Tropical Graph Signal Processing

Vincent Gripon
IMT Atlantique
655 Avenue du Technopôle, 29280 Plouzané, France
name.surname@imt-atlantique.fr

***Abstract*—For the past few years, the domain of graph signal processing has extended classical Fourier analysis to domains described by graphs. Most of the results were obtained by analogy with the study of heat propagation. We propose to perform a similar analysis in the context of tropical algebra, widely used in theoretical computer science to monitor propagation processes over graphs of distances. We introduce a Tropical Graph Fourier Transform and prove a few results on graph inference and the existence of a tropical uncertainty principle.**

***Index Terms*—graph signal processing, tropical algebra, graph inference, uncertainty principle**

I. Introduction

For the past few years, the emerging field of graph signal processing [1] has proposed a framework to extend classical signal processing tools to domains with a complex topology described by a graph. The key idea is to proceed by analogy with the study of heat diffusion, where on graphs diffusion correspond to the product of a matrix, representing the dependencies of signal components, and a vector, representing the intensity of the signal at the locations corresponding to vertices in the graph.

Graph signal processing tools offer perspectives to exploit structures of signals, when available. Examples include the locations of electrodes in electroencephalography [2], that of sensors in networks [3], or even more abstract ones like social networks [4]. However, the diffusion operator which was introduced to study heat propagation does not necessarily match well these applications where linearity of processes is debatable.

The purpose of our paper is to show it is possible to propose a graph signal processing framework for different algebras, conducive to better describing some mechanisms. More precisely, we focus on the case of tropical algebra, widely used in theoretical computer science to study propagation in graphs of distances or social networks.

We derive definitions analogous to those of classical graph signal processing, and introduce a Tropical Graph Fourier Transform (TGFT). We propose a simple algorithm to infer a graph from smooth signals and discuss the existence of a tropical graph uncertainty principle. Interestingly, all the proposed definitions and results find elegant echos in the classical graph signal processing setting.

The outline of the paper is as follows: Section 2 introduces notations and some definitions. Section 3 contains the definition of Tropical Graph Fourier Transform. Section 4 introduces a few results on graph inference and the tropical graph uncertainty principle. A few experiments are presented in Section 5 and Section 6 is a conclusion.

II. Notations and definitions

Let us work on the semiring $(\bar{\mathbb{R}}, \oplus, \otimes, +\infty, 0)$, where $\bar{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$, $\oplus = \min$ and $\otimes = +$. Hence, for two matrices in $M_n(\bar{\mathbb{R}})$ A and B, it holds that:

$$(A \cdot B)_{ij} = \min_k (A_{ik} + B_{kj}) ,$$

where $\cdot$ denotes the matrix product and A_{ij} is the coefficient at coordinates (i, j) in matrix A.

We term *graph* a couple $\mathcal{G} = \langle V, E \rangle$ where V is the finite set of vertices and $E \subseteq V \times V$ is the set of edges. We restrict our study to graphs for which each vertex is connected to itself ($\forall v \in V, (v, v) \in E$). Without loss of generality, we consider $V = \{1, \dots, n\}$. As a consequence, a graph is entirely specified by an adjacency matrix $A \in M_n(\bar{\mathbb{R}})$ such that

$$A_{ij} = \begin{cases} 0 & \text{if } i = j \\ 1 & \text{if } (i,j) \in E \wedge i \neq j \\ +\infty & \text{otherwise} \end{cases} .$$

We denote by $A^\top$ the transpose of A.

We term *weighted graph* a graph for which each edge is associated with a weight (but self loops), or equivalently, a graph which adjacency matrix can contain any nonnegative value instead of 1 where $(i, j) \in E, i \neq j$.

A *signal* is a collection of values in $\bar{\mathbb{R}}$ for each vertex in the graph. We denote it by $\mathbf{x} \in \bar{\mathbb{R}}^n$. The hermitian of a signal $\mathbf{x}$, denoted $\mathbf{x}^*$ is the signal $-\mathbf{x}^\top$, where $-$ is extended to $+\infty$ with $-(+\infty) = +\infty$.

As an example, consider a signal as a collection of dates corresponding to when vertices acquire some information. The entry A_{ij} should be interpreted as the duration it takes to communicate the piece of information from vertex j to vertex i (note the inversion of i and j).

Definition: 1 (smoothness). *Given a graph $\mathcal{G}$ and a signal $\mathbf{x}$, the smoothness of $\mathbf{x}$ over $\mathcal{G}$ is given by $\mathcal{L}(\mathbf{x}) = (\mathbf{x}^* \cdot A \cdot \mathbf{x})^*$.*

Proposition: 1. *Smoothness is a nonnegative value, reaching 0 only for signals $\mathbf{x}$ such that:*

$$\forall i, j, \mathbf{x}_i \leq \mathbf{x}_j + A_{ij} .$$

Proof: We rewrite:

$$\mathcal{L}(\mathbf{x}) = \min_{ij} (A_{ij} + \mathbf{x}_j - \mathbf{x}_i) .$$

In particular for $i = j = 1$, we obtain that $\mathcal{L}(\mathbf{x}) \leq 0$. Also, this quantity is nonnegative if and only if all terms are nonnegative. ■

A signal with smoothness 0 is thus such that the graph cannot reduce the dates of acquisition of signals for any vertex. More generally, the smoothness of a signal measures the maximal gain in dates of acquisition of information it is possible to obtain from the graph structure.

Definition: 2 (diffusion)**.** *Given a graph $\mathcal{G}$ and a signal $\mathbf{x}$, the diffusion of $\mathbf{x}$ over $\mathcal{G}$ is the signal $A \cdot \mathbf{x}$.*

Lemma: 1. *Consider a graph $\mathcal{G}$, then for any vertices i and j, $A^t_{ij}, t \in \mathbb{N}$ is the length of a shortest path in $\mathcal{G}$ from vertex i to vertex j using at most $t+1$ distinct vertices. Note that we choose the convention that $A^0 = I$ is the matrix with 0 on its diagonal and $+\infty$ everywhere else.*

Proof: We proceed by induction on t. The result is immediate for $t = 0$. Now suppose that for some t, any vertices i and j are such that A^t_{ij} is the length of a shortest path in $\mathcal{G}$ from vertex i to vertex j containing at most $t+1$ distinct vertices. Now let us fix i, j and a shortest path p containing at most $t+2$ edges from i to j. We consider three cases:

1) Case $i = j$. Then we observe that $\forall i, A_{ii} = 0$ as a consequence of the null diagonal in A.
2) Otherwise, p can be split into two subpaths p_1 and p_2 such that p_2 is of length exactly 1. Denote by k the intermediary vertex which is the end of p_1 and the beginning of p_2. These two subpaths are shortest paths, as by contradiction p would not be a shortest path. Note that they both contain at most p edges. Finally, we have $(A^{t+1})_{ij} = \min_{k'} \left(A^t_{ik'} + A^t_{k'j}\right) \leq A^t_{ik} + A^t_{kj}$.

■

Proposition: 2. *The sequence $\left(A^t \cdot \mathbf{x}\right), t \in \mathbb{N}$ is stationary.*

Proof: As a direct consequence of Lemma 1, and since A only contain nonnegative entries, we conclude that $\forall t \in \mathbb{N}, A^{t+n} = A^n$. ■

We denote by $A^\infty = \lim\limits_{t \to +\infty} A^t$ and $\mathbf{x}^\infty = \lim\limits_{t \to +\infty} A^t \cdot \mathbf{x}$.

Proposition: 3. *For any signal and graph, the smoothness of $\mathbf{x}^\infty$ is 0.*

Proof: Using Proposition 1, we proceed by contradiction. Suppose that $\mathcal{L}(\mathbf{x}^\infty) < 0$. Then it holds that: $\exists i, j, A_{ij} + \mathbf{x}^\infty_j < \mathbf{x}^\infty_i$. But $\mathbf{x}^\infty_i = (A \cdot \mathbf{x}^\infty_i) = \min_{j'} A_{ij'} + \mathbf{x}^\infty_{j'} \leq A_{ij} + \mathbf{x}^\infty_j$. ■

For our example, the signal $\mathbf{x}^\infty$ corresponds to the best dates for which vertices can acquire information given initial dates given by $\mathbf{x}$ and the duration graph $\mathcal{G}$.

Definition: 3 (Laplacian)**.** *The Laplacian of a graph $\mathcal{G}$ with adjacency matrix A is the matrix $L = I - A$ (recall that I is the matrix with a null diagonal and $+\infty$ everywhere else).*

Proposition: 4. *The Laplacian can admit only one eigenvalue which is 0.*

Proof: Denote $\mathbf{x}$ an eigenvector of L and μ the corresponding eigenvalue, i.e. $L \cdot \mathbf{x} = \mu \otimes \mathbf{x}$. Here, the $\otimes$ operator is extended such that $\forall i, (\mu \otimes \mathbf{x})_i = \mu + \mathbf{x}_i$. Then $A^t \cdot \mathbf{x} = (t\mu) \otimes \mathbf{x}$, which converges if and only if $\mu = 0$. Proposition 2 concludes. ■

Proposition: 5. *The eigenvectors of L are the signals for which $\mathbf{x} = \mathbf{x}^\infty$. Moreover, they are such that $\forall i, \forall j, \mathbf{x}_i \leq A_{ij} + \mathbf{x}_j$.*

Proof: Consider $\mathbf{x}$ an eigenvector of L. Proposition 4 gives us that $\forall i, \min_j \left(A_{ij} + \mathbf{x}_j\right) = \mathbf{x}_i$, leading to the second part of the proposition. Moreover, $L \cdot \mathbf{x} = x - A \cdot \mathbf{x} = 0$ gives us the first part. ■

III. Tropical Graph Fourier Transform

In this section we show it is possible to rewrite any signal $\mathbf{x}^\infty$ as a linear combination of eigenvectors of L. We call this representation Tropical Graph Fourier Transform of the signal $\mathbf{x}$.

We denote by $\mathbf{e}^i$ a signal with only $+\infty$ values but the i-th coordinate which is 0.

We first introduce the following result, which is a direct corollary of Lemma 1.

Corollary: 1. *For any signal $\mathbf{x}$, it holds that $\forall i, j, \mathbf{x}^\infty_i - \mathbf{x}^\infty_j \leq d(i,j)$, where $d(i,j)$ is the length of a shortest path between vertices i and j in the graph $\mathcal{G}$.*

We also introduce the following algorithm to compute the Tropical Graph Fourier Transform (TGFT) of a signal $\mathbf{x}$, denoted $\hat{\mathbf{x}}$:

Data: Graph adjacency matrix A, signal $\mathbf{x}$.
Result: $\hat{\mathbf{x}}$
Initialize each coordinate of $\hat{\mathbf{x}}$ as $+\infty$
while $\mathbf{x}^\infty \neq \hat{\mathbf{x}}^\infty$ **do**
  $i^* \leftarrow \arg \min\limits_{\substack{i \\ \mathbf{x}^\infty_i \neq \hat{\mathbf{x}}^\infty_i}} \{\mathbf{x}^\infty_i\}$
  $\hat{\mathbf{x}}_{i^*} \leftarrow \mathbf{x}_{i^*}$
end
Return $\hat{\mathbf{x}}$

Algorithm 1: Algorithm to compute the TGFT of a signal $\mathbf{x}$ over a graph $\mathcal{G}$.

Let us prove that Algorithm 1 ends:

Proposition: 6. *For any graph $\mathcal{G}$ and any signal $\mathbf{x}$, Algorithm 1 ends.*

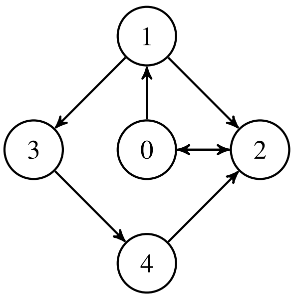
Proof: Denote by t the number of times the `while` loop has been visited in the algorithm. We denote by $\hat{\mathbf{x}}^t$ the corresponding value of $\hat{\mathbf{x}}$ in the algorithm and by i^t the chosen vertex i^*.

We denote $H(t)$ the hypothesis that

1) $\forall 1 \leq \tau < t, \mathbf{x}_{i^\tau} \leq \mathbf{x}_{i^{\tau+1}}$,
2) $\forall 1 \leq \tau, \tau' \leq t, i^\tau = i^{\tau'} \Rightarrow \tau = \tau'$.

First, note that $H(1)$ is trivially correct.

Now suppose $H(1), \ldots, H(t)$ for some t. There are two cases: $\mathbf{x}^\infty = \hat{\mathbf{x}}^{t\,\infty}$, in which case the algorithm ends, or $\exists j, \mathbf{x}^\infty_j \neq \hat{\mathbf{x}}^{t\,\infty}_j$. In the latter case, we want to show $H(t+1)$.



$$\mathbf{x} = (2, 9, 8, 1, 3)$$

$$\mathbf{x}^\infty = (2, 2, 3, 1, 3)$$

$$\hat{\mathbf{x}} = (2, +\infty, +\infty, 1, +\infty)$$

Figure 1. Example of a nonweighted graph, of a signal $\mathbf{x}$, of its diffusion limit $\mathbf{x}^\infty$, and of its TGFT $\hat{\mathbf{x}}$.

First let us show that $\forall \tau \leq t, i^\tau \neq i^{t+1}$. Suppose by contradiction that $\exists \tau \leq t, i^\tau = i^{t+1}$. Recall we have $\mathbf{x}^\infty_{i^\tau} \neq \hat{\mathbf{x}}^{t}_{i^\tau}{}^{\infty}$. Because $\hat{\mathbf{x}}^t_{i^\tau} = \mathbf{x}_{i^\tau}$, we more precisely have $\mathbf{x}^\infty_{i^\tau} > \hat{\mathbf{x}}^{t}_{i^\tau}{}^{\infty}$. This is not possible because of Lemma 1 and the fact the non $+\infty$ values in $\hat{\mathbf{x}}^t$ are identical to those in $\mathbf{x}$.

Second, let us show that $\mathbf{x}_{i^{t+1}} \geq \mathbf{x}_{i^t}$. Suppose by contradiction that $\mathbf{x}_{i^{t+1}} < \mathbf{x}_{i^t}$. At step t, we also have $\mathbf{x}^\infty_{i^t} \neq \hat{\mathbf{x}}^{t-1}_{i^t}{}^{\infty}$. Since i^{t+1} has not been chosen at step t, it must holds that $\mathbf{x}^\infty_{i^{t+1}} = \hat{\mathbf{x}}^{t-1}_{i^{t+1}}{}^{\infty}$. But we have $\mathbf{x}^\infty_{i^{t+1}} \neq \hat{\mathbf{x}}^{t}_{i^{t+1}}{}^{\infty}$, which means that $\hat{\mathbf{x}}^{t}_{i^{t+1}}{}^{\infty} \neq \hat{\mathbf{x}}^{t-1}_{i^{t+1}}{}^{\infty}$. More precisely $\hat{\mathbf{x}}^{t}_{i^{t+1}}{}^{\infty} < \hat{\mathbf{x}}^{t-1}_{i^{t+1}}{}^{\infty}$ because of Lemma 1 and $\hat{\mathbf{x}}^t \leq \hat{\mathbf{x}}^{t-1}$. Since the only difference between $\hat{\mathbf{x}}^t$ and $\hat{\mathbf{x}}^{t-1}$ is the i^t-th coordinate, by Lemma 1 it means that $\mathbf{x}_{i^{t+1}} \geq \mathbf{x}_{i^t}$.

We conclude by induction. Note that the set $\{i^1, \dots, i^t\}$ is increasing. Since there is a finite number of vertices, we conclude. ■

We conclude by observing that the output of the algorithm is a signal $\hat{\mathbf{x}}$ such that $\hat{\mathbf{x}}^\infty = \mathbf{x}^\infty$. We point out that $\hat{\mathbf{x}} = \bigoplus_{i \in S} \mathbf{x}_i \otimes \mathbf{e}^i$, where S is the set of marked vertices at the end of Algorithm 1.

Definition: 4. *The signal $\hat{\mathbf{x}}$ result of Algorithm 1 is called the* Tropical Graph Fourier Transform *of $\mathbf{x}$ over $\mathcal{G}$.*

Note the immediate corollary:

Corollary: 2. *For any signal and any graph, it holds that $\hat{\mathbf{x}} = \widehat{A \cdot \mathbf{x}}$.*

Proof: This result is directly derived from the proof of Proposition 6. ■

Figure 1 depicts an example of a graph, of a signal, of its diffusion limit and of its TGFT.

IV. Graph inference and uncertainty principle

An important question when it comes to classical graph signal processing is how to infer a graph from signal observations. A common way to answer it consists of supposing signals to be smooth [] or stationary [5], [6] over the graph. In the latter case, one can show that the eigenvectors of the covariance matrix are identical to those of the graph.

Following the same vein, we consider m observations of signals with dimension n, grouped in the matrix $\mathcal{X}$ with n lines and m columns. We introduce the following definition:

Definition: 5 (induced graph)**.** *The induced graph of a set of signals $\mathcal{X} \in \bar{\mathbb{R}}^{mn}$ is the graph which adjacency matrix is $h((\mathcal{X} \cdot \mathcal{X}^*)^*)$, where h is applied component-wise and is such that*

$$h : x \mapsto \begin{cases} +\infty & \textit{if } x < 0 \\ x & \textit{otherwise} \end{cases} .$$

When signals are only containing 0 and $+\infty$ values, obtained graphs are similar to binary associative memories [7].

Note that an induced graph is more easily recovered with signals containing only a few coefficients that are not $+\infty$ in their TGFT, as many of the difference of their coordinates correspond to weights of the targeted adjacency matrix.

Also, in classical graph signal processing, the eigenvectors of a matrix or any of its powers are identical, such that retrieving a graph is up to the power of its adjacency matrix. The same result holds here, where the induced graph is the transitive closure of the targeted graph. This motivates for the following definition:

Definition: 6 (binary induced subgraph)**.** *The binary induced subgraph of a set of signals $\mathcal{X} \in \{0, +\infty\}^{mn}$ is obtained by taking the corresponding induced subgraph and keeping only its weights that are lesser or equal to 1, the others being put to $+\infty$.*

Proposition: 7. *Consider observed signals to be with smoothness 0. Then the corresponding binary induced graph may contain spurious edges but does not miss any one.*

Proof: First note that the signal $\mathbf{0}$ is of smoothness 0 for any graph. The corresponding binary induced graph is containing only 0s in its adjacency matrix, proving that some edges might be spurious.

Now consider by contradiction a missing edge from vertex i to vertex j. This means that there is some observed signal $\mathbf{x}$ for which $\mathbf{x}_i > \mathbf{x}_j + 1$ despite there is an edge from i to j in the initial graph. This contradicts Lemma 1. ■

Next, we propose definitions loosely based on those in [8]. We now consider a connected graph (A^n contains no $+\infty$ coefficient).

Definition: 7 (spatial spread)**.** *The spatial spread of a signal $\mathbf{x}$ is defined as $\Lambda(\mathbf{x}) = -\mathbf{x}^* \cdot \bar{I} \cdot \mathbf{x}$, where $\bar{I}$ is the matrix which diagonal is $+\infty$ and all other coefficients are 0.*

Note that the spatial spread of a signal is the maximum difference between two values in $\mathbf{x}$.

Definition: 8 (spectral spread)**.** *The spectral spread of a signal $\mathbf{x}$, denoted $\Delta(\mathbf{x})$, is the number of coordinates that are not $+\infty$ in its TGFT.*

Definition: 9 (uncertainty domain and uncertainty compromise)**.** *The uncertainty domain of a graph $\mathcal{G}$ is the set of couples $\{(\Delta(\mathbf{x}^\infty), \Lambda(\mathbf{x}^\infty)), \mathbf{x} \in \bar{\mathbb{R}}^n - \{+\infty\}^n\}$. The uncertainty compromise is the function $\Gamma(s) = \min_{\substack{\mathbf{x} \\ \Delta(\mathbf{x}^\infty) = s}} \Lambda(\mathbf{x}^\infty)$.*

We have the following results:

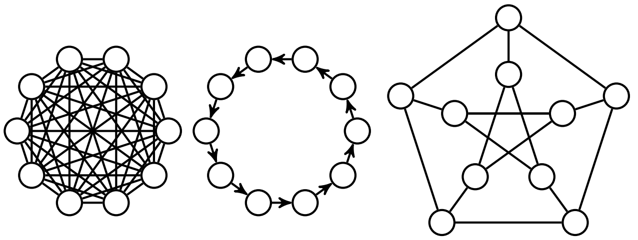


Figure 2. Graphs used for our simulations: complete, cycle and Petersen.

Proposition: 8. *For any graph,* $\Gamma(1)$ *is the radius of the graph.*

Proof: Note that a signal with spectral spread 1 is necessarily of the form $\mathbf{e}^i$. Lemma 1 concludes. ■

Proposition: 9. *For any unweighted graph with order* n*,* $\forall s < n, \Gamma(s) \geq 1$.

Proof: Note that $\Gamma(s) = 0$ gives that $\mathbf{0}$ has spectral spread s. Suppose by contradiction that $s < n$. Consider a vertex i in $\hat{\mathbf{x}}$ which value is $+\infty$. Using Lemma 1 gives us that there is a shortest path with length 0 from i to some other vertex j, contradicting the fact the graph is unweighted. ■

Proposition: 10. *For any graph,* $\Gamma(n) = 0$.

Proof: Consider $\mathbf{0}$. ■

V. Experiments

Throughout this section, we consider several unweighted graphs. Namely, we consider the graph from Figure 1, and three graphs of order 10. The latter are: a) a complete symmetric graph, b) a directed cycle graph and c) the famous Petersen graph. These graphs are depicted in Figure 2.

A. Graph inference from smooth signals

In our first experiments, we have generated random signals in $\{1, \ldots, 20\}^n$ where each value is drawn independently from each other with a uniform probability. These signals have then been diffused up to convergence and processed to obtain the corresponding binary induced graph.

Because of Proposition 7, we are only interested in measuring the expected number of spurious edges in the reconstructed graphs. We thus derive this empirical probability using Monte-Carlo simulations. The results are depicted in Figure 3 as a function of the number of observations m.

B. Uncertainty compromise of classical graphs

We also generated the uncertainty compromise corresponding to these graphs using an exhaustive approach. The results are depicted in Figure 4. As we can see, some graphs offer better compromise than others, and the best one is achieved with the complete graph.

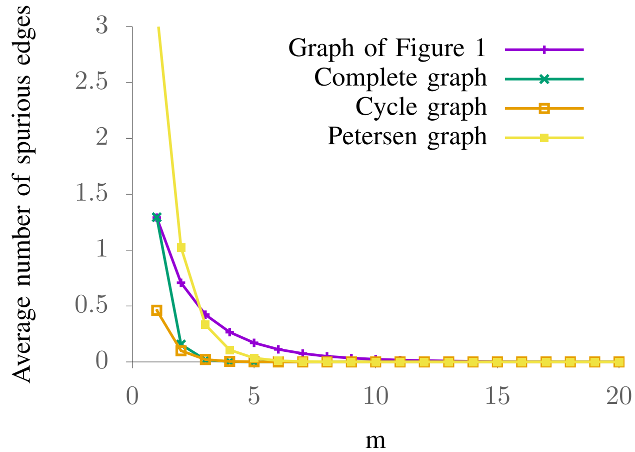


Figure 3. Average number of spurious edges when reconstructing a graph from random smooth signals, as a function of the number of observations m.

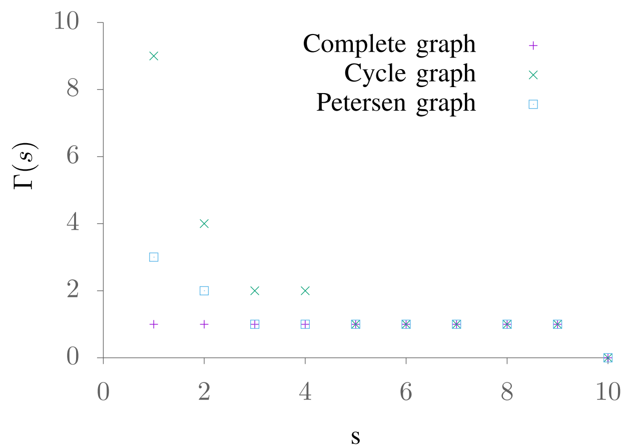


Figure 4. Uncertainty compromise for various graphs of order 10.

VI. Conclusion

We have introduced the notion of tropical graph signal processing. The proposed definitions are loosely based on those of classical signal processing. Interestingly, many classical results find an echo in the tropical setting, leading to possible new ways to confront real world applications.

These definitions are merely an introduction to the domain, and are debatable. We hope these ideas will help get a better understanding of graph signal processing, and provide a socle for extending it to other algebras.

Acknowledgements

This work has been funded in part by the CominLabs project Neural Communications.

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 Processing Magazine*, vol. 30, no. 3, pp. 83–98, 2013.

[2] W. Huang, L. Goldsberry, N. F. Wymbs, S. T. Grafton, D. S. Bassett, and A. Ribeiro, "Graph frequency analysis of brain signals," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 7, pp. 1189–1203, 2016.

[3] A. Sandryhaila and J. M. Moura, "Discrete signal processing on graphs," *IEEE transactions on signal processing*, vol. 61, no. 7, pp. 1644–1656, 2013.

[4] ——, "Big data analysis with signal processing on graphs: Representation and processing of massive data sets with irregular structure," *IEEE Signal Processing Magazine*, vol. 31, no. 5, pp. 80–90, 2014.

[5] B. Pasdeloup, V. Gripon, G. Mercier, D. Pastor, and M. G. Rabbat, "Characterization and inference of graph diffusion processes from observations of stationary signals," *Arxiv preprint:1605:02569v2*, 2016.

[6] A. G. Marques, S. Segarra, G. Leus, and A. Ribeiro, "Stationary graph processes and spectral estimation," *arXiv preprint arXiv:1603.04667*, 2016.

[7] D. J. Willshaw, O. P. Buneman, and H. C. Longuet-Higgins, "Non-holographic associative memory." *Nature*, 1969.

[8] A. Agaskar and Y. M. Lu, "A spectral graph uncertainty principle," *IEEE Transactions on Information Theory*, vol. 59, no. 7, pp. 4338–4356, 2013.