Biostatistics (2021), **0**, 0, *pp.* 1–30
doi:10.1093/biostatistics/output

Interpretable Principal Component Analysis for Multilevel Multivariate Functional Data Supplementary Materials

Jun Zhang

Department of Biostatistics, University of Pittsburgh, Pittsburgh, PA, USA

Greg J. Siegle

Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA, USA

Tao Sun

School of Statistics, Renmin University, Beijing, China

Wendy D'Andrea

Department of Psychology, New School for Social Research, New York, NY, USA

Robert T. Krafty*

Department of Biostatistics and Bioinformatics, Emory University, Atlanta, GA, USA

rkrafty@emory.edu

Summary

This document contains supplemental materials to the article "Interpretable Principal Components Analysis for Multilevel Multivariate Functional Data." Section 1 presents additional methodological details, including the estimation of within-subject correlation ρ_{ij} (Section 1.1), the selection of tuning parameters (Section 1.2), the estimation of eigenvalues and principal

*To whom correspondence should be addressed.

component scores (Section 1.3) and the form of the Fantope projection operator (Section 1.4). Section 2 presents additional details with regards to the application, including a description of data processing (Section 2.1), a presentation of all estimated regression coefficients (Section 2.2), additional visual representations of estimated quantities (Section 2.3), and a comparison of performance of the proposed procedure to the naive procedure often used in the applied literature (Section 2.4). Section 3 presents results from additional simulation studies with varying number of subjects (Section 3.1) and varying values for the tuning parameter δ (Section 3.2).

Key words:

1. Additional Methodological Details

1.1 *Estimation of* ρ_{jk}

The MoM estimator discussed in Section 4.1 in the main text depends on an estimator of the within-subject-between-electrode correlation ρ_{jk}. Here, we consider an estimator of ρ_{jk} that takes advantage both of the separability of the within-subject electrode and temporal effects, and of the sparsity of ρ_{jk}. This estimator can be viewed as a modified version of the estimator considered by Staicu *and others* (2010) to our setting. Define the $M \times M$ matrix $f_{jk}(t,s) = 2\left\{K_w(t,s)\left(1-\rho_{jk}\right) + \sigma^2 I(t=s) I_M\right\}$, which can be consistently estimated as

$$\hat{f}_{jk}(t,s) = \frac{1}{N}\sum_{i=1}^{N}\left\{\boldsymbol{Y}_{ij}(t) - \boldsymbol{Y}_{ik}(t)\right\}\left\{\boldsymbol{Y}_{ij}(s) - \boldsymbol{Y}_{ik}(s)\right\}^T.$$

Note that, due to separability of the within-subject electrode and temporal effects, we can define a function $F_{jk} = \int_{t,s\in\mathcal{T}} \mathbf{1}_M^T f_{jk}(t,s)\mathbf{1}_M \, dt\, ds$ which, since $F_{jk} \propto (1-\rho_{jk})$, provides a measure of the disassociation of electrodes in that F_{jk} is large for electrodes for which $\rho_{jk} = 0$. A set of electrodes for which $\rho_{jk} = 0$ can be estimated by thresholding the estimator $\widehat{F}_{jk} = \sum_{p=1}^{P}\sum_{q\neq p} \mathbf{1}_M^T \hat{f}_{jk}\left(t_p, t_q\right)\mathbf{1}_M / P(P-1)$.

For $\delta \in (0,1)$, define the set of pairs of electrodes $\Delta = \left\{(j,k) \mid \widehat{F}_{jk} > \text{ upper } \delta \text{ quantile of } \widehat{F}_{jk}\right\}$.

Our goal is to identify a subset of electrodes that are uncorrelated adjusting for subject-level deviations, and not necessarily the entire set. Consequently, δ should be selected in a conservative manner relative to the anticipated percentage of uncorrelated electrodes. The parameter δ was set at 30% for methods $1-4$ in our simulation and was set at 20% in the real data analysis. In practice, we suggest plotting all the value of $\hat{F}_{jk}$ and select somewhere below the changing point of $\hat{F}_{jk}$ to decide δ. Results for varying levels of δ are provided in Section 3.2.

Given this set Δ, we then define $\tilde{f}_{\Delta}(t,s) = \sum_{(j,k)\in\Delta} \hat{f}_{jk}(t,s)/\left|\Delta\right|$, where $|\Delta|$ is the number of pairs of electrodes in Δ, which is a consistent estimator of $2K_w(t,s)$ for $t \neq s$. Since $\tilde{f}_{\Delta}(t,s) - \hat{f}_{jk}(t,s)$ is a consistent estimator of $2\rho_{jk}K_W(t,s)$, we can construct a consistent estimator of ρ_{jk} as

$$\hat{\rho}_{jk} = \left\{ \sum_{p=1}^{P} \sum_{q\neq p} 1_M^T \left[\tilde{f}_{\Delta}(t_p,t_q) - \hat{f}_{jk}(t_p,t_q) \right] 1_M \right\} \Big/ \left\{ \sum_{p=1}^{P} \sum_{q\neq p} 1_M^T \left[\tilde{f}_{\Delta}(t_p,t_q) \right] 1_M \right\}.$$

1.2 *Tuning Parameter Selection*

Our optimization procedure (3) in the main text involves three tuning parameters: γ controls smoothness, α controls the among variate sparsity, and λ controls the within time localization. We will first select a common γ for all eigenfunctions $\phi_r, r = 1, ..., R$ using cross-validation (Rice and Silverman, 1991), then fix γ and select α_r and λ_r sequentially using either cross-validation or fraction of variance explained, depending the goal of the analysis.

To select γ, we employ five-fold cross-validation. The parameter γ is chosen among a set of candidates γ_s such that the estimated covariance $K^{(\nu)}$ from the validation dataset, and the estimated $H_{r=1}^{(-\nu)}(\gamma,0,0)$ from the training dataset with α and λ being 0, have the largest cross-validated inner product. Formally:

$$\hat{\gamma} = \underset{\gamma\in\Upsilon_1}{\operatorname{argmax}} \sum_{\nu=1}^{5} \langle H_{r=1}^{(-\nu)}(\gamma,0,0), K^{(\nu)} \rangle$$

where Υ_1 is a candidate set of γ, for which we used a sequence between 0 and P times the largest eigenvalue of K.

Similarly, (α_r, λ_r) as a combination can be chosen by maximizing the cross-validated inner product of $H_{r=1}^{(-\nu)}(\hat{\gamma}, \alpha_r, \lambda_r)$ and $K^{(\nu)}$:

$$(\hat{\alpha}_r, \hat{\lambda}_r) = \underset{(\alpha_r,\lambda_r)\in\Upsilon_{2,r}}{\operatorname{argmax}} \sum_{\nu=1}^{5} \langle H_r^{(-\nu)}(\hat{\gamma}, \alpha_r, \lambda_r), K^{(\nu)} \rangle$$

where $\Upsilon_{2,r}$ is a candidate set of (α_r, λ_r). Both the candidate sequences of α_r and λ_r are between 0 and the 95% quantile of absolute values of off-diagonal entries in $K_r = (I - \widehat{\Pi}_{r-1})K(I - \widehat{\Pi}_{r-1})$. In our simulation we used coordinate descent to find the maximum cross-validated inner product.

The $(\hat{\alpha}_r, \hat{\lambda}_r)$ found by the cross-validation approach minimizes the bias of estimating the eigenfunction ϕ_r. When ϕ_r is truly localized either within variates or among variates, cross-validation would be a desirable approach to reveal the true level of sparsity in ϕ_r. We adopted this method in our simulation analysis in Section 5 in the main text.

Often, rather than an accurate estimate that is closest to the true ϕ_r, we are more interested in an interpretable estimate $\widehat{\phi}_r$ that highlights variates and time points with dominant variation, even with some sacrifice of the fraction of variance explained (FVE). The second method of choosing (α_r, λ_r) is designed to provide such interpretable $\widehat{\phi}_r$. Define $\text{FVE}(\phi) = \frac{\phi^T K \phi}{\text{totv}(K-\gamma D)}$, where $\text{totv}(K-\gamma D)$ is the sum of all positive eigenvalues of $K-\gamma D$, which is an approximation of the total variation removing the noise σ^2. Also define relative FVE as $\text{rFVE}(\alpha_r, \lambda_r) = \frac{\text{FVE}(\widehat{\phi}_r(\hat{\gamma},\lambda_r,\alpha_r))}{\text{FVE}(\widehat{\phi}_r(\hat{\gamma},0,0))}$ where $\widehat{\phi}_r(\hat{\gamma}, \lambda_r, \alpha_r)$ is the r^{th} estimated eigenfunction with $\hat{\gamma}, \alpha_r$, and λ_r. Then we select the largest localization under the condition that rFVE is larger than some proportion $b \in (0, 1]$ that one choose to guarantee:

$$(\hat{\alpha}_r, \hat{\lambda}_r) = \underset{(\alpha_r,\lambda_r)\in\Upsilon_{2,r}}{\operatorname{argmax}} \{\alpha_r + \lambda_r : \text{rFVE}_r(\alpha_r, \lambda_r) \geqslant b\}. \tag{1.1}$$

If there are more than one combination providing largest $\alpha_r + \lambda_r$, we will choose the combination with largest α_r to provide a more parsimonious eigenfunction. We will illustrate this method in our real data analysis in Section 6 in the main text.

1.3 *Estimation of Eigenvalues and Principal Component Scores*

For each subject i, we re-write Equation 1 in the manuscript using the mixed model format

$$Y_i = Z_i u_i + e_i$$

where $Y_i = [Y_{i1}^T, Y_{i2}^T, ..., Y_{iJ}^T]^T, i = 1, ..., N$ is a $JMP \times 1$ vector. The design matrix Z_i is an $JMP \times (R_1 + JR_2)$ matrix containing the subject-level and electrode-level eigenvectors

$$Z_i = \begin{bmatrix} \begin{bmatrix} \phi_1^z & ... & \phi_{R_1}^z \\ \phi_1^z & ... & \phi_{R_1}^z \\ \vdots & ... & \vdots \\ \phi_1^z & ... & \phi_{R_1}^z \end{bmatrix} \begin{bmatrix} \phi_1^w & & & \\ & \phi_1^w & & \\ & & \ddots & \\ & & & \phi_1^w \end{bmatrix} ... \begin{bmatrix} \phi_{R_2}^w & & & \\ & \phi_{R_2}^w & & \\ & & \ddots & \\ & & & \phi_{R_2}^w \end{bmatrix} \end{bmatrix}.$$

The vector $\mathbf{u_i}$ contains R_1 subject level and $J \times R_2$ electrode level principal component scores to be estimated, and takes the form $u_i = [\xi_{i1}^z, ..., \xi_{iR_1}^z, [\xi_{i1,j=1}^w, ..., \xi_{i1,j=J}^w], ..., [\xi_{iR_2,j=1}^w, ..., \xi_{iR_2,j=J}^w]]^T$. The covariance matrix of u_i is

$$G_i = \begin{bmatrix} \theta_1^z & & & & & \\ & \ddots & & & & \\ & & \theta_{R_1}^z & & & \\ & & & \begin{bmatrix} \theta_1^w & & \\ & \ddots & \\ & & \theta_{R_2}^w \end{bmatrix} \otimes \begin{bmatrix} 1 & \rho_{12} & ... & \rho_{1J} \\ \vdots & \vdots & ... & \vdots \\ \rho_{J1} & \rho_{J2} & ... & 1 \end{bmatrix} \end{bmatrix}.$$

The error term e_i is a vector of length JMP, and the covariance matrix of e_i is $R_i = \sigma^2 I$.

Following Robinson (1991), the best linear unbiased predictor (BLUP) of the principal component scores has the form $\hat{u}_i = G_i Z_i^T (Z_i G_i Z_i^T + R_i)^{-1} Y_i$, Z_i, G_i and R_i are assumed to be known in the BLUP, which involves the estimated $\hat{\rho}_{jk}$, $\hat{\sigma}^2$, $\hat{\phi}_r^z, \hat{\phi}_r^w$, and $\hat{\theta}_r^z, \hat{\theta}_r^w$. As previously described, $\hat{\rho}_{jk}$ is obtained from the MoM estimator, $\hat{\phi}_r^z$ and $\hat{\phi}_r^w$ can be obtained by interpolating eigenvectors estimated from LVPCA. To estimate θ_r^z, θ_r^w and σ^2, we follow two steps in below:

1. Estimate σ^2: Before introducing any localization, we first introduced a roughness term to smooth the eigenvectors by maximizing the problem $\phi^T(K - \gamma D)\phi$ such that $\|\phi\|_{12} = 1$ for

some given $\gamma > 0$. As this is just the form of singular value decomposition, the middle part of the quadratic form can be viewed as a smoothed covariance matrices $\widetilde{K}_z = \widehat{K}_z - \gamma_z D$ and $\widetilde{K}_w = \widehat{K}_w - \gamma_w D$. Recall the electrode level $E\{\widehat{K}_w(t,s)\} = K_w(t,s) + \frac{1}{c}\sigma^2 I(t=s)$. We estimate σ^2 by

$$\hat{\sigma}^2 = \frac{c}{MP}\left\{\text{trace}\left[\widehat{K}_w\right] - \sum \text{all the positive eigenvalues of } \widetilde{K}_w\right\}$$

2. Estimate θ_r: Recall $\theta_r = \int_{t\in\mathcal{T}}\int_{s\in\mathcal{T}} \phi_r^T(t)K(t,s)\phi_r(s)dtds$, we estimate the subject-level θ_r^z by $\hat{\theta}_r^z = \hat{\phi}_r^{zT}\widehat{K}_z\hat{\phi}_r^z/MP$, and estimate the electrode-level θ_r^w by $\hat{\theta}_r^w = \hat{\phi}_r^{wT}\widehat{K}_w\hat{\phi}_r^w/MP$.

1.4 *Frobenius projection operator* $\mathcal{P}_{\mathcal{D}_\Pi(B)}$

Lemma (Chen and Lei, 2015): Let $\Pi = VV^T$, where the $p \times d$ matrix V contains d orthogonal basis. Let the $p \times (p-d)$ matrix U be the orthogonal complement basis of V.

Suppose $U^TBU = \sum_{i=1}^{p-d}\gamma_i\eta_i\eta_i^T$ is a spectral decomposition of U^TBU, then $U^T\mathcal{P}_{\mathcal{D}_\Pi(B)}U = \sum_{i=1}^{p-d}\gamma_i^+(\theta)\eta_i\eta_i^T$, where $\gamma_i^+(\theta) = \min\left[\max(\gamma_i-\theta,0),1\right]$ and $\sum_{i=1}^{p-d}\gamma_i^+(\theta) = 1$. So the Frobenius projection operator of symmetric matrix B is

$$\mathcal{P}_{\mathcal{D}_\Pi(B)} = U\left[\sum_{i=1}^{p-d}\gamma_i^+(\theta)\eta_i\eta_i^T\right]U^T.$$

2. Additional Details for the Analysis of Data from the BADA Study

2.1 *Data Processing*

The primary task, preprocessing, and conventional analysis of data from the task on a different sample are fully described in Kerr *and others* (2019). Briefly, temporal artifacts likely to be caused by blinks were rescaled, electrodes with outlying time domain characteristics (mean or standard deviation outside the Turkey hinges) were interpolated, within participants, from other electrodes, movement derived from raw and smoothed (4s kernel), and accelerometer data were regressed out of each channel independently. Following transformation to the frequency domain, frequency

data were linearly rescaled by dividing by the base frequency to make power at each frequency roughly consistent, and were log transformed. Temporal outliers in the frequency domain time series were Windsorized within electrodes. Frequency domain data were temporally smoothed (10 second kernel) and normalized to have initial values of zero at each electrode by subtracting the time series at each electrode from a baseline calculated from the first 100 millisecond of the trial. Participants did the 10 second rumination task twice, separated by a 10 second rest period. The processed 10 second frequency domain time series was averaged for the two runs.

2.2 *False Discovery Rate Control*

To adjust for multiple testing, all reported p-values are adjusted to control the false discovery rate (FDR) at 0.05. We utilize the adaptive group Benjamin and Hochberg (GBH) procedure (Hu *and others*, 2010) for the 14 hypothesis tests grouped within anatomical regions for each electrode-level principal components, controlling the FDR for each component at $.00625 = .05/8$. Specifically, based on the correlation structure the 14 electrodes are grouped into 5 functionally distinct regions across the scalp: the right frontal (AF4, F4, F8, FC6), the right temporoparietal (T8, P8), the left frontal (AF3, F3, F7, FC5), the left temporoparietal (T7, P7), and the occipital (O1, O2) region. The proportions of true null hypothesis, which are estimated through the two-stage (TST) method (Benjamini *and others*, 2006), are assumed to be dissimilar between the five groups and the signals are more likely to appear together in these groups.

2.3 *Complete Regression Results*

Table 1 presents estimates, standard errors and adjusted p-values from the univariate regression models for all 60 principal component scores.

2.4 *Additional Visual Representations*

2.4.1 *Scatter plots of principal component scores vs. transformed DES* Figure 1 shows scatter plots of principal component scores vs. transformed DES along with fitted values for the models presented in Table 4 of the manuscript.

2.4.2 *Topologies of η_j* Figure 2 shows the mean topologies across 14 of the electrode-specific shifts from the overall mean function (maps for $\widehat{\eta}_j$ from Equation 1 in the manuscript) across all frequency bands.

2.5 *Regression between Average Activity over Time and Transformed DES Scores*

One of the traditional approaches for EEG analysis is to consider temporally averaged activity during a task. In this subsection, we show that the associations identified with average activity are mostly consistent with that identified from the multilevel LVPCA. However, the LVPCA reveals more information, with regards to both dynamics of brain activity and connections across frequency bands.

Instead of using principal component scores ξ^z_{ir} and ξ^w_{ijr} as predictors, consider a simpler way of defining subject-level and electrode-level scores. Let $\zeta^z_{m,i} = \frac{1}{p}\sum_{t=1}^{p} \bar{Y}^{(m)}_{i\cdot}(t_p)$ be the average activity across all the time points and all the electrodes for subject i, $m = 1, ..., 4$ for theta, alpha, beta and gamma frequency band. Let $\zeta^w_{m,ij} = \frac{1}{p}\sum_{t=1}^{p}[Y^{(m)}_{ij}(t_p) - \bar{Y}^{(m)}_{i\cdot}(t_p)]$ be the average activity across all the time points for subject i at electrode j with subject mean removed. Here $\zeta^z_{m,i}$ and $\zeta^w_{m,ij}$ are new subject-level and electrode-level scores and will be used to fit individual regression models with the square-root transformed DES scores same as in Section 6.2 in the main text.

Table 2 displays estimates from five models with significant effects or main score effects. The significant interaction effects between a history of trauma and ζ^z_{theta}, ζ^w_{beta} at O_1 and ζ^w_{gamma} at

O_1 indicates that, among participants with a history of trauma, higher whole-brain theta power, and higher beta and gamma power within the occipital cortex are associated with lower level of dissociation. The significant main effect of ζ^w_{beta} at O_1 and O_2 indicates that, among participants without a history of trauma, higher beta power within the occipital cortext is associated with higher levels of dissociation. These findings are consistent with those from the multilevel LVPCA analysis. However, the LVPCA provides more detail. For example, from the LVPCA we know that higher whole-brain theta power with an emphasis on power in the middle of the trial is associated with dissociation, and that higher beta and gamma power with an early increasing trend, usually occurs together and jointly associate with dissociation. In addition, coefficients of the interactions in this analysis are shrunk towards zero relative to those from the LVPCA, which may due to loss of information incurred by averaging over time. It should be noted that the only significant association that is found with this simple approach but not with LVPCA is the interaction effect between ζ^w_{alpha} at O_1.

3. Additional Simulation Results

3.1 *Simulation Results with Varying Numbers of Subjects*

In this section, we present additional simulation results with varying numbers of subjects, $N = 50$ and 200, and with $P = 100$, $\sigma = 1$ and other parameters same as the setting shown in the main text, to evaluate the effect of sample size on empirical performance.

Figures 3 and 4 display estimated eigenfunctions ϕ^z_r and $\phi^w_r, r = 1, 2, 3$, from one simulated data set with $N = 50$ and $N = 200$ subjects, respectfully. They display similar characteristics as in Figure 2 in the manuscript with $N = 100$ subjects in that within-variate localization, between-variate sparseness penalties and accounting for within-subject correlation between electrode-specific deviations leads to favorable recovery of eigenfunctions. The medians of the errors $\|\phi - \hat{\phi}\|_2$ over 200 simulations with $N = 50$ and $N = 200$ subjects are reported in Tables 3 and 4,

respectfully. Together with Table 2 in the manuscript with $N = 100$, it can be seen that the errors in estimating eigenfunctions decrease with increasing sample size for each of the eight estimation methods. As seen for $N = 100$ in the manuscript, for both $N = 50$ and $N = 200$ the proposed LVPCA outperforms the other methods when estimating subject-level eigenfunctions, and performs favorably when estimating electrode-level eigenfunctions.

In terms of identifying areas of signal, Tables 5 and 6 list the median specificity and the median sensitivity for estimating nonzero eigenvector elements with $N = 50$ and $N = 200$ subjects, respectively. As seen in the manuscript with $N = 100$, here we see that the proposed LVPCA has the highest specificity and a reasonable level of sensitivity for $N = 50$ and $N = 200$. Median specificity and sensitivity increased slightly with increasing sample size for all methods, with two exceptions. The median sensitivity in estimating zero elements of $\hat{\phi}_3^z$ for the two methods $(\hat{\alpha}, \hat{\lambda}, 0)$ and $(0, \hat{\lambda}, 0)$, which have wider quantile ranges, decrease as sample size increases.

Figure 5 displays boxplots of estimated subject-level and electrode-level eigenvalues with $N = 50, 100$ and 200. The four methods that adjust for within-subject correlation between electrode-specific deviations can recover eigenvalues with relative little bias compared to the four methods without adjusting for within-subject correlation. Variance decreases for all components and for all methods with increasing sample size.

We also present empirical performance in estimating principal component scores, noise level and within-subject correlation between electrode specific deviations. Table 9 lists the average root mean square errors (RMSE) for estimating the principal component scores with $N = 50, 100$ and 200 subjects. The proposed LVPCA has minimum RMSEs among the methods, and all RMSEs reduce with increased sample size. With the setting $\sigma^2 = 1$, the average $\hat{\sigma}^2$ is 0.932 (SD = 0.007) when $N = 50$, 0.945 (SD=0.005) when $N = 100$, and 0.950 (SD=0.003) when $N = 200$. The empirical performance in estimating the within-subject correlation between electrode specific deviations is quantified through the average of errors $\|\hat{\Sigma} - \Sigma\|_2$ where Σ is the $J \times J$ correlation

matrix. The average of $\|\hat{\Sigma} - \Sigma\|_2$ is 0.496 (SD=0.181) when $N = 50$, 0.347 (SD=0.145) when $N = 100$, and 0.253 (SD=0.181) when $N = 200$, compared to $\|I - \Sigma\|_2 = 1.60$ without accounting for within-subject correlation between electrode specific deviations and restricting all $\hat{\rho}_{jk} = 0$.

3.2 *Simulation Results with Varying Values of δ*

In this section, we present results with varying values of δ. In Section 4.4 of the main text we discussed the estimation of ρ_{jk}, which requires to first identify a set Δ of pairs of electrodes for which $\rho_{jk} = 0$. We proposed to estimate the set Δ by thresholding the estimator $\hat{F}_{jk}$, i.e. $\Delta = \{(j,k)|\hat{F}_{jk} >$ upper δ quantile of $\hat{F}_{jk}\}$. In the previous simulations we assumed we know that 30% pairs of electrodes are uncorrelated adjusting for the subject-level deviation, i.e. $\delta = 30\%$. Here, we consider simulations with δ set to be 20% and 50% in order to explore robustness to selection of δ.

Table 8 shows the median of errors $\|\phi - \widehat{\phi}\|_2$ for $\phi_r^z, \phi_r^w, r = 1, 2, 3,$ (with median absolute deviations in parenthesis) with $\delta = 20\%, 30\%$ and 50% over 200 simulation runs, $N = 100, P = 100, \sigma = 1$. The middle part of the table with $\delta = 30\%$ is the same as the table 2 of the manuscript, but is displayed for comparison. Here we see that when $\delta = 20\%$, the median error for $\widehat{\phi}_r^z$ and $\widehat{\phi}_r^w$ are similar as when $\delta = 30\%$. When $\delta = 50\%$, some median error for $\widehat{\phi}_r^z$ becomes a little larger, especially for $\widehat{\phi}_3^z$, while the median error for $\widehat{\phi}_r^w$ still remains the same. This trend is consistent with what we observed in the table 2 of the manuscript, where methods 5 to 8 with $\hat{\rho} = 0$ (can be seen as $\delta = 100\%$) lead to higher median error for $\widehat{\phi}_r^w$ especially $\widehat{\phi}_3^w$ in a larger scale. Table 9 lists the median specificity and the median sensitivity for estimating nonzero eigenvector elements. the values are almost the same for $\delta = 20\%, 30\%$ and 50%.

Figure 6 displays boxplots of estimated subject-level and electrode-level eigenvalues. When δ is under-specified as 20%, θ_1^w, θ_2^w and θ_3^w are over-estimated with median bias ranging from 0.012 to 0.036. When δ is over-specified as 50%, θ_3^z is over-estimated, with median bias ranging from

0.018 to 0.022, and θ_1^w, θ_2^w and θ_3^w are under-estimated with median bias ranging from -0.108 to -0.025. This trend is consistent with what we have observed in Figure 5 where methods 5 to 8 with $\delta = 100\%$ lead to over-estimated θ_3^z and under-estimated θ_r^w in a larger scale.

In practice we suggest plotting all the value of $\hat{F}_{jk}$. Since $\hat{F}_{jk}$ should be large for electrodes pairs with $\rho_{jk} = 0$ compared to electrode pairs with correlation, we may select δ somewhere below the change in the velocity of $\hat{F}_{jk}$. For example, Figure 7 shows $\hat{F}_{jk}$ as a function of δ from one of the simulation datasets. The "elbow" of the line suggests picking $\delta = 30\%$.

3.3 *Simulation Results with larger number of variates*

In this section, we present results with larger number of variates $M = 15$. Specifically, we expanded the original eigenfunctions of $M = 3$ in Table 1 in the manuscript to $M = 15$ by replicating the three original eigenfunctions by 5 times. Other parameters same as the setting shown in the main text. We also considered varying number of subjects $N = 50, 100, 200$. The bias or error level for estimating the eigenvalues θ, PC scores ξ, and the correlation ρ remains the same compared to that of $M = 3$ in the manuscript, while the bias of eigenfunctions ϕ are slightly larger than that of $M = 3$.

Table 10 reports the medians of the errors $\|\phi - \hat{\phi}\|_2$ over 200 simulations with $N = 50, 100$ and 200 subjects. Overall the bias are slightly larger than the $M = 3$ case in the manuscript, while the overall trends are similar that the proposed LVPCA has the smallest bias. Table 11 reports the median specificity and sensitivity. Similar as in the $M = 3$ case, LVPCA has the highest specificity among all the methods.

Figure 8 displays boxplots of estimated subject-level and electrode-level eigenvalues with $N = 50, 100$ and 200. The trends are similar that the four methods that adjust for within-subject correlation between electrode-specific deviations can recover eigenvalues with relative little bias (ranged from -0.01 to 0.03 for $N = 100$, from -0.02 to 0.02 for $N = 50$ and from -0.01 to 0.02 for

$N = 200$) compared to the four methods without adjusting for within-subject correlation. The overall level of bias of $M = 15$ is comparable to that of $M = 3$.

Table 12 lists the average root mean square errors (RMSE) for estimating the principal component scores with $M = 15$. The proposed LVPCA has minimum RMSEs among the methods, and the overall error level is comparable to that of $M = 3$. For the estimated noise, the average $\hat{\sigma}^2$ is 0.946 (SD=0.004) when $N = 50$, 0.96 (SD = 0.003) when $N = 100$, and 0.966 (SD=0.002) when $N = 200$. The empirical performance in estimating the within-subject correlation between electrode specific deviations is quantified through the average of errors $\|\hat{\Sigma} - \Sigma\|_2$ where Σ is the $J \times J$ correlation matrix. The average of $\|\hat{\Sigma} - \Sigma\|_2$ is 0.532 (SD=0.182) when $N = 50$, 0.383 (SD=0.119) when $N = 100$, and 0.267 (SD=0.1) when $N = 200$, compared to $\|I - \Sigma\|_2 = 1.60$ without accounting for within-subject correlation between electrode specific deviations and restricting all $\hat{\rho}_{jk} = 0$. The overall level of error is comparable to that of $M = 3$, which is 0.496 (SD=0.181) when $N = 50$, 0.347 (SD=0.145) when $N = 100$, and 0.253 (SD=0.181) when $N = 200$.

References

Benjamini, Y., Krieger, M. A. and Yekutieli, D. (2006). Adaptive linear step-up procedures that control the false discovery rate. *Biometrika* **93**(3), 491–507.

Hu, J. X., Zhao, H. and Zhou, H. H. (2010). False discovery rate control with groups. *Journal of the American Statistical Association* **105**(491), 1215–1227.

Kerr, Margee, Siegle, Greg J. and Orsini, Jahala. (2019). Voluntary arousing negative experiences (vane): Why we like to be scared. *Emotion* **19**, 682–698.

Rice, J. A. and Silverman, B. W. (1991). Estimating the mean and covariance structurenon-

parametrically when the data are curves. *Journal of the Royal Statistical Society, Series B* **53**, 233–243.

Staicu, A. M., Crainiceanu, C. M. and Carroll, R. J. (2010). Fast methods for spatially correlated multilevel functional data. *Biostatistics* **11**(2), 177–194.

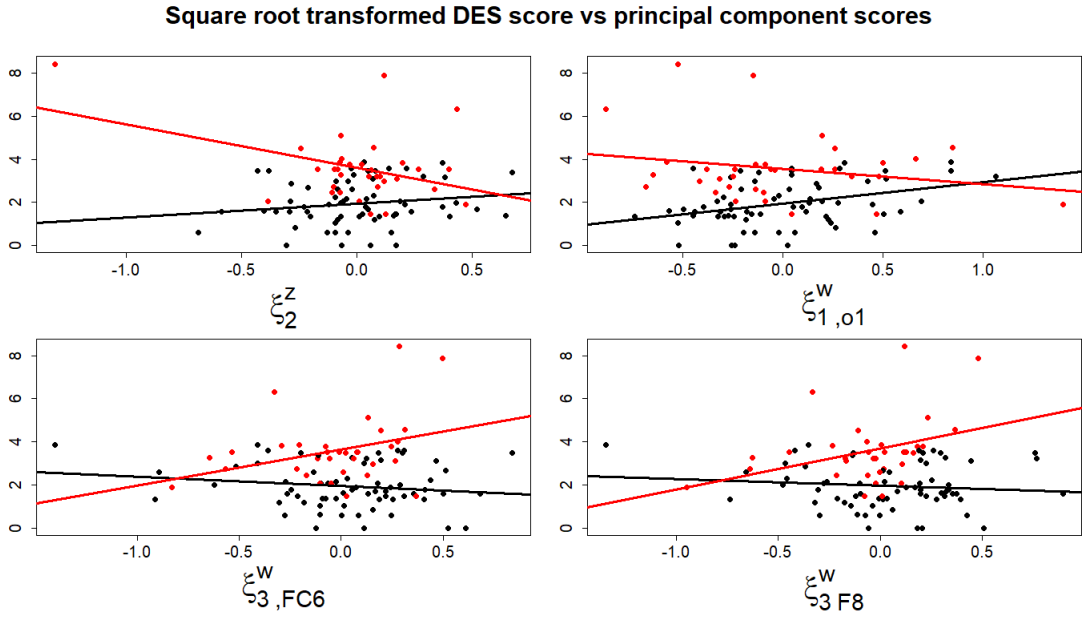
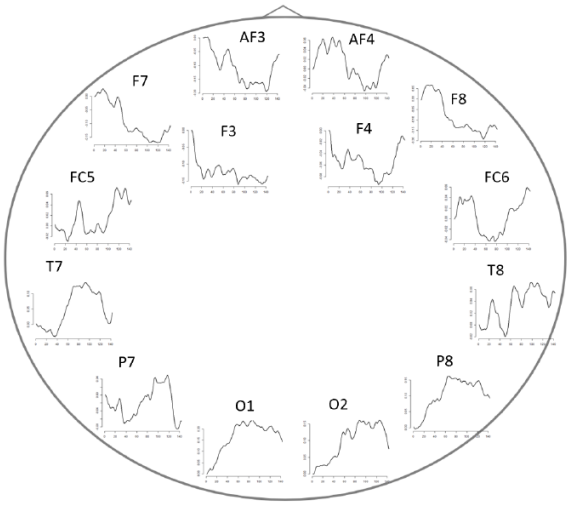
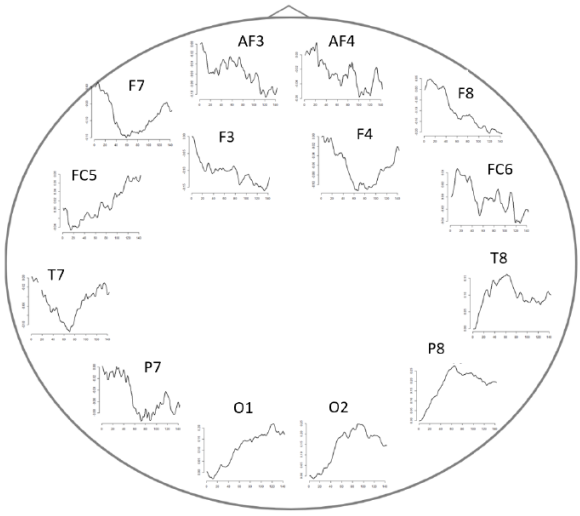


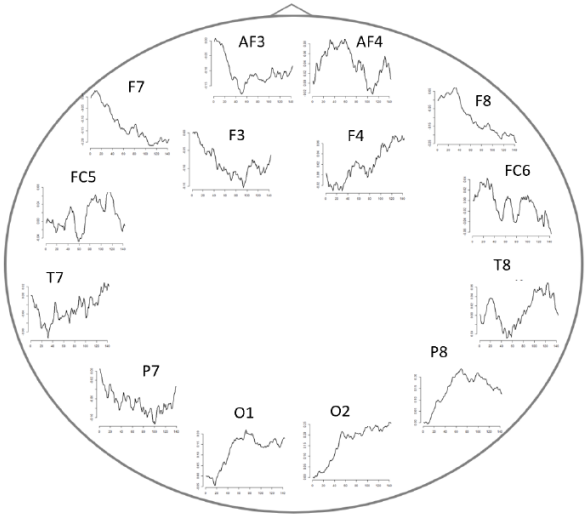
Fig. 1: Scatter plots and fitted regression lines of square root transformed DES score on ξ_2^z, $\xi_{1,\ \mathrm{O1}}^w$, $\xi_{3,\ \mathrm{FC6}}^w$ and $\xi_{1,\ \mathrm{F8}}^w$. Red represents participants with a history of trauma and black those without.



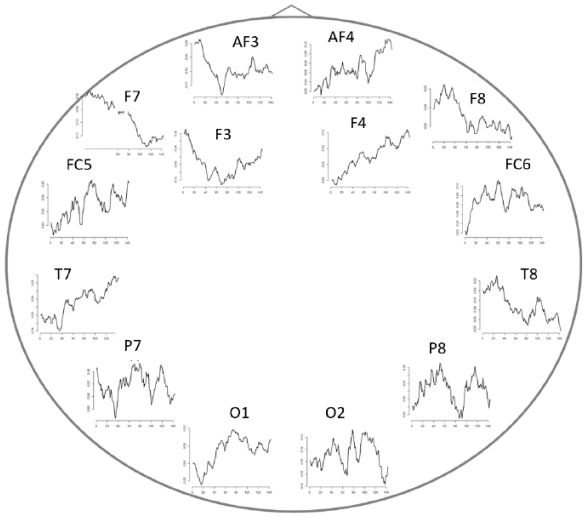
(a) $\widehat{\boldsymbol{\eta}}_j(t)$ for theta



(b) $\widehat{\boldsymbol{\eta}}_j(t)$ for alpha



(c) $\widehat{\boldsymbol{\eta}}_j(t)$ for beta



(d) $\widehat{\boldsymbol{\eta}}_j(t)$ for gamma

Fig. 2: $\widehat{\eta}_j(t)$ for four frequency bands.

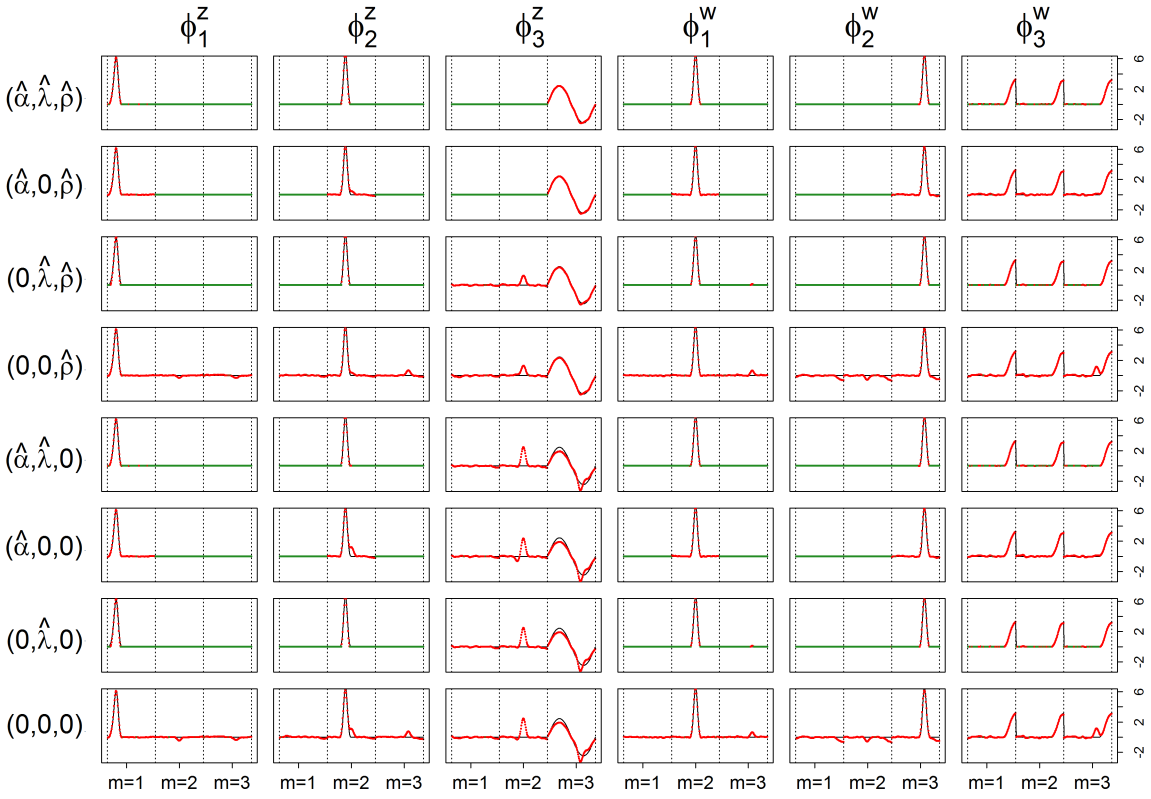


Fig. 3: True (black solid) and estimated (green and red dot) eigenfunctions from one simulated data set with $N = 50, P = 100$ and $\sigma = 1$, by the described eight estimation procedures. Solid black lines are the true eigenfunctions, green lines indicate estimated zero elements and red lines indicate estimated nonzero elements.

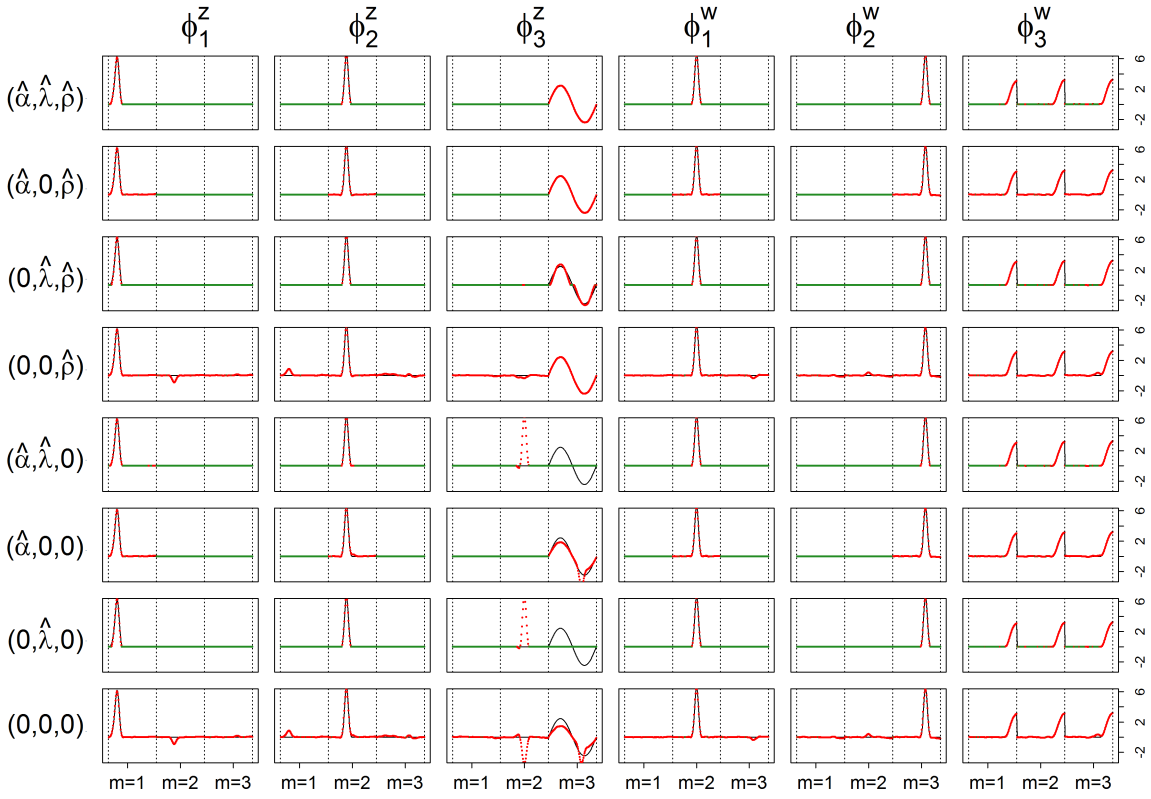


Fig. 4: True (black solid) and estimated (green and red dot) eigenfunctions from one simulated data set with $N = 200, P = 100$ and $\sigma = 1$, by the described eight estimation procedures. Solid black lines are the true eigenfunctions, green lines indicate estimated zero elements and red lines indicate estimated nonzero elements.

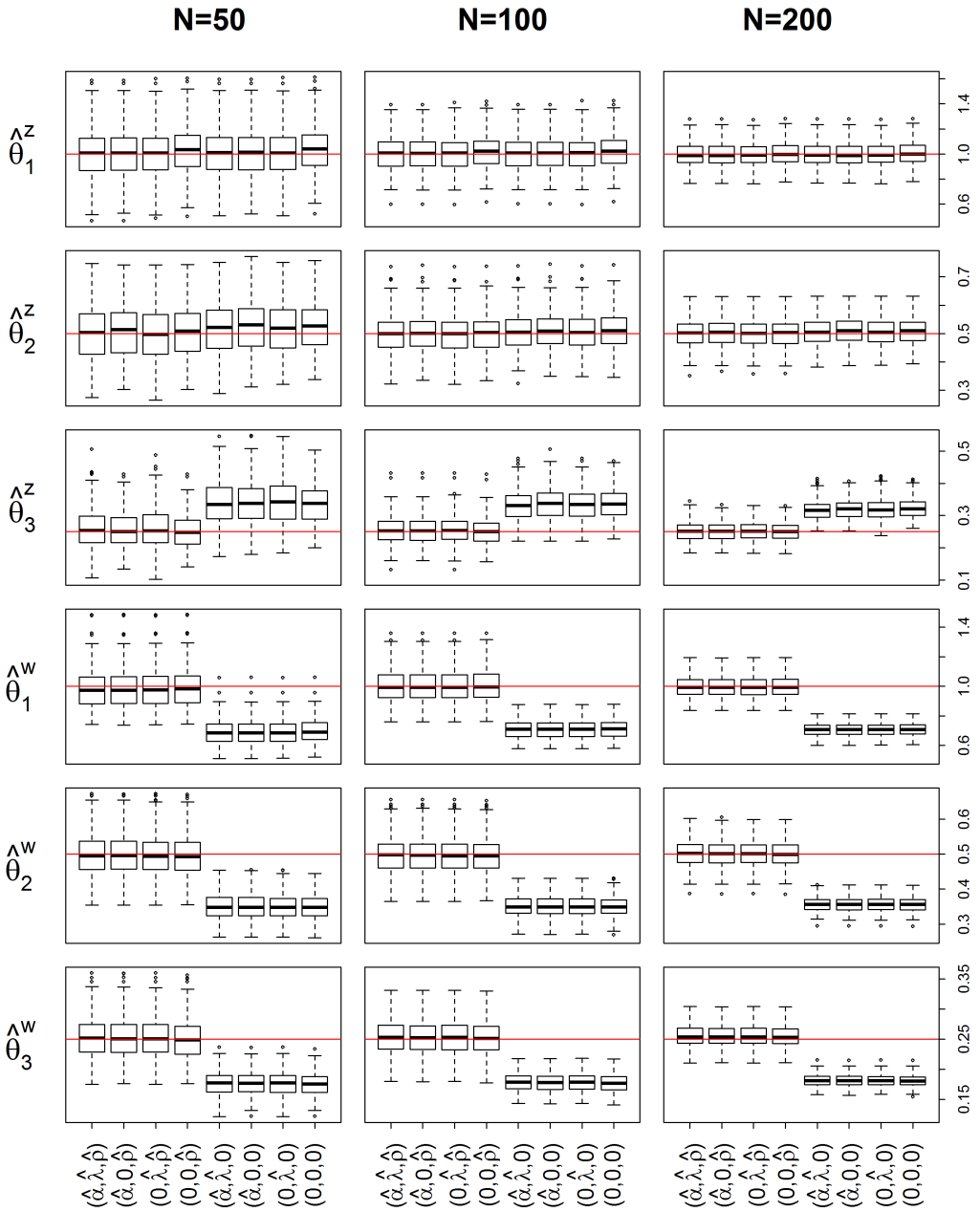


Fig. 5: Boxplots of eigenvalues estimated by the eight described estimation methods with $N = 50, 100,$ and $200, P = 100$ and $\sigma = 1$, over 200 simulation runs. Red solid lines indicate true eigenvalues.

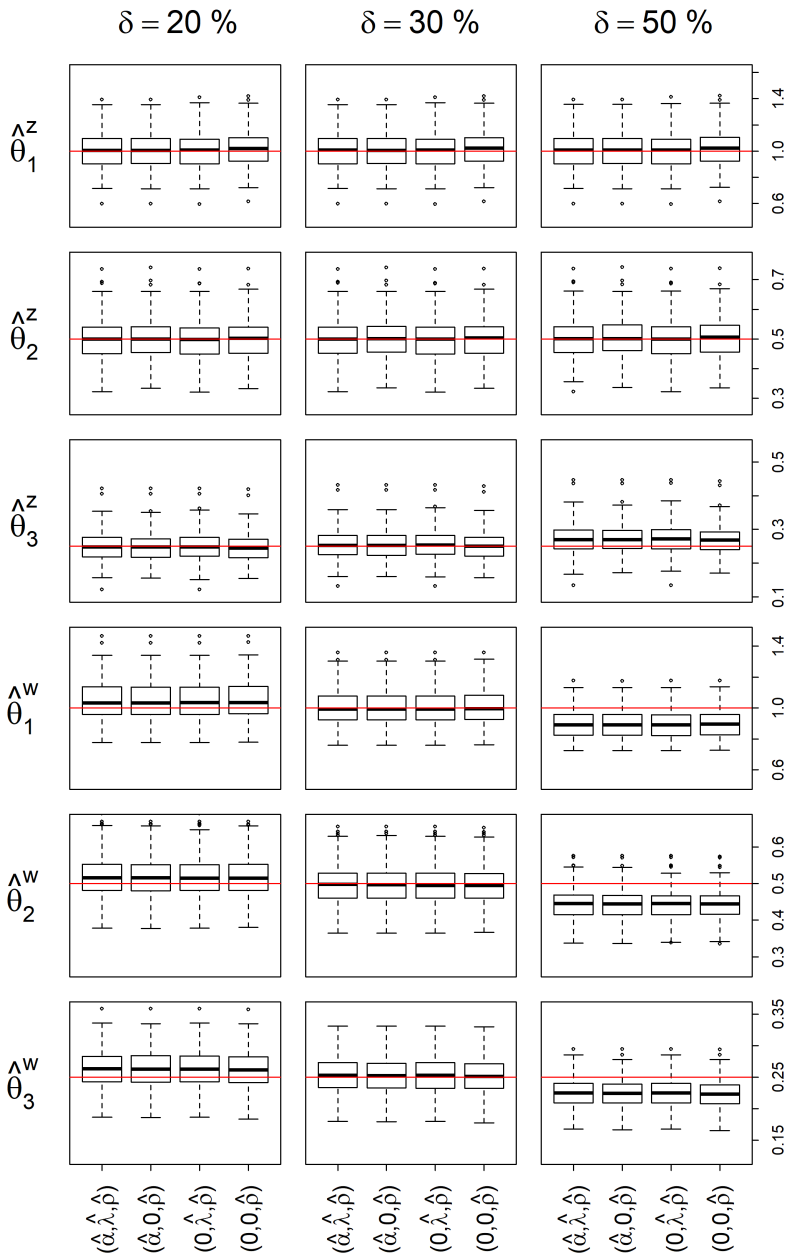


Fig. 6: Boxplots of eigenvalues estimated by the top four estimation methods, with $\delta = 20\%, 30\%$ and 50% over 200 simulation runs. Other parameters $N = 100, P = 100$ and $\sigma = 1$.

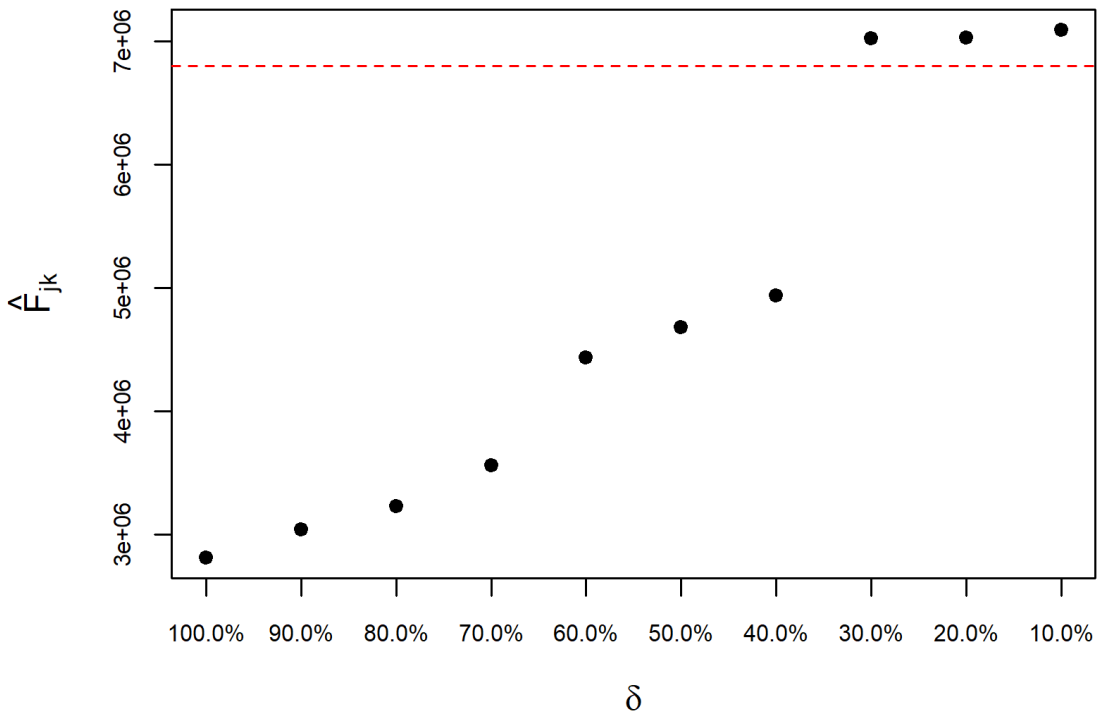


Fig. 7: Choosing δ by looking for the changing point of $\hat{F}_{jk}$ from one simulation dataset

Table 1: BADA Study - Univariate regression models for all 60 PC scores.

	Score		Trauma		Score × Trauma	
	β_1 (SE)	p-value	β_2 (SE)	p-value	β_3 (SE)	p-value
ξ^z_1	0.57 (0.35)	0.823	1.65 (0.26)	< 0.001	-1.11 (0.71)	0.967
ξ^z_2	0.65 (0.56)	0.999	1.66 (0.25)	< 0.001	-2.67 (0.89)	0.027
ξ^z_3	-0.44 (0.53)	0.999	1.62 (0.26)	< 0.001	1.78 (0.96)	0.526
ξ^z_4	0.52 (1.2)	0.999	1.64 (0.26)	< 0.001	-2.70 (1.71)	0.939
$\xi^w_{1,\text{AF3}}$	-0.47 (0.41)	0.999	1.62 (0.27)	< 0.001	0.48 (0.60)	0.999
$\xi^w_{1,\text{ F7}}$	-0.06 (0.39)	0.999	1.66 (0.27)	< 0.001	0.35 (0.57)	0.999
$\xi^w_{1,\text{ F3}}$	0.06 (0.38)	0.999	1.65 (0.27)	< 0.001	0.23 (0.68)	0.999
$\xi^w_{1,\text{FC5}}$	0.31 (0.42)	0.999	1.64 (0.26)	< 0.001	0.02 (0.62)	0.999
$\xi^w_{1,\text{ T7}}$	0.05 (0.35)	0.999	1.66 (0.26)	< 0.001	-0.55 (0.52)	0.999
$\xi^w_{1,\text{ P7}}$	-0.11 (0.32)	0.999	1.63 (0.27)	< 0.001	0.10 (0.68)	0.999
$\xi^w_{1,\text{ O1}}$	1.01 (0.40)	0.098	1.62 (0.25)	< 0.001	-1.73 (0.59)	0.036
$\xi^w_{1,\text{ O2}}$	0.68 (0.38)	0.313	1.67 (0.26)	< 0.001	-0.69 (0.59)	0.999
$\xi^w_{1,\text{ P8}}$	0.20 (0.34)	0.999	1.64 (0.26)	< 0.001	-0.25 (0.59)	0.999
$\xi^w_{1,\text{ T8}}$	0.30 (0.37)	0.999	1.64 (0.27)	< 0.001	-0.38 (0.61)	0.999
$\xi^w_{1,\text{FC6}}$	0.20 (0.32)	0.999	1.65 (0.27)	< 0.001	0.06 (0.55)	0.999
$\xi^w_{1,\text{ F4}}$	-0.43 (0.42)	0.999	1.62 (0.26)	< 0.001	0.46 (0.60)	0.999
$\xi^w_{1,\text{ F8}}$	0.02 (0.35)	0.999	1.65 (0.26)	< 0.001	0.72 (0.51)	0.999
$\xi^w_{1,\text{AF4}}$	-0.06 (0.34)	0.999	1.64 (0.27)	< 0.001	0.09 (0.45)	0.999
$\xi^w_{2,\text{AF3}}$	-0.72 (0.53)	0.999	1.64 (0.26)	< 0.001	0.82 (0.95)	0.999
$\xi^w_{2,\text{ F7}}$	-0.35 (0.47)	0.999	1.63 (0.26)	< 0.001	-0.64 (0.77)	0.999
$\xi^w_{2,\text{ F3}}$	-0.07 (0.50)	0.999	1.64 (0.27)	< 0.001	-0.13 (0.85)	0.999
$\xi^w_{2,\text{FC5}}$	-0.60 (0.50)	0.999	1.66 (0.27)	< 0.001	0.43 (0.80)	0.999
$\xi^w_{2,\text{ T7}}$	-0.22 (0.47)	0.999	1.67 (0.27)	< 0.001	-0.40 (0.78)	0.999
$\xi^w_{2,\text{ P7}}$	-0.15 (0.41)	0.999	1.67 (0.26)	< 0.001	-1.02 (0.82)	0.999
$\xi^w_{2,\text{ O1}}$	-0.25 (0.42)	0.999	1.64 (0.27)	< 0.001	0.48 (0.73)	0.999
$\xi^w_{2,\text{ O2}}$	-0.07 (0.48)	0.999	1.64 (0.27)	< 0.001	-0.08 (0.86)	0.999
$\xi^w_{2,\text{ P8}}$	0.11 (0.45)	0.999	1.63 (0.27)	< 0.001	-0.17 (0.85)	0.999
$\xi^w_{2,\text{ T8}}$	-0.08 (0.49)	0.999	1.63 (0.26)	< 0.001	0.76 (0.86)	0.999
$\xi^w_{2,\text{FC6}}$	0.24 (0.42)	0.999	1.68 (0.25)	< 0.001	1.76 (0.83)	0.999
$\xi^w_{2,\text{ F4}}$	-0.44 (0.49)	0.999	1.63 (0.26)	< 0.001	0.70 (0.80)	0.999
$\xi^w_{2,\text{ F8}}$	0.11 (0.44)	0.999	1.69 (0.26)	< 0.001	1.27 (0.84)	0.999
$\xi^w_{2,\text{AF4}}$	-0.23 (0.48)	0.999	1.64 (0.27)	< 0.001	0.17 (0.77)	0.999

Table 1 Continued

	Score		Trauma		Score × Trauma	
	β_1 (SE)	p-value	β_2 (SE)	p-value	β_3 (SE)	p-value
$\xi^w_{3,\text{AF3}}$	-0.48 (0.42)	0.999	1.65 (0.26)	< 0.001	1.40 (0.84)	0.999
$\xi^w_{3,\text{ F7}}$	-0.36 (0.41)	0.999	1.62 (0.26)	< 0.001	-0.35 (0.77)	0.999
$\xi^w_{3,\text{ F3}}$	-0.67 (0.48)	0.999	1.57 (0.26)	< 0.001	-0.70 (1.12)	0.999
$\xi^w_{3,\text{FC5}}$	-1.01 (0.45)	0.999	1.66 (0.26)	< 0.001	0.65 (0.76)	0.999
$\xi^w_{3,\text{ T7}}$	-0.73 (0.44)	0.999	1.65 (0.26)	< 0.001	0.25 (0.88)	0.999
$\xi^w_{3,\text{ P7}}$	-0.69 (0.38)	0.999	1.57 (0.26)	< 0.001	-0.48 (0.82)	0.999
$\xi^w_{3,\text{ O1}}$	-0.03 (0.37)	0.999	1.62 (0.26)	< 0.001	-0.66 (0.59)	0.999
$\xi^w_{3,\text{ O2}}$	-0.23 (0.37)	0.999	1.64 (0.26)	< 0.001	0.07 (0.71)	0.999
$\xi^w_{3,\text{ P8}}$	-0.24 (0.40)	0.999	1.64 (0.27)	< 0.001	0.75 (1.01)	0.999
$\xi^w_{3,\text{ T8}}$	-0.56 (0.40)	0.999	1.64 (0.27)	< 0.001	0.96 (0.89)	0.999
$\xi^w_{3,\text{FC6}}$	-0.43 (0.39)	0.999	1.68 (0.26)	< 0.001	2.10 (0.80)	0.082
$\xi^w_{3,\text{ F4}}$	-0.72 (0.41)	0.999	1.66 (0.27)	< 0.001	1.60 (0.86)	0.359
$\xi^w_{3,\text{ F8}}$	-0.31 (0.40)	0.999	1.74 (0.26)	< 0.001	2.21 (0.84)	0.082
$\xi^w_{3,\text{AF4}}$	-0.48 (0.43)	0.999	1.63 (0.26)	< 0.001	0.64 (0.80)	0.999
$\xi^w_{4,\text{AF3}}$	-0.41 (0.64)	0.999	1.52 (0.26)	< 0.001	-1.95 (1.05)	0.999
$\xi^w_{4,\text{ F7}}$	0.36 (0.66)	0.999	1.57 (0.27)	< 0.001	-1.58 (0.97)	0.999
$\xi^w_{4,\text{ F3}}$	0.27 (0.67)	0.999	1.54 (0.27)	< 0.001	-1.28 (0.97)	0.999
$\xi^w_{4,\text{FC5}}$	1.03 (0.72)	0.999	1.68 (0.26)	< 0.001	-0.32 (1.08)	0.999
$\xi^w_{4,\text{ T7}}$	0.62 (0.79)	0.999	1.66 (0.26)	< 0.001	0.25 (1.17)	0.999
$\xi^w_{4,\text{ P7}}$	-0.07 (0.66)	0.999	1.64 (0.27)	< 0.001	-0.41 (1.13)	0.999
$\xi^w_{4,\text{ O1}}$	0.39 (0.67)	0.999	1.53 (0.27)	< 0.001	-1.63 (1.00)	0.999
$\xi^w_{4,\text{ O2}}$	-0.01 (0.61)	0.999	1.65 (0.26)	< 0.001	-1.48 (1.02)	0.999
$\xi^w_{4,\text{ P8}}$	-0.01 (0.54)	0.999	1.64 (0.26)	< 0.001	-0.63 (0.89)	0.999
$\xi^w_{4,\text{ T8}}$	0.10 (0.56)	0.999	1.61 (0.26)	< 0.001	-1.38 (1.00)	0.999
$\xi^w_{4,\text{FC6}}$	0.10 (0.59)	0.999	1.66 (0.26)	< 0.001	-1.26 (1.01)	0.999
$\xi^w_{4,\text{ F4}}$	0.09 (0.67)	0.999	1.60 (0.26)	< 0.001	-1.61 (0.96)	0.999
$\xi^w_{4,\text{ F8}}$	0.03 (0.64)	0.999	1.63 (0.26)	< 0.001	-1.31 (1.02)	0.999
$\xi^w_{4,\text{AF4}}$	0.54 (0.65)	0.999	1.59 (0.26)	< 0.001	-1.62 (0.92)	0.999

Table 2: Coefficients, standard errors and adjusted p-values from univariate models on $\zeta^z_{m,i}$ and $\zeta^w_{m,ij}$ with significant interaction effects or main score effects.

	Score		Trauma		Score × Trauma	
	β_1 (SE)	p-value	β_2 (SE)	p-value	β_3 (SE)	p-value
ζ^z_{theta}	0.32 (0.31)	0.999	1.68 (0.25)	< 0.001	-1.56 (0.49)	0.015
ζ^w_{beta} at O_1	0.78 (0.33)	< 0.001	1.65 (0.26)	< 0.001	-1.38 (0.48)	0.036
ζ^w_{beta} at O_2	0.69 (0.34)	< 0.001	1.68 (0.26)	< 0.001	-0.41 (0.59)	0.999
ζ^w_{gamma} at O_1	0.62 (0.29)	0.999	1.57 (0.25)	< 0.001	-1.50 (0.49)	0.021
ζ^w_{alpha} at O_1	0.57 (0.32)	0.999	1.65 (0.26)	< 0.001	-1.252 (0.445)	0.048

Table 3: Median of errors $\|\phi - \widehat{\phi}\|_2$ for $\phi_r^z, \phi_r^w, r = 1, 2, 3$, (with median absolute deviations in parenthesis) with $N = 50, P = 100$ and $\sigma = 1$, over 200 simulation runs.

	ϕ_1^z	ϕ_2^z	ϕ_3^z	ϕ_1^w	ϕ_2^w	ϕ_3^w
$(\hat{\alpha}, \hat{\lambda}, \hat{\rho})$	0.72 (0.26)	1.23 (0.64)	2.99 (1.49)	0.35 (0.05)	0.44 (0.06)	0.84 (0.13)
$(\hat{\alpha}, 0, \hat{\rho})$	0.92 (0.36)	3.12 (1.32)	4.00 (2.03)	0.75 (0.10)	1.24 (0.20)	1.22 (0.12)
$(0, \hat{\lambda}, \hat{\rho})$	1.47 (0.41)	1.35 (0.74)	4.96 (1.81)	0.63 (0.22)	0.45 (0.08)	0.82 (0.14)
$(0, 0, \hat{\rho})$	3.78 (1.17)	5.79 (1.63)	6.28 (1.51)	1.92 (0.68)	2.90 (0.78)	2.23 (0.64)
$(\hat{\alpha}, \hat{\lambda}, 0)$	0.71 (0.25)	1.94 (1.44)	24.49 (0.00)	0.35 (0.05)	0.44 (0.06)	0.84 (0.13)
$(\hat{\alpha}, 0, 0)$	0.93 (0.36)	6.00 (2.78)	14.29 (7.77)	0.75 (0.10)	1.24 (0.19)	1.22 (0.12)
$(0, \hat{\lambda}, 0)$	1.50 (0.40)	1.99 (1.47)	24.41 (1.08)	0.63 (0.22)	0.45 (0.08)	0.82 (0.14)
$(0, 0, 0)$	4.39 (1.35)	7.99 (2.63)	14.52 (4.15)	1.92 (0.68)	2.90 (0.78)	2.23 (0.64)

Table 4: Median of errors $\|\phi - \widehat{\phi}\|_2$ for $\phi_r^z, \phi_r^w, r = 1, 2, 3$, (with median absolute deviations in parenthesis) with $N = 200, P = 100$ and $\sigma = 1$, over 200 simulation runs.

	ϕ_1^z	ϕ_2^z	ϕ_3^z	ϕ_1^w	ϕ_2^w	ϕ_3^w
$(\hat{\alpha}, \hat{\lambda}, \hat{\rho})$	0.36 (0.04)	0.57 (0.18)	1.40 (0.77)	0.35 (0.04)	0.39 (0.02)	0.47 (0.06)
$(\hat{\alpha}, 0, \hat{\rho})$	0.45 (0.05)	1.42 (0.40)	1.62 (0.91)	0.59 (0.04)	1.02 (0.08)	0.65 (0.06)
$(0, \hat{\lambda}, \hat{\rho})$	0.93 (0.25)	0.65 (0.25)	2.41 (0.72)	0.33 (0.03)	0.39 (0.03)	0.43 (0.06)
$(0, 0, \hat{\rho})$	2.08 (0.60)	2.89 (0.71)	3.01 (0.84)	1.09 (0.31)	1.71 (0.38)	1.12 (0.32)
$(\hat{\alpha}, \hat{\lambda}, 0)$	0.35 (0.04)	0.66 (0.26)	24.49 (0.00)	0.31 (0.02)	0.39 (0.02)	0.48 (0.06)
$(\hat{\alpha}, 0, 0)$	0.45 (0.05)	2.81 (1.35)	10.42 (5.20)	0.59 (0.04)	1.02 (0.09)	0.65 (0.06)
$(0, \hat{\lambda}, 0)$	0.97 (0.27)	0.82 (0.41)	23.75 (0.75)	0.33 (0.03)	0.39 (0.03)	0.43 (0.06)
$(0, 0, 0)$	2.26 (0.62)	4.07 (1.11)	11.52 (4.40)	1.09 (0.31)	1.71 (0.38)	1.12 (0.32)

Table 5: Median of specificity (proportion of zero elements correctly estimated as zero) and sensitivity (proportion of nonzero elements estimated as nonzero) for $\phi_r^z, \phi_r^w, r = 1, 2, 3$, with $N = 50, P = 100$ and $\sigma = 1$, over 200 simulation runs.

	Specificity						Sensitivity					
	ϕ_1^z	ϕ_2^z	ϕ_3^z	ϕ_1^w	ϕ_2^w	ϕ_3^w	ϕ_1^z	ϕ_2^z	ϕ_3^z	ϕ_1^w	ϕ_2^w	ϕ_3^w
$(\hat{\alpha}, \hat{\lambda}, \hat{\rho})$	0.99	0.99	1.00	0.99	1.00	0.70	0.87	0.85	1.00	0.88	0.88	0.84
$(\hat{\alpha}, 0, \hat{\rho})$	0.75	0.72	1.00	0.73	0.73	0.01	1.00	1.00	1.00	1.00	1.00	1.00
$(0, \hat{\lambda}, \hat{\rho})$	0.97	0.99	0.78	1.00	1.00	0.70	0.71	0.83	0.99	0.76	0.80	0.85
$(0, 0, \hat{\rho})$	0.01	0.01	0.01	0.01	0.01	0.01	1.00	1.00	1.00	1.00	1.00	1.00
$(\hat{\alpha}, \hat{\lambda}, 0)$	0.99	0.97	0.86	0.99	1.00	0.70	0.87	0.88	0.13	0.88	0.92	0.85
$(\hat{\alpha}, 0, 0)$	0.75	0.72	0.50	0.73	0.73	0.01	1.00	1.00	1.00	1.00	1.00	1.00
$(0, \hat{\lambda}, 0)$	0.97	0.97	0.80	1.00	1.00	0.70	0.71	0.83	0.21	0.76	0.80	0.85
$(0, 0, 0)$	0.01	0.01	0.01	0.01	0.01	0.01	1.00	1.00	1.00	1.00	1.00	1.00

Table 6: Median of specificity (proportion of zero elements correctly estimated as zero) and sensitivity (proportion of nonzero elements estimated as nonzero) for $\phi_r^z, \phi_r^w, r = 1, 2, 3$, with $N = 200, P = 100$ and $\sigma = 1$, over 200 simulation runs.

	Specificity						Sensitivity					
	ϕ_1^z	ϕ_2^z	ϕ_3^z	ϕ_1^w	ϕ_2^w	ϕ_3^w	ϕ_1^z	ϕ_2^z	ϕ_3^z	ϕ_1^w	ϕ_2^w	ϕ_3^w
$(\hat{\alpha}, \hat{\lambda}, \hat{\rho})$	1.00	1.00	1.00	1.00	1.00	0.79	0.87	0.88	1.00	0.92	0.88	0.84
$(\hat{\alpha}, 0, \hat{\rho})$	0.75	0.73	1.00	0.73	0.73	0.02	1.00	1.00	1.00	1.00	1.00	1.00
$(0, \hat{\lambda}, \hat{\rho})$	0.97	1.00	0.04	1.00	1.00	0.80	0.74	0.83	1.00	0.76	0.80	0.84
$(0, 0, \hat{\rho})$	0.03	0.02	0.02	0.03	0.02	0.02	1.00	1.00	1.00	1.00	1.00	1.00
$(\hat{\alpha}, \hat{\lambda}, 0)$	0.99	0.98	0.88	1.00	1.00	0.79	0.87	0.88	0.00	0.88	0.92	0.84
$(\hat{\alpha}, 0, 0)$	0.75	0.73	0.51	0.73	0.73	0.02	1.00	1.00	1.00	1.00	1.00	1.00
$(0, \hat{\lambda}, 0)$	0.97	1.00	0.81	1.00	1.00	0.80	0.74	0.83	0.18	0.76	0.80	0.84
$(0, 0, 0)$	0.03	0.02	0.02	0.03	0.02	0.02	1.00	1.00	1.00	1.00	1.00	1.00

Table 7: Average RMSEs (root mean square error) for estimating the principal component scores by the eight described estimation methods with $N = 50, 100, 200, P = 100$ and $\sigma = 1$, over 200 simulation runs. The values are augmented 100 times for better presentation

		ξ_1^z	ξ_2^z	ξ_3^z	ξ_1^w	ξ_2^w	ξ_3^w
N=50	$(\hat{\alpha}, \hat{\lambda}, \hat{\rho})$	15.1	16.4	56.7	19.7	16.0	10.3
	$(\hat{\alpha}, 0, \hat{\rho})$	15.2	18.8	57.0	21.3	16.1	11.5
	$(0, \hat{\lambda}, \hat{\rho})$	16.0	17.9	57.6	22.4	17.3	11.3
	$(0, 0, \hat{\rho})$	20.5	26.0	63.7	25.8	21.4	15.7
	$(\hat{\alpha}, \hat{\lambda}, 0)$	15.4	27.9	87.8	44.0	28.5	11.7
	$(\hat{\alpha}, 0, 0)$	16.3	36.9	92.3	54.2	26.5	13.3
	$(0, \hat{\lambda}, 0)$	17.0	30.8	91.6	54.0	29.3	12.5
	$(0, 0, 0)$	24.6	45.3	103.2	71.0	31.5	19.2
N=100	$(\hat{\alpha}, \hat{\lambda}, \hat{\rho})$	10.2	9.3	63.8	14.2	12.1	8.7
	$(\hat{\alpha}, 0, \hat{\rho})$	10.5	10.4	61.3	14.9	12.2	9.4
	$(0, \hat{\lambda}, \hat{\rho})$	10.9	10.4	64.2	15.2	12.7	9.2
	$(0, 0, \hat{\rho})$	14.0	17.0	60.6	17.6	15.4	12.3
	$(\hat{\alpha}, \hat{\lambda}, 0)$	10.1	16.5	93.3	43.0	25.2	10.2
	$(\hat{\alpha}, 0, 0)$	10.8	20.9	107.1	53.4	25.1	11.1
	$(0, \hat{\lambda}, 0)$	11.2	18.4	97.1	48.5	26.8	10.8
	$(0, 0, 0)$	16.9	31.4	116.9	70.4	28.3	15.3
N=200	$(\hat{\alpha}, \hat{\lambda}, \hat{\rho})$	7.1	6.5	59.0	10.3	9.3	7.1
	$(\hat{\alpha}, 0, \hat{\rho})$	7.0	7.0	59.0	11.1	9.6	7.5
	$(0, \hat{\lambda}, \hat{\rho})$	7.6	7.6	59.3	11.2	9.9	7.5
	$(0, 0, \hat{\rho})$	10.1	12.3	58.2	13.1	11.9	9.4
	$(\hat{\alpha}, \hat{\lambda}, 0)$	7.1	8.7	89.1	38.5	23.3	8.8
	$(\hat{\alpha}, 0, 0)$	7.1	13.1	99.3	42.5	22.7	8.8
	$(0, \hat{\lambda}, 0)$	7.9	11.8	91.1	47.6	23.9	9.2
	$(0, 0, 0)$	11.5	20.2	111.6	60.5	23.8	11.2

Table 8: Median of errors $\|\phi - \widehat{\phi}\|_2$ for $\phi_r^z, \phi_r^w, r = 1, 2, 3$, (with median absolute deviations in parenthesis) with $\delta = 20\%, 30\%$ and 50% over 200 simulation runs. Other parameters $N = 100, P = 100$ and $\sigma = 1$.

		ϕ_1^z	ϕ_2^z	ϕ_3^z	ϕ_1^w	ϕ_2^w	ϕ_3^w
$\delta = 20\%$	$(\hat{\alpha}, \hat{\lambda}, \hat{\rho})$	0.48 (0.14)	0.87 (0.38)	2.14 (1.03)	0.35 (0.05)	0.41 (0.04)	0.66 (0.10)
	$(\hat{\alpha}, 0, \hat{\rho})$	0.60 (0.15)	1.77 (0.58)	2.42 (1.40)	0.66 (0.08)	1.11 (0.11)	0.91 (0.10)
	$(0, \hat{\lambda}, \hat{\rho})$	1.25 (0.31)	0.97 (0.46)	3.26 (1.13)	0.41 (0.09)	0.41 (0.05)	0.60 (0.09)
	$(0, 0, \hat{\rho})$	2.61 (0.79)	3.76 (0.91)	4.27 (1.09)	1.43 (0.43)	2.14 (0.58)	1.57 (0.44)
$\delta = 30\%$	$(\hat{\alpha}, \hat{\lambda}, \hat{\rho})$	0.49 (0.15)	0.91 (0.41)	2.18 (1.15)	0.34 (0.04)	0.41 (0.04)	0.66 (0.10)
	$(\hat{\alpha}, 0, \hat{\rho})$	0.60 (0.15)	1.89 (0.68)	2.54 (1.50)	0.66 (0.08)	1.11 (0.11)	0.91 (0.10)
	$(0, \hat{\lambda}, \hat{\rho})$	1.25 (0.31)	0.96 (0.44)	3.49 (1.13)	0.41 (0.09)	0.41 (0.05)	0.60 (0.09)
	$(0, 0, \hat{\rho})$	2.67 (0.78)	3.89 (1.00)	4.46 (1.23)	1.43 (0.43)	2.14 (0.58)	1.57 (0.44)
$\delta = 50\%$	$(\hat{\alpha}, \hat{\lambda}, \hat{\rho})$	0.48 (0.13)	0.88 (0.38)	3.39 (1.86)	0.34 (0.04)	0.41 (0.03)	0.66 (0.10)
	$(\hat{\alpha}, 0, \hat{\rho})$	0.60 (0.15)	2.18 (0.91)	3.65 (2.09)	0.66 (0.08)	1.11 (0.11)	0.91 (0.10)
	$(0, \hat{\lambda}, \hat{\rho})$	1.25 (0.31)	0.97 (0.46)	4.83 (1.58)	0.41 (0.09)	0.41 (0.05)	0.60 (0.05)
	$(0, 0, \hat{\rho})$	2.78 (0.81)	4.38 (1.21)	5.53 (1.60)	1.43 (0.43)	2.14 (0.58)	1.57 (0.44)

Table 9: Median of specificity (proportion of zero elements correctly estimated as zero) and sensitivity (proportion of nonzero elements estimated as nonzero) for $\phi_r^z, \phi_r^w, r = 1, 2, 3$, with $\delta = 20\%, 30\%$ and 50% over 200 simulation runs. Other parameters $N = 100, P = 100$ and $\sigma = 1$.

		Specificity						Sensitivity					
		ϕ_1^z	ϕ_2^z	ϕ_3^z	ϕ_1^w	ϕ_2^w	ϕ_3^w	ϕ_1^z	ϕ_2^z	ϕ_3^z	ϕ_1^w	ϕ_2^w	ϕ_3^w
$\delta = 20\%$	$(\hat{\alpha}, \hat{\lambda}, \hat{\rho})$	0.99	1.00	1.00	0.99	1.00	0.75	0.87	0.83	1.00	0.92	0.92	0.85
	$(\hat{\alpha}, 0, \hat{\rho})$	0.75	0.73	1.00	0.73	0.73	0.01	1.00	1.00	1.00	1.00	1.00	1.00
	$(0, \hat{\lambda}, \hat{\rho})$	0.97	1.00	0.02	1.00	1.00	0.76	0.71	0.83	1.00	0.76	0.80	0.85
	$(0, 0, \hat{\rho})$	0.02	0.01	0.02	0.02	0.02	0.01	1.00	1.00	1.00	1.00	1.00	1.00
$\delta = 30\%$	$(\hat{\alpha}, \hat{\lambda}, \hat{\rho})$	0.99	0.99	1.00	0.99	1.00	0.75	0.87	0.83	1.00	0.92	0.92	0.85
	$(\hat{\alpha}, 0, \hat{\rho})$	0.75	0.73	1.00	0.73	0.73	0.01	1.00	1.00	1.00	1.00	1.00	1.00
	$(0, \hat{\lambda}, \hat{\rho})$	0.97	1.00	0.03	1.00	1.00	0.75	0.71	0.83	1.00	0.76	0.80	0.85
	$(0, 0, \hat{\rho})$	0.02	0.01	0.02	0.02	0.02	0.01	1.00	1.00	1.00	1.00	1.00	1.00
$\delta = 50\%$	$(\hat{\alpha}, \hat{\lambda}, \hat{\rho})$	0.99	0.99	1.00	1.00	1.00	0.74	0.84	0.88	1.00	0.88	0.92	0.85
	$(\hat{\alpha}, 0, \hat{\rho})$	0.75	0.73	1.00	0.73	0.73	0.01	1.00	1.00	1.00	1.00	1.00	1.00
	$(0, \hat{\lambda}, \hat{\rho})$	0.97	1.00	0.03	1.00	1.00	0.76	0.71	0.83	1.00	0.76	0.80	0.85
	$(0, 0, \hat{\rho})$	0.02	0.01	0.02	0.02	0.02	0.01	1.00	1.00	1.00	1.00	1.00	1.00

Table 10: Median of errors $\|\phi - \widehat{\phi}\|_2$ for $\phi_r^z, \phi_r^w, r = 1, 2, 3$, (with median absolute deviations in parenthesis) with $M = 15$ over 200 simulation runs for $N = 50, 100$ and 200. Other parameters $P = 100$ and $\sigma = 1$.

		ϕ_1^z	ϕ_2^z	ϕ_3^z	ϕ_1^w	ϕ_2^w	ϕ_3^w
$N = 50$	$(\hat{\alpha}, \hat{\lambda}, \hat{\rho})$	1.57 (0.54)	3.59 (2.37)	6.71 (3.10)	0.78 (0.09)	1.26 (0.29)	1.48 (0.21)
	$(\hat{\alpha}, 0, \hat{\rho})$	1.98 (0.75)	6.69 (2.69)	7.24 (3.44)	1.71 (0.31)	2.77 (0.62)	2.37 (0.32)
	$(0, \hat{\lambda}, \hat{\rho})$	2.96 (1.18)	5.29 (3.56)	9.53 (3.36)	0.86 (0.19)	1.63 (0.74)	1.52 (0.17)
	$(0, 0, \hat{\rho})$	8.56 (2.68)	13.08 (3.56)	12.54 (3.91)	4.63 (1.46)	6.45 (1.87)	4.75 (1.43)
	$(\hat{\alpha}, \hat{\lambda}, 0)$	1.57 (0.55)	7.64 (6.24)	32.92 (21.85)	0.78 (0.08)	1.26 (0.29)	1.49 (0.21)
	$(\hat{\alpha}, 0, 0)$	2.10 (0.80)	13.04 (6.72)	29.27 (16.99)	1.71 (0.31)	2.76 (0.61)	2.34 (0.30)
	$(0, \hat{\lambda}, 0)$	3.53 (1.70)	9.14 (7.07)	32.42 (17.46)	0.86 (0.19)	1.63 (0.74)	1.52 (0.17)
	$(0, 0, 0)$	9.53 (3.15)	18.71 (5.36)	32.22 (10.39)	4.63 (1.46)	6.45 (1.87)	4.75 (1.43)
$N = 100$	$(\hat{\alpha}, \hat{\lambda}, \hat{\rho})$	1.06 (0.30)	1.87 (0.95)	4.62 (2.01)	0.72 (0.05)	1.44 (0.29)	1.05 (0.14)
	$(\hat{\alpha}, 0, \hat{\rho})$	1.37 (0.41)	4.20 (1.55)	4.93 (2.22)	1.43 (0.17)	2.50 (0.36)	1.64 (0.20)
	$(0, \hat{\lambda}, \hat{\rho})$	2.25 (0.80)	2.80 (1.65)	7.03 (2.19)	0.77 (0.13)	1.55 (0.43)	1.09 (0.11)
	$(0, 0, \hat{\rho})$	6.33 (1.52)	8.64 (2.23)	9.02 (2.29)	3.13 (1.04)	4.74 (1.21)	3.32 (0.87)
	$(\hat{\alpha}, \hat{\lambda}, 0)$	1.15 (0.36)	3.61 (2.65)	25.72 (15.50)	0.72 (0.05)	1.44 (0.30)	1.07 (0.16)
	$(\hat{\alpha}, 0, 0)$	1.46 (0.47)	8.32 (4.08)	21.79 (10.59)	1.41 (0.15)	2.50 (0.36)	1.64 (0.18)
	$(0, \hat{\lambda}, 0)$	2.37 (0.92)	3.43 (2.32)	28.74 (14.01)	0.77 (0.13)	1.57 (0.43)	1.09 (0.11)
	$(0, 0, 0)$	0.68 (1.73)	12.28 (3.47)	25.93 (8.93)	3.12 (1.03)	4.74 (1.17)	3.29 (0.83)
$N = 200$	$(\hat{\alpha}, \hat{\lambda}, \hat{\rho})$	0.83 (0.13)	1.43 (0.55)	3.20 (1.51)	0.76 (0.04)	1.60 (0.31)	0.78 (0.14)
	$(\hat{\alpha}, 0, \hat{\rho})$	1.05 (0.14)	3.28 (0.96)	3.34 (1.63)	1.32 (0.11)	2.32 (0.24)	1.27 (0.22)
	$(0, \hat{\lambda}, \hat{\rho})$	1.42 (0.29)	1.78 (0.76)	4.89 (1.84)	0.79 (0.07)	1.54 (0.19)	0.80 (0.10)
	$(0, 0, \hat{\rho})$	4.49 (1.17)	6.43 (1.49)	6.37 (1.88)	2.37 (0.64)	3.50 (0.75)	2.54 (0.81)
	$(\hat{\alpha}, \hat{\lambda}, 0)$	0.81 (0.12)	1.45 (0.65)	30.25 (19.61)	0.76 (0.04)	1.60 (0.31)	0.78 (0.14)
	$(\hat{\alpha}, 0, 0)$	1.04 (0.13)	5.57 (2.46)	18.18 (8.31)	1.32 (0.11)	2.23 (0.24)	1.27 (0.23)
	$(0, \hat{\lambda}, 0)$	1.58 (0.33)	2.32 (1.44)	35.46 (19.31)	0.79 (0.07)	1.54 (0.19)	0.80 (0.10)
	$(0, 0, 0)$	5.05 (1.35)	9.25 (2.96)	24.43 (9.75)	2.37 (0.64)	3.50 (0.75)	2.54 (0.81)

Table 11: Median of specificity (proportion of zero elements correctly estimated as zero) and sensitivity (proportion of nonzero elements estimated as nonzero) for $\phi_r^z, \phi_r^w, r = 1, 2, 3$, with $M = 15$ for $N = 50, 100$ and 200 over 200 simulation runs. Other parameters $P = 100$ and $\sigma = 1$.

		Specificity						Sensitivity					
		ϕ_1^z	ϕ_2^z	ϕ_3^z	ϕ_1^w	ϕ_2^w	ϕ_3^w	ϕ_1^z	ϕ_2^z	ϕ_3^z	ϕ_1^w	ϕ_2^w	ϕ_3^w
$N = 50$	$(\hat{\alpha}, \hat{\lambda}, \hat{\rho})$	1	0.98	1	0.99	1	0.71	0.84	0.93	1	0.85	0.91	0.92
	$(\hat{\alpha}, 0, \hat{\rho})$	0.86	0.84	1	0.84	0.84	0.5	1	1	1	1	1	1
	$(0, \hat{\lambda}, \hat{\rho})$	0.96	0.96	0.81	0.97	0.97	0.5	0.74	0.93	1	0.81	0.89	0.91
	$(0, 0, \hat{\rho})$	0.02	0.01	0.01	0.02	0.01	0.01	1	1	1	1	1	1
	$(\hat{\alpha}, \hat{\lambda}, 0)$	1	0.97	0.9	1	0.99	0.71	0.83	0.97	1	0.84	0.91	0.92
	$(\hat{\alpha}, 0, 0)$	0.86	0.84	0.75	0.84	0.84	0.5	1	1	1	1	1	1
	$(0, \hat{\lambda}, 0)$	0.96	0.95	0.76	0.97	0.97	0.5	0.74	0.93	0.99	0.81	0.89	0.91
	$(0, 0, 0)$	0.02	0.01	0.01	0.02	0.01	0.01	1	1	1	1	1	1
$N = 100$	$(\hat{\alpha}, \hat{\lambda}, \hat{\rho})$	0.99	0.99	1	1	0.99	0.73	0.86	0.92	1	0.87	0.97	0.93
	$(\hat{\alpha}, 0, \hat{\rho})$	0.86	0.84	1	0.84	0.65	0.5	1	1	1	1	1	1
	$(0, \hat{\lambda}, \hat{\rho})$	0.96	0.96	0.84	0.97	0.97	0.51	0.76	0.92	0.99	0.87	0.96	0.92
	$(0, 0, \hat{\rho})$	0.03	0.02	0.02	0.03	0.02	0.02	1	1	1	1	1	1
	$(\hat{\alpha}, \hat{\lambda}, 0)$	0.99	0.98	0.93	1	0.99	0.73	0.85	0.92	1	0.87	0.96	0.93
	$(\hat{\alpha}, 0, 0)$	0.86	0.84	0.75	0.84	0.65	0.5	1	1	1	1	1	1
	$(0, \hat{\lambda}, 0)$	0.96	0.97	0.81	0.97	0.97	0.51	0.75	0.91	0.98	0.87	0.96	0.92
	$(0, 0, 0)$	0.03	0.02	0.02	0.03	0.02	0.02	1	1	1	1	1	1
$N = 200$	$(\hat{\alpha}, \hat{\lambda}, \hat{\rho})$	1	0.99	1	1	0.98	0.74	0.87	0.93	1	0.89	0.99	0.93
	$(\hat{\alpha}, 0, \hat{\rho})$	0.86	0.84	1	0.84	0.51	0.5	1	1	1	1	1	1
	$(0, \hat{\lambda}, \hat{\rho})$	0.97	0.97	0.91	0.98	0.97	0.54	0.77	0.92	0.99	0.89	0.97	0.91
	$(0, 0, \hat{\rho})$	0.04	0.03	0.02	0.04	0.03	0.02	1	1	1	1	1	1
	$(\hat{\alpha}, \hat{\lambda}, 0)$	1	0.98	0.93	1	0.98	0.74	0.87	0.92	1	0.89	0.99	0.93
	$(\hat{\alpha}, 0, 0)$	0.86	0.84	0.76	0.84	0.51	0.5	1	1	1	1	1	1
	$(0, \hat{\lambda}, 0)$	0.97	0.97	0.89	0.98	0.97	0.54	0.77	0.92	0.98	0.89	0.97	0.91
	$(0, 0, 0)$	0.04	0.03	0.03	0.04	0.03	0.02	1	1	1	1	1	1

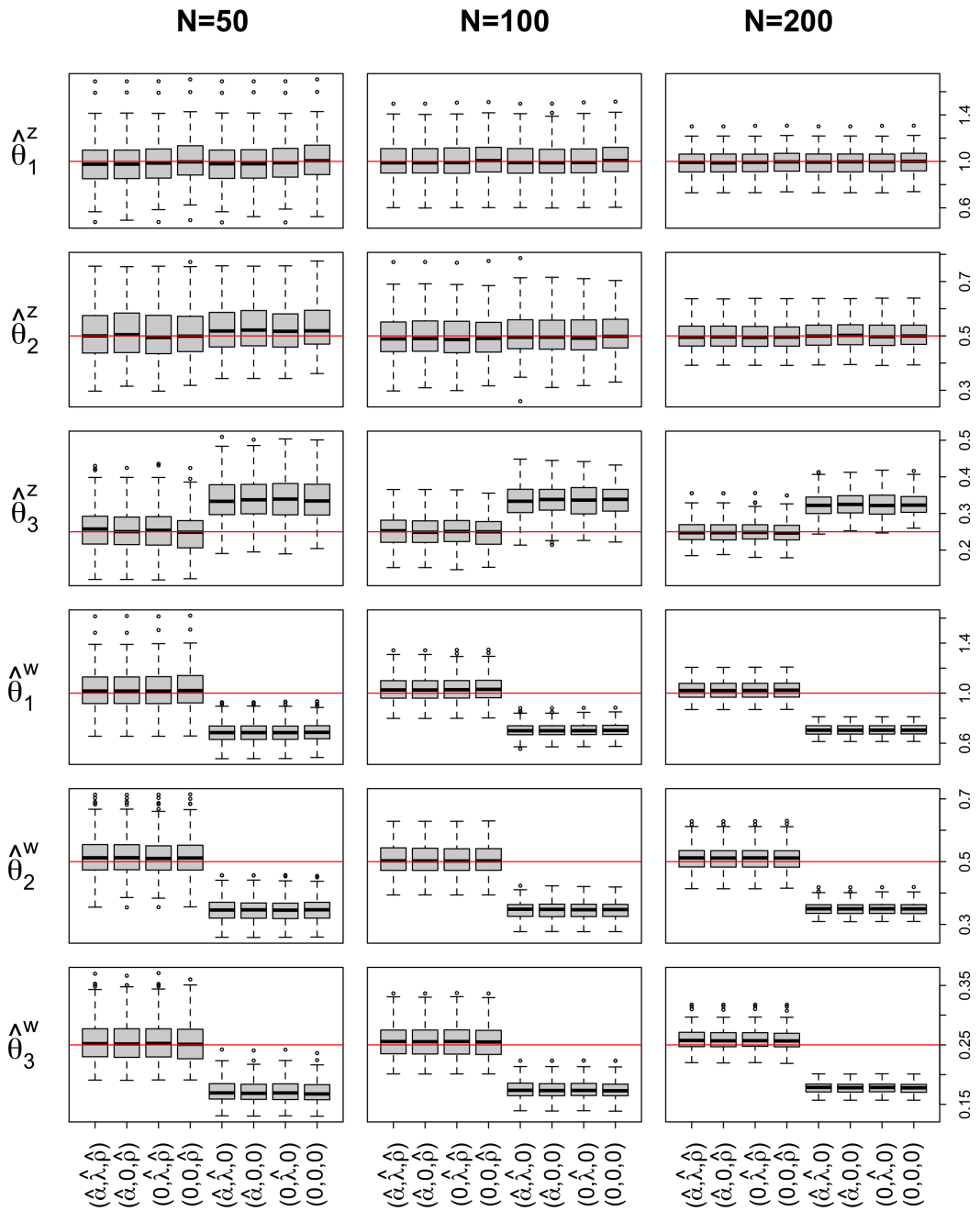


Fig. 8: Boxplots of eigenvalues estimated by the top four estimation methods, with $M = 15$ for $N = 50, 100$ and 200 over 200 simulation runs. Other parameters $P = 100$ and $\sigma = 1$.

Table 12: Average RMSEs (root mean square error) for estimating the principal component scores by the eight described estimation methods with $M = 15, N = 50, 100, 200, P = 100$ and $\sigma = 1$, over 200 simulation runs. The values are augmented 100 times for better presentation

		ξ_1^z	ξ_2^z	ξ_3^z	ξ_1^w	ξ_2^w	ξ_3^w
N=50	$(\hat{\alpha}, \hat{\lambda}, \hat{\rho})$	15.1	14.1	46.0	18.7	15.0	9.5
	$(\hat{\alpha}, 0, \hat{\rho})$	15.0	15.0	46.7	19.7	15	10.3
	$(0, \hat{\lambda}, \hat{\rho})$	16.5	17.0	45.7	20.3	16.4	10.4
	$(0, 0, \hat{\rho})$	20.7	25.0	48.5	23.6	19.9	15.0
	$(\hat{\alpha}, \hat{\lambda}, 0)$	15.6	27.5	101.6	54.2	28.2	11.2
	$(\hat{\alpha}, 0, 0)$	16.2	29.2	102.8	55.5	26.1	12.5
	$(0, \hat{\lambda}, 0)$	17.8	30.4	102.4	59.5	29.8	12.5
	$(0, 0, 0)$	25.5	40.6	113.1	72.4	31.3	18.6
N=100	$(\hat{\alpha}, \hat{\lambda}, \hat{\rho})$	8.9	8.6	55.0	12.6	10.8	6.8
	$(\hat{\alpha}, 0, \hat{\rho})$	8.9	9.2	55.0	13.3	10.8	7.3
	$(0, \hat{\lambda}, \hat{\rho})$	9.9	10.3	56.4	14.6	11.9	7.4
	$(0, 0, \hat{\rho})$	13.2	16.3	55.2	16.3	14.5	10.7
	$(\hat{\alpha}, \hat{\lambda}, 0)$	9.1	15.6	102.8	47.9	26.1	8.3
	$(\hat{\alpha}, 0, 0)$	9.7	19.2	113.1	54.5	25.3	9.2
	$(0, \hat{\lambda}, 0)$	10.4	16.7	104.7	54.4	29.3	9.0
	$(0, 0, 0)$	16.2	31.5	122.1	70.6	28.7	13.6
N=200	$(\hat{\alpha}, \hat{\lambda}, \hat{\rho})$	6.5	6.2	49.3	8.9	7.9	4.9
	$(\hat{\alpha}, 0, \hat{\rho})$	6.5	6.5	49	9.5	8.0	5.3
	$(0, \hat{\lambda}, \hat{\rho})$	7.0	7.0	50.0	9.5	8.5	5.4
	$(0, 0, \hat{\rho})$	9.5	11.7	51.3	11.3	10.4	8.0
	$(\hat{\alpha}, \hat{\lambda}, 0)$	6.6	10.2	100.2	41.4	24.9	6.8
	$(\hat{\alpha}, 0, 0)$	7.0	12.3	110.9	45.6	23.5	6.8
	$(0, \hat{\lambda}, 0)$	7.3	15.6	107.4	60.9	26.4	7.4
	$(0, 0, 0)$	11.3	22.2	127.7	70.8	25.3	10.6