Inverse problems in functional brain imaging

Ph. Ciuciu[1,2]

philippe.ciuciu@cea.fr www.lnao.fr

1: CEA/NeuroSpin/LNAO

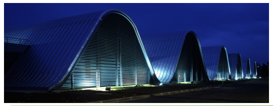
2: IFR49

July 24, 2009

GDR -ISIS Summer school -Peyresq

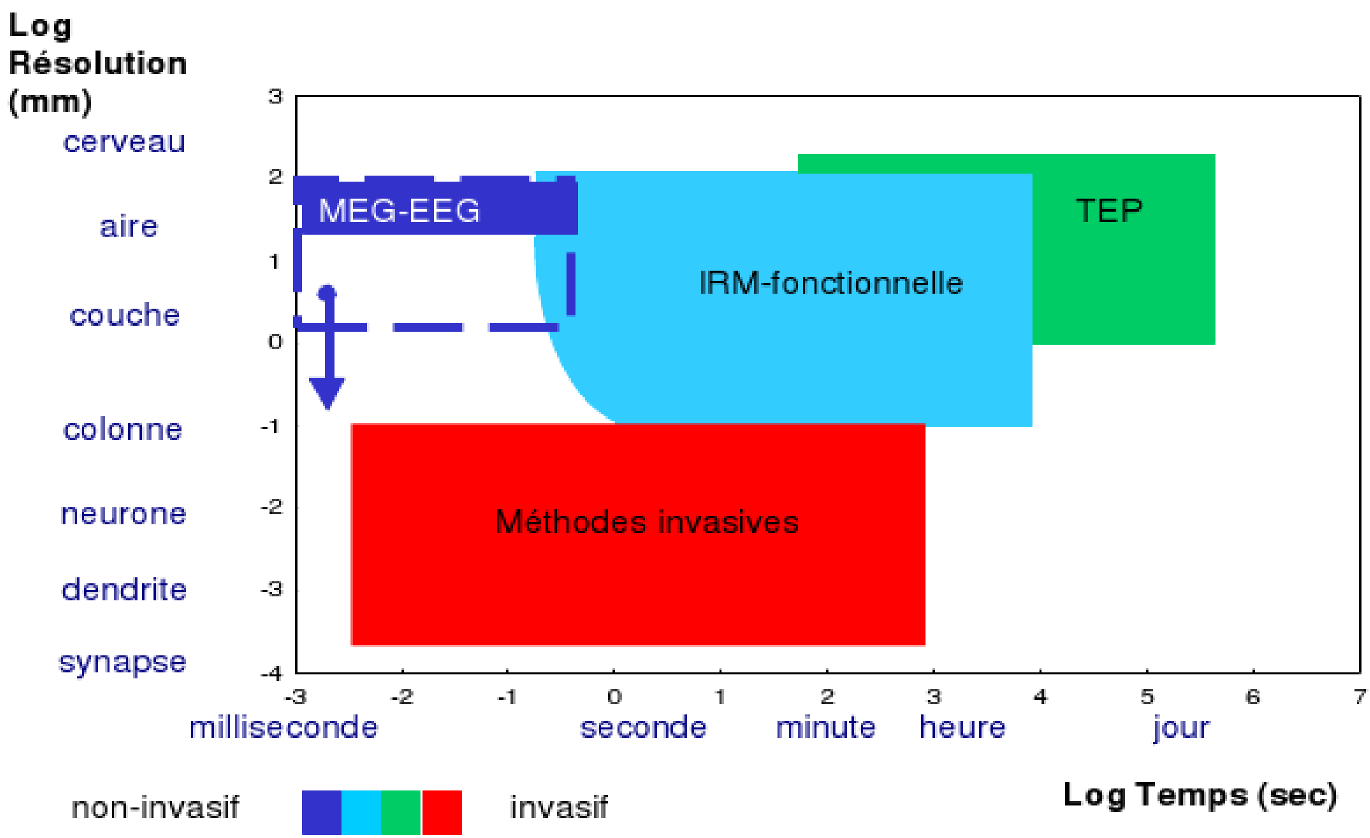
Outline

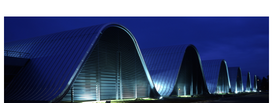
- Introduction to functional brain imaging
  - Magneto/Electro-encephalography
  - functional MRI
- Mapping brain activity
  - GLM framework (Statistical Parametric Mapping)
  - Spatial regularization [VB optimization]
- Probe brain dynamics in fMRI
  - FIR modeling
  - Temporal regularization [EM/SAEM algorithms]
- Joint detection-estimation
  - Unify both questions
  - Spatio-temporal regularization [MCMC methods]



Functional brain imaging

- Imagerie Fonctionnelle Cérébrale
    - Etude du cerveau en action
    - Nombreuses applications cliniques et en Sciences Cognitives
    - À l 'interface entre les Sciences de la Vie et les Sciences pour l 'Ingénieur
- Techniques
    - Où et Quand ?





Functional brain imaging

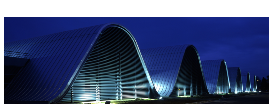
Probe brain dynamics
non-invasively

[Ogawa et al, 1990,1992]

stimulus → neuronal activity → metabolism + hemodynamics

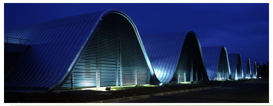
neuronal activity → electromagnetic activity → **EEG / MEG**

metabolism + hemodynamics → **BOLD fMRI**



Outline

- Introduction to functional brain imaging
  - Magneto/Electro-encephalography
  - functional MRI
- Mapping brain activity
  - GLM framework (Statistical Parametric Mapping)
  - Spatial regularization [VB optimization]
- Probe brain dynamics in fMRI
  - FIR modeling
  - Temporal regularization [EM/SAEM algorithms]
- Joint detection-estimation
  - Unify both questions
  - Spatio-temporal regularization [MCMC methods]



MEG/EEG

Électroencéphalographie (EEG)

Activité électrique neuronale

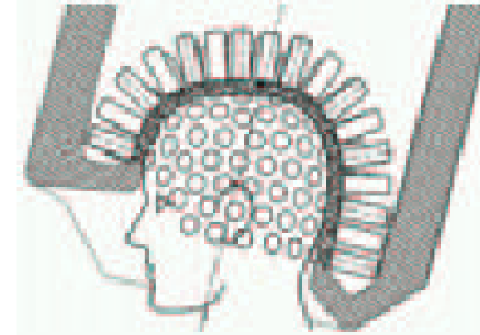
Résolution temporelle : ~ 1ms

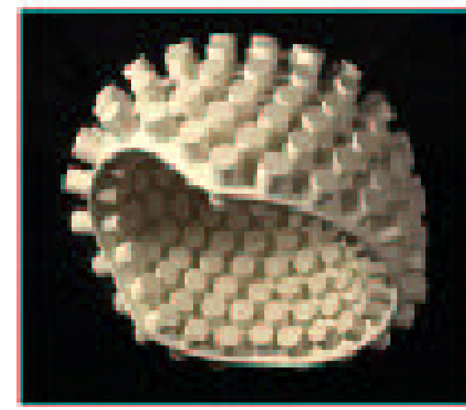
EEG : mesure du potentiel électrique

Ordre de grandeur : qq μ-volts

Capteurs : électrodes

Magnétoencéphalographie (MEG)

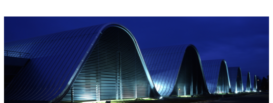




MEG : mesure du champ magnétique

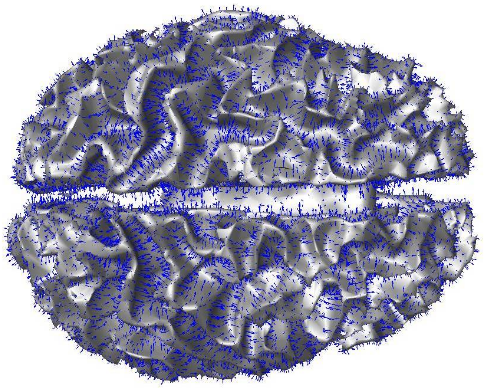
Ordre de grandeur : 10^{-13} Tesla

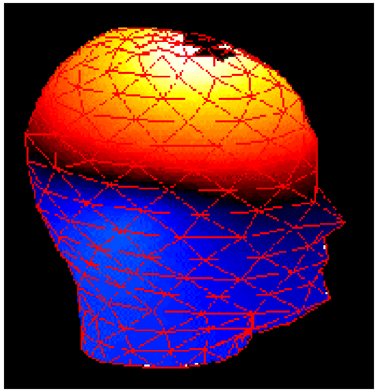
Capteurs SQUID couplés à des bobines



MEG/EEG source reconstruction

Inverse procedure





Distributed Source model

Forward modelling

Data

J K $Y = K * J + E$

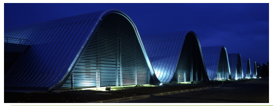
$$Y = KJ + E$$

[nxt] [nxp][pxt] [nxt]

n : number of sensors
p : number of dipoles
t : number of time samples

- under-determined system
- priors required

Bayesian framework



[Mattout et al., Neuroimage, 2006]

MEG/EEG source reconstruction

posterior — likelihood — prior

$$p(J|Y) \propto p(Y|J)\,p(J)$$

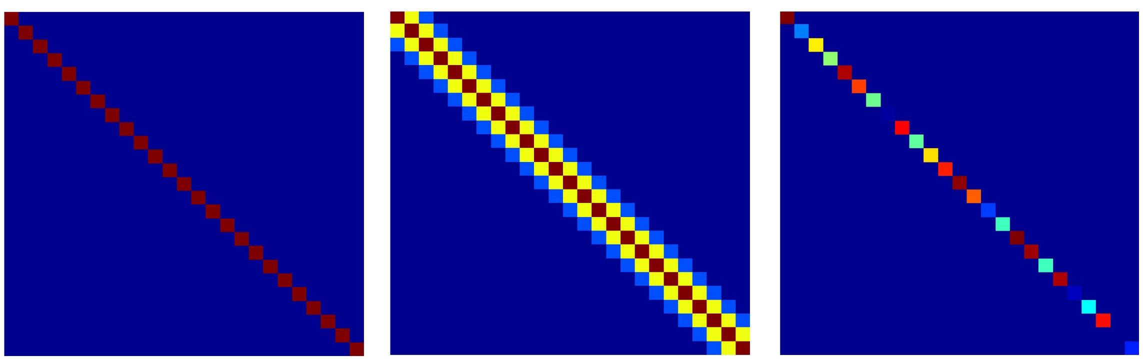
$$U_{MAP}(J) = \left\| C_e^{-1/2}(Y - KJ) \right\|^2 + \lambda \left\| WJ \right\|^2$$

likelihood — WMN prior

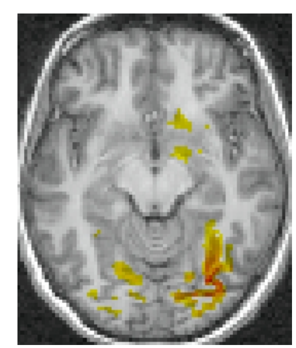
$$p(J) \sim N(0, C_j) \qquad C_j^{-1} = \lambda W^T W$$

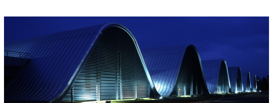
2-level hierarchical model:

$$Y = KJ + E_1 \qquad E_1 \sim N(0, C_e)$$
$$J = 0 + E_2 \qquad E_2 \sim N(0, C_p)$$



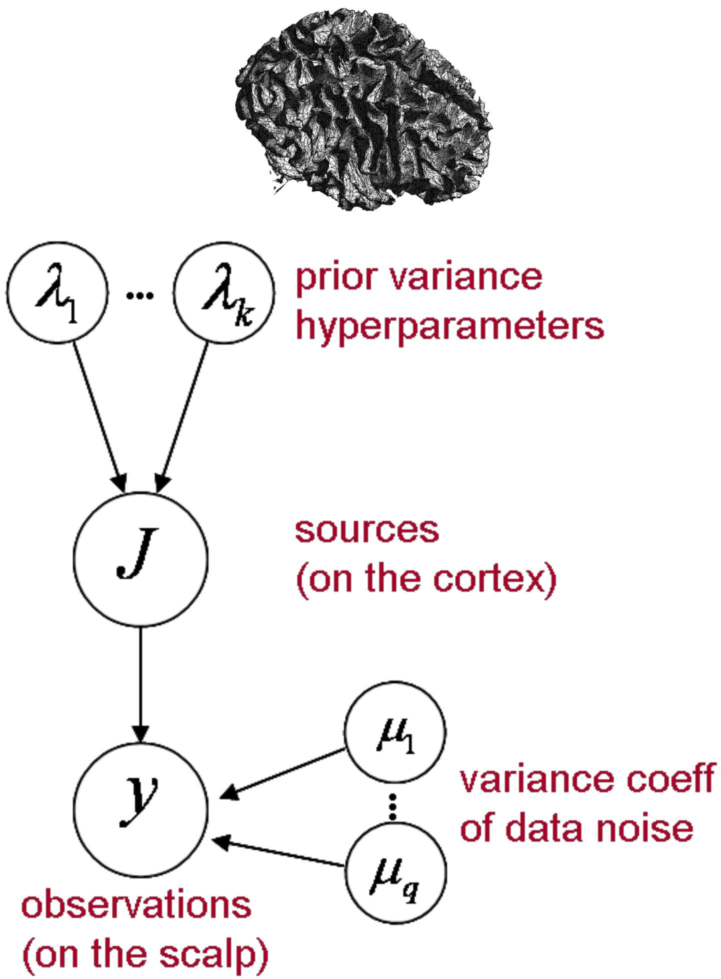
minimum norm — smoothness prior — functional prior

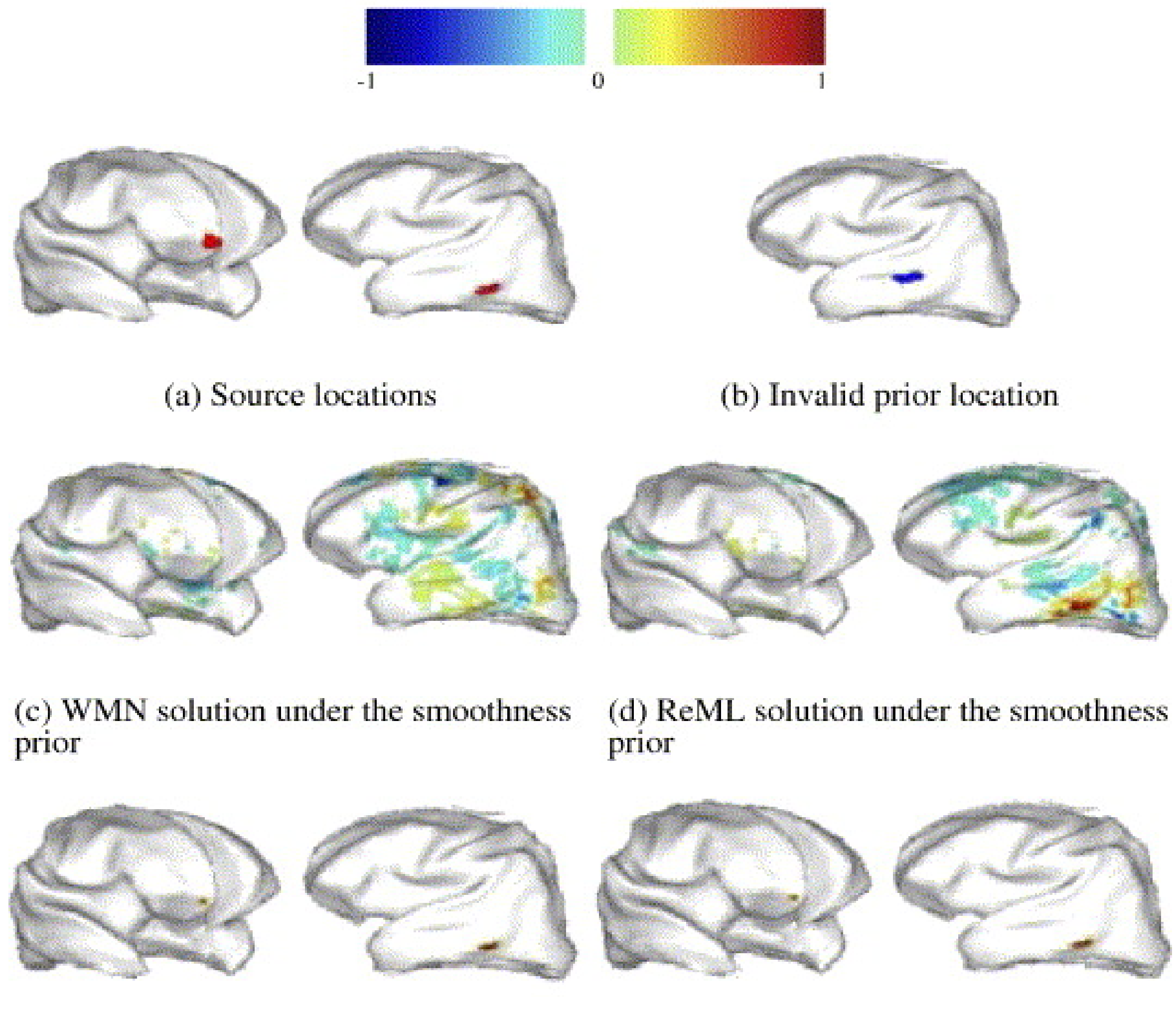




[Mattout et al., Neuroimage, 2006]

MEG/EEG source reconstruction





(a) Source locations

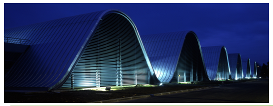
(b) Invalid prior location

(c) WMN solution under the smoothness prior

(d) ReML solution under the smoothness prior

(e) ReML solution under the smoothness and valid priors

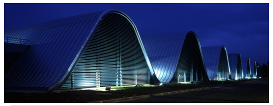
(f) ReML solution under the smoothness, valid and invalid priors



[Mattout et al., Neuroimage, 2006]

Outline

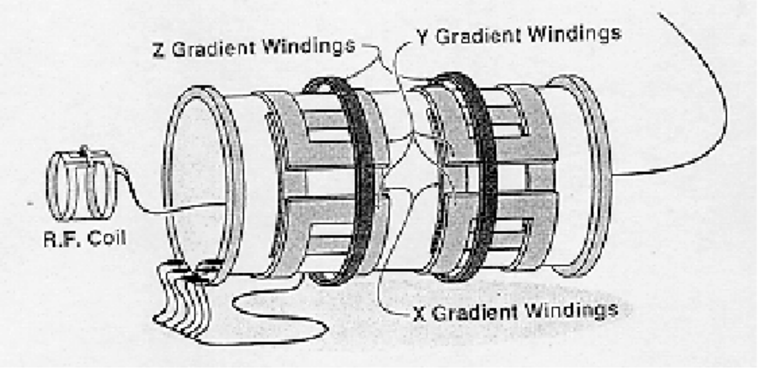
- Introduction to functional brain imaging
  - Magneto/Electro-encephalography
  - functional MRI
- Mapping brain activity
  - GLM framework (Statistical Parametric Mapping)
  - Spatial regularization [VB optimization]
- Probe brain dynamics in fMRI
  - FIR modeling
  - Temporal regularization [EM/SAEM algorithms]
- Joint detection-estimation
  - Unify both questions
  - Spatio-temporal regularization [MCMC methods]



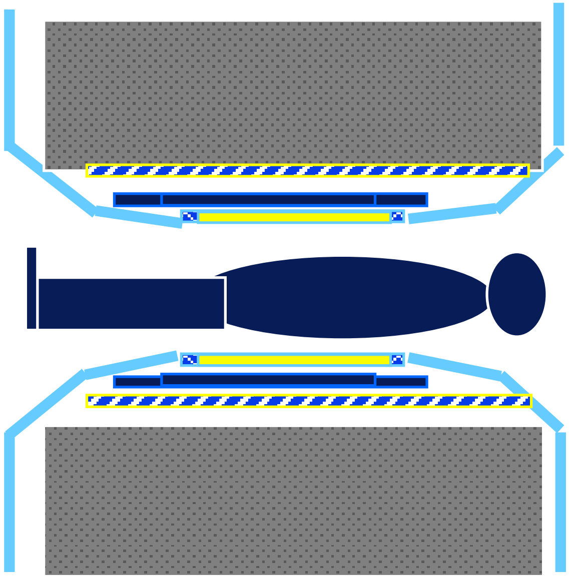
Magnetic Resonance Imaging

3T

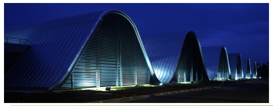
⇨ High magnetic field



⇨ Auxiliary coils: the « gradients »



⇨ Émetteur/récepteur RF

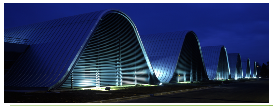


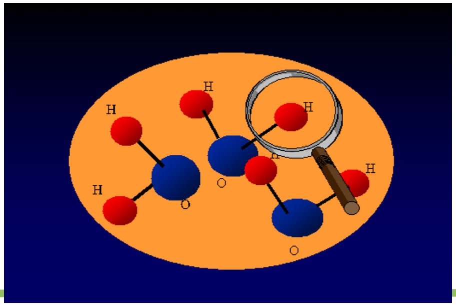
Atomes en RMN

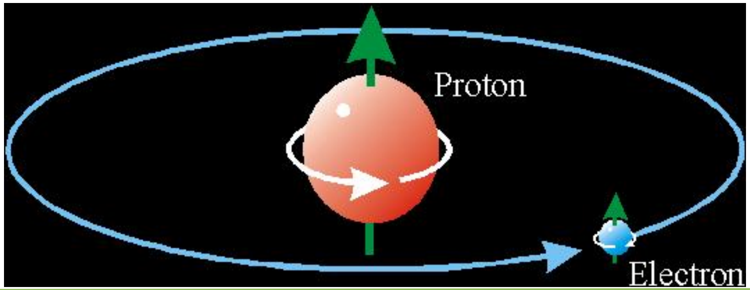
Éléments	Abondance biologique
Hydrogen (H)	0.63
Sodium (Na)	0.00041
Phosphorus (P)	0.0024
Carbon (C)	0.094
Oxygen (O)	0.26
Calcium (Ca)	0.0022
Nitrogen (N)	0.015

Eléments Utilisé en RMN	Symbole	Abondance dans le corps humain
Hydrogen	^{1}H	99.985
	^{2}H	0.015
Carbon	^{13}C	1.11
Nitrogen	^{14}N	99.63
	^{15}N	0.37
Sodium	^{23}Na	100
Phosphorus	^{31}P	100
Potassium	^{39}K	93.1
Calcium	^{43}Ca	0.145

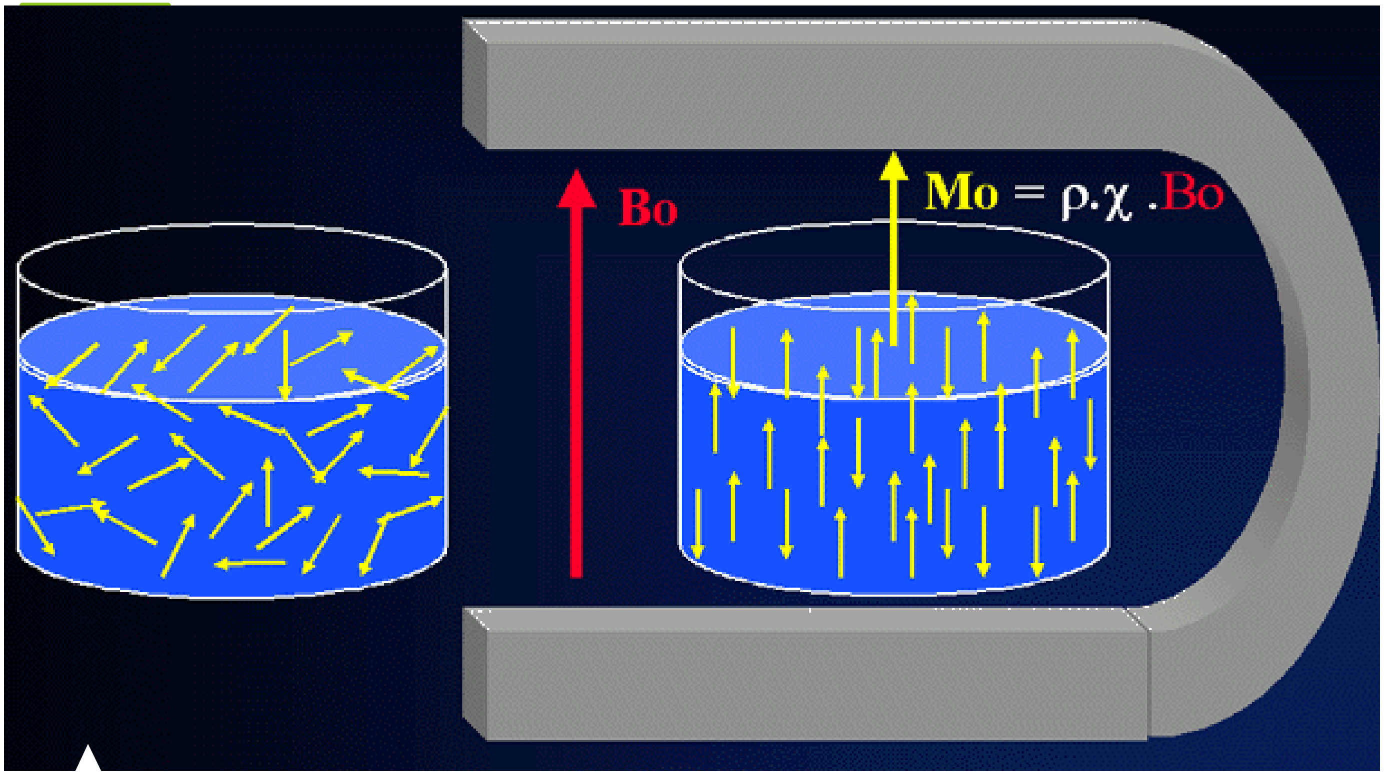
Choix des **protons des atomes d'hydrogène** de l'eau pour l'imagerie



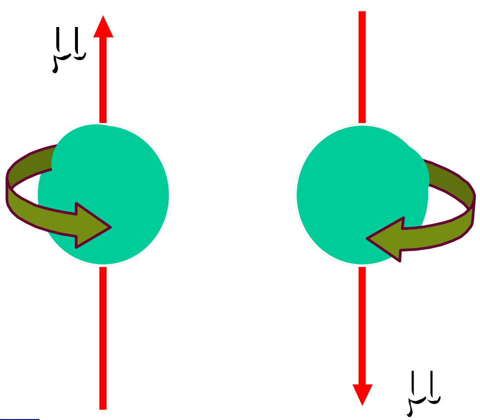




Choix de l'aimant statique



Alignement des moments magnétiques suivant 2 directions:

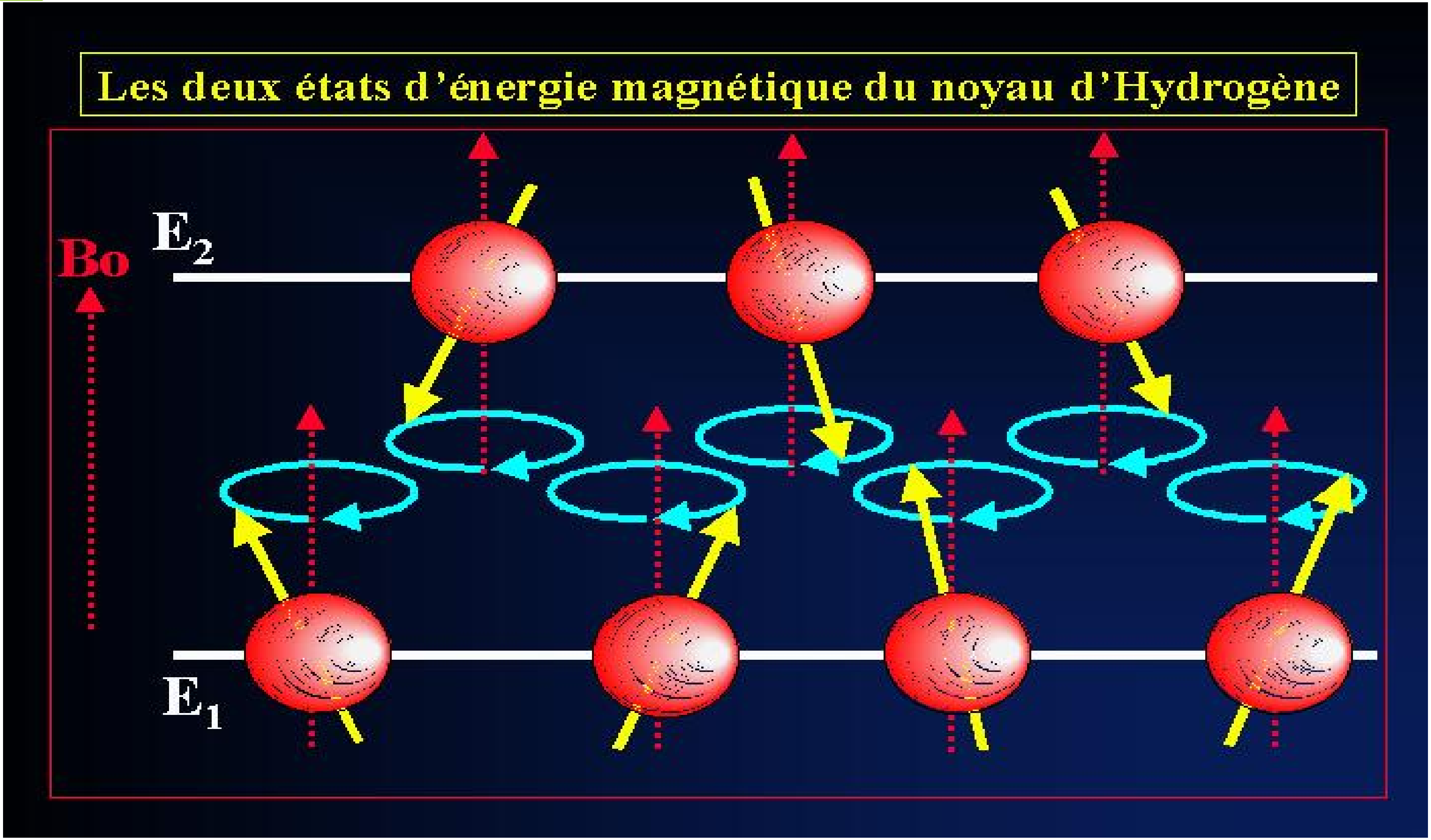


- direction parallèle: orientation dans la direction de B0
- direction antiparallèle: orientation dans la direction opposée à B0

Mouvement de précession $\omega = \gamma B$

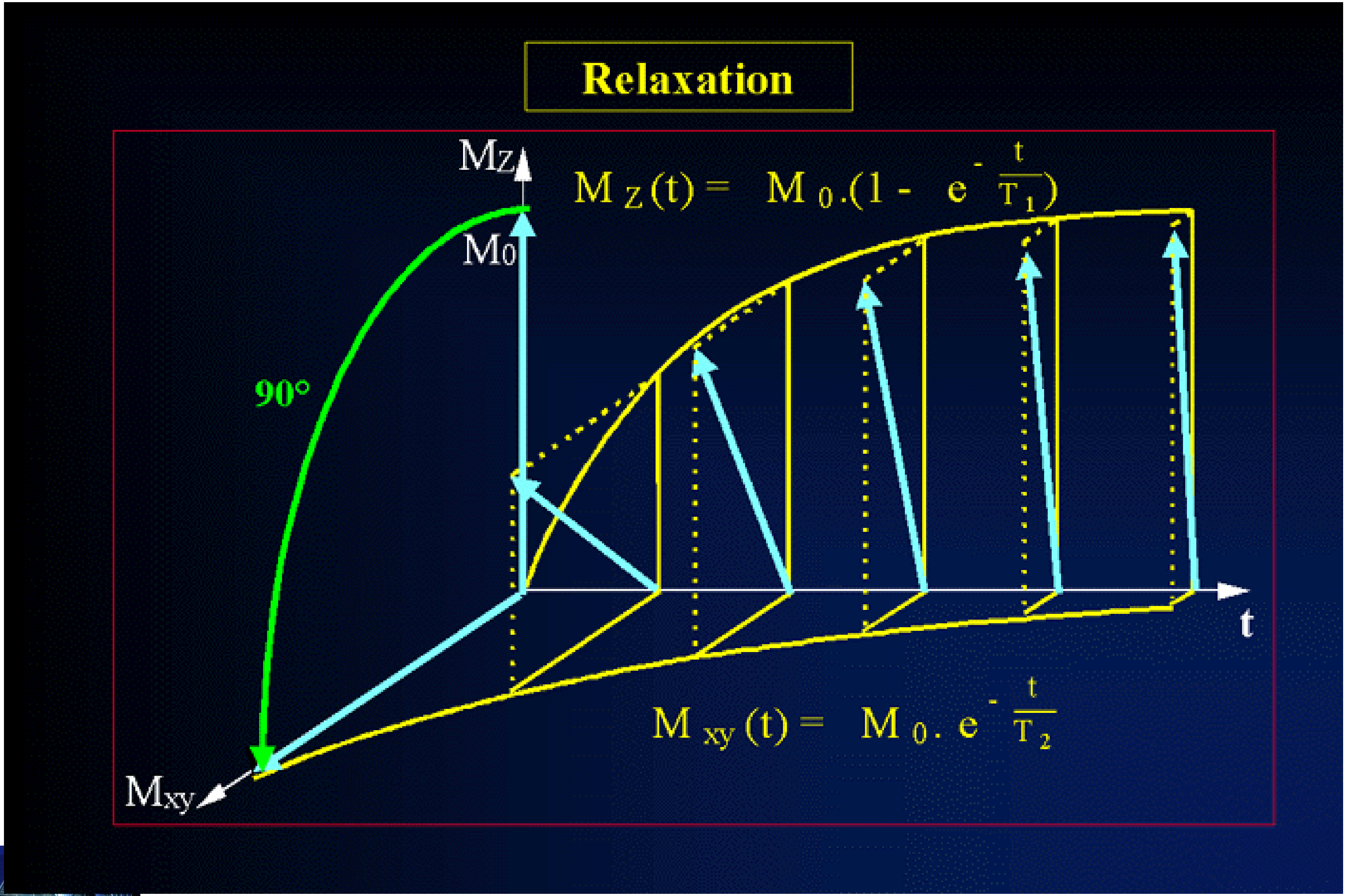
Pour les protons 42.58 MHz/T

Spins des protons

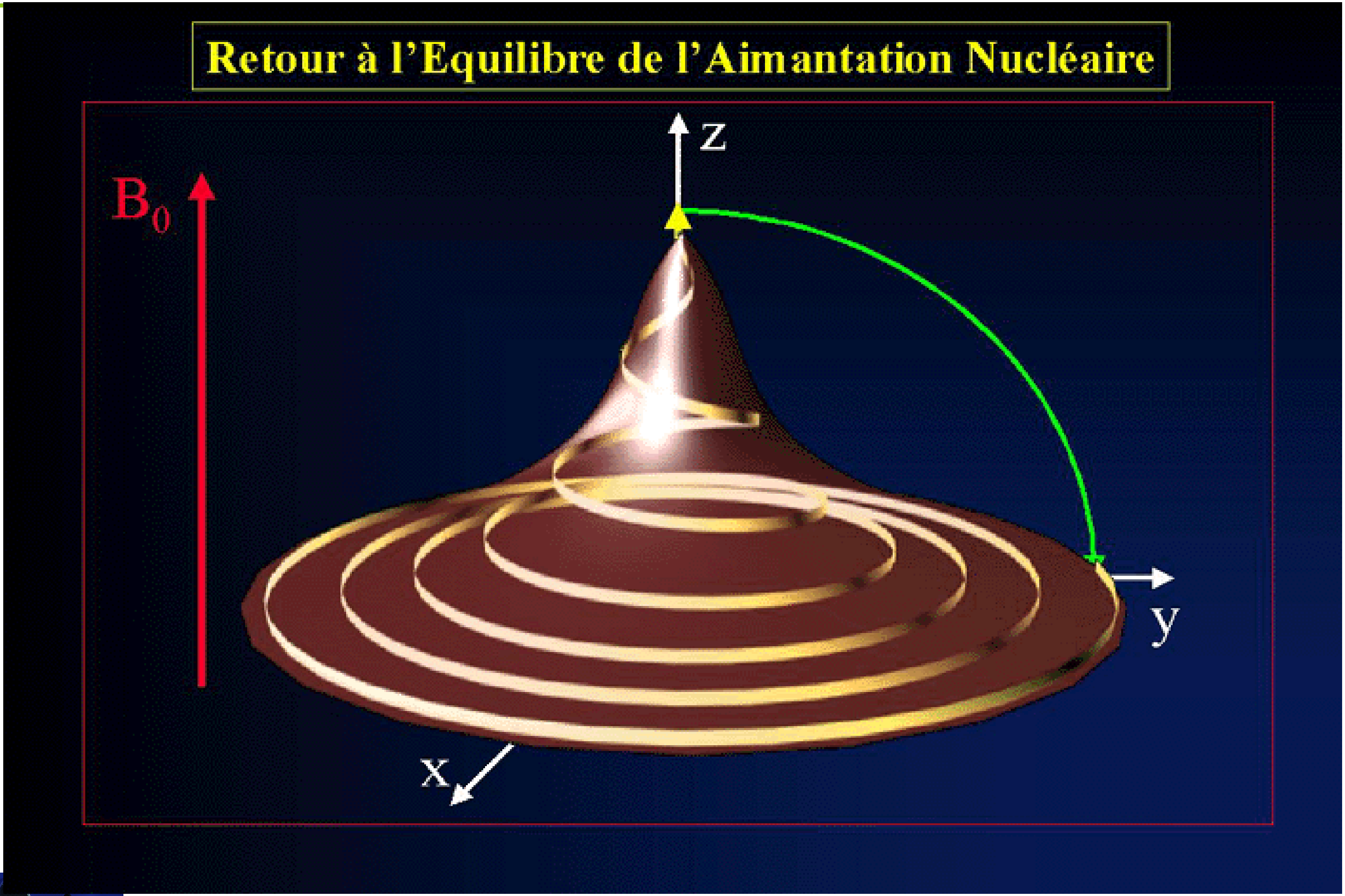


en présence d'un champ magnétique B0

Phénomène de RMN



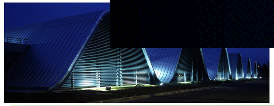
Phénomène de RMN



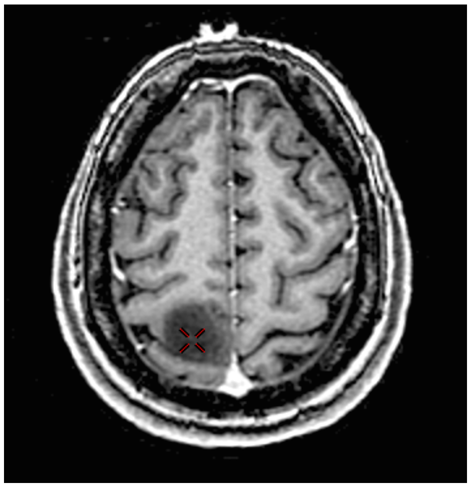
Temps de relaxation

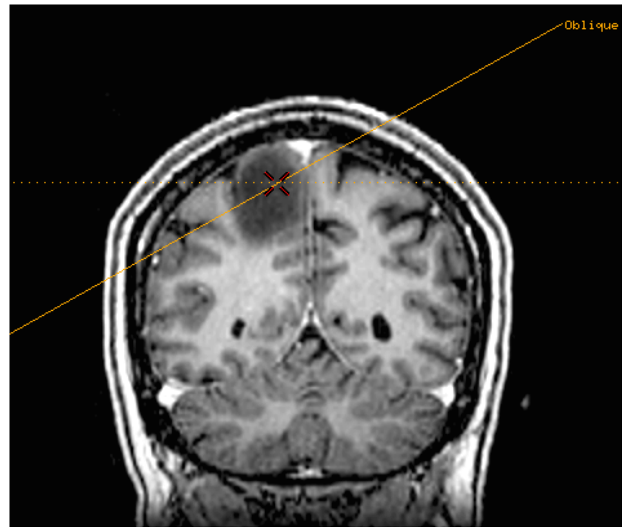
Ordres de Grandeurs des Temps de Relaxation à 1,5 T

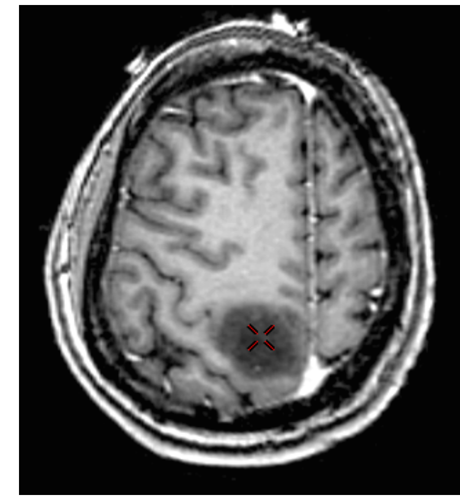
Temps de Relaxation / Tissus Humains	T1	T2
Liquide Céphalo-rachidien	2500 ms	2000 ms
Substance Grise	900 ms	90 ms
Substance Blanche	750 ms	80 ms
Graisse	300 ms	40 ms

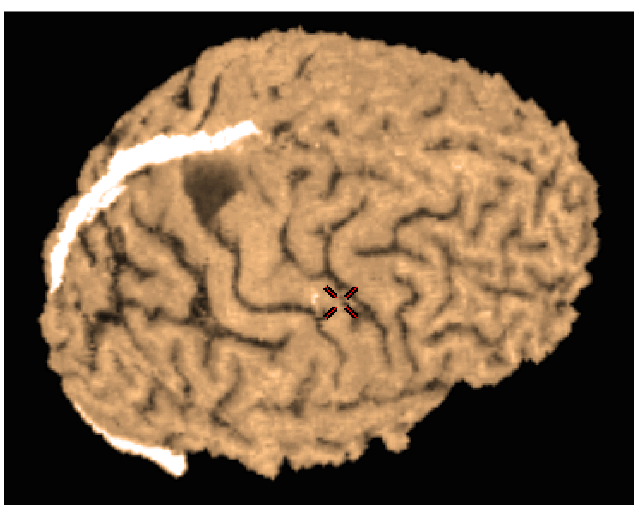


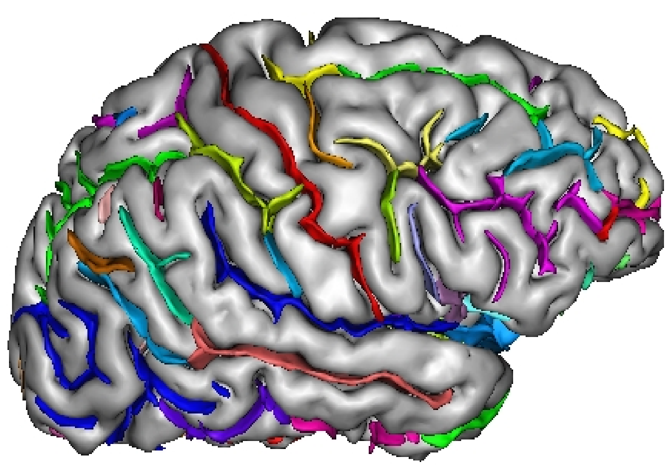
Anatomical MRI









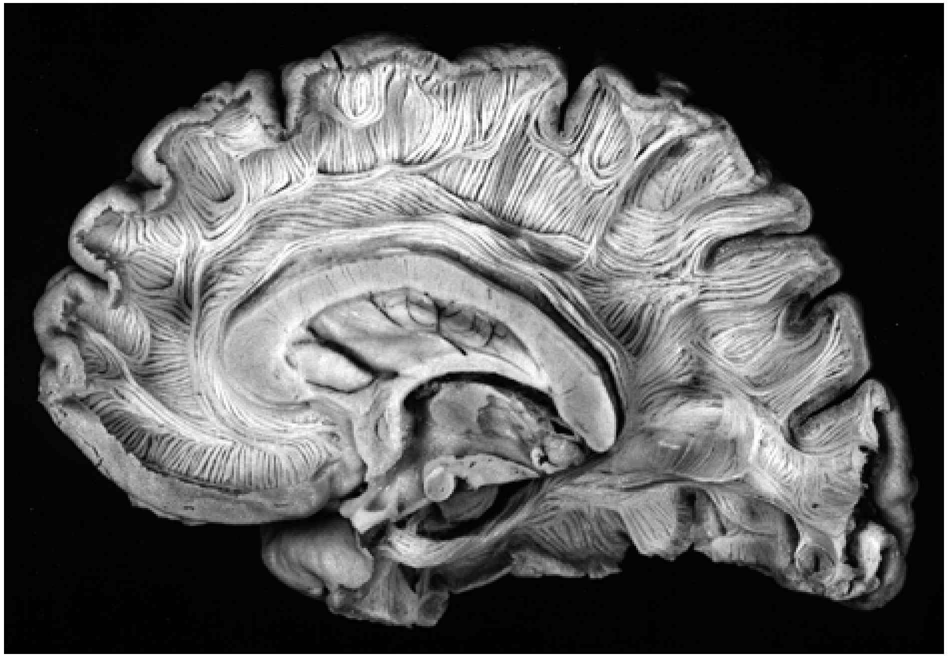


Brain tumor

Post-op checking

Sulci recognition

diffusion MRI

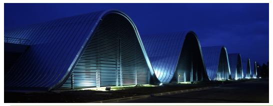


[Virtual Hospital, 1998]

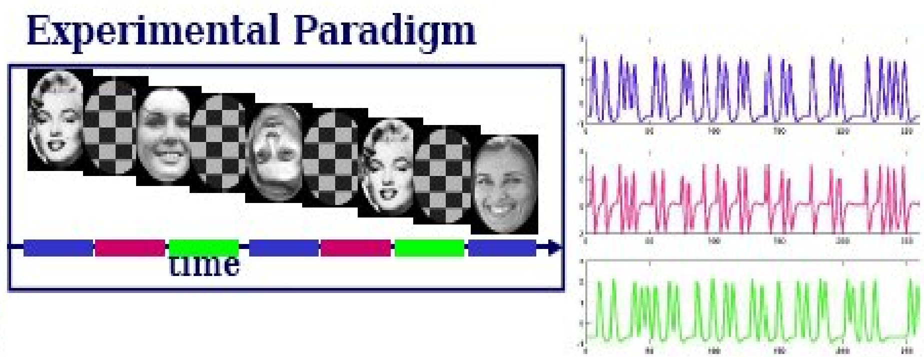
[Elkouby, 2005]

- ✘ in *vivo* imaging of the water brownian motion
- ✘ Get access to the structural connectivity

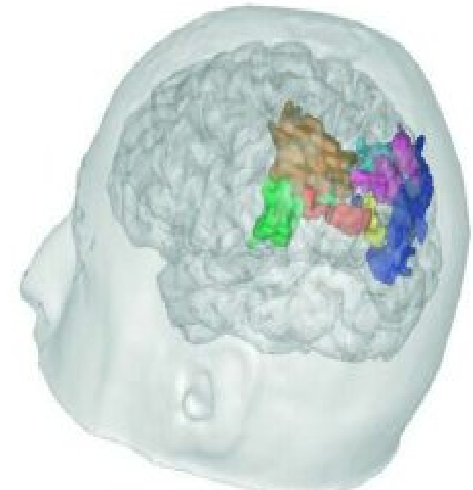
Mouvement brownien [Einstein, 1905] → Spectro diffusion [Stejskal-Tanner, 1965] → IRM de diffusion [Le Bihan, 1986]

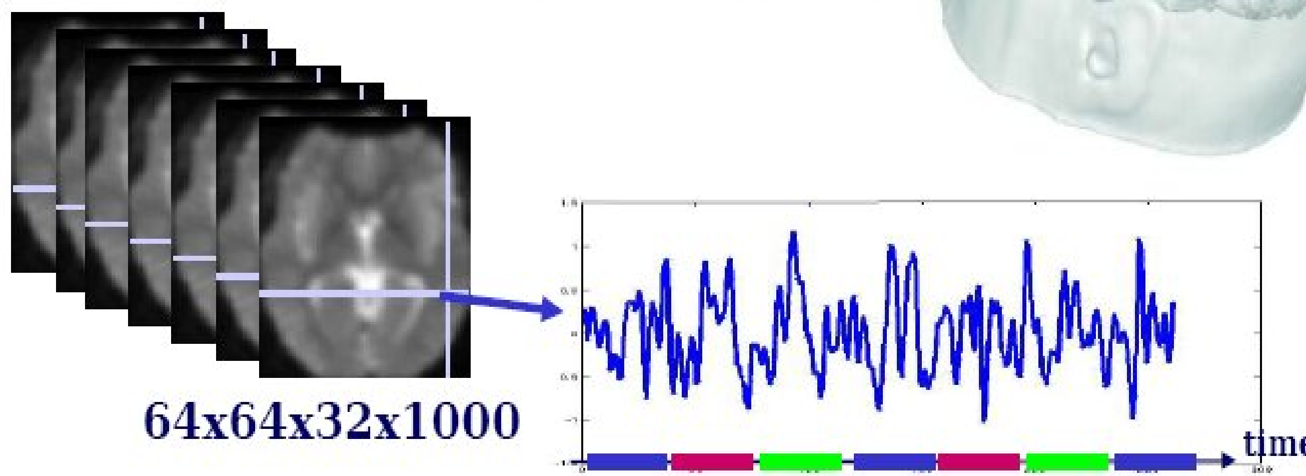


Functional MRI



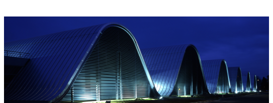
Neuronal activation





Measured 4D signal :

BOLD = Blood Oxygenation Level Dependent signal

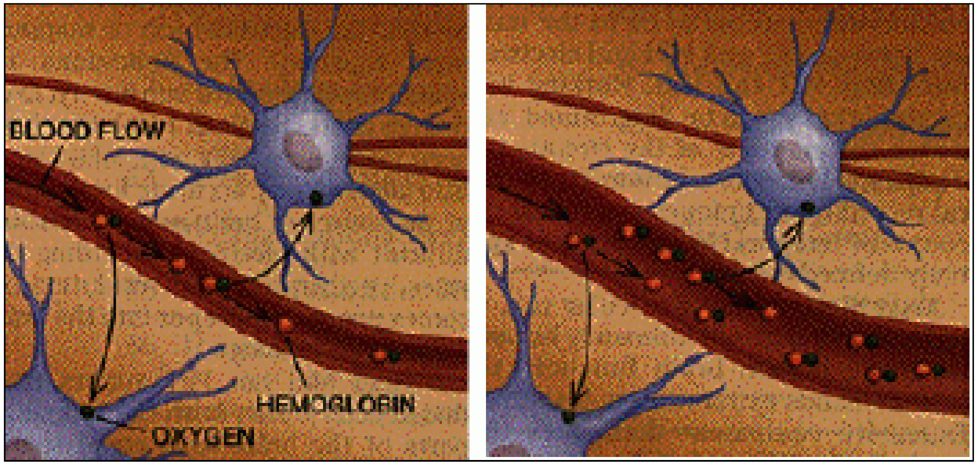


Le signal BOLD

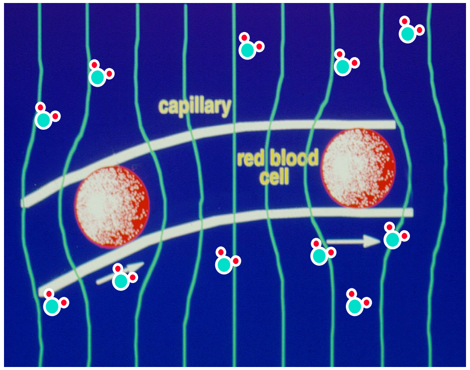
[Ogawa et al, 1990,1992]

Produit de contraste intrinsèque :
Oxyhémoglobine (HbO2) : diamagnétique
Désoxyhémoglobine (Hb) : paramagnétique

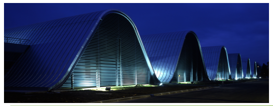
Détectable en IRMf



Activation cérébrale → Légère augmentation de la consommation O2, accompagnée d'un fort afflux de sang oxygéné

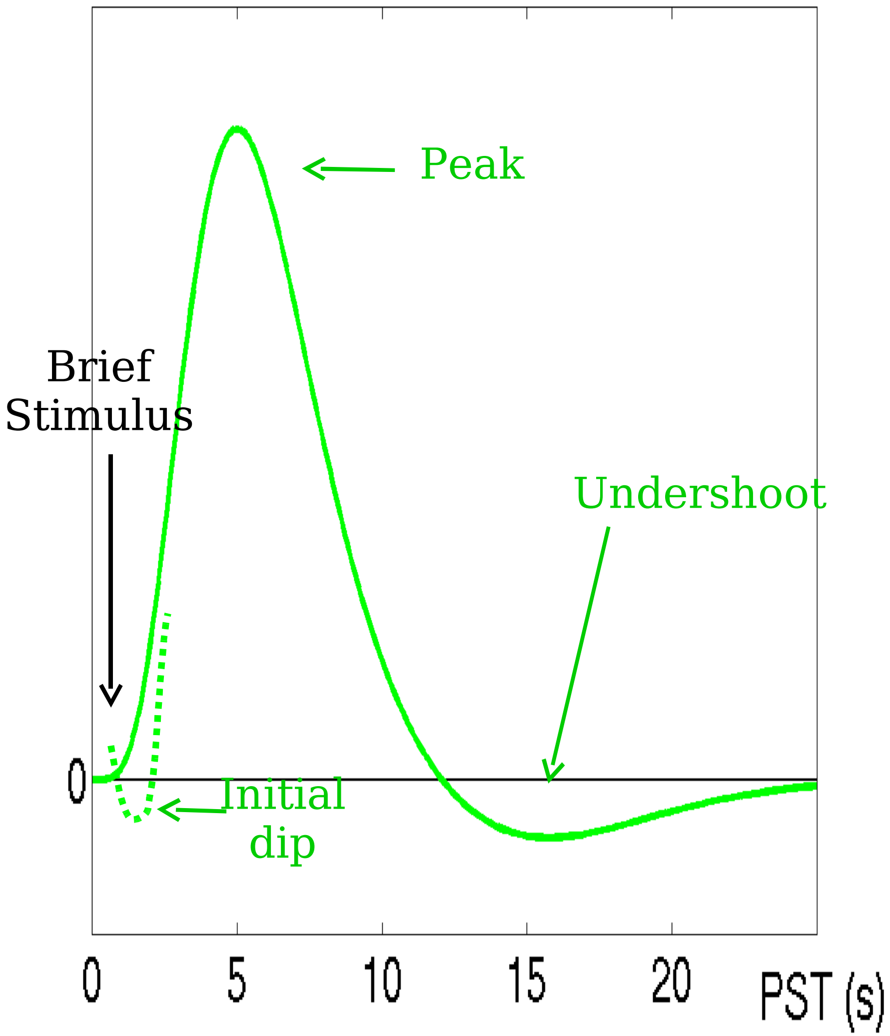


Conséquence : augmentation de la concentration en sang oxygéné (Hb02) des vaisseaux proches des neurones actifs



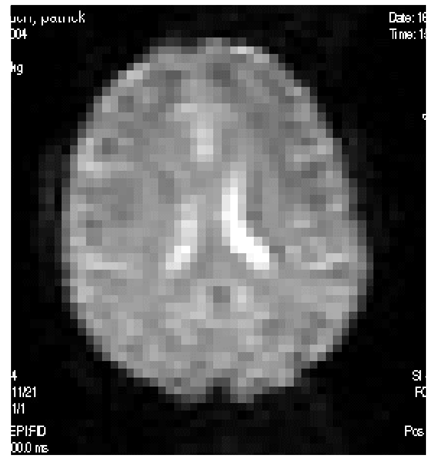
Hemodynamic Response Function

- Function of blood oxygenation, flow, volume **[Buxton et al. 98]**
- Peak (max. oxygenation) 4-6s poststimulus;
- Baseline after 20-30s
- Initial undershoot can be observed **[Malonek &Grinvald. 96]**
- … but differences across: other regions **[Schacter et al. 97]**
- individuals **[Aguirre et al. 98]**

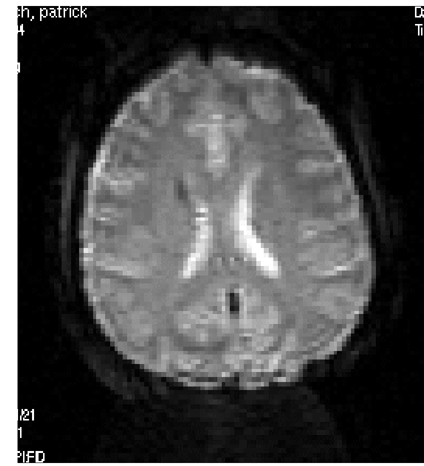


The truth about fMRI data

64x64 Pixels ~ 3 x 3 mm



128x128 Pixels ~ 1.5 x 1.5mm

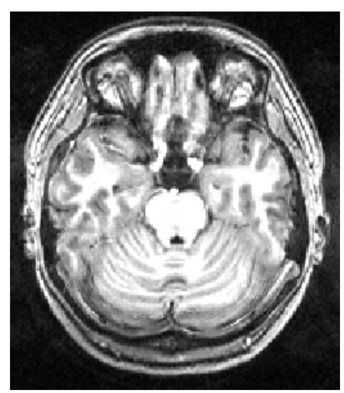


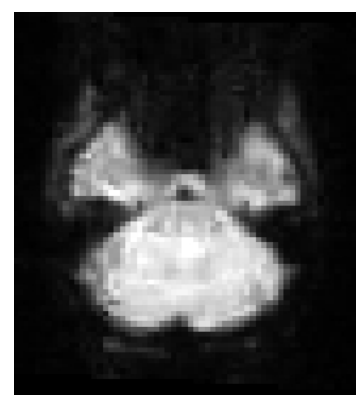
I. They are distorted

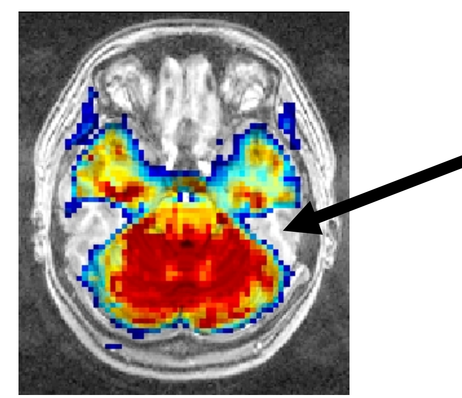
II. They are noisy

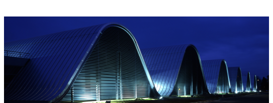
III. They don't have signal everywhere in the brain

IV. They depend on many parameters: T2*, B0, TE, …



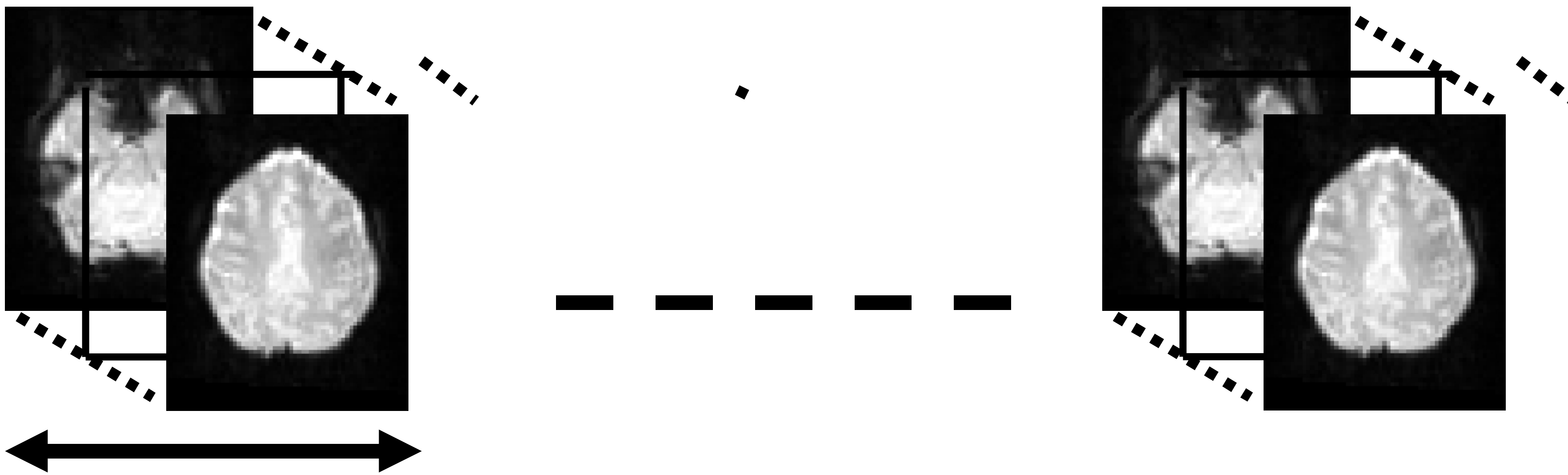






The truth about fMRI data

V. They are big … and are getting bigger



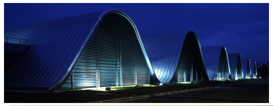
1 brain volume (64x64x30)
in 3 sec (TR =1)

scan =200

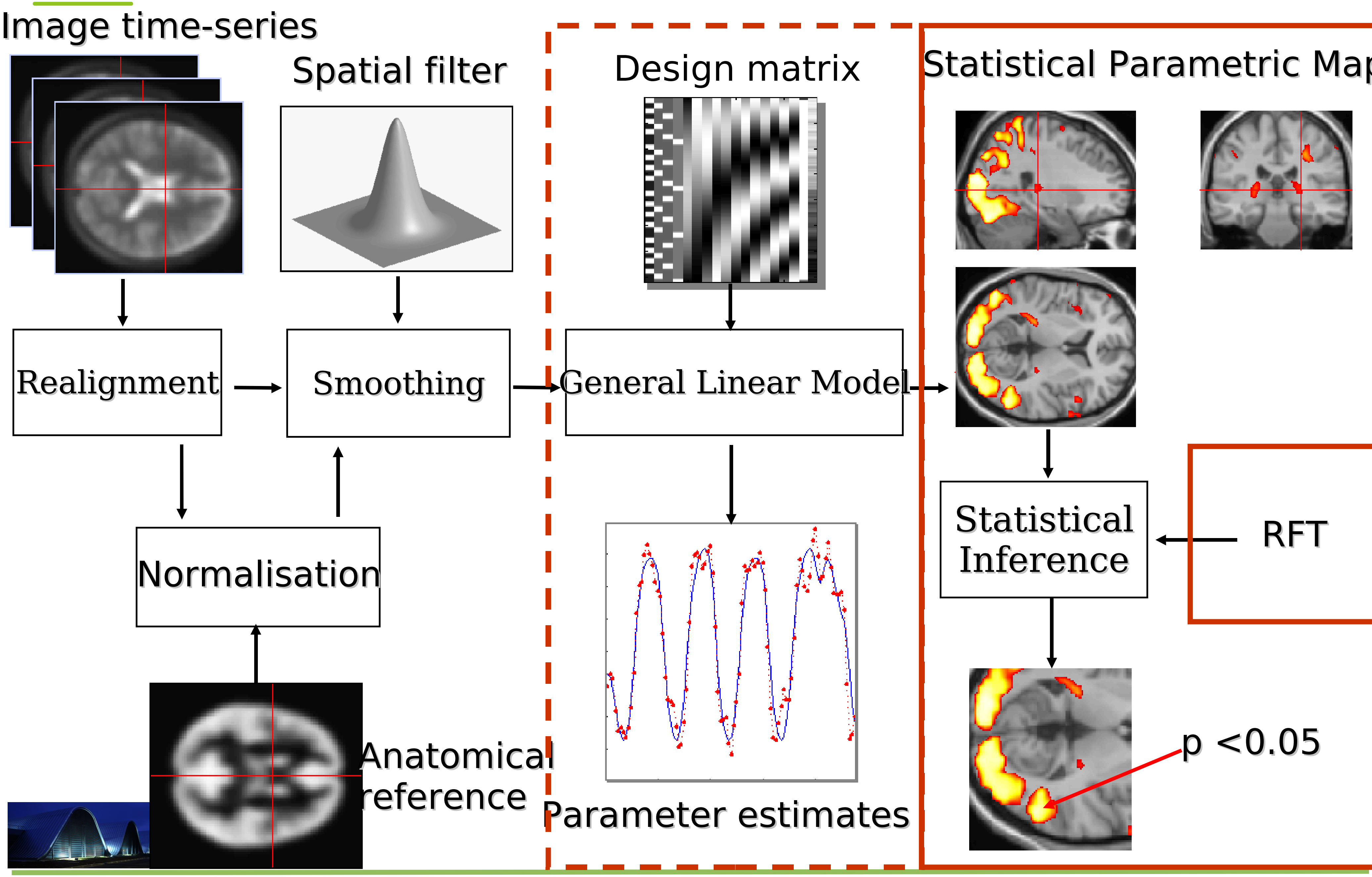
This is ONE run; often 3-8 runs X 15 subjects
6D Data (~20 Go)

Outline

- Introduction to functional brain imaging
  - Magneto/Electro-encephalography
  - functional MRI
- Mapping brain activity
  - GLM framework (Statistical Parametric Mapping)
  - Spatial regularization [VB optimization]
- Probe brain dynamics in fMRI
  - FIR modeling
  - Temporal regularization [EM/SAEM algorithms]
- Joint detection-estimation
  - Unify both questions
  - MCMC methods



fMRI processing pipeline



fMRI example

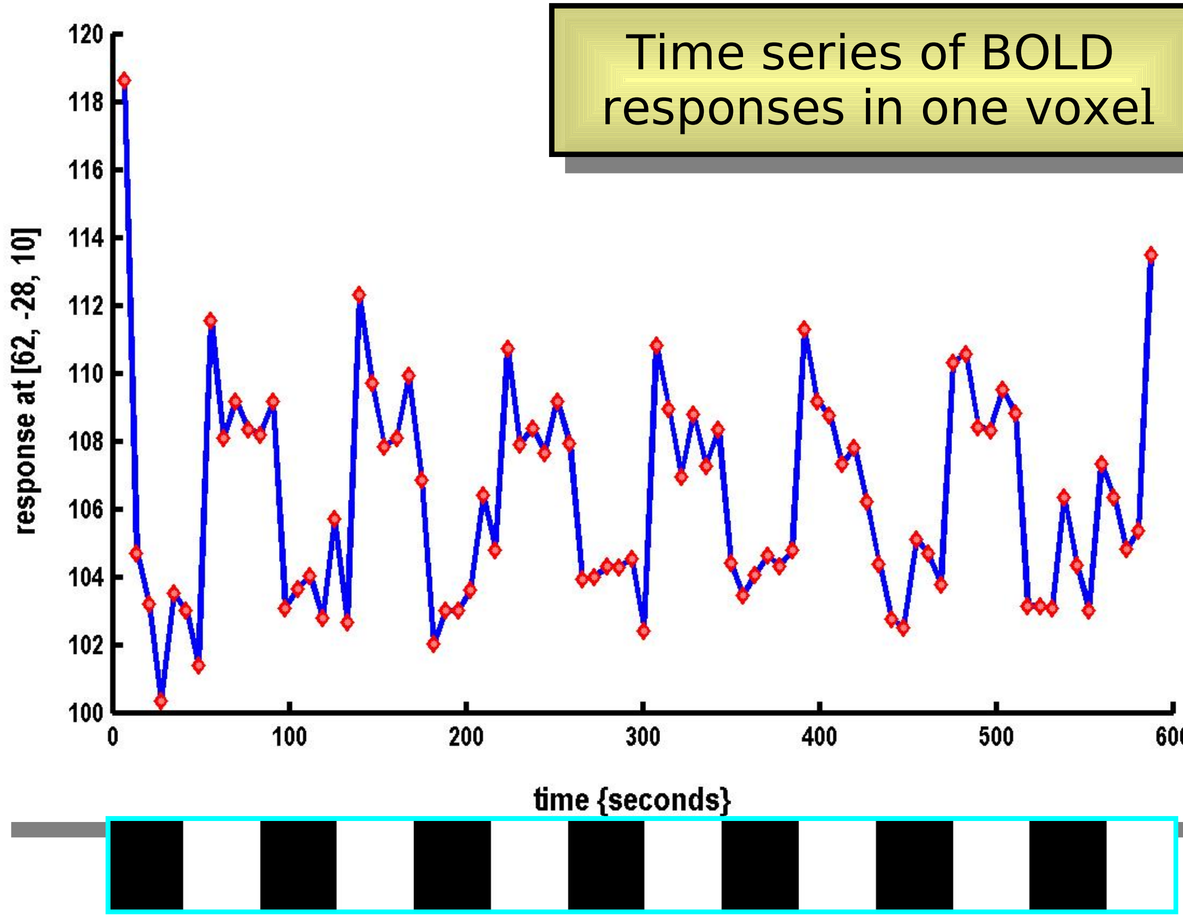
One session

Passive word listening versus rest

7 cycles of rest and listening

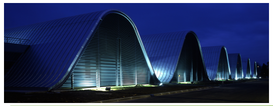
Each epoch 6 scans with 7 sec TR

Time series of BOLD responses in one voxel



Stimulus function

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

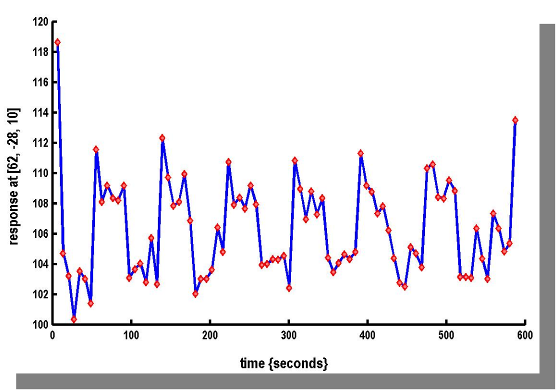


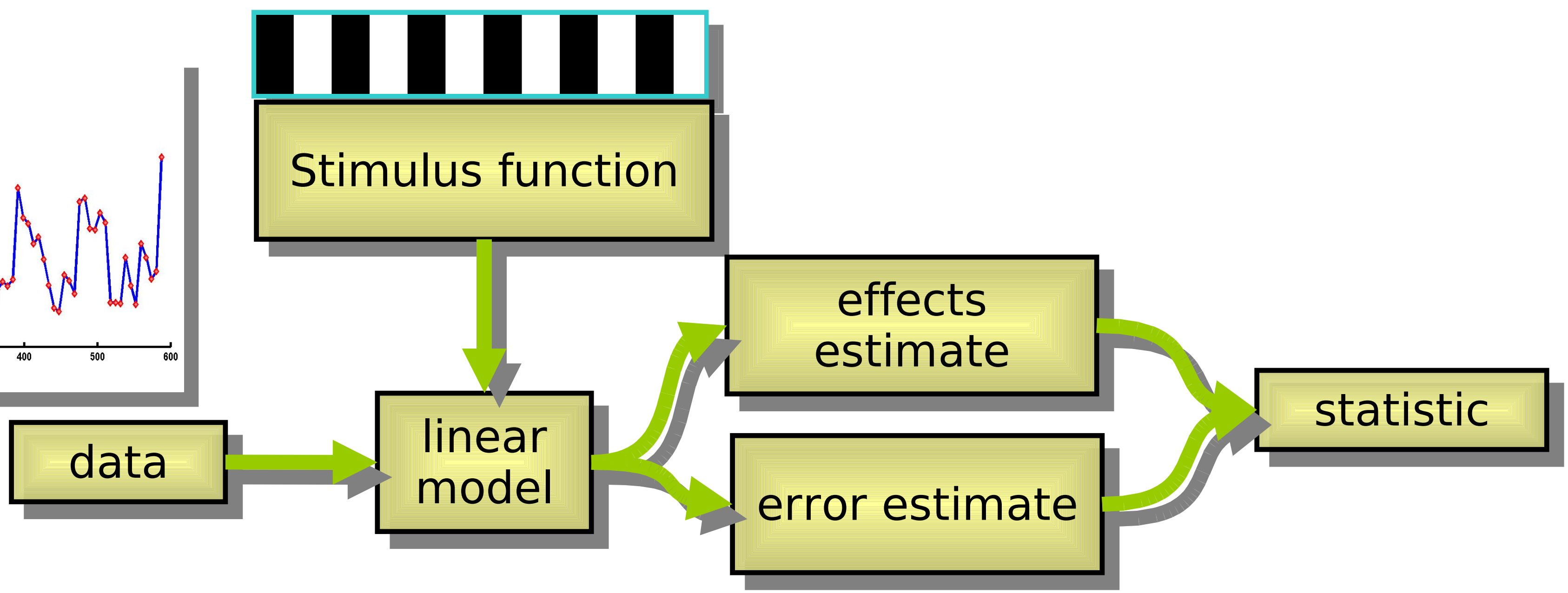
Why modelling?

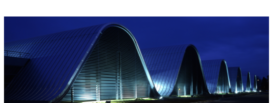
Why? Make inferences about effects of interest

How?

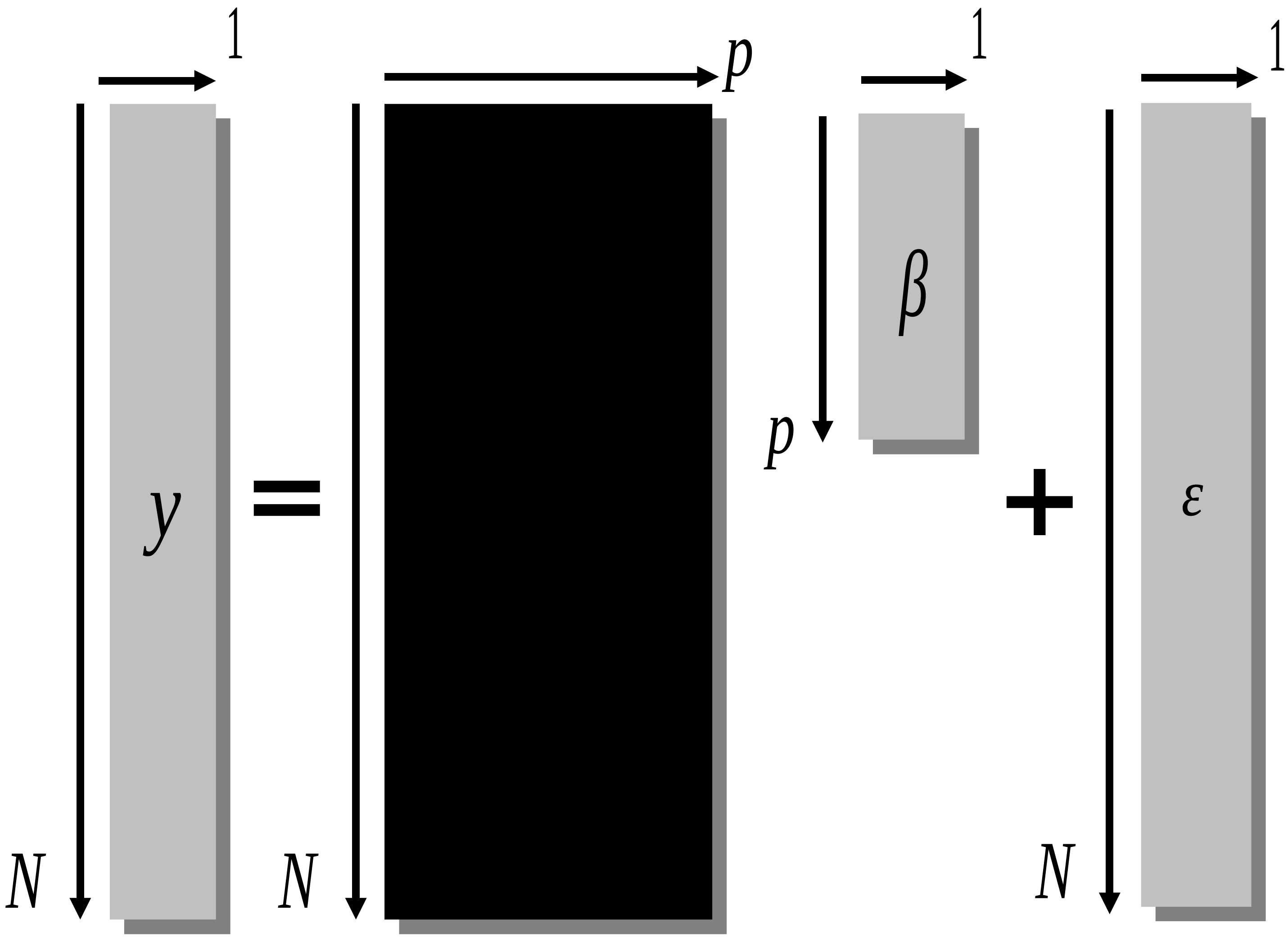
1. Decompose data into effects and error
2. Form statistic using estimates of effects and error







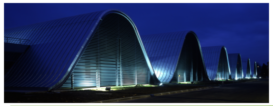
General Linear Model



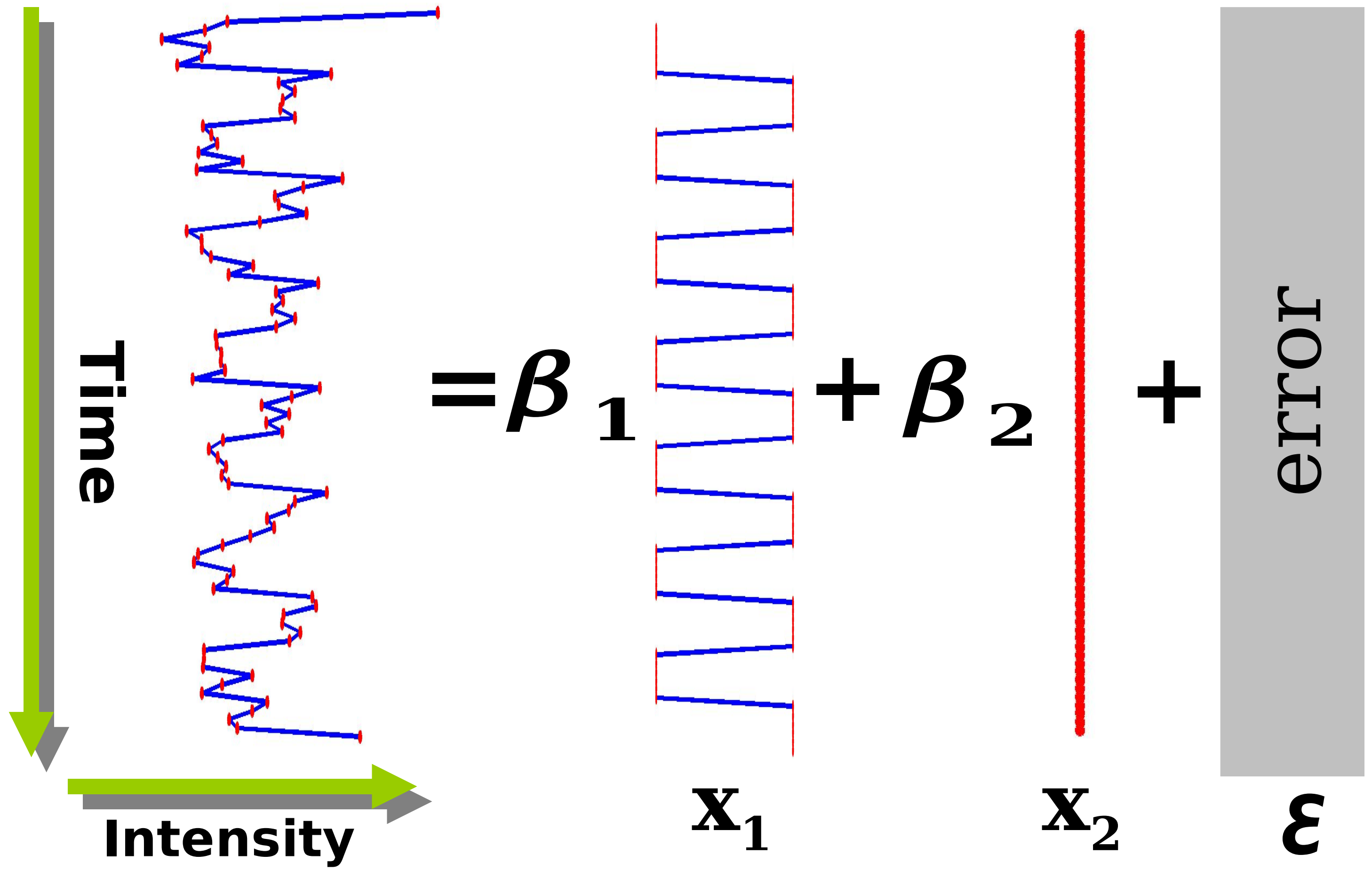
$$y = X\beta + \varepsilon$$

N: number of scans
p: number of regressors

Model is specified by
- Design matrix X
- Assumptions about ε



General Linear Model

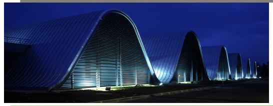


$$\varepsilon \sim N(0, \sigma^2 I)$$

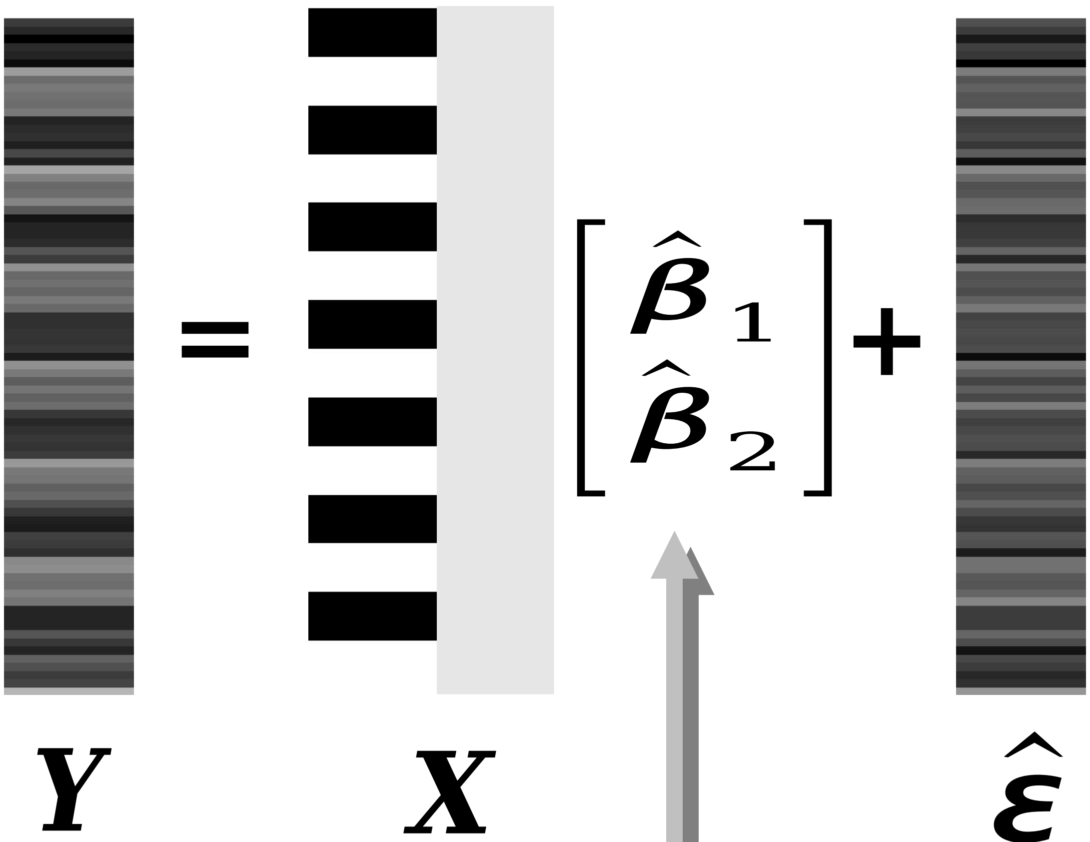
(error is normal and iid)

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

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



Parameter estimation



$$\begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix}$$

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

Assume iid error

$$\hat{\boldsymbol{\varepsilon}} = \boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}$$

residuals

$$\hat{\boldsymbol{\beta}} = \left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1}\boldsymbol{X}^T\boldsymbol{Y}$$

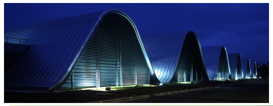
Estimate parameters

such that

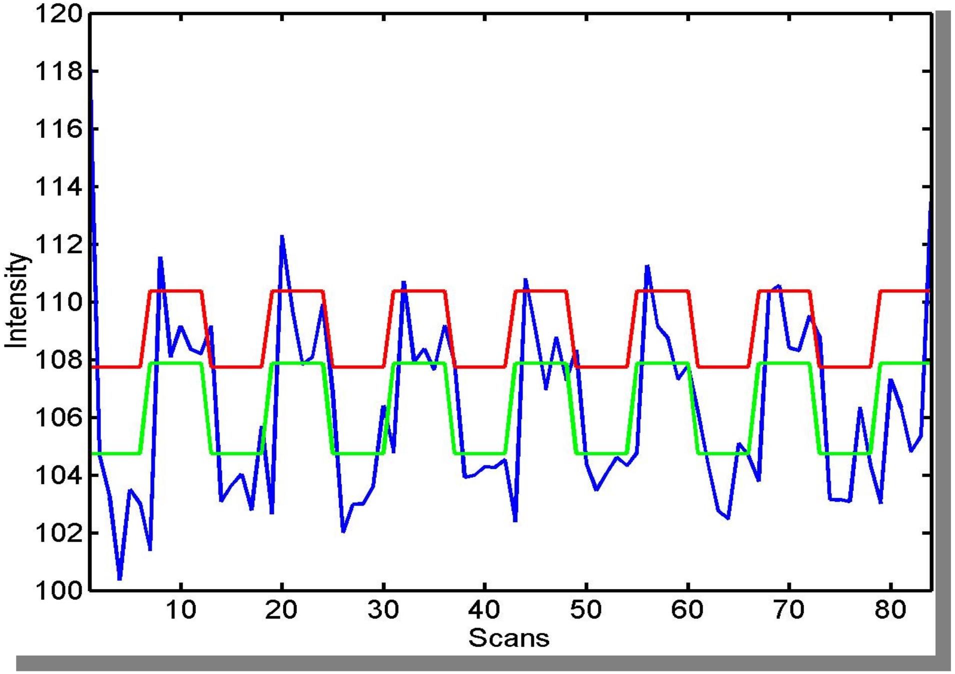
$$\sum_{t=1}^{N} \hat{\varepsilon}_t^2$$

minimal

Least squares parameter estimate



Estimation, example

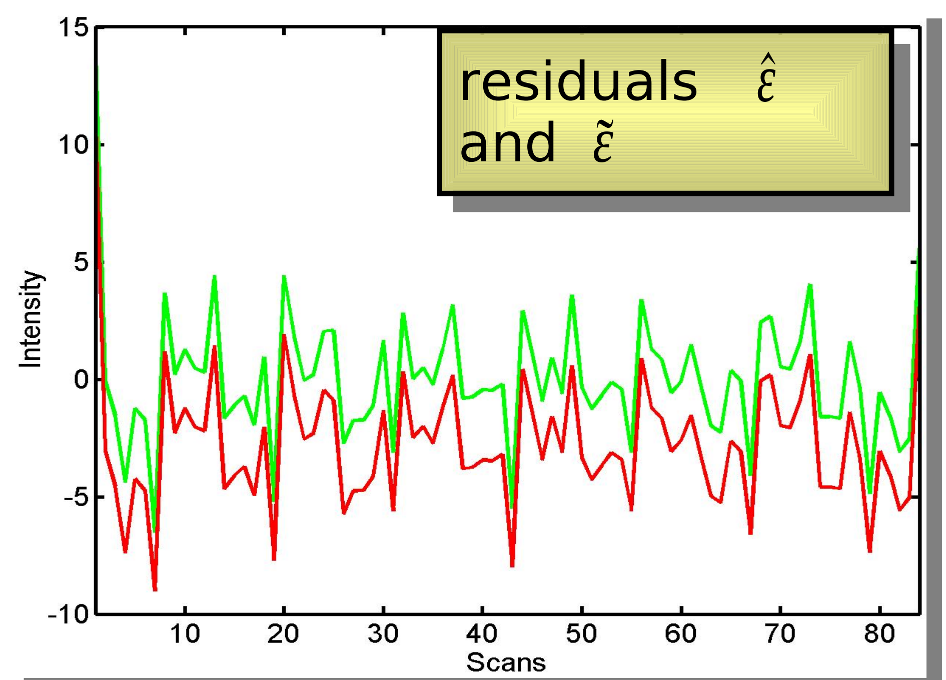


$$Y = X\beta + \varepsilon$$

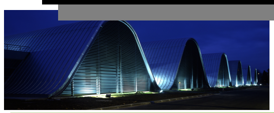
Assume iid error

$$\hat{\beta} = \left(X^T X\right)^{-1} X^T Y$$

Least squares estimate	$\hat{\beta} = \begin{bmatrix} 3.15 \\ 104.74 \end{bmatrix}$	
Another estimate	$\tilde{\beta} = \begin{bmatrix} 2.65 \\ 107.74 \end{bmatrix}$	

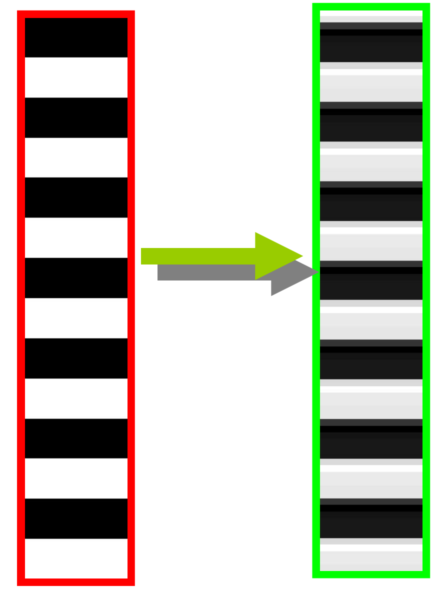


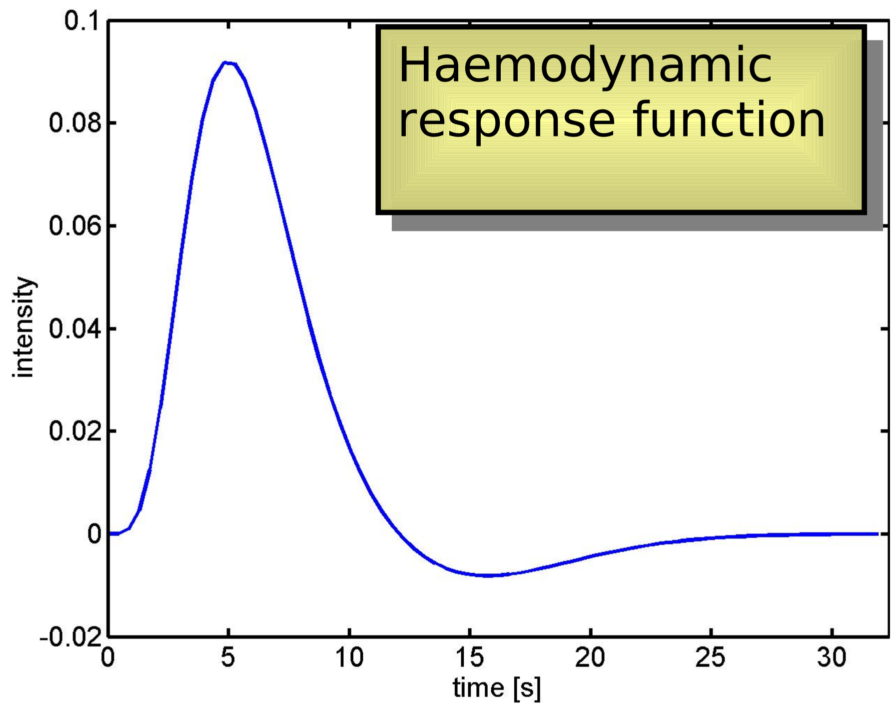
residuals $\hat{\varepsilon}$ and $\tilde{\varepsilon}$

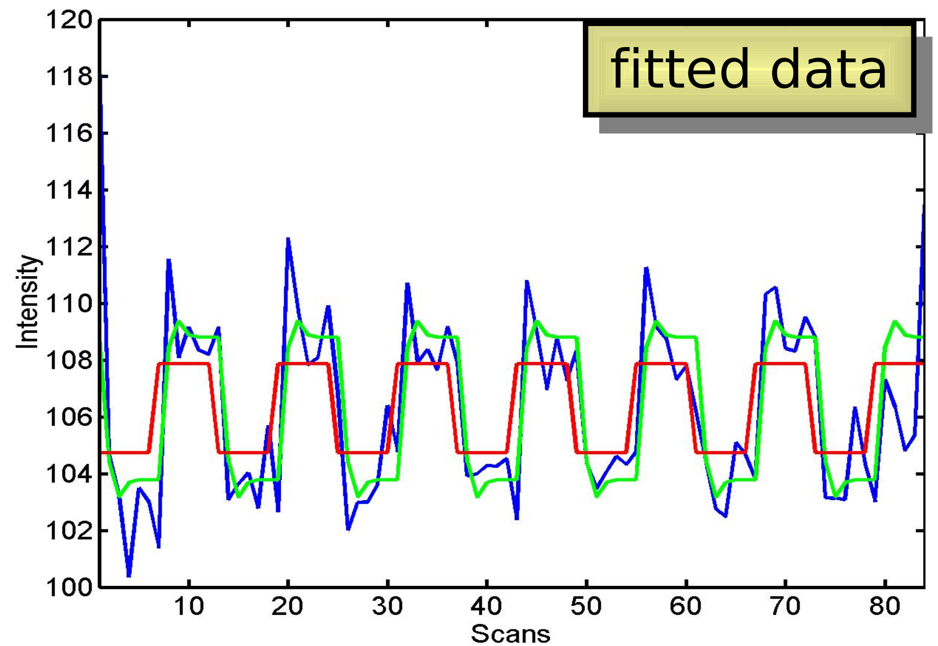


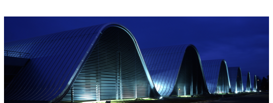
Improved model

Convolve stimulus function with model of BOLD response



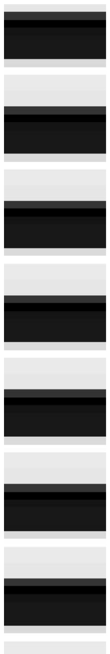


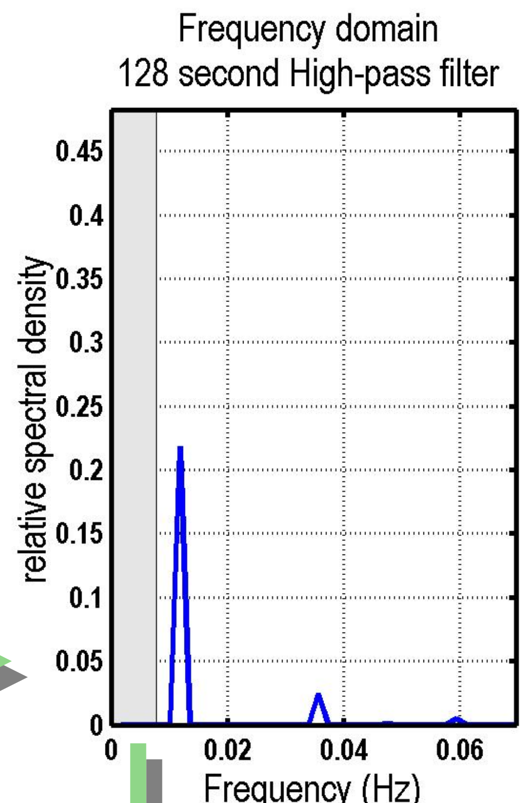




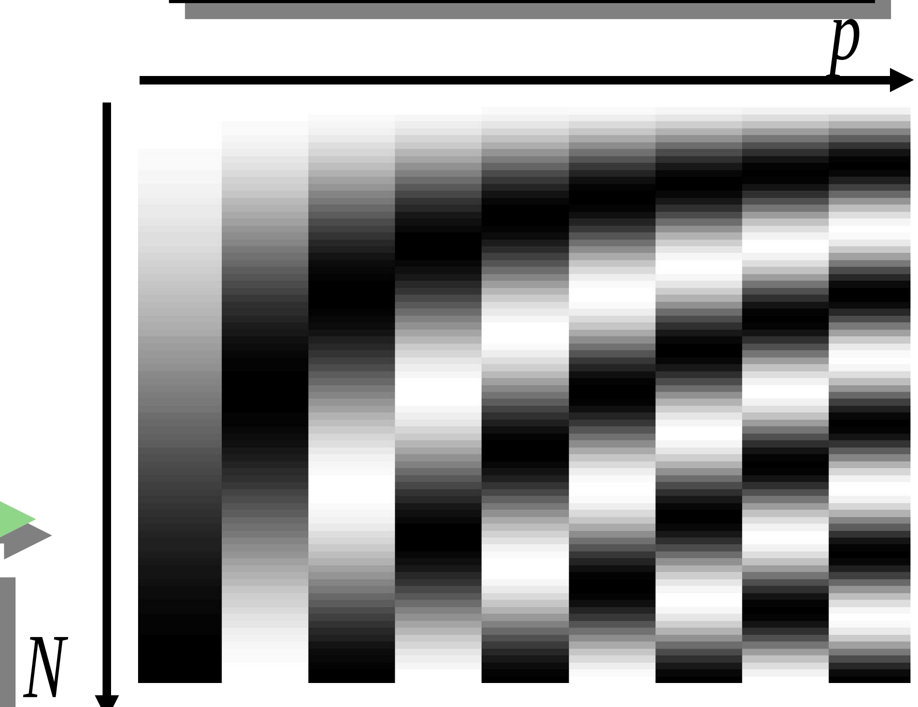
High-pass filtering

$$Y = X\beta + \varepsilon$$

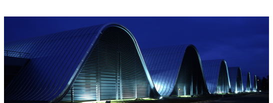




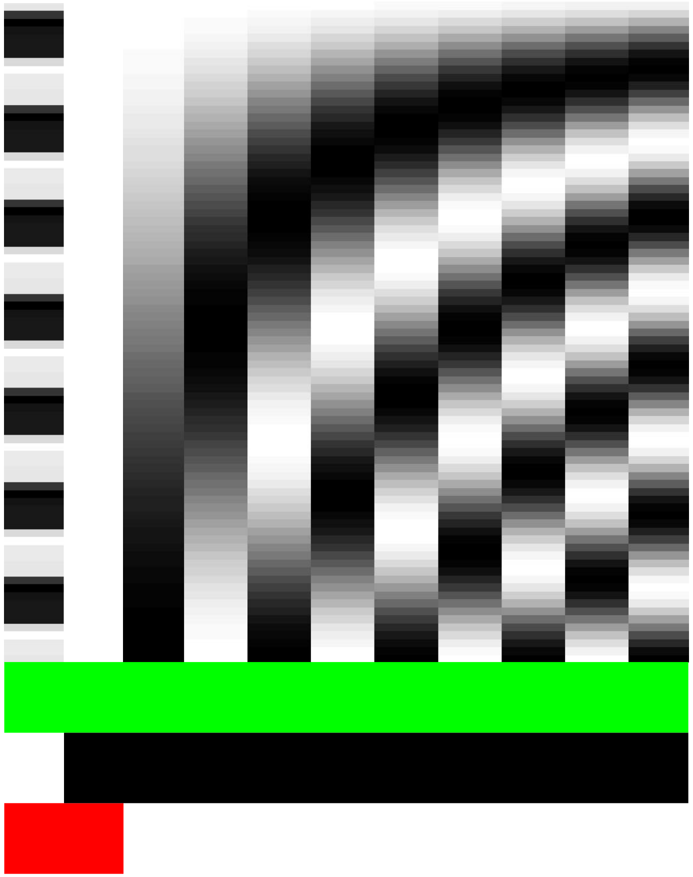
High-pass filter implemented by modelling low frequencies



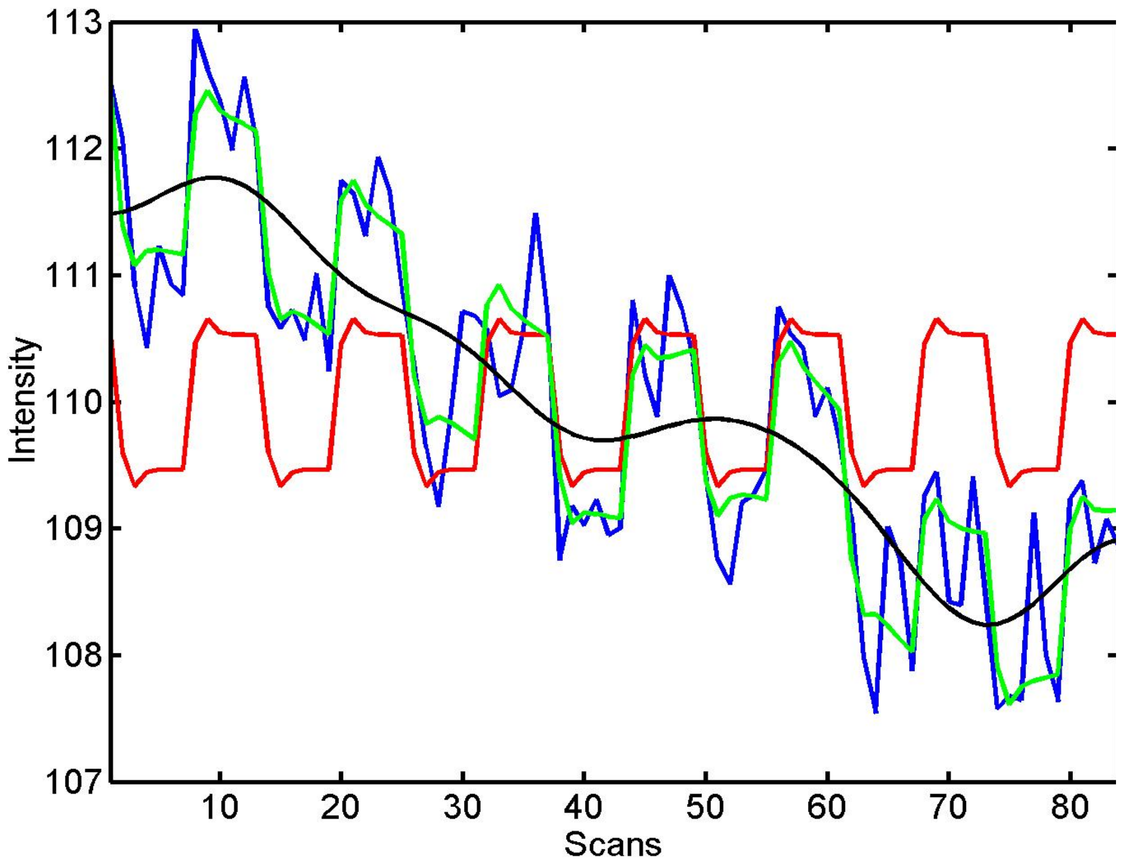
discrete cosine transform set



High-pass filtering



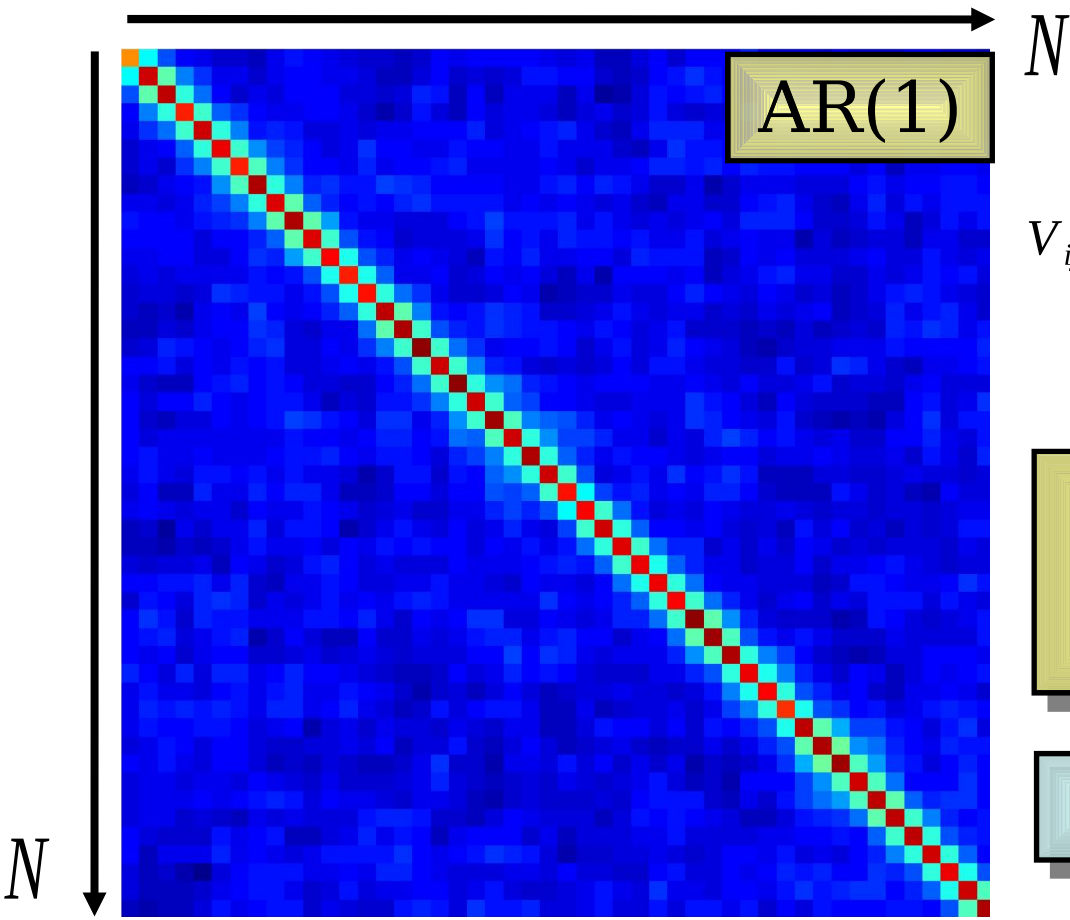
$$Y = X\beta + \varepsilon$$



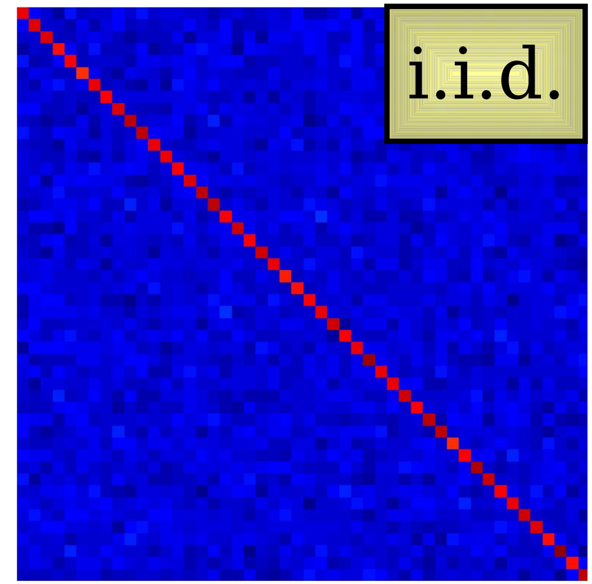
the data and three different models

Error covariance matrix

$Y = X\beta + \varepsilon$ → $Cov(\varepsilon)$

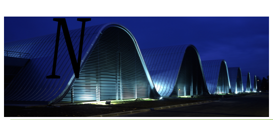


$$V_{ij} = \sigma^2 \frac{a^{|i-j|}}{1-a^2}$$



sampled error covariance matrix

Serial correlations



Maximum likelihood solution

- Least Mahalanobis distance (Gaussian assumption)
  - Assume V is known up to a scalar factor: $V = \sigma^2 W$
  - The ML effect estimator minimizes the Mahalanobis distance

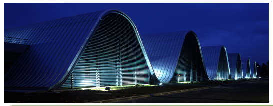
$$d^2_{maha} = (Y - X\beta)^t W^{-1} (Y - X\beta) = \| W^{-\frac{1}{2}} (Y - X\beta) \|^2$$

  - To see this, note that:

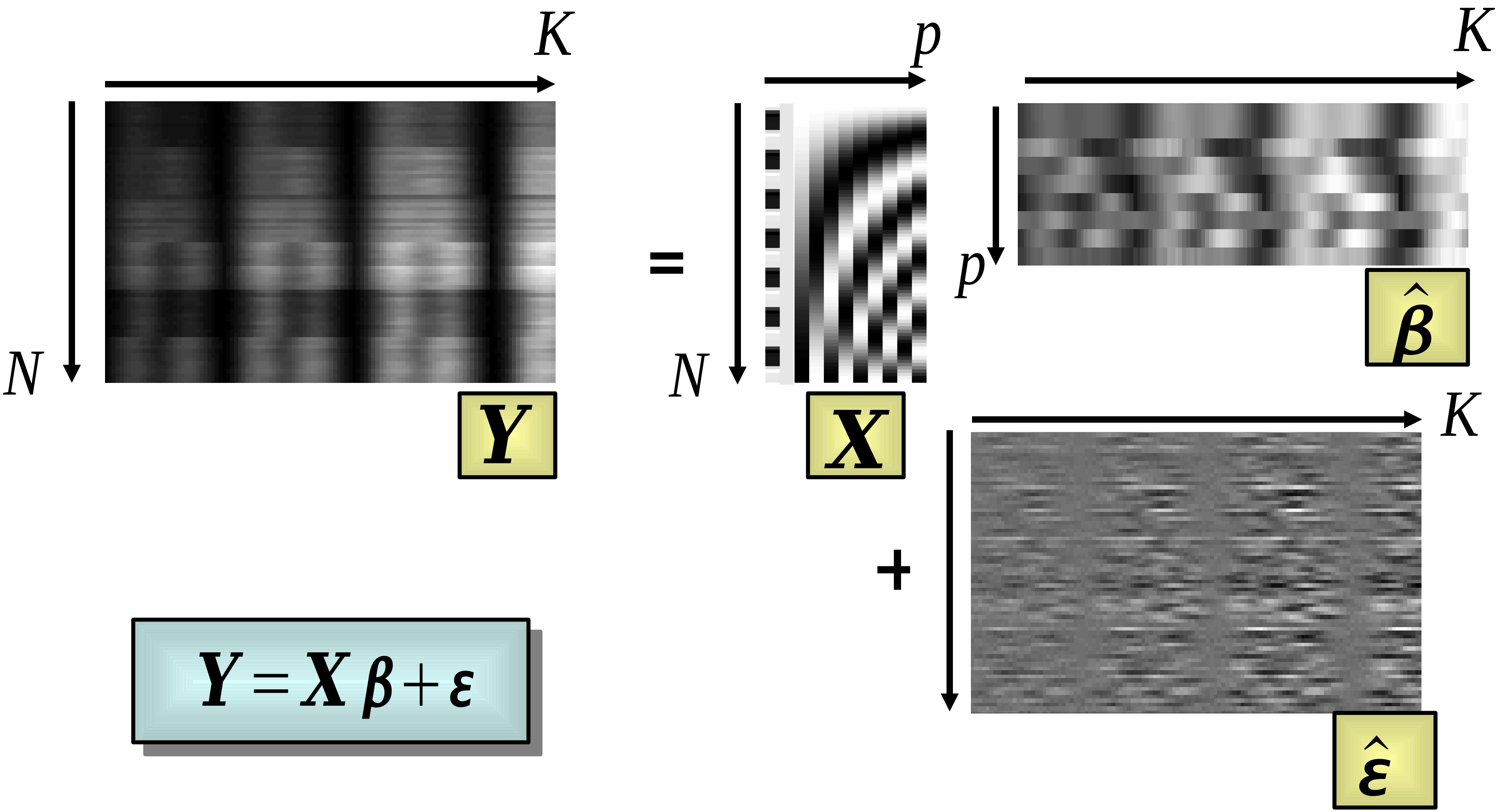
$$W^{-\frac{1}{2}} Y = W^{-\frac{1}{2}} X\beta + \sigma^2 N(0, I_n)$$

  - Therefore, only need to pre-whiten the data and the design matrix

$$Y \leftarrow W^{-\frac{1}{2}} Y, \qquad X \leftarrow W^{-\frac{1}{2}} X$$



Mass-univariate approach

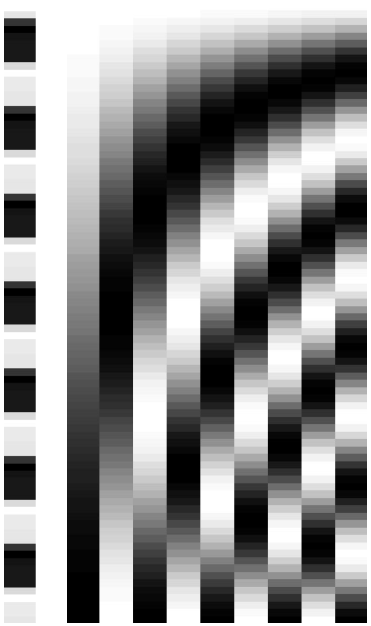


$$Y = X\beta + \varepsilon$$

Inference - t-statistic

$$Y = X\beta + \varepsilon$$

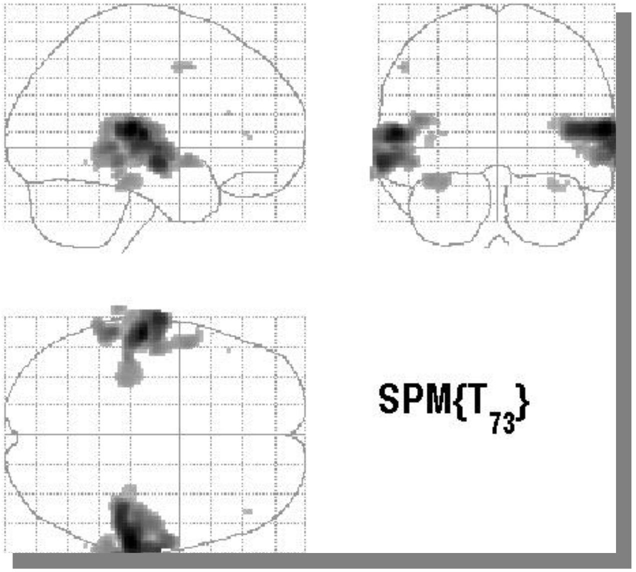
c = +1 0 0 0 0 0 0 0
0 0 0



boxcar parameter > 0 ?

Null hypothesis: $\beta_1 = 0$

$$t = \frac{c^t \hat{\beta}}{\sqrt{Var(c^t \hat{\beta})}}$$



t-statistic - Computations

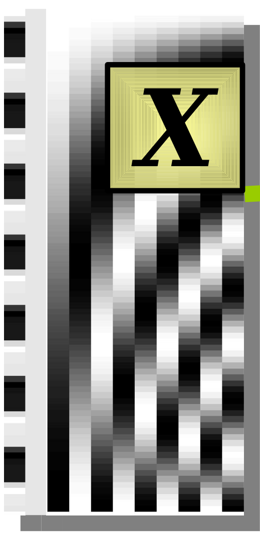
$$Y = X\beta + \varepsilon$$

$\hat{\beta}$ least squares estimates

c = +1 0 0 0 0 0 0 0 0 0 0

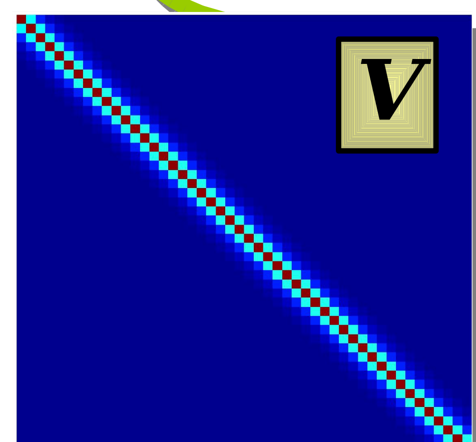
$$t = \frac{c^t \hat{\beta}}{\sqrt{Var(c^t \hat{\beta})}}$$

$$Var(c^t \hat{\beta}) = \hat{\sigma}^2 c^t X V X^{-t} c$$

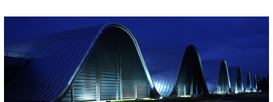


$$\hat{\sigma}^2 = \frac{\sum_{t=1}^{N} \left(y_t - \left[X \hat{\beta} \right]_t \right)}{df}$$

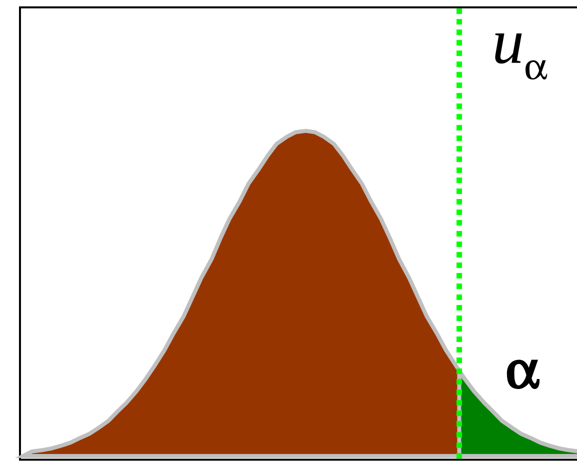
compute ***df*** using Satterthwaite approximation



ReML



Hypothesis Testing



Null Distribution of T

- ***Type I Error α:***

  Acceptable *false positive rate α*.

  Level ⇨ threshold u_α

  Threshold u_α controls the false positive rate

$$\alpha = p\left(T > u_\alpha \middle| H_0\right)$$

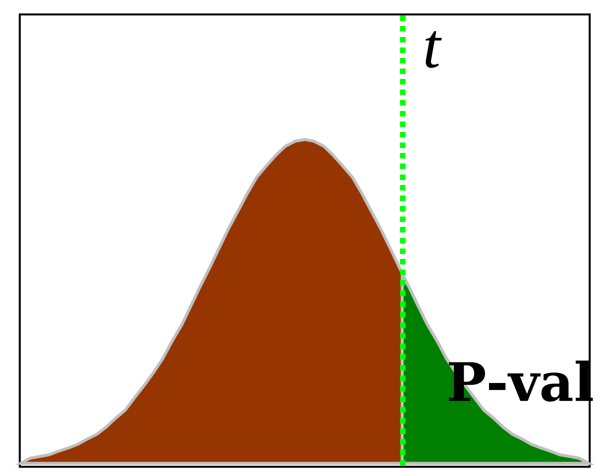
Observation of test statistic t, a realisation of T

- ***P*-value:**

  A *p*-value summarises evidence against H_0.

  This is the change of observing value more extreme than *t* under the null hypothesis.

$$p\left(T > t \middle| H_0\right)$$



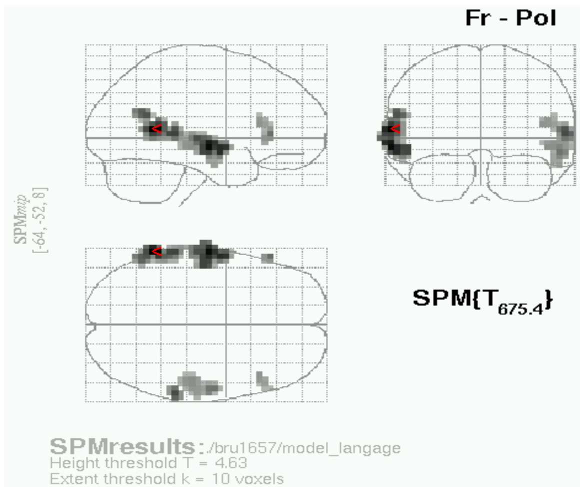
Null Distribution of T

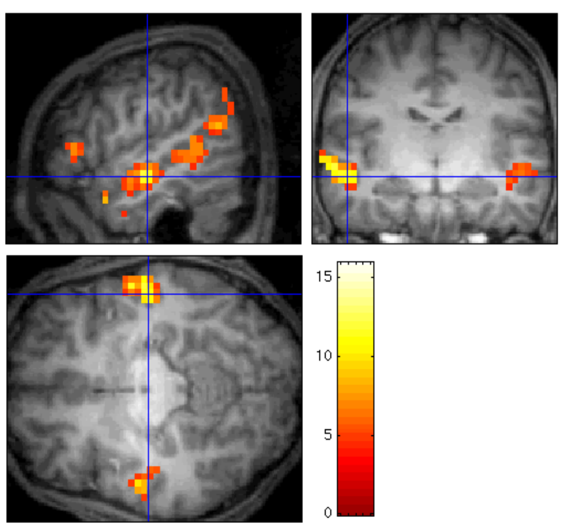
- **The conclusion about the hypothesis:**

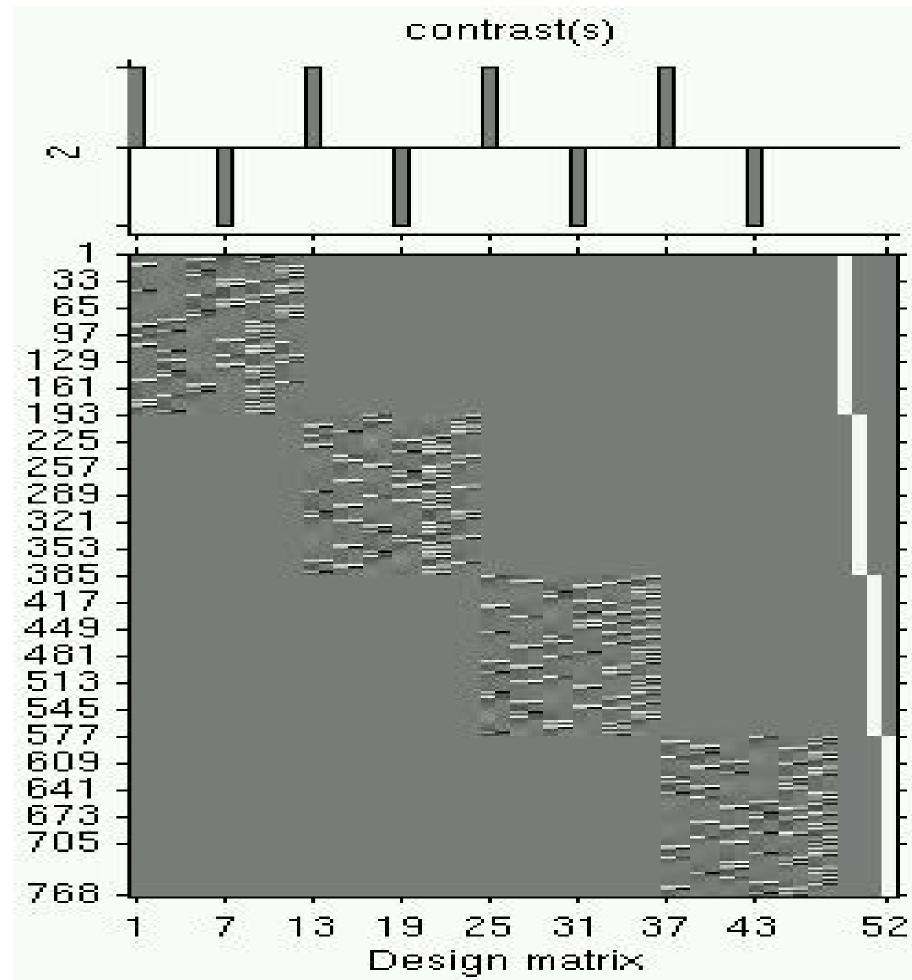
  We reject the null hypothesis in favour of the alternative hypothesis if $t > u_\alpha$

Example of fMRI model

- A language comprehension study **[Pallier et al, 2002]**





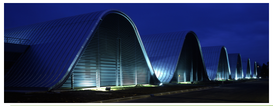


Part I – Mapping brain activity

A) A tour about the GLM framework

B) What kind of regularization?

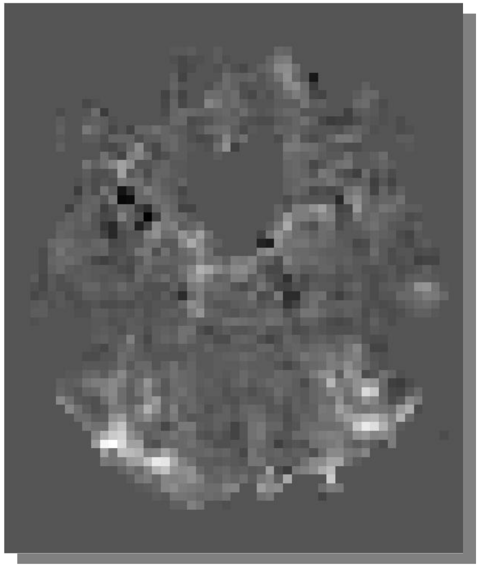
C) Numerical Bayesian inference methods



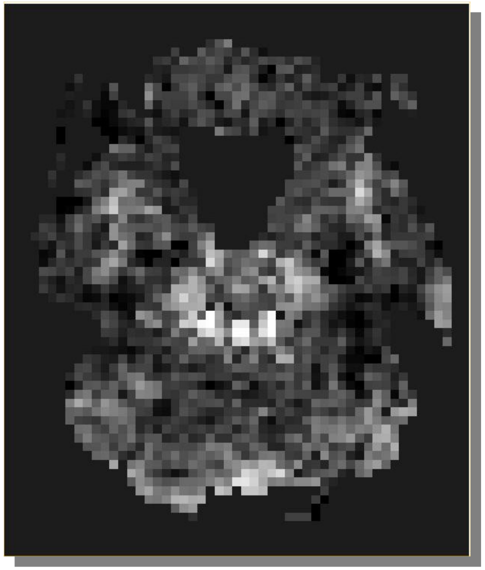
Experimental evidence

Even without applied spatial smoothing, activation maps (and maps of eg. AR coefficients) have spatial structure.

Contrast



AR(1)



⇨ Definition of a spatial prior via Gaussian Markov Random Field
⇨ Automatic spatial regularisation of Regression coefficients and AR coefficients

Bayesian fMRI

Bayesian fMRI

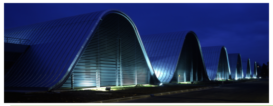
General Linear Model:

$$Y = X\beta + \varepsilon \quad \text{with} \quad \varepsilon \propto N\left(0, C_{\varepsilon}\right)$$

What are the priors?

- In *"classical"* SPM, no (flat) priors
- In *"full"* Bayes, priors might be from theoretical arguments or from independent data
- In *"empirical"* Bayes, priors derive from the same data, assuming a hierarchical model for generation of the data

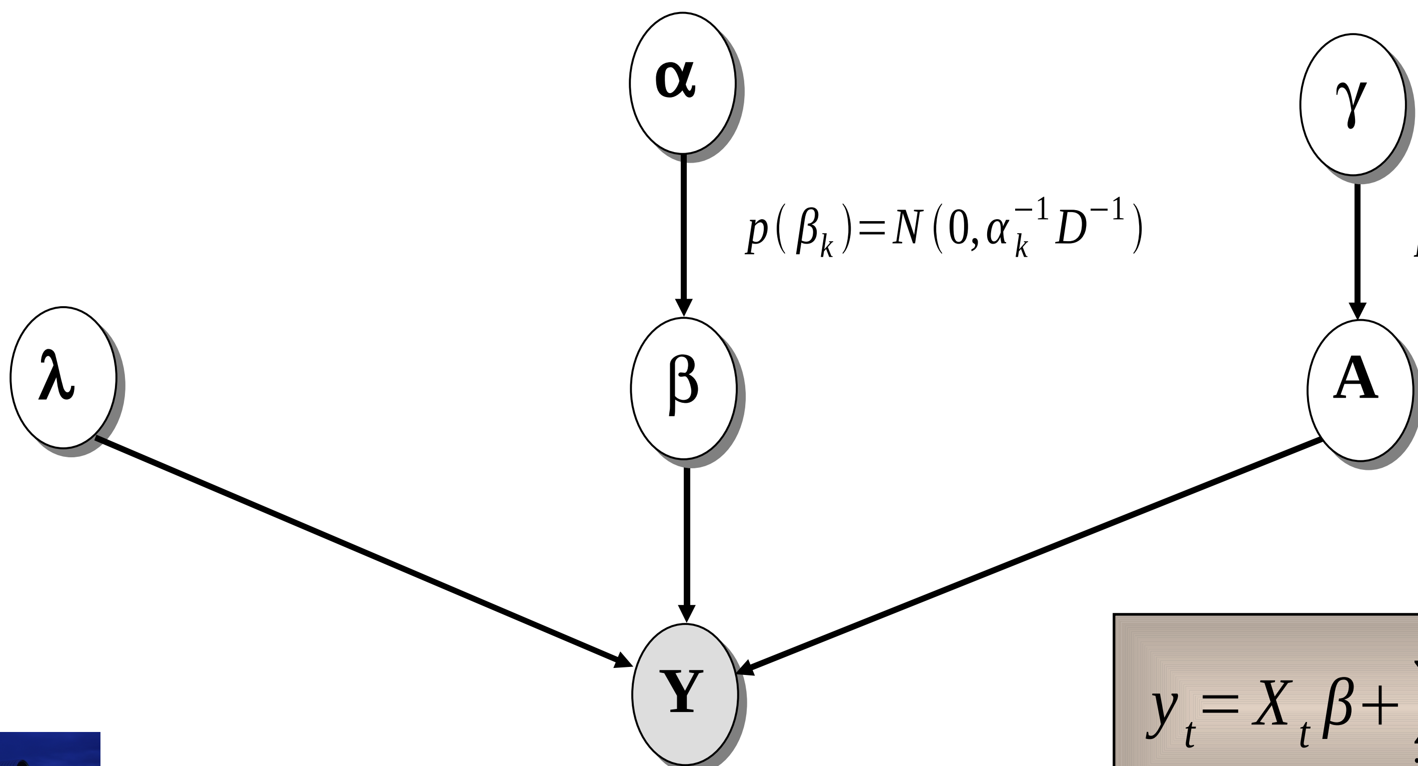
→ Parameters of one level can be made priors on distribution of parameters at lower level



The generative model

General Linear Model with Auto-Regressive error terms (GLM-AR):

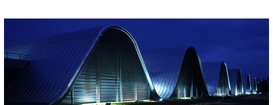
$Y = X\beta + E$ where E is an AR(p)



$$p(\beta_k) = N(0, \alpha_k^{-1} D^{-1})$$

$$p(a_p) = N(0, \gamma_p^{-1} D^{-1})$$

$$y_t = X_t \beta + \sum_{i=1}^{p} a_i e_{t-i} + \varepsilon_t$$



Spatial prior

Over the regression coefficients: [Penny et al, NeuroImage,2003, 2005]

$$p(\beta_k) = N\left(0, \alpha_k^{-1} D^{-1}\right)$$

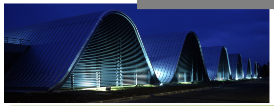
Shrinkage prior → α_k … Spatial precison: determines the amount of smoothness → α_k … Spatial kernel matrix → D

- Shrinkage prior
- Spatial precison: determines the amount of smoothness
- Spatial kernel matrix

Gaussian Markov Random Field priors D

$$D = \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & \ddots & d_{ij} & \\ & & d_{ji} & 1 & \\ & & & & 1 \end{bmatrix}$$

$$\begin{cases} 1 \text{ on diagonal elements } d_{ii} \\ d_{ij} > 0 \text{ if voxels } i \text{ and } j \text{ are neighbors.} \\ 0 \text{ elsewhere} \end{cases}$$



Same prior on the AR coefficients.

Prior, Likelihood and posterior

The prior:

$$p(\beta, A, \lambda, \alpha, \gamma) = \left(\prod_k p(\beta_k | \alpha_k) p(\alpha_k | q_1, q_2) \right) \left(\prod_p p(a_p | \gamma_p) p(\gamma_p | r_1, r_2) \right) \left(\prod_n p(\lambda_n | u_1, u_2) \right)$$

The likelihood:

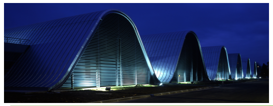
$$p(Y | \beta, A, \lambda) = \prod_n p(y_n | \beta_n, a_n, \lambda_n)$$

The posterior? $p(\beta | Y)$?

The posterior over β doesn't factorise over k or n.

⇨ Exact inference requires sampling techniques

⇨ Variational approximation achievable at lower cost

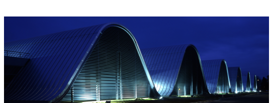


Part I – Mapping brain activity

A) A tour about the GLM framework

B) What kind of regularization?

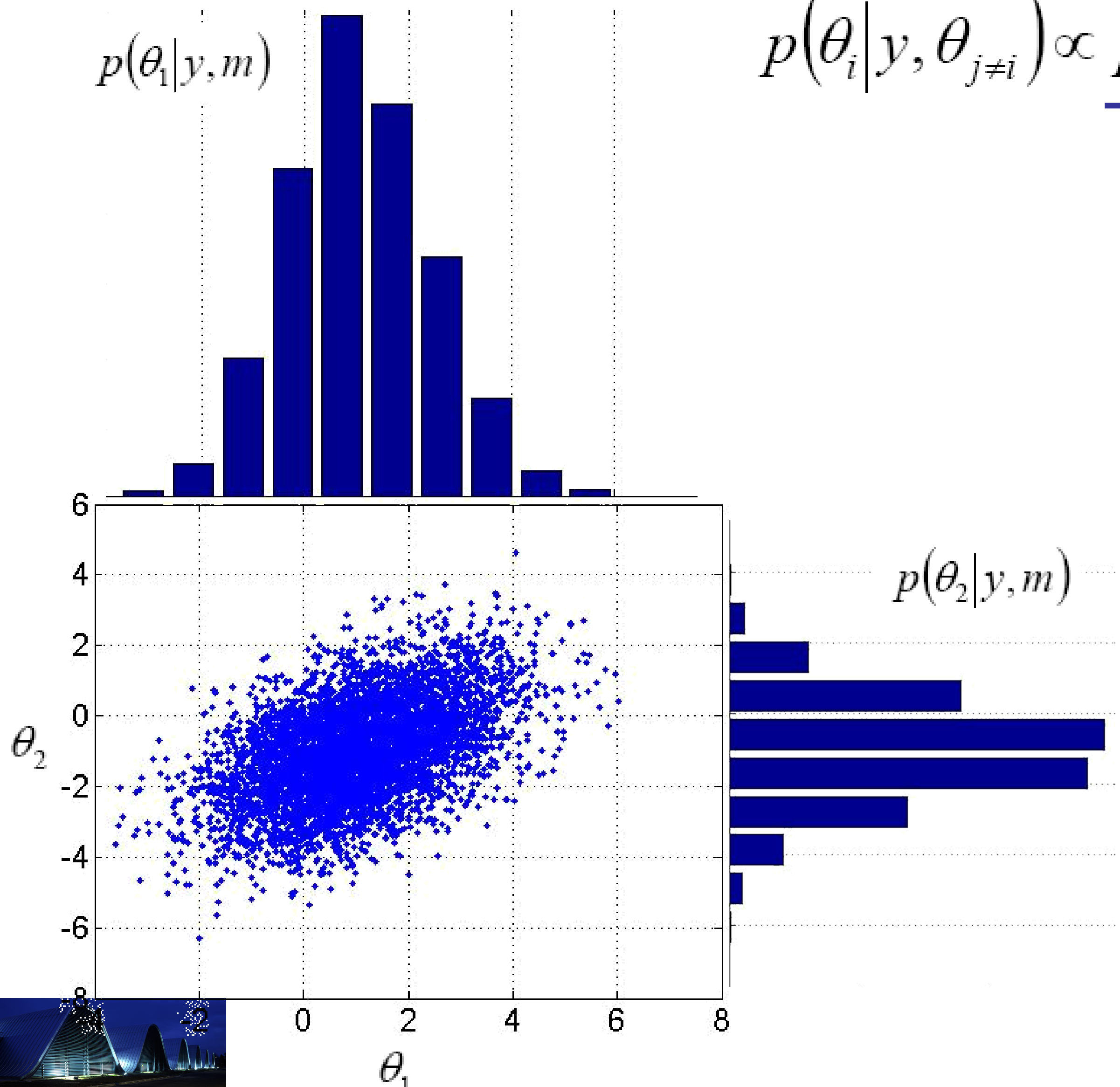
C) Numerical Bayesian inference methods

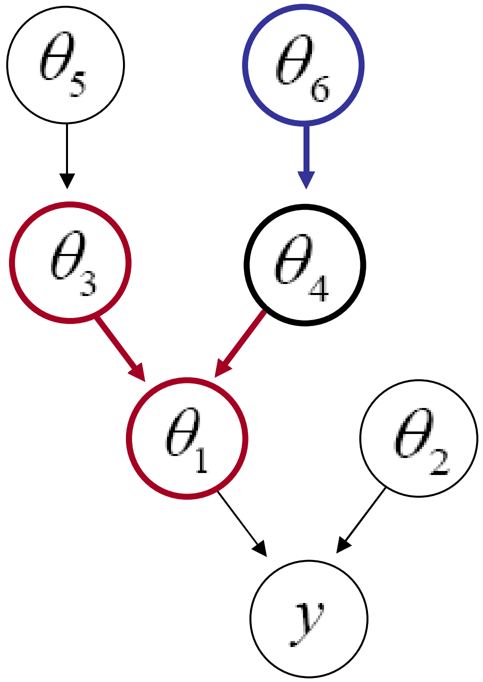


Sampling methods

MCMC example: Gibbs sampling

$$p\left(\theta_i \middle| y, \theta_{j\neq i}\right) \propto \underline{p\left(\theta_i \middle| par(\theta_i)\right)} \underline{\prod_{j=ch(i)} p\left(\theta_j \middle| par(\theta_j)\right)}$$





Variational Bayes

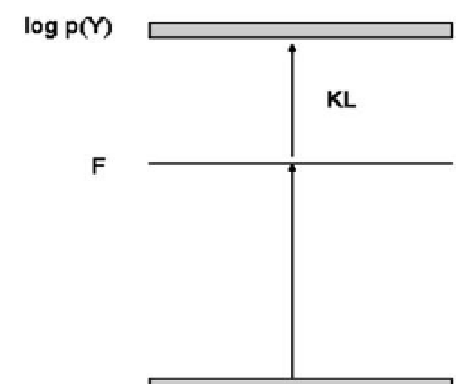
- Central quantity of Bayesian learning $p(\boldsymbol{\theta} \,|\, \boldsymbol{Y})$ with $\boldsymbol{\theta} = \{\beta, A, \lambda\}$

- Log-evidence of the model or integrated likelihood

$$
\begin{aligned}
\forall \text{ pdf } q, \log p(\boldsymbol{Y}) &= \int q(\boldsymbol{\theta} \,|\, \boldsymbol{Y}) \log p(\boldsymbol{Y})\, d\boldsymbol{\theta} \\
&= \int q(\boldsymbol{\theta} \,|\, \boldsymbol{Y}) \log \frac{p(\boldsymbol{Y}, \boldsymbol{\theta})}{p(\boldsymbol{\theta} \,|\, \boldsymbol{Y})}\, d\boldsymbol{\theta} \\
&= \int q(\boldsymbol{\theta} \,|\, \boldsymbol{Y}) \log [\frac{q(\boldsymbol{\theta} \,|\, \boldsymbol{Y}) p(\boldsymbol{Y}, \boldsymbol{\theta})}{p(\boldsymbol{\theta} \,|\, \boldsymbol{Y}) q(\boldsymbol{\theta} \,|\, \boldsymbol{Y})}]\, d\boldsymbol{\theta} \\
&= \int q(\boldsymbol{\theta} \,|\, \boldsymbol{Y}) \log \frac{p(\boldsymbol{\theta}, \boldsymbol{Y})}{q(\boldsymbol{\theta} \,|\, \boldsymbol{Y})}\, d\boldsymbol{\theta} + \int q(\boldsymbol{\theta} \,|\, \boldsymbol{Y}) \log \frac{q(\boldsymbol{\theta} \,|\, \boldsymbol{Y})}{p(\boldsymbol{\theta} \,|\, \boldsymbol{Y})}\, d\boldsymbol{\theta} \\
&= \mathcal{F} + D(q || p_{\theta \,|\, \boldsymbol{Y}})
\end{aligned}
$$

- $\mathcal{F}$ is a lower bound of the model evidence

$$\log p(\boldsymbol{Y}) = \mathcal{F} \Longleftrightarrow q(\theta \,|\, \boldsymbol{Y}) = p(\theta \,|\, \boldsymbol{Y})$$



Variational Bayes (cont'd)

- Aim of VB : maximize $\mathcal{F}$
  - Make the approximate posterior $q(\boldsymbol{\theta} \,|\, \boldsymbol{Y})$ as close as possible to the true posterior $p(\boldsymbol{\theta} \,|\, \boldsymbol{Y})$

- Practical efficient algorithm:
  - Ensure tractability of integrals in $\mathcal{F}$
  - Generic procedure: mean-field approximation

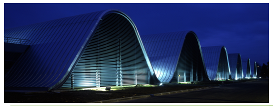
$$q(\boldsymbol{\theta} \,|\, \boldsymbol{Y}) = \prod_i q(\theta_i \,|\, \boldsymbol{Y})$$

$\theta_i =$ ith group of parameters

- Maximizers of $\mathcal{F}$ **[Lappalainen and Miskin, 2000]**

$$q(\theta_i \,|\, \boldsymbol{Y}) = \frac{\exp\left[I(\theta_i)\right]}{\int \exp\left[I(\theta_i)\right] d\theta_i}$$

$$\text{with } I(\theta_i) = \int q(\boldsymbol{\theta}^{\setminus i} \,|\, \boldsymbol{Y}) \log p(\boldsymbol{Y}, \boldsymbol{\theta}) \, d\boldsymbol{\theta}^{\setminus i}$$



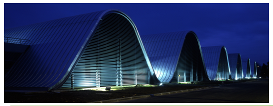
Variational Bayes (exemple)

Approximate posteriors that allows for *factorisation*

$$q(\beta, A, \lambda, \alpha, \gamma) = \left(\prod_k q(\alpha_k | Y)\right)\left(\prod_p q(\gamma_p | Y)\right)\left(\prod_n q(\beta_n | Y)\, q(a_n | Y)\, q(\lambda_n | Y)\right)$$

Variational Bayes Algorithm

Initialisation
While (ΔF > tol)
    Update Suff. Stats. for β
    Update Suff. Stats. for A
    Update Suff. Stats. for λ
    Update Suff. Stats. for α
    Update Suff. Stats. for γ
End



Variational Bayes (exemple)

Approximate posteriors that allows for *factorisation*

$$q(\beta, A, \lambda, \alpha, \gamma) = \left(\prod_k q(\alpha_k|Y)\right)\left(\prod_p q(\gamma_p|Y)\right)\left(\prod_n q(\beta_n|Y)\,q(a_n|Y)\,q(\lambda_n|Y)\right)$$

Regression coefficients

$$q(\mathbf{w}_n) = N\left(\mathbf{w}_n; \hat{\mathbf{w}}_n, \hat{\boldsymbol{\Sigma}}_n\right)$$
$$\hat{\mathbf{w}}_n = \hat{\boldsymbol{\Sigma}}_n\left(\bar{\lambda}_n \tilde{\mathbf{b}}_n^{\,T} + \mathbf{r}_n\right)$$
$$\hat{\boldsymbol{\Sigma}}_n = \left(\bar{\lambda}_n \tilde{\mathbf{A}}_n + \mathbf{B}_{nn}\right)^{-1}$$
$$\mathbf{B} = \mathbf{H}\left(diag(\bar{\boldsymbol{\alpha}}) \otimes \mathbf{S}^T\mathbf{S}\right)\mathbf{H}^T$$
$$\mathbf{r}_n = -\sum_{i=1, i\neq n}^{N} \mathbf{B}_{ni}\hat{\mathbf{w}}_i$$

AR coefficients

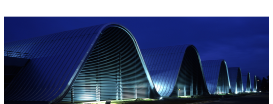
$$q(\mathbf{a}_n) = N(\mathbf{a}_n; \mathbf{m}_n, \mathbf{V}_n)$$
$$\mathbf{V}_n = \left(\bar{\lambda}_n \tilde{\mathbf{C}}_n + \beta \mathbf{I}_p\right)^{-1}$$
$$\mathbf{m}_n = \bar{\lambda}_n \tilde{\mathbf{D}}_n \mathbf{V}_n$$

Observation noise

$$q(\lambda_n) = Ga(\lambda_n; b_n, c_n)$$
$$\frac{1}{b_n} = \frac{\tilde{G}_n}{2} + \frac{1}{u_1}$$
$$c_n = \frac{T}{2} + u_2$$

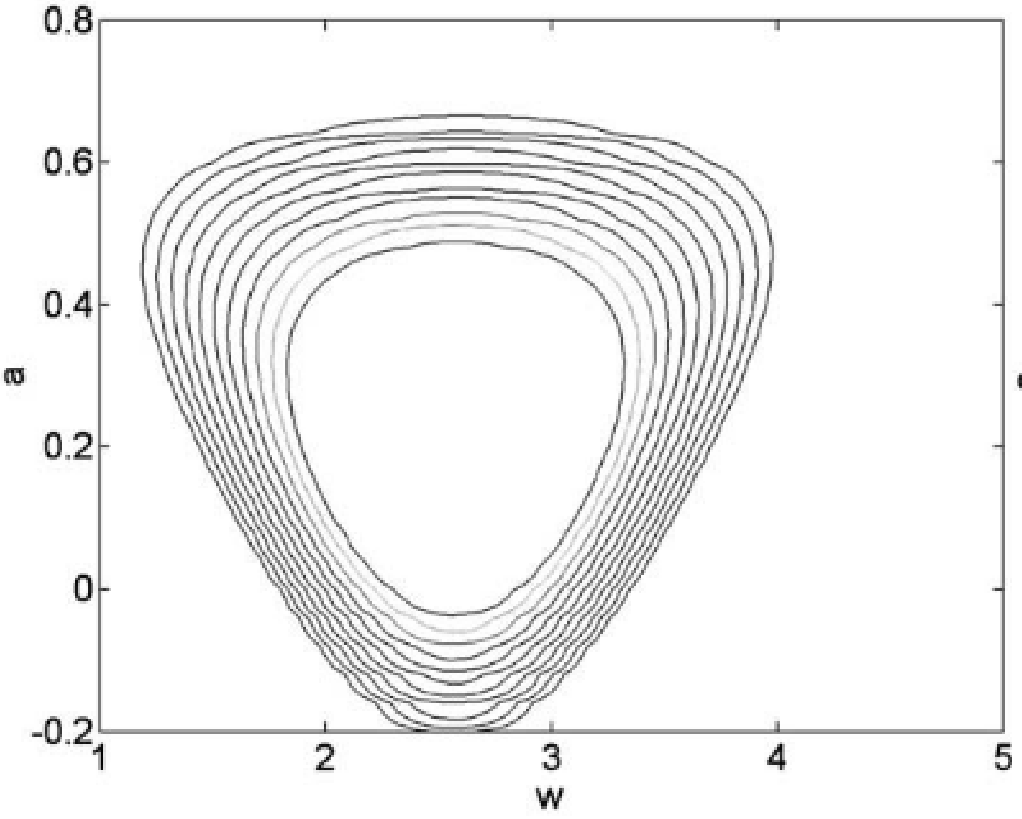
Spatial precisions

$$q(\boldsymbol{\alpha}) = \prod_{k=1}^{K} q(\alpha_k)$$
$$q(\alpha_k) = Ga(\alpha_k; g_k, h_k)$$
$$\frac{1}{g_k} = \frac{1}{2}\left[Tr\left(\hat{\boldsymbol{\Sigma}}_k \mathbf{S}^T\mathbf{S}\right) + \hat{\mathbf{w}}_k^{\,T}\mathbf{S}^T\mathbf{S}\hat{\mathbf{w}}_k\right] + \frac{1}{q_1}$$
$$h_k = \frac{N}{2} + q_2$$
$$\bar{\alpha}_k = g_k h_k$$

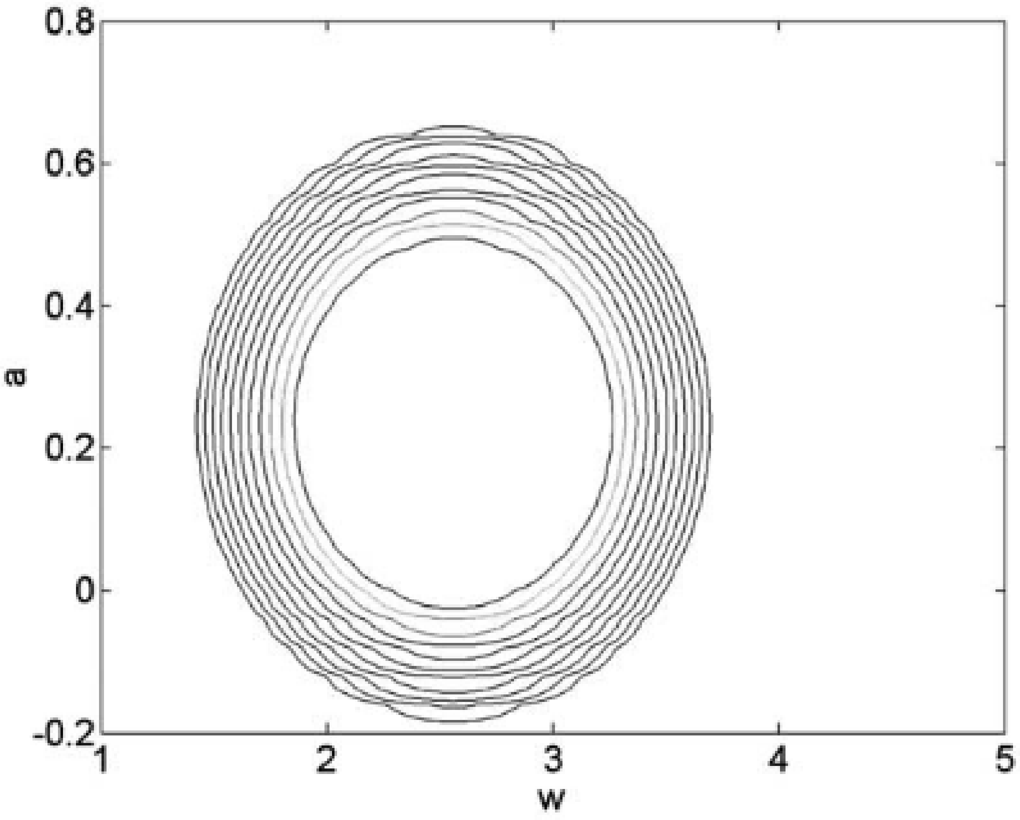


VB: approximation effect

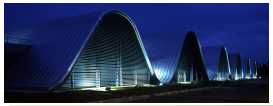
iso-levels of the posterior $p(\boldsymbol{\theta} \mid \boldsymbol{Y})$



iso-levels of the posterior $q(\boldsymbol{\theta} \mid \boldsymbol{Y})$

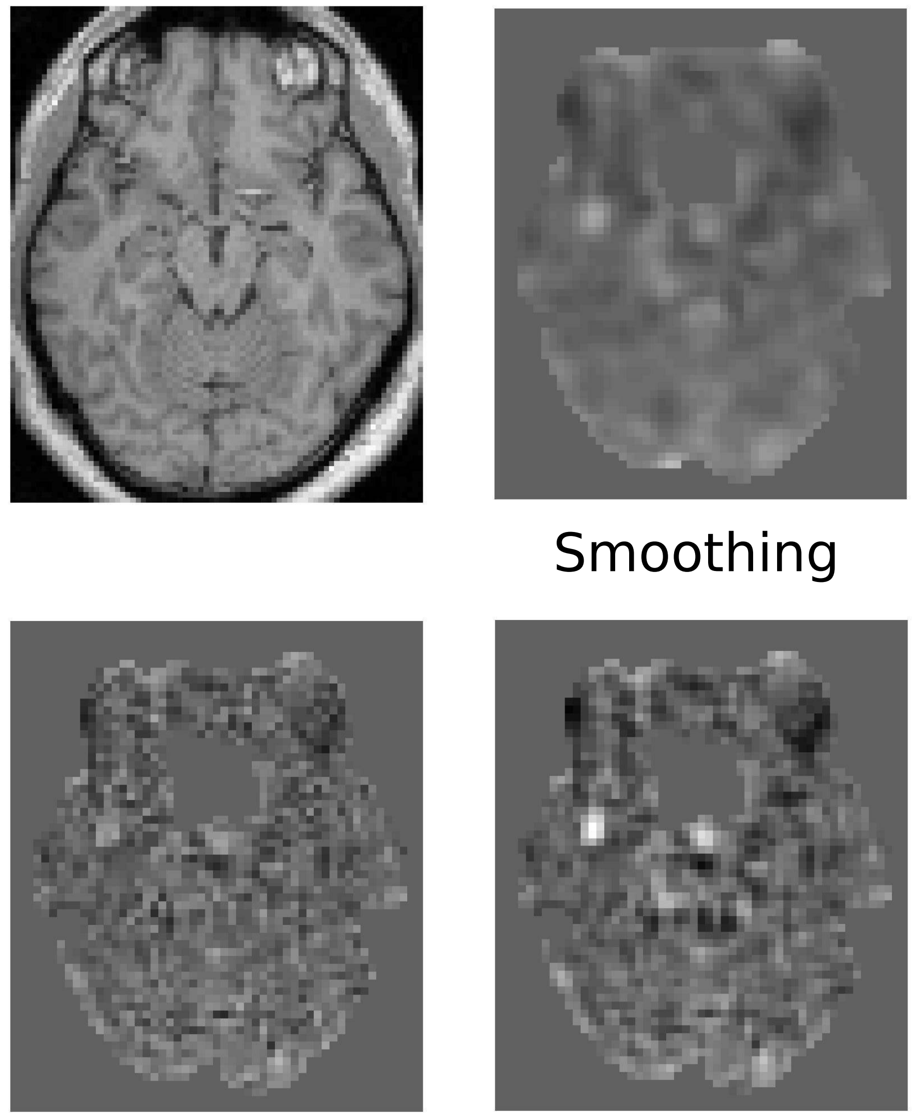


Accurate mode approximation
Error hidden in the higher order moments



Event-related fMRI

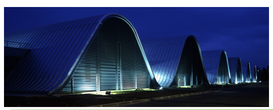
Familiar vs.
unfamiliar faces



Smoothing

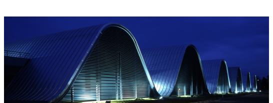
Global prior

Spatial Prior

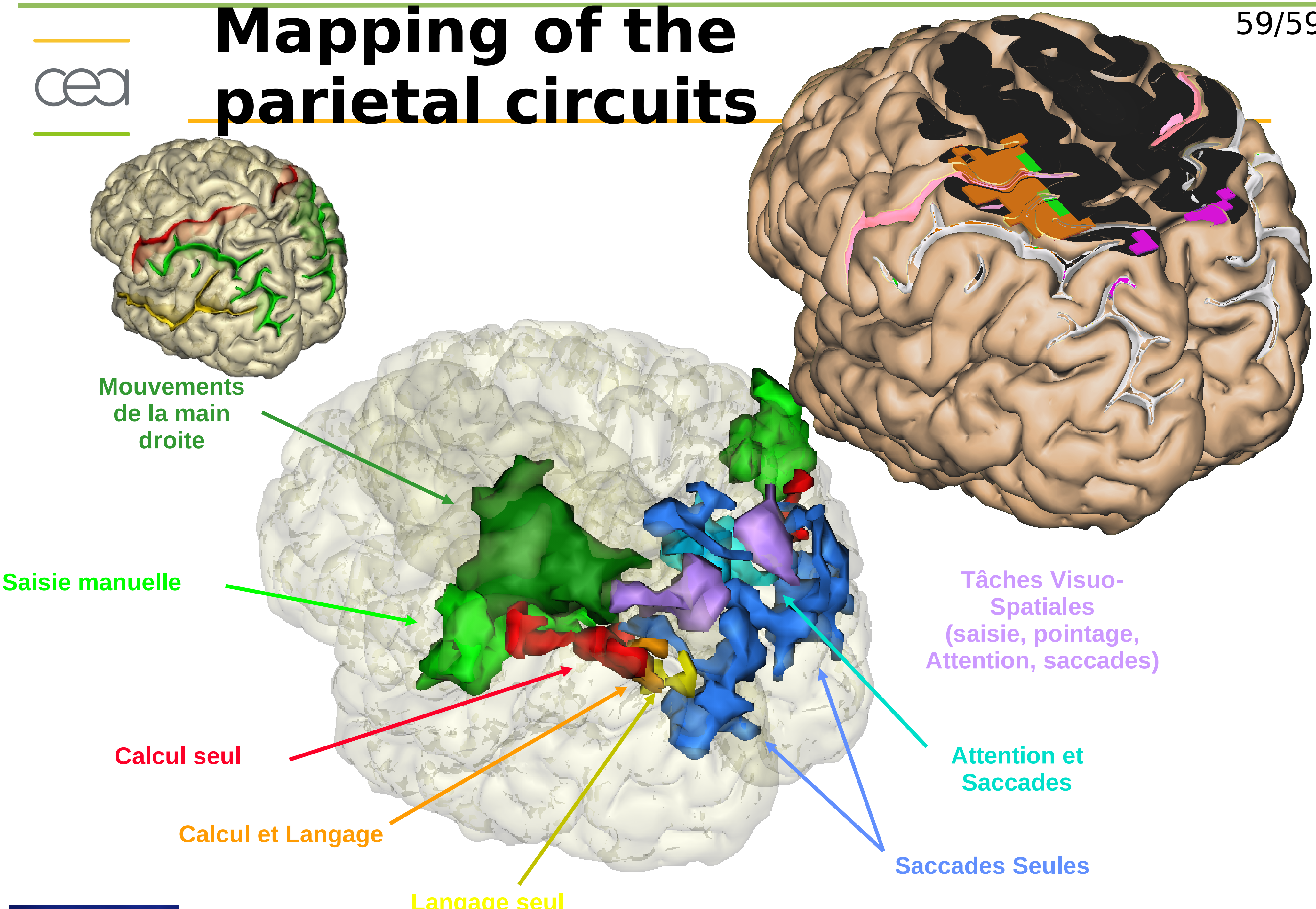


Summary

- Activation detection in fMRI
  - Preprocessings
  - A whole brain model of the BOLD signal
  - Statistical tests
- Bayesian gain
  - Don't smooth the data
  - Prefer spatial regularization



Mapping of the parietal circuits



O. Simon, D. Rivière (2000)