Model-free Functional Data Analysis

MELODIC

Multivariate Exploratory Linear Optimised Decomposition into Independent Components

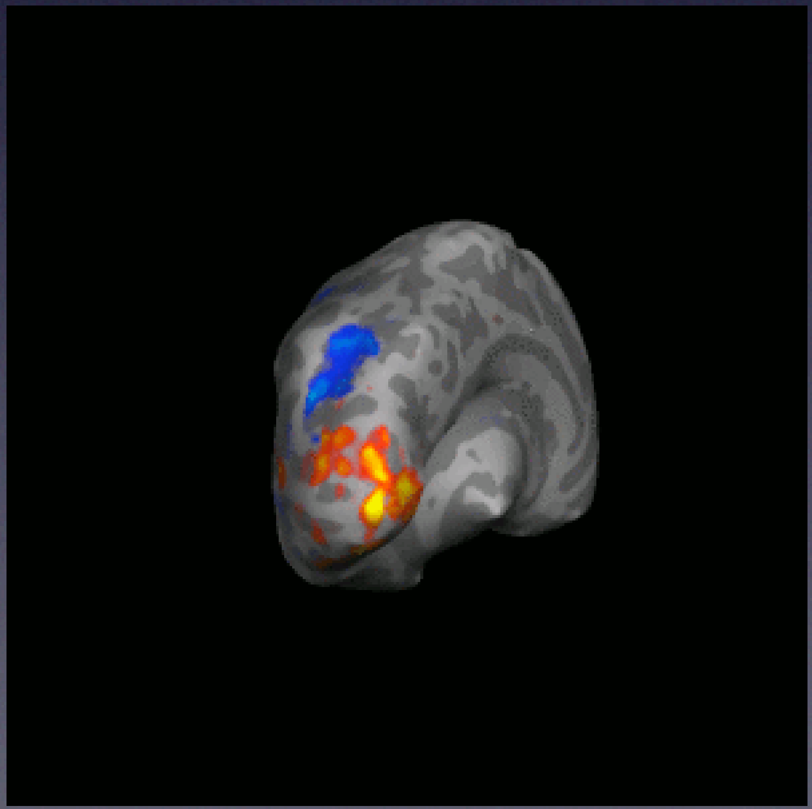
- decomposes data into a set of statistically independent spatial component maps and associated time courses
- fully automated (incl. estimation of the number of components)
- inference on IC maps using alternative hypothesis testing

Model-free Functional Data Analysis

MELODIC

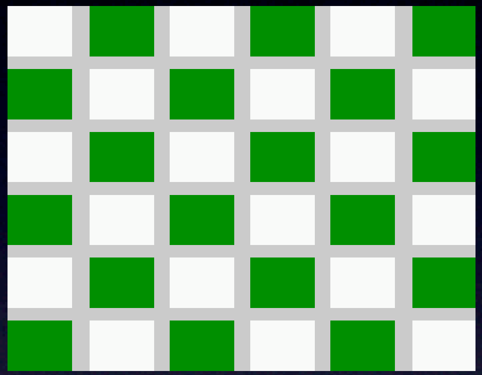
Multivariate Exploratory Linear Optimised Decomposition into Independent Components

- decomposes data into a set of statistically independent spatial component maps and associated time courses
- fully automated (incl. estimation of the number of components)
- inference on IC maps using alternative hypothesis testing



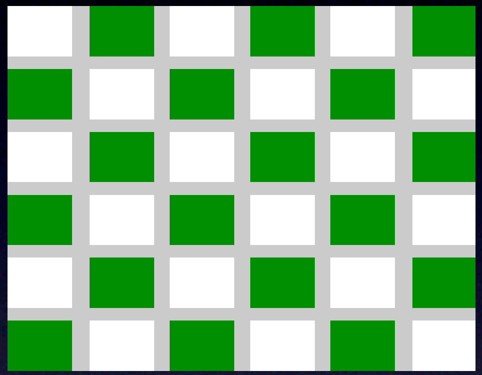
The FMRI inferential path

Experiment

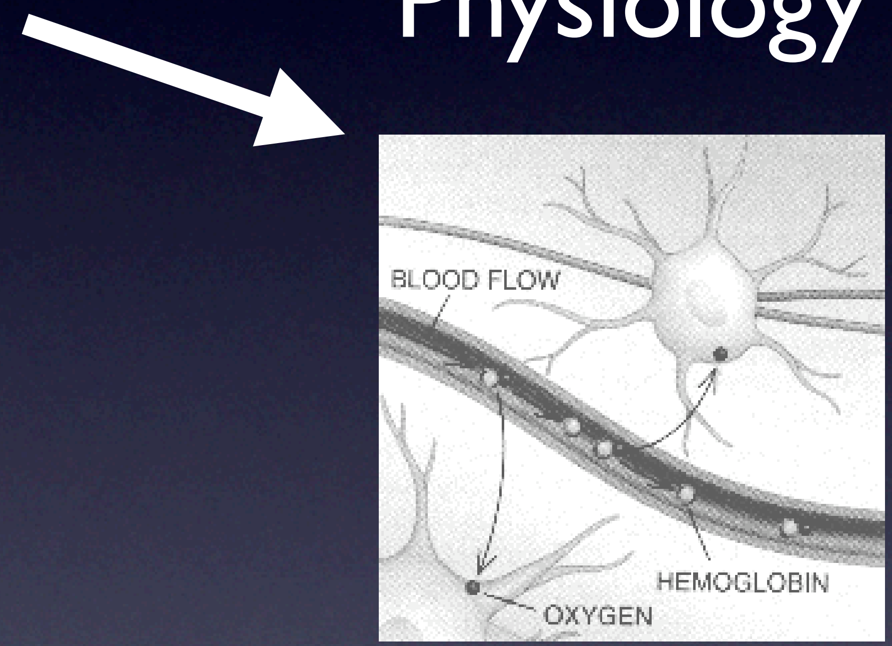


The FMRI inferential path

Experiment

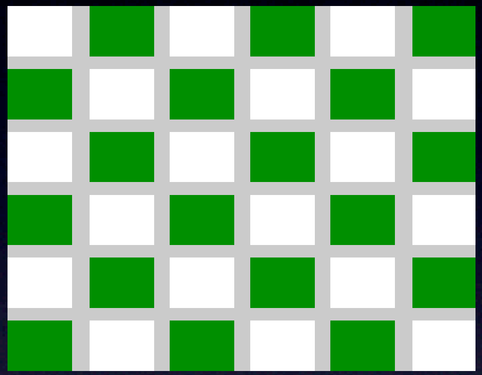


Physiology

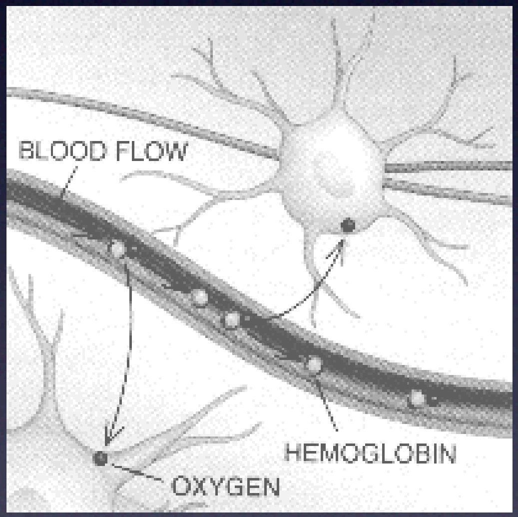


The FMRI inferential path

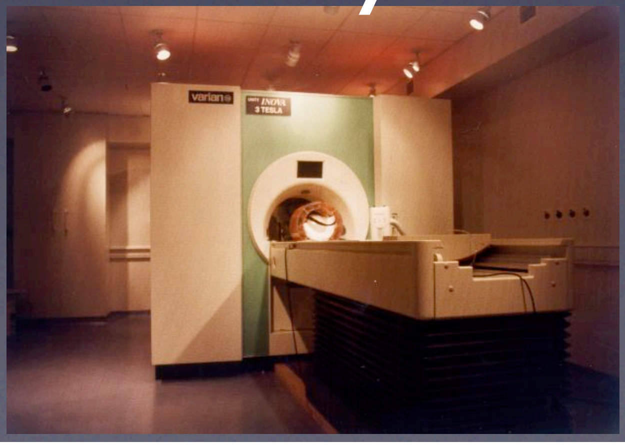
Experiment



Physiology

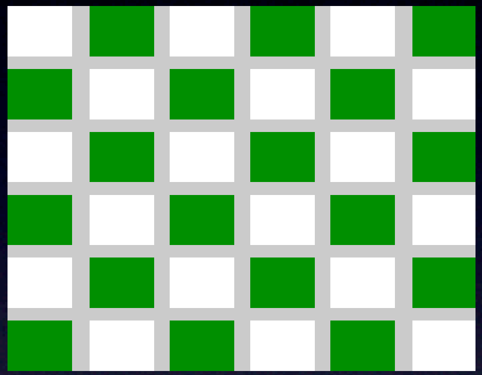


MR Physics

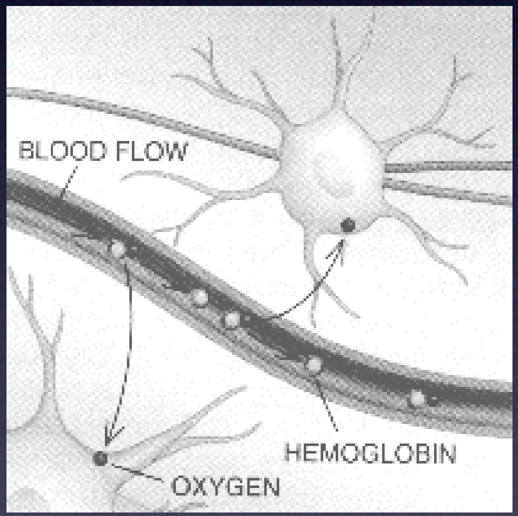


The FMRI inferential path

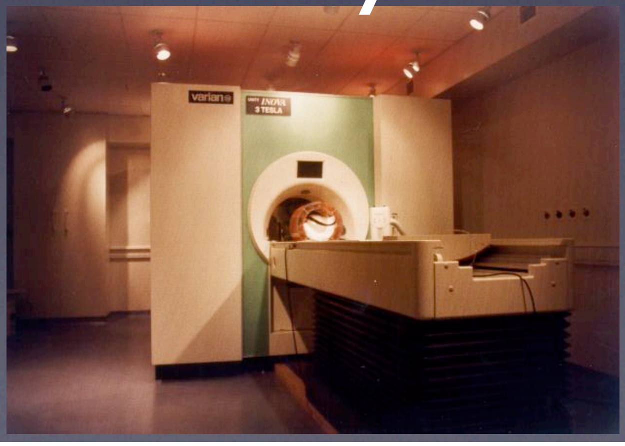
Experiment



Physiology



MR Physics

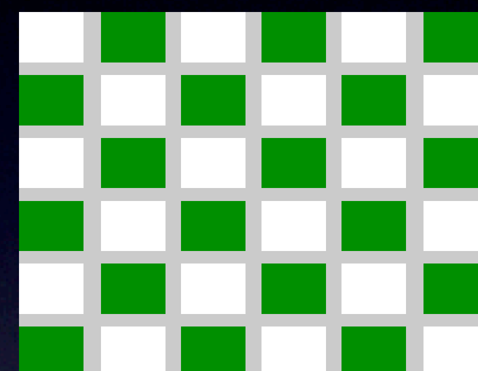


Analysis

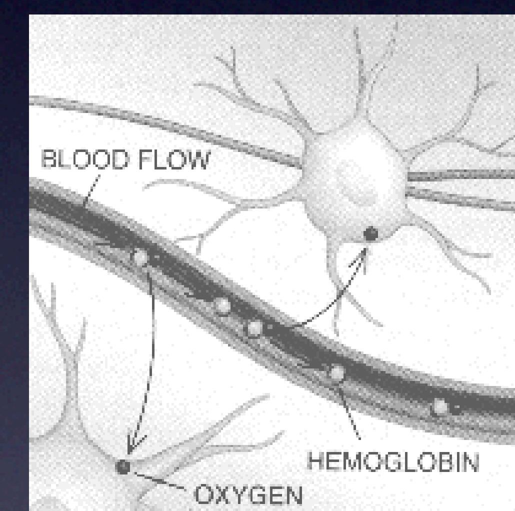
FSL

The FMRI inferential path

Experiment

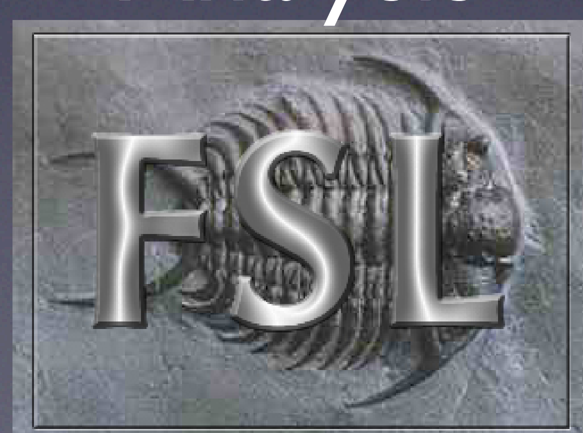


Physiology

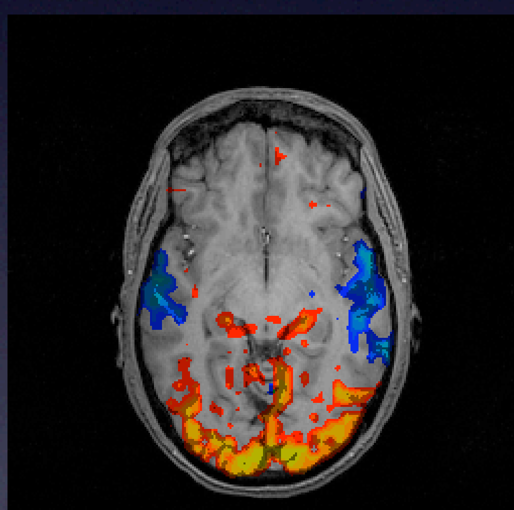


MR Physics

Analysis



Interpretation of final results

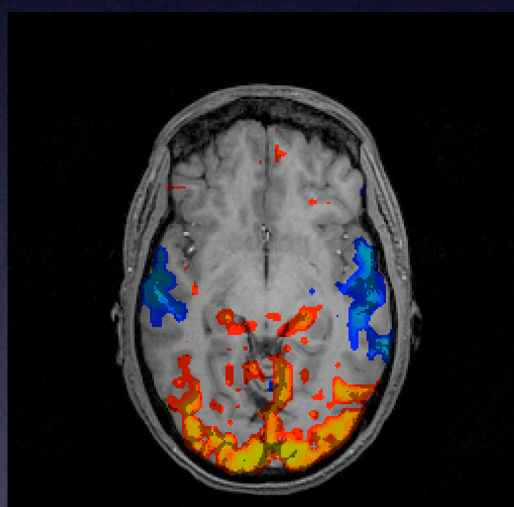


Variability in FMRI

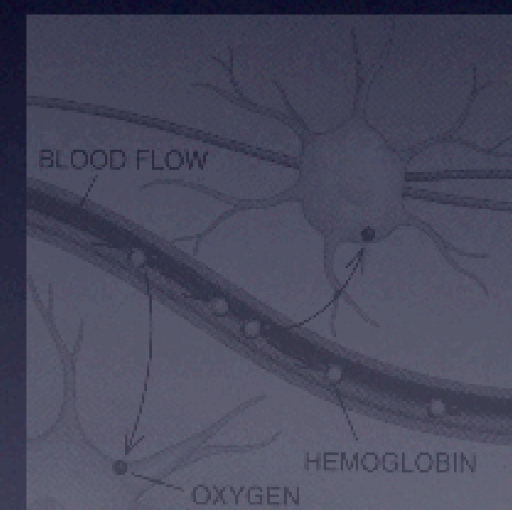
Experiment

suboptimal event timing,
non-efficient design, etc.

Interpretation of final results



Physiology



Analysis

MR Physics

Variability in FMRI

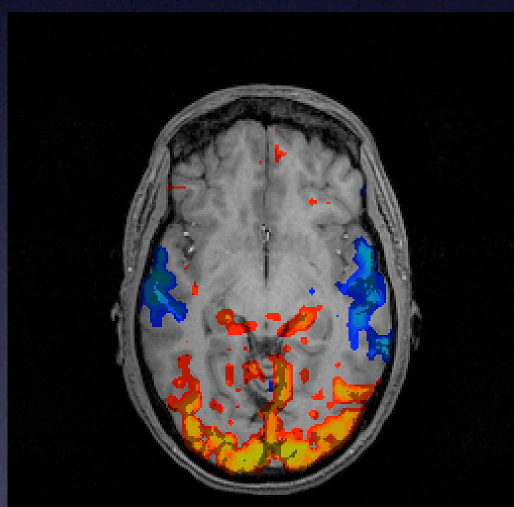
Experiment

suboptimal event timing,
non-efficient design, etc.

Physiology

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

Interpretation
of final results



Analysis

MR Physics

Variability in FMRI

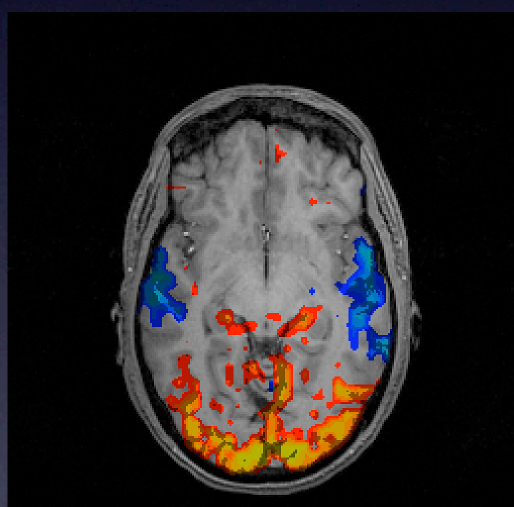
Experiment

suboptimal event timing,
non-efficient design, etc.

Physiology

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

Interpretation
of final results



MR Physics

MR noise,
field inhomogeneity,
MR artefacts etc.

Analysis

Variability in FMRI

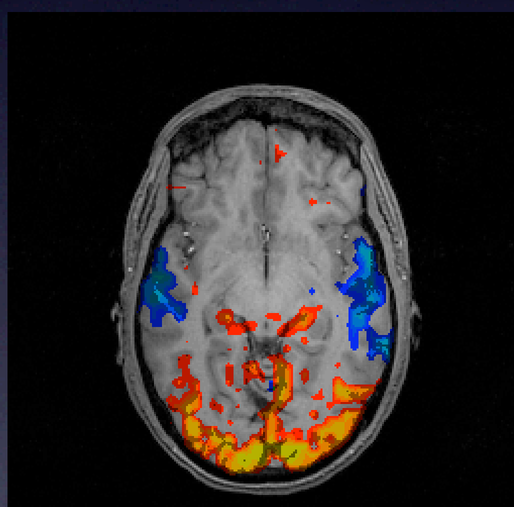
Experiment

suboptimal event timing,
non-efficient design, etc.

Physiology

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

Interpretation
of final results

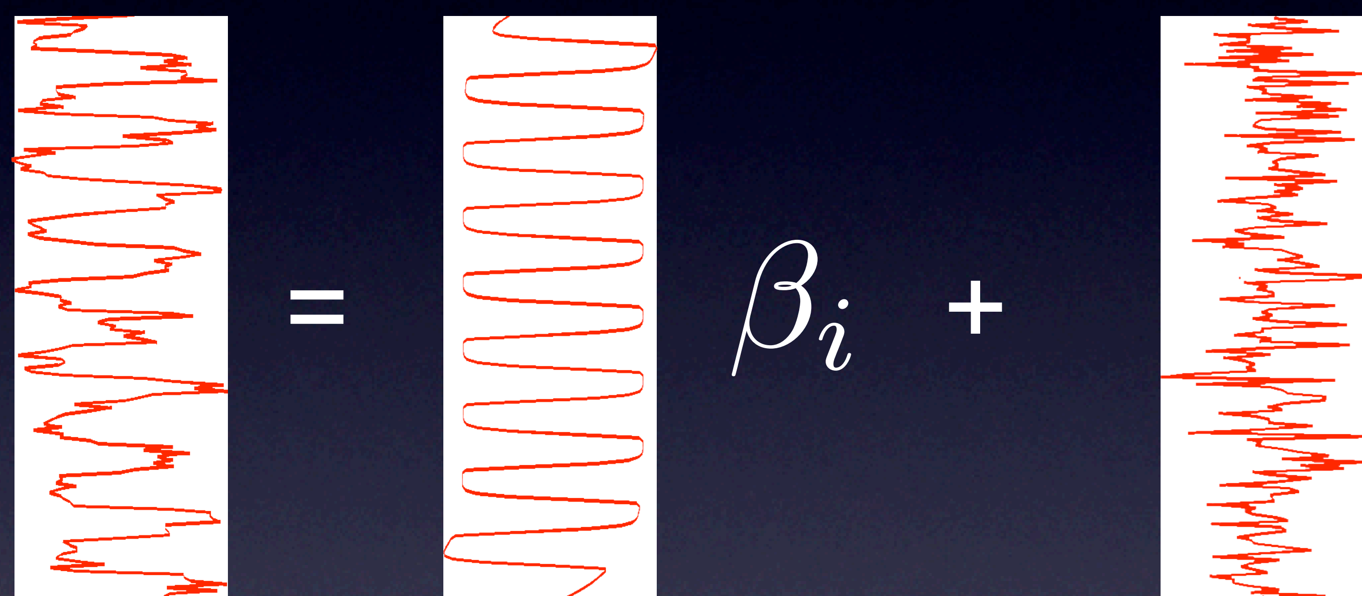


filtering & sampling
artefacts, design
misspecification, stats &
thresholding issues etc.

MR Physics

MR noise,
field inhomogeneity,
MR artefacts etc.

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

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

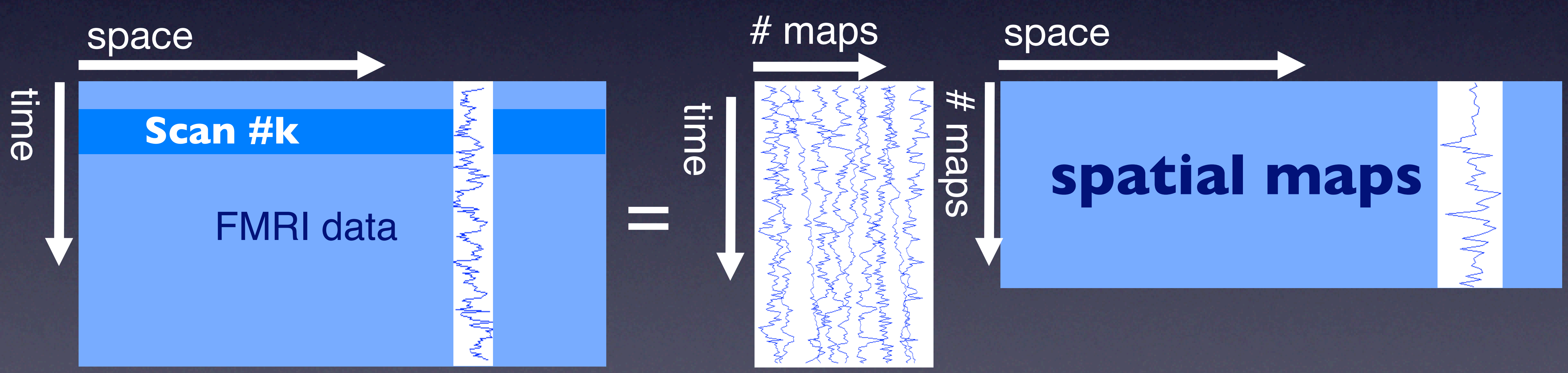
- can give unexpected results

EDA 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)
- typically are *multivariate* and *linear*

EDA techniques for FMRI

- are mostly multivariate
- often provide a multivariate linear decomposition:

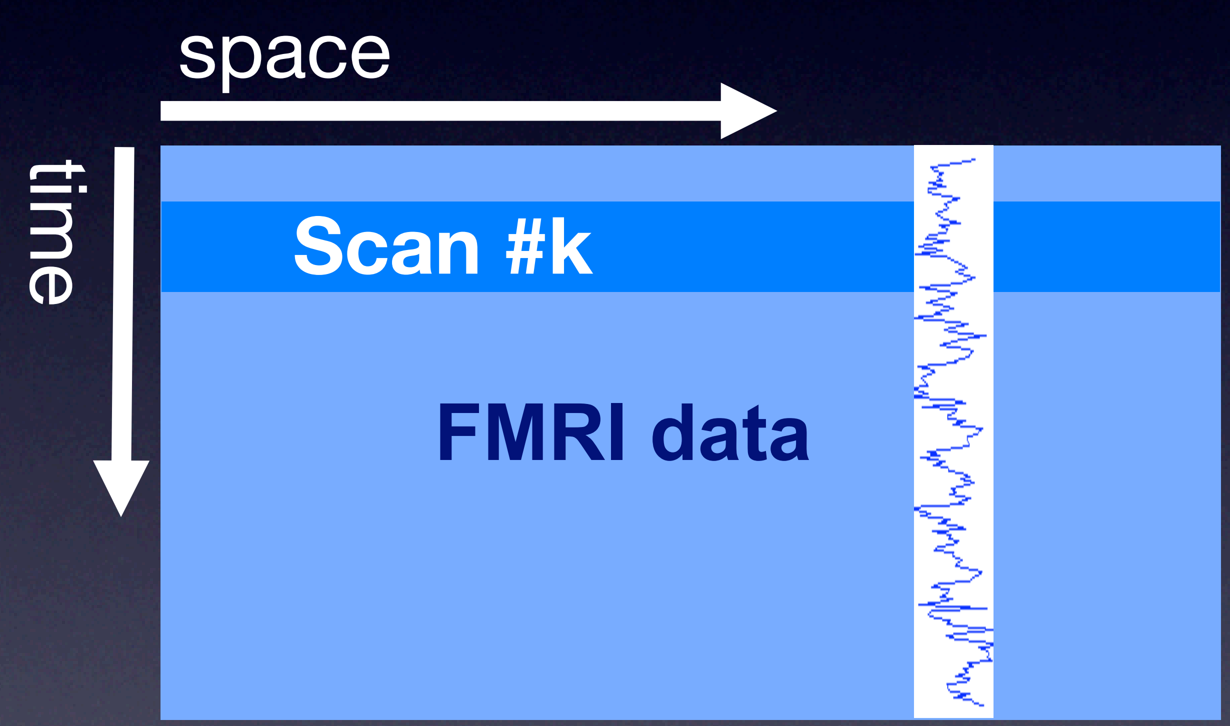


Data is represented as a 2D matrix and decomposed into factor matrices (or modes)

PCA for FMRI

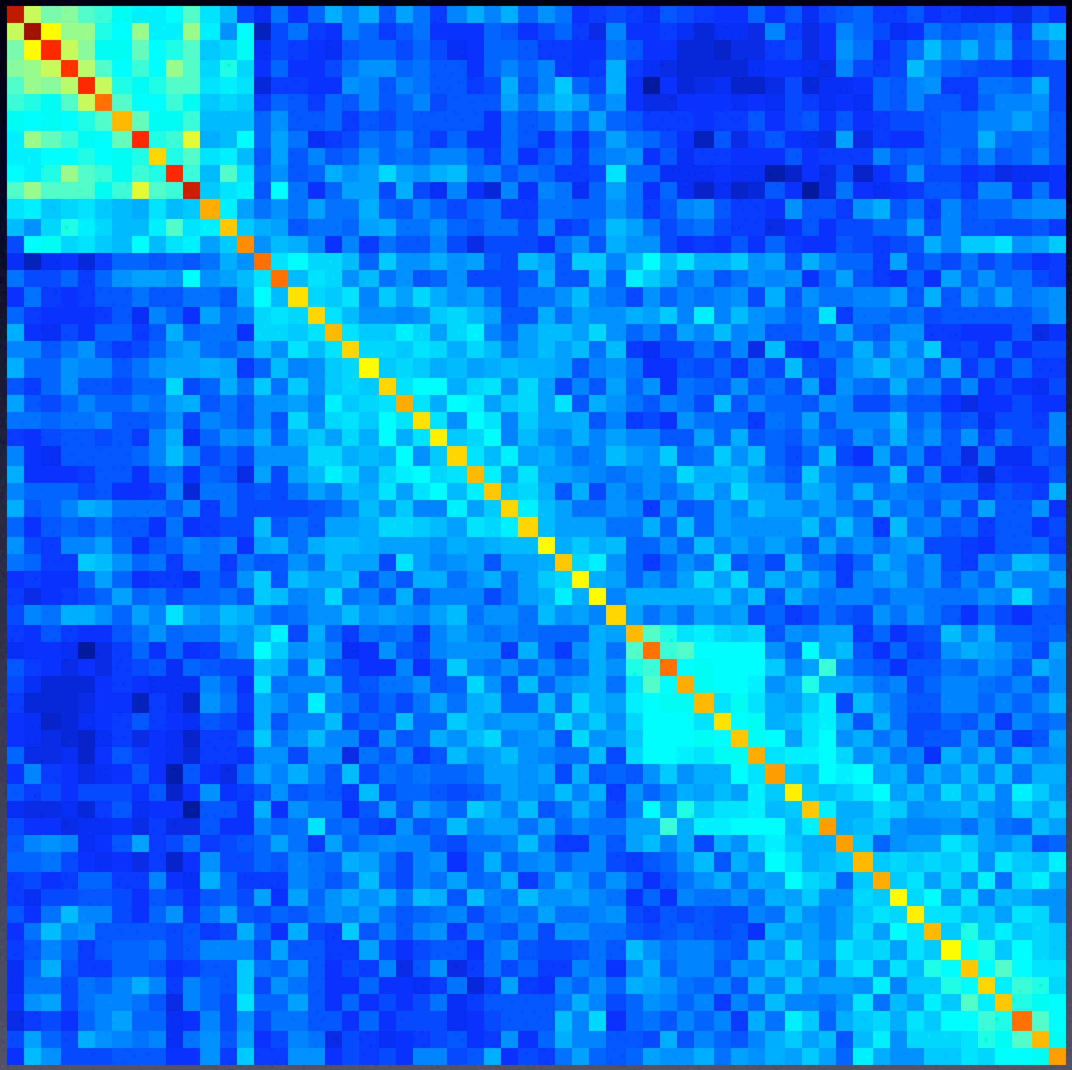
PCA for FMRI

1. assemble all (de-meaned) data as a matrix



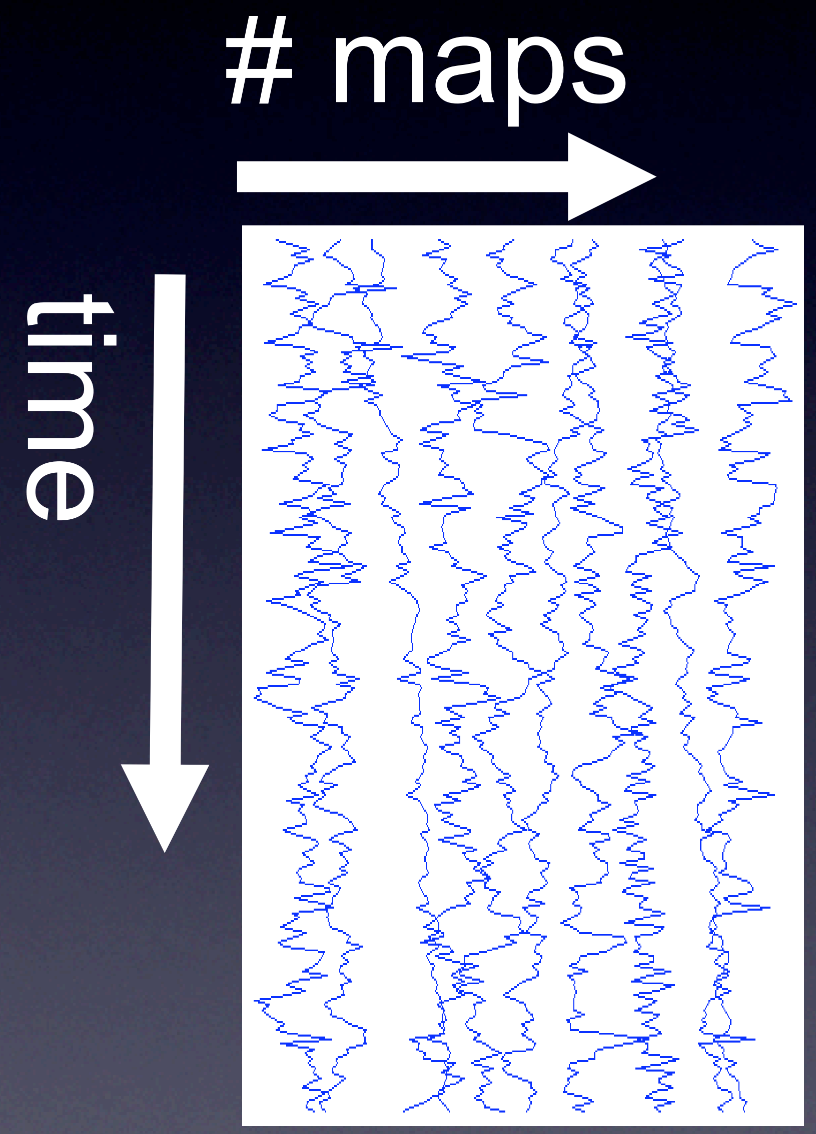
PCA for FMRI

1. assemble all (de-meaned) data as a matrix
2. calculate the data covariance matrix



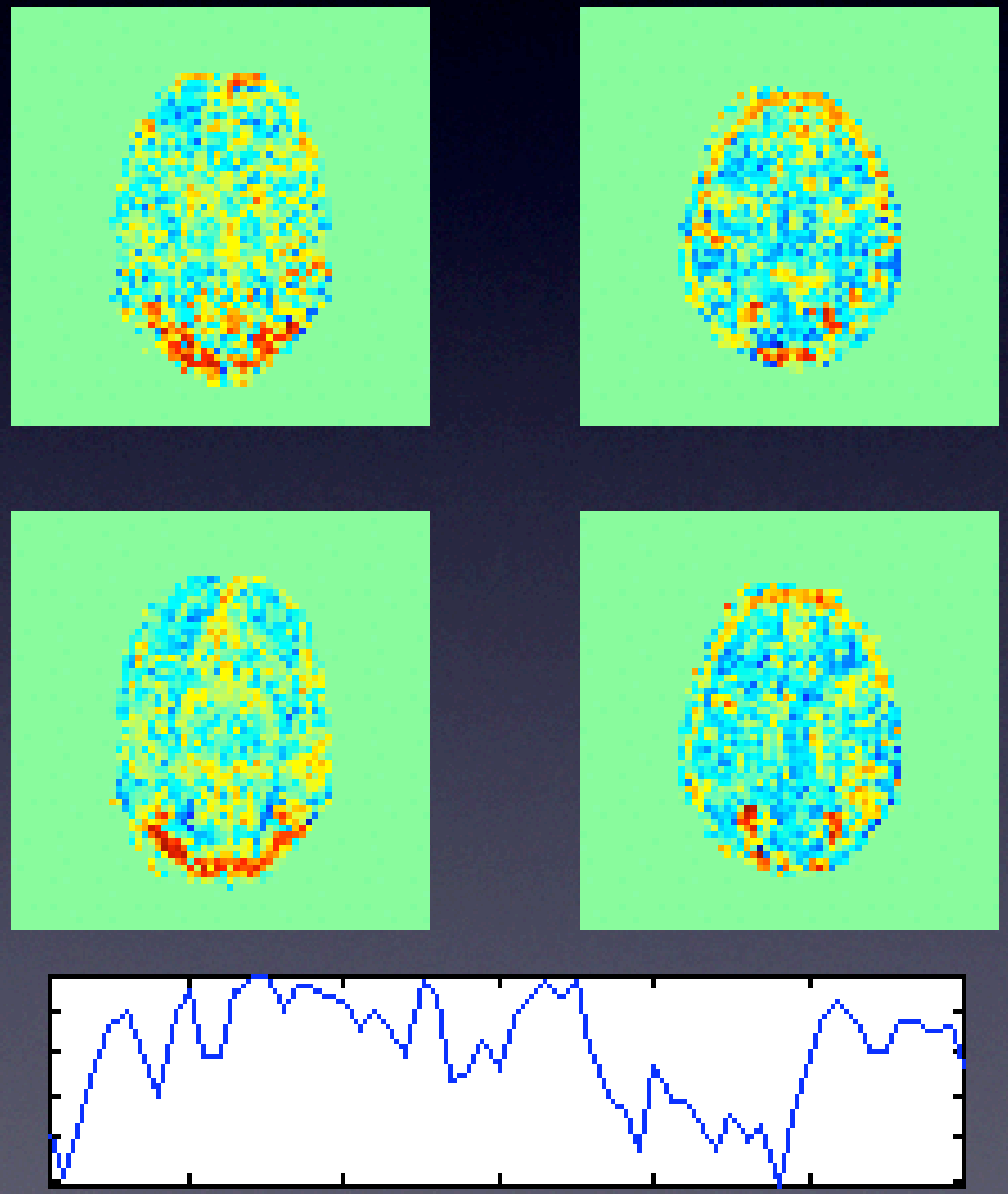
PCA for FMRI

1. assemble all (de-meaned) data as a matrix
2. calculate the data covariance matrix
3. calculate SVD



PCA for FMRI

1. assemble all (de-meaned) data as a matrix
2. calculate the data covariance matrix
3. calculate SVD
4. project data onto the Eigenvectors

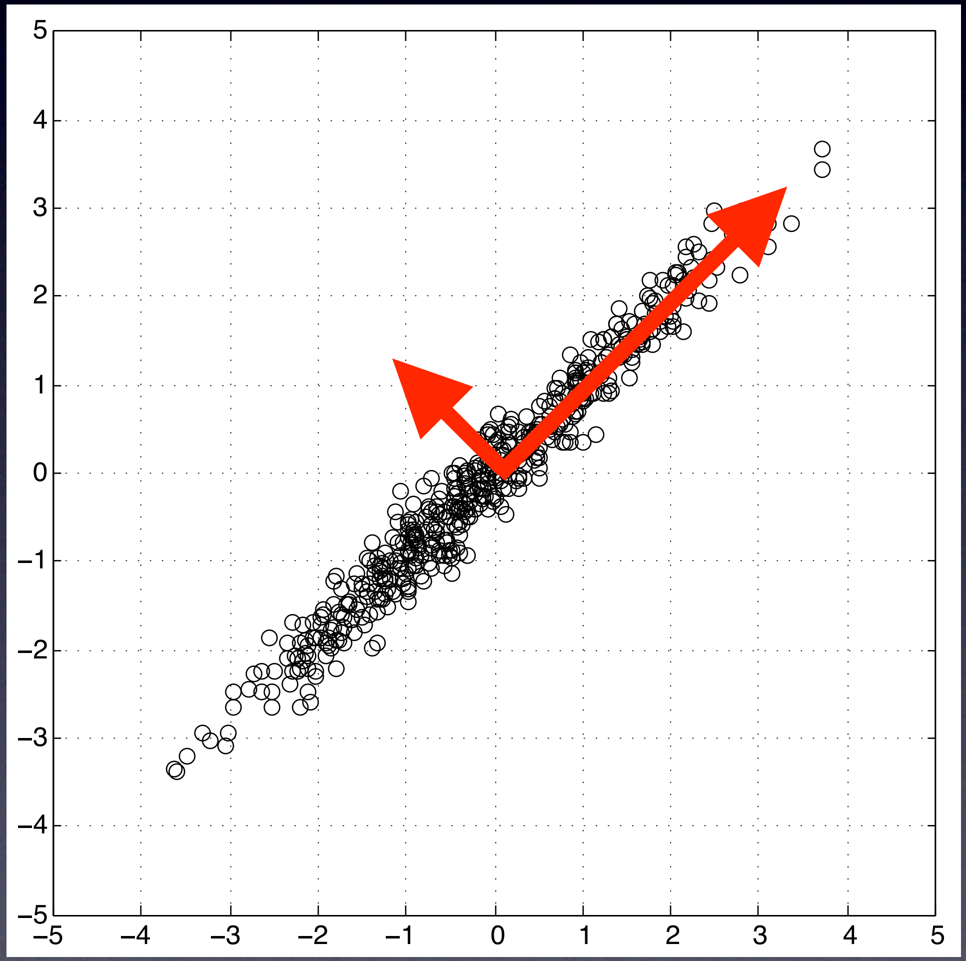


PCA decompositions: Implications

PCA decompositions: Implications

- spatial maps *and* time courses are orthogonal (uncorrelated)

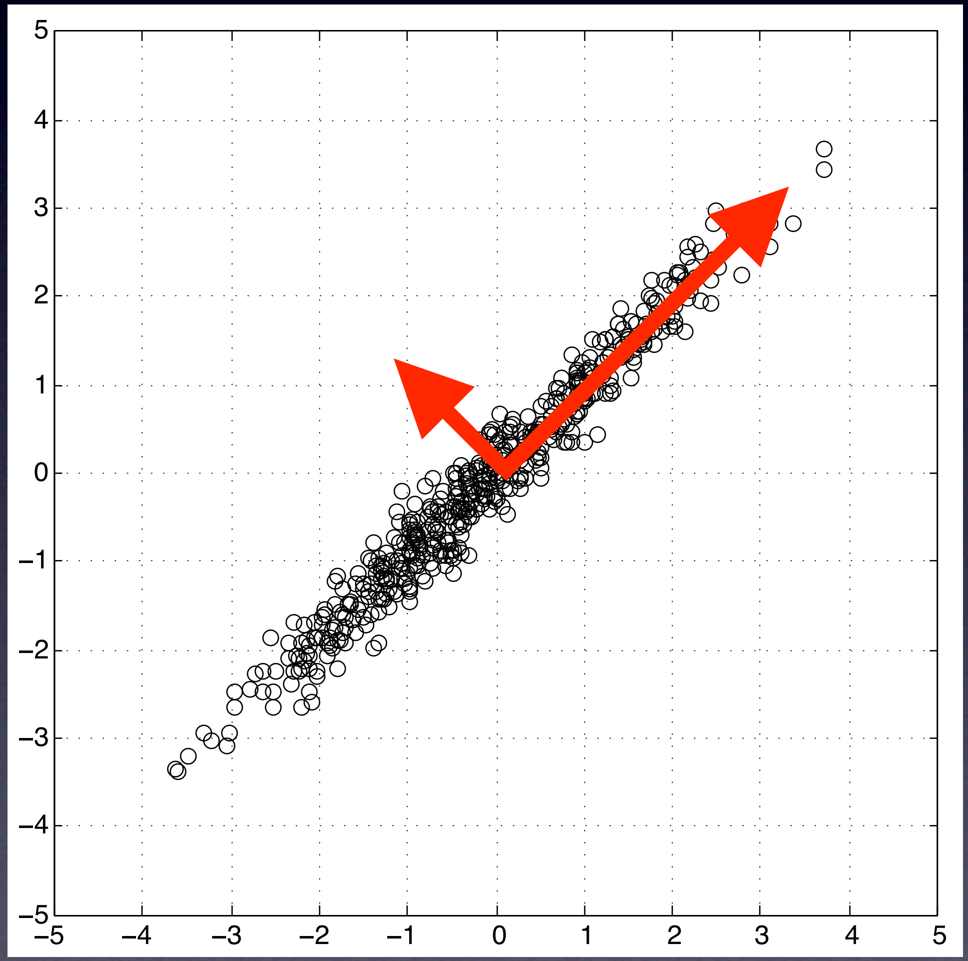
PCA vs. ICA



Gaussian data

PCA vs. ICA

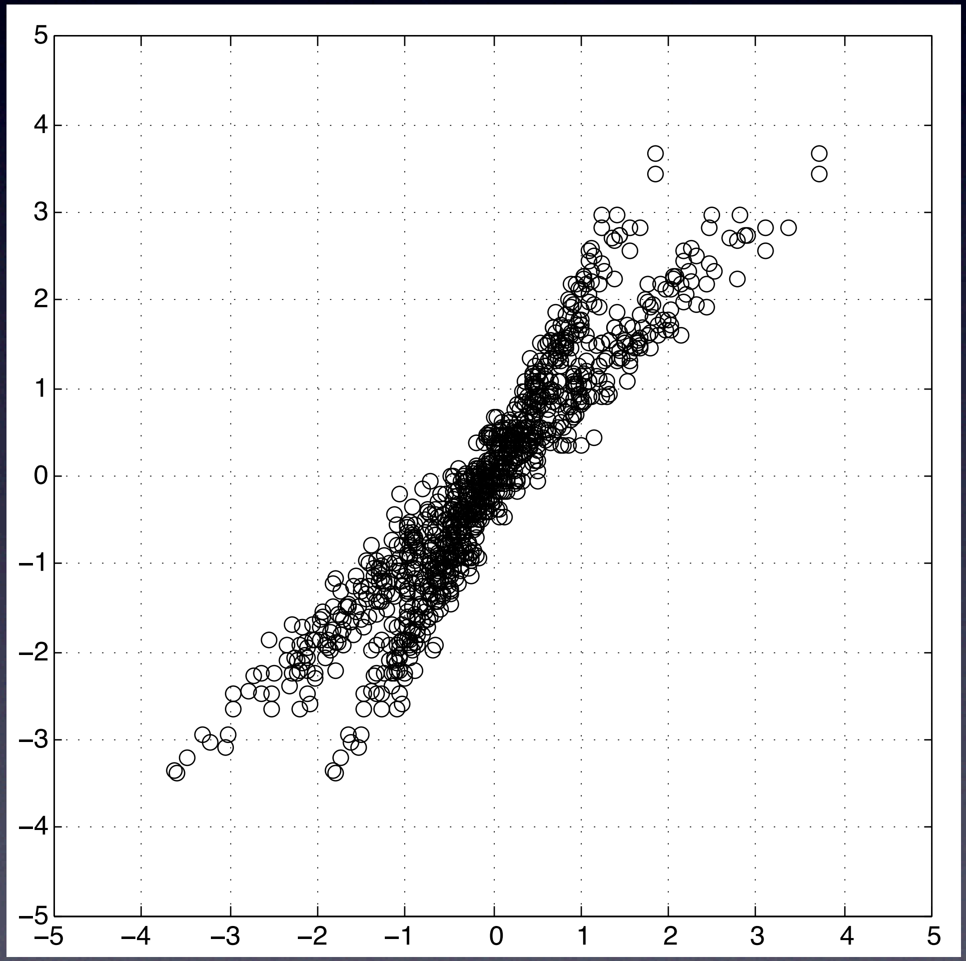
- PCA finds projections of maximum amount of variance in Gaussian data (uses 2nd order statistics only)



Gaussian data

PCA vs. ICA

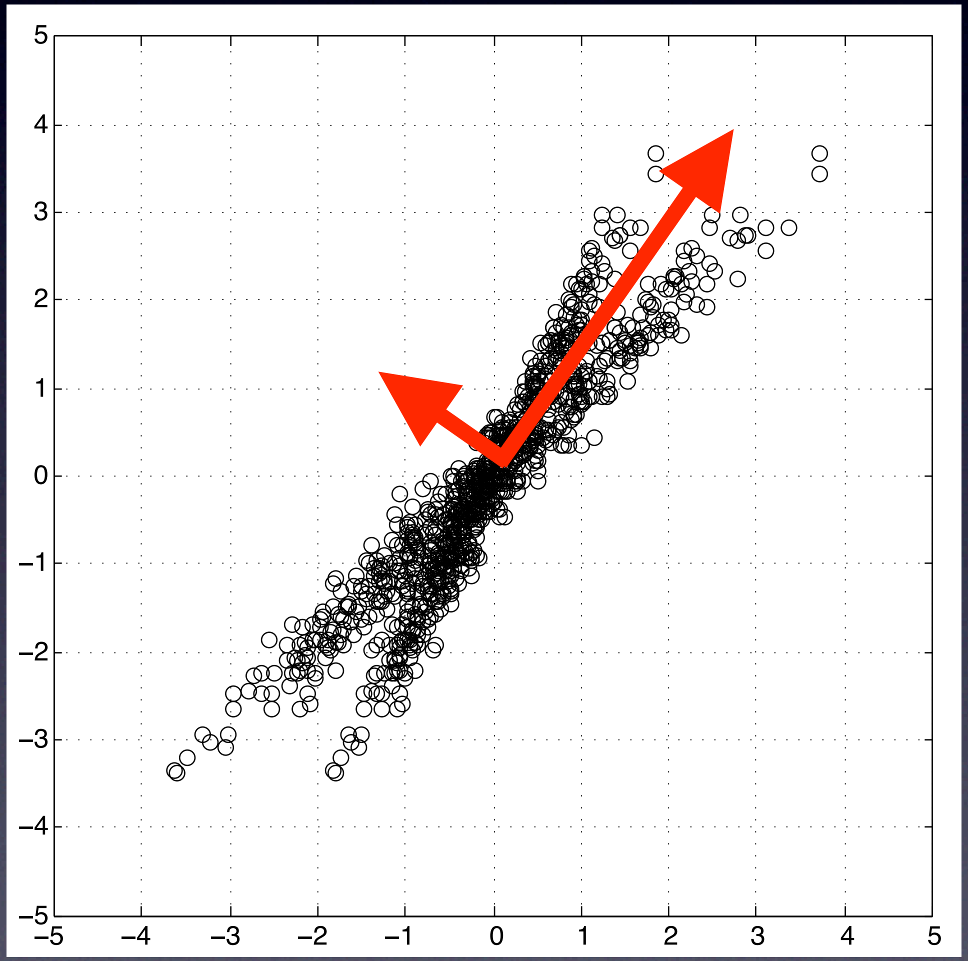
- PCA finds projections of maximum amount of variance in Gaussian data (uses 2nd order statistics only)



non-Gaussian data

PCA vs. ICA

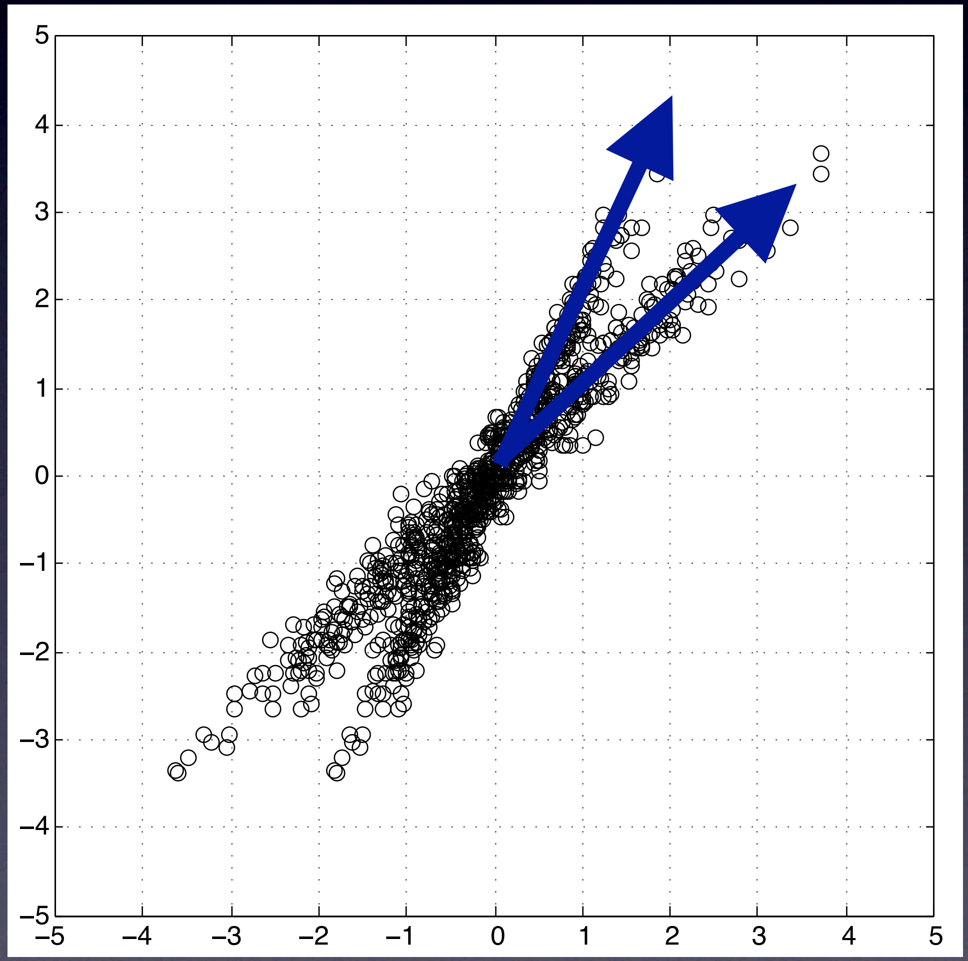
- PCA finds projections of maximum amount of variance in Gaussian data (uses 2nd order statistics only)



non-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

PCA decompositions: Implications

- spatial maps *and* time courses are orthogonal (uncorrelated)

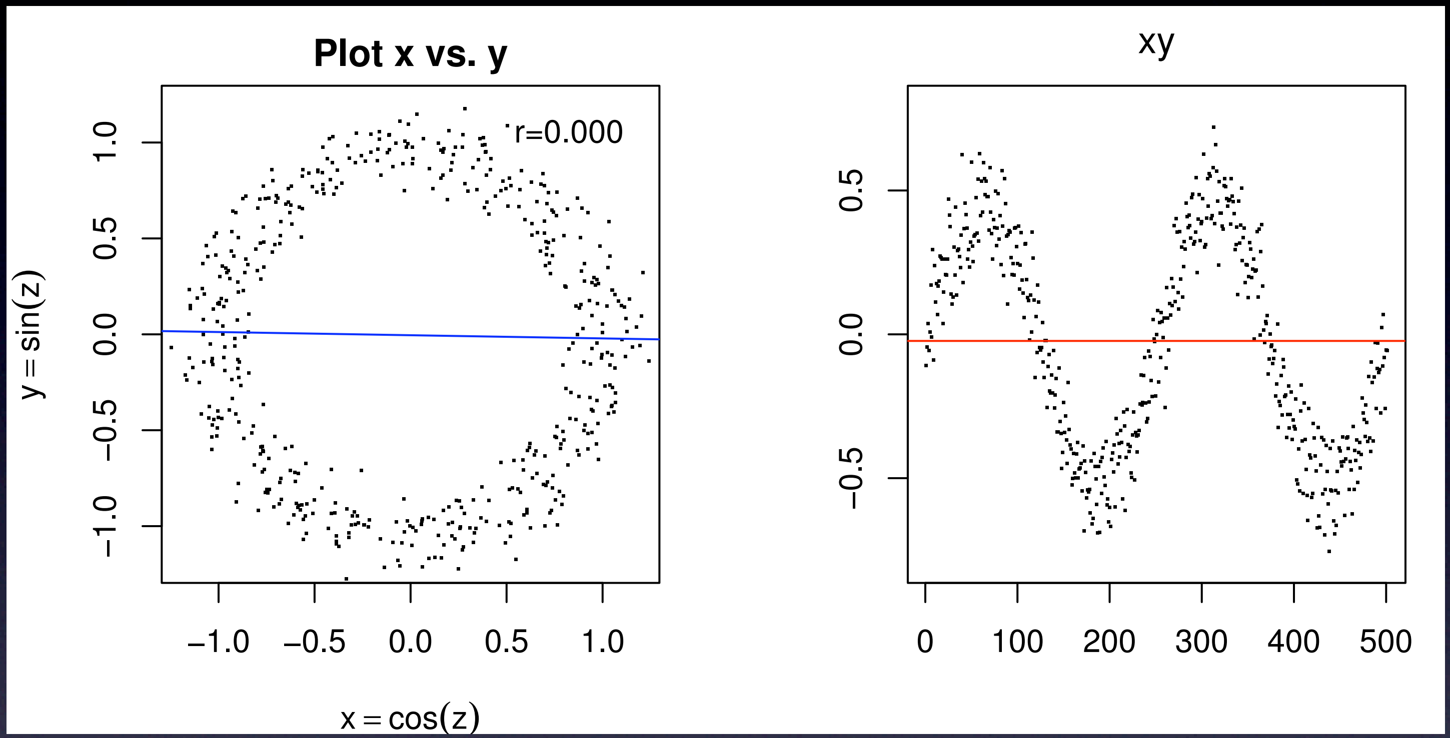
PCA decompositions: Implications

- spatial maps *and* time courses are orthogonal (uncorrelated)
- uncorrelated ≠ un-related

Correlation vs. independece

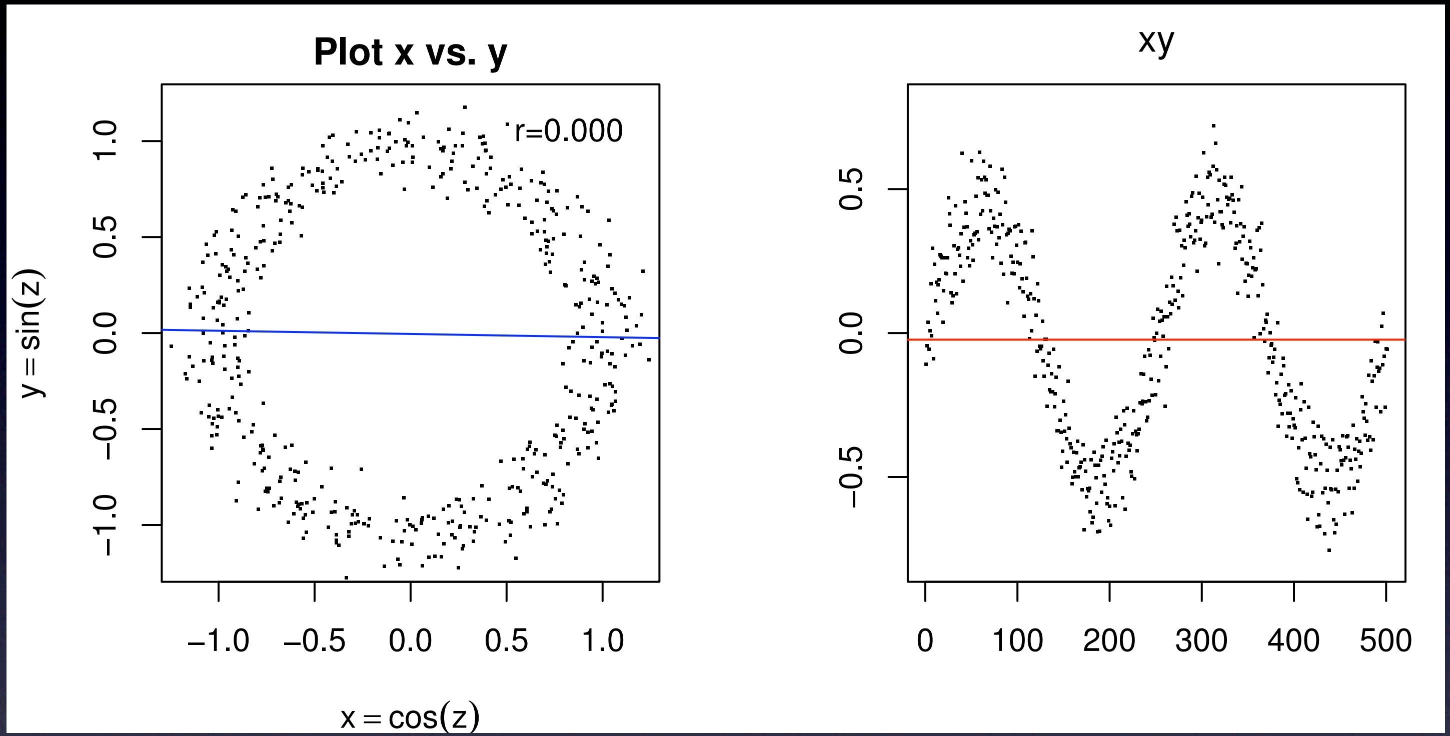
Correlation vs. independece

- de-correlated signals can still be dependent

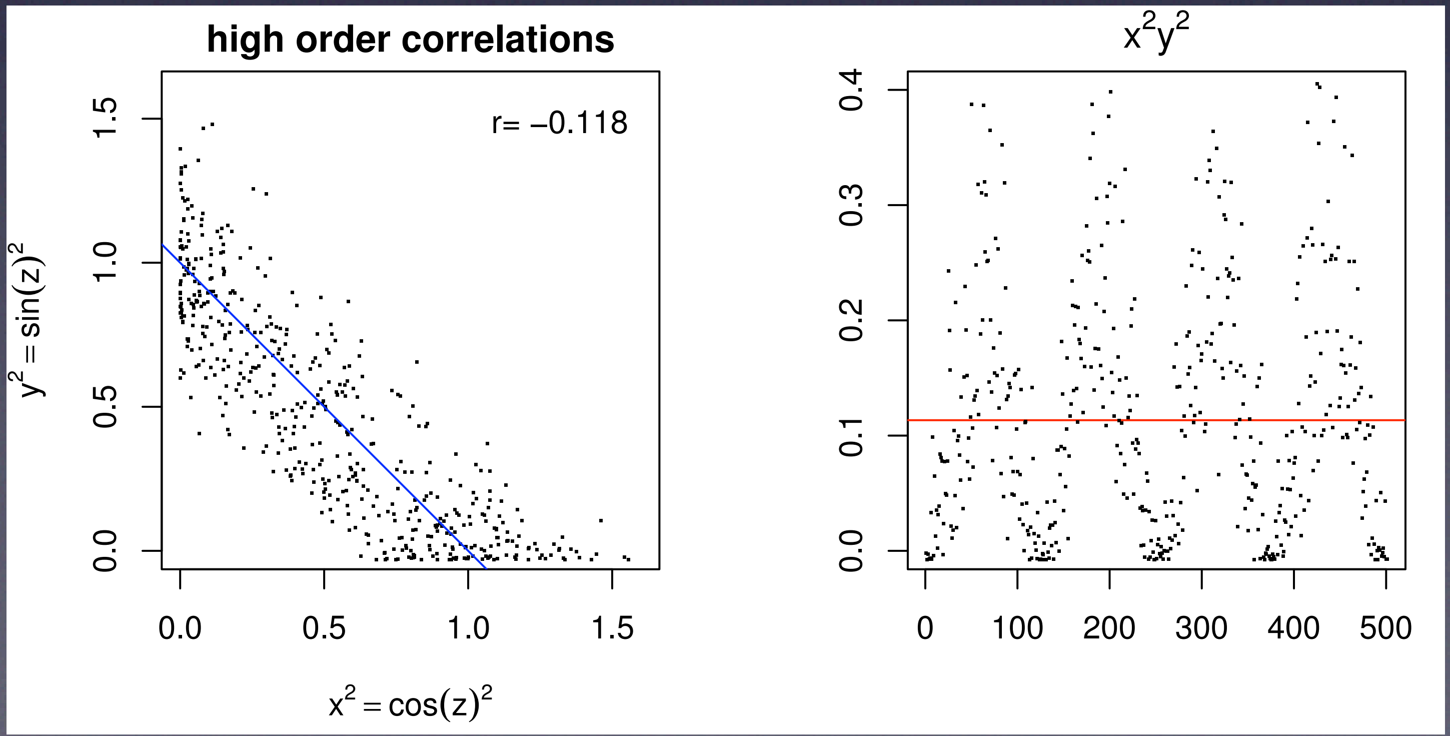


Correlation vs. independece

- de-correlated signals can still be dependent



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



Stone et al. 2002

PCA decompositions: Implications

PCA decompositions: Implications

- spatial maps *and* time courses are orthogonal (uncorrelated)
- uncorrelated ≠ un-related
- only looks at 2nd order statistics: problem with non-Gaussian sources
- PCA is *rotationally invariant*

Rotational invariance

Rotational invariance

- consider de-correlated signals Y , i.e. $R_Y = I$

Rotational invariance

- consider de-correlated signals Y, i.e. $R_Y = I$
- consider projecting the data onto a basis Q

$$R_{QY} \propto QYY^TQ^T = QQ^T$$

Rotational invariance

- consider de-correlated signals Y , i.e. $R_Y = I$
- consider projecting the data onto a basis Q

$$R_{QY} \propto QYY^TQ^T = QQ^T$$

- then , $R_{QY} = R_Y = I$ for any Q s.t. $QQ^T = I$

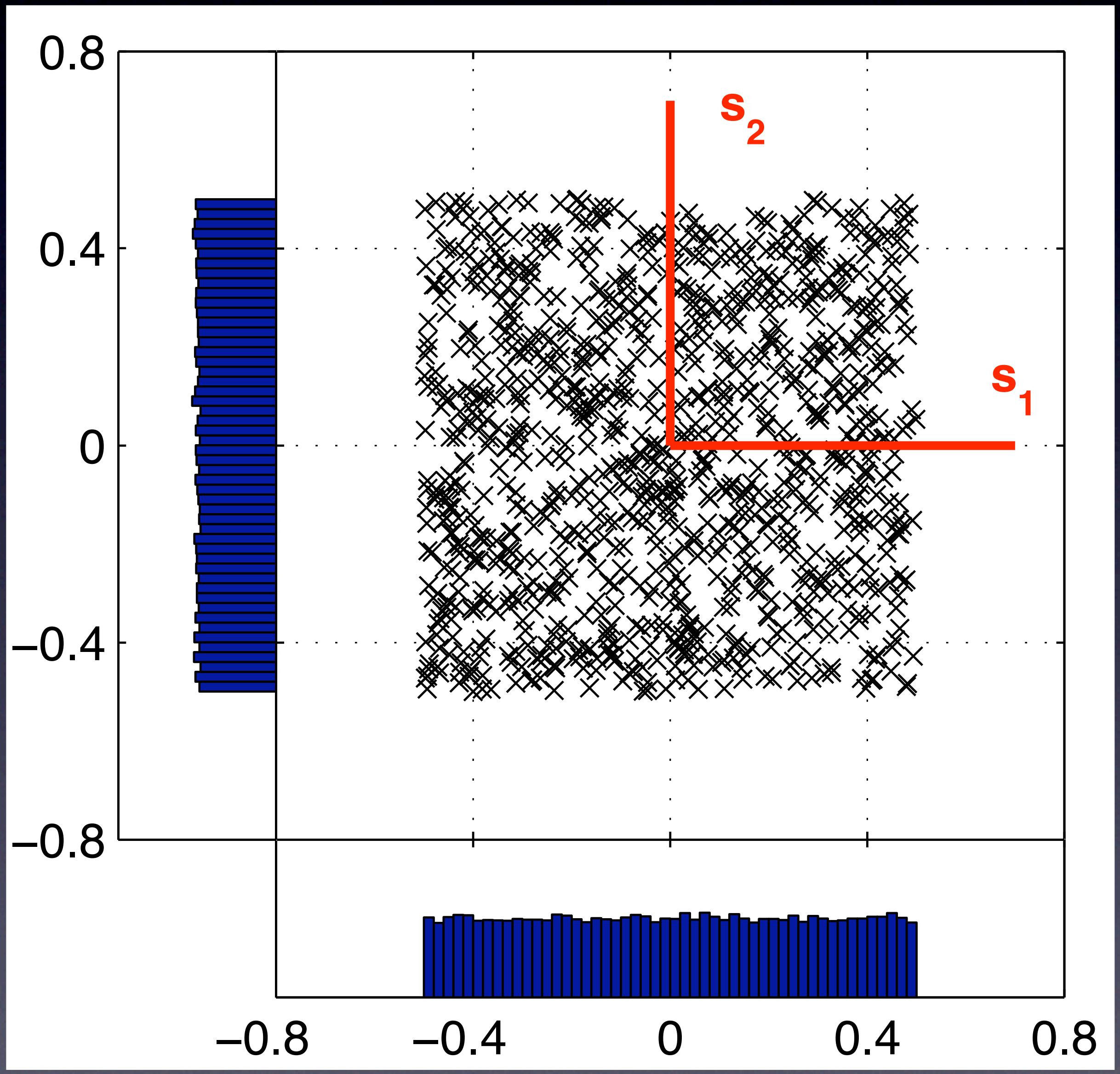
Rotational invariance

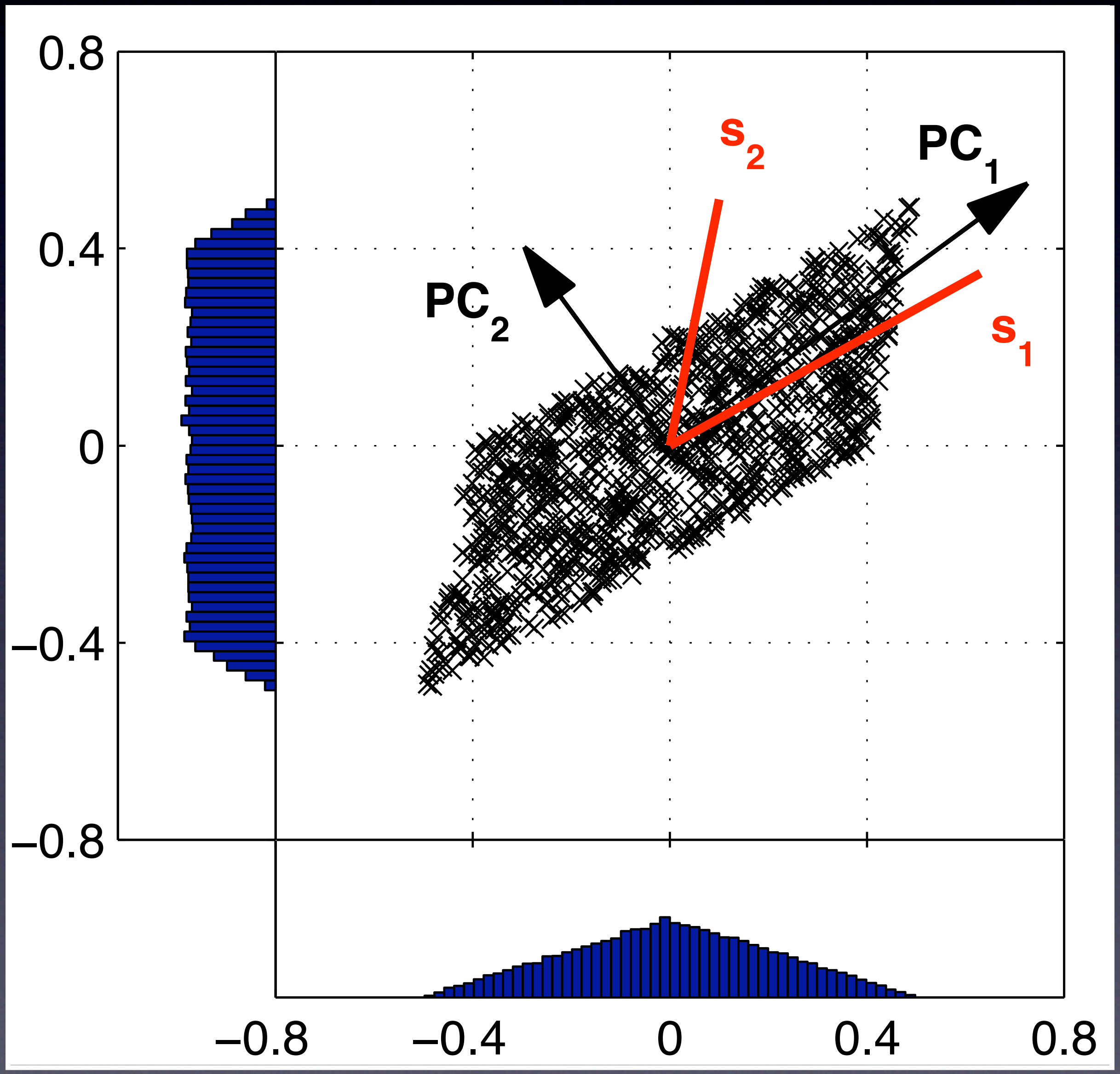
- consider de-correlated signals Y , i.e. $R_Y = I$
- consider projecting the data onto a basis Q

$$R_{QY} \propto QYY^TQ^T = QQ^T$$

- then , $R_{QY} = R_Y = I$ for any Q s.t. $QQ^T = I$
- $QQ^T = I \;:\; Q$ is a *rotation matrix*

The Geometry of PCA and ICA

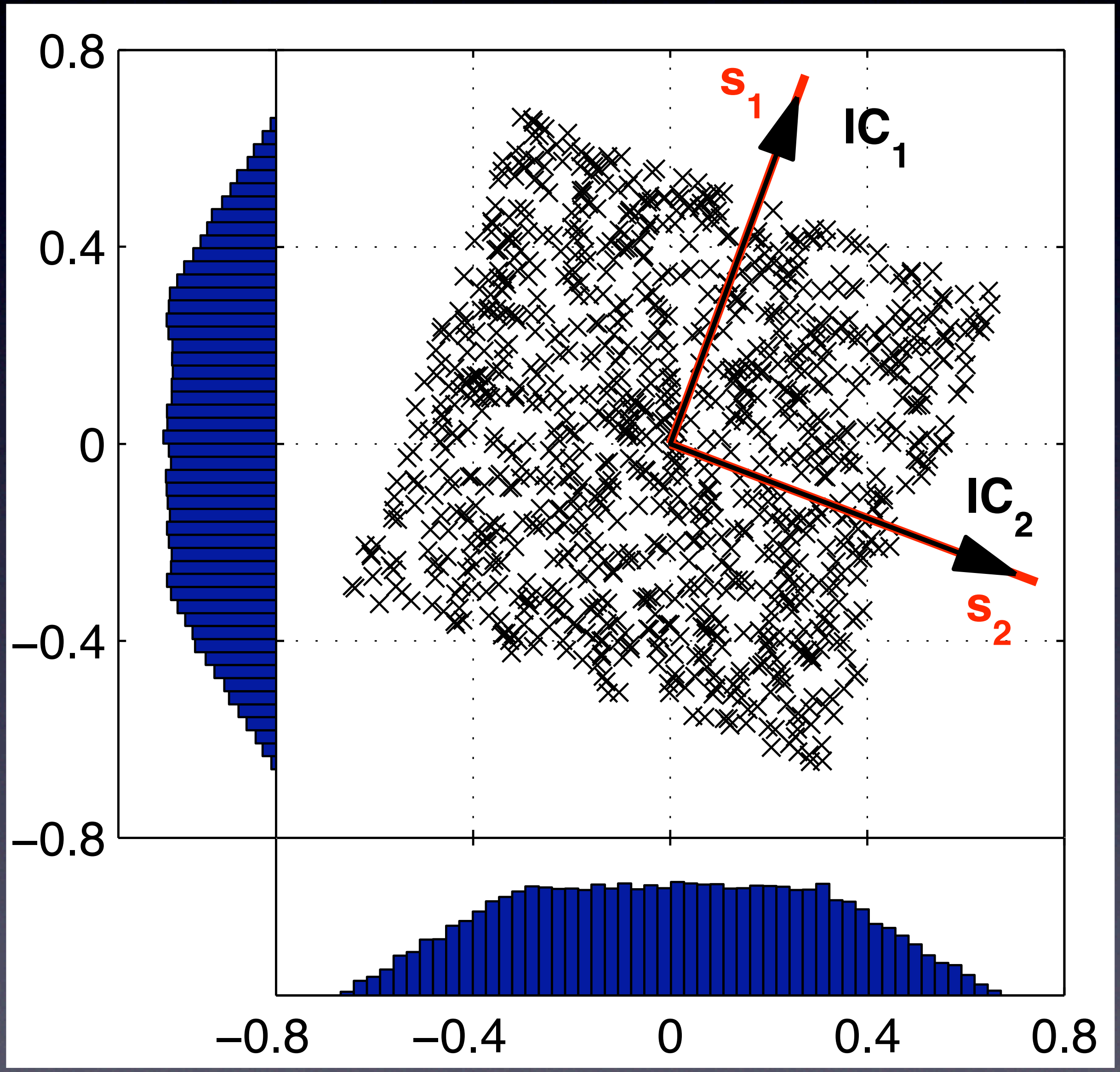




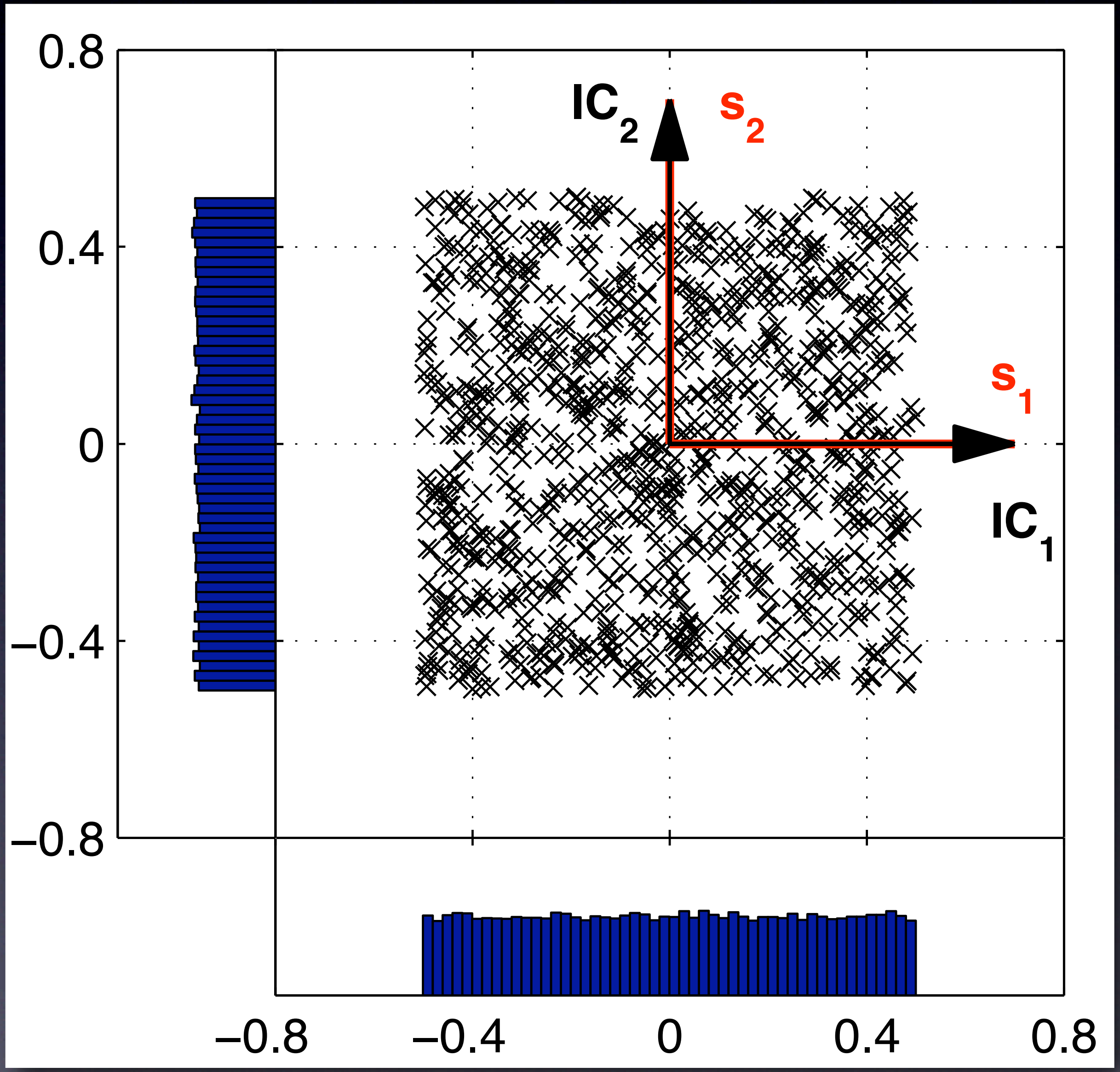
2 independent, uniformly distributed sources

linear mixtures of sources

The Geometry of PCA and ICA



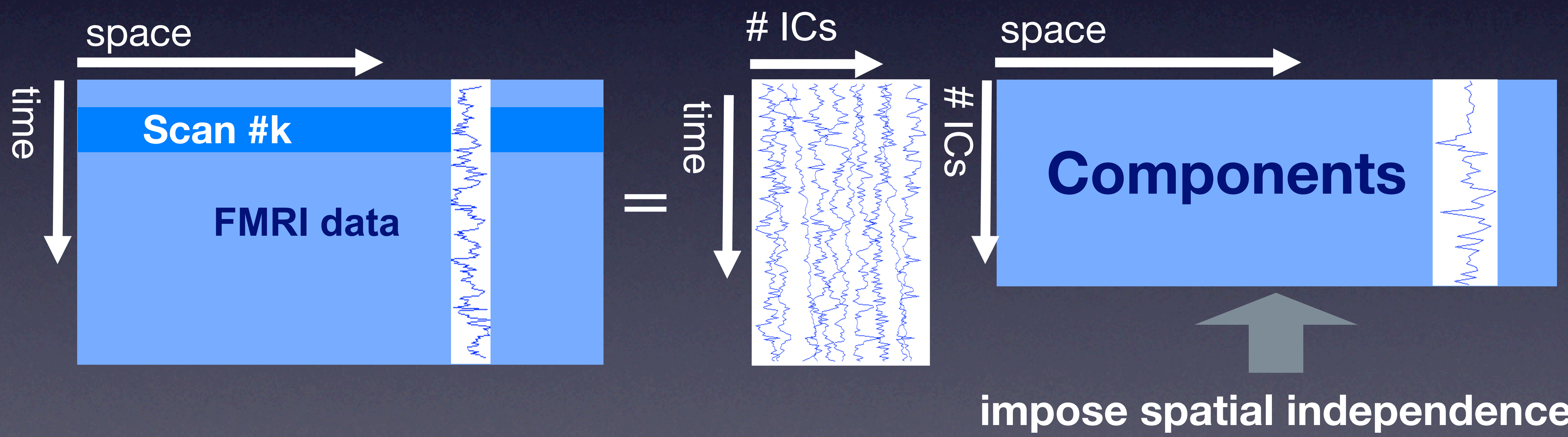
PCA solution



ICA solution

spatial ICA for FMRI

- McKeown et al. (1998): data is decomposed into a set of spatially independent component maps and a set of time-course



McKeown et al.
HBM 1998

PCA pre-processing for ICA

PCA pre-processing for ICA

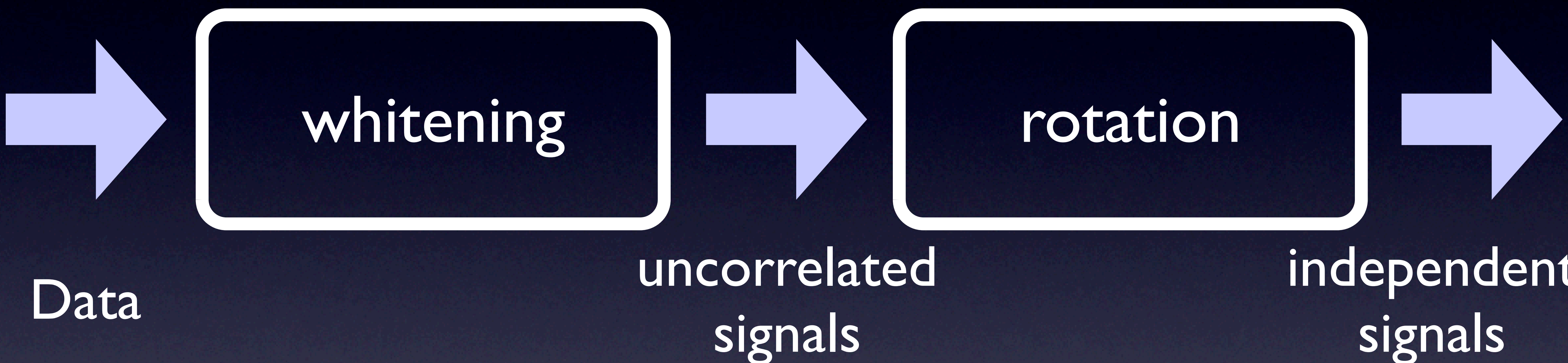
Data

uncorrelated signals

2-stage approach to ICA:

- PCA to obtain whitened (uncorrelated) signals using 2nd order statistics

PCA pre-processing for ICA



2-stage approach to ICA:

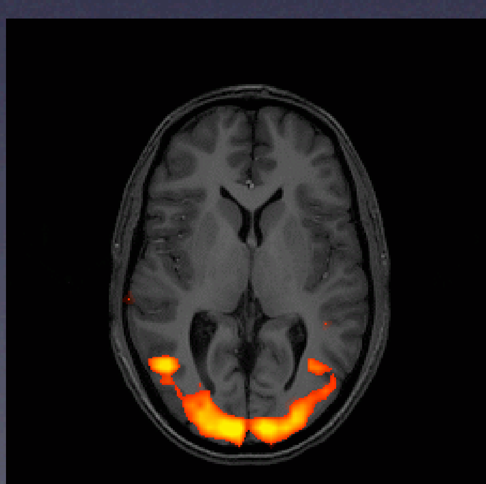
- PCA to obtain whitened (uncorrelated) signals using 2nd order statistics
- estimate rotation which minimises higher-order dependencies

ICA - the ‘overfitting’ problem

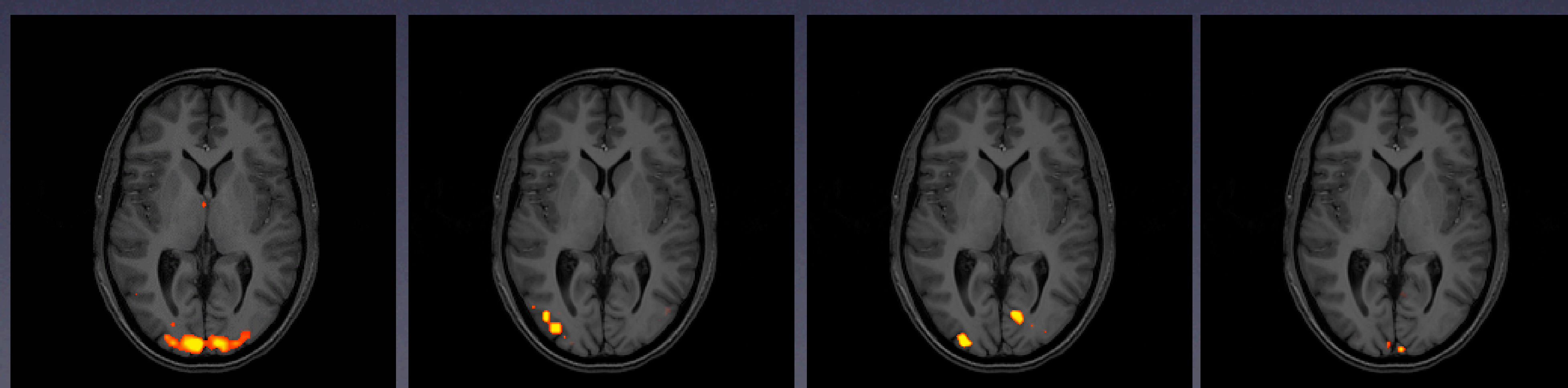
- fitting a noise-free model to noisy observations:
  - no control over signal vs. noise (non-interpretable results)
  - statistical significance testing not possible

ICA - 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

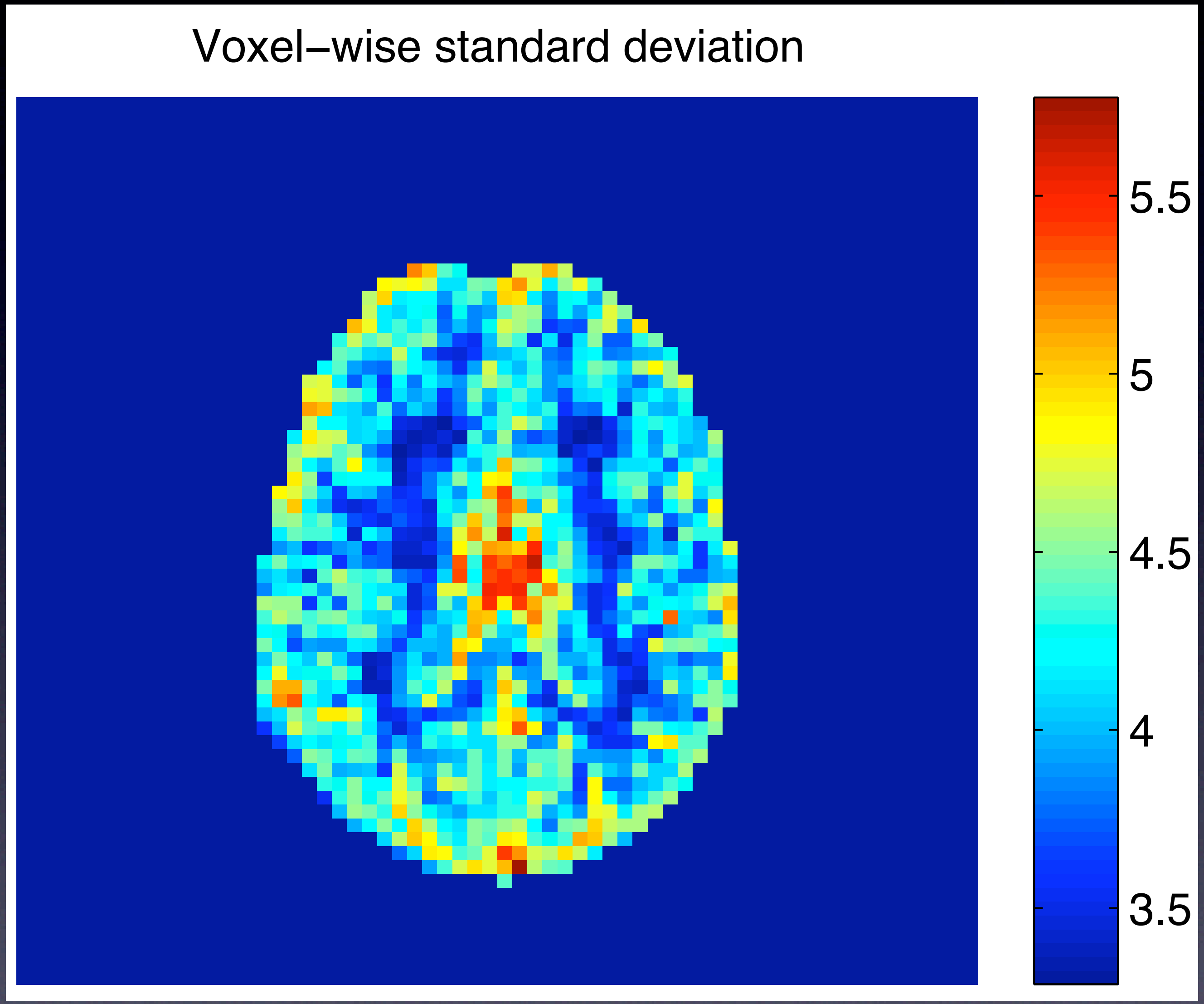
- statistical "latent variables" model: we observe linear mixtures of hidden sources

$$y_i = X\beta_i + \epsilon_i \xrightarrow{\text{for all i}} Y = XB + E$$

- we can estimate the model order from the Eigenspectrum of the data covariance matrix
- need to account for differences in voxel-wise variance (can use a pre-whitening approach (e.g. Woolrich et al., NeuroImage 14(6) 2001)

Variance-normalisation

- we might choose to ignore temporal auto-correlation in the EPI time-series but always need to normalise by the voxel-wise variance
- this amounts to modelling the spatial covariance matrix as a diagonal matrix

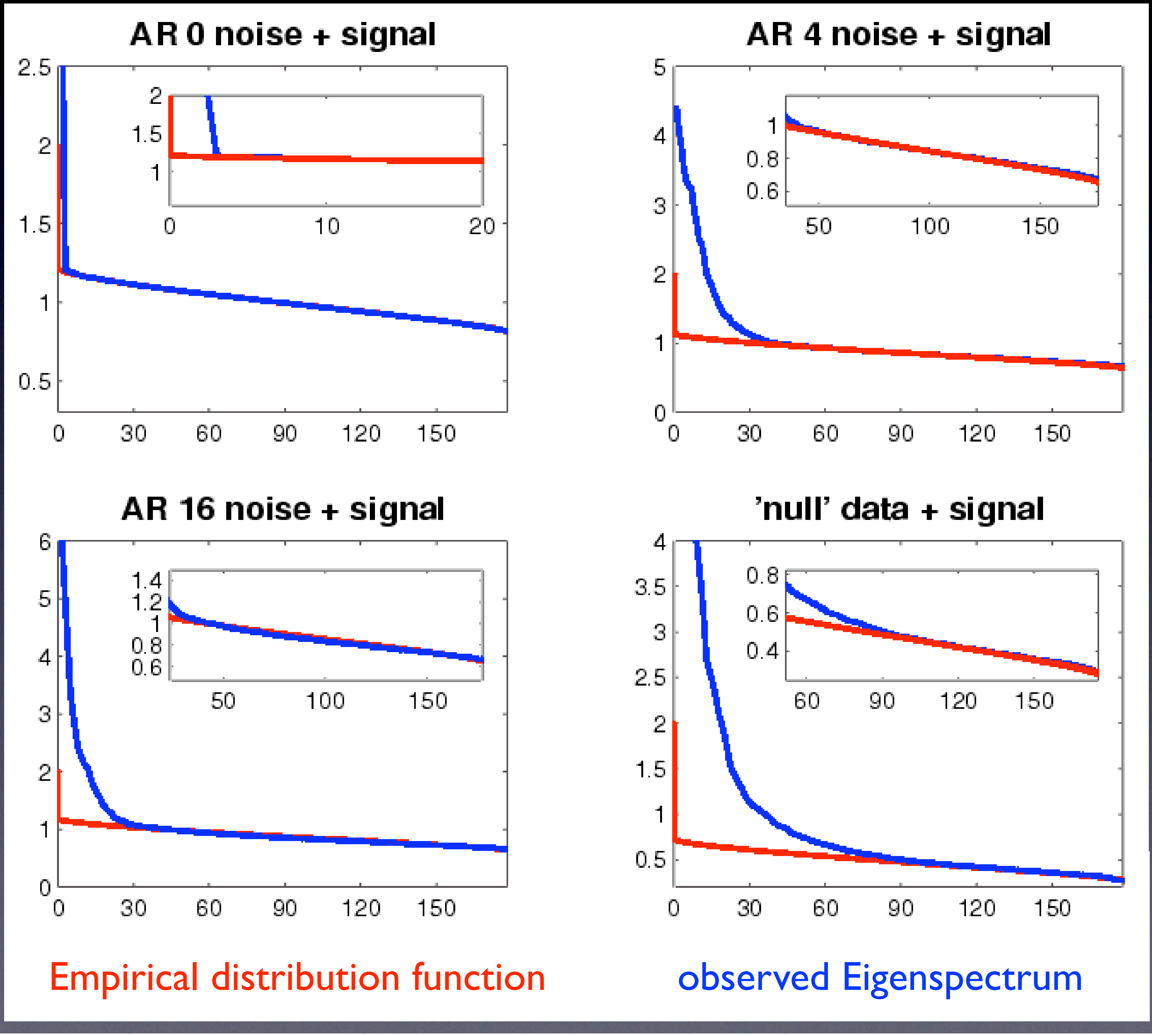


Estimated voxel-wise std. deviation (log-scale) from FMRI data obtained under rest condition.

$$\mathbf{V}^{-1/2} = \mathrm{diag}(\sigma_1, \ldots, \sigma_N)$$

Model Order Selection

- The sample covariance matrix has a Wishart distribution and we can calculate the empirical distribution function for the eigenvalues

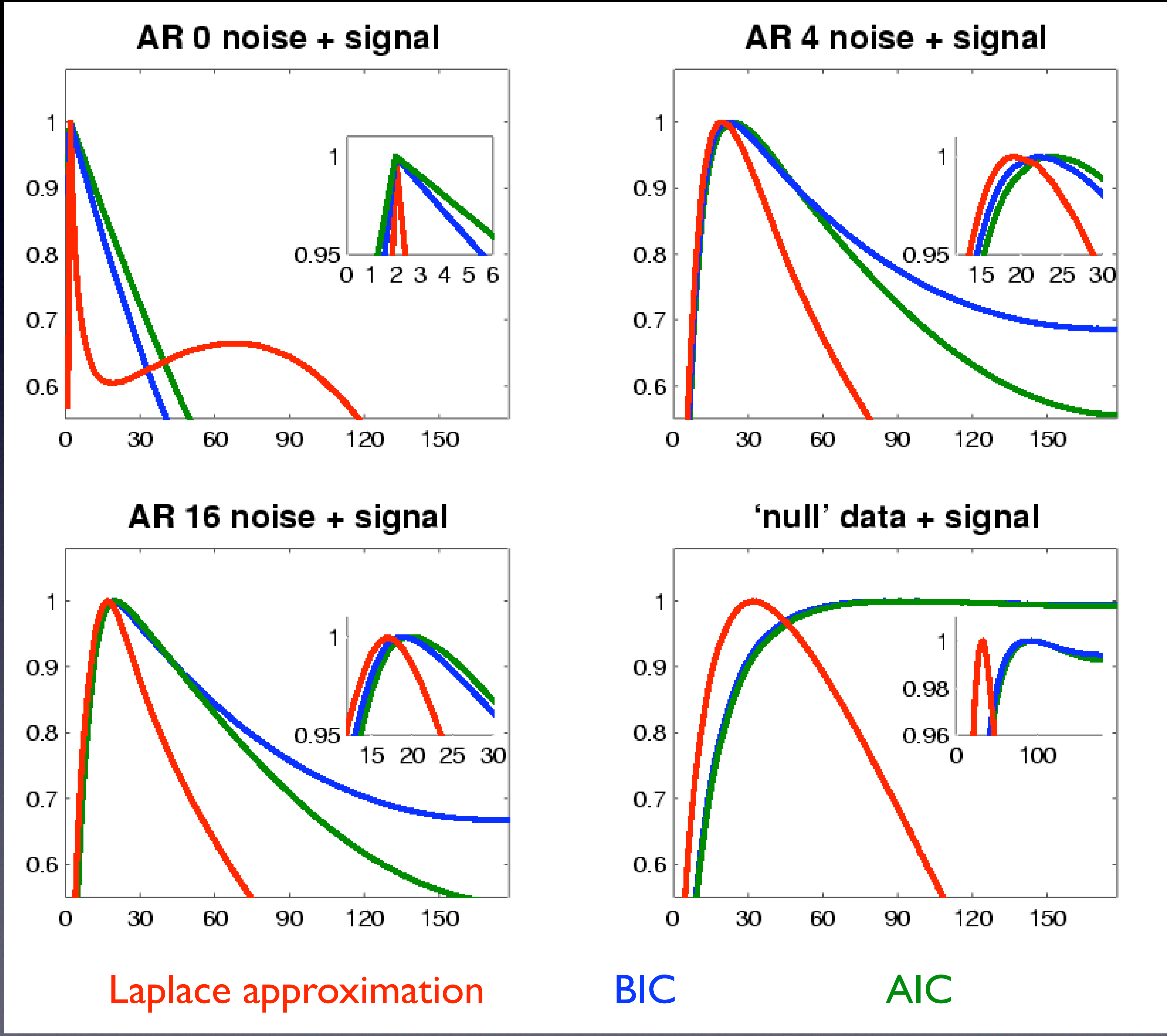


Everson & Roberts, IEEE Trans. Sig. Proc. 48(7), 2000

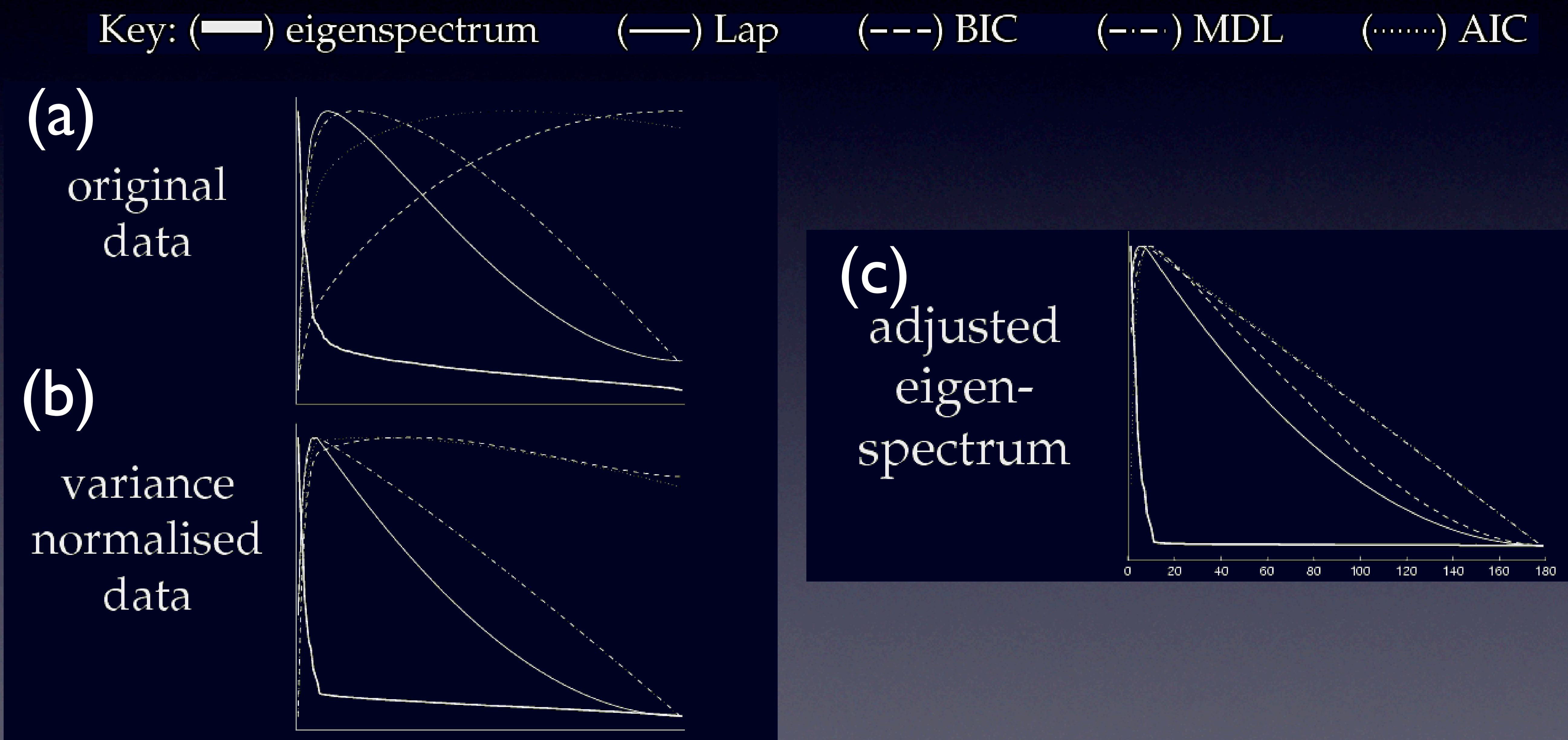
Model Order Selection (PPCA)

- use a probabilistic PCA model and calculate (approximate) the Bayesian evidence for the model order

Minka, TR 514 MIT Media Lab 2000



Conditioning the Eigenspectrum

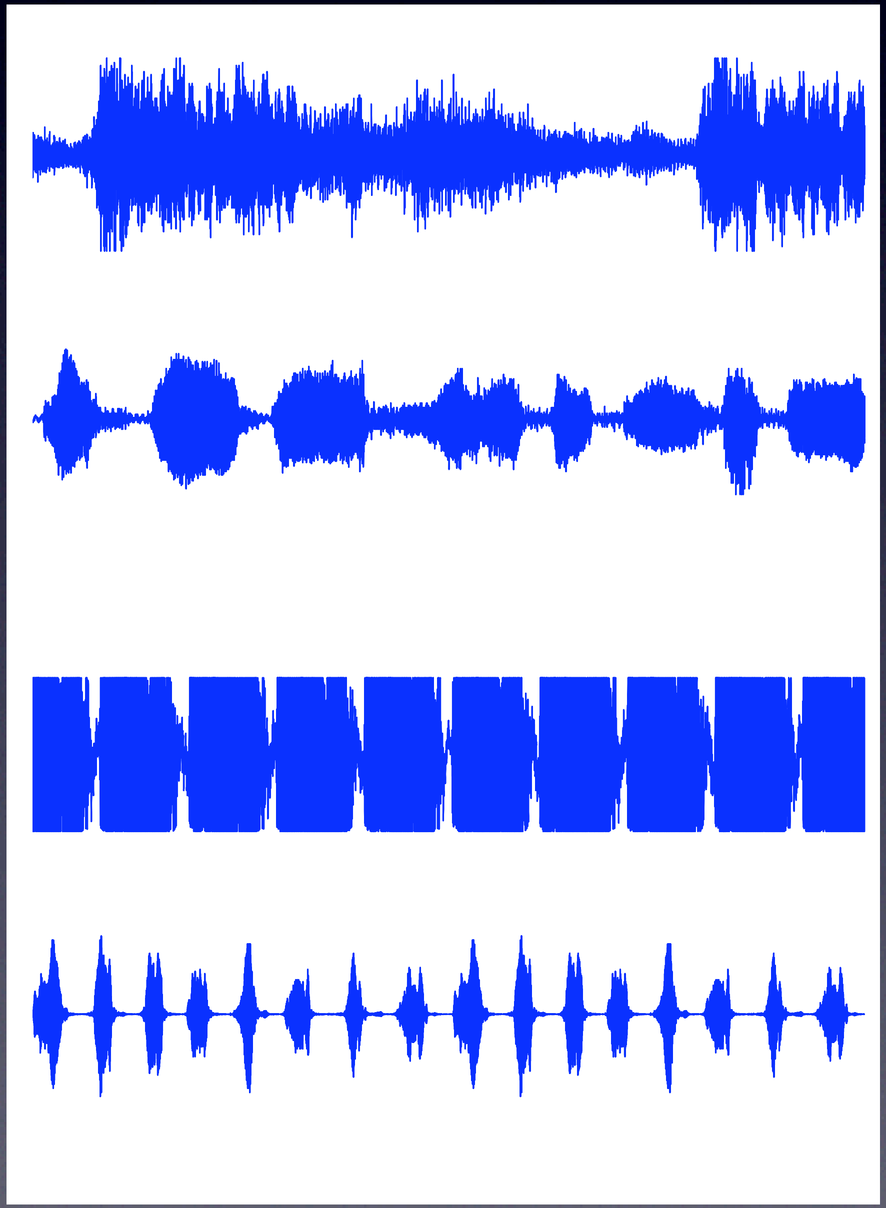


10 sources introduced into white noise (AR 0)

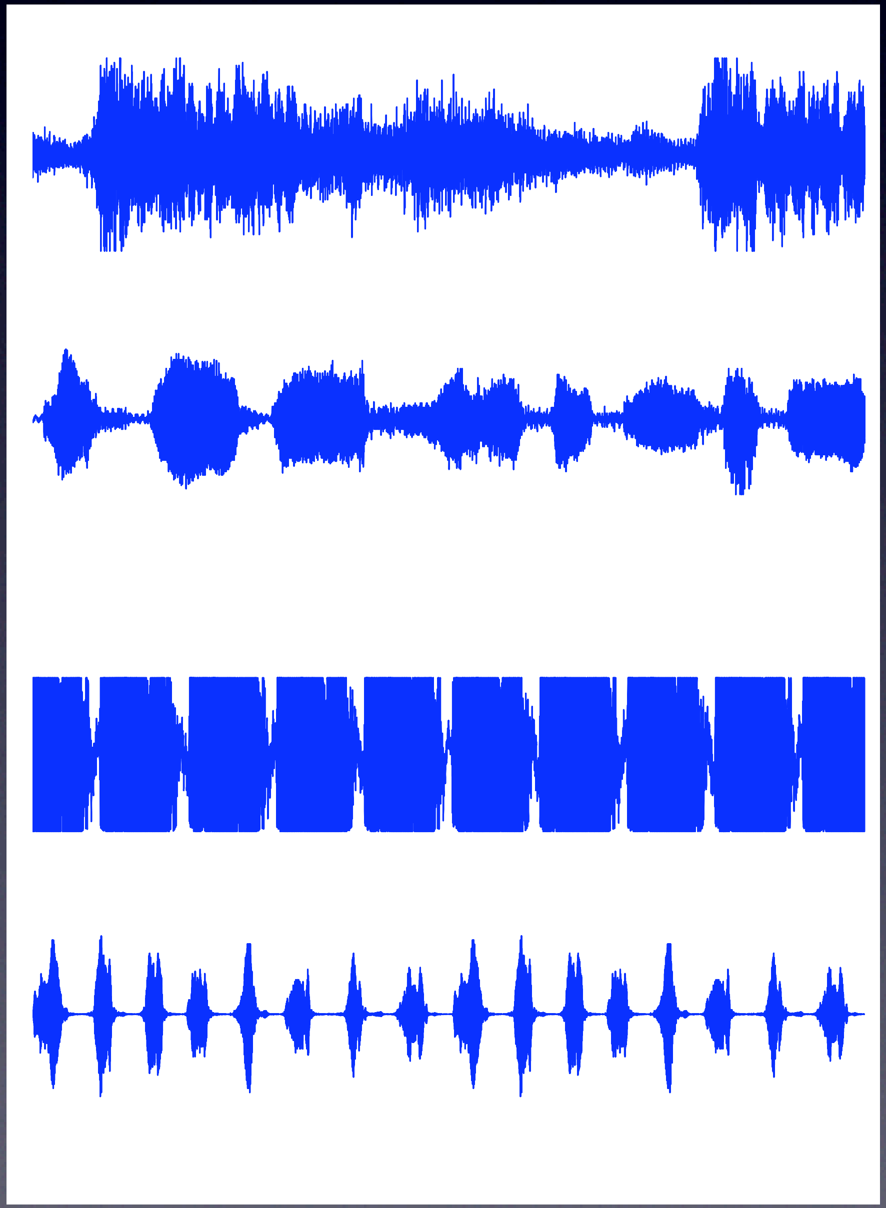
ICA estimation

- need to find an ‘unmixing matrix’ such that the dependency between estimated sources is minimised
- need (i) a *contrast (objective/cost) function* 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)

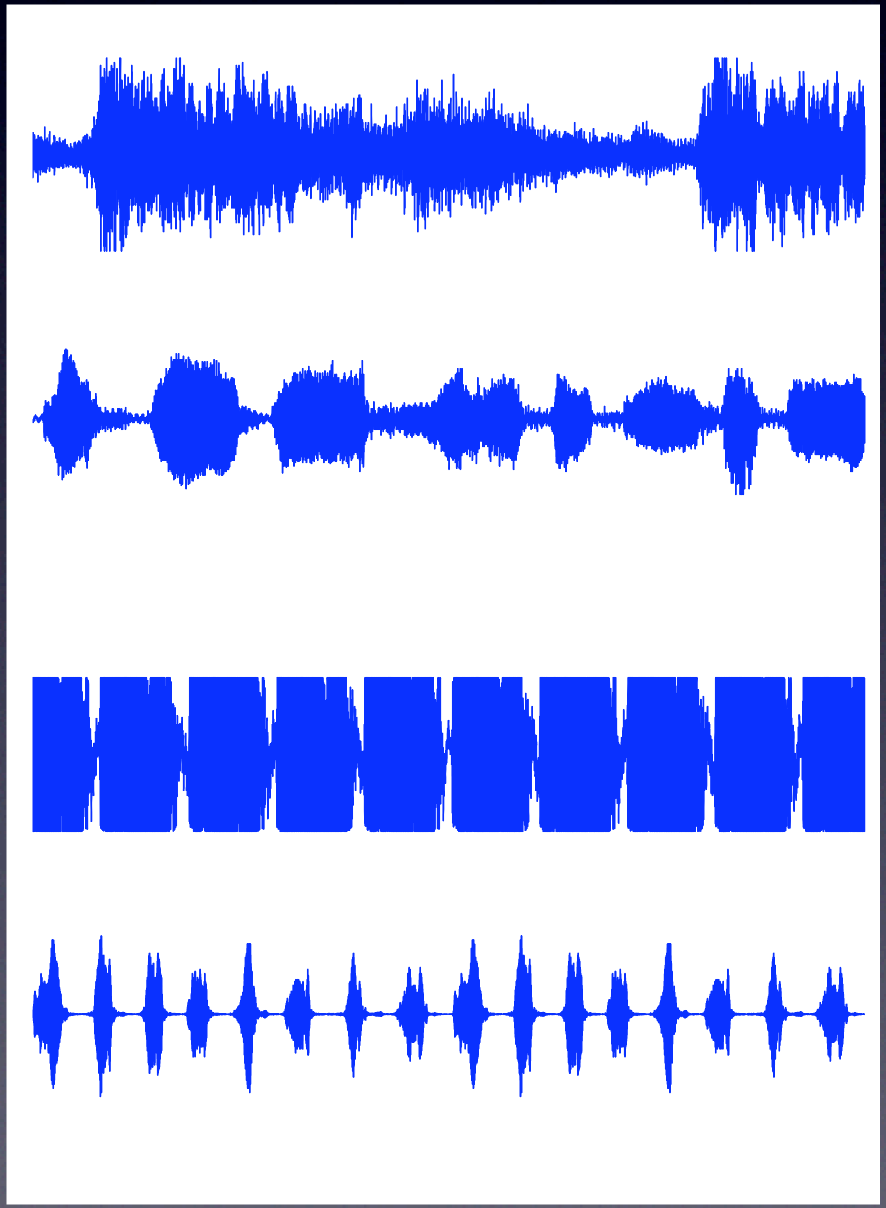
Non-Gaussianity is 'interesting'



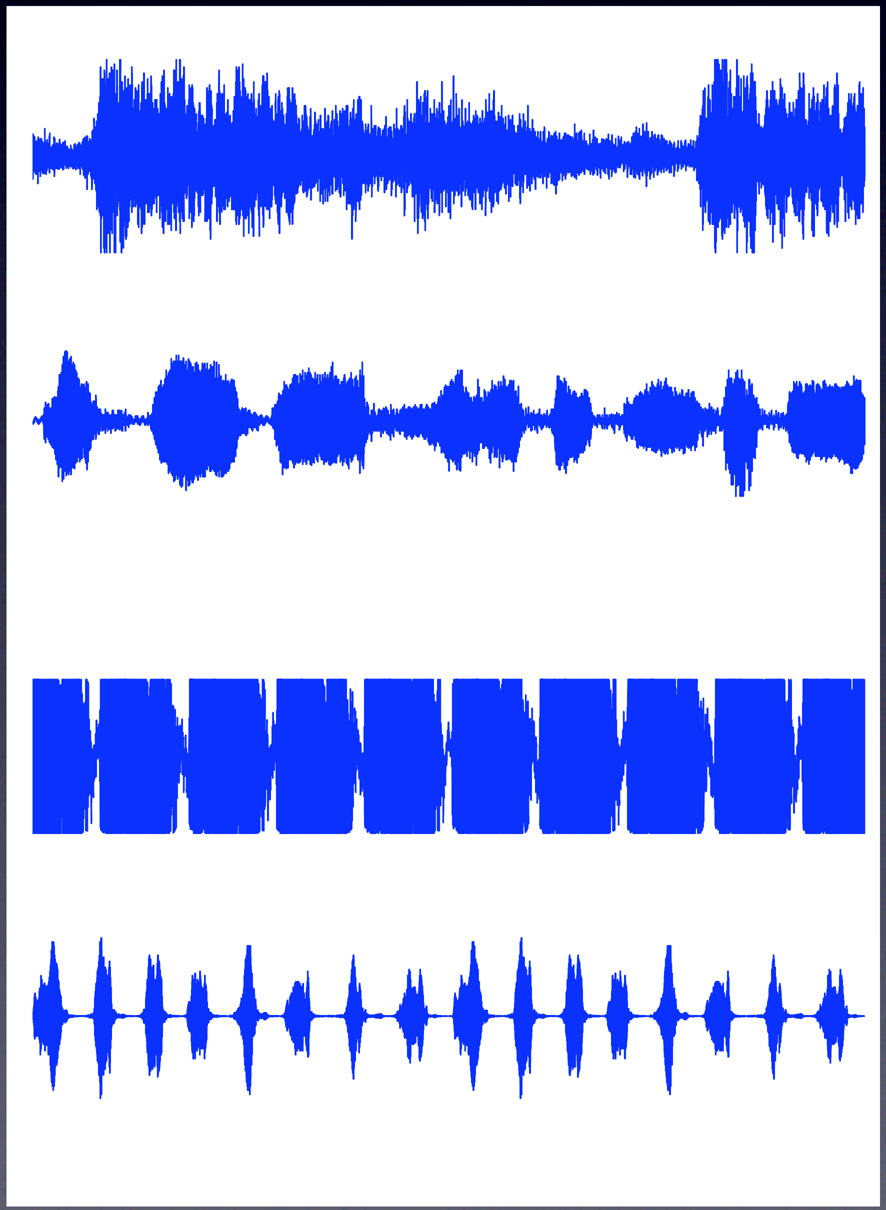
Non-Gaussianity is ‘interesting’



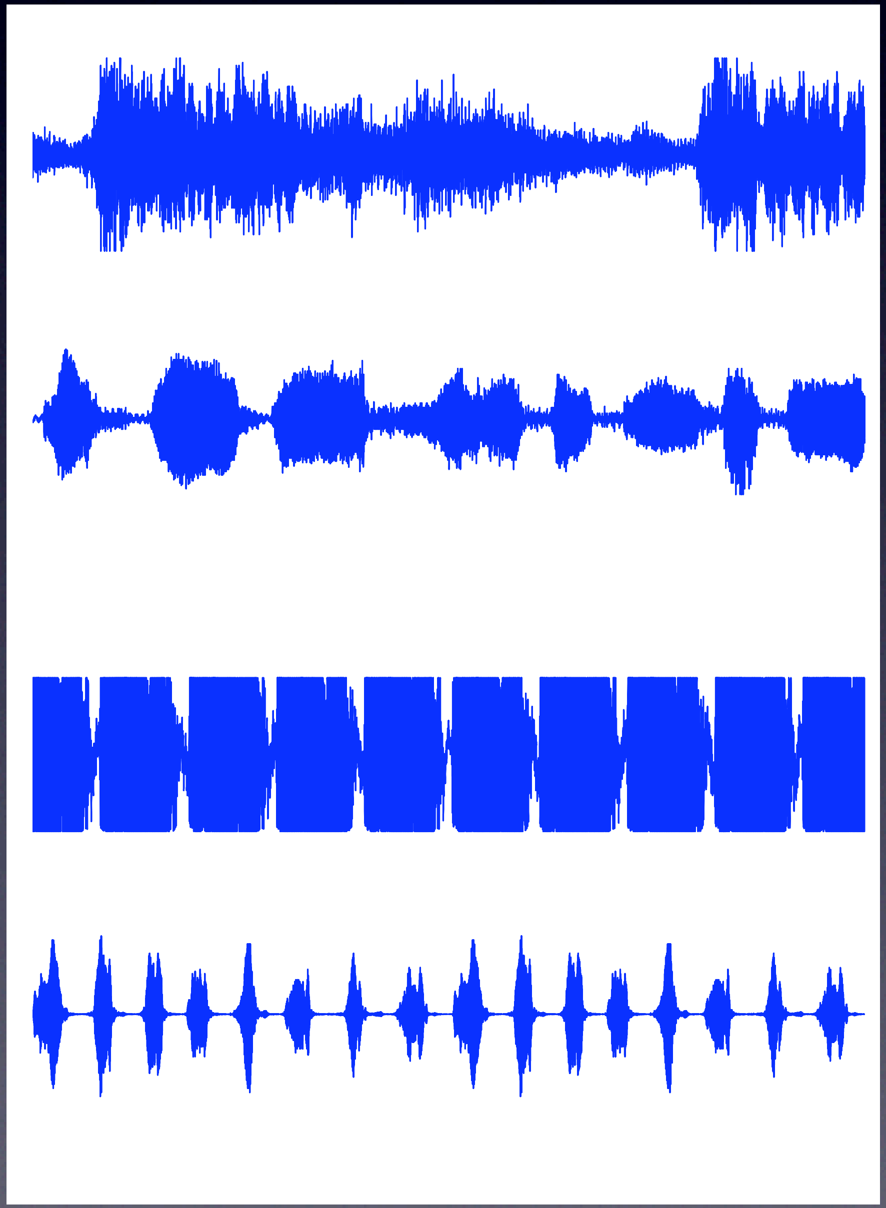
Non-Gaussianity is 'interesting'



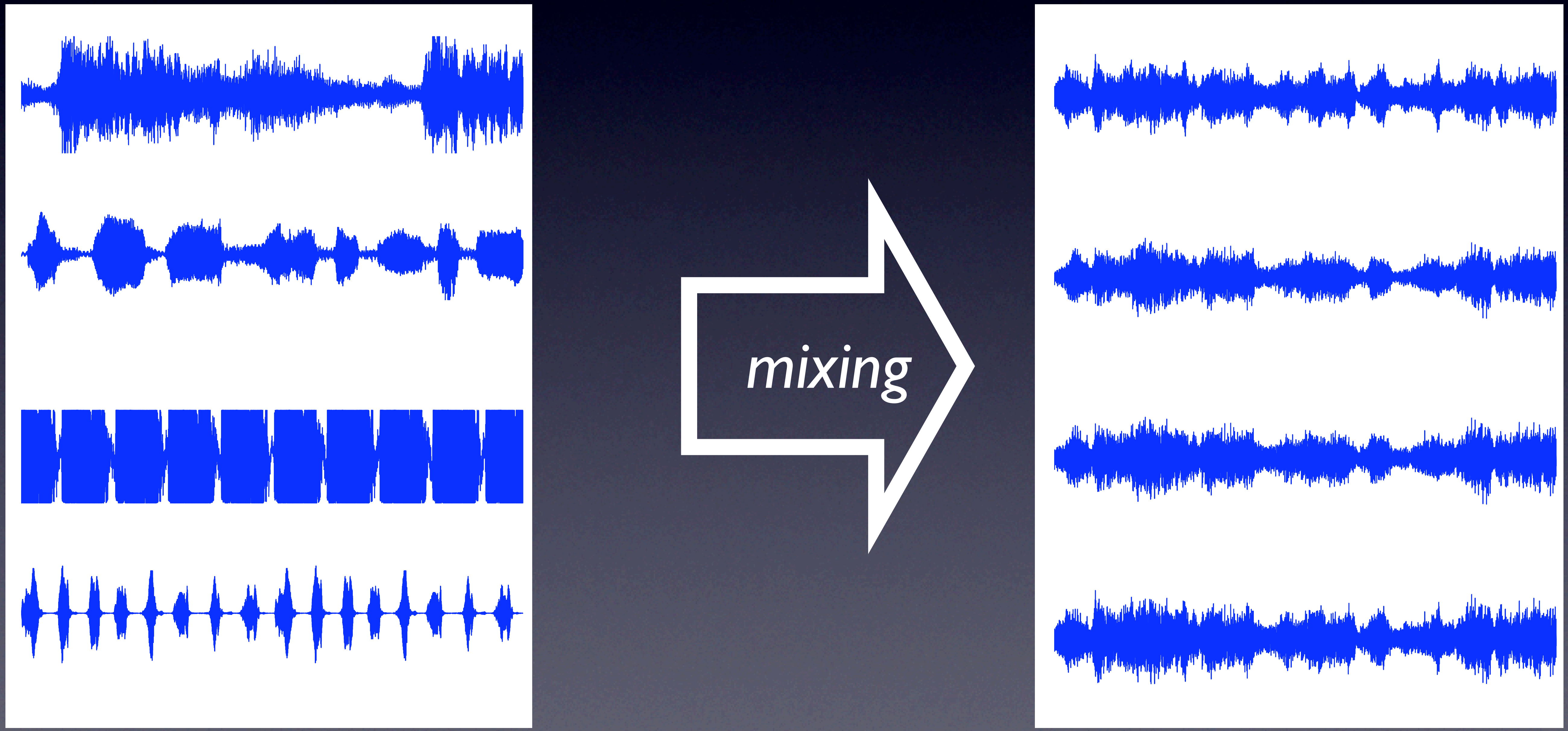
Non-Gaussianity is 'interesting'



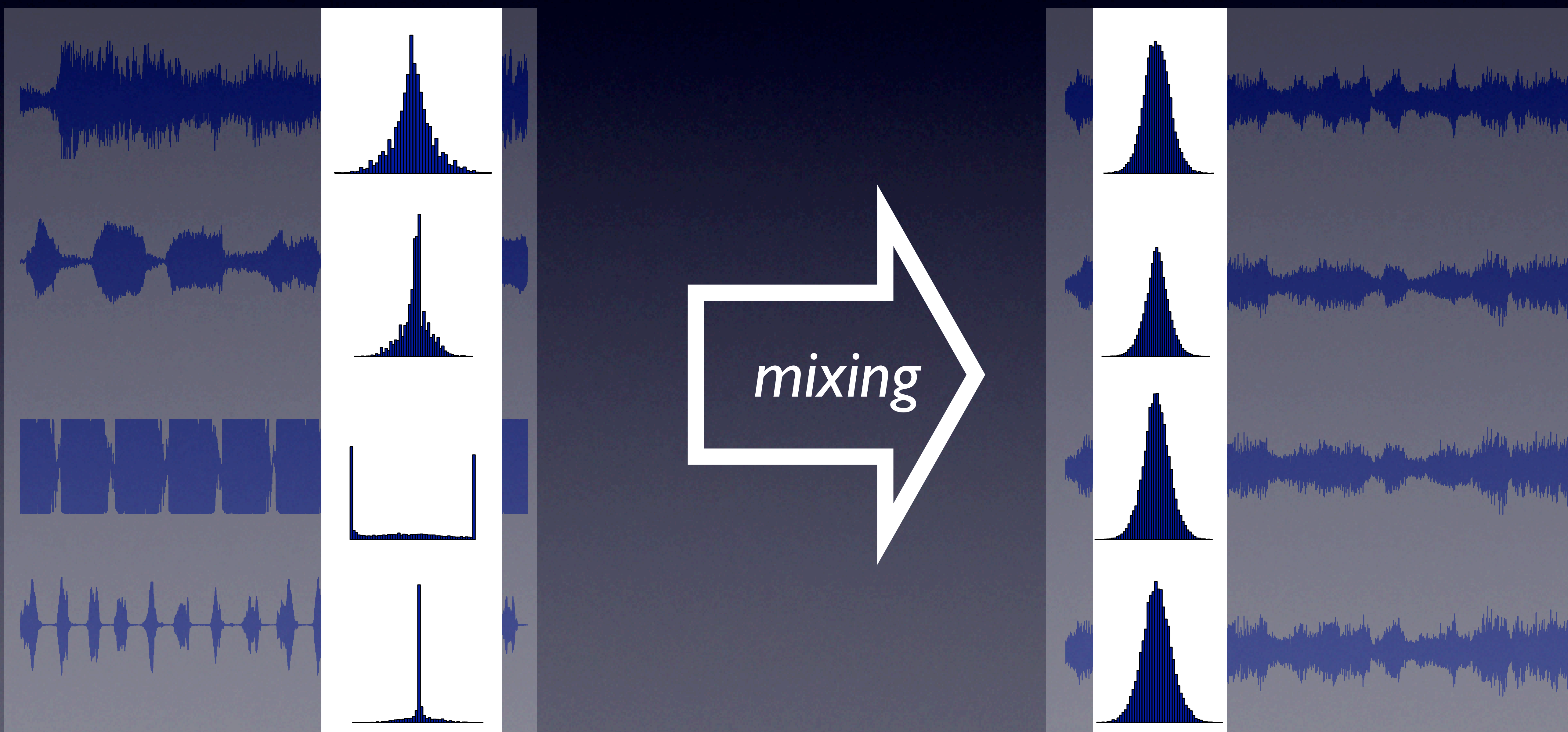
Non-Gaussianity is ‘interesting’



Non-Gaussianity is ‘interesting’



Non-Gaussianity is ‘interesting’



ICA estimation

ICA estimation

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

ICA estimation

- Random mixing results in more Gaussian shapes PDFs (Central Limit Theorem)
- conversely: if an unmixing matrix produces less Gaussian shaped PDFs this is unlikely to be a random result

ICA estimation

- Random mixing results in more Gaussian shapes PDFs (Central Limit Theorem)
- conversely: if an unmixing matrix produces less Gaussian shaped PDFs this is unlikely to be a random result
    - measure non-Gaussianity

ICA estimation

- Random mixing results in more Gaussian shapes PDFs (Central Limit Theorem)
- conversely: if an unmixing matrix produces less Gaussian shaped PDFs this is unlikely to be a random result
    - measure non-Gaussianity
- can use neg-entropy $\mathcal{J}(\mathrm{s}) = \mathcal{H}(\mathrm{s}_{\mathrm{gauss}}) - \mathcal{H}(\mathrm{s})$

ICA estimation

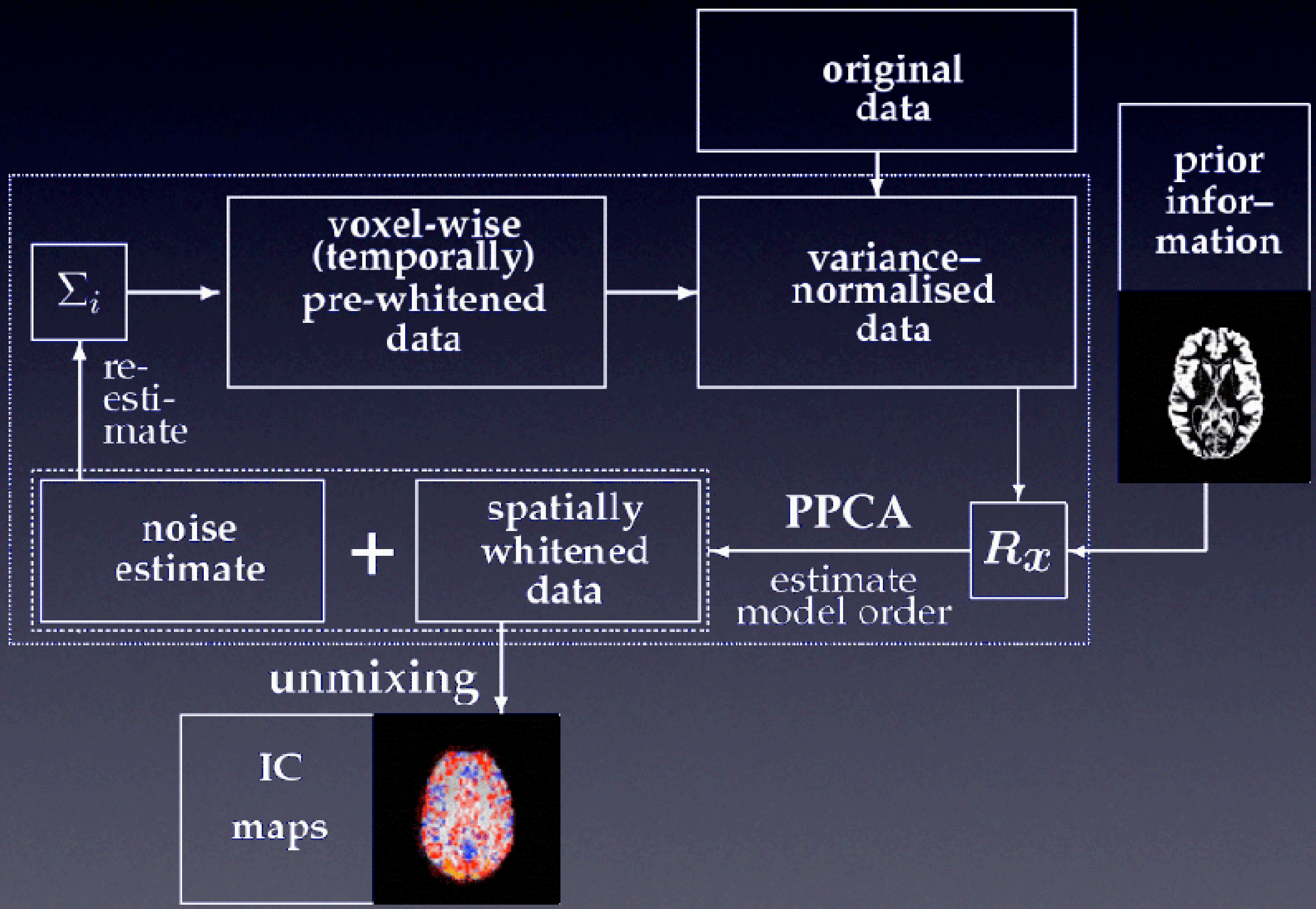
- Random mixing results in more Gaussian shapes PDFs (Central Limit Theorem)
- conversely: if an unmixing matrix produces less Gaussian shaped PDFs this is unlikely to be a random result
    - measure non-Gaussianity
- can use neg-entropy $\mathcal{J}(\mathrm{s}) = \mathcal{H}(\mathrm{s}_{\mathrm{gauss}}) - \mathcal{H}(\mathrm{s})$
- Fast approximations:

$$\mathcal{J}(\mathrm{s}) \simeq \sum_{i}^{p} \kappa_i \left[\mathsf{E}(g_i(\mathrm{s})) - \mathsf{E}(g_i(\mathrm{s}_{\mathrm{gauss}})) \right]^2$$

Hyvärinen & Oja 1997

Probabilistic ICA (I)

- Data pre-processing: voxel wise de-meaning and variance normalization
- estimate intrinsic dimensionality (PCA)
- unmix source signals using negentropy as a measure of non-Gaussianity (Hyvärinen, 1999)



ICA and spatially correlated sources

- Spatial correlation:

$$\rho(s_1, s_2) = \frac{s_1^t s_2}{N\sqrt{\mathrm{Var}(s_1)}\sqrt{\mathrm{Var}(s_2)}}$$

- In the presence of noise:

$$\rho(s_1 + \eta_1, s_2 + \eta_2) = \frac{s_1^t s_2}{N\sqrt{\mathrm{Var}(s_1) + \sigma^2}\sqrt{\mathrm{Var}(s_2) + \sigma^2}}$$

ICA and spatially correlated sources

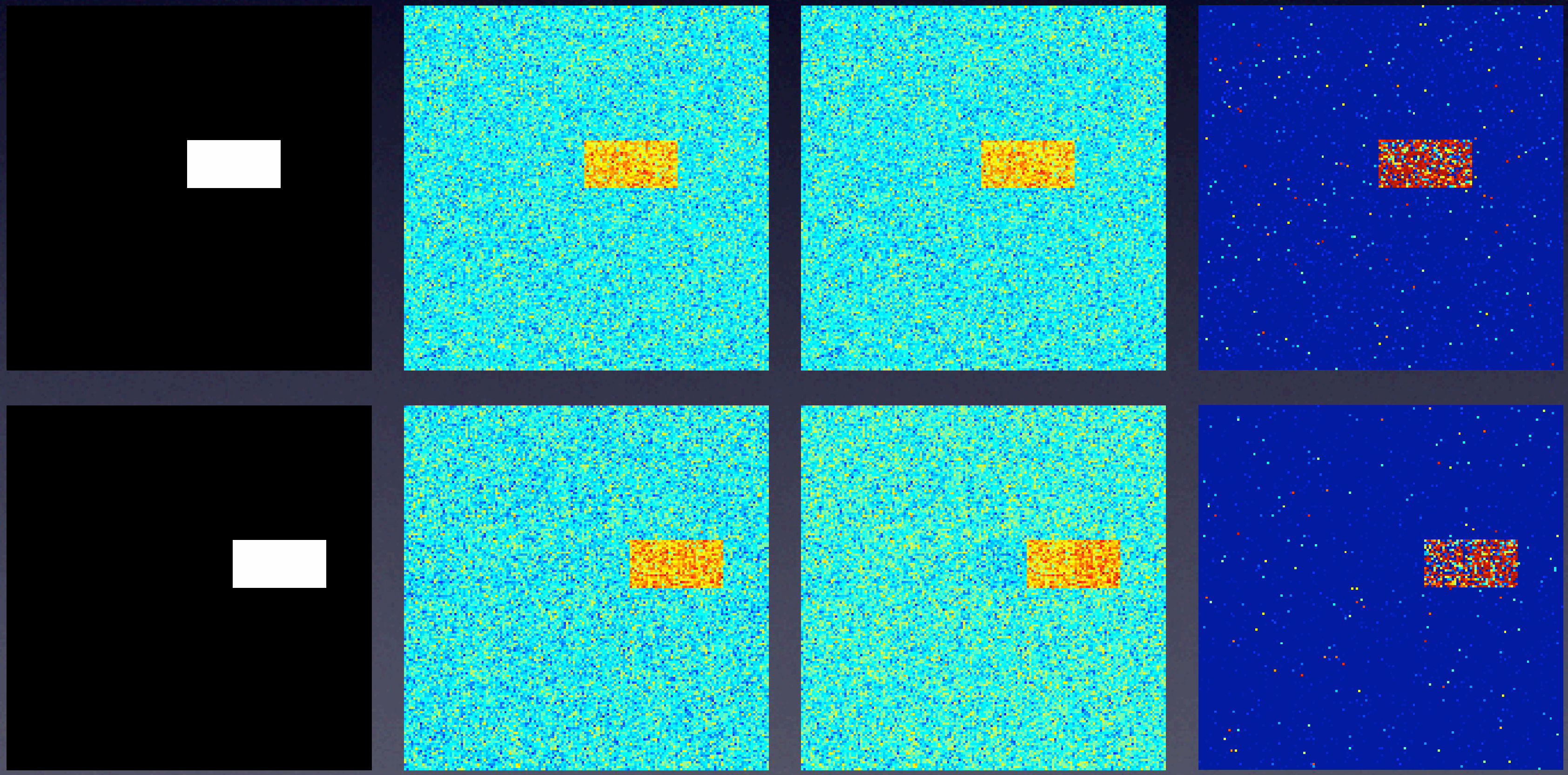
Example:

(a) 2 sources correlated (ρ=0.5);

(b) in the presence of noise, ρ<0.1 and

(c) de-correlating the two maps preserves the structure;

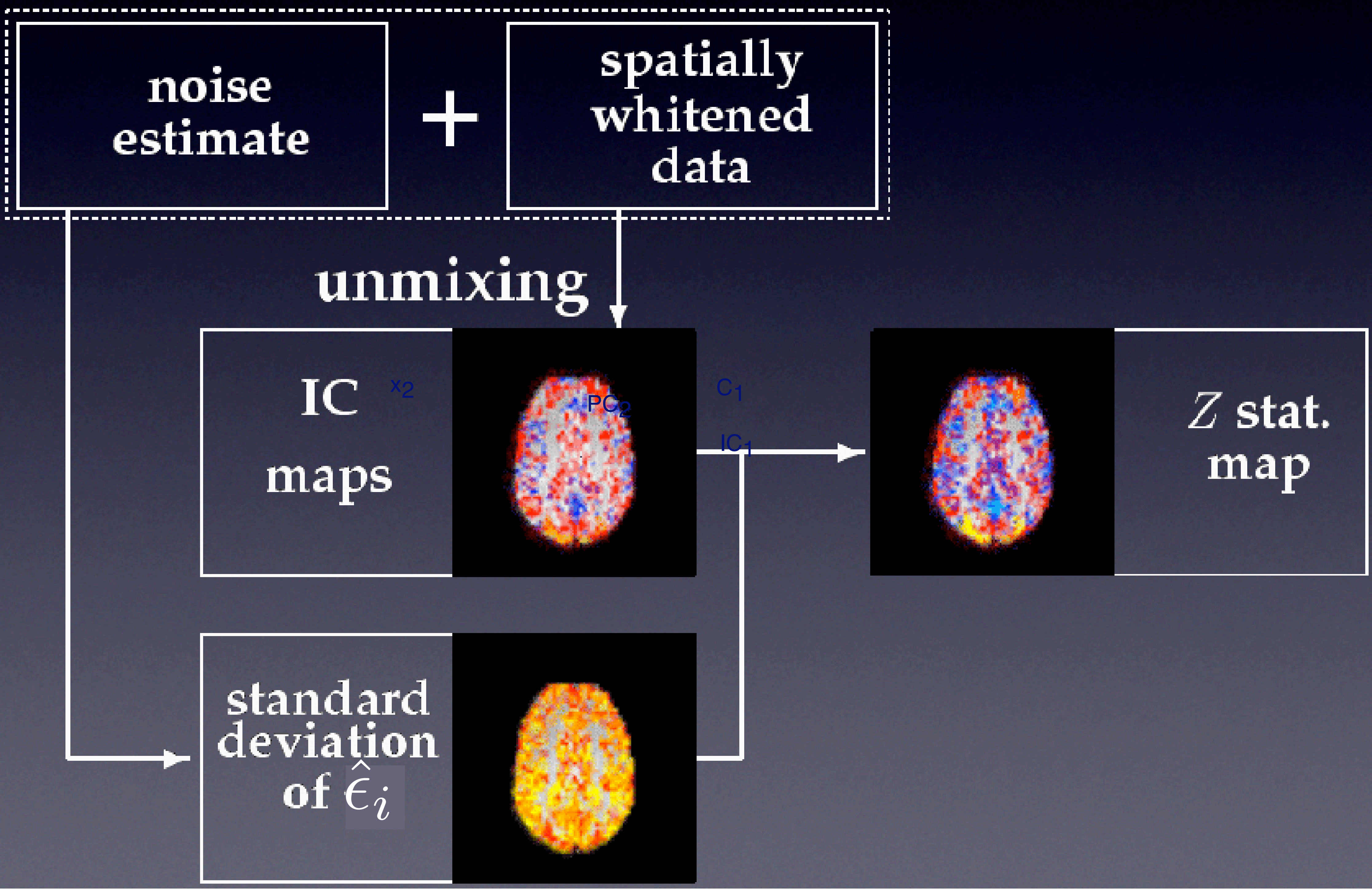
(d) thresholded maps represent the sources well



(a) (b) (c) (d)

Probabilistic ICA (II)

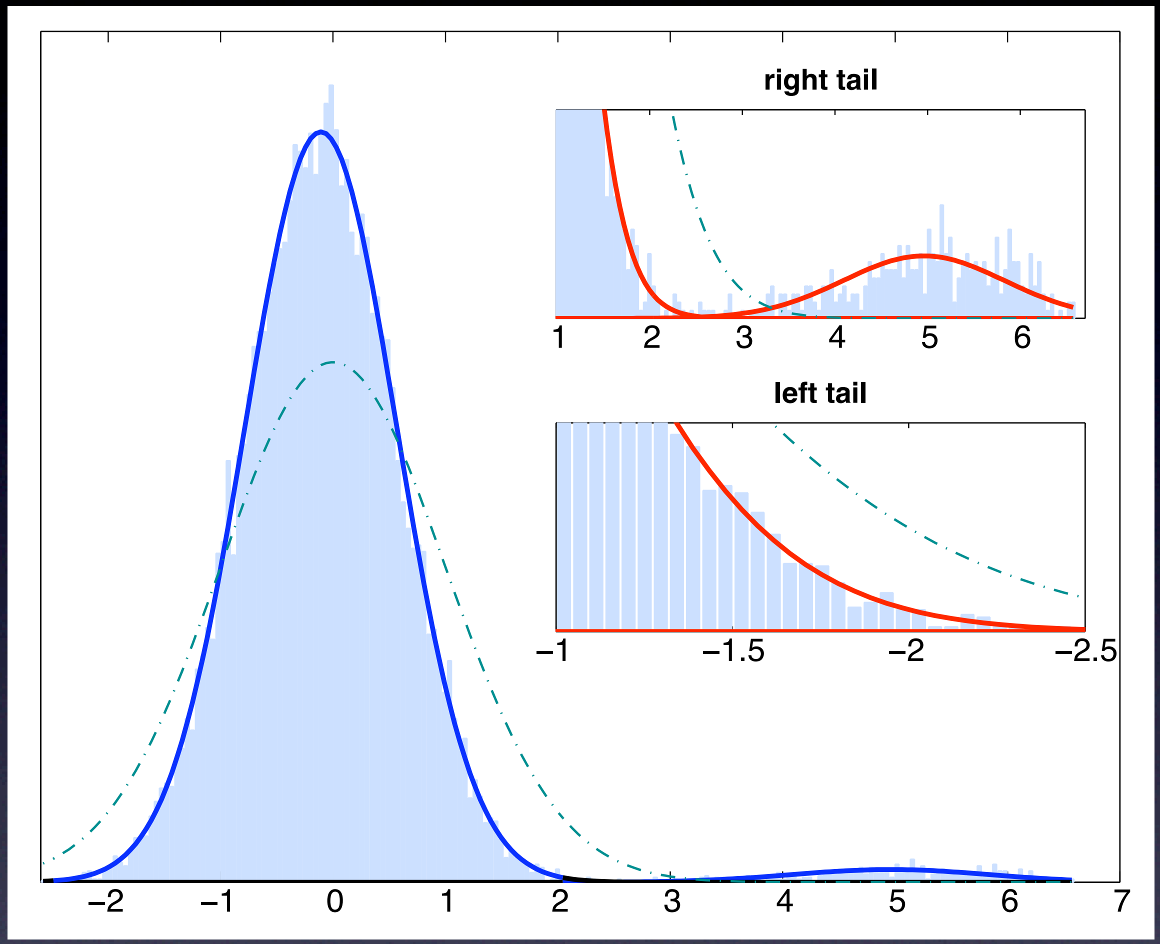
- form voxel-wise Z-statistics using the estimated standard deviation of the noise



Thresholding

- Null-hypothesis testing is invalid:

  IC-map thresholding based on Z-transforming across the spatial domain gives wrong false-positives rate!



- Under the model:

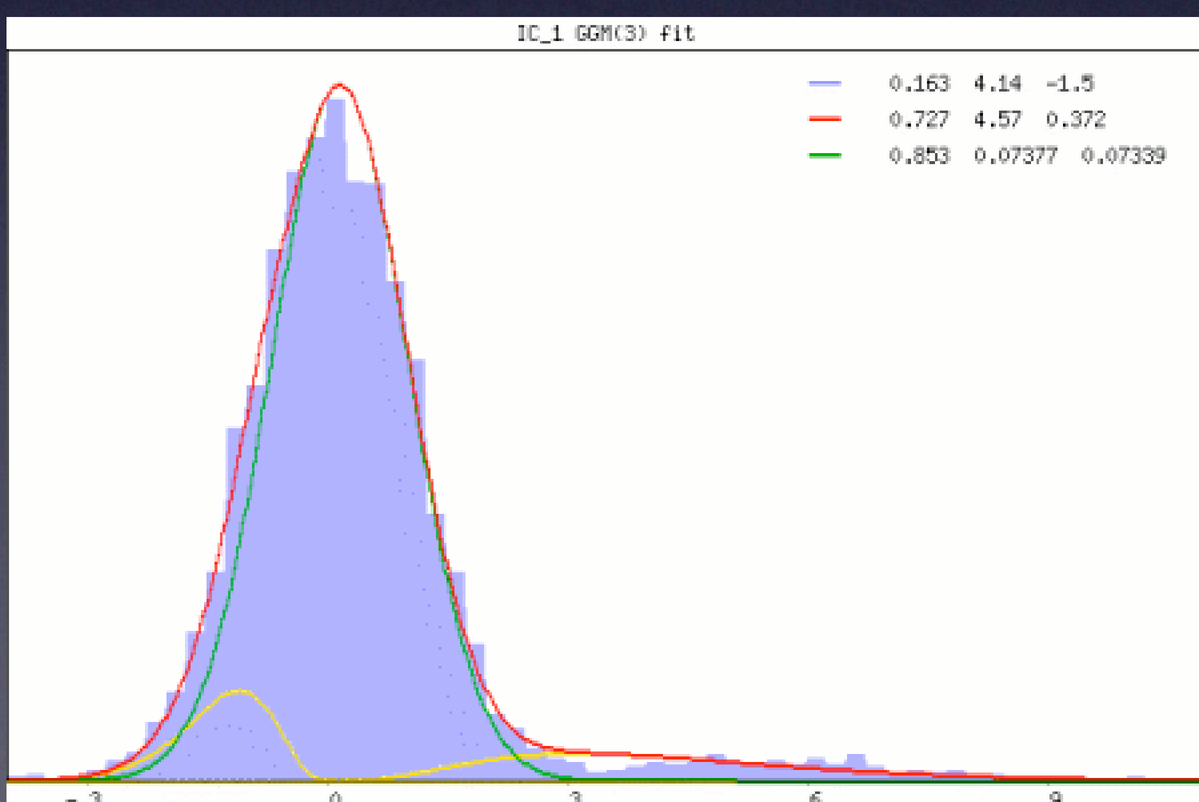
$$\widehat{B} = WY = W(XB + E)$$

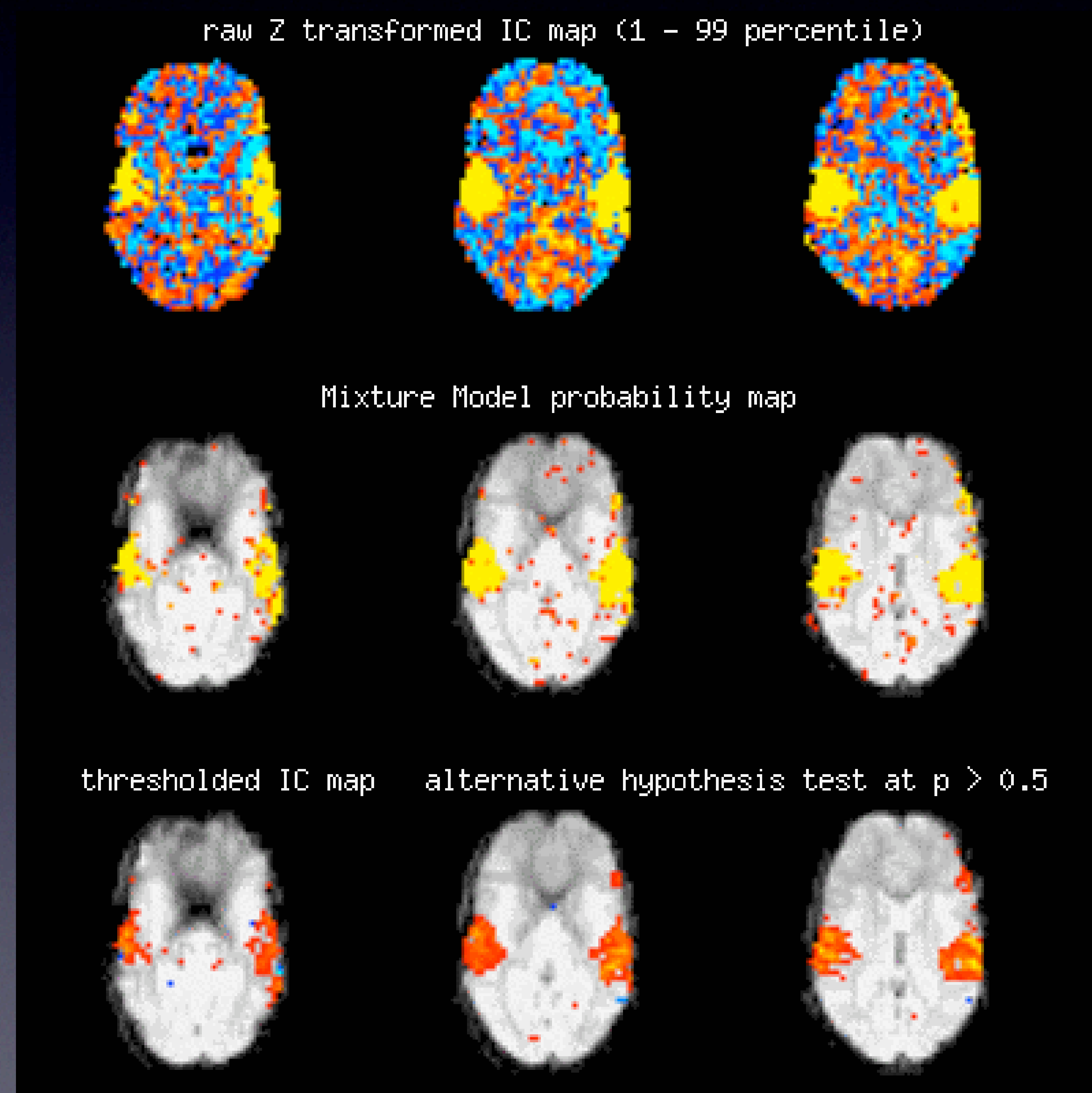
  i.e. the estimated spatial maps contain a linear projection of the noise term.

- the distribution of the estimated spatial maps is a mixture distribution.

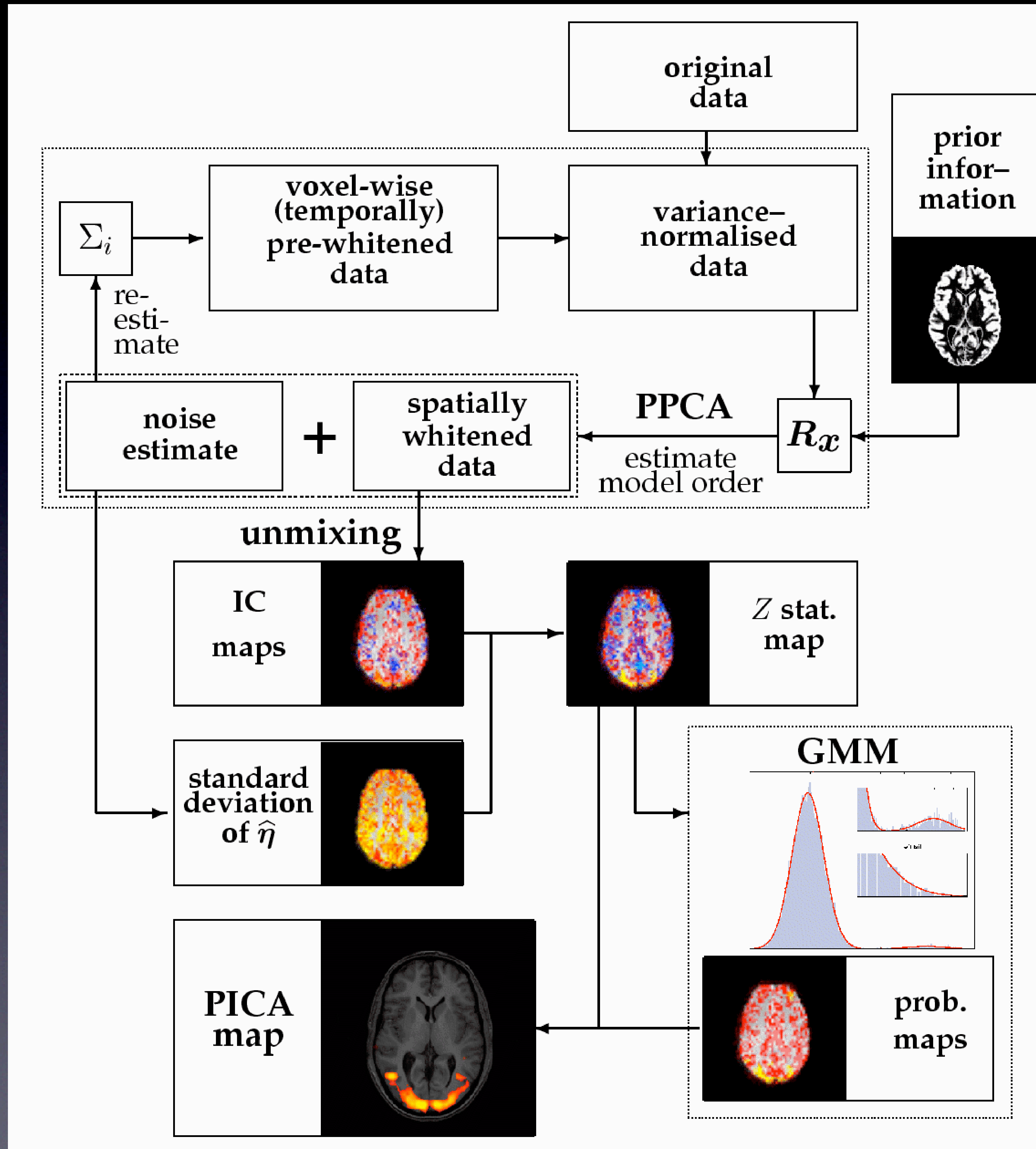
Alternative Hypothesis Test

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





Full PICA model

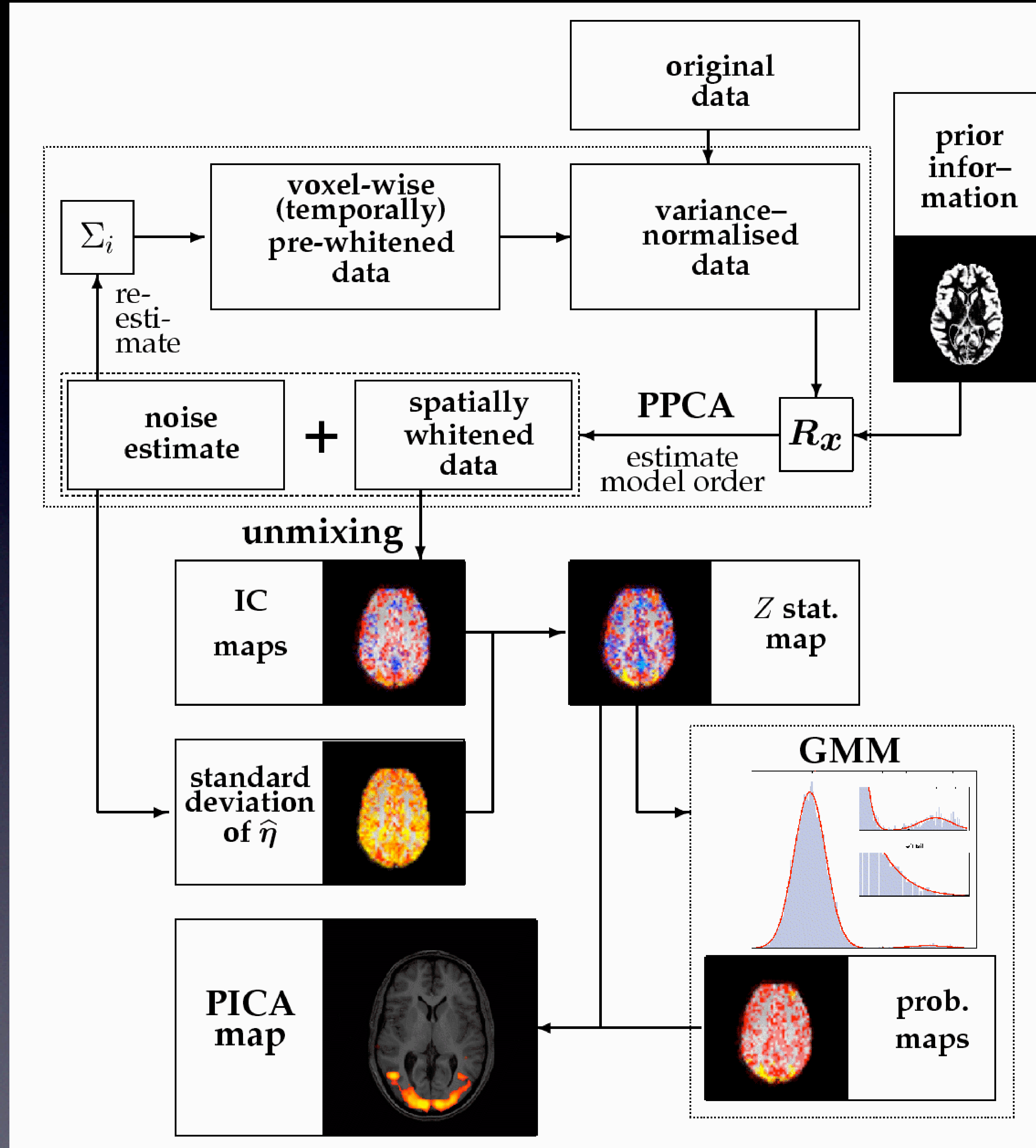


MELODIC

Full PICA model

MELODIC

Beckmann and Smith IEEE TMI 2004

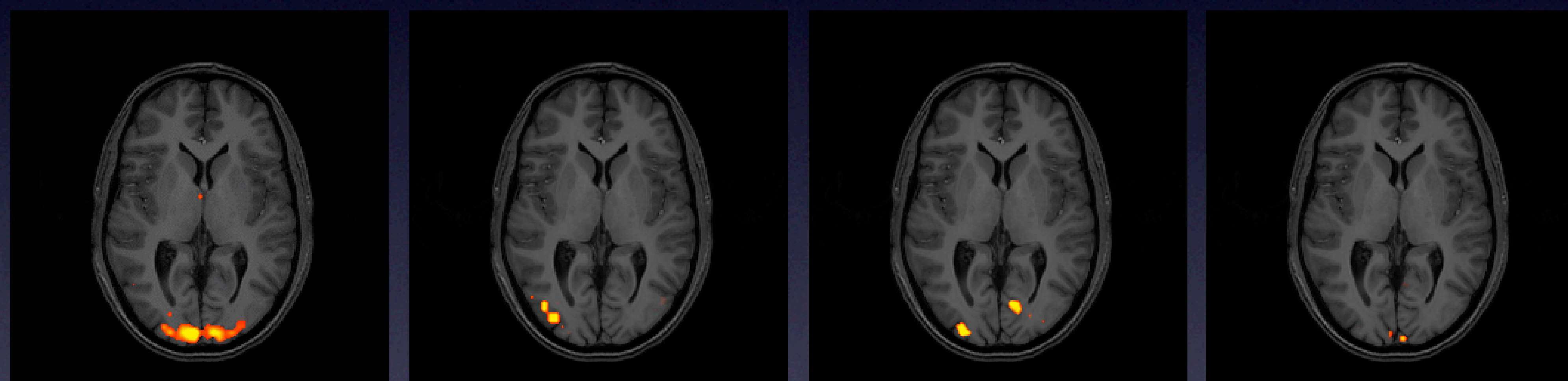


Probabilistic ICA

designed to address the ‘overfitting problem’:

GLM analysis

standard ICA (unconstrained)

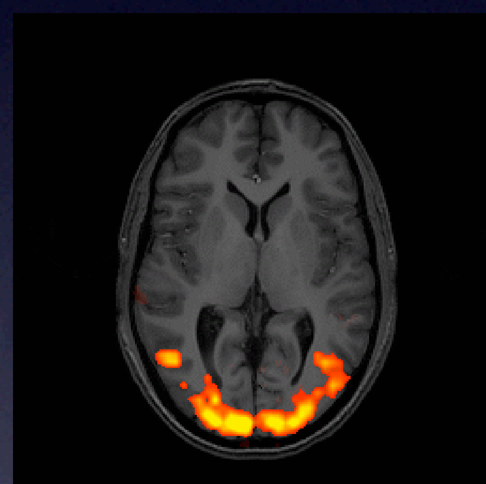


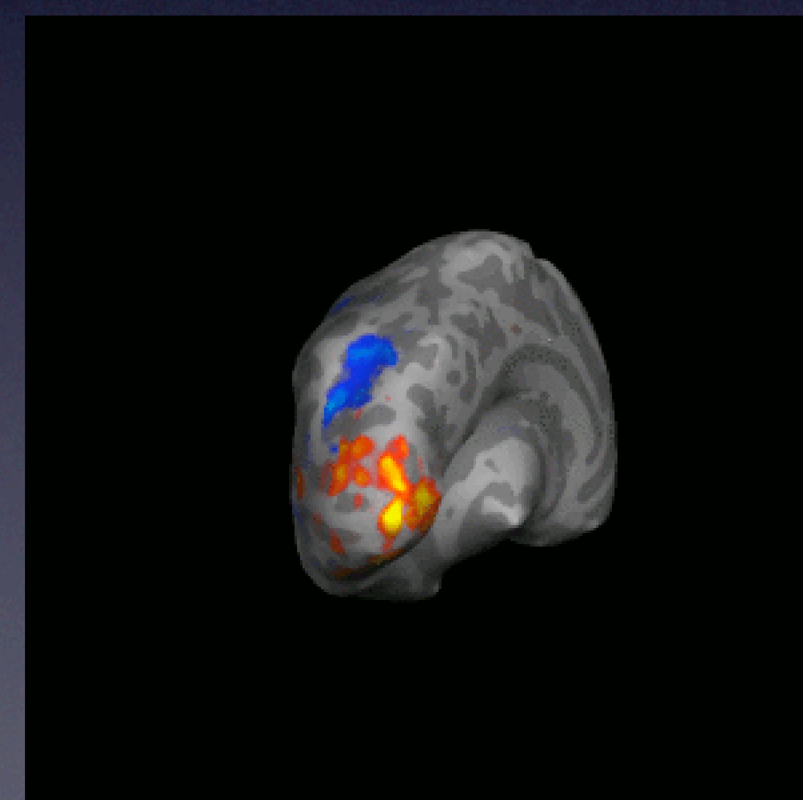
Probabilistic ICA

designed to address the 'overfitting problem':

GLM analysis

probabilistic ICA (constrained)

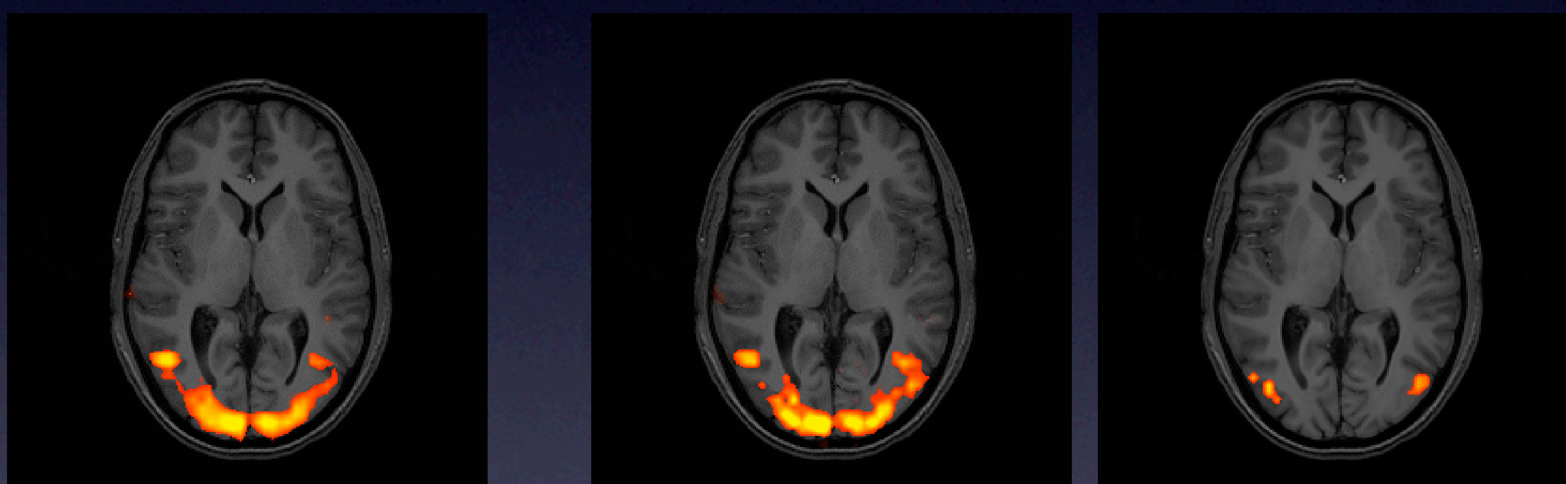


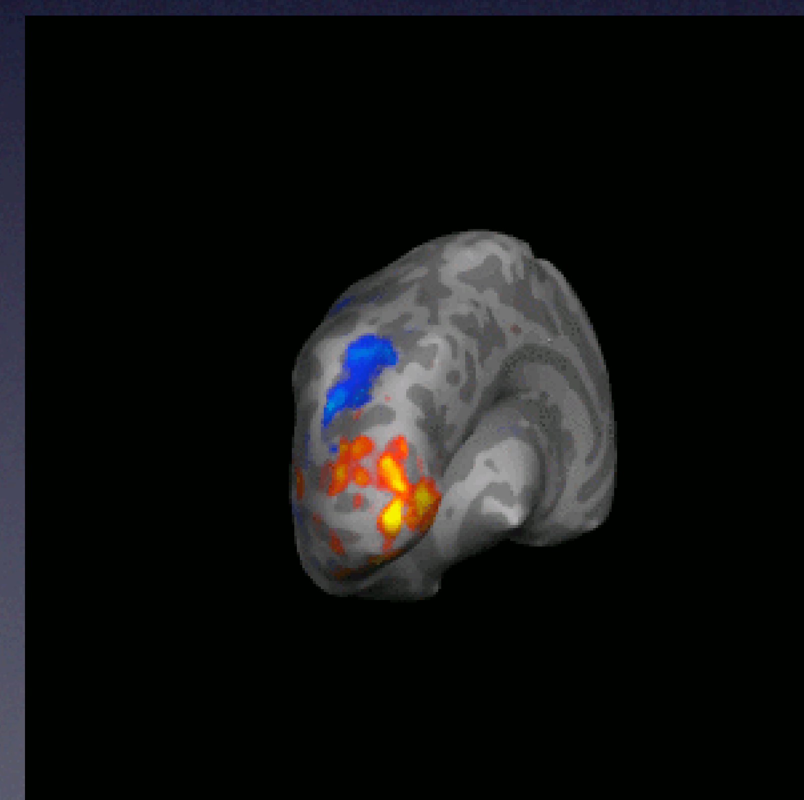


Probabilistic ICA

designed to address the 'overfitting problem':

GLM analysis **probabilistic ICA (constrained)**



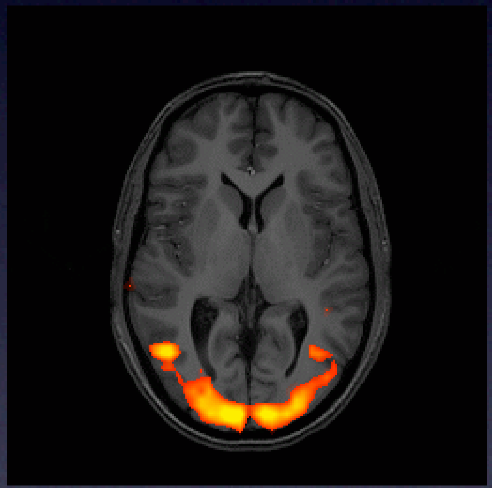


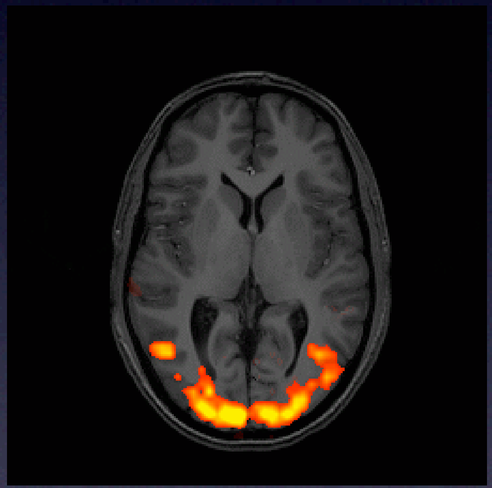
- avoids generation of 'spurious' results

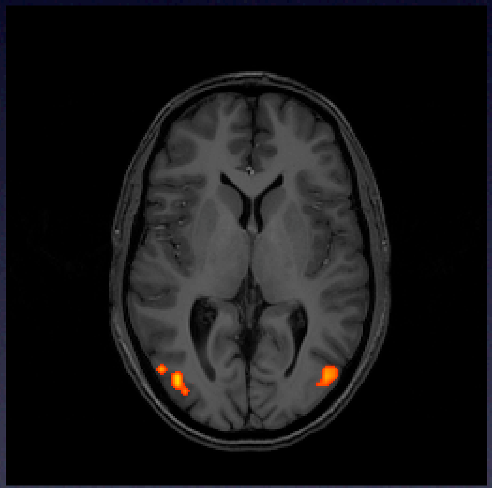
Probabilistic ICA

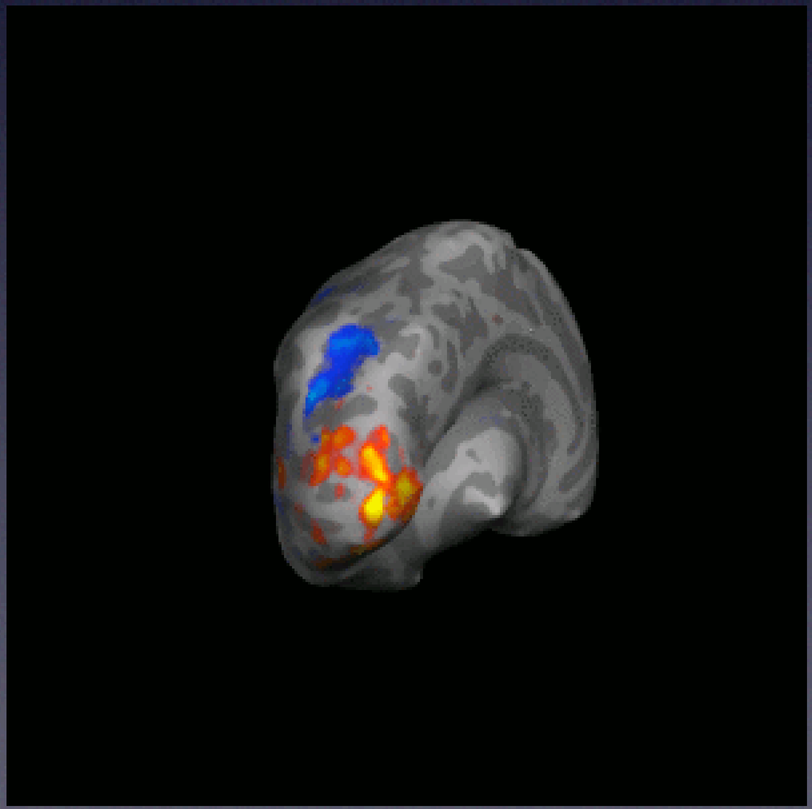
designed to address the 'overfitting problem':

GLM analysis **probabilistic ICA (constrained)**







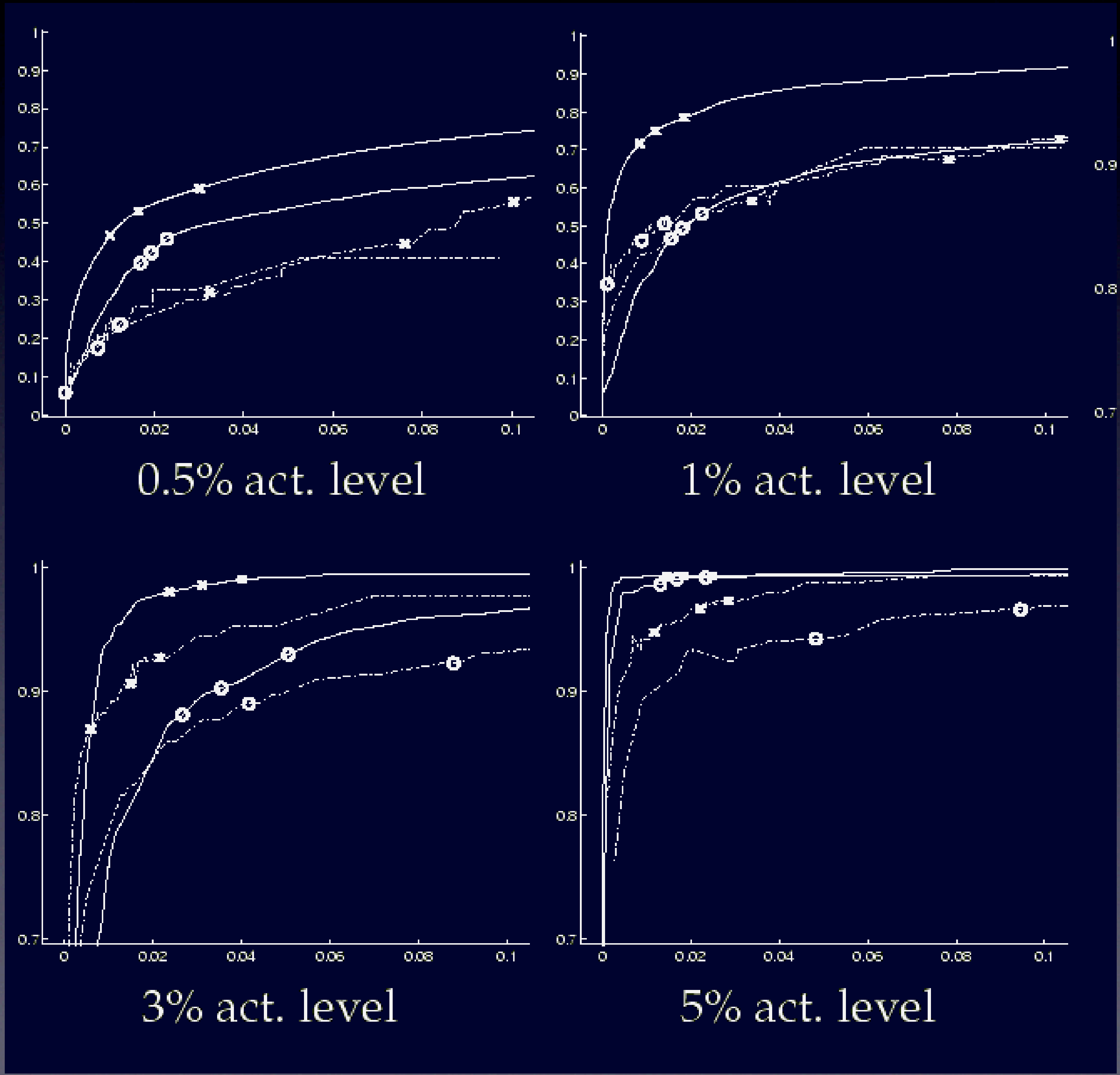


- avoids generation of 'spurious' results
- high spatial sensitivity and specificity

Simulated data

- Receiver - Operator Characteristics:

  PICA vs. GLM at different `activation` levels and different thresholds

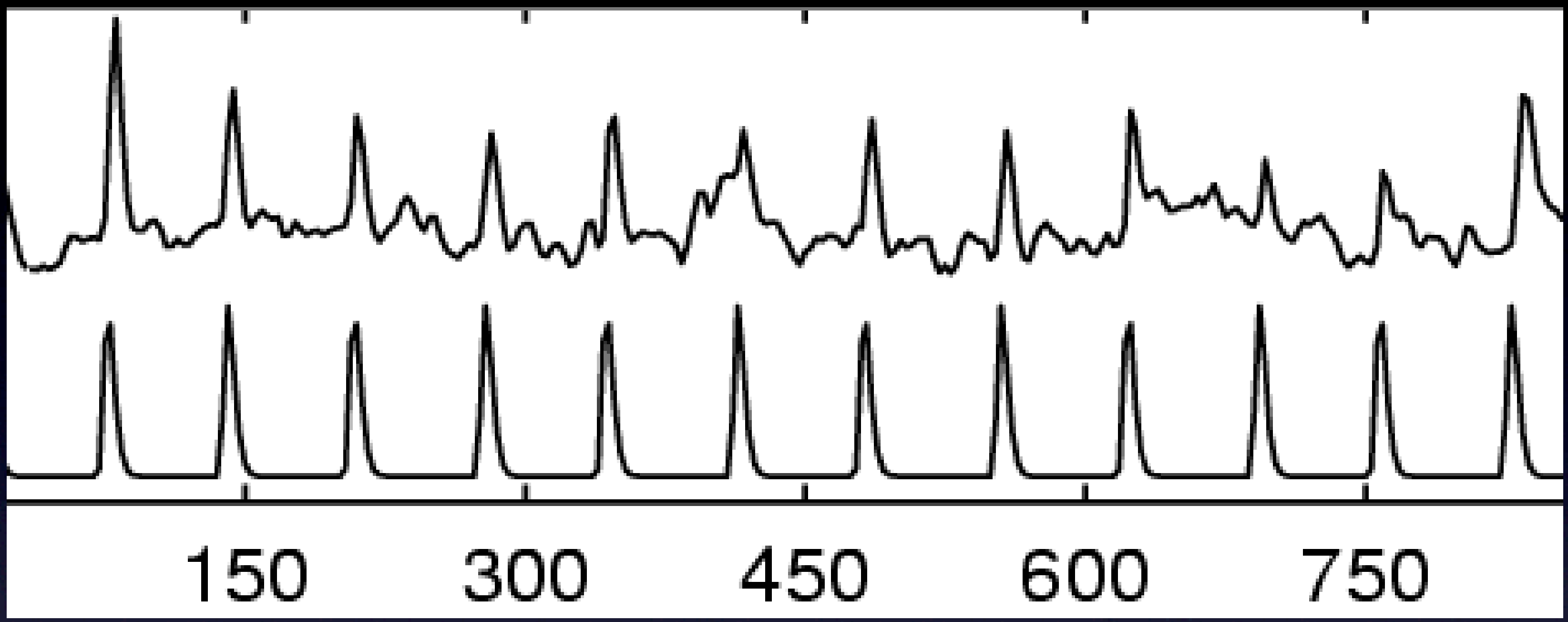


Applications

EDA techniques can be useful to

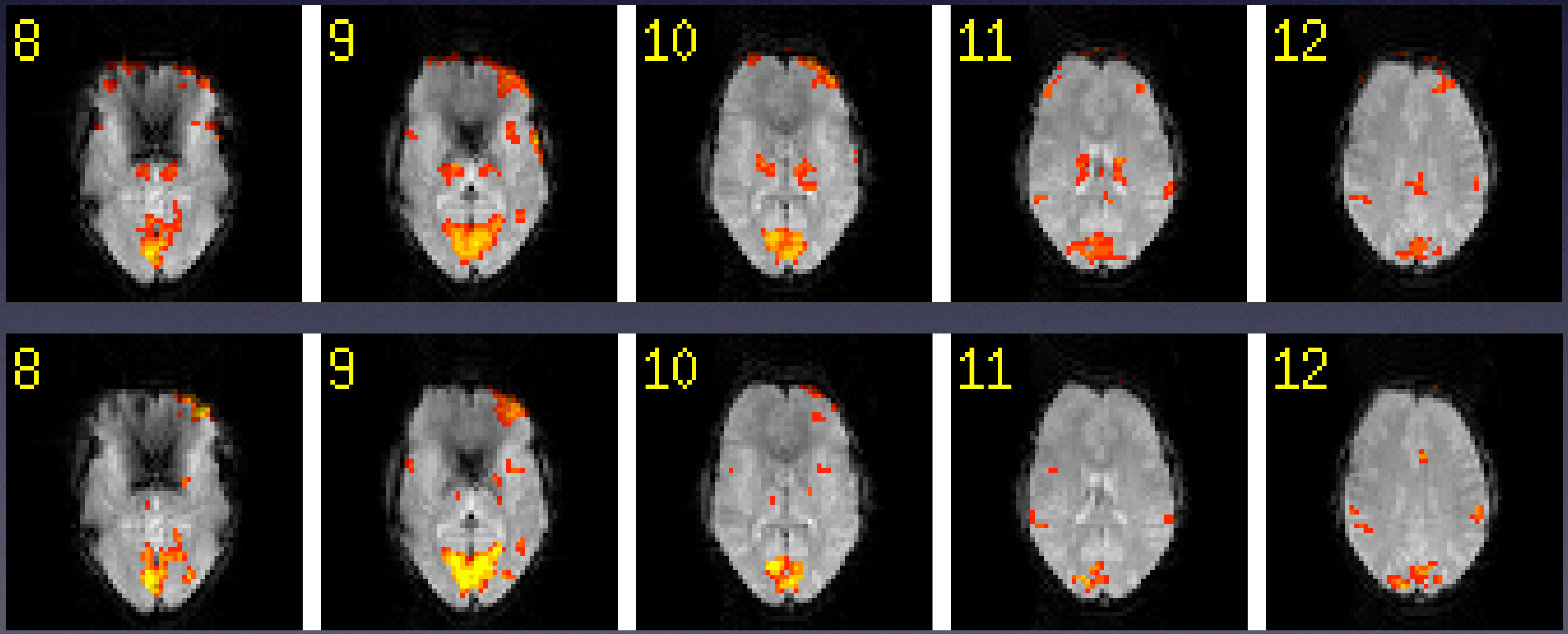
- investigate the BOLD response
- estimate artefacts in the data
- find areas of 'activation' which respond in a non-standard way
- analyse data for which no model of the BOLD response is available

Investigate BOLD



estimated signal time course

standard hrf model



Wise & Tracey

Applications

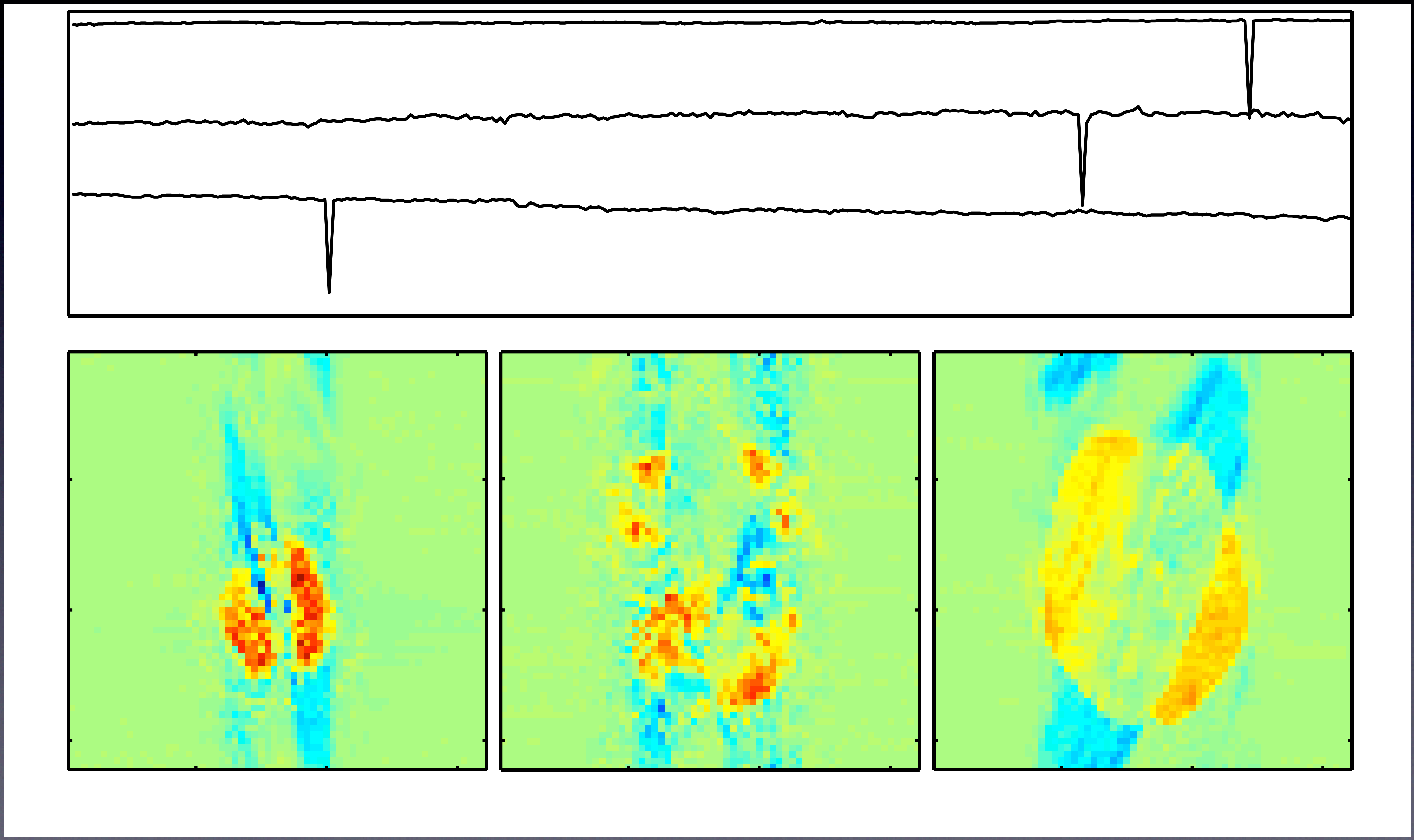
EDA techniques can be useful to

- investigate the BOLD response
- estimate artefacts in the data
- find areas of ‘activation’ which respond in a non-standard way
- analyse data for which no model of the BOLD response is available

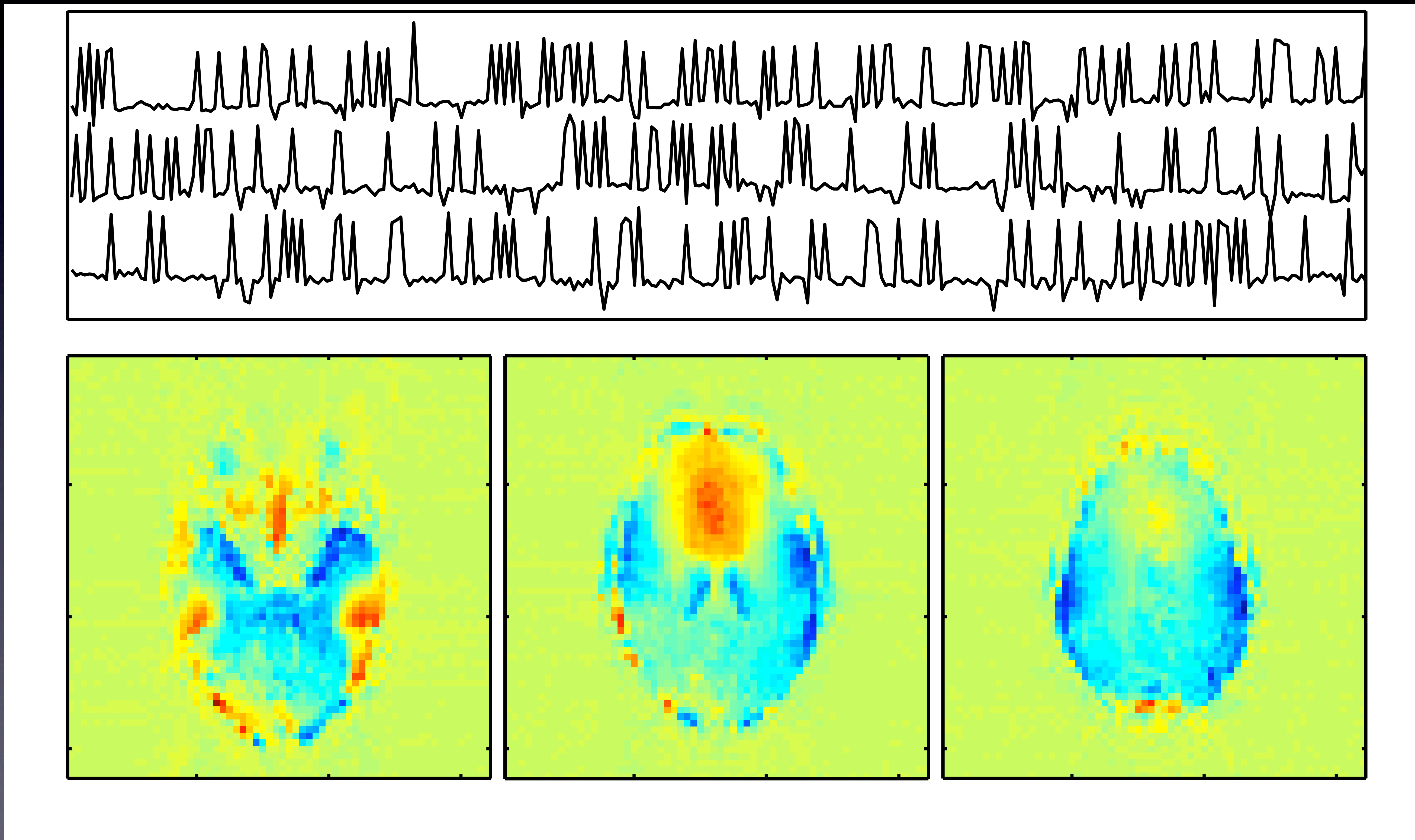
Artefact detection

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

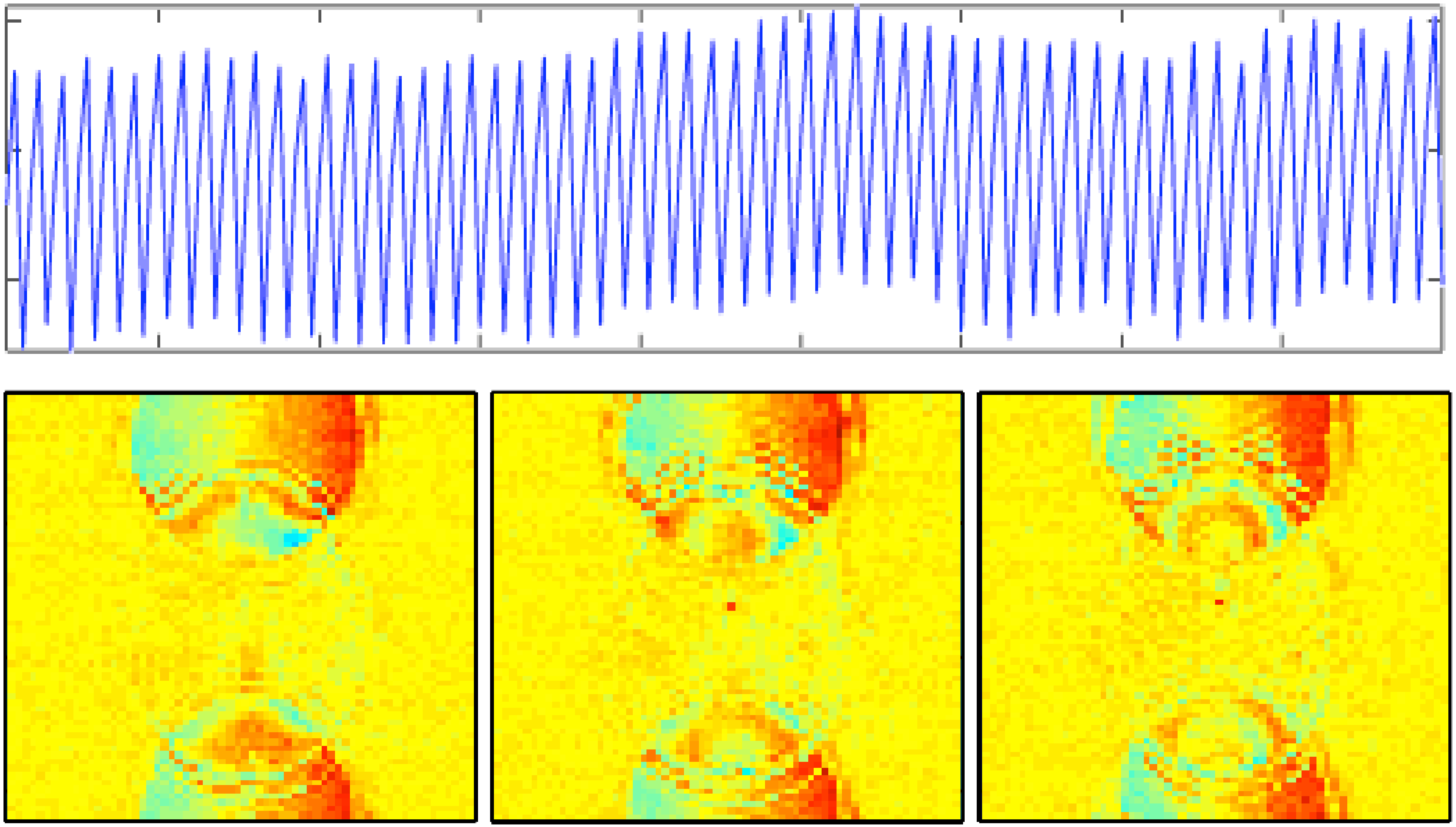
slice drop-outs



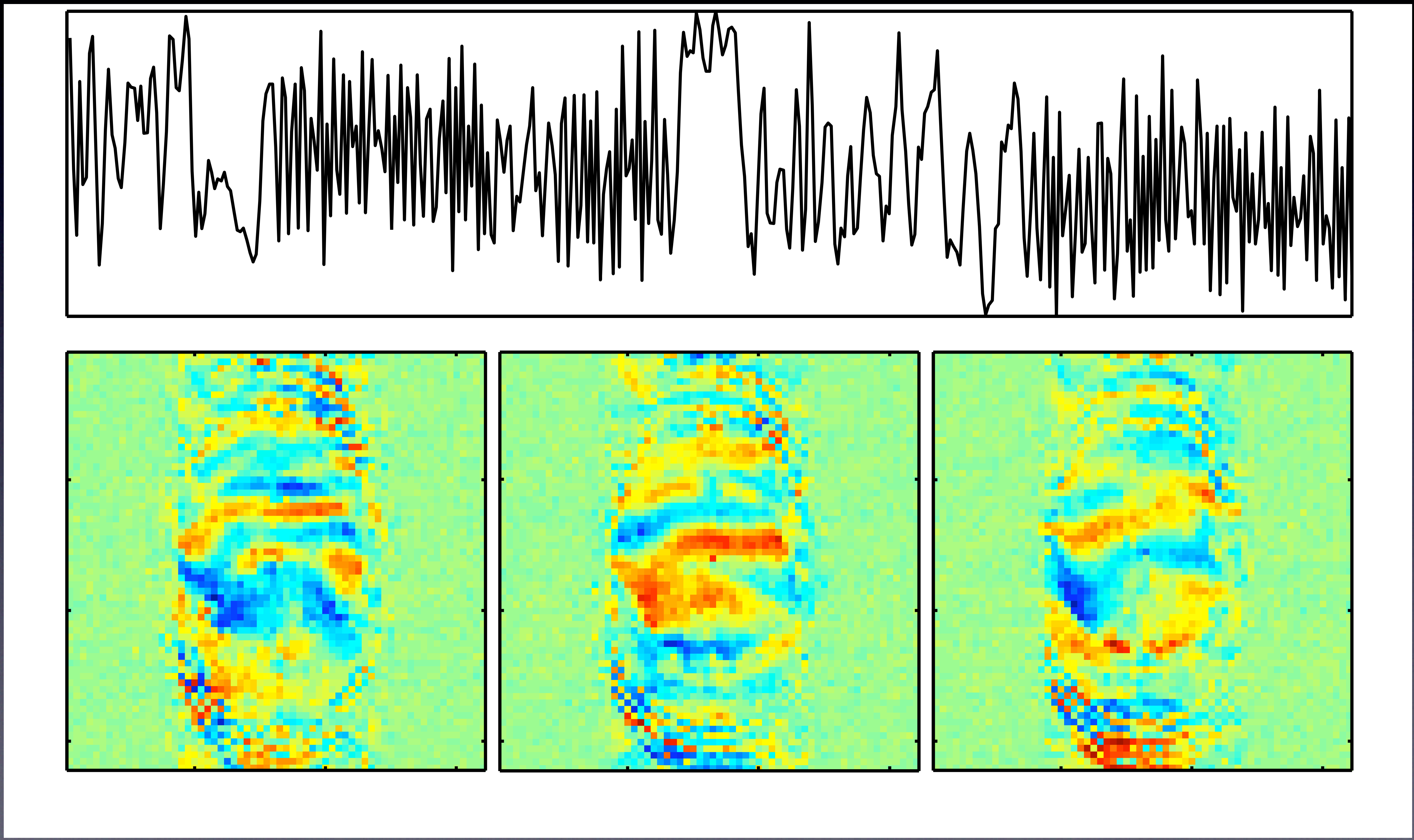
gradient instability



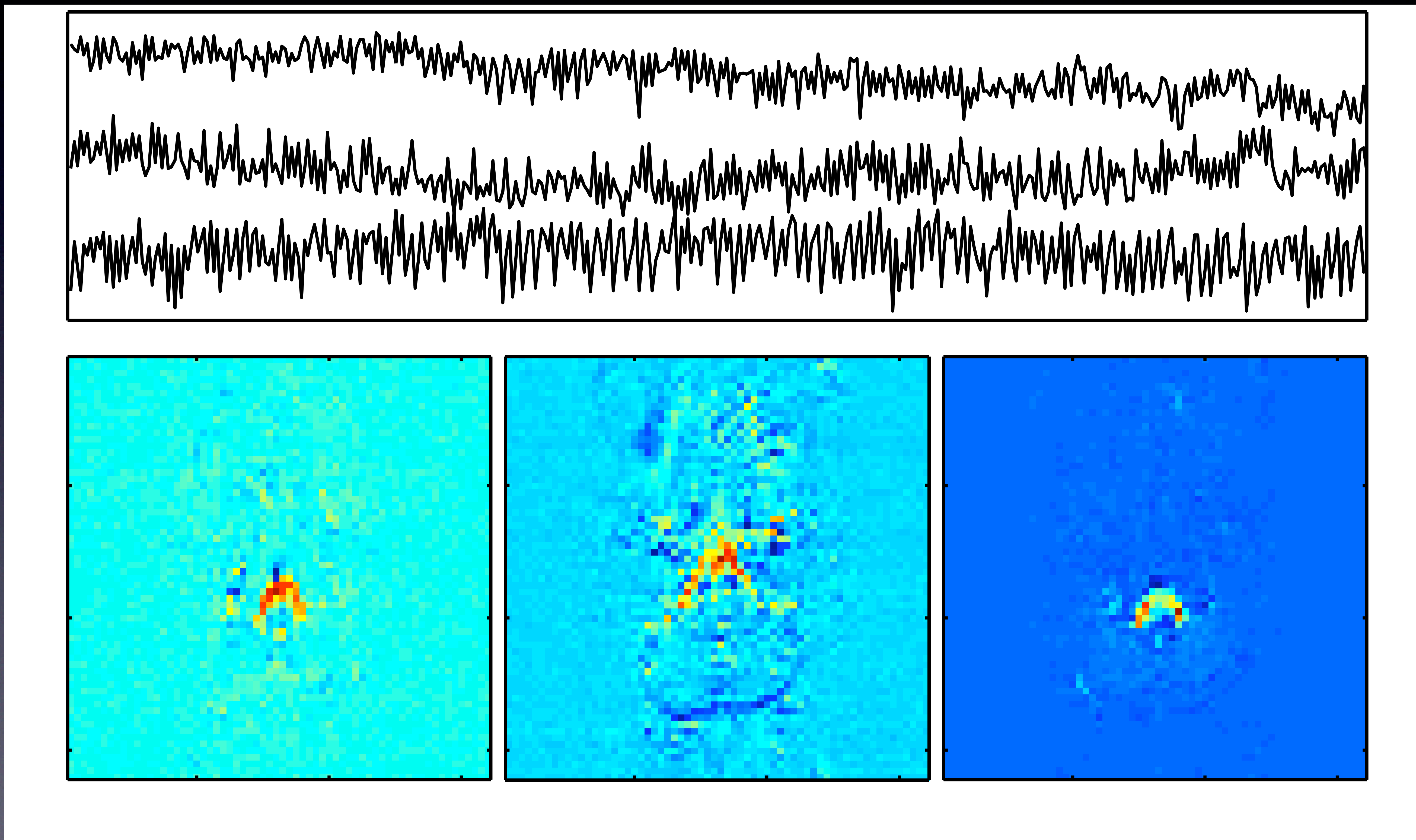
EPI ghost



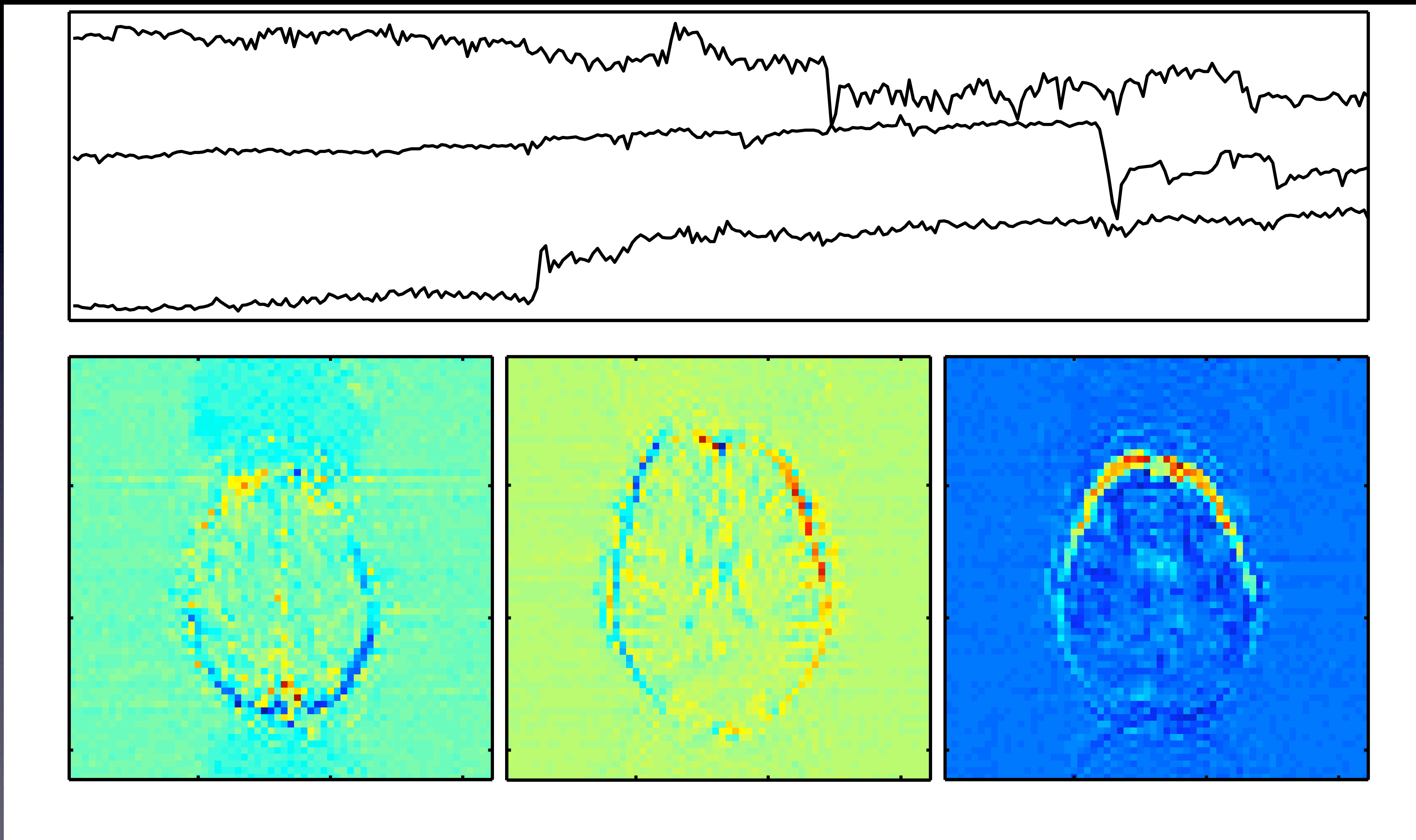
EPI ghost



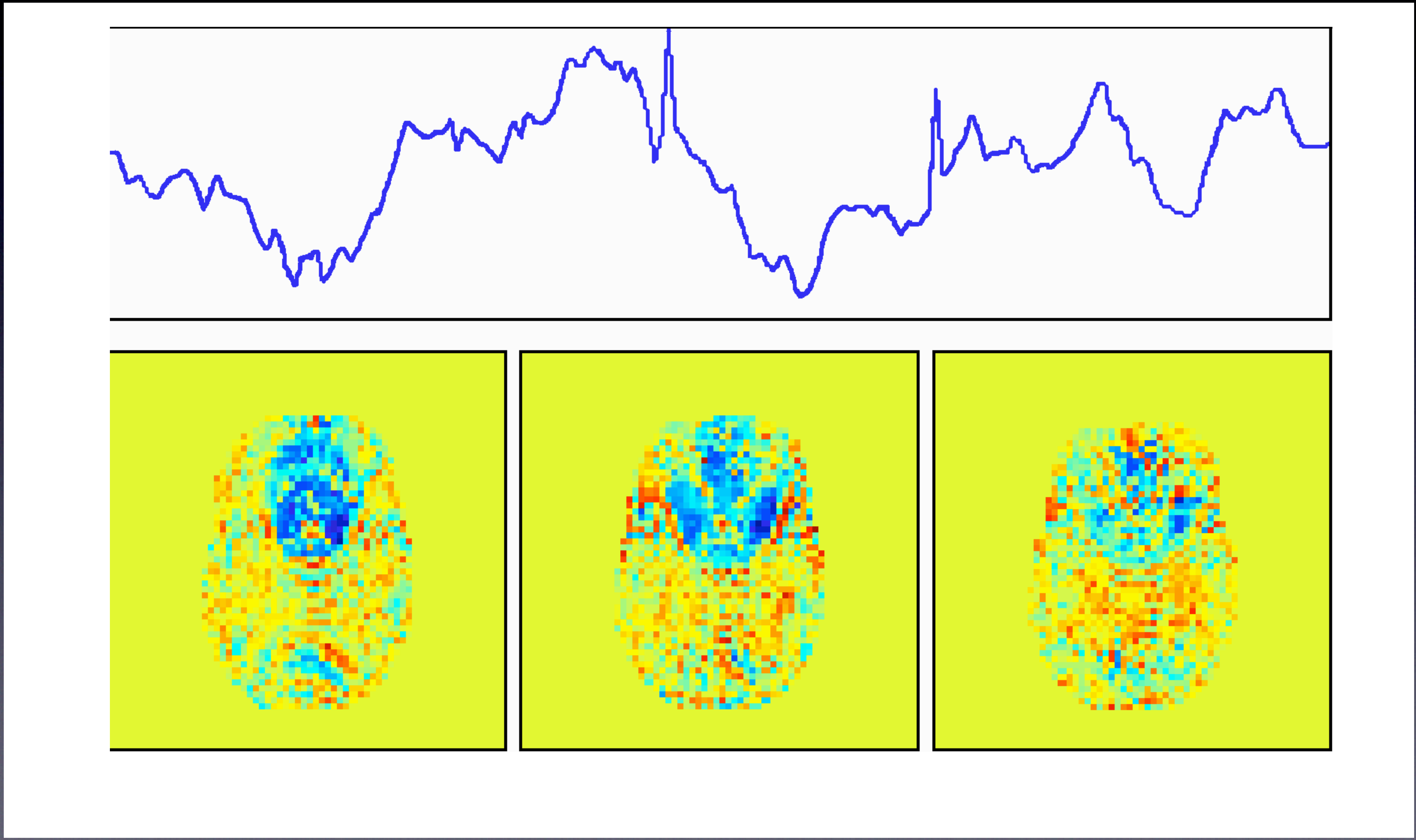
high-frequency noise



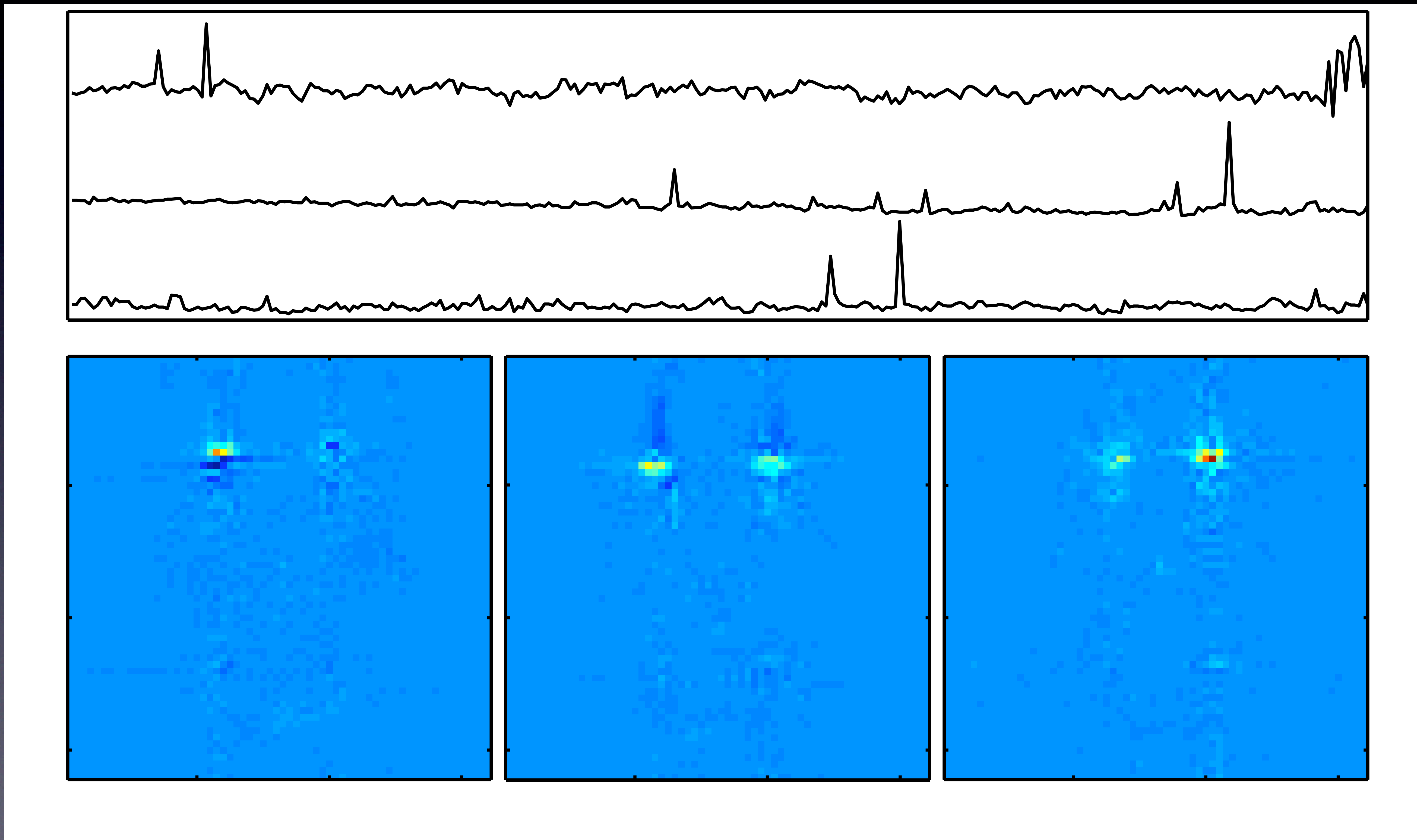
head motion



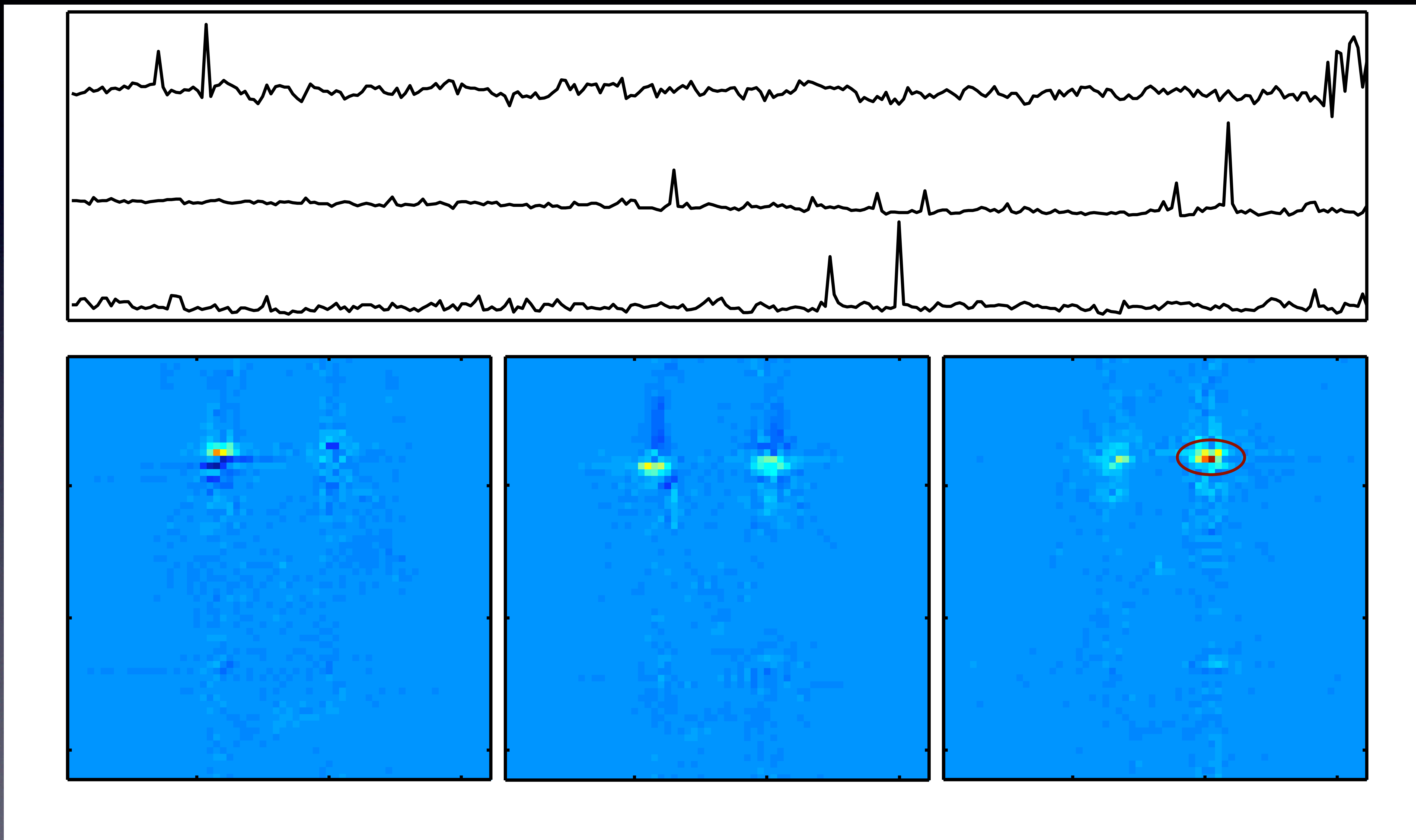
field inhomogeneity



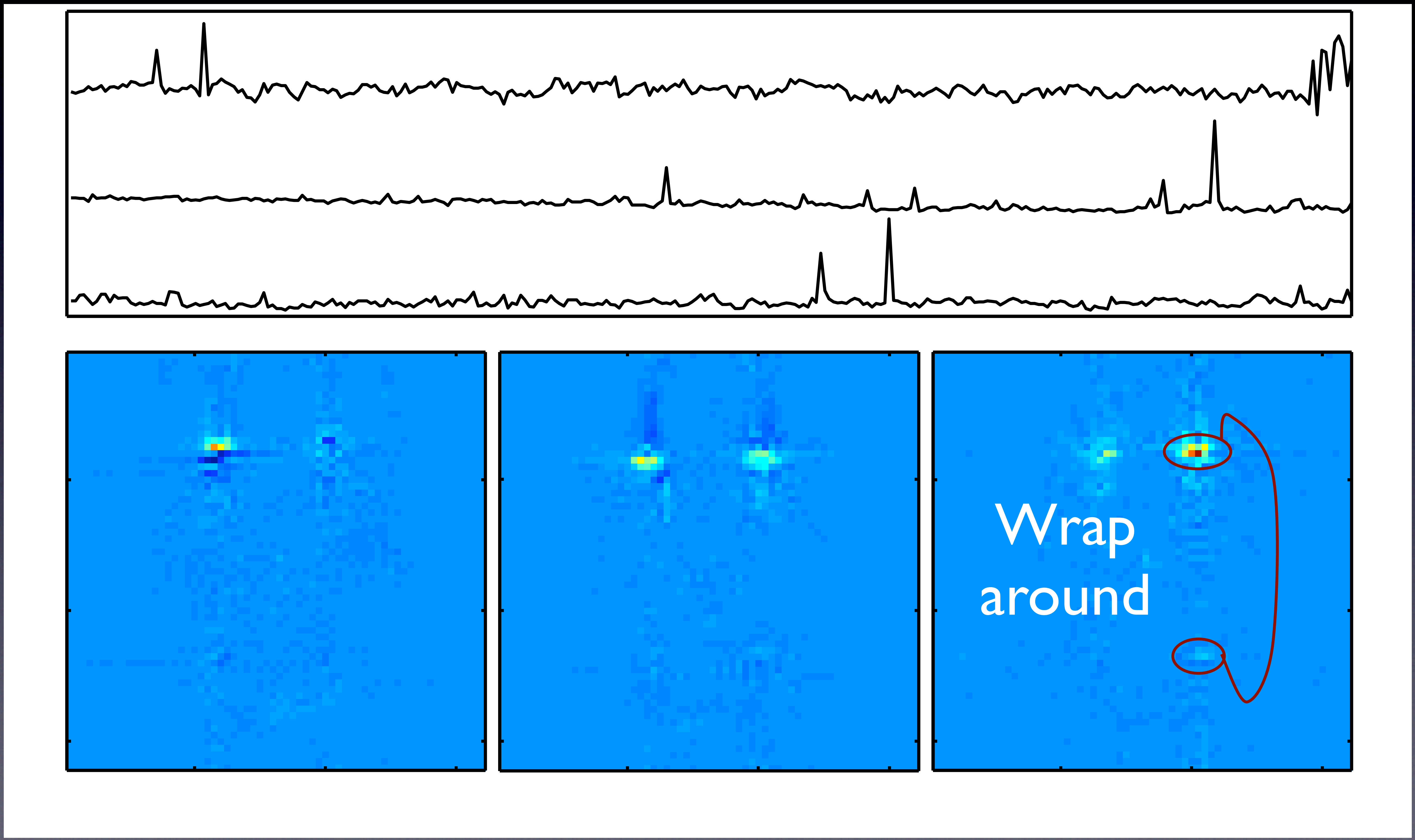
eye-related artefacts



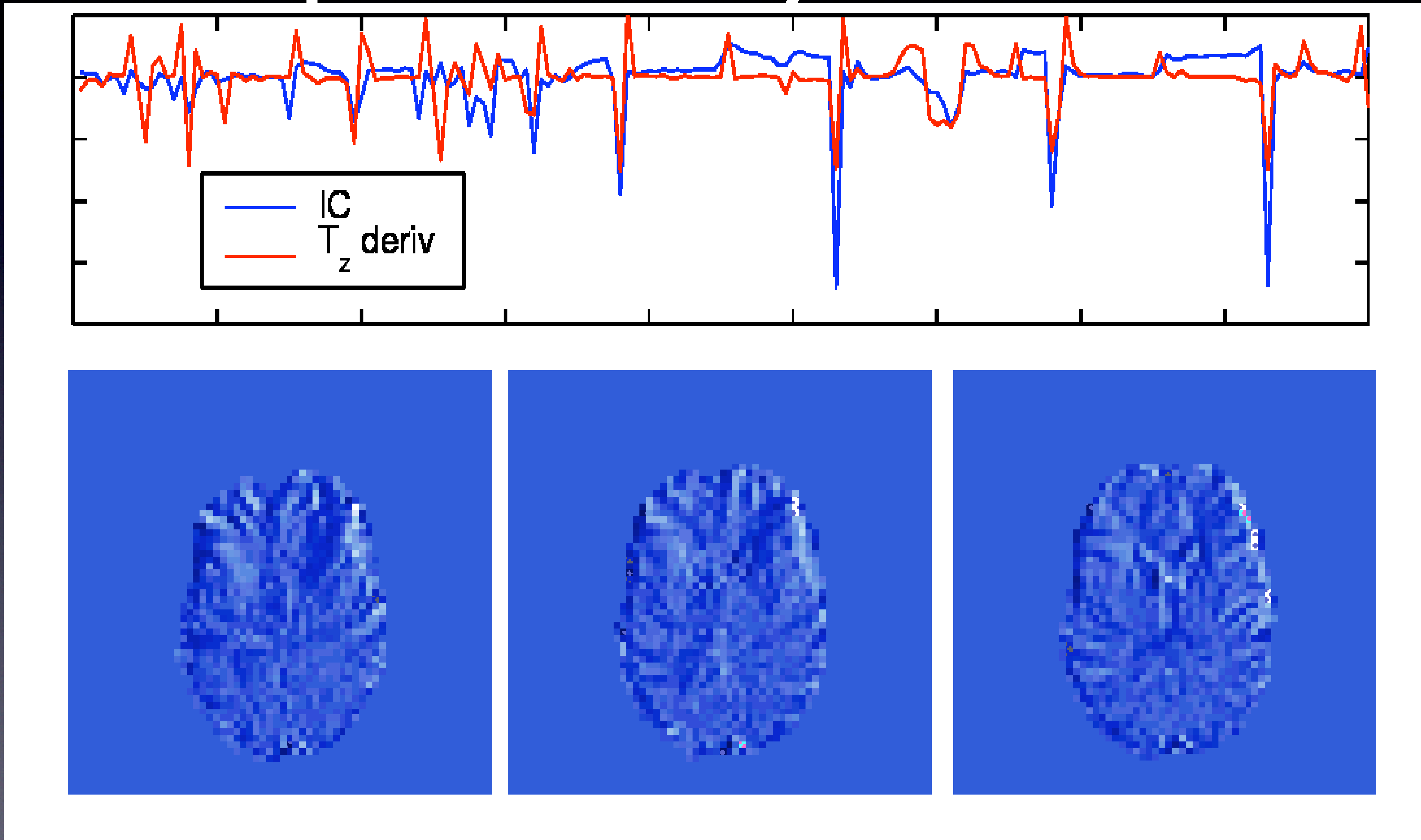
eye-related artefacts



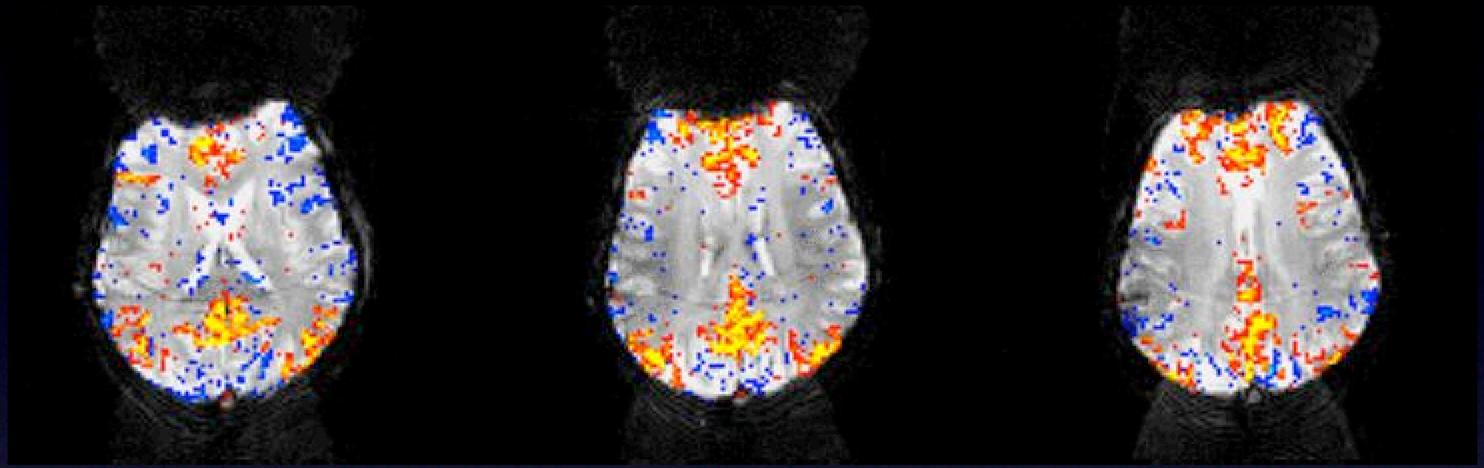
eye-related artefacts



spin-history effects



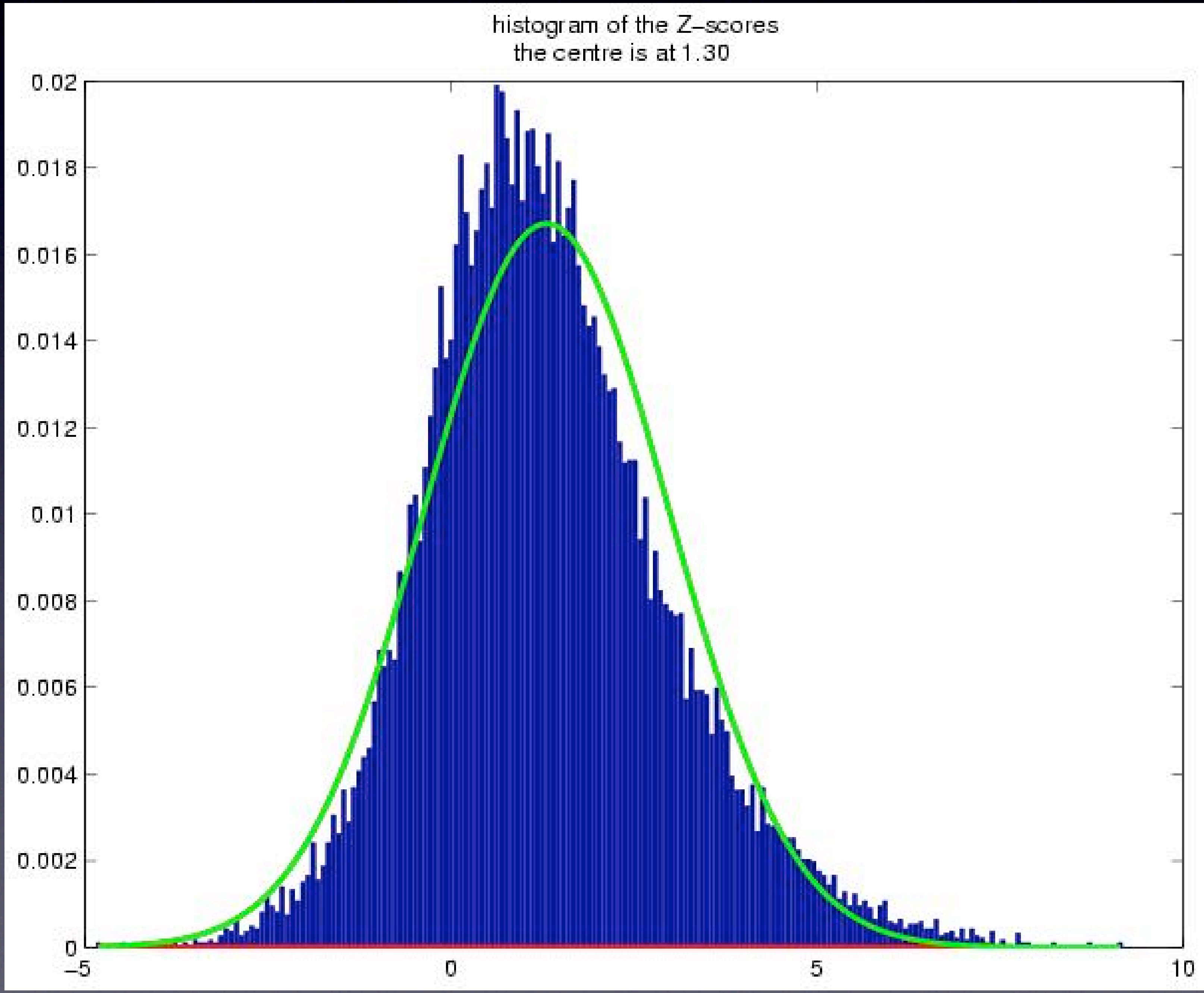
Implication for the GLM



- 'structured noise' appears:
  - in the GLM residuals and inflate variance estimates *(more false negatives)*
  - in the parameter estimates *(more false positives and/or false negatives)*
- In either case lead to suboptimal estimates and wrong inference!

Structured noise and GLM Z-stats bias

- Correlations of the noise time courses with ‘typical’ FMRI regressors can cause a shift in the histogram of the Z-statistics
- Thresholded maps will have wrong false-positive rate

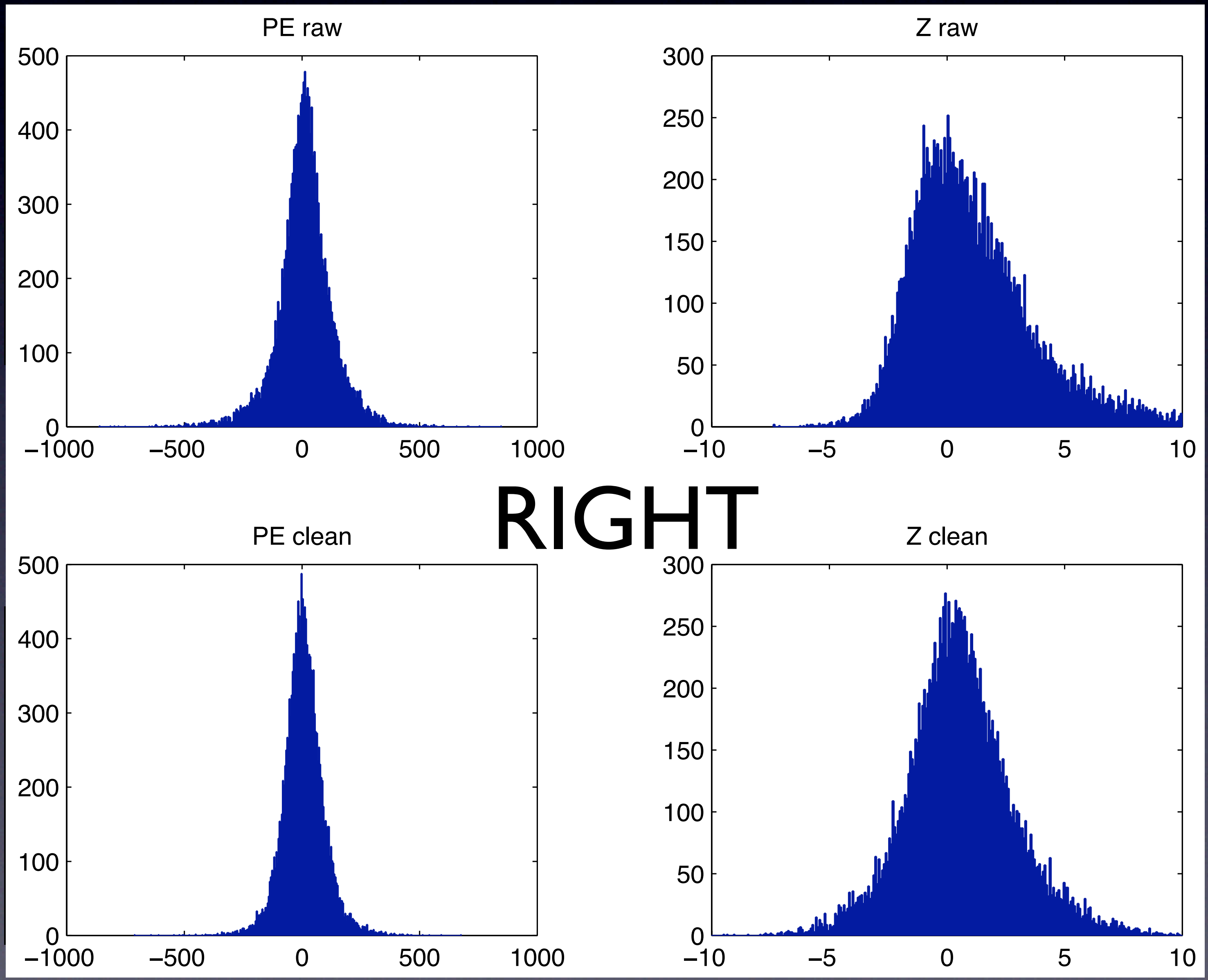


Denoising FMRI

- Example: left vs right hand finger tapping

Johansen-Berg et al. PNAS 2002

before denoising



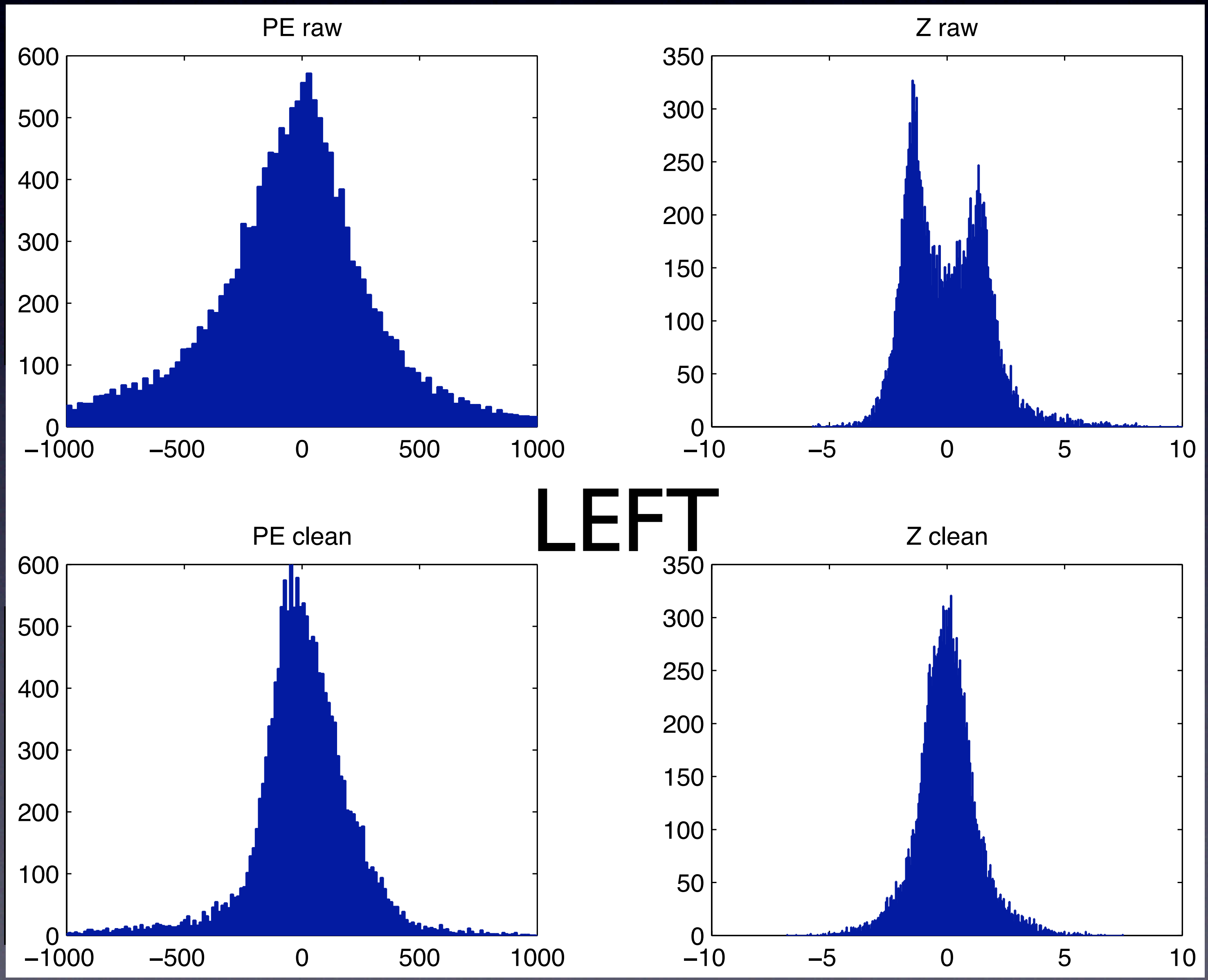
after denoising

Denoising FMRI

- Example: left vs right hand finger tapping

Johansen-Berg et al. PNAS 2002

before denoising



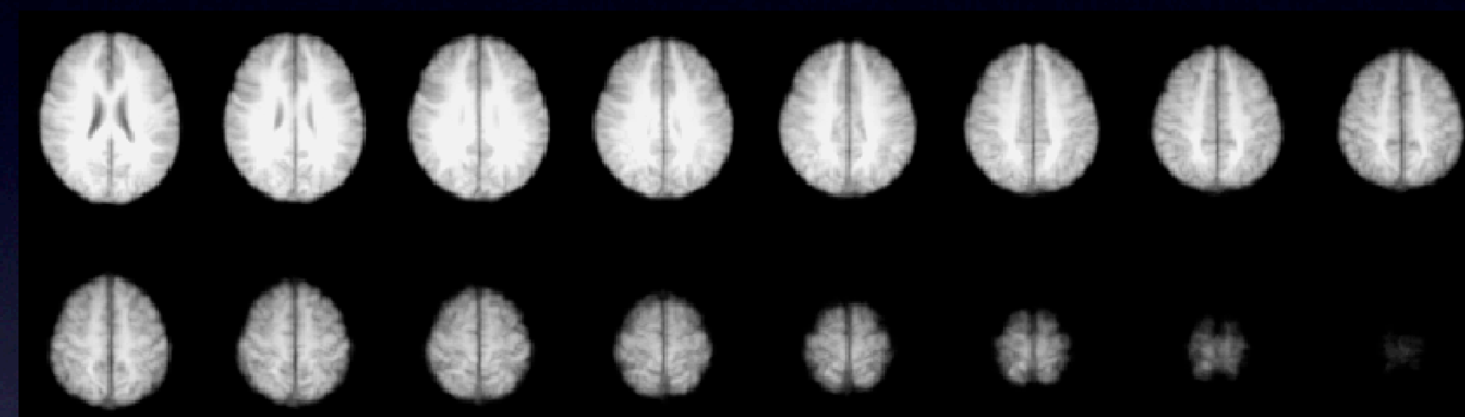
after denoising

Denoising FMRI

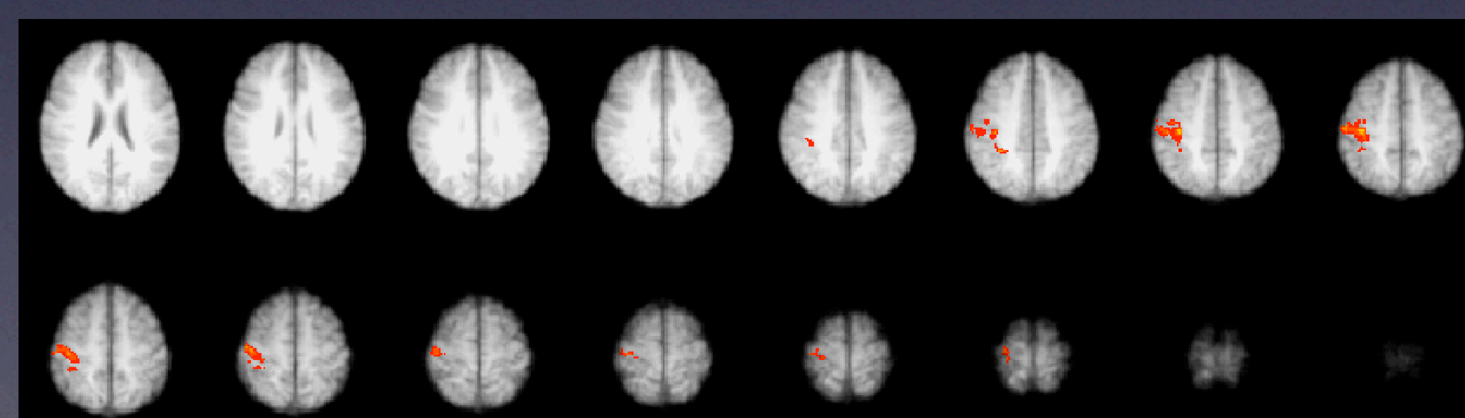
- Example: left vs right hand finger tapping

Johansen-Berg et al. PNAS 2002

before denoising

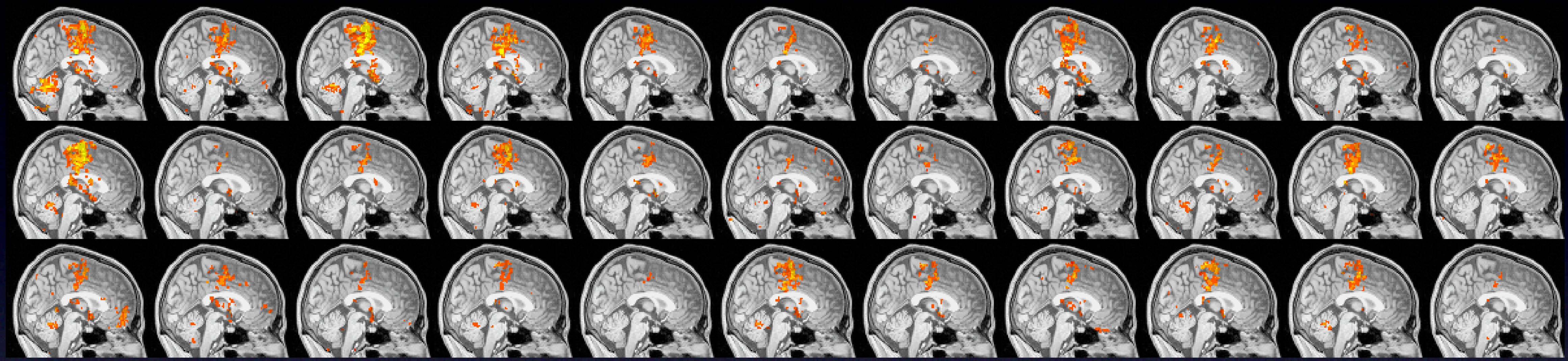


LEFT - RIGHT contrast

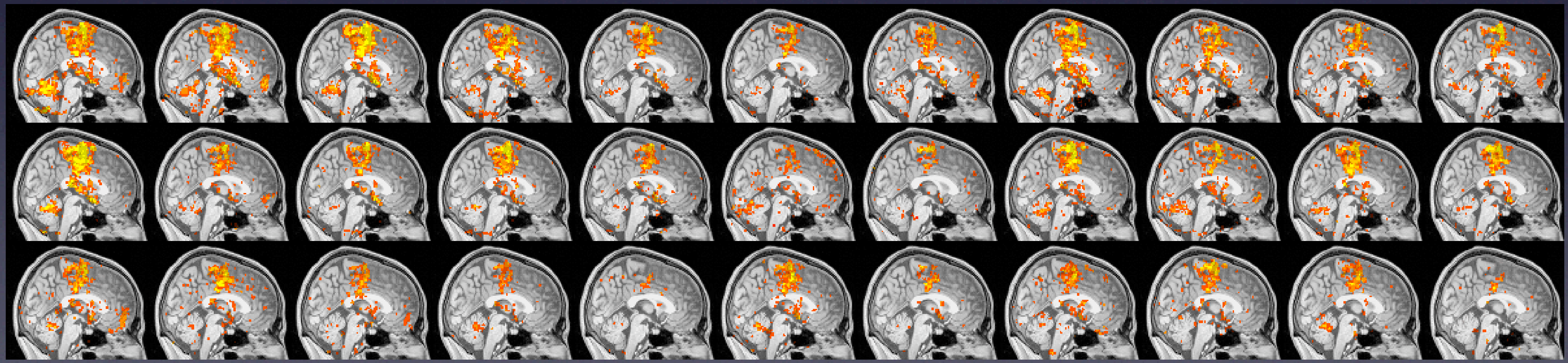


after denoising

Apparent variability



McGonigle et al.: 33 Sessions under motor paradigm



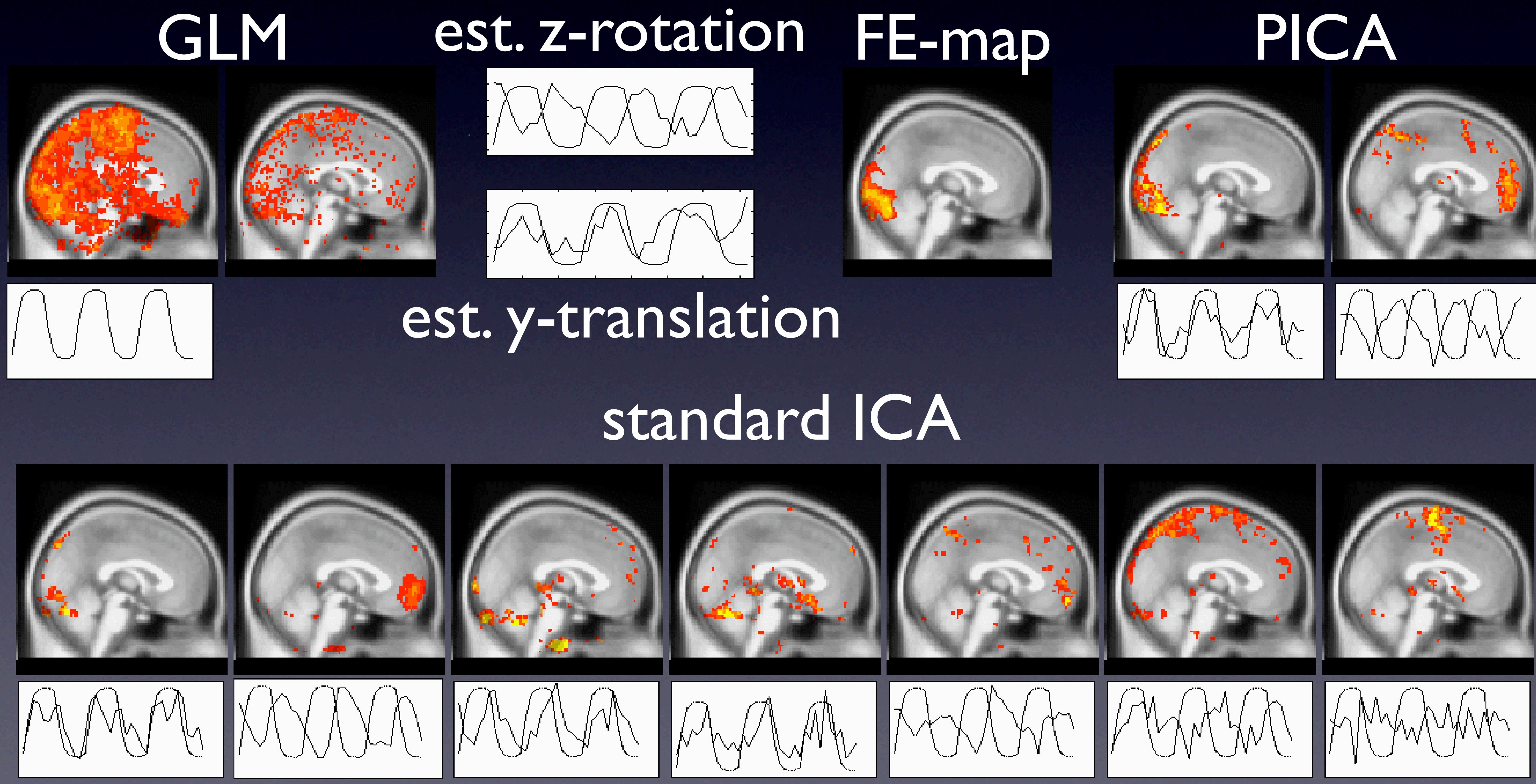
'de-noising' data by regressing out noise: reduced 'apparent' session variability

Applications

EDA techniques can be useful to

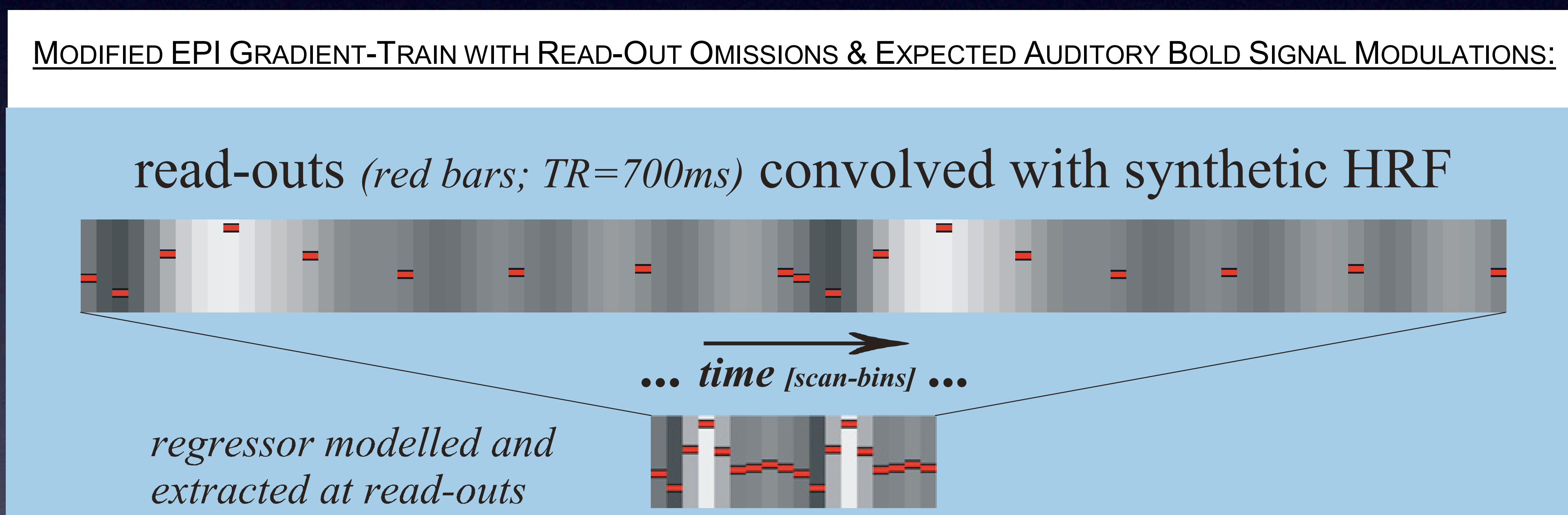
- investigate the BOLD response
- estimate artefacts in the data
- find areas of 'activation' which respond in a non-standard way
- analyse data for which no model of the BOLD response is available

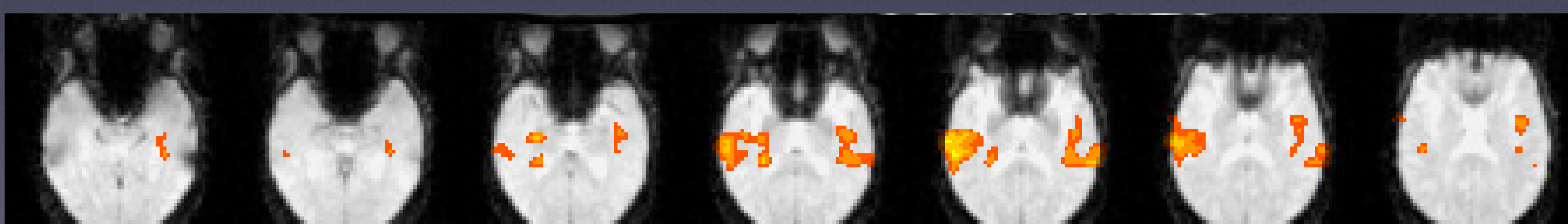
Example: stim. correl. motion



Clinical example

- use EPI readout gradient noise to evoke auditory responses in patients before cochlear implant surgery

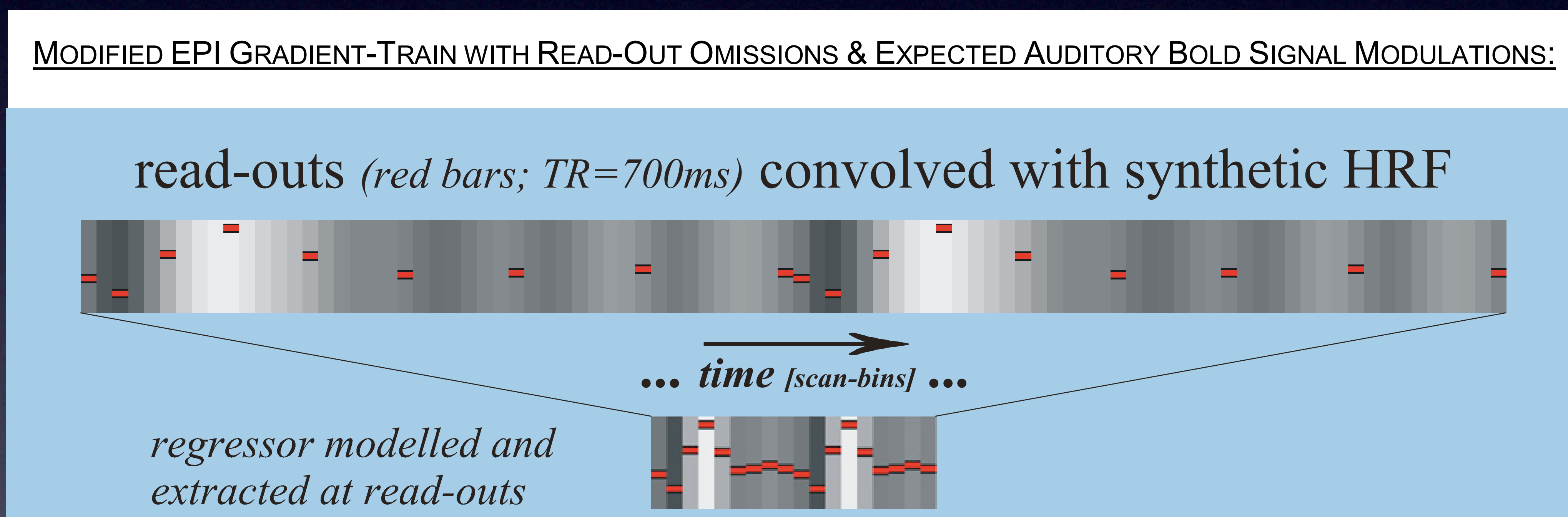


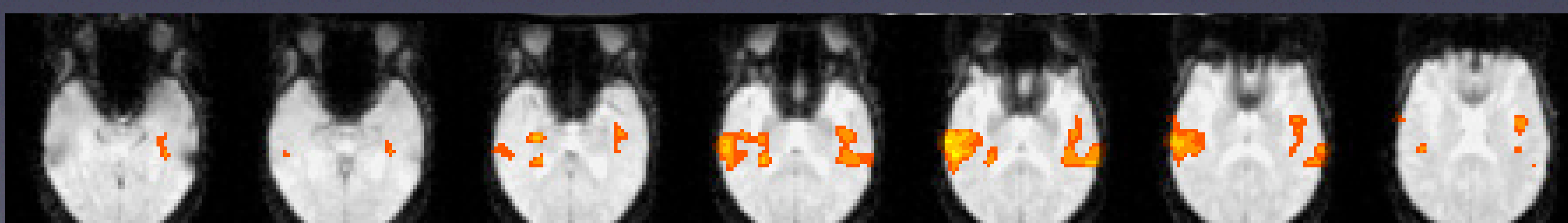


Bartsch et al. HBM 2004

Clinical example

- use EPI readout gradient noise to evoke auditory responses in patients before cochlear implant surgery





Bartsch et al. HBM 2004

Applications

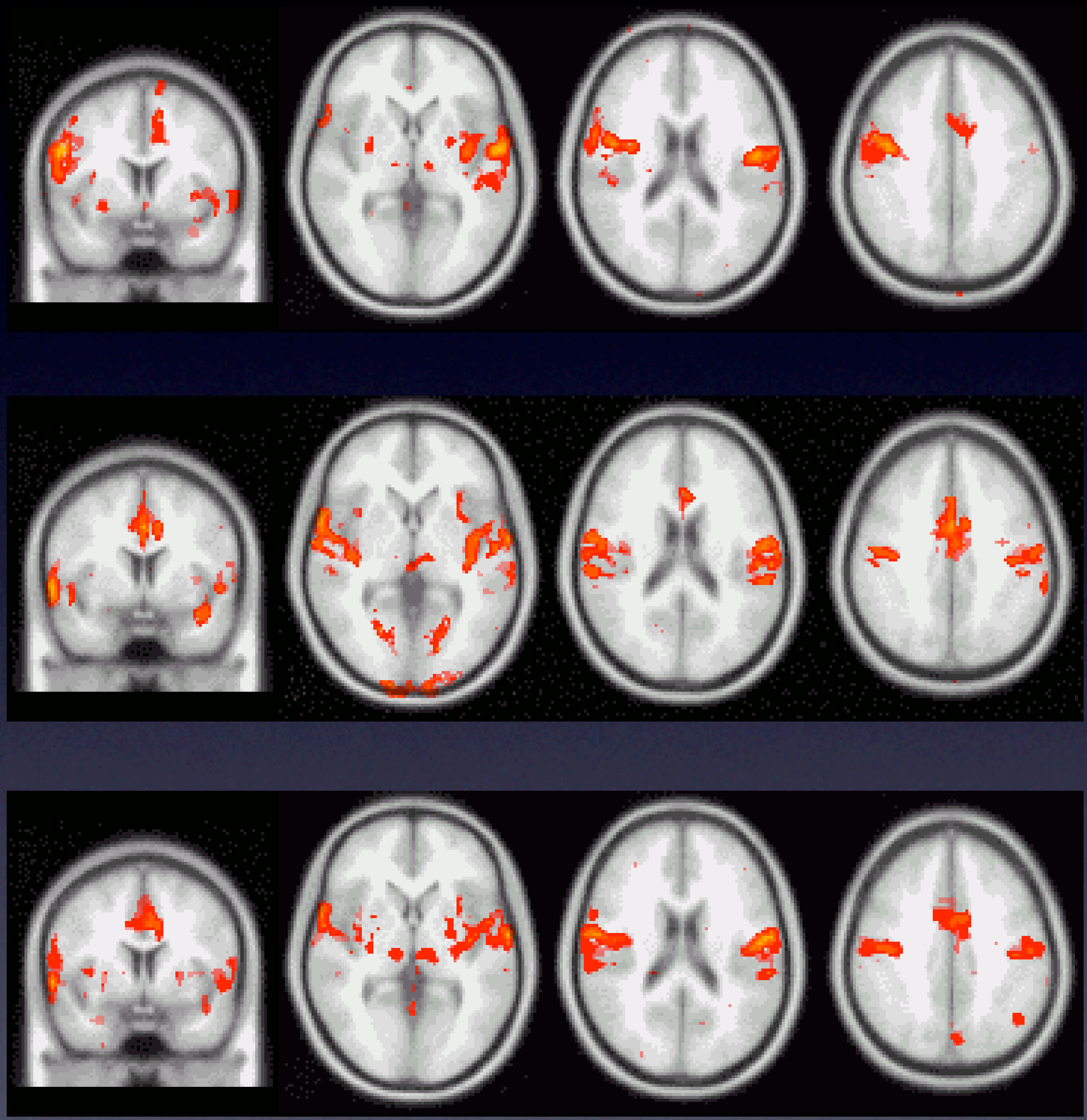
EDA techniques can be useful to

- investigate the BOLD response
- estimate artefacts in the data
- find areas of ‘activation’ which respond in a non-standard way

▸ analyse data for which no model of the BOLD response is available

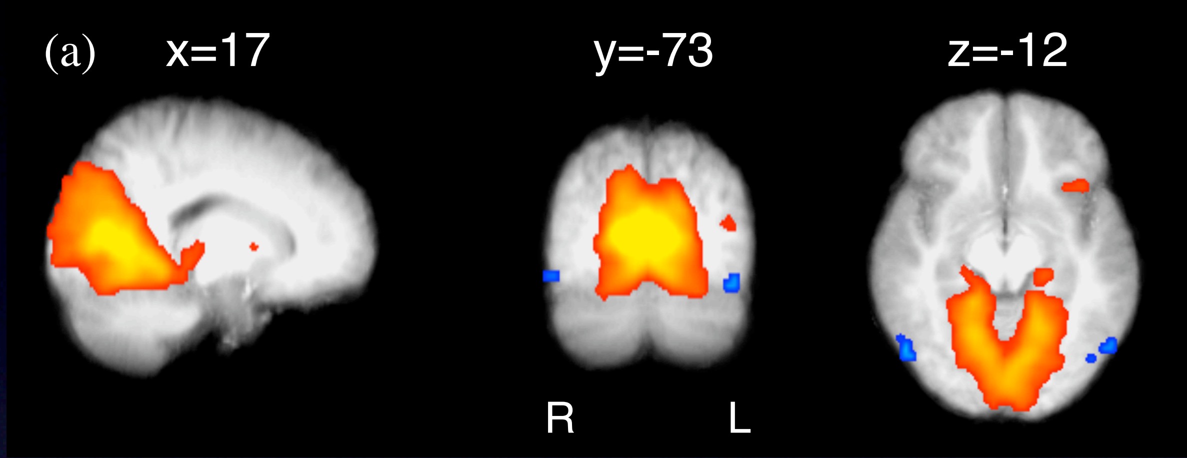
PICA on resting data

- perform ICA on null data and compare spatial maps between subjects/scans
- ICA maps depict spatially localised and temporally coherent signal changes

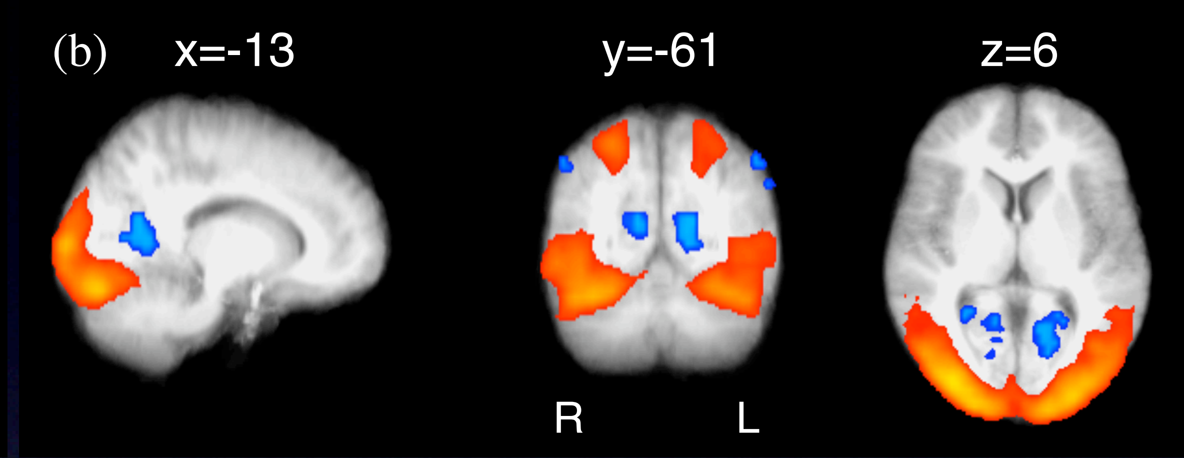


Example: ICA maps - 1 subject at 3 different sessions

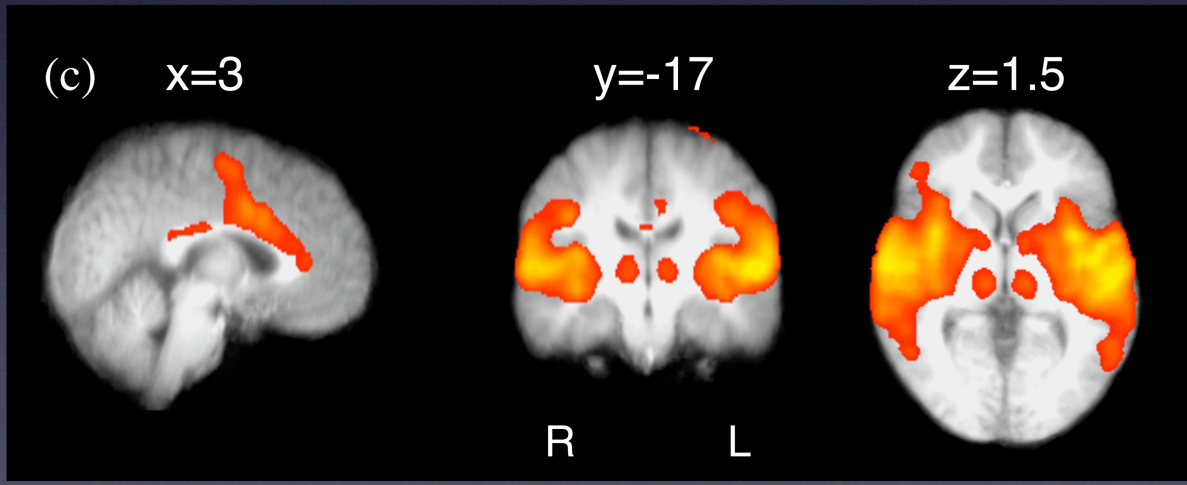
Spatial characteristics



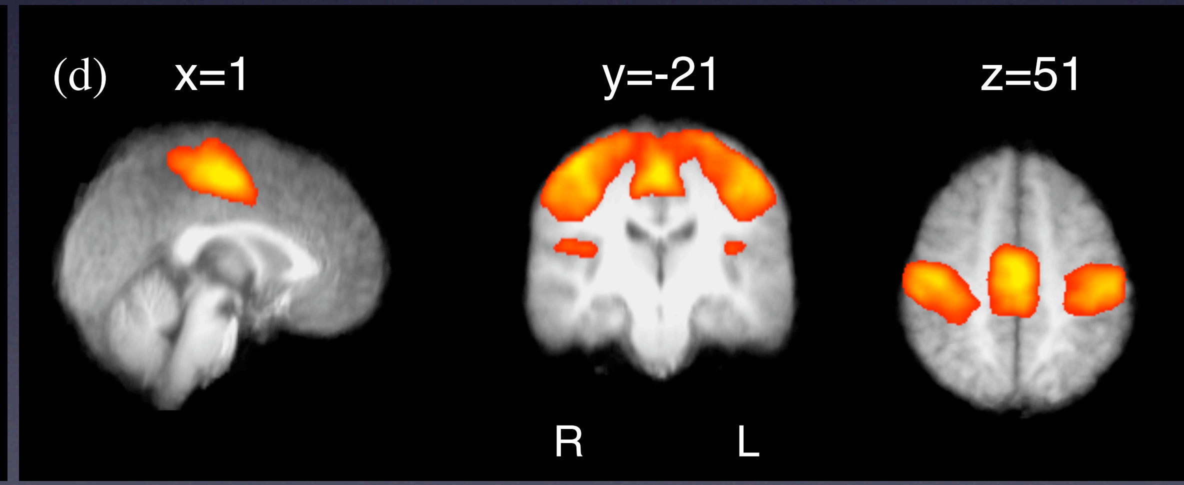
Medial visual cortex



Lateral Visual Cortex

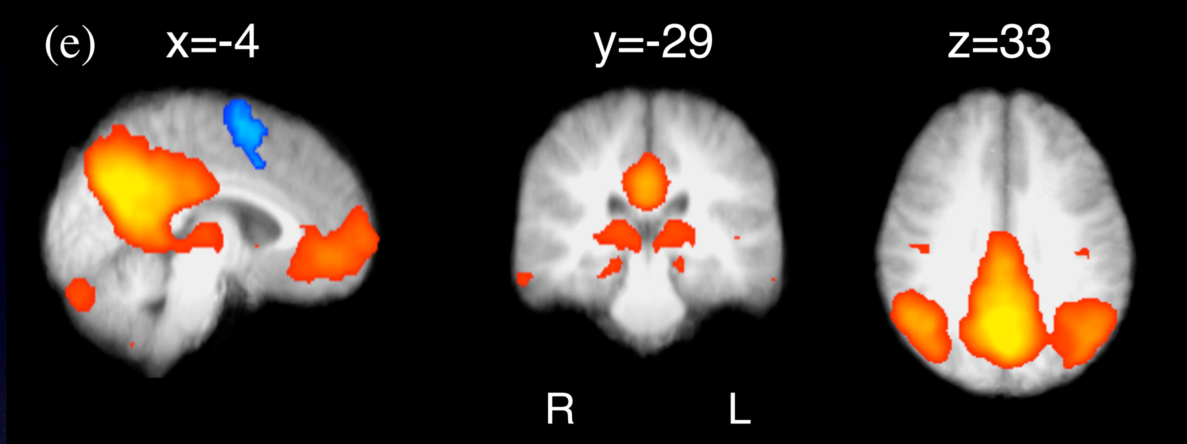


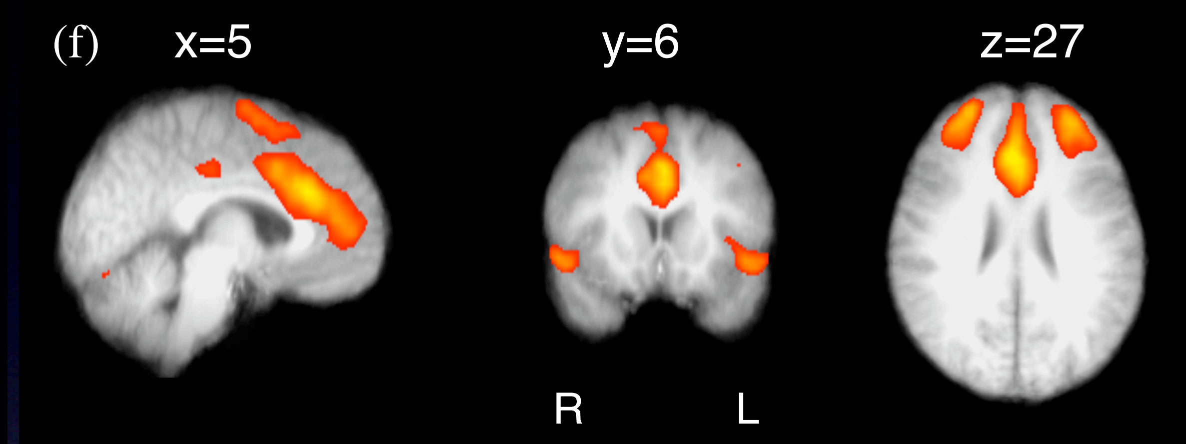
Auditory system



Sensori-motor system

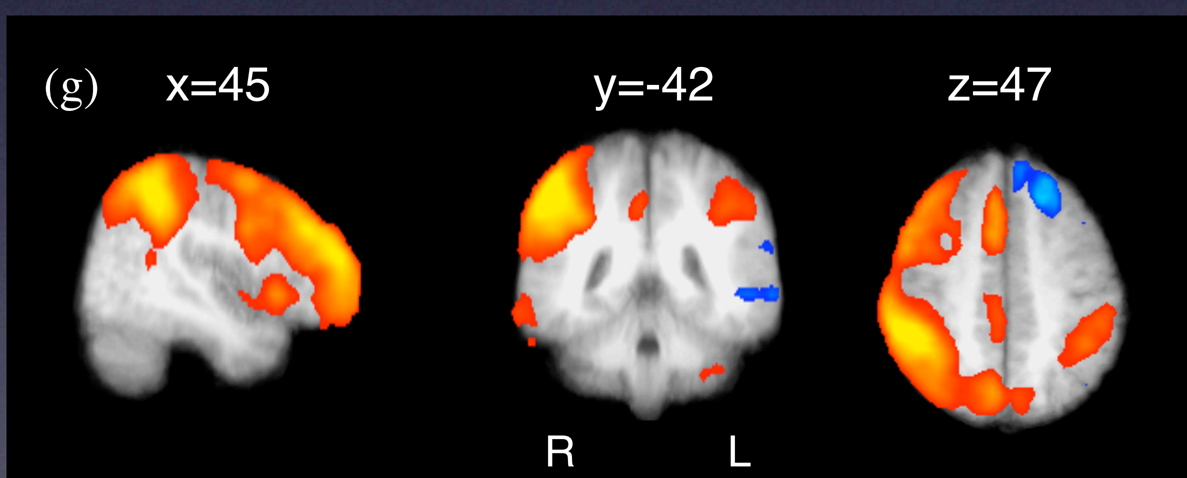
Spatial characteristics

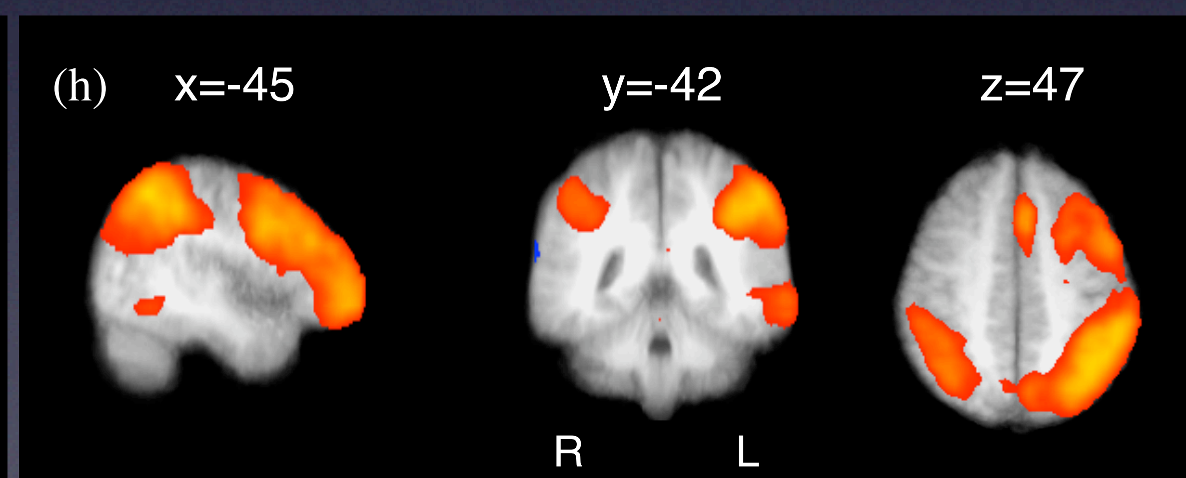




Visuospatial system

Executive control

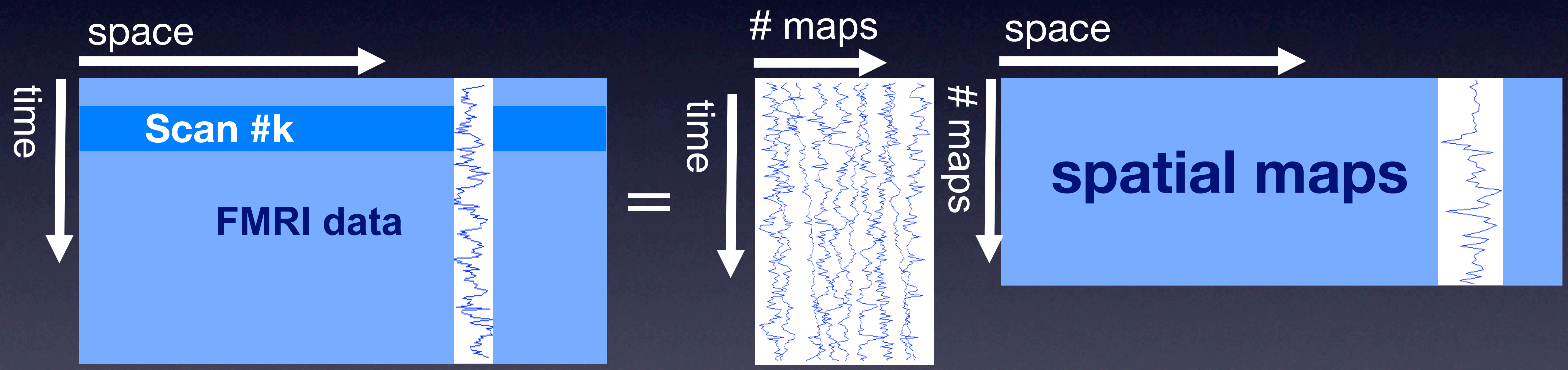




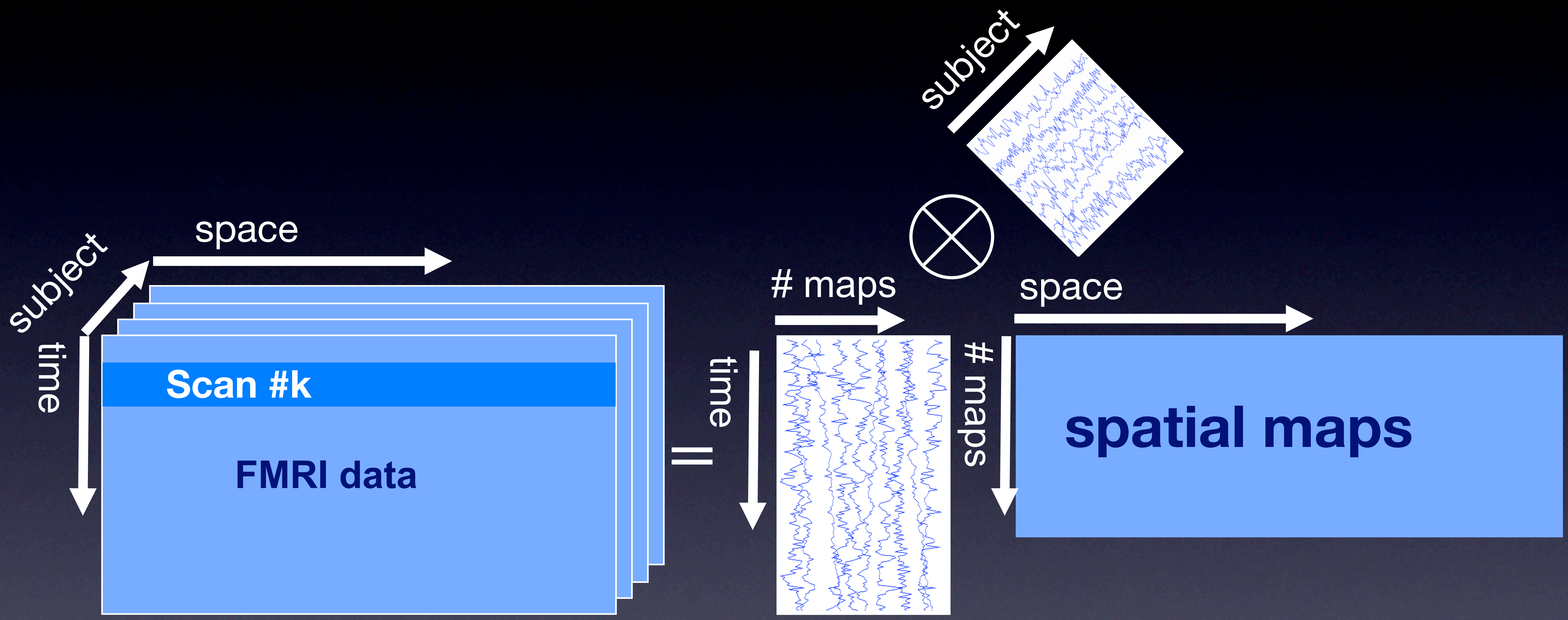
Visual Stream

ICA Group analysis

Extend single ICA to higher dimensions



ICA Group analysis



- Factors characterise processes across space, time and sessions/subjects

$$x_{ijk} = \sum_r a_{ir} b_{jr} c_{kr} + \epsilon_{ijk}$$

Beckmann and Smith NI 2005

PARAFAC

PARAFAC

- as a symmetric least-square problem this is known as *PARAFAC (Parallel Factor Analysis; Harshman 1970)*
- can be solved using Alternating Least Squares (ALS), i.e. by iterating least-squares solutions for

$$\mathbf{X}_{i..} = \mathbf{B}\mathrm{diag}(\mathbf{a}_i)\mathbf{C}^t + \mathbf{E}_{i..} \quad i = 1, \ldots, I$$
$$\mathbf{X}_{.j.} = \mathbf{C}\mathrm{diag}(\mathbf{b}_j)\mathbf{A}^t + \mathbf{E}_{.j.} \quad j = 1, \ldots, J$$
$$\mathbf{X}_{..k} = \mathbf{A}\mathrm{diag}(\mathbf{c}_k)\mathbf{B}^t + \mathbf{E}_{..k} \quad k = 1, \ldots, K$$

PARAFAC

- as a symmetric least-square problem this is known as *PARAFAC (Parallel Factor Analysis; Harshman 1970)*
- can be solved using Alternating Least Squares (ALS), i.e. by iterating least-squares solutions for

$$\mathbf{X}_{i..} = \mathbf{B}\mathrm{diag}(\mathbf{a}_i)\mathbf{C}^t + \mathbf{E}_{i..} \quad i = 1, \ldots, I$$
$$\mathbf{X}_{.j.} = \mathbf{C}\mathrm{diag}(\mathbf{b}_j)\mathbf{A}^t + \mathbf{E}_{.j.} \quad j = 1, \ldots, J$$
$$\mathbf{X}_{..k} = \mathbf{A}\mathrm{diag}(\mathbf{c}_k)\mathbf{B}^t + \mathbf{E}_{..k} \quad k = 1, \ldots, K$$

- treats all modes the same

PARAFAC

- as a symmetric least-square problem this is known as *PARAFAC (Parallel Factor Analysis; Harshman 1970)*
- can be solved using Alternating Least Squares (ALS), i.e. by iterating least-squares solutions for

$$
\begin{aligned}
\mathbf{X}_{i..} &= \mathbf{B}\mathrm{diag}(\mathbf{a}_i)\mathbf{C}^t + \mathbf{E}_{i..} \quad & i = 1, \ldots, I \\
\mathbf{X}_{.j.} &= \mathbf{C}\mathrm{diag}(\mathbf{b}_j)\mathbf{A}^t + \mathbf{E}_{.j.} \quad & j = 1, \ldots, J \\
\mathbf{X}_{..k} &= \mathbf{A}\mathrm{diag}(\mathbf{c}_k)\mathbf{B}^t + \mathbf{E}_{..k} \quad & k = 1, \ldots, K
\end{aligned}
$$

- treats all modes the same
- requires *system variation* (no co-linearity in modes)

Tensor-PICA

Tensor-PICA

- rewrite: $\mathbf{X}_{IK \times J} = (\mathbf{C}|\otimes|\mathbf{A})\mathbf{B}^t + \mathbf{E}$

Tensor-PICA

- rewrite: $\mathbf{X}_{IK \times J} = (\mathbf{C}|\otimes|\mathbf{A})\mathbf{B}^t + \mathbf{E}$
- can be treated as a 2-stage estimation problem:

Tensor-PICA

- rewrite: $\mathbf{X}_{IK \times J} = (\mathbf{C}|\otimes|\mathbf{A})\mathbf{B}^t + \mathbf{E}$
- can be treated as a 2-stage estimation problem:
  1. PICA estimation of B from $\mathbf{X}_{IK \times J}$ by estimating M_c as the mixing matrix

Tensor-PICA

- rewrite: $\mathbf{X}_{IK\times J} = (\mathbf{C}|\otimes|\mathbf{A})\mathbf{B}^t + \mathbf{E}$
- can be treated as a 2-stage estimation problem:

1. PICA estimation of $\mathbf{B}$ from $\mathbf{X}_{IK\times J}$ by estimating $\mathbf{M}_c$ as the mixing matrix
2. rank-1 Eigen-decomposition of each column $\mathbf{M}_c^r$, reshaped into a $I \times K$ matrix, in order to find underlying factors such that $\mathbf{M}_c = (\mathbf{C}|\otimes|\mathbf{A})$

Tensor-PICA

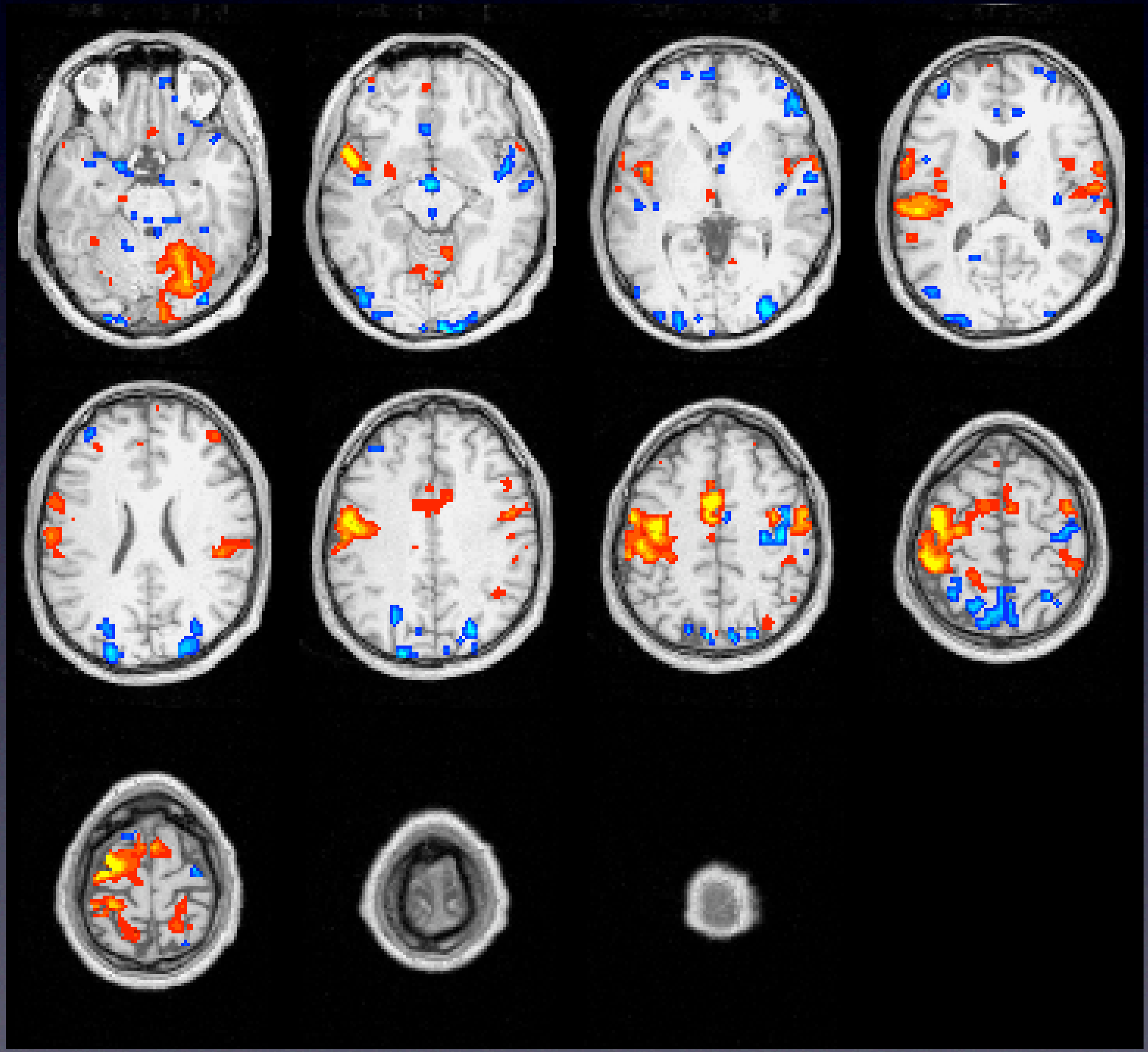
- rewrite: $\mathbf{X}_{IK \times J} = (\mathbf{C}|\otimes|\mathbf{A})\mathbf{B}^t + \mathbf{E}$
- can be treated as a 2-stage estimation problem:

1. PICA estimation of $\mathbf{B}$ from $\mathbf{X}_{IK \times J}$ by estimating $\mathbf{M}_c$ as the mixing matrix
2. rank-1 Eigen-decomposition of each column $\mathbf{M}_c^r$, reshaped into a $I \times K$ matrix, in order to find underlying factors such that $\mathbf{M}_c = (\mathbf{C}|\otimes|\mathbf{A})$

- jointly estimate modes which describe signal in the temporal/ spatial and subject domain

Tensor-ICA

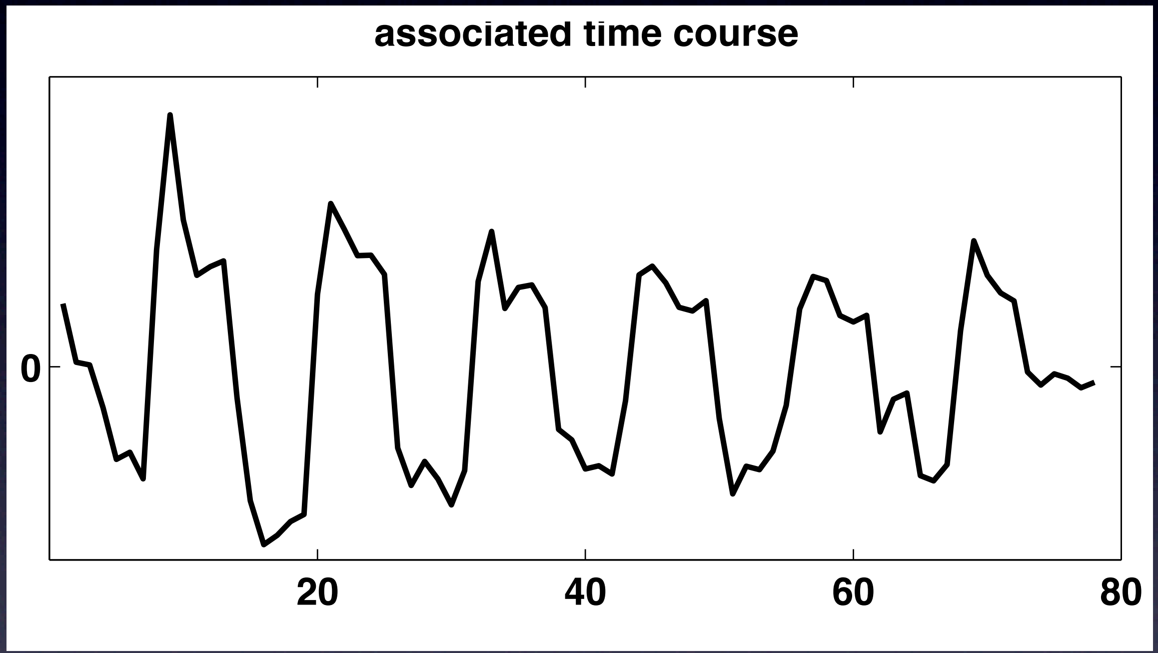
- Higher-level GLM designs are even more simplistic (e.g. assumption of constant within group activation strength)
- 10 sessions under motor paradigm (right index finger tapping)
- Group-level GLM results (mixed-effects)

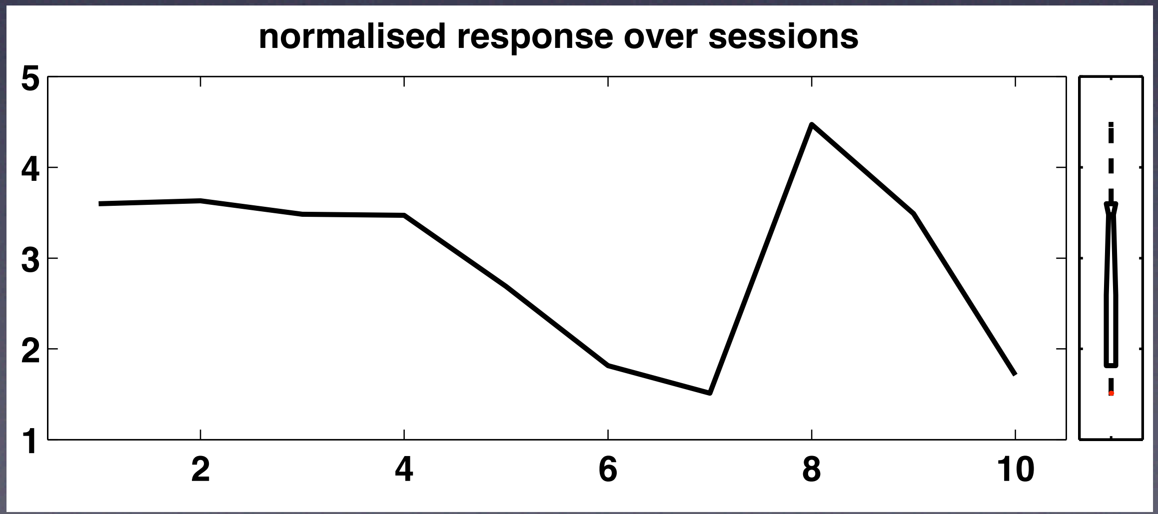


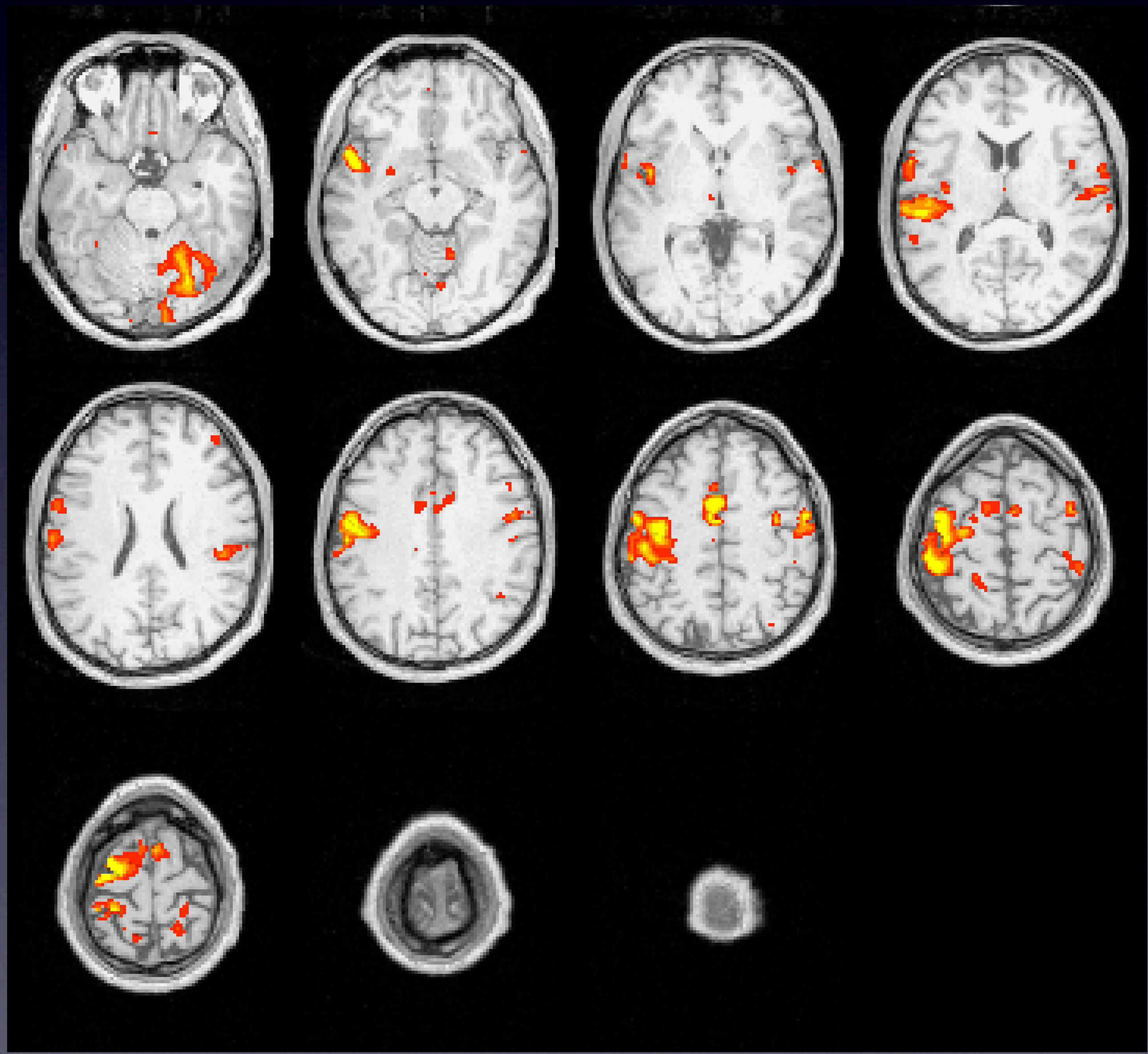
McGonigle et al. NI 2000

tensor-PICA: example

- tensor-PICA map

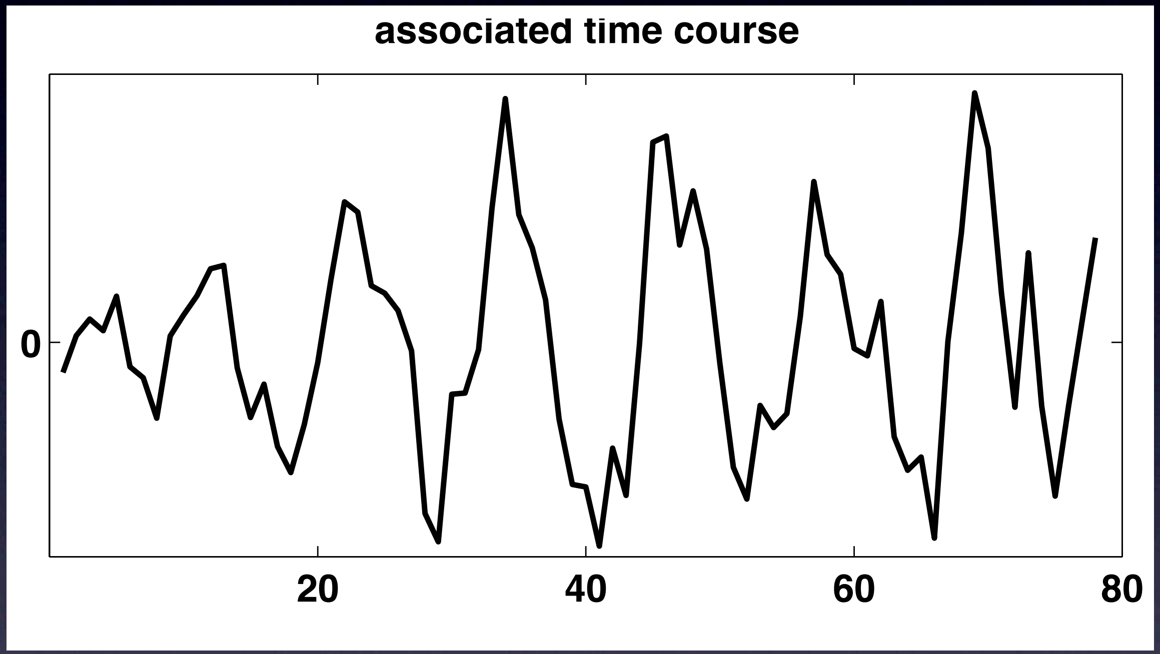


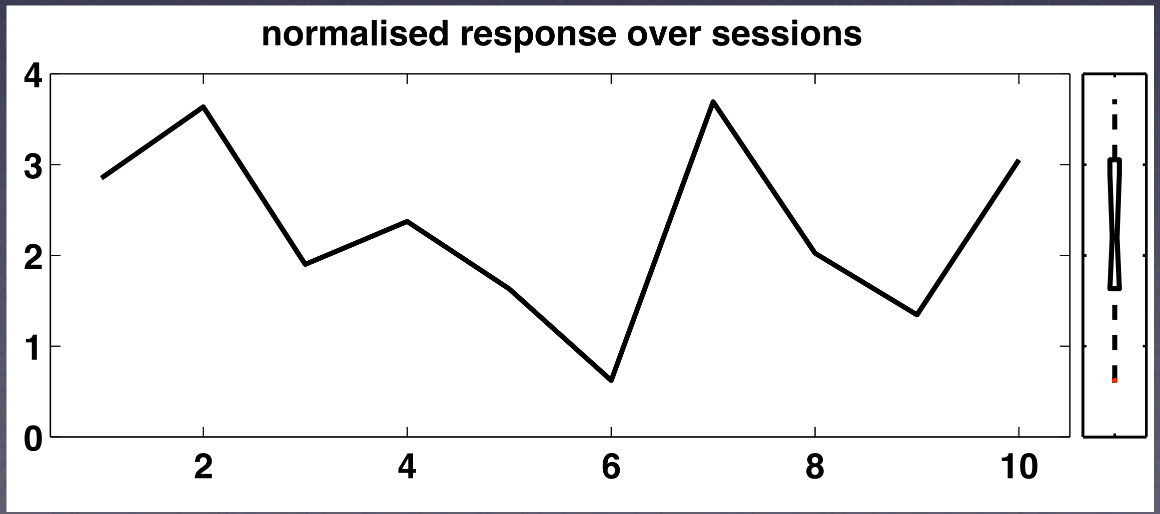


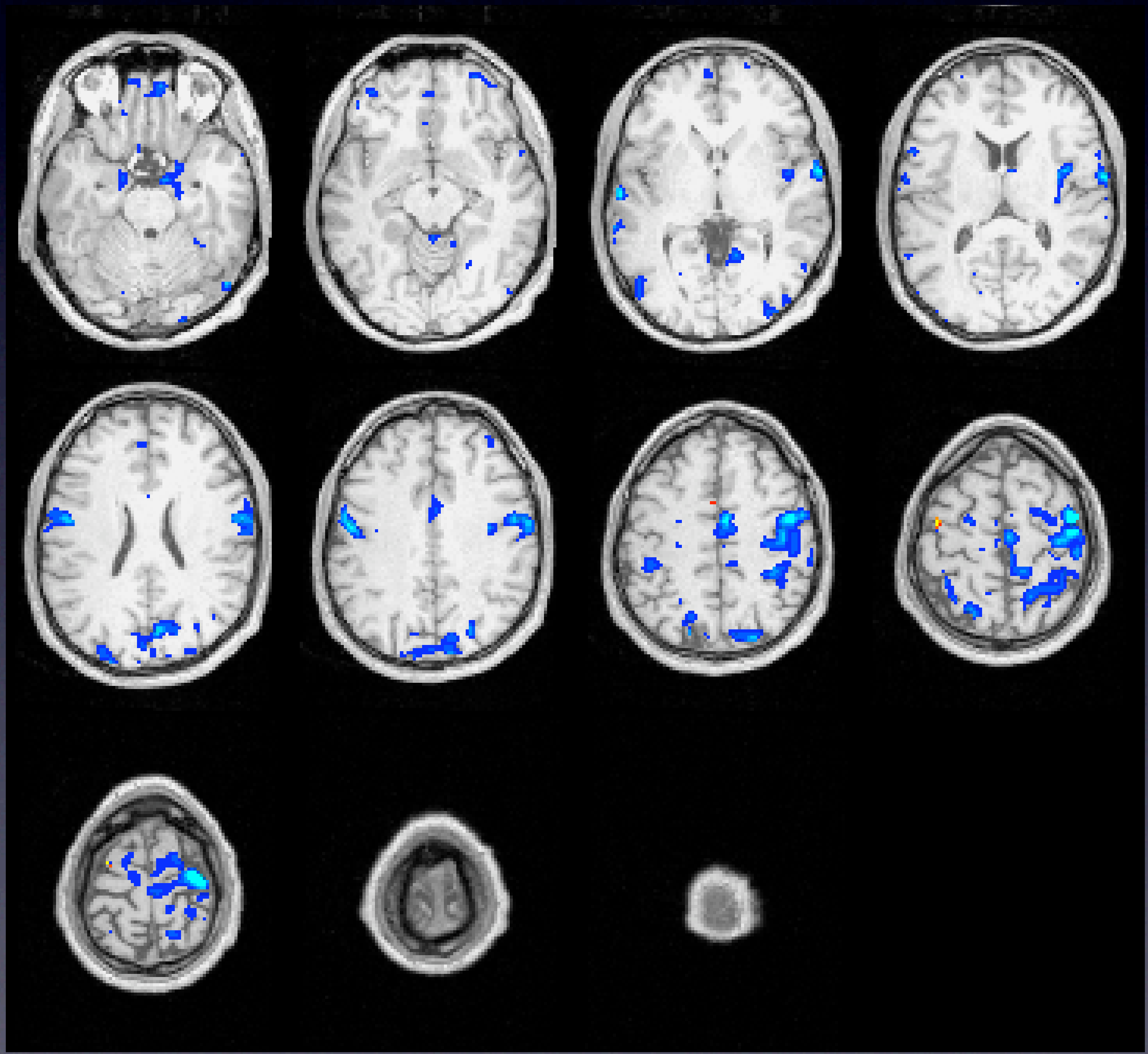


tensor-PICA: example

- estimated 'de-activation'

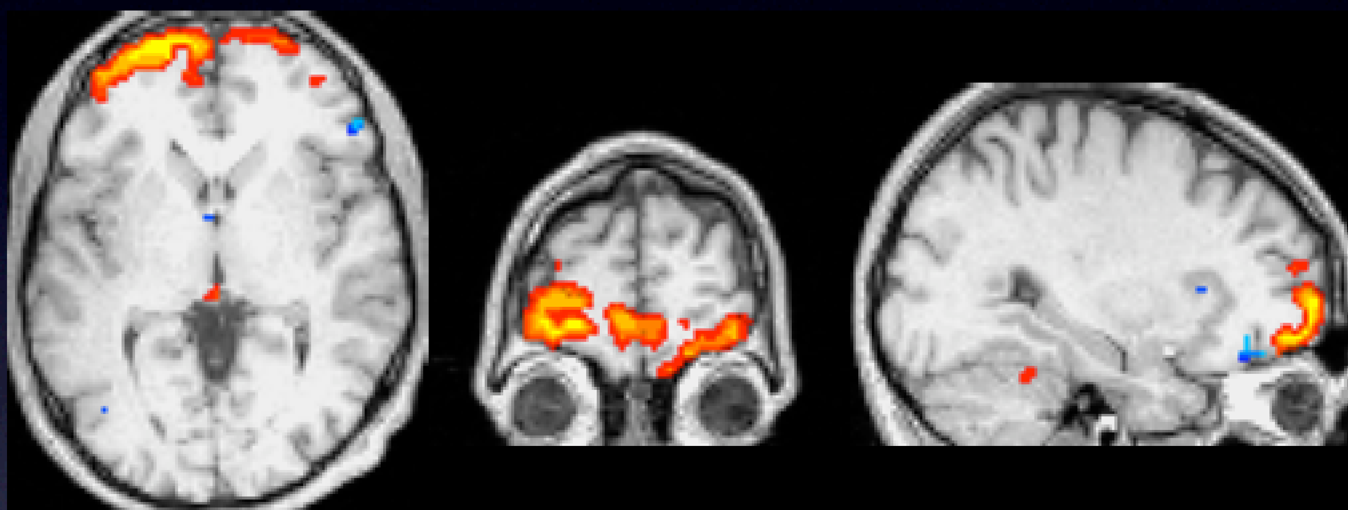




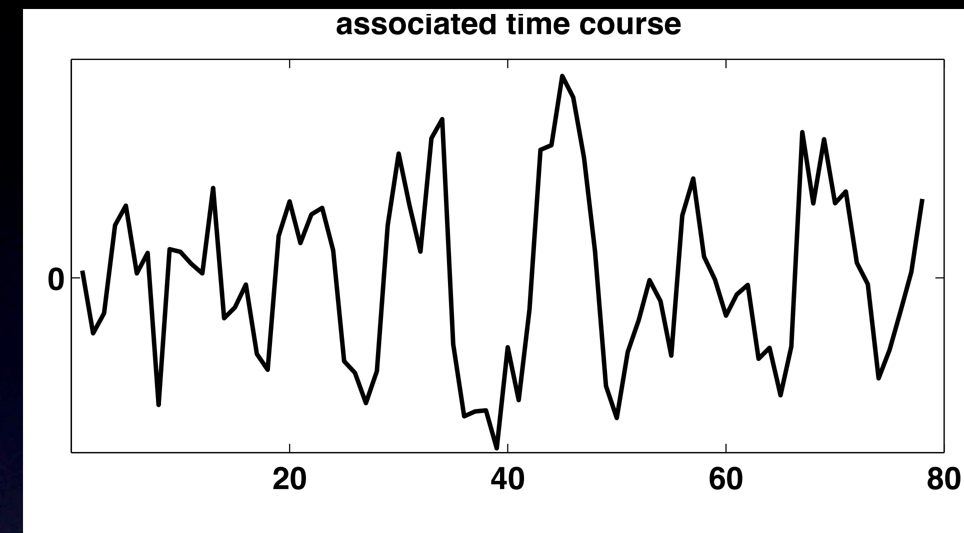


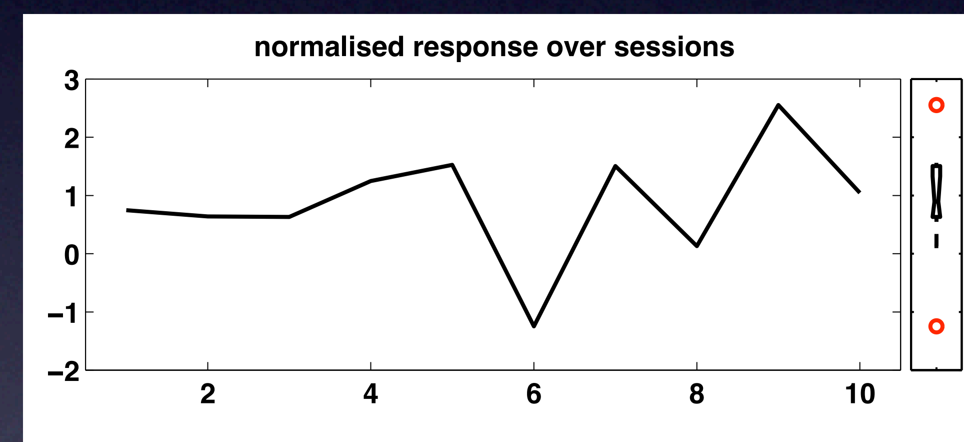
tensor-PICA: example

- residual stimulus correlated motion



tensor-PICA map





Z-stats map (session 9)

Group-GLM (ME) Z-stats map