Total-activation regularized deconvolution of resting-state fMRI leads to reproducible networks with spatial overlap

F. Işık Karahanoğlu

MGH/HST Athinoula A. Center for
Biomedical Imaging, Harvard Medical School
Boston, MA, USA
fkarahanoglu@mgh.harvard.edu

Dimitri Van De Ville

Medical Image Processing Lab
IBI-STI, Ecole Polytechnique Fédérale de Lausanne
University of Geneva, Switzerland
dimitri.vandeville@epfl.ch

Abstract—**Spontaneous activations in resting-state fMRI have been shown to corroborate recurrent intrinsic functional networks. Recent studies have explored integration of brain function in terms of spatially overlapping networks. We have proposed a method to recover not only spatially but also temporally overlapping networks, which we named innovation-driven co-activation patterns (iCAPs). These networks are driven by the sparse innovation signals recovered from Total Activation (TA), a spatiotemporal regularization framework for fMRI deconvolution. The fMRI data is processed with TA, which uses the inverse of the hemodynamic response function—as a linear differential operator—combined with the derivative in the regularization with ℓ_1-norm. As a result, sparse innovation signals are reconstructed as the deconvolved fMRI time series. Temporal clustering of innovation signals lead to iCAPs. In this work, we investigate the reproducible iCAPs in individuals with relapsing-remitting multiple sclerosis and healthy volunteers.**

I. INTRODUCTION

Functional magnetic resonance imaging (fMRI) measures the blood-oxygen-level-dependent (BOLD) signal that reflects the (de)oxygen concentration in the brain [1]. The relation between the measured BOLD signal and the neural activity can be explained through neurovascular coupling, which accounts for the changes in blood volume, blood flow, and oxygen consumption in the vessels. The balloon model introduced by Buxton et al. explained this link through a non-linear differential system [2], whose simplified form, hemodynamic response function (HRF), allows for linear, time-invariant analysis [3]–[5].

In conventional task-related fMRI analysis, the experimenter posits a hypothesis and designs an experimental paradigm during which fMRI data is acquired; e.g., watching a movie. Then, the measured BOLD signal is fitted to this a priori model, representing the exact timing of the experimental paradigm convolved with HRF, followed by statistical testing. The resulting activation patterns then relate solely to the designed experimental condition. Instead, during resting-state fMRI (rs-fMRI) acquisition subjects are not asked to perform a specific task but to relax without falling asleep. The spontaneous activity acquired during rs-fMRI reveals functional coordination between different brain regions in the absence of any extrinsic stimuli, and this coordination is very well represented in terms of large scale functional networks, referred to as resting-state networks (RSNs).

Since there is no a priori temporal model in rs-fMRI, data-driven methods have been adopted including seed correlation analysis [6], clustering analysis [7], [8] principal and independent component analysis (ICA) [9], [10], canonical correlation analysis [11]. Recent studies of functional connectivity have presented evidence towards the non-stationary brain dynamics [12], [13]. Existing rs-fMRI methods have been revisited; i.e., ICA and seed-based connectivity analysis have been adapted to allow for spatially overlapping patterns [14]–[18]. This new avenue led to a new understanding of brain processing, thereby, has potential to contribute to our current understanding of underlying brain dynamics in neurological diseases.

In order to investigate spontaneous brain dynamics in this emerging point of view, we proposed a new method to decompose the fMRI data into spatially and temporally overlapping activation patterns, which we named innovation-driven co-activation patterns (iCAPs) [19]. These activation patterns are the building blocks of rs-fMRI, where the activation of pattern at each time instance is represented as specific combination of iCAPs.

In order to extract the iCAPs, we first employ the Total Activation (TA) framework to detect intrinsic transient activations, represented as a sparse signal, for each voxel [20]. TA is cast as an optimization problem using regularization terms with spatial and temporal priors specifically tuned for fMRI data analysis. The spatial prior ensures coherence of activation within a given region of interest, and temporal prior acts as a deconvolution operator, assuming a block-like activity-inducing signal, with a generalized derivative operator representing the inverse of HRF [21]. TA reveals the 1) denoised, 2) deconvolved activity-inducing signals and 3) derivatives of activity-inducing signals; i.e., the sparse transients. The temporal clustering of sparse transient signals allows defining iCAPs.

In our previous work [19], we have found stable iCAPs

representations of rs-fMRI data in healthy volunteers. In this work, we elaborate on the stability and reproducibility of the spatial iCAP patterns in healthy individuals and individuals with relapsing-remitting multiple sclerosis (RRMS).

II. METHODS

A. Total activation

The TA framework assumes a generative model for the measured BOLD signal of ith voxel, $x(i, t)$ as

$$y(i,t) = x(i,t) + \epsilon(i,t) = u(i,t) * h(t) + \epsilon(i,t), \quad (1)$$

where $h(t)$ is the HRF, $\epsilon \sim \mathcal{N}(0, \sigma_i^2)$ is Gaussian noise, $u(i,t) = \sum_k c_k(i) b(t/a_k - t_k)$, is the block-like activity-inducing signal with weights c and step function $b(t)$.

TA aims to recover the underlying piecewise-constant activity-inducing signals, $u(i,t)$, by incorporating the inverse of the HRF as a linear differential operator. The final regularization problem is formulated as

$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{x}\|_F^2 + \mathcal{R}_T(\mathbf{x}) + \mathcal{R}_S(\mathbf{x}), \quad (2)$$

where

$$\mathcal{R}_T(\mathbf{x}) = \sum_{i=1}^{V} \lambda_1(i) \left\| \Delta_L \{\mathbf{x}(i, \cdot)\} \right\|_1, \text{ and} \quad (3)$$

$$\mathcal{R}_S(\mathbf{x}) = \sum_{t=1}^{N} \lambda_2(t) \left\| \Delta_{\text{Lap}} \{\mathbf{x}(\cdot, t)\} \right\|_{(2,1)}, \quad (4)$$

$\Delta_L = \Delta\Delta_{HRF}$ is the generalized derivative operator representing the inverse of HRF obtained through Balloon model, Δ_{HRF}, combined with a first-order derivative operator Δ, Δ_{Lap} is the Laplacian operator, N is the number of time points and V is the number of voxels. The temporal regularization inverts the effect of the HRF by promoting sparsity on the derivative of the activity-inducing signal, $\mathbf{u}$, thereby, the activity-inducing signal becomes a linear combination of step functions. This formulation is a generalized form of the total variation regularization in the presence of a linear time-invariant system. The spatial regularization implements the group-sparsity through $\ell_{(2,1)}$-norm within the predefined regions. Specifically, the spatial regularization promotes coherent activation within the regions of a predefined structural atlas while enforcing sparsity on the amount of active regions. We used the generalized forward-backward splitting algorithm in order to solve the regularization problem in temporal and space domains [22]. The algorithm was presented in detail in our previous work [20].

B. iCAPs

The sparse innovations: i.e., transient signals, provide useful information about the onsets and offsets of the underlying activation timeseries. In order to recover the activation patterns that share the concurrent transients, we clustered the transient signals with k-means algorithm. Specifically, we first determined the significant transients for each subject through two-step criteria: 1) we ran the TA algorithm once more on surrogate data, that is each subject's phase randomized timeseries, and computed a subject-specific threshold (1% confidence interval) from the transients in the surrogate data, and 2) we used a general threshold of at least 200 voxels that show the same transient ($\sim$2.5% of all voxels). The significant transients constitute 3200 out of 5280 total volumes in healthy controls (60.6%, 28.8% positive 31.8% negative) and 3250 out of 6160 total volumes in individuals with RRMS (52.8%, 24.6% positive 28.2% negative). Finally, these significant transients are included in the k-means algorithm.

C. Group-level clustering

We ran the k-means algorithm with number of cluster varying between [5,30], using cosine distance as the similarity measure. For each number of cluster, we performed 30 replicates with random initialization, and computed the sum of distance between each sample and the corresponding cluster centroid. The total cost function was computed as the sum of average distance within each cluster. Increasing the number of clusters will lead to overfitting and the cost function will slowly drop to 0. Therefore, we evaluated the difference of the total cost function, and look for the local minima, indicated by arrows in Fig. 1 (a).

Furthermore, we generated bootstrap samples (10 samples, with replacement), applied k-means on a training dataset (60%), and computed the cost function in a test dataset (40%) in Fig. 1 (b). Each bootstrap also used 10 random initializations where the minimum cost solution was picked as the most stable solution.

Finally, we ran the k-means algorithm opting for 20 clusters 10 times, each k-means had 30 random initialization where the best solution was picked as above. We matched the group centroids between 10 solutions using Hungarian algorithm [23]. We explored the reproducibility and stability of each cluster to define common-level iCAPs amongst our samples.

III. RESULTS

A. Data acquisition

The imaging data were acquired from fourteen healthy individuals (age: 38.4±6, 9F/5M) and twenty-two individuals with relapsing-remitting multiple sclerosis (RRMS, age: 36.8±8, 14F/8M), the subjects were the same as in [24]. Participants were scanned with a Siemens 3T TrioTIM (Erlangen, Germany) using a 32-channel head coil. The rs-fMRI data acquisition used gradient-echo echo-planar imaging and lasted around 8 minutes (TE/TR/FA=27ms/1.1s/90°, matrix = 64×64, 21 transverse slices, voxel size = 3.75×3.75×5.625mm^3, 450 volumes). T1-weighted anatomical scans were collected using an MPRAGE sequence (TR/TE/FA=2.98ms/2.4s/9°, 160 slices, voxel size = 1×1×1.2mm^3).

B. Preprocessing

The fMRI data were preprocessed using custom MATLAB code combined with SPM8 (FIL, UCL, UK) and IBASPM toolboxes [25]. First, fMRI volumes were realigned to the

mean scan and spatially smoothed with Gaussian filter (full width half maximum=5mm). We used motion estimation to mark the time points with high frame-wise displacement (≥ 0.5) [26]. Marked frames were not removed as TA relies on the full fMRI time courses to deconvolve, but we performed cubic-spline interpolation instead. Two healthy subjects and eight individuals with RRMS , who moved more than 3mm, were excluded from further analysis, therefore our analysis consisted twelve healthy individuals and fourteen individuals with RRMS. The anatomical images are coregistered onto the functional mean image and segmented (NewSegment, SPM8) for the six different MNI templates. The anatomical automatic labelling atlas, composed of 90 regions without the cerebellum, was mapped onto each subject's coregistered anatomical image and further downsampled to match the functional images. The atlasing was only used to guide TA spatial regularization. After TA analysis, each subjects fMRI volumes are normalized to MNI space (original fMRI data resolution) using the deformation matrix in the segmentation step in order to derive group level iCAPs.

C. Group-level iCAPs

The cost functions in Fig.1 suggest an optimal cluster number in the range of 15–20. Considering the subject-specific variability and maximum flexibility we adopted k=20. Fig. 2 shows the consistency and reproducibility of the clusters for each initialization (for 10 initializations we have $\binom{10}{2} = 45$ metrics for each cluster). The cluster similarity measurements provide an insight on the spatial reproducibility in Fig. 2 (top), however, they are not sufficient to define the most representative clusters; i.e., one subject might always drive one cluster. Therefore, we counted the number of instances assigned to each cluster per subject, and computed the mean and variation across subjects in Fig. 2 (bottom). For example, cluster 3 has relatively high reproducibility (Fig. 2 (top)), however it has driven only by a few patients and almost no controls (Fig. 2 (bottom)). Unlike cluster 3, the spatial similarity of cluster 13 is more variable in patients, but it has been consistently represented in both groups. We defined the representative iCAPs as the clusters that have high stability ($\geq$0.5 in Fig. 2 (top)) and similar reproducibility in each group (Fig. 2 (bottom)). The stable clusters are indicated by ($\ast$).

The final eleven group-level iCAPs are depicted in Fig. 3. The iCAP 1 shows coordinated precuneus and thalamus activity, iCAP 2 contains primary motor regions. The iCAP 3 includes the anterior salience network with dominant middle-inferior frontal gyrus and anterior cingulate activations. The iCAP 4 highlights the regions of attention network as well as precuneus, anterior and posterior cingulate. iCAP 5 and 6 represents spatiovisual and secondary visual areas, respectively. iCAP 7 is the default-mode network, and iCAP 8 shows the anterior cingulate. The iCAP 9 contains auditory regions, iCAP 10 represents the primary visual network. iCAP 11 covers the anterior executive network with superior frontal gyrus.
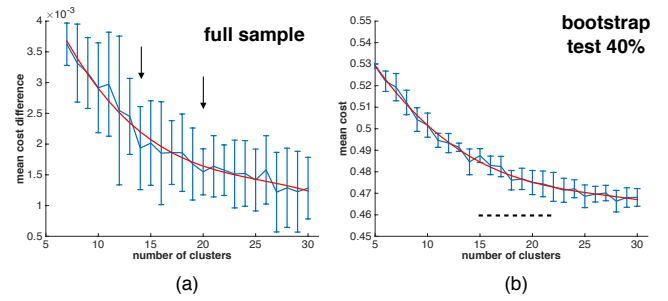

Fig. 1. The cost function for different number of clusters. The cost function, 1 - cosine distance, is computed as the total distance of samples to the cluster centroid within each cluster. (a) The difference of cost as a function of cluster number, each algorithm was ran with 30 random initialization using the full dataset. Local minima indicate stable solutions at k=14 and k=20 (with minimum variation). (b) The cost function of the bootstrap sample (training 60%, test 40%, 10 bootstraps, each with 10 random initialization). The cost function starts flattening between k=15–20.

D. Group-specific iCAPs

Fig. 4 depicts the group-specific iCAPs. We selected the most stable clustering replica and computed the mean z-normalized maps in the groups. In general, the spatial content of all group-specific iCAPs are consistent with the group-level iCAPs in Fig 3. The differences between the group-specific maps were focussed in very specific regions (Fig. 4 green arrows). Specifically in iCAP 1 controls have increased activation in thalamus in conjunction with precuneus with respect to RRMS patients ($t \geq 3$, $p \leq 10^{-4}$ uncorrected). Moreover, in the anterior executive control network (iCAP 11), RRMS subjects show increased and decreased activations in the superior frontal region's inferior and superior sections, respectively (($t \geq 3$, $p \leq 10^{-5}$ uncorrected)). Although the group-specific maps suggest spatial differences in middle frontal gyrus and fontal eye fields in iCAPs 3 and 5, respectively, the t-maps do not suggest significant differences.

IV. CONCLUSION

In this work, we have investigated the reproducible patterns of iCAPs in individuals with RRMS and healthy controls. We recovered the transient activity; i.e, the sparse innovation signals as the deconvolved fMRI time series through TA framework. The temporal clustering of innovation signals led to iCAPs, of which we investigated the reproducibility and consistency in the whole group. Finally, we found specific differences in thalamus and precuneus network (iCAP 1) as well as in superior frontal region (iCAP 11) between the RRMS patients and healthy volunteers.
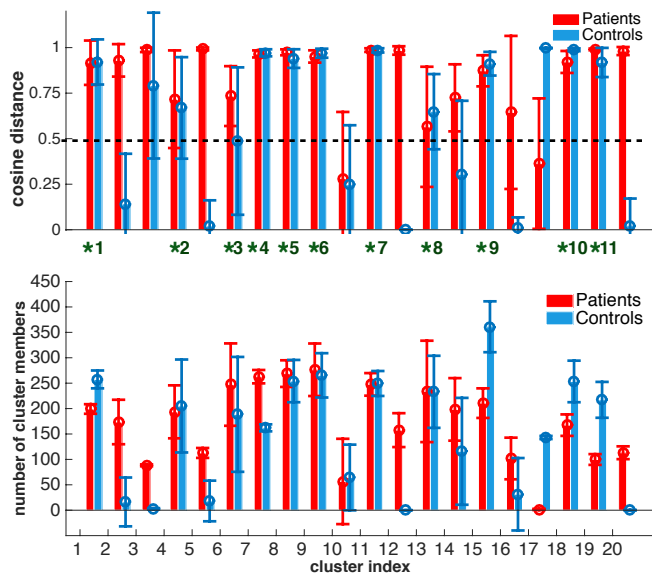
ACKNOWLEDGMENT

Fig. 2. The stability and reproducibility of iCAPs. We computed the average similarity, cosine distance, and standard deviation between each matching pair of cluster centroids; i.e., there are 45 pairs for 10 solutions (top). In order to insure that these clusters are consistently represented in every subject and group, we computed the number of volumes in each cluster for every subject, and computed the group mean and standard deviation (bottom). Having high variations and/or low values in either of these metrics indicate instability (∗ indicate the most stable configurations).
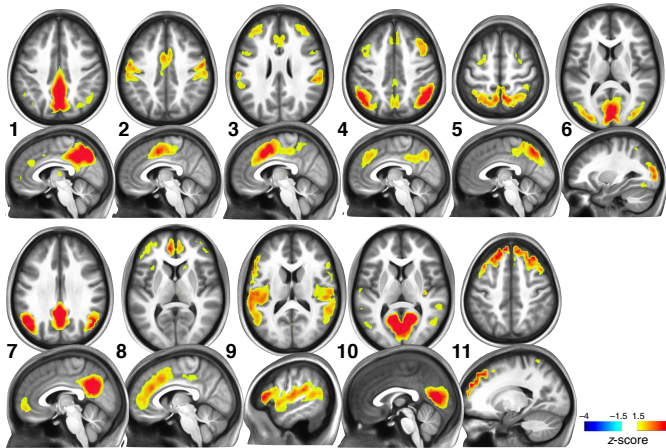

Fig. 3. The group-level iCAPs (z-normalized)

REFERENCES

[1] S. Ogawa, R. S. Menon, D. W. Tank, S. G. Kim, H. Merkle, J. M. Ellermann, and K. Ugurbil, "Functional brain mapping by blood oxygenation level-dependent contrast magnetic resonance imaging. a comparison of signal characteristics with a biophysical model," *Journal of Biophysics*, vol. 64, no. 3, pp. 803–12, Mar 1993.

[2] R. B. Buxton, E. C. Wong, and L. R. Frank, "Dynamics of blood flow and oxygenation changes during brain activation: The Balloon model." *Magnetic Resonance in Medicine*, vol. 39, no. 6, pp. 855–864, Jun 1998.
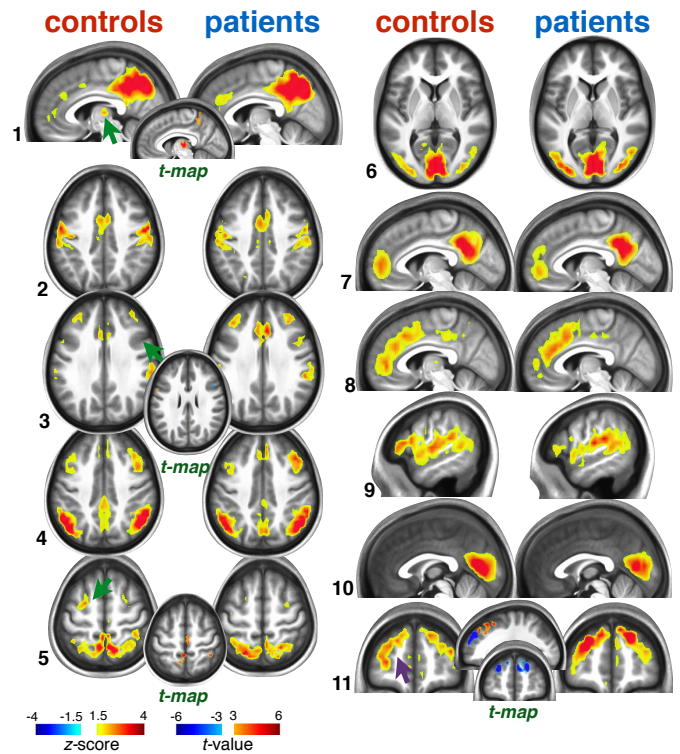
Fig. 4. The group-specific iCAPs (z-normalized). The group-specific maps of healthy volunteers and individuals with RRMS reveal slight differences between the groups (indicated by green arrows). The controls show higher coordination between thalamus and precuneus (iCAP 1) than patients ($t \geq 3$, $p \leq 10^{-4}$ uncorrected). In iCAP 3, the controls have lower right middle frontal gyrus, however the t-maps do not suggest significance. The activation in areas surrounding the frontal eye fields in iCAP5 has lower but not significant intensity in patients with respect to controls ($t \leq 3$). Patients have increased activity in the inferior sections of superior frontal gyrus whereas controls have increased activation in the superior sections of superior frontal gyrus.

[3] K. J. Friston, P. Fletchera, O. Josephs, A. Holmes, M. D. Rugg, and R. Turner, "Event-Related fMRI: Characterizing Differential Responses," *NeuroImage*, vol. 7, no. 1, pp. 30–40, Jan. 1998.

[4] K. J. Friston, A. Mechelli, R. Turner, and C. J. Price, "Nonlinear responses in fMRI: The balloon model, Volterra kernels, and other hemodynamics," *NeuroImage*, vol. 12, no. 4, pp. 466 – 477, 2000.

[5] I. Khalidov, D. Van De Ville, J. Fadili, and M. Unser, "Activelets and sparsity: A new way to detect brain activation from fMRI data," in *Proceedings of the SPIE Conference on Mathematical Imaging: Wavelet XII*, vol. 6701, San Diego CA, USA, August 26-29, 2007, pp. 1–8.

[6] B. Biswal, F. Z. Yetkin, V. Haughton, and J. Hyde, "Functional connectivity in the motor cortex of resting human brain using echo-planar MRI." *Magnetic Resonance in Medicine*, vol. 34, no. 4, pp. 537–541, Oct. 1995.

[7] R. Baumgartner, L. Ryner, W. Richter, R. Summers, M. Jarmasz, and R. Somorjai, "Comparison of two exploratory data analysis methods for fMRI: fuzzy clustering vs. principal component analysis," *Magnetic Resonance Imaging*, vol. 18, no. 1, pp. 89 – 94, 2000.

[8] V. L. Morgan, Y. Li, B. Abou-Khalil, and J. C. Gore, "Development of 2dTCA for the detection of irregular, transient BOLD activity," *Human Brain Mapping*, vol. 29, no. 1, pp. 57–69, 2008.

[9] M. J. McKeown, S. Makeig, G. G. Brown, T. P. Jung, S. S. Kindermann, A. J. Bell, and T. J. Sejnowski, "Analysis of fMRI data by blind separation into independent spatial components," *Human Brain Mapping*, vol. 6, no. 3, pp. 160–88, Jan 1998.

[10] C. F. Beckmann and S. M. Smith, "Probabilistic independent component analysis for fMRI," *IEEE Transactions on Medical Imaging*, vol. 23, no. 2, pp. 137 –152, feb. 2004.

[11] B. Afshin-Pour, G.-A. Hossein-Zadeh, S. C. Strother, and H. Soltanian-Zadeh, "Enhancing reproducibility of fMRI statistical maps using generalized canonical correlation analysis in NPAIRS framework," *NeuroImage*, vol. 60, no. 4, pp. 1970 – 1981, 2012.

[12] E. A. Allen, E. Damaraju, S. M. Plis, E. B. Erhardt, T. Eichele, and V. D. Calhoun, "Tracking whole-brain connectivity dynamics in the resting state," *Cereb Cortex*, vol. 24, no. 3, pp. 663–76, 2014.

[13] N. Leonardi, J. Richiardi, M. Gschwind, S. Simioni, J.-M. Annoni, M. Schluep, P. Vuilleumier, and D. V. D. Ville, "Principal components of functional connectivity: A new approach to study dynamic brain connectivity during rest," *NeuroImage*, vol. 83, no. 0, pp. 937 – 950, 2013.

[14] S. M. Smith, K. L. Miller, S. Moeller, J. Xu, E. J. Auerbach, M. W. Woolrich, C. F. Beckmann, M. Jenkinson, J. Andersson, M. F. Glasser, D. C. V. Essen, D. A. Feinberg, E. S. Yacoub, and K. Ugurbil, "Temporally-independent functional modes of spontaneous brain activity," *Proceedings of the National Academy of Sciences*, vol. 8, pp. 3131–3136, 2012.

[15] B. T. T. Yeo, F. M. Krienen, M. W. L. Chee, and R. L. Buckner, "Estimates of segregation and overlap of functional connectivity networks in the human cerebral cortex." *NeuroImage*, vol. 88C, pp. 212–227, Nov. 2013.

[16] 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*, vol. 110, no. 11, pp. 4392–4397, 2013.

[17] X. Liu, C. Chang, and J. H. Duyn, "Decomposition of spontaneous brain activity into distinct fMRI co-activation patterns," *Frontiers in Systems Neuroscience*, vol. 7, no. 101, 2013.

[18] S. J. Harrison, M. W. Woolrich, E. C. Robinson, M. F. Glasser, C. F. Beckmann, M. Jenkinson, and S. M. Smith, "Large-scale probabilistic functional modes from resting state fMRI." *NeuroImage*, vol. 109, pp. 217–231, Apr. 2015.

[19] F. I. Karahanoğlu and D. Van De Ville, "Transient brain activity disentangles fMRI resting-state dynamics in terms of spatially and temporally overlapping networks." *Nature Communications*, vol. 6, p. 7751, 2015.

[20] F. I. Karahanoglu, C. Caballero-Gaudes, F. Lazeyras, and D. Van De Ville, "Total activation: fMRI deconvolution through spatio-temporal regularization," *NeuroImage*, vol. 73, pp. 121 – 134, 2013.

[21] F. I. Karahanoglu, I. Bayram, and D. Van De Ville, "A signal processing approach to generalized 1-D total variation," *IEEE Transactions on Signal Processing*, vol. 59, no. 11, pp. 5265 –5274, November 2011.

[22] H. Raguet, J. Fadili, and G. Peyre, "A generalized forward-backward splitting," *SIAM Journal on Imaging Sciences*, vol. 6, no. 3, pp. 1199–1226, 2013.

[23] H. W. Kuhn and B. Yaw, "The hungarian method for the assignment problem," *Naval Res. Logist. Quart*, pp. 83–97, 1955.

[24] J. Richiardi, M. Gschwind, S. Simioni, J.-M. Annoni, B. Greco, P. Hagmann, M. Schluep, P. Vuilleumier, and D. Van De Ville, "Classifying minimally-disabled multiple sclerosis patients from resting-state functional connectivity," *NeuroImage*, vol. 62, pp. 2021–2033, 2012. [Online]. Available: http://dx.doi.org/10.1016/j.neuroimage.2012.05.078

[25] Y. Alemán-Gómez, L. Melie-Garćia, and P. Valdés-Hernandez, "IBASPM: Toolbox for automatic parcellation of brain structures," *12th Annual Meeting of the Organization for Human Brain Mapping*, vol. 27, 2006.

[26] J. Power, A. Kelly, Z. Abraham, L. Bradley, and E. Steven, "Spurious but systematic correlations in functional connectivity mri networks arise from subject motion," *NeuroImage*, vol. 59, no. 3, pp. 2142–2154, 2012.