



Chapitre d'actes

2014

Published version

Open Access

This is the published version of the publication, made available in accordance with the publisher's policy.

Learning with multi-site fMRI graph data

Castrillon, J. Gabriel; Ahmadi, Ahmad; Navab, Nassir; Richiardi, Jonas

How to cite

CASTRILLON, J. Gabriel et al. Learning with multi-site fMRI graph data. In: 2014 48th Asilomar Conference on Signals, Systems and Computers. Pacific Grove (CA, USA). Pacific Grove (CA, USA) : IEEE, 2014. doi: 10.1109/ACSSC.2014.7094518

This publication URL: <https://archive-ouverte.unige.ch/unige:77751>

Publication DOI: [10.1109/ACSSC.2014.7094518](https://doi.org/10.1109/ACSSC.2014.7094518)

Learning with multi-site fMRI graph data

J. Gabriel Castrillon*, Ahmad Ahmadi*[†], Nassir Navab*, and Jonas Richiardi^{‡§}

*Computer Assisted Medical Procedures

Technische Universität München

Germany

[†]Klinikum Grosshadern, Department of Neurology

Ludwig-Maximilians Universität München

Germany

[‡]Functional Imaging in Neuropsychiatric Disorders Laboratory

Stanford University

USA

[§]Laboratory for Neurology and Imaging of Cognition

University of Geneva

Switzerland

Abstract—Neuroimaging data collection is very costly, and acquisition is commonly distributed across multiple sites. However, factors such as different noise characteristics or inhomogeneities make it difficult to successfully combine multi-site functional imaging data. Here, we show that the distribution of signal quality measures across scanners can be significantly different, and that this will have an impact on correlation estimators necessary for computing functional connectivity graphs as well as topological features extracted from the graphs. We propose to find a stable subspace by using a discriminative projection that does not only minimise site differences, but also preserves discriminative class information. We compare our method with the “regressing-out” approach in a cross-validation setting and show that regressing out can yield very poor results.

Keywords—Brain graphs, multi-centric studies, brain connectivity, resting-state.

I. INTRODUCTION

Brain graphs [1], representing interactions between different parts of the brain based on neuroimaging data, are a useful representation of resting-state functional MRI (rs-fMRI) data. They can be used to compare groups of subjects or for prognosis in clinical applications [2], or to perform inference about complex brain states [3].

Since data acquisition with human subjects is costly in time, there is an increased tendency to share data and distribute dataset collections across multiple sites, with the promise of improved statistical power. Indeed, most large-scale fMRI datasets (including ADHD-200, ABIDE, ADNI GO, and fBIRN) are multisite, ranging from 7 to 50 sites.

However, MR physics is such that even two scanners of the same brand or model, using the same operating system and acquisition sequence will yield different images. Signal fluctuations, spiking patterns, and uneven sensitivity profiles will contribute to significantly altering signal dynamics and large-scale spatio-temporal covariance. The computation of brain graph rests on several steps, including spatial and temporal data preprocessing and edge weight estimation, all of which can be affected by the quality of MR acquisition.

While good cross-site reproducibility has been reported for structural imaging classification [4], evidence supporting the applicability of predictive modeling methods to multisite connectivity datasets is non-existent, although connectivity estimates themselves may be resilient [5].

In the present work, we wanted to verify if site differences could indeed cause significant shifts in signal quality measures likely to affect brain graphs, what the effect of site differences is on graph representations and cross-site classification performance, and propose a projection method to diminish site effects.

II. IMAGING DATA AND PROCESSING

A. Datasets

In this paper we will present results on two open multi-site datasets: fBIRN and ABIDE.

The fBIRN Phase I Traveling Subjects dataset [6] includes five healthy males (mean age: 25.2, range = 20.2 to 29). All subjects traveled to 9 sites (10 scanners), where they were scanned twice over a period of 2 days for a total of 20 scans per participant. However, all the datasets without a T1 acquisition were excluded from the analysis, and our results are based on 4 scanners and 4 subjects, shown in table I. Two rs-fMRI acquisitions (TR/TE = 3 s/30 - 40ms, FOV = 22 cm, thickness = 5 mm, number of volumes = 85, slice order = interleaved) and one anatomical T1 scan (TR/TE = 9.8 ms/min, FOV = 22 cm, thickness = 4 mm) were acquired in most of the sites.

Site (ID code)	Field strength	Brand	TR/TE (ms)	Volumes
Duke/UNC (3.15)	1.5	GE	3000/30-40	85
Duke/UNC (3.4)	4.0	GE	3000/30-40	85
UCIR (9.15)	1.5	Philips	3000/30-40	85
Stanford (11.3)	3.0	GE	3000/30-40	85
KKI	3	Philips	2500/30	156
Leuven	3	Philips	1600/33	250
Trinity	3	Philips	2000/28	150
UCLA	3	Siemens	3000/28	120
USM	3	Siemens	2000/28	240

TABLE I. SELECTED fBIRN (TOP) AND ABIDE (BOTTOM) SITES CHARACTERISTICS

The Autism Brain Imaging Data Exchange (ABIDE) provides rs-fMRI datasets from 539 individuals with autism spectrum disorder (ASD) and 573 typical development (TD) subjects acquired in 16 sites, publicly available for the scientific community. Based on the ABIDE image quality check reported by Abraham [7], 5 of 16 sites were selected, consisting of 20 subjects per site (10 ASD and 10 TD paired by age).

All ABIDE sites had their own rs-fMRI protocol, therefore acquisition parameters were different across sites. Principal site

characteristics are listed in table I. KKI, UCLA, and USM had eyes-opened acquisition, while Leuven and Trinity had eyes-closed acquisition.

B. Data processing and graph computation

Standard SPM8 (<http://www.fil.ion.ucl.ac.uk/spm/>) fMRI preprocessing (slice timing correction, realignment, coregistration, normalisation) was performed on the data using all default settings of SPM 8. For the extraction of resting state functional connectivity matrix, we followed an atlas-based pipeline [8].

For spatial processing, the structural image was coregistered to the mean image of the functional data. Then, an individual brain atlas containing 116 cortical and sub-cortical regions of interest (ROIs) was then computed with a modified version of the IBASPM toolbox [9] and the AAL atlas [10]. Finally, the spatially-averaged time courses were calculated from the voxels corresponding to these regions in the functional space. We restricted the analysis to the 90 non-cerebellum regions, and further excluded the bilateral globus pallidus due to signal dropout, resulting in 88 regions per brain, each associated with a regional time course.

For temporal processing, the 6 realignment parameters were regressed out of each regional time courses. The regional time courses were then filtered into four orthogonal frequency subbands using a wavelet transform (if the TR was the same across sites), or bandpass-filtered between 0.01 and 0.1Hz (if the TR was different across sites). Then, the pairwise Pearson correlation coefficients between all pairs of filtered regional time courses were computed to obtain the edge weights of the edges linking each of the 88 regions represented by a graph vertex, leading to an 88 x 88 adjacency matrix.

III. FUNCTIONAL CONNECTIVITY-RELATED SIGNAL QUALITY MEASURES ACROSS SITES

If image quality differs between sites, brain graphs can be impacted. Here we focus on only two measures of fMRI signal quality.

A. Structural quality measure

The number of voxels per atlas region M (in native space) depends largely on the segmentation step in preprocessing, because it is computed from the intersection of a gray matter mask (itself obtained from a thresholded tissue probability map) and an atlas mask. The segmentation, in turn, depends on several image factors, in particular local inhomogeneities and how bias field correction is performed. Thus, M can be expected to vary across sites.

In turn, we have shown previously [11] that the number of voxels in two regions will bias the inter-regional correlation estimator depending on the correlation within each of the two regions. Assuming a true correlation value $cor_{X,Y}$ between all voxels of regions X and Y , then the estimator of the correlation between the regions' average time courses $\bar{X}$ and $\bar{Y}$, $\widehat{cor}_{\bar{X},\bar{Y}}$, will behave as $\widehat{cor}_{\bar{X},\bar{Y}} \rightarrow cor_{X,Y}$ if the within-region correlation between voxels is 1, and as $\widehat{cor}_{\bar{X},\bar{Y}} \rightarrow \sqrt{PM} cor_{X,Y}$ if the within-region correlation between voxels is low (inhomogeneous region).

Therefore, large regions will tend to have higher correlation to the rest of the brain, as illustrated in Fig. 1.

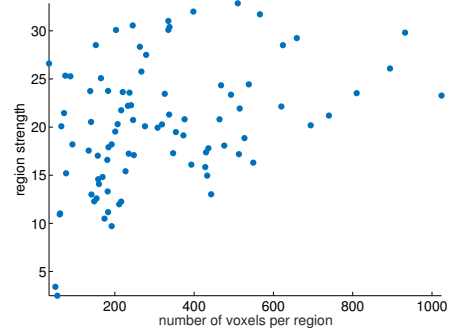


Fig. 1. Number of voxels in a region underlying a graph vertex versus that vertex's strength (sum of edge weights), computed on one random control of the ABIDE dataset.

B. Functional quality measure

Functional connectivity estimates will also be impacted by site-specific temporal noise. Here, we use Signal to Fluctuation Noise Ratio (SFNR) [12] a measure useful to account for brain-wide BOLD signal activation differences in multi-site studies [13].

Here, we defined a spatially localised version: for each atlas region, we computed an SFNR value by dividing signal image (average of the time series) by the temporal fluctuation noise image (standard deviation of the time series detrending with a second-order polynomial).

C. Empirical results

For fBIRN and ABIDE we wrote a mixed-effect model $qm_i \sim 1 + scannerID + (1|subjectID)$, where qm_i is a quality measure (either number of voxels or SFNR) for region i , with $subjectID$ as random effect to account for the correlation of regional measures within each subject. We then tested the main effect of $scannerID$ using a mixed-effects ANOVA.

The total number of voxels for the whole brain did not differ significantly between sites for the fBIRN dataset ($p = 0.56$), but did for the ABIDE dataset ($p < 10^{-4}$). Testing regionwise showed that 30/88 regions differed significantly in their number of voxels in fBIRN ($p_{FDR} < 0.01$, Fig. 2a), and all regions had significantly different number of voxels in ABIDE ($p_{FDR} < 10^{-4}$, Fig. 2b).

Whole-brain SFNR was significantly different between sites for the fBIRN dataset ($p = 0.01$) (consistent with previous results on this dataset [13]) and the ABIDE dataset ($p < 10^{-4}$). Regionwise, 85/88 regions differed significantly in fBIRN ($p_{FDR} < 0.03$, Fig. 2c), and all regions had significant differences in ABIDE ($p_{FDR} < 10^{-4}$, Fig. 2d).

IV. GRAPH REPRESENTATIONS ACROSS SITES

Graphs can be represented as vectors in different ways in order to use statistical learning methods operating in vector spaces [1]. Here we focus on two techniques: direct embedding, where edge weights are taken as features directly [14], and topological features embedding [15]. Both representations rely on estimates of edge weights (or absence/presence of edges), which is likely to be influenced by properties of the acquisition site.

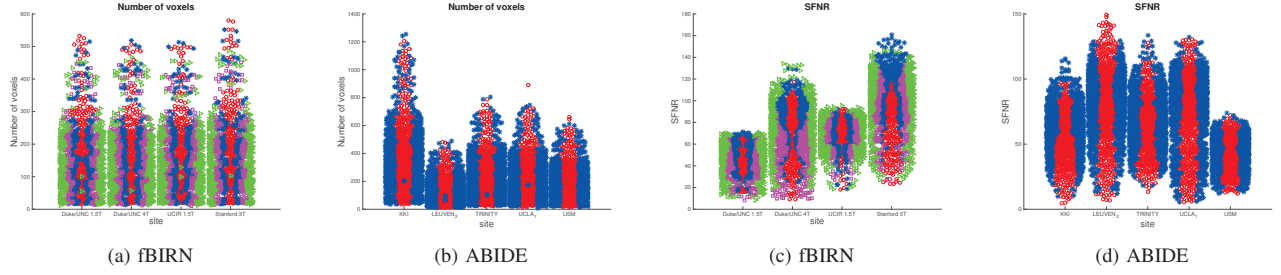


Fig. 2. Site-specific violin plots of the number of voxels per region (two left panels) and regional signal-to-fluctuation noise ratio (two right panels) in the fBIRN and ABIDE datasets. Each dot is the value for one region in one subject. fBIRN data is color-coded by subject (4 subjects shown here), and ABIDE data is colored by diagnosis class, where blue is typically developing (TD) and red is autism spectrum disorder (ASD)

A. Edge weight distributions

Given our results on signal quality measures, we can expect the distribution of edge weights (obtained by Pearson linear correlation coefficient) to shift significantly depending on the acquisition site. Indeed, Fig. 3 shows that this is the case, even when the acquisition protocol and field strengths are the same (as in the case of fBIRN). More worryingly, in the case of clinical applications (ABIDE dataset), the patients and control distributions can shift and overlap to the point where researchers at different sites can reach opposing conclusions: patients can be thought to be under-connected for a UCLA researcher, while they could be described as over-connected for a Trinity researcher.

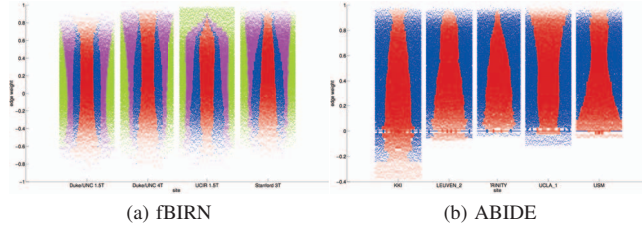


Fig. 3. Site-specific violin plots of the distribution of edge weights in the fBIRN and ABIDE datasets. Color code as in figure 2

B. Topological graph features

Both weighted and unweighted topological graph features rely on edge strength estimation. Unweighted features are computed on a thresholded and binarised version of the graph, meaning that the value of the numerical threshold can have a large influence on the resulting feature.

Here, we first used a fixed edge density threshold for all subjects, keeping 90% of the edges with the largest weight. We then computed six vertex topological features: degree, clustering coefficient [16], unweighted local efficiency [17], unweighted betweenness centrality [18], closeness centrality [19], and pageRank [20].

We then defined a mixed-effect model $vp_i \sim 1 + \text{scannerID} + (1|\text{subjectID})$, where vp_i is a vertex property for vertex i , and subjectID is a random effect to account for the correlation of regional vertex properties within each subject. We then tested the main effect of scannerID using a mixed-effects ANOVA.

Robustness was consistent for the ABIDE and fBIRN datasets:

degree, path length, betweenness centrality and pageRank showed no site effect.

Local efficiency, closeness centrality, and clustering coefficient showed significant effect of site on ABIDE (all at $p < 10^{-4}$), and an insignificant effect on fBIRN.

Our results show that robustness to site differences varies between topological features, and that a consistent acquisition protocol can help reduce the effect of site differences on topological features.

V. MINIMISING SITE INFLUENCE

Given a multi-site dataset, how can we mitigate the confounding effect of having different sites? The dominant approach is to regress out the site identity, and to perform the rest of the modeling using prediction residuals. We evaluate how this approach performs in a cross-validation setting, and propose a projection-based approach that attempts to preserve class discrimination while reducing site effects.

Our graph representation is either a direct embedding (leading to 2828-dimensional feature vectors), or a concatenation of the six topological features mentioned above (leading to a 528-dimensional vectors ($88 \text{ regions} \times 6 \text{ properties}$)).

A. Cross-validation setup

For the ABIDE dataset, we look at an imaging marker scenario, where data from multiple sites is available to train an imaging marker, but we cannot see all subjects ahead of time. We want to evaluate how well the imaging marker is likely to perform on unseen subjects, so we estimate the generalisation ability by leaving a few subjects out for each site. The training data contains a total of N_{tr} subjects coming from a number of sites.

B. Regression approach

The regressing-out approach assumes a linear shift effect of the site, and mean-centers the features using site-specific means. However this is typically performed outside of a cross-validation loop, thereby benefitting from a peek at some testing data.

Here, we learn the parameters of the regression (beta coefficients) strictly on training data, before applying the prediction to both training and testing data. In addition, our approach to regression also includes regional quality measures SFNR and number of voxels, as well as a random effect for subject: we write our model $vp_i \sim 1 + SFNR_i + M_i + \text{scannerID} + (1|\text{subjectID})$.

C. Projection approach

For the projection approach, we modified the Nuisance Attribute Projection (NAP) technique [21], originally developed for channel compensation in speaker verification problems.

We start by defining two binary constraint matrices over the N_{tr} training feature vectors. First, a site constraint matrix $\mathbf{W}_s \in \{0, 1\}^{N_{tr} \times N_{tr}}$, where $\mathbf{W}_{s_{ij}}$ is 0 if feature vectors i and j are from the same site, and 1 if they are from different sites. Second, a class constraint matrix $\mathbf{W}_c \in \{0, 1\}^{N_{tr} \times N_{tr}}$, where $\mathbf{W}_{c_{ij}}$ is 1 if feature vectors i and j are from the same class, and 0 if they are from different classes. We obtained a linear combination of the two constraint matrices $\mathbf{W} = \mathbf{W}_s + \lambda \mathbf{W}_c$.

We then defined an optimisation problem to find the optimal projection matrix $\hat{\mathbf{P}}$ such that

$$\hat{\mathbf{P}} = \arg \min_{\mathbf{P}} \sum_{i \neq j} W_{ij} \|\mathbf{P}(\mathbf{x}_i - \mathbf{x}_j)\|^2, \quad (1)$$

where $\mathbf{x}_i$ are feature vectors. The optimal projection matrix will push feature vectors together if they are from different sites, and pull them apart if they are from different classes.

By writing $\mathbf{P} = \mathbf{I} - \mathbf{V}_{1:k} \mathbf{V}_{1:k}^T$, where $\mathbf{V}$ is a matrix with orthonormal columns and $\mathbf{V}^T \mathbf{V} = \mathbf{I}$, we can solve the optimisation problem as a generalised eigenvalue problem:

$$\mathbf{X}(\text{diag}(\mathbf{W}\mathbf{1}) - \mathbf{W})\mathbf{X}^T \mathbf{V} = \mathbf{V}\Lambda, \quad (2)$$

where the columns of $\mathbf{V}$ contains the eigenvectors sorted by decreasing eigenvalue. Thus, to train this method two free parameters must be tuned: $\lambda \in \mathbb{R}^+$, which selects the trade-off between reducing site differences (low λ) and maximising class differences (high λ), and $k \in \mathbb{N}^+$, which selects the number of leading eigenvectors retained. After training, the projection matrix was applied both to the training data and the testing data.

D. Statistical analysis of multi-site topological feature distributions

We focused the analysis on local efficiency, closeness centrality, and clustering coefficient, which we have shown to be site-dependent. Here, the goal of the regression and projection procedures was to minimise the effect of site. For both fBIRN and ABIDE, the regressing-out procedure succeeds in cancelling the site effect as measured by the mixed-model ANOVA. However the distributions are still mismatched. For fBIRN, the projection procedure succeeds in making the main effect of site insignificant ($p = 0.94, 0.95$ and 0.66 for local efficiency, closeness centrality, and clustering coefficient respectively). For ABIDE, the projection procedure fails to make the site effect insignificant.

E. Prediction accuracy in multi-site data

Using the ABIDE dataset (fBIRN has no class information), we first ran a leave-one-subject-per-group-out cross-validation on single sites (Fig 4). The classifier used was a random forest with 501 trees. Depending on sites and graph representations, accuracies ranged from 0.33 (below chance) to 0.75. While using 10 subjects per class per site yields high confidence intervals, accuracy appears strongly linked to site.

Then, we ran multisite tests with the original data, regression-corrected data and projection-corrected data in a cross-validated fashion (leaving one control and one patient out for each site). We used a nested-cross validation optimising for accuracy to set the λ and k parameters of the projection. In all three cases (original, regression, projection) we used both direct embedding and topological feature embedding. The results are shown in Fig. 4. The accuracy of the two embedding does not differ significantly, and is significantly above chance. 60% accuracy is consistent with the literature [22], but below the state of the art at 67% [23], although comparison is difficult due to the different subsets of subjects. Regression makes prediction random for both embeddings. Projection on topological features also makes prediction random, but a small (non-significant) improvement is obtained on direct embedding.

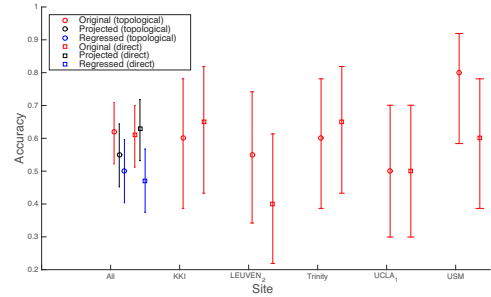


Fig. 4. Prediction accuracy for typically developing versus autism spectrum disorder subjects. The dot shows the point estimate for accuracy, the bars show the 95% binomial confidence interval. The right side of the plot show results for individual sites on the original data, while the left part shows results on all sites combined, for the original, regression, and projection approaches.

VI. CONCLUSIONS

Acquisition sites have significant and large effects on basic image properties. As shown using structural and functional quality measures, these effects can impact functional connectivity, and furthermore impact graph representations. This effect is both global and localised in space, since not all regions are affected equally by site effects.

Different topological features have various degrees of robustness. It is possible that unweighted properties, assorted with an edge-density thresholding procedure, are less susceptible to site effects, but this robustness can come at the expense of discriminative power.

Regressing-out the site identity may harm classification performance when performed in a cross-validated setting. This may be due to mean-alignment that does not rescale the distribution.

Discriminative projection approaches may be better suited to the task, although the current version of our proposed approach is does not perform particularly well. This is probably due to the difficulty of tuning the parameters of the projection within a nested cross-validation loop, where the data is high dimensional.

Meta-analysis techniques, which we have not explored here, may offer a good way forward by abstracting away signal differences and focusing on effect sizes directly.

ACKNOWLEDGMENT

This work was supported in part by a Marie Curie International Outgoing Fellowship of the European Community's 7th

framework programme (PIOF-GA-2011-299500), and in part by the Bavaria-California Technology Center. The fBIRN dataset is provided by the Biomedical Informatics Research Network under grants U24-RR021992 and U24 GM104203. The ABIDE dataset is supported by the NIMH under grant K23MH087770 and the Leon Levy Foundation.

REFERENCES

- [1] J. Richiardi, S. Achard, H. Bunke, and D. Van De Ville, "Machine learning with brain graphs: predictive modeling approaches for functional imaging in systems neuroscience," *IEEE Signal Processing Magazine*, pp. 58–70, May 2013.
- [2] F. X. Castellanos, A. Di Martino, R. C. Craddock, A. D. Mehta, and M. P. Milham, "Clinical applications of the functional connectome," *Neuroimage*, vol. 80, pp. 527–540, Oct 2013. [Online]. Available: <http://dx.doi.org/10.1016/j.neuroimage.2013.04.083>
- [3] W. R. Shirer, S. Ryali, E. Rykhlevskaia, V. Menon, and M. D. Greicius, "Decoding subject-driven cognitive states with whole-brain connectivity patterns," *Cerebral Cortex*, vol. 22, no. 1, pp. 158–165, May 2012.
- [4] S. Klöppel, C. Stonnington, J. Barnes, F. Chen, and C. C. et. al., "Accuracy of dementia diagnosis: a direct comparison between radiologists and a computerized method," *Brain*, vol. 131, no. 11, pp. 2969–2974, 2008.
- [5] D. Bernal-Casas, E. Balaguer-Ballester, and M. G. et. al., "Accuracy of dementia diagnosis: a direct comparison between radiologists and a computerized method," *Neuroimage*, vol. 82, pp. 555–563, 2013.
- [6] L. Friedman, G. H. Glover, and the FBIRN Consortium, "Reducing interscanner variability of activation in a multicenter fMRI study: Controlling for signal-to-fluctuation-noise-ratio (sfmr) differences," *Neuroimage*, vol. 33, pp. 471–481, 2006.
- [7] A. Abraham, "Study of the ABIDE dataset," <http://www.twincet.fr/site/research/blog/ABIDE>, 2013, online; accessed 07.09.2014.
- [8] J. Richiardi, H. Eryilmaz, S. Schwartz, P. Vuilleumier, and D. V. D. Ville, "Decoding brain states from fMRI connectivity graphs," *Neuroimage*, vol. 56, no. 2, pp. 616–626, 2011.
- [9] Y. Aleman-Gomez, L. Melie-Garcia, and P. Valdes-Hernandez, "IBASPM: Toolbox for automatic parcellation of brain structures," in *12th Annual Meeting of the Organization for Human Brain Mapping*, vol. 27, no. 1, 2006.
- [10] N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, and M. Joliet, "Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain," *Neuroimage*, vol. 15, no. 1, pp. 273–289, 2002.
- [11] S. Achard, J. Coeurjolly, R. Marcillaud, and J. Richiardi, "fMRI functional connectivity estimators robust to region size bias," in *Statistical Signal Processing Workshop (SSP), 2011 IEEE*, June 2011, pp. 813–816.
- [12] G. H. Glover and S. Lai, "Self-navigated spiral fMRI: interleaved versus single-shot," *Magn Reson Med*, vol. 39, no. 3, pp. 361–368, Mar 1998.
- [13] L. Friedman, G. H. Glover, and the FBIRN Consortium, "Reducing interscanner variability of activation in a multicenter fmri study: Controlling for signal-to-fluctuation-noise-ratio (sfmr) differences," *Neuroimage*, vol. 33, pp. 471–481, 2006.
- [14] K. Wang, T. Jiang, M. Liang, L. Wang, L. Tian, X. Zhang, K. Li, and Z. Liu, "Discriminative analysis of early Alzheimer's disease based on two intrinsically anti-correlated networks with resting-state fMRI," in *Proc. Int. Conf. on Medical Image Computing and Computer Assisted Intervention*, 2006, pp. 340–7.
- [15] G. A. Cecchi, I. Rish, B. Thyreau, B. Thirion, M. Plaze, M.-L. Paillere-Martinot, C. Martelli, J.-L. Martinot, and J.-B. Poline, "Discriminative network models of schizophrenia," in *Proc. Neural Information Processing Systems (NIPS)*, Vancouver, Canada, 2009, pp. 252–260.
- [16] J.-P. Onnela, J. Saramäki, J. Kertész, and K. Kaski, "Intensity and coherence of motifs in weighted complex networks," *Physical Review E*, vol. 71, no. 6, p. 065103, 2005. [Online]. Available: <http://journals.aps.org/pre/abstract/10.1103/PhysRevE.71.065103>
- [17] V. Latora and M. Marchiori, "Efficient behavior of small-world networks," *Physical Review Letters*, vol. 87, no. 19, Nov 2001. [Online]. Available: <http://dx.doi.org/10.1103/PhysRevLett.87.198701>
- [18] L. C. Freeman, "A set of measures of centrality based on betweenness," *Sociometry*, vol. 40, no. 1, pp. 35–41, Mar. 1977. [Online]. Available: <http://www.jstor.org/stable/3033543>
- [19] L. Freeman, "Centrality in networks: I. conceptual clarification," *Social Networks*, vol. 1, pp. 215–239, 1979.
- [20] S. Brin and L. Page, "The anatomy of a large-scale hypertextual web search engine," *Computer networks and ISDN systems*, vol. 30, no. 1, pp. 107–117, 1998. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S016975529800110X>
- [21] A. Solomonoff, W. M. Campbell, and C. Quillen, "Nuisance attribute projection," *Speech Communication*, 2007.
- [22] J. A. Nielsen, B. A. Zielinski, P. T. Fletcher, A. L. Alexander, N. Lange, E. D. Bigler, J. E. Lainhart, and J. S. Anderson, "Multisite functional connectivity MRI classification of autism: ABIDE results," *Front Hum Neurosci*, vol. 7, p. 599, 2013. [Online]. Available: <http://dx.doi.org/10.3389/fnhum.2013.00599>
- [23] C. P. Chen, C. L. Keown, and R.-A. Muller, "Diagnostic classification of autism using particle swarm optimization for fMRI feature selection," 2013, poster at the 19th Annual Meeting of the Organization for Human Brain Mapping.