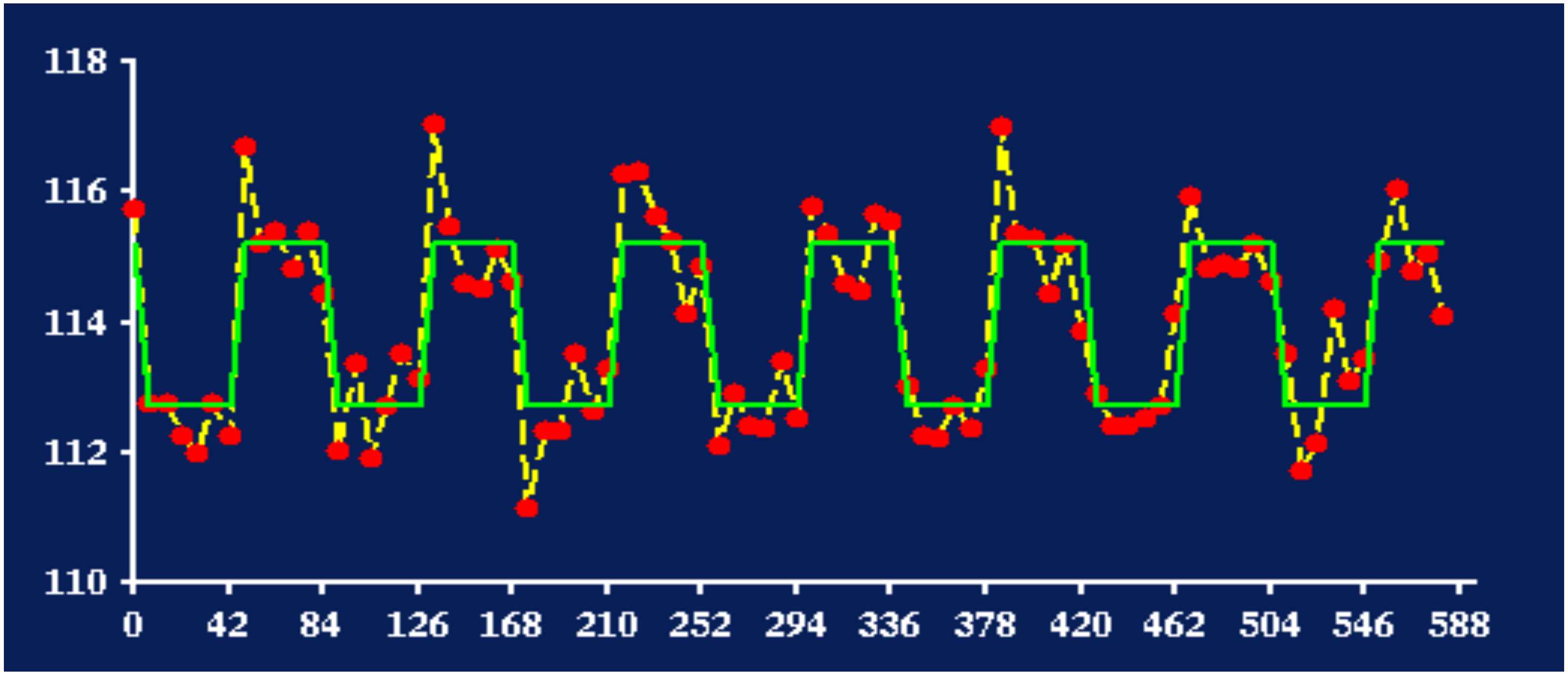
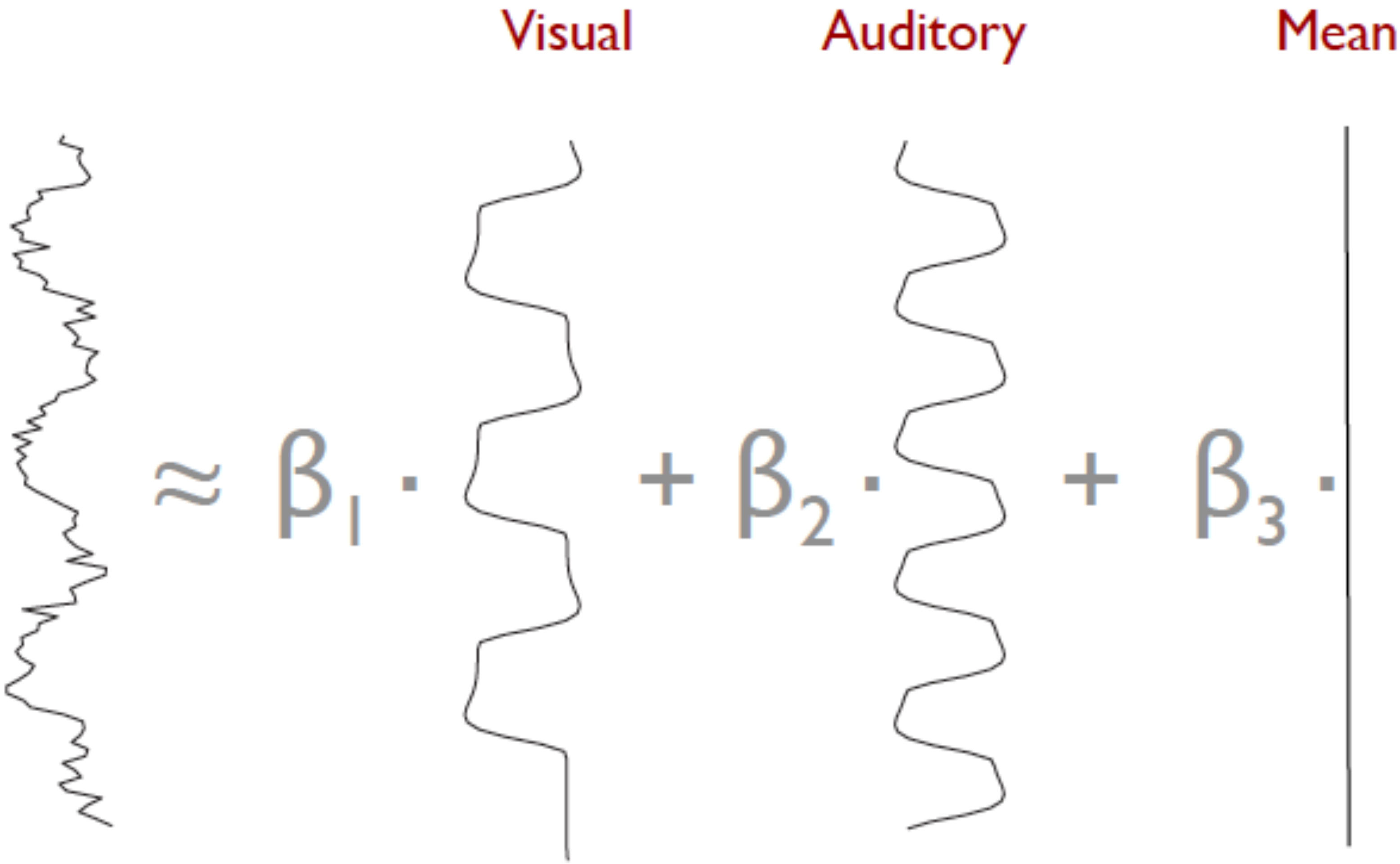
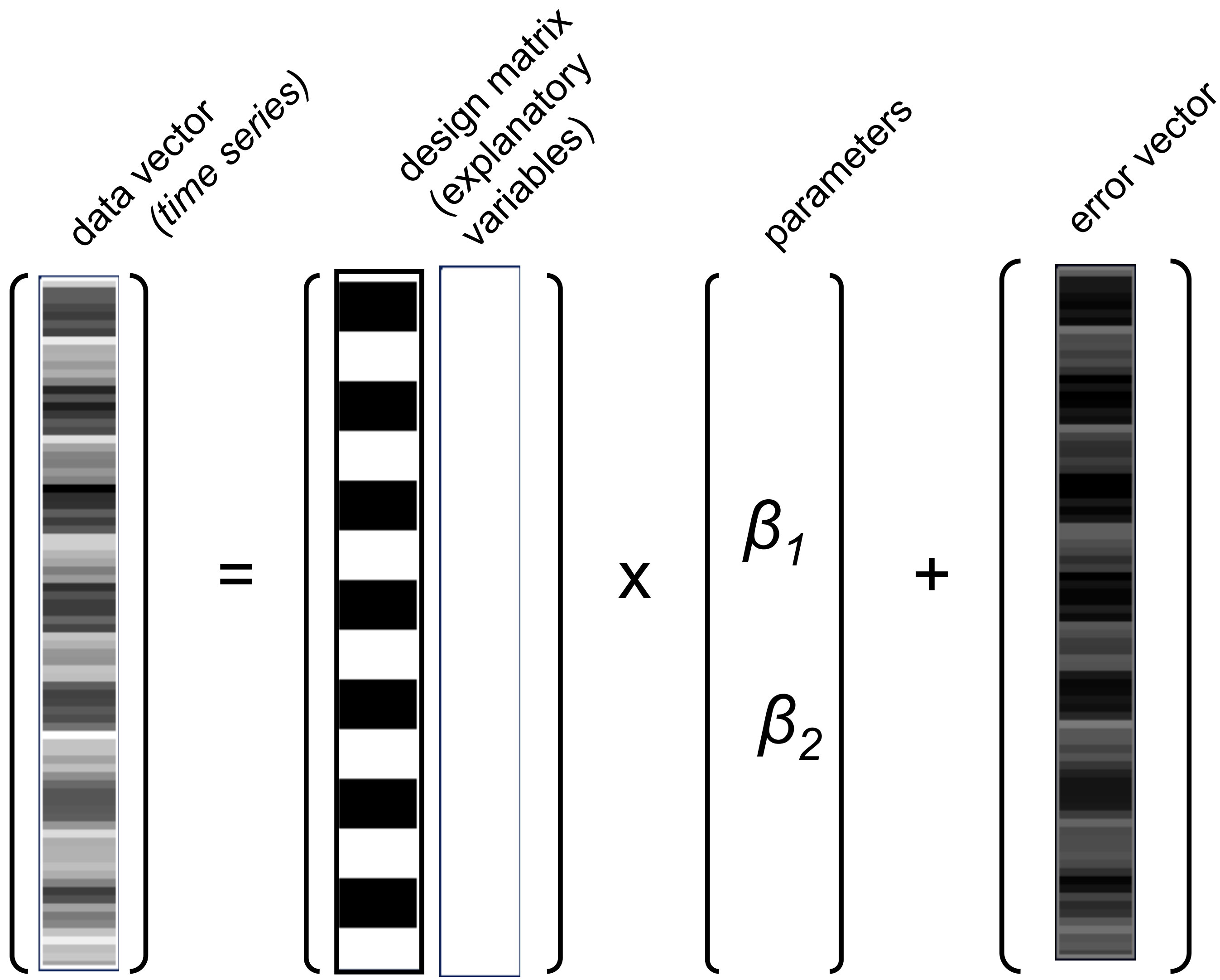

118
116
114
112
110
0
42
84
126
168
210
252
294
336
378
420
462
504
546
588



Visual
Auditory
Mean
$\approx \beta_1 \cdot \quad + \beta_2 \cdot \quad + \beta_3 \cdot$

The “General” Linear Model



$$\underline{y} = X \times \underline{\beta} + \underline{\varepsilon}$$

Matrix Formulation

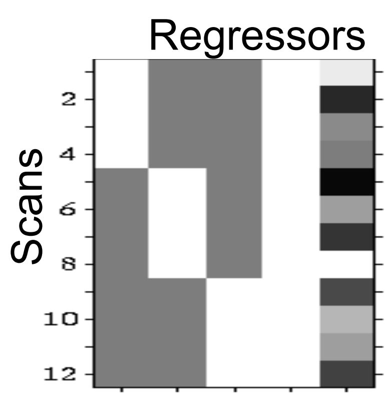
$$\boxed{Y_j} = x_{j1}\boxed{\beta_1} + \ldots + x_{jl}\boxed{\beta_l} + \ldots + x_{jL}\boxed{\beta_L} + \boxed{\epsilon_j}$$

Equation for scan j

Simultaneous equations for scans *1.. J*

$$\begin{aligned}
Y_1 &= x_{11}\boxed{\beta_1} + \ldots + x_{1l}\boxed{\beta_l} + \ldots + x_{1L}\boxed{\beta_L} + \epsilon_1 \\
\vdots &= \vdots \\
Y_j &= x_{j1}\boxed{\beta_1} + \ldots + x_{jl}\boxed{\beta_l} + \ldots + x_{jL}\boxed{\beta_L} + \epsilon_j \\
\vdots &= \vdots \\
Y_J &= x_{J1}\boxed{\beta_1} + \ldots + x_{Jl}\boxed{\beta_l} + \ldots + x_{JL}\boxed{\beta_L} + \epsilon_J
\end{aligned}$$

…that can be solved for parameters $\beta_{1..L}$



$$\begin{pmatrix} Y_1 \\ \vdots \\ Y_j \\ \vdots \\ Y_J \end{pmatrix} = \begin{pmatrix} x_{11} & \cdots & x_{1l} & \cdots & x_{1L} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{j1} & \cdots & x_{jl} & \cdots & x_{jL} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{J1} & \cdots & x_{Jl} & \cdots & x_{JL} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_l \\ \vdots \\ \beta_L \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_l \\ \vdots \\ \epsilon_L \end{pmatrix}$$

$$\underbrace{Y}_{J\times 1} \qquad \underbrace{X}_{J\times L} \qquad \underbrace{\beta}_{L\times 1} \qquad \underbrace{\epsilon}_{J\times 1}$$

Gauss-Markov Solution

- Model Error is $\underline{y} - X\underline{\hat{\beta}}$
- SSE = $(\underline{y}-X\underline{\hat{\beta}})'(\underline{y}-X\underline{\hat{\beta}})$
- SSE = $\underline{yy}' + XX'\underline{\hat{\beta}}^2 - 2X'y\underline{\hat{\beta}}$
- Differentiate respect to $\hat{\beta}$ and solve for 0
- $2XX'\underline{\hat{\beta}} - 2X'\underline{y}=0$
- $XX'\underline{\hat{\beta}} = X'\underline{y}$
- $\underline{\hat{\beta}}$ = inv(X X')X'$\underline{y}$ Ordinary Least Squares Solution
- Fitted response is:

$$\underline{\hat{Y}} = X\underline{\hat{\beta}}$$

Bad Design

A word on correlation/estimatability

b = inv(XX')YX'

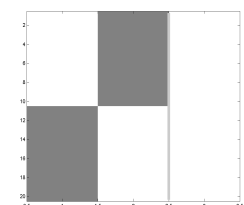
inv(XX')

Assumption 1: (X must be full rank)

Note that X'X is symmetric, and can be rewritten after SVD to:

X'X = QDQ' where the columns of Q are the eigenvectors of X'X

$$D^{-1} = \begin{bmatrix} 1/d_1 & 0 & \dots & 0 \\ 0 & 1/d_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1/d_N \end{bmatrix}$$



d_i is the i^{th} eigenvalue. If any are 0 the matrix has no inverse (singular). If any is small, the matrix is ill conditioned and the inverse is poorly estimated.

Jitter the design matrix
Independent predictor variables
Watch out for overlap with temporal derivatives

Bad Design

Another word on correlation/estimability

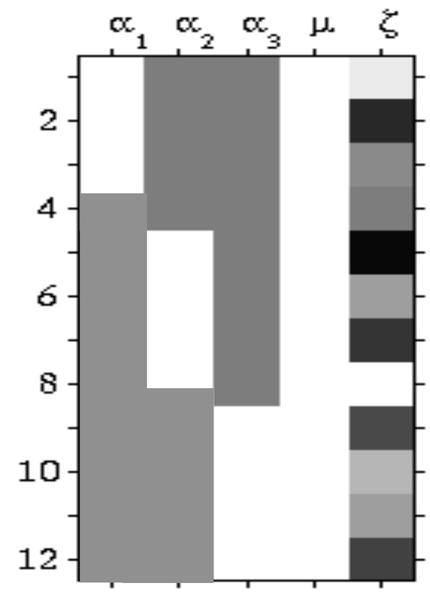
b = inv(XX')YX'

inv(XX')

- When there is high (but not perfect) correlation between regressors, parameters can still be estimated…

…but

- the estimates will be inefficiently estimated (ie highly variable)
- …meaning some contrasts will not lead to very powerful tests



General Linear Model (Estimation)

Estimate parameters from least squares fit to BOLD data, $\underline{B}$:

$\underline{\hat{\beta}}$ = inv(X X')X'B   Ordinary Least Squares Solution

Fitted response is:

$$\underline{\hat{Y}} = X\underline{\hat{\beta}}$$

Estimate Error Variance:

$\sigma_\varepsilon^2 = 1/df_E\,(\underline{B} - \underline{X\hat{\beta}})'(\underline{B} - X\underline{\hat{\beta}})$

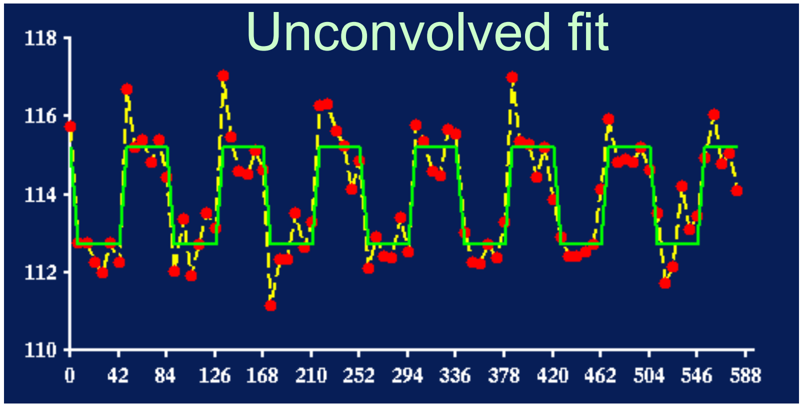
df_E = (# TRs) – (# parameters in $\underline{\hat{\beta}}$)

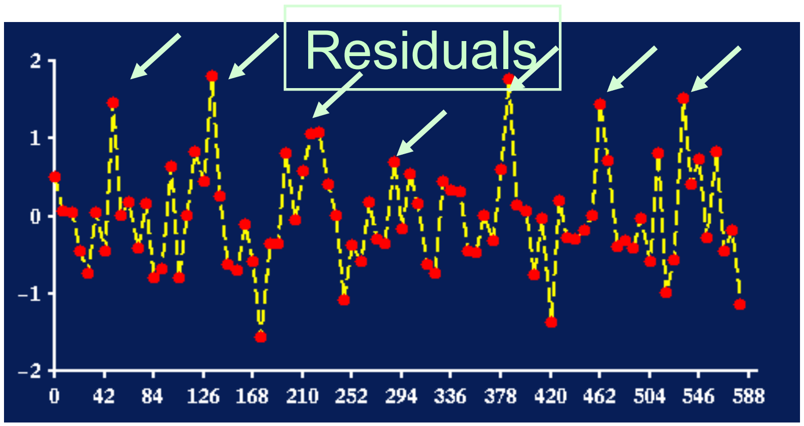
Assumption 2: OLS produces the minimum variance unbiased estimators of β if:

$\Sigma_\varepsilon = \sigma_\varepsilon^2 I$

That is, the error variance is the same on every TR. There is homogeneity of variance. Covariance between errors in any two TRs is 0.

Correlated error:
Structure in the residuals





What are the sources of correlated error?

Hint: $1/f$

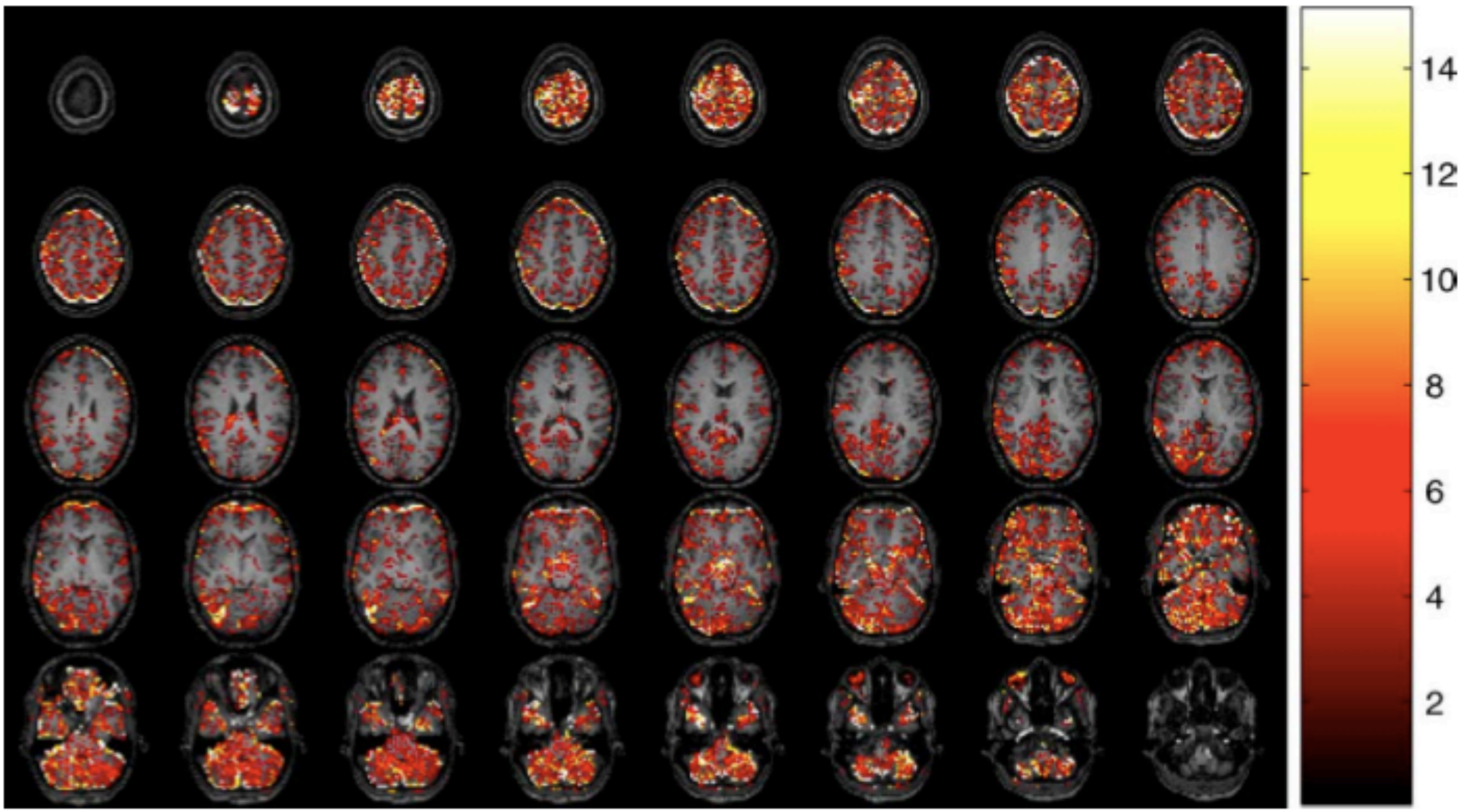
Sources of correlated error

- Hemodynamic response function (smooths Y over time)
  (we will come back to this in a minute)
- Heart Rate
- Respiratory Rate
- Scanner Drift
- Head Motion
- Other

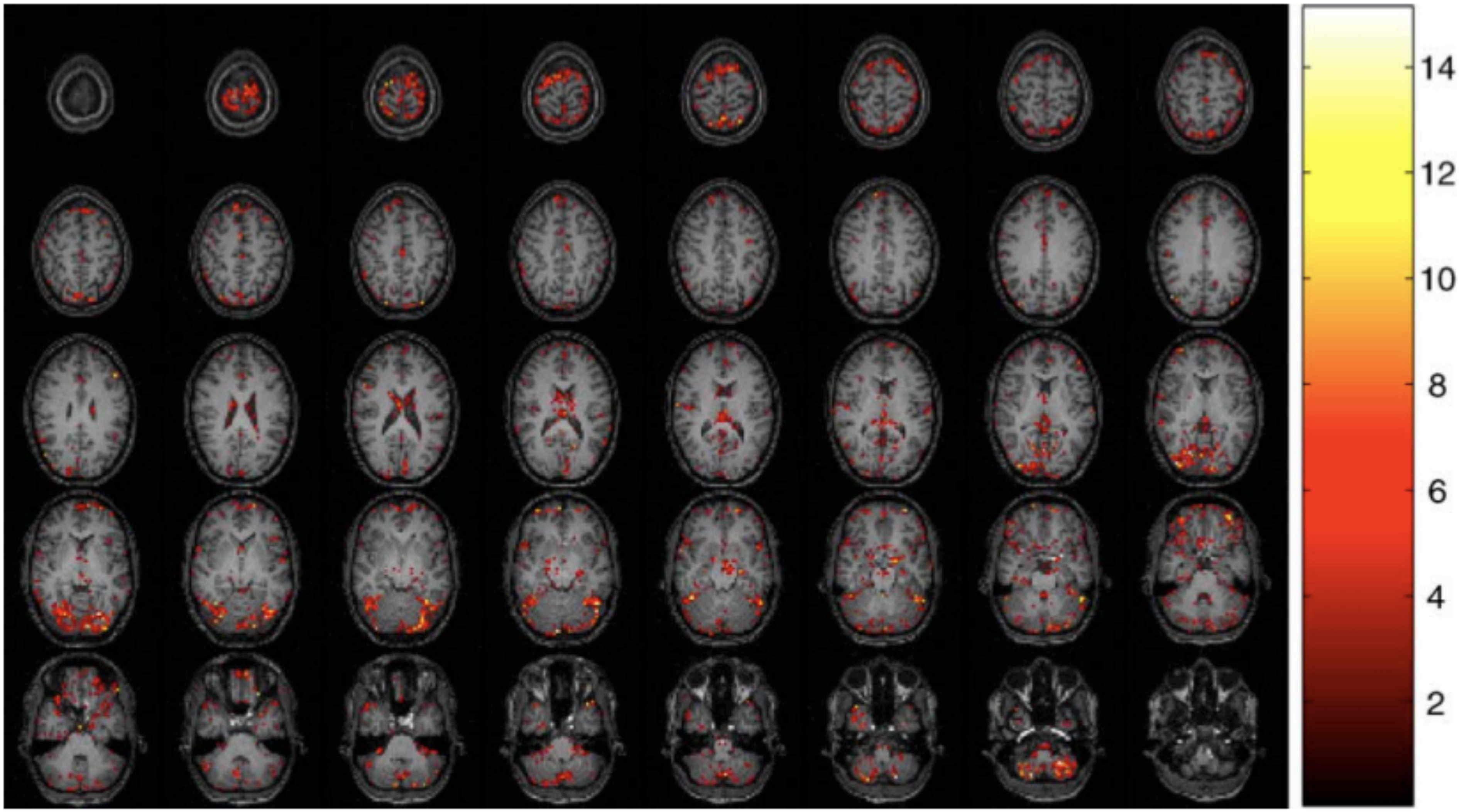
Non-white noise in fMRI: Does modelling have an impact?

Torben E. Lund,[a,*] Kristoffer H. Madsen,[a,b] Karam Sidaros,[a]
Wen-Lin Luo,[c] and Thomas E. Nichols[d]

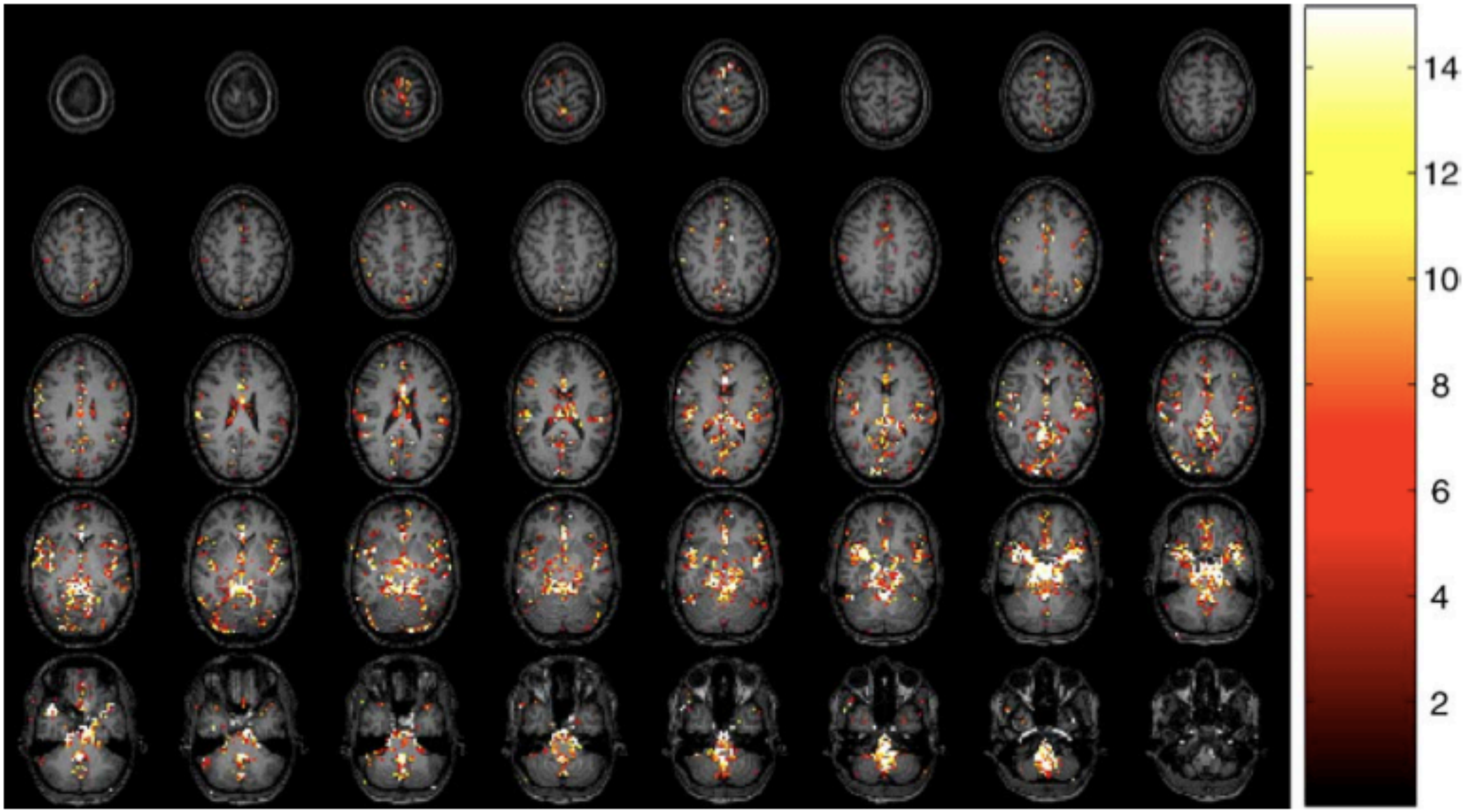
Low Frequency Noise: F-test of basis functions



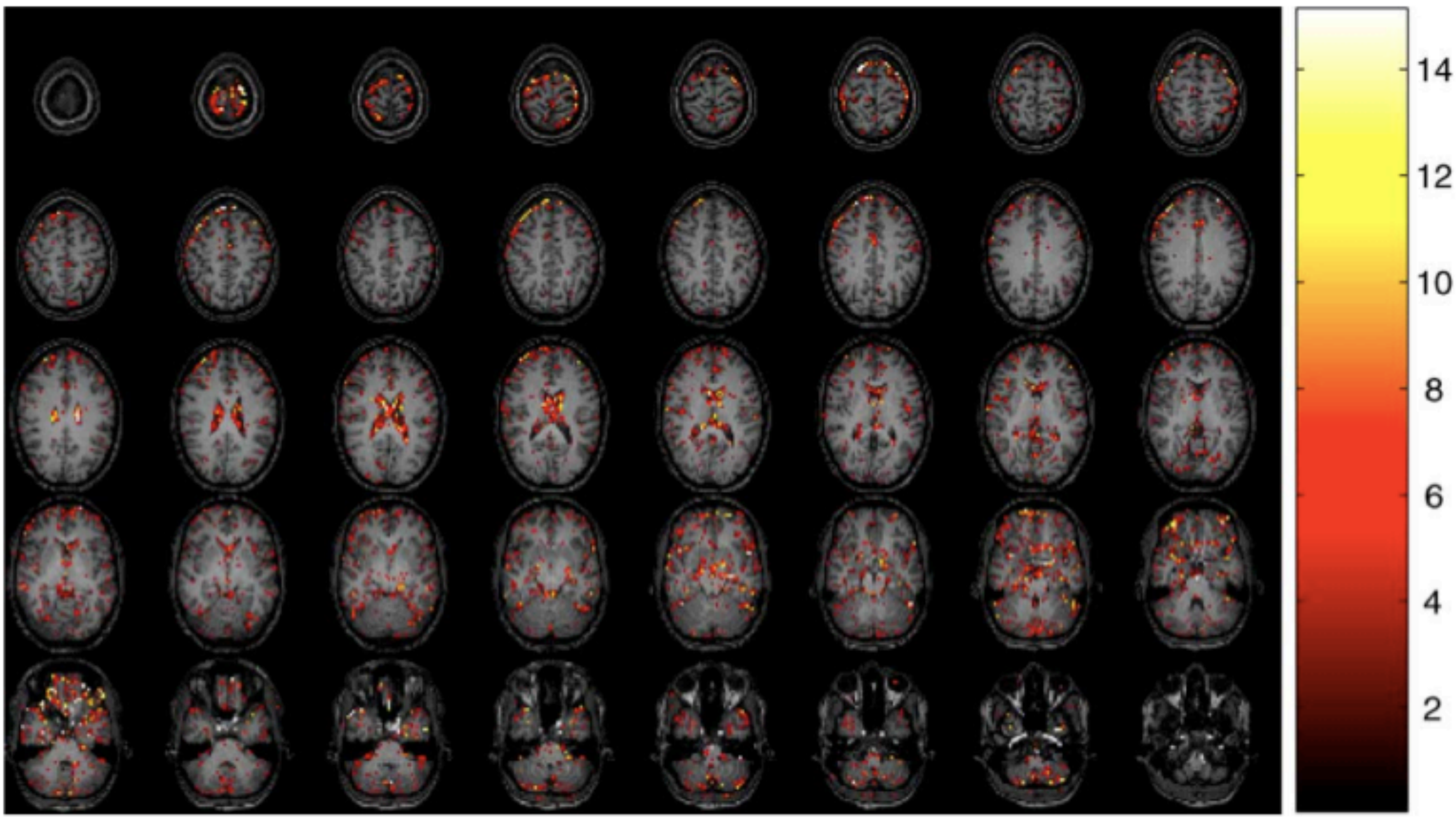
Residual Movement Effects



Cardiac Induced Noise



Respiration Induced Noise



How can we reduce correlated noise/error?

How can we reduce correlated noise/error?

“Color” the data with pink noise

Temporally smooth the data at a known frequency that dominates the correlated noise

Estimate the GLM with weighted least squares

$\underline{\beta} = inv(XW X')X'B$

Temporally smoothing data with poor temporal resolution. A bad idea…

How can we reduce correlated noise/error?

Estimate the degree of error covariance and reestimate with weighted least squares

AR(1) estimation on the residuals

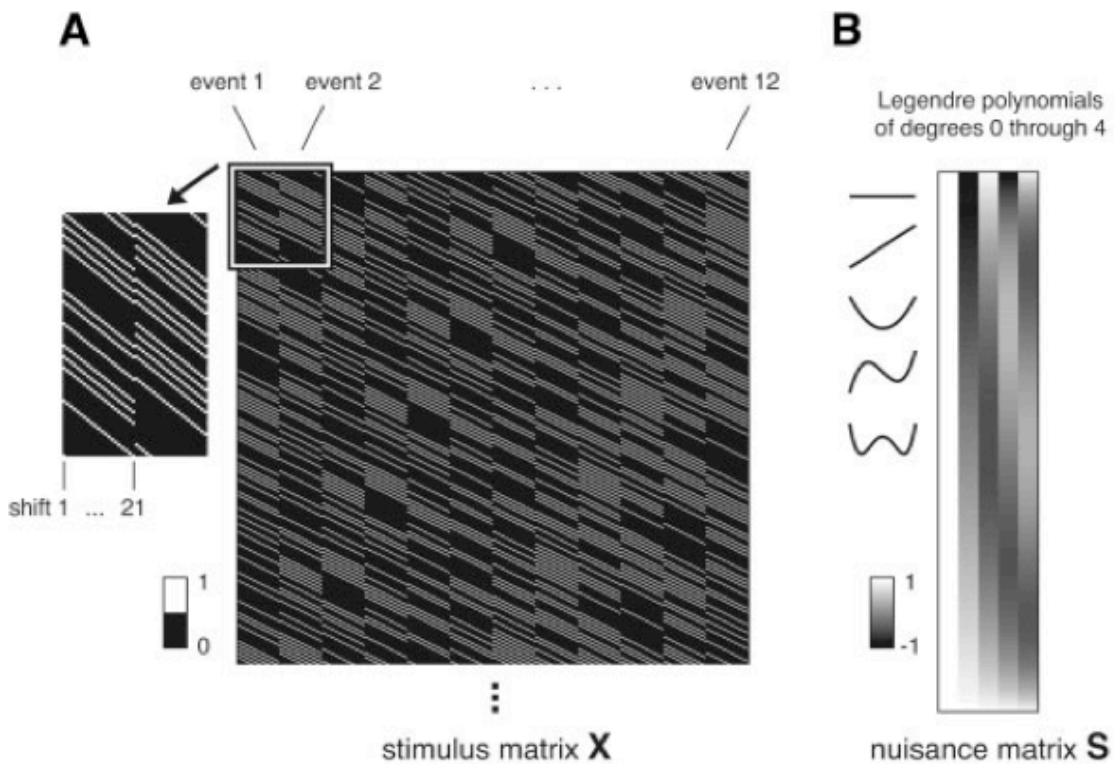
Estimate the GLM with weighted least squares

$\underline{\beta}$ = inv(XW X')X'B

This is in SPM2, and SPM5
FSL "pre-whiten" option does a VERY complicated estimation of the error covariance

How can we reduce correlated noise/error?

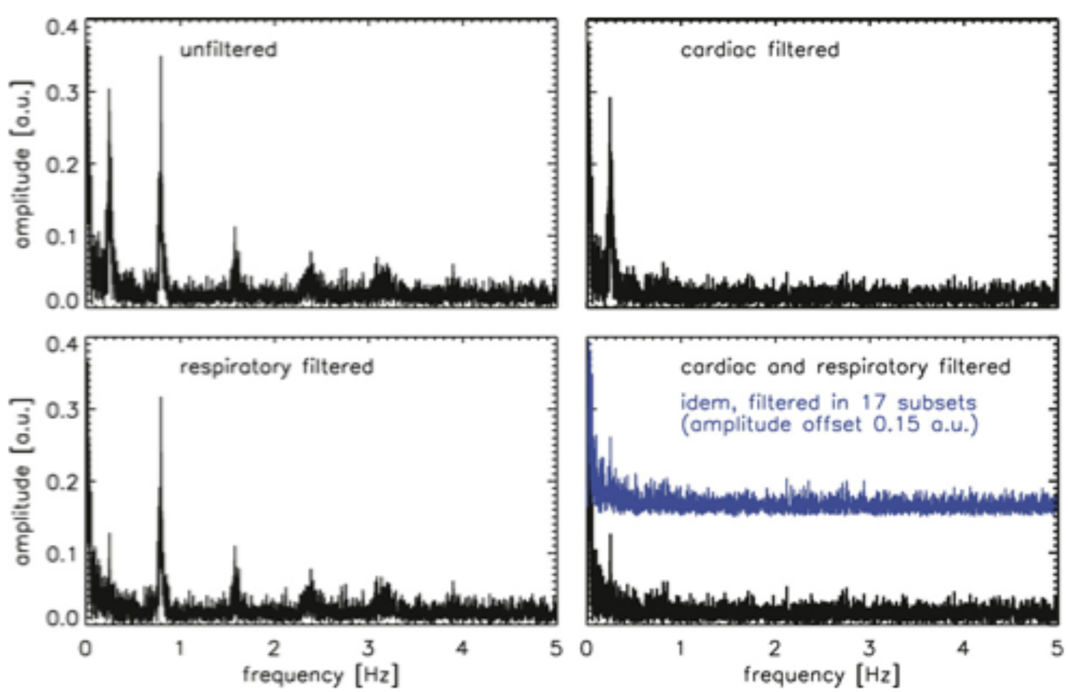
- Make a reasonable guess about noise structure
- Add nuisance parameters to X, with a set of basis functions that cover most of the relevant frequencies:



How can we reduce correlated noise/error?

- Filter Y, the data matrix BEFORE model estimation using physiological information
    HR, RR

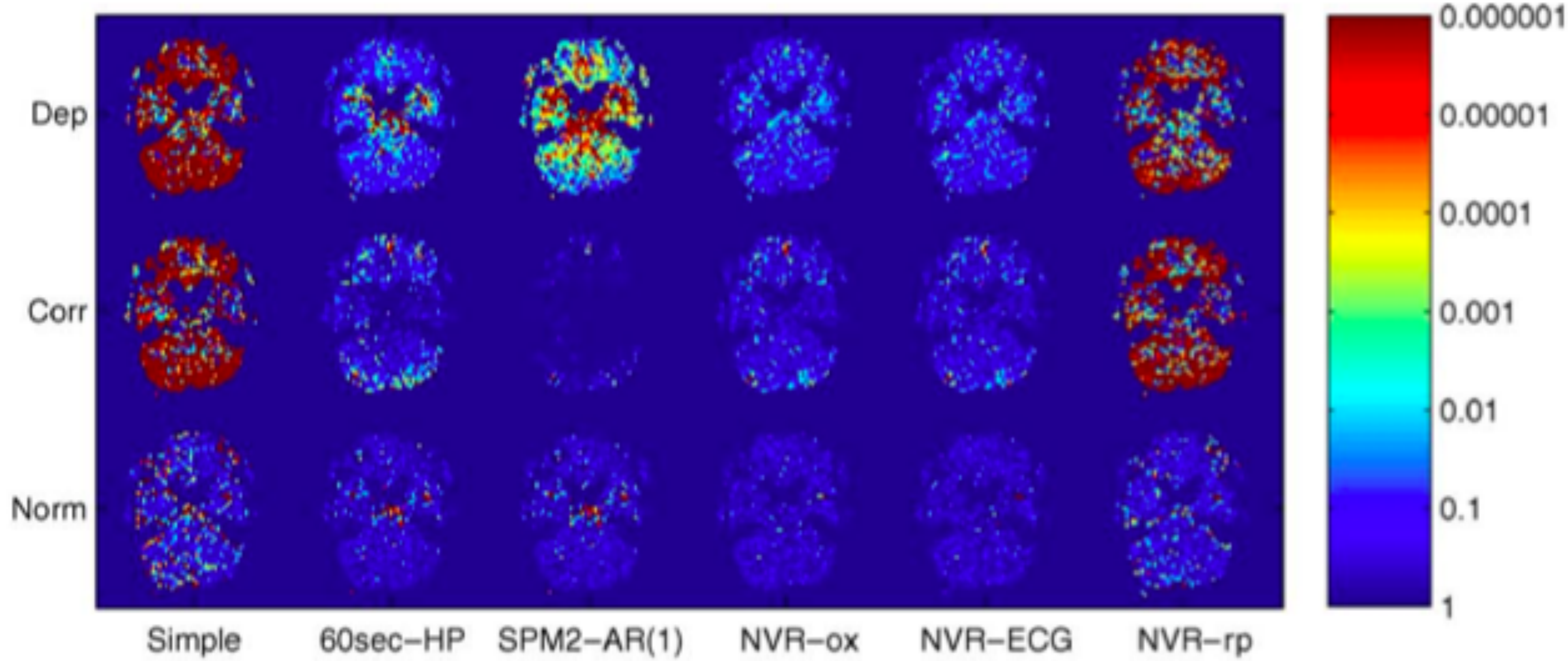
R.H.R. Deckers et al. / NeuroImage 33 (2006) 1072–1081



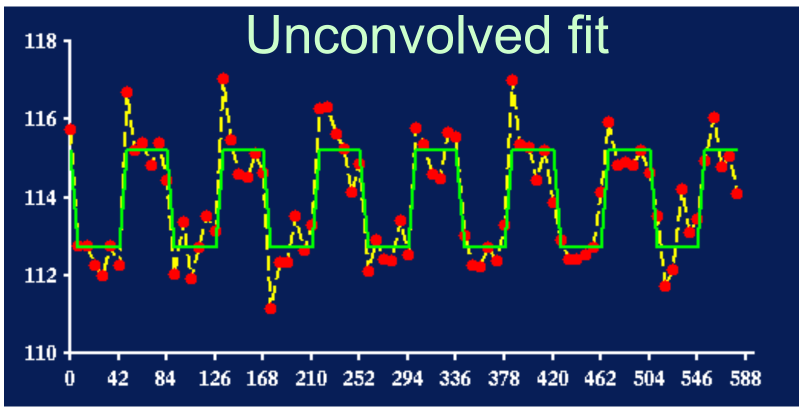
How can we reduce correlated noise/error?

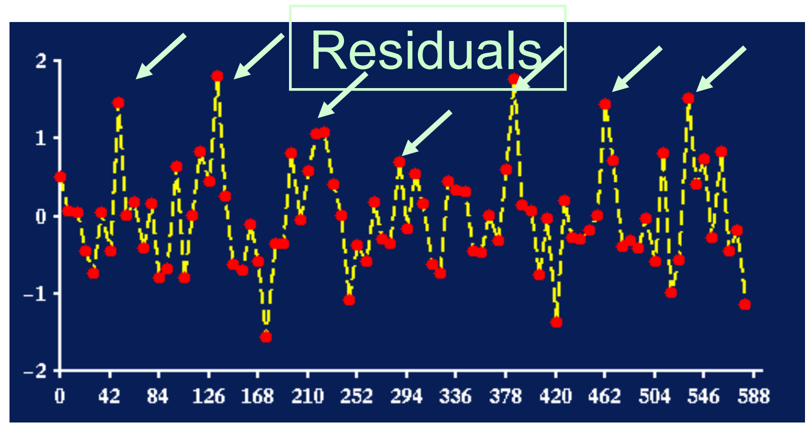
- Add measured physiological parameters to X as nuisance variables

T.E. Lund et al. / NeuroImage xx (2005) xxx–xxx



Last but not least: The hemodynamic response function

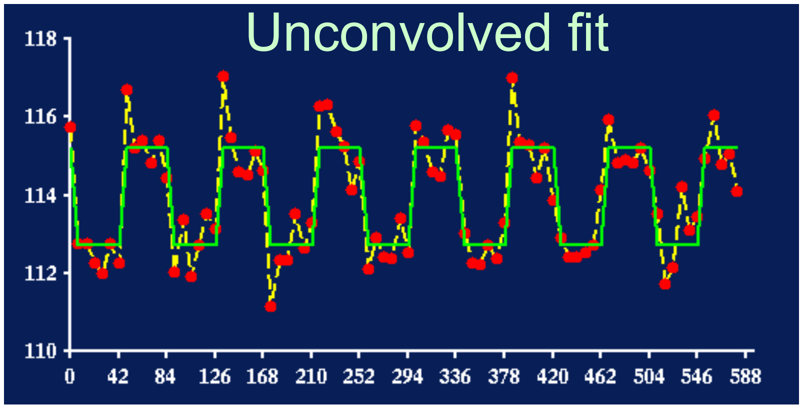


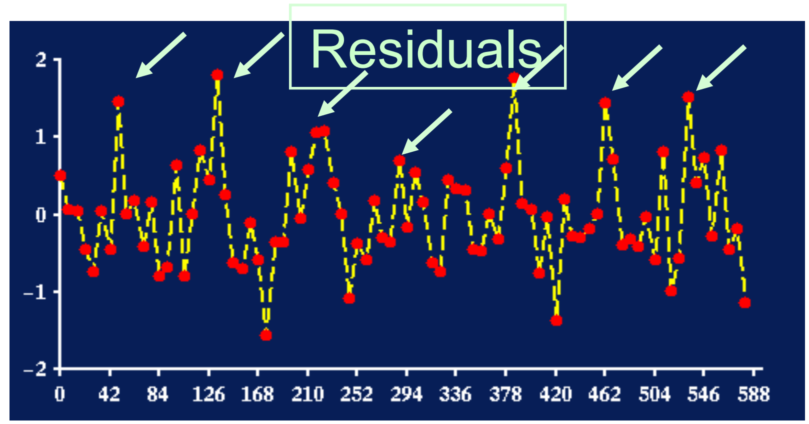


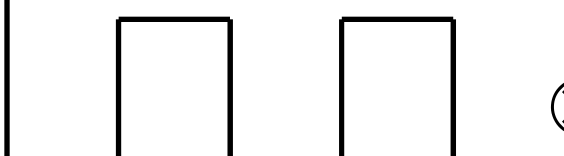
How can we account for it?

Hint: 3 “solutions”

Solution 1: Assume a single shape and convolve with X

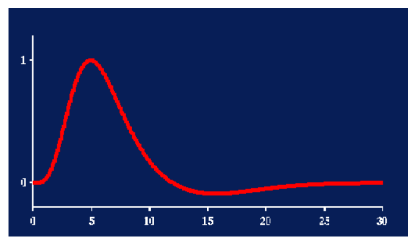






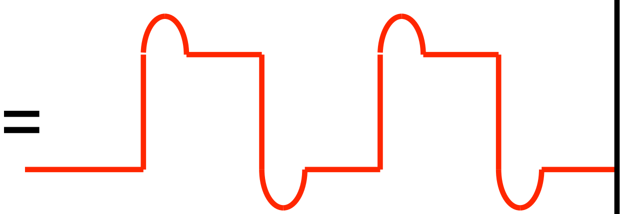
Boxcar function

$\otimes$

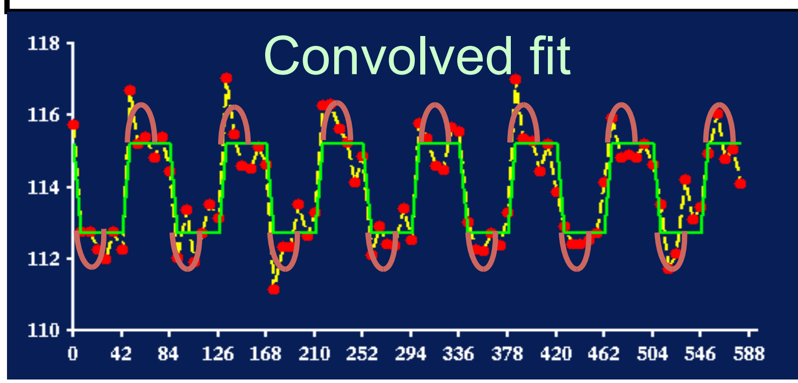


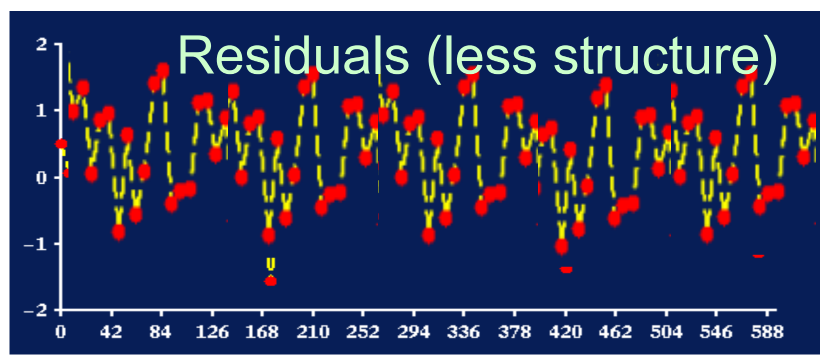
hemodynamic response

=



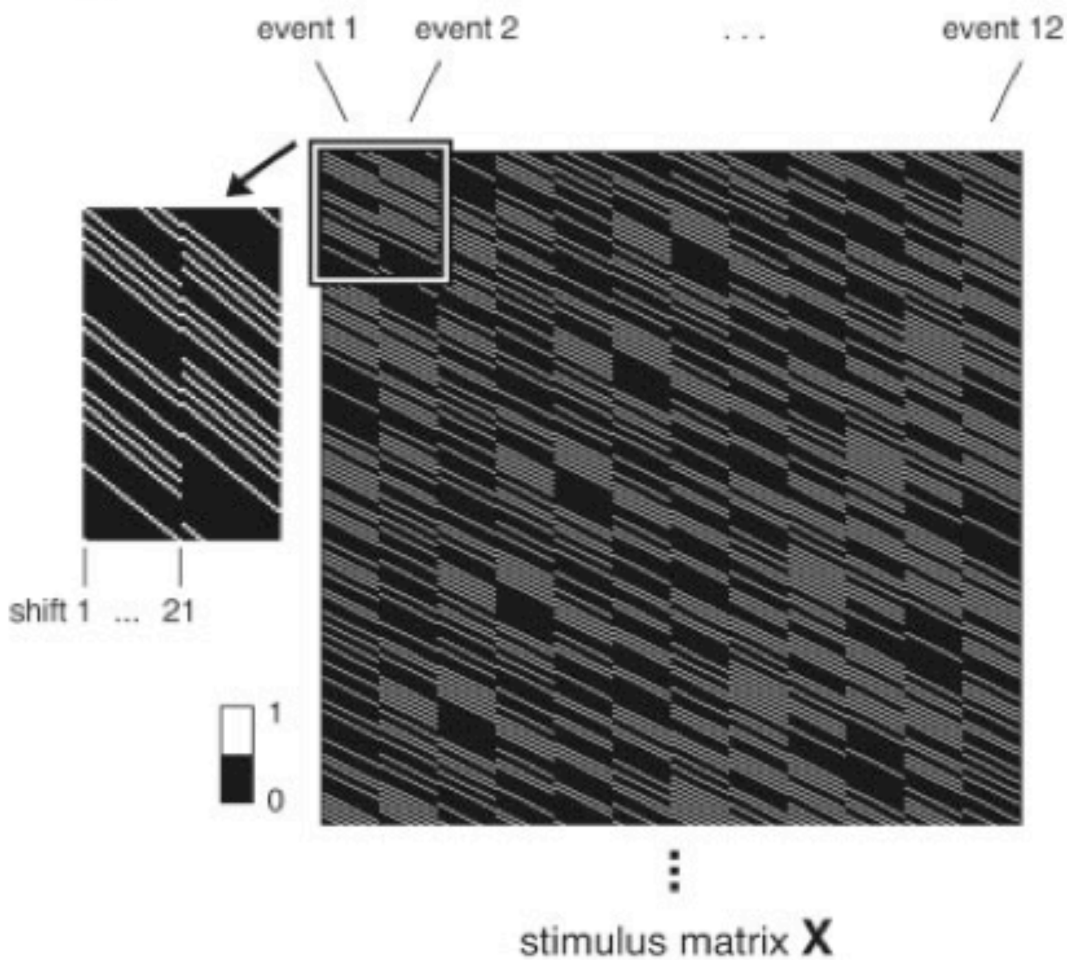
convolved with HRF





Solution 2: Estimate the HRF with no prior assumptions about its shape

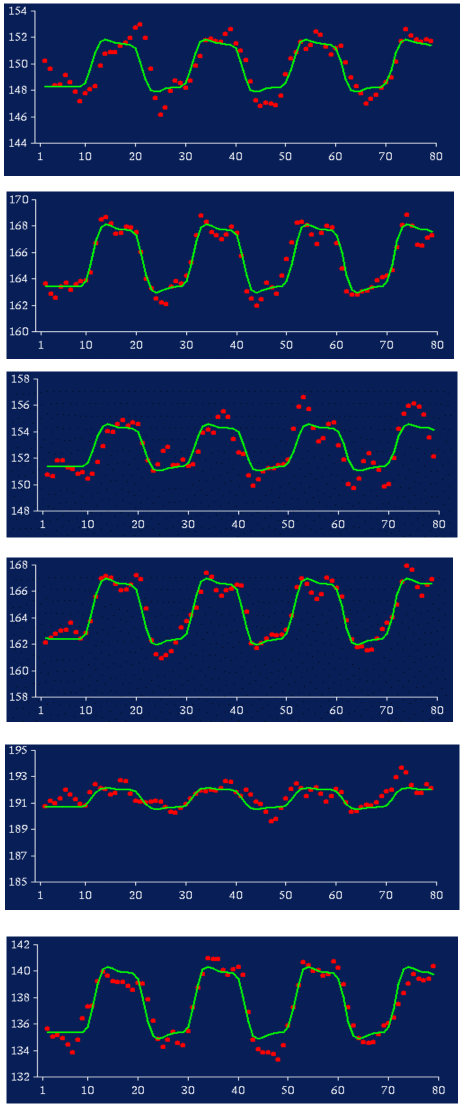
Estimate the Finite BOLD Response or FBR (not FIR)



Solution 3: Assume a family of shapes and estimate with basis functions (such as a family of gamma functions)

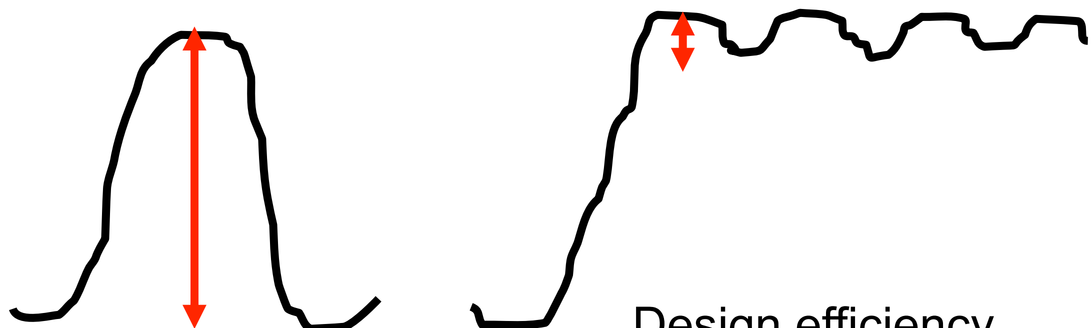
All the methods assume similar shape over all trials of a condition. Correlation assumes same shape over conditions.

T-tests 1st-level (within-subject)



Beta values can be related to size of an effect.

The ‘goodness of fit’ or error term is the residual term and has the same value at a given voxel regardless of which beta(s) is/are being used to create a contrast



Design efficiency

Estimating effect size

Consider an experiment with two explanatory variables (θ), a constant term β, and a linear gradient to account for scanner drift Δ.

The null hypothesis $H_0 : \theta_1 = 0$ against the alternate $H_1 : \theta_1 > 0$

Can be written as:

$$H_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \beta_0 \\ \Delta \end{bmatrix} = 0$$

And the contrast vector is:

$$\underline{c}' = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}$$

Estimate effect size:

$$t = \frac{\underline{c}'\hat{\underline{\beta}}}{\sqrt{Var(c'\hat{\underline{\beta}})}}$$

$$df = \#TRs - \#\underline{\beta}s$$

But how do we calculate $Var(c'\hat{\underline{\beta}})$?

Recall, $\hat{\underline{\beta}}$ = inv(X X')X'B

$$\begin{aligned}
Var(\underline{c}'\hat{\underline{\beta}}) &= Var[\underline{c}'(X'X)^{-1}X'\underline{B}] \\
&= [\underline{c}'(X'X)^{-1}X']\Sigma_{\varepsilon}[\underline{c}'(X'X)^{-1}X']' \\
&= \sigma_{\varepsilon}^{2}\underline{c}'(X'X)^{-1}X'X(X'X)^{-1}\underline{c} \\
&= \sigma_{\varepsilon}^{2}\underline{c}'(X'X)^{-1}\underline{c}
\end{aligned}$$

$$t = \frac{\underline{c}'\hat{\underline{\beta}}}{\sqrt{\hat{\sigma}_{\varepsilon}^{2}\underline{c}'(X'X)^{-1}\underline{c}}}$$

Are two explanatory variables significantly different?

$$\underline{c}' = \begin{bmatrix} 1 & -1 & 0 & 0 \end{bmatrix}$$

Estimation:

$$t = \frac{\underline{c}'\hat{\underline{\beta}}}{\sqrt{\hat{\sigma}_{\varepsilon}^{2}\underline{c}'(X'X)^{-1}\underline{c}}}$$

$$df = \#TRs - \#\underline{\beta}s$$

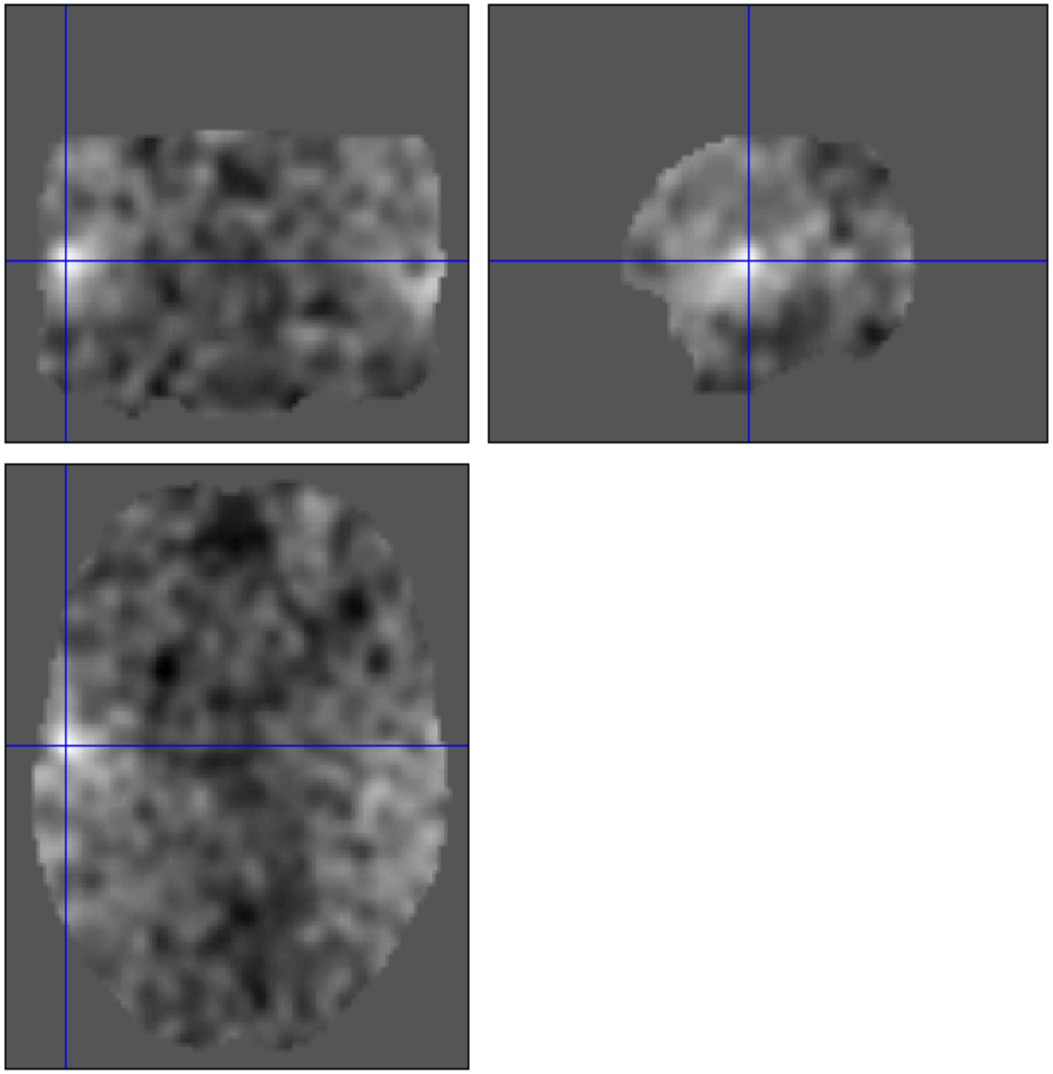
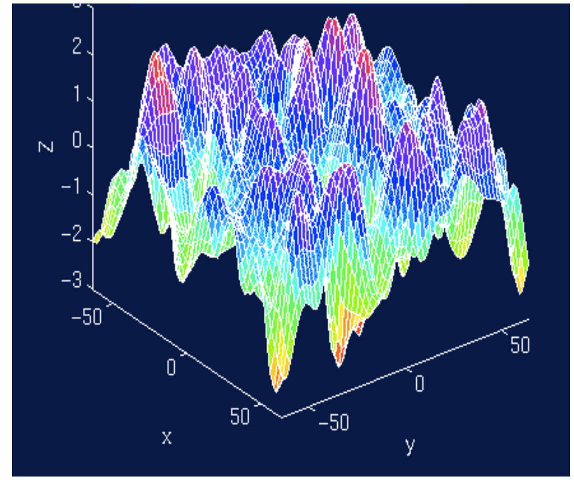
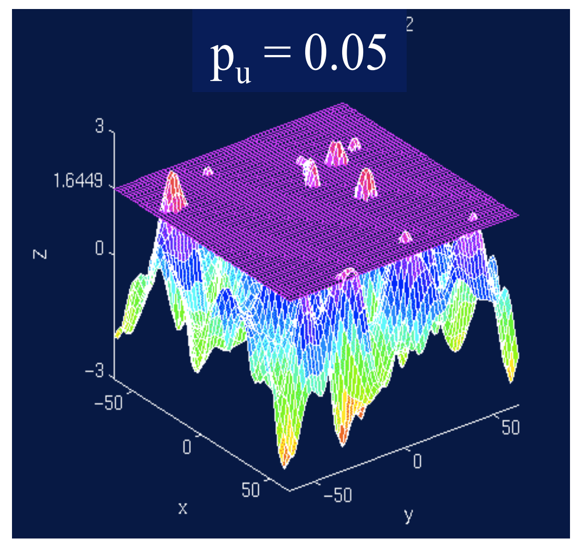


Image of t-values or z-values

Thresholding a statistical image





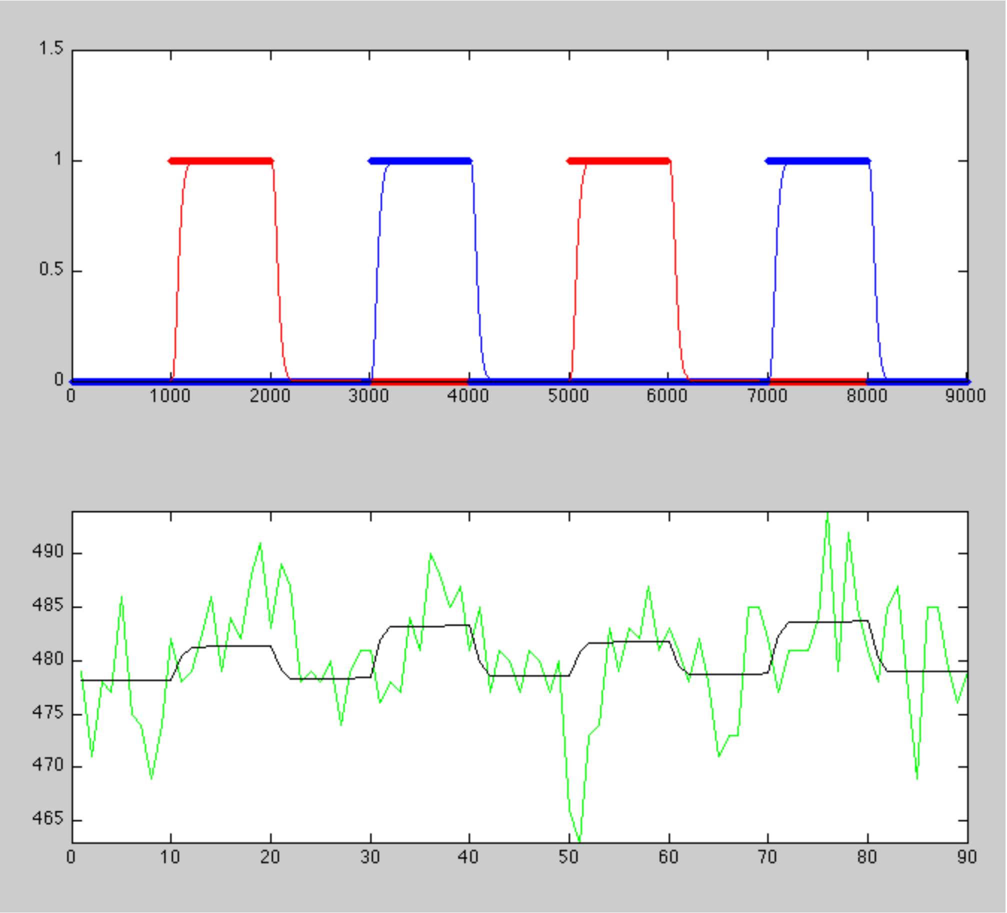
Interim Summary

- beta images contain information about the size of the effect of interest.
- Information about the error variance is held in a separate imagr (SPM: ResMS.img, FSL:)
- beta images are linearly combined to produce relevant contrast images.
- The design matrix, contrast, constant and error.img are subjected to matrix multiplication to produce an estimate of the standard dev associated with each voxel in the contrast image.
- A t image is derived from this and thresholded in the results step.

What constitutes a good design, X?

Depends on what we want to know.

Hint: At least 5 plausible answers



What constitutes a good design, X?

Depends on what we want to know.

1. Appropriate psychological construct. Do we need an unpredictable paradigm?
   Block design bad (low entropy), Event related design good (high entropy)

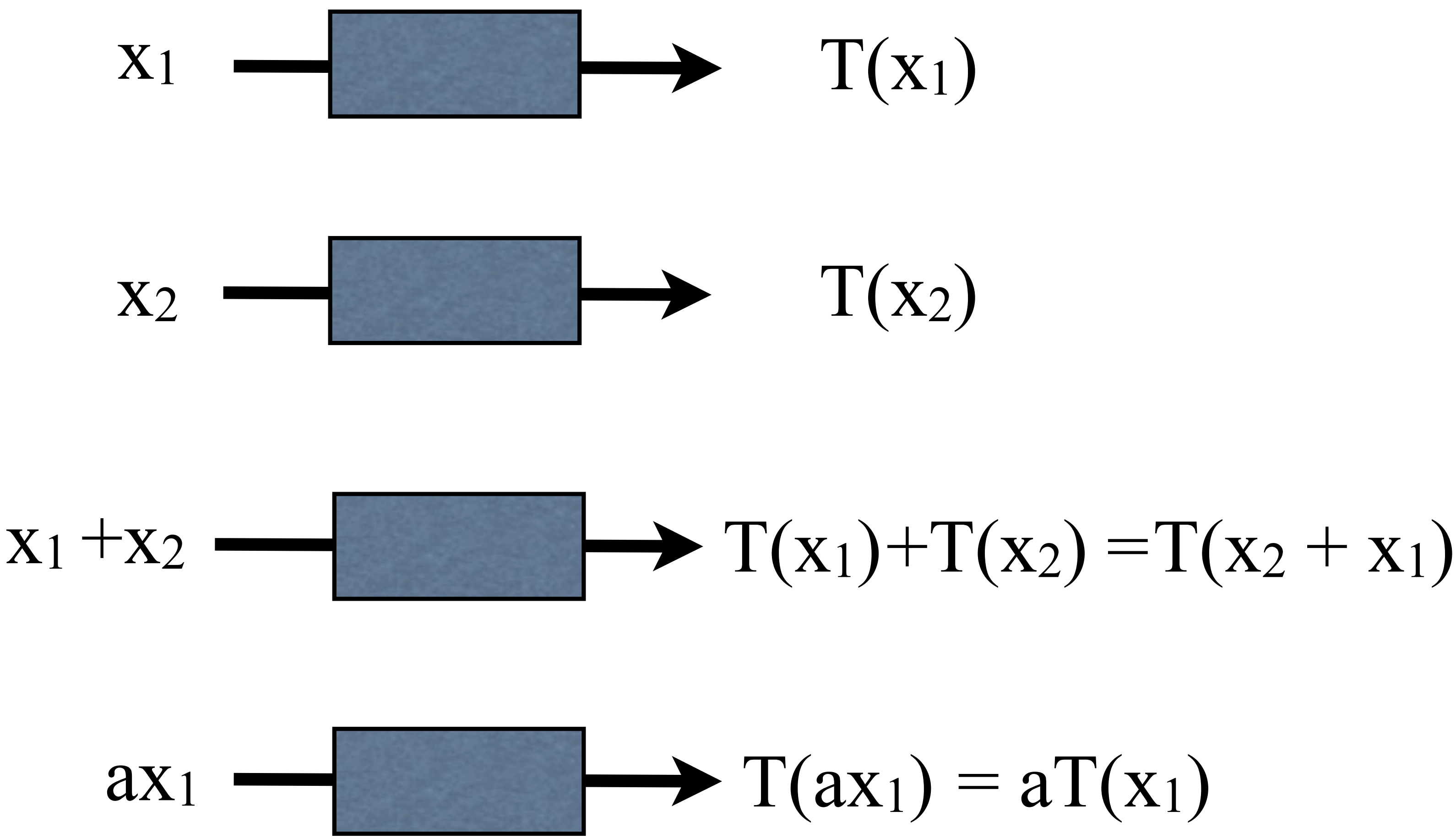
2. Do we want to know the shape of the HRF?
   Then we want FBR or basis function estimation. Jittered slow events

3. De we want to estimate the height of the HRF for all conditions? **MVPA**
   Then HRF correlation method is potentially better. Lots of events

4. Do we want to maximize the POWER to detect a difference between conditions?
   Then block design is the best

5. Do we have a finite amount of time? Maximize(Trials/hour)
   Fast rather than slow event related design

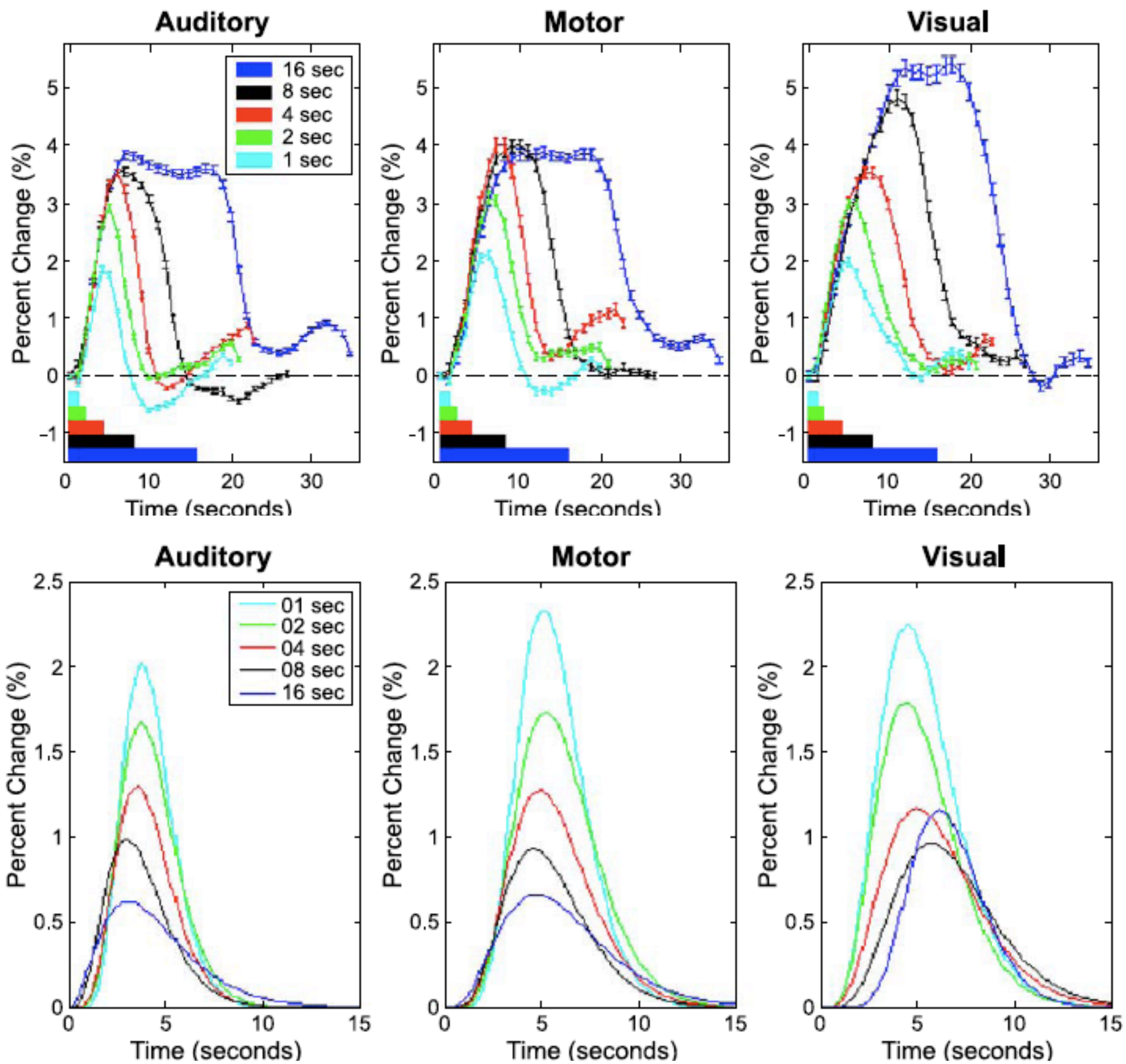
Linear Time Invariant Systems



Comparison of hemodynamic response nonlinearity across primary cortical areas

NeuroImage 22 (2004) 1117–1127

David A. Soltysik,[a,*] Kyung K. Peck,[b,1] Keith D. White,[c]
Bruce Crosson,[d] and Richard W. Briggs[b,2]

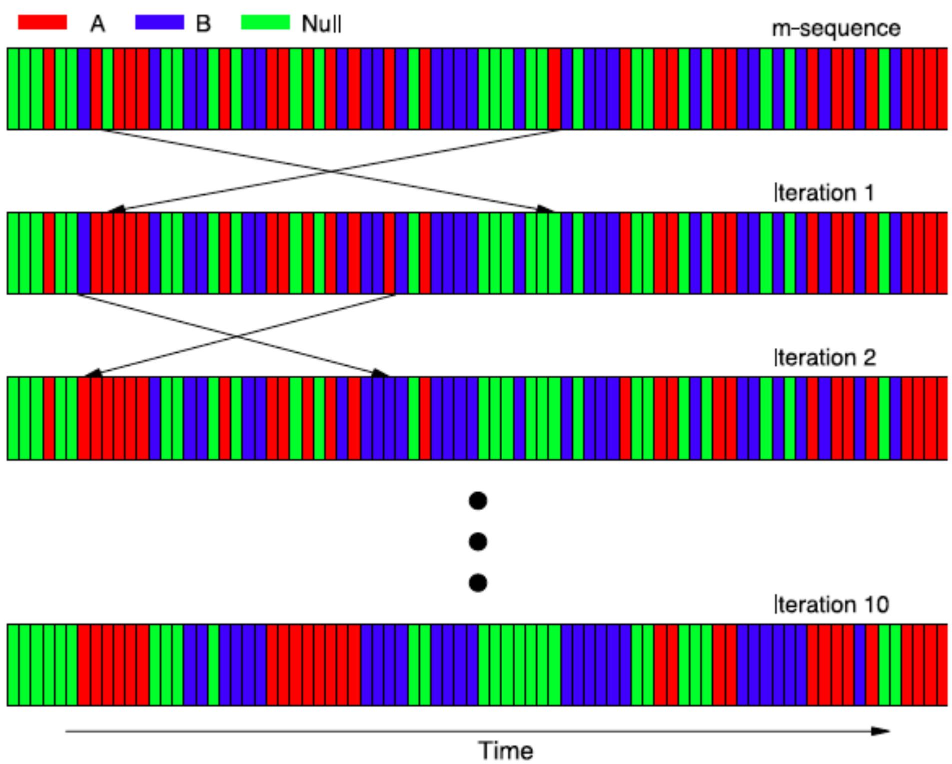


M-sequences … Blocked Design
Maximum length sequences are spectrally “flat”

Of all the "runs" in the sequence of each type (i.e. runs consisting of "1"s and runs consisting of "0"s):

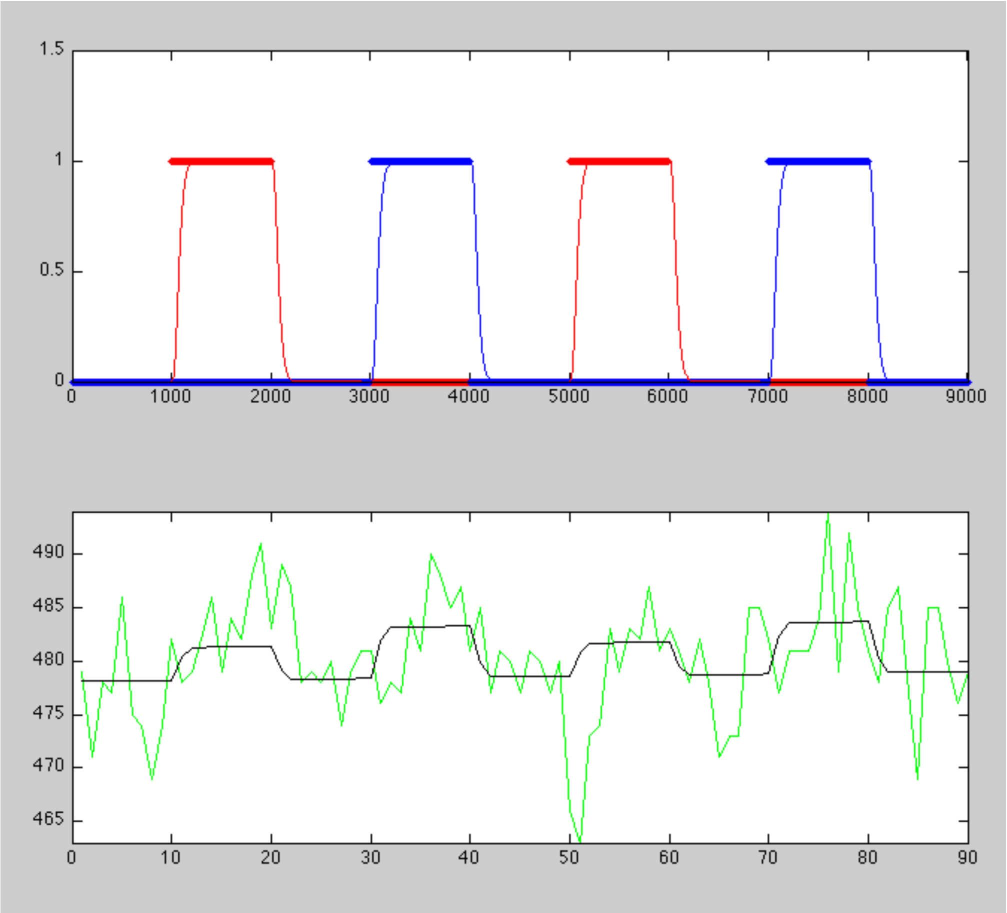
One half of the runs are of length 1.
One quarter of the runs are of length 2.
One eighth of the runs are of length 3.
... etc. ...

T.T. Liu / NeuroImage 21 (2004) 401–413

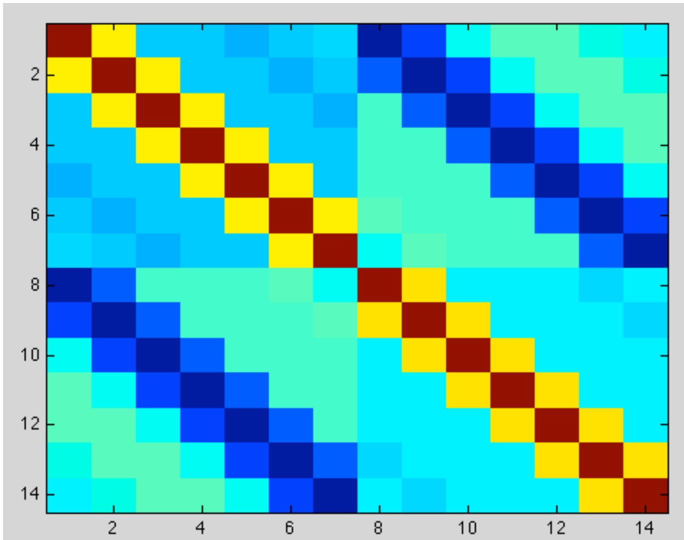


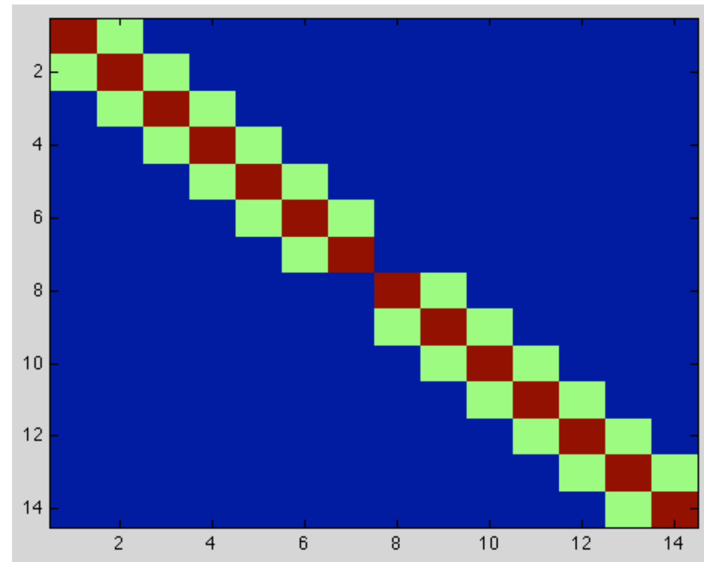
Efficiency for the HRF correlation method

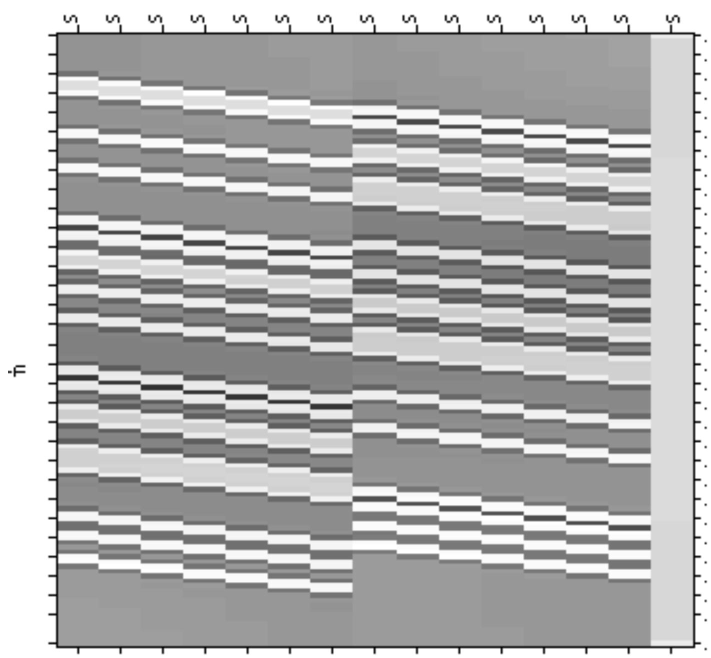
$$\xi = \frac{h_0^T X^T X h_0}{\mathrm{var}(\hat{\beta})}$$

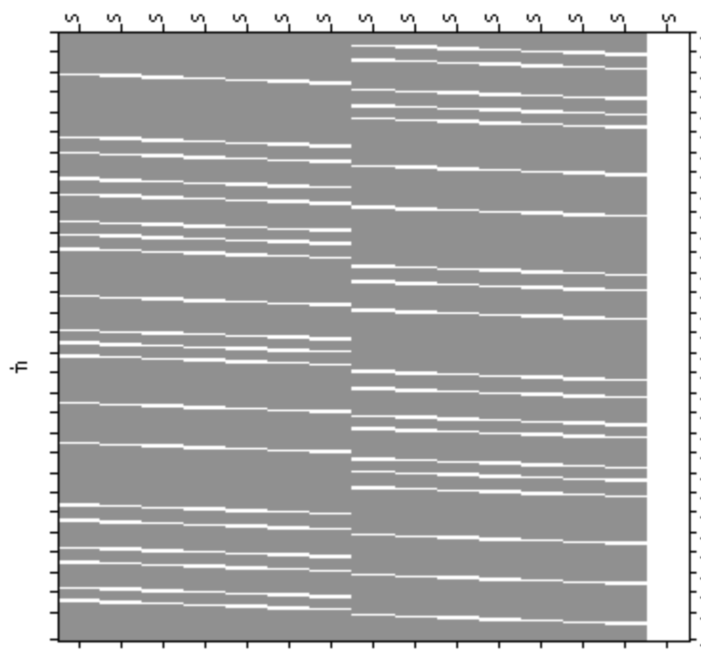


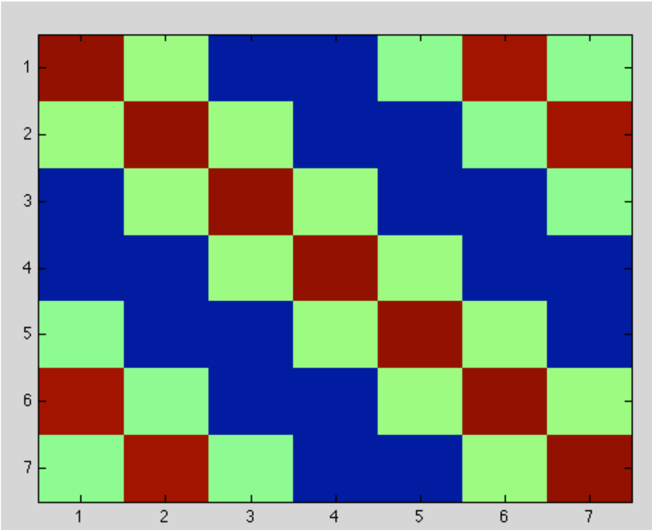
Rapid Q=2, k=7 Slow

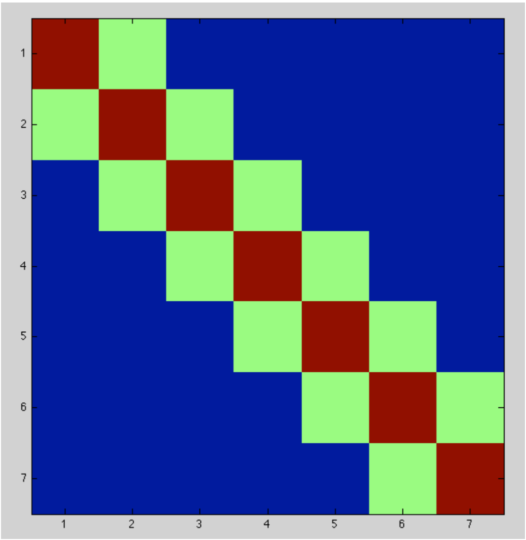


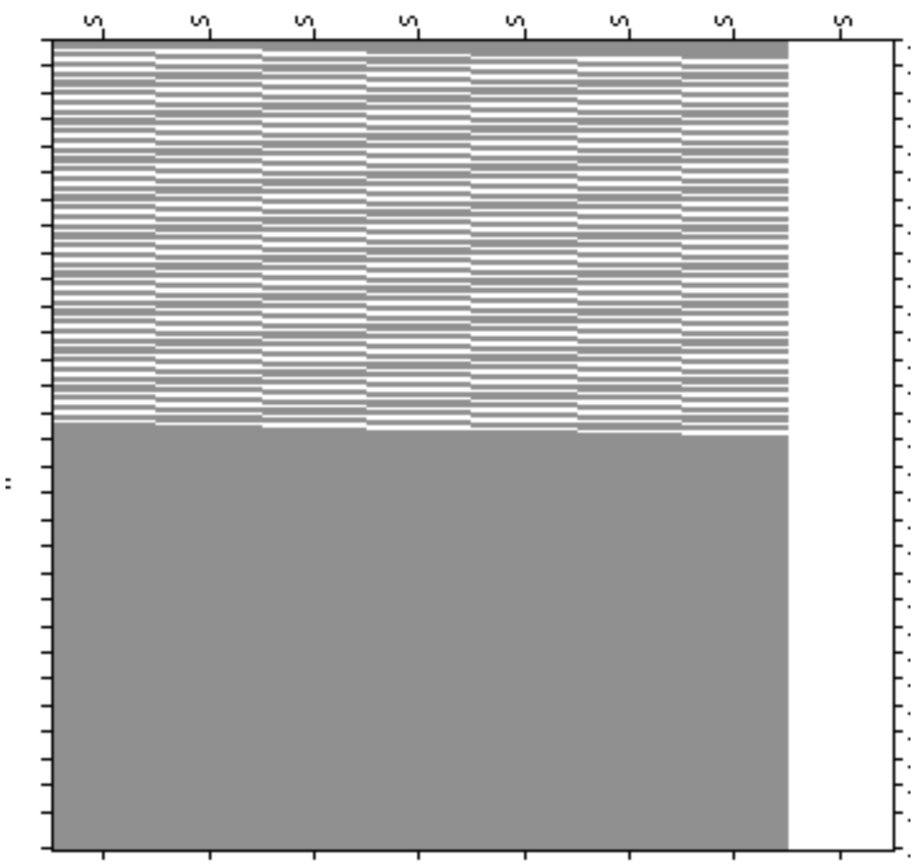


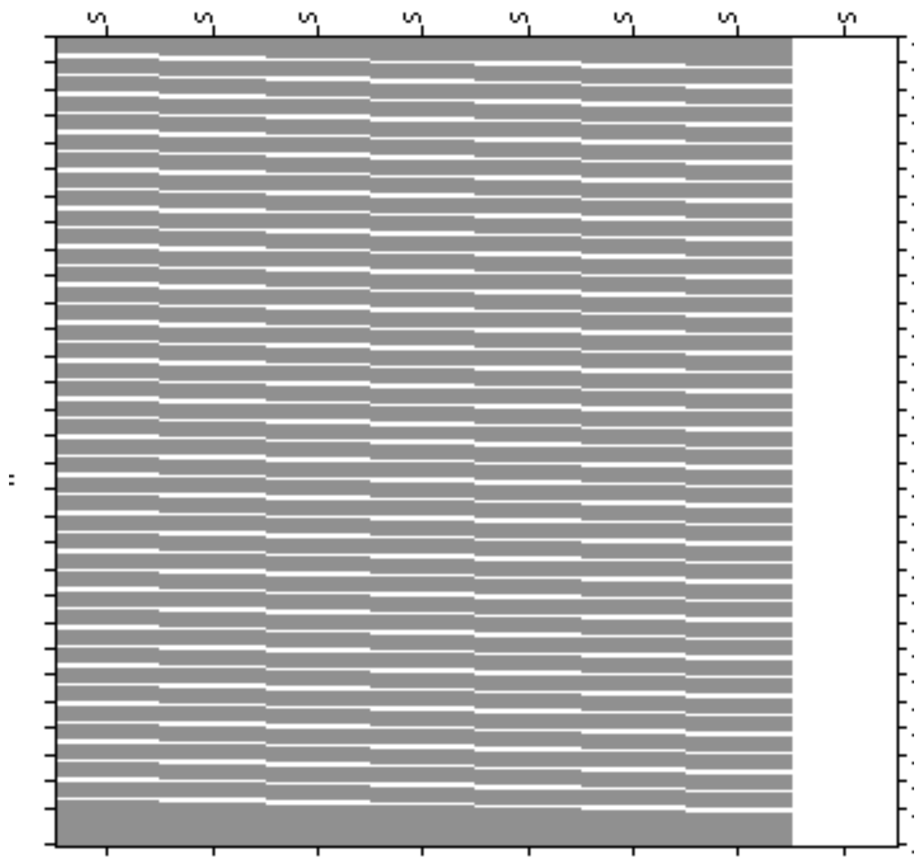





1 2 3 4 5 6 7
1 2 3 4 5 6 7


1 2 3 4 5 6 7
1 2 3 4 5 6 7





Rapid, Q=1, k=7