Biometrics , 1–7

Web-based Supplementary Materials for "Structured Functional Principal Component Analysis" by

Haochang Shou

Department of Biostatistics and Epidemiology, University of Pennsylvania Perelman School of Medicine

email: hshou@mail.med.upenn.edu

and

Vadim Zipunnikov, Ciprian M. Crainiceanu

Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health

and

Sonja Greven

Department of Statistics, Ludwig-Maximilians-Universität München

APPENDIX A: Principal Scores for Three-way Nested and Two-way Crossed Designs

Assuming noise-free scenarios and follow Zipunnikov et al. (2011), we provide details of calculating principal scores in three-way nested model (N3) and two-way crossed design (C2) as listed in Table 1 of the paper.

For model (N3) $Y_{ijk}(t) = X_i(t) + U_{ij}(t) + W_{ijk}(t), i = 1, 2, \cdots, I; j = 1, 2, \cdots, J_i; k = 1, 2, \cdots, n_{ij}$, suppose we have balanced design $n_{i1} = n_{i2} = \cdots = n_{iJ_i}$, and let $\mathbf{Y}_{ijk} = \{Y_{ijk}(t_1), Y_{ijk}(t_2), \cdots, Y_{ijk}(t_p)\}'$. $\mathbf{Y}_i = \{\mathbf{Y}_{i11}, \mathbf{Y}_{i12}, \cdots, \mathbf{Y}_{i1n_{i1}}, \cdots, \mathbf{Y}_{iJ_i1}, \mathbf{Y}_{iJ_i2}, \cdots, \mathbf{Y}_{iJ_in_{iJ_i}}\}$ is the $p \times n_{i\cdot}$ matrix formed by stacking the functions side-by-side. Let $\widetilde{\mathbf{Y}}_i = \text{vec}\{\mathbf{Y}_i\}$, then (N3) can be expressed in matrix notations as $\widetilde{\mathbf{Y}}_i = \mathbf{B}\mathbf{u}_i$, where $\mathbf{B} = [\mathbf{B}_X|\mathbf{B}_U|\mathbf{B}_W] = \begin{pmatrix} \mathbf{\Phi}_X & \mathbf{\Phi}_U & & \mathbf{\Phi}_W & & & \\ \vdots & \vdots & & & \ddots & & \\ \mathbf{\Phi}_X & \mathbf{\Phi}_U & & & & \mathbf{\Phi}_W & \\ \vdots & & \ddots & & & & \ddots \\ \mathbf{\Phi}_X & & \mathbf{\Phi}_U & & & \mathbf{\Phi}_W & \\ \vdots & & \vdots & & & & \ddots \\ \mathbf{\Phi}_X & & \mathbf{\Phi}_U & & & & \mathbf{\Phi}_W \end{pmatrix}$, $\mathbf{B}_X = \mathbf{1}_{n_{i\cdot}} \otimes \mathbf{\Phi}_X$, $\mathbf{B}_U = \mathbf{I}_{J_i} \otimes (\mathbf{1}_{n_{ij}} \otimes \mathbf{\Phi}_U)$, $\mathbf{B}_W = \mathbf{I}_{n_{i\cdot}} \otimes \mathbf{\Phi}_W$ and $\mathbf{u}_i = (\boldsymbol{\xi}_i^X, \boldsymbol{\xi}_i^U, \boldsymbol{\xi}_i^W)'$. Here $\boldsymbol{\xi}_i^X = (\xi_{i1}{}^X, \cdots, \xi_{iN_1}^X)$, $\boldsymbol{\xi}_i^U = (\xi_{i11}{}^U, \cdots, \xi_{i1N_2}{}^U, \cdots, \xi_{iJ_i1}{}^U, \cdots, \xi_{iJ_iN_2}{}^U)$, and $\boldsymbol{\xi}_i^W = (\xi_{i111}{}^W, \cdots, \xi_{i11N_3}{}^W, \cdots, \xi_{iJ_in_{iJ_i}1}{}^W, \cdots, \xi_{iJ_in_{iJ_i}N_3}{}^W)$.

So the BLUP of $\mathbf{u}_i$ would be

$$
\begin{aligned}
\hat{\mathbf{u}}_i &= (\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}'\widetilde{\mathbf{Y}}_i \\
&= \begin{pmatrix} n_{i\cdot}\mathbf{I}_{N_1} & \mathbf{1}'_{J_i} \otimes (n_{iJ_i}\mathbf{\Phi}'_X\mathbf{\Phi}_U) & \mathbf{1}'_{n_{i\cdot}} \otimes (\mathbf{\Phi}'_X\mathbf{\Phi}_W) \\ & \mathbf{I}_{J_i} \otimes (n_{ij}\mathbf{I}_{N_2}) & \mathbf{I}_{J_i} \otimes (\mathbf{1}'_{n_{ij}} \otimes \mathbf{\Phi}'_U\mathbf{\Phi}_W) \\ & & \mathbf{I}_{n_{i\cdot}N_3} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{\Phi}_X^{'}\mathbf{Y}_i\mathbf{1}_{n_{i\cdot}} \\ \text{vec}\{\mathbf{\Phi}_U^{'}\mathbf{Y}_i(\mathbf{1}_{n_{ij}} \otimes \mathbf{I}_{J_i})\} \\ \text{vec}\{\mathbf{\Phi}_W^{'}\mathbf{Y}_i\} \end{pmatrix} \\
&= \begin{pmatrix} n_{i\cdot}\mathbf{I}_{N_1} & \mathbf{1}'_{J_i} \otimes n_{ij}\mathbf{C}_{XU} & \mathbf{1}'_{n_{i\cdot}} \otimes \mathbf{C}_{XW} \\ & \mathbf{I}_{J_i} \otimes (n_{ij}\mathbf{I}_{N_2}) & \mathbf{I}_{J_i} \otimes (\mathbf{1}'_{n_{ij}} \otimes \mathbf{C}_{UW}) \\ & & \mathbf{I}_{n_{i\cdot}N_3} \end{pmatrix}^{-1} \begin{pmatrix} (\mathbf{A}_X^{N_1})'\mathbf{S}^{1/2}\mathbf{U}_i'\mathbf{1}_{n_{i\cdot}} \\ \text{vec}\{(\mathbf{A}_U^{N_2})'\mathbf{S}^{1/2}\mathbf{U}_i'(\mathbf{1}_{n_{ij}} \otimes \mathbf{I}_{J_i})\} \\ \text{vec}\{(\mathbf{A}_W^{N_3})'\mathbf{S}^{1/2}\mathbf{U}_i'\} \end{pmatrix}
\end{aligned}
$$

The second row of the equation corresponds to the estimation obtained based on the fast algorithm of the rank preserved transformation. $\mathbf{A}_X^{N_1}$, $\mathbf{A}_U^{N_2}$ and $\mathbf{A}_W^{N_3}$ are the estimated eigenfunctions in the lower-dimensional space. $\mathbf{Y}_i = \mathbf{V}\mathbf{S}^{1/2}\mathbf{U}_i$ is derived from the SVD of the whole data matrix. $\mathbf{C}_{XU}$ is defined as $(\mathbf{A}_X^{N_1})'\mathbf{A}_U^{N_2}$, $\mathbf{C}_{XW} = (\mathbf{A}_X^{N_1})'\mathbf{A}_W^{N_3}$ and $\mathbf{C}_{UW} = (\mathbf{A}_U^{N_2})'\mathbf{A}_W^{N_3}$.

For two-way crossed model (C2) with balanced design $Y_{ij}(t) = X_i(t) + Z_j(t) + W_{ij}(t), i = 1, 2, \cdots, I; j = 1, 2, \cdots, J$, we again have $\widetilde{\mathbf{Y}} = \mathbf{B}\mathbf{u}$, where $\mathbf{Y} = (\mathbf{Y}_{11}, \cdots, \mathbf{Y}_{ij}, \cdots, \mathbf{Y}_{IJ})$ and $\widetilde{\mathbf{Y}} = \mathrm{vec}(\mathbf{Y})$. $\mathbf{B} = [\mathbf{B}_X|\mathbf{B}_Z|\mathbf{B}_W]$, where $\mathbf{B}_X = \mathbf{I}_I \otimes (\mathbf{1}_J \otimes \mathbf{\Phi}_X)$, $\mathbf{B}_Z = \mathbf{1}_I \otimes (\mathbf{I}_J \otimes \mathbf{\Phi}_Z)$ and $\mathbf{B}_W = \mathbf{I}_{IJ} \otimes \mathbf{\Phi}_W$. $\mathbf{u} = (\mathbf{u}^X, \mathbf{u}^Z, \mathbf{u}^W)'$, and $\mathbf{u}^X = (\boldsymbol{\xi}_1{}^{X\prime}, \cdots, \boldsymbol{\xi}_I{}^{X\prime})$ where $\boldsymbol{\xi}_i{}^X = (\xi_{i1}{}^X, \xi_{i2}{}^X, \cdots, \xi_{iN_1}{}^X)'$; $\mathbf{u}^Z = (\boldsymbol{\xi}_1{}^{Z\prime}, \cdots, \boldsymbol{\xi}_J{}^{Z\prime})$ where $\boldsymbol{\xi}_j{}^Z = (\xi_{j1}{}^Z, \xi_{j2}{}^Z, \cdots, \xi_{jN_2}{}^Z)'$; and $\mathbf{u}^W = (\boldsymbol{\xi}_{11}{}^{W\prime}, \cdots, \boldsymbol{\xi}_{IJ}{}^{W\prime})$ where $\boldsymbol{\xi}_{ij}{}^W = (\xi_{ij1}{}^W, \xi_{ij2}{}^W, \cdots, \xi_{ijN_3}{}^W)'$.

The BLUP gives

$$
\begin{aligned}
\hat{\mathbf{u}} &= (\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}'\widetilde{\mathbf{Y}} \\
&= \begin{pmatrix} J\mathbf{I}_{IN_1} & (\mathbf{1}_I\mathbf{1}_J') \otimes (\mathbf{\Phi}_X'\mathbf{\Phi}_Z) & (\mathbf{I}_I \otimes \mathbf{1}_J') \otimes (\mathbf{\Phi}_X'\mathbf{\Phi}_W) \\ & I\mathbf{I}_{JN_2} & (\mathbf{1}_I' \otimes \mathbf{I}_J) \otimes (\mathbf{\Phi}_Z'\mathbf{\Phi}_W) \\ & & \mathbf{I}_{IJN_3} \end{pmatrix}^{-1} \begin{pmatrix} \mathrm{vec}\{\mathbf{\Phi}_{\mathrm{X}}'\mathbf{Y}(\mathbf{I}_{\mathrm{I}} \otimes \mathbf{1}_{\mathrm{J}})\} \\ \mathrm{vec}\{\mathbf{\Phi}_{\mathrm{Z}}'\mathbf{Y}(\mathbf{1}_{\mathrm{I}} \otimes \mathbf{I}_{\mathrm{J}})\} \\ \mathrm{vec}(\mathbf{\Phi}_{\mathrm{W}}'\mathbf{Y}) \end{pmatrix} \\
&= \begin{pmatrix} J\mathbf{I}_{IN_1} & (\mathbf{1}_I\mathbf{1}_J') \otimes \mathbf{C}_{XZ} & (\mathbf{I}_I \otimes \mathbf{1}_J') \otimes \mathbf{C}_{XW} \\ & I\mathbf{I}_{JN_2} & (\mathbf{1}_I' \otimes \mathbf{I}_J) \otimes \mathbf{C}_{ZW} \\ & & \mathbf{I}_{IJN_3} \end{pmatrix}^{-1} \begin{pmatrix} \mathrm{vec}\{(\mathbf{A}_{\mathrm{X}}^{\mathrm{N}_1})'\mathbf{S}^{1/2}\mathbf{U}'(\mathbf{I}_{\mathrm{I}} \otimes \mathbf{1}_{\mathrm{J}})\} \\ \mathrm{vec}\{(\mathbf{A}_{\mathrm{Z}}^{\mathrm{N}_2})'\mathbf{S}^{1/2}\mathbf{U}'(\mathbf{1}_{\mathrm{I}} \otimes \mathbf{I}_{\mathrm{J}})\} \\ \mathrm{vec}\{(\mathbf{A}_{\mathrm{W}}^{\mathrm{N}_3})'\mathbf{S}^{1/2}\mathbf{U}'\} \end{pmatrix}
\end{aligned}
$$

Again, $\mathbf{C}_{XZ} = \mathbf{\Phi}_X'\mathbf{\Phi}_Z = (\mathbf{A}_X^{N_1})'\mathbf{A}_Z^{N_2}$, and similar definitions apply for $\mathbf{C}_{XW}$ and $\mathbf{C}_{ZW}$.

APPENDIX B: Method of Moment Estimators for Additional Models

In this section, we list the method of moments estimators for covariance matrices of latent processes in the additional models that are listed in Table 1 of the paper 'Structured Functional Principal Component Analysis'.

B.1 Multi-way nested model (NM)

The covariance operators of latent processes in multi-way nested model satisfy

$$
\begin{aligned}
&E\big[\{Y_{i_1 i_2 \cdots i_r}(t) - Y_{h_1 h_2 \cdots h_r}(t)\}\{Y_{i_1 i_2 \cdots i_r}(s) - Y_{h_1 h_2 \cdots h_r}(s)\}^T\big] \\
&= \begin{cases}
2K_r(t,s), & \text{if } i_1 = h_1, \cdots, i_{r-1} = h_{r-1}, i_r \neq h_r \\
2\{K_{r-1}(t,s) + K_r(t,s)\}, & \text{if } i_1 = h_1, \cdots, i_{r-2} = h_{r-2}, i_{r-1} \neq h_{r-1} \\
\cdots & \\
2\{K_1(t,s) + \cdots + K_r(t,s)\}, & \text{if } i_1 \neq h_1
\end{cases}
\end{aligned}
$$

Therefore, we have $H_j (j = 1, 2, \cdots, r)$ operators as

$$
\begin{aligned}
H_1 &= \frac{1}{k_1 - n} \sum_{i_1, i_2, \cdots, i_{r-1}} \sum_{i_r, h_r} \{Y_{i_1 \cdots i_{r-1} i_r} - Y_{i_1 \cdots i_{r-1} h_r}\}\{Y_{i_1 \cdots i_{r-1} i_r} - Y_{i_1 \cdots i_{r-1} h_r}\}^T \\
&= \frac{2}{k_1 - n} \mathbf{Y}(\mathbf{D}_1 - \mathbf{E}_1^T \mathbf{E}_1)\mathbf{Y}^T
\end{aligned}
$$

$$
\begin{aligned}
H_2 &= \frac{1}{k_2 - k_1} \sum_{\substack{i_1, \cdots, i_{r-2} \\ i_{r-1} \neq h_{r-1} \\ i_r, h_r}} (Y_{i_1 \cdots i_r} - Y_{i_1 \cdots i_{r-2} h_{r-1} h_r})(Y_{i_1 \cdots i_r} - Y_{i_1 \cdots i_{r-2} h_{r-1} h_r})^T \\
&= \frac{2}{k_2 - k_1} \mathbf{Y}(\mathbf{D}_2 - \mathbf{E}_2^T \mathbf{E}_2 - \mathbf{D}_1 + \mathbf{E}_1^T \mathbf{E}_1)\mathbf{Y}^T
\end{aligned}
$$

$$
\begin{aligned}
H_3 &= \frac{1}{k_3 - k_2} \{ \sum_{\substack{i_{r-1}, i_r, h_{r-1}, h_r \\ i_{r-2} \neq h_{r-2} \\ i_1, \cdots, i_{r-3}}} (Y_{i_1 \cdots i_r} - Y_{i_1 \cdots i_{r-3} h_{r-2} h_{r-1} h_r})(Y_{i_1 \cdots i_r} - Y_{i_1 \cdots i_{r-3} h_{r-2} h_{r-1} h_r})^T \\
&\quad - \sum_{\substack{i_1, \cdots, i_{r-2} \\ i_{r-1}, i_r, h_{r-1}, h_r}} (Y_{i_1 \cdots i_{r-2} i_{r-1} i_r} - Y_{i_1 \cdots i_{r-2} h_{r-1} h_r})(Y_{i_1 \cdots i_{r-2} i_{r-1} i_r} - Y_{i_1 \cdots i_{r-2} h_{r-1} h_r})^T \} \\
&= \frac{2}{k_3 - k_2} \mathbf{Y}(\mathbf{D}_3 - \mathbf{E}_3^T \mathbf{E}_3 - \mathbf{D}_2 + \mathbf{E}_2^T \mathbf{E}_2)\mathbf{Y}^T \\
&\cdots
\end{aligned}
$$

$$
\begin{aligned}
H_j &= \frac{2}{k_j - k_{j-1}} \mathbf{Y}(\mathbf{D}_j - \mathbf{E}_j^T \mathbf{E}_j - \mathbf{D}_{j-1} + \mathbf{E}_{j-1}^T \mathbf{E}_{j-1})\mathbf{Y}^T \\
&\cdots
\end{aligned}
$$

where $k_j = \sum_{i_1 i_2 \cdots i_{r-j}} n^2_{i_1 i_2 \cdots i_{r-j}}, j = 1, 2, \cdots, r$. The covariance operators are represented as

$$\begin{aligned}\widehat{K}_{r+1-j} &= (H_{j+1} - H_j)/2 \\ &= \mathbf{Y}\{\frac{1}{k_{j+1} - k_j}(\mathbf{D}_{j+1} - \mathbf{E}^T_{j+1}\mathbf{E}_{j+1} - \mathbf{D}_j + \mathbf{E}^T_j\mathbf{E}_j) \\ &\quad - \frac{1}{k_j - k_{j-1}}(\mathbf{D}_j - \mathbf{E}^T_j\mathbf{E}_j - \mathbf{D}_{j-1} + \mathbf{E}^T_{j-1}\mathbf{E}_{j-1})\}\mathbf{Y}^T\end{aligned}$$

B.2 Two-way crossed design with sub-sampling (C2s)

With model $Y_{ijk}(t) = X_i(t) + Z_j(t) + W_{ij}(t) + U_{ijk}(t)$, $i = 1, 2, \cdots, I$; $j = 1, 2, \cdots, J$ and $k = 1, 2, \cdots, n_{ij}$, we have

$$\begin{aligned}&E\{Y_{ijk}(t) - Y_{luv}(t)\}\{(Y_{ijk}(s) - Y_{luv}(s)\}^T \\ = &\begin{cases} 2K_U(t,s), & \text{if } i = l, j = u, k \neq v \\ 2\{K_Z(t,s) + K_W(t,s) + K_U(t,s)\}, & \text{if } i = l, j \neq u \\ 2\{K_X(t,s) + K_W(t,s) + K_U(t,s)\}, & \text{if } i \neq l, j = u \\ 2\{K_X(t,s) + K_Z(t,s) + K_W(t,s) + K_U(t,s)\}, & \text{if } i \neq k, j \neq u \end{cases}\end{aligned}$$

The corresponding H operators are

$$H_U = \frac{1}{k_{12} - n}\sum_{i,j}\sum_{k \neq v}(Y_{ijk} - Y_{ijv})(Y_{ijk} - Y_{ijv})^T = \frac{2}{k_{12} - n}\mathbf{Y}(\mathbf{D}_{12} - \mathbf{E}^T_{12}\mathbf{E}_{12})\mathbf{Y}^T$$

$$\begin{aligned}H_Z &= \frac{1}{k_1 - k_{12}}\sum_{i=1}\sum_{j \neq u}\sum_{k,v}(Y_{ijk} - Y_{iuv})(Y_{ijk} - Y_{iuv})^T \\ &= \frac{2}{k_1 - k_{12}}\mathbf{Y}(\mathbf{D}_1 - \mathbf{E}^T_1\mathbf{E}_1 - \mathbf{D}_{12} + \mathbf{E}^T_{12}\mathbf{E}_{12})\mathbf{Y}^T\end{aligned}$$

$$\begin{aligned}H_X &= \frac{1}{k_2 - k_{12}}\sum_{i \neq l}\sum_{j}\sum_{k,v}(Y_{ijk} - Y_{ljv})(Y_{ijk} - Y_{ljv})^T \\ &= \frac{2}{k_2 - k_{12}}\mathbf{Y}(\mathbf{D}_2 - \mathbf{E}^T_2\mathbf{E}_2 - \mathbf{D}_{12} + \mathbf{E}^T_{12}\mathbf{E}_{12})\mathbf{Y}^T\end{aligned}$$

$$\begin{aligned}H_T &= \frac{1}{n^2 - k_1 - k_2 + k_{12}}\sum_{i \neq l}\sum_{j \neq u}\sum_{k,v}(Y_{ijk} - Y_{luv})(Y_{ijk} - Y_{luv})^T \\ &= \frac{2}{n^2 - k_1 - k_2 + k_{12}}\mathbf{Y}(n\mathbf{I} - \mathbf{1}\mathbf{1}^T - \mathbf{D}_1 - \mathbf{D}_2 + \mathbf{D}_{12} + \mathbf{E}^T_1\mathbf{E}_1 + \mathbf{E}_2\mathbf{E}_2 - \mathbf{E}_{12}\mathbf{E}^T_{12})\mathbf{Y}^T\end{aligned}$$

where $k_1 = \sum_i n^2_{i0}$, $k_2 = \sum_j n^2_{0j}$ and $k_{12} = \sum_{i,j} n^2_{ij}$.

Multi-way crossed design (CM)

The most general model for crossed design is

$$Y_{i_1 i_2 \cdots i_r u}(t) = R^{(1)}_{i_1}(t) + R^{(2)}_{i_2}(t) + \cdots + R^{(r)}_{i_r}(t) + \cdots + R^{(j_1 j_2 \cdots j_q)}_{i_{j_1} i_{j_2} \cdots i_{j_q}}(t) + \cdots + R^{(12\cdots r)}_{i_1 i_2 \cdots i_r}(t) + R_{i_1 i_2 \cdots i_r u}(t)$$

with $i_k = 1, 2, \cdots, m_k$ where $k = 1, 2, \cdots, r$, $u = 1, 2, \cdots, n_{i_1 i_2 \cdots i_r}$; $1 \leqslant q \leqslant r$, $(j_1, j_2, \cdots, j_q) \in \{1, 2, \cdots, r\}$, $j_1 < j_2 < \cdots < j_q$ and $R^{(j_1 j_2 \cdots j_q)}_{i_{j_1} i_{j_2} \cdots i_{j_q}}(t)$ has variance operator $K_{j_1 j_2 \cdots j_q}$, $R_{i_1 i_2 \cdots i_r u}(t)$ has covariance operator K_U. With similar procedure as in the previous sections, we can work out the formula for the covariance operators. We omit the details here.

Appendix C: Additional Simulation Results

In this section, we present additional simulation results for both crossed deisgn (C2) and nested design (N3). We also compare the effect of sample sizes for each latent process and signal-to-noise ratio in estimating the eigenfunctions. We use different smoothing methods for high- and low-dimensional data.

C.1 Two-way crossed designs (C2)

For the two-way crossed design (C2), we generate data from the following model

$$\begin{cases} Y_{ij}(t) = \sum_{k=1}^{N_X} \phi^X_k(t)\xi_{ik} + \sum_{l=1}^{N_Z} \phi^Z_l(t)\zeta_{jl} + \sum_{h=1}^{N_W} \phi^W_h(t)\eta_{ijh}, t \in \mathcal{T} \\ \xi_{ik} \overset{i.i.d}{\sim} N(0, \lambda^X_k), \zeta_{jl} \overset{i.i.d}{\sim} N(0, \lambda^Z_l) \text{ and } \eta_{ijh} \overset{i.i.d}{\sim} N(0, \lambda^W_h) \end{cases} \tag{1}$$

where ξ_{ik}'s, ζ_{jl}'s and η_{ijh}'s are mutually uncorrelated. We choose $N_X = N_Z = N_W = 4$ and assign the true eigenvalues to be $\lambda^X_k = \lambda^Z_k = \lambda^W_k = 0.5^{k-1}, k = 1, 2, 3, 4$. True eigenfunctions are

$\phi^X_1(t) = \sin(2\pi t)$	$\phi^Z_1(t) = \sin(6\pi t)$	$\phi^W_1(t) = \sqrt{3}(2t-1)$
$\phi^X_2(t) = \cos(2\pi t)$	$\phi^Z_2(t) = \cos(6\pi t)$	$\phi^W_2(t) = \sqrt{5}(6t^2 - 6t + 1)$
$\phi^X_3(t) = \sin(4\pi t)$	$\phi^Z_3(t) = \sin(8\pi t)$	$\phi^W_3(t) = \sqrt{5}(20t^3 - 30t^2 + 12t - 1)$
$\phi^X_4(t) = \cos(4\pi t)$	$\phi^Z_4(t) = 1/\sqrt{2}$	$\phi^W_4(t) = 1$

These functions are measured on the grid $\mathcal{T} = 1/p, 2/p, \cdots, 1$. For the low-dimensional case, we choose $p = 100$. Let $I = 200$ and $J = 20$ be the number of categories for the first-level processes $X(t)$ and $Z(t)$, respectively.

We repeat simulation studies 100 times and display the estimated eigenfunctions for the three latent processes in Figure 1. The algorithm estimates $X(t)$ and $W(t)$ very well. Given the small sample size that is observed for $Z(t)$, estimation for K_Z and thus its eigenfunctions are more noisy and less stable compared to those for $X(t)$ and $W(t)$. Estimation precision gets improved when J increases. In fact, we examine another two scenarios where $I = 20$, $J = 200$ and $I = J = 200$ (Figure 2 and 3) to evaluate the effect of sample size on recovering each variance component. The accuracy of estimating the eigenfunctions increases when there are more observations per latent process.

[Figure 1 about here.]

[Figure 2 about here.]

[Figure 3 about here.]

C.2 Three-way nested designs (N3) with white noise for high-dimensional data

Here we show the simulation results for the high-dimensional (N3) model as presented in Section 4 of the paper when $p = 50,000$ and under different varying signal-to-noise ratios. As described in more details in the paper, the true model is

$$\begin{cases} Y_{ijk}(t) = \sum_{l=1}^{N_1} \phi_l^X(t)\xi_{il} + \sum_{m=1}^{N_2} \phi_m^U(t)\zeta_{ijm} + \sum_{h=1}^{N_3} \phi_h^W(t)\eta_{ijkh} + \epsilon_{ijkt}, \\ \xi_{il} \overset{i.i.d}{\sim} N(0, \lambda_l^X), \zeta_{ijm} \overset{i.i.d}{\sim} N(0, \lambda_m^U), \eta_{ijkh} \overset{i.i.d}{\sim} N(0, \lambda_h^W) \text{ and } \epsilon_{ijkt} \overset{i.i.d}{\sim} N(0, \sigma^2) \end{cases}$$

where $i = 1, \cdots, I; j = 1, \cdots, J; k = 1, \cdots, K$; $N_1 = N_2 = N_3 = 4$, $\lambda_k^X = \lambda_k^U = \lambda_k^W = 0.5^{k-1}, k = 1, 2, 3, 4$; $t \in \mathcal{T} = 1/p, 2/p, \cdots, 1$; $I = 50$, $J = 5$ and $K = 5$.

Figure 4, 5, 6 and 7 display the estimated eigenfunctions for the three hierarchical processes

when $\sigma = 0, 0.1, 0.5$ and 1. When p is small, we can directly smooth the resulting covariance matrices instead of smoothing the data. In particular, we have $\widetilde{K}_W(t,s) = K_W(t,s) + \sigma^2\delta_{ts}$, $\widetilde{K}_U(t,s) = K_U(t,s)$, $\widetilde{K}_X(t,s) = K_X(t,s)$. Thus, we can smooth the off-diagonal elements of $\widetilde{K}_W(t,s)$, $\widetilde{K}_U(t,s)$ and $\widetilde{K}_X(t,s)$ and obtain an estimate of σ^2. In high-dimensional case, the simulated curves are smoothed first before applying SFPCA. The estimations (gray curves) in general can well capture the shape of the true eigenfunctions (black curve). However, when the observations are contaminated with large noise ($\sigma = 0.5$ or 1), the estimated eigenfunctions are quite noisy, especially for processes $X(t)$ and $U(t)$ with small sample sizes. In addition, within each process, the 3^{rd} and 4^{th} eigenfunctions are worse estimated than the first two principal components.

[Figure 4 about here.]

[Figure 5 about here.]

[Figure 6 about here.]

[Figure 7 about here.]

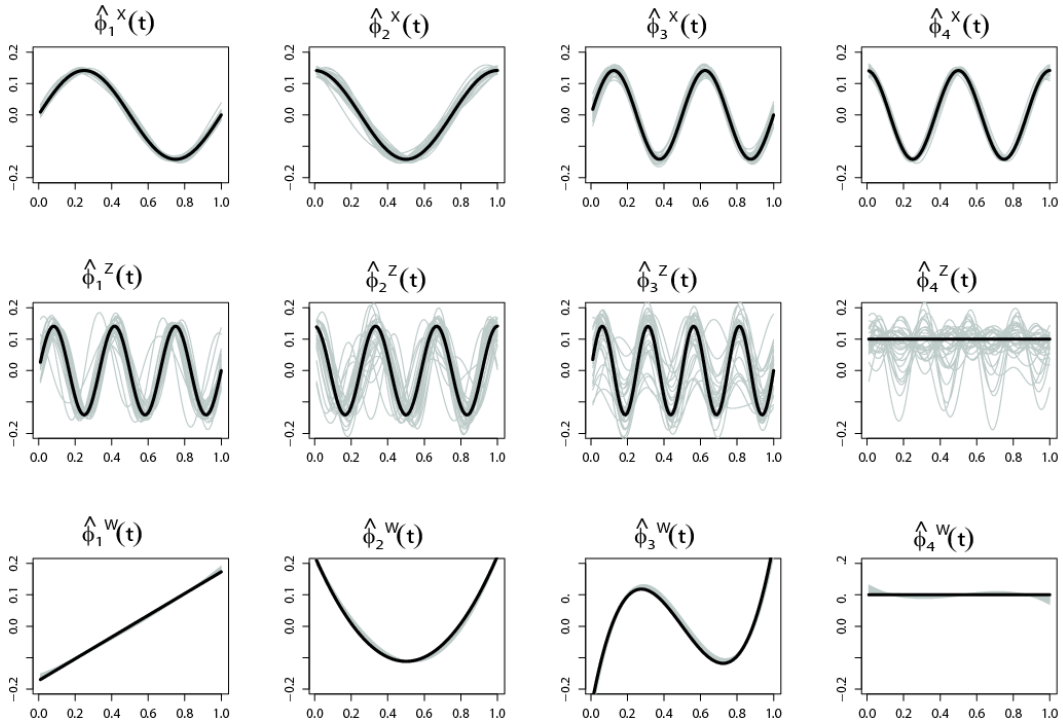

Figure 1. Plots display the estimated eigenfunctions for three latent processes $X(t)$, $Z(t)$ and $W(t)$, of two-way crossed design (C2) when $I = 200$ and $J = 20$. Results are shown as 50 random estimates of the eigenfunctions out of 100 simulations (shown in gray). The true eigenfunctions are displayed in black curves. The eigenfunctions for the two first-level processes $X(t)$ and $Z(t)$ are captured by two sets of trigonometric basis. While the eigenfunctions for the second-level process $W(t)$ are polynomial. Since the process-specific sample sizes for $X(t)$ and $W(t)$ are larger than that of $Z(t)$, eigenfunctions on the first and third rows are better estimated than those on the second row.

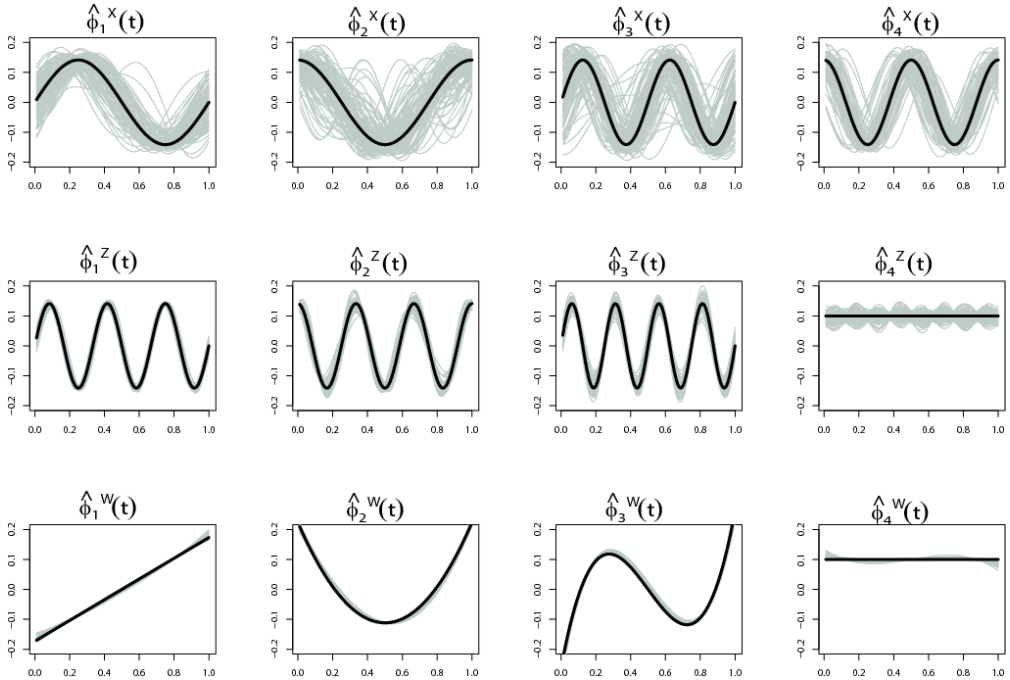

Figure 2. The estimated eigenfunctions when there are 20 samples in $X(t)$ and 200 in $Z(t)$. Since the sample size for $W(t)$ remains unchanged, the third row still demonstrate decent accuracy. However, the eigenfuctions for $X(t)$ are worsely estimated than those for $Z(t)$ in this scenario.

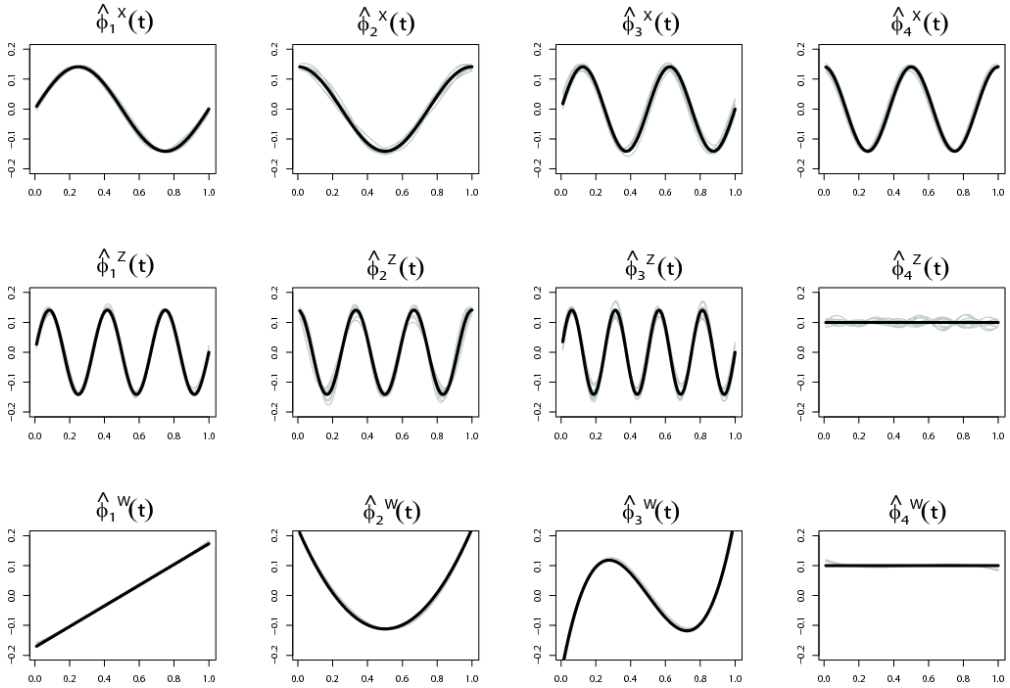

Figure 3. The estimated eigenfuctions for (C2) model when $I = J = 200$. The estimations for all the three processes are significantly improved due to the increased sample sizes.

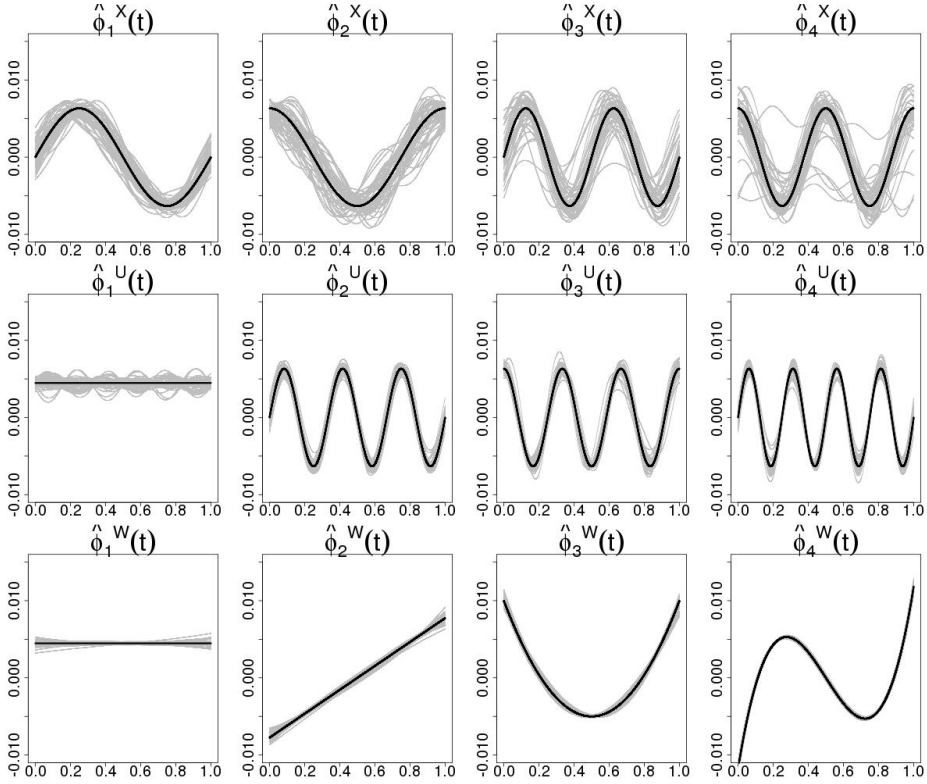

Figure 4. Plots demonstrate the estimated eigenfunctions (gray curves) for the three-way nested design (N3) when $p = 50,000$, $I = 50$, $J = K = 5$ and $\sigma = 0$. The true eigenfunctions are displayed in black curves. The eigenfunctions for the first- and second-level processes $X(t)$ and $U(t)$ are captured by two sets of trigonometric basis. The eigenfunctions for the third-level process $W(t)$ are polynomial. The simulated curves are smoothed first before applying fast SFPCA algorithm.

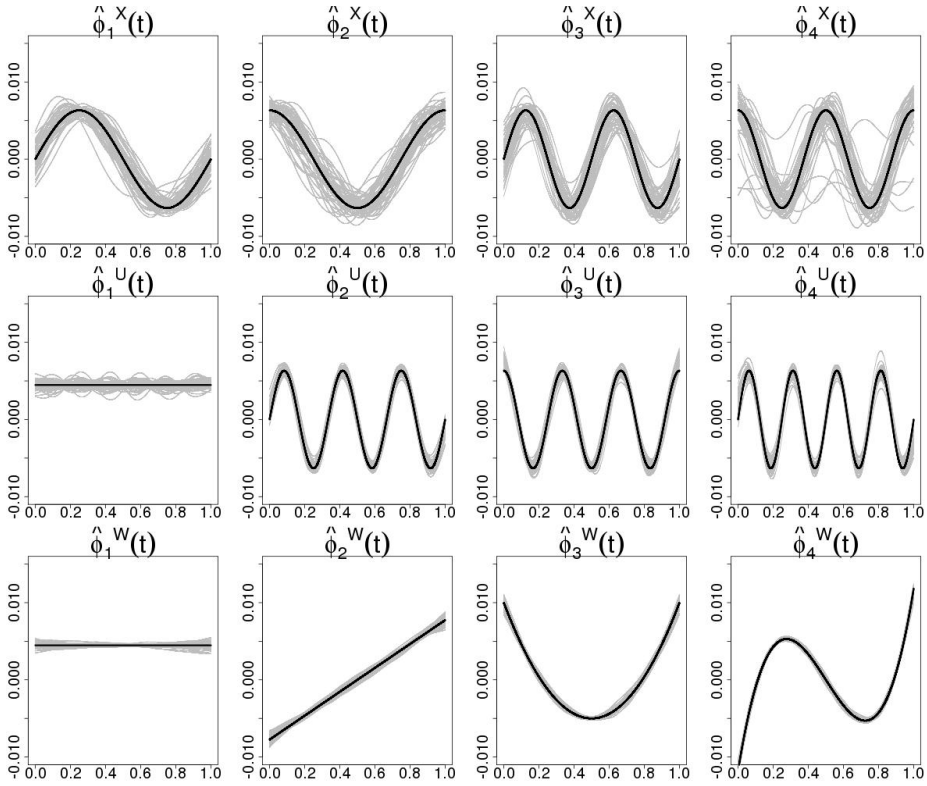

Figure 5. The estimated eigenfunctions for high-dimensional (N3) model when $\sigma = 0.1$.

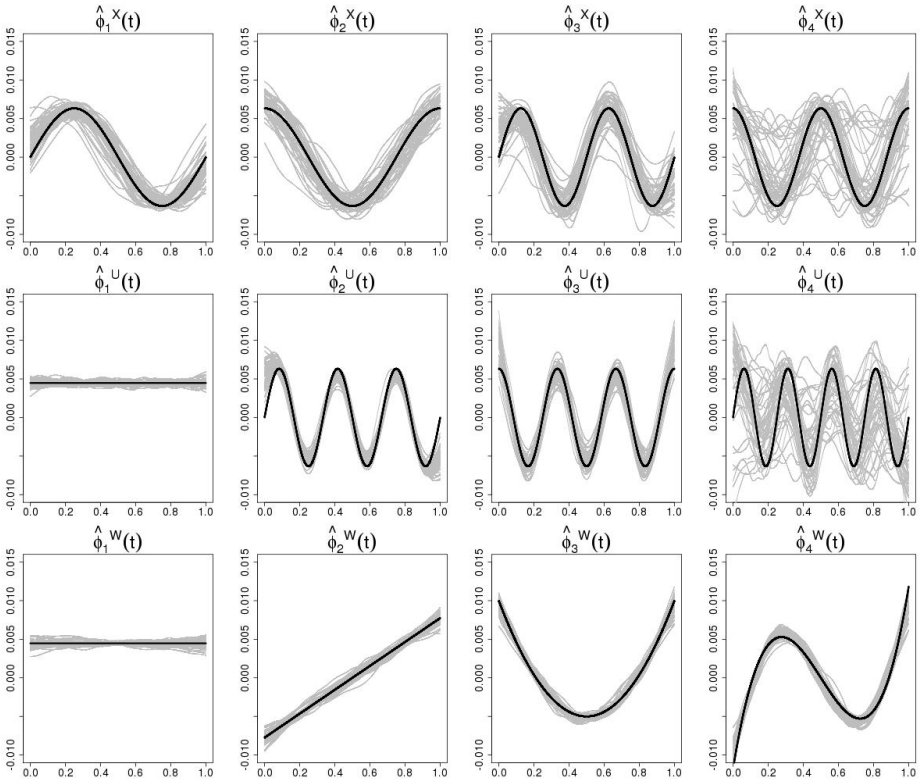

Figure 6. The estimated eigenfunctions for high-dimensional (N3) model when $\sigma = 0.5$. $\phi_4^X(t)$ and $\phi_4^U(t)$ are especially noisy due to the small percentages of variance that these two components explain in the data.

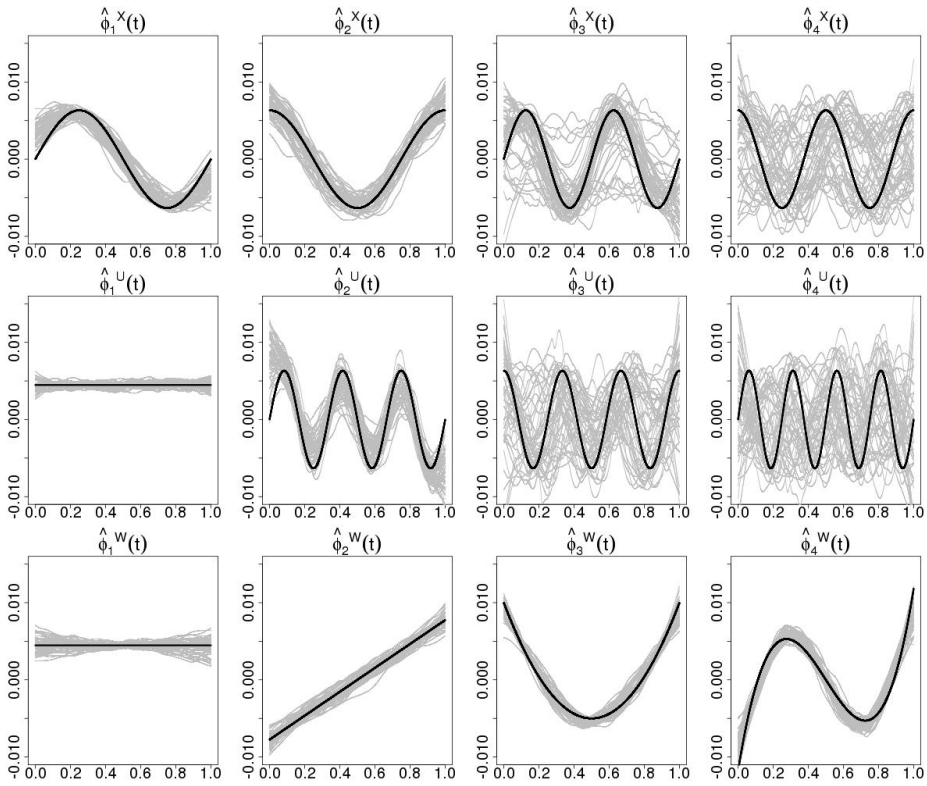

Figure 7. The estimated eigenfunctions for high-dimensional (N3) model when $\sigma = 1$. $\phi_4^X(t)$ and $\phi_4^U(t)$ also appear much less stable compared to the estimates when $\sigma = 0.5$.