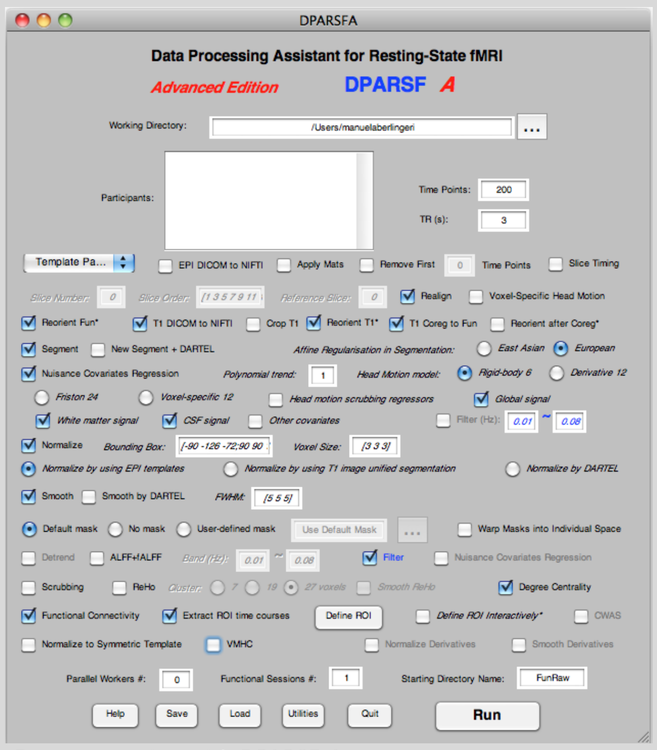
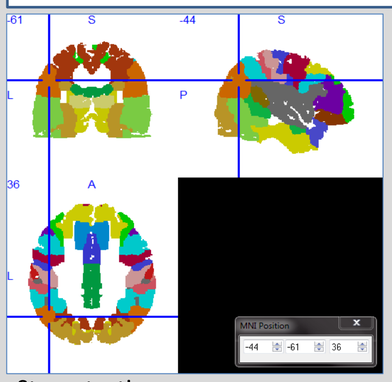
MULTIPLE TESTING IN BRAIN NETWORKS CONSTRUCTION

Agnese M. Di Brisco, Daniela Rigozzi, Manuela Berlingeri, Piero Quatto
Università degli studi Milano-Bicocca



Data processing with DPARSF-A



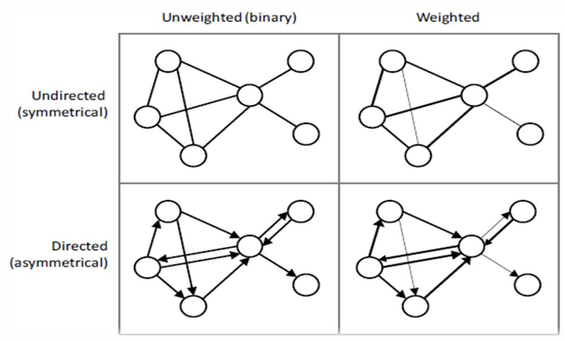
Stereotactic space

Functional Neuroimaging Methods

The functional Magnetic Resonance Imaging (***fMRI***) is a measure of regional neural activity based on the detecion of changes in blood flow, represented by the BOLD (blood oxygenation level dependent) contrast. A set of images covering the whole brain (a brain volume) of a single participant is typically acquired every 2-3 s and, to increase sensitivity, hundreds of brain volumes are typically recorded during either the execution of a complete fMRI scan, or a resting state period (usually from 120-240 volumes for each subject).
Finally, we used the toolbox DPARSF-A to extract 116 non-overlapping anatomical Regions of Interest (ROIs) defined by the automated anatomical labeling (AAL) atlas previously validated by Tzourio-Mazoyer. For the purposes of this study 90 ROIs are taken into account.

Brain Networks

The brain is a complex system that can be represented by a graph: a node corresponds to a brain region, while each link represents the connectivity between two brain regions.



Starting from the fMRI time series, a correlation matrix was computed.

$$\begin{bmatrix} 1 & \cdots & \rho_{1,90} \\ \vdots & \rho_{ij} & \vdots \\ \rho_{90,1} & \cdots & 1 \end{bmatrix}$$

This matrix may be "thresholded" to yield a binary undirected graph; the related adjacency matrix has binary elements that indicate either the presence (value = 1) or absence (value = 0) of a link between pairs of vertices (ROIs).

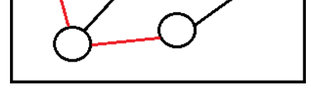
The problem of choosing an appropriate threshold may be dealt with a multiple testing approach

$\binom{90}{2} = 4005$ simultaneous tests

Multiple Testing

m hypotheses to be tested simultaneously: H_0: $H_{01} \Lambda \ldots \Lambda H_{0m}$. *FWER* is the probability of one or more false rejections: *FWER* = P(*V*≥1). If the number of tests is large the procedures that control FWER turn out to be too conservative.
H_j are Bernoulli random variables with P(H_{0j}) = π_0 and P(H_{1j}) = π_1.

Hypotheses	Not rejected	Rejected	Total
H_0	U	V	m_0
H_1	T	S	m_1
	W	R	m



Links between brain regions: black lines represent links correctly accepted, red lines represent links erroneously accepted while green lines represent links erroneously rejected

$$pFDR = E\left[\frac{V}{R} \mid R > 0\right]$$ pFDR is defined as the expected proportion of null hypotheses erroneously rejected., conditioned on the event that positive findings have occurred

$$pFNR = E\left[\frac{T}{m-R} \mid (m-R) > 0\right]$$ pFNR is defined as the expected proportion of null hypotheses erroneously accepted, conditioned on the event that negative findings have occurred

Quantity of interest	*Brief description*	*Estimate*
π_0	a priori probability of a null hypothesis to be true	Let $W(\lambda) = \#\{ p_j > \lambda \}$ where $p_1,\ldots, p_m$ are the observed p-values then, once chosen a good value for the tuning parameter λ, a conservative estimate of π_0 is $\hat{\pi}_0(\lambda) = \frac{W(\lambda)}{(1-\lambda)\,m}$
P($P \le \gamma$)	Probability of p-value r.v. to be less of equal to γ	Let $R(\gamma) = \#\{ p_j \le \lambda \}$ be defined as the number of p-values lower or equal to γ, then a natural estimate of P($P \le \gamma$) is: $\hat{P}(P \le \gamma) = \frac{\max\{R(\gamma),\ 1\}}{m}$
pFDR	Positive False Discovery Rate	A suitable estimate of *pFDR*(γ) could be, according to Storey (2002): $\widehat{pFDR}_\lambda(\gamma) = \frac{\hat{\pi}_0(\lambda)\ \gamma}{\hat{P}(P \le \gamma)\{1-(1-\gamma)^m\}}$
λ	the tuning parameter	An automatic way to estimate the optimal λ: $\hat{\lambda}_1 = \text{argmin}_{\lambda\in[0,1]}\{\text{MSE}(\widehat{pFDR}_\lambda(\gamma))\}$ where $\text{MSE}(\widehat{pFDR}_\lambda(\gamma)) = E[(\widehat{pFDR}_\lambda(\gamma) - pFDR(\gamma))^2]$ In order to estimate the MSE ($\widehat{pFDR}_\lambda(\gamma)$) over a range *R* of λ, for example *R* = {0, 0.05, 0.10, …, 0.95}, the bootstrap version $\widehat{pFDR}_\lambda^{*b}(\gamma)$, for *b* = 1,…, *B*, of $\widehat{pFDR}_\lambda(\gamma)$ is generated for any fixed λ. The bootstrap version of MSE ($\widehat{pFDR}_\lambda(\gamma)$) is: $\widehat{\text{MSE}}(\widehat{pFDR}_\lambda(\gamma)) = \frac{1}{B}\sum_{b=1}^{B}[\widehat{pFDR}_\lambda^{*b}(\gamma) - \min_{\lambda'\in R}\widehat{pFDR}_{\lambda'}(\gamma)]^2$
pFNR	positive False Non-discovery Rate	A suitable estimate of *pFNR*(γ) proposed by Sala et al. (2013) is: $\widehat{pFNR}_\lambda(\gamma) = \frac{(1-\hat{\pi}_0(\lambda))\,m - (1-\widehat{pFDR}_\lambda(\gamma))\,R(\gamma)}{m-R(\gamma)}$ A good estimate of *pFNR* requires the definition of a fixed value for the tuning parameter λ, that can be calculated as follows: $\hat{\lambda}_2 = \text{argmin}_{\lambda\in[0,1]}\{\text{MSE}(\widehat{pFNR}_\lambda(\gamma))\}$
CI	Bootstrap confidence intervals	The bootstrap CI for pFDR is CI $pFDR(\gamma) = [0; \hat{q}_{1-\alpha}]$ where $\hat{q}_{1-\alpha}$ is the 1-α quantile of the ordered distribution of the bootstrap estimates $\widehat{pFDR}_{\lambda_1}^{*1}(\gamma), \ldots, \widehat{pFDR}_{\lambda_1}^{*B}(\gamma)$ Likewise the bootstrap CI for pFNR is CI $pFNR(\gamma) = [0; \hat{q}_{1-\alpha}]$ where $\hat{q}_{1-\alpha}$ is the 1-α quantile of the ordered distribution of the bootstrap estimates $\widehat{pFNR}_{\lambda_2}^{*1}(\gamma), \ldots, \widehat{pFNR}_{\lambda_2}^{*B}$

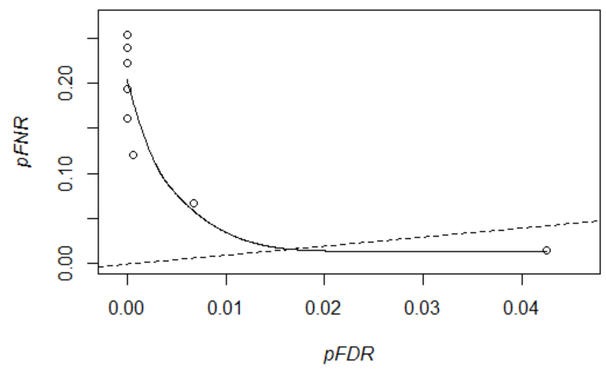
Main results

τ	γ	$_1$	$_0$	(γ)	MSE	CI
0.10	$7.9435\ 10^{-2}$	0.05	0.7131	$1.6438\ 10^{-1}$	$1.2311\ 10^{-5}$	$[0; 1.7031\ 10^{-1}]$
0.15	$1.7002\ 10^{-2}$	0.05	0.7131	$4.2442\ 10^{-2}$	$1.0708\ 10^{-6}$	$[0; 4.4260\ 10^{-2}]$
0.20	$2.2596\ 10^{-2}$	0.05	0.7131	$6.7712\ 10^{-3}$	$3.6478\ 10^{-8}$	$[0; 7.0916\ 10^{-3}]$
0.25	$1.7833\ 10^{-4}$	0.05	0.7131	$6.7010\ 10^{-4}$	$4.8274\ 10^{-10}$	$[0; 7.0640\ 10^{-4}]$
0.30	$7.9485\ 10^{-6}$	0.05	0.7131	$3.7707\ 10^{-5}$	$2.1199\ 10^{-12}$	$[0; 4.0319\ 10^{-5}]$
0.35	$1.8839\ 10^{-7}$	0.05	0.7131	$1.1545\ 10^{-6}$	$2.7303\ 10^{-15}$	$[0; 1.2483\ 10^{-6}]$
0.40	$2.2023\ 10^{-9}$	0.05	0.7131	$1.8830\ 10^{-8}$	$9.9434\ 10^{-19}$	$[0; 2.0621\ 10^{-8}]$
0.45	$1.1534\ 10^{-11}$	0.05	0.7131	$1.3282\ 10^{-10}$	$7.0145\ 10^{-23}$	$[0; 1.4837\ 10^{-10}]$
0.50	$2.3869\ 10^{-14}$	0.05	0.7131	$3.7660\ 10^{-13}$	$7.6153\ 10^{-28}$	$[0; 4.2603\ 10^{-13}]$

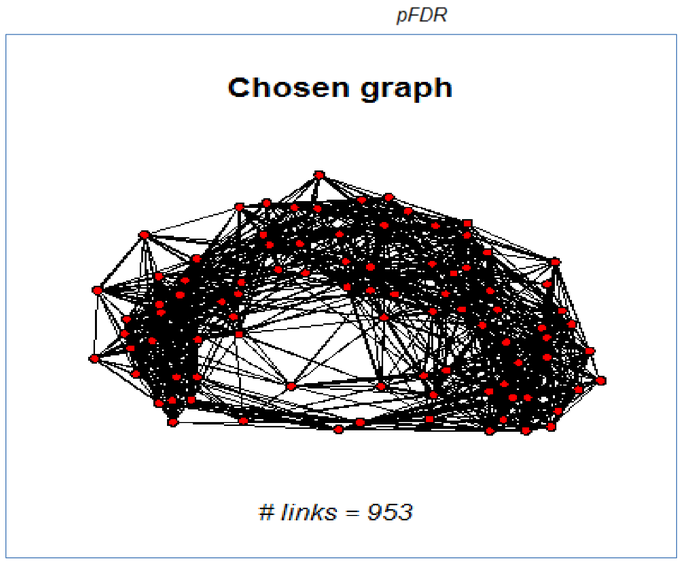
Estimates of *pFDR* and its 95% confidence interval (CI) over a range *T* of τ.

τ	γ	$_2$	$_0$	(γ)	MSE	CI
0.10	$7.9435\ 10^{-2}$	.	.	.	.	.
0.15	$1.7002\ 10^{-2}$	0.10	0.7161	0.0147	$6.8736\ 10^{-5}$	[0; 0.0309]
0.20	$2.2596\ 10^{-2}$	0.05	0.7131	0.0664	$6.7306\ 10^{-5}$	[0; 0.0794]
0.25	$1.7833\ 10^{-4}$	0.05	0.7131	0.1203	$4.5320\ 10^{-5}$	[0; 0.1306]
0.30	$7.9485\ 10^{-6}$	0.05	0.7131	0.1610	$3.2279\ 10^{-5}$	[0; 0.1703]
0.35	$1.8839\ 10^{-7}$	0.05	0.7131	0.1936	$2.2771\ 10^{-5}$	[0; 0.2010]
0.40	$2.2023\ 10^{-9}$	0.05	0.7131	0.2223	$1.3857\ 10^{-5}$	[0; 0.2283]
0.45	$1.1534\ 10^{-11}$	0.05	0.7131	0.2401	$9.6235\ 10^{-6}$	[0; 0.2451]
0.50	$2.3869\ 10^{-14}$	0.05	0.7131	0.2534	$6.4910\ 10^{-6}$	[0; 0.2573]

Estimates of *pFNR* and its 95% confidence interval (CI) over a range *T* of τ.



Trade-off for the subject in study



Graph of the subject in study, once fixed τ = 0.2

Future perspectives

- Construction of group-based brain networks.
- The possible links in the network structure may be weighted in order to obtain a more representative and informative network.
- This approach may be extended also to electrophysiological data collected by means of high-density EEG

Bibliography

J.R. Chumbley, K.J. Worsley, G. Flandin and K.J. Friston (2010). "Topological FDR for NeuroImaging". NeuroImage, 49(4):3057-3064.
•**Genovese C., Lazar N., Nichols T. (2002).** Thresholding of statistical maps in functional neuroimaging using the false discovery rate. Neuroimage, 15(4):870-878.
•**Eric D. Kolaczyk (2009).** "Statistical analysis of network data". Springer Verlag: New York.
•**Sala S., Quatto P., Valsasina P., Agosta F. and Filippi M. (2013).** pFDR and pFNR estimation for brain networks construction. Stat Med, 15;33(1):158-69.
•**Salvador R., Suckling J., Schwarzbauer C., Bullmore E. (2005).** Undirected graphs of frequency-dependent functional connectivity in whole brain networks. Philos Trans R Soc Lond B Biol Sci, 360:937-946.
•**Storey J. (2002).** A direct approach to false discovery rates. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 64(3):479-49.
•**Yan C. G., Zang Y. F. (2010).** DPARSF: a MATLAB toolbox for "pipeline" data analysis of resting-state fMRI. Frontiers in Systems Neuroscience, Vol.4, Art.13.