Resting state fMRI and ICA

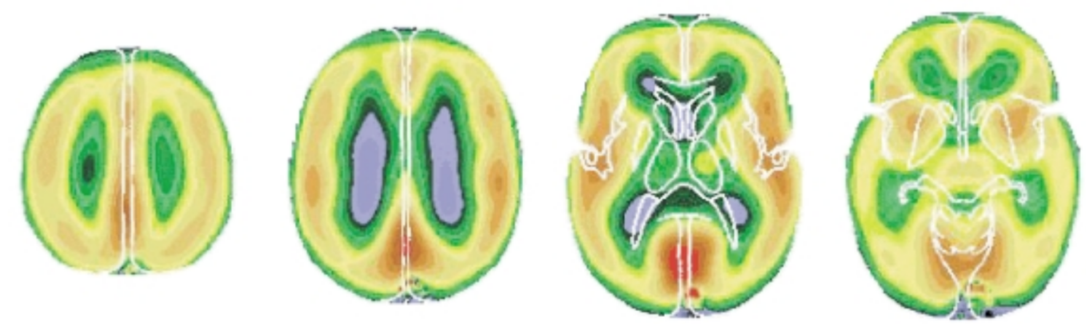
- Introduction to resting state
- Independent Component Analysis
- Single-subject ICA
- Multi-subject ICA
- Dual regression

Energy consumption in the brain

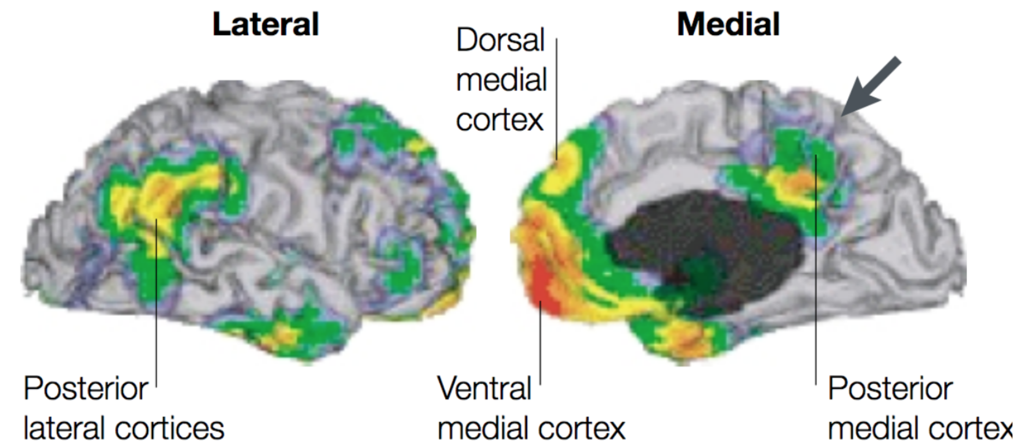
- Brain < 2% body weight but consumes ~20% of total energy
- estimated 60-80% of this energy used to support communication between cells
- task-evoked activity accounts for ~1%

Raichle et al (2001), _Gusnard et al (2001)_

Oxygen consumption

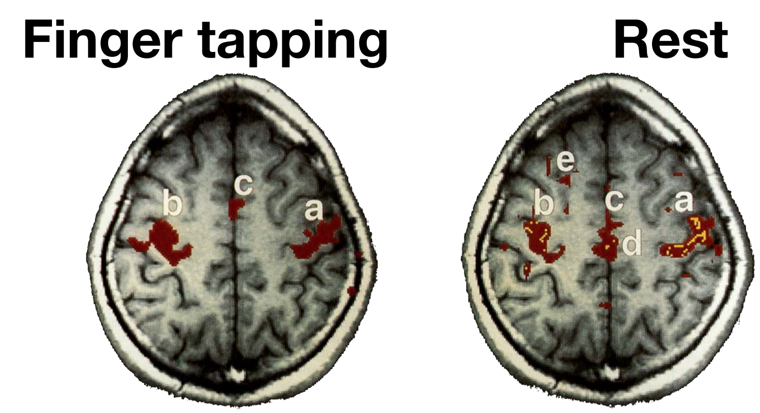


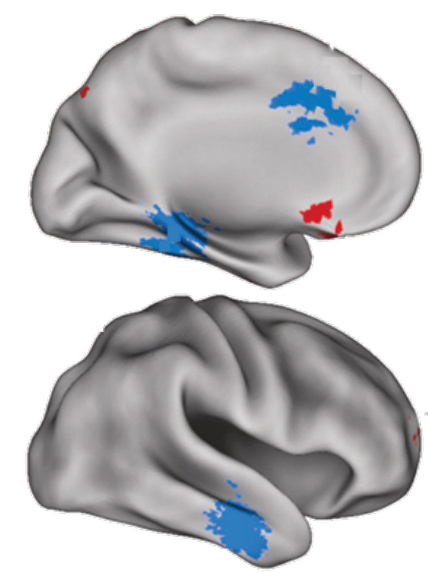
Decreased activity during tasks (PET)



Why study the brain at rest?

- Localisation versus connectivity
- Understand the inherent functional organisation of the brain
- Clinical/ cognitive biomarker
- Pragmatic benefits: can be done in any population, with relatively little setup and expertise required





Biswal et al (1995), *Sheline et al (2010)*

Principles of resting state analysis

- Many different methods available for analysis

- All have one assumption in common:

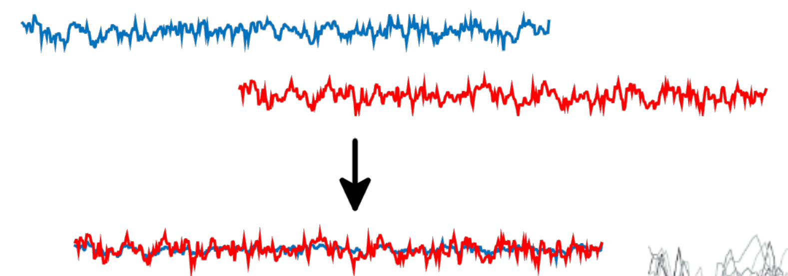
- i.e. definition of functional connectivity is based on a statistical dependency between timeseries

- Differences between methods lie in the way these similarities are estimated and/or represented

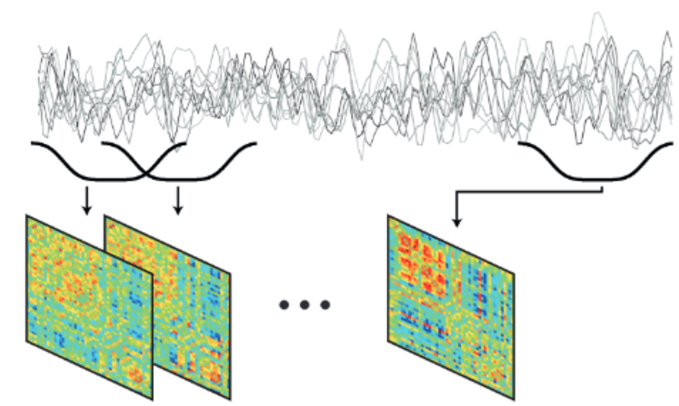
If two brain regions show similarities in their BOLD timeseries, they are functionally connected

Types of connectivity

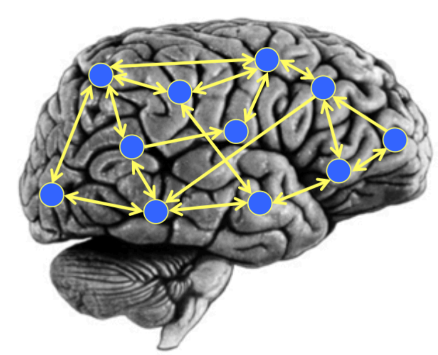
- Functional connectivity
    - Statistical dependency



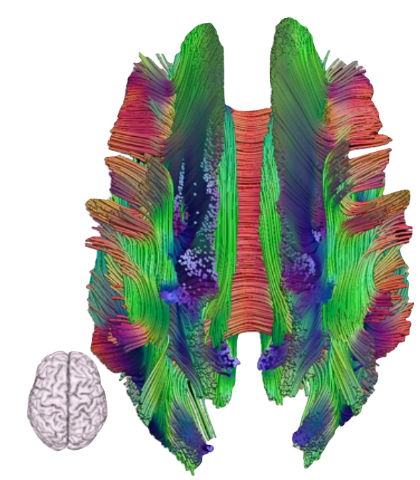
- Dynamic connectivity
    - Changes in functional connectivity over time



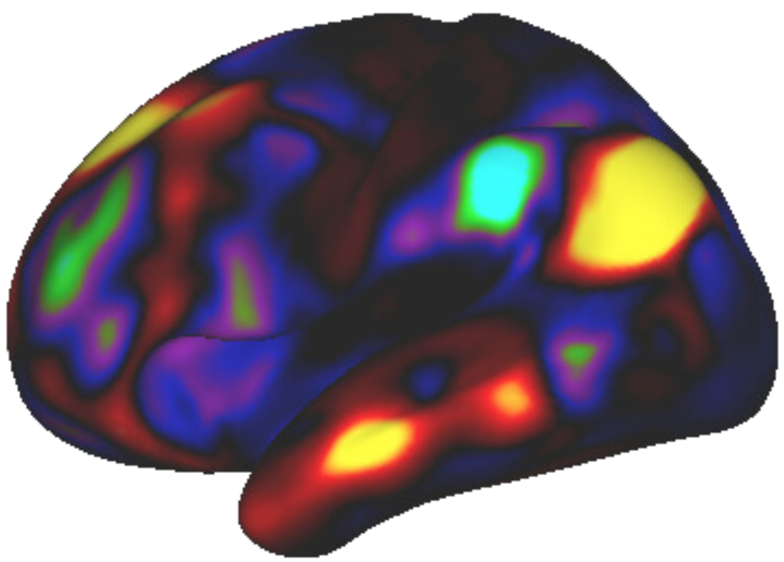
- Effective connectivity
    - Directional influence



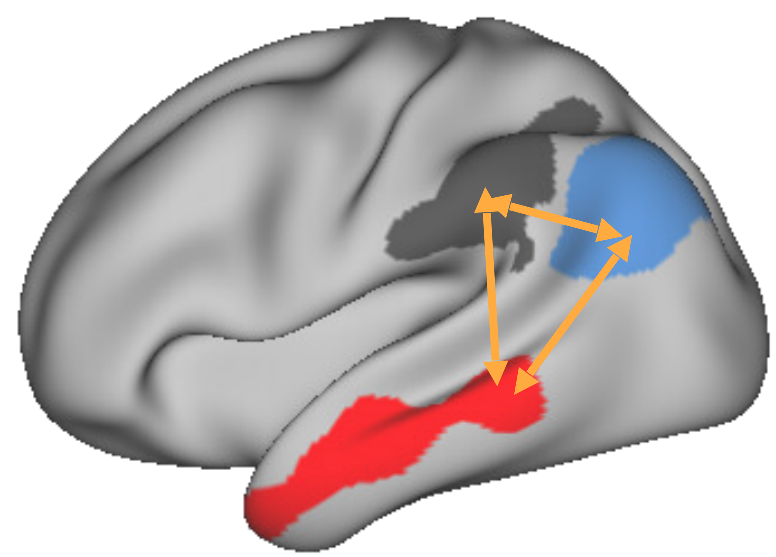
- Anatomical (structural) connectivity
    - Presence of a white matter tract



Overview of resting state methods



Voxel-based methods



Node-based methods

Overview of resting state methods

<u>Voxel-based methods</u>

- Seed-based correlation analysis
  - SCA
- Independent component analysis
  - ICA

<u>Node-based methods</u>

- Network modelling analysis
  - FSLnets
- Graph theory analysis
  - Such as degree, hub, path length

Overview of resting state methods

Voxel-based methods

- Seed-based correlation analysis
  - SCA
- Independent component analysis
  - ICA

Node-based methods

- Network modelling analysis
  - FSLnets
- Graph theory analysis
  - Such as degree, hub, path length

Resting state fMRI and ICA

- Introduction to resting state
- Independent Component Analysis
- Single-subject ICA
- Multi-subject ICA
- Dual regression

Resting state methods

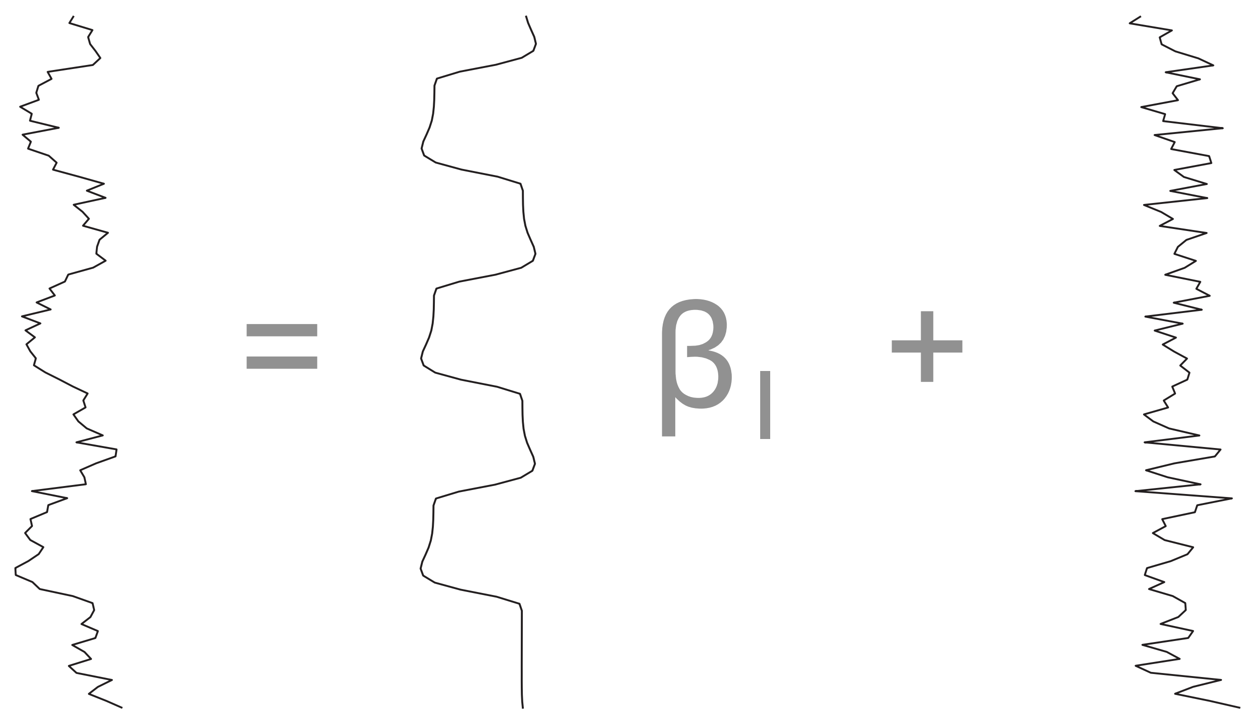
ICA

- Multivariate voxel-based approach
- Finds interesting structure in the data
- Exploratory "model-free" method
- Spatial approach

Network modelling

- Node-based approach (first need to parcellate the brain into functional regions)
- Map connections between specific brain regions (connectomics)
- Temporal approach

Model-based (GLM) analysis



- model each measured time-series as a linear combination of signal and noise
- If the design matrix does not capture every signal, we typically get wrong inferences!

Data Analysis

Confirmatory

- "How well does my model fit to the data?"

Problem → Data →

Model → Analysis

→ Results

- results depend on the model

Exploratory

- "Is there anything interesting in the data?"

Problem → Data →

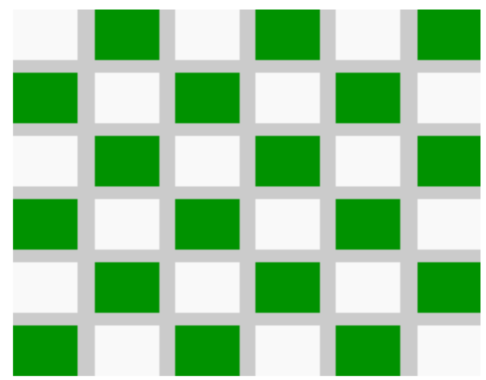
Analysis → Model

→ Results

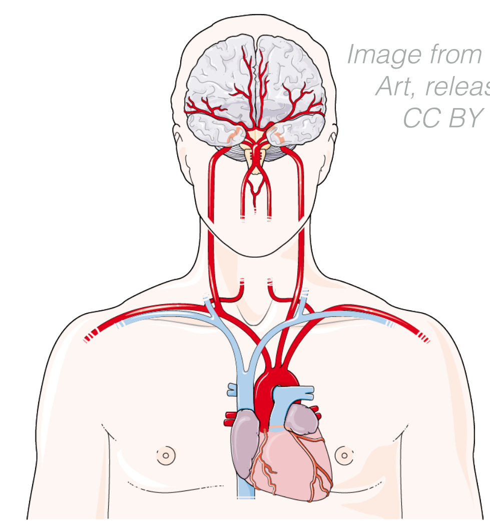
- can give unexpected results

FMRI inferential path

Experiment



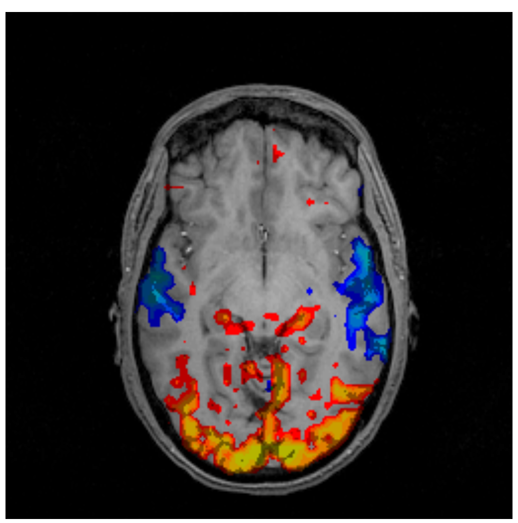
Physiology



MR Physics

Analysis

Interpretation of final results

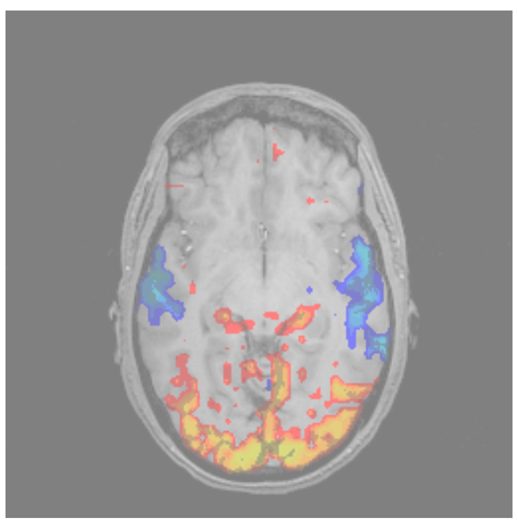


Variability in FMRI

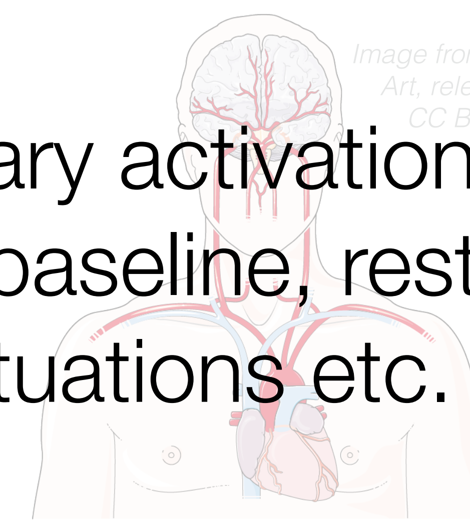
Experiment

suboptimal event timing,
inefficient design, etc.

Interpretation of final results

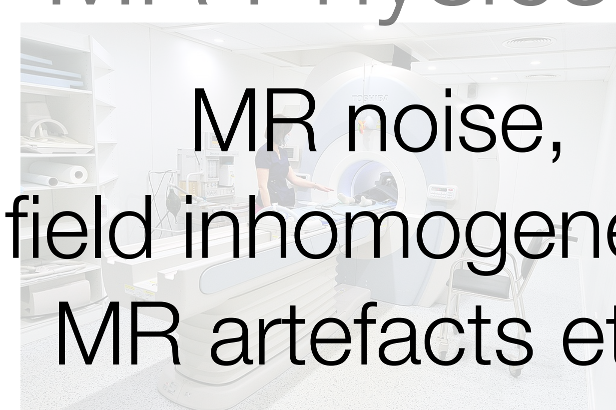


Physiology



secondary activation, ill-defined baseline, resting-fluctuations etc.

MR Physics



MR noise,
field inhomogeneity,
MR artefacts etc.

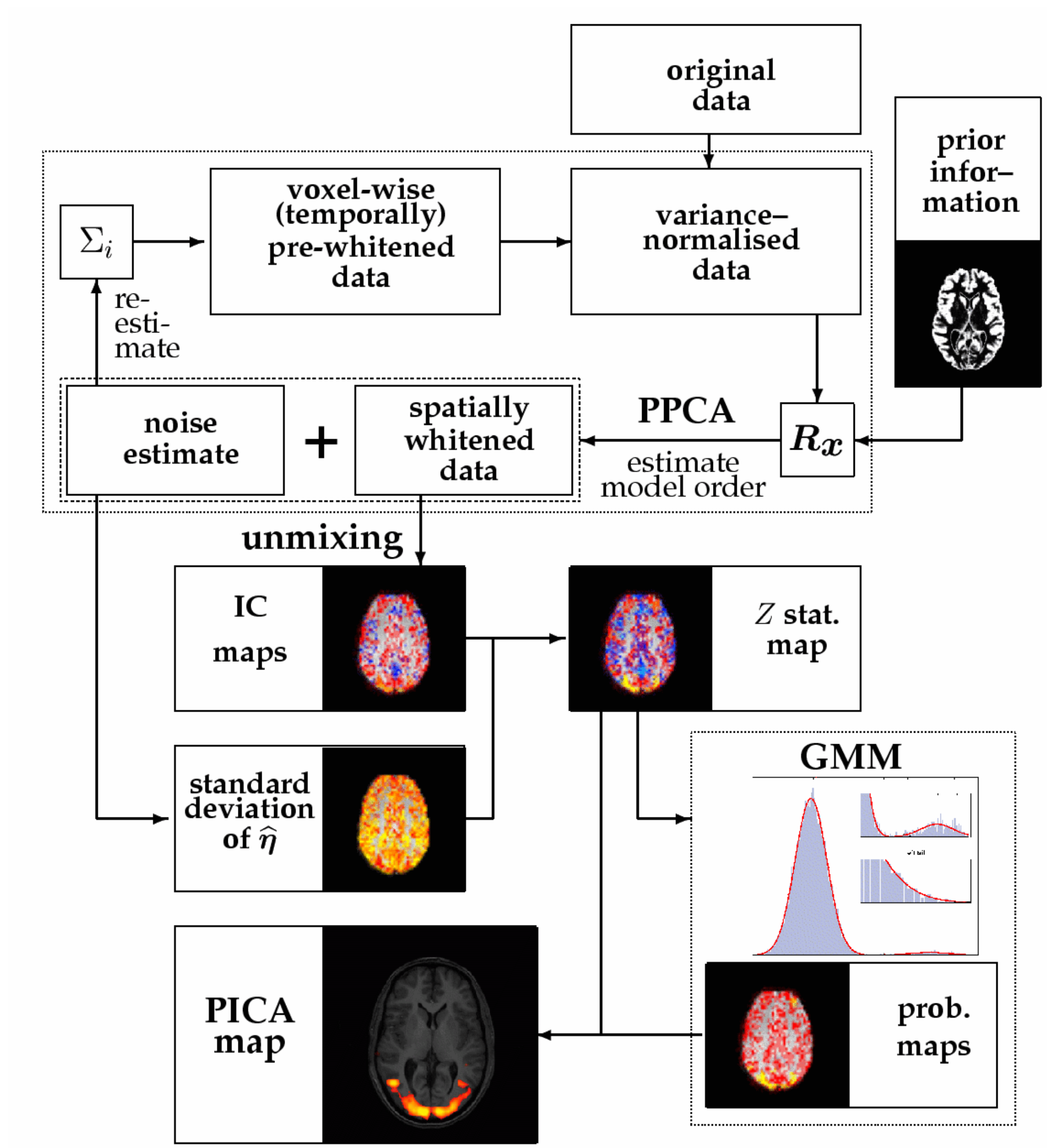
Analysis

filtering & sampling artefacts, design misspecification, stats & thresholding issues etc.

Model-free?

There is no explicit time-series model of assumed ‘activity’

Model-free?



$$Y^i = S^i A^i + E^i, \quad \text{where} \quad E^i_{\cdot j} \sim \mathcal{N}(0, \sigma^2_Y \boldsymbol{I})$$

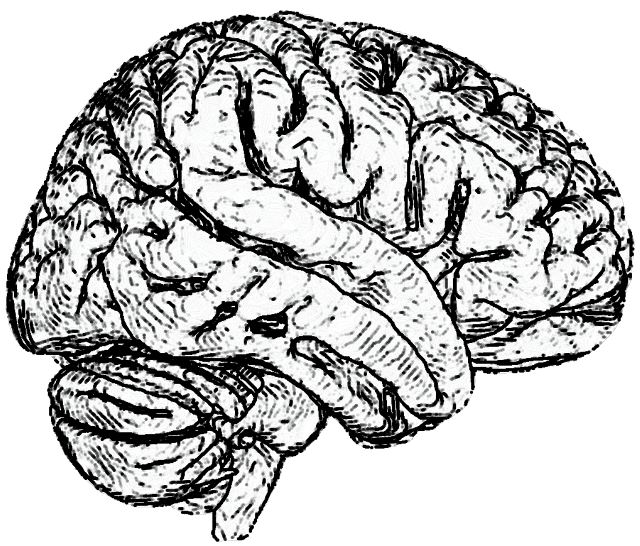
There is an underlying mathematical (generative) model

Decomposition techniques

- try to ‘explain’ / represent the data
  - by calculating quantities that summarise the data
  - by extracting underlying ‘hidden’ features that are ‘interesting’
- differ in what is considered ‘interesting’
  - are localised in time and/or space (Clustering)
  - explain observed data variance (PCA, FDA, FA)
  - are maximally independent (ICA)

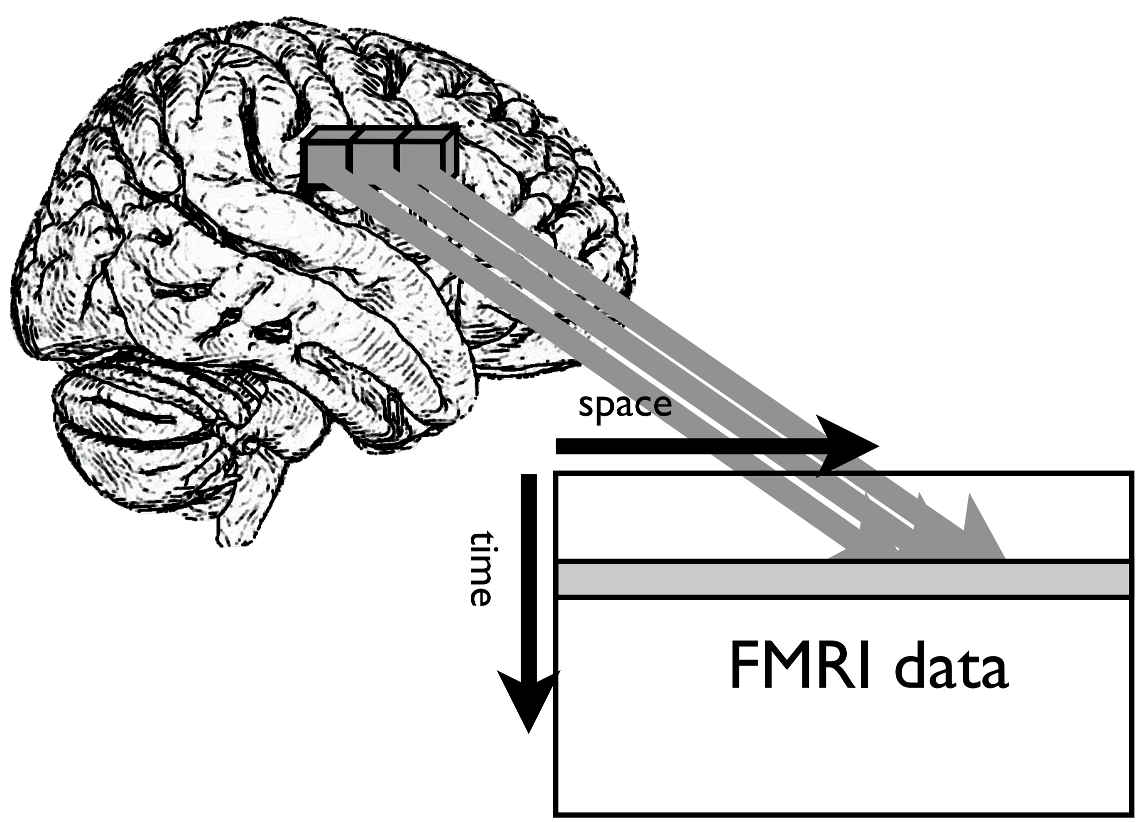
Melodic

multivariate linear decomposition:



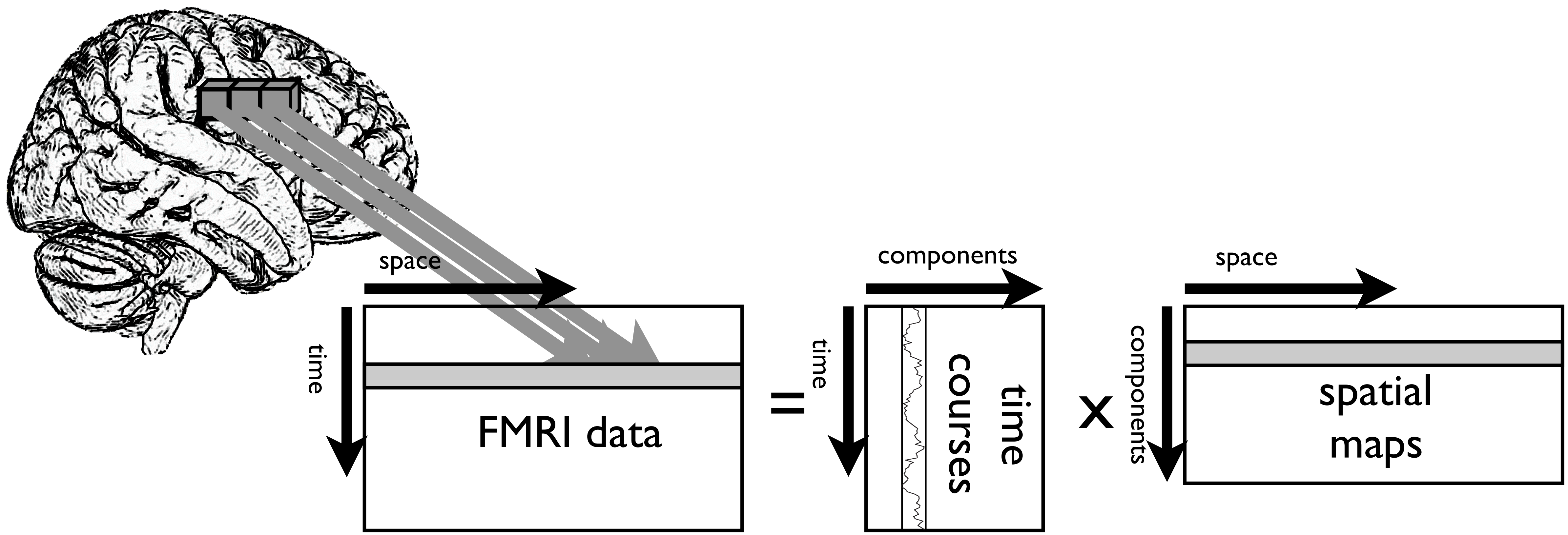
Melodic

multivariate linear decomposition:



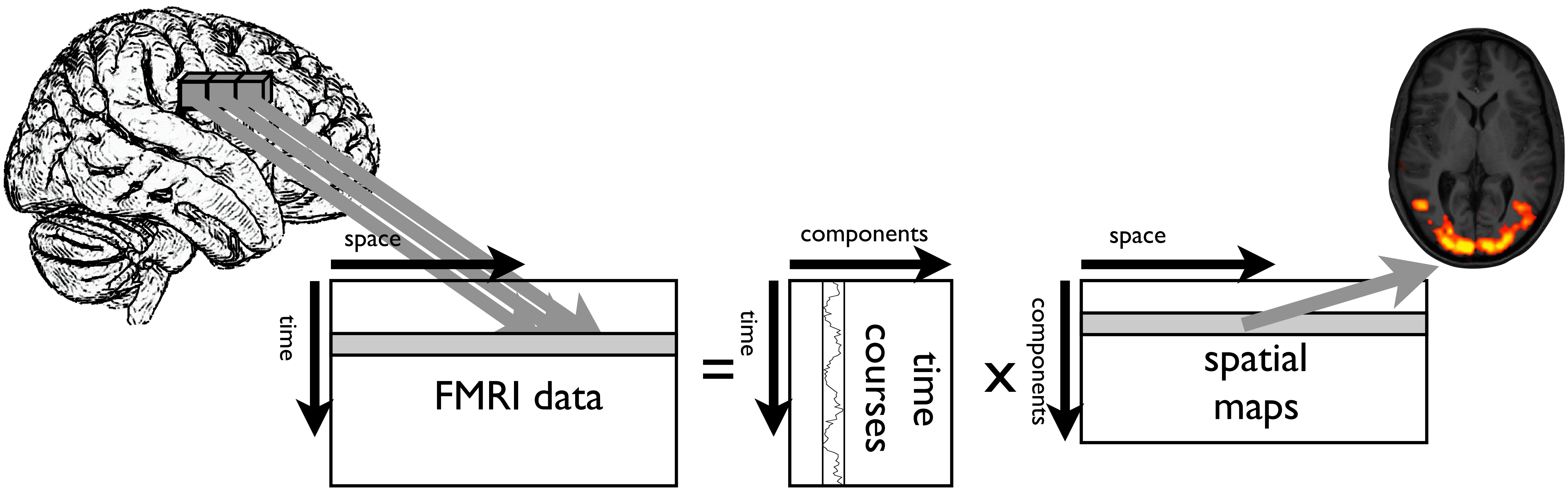
Melodic

multivariate linear decomposition:



Melodic

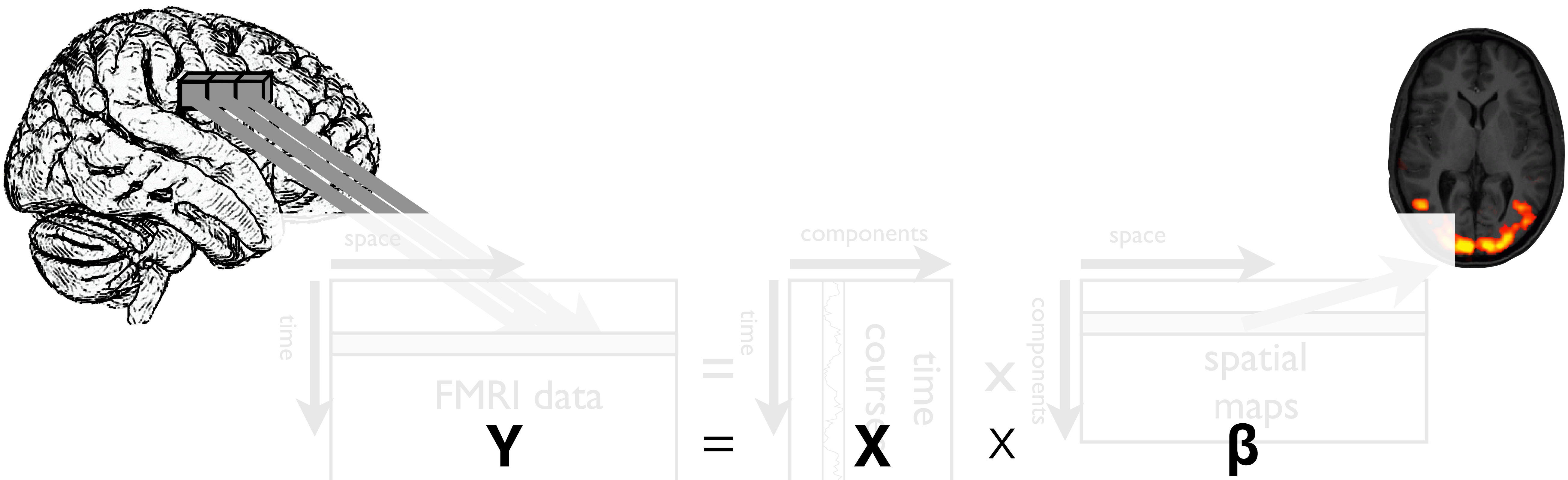
multivariate linear decomposition:



Data is represented as a 2D matrix and decomposed into components

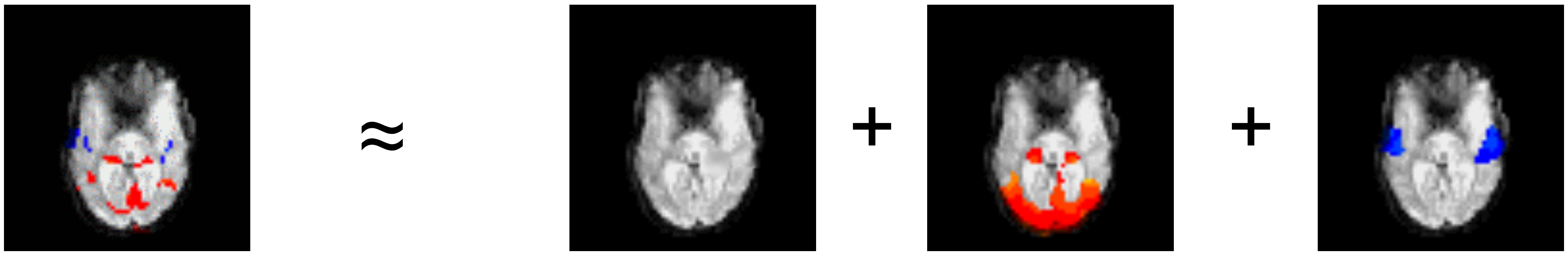
Melodic

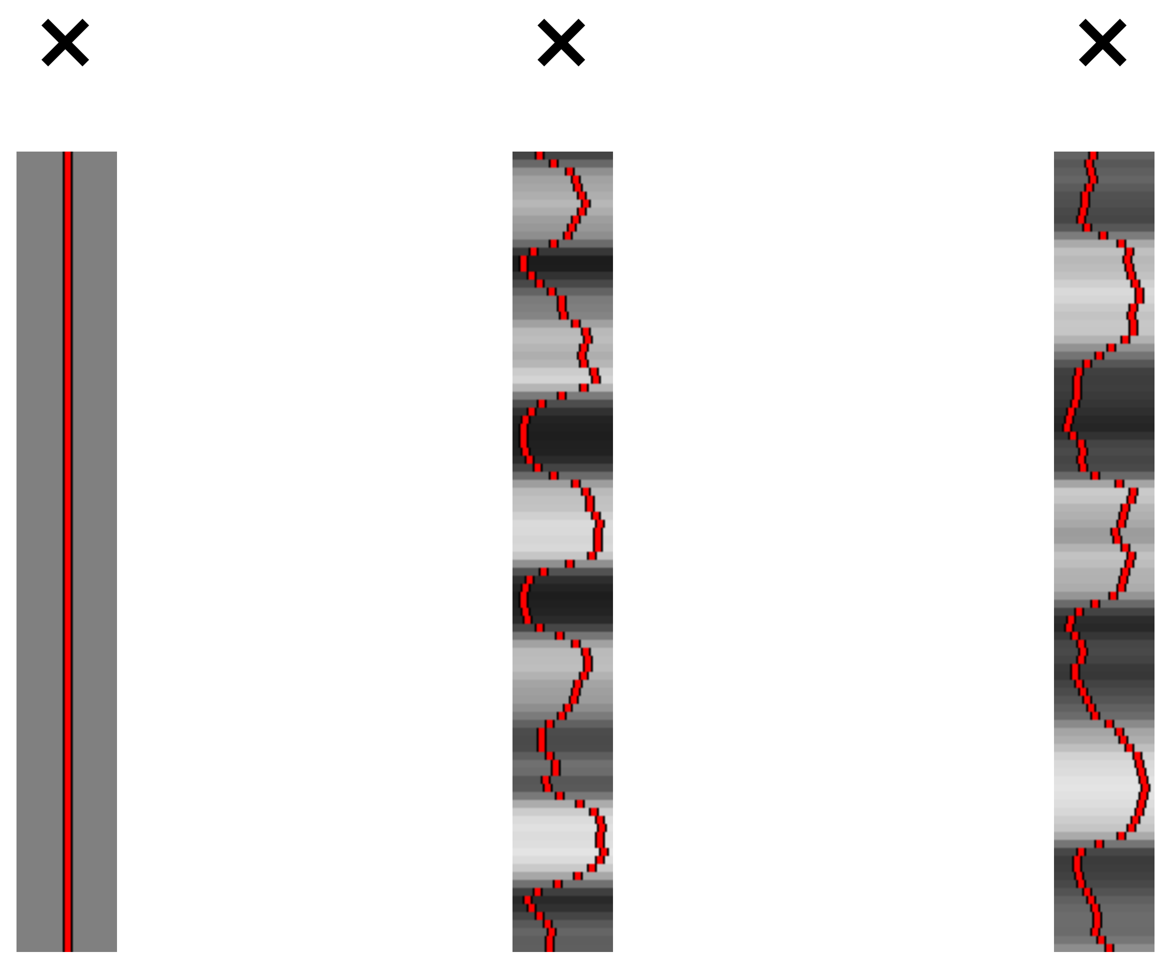
multivariate linear decomposition:



Data is represented as a 2D matrix and decomposed into components

What are components?

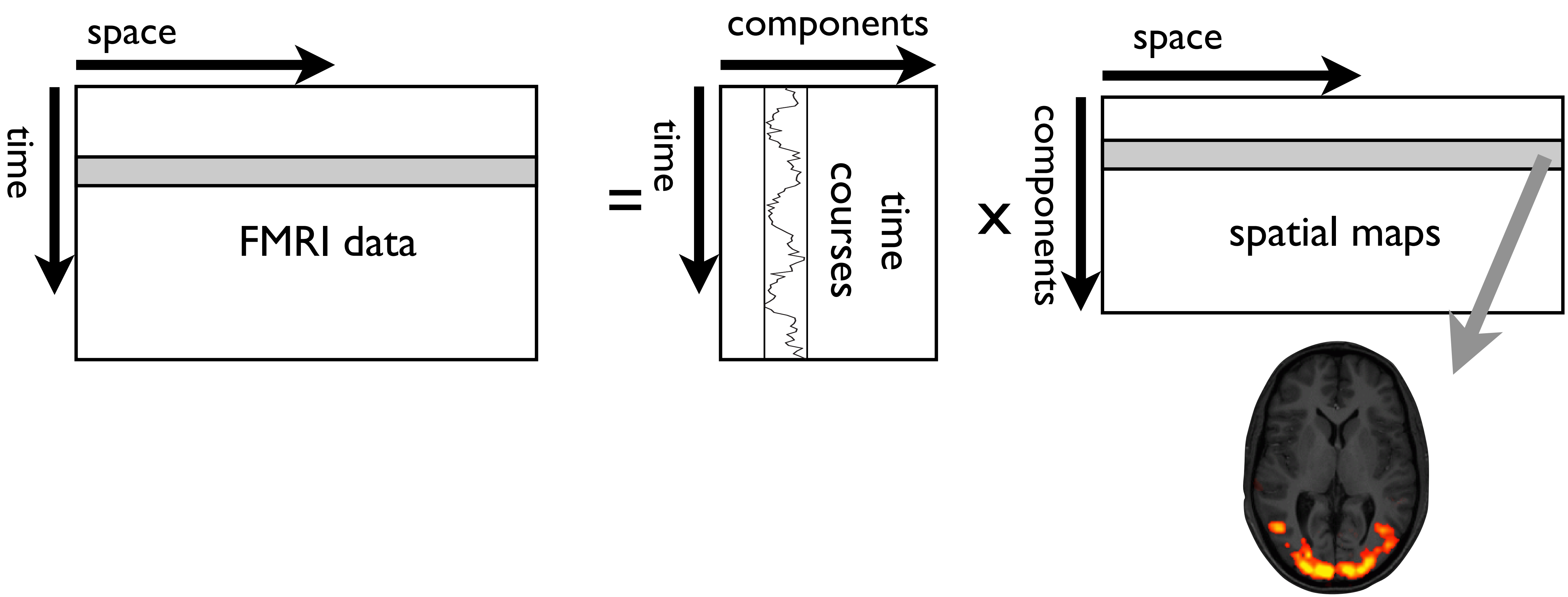




- express observed data as linear combination of spatio-temporal processes

- techniques differ in the way data is represented by components

Spatial ICA for FMRI



– data is decomposed into a set of spatially independent maps and a set of time-courses

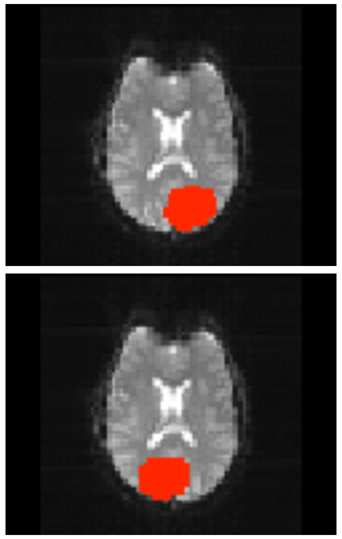
McKeown et al.
HBM 1998

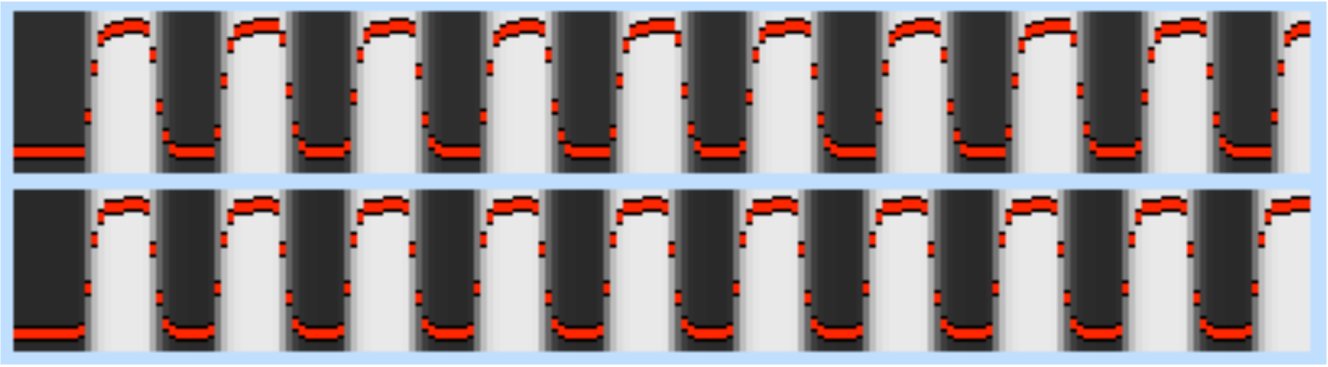
Independence

PCA vs. ICA ?

Simulated Data

(2 components, slightly different timecourses)

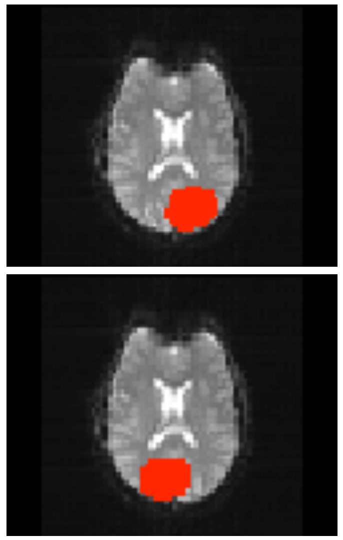


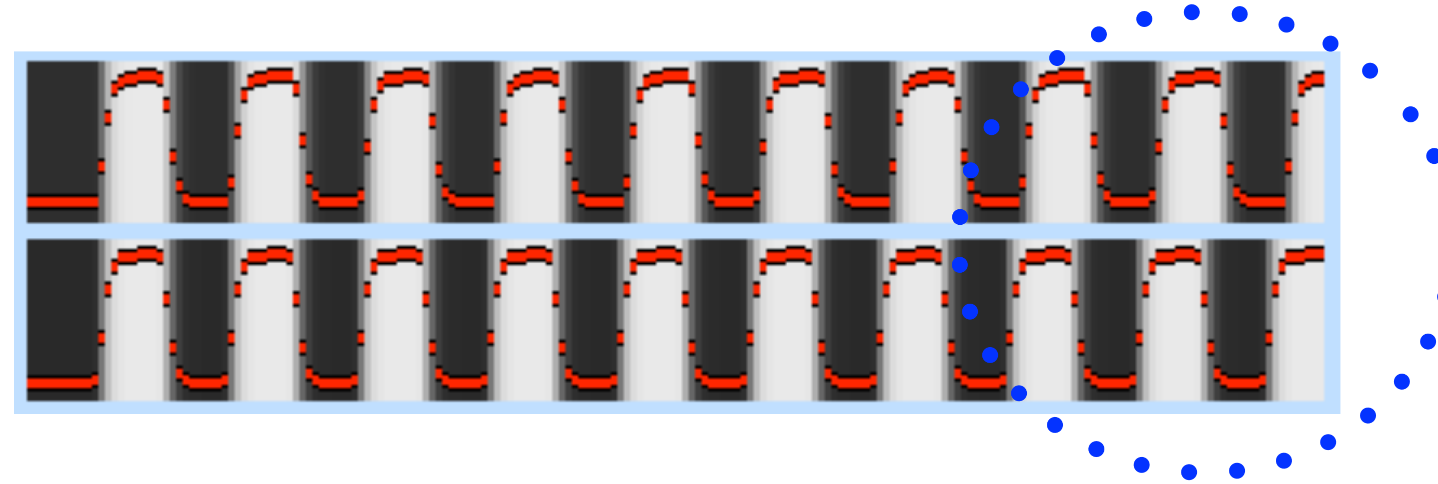


PCA vs. ICA ?

Simulated Data

(2 components, slightly different timecourses)

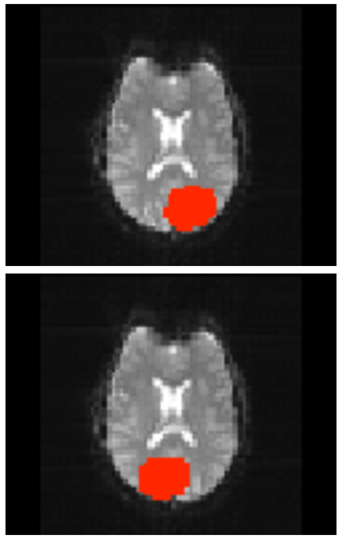


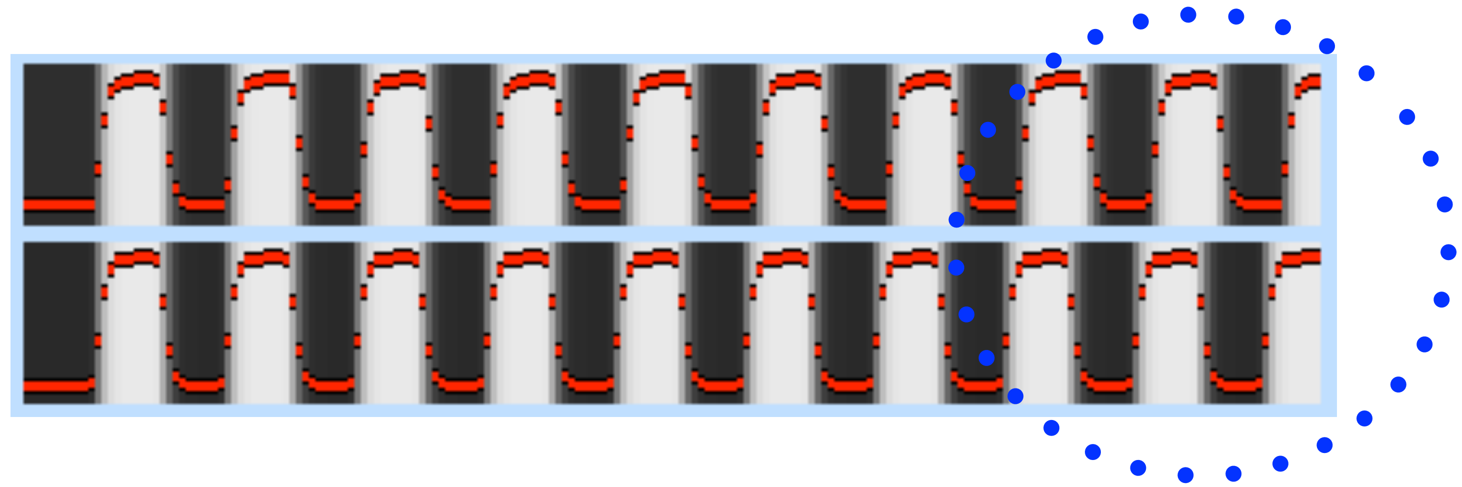


PCA vs. ICA ?

Simulated Data

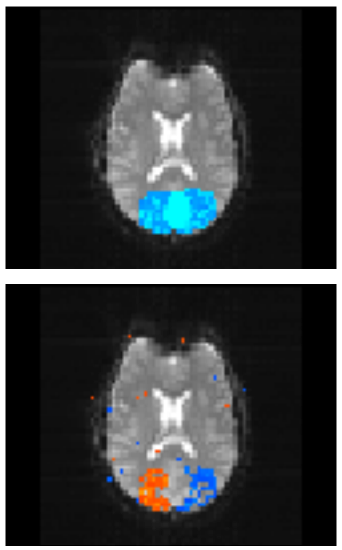
(2 components, slightly different timecourses)

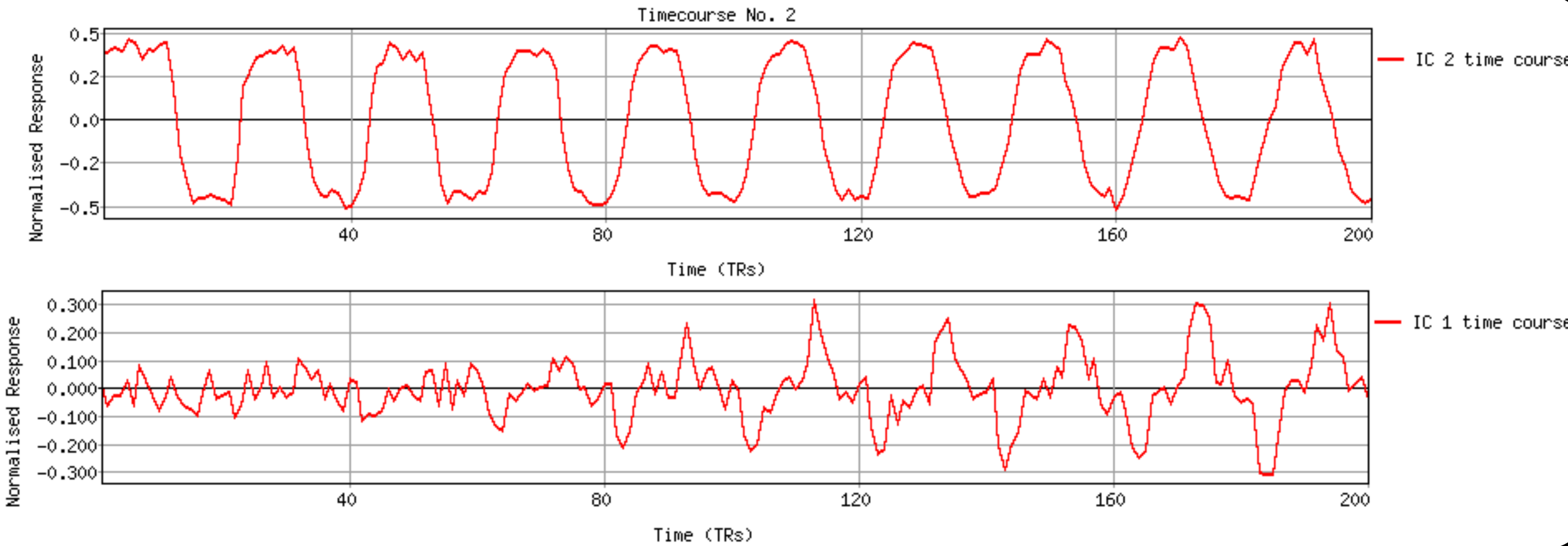




PCA

- Timecourses orthogonal
- Spatial maps and timecourses "wrong"

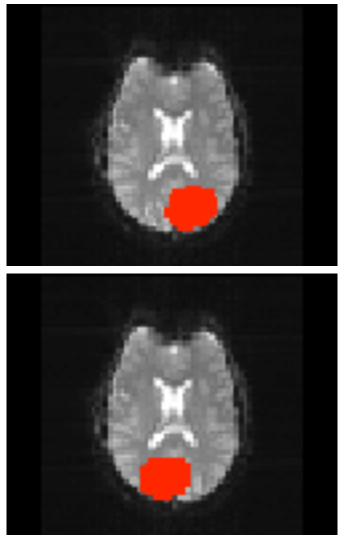


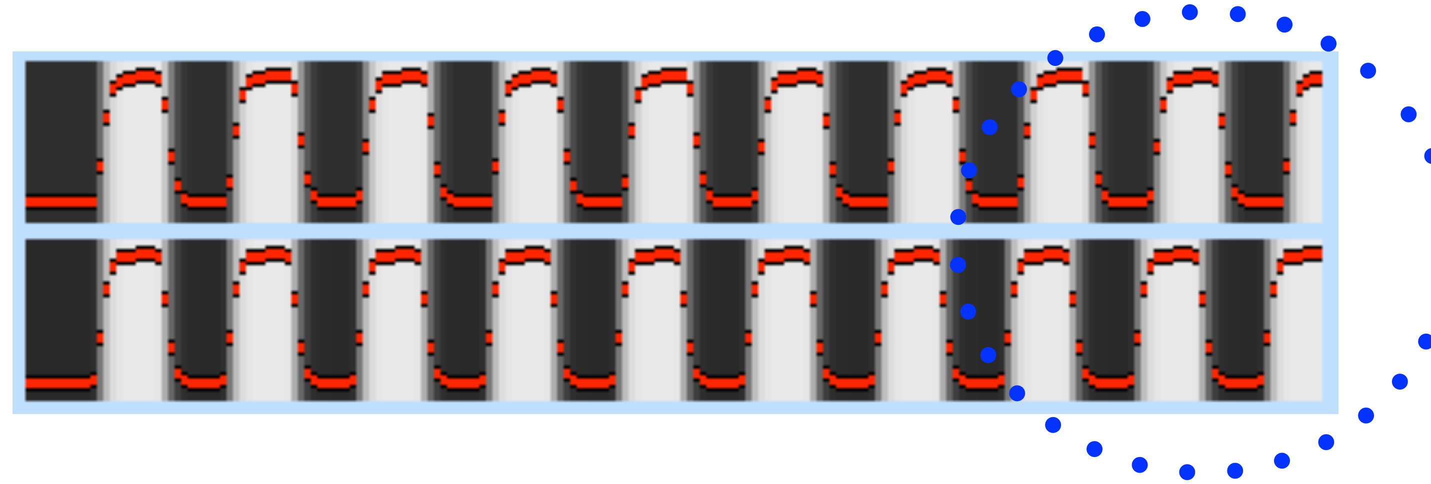


PCA vs. ICA ?

Simulated Data

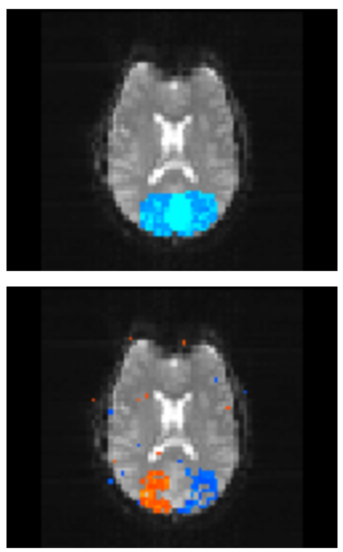
(2 components, slightly different timecourses)

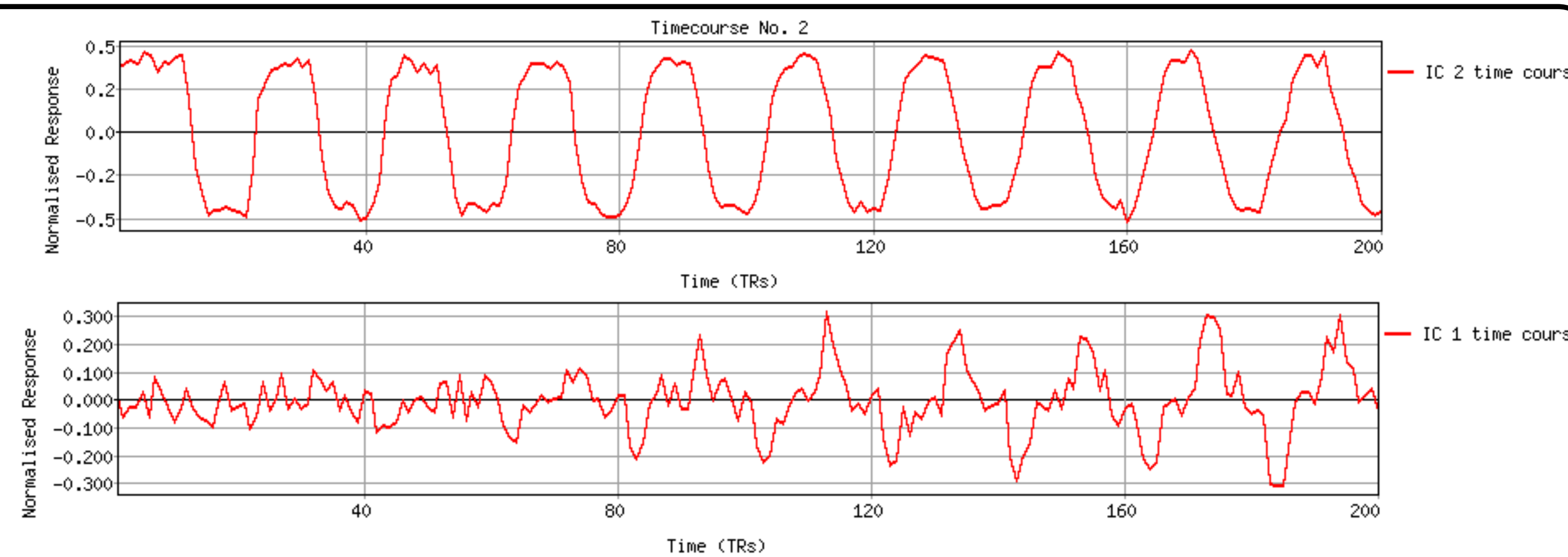




PCA

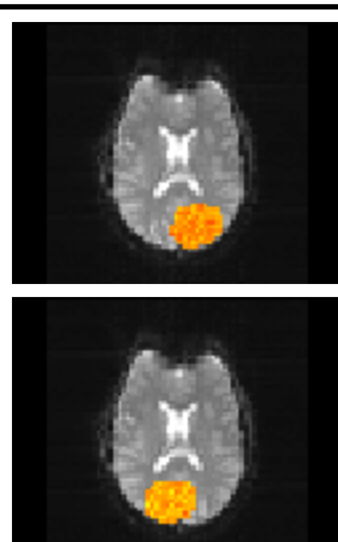
- Timecourses orthogonal
- Spatial maps and timecourses "wrong"

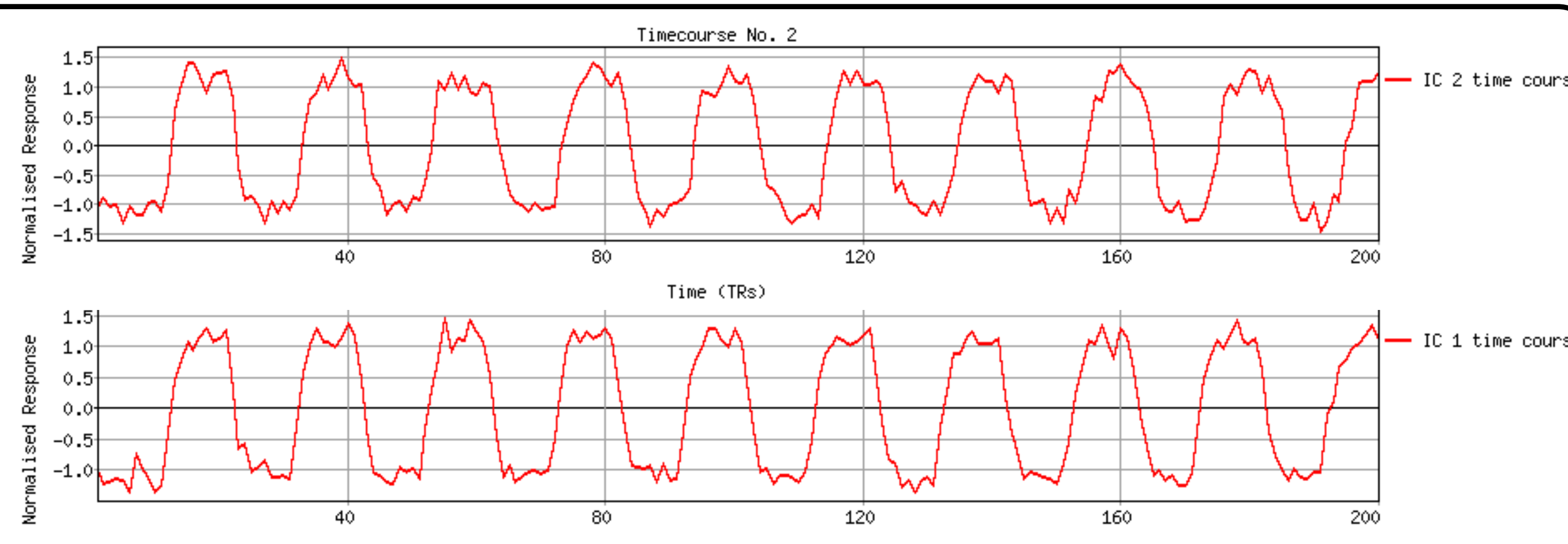




ICA

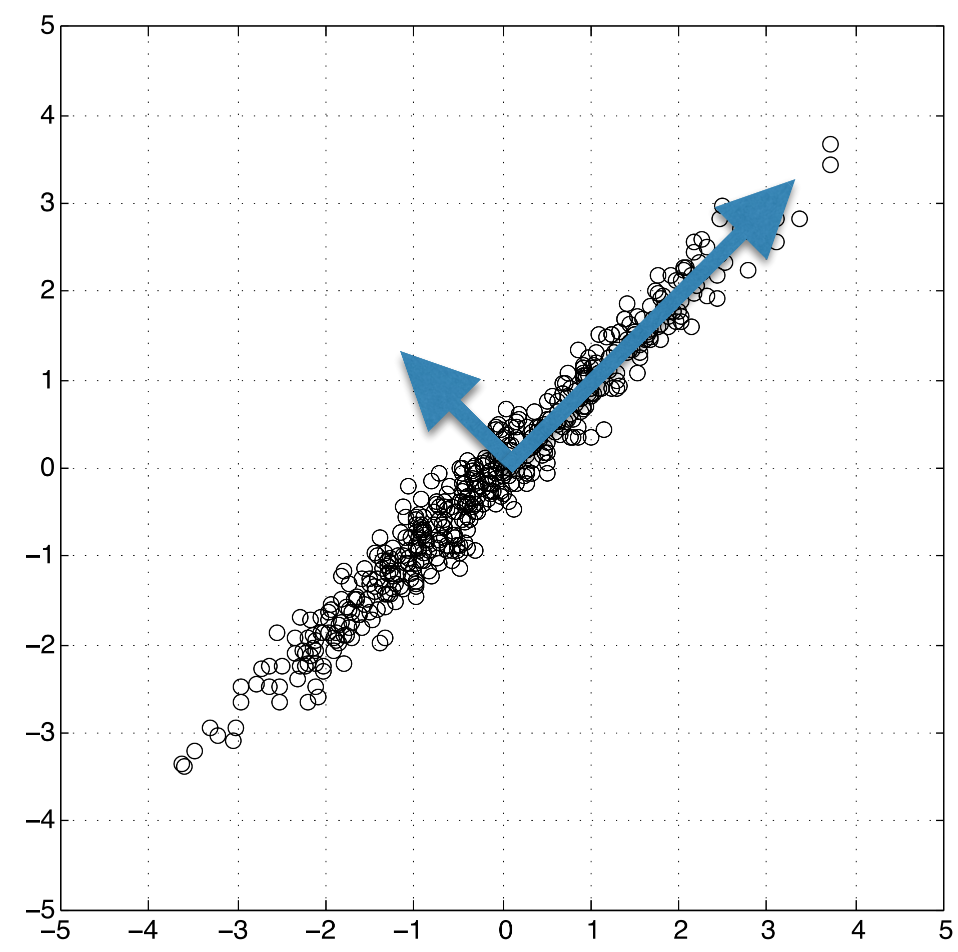
- Timecourses non-co-linear
- Spatial maps and timecourses "right"





PCA vs. ICA

- PCA finds projections of maximum amount of variance in Gaussian data (uses 2nd order statistics only)

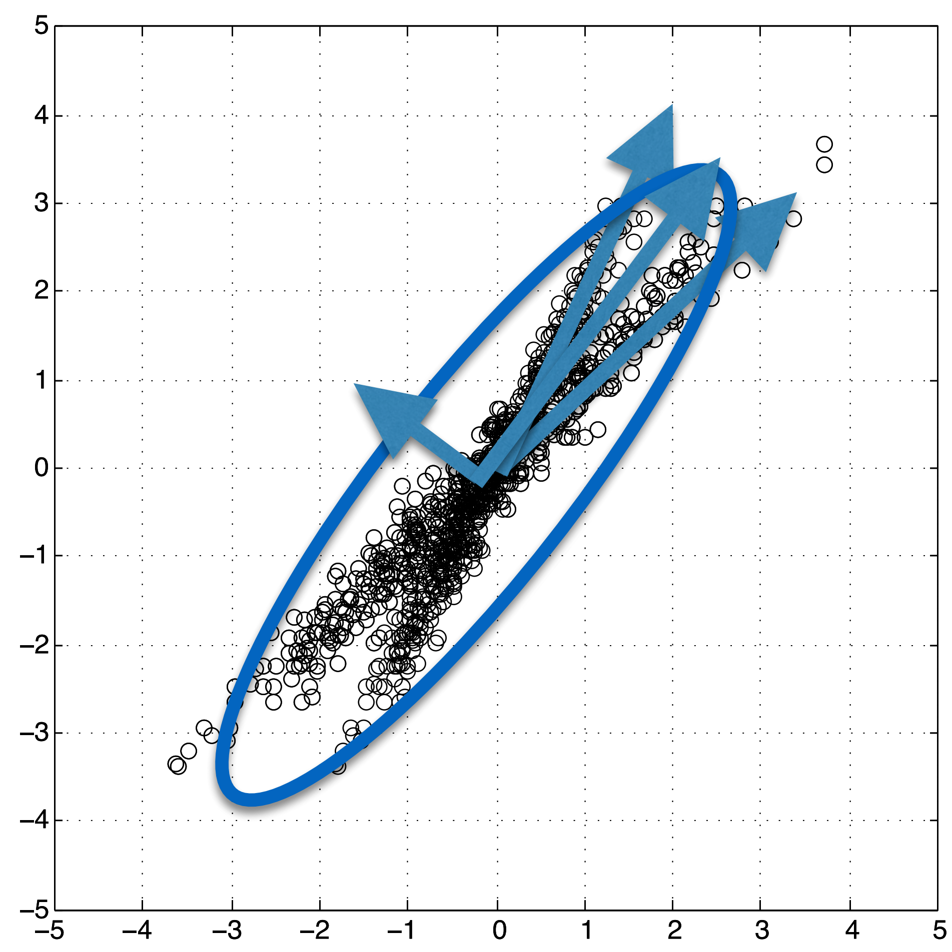


Gaussian data

PCA vs. ICA

- PCA finds projections of maximum amount of variance in Gaussian data (uses 2nd order statistics only)

- Independent Component Analysis (ICA) finds projections of maximal independence in non-Gaussian data (using higher-order statistics)



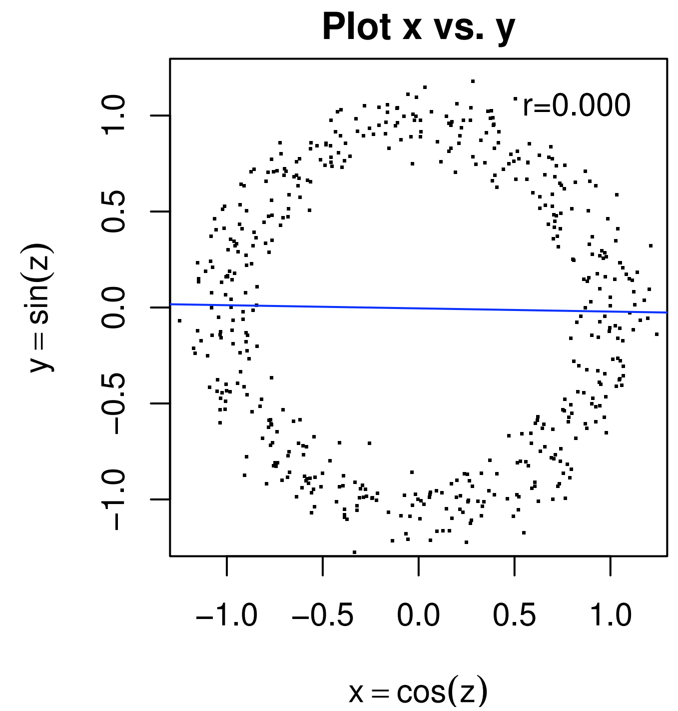
non-Gaussian data

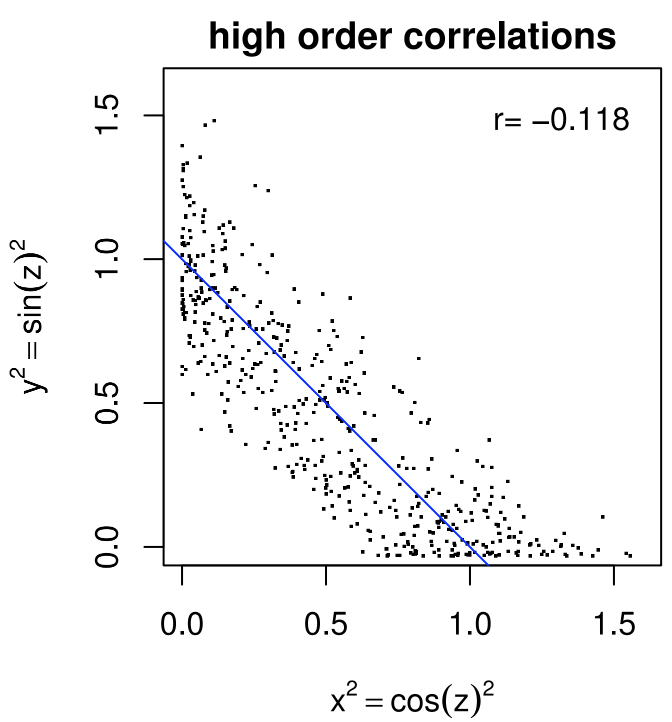
Correlation vs. independence

- de-correlated signals can still be dependent

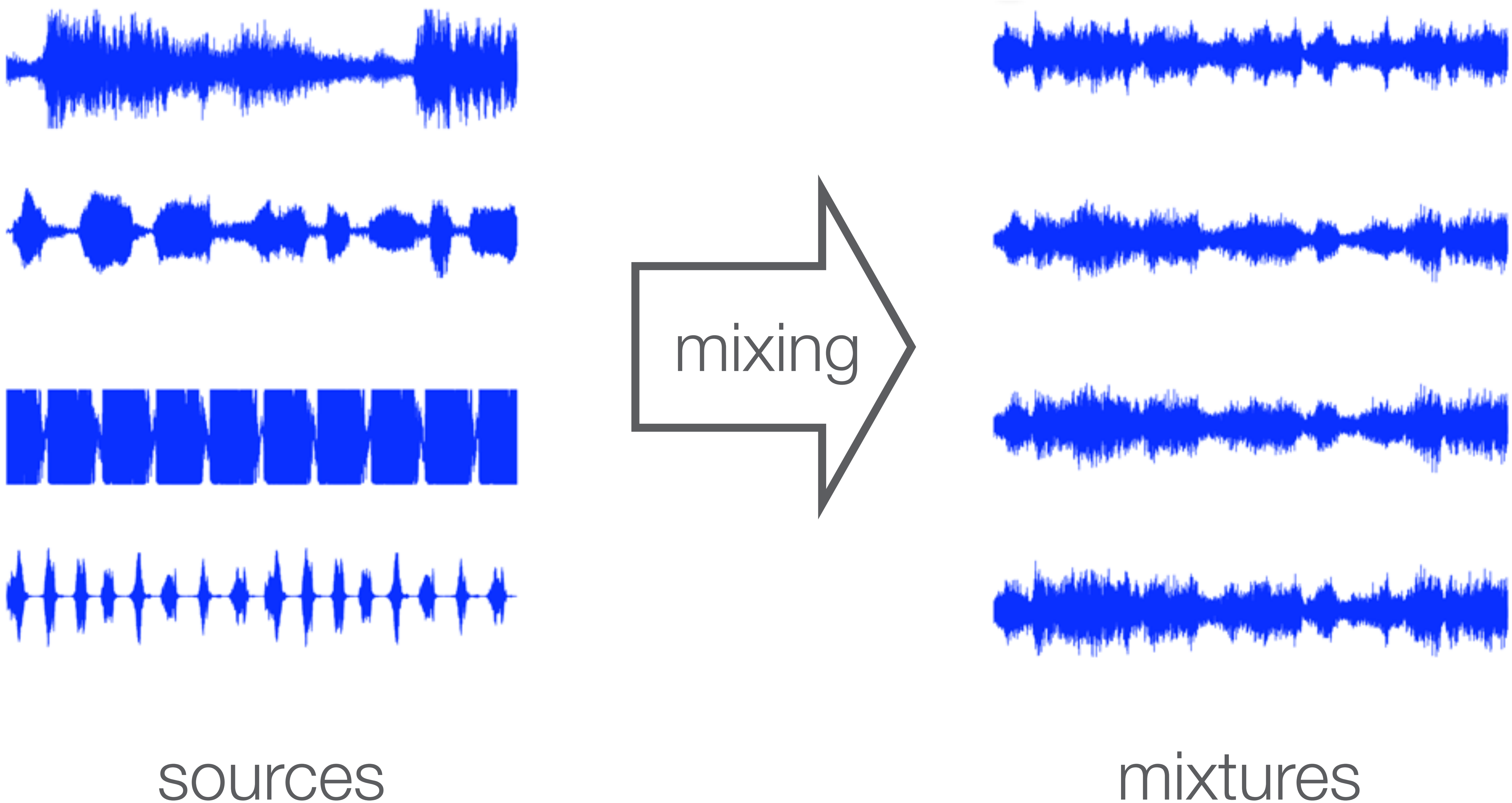
- higher-order statistics (beyond mean and variance) can reveal these dependencies

Stone et al. 2002

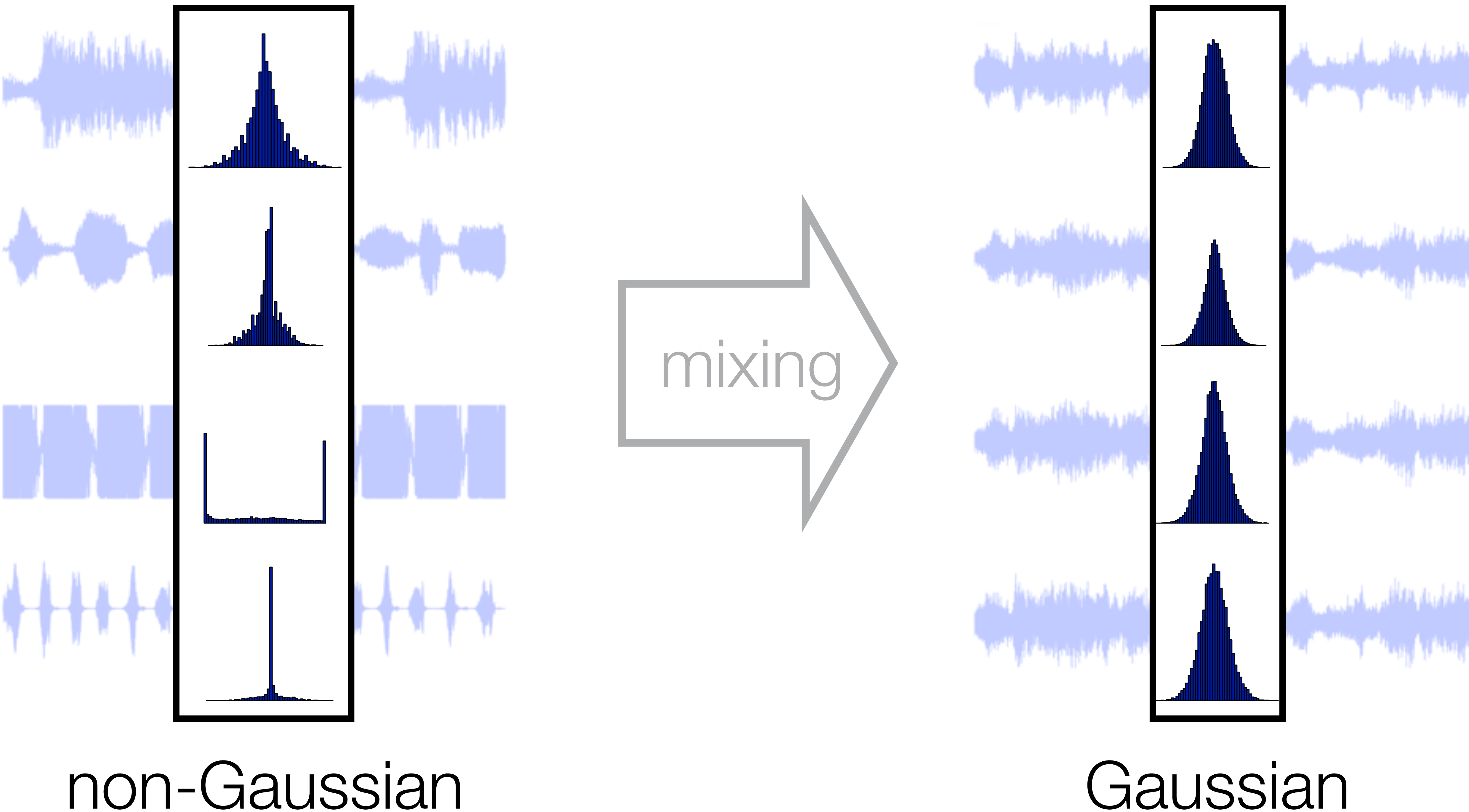




Non-Gaussianity



Non-Gaussianity



ICA estimation

- Random mixing results in more Gaussian-shaped PDFs (Central Limit Theorem)

- conversely:

  if mixing matrix produces less Gaussian-shaped PDFs this is unlikely to be a random result

  ➡ measure non-Gaussianity

- can use neg-entropy as a measure of non-Gaussianity

Hyvärinen & Oja 1997

ICA estimation

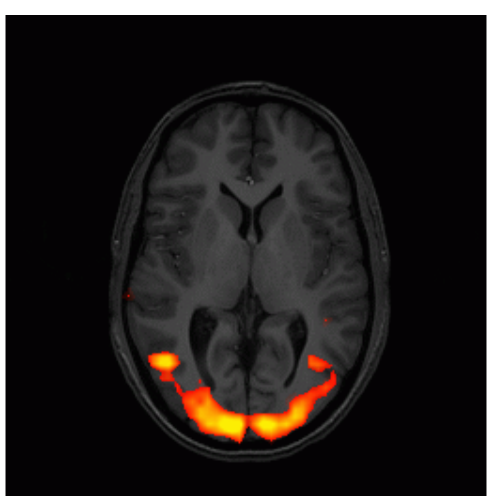
- need to find an unmixing matrix such that the dependency between estimated sources is minimised
- need (i) a contrast (objective/cost) function to drive the unmixing which measures statistical independence and (ii) an optimisation technique:
- kurtosis or cumulants & gradient descent (Jade)
- maximum entropy & gradient descent (Infomax)
- neg-entropy & fixed point iteration (FastICA)

Overfitting & thresholding

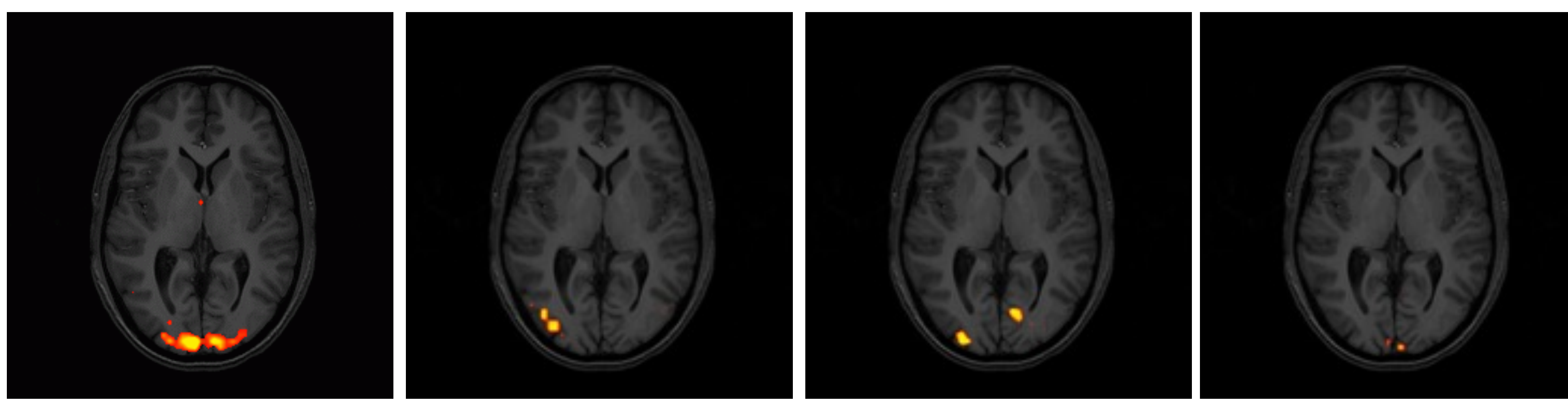
The ‘overfitting’ problem

fitting a noise-free model to noisy observations:

- no control over signal vs. noise (non-interpretable results)
- statistical significance testing not possible



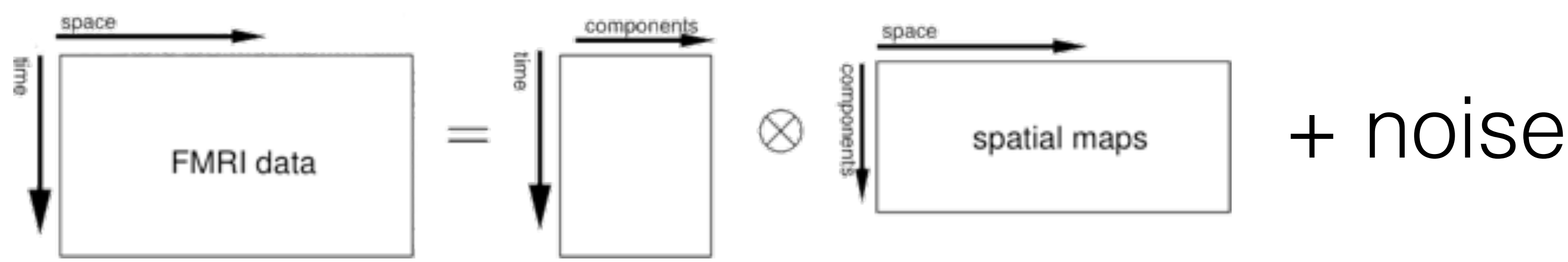
GLM analysis



standard ICA (unconstrained)

Probabilistic ICA model

statistical "latent variables" model: we observe linear mixtures of hidden sources in the presence of Gaussian noise



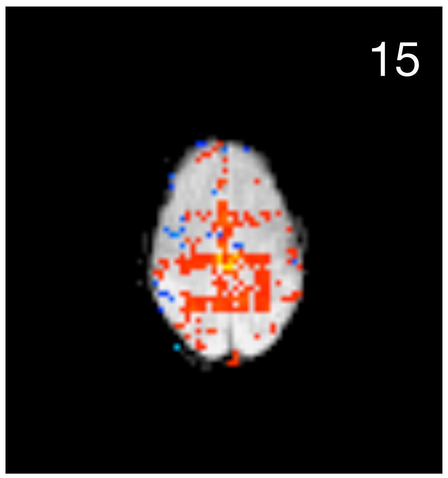
$$\mathbf{Y} = \mathbf{X} \quad \boldsymbol{\beta} + \mathbf{E}$$

Issues:

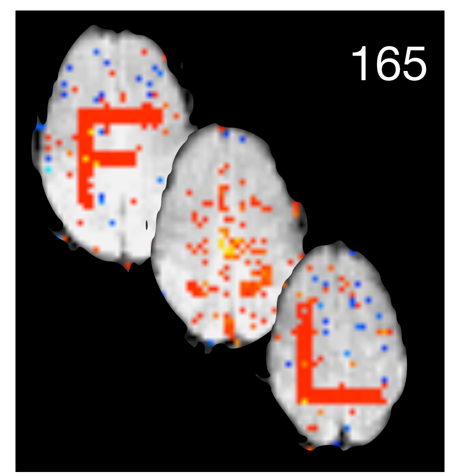
- Model Order Selection: how many components?
- Inference: how to threshold ICs?

Model Order Selection

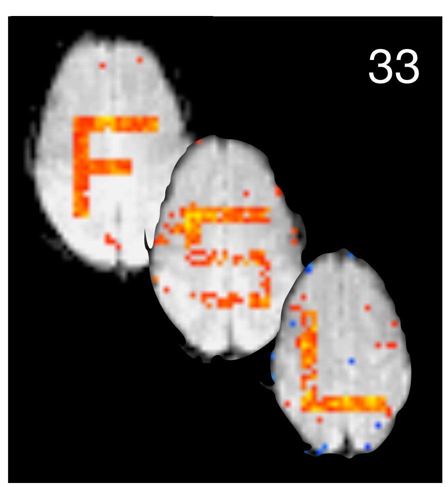
‘How many components’?



under-fitting: the amount of explained data variance is insufficient to obtain good estimates of the signals

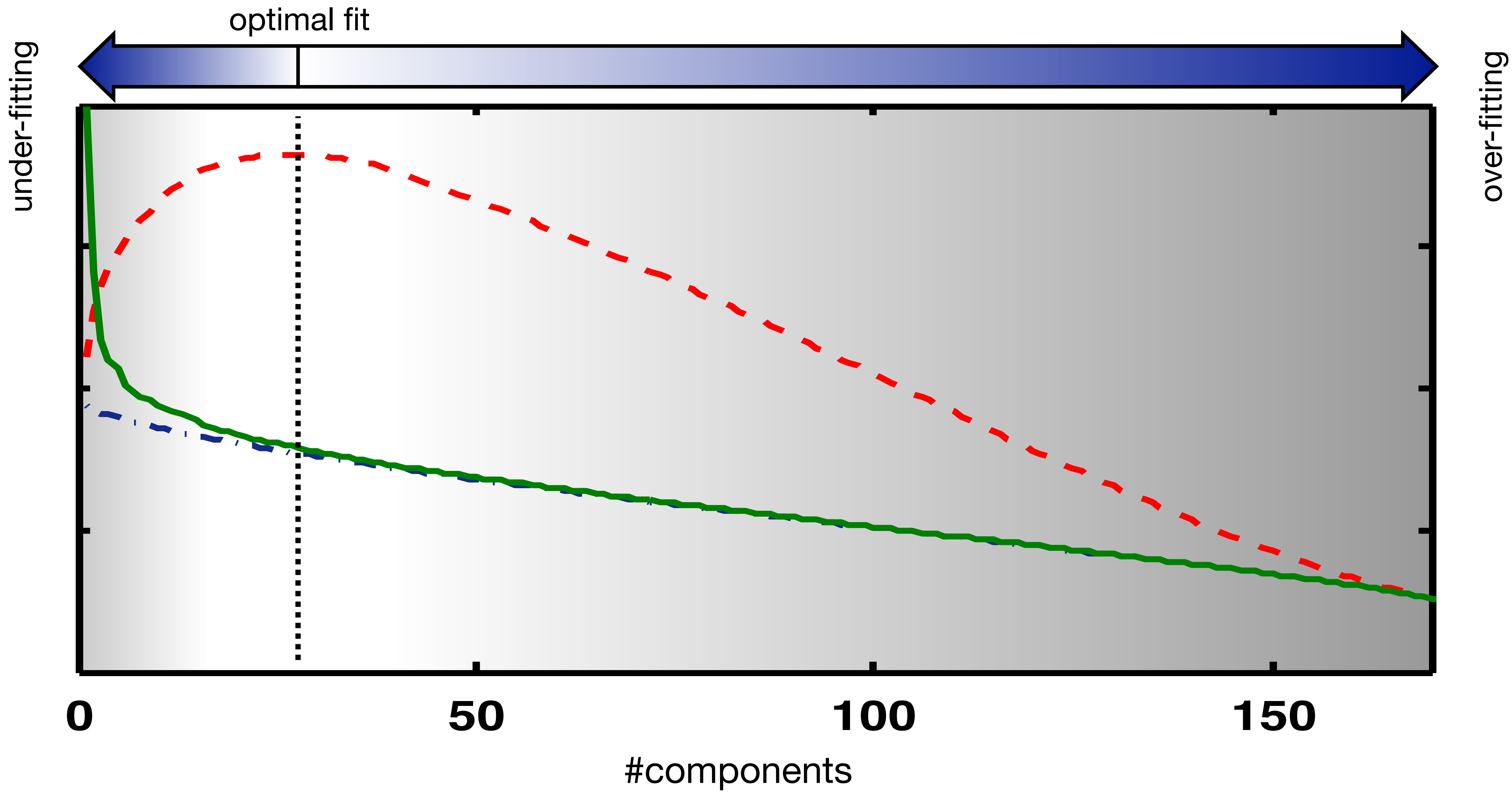


over-fitting: the inclusion of too many components leads to fragmentation of signal across multiple component maps, reducing the ability to identify the signals of interest



optimal fitting: the amount of explained data variance is sufficient to obtain good estimates of the signals while preventing further splits into spurious components

Model Order Selection

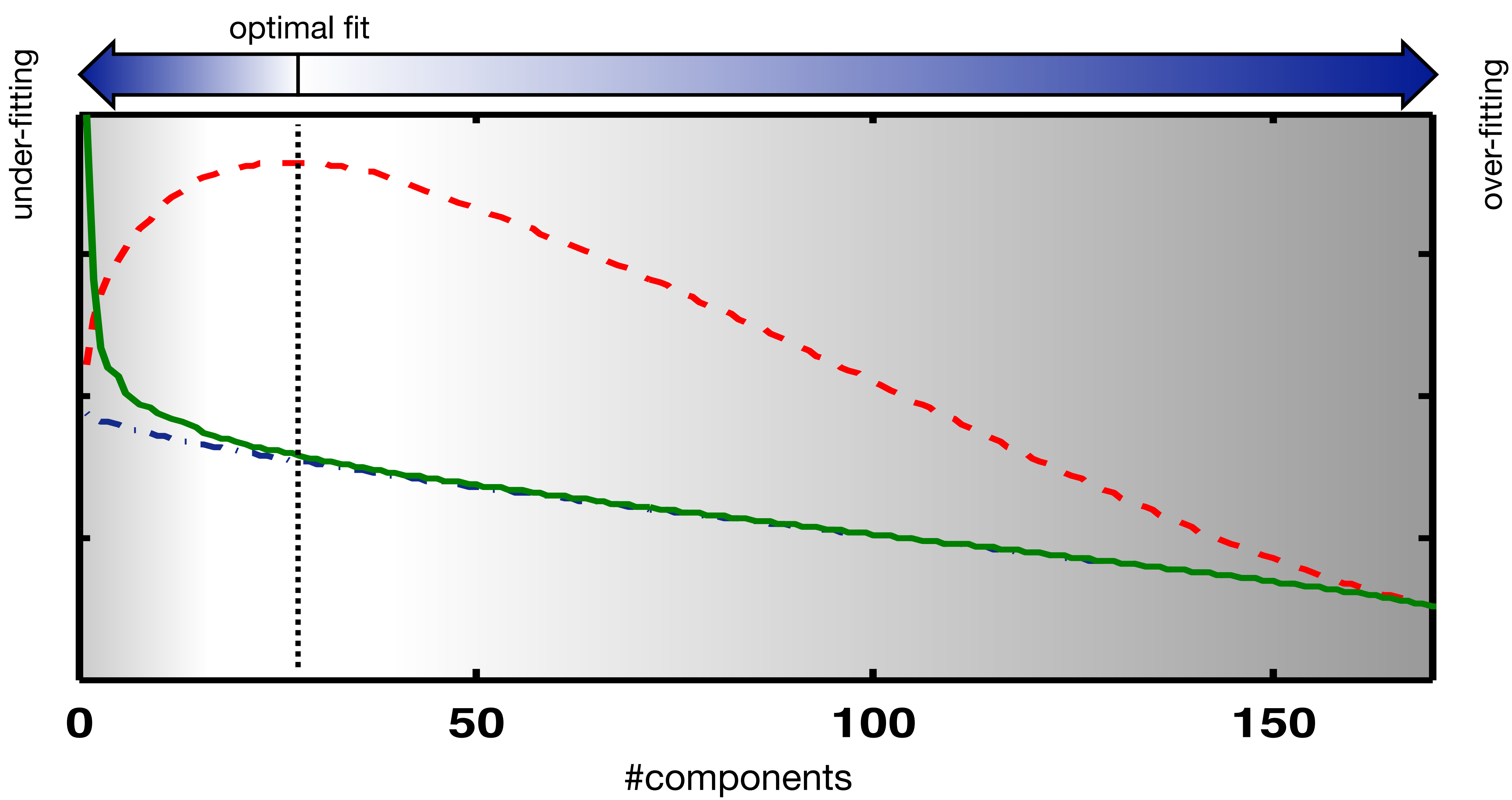


observed Eigenspectrum of the data covariance matrix

Laplace approximation of the posterior probability of the model order

theoretical Eigenspectrum from Gaussian noise

Model Order Selection

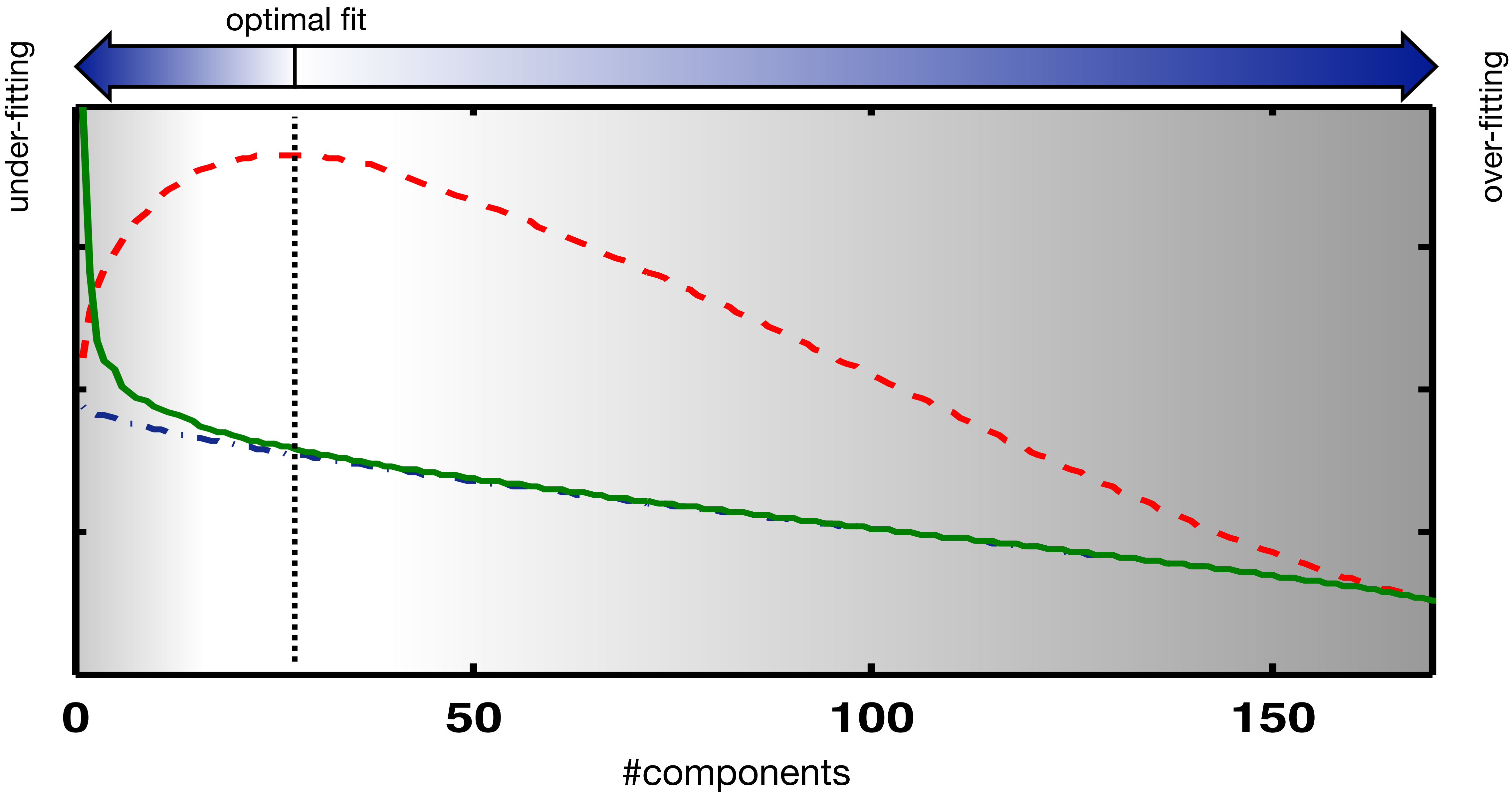


observed Eigenspectrum of the data covariance matrix

Laplace approximation of the posterior probability of the model order

theoretical Eigenspectrum from Gaussian noise

Model Order Selection

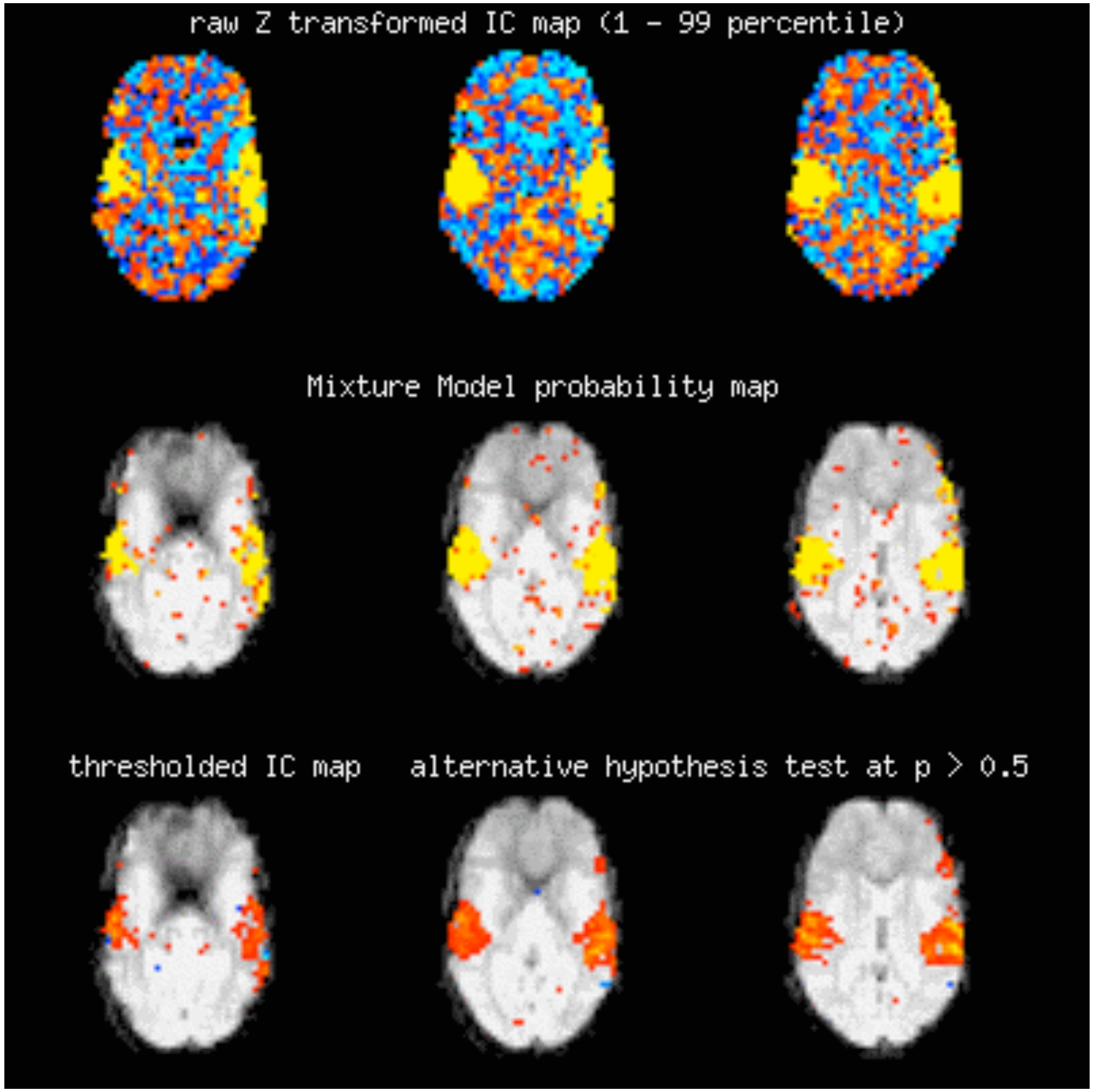


observed Eigenspectrum of the data covariance matrix

Laplace approximation of the posterior probability of the model order

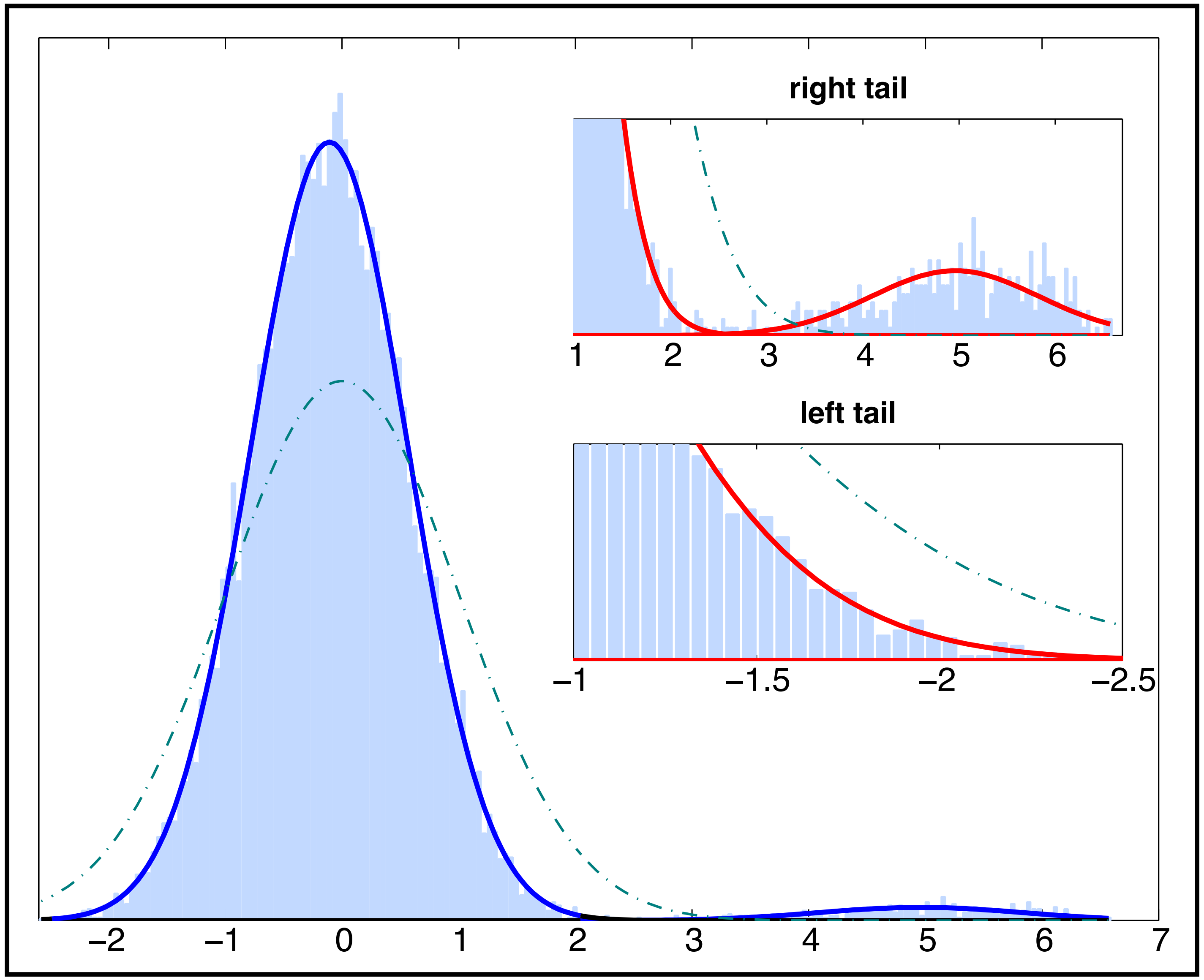
theoretical Eigenspectrum from Gaussian noise

Thresholding

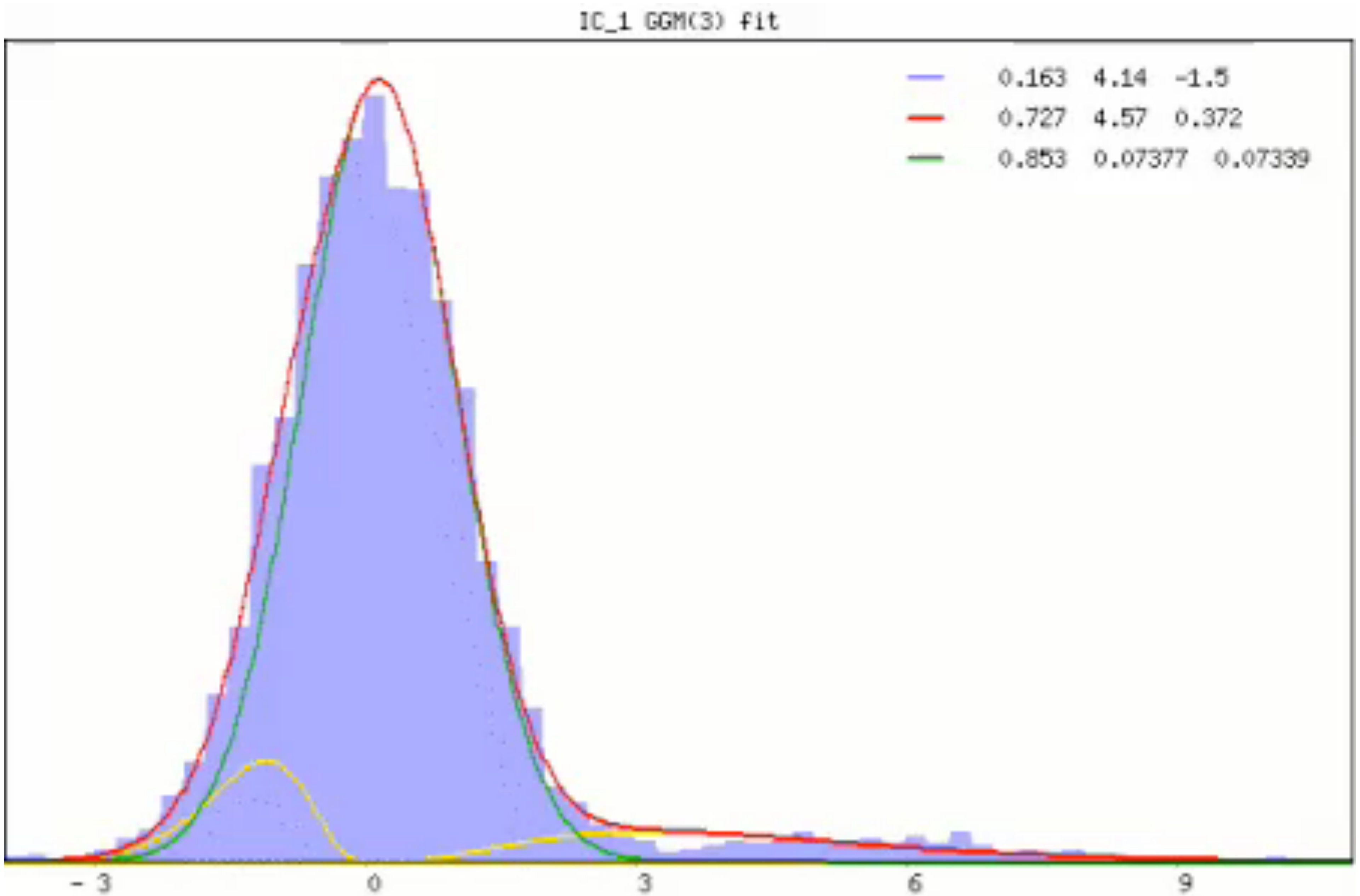


Thresholding

- classical null-hypothesis testing is invalid
- data is assumed to be a linear combination of signals and noise
- the distribution of the estimated spatial maps is a mixture distribution!

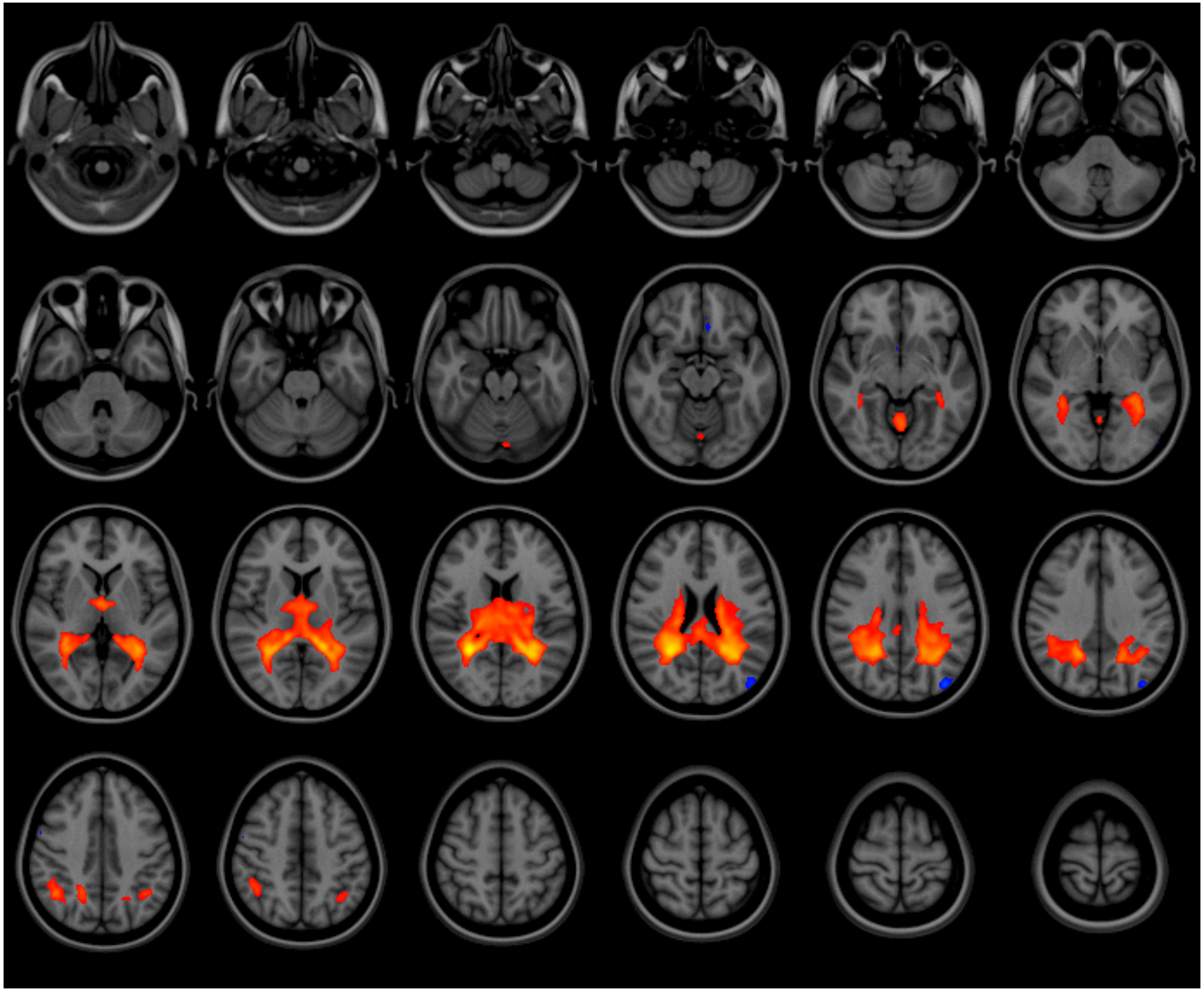


Alternative Hypothesis Test



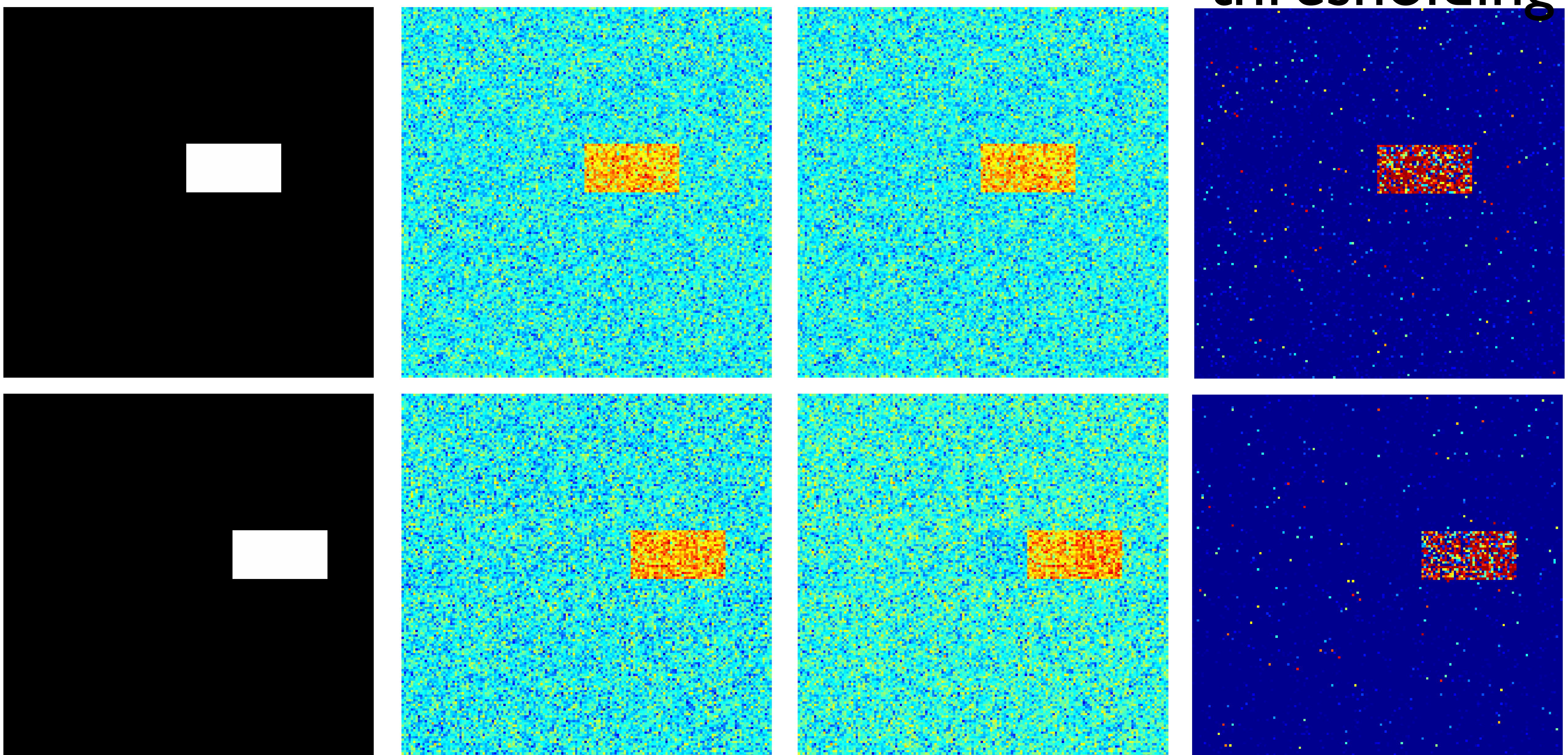
- use Gaussian/Gamma mixture model fitted to the histogram of intensity values (using EM)

What about overlap?



What about overlap?

Sources	Sources + noise	ICA solution	after thresholding



$\rho = 0.5$	$\rho < 0.1$	$\rho = 0$	$\rho \approx 0.5$

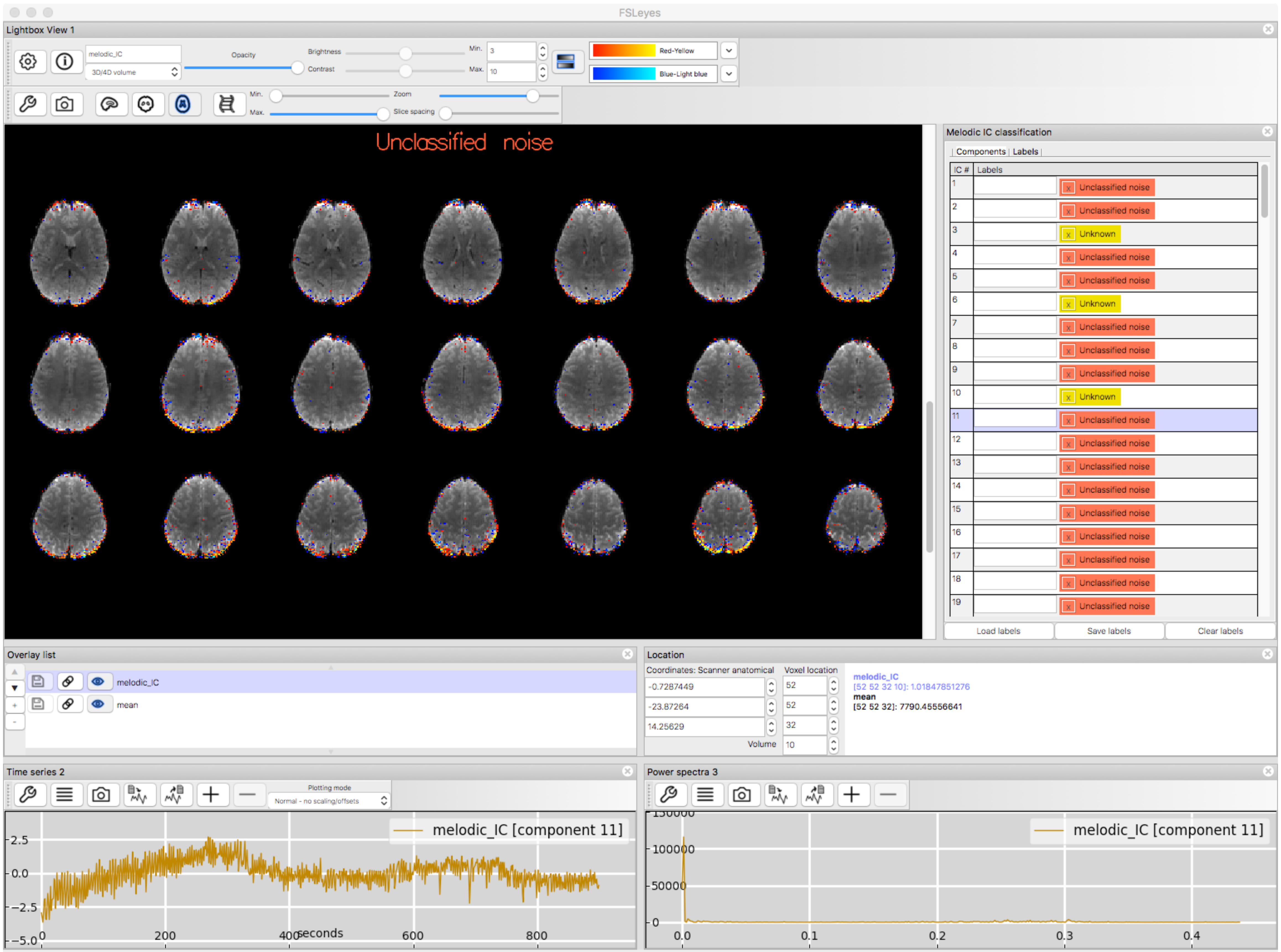
Resting state fMRI and ICA

- Introduction to resting state
- Independent Component Analysis
- Single-subject ICA
- Multi-subject ICA
- Dual regression

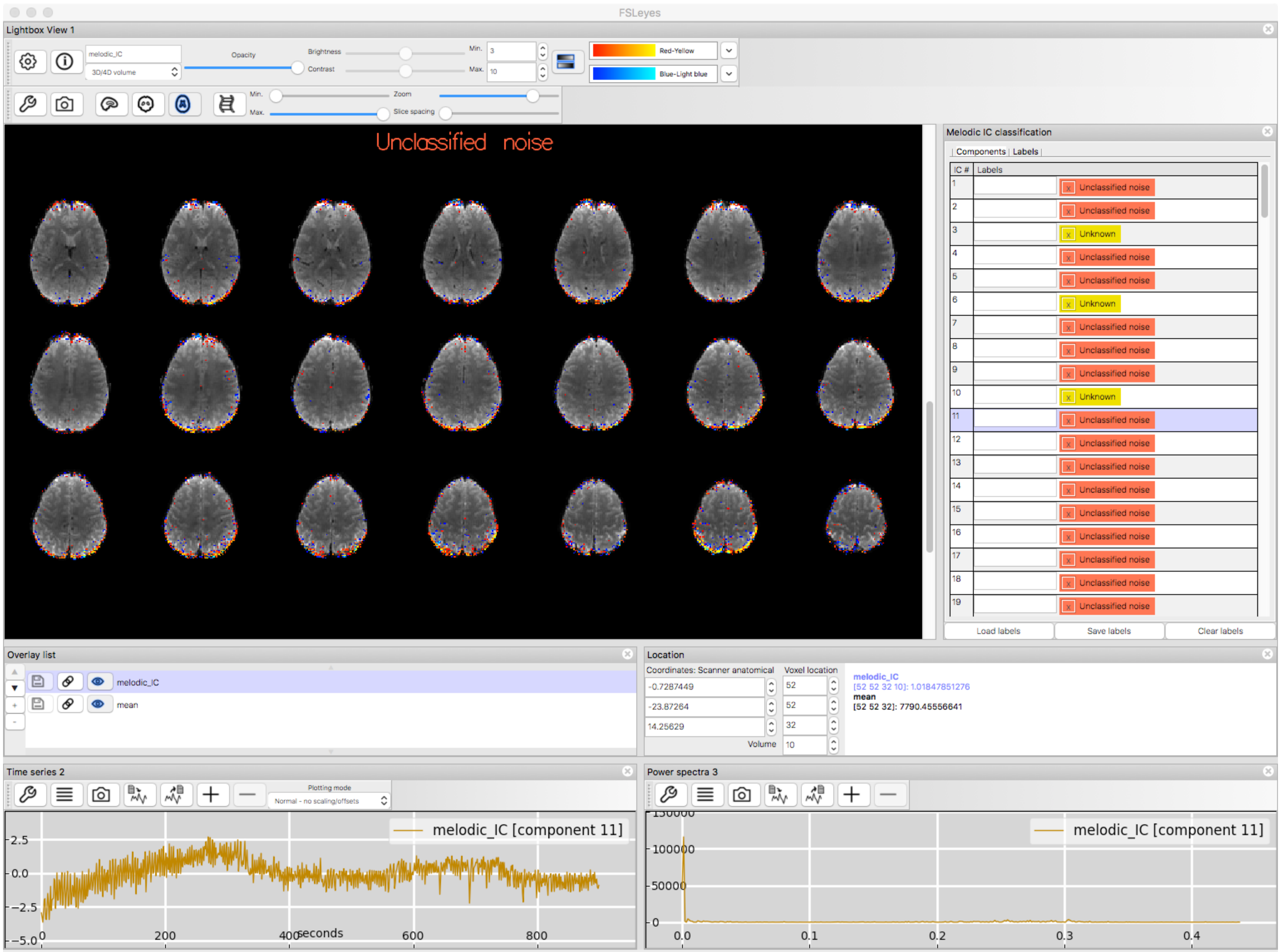
Artefact detection

- FMRI data contain a variety of source processes
- Artifactual sources typically have unknown spatial and temporal extent and cannot easily be modelled accurately
- Exploratory techniques do not require a priori knowledge of time-courses and spatial maps

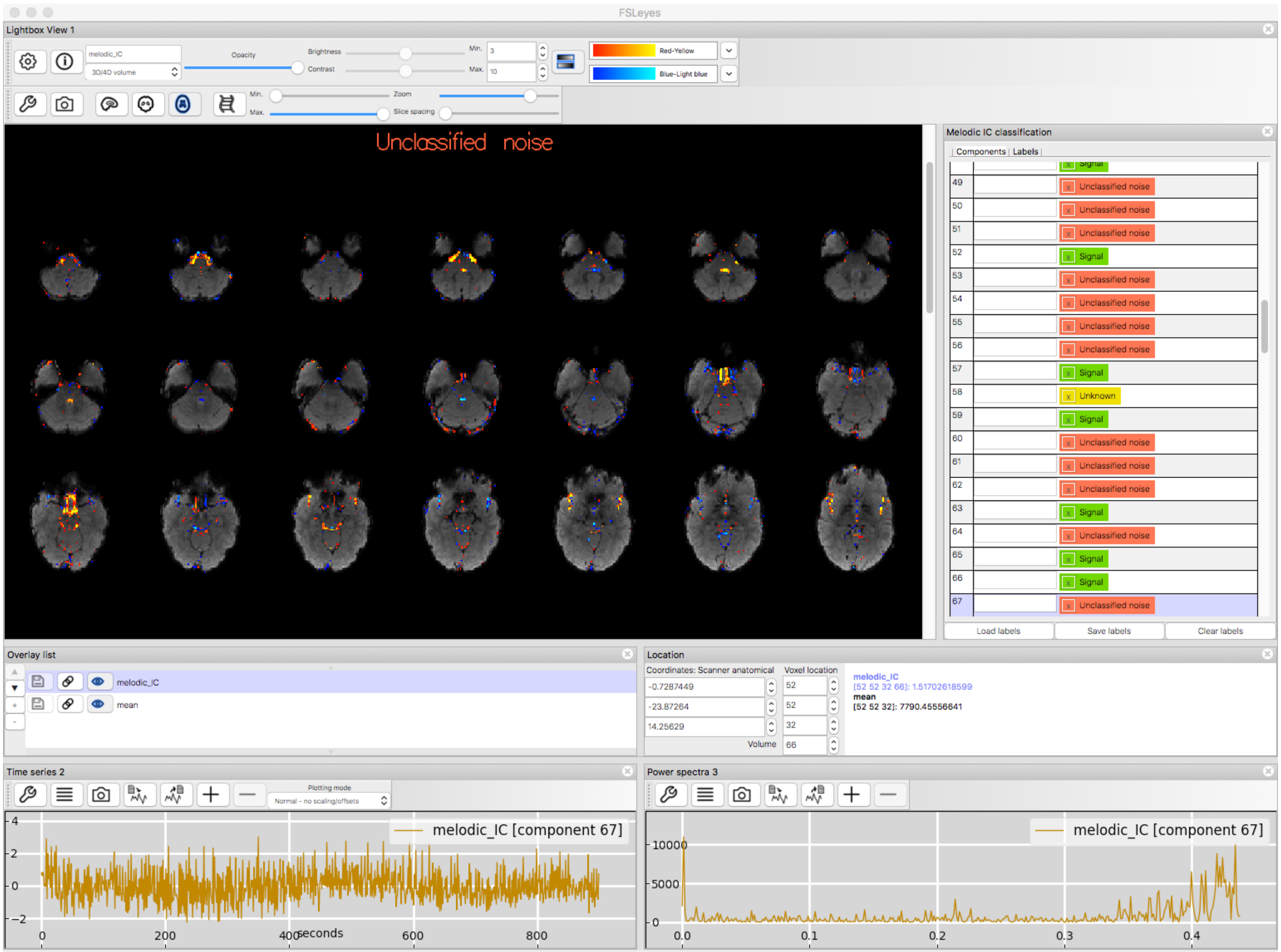
FSLeyes Melodic Mode



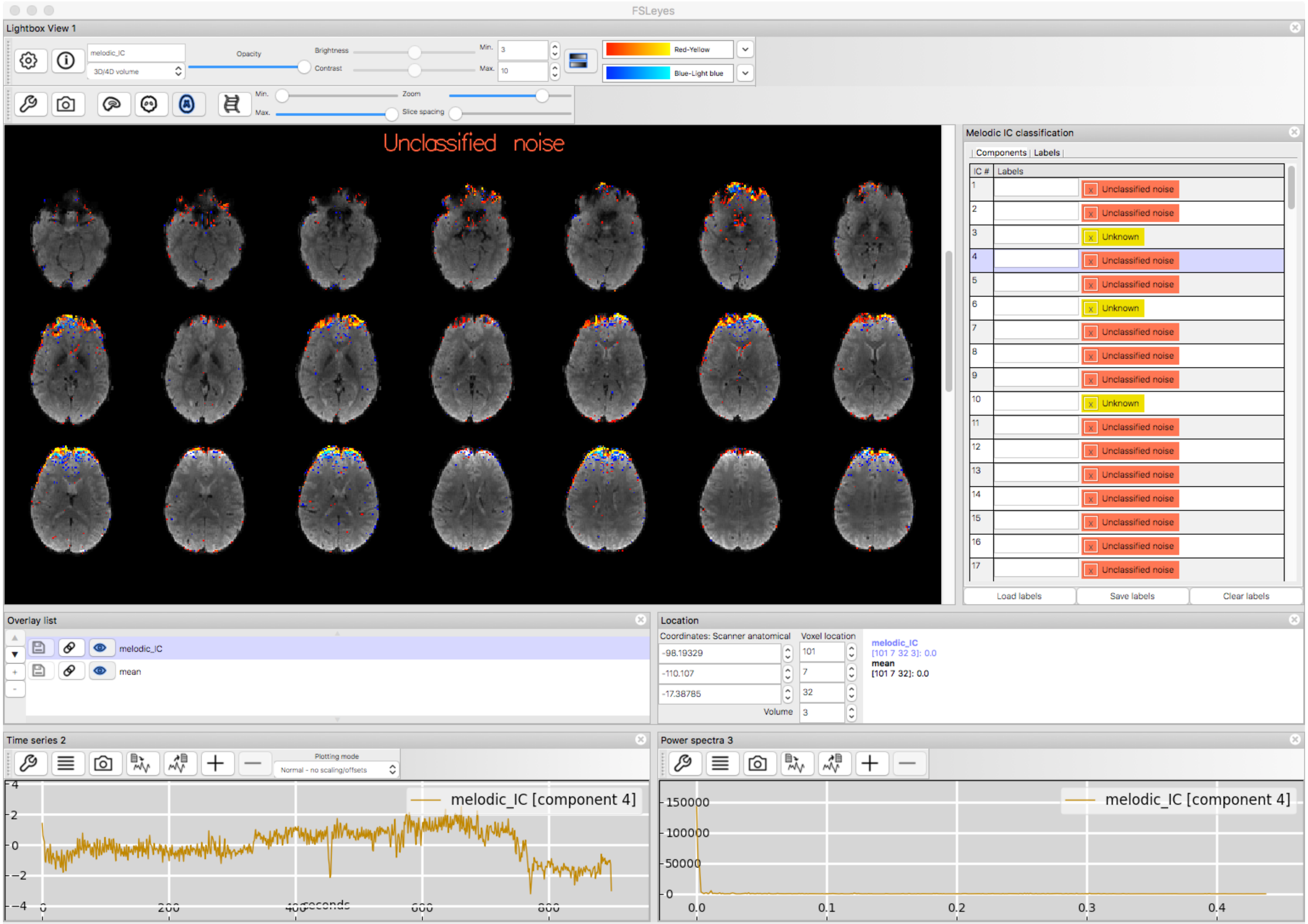
motion


FSLeyes
Lightbox View 1
Unclassified noise
Melodic IC classification
Components | Labels
IC # Labels
Load labels
Save labels
Clear labels
Overlay list
melodic_IC
mean
Location
Time series 2
melodic_IC [component 11]
seconds
Power spectra 3
melodic_IC [component 11]

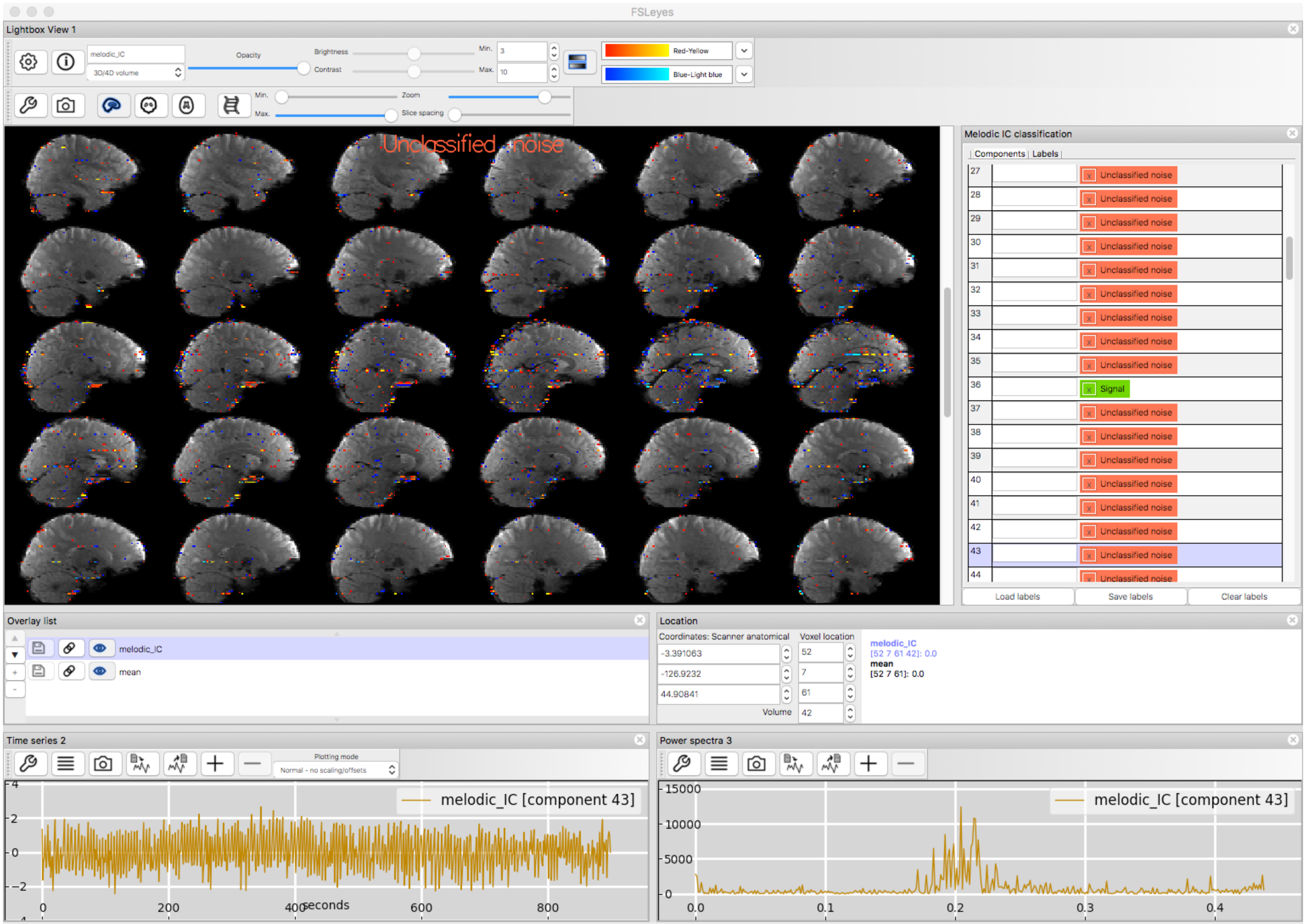
cardiac


FSLeyes
Lightbox View 1
Unclassified noise
Melodic IC classification
Components | Labels
49 Unclassified noise
50 Unclassified noise
51 Unclassified noise
52 Signal
53 Unclassified noise
54 Unclassified noise
55 Unclassified noise
56 Unclassified noise
57 Signal
58 Unknown
59 Signal
60 Unclassified noise
61 Unclassified noise
62 Unclassified noise
63 Signal
64 Unclassified noise
65 Signal
66 Signal
67 Unclassified noise
Load labels
Save labels
Clear labels
Overlay list
melodic_IC
mean
Location
melodic_IC
[52 52 32 66]: 1.51702618599
mean
[52 52 32]: 7790.45556641
Time series 2
melodic_IC [component 67]
seconds
Power spectra 3
melodic_IC [component 67]

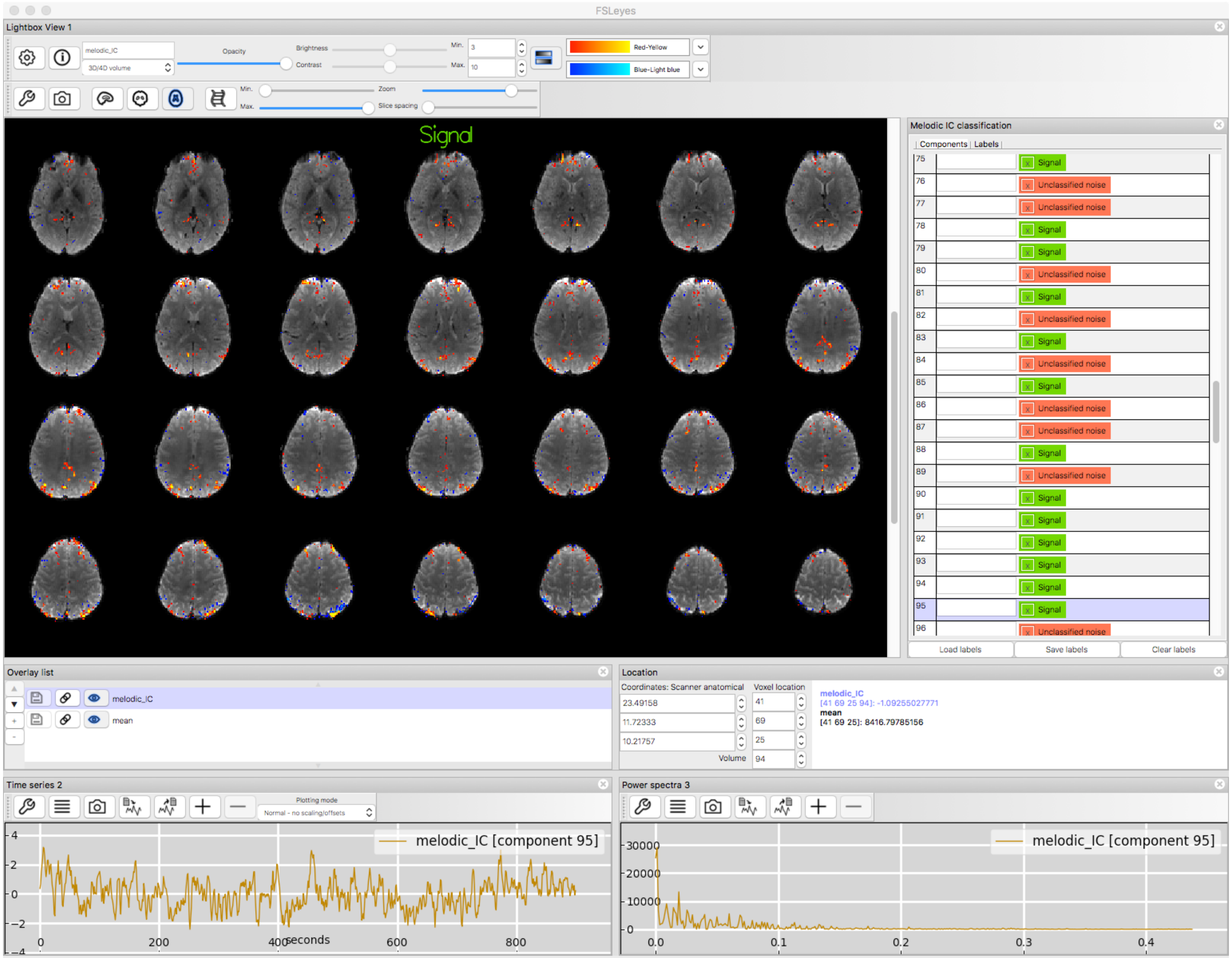
susceptibility motion



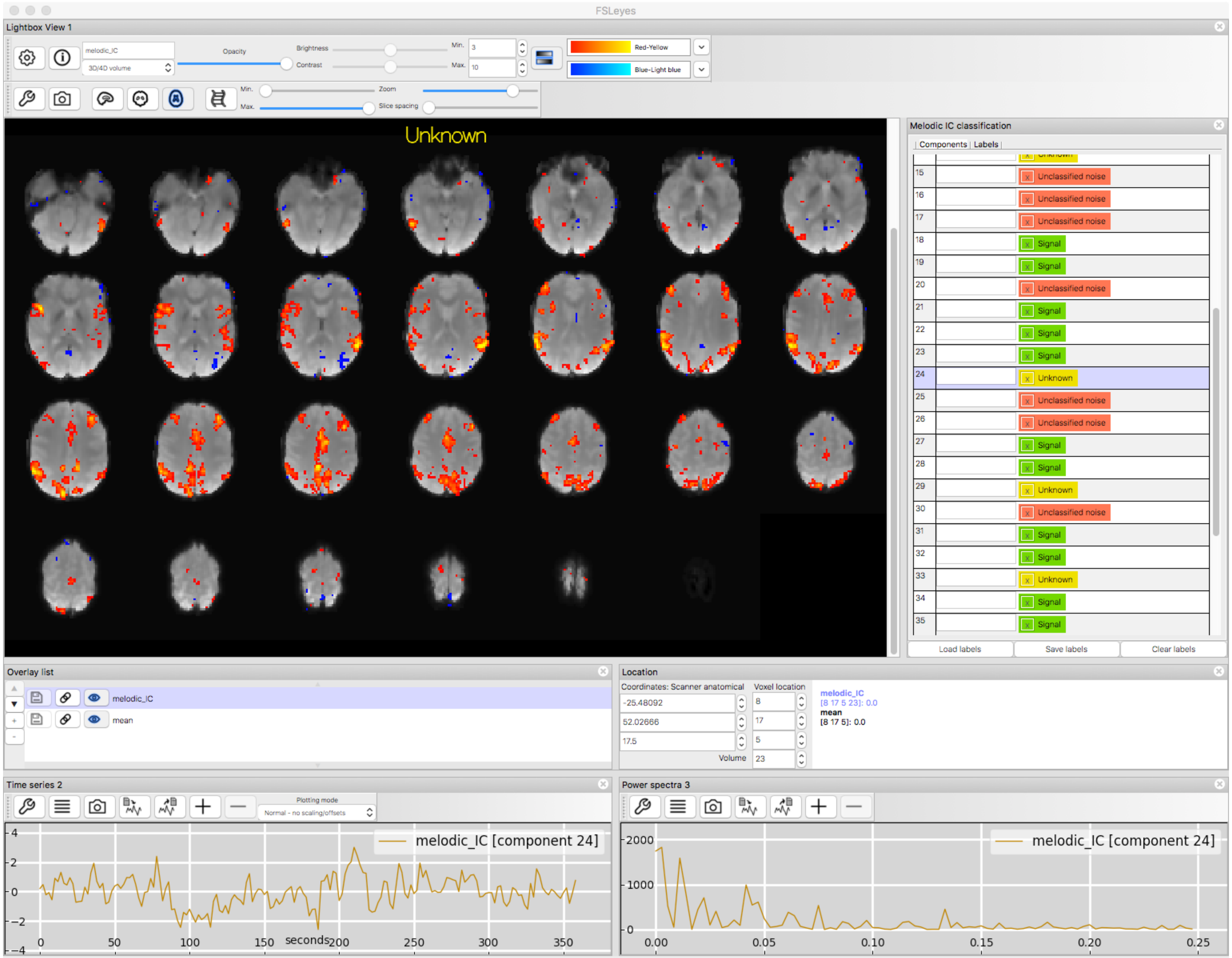
multiband


FSLeyes
Lightbox View 1
melodic_IC
3D/4D volume
Opacity
Brightness
Contrast
Min.
Max.
Red-Yellow
Blue-Light blue
Zoom
Slice spacing
Unclassified noise
Melodic IC classification
Components
Labels
27 Unclassified noise
28 Unclassified noise
29 Unclassified noise
30 Unclassified noise
31 Unclassified noise
32 Unclassified noise
33 Unclassified noise
34 Unclassified noise
35 Unclassified noise
36 Signal
37 Unclassified noise
38 Unclassified noise
39 Unclassified noise
40 Unclassified noise
41 Unclassified noise
42 Unclassified noise
43 Unclassified noise
44 Unclassified noise
Load labels
Save labels
Clear labels
Overlay list
melodic_IC
mean
Location
Coordinates: Scanner anatomical
Voxel location
-3.391063
-126.9232
44.90841
52
7
61
Volume
42
melodic_IC
[52 7 61 42]: 0.0
mean
[52 7 61]: 0.0
Time series 2
Plotting mode
Normal - no scaling/offsets
melodic_IC [component 43]
seconds
Power spectra 3
melodic_IC [component 43]

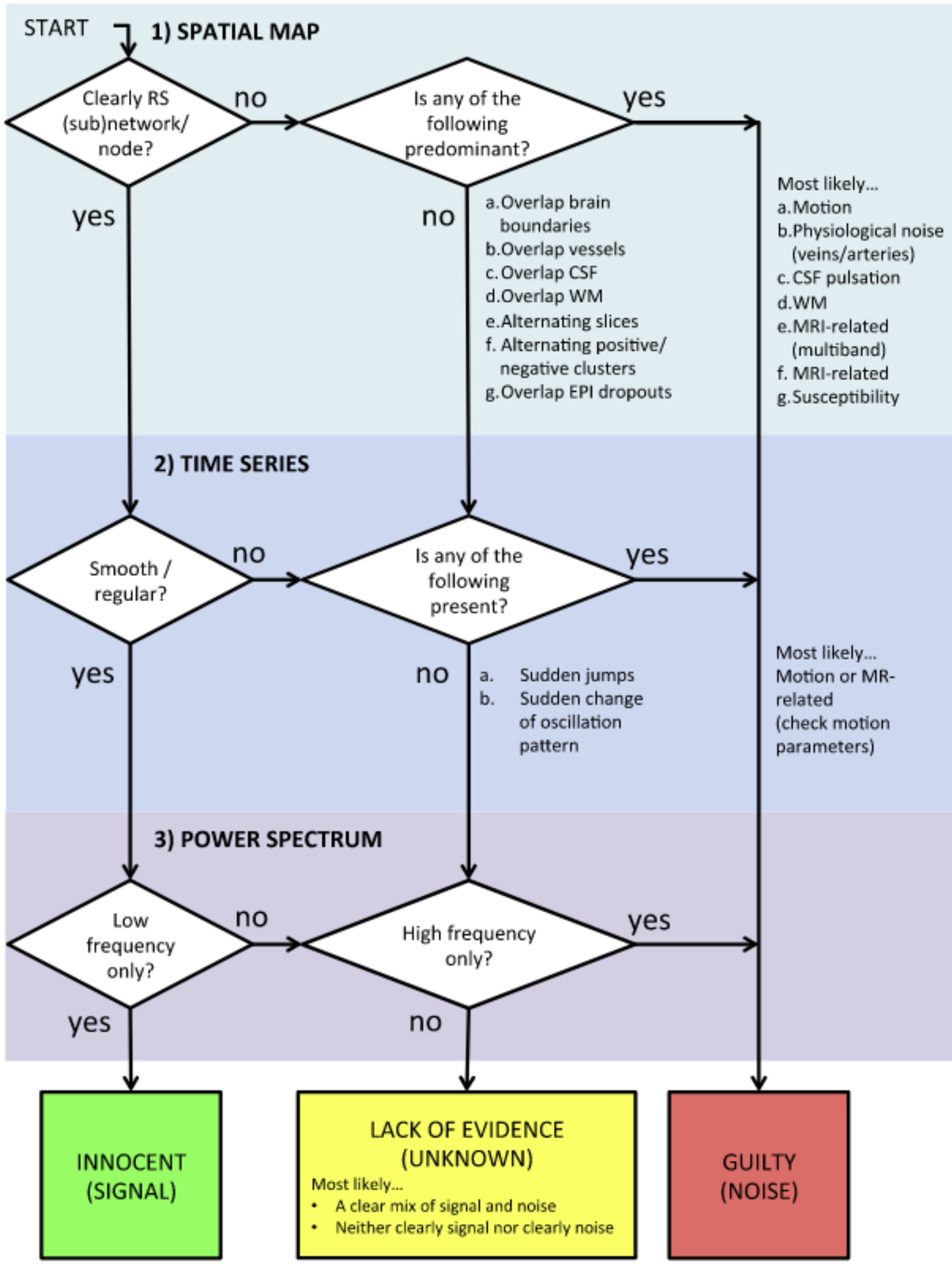
signal


FSLeyes
Lightbox View 1
Signal
Melodic IC classification
Components
Labels
Load labels
Save labels
Clear labels
Overlay list
melodic_IC
mean
Location
Time series 2
Power spectra 3
melodic_IC [component 95]
seconds

effects of scan parameters



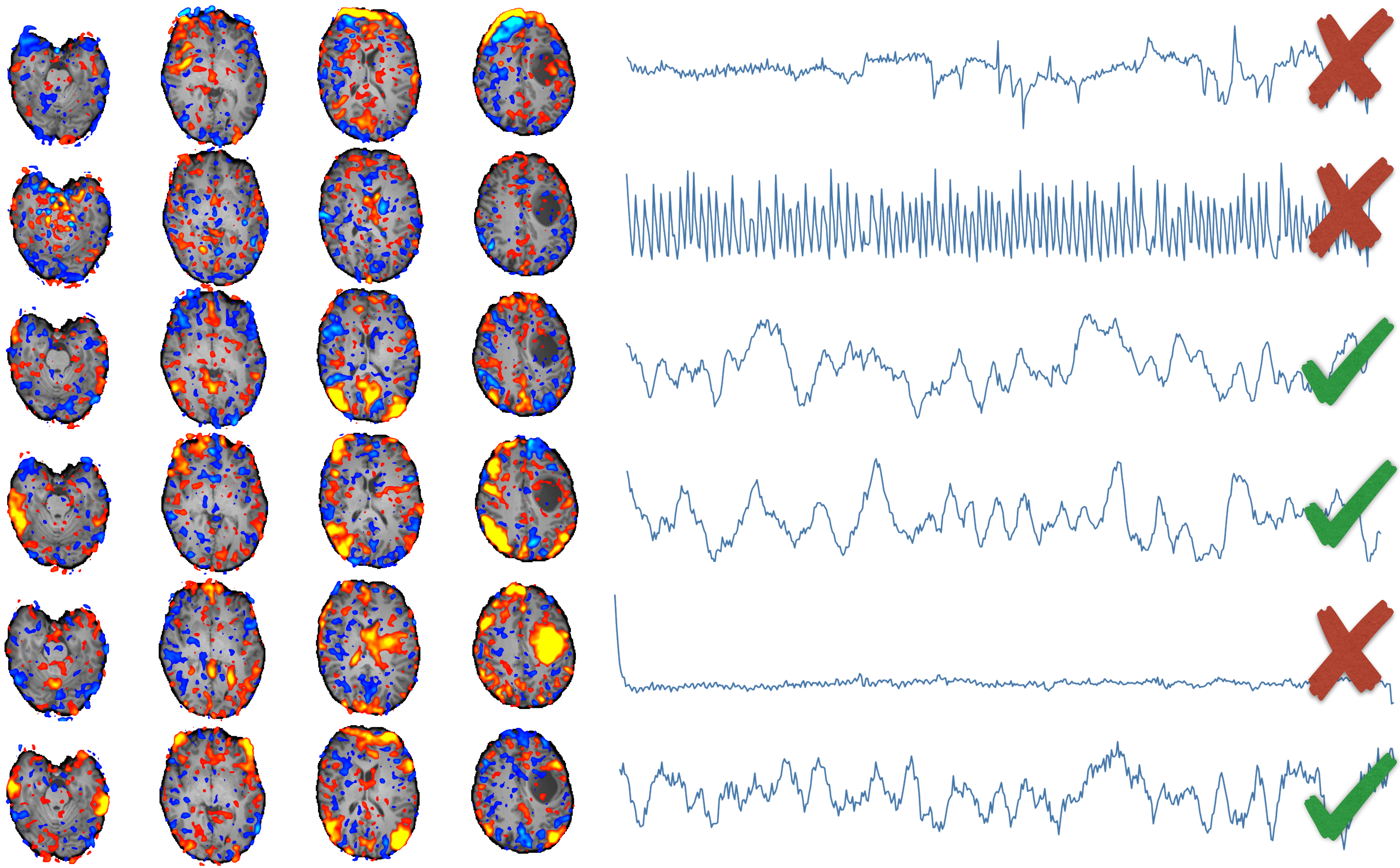
manual classification



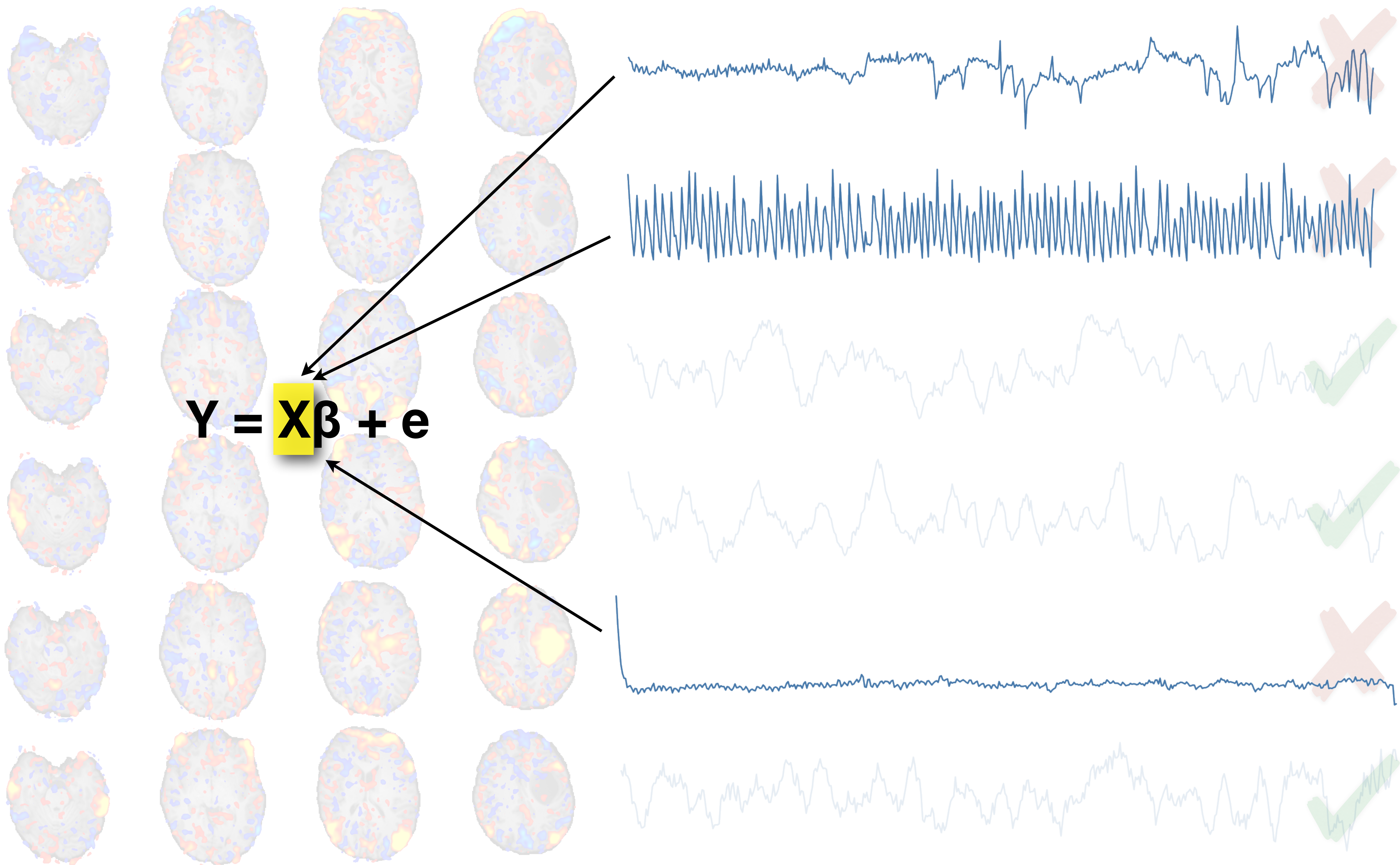
Griffanti et al (2016).

https://doi.org/10.1016/j.neuroimage.2016.12.036

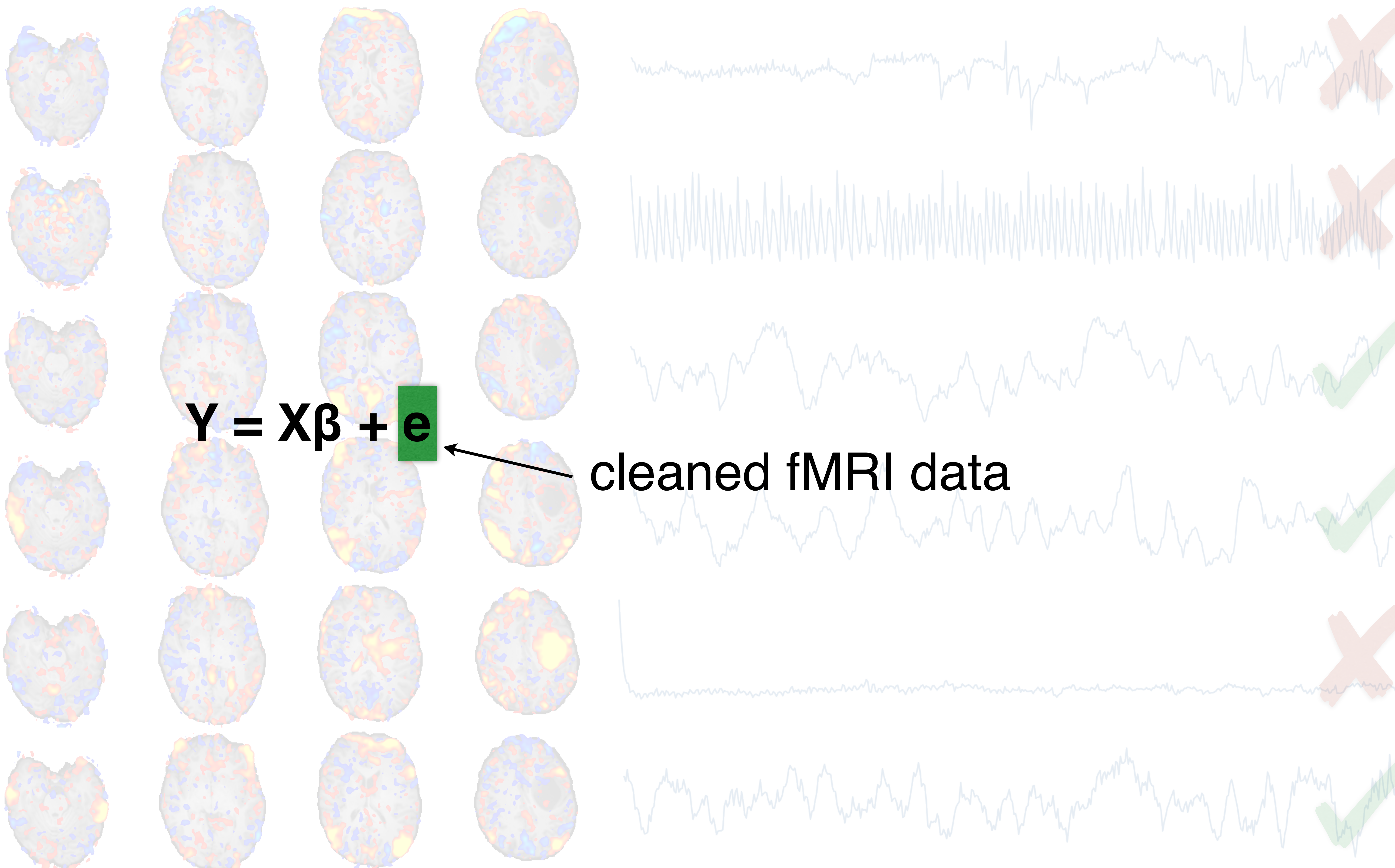
ICA-based denoising



ICA-based denoising



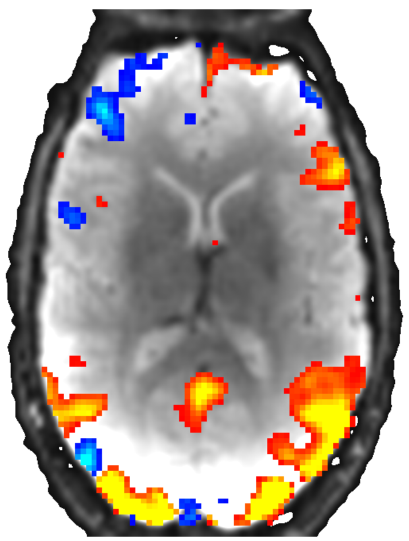
ICA-based denoising



$$Y = X\beta + \boxed{e}$$

cleaned fMRI data

semi-automatic classification



component

classifier

signal

-or-

noise

label

semi-automatic classification

- FIX (fsl.fmrib.ox.ac.uk/fsl/fslwiki/FIX)
  - Classifier with many features
  - Requires manually labelled training data
  - 99% accuracy on high-quality data

semi-automatic classification

- FIX (fsl.fmrib.ox.ac.uk/fsl/fslwiki/FIX)
  - Classifier with many features
  - Requires manually labelled training data
  - 99% accuracy on high-quality data
- ICA-AROMA (github.com/rhr-pruim/ICA-AROMA)
  - Simple classifier with only 4 features
  - No training data required
  - Mainly designed for motion artefacts

Resting state fMRI and ICA

- Introduction to resting state
- Independent Component Analysis
- Single-subject ICA
- Multi-subject ICA
- Dual regression

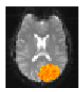
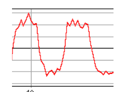
Resting state fMRI and ICA

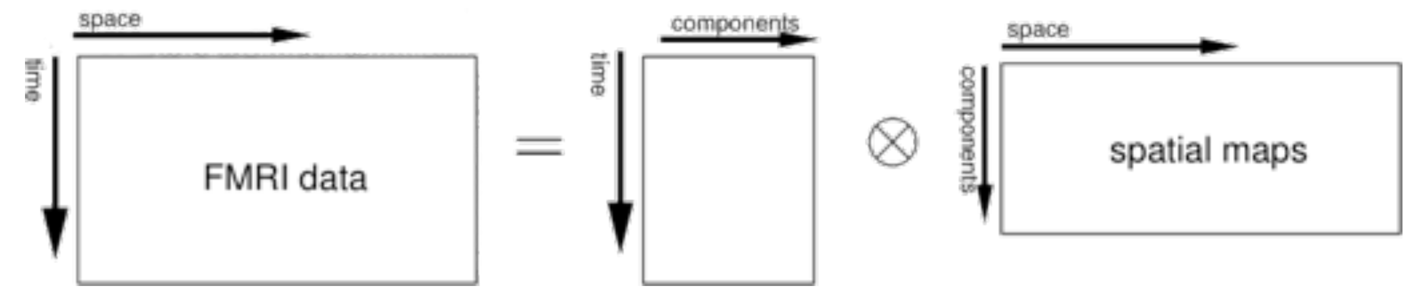
- Introduction to resting state
- Independent Component Analysis
- Single-subject ICA
- Multi-subject ICA
- Dual regression

Different ICA models

Single-Session ICA

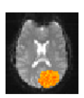
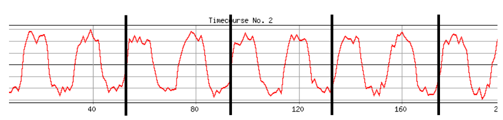
each ICA component comprises:

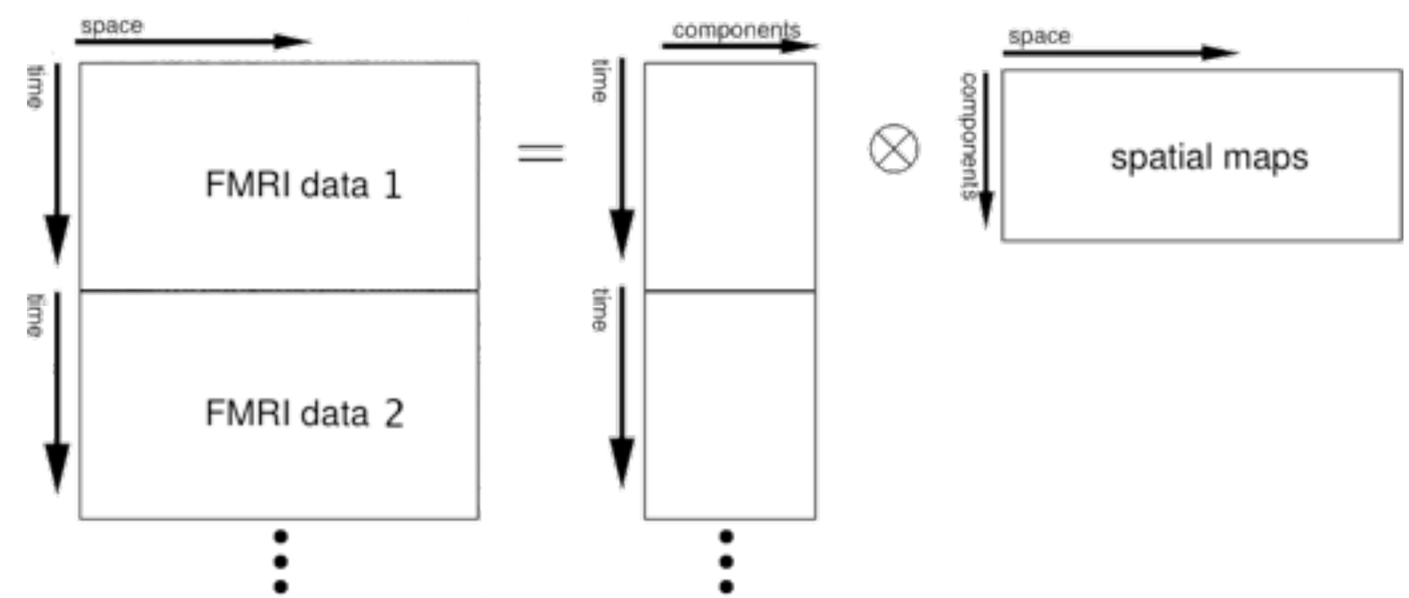
 spatial map & timecourse 



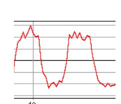
Multi-Session or Multi-Subject ICA: **Concatenation approach**

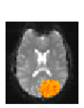
each ICA component comprises:

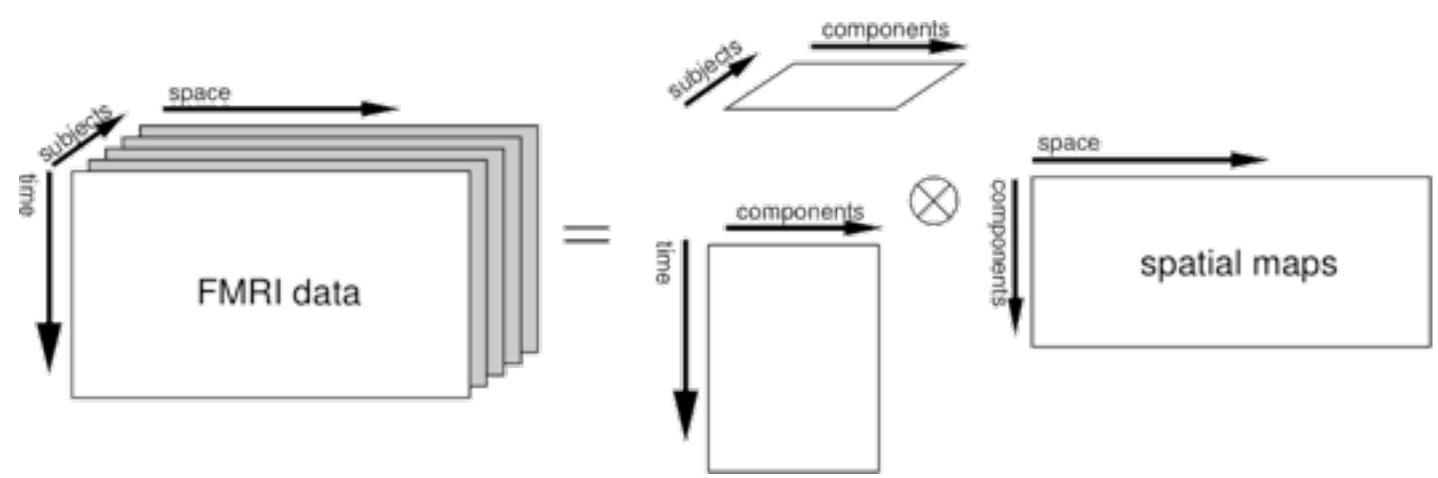
 spatial map & timecourse
(that can be split up into subject-specific chunks) 



Multi-Session or Multi-Subject ICA: **Tensor-ICA approach**

each ICA component comprises: 

 spatial map, session-long-timecourse
& subject-strength plot 

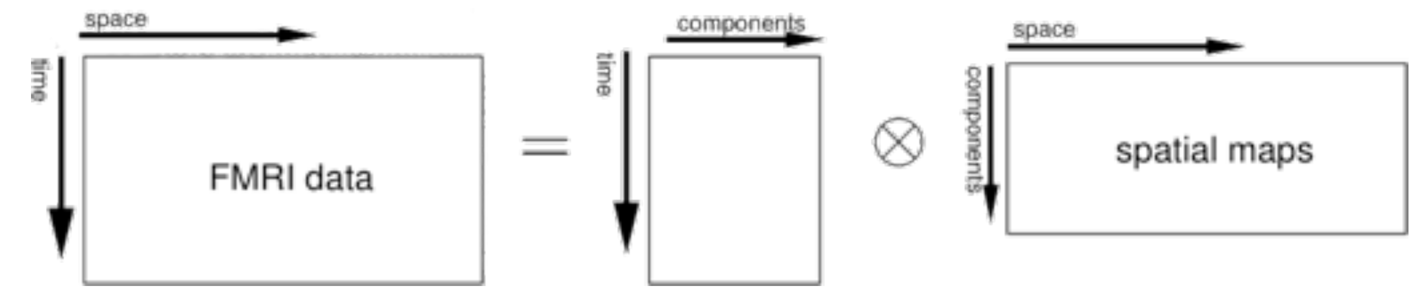


Different ICA models

Single-Session ICA

each ICA component comprises:

spatial map & timecourse



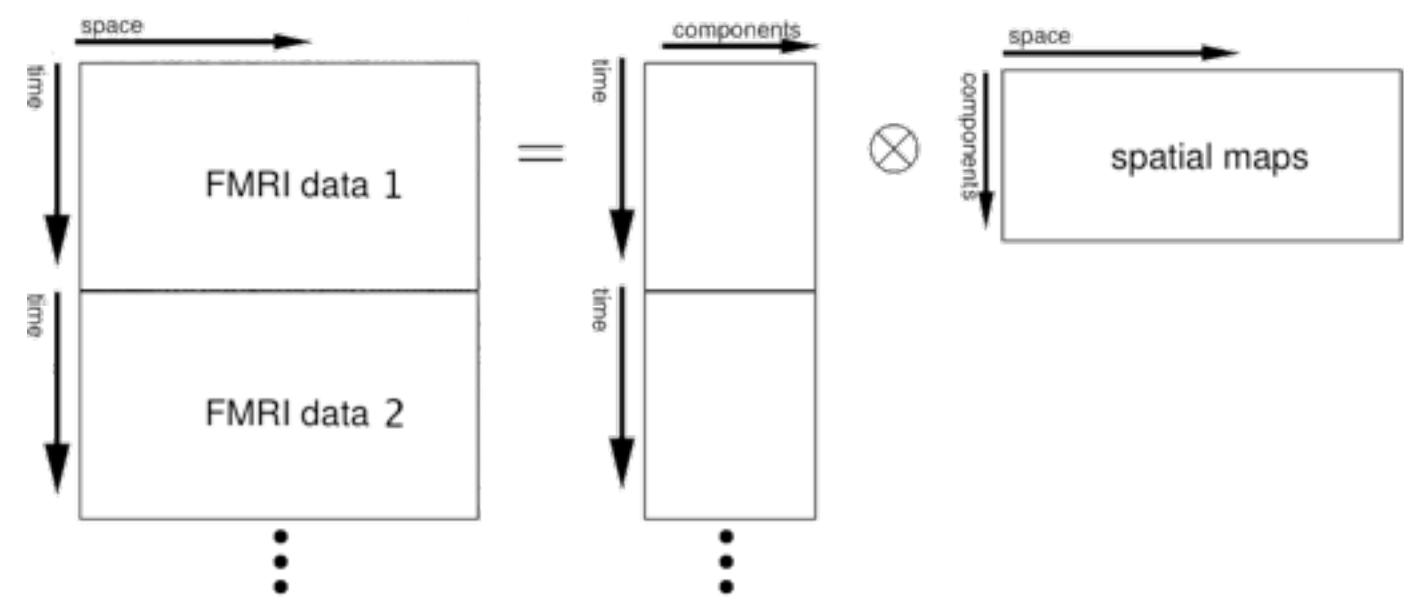
Multi-Session or Multi-Subject ICA:

Concatenation approach

good when:

each subject has DIFFERENT timeseries

e.g. resting-state FMRI



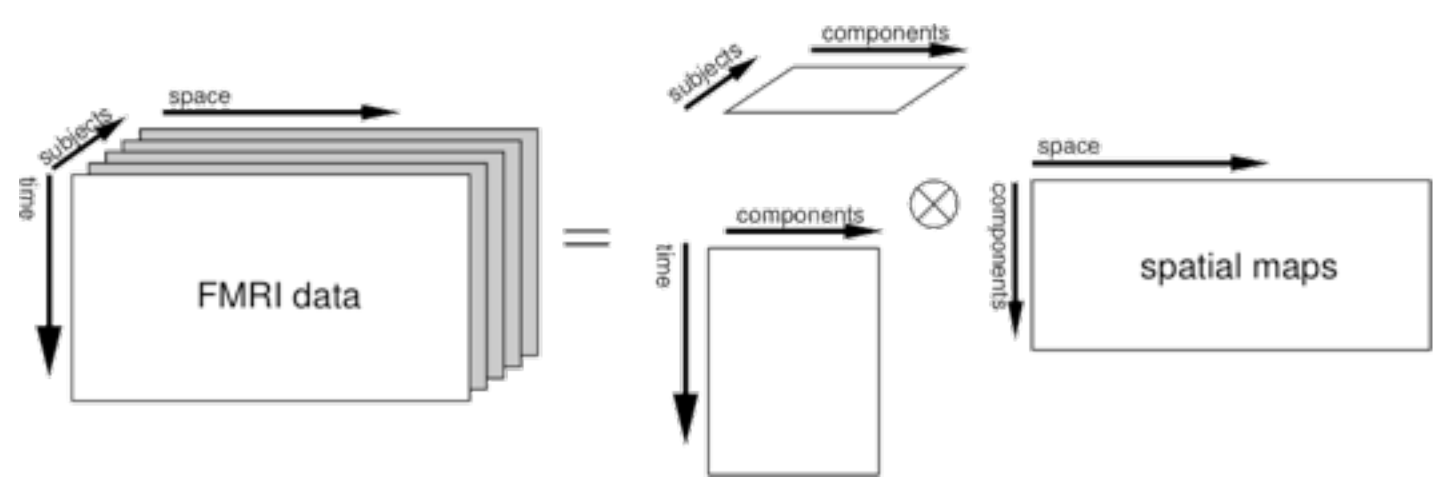
Multi-Session or Multi-Subject ICA:

Tensor-ICA approach

good when:

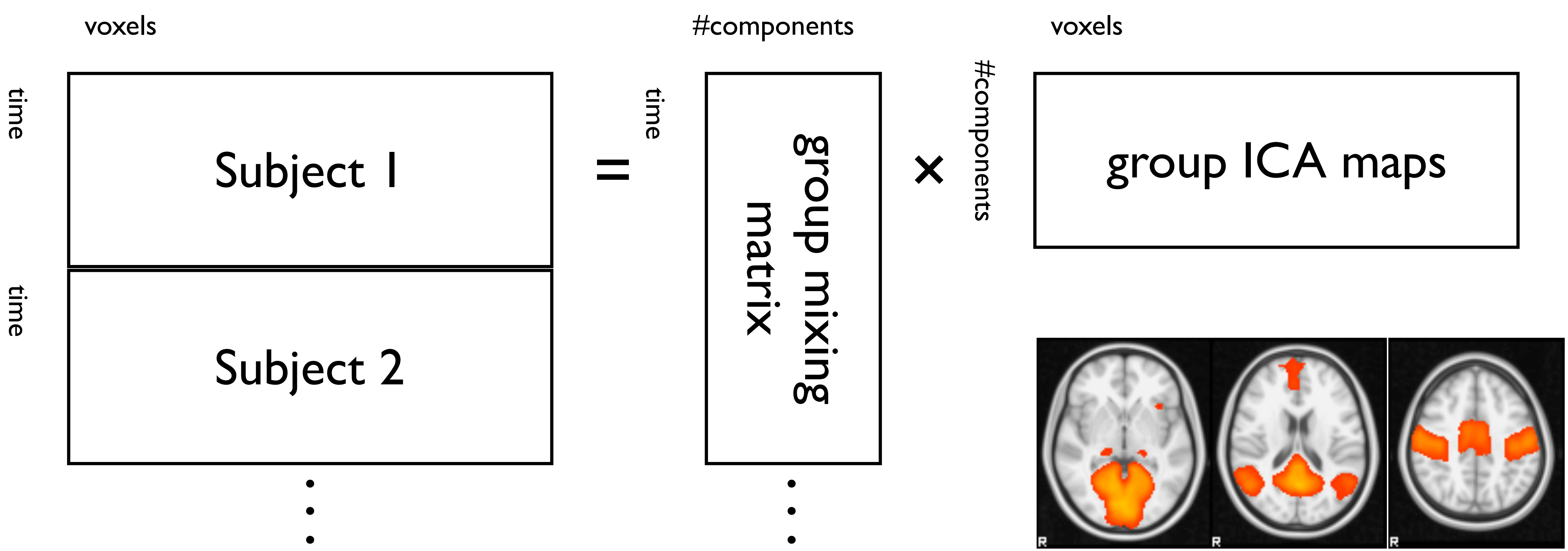
each subject has SAME timeseries

e.g. activation FMRI



Concatenated ICA

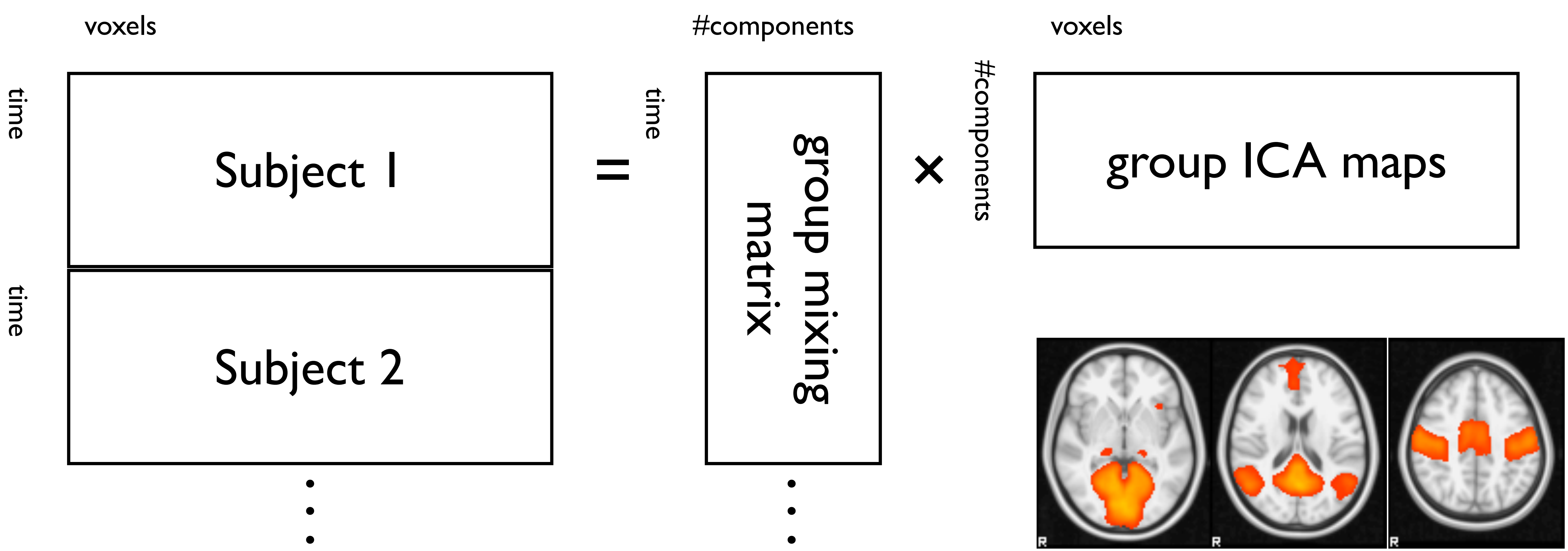
- Concatenate all subjects’ data temporally
- Then run ICA
- More appropriate than tensor ICA (for RSNs)



Concatenated ICA

- Data sets must be registered to a common space (anatomical alignment)

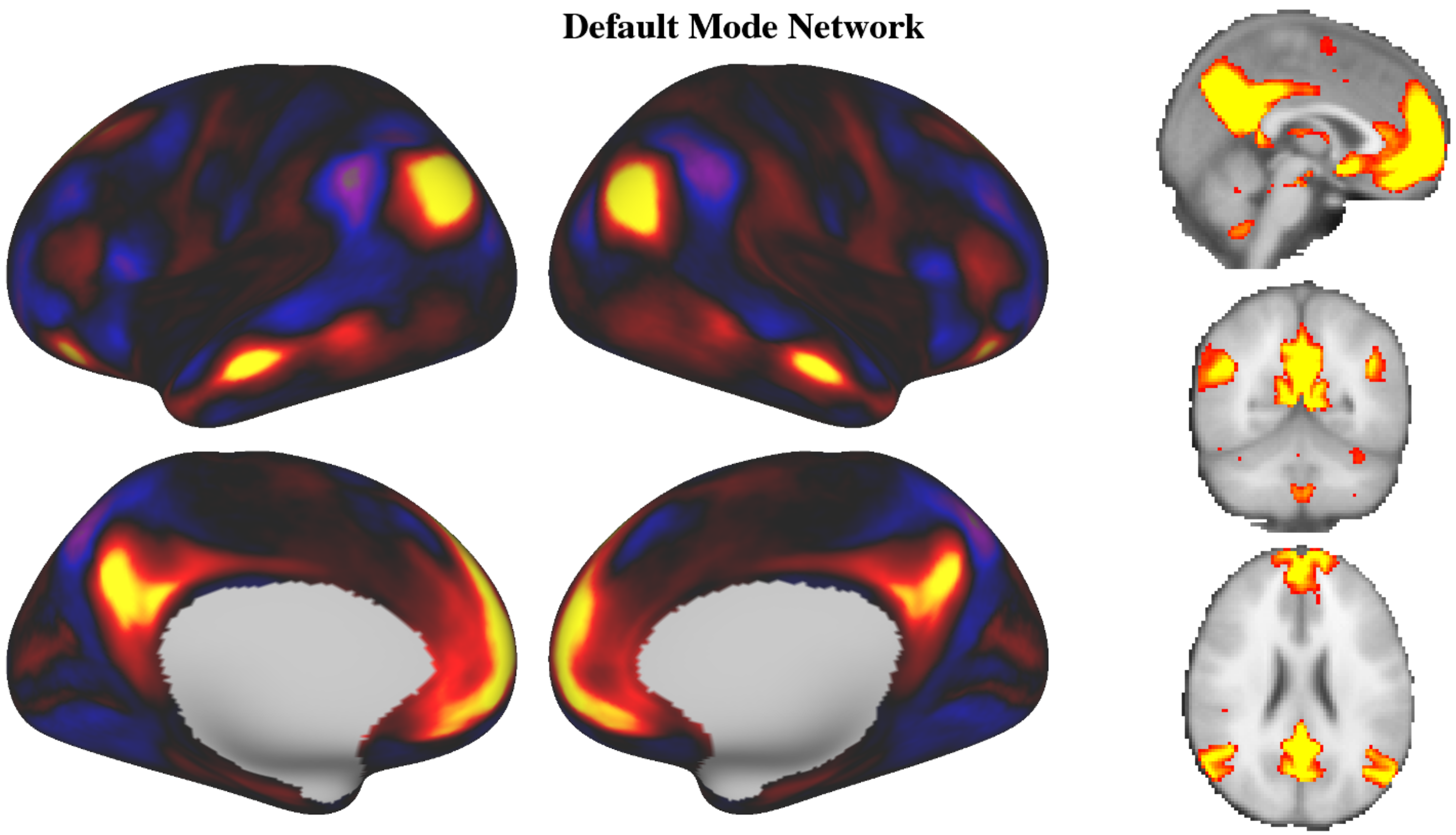
- Memory optimisation trick (called MIGP) means that time courses are not interpretable

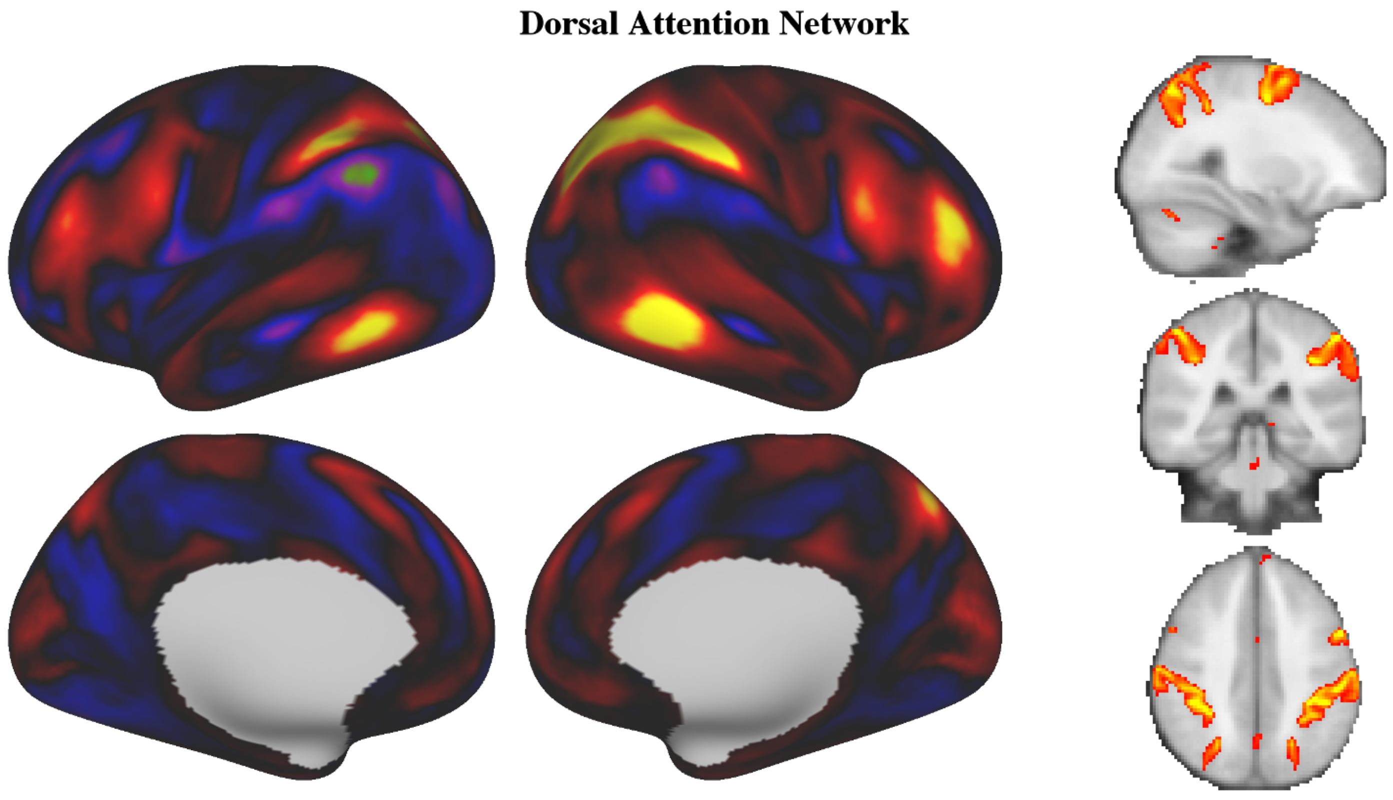


Resting state multi-subject ICA

- Why not just run ICA on each subject separately?
  - Correspondence problem (eg RSNs across subjects)
  - Different splittings sometimes caused by small changes in the data (naughty ICA!)
- Instead - start with a "group-average" ICA
  - But then need to relate group maps back to the individual subjects

Resting state networks

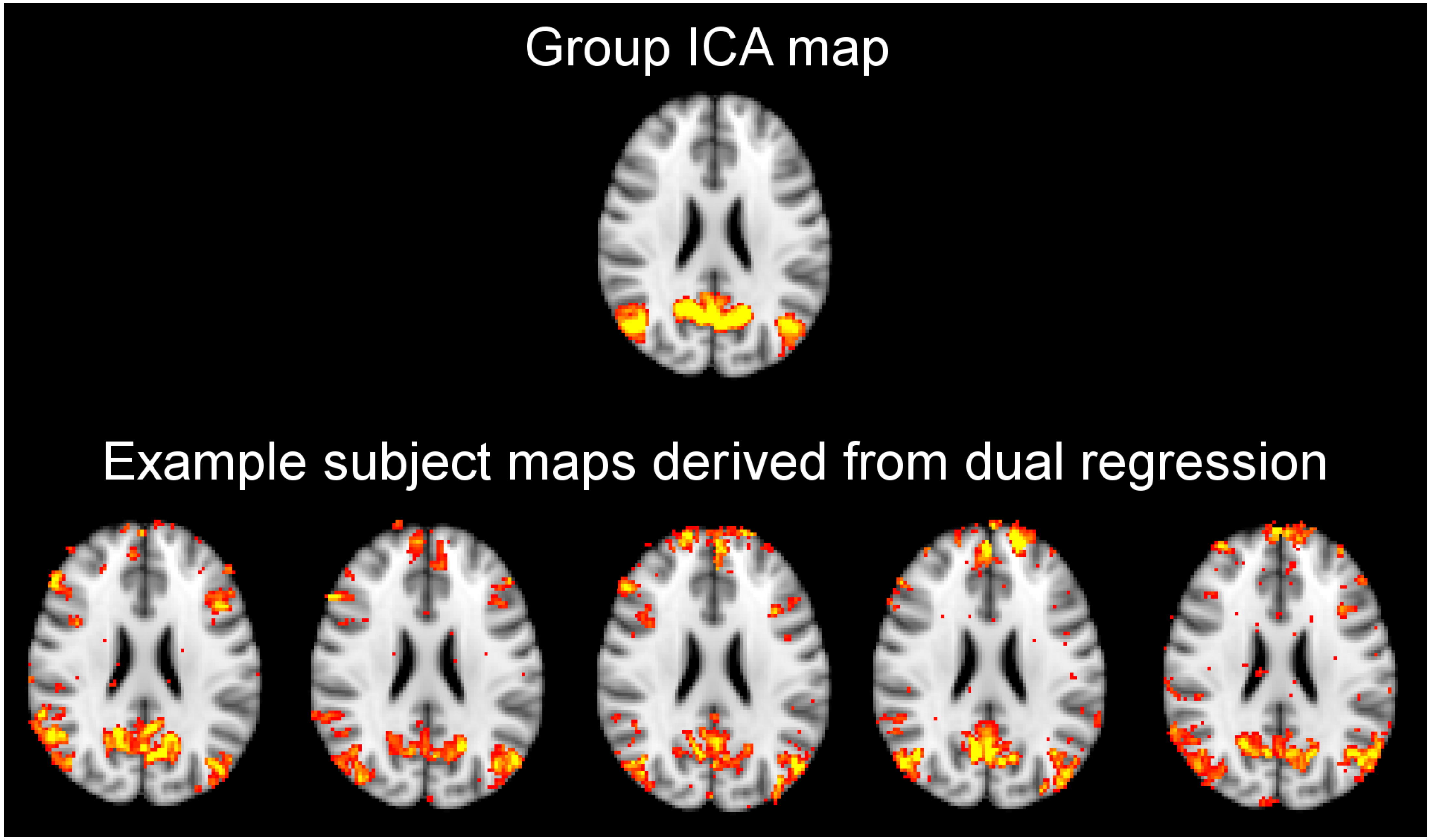




Resting state fMRI and ICA

- Introduction to resting state
- Independent Component Analysis
- Single-subject ICA
- Multi-subject ICA
- Dual regression

Resting state multi-subject ICA

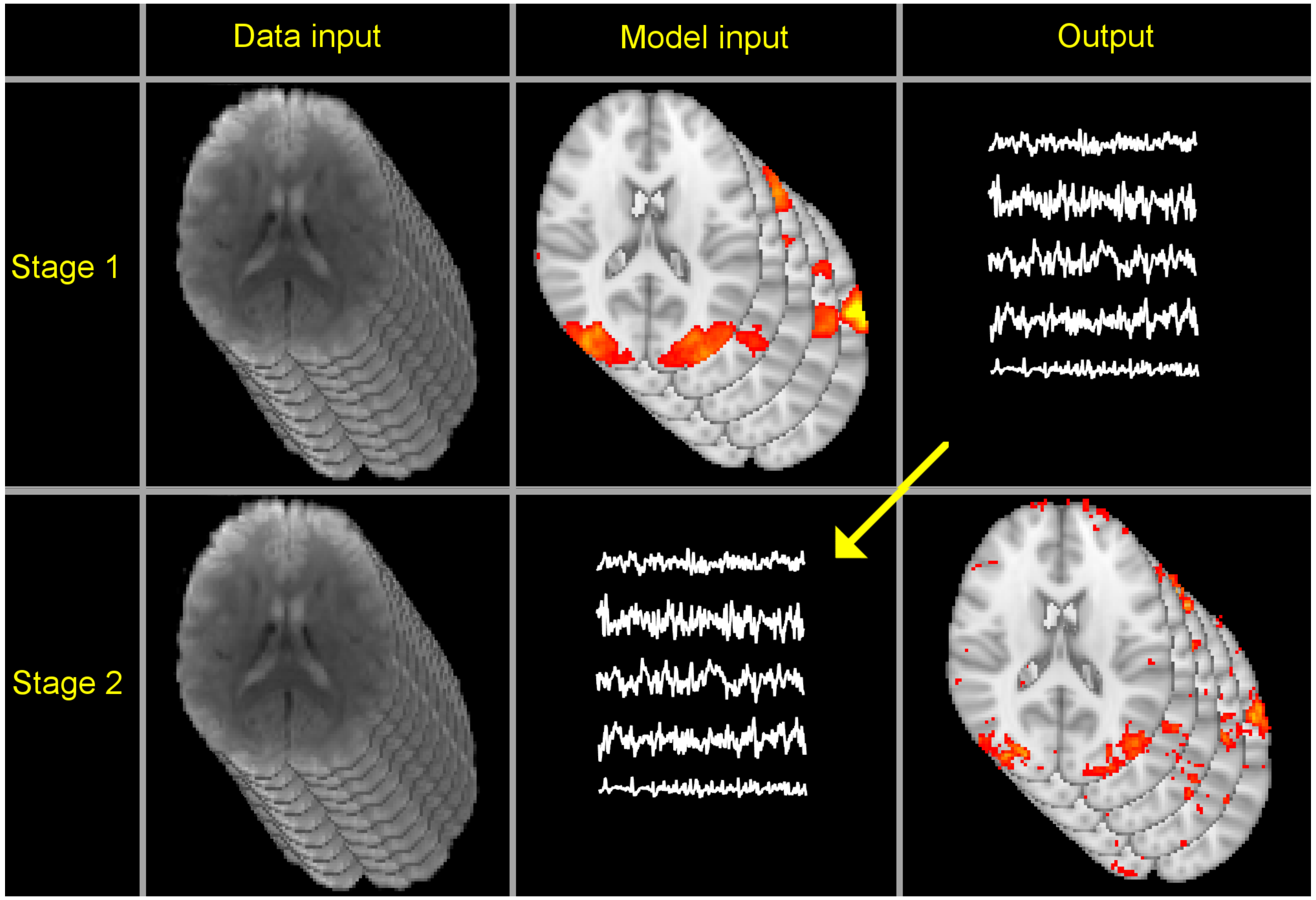


Dual Regression

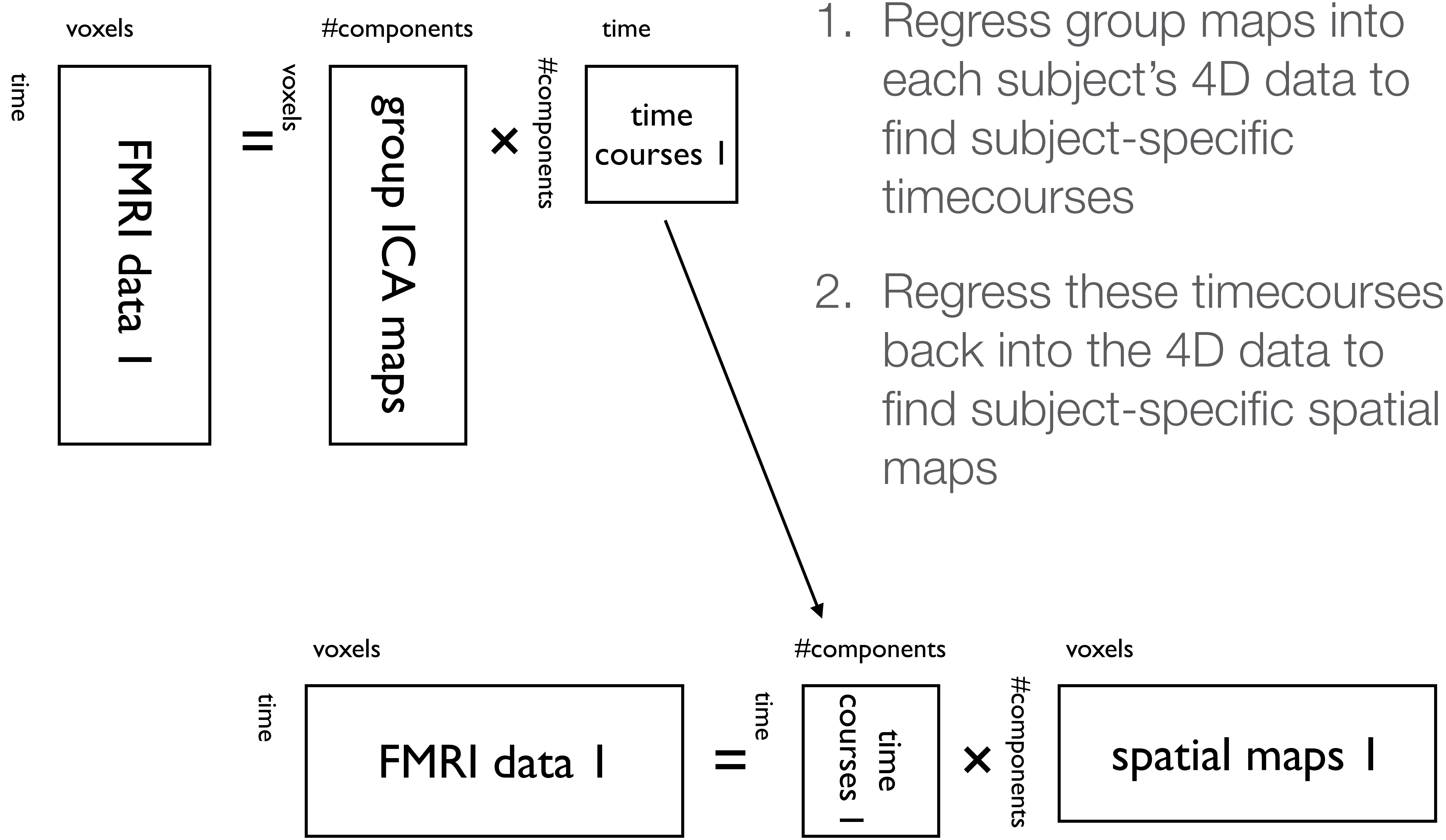
Two steps that both involve multiple regression:

1. Extract subject timeseries

2. Extract subject maps



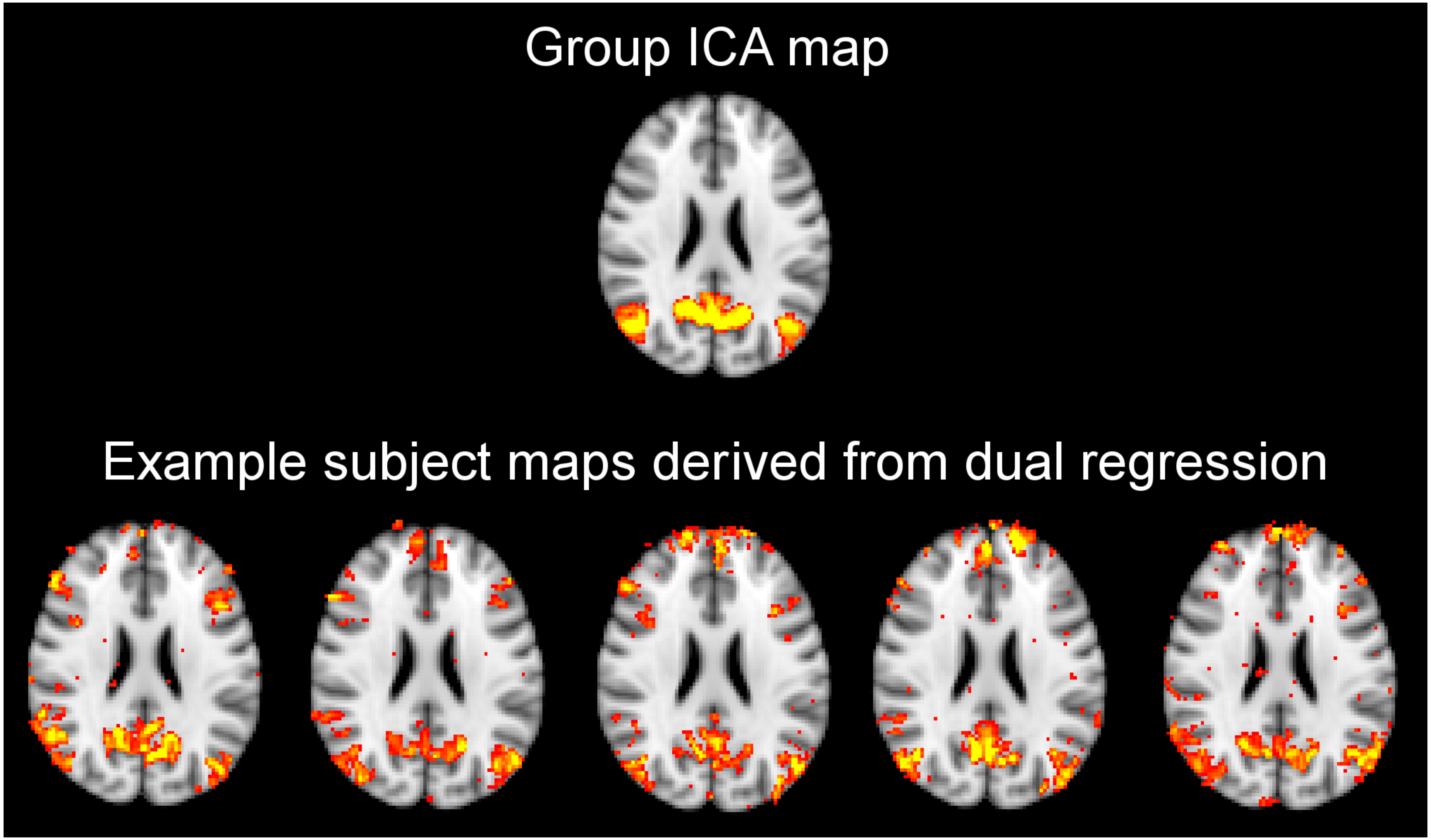
Dual Regression



1. Regress group maps into each subject's 4D data to find subject-specific timecourses

2. Regress these timecourses back into the 4D data to find subject-specific spatial maps

Dual Regression



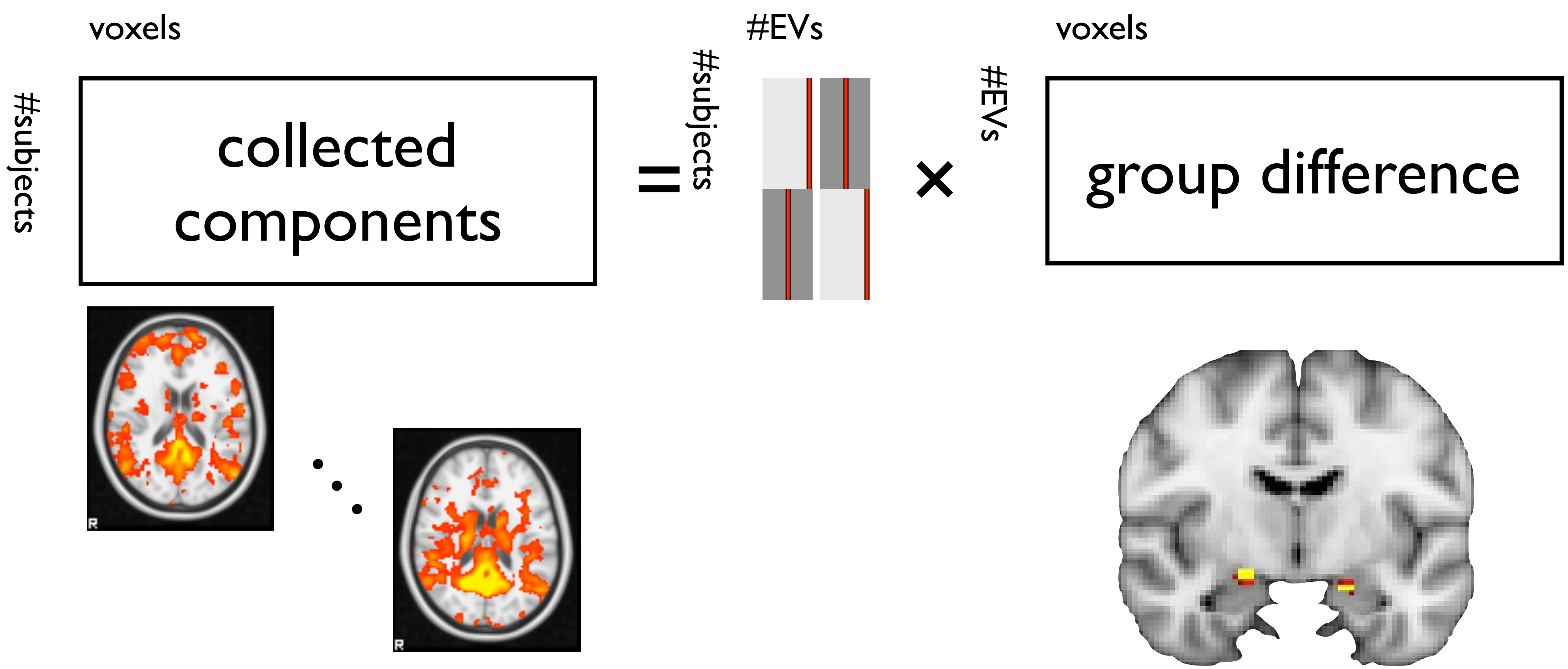
Running dual_regression

```
beckmann — bash — bash — 142×23
[islay:~] dual_regression.sh

dual_regression v0.5 (beta)

***NOTE*** ORDER OF COMMAND-LINE ARGUMENTS IS DIFFERENT FROM PREVIOUS VERSION

Usage: dual_regression <group_IC_maps> <des_norm> <design.mat> <design.con> <n_perm> <output_directory> <input1> <input2> <input3> .........
e.g.   dual_regression groupICA.gica/groupmelodic.ica/melodic_IC 1 design.mat design.con 500 grot `cat groupICA.gica/.filelist`

<group_IC_maps_4D>            4D image containing spatial IC maps (melodic_IC) from the whole-group ICA analysis
<des_norm>                    0 or 1 (1 is recommended). Whether to variance-normalise the timecourses used as the stage-2 regressors
<design.mat>                  Design matrix for final cross-subject modelling with randomise
<design.con>                  Design contrasts for final cross-subject modelling with randomise
<n_perm>                      Number of permutations for randomise; set to 1 for just raw tstat output, set to 0 to not run randomise at all.
<output_directory>            This directory will be created to hold all output and logfiles
<input1> <input2> ...         List all subjects' preprocessed, standard-space 4D datasets

<design.mat> <design.con>     can be replaced with just
-1                            for group-mean (one-group t-test) modelling.
If you need to add other randomise option then just edit the line after "EDIT HERE" below

[islay:~]
```

- FSL command line tool, combining:
  - DR to create subject-wise estimates (stage 1 + stage 2)
  - Group comparison using randomise (stage 3)

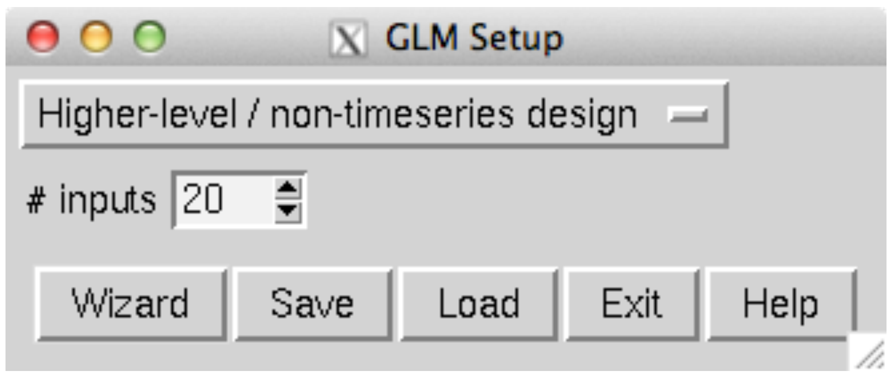
Group comparison

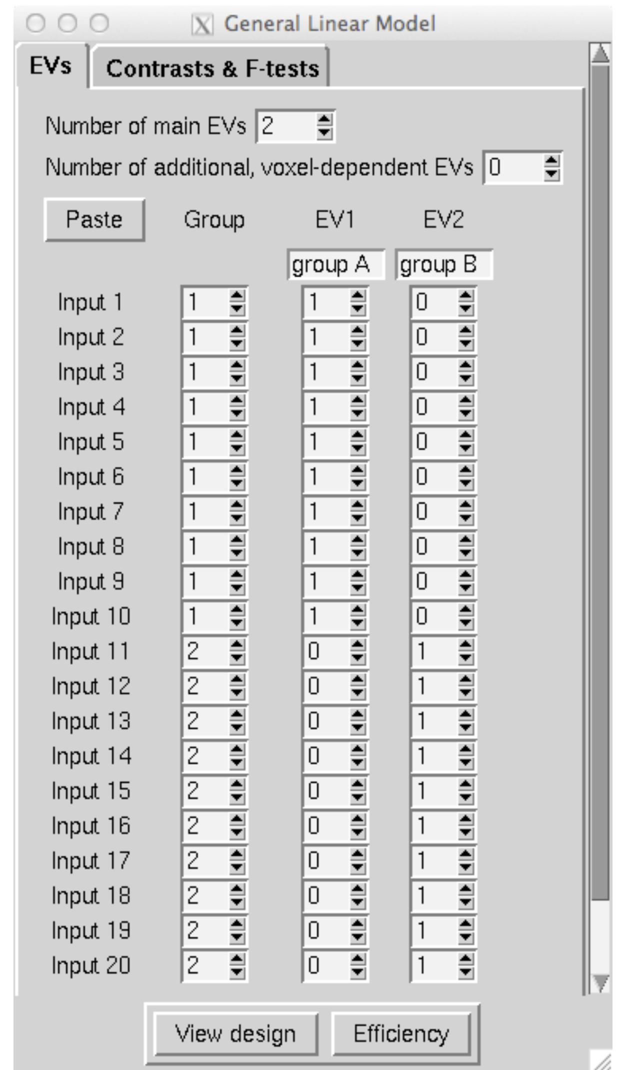
- Collect maps and perform voxel-wise test (e.g. randomisation test on GLM)

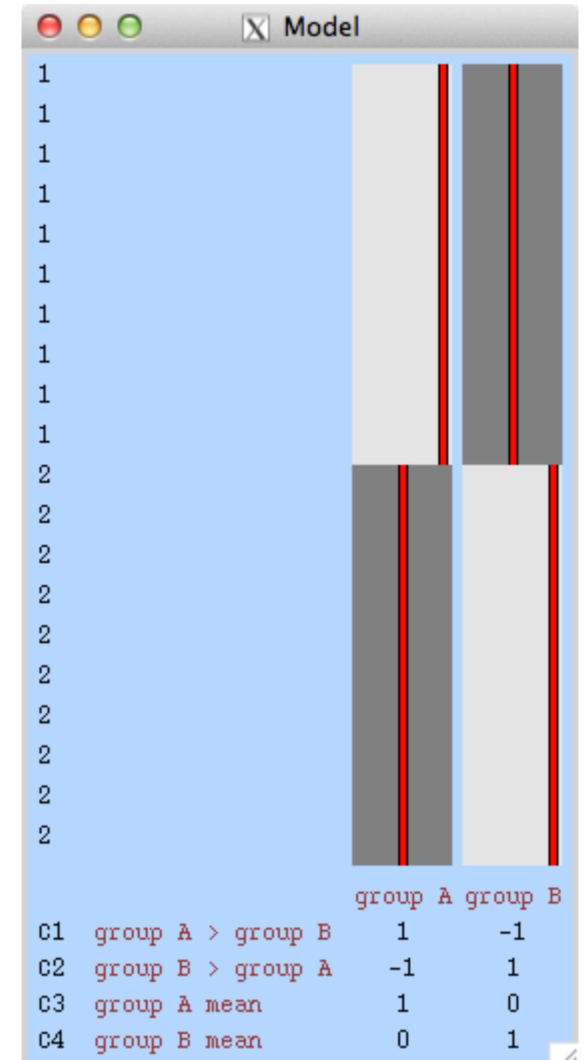


- Can now do voxelwise testing across subjects, separately for each original group ICA map
- Can choose to look at strength-and-shape differences

Group analysis on maps



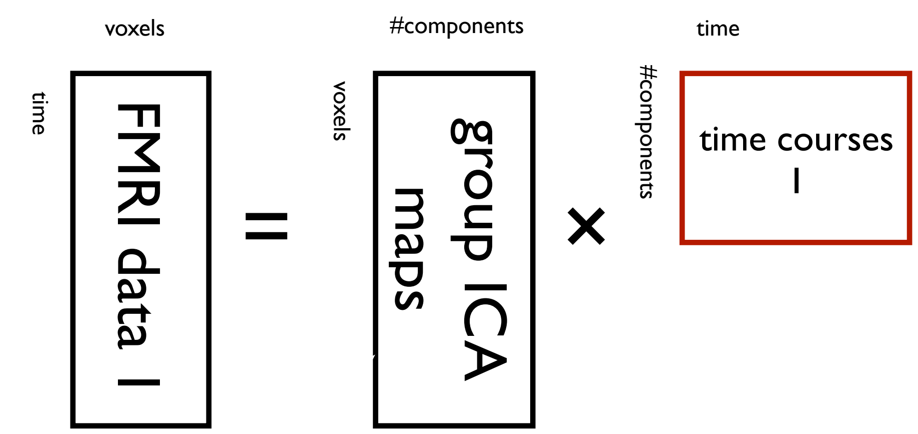




- can use the Glm tool (Glm_gui on mac) to create GLM design and contrast matrices

Dual regression outputs

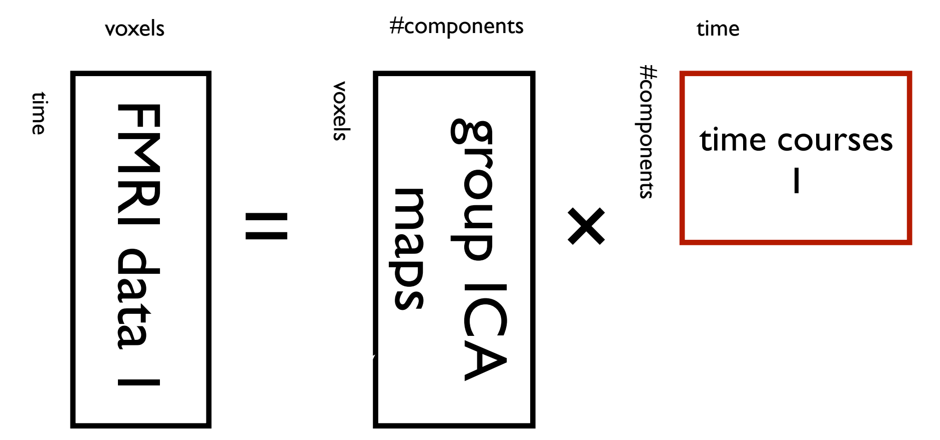
- dr_stage1_subject[#SUB].txt - the timeseries outputs of stage 1 of the dual-regression.

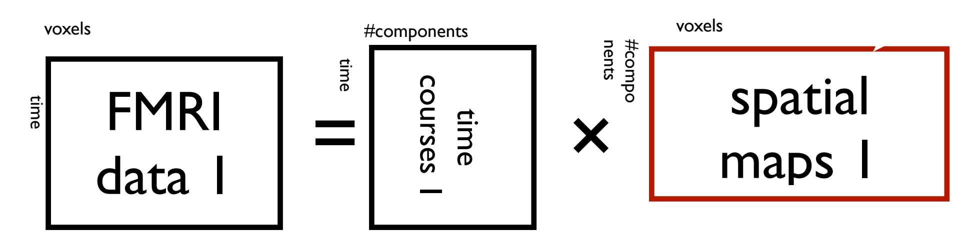


Dual regression outputs

- dr_stage1_subject[#SUB].txt - the timeseries outputs of stage 1 of the dual-regression.

- dr_stage2_subject[#SUB].nii.gz - the spatial maps outputs of stage 2 of the dual-regression.



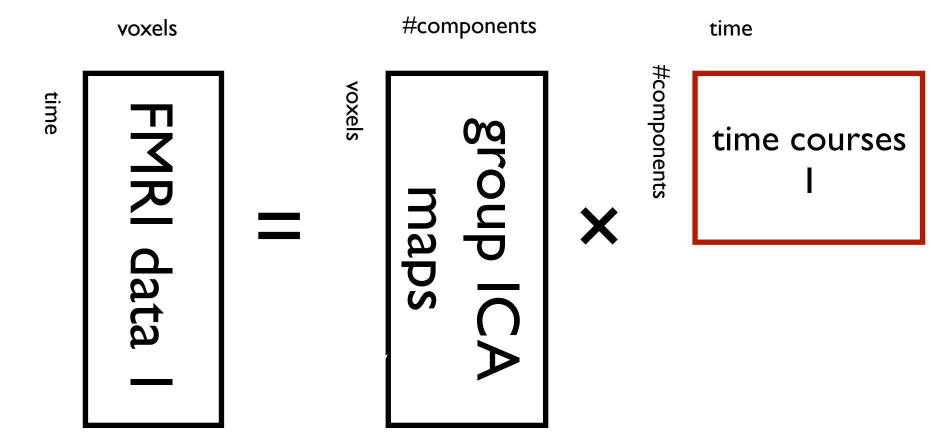


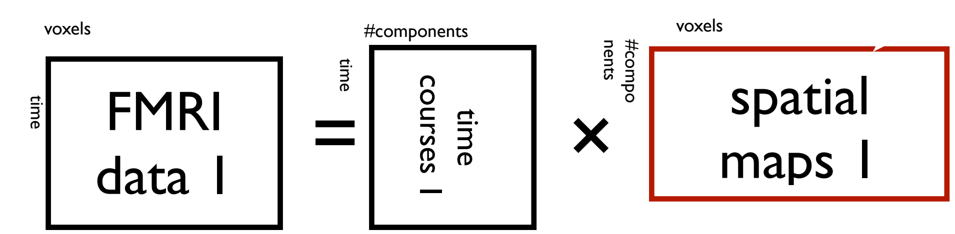
Dual regression outputs

- dr_stage1_subject[#SUB].txt - the timeseries outputs of stage 1 of the dual-regression.

- dr_stage2_subject[#SUB].nii.gz - the spatial maps outputs of stage 2 of the dual-regression.

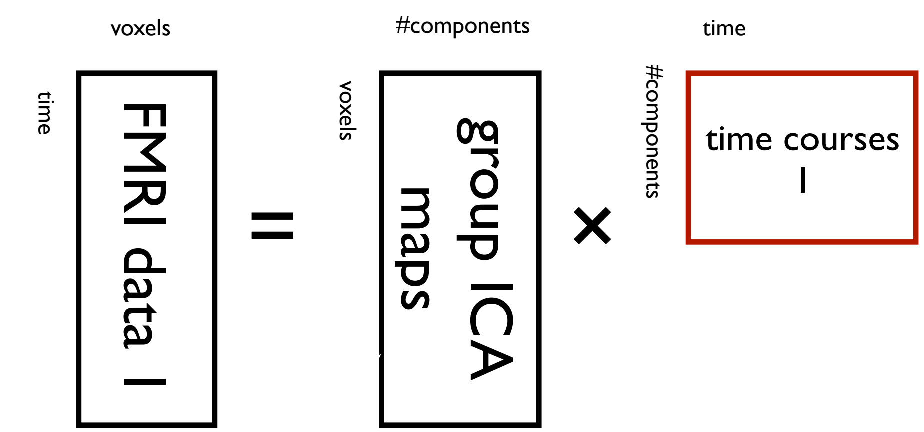
- dr_stage2_ic[#ICA].nii.gz - the re-organised parameter estimate images





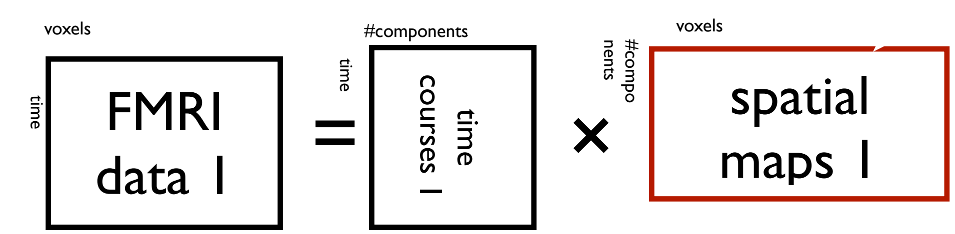
Dual regression outputs

- dr_stage1_subject[#SUB].txt - the timeseries outputs of stage 1 of the dual-regression.



- dr_stage2_subject[#SUB].nii.gz - the spatial maps outputs of stage 2 of the dual-regression.

- dr_stage2_ic[#ICA].nii.gz - the re-organised parameter estimate images



- dr_stage3_ic[#ICA]_tstat[#CON].nii.gz - the output from randomise

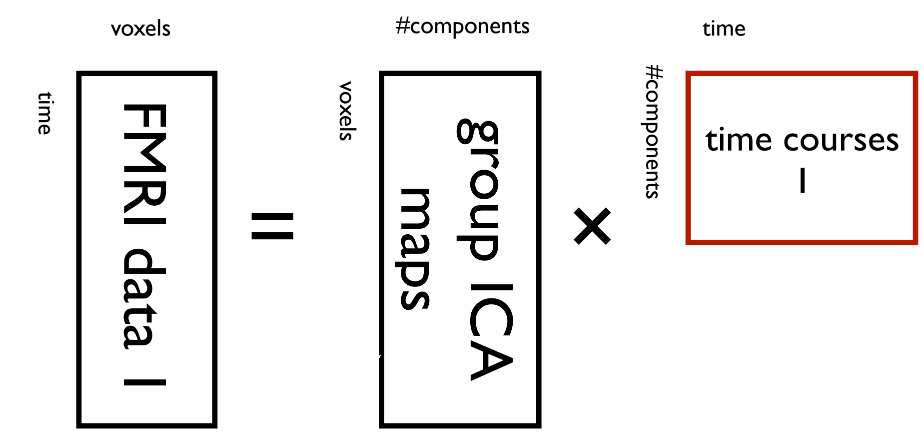
Dual regression outputs

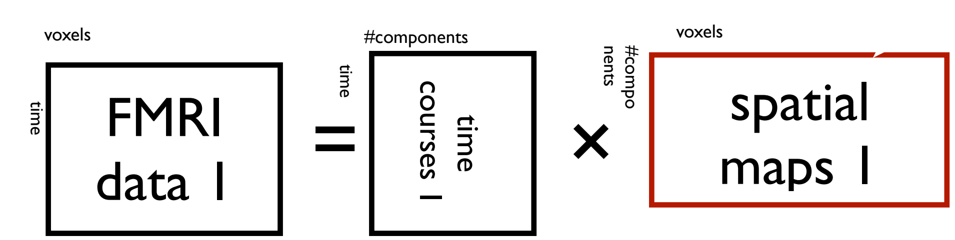
- dr_stage1_subject[#SUB].txt - the timeseries outputs of stage 1 of the dual-regression.

- dr_stage2_subject[#SUB].nii.gz - the spatial maps outputs of stage 2 of the dual-regression.

- dr_stage2_ic[#ICA].nii.gz - the re-organised parameter estimate images

- dr_stage3_ic[#ICA]_tstat[#CON].nii.gz - the output from randomise

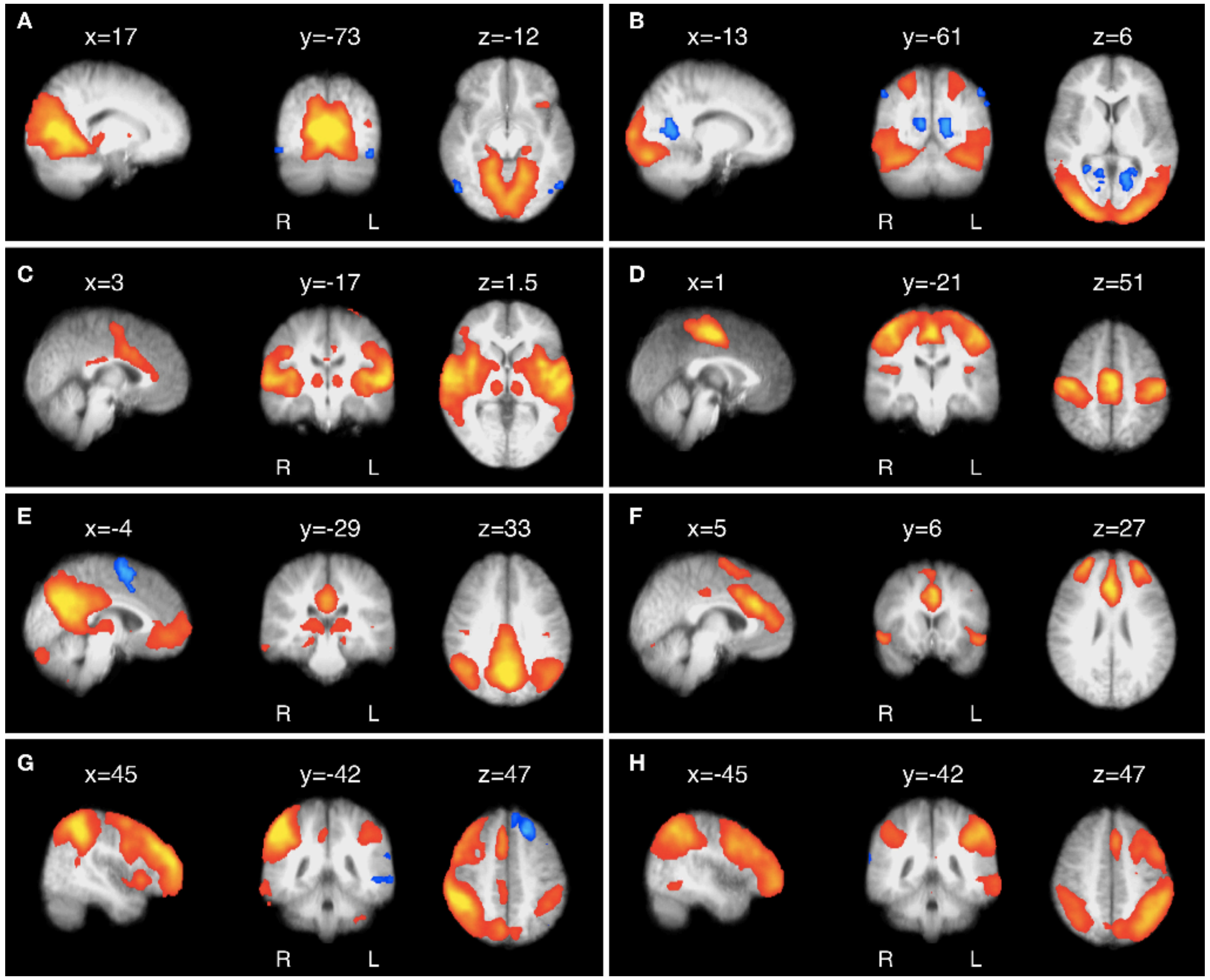




(corrected for multiple comparisons across voxels but not across #components!!)

Group template maps

- Generate from the data using ICA
    - use all data to get unbiased templates

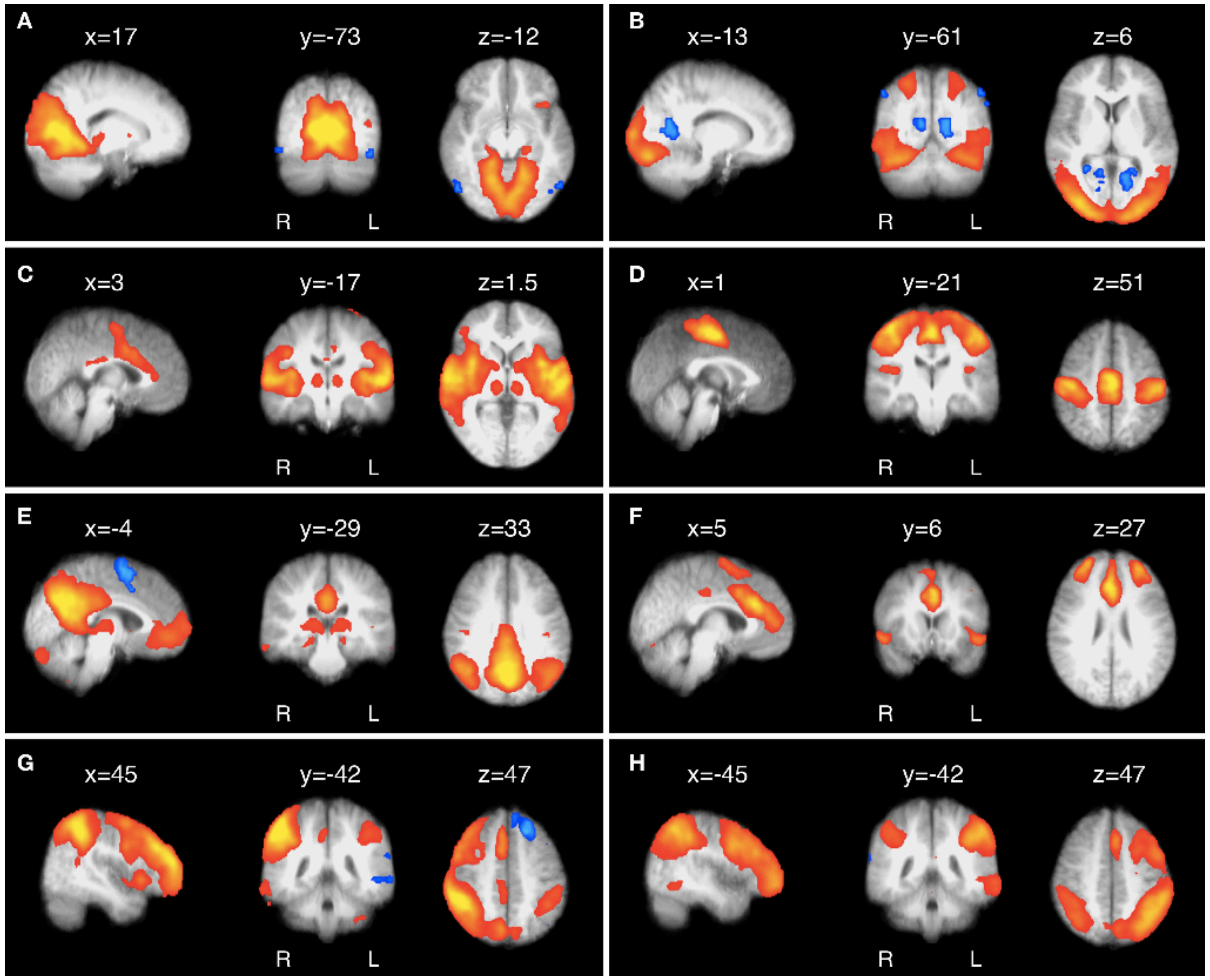


template RSNs

https://www.fmrib.ox.ac.uk/datasets/royalsoc8/

Group template maps

- Generate from the data using ICA
    - use all data to get unbiased templates
    - use independent control group
    - will model signals and artefacts

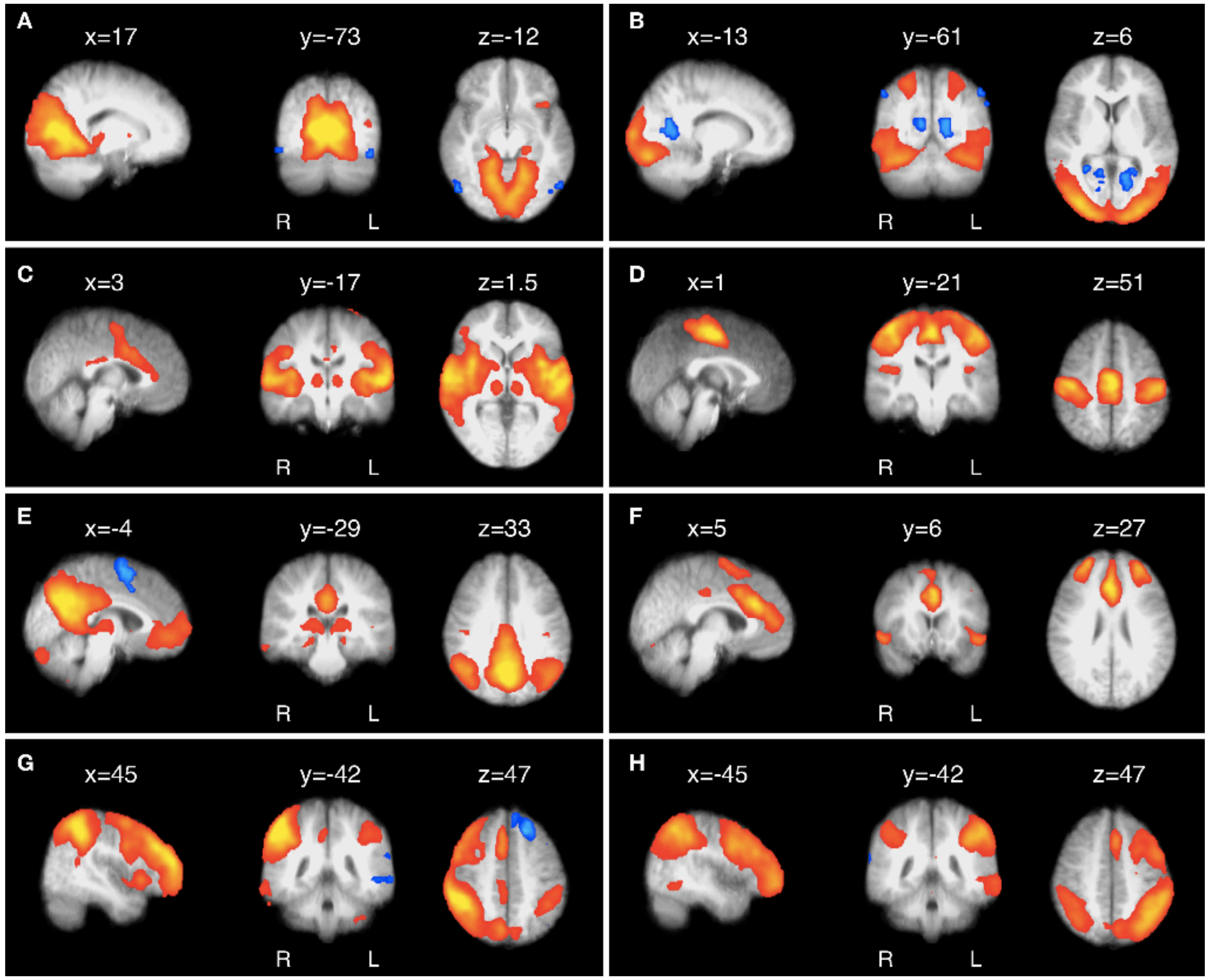


template RSNs

https://www.fmrib.ox.ac.uk/datasets/royalsoc8/

Group template maps

- Generate from the data using ICA
  - use all data to get unbiased templates
  - use independent control group
  - will model signals and artefacts
- use existing template



template RSNs

https://www.fmrib.ox.ac.uk/datasets/royalsoc8/

Resting state fMRI and ICA

- Introduction to resting state
- Independent Component Analysis
- Single-subject ICA
- Multi-subject ICA
- Dual regression

Resting state fMRI and ICA

Available from:

- [Oxford University Press]
- [Amazon]