SYSTEM IDENTIFICATION

Prof. Alexander Huth

9/14/2017

RECAP

* Natural stimuli are your friend because:
  * Generalization
  * Efficiency

RECAP

* But our variables of interest are correlated/confounded in natural stimuli!

  * ***REGRESSION***

TODAY

* Overview of system identification

* Spatiotemporal models

SYSTEM IDENTIFICATION

$$Y = f(X)$$

* What kind of a function is f?

SYSTEM IDENTIFICATION

READ THIS PAPER for next Tuesday (9/19):

Complete Functional Characterization of Sensory Neurons by System Identification

Michael C.-K. Wu,[1] Stephen V. David,[2] and Jack L. Gallant[3,4]

https://github.com/alexhuth/n4cs-fa2017/

SYSTEM IDENTIFICATION

* Linear model $$Y = X\beta$$

* Linearized model $$Y = \mathbb{L}(X)\beta$$

* Nonlinear model $$Y = \Theta(X)$$

LINEAR MODELS

$$Y = \underset{\text{image pixels}}{X}\beta$$

X1, Y=0.7

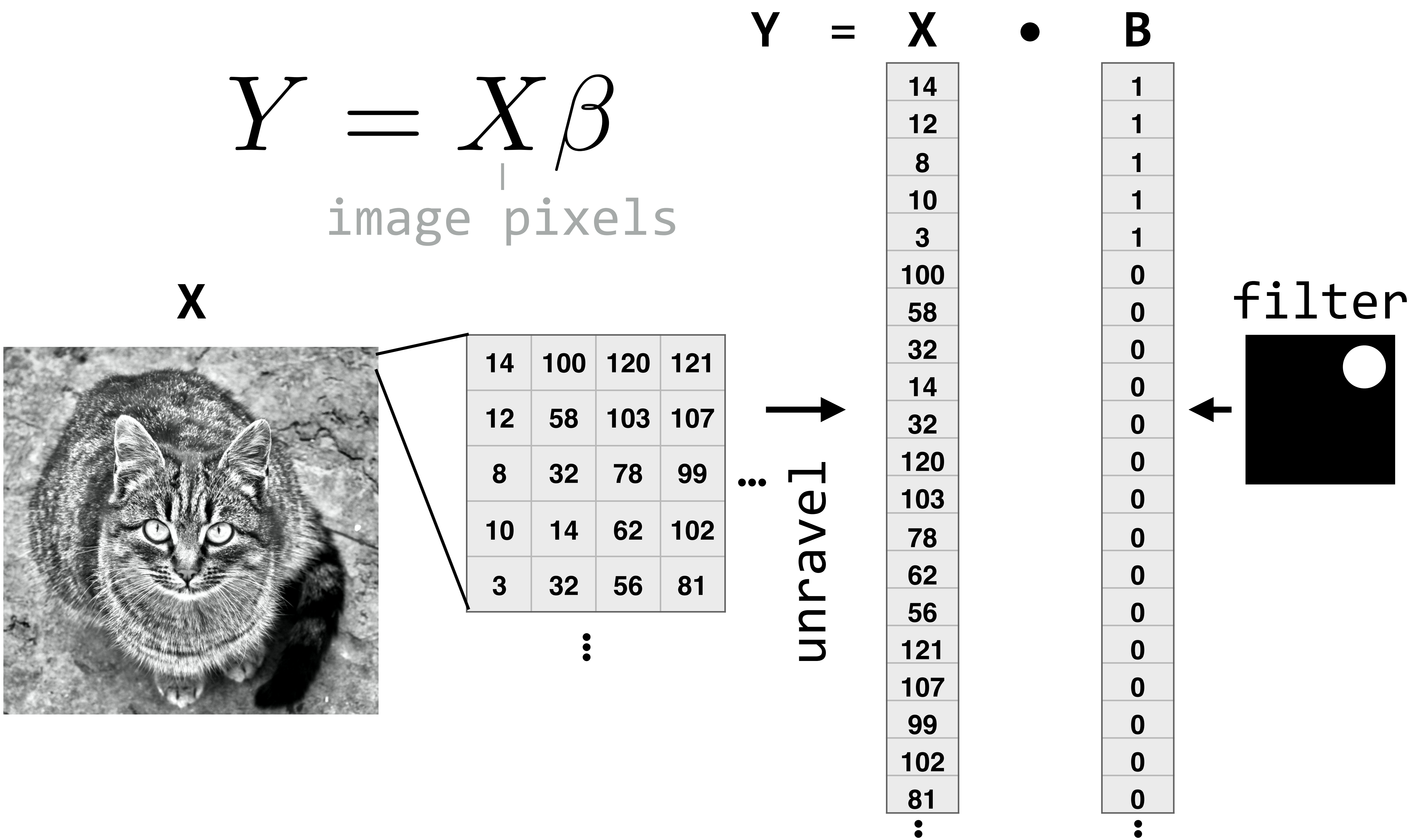
X2, Y=0.3

X3, Y=0.0

LINEAR MODELS

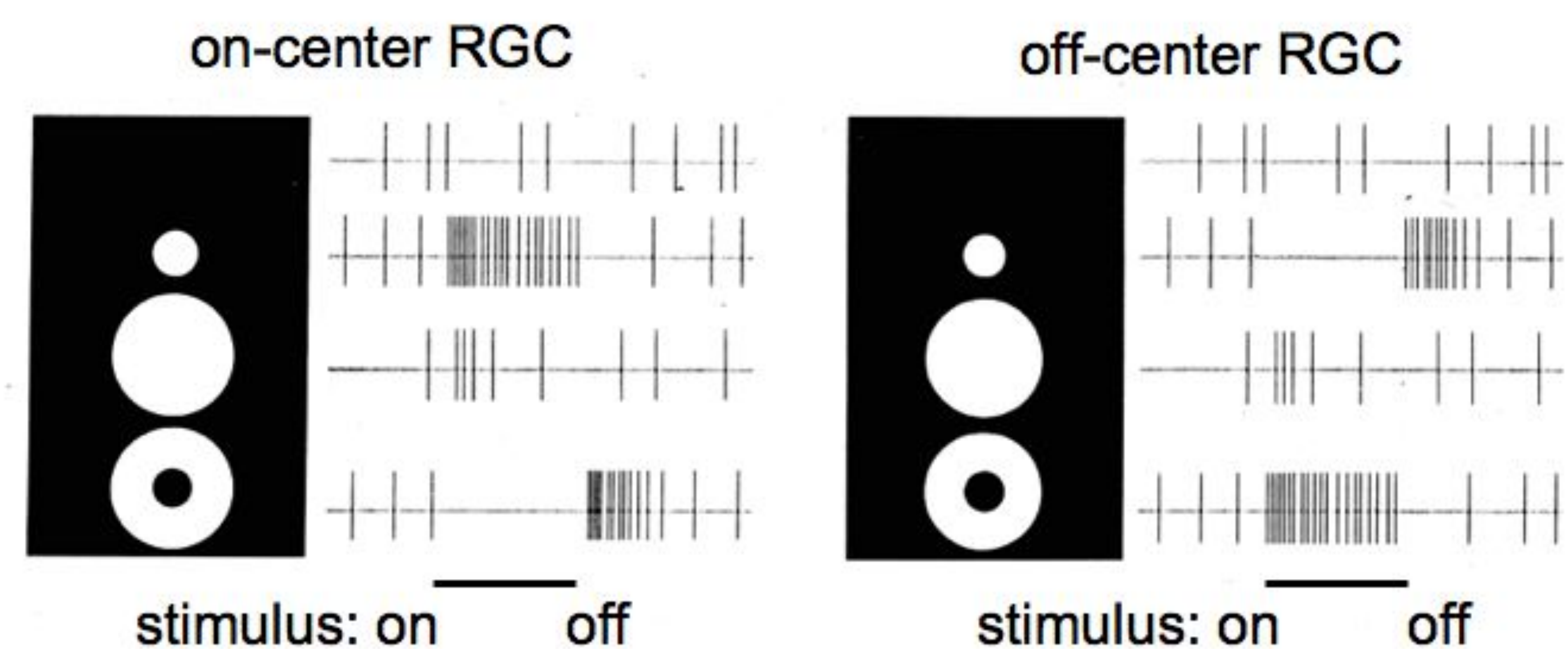
$$Y = X\beta$$

image pixels

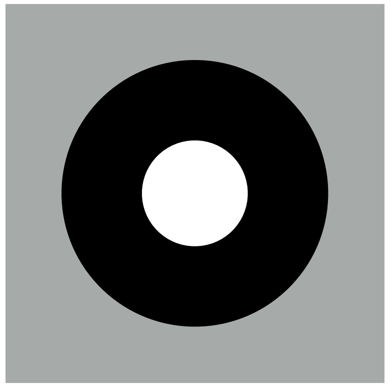


LINEAR MODELS

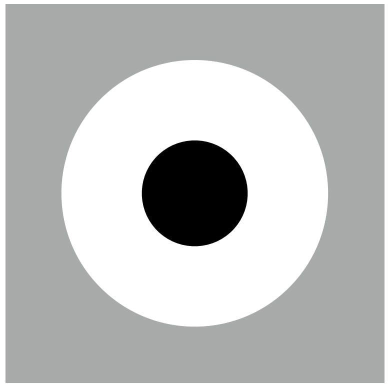
Retinal ganglion cell responses



Beta
(on-center)



Beta
(off-center)



http://www.cns.nyu.edu/~david/courses/perception/lecturenotes/ganglion/ganglion.html

LINEARIZED MODELS

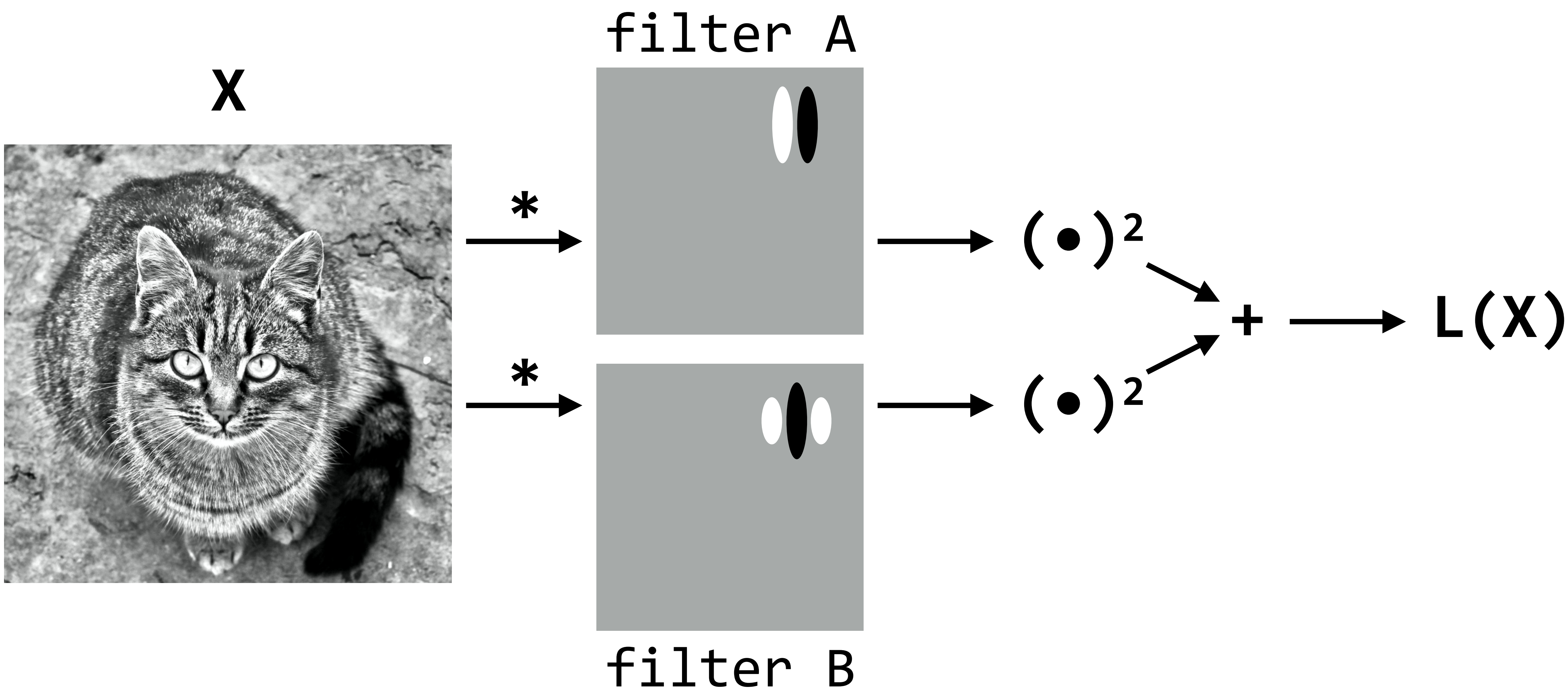
$$Y = \mathbb{L}(X)\beta$$

* **L** is some non-linear function of the stimulus ***X*** that gives us *features*

* ***Beta*** is a linear weighting of the *features* that gives us the response ***Y***

LINEARIZED MODELS

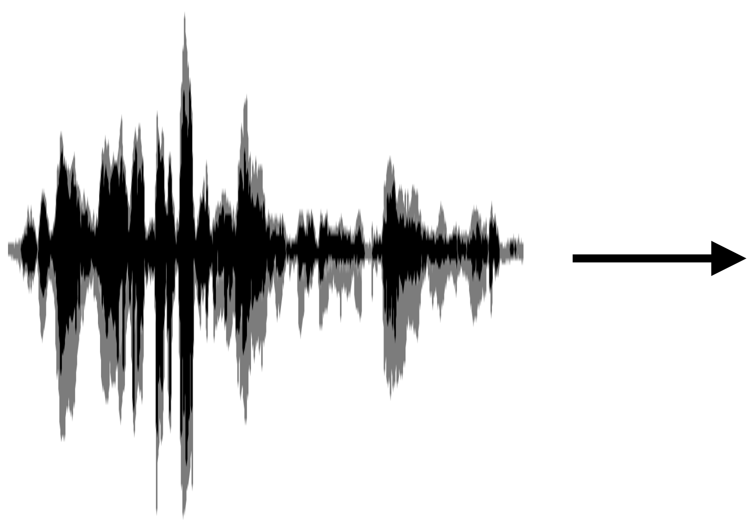
$$Y = \mathbb{L}(X)\beta$$



LINEARIZED MODELS

$$Y = \mathbb{L}(X)\beta$$

story



word matrix

time

word				
now	1	0	0	
this	1	0	0	
is	0	1	0	…
a	0	1	0	
story	0	0	1	
⋮				

NONLINEAR MODELS

$$Y = \Theta(X)$$

X1, Y="cat"

X2, Y="dog"

X3, Y="mate"

NONLINEAR MODELS

$$Y = \Theta(X)$$

X1, Y=[1,0,0]

X2, Y=[0,1,0]

X3, Y=[0,0,1]

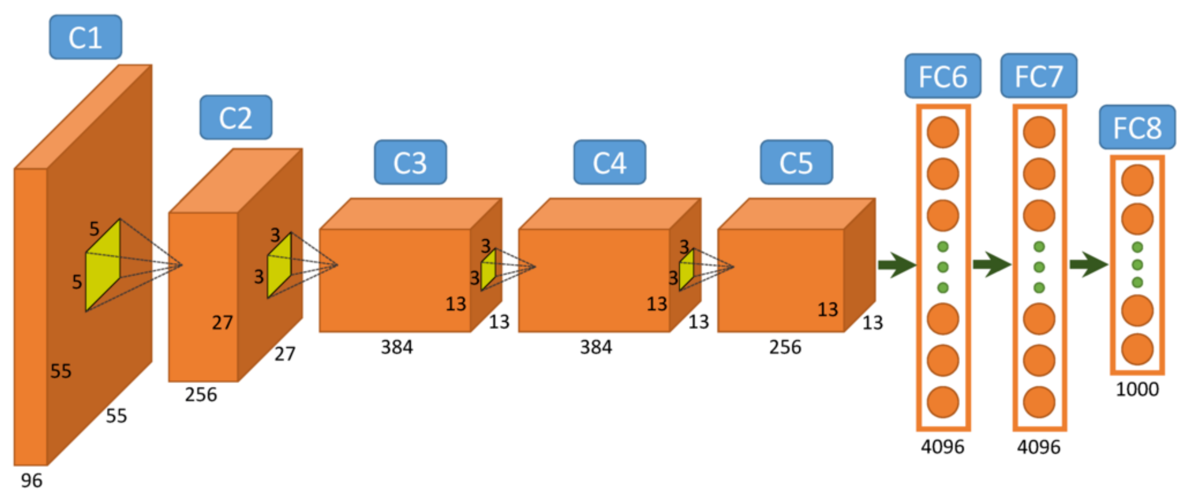
NONLINEAR MODELS

$$Y = \Theta(X)$$

image pixels

X **Theta** **Y**

(deep neural network)



"cat"

https://www.saagie.com/blog/object-detection-part1

SYSTEM IDENTIFICATION

* **Linear model**
  * easy, usually pointless
* **Linearized model**
  * sweet spot, but requires **hypothesis**!
* **Nonlinear model**
  * very expensive, need lots of data

LINEARIZING TRANSFORMATION

=

FEATURE SPACE

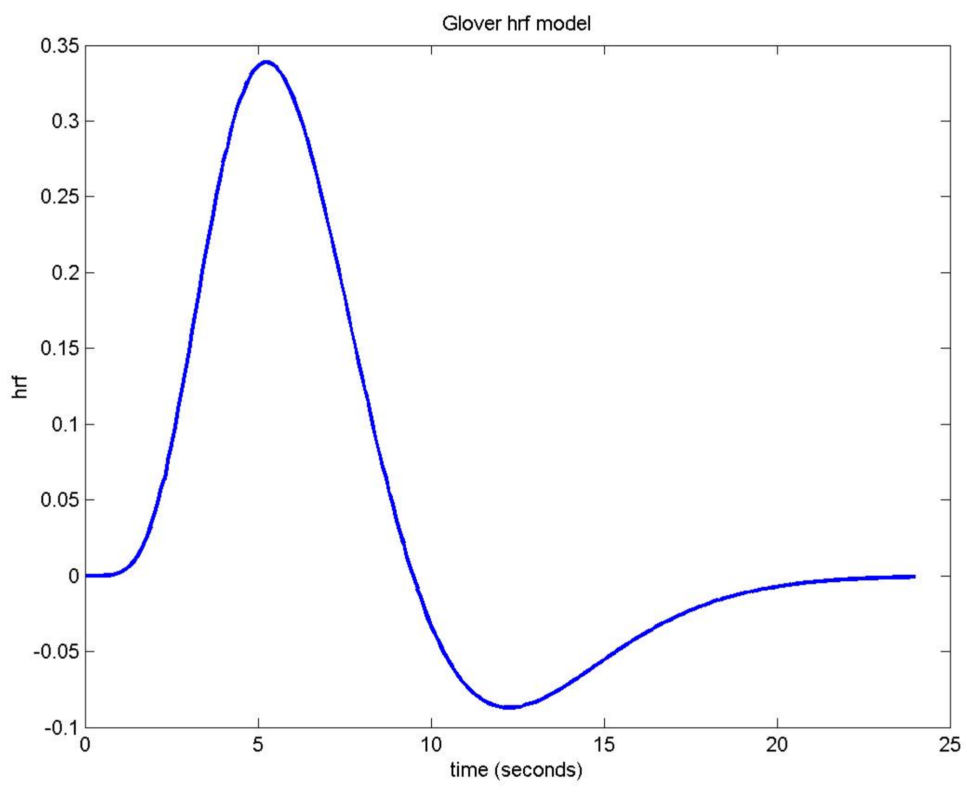
=

HYPOTHESIS

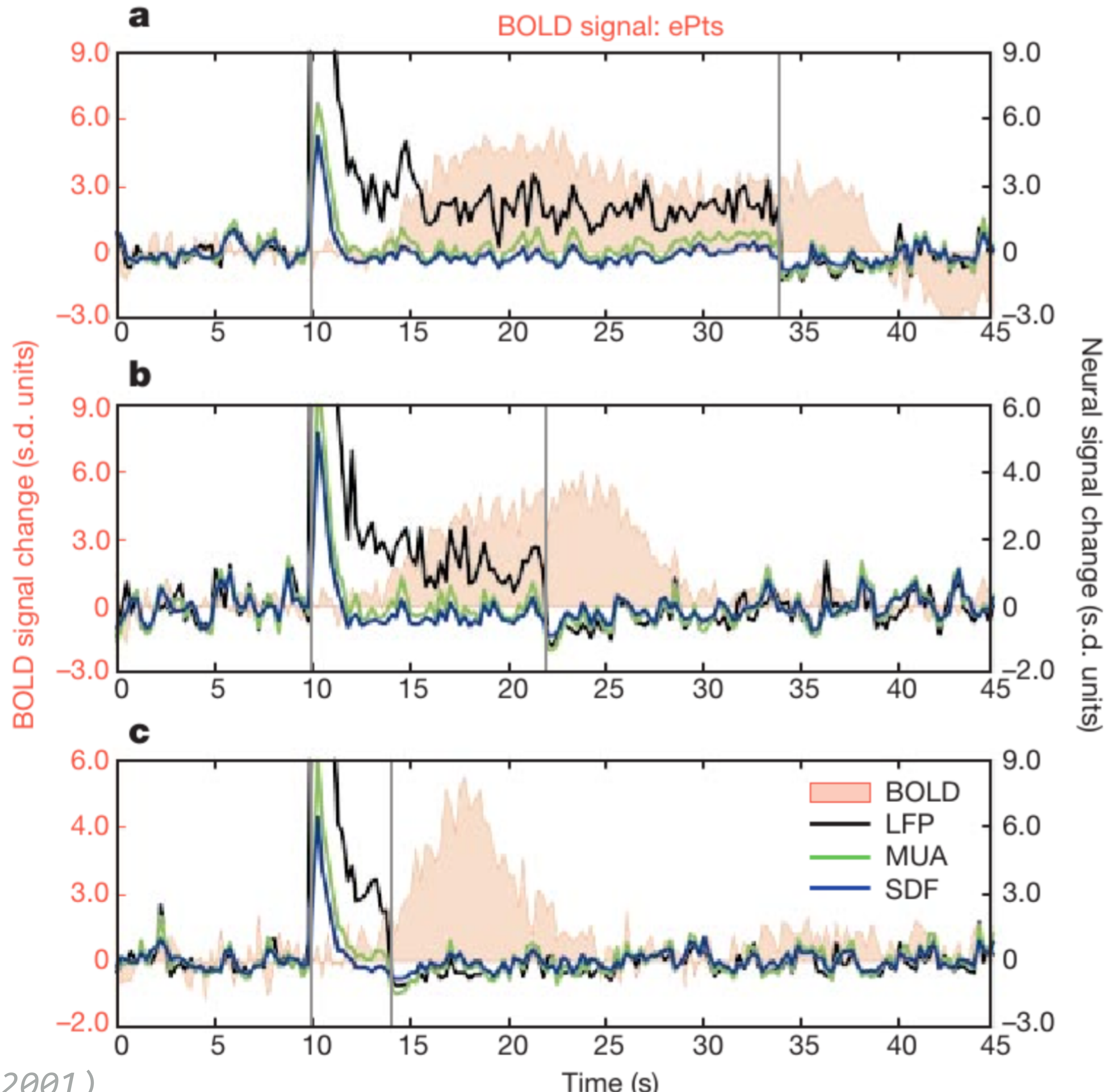
BOLD & HRF

* BOLD = blood-oxygen level dependent

* HRF = hemodynamic response function



typical model HRF



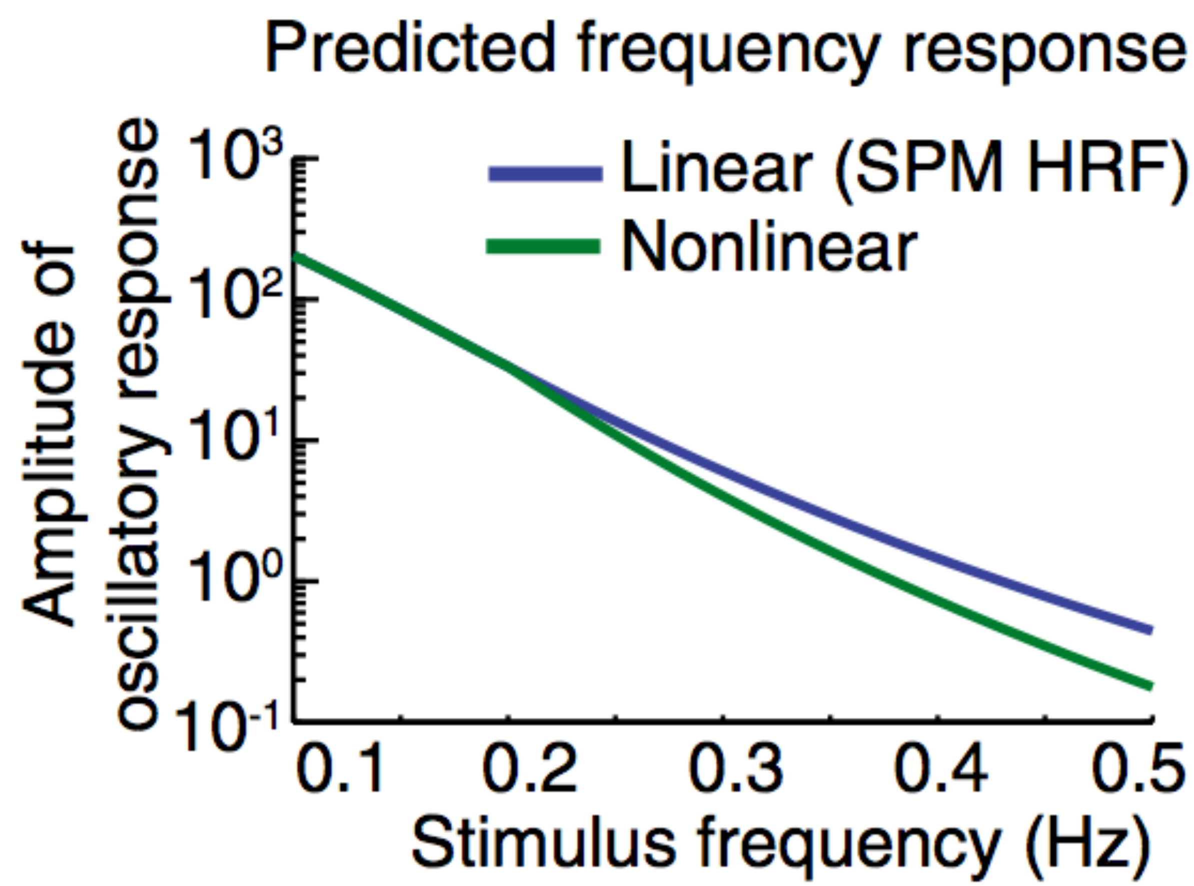
Logothetis et al. (2001)

BOLD & HRF

* HRF is approximately linear

* …but not perfectly

HRF frequency response



Lewis, Setsompop, Rosen, & Polimeni (2016)

SPATIOTEMPORAL MODEL

* HRF needs to be accounted for in model
* Three approaches:
    * assume HRF 👎 👎 👎 🙅‍♀️
    * fit **space-time separable** model
    * fit **space-time inseparable** model

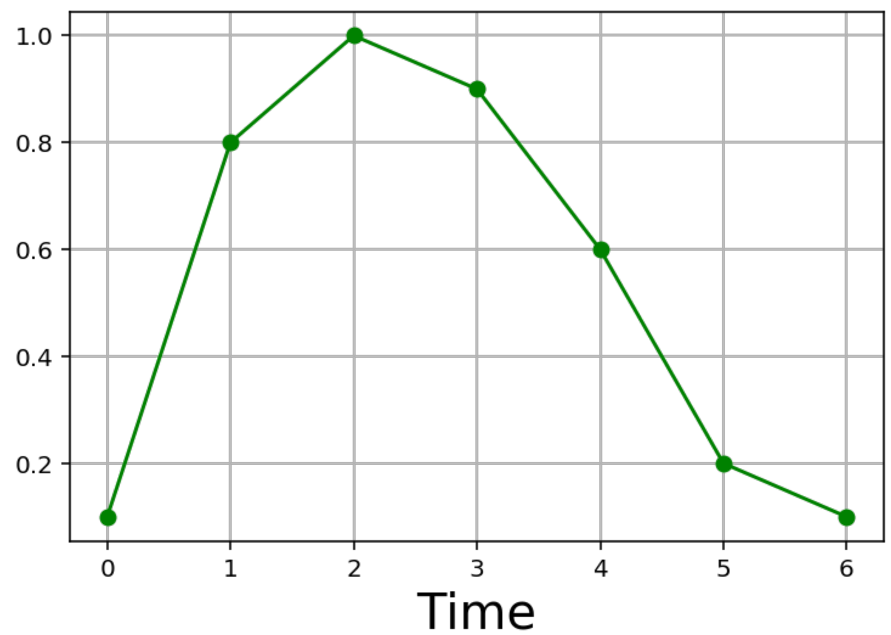
SPACE-TIME SEPARABLE

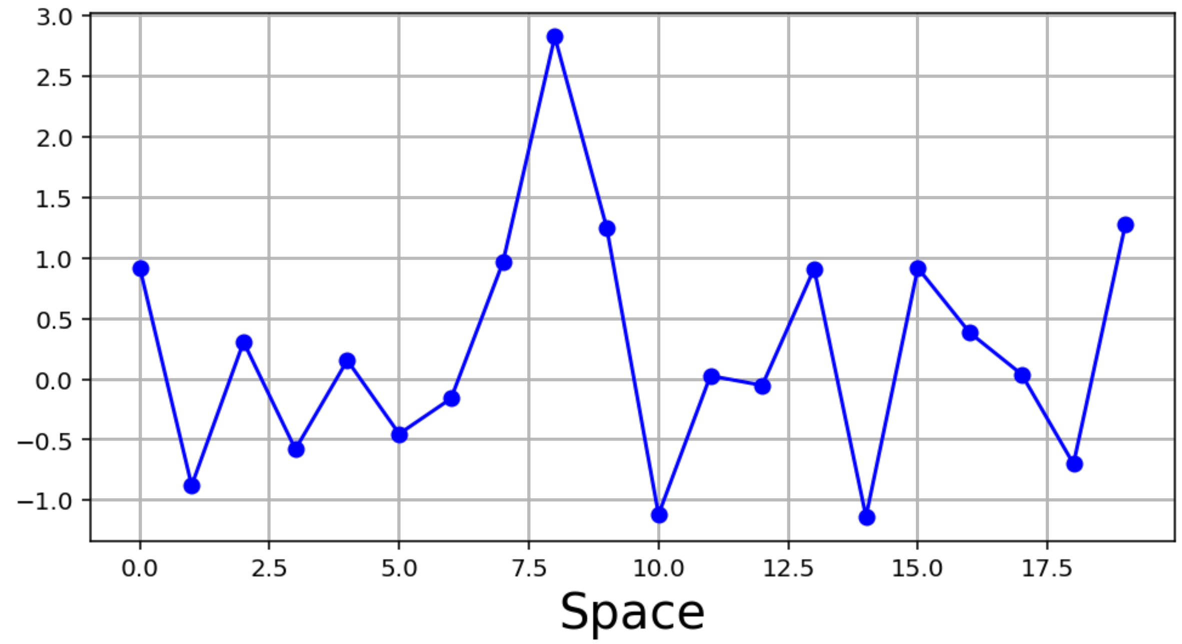
* Model is outer product of one spatial kernel & one temporal kernel

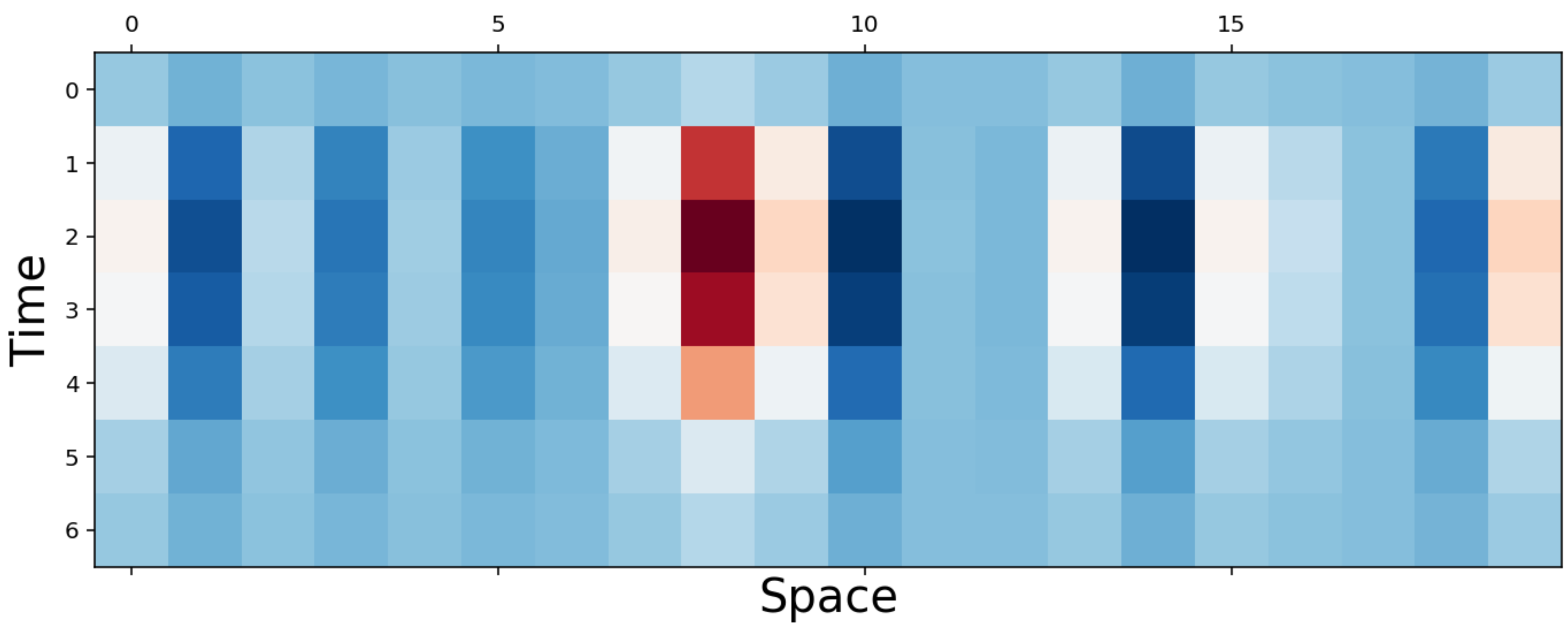
* Fit using e.g. generalized least squares (GLS)

  * Each model (i.e. each voxel) must be fit separately

SPACE-TIME SEPARABLE







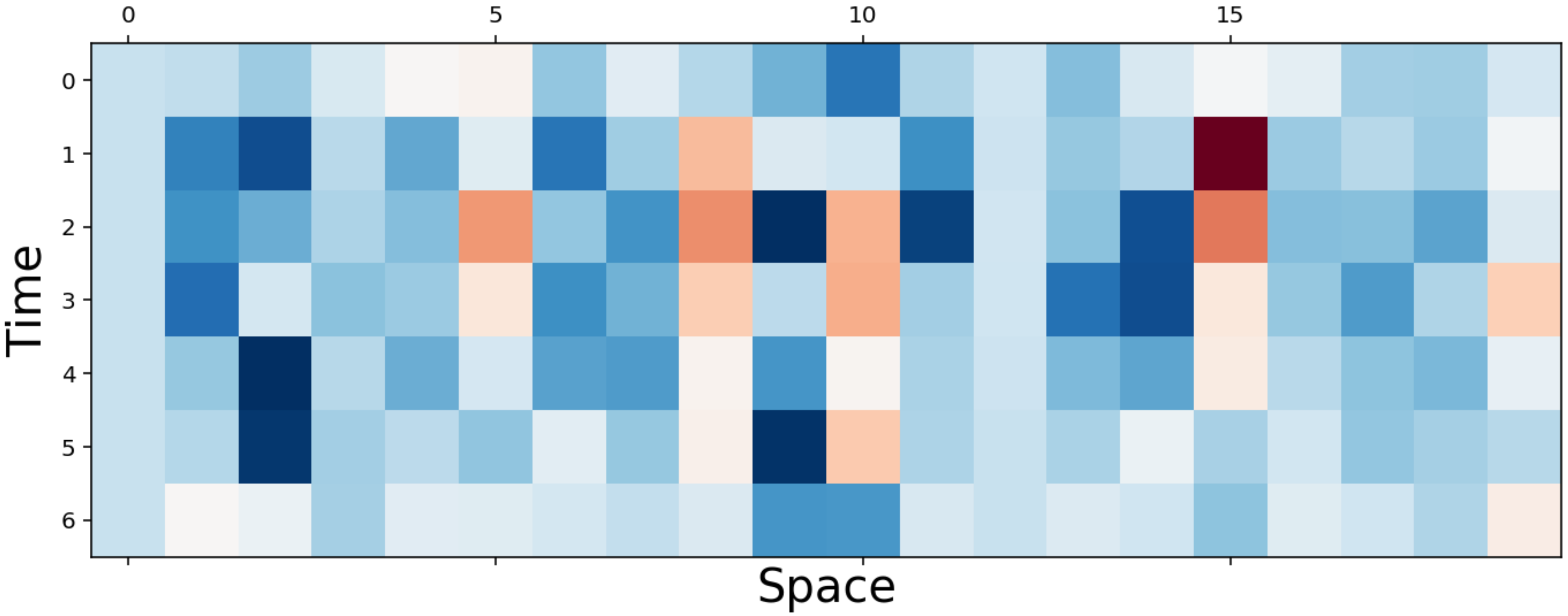
rank=1

SPACE-TIME INSEPARABLE

* Different temporal kernel for each spatial dimensions (i.e. each feature)

* Fit using ordinary methods

SPACE-TIME INSEPARABLE

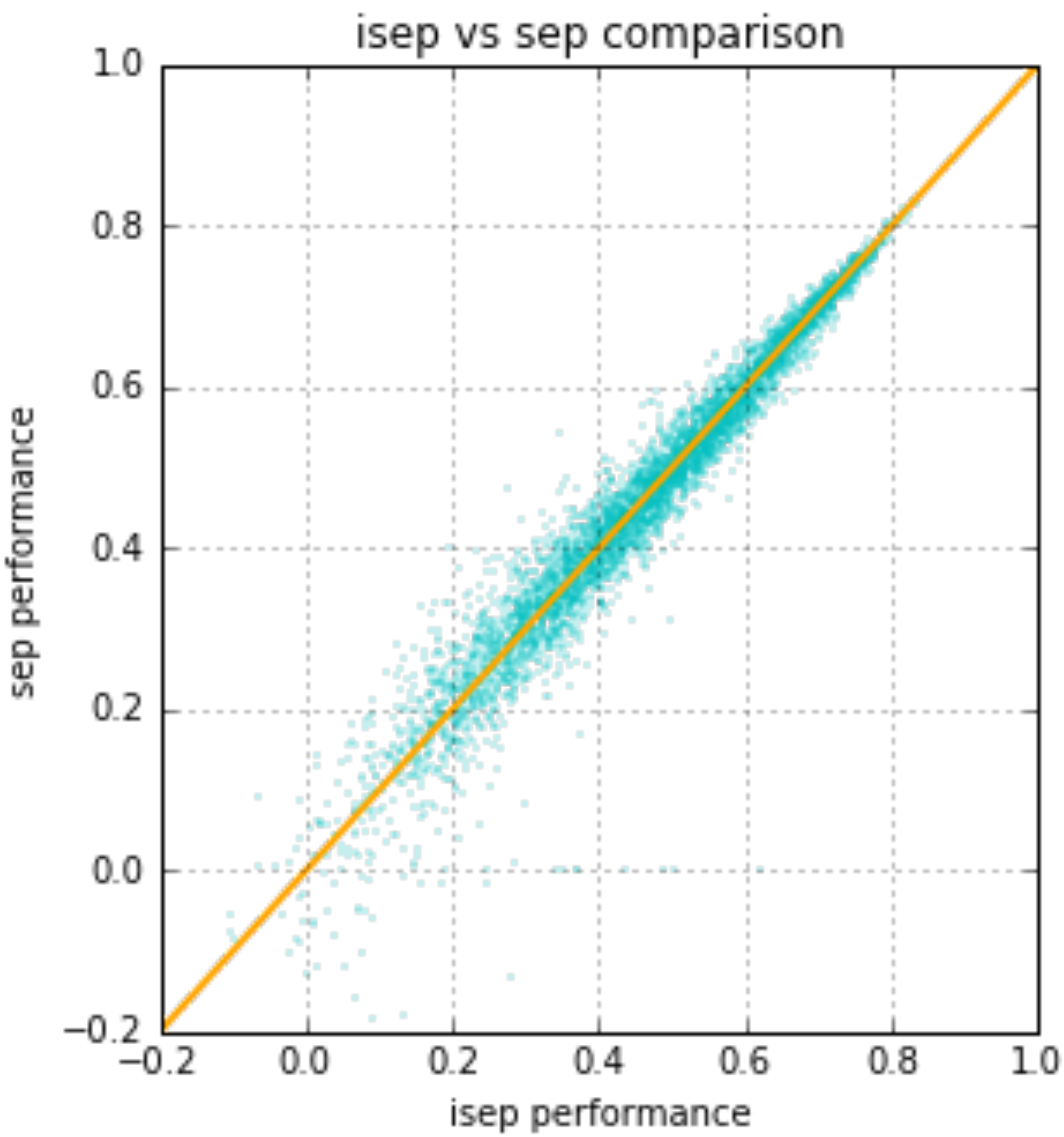


SPATIOTEMPORAL MODEL

- **Separable**
  - Expensive
  - Highly constrained
- **Inseparable**
  - Cheap
  - Unconstrained

SPATIOTEMPORAL MODEL

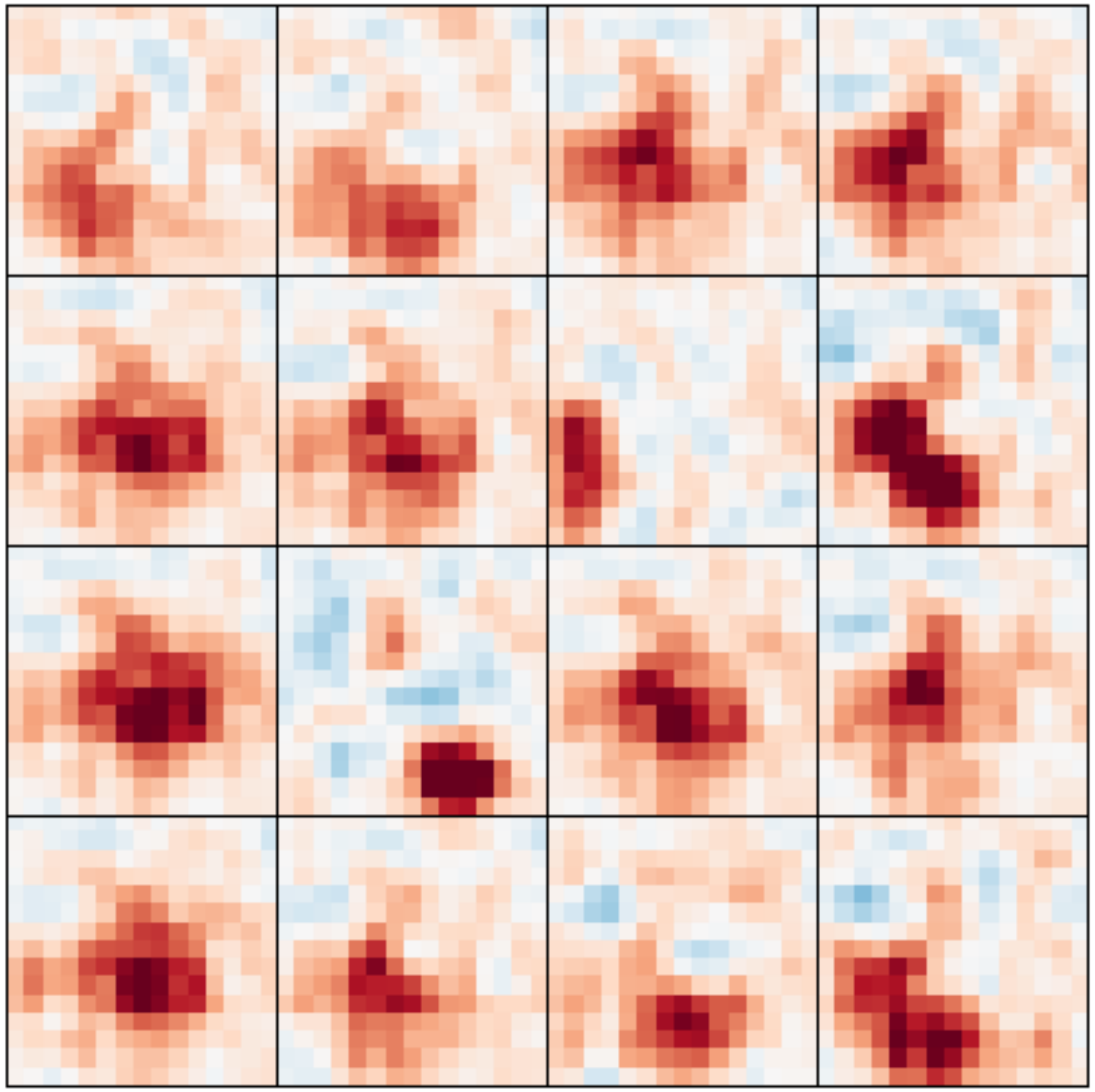
* In practice, both perform very similarly, so I prefer inseparable models



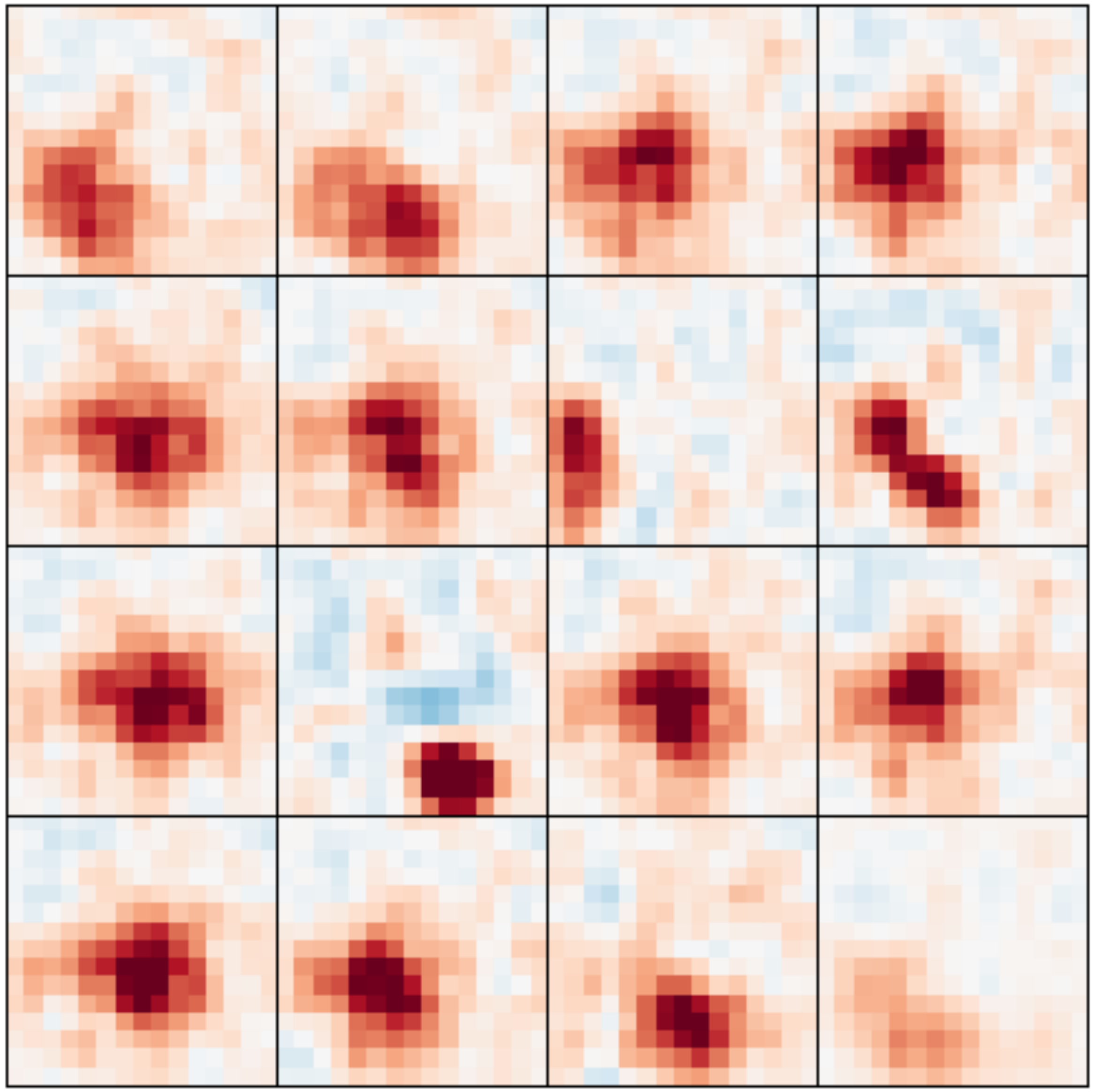
h/t Anwar Nuñez

SPATIOTEMPORAL MODEL

inseparable model weights



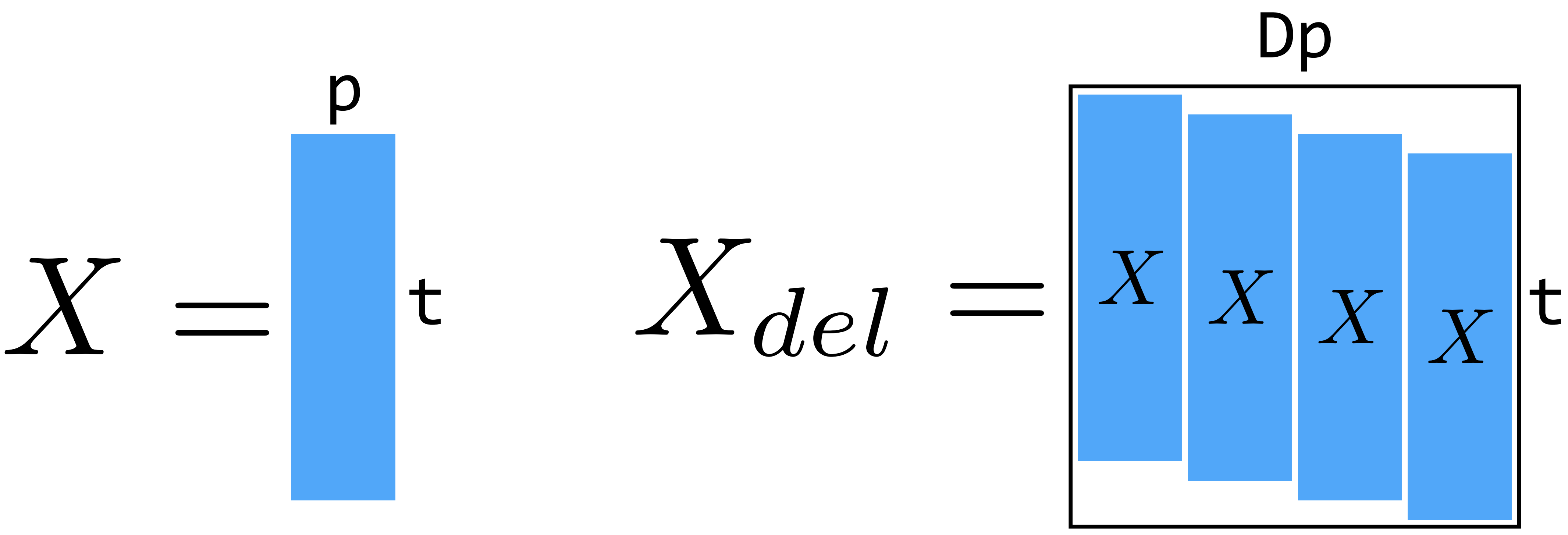
separable model weights



h/t Anwar Nuñez

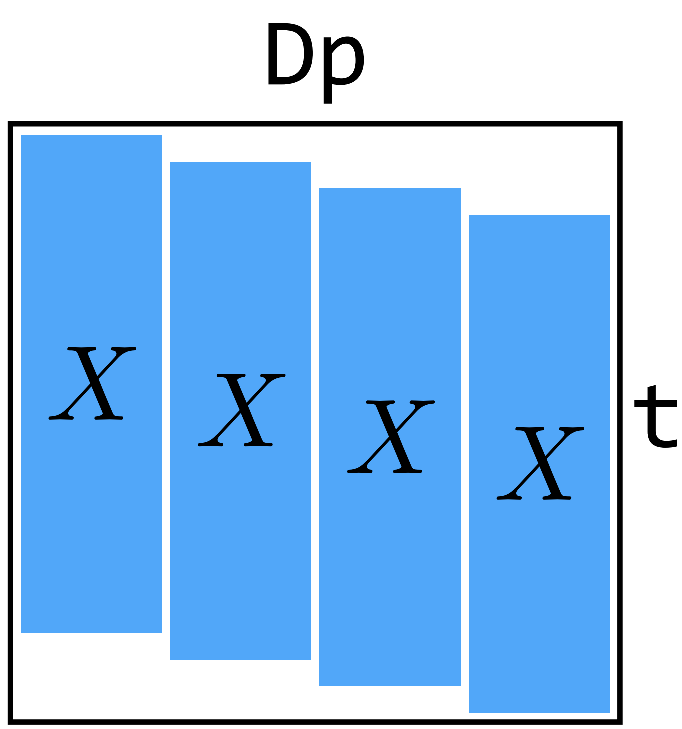
FINITE IMPULSE RESPONSE (FIR)

* Concatenate delayed copies of the stimulus matrix



FINITE IMPULSE RESPONSE (FIR)

$$X_{del} =$$



$$Y = X_{del} \begin{bmatrix}\beta_1\\ \beta_2\\ \beta_3\\ \beta_4\end{bmatrix}$$

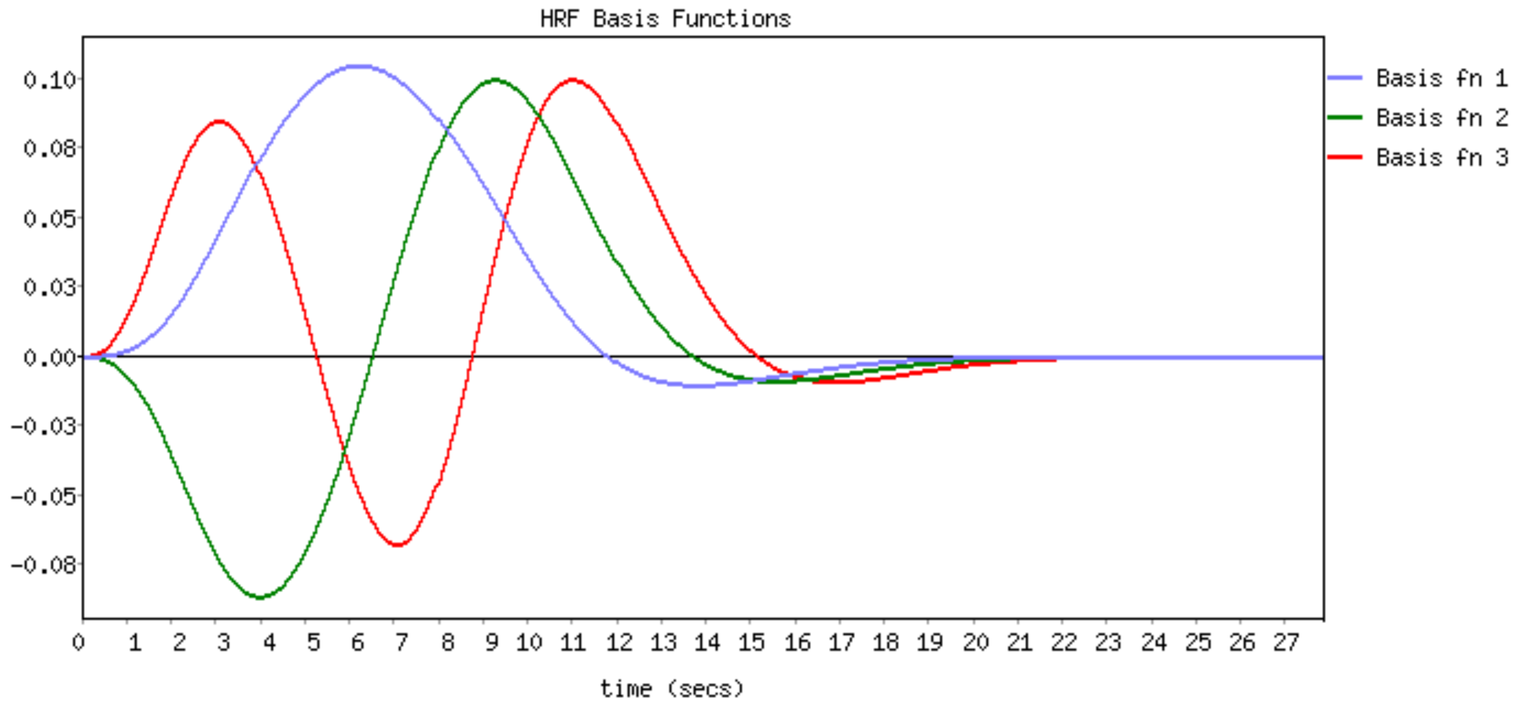
$$Y(t) = \beta_1 X(t) + \beta_2 X(t-1) + \beta_3 X(t-2) + \beta_4 X(t-3)$$

Y(t) = \beta_1 X(t) + \beta_2 X(t-1) + \beta_3 X(t-2) + \beta_4 X(t-3)

Y = X_{del} \begin{bmatrix}\beta_1\\

HRF BASIS

* Instead of delaying, convolve stimuli with a set of filters



* (Although delaying can be thought of as convolving with impulse delay filters!)

https://surfer.nmr.mgh.harvard.edu/

NEXT TIME

* Fitting linear models