Multilevel Dynamic Generalized Structured Component Analysis for Brain Connectivity Analysis in Functional Neuroimaging Data

Kwanghee Jung[1], Yoshio Takane[2], Heungsun Hwang[3], Todd S. Woodward[4].

1. University of Texas Health Science Center at Houston
2. University of Victoria
3. McGill University
4. University of British Columbia

Correspondence regarding this article should be sent to Kwanghee Jung, Department of Pediatrics, The University of Texas Health Science Center at Houston, Children's Learning Institute, 7000 Fannin UCT 2373J, Houston, Texas 77030, USA. Matlab programs that carried out the computations reported in the paper are available upon request.

E-Mail: projectionsvd@gmail.com Phone: 713-500-3727 Fax: 713-500-0386

Multilevel Dynamic Generalized Structured Component Analysis for Brain Connectivity Analysis in Functional Neuroimaging Data

Abstract

We extend Dynamic GSCA (Generalized Structured Component Analysis) to enhance its data-analytic capability in structural equation modeling of multi-subject time series data. Time series data of multiple subjects are typically hierarchically structured, where time points are nested within subjects who are in turn nested within a group. The proposed approach, named Multilevel Dynamic GSCA, accommodates the nested structure in time series data. Explicitly taking the nested structure into account, the proposed method allows investigating subject-wise variability of the loadings and path coefficients by looking at the variance estimates of the corresponding random effects, as well as fixed loadings between observed and latent variables and fixed path coefficients between latent variables. We demonstrate the effectiveness of the proposed approach by applying the method to the multi-subject functional neuroimaging data for brain connectivity analysis, where time series data-level measurements are nested within subjects.

Key words: generalized structured component analysis, multilevel analysis, structural equation modeling, alternating least squares (ALS) algorithm, brain connectivity analysis, time series data, functional neuroimaging, multi-subject data.

1. Introduction

One goal of brain connectivity analysis is to quantify the directionality as well as the magnitude of the influence that one neural system exerts on another (Friston, 1994). Specific brain regions of interest (ROIs) can be selected a priori, and their directional relationships modeled and tested. For brain connectivity analysis, Dynamic GSCA (Generalized Structured Component Analysis) (Jung et al., 2012) has extended the original GSCA (Hwang and Takane, 2004, 2014) by incorporating a multivariate autoregressive model to deal with longitudinal/time series data (e.g., functional MRI data). Dynamic GSCA is equipped with both measurement and structural models, thereby enabling to define latent variables (i.e., ROIs; Regions of Interest) by simultaneously taking into account both indicator variables (e.g., BOLD signals in fMRI) specific to ROIs and the relationships among ROIs. Furthermore, Dynamic GSCA incorporates new features such as time lagged and stimulus effects. Time lagged effects denote that an earlier state of ROIs exert influence on a later state. Stimulus effects include direct effects of experimental stimuli on specific ROIs and modulating effects on connections between ROIs. These new features are of great help in capturing the dynamic nature of time series data and accommodating more elaborate research designs such as event-related designs in functional neuroimaging studies. For parameter estimation, Dynamic GSCA adopts a least squares estimation procedure, so that it does not require rigid distributional assumptions and can avoid model identification problems and improper solutions. Note that Dynamic GSCA subsumes the existing path-analytic approaches to brain connectivity analysis such as unified SEM (Kim et al., 2007), the extended unified SEM (Gates et al, 2011), and structural vector autoregression (SVAR) method (Chen et al., 2011) as a special case by incorporating both measurement and structural models in a unified framework.

Despite its theoretical and empirical contributions, Dynamic GSCA is currently capable of analyzing a single sample of data at a time. However, in functional neuroimaging data, neuronal activities of each subject are measured repeatedly over time in voxels (e.g., BOLD signals in fMRI) and subjects usually belong to an experimental group. As such, functional neuroimaging data are typically hierarchically structured, where time points are nested within subjects who are in turn nested across experimental groups. Such a nested structure is likely to lead to dependence in the data. Unless this dependence is taken into account properly, solutions obtained tend to be inaccurate (Snijders and Bosker, 1999). It is worth noting that BOLD signals are vascular responses to neuronal activities rather than direct neuronal signals (Greve et al, 2013). The relationship between BOLD signals and neuronal activities are often modeled by a hemodynamic response function. An autoregressive based method like the proposed method needs to assume that the hemodynamic response function is invariant across different brain regions.

In this paper, we extend Dynamic GSCA to the brain connectivity analysis of multi-subject functional neuroimaging data. The proposed approach, named Multilevel Dynamic GSCA,

explicitly accommodates the nested structure of both indicators (e.g., BOLD signals in voxels) and latent variables (e.g., ROIs) in functional neuroimaging data.

This paper is organized as follows. In Section 2, we present the model for two-level Dynamic GSCA in detail. In Section 3, we illustrate the empirical usefulness of Multilevel Dynamic GSCA with a real functional neuroimaging example. In the final section, we summarize the previous sections and discuss further prospects for Multilevel Dynamic GSCA.

2. The Two-Level Dynamic GSCA

We begin with the measurement model. Let Z_{ij} indicate a T by V_i matrix of observed variables (e.g., voxels) pertaining to the i-th latent variable (e.g., ROIs) ($i = 1, \dots, P$; P is the number of latent variables) for the j-th subject ($j=1,\dots J$; J is the number of subjects), where T and V_i indicate the number of time points and the number of observed variables for the i-th latent variable, respectively. We assume that Z_{ij} is columnwise standardized, although this is not essential (i.e., an option can be made not to standardize the data). We also define two matrices Z_j and Z_i based on Z_{ij}. The former is a row block matrix of Z_{ij} for the j-th subject, namely

$$Z_j = [Z_{1j},\ Z_{2j}, \dots, Z_{Pj}], \tag{1}$$

and the latter is a column block matrix for the i-th latent variable, i.e.,

$$Z_i = \begin{bmatrix} Z_{i1} \\ \vdots \\ Z_{iJ} \end{bmatrix}. \tag{2}$$

Let γ_{ij} denote T-component vector of the i-th latent variable for the j-th subject. Let Γ_j denote a row block matrix with γ_{ij} as the i-th column vector, namely

$$\Gamma_j = [\gamma_{1j}, \gamma_{2j}, \dots, \gamma_{Pj}], \tag{3}$$

and let Γ_i denote a tall supervector of γ_{ij} placed on top of each other, i.e.,

$$\Gamma_i = \begin{bmatrix} \gamma_{i1} \\ \vdots \\ \gamma_{iJ} \end{bmatrix}. \tag{4}$$

This vector is defined as an exact linear combination of the observed variables, i.e., $\Gamma_i = Z_i w_i$, where $w_i (i = 1, \dots, P)$ is a vector of component weights for a latent variable (Γ_i), and it is scaled to have unit variance with the normalization restriction that

$$(1/JT)\ \Gamma_i' \Gamma_i = (1/JT)\ w_i' Z_i' Z_i w_i = 1 \tag{5}$$

for each *i*. Also, we define a block diagonal matrix D_W with w_i as the *i*-th diagonal block,

$$D_W = \text{bdiag}([w_1, w_2, \ldots, w_P]). \tag{6}$$

(the operator bdiag forms a block matrix with vectors/matrices in its argument as diagonal blocks). A similar block diagonal matrix for the *j*-th subject with c_{ij} as the *i*-th diagonal block,

$$D'_{\mathrm{c}_j} = \text{bdiag}([c_{1j}, c_{2j}, \ldots, c_{Pj}]) \,, \tag{7}$$

where c_{ij} is the vector of loadings to applied to γ_{ij} to best approximate Z_j for the latent variable *i* in the *j*-th subject. Then, the Level-1 measurement model can be written as

$$Z_j = \Gamma_j D'_{\mathrm{c}_j} + E_{Mj} = Z_j D_W D'_{\mathrm{c}_j} + E_{Mj}, \tag{8}$$

where $\Gamma_j = Z_j D_W$. Let D'_C denote a matrix of Level-2 fixed loadings. Let D'_{λ_j} indicate a matrix of Level-2 random loadings, which are assumed to vary across subjects. The level-2 model is given by

$$D'_{Cj} = D'_C + D'_{\lambda_j}, \tag{9}$$

where $D'_C = \sum_{j=1}^{J} D'_{Cj} \,/J$ and $D'_{\lambda_j} = D'_{Cj} - D'_C$ (*j=1*,…*J*). The Level-2 model indicate that Level-1 loadings (D'_{c_j}) are conceived as varying over the population of a Level-2 unit, and the variance of each random effect represents the inter-subject variability of the corresponding loading. Thus, the measurement model for two-level Dynamic GSCA is as follows:

$$Z_j = \Gamma_j D'_C + \Gamma_j D'_{\lambda_j} + E_{Mj} = Z_j D_W D'_C + Z_j D_W D'_{\lambda_j} + E_{Mj}. \tag{10}$$

We now specify the structural model for two-level Dynamic GSCA. For simplicity, stimulus effects will not be considered in this paper. The example in this paper does not need them, and inclusion of these effects substantially complicates the derivation of the model and an estimation procedure. Note that a large portion of fMRI studies are resting-state fMRI studies, where participants are not engaged in explicit tasks, and thus no stimulus effects are necessary (van den Heuvel and Hulshoff Pol, 2010). As in Dynamic GSCA, shift matrices are introduced to capture time lagged effects. A series of shift matrices represent both contemporaneous and lagged effects. Contemporaneous effects denote concurrent relations between latent variables at the same time point. The shift matrix with time lag 0 ($S_0 = I_T$; the identity matrix of order *T*) represents contemporaneous effect. The matrix S_l ($l = 1, \ldots, q$) represents time lagged effects among latent variables. The subscript l indexes the order of lags. The matrix S_1, for example, represents the effect of time $t-1$ on time *t*. Premultiplying Γ_j by S_l shifts down the rows of Γ_j by l rows and defines the matrix of effect of latent variables at time $t-l$ on time *t*. Let A'_{lj} denote a matrix of

Level-1 path coefficients of latent variables, which are conceived to be different across subjects. Then the Level-1 structural model is given by

$$\Gamma_j = \sum_{l=0}^{q}(S_l\,\Gamma_j A'_{lj}) + E_{Sj}. \tag{11}$$

Let A'_l denote a matrix of Level-2 fixed path coefficients, and let Θ'_{lj} denote a matrix of Level-2 random effects of path coefficients, which may vary across subjects. The Level-2 model is given by

$$A'_{lj} = A'_l + \Theta'_{lj}, \tag{12}$$

where $A'_l = \sum_{j=1}^{J} A'_{lj}\,/J$ and $\Theta'_{lj} = A'_{lj} - A'_l$ (*j=1,…J*). Thus, the structural model for two-level Dynamic GSCA is as follows:

$$\Gamma_j = \sum_{l=0}^{q}(S_l\,\Gamma_j A'_l + S_l\Gamma_j\Theta'_{lj}) + E_{Sj}. \tag{13}$$

***** Insert Figure 1 *****

For illustration, an exemplary path diagram of a Level-1 structural model is displayed in Figure 1, in which we assume that there are three latent variables. In this model, we also assume that the three latent variables are fully and reciprocally (bidirectionally) connected. Solid arrows indicate contemporaneous effects between latent variables, while dashed arrows indicate the time lagged effects. We further assume that there are autoregressive effects of lag *1* of latent variables on themselves. The Level-1 structural model can be written as

$$\Gamma_j = S_0\Gamma_j A'_{0j} + S_1\Gamma_j A'_{1j} + E_{Sj}. \tag{14}$$

Then, the structural model for two-level Dynamic GSCA can be written as follows:

$$\Gamma_j = S_0\Gamma_j A'_0 + S_0\Gamma_j\Theta'_{0j} + S_1\Gamma_j A'_1 + S_1\Gamma_j\Theta'_{1j} + E_{Sj}, \tag{15}$$

where

$$A'_0 = \begin{bmatrix} 0 & a_4 & a_7 \\ a_2 & 0 & a_8 \\ a_3 & a_6 & 0 \end{bmatrix}, \tag{16}$$

$$\Theta'_{0j} = \begin{bmatrix} 0 & \theta_{4j} & \theta_{7j} \\ \theta_{2j} & 0 & \theta_{8j} \\ \theta_{3j} & \theta_{6j} & 0 \end{bmatrix}, \tag{17}$$

$$A_1' = \begin{bmatrix} a_1 & 0 & 0 \\ 0 & a_5 & 0 \\ 0 & 0 & a_9 \end{bmatrix}, \tag{18}$$

$$\Theta_{1j}' = \begin{bmatrix} \theta_{1j} & 0 & 0 \\ 0 & \theta_{5j} & 0 \\ 0 & 0 & \theta_{9j} \end{bmatrix}. \tag{19}$$

In general, the rows of A_l' and Θ_{lj}' represent latent variables exerting influence, whereas the columns represent latent variables being influenced.

In Multilevel Dynamic GSCA, the following least squares (LS) criterion is minimized for parameter estimation.

$$\emptyset = \sum_{j=1}^{J} \mathrm{SS}(E_j) = \sum_{j=1}^{J} \mathrm{tr}(E_j{}' E_j) \tag{20}$$

with respect to model parameters, where $E_j = [E_{Mj}, E_{Sj}]$. We use an alternating least squares (ALS) algorithm (e.g., de Leeuw, Young, & Takane, 1976) to obtain the LS estimates of parameters that minimize (16). A technical description of the ALS algorithm is provided in the Appendix.

Similarly to Dynamic GSCA, Multilevel Dynamic GSCA measures the overall goodness of fit of a hypothesized model by the portion of the total variance of all endogenous variables explained by model specifications. This is given by

$$\mathrm{FIT} = 1 - \sum_{j=1}^{J} \mathrm{SS}(E_j) \,/\, \sum_{j=1}^{J} \mathrm{SS}(Z_j). \tag{21}$$

This index ranges from *0* to *1*. The larger the FIT value, the larger the proportion of the variance in the endogenous variables explained the model.

Multilevel Dynamic GSCA employs the bootstrap method (Efron, 1982) in order to estimate the standard errors of parameter estimates. In this method, random samples (bootstrap samples) of Z_j are repeatedly sampled from the original data matrix with replacement. Multilevel Dynamic GSCA is applied to each bootstrap sample to obtain the estimates of parameters. Then, the bootstrapped standard errors of the estimates are calculated across entire bootstrap samples. The bootstrapped standard errors are used to assess the reliability of the estimates. The critical ratios (i.e., the parameter estimates divided by their standard errors) can be used to examine the significance of the parameter estimates (e.g.., a parameter estimate having a critical ratio greater than two in absolute value is considered significant at *.05* level).

3. Applications of Dynamic GSCA to Real Functional Neuroimaging Data

In this section, we applied two-level Dynamic GSCA to part of the functional MRI data from a Sternberg working memory task to demonstrate the effectiveness of the proposed approach. In the working memory study, thirty subjects (fifteen normal and fifteen patients with schizophrenia) performed two runs (*10 min, 53 sec*) of a variable load working memory task. During a single trial of this task, subjects saw a string of *2, 4, 6,* and *8* different uppercase consonants for *4 sec*, and were subsequently instructed to remember these consonants over a short delay (*6 sec*). A single consonant in lowercase was presented for *1* sec after this delay, and subjects were asked to indicate if this letter had been present in the preceding string. Functional image volumes were collected with a gradient echo (GRE) sequence (TR/TE *3000/40ms, 90º* flip angle, FOV *24×24 cm, 64×64* matrix, *62.5 kHz* bandwidth, *3.75×3.75 mm* in plane resolution, *5.00 mm* slice thickness, *29* slices, *145 mm* axial brain coverage). Statistical Parametric Mapping Software (SPM99) was used for image realignment, normalization into modified Talairch stereotaxic anatomical space, and smoothing with a Gaussian kernel to compensate for inter-subject anatomical differences and optimize the signal to noise ratio (for details of the study, see Cairo et al., 2006).

For brain connectivity analysis, the time series data-level measures of thirty five observed BOLD signals (i.e., five BOLD signals per ROI) were used, which were obtained from fifteen normal subjects. The seven ROIs used in the analysis were: Left Precentral Gyrus (PreCG_L), Right Precentral Gyrus (PreCG_R), Left Inferior Parietal Lobule (IPL_L), Right Inferior Parietal Lobule (IPL_R), Left Cerebellum (CL_L), Right Cerebellum (CL_R), Supplementary Motor Area (SMA).

The seven ROIs were computed using constrained principal component analysis (CPCA) for fMRI (fMRI-CPCA; Woodward et al., 2006). CPCA is a general method for structural analysis of multivariate data that combines regression analysis and principal component analysis into a unified framework (Takane, 2013; Takane and Hunter, 2001; Takane and Shibayama, 1991). fMRI-CPCA is a special application of CPCA to analyze fMRI data and derive functional images of functional neural networks from singular value decomposition of BOLD signal time series with the analyzed variance constrained to that which is predictable from stimulus timing and other constraints. For the current data, we first carried out a multivariate multiple regression analysis using a finite impulse response (FIR) model of the expected BOLD response to the timing of stimulus presentations as input *(G): $Z = GC + E$*, where Z is a matrix of the original data, C is a matrix of regression coefficients, and E is a matrix of residuals. Note that GC contains variability in Z that was predictable from the FIR model (G). Following this, we extracted the first component from GC using PCA, representing a major functional neural network consisting of several brain regions. The seven clusters were based on a cut-off of the highest *5%* of the first component loadings, and defined ROIs as having greater than *2500 mm^3* in cluster volume. Here, the ROIs represent the clusters surrounding the peak coordinates. The neural image of the clusters

consisting of the dominant *5%* of component loadings are displayed in Figure 2, with anatomical descriptions of these presented in Table 1. The number of voxels containing BOLD signal was *180* in SMA (Supplementary Motor Area), *157* in CL_L (Left Cerebellum), *144* in PreCG_R (Right Precentral Gyrus), *135* in IPL_R (Right Inferior Parietal Lobule), *68* in IPL_L (Left Inferior Parietal Lobule), *53* in PreCG_L (Left Precentral Gyrus), and *40* in CL_R (Right Cerebellum). There are simply too many BOLD signals to be processed for each ROI. Thus, they were aggregated into five distinct signals for each ROI, capturing representative patterns of BOLD signals within the ROI. We calculated the centroids of clusters, which were used as indicators of ROIs in the two-level Dynamic GSCA. For the cluster analysis, the *K*-means algorithm in MATLAB was used (*www.mathworks.com*). The choice of five clusters may sound rather arbitrary. However, we examined the effects of the number of clusters by systematically varying its value from *1* to *15*, and found that parameter estimates were virtually unchanged beyond five (i.e., correlations among the estimates of structural parameters were all higher than *.95*). We thus considered that five clusters were a good compromise, not too few to potentially miss important information and not too many to waste precious computational resources.

***** Insert Figure 2 *****

***** Insert Table1 ******

The BOLD time series records were nested within the subjects. The number of time points of BOLD signals was *214*. We specified a fully and bidirectionally connected structural model with time-lagged effects. Then, we applied two-level Dynamic GSCA to fit the specified model to the data. More specifically, the loadings for time series data-level observed BOLD signals in the voxels within the ROIs and the path coefficients of time series data-level latent variables (i.e., ROIs) were assumed to vary freely across the subjects. The hypothesized structural model with the seven ROIs is depicted in Figure 3.

***** Insert Figure 3 *****

The specified two-level model provided an overall goodness of fit (FIT) of *.812*, indicating that it accounted for about *81.2%* of the total variance of all endogenous variables. Table 2 provides weight (*w*'s) estimates for the observed variables in the model. It also presented the standard errors of the parameter estimates, calculated based on *100* bootstrapped samples. Note that C.R. denotes a critical ratio and significant parameter estimates at α*=.05* (i.e., C.R. greater than *2*) are indicated by an asterisk on estimates. It was shown that the weight estimates for each latent variable were similar to each other and all turned out to be statistically significant.

This indicates that all observed variables contributed equally well to determining their latent variables in the model.

***** Insert Table2 ******

Table 3 presents the fixed (*c*'s) loadings for the observed variables. The estimated loadings appeared high and were statistically significant. This suggests that the latent variables seemed to be well constructed such that they accounted for a large portion of the variances of the observed variables. Table 4 provides the fixed (*a*'s) path coefficient estimates and their standard errors. About *45%* of path coefficients turned out to be statistically significant. This denotes that the interpretations of the fixed effects do not appear to be consistent with the relationships among the latent variables hypothesized in the model. That is, we could have a more parsimonious model rather than a fully connected model with bidirectional relationships among ROIs.

***** Insert Table3 ******

***** Insert Table4 ******

Table 5 shows that the variance estimate of each level-2 random loading ($\lambda_j's$) across the subjects and its standard error obtained from bootstrap samples. Here, the random loading denotes the deviation between the fixed loading and individual subject's estimate of the loading. Thus, the variance estimate of each level-2 random loading represents inter-subject variability in the loading. It was found that the variances of the random loadings for all observed variables turned out to be statistically significant. This suggests that there existed substantial subject-wise differences in each of the loadings. Table 6 exhibits the variance estimate of the Level-2 random path coefficient ($\theta_j's$) of each latent variable and its standard error obtained from bootstrap samples. About *87%* of the variance estimates of the random path coefficients turned out to be statistically significant, suggesting substantial differences in each of the path coefficients across the subjects.

***** Insert Table5 ******

***** Insert Table6 ******

4. Summary and Discussion

In this paper, we proposed Multilevel Dynamic GSCA for the simultaneous analysis of multi-subject functional neuroimaging data. Multilevel Dynamic GSCA extends the single-level Dynamic GSCA to take into account the nested structure in functional neuroimaging data.

We demonstrated the effectiveness of Multilevel Dynamic GSCA by applying the method to the two-level Sternberg working memory task data, where time series data-level measurements were nested within different subjects. Explicitly taking the nested structure into account, the proposed method allowed investigating subject-wise variability of the loadings and path coefficients by looking at the variance estimates of the corresponding random effects, as well as fixed loadings between observed and latent variables and fixed path coefficients between latent variables. Substantial inter-subject variations of all loadings and path coefficients were revealed form these data.

Multilevel Dynamic GSCA may be extended in a variety of ways to enhance its data-analytic capability and applicability. First, two-level Dynamic GSCA can be extended to a higher-level (e.g., three-level) Dynamic GSCA. For example, if we have information on a higher-level unit (e.g., sites) that subjects belong to, observed data can be further split into between- and within- the higher unit samples data, each of which is separately modeled by Dynamic GSCA. Another possible extension of Multilevel Dynamic GSCA involves combining Multilevel Dynamic GSCA and cluster analysis to accommodate higher-level (e.g., groups) heterogeneity of subjects. In many instances, higher level heterogeneity can be present in hierarchically structured data. The information on higher-level heterogeneity can be obtained by identifying clusters of subjects through post-hoc approaches such as clustering analysis. Another possible extension of Multilevel Dynamic GSCA involves incorporating latent interactions. A latent interaction is defined as a product of interacting latent variables (Hwang et al., 2010). This extension may be particularly useful for brain connectivity analysis to capture important aspects of neuronal interactions by modeling changes in the magnitude of path coefficients between ROIs as a function of activities of different latent ROIs.

Appendix: The ALS algorithm for Two-Level Dynamic GSCA

The level-2 model parameters (D_C, A_l's,D_{λ_j}'s, Θ_{lj}'s, and D_W) in (10) and (13) are estimated using an alternation least squares (ALS) algorithm. Here, instead of updating the level-2 model parameters (D_C, A_l's,D_{λ_j}'s, Θ_{lj}'s) separately, we take a much simpler way to obtain the level-2 model parameters by updating first the level-1 model parameters D'_{c_j}'s and A_{lj}'s in (8) and (11), and then split them into D_C and A_l's, and D_{λ_j}'s and Θ_{lj}'s, respectively, after convergence in the way that $D'_C = \sum_{j=1}^{J} D'_{Cj}/J$, $D'_{\lambda_j} = D'_{Cj} - D'_C$, $A'_l = \sum_{j=1}^{J} A'_{lj}/J$, and $\Theta'_{lj} = A'_{lj} - A'_l$. The proposed ALS algorithm consists of two major steps:

Step I. Update D'_{c_j}'s and A_{lj}'s while D_W is fixed. The error terms for the measurement model (8) can be rearranged as

$$\sum_{j=1}^{J} E_{Mj} = \sum_{j=1}^{J} \left[Z_j - Z_j D_W D'_{c_j} \right]. \tag{22}$$

Let $D_{Z_j} = bdiag\left(\left[Z_{1j},\ Z_{1j}, \ldots, Z_{Pj}\right]\right)$ *and* $c_j = D'_{c_j} 1_P$. Criterion (22) can then be written as

$$e_M = z - D_{X_M} c, \tag{23}$$

where

$$e_M = \text{vec}([E_{M1}, \ldots, E_{MJ}]), \tag{24}$$

$$z = \text{vec}([Z_1, \ldots, Z_J]), \tag{25}$$

$$D_{X_M} = \text{bdiag}\begin{pmatrix} I_V \otimes D_{Z_1} w \\ \vdots \\ I_V \otimes D_{Z_J} w \end{pmatrix}. \tag{26}$$

$$c = \text{vec}(c_1, \ldots, c_J) \tag{27}$$

Next, the error term for the structural model (11), on the other hand, can be rearranged as

$$\sum_{j=1}^{J} E_{Sj} = \sum_{j=1}^{J} \left[\Gamma_j - \sum_{l=0}^{q} (S_l \Gamma_j A'_{lj}) \right]. \tag{28}$$

Criterion (28) can be written as

$$e_S = \gamma - X_S h_1^*, \tag{29}$$

where

$$e_S = \text{vec}([E_{S1}, \ldots, E_{SJ}]), \tag{30}$$

$$\gamma = \begin{pmatrix} \text{vec}([\Gamma_1, \ldots, \Gamma_1]) \\ \vdots \\ \text{vec}([\Gamma_J, \ldots, \Gamma_J]) \end{pmatrix}, \tag{31}$$

$$X_S = \begin{pmatrix} \left[I_P \otimes S_0 \Gamma_1, \ldots, I_P \otimes S_q \Gamma_1 \right] \\ \vdots \\ \left[I_P \otimes S_0 \Gamma_J, \ldots, I_P \otimes S_q \Gamma_J \right] \end{pmatrix}, \tag{32}$$

$$h_1^* = \begin{pmatrix} \mathrm{vec}\left(\left[\mathrm{A}'_{01}, \ldots, \mathrm{A}'_{q1}\right]\right) \\ \vdots \\ \mathrm{vec}\left(\left[\mathrm{A}'_{0J}, \ldots, \mathrm{A}'_{qJ}\right]\right) \end{pmatrix}. \tag{33}$$

The vector h_1^* has many 0 elements. Let $\tilde{h}_1$ denote the vector formed from h_1^* by eliminating all of its 0 elements, and let $\tilde{X}_S$ denote the matrix form X_S by eliminating all the corresponding columns of X_S. Then (35) and (39) can be collectively written as

$$e = f_1 - X_1 h_1, \tag{34}$$

where

$$f_1 = \begin{pmatrix} z \\ \gamma \end{pmatrix}, \tag{35}$$

$$X_1 = \mathrm{bdiag}\left(D_{X_M}, \tilde{X}_S\right), \tag{36}$$

and

$$h_1 = \begin{pmatrix} c \\ \tilde{h}_1 \end{pmatrix}. \tag{37}$$

Thus, the update of h_1 comprising D'_{c_j}'s and A_{lj}'s that minimizes $e'e$ is found by

$$\hat{h}_1 = (X_1{}'X_1)^{-1}X_1{}'f_1. \tag{38}$$

Note that in (38) we assume that X_1 has full column rank, which ensures uniqueness of $\hat{h}_1$.

Step II. Update D_W while D'_{c_j}'s and A_{lj}'s are fixed. Let $w = D_W 1_P$. The e_M in (23) can be rewritten as

$$e_M = z - D_{ZC} w, \tag{39}$$

where

$$D_{ZC} = \begin{pmatrix} D_{c_1} \otimes D_{Z_1} \\ \vdots \\ D_{c_j} \otimes D_{Z_J} \end{pmatrix}. \tag{40}$$

Next, the error term for the structural model (11), on the other hand, can be rearranged as

$$\sum_{j=1}^{J} E_{Sj} = \sum_{j=1}^{J}\left[\Gamma_j - \sum_{l=0}^{q}(S_l \Gamma_j \mathrm{A}'_{lj})\right]$$

$$= \sum_{j=1}^{J}\left[Z_j D_W - \sum_{l=0}^{q}(S_l Z_j D_W \mathrm{A}'_{lj})\right]$$

$$= \sum_{j=1}^{J}\left\{I - \sum_{l=0}^{q}(A'_{lj} \otimes S_l)\right\} D_{Z_j} w. \quad (41)$$

Criterion (41) can be written as

$$e_S = 0 - X_w w, \quad (42)$$

where

$$X_w = \begin{pmatrix} \left\{I - \sum_{l=0}^{q}(A'_{l1} \otimes S_l)\right\} D_{Z_1} \\ \vdots \\ \left\{I - \sum_{l=0}^{q}(A'_{lj} \otimes S_l)\right\} D_{Z_J} \end{pmatrix}. \quad (43)$$

Then (39) and (42) can be collectively written as

$$e = f_2 - X_2 w, \quad (44)$$

where

$$f_2 = \begin{pmatrix} z \\ 0 \end{pmatrix}, \quad (45)$$

$$X_2 = \begin{pmatrix} D_{ZC} \\ X_w \end{pmatrix}, \quad (46)$$

In this step, $e'e$ should be minimized with respect to w_i separately and sequentially subject to the normalization restriction presented in (5) (page. 3) for each i, where latent variables are scaled to have unit variance. For more details about the sequential updating of w_i, see Jung et al. (2012).

The two steps described above are repeatedly applied until the change in $e'e$ from one iteration to the next gets smaller than a certain prescribed value, say *1.0E-6*. After convergence the model parameters (D_C, D_{λ_j}'s, A_l's,Θ_{lj}'s) are obtained by splitting D_{λ_j}'s, and Θ_{lj}'s into D_C and A_l's, and D_{λ_j}'s and Θ_{lj}'s, respectively, in the way that $D'_C = \sum_{j=1}^{J} D'_{Cj}/J$, $D'_{\lambda_j} = D'_{Cj} - D'_C$, $A'_l = \sum_{j=1}^{J} A'_{lj}/J$, and $\Theta'_{lj} = A'_{lj} - A'_l$.

The overall goodness of fit of a hypothesized model is given by

$$\mathrm{FIT} = 1 - SS(f_2 - X_2\hat{w})/SS(f_2), \quad (48)$$

where $\hat{w}$ is the estimate of w at a convergence point of the ALS algorithm.

References

Cairo, T. A., Woodward, T. S., & Ngan, E. T. (2006). Decreased encoding efficiency in schizophrenia. *Biological Psychiatry, 59,* 740-746.

deLeeuw, J., Young, F.W., & Takane, Y. (1976). Additive structure in qualitative data: an alternating least squares method with optimal scaling features. *Psychometrika, 41,* 471–503.

Efron, B. (1982). *The jackknife, the bootstrap and other resampling plans*. Philadelphia: SIAM.

Friston, K. J. (1994). Functional and effective connectivity in neuroimaging: a synthesis. *Human Brain Mapping, 2,* 56-78.

Greve, D. N., Brown, G. G., Mueller, B. A., Glover, G., & Liu, T. T. (2013). A survey of the sources of noise in fMRI. *Psychometrika, 78,* 396-416.

Hwang, H., & Takane, Y. (2004). Generalized structured component analysis. *Psychometrika, 69,* 81–99.

Hwang, H. & Takane, Y. (2014). *Generalized structured component analysis: A component-based approach to structural equation modeling*, Chapman & Hall/CRC Press. (forthcoming)

Hwang, H., Takane, Y., & Malhotra, N.K. (2007). Multilevel generalized structured component analysis. *Behaviormetrika, 34,* 95–109.

Hwang, H., Ho, R. M., & Lee, J. (2010). Generalized structured component analysis with latent interactions. *Psychometrika, 75,* 228-242.

Jung, K., Takane, Y., Hwang, H., & Woodward, T. S. (2012). Dynamic GSCA (Generalized Structured Component Analysis) with applications to the analysis of effective connectivity in functional neuroimaging data. *Psychometrika, 77,* 827-848.

Snijders, T., & Bosker, R. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*, London: Sage Publication.

Takane, Y. (2013). *Constrained principal component analysis and related techniques*. Boca Raton, FL: Chapman and Hall/CRC Press.

Takane, Y., & Hunter, M. A. (2001).Constrained principal component analysis: a comprehensive theory. *Applicable Algebra in Engineering, Communication and Computing, 12,* 391-419.

Takane, Y., & Shibayama, T. (1991). Principal component analysis with external information on both subjects and variables. *Psychometrika, 56,* 97-120.

Van Den Heuvel, M. P., & Hulshoff Pol, H. E. (2010). Exploring the brain network: a review on resting-state fMRI functional connectivity. *European Neuropsychopharmacology, 20,* 519-534.

Woodward, T. S., Cairo, T. A., Ruff, C. C., Takane, Y., Hunter, M. A. & Ngan, E. T. C. (2006). Functional connectivity reveals load dependent neural systems underlying encoding and maintenance in verbal working memory. *Neuroscience, 139*, 317-325.

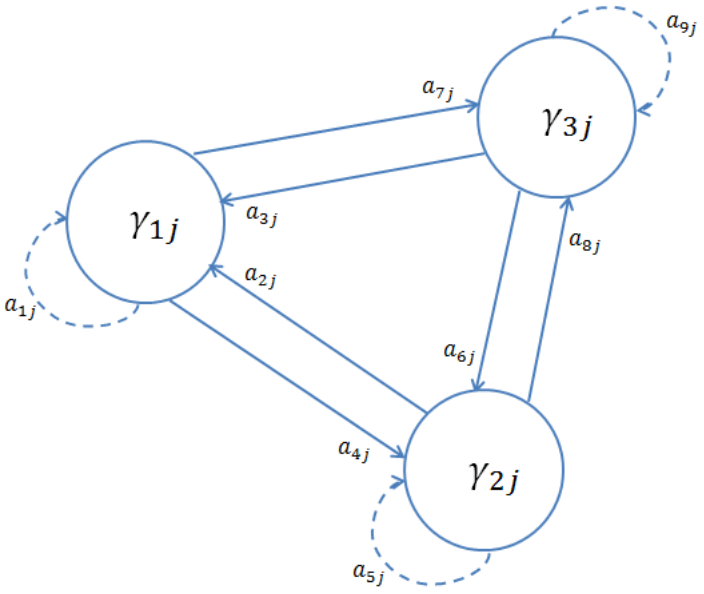

Figure 1. A hypothesized Level-1 structural model among three ROIs for the *j*-th subject

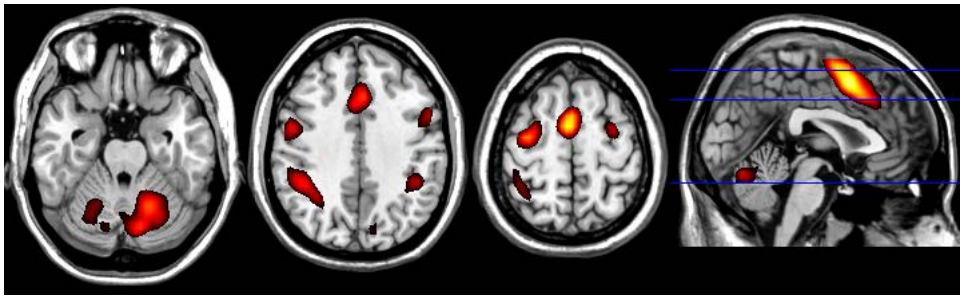

Figure 2. The neural image of the clusters consisting of the dominant *5%* of component loadings obtained from fMRI-CPCA (slice numbers: 48,110,132)

.

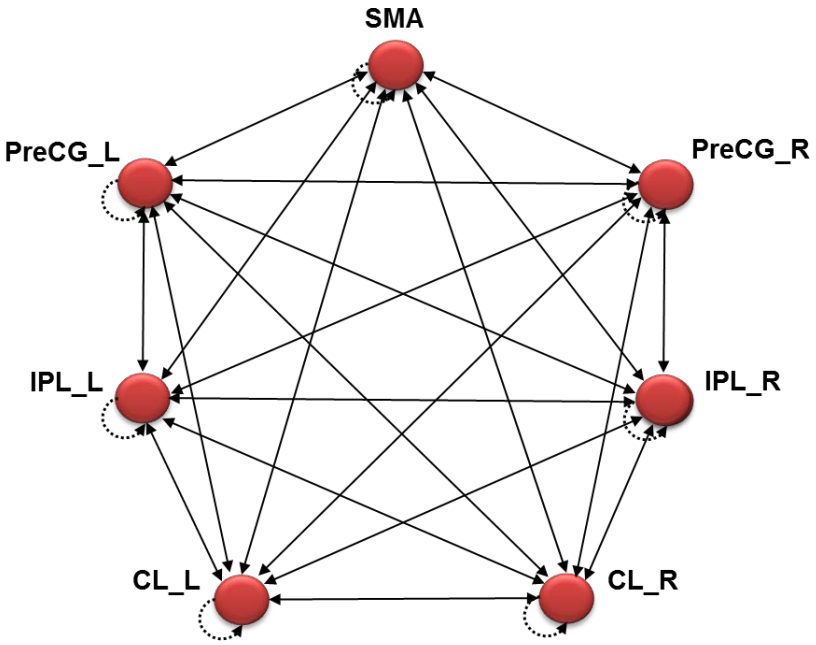

Fig 3: A Hypothesized Structural Model

Note. PreCG_L=Left Precentral Gyrus; PreCG_R = Right Precentral Gyrus; IPL_L = Left Inferior Parietal Lobule; IPL_R = Right Inferior Parietal Lobule; CL_L = Left Cerebellum; CL_R= Right Cerebellum; SMA = Supplementary Motor Area

Table 1. Descriptions of ROI clusters: the number of voxels within clusters and cluster volumes for most extreme *5%* of the first component obtained from fMRI-CPCA, and Montreal Neurological Institute (MNI) coordinates for peak location within each cluster.

ROIs	Voxels	Cluster Volume (mm^3)	MNI coordinate (x, y, z) for peak locations
SMA	180	11520	4 8 52
CL_L	157	10048	-28 -64 -28
PreCG_R	144	9216	48 4 28
IPL_R	135	8640	48 -40 48
IPL_L	68	4352	-44 -40 40
PreCG_L	53	3392	-52 8 32
CL_R	40	2560	32 -60 -32

Note. SMA = Supplementary Motor Area; CL_L = Left Cerebellum; PreCG_R = Right Precentral Gyrus; IPL_R = Right Inferior Parietal Lobule; IPL_L = Left Inferior Parietal Lobule; PreCG_L=Left Precentral Gyrus; CL_R= Right Cerebellum

Table 2. The weight estimates and their standard errors.

ROIs	*w*'s	Estimate	S.E.	C.R.
IPL_R	1	0.231*	0.014	16.31
	2	0.213*	0.016	13.29
	3	0.198*	0.018	10.88
	4	0.234*	0.021	11.35
	5	0.206*	0.012	17.36
PreCG_R	1	0.226*	0.018	12.28
	2	0.229*	0.018	12.41
	3	0.222*	0.035	6.37
	4	0.251*	0.025	9.93
	5	0.206*	0.018	11.31
CL_R	1	0.193*	0.013	14.96
	2	0.225*	0.012	19.10
	3	0.261*	0.014	18.56
	4	0.203*	0.012	16.83
	5	0.197*	0.011	17.94
CL_L	1	0.221*	0.020	10.79
	2	0.243*	0.014	17.95
	3	0.186*	0.014	12.91
	4	0.210*	0.022	9.44
	5	0.237*	0.018	12.85
IPL_L	1	0.206*	0.013	15.76
	2	0.222*	0.021	10.68
	3	0.211*	0.017	12.53
	4	0.243*	0.018	13.50
	5	0.210*	0.012	17.96
PreCG_L	1	0.230*	0.015	14.94
	2	0.244*	0.014	17.25
	3	0.209*	0.010	21.02
	4	0.202*	0.016	12.26
	5	0.232*	0.013	18.18
SMA	1	0.190*	0.016	11.54
	2	0.217*	0.012	18.87
	3	0.244*	0.018	13.41
	4	0.199*	0.010	20.15
	5	0.245*	0.013	19.54

Table 3. The Level-2 fixed loading estimates and their standard errors

ROIs	*c*'s	Estimate	S.E.	C.R.
IPL_R	1	0.935*	0.013	73.43
	2	0.923*	0.018	50.27
	3	0.911*	0.018	50.67
	4	0.935*	0.011	85.21
	5	0.914*	0.015	60.69
PreCG_R	1	0.883*	0.025	34.93
	2	0.894*	0.021	42.95
	3	0.858*	0.052	16.45
	4	0.901*	0.014	66.53
	5	0.868*	0.031	28.38
CL_R	1	0.931*	0.009	108.58
	2	0.934*	0.009	100.80
	3	0.938*	0.011	87.68
	4	0.924*	0.018	51.63
	5	0.910*	0.018	50.29
CL_L	1	0.904*	0.021	42.19
	2	0.920*	0.018	50.83
	3	0.915*	0.017	52.51
	4	0.903*	0.019	47.92
	5	0.916*	0.009	102.22
IPL_L	1	0.911*	0.016	57.37
	2	0.927*	0.024	38.80
	3	0.929*	0.008	114.66
	4	0.915*	0.013	72.83
	5	0.898*	0.022	40.18
PreCG_L	1	0.902*	0.017	51.66
	2	0.890*	0.018	48.77
	3	0.908*	0.016	55.79
	4	0.885*	0.028	31.27
	5	0.896*	0.016	54.71
SMA	1	0.882*	0.021	41.08
	2	0.932*	0.010	97.56
	3	0.928*	0.016	56.51
	4	0.901*	0.025	36.08
	5	0.917*	0.016	58.59

Table 4. The Level-2 fixed path coefficient estimates and their standard errors

	a's		Estimate	S.E.	C.R.
IPL_R	←	IPL_R	-0.010	0.023	0.44
		PreCG_R	0.352*	0.039	9.11
		CL_R	0.067	0.078	0.85
		CL_L	0.123	0.067	1.84
		IPL_L	0.284*	0.052	5.49
		PreCG_L	-0.028	0.054	0.52
		SMA	0.202*	0.044	4.60
PreCG_R	←	IPL_R	0.243*	0.043	5.61
		PreCG_R	0.022	0.021	1.06
		CL_R	-0.034	0.043	0.78
		CL_L	0.063	0.067	0.95
		IPL_L	0.061	0.037	1.66
		PreCG_L	0.153*	0.065	2.34
		SMA	0.493*	0.054	9.10
CL_R	←	IPL_R	0.019	0.102	0.19
		PreCG_R	0.028	0.087	0.32
		CL_R	0.055	0.031	1.78
		CL_L	0.721*	0.065	11.02
		IPL_L	0.154*	0.061	2.54
		PreCG_L	-0.045	0.042	1.07
		SMA	0.061	0.077	0.80
CL_L	←	IPL_R	0.142*	0.058	2.47
		PreCG_R	0.029	0.094	0.31
		CL_R	0.654*	0.052	12.52
		CL_L	0.136*	0.025	5.49
		IPL_L	-0.023	0.046	0.51
		PreCG_L	0.118*	0.037	3.16
		SMA	-0.019	0.070	0.27
IPL_L	←	IPL_R	0.409*	0.071	5.79
		PreCG_R	0.106	0.069	1.55
		CL_R	0.219*	0.070	3.12
		CL_L	-0.085	0.051	1.67
		IPL_L	0.104*	0.027	3.91
		PreCG_L	0.350*	0.059	5.92
		SMA	-0.117	0.092	1.27

(Continued)

	a's		Estimate	S.E.	C.R.
PreCG_L	←	IPL_R	0.026	0.104	0.25
		PreCG_R	0.334*	0.144	2.31
		CL_R	-0.116	0.072	1.60
		CL_L	0.120	0.086	1.39
		IPL_L	0.371*	0.068	5.48
		PreCG_L	0.054	0.029	1.86
		SMA	0.182	0.104	1.74
SMA	←	IPL_R	0.171*	0.043	3.95
		PreCG_R	0.645*	0.075	8.61
		CL_R	0.141	0.075	1.88
		CL_L	-0.114	0.066	1.72
		IPL_L	-0.077	0.068	1.13
		PreCG_L	0.135*	0.064	2.11
		SMA	0.074*	0.021	3.62

Table 5. The variance estimates of the Level-2 random loading estimates and their standard errors

The variance of $\lambda_j's$		Estimate	S.E.	C.R.
IPL_R	1	0.001*	0.000	3932.79
	2	0.002*	0.001	1317.46
	3	0.002*	0.001	1192.64
	4	0.001*	0.000	2901.17
	5	0.002*	0.001	997.75
PreCG_R	1	0.008*	0.002	389.32
	2	0.003*	0.001	1084.31
	3	0.018*	0.016	52.23
	4	0.003*	0.001	847.23
	5	0.010*	0.003	285.46
CL_R	1	0.001*	0.000	2600.30
	2	0.002*	0.001	1145.35
	3	0.000*	0.000	4369.05
	4	0.001*	0.001	1541.32
	5	0.002*	0.001	685.63
CL_L	1	0.003*	0.002	451.20
	2	0.001*	0.001	979.20
	3	0.004*	0.001	720.98
	4	0.001*	0.001	1541.05
	5	0.001*	0.000	2994.64
IPL_L	1	0.003*	0.001	882.81
	2	0.004*	0.003	360.93
	3	0.001*	0.000	3868.73
	4	0.001*	0.000	2582.41
	5	0.004*	0.002	559.30
PreCG_L	1	0.002*	0.001	1593.66
	2	0.002*	0.001	1250.32
	3	0.006*	0.001	661.50
	4	0.006*	0.005	176.37
	5	0.003*	0.001	1052.30
SMA	1	0.003*	0.002	456.22
	2	0.003*	0.001	1111.56
	3	0.003*	0.002	563.89
	4	0.007*	0.002	375.74
	5	0.001*	0.001	1481.44

Table 6. The variance estimates of the Level-2 random path coefficient estimates and their standard errors

The variance of $\theta_{lj}'s$			Estimate	S.E.	C.R.
IPL_R	←	IPL_R	0.009*	0.004	2.37
		PreCG_R	0.029*	0.013	27.09
		CL_R	0.105*	0.030	2.24
		CL_L	0.082*	0.034	3.57
		IPL_L	0.043*	0.017	16.86
		PreCG_L	0.043*	0.011	2.47
		SMA	0.030*	0.011	18.78
PreCG_R	←	IPL_R	0.025*	0.009	28.38
		PreCG_R	0.006*	0.003	8.66
		CL_R	0.028*	0.008	4.50
		CL_L	0.075*	0.030	2.13
		IPL_L	0.028*	0.007	8.38
		PreCG_L	0.058*	0.017	9.06
		SMA	0.047*	0.016	31.13
CL_R	←	IPL_R	0.150	0.050	0.38
		PreCG_R	0.100	0.045	0.62
		CL_R	0.013*	0.003	16.37
		CL_L	0.049*	0.019	37.85
		IPL_L	0.050*	0.015	10.12
		PreCG_L	0.030*	0.010	4.46
		SMA	0.073*	0.025	2.46
CL_L	←	IPL_R	0.044*	0.014	9.82
		PreCG_R	0.096	0.077	0.37
		CL_R	0.034*	0.013	50.15
		CL_L	0.009*	0.003	52.98
		IPL_L	0.035	0.014	1.69
		PreCG_L	0.020*	0.006	20.77
		SMA	0.052	0.024	0.78
IPL_L	←	IPL_R	0.091*	0.033	12.52
		PreCG_R	0.089*	0.024	4.52
		CL_R	0.072*	0.026	8.50
		CL_L	0.052*	0.016	5.19
		IPL_L	0.012*	0.003	32.19
		PreCG_L	0.053*	0.017	20.95
		SMA	0.129*	0.038	3.08

(Continued)

The variance of $\theta_{lj}'s$			Estimate	S.E.	C.R.
PreCG_L	←	IPL_R	0.154	0.046	0.55
		PreCG_R	0.304*	0.095	3.51
		CL_R	0.066*	0.025	4.72
		CL_L	0.097*	0.051	2.37
		IPL_L	0.053*	0.013	27.74
		PreCG_L	0.012*	0.006	8.60
		SMA	0.156*	0.058	3.16
SMA	←	IPL_R	0.025*	0.013	13.44
		PreCG_R	0.080*	0.023	28.19
		CL_R	0.081*	0.030	4.66
		CL_L	0.068*	0.029	3.90
		IPL_L	0.081*	0.022	3.43
		PreCG_L	0.054*	0.016	8.47
		SMA	0.008*	0.002	35.21