Modelling temporal structure

(in noise *and* signal)

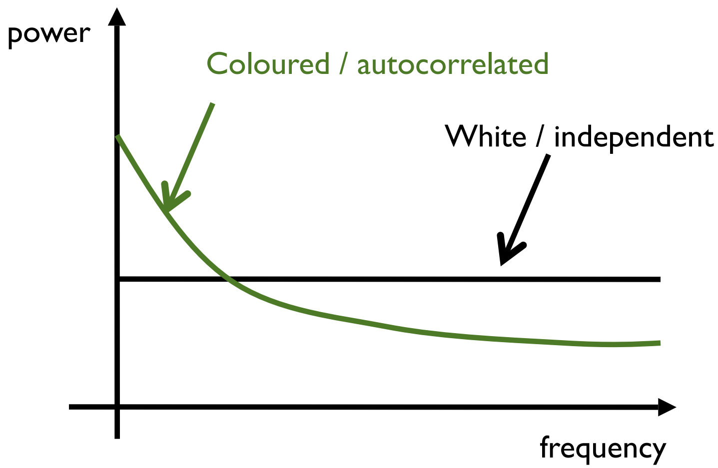
Mark Woolrich, Christian Beckmann*, Salima Makni & Steve Smith
FMRIB, Oxford *Imperial/FMRIB

- temporal noise: modelling temporal autocorrelation

- temporal signal: FLOBS HRF optimal basis functions
- temporal signal: HRF deconvolution

- spatiotemporally structured signal / noise: ICA
- "functional grand-plan": integrating ICA+GLM

Non-independent/Autocorrelation/ Coloured FMRI noise



Even after high-pass filtering, FMRI noise has extra power at low frequencies (positive autocorrelation or temporal smoothness)

Uncorrected, this causes:

- biased stats (increased false positives)
- decreased sensitivity

FMRIB's Improved Linear Modelling (FILM)

- FILM is used to fit the GLM voxel-wise in FEAT
- Deals with the autocorrelation ***locally*** and uses prewhitening

FILM estimates autocorrelation by looking at the residuals of the GLM fit:

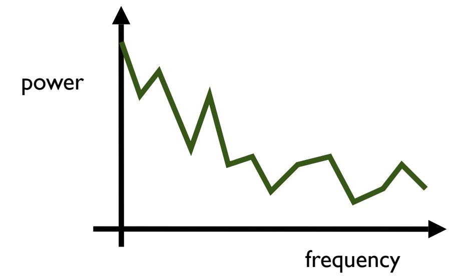
$$Y = X\beta + \epsilon$$

$$\text{residuals} = Y - X\widehat{\beta}$$

FMRIB's Improved Linear Modelling (FILM)

1) Fit the GLM and estimate the autocorrelation on the residuals

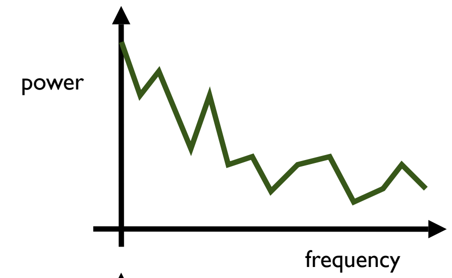
Power vs. freq in the residuals



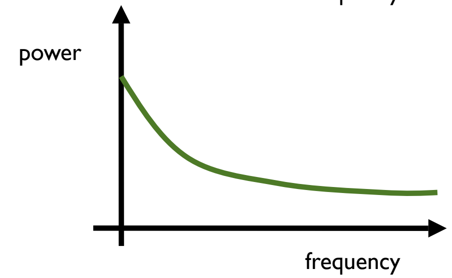
FMRIB's Improved Linear Modelling (FILM)

Power vs. freq in the residuals

1) Fit the GLM and estimate the autocorrelation on the residuals

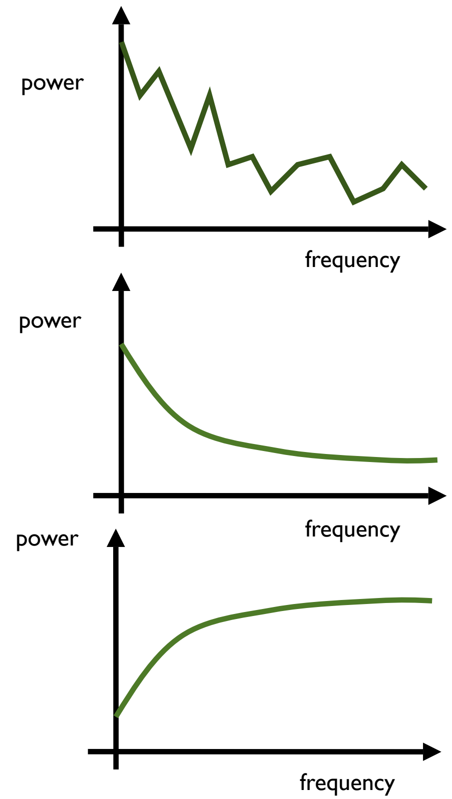


2) Spatially and spectrally smooth the data



FMRIB's Improved Linear Modelling (FILM)

Power vs. freq in the residuals



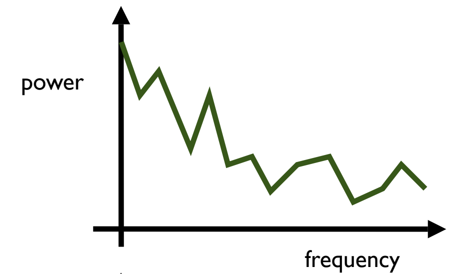
1) Fit the GLM and estimate the autocorrelation on the residuals

2) Spatially and spectrally smooth the data

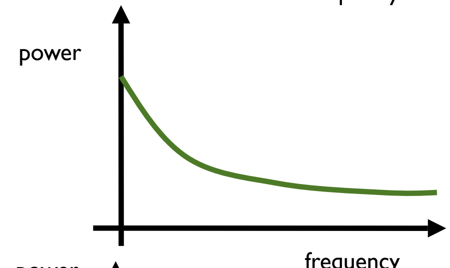
3) Construct prewhitening filter to "undo" autocorrelation

FMRIB's Improved Linear Modelling (FILM)

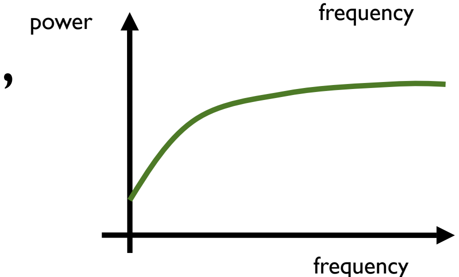
1) Fit the GLM and estimate the autocorrelation on the residuals



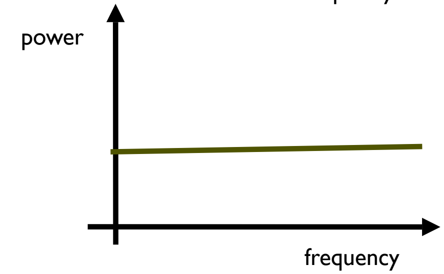
2) Spatially and spectrally smooth the data



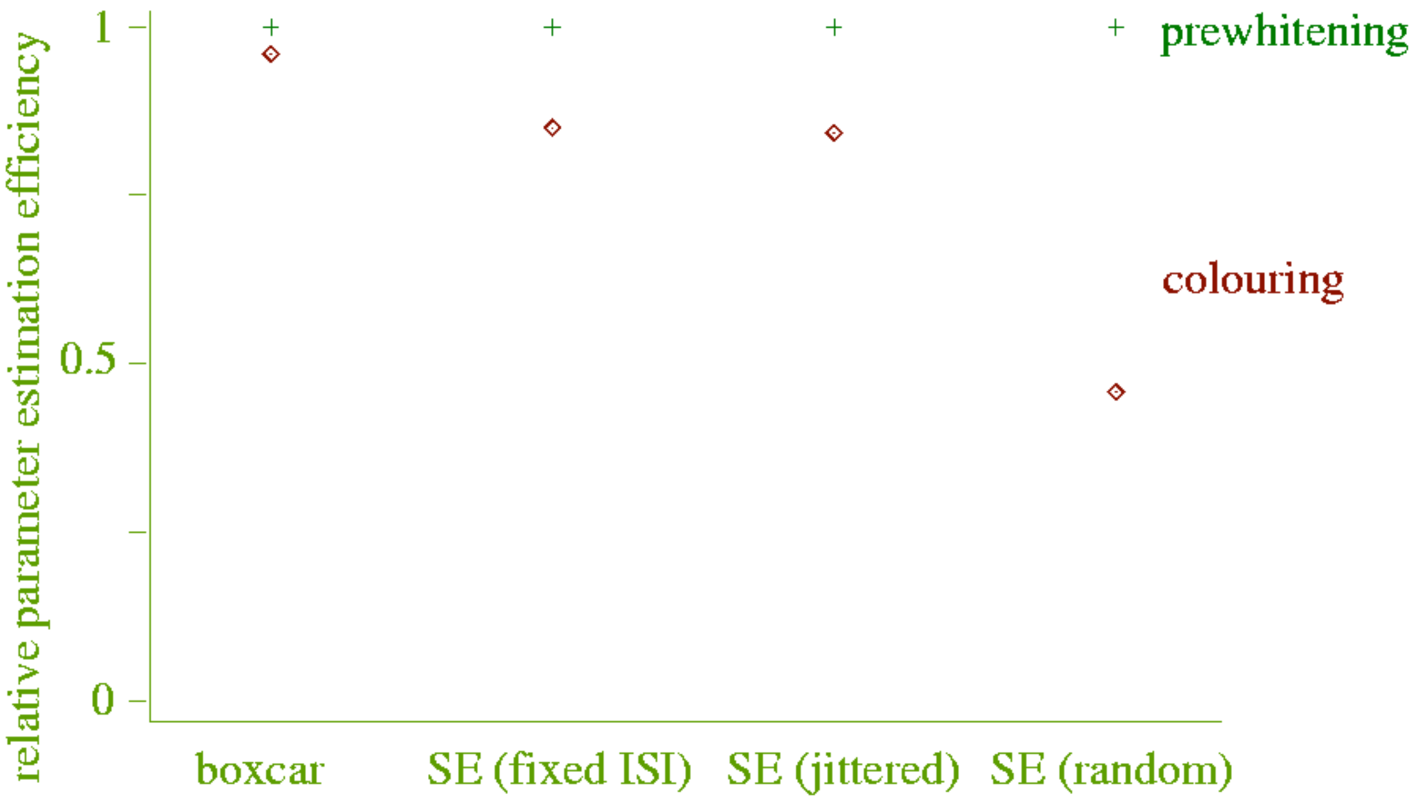
3) Construct prewhitening filter to "undo" autocorrelation



4) Apply filter to data ***and*** design matrix and refit

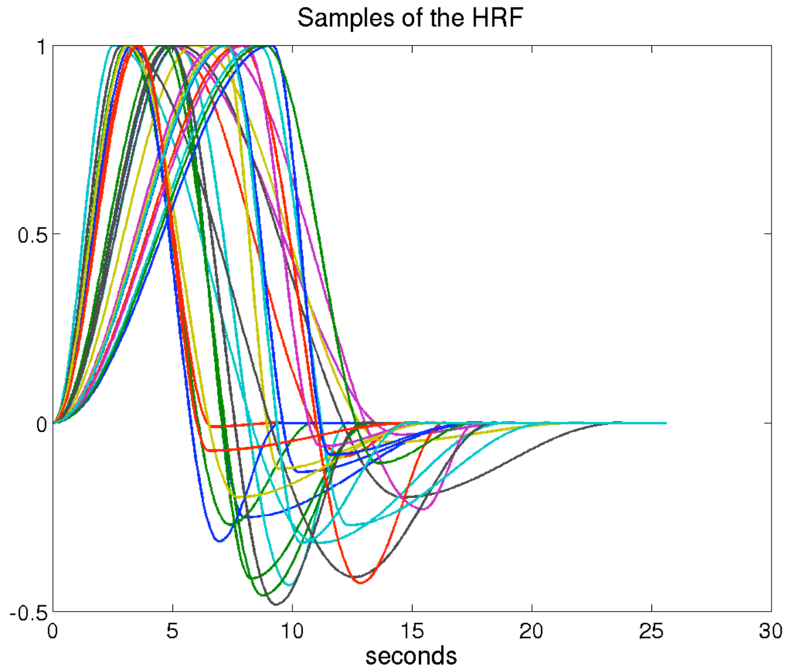


FMRIB's Improved Linear Modelling (FILM)



Dealing with Variations in Haemodynamics

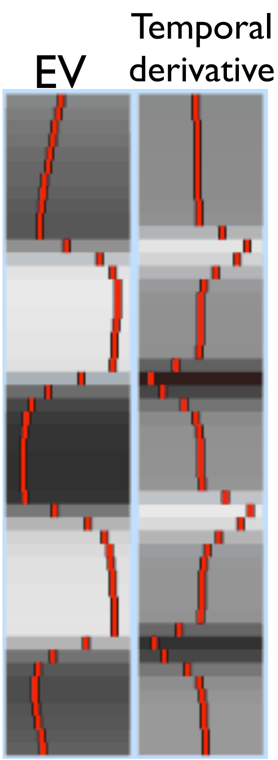
- The haemodynamic responses vary between subjects and areas of the brain
- How do we allow haemodynamics to be flexible but remain plausible?

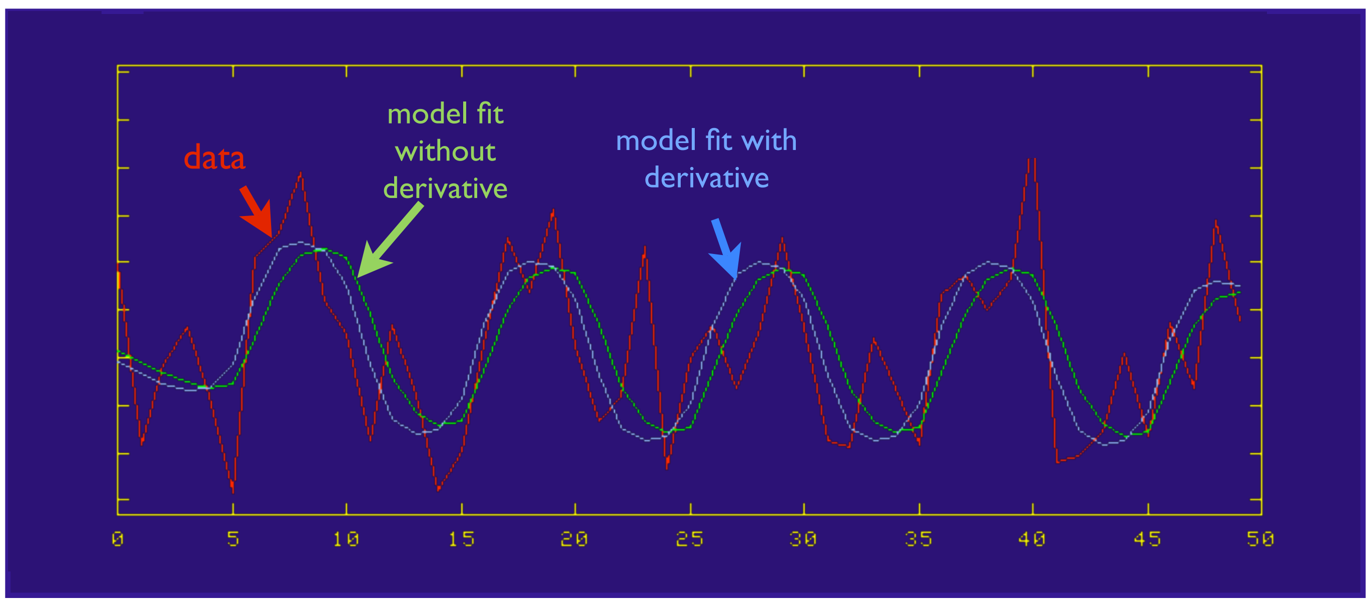


Reminder: the haemodynamic response function (HRF) describes the BOLD response to a short burst of neural activity

Temporal Derivatives

- A very simple approach to providing HRF variability is to include (alongside each EV) the EV temporal derivative
- Including the temporal derivative of an EV allows for a small shift in time of that EV
- This is based upon a first order Taylor series expansion

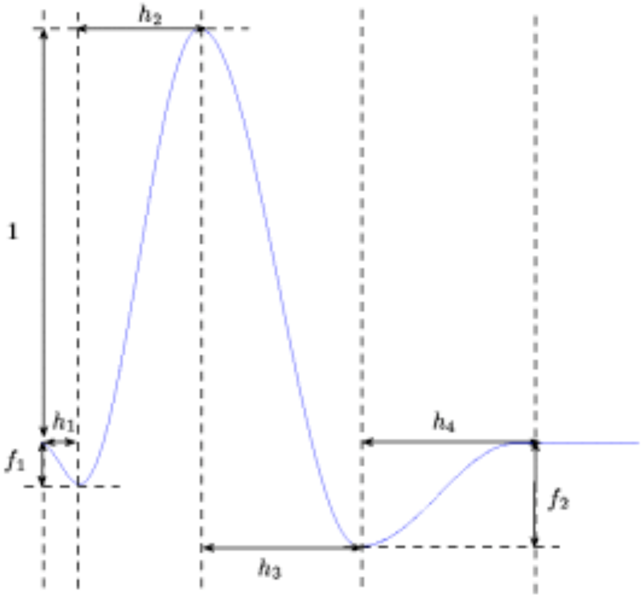




Using Parameterised HRFs

- We need to allow flexibility in the shape of the fitted HRF

Parameterise HRF shape and fit shape parameters to the data

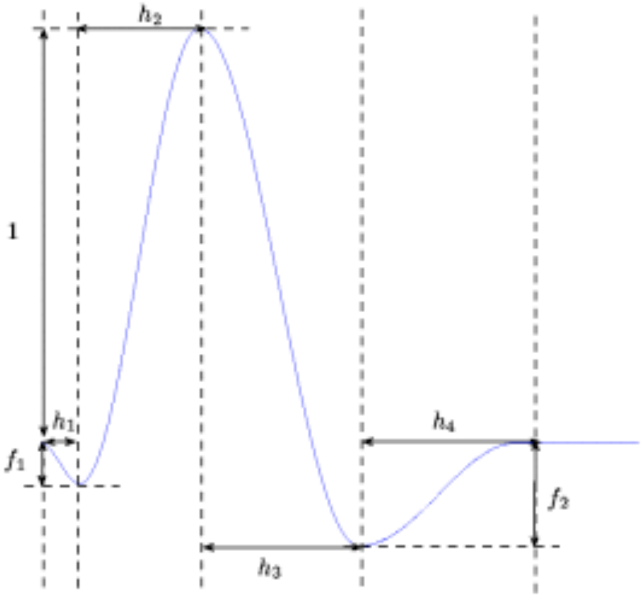


Needs nonlinear fitting - HARD

Using Basis Sets

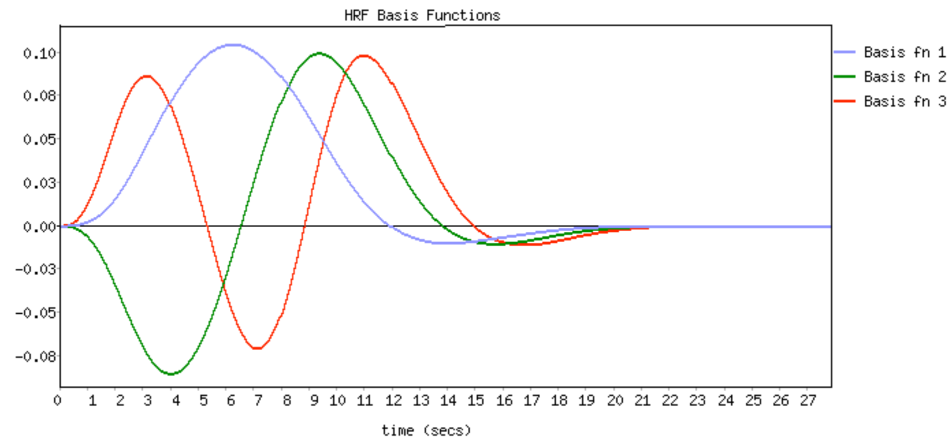
- We need to allow flexibility in the shape of the fitted HRF

Parameterise HRF shape and fit shape parameters to the data



Needs nonlinear fitting - HARD

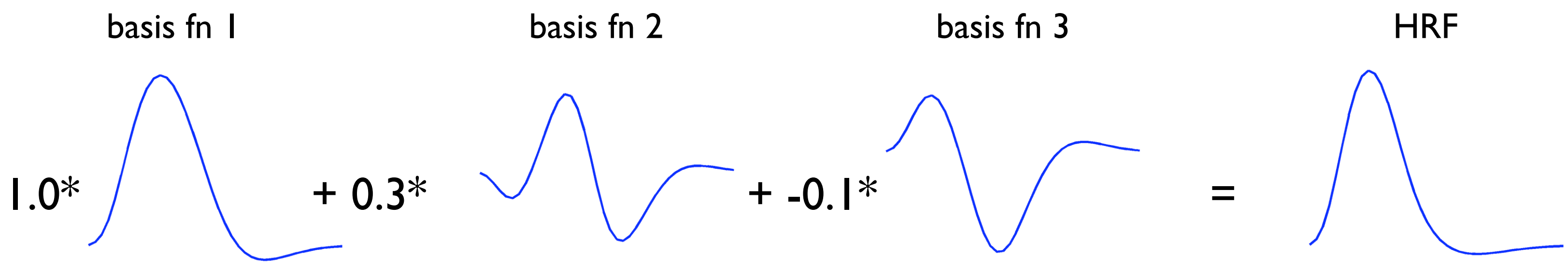
We can use **linear basis sets** to span the space of expected HRF shapes



Linear fitting (use GLM) - EASY

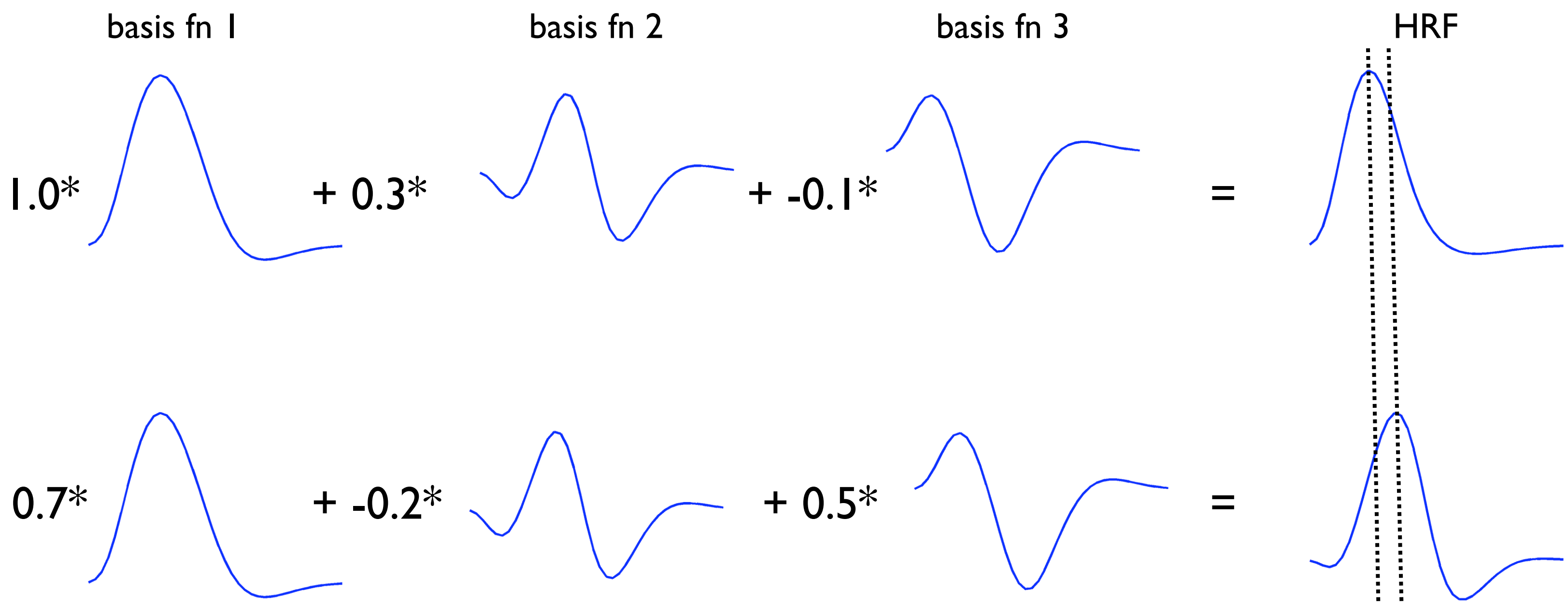
How do HRF Basis Sets Work?

Different linear combinations of the basis functions can be used to create different HRF shapes



How do HRF Basis Sets Work?

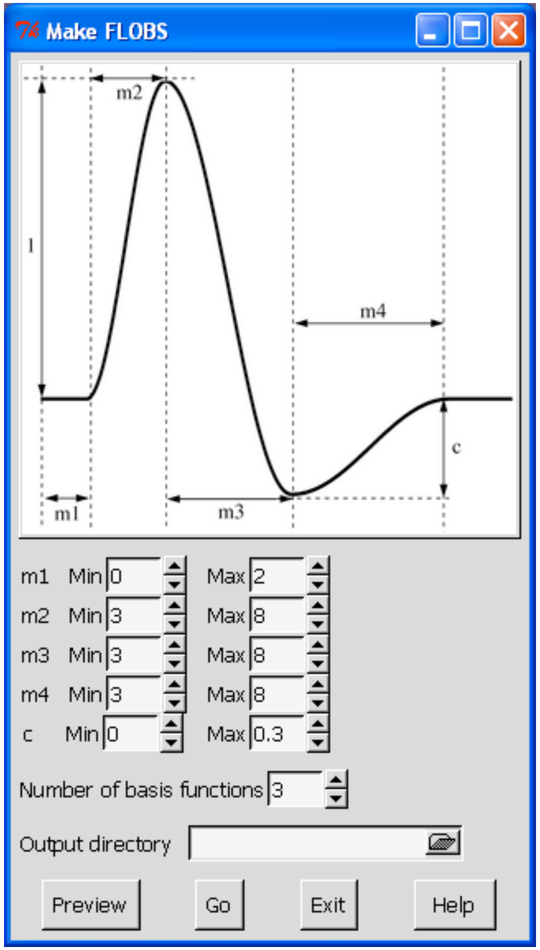
Different linear combinations of the basis functions can be used to create different HRF shapes



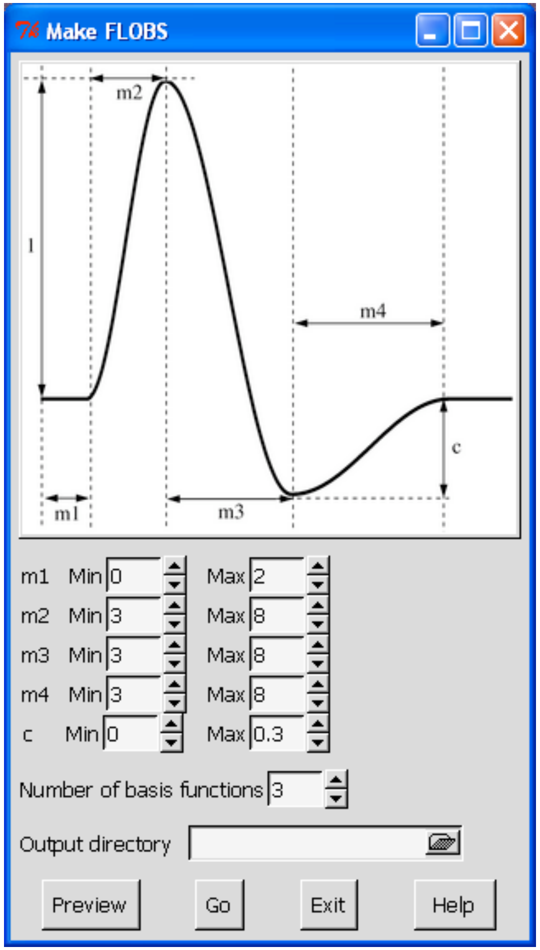
FMRIB's Linear Optimal Basis Set (FLOBS)

Using FLOBS we can:

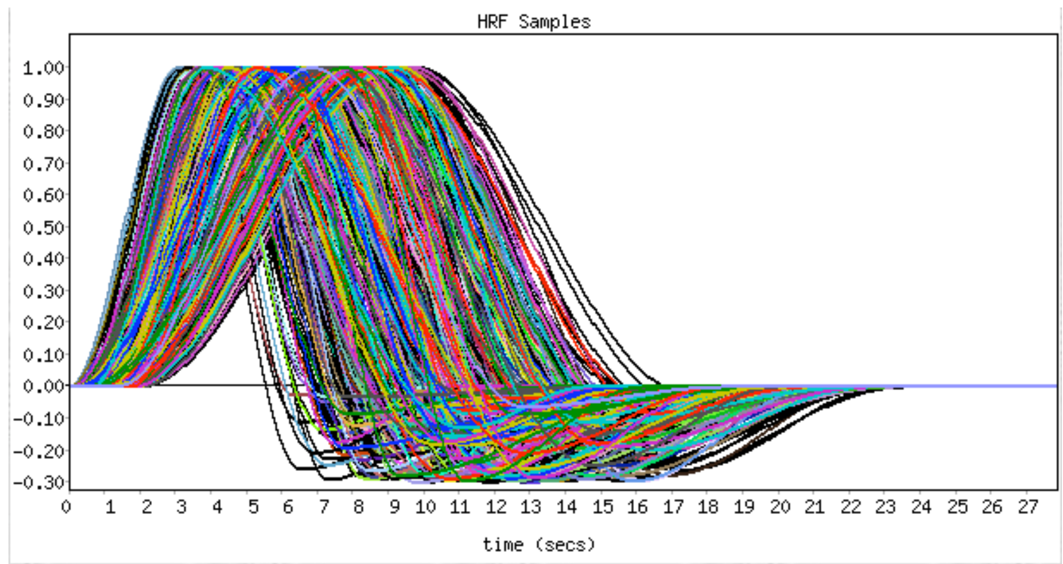
- Specify a priori expectations of parameterised HRF shapes
- Generate an appropriate basis set



Generating FLOBs

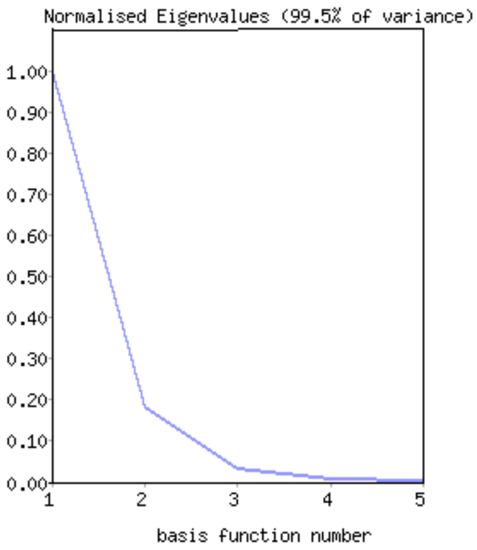


(1) Take samples of the HRF

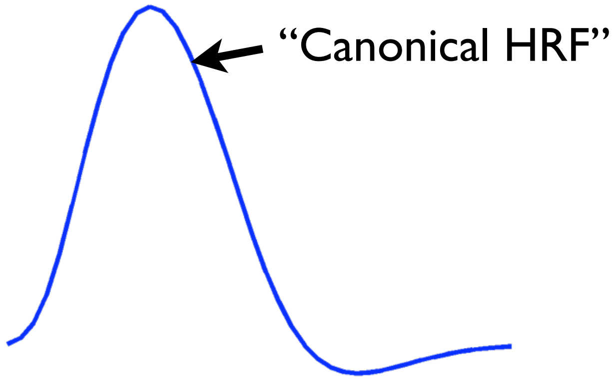


Generating FLOBs

(2) Perform SVD

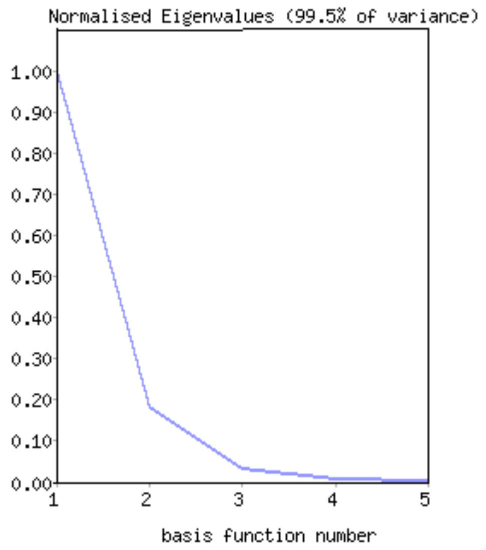


(3) Select the top eigenvectors as the optimal basis set

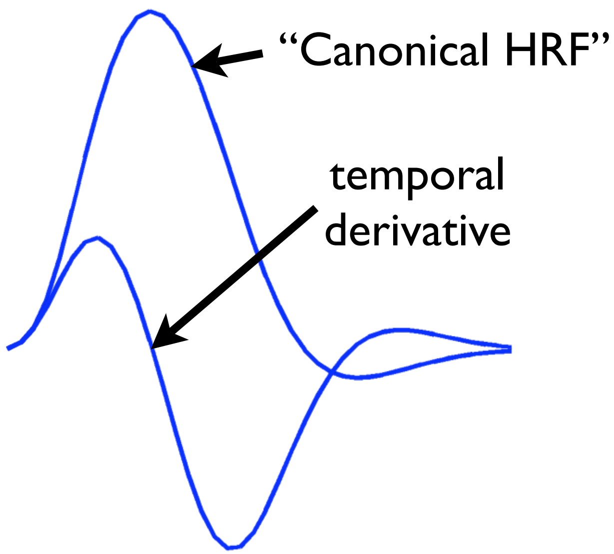


Generating FLOBs

(2) Perform SVD

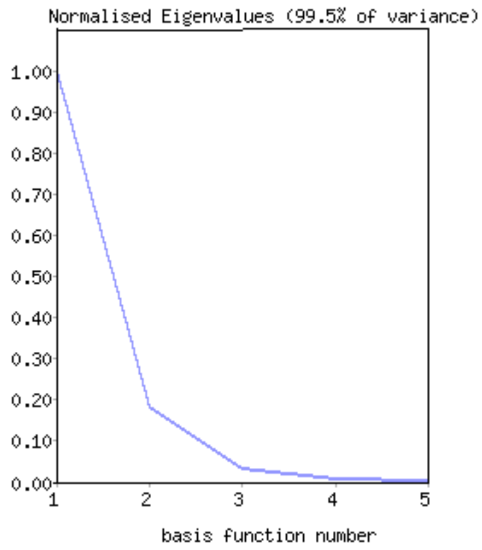


(3) Select the top eigenvectors as the optimal basis set

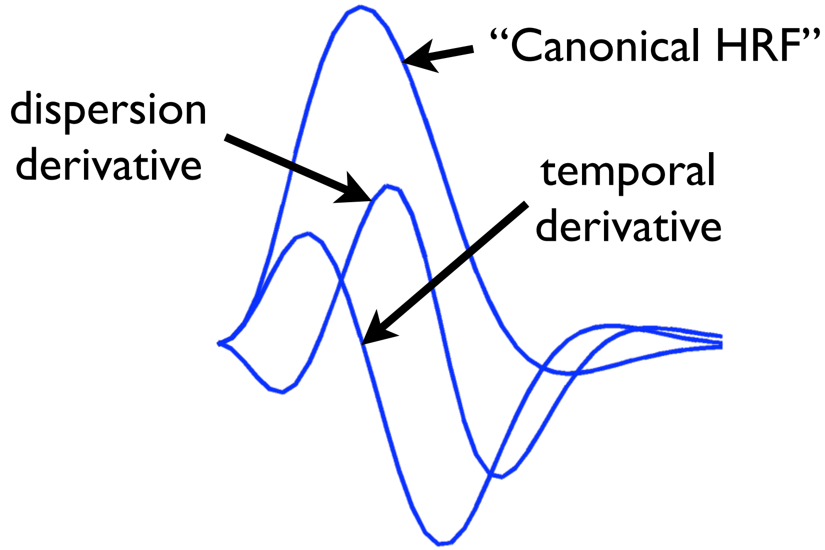


Generating FLOBs

(2) Perform SVD

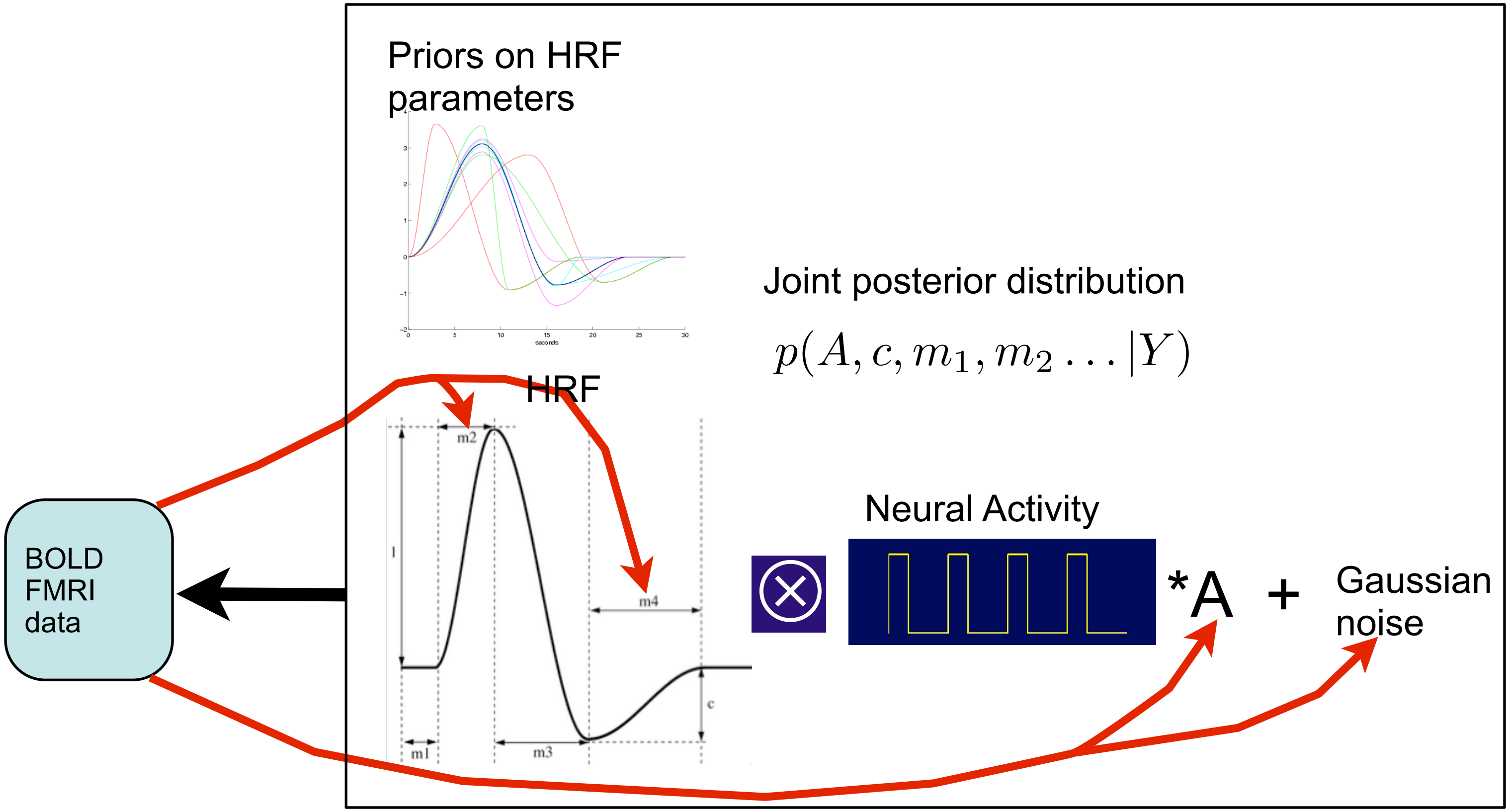


(3) Select the top eigenvectors as the optimal basis set



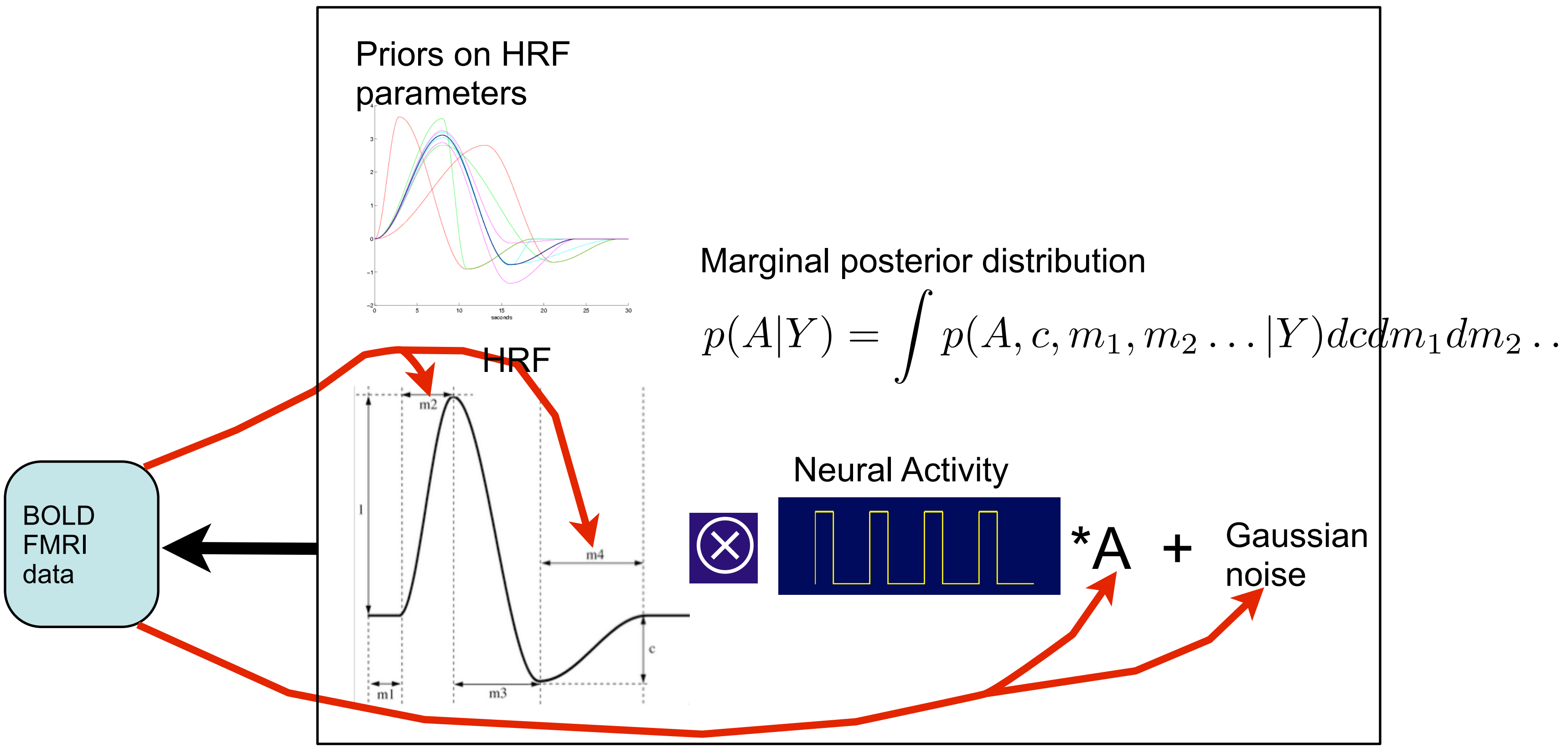
The resulting basis set can then be used in FEAT

Bayesian Inference



Woolrich et al. , TMI, 2004

Bayesian Inference



Infer using MCMC

Woolrich et al. , TMI, 2004

Temporal deconvolution of FMRI timecourses

- Inputs are raw paradigm (stimulation and "modulation") timecourses
- Model based on Bilinear Dynamical Systems (Penny 2005), where modulatory input changes neural response to stimulation
- What's new:
  - estimate HRF from data
  - full Bayesian inference on model, using VB

Makni, NeuroImage 2008

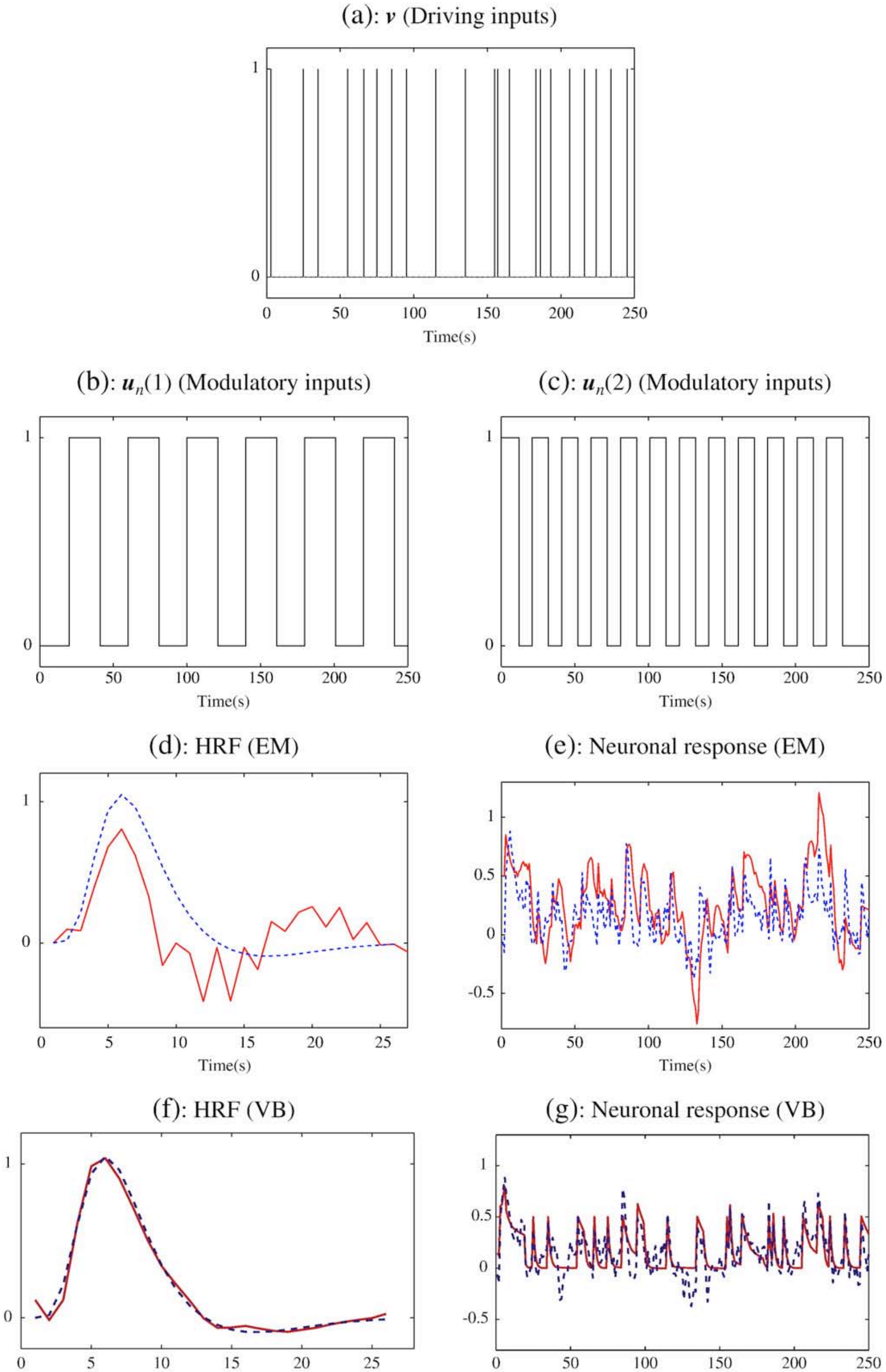
$$s_n = \left(a + \boldsymbol{b}^T \boldsymbol{u}_n\right) s_{n-1} + \boldsymbol{d}^T \boldsymbol{v}_n + w_n,$$
$$\boldsymbol{x}_n = \left[s_n, s_{n-1}, \ldots, s_{n-L+1}\right]^T,$$
$$y_n = \boldsymbol{h}^T \boldsymbol{x}_n + e_n.$$

(1) Initialisation: choose $\boldsymbol{s}^0$ and Θ^0.

(2) Iteration k:

- Update $q(a)$ using C.2, a is the intrinsic connection coefficient
- Update $q(\boldsymbol{b})$ using C.3, $\boldsymbol{b}$ is the modulatory coefficient
- Update $q(\boldsymbol{d})$ using C.4, $\boldsymbol{d}$ is the driving coefficient
- Update $q(\boldsymbol{h})$ using C.5, $\boldsymbol{h}$ is the HRF
- Update $q(\boldsymbol{s})$ using C.6, $\boldsymbol{s}$ is the neuronal response
- Update $q(s_1)$ using C.7, s_1 is the neuronal response at $t = 1$
- Update $q(\phi_w^{-1})$ using C.8, ϕ_w^{-1} is the inverse state noise precision
- Update $q(\phi_e^{-1})$ using C.9, ϕ_e^{-1} is the inverse space noise precision
- Compute F^k using D.1.

(3) Stop when $F^k - F^{k-1}$ < tolerance value



To do the Bayes, either:

- MCMC (computer takes ages)
- Variational Bayes (maths takes ages)

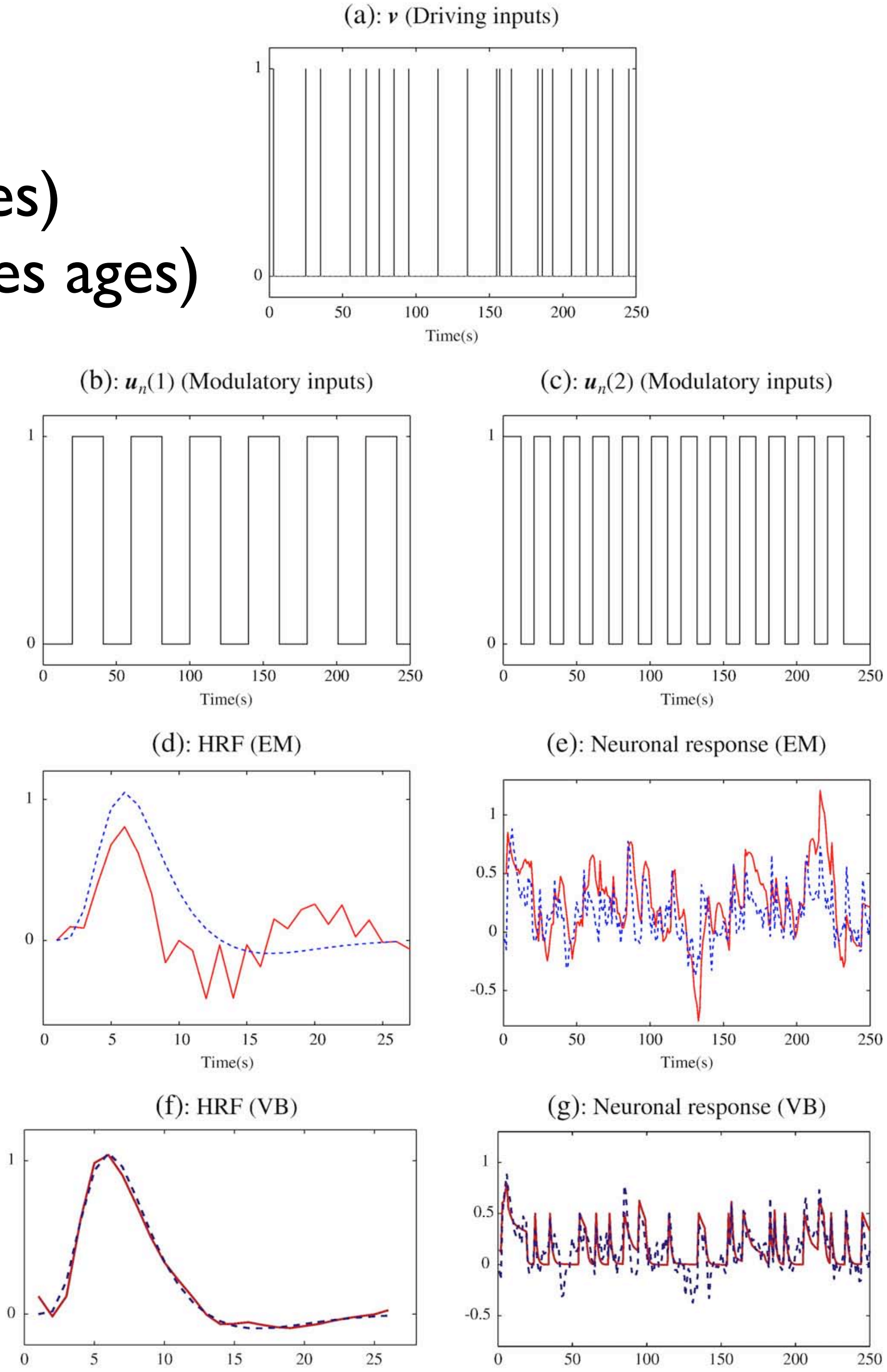
$$s_n = \left(a + \boldsymbol{b}^T \boldsymbol{u}_n\right) s_{n-1} + \boldsymbol{d}^T \boldsymbol{v}_n + w_n,$$
$$\boldsymbol{x}_n = \left[s_n, s_{n-1}, \ldots, s_{n-L+1}\right]^T,$$
$$y_n = \boldsymbol{h}^T \boldsymbol{x}_n + e_n.$$

(1) Initialisation: choose $\boldsymbol{s}^0$ and Θ^0.

(2) Iteration k:

* Update $q(a)$ using C.2, a is the intrinsic connection coefficient
* Update $q(\boldsymbol{b})$ using C.3, $\boldsymbol{b}$ is the modulatory coefficient
* Update $q(\boldsymbol{d})$ using C.4, $\boldsymbol{d}$ is the driving coefficient
* Update $q(\boldsymbol{h})$ using C.5, $\boldsymbol{h}$ is the HRF
* Update $q(\boldsymbol{s})$ using C.6, $\boldsymbol{s}$ is the neuronal response
* Update $q(s_1)$ using C.7, s_1 is the neuronal response at $t = 1$
* Update $q(\phi_w^{-1})$ using C.8, ϕ_w^{-1} is the inverse state noise precision
* Update $q(\phi_e^{-1})$ using C.9, ϕ_e^{-1} is the inverse space noise precision
* Compute F^k using D.1.

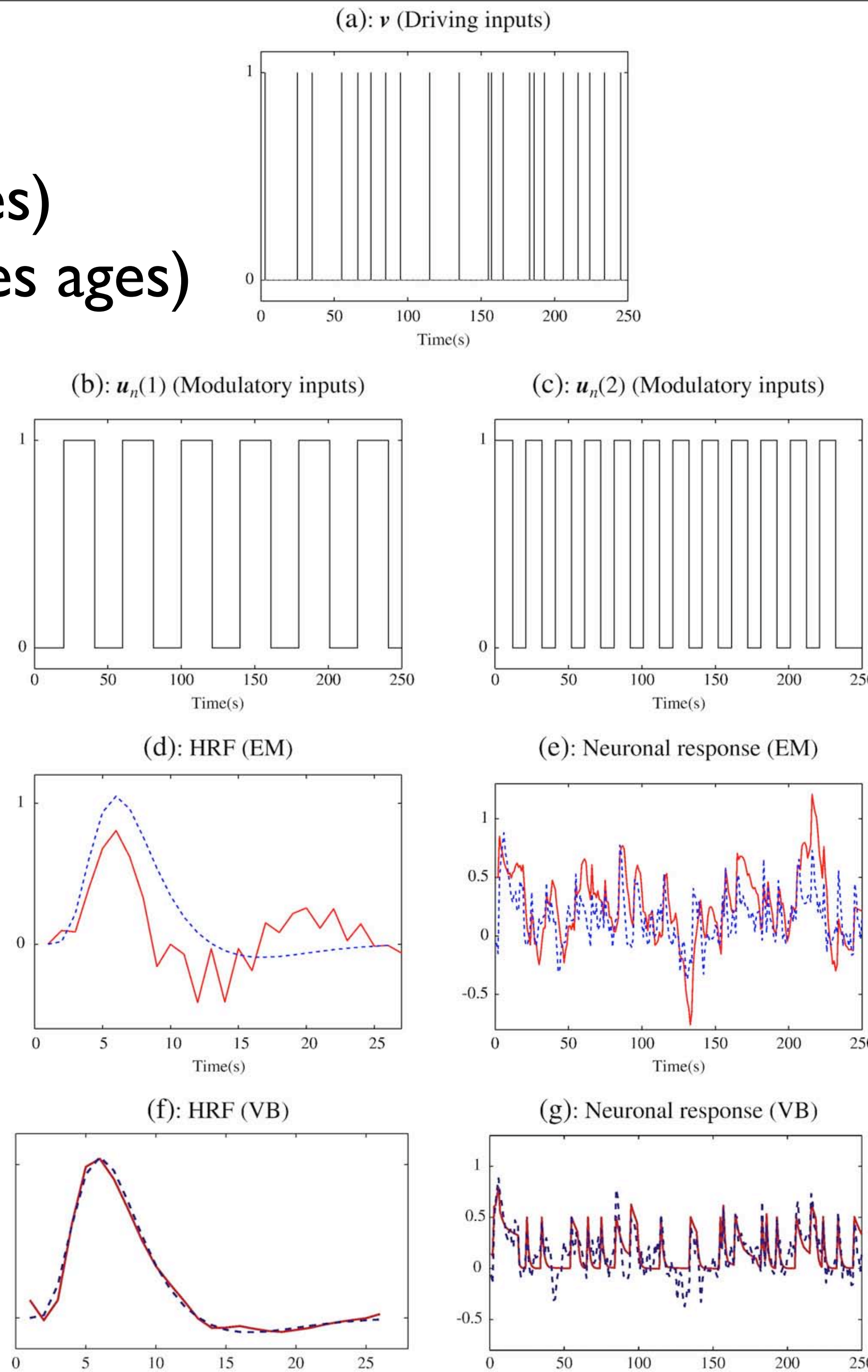
(3) Stop when $F^k - F^{k-1}$ < tolerance value



To do the Bayes, either:

- MCMC (computer takes ages)
- Variational Bayes (maths takes ages)

$$
\begin{aligned}
\mathcal{F} = {} & \frac{\boldsymbol{N}}{2}[\Psi(\tau_e) + \log(\alpha_e) - \log(2\pi)] - \frac{\alpha_e \tau_e}{2}\xi_1 \\
& + \frac{1}{2}\log|\Sigma_a \phi_a^{-1}| + \frac{1}{2}\left[1 - (\Sigma_a + m_a^2)\phi_a^{-1}\right] \\
& + \frac{1}{2}\log|\boldsymbol{\Sigma_b} diag(\boldsymbol{\Phi}_b)^{-1}| + \frac{1}{2} tr \\
& \times \left[\boldsymbol{I} - (\boldsymbol{\Sigma_b} + \boldsymbol{m_b m_b^T})\mathrm{diag}(\boldsymbol{\Phi}_b)^{-1}\right] + \frac{1}{2}\log \\
& \times |\boldsymbol{\Sigma_d}\mathrm{diag}(\boldsymbol{\Phi}_d)^{-1}| + \frac{1}{2} tr\left[\boldsymbol{I} - (\boldsymbol{\Sigma_d} + \boldsymbol{m_d m_d^T}) diag(\boldsymbol{\Phi}_d)^{-1}\right] \\
& \times \frac{1}{2}\sum_{n=1}^{N} \log|\Sigma_{s_n}(\beta)_n^{-1}| + \frac{1}{2}\left[1 - \left(\Sigma_{s_n} + m_{s_n}^2\right)(\beta)_n^{-1}\right] \\
& + \alpha_w^0 \log(\tau_w^0) - \alpha_w \log(\tau_w) + \log\frac{\Gamma(\alpha_w)}{\Gamma(\alpha_w^0)} + (\alpha_w^0 - \alpha_w) \\
& \times (\Psi(\alpha_w) - \log(\tau_w))\alpha_w\left(1 - \frac{\tau_w^0}{\tau_w}\right) + \frac{1}{2}\log \\
& \times |\boldsymbol{\Sigma_h R}^{-1}|\phi_h^{-N_h} + \frac{1}{2}\mathrm{tr}\left[\boldsymbol{I} - (\boldsymbol{\Sigma_h} + \boldsymbol{m_h m_h^T})\boldsymbol{R}^{-1}\phi_h^{-1}\right] \\
& + \alpha_e^0 \log(\tau_e^0) - \alpha_e \log(\tau_e) + \log\frac{\Gamma(\alpha_e)}{\Gamma(\alpha_e^0)} \\
& + (\alpha_e^0 - \alpha_e)[\Psi(\alpha_e) - \log(\tau_e)] + \alpha_e\left(1 - \frac{\tau_e^0}{\tau_e}\right)
\end{aligned}
$$



Model-free Functional Data Analysis

MELODIC

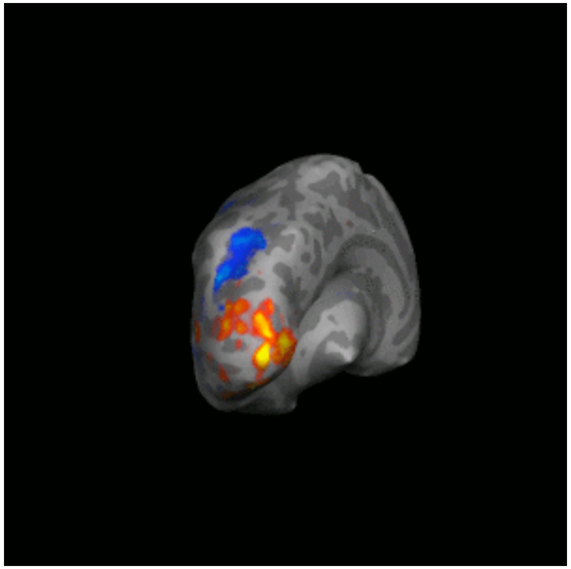
Multivariate Exploratory Linear Optimised Decomposition into Independent Components

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

Model-free Functional Data Analysis

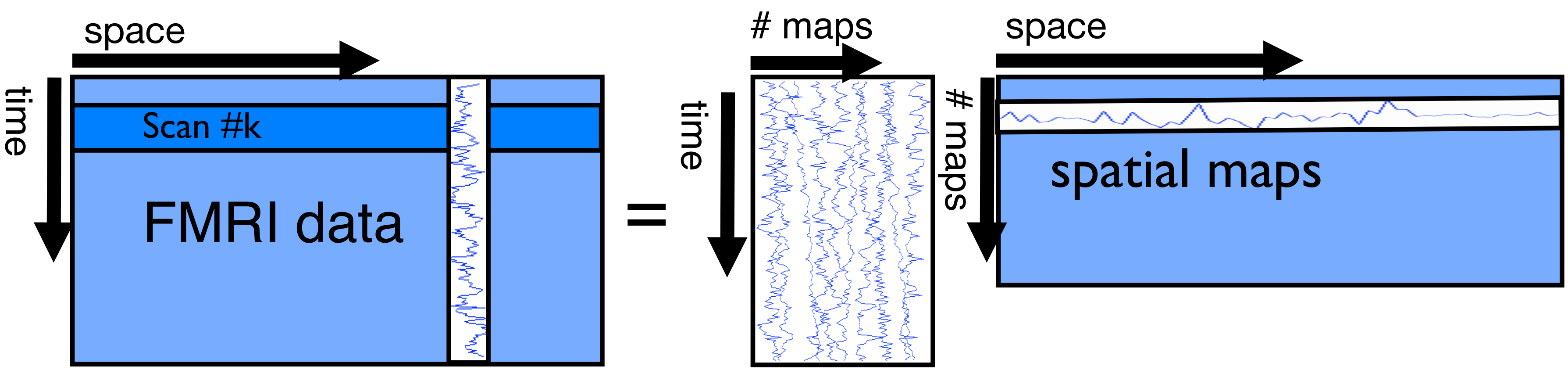
MELODIC
Multivariate Exploratory Linear Optimised Decomposition into Independent Components

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



EDA techniques for FMRI

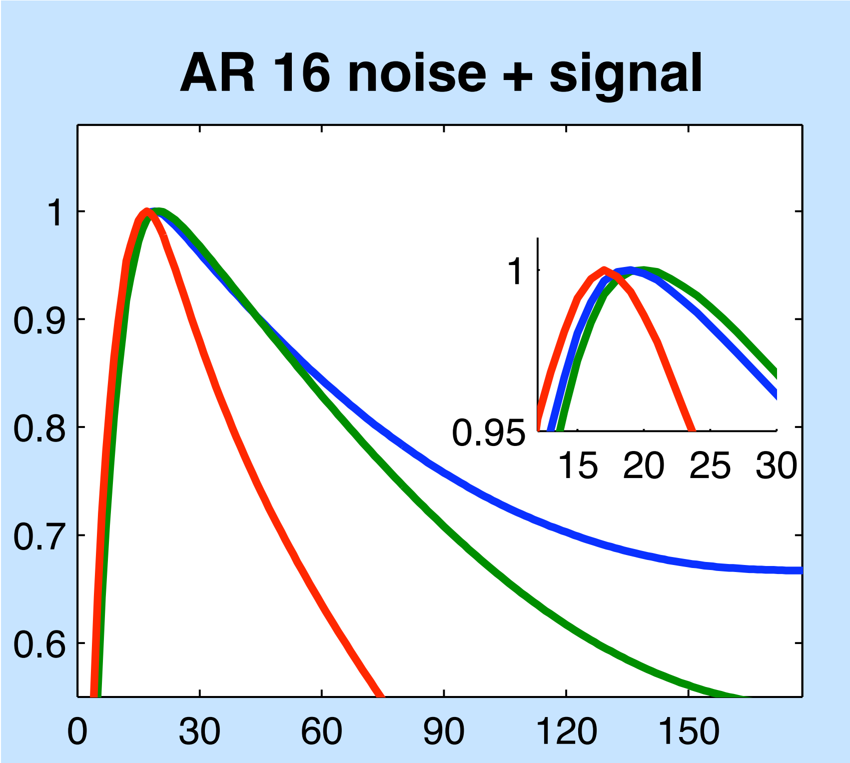
- are mostly multivariate
- often provide a multivariate linear decomposition:



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

Model Order Selection

- can estimate the model order from the Eigenspectrum of the data covariance matrix (corrected using Wishart random matrix theory)
- approximate the Bayesian evidence for the model order for a probabilistic PCA model (PPCA)

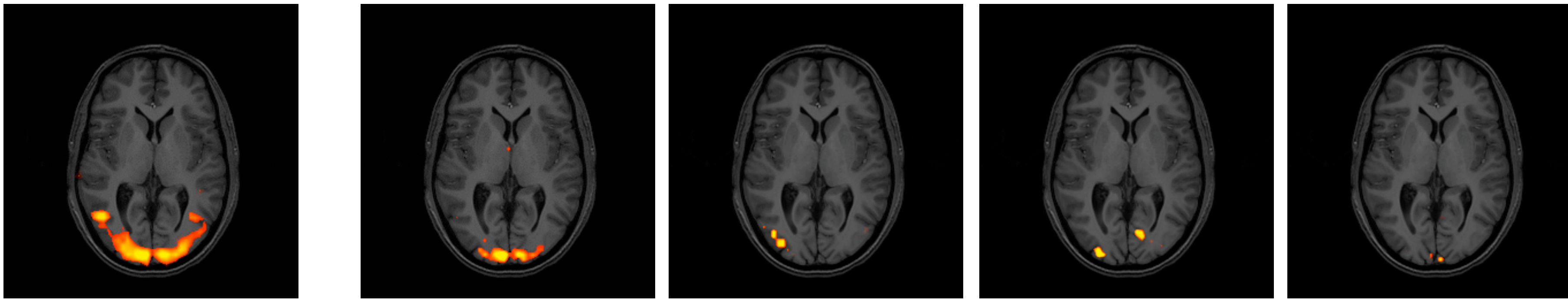


Laplace approximation BIC AIC

Minka, TR 514 MIT Media Lab 2000

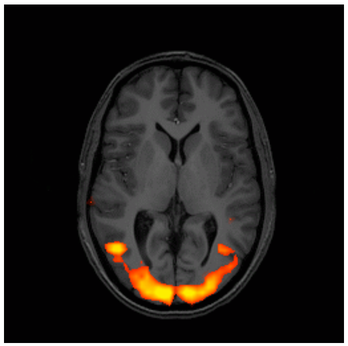
Probabilistic ICA

GLM analysis standard ICA (unconstrained)

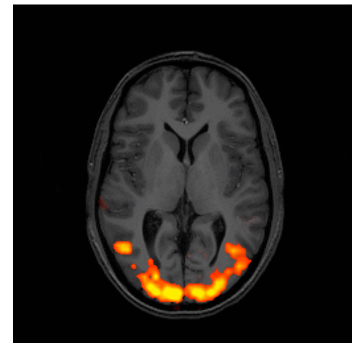


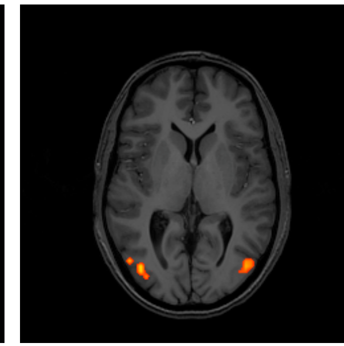
Probabilistic ICA

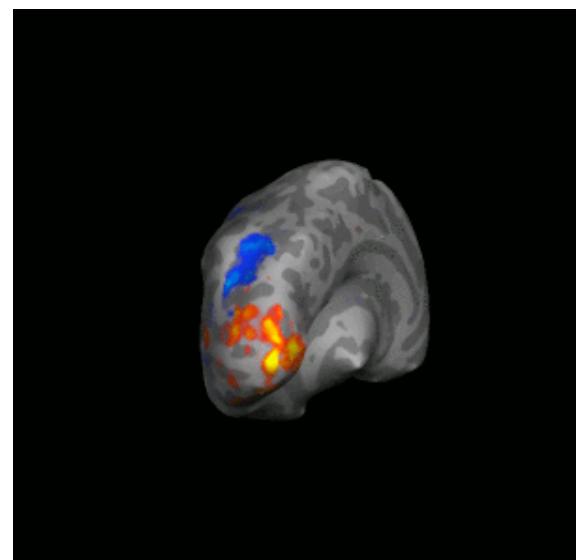
GLM analysis



probabilistic ICA

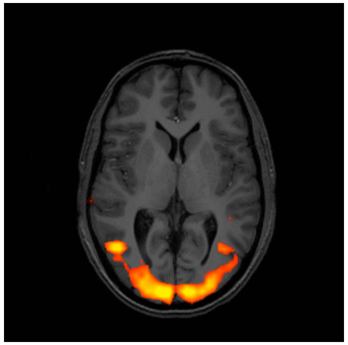




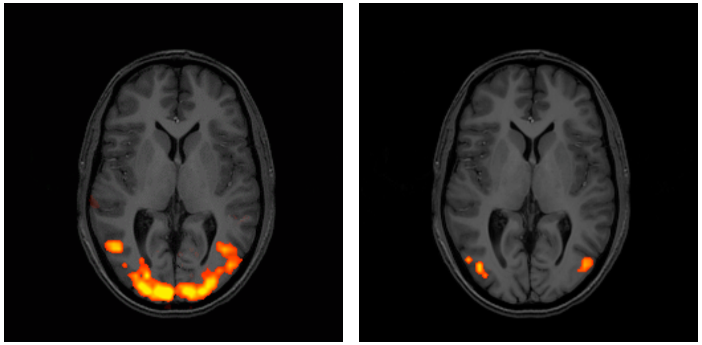


Probabilistic ICA

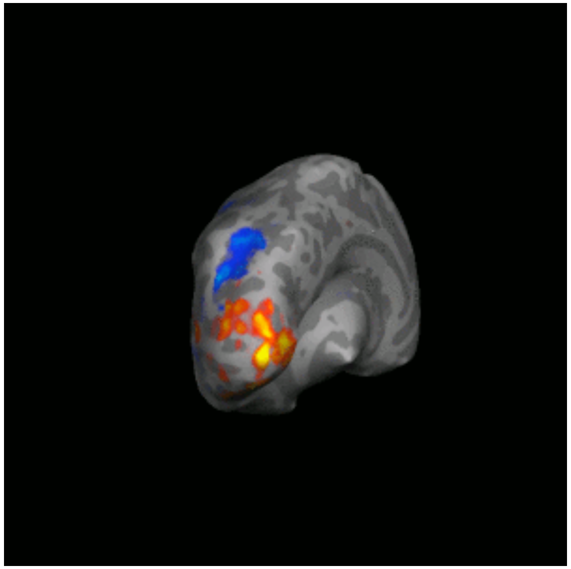
GLM analysis



probabilistic ICA



- designed to address the 'overfitting problem':
  - tries to avoid generation of 'spurious' results
  - high spatial sensitivity and specificity

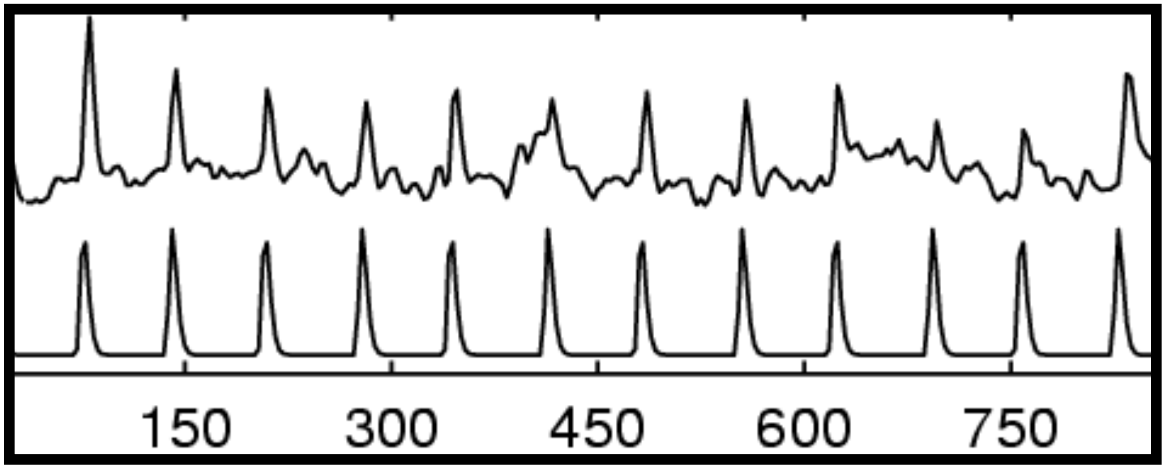


Applications

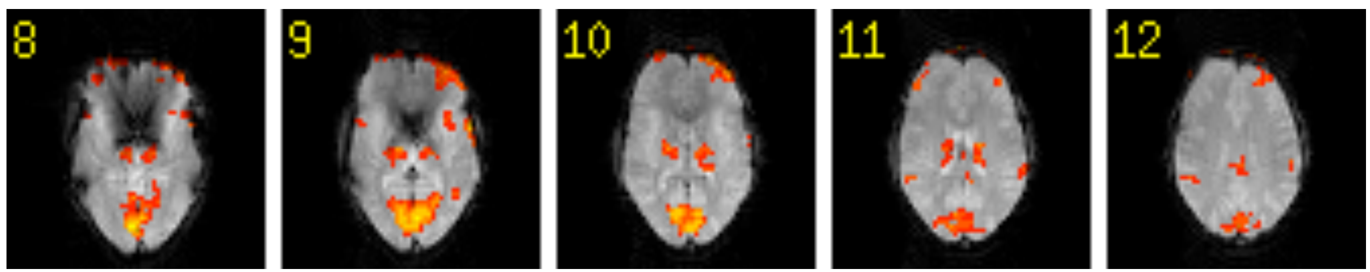
EDA techniques can be useful to

- 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

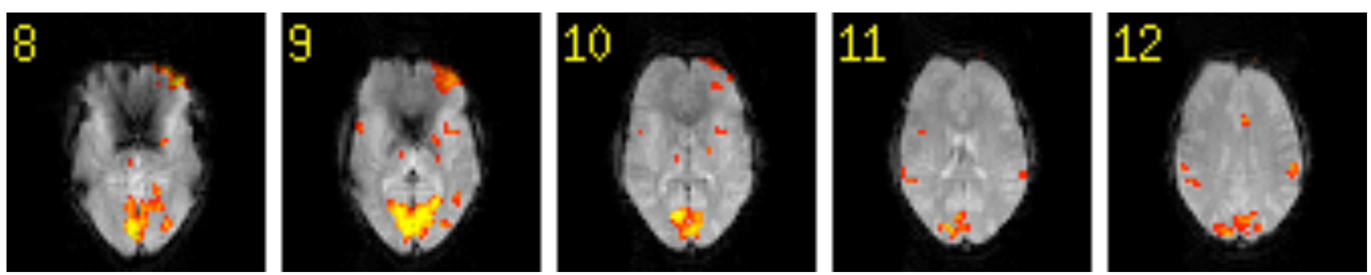
Investigate BOLD response



estimated signal time course



standard hrf model

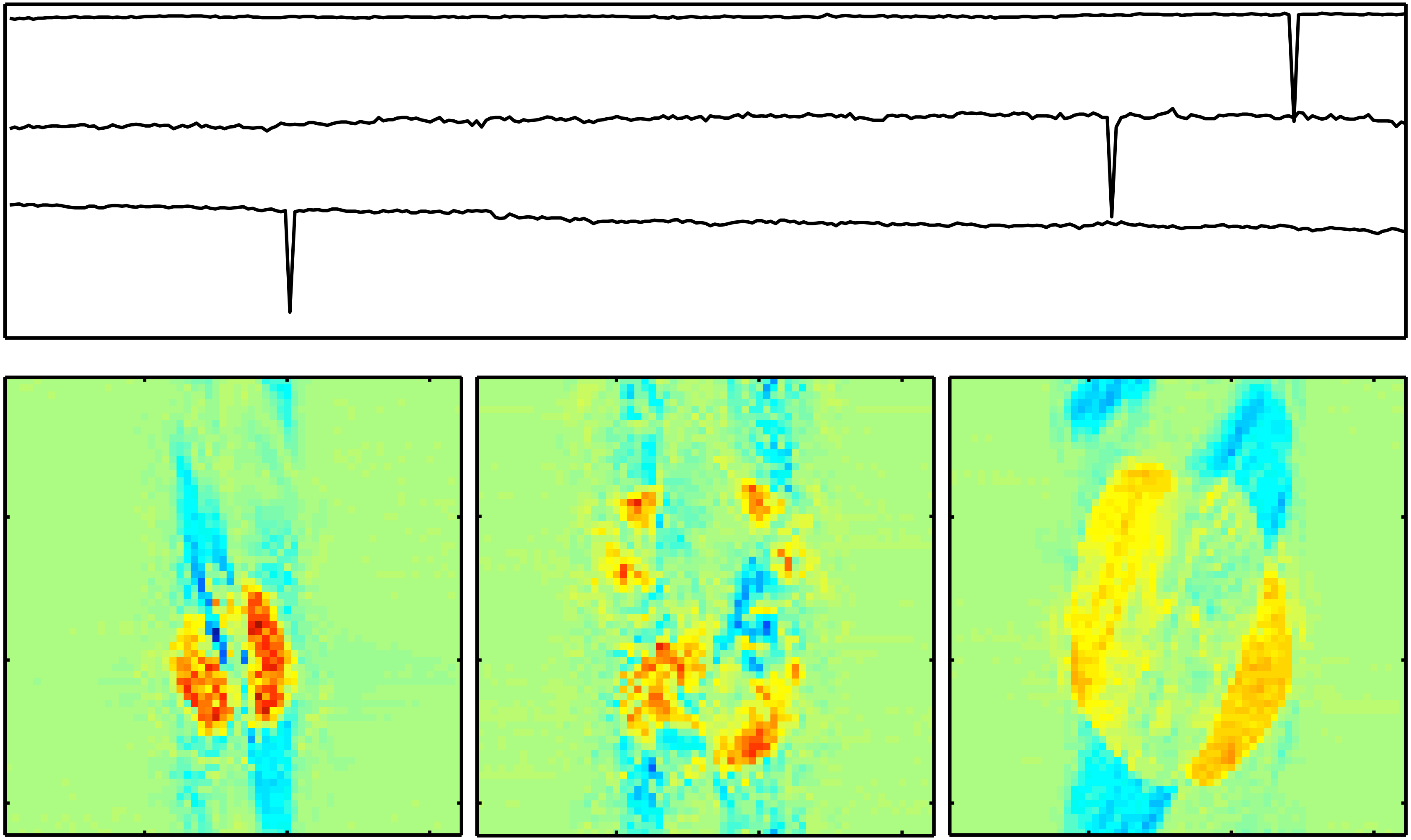


Applications

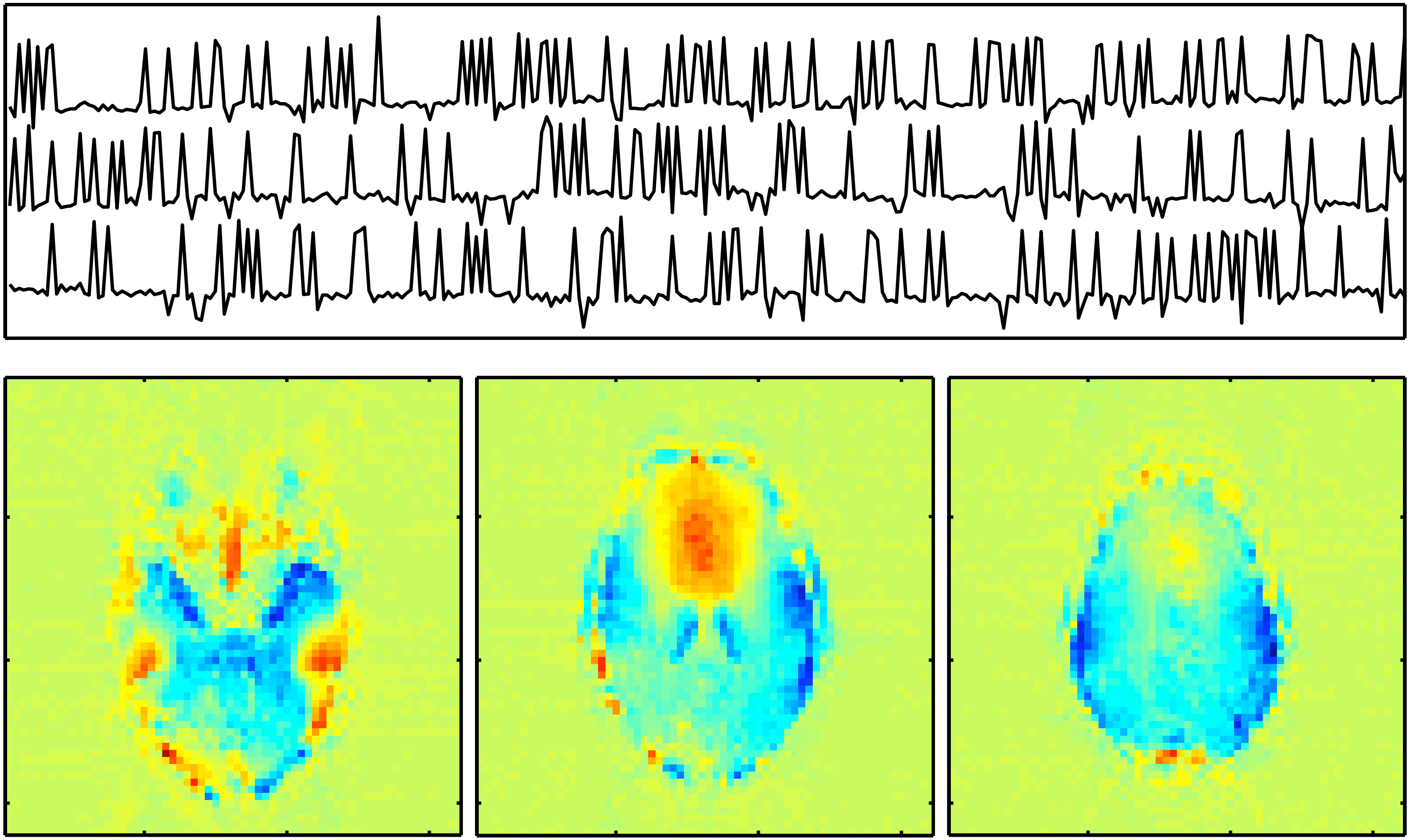
EDA techniques can be useful to

- 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

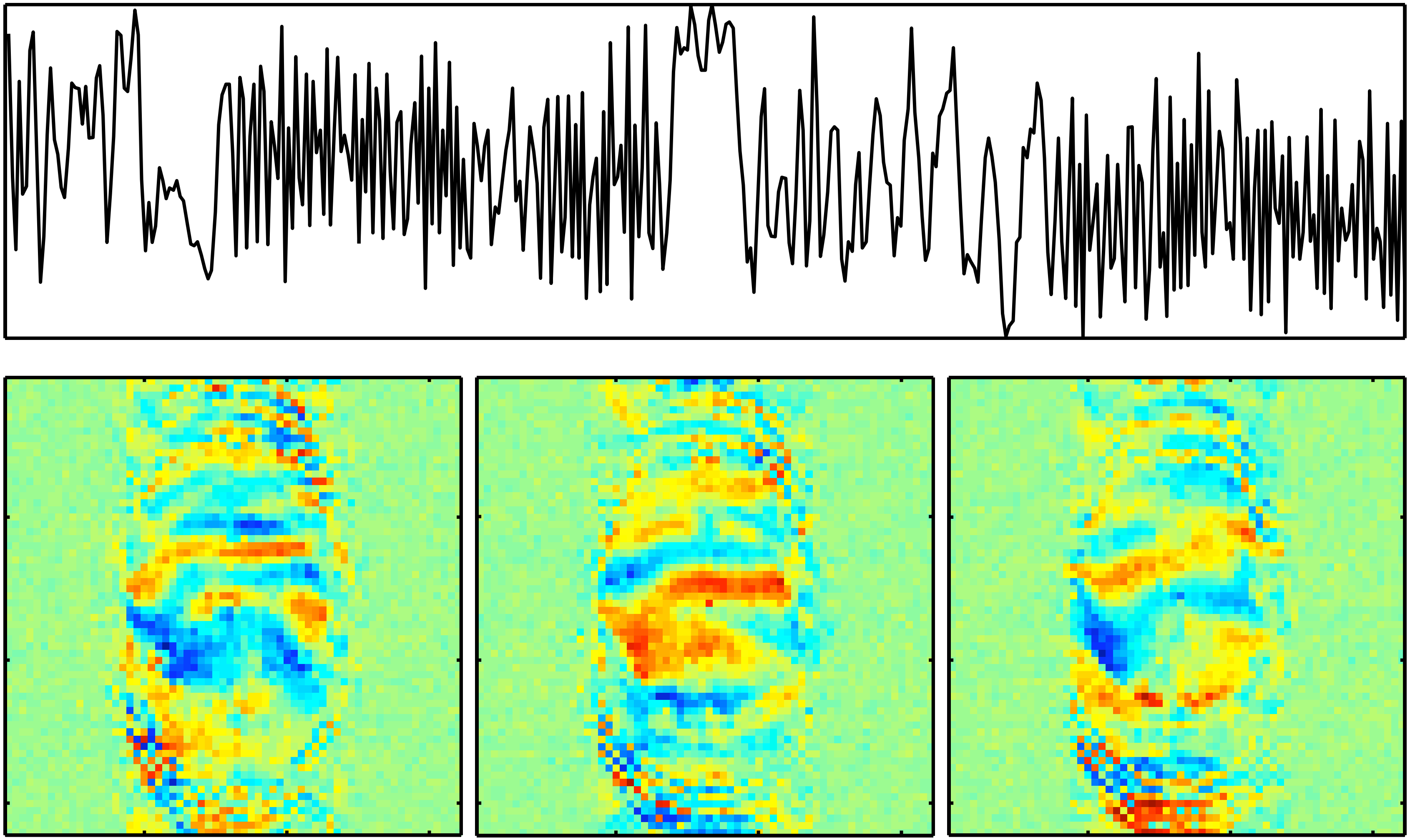
slice drop-outs



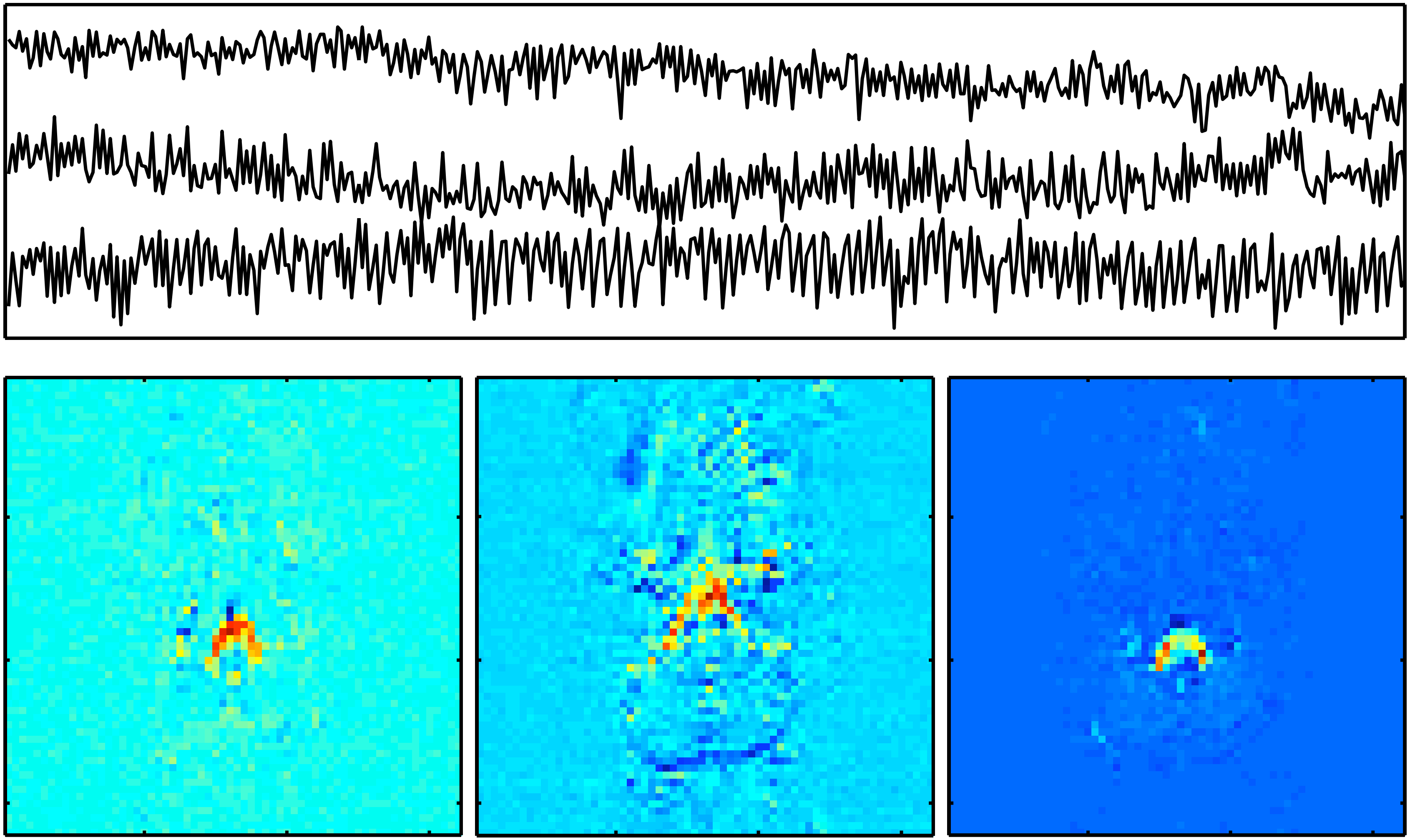
gradient instability



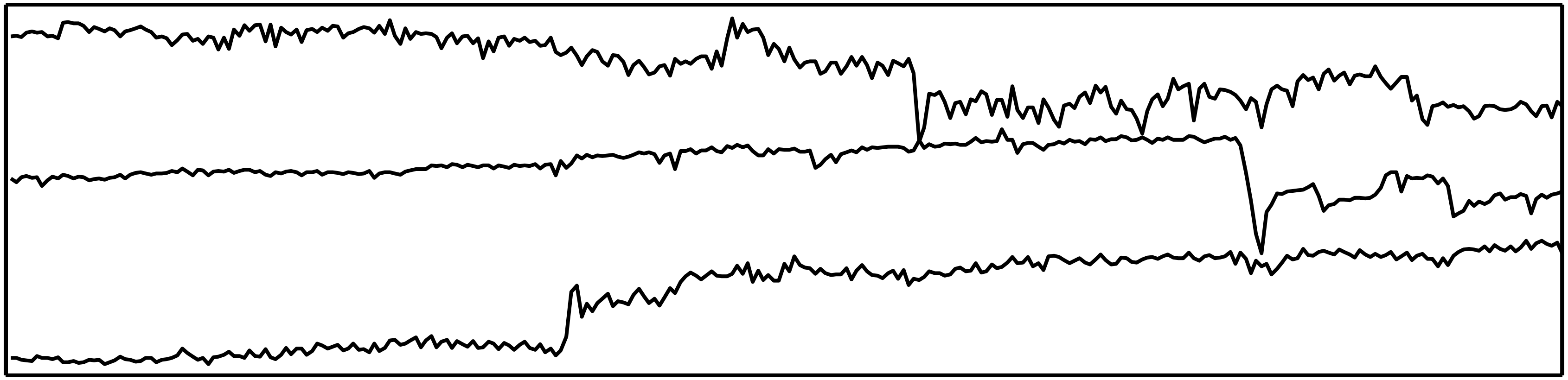
EPI ghost

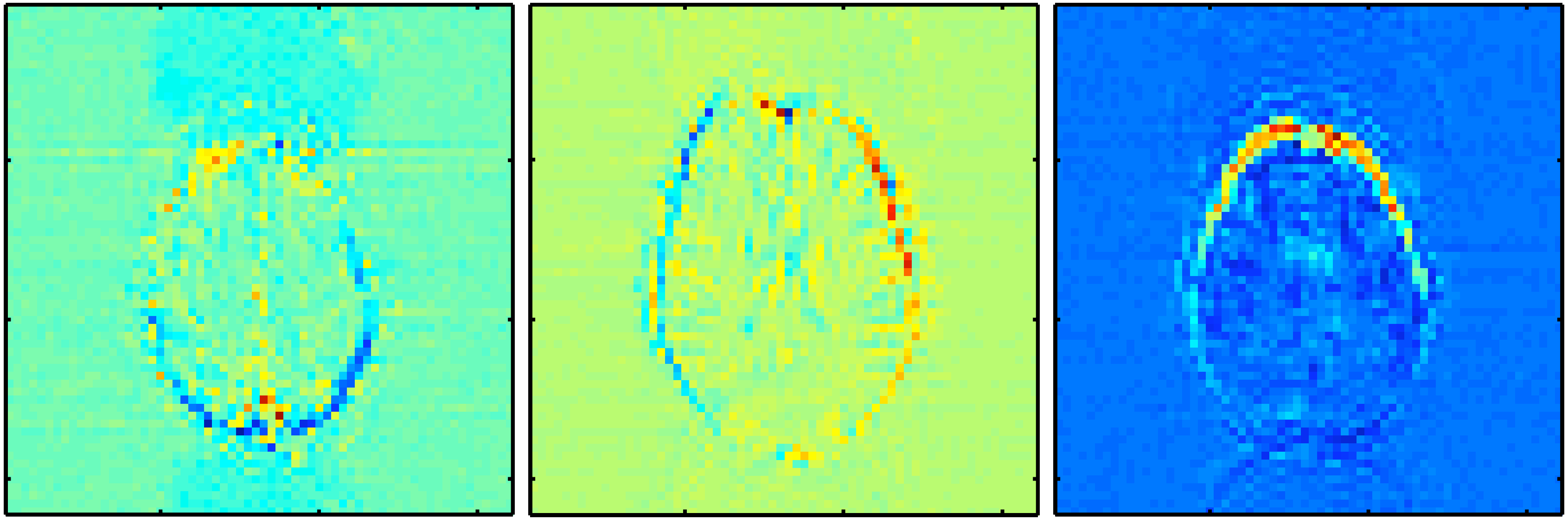


high-frequency noise

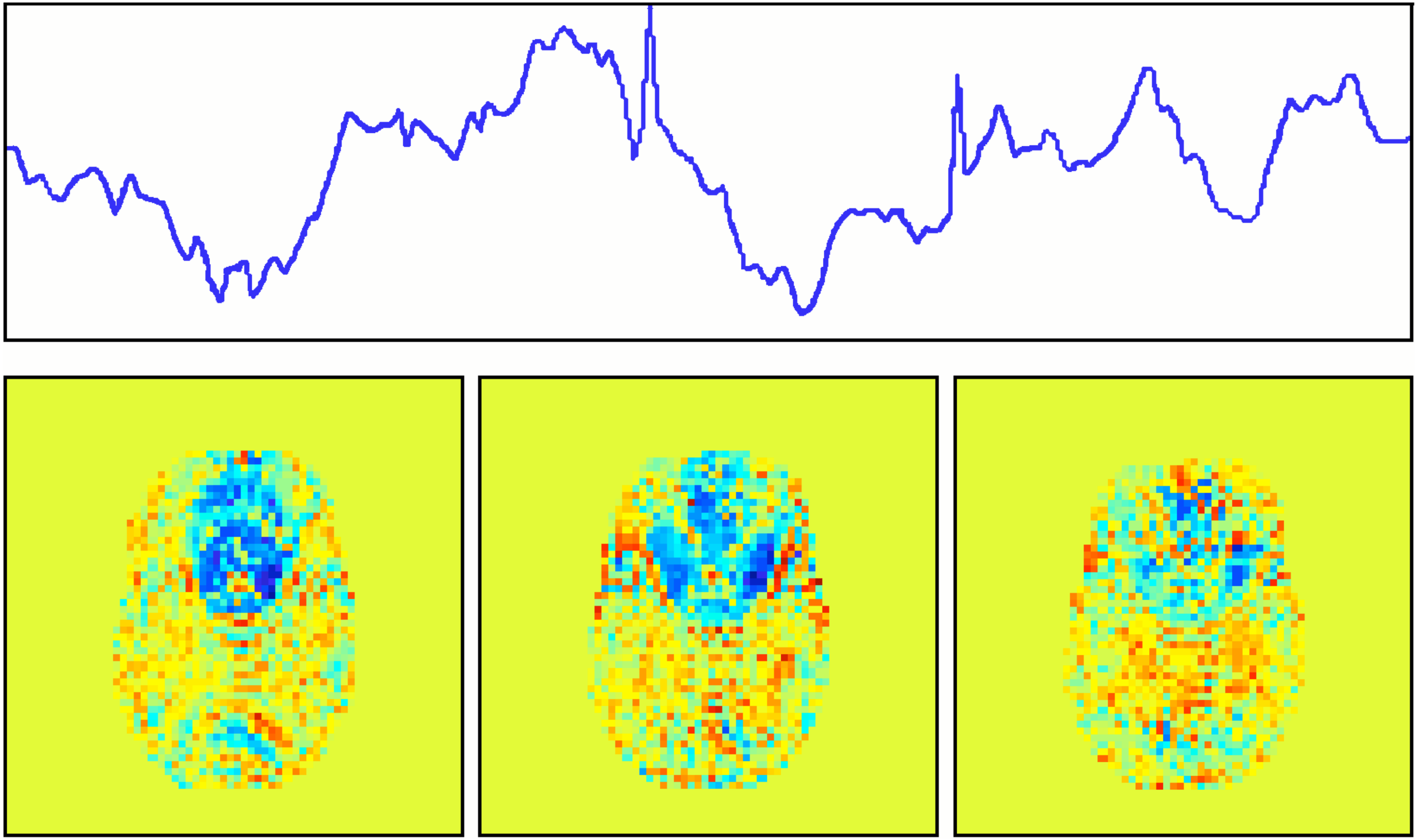


head motion

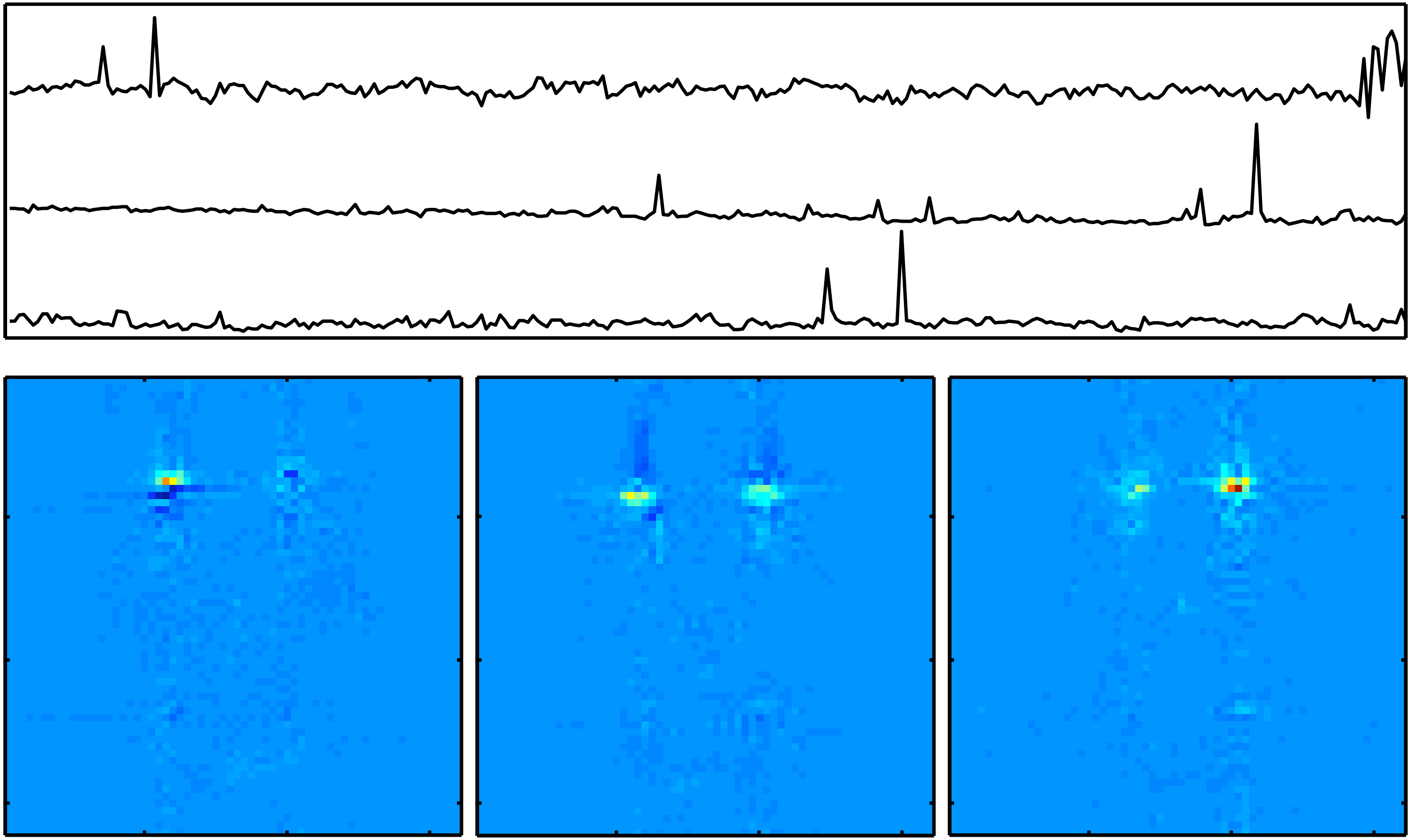




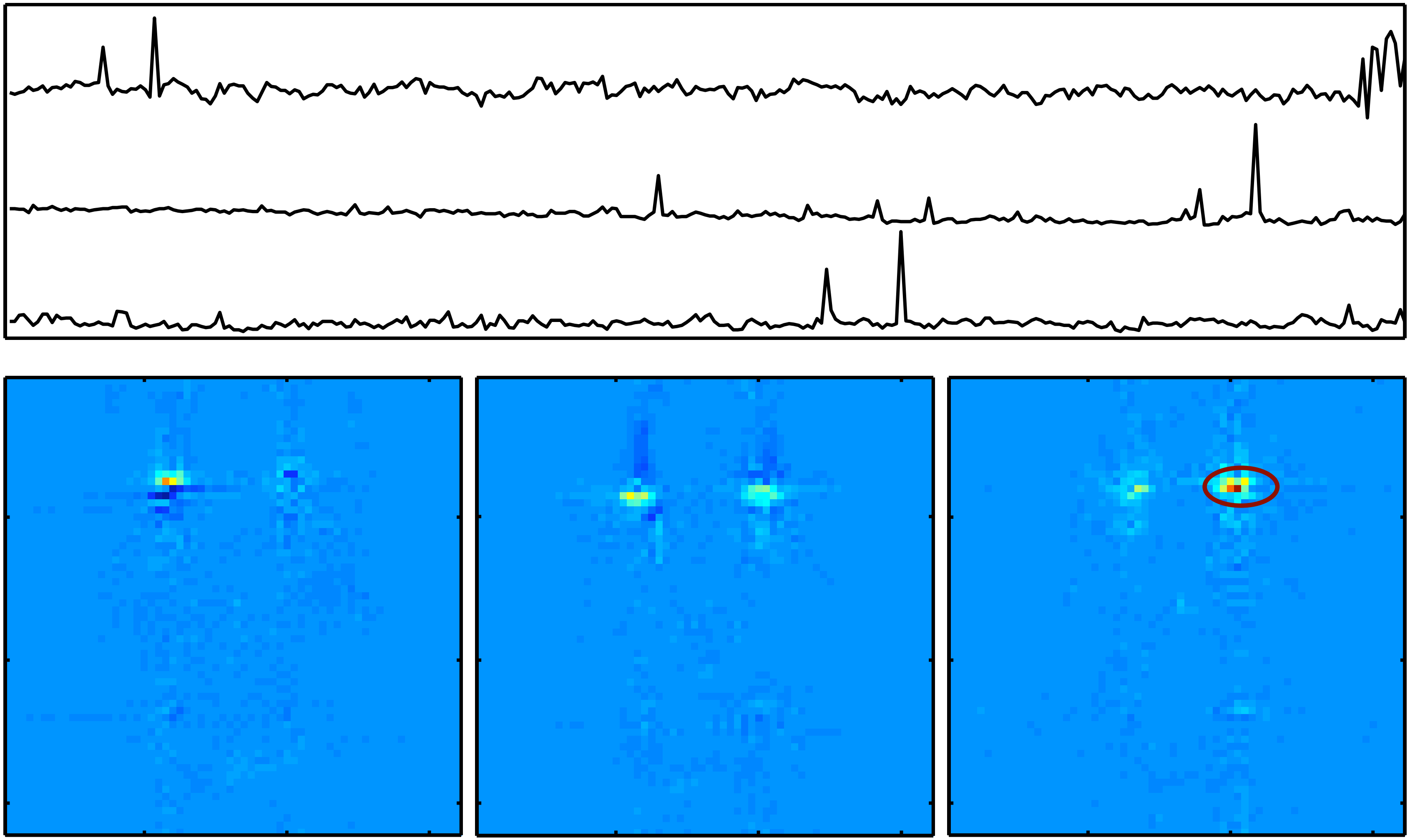
field inhomogeneity



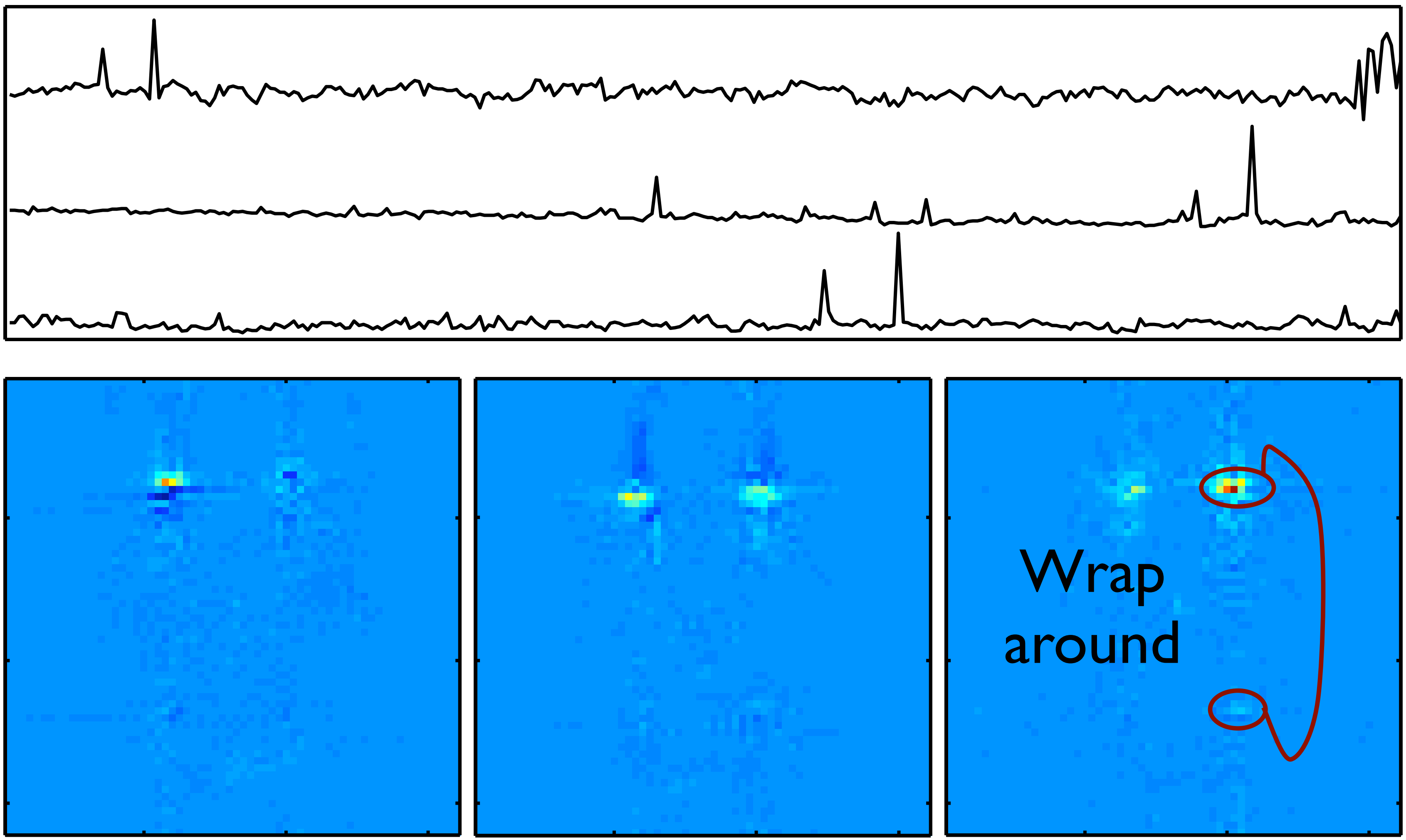
eye-related artefacts



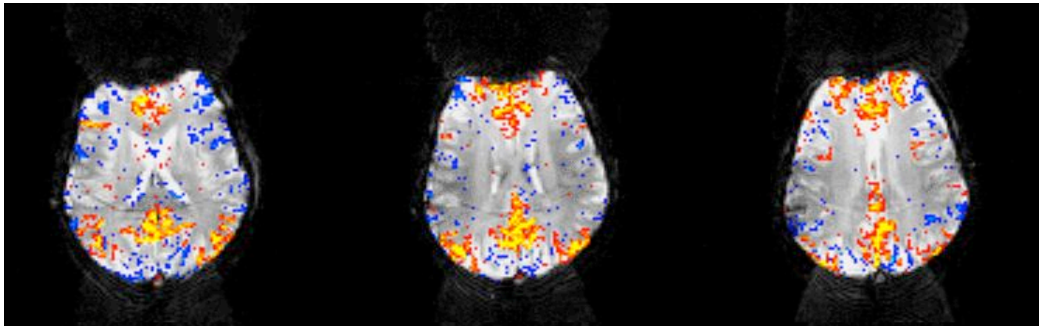
eye-related artefacts



eye-related artefacts



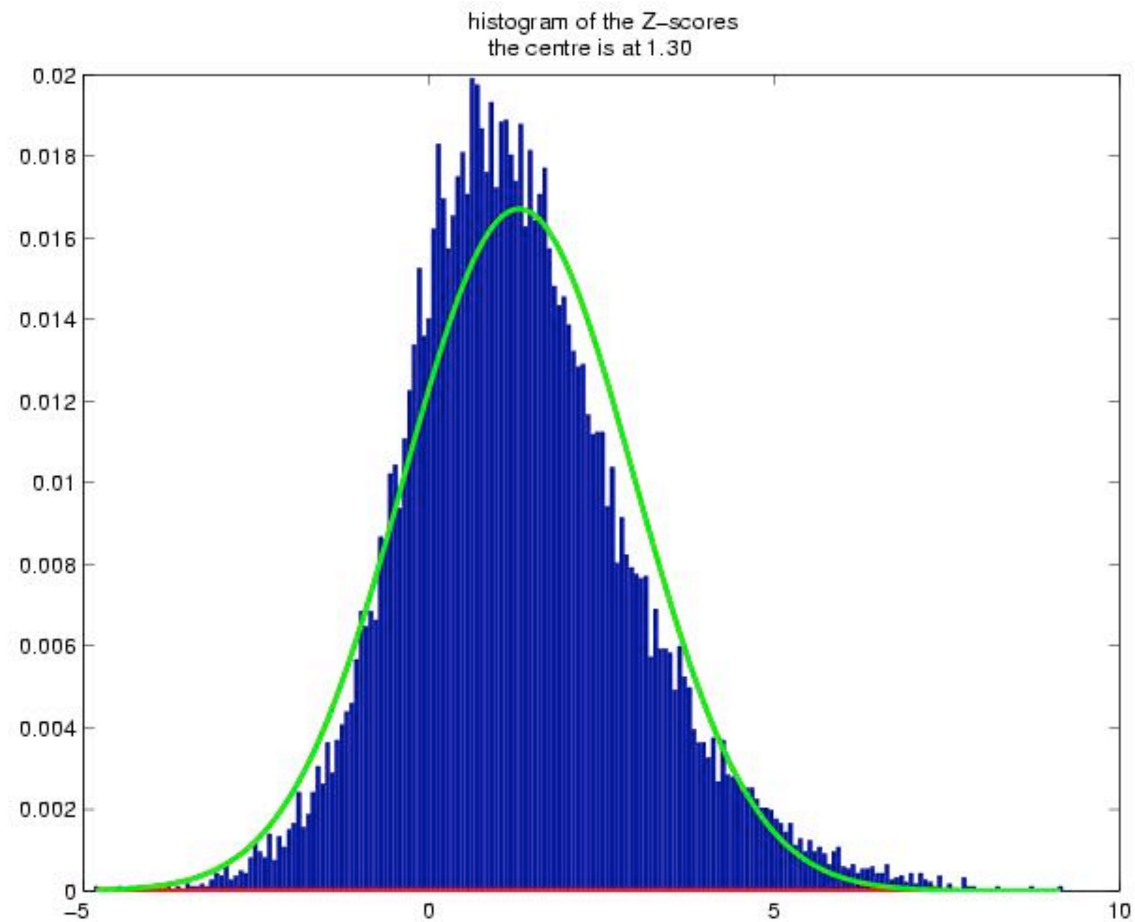
Structured Noise and the GLM



- ‘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

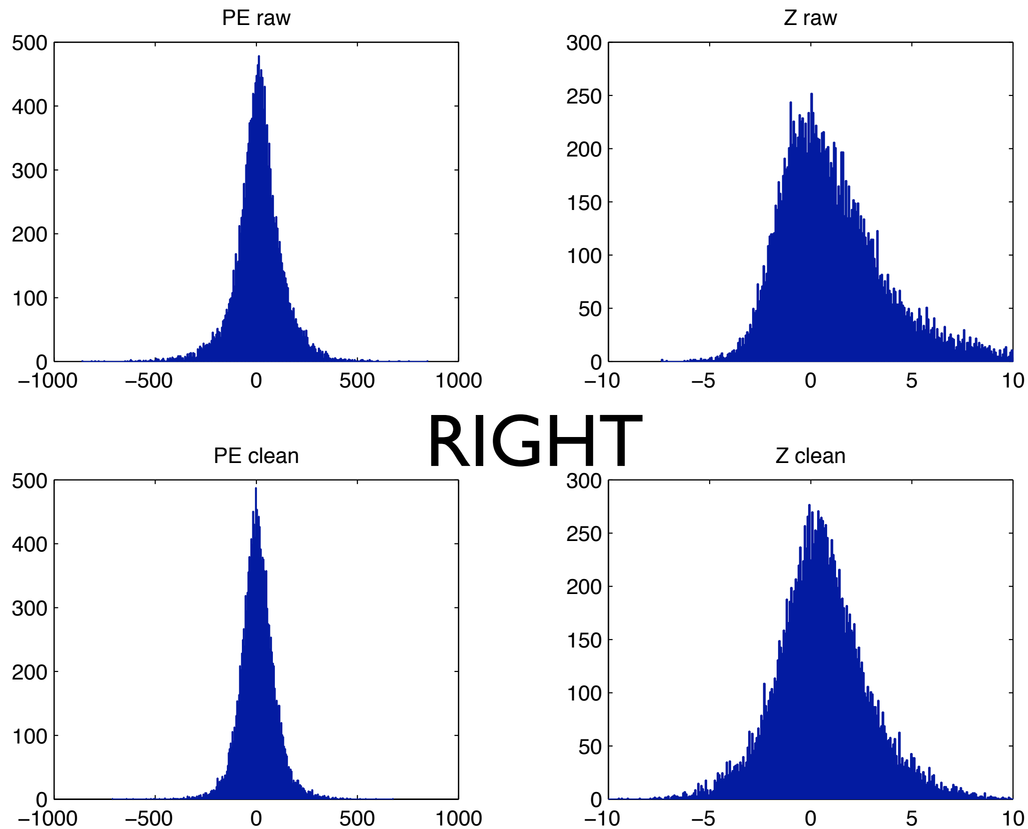
- Correlations of the noise time courses with 'typical' FMRI regressors can cause a shift in the histogram of the Z-statistics
- Thresholded maps will have wrong false-positive rate



Denoising FMRI

- Example: left vs right hand finger tapping

before denoising



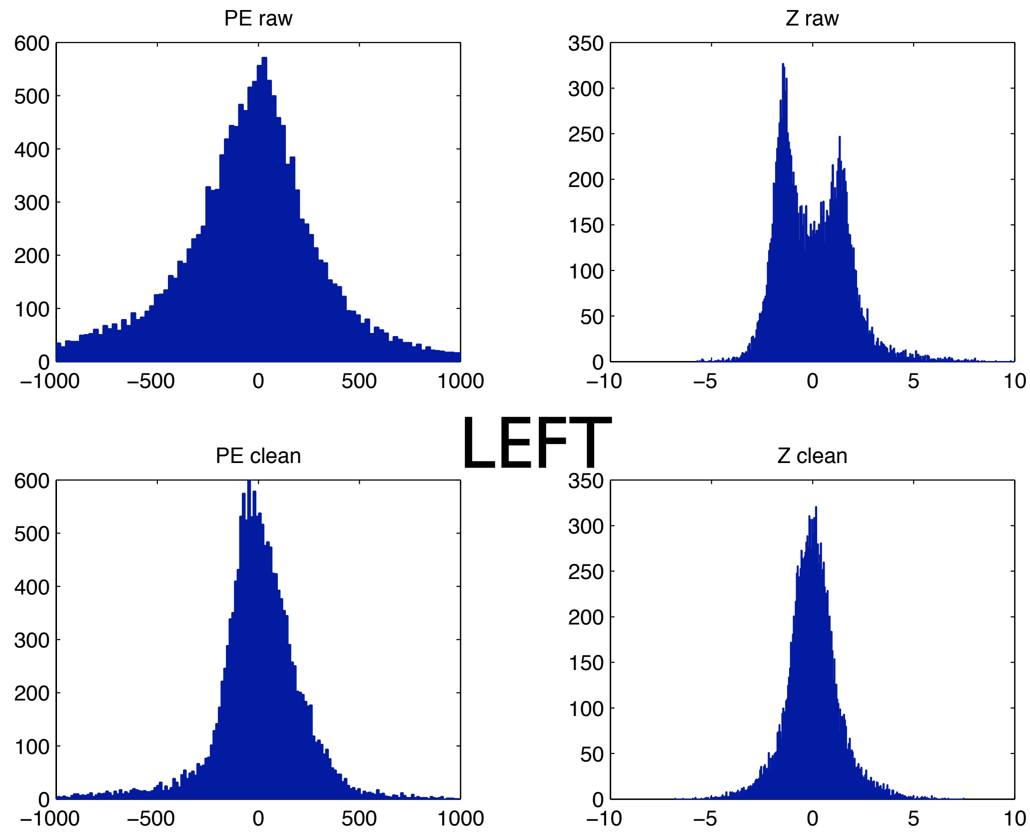
after denoising

Johansen-Berg et al. PNAS 2002

Denoising FMRI

- Example: left vs right hand finger tapping

before denoising



LEFT

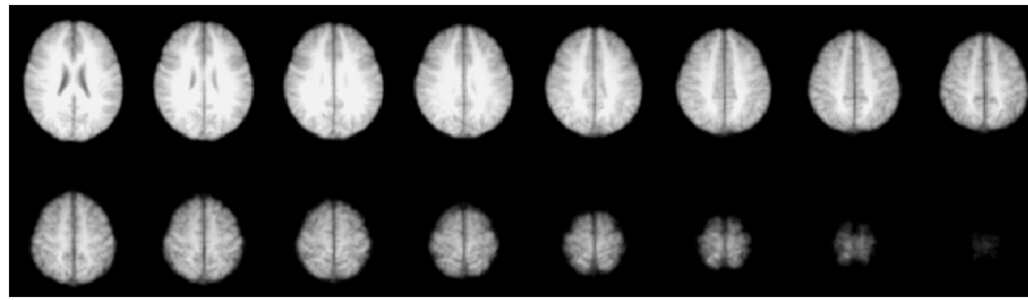
after denoising

Johansen-Berg et al. PNAS 2002

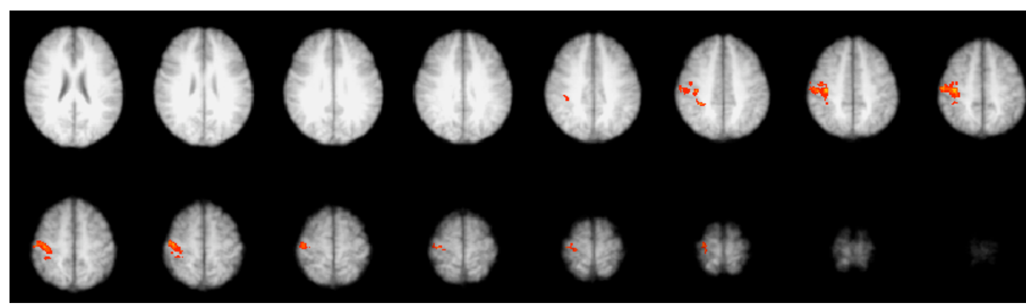
Denoising FMRI

- Example: left vs right hand finger tapping

before denoising



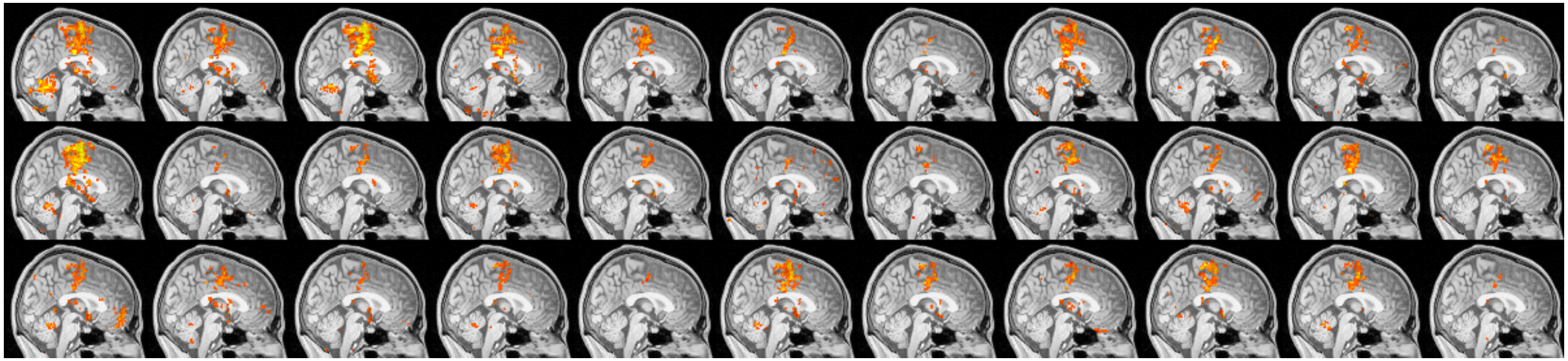
LEFT - RIGHT contrast



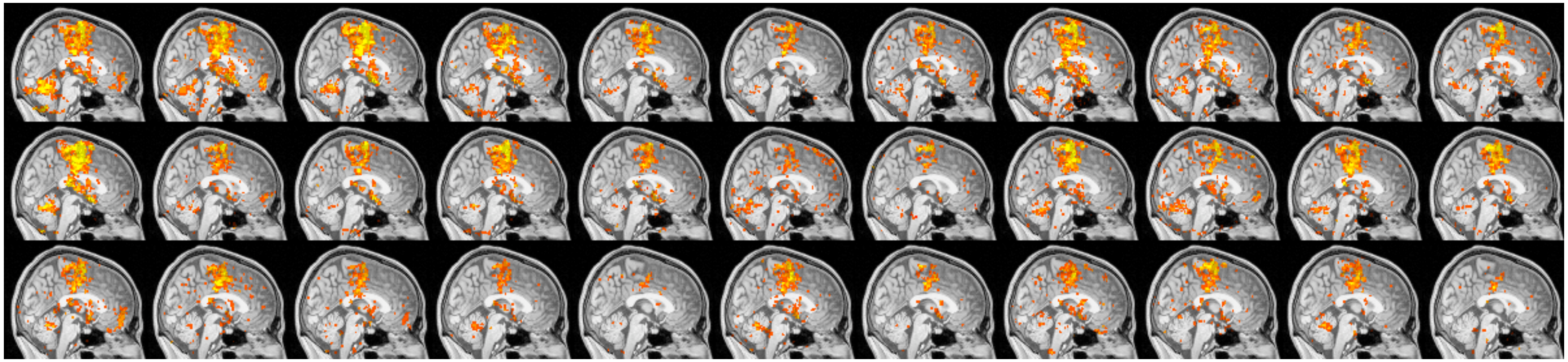
after denoising

Johansen-Berg et al.
PNAS 2002

Apparent variability



McGonigle et al.: 33 Sessions under motor paradigm



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

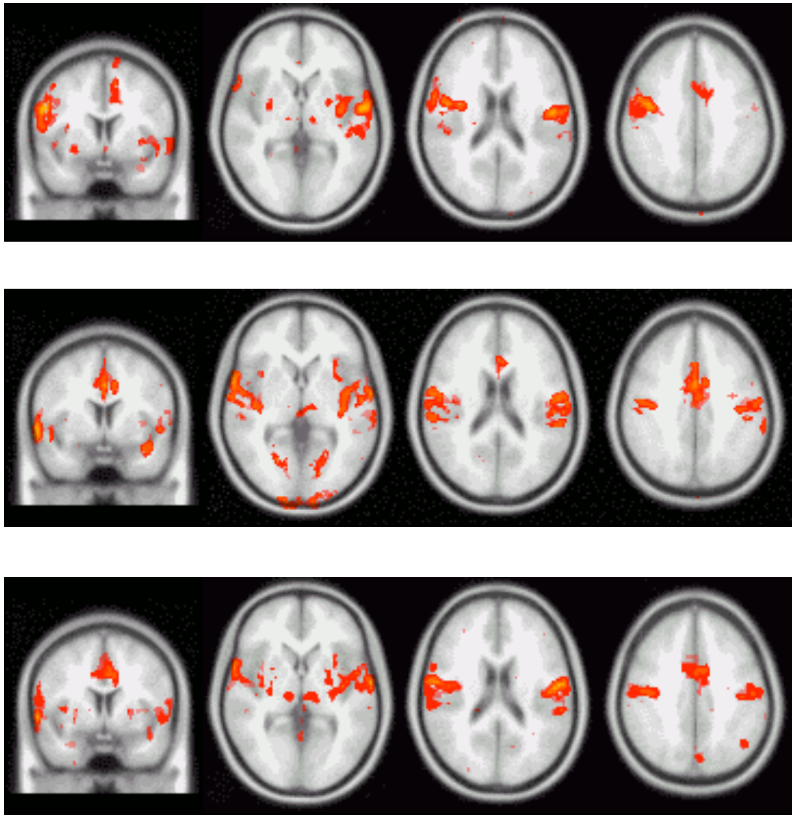
Applications

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

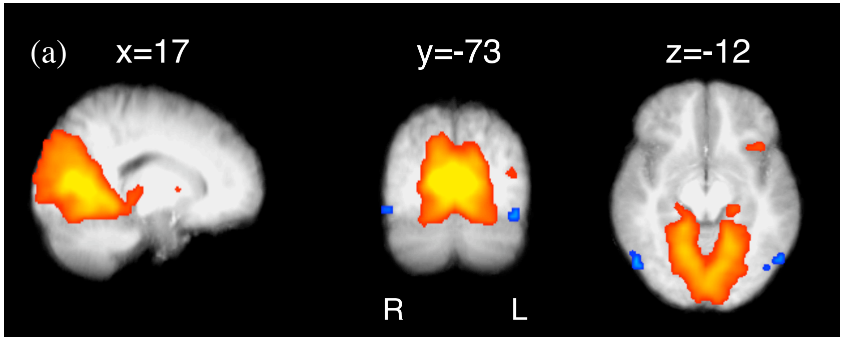
PICA on resting data

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

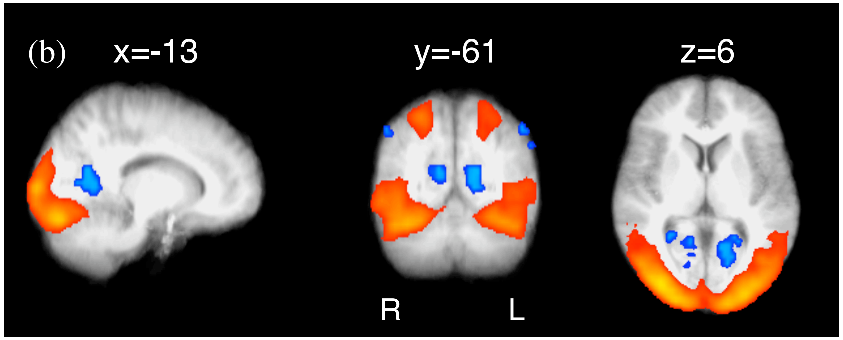


Example: ICA maps - 1 subject at 3 different sessions

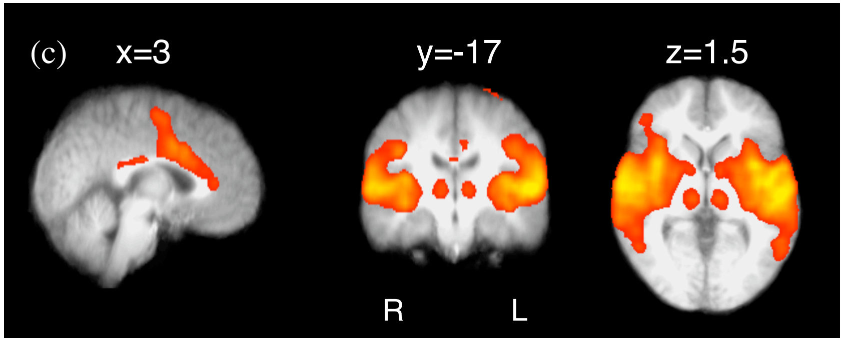
Spatial characteristics



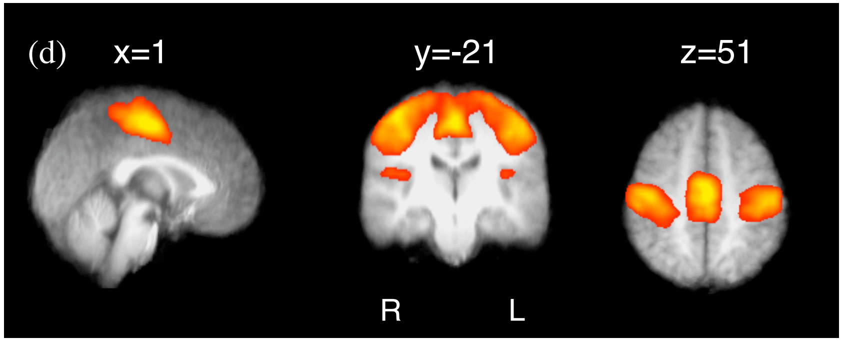
Medial visual cortex



Lateral Visual Cortex

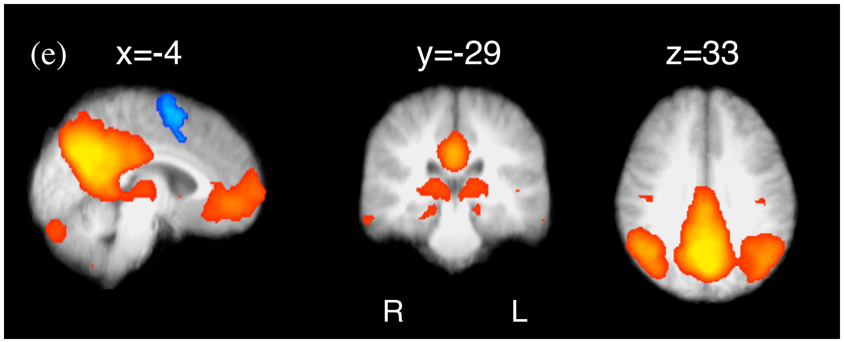


Auditory system

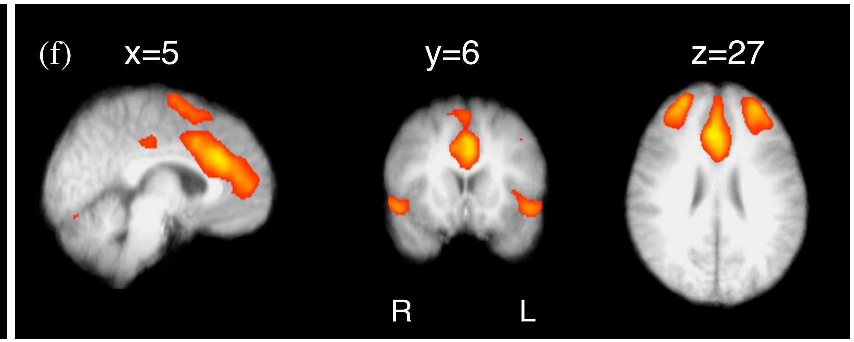


Sensori-motor system

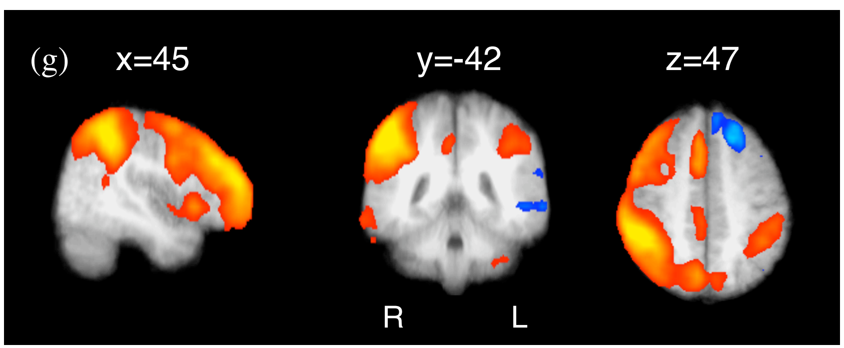
Spatial characteristics

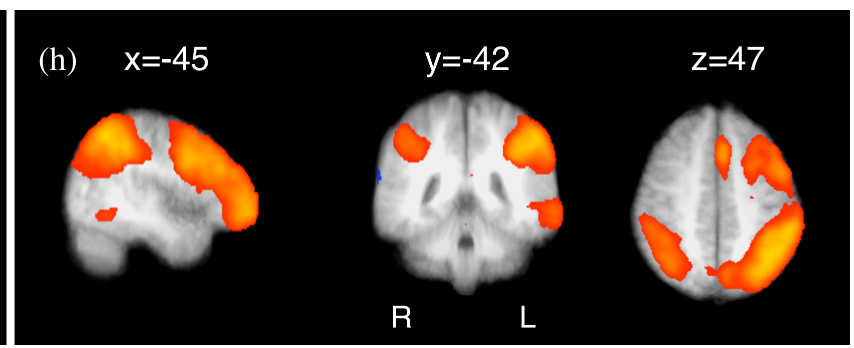


Visuospatial system



Executive control

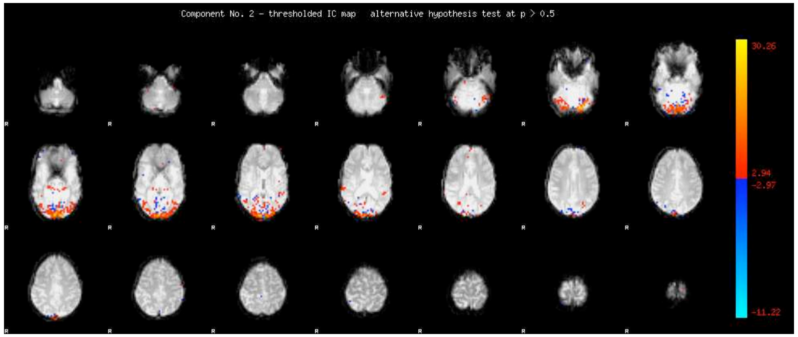


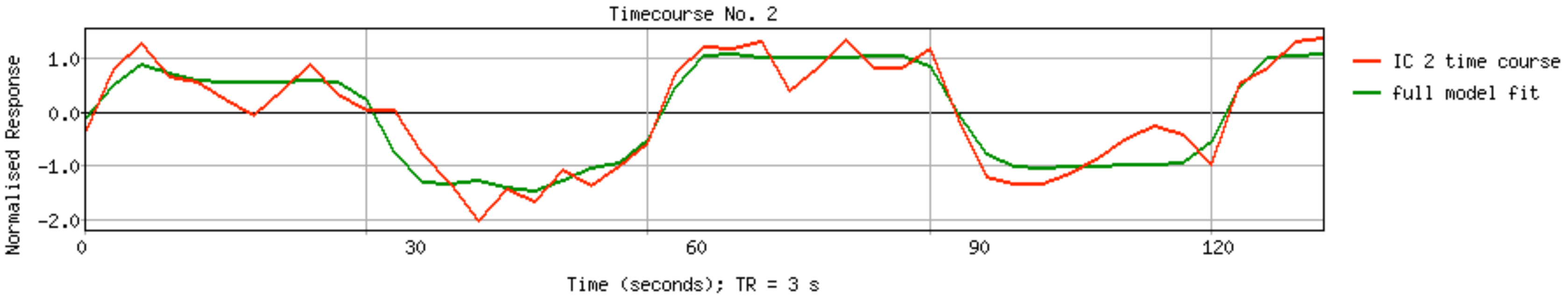


Visual Stream

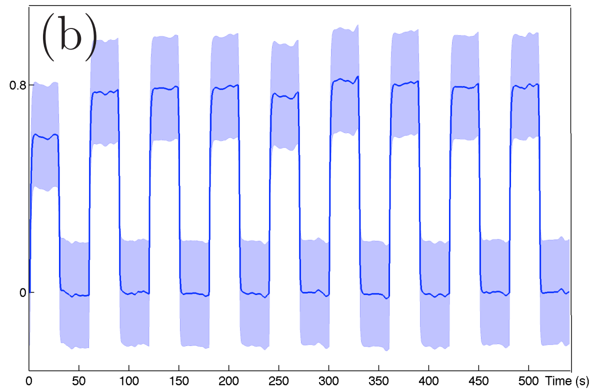
Temporal deconvolution of ICA timecourses

- ‘What are the “task-related” components?
- Use explicit time series model on the IC-generated temporal modes (using BDS)

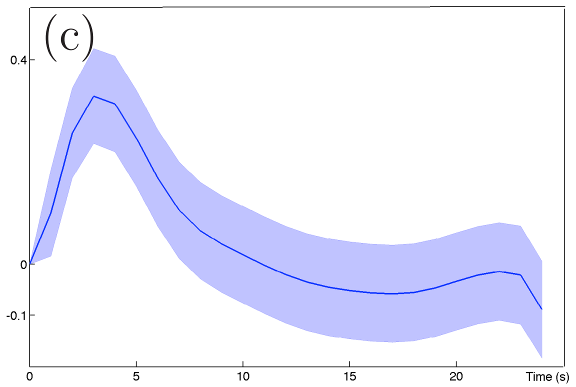




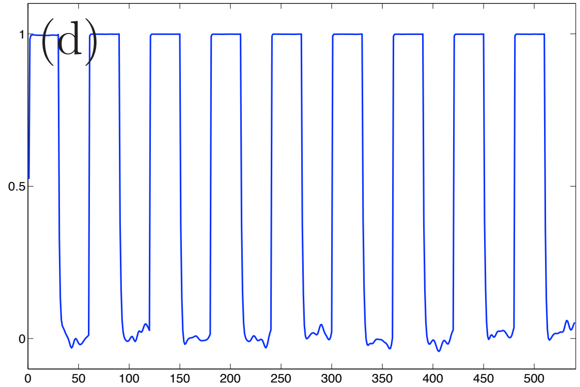
Example: BDS/ICA integration



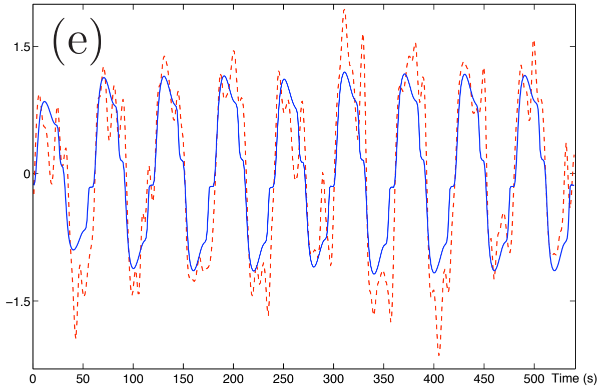
- Estimate the mean of the posterior distribution over the neuronal response



- Estimate the mean of the Gaussian posterior distribution of the HRF



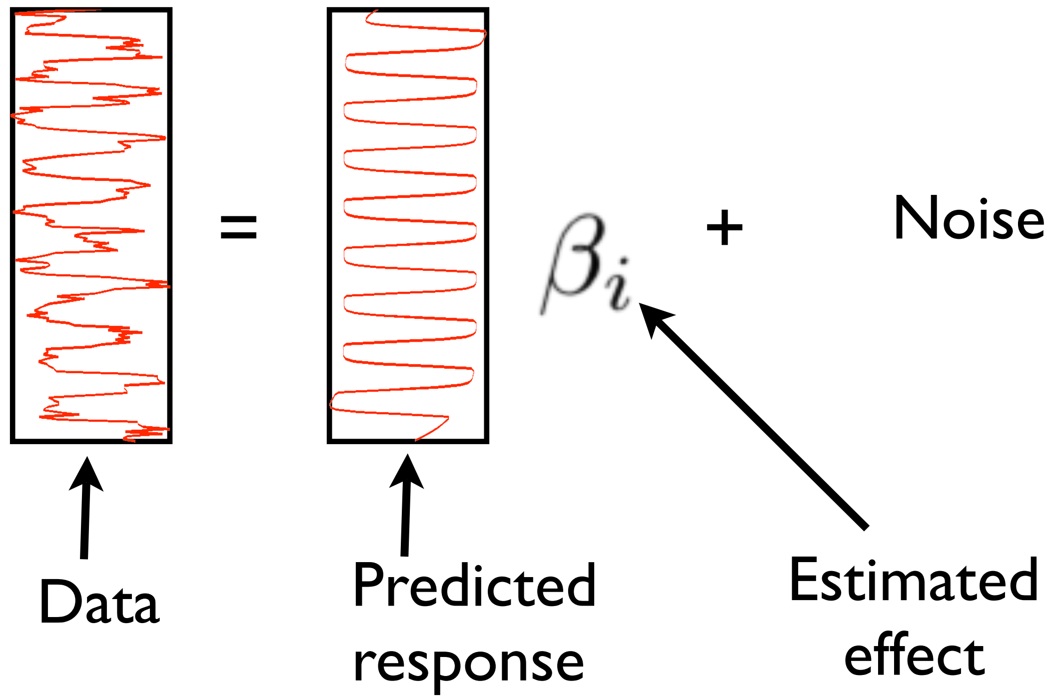
- posterior probability of neuronal response > 0 for the IC



- Assess the total model fit

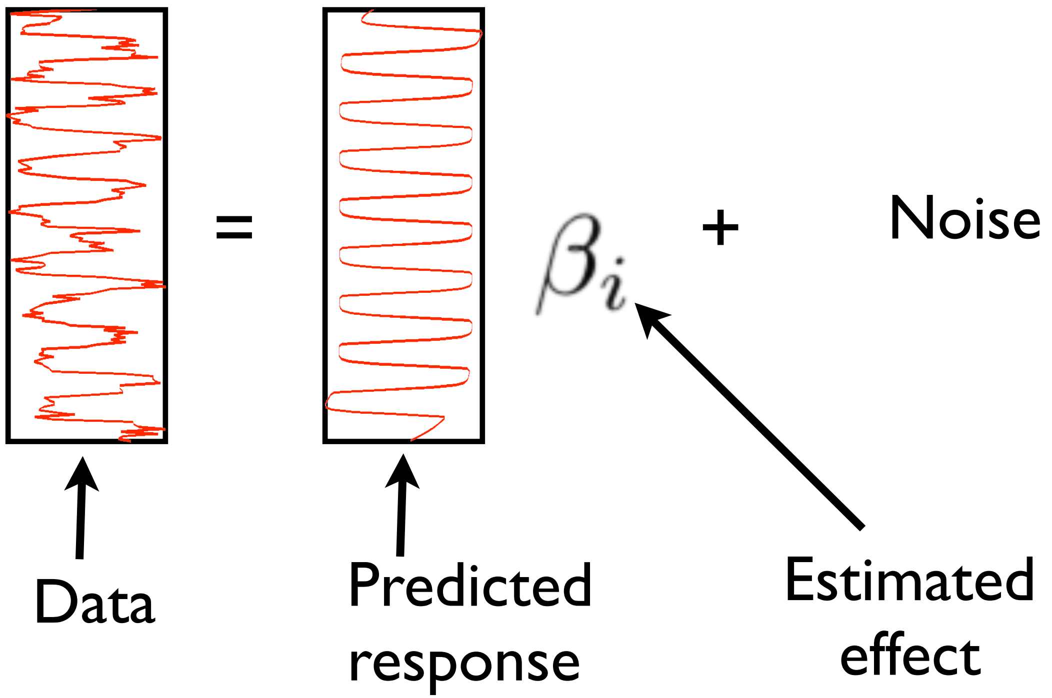
Integrating GLM and ICA for FMRI analysis

GLM



Integrating GLM and ICA for FMRI analysis

GLM

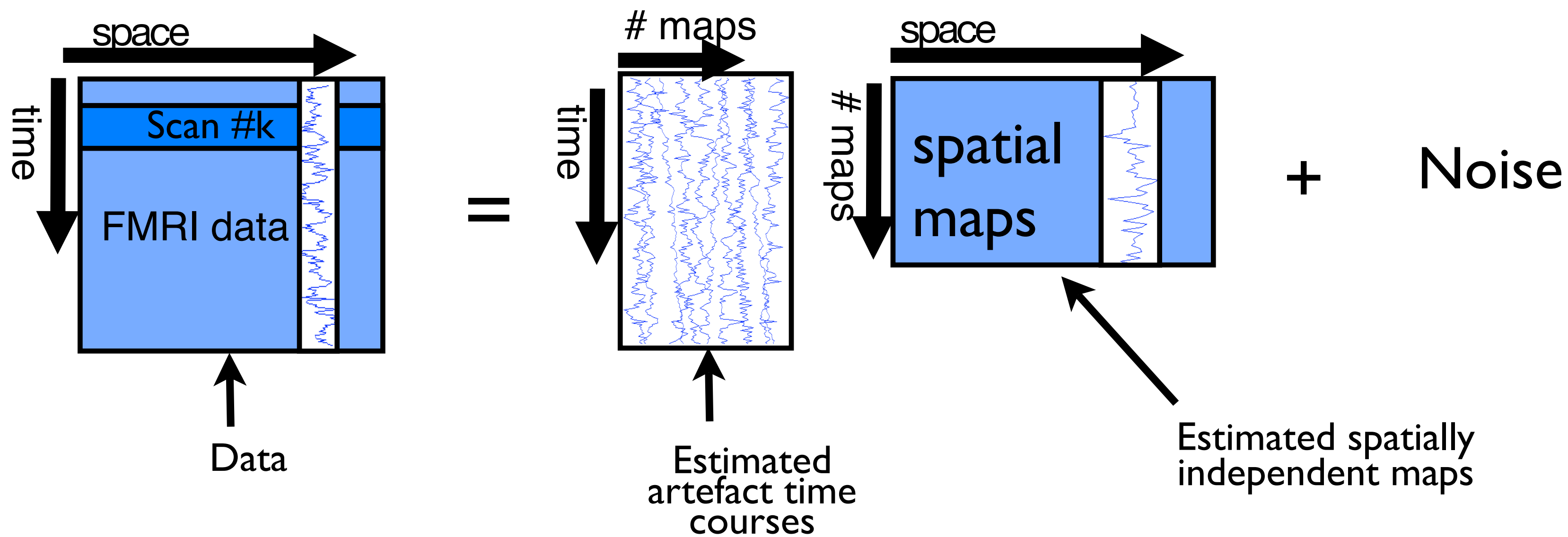


+ clear conclusions on a particular question

- results depend on the model

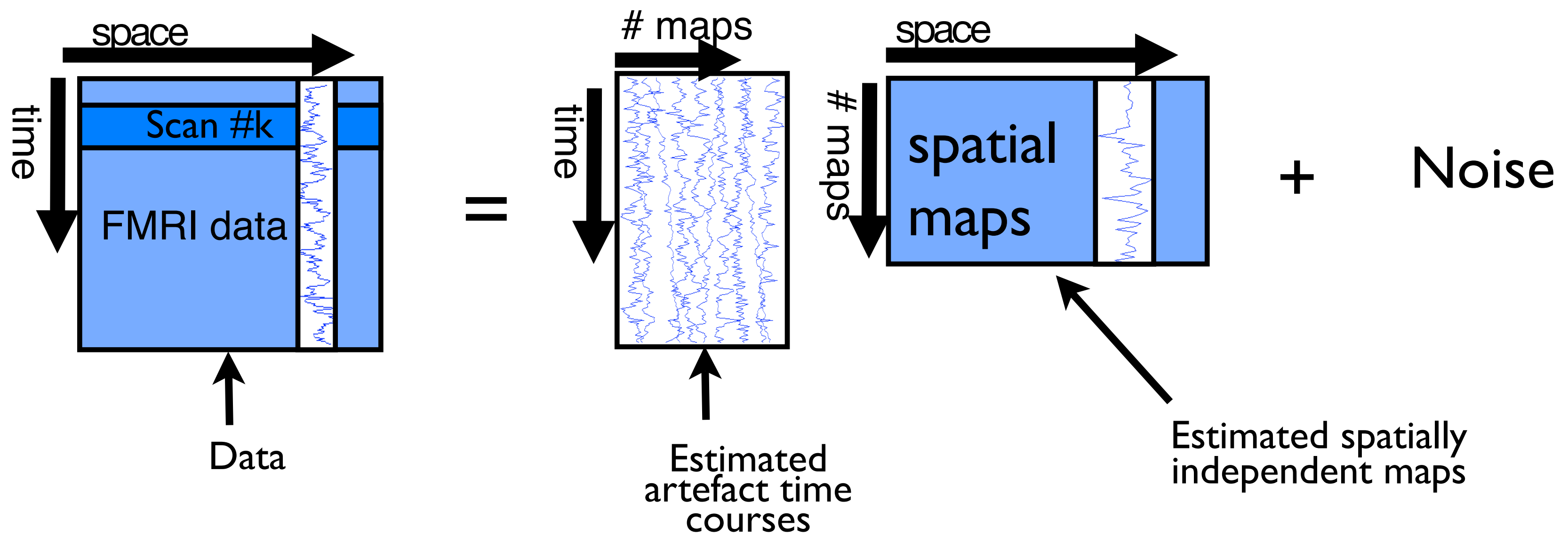
Integrating GLM and ICA for FMRI analysis

PICA



Integrating GLM and ICA for FMRI analysis

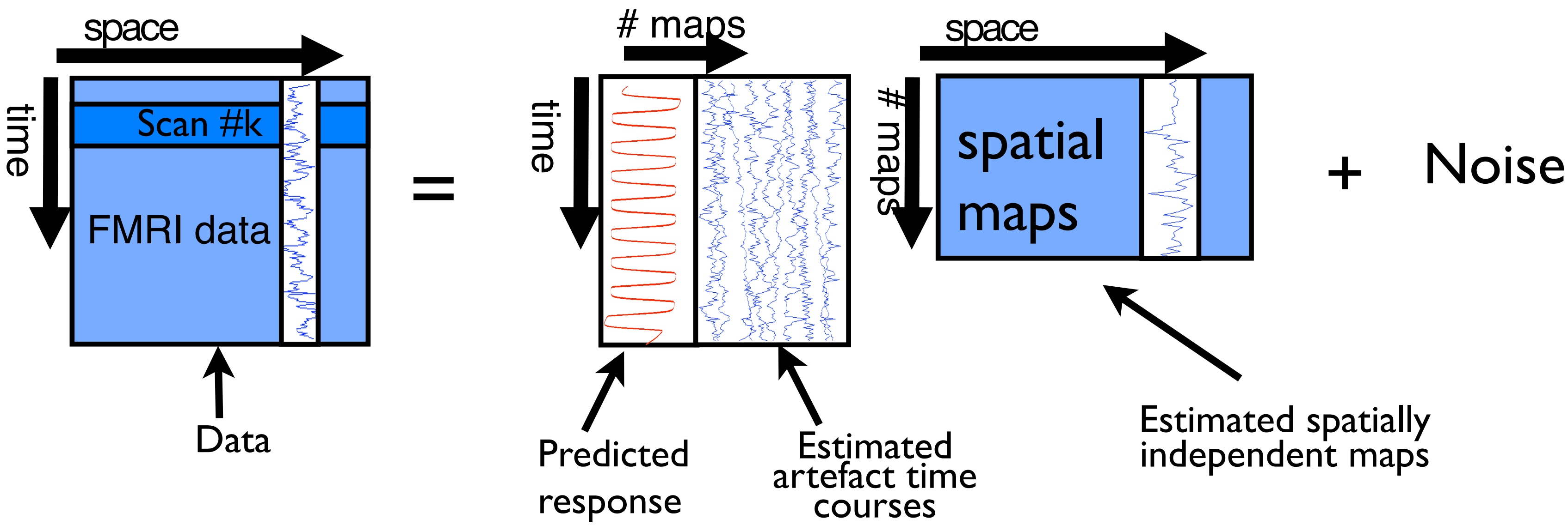
PICA



+ data driven and multivariate approach
- no knowledge about the fMRI paradigm is used
- can be hard to interpret activation results

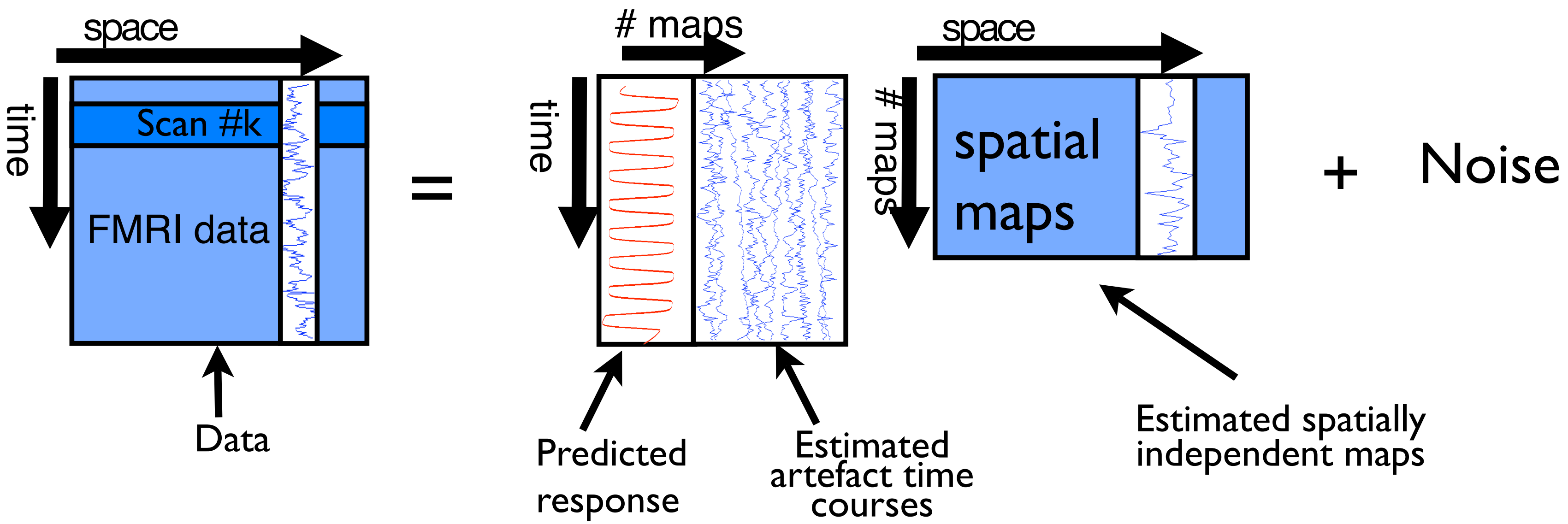
Integrating GLM and ICA for FMRI analysis

GLM + ICA



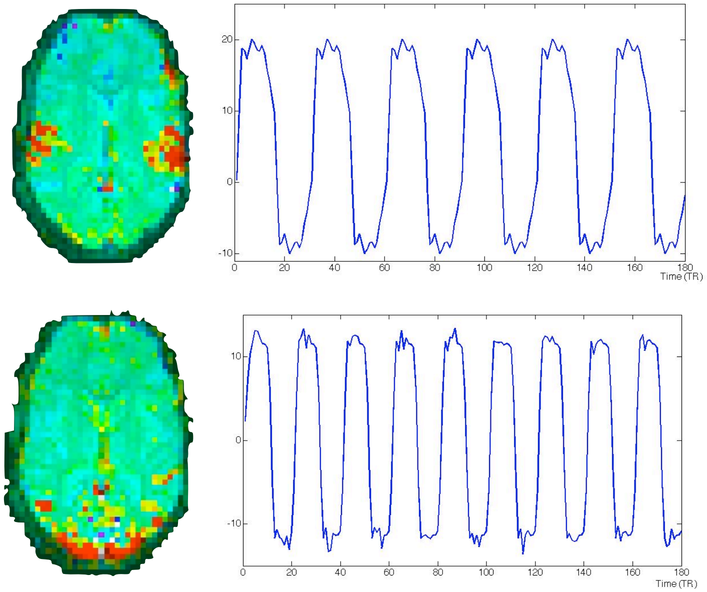
Integrating GLM and ICA for FMRI analysis

GLM + ICA



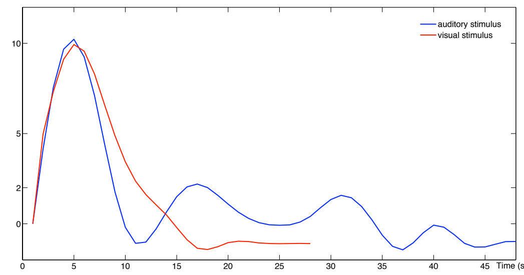
- Stimulus model-based hypothesis testing
- Adaptive model-free artefact modelling
  - e.g. stimulus correlated motion, physiological noise, networks of spontaneous neuronal activity

Modelled ICs

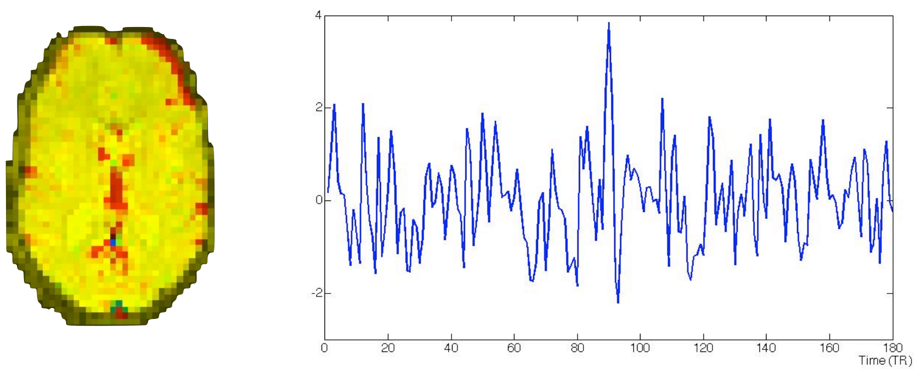


Spatial map Time course

Estimated HRFs



Example “model free” IC



Spatial map Time course