Investigation of large-scale extended Granger causality (lsXGC) on synthetic functional MRI data

Axel Wismüller\textsuperscript{a,b,c,d}, M. Ali Vosoughi\textsuperscript{a}, Adora DSouza\textsuperscript{a}, and Anas Z. Abidin\textsuperscript{c}

\textsuperscript{a}Department of Electrical and Computer Engineering, University of Rochester, NY, USA
\textsuperscript{b} Department of Imaging Sciences, University of Rochester, NY, USA
\textsuperscript{c} Department of Biomedical Engineering, University of Rochester, NY, USA
\textsuperscript{d} Faculty of Medicine and Institute of Clinical Radiology, Ludwig Maximilian University, Munich, Germany

\textbf{ABSTRACT}

It is a challenging research endeavor to infer causal relationships in multivariate observational time-series. Such data may be represented by graphs, where nodes represent time-series, and edges directed causal influence scores between them. If the number of nodes exceeds the number of temporal observations, conventional methods, such as standard Granger causality, are of limited value, because estimating free parameters of time-series predictors lead to underdetermined problems. A typical example for this situation is functional Magnetic Resonance Imaging (fMRI), where the number of nodal observations is large, usually ranging from $10^2$ to $10^5$ time-series, while the number of temporal observations is low, usually less than $10^3$. Hence, innovative approaches are required to address the challenges arising from such data sets. Recently, we have proposed the large-scale Extended Granger Causality (lsXGC) algorithm, which is based on augmenting a dimensionality-reduced representation of the system’s state-space by supplementing data from the conditional source time-series taken from the original input space. Here, we apply lsXGC on synthetic fMRI data with known ground truth and compare its performance to state-of-the-art methods by leveraging the benefits of information-theoretic approaches. Our results suggest that the proposed lsXGC method significantly outperforms existing methods, both in diagnostic accuracy with Area Under the Receiver Operating Characteristic (AUROC = 0.849 vs. $[0.727,0.762]$ for competing methods, $p < 10^{-8}$), and computation time (3.4 sec vs. $[9.7, 4.8 \times 10^3]$ sec for competing methods) benchmarks, demonstrating the potential of lsXGC for analyzing large-scale networks in neuroimaging studies of the human brain.

\textbf{Keywords:} Resting-state MRI (fMRI), Large-Scale Extended Granger Causality (lsXGC), functional connectivity, graph learning, relation discovery, machine learning

Further author information: (Send correspondence to Ali Vosoughi)
Ali Vosoughi: E-mail: mvosough@ece.rochester.edu
1. INTRODUCTION

For studying the underlying pathophysiology of neurological and psychiatric disease, functional connectivity among various cortical regions has been identified as an important research subject [1]. By providing images with sufficiently high spatial resolution and the hemodynamic response dynamics over a temporal axis, fMRI studies have demonstrated a tremendous potential to serve as a biomarker for neurologic and psychiatric disease [2–5]. Currently, most of the diagnosis in brain-related disorders relies on clinical symptom evaluations, such as based on neuropsychological testing. However, there is a need for more objective biomarkers. To this end, more recently, studies have investigated, if such diagnostic information can be extracted non-invasively from brain activity data. Despite promising results, there is scope for improvement, specifically with regards to using more meaningful connectivity analysis approaches [6].

Here, biomarkers from brain imaging using resting-state functional MRI (rs-fMRI) can be derived using Multi-Voxel Pattern Analysis (MVPA) techniques [7]. MVPA is a machine-learning framework that extracts differences in patterns of brain connectivity for discriminating between connectivity profiles of individuals with neurological disorders and healthy individuals. Results of such studies demonstrate that clinically meaningful information can be learned from fMRI data. However, the simple conventional approach of using cross-correlation between time-series does not measure directed connectivity. Hence, more relevant information in the fMRI data may be captured by more sophisticated connectivity measures which retain information about the direction of interdependence between time-series.

Various methods have been proposed to obtain such directional relationships in multivariate time-series data, including transfer entropy [8] and mutual information [9]. However, in high-dimensional systems, estimating the system’s underlying density function becomes computationally expensive [10]. Under the Gaussian assumption, transfer entropy is equivalent to Granger causality [11]. However, the computation of multivariate Granger causality for short time series in large-scale problems is challenging [2, 12–15]. Addressing this challenge, we recently proposed large-scale Extended Granger Causality (lsXGC), which is a method that combines the advantages of dimensionality reduction with the augmentation of a conditional source time-series adopted from the original space. The lsXGC method uses predictive time-series modeling for estimating directed causal relationships among fMRI time-series [2], which makes it a suitable method to quantify the uncertainty associated with the stochastic process. This work extends our previous investigation on lsXGC and evaluates its ability to accurately recovering directional information from synthetic rs-fMRI data with known ground truth (GT). Specifically, we compare the performance of lsXGC to conventional multivariate Granger causality (GC) [16], mutual information (MI) [9], and transfer entropy (TE) [8]. Based on extensive simulations, we investigate the
performance of lsXGC at correctly unveiling the underlying dynamics of the synthetic fMRI dataset, even at a significantly lower computational expense when compared to competing methods. Our results qualify lsXGC as a promising candidate for serving as a potential biomarker for brain disease in clinical fMRI studies. This work is embedded in our group’s endeavor to expedite artificial intelligence in biomedical imaging by means of advanced pattern recognition and machine learning methods for computational radiology and radiomics, e.g., [17–85].

2. METHODS

2.1 Large-scale Extended Granger Causality (lsXGC)

Large-scale Extended Granger Causality (lsXGC) has been developed based on 1) the principle of original Granger causality, which quantifies the causal influence of time-series $x_i$ on time-series $x_t$ by quantifying the amount of improvement in the prediction of $x_t$ in presence of $x_s$. 2) the idea of dimension reduction, which resolves the problem of the tackling a under-determined system, which is frequently faced in fMRI analysis, since the number of acquired temporal samples usually is not sufficient for estimating the model parameters [83, 84].

Consider the ensemble of time-series $X \in \mathbb{R}^{N \times T}$, where $N$ is the number of time-series (Regions Of Interest – ROIs) and $T$ the number of temporal samples. Let $X = (x_1, x_2, \ldots, x_N)^T$ be the whole multidimensional system and $x_i \in \mathbb{R}^{1 \times T}$ a single time-series with $i = 1, 2, \ldots, N$, where $x_i = (x_i(1), x_i(2), \ldots, x_i(T))$. In order to overcome the under-determined problem, first $X$ will be decomposed into its first $p$ high-variance principal components $Z \in \mathbb{R}^{p \times T}$ using Principal Component Analysis (PCA), i.e.,

$$Z = WX,$$  \hfill (1)

where $W \in \mathbb{R}^{p \times N}$ represents the PCA coefficient matrix. Subsequently, the dimension-reduced time-series ensemble $Z$ is augmented by one original time-series $x_s$ yielding a dimension-reduced augmented time-series ensemble $Y \in \mathbb{R}^{(p+1) \times T}$ for estimating the influence of $x_s$ on all other time-series.

Following this, we locally predict $X$ at each time sample $t$, i.e. $X(t) \in \mathbb{R}^{N \times 1}$ by calculating an estimate $\hat{X}_{x_s}(t)$. To this end, we fit an affine model based on a vector of $m$ vector of $m$ time samples of $Y(\tau) \in \mathbb{R}^{(p+1) \times 1}$ ($\tau = t - 1, t - 2, \ldots, t - m$), which is $y(t) \in \mathbb{R}^{m \times (p+1) \times 1}$, and a parameter matrix $A \in \mathbb{R}^{N \times m \times (p+1)}$ and a constant bias vector $b \in \mathbb{R}^{N \times 1}$,

$$\hat{X}_{x_s}(t) = Ay(t) + b, \quad t = m + 1, m + 2, \ldots, T.$$  \hfill (2)
Now $\hat{X}_{x_s}(t)$, which is the prediction of $X(t)$ without the information of $x_s$, will be estimated. The estimation processes is identical to the previous one, with the only difference being that we have to remove the augmented time-series $x_s$ and its corresponding column in the PCA coefficient matrix $W$.

The computation of a lsXGC index is based on comparing the variance of the prediction errors obtained with and without consideration of $x_s$. The lsXGC index $f_{x_s \rightarrow x_t}$, which indicates the influence of $x_s$ on $x_t$, can be calculated by the following equation:

$$f_{x_s \rightarrow x_t} = \log \frac{\text{var}(e_{x_s})}{\text{var}(e_{x_t})},$$

where $e_{x_s}$ is the error in predicting $x_t$ when $x_s$ was not considered, and $e_s$ is the error, when $x_s$ was used. In this study, we set $p = 1$ and $m = 2$.

3. SIMULATION ON SYNTHETIC FMRI DATASETS

Quantitative evaluation of lsXGC and competing algorithms was performed on semi-realistic data of rs-fMRI brain recordings [86, 87], in which noise and the hemodynamic response were included in the data generation process. These data were adopted from Netsim [86], and the network 'sim3' with $N = 15$ was used in our simulations. The hemodynamic responses were generated to model the time-delayed interactions among different brain regions, and, as a result, a smooth signal was generated that resembles recorded fMRI brain signals with an inherent signal-to-noise ratio (SNR) of 20 $dB$. The advantage of the simulated fMRI is that the ground truth (GT) for network connectivity is known, which is not available for real fMRI datasets. The graph and GT are shown in Fig. 1. Hence, the test results can be used to quantitatively compare different algorithms in the presence of noise hemodynamic response condition, which is not possible for real-world datasets. In line with the literature, we use the Area Under the Receiver Operating Characteristic (AUROC) as a comparison metric. Here, the task is to infer the underlying connectivity between 15 brain regions across 50 network simulation instances.

Results are shown in Fig. 2 and Table. 1. In Fig. 2, the AUROC of the proposed lsXGC algorithm significantly outperforms the competing methods, affirming its ability at accurately detecting the underlying directed network connectivity structure in simulated fMRI signals under conditions of noise and smoothing based on hemodynamic response, resembling the real-world situation. Simulation time and diagnostic accuracy results are listed in Table. 1.

4. CONCLUSIONS

This paper investigates a novel method, large-scale Extended Granger Causality (lsXGC), for discovering relations in high-dimensional dynamical systems involving observational data. The lsXGC method addresses the
Figure 1: Netsim network 'sim3' adopted from [86] and the ground truth of one of the networks is shown in this figure. The network simulations model hemodynamic response and noise, which resembles real-world resting-state fMRI data, with signal-to-noise ratio of 20 $dB$. The network has 15 nodes, and 50 different realizations of the network were used in our evaluations.

Figure 2: Performance of various algorithms, namely mutual information (MI) [9], transfer entropy (TE) [8], multivariate Granger causality (GC) [16], and our proposed method (lsXGC, [2]) on the Netsim dataset [86]. Areas under the Receiver Operating Characteristics (AUROC) are shown in the figure, where each column represents a different algorithm (see titles). Boxplots represent $[Q1, Q3]=[0.25, 0.75]$ quartiles and median, and the cap lines represent $[minimum, maximum]=[Q1-1.5 \times (Q3-Q1), Q3+1.5 \times (Q3-Q1)]$. As can be seen, the proposed algorithm (lsXGC) significantly outperforms other methods from the literature, suggesting robustness of the method against noise and hemodynamic response effects encountered in fMRI data. Wilcoxon test $p$-values of the proposed lsXGC method as compared to MI, TE, and GC are less than $\{<10^{-9}, <10^{-9}, <10^{-8}\}$, correspondingly.
Table 1: Simulation time and performance on Netsim dataset [86] in identical conditions, except TE and MI, which were performed on Nvidia GeForce 1080-Ti GPU, with the rest based on CPU operations. All simulations were performed in Python 3.8.

| Method | Time (seconds) | AUROC       |
|--------|----------------|-------------|
| MI     | 4753           | 0.728 ± 0.010 |
| TE     | 2806           | 0.727 ± 0.006 |
| GC     | 9.7            | 0.762 ± 0.038 |
| lsXGC  | 3.4            | 0.849 + 0.032 |

 curse of dimensionality in large-scale dynamic systems by combining dimensionality reduction with data augmentation, namely by augmenting data from the conditional source time-series in the original input space onto a dimensionally reduced state-space representation of the multidimensional time-series system, thus improving prediction quality in a Granger causality setting. Here, we focus on numerical results for learning meaningful representations from synthetic resting-state fMRI data, although there may be a wide scope of other application domains, ranging from neuroscience to climatology and finance. Our results suggest competitive performance of the proposed method in comparison to other widely used techniques from the literature, namely conventional Granger causality, transfer entropy, and mutual information. Most importantly, our method infers network connectivity at a significantly lower computational expense, while accurately identifying causal representations from multidimensional time-series data. We conclude that lsXGC exhibits important advantages over competing methods from the existing literature for inferring directed causal network connectivity in large-scale systems with limited observational data. Specifically, our results suggest that lsXGC may qualify as a promising candidate for serving as a potential biomarker for brain disease in clinical fMRI studies.

ACKNOWLEDGMENTS

This research was partially funded by the American College of Radiology (ACR) Innovation Award “AI-PROBE: A Novel Prospective Randomized Clinical Trial Approach for Investigating the Clinical Usefulness of Artificial Intelligence in Radiology” (PI: Axel Wismüller) and an Ernest J. Del Monte Institute for Neuroscience Award from the Harry T. Mangurian Jr. Foundation (PI: Axel Wismüller). This work was conducted as a Practice Quality Improvement (PQI) project related to American Board of Radiology (ABR) Maintenance of Certificate (MOC) for A.W.

REFERENCES

[1] Norman, K. A., Polyn, S. M., Detre, G. J., and Haxby, J. V., “Beyond mind-reading: multi-voxel pattern analysis of fMRI data,” Trends in cognitive sciences 10(9), 424–430 (2006).
[2] Wismüller, A. and Vosoughi, M. A., “Classification of schizophrenia from functional MRI using large-scale extended Granger causality,” International Society for Optics and Photonics (2021).
[3] Chockanathan, U., D’Souza, A. M., Abidin, A. Z., Schifitto, G., and Wismüller, A., “Automated diagnosis of HIV-associated neurocognitive disorders using large-scale Granger causality analysis of resting-state functional MRI,” Computers in Biology and Medicine 106, 24–30 (2019).
[4] Abidin, A. Z., D’Souza, A. M., Nagarajan, M. B., Wang, L., Qiu, X., Schifitto, G., and Wismüller, A., “Alteration of brain network topology in HIV-associated neurocognitive disorder: A novel functional connectivity perspective,” NeuroImage: Clinical 17, 768–777 (2018).
[5] D’Souza, A. M., Abidin, A. Z., and Wismüller, A., “Classification of autism spectrum disorder from resting-state fMRI with mutual connectivity analysis,” in [Medical Imaging 2019: Biomedical Applications in Molecular, Structural, and Functional Imaging], 10953, 109531D, International Society for Optics and Photonics (2019).
[6] Dadi, K., Rahim, M., Abraham, A., Chyzhyk, D., Milham, M., Thirion, B., Varoquaux, G., Initiative, A. D. N., et al., “Benchmarking functional connectome-based predictive models for resting-state fMRI,” Neuroimage 192, 115–134 (2019).
Wismüller, A., “The exploration machine—a novel method for structure-preserving dimensionality reduction.,” in [ESANN], Citeseer (2009).

Meyer-Bäse, A., Schlossbauer, T., Lange, O., and Wismüller, A., “Small lesions evaluation based on unsupervised cluster analysis of signal-intensity time courses in dynamic breast MRI,” International journal of biomedical imaging 2009 (2009).

Wismüller, A., Lange, O., Auer, D., and Leinsinger, G., “Model-free functional MRI analysis for detecting low-frequency functional connectivity in the human brain,” in [Medical Imaging 2010: Computer-Aided Diagnosis], 7624, 76241M, International Society for Optics and Photonics (2010).

Meyer-Bäse, A., Saalbach, A., Lange, O., and Wismüller, A., “Unsupervised clustering of fMRI and MRI time series,” Biomedical Signal Processing and Control 2(4), 295–310 (2007).

Wismüller, A., Vietze, F., Dersch, D. R., Hahn, K., and Ritter, H. J., “A neural network approach to adaptive pattern analysis—the deformable feature map.,” in [ESANN], 189–194 (2000).

Meyer-Bäse, A., Lange, O., Wismüller, A., and Hurdal, M. K., “Analysis of dynamic susceptibility contrast MRI time series based on unsupervised clustering methods,” IEEE Transactions on Information Technology in Biomedicine 11(5), 563–573 (2007).

Wismüller, A., Behrends, J., Hoole, P., Leinsinger, G. L., Reiser, M. F., and Westesson, P.-L., “Human vocal tract analysis by in vivo 3d MRI during phonation: a complete system for imaging, quantitative modeling, and speech synthesis,” in [International Conference on Medical Image Computing and Computer-Assisted Intervention], 306–312, Springer (2008).

Wismüller, A., Dersch, D. R., Lipinski, B., Hahn, K., and Auer, D., “Hierarchical clustering of functional MRI time-series by deterministic annealing,” in [International Symposium on Medical Data Analysis], 49–54, Springer (2000).

Wismüller, A., “Method and device for representing multichannel image data,” (Nov. 17 2015). US Patent 9,189,846.

Huber, M. B., Bunte, K., Nagarajan, M. B., Biehl, M., Ray, L. A., and Wismüller, A., “Texture feature ranking with relevance learning to classify interstitial lung disease patterns,” Artificial intelligence in medicine 56(2), 91–97 (2012).

Otto, T. D., Meyer-Bäse, A., Hurdal, M., Sumners, D., Auer, D., and Wismüller, A., “Model-free functional MRI analysis using cluster-based methods,” in [Intelligent Computing: Theory and Applications], 5103, 17–24, International Society for Optics and Photonics (2003).

Varini, C., Nattkemper, T. W., Degenhard, A., and Wismüller, A., “Breast MRI data analysis by lle,” in [2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No. 04CH37541)], 3, 2449–2454, IEEE (2004).

Meyer-Bäse, A., Wismüller, A., Lange, O., and Leinsinger, G., “Computer-aided diagnosis in breast MRI based on unsupervised clustering techniques,” in [Intelligent Computing: Theory and Applications II], 5421, 29–37, International Society for Optics and Photonics (2004).

Meyer-Base, A., Pilyugin, S. S., and Wismüller, A., “Stability analysis of a self-organizing neural network with feedforward and feedback dynamics,” in [2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No. 04CH37541)], 2, 1505–1509, IEEE (2004).

Wismüller, A., Dersch, D., Lipinski, B., Hahn, K., and Auer, D., “Hierarchical clustering of fMRI time-series by deterministic annealing,” Neuroimage 7(4), S593 (1998).

Meyer-Bäse, A., Lange, O., Schlossbauer, T., and Wismüller, A., “Computer-aided diagnosis and visualization based on clustering and independent component analysis for breast MRI,” in [2008 15th IEEE International Conference on Image Processing], 3000–3003, IEEE (2008).

Wismüller, A., De, T., Lochmüller, E., Eckstein, F., and Nagarajan, M. B., “Introducing anisotropic minkowski functionals and quantitative anisotropy measures for local structure analysis in biomedical imaging,” in [Medical Imaging 2013: Biomedical Applications in Molecular, Structural, and Functional Imaging], 8672, 86720I, International Society for Optics and Photonics (2013).

Bhole, C., Pal, C., Rim, D., and Wismüller, A., “3d segmentation of abdominal ct imagery with graphical models, conditional random fields and learning,” Machine vision and applications 25(2), 301–325 (2014).

Nagarajan, M. B., Coan, P., Huber, M. B., Diemoz, P. C., Glaser, C., and Wismüller, A., “Computer-aided diagnosis in phase contrast imaging x-ray computed tomography for quantitative characterization of ex vivo human patellar cartilage,” IEEE Transactions on Biomedical Engineering 60(10), 2896–2903 (2013).

Wismüller, A., Behrends, J., Dersch, D., Leinsinger, G., Vietze, F., and Hahn, K., “Automatic segmentation of cerebral contours in multispectral MRI data sets of the human brain by self-organizing neural networks,” in [Radiology], 221, 461–461 (2001).

Wismueller, A., Vietze, F., Dersch, D., Leinsinger, G., Ritter, H., and Hahn, K., “Adaptive self-organized template matching of the gray-level feature space for automatic segmentation of multispectral MRI data of the human brain,” in [Radiology], 213, 364–364 (1999).

Meyer-Bäse, A., Lange, O., Wismüller, A., and Leinsinger, G., “Computer-aided diagnosis in breast MRI based on ICA and unsupervised clustering techniques,” in [Independent Component Analyses, Wavelets, Unsupervised Smart Sensors, and Neural Networks III], 5818, 38–49, International Society for Optics and Photonics (2005).

Nagarajan, M. B., Coan, P., Huber, M. B., Diemoz, P. C., Glaser, C., and Wismüller, A., “Computer-aided diagnosis for phase-contrast x-ray computed tomography: quantitative characterization of human patellar cartilage with high-dimensional geometric features,” Journal of digital imaging 27(1), 98–107 (2014).

Pester, B., Leistritz, L., Witte, H., and Wismueller, A., “Exploring effective connectivity with a Granger causality approach with embedded dimension reduction,” Biomedical Engineering/Biomedizinische Technik 58(SI-1-Track-G), 000010151520134172 (2013).
Abidin, A. Z., Dar, I., D’Souza, A. M., Lin, E. P., and Wismüller, A., “Investigating a quantitative radiomics approach for brain tumor classification,” in [Medical Imaging 2019: Biomedical Applications in Molecular, Structural, and Functional Imaging], 10953, 109530B, International Society for Optics and Photonics (2019).

Abidin, A. Z., D’Souza, A. M., Schifitto, G., and Wismüller, A., “Detecting cognitive impairment in HIV-infected individuals using mutual connectivity analysis of resting state functional MRI,” Journal of neurovirolology 26(2), 188–200 (2020).

Wismüller, A. and Stockmaster, L., “A prospective randomized clinical trial for measuring radiology study reporting time on artificial intelligence-based detection of intracranial hemorrhage in emergent care head ct.” in [Medical Imaging 2020: Biomedical Applications in Molecular, Structural, and Functional Imaging], 11317, 113170M, International Society for Optics and Photonics (2020).

Wismüller, A., D’Souza, A. M., Abidin, A. Z., and Vosoughi, M. A., “Large-scale nonlinear Granger causality: A data-driven, multivariate approach to recovering directed networks from short time-series data,” arXiv preprint arXiv:2009.04681 (2020).

Vosoughi, M. A. and Wismüller, A., “Large-scale kernelized Granger causality to infer topology of directed graphs with applications to brain networks,” arXiv preprint arXiv:2011.08261 (2020).

Vosoughi, M. A. and Wismüller, A., “Large-scale extended granger causality for classification of marijuana users from functional mri,” in [Medical Imaging 2021: Biomedical Applications in Molecular, Structural, and Functional Imaging], 11600, 116000D, International Society for Optics and Photonics (2021).

Smith, S. M., Miller, K. L., Salimi-Khorshidi, G., Webster, M., Beckmann, C. F., Nichols, T. E., Ramsey, J. D., and Woolrich, M. W., “Network modelling methods for fmri,” Neuroimage 54(2), 875–891 (2011).

Löwe, S., Madras, D., Zemel, R., and Welling, M., “Amortized causal discovery: Learning to infer causal graphs from time-series data,” arXiv preprint arXiv:2006.10833 (2020).