fMRI 2: single participant GLM

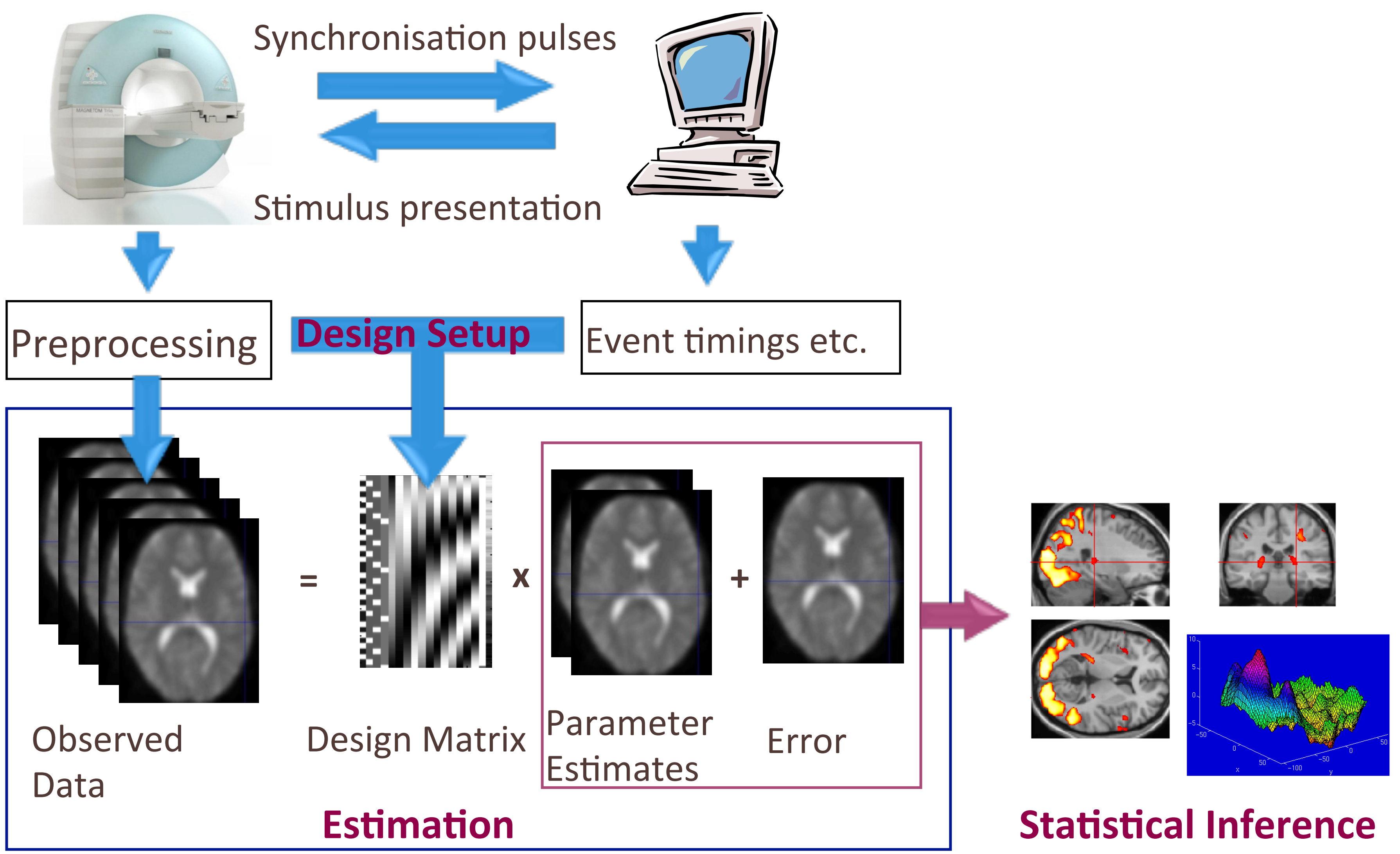
Johan Carlin
MRC Cognition and Brain Sciences Unit, Methods Group
johan.carlin@mrc-cbu.cam.ac.uk

Acknowledgment: Tibor Auer, Kendrick Kay

MRC Cognition and Brain Sciences Unit

Today

- How to fit fMRI responses with a general linear model
- Workshop: Building an fMRI model by hand in Matlab
- Implications for experimental design



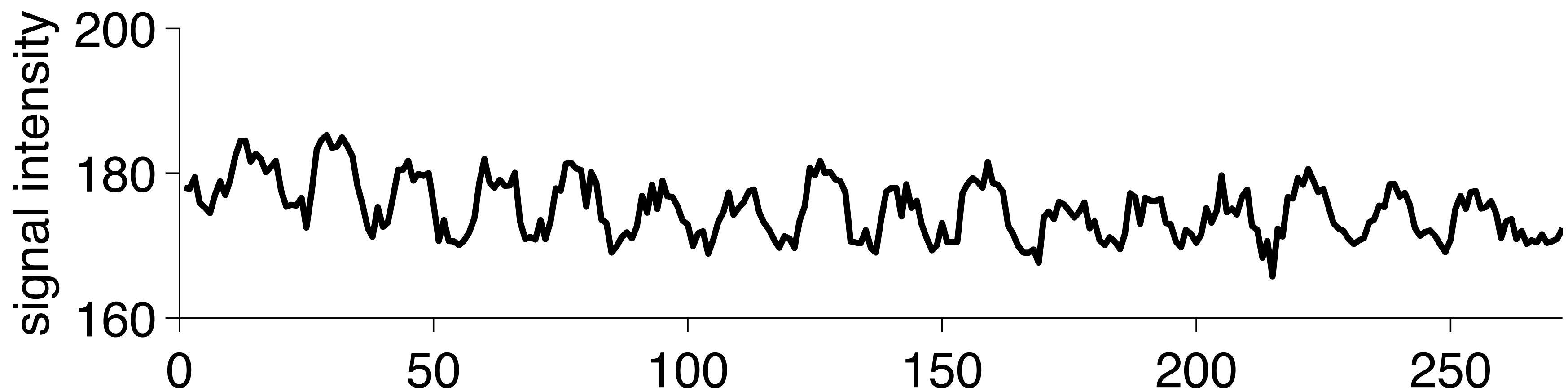
last session today next session

The general linear model

$$\text{i observations}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \dots \\ y_i \end{bmatrix} = \begin{array}{c} \textit{j predictor variables} \\ \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1j} \\ x_{21} & x_{22} & \dots & x_{2j} \\ x_{31} & x_{32} & \dots & x_{3j} \\ \dots & \dots & \dots & \dots \\ x_{i1} & x_{i2} & \dots & x_{ij} \end{bmatrix} \end{array} \text{x} \quad \textit{j predictor weights}\begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \dots \\ \beta_j \end{bmatrix} + \textit{i noises}\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \dots \\ \varepsilon_i \end{bmatrix}$$

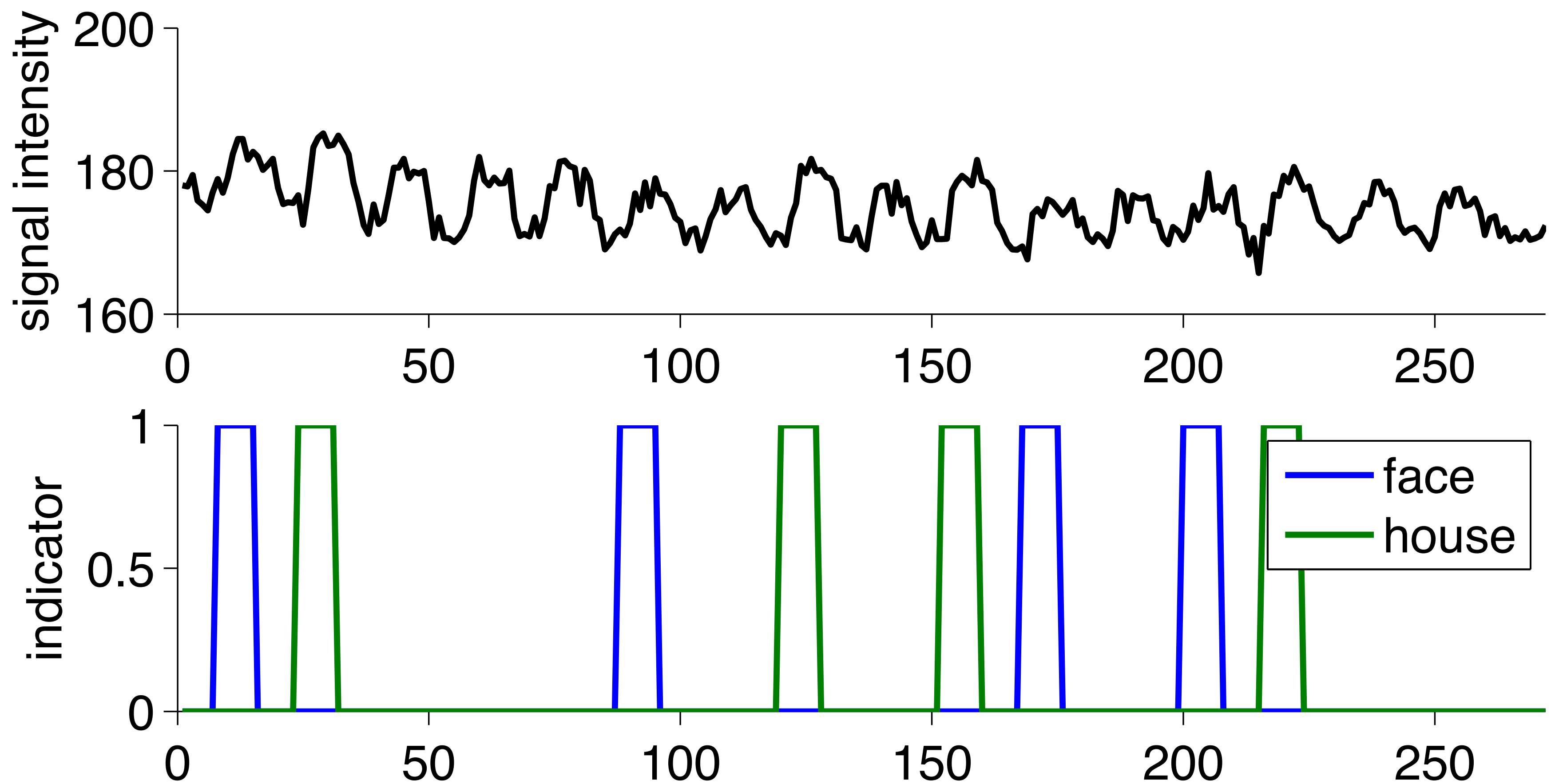
$$\mathbf{Y = X\beta + \varepsilon}$$

A typical fMRI task



- Face blocks appeared at volumes 8, 88, 168, 200
- House blocks at volumes 24, 120, 152, 216
- All blocks lasted 8 volumes

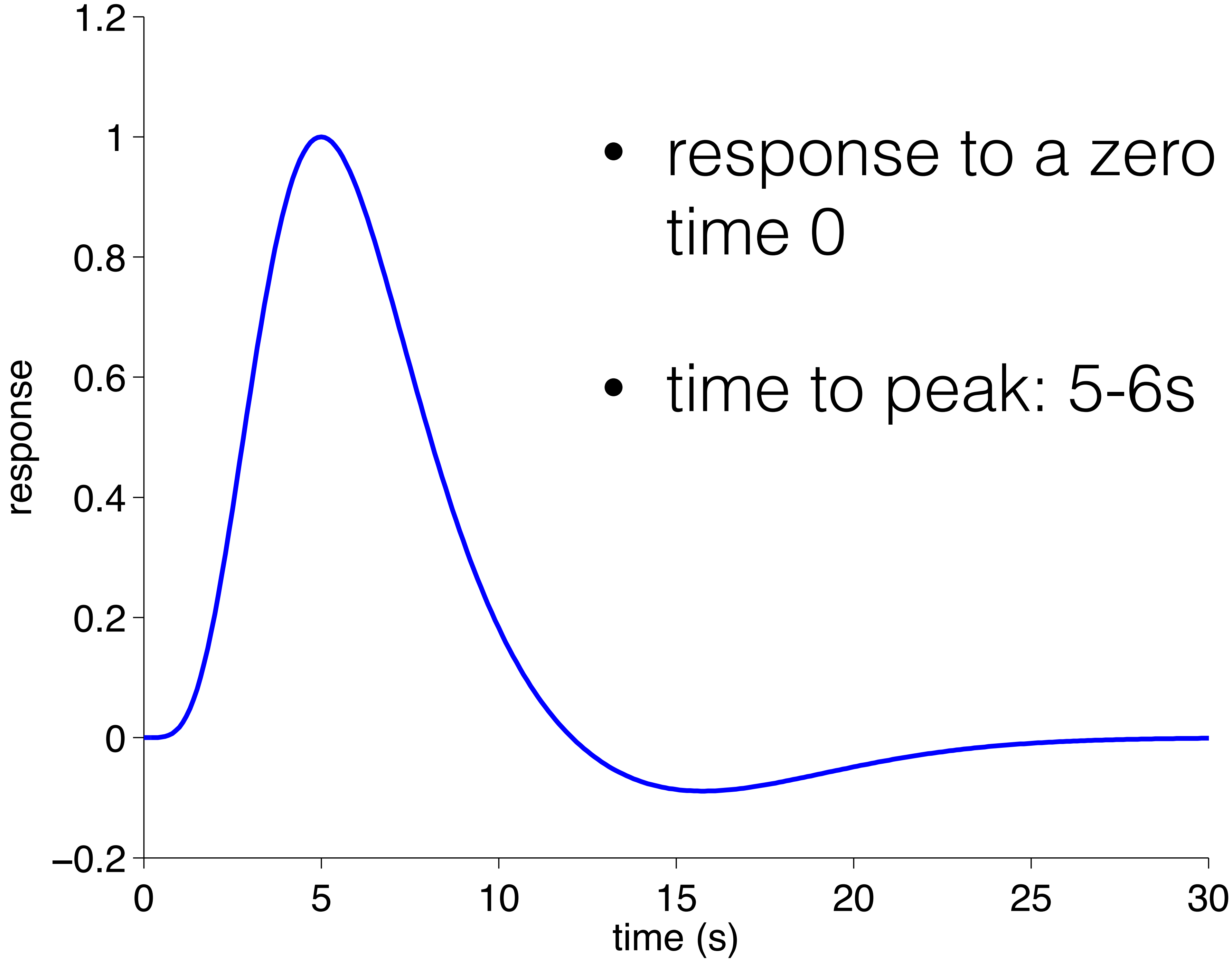
A typical fMRI task



But fMRI responses are not instantaneous…

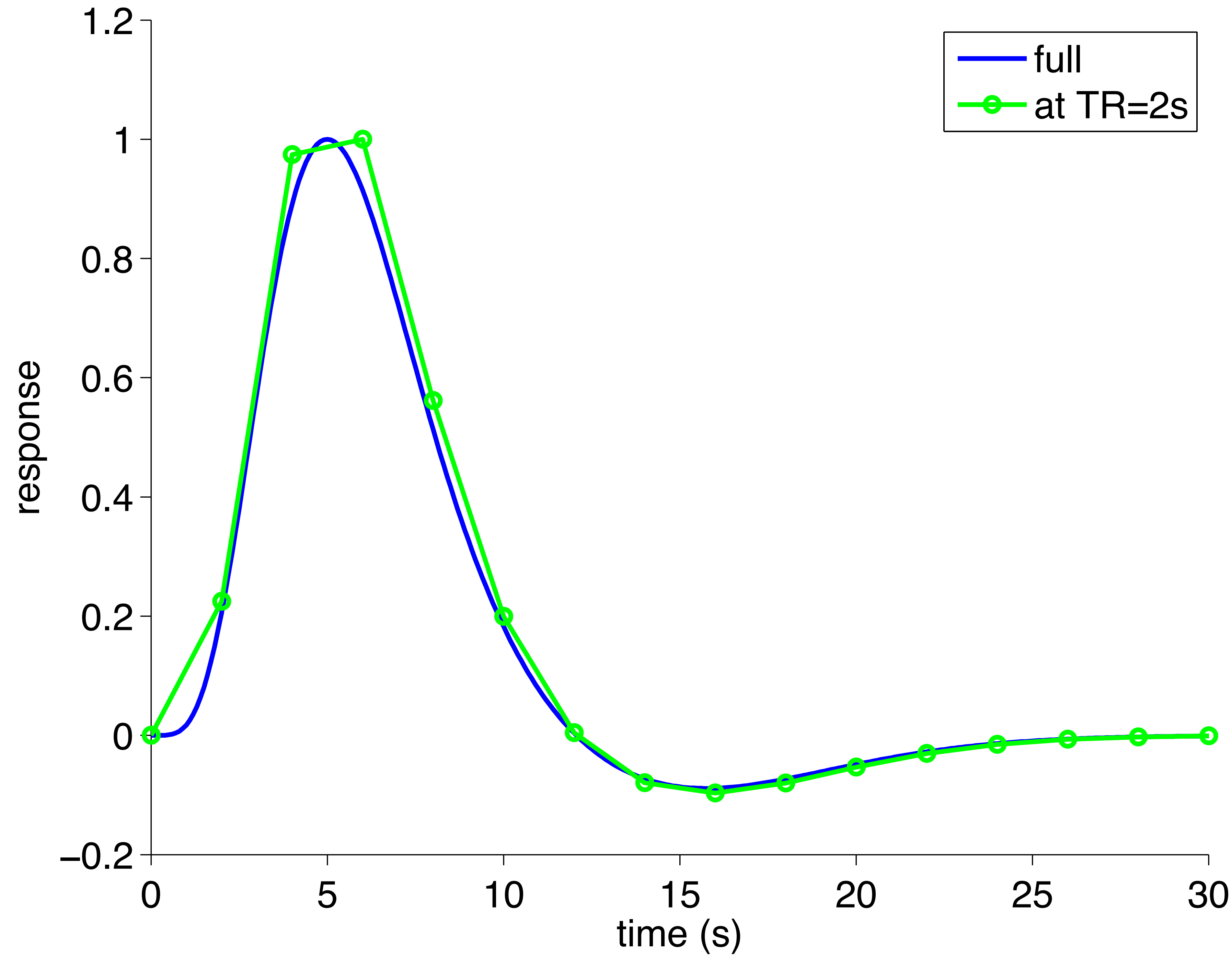
time (scans)

The canonical SPM HRF

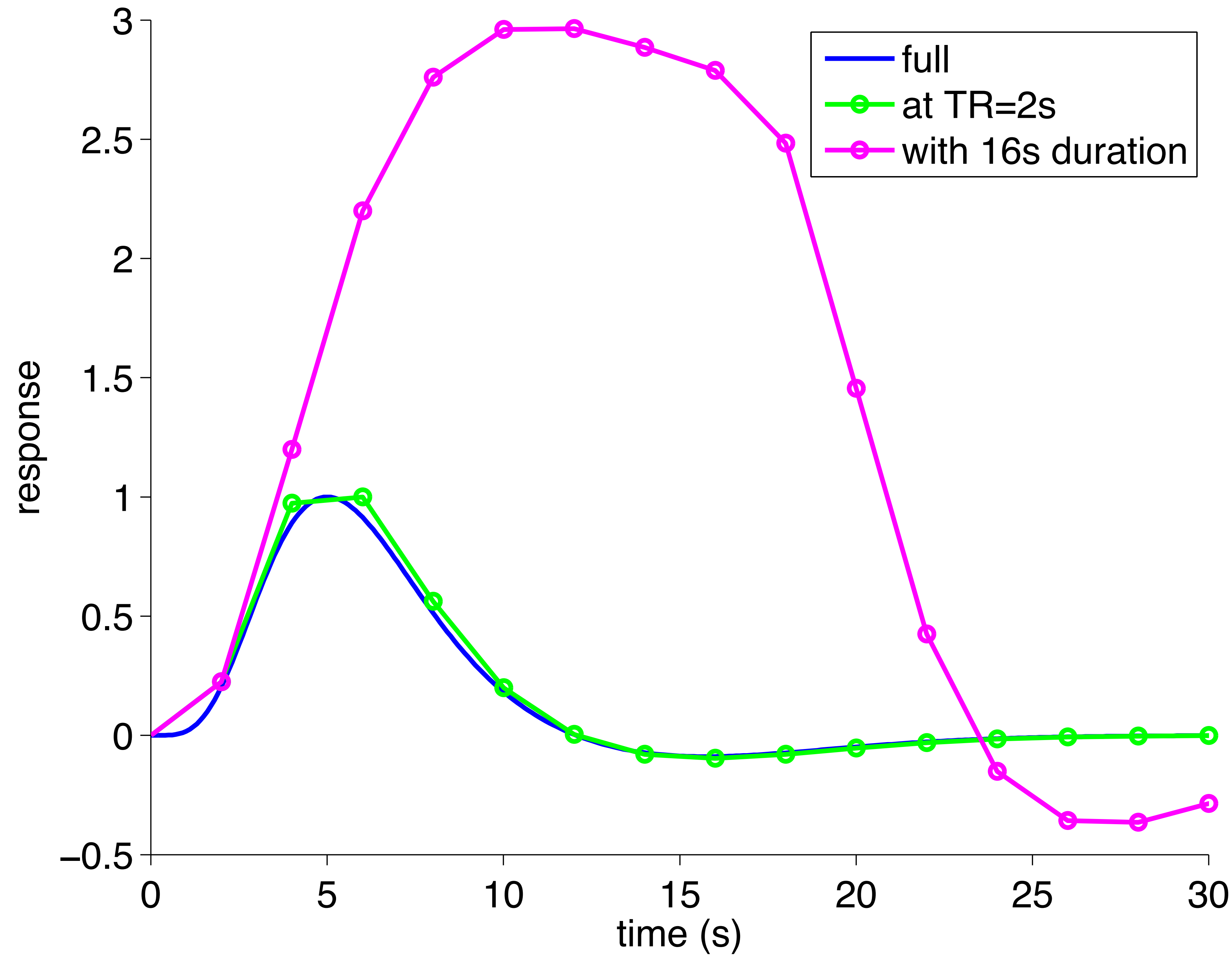


- response to a zero duration stimulus at time 0
- time to peak: 5-6s

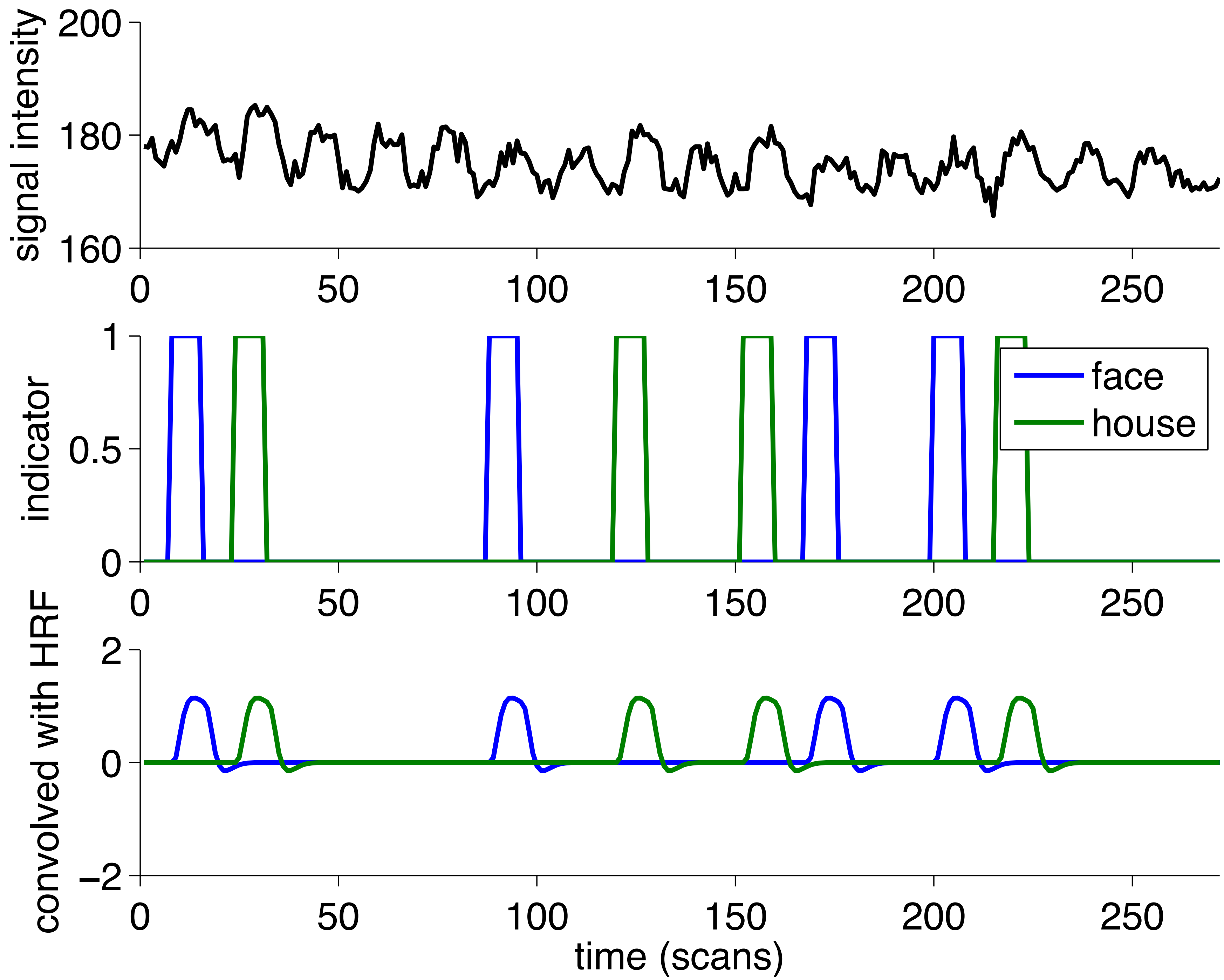
The canonical SPM HRF



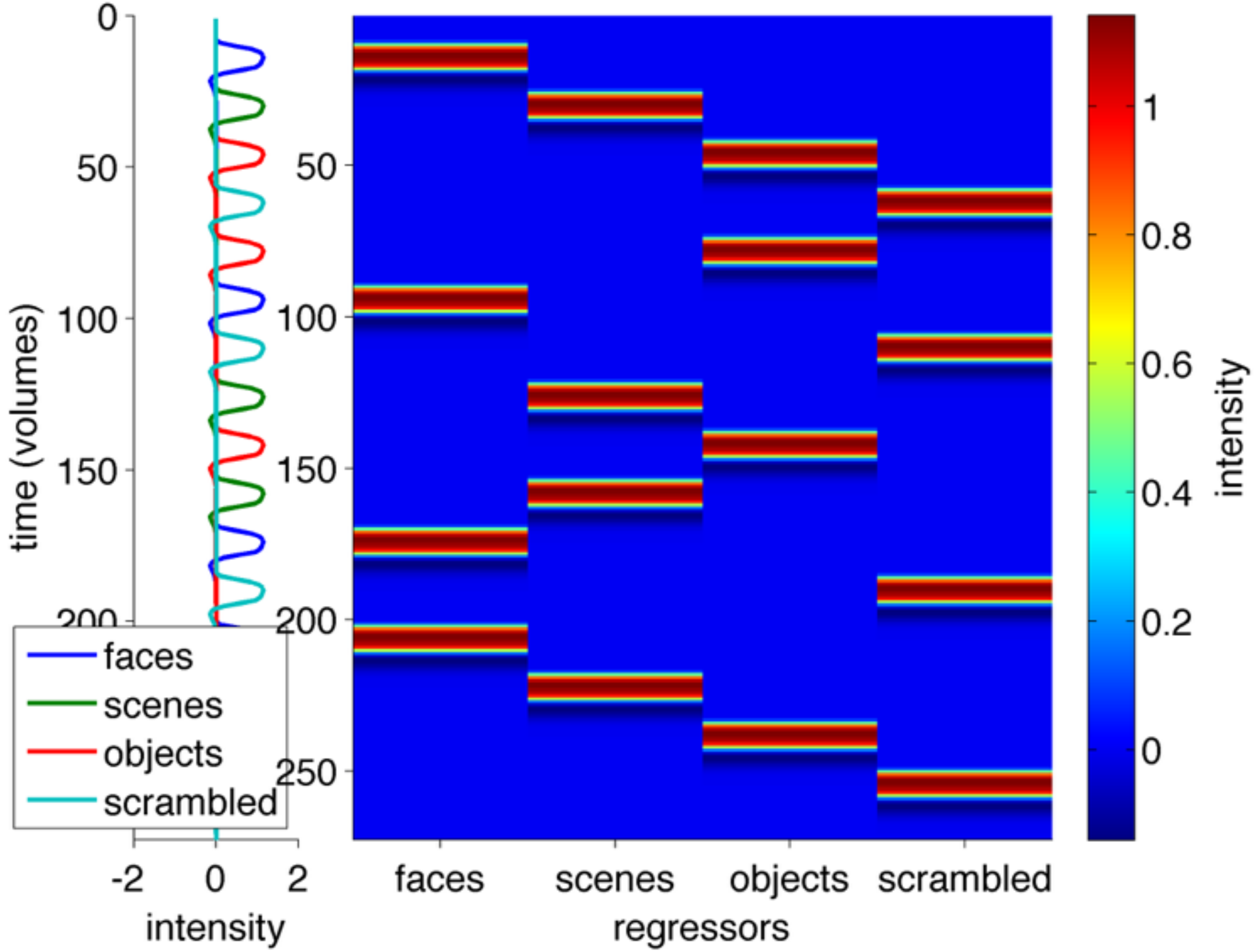
The canonical SPM HRF



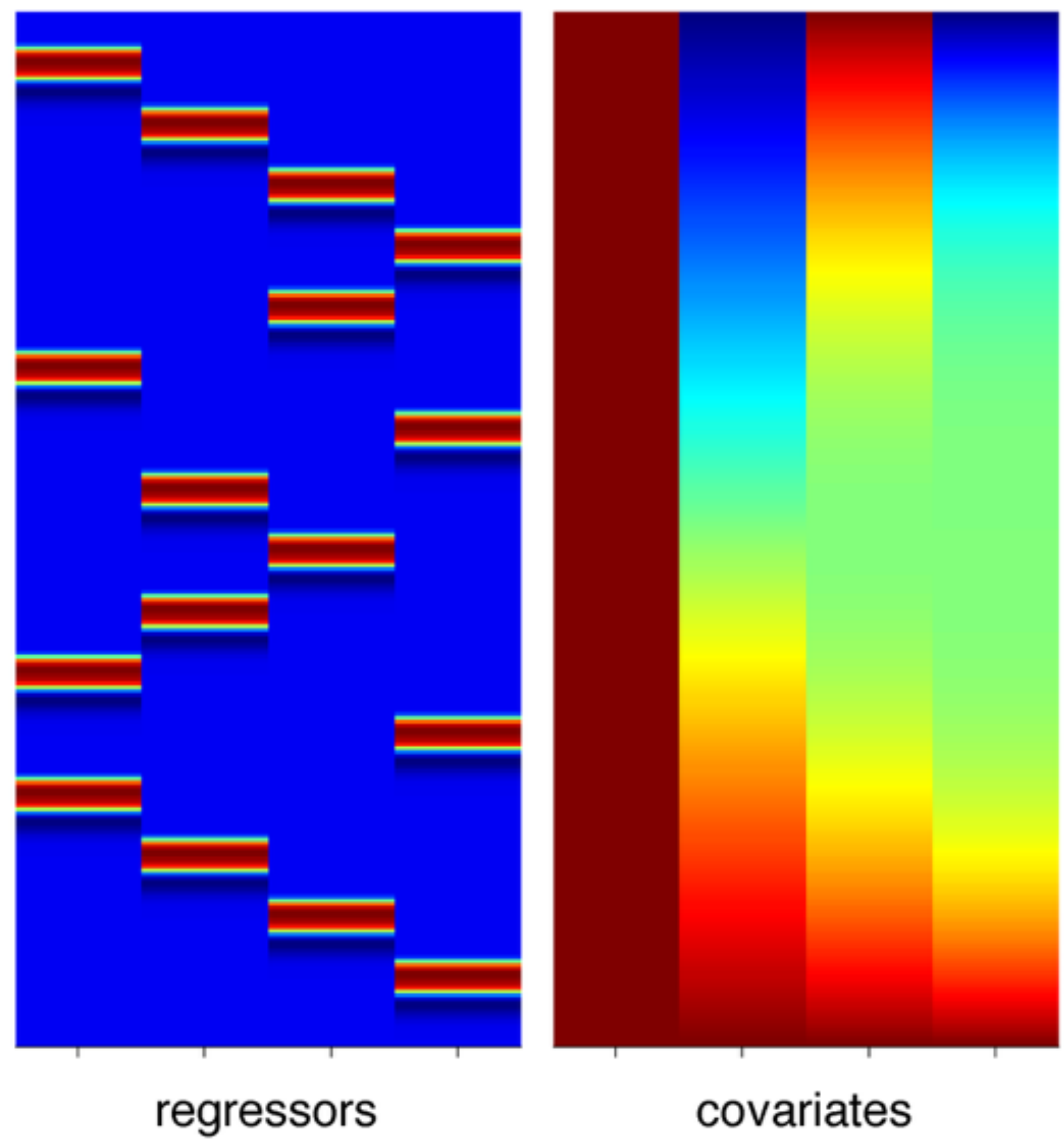
A typical fMRI task



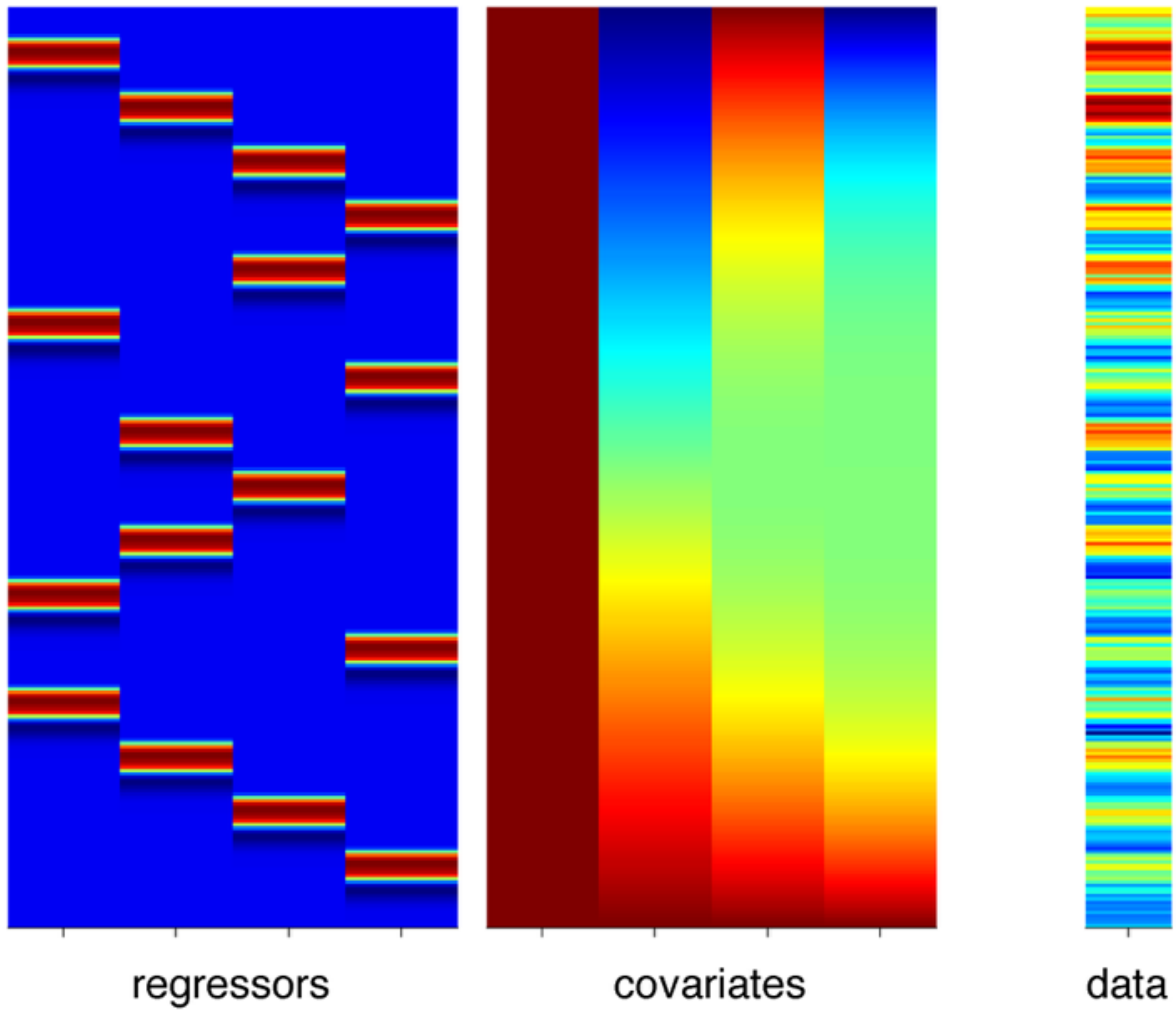
The convolved design matrix

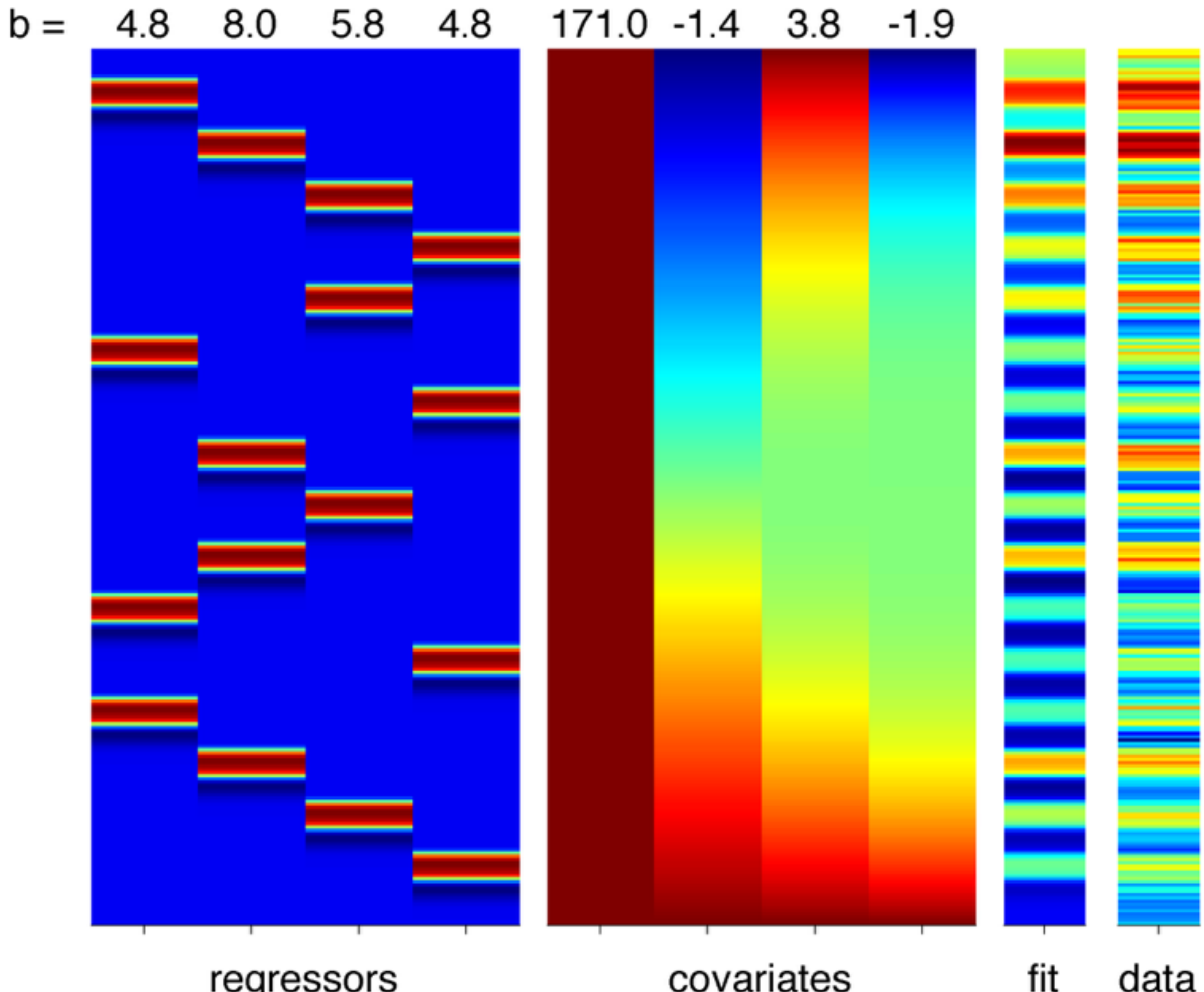


Trend removal



What weights should we use to obtain a fitted timecourse that is as close as possible to the data (ie, smallest squared deviation)?

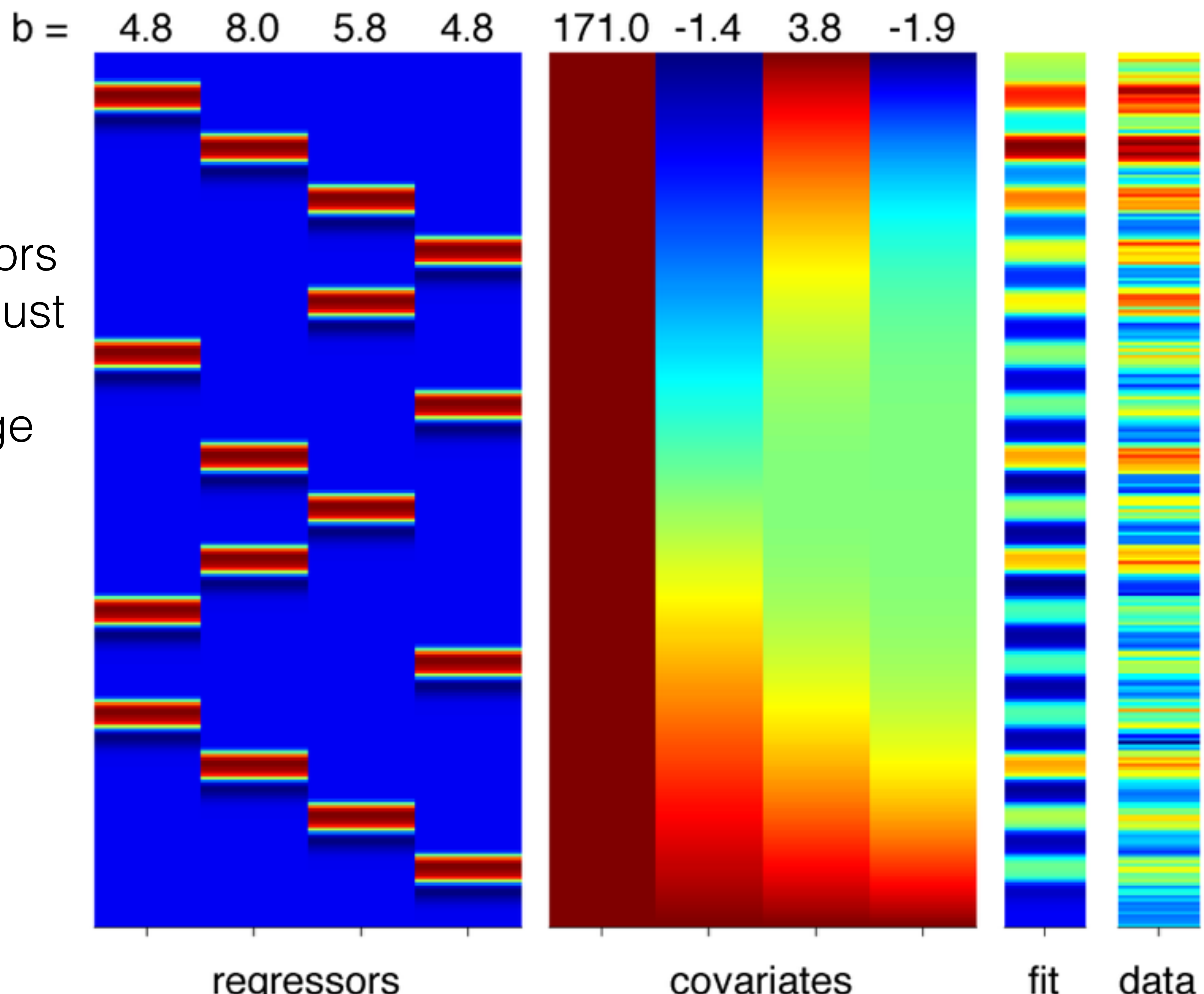



b =
4.8
8.0
5.8
4.8
171.0
-1.4
3.8
-1.9
regressors
covariates
fit
data

b * [-1 1 0 0 0 0 0 0]' = 3.2 = 8.0-4.8

b * [.25 .25 .25 .25 0 0 0 0]' = 5.8 = mean(b_regressors)

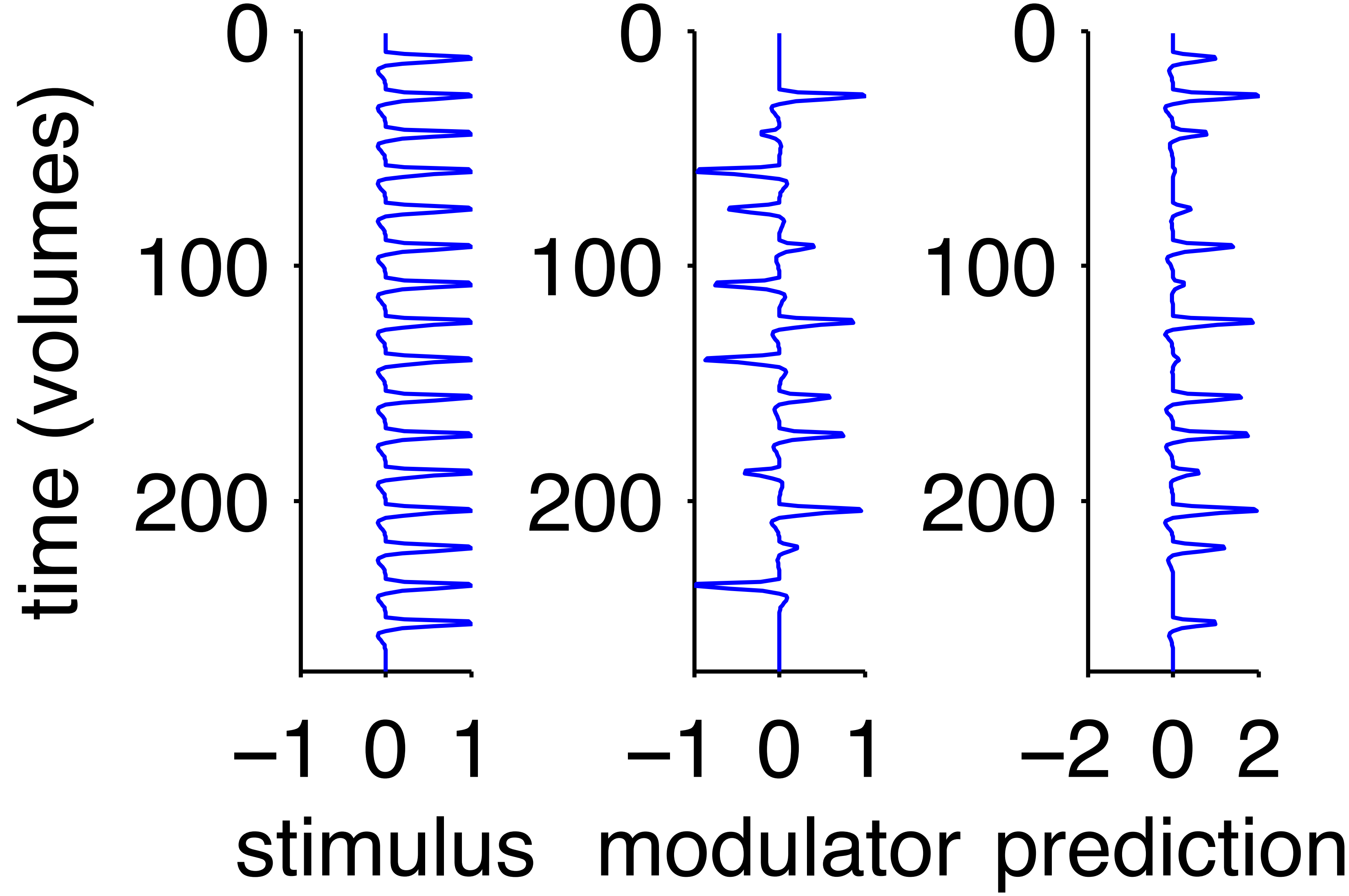
Contrast vectors are basically just a convenient way to average or subtract parameter estimates



Serial autocorrelation

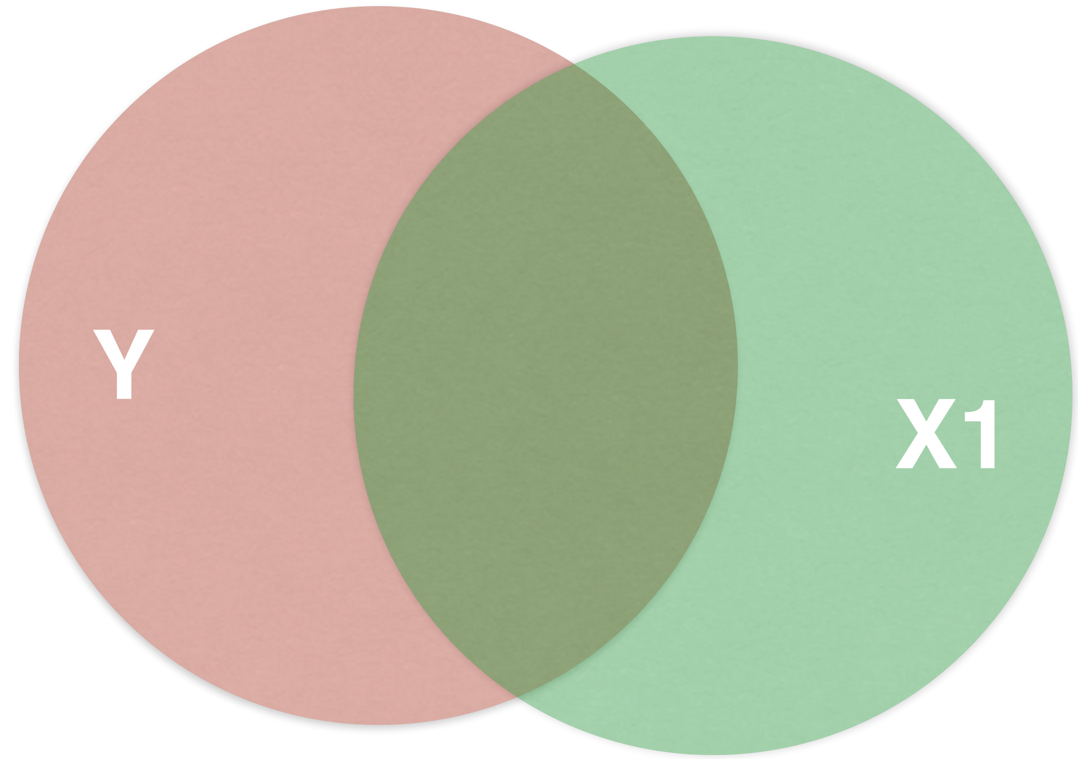
- fMRI data residuals are not independent and identically distributed (iid)
- Why not? Breathing, heartbeat cycle, unmodeled neuronal activity (remember, the BOLD response is temporally smooth)
- This invalidates the error term, which is used for parametric stats inference (T tests, standard errors etc, p values)
- SPM attempts to correct this by estimating (1st-order) autocorrelation and whitening data and design by this
- This works to some extent, but not perfectly (Eklund et al., 2012, NeuroImage) - if your analysis depends on single-participant parametric p values you may want to read this ref and consider alternative (permutation test) approaches
- But for the typical group analysis case, problems with AR modelling are not going to bias your inferences (more on this next time)

Parametric modulators



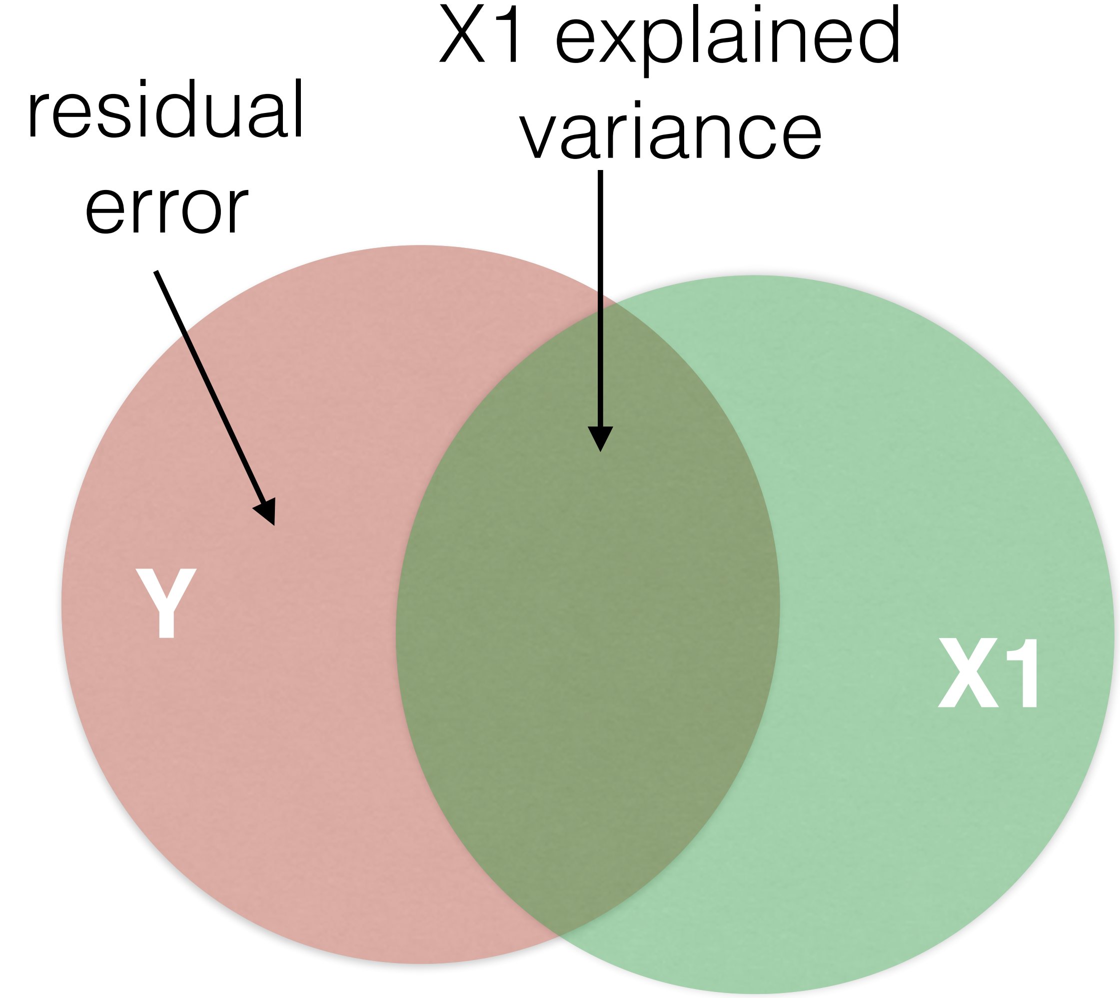
- Encode modulations of stimulus responses by continuous variables
- SPM solution is one regressor for the stimulus effect and another mean-centered regressor for each modulator on that response
- Why not just the modulator? Because we don't want to assume zero response when modulator=0
- Typical applications: Reinforcement learning (Dolan, O'Doherty), visual coding (Huth/Gallant, Kay), 'carry-over' fMRI adaptation (Aguirre), fancy grid cell stuff (Behrens)

Orthogonalisation



Mumford et al., 2015, *PLoS ONE*

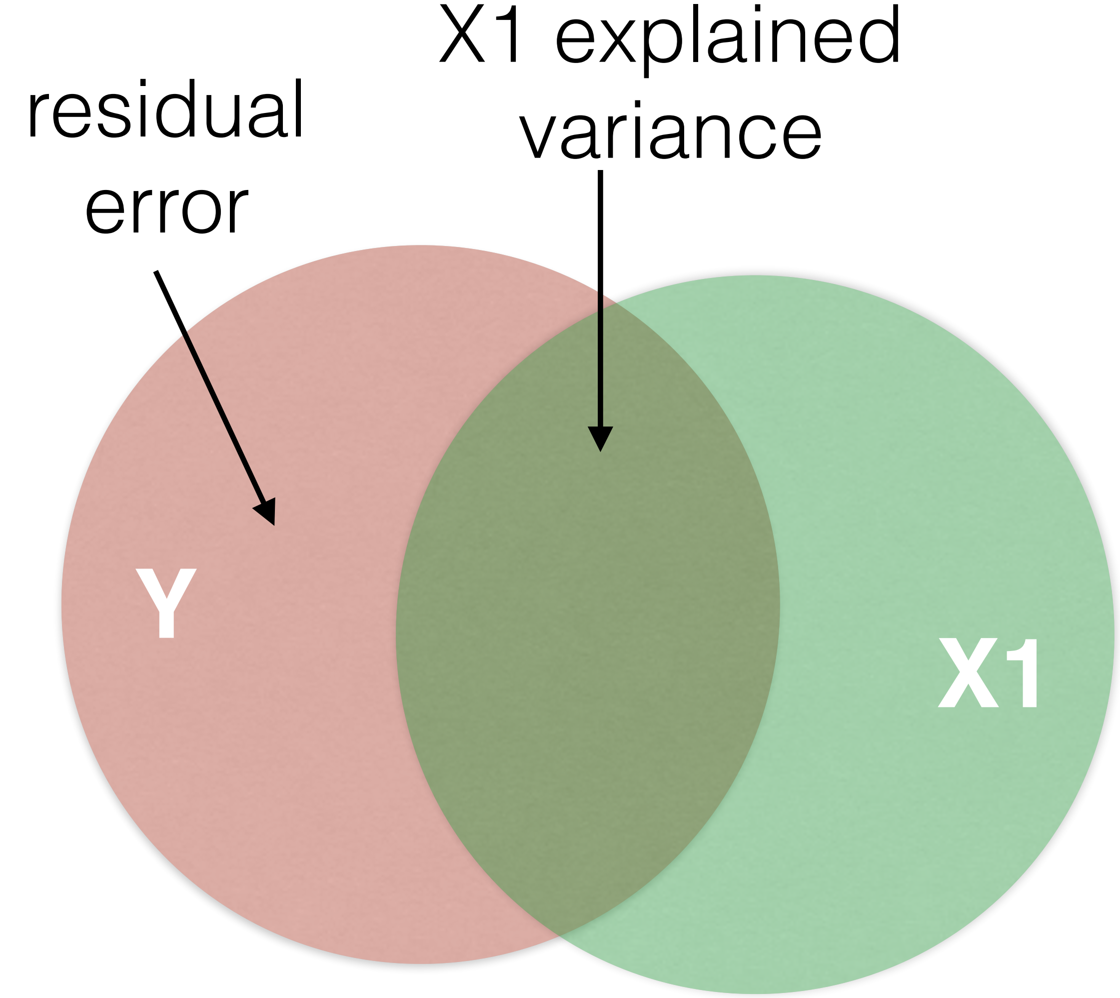
Orthogonalisation



Mumford et al., 2015, *PLoS ONE*

Orthogonalisation

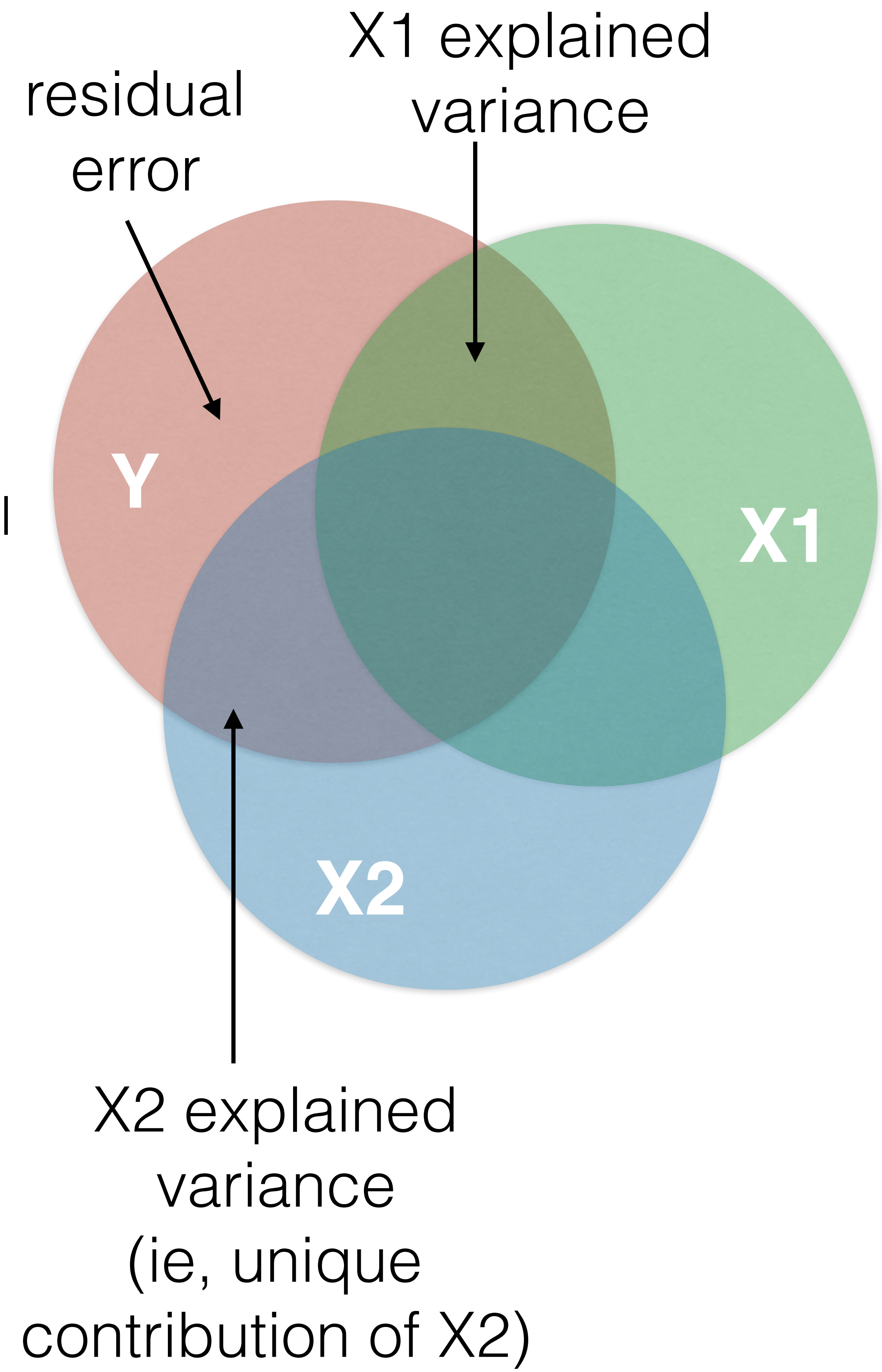
- If we compare the models Y = X1 and Y = X1 + X2, the beta(X1) will explain more variance when it is the only predictor in the model



Mumford et al., 2015, *PLoS ONE*

Orthogonalisation

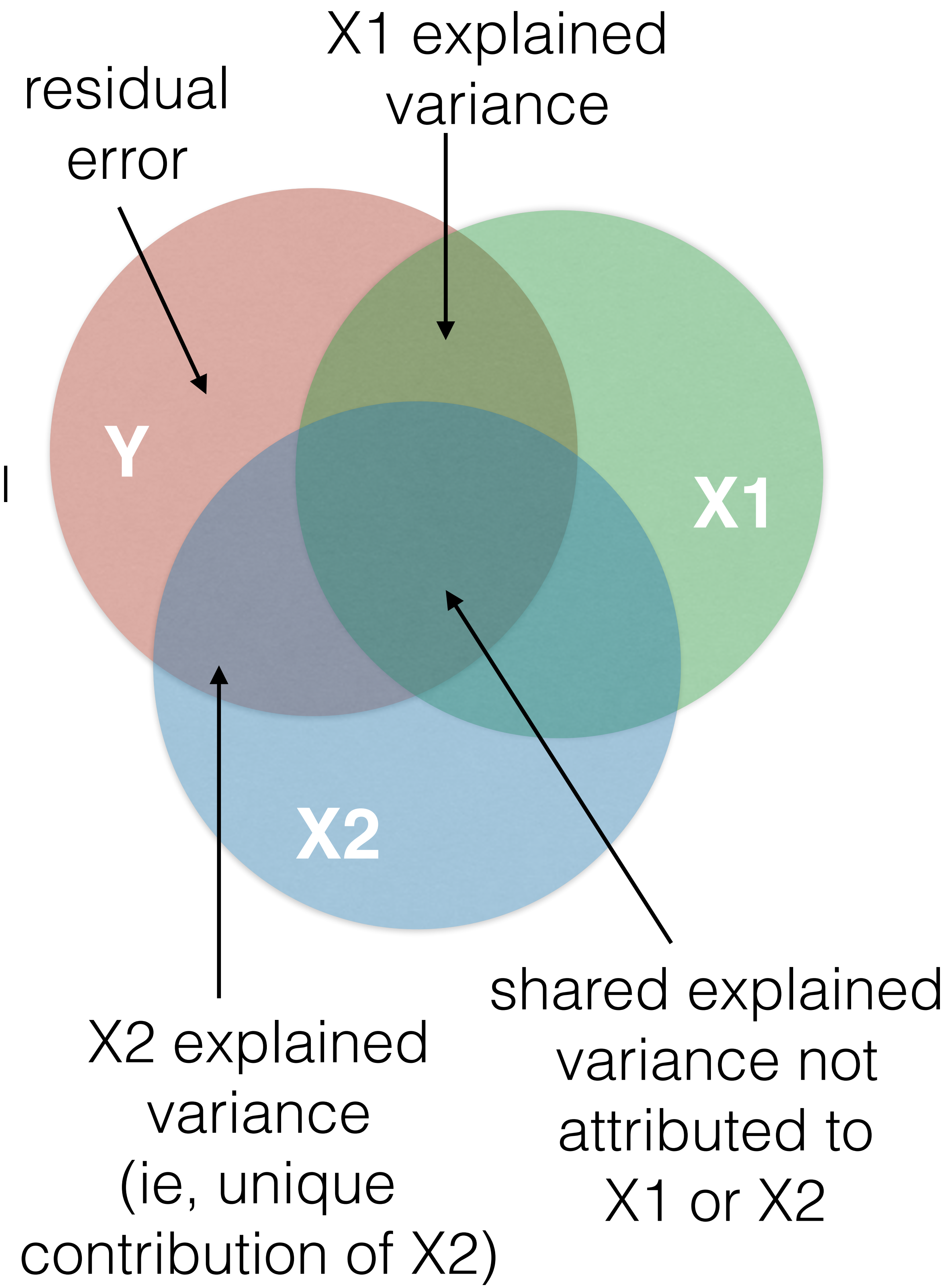
- If we compare the models Y = X1 and Y = X1 + X2, the beta(X1) will explain more variance when it is the only predictor in the model
    - We can think of beta(X1) as being ‘adjusted for X2’ when X2 is in the model



Mumford et al., 2015, *PLoS ONE*

Orthogonalisation

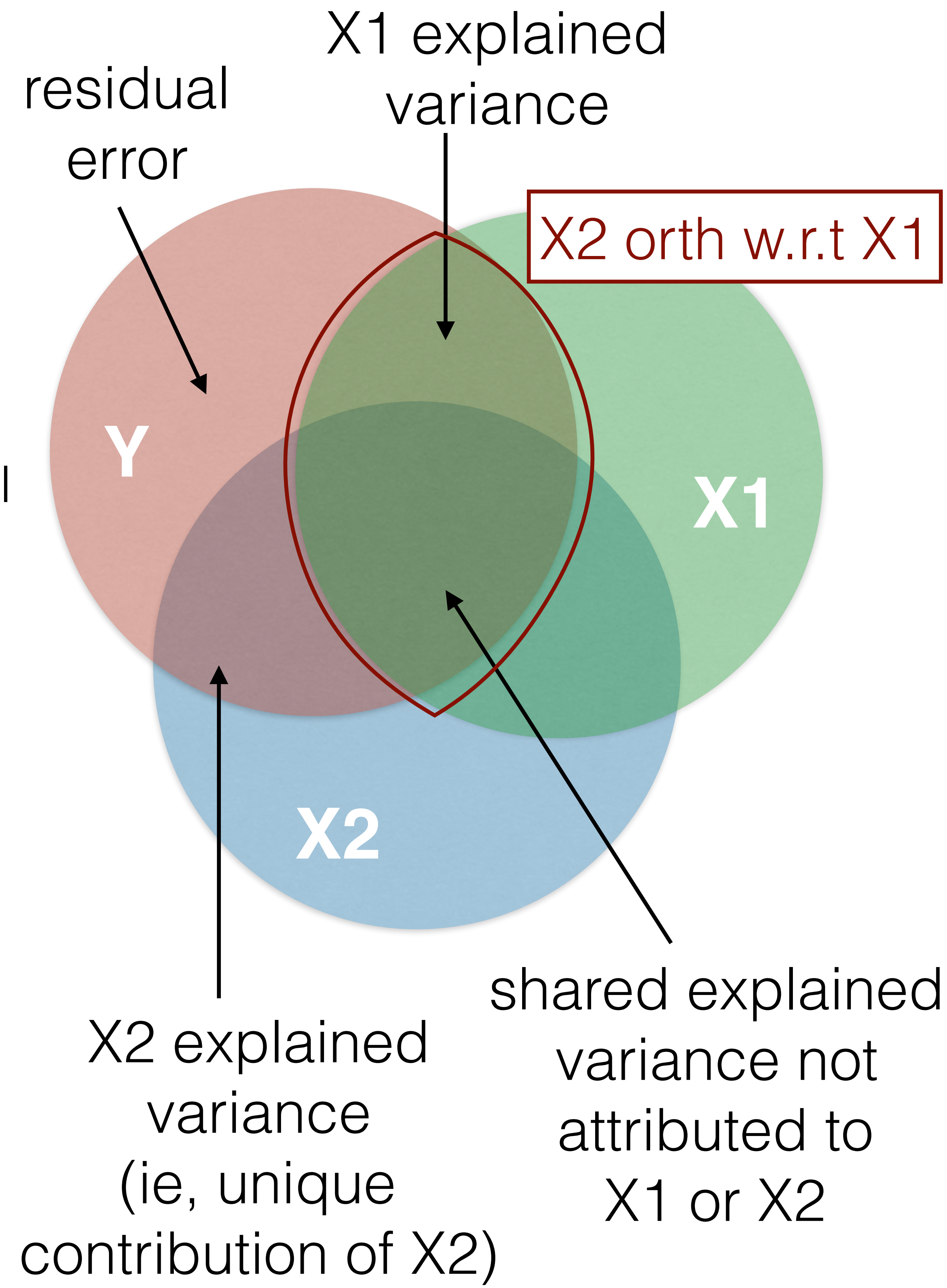
- If we compare the models Y = X1 and Y = X1 + X2, the beta(X1) will explain more variance when it is the only predictor in the model
    - We can think of beta(X1) as being 'adjusted for X2' when X2 is in the model



Mumford et al., 2015, *PLoS ONE*

Orthogonalisation

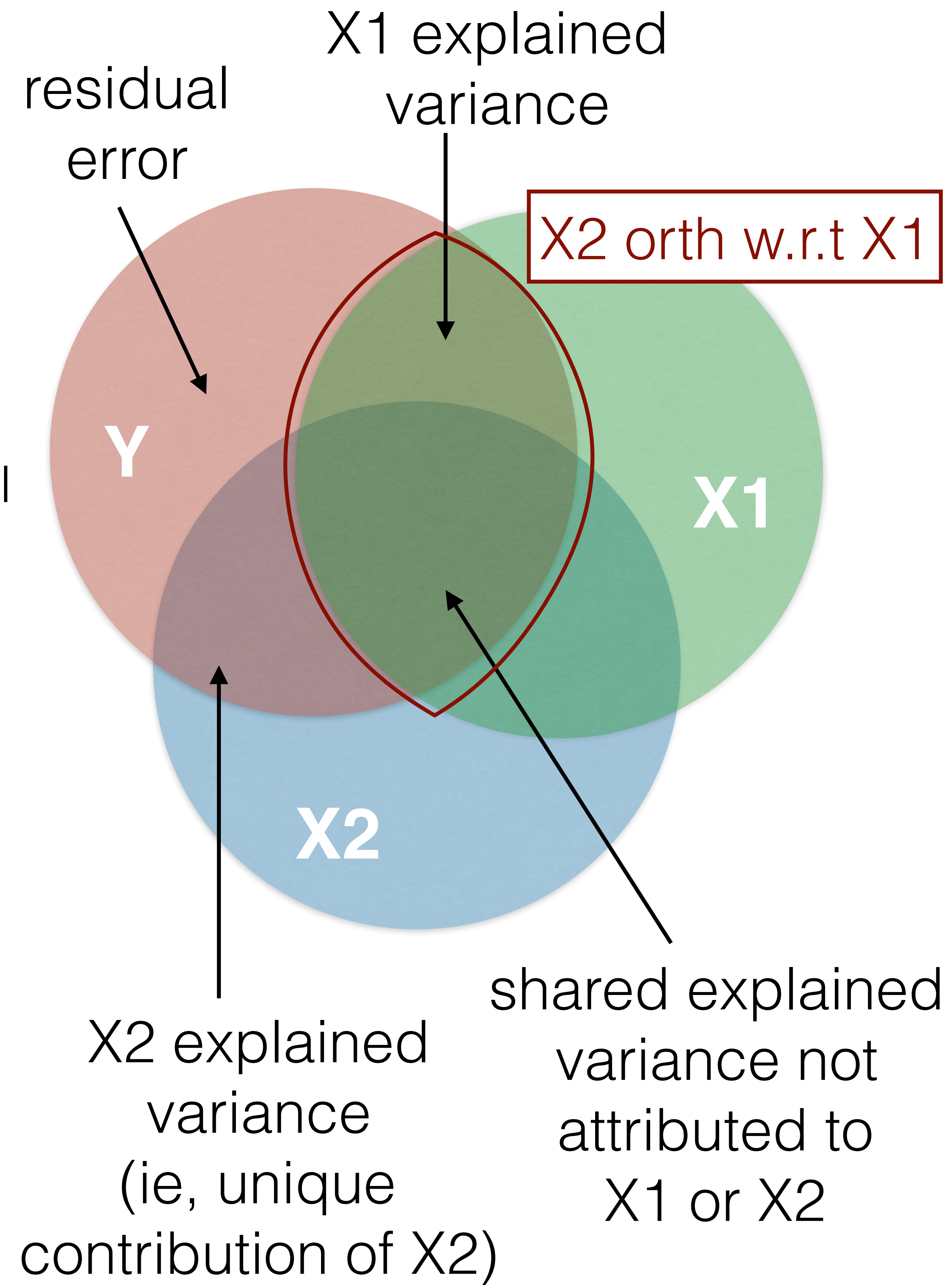
- If we compare the models Y = X1 and Y = X1 + X2, the beta(X1) will explain more variance when it is the only predictor in the model
    - We can think of beta(X1) as being 'adjusted for X2' when X2 is in the model
- What about if we orthogonalise X2 with respect to X1, ie, we regress out the contribution of X1 from X2?



Mumford et al., 2015, *PLoS ONE*

Orthogonalisation

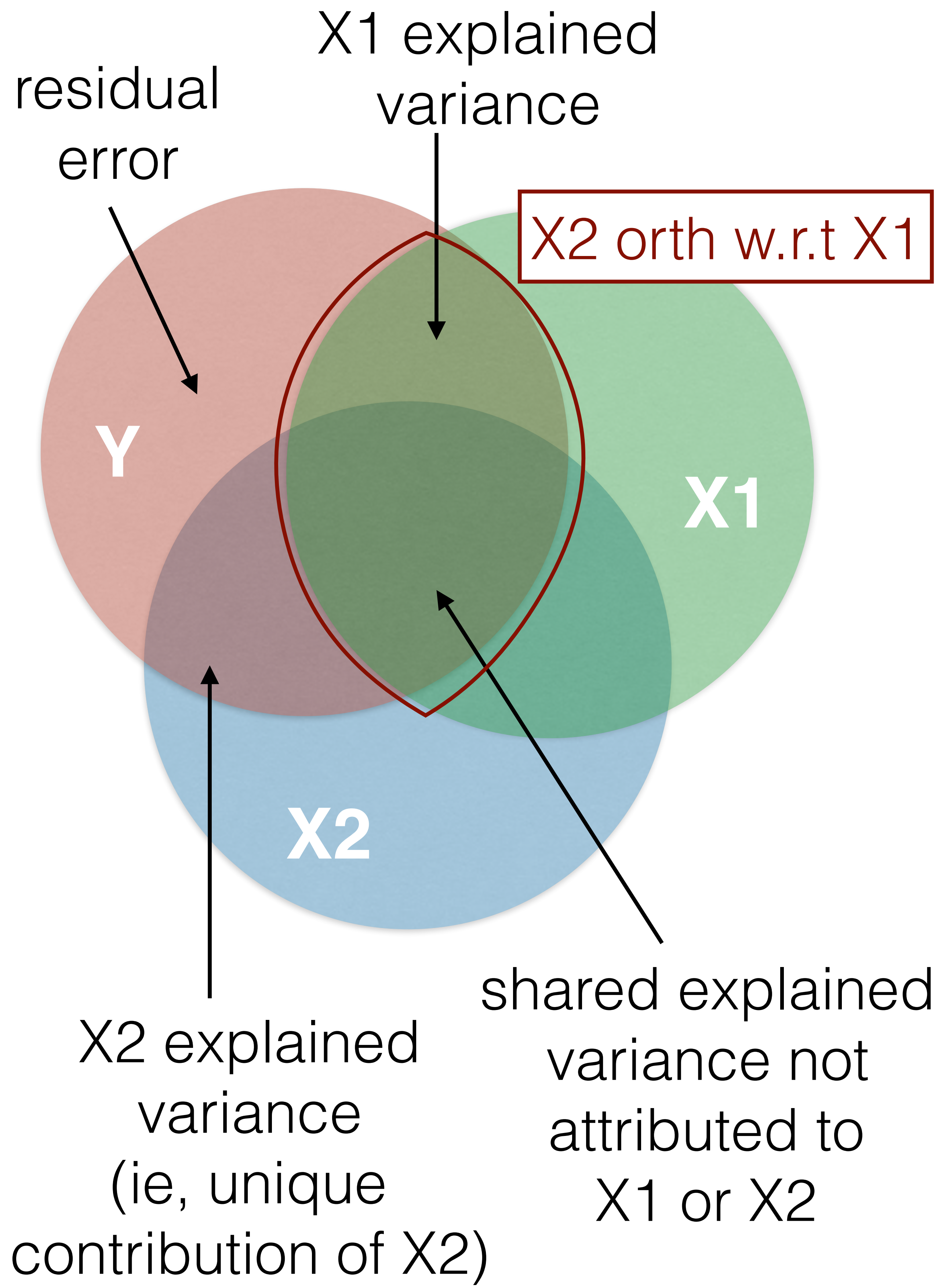
- If we compare the models Y = X1 and Y = X1 + X2, the beta(X1) will explain more variance when it is the only predictor in the model
  - We can think of beta(X1) as being 'adjusted for X2' when X2 is in the model
- What about if we orthogonalise X2 with respect to X1, ie, we regress out the contribution of X1 from X2?
  - The X2 predictor will change - *but beta(X2) will stay exactly the same* (counterintuitive but true!)
  - The X1 predictor stays the same - *but beta(X1) will change*, because all the shared variance now goes to X1



Mumford et al., 2015, *PLoS ONE*

Orthogonalisation

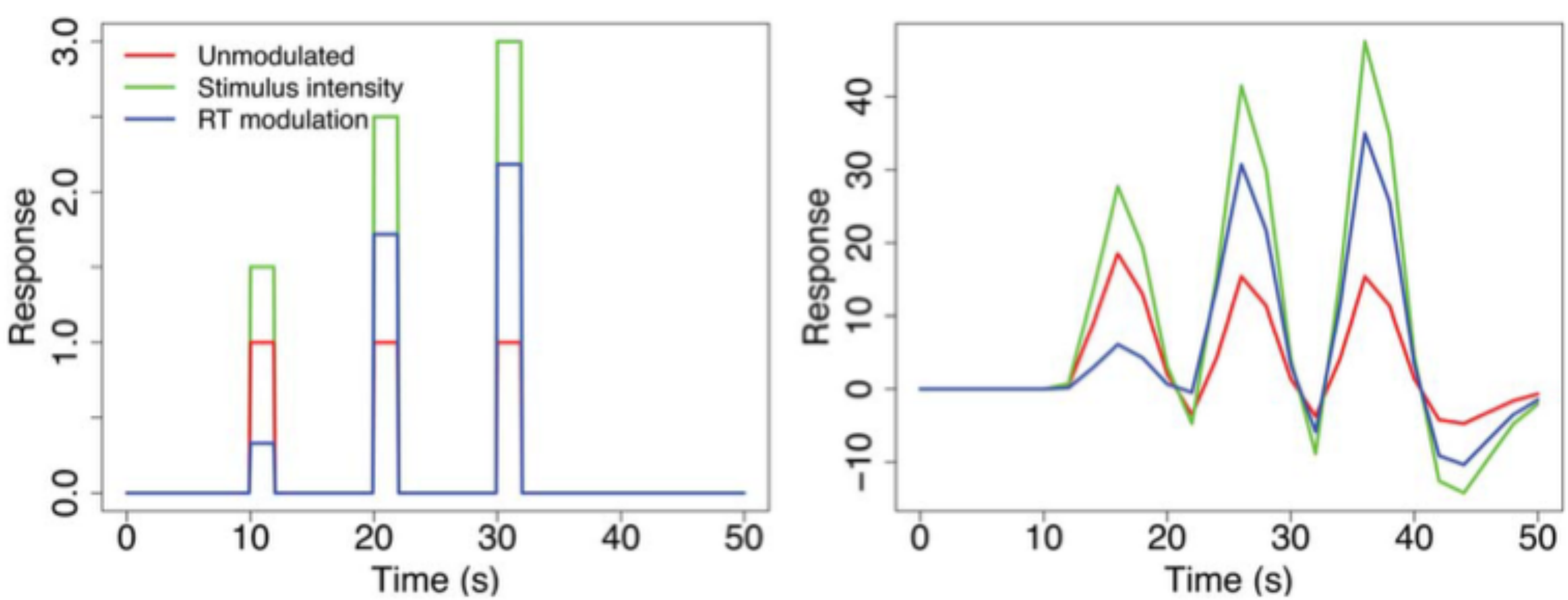
- If we compare the models Y = X1 and Y = X1 + X2, the beta(X1) will explain more variance when it is the only predictor in the model
    - We can think of beta(X1) as being ‘adjusted for X2’ when X2 is in the model
- What about if we orthogonalise X2 with respect to X1, ie, we regress out the contribution of X1 from X2?
    - The X2 predictor will change - *but beta(X2) will stay exactly the same* (counterintuitive but true!)
    - The X1 predictor stays the same - *but beta(X1) will change*, because all the shared variance now goes to X1
- In effect, beta(X1) from the model Y = X1 + X2(orth(X1)) will be similar to the model Y = X1, but with less residual error
    - Same beta but less error = ‘better’ stats



Mumford et al., 2015, *PLoS ONE*

When orthogonalisation makes sense

- We want the main regressor to reflect the mean response to the stimulus, and the modulators to reflect deviations from that mean
- Solution: orthogonalize the two modulators (‘intensity’, ‘RT’) with respect to the main task regressor (‘unmodulated’). Changes ‘unmodulated’ estimates, but not the modulator estimates
- What happens if we don’t do this? unmodulated response is response when intensity=0, RT=0. Hard to interpret.



Mumford et al., 2015, *PLoS ONE*

When orthogonalisation make sense

Careful - colour scheme changes!

- If you want the ‘correct’ case in SPM you will have to hand code the modulators (with appropriate orth) and add them (disabling SPM’s own orth - only available in SPM12).
- Or fit the model twice, taking the second modulator each time
- Or switch to FSL…

Regressors without orthogonalization

Correct orthogonalization for interpretable Unmodulated

RT wrt **Unmod**
Intensity wrt **Unmod**

SPM
RT first
Intensity second

RT wrt **Unmod**
Intensity wrt **Unmod** and **RT**

Mumford et al., 2015, *PLoS ONE*

Workshop time

Setup (if running in your own time)

- Open a terminal:
  - cd /imaging/[yourusername]
  - scp -r /imaging/jc01/practicals/firstlevelmodel ./
  - cd firstlevelmodel
  - matlab_r2012b
- In Matlab:
  - edit firstlevelpractical
  - other code to explore: visualisedesign.m, loadfmridata.m, firstlevelpractical_exampleanswer.m

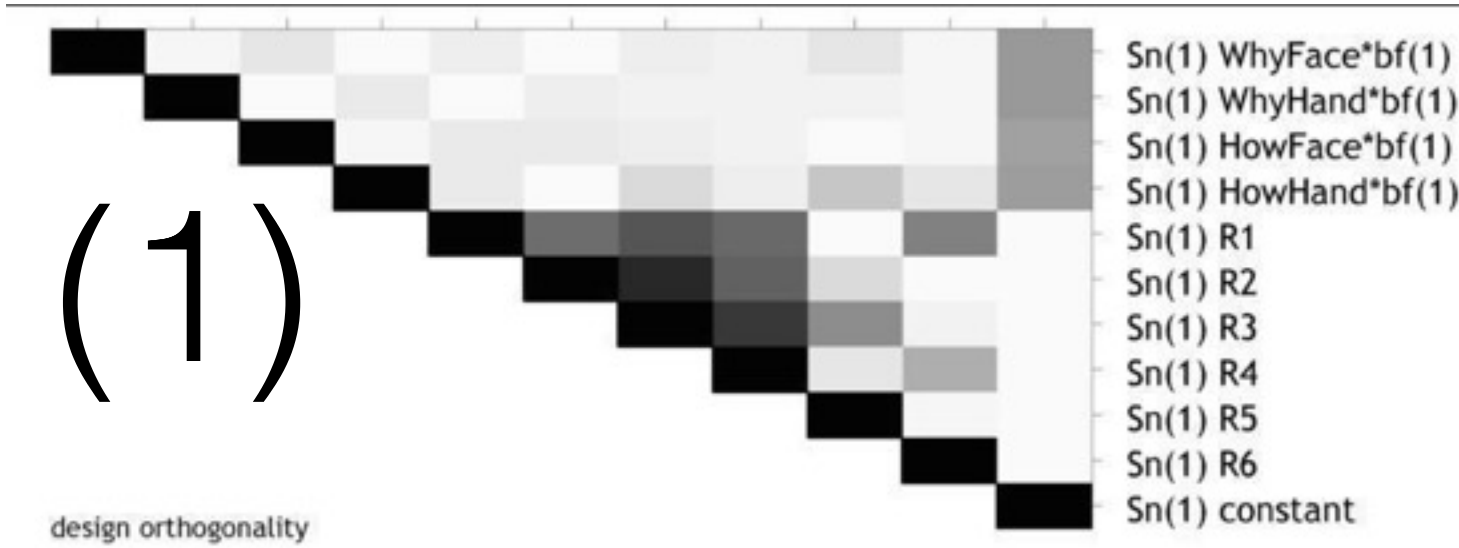
Break

Designing fMRI experiments

Given what we now know about the assumed HRF shape and the noise model, what kind of design is most efficient for detecting the hypothesised effect?

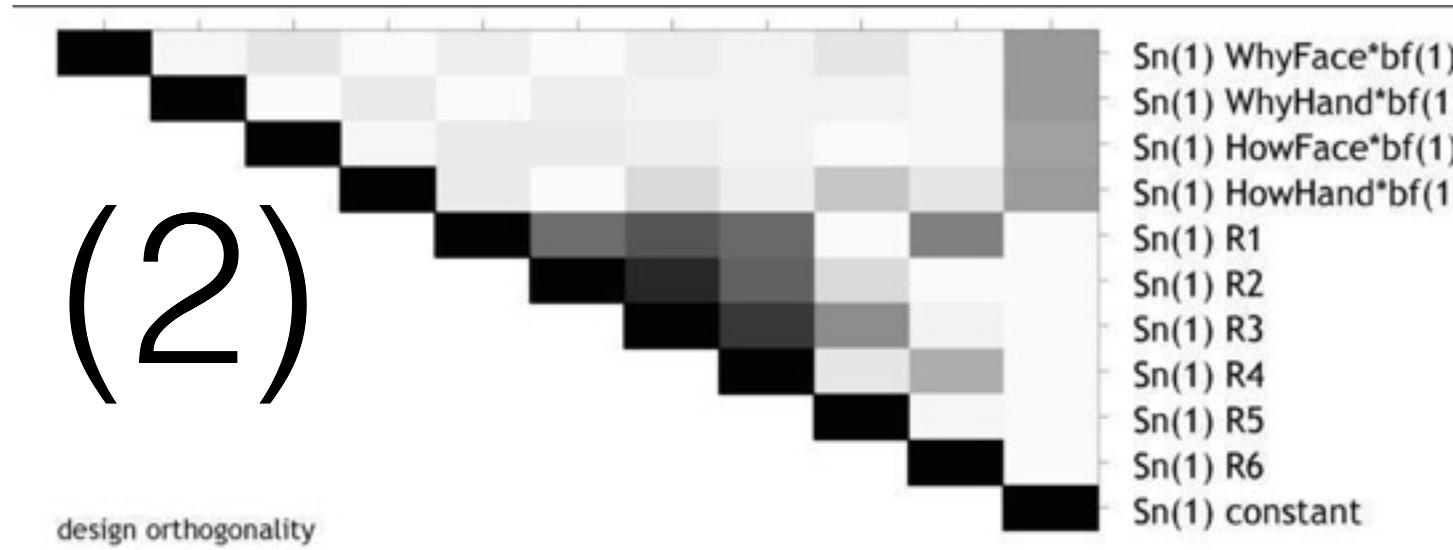
Experimental design has a *huge* effect on detection power in fMRI — this can make or break your study

Collinearity (1)



- Dependencies between convolved regressors increases the variance of parameter estimates
    - NB, does *not* bias the fit - but can make it almost impossible to detect effects (e.g. the single-participant betas that go into group analysis will be highly variable)
- Big problem in fMRI, since convolution with HRF introduces dependencies between neighbouring events (e.g., encoding and recall phase in memory experiment)
- SPM outputs collinearity estimates (see above - basically predictor correlation matrix). Useful for finding pairs of dependent conditions. But a bit late to find this out at model fit stage!

Collinearity (2)

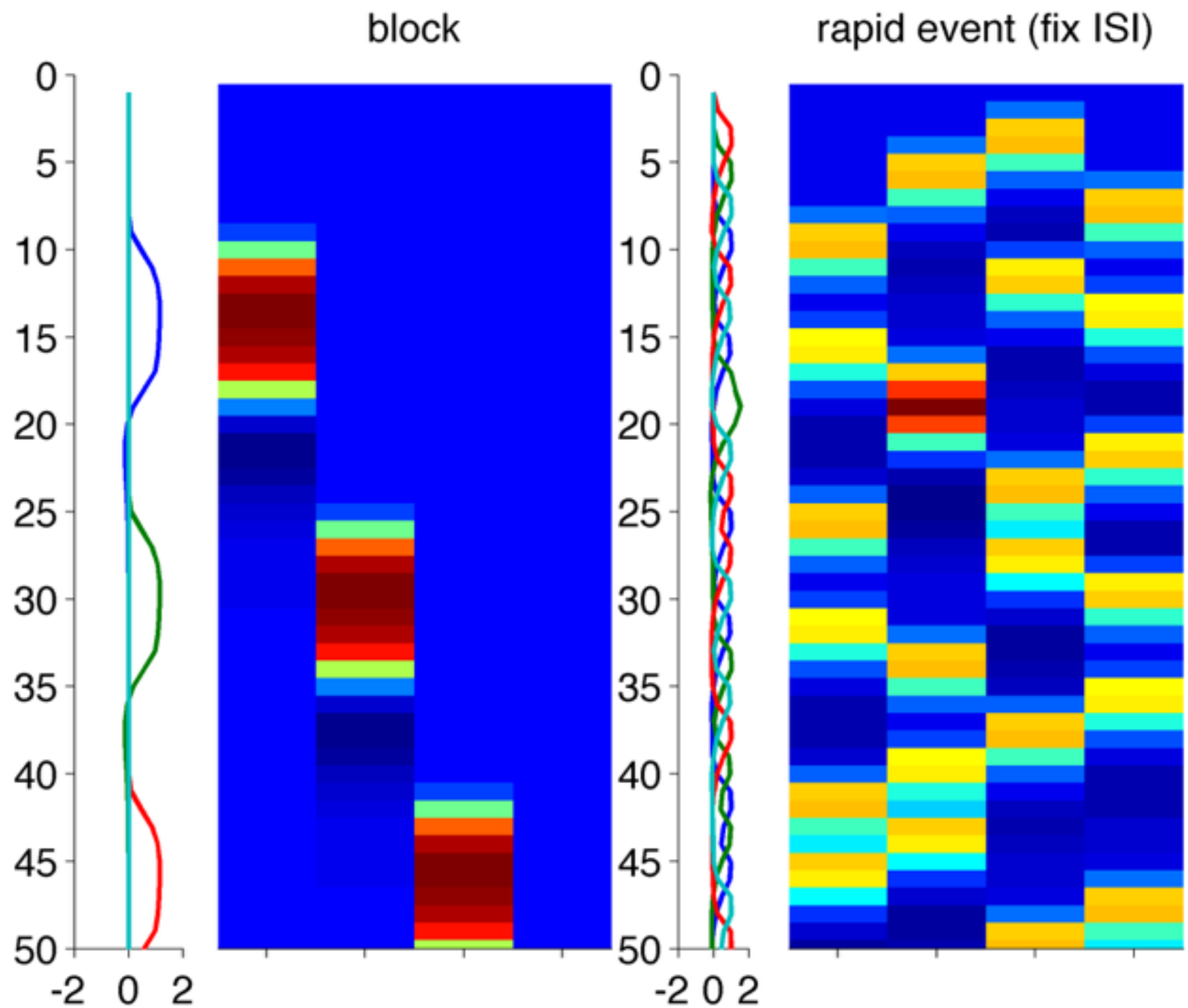


- Collinearity can also arise over sets of regressors - consider using variance inflation factor (VIF) to test for this *at experimental design stage*
- VIF = 1 / (1-R2) where R2 comes from using all regressors but one to predict the final regressor
- Typical values:
    - VIF=1 for completely orthogonal designs (zero correlation between prediction and left-out regressor)
    - VIF=Inf for rank deficient designs (perfect correlation between prediction and left-out regressor)
    - By convention, VIF>5 indicates a problem (but lower is better, always)

… Can we fix collinearity by orthogonalizing regressors?

- No.

Typical fMRI designs



Very efficient

Very inefficient

Rules of thumb for fMRI design

1. Randomise trial order for each run to minimise collinearity
2. Cluster trials (pseudorandom event-related design or just block) to keep signal in low frequency band (the HRF convolution basically *low-pass filters* the regressor)
3. Don’t put conditions you want to compare too far apart (>60s) (the de-trend *high-pass filters* the regressor)
4. Keep the number of conditions as small as possible to make the above easier (and to enable shorter runs)
5. For differential effects (ie, what you usually care about), fixed ISI works best
6. For much more on this, see Rik’s SPM lectures, or CBU imaging wiki entry on design efficiency

Useful references

- Rik's design efficiency wiki: http://imaging.mrc-cbu.cam.ac.uk/imaging/DesignEfficiency
- Jeanette Mumford's brain stats blog: mumfordbrainstats.tumblr.com (see also facebook group)
- The SPM mailing list: https://www.jiscmail.ac.uk/lists/SPM.html (vast searchable archive)
- Kendrick Kay's course on Statistics and Data Analysis in MATLAB: http://kendrickkay.net/psych5007/ (if you want to roll your own GLM)