Group Analysis

Alexa Morcom

Edinburgh SPM course 2019

Thanks to Jesper Anderson, Tom Nichols, Jean Daunizeau, Stephan Kiebel & other SPM authors for slides

Overview

Making the group-level inferences we want

- Optimising the first-level GLM
- Two methods of RFX inference

2-stage GLM

Single subject

Each subject's scans are modelled separately
Single subject parameter estimates

1st level

Single subject **contrasts of parameter estimates** represent different hypothesis tests

Group/s of subjects

A group model is made using the contrasts
Parameter estimates apply to group effect/s

2nd level

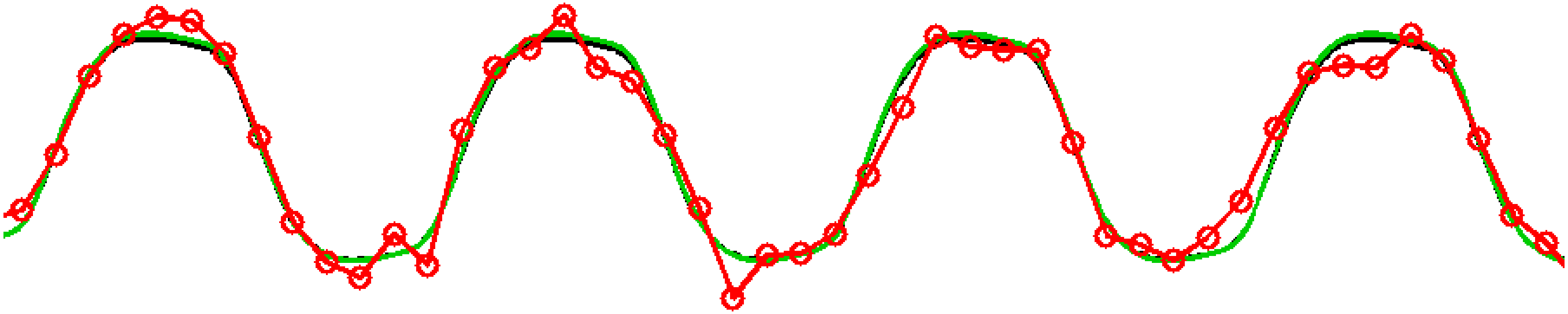
Group level **contrasts of 2nd level parameter estimates** are used to form statistics

- Hierarchical models
- Mixed-effects models
- Random effects (RFX) models
- Variance components

… All the same

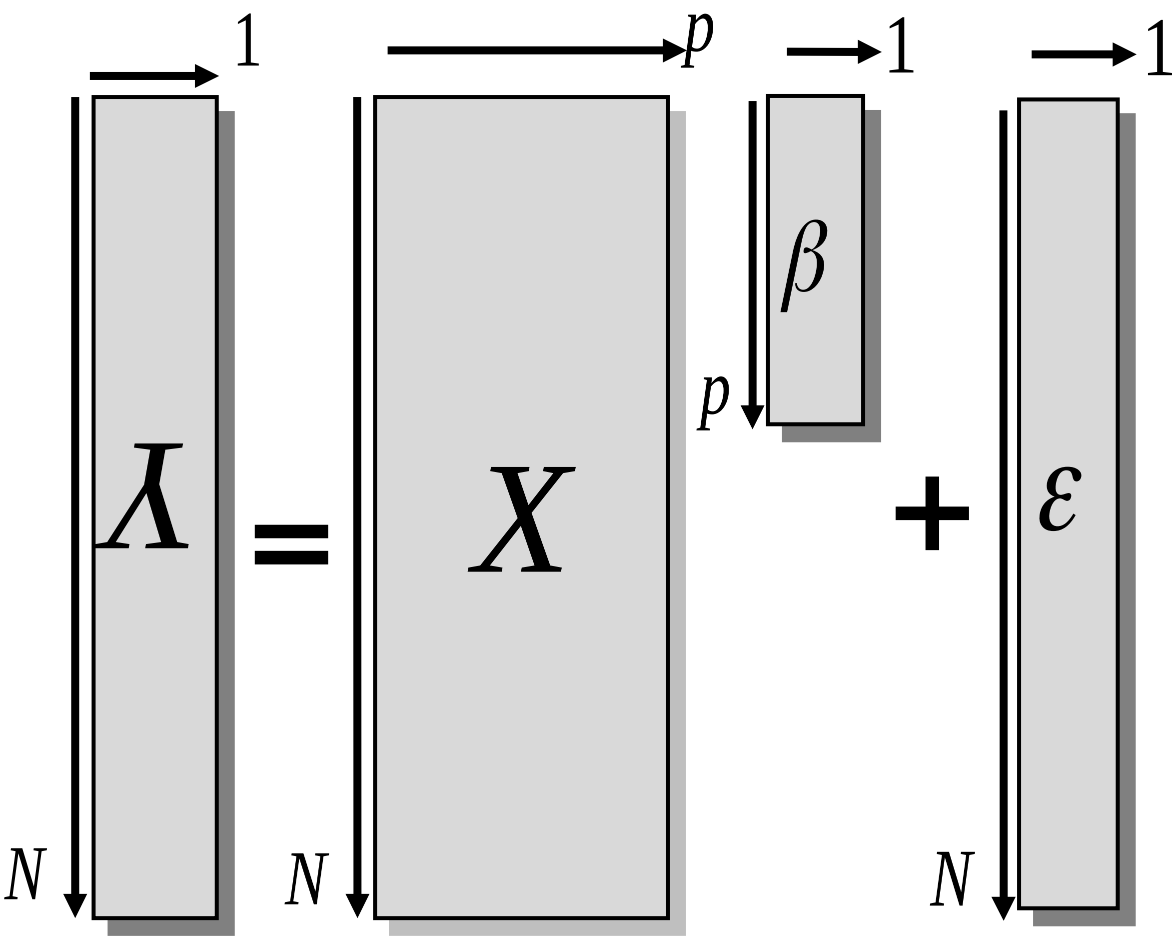
Refer to dealing with multiple sources of variance to make the inferences we want, i.e. generalising to a population

Why isn't this easy?



- Multiple levels of model
- Multiple components of error at each
- To estimate a model's parameters we need to know about the error

The GLM revisited



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

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

Estimate with Ordinary Least Squares (OLS)

N: number of scans
p: number of regressors

Model is specified by
1. Design matrix X
2. Assumptions about ε

Ordinary Least Squares revisited

Find $\hat{\beta}$ that minimises $\|y - X\beta\|^2 = \varepsilon^T \varepsilon$

The Ordinary Least Squares parameter estimates obtained directly:

$$\hat{\beta} = (X^T X)^{-1} X^T y$$

Assumes errors are Independent Identically distributed (iid) – i.e. there is a single error covariance component, the variance s².

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

If not:

- estimates (betas) non optimal
- statistics invalid/ biased, because covariance affects the statistics…

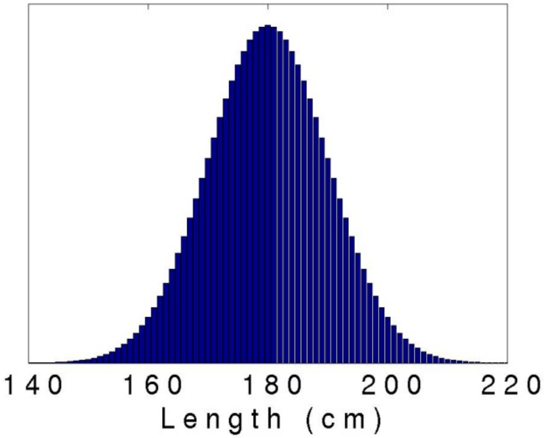
Error covariance and statistics

Classical inference is about what is surprising

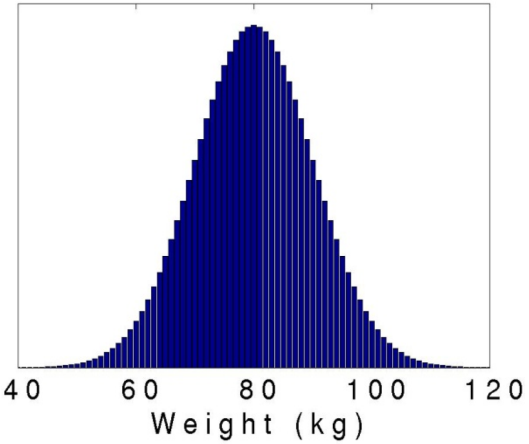
- Use statistic to test an effect's size relative to its expected behaviour under the null hypothesis
- The degrees of freedom must reflect how related (correlated) different observations are
- If observations covary, there are fewer independent observations than we think, so significance of statistics can be overrated

Variance

Length of men



Weight of men



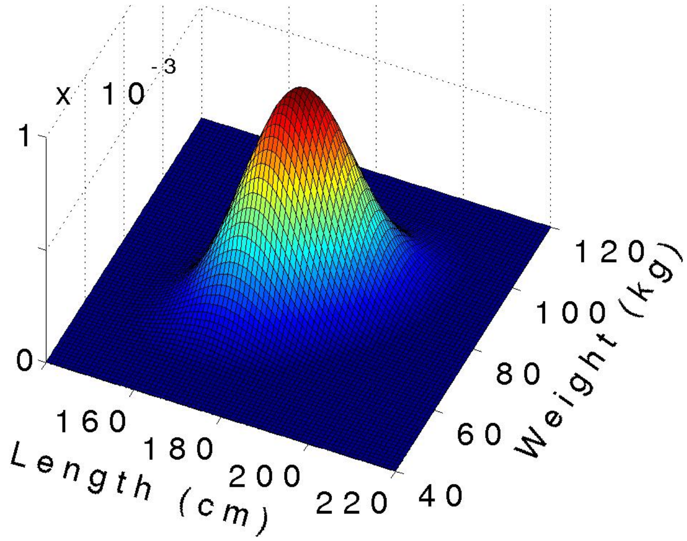
μ=180cm, σ=14cm (σ^2=200)

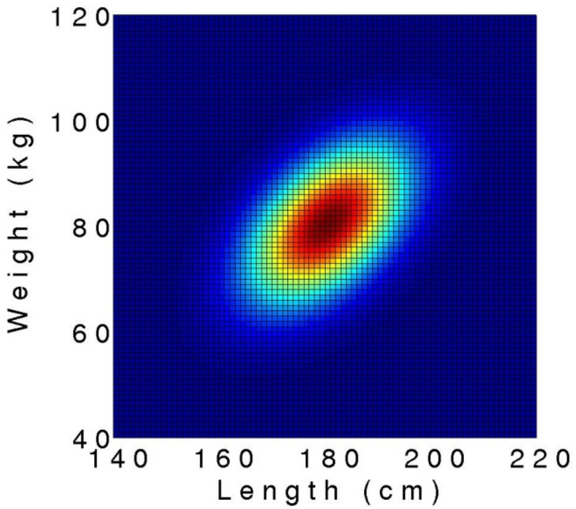
μ=80kg, σ=14kg (σ^2=200)

Each 1-dimensional variable is completely characterised by μ (mean) and σ^2 (variance)

Variance-covariance matrix

- Can also view length and weight as a 2-dimensional random variable





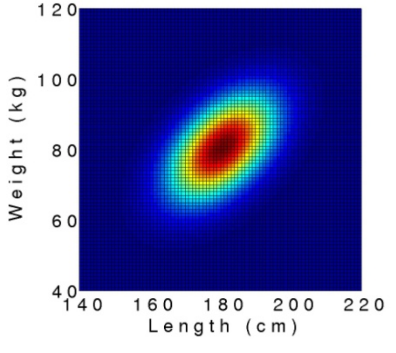
$$\mu = \begin{bmatrix} 180 \\ 80 \end{bmatrix} \qquad \Sigma = \begin{bmatrix} 200 & 100 \\ 100 & 200 \end{bmatrix}$$

Length and weight are related – i.e., covary

Covariance and degrees of freedom

- Measure departure from sphericity (episilon)
- Evaluate significance of sum of squares ratios using F with fewer df (approx); Greenhouse-Geisser

Heights & weights



$$\boldsymbol{\Sigma} = \begin{bmatrix} 200 & 100 \\ 100 & 200 \end{bmatrix} \quad \varepsilon = 0.8$$

= Satterthwaite correction

(in theory sl. liberal – but see Mumford & Nichols, 2009)

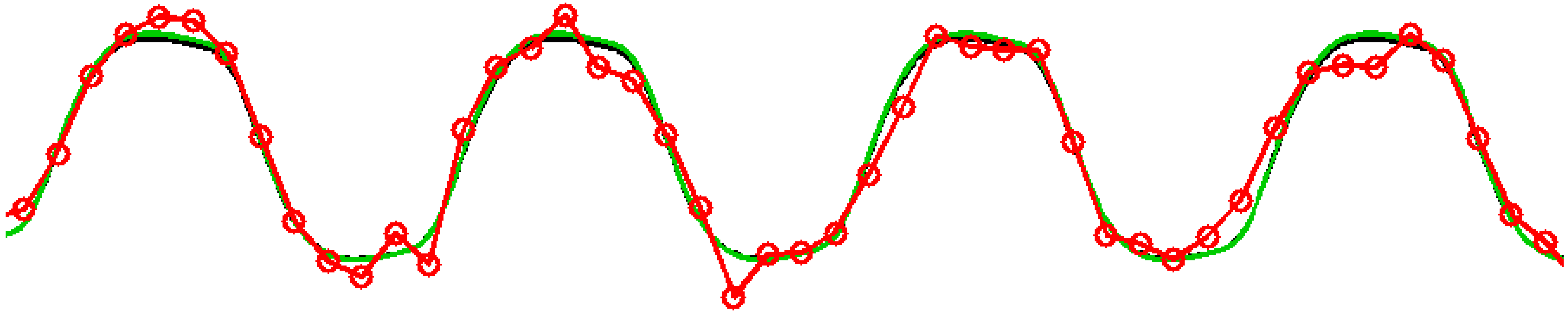
Interim summary

- We want to use Ordinary Least Squares to estimate our model parameters as it's direct (quick, simple)

BUT

- If assumption of sphericity of errors not met, parameter estimates (betas) not optimal
- And statistical tests may be biased

Why isn't this easy?



- Multiple levels of model
- Multiple components of error in each
- First level: scan to scan

The rain in Bergen

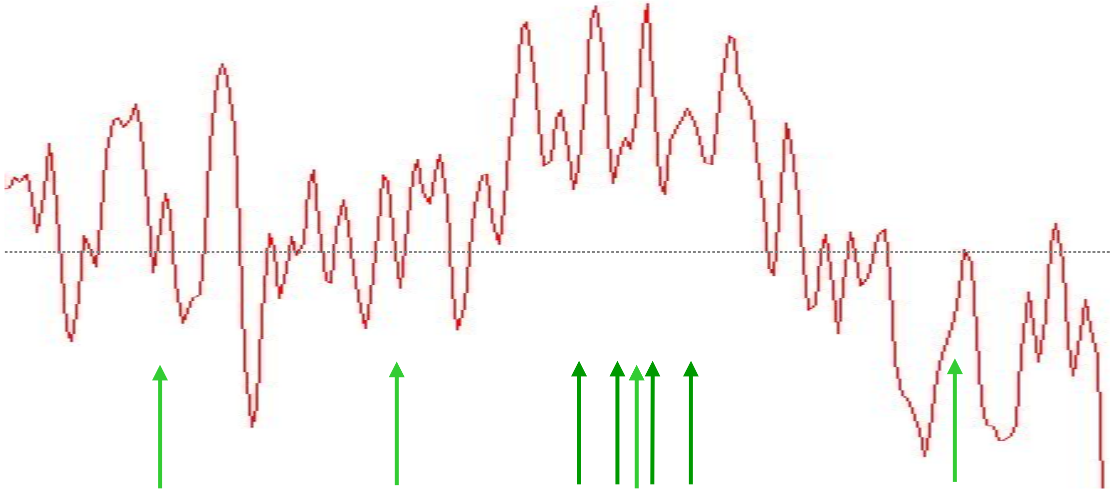
How much do the following observations tell us?

Rain on 4 consecutive days in June

Rain on the same day in May, June, July and August

…which is more likely to indicate a wet summer?

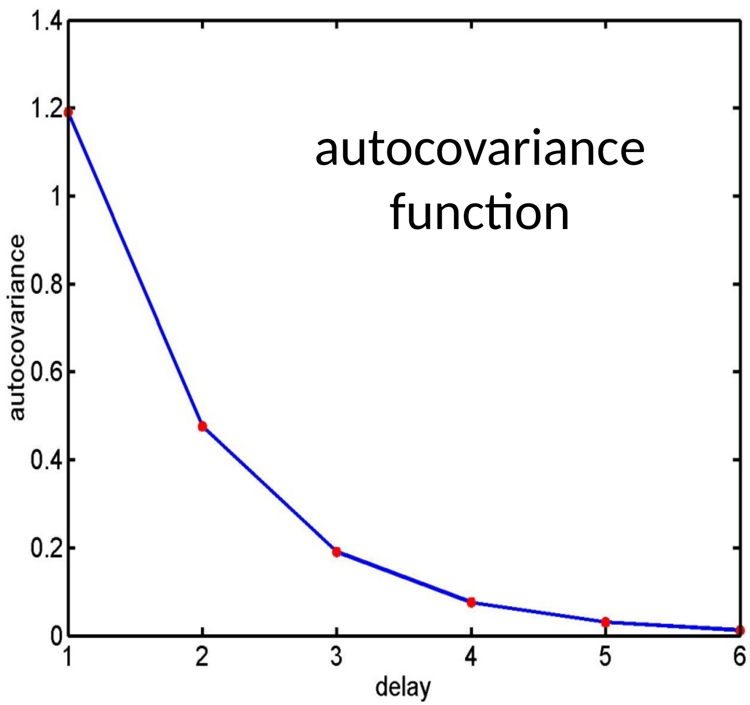
Can we determine the patterns of temporal autocorrelation?

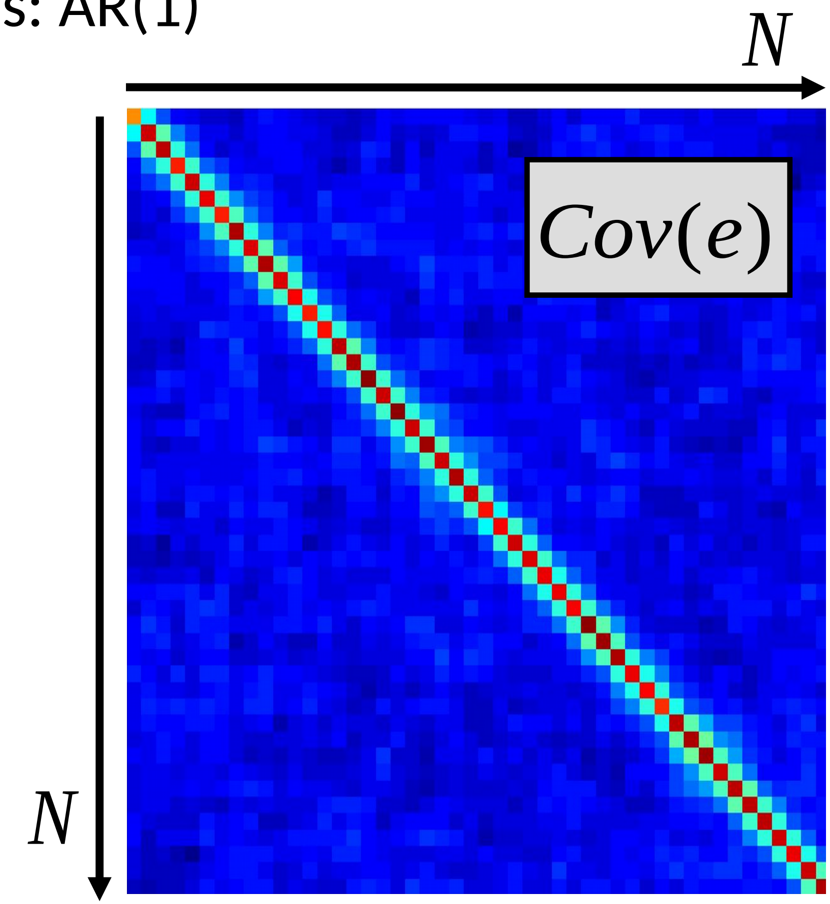


Serial correlations in fMRI

$$e_t = a e_{t-1} + \varepsilon_t \text{ with } \varepsilon_t \sim N(0, \sigma^2)$$

1st order autoregressive process: AR(1)





AND also high-pass filtering

Serial correlations in fMRI

Pre-whitening

- Use an enhanced noise model with multiple error covariance components
- Two components AR (1) + white noise
- Estimate these
- Use resulting matrix *W* to whiten the data – '*undoing*' the serial correlations

$$Wy = WX\beta + We \qquad We \sim N(0, \sigma^2 W^2 V)$$

Serial correlations in fMRI

SPM12 prewhitening model: AR(1) + white noise*

- AR(1) cannot be estimated precisely at each voxel
- But precision is critical, or eventual estimates may be worse than OLS – biased as well as imprecise
- Use spatial regularisation: pool estimation over ‘active voxels‘ (1st pass OLS estimate at p < .001)
- + White noise – voxel-specific variance s^2

*Bayesian estimation option: AR(3) with spatial priors

$$e_t = ae_{t-1} + \varepsilon_t \text{ with } \varepsilon_t \sim N(0, \sigma^2)$$

Serial correlations in fMRI

Once data are 'pre-whitened', estimation can proceed using Ordinary Least Squares

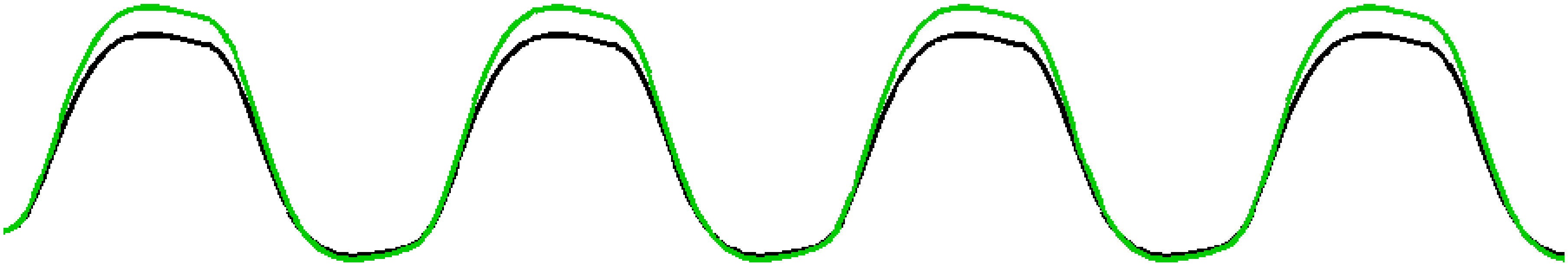
- The parameter estimates are again optimal – unbiased and minimum variance (BLUE estimator)
- The df are also correct
- (if we want to do our statistical inference at the first level)

Take-home message (1)

- If *'error structure'* is complex with multiple components of covariance – not just i.i.d. – our inference depends on modelling the error structure

- What does this have to do with 2-level models?

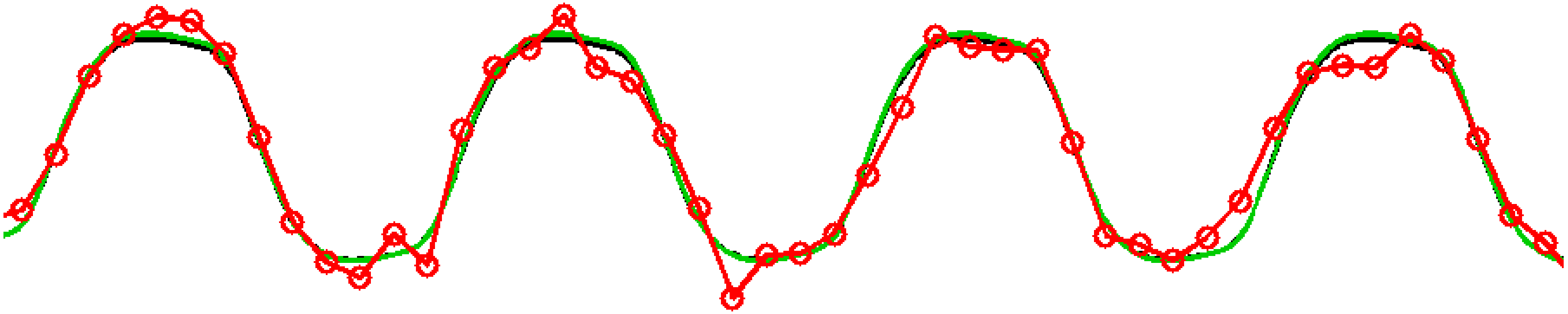
Why isn't this easy?



- Multiple levels of model
  - measurement errors
  - response magnitude random (over subjects)

  - population mean magnitude is *fixed*

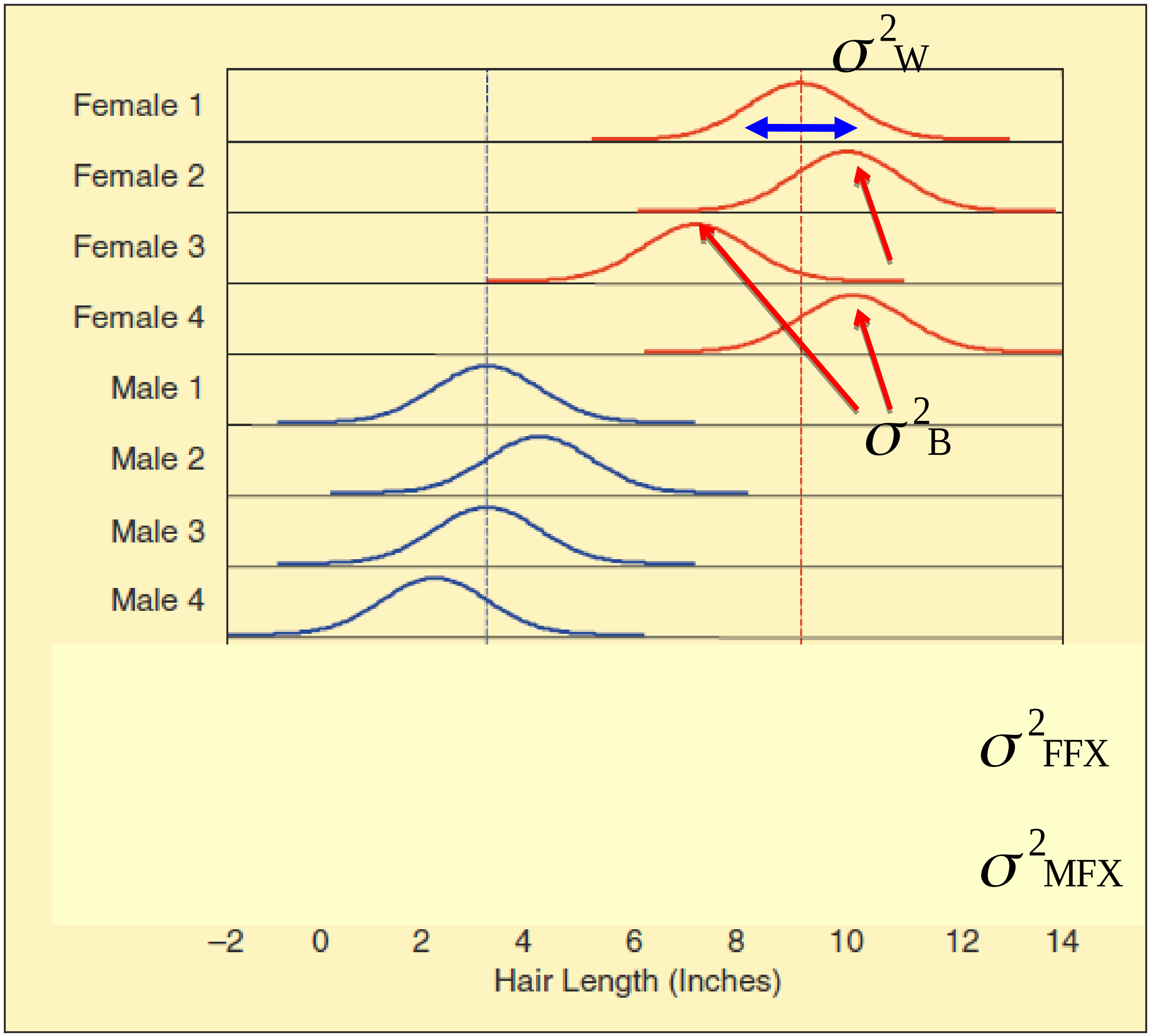
Why isn’t this easy?



- Multiple levels of model
- Multiple components of error in each

- Second level: errors from hierarchical structure

Hierarchical model: worked example



Does hair length differ by gender?

2 sources of variability

Within-subject: σ^2_W

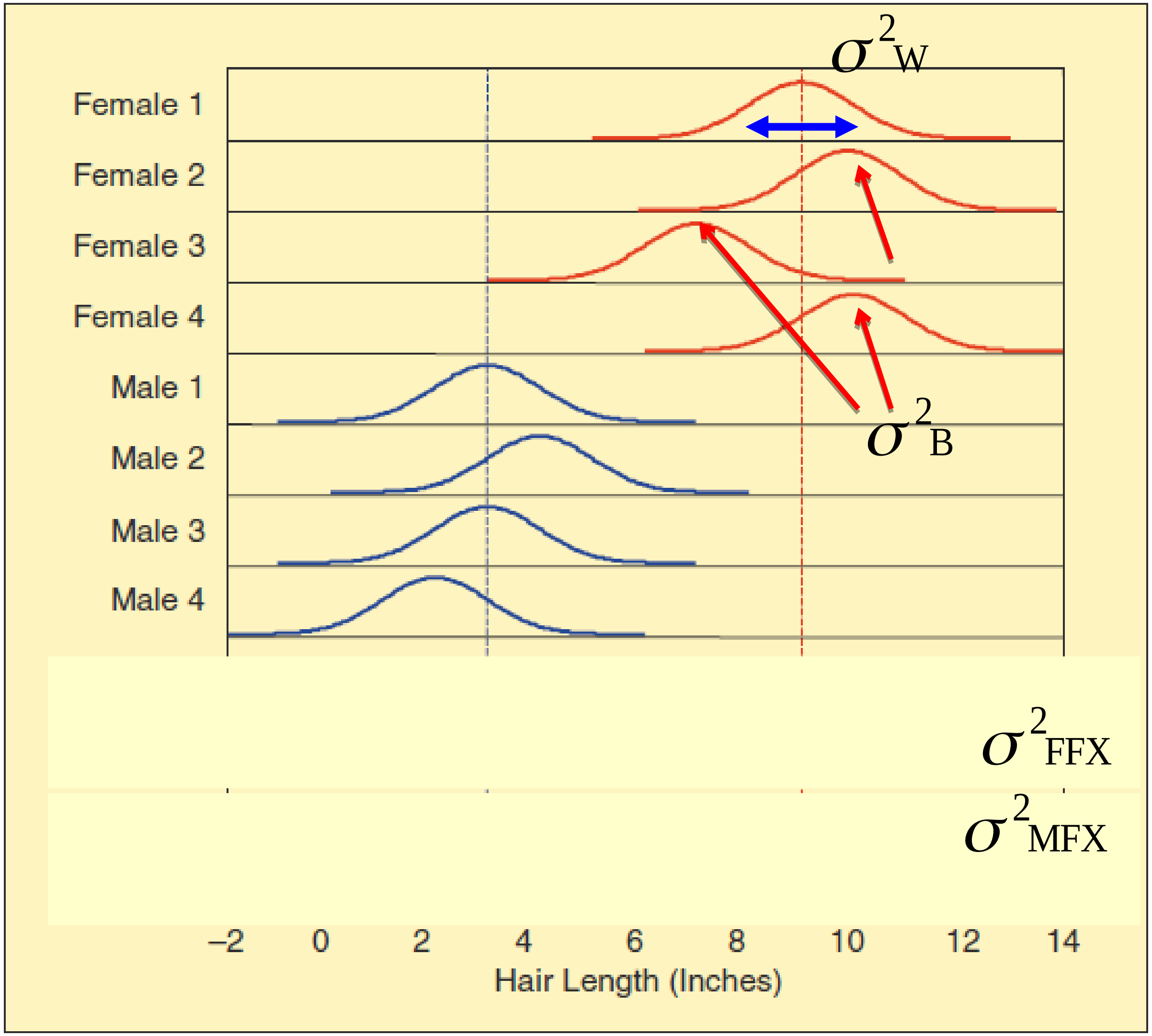
Between-subjects: σ^2_B

To generalise to population, use estimates of hair length for each subject, get σ^2_{MFX}

MIX of between/ within subject variability

Mumford & Nichols (2006)

Hierarchical model: worked example



Does hair length differ by gender?

2 sources of variability

Within-subject ($\sigma^2_W=1$)

Between-subjects ($\sigma^2_B=49$)

To generalise to population:

Combine within- and between- variances (weighting by N=4 subj. and k=25 trials):

$$\sigma^2_{MFX} = \frac{1}{4} * \frac{\sigma^2_W}{25} + \boxed{\frac{1}{4}\sigma^2_B} = 12.26$$

Mumford & Nichols (2006)

Implementing the group models

- Hierarchical models must deal (in some way) with the mixed sources of variance, not just between-subject variance
- In fMRI, we have both scan-to-scan and subject-to-subject variability
- In theory, there is therefore more than 1 variance component (i.e., nonsphericity) at the group level

Interim summary

- We (almost always*) need to make a Random effect inference (over participants):

  "If I were to sample a new cohort from the same population I would get the same result"

- How does SPM address the multiple variance components from the multi-level hierarchical model?

*sometimes ignore this and do Fixed effect analysis e.g. PET/ early fMRI studies; not 'wrong', but not usually of interest

Two-stage models in SPM

Two* approaches

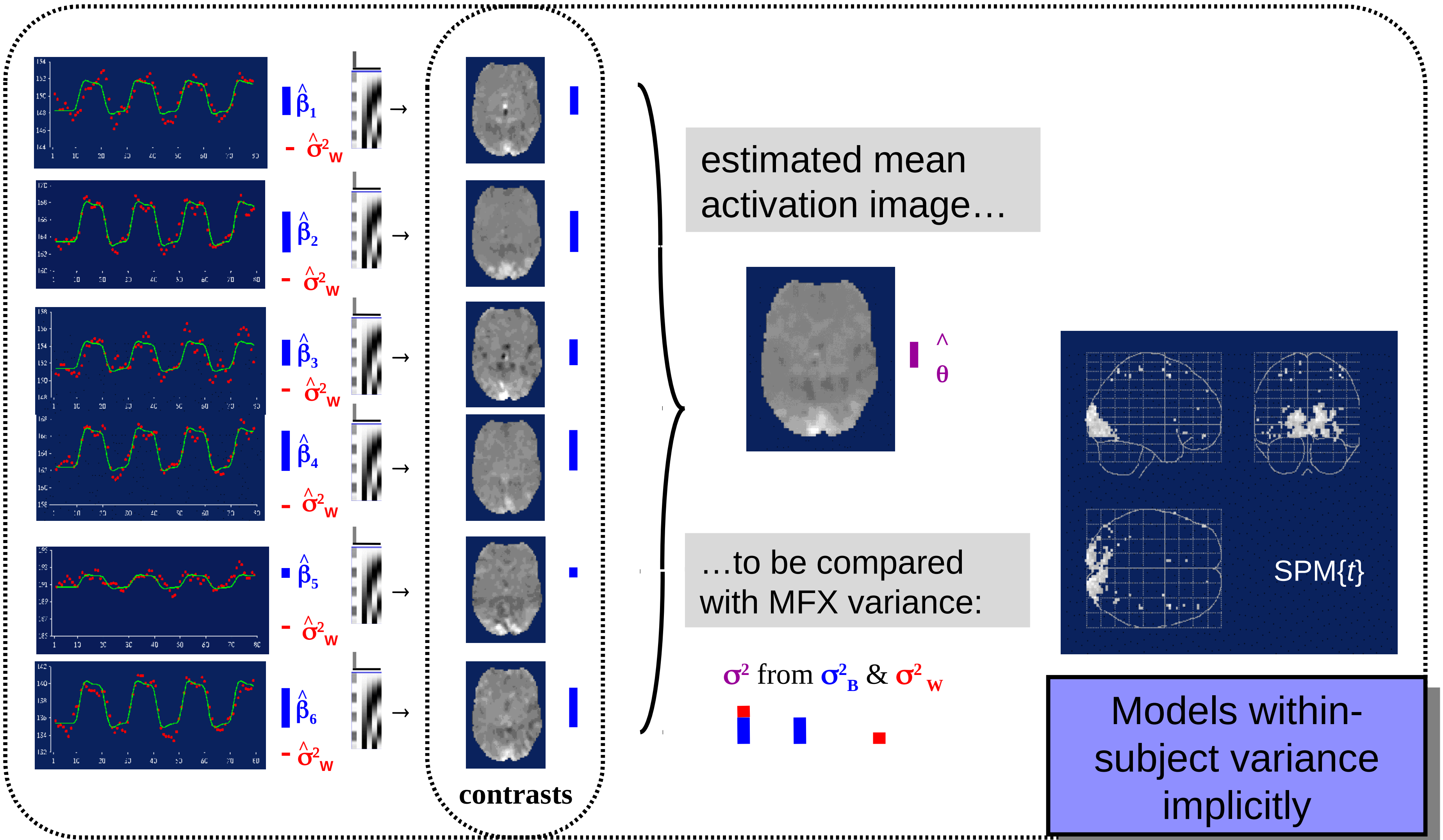
1) Simple summary statistic – Holmes & Friston (HF)
2) Non-sphericity modelling at group level

- Pros and cons – assumptions vs. flexibility
  - Subject variances equivalent
  - Subject *design matrices* equivalent
  - (2) enables a wide range of 2nd level models

*Actually 3, but we talk about 2 here; see Friston et al., 2005 re. full MFX/ LMM

Simple summary statistic approach ('HF')

1st level (within subjects) 2nd level (between-subjects)



Simple summary statistic approach (‘HF’)

Assumptions

- Distribution normal over independent subjects
- Homogeneous variance
  - Subjects’ residual errors same
  - Subjects’ design matrices same

  - Our 2 MFX covariance components can be collapsed into 1 if these elements of the group level covariance are homogenous over subjects

Simple summary statistic approach (‘HF’)

Use only a single image per subject

- Limited to 1- or 2-sample t-tests at the 2nd level
- Balanced designs

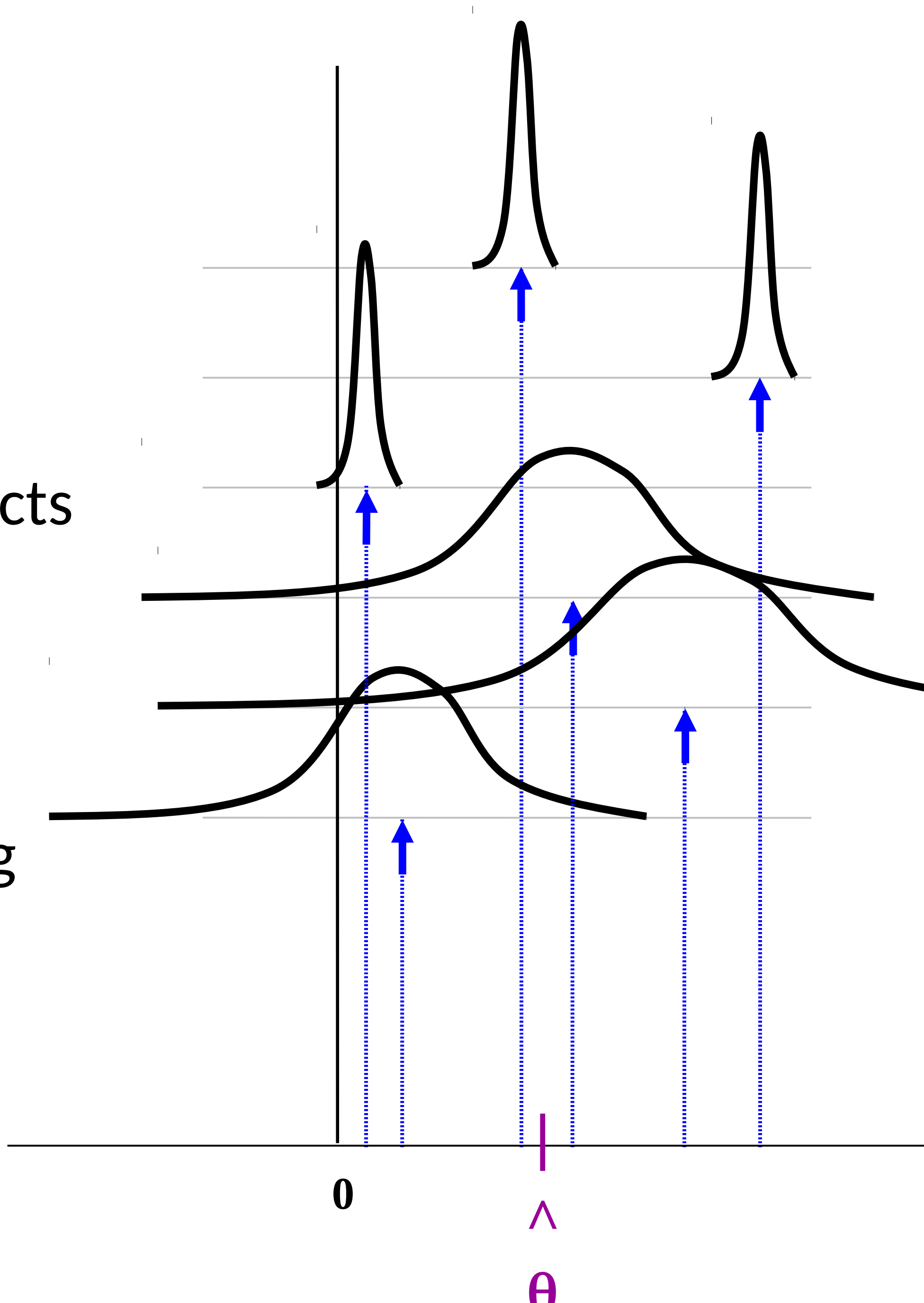
- Limitation = strength
  - No 2nd level sphericity assumption
  - ‘Partitioned’ error term @ 2nd level

  c.f. main effects, interactions in ANOVA have separate error terms

HF - efficiency

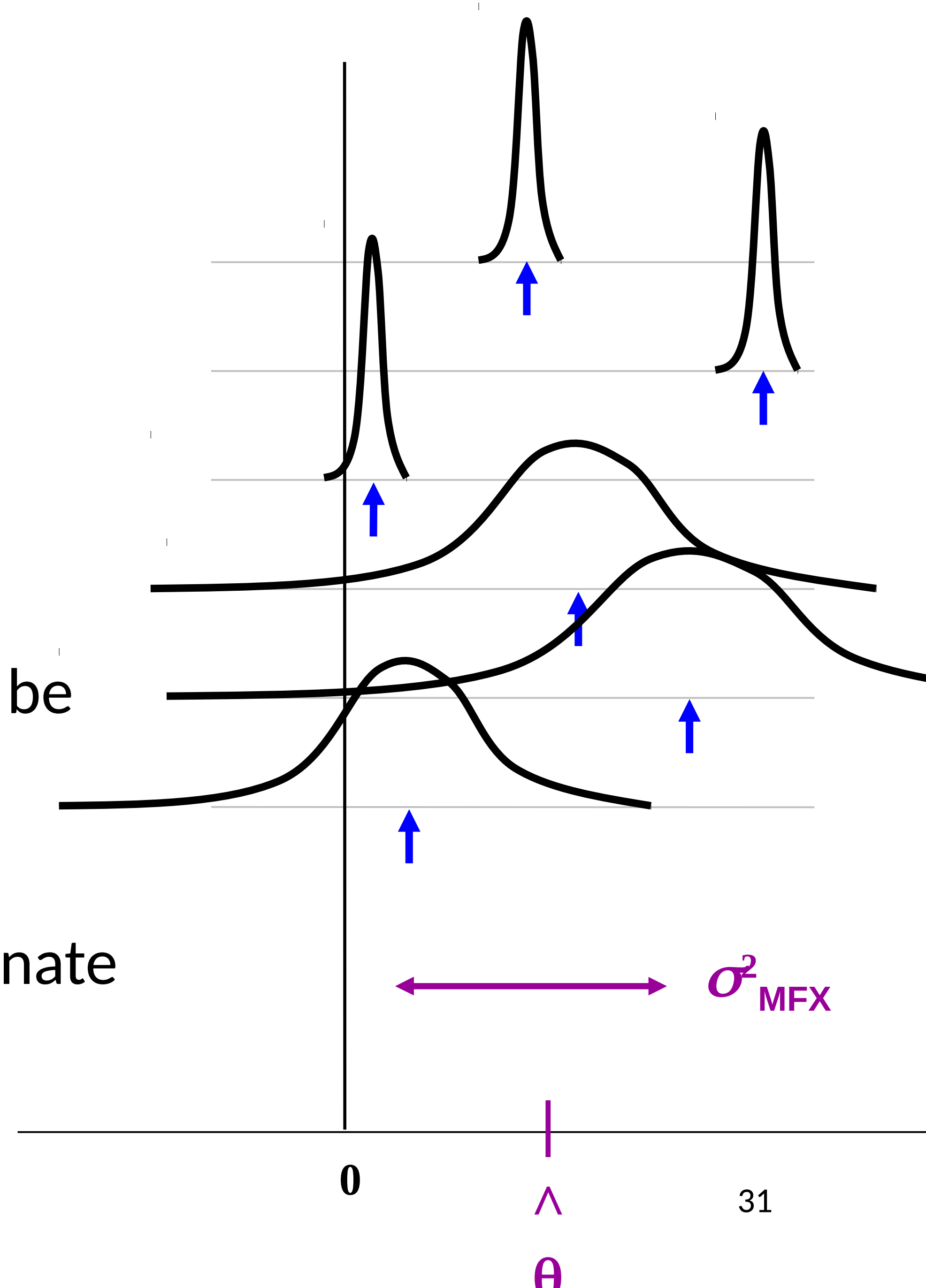
- If assumptions true
  - Optimal, fully efficient
- If σ^2_w differs between subjects
  - Reduced efficiency
  - Here, optimal group parameter estimate requires down-weighting the 3 highly variable subjects

$\hat{\theta}$



HF - validity

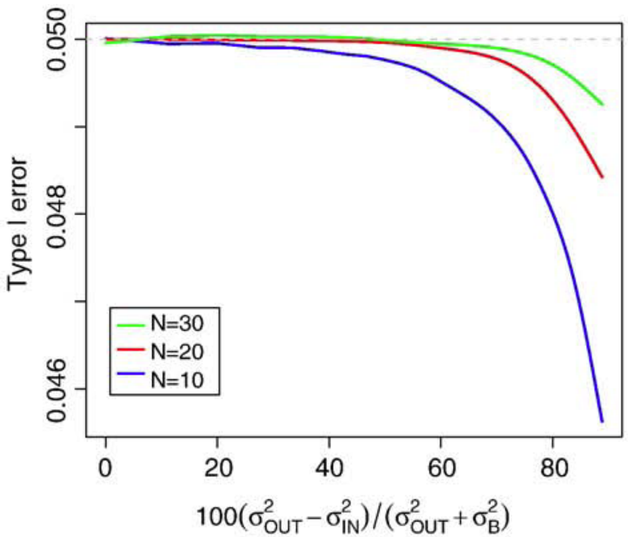
- If assumptions true
  - Exact *P*-values
- If σ^2_w differs btw subj.
  - Standard errors not OK
    - Estimate of σ^2_{MFX} may be biased
  - df not OK
    - Here, 3 subjects dominate
    - df actually < 5 (= 6-1)



HF – robustness

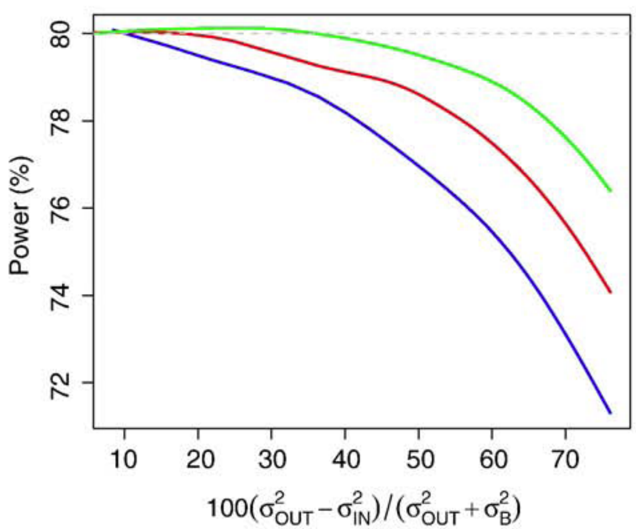
- In practice, validity & efficiency are excellent
  - For the one sample case, HF is very robust

False Positive Rate



(outlier severity)

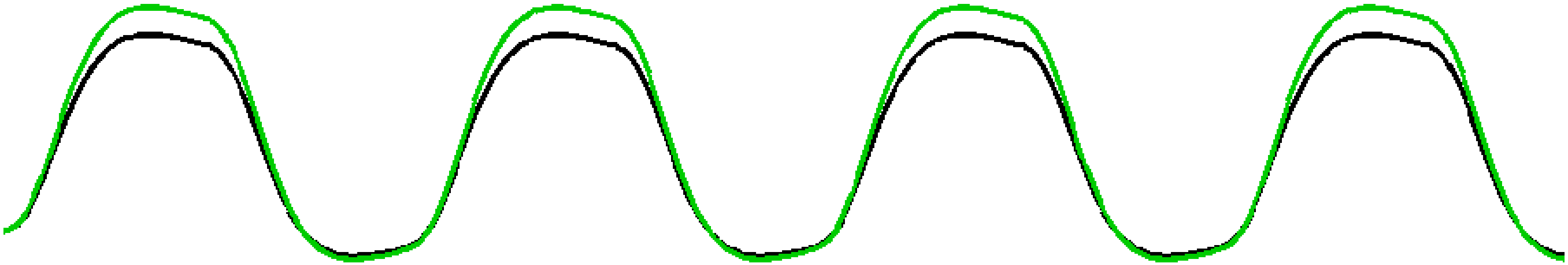
Power Relative to Optimal



(outlier severity)

Mumford & Nichols (2006)

Why isn't this easy?



- Two sources of variation
- Multiple components of error in each

Modelling 2nd level non-sphericity

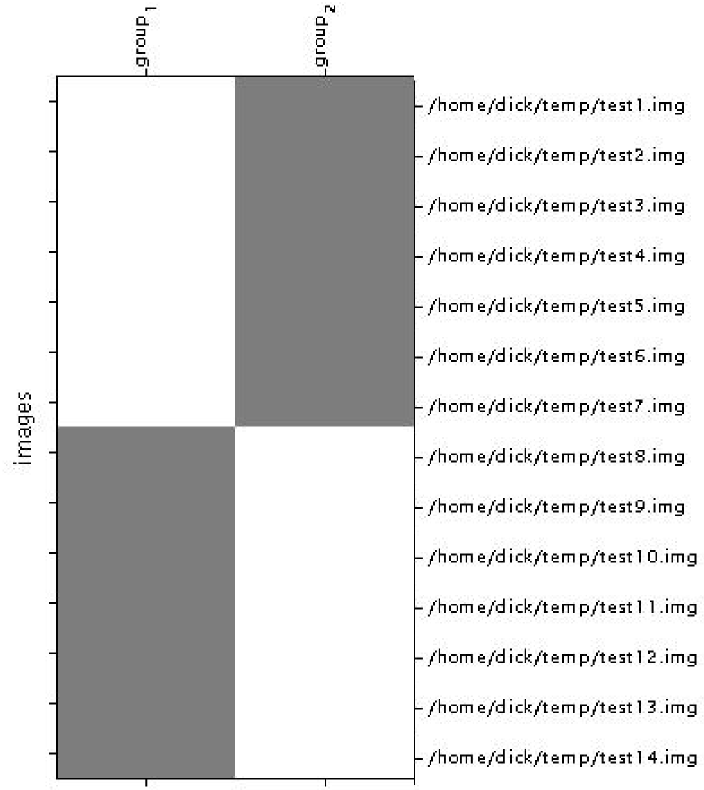
A more flexible summary statistic approach

- 1st level model is just the same
- At 2nd level, we model multiple covariance components
- Use linear combination of basis functions to represent different sources of covariance
- Same estimation using prewhitening approach, and spatial regularisation (cross-voxel pooling)

Example: 2-sample t-test

Model includes one contrast but from different groups – 2-sample t-test e.g. patients vs. controls, young vs. older

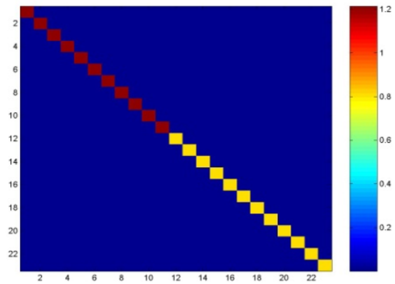
2-sample t-test

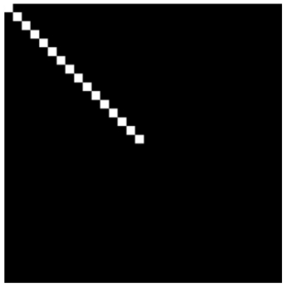
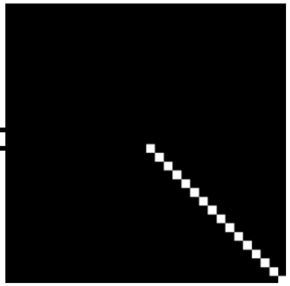


Example: 2-sample t-test

Errors can be Independent but Non-Identical when…

1) Model includes one contrast but from different groups – 2-sample t-test e.g. patients and control groups



$Q_1 =$  $Q_2 =$ 

Variances are likely different between groups

Modelling 2nd level non-sphericity

Error can be Non-Independent and Non-Identical when…

2) Several contrasts per subject are taken to 2nd level i.e., Repeated Measures/ Mixed ANOVA

3) Omnibus test is needed across several basis functions characterising the hemodynamic response

e.g. F-test combining HRF, temporal derivative and dispersion regressors

Example: within-subjects ANOVA

Block design study of semantic processing

- Auditory presentation (SOA = 4 sec)
- Words:

Motion	Sound	Visual	Action
“jump”	“click”	“pink”	“turn”

N = 12

Question:

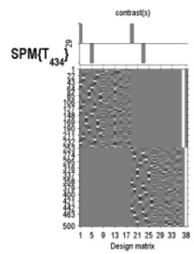
- What regions are generally affected by the semantic content of the words?

Noppeney et al. (2003)

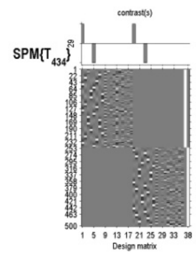
Example: within-subjects ANOVA

1: motion **2: sounds** **3: visual** **4: action**

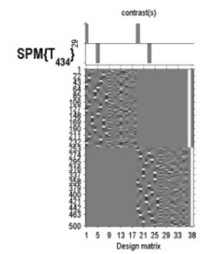
1st level



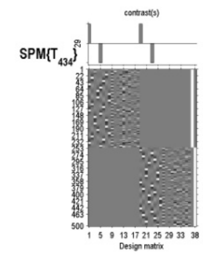
? =



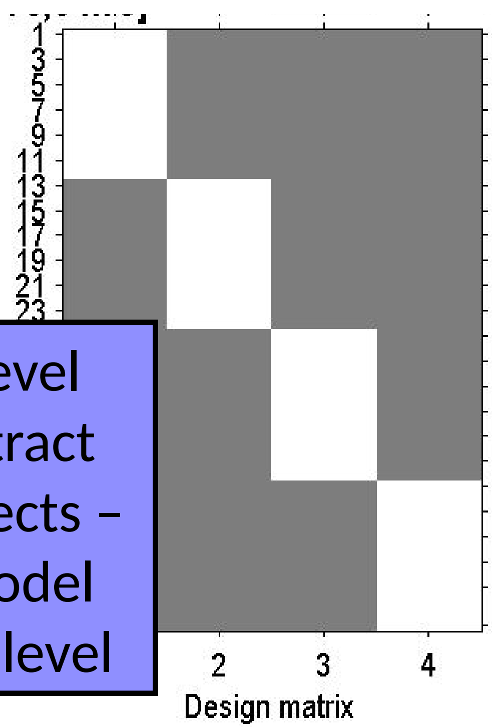
? =

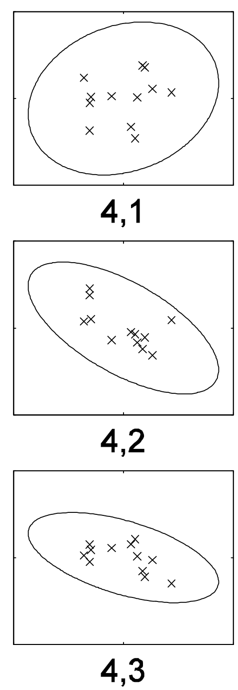


? =



2nd level





N.B. The 1st level contrasts 'subtract out' subject effects – if not, must model these at the 2nd level

For each condition subtraction done at 1st level, taken to group level

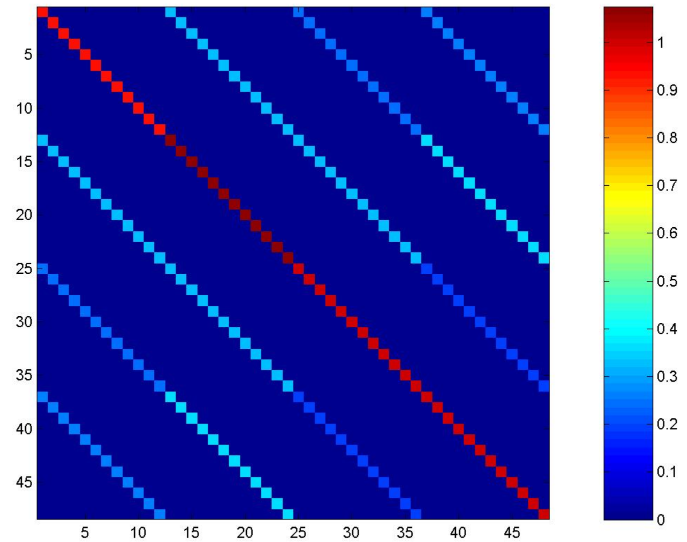
Tests use pooled error term over conditions (more df)

Noppeney et al. (2003)

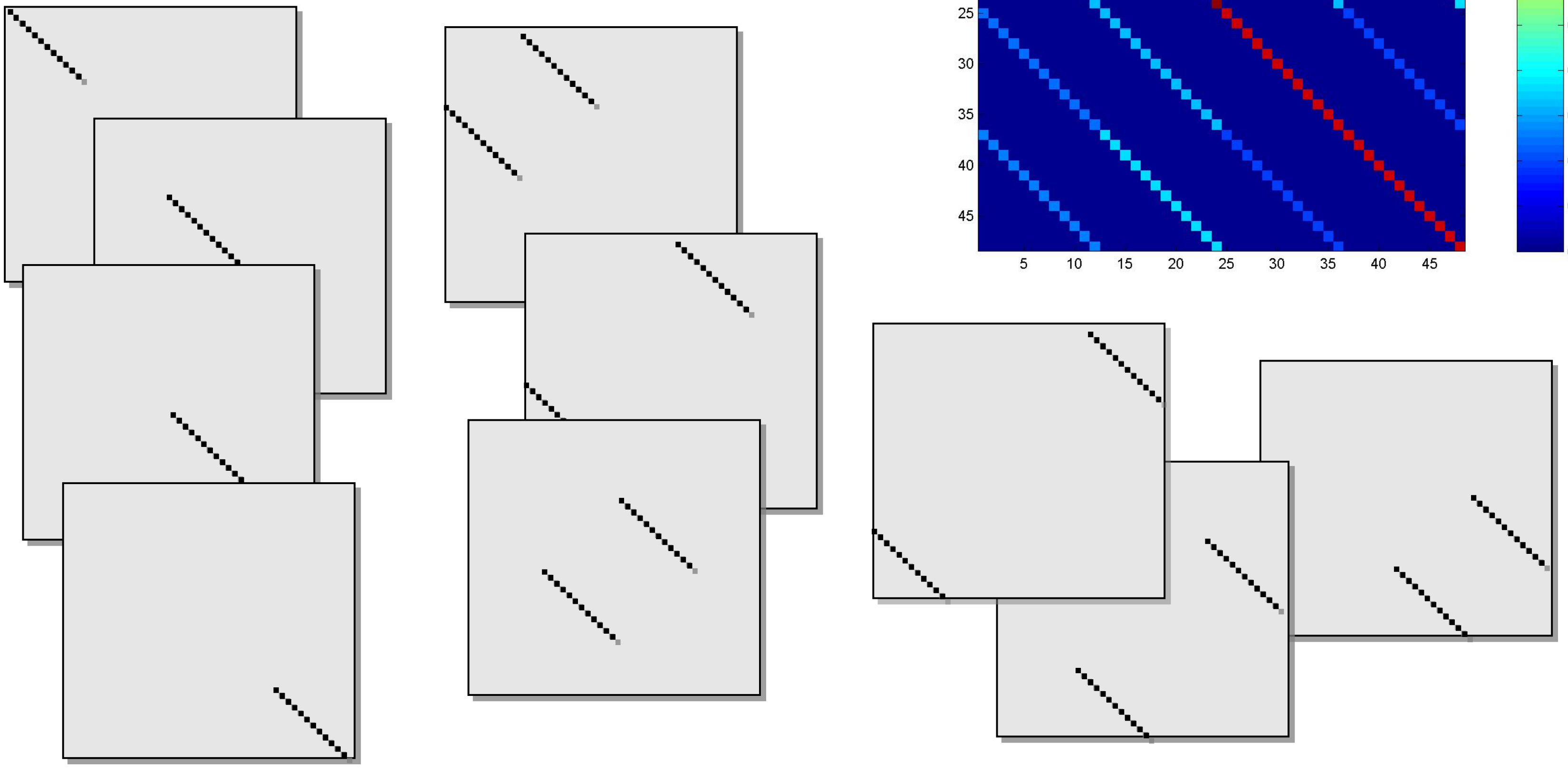
Example: within-subjects ANOVA

Errors are not independent and not identical (4 conditions)

residuals covariance matrix



Q_k's:



Modelling 2nd level non-sphericity

Assumptions

- As at 1st level, assumptions of cross-voxel pooling: homogenous across 'active' voxels
- Within subject covariance still homogenous
- Unlike HF: pooled variance at 2nd level

Advantages/ disadvantages

- Fast relative to 'full' linear mixed models
- May be more sensitive than HF if assump. met
- Flexibility of possible 2nd level models

Summary

fMRI models need to take account of

- Hierarchical nature of data
- Multiple sources of variability at each level

Estimation & correction for resulting nonsphericity

- Some assumptions
- If correct, optimise estimation & inference
- SPM enables very flexible 2nd level models

2-stage GLM

Single subject

Each has an independently acquired set of data
These are modelled separately
Models account for **within subjects variability**
Parameter estimates apply to individual subjects

1st level

Single subject **contrasts of parameter estimates** taken forward to 2nd level as (spm_con*.img) **'con images'**

Group/s of subjects

To make population inferences, 2nd level models account for **between subjects variability**
Parameter estimates apply to group effect/s

2nd level

Statistics compare **contrasts of 2nd level parameter estimates to a 2nd level error**

References

- Simple group fMRI modeling and inference. Mumford & Nichols, Neuroimage, 2009
- *Statistical Parametric Mapping: The Analysis of Functional Brain Images*, 2007.
- *Generalisability, Random Effects & Population Inference*. Holmes & Friston, NeuroImage, 1999.
- *Classical and Bayesian inference in neuroimaging: theory*. Friston et al., NeuroImage, 2002.
- *Classical and Bayesian inference in neuroimaging: variance component estimation in fMRI*. Friston et al., NeuroImage, 2002.
- *Mixed-effects and fMRI studies*. Friston et al. Neuroimage, *2005*
- *Analysis of Variance*. Penny. Chapter in SPM book, copy at: https://www.fil.ion.ucl.ac.uk/~wpenny/drafts/spm-book/anova.pdf
- *Flexible factorial tutorial* by Glascher and Gitelman, 2008, copy at: http://www.sbirc.ed.ac.uk/cyril/download/Contrast_Weighting_Glascher_Gitelman_2008.pdf

SPM