Pairwise Functional Connectivity Estimation in Spinocerebellar Ataxia Type 3 Using Sparse Gaussian Markov Network: Integrating Group and Individual Analyses of rs-fMRI

Faezeh Moradi
Dalhousie University
Faculty of Computer Science
Halifax, Canada
faezeh.moradi@dal.ca

Jennifer Faber
German Center for Neurodegenerative Diseases (DZNE)
Department of Neurology
University Hospital Bonn
Bonn, Germany
jennifer.faber@dzne.de

Carlos R. Hernandez-Castillo
Dalhousie University
Faculty of Computer Science
Halifax, Canada
carlos.hernandez@dal.ca

***Abstract*—Functional connectivity (FC) patterns from resting-state fMRI data provide relevant information for understanding brain function. In this research, we propose a representation of FC patterns in the form of a sparse graphical model that characterizes brain dynamics. Given that neurological processes typically involve specific brain regions interacting with only a few other regions for given tasks, it is reasonable to introduce sparsity into a graphical model to describe brain dynamics better. However, a challenge in constructing individual Gaussian Markov networks for each subject's rs-fMRI data set is the small number of time points in clinical data. This limitation can lead to unreliable and inaccurate estimation of precision metrics. Therefore, we propose a novel three-step approach that reduces the spatial information of the individual-level based on propagating information from group-level to individual-level analysis. In addition, this approach simultaneously considers different sparsity patterns outside each group and similar sparsity patterns inside each group, considering individual variability. We evaluate the effectiveness of our approach by training the model to differentiate between spinocerebellar ataxia patients (SCA3) and healthy controls. Our results show that our proposed method outperforms other state-of-the-art methods in performance and interoperability. Moreover, our approach gives an explainable subset of FC patterns, which includes information on regions and connections that are conditionally independent and can be used for future studies of SCA type 3 patients.**

***Index Terms*—Functional connectivity, Graphical LASSO, rs-fMRI, Gaussian Markov Network**

I. Introduction

Resting-state functional magnetic resonance imaging (rs-fMRI) studies a subject's spontaneous neural activity while resting and performing no predefined task. The rs-fMRI data comprises a time series of neural activity measurements across three-dimensional spatial coordinates. Functional connectivity (FC) networks refer to regions of the brain that are spatially apart but temporarily correlated [1]. Previous studies have underlined the potential of FC patterns derived from rs-fMRI data as imaging biomarkers in neurodegenerative diseases [2].

Spinocerebellar ataxia type 3 (SCA3) is a neurodegenerative disease and the most common autosomal dominantly inherited ataxia worldwide [3], [4]. The research in [5] reveals the potential for FC being a biomarker for SCA3. In addition, the accurate estimation of FC on an individual subject level is a crucial prerequisite in recognizing disease-specific FC patterns.

Model-driven approaches like seed-based analysis require prior knowledge to define the seed [6]. However, selecting the appropriate seed for rs-fMRI studies poses a considerable challenge. Data-driven approaches such as independent components analysis [7] make it difficult to interpret each component's latent spatial map and determine pairwise FC between specific brain regions. Existing methods that estimate pairwise FC based on statistical conditional dependencies between regions often fail to adjust the sparsity in FC patterns to account for inter-subject variability within each group. Bearing these challenges in mind, our work employs probabilistic graphical modeling to estimate pairwise FC networks for each subject.

Due to the complexity of the neural activity, the exploration of statistical dependencies are usually modeled probabilistically under the assumption of Gaussian distribution [8].

In our proposed method, based on the assumption that neural activity across multiple regions of the brain follows a multivariate Gaussian distribution, we use a sparse Gaussian Markov network (SGMN) model to represent the pairwise FC estimation for each subject. We evaluate this new representation of rs-fMRI data by training an individual-level SGMN for SCA3 detection.

In this paper, we make the following contributions:

1) Introducing a method that enables the application of SGMN models to rs-fMRI clinical data, particularly with short time series.
2) Evaluating SGMN for estimating individual-level pairwise FC with the aim of SCA3 detection.

II. Related Work

Functional connectivity patterns allow us to study the temporal statistical dependencies between brain regions, commonly defined by correlation-based statistics such as Pearson correlation coefficients. However, this pairwise approach fails to capture the conditional dependency between regions when determining the correlation value of two regions without considering information from other regions.

Probabilistic graphical models provide a robust approach for capturing complex dependencies between regions of the brain [9]. Therefore, we model the brain network as a graph where the nodes represent the regions of interest, and the edges represent functional connectivity between regions of interest. The inverse covariance matrix of multivariate normal distributions is advantageous since it allows the imposition of sparsity in a graph [10]. Furthermore, it is a well-studied phenomenon that during a neurological process, specific brain regions interact with only a few other regions [11], [12].

In the group graphical Least Absolute Shrinkage and Selection Operator (LASSO) approach [13], an additional regularization parameter imposes the sparsity patterns that are shared among different groups. However, this approach does not consider both different sparsity patterns outside each group and similar sparsity patterns inside each group.

As an alternative, previous reports have demonstrated the use of Markov networks on fMRI data. In [14], Huang et al. use graphical LASSO in a sparse graphical model to estimate FC as a biomarker for Alzheimer's disease. Narayan et al. [15] used a Markov network to model the brain functional connectivity of rs-fMRI data to investigate how FC patterns differ between healthy subjects and patients with Autism. However, these approaches do only consider individual variability, and do not take group-level information into account. In our approach, we incorporated sparsity into the graphical model of FC using the Graphical LASSO. Additionally, we incorporate the group-level information by pruning connections and regions on the Individual-level graphical model, which allows for further improvement in the reliable estimation of FC.

III. Methodology

We use the undirected graph $G = (\nu, \varepsilon)$ as a mathematical representation to model our network [16]. The nodes (ν) represent brain regions, and the edges (ε) represent the FC strength between regions. We built our model based on two main assumptions: (1) each time point of the rs-fMRI time series is independent and identically distributed, and (2) the neural activity across various brain regions follows a multivariate Gaussian distribution.

A. Sparse Gaussian Markov Network (SGMN) on rs-fMRI Data

For the given undirected graph G, a set of variables $\{x_1, ..., x_{ROI}\}$, representing the time signals of regions of interest, form a Markov network when they satisfy the pairwise Markov property, which says that any two non-adjacent random variables are conditionally independent given all other variables:

$$x_u \perp x_v | x_{\nu - \{u,v\}} \tag{1}$$

Assuming that neural activity of brain regions follows a multivariate Gaussian distribution $x \sim N(\mu, \Sigma)$ defined by n dimensional mean vector μ and a symmetric $n \times n$ covariance matrix Σ. Here n represents the number of regions, and in the precision matrix (i.e., inverse covariance matrix) Σ^{-1}, zero values in the off-diagonal entries represent conditional independence. With this in mind, the problem of determining the Gaussian Markov network will be diminished to estimating the precision matrix. A straightforward metric to estimate our Gaussian Markov network is to measure the probability of our data given the model with likelihood function Eq. (2), where the goal will be to maximize the penalized likelihood function, formulated as:

$$\max_{\Omega} log(|\Omega|) - tr(S\Omega) - \lambda ||\Omega|| \tag{2}$$

Where S is the empirical covariance matrix, Ω is the precision matrix, $tr()$ is the trace function, and λ is the sparse regularization parameter [17].

To solve this optimization problem, we use a single graphical LASSO (SGL) [18] to estimate the precision matrix. SGL ensures a positive and symmetric precision matrix, which we utilize in its normalized form to represent the estimated FC of the SGMN model.

B. The Proposed Three-Step Model

The main limitation to apply Gaussian Markov networks on individual rs-fMRI datasets is the low number of temporal points typical of clinical data. The ratio between sample size and the number of variables significantly impacts graphical LASSO's performance. Consequently, this can result in unreliable and inaccurate estimations of precision metrics. Therefore, we propose a novel three-step approach that reduces the spatial information at the individual-level by propagating information from group-level to individual-level analysis. Fig. 1 illustrates a general overview of the proposed three-step method.

1) Pre-processing: Working at the voxel scale, given the low temporal resolution of the subject's rs-fMRI scan, is computationally challenging when using machine learning techniques. This is well-known as the small n, large p problem, where n is the number of instances and p is the number of features [19]. The Automated anatomical labeling (AAL) atlas provides a well-established anatomical parcellation of the brain designed for MRI analysis [20]. Therefore, we use the AAL [21] atlas that subdivides the brain into 120 regions of interest (ROIs) to extract the time series for each ROI. With 340 time steps for each subject, the raw rs-fMRI data results in a matrix of 340×120 for each subject's record. Fig. 2 illustrates an overview of the preprocessing step.

We normalize the data to a mean of zero and a standard deviation of one for each subject over the time series of each

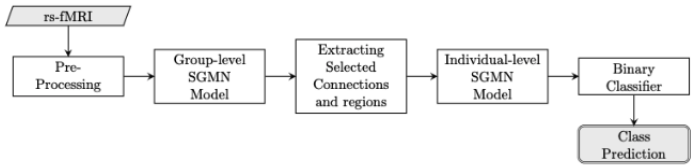


Fig. 1: Building the individual-level Sparse Gaussian Markov Network (SGMN) model for the rs-fMRI classifier based on the information we pass from the group-level SGMN model in three steps: First, the group-level SGMN model is built by aggregating the rs-fMRI data of all subjects in each group. Second, the selected connections and regions that best represent the group-level SGMN model are extracted using a thresholding technique. Finally, the individual-level SGMN is constructed based on the pruned regions and connections.

region separately. We then randomly split the data into training and test sets, in which 70% of the data, including patients and controls, is used for training. Given the assumption that time series data points are independent, we have $340 \times X$ samples for our training data set and $340 \times Y$ samples for the test data set (where 340 is the number of time steps in each record, X is the number of subjects in the training set, and Y is the number of subjects in the test set).

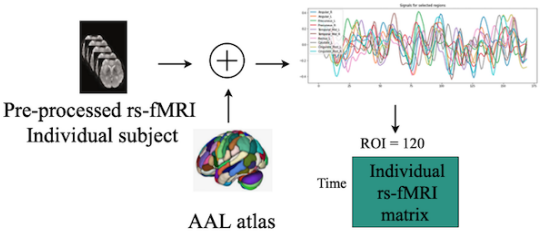


Fig. 2: We used the AAL atlas within the rs-fMRI preprocessing pipeline for the definition of in total 120 anatomical ROIs

2) Group-level SGMN model: This step aims to build a group-level SGMN and extract the subset of regions and connections that best represent each group. For this reason, we aggregate the rs-fMRI data of all subjects in each group, known as the group-level data, and train the SGMN model on 70% of the full data. The aggregated data will have the size $340N \times 120$, where N depends on whether we are building the training or test samples of the groups. We build the graphical model by training two SGL models, one for the healthy control group (HC) and one for the patient group (SCA3). For testing the model, we calculate the likelihood of each new subject's data belonging to each of the built group-level SGMN models. Fig. 3 illustrates the steps of the group-level analysis.

3) Extracting selected connections and regions: From the group-level SGMN model, through a connection-pruning block, we acquire a subset of connections that strongly distinguishes between two groups. This process has two main steps: (1) removing weak connections and (2) applying thresholding the difference matrix of the final FC patterns of the two

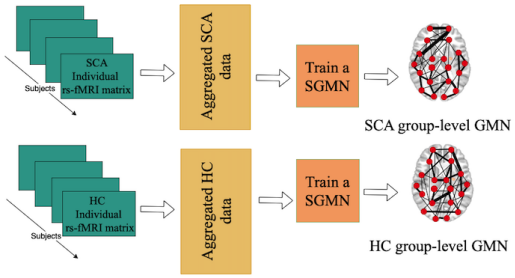


Fig. 3: The group-level Sparse Gaussian Markov Network (SGMN) model aggregates the data for each group. Subsequently, based on this aggregation, one SGMN model is built for the healthy control group (HC), and another is built for the patient group (SCA3).

groups. We tune the threshold value based on the individual-level classification performance. Based on the threshold value, connections that did not allow differentiation between the two groups are pruned from both group-level SGMN models. After the connection-pruning step, the resulting regions and connection of the HC and SCA group were combined, and this combined set is used for further analyses. In this approach, we impose sparsity at the individual level based on existing sparsity patterns between groups and similar sparsity patterns inside groups. We pass the summary of group-level information as an explainable subset of FC patterns to build our individual-level SGMN model.

4) Individual-level SGMN model: In this step, we construct one SGMN model for each subject to represent FC patterns to train our classifier. The graphical model of each subject is built based on the regions selected in the group-level SGMN model. On the updated subject's data, we train a separate SGMN for each subject and estimate the FC with a precision matrix as its new representation. Next, we mask the precision matrix to the selected number of connections to impose final sparsity patterns inside each group (Fig. 5). Finally, this allows the establishment of individual FC pattern representations, which can subsequently be used for the classification task.

After this step, the new representation of the individual-level FC pattern is ready for SCA3 detection.

We consider the same sparsity level when training the

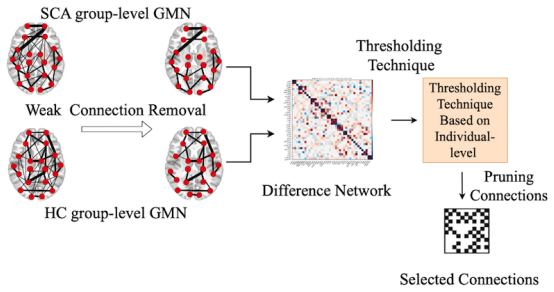


Fig. 4: The Connection-pruning block, in which selected connections and regions are extracted based on the group-level model. Connections with insignificant roles in differentiating the two groups are pruned. The union of the remaining connections is then used to build the explainable subset of FC patterns, which is then used in turn to build the individual-level model.

SGMN for each subject. We tune the sparsity regularization parameter on the training data set over the same range we used for our group-level analysis. After extracting the new representation for each subject and regularization parameter, we evaluate the sparsity level with an external classifier.

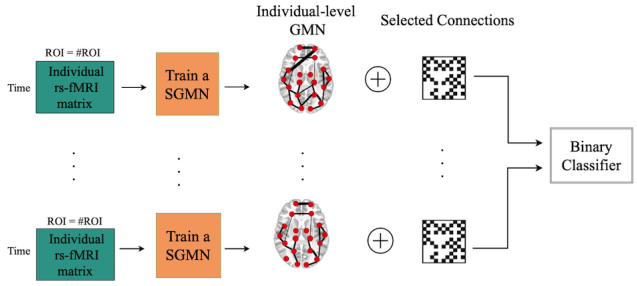


Fig. 5: The Individual-level Sparse Gaussian Markov Network model, in which each subject's data is constrained based on the selected regions, to train for FC estimation. We use the selected connections to impose the final sparsity pattern. The data is ready for training with a binary classifier at this step.

IV. Experiments and Results

A. rs-fMRI Data

Using data from the European Spinocerebellar Ataxia Type 3/Machado-Joseph Disease Initiative (ESMI) cohort study, we performed our analysis on datasets from 42 ataxic SCA3 mutation carriers and 36 age- and sex-matched healthy controls from eleven European and US sites. MRI datasets were acquired on 3T Siemens (Erlangen, Germany) scanners (5 Prisma, 1 Skyra) operating under Syngo MR E11 software and using body coil transmission and a 32-channel receive array.

B. Baseline

To validate our approach, we compare the results with two baselines. These baselines differ in how they define the representation of the FC pattern of rs-fMRI data for training a classifier to distinguish between healthy subjects and patients. In the first baseline, we use the rs-fMRI data, assuming that each time point is an independent sample; we will have a training data size of $340 \times X$ samples with 120 feature sets. For the second baseline method, we calculated a FC matrix using a pairwise Pearson correlation. Currently, this is a commonly used approach to represent functional connectivity. In this approach, we have a training data set of X samples with 7,140 features.

C. Classifiers

In order to evaluate the different approaches, we considered four different traditional machine learning techniques: Logistic Regression (LR), Random Forest (RF), Support Vector Machine (SVM) with both linear and Radial Basis Function (RBF) kernel, and K-nearest neighbors (KNN). To implement these models, we use the default implementation provided by scikit-learn [22]. We calculate the sensitivity, specificity, and overall performance metrics to evaluate different approaches. We define the performance as the average of sensitivity and specificity.

D. Hyper-Parameters

The threshold hyper-parameter controls the degree of sparsity based on different FC patterns between the two groups. In the group-level approach, we consider separate sparsity regularization parameters (λ_c for the HC group and λ_p for the SCA group) to control the sparsity level of each group's graphical model. In the hyper-parameter tuning algorithm, we perform five-fold cross-validation on training subset data of five subjects. We selected the best model based on the performance among all possible combination values. Among all possible pairwise values in the range [0.001, 0.003, 0.005, 0.007, 0.01, 0.03, 0.05, 0.07, 0.1, 0.3, 0.5, 0.7] for sparsity regularization parameters, the selected model has the value of 0.07 for λ_p and value of 0.07 for λ_c.

In the connection-pruning block, after normalizing the estimated precision matrix to -1.0 and $+1.0$, giving the partial correlation values, we remove the weak connections with a threshold of 0.1 on absolute values. The threshold value was determined based on the data variance. For the threshold mechanism, we decide on the threshold value based on its performance on the individual-level GMN model with fixed sparse regularization of 0.01 for both groups.

Table I shows the effect of different threshold changes on the number of pruned connections and their influence on the performance of the individual-level SGMN model classification. The results are averaged over 40 runs on the validation data set. The highest performance for both groups was achieved at a threshold of 0.11. Based on this step, the total number of 60 connections and 70 regions is defined as the union of connections across all connections of two groups. On the updated subject's data, now containing 70 selected ROIs, a separate SGMN is trained for each subject. Each subject's data has a size of 340×70, where only the 60 selected connections

out of the 2,415 possible connectivity values between the 70 regions are considered.

TABLE I: The effect of the different threshold values of the difference matrix on the selected connections (CONN) and regions (ROI). The model is evaluated based on its performance of the individual-level SGMN model on a fixed sparse regularization parameter of 0.07 for both groups.

Model	Threshold	CONN	ROI	Sensitivity	Specificity
1	0.10	61	72	0.71 ± 0.10	0.66 ± 0.13
2	**0.11**	**60**	**70**	**0.73 ± 0.09**	**0.80 ± 0.11**
3	0.12	42	56	0.78 ± 0.11	0.53 ± 0.17
4	0.13	29	47	0.77 ± 0.10	0.63 ± 0.12

To build the individual-level SGMN model, we validate the model using 70 percent of the data randomly selected from each class, across a range of different values for sparse regularization parameters. The best model's sparse regularization parameter has the values of 0.3 and 0.1 for λ_p and λ_c, respectively with a sensitivity of 0.95 ± 0.05 and specificity of 0.92 ± 0.07 on validation data set. We consider a sparser graphical model of all subjects within each group to have the same sparse regularization parameter for estimating each individual's SGMN model.

TABLE II: The effect of different sparsity parameter on Individual-level SGMN model on the validation data set. Results show that the sparser model for the SCA group will perform better in classification.

Model	λ_p	λ_c	Sensitivity	Specificity	Performance
1	0.1	0.3	0.88 ± 0.07	0.93 ± 0.06	0.90 ± 0.04
2	0.3	0.3	0.86 ± 0.07	0.84 ± 0.01	0.85 ± 0.05
3	**0.3**	**0.1**	**0.95 ± 0.05**	**0.92 ± 0.07**	**0.93 ± 0.05**
4	0.1	0.1	0.83 ± 0.11	0.84 ± 0.11	0.84 ± 0.05

E. Results

In the last step of our proposed method, where we build the final model, we use the selected model and train it on 70% of the entire data set after validation. To evaluate our model, we used the classifiers described in Section IV (part c) and presented the best results for each classifier. Among these classifiers, SVM with a linear kernel has the highest performance, except in the second baseline, where linearity was not viable, resulting in SVM with an RBF kernel demonstrating superior performance as the best alternative.

Table III shows a comparison of the proposed method to the baseline models. For the baseline method, from the full Pearson correlation matrix, we used the first 60 selected connections based on the recursive feature elimination (RFE) technique with a linear SVM model for the classifier.

TABLE III: Comparison of different models for SCA3 detection on the test set

Model	Sensitivity	Specificity	Performance
Individual-level SGMN model	0.93 ± 0.06	0.95 ± 0.07	0.93 ± 0.055
Group-level SGMN model	0.71 ± 0.17	0.96 ± 0.07	0.83 ± 0.12
First baseline model with rs-fMRI	0.72 ± 0.03	0.55 ± 0.03	0.63 ± 0.03
Second baseline model with Pearson correlation	0.82 ± 0.10	0.82 ± 0.08	0.82 ± 0.09

DISCUSSION

This paper demonstrates that SGMN outperforms other FC representations in terms of overall classification performance. In the experiments where we compare the different FC estimations on the test set (Table II), the overall performance of the individual-level SGMN is the highest. The performance of the second baseline is as close as the performance of the group-level SGMN. However, results from the second baseline do not exclude the confounding effect of another brain region.

As shown in Table III, the SVM with a linear kernel performed the best among all classifiers for both individual-level and group-level SGMN, as well as the second baseline with Pearson correlation. However, with the first baseline based on rs-fMRI data, the data is not linearly separable, and the SVM with RBF kernel performs best among other classifiers.

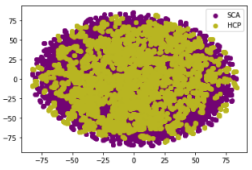
Table II shows the importance of different sparsity levels for each group in improving performance. Of particulate note is that the related performance (sensitivity or specificity) is higher for the group with a higher sparsity. Another strength of our proposed method lies in its ability to control each group's sparsity level independently for considering the cases in which patients and controls have different underlying sparsity patterns.

In Figs. 6 and 7, we use the t-distributed Stochastic Neighbor Embedding (t-SNE) [23] to compare the different representations of FC. In Fig. 6, each point represents a single time point from the rs-fMRI: in this case, it is very hard to separate the two groups since they overlap significantly. In Fig. 7, each point represents a subject after applying our method with 60 selected group-level functional connections: in this case, the two groups are more likely to be easily separated. Also, we can observe that the individual-level SGMN representation of FC is more effective at providing separability than the Pearson correlation representation.

As a by-product of our research, our approach produces an explainable subset of FC patterns, which includes information on regions and connections that are conditionally independent and can be used for future studies of SCA3. However, discussing these connections is out of the scope of this report.

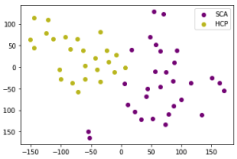
The t-SNE visualization of samples for various FC representations highlights the inseparability of SCA (purple) and HC (yellow) groups. Each point depicts the FC pattern of the individual subject's rs-fMRI data. Notably, FC estimation with individual-level SGMN (b) is more effective in delineating separable groups in high-dimensional data than Pearson

Fig. 6: The t-SNE representation of the first baseline on rs-fMRI data, illustrating the inseparability of SCA (purple) and HC (yellow) groups. Each point represents a single time point.

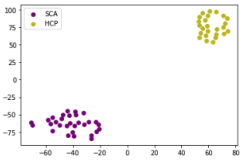


correlation (a).

Fig. 7: The t-SNE visualization of samples for different FC representations, illustrating the inseparability of SCA (purple) and HC (yellow) groups. Each point depicts the FC pattern of each subject's rs-fMRI data. Note that FC estimation with individual-level SGMN (b) is more effective in delineating separable groups in high-dimensional data than Pearson correlation (a).



(a) Second baseline: Pearson correlation



(b) Individual-level SGMN representation

Conclusion

This paper proposes a three-step approach for FC estimation in a sparse Gaussian Markov Network framework by enforcing group-level sparsity patterns. We evaluated our approach by training a SGMN model for SCA3 detection. Our method finds the most critical connections between brain regions to make detecting the disease easier for a machine learning classifier.

As shown by our experiments, our approach displayed better performance when compared to other state-of-the-art methods for detecting SCA3 directly from the patients' rs-fMRI, even with a low number of time points.

An essential finding of this work was the importance of imposing sparsity inside the GMNs: this made our method much more efficient than other state-of-the-art methods while displaying superior performance. Another advantage of sparse graphic modeling is that it is closer to the actual biological behavior of the brain and leads to fewer connections, making it easier to interpret. In addition, the final sparsity pattern takes into account distinct sparsity patterns across groups, while also accounting for shared sparsity patterns within each group, with the consideration of individual variability.

References

[1] Ferreira, Luiz Kobuti, and Geraldo F. Busatto. "Resting-state functional connectivity in normal brain aging." Neuroscience and Biobehavioral Reviews 37, no. 3 (2013): 384-400.

[2] Hohenfeld, Christian, Cornelius J. Werner, and Kathrin Reetz. "Resting-state connectivity in neurodegenerative disorders: Is there potential for an imaging biomarker?." NeuroImage: Clinical 18 (2018): 849-870.

[3] Hernandez-Castillo, C. R., Diaz, R., Campos-Romo, A., & Fernandez-Ruiz, J. (2017). Neural correlates of ataxia severity in spinocerebellar ataxia type 3/Machado-Joseph disease. Cerebellum & ataxias, 4(1), 7.

[4] Schöls, Ludger, Peter Bauer, Thorsten Schmidt, Thorsten Schulte, and Olaf Riess. "Autosomal dominant cerebellar ataxias: clinical features, genetics, and pathogenesis." The Lancet Neurology 3, no. 5 (2004): 291-304.

[5] Guo, Jing, Zhouyu Jiang, Xinyuan Liu, Haoru Li, Bharat B. Biswal, Bo Zhou, Wei Sheng et al. "Cerebello-cerebral resting-state functional connectivity in spinocerebellar ataxia type 3." Human Brain Mapping 44, no. 3 (2023): 927-936.

[6] Biswal, Bharat, F. Zerrin Yetkin, Victor M. Haughton, and James S. Hyde. "Functional connectivity in the motor cortex of resting human brain using echo-planar MRI." Magnetic resonance in medicine 34, no. 4 (1995): 537-541.

[7] V.D. Calhoun, T. Adali, G.D. Pearlson, J.J. Pekar. (2001) Spatial and temporal independent component analysis of functional MRI data containing a pair of task-related waveforms. Hum.Brain Mapp. 13, 43-53.

[8] Friston, K. J. (2011). Functional and effective connectivity: a review. Brain connectivity, 1(1), 13-36.

[9] M. J. Wainwright, M. I. Jordan, et al. Graphical models, exponential families, and variational inference. Foundations and Trends® in Machine Learning, 1(1–2):1–305, 2008. Irina Rish and Genady Ya Grabarnik. Sparse modeling : theory, algorithms, and applications. Chapman & Hall/CRC machine learning & pattern recognition series. Boca Raton, FL : CRC Press : Taylor & Francis Group, 2015., 2015.

[10] N. Meinshausen and P. Bu€hlmann, "High-dimensional graphs and variable selection with the lasso," Ann. Statist., vol. 34, no. 3, pp. 1436–1462, 2006.

[11] Stam, Cornelis J., B. F. Jones, Guido Nolte, Michael Breakspear, and Ph Scheltens. "Small-world networks and functional connectivity in Alzheimer's disease." Cerebral cortex 17, no. 1 (2007): 92-99.

[12] E. T. Rolls and A. Treves, "The relative advantages of sparse ver- sus distributed encoding for associative neuronal networks in the brain," Netw. Comput. Neural Syst., vol. 1, no. 4, pp. 407–421, 1990.

[13] Danaher, Patrick, Pei Wang, and Daniela M. Witten. "The joint graphical lasso for inverse covariance estimation across multiple classes." Journal of the Royal Statistical Society Series B: Statistical Methodology 76, no. 2 (2014): 373-397.

[14] S. Huang et al., "Learning brain connectivity of alzheimer's disease by sparse inverse covariance estimation," NeuroImage, vol. 50, no. 3, pp. 935–949, 2010.

[15] Narayan, Manjari, Genevera I. Allen, and Steffie Tomson. "Two sample inference for populations of graphical models with applications to functional connectivity." arXiv preprint arXiv:1502.03853 (2015).

[16] Koller, Daphne, and Nir Friedman. Probabilistic graphical models: principles and techniques. MIT press, 2009.

[17] Friedman, J., Hastie, T., and Tibshirani, R. (2007). Sparse inverse covariance estimation with the Graphical Lasso. Biostatistics, 9(3):432–44b.

[18] Friedman, Jerome, Trevor Hastie, and Robert Tibshirani. "Sparse inverse covariance estimation with the graphical lasso." Biostatistics 9, no. 3 (2008): 432-441.

[19] Friedman J, Hastie T, Tibshirani R. The elements of statistical learning. New York: Springer series in statistics; 2001.

[20] Rolls, Edmund T., Chu-Chung Huang, Ching-Po Lin, Jianfeng Feng, and Marc Joliot. "Automated anatomical labelling atlas 3." Neuroimage 206 (2020): 116189.

[21] Tzourio-Mazoyer, Nathalie, Brigitte Landeau, Dimitri Papathanassiou, Fabrice Crivello, Octave Etard, Nicolas Delcroix, Bernard Mazoyer, and Marc Joliot. "Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain." Neuroimage 15, no. 1 (2002): 273-289.

[22] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: Machine learning in Python, the Journal of machine Learning research 12 (2011) 2825–2830.

[23] Van der Maaten, Laurens, and Geoffrey Hinton. "Visualizing data using t-SNE." Journal of machine learning research 9, no. 11 (2008).