

Community-Aware Graph Signal Processing

Modularity defines new ways of processing graph signals



©ISTOCKPHOTO.COM/ALISEFOX

The emerging field of graph signal processing (GSP) allows one to transpose classical signal processing operations (e.g., filtering) to signals on graphs. The GSP framework is generally built upon the graph Laplacian, which plays a crucial role in studying graph properties and measuring graph signal smoothness. Here, instead, we propose the graph modularity matrix as the centerpiece of GSP to incorporate knowledge about graph community structure when processing signals on the graph but without the need for community detection. We study this approach in several generic settings, such as filtering, optimal sampling and reconstruction, surrogate data generation, and denoising. Feasibility is illustrated by a small-scale example and a transportation network data set as well as one application in human neuroimaging where community-aware GSP reveals relationships between behavior and brain features that are not shown by Laplacian-based GSP. This work demonstrates how concepts from network science can lead to new, meaningful operations on graph signals.

Introduction

Network science is a multidisciplinary field that accounts for complex structures of data, providing new interpretations of datasets in diverse scientific disciplines, ranging from the humanities to physics and biomedicine. Naturally, the analysis of network data relies on methods from graph theory but also from statistical mechanics, statistical inference, advanced visualization, and domain knowledge from applied fields. More recently, GSP emerged as a new research theme at the intersection between signal processing and graph theory, with a particular focus on processing graph signals that associate values to the nodes of the graph. In many cases, the graph Fourier transform (GFT) was defined by the eigendecomposition of the graph Laplacian; i.e., the eigenvectors of the Laplacian are considered graph Fourier basis vectors, and the associated eigenvalues are graph frequencies [1]. Such a GFT can then generalize various classical signal processing tools to graphs [1], [2], such as the wavelet transform [3], as well as theoretical considerations about graph uncertainty principles [4].

The graph Laplacian defines the second-order derivative on the graph and, therefore, is linked to smoothness, but alternative operators can explore other properties of graphs and graph signals. For example, community structure is a particularly interesting concept from network science where nodes inside a community are more strongly interconnected than with the rest of the graph [5], [6]. Community structure turned out to be present and relevant for a broad range of applications in sociology [7], transportation [8], biology [9], and neuroscience [10]. In practice, communities can be found by maximizing the modularity index that evaluates the density of connections within clusters against a degree-matched graph where no cluster preference exists [11]. In Laplacian-based spectral clustering, the Laplacian eigenvectors with the smallest nonzero eigenvalues are considered because they optimize the convex relaxation of the graph cut criterion [12]. Similar to this, one approach for community detection is to compute the eigendecomposition of the modularity operator and then consider the eigenvectors with the largest eigenvalues [5].

In this article, we set the foundations for community-aware GSP by introducing the modularity operator at the heart of the framework. This allows one to define GSP operations that are aware of the graph community structure but without the need of explicit community detection.

GSP

We consider an undirected graph, $\mathcal{G} = (\mathcal{N}, \mathcal{E})$, with node set \mathcal{N} of cardinality N and edge set \mathcal{E} . \mathcal{G} can also be represented by the $N \times N$ weighted adjacency matrix \mathbf{A} , whose entry $a_{i,j}$ is nonzero and indicates the edge weight for an edge $(i, j) \in \mathcal{E}$ that runs from node i to node j . For an undirected graph, \mathbf{A} is symmetric; i.e., it holds that $a_{i,j} = a_{j,i}$ and $\mathbf{A} = \mathbf{A}^\top$. We will refer to a subgraph, $\mathcal{G}_S = (\mathcal{N}_S, \mathcal{E}_S)$, by its node set $\mathcal{N}_S \subset \mathcal{N}$ and assume \mathcal{E}_S containing all edges (i, j) between nodes in \mathcal{N}_S . A graph signal associated with \mathcal{G} is a vector, $\mathbf{x} \in \mathbb{R}^N$, that attributes values x_i to the nodes $i = 1, 2, \dots, N$. The neighborhood of node i is defined as the set of nodes \mathcal{N}_i connected to it. A graph shift operator is defined as a linear operator on the space of signals, such that each entry of the shifted graph signal is a linear combination of input signal values, which often only involves neighboring entries to the one at hand [2]. Therefore, the shift operator can be represented by a symmetric matrix, $\mathbf{S} \in \mathbb{R}^{N \times N}$, that associates weights $s_{i,j}$ to edges (i, j) such that $\mathbf{x}_{\text{shift}} = \mathbf{S}\mathbf{x}$. We will consider graph operators \mathcal{H} that are shift-invariant under \mathbf{S} and, thus, satisfy $\mathcal{H}\mathbf{S}\mathbf{x} = \mathbf{S}\mathcal{H}\mathbf{x}$ and can be represented as a matrix polynomial of \mathbf{S} [13], i.e., $\mathcal{H} = p(\mathbf{S}) = \sum_{k=0}^K h_k \mathbf{S}^k$, with maximum degree of $N - 1$ due to the Cayley–Hamilton theorem.

The eigendecomposition of the shift operator provides the factorization

$$\mathbf{S} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^\top, \quad (1)$$

where $\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_N]$ contains the N eigenvectors and $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_N)$ is a diagonal matrix with the corresponding eigenvalues. This allows us to write the graph operator \mathcal{H} alternatively as $\mathcal{H} = p(\mathbf{S}) = \mathbf{U}\text{diag}(\tilde{\mathbf{h}})\mathbf{U}^\top = \mathbf{U}\tilde{\mathbf{H}}\mathbf{U}^\top$, where the entries $\tilde{h}_i = p(\lambda_i) = \sum_{k=0}^K h_k \lambda_i^k$ of $\tilde{\mathbf{H}}$ yield the spectral characterization of the graph operator. For the perspective of

GSP, a given shift operator \mathbf{S} defines the GFT of the graph signal \mathbf{x} as (1):

$$\hat{\mathbf{x}} = \mathbf{U}^\top \mathbf{x}, \quad \text{and} \quad \mathbf{x} = \mathbf{U}\hat{\mathbf{x}}, \quad (2)$$

where \mathbf{U} is defined as in (1) and $\hat{\mathbf{x}}$ contains the spectral coefficients of the GFT. The graph operator \mathcal{H} can then be implemented elegantly in the graph Fourier domain as

$$\mathbf{x}_{\text{out}} = \mathcal{H}\mathbf{x} = p(\mathbf{S})\mathbf{x} = \mathbf{U}p(\mathbf{\Lambda})\mathbf{U}^\top \mathbf{x} = \mathbf{U}\tilde{\mathbf{H}}\hat{\mathbf{x}}, \quad (3)$$

which allows one to directly specify $\tilde{\mathbf{H}}$ in terms of a spectral window (e.g., low-pass, bandpass, and high-pass) for graph filtering operations [13]. Beyond filtering, other operations have been extended to the graph domain, such as stationarity analysis [14], wavelet transforms [3], and convolutional neural networks [15].

One common choice for \mathbf{S} is the weighted graph adjacency matrix \mathbf{A} [2], [16]. Another one is the Laplacian matrix $\mathbf{L} = \mathbf{D} - \mathbf{A}$, where $\mathbf{D} = \text{diag}(k_1, k_2, \dots, k_N)$ is the degree matrix with $k_i = \sum_{j=1}^N a_{i,j}$, the weighted degree [3], [17]. For the latter, the eigenvalues are sometimes referred to as *graph frequencies* and reflect smoothness in terms of the signal variation norm of the corresponding eigenvectors [13]. For a graph signal \mathbf{x} , its smoothness is measured by the quadratic form

$$q_L(\mathbf{x}) = \sum_{i \neq j} a_{i,j} (x_i - x_j)^2 = \mathbf{x}^\top \mathbf{L}\mathbf{x} = \sum_{i=1}^N \lambda_i \hat{x}_i^2, \quad (4)$$

which shows that measuring smoothness in the spectral domain can be done by weighting with the graph frequencies. The example in Figure 1 illustrates the Laplacian eigendecomposition for a simple graph and will be discussed in more detail later.

Community structure

Communities refer to dense subgraphs $\mathcal{P}_c \subset \mathcal{N}$, $c = 1, \dots, C$, which are well separated from each other, and manifested at the “mesoscale” level between local nodal and global graph properties [5]. A large number of measures have been proposed with the purpose of discovering the community structure of an observed graph. For our aim, it is insightful to first revisit the graph Laplacian as it relates to one aspect of community structure, which is quantifying the separation between subgraphs. Specifically, the splitting of a graph into two mutually exclusive subgraphs \mathcal{P}_1 and \mathcal{P}_2 can be encoded by a vector, \mathbf{s} , whose entry $s_i = +1$ or -1 indicates whether a node, i , belongs to the first or second subgraph, respectively. The graph cut size—the number of connections running between the two subgraphs—can then be related to the Laplacian as

$$R = \frac{1}{2} \sum_{\substack{i,j \\ s_i \neq s_j}} a_{i,j} = \frac{1}{2} \sum_{i,j} \left(\frac{1 - s_i s_j}{2} \right) a_{i,j} = \frac{1}{4} \mathbf{s}^\top \mathbf{L}\mathbf{s}. \quad (5)$$

Optimizing R by convex relaxation of \mathbf{s} (i.e., allowing the entries to take any value) leads to the well-known spectral clustering [12]. The eigenvector of \mathbf{L} with the smallest nonzero eigenvalue (also known as the Fiedler vector) provides the

solution to the bipartition problem. Recent work has also used graph wavelets to enable multiscale subgraph discovery [18].

The network-science view on community structure considers the adjacency matrix as a realization of an underlying stochastic model that defines edge probabilities within and between subgraphs. Stochastic block models (SBMs) [19] are the best-known generative models that can express assortativity (preferential connectivity within a node’s subgraph, leading to community structure) but also disassortativity (preferential connectivity to a subgraph to which the node does not belong) and core–periphery structure (densely interconnected core and periphery to the core). SBMs can be fitted by statistical inference to an observed graph or can generate random graphs with predefined structure. *Modularity*, denoted by \mathbf{Q} , is a specific graph measure derived from stochastic considerations that quantifies the density of subgraphs by comparison to a null model:

$$Q = \frac{1}{2} \sum_{s_i=s_j} (a_{i,j} - z_{i,j}) = \frac{1}{2} \sum_{i,j} \left(\frac{1 + s_i s_j}{2} \right) (a_{i,j} - z_{i,j})$$

$$= \frac{1}{4} \mathbf{s}^\top \underbrace{\left(\mathbf{A} - \frac{\mathbf{k}\mathbf{k}^\top}{2M} \right)}_{\mathbf{Q}} \mathbf{s}, \quad (6)$$

where \mathbf{s} is a vector encoding the graph partition into two communities, $z_{i,j} = k_i k_j / 2M$ is the edge probability between nodes i and j according to the null model, $M = \sum_{i=1}^N k_i / 2$ is the total edge weight, and \mathbf{Q} is the modularity matrix. Choosing this

null model allows for comparisons against a reference that preserves the graph degree distribution (i.e., $\sum_{j=1}^N a_{i,j} = \sum_{j=1}^N z_{i,j}$ for $i = 1, 2, \dots, N$), with edges placed evenly [11]. Hence, \mathbf{Q} encodes the difference between edge densities in the original graph and in a degree-matched null model. This model is known as the configuration model and is commonly used to define \mathbf{Q} , but other null models can be considered [11], [20], [21].

The solution to maximizing Q is found by spectral clustering using the eigenvectors of \mathbf{Q} with the largest eigenvalues, identifying “modules” with high assortativity. Similarly, “antimodules” with high disassortativity can be found by minimizing modularity. In fact, \mathbf{Q} is a rank-one perturbation of \mathbf{A} , and, consequently, Weyl’s inequality informs us that eigenvalues $\lambda_i^{(\mathbf{Q})}$ and $\lambda_i^{(\mathbf{A})}$ of \mathbf{Q} and \mathbf{A} are interleaved, i.e., $\lambda_1^{(\mathbf{A})} \geq \lambda_1^{(\mathbf{Q})} \geq \lambda_2^{(\mathbf{A})} \geq \lambda_2^{(\mathbf{Q})} \geq \dots \geq \lambda_N^{(\mathbf{A})} \geq \lambda_N^{(\mathbf{Q})}$, where the sequences of eigenvalues are in descending order. This result confirms that the modularity matrix of any simple undirected graph (that is, without self-loops) has both positive and negative eigenvalues [22]. The existence of both positive and negative eigenvalues implies that such a graph can be analyzed in terms of modular and antimodular spectral components of \mathbf{Q} . Eigenvectors with zero eigenvalues are modularity neutral; e.g., the constant vector $\mathbf{1}$ is in the kernel of \mathbf{Q} due to $\mathbf{Q} \cdot \mathbf{1} = \mathbf{0}$.

Further illustration of the differences between spectral properties of \mathbf{L} and \mathbf{Q} is provided in Figure 1 for a toy graph

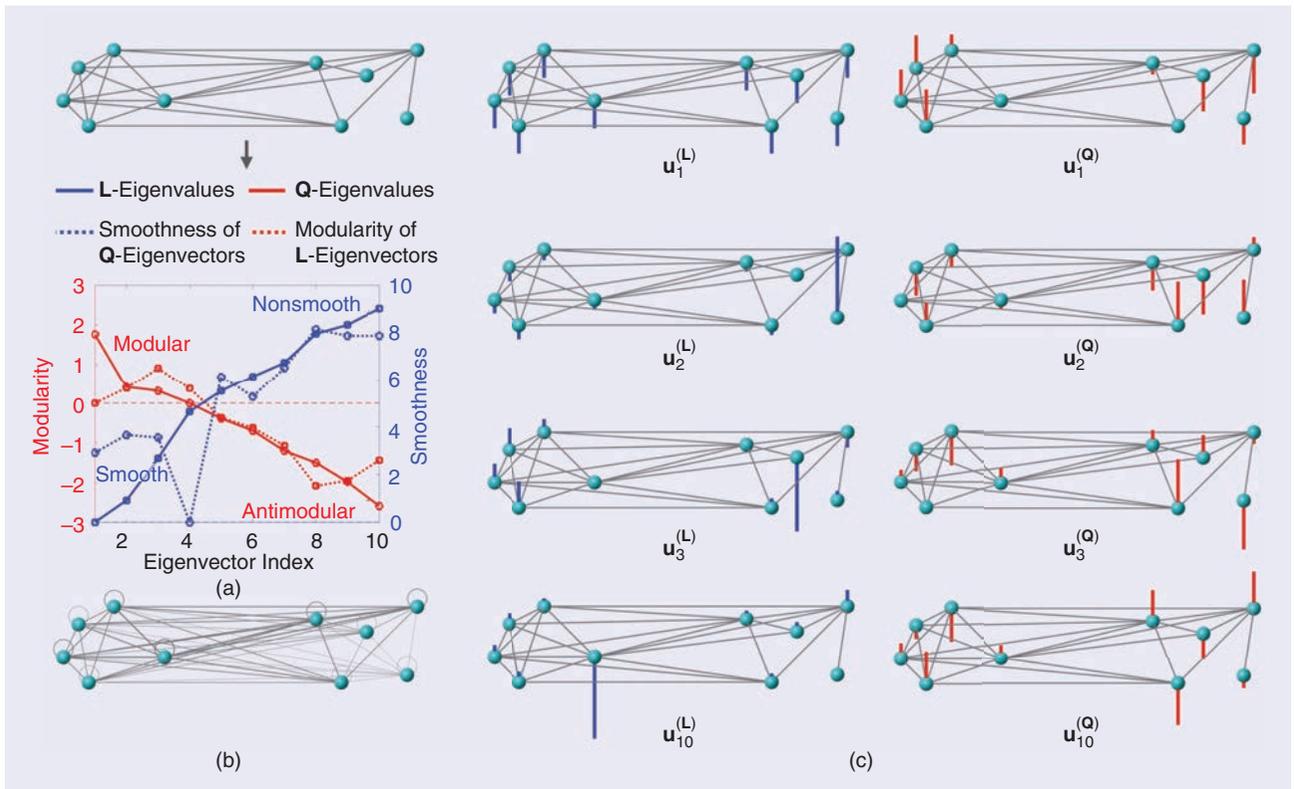


FIGURE 1. The eigenvalues and eigenvectors of the graph Laplacian (\mathbf{L}) and modularity (\mathbf{Q}) matrices. (a) Underlying graph structure (top), corresponding \mathbf{L} (blue) and \mathbf{Q} (red) eigenspectra (solid lines), and quadratic forms of smoothness and modularity (dotted lines). (b) Corresponding degree-matched null model used to compute \mathbf{Q} (6). (c) Selected eigenvectors of \mathbf{L} (blue) and \mathbf{Q} (red) matrices. The value and sign of eigenvectors’ entries are reflected by the height and up-down direction, respectively, of the vertical bars.

with $N = 10$ nodes, 5 of which form a strong (fully connected) community and the others are weakly connected. Figure 1(a) shows the graph and a plot of eigenvalues $\lambda_i^{(L)}$ (blue solid line) and $\lambda_i^{(Q)}$ (red solid line) in the conventional ascending and descending order, respectively. To better highlight the differences between the corresponding eigenvectors $\mathbf{u}_i^{(L)}$ and $\mathbf{u}_i^{(Q)}$, respectively, we also plot smoothness $\mathbf{u}_i^{(Q)\top} \mathbf{L} \mathbf{u}_i^{(Q)}$ of modularity eigenvectors (blue dotted line) and modularity $\mathbf{u}_i^{(L)\top} \mathbf{Q} \mathbf{u}_i^{(L)}$ of Laplacian eigenvectors (red dotted line) (the quadratic form associated with the modularity matrix will be formally introduced in the ‘‘Community-aware GSP’’ section). Figure 1(b) shows the degree-matched null model graph $\mathbf{k}\mathbf{k}^\top/2M$. Several eigenvectors $\mathbf{u}_i^{(Q)}$ and $\mathbf{u}_i^{(L)}$ ($i = 1, 2, 3, 10$) are presented in Figure 1(c). Although $\mathbf{u}_i^{(L)}$ values are optimized for smoothness, they show high values at specific nodes [except the constant eigenvector $\mathbf{u}_1^{(L)}$, with zero eigenvalue]. Therefore, the Fiedler vector, $\mathbf{u}_2^{(L)}$, with lowest graph cut size, does not provide a correct partitioning. The eigenvectors of \mathbf{Q} , though, are optimized for the modularity index, and $\mathbf{u}_1^{(Q)}$ provides a conspicuous split between the communities. \mathbf{Q} also has a constant eigenvector, $\mathbf{u}_4^{(Q)}$, with zero smoothness. Curiously, the modularity of $\mathbf{u}_3^{(L)}$ is actually the highest among the Laplacian eigenvectors but still does not provide a convincing partitioning. Eigenvectors of \mathbf{Q} with negative eigenvalues, such as $\mathbf{u}_{10}^{(Q)}$, are driven by smoothness across modules and signal variability within modules.

Community-aware GSP

The Laplacian operator \mathbf{L} is the common choice of shift operator in GSP [1], [15], [16] from which the GFT and all operations are derived. Instead, we propose to use the modularity matrix \mathbf{Q} as the shift operator. Interestingly, the modularity matrix is a nonlocal operator, since the second term that originates from the null model ‘‘spreads out’’ the signal over the whole graph according to the degree distribution—and not only the local neighborhood. Based on this generalized notion of shift operator, we will obtain GSP operations that become aware of the graph community structure but without the need for explicit community detection. Given a graph signal \mathbf{x} , its modularity is computed by the quadratic form

$$q_Q(\mathbf{x}) = \sum_{i,j} a_{i,j} x_i x_j - \frac{\left(\sum_{i,j} a_{i,j} x_i\right)^2}{\sum_{i,j} a_{i,j}} = \mathbf{x}^\top \mathbf{Q} \mathbf{x}. \quad (7)$$

Since \mathbf{Q} is not positive semidefinite, $q_Q(\mathbf{x})$ can take positive and negative values, depending on whether signal variations follow modular or antimodular organization [11]. Thus, the quadratic form $q_Q(\mathbf{x})$ is not a variation norm of the graph signal \mathbf{x} , which is needed for some GSP operations, such as regularization. We overcome this limitation by introducing

$$q_{Q^+}(\mathbf{x}) = \mathbf{x}^\top \mathbf{Q}^+ \mathbf{x}, \quad (8)$$

based on $\mathbf{Q}^+ = \lambda_{\max}^{(Q)} \cdot \mathbf{I} - \mathbf{Q}$, where $\lambda_{\max}^{(Q)}$ is the largest eigenvalue of \mathbf{Q} and \mathbf{I} is the identity matrix. Since \mathbf{Q}^+ is positive semidefinite, $q_{Q^+}(\mathbf{x})$ is a nonnegative function of \mathbf{x} . A low value of $q_{Q^+}(\mathbf{x})$ reflects that the graph signal \mathbf{x} follows the

modular organization of the graph. On the contrary, a high value of $q_{Q^+}(\mathbf{x})$ is obtained for graph signals reflecting the antimodular organization. In other words, $q_{Q^+}(\mathbf{x})$ can be interpreted as the modularity-based graph signal variation of \mathbf{x} . Minimization of this metric is achieved by the eigenvectors of \mathbf{Q}^+ that define the spectral basis of a GFT exploiting the modularity of graph signals. Since \mathbf{Q}^+ and \mathbf{Q} have the same eigenvectors with eigenvalues that are reversed and shifted, the eigenvectors of \mathbf{Q} define a proper GFT basis that is built up according to modularity/antimodularity. Similarly, denoting $\mathbf{Q}^- = \mathbf{Q} - \lambda_{\min}^{(Q)} \cdot \mathbf{I}$ allows one to define a variation norm, $q_{Q^-}(\mathbf{x})$, that encodes antimodular organization in low values while exploiting the same spectral basis, since \mathbf{Q}^- also has the same eigenvectors as \mathbf{Q} .

We now illustrate the utility of community-aware GSP tools using data from the OpenFlights Airports Database (<https://openflights.org/data.html>), which consists of 3,281 airports and 67,202 routes [Figure 2(a)]. Graph nodes denote airports that are connected by an undirected binary edge if there exists an airline route between them. Node colors reflect a graph signal, computed as the sum of both departing and incoming flights at each airport, which was then demeaned and scaled to unit variance. We considered a ground truth community structure based on the continent to which each airport belongs, resulting in a partition of the nodes into the following six communities: Europe, Africa, Asia, Oceania, North America, and South America [8]. The inset of Figure 2(a) depicts the total number of flights leaving from or arriving at the airports in the eastern part of the North American continent. It can be seen that the Atlanta airport (ATL) has more traffic than John F. Kennedy Airport (JFK) in New York and that the vast majority of airports have very low traffic.

Filtering

From the general definition of GSP filtering proposed in (3), community-aware filtering uses the modularity-based spectral domain with a spectral window $\tilde{\mathbf{h}}$:

$$\mathbf{x}_{\text{out}} = \underbrace{\mathbf{U} \text{diag}(\tilde{\mathbf{h}})}_{=\tilde{\mathbf{h}}} \mathbf{U}^\top \mathbf{x}, \quad (9)$$

where \mathbf{U} contains the eigenvectors of \mathbf{Q} . While low- and high-pass filtering are natural operations when using the Laplacian GFT, modularity-based GFT allows one to define a modular filter (i.e., $\tilde{\mathbf{h}}$ has nonzero weights on spectral components with positive eigenvalues) or an antimodular filter (i.e., $\tilde{\mathbf{h}}$ only has nonzero weights on spectral components with negative eigenvalues).

The community-aware filtering was applied to the graph signal of Figure 2(a) and was compared to a Laplacian-based filtering. The passband, that is, the range of eigenvalue indices with nonzero filter weights, of the modular (antimodular) filter includes all 1,125 (1,159) strictly positive (negative) eigenvalues of \mathbf{Q} , and the smooth (nonsmooth) filter was matched so as to capture the same number of spectral Laplacian components [Figure 2(b)]. Within a passband $[N_1, N_2]$, the i th entry of $\tilde{\mathbf{h}}$ was set to $|\lambda_i|/\sum_{k=N_1}^{N_2} |\lambda_k|$ for modular, antimodular, and nonsmooth filterings and to $1 - |\lambda_i|/\sum_{k=N_1}^{N_2} |\lambda_k|$ for the smooth filtering, which accounts for the strength of modularity or the

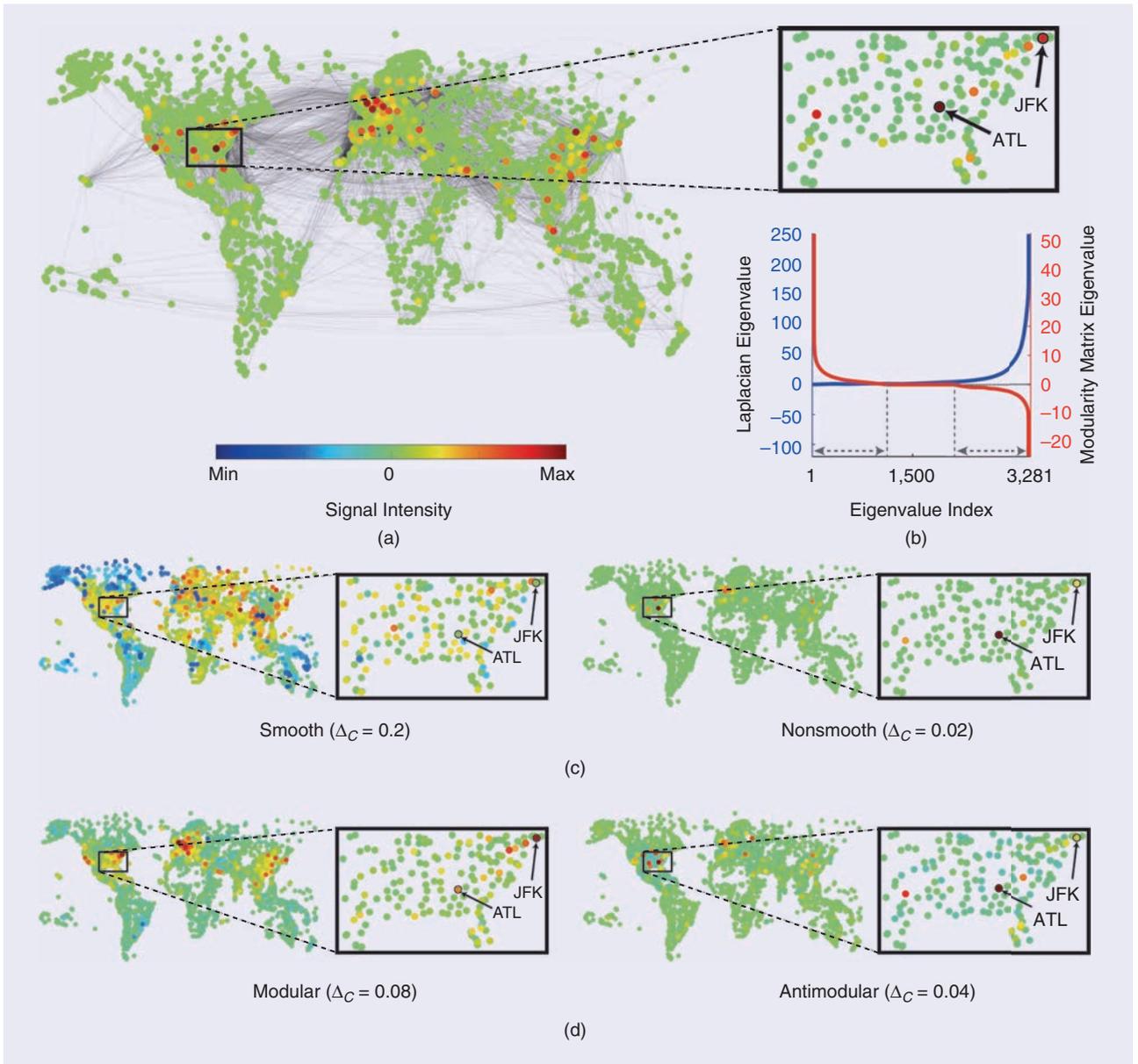


FIGURE 2. The application of the GSP framework on OpenFlights. (a) Graph nodes correspond to airports, and an edge connects two nodes when at least one flight connects the two corresponding airports. Graph signal (number of flights at each airport) is reflected in nodes' colors. The inset shows a zoom on JFK and ATL airports. (b) Eigenvalues of the graph Laplacian (blue) and modularity matrix (red). Dashed gray lines and arrows represent limits of the filtering passbands, left for modular and smooth, right for antimodular and nonsmooth. (c) Laplacian filtering of the graph signal shown in panel (a) yields smooth and nonsmooth signals, and (d) community-aware filtering yields modular and antimodular signals. The value of within-community variability (Δ_c) is shown for the four filtered signals.

smoothness of the components. Finally, for each filtered signal, we computed a measure of within-community variability, denoted Δ_c , defined as the standard deviation of the filtered signal values within a ground truth community, averaged over the six communities.

As shown in Figure 2(c), the Laplacian-based filtering extracts smooth and nonsmooth parts of the graph signal. The smooth signal tends to capture widespread fluctuations over the graph, whereas the nonsmooth signal contains rather localized peaks that partially correspond to the extreme values in the original signal (e.g., ATL and JFK airports), suggesting

that the underlying community structure is not a predominant feature encoded in Laplacian-filtered signals. In contrast, the modular signal detailed in Figure 2(d) reflects the community structure of the underlying graph by the clusters of high values in North America, Europe, and Asia. This effect is further supported by within-community variability (Δ_c) that is lower in the modular signal than in the smooth one. In other words, modular-based filtering can be seen as promoting smoothness within communities. On the contrary, the antimodular signal promotes variability within communities, as this signal shows higher Δ_c than the nonsmooth signal.

To further explore the roles of particular nodes in the different filtering operations, we focus on two airports: ATL and JFK. While these two airports are both highly connected, as seen from Figure 2(a), they play different roles in the graph community structure. Indeed, ATL has a within-community z -score degree (Z_{in}) [9] of 8.98 and an outside-community z -score degree (Z_{out}) of 5.97, while for JFK, we have $Z_{\text{in}} = 4.22$ and $Z_{\text{out}} = 15.29$. Therefore, ATL has stronger connections within its community than between communities, and vice versa for JFK. For Laplacian filtering, the signal values of both JFK and ATL are evened out in the smooth signal and stand out in the nonsmooth signal [insets of Figure 2(c)]. However, community-aware filtering picks up differences between these airports by a relatively stronger value of JFK in the modular signal and of ATL in the antimodular signal [insets of Figure 2(d)]. Since modularity-promoting filtering favors smoothing within the communities, the value of the strongly within-community connected ATL will be more reduced than that for JFK. The large value of ATL is captured in the antimodular signal, as it stands out with respect to values of its within-community neighbors. This suggests that the modular/antimodular signal identifies nodes with high values and strong intermodular/intramodular connectivity. Overall, the results reveal that community-aware filtering can attenuate or enhance values of nodes according to their connectivity within or between communities.

Finally, the need for the eigendecomposition of \mathbf{Q} in (9) can be circumvented by implementing the filtering operation in the vertex domain by a polynomial matrix function, $p(\mathbf{Q})$, as suggested in (3), which is equivalent to applying the spectral window $\tilde{\mathbf{H}} = p(\mathbf{\Lambda})$. To further improve computational efficiency of filtering for large-scale but sparse graphs, one can break down the operation $\mathbf{Q}\mathbf{x}$ into $\mathbf{Q}\mathbf{x} = \mathbf{A}\mathbf{x} - (1/2M)\mathbf{k}\mathbf{k}^T\mathbf{x}$, where the first term is a sparse matrix–vector multiplication and the second term can be evaluated by consecutively computing $\mathbf{k}^T\mathbf{x}$ and then multiplying the resulting scalar with $\mathbf{k}/2M$. Therefore, the dense matrix \mathbf{Q} never needs to be stored explicitly. For an undirected graph with M' edges and N nodes, computing $\mathbf{A}\mathbf{x}$ takes $\mathcal{O}(2M')$ time and $\mathbf{L}\mathbf{x}$ takes $\mathcal{O}(2M' + N)$, (in big O notation). The term $(1/2M)\mathbf{k}\mathbf{k}^T\mathbf{x}$ has complexity of $\mathcal{O}(N)$. Consequently, $\mathbf{Q}\mathbf{x}$ takes $\mathcal{O}(2M' + N)$, identical to $\mathbf{L}\mathbf{x}$. For a polynomial filter of order K applied to a large sparse graph, this reverts to $\mathcal{O}(KM')$.

Optimal sampling and reconstruction

Finding the subset of nodes from which a signal can be optimally reconstructed has been extended to the graph domain in the context of band-limited graph signals $\mathbf{x} = \mathbf{B}\mathbf{x} = \mathbf{U}\mathbf{\Sigma}\mathbf{U}^T\mathbf{x}$ [4], where \mathbf{U} contains the eigenvectors of the shift operator and $\mathbf{\Sigma}$ is a diagonal matrix indicating the passband. The noisy graph signal $\mathbf{y} = \mathbf{x} + \mathbf{n}$, with \mathbf{n} additive independent and identically distributed noise, is sampled into $\mathbf{x}_s = \mathbf{R}\mathbf{y}$, where the diagonal matrix \mathbf{R} indicates with 0's and 1's the sampled nodes. Reconstruction denotes the procedure of finding \mathbf{x}_{rec} from \mathbf{x}_s , such that the mean squared error $\mathbb{E}[\|\mathbf{x}_{\text{rec}} - \mathbf{x}\|_2^2]$ is minimized [4]. The minimization condition further extends to the choice of optimal sampling procedure, since sampling at

specific nodes can limit the potential performance of the subsequent reconstruction. One of the solutions [4] to the problem of finding (sub)optimal sampling and reconstruction defines sampling as finding \mathbf{R}^* via

$$\mathbf{R}^* = \underset{\mathbf{R}}{\operatorname{argmax}} \|\mathbf{\Sigma}\mathbf{U}^T\mathbf{R}\|_F. \quad (10)$$

Solving (10) amounts to selecting nodes for the optimal sampling subset for which the columns of $\mathbf{\Sigma}\mathbf{U}^T$ have the highest l_2 -norm. Given the graph signal \mathbf{x}_s sampled at nodes defined in \mathbf{R}^* , the reconstruction follows:

$$\mathbf{x}_{\text{rec}} = \mathbf{V}\mathbf{\Psi}^{-1}\mathbf{V}^T\mathbf{x}_s, \quad (11)$$

where \mathbf{V} and $\mathbf{\Psi}$ contain the eigenvectors and eigenvalues of $\mathbf{B}\mathbf{R}^*\mathbf{B}^T$.

We explore how well the graph signal presented in Figure 2(a) can be reconstructed using this framework and considering either \mathbf{L} or \mathbf{Q} as the shift operator. We set to 500 the number of nodes to be sampled and use a spectral band including 200 components with lowest (\mathbf{L}) or highest (\mathbf{Q}) positive eigenvalues. The set of optimal nodes in these two cases is given in Figure 3.

The Laplacian-based sampling subset is composed of more peripheral nodes with low degree (1.78 ± 1.01), whereas the modularity-based sampling contains nodes with high degree (53.12 ± 43.61), which are important for inter- and intracommunity connectivity. Only two nodes were found to belong to both subsets. An interpretation for this is that while the Laplacian framework focuses on preserving values on nodes where that value is hard to predict due to their low connectivity, the modularity framework maximizes predictability of all nodes by selecting nodes with high degree. This is in accordance with the assumption that traffic at a well-connected airport could be a good predictor of the traffic at airports connected to it. Finally, the average reconstruction error is found to be significantly lower ($p < 0.01$, paired t test over nodes) for the modular-based framework than for the Laplacian-based framework (0.0001 ± 0.0002 versus 0.0006 ± 0.0037). This result supports the relevance of the modularity matrix as the shift operator rather than the Laplacian in applications where community structure is pertinent.

Surrogate data generation

Surrogates play an essential role in nonparametric statistical testing to provide data under the null hypothesis, i.e., randomizing measurements while also preserving some properties. For instance, phase randomization preserves the moduli of the Fourier coefficients while their phases are randomized, leading to surrogate data with the same autocorrelation properties as the original data. This framework was extended to graph signals using the Laplacian, yielding surrogate data that preserve the smoothness of the original graph signal [23]. We propose to transpose this method to community-aware representations to preserve modular organization of a given graph signal. In particular, a community-aware surrogate signal \mathbf{x}_{Surr} of the graph signal \mathbf{x} is given by

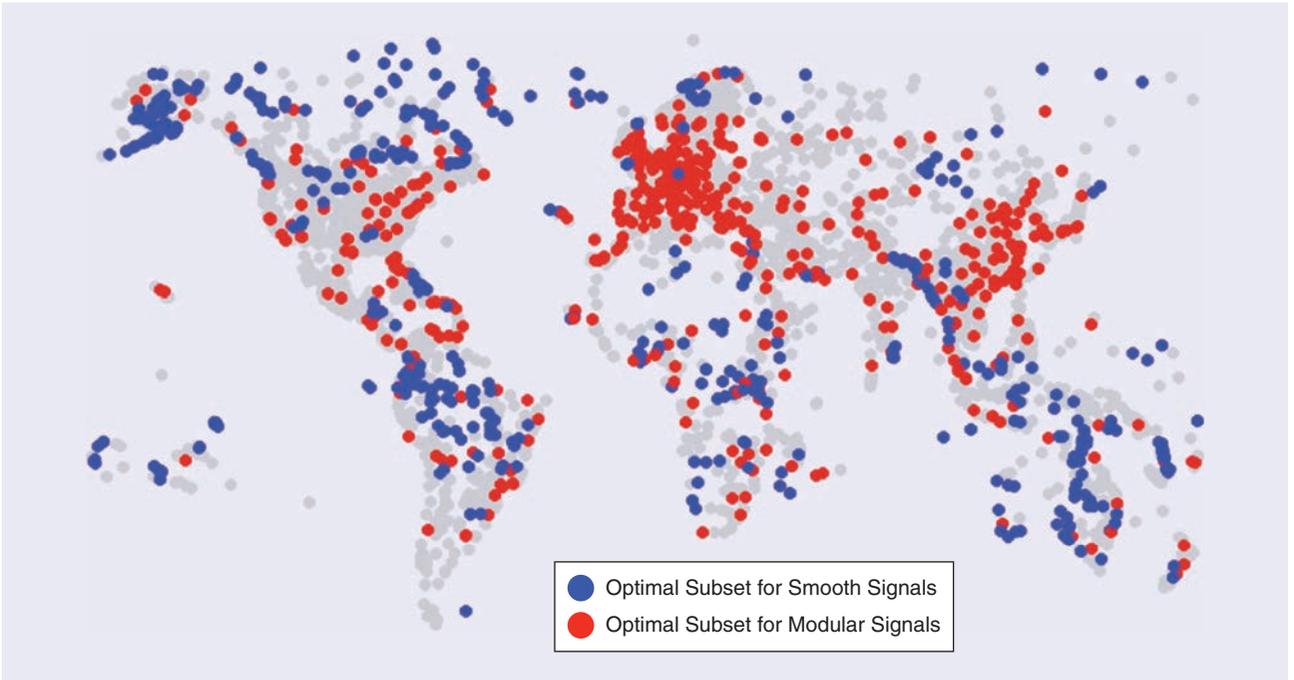


FIGURE 3. The optimal subset of 500 nodes for subsampling smooth (blue) or modular (red) signals of bandwidth 200.

$$\mathbf{x}_{\text{surr}} = \mathbf{UC}\hat{\mathbf{x}}, \quad (12)$$

where $\hat{\mathbf{x}}$ is the modularity-based GFT of \mathbf{x} and \mathbf{C} is a diagonal matrix with random entry 1 or -1 , thereby preserving the modularity index of the original signal. The null distribution of any test statistic can then be obtained from multiple realizations of \mathbf{x}_{surr} and compared against its value for the empirical signal \mathbf{x} . This could be refined to a more specific null model by only changing the signs of (anti)modular components, i.e., entries in \mathbf{C} corresponding to positive (negative) eigenvalues of \mathbf{Q} .

We applied this framework to the signal of Figure 2(a) by permuting signs of 1) all Laplacian-based, 2) all modularity-based, 3) only modular, and 4) only antimodular spectral coefficients. For each case, we generated 10,000 surrogate samples that were used to test whether the original signal value is higher than expected under the null hypothesis. The test used an α level of 0.05, Bonferroni corrected for multiple comparisons. In 1) and 2), no nodes were found with values significantly different from their surrogates, but when only randomizing modular (antimodular) components, 16 (2) airports revealed higher values than expected.

These airports had lower values of $Z_{\text{in}} (-0.35 \pm 0.16)$ than $Z_{\text{out}} (-0.18 \pm 0.03)$, indicating these nodes have stronger connectivity with other communities. Considering the results of the filtered signal values of JFK and ATL [Figure 2(d)], one could expect that high signal values at these nodes can be explained by the underlying community structure. However, surrogate testing showed they cannot be explained by community structure alone. Similarly, two “outlier” airports are identified when only randomizing signs of antimodular spectral coefficients. The results illustrate the complementary roles of modular and antimodular

parts to describe a graph signal. In the context of OpenFlights, this corroborates the assumption of relevant community structure being present in the graph that is only accounted for by modularity-based GSP and can then be used to assess to what extent graph signals follow this underlying graph structure.

Denoising

Another generic GSP operation is the recovery of the graph signal \mathbf{x} from its noisy observation $\mathbf{y} = \mathbf{x} + \mathbf{n}$. The variational formulation puts forward a data-fitting term and a regularization term:

$$\underset{\mathbf{x}}{\text{argmin}} \|\mathbf{x} - \mathbf{y}\|_2^2 + \mu \cdot \mathbf{x}^\top \mathbf{P} \mathbf{x}, \quad (13)$$

where μ is the regularization tuning parameter and the quadratic form of \mathbf{P} reflects prior knowledge about \mathbf{x} . A classical choice is $\mathbf{P} = \mathbf{L}$, which corresponds to assuming that the graph signal \mathbf{x} should be smooth on the graph. Since \mathbf{L} is positive semidefinite, the cost function in (13) is convex and has a unique optimal solution. The same is true if \mathbf{P} is chosen as \mathbf{Q}^+ and \mathbf{Q}^- [cf. (8)] to favor modular or antimodular organization of \mathbf{x} , respectively.

Performance of these reconstruction approaches is illustrated using the original signal of Figure 2(a). This signal was normalized to unit norm, corrupted with additive Gaussian noise of different variance, σ^2 , ranging between 0.01 and 1, and the optimal value of μ was determined using an oracle approach. For small to intermediate noise levels, we found that imposing a modular structure on \mathbf{x} (i.e., $\mathbf{P} = \mathbf{Q}^+$) yielded the best performance [root-mean-square (RMS) error is 0.0048 for $\sigma^2 = 0.01$ and 0.0083 for $\sigma^2 = 0.25$]. The error increases by an order of magnitude (0.0106 for $\sigma^2 = 0.01$ and 0.0168 for $\sigma^2 = 0.25$) when using a Laplacian regularizer (i.e., $\mathbf{P} = \mathbf{L}$),

whereas for antimodular regularization (i.e., $\mathbf{P} = \mathbf{Q}^-$) similar values of RMS are reached (0.0096 for $\sigma^2 = 0.01$ and 0.0145 for $\sigma^2 = 0.25$). The advantage of modular regularization decreases for larger noise, and the reconstruction errors become comparable when $\sigma^2 = 1$ (all errors above 0.15).

The assumptions of the different regularizers can be summarized as follows: Laplacian (\mathbf{L}) favors smoothness of the graph signal by measuring differences between adjacent nodes; modularity (\mathbf{Q}^+) favors smoothness of the graph signal between nodes weighted by their closeness community-wise (assortative mixing); and antimodularity (\mathbf{Q}^-) favors smoothness of the graph signal between nodes weighted inversely by this closeness (disassortative mixing). An explanation of the superior performance of modularity-based regularization in the present example is that similarly high air traffic is more bound to modular organization than to neighborhood relationships; e.g., low traffic of a small island airport connected to several high-traffic mainland hubs of different communities would lead to a high signal variation through the Laplacian lens but not through the modularity lens, since the island is not close to the large communities. In the end, it is the nature of the graph signal and how it relates to the underlying graph structure that will motivate the use of one or the other regularizer.

Validation for neuroimaging

The results presented in the previous section are built from a flight network with known community structure. Likewise, many real-world networks exhibit community structure; therefore, we expect the proposed framework to provide a more appropriate way to analyze the corresponding graph signals. We further illustrate the benefits of community-aware GSP in a validation experiment using brain anatomical and functional data from the Human Connectome Project [24] using a parcellation of the cerebral cortex ($N = 360$). The graph structure was defined by counting the number of fiber tracts in diffusion-weighted magnetic resonance imaging (MRI) [25], and the graph signals are the activity patterns at different time-points obtained from functional MRI (fMRI) time series reflecting activity in each brain region [24]. For each region, the timecourse was z -scored (centered and unit variance).

The experiment consisted of exploring the link between brain imaging data and 62 behavioral scores for 181 healthy volunteers. To that aim, functional time series were filtered using either the anatomical graph Laplacian or modularity matrix by following the procedure described in the “Filtering” section, and the filtered time series yielded, for each subject, a power metric reflecting (non)smoothness or (anti)modular structure of brain function in each brain region [26]. The link between these measures and the 62 behavioral scores was then computed using a nested cross-validation scheme, and the R^2 coefficient of determination was used to quantify the strength of the association between brain function and behavior.

Figure 4 shows the values of R^2 for the 62 behavioral measures and different bandwidths. It can first be seen that in most cases, using a narrow bandwidth yields stronger R^2 , which

suggests that the information of interest is captured in the very first (non)smooth or (anti)modular eigenvectors. We also observe that community-aware filtering reveals links between brain function and behavior that are not captured by Laplacian-based filterings. For example, the antimodular functional signal shows strong links with various personal character traits, while the modular signal mainly captures information about cognitive features.

Conclusions and outlook

Measures of community structure have been extensively used in network science to probe the organization of complex networks. Importantly, the tools that have been developed for processing graph signals expressed on these networks are based on the graph Laplacian and, thus, are blind to the underlying community structure. We proposed making GSP community-aware, not by detecting communities but by defining operations based on the modularity matrix. This provides a natural interpretation of the modularity spectrum in terms of modular and antimodular contributions, although it requires adaptation when a variation metric is needed. We showed, using several examples, that community-aware GSP acts meaningfully differently compared to classical GSP. Considering the variety of datasets with community structure, the proposed framework will find its use in a wide range of fields and applications.

One extension of Laplacian- and modularity-based GSP is to account for directed graphs. We believe this is beyond the scope of this article, but the interested reader is referred to [27] for defining Laplacian-based spectral bases of directed graphs or modularity matrices for directed graphs [28] using in- and out-degrees of nodes. Another extension of the community-aware GSP framework could include different null models in the modularity criterion. Specifically, the Bernoulli model preserves the average degree [11], whereas the configuration model can be modified to exclude self-loops [20] or to consider possible correlation between degrees of nodes [21]. Finally, communities could also be defined at the level of edges instead of nodes to deal with overlapping communities [29] or even triangles and higher-order simplicial complexes, as in higher-order Laplacian-based topological GSP [30].

Acknowledgments

This work was supported in part by the Swiss National Science Foundation (200021_17 5506), the CHIST-ERA Interactive and Visual Analysis of Networks project (20CH21_174081), a Japan Science and Technology Agency Exploratory Research for Advanced Technology grant (JPMJER1801), and the Center for Biomedical Imaging. Miljan Petrović and Raphaël Liégeois contributed equally to this article.

Authors

Miljan Petrović (miljan.petrovic@epfl.ch) received his M.Sc. degree in electrical engineering and computing from the University of Nis, Serbia. He joined the Medical Image Processing Lab (MIPLab), headed by Prof. Van De Ville, as a Ph.D. student in April 2018. His research focuses on bridging

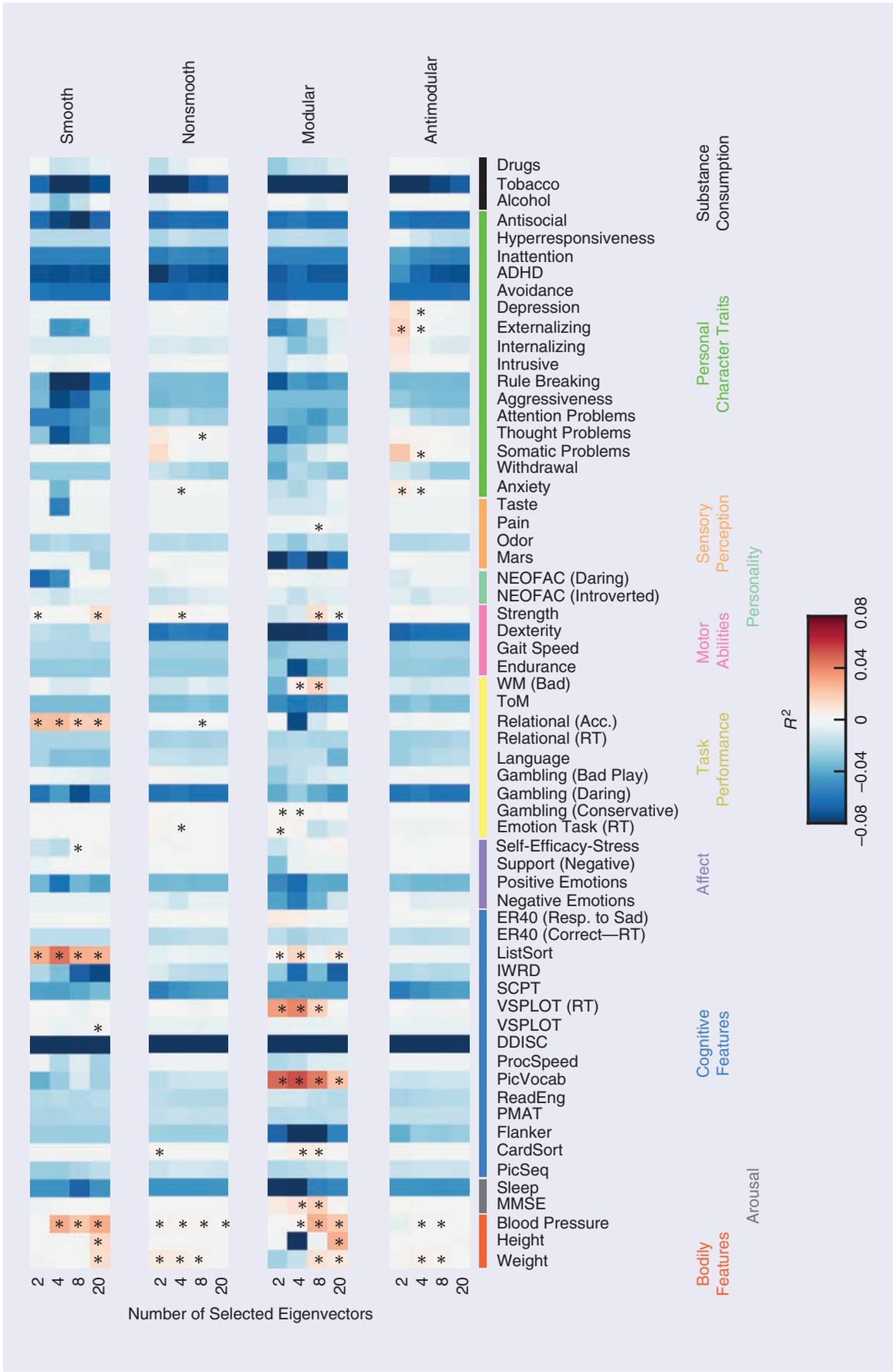


FIGURE 4. The link between brain function and behavior. Filtering schemes of fMRI time series: smooth and nonsmooth using the anatomical graph Laplacian and modular and antimodular using the modularity matrix. Results are shown for bandwidths of 2, 4, 8, or 20. The symbol (*) denotes significance on the whole population upon permutation testing ($p < 0.01$), with $R^2 > 0$. MMSE: mini mental state examination; PicSeq: picture sequence memory; CardSort: dimensional change card sort; PMAT: Penn progressive matrices; ReadEng: oral reading recognition; PicVocab: picture vocabulary; ProcSpeed: pattern completion processing speed; DDISC: delay discounting; VSPLOT: variable short Penn line orientation test; RT: response time; SCPT: short Penn continuous performance test; Perf.: performance; IWRD: Penn word memory test; ListSort: list sorting; ER40: Penn emotion recognition test; Acc.: accuracy; ToM: theory of mind; WM: working memory; Mars: contrast sensitivity; ADHD: attention-deficit hyperactivity disorder.

the gap between graph signal processing and network science. He recently developed, together with members of the MIPLab and including the authors of this article, a novel graph Slepian design that provides a smooth embedding of graph signals concentrated on a given subgraph. He is the corresponding author of this article.

Raphaël Liégeois (raphael.liegeois@epfl.ch) received his Ph.D. degree from the University of Liège, Belgium, and joined the Medical Image Processing Lab at École polytechnique fédérale de Lausanne in March 2018 as a postdoctoral fellow. His expertise lies at the crossroads of neuroscience, system identification, and graph theory. Most relevant to the present article are his works on the identification of sparse plus low-rank graphical models from neuroimaging time series and on the definition of alternative graph representations to explore the dynamics of signals defined on graphs.

Thomas A.W. Bolton (thomas.bolton@epfl.ch) received his M.Sc. degree from École polytechnique fédérale de Lausanne and has recently completed his Ph.D. degree in the Medical Image Processing Lab under the supervision of Prof. Van De Ville. His research interests cover various topics relating to models of neuroimaging data. In particular, he recently jointly led a project aiming at merging anatomical and functional MRI data using graph signal processing filtering that has inspired the framework we propose to utilize in the validation experiment of the present article. He is a Student Member of IEEE.

Dimitri Van De Ville (dimitri.vandeville@epfl.ch) is a professor of bioengineering at the École polytechnique fédérale de Lausanne and the University of Geneva, Switzerland. He is cochair of the biennial Wavelets & Sparsity series conferences, together with V. Goyal, Y. Lu, and M. Papadakis. He was a recipient of the Pfizer Research Award 2012, the National Alliance for Research on Schizophrenia and Depression Independent Investigator Award 2014, and the Leenaards Foundation Award 2016. His research interests include wavelets, sparsity, pattern recognition, and their applications in computational neuroimaging. He also explored the properties of alternative graph representations and developed various methodological frameworks exploiting the spectral properties of multivariate signals. He is a Fellow of IEEE.

References

- [1] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *IEEE Signal Process. Mag.*, vol. 30, no. 3, pp. 83–98, May 2013. doi: 10.1109/MSP.2012.2235192.
- [2] A. Sandryhaila and J. M. F. Moura, "Discrete signal processing on graphs," *IEEE Trans. Signal Process.*, vol. 61, no. 7, pp. 1644–1656, Apr. 2013. doi: 10.1109/TSP.2013.2238935.
- [3] D. K. Hammond, P. Vandergheynst, and R. Gribonval, "Wavelets on graphs via spectral graph theory," *Appl. Comput. Harmon. Anal.*, vol. 30, no. 2, pp. 129–150, Mar. 2011. doi: 10.1016/j.acha.2010.04.005.
- [4] M. Tsitsvero, S. Barbarossa, and P. Di Lorenzo, "Signals on graphs: Uncertainty principle and sampling," *IEEE Trans. Signal Process.*, vol. 64, no. 18, pp. 4845–4860, 2016. doi: 10.1109/TSP.2016.2573748.
- [5] S. Fortunato and D. Hric, "Community detection in networks: A user guide," *Phys. Rep.*, vol. 659, pp. 1–44, Nov. 2016. doi: 10.1016/j.physrep.2016.09.002.
- [6] M. E. J. Newman, "Fast algorithm for detecting community structure in networks," *Phys. Rev. E*, vol. 69, no. 6, pp. 066133-1–066133-5, June 2004. doi: 10.1103/PhysRevE.69.066133.

- [7] L. G. S. Jeub, P. Balachandran, M. A. Porter, P. J. Mucha, and M. W. Mahoney, "Think locally, act globally: Detection of small, medium-sized, and large communities in large networks," *Phys. Rev. E*, vol. 91, no. 1, pp. 012821-1–012821-29, Jan. 2015. doi: 10.1103/PhysRevE.91.012821.
- [8] R. Guimera, S. Mossa, A. Turtchi, and L. A. N. Amaral, "The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles," *Proc. Natl. Acad. Sci.*, vol. 102, no. 22, pp. 7794–7799, May 2005. doi: 10.1073/pnas.0407994102.
- [9] R. Guimera and L. A. N. Amaral, "Functional cartography of complex metabolic networks," *Nature*, vol. 433, no. 7028, pp. 895–900, Feb. 2005. doi: 10.1038/nature03288.
- [10] D. Meunier, R. Lambiotte, and E. T. Bullmore, "Modular and hierarchically modular organization of brain networks," *Front. Neurosci.*, vol. 4, pp. 200-1–200-11, Dec. 2010. doi: 10.3389/fnins.2010.00200.
- [11] M. E. J. Newman, "Finding community structure in networks using the eigenvectors of matrices," *Phys. Rev. E*, vol. 74, no. 3, pp. 036104-1–036104-19, Sept. 2006. doi: 10.1103/PhysRevE.74.036104.
- [12] U. von Luxburg, "A tutorial on spectral clustering," *Statist. Comput.*, vol. 17, no. 4, pp. 395–416, Aug. 2007. doi: 10.1007/s11222-007-9033-z.
- [13] A. Ortega, P. Frossard, J. Kovačević, J. M. F. Moura, and P. Vandergheynst, "Graph signal processing: Overview, challenges, and applications," *Proc. IEEE*, vol. 106, no. 5, pp. 808–828, May 2018. doi: 10.1109/JPROC.2018.2820126.
- [14] A. G. Marques, S. Segarra, G. Leus, and A. Ribeiro, "Stationary graph processes and spectral estimation," *IEEE Trans. Signal Process.*, vol. 65, no. 22, pp. 5911–5926, Nov. 2017. doi: 10.1109/TSP.2017.2739099.
- [15] M. Defferrard, X. Bresson, and P. Vandergheynst, "Convolutional neural networks on graphs with fast localized spectral filtering," in *Advances in Neural Information Processing Systems 29*, D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett, Eds. Red Hook, NY: Curran Associates, 2016, pp. 3844–3852.
- [16] W. Huang, T. A. W. Bolton, J. D. Medaglia, D. S. Bassett, A. Ribeiro, and D. V. D. Ville, "A graph signal processing perspective on functional brain imaging," *Proc. IEEE*, vol. 106, no. 5, pp. 868–885, May 2018. doi: 10.1109/JPROC.2018.2798928.
- [17] N. Leonardi and D. Van De Ville, "Tight wavelet frames on multislice graphs," *IEEE Trans. Signal Process.*, vol. 61, no. 13, pp. 3357–3367, 2013. doi: 10.1109/TSP.2013.2259825.
- [18] N. Tremblay and P. Borgnat, "Graph wavelets for multiscale community mining," *IEEE Trans. Signal Process.*, vol. 62, no. 20, pp. 5227–5239, 2014. doi: 10.1109/TSP.2014.2345355.
- [19] C. Lee and D. J. Wilkinson, "A review of stochastic block models and extensions for graph clustering," *Appl. Netw. Sci.*, vol. 4, no. 1, pp. 122-1–122-50, Dec. 2019. doi: 10.1007/s41109-019-0232-2.
- [20] C. P. Massen and J. P. K. Doye, "Identifying communities within energy landscapes," *Phys. Rev. E*, vol. 71, no. 4, Apr. 2005. doi: 10.1103/PhysRevE.71.046101.
- [21] M. E. J. Newman, "Assortative mixing in networks," *Phys. Rev. Lett.*, vol. 89, no. 20, pp. 208701-1–208701-4, Oct. 2002. doi: 10.1103/PhysRevLett.89.208701.
- [22] M. Bolla, B. Bullsins, S. Chaturapruek, S. Chen, and K. Friedl, "Spectral properties of modularity matrices," *Linear Algebra Appl.*, vol. 473, pp. 359–376, May 2015. doi: 10.1016/j.laa.2014.10.039.
- [23] E. Pirondini, A. Vybormova, M. Coscia, and D. V. D. Ville, "A spectral method for generating surrogate graph signals," *IEEE Signal Process. Lett.*, vol. 23, no. 9, pp. 1275–1278, Sept. 2016. doi: 10.1109/LSP.2016.2594072.
- [24] D. C. V. Essen, S. M. Smith, D. M. Barch, T. E. Behrens, E. Yacoub, and K. Ugurbil, "The WU-Minn human Connectome project: An overview," *NeuroImage*, vol. 80, pp. 62–79, Oct. 2013. doi: 10.1016/j.neuroimage.2013.05.041.
- [25] M. G. Preti and D. V. D. Ville, "Decoupling of brain function from structure reveals regional behavioral specialization in humans," *Nature Commun.*, vol. 10, no. 1, pp. 4747-1–4747-7, Oct. 2019. doi: 10.1038/s41467-019-12765-7.
- [26] J. D. Medaglia, W. Huang, E. A. Karuza, A. Kelkar, S. L. Thompson-Schill, A. Ribeiro, and D. S. Bassett, "Functional alignment with anatomical networks is associated with cognitive flexibility," *Nat Hum Behav.*, vol. 2, no. 2, pp. 156–164, 2018. doi: 10.1038/s41562-017-0260-9.
- [27] S. Sardellitti, S. Barbarossa, and P. D. Lorenzo, "On the graph Fourier transform for directed graphs," *IEEE J. Sel. Topics Signal Process.*, vol. 11, no. 6, pp. 796–811, Sept. 2017. doi: 10.1109/JSTSP.2017.2726979.
- [28] F. D. Malliaros and M. Vazirgiannis, "Clustering and community detection in directed networks: A survey," *Phys. Rep.*, vol. 533, no. 4, pp. 95–142, Dec. 2013. doi: 10.1016/j.physrep.2013.08.002.
- [29] Y.-Y. Ahn, J. P. Bagrow and S. Lehmann, "Link communities reveal multiscale complexity in networks," *Nature*, vol. 466, no. 7307, pp. 761–764, Aug. 2010. doi: 10.1038/nature09182.
- [30] S. Barbarossa and S. Sardellitti, "Topological signal processing over simplicial complexes," *IEEE Trans. Signal Process.*, vol. 68, pp. 2992–3007, Mar. 2020. doi: 10.1109/TSP.2020.2981920.