Visibility graphs for fMRI data: multiplex temporal graphs and their modulations across resting state networks

Speranza Sannino¹,⁴, Sebastiano Stramaglia², Lucas Lacasa³, Daniele Marinazzo⁴,*

1 Department of Electric and Electronic Engineering, University of Cagliari, Italy
2 Department of Physics, University of Bari and INFN Section of Bari, Italy
3 School of Mathematical Sciences, Queen Mary University of London, United Kingdom
4 Department of Data Analysis, Faculty of Psychology and Educational Sciences, University of Ghent, Belgium

* daniele.marinazzo@ugent.be

Abstract

Visibility graphs have proved a convenient tool to describe the dynamics of time series, finding applications across several disciplines. Recently an approach has been proposed to extend this framework to multivariate time series, allowing a novel way to describe collective dynamics. Here we test their application to fMRI time series and give some explanation on why we think that it’s an interesting thing to do.

Introduction

Visibility graphs (VG) were recently introduced as a method to map time series into networks [22,31], with the aims of using the tools of Network Science [5,35] to describe the structure of time series and their underlying dynamics, thereby enabling the possibility of performing graph-theoretical time series analysis [12]. Research on VG has since then focused essentially on two separated avenues. First, analytic studies have primarily explored the foundations of this mapping [11,16,20,30] and elaborated on mathematical methods [19] to extract rigorous results on the topology of visibility graphs associated to canonical dynamics such as stochastic or chaotic processes [6,13,23,29,32] and to obtain combinatoric analogues of different dynamical quantities [25]. The second avenue deals with applications of this machinery, primarily by using this method as a feature extraction procedure with which build feature vectors which can properly characterize time series with the purpose of making statistical learning (see [4,15,28,41] for a few examples in the life sciences).

In this latter context, the application to neuroscience is on its infancy, and has essentially limited so far to the analysis of electroencephalogram (EEG) data (see [1,14,34] for a few examples). The study of fmRI recordings under these lens has been scarce, and in this work we would like to motivate and justify why we think this is a promising endeavor, both from a univariate and -perhaps more interestingly- from a multivariate time series perspective [24]. In order to do so, we will choose a bigghish, high quality public dataset of resting state fmRI data [36], and will make use of the family of visibility algorithms to build a multilevel graph of temporal networks, where each node represents a time point, and two nodes are connected if they are visible to each other, according to the algorithm explained below. In the case of multivariate time series -as the ones acquired in neuroimaging- each of these networks is the layer of a multiplex network, what enables both the intralayer (univariate) and the interlayer (multivariate) analysis simultaneously. We will show that a direct analysis of this network provides genuine and nontrivial information on fmRI data, showcasing its relevance and usefulness for neuroscience.

The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.
Materials and Methods

fMRI data  We used the public dataset described in [36]. This data was obtained from the OpenfMRI database. Its accession number is ds000030. We use resting state fMRI data from 121 healthy controls, 50 individuals diagnosed with schizophrenia, 49 individuals diagnosed with bipolar disorder and 40 individuals diagnosed with ADHD. The demographics are reported in the original paper, and they can additionally be found in the GitHub page containing the results of this study (https://github.com/danielemarinazzo/Visibility_LA5C_data). The fMRI data was preprocessed with FSL (FMRIB Software Library v5.0). The volumes were corrected for motion, after which slice timing correction was applied to correct for temporal alignment. All voxels were spatially smoothed with a 6mm FWHM isotropic Gaussian kernel and after intensity normalization, a band pass filter was applied between 0.01 and 0.08 Hz. In addition, linear and quadratic trends were removed. We next regressed out the motion time courses, the average CSF signal and the average white matter signal. Global signal regression was not performed. Data were transformed to the MNI152 template, such that a given voxel had a volume of 3mm x 3 mm x 3mm. Finally we averaged the signal in 278 regions of interest (ROIs) using the template described in [42].

In order to localize the results within the intrinsic connectivity network of the resting brain, we assigned each of these ROIs to one of the 9 resting state networks (7 cortical networks, plus subcortical regions and cerebellum) as described in [45].

Figure 1. Examples of Natural Visibility (NV, top) and Horizontal Visibility (HV, bottom) for the same BOLD time series. In each case, a time series of \( N \) data map into a graph of \( N \) nodes, where two nodes are linked according to either natural or horizontal visibility criteria (i.e. convexity and ordering criteria respectively, see the text). In the right part of the figure the resulting adjacency matrices are depicted.

Construction of the visibility graphs  The procedure to build up a visibility graph is extensively and clearly described in [22,23,25] for univariate and [24] for multivariate time series. Here we will recall the basic steps and provide a visualization of the application of the methodology to BOLD data. Given a time series of \( N \) data, any two time points \( i \) and \( j \) in which the measured quantity takes the values \( y_i \) and \( y_j \) respectively will have visibility and consequently will become two connected nodes in the associated graph if any other data point \( y_k \) placed between them fulfills the condition:

\[
y_k < y_j + (y_i - y_j) \frac{j-i}{j-i}.
\]

Together with this convexity criterion, named Natural Visibility (NV), an ordering criterion, named Horizontal Visibility (HV) has also been defined [23]. According to the latter, two time points \( i \) and \( j \), in which the measured quantity takes the values \( y_i \) and \( y_j \) respectively, will now have horizontal visibility if any other data point \( y_k \) placed between them is smaller, i.e.
\[ y_k < \inf\{y_j, y_j\}, \quad \forall k : i < k < j. \]

In either case, the resulting graphs have \( N \) nodes, are connected by a trivial Hamiltonian path that induces a natural ordering in the degree sequence, and are undirected (see figure 1 for an illustration). In the event the time arrow turns out to be a relevant aspect, directed graphs can be easily constructed, as detailed in [25]. Note that the resulting Horizontal visibility graph (HVG) is simply a core subgraph of the Natural visibility graphs (NVG), the former being analytically tractable [19]. As a matter of fact, HVG can be understood as an order statistic [21] and therefore filters out any dependency on the series marginal distributions (that’s not true for NVG so in applications where marginal distributions are relevant, one should use NVG over HVG).

Both algorithms are fast: naive implementations of NVGs have a runtime complexity \( O(N^2) \), however a divide-and-conquer strategy already reduces it to \( O(N \log N) \) [26]. Naive implementation of HVG is already \( O(N \log N) \) in most of the cases of practical interest. Finally, these methods are well-suited to handle several degrees of non-stationarity in the associated time series [21].

In this work we will be analyzing BOLD data, and for that task we decided to choose NVG over HVG. This is because NVGs are in principle better suited to handle and extract long range correlations than HVG, as the former naturally allow for the development of hubs -which will be typically associated to extreme events in the data and can correlate with data at all scales-. Correlations in time series are actually inherited in graph space in the degree distribution. It is somewhat easier to find fat-tailed degree distributions in NVGs (which account for hubs with extremely large degrees). On the other hand, HVGs (which have shown to work fine with processes evidencing short-range correlations) typically display exponentially decaying degree distributions: a feature which is linked to short-scale visibility, making this method more local.

For illustration, Figure 1 depicts how the links are established according to both visibility criteria, resulting in an adjacency matrix for each time series. The code used to compute the Visibility Graphs is available at https://github.com/danielemarinazzo/Visibility and it is basically a translation to Matlab of the scripts in Fortran90 available at http://www.maths.qmul.ac.uk/~lacasa/Software.html.

**Figure 2.** Example of the construction of a multiplex visibility graph from a multivariate time series with \( M = 3 \) components. In this cartoon, each layer builds the HVG associated to each series. Adapted from [24].

When it comes to the application to multivariate time series formed by \( M \) series, each time series yields a different visibility graph to begin with, so in principle the multivariate series can always be mapped into a multilayer graph with \( M \) layers [24]. Moreover, since for every node \( i \) there is a correspondence across layers (node \( i \) corresponds to time stamp \( i \) and this is the same time stamp for all series), there exist a natural alignment between every node of each layer, so the multilayer graph is effectively a multiplex network [5,24] (see figure 2 for an illustration).

First, this multiplex visibility graph encodes the complex structure of each time series in the topology of each layer. One can therefore extract any desired topological feature (say, the entropy over the degree distribution,
which would provide a different number for each layer), with which one could build a feature vector that provides
a compact representation of the multivariate time series complexity. A similar procedure was followed for
instance in [1] to extract markers of Alzheimer’s disease from a graph theoretical characterization of the Hurst
index of EEG data.

Second, the complex interdependencies and correlations which might emerge in a multivariate series across
variables could in turn be extracted using similarity measures across layers. An simple example of such a measure
is the so-called interlayer mutual information, recently explored in the context of coupled chaotic maps [24]. This
quantity is computed starting from the degree distributions \( P(\alpha) \) and \( P(\beta) \) at two arbitrary layers \( \alpha \) and \( \beta \) as:

\[
\text{MI}_{\alpha,\beta} = \sum_{k[\alpha]} \sum_{k[\beta]} P(k^{[\alpha]}, k^{[\beta]}) \log \frac{P(k^{[\alpha]}, k^{[\beta]})}{P(k^{[\alpha]})P(k^{[\beta]})}.
\]

As the degree distribution captures the structure of each layer, this measure is in turn capturing the information
shared between the two layers, that is to say, the information shared across time series. Since this is a \( M \times M \)
matrix whose \( ij \) entry provides the mutual information between layers (ROIs) \( i \) and \( j \), one can then (for instance)
average across pairs -that is to say, across ROIs- to find a mean value of the mutual information for each intrinsic
connectivity network. This methodology is schematized in figure 3.

Figure 3. Scheme of the procedure: Within a given region which aggregates a certain number
of ROIs, one constructs a visibility graph per ROI and builds accordingly a multiplex visibility
graph. We then compute the pairwise mutual information between degree distributions across
the multiplex layers (ROIs) and finally average to obtain a value for each RSN.
Results

We start by reporting in Figure 4 the average interlayer mutual information within each of the intrinsic connectivity networks, for the four groups of subjects considered. For each group of subjects, each circle corresponds to the average interlayer mutual information of a given subject, and Random Average Shifted Histograms (RASH) are also provided. This representation is not parametric, and it is bounded. The plots report the 5th decile of the Harrell Davis estimator, and the 95% high density intervals using a Bayesian bootstrap. Additional details can be found at http://neurostatscyrilpernet.blogspot.be/2016/08/show-me-data.html. The outliers are detected based on the distance between each pair of data points do not assume symmetry of distributions.

In order to account for departure for normality of these distributions we used a graphical approach and computed the Kolmogorov-Smirnov distance, using the code available at https://github.com/GRousselet/matlab_stats, obtaining values up to 0.7 (a value of 0.39 would correspond to rejecting the null hypothesis at a level $\alpha < 0.001$ for the smallest population). The number of ROIs constituting each intrinsic state network (thus a proxy for the network size, given that Shen’s parcellation has ROIS of similar size) is not correlated with the average value of the mutual information. In particular, it is interesting to observe that the intrinsic connectivity network called Limbic in Yeo’s parcellation is the smallest one, but nonetheless has a low interlayer mutual information compared to the other networks for all the clinical groups.

![Figure 4](https://github.com/CPernet/Robust_Statistical_Toolbox)

**Figure 4.** For each group and each intrinsic connectivity network, we plot the distribution across subjects of the averaged interlayer mutual information. The code used for the plots is here https://github.com/CPernet/Robust_Statistical_Toolbox/

The network which showed a clearest differentiation in terms of the average interlayer mutual information among the four clinical groups is the Limbic one (Figure 5). This evidence was assessed by means of a multivariate response test with age of the subjects and framewise displacement as covariates. The p-value of 0.005 was corrected for multiple comparisons using the Bonferroni-Holm criterion with $\alpha = 0.05$. The Kolmogorov-Smirnov statistics of the pairwise comparison between the distributions of average interlayer mutual information values for these particular networks ranged from 0.15 to 0.3. The null hypothesis of values for controls and schizophrenics drawn from the same distribution would be rejected with an $\alpha < 0.005$. Figure 5 also
reports the shift functions to visualize the difference between two distributions, in this case controls and schizophrenics. This function \([46]\) does not assume (as t-tests do) that two distributions differ only in the location of the bulk of the observations, and allow to determine how, and by how much, two distributions differ. Here the Harrell-Davis quantile estimator is used. Confidence intervals of the decile differences with a bootstrap estimation of the standard error of the deciles are computed, and one controls for multiple comparisons so that the type I error rate remains around 0.05 across the 9 confidence intervals. The code is described at [https://garstats.wordpress.com/2016/07/12/shift-function/](https://garstats.wordpress.com/2016/07/12/shift-function/) in this specific case we can observe separation for all the quantiles but the ninth one.

To complement this analysis, in figure S1 we further report two really suboptimal -but unfortunately very common- ways in which results of this kind are often represented (mean and standard errors), which provide a naked eye differentiation between clinical groups.

**Figure 5.** Left: The average interlayer mutual information for the intrinsic connectivity network denoted as **Limbic**, for the four groups of subjects. Right: Shift function to visually and statistically compare the distributions for controls and schizophrenics, at different quantiles. The code used for the plots can be found at [https://github.com/CPernet/Robust_Statistical_Toolbox/](https://github.com/CPernet/Robust_Statistical_Toolbox/) and [https://github.com/GRousselet/matlab_visualisation](https://github.com/GRousselet/matlab_visualisation)

**Discussion**

**Why the (Multivariate) Visibility Graph?** These are some of the reasons why we think that visibility graphs are a convenient tool

- **Useful:** visibility graphs have been shown to inherit in their topology the essence of the associated dynamics, including nontrivial fingerprints which turn to be both descriptive and informative for statistical learning purposes.

- **Do not require pre-processing:** Can be used directly in both stationary and non-stationary signals, and do not require *ad hoc* phase partitioning or symbolization. They naturally filter out linear trends so they don’t require such detrending.

- **Computationally easy and efficient:** the method is numerically straightforward to implement and the runtime algorithms are quite decent (varying from \(O(n)\) for so-called visibility sequential motifs \([16]\) to \(O(n \log n)\) for the full adjacency matrices using divide-and-conquer strategy)

- **Amenable to analytical insight:** differently from other strategies for graph-theoretical time series analysis, visibility graphs are not black boxes. More particularly for the HVG (but not only \([16,31]\) ), there exist several theorems available and methods to build rigorous results of HVGs properties \([19,20,23,25,29]\).

- **Versatile:** generally applicable to both univariate and multivariate time series across the disciplines.
Novel: It builds a bridge between time series and networks and thus opens the exciting possibility of exploring the usefulness of a large bunch of new tools in the endeavor of describing and classifying complex signals.

Coming back to the specific reason why we think that Natural Visibility Graphs are particularly suited for BOLD data, it has been shown that relevant information on the time course of the BOLD signal and on correlated activity can be extracted by looking at single frames, corresponding to peaks in the signal [27,43], and that these events could be the proxy for an innovation signal at the neural level [17,47]. In this framework, the degree of the nodes corresponding to the BOLD peaks in the adjacency matrix constructed according to the Natural Visibility emphasizes the functional relevance of the neural events and of the corresponding patterns of coactivation across the brain. Notwithstanding, both NVG and HVG have been shown to be useful in different contexts so there is no general rule of thumb on what method should we use: this choice shall be addressed in a case by case basis.

Finally, what is important/informative when it comes to describe the properties of a certain cognitive state? Is it the complex pattern underlying the structure of individual time series (that is, local activity of ROIs) of different regions? Or are the correlations and interdependencies (understood in a broad sense) between these regions the key aspect to look at? When the latter is the case, then a functional network analysis approach [7] seems to be the appropriate thing to do. In the former case where the nature of local activity across regions already captures information [14,49], then one does not need to resort to functional dependencies and local analysis is the correct thing to do. This is obviously an open question which should be addressed, from a biological point of view, in a case by case basis. A recent study suggests that both conceptual frameworks can indeed be connected [40]. In general, most probably both aspects play a relevant role, and some studies have already successfully merged the two [9,44]. Be that as it may, the multiplex visibility framework offers a compact way of extracting at once both the local temporal structure (via the network intralayer properties) and the global interconnection pattern (via multiplex interlayer similarities).

Spatial similarity with other measures Using the excellent resource that is NeuroVault (http://neurovault.org/) we also looked at the maps depicting the results of other measures, and noticed that the areas belonging to the limbic Yeo network are associated with lower levels of Regional Homogeneity (ReHo) [49], higher coefficient of variation of the BOLD signal [48], and lower value of the fractional amplitude of low-frequency fluctuations (fALFF) [50]. This evidence speaks to the fact that interlayer mutual information in multiplex visibility networks is associated to decreased predictability and increased independence between the degrees of freedom of the measured time series.

So, what about the limbic system and psychosis? In a parallel universe, proud of our surviving p-value (and a corrected one!!) we could have chosen to submit this paper to a glam journal perhaps with the title "Limbic network biomarkers of psychosis". In an admirable effort of HARKing [37] we would then have referred to papers reporting differences in the limbic network associated to mental disorders [18,38,39] in the introduction and not in a post-hoc tiny bit of discussion. Nonetheless we find these studies really interesting and motivating, and we refer to studies much better motivated than this one, aimed to classify mental disorders and to map them on neuroimaging data [8,10,33].

To conclude, given the exposition and results reported in this study, we hope to have motivated our colleagues to consider Visibility Graphs as a valuable Network Neuroscience tool for both exploratory and focused studies.

Supporting Information

The code and data necessary to replicate the results reported here are indicated in the text. For convenience we report here the location of the main repository, linking to the others. https://github.com/danielemarinazzo/Visibility_LA5C_data
Figure S1. Two visually attractive but statistically suboptimal ways to report the group results, based on mean values and bars representing standard errors.

Acknowledgments

We thank Matteo Fraschini (University of Cagliary) for setting up the Erasmus mobility for Speranza. We thank Caroline Garcia Forlim for consulting on the mutual information code. We thank Enzo Nicosia (Queen Mary University of London) for stimulating discussions on visibility graphs.
References


19. L. Lacasa. On the degree distribution of horizontal visibility graphs associated with markov processes and

20. L. Lacasa. Horizontal visibility graphs from integer sequences. Journal of Physics A: Mathematical and
Theoretical, 49(35):35LT01, 2016.


22. L. Lacasa, B. Luque, F. Ballesteros, J. Luque, and J. C. Nuño. From time series to complex networks: the
visibility graph. Proceedings of the National Academy of Sciences of the United States of America,

23. L. Lacasa, B. Luque, J. Luque, and J. C. Nuño. The visibility graph: A new method for estimating the

24. L. Lacasa, V. Nicosia, and V. Latora. Network structure of multivariate time series. Scientific reports,
5:15508, oct 2015.


27. X. Liu and J. H. Duyn. Time-varying functional network information extracted from brief instances of
spontaneous brain activity. Proceedings of the National Academy of Sciences of the United States of


29. B. Luque, F. J. Ballesteros, A. M. Núñez, and A. Robledo. Quasiperiodic graphs: Structural design,

30. B. Luque and L. Lacasa. Canonical horizontal visibility graphs are uniquely determined by their degree

31. B. Luque, L. Lacasa, F. Ballesteros, and J. Luque. Horizontal visibility graphs: Exact results for random

32. B. Luque, L. Lacasa, F. J. Ballesteros, and A. Robledo. Feigenbaum graphs: A complex network
perspective of chaos. PLOS ONE, 6(9):1–8, 09 2011.

Functional-Domain Connectivity Analysis: Application to Schizophrenia Reveals Distorted Directional

34. A. Mira-Iglesias, J. Alberto Conejero, and E. Navarro-Pardo. Natural visibility graphs for diagnosing
attention deficit hyperactivity disorder (ADHD). 2016.


