Graph Theoretic General Linear Model (GTG): a MATLAB toolbox

Jeffrey M. Spielberg
VA Boston Healthcare System & Boston University School of Medicine

Graph Theory & Psychopathology

There is increasing agreement that that psychopathology is best understood as *disturbances in brain networks* (Insel, 2010). The recent application of graph theory to clinical neuroscience allows network pathology to be studied in an increasingly sophisticated manner. Graph-theoretic topological properties provide insight into the organization of networks and the role of regions within a network, and can reduce the incredibly large search space of brain networks in a meaningful manner (Rubinov & Sporns, 2010). Thus, graph theory has the potential to greatly expand knowledge of pathology-related network disturbance, enriching our understanding of etiological pathways.

Categories of topological graph properties include:

Segregation – the degree of optimization for specialized processing to occur within densely interconnected groups of nodes. Properties include: *Clustering Coefficient, Local Efficiency, Transitivity*

Integration – the degree to which a network can rapidly combine specialized information across distributed nodes. Properties include: *Characteristic Path Length, Global Efficiency*

Centrality (Influence) – the degree to which a particular node/edge facilitates intercommunication. Properties include: *Node Degree/ Strength, Node/Edge Betweeness, Eigenvector Centrality, Pagerank Centrality, K-Coreness Centrality, Subgraph Centrality, Diversity Coefficient, Participation Coefficient, Within-Module Degree Z-Score*

Resilience – network (in)vulnerability to insult. Properties include: *Assortativity*

Testing Relationships Between Graph Properties & Pathology

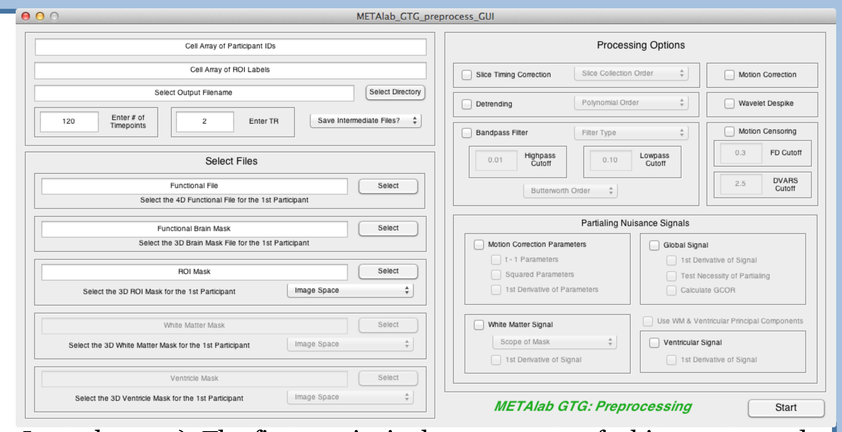
Tools are available to calculate graph theoretic properties (e.g., Brain Connectivity Toolbox, Rubinov & Sporns, 2010). However, there is a dearth of tools available to test hypotheses about the relationship between graph properties and pathology in a (A) flexible and (B) valid manner. By flexible, we mean the ability to use of continuous and/or categorical predictors and between- and/or within-participant measures. By valid, we mean using non-parametric methods to calculate significance. Therefore, we developed the Graph Theoretic General Linear Model (GTG) MATLAB toolbox. The toolbox is freely available on NITRC: www.nitrc.org/projects/metalab_gtg

References

Rubinov, M., Sporns, O. (2010). Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage*, 52, 1059-1069.
Freedman, D., Lane, D. (1983). A nonstochastic interpretation of reported significance levels. *Journal of Business & Economic Statistics*, 1, 292-298.
Power, J.D., Mitra, A., ... Petersen, S.E. (2014). Methods to detect, characterize, and remove motion artifact in resting state fMRI. *NeuroImage*, 84, 320-341.
Patel, A.X., Kundu, P., ... Bullmore, E.T. (2014). A wavelet method for modeling and despiking motion artifacts from resting-state fMRI time series. *NeuroImage*, 95, 287-304.
Jo, H.J., Gotts, S.J., ... Saad, Z.S. (2013). Effective preprocessing procedures virtually eliminate distance-dependent motion artifacts in resting state FMRI. *Journal of Applied Mathematics*, 3, 339-352.
Muschelli, J., Nebel, ... Mostofsky, S.H. (2014). Reduction of motion-related artifacts in resting state fMRI using aCompCor. *NeuroImage*, 96, 22-35.
Chen, G., Chen, G., ... Li, S. (2012). A method to determine the necessity for global signal regression in resting-state fMRI studies. *Magnetic Resonance in Medicine*, 68, 1828-1835.
Saad, Z.S., Reynold, R.C., ... Cox, R.W. (2013). Correcting brain-wide correlation differences in resting-state fMRI. *Brain Connectivity*, 3, 339-352.

Stage 1 – Preprocessing

This stage accepts raw fMRI 4D timeseries data, performs preprocessing, and extracts a processed timeseries for each input ROI. Processing includes standard options: slice-timing correction, motion correction, polynomial detrending. To remove motion variance, Power et al. (2014)'s motion scrubbing and Patel et al. (2014)'s wavelet despiking procedures are available.

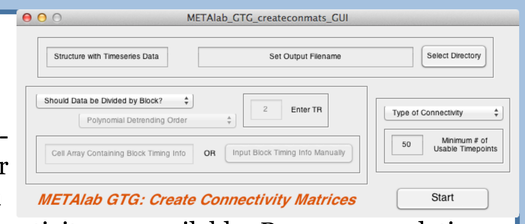


Several options are available for partialing nuisance signals, including mean global, white matter, and ventricular signals, and their (2^{nd} order) 1^{st} derivative. Partialing of motion correction parameters is available, along with the t-1 parameters, squared parameters, and the (2^{nd} order) 1^{st} derivatives. It is also possible to partial signal from only local white matter (within a 45mm sphere, Jo et al., 2013). The first 5 principal components of white matter and ventricular signal (5 components each) can be partialled instead of the mean (Muschelli et al., 2014). Chen et al. (2012)'s Global Negative Index (GNI) can be calculated and (automatically) used determine whether it is necessary to partial global signal. Finally, Saad et al. (2013)'s measure of average brain-wide correlations (GCOR) can be calculated (useful as a GLM covariate).

Stage 2 – Calculate Connectivity Matrices

This stage creates connectivity matrices (one per participant/repeated condition).

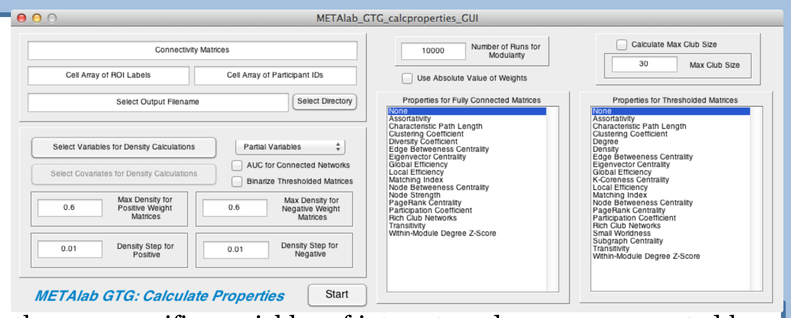
For block-design task fMRI, the toolbox will compute connectivity matrices for each user-specified condition after dividing up the timeseries by condition. In order to compensate for HDR-related delay, timeseries are first deconvolved (using SPM's method), allowing for division at actual onset/offset times. Detrending within each block is available. Four measures of connectivity are available: Pearson correlation, Partial correlation, Mutual information, & Robust (bendcorr) Correlation.



Stage 3 – Calculate Graph Properties

This stage calculates graph theoretic properties for each participant/ repeated level using the Brain Connectivity Toolbox. The measure of connectivity does not matter (e.g., correlation, diffusion tract strength).

Properties are computed for both positive and negative weights from fully-connected and/or thresholded networks. For thresholded networks, properties are computed across a set of density thresholds. The minimum density is chosen such that the presence of disconnected networks is not highly correlated with variables of interest. Specifically, the user specifies variables of interest, and groups are created by stratifying these variables. Mean networks are created for each (stratification) group, and the minimum density at which that network remains connected is identified. This is done for each group (across each variable, across all selected variables) and for the overall mean network, and the maximum of these minima is chosen. An area under the curve (AUC) across density levels is computed for each property.



Stage 4 – Run GLM

This stage calculates GLMs on graph properties. Continuous & categorical between-participant predictors and a categorical within-participant predictor are accepted. When using a within-participant predictor, it is possible to test both main effects and between X within interactions. It is possible to test the significance of individual predictors, contrasts between predictors, or F-tests across a set of predictors.

Along with Ordinary Least Squares, it is possible to use Robust or Trimmed Least Squares GLMs.

Significance is determined via non-parametric permutation tests using the method of Freedman & Lane (1983) to deal with covariates.