Enabling Prediction via Multi-Layer Graph Inference and Sampling

Stefania Sardellitti, Sergio Barbarossa, and Paolo Di Lorenzo
Sapienza Univ. of Rome, DIET Dept., Via Eudossiana 18, 00184, Rome, Italy
E-mail: {stefania.sardellitti; sergio.barbarossa; paolo.dilorenzo}@uniroma1.it

***Abstract*— In this work we propose a novel method to efficiently predict dynamic signals over both space and time, exploiting the theory of sampling and recovery of band-limited graph signals. The approach hinges on a multi-layer graph topology, where each layer refers to a spatial map of points where the signal is observed at a given time, whereas different layers pertain to different time instants. Then, a dynamic learning method is employed to infer space-time relationships among data in order to find a band-limited representation of the observed signal over the multi-layer graph. Such a parsimonious representation is then instrumental to use sampling theory over graphs to predict the value of the signal on a future layer, based on the observations over the past graphs. The method is then tested on a real data-set, which contains the outgoing cellular data traffic over the city of Milan. Numerical simulations illustrate how the proposed approach is very efficient in predicting the calls activity over a grid of nodes at a given daily hour, based on the observations of previous traffic activity over both space and time.**

***Index Terms*—Graph topology inference, multi-layer graphs, sampling over graphs, data traffic interpolation.**

I. Introduction

Over the last years, the research field known as Graph Signal Processing (GSP) has extended classical signal processing tools to the analysis of signals defined over graph, see, e.g., [1]–[3]. A key feature of GSP is that the analysis tools come to depend on the graph topology. As an example, the Graph Fourier Transform (GFT) for undirected graph has been defined as the projection of the observed signal onto the space spanned by the eigenvectors of the graph Laplacian matrix [1]. Consequently, whenever the graph structure is not known a priori, the inference of the topology from the observed data plays a crucial role in determining the properties of the graph signal. There are many works in literature aimed at learning the graph topology from a set of observations [4], [5]. By modelling the observations as random variables or processes, the graph topology typically reflects correlations among signals defined over its vertices. Alternative methods using the partial correlation [4] or Gaussian graphical models [6] have also been deeply investigated. Some GSP-based approaches make assumptions about the graph by enforcing properties such as sparsity and smoothness of the signals [7], [8]. There are also some recent works that focus on learning the graph topology from signals that diffuse over a graph [9], [10], [11]. In [12], Ioannidis *et al.* proposed novel methods based on structural models for joint inference of the network topology and time-evolving processes over the graph. This joint graph-process inference problem has been also deeply investigated in [13], where to reveal additional structural information of the network, the authors modelled data as multilayer networks whose nodes belong to different groups, termed layers [14]. The layers can represent, for example, the spatial variation of a graph signal at a given temporal snapshot, so that different layers represent the spatial-temporal evolution of the observed signal. Then, a key challenge becomes to learn the multi-layer graph topology by catching the time inter-layers relationships.

In our recent work [15], we proposed a method to associate a graph topology with the observed data in order to build a signal that is band-limited over the inferred graph. Enforcing band-limitedness of the signal enables the usage of sampling theory over graphs, see e.g. [16], which is appealing in all real situations where collecting a large number of observations is not possible or economically efficient.

In this work, we hinge on the graph topology inference proposed in [15] and we generalize it to build an efficient prediction strategy for data that vary both along space and time. In particular, the approach consists of two main steps: 1) Learn a multi-layer topology so that the space-time data can be modeled as a band-limited graph signal over the multi-layer graph; 2) Predict the signal over a future graph layer by hinging on sampling theory over the inferred multi-layer topology. As real data application of this framework, we consider the outgoing calls activity in the city of Milan, thus showing how, observing data samples from past layers, we can predict with a good accuracy the value of the signal over space and time.

II. Sampling Theory for Graph Signals

We consider an undirected graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ composed of N nodes, where $\mathcal{V} = \{1, \ldots, N\}$ and $\mathcal{E} = \{a_{ij}\}_{i,j \in \mathcal{V}}$, denote the vertex and edge sets, respectively, with $a_{ij} > 0$, if there is a link between node j and node i, or $a_{ij} = 0$, otherwise. The graph $\mathcal{G}$ can be described using the adjacency matrix $\mathbf{A} = \{a_{ij}\}_{i,j=1}^{N} \in \mathbb{R}^{N \times N}$, which collects all the edge weights a_{ij} for $i, j = 1, \ldots, N$. The combinatorial graph Laplacian is defined as $\mathbf{L} \triangleq \mathbf{D} - \mathbf{A}$, where $\mathbf{D}$ is the degree diagonal matrix containing the node degrees $d_i = \sum_j a_{ij}$ for $i = 1, \ldots, N$ on its diagonal. Since the graph is undirected, the Laplacian matrix is symmetric and positive semidefinite. We write its eigendecomposition as $\mathbf{L} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$, where $\mathbf{U}$

This work was supported by H2020 EUJ Project 5G-MiEdge, Nr. 723171.

is the matrix whose columns are the eigenvectors $\{\boldsymbol{u}_i\}_{i=1}^N$ of $\mathbf{L}$, whereas $\boldsymbol{\Lambda}$ is a diagonal matrix whose diagonal entries are the eigenvalues of $\mathbf{L}$. A signal $\boldsymbol{y}$ defined over a graph $\mathcal{G}$ is as a mapping from the vertex set to the set of real numbers, i.e. $\boldsymbol{y}: \mathcal{V} \rightarrow \mathbb{R}$. For undirected graphs, the GFT $\boldsymbol{s}$ of a graph signal $\boldsymbol{y}$ is defined as the projection of $\boldsymbol{y}$ onto the subspace spanned by the eigenvectors $\mathbf{U} = \{\boldsymbol{u}_i\}_{i=1}^N$ of the Laplacian matrix $\mathbf{L}$, i.e. $\boldsymbol{s} = \mathbf{U}^T\boldsymbol{y}$ [1], [2]. A band-limited graph signal can then be represented as:

$$\boldsymbol{y} = \mathbf{U}\boldsymbol{s}, \tag{1}$$

where the GFT vector $\boldsymbol{s}$ is sparse. Given a subset of indices $\mathcal{K} \subseteq \mathcal{V}$, we define the band-limiting operator over the set $\mathcal{K}$ as $\mathbf{B}_\mathcal{K} = \mathbf{U}\boldsymbol{\Sigma}_\mathcal{K}\mathbf{U}^T$, where $\boldsymbol{\Sigma}_\mathcal{K}$ is a diagonal matrix whose i-th diagonal entry is 1, if $i \in \mathcal{K}$, and 0 otherwise. A signal $\boldsymbol{y}$ is said to be perfectly band-limited, within the index set $\mathcal{K}$, if $\mathbf{B}_\mathcal{K}\,\boldsymbol{y} = \boldsymbol{y}$ [16]. If the signal is band-limited we can apply sampling theory to reconstruct the overall signal by observing only over a subset of nodes $\mathcal{I}$. Let us denote by $\mathbf{r} = \mathbf{G}_\mathcal{I}\,\boldsymbol{y}$ the observed signal, where $\mathbf{G}_\mathcal{I}$ is a vertex-selection diagonal matrix, whose i-th diagonal entry is given by 1 if $i \in \mathcal{I}$, and 0 otherwise. Necessary and sufficient conditions to recover $\boldsymbol{y}$ from the sampled signal $\mathbf{r}$ are that $\boldsymbol{y}$ is band-limited, i.e., $\mathbf{B}_\mathcal{K}\,\boldsymbol{y} = \boldsymbol{y}$, and [16]

$$\|\mathbf{G}_{\overline{\mathcal{I}}}\mathbf{B}_\mathcal{K}\|_2 < 1, \tag{2}$$

where $\mathbf{G}_{\overline{\mathcal{I}}}$ denotes the selection matrix operating over $\overline{\mathcal{I}}$, i.e., the complement set of $\mathcal{I}$; $\|\mathbf{A}\|_2$ denotes the spectral norm of matrix $\mathbf{A}$. A necessary condition for the recovery of the overall signal from a subset of samples is that the number of samples has not to be smaller than the bandwidth K, which is equal to the cardinality of $\mathcal{K}$. Assuming (2) holds true, the entire signal $\boldsymbol{y}$ can be recovered from $\mathbf{r}$ as follows:

$$\boldsymbol{y} = \mathbf{U}_\mathcal{K}\left(\mathbf{U}_\mathcal{K}^T\mathbf{G}_\mathcal{I}\mathbf{U}_\mathcal{K}\right)^{-1}\mathbf{U}_\mathcal{K}^T\,\mathbf{r} \tag{3}$$

where $\mathbf{U}_\mathcal{K}$ is the $N \times K$ matrix whose columns are the eigenvectors of $\mathbf{L}$ associated with the signal bandwidth [17].

III. Multi-layer graph topology inference

Let us assume to observe a set of N time series over T consecutive time instants. We denote the observation at time t as the vector $\boldsymbol{y}_i^t \in \mathbb{R}^{N\times 1}$, where $t = 1,\ldots,T$ is the temporal index within a time frame indexed by i. Our first goal is to capture the relations among the entries of these vectors $\boldsymbol{y}_i^t$ by associating to these values a multi-layer graph composed of T layers, where each layer refers to a generic time instant t and is given by a graph of N nodes. The structure of the multi-graph is not known a priori and we wish to infer its topology by observing M snapshots (frames) $\boldsymbol{y}_i^t$, for $i = 1,\ldots,M$. Given a number M of time frames, we infer the structure of the multi-layer graph by extending the method proposed in [15], which finds a block sparse representation of the data by learning a graph, whose Laplacian matrix admits the sparsifying dictionary as its eigenvectors. By extending our approach in [15] to a multi-layer structure, our goal is to associate a multi-layer graph topology with the data in such a way that the observed signal turns out to be band-limited over the inferred multi-layer graph. The key idea of this paper is that inferring a topology such that the observed signal is band-limited over the multi-layer graph enables the application of graph sampling theory to perform prediction.

Therefore, assuming that at each layer t we observe the same set of N nodes, we introduce the multi-layer graph $\mathcal{G}_T = \{\mathcal{V}_T, \mathcal{E}_T\}$, where the set of vertices $\mathcal{V}_T$ is composed of $N_T = NT$ nodes, with T the number of layers, or, equivalently, the number of time slots in the frame interval. The edge set $\mathcal{E}_T$ consists of all edges connecting nodes inside each layer *and* between layers. Then, we denote by $\mathbf{L}$ the $N_T \times N_T$ dimensional Laplacian matrix associated with $\mathcal{G}_T$. We assume to collect a set of M training vectors $\boldsymbol{y}_i = [\boldsymbol{y}_i^1;\ldots;\boldsymbol{y}_i^T] \in \mathbb{R}^{N_T\times 1}$, $i = 1,\ldots,M$, and we wish to enforce the band-limited structure by finding the orthonormal matrix $\mathbf{U} \in \mathbb{R}^{N_T\times N_T}$ and the sparse vectors $\boldsymbol{s}_i \in \mathbb{R}^{N_T\times 1}$, $i = 1,\ldots,M$, that minimize the fitting error $\sum_{i=1}^{N_T}||\boldsymbol{y}_i - \mathbf{U}\boldsymbol{s}_i||_F^2$. We denote by $\mathbf{S} \triangleq [\boldsymbol{s}_1,\ldots,\boldsymbol{s}_M]$ the GFT coefficient matrix of size $N_T \times M$ collecting all the vectors $\boldsymbol{s}_i$, $i = 1,\ldots,M$, where $\boldsymbol{s}_i$ is the GFT of the graph signal $\boldsymbol{y}_i$. Assuming that the transform coefficients $\boldsymbol{s}_i$ share a common zero support, $\mathbf{S}$ is a block-sparse matrix having multiple null rows and we define the set of K-block sparse matrices as $\mathcal{B}_K \triangleq \{\mathbf{S} = [\boldsymbol{s}_1,\ldots,\boldsymbol{s}_M] \in \mathbb{R}^{N_T\times M} |\, \mathbf{S}(n,:) = \mathbf{0}, \forall\, n \notin \mathcal{K} \subseteq \mathcal{V}_T,\, K = |\mathcal{K}|\}$, where $\mathbf{S}(n,:)$ denotes the nth row of $\mathbf{S}$. Then, defining the observation training matrix as $\mathbf{Y} \triangleq [\boldsymbol{y}_1,\ldots,\boldsymbol{y}_M]$, we can write this matrix in compact form as $\mathbf{Y} = \mathbf{U}\mathbf{S}$.

Our goal is to learn first the orthonormal transform matrix $\mathbf{U}$ and the sparse matrix $\mathbf{S}$ from the observation data-set $\mathbf{Y}$. Then, based on $\mathbf{U}$, we infer the graph topology by recovering the Laplacian matrix $\mathbf{L}$ that admits the columns of $\mathbf{U}$ as its eigenvectors. Proceeding as in [15], the first step is to learn the pair of matrices $(\mathbf{U},\mathbf{S})$ jointly, up to a rotation matrix, by finding the block-sparse columns $\{\boldsymbol{s}_i\}_{i=1}^M$, and the orthonormal vectors $\{\boldsymbol{u}_i\}_{i=1}^{N_T}$ that minimize the fitting error $\|\,\mathbf{U}^T\mathbf{Y} - \mathbf{S}\,\|_F^2$, by solving the optimization problem:

$$\begin{aligned} \min_{\mathbf{U}\in\mathbb{R}^{N_T\times N_T},\,\mathbf{S}\in\mathbb{R}^{N_T\times M}} \quad & \|\,\mathbf{U}^T\mathbf{Y} - \mathbf{S}\,\|_F^2 \qquad (\mathcal{P}_{U,S}) \\ \text{s.t.} \quad & \mathbf{U}^T\mathbf{U} = \mathbf{I},\;\; \boldsymbol{u}_1 = b\mathbf{1}, \\ & \mathbf{S} \in \mathcal{B}_K \end{aligned}$$

where we force $\mathbf{U}$ to be unitary and to contain an eigenvector proportional to the vector of all ones, by a coefficient $b = 1/\sqrt{N_T}$. Although problem $\mathcal{P}_{U,S}$ is non-convex, in [15] we proposed an algorithmic solution that alternates between the minimization with respect to $\mathbf{S}$ and $\mathbf{U}$ at each step k,

iteratively, as follows:

$$1.\quad \hat{\mathbf{S}}^{(k)} \triangleq \underset{\mathbf{S}\in\mathbb{R}^{N_T\times M}}{\arg\min} \quad \|(\hat{\mathbf{U}}^{(k-1)})^T\mathbf{Y}-\mathbf{S}\|_F^2 \qquad (\mathcal{S}_k)$$
$$\text{s.t.} \qquad \mathbf{S}\in\mathcal{B}_K,$$

$$2.\quad \hat{\mathbf{U}}^{(k)} \triangleq \underset{\mathbf{U}\in\mathbb{R}^{N_T\times N_T}}{\arg\min} \quad \|\mathbf{U}^T\mathbf{Y}-\hat{\mathbf{S}}^{(k)}\|_F^2 \qquad (\mathcal{U}_k)$$
$$\text{s.t.} \qquad \mathbf{U}^T\mathbf{U}=\mathbf{I},\ \ \boldsymbol{u}_1=b\mathbf{1},$$

where the superscript $^{(k)}$ denotes the step k of the algorithm. The method iterates until a termination criterion is met, within a prescribed accuracy. As illustrated in [15], even if the two problems $\mathcal{S}_k$ and $\mathcal{U}_k$ are non-convex, they admit a closed form solution, which can be efficiently evaluated with a low complexity cost.

Once we have found an estimate $\hat{\mathbf{U}}_\mathcal{K}$ of the Laplacian eigenvectors and $\hat{\mathbf{S}}$ of the sparse representation, the graph learning problem is formulated as follows

$$\underset{\mathbf{L}\in\mathbb{R}^{N_T\times N_T},\mathbf{C}_\mathcal{K}\in\mathbb{R}^{K\times K}}{\min} \quad f(\mathbf{L},\mathbf{Y},\hat{\mathbf{S}})+\mu||\mathbf{L}||_F^2 \qquad (\mathcal{P}_f)$$
$$\text{s.t.} \qquad \mathbf{L}\in\mathcal{L},\ \ \text{tr}(\mathbf{L})=p$$
$$\mathbf{L}\hat{\mathbf{U}}_\mathcal{K}=\hat{\mathbf{U}}_\mathcal{K}\mathbf{C}_\mathcal{K},\ \ \mathbf{C}_\mathcal{K}\succeq\mathbf{0}$$

where $\mathcal{L}$ represents the class of valid combinatorial Laplacian matrices, i.e.

$$\mathcal{L}\triangleq\left\{\mathbf{L}\in\mathcal{S}_+^{N_T}\mid\mathbf{L1}=\mathbf{0},\ L_{kn}=L_{nk}\leq 0,\forall k\neq n\right\},$$

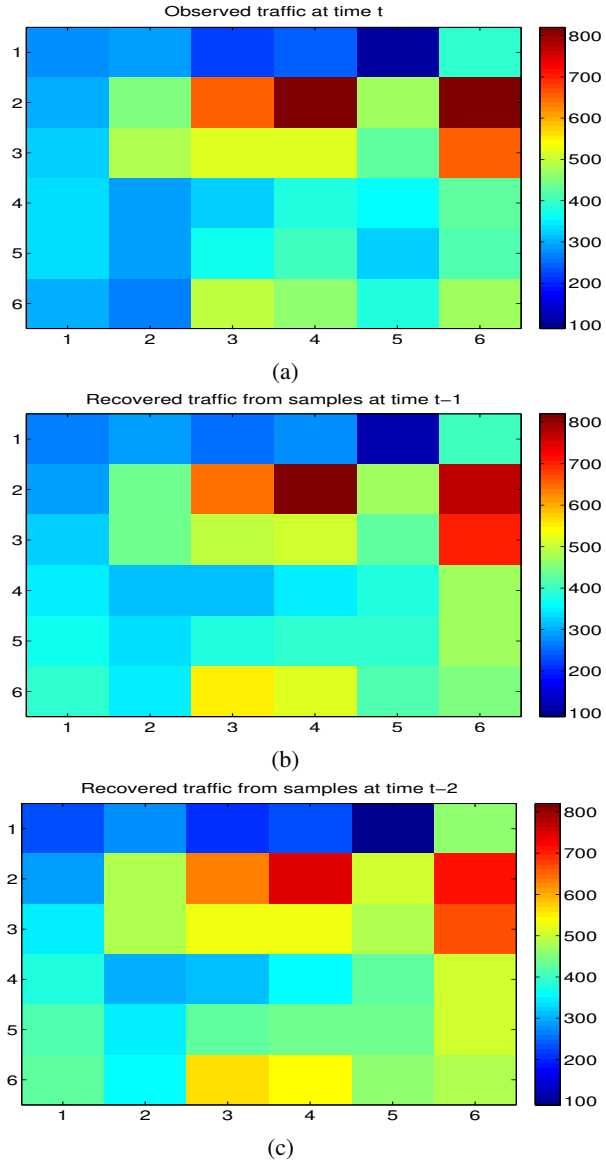
with $\mathcal{S}_+^{N_T}$ the set of real, symmetric and positive semidefinite matrices. In [15], we considered two different choices of the objective function $f(\mathbf{L},\mathbf{Y},\hat{\mathbf{S}})$, leading to the two graph learning strategies: i) the Total Variation based Graph Learning (TV-GL) algorithm where

$$f(\mathbf{L},\mathbf{Y})=\text{TV}(\mathbf{L},\mathbf{Y})=-\sum_{i=1}^{M}\sum_{k,n=1}^{N_T}L_{kn}|y_i(k)-y_i(n)|$$

is the total variation of the graph signals $\mathbf{Y}$, with $y_i(l)$ the lth entry of the observation vector $\boldsymbol{y}_i$. This approach tends to connect nodes where the observed signals are similar and to disconnect nodes with different values; ii) the Estimated-Signal-Aided Graph Learning (ESA-GL) algorithm, where

$$f(\mathbf{L},\hat{\mathbf{S}}_\mathcal{K})=\text{tr}(\mathbf{Y}^T\mathbf{L}\mathbf{Y}^T)=\text{tr}(\hat{\mathbf{S}}_\mathcal{K}^T\mathbf{C}_\mathcal{K}\hat{\mathbf{S}}_\mathcal{K}).$$

In problem $\mathcal{P}_f$, the trace of $\mathbf{L}$ is forced to assume a fixed value $p>0$ to avoid the trivial solution; furthermore, the Frobenius norm penalty in the objective function, controlled by the coefficient $\mu>0$, is added to avoid too sparse solutions. Indeed, increasing μ, the method drives the solution towards denser graphs with a small values of the Frobenius norm. Note that both the TV-GL and ESA-GL problems are convex, so that they can be efficiently solved.



(a)

(b)

(c)

Fig. 1: (a) Observed calls map from hours $9:00-10:00$ a.m. at time t corresponding to day 25-11-2013; recovered map from samples (b) at time $t-1$ and (c) at time $t-2$.

IV. Prediction of the traffic data maps

Once the multi-layer graph has been inferred by applying the strategy described in the previous section, we can reconstruct the overall signal by using (3) to recover the value of the traffic over the vertices belonging on the future layers, depending on the values observed over the past ones. In this work we apply our graph topology inference method to the recovery of the outgoing cellular activity generated by the telecommunication operator Telecom Italia, over the city of Milan, Italy [18]. The activity, expressed in terms of issued calls, is spatially aggregated using a squared grid of $N=36$ nodes, with a spatial resolution of about 235 meters, covering

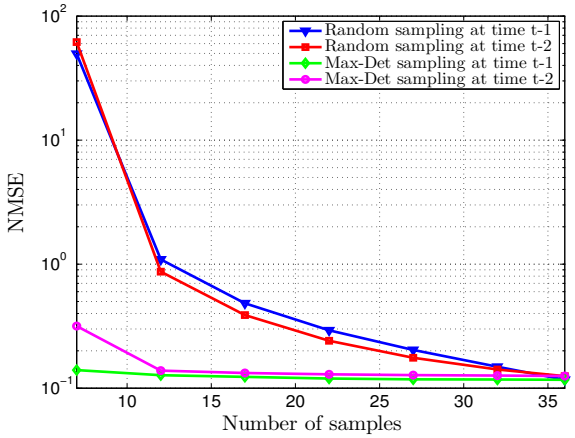


Fig. 2: Normalized mean squared error versus the number of samples.

the area around the historical center of Milan. We observe the calls daily activity map in November 2013 and data are temporally aggregated for each day over an interval of one hour, by observing a time frame of $T = 4$ consecutive days, from hours $7:00$ to $11:00$. The overall number of nodes in the inferred multi-layer graph $\mathcal{G}_T$ is $N_T = N \cdot T = 144$ and the layer t will correspond to day t, with $t = 1, \ldots, T$. Once the multi-layer graph has been learnt, we assume that such relationships hold for a certain amount of time. Exploiting such a structure, we apply sampling theory to predict the signal over future layers by collecting only a subset of samples in the previous layers. Thus, after inferring the graph eigenvectors matrix $\hat{\mathbf{U}}_{\mathcal{K}}$, with a very small bandwidth, i.e. $|\mathcal{K}| = 7$, we apply the Max-Det greedy sampling strategy in [16], [17], to select the subset of nodes over which to collect the signal values. Given this set of samples, we use a training set of $M = 12$ working days to predict from (3) the traffic map on the remaining ones. Then, in Figure 1a we report the aggregated call activity from 9 to 10 a.m., at time t corresponding to November 25th 2013. In Figs. 1b and 1c we illustrate the recovered maps by using 20 samples from the instants $t-1$ and $t-2$, respectively. Comparing these maps with the true one, shown in Fig. 1a, it can be noticed that the prediction is quite accurate even though based only on a limited number of previous observations. To better investigate how much we can predict the future activities from the past, in Fig. 2 we plotted the normalized mean squared error (NMSE) versus the number of samples for the cases where the calls activity at each time t is recovered from samples observed at the time intervals $t-1$ or $t-2$. The numerical results are averaged over 8 working days and 4 consecutive hours. We also compare the results by using the Max-Det sampling selection strategy with random sampling. We can notice that, as expected, the performance improves as the number of observed samples is increased, and random sampling performs worse than the Max-Det sampling.

V. Conclusions

In this work we proposed an efficient signal prediction method to recover signals that vary both over space and time. The approach is based on the inference of a multi-layer graph that describes the causal relations among the data, in such a way that the observed signals appear to be band-limited over the learned graph. Enforcing the signal band-limitedness property enables the use of sampling theory to predict the signal in future times from past observations. Numerical results on the prediction of real data traffic show the good performance of the proposed method.

References

[1] I. Pesenson, "Sampling in Paley-Wiener spaces on combinatorial graphs," *Trans. Amer. Math. Soc.*, vol. 360, no. 10, pp. 5603–5627, Oct. 2008.

[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, no. 3, pp. 83–98, May 2013.

[3] A. Sandryhaila and J. M. F. Moura, "Discrete signal processing on graphs: Frequency analysis," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3042–3054, June 2014.

[4] E. D. Kolaczyk, *Statistical analysis of network data: methods and models*, Springer Series Stat., 2009.

[5] G. Mateos, S. Segarra, and A. G. Marques, "Inference of graph topology," *in Cooperative and Graph Signal Processing: Principles and Applications (P.M. Djuric and C. Richard, Eds.), Amsterdam, Netherlands: Elsevier*, 2018.

[6] J. Friedman, T. Hastie, and R. Tibshirani, "Sparse inverse covariance estimation with the graphical lasso," *Biost.*, vol. 3, pp. 432–441, 2008.

[7] V. Kalofolias, "How to learn a graph from smooth signals," in *Proc. of the 19th Int. Conf. Art. Intel. Stat.*, A. Gretton and C. C. Robert, Eds., Cadiz, Spain, May 2016, vol. 51 of *Proc. of Mach. Lear. Res.*, pp. 920–929, PMLR.

[8] X. Dong, D. Thanou, P. Frossard, and P. Vandergheynst, "Learning Laplacian matrix in smooth graph signal representations," *IEEE Trans. Signal Process.*, vol. 64, no. 23, pp. 6160–6173, Dec. 2016.

[9] S. Segarra, A. G. Marques, G. Mateos, and A. Ribeiro, "Network topology inference from spectral templates," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 3, no. 3, pp. 467–483, Sep. 2017.

[10] B. Pasdeloup, V. Gripon, G. Mercier, D. Pastor, and M. G. Rabbat, "Characterization and inference of graph diffusion processes from observations of stationary signals," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 4, no. 3, pp. 481–496, Sep. 2018.

[11] D. Thanou, X. Dong, D. Kressner, and P. Frossard, "Learning heat diffusion graphs," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 3, no. 3, pp. 484–499, Sep. 2017.

[12] V. N. Ioannidis, Y. Shen, and G. B. Giannakis, "Semi-blind inference of topologies and dynamical processes over dynamic graphs," *IEEE Trans. Signal Process.*, vol. 67, no. 9, pp. 2263–2274, May 2019.

[13] G. B. Giannakis, Y. Shen, and G. V. Karanikolas, "Topology identification and learning over graphs: Accounting for nonlinearities and dynamics," *Proc. of the IEEE*, vol. 106, no. 5, pp. 787–807, 2018.

[14] M. Kivelä, A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno, and M. A. Porter, "Multilayer networks," *J. Complex Netw.*, vol. 2, no. 3, pp. 203–271, 2014.

[15] S. Sardellitti, S. Barbarossa, and P. Di Lorenzo, "Graph topology inference based on sparsifying transform learning," *IEEE Trans. Signal Process.*, vol. 67, no. 7, pp. 1712–1727, Apr. 2019.

[16] 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, Sep. 2016.

[17] P. Di Lorenzo, S. Barbarossa, and P. Banelli, "Sampling and recovery of graph signals," *in Cooperative and Graph Signal Processing: Principles and Applications (P.M. Djuric and C. Richard, Eds.), Amsterdam, Netherlands: Elsevier*, 2018.

[18] *Big data-Dandelion API*, datasets available at: https://dandelion.eu/.