

HHS Public Access

Author manuscript Proc SPIE Int Soc Opt Eng. Author manuscript; available in PMC 2017 November 21.

Published in final edited form as:

Proc SPIE Int Soc Opt Eng. 2016 March ; 9788: . doi:10.1117/12.2217264.

Large-Scale Granger Causality Analysis on Resting-State Functional MRI

Adora M. DSouza1, **Anas Zainul Abidin**2, **Lutz Leistritz***,3, and **Axel Wismüller***,1,2,4,5

¹Department of Electrical Engineering, University of Rochester, NY, USA

²Department of Biomedical Engineering, University of Rochester, NY, USA

³Bernstein Group for Computational Neuroscience Jena, Institute of Medical Statistics, Computer Science, and Documentation, Jena University Hospital, Friedrich Schiller University Jena, Germany

⁴Department of Imaging Sciences, University of Rochester, NY, USA

⁵Faculty of Medicine and Institute of Clinical Radiology, Ludwig Maximilian University, Germany

Abstract

We demonstrate an approach to measure the information flow between each pair of time series in resting-state functional MRI (fMRI) data of the human brain and subsequently recover its underlying network structure. By integrating dimensionality reduction into predictive time series modeling, large-scale Granger Causality (lsGC) analysis method can reveal directed information flow suggestive of causal influence at an individual voxel level, unlike other multivariate approaches. This method quantifies the influence each voxel time series has on every other voxel time series in a multivariate sense and hence contains information about the underlying dynamics of the whole system, which can be used to reveal functionally connected networks within the brain. To identify such networks, we perform non-metric network clustering, such as accomplished by the Louvain method. We demonstrate the effectiveness of our approach to recover the motor and visual cortex from resting state human brain fMRI data and compare it with the network recovered from a visuomotor stimulation experiment, where the similarity is measured by the Dice Coefficient (DC). The best DC obtained was 0.59 implying a strong agreement between the two networks. In addition, we thoroughly study the effect of dimensionality reduction in lsGC analysis on network recovery. We conclude that our approach is capable of detecting causal influence between time series in a multivariate sense, which can be used to segment functionally connected networks in the resting-state fMRI.

Keywords

Resting-state fMRI; functional connectivity; effective connectivity; Granger causality; dimensionality reduction; Principal Component Analysis; non-metric clustering; Louvain method

adora.dsouza@rochester.edu; University of Rochester, NY.

These authors are joint senior authors

1. INTRODUCTION

The development in our understanding of complex systems has directly influenced studies analyzing connectivity between different regions in the brain [1]. Of particular interest is the quantification of causal influence that different regions exert on one another. Methods such as Dynamic Causal Modelling (DCM) [2] and Structural Equation Modelling (SEM) [3] demonstrate how such interactions can be estimated. One common shortcoming of these methods is that they are unable to obtain a causality score between every pair of voxel time series in the brain, i.e., the common situation that the number of time series by far exceeds the number of temporal acquisitions. Instead, such methods are limited to the analysis of a small number of pre-selected regions. This is because the above methods encounter problems of underdetermined systems or exponential growth of models to be tested, if influence scores at a voxel level should be obtained. We demonstrate the use of large-scale Granger Causality (lsGC) [4] to investigate causal influence at a voxel resolution scale in resting-state fMRI, and we illustrate how this information can be used to extract underlying brain network structures using non-metric clustering approaches. This work is embedded in our group's endeavor to expedite 'big data' analysis in biomedical imaging by means of advanced machine learning and pattern recognition methods for computational radiology and radiomics, e.g. [5–29].

Our lsGC analysis method overcomes the problem of an underdetermined system by first incorporating a dimensionality reduction step retaining high variance principal components, followed by modelling the system using Vector Auto-Regression (VAR) [31]. As Principal Component Analysis (PCA) [32] is a linear dimension reduction scheme, we can backproject the prediction errors – obtained from VAR modelling – into the original highdimensional space of voxel time series and thus calculate Granger Causality (GC) indices at a voxel resolution scale. This measure of influence likely contains information about underlying information transfer between functionally connected times series in the brain. This influence flow is conditioned on the information present in all the time series by taking the first few high variance components and performing a multivariate analysis [33], leading to a lsGC index between every pair of voxel time series. Thus, we can represent the human brain as a directed graph with voxels being the nodes and the influence score (lsGC index) being the edge weights. We extract the underlying network structure from the indices by performing graph-theoretical approaches for community detection, such as non-metric clustering by the Louvain method [34]. The resulting network, obtained from the restingstate fMRI time series is compared with the network obtained from the task stimulation sequence using the Dice Coefficient (DC) [35]. We also investigate the effect of the number of retained principal components on the robustness of network recovery.

2. DATA

An MRI data set from a healthy male, aged 52 years, was obtained at the Rochester Center for Brain Imaging (Rochester, NY, USA). The scan was acquired using a 3.0 Tesla, Siemens Magnetom TrioTim scanner, using the fMRI scan parameters echo time $(TE) = 23$ milliseconds, repetition time (TR) = 1650 milliseconds, 96×96 acquisition matrix, flip angle of 84°. Two sequences of images were obtained: 1) Resting-state acquisition, where

DSouza et al. Page 3

the subject was asked to lie down still with his eyes closed, while 250 volumes under resting state were obtained. 2) Visuomotor stimulation, where the subject had to perform a finger tapping task in the presence of an alternating visual stimulus, where 110 volumes were acquired. The task sequence was used to aid in localizing the motor and visual cortex of the subject. The study protocol also included high resolution structural imaging using a T1 weighted Magnetization-Prepared Rapid Gradient Echo (MP-RAGE) sequence (TE = 3.44 ms, TR= 2530 ms, isotropic voxel size of 1mm, flip angle $= 7^{\circ}$). The scanned individual had given written consent as per protocol approved by the IRB.

3. METHODS

3.1 Pre-processing

The effects of initial saturation were eliminated by removing the first 10 volumes of the acquisition. The volumes were then motion-corrected and the brain was extracted. To remove effects of signal drifts, linear detrending was performed by high pass filtering (0.01 Hz). Subsequently, the slices were registered to the standard MNI152 template [36]. In addition, the ventricle mask based on the standard MNI152 template was used to eliminate time series in the corresponding regions. Finally, the time series were normalized to zero mean and unit standard deviation to focus on signal dynamics rather than amplitude [37]. The motor task sequences were further processed using the standard (FEAT) [38] boxcar design in order to test, where the BOLD signal correlated with the stimulation. The zstatistic images obtained were thresholded at z with a corrected cluster significance threshold of $p<0.01$. This results in a map of regions corresponding to the visual and motor cortex for each individual, referred to as 'localization aid' in this study.

3.2 Large-scale Granger causality

We use large-scale Granger Causality [4] to obtain a casual influence score between every pair of voxel time series in the resting state human brain. Large-scale Granger Causality works on the principle of Granger Causality, which obtains the casual influence of time series \mathbf{x}_r on time series \mathbf{x}_s by quantifying the improvement in prediction of \mathbf{x}_s in the presence of \mathbf{x}_r . Large-scale Granger Causality obtains an influence score between every pair of voxels in a multivariate sense. The limitation with using a multivariate approach to Granger Causality, without incorporating dimensionality reduction, on large systems like the brain, is that the prediction model cannot estimate model parameters as the system is undetermined. We show how using lsGC approach, this problem can be overcome.

Consider the ensemble of time series $\mathbf{X} \in \mathbb{R}^{N \times T}$, where N is the number of time series and T the number of temporal samples. Let $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_N)^T$ be the whole multidimensional system, $\mathbf{x}_n \in \mathbb{R}^{1 \times T}$ a single time series with $n \in \{1, 2, ..., N\}$, where $\mathbf{x}_n =$ $(x_n(1), x_n(2), ..., x_n(T))$. IsGC obtains a multivariate GC score by first decomposing **X** into its first p high-variance principal components $\mathbf{Z} \in \mathbb{R}^{p \times T}$ using Principal Component Analysis (PCA), i.e.,

 $Z=$ WX

DSouza et al. Page 4

Following this, we model **Z** using Vector Auto-Regressive (VAR) modelling of order m and compute the estimate \hat{z} by using the m auto-regression $\mathbb{R}^{p \times p}$ parameter matrices obtained from VAR modelling of \mathbb{Z} . To obtain the influence of time series \mathbf{x}_r on all other time series, we remove the information of \mathbf{x}_r from the transformation matrix **W**, obtain $Z_{\setminus {\mathbf{x}_r}}$ and model

 $\hat{Z}_{\setminus \{x_r\}}$ as its VAR estimate. In this manner, lsGC obtains a causality score between every pair of time series in a multivariate sense.

The errors of the two estimates are projected into the original high-dimensional space using the respective inverse PCA transformation. After the errors are obtained, we quantify the prediction quality by comparing the variance of the prediction errors obtained with and without consideration of x_r . If the variance of the prediction error of a given time series, say \mathbf{x}_s , decreases with using \mathbf{x}_r , then we say that \mathbf{x}_r Granger-causes \mathbf{x}_s [4].

$$
f_{\mathbf{x}_r \to \mathbf{x}_s} = \ln \frac{\text{var}\left(\mathbf{e}_{s,\backslash \{\mathbf{x}_r\}}\right)}{\text{var}\left(\mathbf{e}_s\right)}
$$

is the GC index for the influence of \mathbf{x}_r on \mathbf{x}_s , which is stored in the affinity matrix **A** at position $(A)_{s,t}$ $e_{s,\lambda\{x_r\}}$ is the error in predicting \mathbf{x}_s , when \mathbf{x}_r was not considered, and e_s is the error when \mathbf{x}_r was used.

3.3 Non-metric clustering

The affinity matrix **A**, whose coefficients are the large-scale Granger causality indices, can be represented as a network graph, from which the underlying network structure can be extracted using non-metric clustering approaches, such as the Louvain method [34]. Functionally connected regions within the brain can be extracted using this information. This method optimizes the modularity of a network by dividing the network into modules that have strong intra-module and weak inter-module connections.

The underlying network structure information present in the affinity matrix **A**, can be extracted using Louvain method which finds high modularity communities in the network. Modularity, which is a measure of the strength of the intra-module links as compared to the inter-modules links, can be used to decompose a complex network into clusters (or modules).

Modularity is given by:

$$
Q = \frac{1}{2m} \sum_{i,j} \left[\mathbf{A}_{ij} - \frac{k_i k_j}{2m} \right] \delta(C_i, C_j) \tag{4}
$$

Where **Aij** is the coefficient that represents the large-scale Granger causality score from nodes *i* to *j*, $k_i = \sum_j \mathbf{A}_{ij}$ is called the strength of connections to node *i*, *i* belongs to community C_i , $\delta(a,b) = 1$ when $a = b$, and 0 otherwise, and $m = \frac{1}{2} \sum_{i} A_{ij}$.

Optimization of modularity, Q, is carried out in an iterative process as proposed in [34]. This approach merges different nodes of the network into larger modules (communities) for a positive change in modularity. The merging terminates when further addition of nodes to communities decreases the modularity of the network. For fMRI data, the modules resulting after performing non-metric clustering correspond to different functionally connected regions in the resting human brain. The similarity between the motor and visual cortex clustering results and the localization aid obtained from the task sequence was evaluated using the Dice coefficient [35].

All preprocessing steps were carried out using the FMRI Expert Analysis Tool (FEAT) in FSL [38]. The time series analysis procedures were implemented using MATLAB (MathWorks Inc., Natick, MA, 2013). The Louvain method implementation was taken from [39].

4. RESULTS

4.1 Influence of number of principal components used on the Dice coefficient

Figure 1 shows the effect of changing the number of principal components used in obtaining the lsGC coefficients on the network recovery — quantified by the Dice Coefficient (DC). We observe that the DC increases initially with increasing number of principal components, until it reaches a maximum, and decreases subsequently. The initial rise in DC can be accounted to the fact that as the retained number of principal components increases, the amount of variability of the retained information on the data increases and hence the system is better represented with more number of principal components, where the explained variance is the measure of the retained variability of system. However, increasing the number of principal components further decreases the DC. This happens because the quality of the VAR parameter estimation is limited by the length of the time series, and hence, although the system is better represented with more principal components, we do not achieve good network recovery.

4.2 Network Recovery

We compare the best network recovered using the resting-state sequence with the localization aid in Figure 2. We coarsened the resolution of the whole brain data set by averaging the signal intensity information of three adjacent voxels, where a new voxel encompasses $3\times3\times3$ original voxels. The total number of time series in the coarsened data set was 6584. The best network recovery was achieved, when 18 principal components were retained ($DC = 0.59$). As clearly seen in Figure 2, there is a strong visually perceivable agreement of the two networks, i.e., the task-based localization aid and the recovered visual and motor network.

5. NEW AND BREAKTHROUGH WORK

We demonstrate the use of large-scale Granger Causality (lsGC) to first obtain a directed multivariate Granger causality coefficient for every pair of voxel time series in the brain from resting-state fMRI data. Subsequently, we recover functionally connected networks in the brain by a community detection approach using non-metric clustering. Note that such a

whole-brain multivariate connectivity analysis is not feasible using other current approaches. For example, to perform exhaustive multivariate GC analysis for every pair of voxel time series within the brain is not feasible due to the limited number of temporal fMRI acquisitions, resulting in an underdetermined problem for VAR parameter estimation. Alternatively, one would first have to pre-select a much smaller number of brain regions and obtain influence scores within such a markedly simplified network only. In contrast, the work presented here is the first approach to introduce a method to successfully evaluate multivariate effective connectivity at a voxel resolution scale for whole-brain functional MRI

6. CONCLUSION

data.

We present the applicability of our approach, called large-scale Granger Causality (lsGC) analysis, to recover functionally connected networks in whole-brain human resting-state fMRI data. This analysis reveals important information about brain connectivity by quantifying the directed causal influence of every voxel time series on every other voxel time series in a multivariate sense. Using this information, we obtain functionally connected networks by performing a non-metric clustering approach, using the Louvain method. We observe a strong agreement of the visual and motor cortex networks and the localization aid provided by a task-stimulation fMRI experiment. The successful recovery of functionally connected networks demonstrates the applicability of lsGC to explore whole-brain network connectivity in human resting-state fMRI data.

This work is not being and has not been submitted for publication or presentation elsewhere.

Acknowledgments

This research was funded by the National Institutes of Health (NIH) Award R01-DA-034977. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Health. This work was conducted as a Practice Quality Improvement (PQI) project related to American Board of Radiology (ABR) Maintenance of Certificate (MOC) for Prof. Dr. Axel Wismüller. The authors would like to thank Prof. Dr. Dorothee Auer at the Institute of Neuroscience, University of Nottingham, UK, for her assistance with the fMRI data acquisition process. The authors would also like to thank Dr. Lutz Leistritz and Prof. Dr. Herbert Witte of Bernstein Group from Computational Neuroscience Jena, Institute of Medical Statistics, Computer Sciences, and Documentation, Jena University Hospital, Friedrich Schiller University Jena, Germany, Dr. Oliver Lange and Prof. Dr. Maximilian F. Reiser of the Institute of Clinical Radiology, Ludwig Maximilian University, Munich, Germany for their support.

References

- 1. Margulies DS, Böttger J, Long X, Lv Y, Kelly C, Schäfer A, Goldhahn D, Abbushi A, Milham MP, Lohmann G, Villringer A. Resting developments: a review of fMRI post-processing methodologies for spontaneous brain activity. Magnetic Resonance Materials in Physics, Biology and Medicine. 2010; 23(5–6):289–307.
- 2. Friston KJ, Harrison L, Penny W. Dynamic causal modelling. Neuroimage. 2003; 19(4):1273–1302. [PubMed: 12948688]
- 3. Mclntosh AR, Gonzalez-Lima F. Structural equation modeling and its application to network analysis in functional brain imaging. Human Brain Mapping. 1994; 2(1–2):2–22.
- 4. Wismüller, A., Nagarajan, MB., Witte, H., Pester, B., Leistritz, L. SPIE Medical Imaging. International Society for Optics and Photonics; 2014. Pair-wise clustering of large scale Granger causality index matrices for revealing communities; p. 90381R-90381R.

DSouza et al. Page 7

- 5. Hoole P, Wismüller A, Leinsinger G, Kroos C, Geumann A, Inoue M. Analysis of tongue configuration in multi-speaker, multi-volume MRI data. Proc 5th Semin Speech Prod Model Data CREST Work Model Speech Prod Mot Plan Articul Model. 2000:157–160.
- 6. Wismüller, A. Thesis PhD. Technical University of Munich, Department of Electrical and Computer Engineering; 2006. Exploratory Morphogenesis (XOM): a novel computational framework for selforganization.
- 7. Wismüller, A., Dersch, DR., Lipinski, B., Hahn, K., Auer, D. A neural network approach to functional MRI pattern analysis—clustering of time-series by hierarchical vector quantization. Springer; London: 1998. p. 857-862.[ICANN 98]
- 8. Wismüller, A., Vietze, F., Dersch, DR., Behrends, J., Hahn, K., Ritter, H. Neurocomputing. Vol. 48. Elsevier; 2002. The deformable feature map-a novel neurocomputing algorithm for adaptive plasticity in pattern analysis; p. 107-139.
- 9. Behrends, J., Hoole, P., Leinsinger, GL., Tillmann, HG., Hahn, K., Reiser, M., Wismüller, A. Bildverarbeitung für die Medizin 2003. Springer; Berlin, Heidelberg: 2003. A segmentation and analysis method for MRI data of the human vocal tract; p. 186-190.
- 10. Wismüller A, Dersch DR. Neural network computation in biomedical research: chances for conceptual cross-fertilization. Theory in Biosciences. 1997; 116(3):229–240.
- 11. Bunte K, Hammer B, Villmann T, Biehl M, Wismüller A. Exploratory Observation Machine (XOM) with Kullback-Leibler Divergence for Dimensionality Reduction and Visualization. ESANN. 2010; 10:87–92.
- 12. Wismüller, A., Vietze, F., Dersch, DR., Hahn, K., Ritter, H. The deformable feature map—adaptive plasticity for function approximation. Springer; London: 1998. p. 123-128.[ICANN 98]
- 13. Wismüller, A. Advances in Self-Organizing Maps. Springer; Berlin Heidelberg: 2009. The exploration machine–a novel method for data visualization; p. 344-352.
- 14. Wismüller, A. Method, data processing device and computer program product for processing data. US Patent. 7,567,889. 2009.
- 15. Meyer-Bäse, A., Jancke, K., Wismüller, A., Foo, S., Martinetz, T. Eng Appl Artif Intell. Vol. 18. Elsevier; 2005. Medical image compression using topology-preserving neural networks; p. 383-392.
- 16. Huber, MB., Nagarajan, M., Leinsinger, G., Ray, LA., Wismüller, A. SPIE Med Imaging 7624. International Society for Optics and Photonics; 2010. Classification of interstitial lung disease patterns with topological texture features. 762410
- 17. Wismüller A. The exploration machine: a novel method for analyzing high-dimensional data in computer-aided diagnosis. SPIE Med Imaging. 2009:72600G–72600G.
- 18. Bunte, K., Hammer, B., Villmann, T., Biehl, M., Wismüller, A. Neurocomputing. Vol. 74. Elsevier; 2011. Neighbor embedding XOM for dimension reduction and visualization; p. 1340-1350.
- 19. Wismüller, A. Advances in Self-Organizing Maps. Springer; Berlin Heidelberg: 2009. A computational framework for nonlinear dimensionality reduction and clustering; p. 334-343.
- 20. Huber, MB., Nagarajan, MB., Leinsinger, G., Eibel, R., Ray, LA., Wismüller, A. Med Phys. Vol. 38. American Association of Physicists in Medicine; 2011. Performance of topological texture features to classify fibrotic interstitial lung disease patterns; p. 2035-2044.
- 21. Wismüller A, Verleysen M, Aupetit M, Lee JA. Recent Advances in Nonlinear Dimensionality Reduction, Manifold and Topological Learning. ESANN. 2010
- 22. Wismüller, A., Meyer-Baese, A., Lange, O., Reiser, MF., Leinsinger, G. Med Imaging, IEEE Trans. Vol. 25. IEEE; 2006. Cluster analysis of dynamic cerebral contrast-enhanced perfusion MRI timeseries; p. 62-73.
- 23. Otto TD, Meyer-Bäse A, Hurdal M, Sumners D, Auer D, Wismüller A. Model-free functional MRI analysis using cluster-based methods. AeroSense 2003. 2003:17–24.
- 24. Varini C, Nattkemper TW, Degenhard A, Wismüller A. "Breast MRI data analysis by LLE", Neural Networks, 2004. Proceedings 2004 IEEE Int Jt Conf. 2004; 3:2449–2454.
- 25. Huber, MB., Lancianese, SL., Nagarajan, MB., Ikpot, IZ., Lerner, AL., Wismüller, A. Biomed Eng IEEE Trans. Vol. 58. IEEE; 2011. Prediction of biomechanical properties of trabecular bone in MR images with geometric features and support vector regression; p. 1820-1826.

- 26. Meyer-Base A, Pilyugin SS, Wismüller A. "Stability analysis of a self-organizing neural network with feedforward and feedback dynamics", Neural Networks, 2004. Proceedings 2004 IEEE Int Jt Conf. 2004; 2:1505–1509.
- 27. Wismüller, A., Meyer-Base, A., Lange, O., Schlossbauer, T., Kallergi, M., Reiser, MF., Leinsinger, G. J Electron Imaging. Vol. 15. International Society for Optics and Photonics; 2006. Segmentation and classification of dynamic breast magnetic resonance image data. 13020-013020-13
- 28. Nagarajan, MB., Huber, MB., Schlossbauer, T., Leinsinger, G., Krol, A., Wismüller, A. Mach Vis Appl. Vol. 24. Springer; Berlin Heidelberg: 2013. Classification of small lesions in dynamic breast MRI: eliminating the need for precise lesion segmentation through spatio-temporal analysis of contrast enhancement; p. 1371-1381.
- 29. Nagarajan, MB., Huber, MB., Schlossbauer, T., Leinsinger, G., Krol, A., Wismüller, A. J Med Biol Eng. Vol. 33. NIH Public Access; 2013. Classification of Small Lesions in Breast MRI: Evaluating The Role of Dynamically Extracted Texture Features Through Feature Selection.
- 31. Granger CW. Testing for causality: a personal viewpoint. Journal of Economic Dynamics and control. 1980; 2:329–352.
- 32. Wold S, Esbensen K, Geladi P. Principal component analysis. Chemometrics and intelligent laboratory systems. 1987; 2(1):37–52.
- 33. Geweke JF. Measures of conditional linear dependence and feedback between time series. Journal of the American Statistical Association. 1984; 79(388):907–915.
- 34. Blondel VD, Guillame JL, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment. 2008:10008.
- 35. Dice LR. Measures of the Amount of Ecologic Association Between Species. Ecology. 1945; 26(3):297–302.
- 36. Mazziotta J, Toga A, Evans A, Fox P, Lancaster J, Zilles K, Mazoyer B. A four-dimensional probabilistic atlas of the human brain. Journal of the American Medical Informatics Association. 2001; 8(5):401–430. [PubMed: 11522763]
- 37. Wismüller A, Lange O, Dersch DR, Leinsinger GL, Hahn K, Pütz B, Auer D. Cluster analysis of biomedical image time-series. International Journal of Computer Vision. 2002; 46:103–128.
- 38. Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, Matthews PM. Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage. 2004; 23:S208–S219. [PubMed: 15501092]
- 39. Scherrer, A. Community Detection algorithm based on Louvain method. (Version 1.0) [software] Available from [http://perso.uclouvain.be/vincent.blondel/research/Community_BGLL_Matlab.zip](http://perso.uclouvain.be/vincent.blondel/research/Community_BGLL_) (Last accessed July 3, 2014

Figure 1.

Plot showing the variation in Dice coefficient as the number of principal components (explained variance) is increased. It can be noted that the dice coefficient initially increases, reaches a peak and subsequently decreases.

Localization Aid

Clustering Results

Figure 2.

Visual Comparison between clustering results based on large-scale Granger Causality (lsGC) analysis and the localization aid for motor and visual cortices identified by a taskstimulation fMRI experiment. Dice coefficient achieved was 0.59. Recovery of the motor cortex and the visual cortex can be clearly observed.