



Experimental Design for fMRI

OHBM Advanced fMRI Educational Course 2014

Thomas Liu

UCSD Center for Functional MRI

Experimental Design



Condition 1

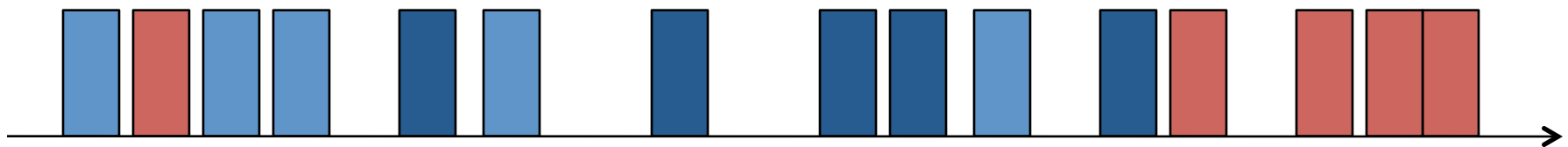
Condition 2

Condition 3

Design 1



Design 2

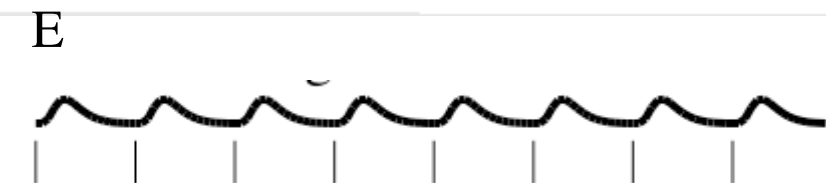
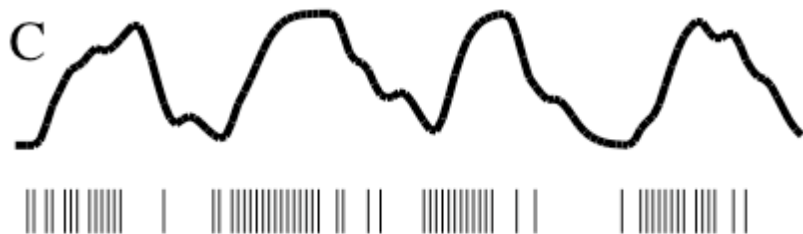
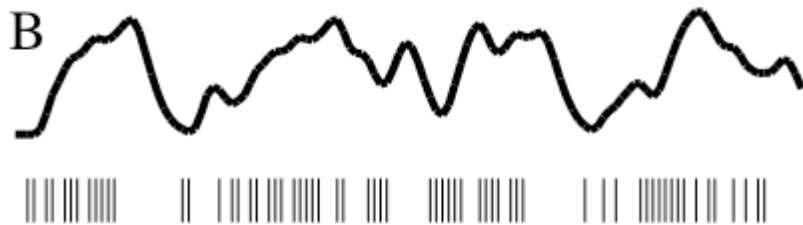


Why worry about design?

- Scans are expensive.
- Subjects can be difficult to find.
- fMRI data are noisy
- A badly designed experiment is unlikely to yield publishable results.

If your result needs a statistician then you should design a better experiment. --*Baron Ernest Rutherford*

Example Stimulus Patterns



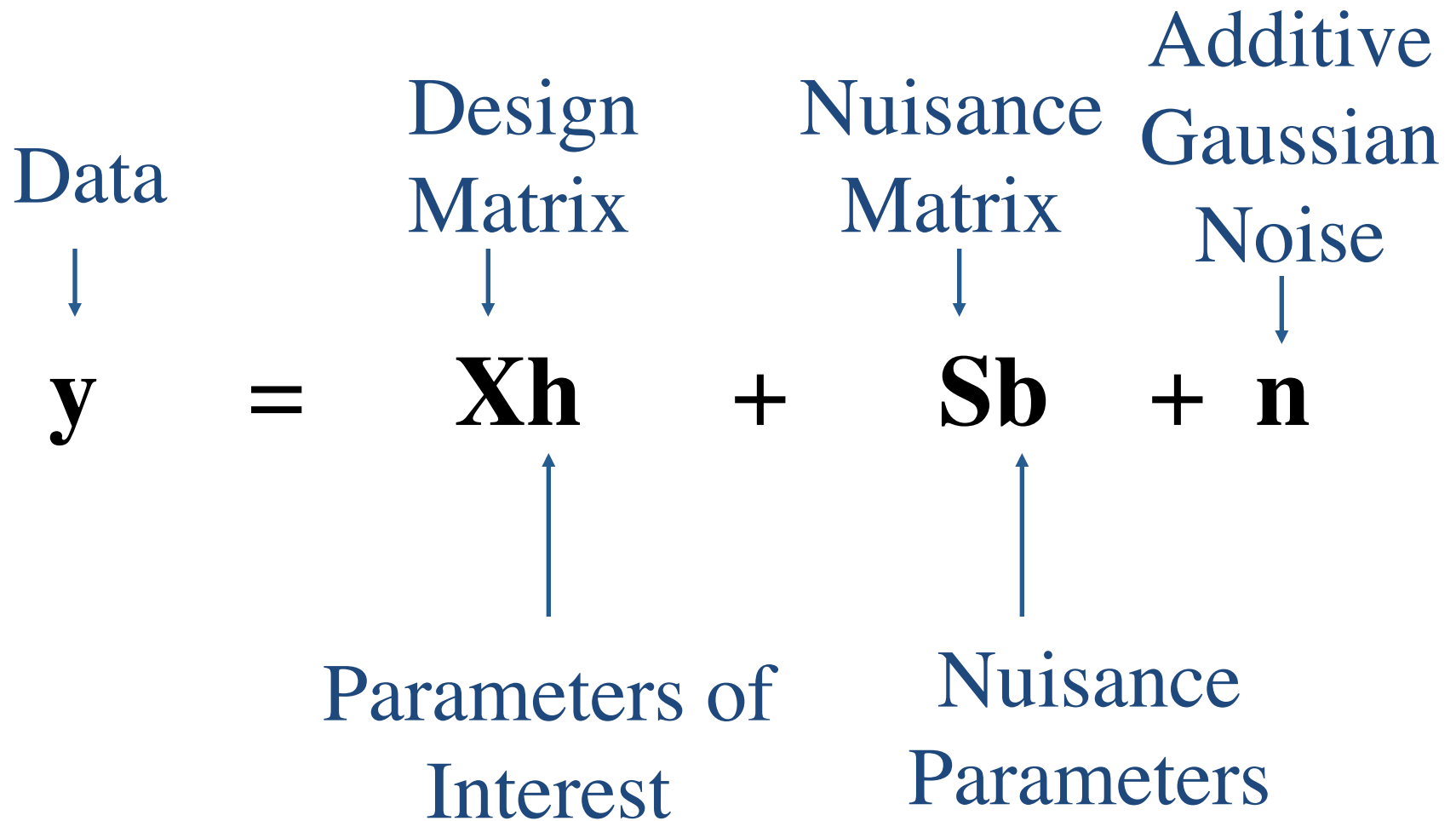
Which is the best design?

It depends on the experimental question.

What to optimize?

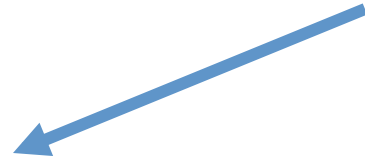
- Statistical Efficiency: maximize contrast of interest versus noise.
- Psychological factors: is the design too boring? Minimize anticipation, habituation, boredom, etc.

General Linear Model

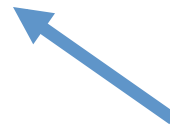


Test Statistic

Stimulus, neural activity, field strength, vascular state



$$t \propto \frac{\text{parameter estimate}}{\sqrt{\text{variance of parameter estimate}}}$$



Thermal noise, physiological noise, low frequency drifts, motion

Also depends on Experimental Design!!!

Hemodynamic
Response
Function (HRF)

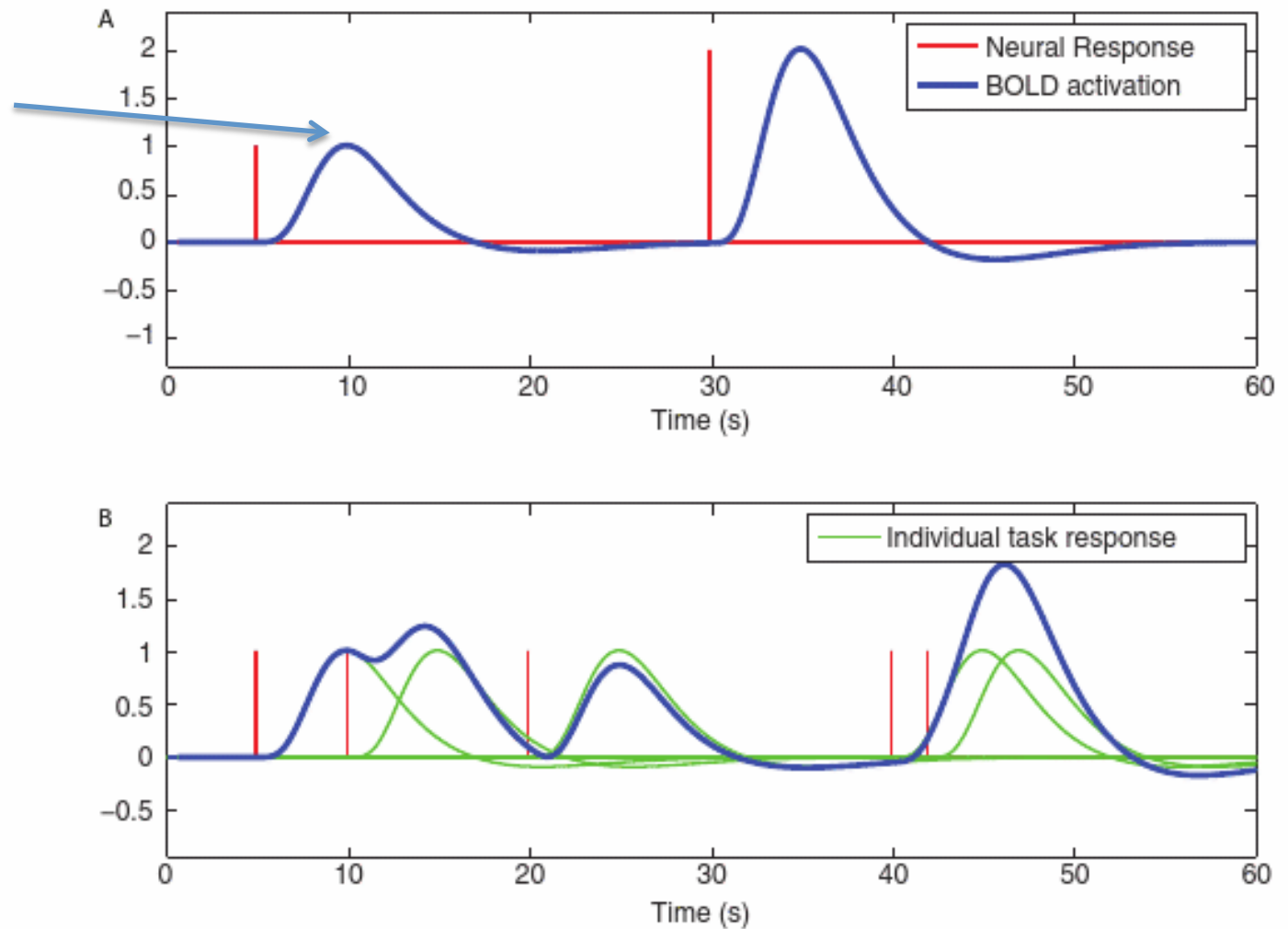


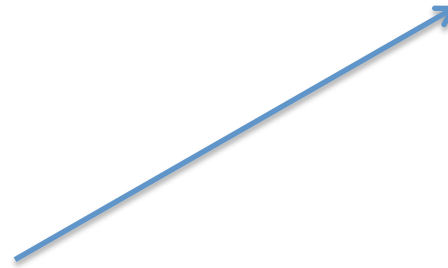
Figure 5.3. Examples of linear time invariance. Panel A illustrates that when a neural signal is twice another, the resulting BOLD activation is also twice as large. Panel B shows how the signals for separate trials, shown in green, add linearly to get the BOLD activation.

From Poldrack et al , 2012

Efficiency

$$t \propto \frac{\text{parameter estimate}}{\sqrt{\text{variance of parameter estimate}}}$$

$$\text{Efficiency} \propto \frac{1}{\text{Variance of Parameter Estimate}}$$

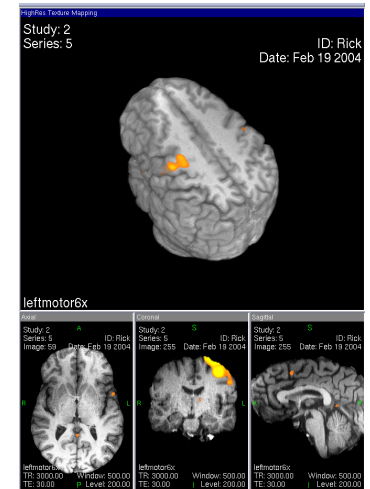


Amplitude of the response
Coefficients of the Hemodynamic Response
Coefficients of Basis Functions

Questions and Assumptions

Where is the activation?

- Assume we know the shape of the HRF but not its amplitude.
- Or sometimes assume something about the shape



What does the HRF look like?

- Assume we know the shape of the HRF but not its amplitude.
- Or sometimes assume something about the shape

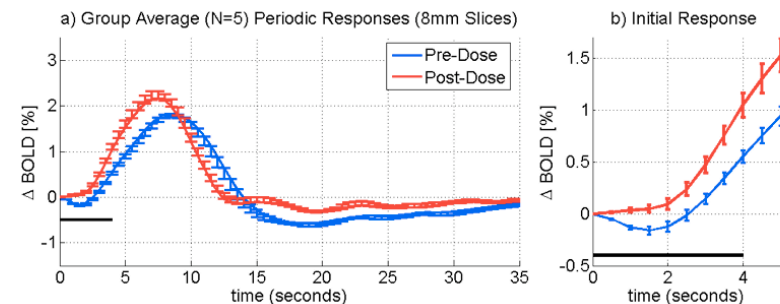


Image-based Example

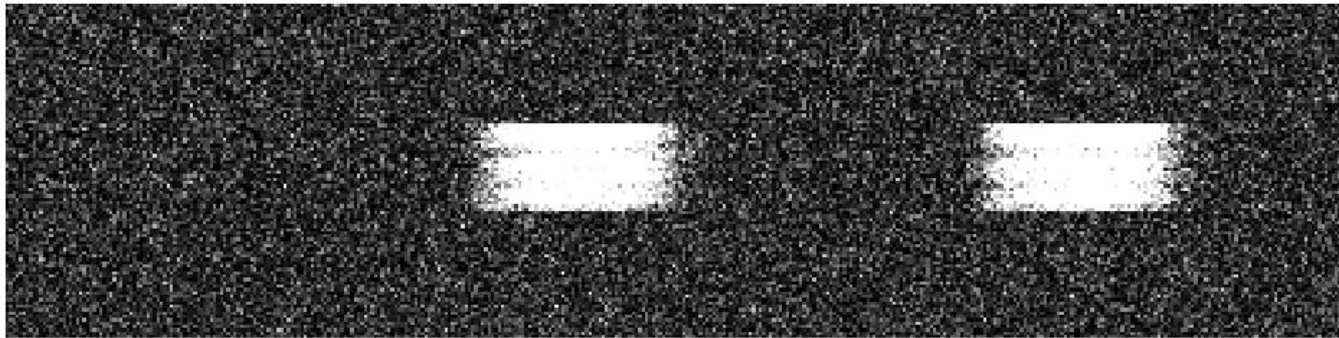
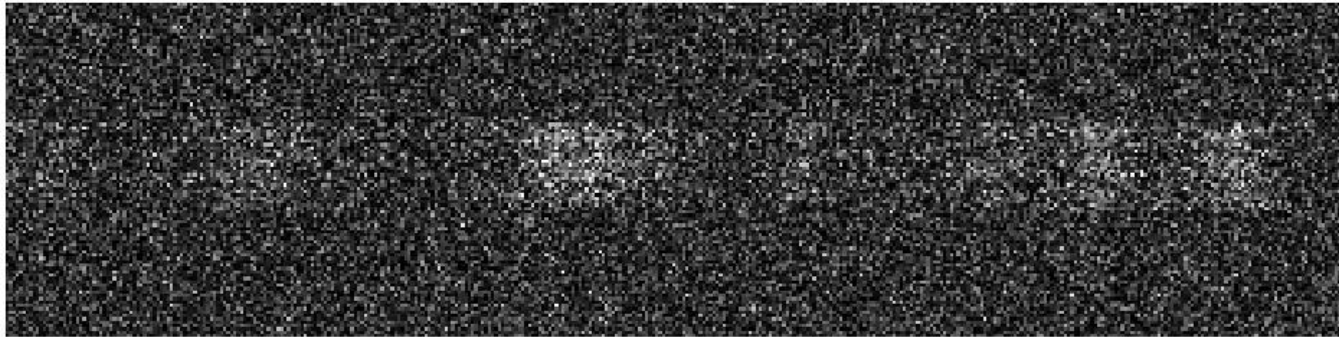
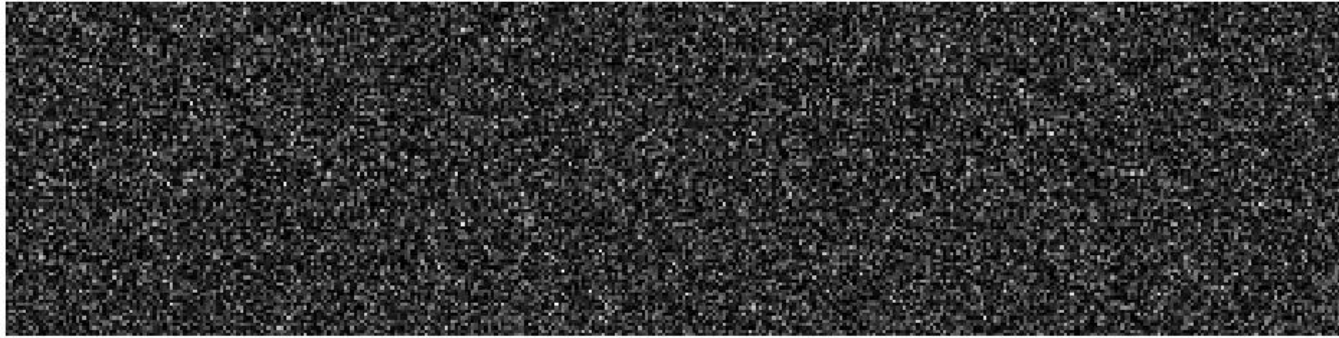


Image-based Example

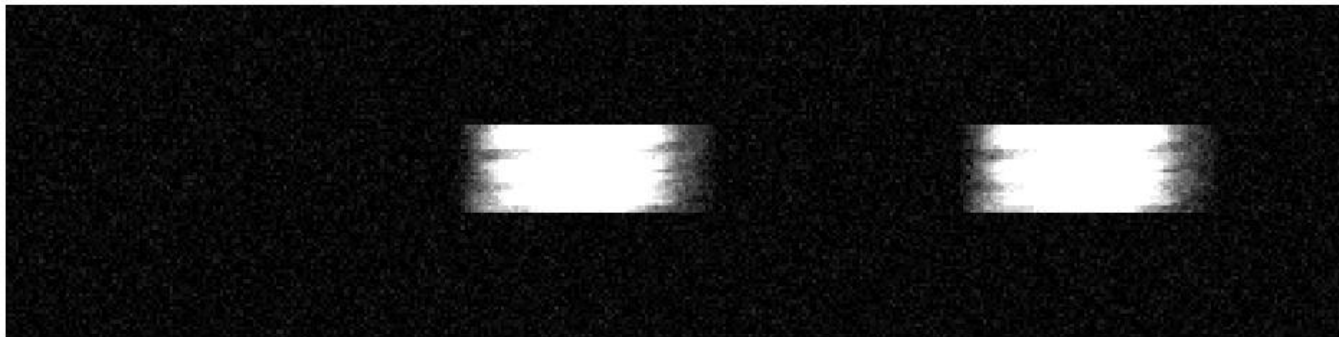
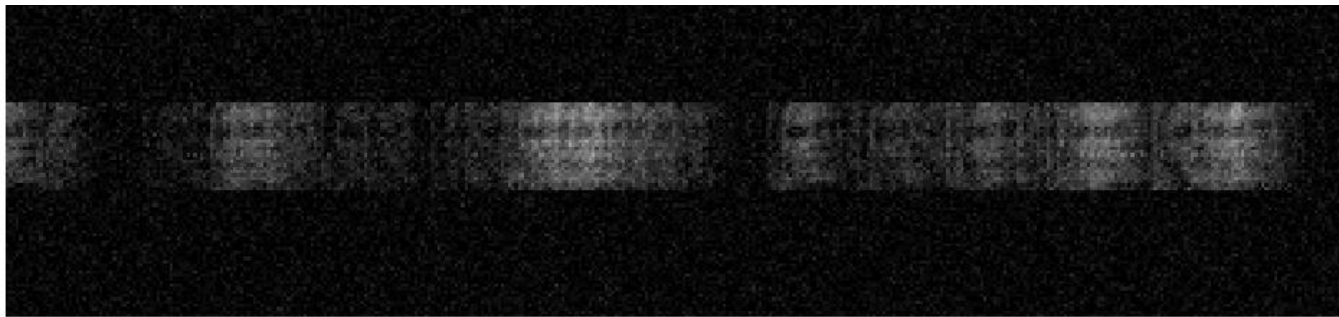
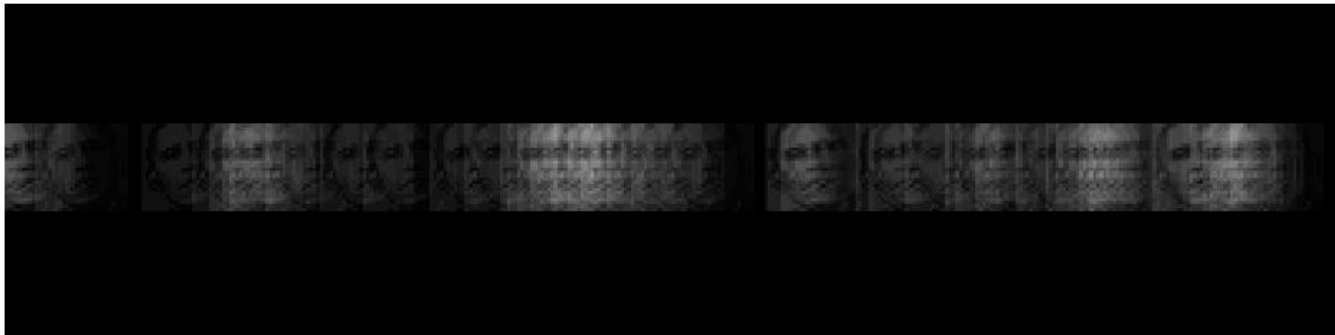
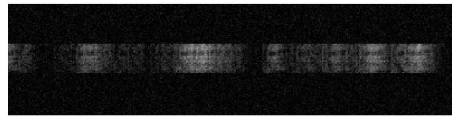


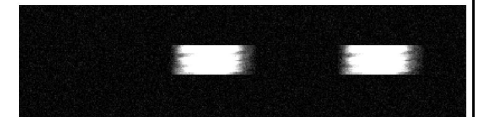
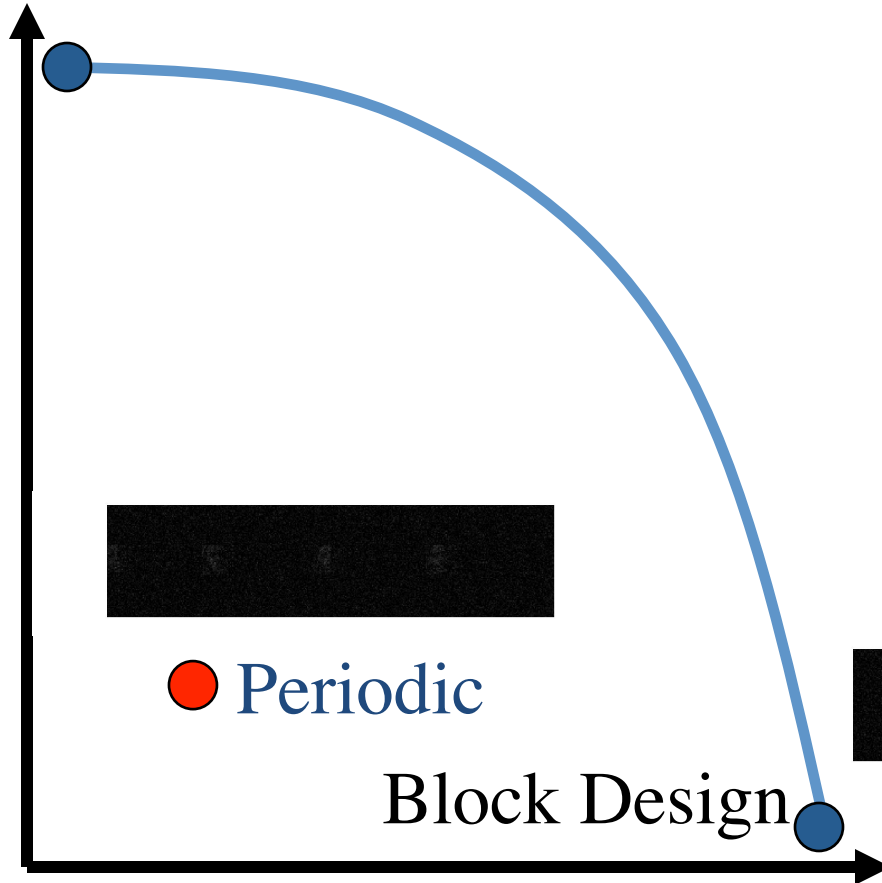
Image-based Example



Fundamental Trade-off

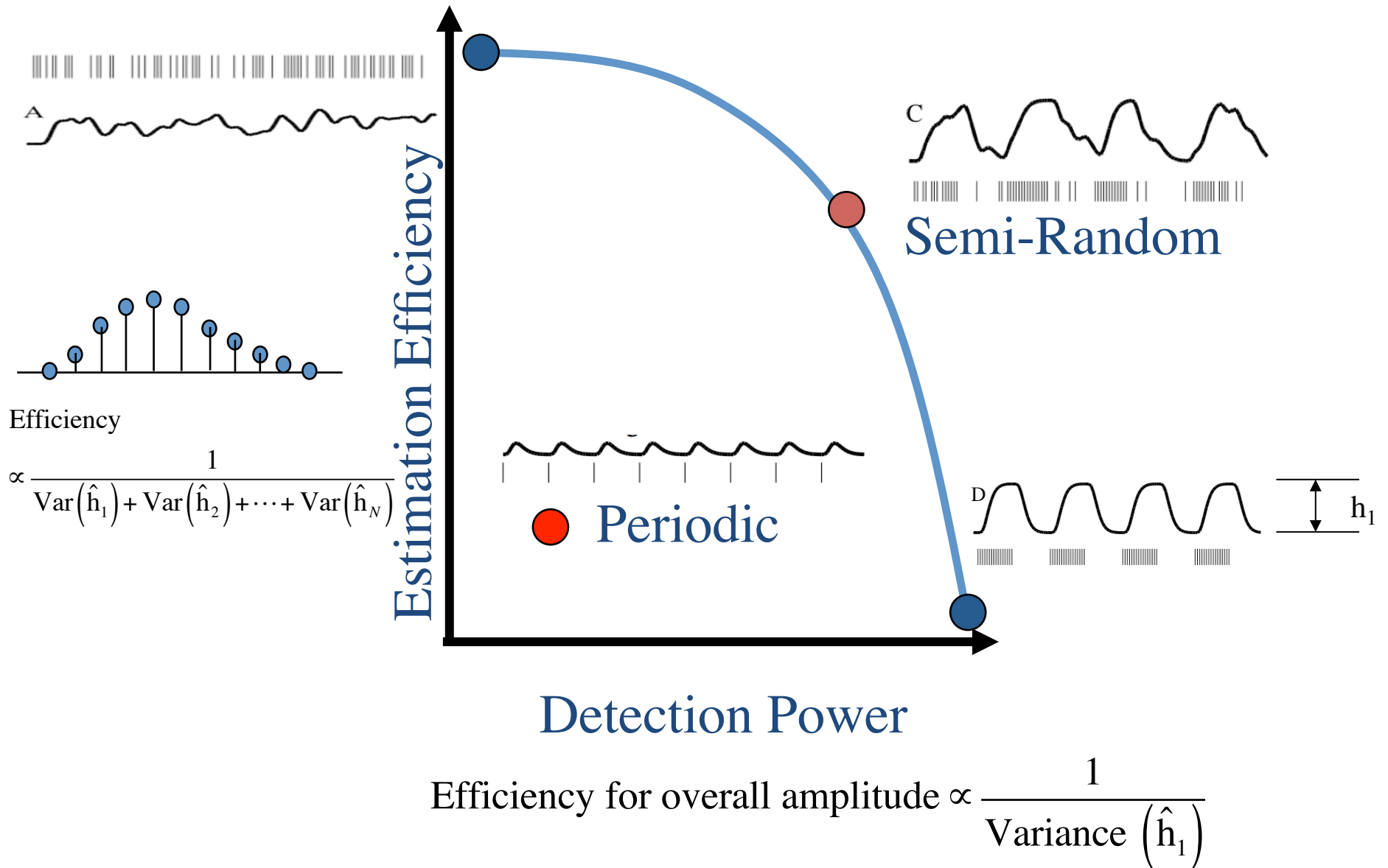


Shape
Estimation
Efficiency



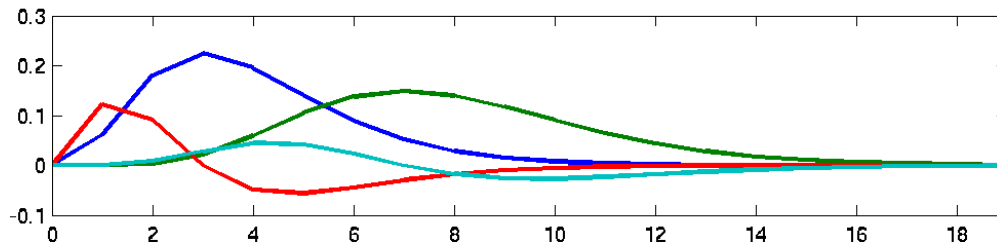
Detection Power
= Estimation Efficiency
for overall amplitude

Fundamental Trade-off

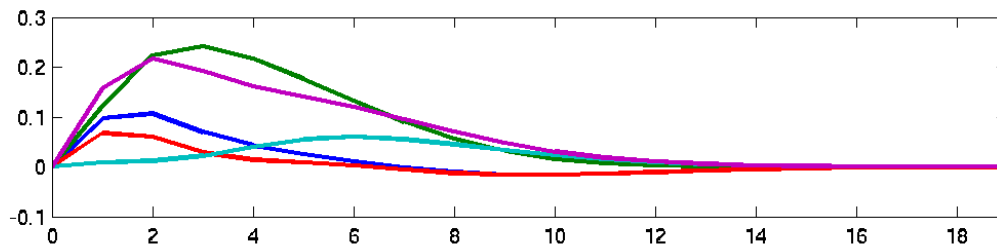


Basis Functions

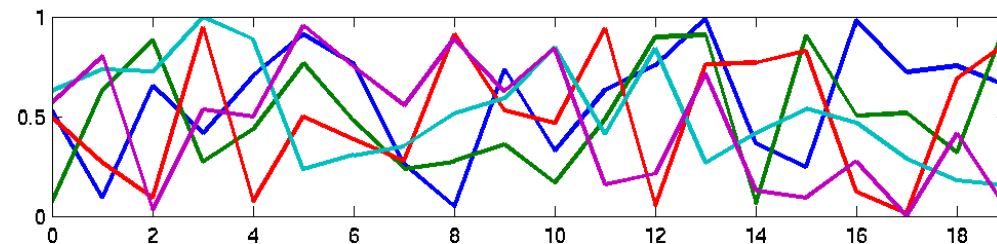
If we know something about the shape, we can use a
basis function expansion : $\mathbf{h} = \mathbf{Bc}$



4 basis functions



5 random HDRs using
basis functions

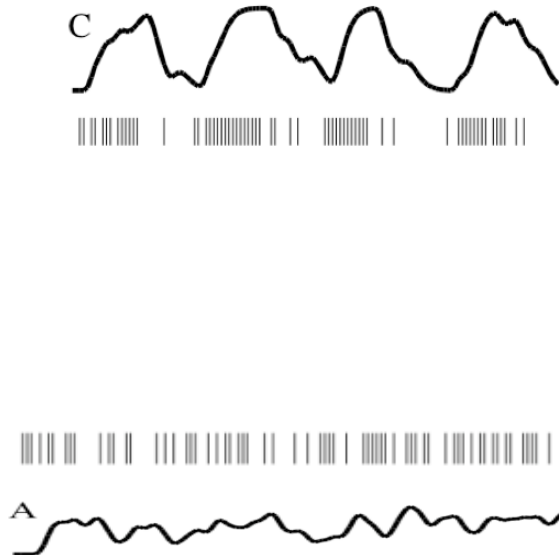


5 random HDRs w/o
basis functions

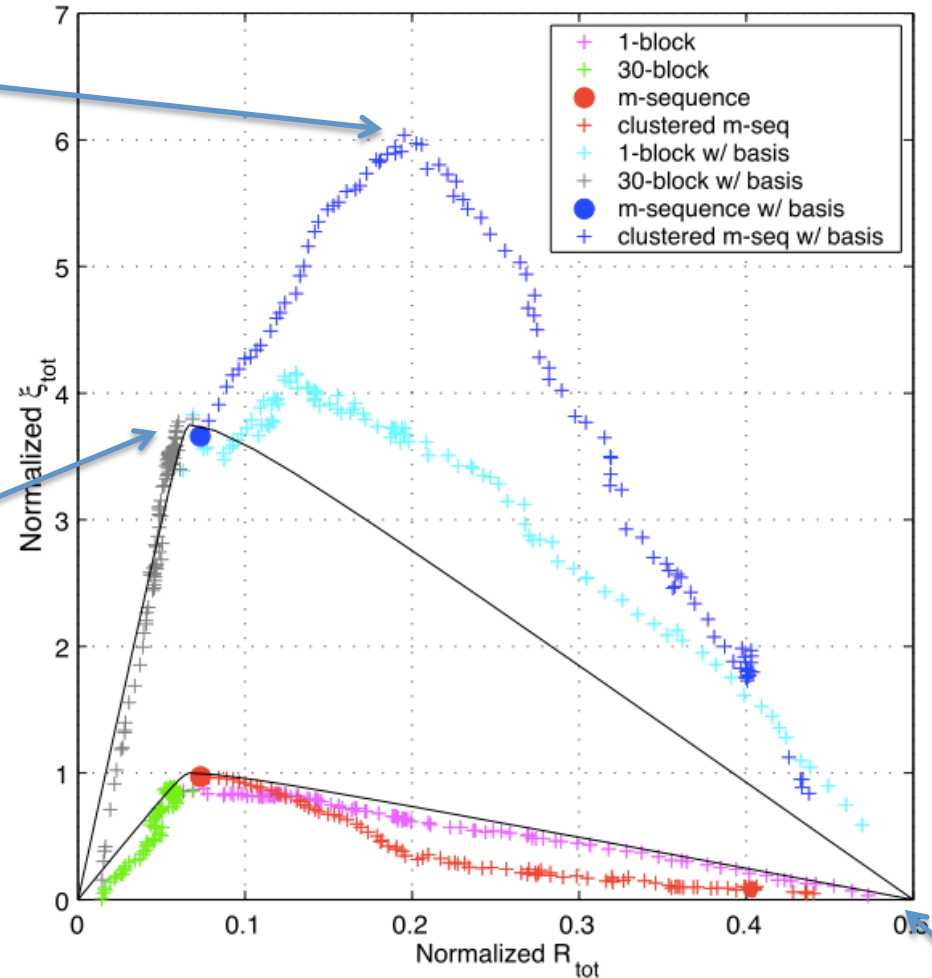
Here if we assume basis functions, we only need to estimate 4 parameters as opposed to 20.

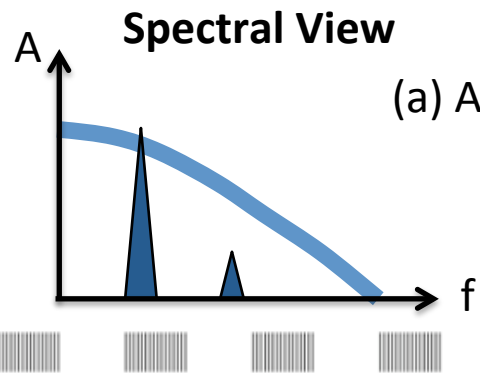
Trade-off w/ basis functions

Semi-Random

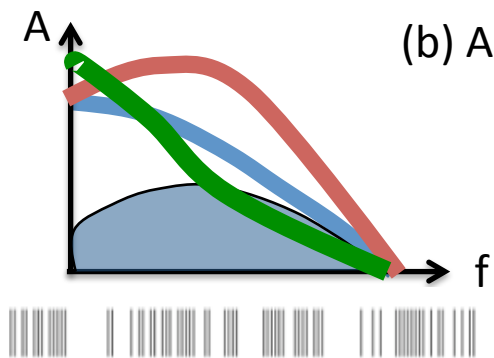
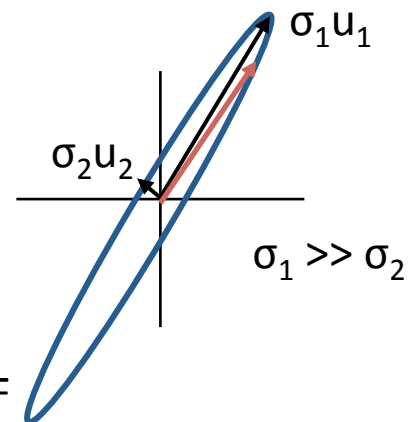
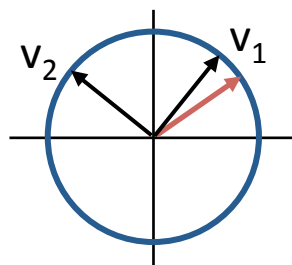


(a) Efficiency and Power for $Q = 2$, with and without basis functions

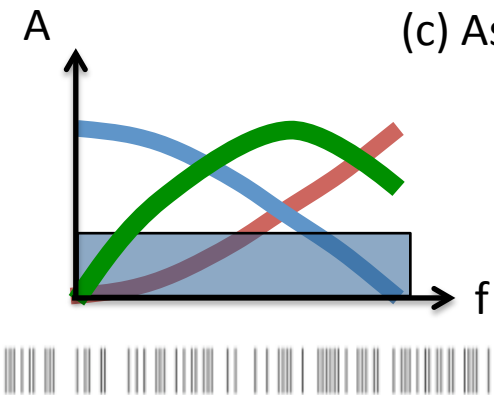
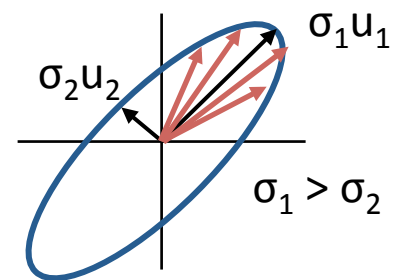
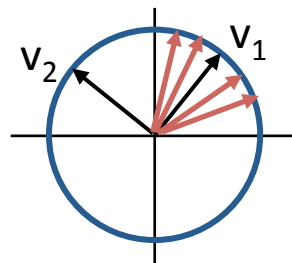




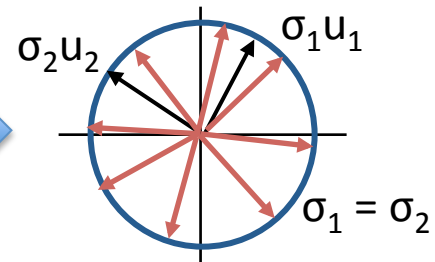
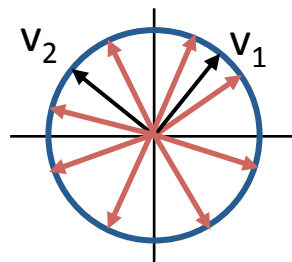
(a) Assume total knowledge about HRF

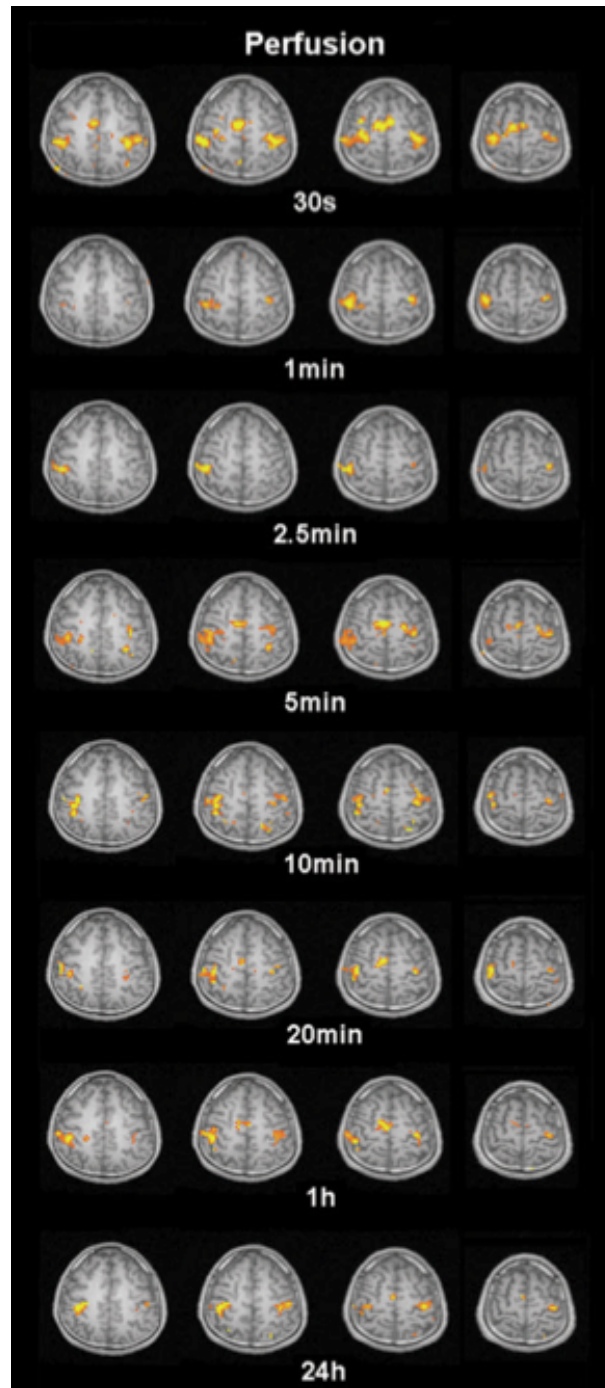
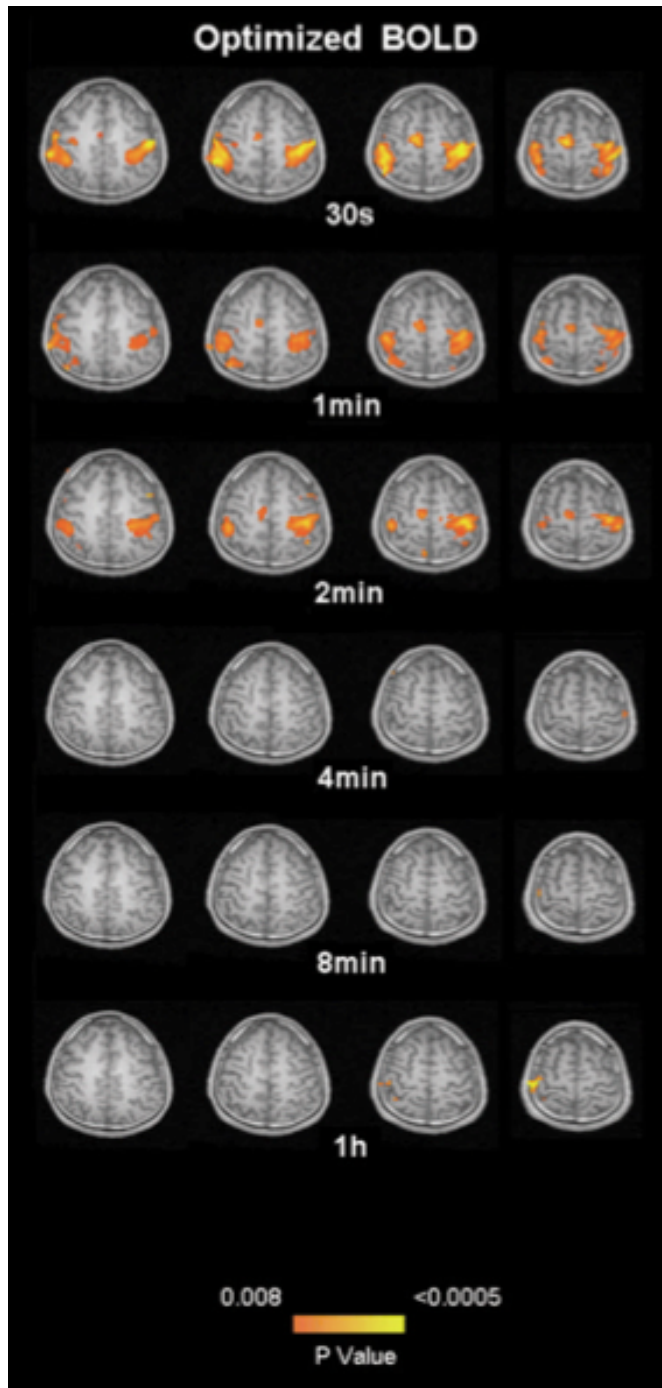


(b) Assume some knowledge about HRF



(c) Assume no knowledge about HRF

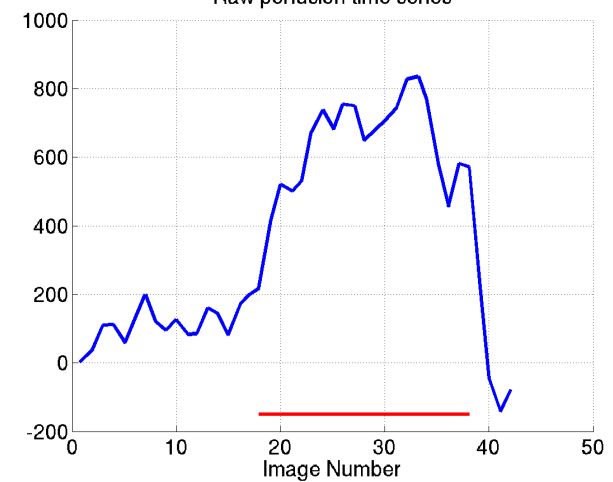
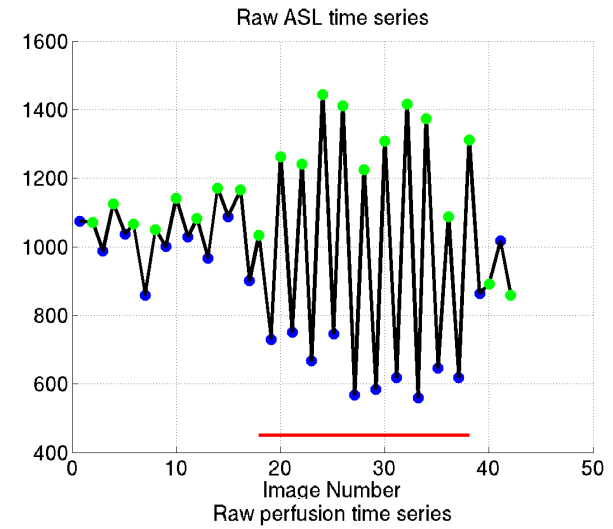
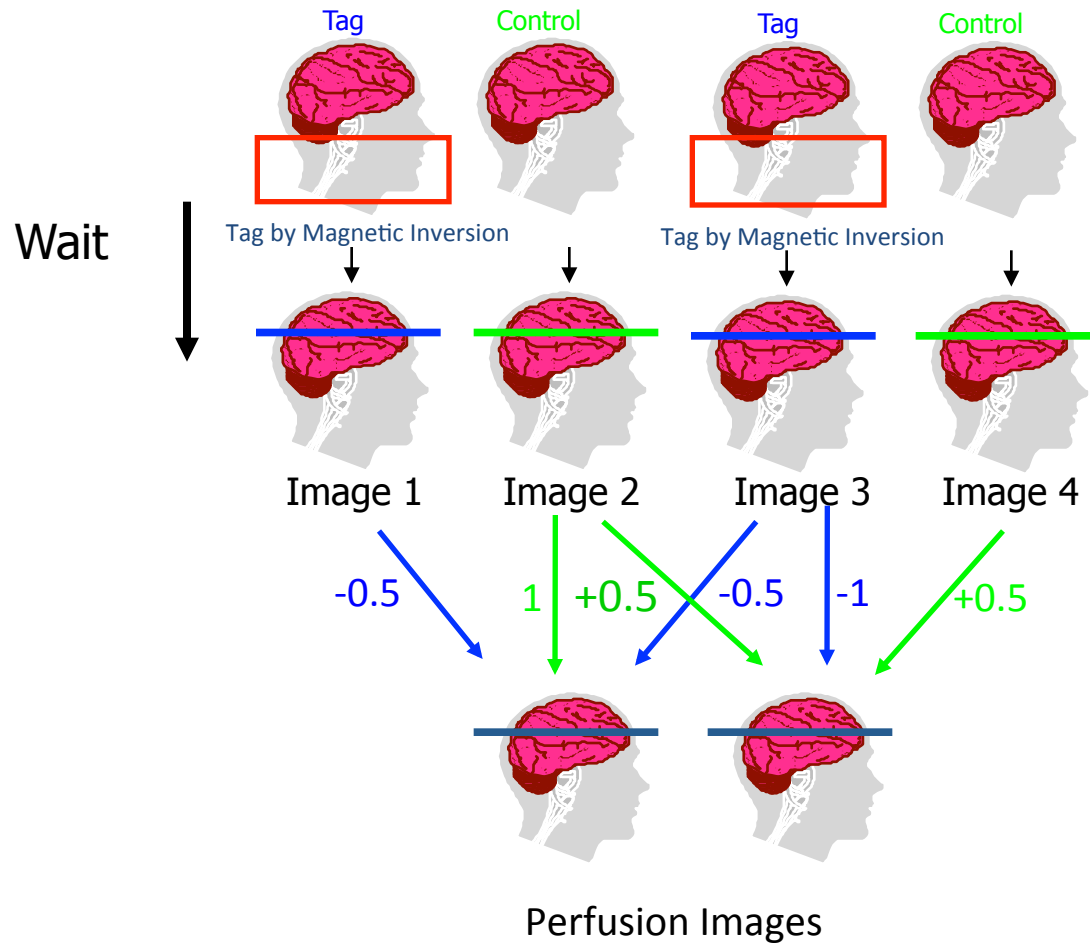




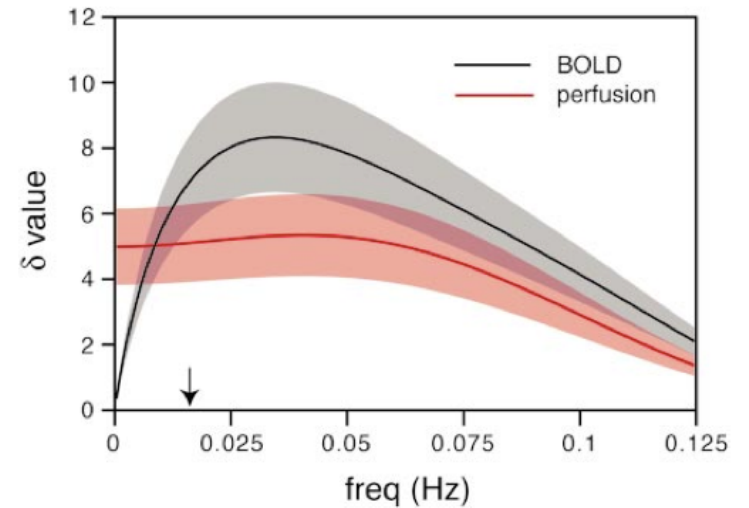
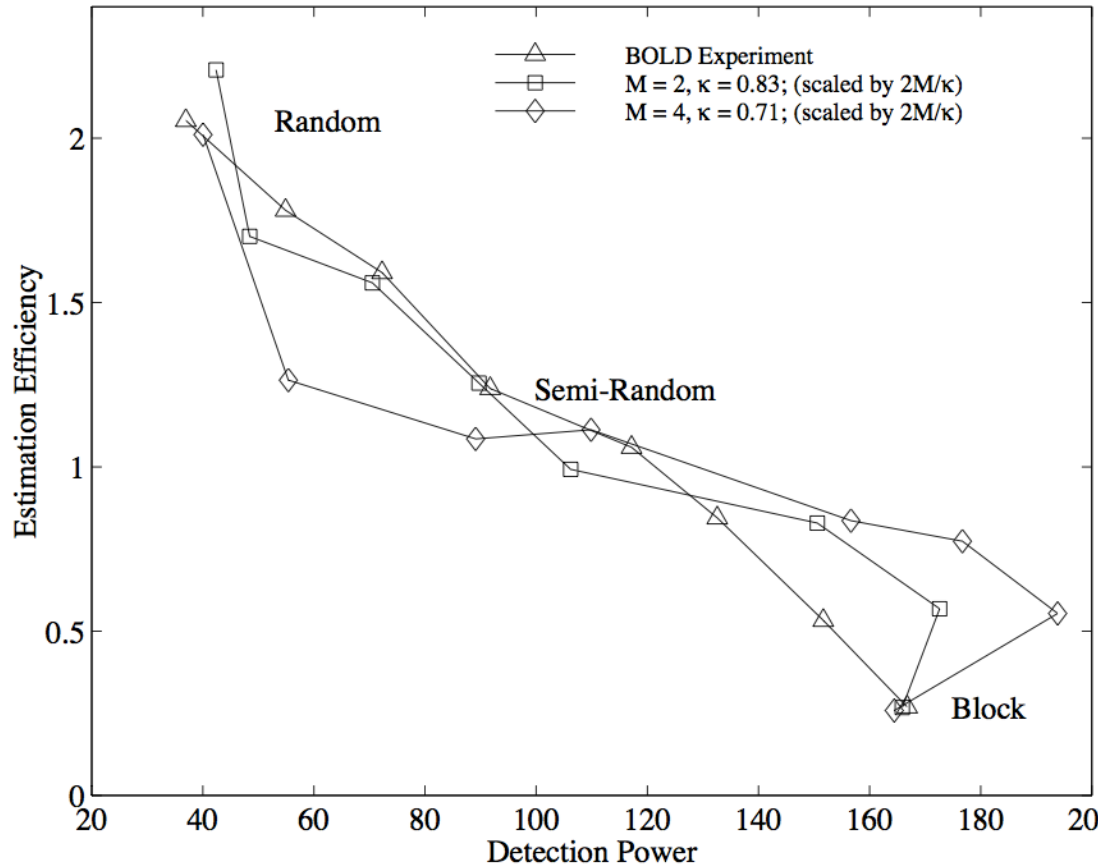
Performance
as function
of task frequency

Wang et al
MRM 2003

Arterial Spin Labeling

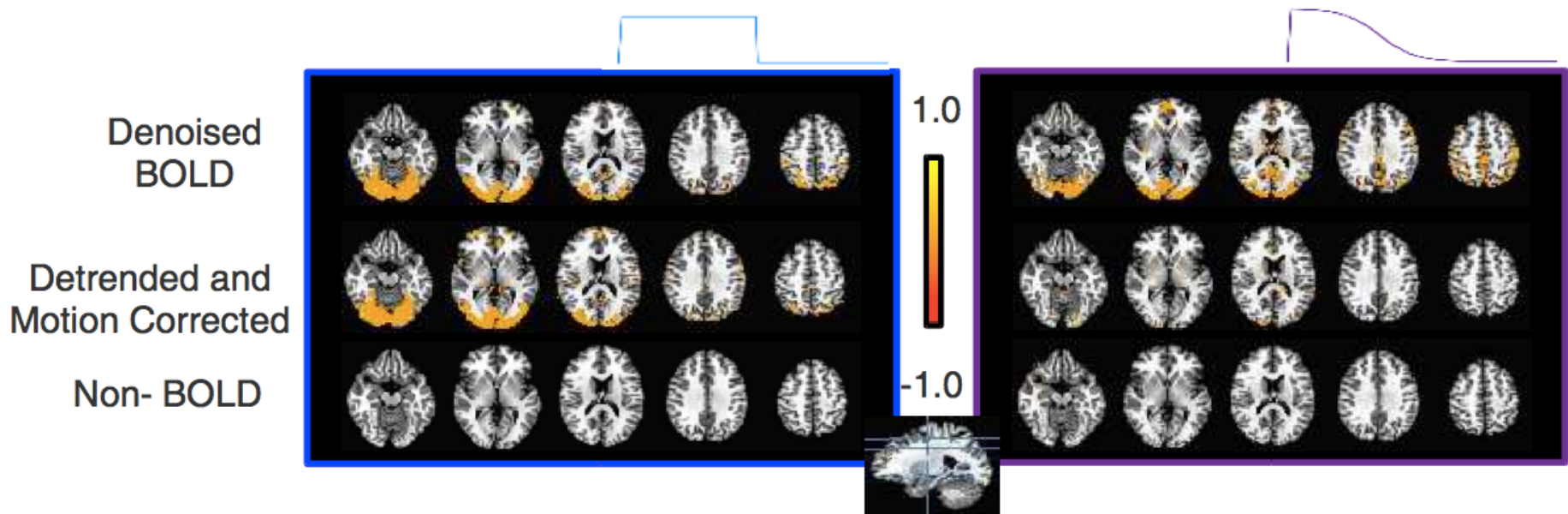
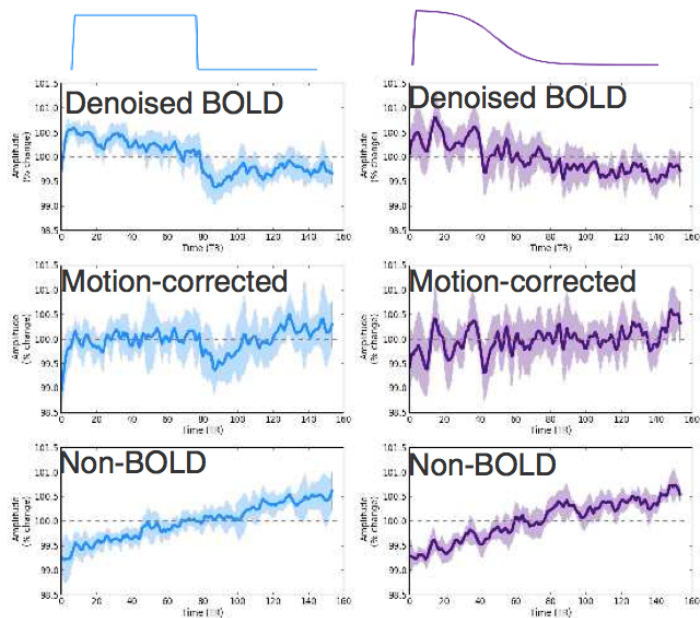


Arterial Spin Labeling



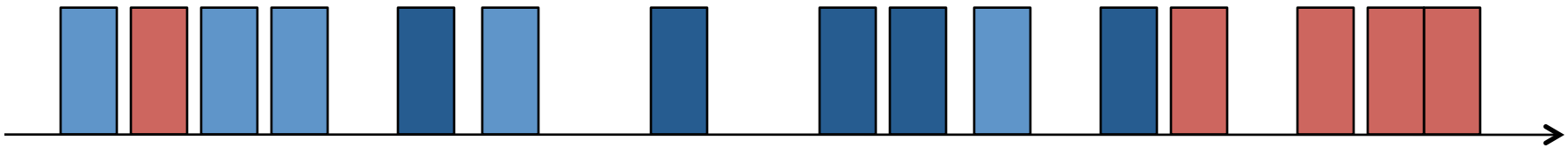
Liu et al 2002; Aguirre et al

Multi-echo BOLD and low-frequency drifts



Evans et al ISMRM 2014; p. 4218; See also Evans et al HBM 2014 Poster 2019

Multiple Trial Types GLM



$$\mathbf{y} = \mathbf{X}\mathbf{h} + \mathbf{S}\mathbf{b} + \mathbf{n}$$

$$\mathbf{X} = [\mathbf{X}_1 \ \mathbf{X}_2 \ \dots \ \mathbf{X}_Q]$$

$$\mathbf{h} = [\mathbf{h}_1^T \ \mathbf{h}_2^T \ \dots \ \mathbf{h}_Q^T]^T$$

Multiple Trial Types Overview

Efficiency includes individual trials and also contrasts between trials.

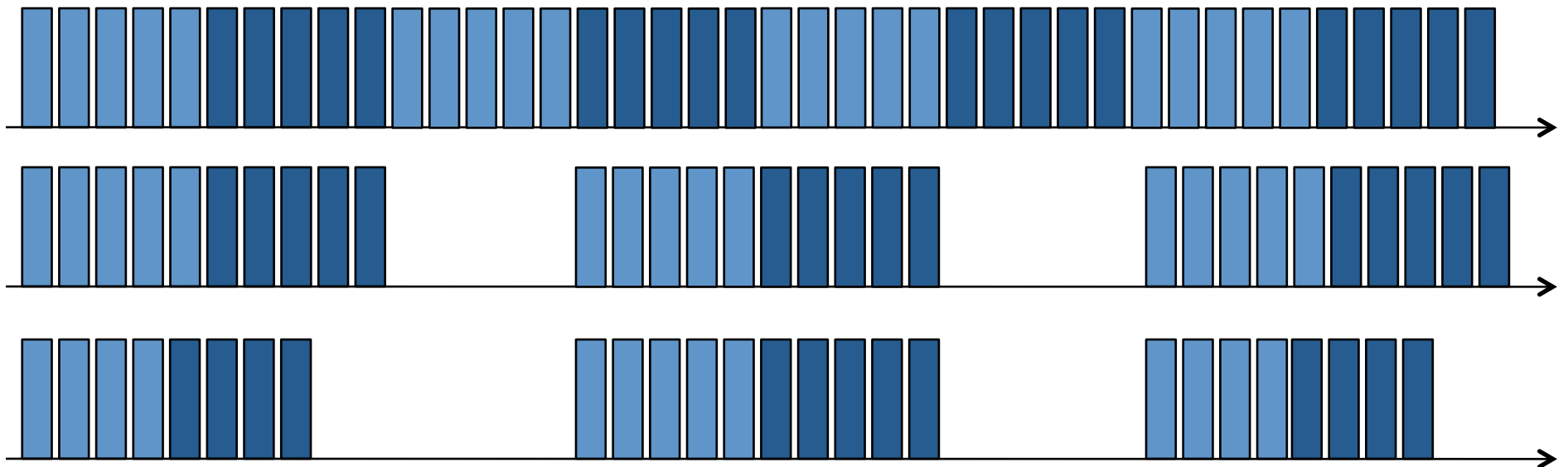
$$R_{tot} = \frac{K}{\left(\begin{array}{l} \text{average variance of HRF amplitude estimates} \\ \text{for all trial types and pairwise contrasts} \end{array} \right)}$$

$$\xi_{tot} = \frac{1}{\left(\begin{array}{l} \text{average variance of HRF estimates} \\ \text{for all trial types and pairwise contrasts} \end{array} \right)}$$

Optimal Frequency

Optimal frequency of occurrence depends on weighting of individual trials and contrasts.

Example: With $Q = 2$ trial types, if only contrasts are of interest $p = 0.5$. If only trials are of interest, $p = 0.2929$. If both trials and contrasts are of interest $p = 1/3$.



Psychological Considerations



Problems with habituation, anticipation, and boredom

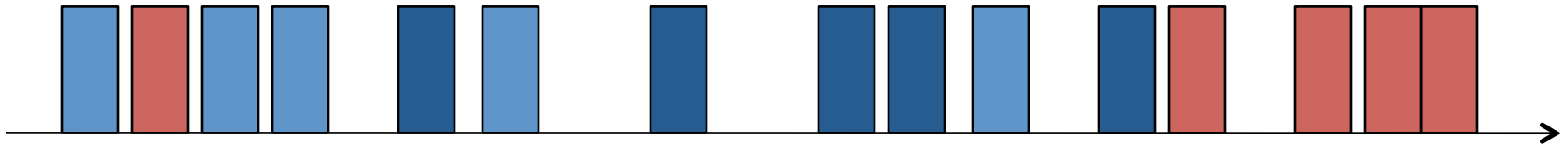


Random



Semi-Random

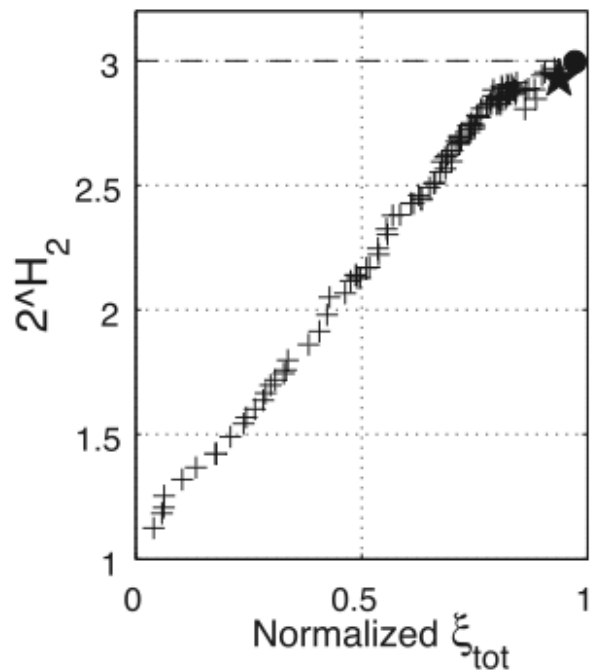
Entropy



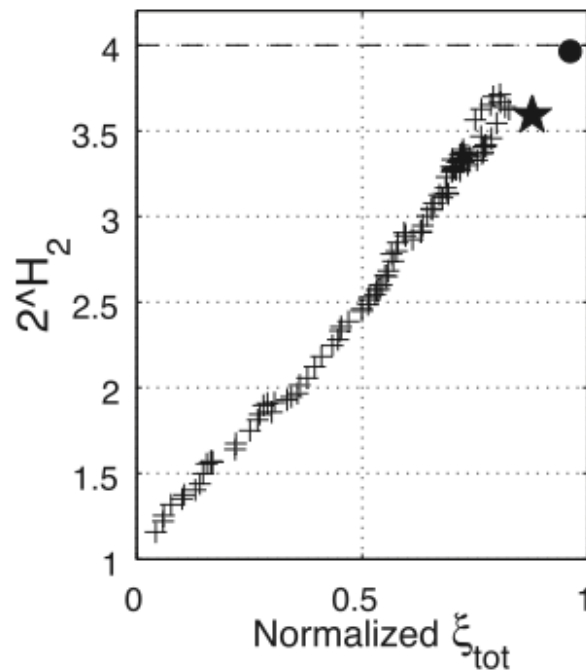
$H = \text{Entropy} = \log_2(\text{number of possible outcomes})$

$2^H = \text{linear measure of randomness} \rightarrow \text{proportional to efficiency}$

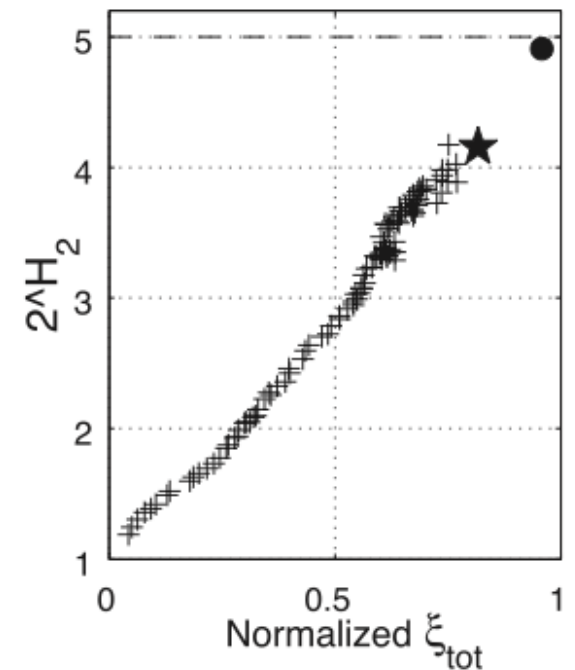
(a) $Q = 2$, 2nd order entropy



(b) $Q = 3$, 2nd order entropy

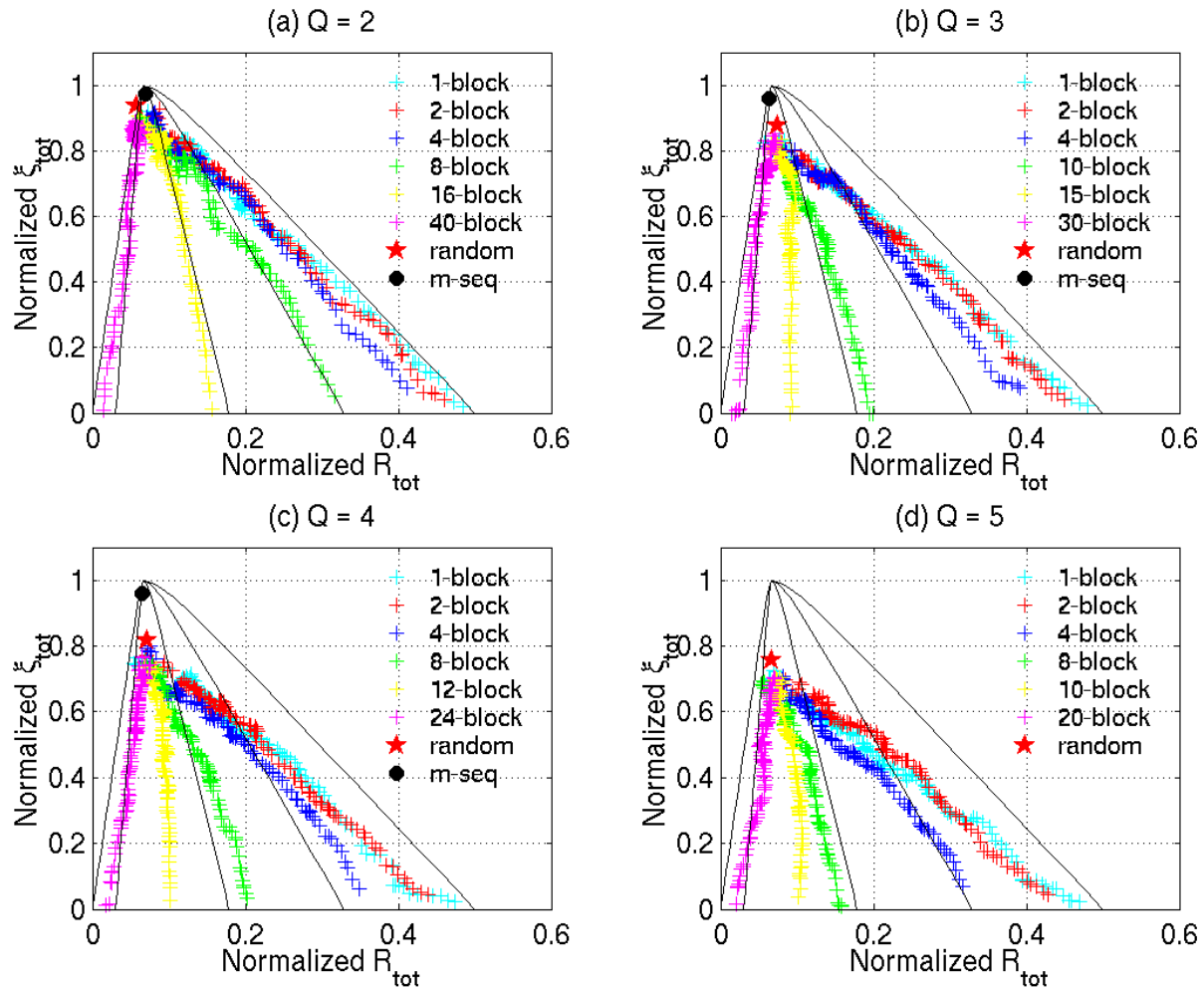


(c) $Q = 4$, 2nd order entropy



Multiple Trial Types Trade-off

Efficiency



Detection Power

Design

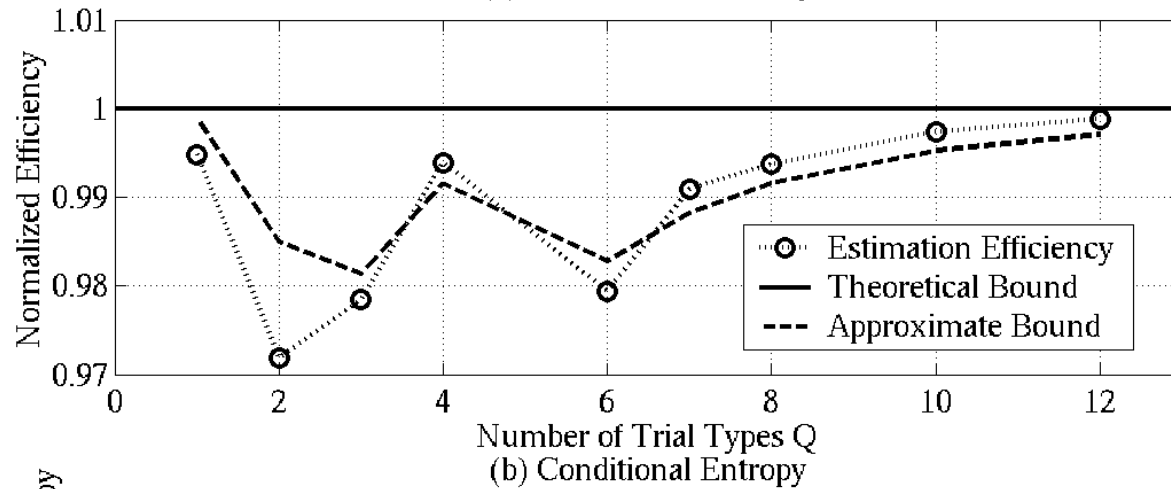
As the number of trial types increases, it becomes more difficult to achieve the theoretical trade-offs. Random search becomes **impractical** and results in **non-optimal** designs.

For unknown HDR, should use an m-sequence based design when possible.

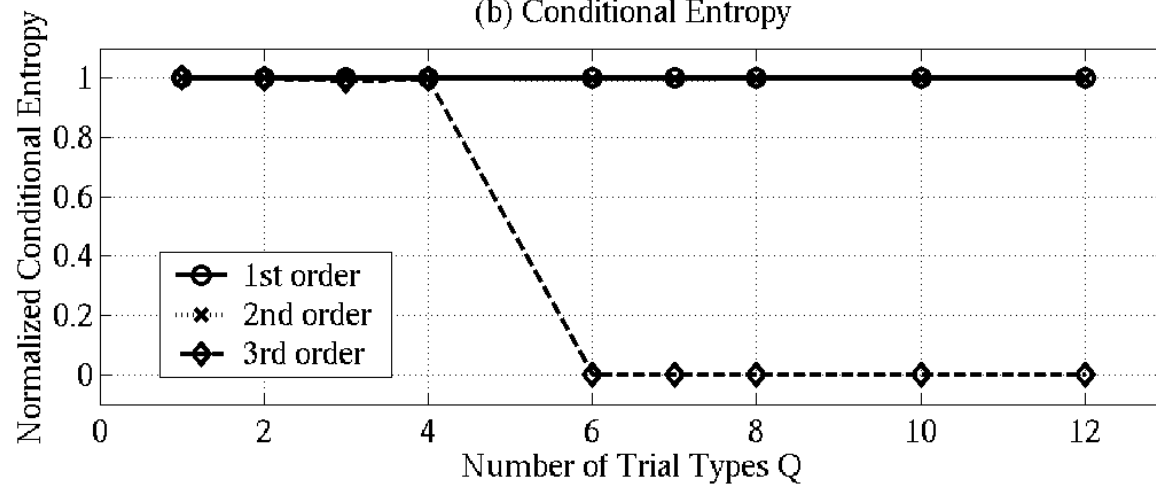
Designs based on block or m-sequences are useful for obtaining intermediate trade-offs or for optimizing with basis functions or correlated noise.

Optimality of m-sequences

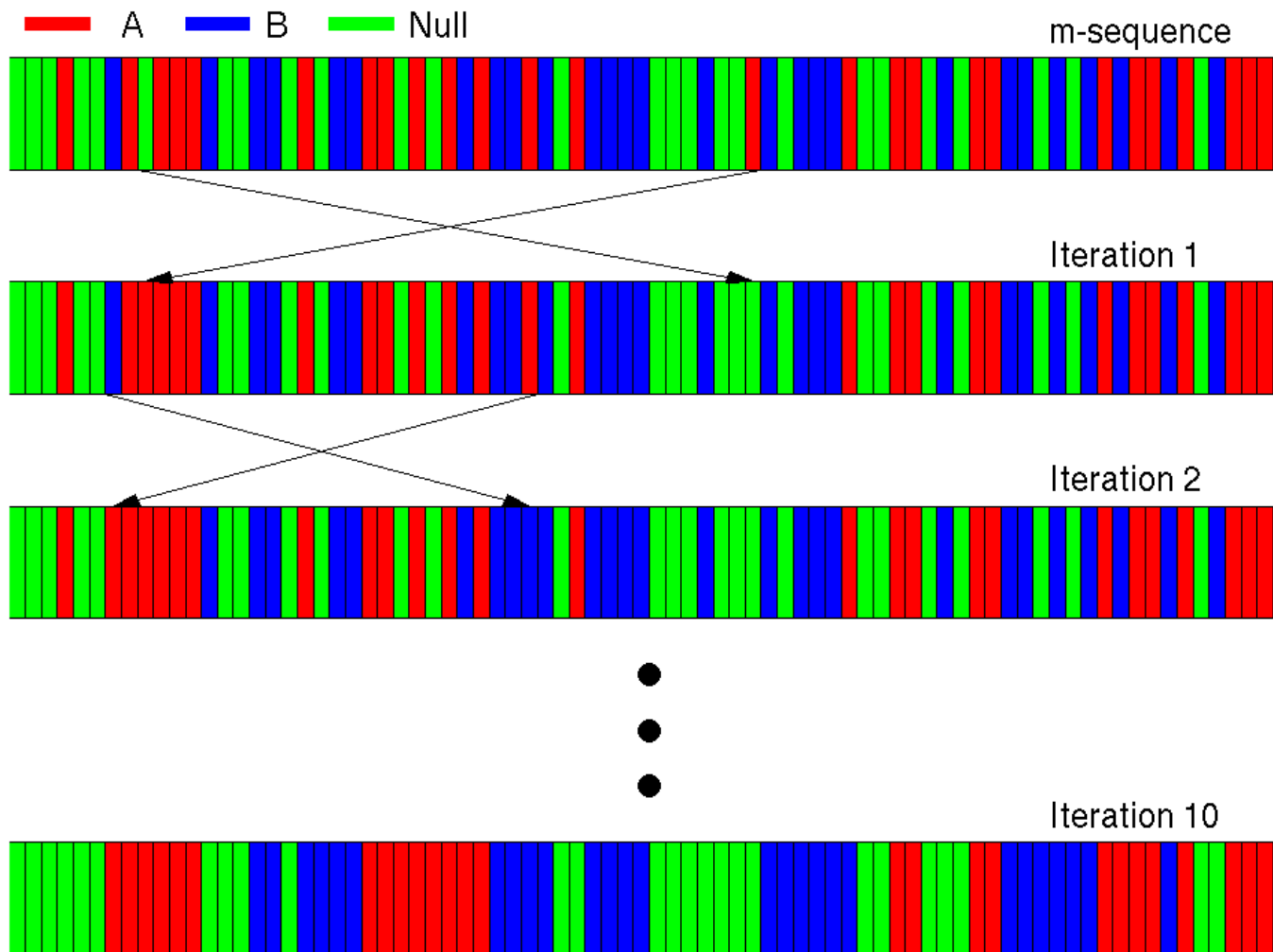
(a) Estimation Efficiency



(b) Conditional Entropy



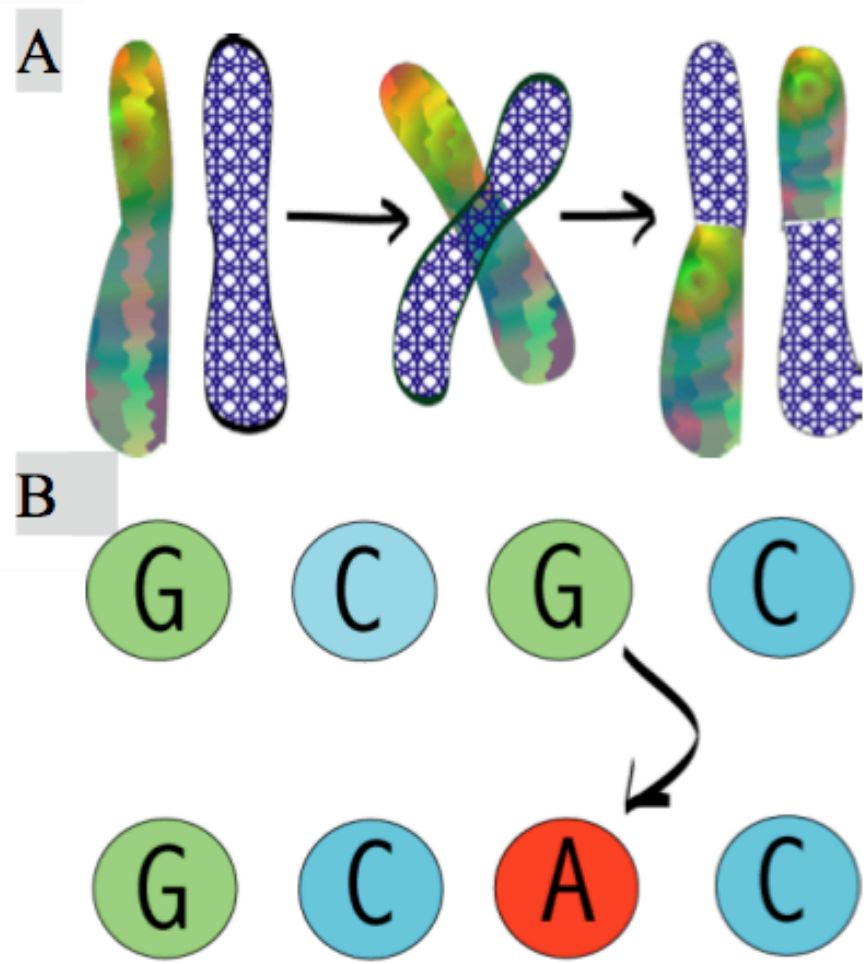
Clustered m-sequences



Additional Complexities

- The impact of low frequency drifts and correlated noise -- this will change the optimal design.
- Impact of nonlinearities in the BOLD response.
- Designs where the timing is constrained by psychology.
- In general, need to search over space of possible solutions, taking into account these practical concerns.

Genetic Algorithms



Wager and Nichols 2003

Genetic Algorithms

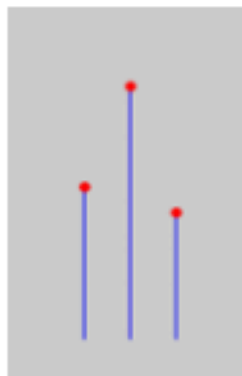
Generate stimulus lists

#1	#2	#3
1	2	1
2	3	4
4	2	3
2	1	4
1	4	3
1	2	4
3	1	2

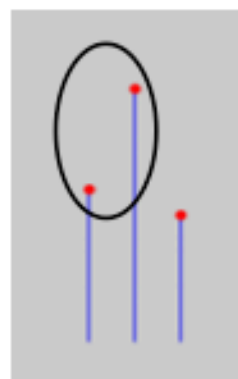
Build design matrices



Test fitness of designs



Select designs



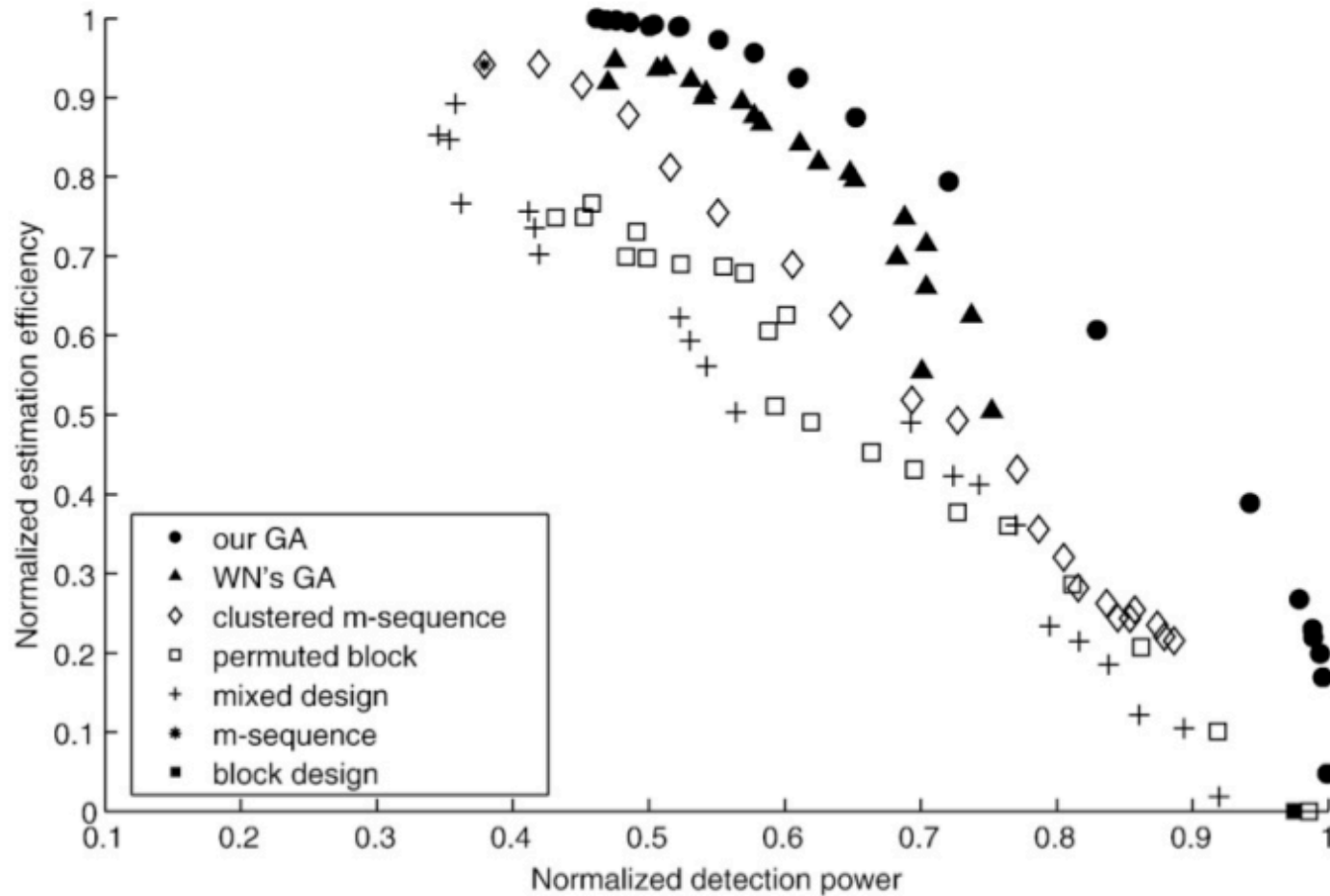
Interbreed stimulus lists

#1	#2	#3
1	2	1
4	3	2
3	2	4
4	1	2
1	4	3
1	2	4
3	1	2

iterate over generations

Wager and Nichols 2003

Genetic Algorithms



Kao et al, NIMG 2009

Genetic Algorithms

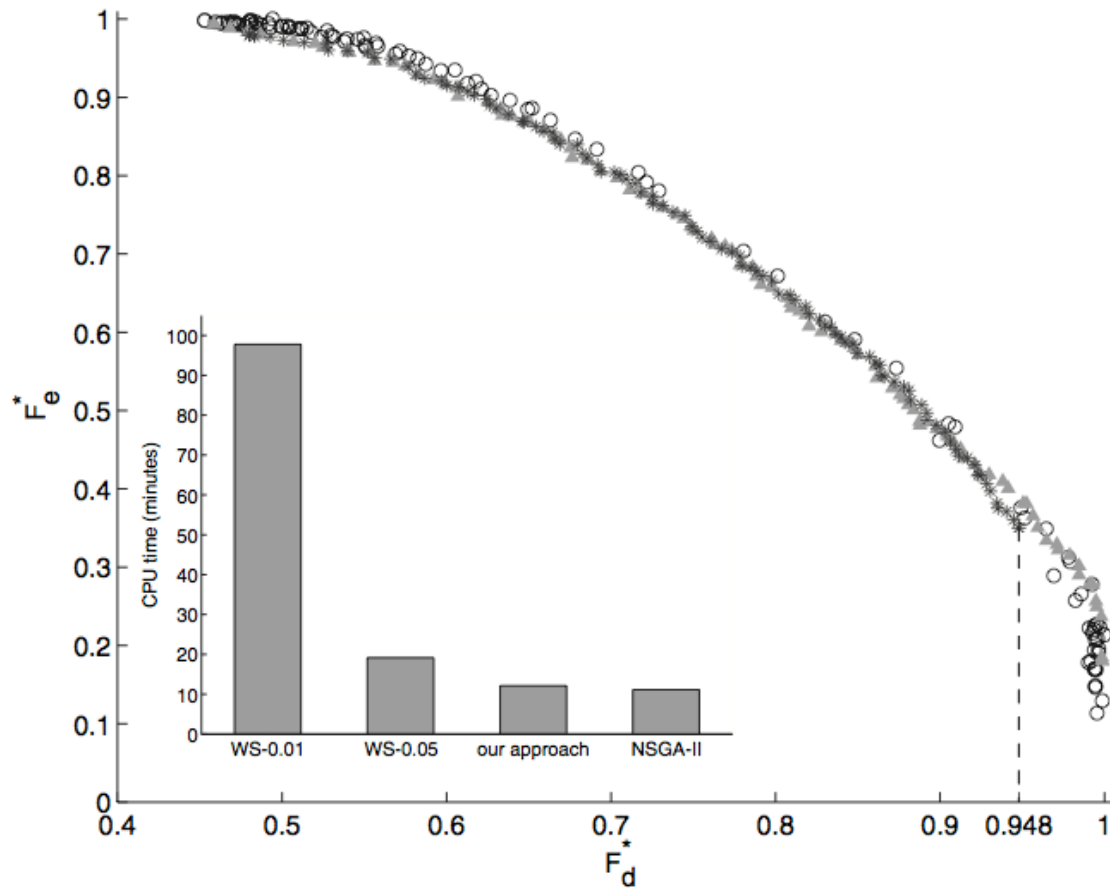
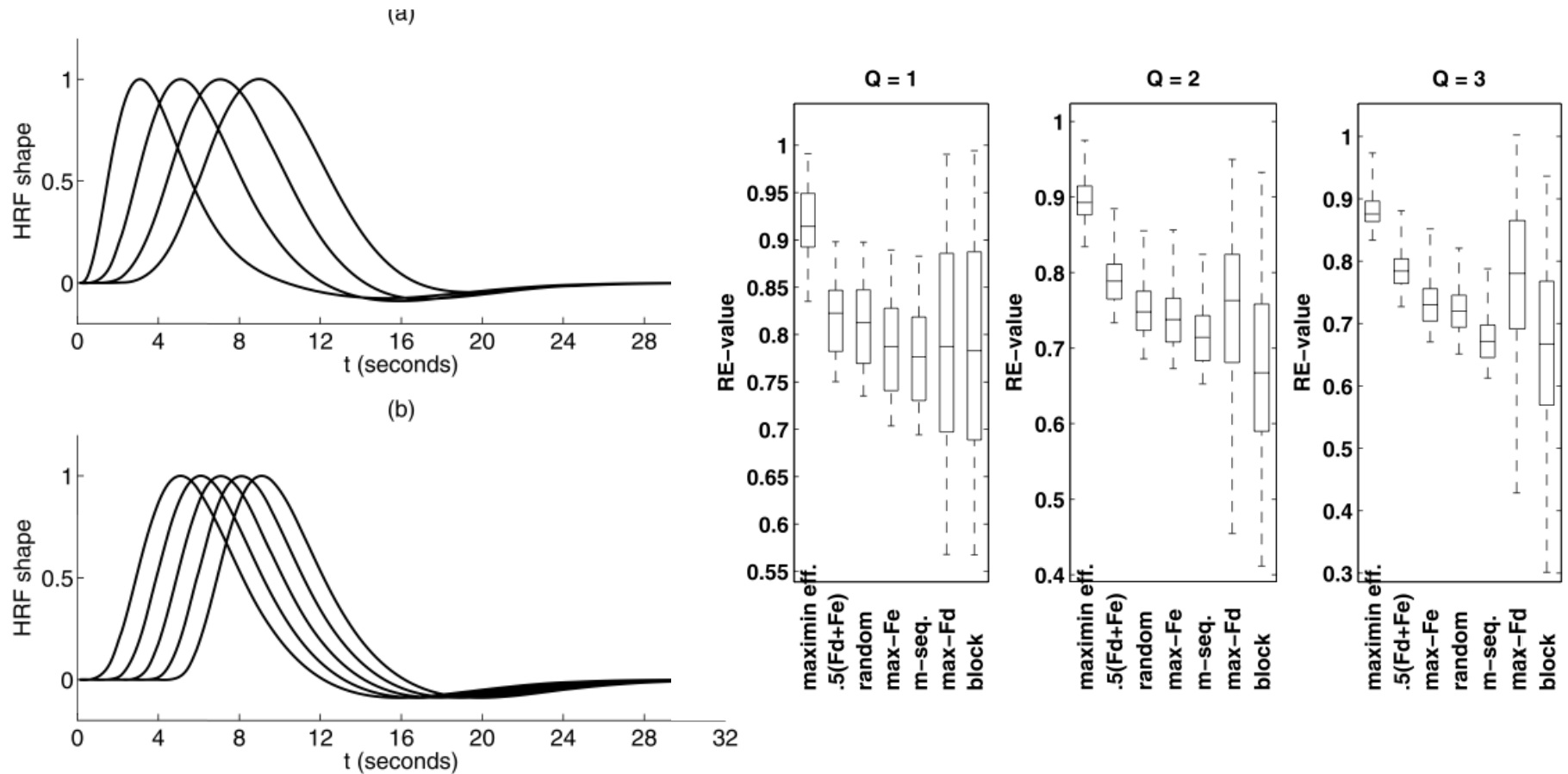


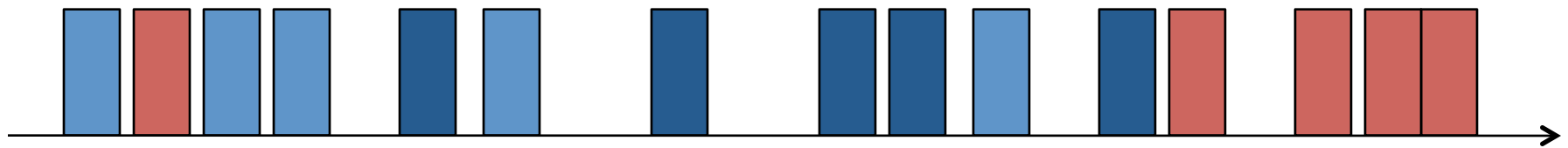
Fig. 3. F_θ^* - against F_d^* -values of the designs of the various approaches and CPU times spent for obtaining these designs (WS-0.05 and WS-0.01 represent the weighted sum methods with mesh sizes 0.05 and 0.01 respectively): \circ , weighted sum; \blacktriangle , our approach; $*$, NSGA II approach; \dagger , reference line

Robust MaxiMin Designs



Kao et al, Ann. Appl. Stat. , 2013; see also Maus et al NIMG 2010

Optimization w/ Design Constraints



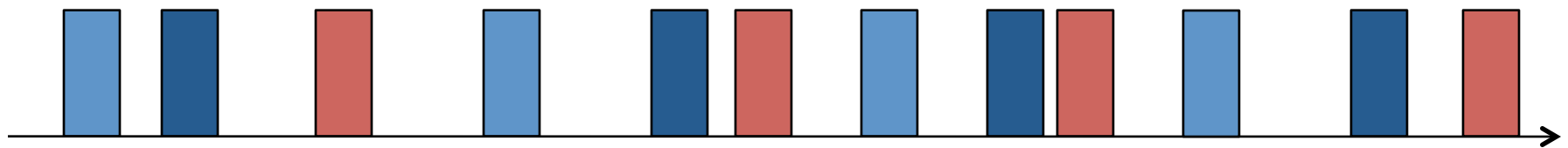
Probe



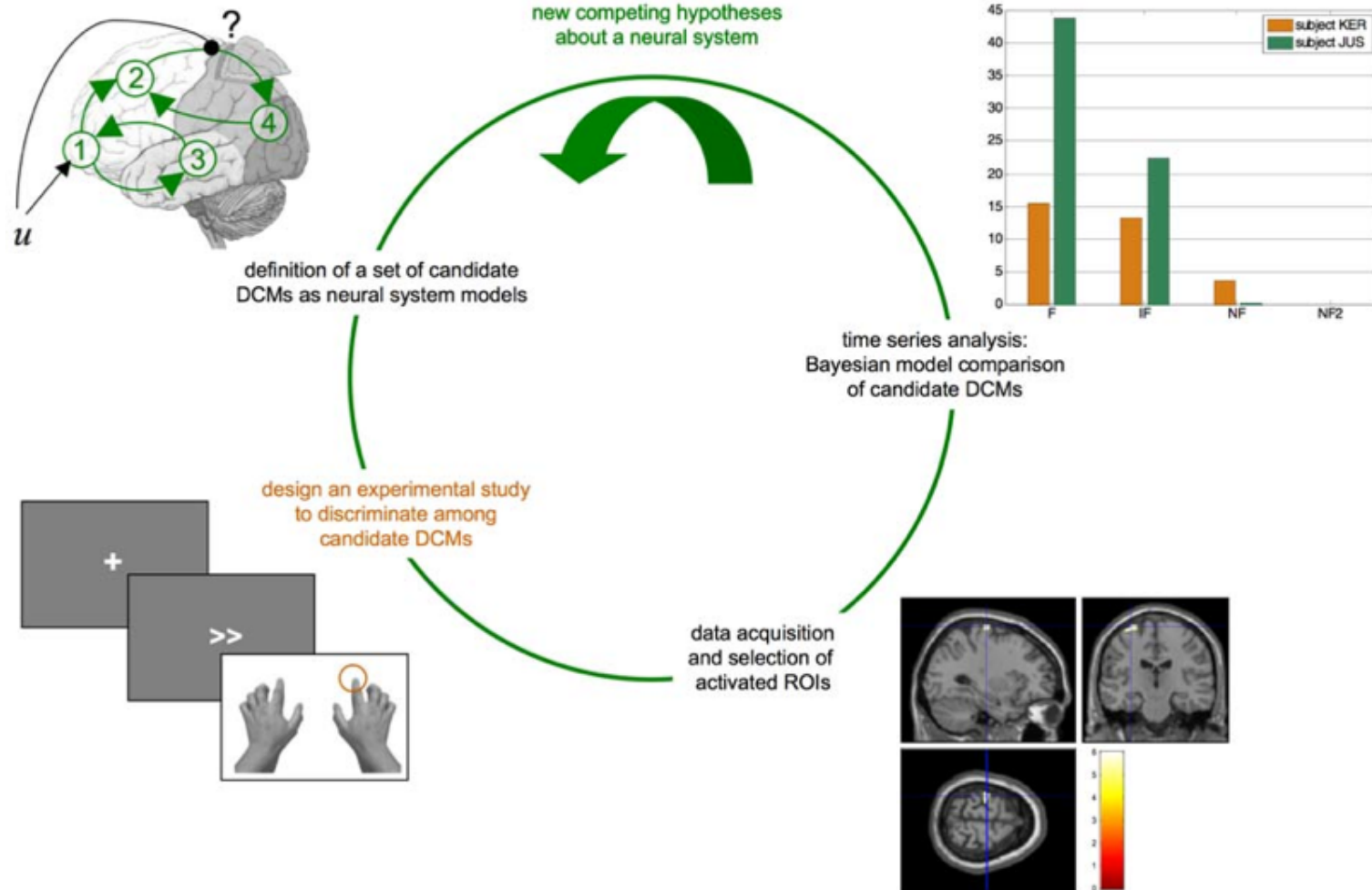
Distractor



Decide

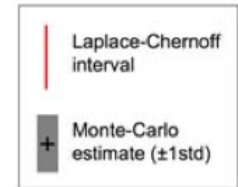
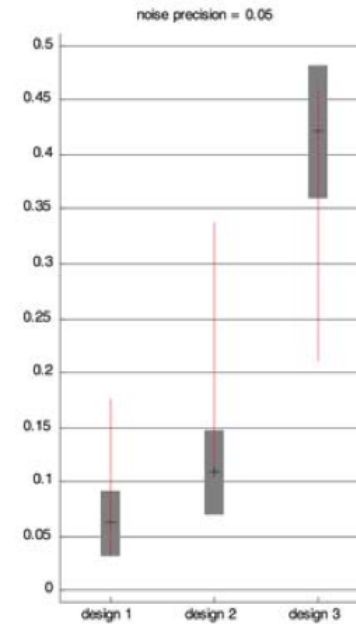
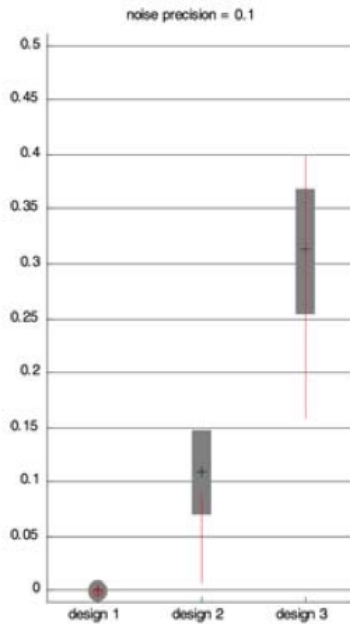
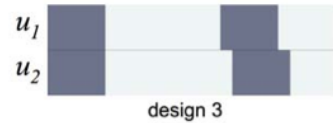
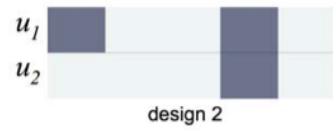
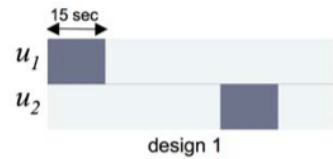
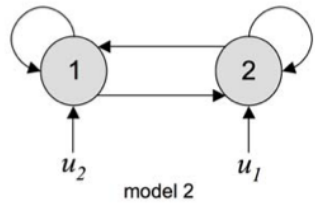
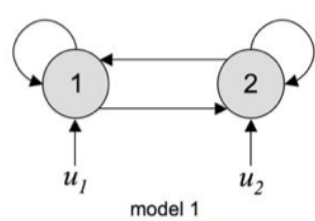


Optimal Design for DCM

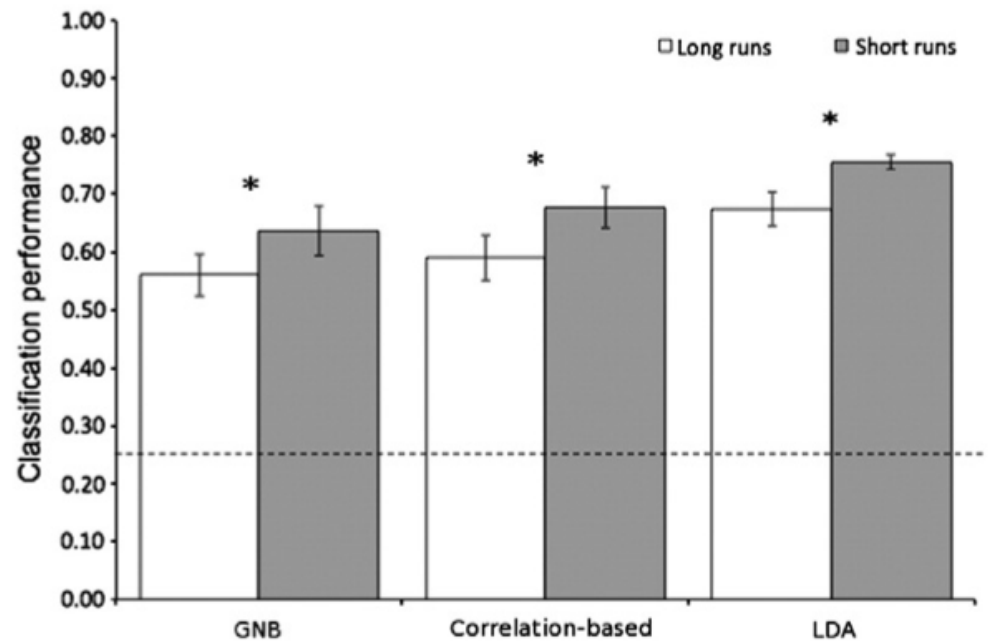
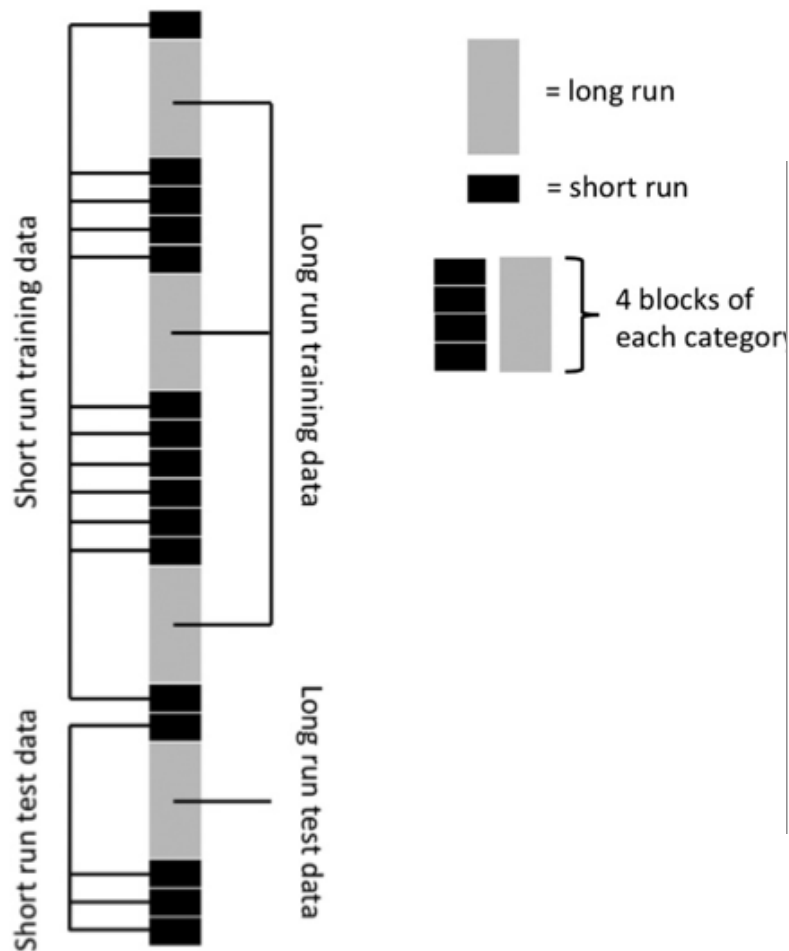


Daunizeau et al, PLOS Comp. Bio 2011

Optimal Design for DCM



Optimal Design for MVPA



Coutanche and Thompson-Schill, NIMG 2012

Software Packages

- AFNI: Rsfgen and 3dDeconvolve – random generation and evaluation of designs
- <http://surfer.nmr.mgh.harvard.edu/optseq/> -- random search over designs
- <http://www.mathworks.com/matlabcentral/fileexchange/authors/3515> -- code for generating m-sequences
- http://cfmriweb.ucsd.edu/ttliu/mttfmri_toolbox.html -- code for clustered m-sequences and other designs
- <http://www.nitrc.org/projects/pobe/> -- optimal designs of multiple-subject block design experiments
- **Genetic Algorithms:**
<http://www.columbia.edu/cu/psychology/tor/software.htm> AND
<http://www.jstatsoft.org/v30/i11/>

Summary

- The “optimal” design depends on both experimental design and assumptions about the hemodynamic response and other factors.
- Theoretical framework provides insight into the fundamental tradeoffs.
- Use search algorithms (such as GA) to find optimal designs under varying assumptions.
- Open questions related to optimization with design constraints.
- Optimization for advanced and emerging analysis methods.