From this morning, you should have learned …

- What is BOLD signal?

- How should we design an fMRI experiment?

In this lecture, you will learn …

- How to analyze fMRI data (the standard way): General Linear Modeling (GLM). What is its theoretical basis? Why is it useful?

After this lecture, you will learn …

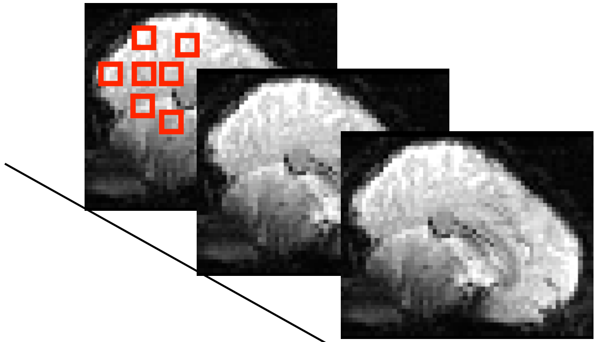
- How to implement a GLM in a standard analysis software (SPM)

General linear model: basic

- Introducing General Linear Model (GLM): Start with an example
- Properties of the BOLD signal
  - Linear Time-Invariant (LTI) system
  - The hemodynamic response function
- (Briefly) Evaluating efficiency of a design
- (Briefly) Going back to the example

Univariate analysis

- Each voxel in the brain is analyzed *separately*



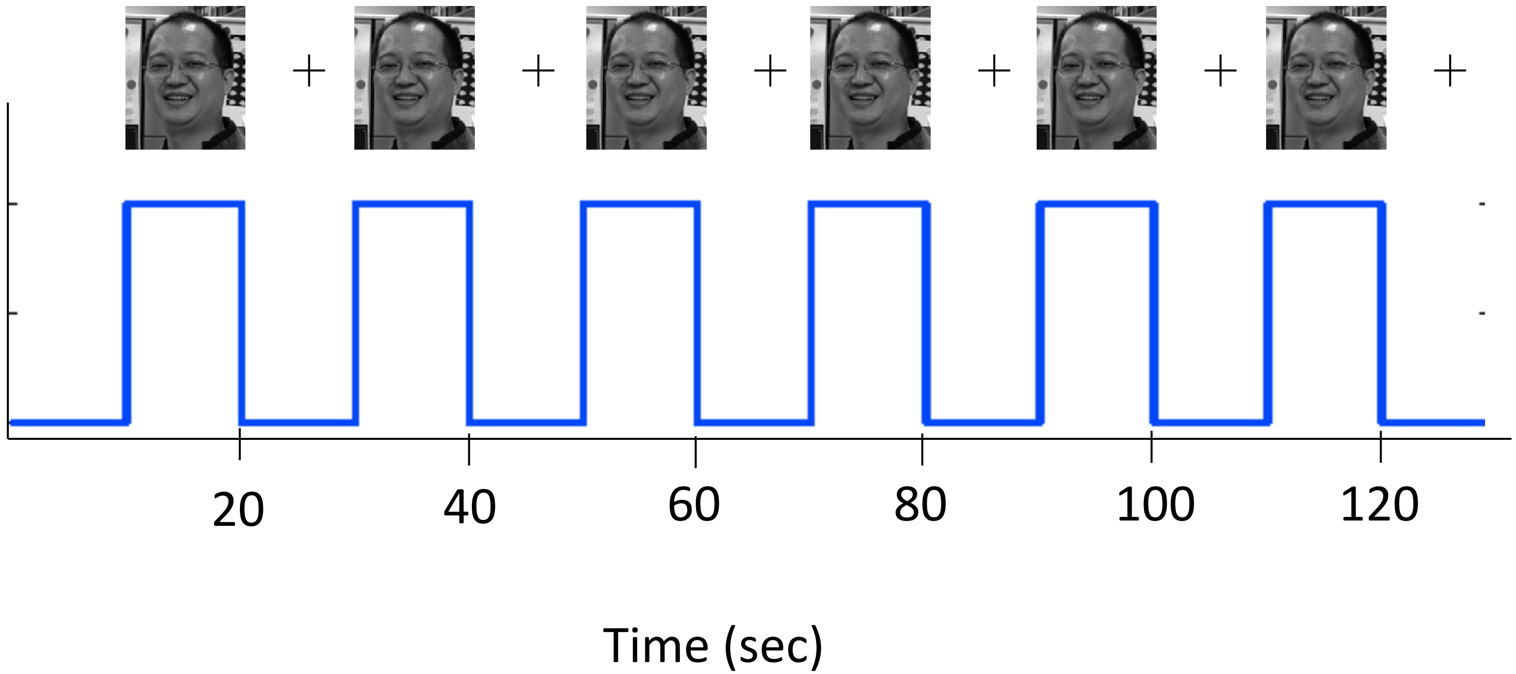
- Each voxel presents a time-series data

...

Time

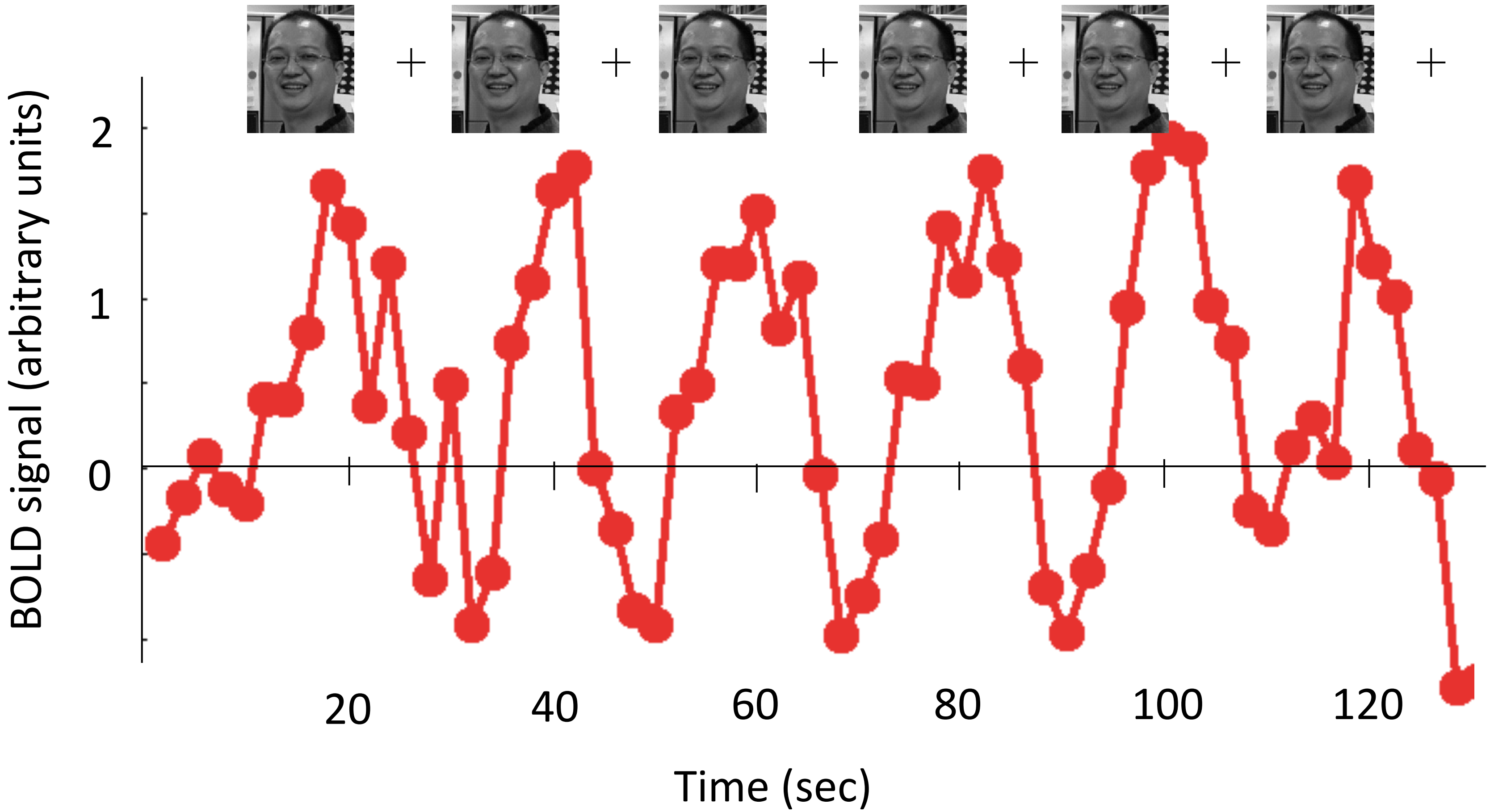
Time-series data

- Suppose you have the following experiment



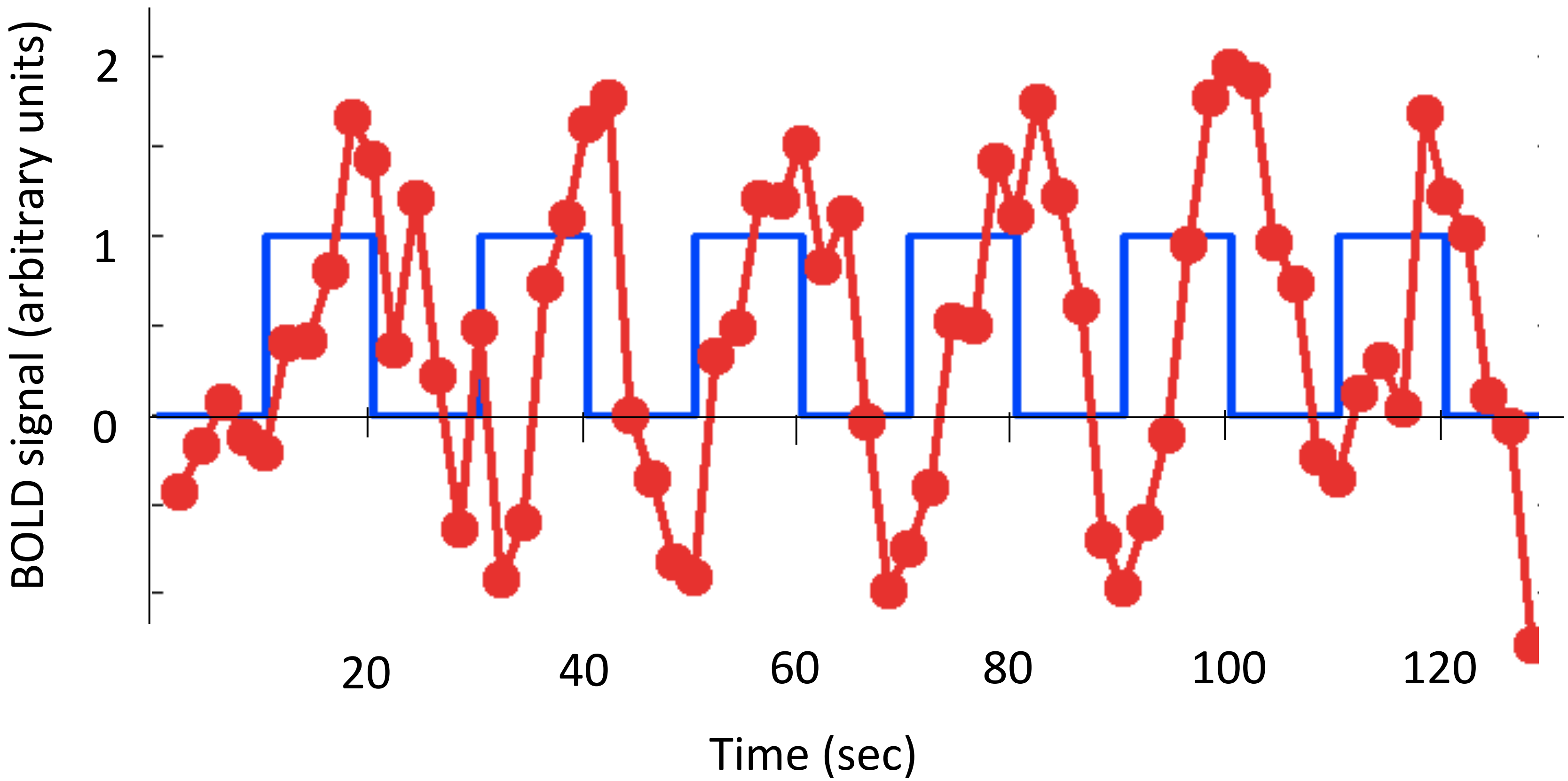
Time-series data

- This is the data you get (from a single voxel)



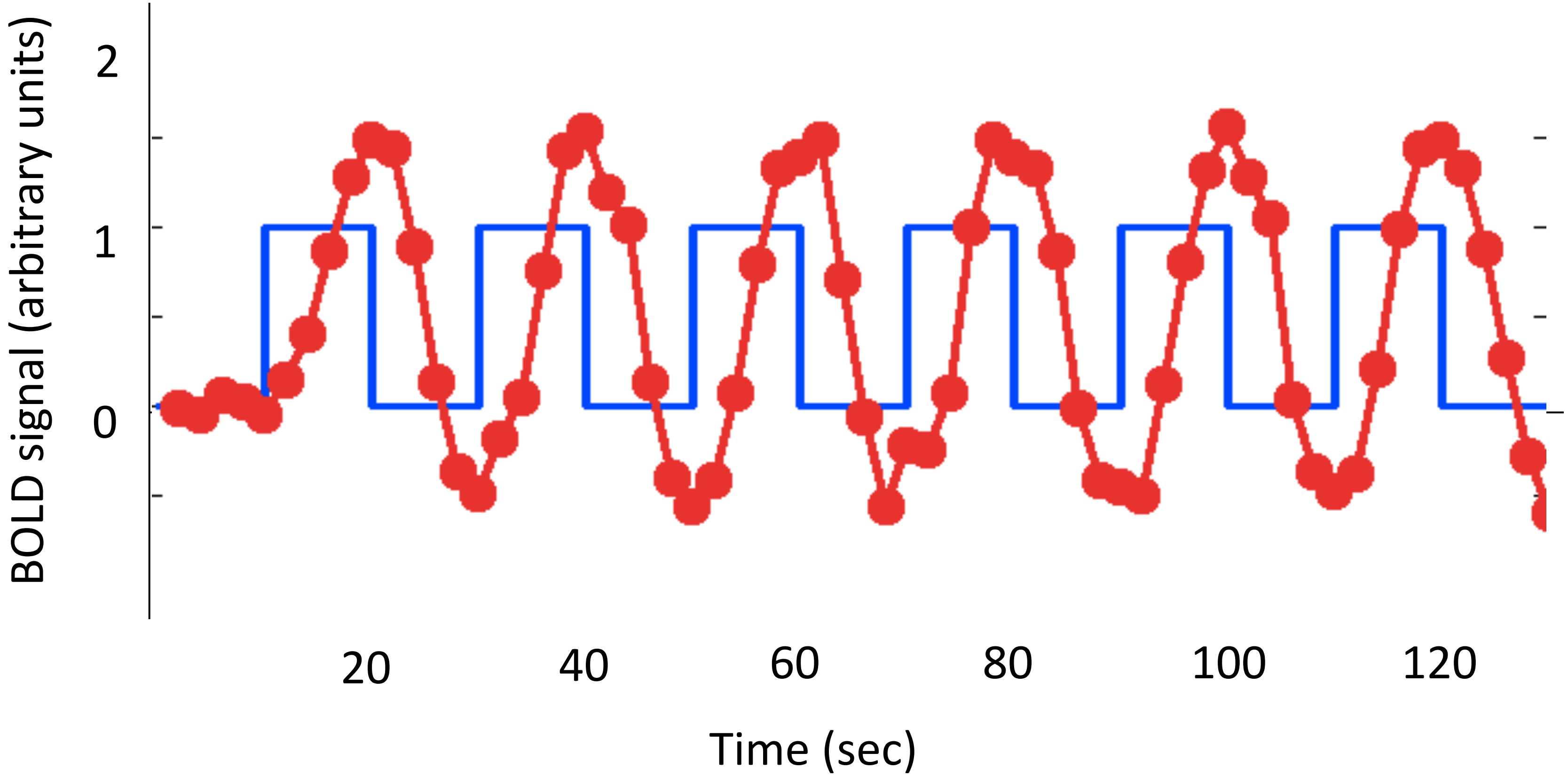
Time-series data

- When you compare prediction (based on your design) and data, you realize that there is somewhat a match, but not close



Time-series data

- What about this one? Which aspect of the comparison is the same, which aspect might be different?



Observation

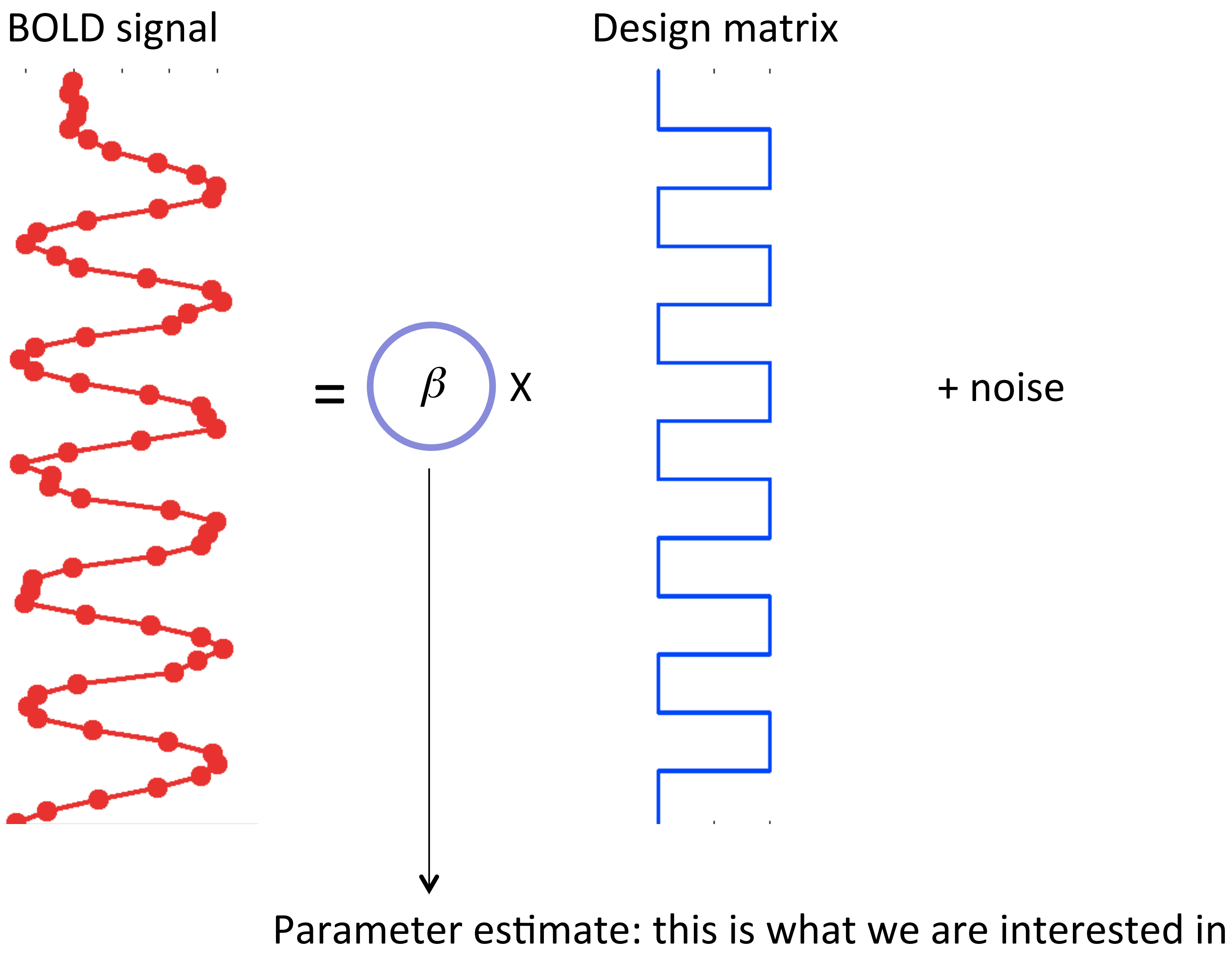
- BOLD signals are noisy

- BOLD signals are *delayed* responses to the events of interest (e.g. the presentation of face)

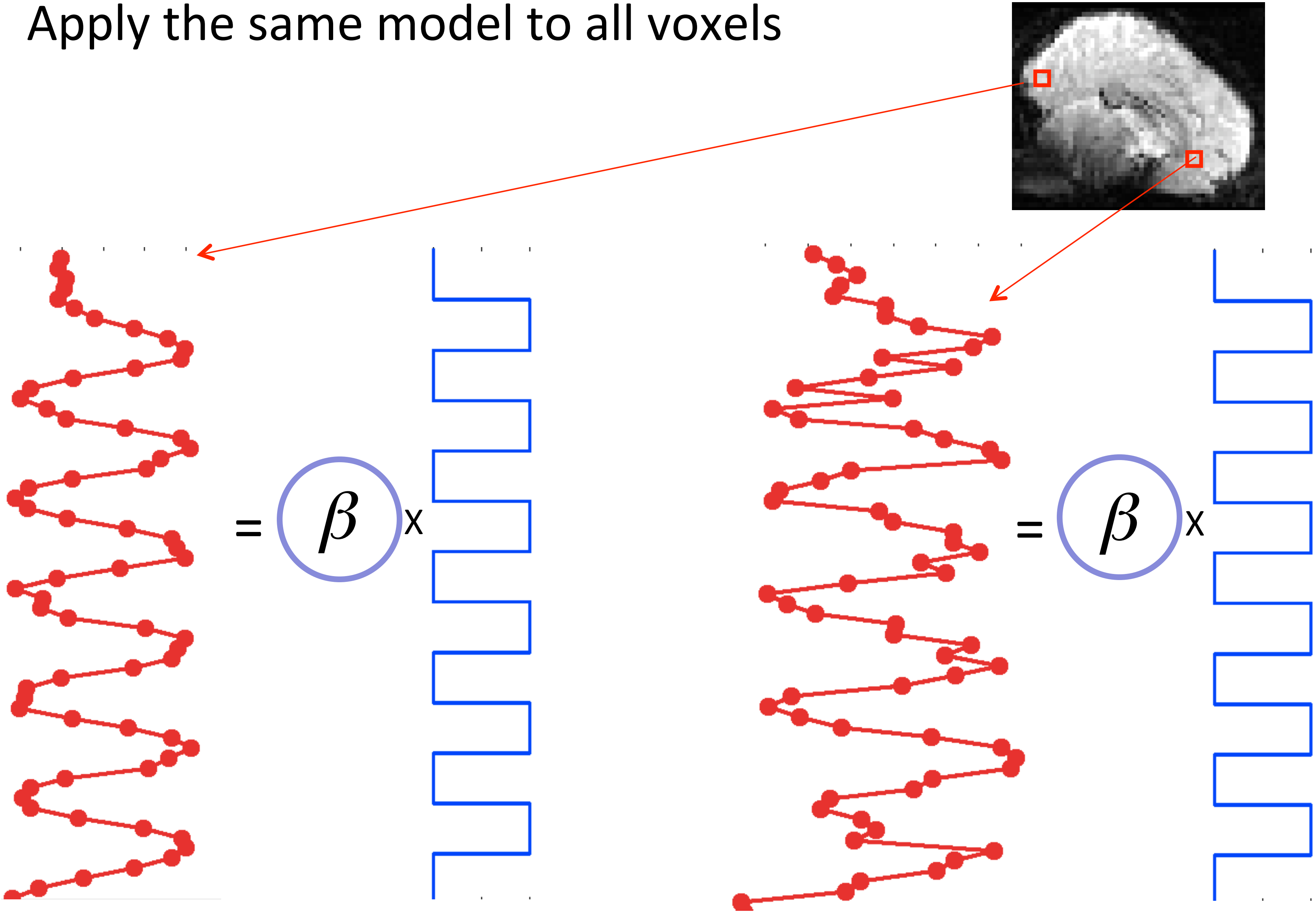
Question

How do we set up the analysis? How do we model the BOLD response?

General Linear Model

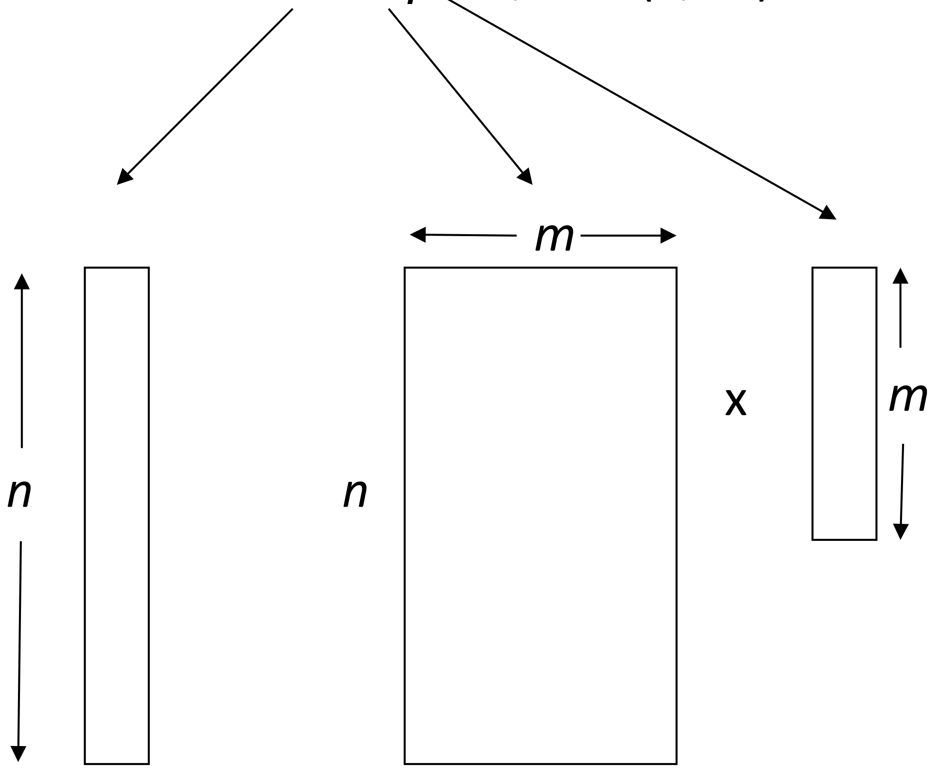


Apply the same model to all voxels



General linear model

$$Y = X\beta + \varepsilon, \varepsilon \sim N(0, \sigma^2)$$



BOLD times series Design matrix Parameter vector

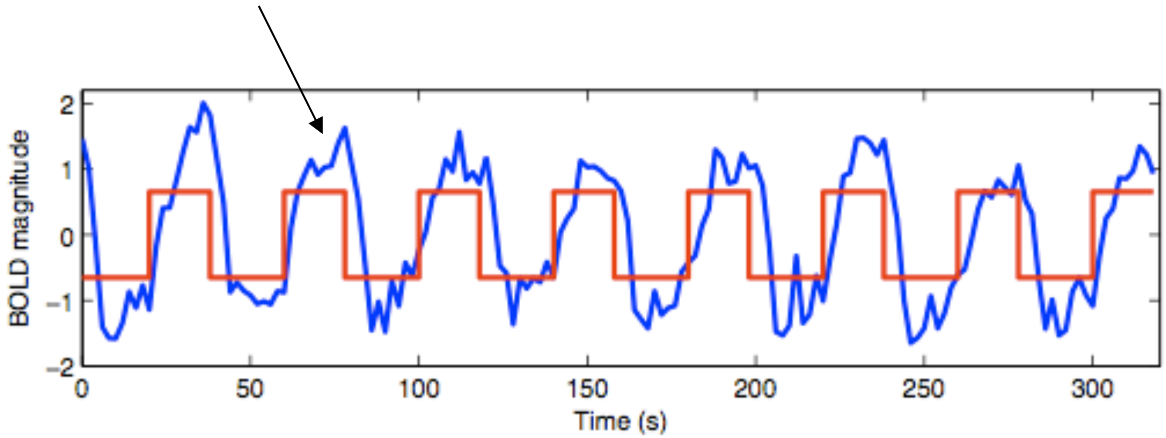
(BOLD: Blood Oxygenation Level Dependent)

Key

- Understanding the relation between BOLD signal and neural activity
- Understanding the properties of the BOLD signal
- Understanding the characteristics of noise in the BOLD signal

The BOLD signal

- Observed BOLD signal (example)



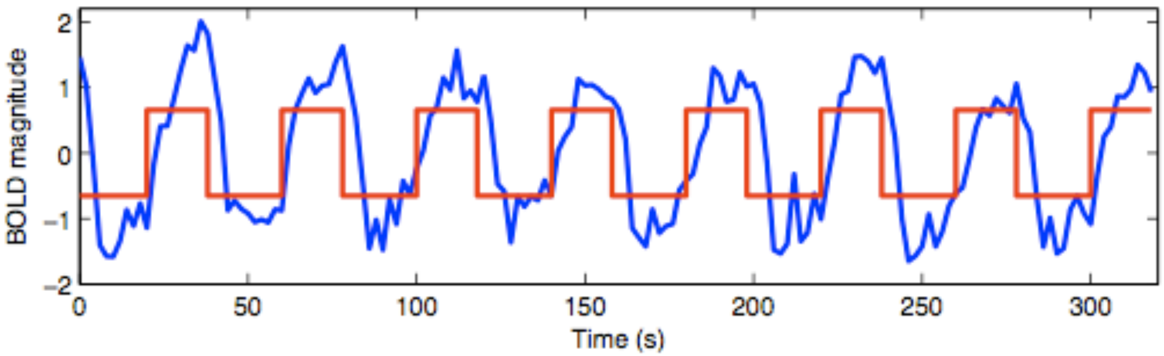
- BOLD signal did not look exactly like the predicted neural activity (in red)

The BOLD signal

- BOLD signal as a transformation of neural activity

$$x \xrightarrow{f(x)} \text{BOLD}$$

(neural activity)



The BOLD signal

- BOLD signal as a transformation of neural activity

$$x \xrightarrow{f(x)} \text{BOLD}$$

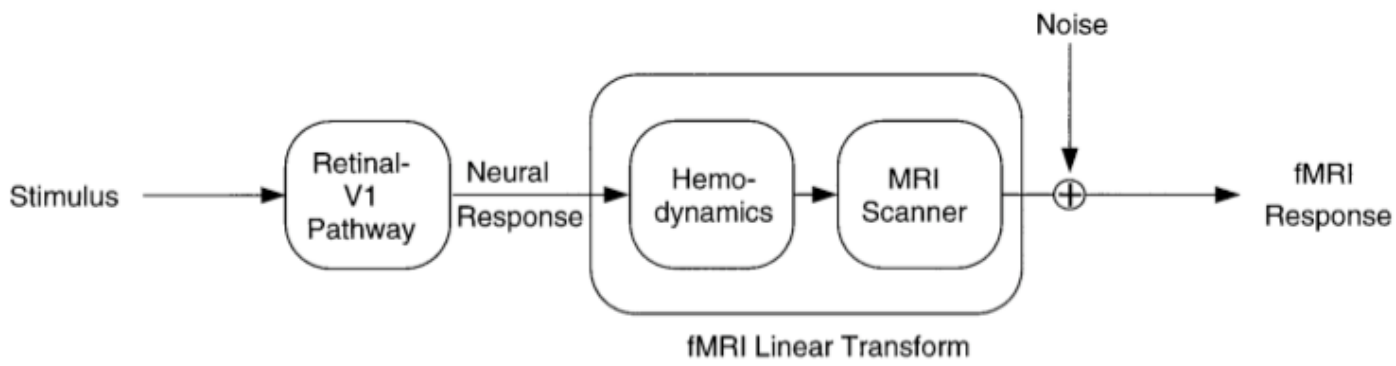
(neural activity)

- Identifying the properties of the transformation function is critical

- The first thing to check is if BOLD is a linear transform of neural activity

Around the mid 1990s …

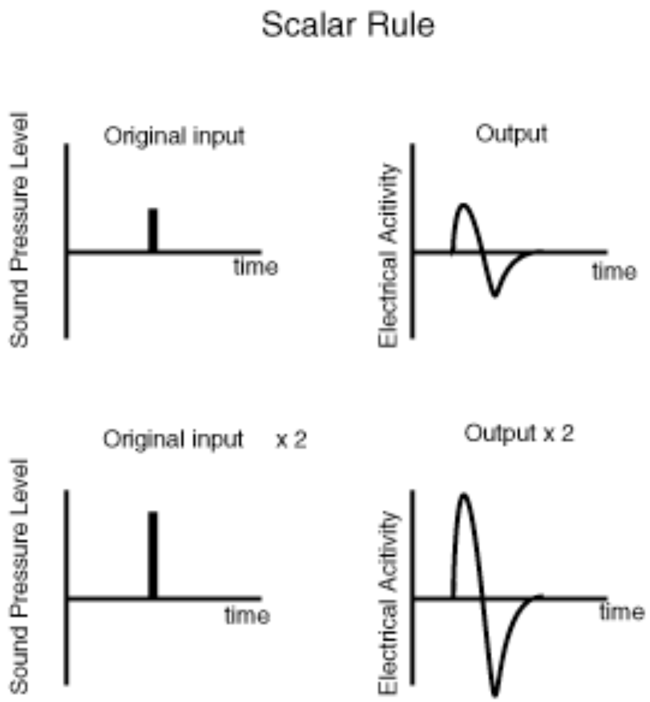
- Boynton et al. (1996, J Neurosci.) tested the linear transform model in primary visual cortex (V1)



Important: It was only assumed that the transformation from neural response to fMRI response is linear.

Properties of a Linear Time-Invariant (LTI) system

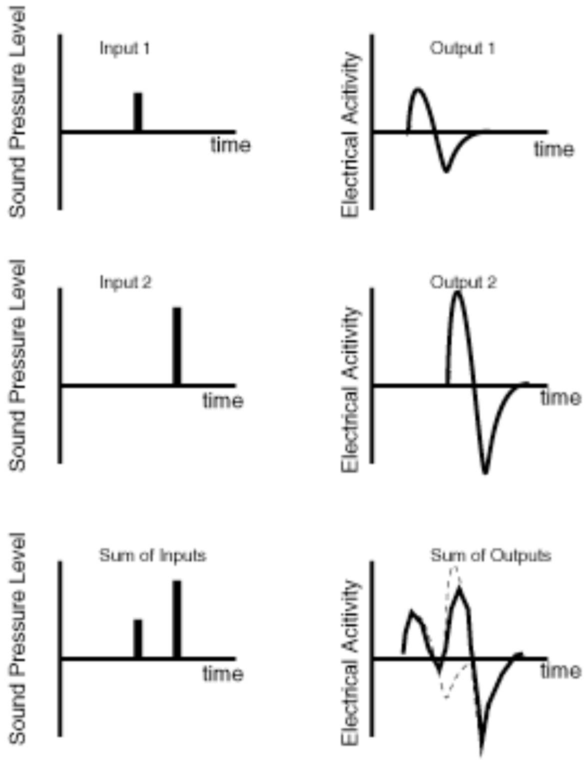
- Homogeneity



When the input magnitude is doubled, the output response is also doubled

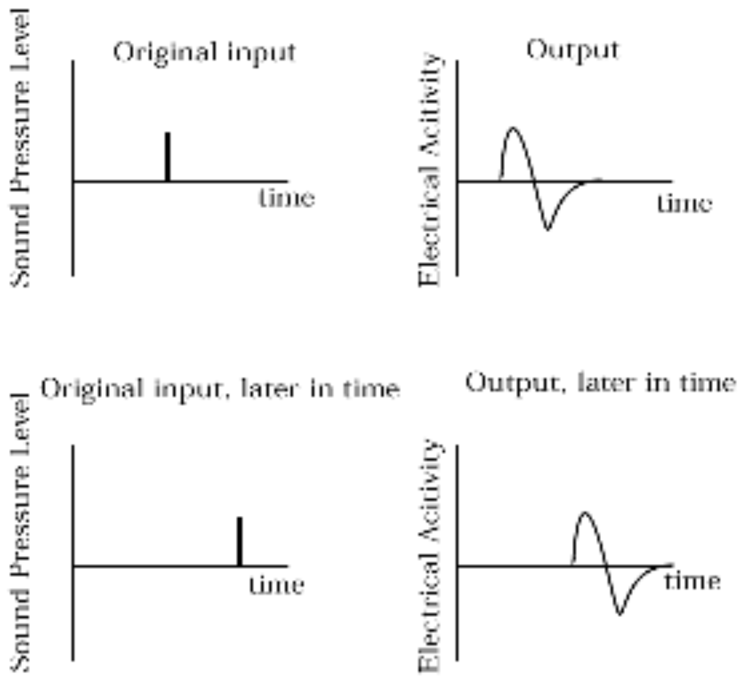
Properties of a Linear Time-Invariant (LTI) system

- Additivity



Properties of a Linear Time-Invariant (LTI) system

- Shift invariance



Shift invariance with respect to time (time invariant)

Why is LTI system important?

If LTI holds,

Convolution

$$(h * f)(t) = \int_0^\infty h(\tau) f(t-\tau) d\tau.$$

The output of a LTI system is simply the convolution of the input and the impulse response function, *h(t)*

If LTI holds,

Convolution

$$(h * f)(t) = \int_0^\infty h(\tau) f(t-\tau)\, d\tau.$$

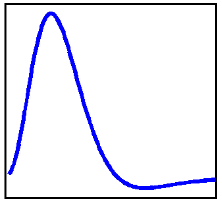
↕

The output of a LTI system is simply the convolution of the input and the impulse response function, *h(t)*

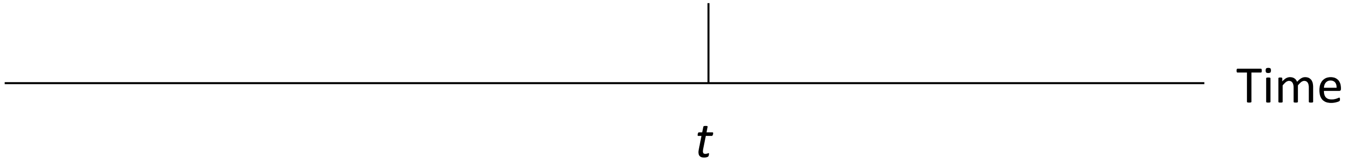
Visualizing convolution

$$(h * f)(t) = \int_0^\infty h(\tau) f(t-\tau) d\tau.$$

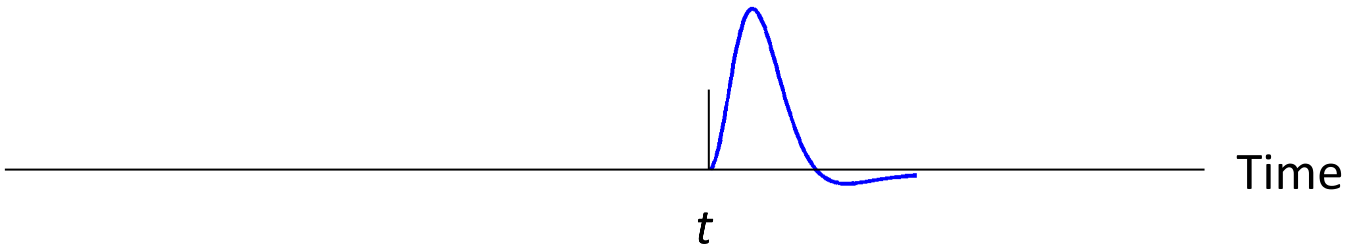
- Suppose *h(t)* looks like



- There is an impulse (an input) at time *t*



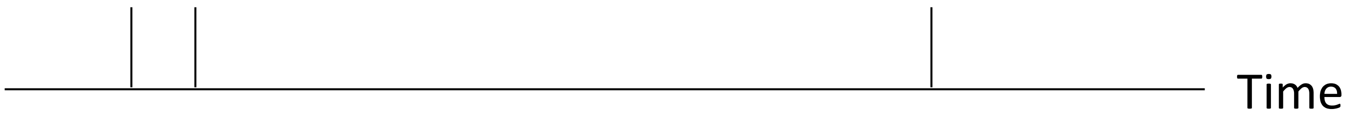
- The impulse response to this input is obtained by convolving the impulse with the impulse response function *h(t)*



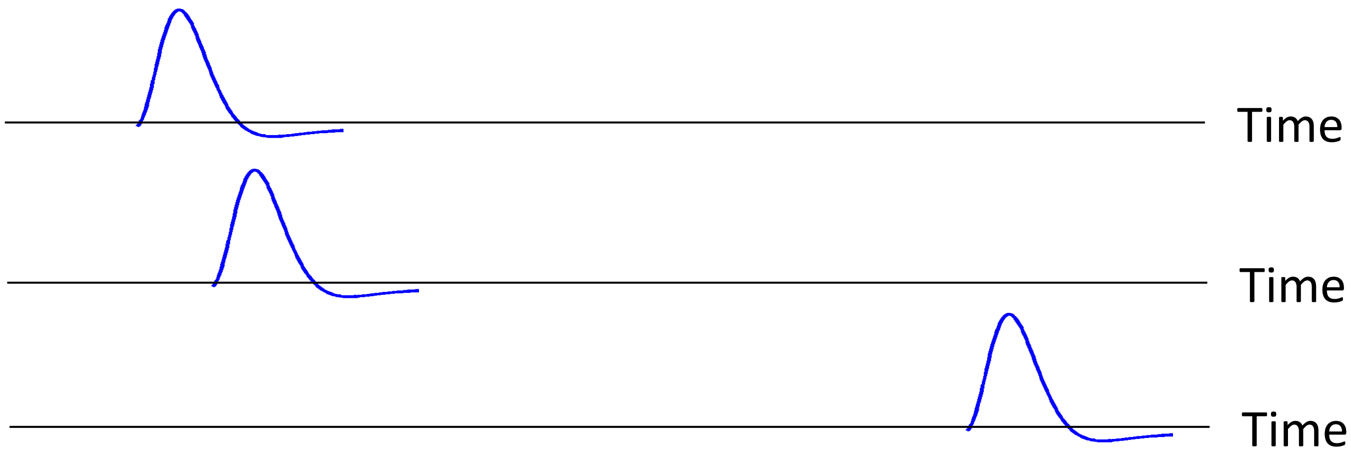
Visualizing convolution

$$(h * f)(t) = \int_0^{\infty} h(\tau) f(t - \tau) d\tau.$$

- Input time series



- Output (response) time series (the sum of the 3 time series)



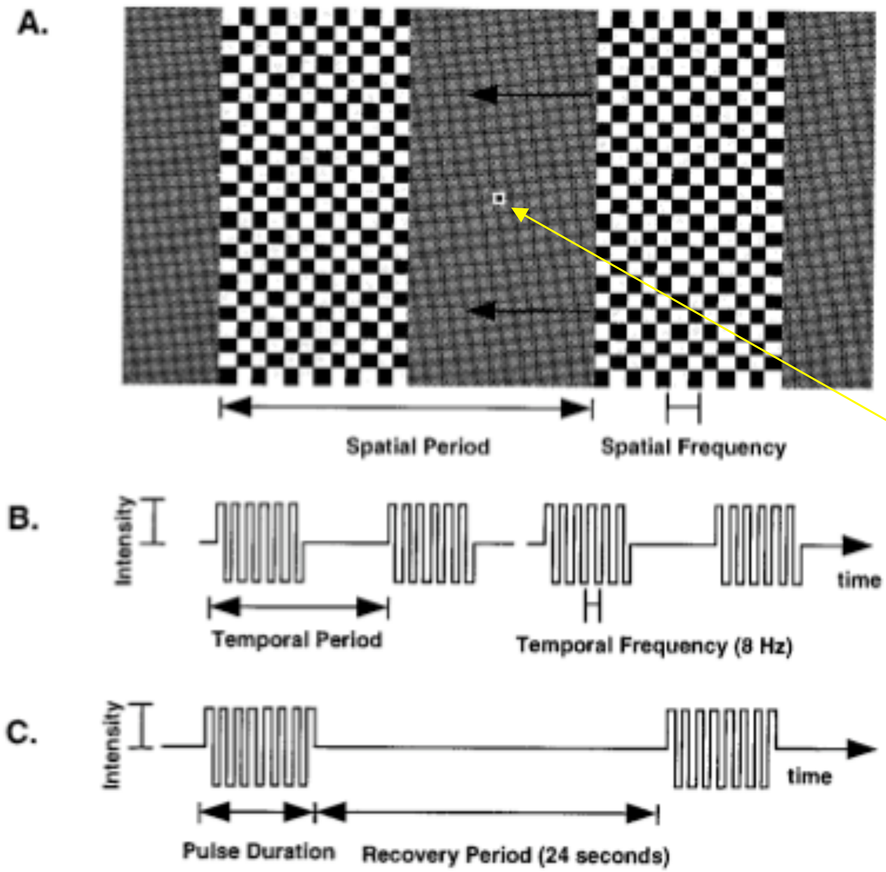
Why is LTI system important?

- The output of a LTI system is simply the convolution of the input and the impulse response function, *h(t)*

- All we need to know is *h(t)*

Boynton et al. (1996)

- flickering checkerboard task



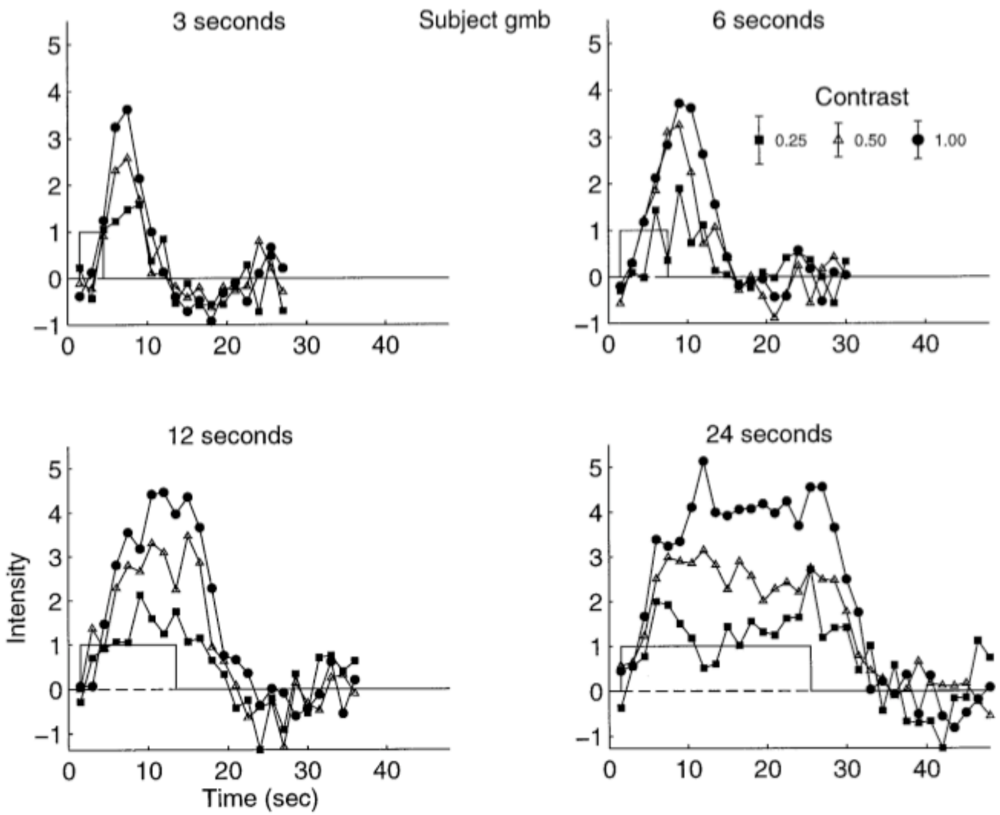
- 4 pulse durations (3, 6, 12, 24 s)
- 4 stimulus contrasts (0, 0.25, 0.5, 1)

Eye fixation

Boynton et al. (1996)

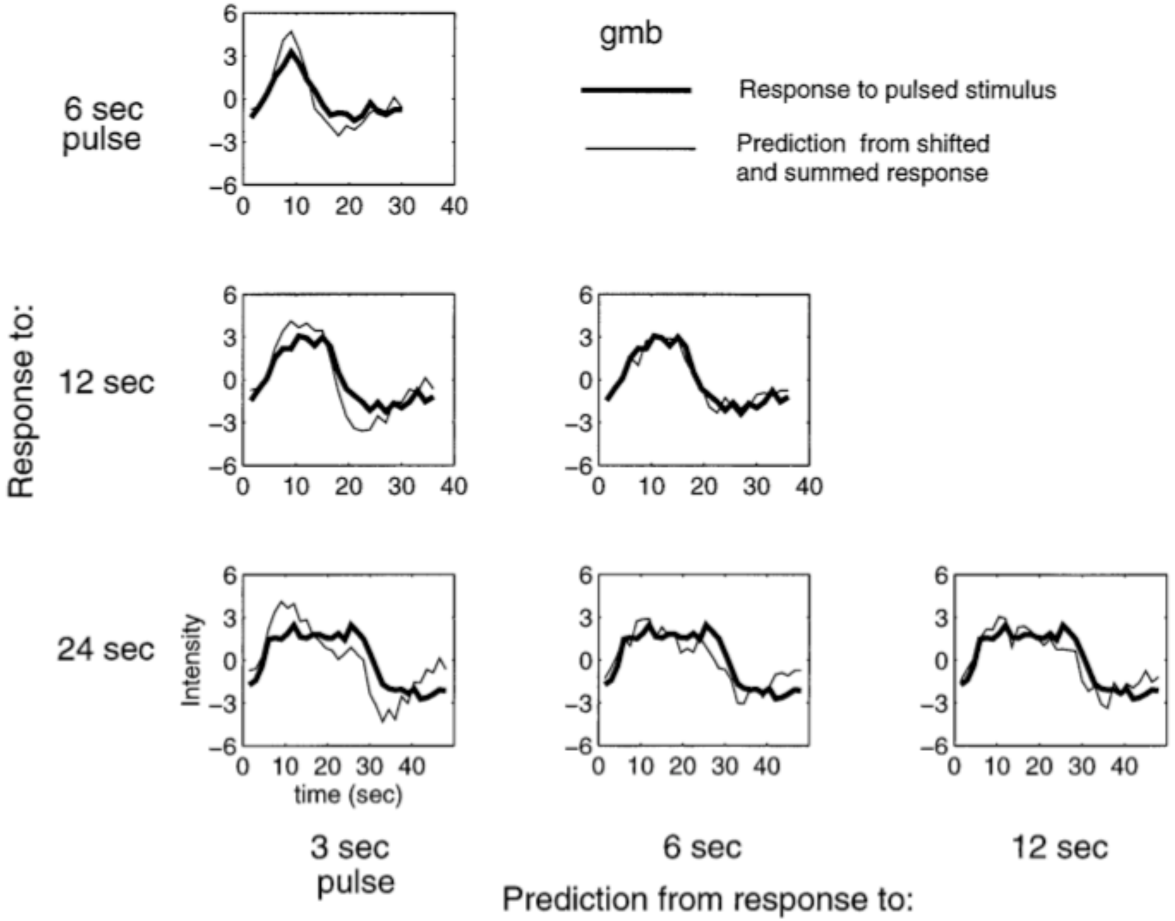
- Results

Look how response varies as a function of contrast and duration; Could you tell if LTI is hold?



Boynton et al. (1996)

- LTI assumption holds in most cases



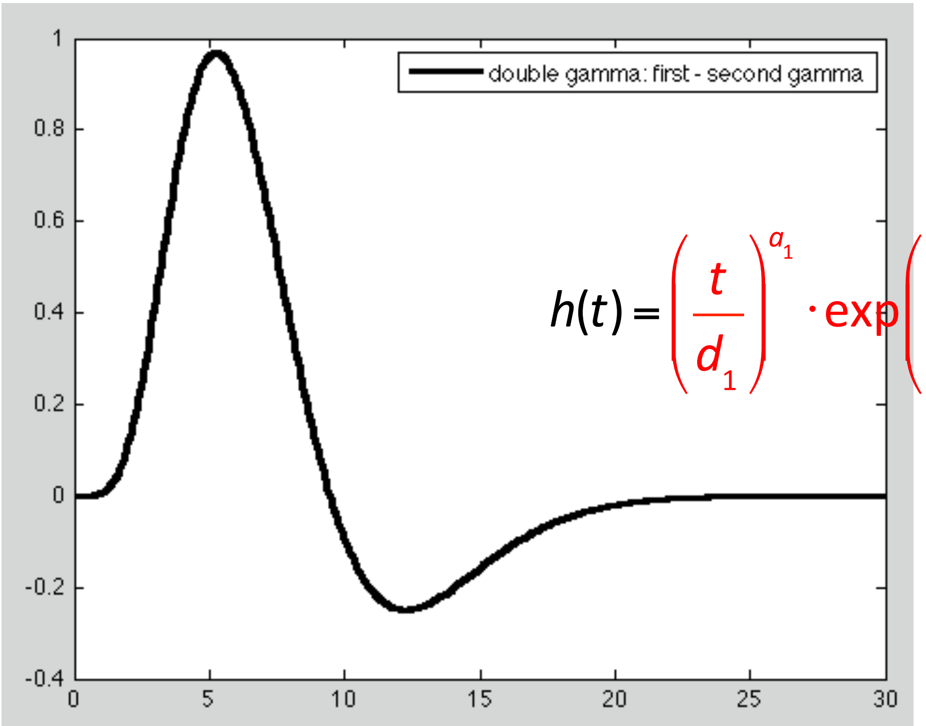
Modeling *h(t)*

We need to have a good estimate of *h(t)*

Modeling *h(t)*

We need to have a good estimate of *h(t)*

Use a double-gamma function to model *h(t)*

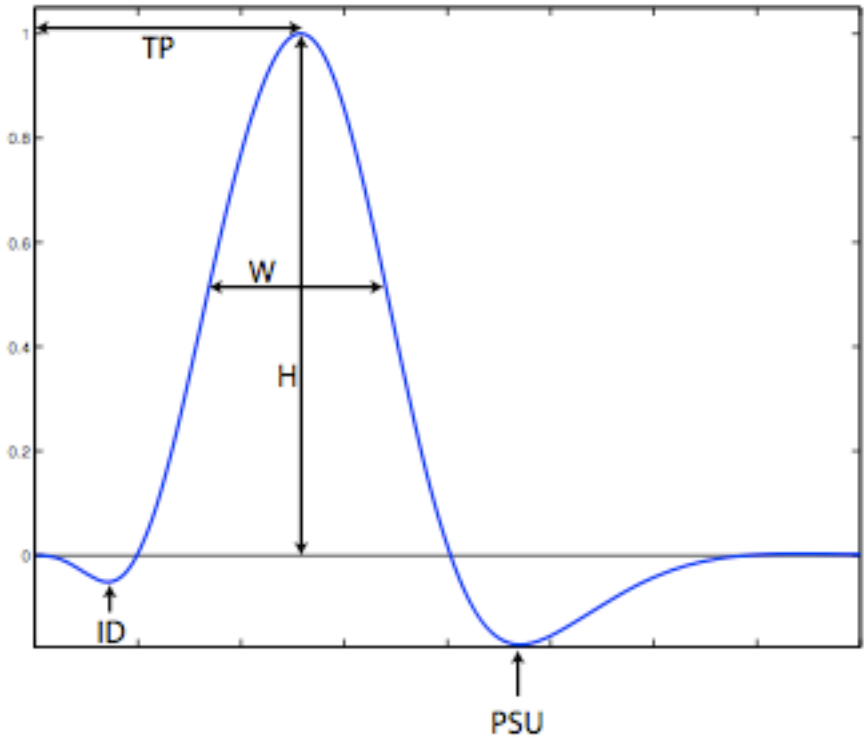


$$h(t) = \left(\frac{t}{d_1}\right)^{a_1} \cdot \exp\left(-\frac{t-d_1}{b_1}\right) - c \cdot \left(\frac{t}{d_2}\right)^{a_2} \cdot \exp\left(-\frac{t-d_2}{b_2}\right)$$

Modeling *h(t)*

We need to have a good estimate of *h(t)*

Canonical HRF (double-gamma function)



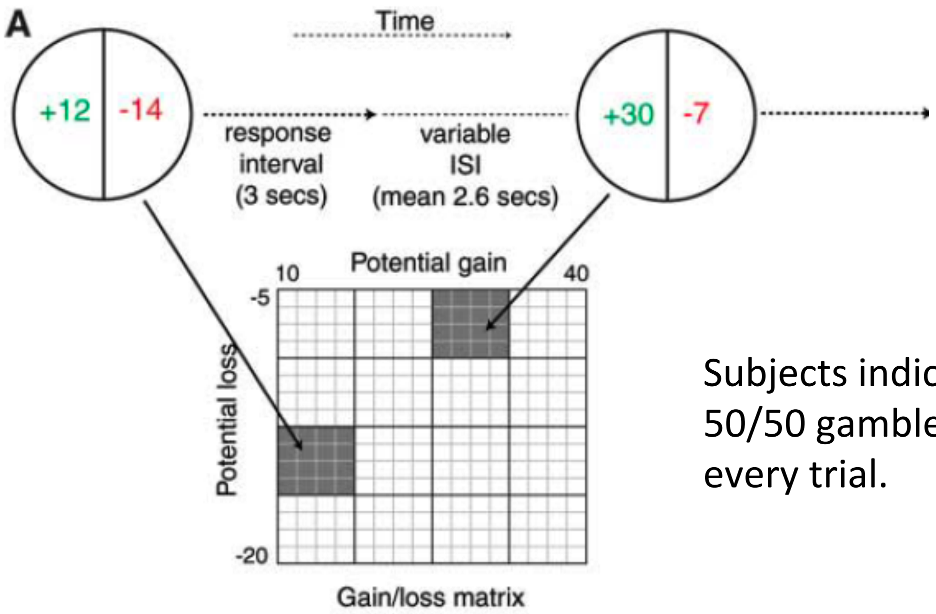
Another example: More complicated design
Tom et al. (2007, Science)

Tom et al. (2007)

Russ Poldrack

- Question: Where and how are monetary gains and losses represented in the brain when people are making risky decisions?

- Experimental design:



Subjects indicate whether s/he wanted to play a 50/50 gamble on either winning $x or losing $y in every trial.

Tom et al. (2007)

- Construct a General Linear Model (GLM) to analyze data

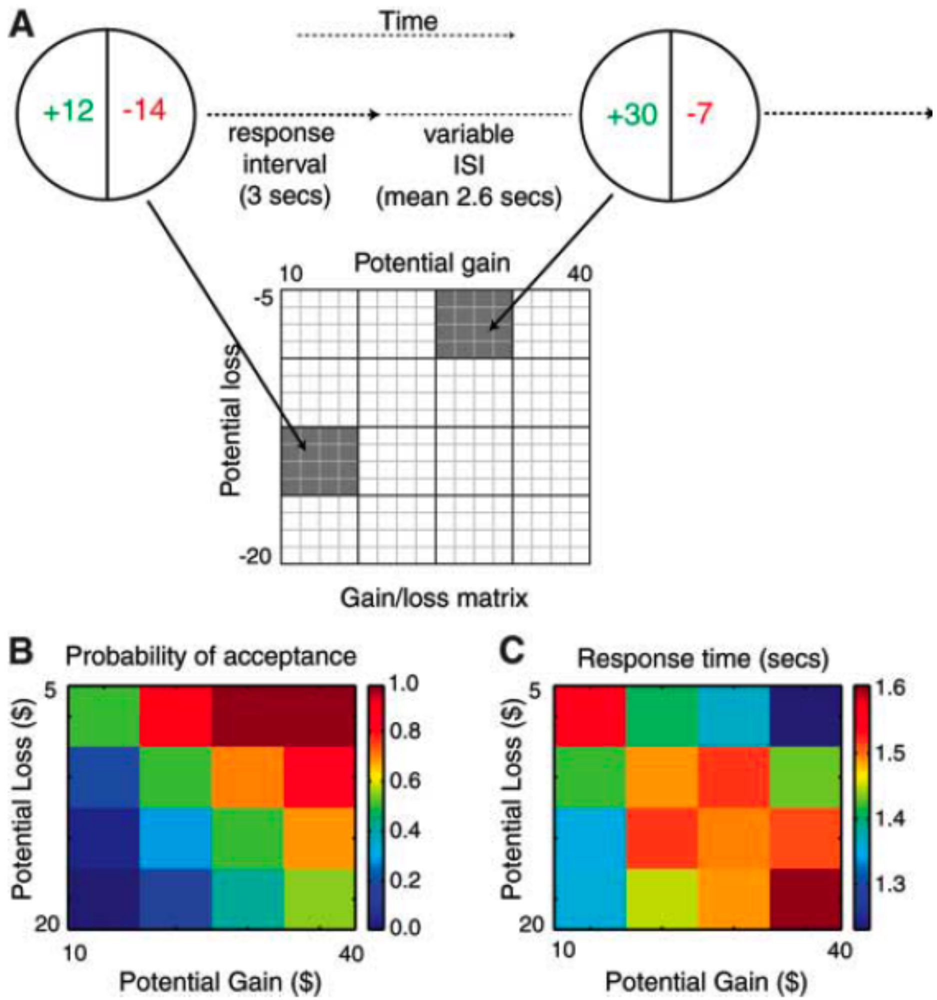
$$Y(t) = \beta_0(t) + \beta_1 x_I(t) + \beta_{Gain} x_{Gain}(t) + \beta_{Loss} x_{Loss}(t) + \beta_{dist} x_{dist}(t)$$

Indicator variable — The amount of gain — The amount of loss — Euclidian distance

By having this model, we can look for brain areas correlated with gains, losses, or both.

Tom et al. (2007)

- Design



Tom et al. (2007)

- Construct General Linear Model (GLM) to analyze data

$$Y(t) = \beta_0(t) + \beta_1 x_I(t) + \beta_{Gain} x_{Gain}(t) + \beta_{Loss} x_{Loss}(t) + \beta_{dist} x_{dist}(t)$$

Indicator variable

The amount of gain

The amount of loss

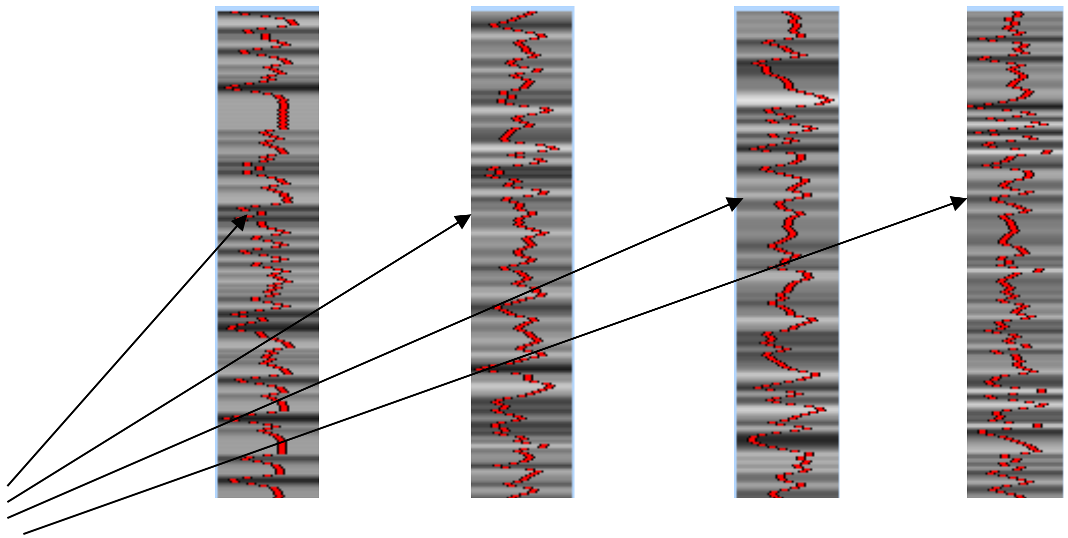
Euclidian distance

Important:
This is the GLM that analyzes data of a single run of a single subject

Tom et al. (2007)

- Construct General Linear Model (GLM) to analyze data

$$Y(t) = \beta_0(t) + \beta_1 x_l(t) + \beta_{Gain} x_{Gain}(t) + \beta_{Loss} x_{Loss}(t) + \beta_{dist} x_{dist}(t)$$

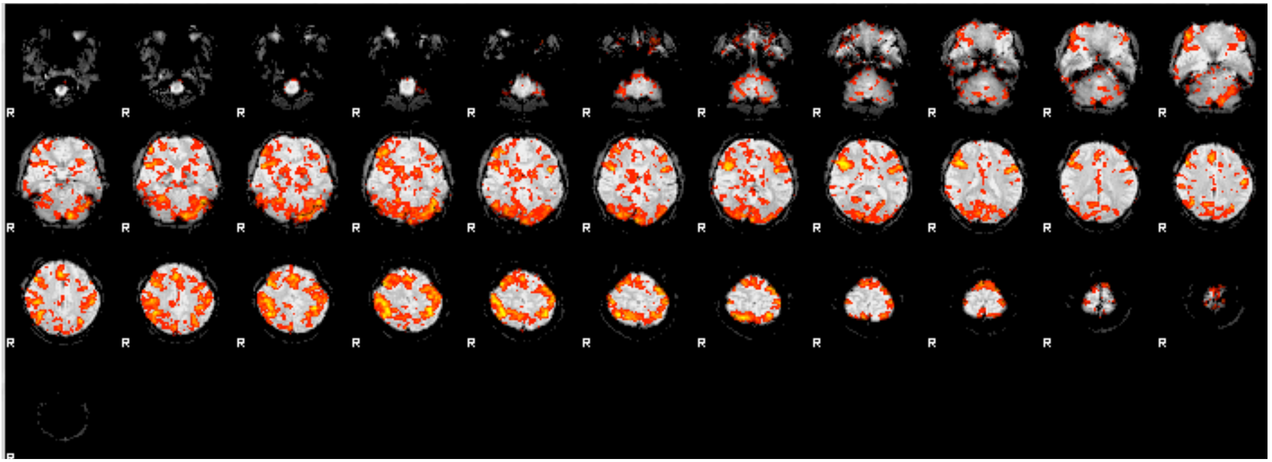


HRF-convolved time course

Example from a single subject's run

Single-subject results (Tom et al. 2007)

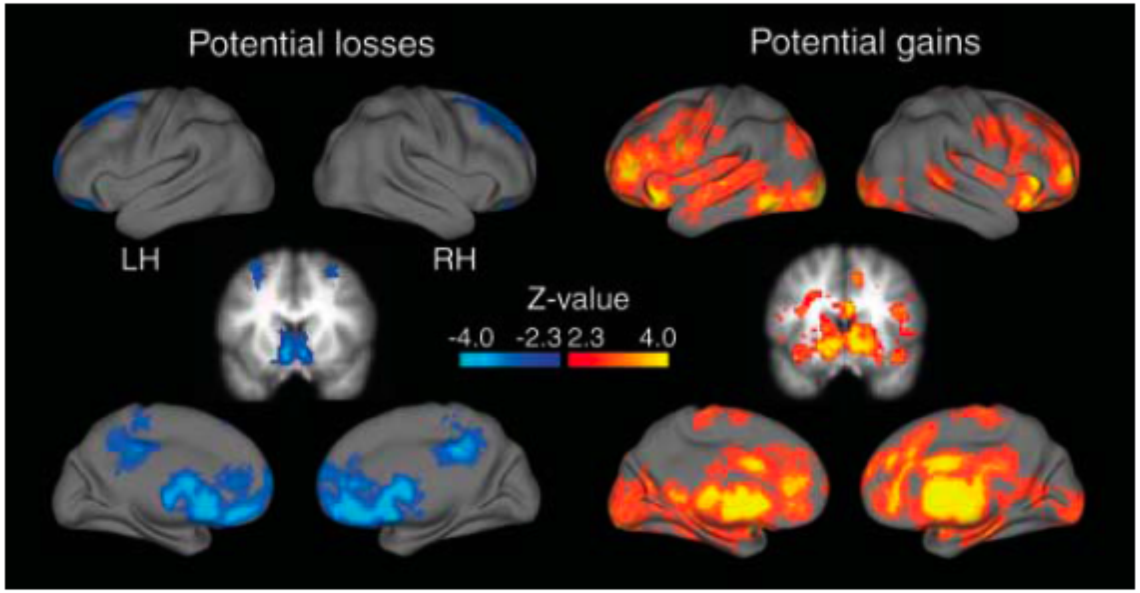
Not the pretty pictures you often see in fMRI papers …



For fMRI experiments, we typically need ~20 subjects' data to obtain meaningful results

Group results (Tom et al., 2007)

• Analysis done at the group level is different from that done at the single-subject level (we won't get into this today)



• Network of regions positively correlated with gains and negatively correlated with losses

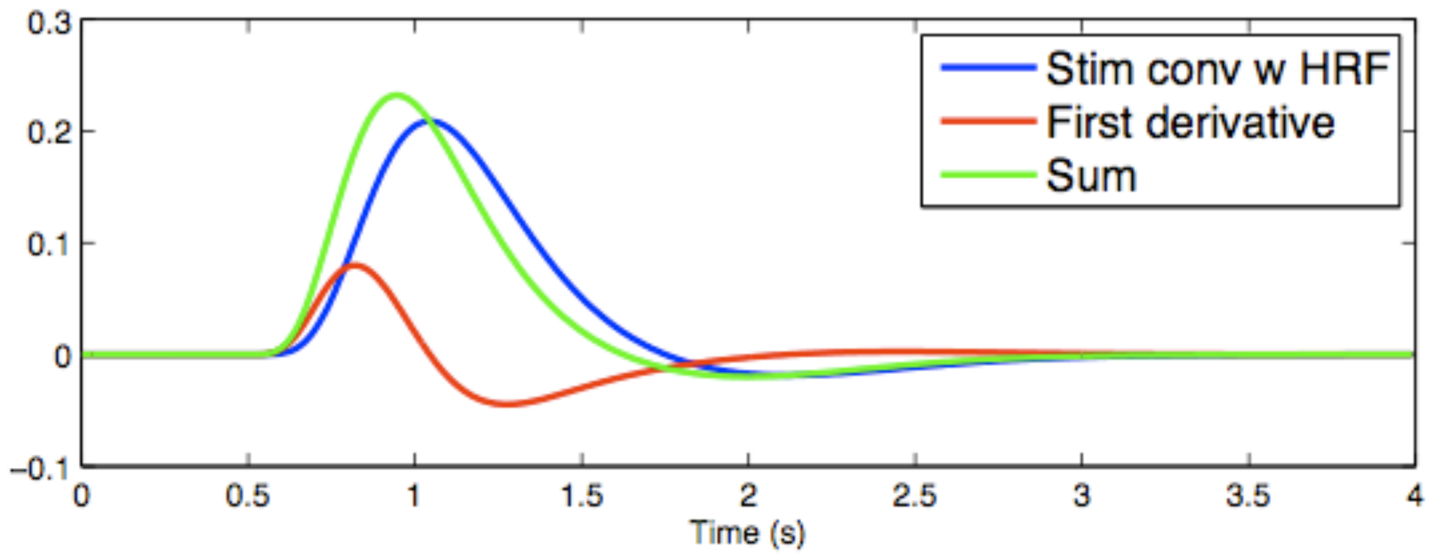
Additional Material

Modeling *h(t)*

We need to have a good estimate of *h(t)*

Beyond canonical HRF

- Modeling the derivative



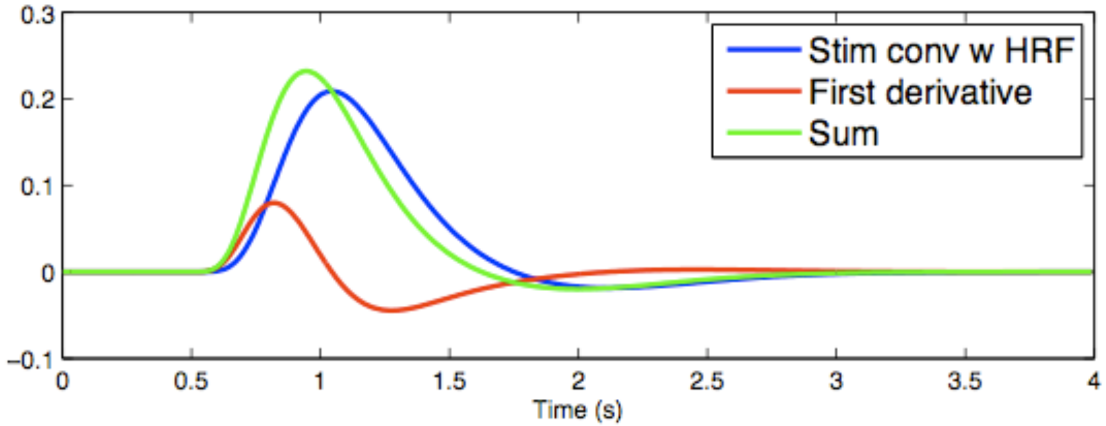
Modeling the derivative

$$Y(t) = \beta X(t)$$

$$Y(t) = \beta X(t + \delta)$$

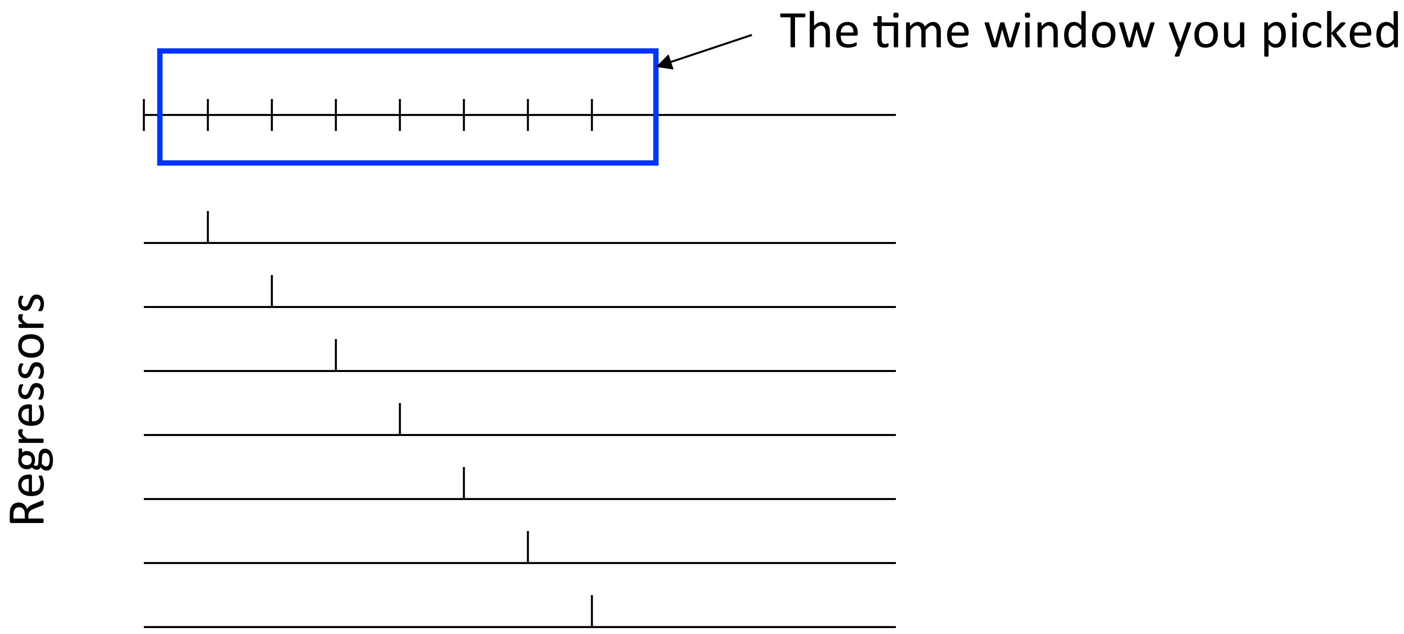
$$X(t + \delta) = X(t) + \delta X^{'}(t) + \ldots$$

$$Y(t) = \beta\left(X(t) + \delta X^{'}(t) + \ldots\right) \approx \beta_1 X(t) + \beta_2 X^{'}(t)$$



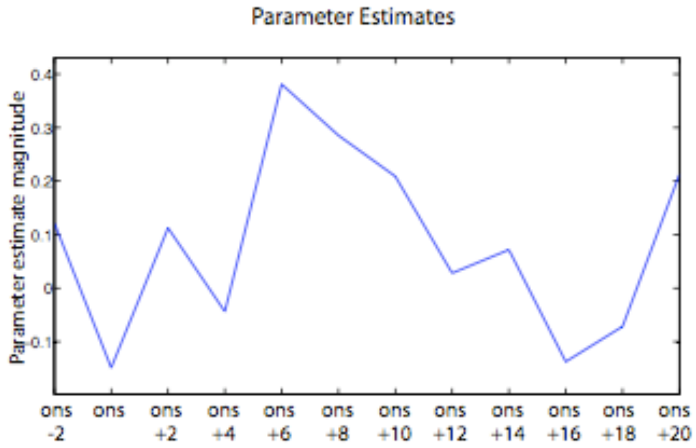
Finite Impulse Response (FIR) models

- Treating each time point within a selected window as a parameter be estimated



Finite Impulse Response (FIR) models

- Treating each time point within a selected window as a parameter be estimated



- Advantage:
    - no assumption needs to be made about the HRF
    - more flexible to estimate the shape of HRF

Finite Impulse Response (FIR) models

- Treating each time point within a selected window as a parameter be estimated

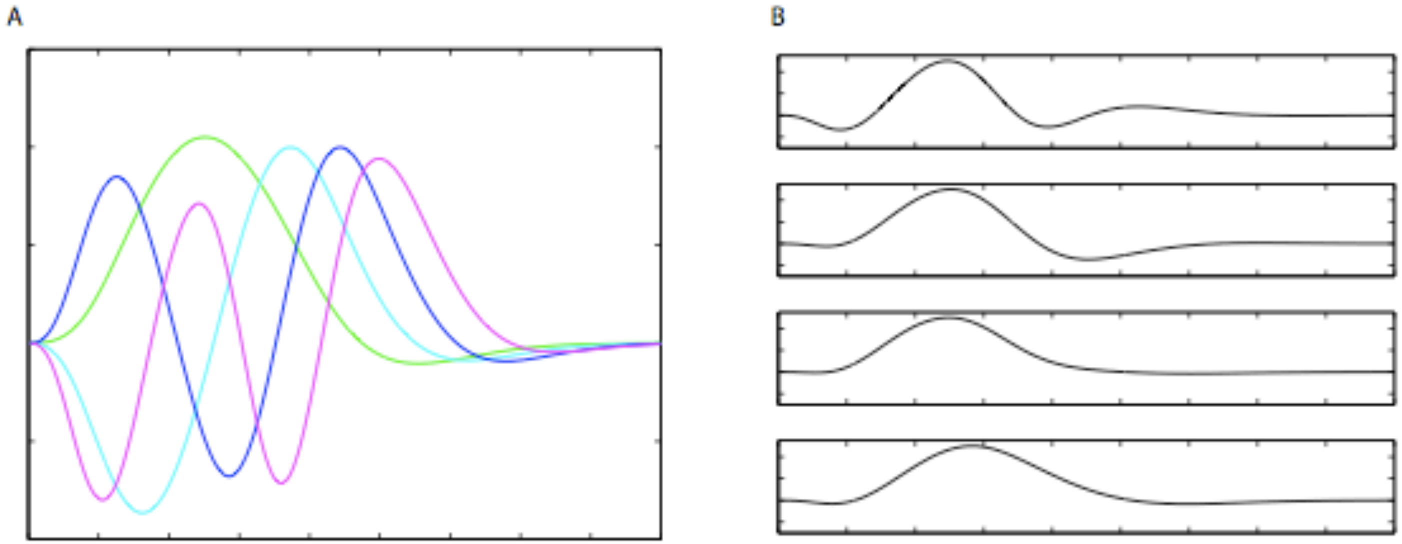
- Advantage:
  - no assumption needs to be made about the HRF
  - more flexible to estimate the shape of HRF

- disadvantage:
  - increase in variability of parameter estimates
  - over-fitting (choice of the length of window critical)
  - not easy to set up group analysis (multivariate problem)

Constrained basis sets

- Picking a number of basis functions known to capture HRF

Example:

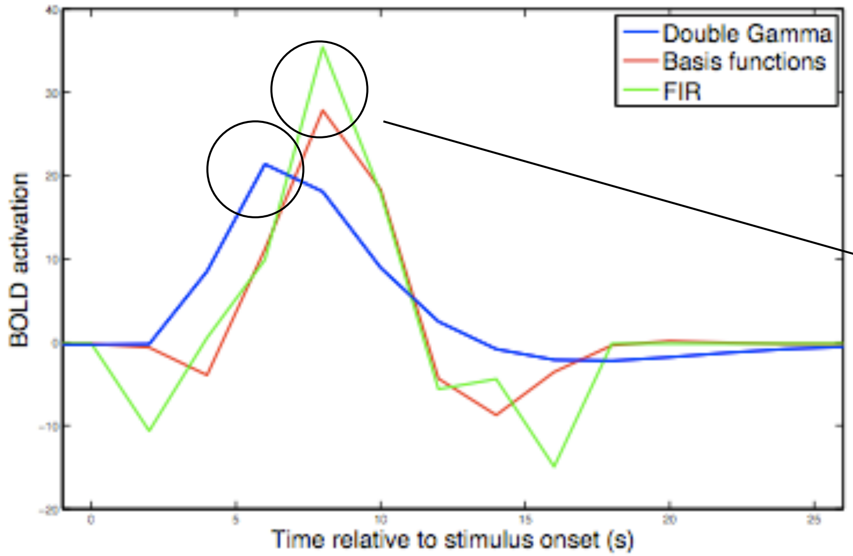


- 4 basis functions
- 4 linear combinations of the basis functions

Constrained basis sets

- Picking a number of basis functions known to capture HRF

- a balance between FIR and canonical HRF;

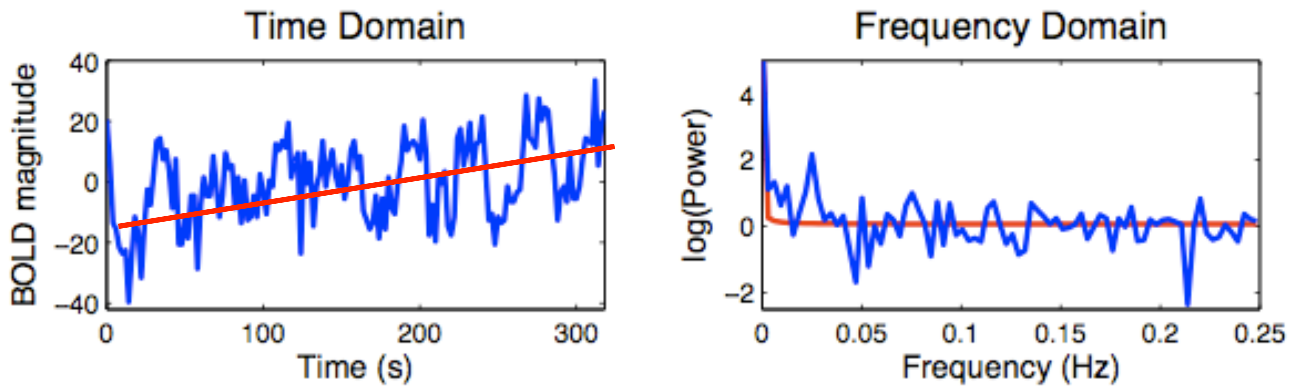


Time to peak slower than Canonical HRF

The BOLD noise

Characterizing the noise

- low frequency drift



Source: scanner as an additional source of structured noise

Important: avoid designing experiment between 0-0.015 Hz (block length no more than 35 s for on-off block design)

Characterizing the noise

- low frequency drift

  Source: scanner as an additional source of structured noise

  Important: avoid designing experiment between 0-0.015 Hz
  (block length no more than 35 s for on-off block design)

- How to remove

  Step 1: high-pass filtering

  Step 2: pre-whitening to remove temporal autocorrelation

Pre-whitening

- Estimate the residuals by ignoring autocorrelation in the data

$$Y = X\beta + \varepsilon, \varepsilon : N(0, \sigma^2)$$

- We get V where $\sigma^2 V$ is the covariance of the matrix

$$V = \begin{pmatrix} 1 & \rho & \cdots & \rho^{n-1} \\ \rho & 1 & \cdots & \rho^{n-2} \\ \cdots & \cdots & \cdots & \cdots \\ \rho^{n-1} & \rho^{n-2} & \cdots & 1 \end{pmatrix}$$

where $Cor(\varepsilon_i, \varepsilon_{i-1}) = \rho^{|l|}$

(l is the time lag)

Pre-whitening

- Estimate the residuals by ignoring autocorrelation in the data

$$Y = X\beta + \varepsilon, \varepsilon : N(0,\sigma^2)$$

- We get V where $\sigma^2 V$ is the covariance of ε

- Find W such that $WVW^{'} = I_T$

- Apply W to both sides of the equation

$$WY = WX\beta + W_\varepsilon$$

- This made sure the estimated noise is i.i.d (no correlation in time)

Pre-whitening

- This made sure the estimated noise is i.i.d (no correlation in time)

- Important: W dependent upon X

  When your GLM changes, so does W

Once we have the pre-whitened data, we are ready to estimate the parameters of interest in our GLM.

But …

- Is my design good?

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

- Covariance

$$\text{cov}(\hat{\beta}) = \left(X^{'}X\right)^{-1}\sigma^{2}$$

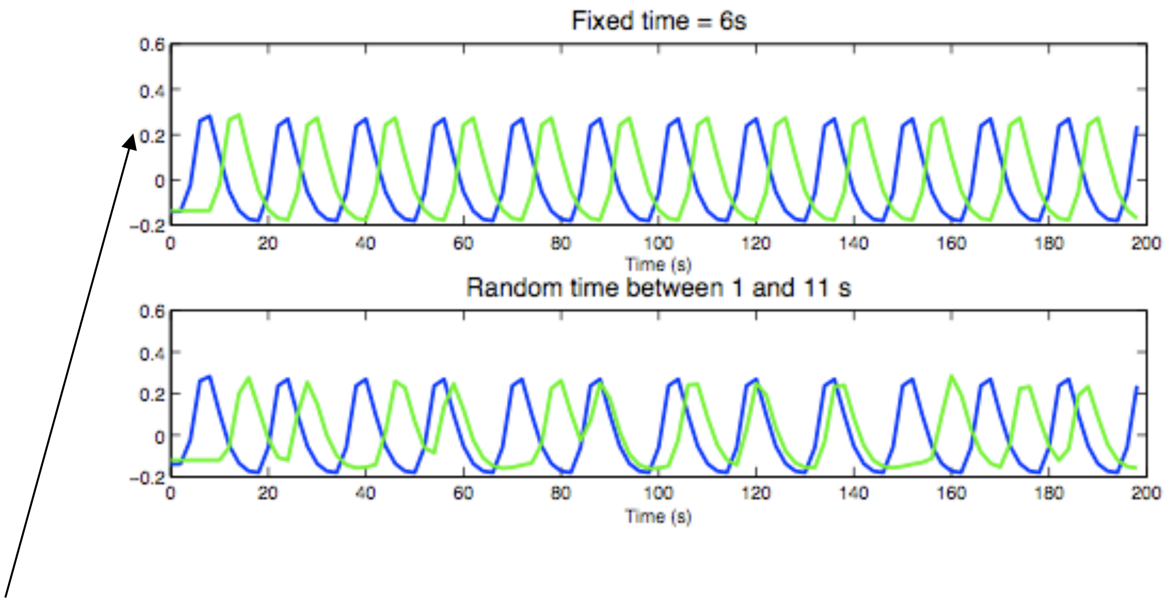
- Design efficiency usually refers to the variance due to the design,
So for a contrast *c*, the efficiency for that contrast is

$$\text{eff}(c\hat{\beta}) = \frac{1}{c\underline{\left(X^{'}X\right)}^{-1}c^{'}}$$

→ Variance of your design matrix

Thinking about timing …

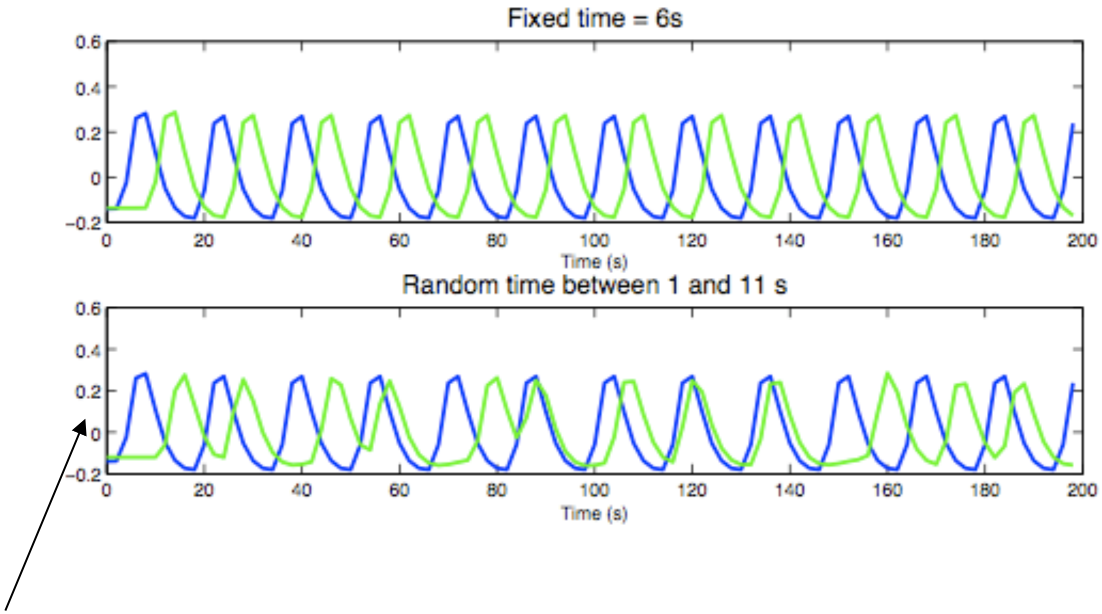
Suppose you have 2 trial types and you want to model them Separately.



When you have a fixed ISI (inter-stimulus interval), the correlation between your trial-type regressors is going to be high.

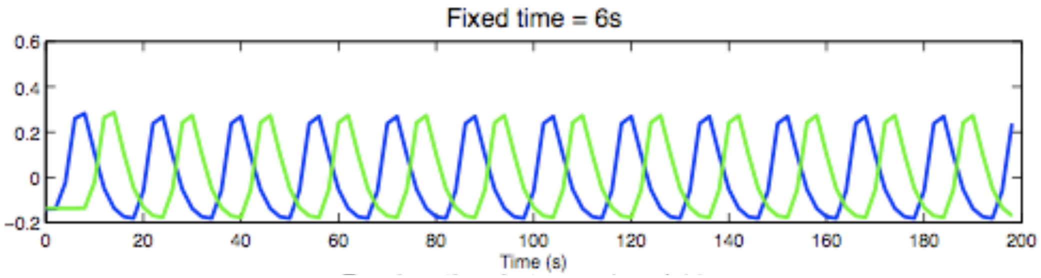
Thinking about timing …

Suppose you have 2 trial types and you want to model them
Separately.



You can reduce that correlation by introducing random timing
Jitters between trials

Calculating efficiency



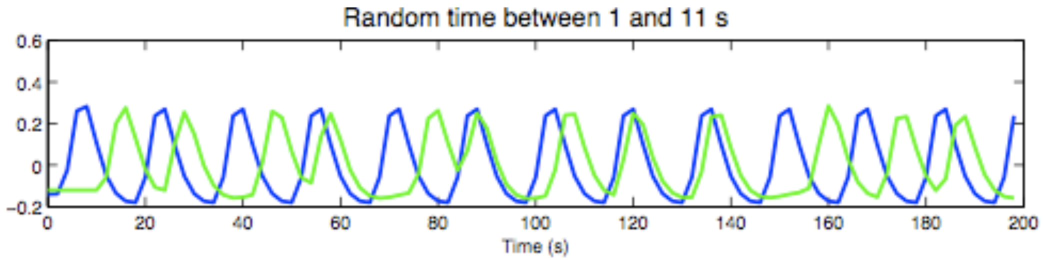
- The correlation between the two time courses was -0.61

$$(X'X)^{-1} = \begin{pmatrix} 0.5632 & 0.3465 \\ 0.3465 & 0.5703 \end{pmatrix}$$

- For contrast c=[1 0],

$$\text{eff}(c\hat{\beta}) = \frac{1}{c(X'X)^{-1}c'} = \frac{1}{0.5632} = 1.76$$

Calculating efficiency



- The correlation between the two time courses was -0.15

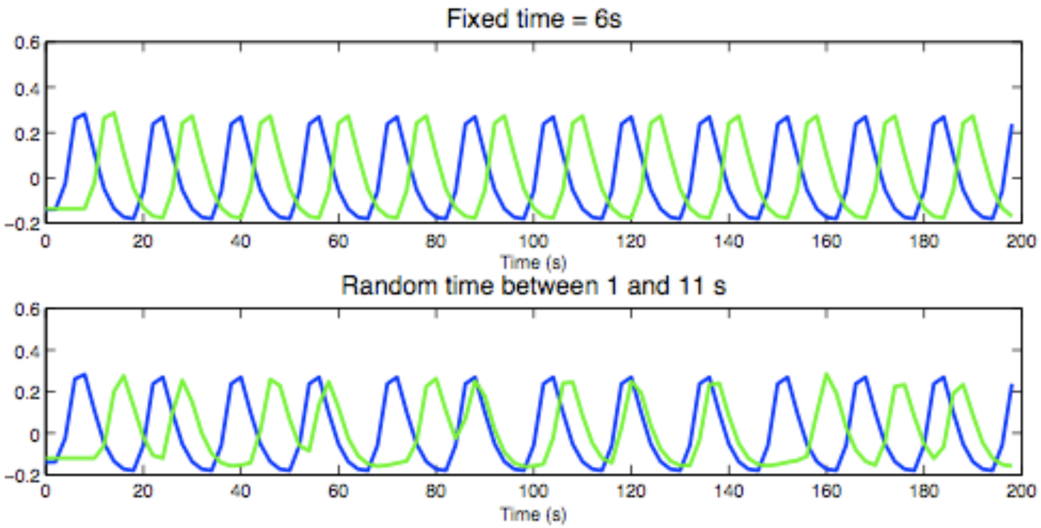
$$(X'X)^{-1} = \begin{pmatrix} 0.3606 & 0.0602 \\ 0.0602 & 0.4640 \end{pmatrix}$$

- For contrast c=[1 0],

$$\mathrm{eff}(c\hat{\beta}) = \frac{1}{c\left(X'X\right)^{-1}c'} = \frac{1}{0.3606} = 2.77$$

Calculating efficiency

- For contrast c=[1 0],



$$\mathrm{eff}(c\hat{\beta}) = 1.76$$

$$\mathrm{eff}(c\hat{\beta}) = 2.77$$

The efficiency is better with the design that jitters timing (in this case from a uniform distribution [1s 11s]

Thinking about timing …

Final note: it's all about how you would like to analyze your data.

- The correlation is introduced because you wish to model the different trial types separately (a legitimate model), and the fact that one trial type always follows the other (in this example)

- It is always a good exercise to come up with a design matrix (or multiple design matrices) for analyzing your data before you even start collecting data!