Subject Section

A Network of Networks Approach for Modeling Interconnected Brain Tissue-Specific Networks

Hideko Kawakubo 1, Yusuke Matsui 2, Itaru Kushima 3,4, Norio Ozaki 4 and Teppei Shimamura 1 *

1 Division of Systems of Biology, Nagoya University Graduate School of Medicine, 65 Tsurumai-cho, Showa-ku, Nagoya, Aichi, 466-8550, Japan 2 Laboratory of Intelligence Healthcare, Nagoya University Graduate School of Medicine, 1-1-20 Daiko-Minami, Hijashiku, Nagoya, Aichi, 461-8673, Japan 3 Institute for Advanced Research, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi, 464-8601, Japan 4 Department of Psychiatry, Nagoya University Graduate School of Medicine, 65 Tsurumai-cho, Showa-ku, Nagoya, Aichi, 466-8550, Japan

* To whom correspondence should be addressed.

Associate Editor: XXXXXXX

Received on XXXXX; revised on XXXXX; accepted on XXXXX

Abstract

Motivation: Recent sequence-based analyses have identified a lot of gene variants that may contribute to neurogenetic disorders such as autism spectrum disorder and schizophrenia. Several state-of-the-art network-based analyses have been proposed for mechanical understanding of genetic variants in neurogenetic disorders. However, these methods were mainly designed for modeling and analyzing single networks that do not interact with or depend on other networks, and thus cannot capture the properties between interdependent systems in brain-specific tissues, circuits, and regions which are connected each other and affect behavior and cognitive processes.

Results: We introduce a novel and efficient framework, called a “Network of Networks” (NoN) approach, to infer the interconnectivity structure between multiple networks where the response and the predictor variables are topological information matrices of given networks. We also propose Graph-Oriented SParsE Learning (GOSPEL), a new sparse structural learning algorithm for network graph data to identify a subset of the topological information matrices of the predictors related to the response. We demonstrate on simulated data that GOSPEL outperforms existing kernel-based algorithms in terms of F-measure. On real data from human brain region-specific functional networks associated with the autism risk genes, we show that the NoN model provides insights on the autism-associated interconnectivity structure between functional interaction networks and a comprehensive understanding of the genetic basis of autism across diverse regions of the brain.

Availability: Our software is available from https://github.com/infinite-point/GOSPEL.

Contact: kawakubo@med.nagoya-u.ac.jp, shimamura@med.nagoya-u.ac.jp

Supplementary information: Supplementary data are available at Bioinformatics online.

1 Introduction

Neurodevelopmental disorders are characterized by impaired functions of the central nervous system that can appear early in development and often persist into adulthood (Tollefsbol, 2017). The spectrum of developmental impairment varies and includes intellectual disabilities, communication and social interaction challenges, and attention and executive function deficits (American Psychiatric Association, 2013). Prototypical examples of neurodevelopmental disorders are intellectual disability, autism spectrum disorder (ASD), epilepsy, and schizophrenia.

Recent sequence-based analyses have unraveled a complex, polygenic, and pleiotropic genetic architecture of neurodevelopmental disorders, and have identified valuable catalogs of genetic variants as genetic risk factors...
Fig. 1. Overview of GOSPEL, an example for the case where \( n = 4, p = 8 \). Assume that we are given the human brain region-specific networks associated with a disease. As input, GOSPEL requires \( p \) adjacency matrices generated from \( p \) network graphs with \( n \) nodes, where \( n \) and \( p \) indicate the number of nodes (genes) and features (brain regions) respectively. Block ‘A’ shows that GOSPEL estimates the brain regions which are related to ‘Brain region 1’ by performing a graph-oriented sparse regression. In this example, ‘Brain region 2’ and ‘Brain region 8’ are related to ‘Brain region 1’. Block ‘B’ illustrates that a regression is performed on each of the brain regions. As with block ‘A’, block ‘B’ estimates the relationship between the target brain region and the other brain regions. Block ‘C’ expresses that the output of GOSPEL, the “Network of Networks” (NoN) model related to the disease genes, is constructed from the obtained regression coefficients.

for neurodevelopmental disorders (Gratten et al., 2014). However, it remains unknown if and how genetic variants interact with environmental and epigenetic risk factors to impart brain dysfunction or pathology.

For a mechanical understanding of specific genetic variants in neurodevelopmental disorders, integrative network approaches have attracted much attention in recent years due to their interdisciplinary applications. Several state-of-the-art network-based analyses provide an organizational framework of functional genomics and demonstrate that they will enable the investigation of relationships that span multiple levels of analysis (Parikshak et al., 2013; Krishnan et al., 2016; Gandal et al., 2018). These methods were mainly designed for modeling and analyzing single networks that do not interact with or depend on other networks. However, the brain consists of a system of multiple interacting networks and must be treated as such. In multiple interacting networks, the failure of nodes in one network generally leads to the failure of dependent nodes in other networks, which in turn may cause further damage to the first network, leading to cascading failures and catastrophic consequences (Gao et al., 2012). It is known, for example, that different kinds of brain-specific tissues, circuits, and regions are also coupled together and affect behavior and cognitive processes, and thus dysfunctions of the central nervous system in neurodevelopmental disorders have been the result of cascading failures between interdependent systems in the brain. However, no systematic mathematical framework is currently available for adequately modeling and analyzing the consequences of disruptions and failures occurring simultaneously in interdependent networks.

We address this limitation by developing a novel and efficient framework, called the “Network of Networks” (NoN) approach, that will provide useful insights on the properties and topological structure of the inter-correlations between functional interaction networks (Figure 1). Motivated by a perspective on structural equation models, we model the topological information of each network as a weight sum of the topological information of all other networks. Our NoN model enables the exploitation of the interconnectivity structure between complex systems. It has shown to be effective in aiding the comprehensible understanding of the genetic basis of neurodevelopmental disorders across diverse tissues, circuits, and regions of the brain.

Our main contributions are summarized as follows:

1. We define a statistical framework of structural equation models for inferring the interconnectivity structure between multiple networks where the response and the predictor variables are given networks which have topological information. Structural equation modeling is a statistical method used to test the relationships between observed and latent variables (Civelek, 2018). We extend the structural equation models for modeling the effects of network-network interactions.

2. In order to accomplish this, we propose a sparse learning algorithm for network graph data, called Graph-Oriented SPArE Learning (GOSPEL), to find a subset of the topological information matrices of the predictor variables (networks) related to the response variable (network). More specifically, we propose to use particular forms of diffusion kernel-based centered kernel alignment (Cortes et al., 2012) as a measure of statistical correlation between graph Laplacian matrices, and solve the optimization problem with a novel graph-guided generalized fused lasso. This new formulation allows the identification of all types of correlations, including non-monotone and non-linear relationships, between two topological information matrices.

3. We use a Bayesian optimization-based approach to optimize the tuning parameters of the graph-guided generalized fused lasso and automatically find the best fitting NoN model with an acquisition function. The software package that implements the proposed method in the R environment is available from https://github.com/infinite-point/GOSPEL.

We describe our proposed framework and algorithm, and discuss properties in Section 2. Section 3.1 contains a simulation study which demonstrates the performance of the proposed method. We use human brain-specific functional interaction networks and known risk genes with strong prior genetic evidence of ASD and identify the interconnectivity
A Network of Networks Approach

between these networks in Section 3.2. Section 4 provides concluding remarks.

2 Method

Our goal is to infer the interconnectivity structure between multiple networks from topological information matrices of given networks. To do this, we make the assumption that each topological information matrix for a given network can be expressed by the linear combination of the topological information matrices of the other given networks. Sparse regression is performed on each of the given networks in order to identify a subset of the topological information matrices of the predictors related to the response. After this computation, NoNode model is constructed from the obtained regression coefficients. In this section, we first explain the problem setting, and then present our method, GOSPEL.

2.1 Problem Setting

Suppose that we are given \( p \) undirected network graphs consisting of \( n \) vertices (nodes) \( V^{(i)} = \{ v_1^{(i)}, \ldots, v_n^{(i)} \} (i = 1, \ldots, p) \) linked by edges. The \( i \)-th adjacency matrix \( A^{(i)} \in \mathbb{R}^{n \times n} \) associated with the \( i \)-th undirected network graph is defined as

\[
A^{(i)}_{j,k} = \begin{cases} 
  w^{(i)}_{j,k} & \text{if } j \neq k \text{ and } v_j \text{ links with } v_k, \\
  0 & \text{otherwise,}
\end{cases}
\]

where \( w^{(i)}_{j,k} \in [0,1] \) denotes the probability of connectivity between the \( v_j \) and \( v_k \) in the \( i \)-th network graph. Here, we compute graph Laplacian matrix \( L^{(i)} \),

\[
L^{(i)}_{j,k} = \begin{cases} 
  \deg(v_j^{(i)}) & \text{if } j = k, \\
  -w^{(i)}_{j,k} & \text{if } j \neq k \text{ and } v_j \text{ links with } v_k, \\
  0 & \text{otherwise,}
\end{cases}
\]

where \( \deg(v_j^{(i)}) \) denotes the degree of \( v_j^{(i)} \). Then let us kernelize the graph Laplacian matrix. Let \( K^{(i)} \) be the diffusion kernel matrix for \( L^{(i)} \):

\[
K^{(i)} = \exp[-(L^{(i)}/\gamma^{(i)})],
\]

where \( \gamma^{(i)} \) is a kernel parameter. This kernel matrix is centered and normalized as follows:

\[
\tilde{K}^{(i)} = K^{(i)}/\|K^{(i)}\|_F, \\
K^{(i)} = H\tilde{K}^{(i)}H,
\]

where \( \| \cdot \|_F \) denotes the Frobenius norm, \( H \) indicates the centering matrix \( H = I_n - \frac{1}{n}1_n1_n^T \), \( I_n \) is an \( n \times n \) identity matrix and \( 1_n \) is an \( n \)-dimensional vector with all ones.

We assume that the diffusion kernel matrix of the \( i \)-th network graph can be represented by linear combinations of the diffusion kernel matrices of the other network graphs as follows:

\[
K^{(i)} = \sum_{j=1}^{p} \beta^{(i)}_{j} K^{(j)} + \epsilon^{(i)}.
\]

where \( \{ \beta^{(i)}_{j} \}_{j=1}^{p} \) denotes a regression coefficient corresponding to predictor \( K^{(j)} \) and response \( K^{(i)} \), and \( \epsilon^{(i)} \in \mathbb{R}^{n \times 1} \) is a Gaussian noise matrix whose elements follow \( N(0, \sigma^2_{\epsilon}) \).

Table 1. Behavior of \( \beta^{(i)}_{j} \) and \( \beta^{(i)}_{k} \) in GOSPEL optimization. When \( K^{(i)} \) and \( K^{(k)} \) are uncorrelated, i.e. \( R^{(i)}_{jk} = 0 \), the value of \( \beta^{(i)}_{j} \) is estimated depending on the correlation between response \( K^{(j)} \) and predictor \( K^{(k)} \). Similarly, the value of \( \beta^{(i)}_{k} \) is computed depending on the correlation between the response and the \( k \)-th predictor. On the other hand, when \( K^{(i)} \) and \( K^{(k)} \) are correlated, i.e. \( R^{(i)}_{jk} = 1 \), \( \beta^{(i)}_{j} \) and \( \beta^{(i)}_{k} \) tend to take similar values depending on the correlation between the response and the \( k \)-th predictor.

<table>
<thead>
<tr>
<th>Correlation</th>
<th>Regression Coefficients</th>
<th>( \beta^{(i)}_{j} )</th>
<th>( \beta^{(i)}_{k} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>uncorrelated</td>
<td>uncorrelated</td>
<td>( \text{a real value} )</td>
<td>( \text{a real value} )</td>
</tr>
<tr>
<td>correlated</td>
<td>correlated</td>
<td>( \text{a real value} )</td>
<td>( \text{a real value} )</td>
</tr>
</tbody>
</table>

2.2 Graph Oriented Sparse Learning (GOSPEL)

The optimization problem of GOSPEL is as follows:

\[
\begin{align*}
\min_{\beta^{(1)}, \ldots, \beta^{(p)}} & \| K^{(1)} - \sum_{j=1}^{p} \beta^{(j)} K^{(j)} \|_F^2 \\
+ & \lambda_1 \sum_{j,k=1}^{p} R^{(1)}_{jk} |\beta^{(j)} - \beta^{(k)}| + \lambda_0 \sum_{j=1}^{p} |\beta^{(j)}|, 
\end{align*}
\]

where \( \lambda_1, \lambda_0 \) are regularization parameters and \( | \cdot | \) indicates the \( \ell_1 \) norm. \( \mathbf{R}^{(1)} \in \mathbb{R}^{p \times p} \) expresses a matrix whose elements consist of correlations between predictors, where

\[
R^{(1)}_{jk} = \begin{cases} 
  1 & \text{if } [\text{CAK}(K^{(j)}, K^{(k)})] \geq \epsilon \text{ and } j \neq k, \\
  0 & \text{otherwise.}
\end{cases}
\]

CAK(\( K^{(j)}, K^{(k)} \)) denotes a correlation between kernel matrices \( K^{(j)} \) and \( K^{(k)} \); this measure is called the Centred Kernel Alignment (CKA) (Cortes et al., 2012), and \( \epsilon \) indicates a threshold. CKA captures the non-linear relationship between two matrices if such a relationship exists. The definition of CKA is as follows:

\[
\text{CAK}(K^{(j)}, K^{(k)}) = \frac{\mathbf{K}^{(j)\top}\mathbf{K}^{(k)}\mathbf{y}}{\|\mathbf{K}^{(j)}\|_F \|\mathbf{K}^{(k)}\|_F},
\]

where \( (\cdot, \cdot)_F \) indicates the Frobenius inner product. The Frobenius inner product can be interpreted as an inner product of two vectorized matrices, and thus we can apply the properties of Pearson’s correlation coefficient (Sharma, 2005) to CKA. Unless the elements of \( K^{(j)} \) or \( K^{(k)} \) are all zero (we omit such cases in the computation of GOSPEL), this definition implies that the value of CKA becomes zero when \( K^{(j)} \) and \( K^{(k)} \) have no correlation and the CKA value takes \( \pm 1 \) when the two matrices are strongly correlated. In practice, the value of the diffusion kernel based CKA ranges from \(-1\) to \(1\) because of the positive semi-definiteness of the diffusion kernel matrix (Lafferty and Kondor, 2002).

GOSPEL optimization Eq. (1) consists of the squared Frobenius norm term, the graph-guided-fused-lasso regularization term (Chen et al., 2012) and the lasso regularization term (Tibshirani, 1996). If \( R^{(1)}_{jk} \) of the graph-guided-fused-lasso regularization term is zero, the equation becomes solely dependent on the lasso regularization term. Table 1 summarizes the behavior of \( \beta^{(i)}_{j} \) and \( \beta^{(i)}_{k} \) in GOSPEL optimization. The table shows that GOSPEL estimates all the relevant predictor networks to the response network, and also eliminates irrelevant predictor networks to the response network. For more detail on the behavior of the regression coefficients, see section 1 in the supplement. The sparsity of the elements of \( \beta^{(i)} \) helps to facilitate the interpretation of the computation results. By extension, the interpretation of the network structure of given networks is also facilitated.
Finally, we construct a NoN model utilizing the obtained regression coefficients \( \{ \beta_{i,j}^{(0)} \}_{j=0}^{p} \). Let \( E \subseteq \{1, \ldots, p\} \times \{1, \ldots, p\} \) be an edge set for a NoN model, where \( E = \{1, \ldots, p\} \). We employ the edge set estimation defined by Meinshausen and Bühlmann (2006):

\[
E \equiv \{(i,j) \mid \beta_{i,j}^{(0)} \neq 0 \lor \beta_{j,i}^{(0)} \neq 0\},
\]

where \( (i,j) \) indicates the pair of the \( i \)-th and the \( j \)-th nodes. Based on the graphical model \( G = (E, \Gamma) \), the NoN model is constructed as the output of GOSPEL.

### 2.3 Computation of GOSPEL

To solve the GOSPEL optimization problem, Eq (1), we first vectorize all the kernel matrices. This produces an \( n^2 \)-dimensional vector associated with the response network, and \( n^2 \)-dimensional vectors corresponding to the predictor networks. This is the same problem setting of the graph guided generalized fused lasso (Chen et al., 2012). G3FL, with \( n^2 \) samples and \( p \) features. Therefore, we employ G3FL to solve our optimization problem.

Regularization parameters \( \lambda_1^{(k)} \), \( \lambda_2^{(k)} \), threshold \( \tau \), and kernel parameter \( \gamma \) are decided by the Bayesian Optimization (Mockus, 2012). We apply the Bayesian Information Criterion (BIC) (Schwarz, 1978) as an acquisition function of the Bayesian Optimization. The BIC score for the case where the response is the \( i \)-th network graph is defined as

\[
\text{BIC}^{(i)} = -2\mathcal{L}^{(i)} + \hat{d}^{(i)} \times \log(n^2),
\]

where \( \mathcal{L}^{(i)} \) is the log-likelihood function:

\[
\mathcal{L}^{(i)} = \frac{n^2}{2} \left( \log \left( \frac{2\pi}{n^2} \mathcal{K}^{(i)} \right) - \sum_{j=1}^{p} \mathcal{B}^{(j)} \mathcal{K}^{(i)} \mathcal{B}^{(j)} \right) + 1,
\]

and \( \hat{d}^{(i)} \) is the degree of freedom of the fused lasso (Tibshirani et al., 2005):

\[
\hat{d}^{(i)} = p - \sum \left( \beta_j^{(i)} = 0 \right) - \sum \left( \beta_j^{(i)} = 0 \right) + 1.\]

In the Bayesian Optimization, we select the set of parameter values which minimize the BIC score.

### 3 Results

#### 3.1 Simulations

We generate synthetic data and evaluate the performance of GOSPEL in order to gain insight into feature selection in the regression problem for network graph data. As synthetic data, we prepare three representative complex network models which have different structures: random networks ( Erdös and Rényi, 1959), scale-free networks (Barabási and Albert, 1999) and small-world networks (Watts and Strogatz, 1998). For each network model, 30 predictor networks are prepared so that the first 15 predictors have the following non-linear relationships:

\[
A_{i,j}^{(k)} = \begin{cases} 
\sin \left( A_{i,j}^{(k)} \right) & j = 6, \ldots, 10 \text{ and } k \neq l, \\
\exp \left( -2A_{i,j}^{(k)} \right) & j = 11, \ldots, 15 \text{ and } k \neq l, \\
0 & j = 1, \ldots, 30 \text{ and } k = l.
\end{cases}
\]

Using the 15 predictors, we generate the following two signal types of response networks with signal-to-noise ratio equal to 1.

- Additive type: \( A^{(i)} = \sum_{j=1}^{15} A^{(j)} + \varepsilon^{(i)} \),
- Non-additive type: \( A^{(i)} = \sum_{j=1}^{5} A^{(j)} + \sum_{j=6}^{10} A^{(j)} + A^{(j=1)} + \varepsilon^{(i)} \),

where \( \varepsilon^{(i)} \) indicates the element-wise product.

We run the simulations 100 times for each combination of the network model and the signal type, varying the number of vertices as \( n = \{500, 1000, 2000\} \). We compare GOSPEL to HSIC Lasso (Yamada et al., 2014), one of the feature selection methods by the feature-wise kernelized lasso. We note that HSIC Lasso is not designed for network graph data and cannot be directly applied. In these simulations, HSIC Lasso is applied to the centered and normalized kernel matrices as follows:

\[
\min_{\alpha} \left\{ \mathcal{R} - \sum_{j=1}^{p} \alpha_j \mathcal{K}^{(j)} \mathcal{K}^{(j)} \right\} + \lambda \sum_{j=1}^{p} |\alpha_j|,
\]

where \( \alpha_j \) is the non-zero coefficients estimated by GOSPEL and HSIC Lasso. The results were analyzed for precision, recall and F-measure. Table 2 shows the F-measures calculated for the 18 different settings with varying sample size and network types. Regarding the precision and the recall of the simulation results, see Tables 1 and 2 in the supplement, respectively. The results highlight the efficacy of the graph-guided fused-lasso regularization.

In the cases of random networks, GOSPEL’s estimation performance remains high regardless of the sample size or the signal type. This result may come from the fact that the random network is the simplest network of the three. Compared with the random network, the scale-free and small-world networks are difficult to estimate. Since the scale-free and small-world networks have distinctive structures, all the predictor networks are similar to each other within their group. In the case of the scale-free network, when the signal type is additive, the performance of GOSPEL becomes better as \( n \) grows. On the other hand, when the signal type is non-additive, GOSPEL and the variant of HSIC Lasso perform almost at the same level. The reason for this is that most elements of the response network take similar values due to the element-wise product. Since the operation tends to break the structure of the response network in spite of the predictor networks keeping their structures, estimation becomes excessively difficult. Interestingly, the performance of the non-additive cases is better than that of the additive cases in the simulations of small-world networks. In the non-additive cases, the operation of the element-wise product may work to emphasize the structure of the response network, and this may improve performance here while leading to opposite results in the scale-free cases.

Our simulation results demonstrate that GOSPEL is able to recover the true network structure from given network graph data, and outperforms HSIC Lasso in terms of precision, recall and F-measure in the different settings with varying sample size and network types.

#### 3.2 Real data

ASD is a complex neurodevelopmental disorder driven by a multitude of genetic variants across the genome that appear as a range of developmental and functional perturbations, often in specific tissues and cell types (Vorstman et al., 2017). To construct human brain-region-specific networks associated with the ASD risk genes, we adopt a manner of data-construction introduced in recent studies (Krishnan et al., 2016; Duda et al., 2014). One of the feature selection methods by the feature-wise kernelized lasso
Next, in order to reconstruct the measures the proximity of the ASD risk genes in the neighborhood of all nodes (princeton.edu/), we use a regularized network approach. This network information can be downloaded from the Genome-scale Integrated Analysis of Gene Networks in Tissues (GIANT) web site (princeton.edu/). We use the GIANT network information to evaluate our proposed method. The purpose of our analysis is to investigate how the ASD risk genes may be coupled in each brain region-specific networks and what inter- connectivity structure between these networks can be formed.

The 17 human brain regions are taken from the whole brain: the frontal lobe (the cerebral cortex), the parietal lobe (the cerebral cortex), the temporal lobe (the cerebral cortex), the occipital lobe (the cerebral cortex), the subthalamic nucleus (the basal ganglia), the caudate nucleus (the basal ganglia), the putamen nucleus (the basal ganglia), the amygdala (the limbic system), the nucleus accumbens (the limbic system), the hippocampus (the limbic system), the thalamus (the limbic system), the hypothalamus, the amygdala and the thalamus. The amygdala has the largest number of edges, and thus the thalamus is suggested to be the center of functional connectivity. The thalamus and the subthalamus are highly connected to the amygdala, the hippocampus, the thalamus and the caudate nucleus. The amygdala is suggested to be the center of functional connectivity in Yahata et al. (2016) but is also well studied as a hub in ASD research. Table 3 indicates the enrichment scores and the community ID for each subregion. Enrich score and Com.ID indicate the mean value of enrichment score and the community ID, respectively.

Table 2 shows the consistency between our resulting model and the abnormal functional connections in Table 1 of Yahata et al. (2016), and samples of the evidence associated with each subregion and ASD. See Table 3 in the supplement for further information. For the classification of ASD and typically developed (TD) persons, Yahata et al. (2016) identified the 16 abnormal functional connections using MRI. We investigate the consistency of the subregions studied in Yahata et al. (2016) and those in our experiment. There are five connections corresponding to our experiment: the caudate nucleus and the amygdala, the frontal lobe and the occipital lobe, the hippocampus and the frontal lobe, and the thalamus and the parietal lobe. As shown in Table 4, four out of the five connections are shared.

In the first group, the amygdala has the largest number of the thickest edges, and thus the amygdala is suggested to be the center of functional abnormality in this group, whereas the thalamus acts as the hub. Table 4 supports this suggestion; the amygdala is not only included in the abnormal functional connection in Yahata et al. (2016) but is also well studied as a subregion strongly related to ASD. As stated, the hub of the first group is the thalamus, and this result is analogous to the medical knowledge that the thalamus is an information relay station (hub) between the subcortical areas and the cerebral cortex (Gazzaniga et al., 2009). In addition, some links in the first group, such as hippocampus and the temporal lobe, and the caudate nucleus and the thalamus, are physically close even though we did not consider locational information in this experiment.

The subregion which seems to be representative of the second group is the frontal lobe. As shown in Table 4, the frontal lobe is a well-studied subregion in ASD research. In addition, four out of the five abnormal connections corresponding to our experiment shown above include the frontal lobe. It is remarkable that the resulting model shares three out of the four abnormal connections.
NoN model related to the ASD risk genes

![Diagram of brain regions and their connections related to ASD risk genes.]

The analysis with a real example thus shows that GOSPEL is able to identify the ASD-associated interconnectivity structure between given functional interaction networks.

4 Discussion

In order to improve understanding of brain-specific complex systems related to a disease and to break through the limitation of the network-based analyses which estimate functional single networks, we proposed a "Network of Networks" (NoN) approach inspired by structural equation models. In this paper, we sought to estimate the topological structure of functional interaction networks of human brain region-specific networks associated with ASD risk genes.

To the best of our knowledge, the sparse regression for network graphs in GOSPEL is the first feature selection method where the features are not vectors but instead are network graphs. In order to construct a NoN model, GOSPEL estimates all the predictor networks relevant to the response network even when they have non-linear correlations. All the parameters in GOSPEL are automatically optimized by the Bayesian Optimization based on the HIC. Though there is room for improvement in that GOSPEL cannot reflect the information of the vertices (nodes), the outputted NoN model is interpretable by combining the information on the vertices and the result of community extraction, as shown in Section 3.2.

We demonstrated the effectiveness of GOSPEL in simulations, and tackled the exploration of the NoN model of human brain region-specific networks related to ASD. The result was the successful production of a NoN model which shows the subregions of the brain which relate to ASD and how they functionally relate to each other. This model is consistent with previous ASD research. Although the accuracy of our model is yet untested, we hope that it will provide useful insight for ASD researchers, and that further research will prove its accuracy.

Finally, while this research was limited to the study of subregions of the brain, we believe that GOSPEL will prove useful to other studies seeking to find relationships between illness and bodily organs or regions, therefore it may be of great use to those studying the systems of biology.
Acknowledgements
We would like to thank Makoto Yamada for helpful discussions.

Funding
This research was supported by AMED under grant No. JP180107087.

References