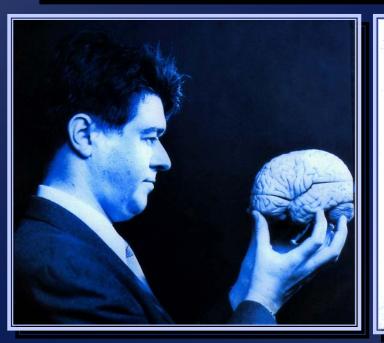


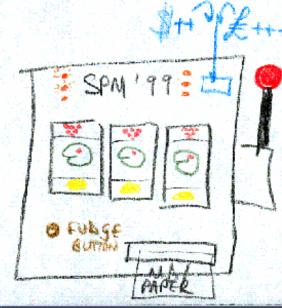
FIL

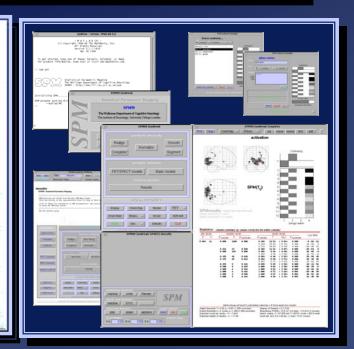
Event-related fMRI

Rik Henson

With thanks to: Karl Friston, Oliver Josephs







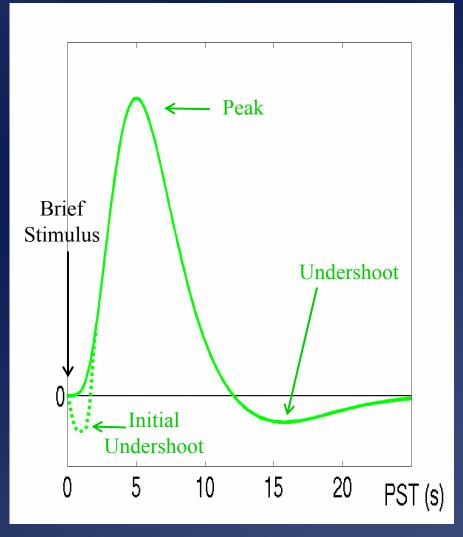
Overview

1. BOLD impulse response

- 2. General Linear Model
- 3. Temporal Basis Functions
- 4. Timing Issues
- 5. Design Optimisation
- 6. Nonlinear Models
- 7. Example Applications

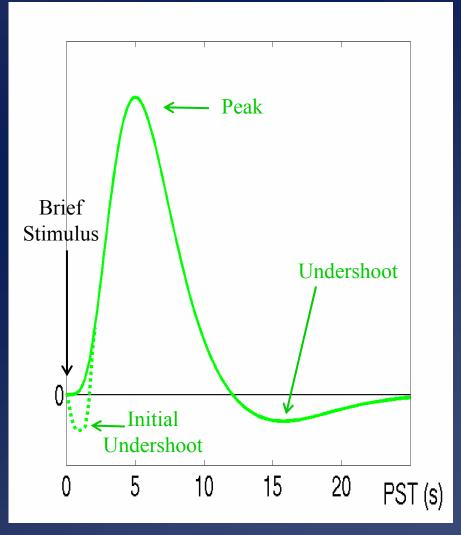
BOLD Impulse Response

- Function of blood oxygenation, flow, volume (Buxton et al, 1998)
- Peak (max. oxygenation) 4-6s poststimulus; baseline after 20-30s
- Initial undershoot can be observed (Malonek & Grinvald, 1996)
- Similar across V1, A1, S1...
- ... but differences across: other regions (Schacter et al 1997) individuals (Aguirre et al, 1998)



BOLD Impulse Response

- Early event-related fMRI studies used a long Stimulus Onset Asynchrony (SOA) to allow BOLD response to return to baseline
- However, if the BOLD response is explicitly modelled, overlap between successive responses at short SOAs can be accommodated...
- ... particularly if responses are assumed to superpose linearly
- Short SOAs are more sensitive...



Overview

1. BOLD impulse response

2. General Linear Model

- 3. Temporal Basis Functions
- 4. Timing Issues
- 5. Design Optimisation
- 6. Nonlinear Models
- 7. Example Applications

General Linear (Convolution) Model

GLM for a single voxel:

 $y(t) = u(t) \otimes h(\tau) + \varepsilon(t)$

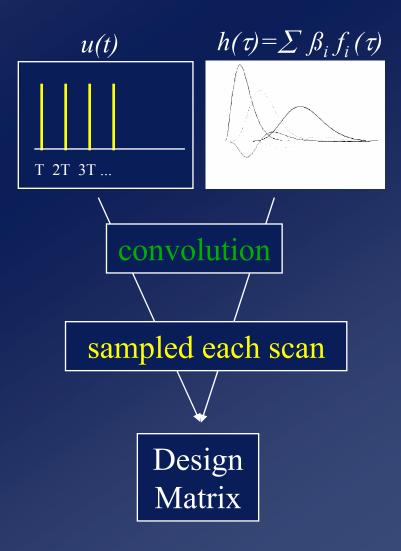
u(t) = neural causes (stimulus train)

 $u(t) = \sum \delta(t - nT)$

 $h(\tau)$ = hemodynamic (BOLD) response

 $h(\tau) = \sum \beta_i f_i(\tau)$

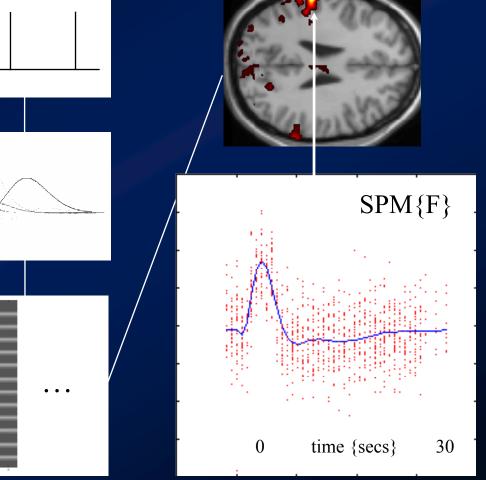
 $f_{i}(\tau) = \text{temporal basis functions}$ $y(t) = \sum \sum \beta_{i} f_{i}(t - nT) + \varepsilon(t)$ $y = X\beta + \varepsilon$



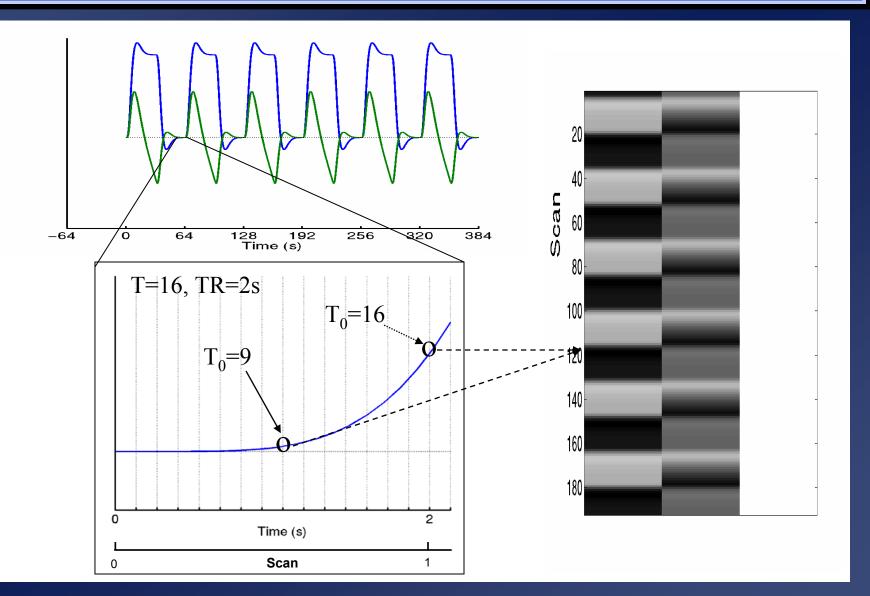
Gamma functions $f_1(\tau)$ of

peristimulus time τ (Orthogonalised)

Sampled every TR = 1.7s Design matrix, **X** $[x(t)\otimes f_1(\tau) | x(t)\otimes f_2(\tau) |...]$



A word about down-sampling



Overview

1. BOLD impulse response

2. General Linear Model

3. Temporal Basis Functions

4. Timing Issues

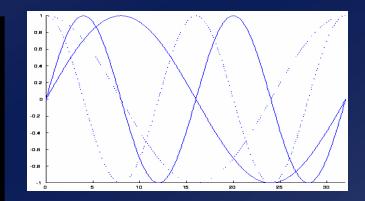
5. Design Optimisation

6. Nonlinear Models

7. Example Applications

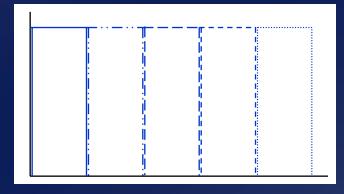
• Fourier Set

Windowed sines & cosines Any shape (up to frequency limit) Inference via F-test



• Finite Impulse Response

Mini "timebins" (selective averaging) Any shape (up to bin-width) Inference via F-test

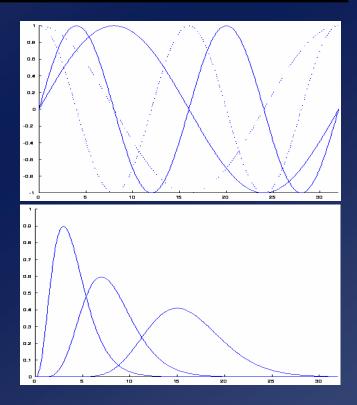


• Fourier Set

Windowed sines & cosines Any shape (up to frequency limit) Inference via F-test

Gamma Functions

Bounded, asymmetrical (like BOLD) Set of different lags Inference via F-test



Fourier Set

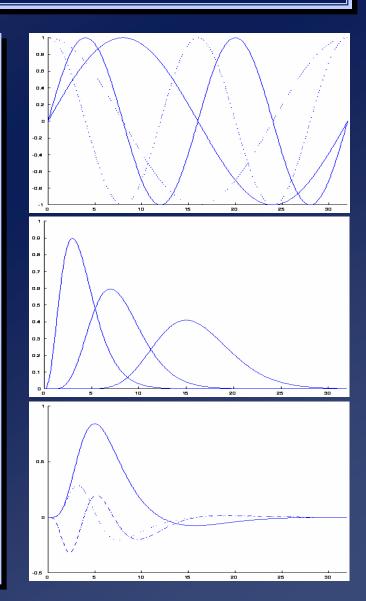
Windowed sines & cosines Any shape (up to frequency limit) Inference via F-test

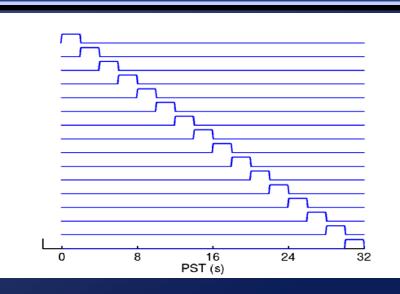
Gamma Functions

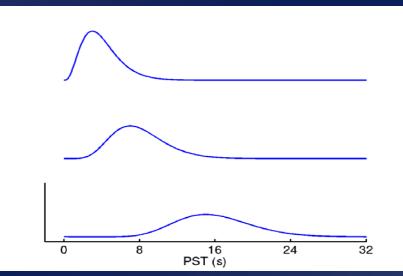
Bounded, asymmetrical (like BOLD) Set of different lags Inference via F-test

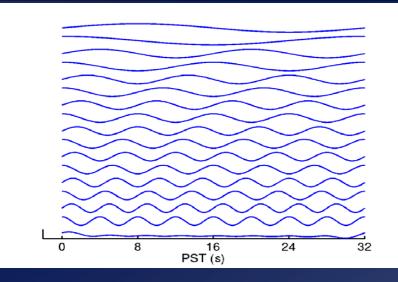
"Informed" Basis Set

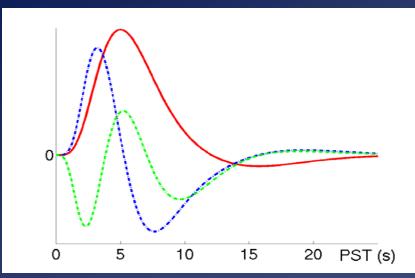
Best guess of canonical BOLD response Variability captured by Taylor expansion "Magnitude" inferences via t-test...?

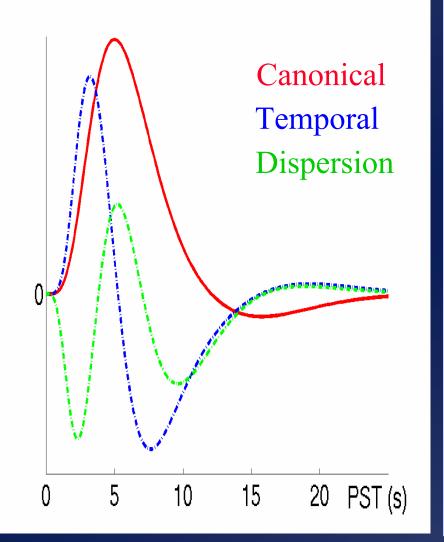












"Informed" Basis Set (Friston et al. 1998)

 Canonical HRF (2 gamma functions)
 plus Multivariate Taylor expansion in: time (Temporal Derivative)
 width (Dispersion Derivative)

• "Magnitude" inferences via t-test on canonical parameters (providing canonical is a good fit...more later)

• "Latency" inferences via tests on *ratio* of derivative : canonical parameters (more later...)

(Other Approaches)

Long Stimulus Onset Asychrony (SOA)

 Can ignore overlap between responses (Cohen et al 1997)
 ... but long SOAs are less sensitive

 Fully counterbalanced designs

 Assume response overlap cancels (Saykin et al 1999)
 Include fixation trials to "selectively average" response

even at short SOA (Dale & Buckner, 1997)

... but unbalanced when events defined by subject

• Define HRF from pilot scan on each subject

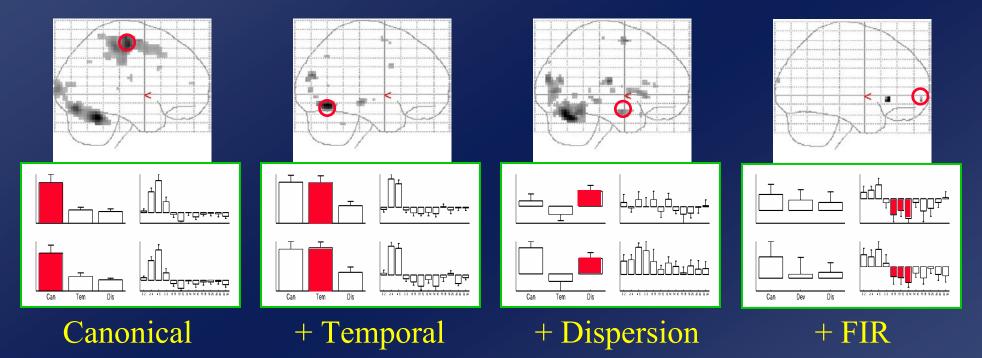
May capture intersubject variability (Zarahn et al, 1997)

... but not interregional variability

• Numerical fitting of highly parametrised response functions Separate estimate of magnitude, latency, duration (Kruggel et al 1999) ... but computationally expensive for every voxel

Temporal Basis Sets: Which One?

In this example (rapid motor response to faces, Henson et al, 2001)...



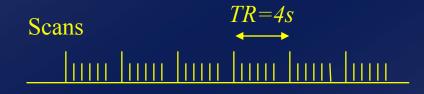
...canonical + temporal + dispersion derivatives appear sufficient
...may not be for more complex trials (eg stimulus-delay-response)
...but then such trials better modelled with separate neural components (ie activity no longer delta function) + constrained HRF (Zarahn, 1999)

Overview

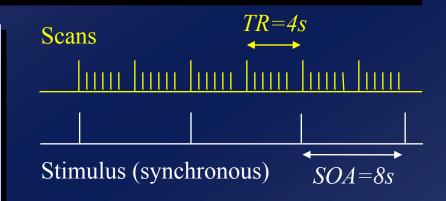
1. BOLD impulse response

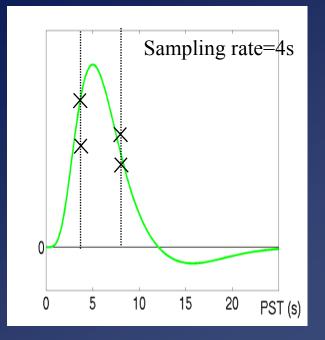
- 2. General Linear Model
- 3. Temporal Basis Functions
- 4. Timing Issues
- 5. Design Optimisation
- 6. Nonlinear Models
- 7. Example Applications

 Typical TR for 48 slice EPI at 3mm spacing is ~ 4s



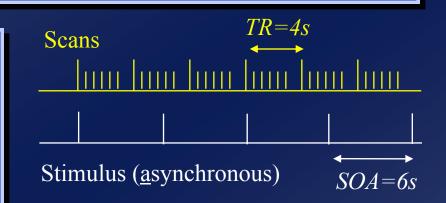
- Typical TR for 48 slice EPI at 3mm spacing is ~ 4s
- Sampling at [0,4,8,12...] poststimulus may miss peak signal

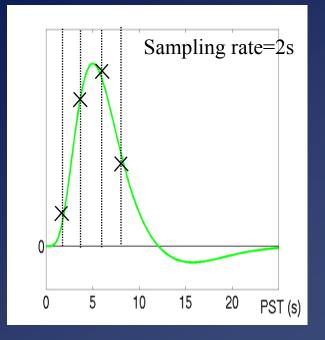




FIL

- Typical TR for 48 slice EPI at 3mm spacing is ~ 4s
- Sampling at [0,4,8,12...] poststimulus may miss peak signal
- Higher effective sampling by: 1. Asynchrony eg SOA=1.5TR

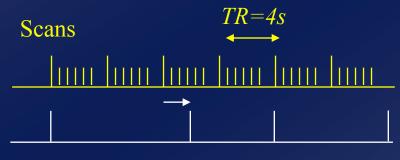




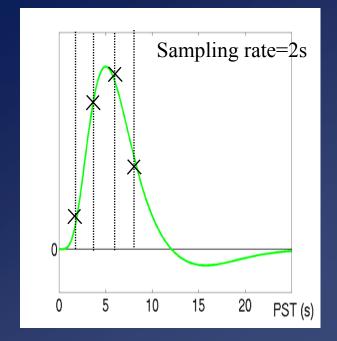
 Typical TR for 48 slice EPI at 3mm spacing is ~ 4s

FIL

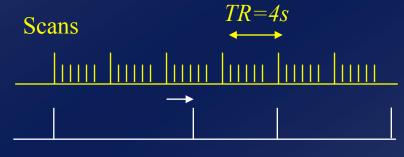
- Sampling at [0,4,8,12...] poststimulus may miss peak signal
- Higher effective sampling by: 1. Asynchrony *eg SOA=1.5TR* 2. Random Jitter *eg SOA=(2±0.5)TR*



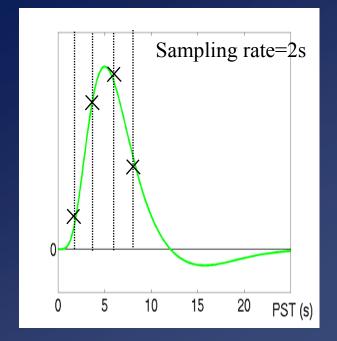
Stimulus (random jitter)



- Typical TR for 48 slice EPI at 3mm spacing is ~ 4s
- Sampling at [0,4,8,12...] poststimulus may miss peak signal
- Higher effective sampling by: 1. Asynchrony *eg SOA=1.5TR* 2. Random Jitter *eg SOA=(2±0.5)TR*
- Better response characterisation (Miezin et al, 2000)



Stimulus (random jitter)

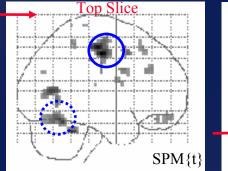


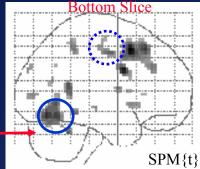
• ...but "Slice-timing Problem" (Henson et al, 1999)

Slices acquired at different times, yet model is the same for all slices

• ...but "Slice-timing Problem" (Henson et al, 1999)

Slices acquired at different times, yet model is the same for all slices => different results (using canonical HRF) for different reference slices



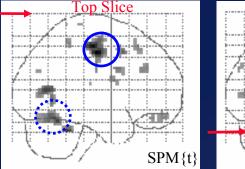


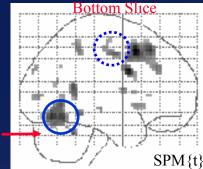
TR=3

• ...but "Slice-timing Problem" (Henson et al, 1999)

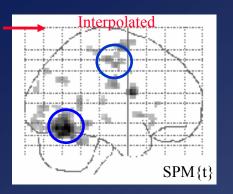
Slices acquired at different times, yet model is the same for all slices => different results (using canonical HRF) for different reference slices

- Solutions:
- 1. Temporal interpolation of data ... but less good for longer TRs





TR=3s



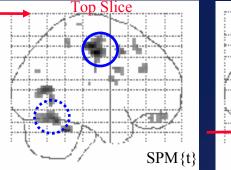
• ...but "Slice-timing Problem" (Henson et al, 1999)

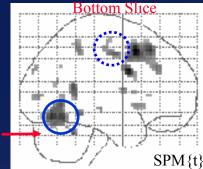
Slices acquired at different times, yet model is the same for all slices => different results (using canonical HRF) for different reference slices

• Solutions:

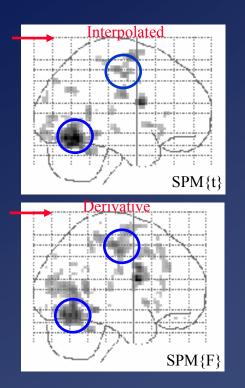
1. Temporal interpolation of data ... but less good for longer TRs

2. More general basis set (e.g., with temporal derivatives) ... but inferences via F-test





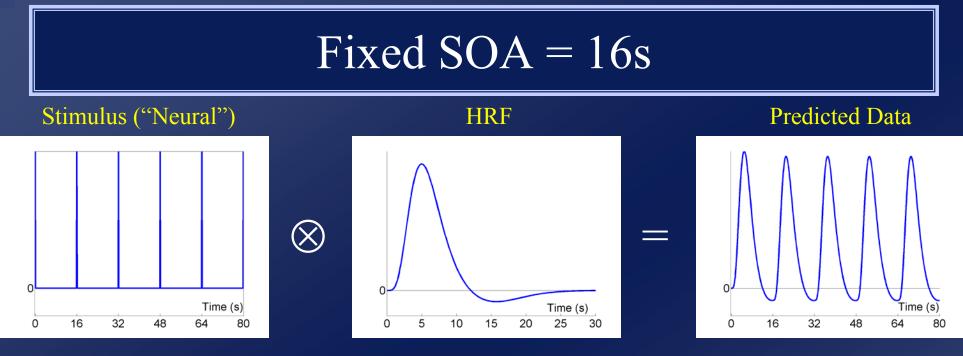
TR=3s



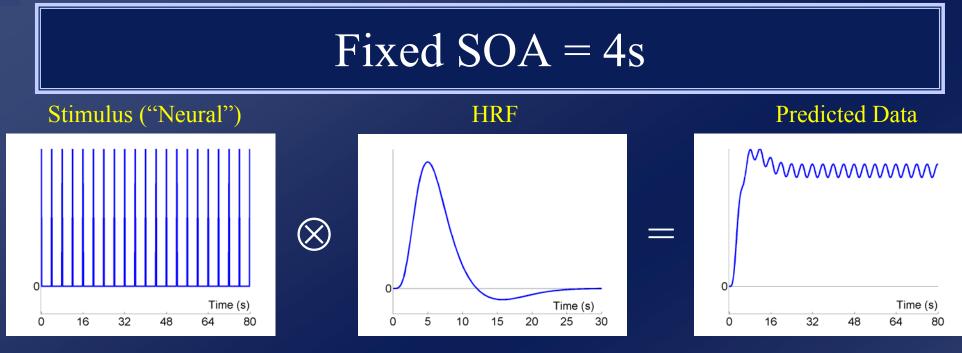
Overview

1. BOLD impulse response

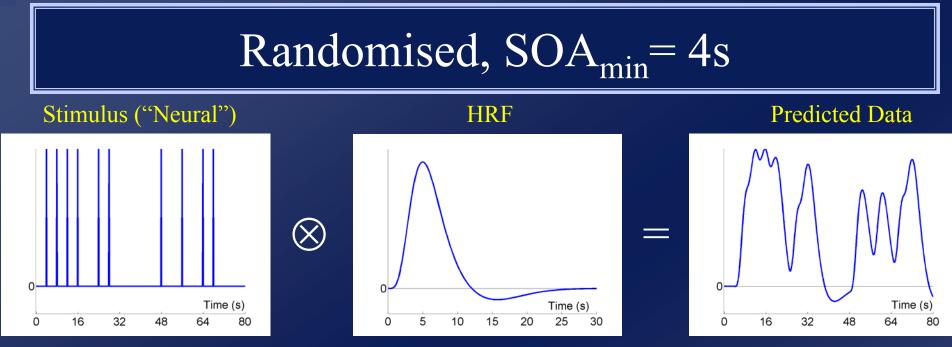
- 2. General Linear Model
- 3. Temporal Basis Functions
- 4. Timing Issues
- 5. Design Optimisation
- 6. Nonlinear Models
- 7. Example Applications



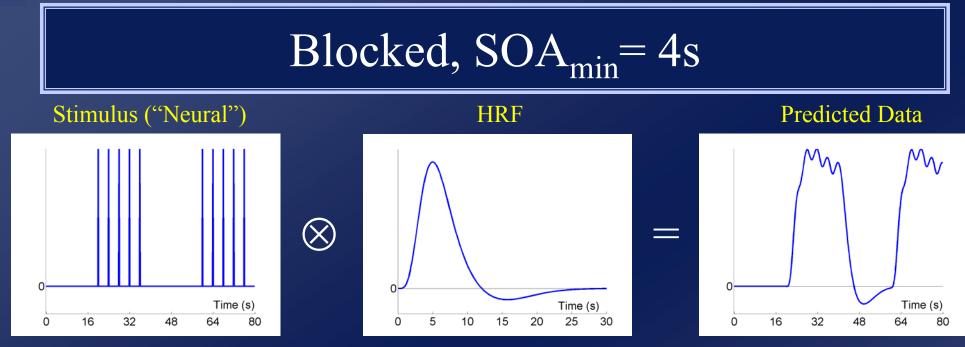
Not particularly efficient...



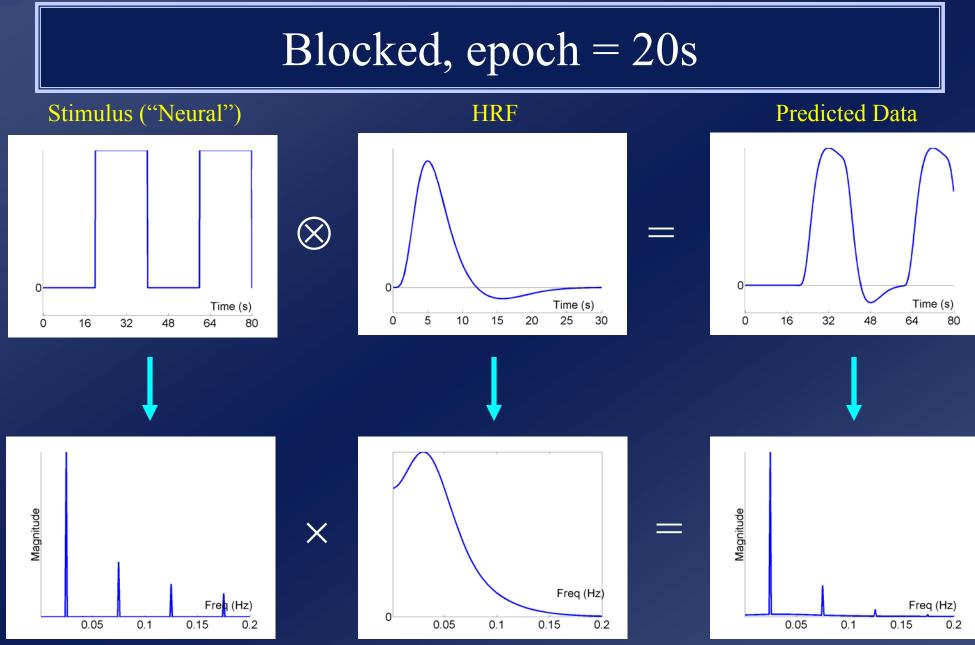
Very Inefficient...



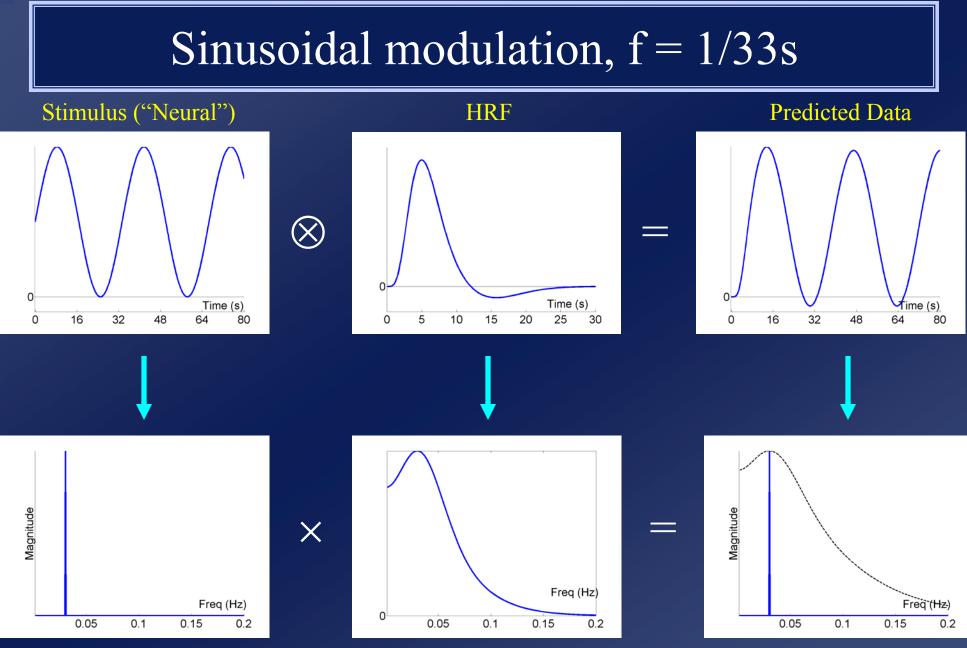
More Efficient...



Even more Efficient...

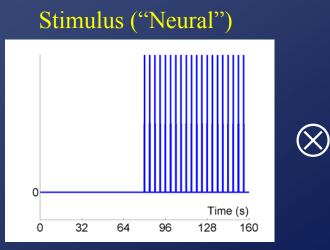


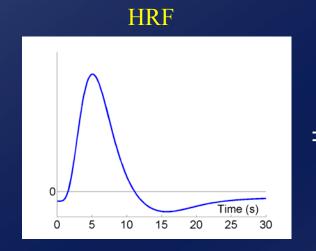
Blocked-epoch (with small SOA) and Time-Freq equivalences

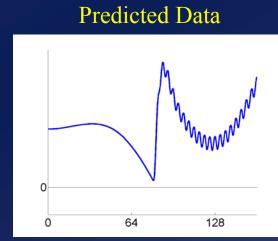


The most efficient design of all!

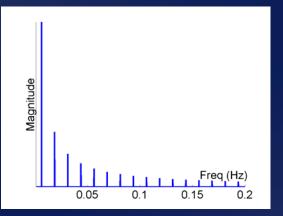
Blocked (80s), SOA_{min} =4s, highpass filter = 1/120s

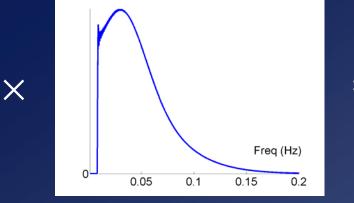


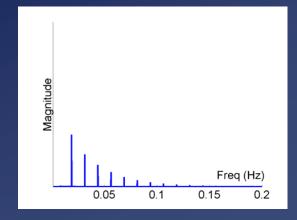




"Effective HRF" (after highpass filtering) (Josephs & Henson, 1999)

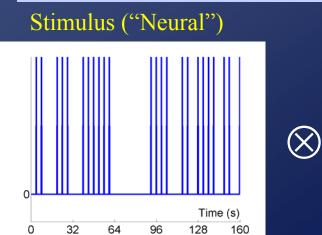


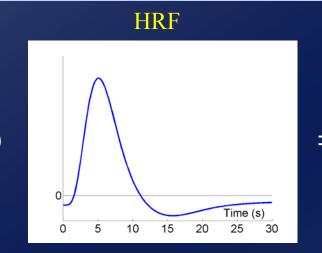




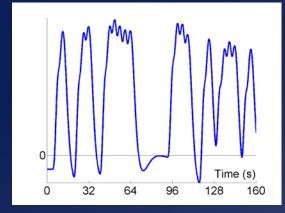
Don't have long (>60s) blocks!

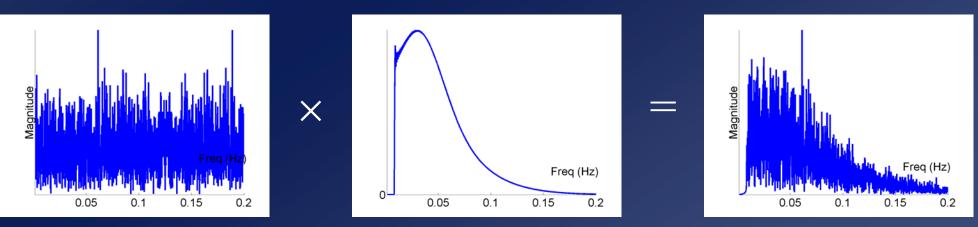
Randomised, SOA_{min} =4s, highpass filter = 1/120s





Predicted Data





(Randomised design spreads power over frequencies)

Design Efficiency

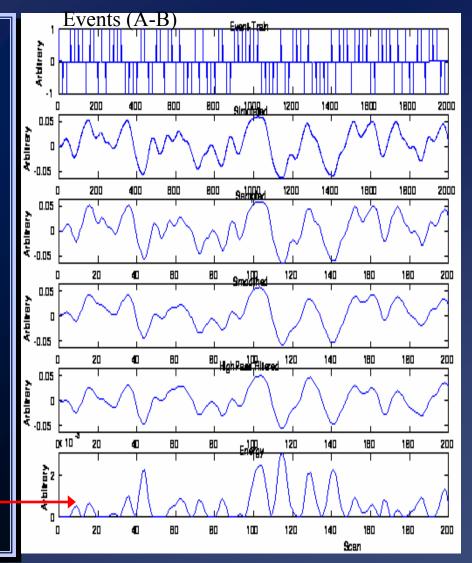
 $\mathbf{T} = \mathbf{c}^{\mathrm{T}}\boldsymbol{\beta} / \operatorname{var}(\mathbf{c}^{\mathrm{T}}\boldsymbol{\beta})$

 $Var(c^{T}\beta) = sqrt(\sigma^{2}c^{T}(X^{T}X)^{-1}c) \quad (i.i.d)$

- For max. T, want min. contrast variability (Friston et al, 1999)
- If assume that noise variance (σ²) is unaffected by changes in X...
- ...then want maximal efficiency, e:

 $e(c,X) = \{ \mathbf{c}^{\mathrm{T}} (\mathbf{X}^{\mathrm{T}} \mathbf{X})^{-1} \mathbf{c} \}^{-1}$

 = maximal bandpassed signal energy (Josephs & Henson, 1999)



- Design parametrised by:
 - SOA_{min} Minimum SOA

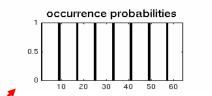
• Design parametrised by:

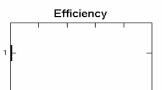
 SOA_{min} Minimum SOAp(t)Probability of eventat each SOA_{min}

• Design parametrised by:

 SOA_{min} Minimum SOAp(t)Probability of eventat each SOA_{min}

• Deterministic p(t)=1 iff t=nT

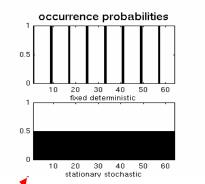


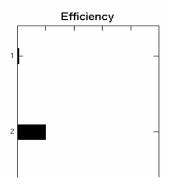


• Design parametrised by:

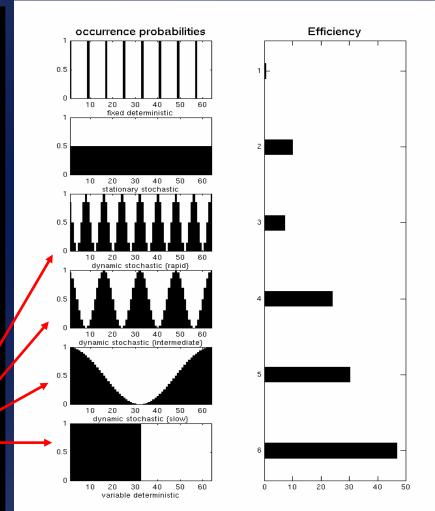
 SOA_{min} Minimum SOAp(t)Probability of eventat each SOA_{min}

- Deterministic $p(t)=1 \text{ iff } t=nSOA_{min}$
- Stationary stochastic p(t)=constant < 1





- Design parametrised by:
 - SOA_{min} Minimum SOAp(t)Probability of eventat each SOA_{min}
- Deterministic p(t)=1 iff t=nT
- Stationary stochastic p(t)=constant
- Dynamic stochastic p(t) varies (eg blocked)



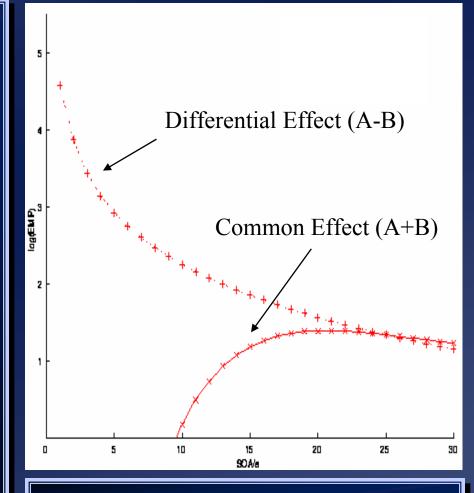
Blocked designs most efficient! (with small SOAmin)

Efficiency - Multiple Event-types

- Design parametrised by: SOA_{min} Minimum SOA
 p_i(h) Probability of event-type
 i given history h of last m events
- With *n* event-types *p_i(h)* is a *n^m × n Transition Matrix*
- Example: Randomised AB

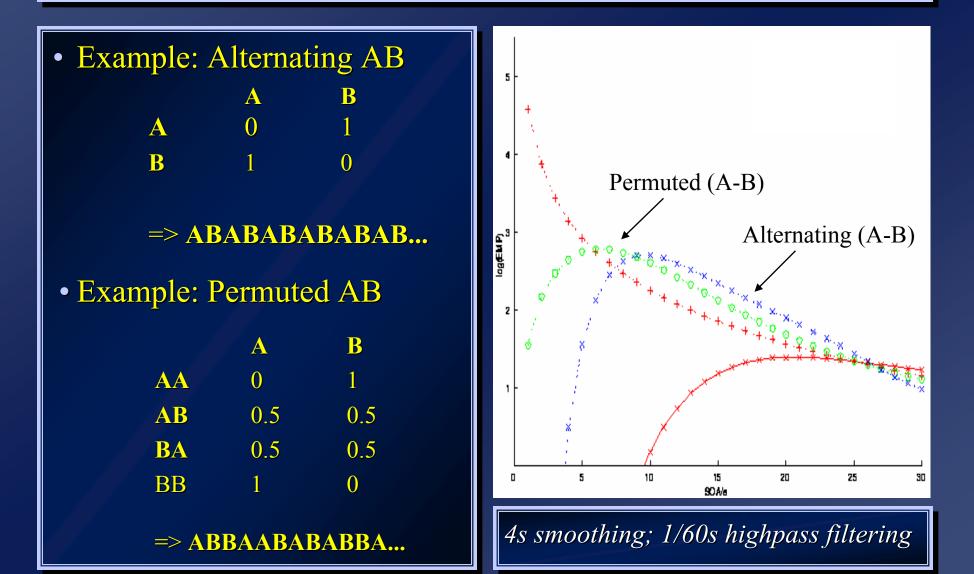
	Α	B
Α	0.5	0.5
B	0.5	0.5

```
=> ABBBABAABABAAAA...
```



4s smoothing; 1/60s highpass filtering

Efficiency - Multiple Event-types



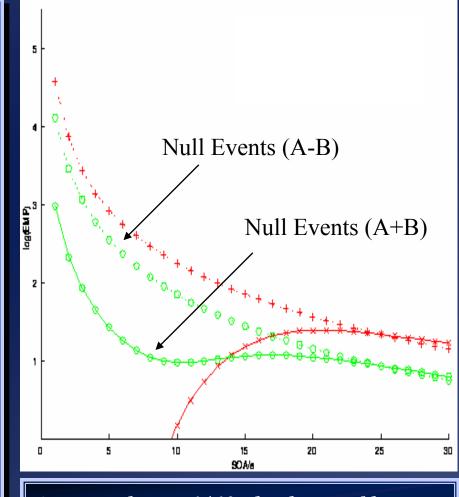
Efficiency - Multiple Event-types

• Example: Null events

	A	B
A	0.33	0.33
B	0.33	0.33

=> **AB-BAA--B---ABB...**

- Efficient for differential *and* main effects at short SOA
- Equivalent to stochastic SOA (Null Event like third unmodelled event-type)
- Selective averaging of data (Dale & Buckner 1997)



4s smoothing; 1/60s highpass filtering

Efficiency - Conclusions

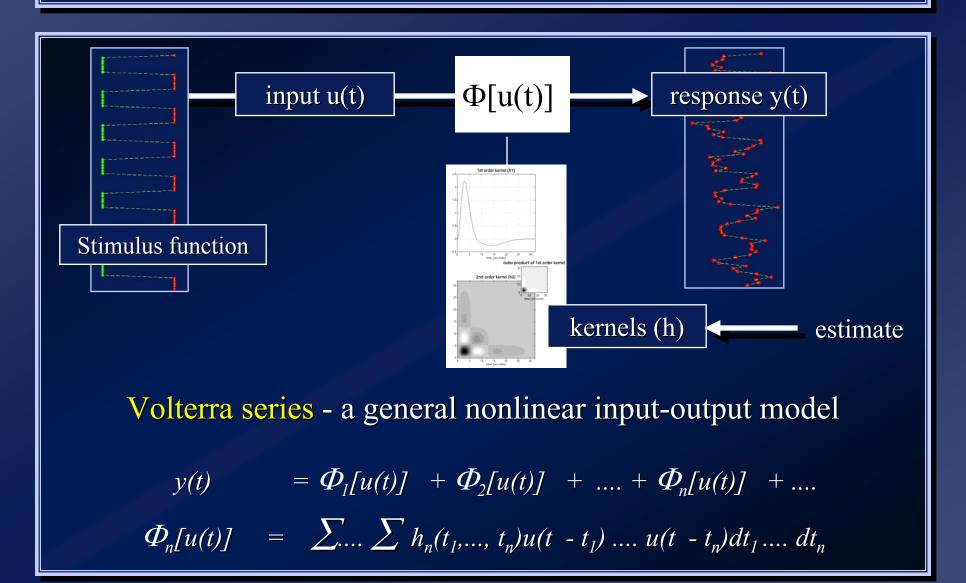
- Optimal design for one contrast may not be optimal for another
- Blocked designs generally most efficient with short SOAs (but earlier restrictions and problems of interpretation...)
- With randomised designs, optimal SOA for differential effect (A-B) is minimal SOA (assuming no saturation), whereas optimal SOA for main effect (A+B) is 16-20s
- Inclusion of null events improves efficiency for main effect at short SOAs (at cost of efficiency for differential effects)
- If order constrained, intermediate SOAs (5-20s) can be optimal; If SOA constrained, pseudorandomised designs can be optimal (but may introduce context-sensitivity)

Overview

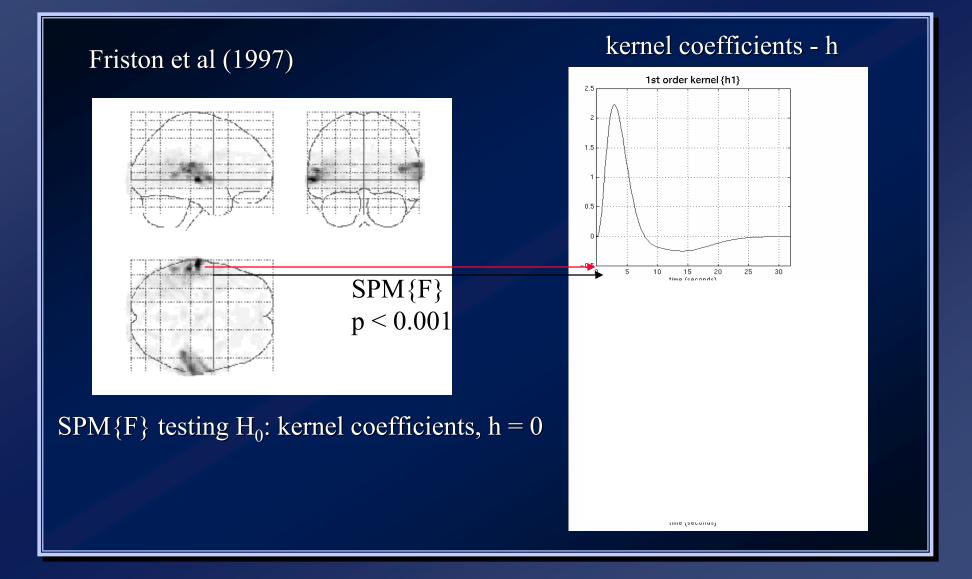
1. BOLD impulse response

- 2. General Linear Model
- 3. Temporal Basis Functions
- 4. Timing Issues
- 5. Design Optimisation
- 6. Nonlinear Models
- 7. Example Applications

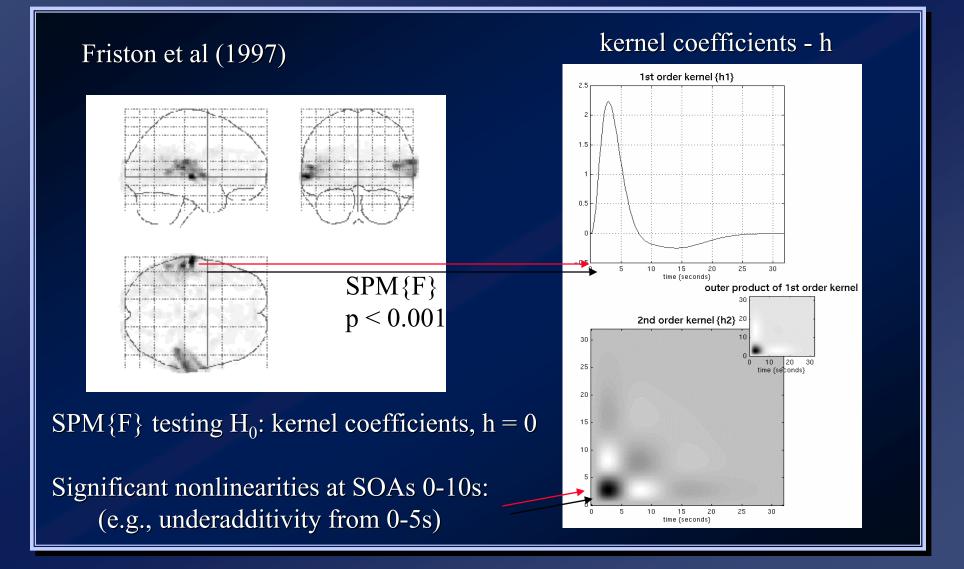
Nonlinear Model



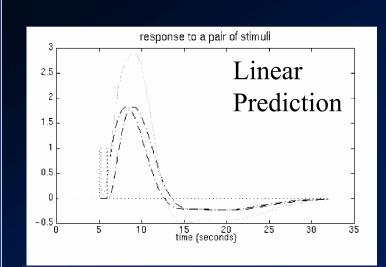
Nonlinear Model



Nonlinear Model

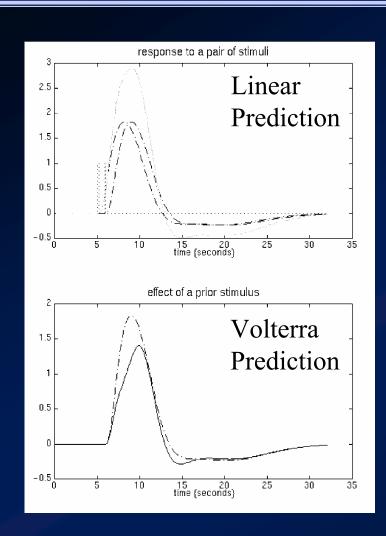


Nonlinear Effects



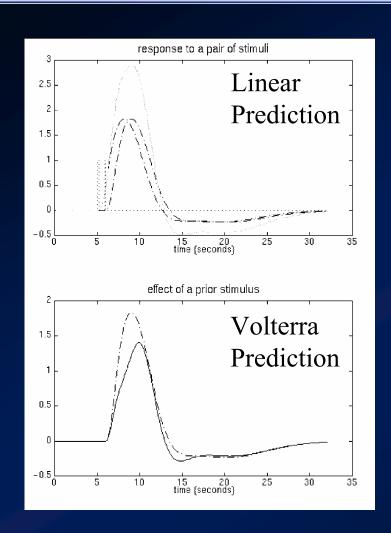
Underadditivity at short SOAs

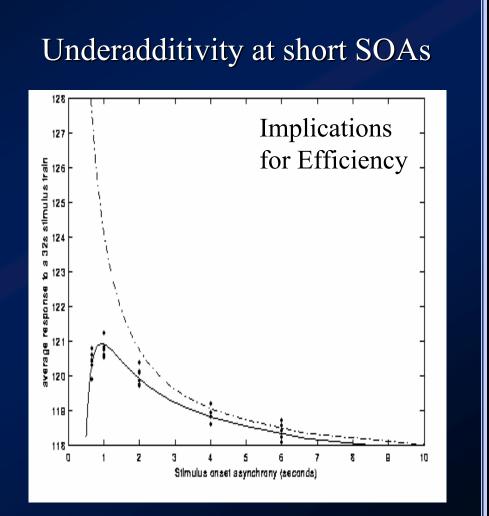
Nonlinear Effects



Underadditivity at short SOAs

Nonlinear Effects





Overview

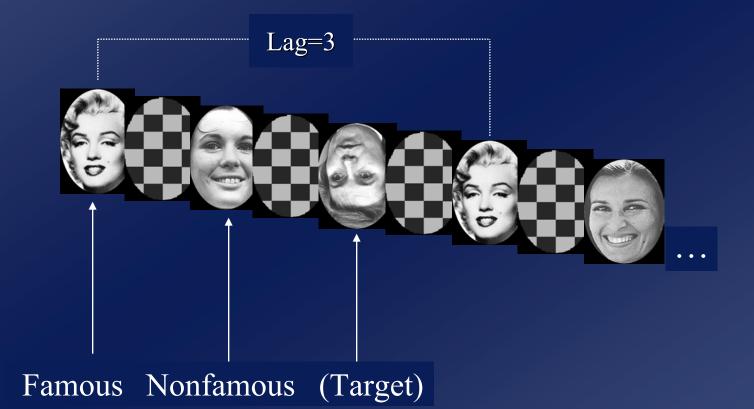
1. BOLD impulse response

- 2. General Linear Model
- 3. Temporal Basis Functions
- 4. Timing Issues
- 5. Design Optimisation
- 6. Nonlinear Models
- 7. Example Applications

Example 1: Intermixed Trials (Henson et al 2000)

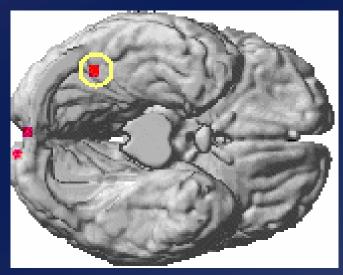
- Short SOA, fully randomised, with 1/3 null events
- Faces presented for 0.5s against chequerboard baseline, SOA=(2 ± 0.5)s, TR=1.4s
- Factorial event-types:
 1. Famous/Nonfamous (F/N)
 - 2. 1st/2nd Presentation (1/2)

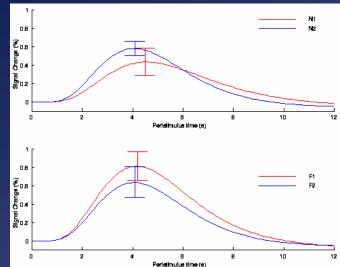
FIL



Example 1: Intermixed Trials (Henson et al 2000)

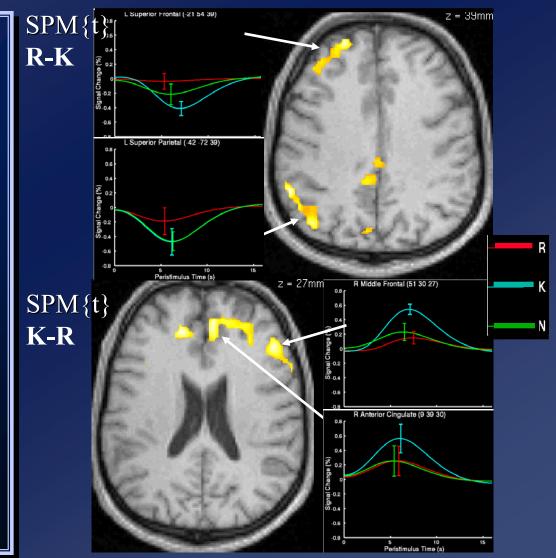
- Short SOA, fully randomised, with 1/3 null events
- Faces presented for 0.5s against chequerboard baseline, SOA=(2 ± 0.5)s, TR=1.4s
- Factorial event-types:
 1. Famous/Nonfamous (F/N)
 2. 1st/2nd Presentation (1/2)
- Interaction (F1-F2)-(N1-N2) masked by main effect (F+N)
- Right fusiform interaction of repetition priming and familiarity





Example 2: Post hoc classification (Henson et al 1999)

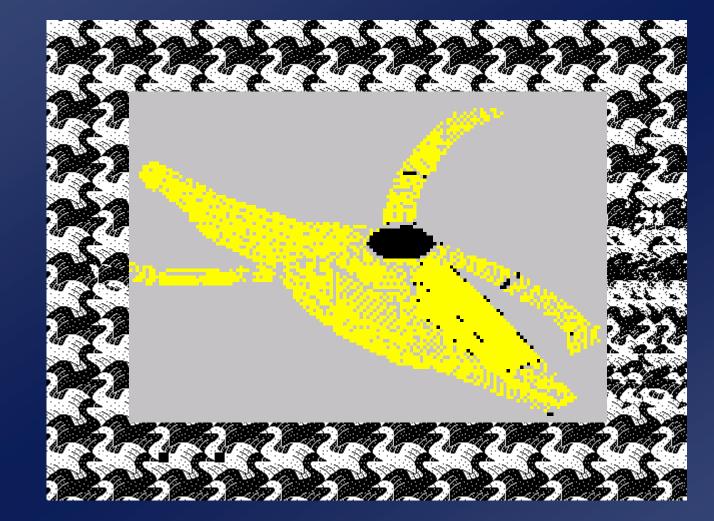
- Subjects indicate whether studied (Old) words:
 i) evoke recollection of prior occurrence (R)
 ii) feeling of familiarity without recollection (K)
 iii) no memory (N)
- Random Effects analysis on canonical parameter estimate for event-types
- Fixed SOA of 8s => sensitive to differential but not main effect (de/activations arbitrary)



FIL

Example 3: Subject-defined events (Portas et al 1999)

 Subjects respond when "pop-out" of 3D percept from 2D stereogram



FIL

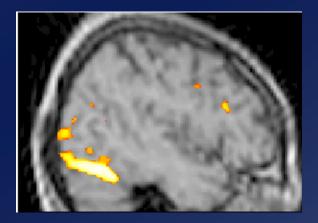
Example 3: Subject-defined events (Portas et al 1999)

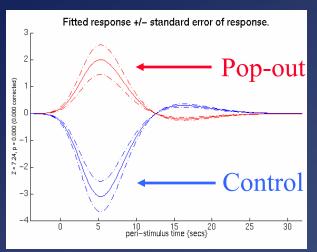
 Subjects respond when "pop-out" of 3D percept from 2D stereogram

FIL

- Popout response also produces tone
- Control event is response to tone during 3D percept

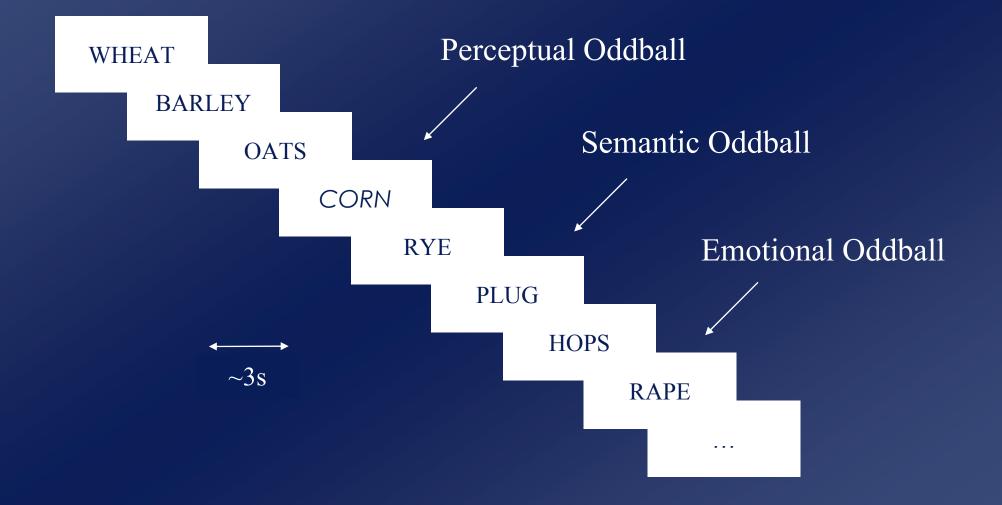
Temporo-occipital differential activation



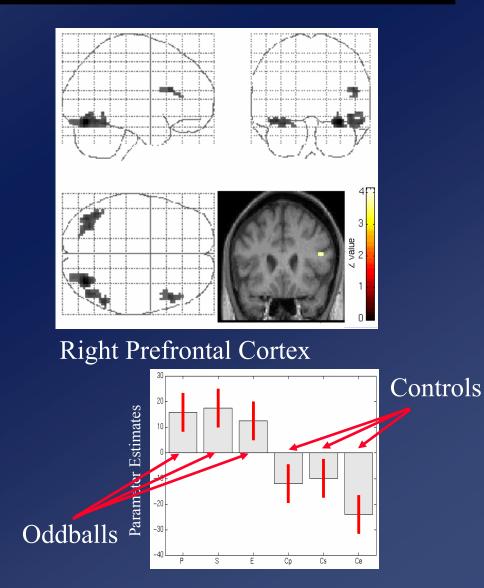


Example 4: Oddball Paradigm (Strange et al, 2000)

- 16 same-category words every 3 secs, plus ...
- ... 1 perceptual, 1 semantic, and 1 emotional oddball



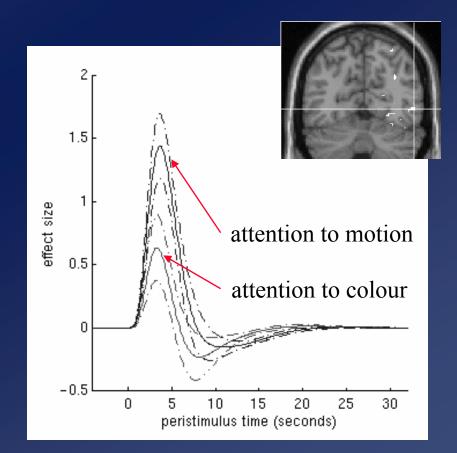
- 16 same-category words every 3 secs, plus ...
- ... 1 perceptual, 1 semantic, and 1 emotional oddball
- 3 nonoddballs randomly matched as controls
- Conjunction of oddball vs. control contrast images: generic deviance detector



Example 5: Epoch/Event Interactions (Chawla et al 1999)

- Epochs of attention to:
 1) motion, or 2) colour
- Events are target stimuli differing in motion or colour
- Randomised, long SOAs to decorrelate epoch and eventrelated covariates
- Interaction between epoch (attention) and event (stimulus) in V4 and V5

Interaction between attention and stimulus motion change in V5

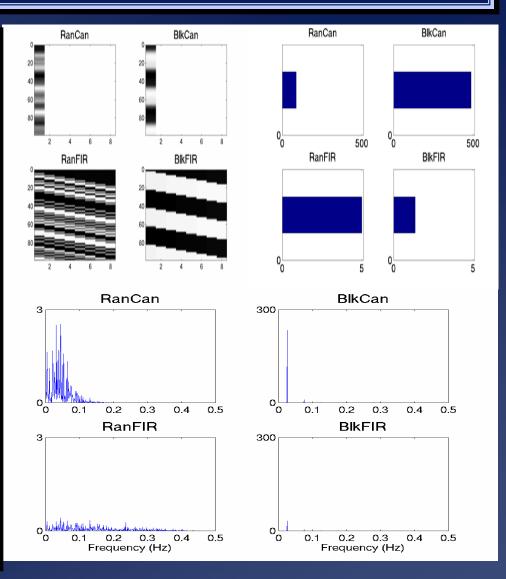


FIL

Efficiency – Detection vs Estimation

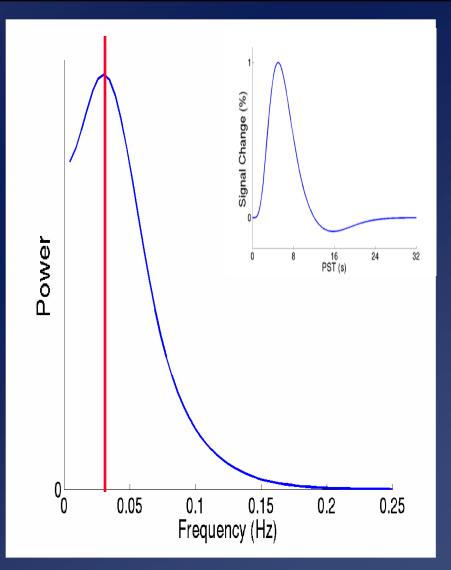
- "Detection power" vs "Estimation efficiency" (Liu et al, 2001)
- Detect response, or characterise shape of response?
- Maximal detection power in blocked designs;
 Maximal estimation efficiency in randomised designs
- => simply corresponds to choice of basis functions:

detection = canonical HRF estimation = FIR



Design Efficiency

- HRF can be viewed as a filter (Josephs & Henson, 1999)
- Want to maximise the signal passed by this filter
- Dominant frequency of canonical HRF is ~0.04 Hz
- So most efficient design is a sinusoidal modulation of neural activity with period ~24s
- (eg, boxcar with 12s on/ 12s off)



Timing Issues : Latency

• Assume the real response, r(t), is a scaled (by α) version of the canonical, f(t), but delayed by a small amount *dt*:

 $r(t) = \alpha f(t+dt) \sim \alpha f(t) + \alpha f'(t) dt$ 1st-order Taylor

- If the fitted response, R(t), is modelled by the canonical + temporal derivative: \mathbf{O} $\underline{R(t)} = \beta_1 f(t) + \beta_2 f'(t)$ GLM fit
- Then canonical and derivative parameter estimates, β_1 and β_2 are such that :

$$\Rightarrow \qquad \alpha = \beta_1 \qquad dt = \beta_2 / \beta_1 \qquad (Henson \ et \ al, \ 2002) \\ (Liao \ et \ al, \ 2002)$$

ie, Latency can be approximated by the ratio of derivative-to-canonical parameter estimates (within limits of first-order approximation, +/-1s)

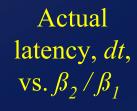
2002)

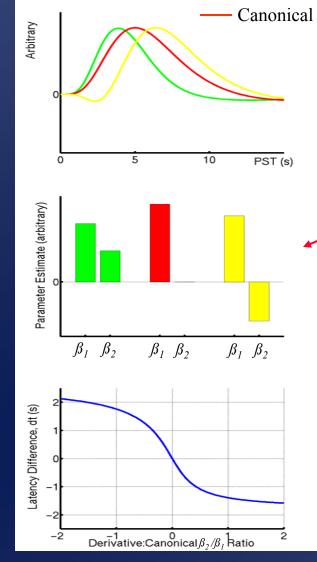
FIL

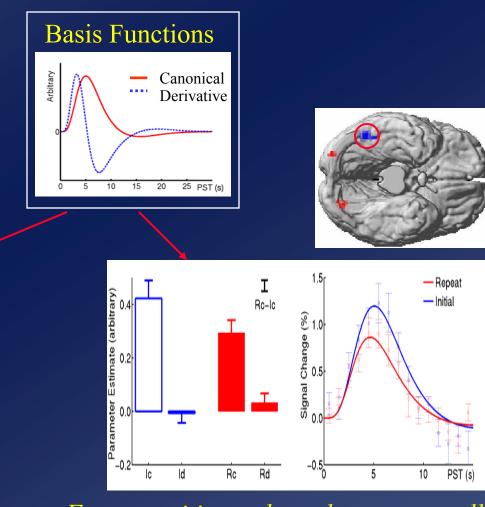
Timing Issues : Latency



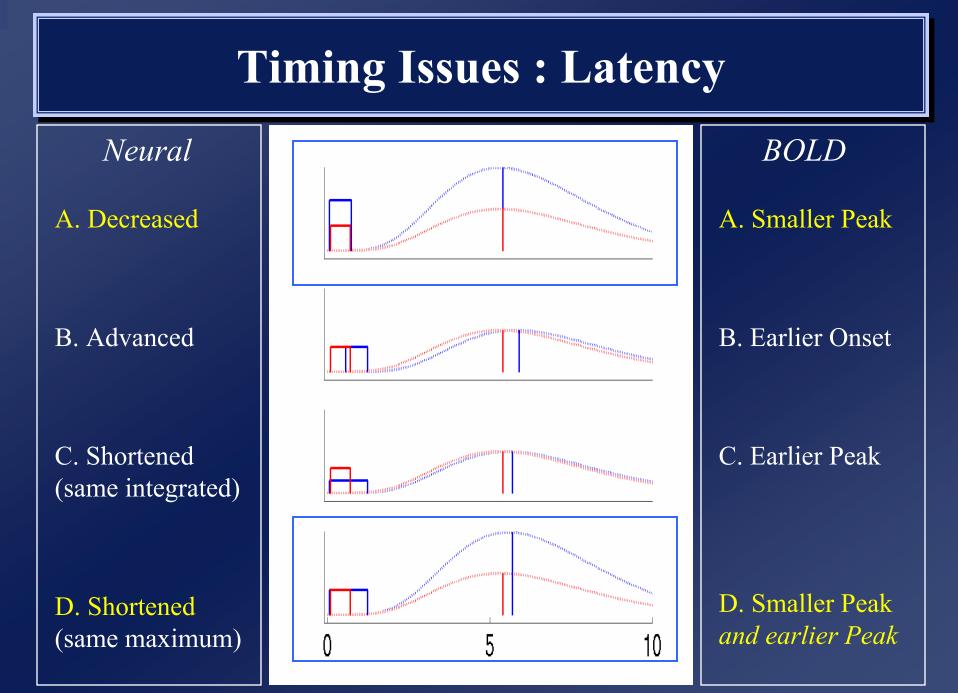
Parameter Estimates



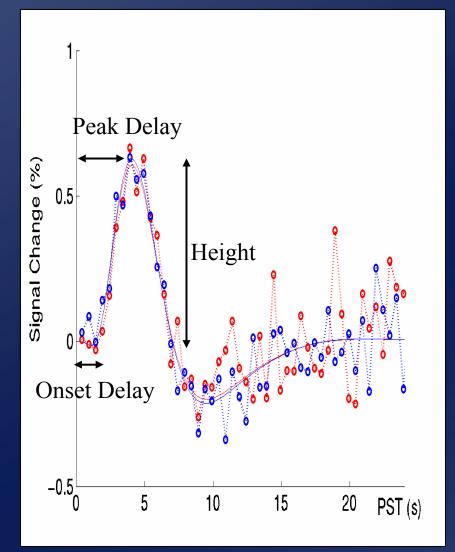




Face repetition reduces latency as well as magnitude of fusiform response

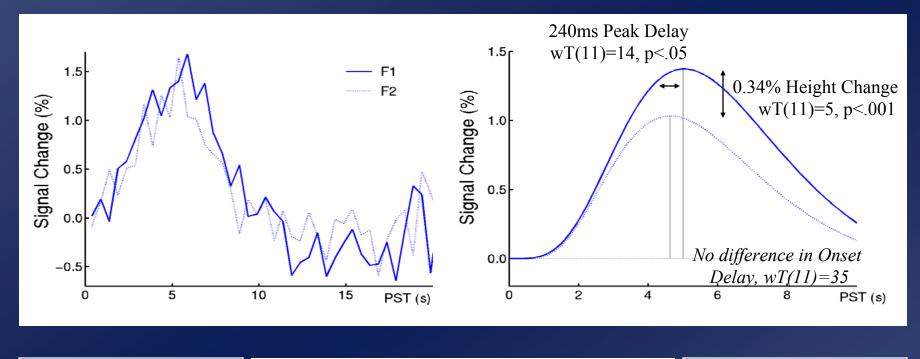


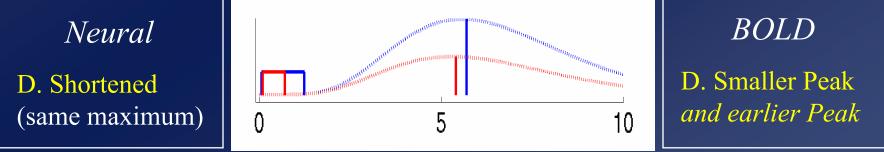
BOLD Response Latency (Iterative)



- Numerical fitting of explicitly parameterised canonical HRF *(Henson et al, 2001)*
- Distinguishes between *Onset* and *Peak* latency...
 - ...unlike temporal derivative...
 ...and which may be important for interpreting neural changes (see previous slide)
- Distribution of parameters tested nonparametrically (Wilcoxon's T over subjects)

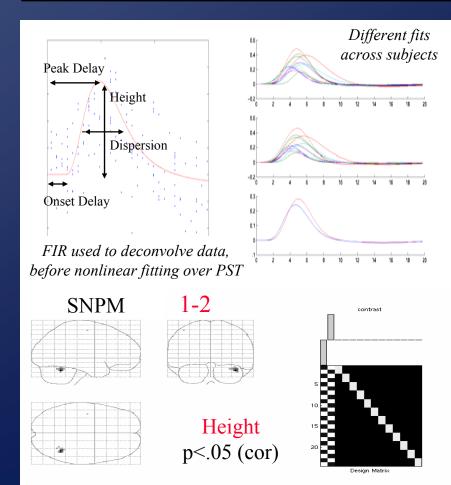
BOLD Response Latency (Iterative)





Most parsimonious account is that repetition reduces duration of neural activity...

BOLD Response Latency (Iterative)



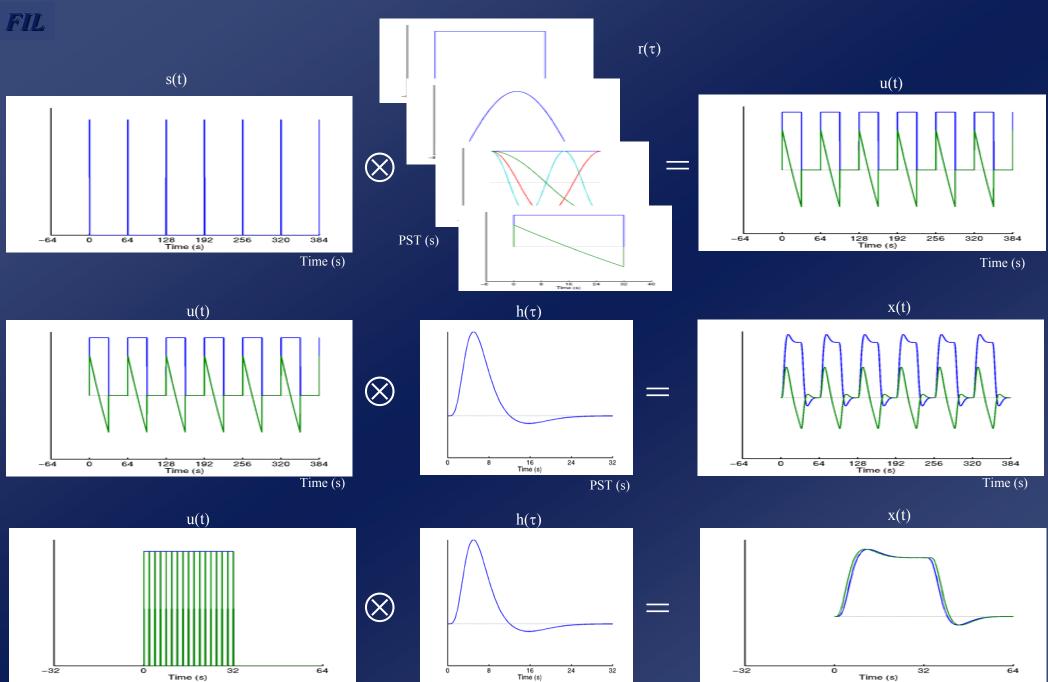
P values & statistics: ./Methods/NonLinFit/SNPM/12/Height

region	size {k}	t	P(T ∎u≩= u)	(uncorrected)	{×,y,z} mm
1	23	10.37	0.001	(0.000000)	39 -54 -27
		6.39	0.020	(0.000026)	33 -60 -24
2	1	5.91	0.033	(0.000051)	-42 -12 60

- Four-parameter HRF, nonparametric Random Effects (SNPM99)
- Advantages of iterative vs linear:
 - 1. Height "independent" of shape Canonical "height" confounded by latency (e.g, different shapes across subjects); no slice-timing error
 - 2. Distinction of onset/peak latency *Allowing better neural inferences?*
- Disadvantages of iterative:
 - 1. Unreasonable fits (onset/peak tension) *Priors on parameter distributions? (Bayesian estimation)*
 - Local minima, failure of convergence?
 CDLL time (2 days for above)
 - 3. CPU time (~3 days for above)

Temporal Basis Sets: Inferences

- How can inferences be made in hierarchical models (eg,
 "Random Effects" analyses over, for example, subjects)?
 - **1. Univariate T-tests on canonical parameter alone?** *may miss significant experimental variability canonical parameter estimate not appropriate index of "magnitude" if real responses are non-canonical (see later)*
- **2. Univariate F-tests on parameters from multiple basis functions?** *need appropriate corrections for nonsphericity (Glaser et al, 2001)*
- **3. Multivariate tests (eg Wilks Lambda, Henson et al, 2000)** not powerful unless ~10 times as many subjects as parameters



PST (s)

Time (s)

Time (s)

Time (s)

Time (s)

