

Dynamic GSCANO (Generalized Structured Canonical Correlation Analysis) with applications to the analysis of effective connectivity in functional neuroimaging data

Lixing Zhou, Yoshio Takane, and Heungsun Hwang

McGill University

February 29, 2016

The authors are grateful to Todd Woodward for providing his data. Requests for reprints should be sent to: Lixing Zhou, Department of Psychology, McGill University, 1205 Dr. Penfield Avenue, Montreal, QC, H3A 1B1, Canada. Email: lixing.zhou@mail.mcgill.ca

Dynamic GSCANO (Generalized Structured Canonical Correlation Analysis) with applications to the analysis of effective connectivity in functional neuroimaging data

Abstract

Effective connectivity in functional neuroimaging studies is defined as the time dependent causal influence that a certain brain region of interest (ROI) exerts on another. A new method of structural equation modeling (SEM) is proposed for analyzing common patterns among multiple subjects' effective connectivity. The proposed method, called Dynamic GSCANO (Generalized Structured Canonical Correlation Analysis) incorporates contemporaneous and lagged effects between ROIs, direct and modulating effects of stimuli, as well as interaction effects among ROIs. An alternating least squares (ALS) algorithm is developed for estimating parameters. Synthetic and real data are analyzed to demonstrate the feasibility and usefulness of the proposed method.

Key words: dynamic generalized structured component analysis (Dynamic GSCA), structural equation modeling (SEM), longitudinal and time series data, alternating least squares (ALS) algorithm, fMRI (functional magnetic resonance imaging) data, brain connectivity

1. Introduction

Functional magnetic resonance imaging (fMRI) has become an important measurement tool in psychological research (e.g., Cacioppo & Decety, 2009; Breiter et al., 2001; Huettel et al., 2006). The basic element of spatial measurement in fMRI is a voxel (i.e., a three dimensional volume element), from which time series of blood oxygenation level dependent (BOLD) signal are recorded. Each brain region of interest (ROI) consists of multiple voxels and thus results in matrix records of BOLD signals, where rows represent time points and column represent voxels. Effective connectivity analysis involves a class of fMRI analysis methods that quantify the influence of one ROI on another (Friston, 1994). In a typical effective connectivity research, a hypothesized connectivity pattern among selected ROIs is expressed as a path diagram, and then a psychometric method is used to analyze the path-analytic relationships. Structural equation modeling is a helpful psychometric methodology in analyzing effective connectivity because of its strength in specifying and testing path-analytic relationships (McIntosh & Gonzalez-Lima, 1994; Büchel & Friston, 1997; Bullmore et al., 2000; Penny et al., 2004).

Recently, two SEMs have been proposed to analyze effective connectivity. They are extended unified SEM (euSEM; Gates, Molenaar, Hillary, & Slobounov, 2011) and Dynamic Generalized Structured Component Analysis (Dynamic GSCA; Jung, Takane, Hwang, & Woodward, 2012). Both approaches incorporate contemporaneous and lagged effects of ROIs, and direct and modulating effects of stimuli. The euSEM obtains parameter estimates under the Gaussian distributional assumptions (Gates et al., 2011), whereas Dynamic GSCA is distribution-free.

Both euSEM and Dynamic GSCA are capable of analyzing only a single subject's effective connectivity at a time. However, it may be more useful to have SEMs that enable simultaneous analysis of multiple subjects and capture common patterns of effective connectivity across subjects. Unsynchronized measurement errors due to various artifacts tend to be cancelled out in simultaneous analysis of multiple-subject data. Examples of capturing population features of effective connectivity are abundant in empirical research (e.g., Luchtman et al., 2012; Stoeckel et al., 2009). Moreover, it has been emphasized by researchers that treatments ignoring simultaneous analysis, such as simply applying a single subject based SEM to BOLD signals of one subject at a time and then representing the common effective connectivity by averaging parameter estimates obtained over multiple subjects, are likely to result in misleading inferences (Eklund et al., 2012; Mecheli et al., 2002).

Despite the prevalence of multiple-subject research in effective connectivity research and the detrimental effects of lacking a systematic approach to simultaneous analysis to statistical inferences, there has been little attempt to develop SEMs to analyze multiple-subject data simultaneously. In this paper, we fill this gap by proposing a new method for simultaneous analysis of multiple subjects' connectivity. The proposed method, named *Dynamic Generalized Structured Canonical Correlation Analysis* (Dynamic GSCANNO), combines generalized canonical correlation analysis (GCANO; Carroll, 1968; Gifi, 1991) with a multivariate autoregressive time series model in a unified framework. GCANO is a natural extension of two-set canonical correlation analysis (Hotelling, 1936) and is able to construct a latent vector which is most representative of columns from multiple data matrices. The conventional GCANO, however, does not model relationships among the extracted components. Dynamic GSCANNO, on the other hand, explicitly models their relationships by postulating an autoregressive time series

regression relationship among them. The autoregressive model subsumes all kinds of time series effects among ROIs and stimuli included in euSEM and Dynamic GSCA, as well as one extra term, the interaction effects between ROIs (Kenny & Judd, 1984; Hwang et al., 2010). In sum, Dynamic GSCANO is an extension of GCANO to the analysis of multiple subjects' time series data.

This paper is organized as follows. In Section 2, we discuss Dynamic GSCANO in detail, including the model (Section 2.1), the parameter estimation algorithm (Section 2.2), and other computational issues (Section 2.3). In Section 3, we conduct simulation studies to investigate the parameter recovery of Dynamic GSCANO. In Section 4, we demonstrate the feasibility and empirical usefulness of the method by applying it to two real data sets. In Section 5, we summarize previous sections and discuss further prospects for Dynamic GSCANO.

2. Dynamic GSCANO

2.1. The Model

Dynamic GSCANO consists of two sub-models: measurement and structural models. For each ROI, the measurement model employs GCANO to relate a common latent variable to multiple subjects' BOLD signal records in the ROI. The latent variable captures the common features of BOLD signals in the ROI across all subjects. The structural model specifies hypothesized intra-dynamics among latent variables for ROIs.

We begin with the measurement model. We assume that there are J ROIs and K subjects. Let \mathbf{Z}_{ki} denote a T by v_i matrix of BOLD signals taken over the i -th ROI on the k -th subject, where v_i indicates the number of voxels in the i -th ROI and T denotes the total number of time points.

The matrix \mathbf{Z}_{ki} is column-wise standardized *a priori*. Following GCANO, a T -element latent variable γ_i is constructed in terms of the following K multiple regression like equations,

$$\gamma_i = \mathbf{Z}_{ki} \mathbf{w}_{ki} + \mathbf{e}_{ki}^{(M)}, \quad k = 1, \dots, K, \quad (1)$$

where \mathbf{w}_{ki} indicates a v_i -element vector of component weights, and $\mathbf{e}_{ki}^{(M)}$ represents a T -element vector of measurement errors for ROI i and subject k . The parenthesized superscript M indicates that $\mathbf{e}_{ki}^{(M)}$ is an error vector in the measurement model. Equation (1) suggests that γ_i captures the common pattern of BOLD signal dynamics across all subjects and voxels in ROI i . Examples of γ_i are shown in Figures 2 and 3. The difference from the usual regression equations is that in (1) not only \mathbf{w}_{ki} but also γ_i are unknown, and the latter is estimated so as to minimize the error over k subject to the standardization constraint, $\frac{1}{T} \gamma_i' \gamma_i = 1$. By concatenating γ_i side by side, we obtain a matrix of latent variables,

$$\mathbf{\Gamma} = [\gamma_1, \gamma_2, \dots, \gamma_J]. \quad (2)$$

We now present the structural model. Let \mathbf{u}_p be a T -element vector, representing the time series records of the p -th stimulus, which is also *a priori* standardized. Assuming that there is a total of P stimulus inputs ($p = 1, \dots, P$), we may express them collectively in matrix notation as follows:

$$\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_P]. \quad (3)$$

Note that in equation (3), P denotes the number of distinct types of stimuli (e.g., the first stimulus is a static image, the second with motion, and the third with attention priming) rather than the number of different onsets of the same stimulus over time (Jung et al. 2012). An example of a stimulus effect is depicted in Figure 1. We use the time series depicted in the bottom panel of the figure as the stimulus input variable, which is obtained by convoluting the raw signals with the hemodynamic response function. This figure contains one stimulus time series ($P = 1$) and onsets of this stimulus at three different occasions.

To place both contemporaneous and lagged effects in a unified framework, a series of T by T shift matrices $\{\mathbf{S}_l\}$ ($l = 0, 1, \dots, L$) are introduced, with the subscript l indicating the order of time lags. Specifically, \mathbf{S}_0 , the shift matrix with lag 0, is defined to be \mathbf{I}_T , the identity matrix of order T . The matrix \mathbf{S}_0 denotes the contemporaneous effects. Matrices \mathbf{S}_l 's with nonzero orders are used to obtain the lag l effects among ROIs. The matrix \mathbf{S}_1 is defined as

$$\mathbf{S}_1 = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}, \quad (4)$$

and in general, $\mathbf{S}_l = \left[s_{ij}^{(l)} \right]$ with l greater than or equal to 2 ($l \geq 2$) is defined as

$$s_{ij}^{(l)} = \begin{cases} 1 & \text{if } j = i - l, \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

Pre-multiplying Γ by \mathbf{S}_l shifts down the rows of Γ by l rows and defines the matrix of the effects of latent variables at time $t - l$ on latent variables at time t . The generic structural model of Dynamic GSCANO is now stated as

$$\Gamma = \sum_{l=0}^L \mathbf{S}_l \Gamma \mathbf{B}_l + \sum_{l=0}^L \mathbf{S}_l \mathbf{U} \mathbf{D}_l + \sum_{l=0}^L \sum_{p=1}^P \mathbf{S}_l \text{diag}(\mathbf{u}_p) \Gamma \mathbf{M}_l^{(p)} + \sum_{l=0}^L \sum_{j=1}^J \mathbf{S}_l [\text{diag}(\boldsymbol{\gamma}_j) \Gamma] \mathbf{Q}_l^{(j)} + \mathbf{E}^{(S)}, \quad (6)$$

where $\text{diag}(\mathbf{u}_p)$ denotes a diagonal matrix with elements of \mathbf{u}_p as the diagonal entries, \mathbf{B}_l is a J by J coefficient matrix for the direct effects between ROIs, \mathbf{D}_l is a P by J coefficient matrix for the direct effects of stimuli on ROIs, $\mathbf{M}_l^{(p)}$ is a J by J matrix of coefficients for the modulating effects of stimuli on connections between ROIs (interactions between stimuli and ROIs), and $\mathbf{Q}_l^{(i)}$'s are J by J matrices of coefficients representing the modulating effects of ROIs on connections between ROIs (interactions between ROIs). All of these effects include both contemporaneous ($l=0$) and time lagged ($l>0$) effects. The matrix $\mathbf{E}^{(S)}$ is a T by J matrix of error terms in the structural model, and L , the upper bound for l , indicates the maximum possible lags.

2.2. Parameter Estimation

We first define criteria for parameter estimation and then present the estimation algorithm.

The criterion for the measurement model is given by

$$\phi^{(M)} = \sum_{i=1}^J \sum_{k=1}^K \|\boldsymbol{\gamma}_i - \mathbf{Z}_{ki} \mathbf{w}_{ki}\|^2, \quad (7)$$

which is, as mentioned above, minimized with respect to \mathbf{w}_{ki} and γ_i subject to the standardization constraint $\frac{1}{T} \boldsymbol{\gamma}'_i \boldsymbol{\gamma}_i = 1$. Here $\|\mathbf{A}\|^2 = \text{tr}(\mathbf{A}'\mathbf{A})$ for an arbitrary matrix \mathbf{A} . The minimization criterion for the structural model is given by

$$\phi^{(S)} = \|\mathbf{E}^{(S)}\|^2, \quad (8)$$

where $\mathbf{E}^{(S)} = \boldsymbol{\Gamma} - \sum_{l=0}^L \mathbf{S}_l \boldsymbol{\Gamma} \mathbf{B}_l - \sum_{l=0}^L \mathbf{S}_l \mathbf{U} \mathbf{D}_l - \sum_{l=0}^L \sum_{p=1}^P \mathbf{S}_l \text{diag}(\mathbf{u}_p) \boldsymbol{\Gamma} \mathbf{M}_l^{(p)} - \sum_{l=0}^L \sum_{i=1}^J \mathbf{S}_l [\text{diag}(\boldsymbol{\gamma}_i) \boldsymbol{\Gamma}] \mathbf{Q}_l^{(i)}$. The

global minimization criterion is now stated as a weighted sum of the two criteria, $\phi^{(M)}$ and $\phi^{(S)}$, as follows:

$$\phi = \alpha \phi^{(M)} + (1 - \alpha) \phi^{(S)}, \quad (9)$$

where α ($0 \leq \alpha \leq 1$) regulates the importance of the two sub-models in Dynamic GSCANO. A larger value of α implies that the fit of the measurement model is more important than that of the structural model, while a smaller value of α indicates the opposite. We normally set $\alpha = 0.5$ (Hwang, et al., 2013).

We develop an alternating least squares (ALS; De Leeuw et al., 1976) algorithm to minimize (9). We outline three essential steps of the ALS algorithm below and leave the detailed updating formulas in the technical appendix. The ALS algorithm repeats the following three steps until convergence.

Step I. Update \mathbf{w}_{ki} to minimize (9) with \mathbf{B}_l 's, \mathbf{D}_l 's, $\mathbf{M}_l^{(p)}$'s, $\mathbf{Q}_l^{(j)}$'s, $\boldsymbol{\Gamma}$, and \mathbf{Z}_{ki} 's assumed temporarily fixed.

Step II. Update \mathbf{B}_i 's, \mathbf{D}_i 's, $\mathbf{M}_i^{(p)}$'s, and $\mathbf{Q}_i^{(j)}$'s minimize (9) with Γ , \mathbf{U} , and \mathbf{Z}_{ki} 's assumed temporarily fixed.

Step III. Update γ_i ($i=1, \dots, J$) sequentially to minimize (9) subject to the standardization constraint with \mathbf{B}_i 's, \mathbf{D}_i 's, $\mathbf{M}_i^{(p)}$'s, $\mathbf{Q}_i^{(j)}$'s, \mathbf{U} and \mathbf{Z}_{ki} 's assumed temporarily fixed.

Closed form solutions can be obtained in the first two steps, whereas the third step is solved with a numerical routine as discussed in the appendix.

The above algorithm is monotonically convergent, since the same criterion (9) is consistently minimized in each step. Note, however, that the monotonic convergence property does not guarantee that the convergence point is what we desire to find. In order to ensure an optimal solution, we may start the algorithm with several different initial estimates, and choose the ones associated with the lowest value of the minimization criterion.

2.3. Additional considerations

2.3.1 Goodness of Fit (GOF) Indices

In Dynamic GSCANNO, the overall fit is measured by the predictability of a given model, stated as

$$\text{FIT} = 1 - \frac{(1-\alpha)\|\boldsymbol{\eta} - \mathbf{X}^* \hat{\mathbf{h}}^*\|^2 + \alpha \sum_{k=1}^K \sum_{i=1}^J \|\gamma_i - \mathbf{Z}_{ki} \hat{\mathbf{w}}_{ki}\|^2}{(1-\alpha)\|\boldsymbol{\eta}\|^2 + \alpha K \sum_{i=1}^J \|\gamma_i\|^2} = 1 - \frac{\hat{\phi}}{(1-\alpha)\|\boldsymbol{\eta}\|^2 + \alpha K \sum_{i=1}^J \|\gamma_i\|^2}, \quad (10)$$

where $\hat{\mathbf{h}}^*$, $\hat{\mathbf{w}}_{ki}$ and $\hat{\phi}$ are the estimates of \mathbf{h}^* , \mathbf{w}_{ki} , and ϕ , respectively. The FIT index (Hwang

& Takane, 2004) is a natural extension of the squared multiple correlation coefficients (R^2) in multiple regression analysis to Dynamic GSCANO. It measures the proportion of the variance in criterion variables that can be explained by the predictor variables. The value of FIT ranges between 0 and 1. In general, the larger the FIT value, the better is the agreement between the observed data (the BOLD signals) and the hypothesized model.

However, the FIT is affected by model complexity. To penalize the FIT by the number of free parameters used in the model, we employ an alternative measure of fit called adjusted goodness of fit (AFIT; Hwang, DeSarbo, & Takane, 2007), which is defined as

$$\text{AFIT} = 1 - (1 - \text{FIT}) \frac{n_0}{n_1}, \quad (11)$$

where $n_0 = TK \sum_{i=1}^J v_i$ is the degree of freedom of the null model (i.e., the model with no explicit structural relations among ROIs) with v_i denoting the number of voxels in the i -th ROI, and where $n_1 = n_0 - r$ is the degree of freedom for the hypothesized model with r indicating the number of free parameters in the model. The AFIT represents a tradeoff between the model fit and model complexity. It tends to favor simpler models over complex ones, given a similar level of model fit.

Note that various local fit indices can also be defined at different levels (Hwang & Takane, 2014, Section 2.9.2). For example, we may take the ratio of the first terms in the numerator and the denominator of the fraction in (10) to indicate the goodness of fit (GOF) of the structural model, and the ratio of the second terms in (10) for the GOF of the measurement model. The former can be further broken down into the GOF of each structural equation, while the latter into subject-wise GOF indices.

2.3.2 *The Bootstrap Method*

The Bootstrap method (Efron, 1982) is utilized to assess the reliability of parameter estimates in Dynamic GSCANO. The entire BOLD signal record of a subject serves as a sampling unit. The units are re-sampled from the original data set K times with replacement, which constitutes one bootstrap sample. Each bootstrap sample is analyzed by Dynamic GSCANO to obtain parameter estimates. This is repeated many times (say, 100 times) to derive empirical distributions of the parameter estimates. We may then calculate means and standard deviations of the parameter estimates. The former provides information about the estimation bias defined as the difference between the estimates from the original data and the mean of the estimates from bootstrap samples (Efron & Tibshirani, 1994, Chapter 10). The latter provides information regarding the reliability of the estimates. To assess the significance of the parameter estimates, in a two-tailed test, 95% confidence intervals may be constructed by taking 2.5% and 97.5% percentile points of the distributions. In a one-tailed test, we simply replace the percentile points from 2.5% or 97.5% to 5% or 95%. If these intervals do not cover 0, the corresponding estimates are regarded statistically significant at the 5% significance level.

A similar bootstrap method may also be used to test the significance of any contrasts between the parameter estimates. Suppose, for example, that we are interested in testing whether the influence of ROI i on ROI j is significantly larger than the other way around. In this case, we can form a contrast, $b_{0,ij} - b_{0,ji}$, whose distributional properties can be obtained in a manner similar to individual parameter estimates.

2.3.3 *Multiple-Group Dynamic GSCANO*

Subjects may have a group structure. These groups may be intact groups such as males and females, or experimentally manipulated treatment groups. Repeated measurement designs may also be considered creating group like structures. In such situations, we may wish to analyze both common and different aspects of the groups. It is not too difficult to estimate parameters in Dynamic GSCANO (Γ , \mathbf{B}_l 's, \mathbf{D}_l 's, $\mathbf{M}_l^{(p)}$'s, and $\mathbf{Q}_l^{(j)}$'s) separately for different groups (to capture distinct aspects of the groups), while optionally keeping some of them equal across the groups (to capture their common aspects). We demonstrate the feasibility of this kind of multiple-group analysis in the second real example in Section 4.

3. Simulation Studies: Recovery of Parameters

We report two of the several Monte Carlo studies conducted (Zhou, 2013). These studies show that Dynamic GSCANO works the way it is supposed to. Specifically, we investigate the parameter recovery capability of Dynamic GSCANO as a function of the number of time points (T), the number of subjects (K), the number of voxels per ROI (v_i), the covariance structure of the BOLD signals ($\Sigma_{kj}^{(M)}$), and presence of a stimulus input. Study 1 roughly emulates the second real data set with no stimulus effects, while Study 2 the first real data set with a stimulus effect. The data generation process is somewhat easier without stimulus effects.

One may wonder why the size of measurement errors is not among the above list (of factors to be examined). In fact, the effects of measurement errors (i.e., σ^2 in $\Sigma_{kj}^{(M)}$, to be explained below) were systematically investigated in Zhou (2013). Specifically, the size of σ^2 was varied at three different levels, which amounted 53%, 59%, and 65% of the total variations in ROIs. We found that the effects of σ^2 were consistently small for T larger than or equal to

100. The recovery rate as measured by the congruence coefficient (to be explained shortly) was all greater than .92 for $K = 30$ and $J = 3$ examined. We thus concluded that Dynamic GSCANO was fairly robust against measurement errors, and the value of σ^2 was set at one level ($\sigma^2 = 0.5$) throughout the rest of the Monte Carlo studies.

3.1. Study 1: A SEM with Seven Latent Variables, Different Covariance Structure of Measurement Errors and No Stimulus Input

In this study, the following structural model was employed:

$$\mathbf{\Gamma} = \mathbf{\Gamma}\mathbf{B}_0 + \mathbf{S}_1\mathbf{\Gamma}\mathbf{B}_1 + \mathbf{E}^{(s)}. \quad (12)$$

The matrix \mathbf{B}_0 is assumed to be hollow (all diagonal entries are zeroes), and \mathbf{B}_1 is assumed to be diagonal, which is equivalent to assuming that the lag 1 effect of a ROI is restricted to within itself. Furthermore, each column of $\mathbf{\Gamma}$ is constrained to have unit variance. The parameters of interest are \mathbf{B}_0 , \mathbf{B}_1 , and $\mathbf{\Sigma}^{(s)}$, where $\mathbf{\Sigma}^{(s)}$ is the covariance matrix among columns of $\mathbf{E}^{(s)}$. However, it is extremely difficult to directly specify \mathbf{B}_0 and \mathbf{B}_1 to generate $\mathbf{\Gamma}$, which is the key step for data generation. Instead, we indirectly specify \mathbf{B}_0 and \mathbf{B}_1 by specifying $\mathbf{C}_{00} = \text{Cov}(\Gamma_t) = \text{Cov}(\Gamma_{t-1})$ where Γ_t and Γ_{t-1} are the t and $(t-1)$ -th rows of $\mathbf{\Gamma}$, respectively.

To explain how, we first rewrite (12) as follows:

$$\begin{aligned} \mathbf{\Gamma} &= \mathbf{S}_1\mathbf{\Gamma}\mathbf{B}_1(\mathbf{I} - \mathbf{B}_0)^{-1} + \mathbf{E}^{(s)}(\mathbf{I} - \mathbf{B}_0)^{-1} \\ &= \mathbf{S}_1\mathbf{\Gamma}\mathbf{C}_1 + \tilde{\mathbf{E}}^{(s)}, \end{aligned} \quad (13)$$

where $\mathbf{C}_1 = \mathbf{B}_1(\mathbf{I} - \mathbf{B}_0)^{-1}$ is the matrix of coefficients for the lag 1 autoregressive effects, and $\tilde{\mathbf{E}}^{(s)} = \mathbf{E}^{(s)}(\mathbf{I} - \mathbf{B}_0)^{-1}$. Given that \mathbf{B}_0 is hollow, and \mathbf{B}_1 is diagonal, they can be uniquely determined by \mathbf{C}_1 , namely $\mathbf{B}_0 = \mathbf{I} - \mathbf{C}_1^{-1}\mathbf{B}_1$ and $\mathbf{B}_1 = (\text{diag}(\mathbf{C}_1^{-1}))^{-1}$, the latter of which can be derived from $\text{diag}(\mathbf{B}_0) = \mathbf{0} = \mathbf{I} - \text{diag}(\mathbf{C}_1^{-1})\mathbf{B}_1$. Let $\tilde{\Sigma}^{(s)}$ denote the covariance matrix of $\tilde{\mathbf{E}}^{(s)}$. The matrix $\Sigma^{(s)}$ can be derived by $\Sigma^{(s)} = (\mathbf{I} - \mathbf{B}'_0)\tilde{\Sigma}^{(s)}(\mathbf{I} - \mathbf{B}_0)$. With \mathbf{C}_{00} being generated, the matrix \mathbf{C}_1 and $\tilde{\Sigma}^{(s)}$ in these formulas can be specified by an algorithm (appendix B1).

To generate Γ , we apply

$$\Gamma_t = \Gamma_{t-1}\mathbf{C}_1 + \tilde{\mathbf{E}}_t^{(s)}, \quad (14)$$

recursively, where $\tilde{\mathbf{E}}_t^{(s)}$ denotes the t -th row of $\tilde{\mathbf{E}}^{(s)}$. The Γ_0 is sampled from $N(\mathbf{0}, \mathbf{C}_{00})$ and then each Γ_t is updated by (14) with the value of $\tilde{\mathbf{E}}_t^{(s)}$ sampled from $N(\mathbf{0}, \tilde{\Sigma}_t^{(s)})$.

We then generate \mathbf{Z}_{kj} by $\mathbf{Z}_{kj} = \gamma_j \mathbf{c}'_j + \mathbf{E}_{kj}^{(M)}$, where $\mathbf{c}_j = (c_{j1}, \dots, c_{jv_j})'$ is a v_j -element vector of component loadings. The loading vector \mathbf{c}_j indicates the strength of the relationships between the BOLD signals and the corresponding latent variable in this study. Each row of the T by v_j measurement error matrix $\mathbf{E}_{kj}^{(M)}$ is randomly sampled from $N(\mathbf{0}, \Sigma_{kj}^{(M)})$. Finally, \mathbf{Z}_{kj} is column-wise standardized.

In this study, the number of ROIs was set to seven ($J = 7$) in accordance with the second real data set. The number of time points were varied at four levels ($T = 50, 100, 200, \text{ and } 500$), the number of subjects at two levels ($K = 15 \text{ and } 30$), the number of voxels at two levels ($v_i = 3$

and 50), and the covariance structure for $\Sigma_{kj}^{(M)}$ at two levels ($\rho = 0$ and $.9$). Other specifications and assumed true values of \mathbf{C}_{00} , \mathbf{c}_j , and $\Sigma_{kj}^{(M)}$ are given in Appendix (B2). To measure the quality of parameter recovery, we used the congruence coefficient. Let \mathbf{h}^* be the column vector of free population path coefficients used to generate the simulation data, and let $\hat{\mathbf{h}}_n^*$ be its estimates from the n -th simulated data set. The congruence coefficient is defined as $\mathbf{h}^{*\prime} \hat{\mathbf{h}}_n^* / \left(\sqrt{\mathbf{h}^{*\prime} \mathbf{h}^*} \sqrt{\hat{\mathbf{h}}_n^{*\prime} \hat{\mathbf{h}}_n^*} \right)$.

We report the mean value and standard deviation of this coefficient over N simulated samples. The mean indicates how well model parameters are recovered on average, while the standard deviation indicates how stable the recovery rate is. The mean value of the congruence coefficient greater than 0.9 is conventionally regarded as an acceptable level of recovery (Muliak, 1972). Table 1 presents a summary of the study. It is observed that there are little effects of K for $v_i = 3$, little effects of v_i , and little effects of ρ for $T \geq 100$. (The parameter recovery is better for $\rho = 0$ than $\rho = 0.9$ when T is small. However, no significant difference is observed for $T \geq 100$.)

Table 1. Means and standard deviations (SD) of the congruence coefficients for the path coefficients.

	#voxels	3	3	50	50
	K	15	30	30	30
T	ρ	0	0	0	0.9
50	Mean	0.7669	0.8021	0.6850	0.6244
	SD	(0.1211)	(0.0894)	(0.2187)	(0.2414)
100	Mean	0.9291	0.9308	0.9312	0.9245
	SD	(0.0250)	(0.0241)	(0.0256)	(0.0309)
200	Mean	0.9663	0.9664	0.9658	0.9666
	SD	(0.0105)	(0.0082)	(0.0093)	(0.0094)
500	Mean	0.9812	0.9812	0.9805	0.9803
	SD	(0.0042)	(0.0038)	(0.0044)	(0.0041)

To sum up, the number of time points T plays an important role in determining the quality of parameter estimates. As long as the sample size T is no less than 100, the effects of the number of subjects between 15 and 30, the number of voxels per ROI between 3 and 50, and the covariance structure of $\Sigma_{kj}^{(M)}$ are negligible.

3.2. Study 2: A SEM with Four Latent Variables and a Stimulus Input

The second simulation study examines the effect of T on parameter recovery in the presence of a stimulus input. In this study, the following structural model was employed:

$$\mathbf{\Gamma} = \mathbf{\Gamma}\mathbf{B}_0 + \mathbf{S}_1\mathbf{\Gamma}\mathbf{B}_1 + \mathbf{U}\mathbf{D} + \mathbf{E}^{(s)}. \quad (15)$$

where \mathbf{U} stands for the T -element vector of the stimulus effect. For a specific value of T , \mathbf{U} is fixed across all simulated data sets. As before, $\mathbf{\Gamma}$ should satisfy the column-wise unit variance constraint. The model (15) can be rewritten as

$$\mathbf{\Gamma} = \mathbf{S}_1\mathbf{\Gamma}\mathbf{B}_1(\mathbf{I} - \mathbf{B}_0)^{-1} + \mathbf{U}\mathbf{D}(\mathbf{I} - \mathbf{B}_0)^{-1} + \mathbf{E}^{(s)}(\mathbf{I} - \mathbf{B}_0)^{-1} = [\mathbf{U}, \mathbf{S}_1\mathbf{\Gamma}] \mathbf{C}_1 + \tilde{\mathbf{E}}^{(s)},$$

where $\mathbf{C}_1 = \begin{bmatrix} \mathbf{D} \\ \mathbf{B}_1 \end{bmatrix} (\mathbf{I} - \mathbf{B}_0)^{-1}$ and $\tilde{\mathbf{E}}^{(s)} = \mathbf{E}^{(s)}(\mathbf{I} - \mathbf{B}_0)^{-1}$. Let $\mathbf{C}_{00} = \text{Cov}([\mathbf{U}_t, \mathbf{\Gamma}_{t-1}])$, where \mathbf{U}_t is the t -

th element of \mathbf{U} . The values of \mathbf{C}_1 and $\tilde{\Sigma}^{(s)}$ are specified from \mathbf{C}_{00} with an algorithm (appendix B1). Then, the path coefficients can be derived similarly as in Study 1. The matrix of latent variables $\mathbf{\Gamma}$ was generated as in (14). In this study, the number of ROIs were set to four ($J = 4$), and the number of stimulus inputs to one ($P = 1$) in accordance with the real data set 1. The number of time points was varied at five levels ($T = 25, 50, 100, 200, \text{ and } 500$), and the number

of voxels per ROI was set to three in all cases. Other relevant specifications as well as assumed true values of \mathbf{C}_{00} , $\tilde{\Sigma}^{(S)}$, \mathbf{c}_j , and $\Sigma_{kj}^{(M)}$ in this study are given in Appendix (B3).

The results are summarized in Table 2. The mean congruence coefficient increases monotonically as T increases, while the opposite is true for the standard deviation. Again, as T goes beyond 100, the mean congruence coefficient goes above .90.

Table 2. Means and standard deviations (SD) of the congruence coefficients for the path coefficients.

T	Mean	SD
25	0.7925	(0.0820)
50	0.8846	(0.0408)
100	0.9195	(0.0242)
200	0.9382	(0.0144)
500	0.9497	(0.0072)

Note that all the structural and measurement errors were assumed to be normally distributed throughout our studies. This assumption may be questionable in some empirical settings, and it will be interesting to see how the fitting criterion and associated algorithm will react to non-normally distributed errors. Another limitation to the present simulation study is that we have not investigated the robustness of the proposed method against misspecifications of the model. Despite these shortcomings, we conclude that the simulation studies are useful to demonstrate that Dynamic GSCANNO performs reasonably well.

4. Applications of Dynamic GSCANNO to Real Functional Neuroimaging Data

4.1. The Visual Task Data

The first example pertains to the “visual task” data reported in Fu et al. (2012), in which a group of fourteen subjects viewed a flickering checkerboard. The experiment comprised three repetitions of a 20-second fixation and 20-second flickering checkerboard viewing trial, followed by another fixation period at the end of the overall experiment. Each of fMRI time series data consists of 225 scans ($T = 255$). The preprocessing of the fMRI data was performed using SPM8 (Statistical Parametric Mapping; <http://www.fil.ucl.ac.uk/spm/>), including motion correction, normalization of the functional images via normalizing the anatomical image and spatial smoothing using 8mm Gaussian kernel. The authors used GLM in SPM to obtain brain regions related to the visual task. They found four activated regions ($J = 4$) which were the left fusiform gyrus (FuG_L; γ_1), the right fusiform gyrus (FuG_R; γ_2), the left middle occipital gyrus (MOG_L; γ_3), and the right middle occipital gyrus (MOG_R; γ_4). Each ROI contains 81 voxels. There is one stimulus input \mathbf{U} ($P = 1$), the flickering checker board, as depicted in Figure 1.

< Insert Figure 1 here >

All four ROIs are considered crucial in completing a visual cognitive task (Grill-Spector, Kourtzi, & Kanwisher, 2001). However, the left and right MOGs are regarded as functionally distinct from the left and right FuG (Kanwisher et al., 1997; Epstein & Kanwisher, 1998; Epstein, Harris, Stanley, & Kanwisher, 1999; Grill-Spector, Kourtzi, & Kanwisher, 2001). It is pointed out that MOG_L and MOG_R can be viewed as a general-purpose system for analyzing object shapes, whereas FuG_L and FuG_R are more specialized in the recognition of human faces (Grill-Spector, Kourtzi, & Kanwisher, 2001). More recently, it is advocated that FuG_L and FuG_R are engaged in different stages of face recognition, that is, the signal is first sent from MOG_L and MOG_R to FuG_L for identifying face like objects (Meng et al., 2012). Only upon

detecting face like signals, they are passed on to FuG_R for further processing. Also, it has been noted that the connections from early visual areas (e.g., MOG_L and MOG_R) to higher visual areas (e.g., FuG_L and FuG_R) are sensitive to the level of stimulus visibility (Haynes et al., 2005). This observation is further supported by the finding of Fu et al. (2012) that, when viewing a flickering checkerboard, the functional connectivity between MOG (left and right) and FUG (left and right) is less stable.

Dynamic GSCANO with one stimulus input was applied to this data set. The structural model can be written as

$$\gamma_j = \sum_{i \neq j} b_{0,ij} \gamma_i + b_{1,j} \mathbf{S}_1 \gamma_i + d_{0,j} \mathbf{u} + \sum_{i \neq j} m_{0,ij} \text{diag}(\mathbf{u}) \gamma_i \text{ for } j = 1, \dots, 4, \quad (16)$$

where $b_{0,ij}$, $b_{1,j}$, $d_{0,j}$ and $m_{0,ij}$ are the coefficients to be estimated. This model includes all possible direct and modulating stimulus effects.

With an Intel 2.50GHz CPU and 8 GB Ram, the average processing time for parameter estimation was 90 seconds, which resulted in a total of 2.5 hours to analyze 100 bootstrap samples. Note, however, that the entire data set contains over 1 million data points. We obtained FIT = 0.750 (with the standard error of 0.021) and AFIT = 0.741 (with the standard error of 0.022), indicating that the model accounts for approximately three fourths of the variability in observed and latent variables. Table 3 gives the estimates of the contemporaneous effects (off-diagonal) and of the lag 1 autoregressive effects (diagonal), their bootstrap standard errors, and the corresponding p -values. The p -value indicates the percentage of bootstrap estimates with the same sign as the one obtained from the original data. Therefore, a p -value of less than 0.01 indicates that an opposite sign has never occurred during the 100 bootstrap replications.

Only a few contemporaneous effects among different ROIs are significant, while all lag 1 autoregressive effects are significant. Table 4 provides the summary statistics for the direct stimulus effects. Only the direct effects of the stimulus on MOG_L and MOG_R are significant. Table 5 gives the modulating effects of the stimulus. Only the effect on the connection from MOG_L to FuG_L is significant.

Table 3. The estimates of path coefficients (top) among ROIs, their bootstrap standard errors (middle), and the p -values (bottom). ROIs in rows exert influence on ROIs in columns. Diagonal entries give the lag 1 autoregressive effects. Path coefficients with one asterisk is significant at $\alpha = 0.05$, and those with two asterisks at $\alpha = 0.01$.

	FuG_L	FuG_R	MOG_L	MOG_R
	-0.777	0.011	0.084	-0.185
FuG_L	(-0.053)	(0.092)	(0.110)	(0.163)
	**<0.01	0.64	0.23	0.18
	-0.240	-0.726	0.122	-0.281
FuG_R	(0.124)	(0.063)	(0.150)	(0.240)
	0.24	**<0.01	0.42	0.42
	0.274	0.186	-0.742	0.471
MOG_L	(0.130)	(0.199)	(0.160)	(0.220)
	*0.05	0.33	**<0.01	*0.02
	-0.089	-0.298	0.762	0.692
MOG_R	(0.183)	(0.250)	(0.262)	(0.200)
	0.16	0.44	**0.01	*0.02

Table 4. The estimates of the direct stimulus effects (top), their bootstrap standard errors (middle), and the p -values (bottom). Path coefficients with one asterisk are significant at $\alpha = 0.05$, and those with two asterisks at $\alpha = 0.01$.

	Estimates	Standard Error	P -value
U->FuG_L	-0.002	(0.110)	0.55

U->FuG_R	0.360	(0.133)	0.10
U->MOG_L	0.138	(0.106)	*0.05
U->MOG_R	0.055	(0.070)	**<0.01

Table 5. The same as in Table 4, but for the modulating effects of stimulus. A path coefficient with two asterisks is significant at $\alpha = 0.01$.

Stimulus	Path	Estimates	Standard Error	P-value
	(FuG_L->FuG_R)	0.335	(0.206)	0.21
U->	(FuG_L->MOG_L)	-0.508	(0.225)	0.13
	(FuG_L->MOG_R)	0.540	(0.270)	0.20
	(FuG_R->FuG_L)	0.365	(0.169)	0.26
U->	(FuG_R->MOG_L)	0.115	(0.196)	0.33
	(FuG_R->MOG_R)	0.036	(0.253)	0.60
	(MOG_L->FuG_L)	-0.310	(0.180)	**<0.01
U->	(MOG_L->FuG_R)	-0.131	(0.170)	0.48
	(MOG_L->MOG_R)	-0.310	(0.160)	0.11
	(MOG_R->FuG_L)	0.490	(0.290)	0.11
U->	(MOG_R->FuG_R)	-0.740	(0.270)	0.14
	(MOG_R->MOG_L)	0.360	(0.230)	0.50

Figure 2 displays the model we arrived at based on the present analysis, in which only significant connections are retained. Several remarks are in order. First of all, the significant bidirectional paths between MOG_L and MOG_R as well as the significant direct effects from stimulus to these two ROIs support Grill-Spector et al.'s (2001) contention that both left and right MOGs are actively engaged in an early stage of visual task. To better illustrate this point, Figure 3 displays the estimated ROI time series. Three-cycle periodic patterns can be clearly observed in the derived ROI time series corresponding to the left and right MOGs. Such a pattern is closely related to the similar three-cycle periodic pattern of the stimulus curve in Figure 1.

< Insert Figures 2 & 3 here >

Secondly, the significant one-way connection from MOG_L to FuG_L provides evidence for Meng et al. (2012) and Grill-Spector et al. (2001) to the effect that FuG_L is nested in a hierarchical network of the visual cognitive system, and receives signals from lower visual cortex. Thirdly, the significant modulating effect of stimulus on the connection from MOG_L to FuG_L justifies the claim by Haynes et al. (2005) that the level of stimulus strength affects the signal transferring from lower visual areas to higher visual areas. Lastly, Dynamic GSCANO asserts that under the visual task of the current experimental design, FuG_R is absent from any connections with the other ROIs. This may be counterintuitive at the first glance. However, it provides the first model-based evidence to Meng et al.'s (2012) finding that FuG_R is specialized in face recognitions and only “faces like” signals are sent to FuG_R. Since a flickering checkerboard is nothing like faces, it is reasonable not to observe a significant connection from FuG_L to FuG_R.

4.2. *The Working Memory Data*

This study pertains to the working memory task data (Metzak, Riley, Wang, Whitman, Ngan, & Woodward, 2011). There are two groups of subjects, the normal control group and the schizophrenic group, and each group consisted of fifteen subjects. Each subject completed a repeated working memory task to judge whether or not consonants were presented previously. During the experimental period, 214 fMRI scans were recorded ($T = 214$). SPM99 was used for image realignment, normalization into modified Talairach stereotaxic anatomical space, and smoothing with a Gaussian kernel (8mm full width at half maximum) to compensate for inter-subject anatomical differences and optimize the signal to noise ratio. Constrained principal

component analysis (CPCA; Takane & Hunter 2001) was first applied to the original BOLD signal data, in which a finite impulse response (FIR) model was fitted with stimulus onset times as the predictor variables. The predictable portions of the BOLD signal data were then subjected to PCA to extract a component representing a major functional network comprising several brain regions. The seven clusters of voxels serving as ROIs in the present study were identified based on a cut-off value of the highest 5% of the first component loadings. The 7 ROIs are: inferior parietal lobule in the left hemisphere (IPL_L; γ_1 ; 135 voxels), precentral gyrus in the left hemisphere (PreCG_L; γ_2 ; 144 voxels), cerebellum in the left hemisphere (CL_L; γ_3 ; 40 voxels), cerebellum in the right hemisphere (CL_R; γ_4 ; 157 voxels), inferior parietal lobule in the right hemisphere (IPL_R; γ_5 ; 68 voxels), precentral gyrus in the right hemisphere (PreCG_R; γ_6 ; 53 voxels), and supplementary motor area (SMA; γ_7 ; 180 voxels). As reported in Metzrak et al. (2011), all of these ROIs are activated during the working memory task, and there are significant discrepancies with respect to the stability of voxel activations between the two groups. We therefore decided to first apply multiple-group Dynamic GSCANNO with completely separate parameters for the two groups, which was essentially the same as separate analyses of the two groups. The hypothesized structural model is

$$\gamma_j^{(g)} = \sum_{i \neq j} b_{0,ij}^{(g)} \gamma_i^{(g)} + b_{1,j}^{(g)} \mathbf{S}_1 \gamma_j^{(g)} + \mathbf{e}_j^{(g)} \text{ for } j = 1, \dots, 7 \text{ and } g = 1, 2, \quad (17)$$

where the parenthesized superscript g indexes a group. The $g = 1$ refers to the normal control group, and $g = 2$ the schizophrenic group. For each g , $b_{1,j}^{(g)}$ indicates the coefficient of the lag 1 autoregressive effect of the j -th ROI on itself, and $b_{0,ij}^{(g)}$ indicates the coefficient of the

contemporaneous effect of ROI i on ROI j . The model hypothesizes that ROIs are fully bidirectionally connected, and that the only autoregressive effects are those from a ROI on itself. This kind of fully connected path model is prevalent in functional neuroimaging studies (Friston, Harrison, & Penny, 2003). It took 340 seconds for parameter estimation in one bootstrap sample and a total of 9 hours to analyze 100 bootstrap samples. Note, however, that the whole data set contains nearly 5 million data points. The overall goodness of fit indices for both groups combined turn out to be $FIT = 0.863$ (with the standard error of 0.007), and $AFIT = 0.861$ (with the standard error of 0.007), indicating about 86% of the total variance is explained by the model.

The time series plots of the estimated latent variables for both groups are presented in Figure 4. It can be observed that the time series plots for the normal control group are generally smoother and display a clearer trend in comparison with those for the schizophrenic group which are more volatile. In the normal control group, the activations are low at the beginning, then go up and then down, and go up again toward the end across all ROIs, while no such clear patterns can be observed in the schizophrenic group. This is consistent with the findings of Metzak et al. (2011).

< Insert Figure 4 here >

Tables 6a and 6b display the estimates of path coefficients, their bootstrap standard errors, and the p -values of the path coefficients obtained for the normal group (a) and the schizophrenic group (b). The diagonal cells of the tables indicate the lag 1 autoregressive effects, whereas the off-diagonal elements the contemporaneous effects. The ROIs in the rows are used as predictor variables to predict the ROIs in the columns. The standard errors of the path coefficients and the

p -values were calculated from 100 bootstrap samples. All contemporaneous effects are significant one way or the other (either positively or negatively). This means that all the ROIs are significantly bi-directionally connected. This is consistent with Metzack et al.'s (2011) remark that all seven ROIs are involved in the working memory task for both groups. On the other hand, none of the autoregressive effects are significant in either group.

Table 6a. The estimates of path coefficients among ROIs (top), their bootstrap standard errors (middle), and the p -values (bottom) for the normal control group obtained from the working memory data. ROIs in rows exert influence on ROIs in columns. Diagonal entries give the lag 1 autoregressive effects. Significant path coefficients are indicated by one asterisk for $\alpha = 0.05$, and by two asterisks for $\alpha = 0.01$.

	IPL_L	PreCG_L	CL_L	CL_R	IPL_R	PreCG_R	SMA
IPL_L	0.007 (0.008) 0.17	0.278 (0.106) **<0.01	-0.306 (0.119) **<0.01	0.379 (0.118) **<0.01	-1.044 (0.431) **<0.01	1.301 (0.661) **<0.01	-0.411 (0.140) **<0.01
PreCG_L	4.071 (1.484) **<0.01	-0.001 (0.002) 0.34	1.104 (0.150) **<0.01	-1.427 (0.315) **<0.01	3.944 (1.494) **<0.01	-5.247 (3.169) **<0.01	1.526 (0.334) **<0.01
CL_L	-3.731 (1.405) **<0.01	0.920 (0.117) **<0.01	0.000 (0.002) 0.51	1.307 (0.310) **<0.01	-3.648 (1.546) **<0.01	4.738 (2.780) **<0.01	-1.409 (0.365) **<0.01
CL_R	2.889 (0.972) **<0.01	-0.743 (0.208) **<0.01	0.817 (0.240) **<0.01	-0.001 (0.002) 0.24	2.800 (1.012) **<0.01	-3.623 (1.805) **<0.01	1.082 (0.174) **<0.01
IPL_R	-1.139 (0.585) **<0.01	0.290 (0.116) **<0.01	-0.321 (0.132) **<0.01	0.394 (0.122) **<0.01	0.005 (0.010) 0.26	1.358 (0.710) **<0.01	-0.427 (0.149) **<0.01
PreCG_R	0.955 (0.477) **<0.01	-0.267 (0.174) **<0.01	0.287 (0.170) **<0.01	-0.358 (0.225) **<0.01	0.930 (0.491) **<0.01	0.003 (0.016) 0.53	0.380 (0.211) **<0.01
SMA	-2.752 (1.112) **<0.01	0.697 (0.206) **<0.01	-0.774 (0.273) **<0.01	0.947 (0.155) **<0.01	-2.661 (1.109) **<0.01	3.407 (1.788) **<0.01	0.001 (0.001) 0.20

Table 6b. The same as in Table 6a, but for the Schizophrenic Group.

	IPL_L	PreCG_L	CL_L	CL_R	IPL_R	PreCG_R	SMA
IPL_L	0.002 (0.002)	0.636 (0.126)	-1.819 (0.842)	-1.916 (0.589)	2.619 (1.594)	-1.762 (0.824)	-1.128 (0.262)
	0.22	**<0.01	**<0.01	**<0.01	**<0.01	**<0.01	**<0.01
PreCG_L	1.646 (0.404)	-0.002 (0.002)	2.979 (1.470)	3.065 (0.926)	-4.029 (2.049)	2.846 (1.379)	1.811 (0.440)
	**<0.01	0.11	**<0.01	**<0.01	**<0.01	**<0.01	**<0.01
CL_L	-0.638 (0.232)	0.406 (0.178)	0.001 (0.005)	-1.227 (0.600)	1.647 (1.140)	-1.032 (0.407)	-0.699 (0.249)
	**<0.01	**<0.01	0.46	**<0.01	**<0.01	**<0.01	**<0.01
CL_R	-0.576 (0.199)	0.355 (0.103)	-1.048 (0.592)	0.001 (0.004)	1.438 (0.829)	-1.008 (0.555)	-0.625 (0.179)
	**<0.01	**<0.01	**<0.01	0.37	**<0.01	**<0.01	**<0.01
IPL_R	0.499 (0.275)	-0.297 (0.119)	0.899 (0.630)	0.916 (0.466)	0.0126 (0.016)	0.899 (0.719)	0.544 (0.264)
	**<0.01	**<0.01	**<0.01	**<0.01	0.15	**<0.01	**<0.01
PreCG_R	-0.660 (0.236)	0.413 (0.151)	-1.102 (0.403)	-1.270 (0.652)	1.780 (1.463)	0.007 (0.008)	-0.732 (0.290)
	**<0.01	**<0.01	**<0.01	**<0.01	**<0.01	0.09	**<0.01
SMA	-0.946 (0.290)	0.585 (0.146)	-1.656 (0.714)	-1.746 (0.567)	2.400 (1.497)	-1.626 (0.785)	0.000 (0.002)
	**<0.01	**<0.01	**<0.01	**<0.01	**<0.01	**<0.01	0.44

It is also of interest to see asymmetry between $b_{0,ij}^{(g)}$ and $b_{0,ji}^{(g)}$. Table 7 gives the directionality tests for the two groups. Each cell in these tables indicates the difference between the effect of a ROI in a row on a ROI in a column and its reverse. Thus, for example, the value of 3.793 in the top row of cell (1, 1) of Table 7 indicates the difference between the effect of PreCG_L on IPL_L and that of IPL_L on PreCG_L. The corresponding p -value indicates the relative frequency with which the former was smaller than the latter in the 100 bootstrap samples. The p -value of 0 shows that this never happened, indicating that the influence of PreCG_L on IPL_L is significantly larger than the other way round, and such asymmetry is likely to be stable

within the normal control group. Similarly, the negative effect of CL_L on IPL_L is significantly larger than the other way round.

Table 7a. The estimates of the differences between the path coefficients in one direction and the opposite direction (top), their bootstrap standard errors (middle), and the p -values (bottom) for the normal control group in working memory data. Significant asymmetric relationships are indicated by one asterisk for $\alpha = 0.05$, and by two asterisks for $\alpha = 0.01$.

	IPL_L	PreCG_L	CL_L	CL_R	IPL_R	PreCG_R
PreCG_L	3.793 (1.578) **<0.01					
CL_L	-3.426 (1.511) **<0.01	-0.184 (0.265) 0.22				
CL_R	2.510 (1.077) **<0.01	0.684 (0.514) 0.08	-0.490 (0.539) 0.15			
IPL_R	-0.094 (0.964) 0.47	-3.654 (1.597) **<0.01	3.327 (1.663) **<0.01	-2.406 (1.123) **<0.01		
PreCG_R	-0.346 (1.085) 0.41	4.981 (2.300) **<0.01	-4.452 (2.915) **<0.01	3.265 (1.981) *0.02	-0.428 (1.142) 0.38	
SMA	-2.341 (1.235) **<0.01	-0.829 (0.531) 0.07	0.635 (0.626) 0.15	-0.135 (0.327) 0.36	-2.234 (1.241) **<0.01	3.028 (1.953) 0.02

Table 7b. The same as in Table 7a, but for the Schizophrenic Group.

	IPL_L	PreCG_L	CL_L	CL_R	IPL_R	PreCG_R
PreCG_L	1.011 (0.524) **<0.01					
CL_L	1.180 (1.046) 0.10	2.573 (1.619) **0.01				
CL_R	1.340 (0.775) *0.02	2.710 (1.024) **<0.01	0.180 (1.138) 0.39			
IPL_R	2.120 (1.105) *0.03	3.733 (2.150) **<0.01	0.748 (1.640) 0.30	0.522 (1.252) 0.36		
PreCG_R	1.102 (1.038) 0.08	2.433 (1.212) **<0.01	0.069 (0.790) 0.48	0.262 (1.146) 0.39	0.884 (0.977) 0.30	
SMA	0.182 (0.539) 0.32	1.226 (0.579) **0.01	0.957 (0.946) 0.14	1.122 (0.537) *0.03	1.858 (1.717) 0.06	0.894 (1.048) 0.20

There were, in total, 12 significantly asymmetric relationships in the normal groups, and 9 in the schizophrenic group. Unfortunately, empirical implication of these findings is yet to be investigated because there is little established evidence regarding the directionality among these ROIs in the working memory task.

By looking at Figure 4 more closely, it may be noticed that for the 4th ROI, namely CL_R, the activation patterns are quite similar across the two groups. This is rather unique in the sense that all other ROIs show distinct intergroup activation patterns. This naturally leads to the question that the activation of CL_R could be regarded homogeneous across the groups. To answer this question, another analysis was conducted in which both groups were analyzed simultaneously under the restriction that $\gamma_4^{(1)} = \gamma_4^{(2)}$. The global goodness of fit indices turned out

to be $FIT = 0.860$ with the standard error of 0.007 computed based on 100 bootstrap data sets, and $AFIT = 0.859$ with the standard error of 0.007. These values are all only negligibly smaller than those obtained in the previous analysis, indicating that equating activations in CL_R across the two groups is at least not harmful. We have also found that the estimated ROI time series are essentially the same as in the previous study except that the activations on CL_R are now identical across the groups. The evidence above shows that the activations of CL_R is virtually identical across two groups, indicating CL_R is relatively unaffected by the disease.

5. Summary and Discussion

We proposed Dynamic GSCANO as a SEM approach to analyzing multiple-subject effective connectivity. Dynamic GSCANO combines multiple-set canonical correlation analysis with a multivariate autoregressive model to deal with time series data. It is capable of analyzing both contemporaneous and time lagged effects of ROIs as well as direct and modulating effects of stimuli. These effects can also be analyzed by euSEM (Gates, et al., 2011) and dynamic GSCA (Jung et al., 2012). However, Dynamic GSCANO is unique in several respects. It is one of the first SEM-based methods for analyzing effective connectivity of multiple subjects simultaneously. The proposed method uses GCANO as the measurement model to relate observed fMRI data to latent variables in such a way that the latent variables capture the most representative activations in the ROIs across subjects. As argued earlier, GCANO is robust against outlying subjects. Such individuals can be easily detected by the local fit indices discussed at the end of Section 2.3.1. The directional associations among the extracted latent variables are captured by the structural model, which is essentially a hypothesized vector (multivariate) autoregressive time series model. Furthermore, the structural model in Dynamic

GSCANNO includes one extra term that is not subsumed by euSEM or dynamic GSCA, namely, the latent interactions among ROIs.

We demonstrated the usefulness of Dynamic GSCANNO with simulated and empirical data sets. In the simulation studies, we found that the number of time points (T) was the key factor in determining the parameter recovery capability of Dynamic GSCANNO. For $T = 100$ or greater, the parameter recovery was good to excellent (i.e., over 90% in the congruence coefficient between parameters and their estimates) and was robust against a variety of factors such as the size of measurement errors, the covariance structure of the measurement errors in the BOLD signals, and the presence of a stimulus input. Furthermore, we showed the usefulness of Dynamic GSCANNO in empirical research involving multiple subjects, who may all be in one group or in different groups. The numerically stable algorithm developed for Dynamic GSCANNO could successfully fit a complex structural model with many reciprocal (bidirectional) connections among ROIs.

Dynamic GSCANNO can be further extended in a variety of directions to enhance its data analytic capability. These include resolving possible heterogeneity among subjects, and dealing with endogeneity in the structural model. They are elaborated below.

At present, we assume that subjects within a group are drawn from a relatively homogeneous population, which favors a unique type of intra-dynamics among ROIs. However, sometimes there may be heterogeneous subgroups of subjects within a group or groups, and that the intra-dynamics among ROIs differ across these subgroups. Indeed, misleading findings may arise if the connectivity maps for individuals vary across samples (Gates & Molenaar, 2012). An appealing approach for addressing subject level heterogeneity is to combine Dynamic GSCANNO

with fuzzy clusterwise linear regression (FCR; Hathaway & Bezdek, 1993) in a unified framework, which accommodates both the clustering of subjects into heterogeneous subgroups and the estimation of clusterwise path coefficients.

Another possible extension is to deal with the endogeneity problem. Endogeneity pertains to the fact that, in the structural model, there are non-trivial correlations between the predictor variables Γ_t and the error terms $E_t^{(s)}$ (Lütkepohl, 2007). Ordinary least squares estimates of parameters could be biased in such situations (Hayashi, 2000). However, the endogeneity problem is unavoidable because of the multiple autoregressive time series nature of the structural model (Lütkepohl, 2007). Indeed, the endogeneity issue exists not only in Dynamic GSCANO but also in euSEM and dynamic GSCA. One possible way to reduce the influence of endogeneity is to employ instrumental variable estimation (Hayashi, 2000; Bowden & Turkington, 1990; Greene, 2011). This means finding a multiple time series record \mathbf{W} which is significantly correlated with $\mathbf{\Gamma}$ and yet exhibits no correlations with $\mathbf{E}^{(s)}$. Once \mathbf{W} is found, we pre-multiply both sides of (6) by $\mathbf{P}_{\mathbf{W}}$, the orthogonal projector onto the linear subspace spanned by the columns of \mathbf{W} . The above approach has been occasionally criticized as easier said than done, since it is not easy to find suitable instrumental variables to be used in practice (Baum, 2008; Berkowitz, Caner & Fang, 2012). Nonetheless, it is worthwhile investigating its feasibility with Dynamic GSCANO.

We have written a MATLAB program for Dynamic GSCANO to produce the results reported in this paper (available upon request). However, this program is not very user-friendly. We realize the importance to develop a general-purpose program for Dynamic GSCANO with interactive input/output features and a graphical interface. Future research on extensions of

Dynamic GSCANO will further enhance its applicability in the investigation of neuronal pathways in the human brain.

Appendices

Appendix A. The alternating least squares (ALS) algorithm for Dynamic GSCANO

We provide derivations of the three steps in the parameter estimation algorithm outlined in Section 2.2. Two linear algebra operations are particularly useful in developing updating formulas: the vectorization operator, vec , and the Kronecker product, \otimes . The vec operator forms a super-vector by stacking the columns of a given matrix. The Kronecker product of two matrices \mathbf{A} and \mathbf{B} , $\mathbf{A} \otimes \mathbf{B}$, is defined as $\mathbf{A} \otimes \mathbf{B} = [a_{ij}\mathbf{B}]$, where $\mathbf{A} = [a_{ij}]$.

In step I, \mathbf{w}_{ki} is related to $\varphi^{(M)}$ only. Hence, it suffices to minimize (7). Such a \mathbf{w}_{ki} can be obtained by $\hat{\mathbf{w}}_{ki} = (\mathbf{Z}'_{ki}\mathbf{Z}_{ki})^{-1}\mathbf{Z}'_{ki}\boldsymbol{\gamma}_i$.

In step II, \mathbf{B}_l 's, \mathbf{D}_l 's, $\mathbf{M}_l^{(p)}$'s, and $\mathbf{Q}_l^{(j)}$ are related to $\varphi^{(S)}$ only. Hence, it suffices to minimize (8), which can be rewritten as

$$\phi^{(S)} = \|\boldsymbol{\eta} - \mathbf{X}\mathbf{h}\|^2, \quad (\text{A1})$$

where $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \mathbf{X}_4]$, in which

$$\mathbf{X}_1 = [\mathbf{I}_J \otimes \mathbf{S}_0\boldsymbol{\Gamma}, \dots, \mathbf{I}_J \otimes \mathbf{S}_L\boldsymbol{\Gamma}],$$

$$\mathbf{X}_2 = [\mathbf{I}_J \otimes \mathbf{S}_0\mathbf{U}, \dots, \mathbf{I}_J \otimes \mathbf{S}_L\mathbf{U}],$$

$$\mathbf{X}_3 = [\mathbf{I}_J \otimes \mathbf{S}_0\text{diag}(\mathbf{u}_1)\boldsymbol{\Gamma}, \dots, \mathbf{I}_J \otimes \mathbf{S}_L\text{diag}(\mathbf{u}_p)\boldsymbol{\Gamma}],$$

$$\mathbf{X}_4 = [\mathbf{I}_J \otimes \mathbf{S}_0 \text{diag}(\gamma_1) \Gamma, \dots, \mathbf{I}_J \otimes \mathbf{S}_L \text{diag}(\gamma_J) \Gamma],$$

and

$$\mathbf{h} = \left[\text{vec}([\mathbf{B}_0, \dots, \mathbf{B}_L])', \text{vec}([\mathbf{D}_0, \dots, \mathbf{D}_L])', \text{vec}([\mathbf{M}_0^{(1)}, \dots, \mathbf{M}_L^{(P)}])', \text{vec}([\mathbf{Q}_0^{(1)}, \dots, \mathbf{Q}_L^{(J)}])' \right]'$$

Let \mathbf{h}^* be the vector constructed from \mathbf{h} by eliminating all zero entries, and let \mathbf{X}^* be the matrix formed from \mathbf{X} by eliminating the corresponding columns. One may rewrite (A1) further as

$$\phi^{(S)} = \|\boldsymbol{\eta} - \mathbf{X}^* \mathbf{h}^*\|^2. \quad (\text{A2})$$

The estimates of \mathbf{h}^* is given by $\hat{\mathbf{h}}^* = (\mathbf{X}^{*'} \mathbf{X}^*)^{-1} \mathbf{X}^{*'} \boldsymbol{\eta}$ from which the estimate of \mathbf{h} can be obtained by adding back the zeroes where appropriate.

In step III, for a specific j , one can rewrite (7) as

$$\phi_j^{(M)} = \|\boldsymbol{\gamma}_j - \mathbf{Z}_j\|^2 + C_j^{(M)}, \quad (\text{A3})$$

where $\mathbf{Z}_j = \frac{1}{K} \sum_{k=1}^K \mathbf{Z}_{ki} \mathbf{w}_{ki}$, and $C_j^{(M)}$ is a constant unrelated to γ_j . Furthermore, (8) can be

rewritten as

$$\phi^{(S)} = \sum_{i=1}^J \|\boldsymbol{\gamma}_i - \mathbf{Y}_1^{(i)} - \mathbf{Y}_2^{(i)} - \mathbf{Y}_3^{(i)} - \mathbf{Y}_4^{(i)}\|^2, \quad (\text{A4})$$

where $\mathbf{B}_l = [b_{l,hi}]$, $\mathbf{D}_l = [\mathbf{d}_{l,1}, \dots, \mathbf{d}_{l,J}]$, $\mathbf{M}_l^{(p)} = [m_{l,hi}^{(p)}]$, $\mathbf{Q}_l^{(j)} = [q_{l,hi}^{(j)}]$, and $\mathbf{Y}_1^{(i)}$ through $\mathbf{Y}_4^{(i)}$

are given by

$$\mathbf{Y}_1^{(i)} = \mathbf{\Pi}_1^{(i,j)} \boldsymbol{\gamma}_j + \mathbf{\Omega}_1^{(i,j)}$$

with $\mathbf{\Pi}_1^{(i,j)} = \sum_{l=0}^L b_{l,ji} \mathbf{S}_l$ and $\mathbf{\Omega}_1^{(i,j)} = \sum_{l=0}^L \sum_{h \neq j} b_{l,hi} \mathbf{S}_l \boldsymbol{\gamma}_h$,

$$\mathbf{Y}_2^{(i)} = \mathbf{\Pi}_2^{(i,j)} \boldsymbol{\gamma}_j + \mathbf{\Omega}_2^{(i,j)}$$

with $\mathbf{\Pi}_2^{(i,j)} = 0$ and $\mathbf{\Omega}_2^{(i,j)} = \sum_{l=0}^L \mathbf{S}_l \mathbf{U} \mathbf{d}_{l,i}$,

$$\mathbf{Y}_3^{(i)} = \mathbf{\Pi}_3^{(i,j)} \boldsymbol{\gamma}_j + \mathbf{\Omega}_3^{(i,j)}$$

with $\mathbf{\Pi}_3^{(i,j)} = \sum_{l=0}^L \sum_{p=1}^P m_{l,ji}^{(p)} \mathbf{S}_l \text{diag}(\mathbf{u}_p)$ and $\mathbf{\Omega}_3^{(i,j)} = \sum_{l=0}^L \sum_{p=1}^P \sum_{h \neq j} m_{l,hi}^{(p)} \mathbf{S}_l \text{diag}(\mathbf{u}_p) \boldsymbol{\gamma}_h$, and

$$\mathbf{Y}_4^{(i)} = \mathbf{\Pi}_4^{(i,j)} \boldsymbol{\gamma}_j + \mathbf{\Omega}_4^{(i,j)}$$

with $\mathbf{\Pi}_4^{(i,j)} = \sum_{l=0}^L \mathbf{S}_l \left(\sum_{g=1}^j q_{l,ji}^{(g)} \text{diag}(\boldsymbol{\gamma}_g) + \sum_{h=j+1}^J q_{l,hi}^{(j)} \text{diag}(\boldsymbol{\gamma}_h) \right)$, and $\mathbf{\Omega}_4^{(i,j)} = \sum_{l=0}^L \sum_{\substack{1 \leq g < h \leq J \\ g \neq j, h \neq j}} q_{l,hi}^{(g)} \mathbf{S}_l \text{diag}(\boldsymbol{\gamma}_g) \boldsymbol{\gamma}_h$.

Define further

$$\mathbf{\Pi}_i = \begin{cases} \mathbf{I}_J - \sum_{c=1}^4 \mathbf{\Pi}_c^{(j,j)} & \text{for } i = j, \\ \sum_{c=1}^4 \mathbf{\Pi}_c^{(i,j)} & \text{otherwise,} \end{cases}$$

and

$$\mathbf{\Omega}_i = \begin{cases} \sum_{c=1}^4 \Omega_c^{(j,j)} & \text{if } i = j, \\ \gamma_i - \sum_{c=1}^4 \Omega_c^{(i,j)} & \text{otherwise.} \end{cases}$$

The estimates of γ_j is the solution of the optimization problem

$$\hat{\gamma}_j = \operatorname{argmin} \gamma' \mathbf{\Pi}' \mathbf{\Pi} \gamma - 2\gamma' \mathbf{\Pi}' \mathbf{\Omega} \quad \text{with respect to} \quad \frac{1}{T} \gamma' \gamma = 1, \quad (\text{A5})$$

where

$$\mathbf{\Pi} = [\sqrt{\alpha K} \mathbf{I}_J, \sqrt{\beta} \mathbf{\Pi}'_1, \dots, \sqrt{\beta} \mathbf{\Pi}'_J]',$$

and

$$\mathbf{\Omega} = [\sqrt{\alpha K} \mathbf{Z}'_j, \sqrt{\beta} \mathbf{\Omega}'_1, \dots, \sqrt{\beta} \mathbf{\Omega}'_J]'$$

This optimization problem is so-called Mosier's oblique Procrustes rotation, which has been solved by an algorithm proposed by ten Berge and Nevels (1977). We incorporate their algorithm in Dynamic GSCANO to sequentially update γ_j in step III.

An initial estimate of $\mathbf{\Gamma}$ is needed before getting into the iterative steps. It may be generated randomly or rationally. The latter can be obtained by applying GCANO to \mathbf{Z}_{kj} ($k = 1, \dots, K$) for each ROI ($j = 1, \dots, J$), temporarily ignoring the structural model. Each problem reduces to an eigen-decomposition of the sum of orthogonal projectors defined by \mathbf{Z}_{ki} over k (Gifi, 1991).

Appendix B. Assumed values of parameters in the model in Monte Carlo studies

(B1) Generation of \mathbf{C}_{00} , \mathbf{C}_1 and $\tilde{\Sigma}^{(s)}$

In both simulation studies, \mathbf{C}_{00} was generated as $\mathbf{K}'\mathbf{K}$ where \mathbf{K} is a square matrix of the same size as \mathbf{C}_{00} . The entries of \mathbf{K} were independently sampled from the standard normal distribution. Each column of \mathbf{K} was then normalized before $\mathbf{K}'\mathbf{K}$ was taken.

For \mathbf{C}_1 and $\tilde{\Sigma}^{(s)}$, we first decompose \mathbf{C}_{00} as $\mathbf{R}'\mathbf{V}\mathbf{R}$ (\mathbf{R} non-singular and \mathbf{V} diagonal). We then define $\mathbf{C}_{01} = E(\Gamma_t' \Gamma_{t-1})$ as $\mathbf{R}'\mathbf{V}^{1/2}\mathbf{Q}\mathbf{V}^{1/2}\mathbf{R}$, where \mathbf{Q} has the effect of shrinking the covariances in \mathbf{C}_{00} . We then obtain $\mathbf{C}_1 = \mathbf{C}_{00}^{-1}\mathbf{C}_{01}'$ (\mathbf{C}_1 being the matrix of regression coefficients of Γ_t onto Γ_{t-1} ; Note that $\mathbf{C}_{11} = E(\Gamma_{t-1}' \Gamma_{t-1}) = \mathbf{C}_{00}$), and $\tilde{\Sigma}^{(s)} = \mathbf{C}_{00} - \mathbf{C}_{01}\mathbf{C}_{00}^{-1}\mathbf{C}_{01}'$ ($\tilde{\Sigma}^{(s)}$ being the error covariance matrix in the regression of Γ_t onto Γ_{t-1}). In study 1, we define $\mathbf{C}_{00} = \mathbf{R}'\mathbf{V}\mathbf{R}$ as the singular value decomposition and $\mathbf{Q} = (\mathbf{I}_K - 0.5 * \mathbf{V}^{-1})^{1/2}$. For study 2, we specify $\mathbf{C}_{00} = \mathbf{R}'\mathbf{V}\mathbf{R}$ as the Cholesky decomposition (\mathbf{R} upper triangular and $\mathbf{V} = \mathbf{I}_K$) and $\mathbf{Q} = \text{diag}([0.5, -0.6, 0.6, -0.7, 0.7])$, from which the stimulus \mathbf{u}_t is constructed as an univariate time series and Γ_t as a multiple time series. This allows the stimulus record \mathbf{U} generated and then fixed across different simulated data sets prior to the generation of Γ . In principle, \mathbf{Q} could be any non-singular symmetric matrix with eigen-values strictly between -1 and 1. By knowledge of time series analysis (Lütkepohl, 2007), $E(\Gamma_t' \Gamma_t) = \sum_{i=0}^{\infty} \mathbf{C}_1^i \tilde{\Sigma}^{(s)} \mathbf{C}_1^i$. Under the above specifications, simple

calculation reveals that $E(\Gamma_t' \Gamma_t) = (\mathbf{R}' \mathbf{V}^{1/2}) \left(\left(\sum_{i=0}^{\infty} \mathbf{Q}^{2i} \right) (\mathbf{I}_K - \mathbf{Q}^2) \right) (\mathbf{V}^{1/2} \mathbf{R})$. Notice that $\sum_{i=0}^{\infty} \mathbf{Q}^{2i} = (\mathbf{I}_K - \mathbf{Q}^2)^{-1}$, leading to $E(\Gamma_t' \Gamma_t) = \mathbf{R}' \mathbf{V} \mathbf{R} = \mathbf{C}_{00}$.

(B2) Study 1:

The generated values of \mathbf{C}_{00} were given as follows:

$$\mathbf{C}_{00} = \begin{bmatrix} 1 & -0.0021 & 0.2133 & -0.0120 & 0.0411 & 0.0028 & 0.0708 \\ -0.0021 & 1 & -0.0091 & 0.1483 & -0.0189 & 0.044 & 0.0339 \\ 0.2133 & -0.0091 & 1 & 0.0295 & 0.2742 & 0.0431 & 0.1584 \\ -0.012 & 0.1483 & 0.0295 & 1 & 0.1119 & 0.2132 & -0.0203 \\ 0.0411 & -0.0189 & 0.2742 & 0.1119 & 1 & 0.1002 & 0.1813 \\ 0.0028 & 0.044 & 0.0431 & 0.2132 & 0.1002 & 1 & 0.0786 \\ 0.0708 & 0.0339 & 0.1584 & -0.0203 & 0.1813 & 0.0786 & 1 \end{bmatrix},$$

The covariance matrix of the measurement error ($\Sigma_{kj}^{(M)}$) was postulated as $\Sigma_{kj}^{(M)} = \sigma^2 \mathbf{I}_3$ for

$$v_i = 3, \text{ and } \Sigma_{kj}^{(M)} = \sigma^2 \begin{bmatrix} 1 & \rho & \dots & \rho^{49} \\ \rho & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \rho^{49} & \dots & \dots & 1 \end{bmatrix} \text{ for } v_j = 50, \text{ where } \rho \text{ is varied at two levels, } \{0, 0.9\}.$$

In all cases, σ^2 is set to 0.5 as in Jung et al. (2012). The case of $\rho = 0$ above is equivalent to $\Sigma_{kj}^{(M)} = \sigma^2 \mathbf{I}_{50}$, while $\rho = 0.9$ assumes that there are excess correlations among voxels left unaccounted for by γ_j . We further prescribed $\mathbf{c}_j = [0.7, 0.8, 0.9]$ for $v_i = 3$, and for $v_j = 50$, $\mathbf{c}_j = [0.7, 0.7041, 0.7082, \dots, 0.9]'$, a 50-element vector with elements equally spaced between 0.7 and 0.9. For each condition, 100 data sets were generated, and each data set was analyzed by Dynamic GSCANO.

(B3) Study 2:

In this study, the generated values of \mathbf{C}_{00} were given as:

$$\mathbf{C}_{00} = \begin{bmatrix} 1 & -0.0907 & 0.1314 & -0.0340 & -0.1711 \\ -0.0907 & 1 & -0.2058 & -0.0346 & 0.2177 \\ 0.1314 & -0.2058 & 1 & 0.1876 & -0.1553 \\ -0.0340 & -0.0346 & 0.1876 & 1 & 0.0805 \\ -0.1711 & 0.2177 & -0.1553 & 0.0805 & 1 \end{bmatrix},$$

For each level of T , one copy of \mathbf{U} was generated recursively as $U_t = 0.5U_{t-1} + \varepsilon_t$, where U_t denotes the t -th element of the vector \mathbf{U} and the measurement error ε_t is sampled from $N(0, 0.75)$. Once \mathbf{U} was generated, we fixed it across all simulation data sets within the same T .

Finally, the BOLD signal data were generated, using $\mathbf{c}_j = [0.7, 0.8, 0.9]$ and $\boldsymbol{\Sigma}_{kj}^{(M)} = \delta^2 \mathbf{I}_3$.

References

- Absil, P. A., Sepulchre, R., and Mahony, R. (2008). Continuous-time subspace flows related to the symmetric eigenvalue problem. *Pacific Journal of Optimization*, 4, 179-194.
- Berkowitz, D., Caner, M., & Fang, Y. (2012). The validity of instruments revisited. *Journal of Econometrics*, 166(2), 255-266.
- Baum, C. F. (2008). Using instrumental variables techniques in economics and finance. A slide presentation at German Stata Group Meetings, Berlin. <http://fmwww.ba.edu/repec/dsug2008/DESUG8621.beamer.pdf>
- Bowden, R. J., and Turkington, D. A. (1990). *Instrumental variables*. Cambridge: Cambridge University Press.

- Breiter, H. C., Aharon, I., Kahneman, D., Dale, A., and Shizgal, P. (2001). Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron*, *30*, 619-639.
- Büchel, C., and Friston, K. J. (1997). Modulation of connectivity in visual pathways by attention: cortical interactions evaluated with structural equation modeling and fMRI. *Cerebral Cortex*, *7*, 768-778.
- Bullmore, E., Horwitz, B., Honey, G., Brammer, M., Williams, S., and Sharma, T. (2000). How good is good enough in path analysis of fMRI data? *NeuroImage*, *11*, 289-301.
- Cacioppo, J. T., and Decety, J. (2009). What are the brain mechanisms on which psychological processes are based? *Perspectives on Psychological Science*, *4*, 10-18.
- Carroll, J. D. (1968). Generalization of canonical correlation analysis to three or more sets of variables. In *Proceedings of the 76th Annual convention of the American Psychological Association* *3*, 227-228.
- De Leeuw, J., Young, F. W., and Takane, Y. (1976). Additive structure in qualitative data: An alternating least squares method with optimal scaling features. *Psychometrika*, *41*, 471-514.
- Efron, B. (1982). *The jackknife, the bootstrap and other resampling plans*. Philadelphia: SIAM.
- Efron, B., and Tibshirani, R. J. (1993). *An introduction to bootstrap*. Boca Raton, FL: Chapman and Hall/CRC Press.

- Eklund, A., Andersson, M., Josephson, C., Johansson, M., and Knutsson, H. (2012). Does parametric fMRI analysis with SPM yield valid results? An empirical study of 1484 rest datasets. *NeuroImage*, *61*, 565-578.
- Epstein, R., Harris, A., Stanley, D., and Kanwisher, N. G. (1999). The parahippocampal place area: Recognition, navigation, or encoding? *Neuron*, *23*, 115-125.
- Epstein, R., and Kanwisher, N. G. (1998). A cortical representation of the local visual environment. *Nature*, *392*, 598-601.
- Friston, K. J. (1994). Functional and effective connectivity in neuroimaging: A synthesis. *Human Brain Mapping*, *2*, 56-78.
- Fu, Z., Di, X., Chan, S. C., Hung, Y. S., Biswal, B. B., and Zhang, Z. (2013). Time-varying correlation coefficients estimation and its application to dynamic connectivity analysis of fMRI. *Engineering in Medicine and Biology Society (EMBC) 35th Annual International Conference of IEEE*, 2944-2947.
- Gates, K. M., and Molenaar, P. (2012). Group search algorithm recovers effective connectivity maps for individuals in homogeneous and heterogeneous samples. *NeuroImage*, *63*, 310-319.
- Gates, K. M., Molenaar, P., Hillary, F. G., and Slobounov, S. (2011). Extended unified SEM approach for modeling event-related fMRI data. *NeuroImage*, *54*, 1151-1158.
- Gifi, A. (1991). *Nonlinear multivariate analysis*. Chichester: Wiley.
- Greene, W. H. (2011). *Econometric analysis (7th Edition)*. New York: Prentice Hall.

- Grill-Spector, K., Kourtzi, Z., and Kanwisher, N. G. (2001). The lateral occipital complex and its role in object recognition. *Visual Research*, 41, 1409-1422.
- Harville, D. A. (1997). *Matrix Algebra from a Statistician's Perspective*. New York: Springer
- Hathaway, R. J., and Bezdek, J. C. (1993). Switching regression models and fuzzy clustering. *IEEE Transactions on Fuzzy Systems*, 1, 195-204.
- Hayashi, F. (2000). *Econometrics*. Princeton, NJ: Princeton University Press.
- Haynes, J. d., Driver, J., and Rees, G. (2005). Visibility reflects dynamic changes of effective connectivity between V1 and fusiform cortex. *Neuron*, 46, 811-821.
- Hotelling, H. (1936). Relations between two sets of variates. *Biometrika*, 28, 321-377.
- Huettel, S. A., Stowe, C. J., Gordon, E. M., Warner, B. T., and Platt, M. L. (2006). Neural Signatures of Economic Preferences for Risk and Ambiguity, *Neuron*, 49, 765-775.
- Hwang, H., and Takane, Y. (2004). Generalized structured component analysis. *Psychometrika*, 69, 81-99.
- Hwang, H., and Takane, Y. (2014). *Generalized structured component analysis: A component-based approach to structural equation modeling*. Boca Raton, FL: Chapman and Hall/CRC Press.
- Hwang, H., Desarbo, W. S., and Takane, Y. (2007). Fuzzy clusterwise generalized structured component analysis. *Psychometrika*, 72, 181-198.

- Hwang, H., Ho, M. H. R., and Lee, J. (2010). Generalized structured component analysis with latent interactions. *Psychometrika*, 75, 228-242.
- Hwang, H., Jung, K., Takane, Y., and Woodward, T. S. (2013). A unified approach to multiple-set canonical correlation analysis and principal components analysis. *British Journal of Mathematical and Statistical Psychology*, 66, 308-321.
- Jung, K., Takane, Y., Hwang, H., and 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.
- Kanwisher, N. G., McDermott, J., and Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, 17, 4302-4311.
- Kenny, D. A., and Judd, C. M. (1984). Estimating the nonlinear and interactive effects of latent variables. *Psychological Bulletin*, 96, 201-210.
- Luchtman, M., Jachau, K., Adolf, D., Baecke, S., Lützkendorf, R., Müller, C., and Kim, J. S., Jung, W. H., Kang, D. H., Park, J. Y., Jang, J. H., Choi, J. S., and Kwon, J. S. (2012). Changes in Effective Connectivity According to Working Memory Load: An fMRI Study of Face and Location Working Memory Tasks. *Psychiatry Investigation*, 9, 283-292.
- Lütkepohl, H. (2007). *New introduction to multivariate time series analysis*. Berlin and Heidelberg: Springer.
- McIntosh, A. R., and Gonzalez-Lima, F. (1994). Structural equation modeling and its application to network analysis in functional brain imaging. *Human Brain Mapping*, 2, 2-22.

- Mechelli, A., Penny, W. D., Price, C. J., Gitelman, D. R., and Friston, K. J. (2002). Effective connectivity and intersubject variability: using a multisubject network to test differences and commonalities. *NeuroImage*, *17*, 1459-1469.
- Meng, M., Cherian, T., Singal, G., and Sinha, P. (2012). Lateralization of face processing in the human brain. *Proceedings of the Royal Society B: Biological Sciences*, *279*, 2052-2061.
- Metzak, P. D., Riley, J. D., Wang, L., Whitman, J. C., Ngan, E. T., and Woodward, T. S. (2011). Decreased efficiency of task-positive and task-negative networks during working memory in schizophrenia. *Schizophrenia Bulletin*, *38*, 803-813.
- Mulaik, S. A. (1972). *The foundations of factor analysis*. New York: McGraw-Hill.
- Penny, W. D., Stephan, K. E., Mechelli, A., and Friston, K. J. (2004). Modeling functional integration: a comparison of structural equation and dynamic causal models. *NeuroImage*, *23*, 264-274.
- Stoeckel, L. E., Kim, J., Weller, R. E., Cox, J. E., Cook III, E. W., and Horwitz, B. (2009). Effective connectivity of a reward network in obese women. *Brain Research Bulletin*, *79*, 388-395.
- Takane, Y., and Hunter, M. A. (2001). Constrained principal component analysis: A comprehensive theory. *Applicable Algebra in Engineering, Communication, and Computing*, *12*, 391-419.
- ten Berge, J. M., and Nevels, K. (1977). A general solution to Moisiier's oblique procrustes problem. *Psychometrika*, *42*, 593-600.

Zhou, L. (2013). Dynamic generalized (multiple-set) structured canonical correlation analysis (Dynamic GCANO): A structural equation model for the analysis of multiple-subject effective connectivity in functional neuroimaging studies. *Unpublished Ph.D. Dissertation, McGill University.*

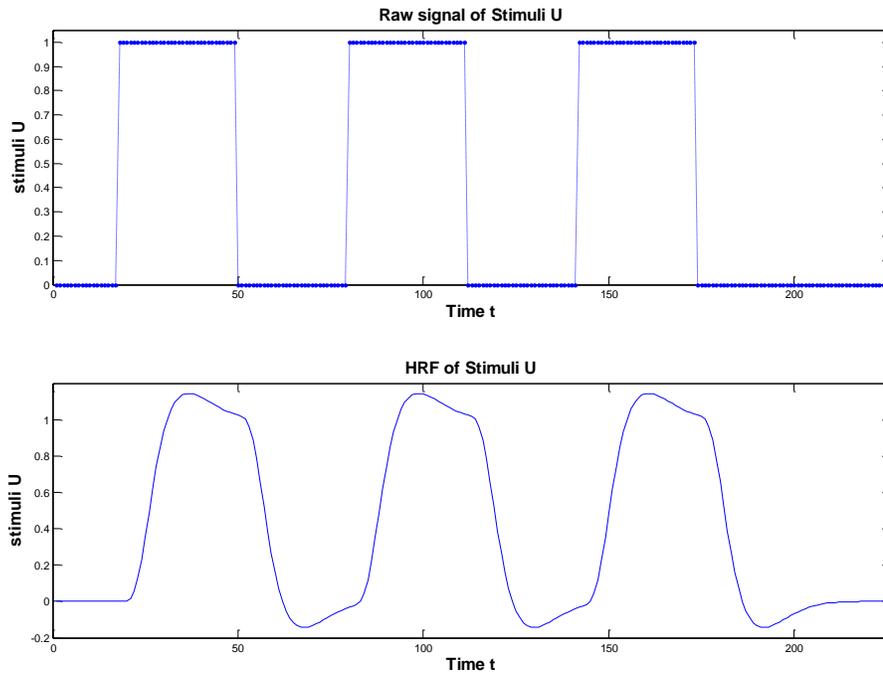


Figure 1. The time series plot of the stimulus input (U) in the first empirical data set: The raw signal (the upper panel) and its impact on dynamics of fMRI obtained by convoluting the raw signal with the hemodynamic response function (the bottom panel).

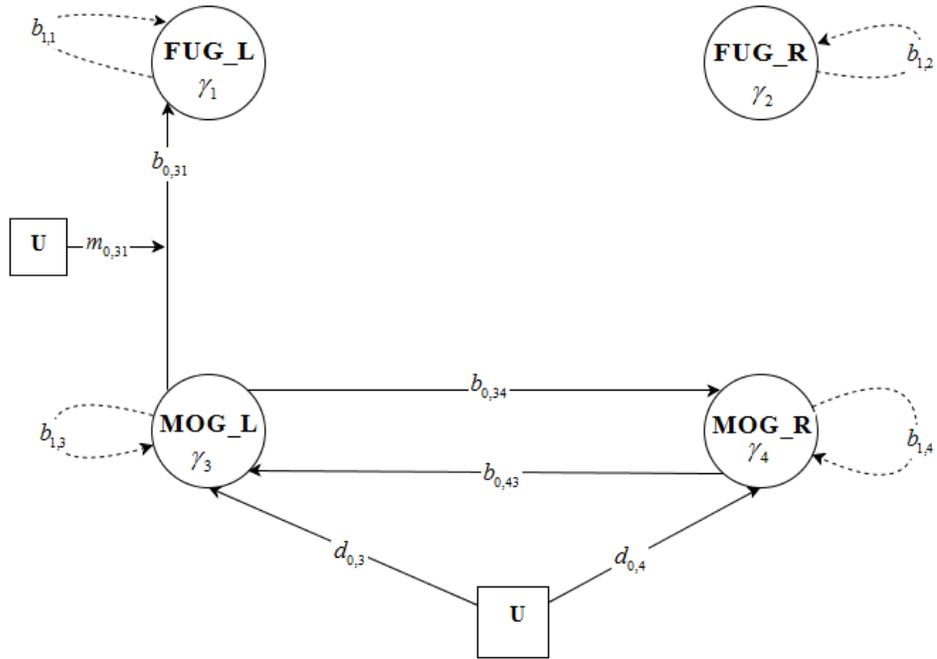


Figure 2. The significant connections found by Dynamic GSCANO for the first empirical data set.

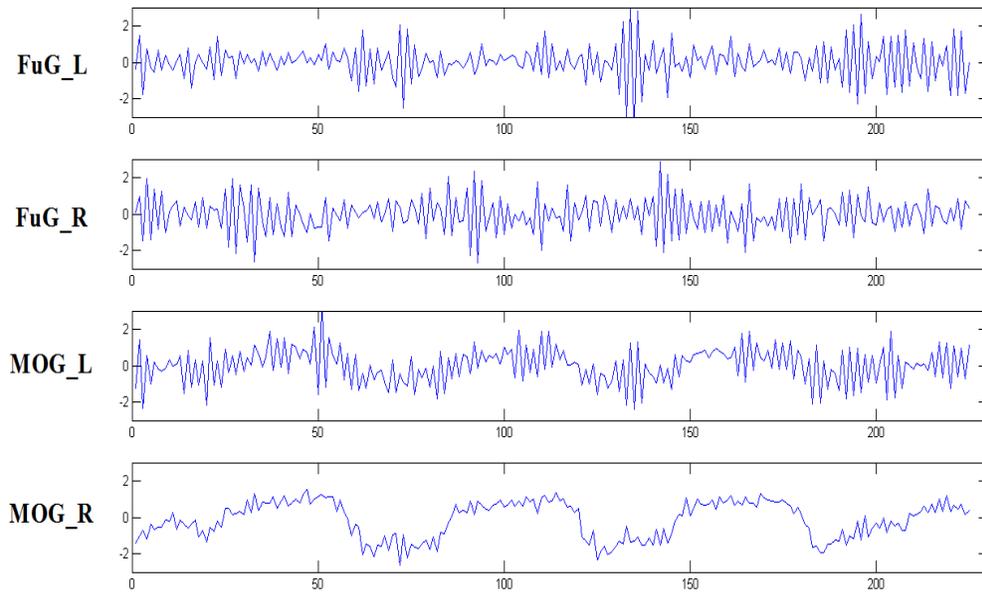


Figure 3. The time series plots of the most representative activations at the four ROIs derived from the first empirical data set.

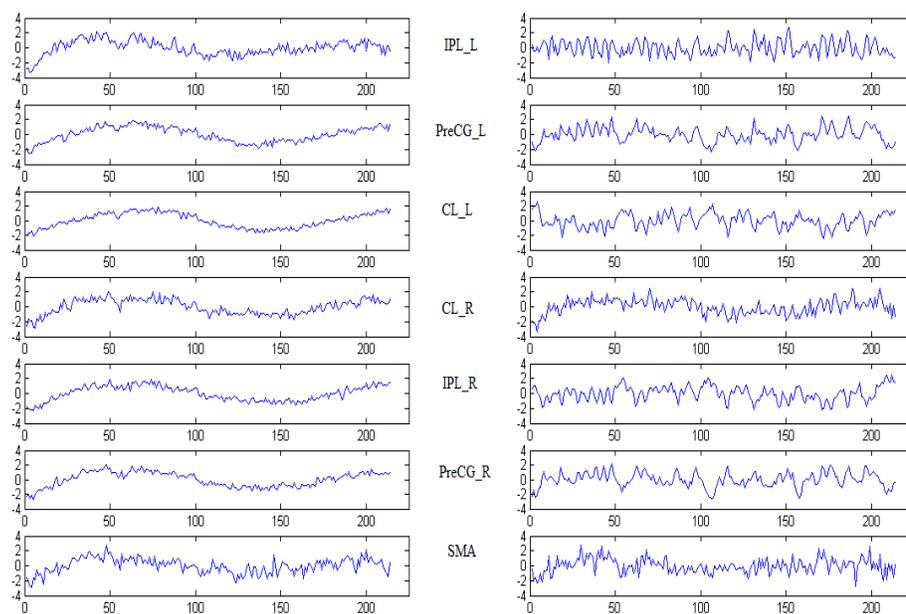


Figure 4. The time series plots of the most representative activations at the seven ROIs extracted from the second empirical data set. The normal control group is in the left column, and the schizophrenia group in the right column. The ROIs are, from the top to the bottom, IPL_L, PreCG_L, CL_L, CL_R, IPL_R, PreCG_R and SMA.