

NeuroImaging Data Processing

aka. Statistical Parametric Mapping short course

Course 3:

General Linear Model, p.1

Content

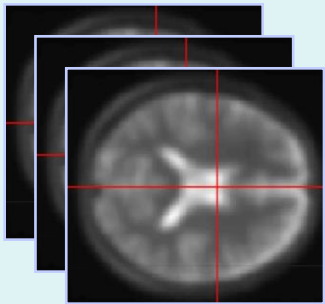
- **Introduction**
- **General Linear Model**
- **Parameter estimation**
- **Improved model**
- **Conclusion**

Content

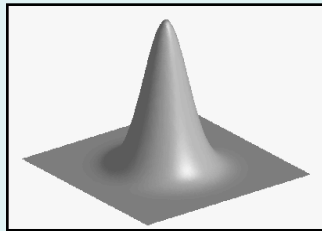
- **Introduction**
- **General Linear Model**
- **Parameter estimation**
- **Improved model**
- **Conclusion**

SPM work flow

Image time-series



Spatial filter

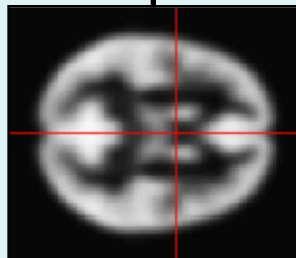


Realignment

Smoothing

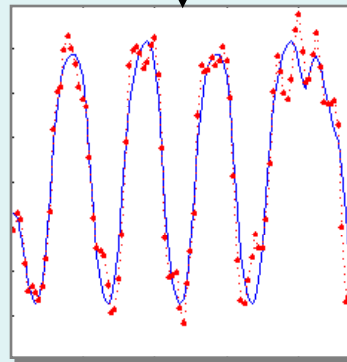
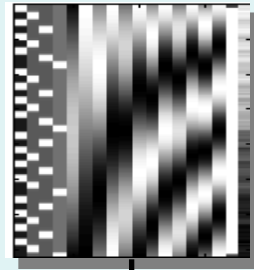
General Linear Model

Normalisation



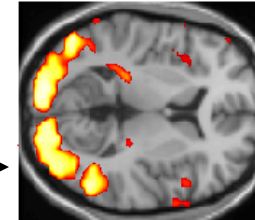
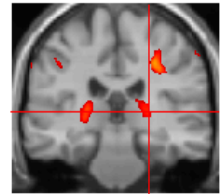
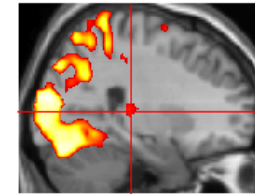
Anatomical
reference

Design matrix



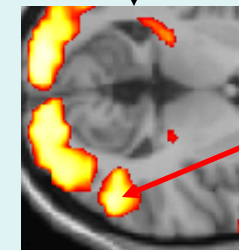
Parameter estimates

Statistical Parametric Map



Statistical
Inference

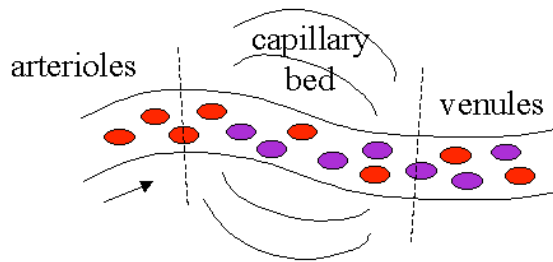
RFT



$p < 0.05$

fMRI & BOLD signal

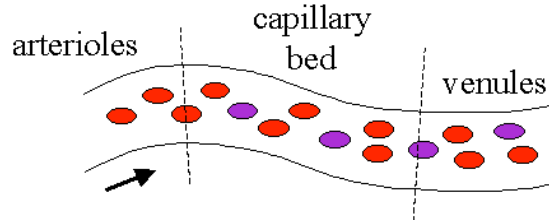
Basal state



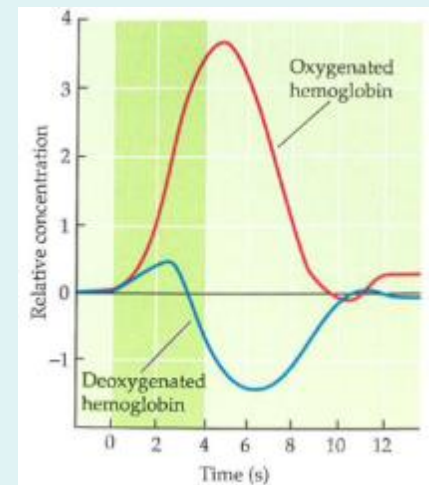
- normal flow
- basal level [Hbr]
- basal CBV
- normal MRI signal

● = HbO₂
● = Hbr

Activated state



- increased flow
- decreased [Hbr] (*lower field gradients around vessels*)
- increased CBV
- increased MRI signal (*from lower field gradients*)

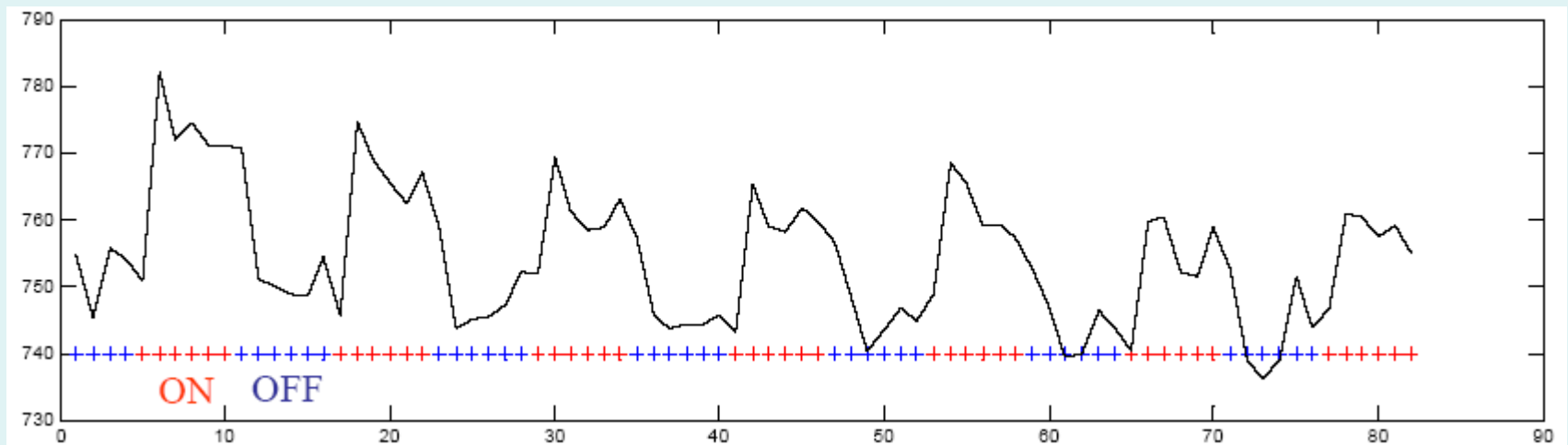


A simple fMRI experiment

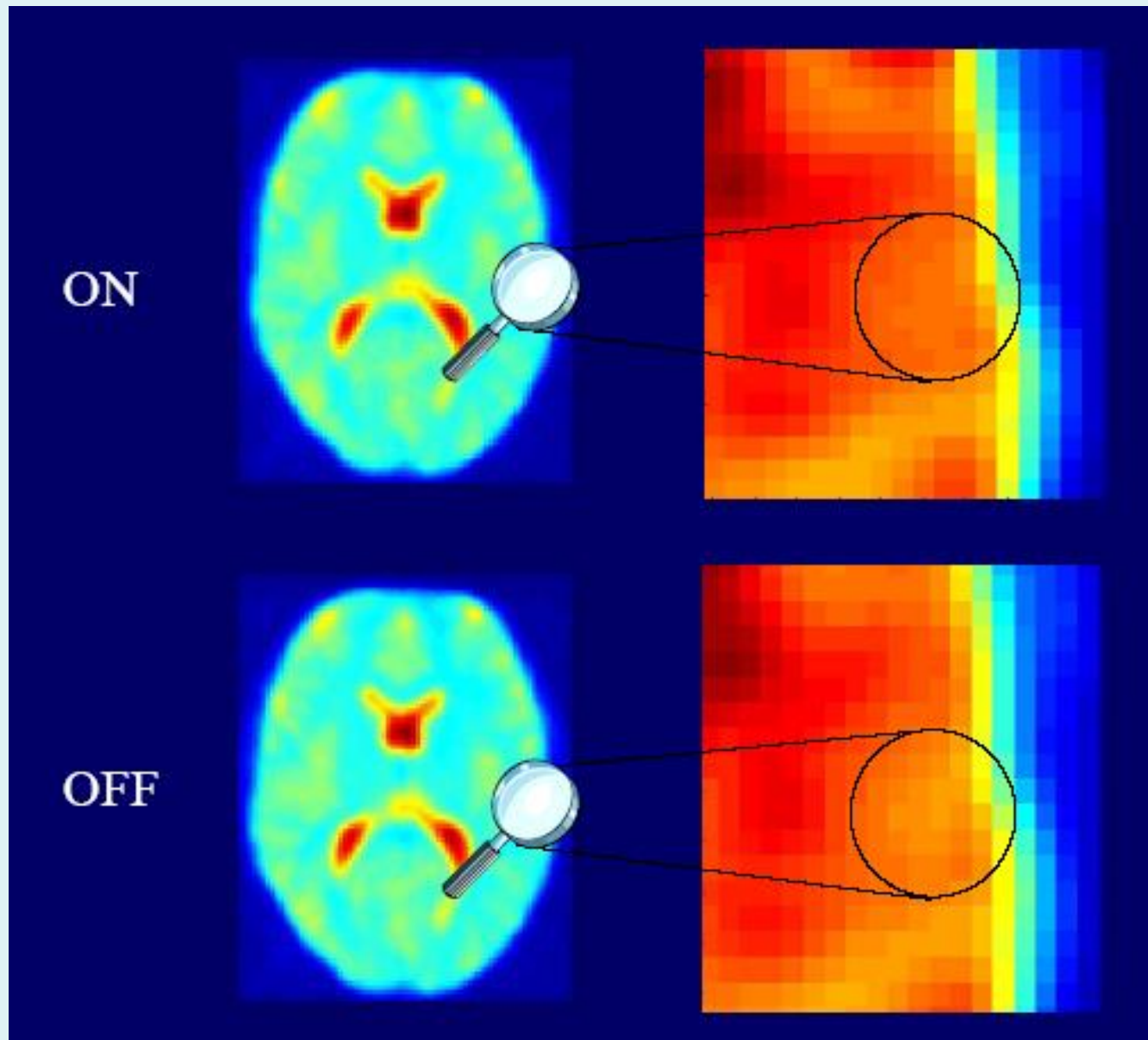
Stimuli: passive word listening versus rest



BOLD response in the primary auditory cortex

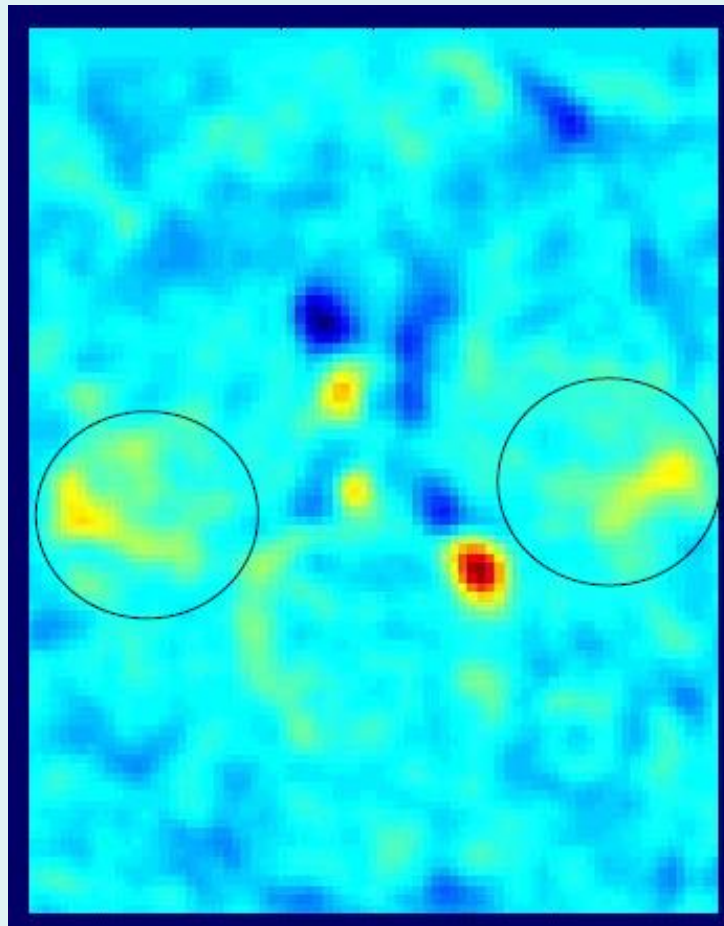


Looking at 2 scans



Looking at 2 scans

ON-OFF, just one scan per condition



Simple fMRI example dataset

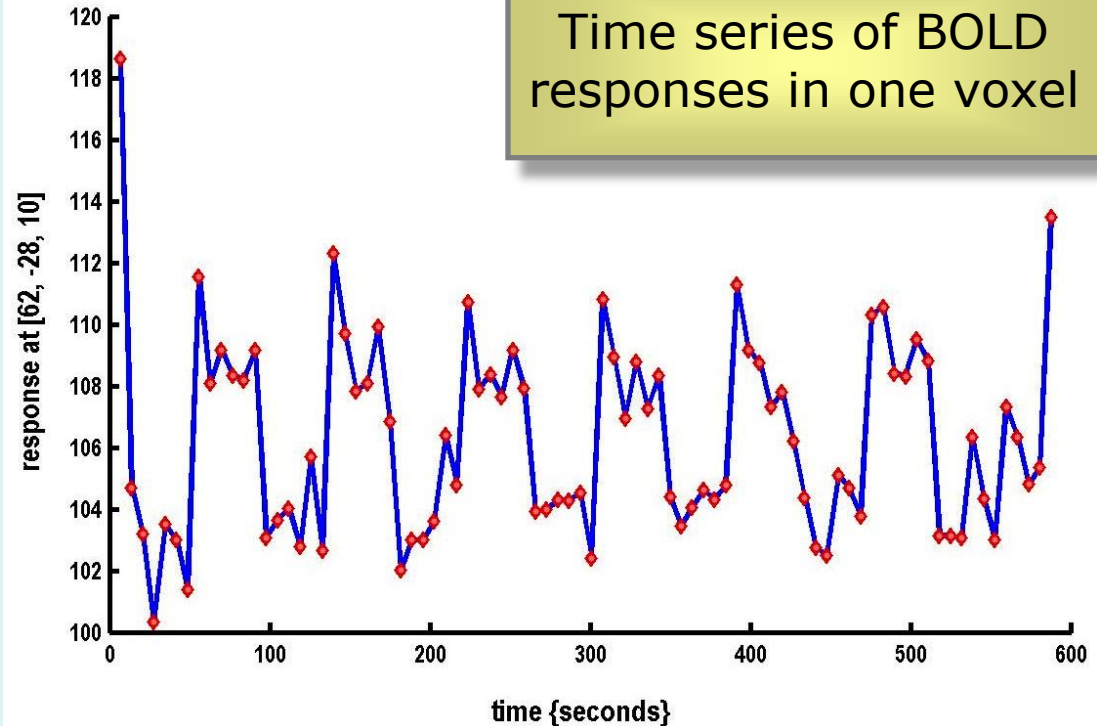
One session, one subject

Passive word listening versus rest

7 cycles of rest and listening

Each epoch 6 scans with 7 sec TR

Time series of BOLD responses in one voxel

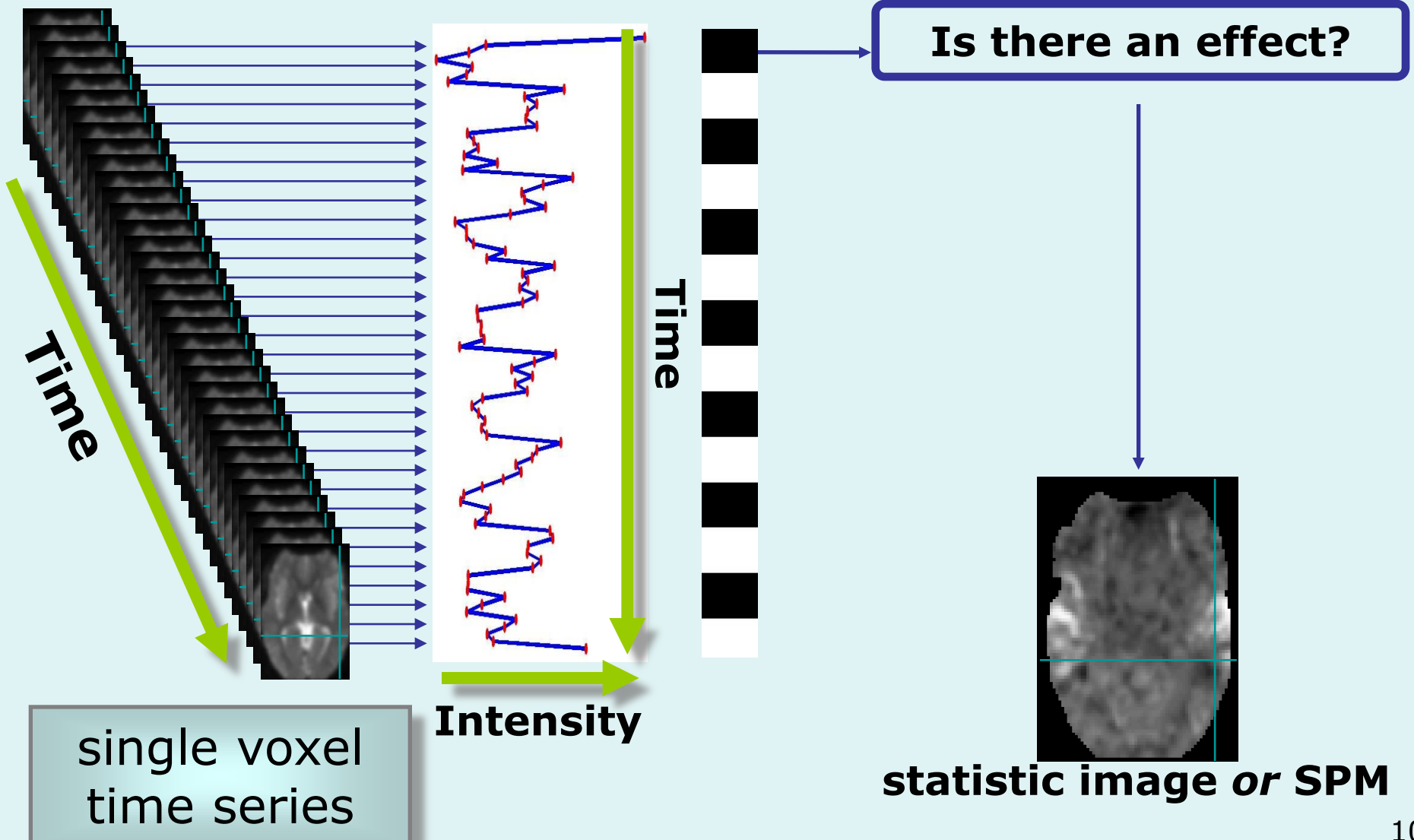


time {seconds}

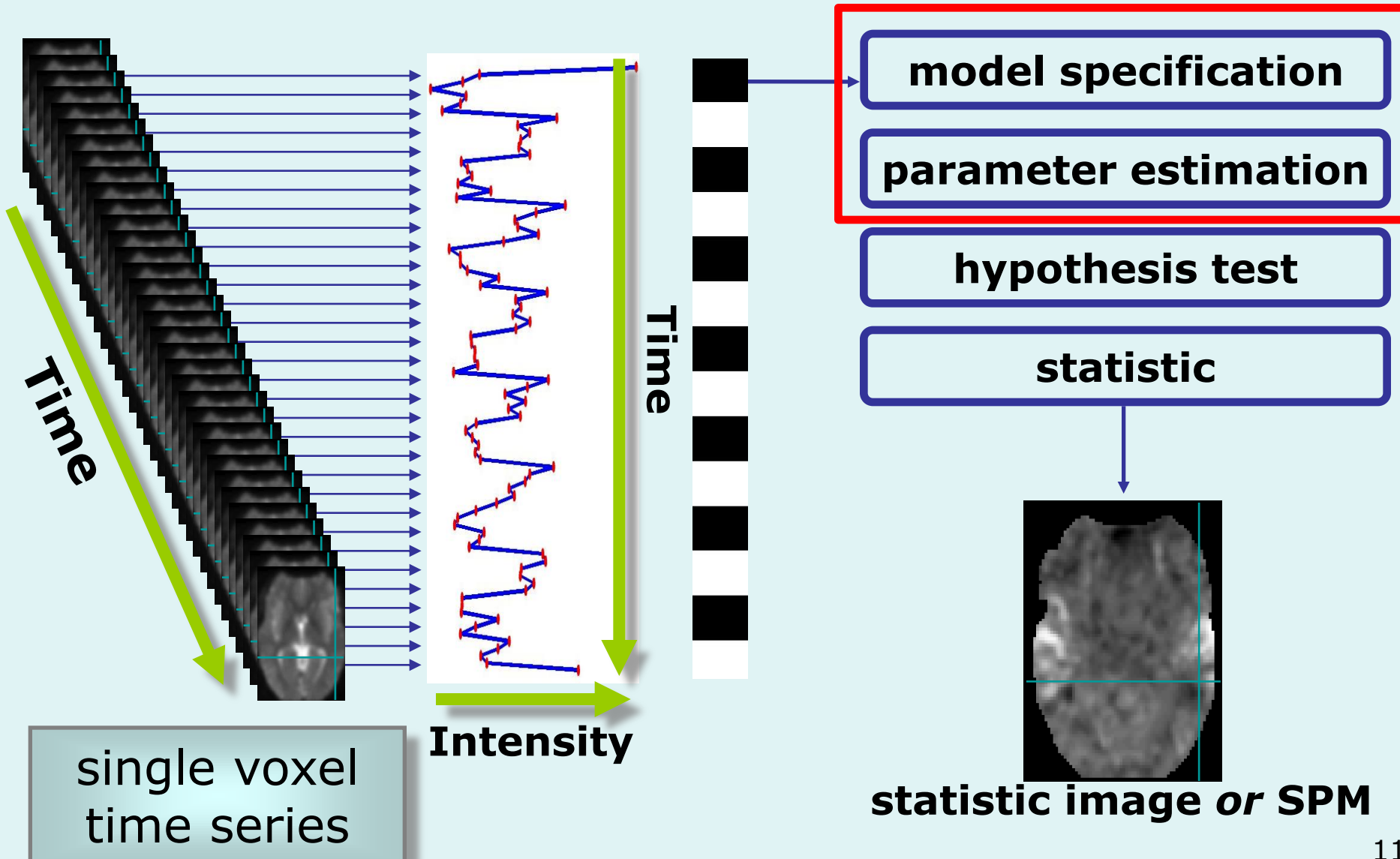
Stimulus function

Question: Is there a change in the BOLD response between listening and rest?

Voxel by voxel statistics



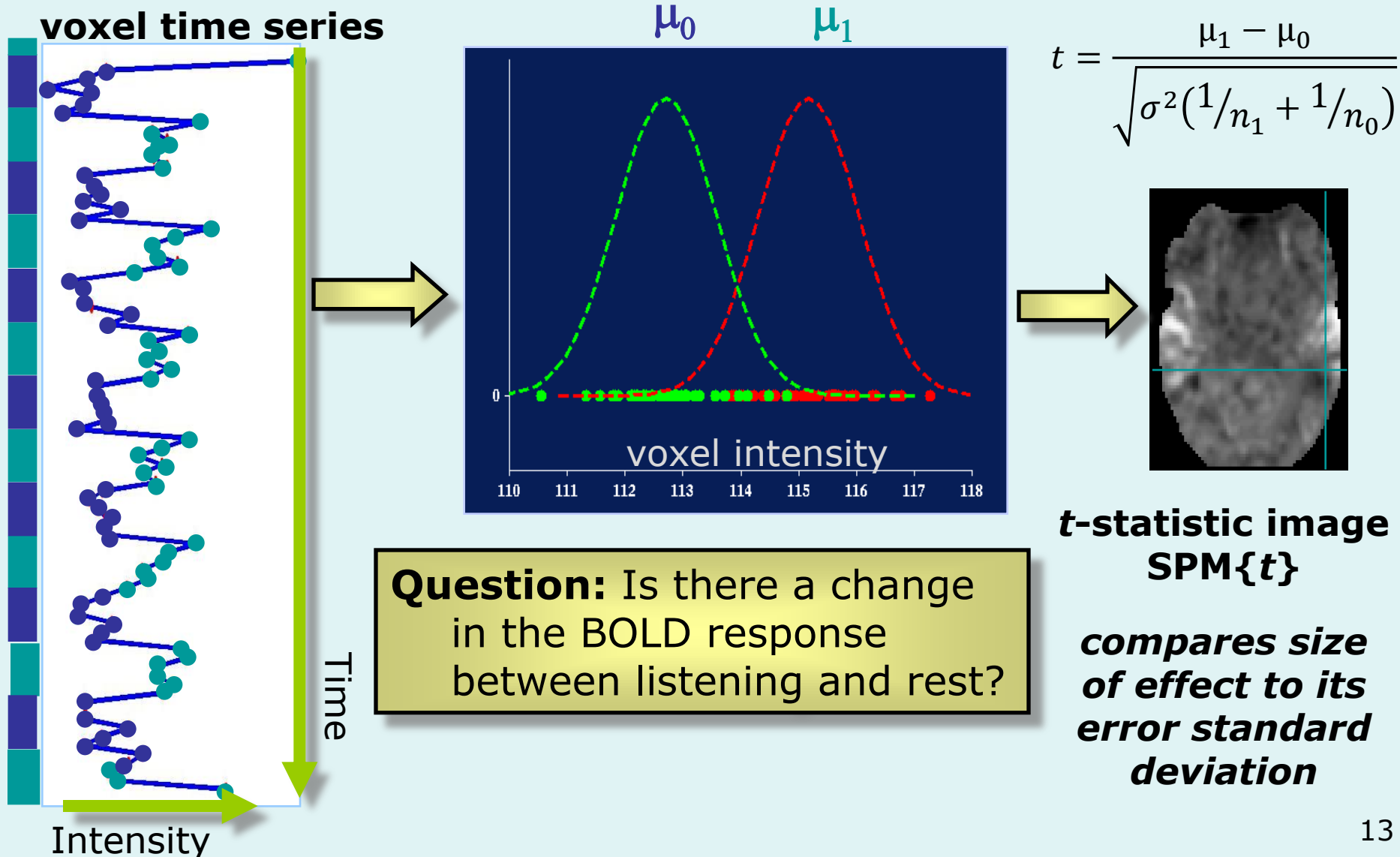
Voxel by voxel statistics



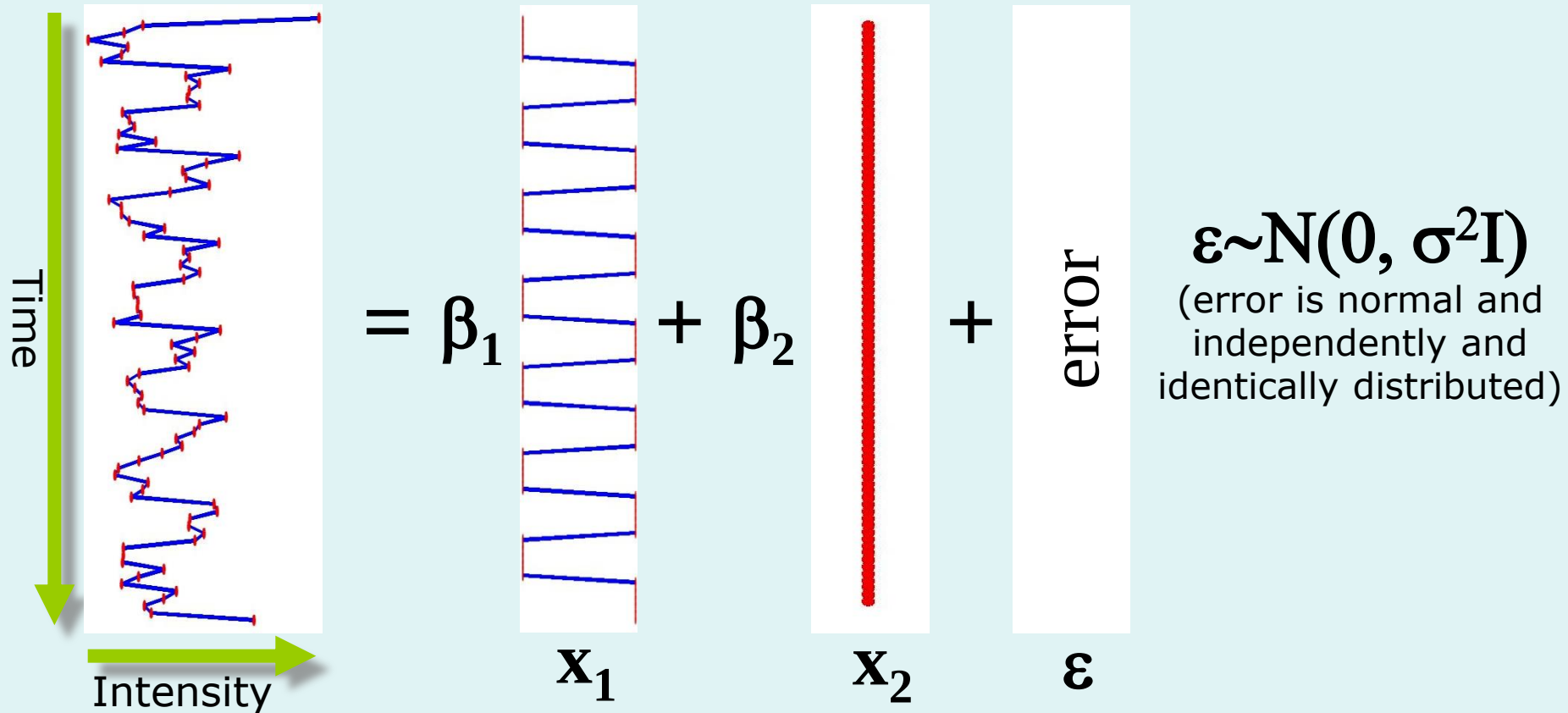
Content

- **Introduction**
- **General Linear Model**
- **Parameter estimation**
- **Improved model**
- **Conclusion**

Single voxel, two-sample t-test



Single voxel, regression model

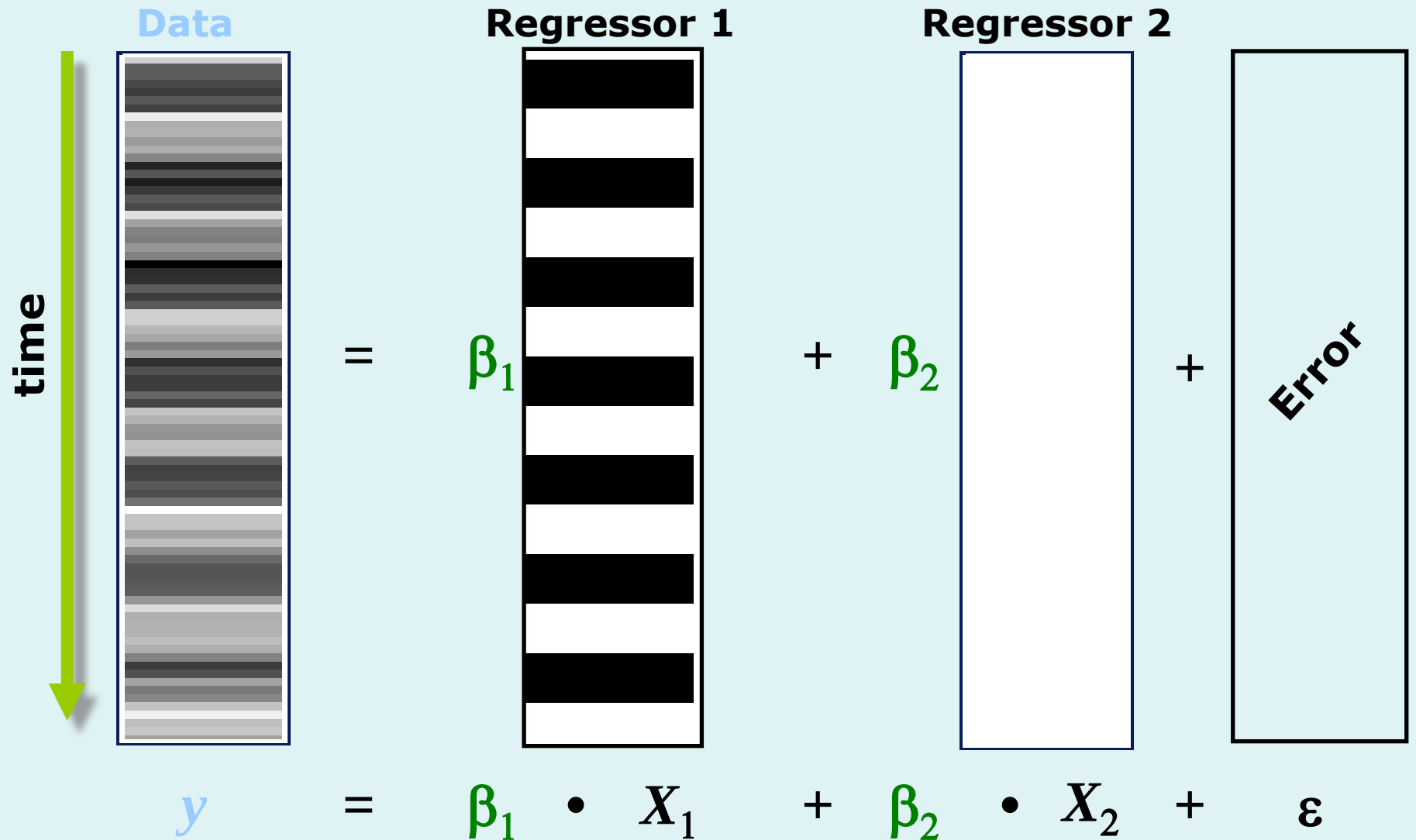


Question: Is there a change in the BOLD response between listening and rest?

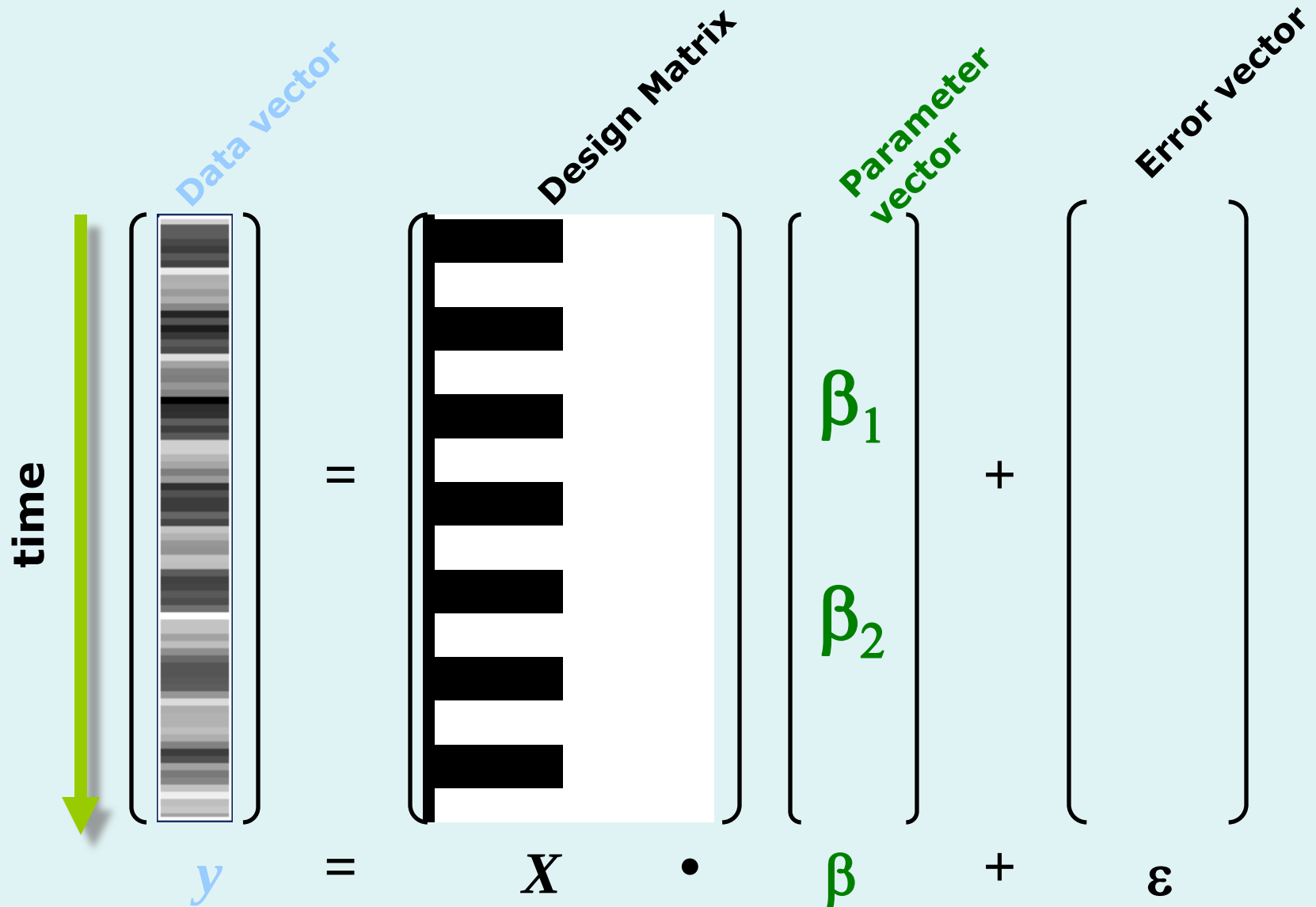


Hypothesis test:
 $\beta_1 = 0$?
(using t-statistic)

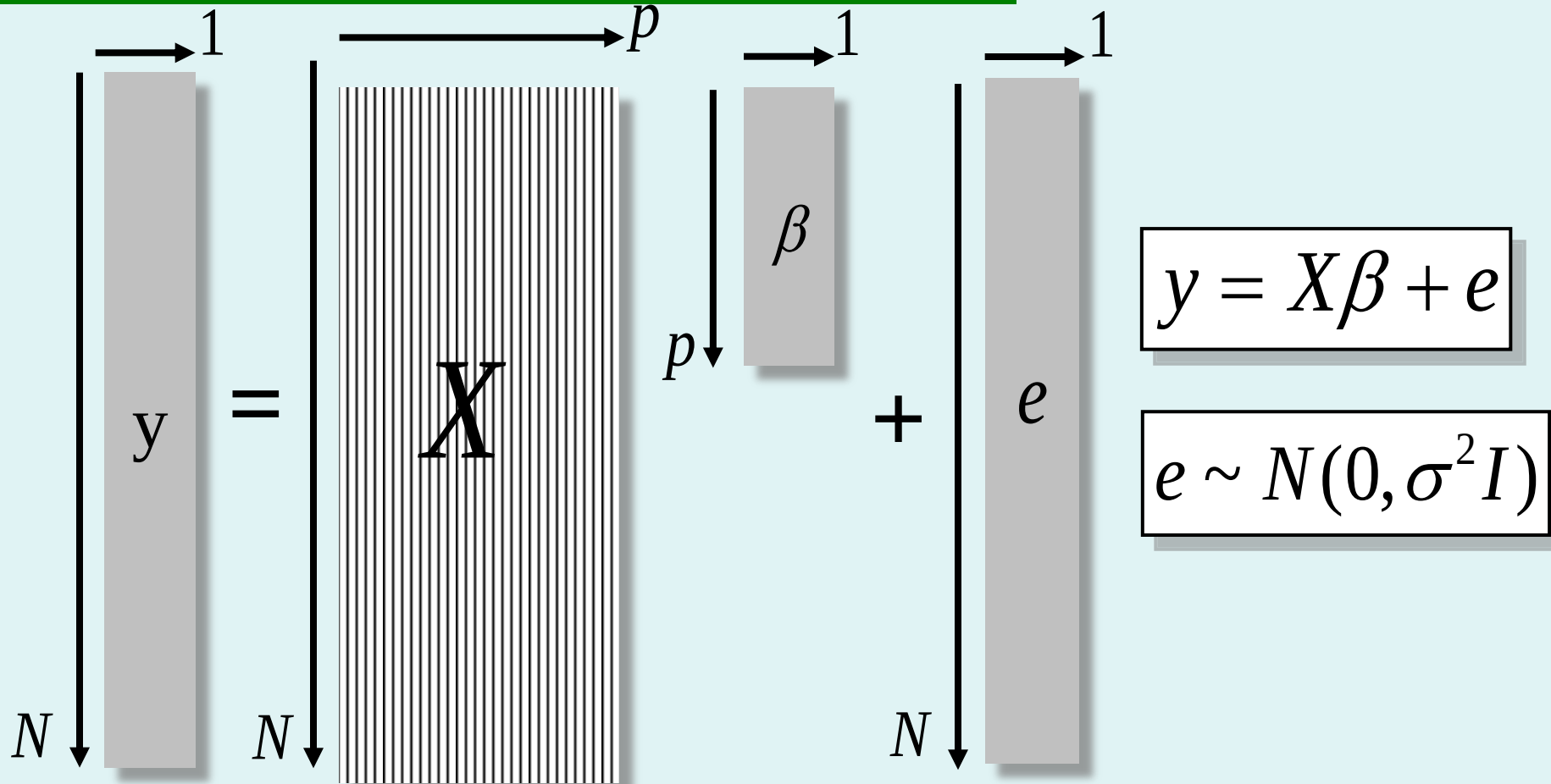
Model as basis functions



Design matrix



General Linear Model

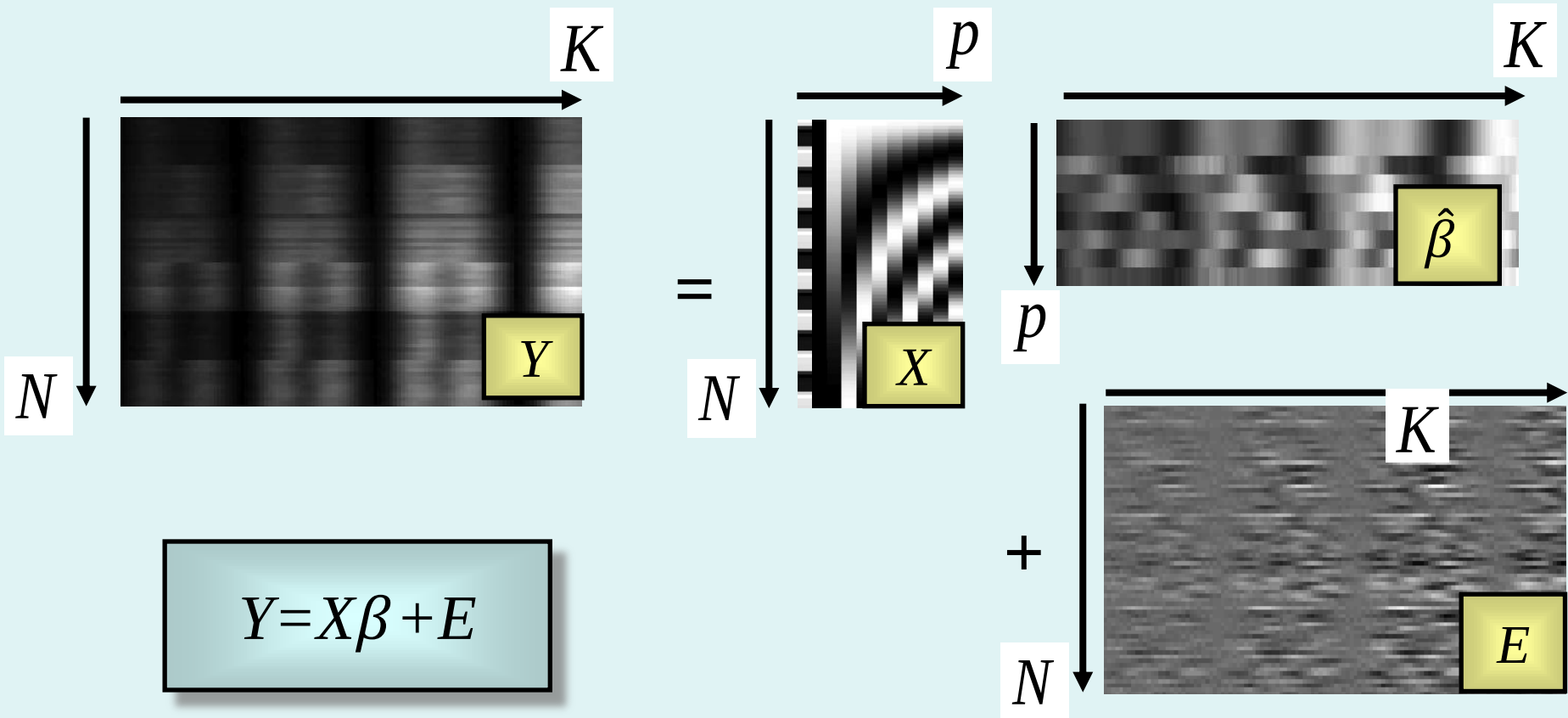


N : number of scans
 p : number of regressors

Model is specified by

1. Design matrix X
2. Assumptions about ε

GLM & Mass univariate approach



The design matrix embodies all available knowledge about experimentally controlled factors and potential confounds.

Classical statistics

- parametric
 - one sample t -test
 - two sample t -test
 - paired t -test
 - Anova
 - AnCova
 - correlation
 - linear regression
 - multiple regression
 - F -tests
 - etc...

all cases of the

General Linear Model

assume normality

to account for serial correlations:

Generalised Linear Model

- non-parametric?

→ **SnPM**

Content

- **Introduction**
- **General Linear Model**
- **Parameter estimation**
- **Improved model**
- **Conclusion**

Parameter estimation

$$y = X \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + e$$

y X e

$$y = X\beta + e$$

Objective:
estimate
parameters to
minimize

$$\sum_{t=1}^N e_t^2$$



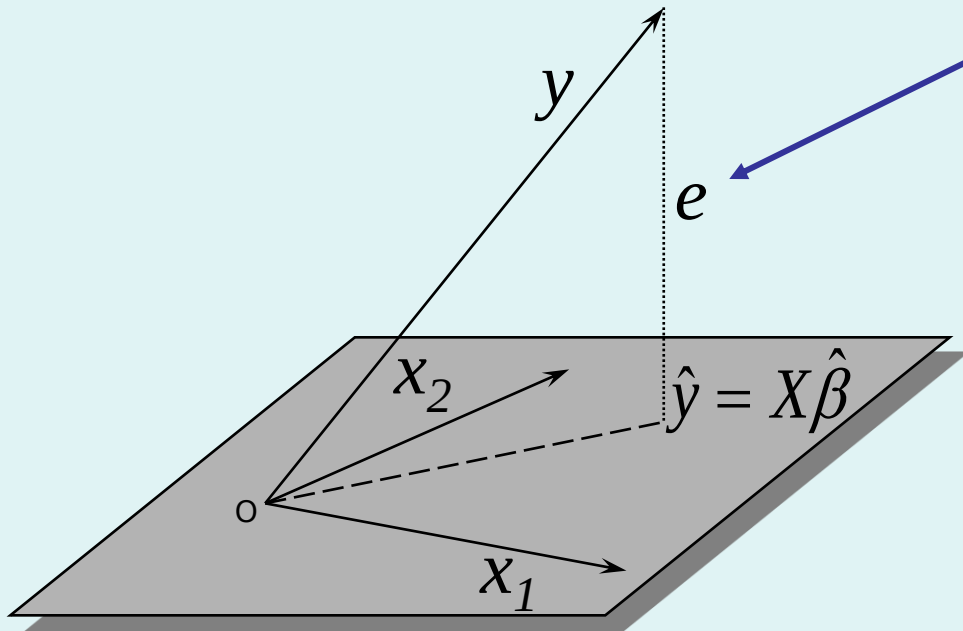
Ordinary least squares
estimation (OLS) (assuming
i.i.d. error):

$$\hat{\beta} = (X^T X)^{-1} X^T y$$

$$\hat{\beta} \sim N(\beta, \sigma^2 (X^T X)^{-1})$$

Geometric perspective on the GLM

Ordinary Least Squares (OLS)



Smallest errors (shortest error vector) when e is *orthogonal* to X

$$X^T e = 0$$

$$X^T (y - X\hat{\beta}) = 0$$

$$X^T y = X^T X \hat{\beta}$$

$$\hat{\beta} = (X^T X)^{-1} X^T y$$

Design space
defined by X

N data points $\rightarrow$ N dimension space !

Content

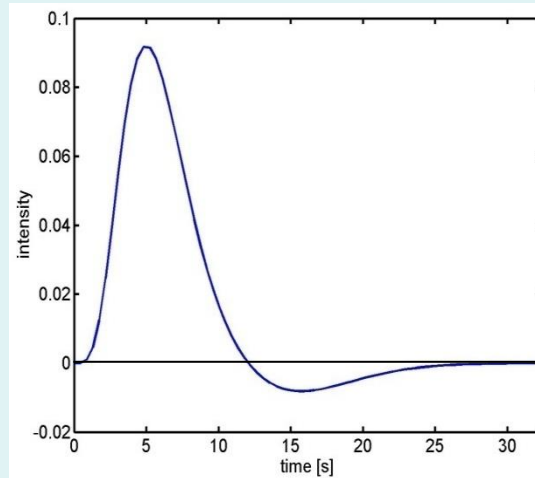
- **Introduction**
- **General Linear Model**
- **Parameter estimation**
- **Improved model**
- **Conclusion**

Problems with fMRI time series

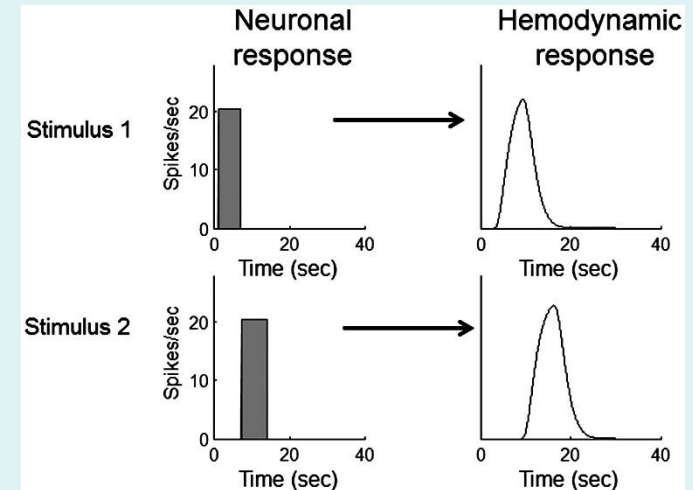
1. The **BOLD response** has a delayed and dispersed shape.
2. The BOLD signal includes substantial amounts of **low-frequency noise** (e.g. due to scanner drift).
3. Due to breathing, heartbeat & unmodeled neuronal activity, the **errors are serially correlated**. This violates the assumptions of the noise model in the GLM.

Problem 1: BOLD response

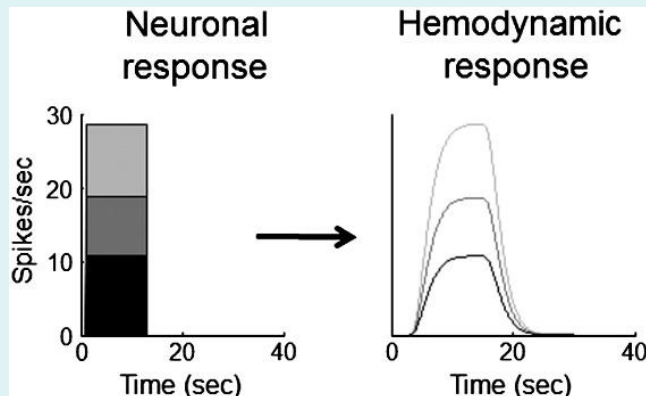
Hemodynamic response function (HRF):



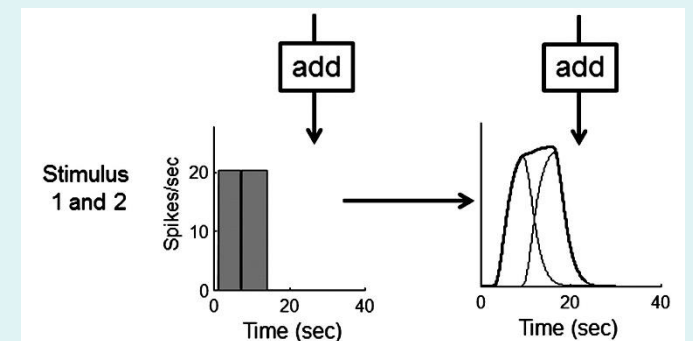
Shift invariance



Scaling

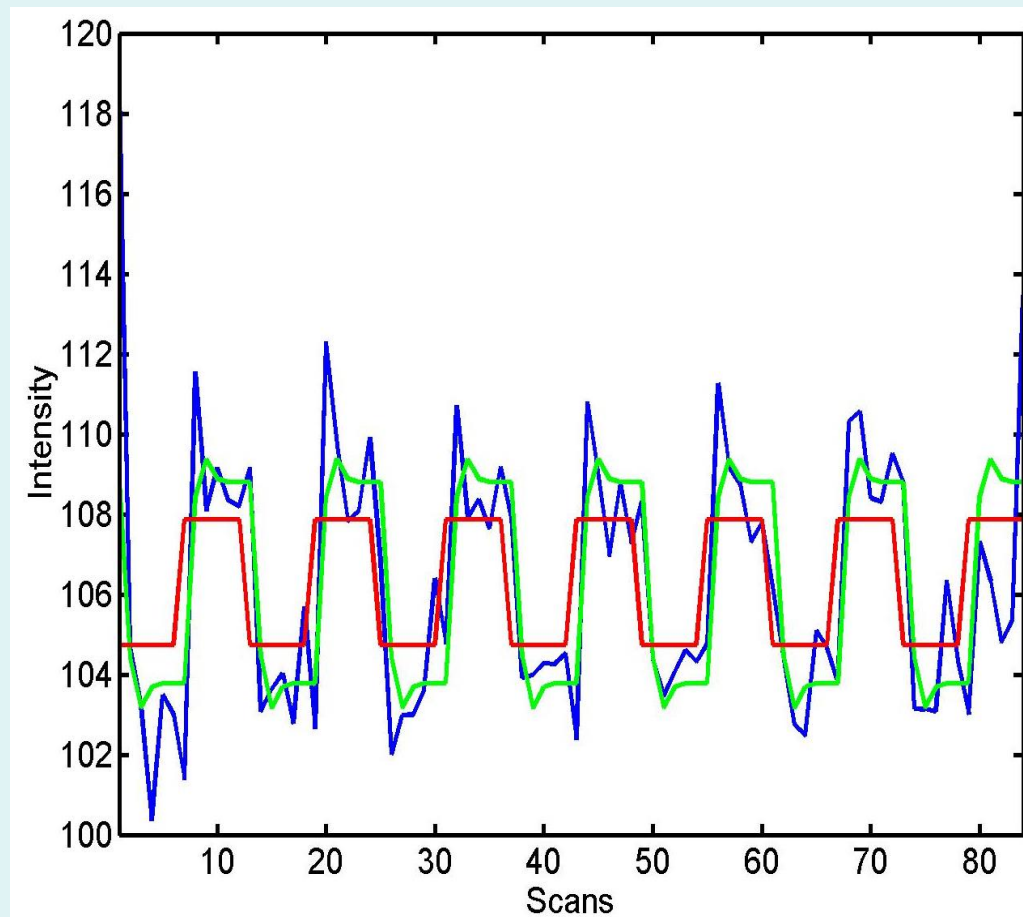
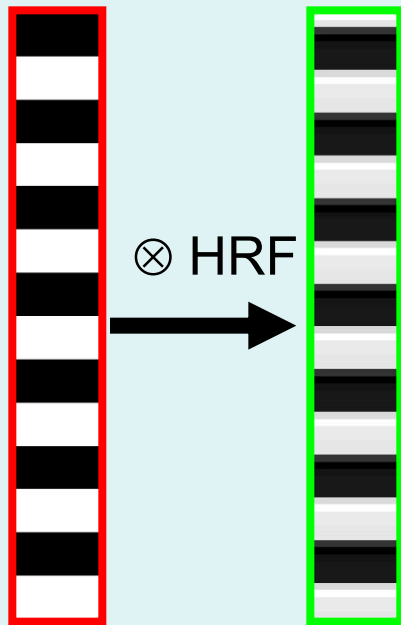


Additivity



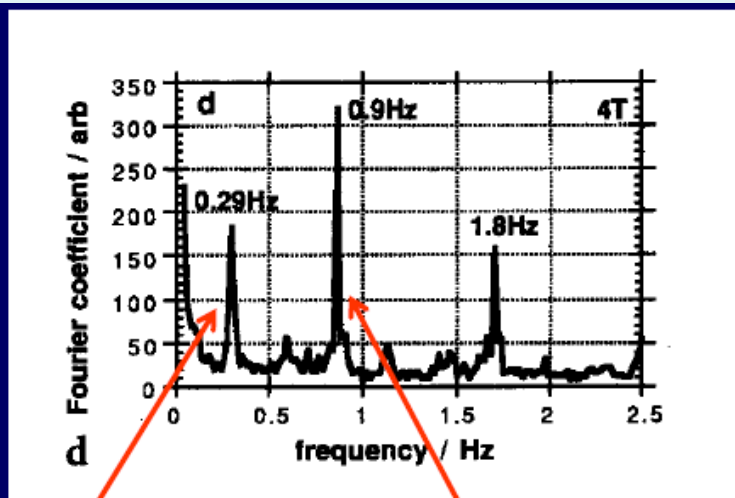
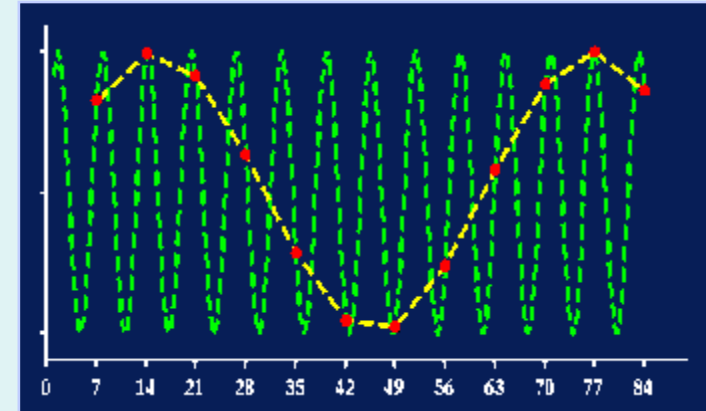
Solution for the BOLD response

Convolve stimulus function with a canonical hemodynamic response function (HRF):



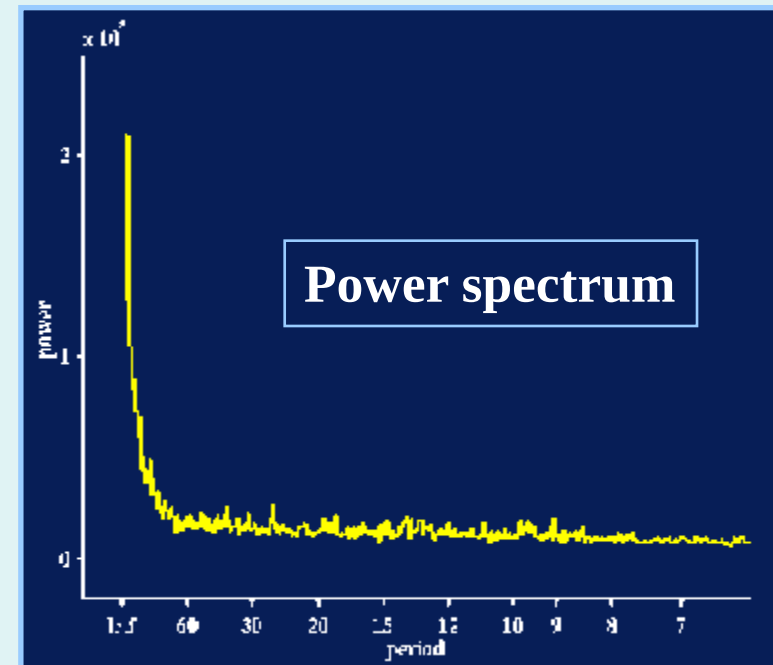
Problem 2: Low frequency noise

- Physiological noise + scanner drift
 - Aliased high frequency effects
- ⇒ Power in the low frequencies

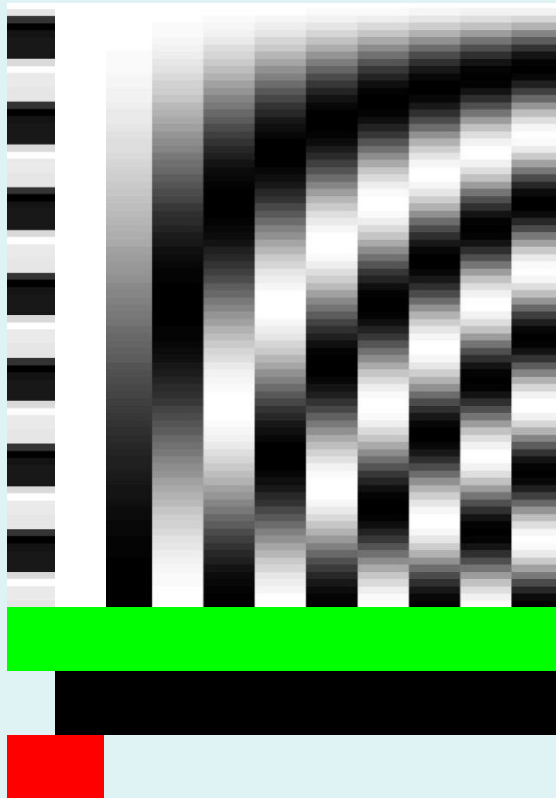


Respiration
every 4-10 s (0.3 Hz)

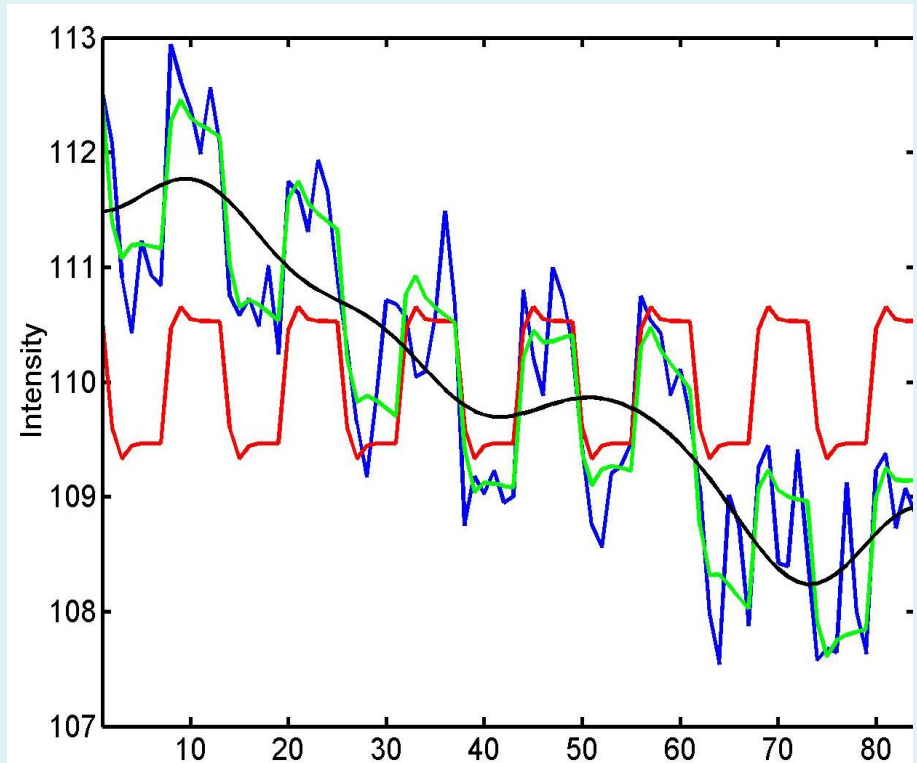
Cardiac cycle
every ~1 s (0.9 Hz)



Solution with high pass filtering



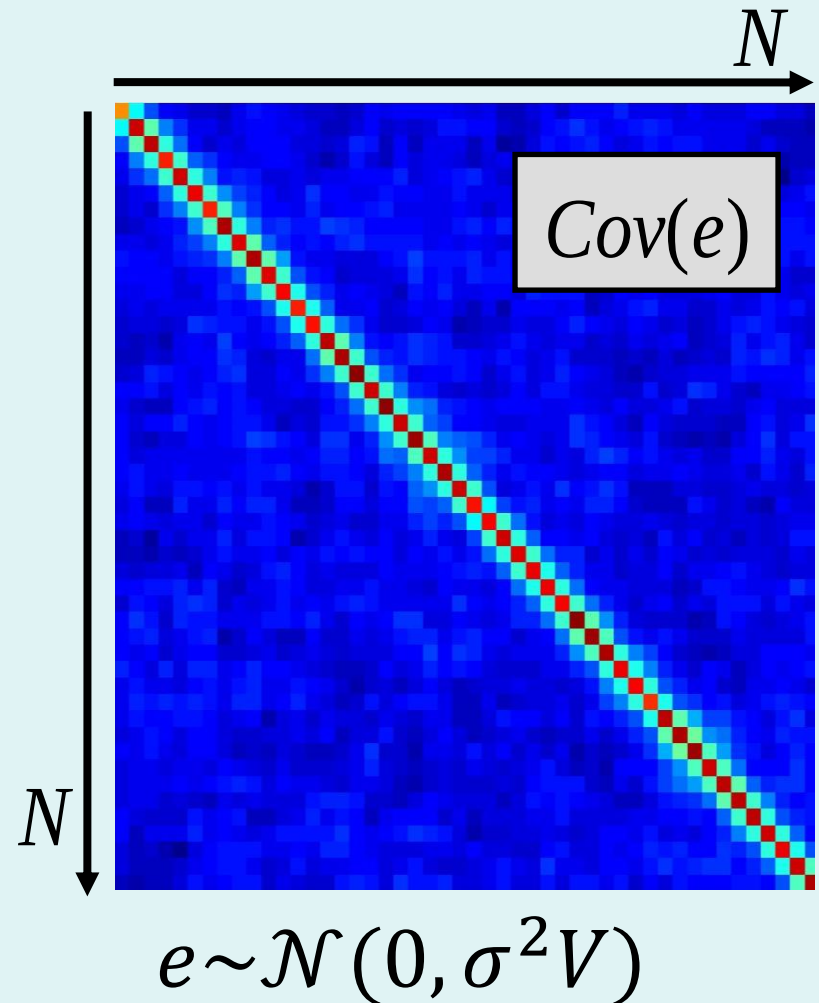
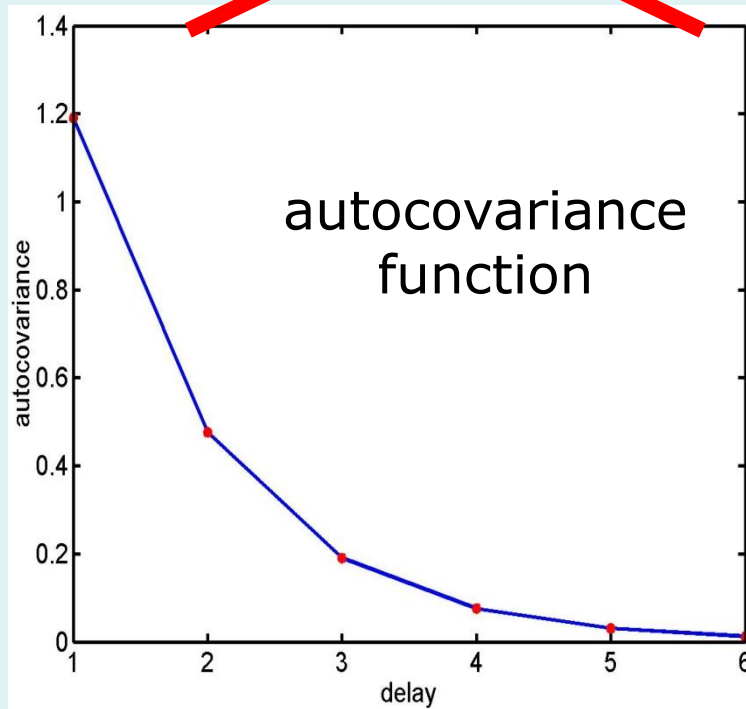
discrete cosine
transform (DCT) set



blue = data
black = mean + low-frequency drift
green = predicted response, taking into account low-frequency drift
red = predicted response, NOT taking into account low-frequency drift

Problem 3: Serial correlations

~~i.i.d: $e \sim \mathcal{N}(0, \sigma^2 I)$~~

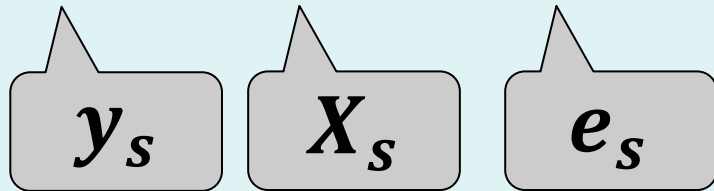


Solution for serial correlations

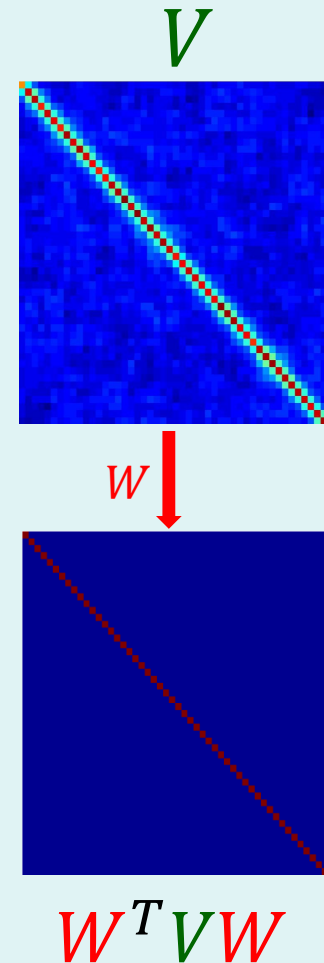
$$y = X\beta + e \quad e \sim \mathcal{N}(0, \sigma^2 V)$$

$$\text{Let } W^T W = V^{-1}$$

$$Wy = WX\beta + We \quad We \sim \mathcal{N}(0, \sigma^2 \underbrace{W^T V W}_I)$$



Solution : Whitening the data
BUT this requires an estimation of V



Equivalent to the Weighted Least Square estimator

Multiple covariance components

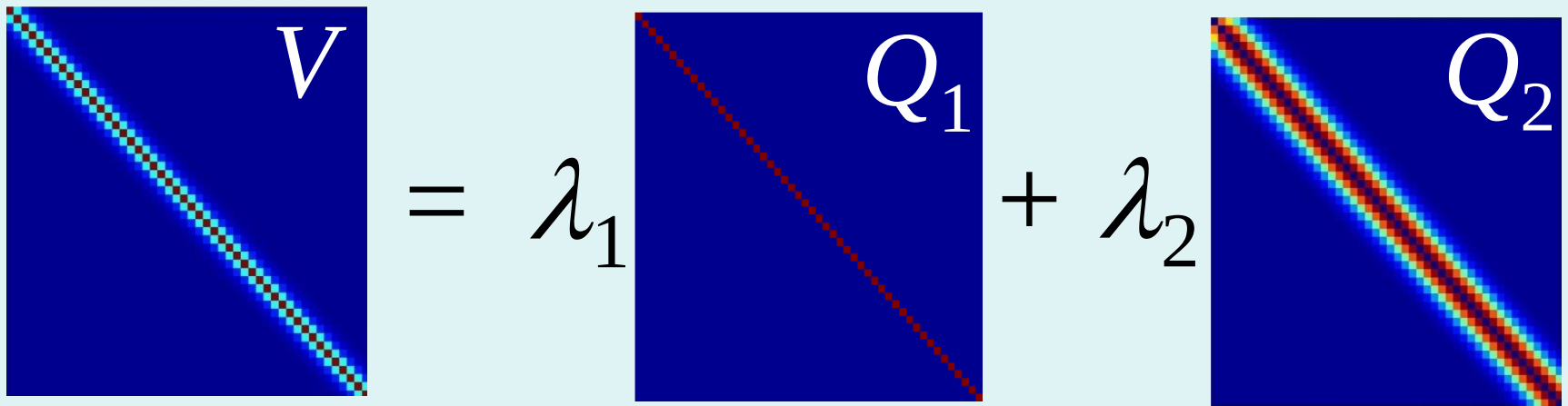
enhanced noise model at voxel i

$$e_i \sim N(0, C_i)$$

$$C_i = \sigma_i^2 V$$

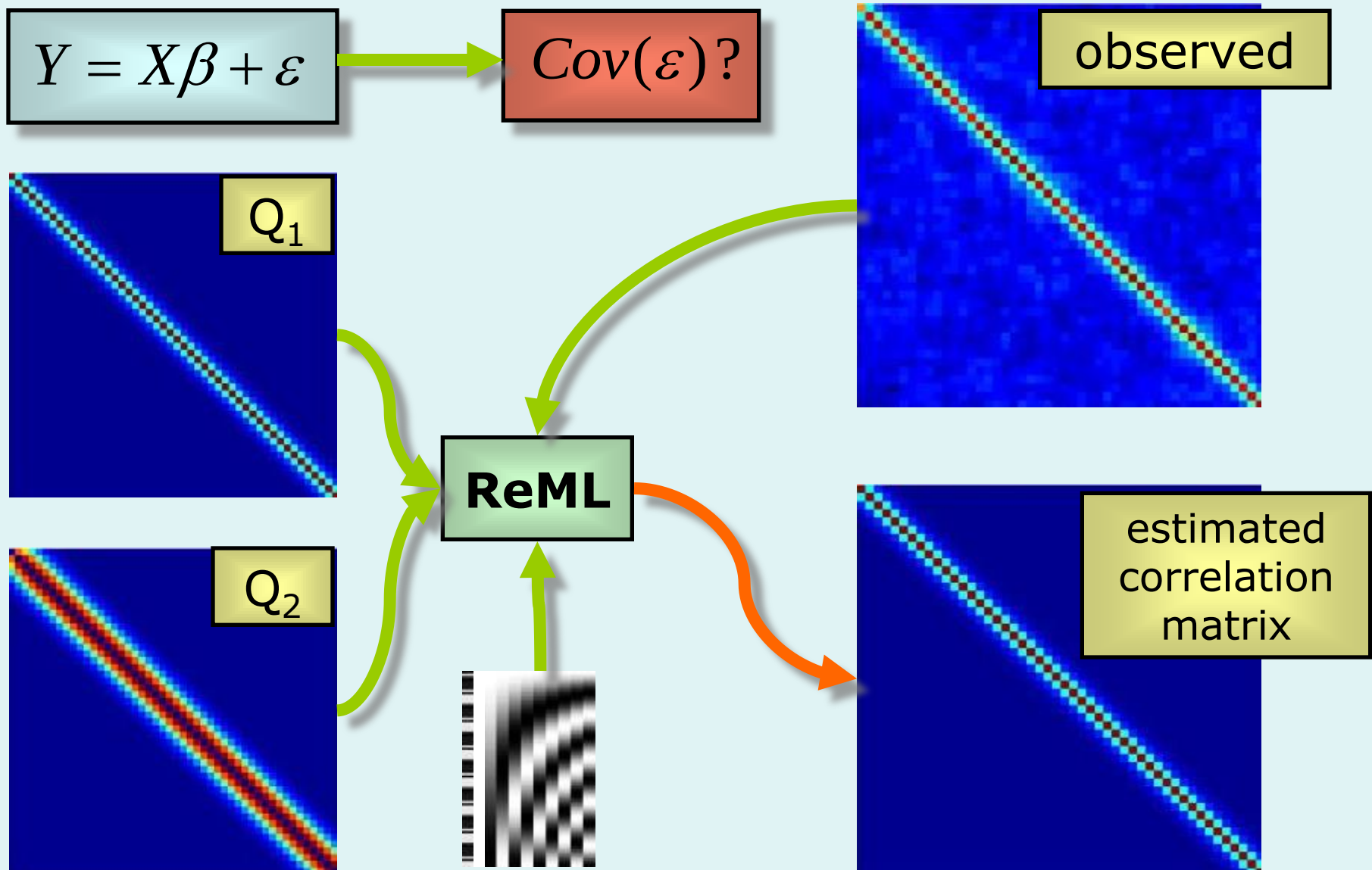
$$V = \sum \lambda_j Q_j$$

error covariance components
 Q and hyperparameters λ

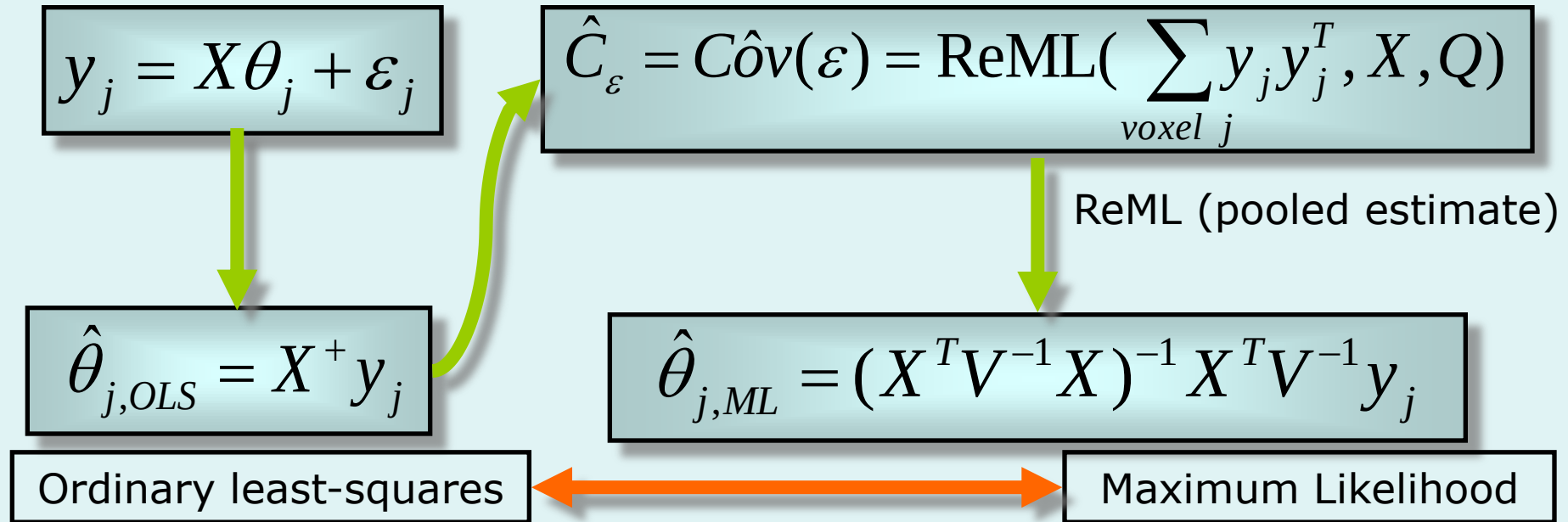


Estimation of hyperparameters λ with ReML (Restricted Maximum Likelihood).

Restricted Maximum Likelihood



Estimation in SPM



- 2 passes (first pass for selection of voxels)
 - more accurate estimate of V
- Assume, at voxel j : $C_{\varepsilon,j} = \sigma_j^2 V$

$$t = \frac{c^T \theta}{\text{SE}(c^T \theta)}$$

$$\text{SE}(c^T \theta) = \sqrt{\hat{\sigma}^2 c^T (V^{-1/2} X)^- (V^{-1/2} X)^{-T} c}$$

Limitations



The AR(1)+white noise model may not be enough for short TR (<1.5 s)

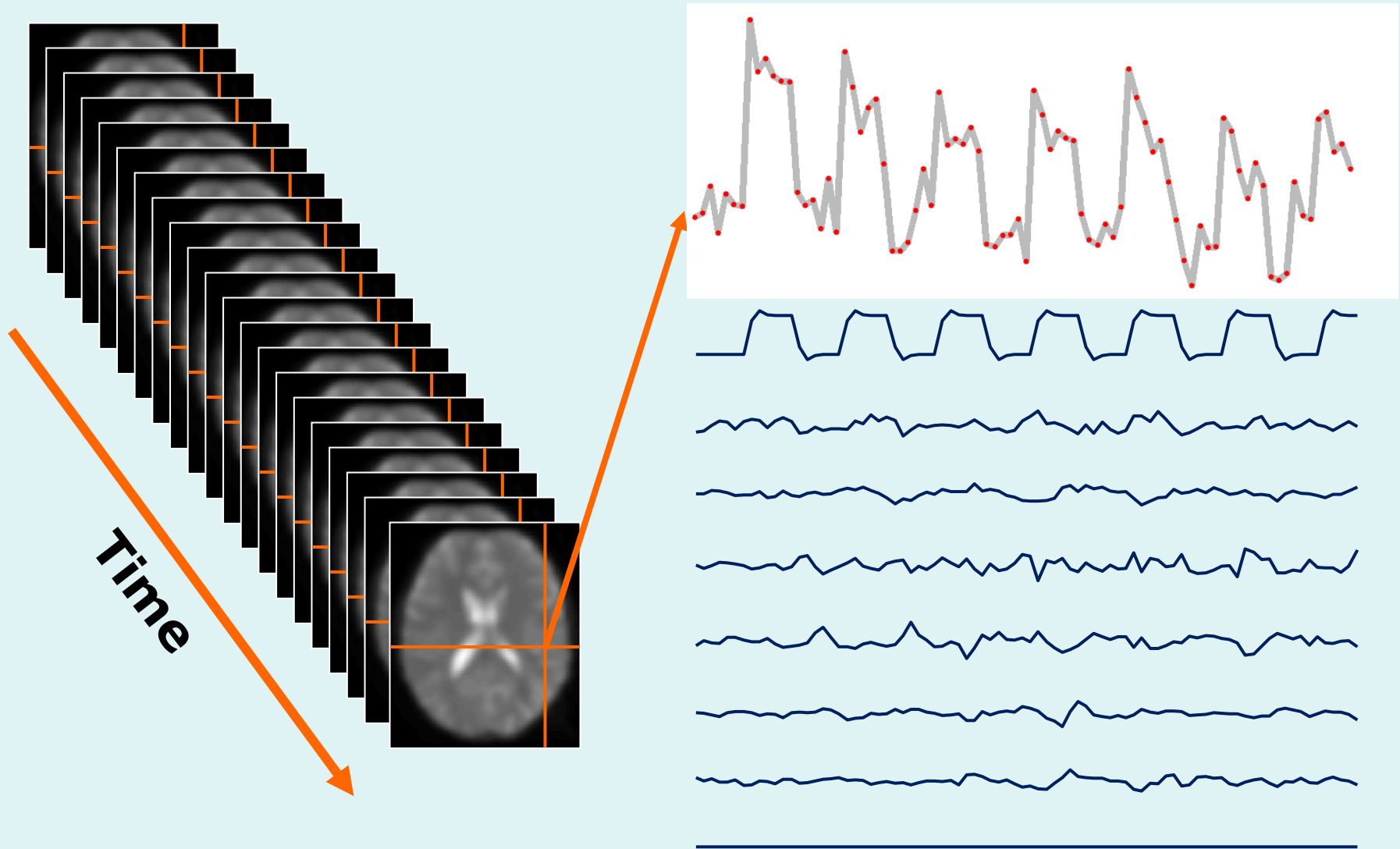
$$V = \lambda_1 Q_1 + \lambda_2 Q_2 + \lambda_3 Q_3 + \lambda_4 Q_4 + \lambda_5 Q_5 + \dots$$

The flexibility of the ReML enables the use of any number of components of any shape

Content

- **Introduction**
- **General Linear Model**
- **Parameter estimation**
- **Improved model**
- **Conclusion**

A mass univariate approach

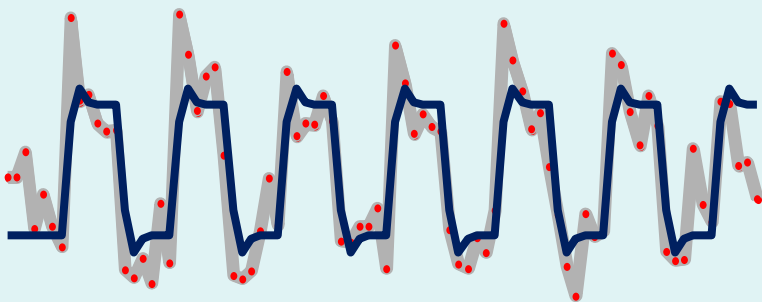


Summary

Mass univariate approach:

- Fit GLMs with
 - design matrix, X ,
 - to data at different points in space
 - to estimate local effect sizes, β
- GLM, a very general approach that accommodates
 - Hemodynamic Response Function
 - Nuisance effects, e.g. high pass filtering
 - Error term covariance, e.g. temporal autocorrelation

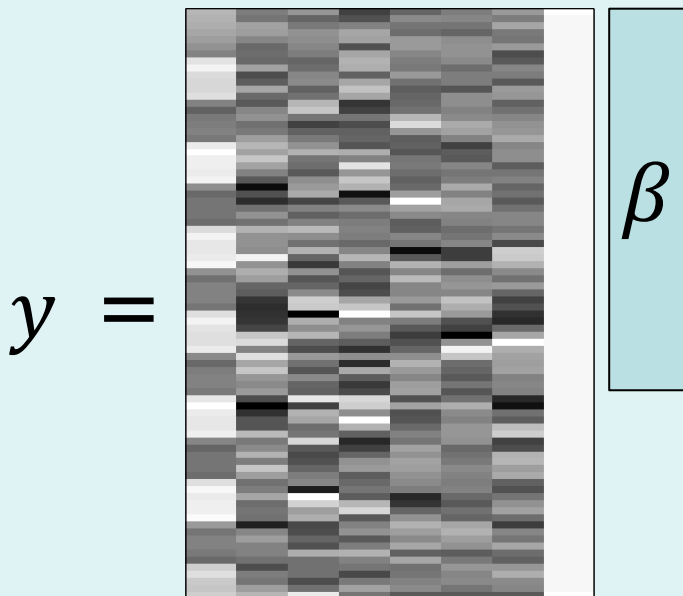
Summary



noise assumptions: $\varepsilon \sim N(0, \sigma^2 V)$

Pre-whitening: $X_S = WX$ $y_S = Wy$ $\varepsilon_S = W\varepsilon$

$$\hat{\beta} = (X_S^T X_S)^{-1} X_S^T y_S$$



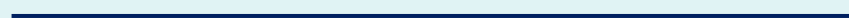
$$\hat{\beta}_1 = 3.9831$$



$$\hat{\beta}_{2-7} = \{0.6871, 1.9598, 1.3902, 166.1007, 76.4770, -64.8189\}$$



$$\hat{\beta}_8 = 131.0040$$



$+\varepsilon$



$$\hat{\beta} \sim N(\beta, \sigma^2 (X_S^T X_S)^{-1})$$

$$\hat{\sigma}^2 = \frac{\widehat{\varepsilon}_S^T \widehat{\varepsilon}_S}{N-p}$$

Why modelling?

Why?

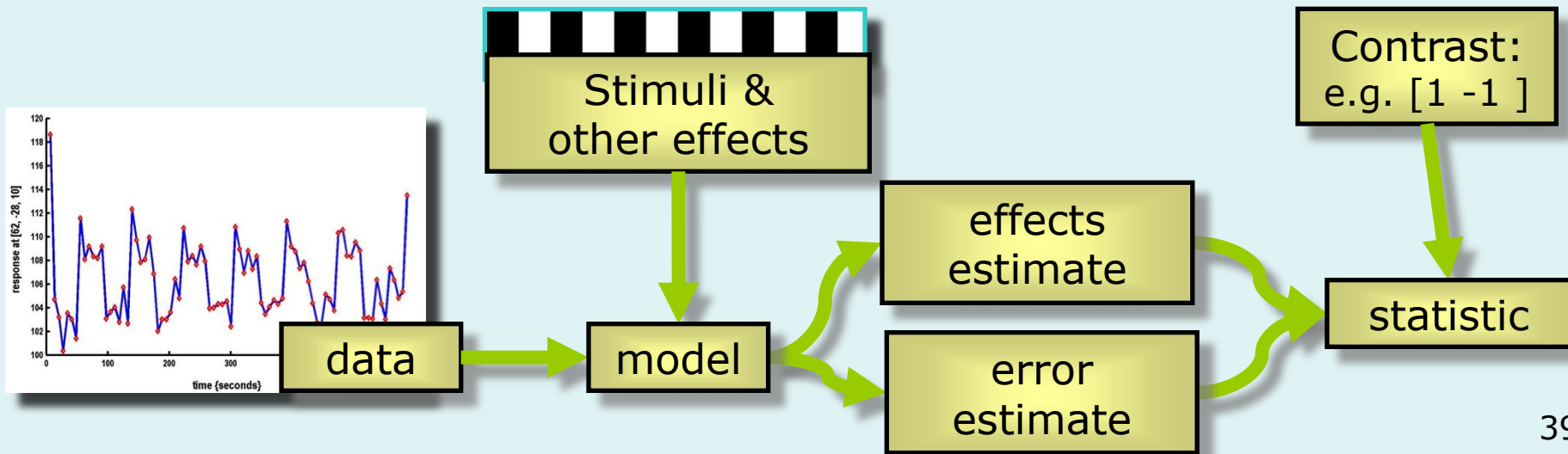
Make inferences about effects of interest

How?

1. Decompose data into effects and error
2. Form statistic using estimates of effects and error

Model?

Use any available knowledge



References

- Statistical parametric maps in functional imaging: a general linear approach, K.J. Friston et al, Human Brain Mapping, 1995.
- Analysis of fMRI time-series revisited – again, K.J. Worsley and K.J. Friston, NeuroImage, 1995.
- The general linear model and fMRI: Does love last forever?, J.-B. Poline and M. Brett, NeuroImage, 2012.
- Linear systems analysis of the fMRI signal, G.M. Boynton et al, NeuroImage, 2012.

