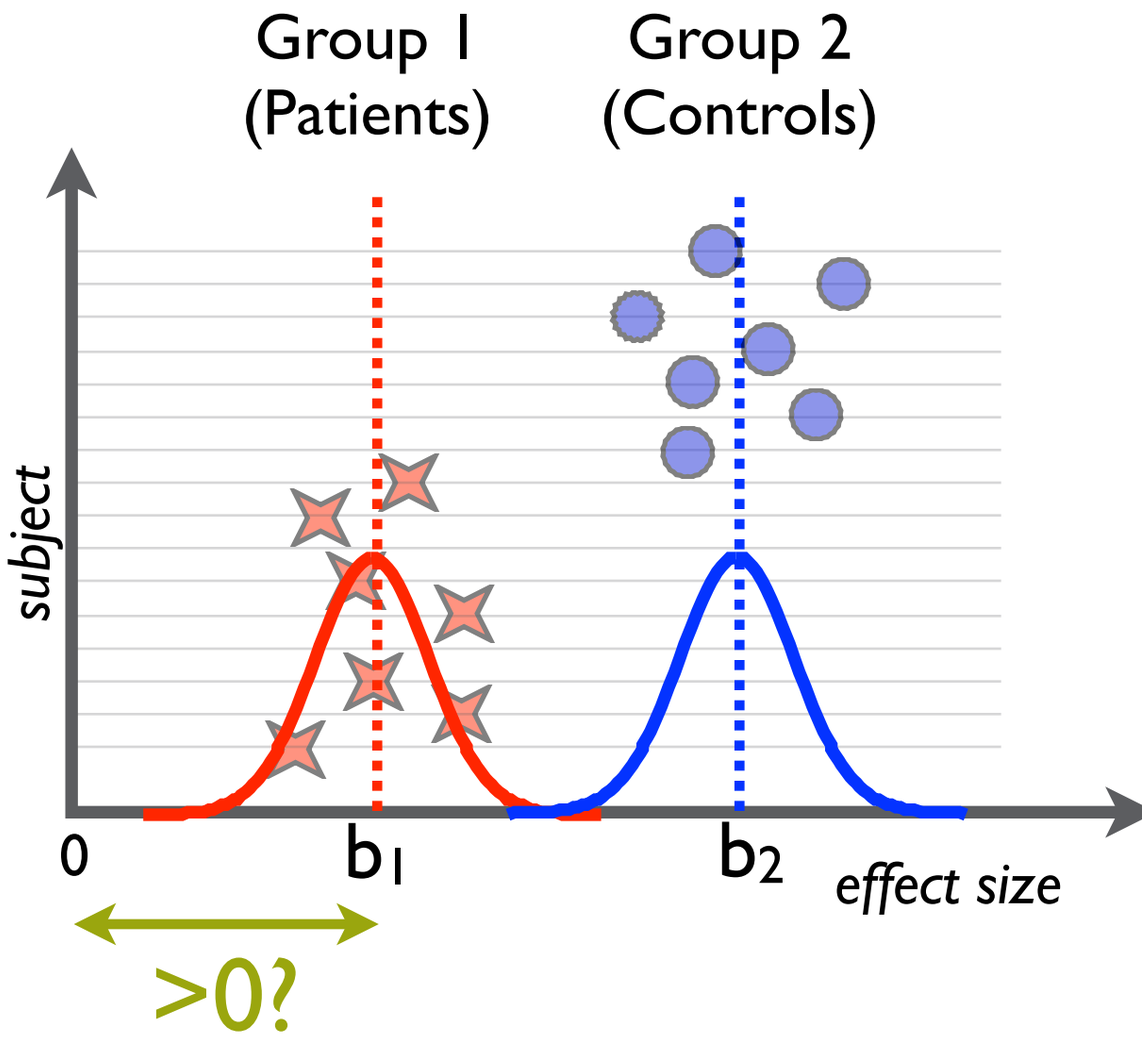


first-level COPE (e.g. [1 -1], "faces-motorbikes") at a voxel and timepoint-within-trial

Subject-wise GLM (Multiple Regression)

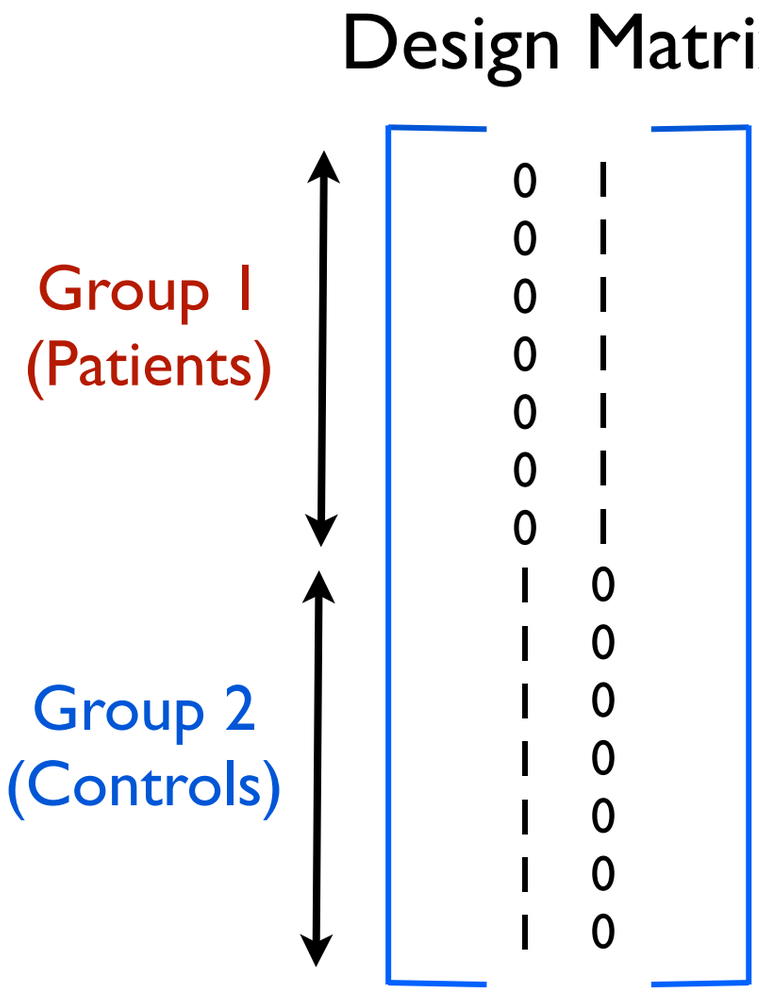
- We have two groups (e.g. 7 patients, 7 controls)





Subject-wise GLM (Multiple Regression)

- We have two groups (e.g. 7 patients, 7 controls)

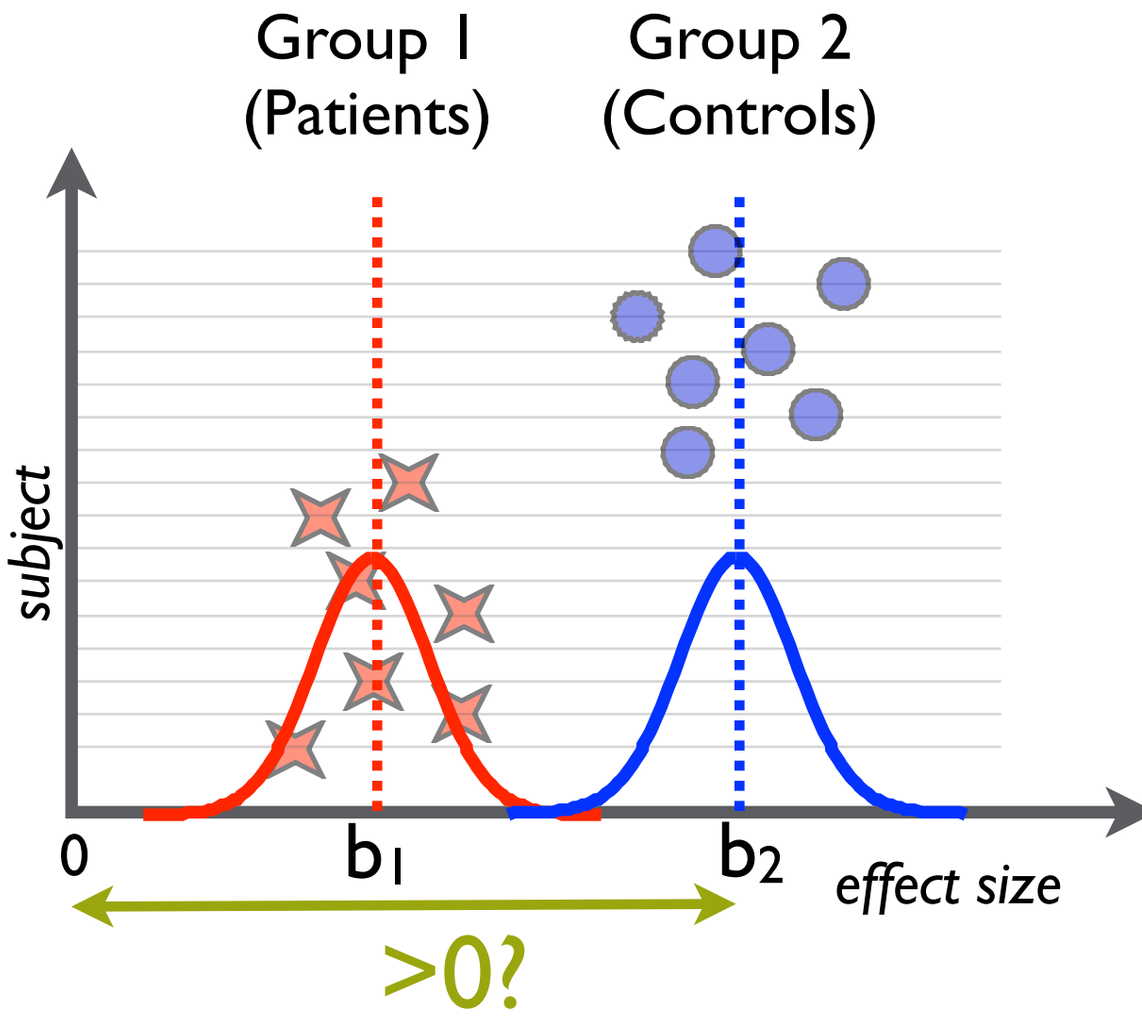


Contrast 1: [1 0] (Patients)

Contrast 2: [0 1] (Controls)

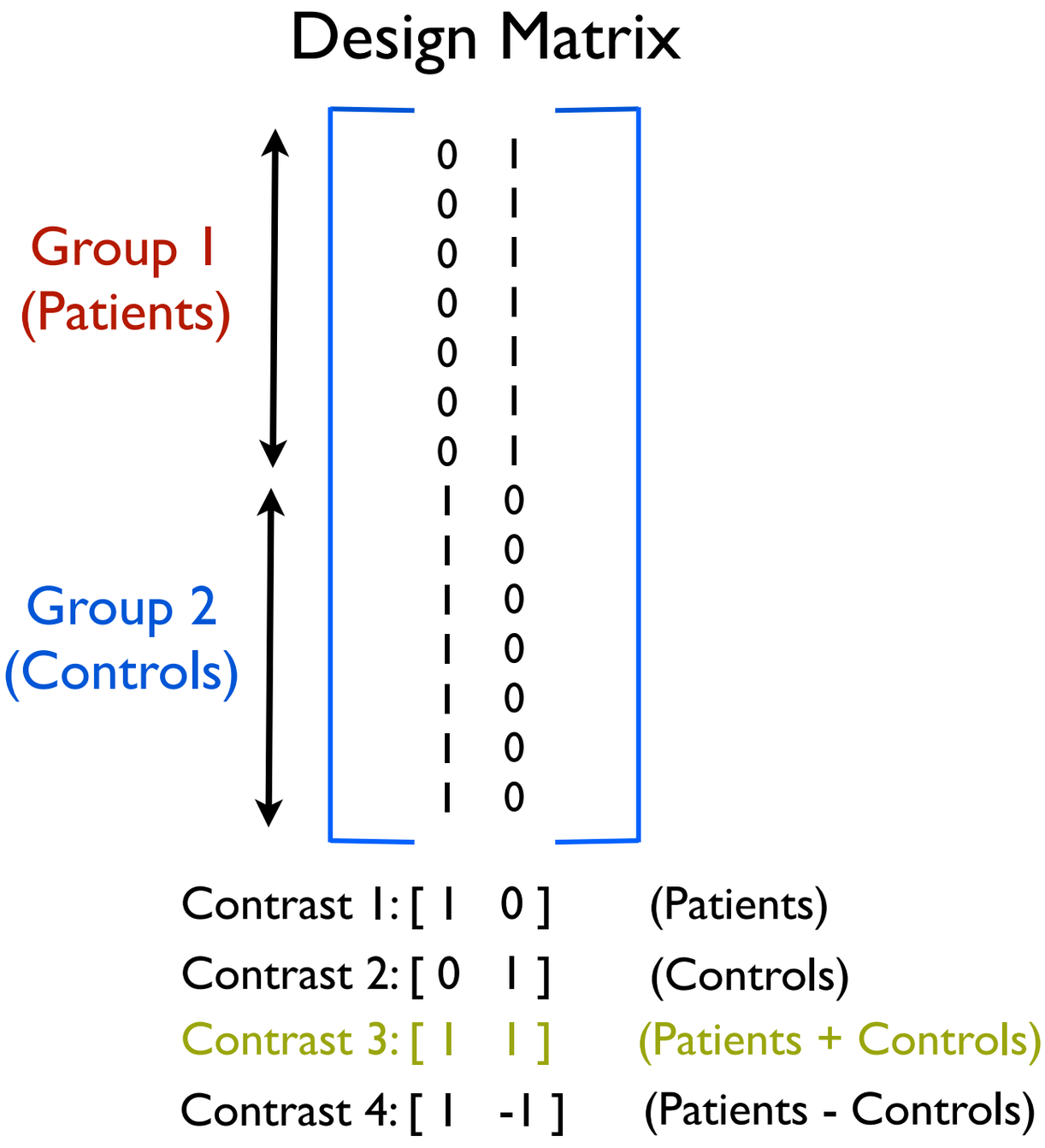
Contrast 3: [1 1] (Patients + Controls)

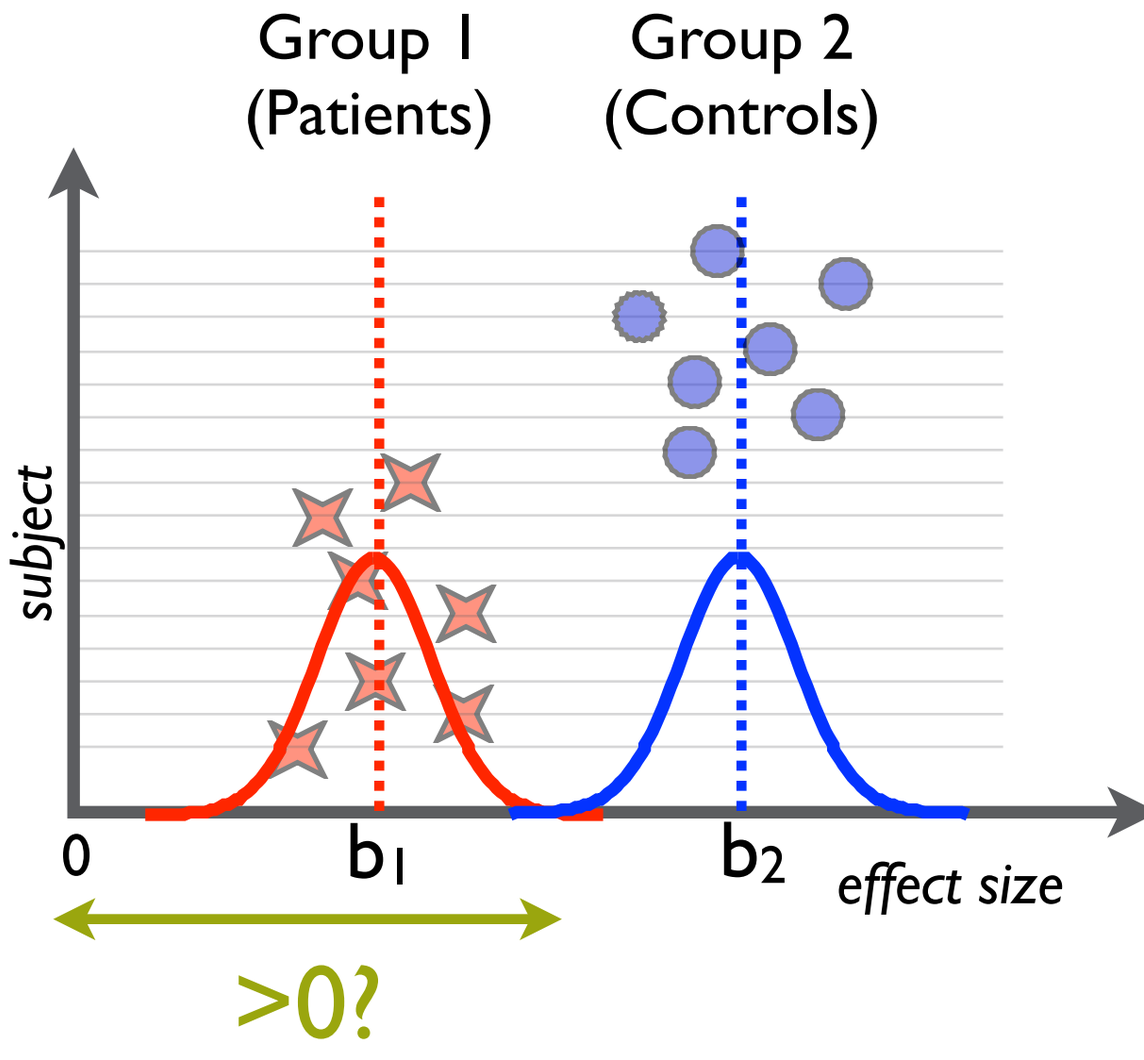
Contrast 4: [1 -1] (Patients - Controls)



Subject-wise GLM (Multiple Regression)

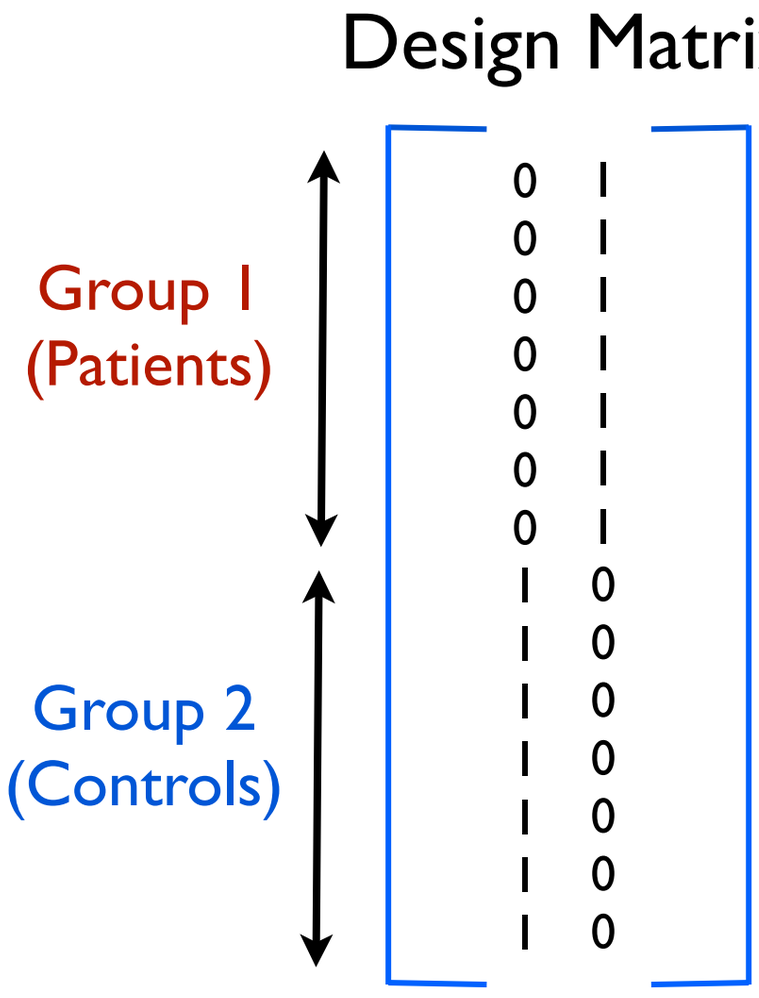
- We have two groups (e.g. 7 patients, 7 controls)





Subject-wise GLM (Multiple Regression)

- We have two groups (e.g. 7 patients, 7 controls)

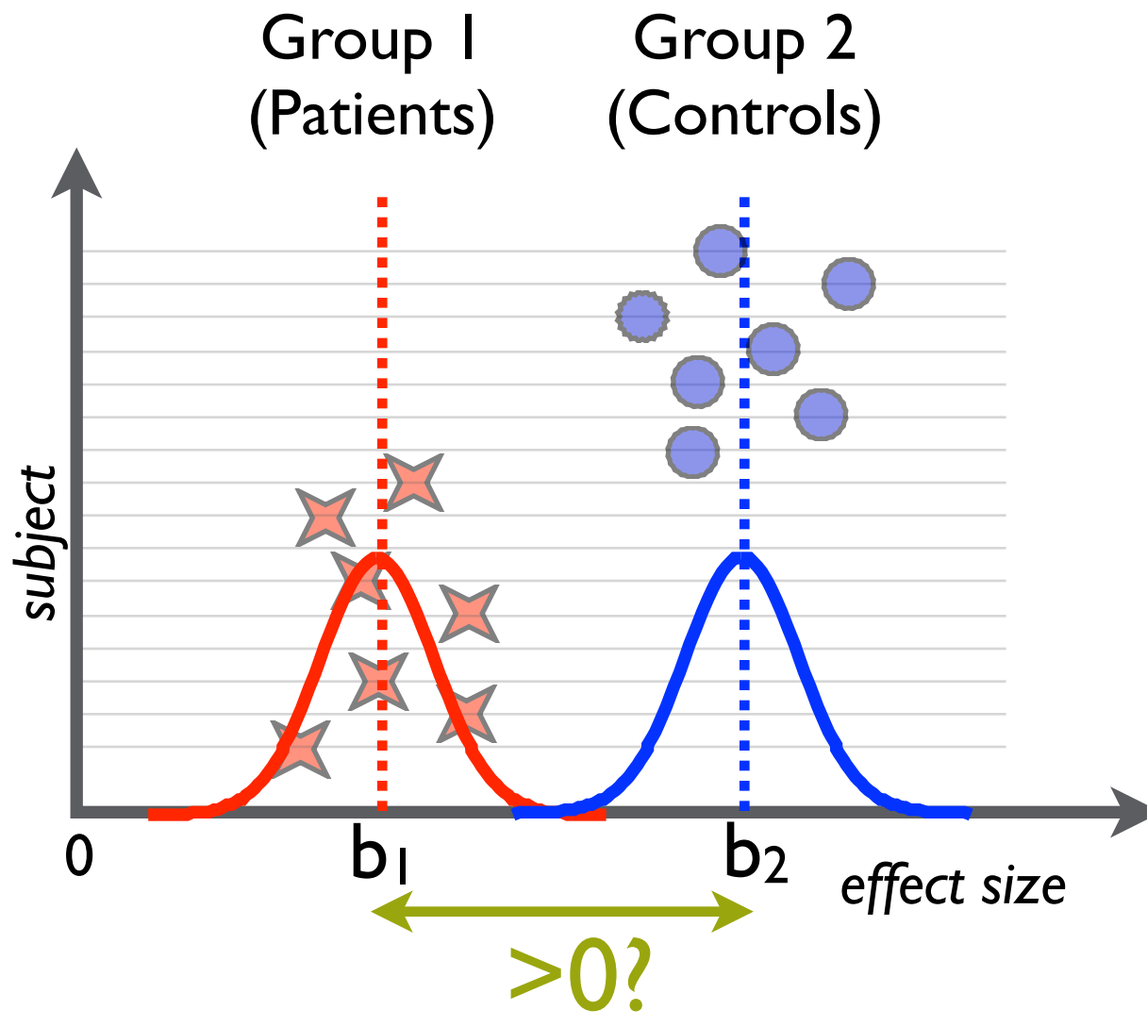


Contrast 1: [1 0] (Patients)

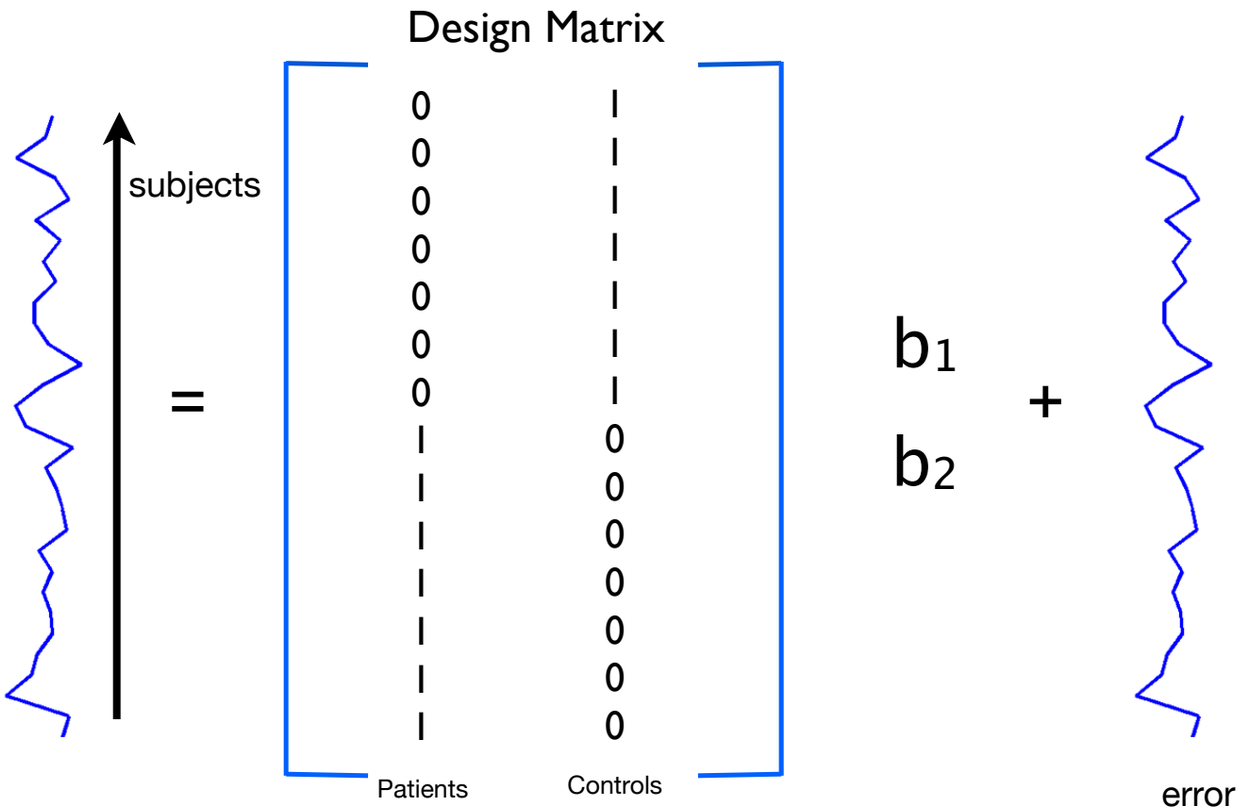
Contrast 2: [0 1] (Controls)

Contrast 3: [1 1] (Patients + Controls)

Contrast 4: [1 -1] (Patients - Controls)



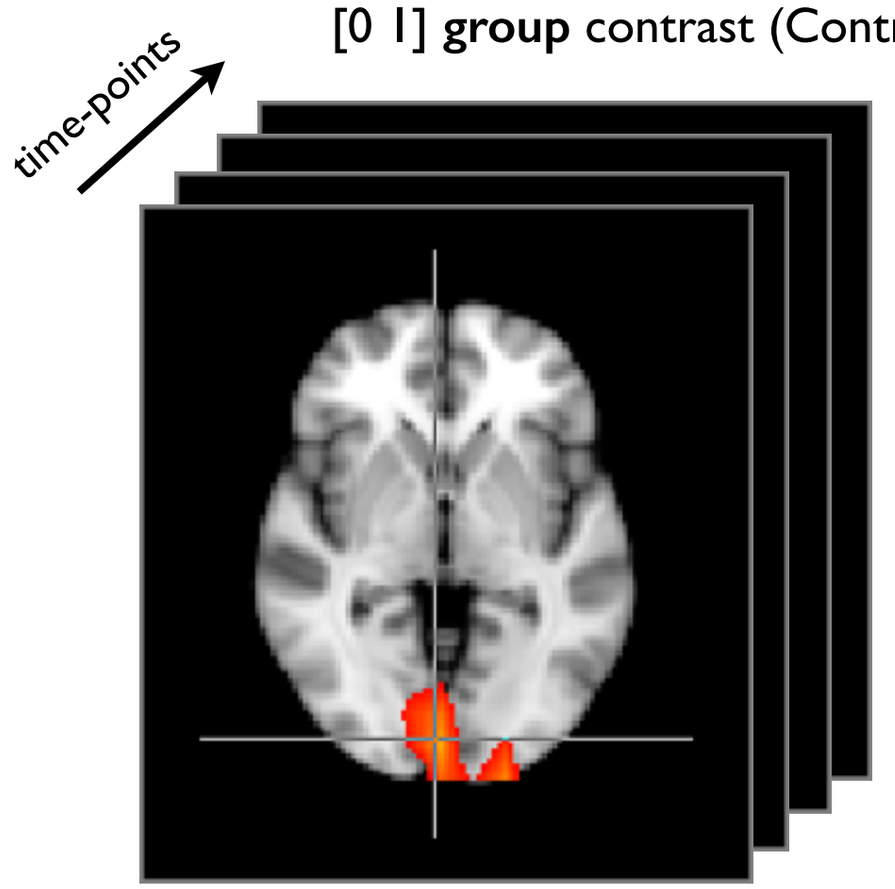
Subject-wise GLM (Multiple Regression)



Y = X b + e

First-level COPE (e.g. [1 -1], “faces-motorbikes”) at a voxel and timepoint-within-trial

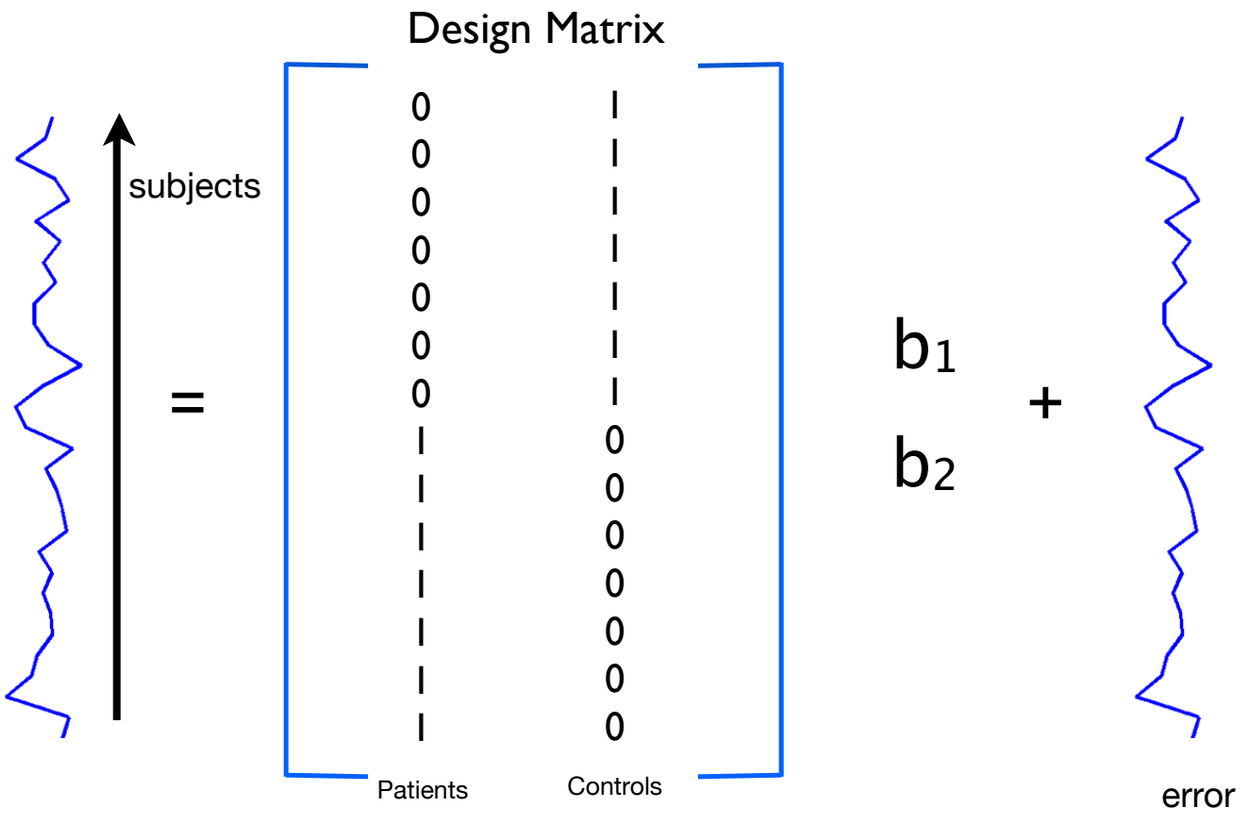
[0 1] **group** contrast (Controls)

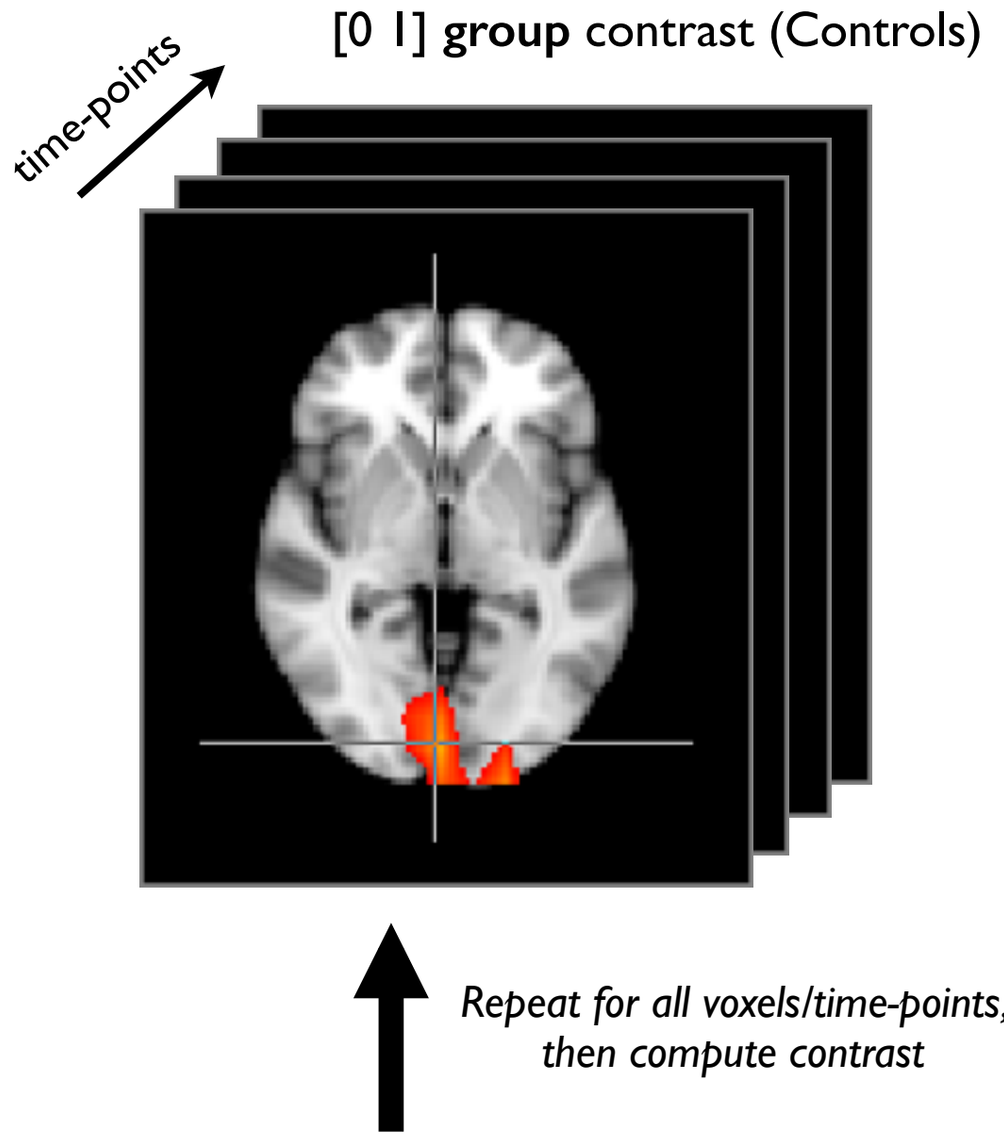


Repeat for all voxels/time-points, then compute contrast

Y=Xb+e → Fit GLM

Subject-wise GLM (Multiple Regression)



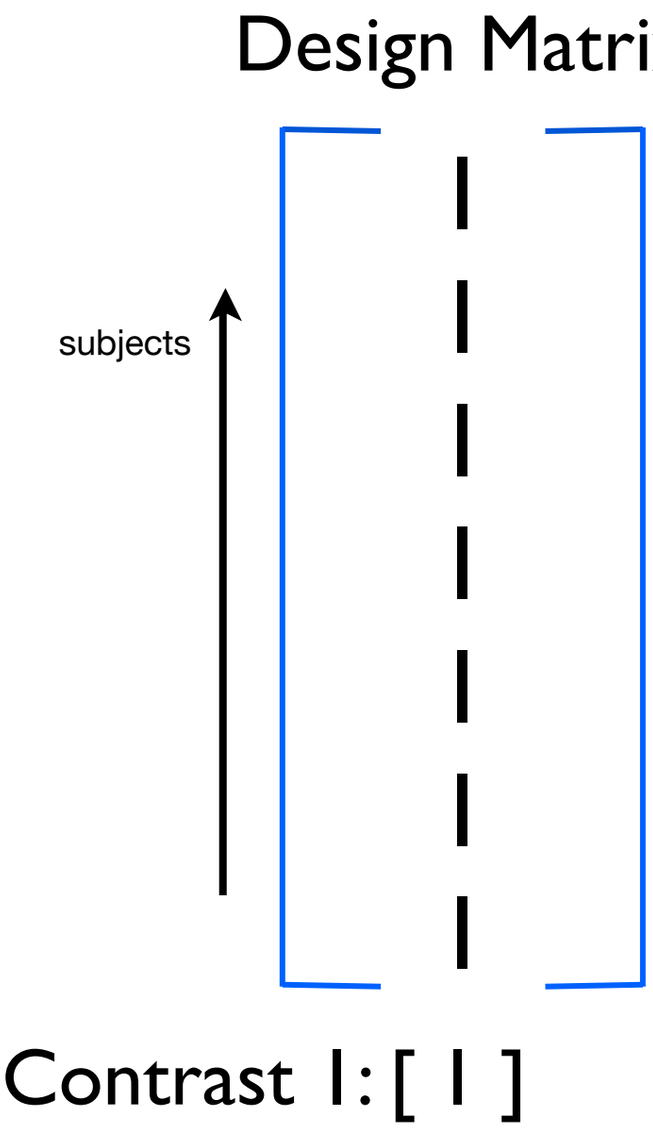


Y=Xb+e → Fit GLM

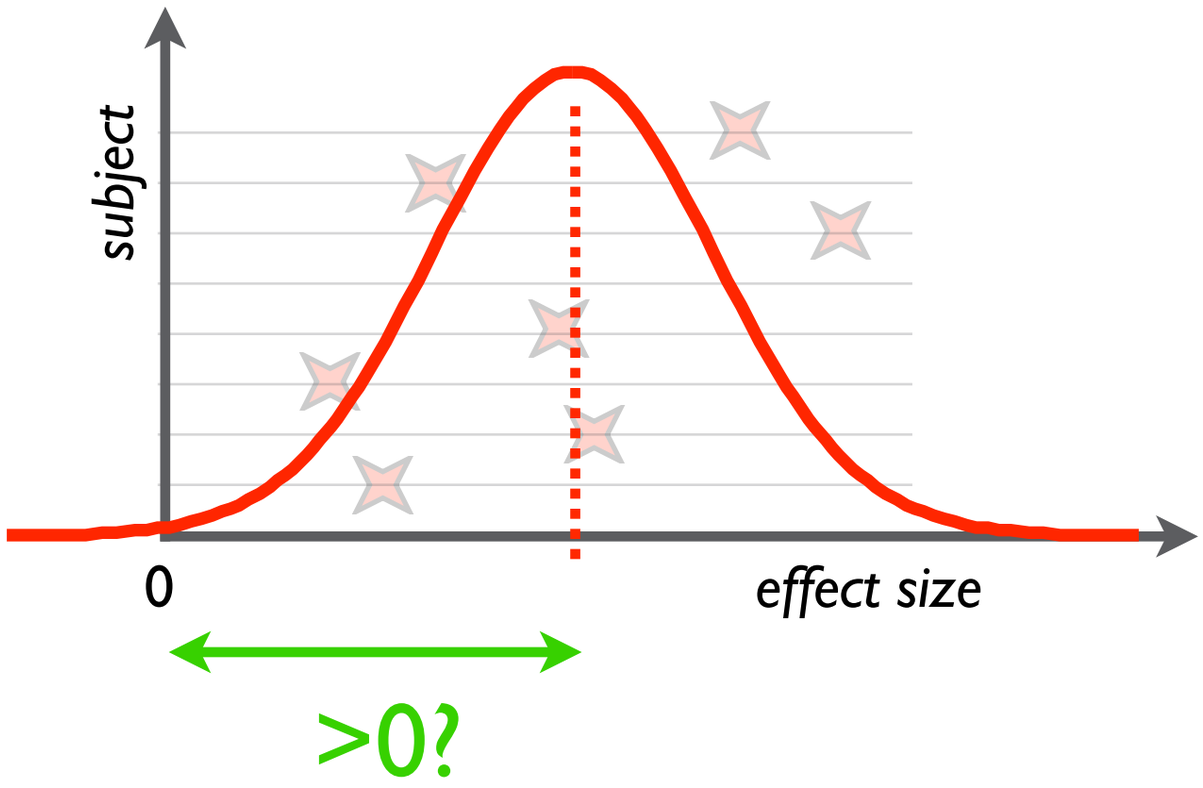
- **The whole group analysis can be repeated for ANY first-level contrast**

Single Group Average

- We have 7 subjects - all in one group - and want the mean group average:

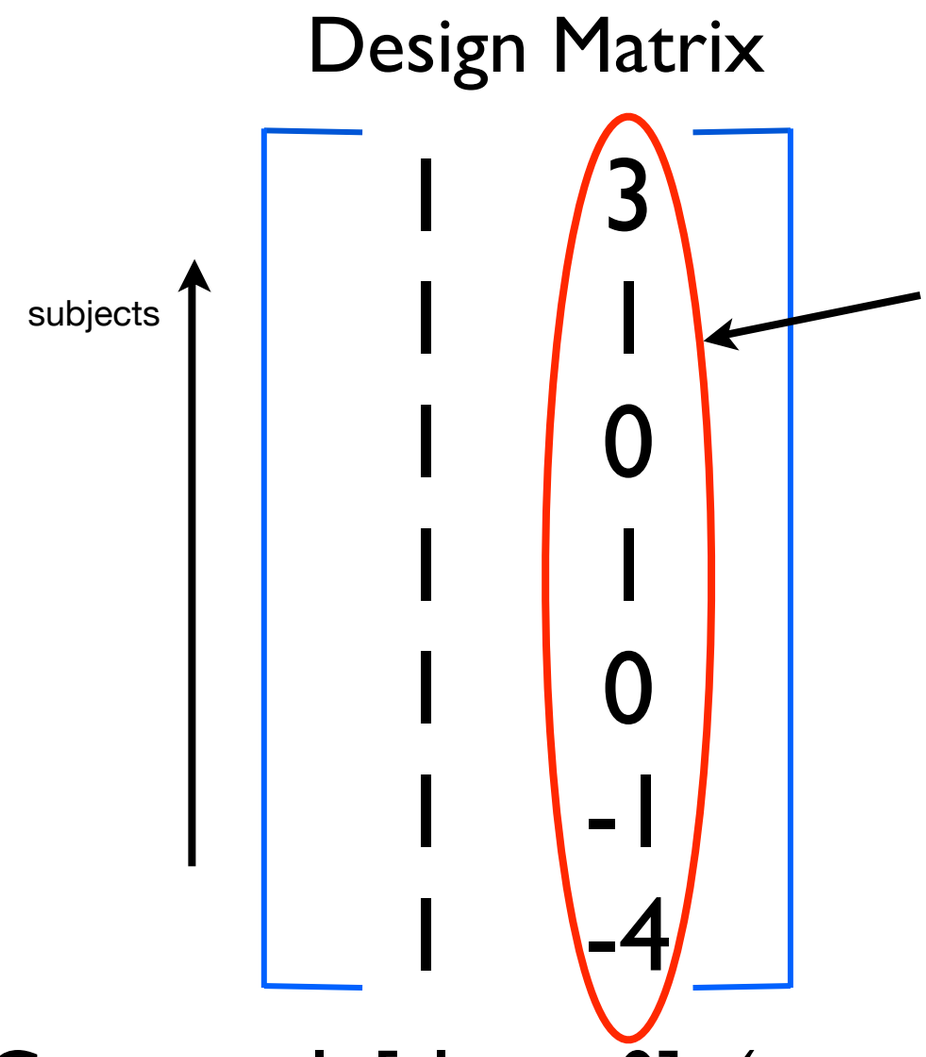


Does the group activate on average?



Behavioural Variables

- We have 7 subjects - all in one group - and want to investigate a behavioural effect



- Behavioural scores for each subject
- (Make sure this regressor is demeaned if you want to interpret the first regression parameter as the group average)

Contrast 1: [1 0] (group average)

Contrast 2: [0 1] (behavioural effect)

Other things you can do with the Group GLM

- Paired t-tests
- Behavioural Regressors
- Interactions
- For more see:
  - *http://www.fmrib.ox.ac.uk/fslcourse/lectures/inference.pdf*

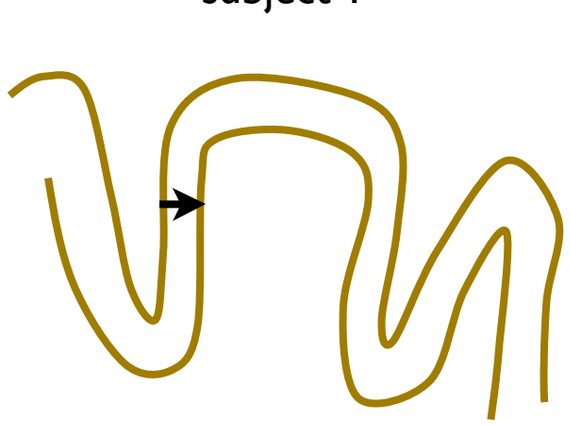
ERF rectification

- Problem:

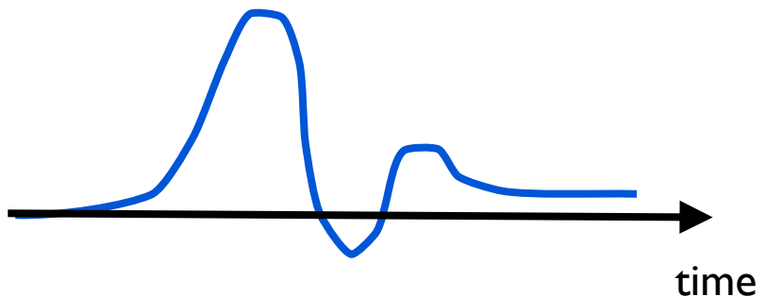
➡ there is an ***ambiguity*** between the reconstructed dipole direction and the sign of the reconstructed time series

➡ not trivial to resolve this, e.g.

subject 1

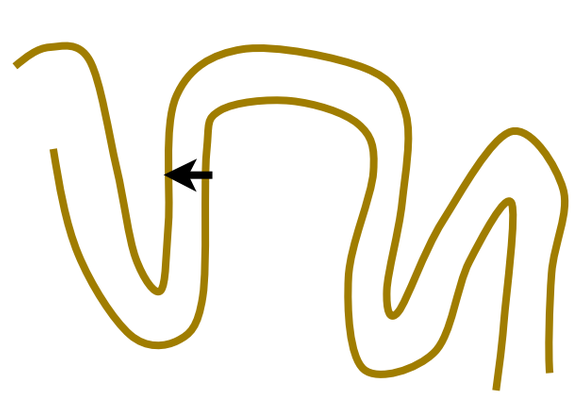


raw COPE estimate:

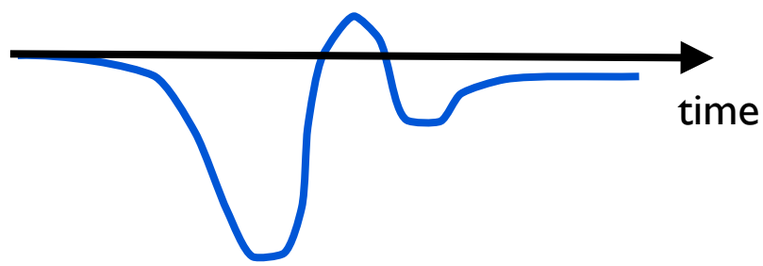


will cancel out

subject 2



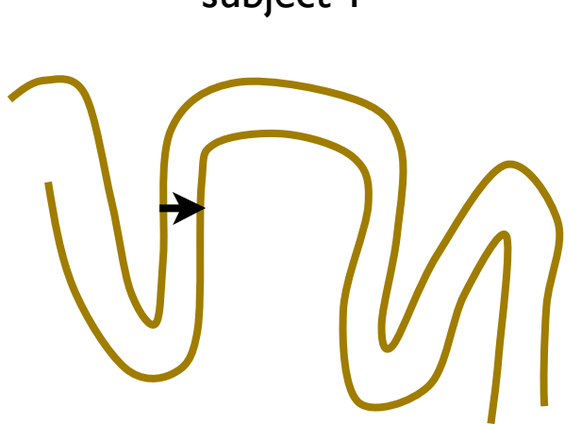
raw COPE estimate:



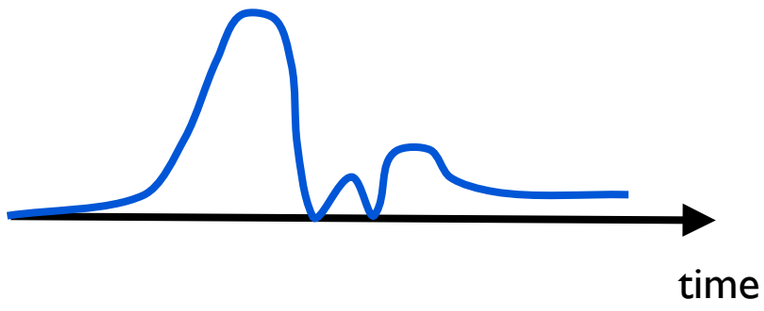
ERF rectification

- Solution: use abs(COPE)

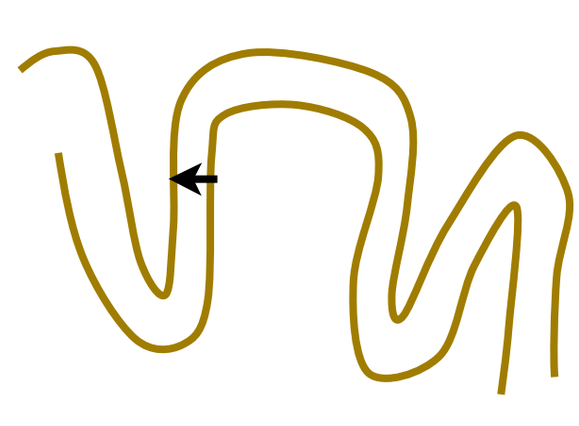
subject 1



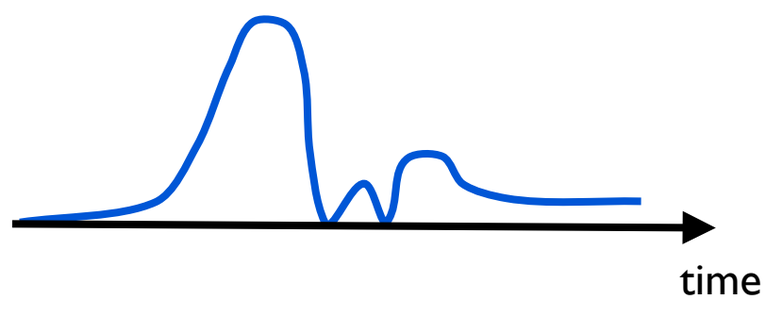
abs(COPE) estimate:



subject 2



abs(COPE) estimate:



Multiple Comparison Problem

- We could carry out a null hypothesis test for each voxel using a chosen false positive rate (FPR)
- However, if we used FPR=0.05 with 20,000 brain voxels then we would get 1000 FPs

Multiple Comparison Problem

- We could carry out a null hypothesis test for each voxel using a chosen false positive rate (FPR)
- However, if we used FPR=0.05 with 20,000 brain voxels then we would get 1000 FPs
- Not good - we would really like the FPR to correspond to the probability of getting one FP in the entire brain

Bonferroni Correction

- Bonferroni: divide uncorrected p-threshold by number of voxels before thresholding

- E.g. 20,000 brain voxels:

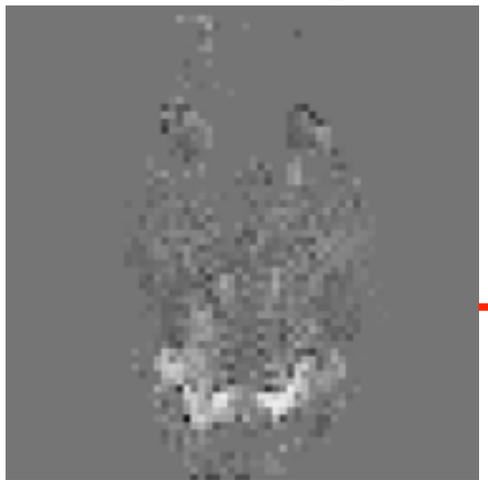
an uncorrected p-threshold of 0.05 becomes

0.05 / 20,000 = 0.0000025

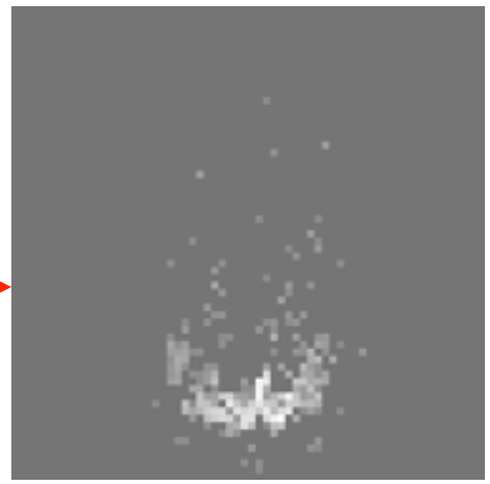
Thresholding: Clustering

What about testing significance on clusters?

stat image



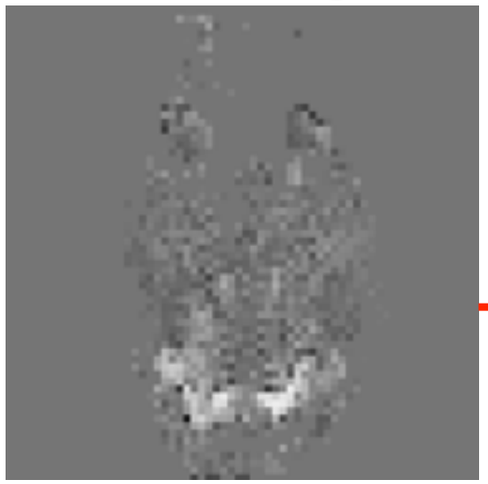
Threshold at
(arbitrary!) level



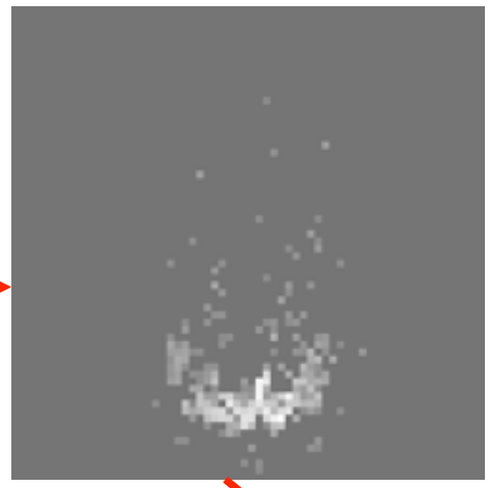
Thresholding: Clustering

What about testing significance on clusters?

stat image



Threshold at (arbitrary!) level



- Form clusters from surviving voxels.

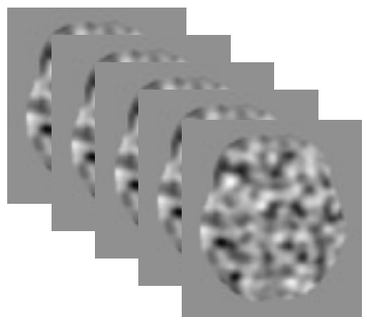
- **BUT**, what is the probability of getting a cluster, given its spatial extent (no. of voxels in cluster) and threshold, under the null hypothesis?
- Can use Random Field Theory, but involves dodgey assumptions, instead ...

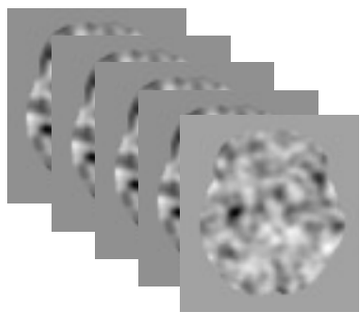
Permutation Testing

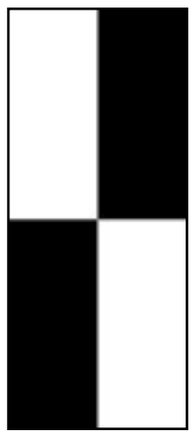
We can record the cluster sizes from our dataset

Group 1

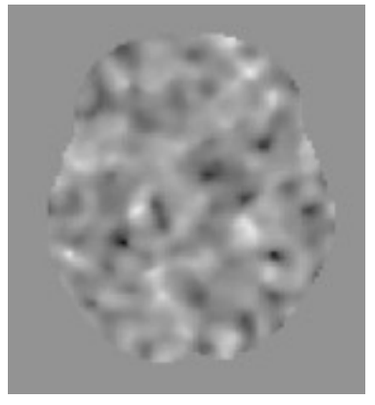
Group 2







Our correct group model

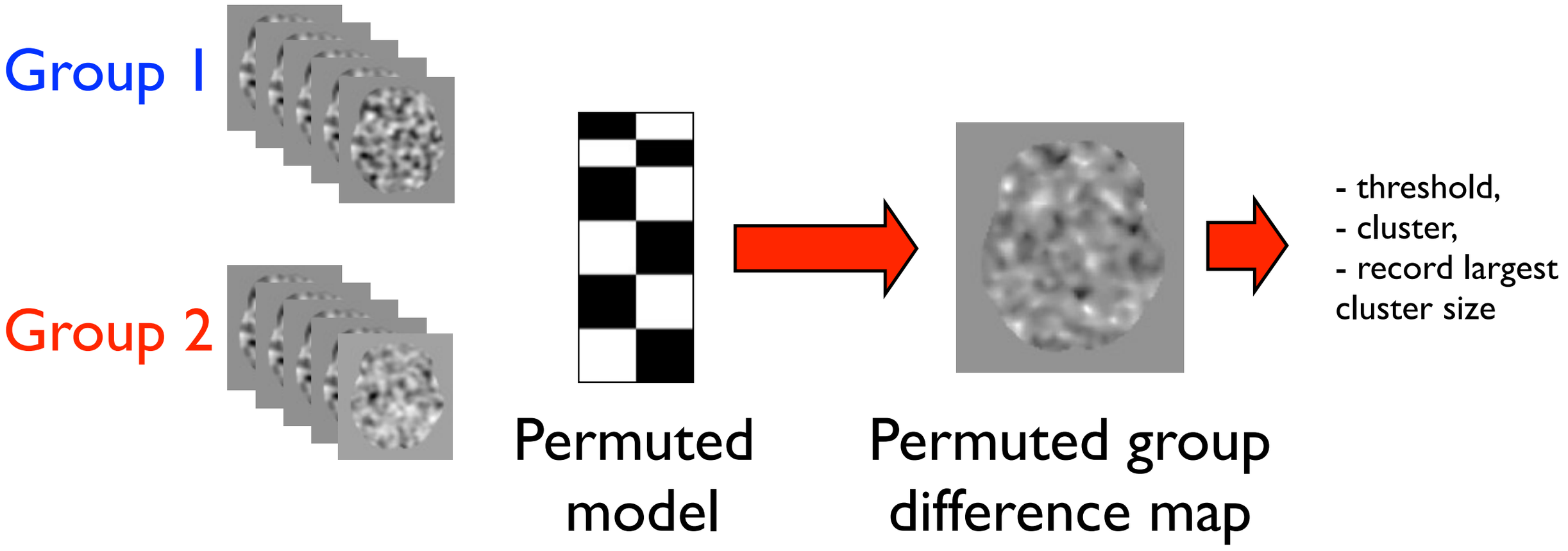


Our actual group difference map

- threshold,
- cluster,
- record **our actual** cluster sizes (but are they significant?)

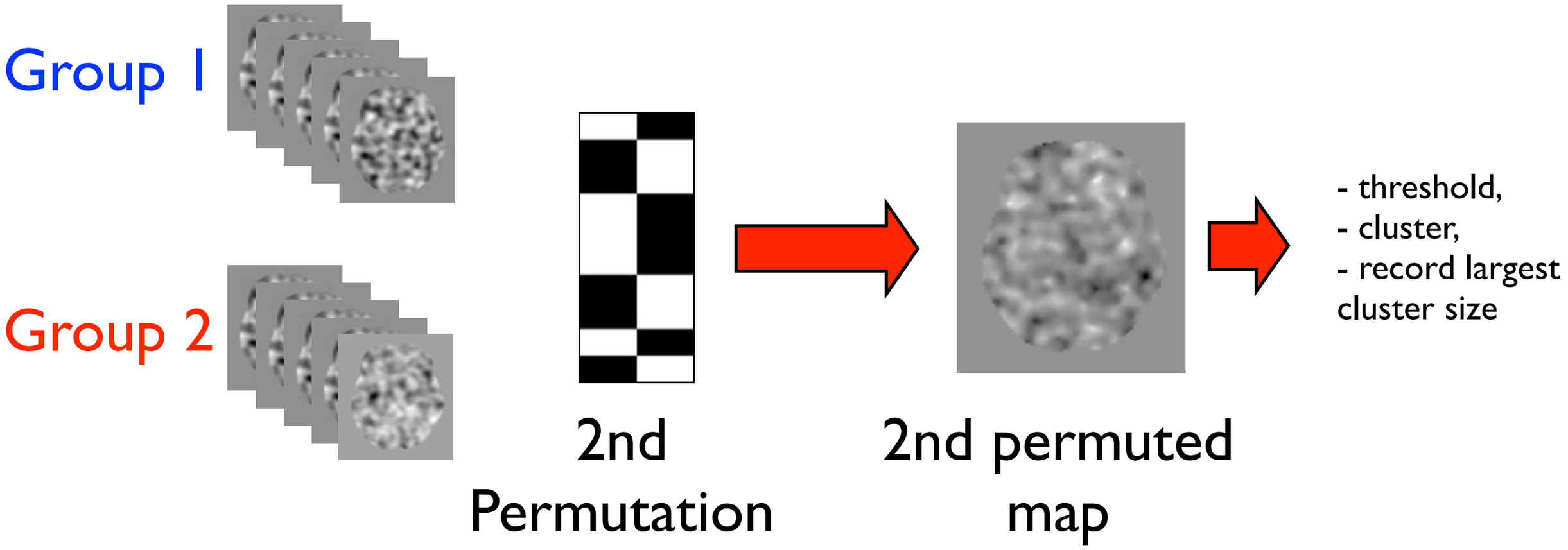
Permutation Testing

We can then permute the design matrix group labellings to get the null distribution of the maximum cluster size



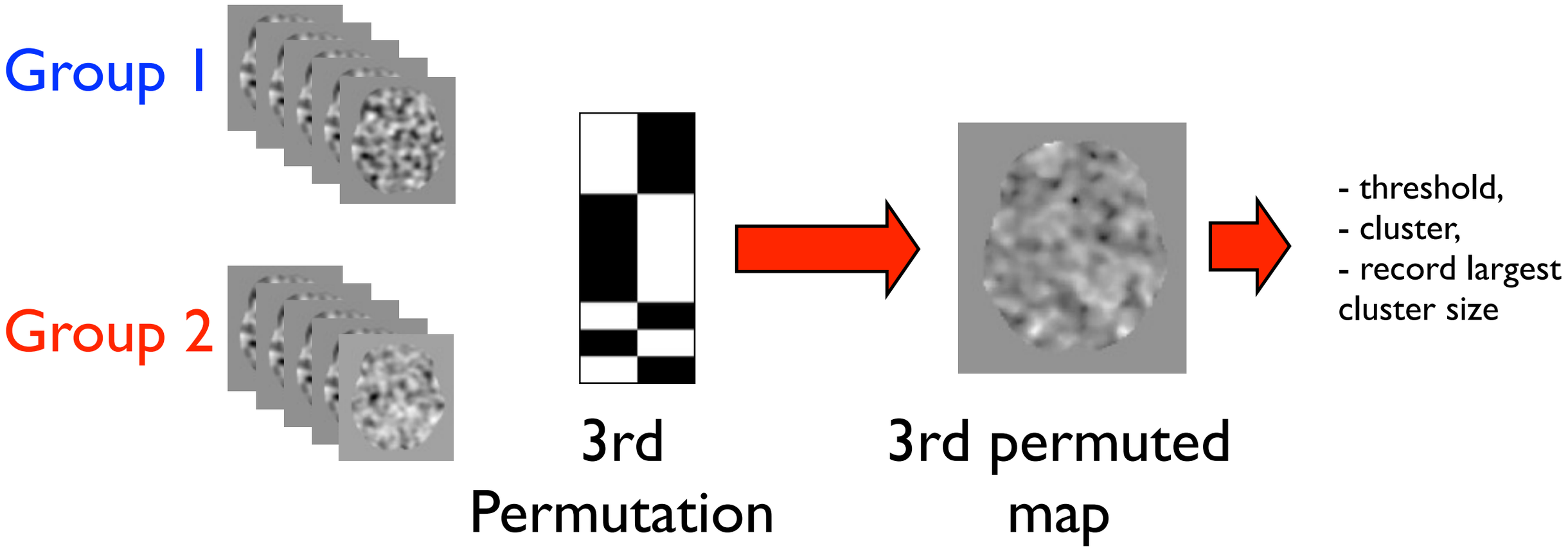
Permutation Testing

We can then permute the design matrix group labellings to get the null distribution of the maximum cluster size



Permutation Testing

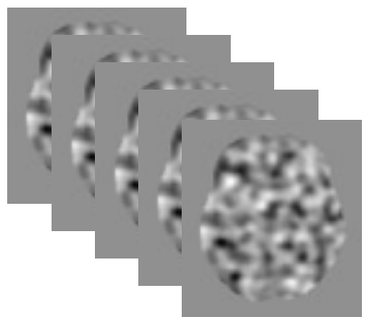
We can then permute the design matrix group labellings to get the null distribution of the maximum cluster size



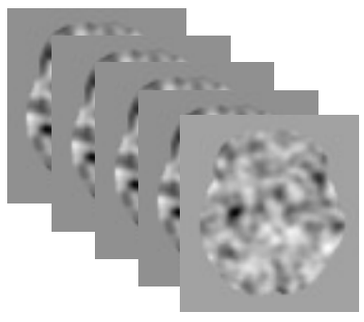
Permutation Testing

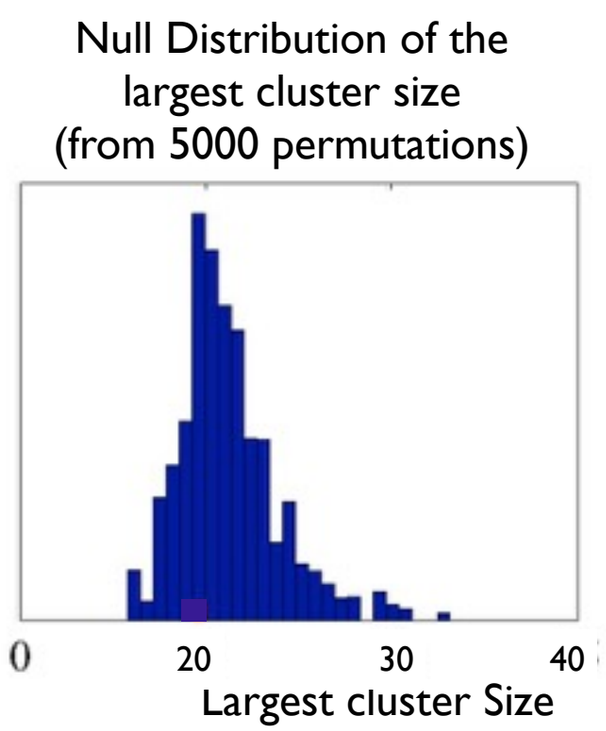
We can then permute the design matrix group labellings to get the null distribution of the maximum cluster size

Group 1



Group 2

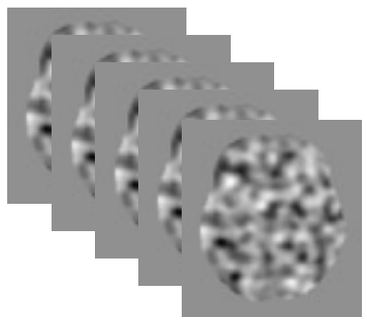




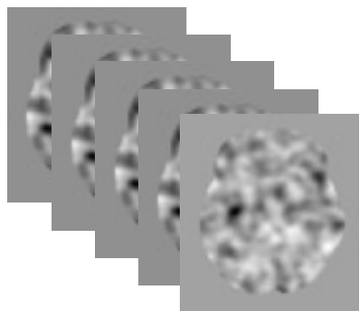
Permutation Testing

We can then permute the design matrix group labellings to get the null distribution of the maximum cluster size

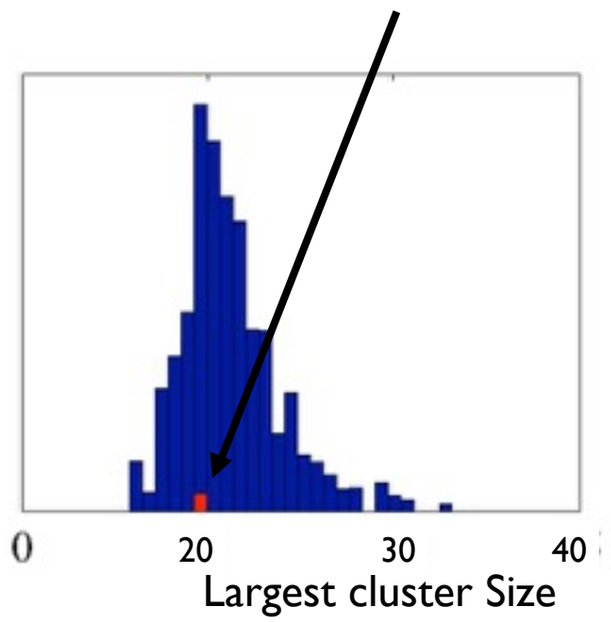
Group 1



Group 2



Cluster size from our correct group model



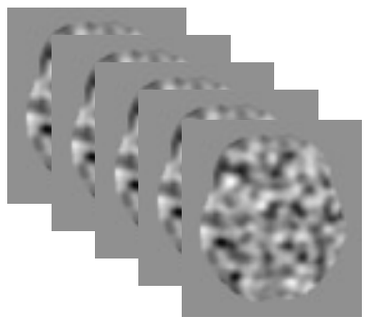
- 3925 permutations yielded larger clusters than original group labellings.

- We CANNOT reject the null-hypothesis in this case

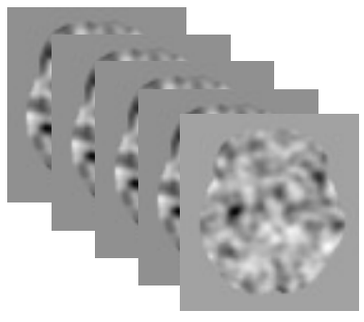
Permutation Testing

We can then permute the design matrix group labellings to get the null distribution of the maximum cluster size

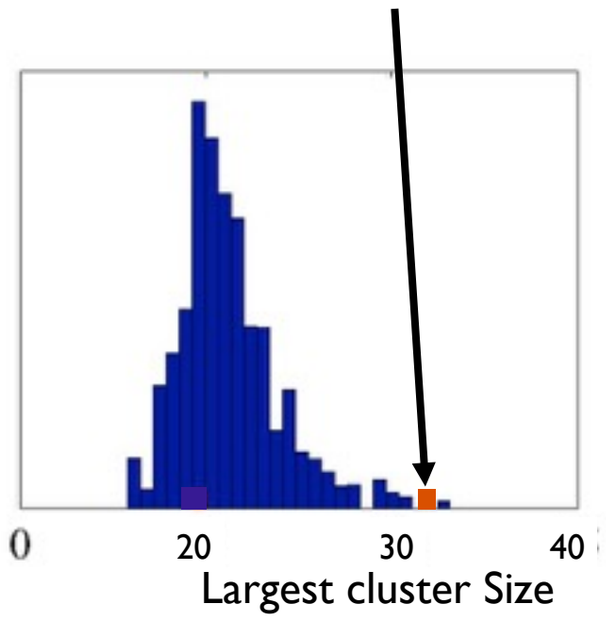
Group 1



Group 2



Cluster size from our correct group model



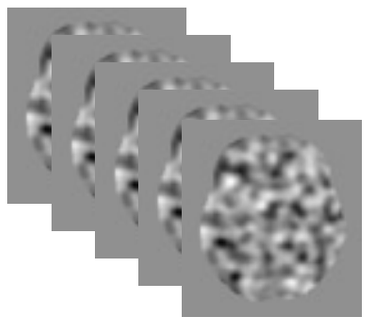
- Only 5 permutations yielded larger clusters than original group labellings:
[Corrected P-Value=5/5000=0.001]

- We CAN reject the null-hypothesis in this case

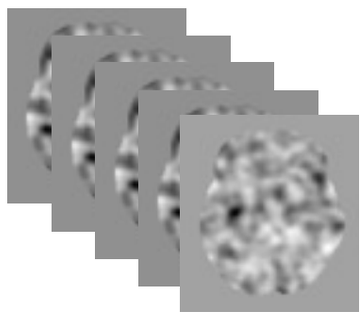
Permutation Testing

We can then permute the design matrix group labellings to get the null distribution of the maximum cluster size

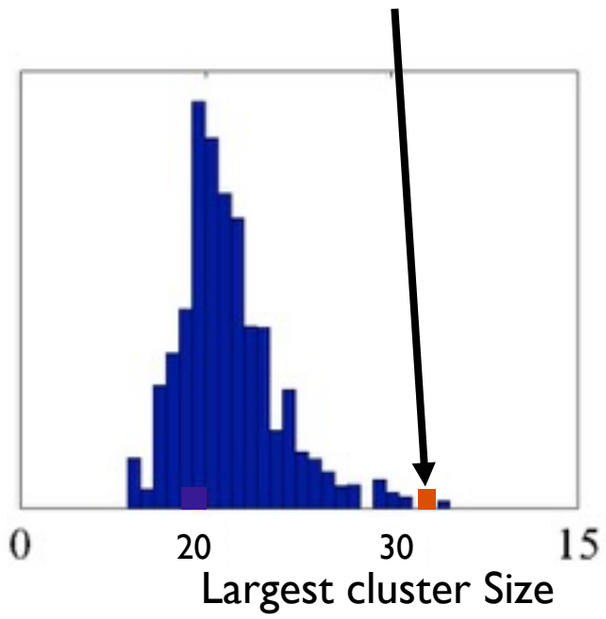
Group 1



Group 2



Cluster size from our correct group model



- OSL has the facility to test clusters in:

  - 1-D (time)
  - 2-D (time-frequency)
  - 2-D/3-D (space)
  - 3-D/4-D (space and time)

- Note: 4D is very slow!!

Variance Smoothing

- It is advantageous to smooth the group (between-subject) variance (VARCOPE)
- This preserves the effect size (COPE) spatial resolution, while increases the degrees of freedom for the VARCOPE estimation
- Permutation testing adjusts accordingly to keeps stats valid

$$\text{tstat} = \frac{\text{COPE}}{\text{sqrt(VARCOPE)}}$$

Practical

Beamformer group analysis, for which source_recon, first_level and subject_level OAT stages have already been run.

Includes:

a) Wholebrain (ERF) analysis
b) Spatial ROI analysis
c) Time window (spatial map) analysis
- Using (3D) permutation testing
d) ROI time-freq analysis using (2D) perm testing