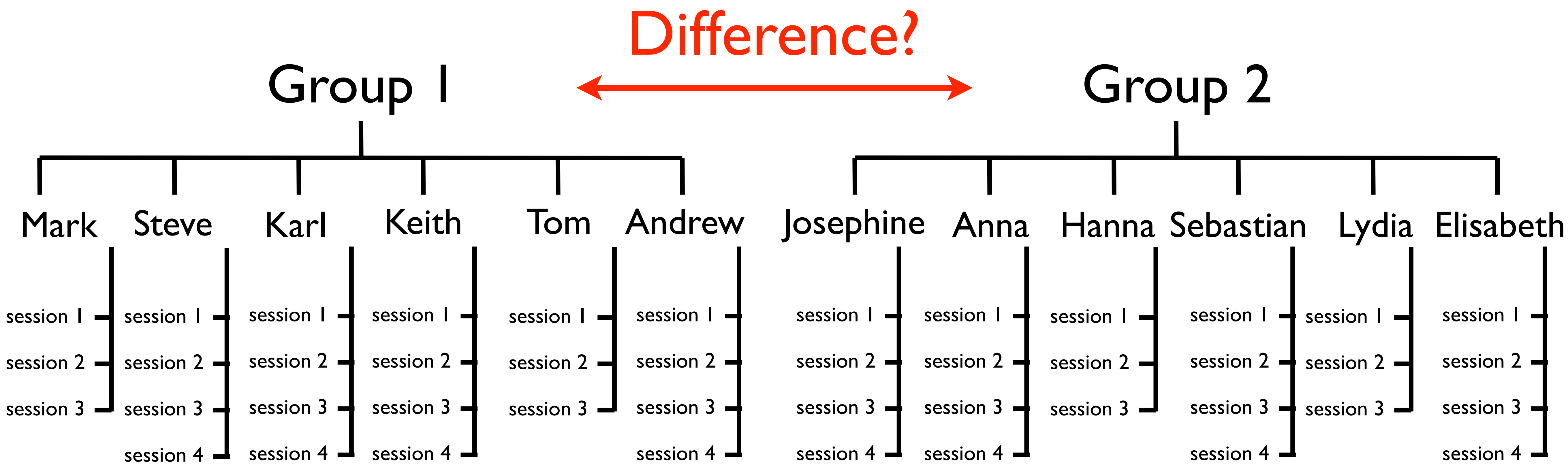
Group-level Analysis in OSL

OHBA Analysis Workshop

Group Analysis

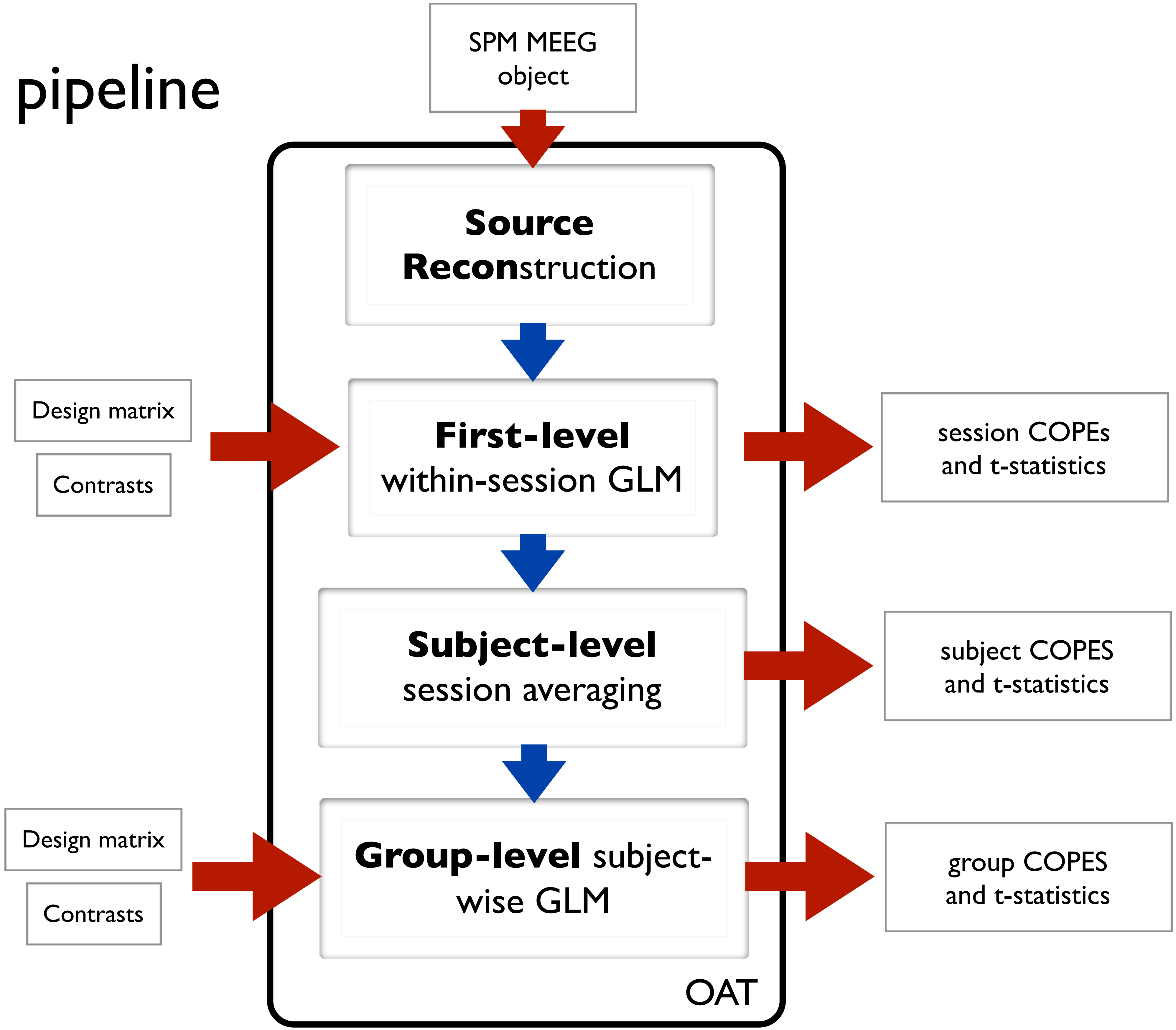
- typically need to infer across multiple subjects, sometimes multiple groups and/or multiple sessions



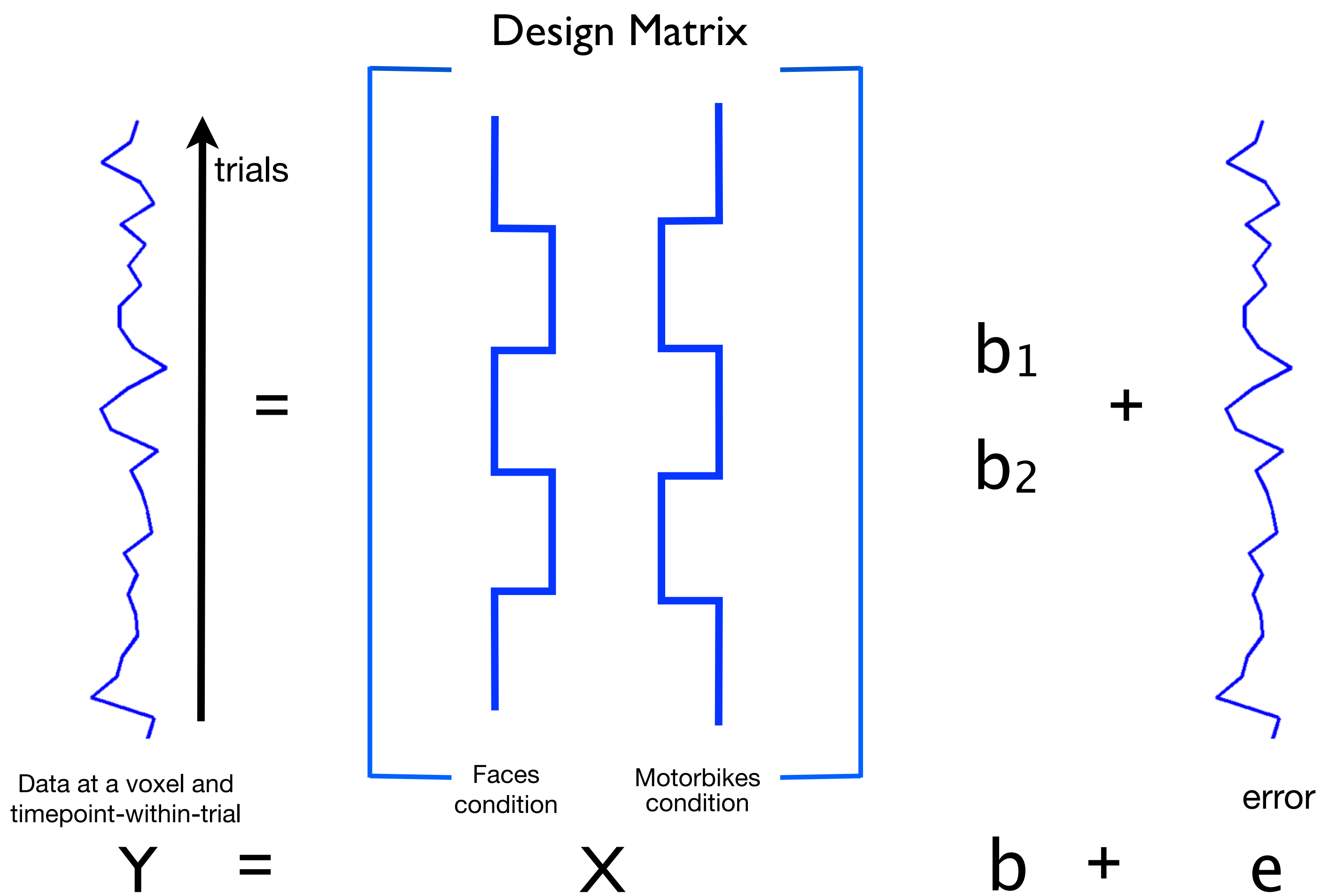
- questions of interest involve tests/comparisons at the group level

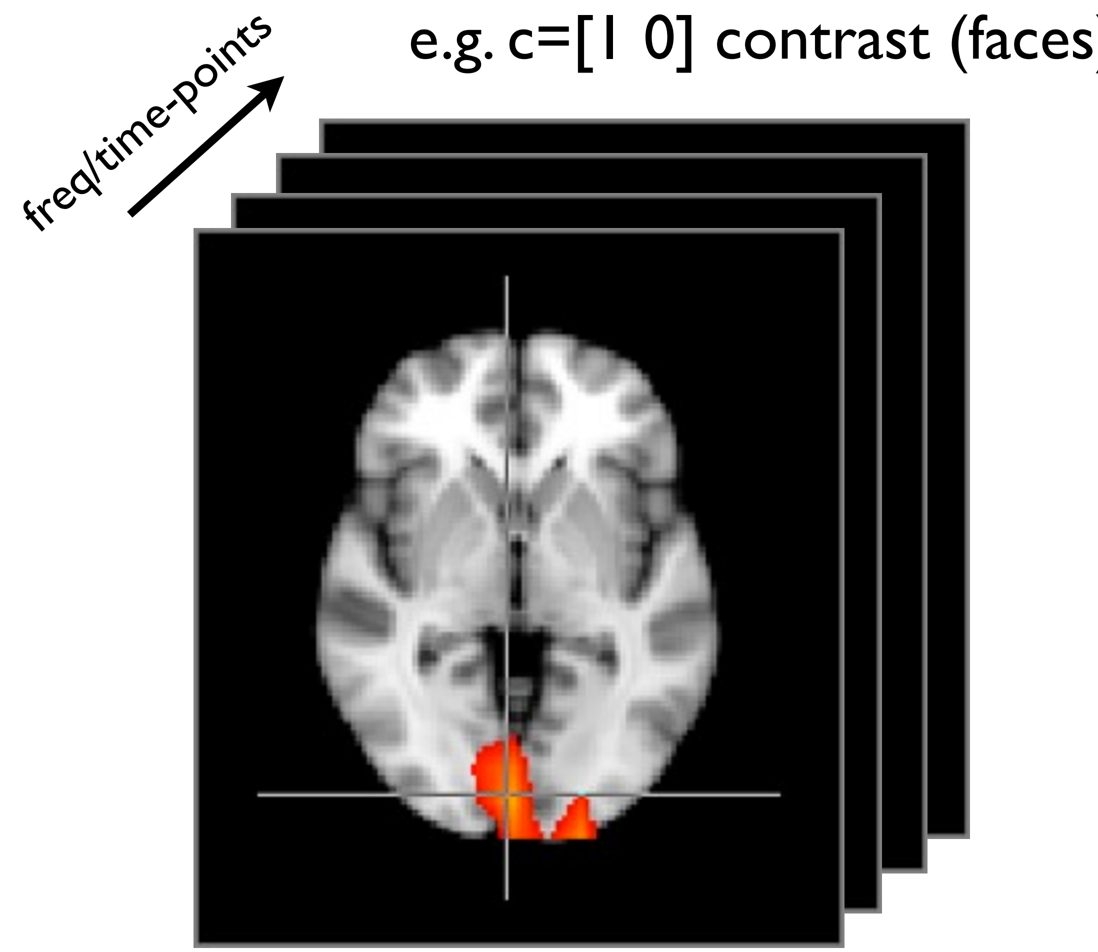
OAT Pipeline Stages

- 4 distinct pipeline stages:



First-Level (Trial-wise) GLM



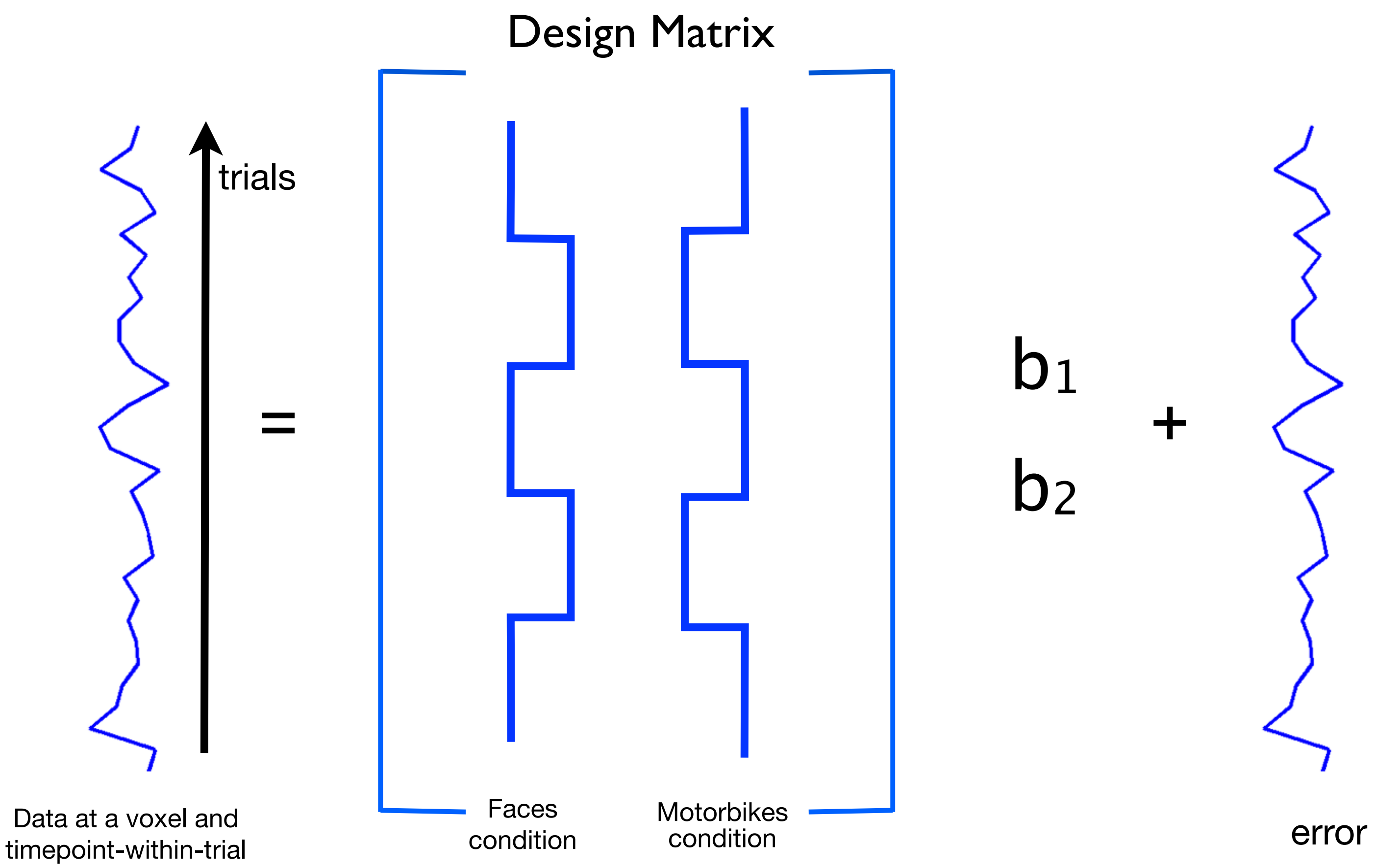


Repeat for all voxels/ freq/time-points, then compute contrasts: COPE=c*B

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

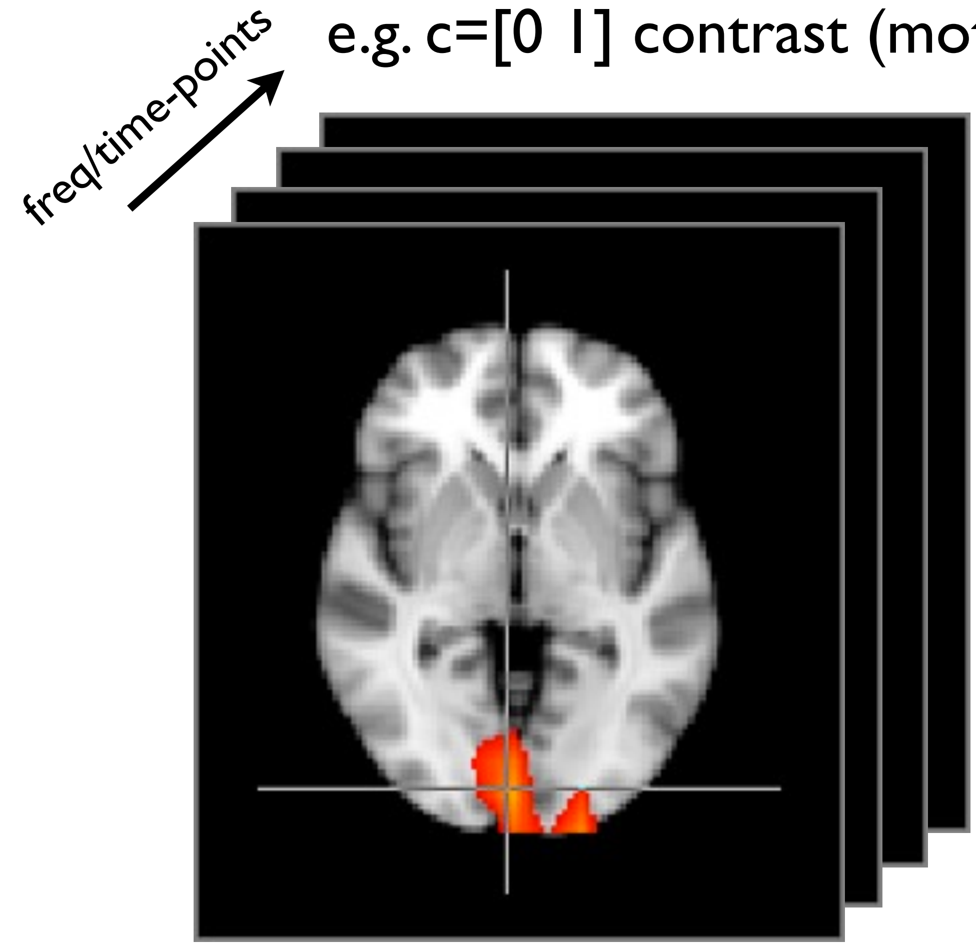
B is the Parameter Estimates (PEs) of b

First-Level (Trial-wise) GLM



$Y \quad = \quad X \quad b \quad + \quad e$

e.g. c=[0 1] contrast (motorbikes)

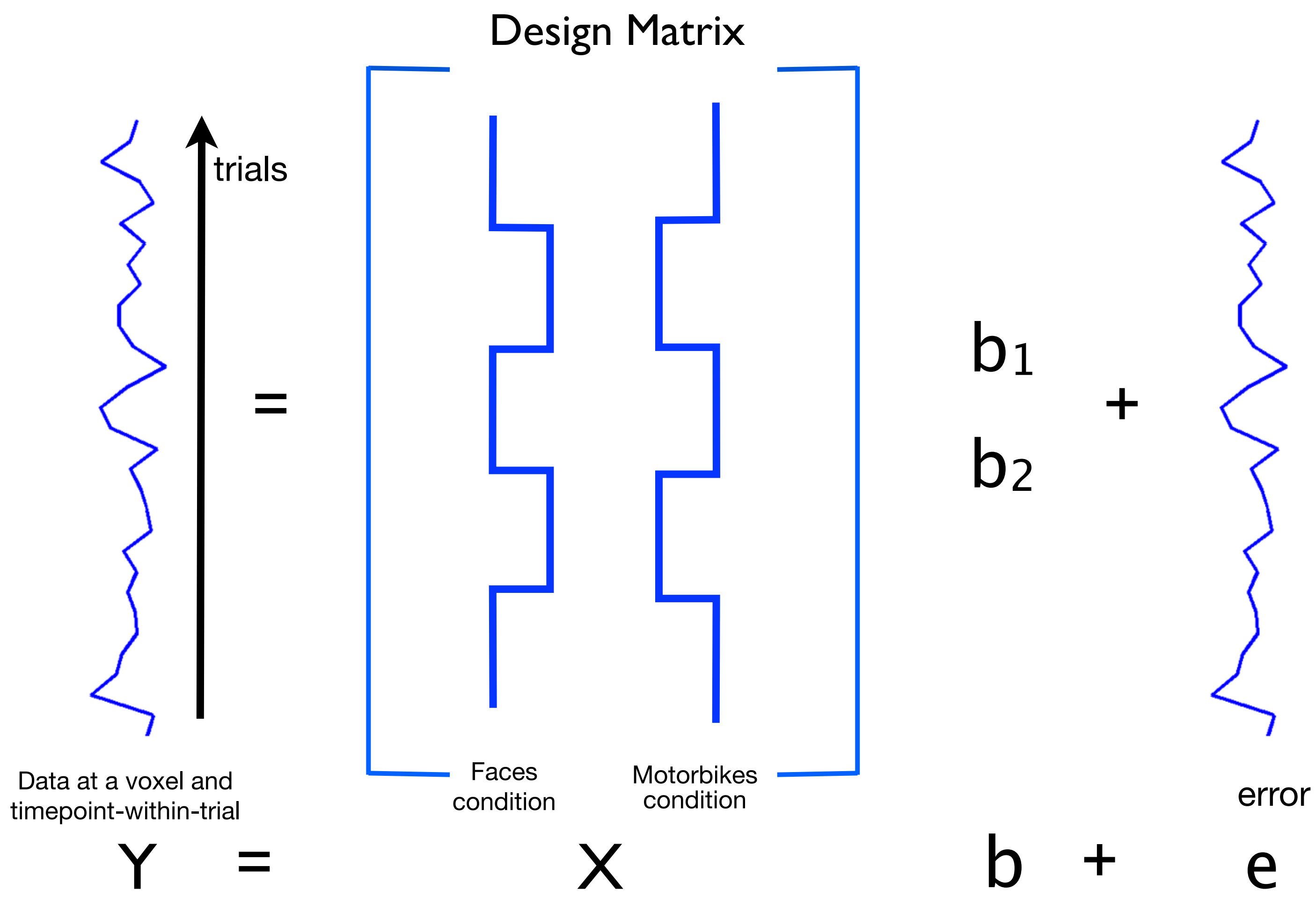


Repeat for all voxels/ freq/time-points, then compute contrasts: COPE=c*B

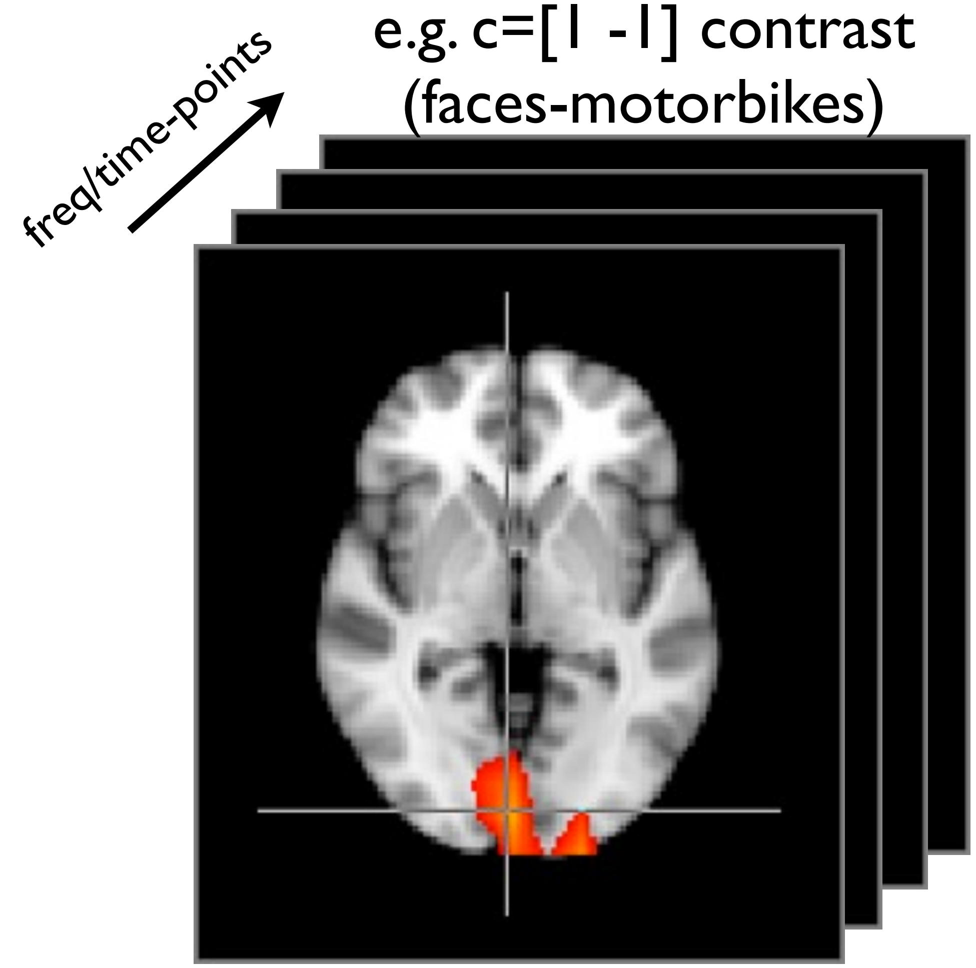
$$Y=Xb+e \rightarrow B=pinv(X)Y$$

B is the Parameter Estimates (PEs) of b

First-Level (Trial-wise) GLM



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



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

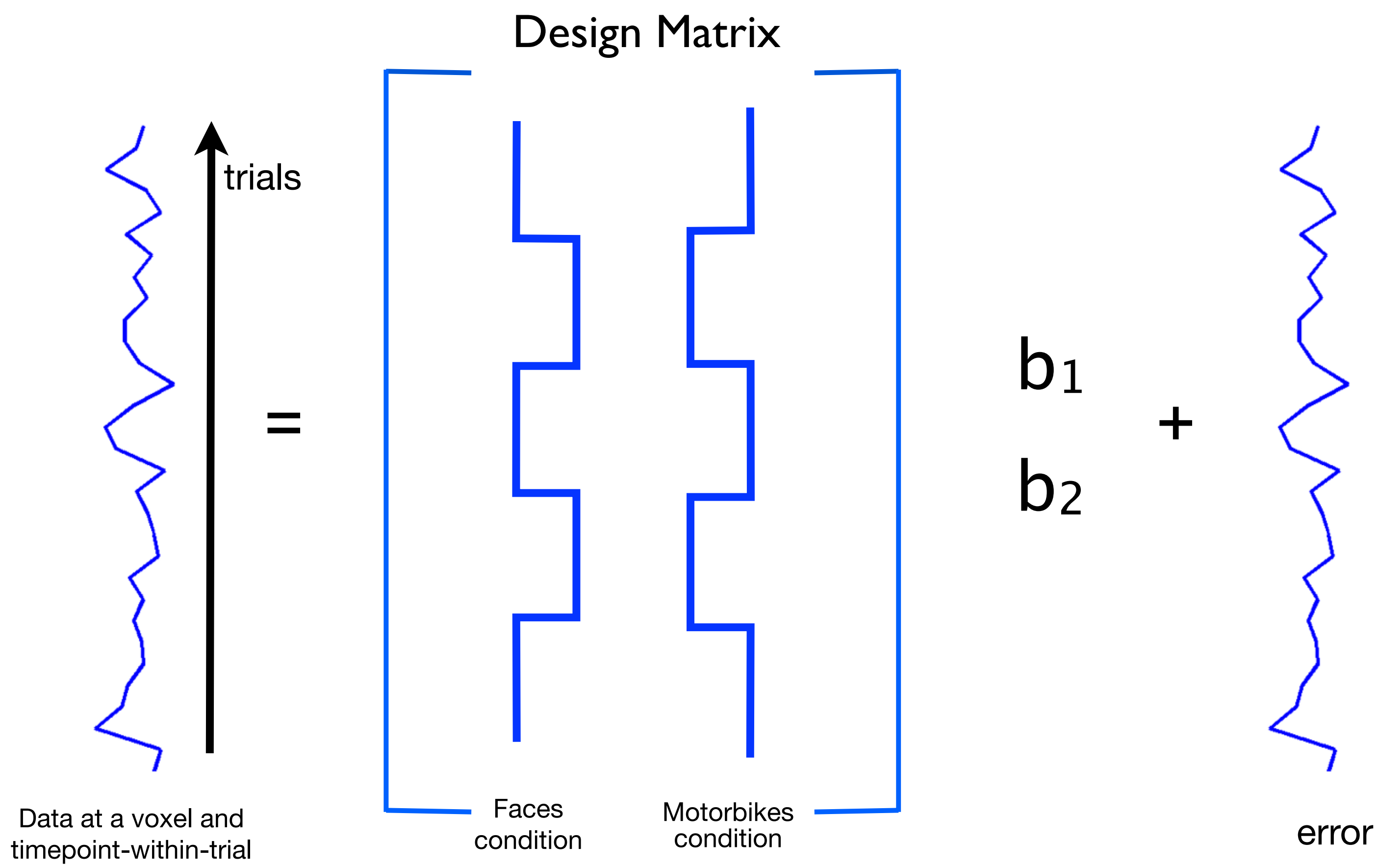
$$Y=Xb+e \rightarrow B=pinv(X)Y$$

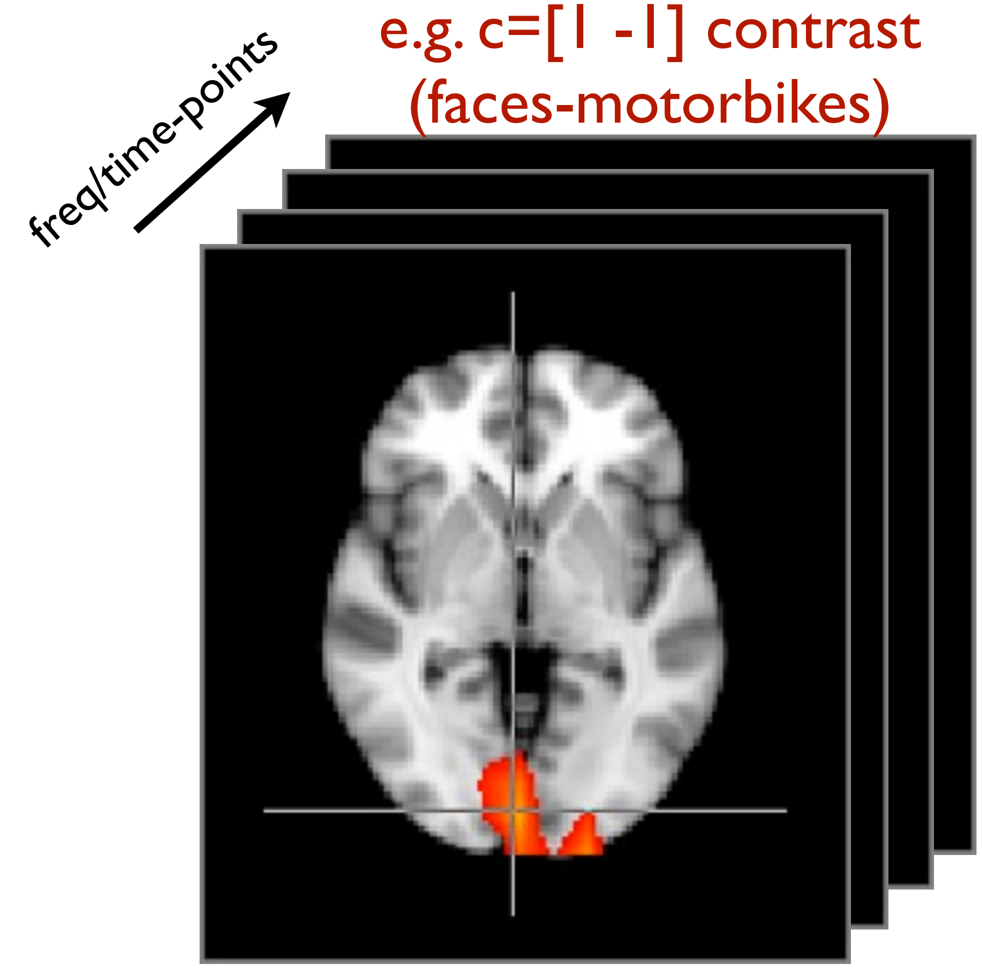
B is the Parameter Estimates (PEs) of b

First-Level (Trial-wise) GLM

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

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





Y = X b + e

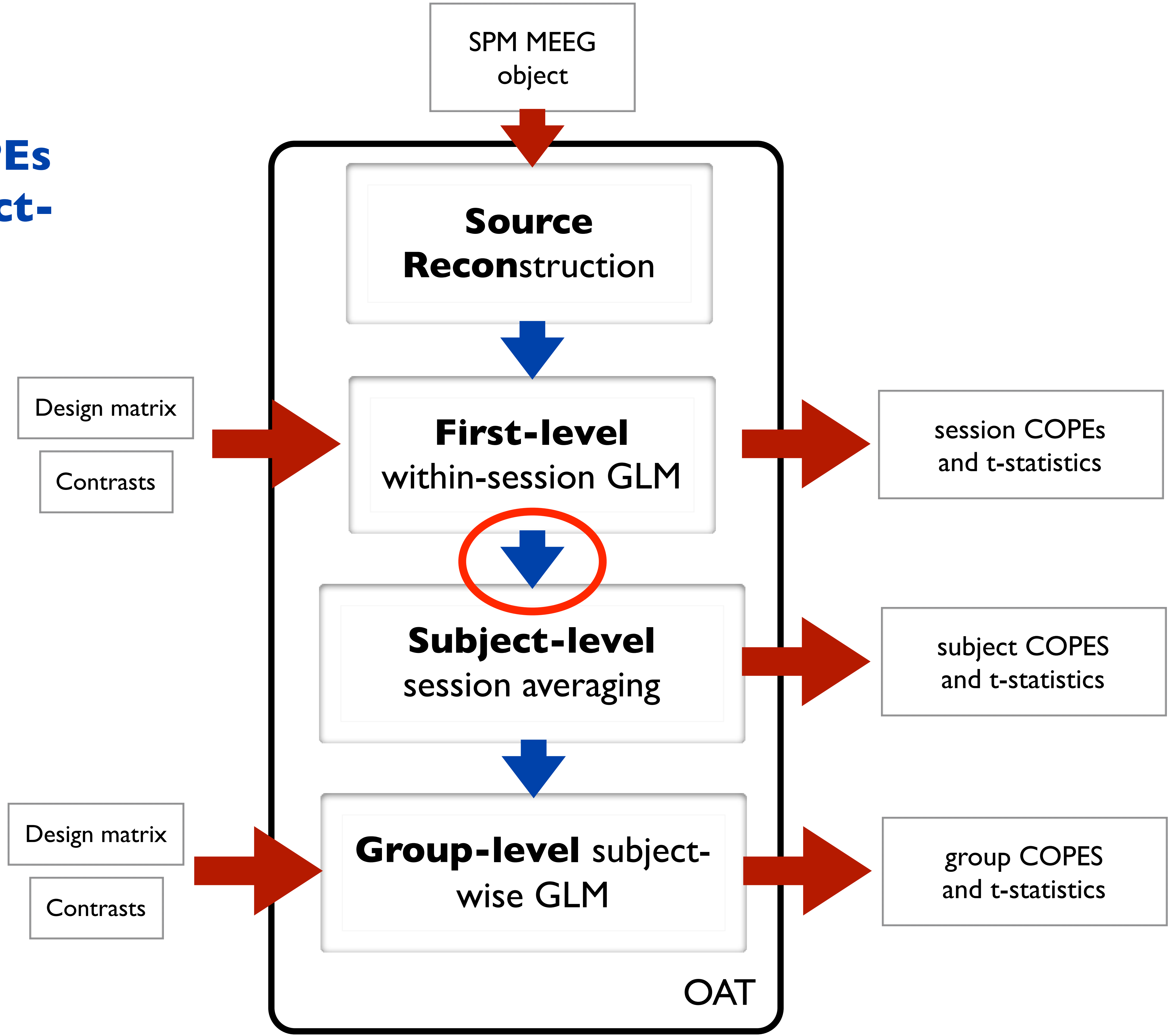
Repeat for all voxels/
freq/time-points, then
compute contrasts:
COPE=c*B

$$Y=Xb+e \longrightarrow B=\mathrm{pinv}(X)Y$$

B is the Parameter Estimates (PEs) of b

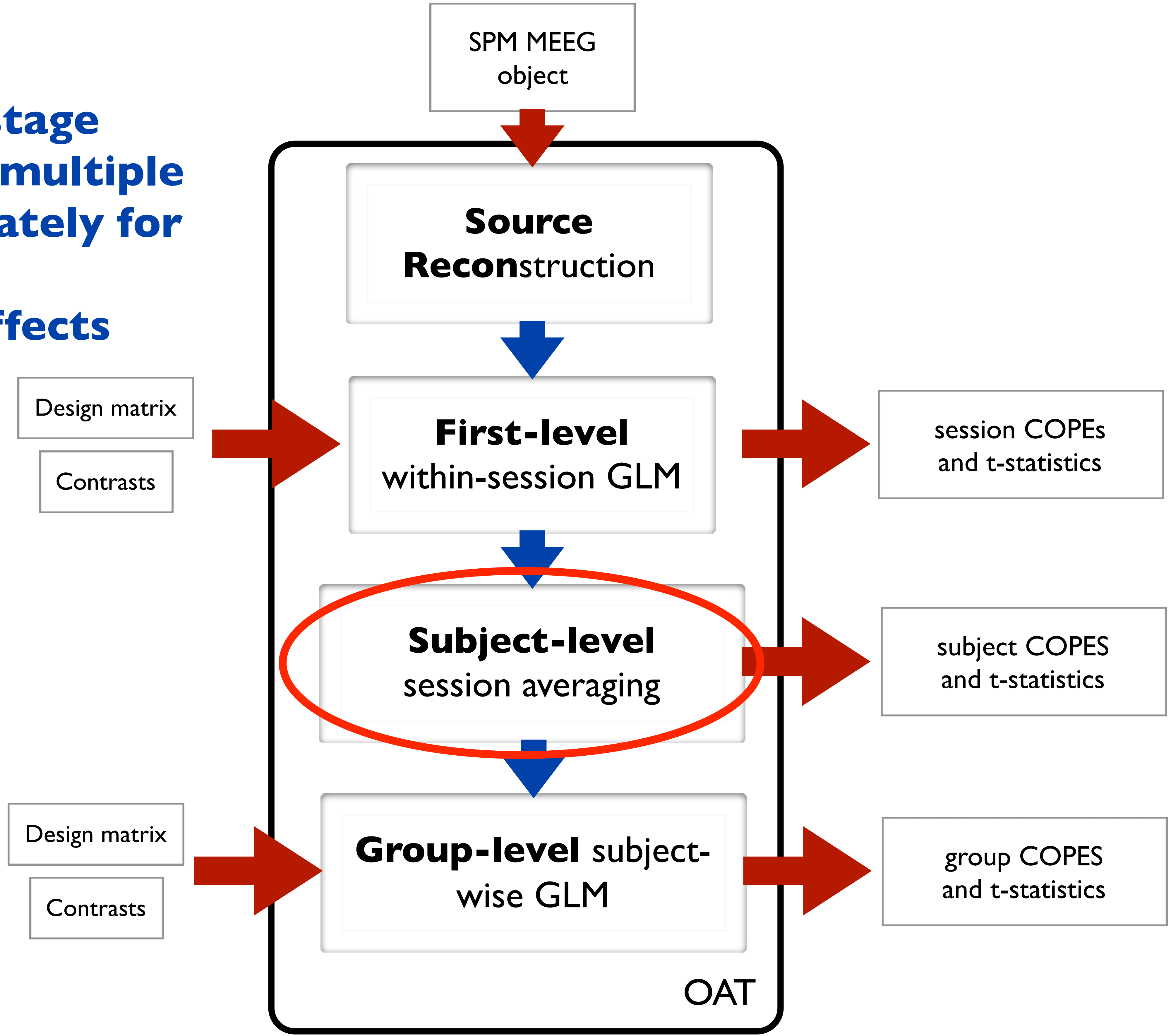
Multiple Session Analysis

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



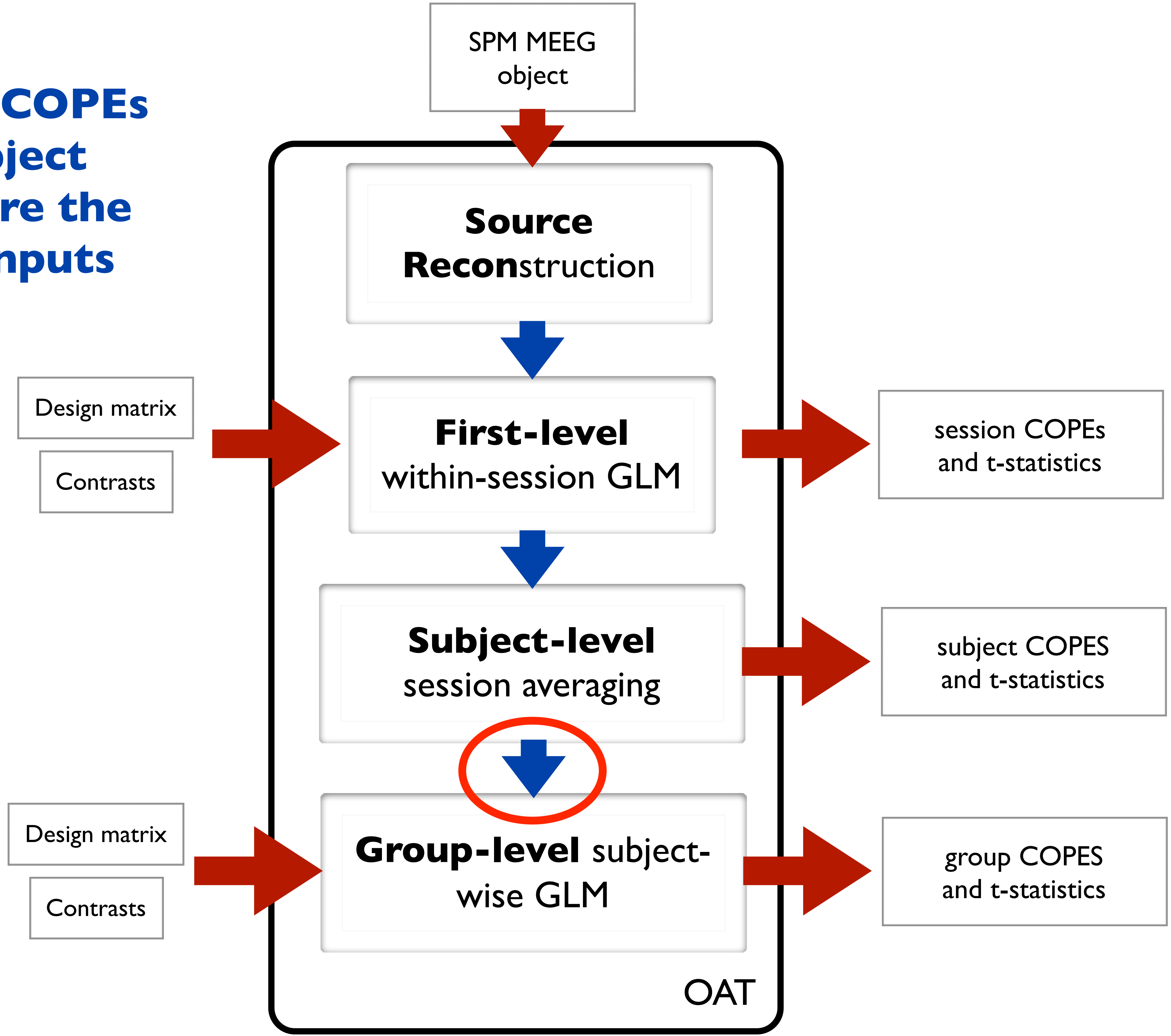
Multiple Session Analysis

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



Multiple Subject Analysis

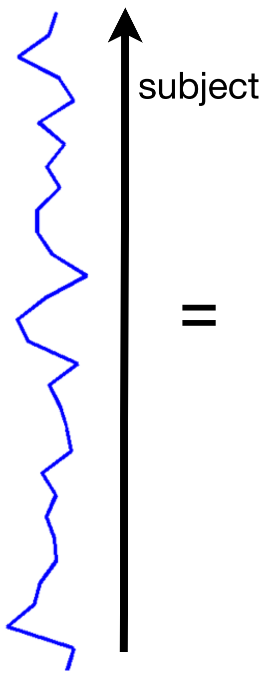
Subject-level COPEs (i.e. mean subject effect sizes) are the Group-level Inputs



Subject-wise GLM (Multiple Regression)

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

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



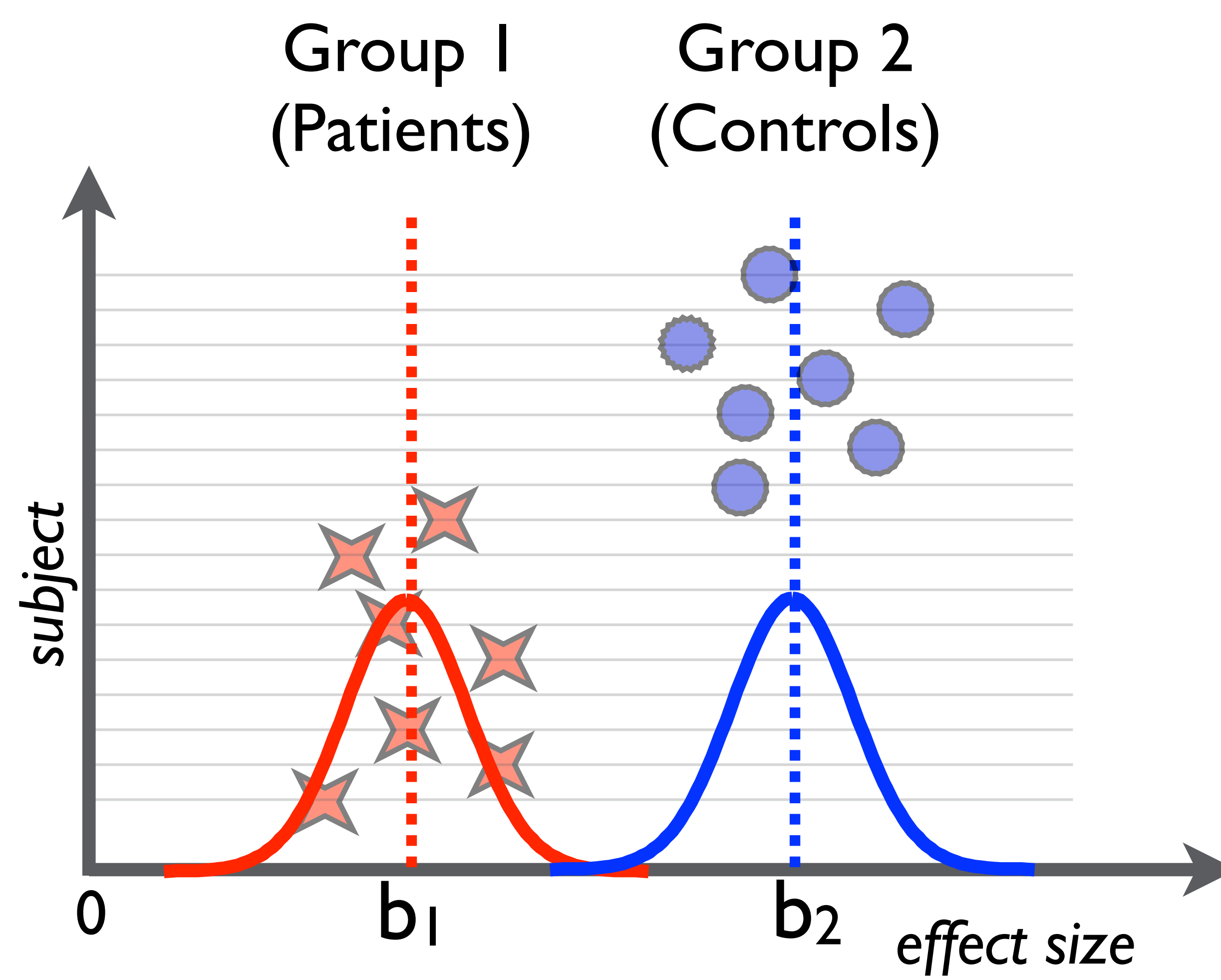
$$
= \begin{array}{c}
\text{Design Matrix} \\
\left[\begin{array}{cc}
1 & 0 \\
1 & 0 \\
1 & 0 \\
1 & 0 \\
1 & 0 \\
1 & 0 \\
1 & 0 \\
0 & 1 \\
0 & 1 \\
0 & 1 \\
0 & 1 \\
0 & 1 \\
0 & 1 \\
0 & 1
\end{array}\right] \\
\text{Patients} \quad \text{Controls}
\end{array}
\begin{array}{c} b_1 \\ b_2 \end{array} +
$$

error

$$Y \quad = \quad X \quad b \quad + \quad e$$

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

$$Y=Xb+e$$



Subject-wise GLM (Multiple Regression)

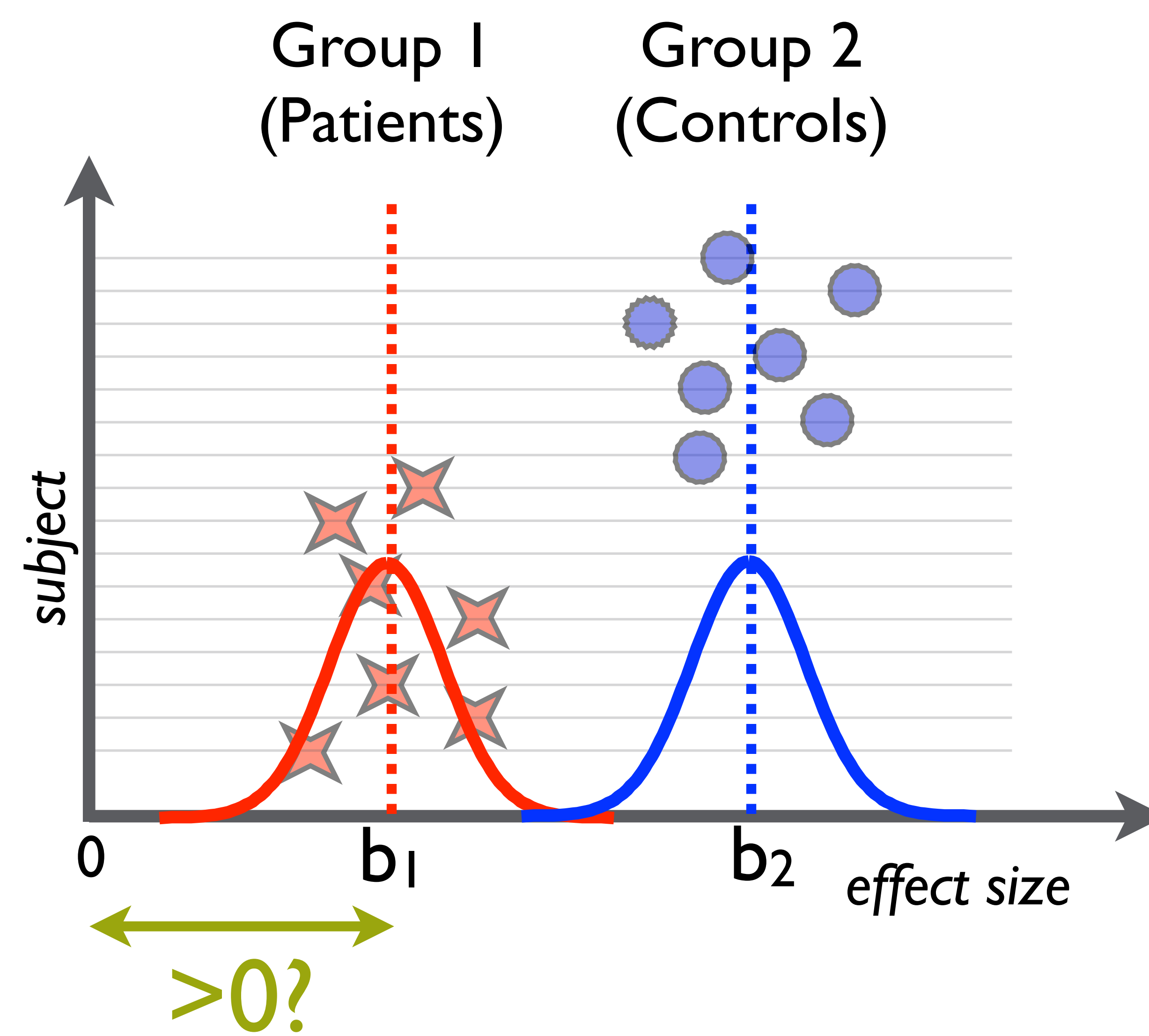
- We have two groups (e.g. 7 patients, 7 controls)

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

Design Matrix

$$
\begin{array}{c}
\text{Group 1 (Patients)} \\ \\ \\ \\ \\ \\ \\
\text{Group 2 (Controls)} \\ \\ \\ \\ \\ \\ \\
\end{array}
\begin{bmatrix}
1 & 0 \\
1 & 0 \\
1 & 0 \\
1 & 0 \\
1 & 0 \\
1 & 0 \\
1 & 0 \\
0 & 1 \\
0 & 1 \\
0 & 1 \\
0 & 1 \\
0 & 1 \\
0 & 1 \\
0 & 1
\end{bmatrix}
$$

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)

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

Design Matrix

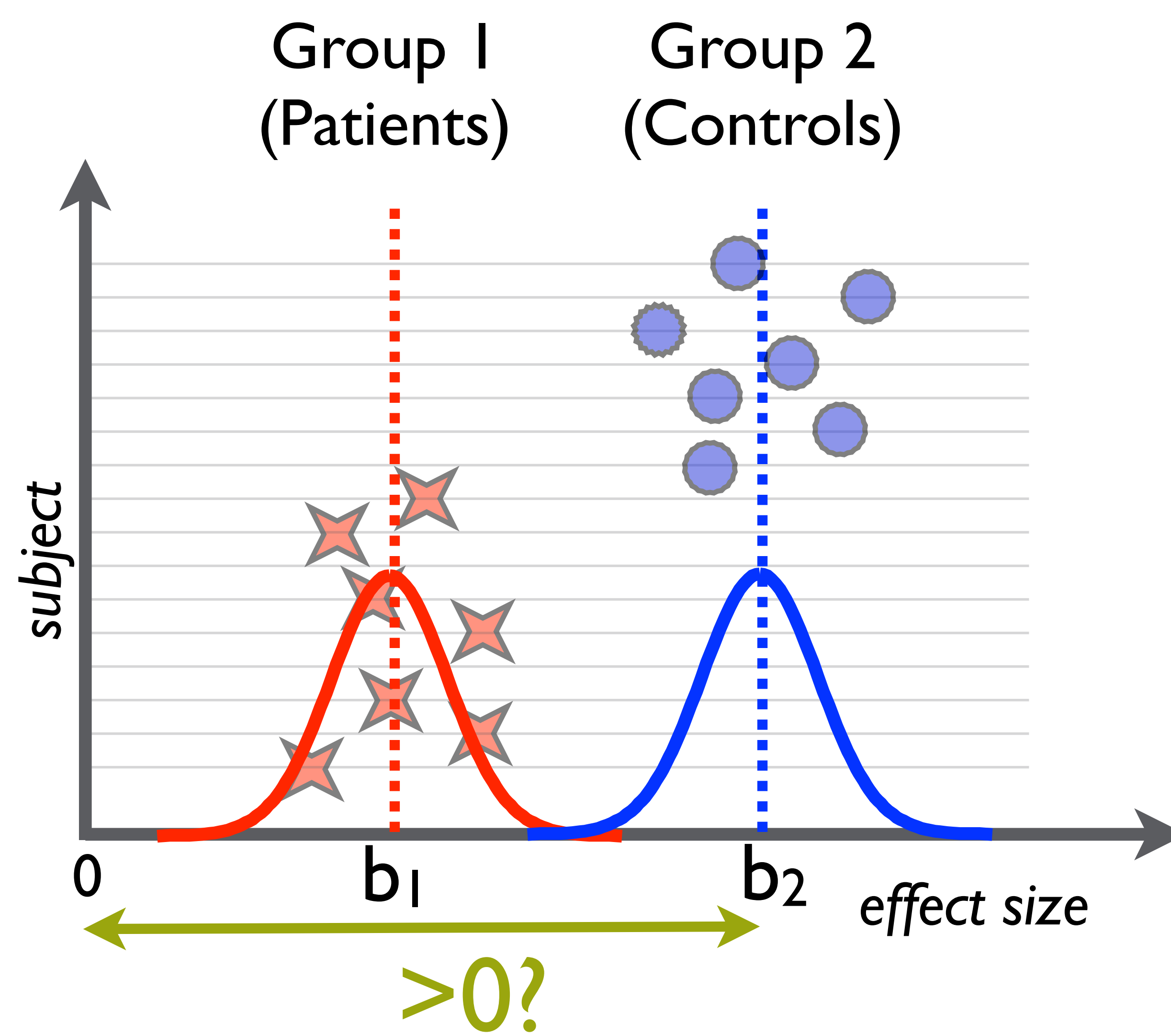
$$\begin{array}{l} \text{Group 1 (Patients)} \\ \\ \text{Group 2 (Controls)} \end{array} \quad \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$$

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)

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

Design Matrix

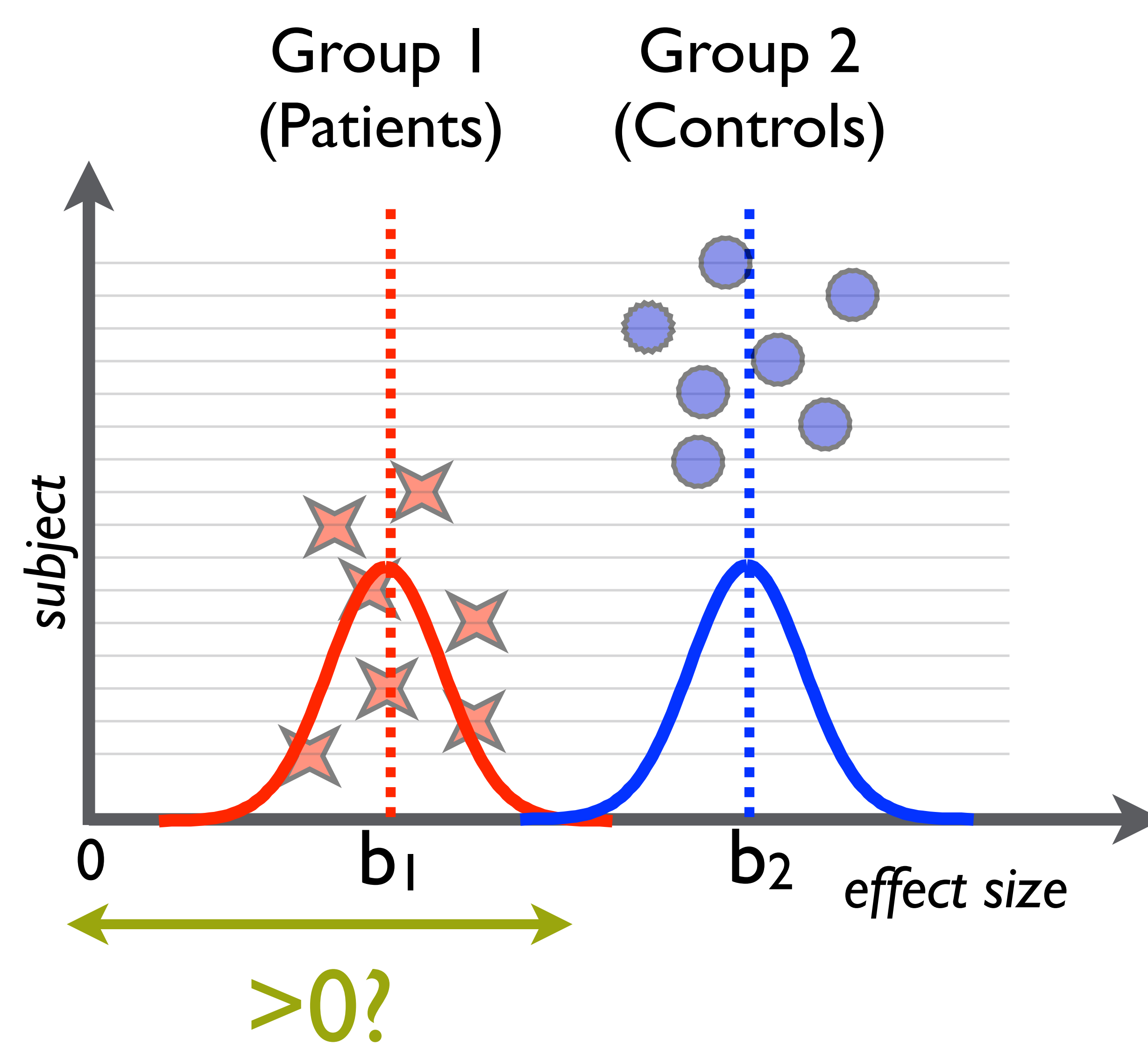
$$
\begin{array}{l}
\text{Group 1 (Patients)} \\
\\
\text{Group 2 (Controls)}
\end{array}
\begin{bmatrix}
1 & 0 \\
1 & 0 \\
1 & 0 \\
1 & 0 \\
1 & 0 \\
1 & 0 \\
1 & 0 \\
0 & 1 \\
0 & 1 \\
0 & 1 \\
0 & 1 \\
0 & 1 \\
0 & 1 \\
0 & 1
\end{bmatrix}
$$

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)

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

Design Matrix

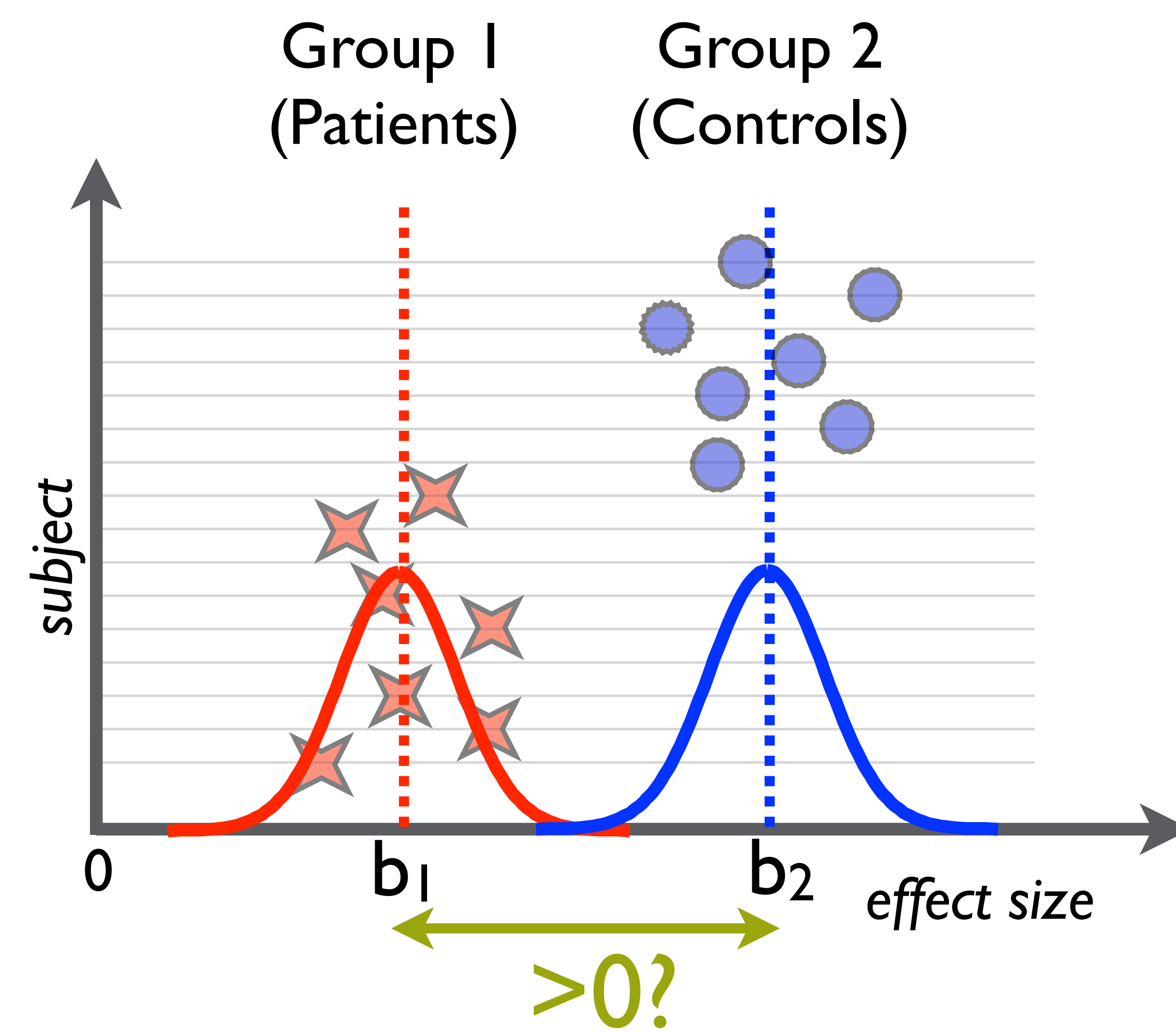
$$
\begin{array}{l}
\text{Group 1 (Patients)} \\
\\
\text{Group 2 (Controls)}
\end{array}
\quad
\begin{bmatrix}
1 & 0 \\
1 & 0 \\
1 & 0 \\
1 & 0 \\
1 & 0 \\
1 & 0 \\
1 & 0 \\
0 & 1 \\
0 & 1 \\
0 & 1 \\
0 & 1 \\
0 & 1 \\
0 & 1 \\
0 & 1
\end{bmatrix}
$$

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)

Design Matrix

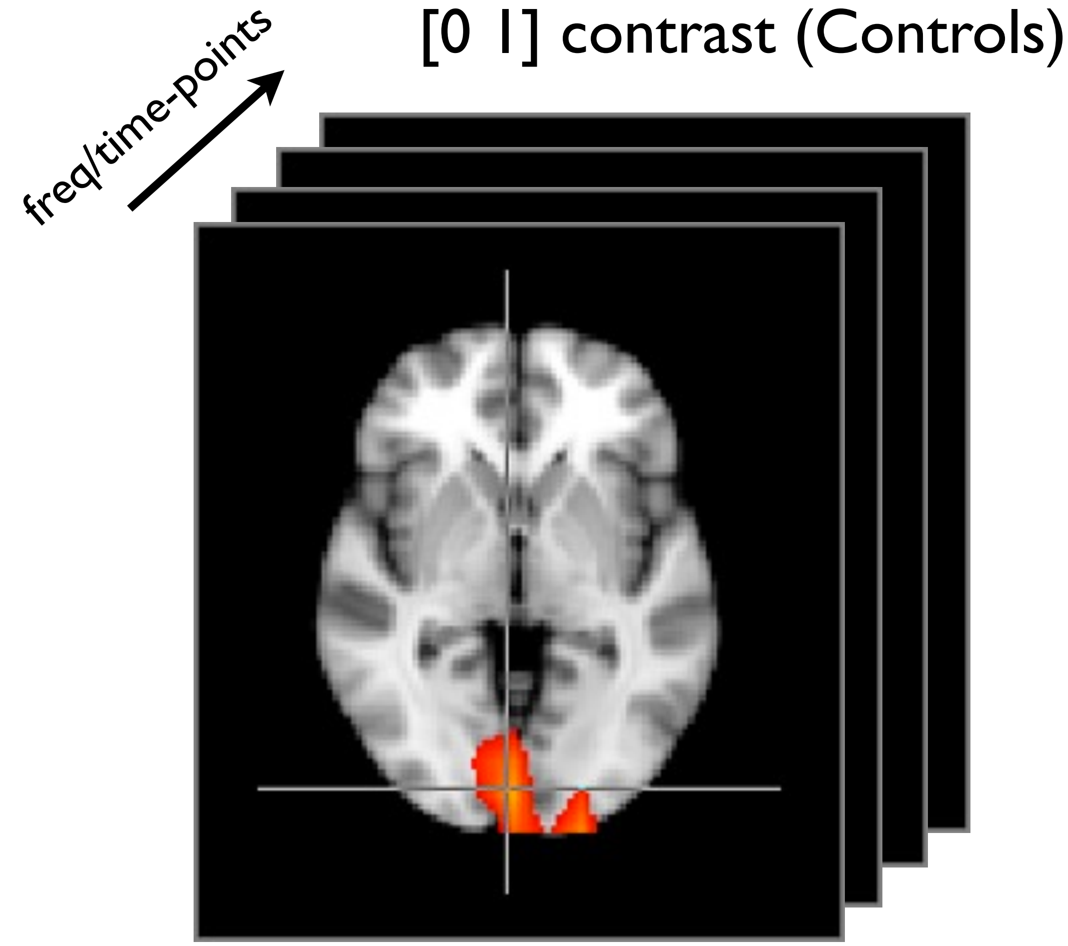
$$
Y = \begin{array}{c} \\ \left[\begin{array}{cc} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{array}\right] \\ \text{Patients} \quad \text{Controls} \end{array} \begin{array}{c} b_1 \\ b_2 \end{array} + e
$$

subjects

error

$Y \;=\; X \;\; b \; + \; e$

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



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

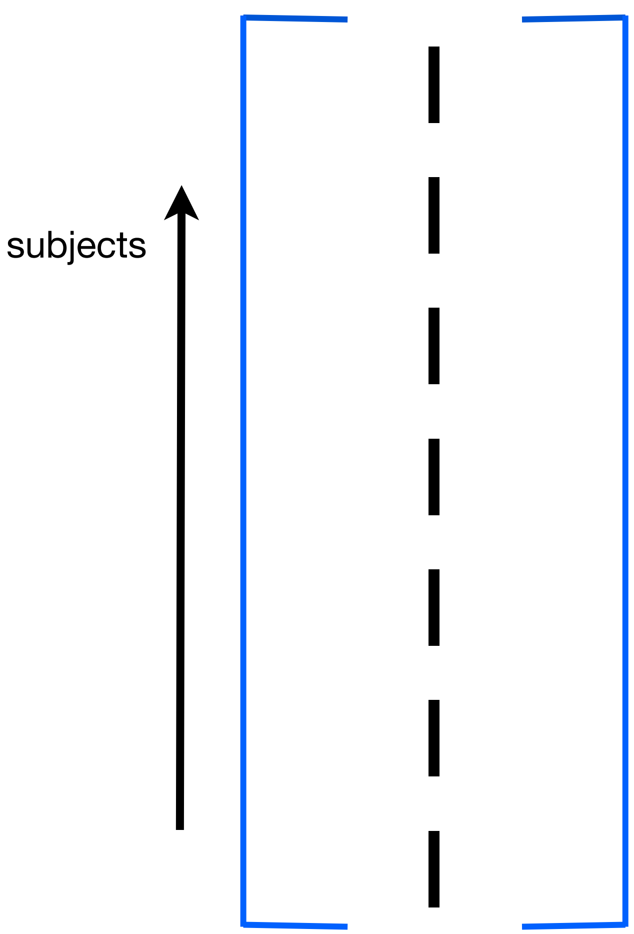
$Y=Xb+e$ → Fit GLM

- This is also all repeated for ALL first-level contrasts (COPES)

Single Group Average

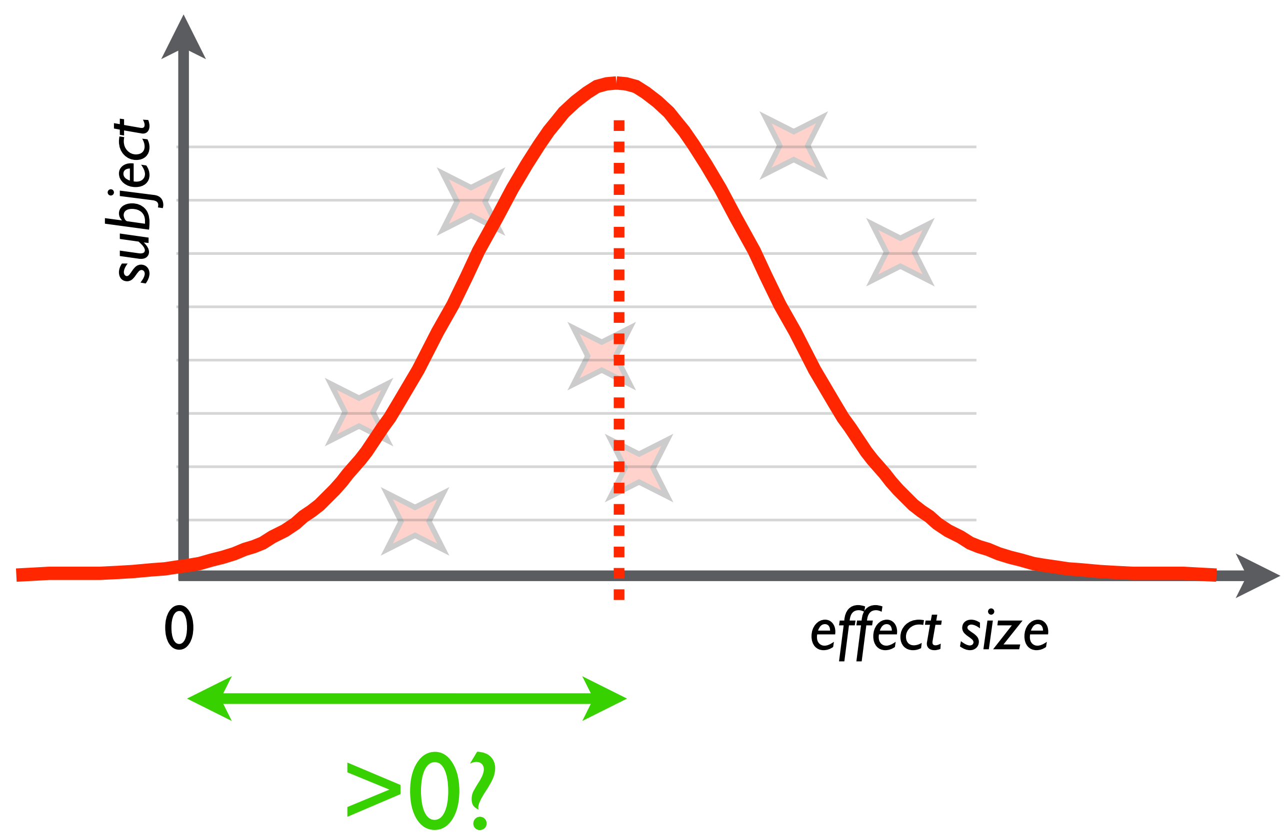
- We have 7 subjects - all in one group - and want the mean group average:

Design Matrix



Contrast 1: [1]

Does the group activate on average?



Behavioural Variables

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

Design Matrix

subjects

$$\begin{bmatrix} 1 & 3 \\ 1 & 1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 0 \\ 1 & -1 \\ 1 & -4 \end{bmatrix}$$

- 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)

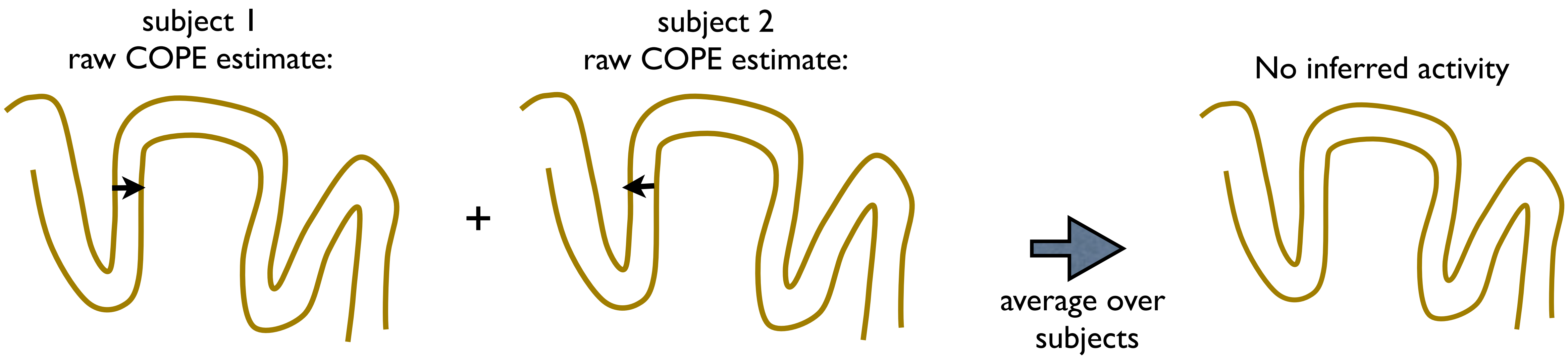
Other things you can do with the Group GLM

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

- Time-frequency power analysis. Note: if whole brain then do these a band at a time (for speed)
- Sensor space

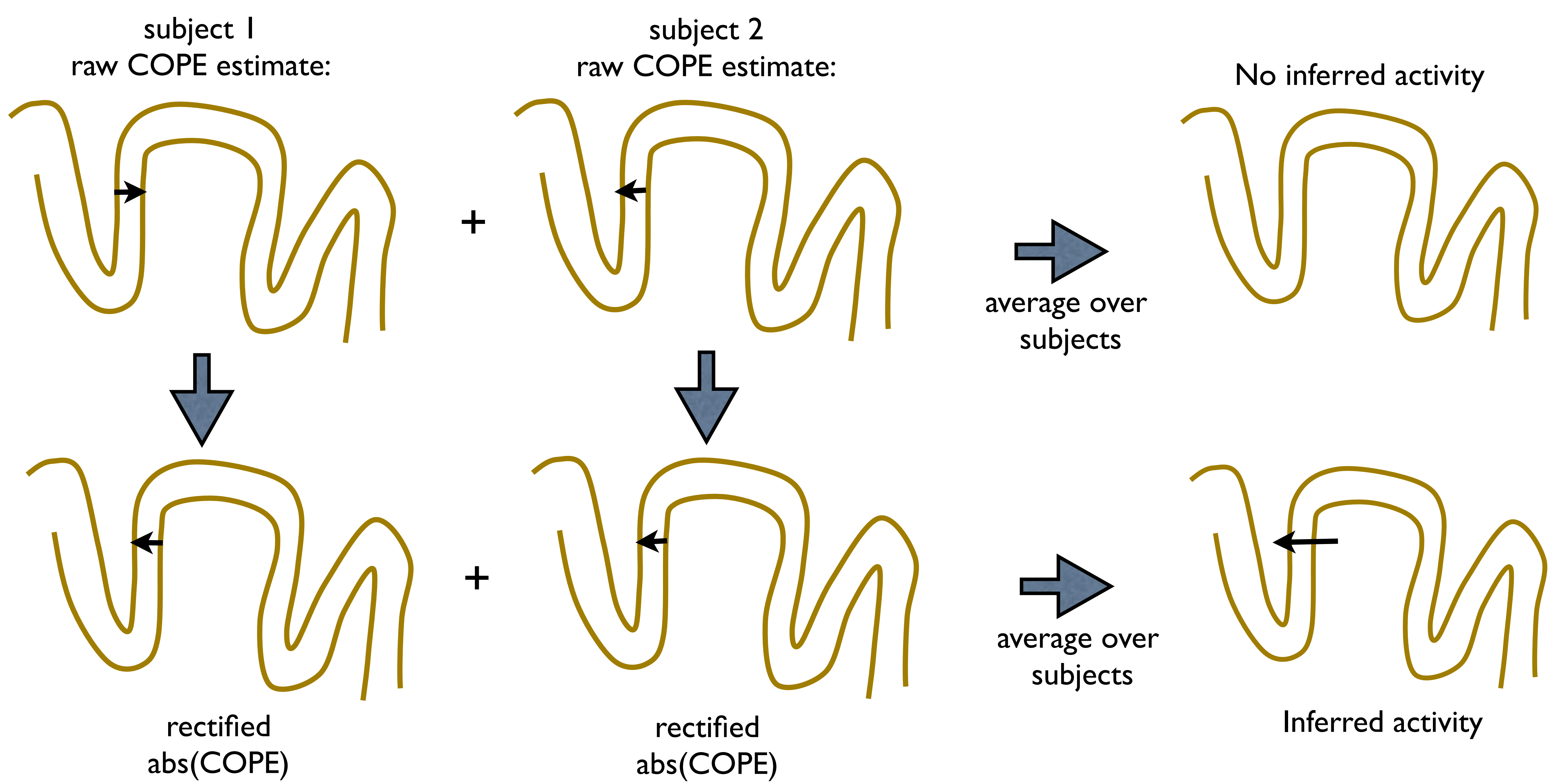
Note on ERF “rectification”

For ERF analysis, in OSL we rectify (take the absolute value of) the first-level COPEs and baseline correct



Note on ERF "rectification"

For ERF analysis, in OSL we rectify (take the absolute value of) the first-level COPEs and baseline correct



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

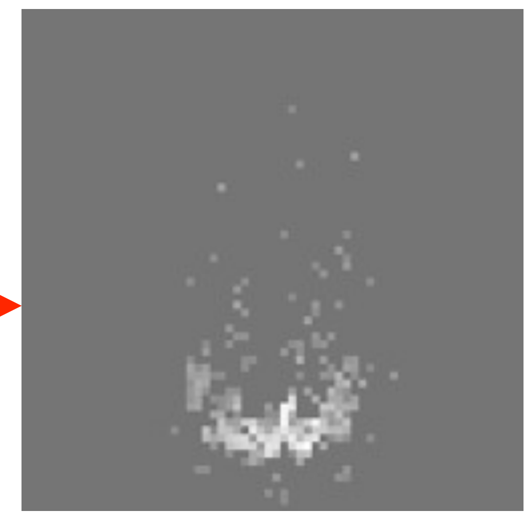
  $0.05 / 20{,}000 = 0.0000025$

Thresholding: Clustering

What about testing significance on clusters?

stat image

Threshold at
(arbitrary!) level

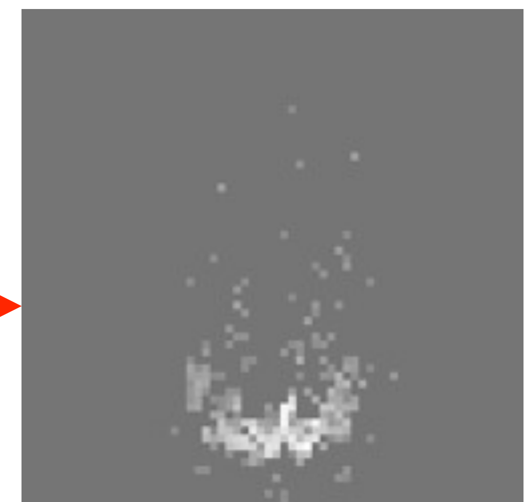


Thresholding: Clustering

What about testing significance on clusters?

stat image

Threshold at
(arbitrary!) level



- Form clusters from surviving voxels.

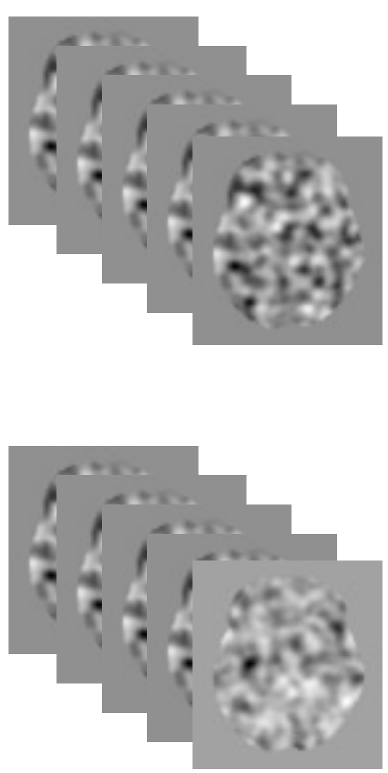
- **BUT**, what is the probability of getting a cluster, given its spatial extent (no. of voxels in cluster) and z threshold, under the null hypothesis?
- Can use Random Field Theory, but involves dodgey assumptions, instead ...

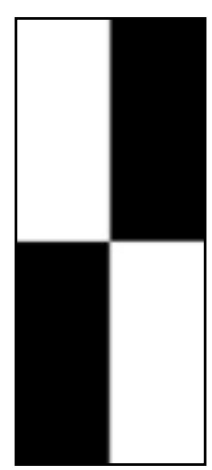
Permutation Testing

We can record the cluster sizes from our dataset

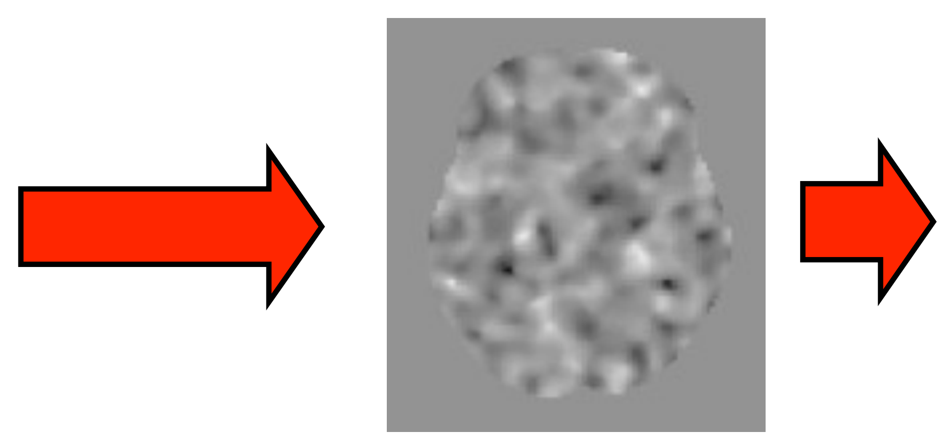
Group 1

Group 2





Our actual 2nd level model

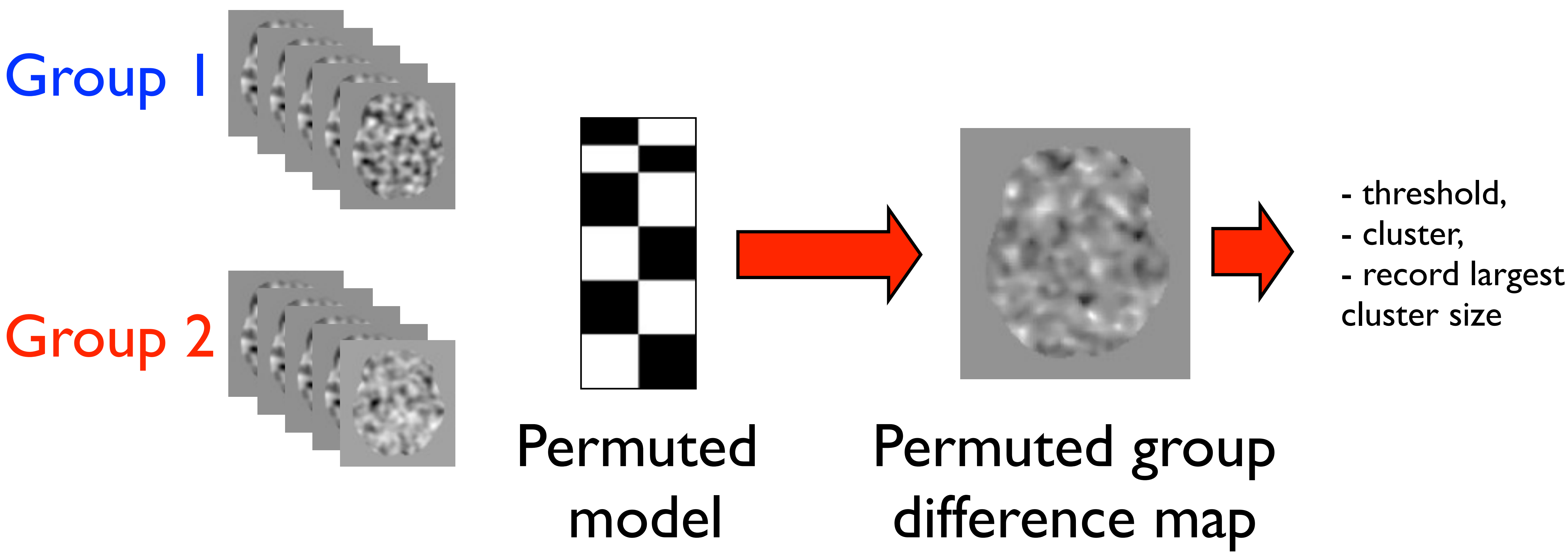


Our actual group difference map

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

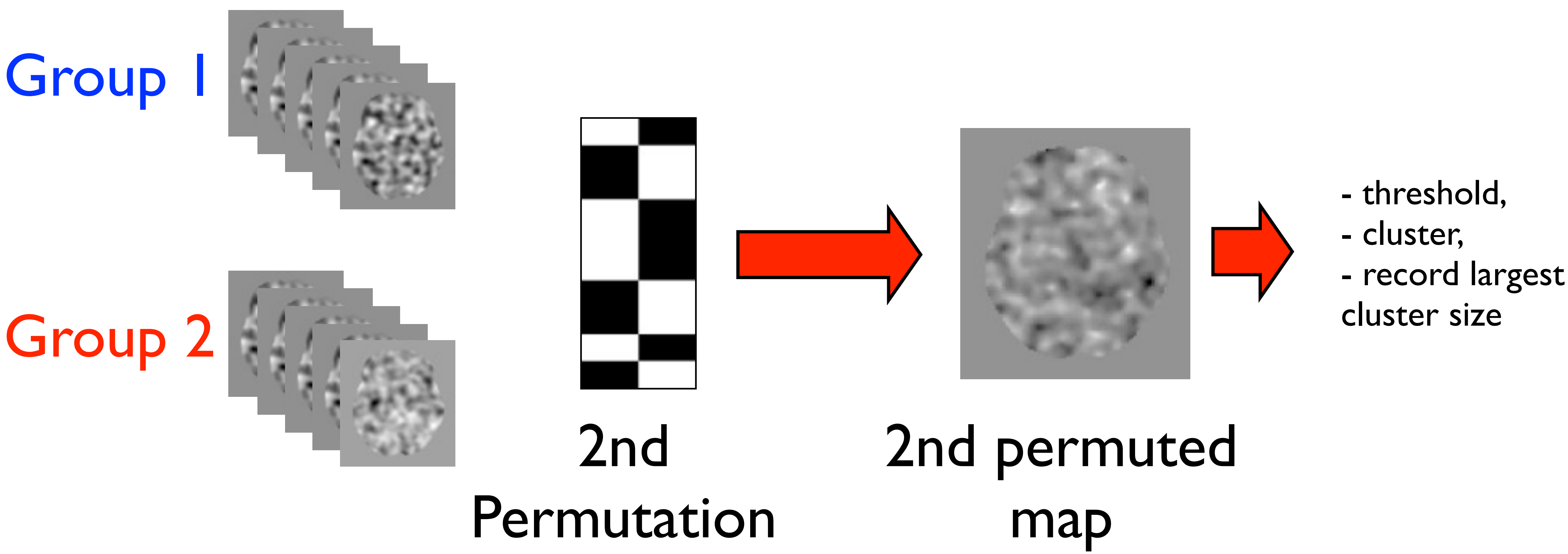
Permutation Testing

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



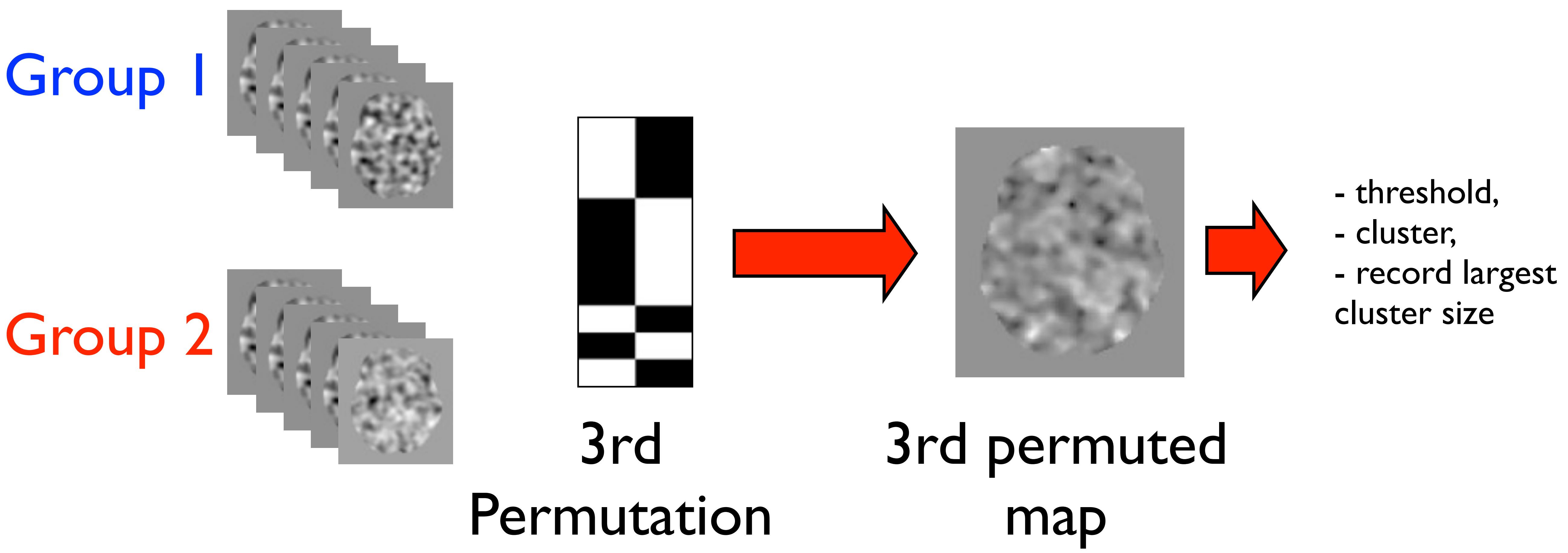
Permutation Testing

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



Permutation Testing

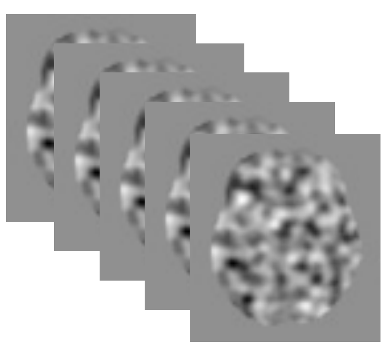
We can then permute the design matrix group labellings to get the null distribution of the maximum cluster size



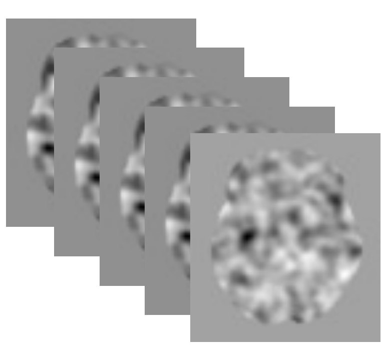
Permutation Testing

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

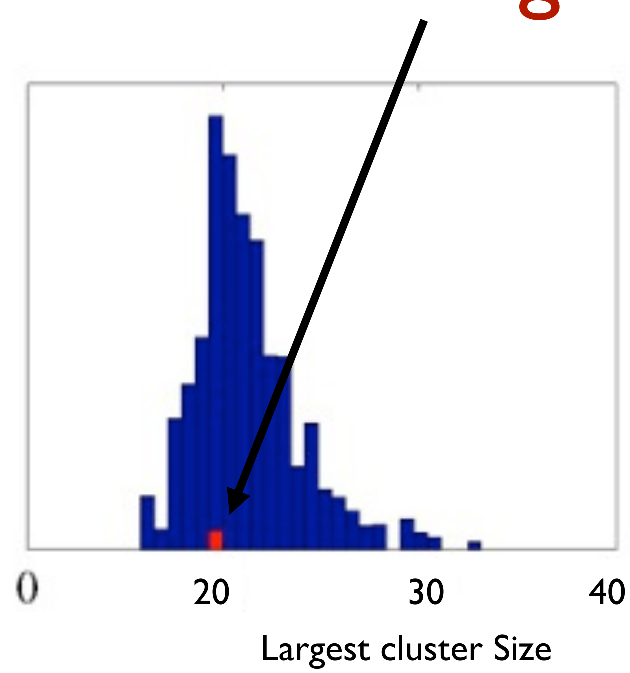
Group 1



Group 2



Original group labellings



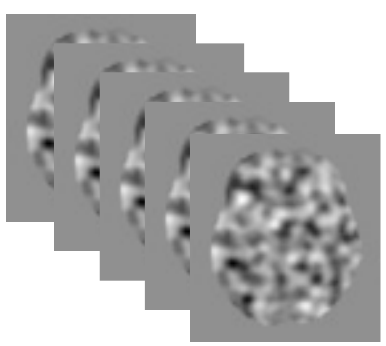
- 3925 permutations yielded larger clusters than original group labellings.
- We cannot reject the null-hypothesis in this case

5000 permutations

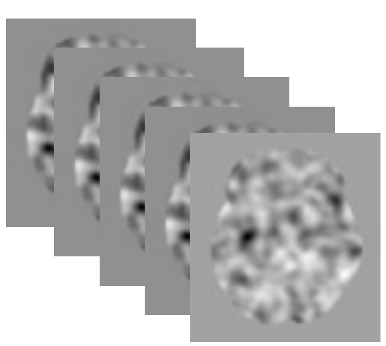
Permutation Testing

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

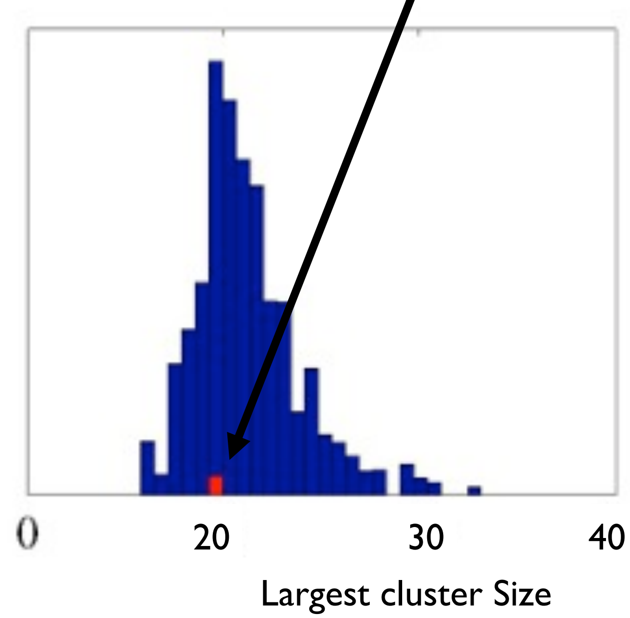
Group 1



Group 2



Original group labellings

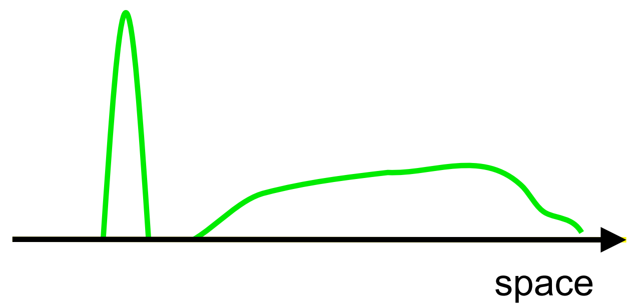


5000 permutations

- OSL has the facility to test clusters in:
  - 2-D (time-frequency)
  - 3-D (space)
  - 4-D (space and e.g. time)
- Note: 4D is very slow!!

How do we choose the (arbitrary!) cluster forming threshold?

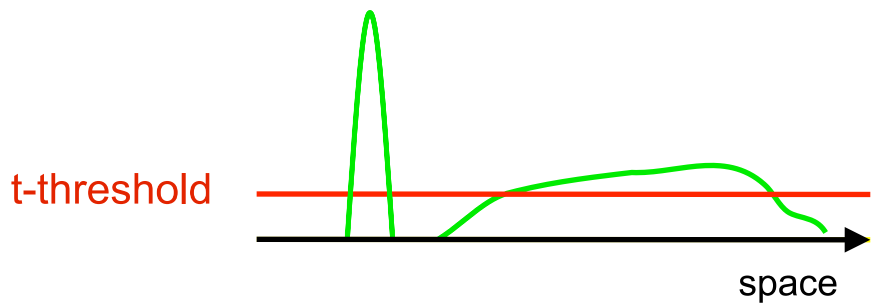
This is arbitrary and a trade-off



How do we choose the (arbitrary!) cluster forming threshold?

This is arbitrary and a trade-off

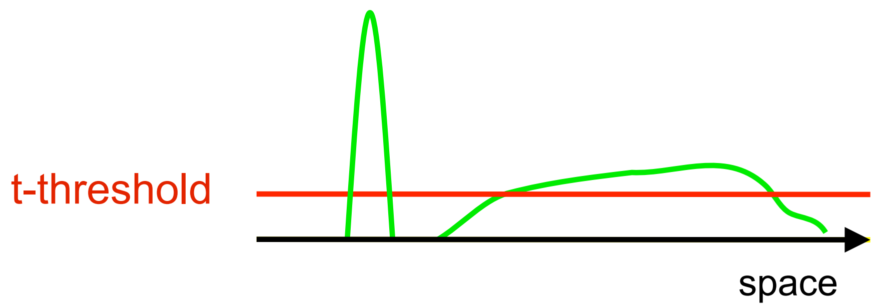
1. **Low threshold** - can detect clusters with large spatial extent and low t-statistic



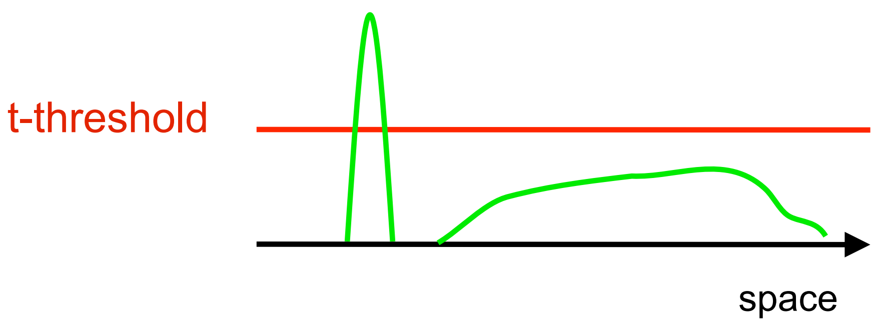
How do we choose the (arbitrary!) cluster forming threshold?

This is arbitrary and a trade-off

1. **Low threshold** - can detect clusters with large spatial extent and low t-statistic



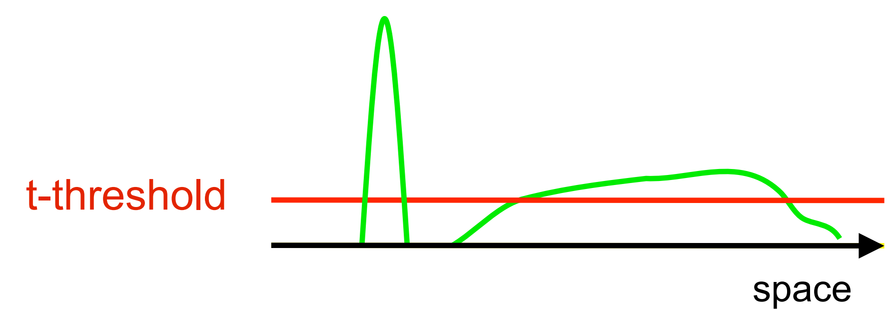
2. **High threshold** - gives more power to clusters with small spatial extent and high t-statistic



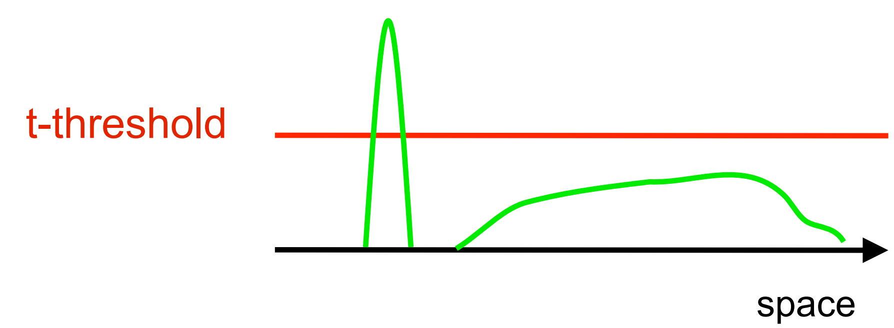
How do we choose the (arbitrary!) cluster forming threshold?

This is arbitrary and a trade-off

1. **Low threshold** - can detect clusters with large spatial extent and low t-statistic



2. **High threshold** - gives more power to clusters with small spatial extent and high t-statistic



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