# The Geometry of fMRI Statistics: Models, Efficiency, and Design

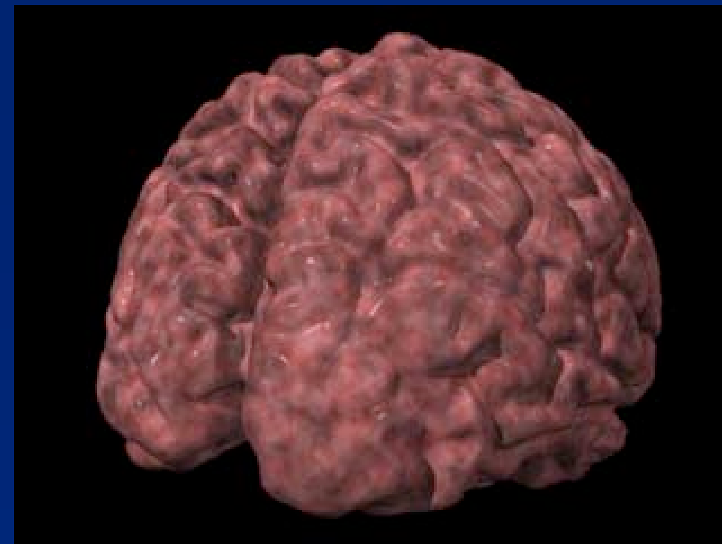
Thomas Liu
UCSD Center for Functional MRI

Mathematics in Brain Imaging Course
July 22, 2004



University of California San Diego

*Conjecture: There are only three things you need for an fMRI experiment.*




University of California
San Diego

The manipulation of statistical formulas (or software) is no substitute for knowing what one is doing. *--Hubert M. Blalock, Jr., Social Statistics*

You should understand what the analysis software is doing -- *Bob Cox, Author of AFNI*

# Overview

Geometric view of basic statistical tests.

Efficiency and the Design of Experiments

# *Why a geometric view?*

1) *Vector space interpretation of linear algebra*
2) *"simpler, more general, more elegant" W.H. Kruskal 1961*
3) *Avoid lots of messy algebra.*

*For a historical account see: D.G. Herr, The American Statistician, 34:1 1980.*

# General Linear Model

Data

Design Matrix

Nuisance Matrix

Additive Gaussian Noise

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

Hemodynamic Response

Nuisance Parameters



University of California San Diego

# Example 1

$$\mathbf{y} = \begin{bmatrix} 0 \\ 0 \\ 0.5 \\ 1 \\ 1 \\ 1 \\ 0.5 \\ 0 \\ 0 \end{bmatrix} h_1 + \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \\ 1 & 6 \\ 1 & 7 \\ 1 & 8 \\ 1 & 9 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} + \begin{bmatrix} -1 \\ -2 \\ 0 \\ 3 \\ -1 \\ 1 \\ 2 \\ .5 \\ -.2 \end{bmatrix}$$

# Example 2

$$
\mathbf{y} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \\ 1 & 6 \\ 1 & 7 \\ 1 & 8 \\ 1 & 9 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} + \begin{bmatrix} -1 \\ -2 \\ 0 \\ 3 \\ -1 \\ 1 \\ 2 \\ .5 \\ -.2 \end{bmatrix}
$$



University of California San Diego

# Simplest Case

$$\mathbf{y} = \mathbf{x}h_1 + \mathbf{s}b_1 + n$$

$$= \begin{bmatrix} 0 \\ 0 \\ 0.5 \\ 1 \\ 1 \\ 1 \\ 0.5 \\ 0 \\ 0 \end{bmatrix} h_1 + \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} s_1 + \begin{bmatrix} -1 \\ -2 \\ 0 \\ 3 \\ -1 \\ 1 \\ 2 \\ .5 \\ -.2 \end{bmatrix}$$

Correlation Coefficient

$$r = \frac{(\mathbf{y}-\bar{\mathbf{y}})^T(\mathbf{x}-\bar{\mathbf{x}})}{\sqrt{(\mathbf{y}-\bar{\mathbf{y}})^T(\mathbf{y}-\bar{\mathbf{y}})}\sqrt{(\mathbf{x}-\bar{\mathbf{x}})^T(\mathbf{x}-\bar{\mathbf{x}})}}$$



University of California San Diego

Plan of attack

1. First derive statistics assuming there are no nuisance functions
2. Then add in nuisance functions.

# General Linear Model

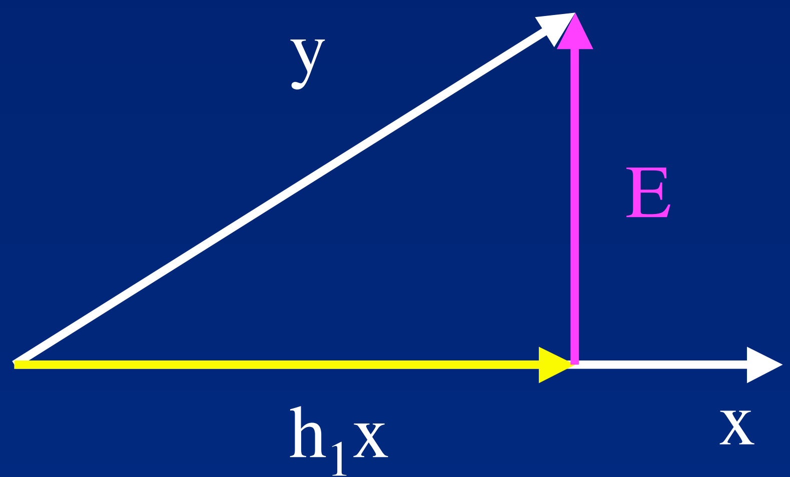
Data → $\mathbf{y}$

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

Design Matrix → $\mathbf{X}$

Hemodynamic Response → $\mathbf{h}$

# Principle of Orthogonality



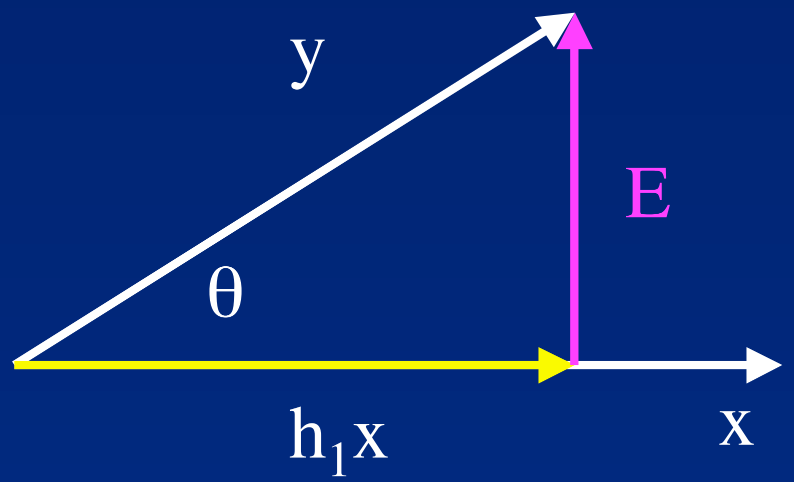

$$\mathrm{E}^{\mathrm{T}}\mathbf{x} = 0$$

$$(\mathbf{y} - h_1\mathbf{x})^T\mathbf{x} = 0$$

$$h_1 = \frac{\mathbf{y}^T\mathbf{x}}{\mathbf{x}^T\mathbf{x}}$$

Minimum error vector is orthogonal to the model space.

# Correlation Coefficient

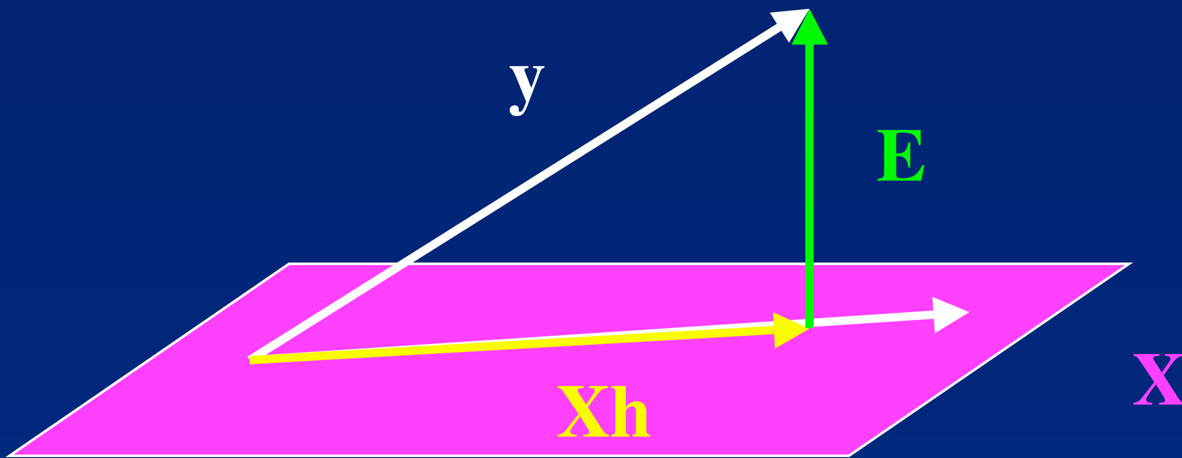



$$
\begin{aligned}
\mathrm{r} &= \cos\theta \\
&= \frac{\left\| h_1 \mathbf{x} \right\|}{\left\| \mathbf{y} \right\|} \\
&= \frac{\left\| \frac{\mathbf{y}^T \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \mathbf{x} \right\|}{\left\| \mathbf{y} \right\|} \\
&= \frac{\mathbf{y}^T \mathbf{x}}{\left\| \mathbf{y} \right\| \left\| \mathbf{x} \right\|}
\end{aligned}
$$

From a previous slide …

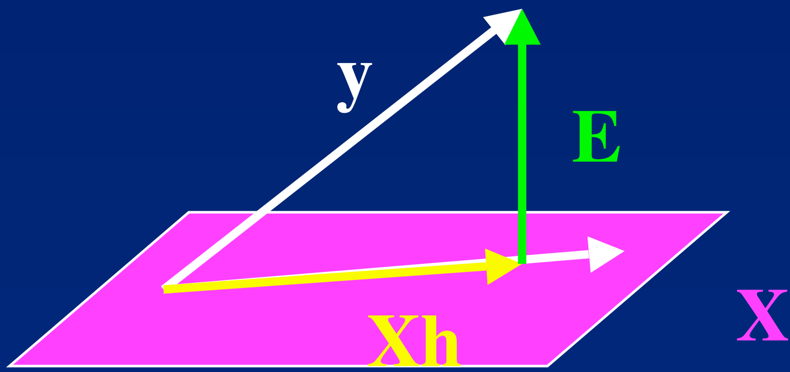
$$
r = \frac{(\mathbf{y} - \overline{\mathbf{y}})^T (\mathbf{x} - \overline{\mathbf{x}})}{\sqrt{(\mathbf{y} - \overline{\mathbf{y}})^T (\mathbf{y} - \overline{\mathbf{y}})} \sqrt{(\mathbf{x} - \overline{\mathbf{x}})^T (\mathbf{x} - \overline{\mathbf{x}})}}
$$

# Principle of Orthogonality




$$\mathbf{X}^T\left(\mathbf{y}-\mathbf{X}\mathbf{h}\right)=\mathbf{0}\Rightarrow\mathbf{h}=\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{y}$$

# Projection Matrices




$$
\begin{aligned}
\mathbf{E} &= \mathbf{y} - \mathbf{P_X y} \\
&= \left(\mathbf{I} - \mathbf{P_X}\right)\mathbf{y} \\
&= \mathbf{P_X^{\perp} y}
\end{aligned}
$$

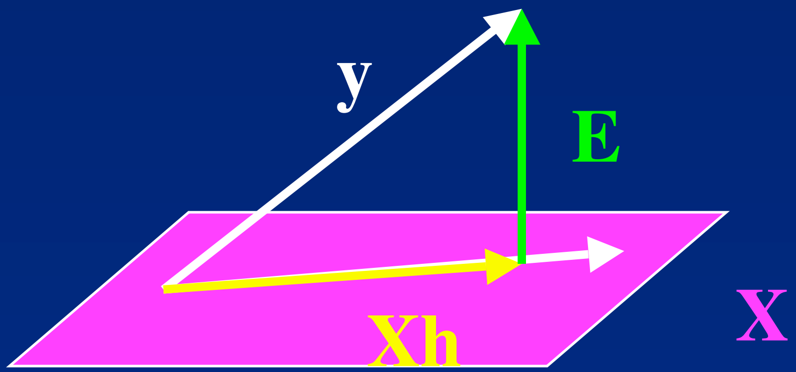
$$
\begin{aligned}
\mathbf{h} &= \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{y} \\
\mathbf{Xh} &= \mathbf{X}\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{y} \\
&= \mathbf{P_X y}
\end{aligned}
$$

Useful Facts

$$
\begin{aligned}
\mathbf{P_X P_X} &= \mathbf{P_X} \\
\mathbf{P_X^T} &= \mathbf{P_X}
\end{aligned}
$$

University of California San Diego

# Orthogonality again

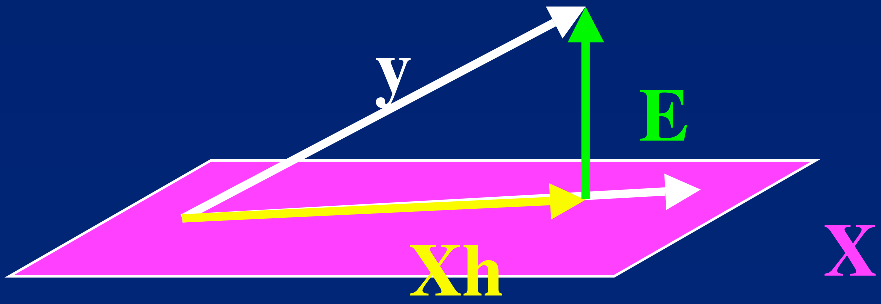



$$\mathbf{h}^T\mathbf{X}^T\mathbf{E} = 0$$

$$\mathbf{y}^T\mathbf{P}_X\mathbf{P}_X^{\perp}\mathbf{y} = 0$$

$$\mathbf{y}^T\mathbf{P}_X\left(\mathbf{I}-\mathbf{P}_X\right)\mathbf{y} = 0$$

# F-statistic



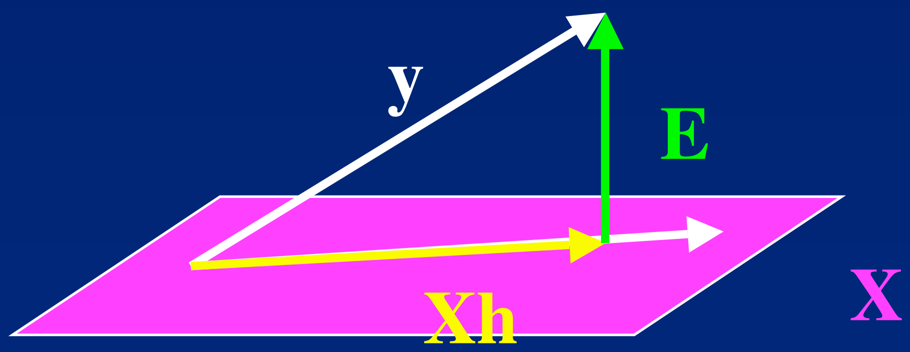

$$F = \frac{\left\|\mathbf{Xh}\right\|^2 \big/ (\#\text{ of model functions})}{\left\|\mathbf{E}\right\|^2 \big/ (\#\text{ of datapoints } - \#\text{ of model functions})}$$

$$= \frac{N-k}{k}\frac{\mathbf{y}^T\mathbf{P}_\mathbf{X}^T\mathbf{P}_\mathbf{X}\mathbf{y}}{\mathbf{y}^T(\mathbf{I}-\mathbf{P}_\mathbf{X})^T(\mathbf{I}-\mathbf{P}_\mathbf{X})\mathbf{y}}$$

$$= \frac{N-k}{k}\frac{\mathbf{y}^T\mathbf{P}_\mathbf{X}\mathbf{y}}{\mathbf{y}^T(\mathbf{I}-\mathbf{P}_\mathbf{X})\mathbf{y}}$$

$$= \frac{N-k}{k}\cot^2\theta$$

# Coefficient of Determination




$$R^2 = \frac{\|\mathbf{Xh}\|^2}{\|\mathbf{y}\|^2} = \frac{\mathbf{y}^T\mathbf{P_X}\mathbf{y}}{\mathbf{y}^T\mathbf{y}} = \cos^2\theta$$

Easy to show that

$$F = \frac{N-k}{k}\frac{R^2}{1-R^2}$$

# General Linear Model

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

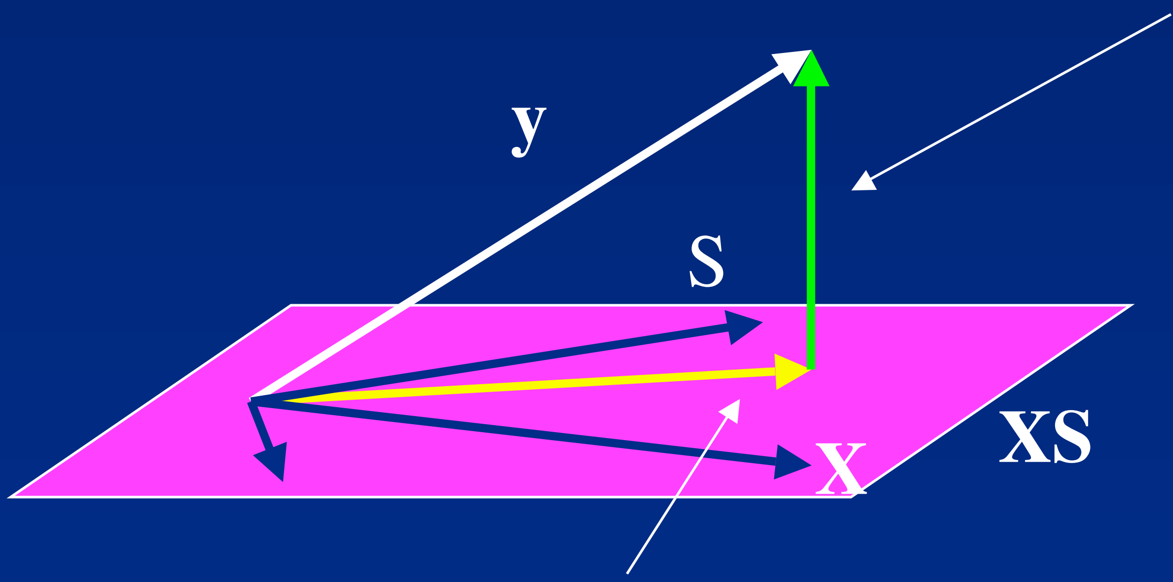
- Data → $\mathbf{y}$
- Design Matrix → $\mathbf{X}$
- Hemodynamic Response → $\mathbf{h}$
- Nuisance Matrix → $\mathbf{S}$
- Nuisance Parameters → $\mathbf{b}$
- Additive Gaussian Noise → $\mathbf{n}$



University of California San Diego

# Nuisance Functions

$\mathbf{E} = \left(\mathbf{I} - \mathbf{P}_{XS}\right)\mathbf{y}$ is the residual error

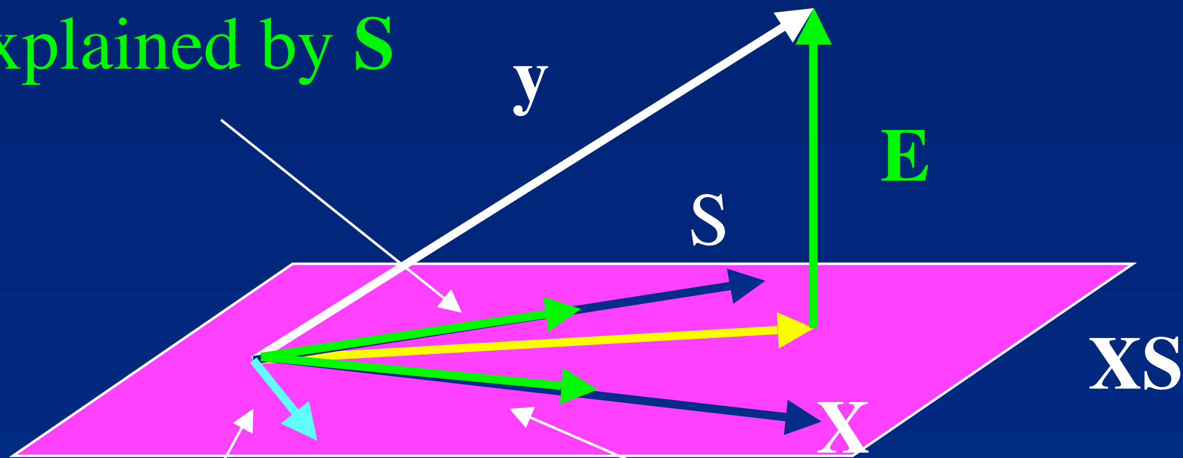



$\mathbf{P}_{XS}\mathbf{y}$ is the data explained by both **X** and **S**



University of California San Diego

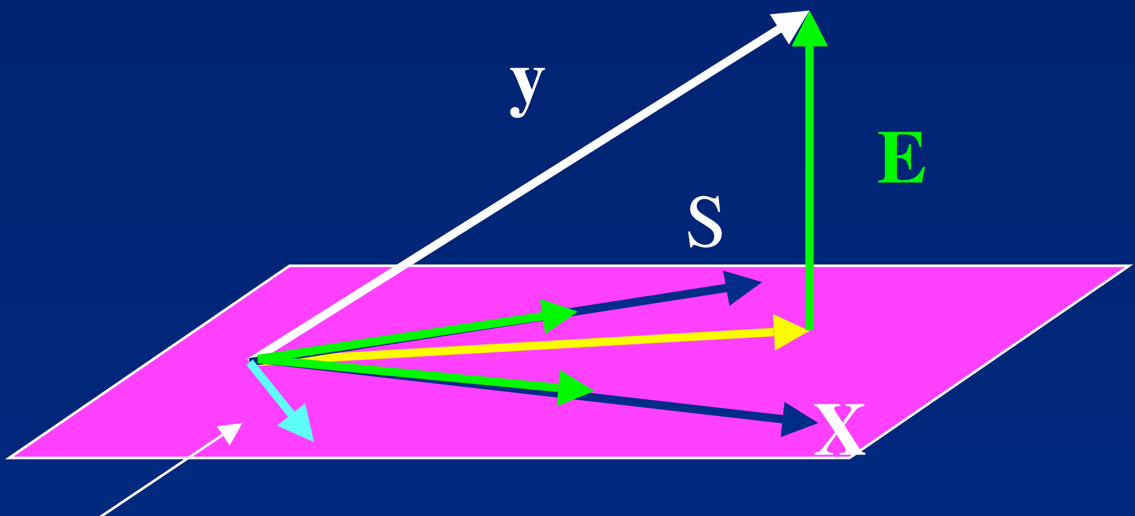
# Nuisance Functions




Data explained by **X** that can't be explained by **S**

Data explained by **X**

# Nuisance Functions

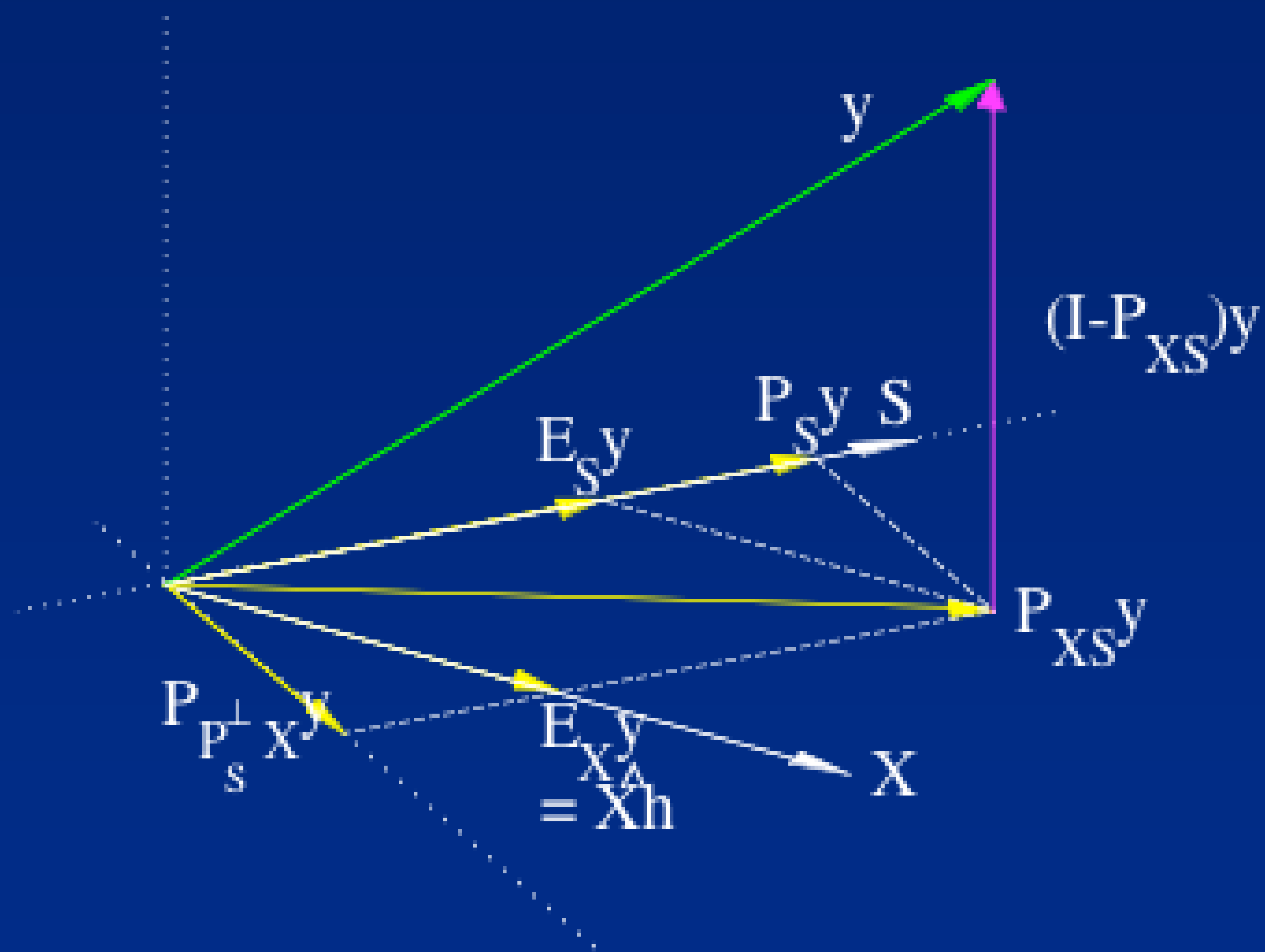



$$\mathbf{P}_{\mathbf{S}}^{\perp} = \mathbf{I} - \mathbf{P}_{\mathbf{S}}$$

The space spanned by the columns of $\mathbf{P}_{\mathbf{S}}^{\perp}\mathbf{X}$ is the part of the model space that is orthogonal to **S**.

$\mathbf{P}_{\mathbf{P}_{\mathbf{S}}^{\perp}\mathbf{X}}\mathbf{y}$ is the projection of the data onto that space, and is therefore the data explained by the model that can't be explained by **S**.

# Geometric Picture



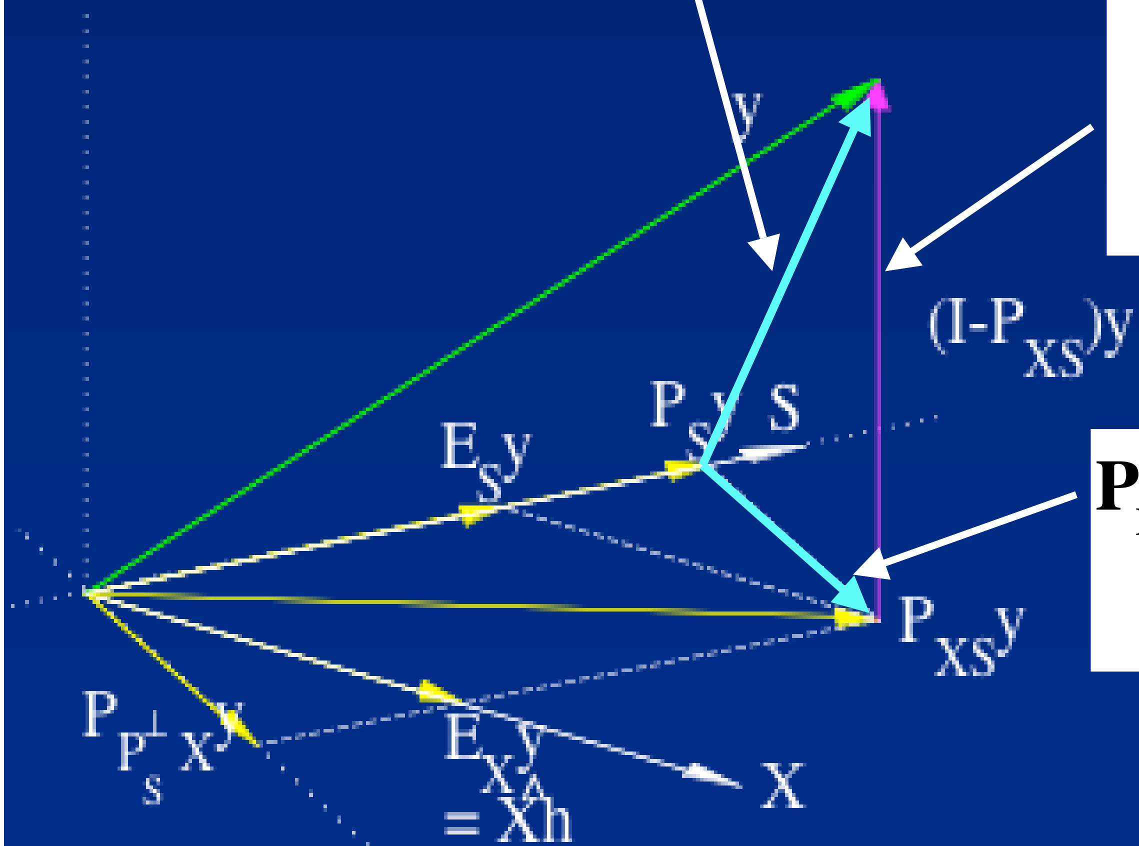

University of California San Diego

# Pythagorean Relation

$\mathbf{P}_S^{\perp}\mathbf{y} =$ data not explained by **S**

$\left(\mathbf{I} - \mathbf{P}_{XS}\right)\mathbf{y} =$ data explained by neither **X** nor **S**

$\mathbf{P}_{\mathbf{P}_S^{\perp}\mathbf{X}}\mathbf{y} =$ data explained by **X** but not by **S**

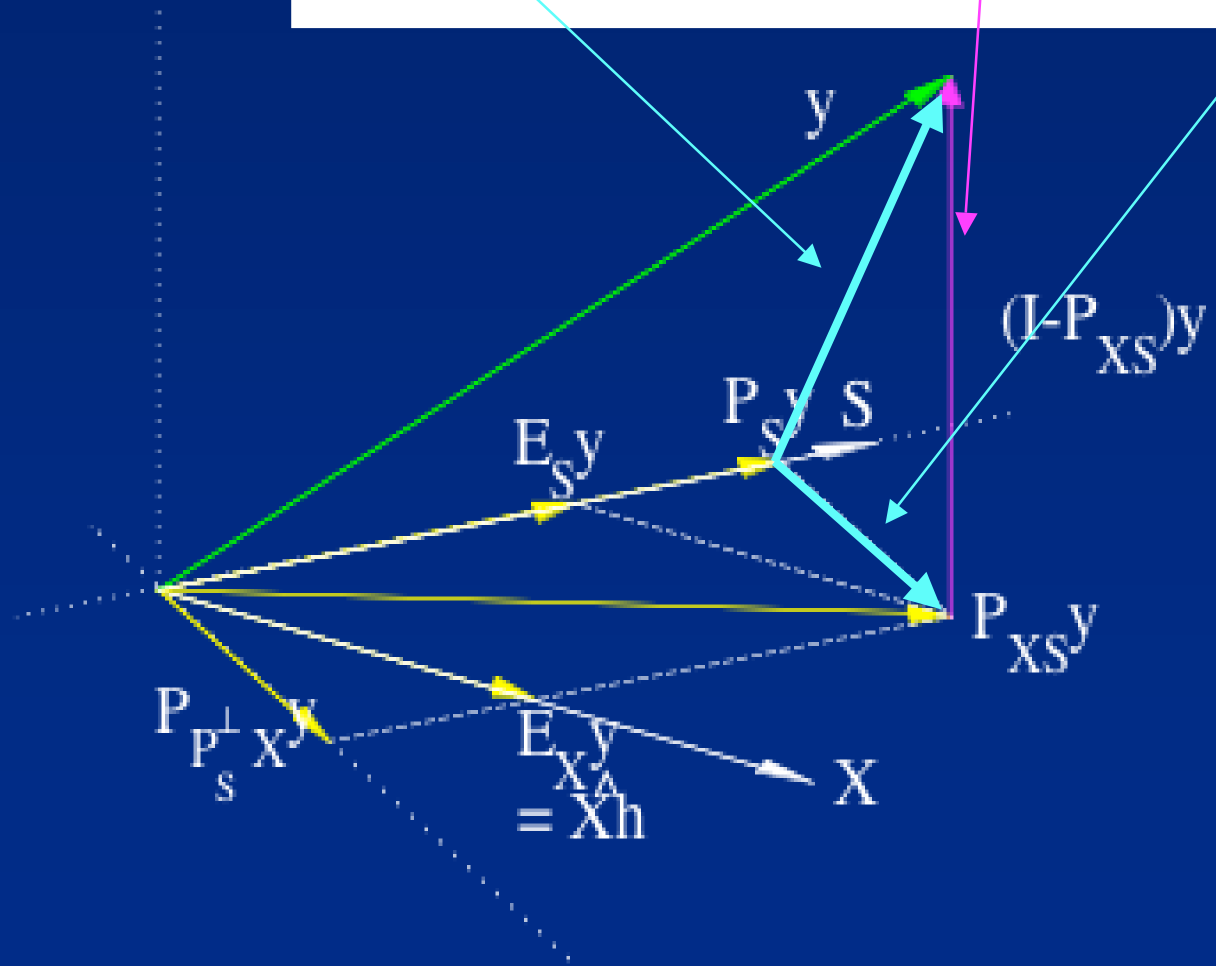





University of California San Diego

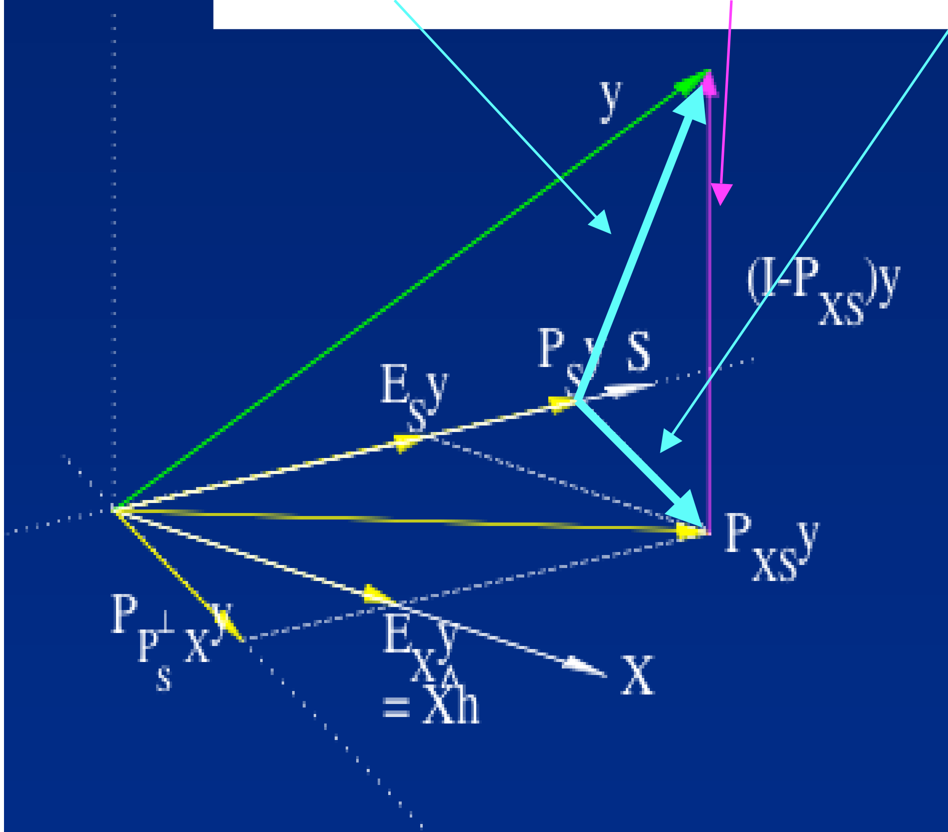
# Pythagorean Relation

$$\mathbf{y}^T\mathbf{P}_\mathbf{S}^{\perp}\mathbf{y} = \mathbf{y}^T\left(\mathbf{I}-\mathbf{P}_{\mathbf{XS}}\right)\mathbf{y} + \mathbf{y}^T\mathbf{P}_{\mathbf{P}_\mathbf{S}^{\perp}\mathbf{X}}\mathbf{y}$$



# F-statistic

$$\mathbf{y}^T\mathbf{P}_\mathbf{S}^\perp\mathbf{y} = \mathbf{y}^T\left(\mathbf{I}-\mathbf{P}_{\mathbf{XS}}\right)\mathbf{y} + \mathbf{y}^T\mathbf{P}_{\mathbf{P}_\mathbf{S}^\perp\mathbf{X}}\mathbf{y}$$



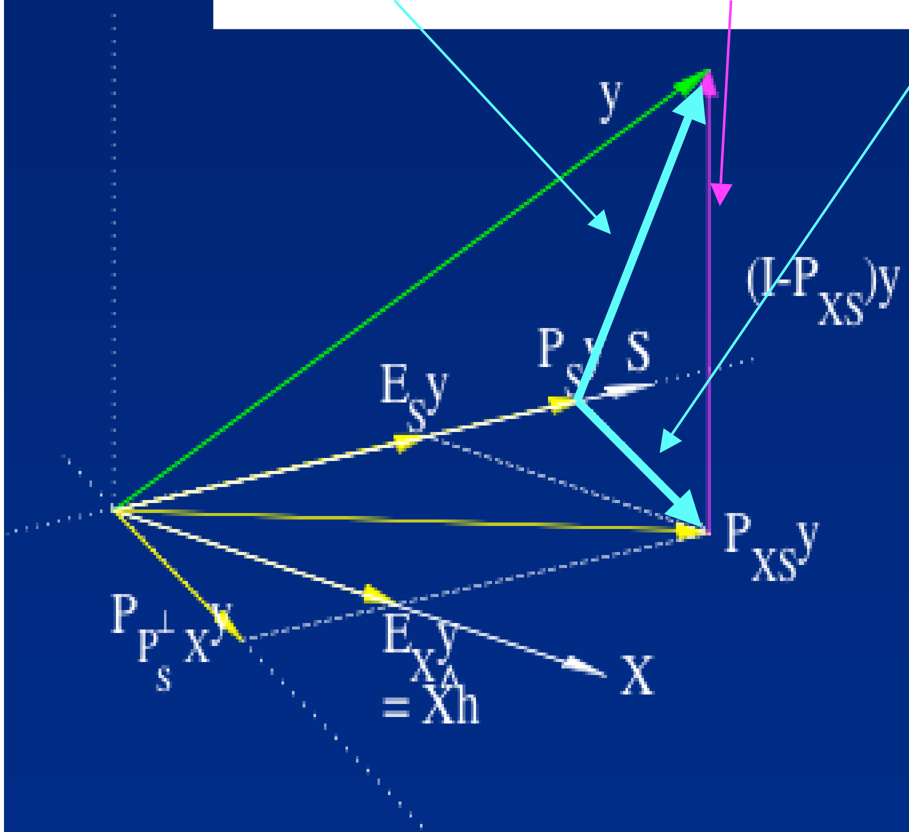

$$F = \frac{N-k-l}{k}\frac{\mathbf{y}^T\mathbf{P}_{\mathbf{P}_\mathbf{s}^\perp\mathbf{X}}\mathbf{y}}{\mathbf{y}^T\left(\mathbf{I}-\mathbf{P}_{\mathbf{XS}}\right)\mathbf{y}}$$

$$= \frac{N-k-l}{k}\cot^2\theta$$

$k$ = # model functions, $l$ = # of nuisance functions

# Coeffcient of Determination $R^2$

$$\mathbf{y}^T\mathbf{P}_S^{\perp}\mathbf{y} = \mathbf{y}^T\left(\mathbf{I}-\mathbf{P}_{XS}\right)\mathbf{y} + \mathbf{y}^T\mathbf{P}_{\mathbf{P}_S^{\perp}X}\mathbf{y}$$




$$R^2 = \frac{\mathbf{y}^T\mathbf{P}_{\mathbf{P}_s^{\perp}\mathbf{x}}\mathbf{y}}{\mathbf{y}^T\mathbf{P}_S^{\perp}\mathbf{y}} = \cos^2(\theta)$$

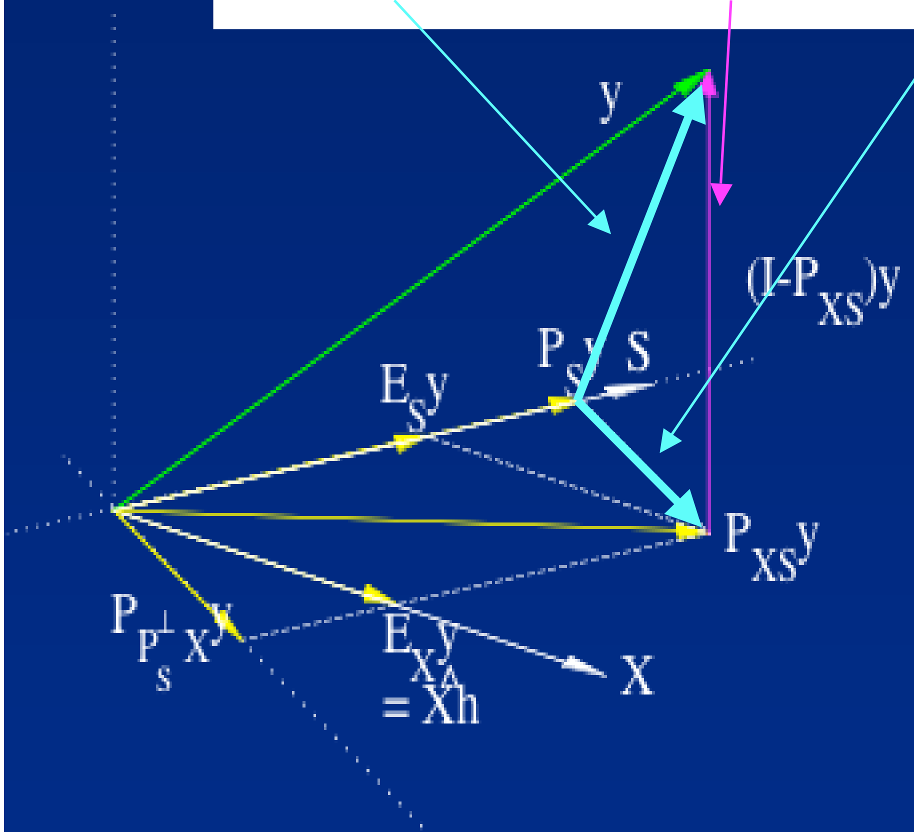
$k$ = # model functions, $l$ = # of nuisance functions



University of California San Diego

# F and $R^2$

$$\mathbf{y}^T\mathbf{P}_S^{\perp}\mathbf{y} = \mathbf{y}^T\left(\mathbf{I}-\mathbf{P}_{XS}\right)\mathbf{y} + \mathbf{y}^T\mathbf{P}_{\mathbf{P}_S^{\perp}X}\mathbf{y}$$




$$F = \frac{N-k-l}{k}\,\frac{\mathbf{y}^T\mathbf{P}_{\mathbf{P}_s^{\perp}X}\mathbf{y}}{\mathbf{y}^T\left(\mathbf{I}-\mathbf{P}_{XS}\right)\mathbf{y}} = \frac{N-k-l}{k}\cot^2\theta$$

$$R^2 = \frac{\mathbf{y}^T\mathbf{P}_{\mathbf{P}_s^{\perp}X}\mathbf{y}}{\mathbf{y}^T\mathbf{P}_S^{\perp}\mathbf{y}} = \cos^2\theta$$

$$F = \frac{N-k-l}{l}\,\frac{R^2}{1-R^2}$$

$k$ = # model functions, $l$ = # of nuisance functions

# Multiple Correlation Coefficient

After a bit of algebra, we can show that…

$$R = \frac{(\mathbf{y}-\mathbf{P}_S\mathbf{y})^T(\hat{\mathbf{y}}-\mathbf{P}_S\mathbf{X}\hat{\mathbf{h}})}{\sqrt{(\mathbf{y}-\mathbf{P}_S\mathbf{y})^T(\mathbf{y}-\mathbf{P}_S\mathbf{y})}\sqrt{(\hat{\mathbf{y}}-\mathbf{P}_S\mathbf{X}\hat{\mathbf{h}})^T(\hat{\mathbf{y}}-\mathbf{P}_S\mathbf{X}\hat{\mathbf{h}})}}$$

For 1 model function and 1 constant nuisance function, this reduces to the familiar

$$R = \frac{(\mathbf{y}-\overline{\mathbf{y}})^T(\mathbf{x}-\overline{\mathbf{x}})}{\sqrt{(\mathbf{y}-\overline{\mathbf{y}})^T(\mathbf{y}-\overline{\mathbf{y}})}\sqrt{(\mathbf{x}-\overline{\mathbf{x}})^T(\mathbf{x}-\overline{\mathbf{x}})}}$$

# Application

In the analysis of most fMRI experiments, we need to properly deal with nuisance terms, such as low frequency drifts. A reasonable approach is to project out the nuisance terms and then correlate the detrended data with a reference function. Does this give us the correct correlation coefficient?

# Multiple Correlation Coefficient

For 1 model function and multiple nuisance function, we obtain

$$R = \frac{(\mathbf{y} - \mathbf{P}_S\mathbf{y})^T(\mathbf{x} - \mathbf{P}_S\mathbf{x})}{\sqrt{(\mathbf{y} - \mathbf{P}_S\mathbf{y})^T(\mathbf{y} - \mathbf{P}_S\mathbf{y})}\sqrt{(\mathbf{x} - \mathbf{P}_S\mathbf{x})^T(\mathbf{x} - \mathbf{P}_S\mathbf{x})}}$$

So, this tells us that we should detrend both the data and the reference function before correlating.

# One more thing

When there is only one model function ($k=1$) and $l$ nuisance functions, the F-statistic is simply the squared of the t-statistic with $N$-$l$-$1$ degrees of freedom.
So, we also have the useful relation

$$t = \sqrt{N - l - 1}\,\frac{R}{\sqrt{1 - R^2}}$$

# Efficiency and Design

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



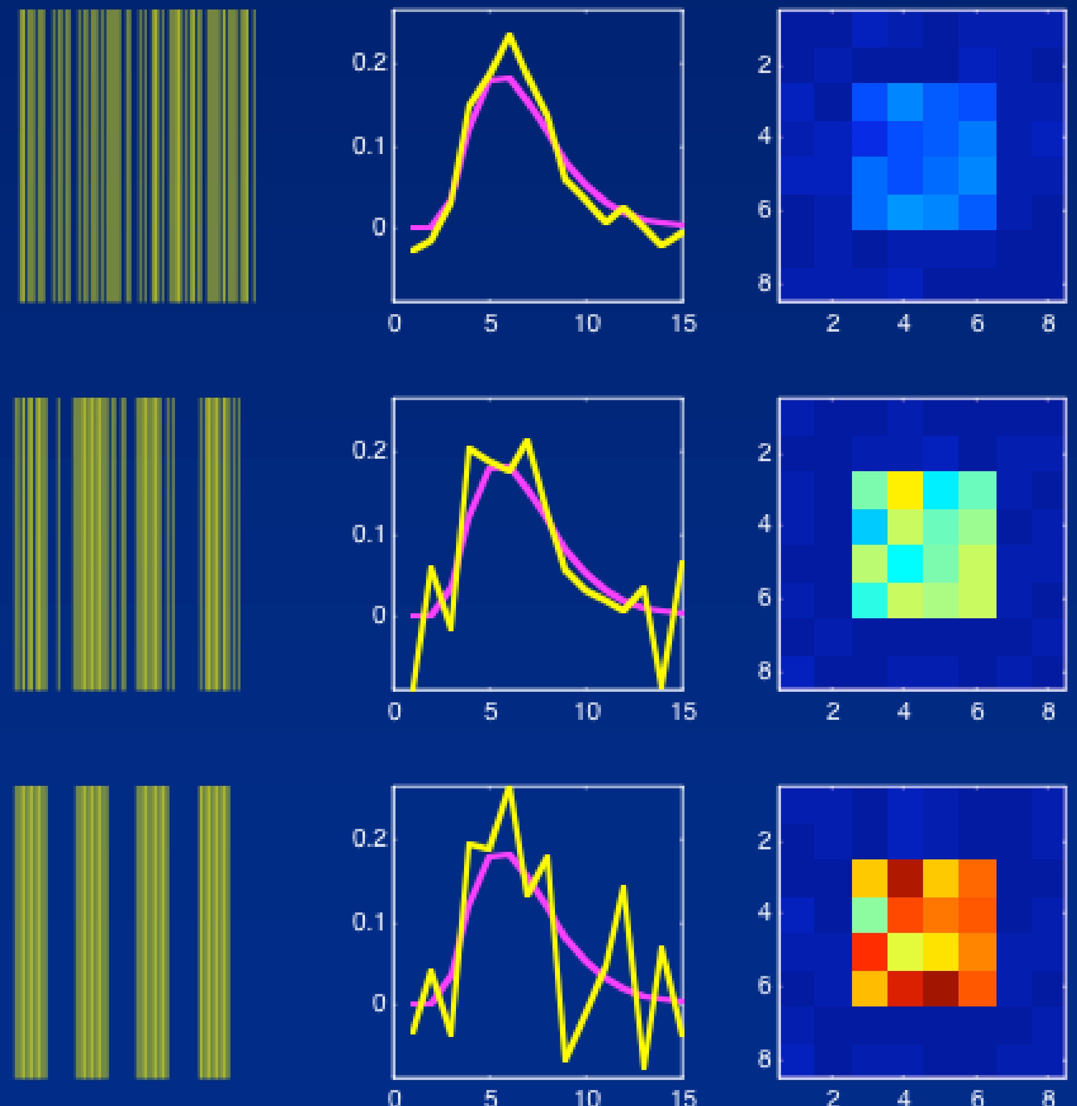
University of California San Diego

# Power, Efficiency, Predictability

Random
P = 0.5

SemiRandom
P = 0.63

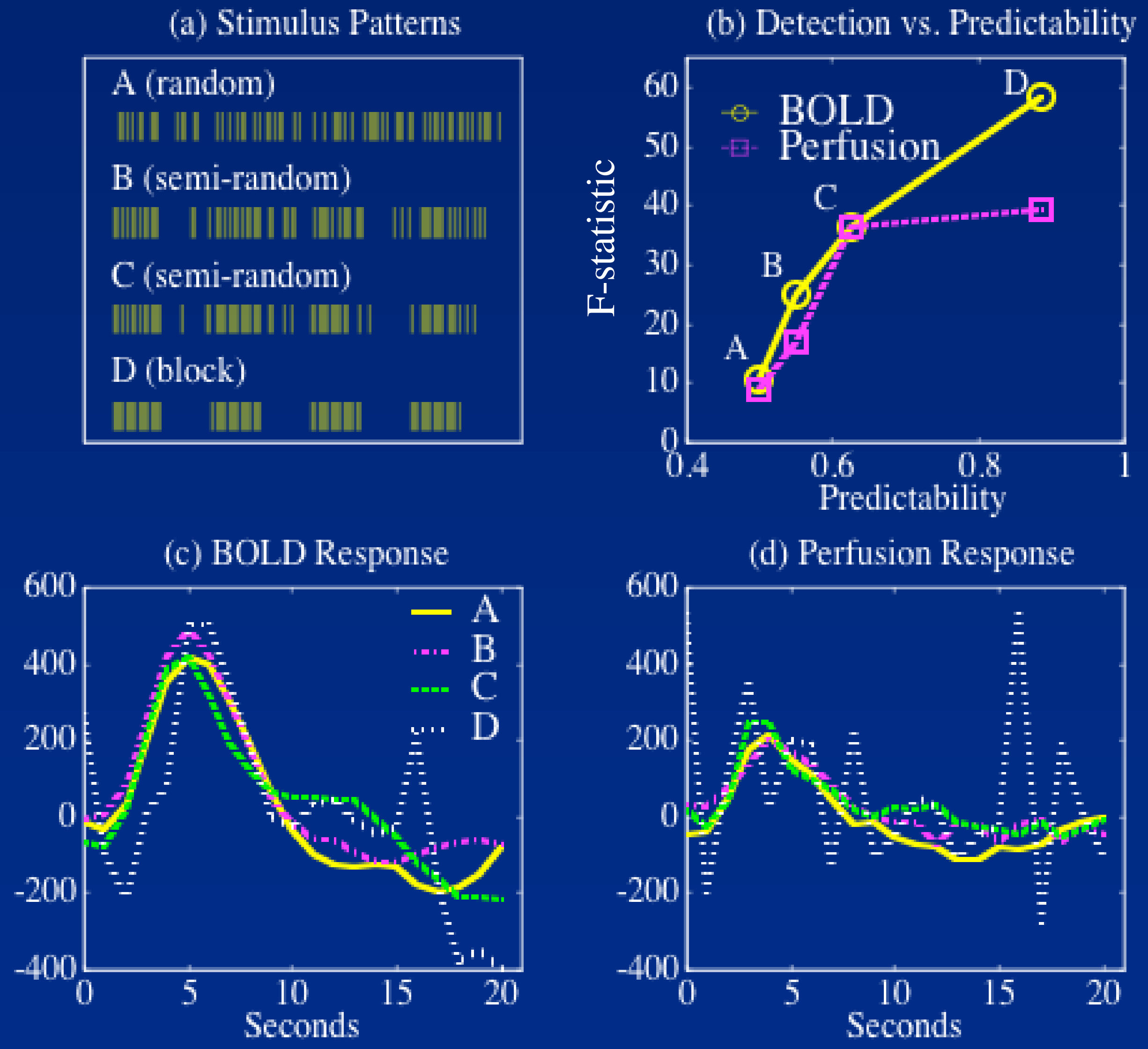
Block
P =0.9






University of California
San Diego

# Experimental Data






University of California San Diego

# General Linear Model

| Data | | Design Matrix | | Nuisance Matrix | | Additive Gaussian Noise |
| --- | --- | --- | --- | --- | --- | --- |
| $\mathbf{y}$ | $=$ | $\mathbf{Xh}$ | $+$ | $\mathbf{Sb}$ | $+$ | $\mathbf{n}$ |
| | | Hemodynamic Response | | Nuisance Parameters | | |



University of California San Diego

# Statistical Efficiency

Least Square Estimate (for now assume white noise)

$$\hat{\mathbf{h}} = \left(\mathbf{X}^T \mathbf{P}_{\mathbf{S}}^{\perp} \mathbf{X}\right)^{-1} \mathbf{X}^T \mathbf{P}_{\mathbf{S}}^{\perp} \mathbf{y}$$

$$= \left(\mathbf{X}_{\perp}^T \mathbf{X}_{\perp}\right)^{-1} \mathbf{X}_{\perp}^T \mathbf{y}$$

$$\xi = Efficiency \propto \frac{1}{\text{variance of } \hat{\mathbf{h}}}$$

# Statistical Efficiency

Efficiency depends on:

1) Model Assumptions
2) Experimental Design

# Model Assumptions

How much do we want to assume about the shape of the hemodynamic response (HDR)?

1) Assume we know nothing about its shape
2) Assume we know its shape completely, but not its amplitude.
3) Assume we know something about its shape.

# Assumptions and Design

Assumption 1: Experiments where you want to characterize in detail the shape of the HDR.

Assumption 2: Experiments where you have a good guess as to the shape (either a canonical form or measured HDR) and want to detect activation.

Assumption 3: A reasonable compromise between 1 and 2. Detect activation when you sort of know the shape. Characterize the shape when you sort of know its properties.

# Assumption 1

If we assume nothing about the shape (except for length) then the GLM is what we had before: $\mathbf{y} = \mathbf{X}\mathbf{h} + \mathbf{S}\mathbf{b} + \mathbf{n}$

$$\text{Covariance}: \ \mathbf{C}_{\hat{\mathbf{h}}} = \sigma^2\left(\mathbf{X}_\perp^T\mathbf{X}_\perp\right)^{-1}$$

$$Efficiency: \xi \propto \frac{1}{\text{average variance}} = \frac{1}{\sigma^2 Trace\left[\left(\mathbf{X}_\perp^T\mathbf{X}_\perp\right)^{-1}\right]}$$

# Assumption 2

Assume we know the HDR shape but not the amplitude

$$\mathbf{h} = \mathbf{h}_0 c$$

$GLM:$

$$\begin{aligned}\mathbf{y} &= \mathbf{X}\mathbf{h}_0 c + \mathbf{S}\mathbf{b} + \mathbf{n} \\ &= \tilde{\mathbf{X}} c + \mathbf{S}\mathbf{b} + \mathbf{n}\end{aligned}$$

$Efficiency:$

$$\xi \propto \frac{1}{\mathrm{var}(\text{amplitude estimate})} = \frac{1}{\mathrm{var}(\hat{c})} = \frac{\mathbf{h}_0^T \mathbf{X}_\perp^T \mathbf{X}_\perp \mathbf{h}_0}{\sigma^2}$$



University of California San Diego

# Assumption 3

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

$$GLM : \mathbf{y} = \mathbf{X}\mathbf{B}\mathbf{c} + \mathbf{S}\mathbf{b} + \mathbf{n} = \tilde{\mathbf{X}}\mathbf{c} + \mathbf{S}\mathbf{b} + \mathbf{n}$$

$$Estimate : \hat{\mathbf{c}} = \left(\mathbf{B}^T \mathbf{X}_\perp^T \mathbf{X}_\perp \mathbf{B}\right)^{-1} \mathbf{B}^T \mathbf{X}_\perp^T \mathbf{y}$$

$$\hat{\mathbf{h}} = \mathbf{B}\hat{\mathbf{c}} = \mathbf{B}\left(\mathbf{B}^T \mathbf{X}_\perp^T \mathbf{X}_\perp \mathbf{B}\right)^{-1} \mathbf{B}^T \mathbf{X}_\perp^T \mathbf{y}$$

$$Efficiency : \xi = \frac{1}{\sigma^2 Trace\left[\mathbf{B}\left(\mathbf{B}^T \mathbf{X}_\perp^T \mathbf{X}_\perp \mathbf{B}\right)^{-1} \mathbf{B}^T\right]}$$

# Summary

$$\text{No assumed shape}: \ \xi = \frac{1}{\sigma^2 Trace\left[\left(\mathbf{X}_\perp^T \mathbf{X}_\perp\right)^{-1}\right]}$$

$$\text{Assume basis functions}: \ \xi = \frac{1}{\sigma^2 Trace\left[\mathbf{B}\left(\mathbf{B}^T \mathbf{X}_\perp^T \mathbf{X}_\perp \mathbf{B}\right)^{-1} \mathbf{B}^T\right]}$$

$$\text{Assume known shape}: \xi = \frac{\mathbf{h}_0^T \mathbf{X}_\perp^T \mathbf{X}_\perp \mathbf{h}_0}{\sigma^2}$$

# Impact on Design

Definition of efficiency depends on model assumptions.

The design that achieves optimal efficiency depends on the definition of efficiency and therefore also depends on the model assumptions.

# What does a matrix do?

$$\mathbf{y} = \mathbf{X}\mathbf{h}$$

N-dimensional vector

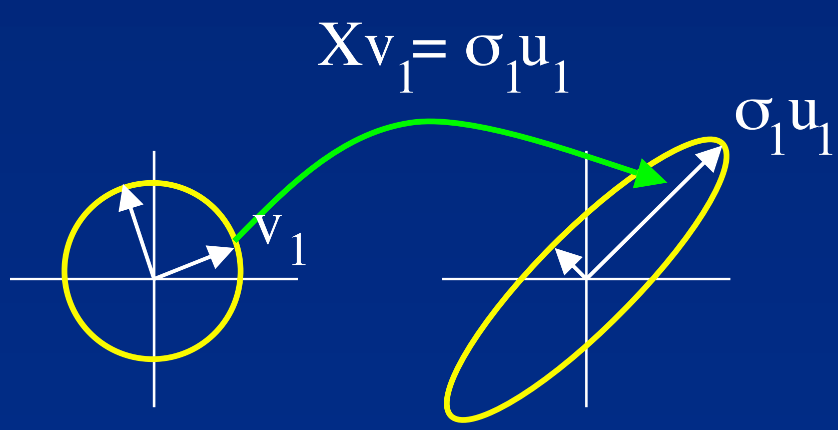
k-dimensional vector

N x k matrix

The matrix maps from a k-dimensional space to a N-dimensional space.

# Matrix Geometry

Geometric fact: The image of the a k-dimensional unit sphere under any *N* x *k* matrix is an N-dimensional hyperellipse.

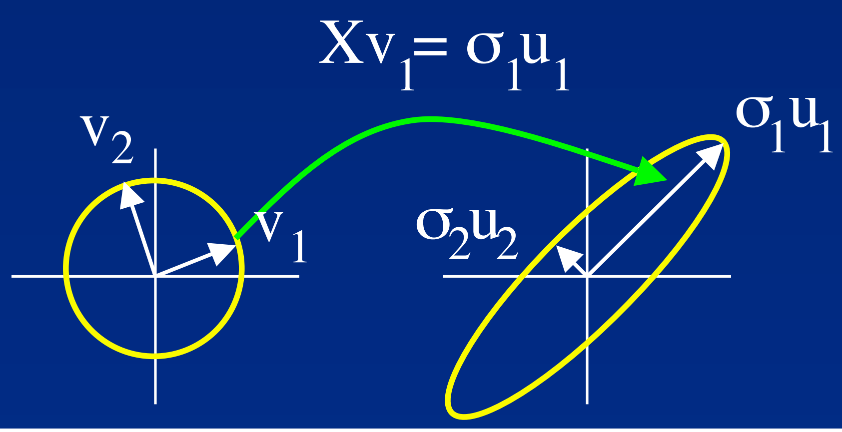





University of California San Diego

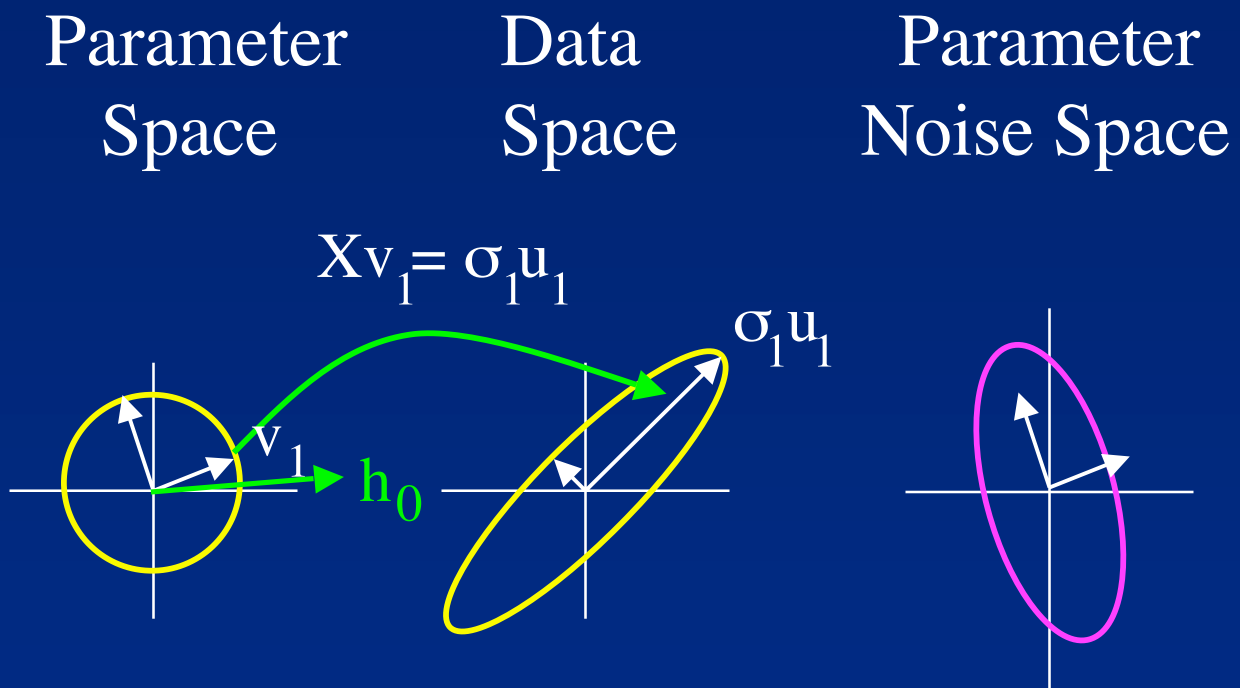
# Singular Value Decomposition

The right singular vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ are transformed into scaled vectors $\sigma_1\mathbf{u}_1$ and $\sigma_2\mathbf{u}_2$, where $\mathbf{u}_1$ and $\mathbf{u}_2$ are the left singular vectors and $\sigma_1$ and $\sigma_2$ are the singular values.




The singular values are the $k$ square roots of the eigenvalues of $kxk$ matrix $\mathbf{X}^T\mathbf{X}$.

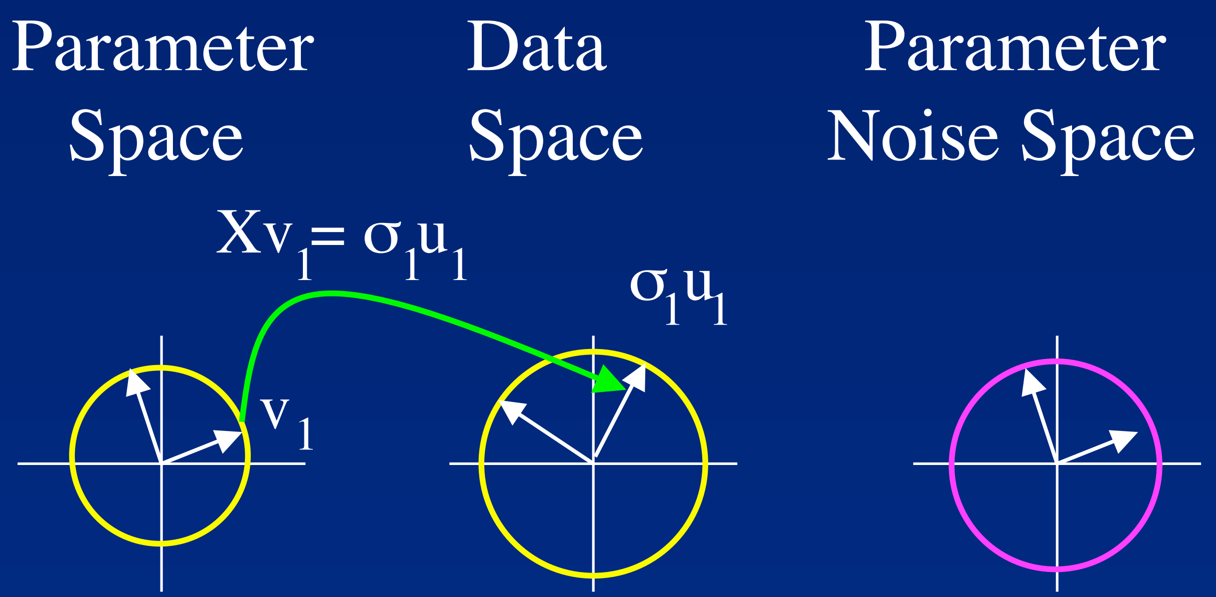
# Assumed HDR shape




Good for Detection

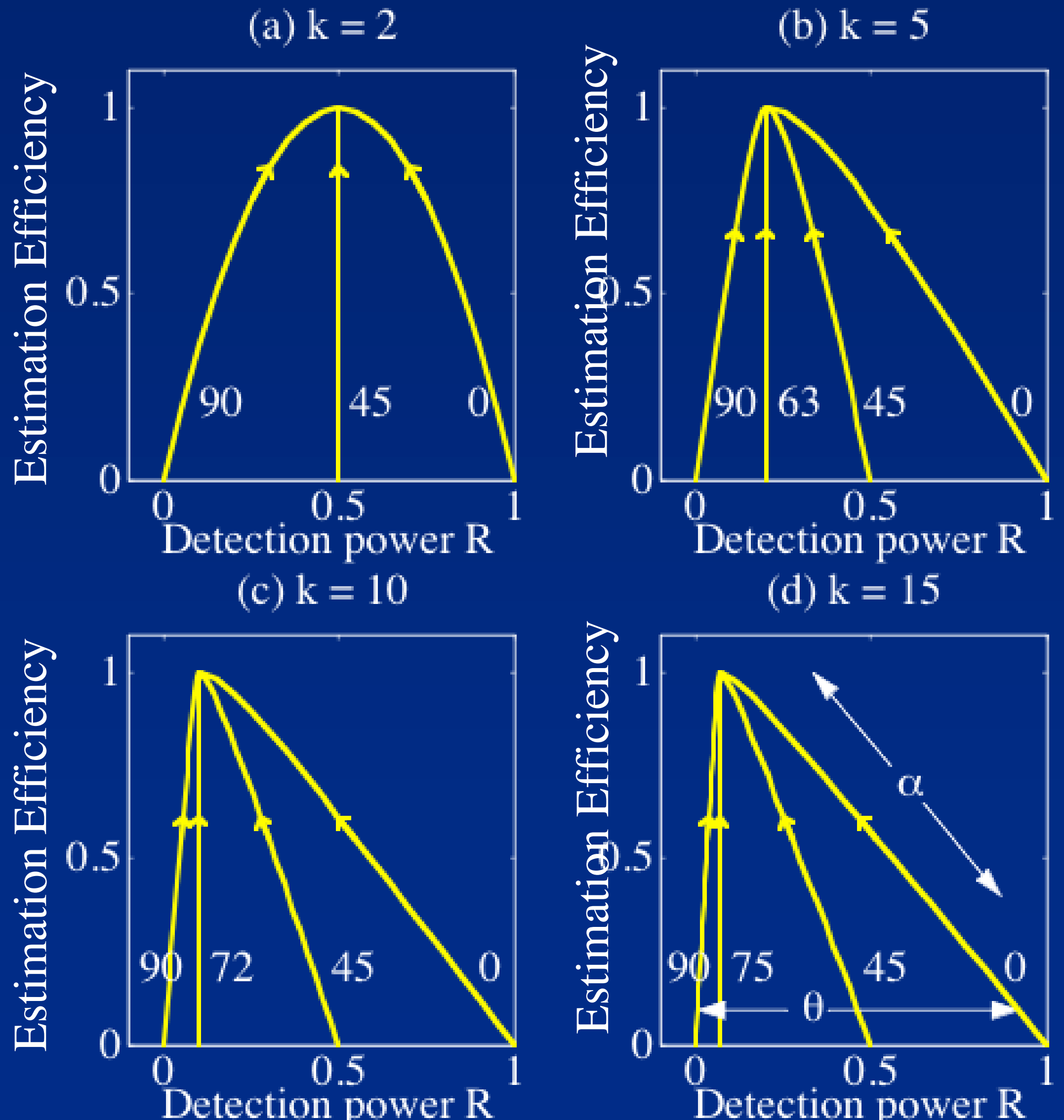
Efficiency here is optimized by amplifying the singular vector closest to the assumed HDR. This corresponds to maximizing one singular value while minimizing the others.

# No assumed HDR shape

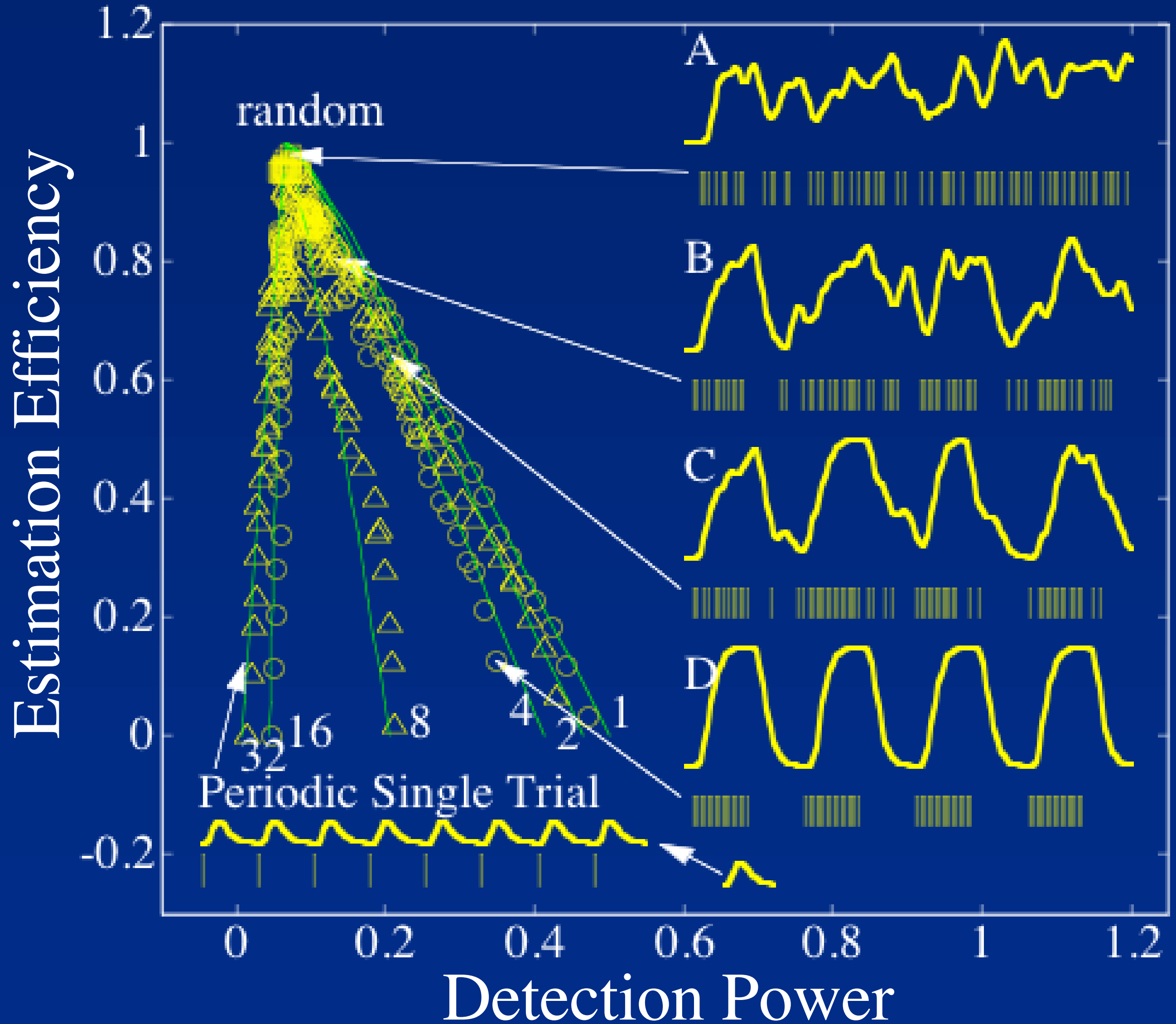



Here the HDR can point in any direction, so we don't want to preferentially amplify any one singular value. This corresponds to an equal distribution of singular values.

# Theoretical Curves



# Efficiency vs. Power

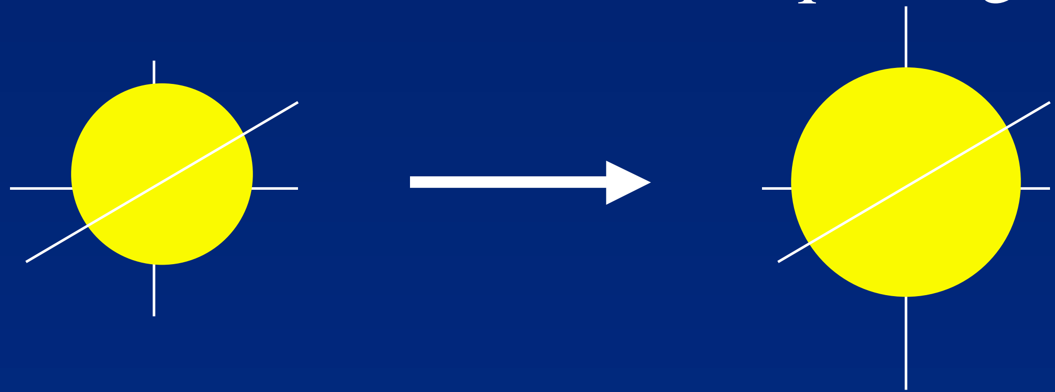




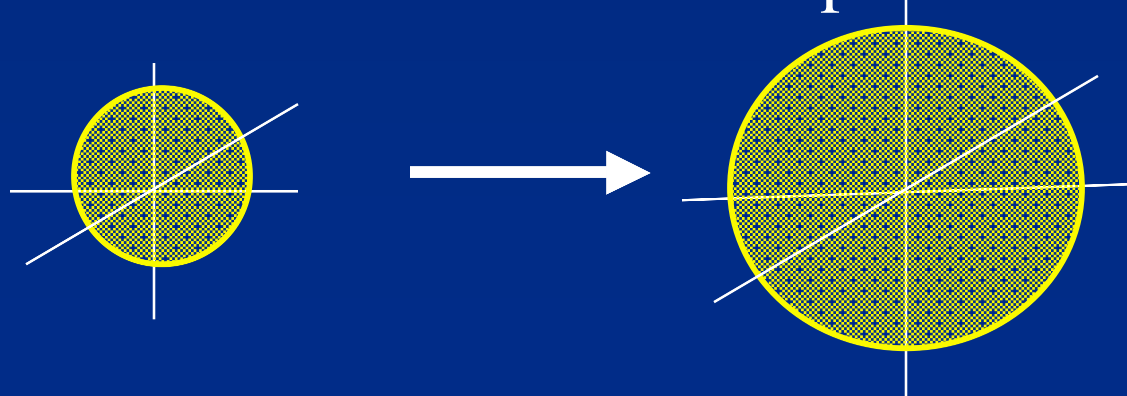

University of California San Diego

# Basis Functions

If no basis functions, then use equal eigenvalues.



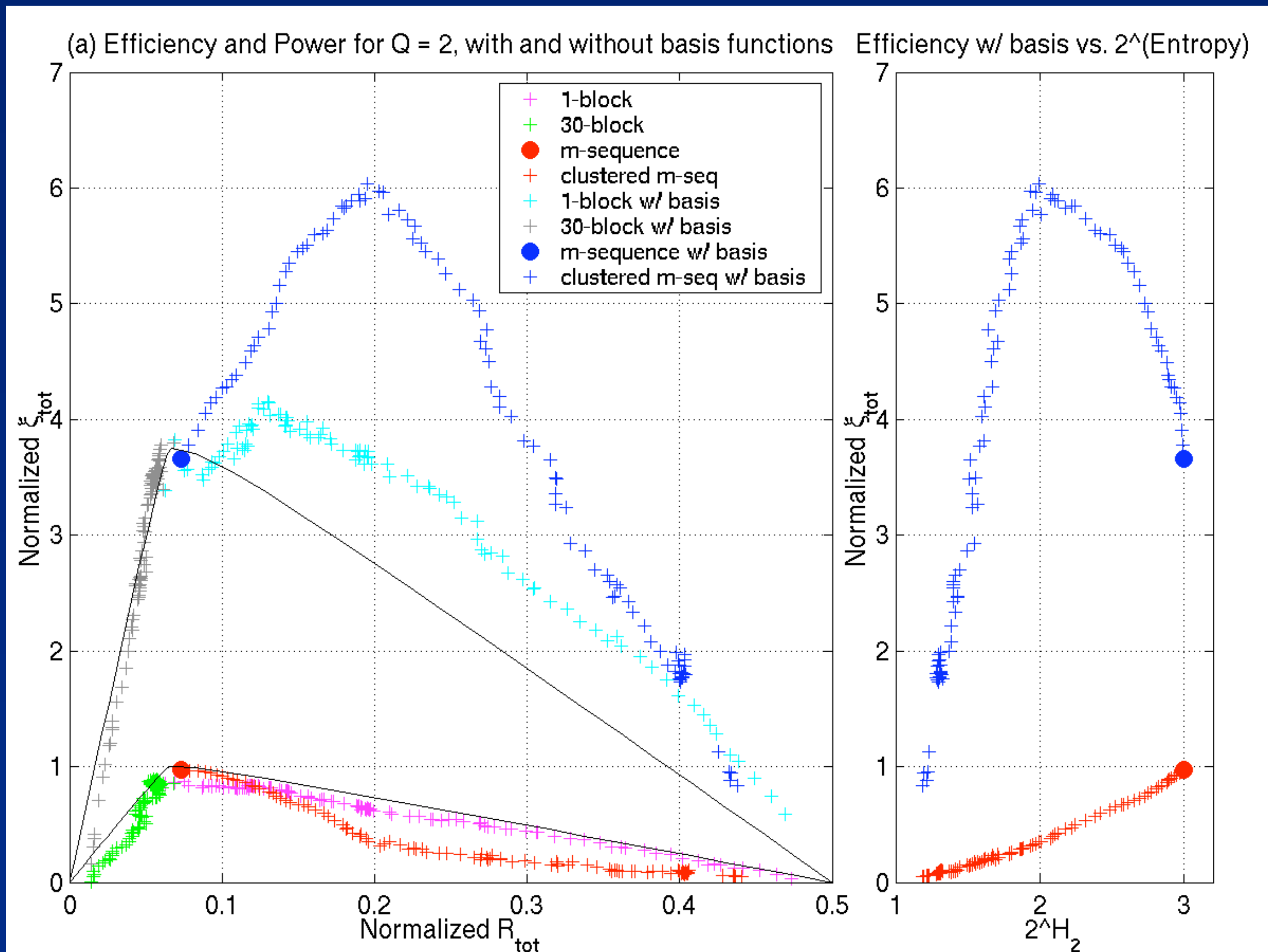
If we know the the HDR lies within a  subspace, we should maximize the singular values in this subspace and minimize outside of this subspace.





University of California San Diego

# Basis Functions



# Overview of designs

Known HDR: Maximize one dominant singular value -- block designs.

Unknown HDR: Equalize singular values -- randomized designs, m-sequences.

Somewhat known HDR:  Amplify singular values within the subspace of interest -- semi-random designs, permuted block, clustered m-sequences. (also good in presence of correlated noise).

# Multiple Trial Types

Previously: 1 trial type + control (null)

A A N A A N A A A N

Extend to experiments with multiple trial types

A B A B N N A N B B A N A N A

B A D B A N D B C N D N B C N

# Multiple Trial Types GLM

$$\mathbf{y} = \mathbf{Xh} + \mathbf{Sb} + \mathbf{n}$$

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

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

# Multiple Trial Types Overview

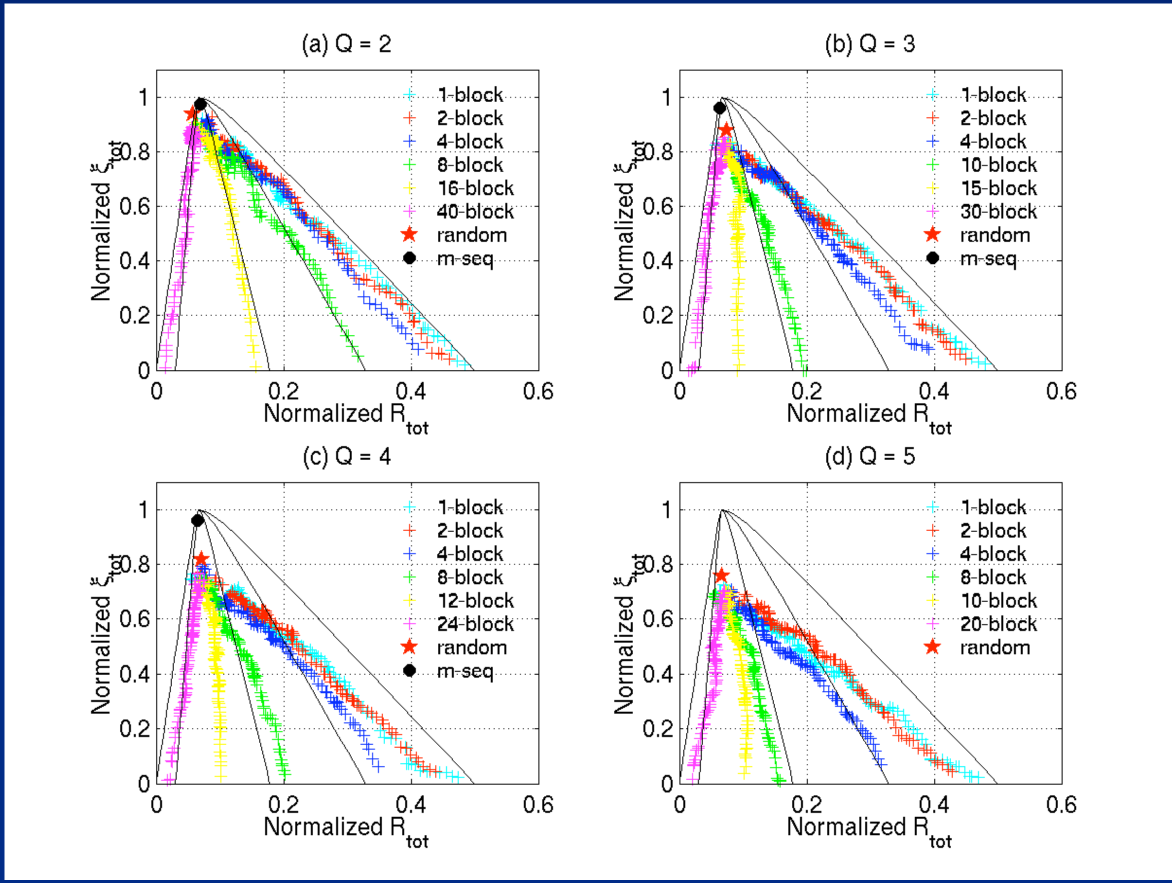
Efficiency includes individual trials and also contrasts between trials.

$$R_{tot} = \frac{K}{\begin{pmatrix}\text{average variance of HRF amplitude estimates} \\ \text{for all trial types and pairwise contrasts}\end{pmatrix}}$$

$$\xi_{tot} = \frac{1}{\begin{pmatrix}\text{average variance of HRF estimates} \\ \text{for all trial types and pairwise contrasts}\end{pmatrix}}$$

# Multiple Trial Types Trade-off

Can show that the same geometric intuition about singular values applies.






University of California San Diego

# Optimal Frequency

Can also weight how much you care about individual trials or contrasts. Or all trials versus events.
Optimal frequency of occurrence depends on weighting.
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.

$$p = \frac{Q(2k_1 - 1) + Q^2(1 - k_1) + k_1^{1/2}\left(Q(2k_1 - 1) + Q^2(1 - k_1)\right)^{1/2}}{Q(Q-1)(k_1 Q - Q - k_1)}$$

# Design

As the number of trial types increases, it becomes more difficult to achieve the theoretical trade-offs. Random search becomes impractical.

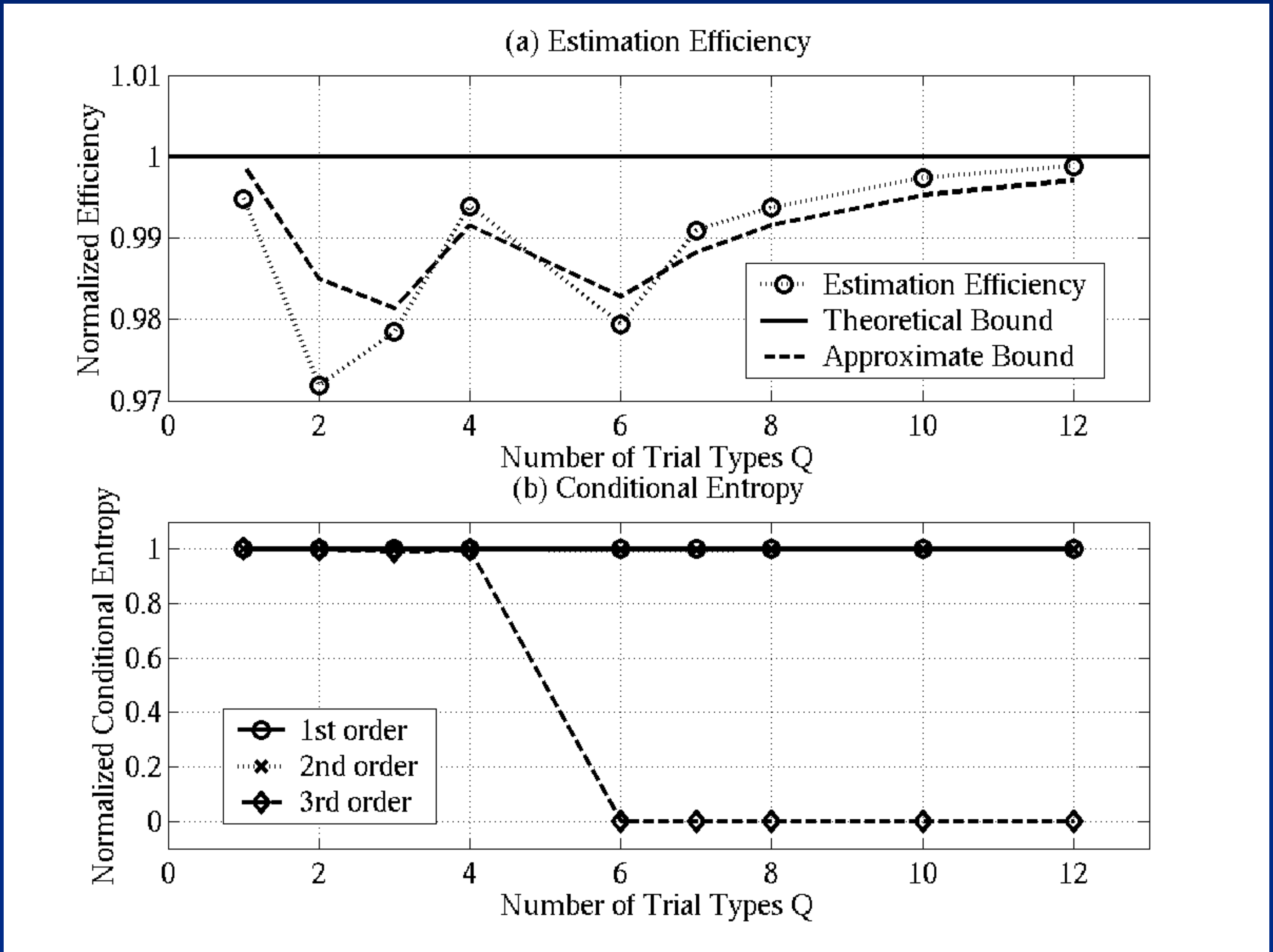
For unknown HDR, should use an m-sequence based design when possible.

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



University of California San Diego

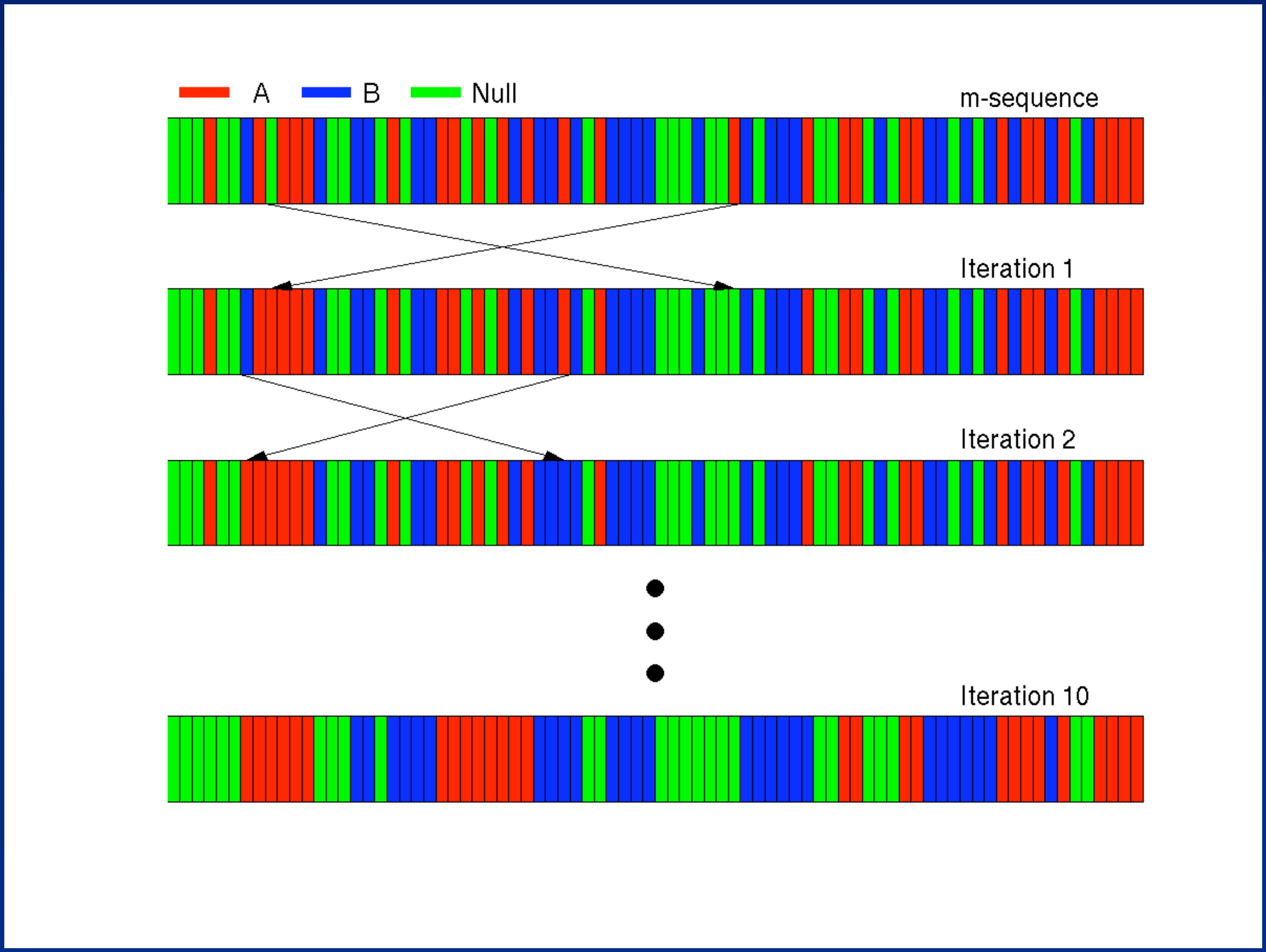
# Optimality of m-sequences






University of California San Diego

# Clustered m-sequences






University of California San Diego

# Topics we haven't covered.

The impact of correlated noise -- this will change the optimal design. Can  you the geometric intuition from singular values to gain some understanding.

Entropy of designs.

# Concluding remarks

Geometric view is useful for developing intuition into the meaning of basis statistical measures and design principles.

It is also very useful as a sanity check of one's theoretical results.