

Diffusion-based Virtual Graph Adjacency for Fourier Analysis of Network Signals

Vitor Rosa Meireles Elias, Wallace Alves Martins, and Stefan Werner

Abstract—This work proposes a graph model for networks where node collaborations can be described by the Markov property. The proposed model augments an initial graph adjacency using diffusion distances. The resulting virtual adjacency depends on a diffusion-scale parameter, which leads to a controlled shift in the graph-Fourier-transform spectrum. This enables a frequency analysis tailored to the actual network collaboration, revealing more information on the graph signal when compared to traditional approaches. The proposed model is employed for anomaly detection in real and synthetic networks, and results confirm that using the proposed virtual adjacency yields better classification than the initial adjacency.

Keywords—diffusion distances, virtual adjacency matrix, graph signal processing (GSP), graph Fourier transform (GFT).

I. INTRODUCTION

The connectivity of real-world elements and the amount of data generated in networks have been increasing consistently [1], [2]. Real networks and their corresponding data come in vastly different shapes and applications, ranging from genetic interaction networks [3] and the human brain [4] to sensor networks and smart cities [5]. Graph signal processing (GSP) explores pairwise relations between elements of a network to construct tools suitable for the processing of network data [1], [2], [6]–[12]. In GSP, networks are modeled as graphs and data defined over, or generated by, elements of these networks are modeled as a graph signal—a mapping from the set of vertices into the set of complex numbers. The relations between elements of the real network are embedded into the edges of the graph, connecting pairs of vertices.

Several GSP tools are functions of a graph-shift operator (GSO), which carries the network-structure information. Thus, the choice of the GSO and its properties impact the outcomes of many GSP tools [13]–[16]. Different approaches for GSP and the definition of the GSO have emerged over the last years [2], [6], [17] and many works are devoted to improve the framework [7], [18]. The two main approaches define the

GSO either as the adjacency matrix of the graph [6] or the Laplacian matrix [2].

GSP is a highly application-dependent framework. This dependency starts in the choice of the mapping from the real network into the graph [19], [20]. Most works rely on generic approaches for modeling the network. For example, defining the structure for physical networks in terms of distance between network elements, and adopting a GSO equal to the Laplacian matrix. If the network and application are known, a GSO can be tailored to the application at hand. Here, we consider Markov networks and the frequency analysis of the associated signals. In a Markov network, nodes collaborate with each other according to a defined Markov property, which defines the initial network adjacency.

We propose the use of diffusion distances (DDs) to incorporate the Markov property into the GSO. DD is a concept within the diffusion maps (DMs) framework proposed in [21]. The DM framework is conceived as a tool for uncovering a hidden geometry of the dataset by exploring properties of the eigenfunctions of the Markov matrix associated with the network states [22], [23]. The DDs serve as a metric for the diffusion-probability-based relation between two states of data. One parameter on the computation of DDs is the number of transition steps of the Markov chain, which corresponds to the stage or level of the collaboration. Hence, the proposed model depends on this number of steps. As graph spectrum depends on graph connectivity, the corresponding Fourier analysis adapts to the collaboration embedded in the model.

The combination of DM and GSP has been considered in [24], [25]. These works proposed the use of Markov matrices as GSO. The Markov matrix has desirable properties for GSP, such as being diagonalizable, and allows the use of DM-based tools, such as dimensionality reduction and clustering [24]. In contrast, we model the relation between elements of Markov network as a function of DDs. This yields an adjacency matrix that captures virtual dependencies between non-adjacent elements of the network. This is particularly relevant when nodes collaborate with neighbors over a sequence of interactions. The virtual adjacency reveals additional spectral content, which can be exploited by applications that make decisions based on frequency features, such as classifiers and detectors.

The paper is organized as follows: Section II reviews concepts of GSP and DM. Section III presents the proposed virtual-adjacency matrix and its effects on the graph Fourier analysis. In Section IV, we analyze the proposed model using numerical experiments and use it together with spectral analysis for anomaly detection in synthetic and real networks. Section V concludes the paper.

Mr. Elias is with the Department of Electronic Systems, NTNU – Norwegian University of Science and Technology (vitor.elias@ntnu.no); Dr. Martins is with the Interdisciplinary Centre for Security, Reliability and Trust, SnT, UniLu – University of Luxembourg (wallace.alvesmartins@uni.lu); Dr. Werner is with the Department of Electronic Systems, NTNU – Norwegian University of Science and Technology (stefan.werner@ntnu.no). Mr. Elias and Dr. Martins are also with the Electrical Engineering Program, PEE/COPPE, UFRJ – Federal University of Rio de Janeiro.

The work of Mr. Elias was supported, in part, by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Grant number: 88887.310189/2018-00 and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). The work of Dr. Martins was supported, in part, by the ERC project AGNOSTIC and, in part, by FAPERJ. The work of Dr. Werner was supported, in part, by the Research Council of Norway.

II. BACKGROUND AND NOTATION

A. Graph signal processing

The graph is denoted by $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, where $\mathcal{V} = \{v_1, \dots, v_N\}$ is the set of vertices (or nodes) and $\mathcal{E} = \{e_{11}, \dots, e_{NN}\}$ is the set of edges. Each vertex corresponds to one element of the network being modeled. Elements e_{ij} indicate pairwise relations between nodes v_i and v_j of the graph. An edge e_{ij} exists if and only if v_i and v_j are related (adjacent). These relations may incorporate functional properties of the network, based on the network data, or structural characteristics of the network, yielding relations based on the network elements. We represent the set of, possibly weighted, edges in the adjacency matrix \mathbf{A} .

The degree matrix \mathbf{D} is a diagonal matrix such that $D_{jj} = \text{deg}(v_j)$, where $\text{deg}(v_j) = \sum_{i \in \mathcal{N}_j} e_{ij}$ and \mathcal{N}_j is the set of vertices that are adjacent to v_j , referred to as neighborhood of v_j . Assuming a symmetric adjacency matrix, the graph Laplacian is the positive semidefinite matrix $\mathbf{L} = \mathbf{D} - \mathbf{A}$.

In the GSP framework [2], [6], the signal on a graph is given by the mapping $s : \mathcal{V} \rightarrow \mathbb{C}$ and is usually represented by a vector \mathbf{s} , such that the i th entry of \mathbf{s} is $s_i = s(v_i)$. The graph signal represents a snapshot of the network state at a given time instant. Let $\mathbf{L} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$ be the eigendecomposition of the graph Laplacian, where $\mathbf{\Lambda}$ is a diagonal matrix whose elements are the real eigenvalues λ_i and \mathbf{U} has the eigenvectors \mathbf{u}_i as columns, with $i \in \{1, \dots, N\}$. The graph Fourier transform (GFT) of a graph signal \mathbf{s} is given by $\hat{\mathbf{s}} = \mathbf{U}^T \mathbf{s}$ and the signal can be recovered from the GFT coefficients via the inverse GFT as $\mathbf{s} = \mathbf{U}\hat{\mathbf{s}}$ [2]. Here, \mathbf{L} is taken as the GSO. The eigenvalues carry a notion of graph frequency that quantifies the intensity of the signal variation across the graph nodes, with larger eigenvalues indicating higher variations [2]. We note that several definitions of the GFT exist in literature. For instance, [6] defines the GFT in terms of the eigendecomposition of \mathbf{A} , with \mathbf{A} taken as the GSO.

B. Diffusion maps and distances

In this section, we consider data-state-wise graphs, instead of network-wise graphs from GSP. Let $\mathbf{X} \in \mathbb{R}^{N \times K}$ be a data matrix that represents a set of K data points of dimension N , also called states, and assume that there is an underlying (hidden) process that relates the data points and possibly influences the data generation. The DM framework aims to make this process explicit [21]–[24]. Our focus is on one of the subproducts of the DM framework, which is the concept of *diffusion distances*. For a detailed description of DMs, see [21]–[24].

The first step of the DM framework is to create a graph in which nodes correspond to the columns of \mathbf{X} . The resulting graph is not intended to model a network, but to capture relations between data states. One may interpret it as a state-wise graph instead of the traditional element-wise graph described in Section II-A. The edges are computed via a symmetric non-negative kernel that maps the affinity between two states into a real value, thereby defining an adjacency matrix \mathbf{A} [21]. By normalizing the adjacency matrix, a probability of transitioning from state \mathbf{x}_i to \mathbf{x}_j is given by $p(\mathbf{x}_j|\mathbf{x}_i) = A_{ij}/\text{deg}(v_i)$.

A right-stochastic (Markov) matrix $\mathbf{M} = \mathbf{D}^{-1}\mathbf{A}$ comprises all transition probabilities [26]. Powers \mathbf{M}^t are associated with t steps of the random walk and $p_t(\mathbf{x}_j|\mathbf{x}_i) = M_{ij}^{(t)}$, the (i, j) th element of \mathbf{M}^t , denotes the probability of starting at \mathbf{x}_i and reaching \mathbf{x}_j after t steps.

The DD between two states is computed as [21], [23], [24]

$$D_t^2(\mathbf{x}_i, \mathbf{x}_j) = \sum_{z=1}^K \frac{(p_t(\mathbf{x}_z|\mathbf{x}_i) - p_t(\mathbf{x}_z|\mathbf{x}_j))^2}{\phi_{1,z}}, \quad (1)$$

where $\phi_{1,z}$ is the z th entry of ϕ_1 , the top left eigenvector of \mathbf{M} . The DD extends relations from the local structure given by \mathbf{A} into a global metric by assimilating probabilities of diffusion paths through the graph. Two points with similar posterior distributions have small DD, even if they are not initially adjacent. This can be seen in terms of paths connecting the two initial points through the end-points in the posterior distributions. The computation of the DD depends on the parameter t , which we denote as the diffusion-scale parameter. It represents the number of steps taken in the random walk. Increasing t allows an initial state to reach ending states that are further away in terms of steps of the random walk.

III. GSP FOR MARKOV NETWORKS

Markov networks inherently allow a modeling of the relation between its elements in terms of a stochastic matrix. Some networks that admit this modeling are: consensus networks [27]–[29], which perform a possibly weighted average of neighboring nodes in order to reach consensus through the entire network; conservative diffusion networks [30]–[32]; and random-walk driven networks [26], [33]. We propose to use DDs to incorporate the Markov property into the graph model. This approach renders a GSO tailored for Markov networks and provides a graph-frequency-analysis tool that offers more information on the signal when compared to traditional approaches.

A. Virtual adjacency

Modeling of networks for GSP usually relies on some strict constraints of adjacency between nodes. For instance, a wireless sensor network (WSN) is usually modeled in terms of the direct communication capabilities between sensors, given limitations imposed by the physical distance between them. If it is a consensus network, however, it operates on the data through iterative steps according to a stochastic matrix [27]. Thus, a node of the network is related to another node that is not adjacent through collaboration. We model this relation by adapting the concept of DDs and derive the virtual-adjacency matrix. Let the graph $\mathcal{G} = \{\mathcal{V}, \mathbf{B}\}$ model the network with an initial symmetric Markov-like adjacency matrix \mathbf{B} . The DD between nodes v_i and v_j (elements of the network) is defined by

$$D_t^2(v_i, v_j) = \sum_{n=1}^N \frac{(B_{in}^{(t)} - B_{jn}^{(t)})^2}{(1/N)}, \quad (2)$$

where $(1/N)$ corresponds to the elements of the top left eigenvector of \mathbf{B} and $B_{ij}^{(t)}$ denotes the (i, j) th element of \mathbf{B}^t .

The DD in (2) extends the relation of nodes to other nodes that are outside their original neighborhoods and this extension is dependent on the scale parameter t . From this definition of DD between elements of the network, we propose the virtual adjacency matrix $\mathbf{A}(t)$ defined by

$$A_{ij}(t) = \begin{cases} B_{ij} + \exp\left(-\frac{D_i^2(v_i, v_j)}{\rho N}\right) & i \neq j \\ 0 & i = j. \end{cases} \quad (3)$$

where ρ is a free parameter, N is the size of the network, and ρN prevents the diffusion-related term from quickly fading to zero as network size increases. Note that N cancels out with the denominator in (2). The parameter ρ provides sufficient adjustment of the function, meeting possible requirements of different applications.

It must be noted that DDs are always real positive values, yielding $A_{ij}(t) > 0$ for every $i \neq j$. Thus, instead of using the expression in (3) we actually consider edges to exist only if $A_{ij}(t)$ is greater than a given threshold; otherwise, $A_{ij}(t)$ is set to zero. For increasing t , the DD between nodes is non-decreasing. This leads to non-decreasing edge values and new edges possibly appearing as t increases. As t tends to infinity, the graph is expected to become fully connected. GSOs associated with different scales t model different stages of node collaboration.

B. Adaptable Fourier Analysis

As presented in Section II-A, the GFT is defined as the expansion of a graph signal in terms of the eigenvectors of the Laplacian matrix, which in turn depends on the adjacency matrix. Hence, the dependence on the diffusion-scale parameter t in the definition of $\mathbf{A}(t)$ is carried over to the GFT, since the diffusion-scale-dependent Laplacian matrix is

$$\mathbf{L}(t) = \mathbf{D}(t) - \mathbf{A}(t), \quad (4)$$

where $\mathbf{D}(t)$ is the degree matrix associated with $\mathbf{A}(t)$. Given the eigendecomposition of the Laplacian as $\mathbf{L}(t) = \mathbf{U}(t)\mathbf{\Lambda}(t)\mathbf{U}^T(t)$, the graph Fourier transform can be written as

$$\hat{\mathbf{x}}(t) = \mathbf{U}^T(t)\mathbf{x}, \quad (5)$$

which, now, depends on t . Laplacian eigenvalues of connected graphs are non-decreasing with addition of edges [34]. As the number of edges in the virtual-adjacency matrix is non-decreasing with t , increasing t also increases the eigenvalues of $\mathbf{L}(t)$. This yields a frequency analysis tailored for each stage of node collaboration. Intuitively, more collaboration results in more connected nodes. Therefore, variations in graph signals are perceived by more pairs of nodes, thus corresponding to larger graph frequencies if compared to the frequency content obtained using the initial adjacency. Hence, by incorporating node collaboration into the graph model, we provide a frequency analysis that reveals more information on the network signal than that offered by the conventional GFT. Therefore, applications that use spectrum-related features, such as classifiers and detectors, can benefit from the proposed methodology.

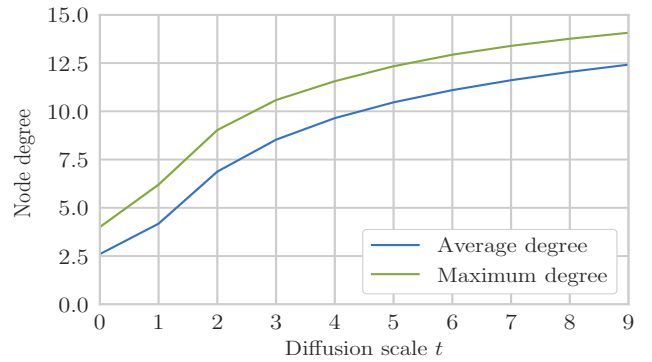


Fig. 1. Average and maximum node degree versus diffusion scale t . For $t = 0$, the values correspond to the average and maximum degrees of the original κ NN network.

IV. EXPERIMENTS

In this section, we present numerical experiments to showcase properties of the virtual-adjacency matrix and its effects on the Fourier analysis. Experiments are conducted over a sensor network that is modeled as a κ -nearest-neighbors (κ NN) graph $\mathcal{G} = \{\mathcal{V}, \mathbf{B}\}$, such that each sensor is bidirectionally connected to the κ nearest sensors. A sensor can be a nearest neighbor of a sensor that is not one of its nearest neighbors. Note that this allows a sensor to be connected to more than κ sensors. Let the symmetric matrix \mathbf{A} indicate connections between nodes, such that $A_{ij} = 1$ if and only if nodes v_i and v_j are connected. A consensus algorithm is implemented over this network, such that the updated network state at iteration $k + 1$ is given by $\mathbf{x}(k + 1) = \mathbf{B}\mathbf{x}(k)$, where the stochastic matrix \mathbf{B} is defined as $\mathbf{B} = \mathbf{I} - \epsilon\mathbf{L}$, where \mathbf{L} is the Laplacian of \mathbf{A} [27]. The parameter ϵ affects the convergence of the consensus algorithm and is set to $\epsilon = 1/(1.25\Delta)$, where Δ is the maximum degree in \mathcal{G} . Given \mathbf{B} , we can associate the network with a Markov chain and, hence, the techniques in Section III can be applied.

A. Increasing connectivity

From the initial stochastic adjacency matrix \mathbf{B} , we compute the diffusion distance between sensors in the network and the virtual-adjacency matrix is constructed according to (3). We consider a κ NN network with $N = 20$ sensors and $\kappa = 2$ nearest sensors. The parameter ρ in (3) is fixed at 0.35. Fig. 1 shows the graph average and maximum degrees of the virtual adjacency matrix versus diffusion scale, with scales varying from $t = 1$ to $t = 9$. As t is increased, more steps of the random walk are taken in the network and, given a reference node, the distance to nodes that are further away is expected to decrease. This results in more edges being created and, therefore, the connectivity of the graph increases. In fact, as t tends to infinity, the diffusion distance between any two nodes in a connected graph tends to zero and $\mathbf{A}(t)|_{t \rightarrow \infty}$ tends to the adjacency matrix of a fully connected graph with no self loops, i.e., each node is connected to every other node.

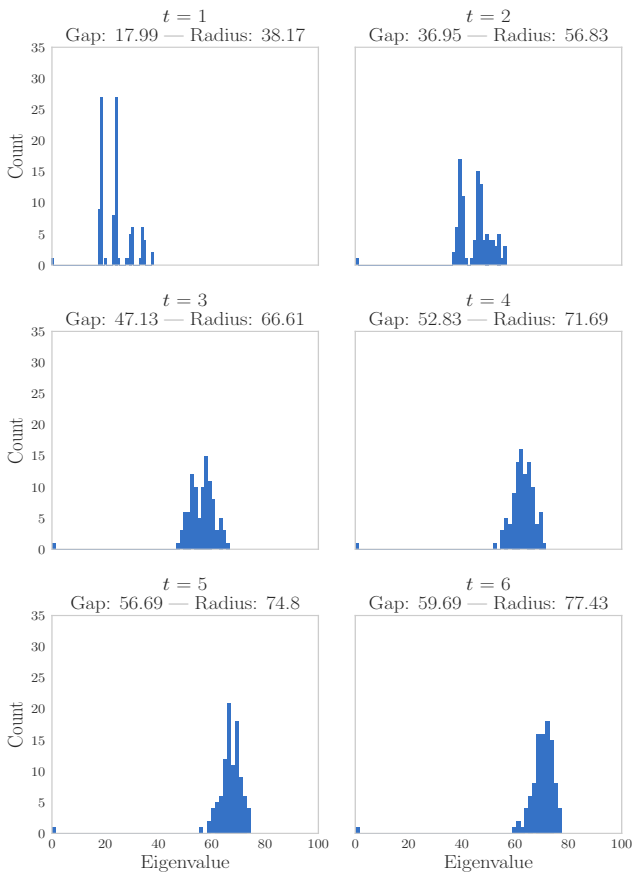


Fig. 2. Histogram of eigenvalues of diffusion Laplacian matrices \mathbf{L}_t , for $t \in \{1, \dots, 6\}$. Spectral “gap” denotes the smallest non-zero eigenvalue and spectral “radius” denotes the maximum eigenvalue.

B. Spectrum analysis

For a network with $N = 100$ sensors and $\kappa = 4$, we analyze the behavior of the GFT spectrum for different diffusion scales, for $\mathbf{A}(t)$ constructed with $\rho = 0.4$. The histograms of the eigenvalues of $\mathbf{L}(t)$ for $t \in \{1, \dots, 6\}$ are shown in Fig. 2. Increasing the diffusion scale and connectivity produces a shift on the graph-frequency spectral range into higher frequency ranges. In fact, as $t \rightarrow \infty$, the graph becomes fully connected and the maximum eigenvalue tends to $\{N + \theta_i\}_{i=1}^N$, where $\{\theta_i\}_{i=1}^N$, with $|\theta_i| \leq 1$, are the eigenvalues of the Laplacian of the stochastic matrix \mathbf{B} . Each diffusion scale t yields a frequency analysis associated with different collaboration stages. Thus, the virtual adjacency allows for additional spectrum information when compared to the initial adjacency. This additional information can be used by applications that make decisions based on the signal’s frequency content, as exemplified in the next section.

C. Application

The virtual adjacency is implemented in a synthetic consensus network and in a real weather-station network, wherein spectral analysis is used for anomaly detection [38]. Different diffusion scales and the parameter ρ allow for a tailored frequency decomposition. We compare detectors based the spectrum yielded by virtual adjacency matrices against the

TABLE I
F1 SCORES — SYNTHETIC DATA

| | 0 step | 1 step | 2 steps |
|---------------------------|-------------|-------------|-------------|
| Virtual Adjacency $t = 1$ | 0.70 | 0.62 | 0.53 |
| Virtual Adjacency $t = 2$ | 0.73 | 0.69 | 0.62 |
| Virtual Adjacency $t = 3$ | 0.74 | 0.66 | 0.61 |
| Initial Adjacency | 0.66 | 0.60 | 0.57 |

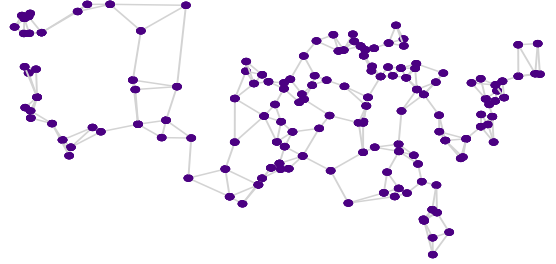


Fig. 3. Graph model for weather-stations network.

detector based on that of the initial adjacency. A similar task using GFT was applied in [35]–[37]. We evaluate the f1 score, given by the harmonic mean between precision and recall of the detectors.

We use a network with $N = 150$ sensors measuring a synthetic healthy signal drawn from a normal distribution with expected value equal to 20 and variance of 0.4. The anomaly is injected in up to 2 sensors that are randomly chosen as anomalous sensors. Anomalous sensor measurements are drawn from a uniform integer distribution in the interval [15, 25].

The detector is based on applying a threshold on the coefficients of the graph signal after a high-pass filter, assuming that anomalies induce high frequencies on the signal. We use grid search to optimize the filter cut-off frequency and the detection threshold for the conventional GFT. For the GFT based on virtual adjacency matrices, ρ is also optimized. We treat three scales $t \in \{1, 2, 3\}$ separately. This simulation is performed for three different stages of the consensus algorithm: raw data following previously described distributions; data after one consensus step; and data after two consensus steps. Results for the f1 score obtained by different detectors are presented in Table I, showing that the GFT using virtual adjacency achieves better results than those obtained by the conventional approach.

For anomaly detection in a real network, we use 150 randomly-selected weather stations, modeled as the graph in Fig. 3, and temperature data from the database in [38]. It must be noted that a Markovian relation is not initially defined for the nodes of this network. However, we are able to generate a Markovian relation by constructing a matrix $\mathbf{B} = \mathbf{I} - \epsilon\mathbf{L}$ based on physical adjacency. This example serves to show that the proposed methodology only requires a form of Markov model, but it is not restricted to Markov networks. Original data are converted from degrees Fahrenheit to degrees Celsius and range from -29.4°C to 38.6°C . We introduce anomaly in up to 7 stations. Anomaly is given by additive white Gaussian noise with variance equal to 1°C^2 and mean drawn from

the integer uniform distribution $[-7^\circ\text{C}, +7^\circ\text{C}]$. Conventional GFT achieves f1 score of 0.50, while GFT using the proposed model achieves better results with f1 score equal to 0.55 for $t = 1$ and 0.64 for $t = 2$.

V. CONCLUSION

We proposed a virtual-adjacency matrix which adapts Fourier analysis to node collaboration in Markov networks. We construct the adjacency matrix as a function of diffusion distances between elements of the network. The obtained virtual adjacency depends on the diffusion scale, given by the number of diffusion steps under consideration, such that increasing diffusion scales increases the connectivity of the graph. We showed that the virtual adjacency allows for an adaptable graph-frequency analysis considering different levels of collaboration between nodes. Changing the diffusion scale in the construction of the virtual-adjacency matrix shifts the range of frequencies discriminated by the GFT. Tools that operate on the graph spectrum can leverage on the additional information. For instance, we employed the proposed model for anomaly detection using spectral information. The resulting detectors leveraged on the proposed graph-frequency representation associated with different collaboration stages to obtain better results than those achieved by the GFT. As future work, we aim at exploring other applications for the proposed methodology and investigating efficient ways to determine the optimal diffusion-scale parameter.

REFERENCES

- [1] A. Sandryhaila and J. M. F. Moura, "Big data analysis with signal processing on graphs: Representation and processing of massive data sets with irregular structure," *IEEE Signal Process. Mag.*, vol. 31, pp. 80–90, Sep. 2014.
- [2] 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, pp. 83–98, May 2013.
- [3] B. Boucher and S. Jenna, "Genetic interaction networks: better understand to better predict," *Front. in Genet.*, vol. 4, pp. 1–16, Dec. 2013.
- [4] W. Huang, T. A. W. Bolton, J. D. Medaglia, D. S. Bassett, A. Ribeiro, and D. Van De Ville, "A graph signal processing perspective on functional brain imaging," *Proc. IEEE*, vol. 106, pp. 868–885, May 2018.
- [5] I. Jablonski, "Graph signal processing in applications to sensor networks, smart grids, and smart cities," *IEEE Sensors J.*, vol. 17, pp. 7659–7666, Dec. 2017.
- [6] A. Sandryhaila and J. M. F. Moura, "Discrete signal processing on graphs," *IEEE Trans. Signal Process.*, vol. 61, pp. 1644–1656, Apr. 2013.
- [7] A. Ortega, P. Frossard, J. Kovacevic, J. M. F. Moura, and P. Vandergheynst, "Graph signal processing: Overview, challenges, and applications," *Proc. IEEE*, vol. 106, pp. 808–828, May 2018.
- [8] G. Ribeiro and J. Lima, "Graph signal processing in a nutshell," *J. of Commun. and Inf. Syst. (JCIS)*, vol. 33, no. 1, pp. 219–233, July 2018.
- [9] V. R. M. Elias and W. A. Martins, "Graph Fourier transform for light field compression," in *Proc. XXXV Simpósio Brasileiro de Telecomunicações e Processamento de Sinais (SBRT)*, Sep. 2017, pp. 881–885.
- [10] V. R. M. Elias and W. A. Martins, "On the use of graph Fourier transform for light-field compression," *J. of Commun. and Inf. Syst. (JCIS)*, vol. 33, no. 1, pp. 92–103, May 2018.
- [11] G. Lewenfus, W. A. Martins, S. Chatzinotas, and B. Ottersten, "On the use of vertex-frequency analysis for anomaly detection in graph signals," in *Proc. XXXVII Simpósio Brasileiro de Telecomunicações e Processamento de Sinais (SBRT)*, Sep. 2019, pp. 1–5.
- [12] M. J. M. Spelta and W. A. Martins, "Normalized LMS algorithm and data-selective strategies for adaptive graph signal estimation," *Signal Process.*, vol. 167, pp. 107326, Feb. 2020.
- [13] A. Sandryhaila and J. M. F. Moura, "Discrete signal processing on graphs: Graph filters," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 2013, pp. 6163–6166.
- [14] O. Teke and P. P. Vaidyanathan, "Extending classical multirate signal processing theory to graphs - Part I: Fundamentals," *IEEE Trans. Signal Process.*, vol. 65, pp. 409–422, Jan. 2017.
- [15] O. Teke and P. P. Vaidyanathan, "Extending classical multirate signal processing theory to graphs - Part II: M-channel filter banks," *IEEE Trans. Signal Process.*, vol. 65, no. 2, pp. 423–437, Jan. 2017.
- [16] B. Girault, "Stationary graph signals using an isometric graph translation," in *Proc. 23rd Eur. Signal Process. Conf.*, Aug. 2015, pp. 1516–1520.
- [17] A. Gavili and X.-P. Zhang, "On the shift operator, graph frequency, and optimal filtering in graph signal processing," *IEEE Trans. Signal Process.*, vol. 65, pp. 6303–6318, Dec. 2017.
- [18] P. Djuric and C. Richard, *Cooperative and graph signal processing: principles and applications*. Orlando, FL, USA: Academic Press, Inc., 1st ed., 2018.
- [19] G. B. Giannakis, Y. Shen, and G. V. Karanikolas, "Topology identification and learning over graphs: Accounting for nonlinearities and dynamics," *Proc. IEEE*, vol. 106, pp. 787–807, May 2018.
- [20] G. Mateos, S. Segarra, A. G. Marques, and A. Ribeiro, "Connecting the dots: Identifying network structure via graph signal processing," *IEEE Signal Process. Mag.*, vol. 36, pp. 16–43, May 2019.
- [21] R. R. Coifman and S. Lafon, "Diffusion maps," *Appl. Comput. Harmon. Anal.*, vol. 21, pp. 5–30, July 2006.
- [22] R. R. Coifman, S. Lafon, A. B. Lee, M. Maggioni, B. Nadler, F. Warner, and S. W. Zucker, "Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps," *Proc. Natl. Acad. Sci.*, vol. 102, pp. 7426–7431, May 2005.
- [23] R. Talmon, I. Cohen, S. Gannot, and R. R. Coifman, "Diffusion maps for signal processing: A deeper look at manifold-learning techniques based on kernels and graphs," *IEEE Signal Process. Mag.*, vol. 30, no. 4, pp. 75–86, 2013.
- [24] A. Heimowitz and Y. C. Eldar, "A unified view of diffusion maps and signal processing on graphs," in *Proc. 12th Int. Conf. on Sampl. Theory Appl.*, July 2017, pp. 308–312.
- [25] A. Heimowitz and Y. C. Eldar, "The Nystrom extension for signals defined on a graph," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Apr. 2018, pp. 4199–4203.
- [26] L. Lovasz, "Random walks on graphs: A survey," *Combinatorics, Paul Erdos is Eighty*, vol. 2, pp. 1–46, 1993.
- [27] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proc. IEEE*, vol. 95, pp. 215–233, Jan. 2007.
- [28] J. Qin, Q. Ma, Y. Shi, and L. Wang, "Recent advances in consensus of multi-agent systems: A brief survey," *IEEE Trans. Ind. Electron.*, vol. 64, pp. 4972–4983, June 2017.
- [29] B. Kailkhura, S. Brahma, and P. K. Varshney, "Data falsification attacks on consensus-based detection systems," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 3, pp. 145–158, Mar. 2017.
- [30] D. Thanou, X. Dong, D. Kressner, and P. Frossard, "Learning heat diffusion graphs," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 3, pp. 484–499, Sep. 2017.
- [31] O. Teke and P. P. Vaidyanathan, "Time estimation for heat diffusion on graphs," in *Proc. 51st Asilomar Conf. Signals, Syst. Comput.*, Oct. 2018, pp. 1963–1967.
- [32] N. Masuda, M. A. Porter, and R. Lambiotte, "Random walks and diffusion on networks," *Phys. Rep.*, vol. 716–717, pp. 1–58, Nov. 2017.
- [33] C. Gkantsidis, M. Mihail, and A. Saberi, "Random walks in peer-to-peer networks," in *Proc. IEEE Conf. Comput. Commun.*, vol. 1, Mar. 2004, pp. 120–130.
- [34] B. Mohar, "The Laplacian spectrum of graphs," in *Graph Theory, Combinatorics, and Applications*, vol. 2, Wiley, 1991, pp. 871–898.
- [35] A. Sandryhaila and J. M. F. Moura, "Discrete signal processing on graphs: Frequency analysis," *IEEE Trans. Signal Process.*, vol. 62, pp. 3042–3054, June 2014.
- [36] E. Drayer and T. Routtenberg, "Detection of false data injection attacks in power systems with graph Fourier transform," in *Proc. IEEE Global Conf. Signal. Inf. Process.*, Nov. 2018, pp. 890–894.
- [37] M. Khatua, S. H. Safavi, and N. M. Cheung, "Sparse Laplacian component analysis for internet traffic anomalies detection," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 4, no. 4, pp. 697–711, Dec. 2018.
- [38] "Global Surface Summary of the Day - GSOD - Data.gov," <https://catalog.data.gov/dataset/global-surface-summary-of-the-day-gsod>. Accessed: 2019-07-29.