John Shamshoian, Nicholas Marco, Damla Şentürk, Shafali Jeste, and Donatello Telesca*

Supplementary Material for the Article: Bayesian Covariance Regression in Functional Data Analysis with Applications to Functional Brain Imaging

1 Appendix S1: Markov-chain monte carlo sampling algorithm

In this section we give a detailed Markov-Chain Monte Carlo (MCMC) algorithm to sample from the posterior. Let N be the number of independent functional responses and assume all response functions are observed on a common grid $T = \{t_1, \ldots, t_n\}$. Let B be an $n \times p$ matrix with $B_{ij} = b_j(t_i)$. Let X be an $N \times r(d_1)$ matrix with row i equal to $\mathbf{b}^x(\mathbf{x}_i)$. Let Y be an $N \times n$ matrix with $Y_{ij} = y_i(t_j)$, so that each row represents one discretized functional response. Let Γ_j be an $N \times N$ diagonal matrix with rth diagonal element equal to η_{rj} for $j = 1, \ldots, k$. In the following, the notation $\Omega = \text{blkdiag}(\Omega_1, \ldots, \Omega_R)$ is shorthand for

$$\Omega = \begin{bmatrix} \Omega_1 & 0 & \cdots & 0 \\ 0 & \Omega_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & \Omega_R \end{bmatrix}.$$

The MCMC transition sequence iterates between the following steps:

1. Update $\boldsymbol{\beta}$:
   Let $\Omega_r = \tau_{1xr}\tilde{K}_r + \tau_{1tr}\tilde{K}$ if $p_r > 1$. Otherwise set $\Omega_r = \tau_{1tr}K$. Construct
   $\Omega = \text{blkdiag}(\Omega_1, \ldots, \Omega_R)$.
   Let $C = \sigma^{-2}X^\top X \otimes B^\top B + \Omega$
   Let $A = \sigma^{-2}\text{vec}[\{B^\top Y^\top - B^\top B(\sum_{j=1}^k \Lambda_j X^\top \Gamma_j)\}X]$
   Sample $\text{vec}(\boldsymbol{\beta}) \sim N(C^{-1}A, C^{-1})$
2. Update Λ_j:
   Let $\Omega_r = \tau_{2xr}\tilde{K}_r + \tau_{2tr}\tilde{K} + \tau^*_{rj}\phi_{rj}$ if $p_r > 1$. Otherwise set $\Omega_r = \tau_{2tr}K + \tau^*_{rj}\phi_{rj}$.
   Construct $\Omega = \text{blkdiag}(\Omega_1, \ldots, \Omega_R)$
   Let $C = \sigma^{-2}X^\top\Gamma_j^2 X \otimes B^\top B + \Omega$
   Let $A = \sigma^{-2}\text{vec}[\{B^\top Y^\top - B^\top B(\boldsymbol{\beta} + \sum_{j' \neq j} \Lambda_{j'} X^\top \Gamma_{j'})\}\Gamma_j X]$
   Sample $\Lambda_j \sim N(C^{-1}A, C^{-1})$
3. Update η_{ij}:
   Let $\eta_i = (\eta_{i1}, \ldots, \eta_{ik})$.
   Let $\ddot{X}_i$ be a $p \times k$ matrix with column j equal to $\Lambda_j \mathbf{b}^x(\mathbf{x}_i)$.
   Let $C = \sigma^{-2}\ddot{X}_i^\top B^\top B\ddot{X}_i + I_k$, where I_k is the $k \times k$ identity matrix.
   Let $A = \sigma^{-2}\ddot{X}_i^\top\{B^\top Y_{i\cdot} - B^\top B\boldsymbol{\beta}\mathbf{b}^x(\mathbf{x}_i)\}$
   Sample $\eta_i \sim N(C^{-1}A, C^{-1})$

John Shamshoian, Nicholas Marco, Damla Şentürk, Donatello Telesca, Department of Biostatistics, University of California, Los Angeles, California, USA, email: jshamsho@gmail.com, ndmarco96@g.ucla.edu, dsenturk@g.ucla.edu, dtelesca@ucla.edu

Shafali Jeste, Division of Neurology and Neurological Institute, Children's Hospital Los Angeles, Los Angeles, USA. email: sjeste@chla.usc.edu

4. Update σ^2:
   Let $A = Y^\top - B\boldsymbol{\beta}X^\top + \sum_{j=1}^{k} B\Lambda_j X^\top \Gamma_j$
   Sample $\sigma^{-2} \sim \text{Gamma}(Nn/2 + a_\epsilon, A \odot A/2 + b_\epsilon)$, where $\odot$ denotes element-wise multiplication.
5. Update τ_{1tr}, τ_{1xr}, τ_{2tr}, τ_{2xr}:
   Sample $\tau_{1tr} \sim \text{Gamma}\{\text{rank}(\tilde{K})/2 - 0.5, \text{vec}(\boldsymbol{\beta}_r)^\top \tilde{K} \text{vec}(\boldsymbol{\beta}_r)/2\}$
   Sample $\tau_{1xr} \sim \text{Gamma}\{\text{rank}(\tilde{K}_r)/2 - 0.5, \text{vec}(\boldsymbol{\beta}_r^\top)\tilde{K}_r \text{vec}(\boldsymbol{\beta}_r)/2\}$ (if $p_r > 1$)
   Sample $\tau_{2trj} \sim \text{Gamma}\{\text{rank}(\tilde{K})/2 - 0.5, \text{vec}(\Lambda_{rj})^\top \tilde{K} \text{vec}(\Lambda_{rj})/2\}$
   Sample $\tau_{2xrj} \sim \text{Gamma}\{\text{rank}(\tilde{K}_r)/2 - 0.5, \text{vec}(\Lambda_{rj}^\top)\tilde{K}_r \text{vec}(\Lambda_{rj})/2\}$ (if $p_r > 1$)
6. Update ϕ_{rj}:
   Let ϕ_{irj} denote the ith diagonal element of ϕ_{rj}.
   Let λ_{irj} be the ith element of $\text{vec}(\Lambda_{rj})$.
   Sample $\phi_{irj} \sim \text{Gamma}(a_\phi + 0.5, \tau_{rj}^* \lambda_{irj}^2/2 + b_\phi)$
7. Update δ_{r1}:
   Let $A = \text{vec}(\Lambda_{r1}^\top)\phi_{r1}\text{vec}(\Lambda_{r1})$
   Let $B = \sum_{j=2}^{k} \tau_{rj}^* \text{vec}(\Lambda_{rj})^\top \phi_{rj} \text{vec}(\Lambda_{rj})$
   Sample $\delta_{r1} \sim \text{Gamma}\{kp_r p/2 + a_{r0}, (A + B)/2 + 1\}$
8. Update δ_{rj}:
   Let $A = \sum_{j'=1}^{k} \tau_{rj'}^{*(j)} \text{vec}(\Lambda_{rj'})^\top \phi_{rj'} \text{vec}(\Lambda_{rj})$, where $\tau_{rj'}^{*(j)} = \tau_{rj'}^*$ if $j \neq j'$ and 1 otherwise.
   Sample $\delta_{rj'} \sim \text{Gamma}\{p_r p(k - j' + 1)/2 + a_{r1}, A/2 + 1\}$
9. Update a_{r0}:
   Let $\text{Gamma}^*(x, a, b)$ denote the Gamma density evaluated at x with shape a and rate b.
   Let $\phi(x)$ denote the standard normal cumulative distribution function evaluated at x. Sample candidate $a_{r0}^* \sim N(a_{r0}, 1)$ until $a_{r0}^* > 0$.
   Compute $A = \dfrac{\text{Gamma}^*(\delta_{r1}, a_{r0}^*, 1) \cdot \text{Gamma}^*(a_{r0}^*, 2, 1) \cdot \phi(a_{r0})}{\text{Gamma}^*(\delta_{r1}, a_{r0}, 1) \cdot \text{Gamma}^*(a_{r0}, 2, 1) \cdot \phi(a_{r0}^*)}$
   Sample $U \sim \text{Unif}(0, 1)$
   If $U \leqslant A$, accept candidate a_{r0}^*.
10. Update a_{r1}:
    Let $\delta_{r2}^* = \prod_{j=2}^{k} \delta_{rj}$
    Let $\text{Gamma}^*(x, a, b)$ denote the Gamma density evaluated at x with shape a and rate b.
    Let $\phi(x)$ denote the standard normal cumulative distribution function evaluated at x. Sample candidate $a_{r1}^* \sim N(a_{r1}, 1)$ until $a_{r1}^* > 0$.
    Compute $A = \dfrac{\text{Gamma}^*(\delta_{r2}^*, a_{r1}^*, 1) \cdot \text{Gamma}^*(a_{r1}^*, 2, 1) \cdot \phi(a_{r1})}{\text{Gamma}^*(\delta_{r2}^*, a_{r1}, 1) \cdot \text{Gamma}^*(a_{r1}, 2, 1) \cdot \phi(a_{r1}^*)}$
    Sample $U \sim \text{Unif}(0, 1)$
    If $U \leqslant A$, accept candidate a_{r1}^*.
11. Update missing values of Y:
    Suppose $y_i(t_j)$ is missing.
    Let $\mu = \mathbf{b}(t_j)\boldsymbol{\beta}\mathbf{b}^x(\mathbf{x}_i) + \sum_{j=1}^{k} \mathbf{b}(t_j)\Lambda_j \mathbf{b}^x(\mathbf{x}_i)\eta_{ij}$
    Sample $y_i(t_j) \sim N(\mu, \sigma^2)$

For convenience, Table 1 contains a list of notation with explanation and dimensions.

2 Appendix S2: Additional details on posterior inference

In this section we give details on post-processing MCMC samples to obtain eigenfunctions as in a usual FPCA. Latent functional factors $\psi_j(t, \mathbf{x})$ are not orthonormal and cannot be directly interpreted as function principal components. To orthonormalize $\psi_j(t, \mathbf{x})$, one may evaluate posterior draws of $c(t, t', \mathbf{x})$ on an arbitrary domain $t, t' \in \mathcal{T}$ followed by a spectral decomposition to yield orthonormal modes of variation.

Tab. 1: Notation with description and dimensions.

Name	Description	Dimension
$y_i(t)$	ith functional response at t	
$\tilde{y}_i(t)$	ith latent trajectory uncontaminated without measurement error	
$\mathbf{x}_i$	Covariate vector for the ith functional response	$d_1 \times 1$
$\mathbf{x}_{ir}$	rth group of covariates associated with the ith functional response	
$\mathbf{b}^r(\mathbf{x}_{ir})$	rth group of covariates expanded into basis functions evaluated at $\mathbf{x}_{ir}$	$p_r \times 1$
$\mathbf{b}^x(\mathbf{x}_i)$	Concatenated $\mathbf{b}^r(\mathbf{x}_{ir})$, $r = 1, \ldots, R$	$r(d_1) \times 1$
p_r	# of basis functions for rth covariate group expansion	
$r(d_1)$	Total number of covariate basis functions used equal to $\sum_{r=1}^{R} p_r$	
$\mathbf{b}(t)$	Basis functions for functional dimension evaluated at t	$p \times 1$
$\mu(t, \mathbf{x}_i)$	Covariate-adjusted functional mean equal to $\mathbf{b}(t)^\top \beta \mathbf{b}^x(\mathbf{x}_i)$	
$c(t, t', \mathbf{x}_i)$	Covariate-adjusted functional covariance	
β_r	Fixed effect parameter matrix associated with $\mathbf{b}^r(\mathbf{x}_{ir})$	$p \times p_r$
β	Fixed effect parameter submatrix associated with $\mathbf{b}^x(\mathbf{x}_i)$	$p \times r(d_1)$
Λ_{rj}	jth latent factor submatrix associated with $\mathbf{b}^r(\mathbf{x}_{ir})$	$p \times p_r$
Λ_j	jth latent factor matrix associated with $\mathbf{b}^x(\mathbf{x}_i)$	$p \times r(d_1)$
$\psi_j(t, \mathbf{x}_i)$	jth latent covariate-adjusted functional factor	
$r_i(t, \mathbf{x}_i)$	Covariate-adjusted random function for ith functional response	
δ_{rj}	Global shrinkage parameter for rth covariate group of Λ_j	
ϕ_{lrj}	Local shrinkage for rth covariate group of Λ_j	
τ^*_{rj}	Cumulative product shrinkage for rth covariate group of Λ_j	
τ_{1tr}, τ_{2jtr}	Smoothing parameter over functional arugment	
τ_{1xr}, τ_{2jxr}	Smoothing parameter over rth covariate argument	
$\tilde{K}, K$	Smoothing penalty matrix over functional argument	$p \cdot p_r \times p \cdot p_r$
$\tilde{K}_r$	Smoothing penalty matrix over covariate argument	$p \cdot p_r \times p \cdot p_r$

However, this procedure may be computationally intensive because $\mathcal{T}$ is in theory infinite-dimensional. To alleviate this burden, we borrow methods from [1] to obtain orthonormal modes of variation by performing a spectral decomposition on a much lower dimensional matrix. To obtain $\tilde{\psi}^{(m)}(t, \mathbf{x})$, the mth draw of orthonormal posterior modes of variation, we

1. Compute $p \times p$ matrix Ψ, where $\Psi_{ij} = \int_{\mathcal{T}} b_i(t) b_j(t) dt$.
2. Compute $\tilde{\Lambda}_{\mathbf{x}} = \sum_{j=1}^{k} \Lambda_j \mathbf{b}^x(\mathbf{x}) \mathbf{b}^x(\mathbf{x})^\top \Lambda_j^\top$
3. Set $\tilde{\Lambda}_{\mathbf{x}} = \Psi^{1/2} \tilde{\Lambda}_{\mathbf{x}} \Psi^{1/2}$
4. Perform a spectral decomposition on $\tilde{\Lambda}_{\mathbf{x}}$ to obtain $\tilde{\psi}_j(\mathbf{x})$
5. Set $\tilde{\psi}_j(t, \mathbf{x}) = \mathbf{b}(t) \Psi^{-1/2} \tilde{\psi}_j(\mathbf{x})$
6. If desired, postprocessed principal scores are equal to $\int_{\mathcal{T}} \tilde{\psi}_j(t, \mathbf{x}_i) r_i(t, \mathbf{x}_i)\, dt$. For sparse functional response settings, principal scores are based on conditional expectation [2].

Step 1 may be pre-computed before any posterior samples are evaluated. We compare the mth posterior eigenfunction to the running average in l_2 norm. If the norm is smaller after multiplying the estimated eigenfunction by -1, we multiply the estimated eigenfunction by -1. This process ensures all posterior samples of eigenfunctions are oriented correctly so that means and interval calculations are sensible.

Up until this point we have only discussed pointwise credible intervals. However, pointwise intervals are not appropriate for making probabilistic statements of an entire function. Simultaneous credible bands are more appropriate for entire function inference and are easily computable using posterior samples assuming normality as detailed in [3] and [4]. Suppose we desire a simultaneous credible band of some functional $f(\cdot)$ observed at points $t_1, \ldots, t_N$. Let $\mu_f(t_i)$ and $\sigma_f(t_i)$ be the estimated pointwise mean and standard deviation respectively. Let $f^{(m)}(\cdot)$ be the mth realization of $f(\cdot)$ drawn from the posterior out of M total

samples. By assuming approximate normality and deriving the $(1-\alpha)$ sample quantity c_b of

$$\max_{i=1,\ldots,N} \left| \frac{f^{(m)}(t_i) - \mu_f(t_i)}{\sigma_f(t_i)} \right|, \ m = 1, \ldots, M$$

a simultaneous credible region is given by the hyperrectangular

$$\big[\mu_f(t_i) - c_b \cdot \sigma_f(t_i), \mu_f(t_i) + c_b \cdot \sigma_f(t_i)\big], \ i = 1 \ldots, N$$

In our implementation on github we include the option to compute pointwise and simultaneous bands for latent subject-specific functions, covariate-adjusted means, and covariate-adjusted eigenfunctions.

3 Appendix S3: Additional details on case studies

In this section we provide details setting up the design matrices and penalty matrices for the ASD case study [5].We also discuss computing low dimensional covariance summaries $g(t, \mathbf{x})$ and associated inference.

The ASD case study [5] has four covariates excluding an intercept term: diagnostic group, age of the child, signal ratio of the child, and a group by age interaction term. In terms of notation used in Web Appendix A, the design matrix X has dimension 97 by 17. The first column of X is an intercept column with repeating ones. The second column is a group indicator with a 1 if the child is diagnosed with ASD and 0 otherwise. The next five columns expand age of child by B-splines. The next five columns expand age of child of child by B-splines row-wise multiplied by 0 if the child is in the TD diagnostic group. The last five columns expand the signal ratio of the child by B-splines. In terms of notation from Sections 2 and 3 of the main manuscript, $\mathbf{x}_1 = \{\text{Intercept}\}$, $\mathbf{x}_2 = \{\text{Group}\}$, $\mathbf{x}_3 = \{\text{Age}\}$, $\mathbf{x}_4 = \{\text{Group by age interaction}\}$, and $\mathbf{x}_5 = \{\text{Signal Ratio}\}$. We also expand the frequency dimension into a set of twelve B-splines. The frequency dimension is associated with a discrete second order 12×12 penalty matrix K with rank 10. The age dimension is associated with a discrete second order 5×5 penalty matrix K_{age} with rank 3. Let M^- denote the generalized inverse of a matrix M. The 12×1 coefficient matrices $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ have priors $N\{0, (\tau_{1t1}K)^-\}$ and $N\{0, (\tau_{1t2}K^-)\}$. The 12×5 coefficient matrices $\boldsymbol{\beta}_3$ and $\boldsymbol{\beta}_4$ have priors $N\{0, (\tau_{1t3}\tilde{K} + \tau_{1x3}\tilde{K}_{\text{age}})^-\}$ and $N\{(0, (\tau_{1x4}\tilde{K}_{\text{age}} + \tau_{1t4}\tilde{K})^-\}$ after vectorization. Here $\tilde{K}_{\text{age}} = K_{\text{age}} \otimes I_{12\times 12}$ and $\tilde{K} = I_{5\times 5} \otimes K$. Priors for Λ_{rj} are identical after replacing τ_{1tr} and τ_{1xr} by τ_{2trj} and τ_{2xrj}.

As a reminder, the low dimensional summaries $g_r(t)$ are designed to help users infer which covariates the functional covariance depends upon. To calculate the affect of age on the covariance function, we approximated the integral by using a sequence of ages from 40 months to 120 months, increasing by 5 months. The integral was then numerically estimated using these points. Similarly the affect of signal ratio on the covariance function was calculated by approximating the integral using a sequence of signal ratios from 0 to 1, increasing by 0.05. The low dimensional summaries can be visualized in Figure 5 in the main manuscript.

Figure 1 displays the top 3 eigenfunctions associated with the plots in Figure 3 of the main manuscript. For TD and ASD, where the signal ratio equal to one, we can see that there is significant phase variation in the location of the alpha peak. Meanwhile, for individuals with signal ratio equal to zero, we can see that most of the variation is in the lower range of frequency. Table 2 depicts the proportion of variation of the top 3 eigenfunctions.

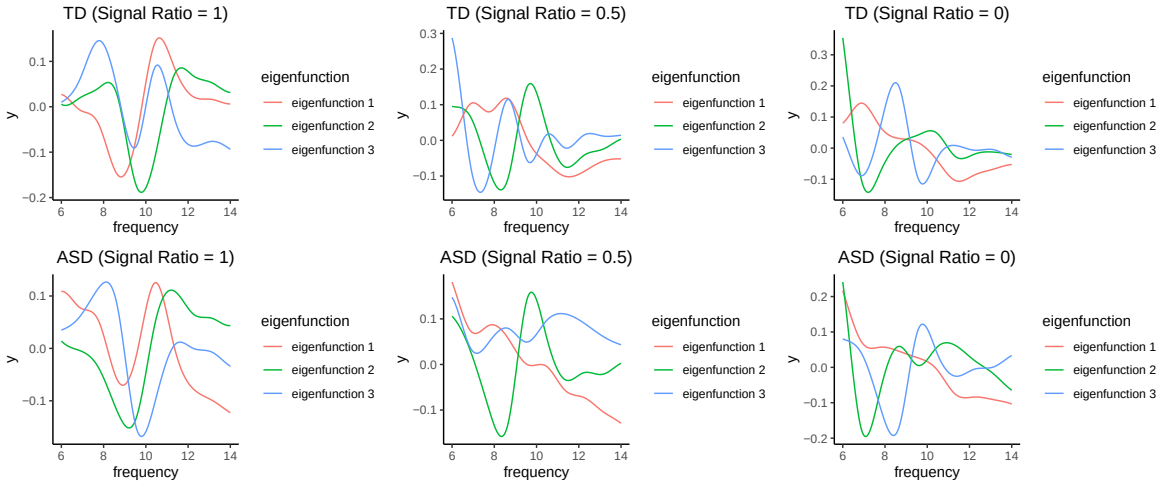

Fig. 1: Posterior mean of the top 3 leading eigenfunctions of the covariance surfaces depicted in Figure 3 of the main manuscript.

	Eigenfunction 1	Eigenfunction 2	Eigenfunction 3
TD (Signal Ratio = 1)	**21.6%**	**9.4%**	**4.5%**
TD (Signal Ratio = 0.5)	**9.7%**	**6.3%**	**3.1%**
TD (Signal Ratio = 0)	**22.5%**	**9.6%**	**6.2%**
ASD (Signal Ratio = 1)	**24.2%**	**12.7%**	**8.2%**
ASD (Signal Ratio = 0.5)	**17.5%**	**6.4%**	**3.4%**
ASD (Signal Ratio = 0)	**32.8%**	**8.8%**	**6.1%**

Tab. 2: Proportion of variance explained by the top 3 eigenfunctions for the plots in Figure 1.

4 Appendix S4: Simulation study and data generation details

We generate data under two scenarios. The first scenario has heteroscedasticity depending on a covariate x generated as independent $U(0,1)$, and the second scenario imposes no such relationship between heteroscedasticity and covariates. Both scenarios consider observations on a uniform grid, $t \in [0,1]$, and a covariate-dependent mean surface. We fit data generated in both scenarios by (1) the proposed method, adjusting for covariate-dependent heteroscedasticity and (2) foregoing this adjustment and assuming the covariate only impacts the mean structure. We generate data in both scenarios by simulating from the model. We generate 300 datasets per scenario and true model parameters are kept fixed over each dataset. We include two sample sizes $N = 100$ and $N = 400$ to numerically verify mean and covariance point estimation convergence.

We begin by using a p-spline of dimension 10 to expand the functional arguement t and a p-spline of dimension 5 (not including an intercept) to expand the covariate argument x. According to notation from Section 2 of the main manuscript, $\mathbf{b}(t) = (b_1(t), \ldots, b_{10}(t))^\top$, $\mathbf{b}^2(x) = (b_1^2(t), \ldots, b_5^2(t))^\top$, and $\mu(t,x) = f_1(t) + f_2(t,x) = \mathbf{b}(t)\,\boldsymbol{\beta}_1 + \mathbf{b}(t)\,\boldsymbol{\beta}_2\,\mathbf{b}^2(x)$. The model parameters $\boldsymbol{\beta}_1$ and $\text{vec}(\boldsymbol{\beta}_2)$ have prior mean zero and precision $\Omega_1 + \epsilon \cdot I_{10\times 10}$ and $\Omega_2 + \epsilon \cdot I_{50\times 50}$ respectively. The nugget term $\epsilon \cdot I$ is used to simulate from a full-rank multivariate normal and in our experiments we set $\epsilon = .1$. To ensure smoothness of the resulting response functions $y_i(t)$, we set $\Omega_1 = \tau_{1t1} K$ and $\Omega_2 = \tau_{1x2}(K_2 \otimes I_{10\times 10}) + \tau_{1t2}(I_{6\times 6} \otimes K)$, where

K and K_2 are second order discrete penalty matrices associated with p-splines. We use $k = 4$ latent functional factors for the random component $r_i(t,x)$, so $r_i(t,x) = \sum_{j=1}^{4} \psi_j(t,x)$. Under data-generating scenario 1, we set $\psi_j(t,x) = l_{1j}(t) + l_{2j}(t,x)$. Under data-generating scenario 2, we set $\psi_j(t,x) = l_{1j}(t)$. Expanding out the terms, $l_{1j}(t) = \mathbf{b}(t)^\top \Lambda_{1j}$ and $l_{2j}(t,x) = \mathbf{b}(t)\,\Lambda_{2j}\,\mathbf{b}^2(x)$. The model parameters Λ_{1j} and Λ_{2j} prior mean zero and prior precision Γ_{j1} and Γ_{j2} respectively. We set $\Gamma_{j1} = \tau_{2t1}K + \tau^*_{1j}$ and $\Gamma_{j2} = \tau_{2x2}(K_2 \otimes I_{10\times 10}) + \tau_{2t2}(I_{6\times 6} \otimes K) + \tau^*_{2j}$. Smoothing and shrinkage parameters are $\tau_{1t1} = \tau_{1t2} = \tau_{2t1} = \tau_{2t2} = 10$, $\tau_{1x2} = \tau_{2x2} = 100$, $\tau^*_{1j} = 2^j$, and $\tau^*_{2j} = 4^j$. Measurement error is $\varphi^2 = .20^2$ and functional data $y_i(t)$ is generated according to Equation (1) in the main manuscript. The functional argument t is evaluated at 100 equally spaced locations within the unit interval and x_i is equal to $(i-1)/(N-1)$.

The methods operating characteristics are assessed by monitoring coverage, average credible interval width, and relative integrated squared error (RISE). Let $t_1, \ldots, t_{100}$ be 100 uniformly placed points along the domain of the functional argument, $x^*_1, \ldots, x^*_{10}$ be uniformly placed points along the interval $[0.1, 0.9]$, and $\tilde{y}_i(t)$ denote the latent response of $y_i(t)$ uncontaminated with random error. Let $\mathbb{1}(\cdot)$ be the indicator function that returns 1 when the argument is true and 0 otherwise. The coverage of $\mu(t,x)$ is $\frac{1}{1000}\sum_{i=1}^{10}\sum_{m=1}^{100}\mathbb{1}\{\mu(t_m,x^*_i) \in \widehat{I}_{mi}\}$, the coverage of $c(t,t',x)$ is $\frac{1}{50500}\sum_{i=1}^{10}\sum_{m=1}^{100}\sum_{m'=1}^{m}\mathbb{1}\{c(t_m,t_{m'},x^*_i) \in \widehat{I}_{mm'i}\}$, and the coverage of $\tilde{y}_i(t)$ is $\frac{1}{N\times 100}\sum_{i=1}^{N}\sum_{m=1}^{100}\mathbb{1}\{\tilde{y}_i(t_m) \in \widehat{I}_{y,mi}\}$. Here $\widehat{I}_{mi}$ is the nominal 95% posterior interval about $\mu(t_m,x_i)$ and $\widehat{I}_{mm'i}$ is the nominal 95% posterior interval about $c(t_m,t_{m'},x_i)$, and $\widehat{I}_{y,mi}$ is the nominal 95% posterior interval about $\tilde{y}_i(t)$. The RISE of $\mu(t,x)$ is $100 \cdot \frac{\sum_{i=1}^{10}\int_{\mathcal{T}}\{\hat{\mu}(t,x^*_i)-\mu(t,x^*_i)\}^2\,dt}{\sum_{i=1}^{10}\int_{\mathcal{T}}\mu(t,x^*_i)^2\,dt}$, the RISE of $c(t,t',x)$ is $100 \cdot \frac{\sum_{i=1}^{10}\int_{\mathcal{T}}\int_{\mathcal{T}}\{\hat{c}(t,t',x^*_i)-c(t,t',x^*_i)\}^2\,dt\,dt'}{\int_{\mathcal{T}}\int_{\mathcal{T}}\sum_{i=1}^{10}c(t,t',x^*_i)^2\,dt\,dt'}$, and the RISE of $\tilde{y}_i(t)$ is $100 \cdot \frac{\sum_{i=1}^{N}\int_{\mathcal{T}}\{\hat{y}_i(t)-\tilde{y}_i(t)\}^2\,dt}{\sum_{i=1}^{N}\int_{\mathcal{T}}\tilde{y}_i(t)^2\,dt}$. In practice we use the trapezoidal rule to evaluate all integrals. We also compute average credible band width for $\mu(t,x)$, $c(t,t',x)$ and $\tilde{y}_i(t)$.

References

[1] Ana M Aguilera and MC Aguilera-Morillo. Penalized pca approaches for b-spline expansions of smooth functional data. *Applied Mathematics and Computation*, 219(14):7805–7819, 2013.

[2] Fang Yao, Hans-Georg Müller, and Jane-Ling Wang. Functional data analysis for sparse longitudinal data. *Journal of the American statistical association*, 100(470):577–590, 2005.

[3] Tatyana Krivobokova, Thomas Kneib, and Gerda Claeskens. Simultaneous confidence bands for penalized spline estimators. *Journal of the American Statistical Association*, 105(490):852–863, 2010. 10.1198/jasa.2010.tm09165. URL https://doi.org/10.1198/jasa.2010.tm09165.

[4] Ciprian M Crainiceanu, David Ruppert, Raymond J Carroll, Adarsh Joshi, and Billy Goodner. Spatially adaptive bayesian penalized splines with heteroscedastic errors. *Journal of Computational and Graphical Statistics*, 16(2):265–288, 2007.

[5] Abigail Dickinson, Charlotte DiStefano, Damla Senturk, and Shafali Spurling Jeste. Peak alpha frequency is a neural marker of cognitive function across the autism spectrum. *European Journal of Neuroscience*, 47(6):643–651, 2018.