ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 41 (2008) 398–407

Dynamic Bayesian network modeling of fMRI: A comparison of group-analysis methods

Junning Li,[a] Z. Jane Wang,[a,*] Samantha J. Palmer,[b] and Martin J. McKeown[b]

[a]*Department of Electrical and Computer Engineering, University of British Columbia, Canada*
[b]*Pacific Parkinson's Research Centre, Brain Research Centre, University of British Columbia, Canada*

Received 4 August 2007; revised 13 January 2008; accepted 24 January 2008
Available online 10 March 2008

Bayesian network (BN) modeling has recently been introduced as a tool for determining the dependencies between brain regions from functional-magnetic-resonance-imaging (fMRI) data. However, studies to date have yet to explore the optimum way for meaningfully combining individually determined BN models to make group inferences. We contrasted the results from three broad approaches: the "virtual-typical- subject" (VTS) approach which pools or averages group data as if they are sampled from a single, hypothetical virtual typical subject; the "individual-structure" (IS) approach that learns a separate BN for each subject, and then finds commonality across the individual structures, and the "common-structure" (CS) approach that imposes the same network structure on the BN of every subject, but allows the parameters to differ across subjects. To explore the effects of these three approaches, we applied them to an fMRI study exploring the motor effect of L-dopa medication on ten subjects with Parkinson's disease (PD), as the profound clinical effects of this medication suggest that fMRI activation in PD subjects after medication should start approaching that of age-matched controls. We found that none of these approaches is generally superior over the others, according to Bayesian-information-criterion (BIC) scores, and that they led to considerably different group-level results. The IS approach was more sensitive to the normalization effect of the L-dopa medication on brain connectivity. However, for the more homogeneous control population, the VTS approach was superior. Group-analysis approaches should be selected carefully with consideration of both statistical and biomedical evidence.

Keywords: Functional magnetic resonance imaging (fMRI); Dynamic Bayesian network; Group analysis; Parkinson's disease

Introduction

Effective brain connectivity, defined as the neural influence that one brain region exerts over another (Friston, 1994; Harrison and Friston, 2004), is important for the assessment of normal brain function, and its impairment is associated with neurodegenerative diseases such as Alzheimer's or Parkinson's disease (PD). Various mathematical methods, such as structural equation modeling (SEM) (McIntosh and Gonzalez-Lima, 1994), multivariate autoregressive modeling (MAR) (Harrison et al., 2003), dynamic causal modeling (DCM) (Friston et al., 2003), bilinear dynamical system (BDS) (Penny et al., 2005) and Bayesian networks (BN) (Zheng and Rajapakse, 2006; Rajapakse and Zhou, 2007), have been proposed for inferring effective connectivity from functional-magnetic-resonance-imaging (fMRI). These models can all be visualized as a graph whose nodes denote brain regions and directed edges denote connections between the regions. Connection parameters are associated with edges, indicating the strength of the connections. MAR, SEM and BNs are ordinary graphical models, but with different constraints on the network structures, regarding whether time lags are considered and whether cycles are allowed on the graph. DCM and BDS add a hidden layer to the graph for the unobserved underlying neural activities and also allow the interaction between two regions to be modulated by the activities of a third region, at the cost of intensive computation.

fMRI experiments are usually performed to infer brain activations consistently shared by a population or to identify their differences between populations. Therefore, it is important to develop group-analysis methods for the aforementioned graphical models if they are to be more fully adopted in group analysis rather than just being used at the individual level (Rajapakse and Zhou, 2007). Here we investigate potential group-analysis approaches, using BNs as a prototype of graphical models. We are especially interested in BN modeling of brain connectivity because it is flexible to handle both categorical and continuous data and as well both linear and nonlinear relationships (Lauritzen, 1996a) and also because plenty of methods about model learning and computation have been developed for BNs in the field of artificial intelligence.

* Corresponding author.
E-mail address: zjanew@ece.ubc.ca (Z.J. Wang).
Available online on ScienceDirect (www.sciencedirect.com).

1053-8119/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2008.01.068

Group analyses of any sort fundamentally involve two factors: common features shared by group members, and the specific features of each individual. The common features seen within a group are typically of ultimate interest, although individually specific features cannot simply be ignored, as they may indicate the presence of outliers, potentially biasing the overall group result. Many studies have found that fMRI activation patterns may considerably differ among subjects even within the same group. For example, Sugihara et al. (2006) found three of the five regions related to writing were inconsistently activated among subjects; in Vandenbroucke et al. (2004)'s study, eighteen of twenty nine subjects showed activation varied in both location and size. These results suggest that the similarity and diversity among group members should be balanced in group models, rather than either over-emphasized or neglected.

There have been several group-analysis techniques implicitly employed in graphical modeling of fMRI data, based on models such as SEM, MAR and BNs. A review of the literature reveals that these techniques can be divided into three broad categories, namely the "virtual-typical-subject" (VTS) approach, the "individual-structure" (IS) approach and the "common-structure" (CS) approach. The VTS approach assumes that every subject within a group performs the same function with exactly the same connectivity network, and it does not generally accommodate inter-subject variability. These methods reconstruct a virtual typical subject to represent the whole group, by pooling or averaging group data as if they were sampled from a single virtual subject (Goncalves et al., 2001; Zheng and Rajapakse, 2006; Rajapakse and Zhou, 2007), and then learn the connectivity network of the virtual typical subject. When subjects are homogeneous within a group or the inter-subject variability follows certain regular distributions such as the Gaussian distribution, the VTS approach could increase sensitivity, because pooling can yield a relatively large data set, and averaging can enhance the signal-to-noise ratio.

At the other extreme, the "individual-structure" (IS) approach learns a network for each subject separately and then performs group analysis on the individually learned networks (Goncalves et al., 2001; Li et al., 2007). The IS approach is consistent with the concept of functional degeneracy, i.e., "the ability of elements that are structurally different to perform the same function or yield the same output" (Edelman and Gally, 2001), or more plainly, "there are multiple ways of completing the same task" (Price and Friston, 2002). The IS approach certainly considers inter-subject variability but may not integrate group data tightly enough to enable correct inferences about statistically significant differences between groups.

The "common-structure" (CS) approach is a trade-off between the two extremes, imposing the same network structure on the statistical graphical models of every subject, while allowing the parameters of the models to differ across subjects (Mechelli et al., 2002; Kim et al., 2007). The CS approach assumes that cognitive functions invoke a similar connectivity pattern for every subject, but the exact details of the connectivity patterns, in terms of connectivity strength (coefficients), differ across subjects. Thus, the CS approach addresses group similarity at the structural level and inter-subject variability at the parameter level.

We investigated the three approaches using data collected from subjects with Parkinson's disease (PD), as the consistent, dramatic effects of L-dopa medication seen in this population provide an additional qualitative "ground truth" to the effects of group inferences. One of the cardinal features of PD that is improved with L-dopa is bradykinesia, a slowness of performed voluntary movements, that represents a major source of disability in PD and is related to impairments in daily activities such as walking and writing. Impaired ramping of force may be fundamentally related to the clinical feature of bradykinesia (Stelmach et al., 1989). Similarly, several studies have demonstrated disturbances of rhythmic movement in PD (Pope et al., 2006) that are likely related to bradykinesia.

However, the bradykinesia is thus far not fully explained in research. A recent study used PET imaging to infer brain areas that appear related to bradykinesia (Turner et al., 2003), such as underactivity in the sensorimotor cortex contralateral to the moving arm, bilateral dorsal premotor cortices, and ipsilateral cerebellum. However, the PET modality did not enable inference about the interaction between these regions. A local problem in the basal ganglia circuit in Parkinson's disease may cause disruption of downstream distributed motor control networks. A greater understanding of the relative contribution of neural regions to bradykinesia in healthy controls and patients may be of diagnostic and therapeutic significance for patients with Parkinson's disease.

The main reason that we chose PD as an example for assessing group analyses is that the standard medication used for this condition, L-dopa, has dramatic effects against bradykinesia and rigidity (although less effect against tremor, balance and gait). The effects of L-dopa on idiopathic disease are so sufficiently dramatic that lack of response to this medication makes the diagnosis of Parkinson's disease questionable (Jankovic and Tolosa, 2002). Thus after introduction of the L-dopa medication, we would the connectivity pattern of PD subjects when perform a paradigm assessing bradykinesia to approach that of normal subjects.

In this paper, we investigate the performances of the three group-analysis approaches (VTS, IS, CS) by applying them to an fMRI study on Parkinson's disease. Broadly speaking, we attempted to answer three fundamental questions: "which approach most accurately reflects the underlying biomedical behavior?" "Do the approaches lead to considerably different analysis results?" "How can the suitable approach be selected?" The three approaches are compared from the aspects of their statistical goodness-of-fit to the data, and more importantly their sensitivity in detecting the effect of the L-dopa medication on the disease. To the best of our knowledge, this is the first study specifically devoted to group-analysis on fMRI with BN modeling.

Materials and methods

fMRI data

The fMRI data were collected from ten healthy people and ten Parkinson's disease (PD) patients, each of whom was asked to squeeze a rubber bulb at three different speeds or at a constant force, as cued by visual instruction. The patients performed the entire task twice, once before and another after the L-dopa medication which is most effective against slowness of movement and rigidity. Six regions were selected for the analysis according to the Talairach atlas (Talairach and Tournoux, 1988): the left and right primary motor cortex (M1) (Brodman area 4), supplementary motor cortex (SMA) (Brodman area 6), and lateral cerebellar hemispheres (CER). fMRI time courses were collected at the sampling frequency of 0.503 Hz for 4 min and 18 s, in total of 130 time points.

Subjects

The study was approved by the University of British Columbia Ethics Board. Subjects gave written informed consent prior to participating. Ten volunteers with clinically diagnosed PD participated in the study (4 men, 6 women, mean age 66±8 years, 8 right-

handed, 2 left-handed). All patients had mild to moderate PD (Hoehn and Yahr stages 2–3) (Hoehn and Yahr, 1967) with mean symptom duration of 5.8±3 years. Exclusion criteria included atypical Parkinsonism, presence of other neurological or psychiatric conditions and use of antidepressants, sleeping tablets, or dopamine blocking agents.

All patients were taking L-dopa with an average daily dose of 685±231 mg. We also recruited ten healthy, age-matched control subjects without active neurological disorders (3 men, 7 women, mean age 57.4±14 years, 9 right-handed, 1 left-handed). All patients stopped their anti-Parkinson medications overnight for a minimum of 12 h before the study. The mean Unified Parkinson Disease Rating Scale (UPDRS) motor score during this off-L-dopa state was 26±8. There were no significant correlations between UPDRS motor scores and age. All patients exhibited some aspects of bradykinesia on examination.

After completing the experiment in an off-medication state, patients were given the equivalent to their usual morning dose of L-dopa in immediate release form (mean 125±35.3 mg L-dopa). They then repeated the same tasks post-medication following an interval of approximately 1 h to allow L-dopa to reach peak dose.

Experimental task

Subjects were instructed to lie on their back in the functional magnetic resonance scanner viewing a computer screen via a projection mirror system. All subjects used an in-house designed response device in their left hand, which was a custom-built MR-compatible rubber squeeze bulb connected to a pressure transducer outside the scanner room. They lay with their forearm resting down in a stable position, and were instructed to squeeze the bulb using an isometric hand grip and to keep their grip constant throughout the study. Each subject had their maximum voluntary contraction (MVC) measured at the start of the experiment and all subsequent movements were scaled to this, so that they had to squeeze at 5–15% of maximal force to accomplish the task. Using the squeeze bulb, subjects were required to control the width of an inflatable ring (shown as a black horizontal bar on the screen) in order to keep the ring within an undulating pathway without scraping the sides. Applying greater pressure to the bulb increased the width of the bar, and releasing pressure from the bulb decreased the width of the bar. To avoid scraping the sides of the tunnel, the required pressure was between 5% and 15% MVC. No additional visual feedback or error reporting was given when subjects went outside the white target lines so subjects had to monitor their own performance carefully. The pathway used a block design with sinusoidal sections in three different frequencies (0.25, 0.5 and 0.75 Hz) in a 25 pseudo-random order, and straight parts in between where the subjects had to keep a constant force of 10% of MVC. The frequencies were chosen based on prior findings and pilot studies were used to determine that PD patients could comfortably perform the required task. Each block lasted 19.85 s (exactly 10 TR intervals), alternating a sinusoid, constant force, sinusoid and so on (Fig. 1) to a total of 4 min and 18 s.

PD patients performed the task both after an overnight withdrawal (minimum of 12 h since last dose of L-dopa) of their anti-Parkinson drugs and repeated the same series 1 h after admission of L-dopa. Before the first scanning session, subjects practiced the task at each frequency until errors stabilized and they were familiar with the task requirements. Custom Matlab software (Mathworks) and the Psychtoolbox (Brainard, 1997; Pelli, 1997) was used to design and present stimuli, and to collect behavioral data from the response devices.

fMRI acquisition

Functional MRI was conducted on a Philips Achieva 3.0-T scanner (Philips, Best, The Netherlands) equipped with a head-coil. We collected echo-planar (EPI) T2*-weighted images with blood oxygenation level-dependent (BOLD) contrast. Scanning parameters were: repetition time 1985 ms, echo time 3.7, flip angle 90°, field of view (FOV) [216.00 143.00 240.00] mm, matrix size=128×128, pixel size 1.9×1.9 mm. Each functional run lasted 4 min and 18 s. Thirty-six axial slices of 3 mm thickness were collected in each volume, with a gap thickness of 1 mm. We selected slices to cover the dorsal surface of the brain and include the cerebellum ventrally. A high-resolution, three-dimensional T1-weighted image consisting of 170 axial slices was acquired of the whole brain to facilitate anatomical localization of activation for each subject. Head motion was minimized by a foam pillow placed around the subjects head within the coil. Subjects also used ear plugs to minimize the noise of the scanner. The subjects constantly viewed visual stimuli on a screen through a mirror built into the head coil.

Preprocessing

The functional MRI data were preprocessed for each subject, using Brain Voyager trilinear interpolation for 3D motion correction and sinc interpolation for slice timing correction. No temporal or spatial smoothing was performed on the data. The data were then further motion corrected with Motion Corrected Independent Component Analysis (MCICA), a computationally expensive but highly accurate method for motion correction (Liao et al., 2005, 2006). The

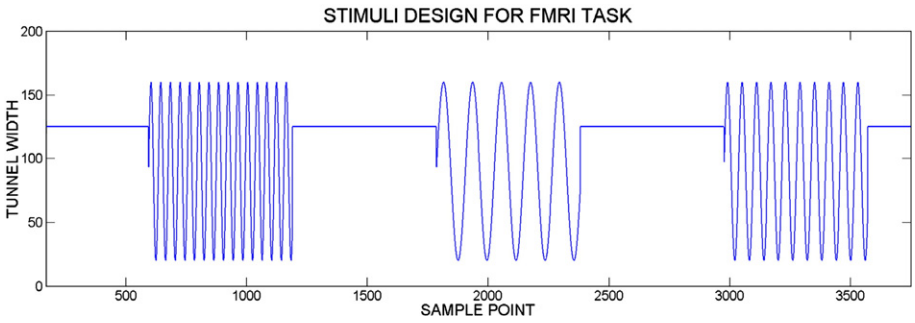

Fig. 1. Graph of the first half of the computer task subjects had to perform. Subjects were either asked to statically squeeze at 10% maximum voluntary contraction (MVC) or squeeze sinusoidally (between 5% and 15% of MVC) at different frequencies in 20 s blocks. All frequencies were performed twice per session.

Brain Extraction Tool in MRIcro (Rorden and Brett, 2000) (http://www.sph.sc.edu/comd/rorden/mricro.html) was used to strip the skull off the anatomical and first functional image from each run, to enable a more accurate alignment of the functional and anatomical scans. Custom scripts in Amira software (Amira 3D Visualization and Volume modeling) were used to co-register the anatomical and functional images. The following ROIs were drawn separately in each hemisphere, based upon anatomical landmarks and guided by the Talairach atlas (Talairach and Tournoux, 1988): primary motor cortex (M1) (Brodman area 4), supplementary motor cortex (SMA) (Brodman area 6), and lateral cerebellar hemispheres (CER). The labels on the segmented anatomical scans were resliced at the fMRI resolution. The raw time courses of the voxels within each ROI were averaged as the overall activity of the ROI, and the averaged time course were then detrended and normalized to unit variance before input into the BN models.

Dynamic Bayesian networks

A dynamic Bayesian network (DBN) (Murphy, 2002) is a graphical model for stochastic processes. The term "dynamic" does not mean that the model itself changes over time but that the Bayesian network (BN) models a dynamic system. As an extension of BN, a DBN follows the same rules as a regular BN does, encoding the conditional independence/dependence relationships among random variables with a directed acyclic graph (DAG). If a random variable x_a directly depends on another random variable x_b, i.e., x_a still depends on x_b even given all the other random variables, then the dependence is encoded as an edge between nodes a and b in the DAG (see Fig. 2 for an example).[1] In this text, nodes a and b are associated with x_a and x_b, respectively. When fMRI signals of ROIs are modeled with BNs, an edge in the DAG implies that two ROIs interact with each other even after removing the influence from all the other ROIs. Regarding the Bayes rule, the joint probability (density) function of all the random variables can be factorized according to the DAG as:

$$f(x) = \prod_{a \in W} f\left(x_a | x_{\mathrm{pa}[a]}\right), \tag{1}$$

where W denotes the set of all the nodes and pa[a] denotes the set of node a's parent node(s).

A multi-channel stochastic process can be modeled with a BN of $C \times T$ nodes, where C denotes the number of channels, T denotes the number of time points, and each node represents the signal of a channel at one time point. Because the future can influence neither the presence nor the past, nodes at time $t+1$ can have only nodes after time t as their children. BNs that account for all time points will become intractably large, so first-order Markov and stationary assumptions are usually applied. In this case, the same dependence relationships repeat time after time and signals at t only depend on signals from $t-1$ to t, so the whole network can be "rolled up" as its DBN representation, a DAG composed of only nodes from $t-1$ to t, as Fig. 2 illustrates.

[1] BNs graphically encode conditional independence theoretically according to Markov properties: if two sets of nodes A and B are d-separated by a third set of nodes C according to the DAG, and the three sets of nodes are disjoint, then x_A and x_B are conditionally independent given x_C (Lauritzen, 1996b), i.e., $P(x_A x_B | x_C) = P(x_A | x_C) P(x_B | x_C)$. The d-separation is a complex concept. For its exact definition, please refer to Lauritzen (1996b).

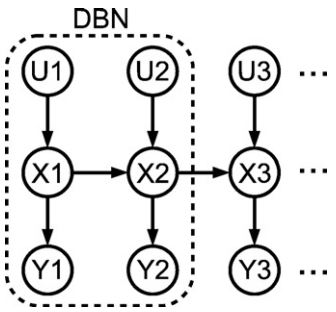

Fig. 2. An example of dynamic Bayesian networks. This DBN is a first-order Kalman filter process. Arrows from X_{t-1} and U_t to X_t (t=2, 3, …) mean that X_t depends on U_t and X_{t-1}, and are associated with the transition distribution $f(X_t|X_{t-1}, U_t)$ which varies according to the input U_t. Arrows from X_t to Y_t (t=1, 2, …) mean that the output Y_t depends on X_t, and is associated with $f(Y_t|X_t)$. As the same dependence relationships repeat, the process can be represented by just the first two time slices circled by dots. The joint probability density function of the process is $f(X, Y, U) = f(U_1) f(X_1|U_1) f(Y_1|X_1) \Pi^T_{t=2} f(U_t) f(X_t|X_{t-1}U_t) f(Y_t|X_t)$ where T is the number of time points.

We modeled fMRI signals with first-order Gaussian DBNs by regarding the signals of ROIs as a multi-channel stochastic Gaussian process. Gaussian DBNs whose conditional probability distributions are all Gaussian are limited to modeling linear relationships between random variables. Discrete DBNs are capable of modeling nonlinear relationships, but at the cost of precision and accuracy. Gaussian DBNs whose parameters explicitly indicate the strength of interactions between ROIs also facilitate the cross-group quantitative comparison. To accommodate nonlinear responses of brain activities to the task frequencies, we represented the bulb-squeezing frequency level (see fMRI data section) as an extra categorical input node L in the network. If node L is a parent of node a, x_a is regressed to other parent variables conditionally on the frequency level x_L, as in Eq. (2)

$$x_{a=} \sum_{b \in \mathrm{pa}[a] \setminus \{L\}} \beta_{ba|l} x_b + \varepsilon_{a|l}, \tag{2}$$

where x_a and x_b are the signals of node a and b, respectively, $\beta_{ba|l}$ is the connection coefficient from node b to node a conditional on the frequency level $x_L = l$, and $\varepsilon_{a|l}$ which follows a zero-mean Gaussian distribution is the regression error conditional on $x_L = l$. If node L is not a parent node of a, $\beta_{ba|l}$ takes the same value at all the four frequency levels. If there is no arrow from b to a, the connection coefficients are set to zero, as $\beta_{ba|l} = 0$.

We imposed on the DBN structures a constraint that there must be at least one path from the input node L to any ROI node. The constraint was imposed to ensure that the response of the ROIs to the bulb-squeezing task was accommodated. It is worth mentioning that in preliminary studies, we had relaxed the constraint so that there must be one path from the input node to at least one ROI, but all the structures learned under the relaxed constraint satisfied the more rigorous one.

Learning DBNs

Because many different DBNs may fit the same data almost equally well, we mixed them together according to their posterior probabilities to provide a more stable approximation to the data

than a single best DBN does. The mixture formula is $M^* = \sum P(M_i|X) M_i$ where M^* is the mixed model, M_i is a possible DBN and $P(M_i|X)$ is its posterior probability given data X. We used its mixed structure and its mixed coefficients to profile the mixture model which was composed of thousands of DBNs. The mixed structure is encoded as a matrix $G^* = \{g^*_{ab}\}$ whose element g^*_{ab} is the posterior probability of the existence of the connection from a to b, defined as Eq. (3)

$$g^*_{ab} = \sum_i P(M_i|X) g_{ab}(i), \tag{3}$$

where the binary variable $g_{ab}(i)$ indicates whether the connection from a to b appears in model M_i or not. The mixed coefficients $\mathcal{B}^* = \{\beta^*_{ab|l}\}$ are the posterior expected values of the connection coefficients, defined as Eq. (4)

$$\beta^*_{ab|l} = \sum_i P(M_i|X) \beta_{ab|l}(i), \tag{4}$$

where $\beta_{ab|l}(i)$ is the coefficient of model M_i.

The mixed structure and the mixed coefficients were estimated with the Bayesian information criterion (BIC) scores (Schwarz, 1978) and Markov chain Monte Carlo (MCMC) (Metropolis et al., 1953). The BIC score is defined as Eq. (5), where N denotes the sample size of data X, and K denotes the number of free parameters θ of DBN model M.

$$\mathrm{BIC}(M|X) = \sup_\theta \ln P(X|M,\theta) - 0.5K\ln N \tag{5}$$

If two models M_1 and M_2 have the same prior probabilities, i.e., $P(M_1) = P(M_2)$ and the prior distributions of their parameters θ are uniform distributions, then the ratio of their posterior probabilities can be asymptotically approximated as $\exp[\mathrm{BIC}(M_1) - \mathrm{BIC}(M_2)]$.[2] In our implementation, 1500 different DBN structures were sampled with MCMC, according to their relative posterior probabilities, i.e., $\exp[\mathrm{BIC}(M_i|X)]$ which was calculated when M_i was sampled, and then all the sampled DBNs were averaged according to their appearance frequencies in MCMC, as the estimation of the mixture model M^*.

Based on the assumption that subjects perform the task with different patterns of brain connectivity, the IS approach learns a mixture DBN model individually for each subject. Thus, each subject is associated with its own G^* and $\mathcal{B}^*$. The group mean structure (GMS) $\bar{G}$ and the group mean coefficients (GMC) $\bar{\mathcal{B}}$ are the mean of G^* and the mean of $\mathcal{B}^*$ over all the subjects in the same group, respectively. The group-level BIC score of applying a combination of subject-specific DBNs $\mathcal{M} = \{M_1, \ldots, M_S\}$ (where S is the number of subjects) respectively to the data of each subject is:

$$\begin{aligned}\mathrm{BIC}(\mathcal{M}|X_1, \ldots, X_S) &= \sum_{s=1}^{S} \mathrm{BIC}(M_s|X_s) \\ &= \sum_{s=1}^{S} \left[\sup_{\theta s} \ln P(X_s|M_s,\theta_s) - 0.5K_s \ln N_s\right],\end{aligned} \tag{6}$$

where all the notations are similar to those in Eq. (5) besides the subscript s, indicating that they are associated with subject s.

Based on the assumption that subjects perform the task with the same pattern of brain connectivity but differently in the details of the interaction between brain regions, the CS approach applies DBN models with the same network structure to the data of every subject in the same group, but optimizes the connection coefficients individually for each subject. Every subject is associated with the same G^*, and each subject is associated with its own $\mathcal{B}^*$. The GMS $\bar{G}$ equals the G^* and the GMC $\bar{\mathcal{B}}$ is the mean of $\mathcal{B}^*$ over all the subjects in the same group. The group-level BIC score of applying a DBN M to the group data in the manner of the CS approach is the sum of its BIC scores over the data of each subject:

$$\mathrm{BIC}(M|X_1, \cdots, X_S) = \sum_{s=1}^{S} \left[\sup_{\theta_s} \ln P(X_s|M,\theta_s) - 0.5K\ln N_s\right], \tag{7}$$

where K is without the subscript s because DBNs with the same structure have the same number of parameters.

Based on the assumption that inter-subject variability is so small that they can be neglected, the VTS approach pools together all subjects' data and learns a mixture model for the whole group. Every subject is associated with the same G^* and $\mathcal{B}^*$, and they are the GMS $\bar{G}$ and the GMC $\bar{\mathcal{B}}$ as well, respectively. Because the data are pooled together, the group-level BIC score of a model M is

$$\begin{aligned}\mathrm{BIC}(M|X_1, \cdots, X_S) = \sup_\theta &\left[\sum_{s=1}^{S} \ln P(X_s|M,\theta)\right] \\ &-0.5K\ln\left(\sum_{s=1}^{S} N_s\right)\end{aligned} \tag{8}$$

Comparison of the three approaches

The IS, VTS and CS approaches were compared from the aspects of the goodness of fit, the similarity between their models at the group level and the sensitivity to the effect of the L-dopa medication. First, their best group-level BIC scores found in the MCMC sampling were compared. As a criterion of model selection, the greater the BIC score is, the better the model fits the data (see Learning DBNs section). More parameters do not guarantee a model a larger BIC score because the penalty term $0.5K\ln N$ in Eq. (5) punishes models with redundant parameters. Group-level BIC scores of the IS approach, the CS approach and the VTS approach are defined in Eqs. (6), (7) and (8), respectively.

To examine the feasibility of using the best BIC scores for selecting the IS, VTS and CS approaches, we carried out simulation studies to investigate the following question: whether the true underlying generating models can be correctly identified with the highest BIC scores. In our simulation, similar to the setting of our real fMRI data set, fMRI time courses of the 6 ROIs at 130 time points were simulated for 10 subjects from mixture DBNs derived from the real fMRI data set. We leaned mixture DBN models from the real fMRI data of a group of 10 subjects (e.g., PD patients before medication) with the IS, VTS and CS approaches respectively. Then the learned mixture models were used as true models of the IS, VTS and CS approaches respectively to generate the simulation data.

Different approaches may fit the data with different BIC scores, but it is still possible for them to yield similar network structures and parameters at the group level. To inspect whether the choice of the analysis approaches impacts the group-level model, we compared the GMSs $\bar{G}$ or GMCs $\bar{\mathcal{B}}$ across approaches. For each of the three comparisons (IS vs. CS, IS vs. VTS and CS vs. VTS), we element-wise plotted the GMSs or GMCs across the two approaches under comparison. Because an element of $\bar{\mathcal{B}}$ will become zero when its corresponding element of $\bar{G}$ is zero, an element of $\bar{\mathcal{B}}$ is plotted only if its corresponding element of $\bar{G}$ is larger than

[2] For rigorous proof, please refer to Schwarz (1978).

1%, to eliminate the information that has already been presented in the comparison of GMSs. If the dots on the plotted graphs are located roughly along the diagonal line, the two approaches yield similar network structures or parameters at the group level.

As a statistical metric, a BIC score alone cannot indicate whether a model approximates the underlying biomedical truth well, so the sensitivity of the three approaches to the possible effect of the L-dopa medication was also compared. Because it has been demonstrated that disturbances of rhythmic movement in PD (Pope et al., 2006) is related to bradykinesia (slowness of movement) and L-dopa is most effective against bradykinesia (Jankovic and Tolosa, 2002), we assume that the L-dopa medication will generally normalize PD patients' brain connectivity in this frequency-related bulb-squeezing task. The overall effect of the medication was verified by checking whether the mixture DBN model of each patient before medication (denoted as the P_{pre} group) changed toward those of the normal group (denoted as the N group) after the patient took the medication (denoted as the P_{post} group).

At the structure level, a connection is considered to have been changed toward the normal group, if elements of $G^*_{Ps_{\text{post}}} - G^*_{Ps_{\text{pre}}}$ and $\bar{G}_N - G^*_{Ps_{\text{pre}}}$ (where subscripts "Ps_{pre}" and "Ps_{post}" indicate that the comparison is about patient s) have the same signs, or equally elements of

$$G'_s = \left(G^*_{Ps_{\text{post}}} - G^*_{Ps_{\text{pre}}}\right) * \text{sign}\left(\bar{G}_N - G^*_{Ps_{\text{pre}}}\right) \tag{9}$$

are positive where an asterisk symbol (∗) denotes element-wise products. At the parameter level, a connection was considered to have been changed toward the normal group, if elements of $\mathcal{B}^*_{Ps_{\text{post}}} - \mathcal{B}^*_{Ps_{\text{pre}}}$ and $\bar{\mathcal{B}}_N - \mathcal{B}^*_{Ps_{\text{pre}}}$ have the same sign, or equally elements of

$$\mathcal{B}'_s = \left(\mathcal{B}^*_{Ps_{\text{post}}} - \mathcal{B}^*_{Ps_{\text{pre}}}\right) * \text{sign}\left(\bar{\mathcal{B}}_N - \mathcal{B}^*_{Ps_{\text{pre}}}\right) \tag{10}$$

are positive. To summarize this patient-matched comparison at the group level, we averaged G'_s and B'_s over all the patients as $G' = \Sigma_{s=1}^{S} G'_s/S$ and $\mathcal{B}' = \Sigma_{s=1}^{S} \mathcal{B}'_s/S$ respectively. If an element of G' or $\mathcal{B}'$ is positive, we call the corresponding connection a "normalized connection" at the structure or the parameter level respectively. The same patient performed the task twice, once before and the other after the medication, so the comparison is patient matched. Because the normal group are healthy people who cannot be individually matched with patients, its group means ($\bar{G}_N$ and $\bar{\mathcal{B}}_N$) are used to represent healthy people. The signs of G'_s and $\mathcal{B}'_s$ are standardized relatively to the signs of $\bar{G}_N - G^*_{Ps_{\text{pre}}}$ and $\bar{\mathcal{B}}_N - \mathcal{B}^*_{Ps_{\text{pre}}}$ respectively, so that positive values indicate that the change is toward that of healthy people.

Based on the "normalized connections", we performed tests between the two hypotheses below:

- H_0: The medication does not functionally change brain connectivity in PD overall.
- H_1: The medication changes the PD subjects' connections between brain regions functionally toward those of normal subjects.

The possibility of a connection appearing to be "normalized" by the medication is parameterized as γ. The value of γ equals 0.5 under H_0, due to the effect of randomness; γ is expected to be greater than 0.5 under H_1. If m out of n connections appear to be normalized by the medication, the right-tailed p-value of this observation under H_0 is $\sum_{i=m}^{n} C_n^i/2^n$.

To further assess the similarity/diversity of the DBN structures learned with the IS approach, the variance of a connection's posterior probabilities, i.e., $\text{Var}[g^*_{ab}]$ was calculated among subjects' mixture DBNs. A small $\text{Var}[g^*_{ab}]$ suggests that the connection from ROI a to ROI b consistently appears (or does not appear) in subjects' mixture DBNs; a large variance suggests that its appearance differs considerably among subjects' mixture DBNs. If a big portion of connections have small $\text{Var}[g^*_{ab}]$, then the DBNs learned individually for each subject are structurally consistent; if a big portion of connections have large $\text{Var}[g^*_{ab}]$, then the DBNs structurally differ among subjects. Please note that the upper bound of the variance is 0.5, not 1, because posterior probabilities are in the range of [0,1].

Group comparison

The group analysis is to identify those connections that appear to be statistically significantly normalized by the medication. Though a connection appears to be normalized if its corresponding element in $\mathcal{B}'$ is positive, the statistical significance of the normalization still needs to be tested. In this section, we employed mixed effect models to verify whether elements of $\mathcal{B}'$ are significantly positive. Because the data of the normal group were best fitted with the VTS approach and the data of the patient groups were best fitted with the IS approach (see Table 2), we applied the $\bar{\mathcal{B}}_N$ of the VTS approach and the $\mathcal{B}^*_{Ps_{\text{pre}}}$ and $\mathcal{B}^*_{Ps_{\text{post}}}$ of the IS approach to Eq. (10) to calculate B'.

An element $\beta'_{ab|l}$ of $\mathcal{B}'$ is associated with a connection from a to b and a frequency level l. When the connection from a to b is under investigation, $\beta'_{ab|l}$ from all the patients and at all the frequency levels can be regressed to the frequency level as in Eq. (11),

$$\beta'_{ab} = c + \text{Freq} + e, \tag{11}$$

where β'_{ab} is a vector of $\beta'_{ab|l}$, c is a constant, and vector e is the regression errors which follow a zero-mean multivariate Gaussian distribution. β'_{ab} can also be just regressed to a constant as in Eq. (12) without the term Freq if the effect of the medication does not differ at different task frequencies.

$$\beta'_{ab} = c + e. \tag{12}$$

To accommodate the variability among the randomly sampled subjects and as well the noise in the fMRI of each subject, the variance of e is composed of two parts: $V = V_a + V_w$ which are the variance among and within subjects, respectively, in the way of the summary statistics approach. Since the variance term includes both among-subject and within-subject terms, this is a mixed effect model that can infer about the population, rather than just the pool of the recruited subjects. V_a has the pattern of $\sigma^2 I$, under the assumption that subjects are independently sampled from the population. V_w is a blocked-diagonal matrix whose diagonal blocks are variances within subjects. The diagonal blocks of V_w are estimated at the stage of calculating $\beta^*_{ab|l}$, as $V_{ws} = V\left(\beta^*_{Ps_{\text{pre}}}\right) + V\left(\beta^*_{Ps_{\text{post}}}\right)$ where the subscript s indicates that the block is associated with patient s. Eqs. (11) and (12) were solved with the restricted maximum likelihood method (Corbeil and Searle, 1976) to ensure an unbiased estimation of V.

The analysis includes three steps. First, nested models Eq. (11) and Eq. (12) are compared with likelihood ratio test, to test whether the medication altered the connections differently for tasks at different frequencies. If Eq. (11) fits the data significantly better

Table 1
Group-level BIC scores of simulated data

		Learning Approaches		
		IS	CS	VTS (pool)
True models	IS	**−5.40e−3**	−5.85e−3	−6.47e−3
	CS	−5.42e−3	**−5.35e−3**	−5.68e−3
	VTS	−6.27e−3	−6.36e−3	**−5.63e−3**

Group-level BIC scores of the IS approach, the CS approach and the VTS approach are defined in Eqs. (6), (7) and (8) respectively. Bold digits are the highest BIC scores found by the three approaches.

than Eq. (12) does, then the medication has different effects for tasks at different frequencies. However, significant results were not discovered for any connection (see Results section), so all the preceded analyses were performed with Eq. (12). Second, whether c is equal to or significantly above zero is tested with right-tailed t-tests. A p-value was calculated for each possible connection in the DBN. Finally, the effect of testing multiple connections simultaneously was adjusted by Storey's positive-false-discovery-rate (pFDR) procedure (Storey, 2002) which controls the pFDR (characterized by a q-value), the expected ratio of falsely rejected hypotheses among all those being rejected. Connections whose q-values were smaller than 5% were selected and regarded to have been statistically significantly normalized by the medication.

Results

Table 1 shows the group-level BIC scores of the simulated data set. The generating models were correctly identified with the highest BIC scores. Table 2 shows the group-level BIC scores of the real data set. For the normal group, the VTS approach yielded the highest group-level BIC score; for the pre-medication and post-medication patient groups, the IS approach yielded the highest group-level BIC scores. The best fitting approaches outperformed the second best fitting approaches with at least 190 BIC scores, which implies that if the conditions of the BIC's asymptomatic approximation to the log posterior probability are satisfied, then the model of the best fitting approach is e^{190} times more likely than that of the second best fitting approach. Though the conditions of the BIC's asymptomatic approximation may not be guaranteed in this study, the sizeable contrasts among the BIC scores still suggest that the goodness-of-fit of the three approaches is considerably different.

Fig. 3 shows the element-wise comparison of the group mean structures learned with the three approaches. If two approaches learn similar networks at the group level, dots on the figures should be located close to the (0, 0) to (1, 1) diagonal lines. However, this phenomenon is not observed. The connections' posterior probabilities estimated by the CS approach tend to cluster at extreme values near 0 and 1, while those estimated by the IS approach cover the range from 0 to 1 more evenly.

Table 2
Group-level BIC scores of real data

	IS	CS	VTS (pool)
N	−5.25e−3	−5.38e−3	**−5.06e−3**
P_{pre}	**−4.99e−3**	−5.21e−3	−5.31e−3
P_{post}	**−6.11e−3**	−6.55e−3	−6.39e−3

Group-level BIC scores of the IS approach, the CS approach and the VTS approach are defined in Eqs. (6), (7) and (8) respectively. Bold digits are the highest BIC scores found by the three approaches. "N", "P_{pre}" and "P_{post}" denote the normal group, the pre-medication group and the post-medication group, respectively.

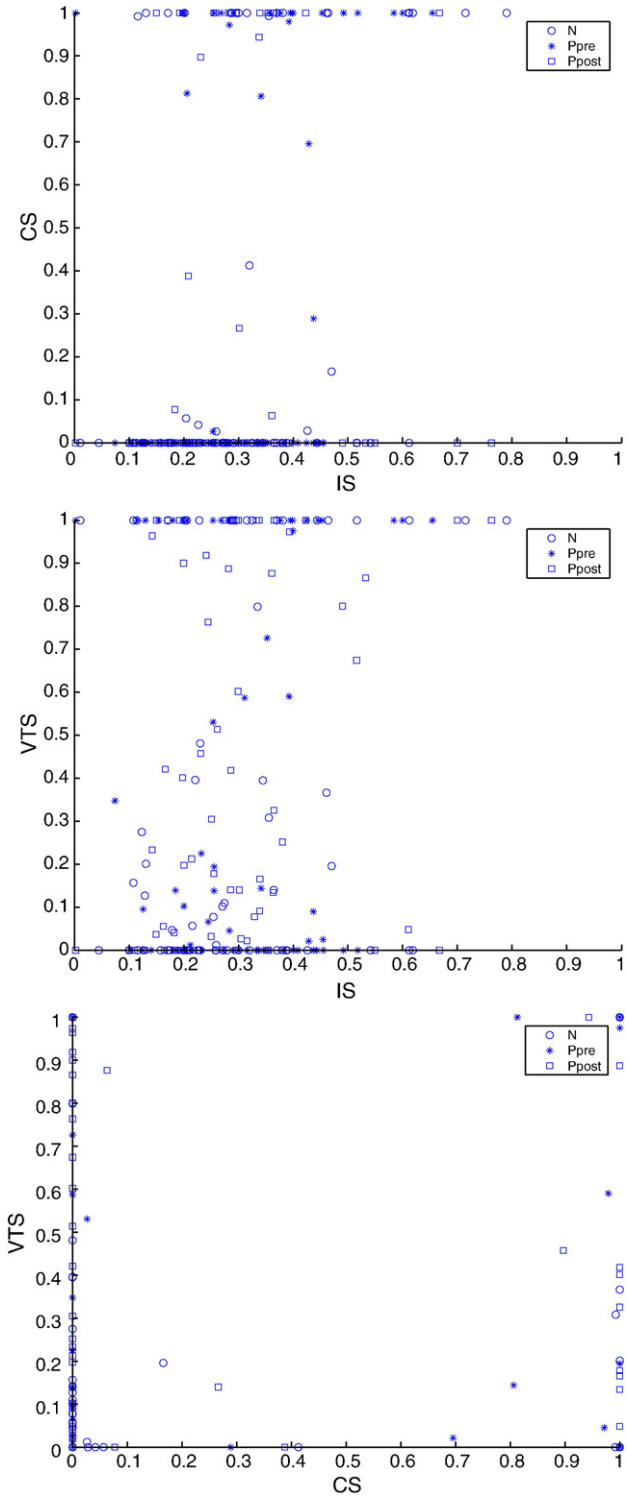

Fig. 3. Comparisons of the group mean structures. The group–mean–structure matrices G of the IS, CS and VTS approaches are plotted element-wise. Plots of the three experimental groups are overlaid with different dot symbols. Legends "N", "P_{pre}" and "P_{post}" denote the normal group, the pre-medication group and the post-medication group, respectively.

Fig. 4 shows the element-wise comparison of the group mean parameters learned with the three approaches. To remove the information that Fig. 3 has showed, an element of $\bar{\mathcal{B}}$ is plotted only if its corresponding element of $\bar{G}$ are larger than 1%. Dots are located roughly along the (−1,−1) to (1,1) diagonal lines, which suggests that the three approaches tend to learn similar coefficients at the group level for those connections that always appear in the learned networks no matter which approach is applied.

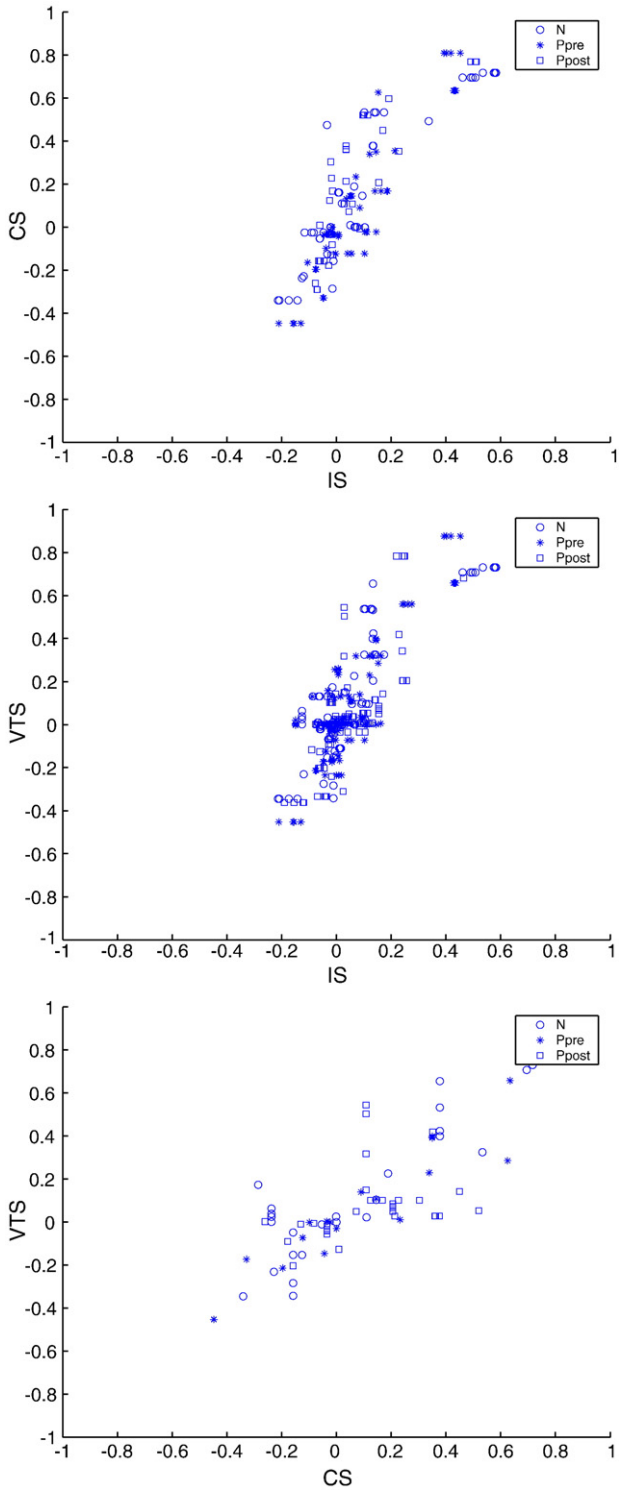

Fig. 4. Comparisons of the group mean coefficients. The group mean coefficients $\bar{\mathcal{B}}$ of the IS, CS, and VTS approaches are plotted element-wise. An element of $\bar{\mathcal{B}}$ is plotted only if its corresponding element of $\bar{G}$ is larger than 1%. Plots of the three experimental groups are overlaid with different dot symbols. Legends "N", "P_{pre}" and "P_{post}" denote the normal group, the pre-medication group and the post-medication group, respectively.

Table 3
The counts of normalized connections

	IS		CS		VTS (pool)	
	Count	p	Count	p	Count	p
Structure	59/66	0.0000	15/66	1.0000	29/66	0.8661
0.00 Hz	64/66	0.0000	27/66	0.9456	35/66	0.3561
0.25 Hz	62/66	0.0000	27/66	0.9456	34/66	0.4511
0.50 Hz	62/66	0.0000	27/66	0.9456	35/66	0.3561
0.75 Hz	64/66	0.0000	27/66	0.9456	34/66	0.4511
Mean	63/66	0.0000	27/66	0.9456	34/66	0.4511

The format of counts is "the number of normalized connections/the number of possible connections". In row "Structure" are the counts of positive elements in G'; in rows from "0.00 Hz" to "0.75 Hz" are the counts of positive elements in $\mathcal{B}'$, given the task frequency level; in row "Mean" are the counts of positive elements in $\mathcal{B}'$, with the elements at the four frequency levels averaged.

Table 3 shows the number of normalized connections detected with the three approaches. The IS approach detected more normalized connections than the other two approaches. It found that at least 59 connections out of all the 66 possible ones were normalized after the medication. The p-value of this observation is smaller than 1.193×10^{-11}, under the hypothesis that the medication does not improve patients' brain connectivity overall.

Fig. 5 shows the histogram of Var[g^*_{ab}], i.e., the variances of connections' posterior probabilities among the mixture DBN models learned with the IS approach. Given the fact that the variance of posterior probabilities cannot exceed 0.5, a large portion of connections vary considerably among the networks of different subjects,

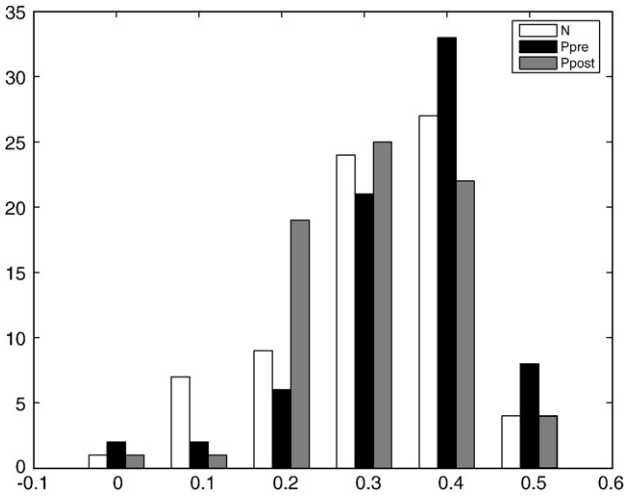

Fig. 5. The histogram of the variances of connections' posterior probabilities estimated with the IS approach. The posterior probability of a connection's existence was estimated individually for each subject, and the variance of a connection's posterior probabilities among different subjects was then calculated. This figure shows the histogram of the posterior probability variances of the connections. Legends "N", "P_{pre}" and "P_{post}" denote the normal group, the pre-medication group and the post-medication group, respectively.

Table 4
Connections whose coefficients were statistically significantly normalized toward those of the normal group

Connection	*t*-Statistics	*df*	*p*	*q*
R_CER↦R_CER	2.5828	18	0.0094	0.0424
R_CER↦L_CER	4.1458	15	0.0004	0.0065
R_M1↦R_M1	3.0563	18	0.0034	0.0275
R_SMA↦R_SMA	2.6912	27	0.0060	0.0362
L_CER↦R_CER	6.6237	15	0.0000	0.0001
L_CER↦L_CER	2.6565	9	0.0131	0.0424
R_CER→R_M1	2.4417	12	0.0155	0.0424
R_M1→R_CER	3.0219	18	0.0037	0.0275
R_M1→L_M1	2.4717	12	0.0147	0.0424
R_SMA→R_CER	2.3685	18	0.0146	0.0424
L_SMA→R_SMA	2.4601	12	0.0150	0.0424

Arrows ↦ denote connections with a time lag; arrows → denote connections without a time lag. "*df*" is short for "degrees of freedom".

which suggests that different subjects' networks learned with the IS approach are not consistent. Even though subjects of the normal group probably share the same brain connectivity pattern, which is suggested by the highest BIC score learned with the VTS approach, networks still vary considerably from subject to subject when the IS approach was applied.

The likelihood-ratio tests show that Eq. (11) does not fit β'_{ab} significantly better than Eq. (12) does for any connection, with the minimum *p*-value larger than 0.0769, before correction for the effect of multiple comparisons. This result does not support that the medication differently altered the connections for tasks at different frequencies. This does not mean that the medication does not alter the connections, but just that the alterations are probably the same at different frequency levels. The connections whose coefficients were significantly normalized according to the right-tailed *t*-test with Eq. (12) are listed in Table 4 and further graphically illustrated in Fig. 6. These results indicate that the right SMA, right M1 and left cerebellum, all part of the motor system expected to be activated with the left hand were modulated by L-dopa mediation.

Discussion

The comparison of the three DBN-based group-analysis approaches (the "virtual-typical-subject" (VTS), "common-structure" (CS) and "individual-structure" (IS) approaches) suggests that no single approach is universally superior over the other two. For this fMRI study on Parkinson's disease that included a motor task at different frequencies, the VTS approach fits the data of the normal group best, while the IS approach fits the data of the pre-medication and post-medication patient groups best, from the view of statistics with the BIC score as the goodness-of-fit metric (see Table 2).

The three approaches led to considerably different group-level results, learning different network structures, and detecting different numbers of connections normalized by the medication, from the same data set. Fig. 3 shows that the probability of a connection's existence may be estimated to be quite high with one approach, but quite low with another approach. The overall effect of the L-dopa medication on the brain connectivity of Parkinson's disease patients is also estimated quite differently with the three approaches, with the proportion of normalized connections varies from 15/66 to 64/66, as shown in Table 3. The results of the IS approach support the assumption better that the L-dopa medication normalizes the brain connectivity for the patients than those of the other two approaches do. The regions showing normalized activity after L-dopa medication are consistent with motor tasks involving the left hand. The sizeable differences among the results of the three approaches suggest that choice over the three approaches influence the analysis considerably. Analysts should not arbitrarily choose an approach without justification. However, as Fig. 4 shows, the three approaches tend to learn similar coefficients at the group level for those connections that always appear in the learned networks no matter which approach is applied.

If the IS approach is applied, the networks tend to vary considerably from subject to subject, even when statistical metrics support that the underlying true models probably share the same structure. For the normal group in this study, its highest BIC score was learned with the VTS approach, suggesting that the data of those normal people are more similar than dissimilar. However, the networks learned with the IS approach still vary considerably from subject to subject, as Fig. 5 shows sizeable variances of connections' posterior probabilities among the individual DBN models. The reason behind this phenomenon probably is that the limited time points of each individual's data do not support an accurate and robust estimation of the true model. The inconsistency among individual networks is not an effective evidence of inter-subject variability, if only the IS approach is applied.

Each of the three approaches has its limitations in practice. The IS approach do not take advantage of the similarity among group members, so when each subject's data is not long enough, it tends to learn non-robust networks for each subject as mentioned above with Fig. 5. The VTS approach does not accommodate differences among subjects, but inter-subject variability is believed to be a common factor in biomedical studies. As shown in Table 3, ignoring the differences among patients may limit the analysis' sensitivity to the effect of the L-dopa medication. The CS approach balances the similarity and diversity among group members at the point that the connectivity structure is the same for all the subjects but details of the connectivity are different. However, it is not sensitive to the effect of the L-dopa medication either, as shown in Table 3. It is possible that the subjects' connectivity networks are the same partially at some connections but differ at other connections, in which case none of the three approaches is suitable. It is ultimately desirable to develop a method that allows researchers to adjust the degree of the balance between the similarity and

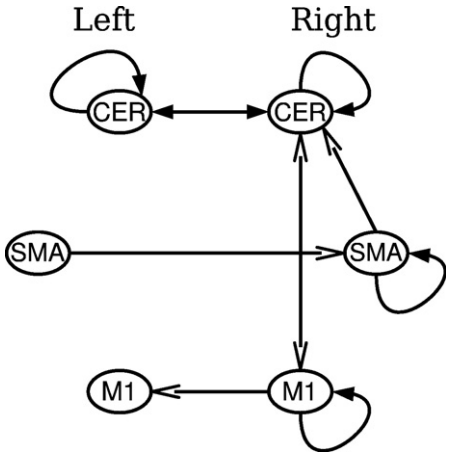

Fig. 6. Connections whose coefficients were statistically significantly normalized toward those of the normal group. Solid arrows denote connections with a time lag; line arrows denote those without a time lag.

diversity among group members. The method that Niculescu-Mizil and Caruana (2007) proposed is a possible solution. It allows network structures to be different while punishes excessive diversity among the structures and controls the degree of the punishment with an adjustable parameter. Detection of sub-groups and outliers can also improve group analysis when inter-subject variability cannot be ignored (Kherif et al., 2003).

Group-analysis approaches should be selected according to comprehensive criteria, such as both statistical and biomedical evidence. In exploratory researches where the ground truth is usually unknown, pieces of statistical evidence are helpful in the selection of group-analysis methods. Besides the BIC score used in the paper, the model evidence calculated with MCMC is a more accurate metric, at the cost of intensive computation, since the BIC score which asymptotically approximates the log posterior probability is inaccurate when the sample size is small. Other goodness-of-fit metrics, such as the Dirichlet Prior Score Metric (DPSM) and the Bayesian Dirichlet metric (BDe), can also be used (Yang and Chang, 2002). The confidence of the statistical evidence can be verified with data simulated from the learned models. Generating data from the models learned from the real data is to make the simulated data as similar to the real data as possible. However, pieces of statistical evidence alone are not sufficient in making decisions, especially in biomedical studies, if heterogeneity of group members is not considered from other aspects such as age, gender, etc. If possible, researchers can design biomedical markers and select group-analysis approaches according to their ability to detect those markers, as we assumed the effect of the L-dopa medication in the study.

Acknowledgments

This work was supported by NSERC Grant CHRPJ 323602-06 (M.J.M.) and CIHR grant CPN-80080 (M.J.M.).

References

Brainard, D.H., 1997. The psychophysics toolbox. Spat. Vis. 10 (4), 433–436.
Corbeil, R.R., Searle, S.R., 1976. A comparison of variance component estimators. Biometrics 32 (4), 779–791 (Dec.).
Edelman, G.M., Gally, J.A., 2001. Degeneracy and complexity in biological systems. PNAS 98 (24), 13763–13768.
Friston, K.J., 1994. Functional and effective connectivity in neuroimaging: a synthesis. Hum. Brain Mapp. 2 (1-2), 56–78.
Friston, K.J., Harrison, L., Penny, W., 2003. Dynamic causal modelling. NeuroImage 19 (4), 1273–1302 (Aug.).
Goncalves, M.S., Hall, D.A., Johnsrude, I.S., Haggard, M.P., 2001. Can meaningful effective connectivities be obtained between auditory cortical regions? NeuroImage 14 (6), 1353–1360 (Dec.).
Harrison, L., Friston, K., 2004. Human Brain Function, 2nd Edition. Academic Press. Ch. 48 and 50.
Harrison, L., Penny, W.D., Friston, K., 2003. Multivariate autoregressive modeling of fMRI time series. NeuroImage 19 (4), 1477–1491 (Aug.).
Hoehn, M.M., Yahr, M.D., 1967. Parkinsonism: onset, progression and mortality. Neurology 17 (5), 427–442 (May).
Jankovic, J., Tolosa, E. (Eds.), 2002. Parkinson's Disease and Movement Disorders, 4th Edition. Williams & Wilkins.
Kherif, F., Poline, J.-B., Meriaux, S., Benali, H., Flandin, G., Brett, M., 2003. Group analysis in functional neuroimaging: selecting subjects using similarity measures. NeuroImage 20 (4), 2197–2208 (Dec.).
Kim, J., Zhu, W., Chang, L., Bentler, P.M., Ernst, T., 2007. Unified structural equation modeling approach for the analysis of multisubject, multivariate functional MRI data. Hum. Brain Mapp. 28 (2), 85–93.
Lauritzen, S. L., 1996a. Graphical Models. Clarendon Press, Oxford University Press, Oxford, New York.
Lauritzen, S.L., 1996b. Graphical Models. Clarendon Press, Oxford University Press, Ch. 3.2.2, pp. 46–52.
Li, J., Wang, Z., McKeown, M., 2007. A multi-subject dynamic Bayesian network (DBN) framework for brain effective connectivity. Acoustics, Speech and Signal Processing, 2007. IEEE International Conference on. Vol. 1, pp. I–429–I–432.
Liao, R., Krolik, J.L., McKeown, M.J., 2005. An information-theoretic criterion for intrasubject alignment of fMRI time series: motion corrected independent component analysis. IEEE Trans. Med. Imag. 24 (1), 29–44 (Jan).
Liao, R., McKeown, M.J., Krolik, J.L., 2006. Isolation and minimization of head motion-induced signal variations in fMRI data using independent component analysis. Magn. Reson. Med. 55 (6), 1396–1413 (Jun.).
McIntosh, A.R., Gonzalez-Lima, F., 1994. Structural equation modeling and its application to network analysis in functional brain imaging. Hum. Brain Mapp. 2, 2–22.
Mechelli, A., Penny, W.D., Price, C.J., Gitelman, D.R., Friston, K.J., 2002. Effective connectivity and intersubject variability: using a multisubject network to test differences and commonalities. NeuroImage 17 (3), 1459–1469 (11).
Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E., 1953. Equation of state calculations by fast computing machines. J. Chem. Phys. 21 (6), 1087–1092.
Murphy, K. P., 2002. Dynamic Bayesian networks: representation, inference and learning. Ph.D. thesis, University of California, Berkeley.
Niculescu-Mizil, A., Caruana, R., 2007. Inductive transfer for Bayesian network structure learning. Proceedings of the 11th International Conference on AI and Statistics (AISTATS '07).
Pelli, D.G., 1997. The videotoolbox software for visual psychophysics: transforming numbers into movies. Spat. Vis. 10 (4), 437–442.
Penny, W., Ghahramani, Z., Friston, K., 2005. Bilinear dynamical systems. Phil. Trans. R. Soc. B 360 (1457), 983–993 (May).
Pope, P.A., Praamstra, P., Wing, A.M., 2006. Force and time control in the production of rhythmic movement sequences in Parkinson's disease. Eur. J. Neurosci. 23 (6), 1643–1650 (Mar).
Price, C.J., Friston, K.J., 2002. Degeneracy and cognitive anatomy. Trends Cogn. Sci. 6 (10), 416–421 (Oct.).
Rajapakse, J.C., Zhou, J., 2007. Learning effective brain connectivity with dynamic Bayesian networks. NeuroImage 37, 749–760.
Rorden, C., Brett, M., 2000. Stereotaxic display of brain lesions. Behav. Neurol. 12 (4), 191–200.
Schwarz, G., 1978. Estimating the dimension of a model. Ann. Stat. 6 (2), 461–464 Short Communications (Mar.).
Stelmach, G.E., Teasdale, N., Phillips, J., Worringham, C.J., 1989. Force production characteristics in Parkinson's disease. Exp. Brain Res. 76 (1), 165–172.
Storey, J.D., 2002. A direct approach to false discovery rates. J. R. Stat. Soc., Ser. B Stat. Methodol. 64 (3), 479–498.
Sugihara, G., Kaminaga, T., Sugishita, M., 2006. Interindividual uniformity and variety of the "writing center": a functional MRI study. NeuroImage 32 (4), 1837–1849 (Oct.).
Talairach, J., Tournoux, P., 1988. Co-Planar Stereotaxic Atlas of the Human Brain. Thieme Medical Publishers.
Turner, R.S., Grafton, S.T., McIntosh, A.R., DeLong, M.R., Hoffman, J.M., 2003. The functional anatomy of Parkinsonian bradykinesia. Neuroimage 19 (1), 163–179 (May).
Vandenbroucke, M.W.G., Goekoop, R., Duschek, E.J.J., Netelenbos, J.C., Kuijer, J.P.A., Barkhof, F., Scheltens, P., Rombouts, S.A.R.B., 2004. Interindividual differences of medial temporal lobe activation during encoding in an elderly population studied by fMRI. NeuroImage 21 (1), 173–180 (Jan.).
Yang, S., Chang, K.-C., 2002. Comparison of score metrics for Bayesian network learning. Systems, Man and Cybernetics: Part A. IEEE Transactions on 32 (3), 419–428.
Zheng, X., Rajapakse, J.C., 2006. Learning functional structure from fMR images. NeuroImage 31 (4), 1601–1613 (Jul.).