Wavelet-Based Functional Mixed Models

Jeffrey S. Morris

UT M.D. Anderson Cancer Center

Raymond J. Carroll

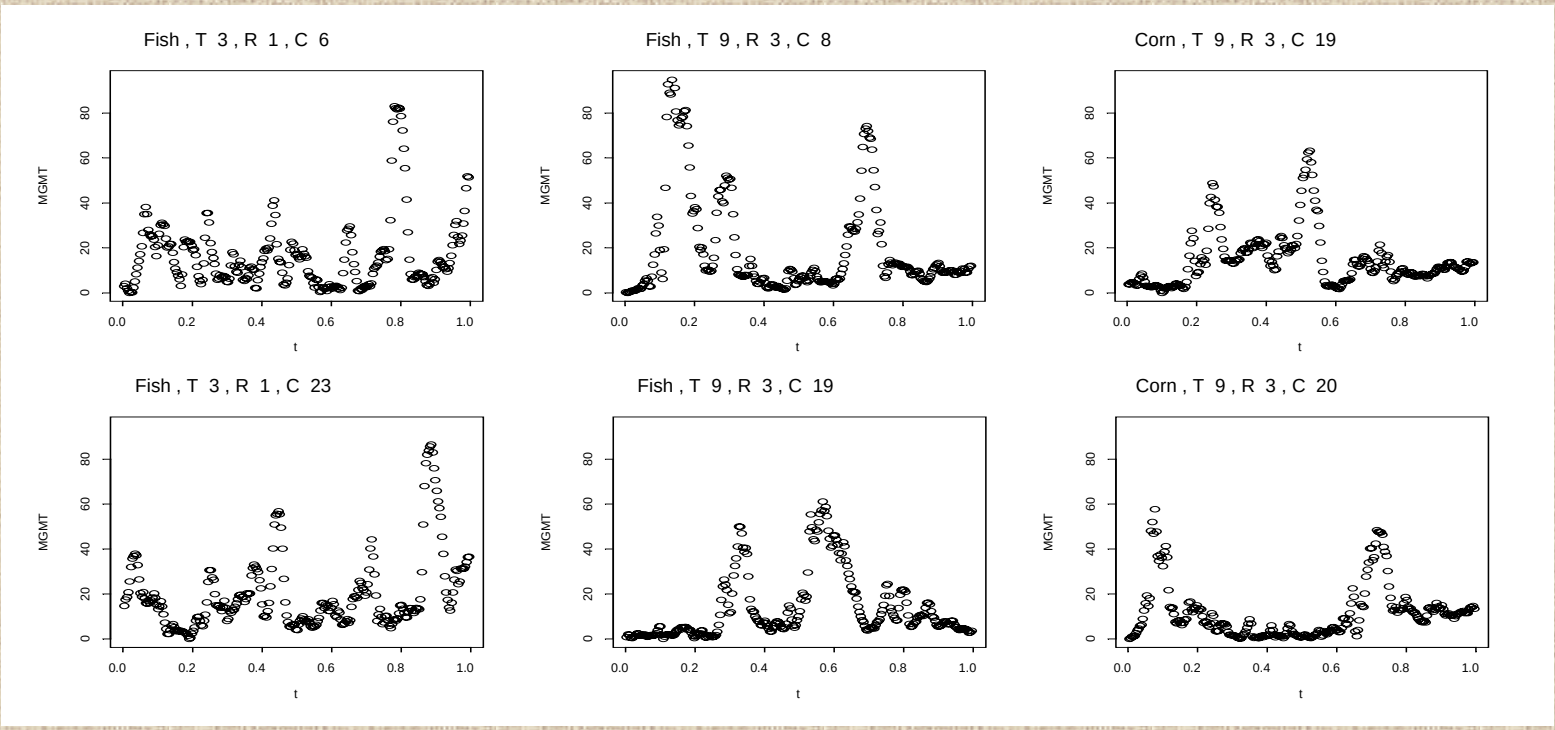
Texas A&M University

http://biostatistics.mdanderson.org/Morris

Functional Data Analysis

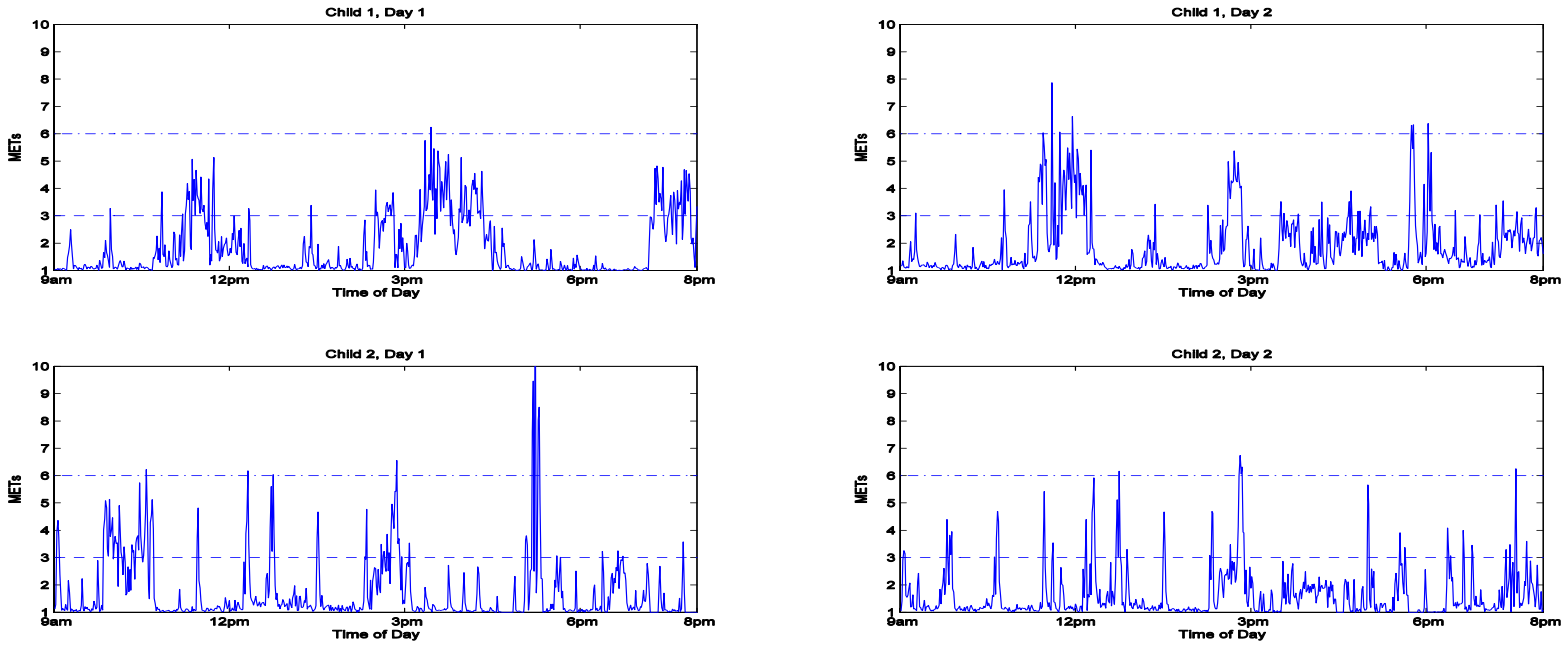
- **Functional Data:**
  - **Ideal units of observation (atoms): curves**
  - **Observed data: curves sampled on fine grid**
- **Increasingly encountered in scientific research**
- **Our Concern Here:**
  - **Functional responses, would like to model relationship with non-functional covariates**
  - **Primary interest: irregular functions with many local features, like peaks.**

Colon Carcinogenesis Data



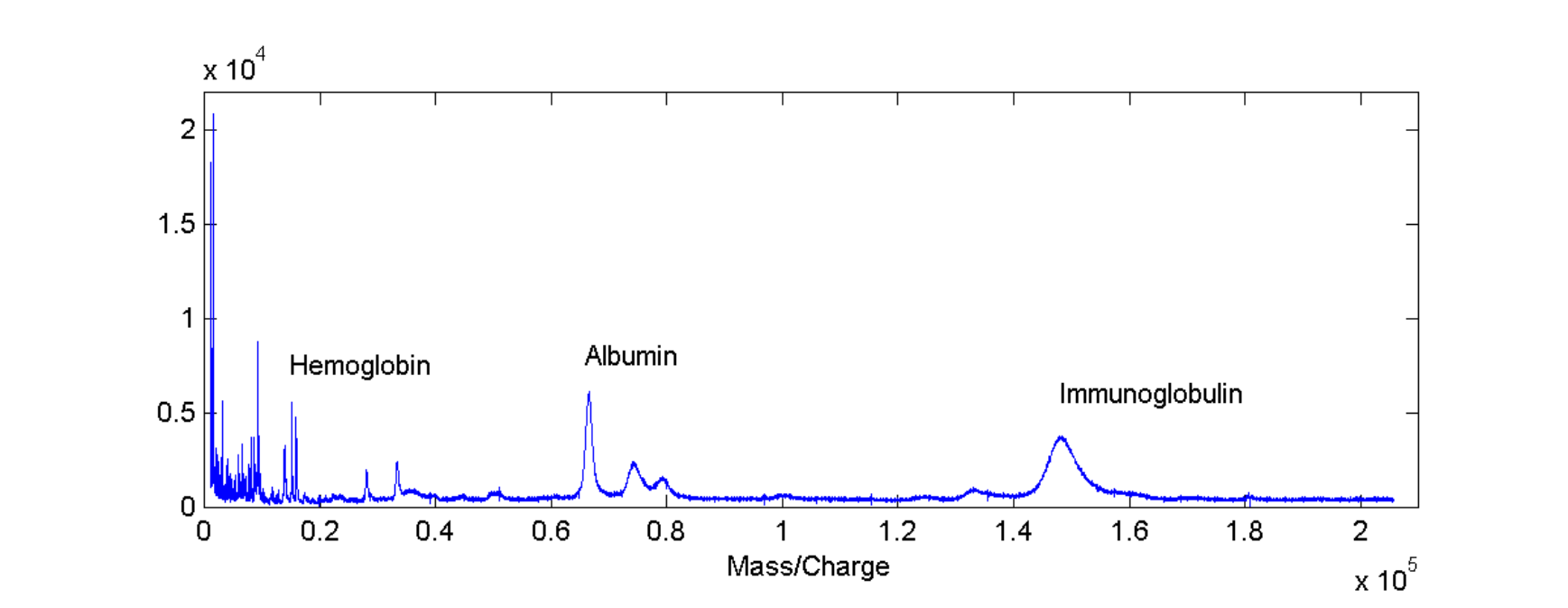
- **Analyzed in Morris, Vannucci, Brown, and Carroll (2003)**
- **30 rats from 2 diets x 5 times, 15 crypts/rat, ~250 obs/crypt**
- **MGMT: Repair enzyme**

Planet Health Study: Accelerometer Data



- **Accelerometer: objectively quantify minute-by-minute activity levels**
- **256 children, 4-8 days/child, 660 obs/day (9am-8pm)**

MALDI-TOF Proteomics Data



- **g(*t*) = intensity of peak at m/z value *t* roughly estimates relative abundance of protein with molecular weight of *t* Daltons.**
- **256 samples (pancreatic cancer/normal), 4 blocks, 20k observations/curve**

Linear Mixed Models

Linear Mixed Model (Laird and Ware, 1982):

$$\underbrace{Y}_{N\times 1} = \underbrace{X}_{N\times p}\overbrace{\beta}^{p\times 1} + \underbrace{Z}_{N\times m}\overbrace{u}^{m\times 1} + \underbrace{e}_{N\times 1}$$

$$u \sim N(0, \overbrace{P}^{m\times m})$$
$$e \sim N(0, \underbrace{R}_{N\times N})$$

- **Fixed effects part, $X\beta$, accommodate a broad class of mean structures, including main effects, interactions, and linear coefficients.**
- **Random effects part, Zu, provide a convenient mechanism for modeling correlation among the N observations.**

Functional Mixed Model (FMM)

Suppose we observe a sample of N curves, $Y_i(t)$, i=1, …, N, on a compact interval $\mathcal{T}$

$$U(t) \sim MGP(P,Q)$$
$$E(t) \sim MGP(R,S)$$

$$\underbrace{Y(t)}_{N \text{ functions}} = \overbrace{X}^{N\times p} \underbrace{B(t)}_{p \text{ functions}} + \overbrace{Z}^{N\times m} \underbrace{U(t)}_{m \text{ functions}} + \underbrace{E(t)}_{N \text{ functions}}$$

- **DEFN: $U(t)$~$MGP(P,Q)$ implies the rows of $P^{-1/2}U(t)$ are ind. mean zero Gaussian Processes with covariance surface $Q(t_1,t_2)$.**
  - **Functional generalization of Matrix Normal (Dawid, 1981).**
  - **Implies Cov$\{U_i(t_1),U_j(t_2)\}=P_{ij} * Q(t_1,t_2)$**

- **P and R are covariance matrices (between-curve)**
- **$Q(t_1,t_2)$ and $S(t_1,t_2)$ are covariance surfaces on $\mathcal{T}\times\mathcal{T}$**

Discrete Version of FMM

Suppose each observed curve is sampled on a common equally-spaced grid of length *T*.

$$\underbrace{Y}_{N\times \mathrm{T}} = \overbrace{X}^{N\times p}\ \underbrace{B}_{p\times \mathrm{T}} + \overbrace{Z}^{N\times m}\ \underbrace{U}_{m\times \mathrm{T}} + \underbrace{E}_{N\times \mathrm{T}}$$

$$U \sim MN(P,Q)$$
$$E \sim MN(R,S)$$

- ***U* and *E* follow the Matrix Normal distn.**
  - ***U*~MN(*P*,*Q*) implies Cov{U_{ij}, $U_{i'j'}$}=$P_{ii'}$*$Q_{jj'}$**
- ***P* and *R* are covariance matrices (*m* × *m* & *N* × *N*)**
- ***Q* and *S* are within-curve covariance matrices (*T* × *T*)**

Functional Mixed Models

- **Key feature of FMM: Does not require specification of parametric form for curves**
- **Much existing literature for nonparametrically modeling functional data is based on kernels or splines.**
- **These methods may not work well for spatially heterogeneous data**

Introduction to Wavelets

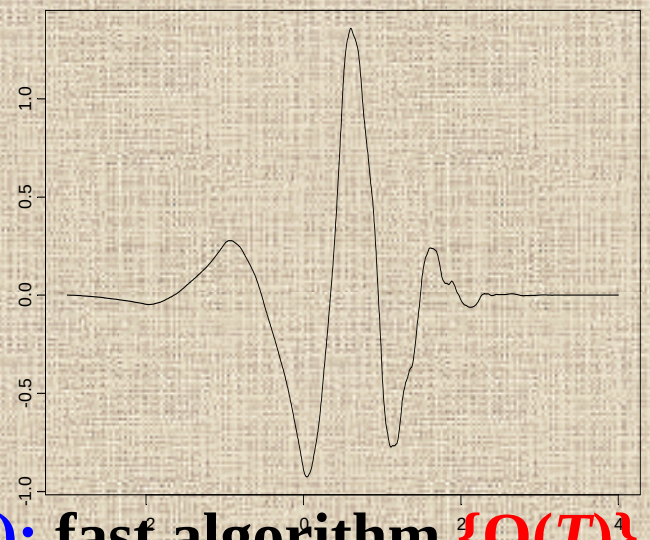
- **Wavelets: families of orthonormal basis functions**

$$g(t) = \sum_{j,k \in \Im} d_{jk} \psi_{jk}(t)$$

$$\psi_{jk}(t) = 2^{-j/2} \psi(2^{-j/2} t - k)$$

$$d_{jk} = \int g(t) \psi_{jk}(t) dt$$

Daubechies Wavelet Function

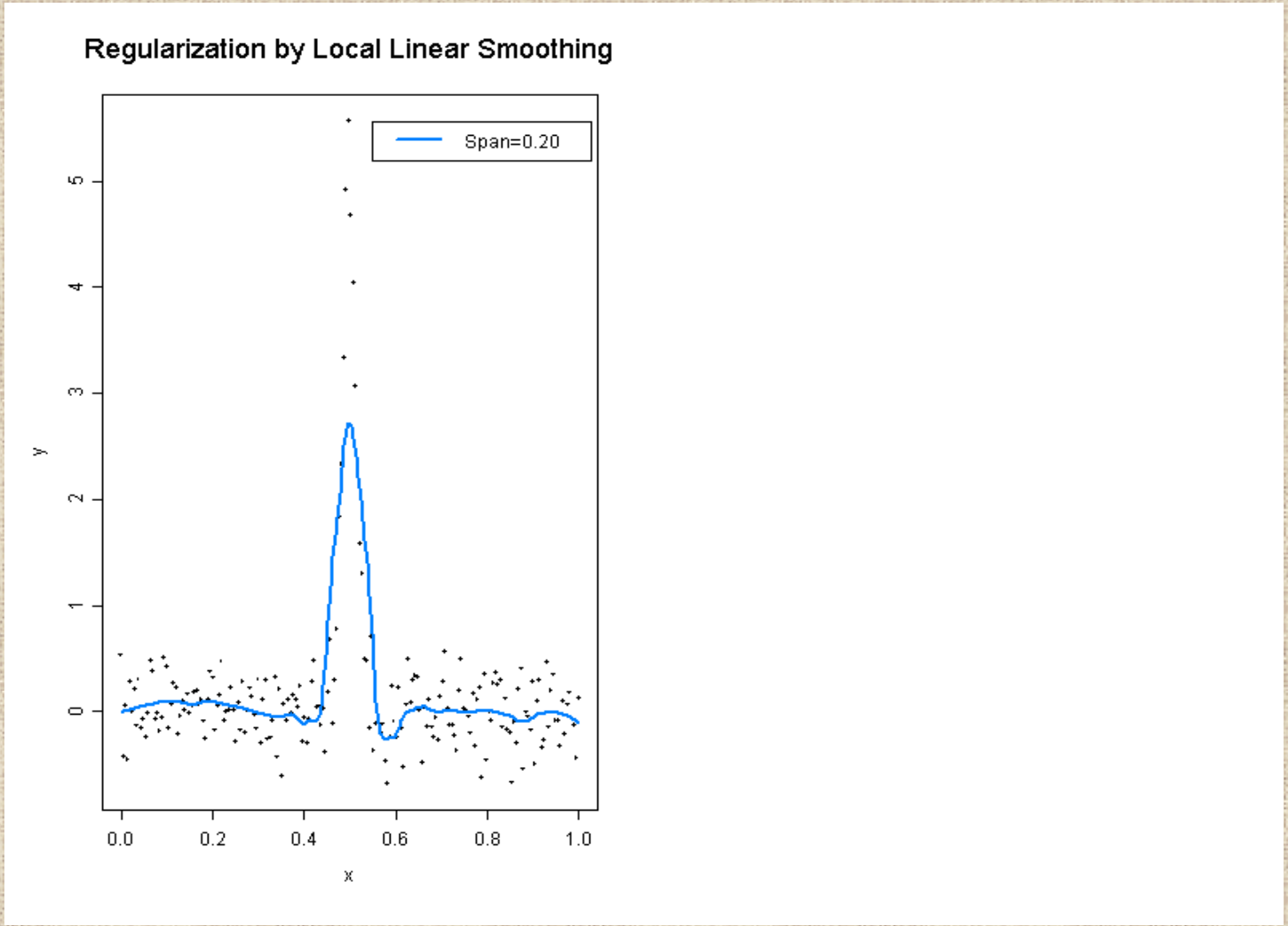


- **Discrete Wavelet Transform (DWT): fast algorithm {O(*T*)} for obtaining *T* empirical wavelet coefficients for curves sampled on equally-spaced grid of length *T*.**
- **Linear Representation:** ***d* = y *W'***
  - ***W'* =*T*-by-*T* orthogonal projection matrix**
  - **Inverse DWT (IDWT):** ***y* = *d W***

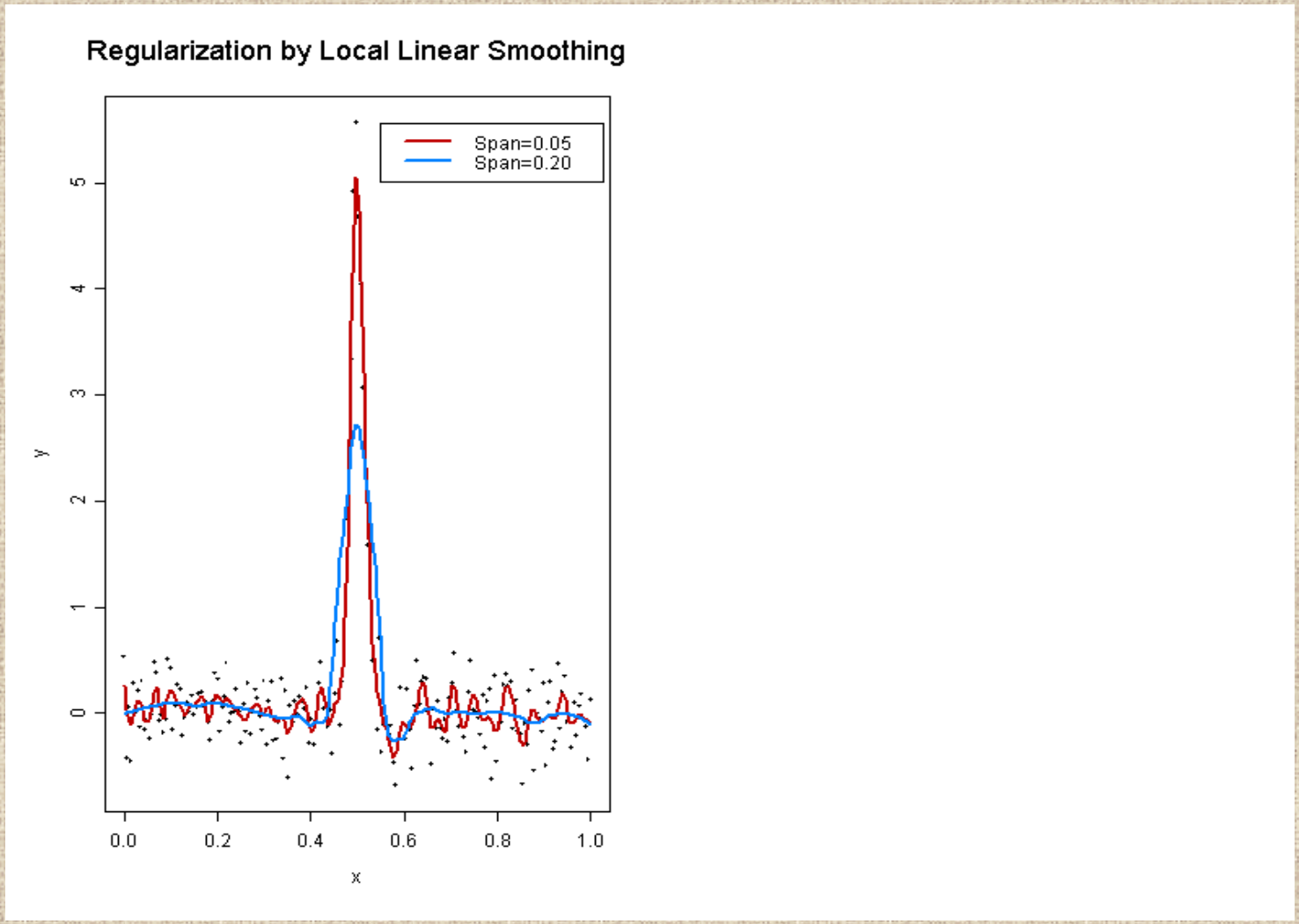
Wavelet Regression

- **Useful properties of wavelets:**
  - **Whitening property**
  - **Compact support**
  - **Parsimonious representation**
- **Wavelet Regression – 3 step process**
  1. **Project data into wavelet space**
  2. **Threshold/shrink coefficients**
  3. **Project back to data space**
- **Yields *adaptively regularized* nonparametric estimates**

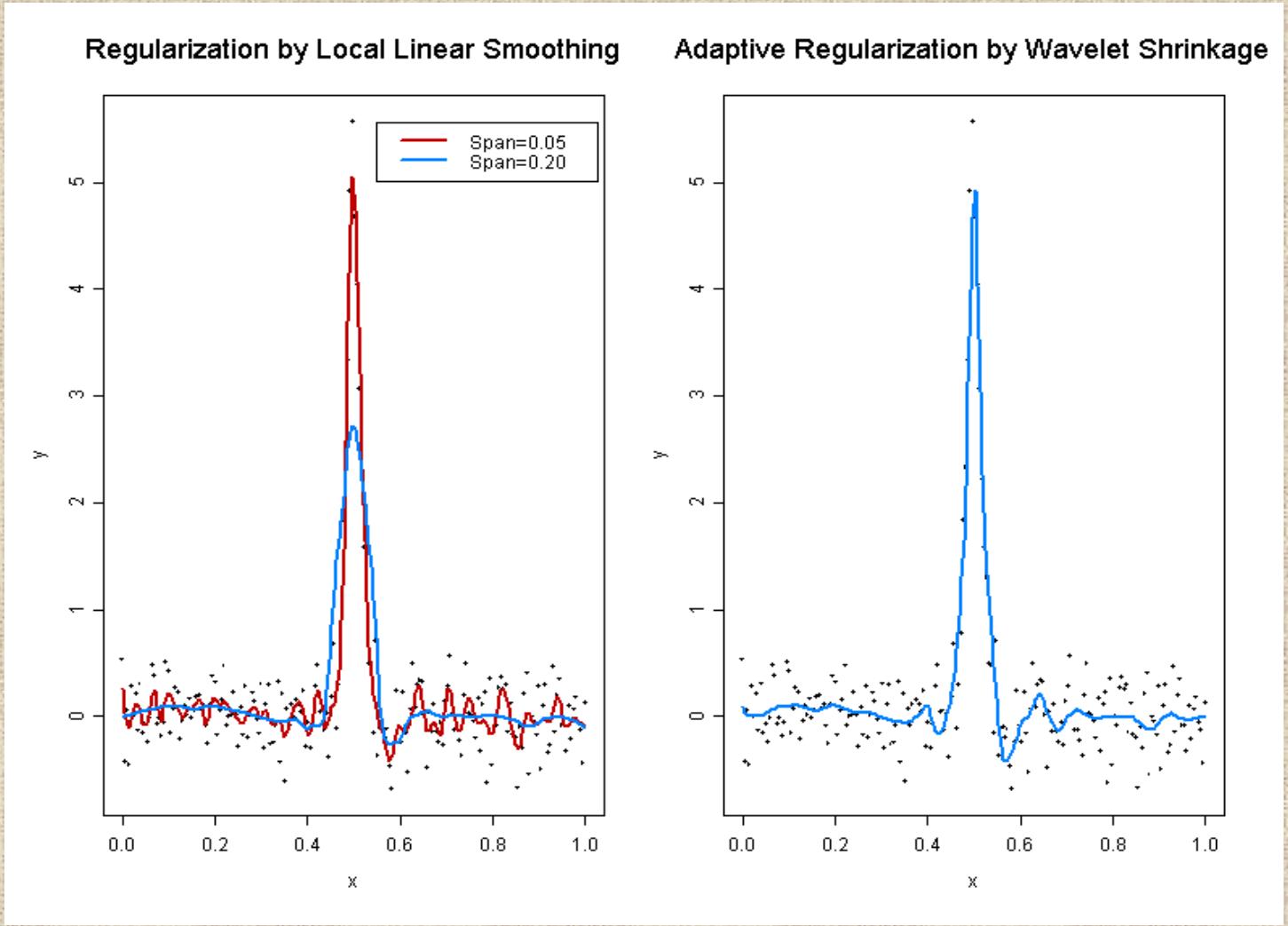
Adaptive Regularization



Adaptive Regularization



Adaptive Regularization



Wavelet-Based FMM:

General Approach

1. **Project observed functions** Y **into wavelet space.**
2. **Fit FMM in wavelet space.**

   (Use MCMC to get posterior samples)
3. **Project wavelet-space estimates (posterior samples) back to data space.**

Wavelet-Based FMM:

General Approach

1. **Project observed functions Y into wavelet space.**

2. **Fit FMM in wavelet space**
   (Use MCMC to get posterior samples)

3. **Project wavelet-space estimates (posterior samples) back to data space.**

Wavelet-Based FMM

1. **Project observed functions** Y **to wavelet space**

- **Apply DWT to rows of** Y **to get wavelet coefficients corresponding to each observed function**

$$\underbrace{D}_{N\times T} = \underbrace{Y}_{N\times T}\ \underbrace{W'}_{T\times T}$$

- **Projects the observed curves into the space spanned by the wavelet bases.**

Wavelet-Based FMM:

General Approach

1. **Project observed functions Y into wavelet space.**
2. **Fit FMM in wavelet space**

   (Use MCMC to get posterior samples)
3. **Project wavelet-space estimates (posterior samples) back to data space.**

Wavelet Space FMM

D : empirical wavelet coefficients for observed curves

Row *i* contains wavelet coefficients for observed curve *i*

Each column double-indexed by wavelet scale *j* and location *k*

$$\underbrace{D}_{N\times \mathrm{T}} = \overbrace{X}^{N\times p}\underbrace{B^*}_{p\times \mathrm{T}} + \overbrace{Z}^{N\times m}\underbrace{U^*}_{m\times \mathrm{T}} + \underbrace{E^*}_{N\times \mathrm{T}}$$

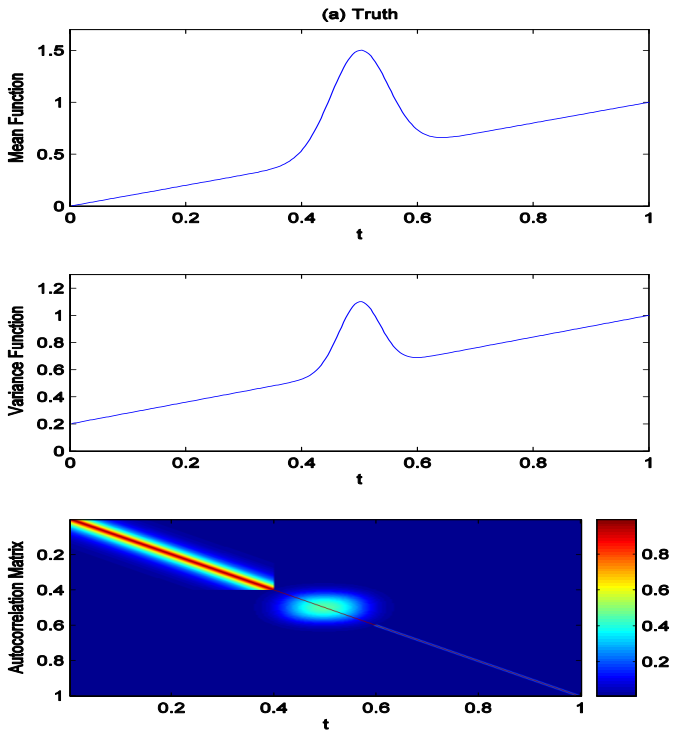
$$U^* \sim MN(P, Q^*)$$
$$E^* \sim MN(R, S^*)$$

- ***B*=BW'*** **&** ***U*=UW'*** **: Rows contain wavelet coefficients for the fixed and random effect functions,**
- ***E*=EW'*** **is the matrix of wavelet-space residuals**
- ***Q*=WQW'*** **and** ***S*=WSW'*** **model the covariance structure between wavelet coefficients for a given function.**
- ***P, Q*, R*** **and** ***S**** **are typically too large to estimate in an unstructured fashion.**

Covariance Assumptions

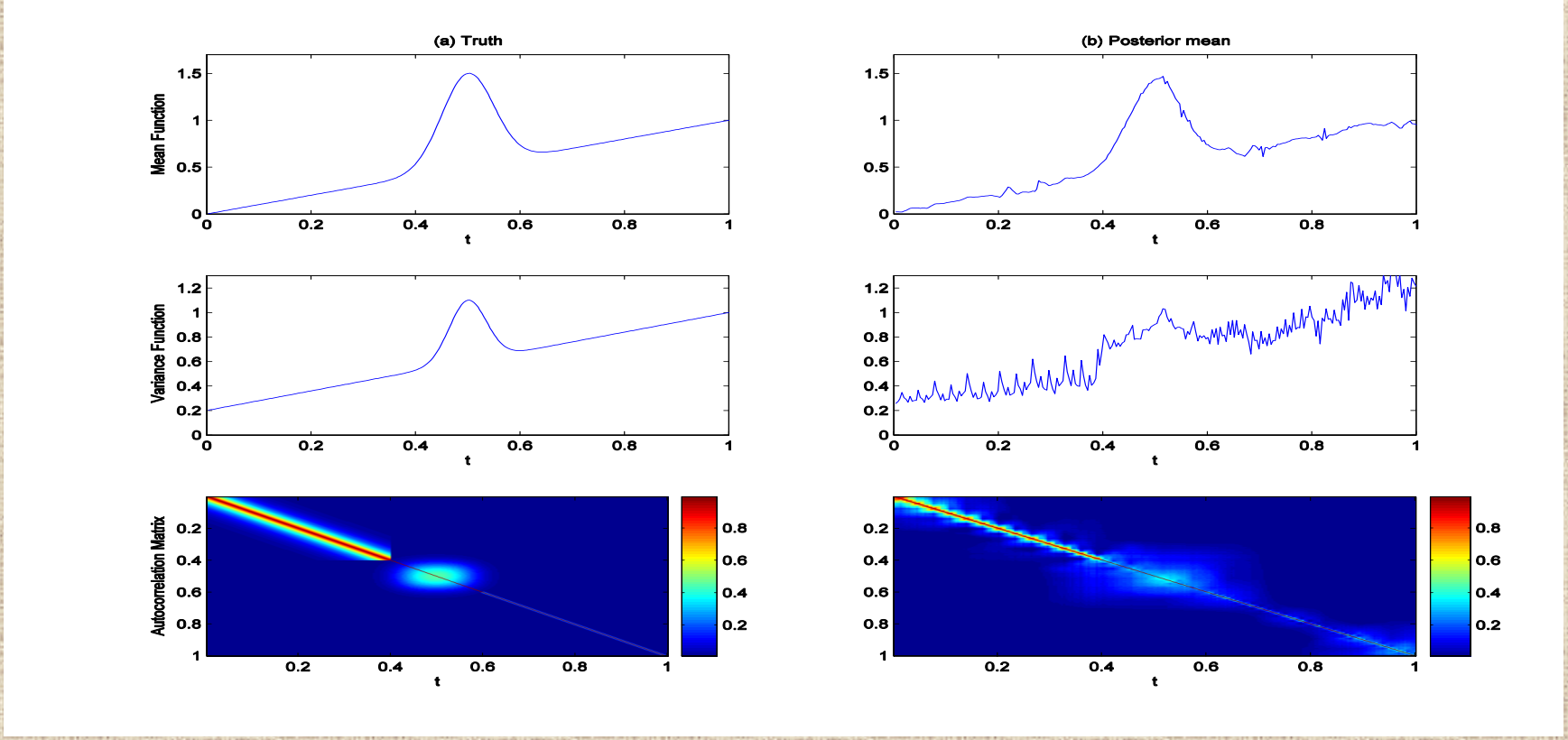
- **We choose parametric structures for *P* and *R* to model the covariance structure between the curves.**
  - **Based on the experimental design**
  - **As in linear mixed models.**
- **We assume the between-wavelet covariance matrices *Q** and *S** are diagonal (working model).**
  - **Wavelet coefficients within given function independent**
  - **Heuristically justified by whitening property of DWT**
  - **Common assumption in wavelet regression**
  - **Is parsimonious in wavelet space (*T* parameters), yet leads to flexible class of covariance structures in data space**

Simulation: Covariance Structure



- **True mean: line plus peak**
- **True variance: increasing in t, with extra var at peak**
- **True autocorrelation: Strong autocorrelation (0.9) at left, weak autocorrelation (0.1) right, extra at peak**

Simulation: Covariance Structure



- **Independence in wavelet space accommodates varying degrees of autocorrelation in data space**
- **Allowing variance components to vary across scale *j* and location *k* accommodates nonstationarities**

Prior Assumptions

Mixture prior on B_{ijk}^*:

$$B_{ijk}^* = \gamma_{ijk}^* N(0, \tau_{ij}) + (1 - \gamma_{ijk}^*)\delta_0$$

$$\gamma_{ijk}^* = \text{Bernoulli}(\pi_{ij})$$

- **Nonlinearly shrinks B_{ijk}^* towards 0, leading to adaptively regularized estimates of $B_i(t)$.**
- **τ_{ij} & π_{ij} are regularization parameters**
  - **Can be estimated from the data using empirical Bayes**
  - **Extend Clyde&George (1999) to functional mixed model**

Model Fitting

- **MCMC to obtain posterior samples of model quantities**
  - **Work with marginal likelihood; U* integrated out;**
- **Let Ω be a vector containing ALL covariance parameters (i.e. for *P*, *Q**, *R*, and *S**).**

MCMC Steps

1. **Sample from $f(B^*|D,\Omega)$:**

   Mixture of normals and point masses at 0 for each *i,j,k*.

2. **Sample from $f(\Omega|D,B^*)$:**

   Metropolis-Hastings steps for each *j,k*

3. **If desired, sample from $f(U^*|D,B^*,\Omega)$:**

   Multivariate normal

Wavelet-Based FMM:

General Approach

1. **Project observed functions Y into wavelet space.**
2. **Fit FMM in wavelet space**
   (Use MCMC to get posterior samples)
3. **Project wavelet-space estimates (posterior samples) back to data space.**

Wavelet-Based FMM

3. Project wavelet-space estimates (posterior samples) back to data space.

- **Apply IDWT to posterior samples of B^* to get posterior samples of fixed effect functions $B_i(t)$ for $i=1,\ldots,p$, on grid t.**
  - **$B=B^*W$**
- **Posterior samples of $U(t)$, P, Q, R, and S are also available, if desired.**
- **Can be used for Bayesian inference/prediction**

Discussion

- **Introduced unified modeling approach for FDA**
  - **Can accommodate very irregular functions**
- **Method based on mixed models; is FLEXIBLE**
  - **Accommodates a wide range of experimental designs**
  - **Addresses large number of research questions**
- **Posterior samples allow Bayesian inference and prediction**
  - **Posterior credible intervals; pointwise or joint**
  - **Predictive distributions for future sampled curves**
  - **Predictive probabilities for classification of new curves**
  - **Bayesian functional inference can be done via Bayes Factors**
- **Since a unified modeling approach is used, all sources of variability in the model propagated throughout inference.**

Discussion

- **Approach is Bayesian. The only informative priors to elicit are regularization parameters, which can be estimated from data using empirical Bayes.**
- **Developed general-use code – reasonably fast and straightforward to use → minimum information to specify is Y, X, Z matrices.**
- **Can be extended to deal with missing data (partially observed functions) or covariate measurement error**
- **Method generalizes to higher dimensional functions, e.g. image data, space/time (fixed domain) data.**

Acknowledgements

- **Co-author: Raymond J. Carroll**
- **Other statistical collaborators on work related to functional mixed models: Phil Brown, Marina Vannucci, Louise Ryan, Brent Coull, Kevin Coombes, Keith Baggerly.**
- **Thanks to Dick Herrick for assistance in optimizing the code for the method, and for converting the Matlab code to C++.**

- **Papers/code/slides can be found on website: http://biostatistics.mdanderson.org/Morris**