Exploring latent networks in resting-state fMRI using voxel-to-voxel causal modeling feature selection

Hassan Baker  
BakerH@udel.edu and Austin J. Brockmeier  
AJbrock@udel.edu

Department of Electrical and Computer Engineering, University of Delaware, Newark, Delaware, USA

Abstract

Functional networks characterize the coordinated neural activity observed by functional neuroimaging. The prevalence of different networks during resting state periods provide useful features for predicting the trajectory of neurodegenerative diseases. Techniques for network estimation rely on statistical correlation or dependence between voxels. Due to the large number of voxels, rather than consider the voxel-to-voxel correlations between all voxels, a small set of seed voxels are chosen. Consequently, the network identification may depend on the selected seeds. As an alternative, we propose to fit first-order linear models with sparse priors on the coefficients to model activity across the entire set of cortical grey matter voxels as a linear combination of a smaller subset of voxels. We propose a two-stage algorithm for voxel subset selection that uses different sparsity-inducing regularization approaches to identify subject-specific causally predictive voxels. To reveal the functional networks among these voxels, we then apply independent component analysis (ICA) to model these voxels’ signals as a mixture of latent sources each defining a functional network. Based on the inter-subject similarity of the sources’ spatial patterns we identify independent sources that are well-matched across subjects but fail to match the independent sources from a group-based ICA. These are resting state networks, common across subjects that group ICA does not reveal. These complementary networks could help to better identify neurodegeneration, a task left for future work.

1. Introduction

Functional magnetic resonance imaging (fMRI) signal is among the most popular techniques for identifying functional connections (FC) across regions of the brain. Resting-state fMRI (rs-fMRI) is a task-free neuroimaging paradigm. It reveals the neural activities while the subject is not doing anything in particular. The importance of studying rs-fMRI lies in its ability to reveal the hidden functional relationships among brain regions, providing baselines or prototypes of the expected neural activities for healthy subjects versus unhealthy subjects who suffer from neurodegenerative diseases such as schizophrenia (Karbasforoushan and Woodward, 2012; Baker et al., 2014) and Alzheimer’s disease (de Vos et al., 2018). A better understanding of rs-fMRI may lead to an early-stage detection for neural diseases and also other kinds of diseases Kucikova et al. (2021).

Often fMRI processing techniques focus on the statistical correlation of the fMRI time series or are based on seed selection. The latter can create variation among the selected seeds (Lv et al., 2018). An alternative approach is to try to model the causal relationship (in the sense of Granger) among voxels directly. There are several works of modeling brain activities using causal structure. Roebroeck et al. (2005) used causality for identifying the influence of any selected regions of interest during complex visuomotor tasks. Liao et al. (2010) were able to find specific causal influences among the detected resting-state networks using causal structure applied on rs-fMRI. Liao et al. (2011) applied multivariate Granger causality analysis and graph theory on rs-fMRI to identify pivotal hubs and their interactions with each other. Their work relied on using the pooled activity of voxels in predefined regions of interest. DSouza et al. (2017) explored causality structure on resting-state data, using principal component analysis (PCA) to overcome the problem of underdeterminedness. Hamilton et al. (2011) and Gao et al. (2020) were able to identify unique neural patterns for major depressive disorder and increases and decreases of Granger causality within different brain regions for Schizophrenic patients.
In this paper, we focus on overcoming the limitations of other techniques mentioned above. In particular, we provide a voxel-level methodology that enables whole-brain voxel subset selection, without either dimensionality reduction, seed selection, or pooling within regions of interest. Firstly, we focus on the voxel subset selection procedure based on the ability of the voxels to represent the dynamics in the brain under a sparsity assumption. We then perform independent component analysis (ICA) on the selected voxels. ICA optimizes a demixing matrix to linearly combine these voxels revealing the underlying sources that account for higher-order dependency among the selected voxels.

2. Data and Methodology

The resting-state fMRI data is obtained from Risk et al. (2021). We choose the dataset with the highest multiband (i.e., MB12) as it gives the highest temporal resolution at ~2 Hz. A subset of 24 healthy subjects are chosen and their preprocessing is done using fmriprep, 6 subjects are held-out for future analysis. The rest of the analysis is done on the grey-matter cortical voxels. To evaluate the causal models, the fMRI time series (688 time points corresponding to ~352 s) for each subject is divided with roughly the first 80% (550 points) for training, the next 10% (68 points) for validation, and the final 10% (70 points) for testing.

Let $S$ denote the set of $|S|$ subjects, where $|\cdot|$ denotes the cardinality of the set. For a subject $s \in S$, the fMRI training data forms a matrix, $X(s) \in \mathbb{R}^{V(s) \times T}$, $V(s)$ is the number of cortical grey voxels per subject and $T$ is the number of time points. When the modeling is applied to any subject we will drop the subject indices and refer to $X(s)$ and $V(s)$ by $X$ and $V$, respectively.

2.1. Voxel Subset Selection

The goal of voxel subset selection is to find a subset of voxels that are predictive of the fMRI signal in the cortical grey matter voxels for a given subject. A key motivation for it is to overcome the computational complexity when performing a voxel-to-voxel analysis that considers the temporal dynamics of a large number of grey-matter voxels. Due to the number of voxels compared with the number of time frames, robust voxel subset selection is non-trivial.

A two-stage approach that combines a fine-grained region based atlas and sparse linear regression is used to find a predictive subset. In the first stage, the 1000 region Schaeffer atlases (Schaefer et al., 2018) is used to organize the cortical regions of the brain into different groups based on their functional connectivity (FC). The 1000 region atlas is spatially fine-grained which helps selecting representative voxels from different regions of the cortical region. Let $R_i \in \{1, \ldots, V\}^V_i \subset \mathbb{N}^V_i, i \in \{1, \ldots, 1000\}$ be a region represented as a vector of voxel indices associated to the $i$-th region of the atlas, and $V_i$ is the number of voxels in the region. We identify a subset of voxels within each region by their ability to predict the activities of the rest of the cortical region. A single-lag predictive model is used $X_{t}^{\text{lag}} \approx W_{R_i}X_{t-1}^{\text{lag}}$, where $X_{t}^{\text{lag}} \in \mathbb{R}^{(V-V_i) \times T-1}$ is a time series starting at the second time index of all voxels except the voxels from the region $R_i$, $X_{t-1}^{\text{lag}} \in \mathbb{R}^{V_i \times T-1}$ is the matrix of delayed time series that starts from the first time index and ends at the second to last for voxels within region $R_i$, and $W_{R_i} \in \mathbb{R}^{(V-V_i) \times V_i}$ is the mapping matrix of the previous neural activities in region $R_i$ to the next neural activities in the rest of the cortical regions. To summarize, we develop predictive models that predict the activity in other cortical regions based on the activity in region $R_i$ at the previous time point.

We assume that $W_{R_i}$ has a special sparsity structure: most of the columns are zeros as they do not contribute to the predictive model. Thus, to estimate the matrix $W_{R_i}$ we add an $\ell_{2,1}$ penalty to the sum of squared errors,

$$W_{R_i}^* = \arg \min_{W \in \mathbb{R}^{(V-V_i) \times V_i}} \|X_t^{\text{lag}} - WX_{t-1}^{\text{lag}}\|^2_F + \lambda_{2,1}\|W\|_{2,1},$$

where $\|\cdot\|^2_F$ denotes the Frobenius norm and $\|W\|_{2,1} = \sum_{ij} \sqrt{\sum_j W_{ij}^2}$. The solution to this optimization problem is obtained using the method detailed by Nie et al. (2010). The selection of the regularization parameter $\lambda_{2,1}$ is done by performing the analysis over 20 equally line spaced points in the interval $[\lambda_{21}^{\text{max}}, 0.001\lambda_{21}^{\text{max}}]$, where $\lambda_{21}^{\text{max}}$ is the regularization that returns a zero solution. After fitting the model, an unshrunk model is obtained by fitting an ordinary least squares model on the nonzero columns. The parameter $\lambda_{21}$ that achieves the minimum validation mean squared error (based on the unshrunk ordinary least square model coefficients) is chosen.

1. Table 1 in the Appendix shows the number of cortical grey matter voxels per subject.

2. Fig. 3 in the Appendix shows a workflow diagram that summarizes the approach.
This voxel selection procedure aims to select voxels within a group. Let supp($W_{R_i}$) be the indices of non-zero columns of $W_{R_i}$. Abusing vector indentation notation, the corresponding set of voxel indices for region $i$ are $R_i(supp(W_{R_i}^*))$. Let $V_{S1} = R_1(supp(W_{R_1}^*)) \cup \cdots \cup R_2(supp(W_{R_{1000}}^*))$ be the set of selected voxels from the first stage with $V_{S1} = |V_{S1}|$.

In the second stage, we assume there are voxels across regions that are redundant or unnecessary for predicting any given voxel. Hence, we solve the following LASSO model to choose a subset of the previously selected voxels to explain each voxel, $j \in \{1, \ldots, V\}$, independently:

$$w_j^* = \arg\min_{w \in \mathbb{R}^{|V_{S1}|}} \|x_j^T - (X_{S1}^T)^T w\|_2^2 + \lambda_j \|w\|_1,$$

where $x_j^T \in \mathbb{R}^{T-1}$ is the fMRI time series for voxel $j$, $X_{S1}^T = ([x_{j(k)}]_{k \in V_{S1}})^T \in \mathbb{R}^{V_{S1} \times T-1}$ is the selected voxels time series from the first stage, and $W^{**} = [w_1^{**}, \ldots, w_V^{**}] \in \mathbb{R}^{V_{S1} \times V}$ is the mapping matrix from the previous activities of the first-stage selected voxels to all grey matter cortical voxels. The same procedure of the first stage is applied here to select the regularization parameter $\lambda_j$, except that only 10 linearly spaced values in the range $[\lambda_{j}^{\text{max}}, 10^{-3}\lambda_{j}^{\text{max}}]$ are tested. Note that each voxel $j \in \{1, \ldots, V\}$ has its own regularization parameter chosen from a voxel-dependent range. The subset of selected voxels from the second stage is $V = \cup_{j=1}^V v_{S1}(supp(w_j^{**}))$, where $v_{S1} = [k]_{k \in V_{S1}}$.

The subsequent analysis is based on the selected voxels from the second stage along with coefficients from a final ridge regression model $W^{***} = [w_1^{***}, \ldots, w_V^{***}] \in \mathbb{R}^{V_{S1} \times V}$ is the mapping matrix across all voxels, where

$$w_j^{***} = \arg\min_{w \in \mathbb{R}^{|V_{S1}|}, \text{supp}(w) \subseteq \text{supp}(w_j^{**})} \|x_j^T - (X_{S1}^T)^T w\|_2^2 + \mu_j \|w\|_2^2,$$

and $\mu_j$ is selected by the validation mean squared error (MSE). To test the significance of the causality, models are built using 100 time-shuffled versions of the training data, and test set MSEs are obtained. Averaged across subjects, 51% of the cortical grey matter voxels have testing MSEs statistically lower then the MSEs of the shuffled models (p-value $\leq 0.05$). The range across subjects is 38%–75%.

2.2. Analysis of Selected Voxels

Although the selected voxels are assumed to capture the dynamics in the brain, it is not guaranteed that these voxels are statistically independent. Thus, we apply ICA to organize the selected voxels into latent/hidden networks per subject. This also enables a direct comparison with group-based ICA, which combines the data from all subjects assuming common networks across subjects.

As an alternative, we apply ICA (in particular, fastICA, with the logcosh contrast function) directly on the matrix of time-series from the subset of predictive voxels: $(X^{S2})^T = \{x_{j(k)}^T\}_{j \in v} = MS \in \mathbb{R}^{T \times |V|}$, where $M \in \mathbb{R}^{T \times |IC|}$ is the mixing matrix, $|IC|$ is the number of components, $S \in \mathbb{R}^{IC \times |V|}$ is the matrix of independent components (ICs). Using the causal model coefficients $W^{***}$, all voxels in the grey matter cortical regions are approximated as a mixture of these ICs, $X \approx W^{***}S^TM^T = QM^T$. For subject $s \in S$, let $Q^{(s)} \in \mathbb{R}^{V^{(s)} \times |IC|}$ denote the spatial pattern of the ICs across the whole set of grey matter cortical voxels. Since each subject has a different set of cortical grey matter voxels $V^{(s)}$, we still need to define a common space to define the ICs. To do this we project each IC back into the MNI152 space and apply a spatial blur to account for the grey-matter misalignment across subjects, with $IC^{(s)} = \text{blur}(\text{map}_\text{MNI152}(Q^{(s)}))$ for $k \in \{1, \ldots, |IC|\}$. Blurring uses a 3D Gaussian kernel with $\sigma = 3$ voxels.

To compare our method, we apply CanICA (Varouquaux et al., 2010) which is a popular group-level ICA approach that combines the subjects’ fMRI data and extracts the ICs based on this group data in the MNI152 space. We also apply the Gaussian kernel to each group IC $IC^{(k)\text{group}}$, $k \in \{1, \ldots, |IC|\}$. We set the number of IC’s to 20 as it is the number used in previous studies (Schultz et al., 2014; Biswal et al., 2010) and is also the default number of components in the Group ICA toolbox such as GIFT (Calhoun et al., 2001) and CanICA in Nilearn framework.

To test the consistency of our approach, we calculate the inter-subject similarity (IS) for each IC, $IC^{(s)}$, and subject $s'$, which is denoted by $IS(IC^{(s)}, s')$. We define it as the maximum absolute cosine similarity of the $k$-th IC from subject $s$ with each IC from subject. Mathematically,

$$IS(IC_k^{(s)}, s') = \max_{k \in \{1, \ldots, |IC|\}} |\cos \angle(IC_k^{(s)}, IC_{k'}^{(s')})|$$

$$IS(IC_k^{(s)}) = [IS(IC_k^{(s)}, 1), \ldots, IS(IC_k^{(s)}, |S|)] \in \mathbb{R}^{|S|},$$

where $|\cos \angle(\cdot, \cdot)|$ is the absolute cosine similarity. Note that when $s = s'$, $IS(IC_k^{(s)}, s) = 1$. 

3
We also calculate the similarity between the individual ICs and the group ICs, and refer to this as the Individual-Group Similarity (IGS):

$$\text{IGS}(IC(s)_k) = \max_{k' \in \{1, \ldots, |IC|\}} \cos \angle(IC(s)_k, IC_{k'}^{\text{group}})$$.

3. Results

We are interested in the ICs that have low IGS and have high values of IS as they explain consistent functional networks identified through the proposed approach that are distinct from the group-based ICA. Initially, we tested whether the subject-specific ICs that have at least half of the values in IS($IC(s)_k$) larger than the value of IGS($IC(s)_k$).

The vast majority of the subjects satisfy this condition (only 2 subjects failed), which means that the majority of similarity values between individual ICs of different subjects are greater than the similarity value with the group ICs.

We then performed a hierarchical clustering (weighted Euclidean linkage) on the individual ICs. Instead of defining the clusters in terms of the similarity of the ICs directly, we concatenate and compare the IS and IGS values, computing a weighted Manhattan distance with a weight of $|S|$ for the IGS entry and a weight of 1 for the IS entries.

![Figure 1: Hierarchical clustering of subject ICs based on their IS and IGS values. Column labeled 24 is the IGS values.](image)

Fig. 1 shows the clustering performed on the feature vector. Column labeled 24 is the IGS values.

![Figure 2: Coronal slices of individual ICs in cluster D (columns 1–4) and group ICs (column 5).](image)

The dendrogram shows four different clusters: A, B, C, D. Cluster D is the set of ICs that have high IS and low IGS. That is, ICs in cluster D identify consistent spatial patterns of the sources associated to selected voxels but are not similar to any group IC. Notably, all subjects but four subjects have at least one IC in cluster D.

We focus on cluster D to investigate if there are features that distinguish individual ICs from group ICs. We show a coronal slice from individual ICs belonging to cluster D from different subjects in Fig. 2. Even in these slices the similarities between ICs are evident, which reflects the high IS values. For contrast, the most similar group ICs are displayed in the last column with green backgrounds. Clearly, the group ICs are much sparser compared with the individual ICs. The individual ICs found through causal modeling may capture more global functional networks rather than the localized functional networks of group ICA.

4. Conclusion and Future Work

We introduced an automated and scalable voxel selection method based on first-order linear causal models with sparsity inducing regularizations. We show that our method is able to recover unique spatial patterns across subjects compared to a group-based method. Whether these patterns are meaningful in distinguishing healthy versus non-healthy subjects is left for future work.
Acknowledgments

Research was carried out with the support of the University of Delaware Research Foundation. This research was supported in part through the use of Data Science Institute (DSI) computational resources at the University of Delaware, specifically the DARWIN system. The DARWIN computing project at the University of Delaware is supported by the National Science Foundation under Grant No. OAC-1919839. The authors thank the support from the University of Delaware IT Research Computing Group.

References

Justin T Baker, Avram J Holmes, Grace A Masters, BT Thomas Yeo, Fenna Krienen, Randy L Buckner, and Dost Öngür. Disruption of cortical association networks in schizophrenia and psychotic bipolar disorder. *JAMA Psychiatry*, 71(2):109–118, 2014.

Bharat B Biswal, Maarten Mennes, Xi-Nian Zuo, Suril Gohel, Clare Kelly, Steve M Smith, Christian F Beckmann, Jonathan S Adelstein, Randy L Buckner, Stan Colcombe, et al. Toward discovery science of human brain function. *Proceedings of the National Academy of Sciences*, 107(10):4734–4739, 2010.

Vince D Calhoun, Tulay Adali, Godfrey D Pearlson, and James J Pekar. A method for making group inferences from functional MRI data using independent component analysis. *Human Brain Mapping*, 14(3):140–151, 2001.

Frank de Vos, Marisa Koini, Tijn M Schouten, Stephan Seiler, Jeroen van der Grond, Anita Lechner, Reinhold Schmidt, Mark de Rooij, and Serge ARB Rombouts. A comprehensive analysis of resting state fMRI measures to classify individual patients with Alzheimer’s disease. *Neuroimage*, 167:62–72, 2018.

Adora M DSouza, Anas Z Abidin, Lutz Leistritz, and Axel Wismüller. Exploring connectivity with large-scale Granger causality on resting-state functional MRI. *Journal of Neuroscience Methods*, 287:68–79, 2017.

Jie Gao, Dongsheng Zhang, Lei Wang, Wei Wang, Yajuan Fan, Min Tang, Xin Zhang, Xiaoyan Lei, Yarong Wang, Jian Yang, and Xiaoling Zhang. Altered effective connectivity in schizophrenic patients with auditory verbal hallucinations: a resting-state fMRI study with Granger causality analysis. *Frontiers in Psychiatry*, 11:575, 2020.

J Paul Hamilton, Gang Chen, Moriah E Thomason, Mirra E Schwartz, and Ian H Gotlib. Investigating neural primacy in major depressive disorder: multivariate Granger causality analysis of resting-state fMRI time-series data. *Molecular Psychiatry*, 16(7):763–772, 2011.

H Karbasforoushan and ND Woodward. Resting-state networks in schizophrenia. *Current Topics in Medicinal Chemistry*, 12(21):2404–2414, 2012.

Ludmila Kucikova, Jantje Goerdtten, Maria-Eleni Dounavi, Elijah Mak, Li Su, Adam D Waldman, Samuel Danso, Graciela Muniz-Terrera, and Craig W Ritchie. Resting-state brain connectivity in healthy young and middle-aged adults at risk of progressive Alzheimer’s disease. *Neuroscience & Biobehavioral Reviews*, 2021.

Wei Liao, Dante Mantini, Zhiqiang Zhang, Zhengyong Pan, Jurong Ding, Qiyong Gong, Yihong Yang, and Huafu Chen. Evaluating the effective connectivity of resting state networks using conditional Granger causality. *Biological Cybernetics*, 102(1):57–69, 2010.

Wei Liao, Jurong Ding, Daniele Marinazzo, Qiang Xu, Zhengge Wang, Cuiping Yuan, Zhiqiang Zhang, Guangming Lu, and Huafu Chen. Small-world directed networks in the human brain: multivariate Granger causality analysis of resting-state fMRI. *Neuroimage*, 54(4):2683–2694, 2011.

Han Lv, Zhenchang Wang, Elizabeth Tong, Leanne M Williams, Greg Zaharchuk, Michael Zeineh, Andrea N Goldstein-Piekarski, Tali M Ball, Chengde Liao, and Max Wintermark. Resting-state functional MRI: everything that nonexperts have always wanted to know. *American Journal of Neuroradiology*, 39(8):1390–1399, 2018.

Feiping Nie, Heng Huang, Xiao Cai, and Chris Ding. Efficient and robust feature selection via joint $\ell_{2,1}$-norms minimization. *Advances in Neural Information Processing Systems*, 23, 2010.

Benjamin B Risk, Raphiel J Murden, Junjie Wu, Mary Beth Nebel, Arun Venkataraman, Zhengwu
Zhang, and Deqiang Qiu. Which multiband factor should you choose for your resting-state fMRI study? *Neuroimage*, 234:117965, 2021.

Alard Roebroeck, Elia Formisano, and Rainer Goebel. Mapping directed influence over the brain using Granger causality and fMRI. *Neuroimage*, 25(1):230–242, 2005.

Alexander Schaefer, Ru Kong, Evan M Gordon, Timothy O Laumann, Xi-Nian Zuo, Avram J Holmes, Simon B Eickhoff, and BT Thomas Yeo. Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. *Cerebral Cortex*, 28(9):3095–3114, 2018.

Aaron P Schultz, Jasmeer P Chhatwal, Willem Huijbers, Trey Hedden, Koene RA Van Dijk, Donald G McLaren, Andrew M Ward, Sarah Wigman, and Reisa A Sperling. Template based rotation: a method for functional connectivity analysis with a priori templates. *Neuroimage*, 102:620–636, 2014.

Gaël Varoquaux, Sepideh Sadaghiani, Philippe Pinel, Andreas Kleinschmidt, Jean-Baptiste Poline, and Bertrand Thirion. A group model for stable multi-subject ICA on fMRI datasets. *Neuroimage*, 51(1):288–299, 2010.

5. Appendix

We show the schematic diagram for the two-stage voxel selection procedure in Fig. 3. We also show the number of cortical grey matter voxels for each subject in Table 1.

Table 1: The number of cortical grey matter voxels $V(s)$ for each subject.

| Subject | Voxels |
|---------|--------|
| 02      | 97837  |
| 03      | 95915  |
| 07      | 102332 |
| 08      | 99330  |
| 10      | 96809  |
| 11      | 98766  |
| 12      | 92915  |
| 13      | 97803  |
| 14      | 100182 |
| 15      | 96888  |
| 16      | 9885   |
| 17      | 97728  |
| 18      | 97810  |
| 19      | 98337  |
| 20      | 99218  |
| 22      | 98011  |
| 23      | 96135  |
| 24      | 98981  |
| 25      | 95911  |
| 26      | 100367 |
| 27      | 94910  |
| 28      | 98893  |
| 29      | 96896  |
| 30      | 95460  |
Figure 3: In the first stage, we select the representative voxels of region $R_i$ by estimating an $\ell_{2,1}$-norm regularized linear model that predicts the activities of other regions based on one time step in advance of neural activities of region $R_i$. The representative voxels are the ones which correspond to the nonzero columns. In the second stage, the goal is to select voxels from the voxels selected from the first stage. We apply a $\ell_1$-norm regularized linear model (LASSO) that predicts the activity of each cortical voxel based on one time step in advance of neural activities of the first stage selected voxels. The final voxels selection is based on the corresponding voxels of nonzero coefficients. In other words, if voxel has at least a nonzero coefficient that is predictive of any voxels activity, then it will be selected.