A Graph Signal Processing Framework for the Classification of Temporal Brain Data

Sarah Itani
Department of Mathematics and Operations Research
University of Mons
Mons, Belgium
sarah.itani@umons.ac.be

Dorina Thanou
Swiss Data Science Center (SDSC)
EPFL and ETH Zürich
Switzerland
dorina.thanou@epfl.ch

***Abstract*—Graph Signal Processing (GSP) addresses the analysis of data living on an irregular domain which can be modeled with a graph. This capability is of great interest for the study of brain connectomes. In this case, data lying on the nodes of the graph are considered as signals (e.g., fMRI time-series) that have a strong dependency on the graph topology (e.g., brain structural connectivity). In this paper, we adopt GSP tools to build features related to the frequency content of the signals. To make these features highly discriminative, we apply an extension of the Fukunaga-Koontz transform. We then use these new features to train a decision tree for the prediction of autism spectrum disorder. Interestingly, our framework outperforms state-of-the-art methods on the publicly available ABIDE dataset.**

***Index Terms*—Graph signal processing; machine learning; explainability; decision trees; functional MRI; autism spectrum disorder.**

I. Introduction

Despite medical progress, Autism Spectrum Disorder (ASD) remains subject to the absence of a recognized etiology. Research continues in quest of explanatory biomarkers that would help for early diagnosis. In this respect, modern data-driven techniques are expected to shed light on interpretable patterns that capture the complexity of this neuropathology [1].

Brain connectivity has been exhaustively explored in an attempt to better understand ASD [2], [3]. This measure of interaction is usually studied from a functional or a structural perspective. Functional Connectivity (FC) typically derives from the correlation of signals such as Blood-Oxygen-Level Dependent (BOLD) ones, measured by resting-state functional Magnetic Resonance imaging (rs-fMRI). Structural Connectivity (SC) is related to the detection of white matter pathways, tracked, for example, by diffusion tensor imaging. Both FC and SC can be analyzed through data mining to reveal ASD-related organizing principles of the brain, e.g., [4], [5].

Given the complex nature of the brain and the paramount importance of both FC and SC, in this work, we associate both information for the extraction of discriminative features. Towards this goal, we borrow tools from Graph Signal Processing (GSP) [6]. GSP allows to integrate both structural and functional brain data by studying the interplay between graphs and signals on graphs. Currently, it has been successfully used for the study of cognitive flexibility [7] and motor skill [8].

In this work, our goal is to classify ASD from neurotypical subjects. The graph in our application consists of a set of brain regions of interest, i.e., graph nodes, that are connected based on the distance between them. This graph structure remains constant over time. In contrast, the data of each subject observed on top of the graph (i.e., graph signals) are of time-varying nature, since they are generated by the BOLD fluctuations. The problem thus boils down to the classification of time-varying graph signals. Our study aims at extending the GSP tools further, by adapting them to the particular settings. For such a purpose, we study the frequency behavior of the BOLD signals by computing their Graph Fourier Transform (GFT). The time-varying GFT coefficients are then merged into a connectivity matrix, from which we compute discriminative graph frequency patterns, by using an extension of the Fukunaga-Koontz transform. We then exploit these features to separate ASD from neurotypical subjects, using an interpretable classification scheme such as the decision tree. Finally, we note that the proposed framework can be advantageously generalized to the classification of time-varying graph signals other than brain data.

The contributions of our study are summarized below.

(1) We extend the GSP framework by proposing a way of building discriminative patterns in the graph Fourier domain. This is convenient when the neuropathology may not be attributed to abnormalities in specific frequency bands, i.e., low, middle and high bands, contrarily to what was proposed in previous GSP-based works, e.g., [7]–[9].
(2) Our discriminative patterns can be classified by a simple decision tree without the necessity of using more complex classification schemes, e.g., through deep learning as was considered in [10]. This facilitates to a certain extent the interpretation of the predictions.
(3) In terms of classification accuracy, our framework outperforms other state-of-the-art methods that are based either on SC or FC, i.e., the GFT [6], [9] and the Spatial Filtering Method (SFM) [11]. Moreover, the interpretation of the results confirms previous findings of the neuroscience literature for the ASD case.

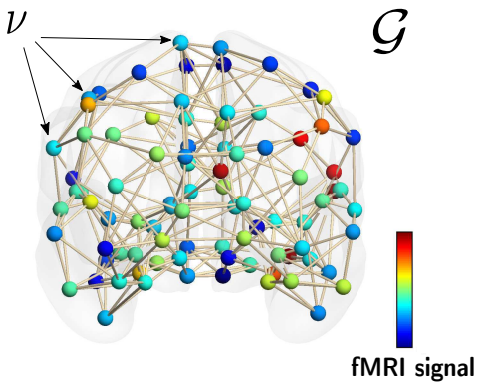

Fig. 1: Representation of a graph signal, at a given time point.

II. Preliminaries

A. Settings and notations

Our classification problem consists in identifying NeuroTypical (NT), and ASD subjects. We denote by n_T, the total number of subjects; n_A, the total number of ASD subjects; n_N, the total number of NT subjects.

Resting-state fMRI signals (i.e., BOLD time-series) are available for each subject, and for a set of r brain Regions of Interest (ROIs). For each subject i ($i = \{1, \dots n_T\}$), we denote by T_i the number of time-points in each fMRI signal, and $\mathbf{X}_i$, the $r \times T_i$ matrix of time-series.

Note that BOLD fluctuations carry a notion of structure, since they are measured for each ROI. However, in this paper, what we denote as the structural dimension of the data relates to the structural connectivity of the brain (i.e., topology). The functional dimension of the data relates to aspects of brain activity, derived from the BOLD fluctuations.

B. Graph-based representation of brain signals

We model the structure of the brain as an undirected, weighted graph $\mathcal{G}$, where the set of nodes ν corresponds to the brain ROIs (see Fig. 1). The edges of the graph are defined by connecting close nodes in terms of their geometrical distance in the brain. In particular, we define the weight $\mathbf{A}_{uv}$ between two nodes u, v of a brain graph $\mathcal{G}$ as the inverse of the Euclidean distance d_{uv} between both nodes. Thus, the adjacency matrix $\mathbf{A}$ is such that:

$$A_{uv} = d_{uv}^{-1} \quad \text{and} \quad A_{uu} = 0 \quad \text{for} \quad u, v = 1, \dots r.$$

For each node, we keep only its K nearest neighbors, while ensuring that the final graph is symmetric; this strategy is a good proxy for representing the brain topology [12], [13].

Fig. 1 illustrates a 5-nearest neighbor topology of the brain, on top of which a fMRI signal is observed at a specific instant of time[1]. This graph signal consists of a vector of r values, which corresponds to a column vector of the matrix $\mathbf{X}_i$ (cf. Subsec. II-A). The spectral domain representation can reveal significant information about the characteristics of those signals. The Graph Fourier Transform (GFT) provides a frequency analysis of the signals that reside on the graph, based on the graph Laplacian operator. The combinatorial Laplacian operator [15] is defined as $\mathbf{L} = \mathbf{D} - \mathbf{A}$ where $\mathbf{A}$ is the graph adjacency matrix and $\mathbf{D}$ is a diagonal matrix including the degree of each node, i.e., $\mathbf{D}_{kk} = \sum_j \mathbf{A}_{kj}$.

The eigenvectors of the Laplacian operators can be used to perform an harmonic analysis of the graph signals, and the corresponding eigenvalues carry a notion of frequency [15]. We thus study the BOLD time-series $\mathbf{X}_i$ related to a patient i in the frequency domain, through their projection on the eigenvectors of a Laplacian matrix $\mathbf{V}$, i.e.,

$$\hat{\mathbf{X}}_i = \mathbf{V}^T \mathbf{X}_i. \tag{1}$$

The columns of $\hat{\mathbf{X}}_i$ thus correspond to the GFT coefficients for each frequency mode at a given instant.

III. Our framework

We propose to handle both the structural and functional information of the BOLD signals to extract discriminative patterns in the graph spectral domain. Our approach consists of three steps detailed in the following subsections.

A. Merging function, structure, and temporality

The first step of our method consists of studying the frequency behavior of the signals on the graph, by computing the GFT coefficients of each subject's BOLD time-series, using Eq. 1. In order to understand the relative importance of each frequency component, we normalize the GFT coefficients at each time point. More specifically, we normalize the columns of $\hat{\mathbf{X}}_i$ by subtracting the mean of each column and dividing by the $L2$ norm. The resulting matrices are denoted by $\mathbf{Y}_i$. We then merge the normalized GFT coefficients over time to reveal some frequency patterns of the time-series in the graph Fourier domain. We achieve that by computing a proxy of the sample covariance matrix, that is given by:

$$\mathbf{S}_i = \frac{\mathbf{Y}_i {\mathbf{Y}_i}^T}{\mathrm{tr}(\mathbf{Y}_i {\mathbf{Y}_i}^T)}, \tag{2}$$

where $\mathrm{tr}()$ denotes the trace operator. The mean joint expectancy matrix over the whole set of patients is:

$$\bar{\mathbf{S}} = \frac{1}{n_T} \sum_{i=1}^{n_T} \frac{\mathbf{Y}_i {\mathbf{Y}_i}^T}{\mathrm{tr}(\mathbf{Y}_i {\mathbf{Y}_i}^T)}. \tag{3}$$

The mean joint expectancy matrices $\bar{\mathbf{S}}^{\mathbf{A}}$ and $\bar{\mathbf{S}}^{\mathbf{N}}$ for all ASD and NT patients respectively are:

$$\bar{\mathbf{S}}^{\mathbf{A}} = \frac{1}{n_A} \sum_{i=1}^{n_A} \mathbf{S}_i \quad \text{and} \quad \bar{\mathbf{S}}^{\mathbf{N}} = \frac{1}{n_N} \sum_{i=1}^{n_N} \mathbf{S}_i.$$

Combining the two last equations, it follows that:

$$\bar{\mathbf{S}} = \frac{n_A}{n_T} \bar{\mathbf{S}}^{\mathbf{A}} + \frac{n_N}{n_T} \bar{\mathbf{S}}^{\mathbf{N}} = \frac{n_A}{n_T} \bar{\mathbf{S}}^{\mathbf{A}} + (1 - \frac{n_A}{n_T}) \bar{\mathbf{S}}^{\mathbf{N}}. \tag{4}$$

[1]The figures presented in this paper were drawn with the BrainNet Software [14].

For the remainder of the development, we denote by α_A and α_N respectively, the factors $\frac{n_A}{n_T}$ and $\frac{n_N}{n_T}$, and with $\alpha_N = 1 - \alpha_A$. Eq. 4 can be reformulated as:

$$\bar{\mathbf{S}} = \alpha_{\mathrm{A}}\bar{\mathbf{S}}^{\mathbf{A}} + \ \alpha_{\mathrm{N}}\bar{\mathbf{S}}^{\mathbf{N}}. \tag{5}$$

The mean joint expectancy matrix is thus expressed as a positive linear combination of both mean joint expectancy matrices of ASD and NT subjects.

B. Finding a discriminative subspace through FKT

After computing a matrix that captures the temporal evolution of the graph spectral components of the BOLD signals for each class (i.e., $\bar{\mathbf{S}}^{\mathbf{A}}, \bar{\mathbf{S}}^{\mathbf{N}}$), we need to classify the subjects in one of the two categories. To achieve this, we follow an approach inspired by the Fukunaga-Koontz Transform (FKT), which allows to find a discriminative subspace to project the GFT coefficients of the BOLD signals. This operation is achieved in three steps which are detailed in the following subsections. A similar approach is followed by [11].

1) Whitening: First, we need to decorrelate the data by means of a whitening operator. Let us consider the eigen-decomposition of the matrix $\bar{\mathbf{S}}$:

$$\mathbf{Q}^{\mathbf{T}}\bar{\mathbf{S}}\mathbf{Q} = \mathbf{\Lambda} \quad \Leftrightarrow \quad \bar{\mathbf{S}} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{\mathbf{T}}.$$

As $\bar{\mathbf{S}}$ is symmetric, it holds that $\mathbf{Q^{-1}} = \mathbf{Q^T}$. It can be shown that $\bar{\mathbf{S}}$ has a zero eigenvalue because of the column-wise normalization of the GFT coefficients [16]. We consider, without loss of generality, that the eigenvalues of $\bar{\mathbf{S}}$ are sorted in ascend order, thus $\mathbf{\Lambda}_{11} = 0$. We then define the diagonal matrix $\mathbf{\Gamma}$ such that:

$$\mathbf{\Gamma}_{11} = 1 \quad \text{and} \quad \mathbf{\Gamma}_{ii} = \mathbf{\Lambda}_{ii}^{-1/2}, \ \text{with} \ i = 2, \dots r.$$

Thus, if we set $\mathbf{Q_2} = \mathbf{\Gamma^T}\mathbf{Q^T}$, we have:

$$\mathbf{Q_2}\bar{\mathbf{S}}\mathbf{Q_2}^{\mathbf{T}} = \text{diagonal}[\mathbf{0}, \mathbf{I_{r-1}}].$$

where diagonal$[\mathbf{0}, \mathbf{I_{r-1}}]$ is a diagonal matrix whose first diagonal element is 0 and the $r-1$ remaining ones equal to one. Using the above developments, Eq. 5 is rewritten as:

$$\begin{aligned}\mathbf{Q_2}\bar{\mathbf{S}}\mathbf{Q_2^T} &= \alpha_{\mathrm{A}}\mathbf{Q_2}\bar{\mathbf{S}}^{\mathbf{A}}\mathbf{Q_2^T} + \ \alpha_{\mathrm{N}}\mathbf{Q_2}\bar{\mathbf{S}}^{\mathbf{N}}\mathbf{Q_2^T} \\ \Leftrightarrow \text{diagonal}[\mathbf{0}, \mathbf{I_{r-1}}] &= \alpha_{\mathrm{A}}\bar{\mathbf{S}}^{\mathbf{A}'} + \ \alpha_{\mathrm{N}}\bar{\mathbf{S}}^{\mathbf{N}'}.\end{aligned} \tag{6}$$

2) Simultaneous diagonalization of the matrices: Then, we have to find a transform which keeps the overall effect of *whitening*, while making the variance of the classes emerge in a complementary way. These operations constitute the simultaneous diagonalization of the joint expectancy matrices ($\bar{\mathbf{S}}$, $\bar{\mathbf{S}}^{\mathbf{A}}$, $\bar{\mathbf{S}}^{\mathbf{N}}$), which are actually related by Eq. 5.

We can show that $\bar{\mathbf{S}}^{\mathbf{A}'}$ is positive semi-definite and has the following structure [16]:

$$\bar{\mathbf{S}}^{\mathbf{A}'} = \left[\begin{array}{c|c} 0 & \dots \ 0 \\ \hline \vdots & \multirow{2}{*}{$\bar{\mathbf{S}}^{\mathbf{A}'}_{\mathbf{r-1}}$} \\ 0 & \end{array}\right].$$

In order to diagonalize $\bar{\mathbf{S}}^{\mathbf{A}'}$, Newcomb [17] proposes to diagonalize $\bar{\mathbf{S}}^{\mathbf{A}'}_{\mathbf{r-1}}$ by an orthogonal transformation $\mathbf{T}'$ which is deduced by eigen-decomposition. The global transformation matrix $\mathbf{T_2}$ constitutes of the following:

$$\mathbf{T_2} = \left[\begin{array}{c|c} 1 & 0 \ \dots \ 0 \\ \hline 0 & \multirow{3}{*}{$\mathbf{T}'$} \\ \vdots & \\ 0 & \end{array}\right].$$

Thus, $\mathbf{T_2^T}\bar{\mathbf{S}}^{\mathbf{A}'}\mathbf{T_2} = \bar{\mathbf{S}}^{\mathbf{A}''}$, where $\bar{\mathbf{S}}^{\mathbf{A}''}$ is a diagonal matrix. Eq. 6 can be reformulated as:

$$\begin{aligned}\mathbf{T_2^T}\ \text{diagonal}[\mathbf{0}, \mathbf{I_{r-1}}]\ \mathbf{T_2} &= \alpha_{\mathrm{A}}\mathbf{T_2^T}\bar{\mathbf{S}}^{\mathbf{A}'}\mathbf{T_2} + \ \alpha_{\mathrm{N}}\mathbf{T_2^T}\bar{\mathbf{S}}^{\mathbf{N}'}\mathbf{T_2} \\ \Leftrightarrow \ \text{diagonal}[\mathbf{0}, \mathbf{I_{r-1}}] &= \alpha_{\mathrm{A}}\bar{\mathbf{S}}^{\mathbf{A}''} + \alpha_{\mathrm{N}}\bar{\mathbf{S}}^{\mathbf{N}''}.\end{aligned}$$

Given that diagonal$[\mathbf{0}, \mathbf{I_{r-1}}]$ and $\bar{\mathbf{S}}^{\mathbf{A}''}$ are diagonal matrices, $\bar{\mathbf{S}}^{\mathbf{N}''}$ is diagonalizable, and it shares the same eigenvectors with $\bar{\mathbf{S}}^{\mathbf{A}''}$. The non-zero eigenvalues of $\bar{\mathbf{S}}^{\mathbf{A}''}$ and $\bar{\mathbf{S}}^{\mathbf{N}''}$, multiplied respectively by α_{A} and α_{N}, are complementary and sum to unity.

3) Computation of the projection matrix: All the above operations can be summarized through a final projection matrix $\mathbf{P}$ such that:

$$\begin{aligned}\mathbf{P}\bar{\mathbf{S}}\mathbf{P^T} &= \alpha_{\mathrm{A}}\mathbf{P}\bar{\mathbf{S}}^{\mathbf{A}}\mathbf{P^T} + \ \alpha_{\mathrm{N}}\mathbf{P}\bar{\mathbf{S}}^{\mathbf{N}}\mathbf{P^T} \\ \Leftrightarrow \text{diagonal}[\mathbf{0}, \mathbf{I_{r-1}}] &= \alpha_{\mathrm{A}}\bar{\mathbf{S}}^{\mathbf{A}''} + \ \alpha_{\mathrm{N}}\bar{\mathbf{S}}^{\mathbf{N}''}.\end{aligned}$$

with

$$\mathbf{P} = \mathbf{T_2^T}\mathbf{Q_2} = \mathbf{T_2^T}\mathbf{\Gamma^T}\mathbf{Q^T}. \tag{7}$$

Thus, we end up with a matrix which can be used to project each patient's matrix of time-varying GFT coefficients in a space where the ASD and NT classes have complementary mean joint expectancy matrices. This explains why the subspace is discriminative: each class may be expressed through a subset of dimensions along which the variance of the related data is high.

C. Performing classification

Now that we have computed the discriminative projection matrix, we can use it to classify our GFT coefficients of the BOLD signals. Thus, we project the GFT coefficients into the discriminative matrix $\mathbf{P}$ (see Eq. 7):

$$\mathbf{Z}_i = \mathbf{P}\mathbf{Y}_i. \tag{8}$$

This projection can be interpreted as filtering of the GFT coefficients by discriminative filters. Classification is then achieved by training a decision tree on the variance of the projected GFT coefficients, as described in the following section. We note that the variance of the elements included in the first row of $\mathbf{Z}_i$ does not carry any discriminative information. Indeed, along this first dimension, both classes are associated with a zero eigenvalue (see Eq. 7). This is due to the singularity of the mean joint expectancy matrix (see Eq. 3).

IV. Experiments

A. Experimental settings

Data – We use the ABIDE I preprocessed dataset [18] which includes BOLD signals preprocessed by the C-PAC pipeline [19]. We selected time-series provided for the Automated Anatomical Labeling atlas on 90 ROIs (AAL90) [20]. We consider patients with eyes opened during the fMRI session; less than 18 years old; less than 0.2 mm in mean framewise displacement. The final data subset includes 251 NT and 201 ASD subjects.

Brain topology – As explained in Sec. II-B, the brain nodes are connected with respect to their geometric distance which is computed as the difference of their coordinates in the MNI space [21]. We focus more particularly on a two-nearest neighbor topology, which roughly divides the brain in two parts: the fronto-temporal areas and the parieto-occipital areas. This choice of topology is meaningful from the neuroscience point of view. Indeed, fronto-temporal areas have been associated to dysfunctions and structural abnormalities in ASD [22]–[24].

Training features and classifier – The initial dataset is randomly split into training (95%) and test sets (5%). We report the average test accuracy over ten partitions of the dataset. We compare our framework with two different approaches: (i) based on the classification of the GFT coefficients (cf. Sec. II-B) and (ii) based on the Spatial Filtering Method (SFM) [11] which, to the best of our knowledge, is the closest approach to ours.

- Concerning GFT, we compute the variance of the normalized GFT coefficients over time, for the whole frequency bands, and for equally-defined frequency bands, i.e., low, middle and high frequency modes.
- For both SFM and our framework, we consider the variance of the projected data related to the m most significant dimensions for each group, with $m \in [2, 5]$. Therefore, the total number of training features is $2m$.

A decision tree [25] is trained on the training sets within each trial. This classifier is convenient for diagnosis prediction, since it ensures the interpretability of the prediction outputs.

B. Results and discussion

1) Comparison with state-of-the-art methods: Table I presents the averaged test accuracies for different assessment modalities (further results are available in [16]). The gain of 4.4% achieved by our framework against SFM indicates that the structure-function interplay improves prediction. In comparison to considering directly the GFT coefficients, it seems that the discriminative information is hidden in more complex patterns, revealed by their combination, as suggested by Eq. 8. Moreover, we analyzed statistically the differences between SFM and our framework, through a Student's t-test. The hypothesis of equal performances is rejected for all the modalities (i.e., values for m), with p-values inferior to 5%.

m	SFM (%)	Ours (%)	Modality	GFT (%)
2	69.1	**73.5**	GFT	53.9
3	70.4	**74.8**	GFT_{LOW}	59.6
4	69.6	**73.0**	GFT_{MID}	62.6
5	67.4	**71.3**	GFT_{HIGH}	59.1

TABLE I: Averaged test accuracy

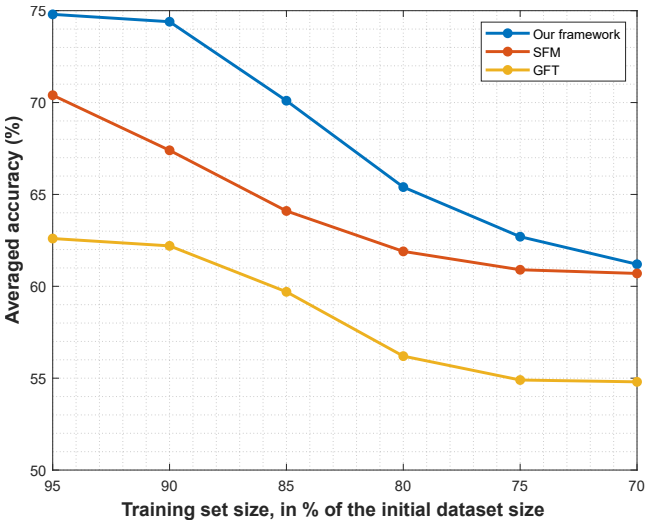

Fig. 2: Performances for different training set sizes

2) Influence of the training set size: Fig. 2 shows the best average test accuracies achieved by each method (across all GSP modalities), according to the procedure described in Sec. IV-A, for different sizes of the initial dataset (from 95% to 70%, per step of 5%). Overall, our method outperforms both SFM and GFT. As expected, as the training set size decreases, the performance deteriorates significantly for all the methods. We note that both SFM and our method are based on the estimation of a covariance matrix through the sample covariance, which requires a large amount of data. This is particularly true for the ASD population which is quite heterogeneous in profiles, given the extent of the spectrum.

3) Influence of the topology: In order to understand the influence of the topology on the classification performance, we perform two additional experiments, by considering a Weighted Fully Connected (WFC) and an Uniformly Connected (UC) topologies. Table II reports the corresponding predictive performances in comparison to those achieved with a 2-NN topology and the SFM method. These results confirm the influence of the structure. In particular, it is worth noting that the performances achieved with the WFC topology are inferior to those achieved in the case of 2-NN. This shows

m	SFM (%)	Our framework (%)		
		2-NN	WFC	UC
2	69.1	**73.5**	67.8	66.1
3	70.4	**74.8**	67.8	69.1
4	69.6	**73.0**	70.0	65.2
5	67.4	**71.3**	65.2	63.9

TABLE II: Influence of the topology on the performances

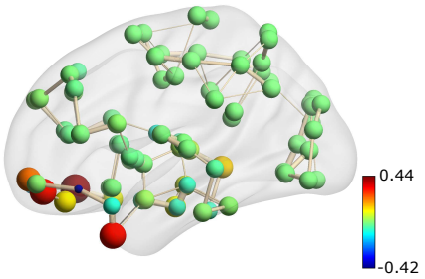

Fig. 3: Mode 44 (dominant in ASD): neutral on parieto-occipital regions, activated on fronto-temporal ones

that probably the edges added to the 2-NN graph, i.e., non-local interactions, add some noise in the process.

4) Interpretation: From the analysis of the projection matrix $\mathbf{P}$, we can identify the GFT modes which are the most weighted in this combination. Thus, we are led to an interesting interpretation of the results. Indeed, by computing the variance of the projected GFT coefficients, we measure the variability over time of the presence of some graph Fourier modes in the fMRI signals. These modes may be seen as *frequency signatures* of the NT/ASD conditions. It appears that most of the signatures related to the NT subjects correspond to a predominant activity in the parieto-occipital regions. The significant modes of the ASD population (e.g., see Fig. 3) are all related to high levels of activity in the fronto-temporal areas in ASD subject. This result is consistent with previous findings reported in the literature of neuroscience (see Sec. IV-A).

V. Conclusion

In this work, we extend the Graph Signal Processing (GSP) framework by introducing a new algorithm which classifies time-varying graph signals, with application to the identification of Autism Spectrum Disorder (ASD). The algorithm aims at exploiting the structure (brain topology) - function (fMRI-based activity) interplay of the brain. To this end, we use the Graph Fourier Transform (GFT) which generalizes the notion of frequency analysis to irregular domains. The resulting graph Fourier features are processed through an extension of the Fukunaga-Koontz transform to build discriminative markers for the classification of ASD and neurotypical patients. Our results are promising in terms of accuracy and interpretation.

VI. Acknowledgements

Sarah Itani is a research fellow of the Fund for Scientific Research (F.R.S.-FNRS), Brussels, Belgium.

References

[1] JM Fellous, G Sapiro, A Rossi, HS Mayberg, and M Ferrante, "Explainable artificial intelligence for neuroscience: Behavioral neurostimulation," *Frontiers in Neuroscience*, vol. 13, pp. 1346, 2019.

[2] O Sporns, "Structure and function of complex brain networks," *Dialogues in clinical neuroscience*, vol. 15, no. 3, pp. 247, 2013.

[3] LQ Uddin, K Supekar, and V Menon, "Reconceptualizing functional brain connectivity in autism from a developmental perspective," *Frontiers in human neuroscience*, vol. 7, pp. 458, 2013.

[4] P Rane, D Cochran, SM Hodge, et al., "Connectivity in autism: a review of MRI connectivity studies," *Harvard review of psychiatry*, vol. 23, no. 4, pp. 223, 2015.

[5] RK Kana, LQ Uddin, T Kenet, et al., "Brain connectivity in autism," *Frontiers in Human Neuroscience*, vol. 8, pp. 349, 2014.

[6] A Ortega, P Frossard, J Kovačević, et al., "Graph signal processing: Overview, challenges, and applications," *Proceedings of the IEEE*, vol. 106, no. 5, pp. 808–828, 2018.

[7] JD Medaglia, W Huang, EA Karuza, et al., "Functional alignment with anatomical networks is associated with cognitive flexibility," *Nature Human Behaviour*, vol. 2, no. 2, pp. 156, 2018.

[8] W Huang, L Goldsberry, NF Wymbs, et al., "Graph frequency analysis of brain signals.," *J. Sel. Topics Signal Processing*, vol. 10, no. 7, pp. 1189–1203, 2016.

[9] M Ménoret, N Farrugia, B Pasdeloup, and V Gripon, "Evaluating graph signal processing for neuroimaging through classification and dimensionality reduction," in *2017 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*. IEEE, 2017, pp. 618–622.

[10] SI Ktena, S Parisot, E Ferrante, et al., "Metric learning with spectral graph convolutions on brain connectivity networks," *NeuroImage*, vol. 169, pp. 431–442, 2018.

[11] V Subbaraju, MB Suresh, S Sundaram, and S Narasimhan, "Identifying differences in brain activities and an accurate detection of autism spectrum disorder using resting state functional-magnetic resonance imaging: A spatial filtering approach," *Medical image analysis*, vol. 35, pp. 375–389, 2017.

[12] AF Alexander-Bloch, PE Vertes, R Stidd, et al., "The anatomical distance of functional connections predicts brain network topology in health and schizophrenia," *Cerebral cortex*, vol. 23, no. 1, pp. 127–138, 2012.

[13] E Bullmore and O Sporns, "Complex brain networks: Graph theoretical analysis of structural and functional systems," *Nature Reviews Neuroscience*, vol. 10, no. 3, pp. 186–198, 2009.

[14] M Xia, J Wang, and Y He, "BrainNet Viewer: A network visualization tool for human brain connectomics," *PloS ONE*, vol. 8, no. 7, pp. e68910, 2013.

[15] DI Shuman, SK Narang, P Frossard, et al., "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *IEEE Signal Processing Magazine*, vol. 30, no. 3, pp. 83–98, 2013.

[16] S Itani and D Thanou, "Combining anatomical and functional networks for neuropathology identification: A case study on autism spectrum disorder," *arXiv preprint arXiv:1904.11296*, 2019.

[17] RW Newcomb, "On the simultaneous diagonalization of two semi-definite matrices," *Quarterly of Applied Mathematics*, vol. 19, no. 2, pp. 144–146, 1961.

[18] A Di Martino, C-G Yan, Q Li, et al., "The autism brain imaging data exchange: Towards large-scale evaluation of the intrinsic brain architecture in autism," *Molecular Psychiatry*, vol. 19, no. 6, pp. 659, 2014.

[19] Preprocessed Connectomes Project, "ABIDE Preprocessed," https://bit.ly/2Z1xKk4, 2014, [Online; accessed 30-05-2020].

[20] N Tzourio-Mazoyer, B Landeau, D Papathanassiou, et al., "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.

[21] Wilkin Chau and Anthony R McIntosh, "The talairach coordinate of a point in the mni space: how to interpret it," *Neuroimage*, vol. 25, no. 2, pp. 408–416, 2005.

[22] K Hirata, K Egashira, K Harada, et al., "Differences in frontotemporal dysfunction during social and non-social cognition tasks between patients with autism spectrum disorder and schizophrenia," *Scientific reports*, vol. 8, no. 1, pp. 3014, 2018.

[23] M-A Lauvin, J Martineau, C Destrieux, et al., "Functional morphological imaging of autism spectrum disorders: Current position and theories proposed," *Diagnostic and interventional imaging*, vol. 93, no. 3, pp. 139–147, 2012.

[24] L Poustka, C Jennen-Steinmetz, R Henze, et al., "Fronto-temporal disconnectivity and symptom severity in children with autism spectrum disorder," *The World Journal of Biological Psychiatry*, vol. 13, no. 4, pp. 269–280, 2012.

[25] JR Quinlan, *C4.5: Programs for machine learning*, San Mateo: Morgan Kaufmann, 1993.