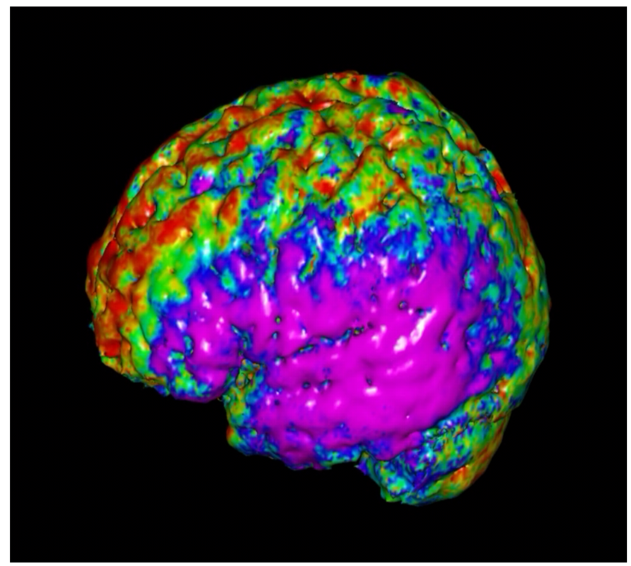
Smoothing Images from Population Studies

Hongtu Zhu, Ph.D.
Department of Biostatistics and Biomedical Research Imaging Center
University of North Carolina at Chapel Hill



Outline

Motivation

Smoothing Strategies

Multiscale Adaptive Smoothing Models

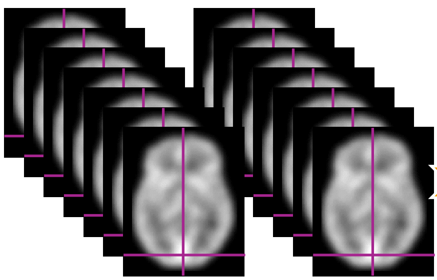
Simulation Studies

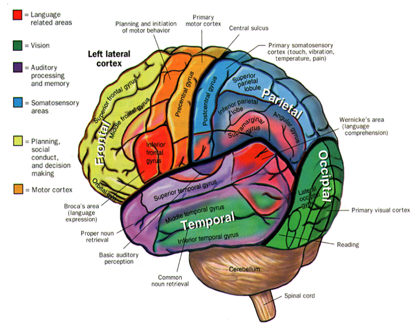
Real Data Analysis

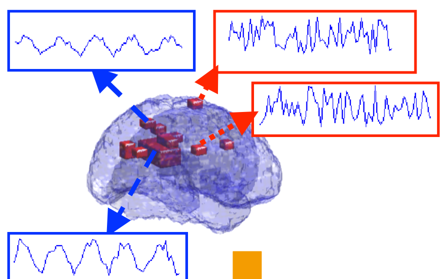
Motivation

Functional Connectivity

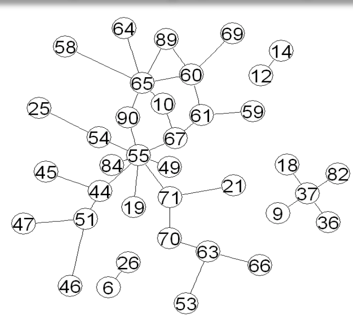
EEG, fMRI, resting fMRI

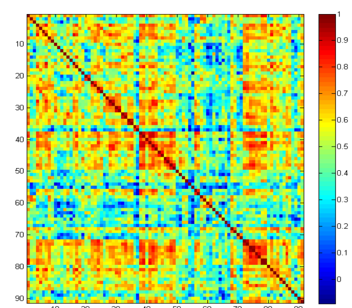






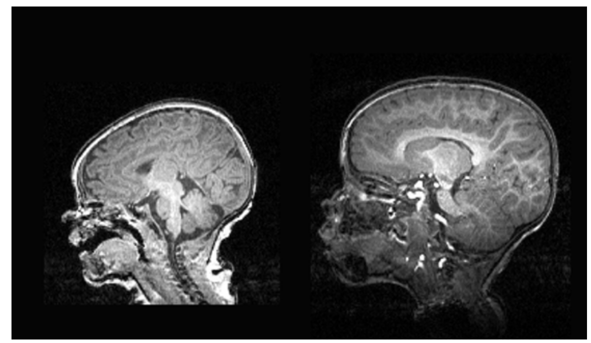
group 1 | **group 2**

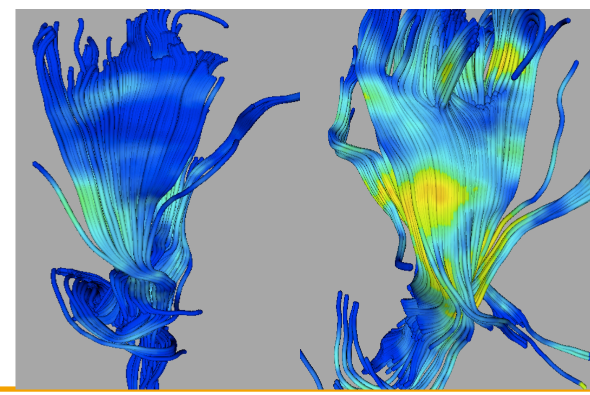




Structural Connectivity

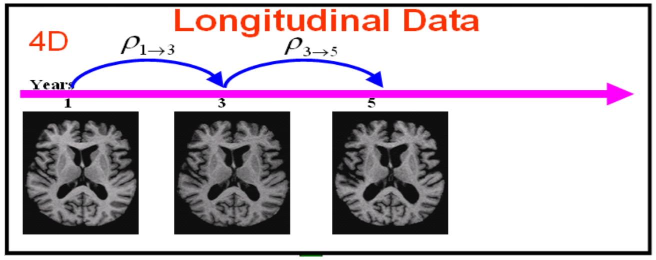
Anatomical MRI, DTI (HARDI)

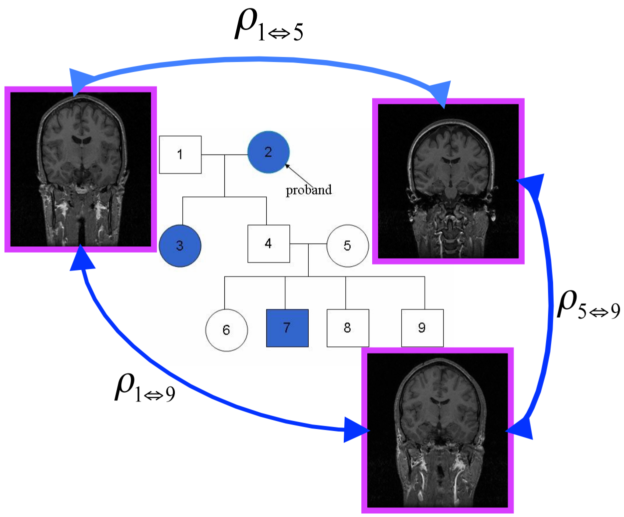




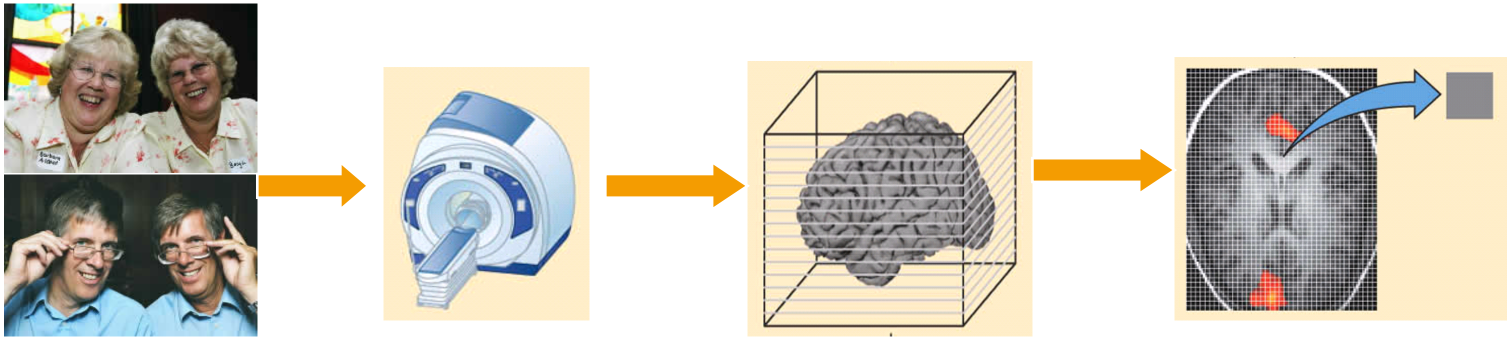
Motivation

Study Design





Neuroimaging Data



www.guysandstthomas.nhs.uk/.../T/Twins400.jpg

Motivation

Raw Images

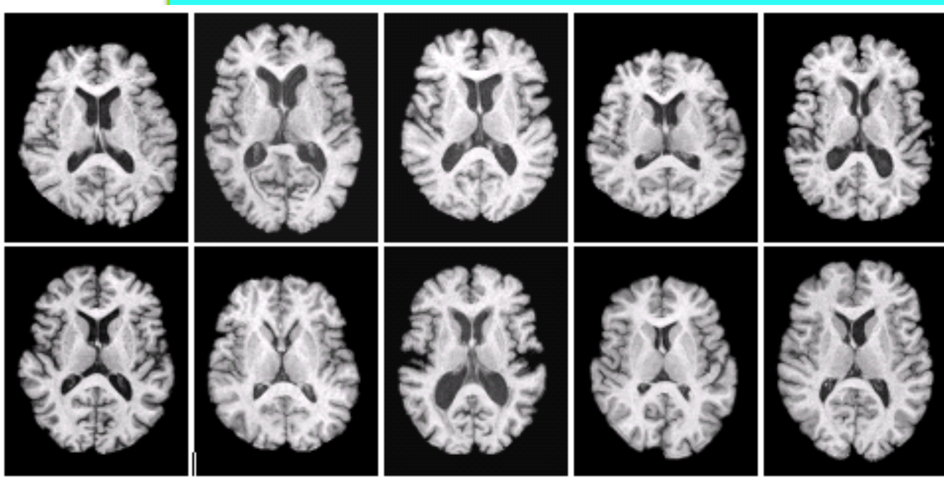
- Image Reconstruction
- Image Registration
- Imaging Smoothing

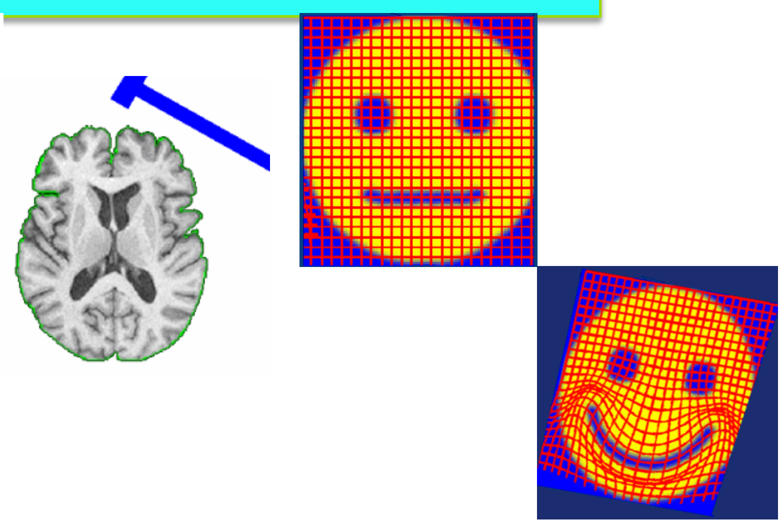
Statistical Analysis

Motivation

Image Registration

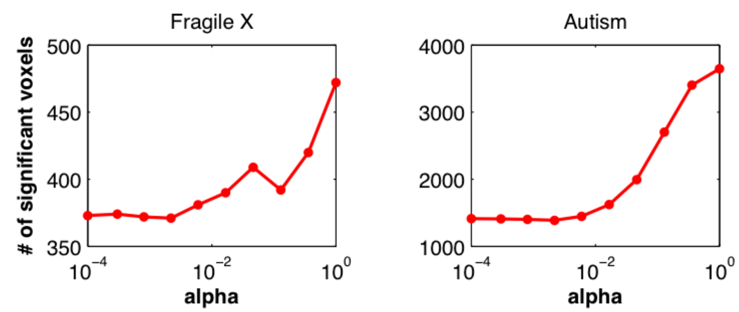
Image registration is the process of transforming different sets of data into *one coordinate system*. Data may be multiple photographs, data from different sensors, from different times, or from different viewpoints.



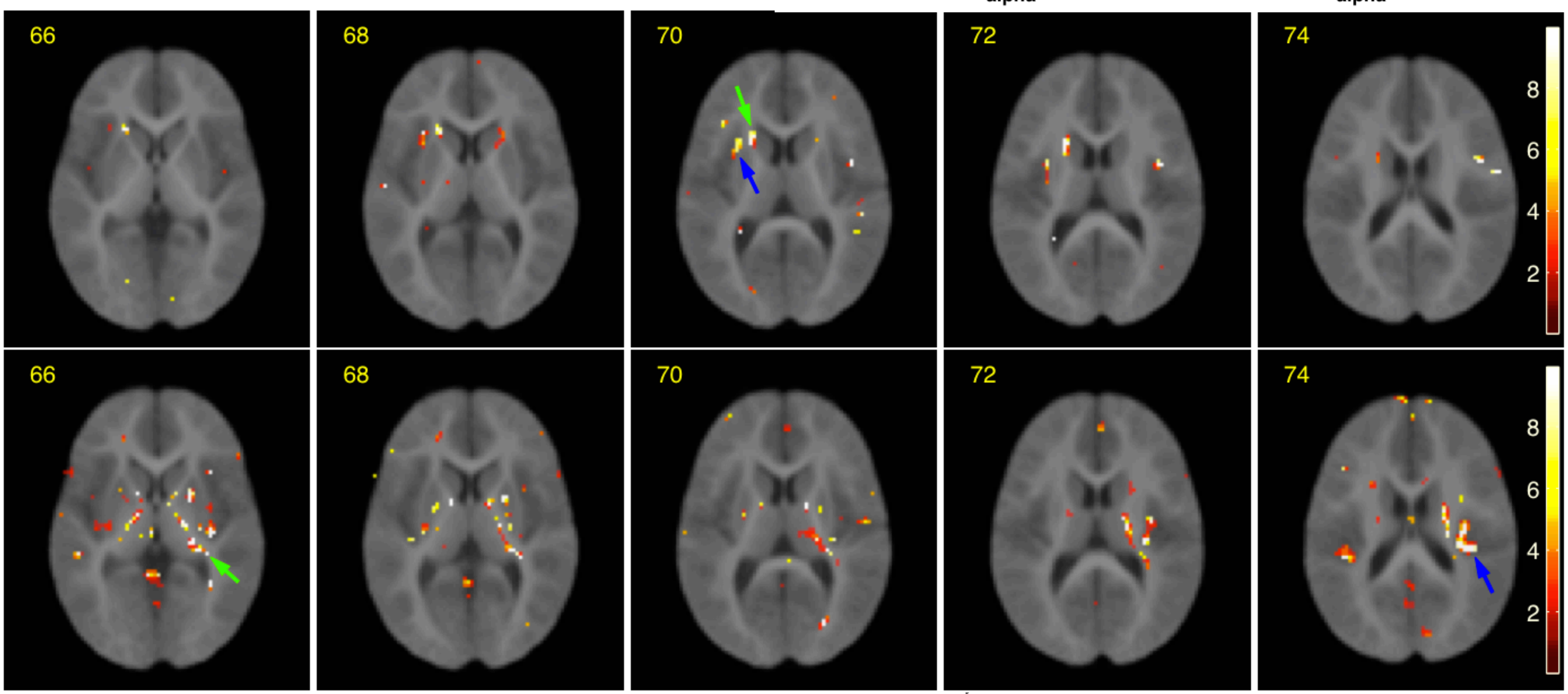


Motivation

25 normal control (NC), 19 fragile X syndrome (FX), and 47 autism (AU)



FX vs NC **AU vs NC**



Motivation

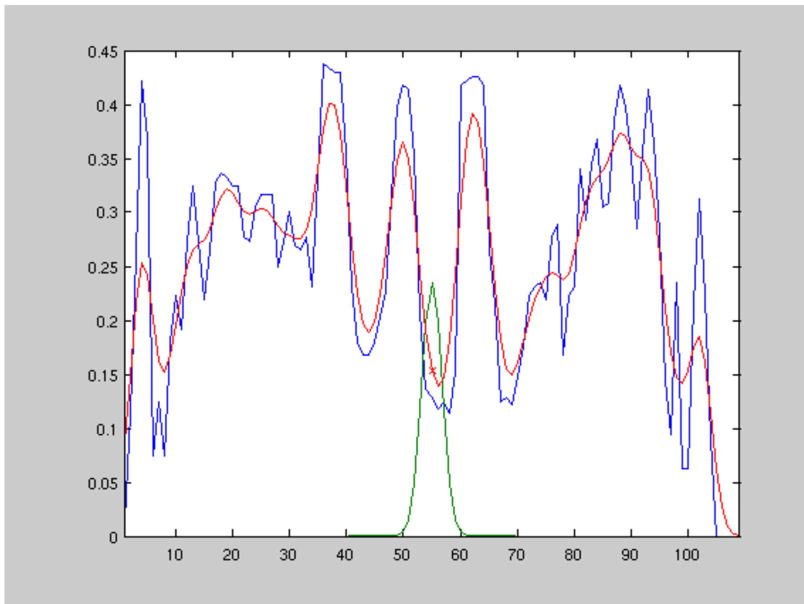
Image Smoothing

- ♦ Registration
- ♦ Signal-to-noise Ratio
- ♦ Gaussian

- How is it implemented?
  - ♦ Convolution with a 3D Gaussian kernel, of specified Full-width at half-maximum (FWHM) in mm

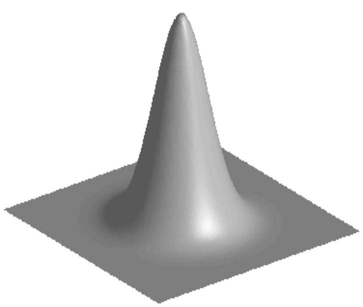
Motivation

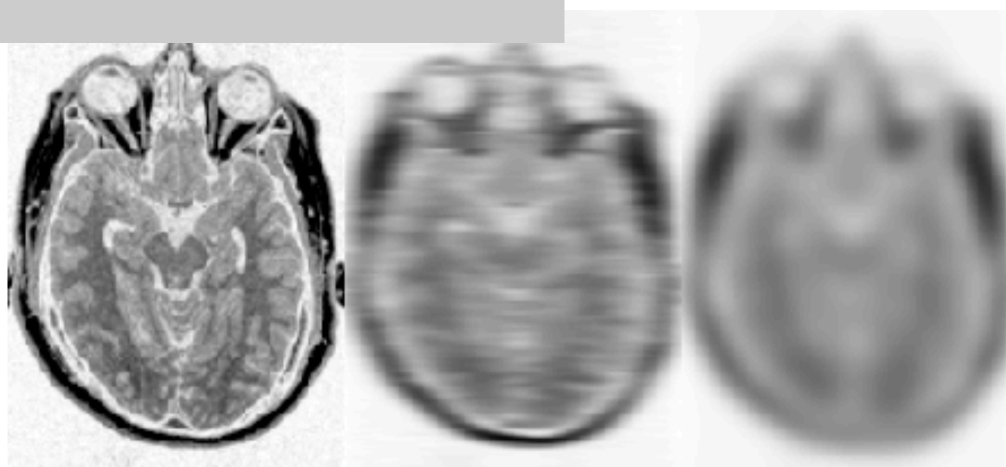
Example of Gaussian smoothing in one-dimension



The Gaussian kernel is **separable** we can smooth 2D data with two 1D convolutions.

SPM training course

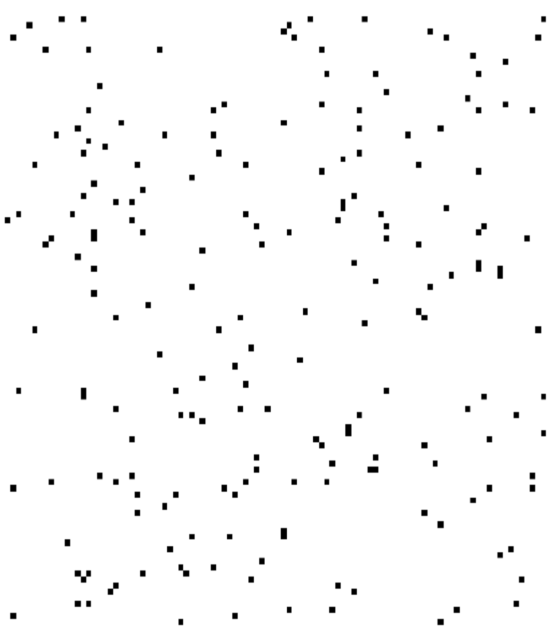




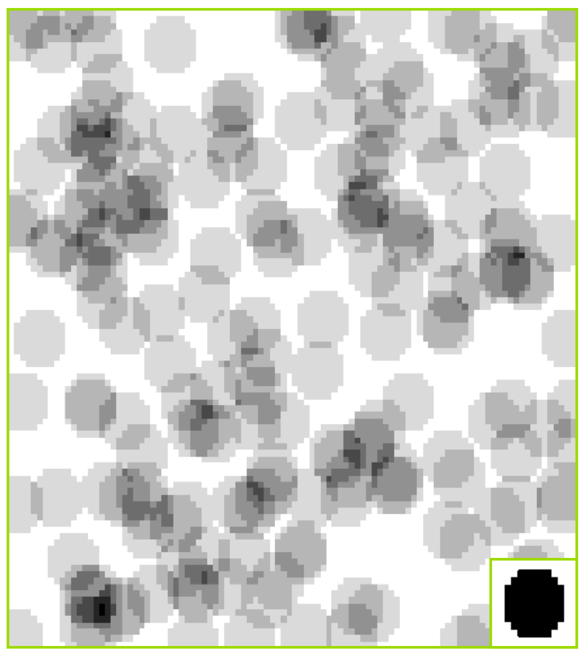
Motivation

Each voxel after smoothing effectively represents a weighted average over its local region of interest (ROI)

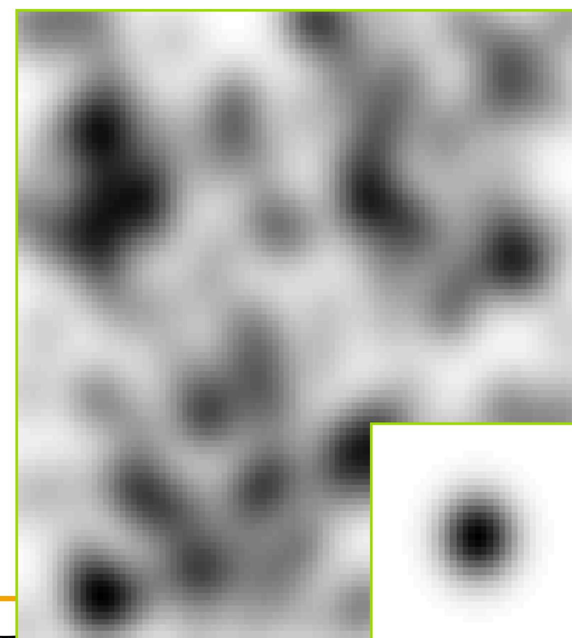
Before convolution



Convolved with a circle

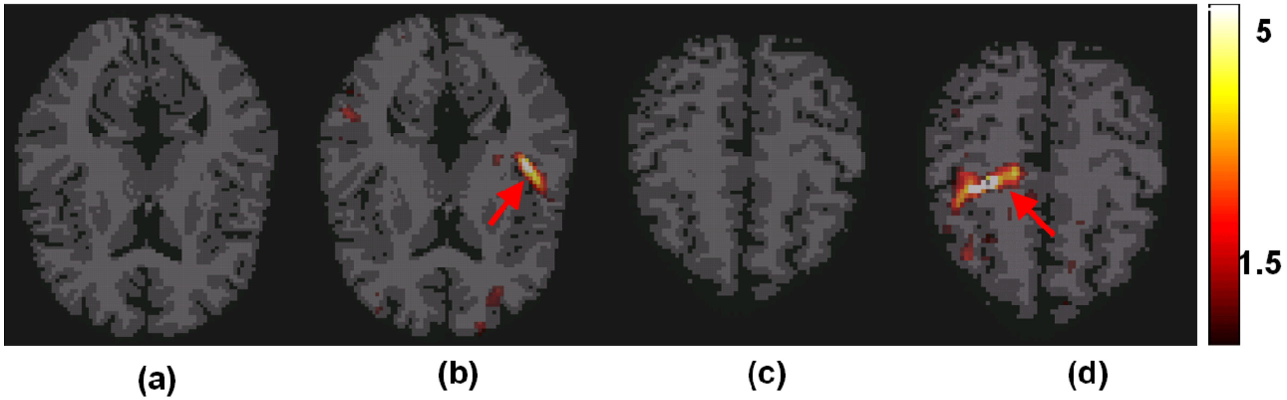


Gaussian convolution



Motivation

- Smoothing method is independent of data
- Degree of smoothness is arbitrary
- Effect of smoothness is profound
- The relationship between smoothing method and study design is unknown



Jones et al. (2006),
Yue et al. (2010)

Smoothing Methods

- **Two Strategies**
  - **Smoothing individual images independently**
    - **Enough measurements**
    - **Mean structure**
  - **Smoothing all images simultaneously**
    - **Independent of data**
    - **Data, design and model driven**

Smoothing Methods

What is image?

an exact replica of the contents of a storage device

an optically formed duplicate or other reproduction of an object

Google/wiki

Smoothing Methods

Mathematics.

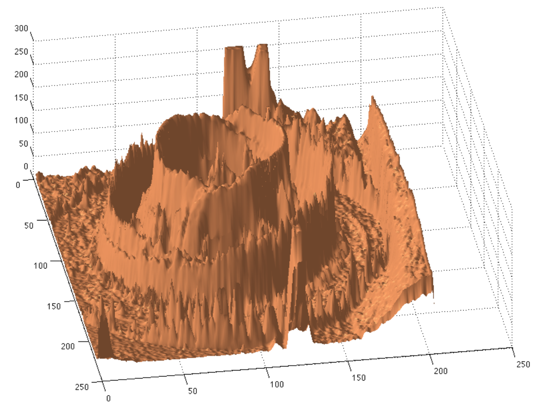


Image is the point or set of points in the range corresponding to a designated point in the domain of a given function.

▲ Ω is a compact set. $\tilde{x} \in \Omega \subseteq R^k$

➡ $f(\tilde{x}) \in M \subseteq R^m$ $\quad f : \Omega \rightarrow M \subseteq R^m$

★ $\int_{\Omega} \| f(\tilde{x}) \|^{k} \, d\tilde{x} < \infty$ for some k>0

Smoothing Methods

Digitized Images

$$f : \Omega_0 \rightarrow \{0,1,\cdots,M_0\}$$

- Sampling (grid points) $\Omega_0 \in \Omega$
- Sampling Rate
- Quantization

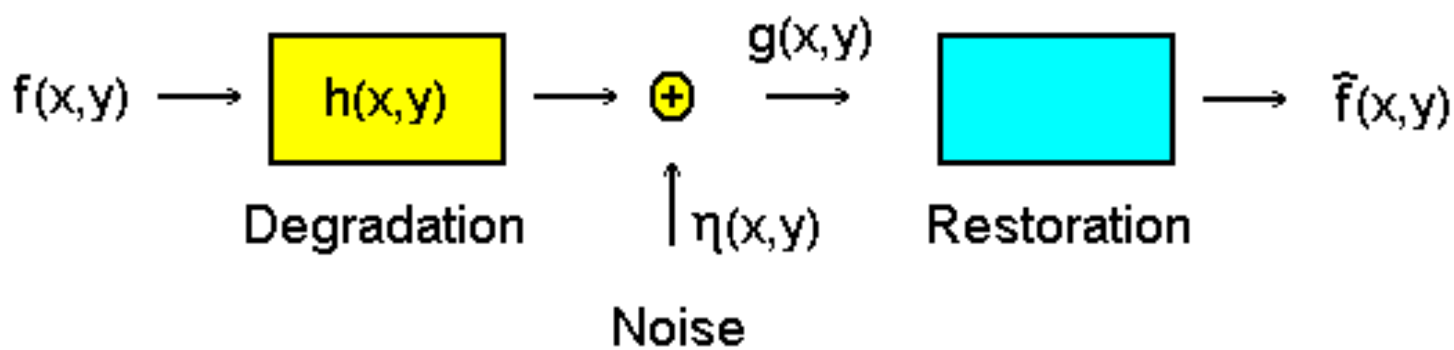
$$0,1,2,\cdots,2^m \text{ for } m = 5 \sim 12, \text{that is } M_0 = 2^{\mathrm{m}}$$

Smoothing Methods

Image Degradation/Restoration Process

The goal of image restoration is to improve a degrade image in some predefined sense. Schematically this process can be visualized as

$$g(x\ ,y) = h(f(x,y)) + \eta(x,y)$$



where f is the original image, g is a degraded/noisy version of the original image and $\tilde{f}$ is a restored version.

Smoothing Methods

Methods

- **Filters**

  Lowpass filtering, Wiener filters, Median filtering
- **Wavelet Shrinkage Denoising**

  Soft and Hard thresholding
- **Variational Denoising based on Bounded Variation Models**
- **Nonlinear Diffusion and Scale-space theory**
- **Bayesian Models**

  Markov random field

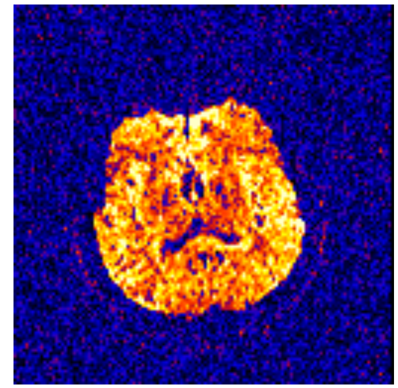
Issues:

- **Noise Distribution**
- **Window Size**
- **Localization**
- **Tuning Parameters**

Chen and Shen (2005)

Smoothing Methods

- An image may be 'dirty' with dots, speckles, stains
- Noise removal
  - Dots can be modeled as impulses (salt-and-pepper or speckle) or continuously varying (Gaussian noise)
  - Low-pass filtering
- Problem with low-pass filtering
  - May blur edges
  - Adaptive, edge preserving



Input:

100	100	100	100	100
100	200	205	203	100
100	195	200	200	100
100	200	205	195	100
100	100	100	100	100

$\frac{1}{9}\times$

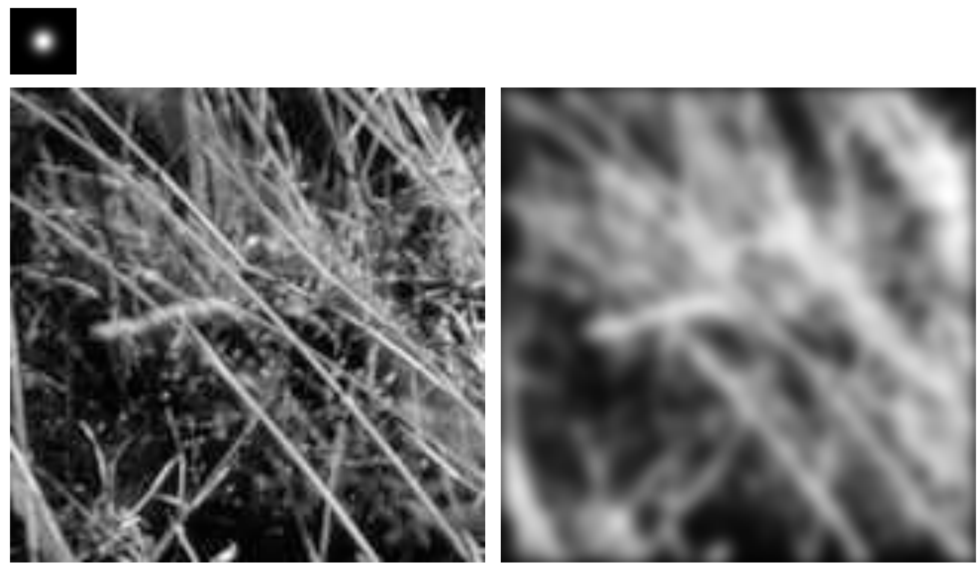
1	1	1
1	1	1
1	1	1

100	100	100	100	100
100	144	167	145	100
100	167	200	168	100
100	144	166	144	100
100	100	100	100	100

$\frac{1}{16}\times$

1	2	1
2	4	2
1	2	1

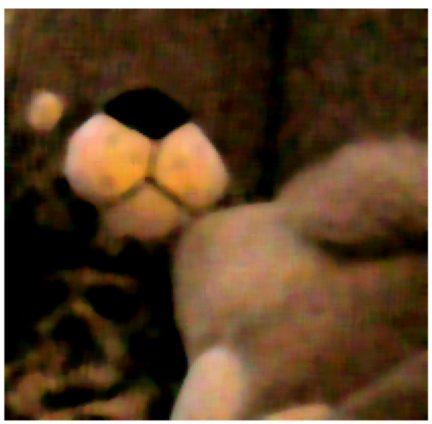
100	100	100	100	100
100	156	176	158	100
100	174	201	175	100
100	156	175	156	100
100	100	100	100	100



Smoothing Methods


original image


1px median filter


3px median filter


10px median filter

Bandwidth Selection

Different Smoothing Methods

Smoothing Methods

Location Adaptation

Different Smoothing Methods

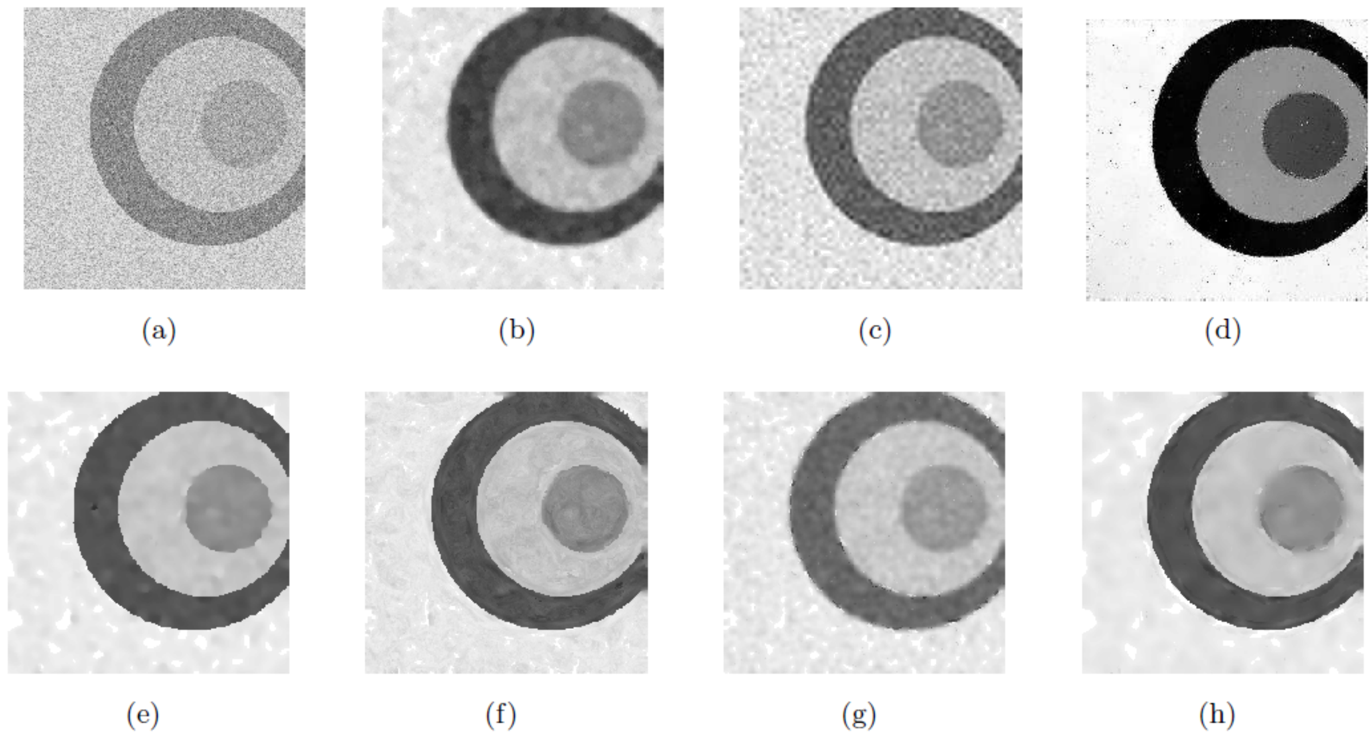


Figure Plot (a): The noisy test image. Plots (b)-(h): the reconstructed images by the local median smoothing procedure, the DWT procedure, the MRF procedure, the AWS procedure, and procedures (6)-(8), respectively.

Qiu (2005)

Smoothing Methods

Feature

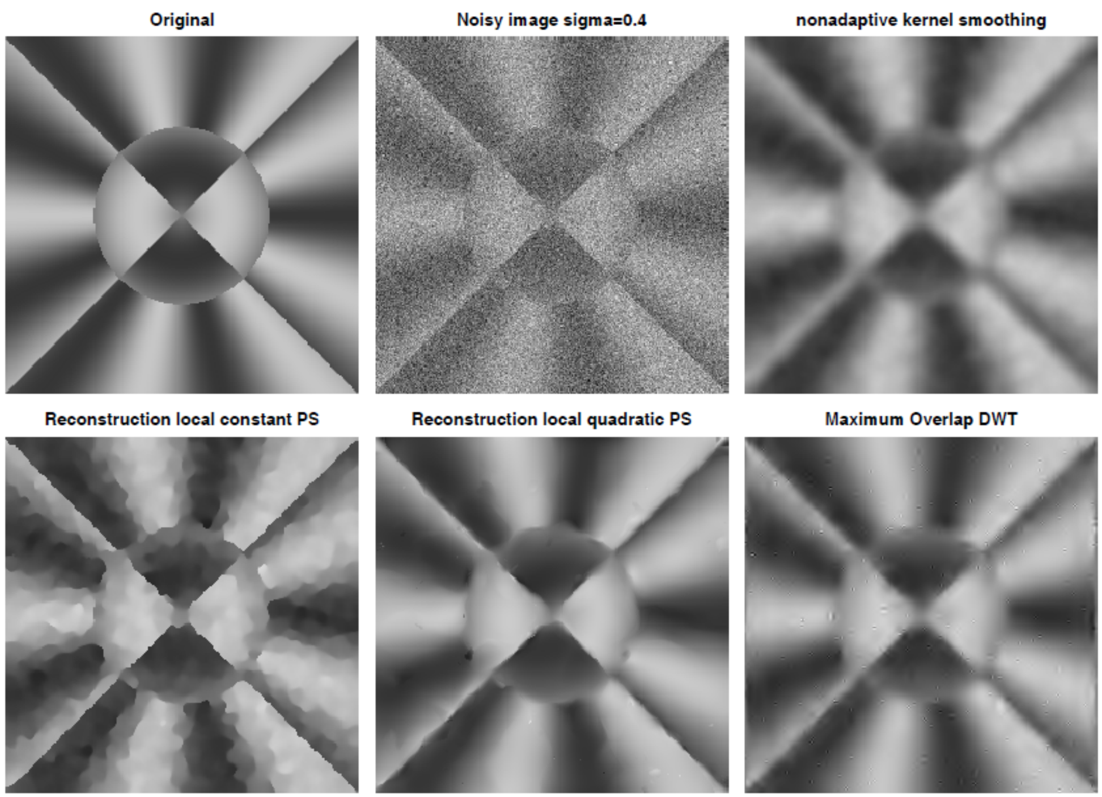
- **Local**
  - Spatial distance
- **Nonlocal**
  - Signal difference

Estimation

- **Pointwise**
  - Single point
- **Multipoint**
  - a set of points

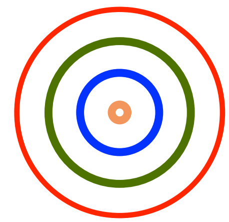
Smoothing Methods

Propogation-Seperation Method



Features

- Increasing Bandwidth



- Adaptive Weights
- Adaptive Estimates

Smoothing Methods

Propogation-Seperation Theory

- **Exponential Family** $Y(d) \sim EF(\theta(d))$

$$\hat{\theta}(d) = \operatorname{argmax}_{\theta(d)} \sum_{d'} w(d,d')\ell(Y(d'),\theta(d)) = \operatorname{argmax}_{\theta(d)} \ell(\sum_{d'} \tilde{w}(d,d')Y(d'),\theta(d))$$

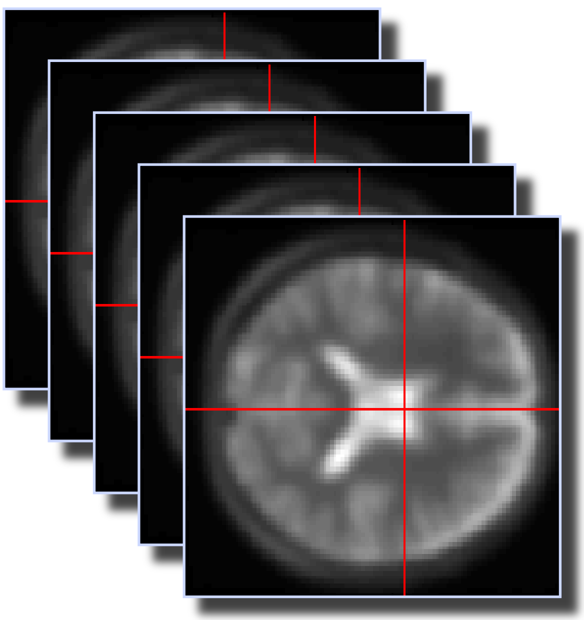
<u>Smoothing Imaging Intensities</u>

- **Under strong conditions,** **Katkovnik,V and Spkoiny, V. (2008)**
  J. Polzehl and V. Spokoiny, (2005)

$$P(K(\hat{\theta}(d),\theta(d))^{1/2} > C(\log(N(D)) / N(D))^{1/(2+c)}) \to 0$$

Smoothing Methods

Images from Multiple Subjects
Multiple Images from a single subject



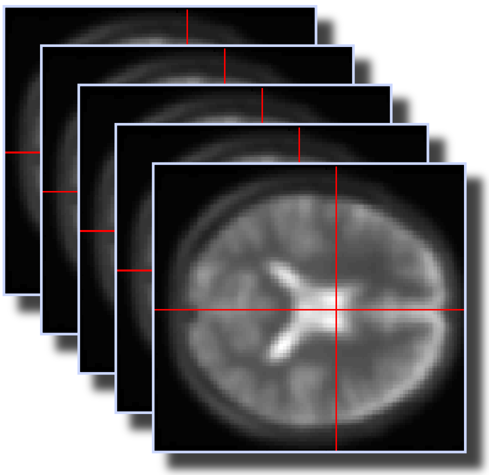
Tabelow et al. (2006, 2008a, 2008b), Polzehl, et al. (2010)

- **Denoising fMRI, DTI**
- **SPM**

$$Y_i(d) = x_i^T \beta(d) + \varepsilon_i(d)\sigma(d)_i$$

Multiscale Adaptive Smoothing Models

Images from Multiple Subjects

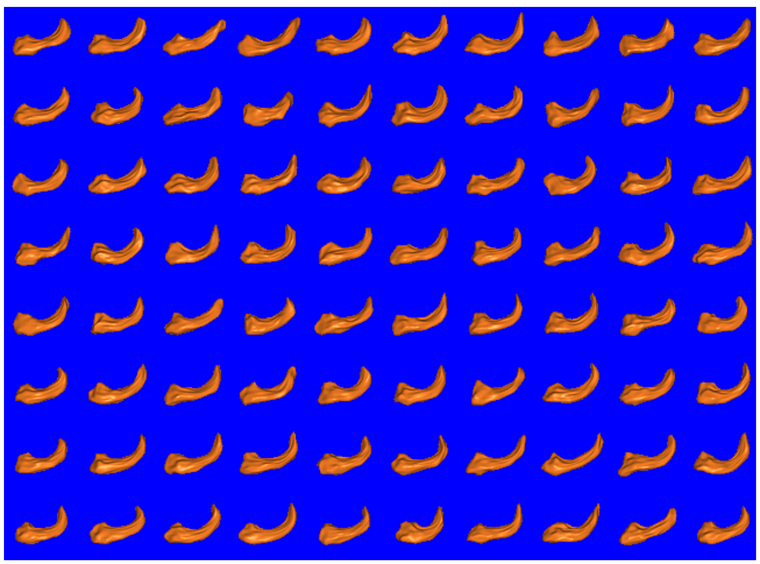


Questions of Interest

- **Complex design**
  Longitudinal, Twin, and family studies
- **Models with parametric and/or nonparametric components**
- **Consistency results**
- **Standard deviation images**
- **Testing theory**

Multiscale Adaptive Smoothing Models

My goal is to develop a class of MASMs with necessary statistical properties for imaging data collected from cross-sectional, longitudinal, twin, and familial studies.



M

$$= g(x, \theta(d), f(d)) \oplus \varepsilon$$

$$x \in R^k, \theta(d) \in \Theta \subset R^p, f(d) \in F$$

$$g: R^k \times R^p \times F \rightarrow M$$

Problems of interest:

$$\{(\theta(d), f(d)) : d \in D\}$$

$$\{\varepsilon(d) : d \in D\}$$

MARM
Multiscale Adaptive Regression Models

- Integrate Parametric Models with PS
- Standard Deviation Image
- Consistency and Asymptotic Distribution

Multiscale Adaptive Regression Model

D: 3D volume or 2D surface.

N_D: the number of points on *D*

d: a voxel in *D*

$\{\mathbf{Y}_{i,\mathcal{D}} : i = 1, \cdots, n\}$: data on *D* for the i^{th} subject

X_i : a $k \times 1$ covariates

$\boldsymbol{\theta}(d)$: unknown parameter

Multiscale Adaptive Regression Model

Voxel-wise Approach

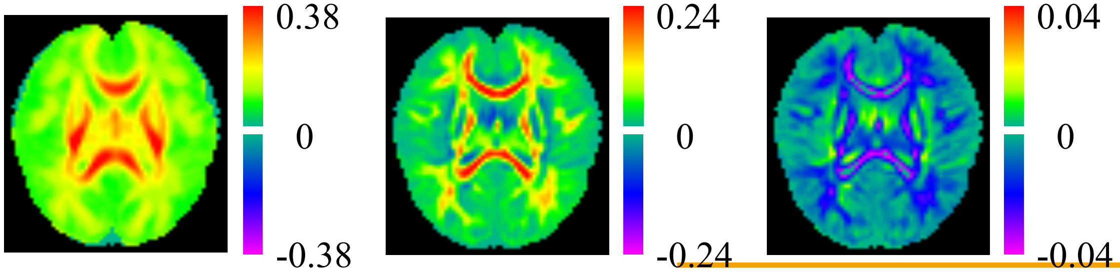
$$p(\mathbf{Y}_{i,\mathcal{D}}|\mathbf{X}_i) = \prod_{d\in\mathcal{D}} p(Y_i(d)|\mathbf{x}_i, \boldsymbol{\theta}(d)),$$

MARM

Being Spatial

$$p(\mathbf{Y}_{i,\mathcal{D}}|\mathbf{X}_i) \approx \prod_{D_k} p(\{Y_i(d') : \quad d' \in D_k\}|\mathbf{x}_i)$$

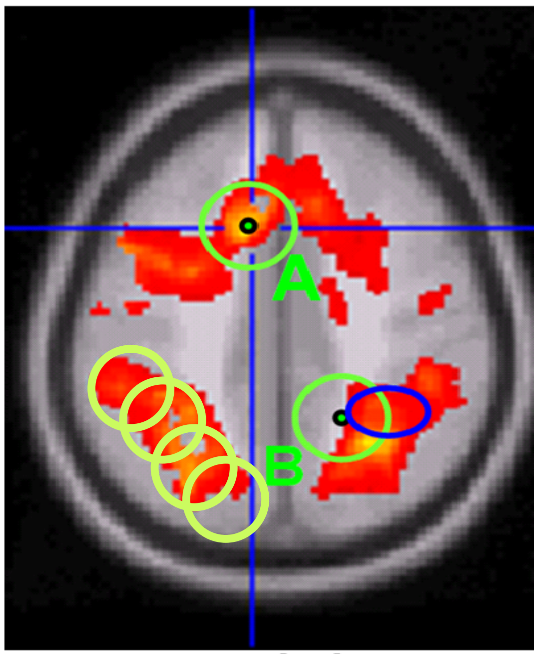
D_k **denotes the set of all voxels in a homogeneous region**



Multiscale Adaptive Regression Model

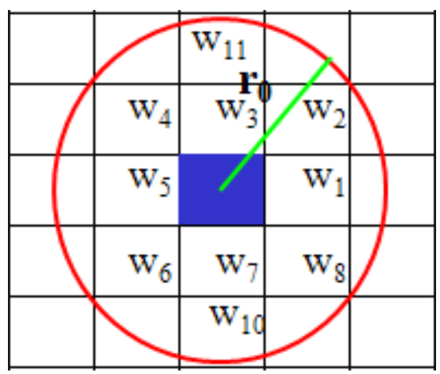
Identifying homogeneous regions

D_k



Drawing a sphere with radius r0 at each voxel

Calculating the similarities between the current voxel and its neighboring voxels.



Multiscale Adaptive Regression Model

Model Specification

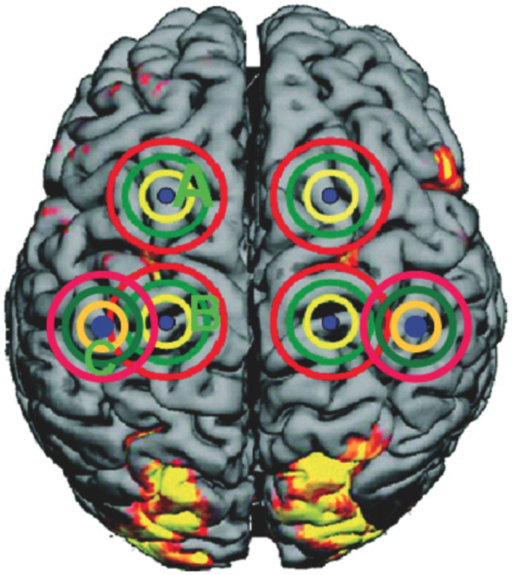
$$p(Y_i \mid x_i) = \prod_{d \in D} p(\{Y_i(d') : d' \in B(d, r_0)\} \mid x_i, \theta(d))$$

$$\ell(\{Y_i(d') : d' \in B(d, r_0)\} \mid x_i) = \sum_{d' \in B(d, r_0)} w(d, d'; r_0)\, \ell(Y_i(d') \mid x_i, \theta(d))$$

$\omega(d, d'; r_0)$ is a **weight** function for characterizing the **similarity** between the data in voxels d and d'.

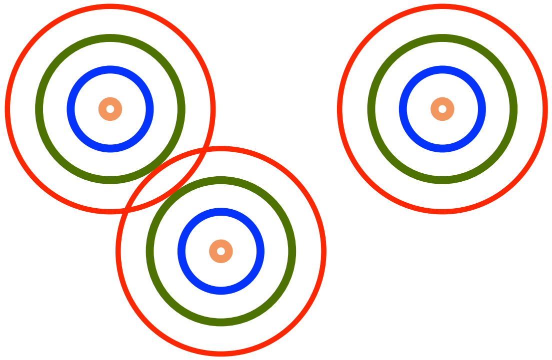
Multiscale Adaptive Regression Model

Being Hierarchical



Drawing nested spheres with increasing radiuses at each voxel

$$h_0 = 0 < h_1 < \cdots < h_S = r_0$$

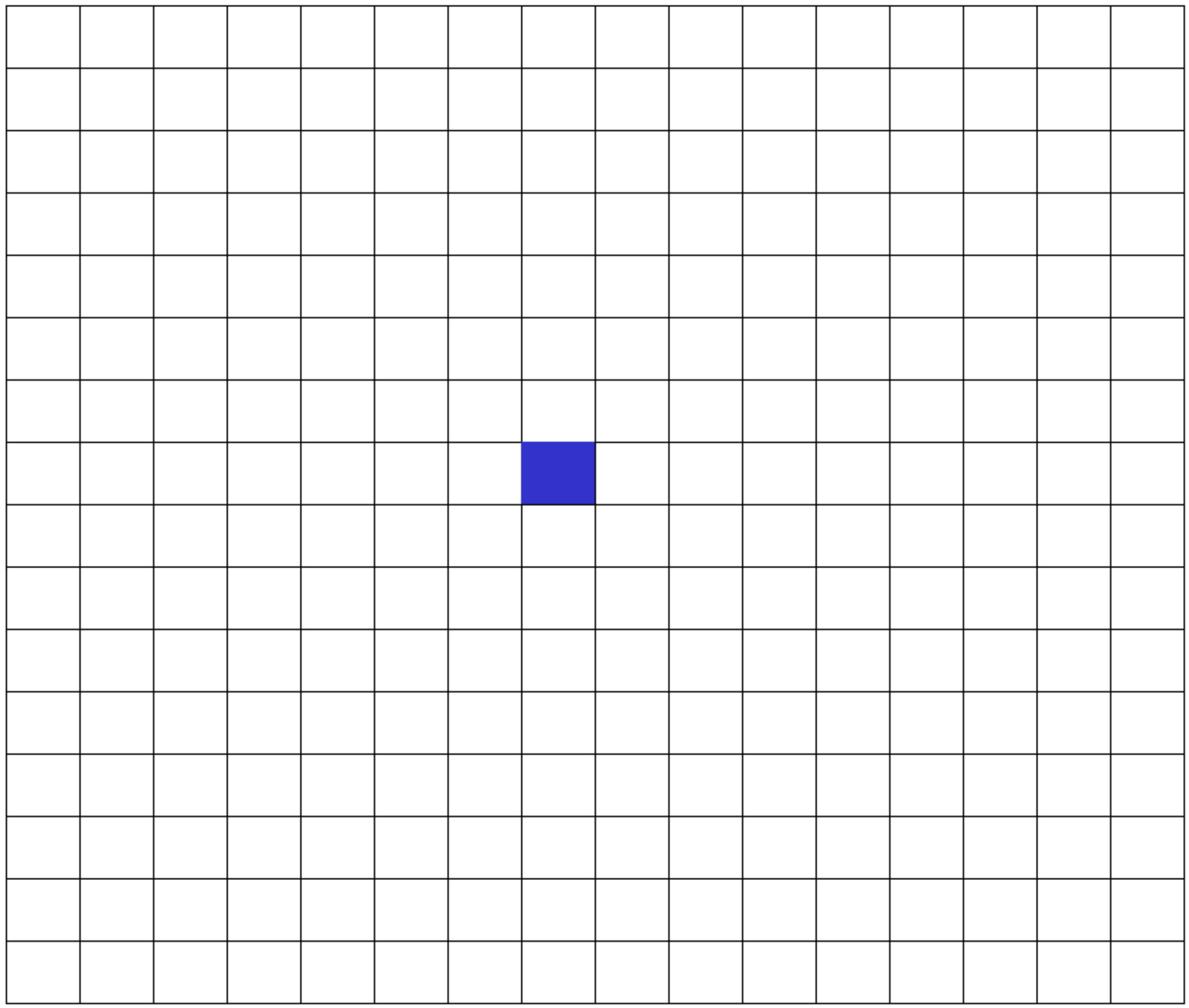


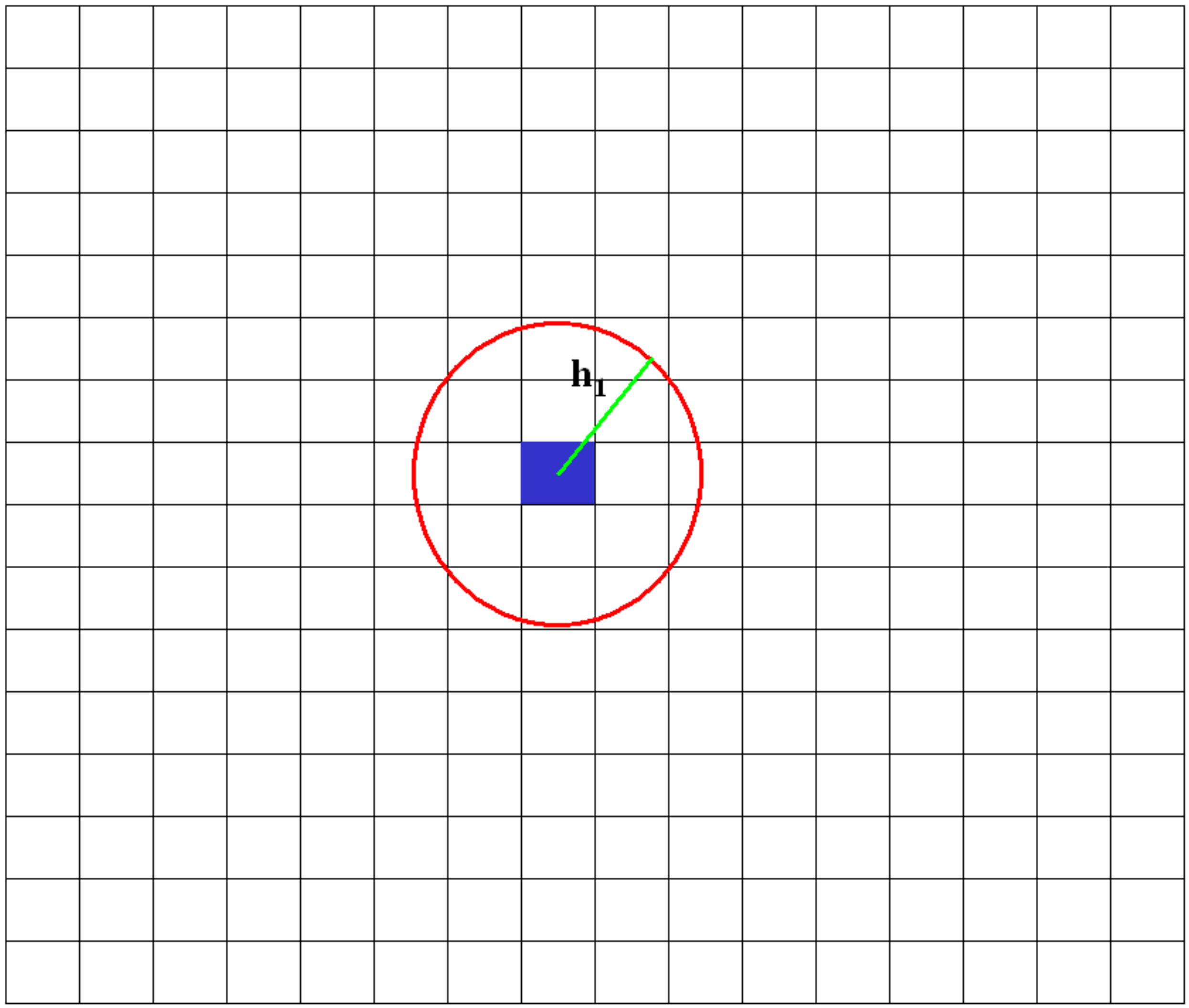
Multiscale Adaptive Regression Model

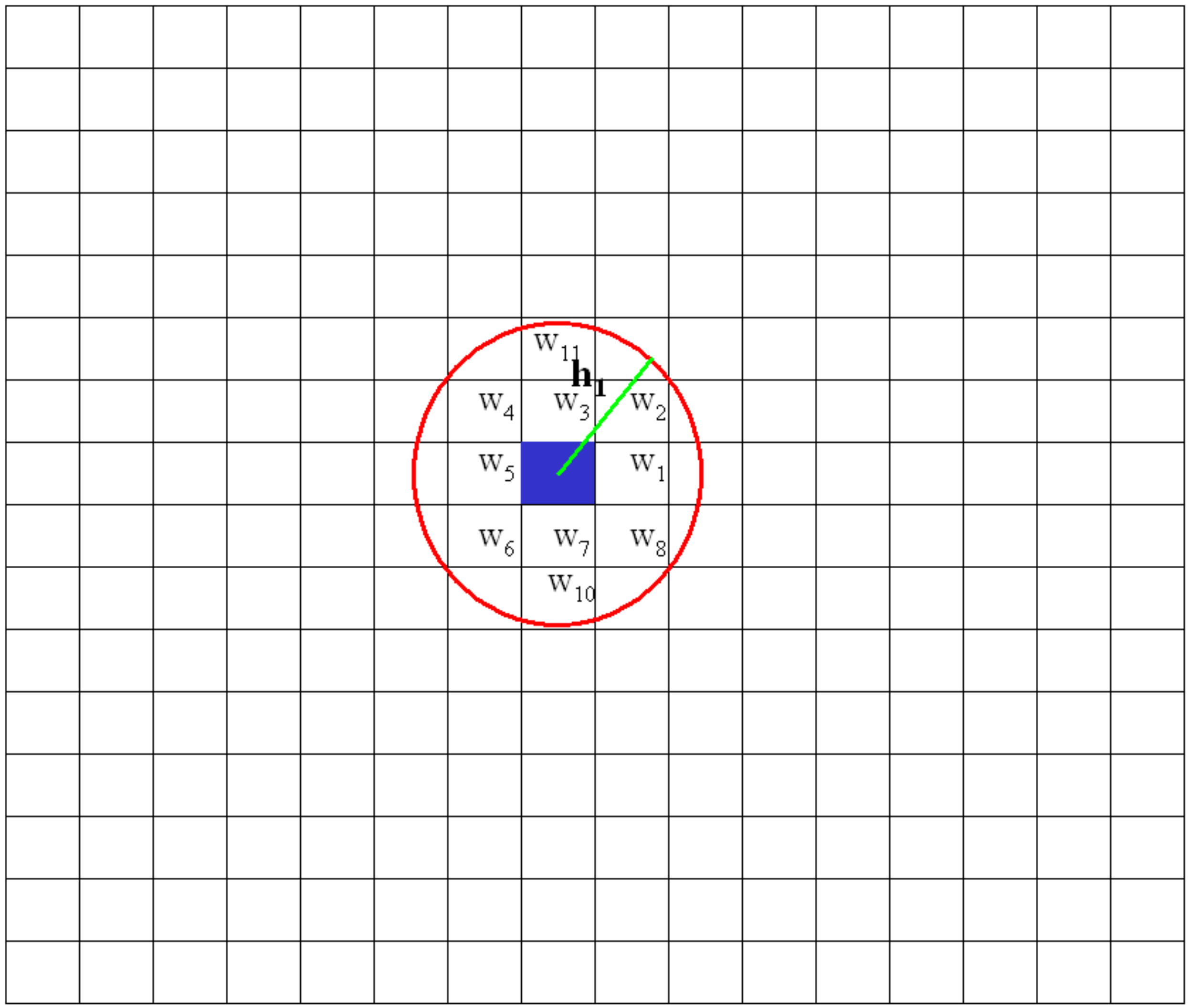
Being Adaptive

Sequentially determine $\omega(d, d'; h)$ **and adaptively updat** $\hat{\boldsymbol{\theta}}(d, h)$

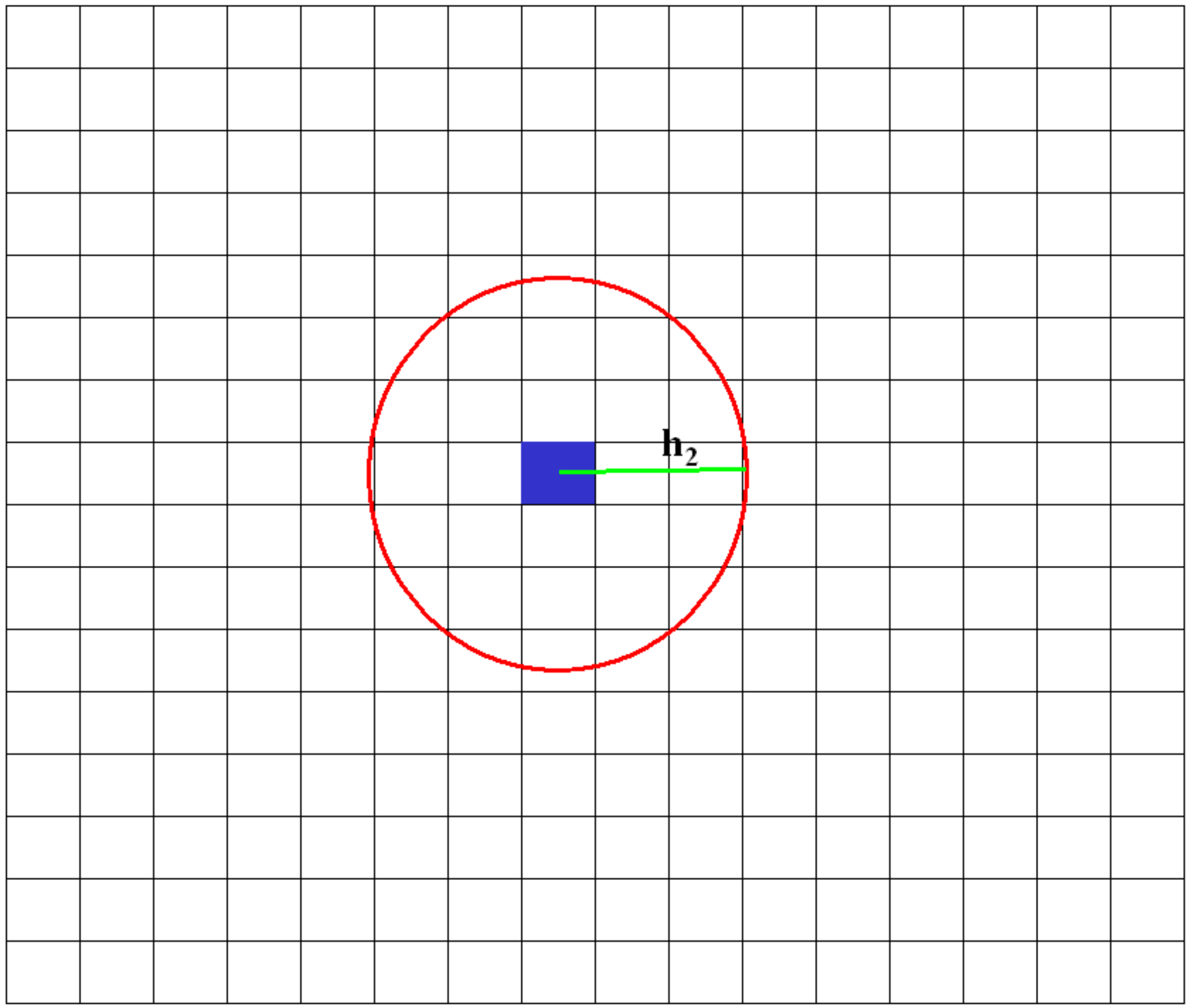
$$
\begin{array}{ccccccc}
\omega(d, d'; h_0) & & \omega(d, d'; h_1) & & \cdots & & \omega(d, d'; h_S = r_0) \\
\Downarrow & \nearrow & \Downarrow & \nearrow & \cdots & \nearrow & \Downarrow \\
\hat{\boldsymbol{\theta}}(d; h_0) & & \hat{\boldsymbol{\theta}}(d; h_1) & & \cdots & & \hat{\boldsymbol{\theta}}(d; h_S) \\
\Downarrow & & \Downarrow & & \cdots & & \Downarrow \\
W_\mu(d; h_0) & & W_\mu(d; h_1) & & \cdots & & W_\mu(d; h_S)
\end{array}
$$

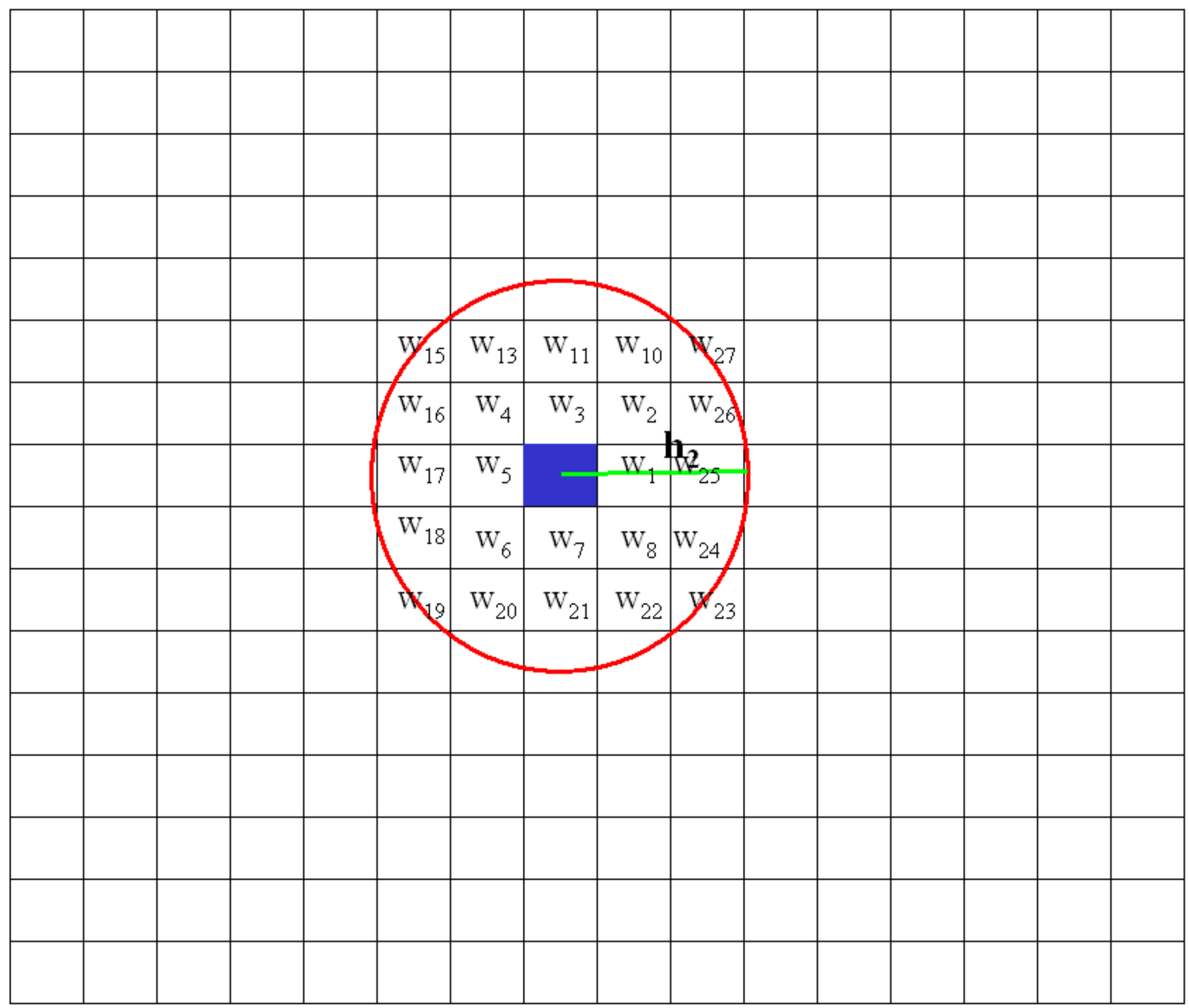


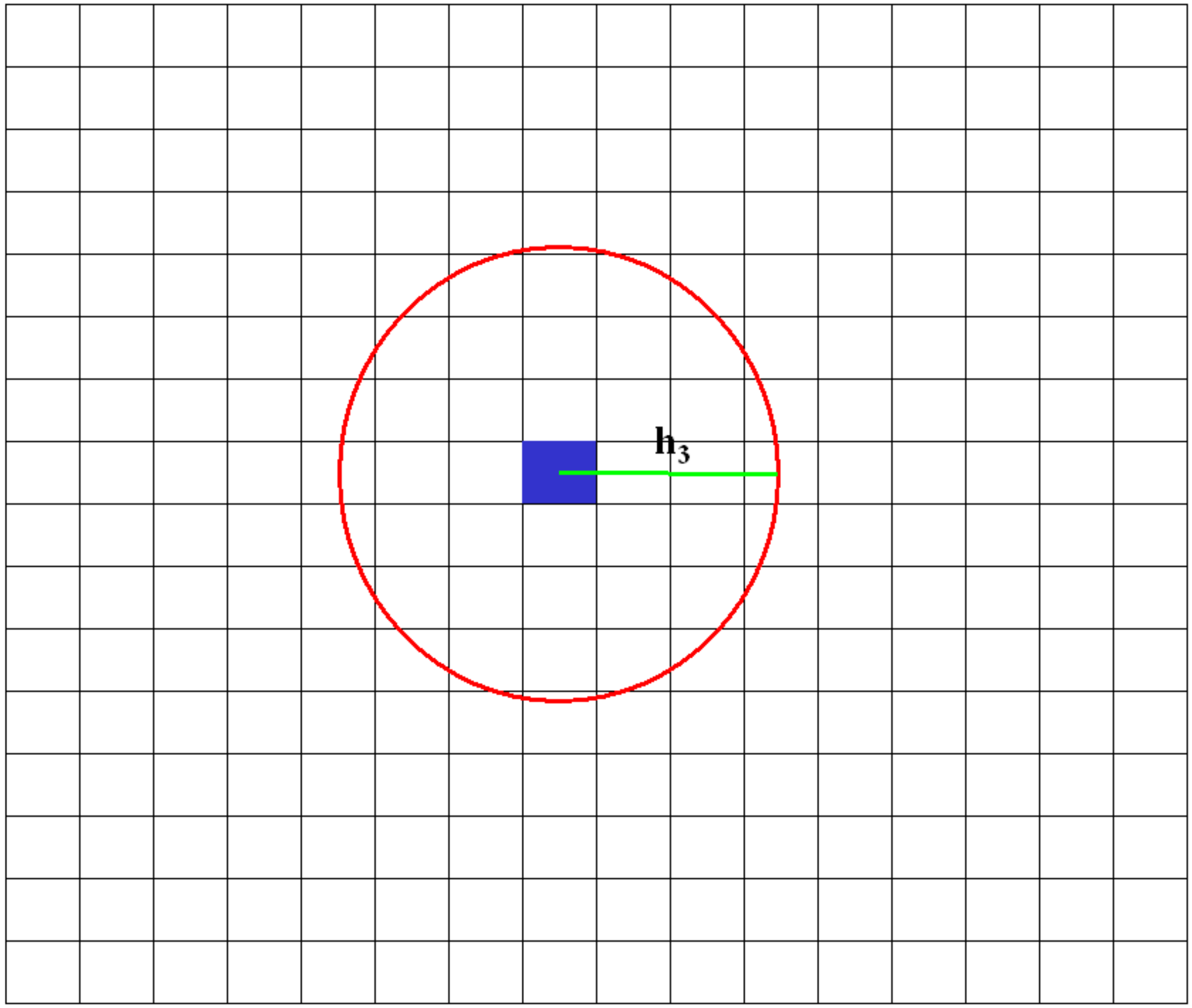

h1

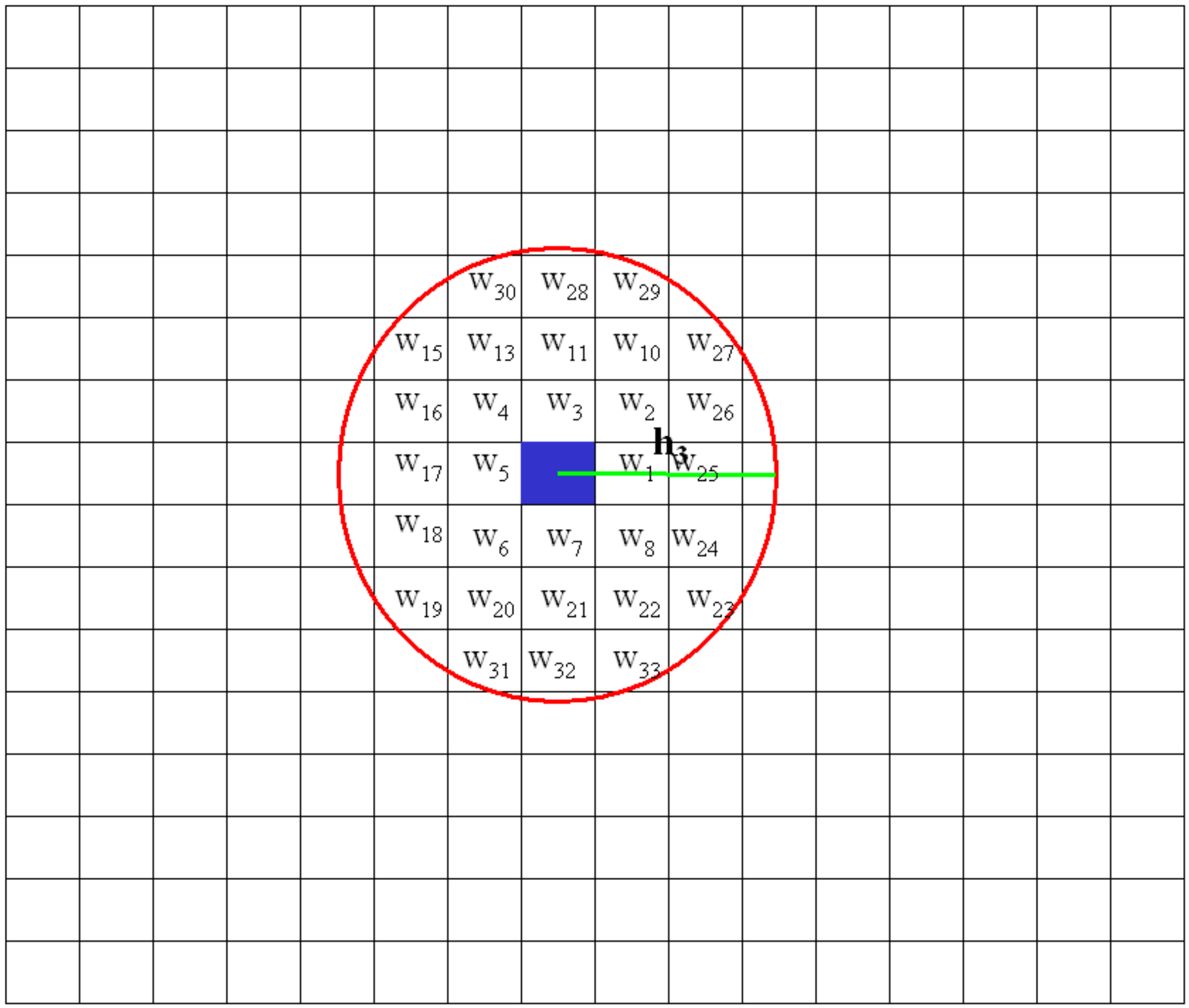


W11
h1
W4
W3
W2
W5
W1
W6
W7
W8
W10

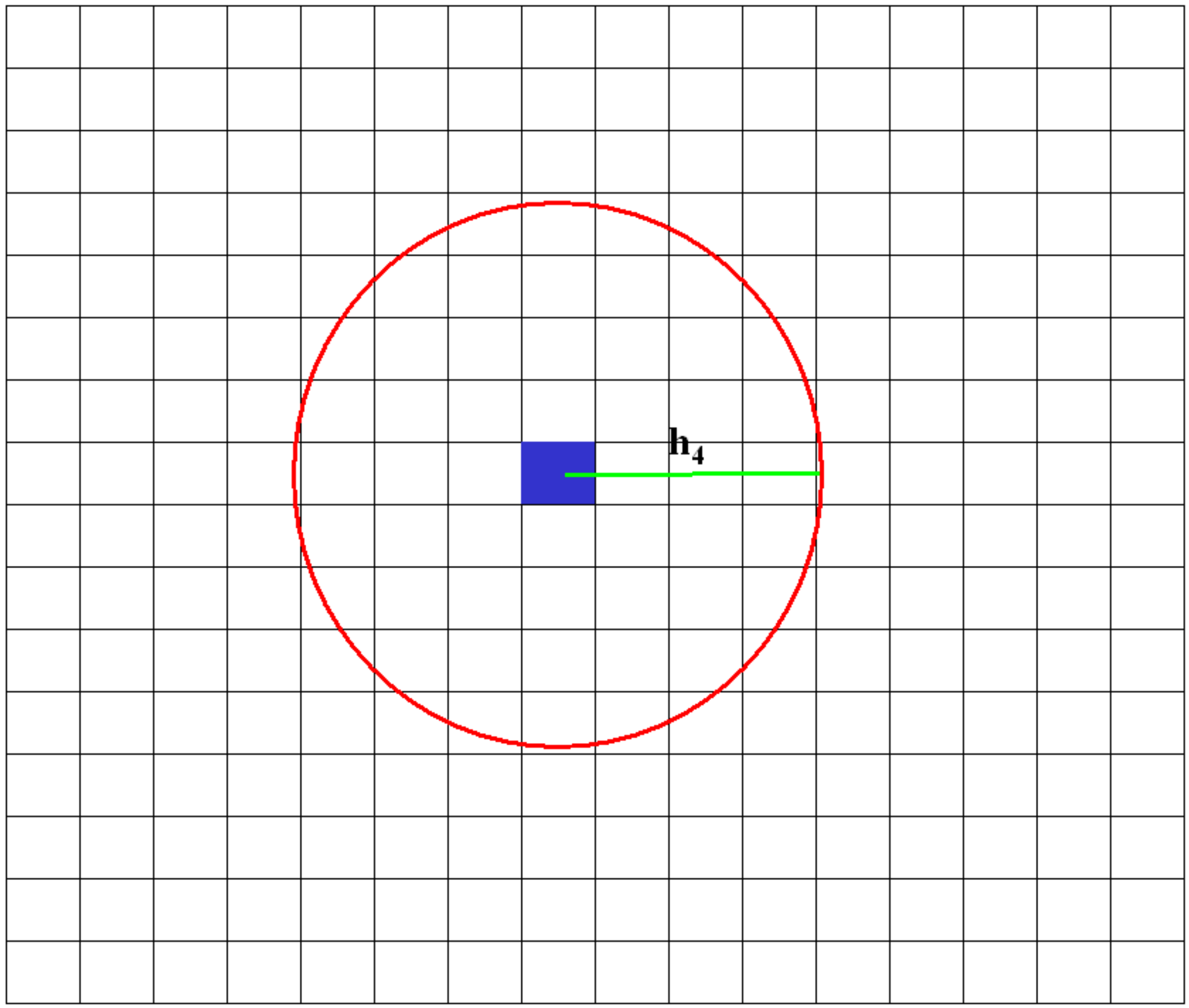

h_2


W_{15} W_{13} W_{11} W_{10} W_{27}
W_{16} W_{4} W_{3} W_{2} W_{26}
W_{17} W_{5} W_{1} W_{25}
h_2
W_{18} W_{6} W_{7} W_{8} W_{24}
W_{19} W_{20} W_{21} W_{22} W_{23}


h_3



w_{30} w_{28} w_{29}
w_{15} w_{13} w_{11} w_{10} w_{27}
w_{16} w_4 w_3 w_2 w_{26}
h_3
w_{17} w_5 w_1 w_{25}
w_{18} w_6 w_7 w_8 w_{24}
w_{19} w_{20} w_{21} w_{22} w_{23}
w_{31} w_{32} w_{33}


h_4

Multiscale Adaptive Regression Model

MARM/PS

- Learning Voxel Feature
- Local Feature Adaptation
- Adaptive Estimation and Testing
- Automatic Stop

Multiscale Adaptive Regression Model

Local Feature Adaptation

- For any radius $h_s > h_0$, define

$$\omega(d,d';h_s) = K_{loc}(||d-d'||_2/h_s)K_{st}(D_{\theta}(d,d';h_{s-1})/C_n)$$

- $K_{loc}(u)$ and $K_{st}(u)$ are two decreasing kernel functions
- Smoothing kernel: $K_{loc}(u) = (1-u^2)_+$
- Similarity kernel: $K_{st}(u) = \exp(-u)\mathbf{1}\left(u \leq \frac{s+2}{s(\log s+2)}\right)$
- Dissimilarity measure:

$$D_{\theta}(d,d';h_{s-1}) = [\hat{\theta}(d;h_{s-1}) - \hat{\theta}(d';h_{s-1})]^T \hat{\Sigma}(\hat{\theta}(d;h_{s-1}))^{-1}[\hat{\theta}(d;h_{s-1}) - \hat{\theta}(d';h_{s-1})].$$

Multiscale Adaptive Regression Model

Adaptive Estimation and Testing

Weighted quasi-likelihood

$$\ell_n(\boldsymbol{\theta}(d); h, \tilde{\boldsymbol{\omega}}) = \sum_{i=1}^{n} \sum_{d' \in B(d,h)} \tilde{\omega}(d, d'; h) \log p(Y_i(d') | \mathbf{x}_i, \boldsymbol{\theta}(d))$$

MWQLE

$$\hat{\boldsymbol{\theta}}(d, h) = \text{argmax}_{\boldsymbol{\theta}(d)} n^{-1} \ell_n(\boldsymbol{\theta}(d); h, \tilde{\boldsymbol{\omega}})$$

Newton-Raphson Algorithm

$$\hat{\boldsymbol{\theta}}(d,h)^{(t+1)} = \hat{\boldsymbol{\theta}}(d,h)^{(t)} + \{-\partial^2_{\boldsymbol{\theta}(d)} \ell_n(\hat{\boldsymbol{\theta}}(d,h)^{(t)}; h, \tilde{\boldsymbol{\omega}})\}^{-1} \partial_{\boldsymbol{\theta}(d)} \ell_n(\hat{\boldsymbol{\theta}}(d,h)^{(t)}; h, \tilde{\boldsymbol{\omega}})$$

Expectation-Maximization Algorithm

Multiscale Adaptive Regression Model

Adaptive Estimation and Testing

Sandwich Estimator

$$\mathrm{Cov}[\hat{\boldsymbol{\theta}}(d,h)] \approx \Sigma_n(\hat{\boldsymbol{\theta}}(d,h)) = [\Sigma_{n,1}(\hat{\boldsymbol{\theta}}(d,h))]^{-1}\Sigma_{n,2}(\hat{\boldsymbol{\theta}}(d,h))[\Sigma_{n,1}(\hat{\boldsymbol{\theta}}(d,h))]^{-1}$$

$$\Sigma_{n,1}(\boldsymbol{\theta}(d)) = -\partial^2_{\boldsymbol{\theta}(d)}\ell_n(\boldsymbol{\theta}(d); h, \tilde{\boldsymbol{\omega}}) \text{ and}$$

$$\Sigma_{n,2}(\boldsymbol{\theta}(d)) = \sum_{i=1}^{n}[\sum_{d' \in B(d,h)} \tilde{\omega}(d,d';h)\partial_{\boldsymbol{\theta}(d)} \log p(Y_i(d')|\mathbf{x}_i, \boldsymbol{\theta}(d))]^{\otimes 2}$$

Wald Test Statistic

$$[R(\hat{\boldsymbol{\theta}}(d;h)) - \mathbf{b}_0]^T[\partial_{\boldsymbol{\theta}(d)}R(\hat{\boldsymbol{\theta}}(d;h))\hat{\Sigma}_n(\hat{\boldsymbol{\theta}}(d;h))\partial_{\boldsymbol{\theta}(d)}R(\hat{\boldsymbol{\theta}}(d;h))^T]^{-1}[R(\hat{\boldsymbol{\theta}}(d;h)) - \mathbf{b}_0]$$

Multiscale Adaptive Regression Model

log(Voxel size)<<Cn << sample size

Kernel functions

Conditions for M-estimators hold uniformly

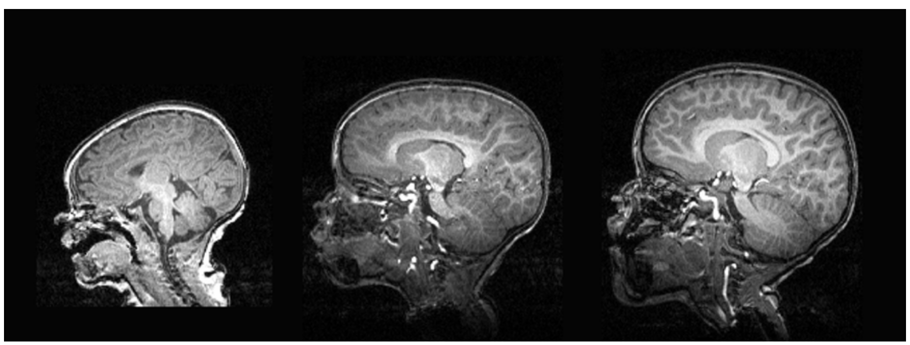
Weak Consistency

Asymptotical Normality

Asymptotically Chi-squared distribution

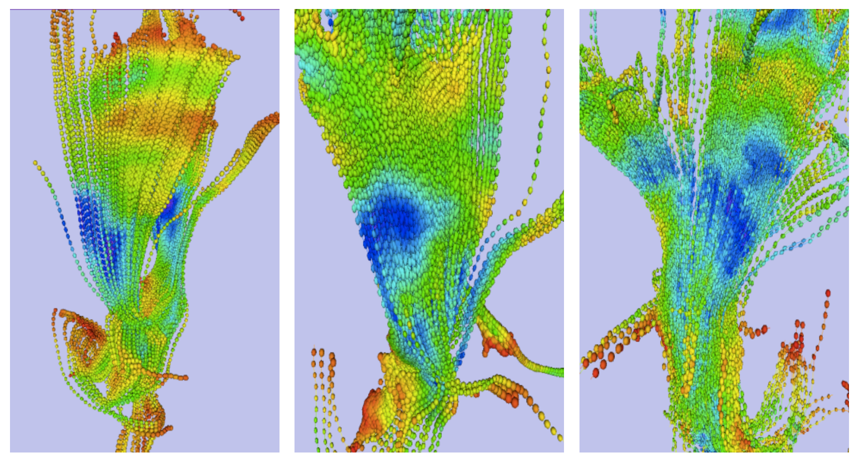
Real Data

- **Early Brain Development Project**
- **Objective:** **We want to assess the brain structural connectivity change in the early brain development.**
- **Subject:** **250 infants.**
- **PI:** **John Gilmore**
- **MRIs:** **DWI, resting fMRI, and T1 MRI were acquired for each subject at 2 weeks, 1, 2, 3, 4 years old.**



2 weeks 1 year 2 year

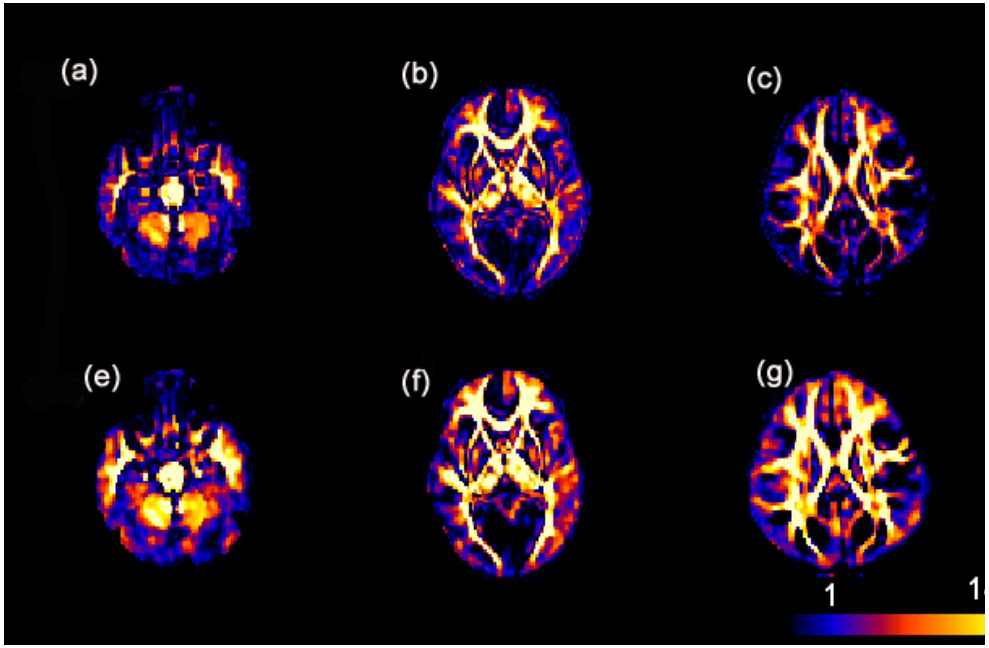
Knickmeyer RC, *et al.* (2008) *J Neurosci* 28: 12176-12182.

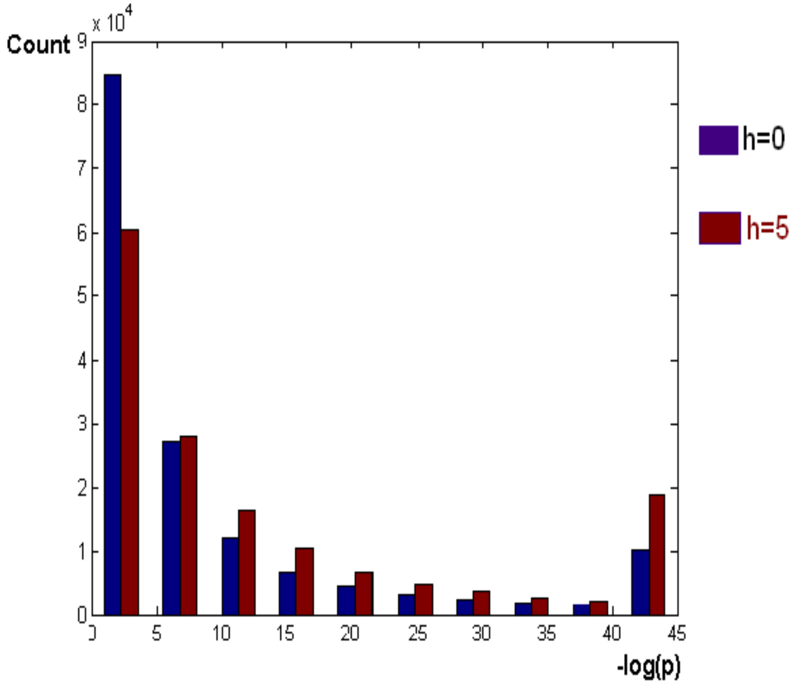


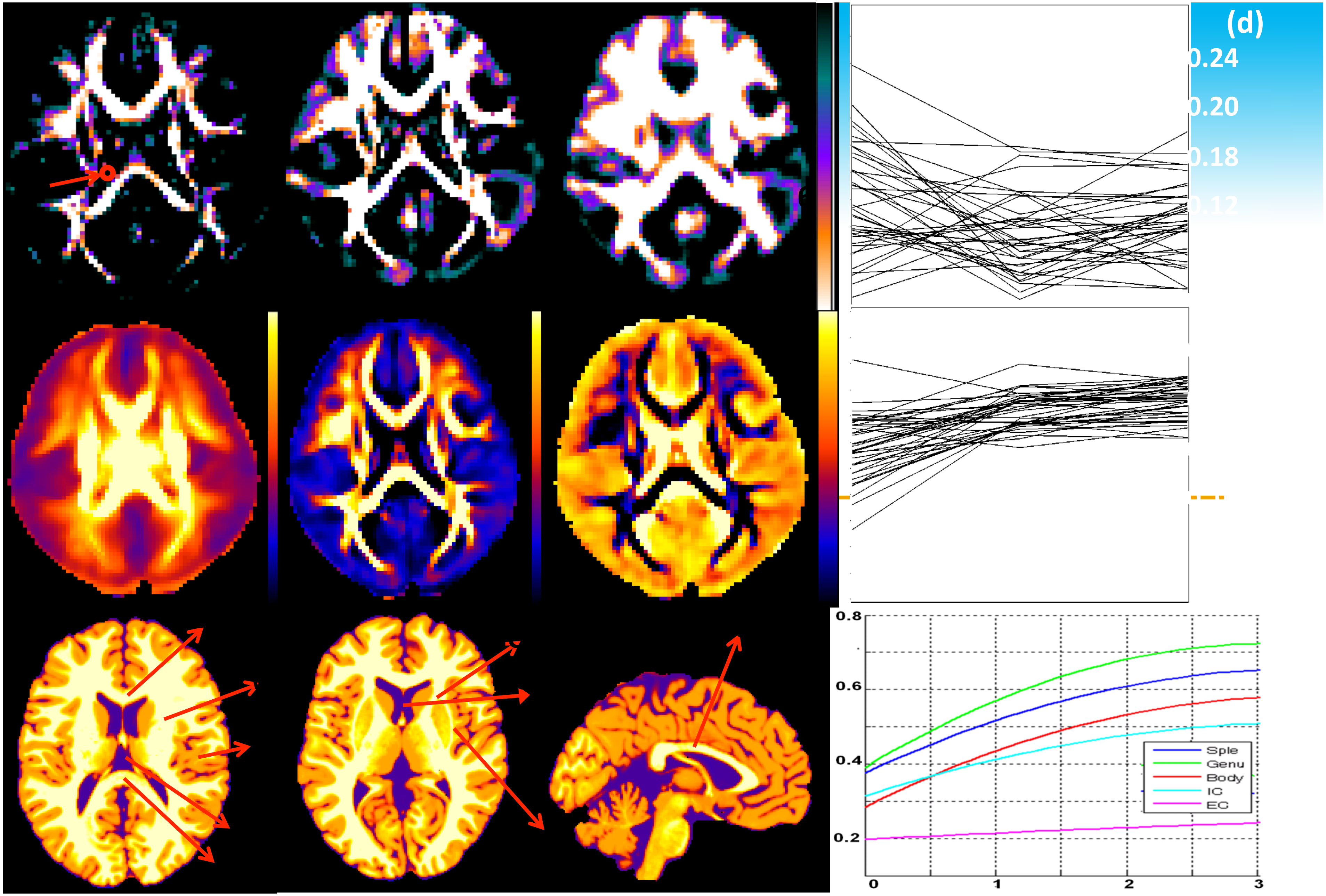
Infant Brain Development Data

- **Objective:** We want to assess the brain structure change in the early brain development.
- **Subject:** 38 infants.
- **Image:** Diffusion-weighted images and T1 weighted images were acquired for each subject at 2 weeks, 1 and 2 years old.
- **Method:** Voxel-wise imaging analysis and MARM.

Time Effect and Comparison



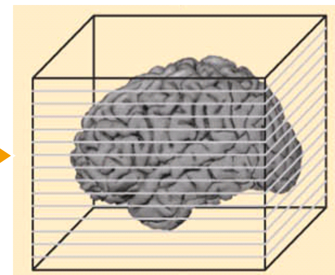


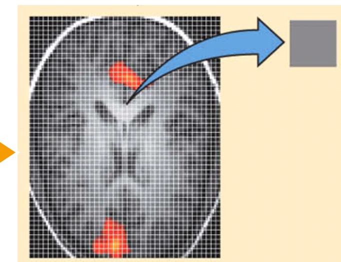

(d)
0.8
0.6
0.4
0.2
0
1
2
3
Sple
Genu
Body
IC
EC

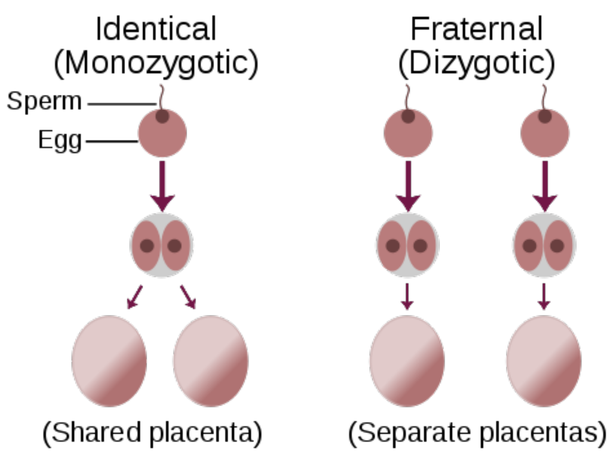
Twin-MARM

- ACE/ADE Models
- Two-stage MARM
- Consistency and Asymptotic Distribution

Twin-MARM







MZ twins share same genetic material

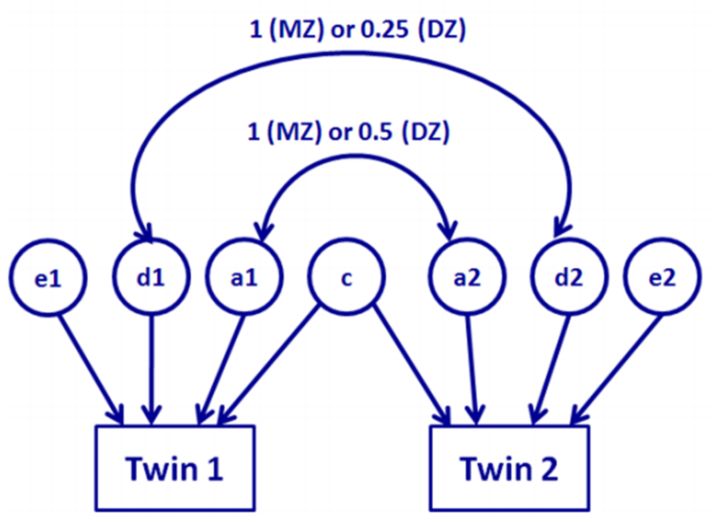
DZ twins share average 50% of their genes

Twin-MARM

At specific voxel v, we consider the structural equation model:

$$y_{ij}(v) = x_{ij}^T \boldsymbol{\beta}(v) + a_{ij}(v) + d_{ij}(v) + c_i(v) + e_{ij}(v)$$

$a_{ij}(v),\ d_{ij}(v),\ c_i(v)$ and $e_{ij}(v)$ **: the additive genetic, dominance genetic, common environmental and residual effects on i-th twin pair. We assume they are independently normally distributed with mean 0 and variance** $\sigma_a(v)^2,\ \sigma_d(v)^2,\ \sigma_c(v)^2$ and $\sigma_e(v)^2$.



Twin-MARM

There are two sets of parameters:
mean structure
variance structure

$$\omega(d,d';h_s) = K_{loc}(||d-d'||_2/h_s)K_{st}(D_{\boldsymbol{\theta}}(d,d';h_{s-1})/C_n)$$

$$D_{\boldsymbol{\theta}}(d,d';h_{s-1}) = [\hat{\boldsymbol{\theta}}(d;h_{s-1}) - \hat{\boldsymbol{\theta}}(d';h_{s-1})]^T \hat{\Sigma}(\hat{\boldsymbol{\theta}}(d;h_{s-1}))^{-1} [\hat{\boldsymbol{\theta}}(d;h_{s-1}) - \hat{\boldsymbol{\theta}}(d';h_{s-1})].$$

Question of interest:

Mean and variance images may have different patterns.

Twin-MARM

Two-stage Approach

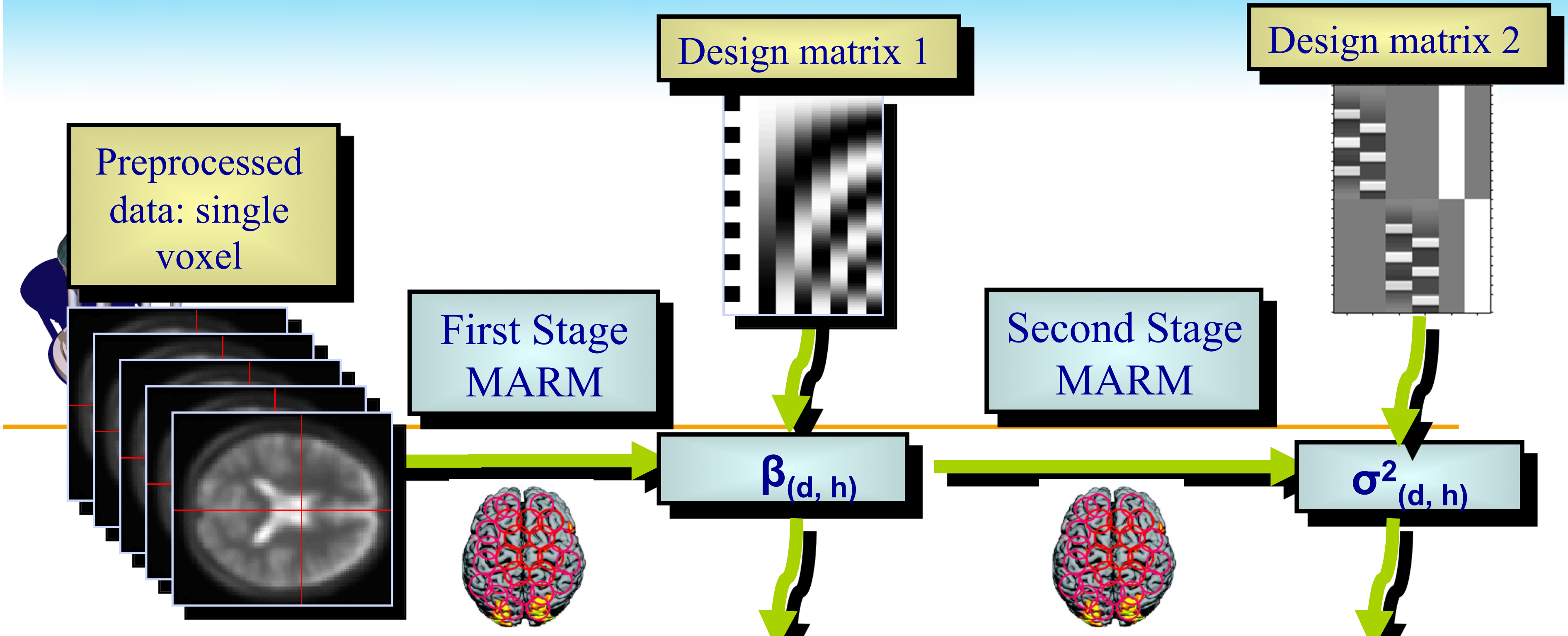
- **Mean structure**

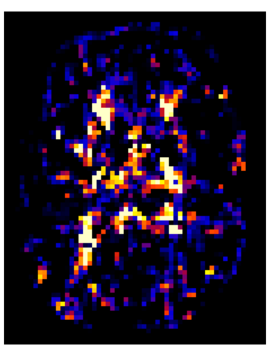
$$Y_{ij}(d) = x_{ij}^T \beta(d) + \varepsilon_{ij}(d) \Rightarrow \{\hat{\beta}(d;h) : d \in D\}$$

- **Variance structure**

$$\{Y_{ij}(d) - x_{ij}^T \hat{\beta}(d;h)\}^2 = z_{ij}^T \rho(d) + \delta_{ij}(d) \Rightarrow \{\hat{\rho}(d;h) : d \in D\}$$

TwinMARM



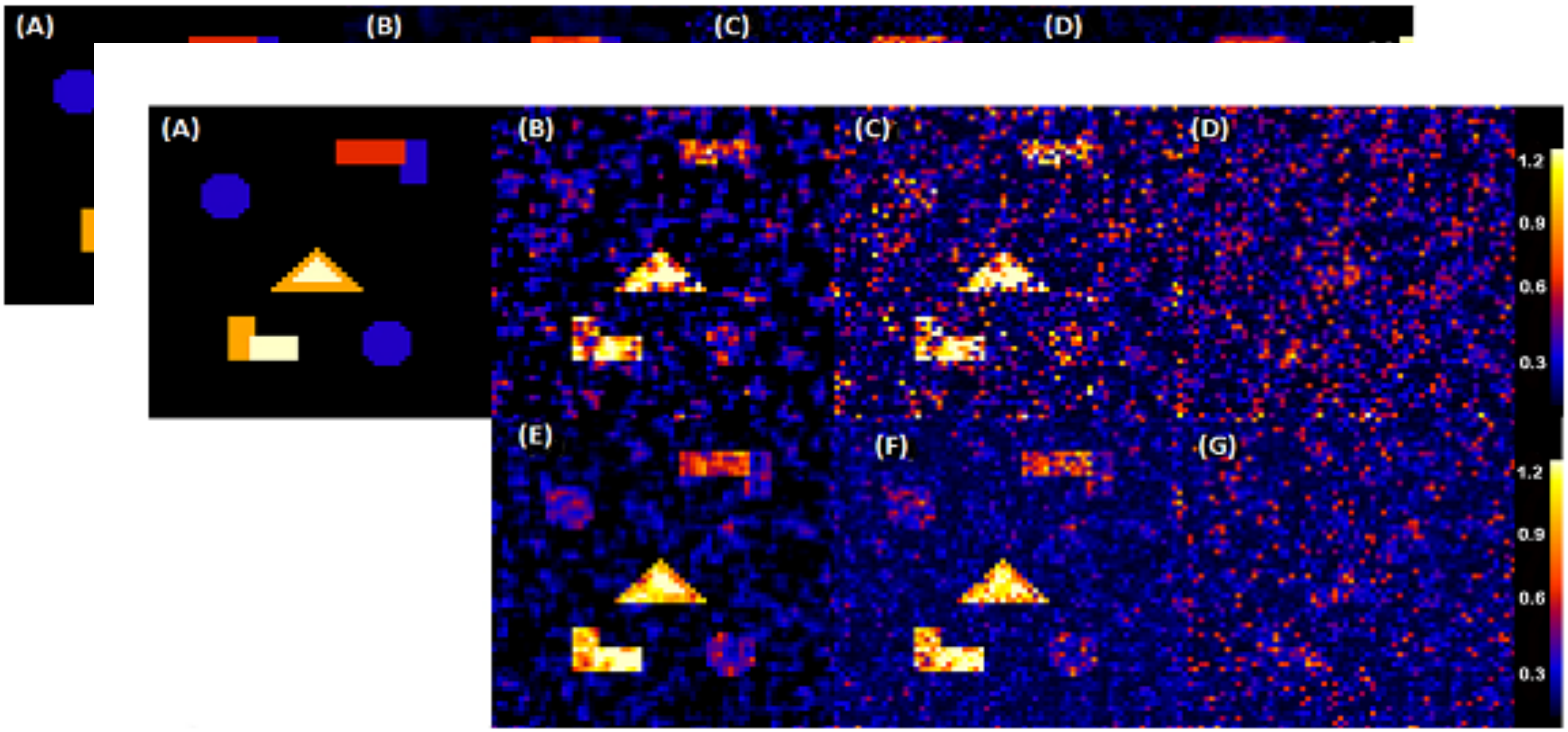


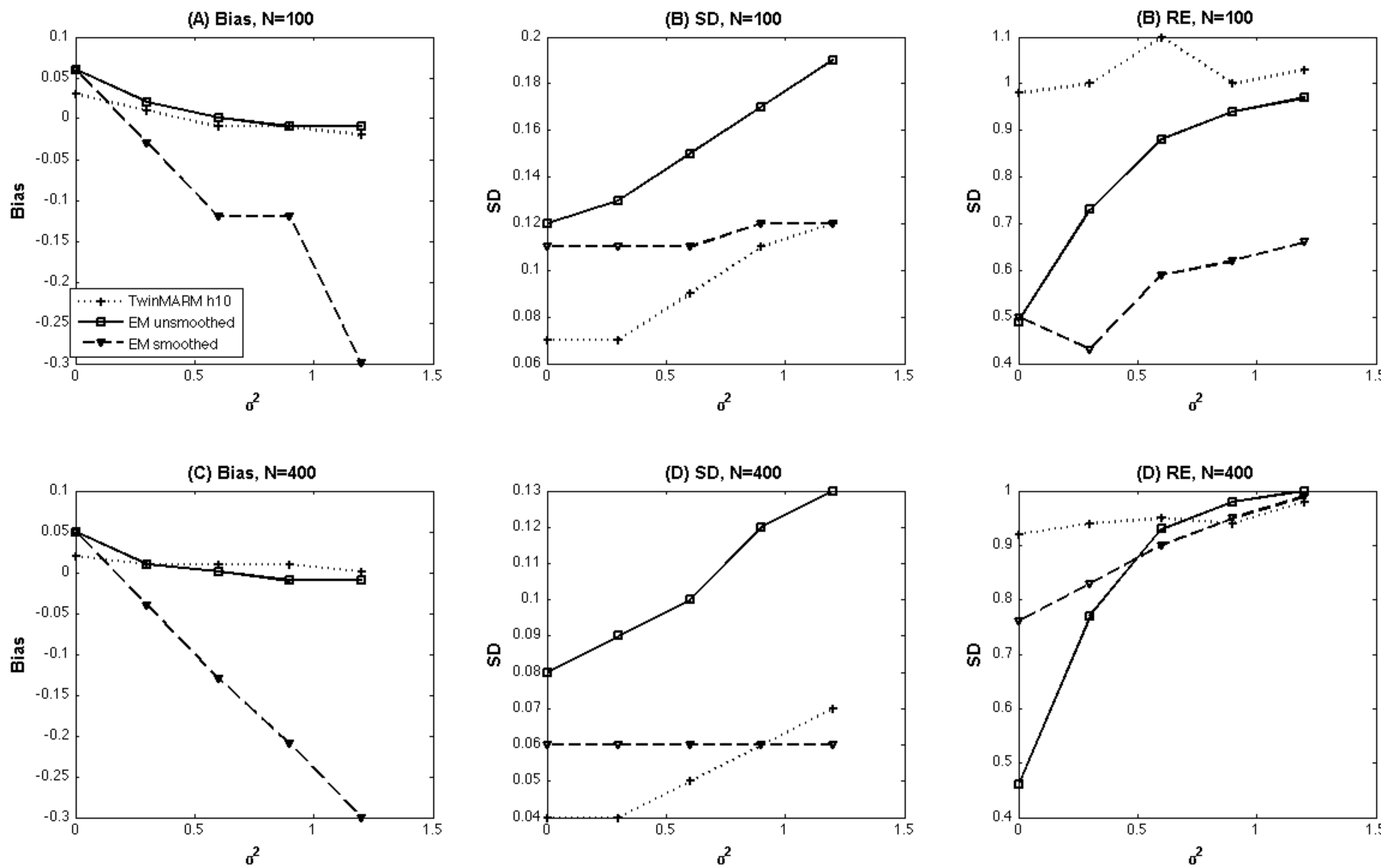
		$n = 100$			$n = 200$			$n = 400$		
$\beta_1(d)$		h_0	h_5	h_{10}	h_0	h_5	h_{10}	h_0	h_5	h_{10}
0	BIAS	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	RMS	0.21	0.12	0.08	0.15	0.09	0.05	0.11	0.06	0.04
	SD	0.21	0.12	0.07	0.15	0.08	0.05	0.11	0.06	0.04
	RE	1.02	1.06	1.06	1.01	1.03	1.03	1.00	1.01	1.01
0.5	BIAS	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	RMS	0.22	0.14	0.09	0.16	0.10	0.06	0.12	0.07	0.04
	SD	0.22	0.13	0.08	0.16	0.09	0.06	0.12	0.07	0.04
	RE	1.01	1.09	1.11	1.01	1.04	1.03	1.00	1.00	0.98

		$n = 100$			$n = 200$			$n = 400$		
$\sigma_a^2(d)$		h_0	h_5	h_{10}	h_0	h_5	h_{10}	h_0	h_5	h_{10}
0	BIAS	0.03	0.03	0.03	0.02	0.02	0.02	0.02	0.02	0.02
	RMS	0.13	0.08	0.07	0.08	0.06	0.05	0.06	0.04	0.04
	SD	0.13	0.08	0.07	0.09	0.06	0.05	0.06	0.04	0.04
	RE	0.99	0.99	0.98	0.99	0.98	0.98	0.96	0.93	0.92
0.3	BIAS	0.03	0.01	0.01	0.02	0.00	0.00	0.02	0.02	0.01
	RMS	0.14	0.09	0.07	0.09	0.06	0.05	0.07	0.04	0.04
	SD	0.14	0.09	0.07	0.10	0.06	0.05	0.07	0.05	0.04
	RE	1.00	1.00	1.00	0.99	0.97	0.97	0.97	0.95	0.94

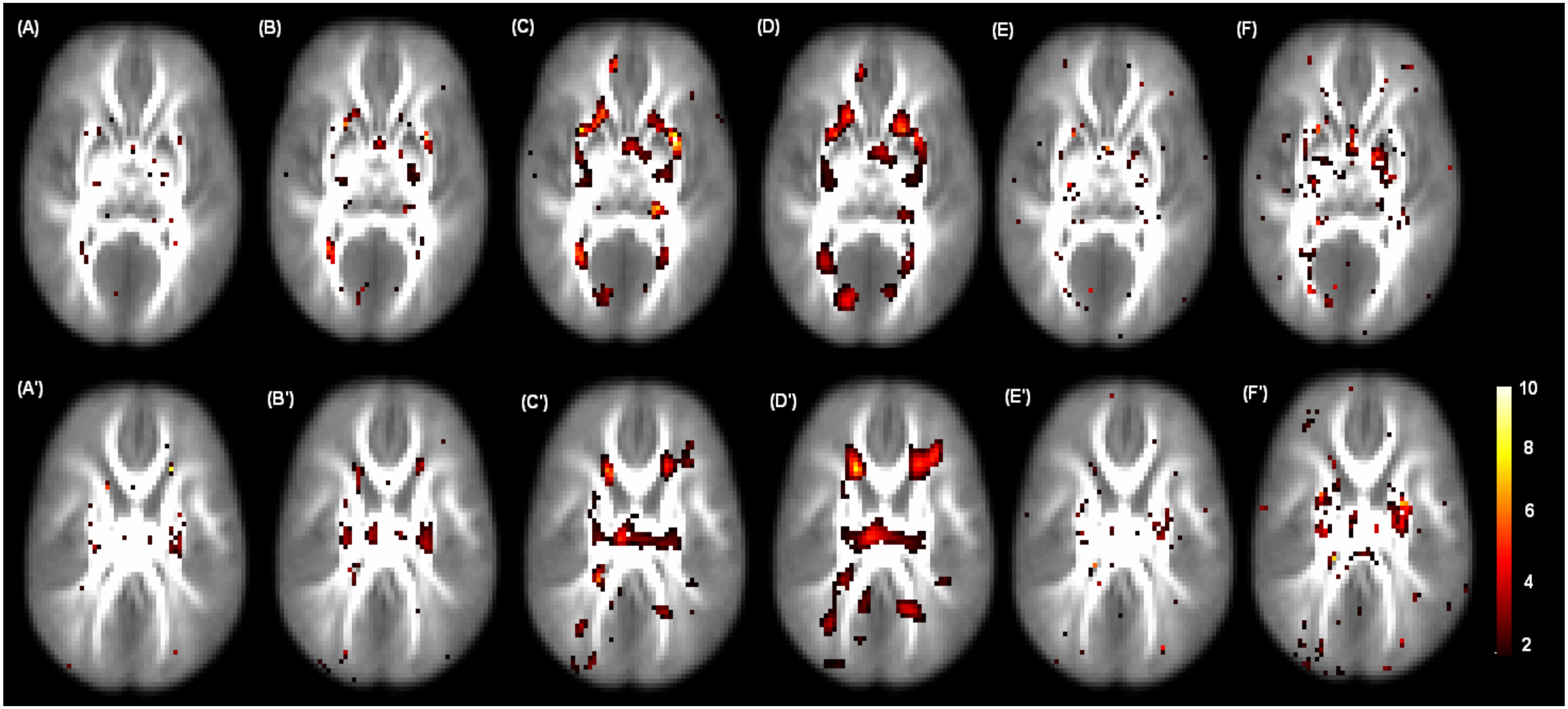
Twin-MARM

It is dangerous to use Gaussian-kernel to smooth imaging data and then carry out twin analysis.





(A) Bias, N=100
(B) SD, N=100
(B) RE, N=100
(C) Bias, N=400
(D) SD, N=400
(D) RE, N=400
Bias
SD
σ^2
TwinMARM h10
EM unsmoothed
EM smoothed


(A)
(B)
(C)
(D)
(E)
(F)
(A')
(B')
(C')
(D')
(E')
(F')
10
8
6
4
2

Multiscale Adaptive Smoothing Models for HRF in fMRI

- Convolution Models in Frequency Domain
- Back Fitting Methods
- Multi-stage MARM

Multiscale Adaptive Smoothing Model

D: 3D volume

N_D: the number of points on *D*

d: a voxel in *D*

$\{Y(t,d): t = 1\times t_{TR},\cdots,T\times t_{TR}, d\in D\}$: spatial-temporal process

$\{X(t): t\in[0,T\times t_{TR}]\}$: external stimulus process

$\{H(t,d): t\in[0,T\times t_{TR}], d\in D\}$: spatial-temporal HRF process

$\{\varepsilon(t,d): t\in[0,T\times t_{TR}], d\in D\}$: error process

Voxel-wise Approach

$$Y(t,d) = H(\bullet,d) \otimes X(t) + \varepsilon(t,d) = \int H(t-u,d)X(u)du + \varepsilon(t,d)$$

time-domain

$$H(t,d) = \sum_{k=1}^{K} \beta_k(d) f_k(t)$$

$$Y(t,d) = \sum_{k=1}^{K} \beta_k(d) \int f_k(t-u,d)X(u)du + \varepsilon(t,d) = \sum_{k=1}^{K} \beta_k(d) x_k(t) + \varepsilon(t,d)$$

frequency-domain

$$F_Y(f,d) = F_H(f,d)F_X(f) + F_\varepsilon(t,d) \text{ for } f \in [0, T \times t_{TR}]$$

$$F_Y(t,d) = \int_0^{T \times t_{TR}} Y(t,d)\exp(-2\pi i f t/(T \times t_{TR}))dt$$

Continuous

$$F_Y(f,d) = F_H(f,d)F_X(f) + F_\varepsilon(t,d) \text{ for } f \in [0, T \times t_{TR}]$$

Discrete

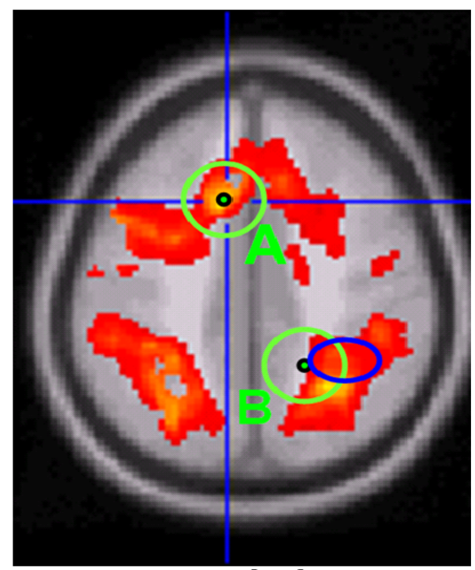
$$\phi_Y(f,d) = \phi_H(f,d)\phi_X(f) + \phi_\varepsilon(f,d) \text{ for } f \in [0, T \times t_{TR}]$$

$$\phi_Y(f,d) = \sum_{t=0}^{T \times t_{TR}} Y(t,d)\exp(-2\pi i f t / (T \times t_{TR}))$$

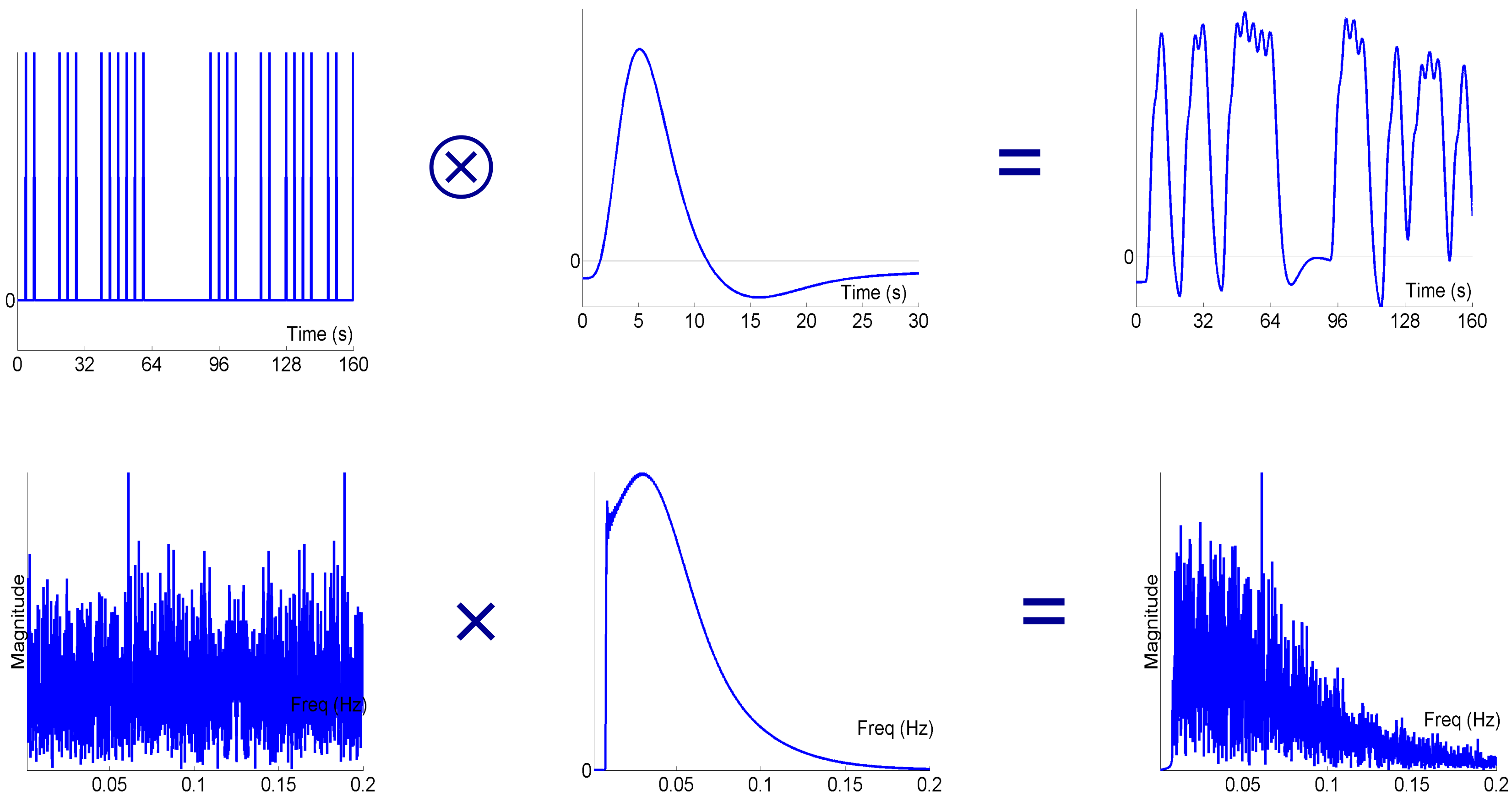
Key Assumptions:

$$\phi_\varepsilon(f,d) \sim (0, 1(f = f')\sigma(f,f';d,d'))$$

$$\phi_H(f,d) \text{ is piecewisely smooth for } (f,d) \in N(f,d)$$



Stimulus (“Neural”) HRF Predicted Data



$$\phi_Y(f,d) = \phi_H(f,d)\phi_X(f) + \phi_\varepsilon(f,d) \text{ for } (f,d) \in \underline{N(f,d)}$$

Unknown

Approximation

$$\begin{aligned}\phi_Y(f_k,d') &= \phi_H(f_k,d')\phi_X(f_k) + \phi_\varepsilon(f_k,d') \\ &\approx \phi_H(f,d)\phi_X(f_k) + \phi_\varepsilon(f_k,d')\end{aligned}$$

$$(f_k,d') \in B((f,d);\varepsilon,r) = (f-\varepsilon, f+\varepsilon) \times B(d',r)$$

Multiple Events: Backfitting Methods

Weighted LSE

$$L(\phi_H(f,d);B((f,d);r,h)) = \sum_{(f_k,d')}[\phi_Y(f_k,d') - \phi_H(f,d)\phi_X(f_k)]^2\, w(f,d,f_k,d';\varepsilon,r)$$

$$\hat{\phi}_H(f,d) = \sum_{(f_k,d')} \phi_Y(f_k,d')\bar{\phi}_X(f_k)\, w(f,d,f_k,d';\varepsilon,r) \Big/ \sum_{(f_k,d')} \phi_X(f_k)\bar{\phi}_X(f_k)\, w(f,d,f_k,d';\varepsilon,r)$$

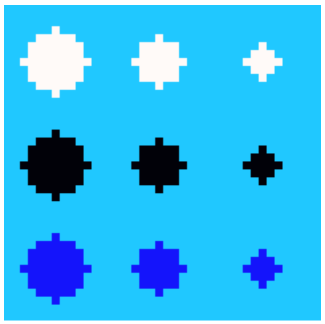
$$\mathrm{Var}\,(\hat{\phi}_H(f,d))$$

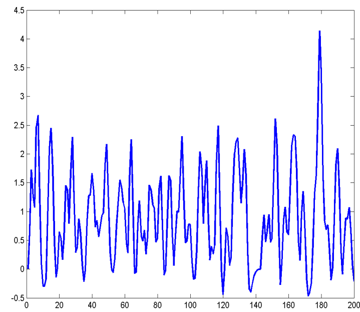
Estimated HRF

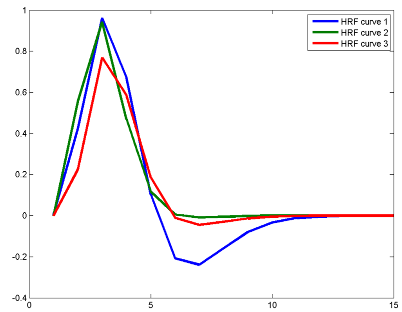
$$\hat{H}(t,d) = \sum_{k=0}^{T-1} \hat{\phi}_H(f,d)\exp(i2\pi t f_k)[1-\cos(2\pi t/T)]/(2\pi^2 t^2/T)$$

Simulation II: Multivariate Case

The background image and the simulated one with their related curves.
In this simulation the smallest SNR is between 0.5 and 0.7.





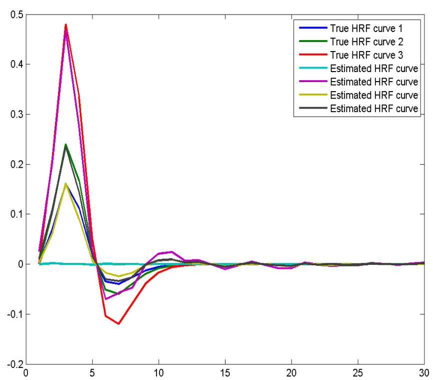


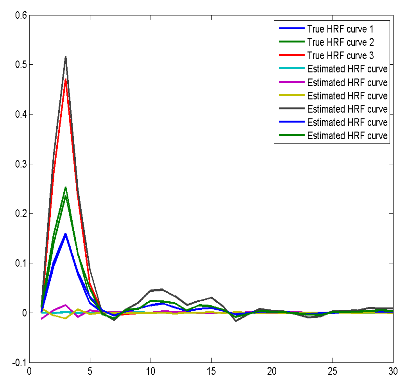
Three activated regions for each sequence of events correspond with three different HRFs: $h_j(t)/2$, $h_j(t)/4$, $h_j(t)/6$, j=1,2,3

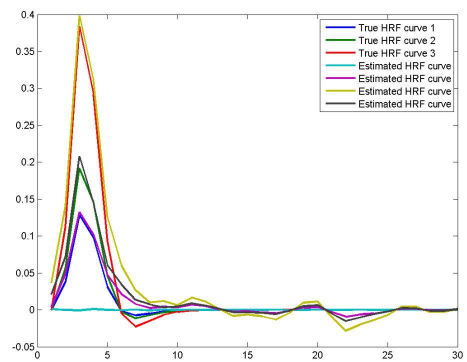
$$h_j(t) = A_j \cdot (t/d_{j1})^{a_{j1}} \exp\left(-(t-d_{j1})/b_{j1}\right) - c(t/d_{j2})^{a_{j2}} \exp\left(-(t-d_{j2})/b_{j2}\right) \qquad j = 1,2,3$$

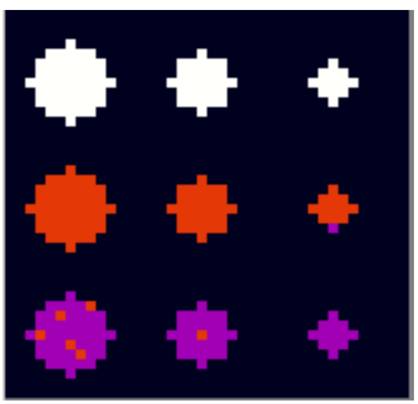
$$\epsilon(t) \sim N(0, \sigma^2) \qquad X_j(t) \sim B(1, 0.15)$$

The estimates of the HRFs, from the left to right, are the 1st, 2nd and 3rd sequences of events.

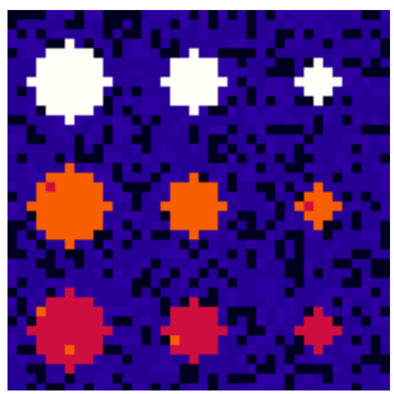




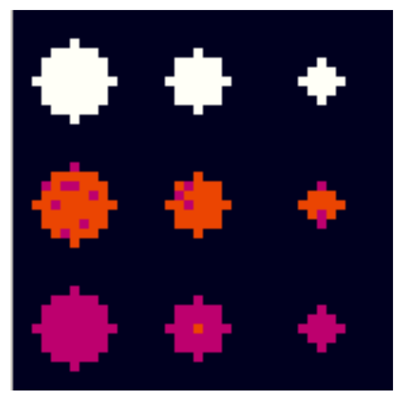




H1(t)



H2(t)



H3(t)

Comparison with Lindquist *et al* (2009)

$D = \frac{1}{n}\frac{1}{m}\sum_{i=1}^{n}\sum_{j=1}^{m}(|\hat{x}_{ij} - x_0| - |\hat{y}_{ij} - x_0|)$ n is the sample size, m is the number of voxels in each active region.

Param		sFIR A1	sFIR A2	sFIR A3	IR A1	IR A2	IR A3	GAM A1	GAM A2	GAM A3
H	C1	-0.02 (0.045)	-0.04 (0.045)	-0.09 (0.060)	-0.11 (0.0819)	-0.16 (0.115)	-0.32 (0.218)	-0.13 (0.039)	-0.20 (0.044)	-0.42 (0.059)
H	C2	-0.02 (0.041)	-0.04 (0.043)	-0.10 (0.057)	-0.09 (0.0699)	-0.14 (0.096)	-0.27 (0.175)	-0.11 (0.041)	-0.17 (0.045)	-0.36 (0.057)
H	C3	-0.01 (0.046)	-0.02 (0.047)	-0.07 (0.064)	-0.07 (0.0673)	-0.10 (0.089)	-0.21 (0.161)	-0.07 (0.038)	-0.11 (0.045)	-0.25 (0.066)
T	C1	-0.08 (0.722)	0.05 (0.326)	0.01 (0.073)	-3.74 (3.313)	-3.49 (3.309)	-3.19 (3.292)	-2.50 (0.425)	-2.66 (0.298)	-2.84 (0.070)
T	C2	-0.05 (0.685)	0.07 (0.292)	0.01 (0.069)	-3.50 (3.440)	-3.34 (3.475)	-2.88 (3.419)	-2.57 (0.452)	-2.75 (0.287)	-2.91 (0.069)
T	C3	-0.55 (1.159)	-0.10 (0.513)	-0.10 (0.513)	-3.54 (3.404)	-3.26 (3.430)	-3.03 (3.415)	-2.46 (0.577)	-2.66 (0.427)	-2.87 (0.254)
W	C1	-0.20 (0.671)	-0.28 (0.596)	-0.42 (0.518)	-1.70 (2.122)	-1.73 (2.127)	-1.66 (2.094)	-3.33 (0.623)	-3.41 (0.576)	-3.46 (0.515)
W	C2	-0.38 (0.760)	-0.41 (0.597)	-0.49 (0.513)	-1.78 (2.143)	-1.80 (2.099)	-1.79 (2.018)	-3.35 (0.634)	-3.41 (0.575)	-3.49 (0.512)
W	C3	-0.32 (0.870)	-0.33 (0.741)	-0.48 (0.658)	-1.79 (2.179)	-1.85 (2.123)	-2.08 (2.221)	-3.30 (0.713)	-3.42 (0.677)	-3.63 (0.550)

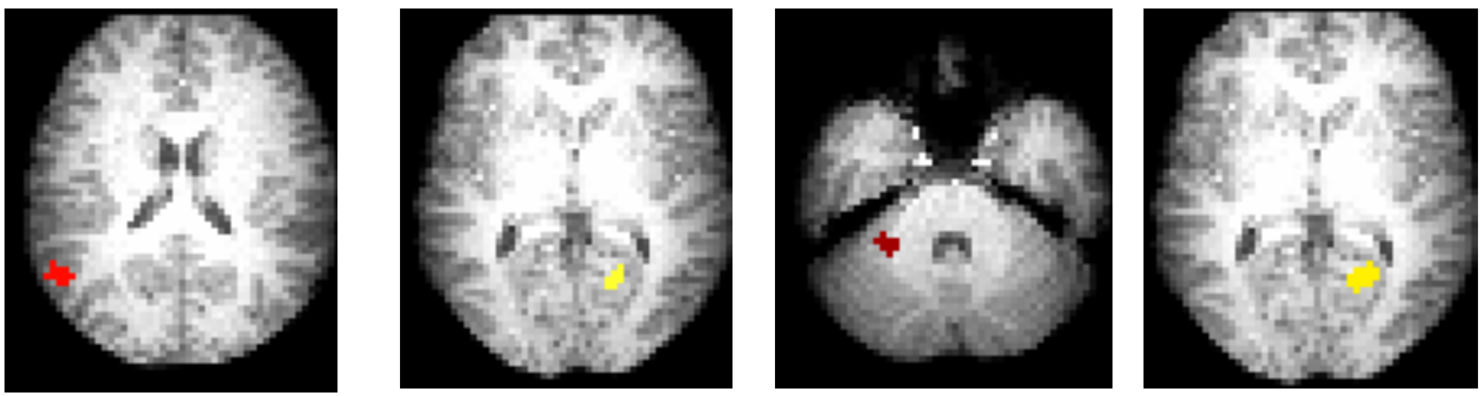
Comparisons of the differences of the absolute errors between our method with smooth finite impulse response (sFIR), inverse logit (IL) and SPM canonical HRF (GAM), respectively. C1, C2 and C3 denotes the 1st, 2nd and 3rdsequences of events, respectively. A1, A2 and A3 denotes the1st, 2nd and 3rd active regions. Values in the blanket are the standard deviations. H=Height, W=Width, T=Time-to-Peak.

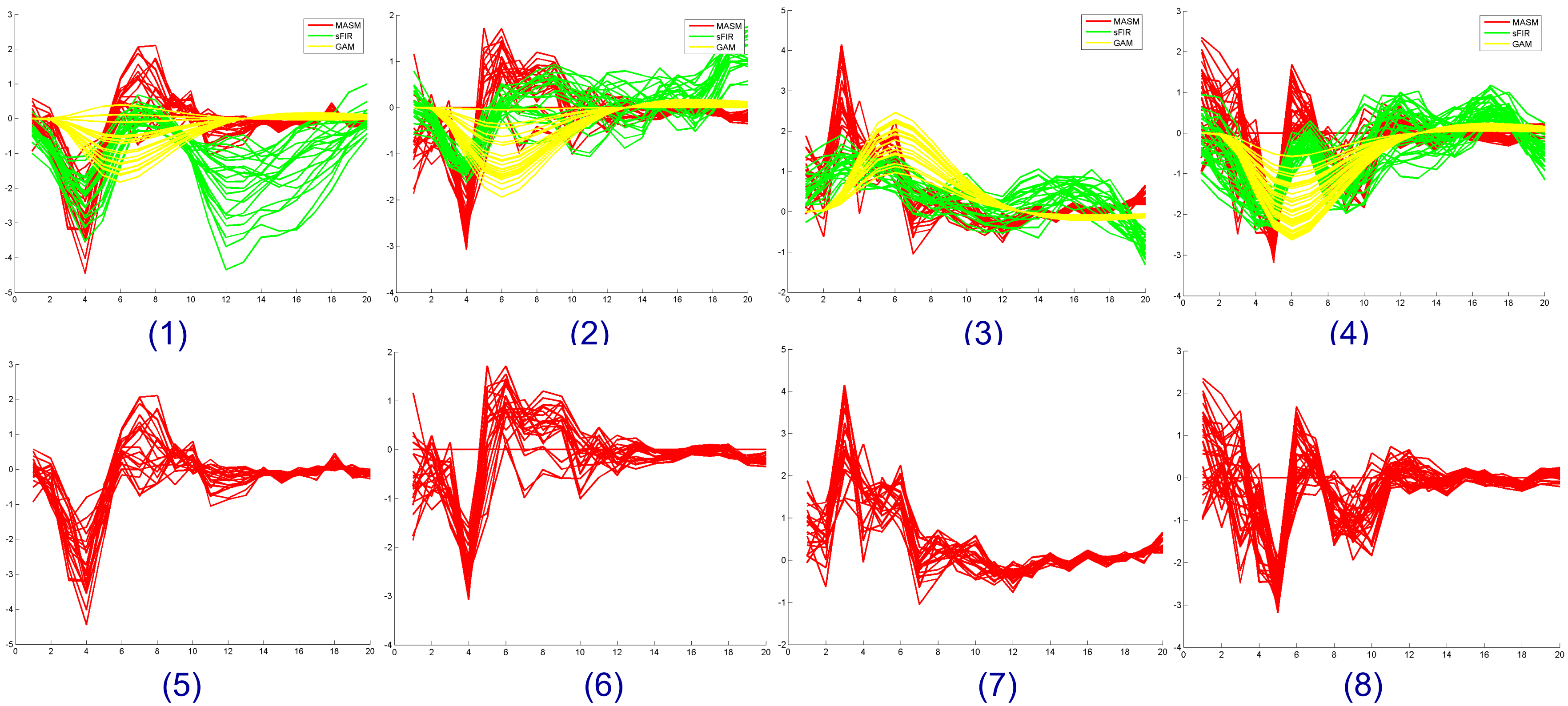
This data set is from a memory related experiment to compare the neural correlates of relational memory during implicit (non-strategic) versus explicit (conscious, strategic) retrieval.

There are four different sequences of stimuli.

We use SPM8 to preprocess the images including the realignment, timing slicing, segmentation, coregistration, normalization and spatial smoothing.

We focus on some significant regions of interest (ROI) detected by SPM to study the HRFs of the voxels by our method. The results are verified by sFIR and GAM.





(1)-(4) Estimated HRFs at the significant ROIs corresponding each condition from MASM (red), sFIR(green) and GAM(yellow); (5)-(8) Estimated HRFs from only MASM in the each ROI.

References

- Li, YM, Zhu, HT, Shen, DG, Lin WL., Gilmore, J, Ibrahim, JG. (2010). Multiscale Adaptive Regression Models for Neuroimaging Data. *JRSSB,* in press.
- Li, YM, Ja-an Lin, M. Styner, John Gilmore, Zhu HT. Two-stage spatial adaptive analysis of twin neuroimaging data. *Submitted.* .
- Li, YM, M. Styner, John Gilmore, Zhu HT. Spatial adaptive generalized estimating equations for longitudinal neuroimaging data. *Submitted.*
- Skup, M, Zhu, HT, Zhang HP. Spatial adaptive generalized Moment Estimation for longitudinal neuroimaging data. *Submitted.*
- Wang, J. Zhu, H.T., Fan, J.Q., Giovanello, K., and Lin, W. (2011).

Multiscale Adaptive Smoothing Models for the Hemodynamic Response Function in fMRI. Submitted.

Acknowledgements

SI and SINS

Establish ASA Section: Statistics in Imaging (SI)

- **Advisory board committee members for SI:**
  Martin Lindquist, Daniel Rowe, Brian Caffo,
  Hernando Ombao, F. Dubois Bowman, Thomas Nichols,
  Ranjan Maitra, Hongtu Zhu, Kary Myers.

Society of Imaging Neuroscience Statisticians (SINS)

http://www.mscs.mu.edu/~dbrowe/sins.html

- **Roundtables in ENAR 2011 and JSM 2011.**

Simulation Studies

$$y_i(d) = x_i^T \beta(d) + \varepsilon_i(d)$$

$$\varepsilon_i(d) \sim N(0,1) \qquad \varepsilon_i(d) \sim \chi^2(3) - 3$$

$$n = 60 \quad \text{or} \quad n = 80$$

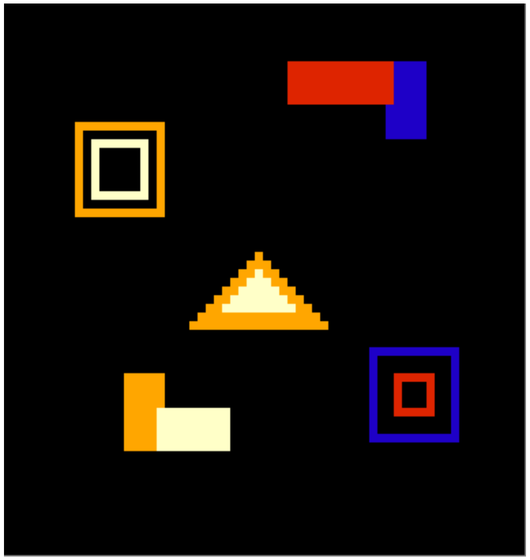
$$x_i = (1, x_{i2}, x_{i3})^T$$

$$x_{i2} \sim \text{Bernoulli}(0.5)$$

$$x_{i3} \sim \text{Uniform}[1,2]$$

$$\beta(d) = (\beta_1(d), \beta_2(d), \beta_3(d))^T$$

$$\beta_1(d) = \beta_2(d) = 0$$



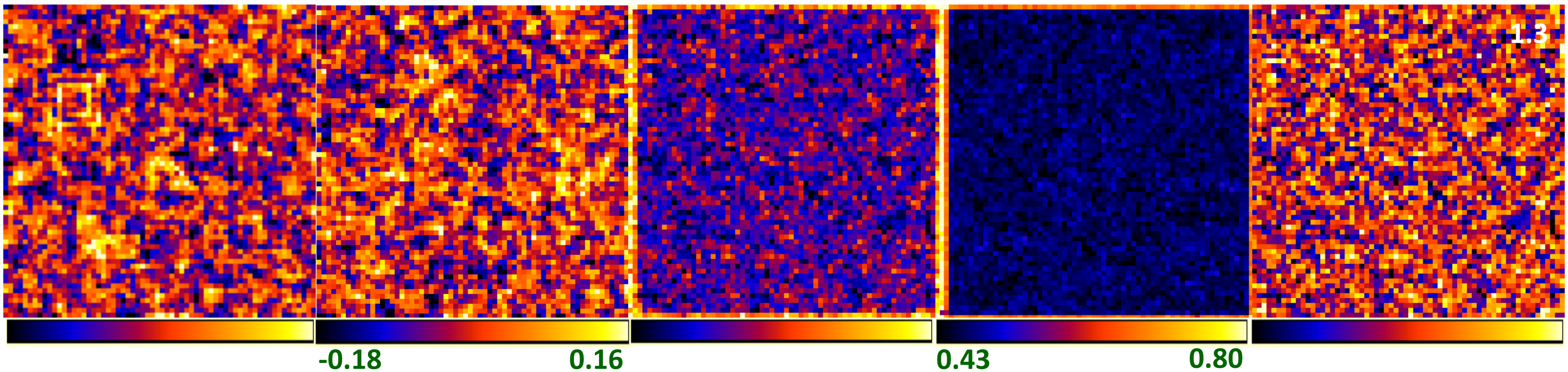
ROI	*black*	*blue*	*red*	*yellow*	*white*
$\beta_3(d)$	0.0	0.2	0.4	0.6	0.8

Simulation Studies

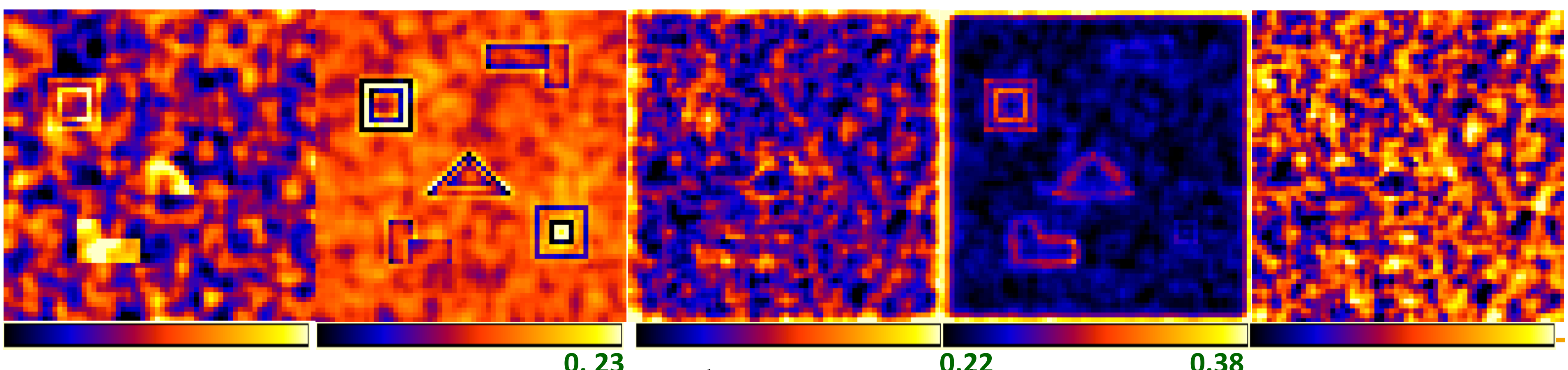
		$\chi^2(3)-3$						$N(0,1)$					
		$n=60$			$n=80$			$n=60$			$n=80$		
$\beta_2(d)$		h_0	h_5	h_{10}	h_0	h_5	h_{10}	h_0	h_5	h_{10}	h_0	h_5	h_{10}
0.0	BIAS	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	RMS	0.48	0.35	0.26	0.41	0.31	0.22	0.20	0.15	0.11	0.17	0.13	0.09
	SD	0.47	0.34	0.24	0.41	0.30	0.21	0.19	0.14	0.10	0.17	0.12	0.09
	RE	1.03	1.05	1.06	1.02	1.03	1.04	1.03	1.05	1.06	1.02	1.03	1.04
0.2	BIAS	0.00	-0.03	-0.07	0.01	-0.02	-0.06	0.00	-0.03	-0.05	0.00	-0.02	-0.05
	RMS	0.46	0.34	0.24	0.39	0.29	0.21	0.19	0.14	0.11	0.16	0.12	0.09
	SD	0.46	0.33	0.24	0.40	0.29	0.21	0.19	0.14	0.10	0.16	0.12	0.09
	RE	1.01	1.01	1.01	0.99	1.00	1.01	1.02	1.04	1.06	1.02	1.02	1.03
0.4	BIAS	-0.01	-0.05	-0.09	0.01	-0.02	-0.06	0.00	0.00	-0.01	0.00	0.00	0.00
	RMS	0.46	0.34	0.25	0.40	0.30	0.22	0.19	0.15	0.12	0.16	0.13	0.10
	SD	0.46	0.33	0.24	0.40	0.29	0.21	0.19	0.14	0.11	0.16	0.12	0.09
	RE	1.01	1.02	1.03	1.01	1.02	1.03	1.03	1.05	1.07	1.00	1.01	1.02
0.6	BIAS	0.00	-0.05	-0.09	0.00	-0.04	-0.07	0.00	0.01	0.02	0.00	0.00	0.01
	RMS	0.46	0.35	0.26	0.40	0.30	0.23	0.19	0.15	0.12	0.16	0.13	0.10
	SD	0.46	0.34	0.25	0.40	0.30	0.22	0.19	0.14	0.11	0.16	0.13	0.10
	RE	1.01	1.03	1.04	1.01	1.02	1.03	1.02	1.04	1.06	1.01	1.03	1.04
0.8	BIAS	0.00	-0.04	-0.06	0.00	-0.02	-0.05	0.00	-0.01	-0.02	0.00	0.00	-0.01
	RMS	0.47	0.35	0.26	0.40	0.30	0.23	0.19	0.15	0.11	0.17	0.13	0.10
	SD	0.46	0.34	0.25	0.40	0.30	0.22	0.19	0.14	0.11	0.16	0.12	0.09
	RE	1.02	1.03	1.04	1.01	1.02	1.03	1.02	1.04	1.05	1.03	1.05	1.06

Simulation Studies

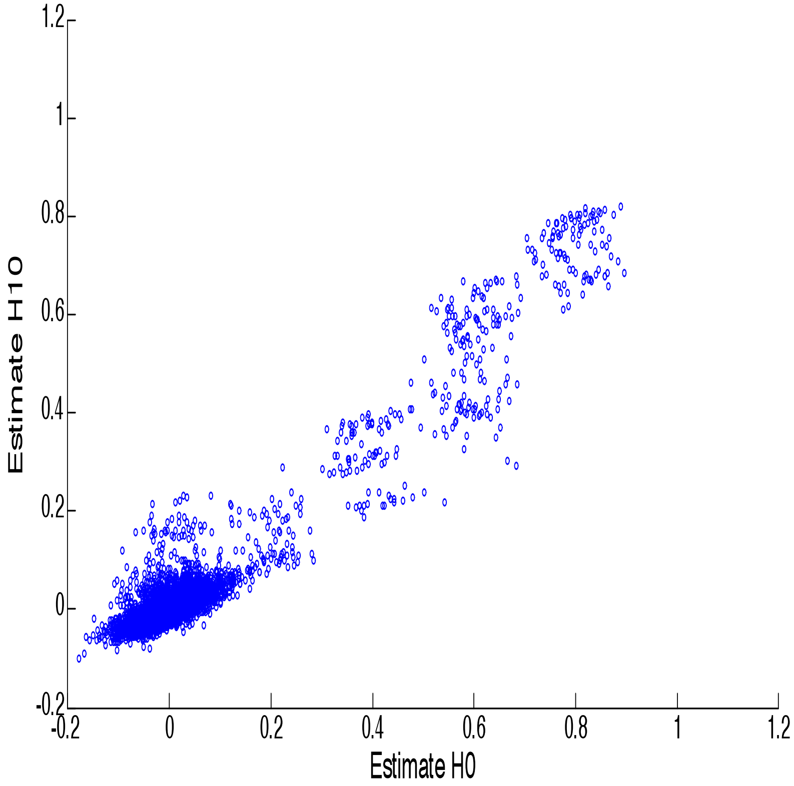
$\hat{\beta}_3(d,h_0)$

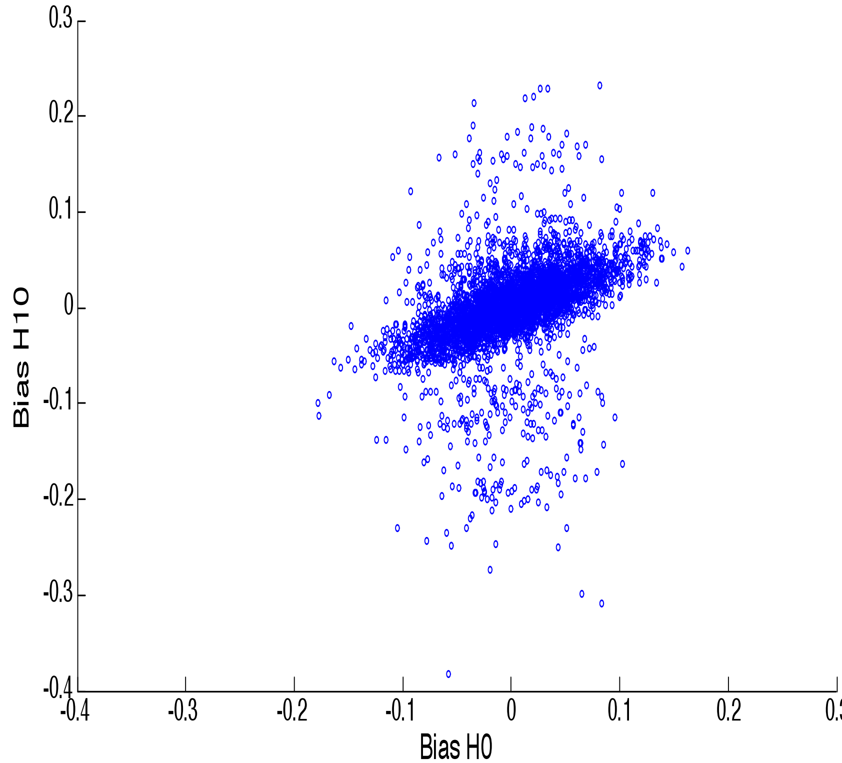


$\hat{\beta}_3(d,h_{10})$

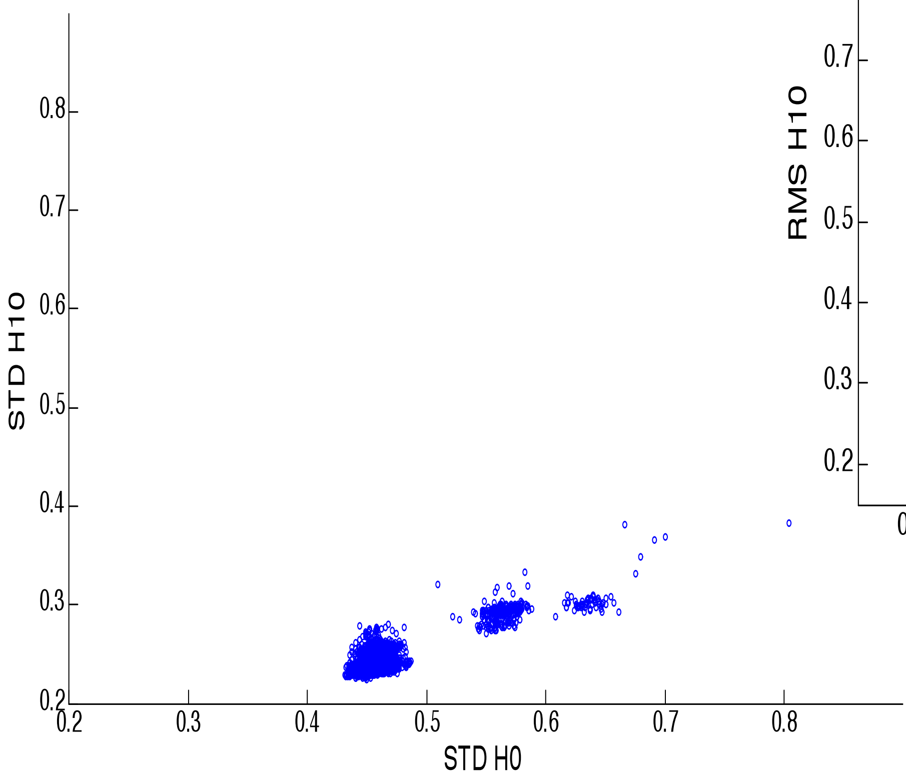


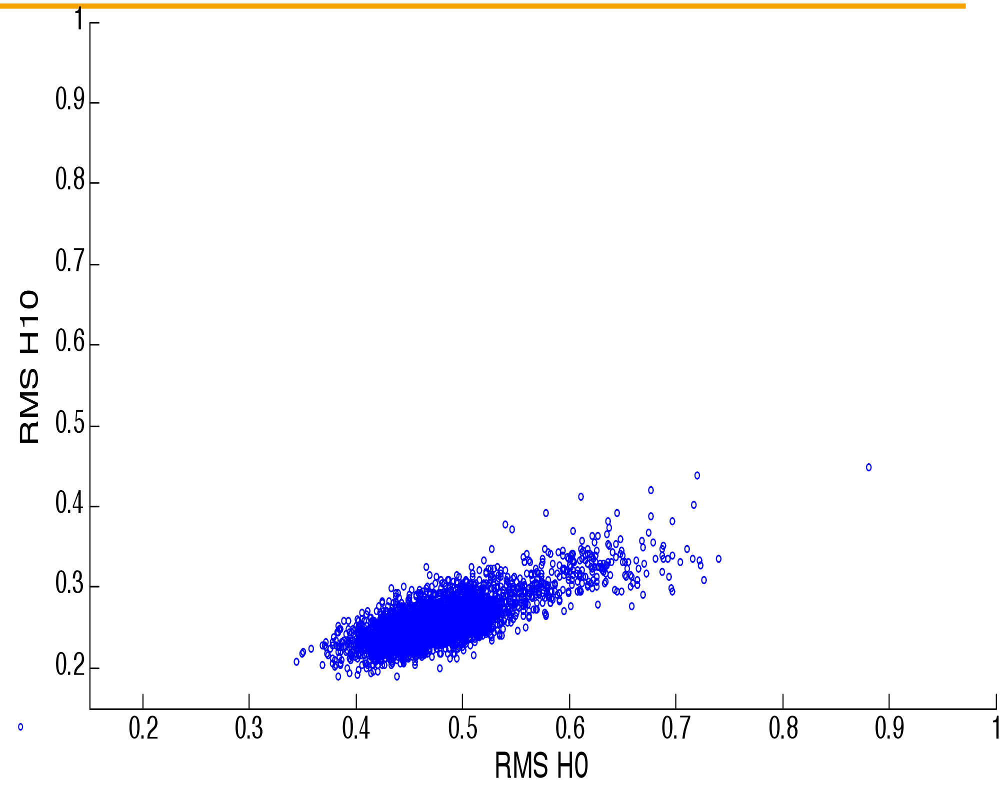
Simulation Studies



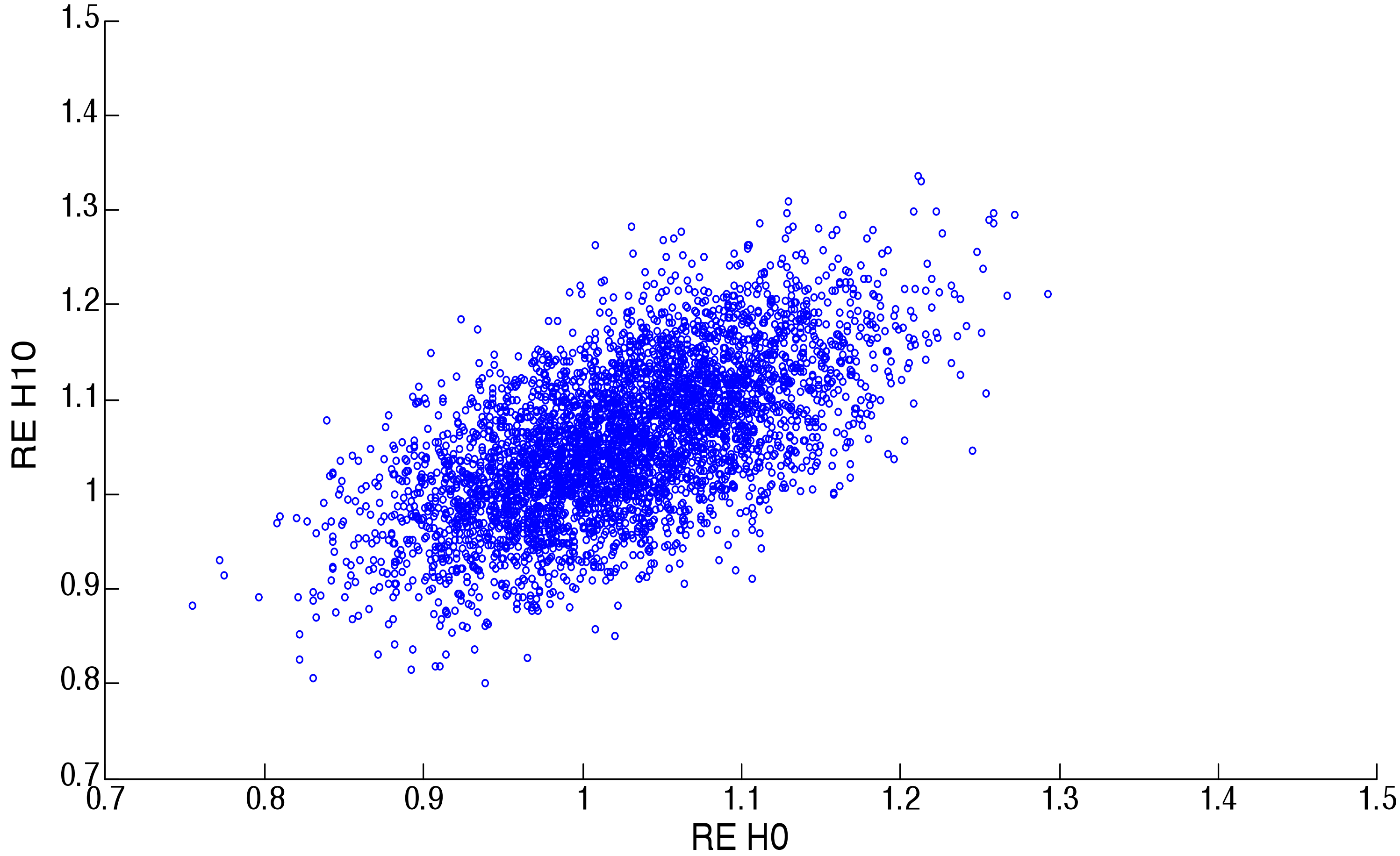


Simulation Studies





Simulation Studies



Simulation Studies

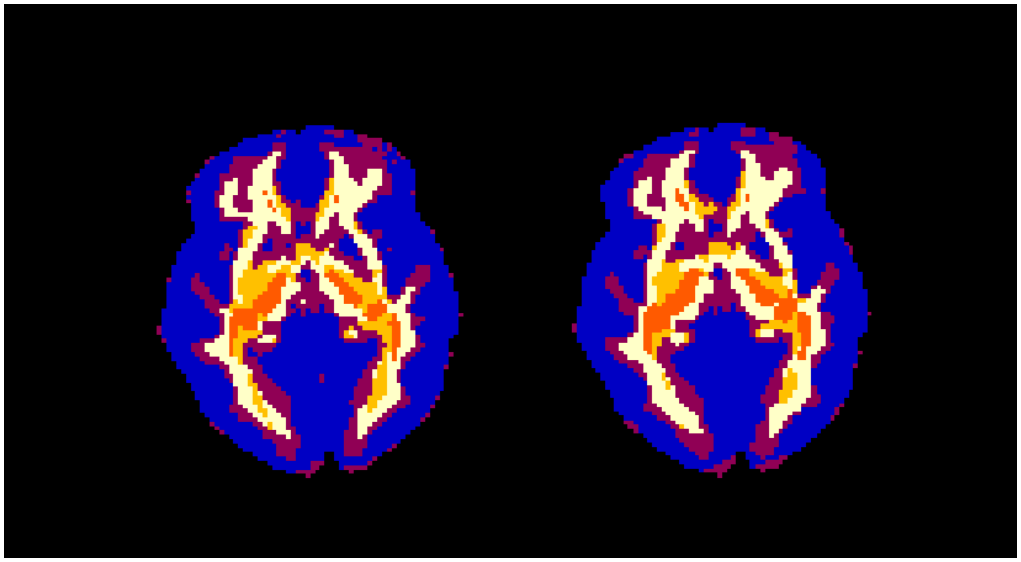
Table 2. Simulation study for $W_\mu(d, h)$: estimates (ES) and standard errors (SE) of rejection rates for pixels inside the five ROIs were reported at 2 different scales (h_0, h_{10}), 2 different distributions ($N(0,1)$ and $\chi^2(3) - 3$), and 2 different sample sizes ($n = 60, 80$) at $\alpha = 5\%$. For each case, 1,000 simulated data sets were used.

		$N(0,1)$				$\chi^2(3)-3$			
		$n=60$		$n=80$		$n=60$		$n=80$	
$\beta_2(d)$	s	ES	SE	ES	SE	ES	SE	ES	SE
0.2	h_0	0.20	0.066	0.24	0.070	0.08	0.038	0.08	0.037
	h_{10}	0.30	0.126	0.38	0.121	0.10	0.069	0.18	0.081
0.4	h_0	0.56	0.090	0.67	0.079	0.15	0.065	0.18	0.070
	h_{10}	0.93	0.051	0.98	0.030	0.26	0.129	0.35	0.159
0.6	h_0	0.88	0.039	0.95	0.024	0.27	0.057	0.33	0.050
	h_{10}	1.00	0.004	1.00	0.004	0.51	0.091	0.63	0.083
0.8	h_0	0.99	0.015	1.00	0.005	0.43	0.080	0.52	0.080
	h_{10}	0.99	0.010	0.99	0.011	0.78	0.099	0.90	0.006
0.0	h_0	0.07	0.006	0.07	0.006	0.06	0.007	0.07	0.006
	h_{10}	0.08	0.011	0.07	0.011	0.07	0.012	0.08	0.012

Early Brain Development: Structural connectivity

Comparison:

MAGEEC(0) versus **MAGEEC(5)**



Model Selection: