

# A SHORT-GRAPH FOURIER TRANSFORM VIA PERSONALIZED PAGERANK VECTORS

Mariano Tepper and Guillermo Sapiro

Department of Electrical and Computer Engineering, Duke University

## ABSTRACT

The short-time Fourier transform (STFT) is widely used to analyze the spectra of temporal signals that vary through time. Signals defined over graphs, due to their intrinsic complexity, exhibit large variations in their patterns. In this work we propose a new formulation for an STFT for signals defined over graphs. This formulation draws on recent ideas from spectral graph theory, using personalized PageRank vectors as its fundamental building block. Furthermore, this work establishes and explores the connection between local spectral graph theory and localized spectral analysis of graph signals. We accompany the presentation with synthetic and real-world examples, showing the suitability of the proposed approach.

**Index Terms**— Graph, localized Fourier transform, personalized PageRank, local spectral graph theory.

## 1. INTRODUCTION

The Fourier transform globally decomposes a temporal signal into its constituting frequencies, identifying their contribution to the signal formation. Often, temporal signals vary their behavior through time; in these cases, the Fourier transform, being global, falls short as a tool to analyze the characteristics of these signals. The short-time Fourier transform (STFT) [1] is used to analyze the Fourier spectrum of temporally localized sections of the signal. It is well studied that there is a trade-off between resolution (sharpness) in time and its counterpart in frequency. There is no way to get arbitrarily sharp analysis in both domains simultaneously [2, Sec. 2.6.2].

Formally, in the STFT a window function  $w$  which is nonzero for only a short period of time is slid along the time axis and multiplied by the input signal  $f$ ; then the Fourier transform of the resulting signal is taken. Formally, for one dimensional signals,

$$\text{STFT}_f(u, \xi) \doteq \int_{-\infty}^{\infty} f(t)w(t-u)e^{-i\xi t}dt. \quad (1)$$

The changing spectra is usually analyzed as a function of the time-shift  $u$  and is well suited to analyze time-varying signals. The above formula can be interpreted as the following three-sets algorithm: (1) translate the window  $w$  by  $u$ , (2) modulate

the result by frequency  $\xi$ , and (3) take the convolution of the result with the signal  $f$ . This can be written as

$$\text{STFT}_f = \langle f, \mathcal{M}_\xi T_u w \rangle, \quad (2)$$

where  $(T_u g)(t) \doteq g(t-u)$  and  $(\mathcal{M}_\xi g)(t) \doteq g(t)e^{i\xi t}$  are the translation and modulation operators, respectively.

Weighted graphs are a natural representational structure in most modern network applications (including, for example, social, energy, transportation, and sensor networks). These graphs are loaded with information, usually in the form of high-dimensional data (i.e., signals) that reside on the vertices (nodes) of graphs. Graph signal processing lies at the intersection of graph theory and computational harmonic analysis and seeks to process such signals on graphs. See [3, 4] for further details and references on this emerging field.

Graph signal processing has been successful at characterizing the equivalent of the Fourier transform in graph domains. Many different types of localized spectral transforms have been proposed in recent years, see [3, Sec. IV] for a thorough discussion. This list includes a windowed Fourier transform [5], later described in this work.

One of the main examples are diffusion wavelets [6], which are based on compressed representations of powers of a graph diffusion operator. In parallel, local spectral techniques, in which personalized PageRank vectors play a prominent role, have become increasingly popular in the field of community detection in graphs [e.g., 7]. As we will see in Section 2, the PageRank equation is defined recursively, and we can consider a single PageRank vector in place of a sequence of random walk vectors, or of powers of a diffusion operator [8].

In this work, we establish and explore for the first time the connection between local spectral graph methods and localized spectral analysis of graph signals. This work is a first step in this exploration, and introduces a short-graph Fourier transform inspired on the ideas of local graph analysis, using personalized PageRank vectors as fundamental building blocks of the method.

The remainder of the paper is organized as follows. In Section 2 we introduce our short-graph Fourier transform. Experimental results on synthetic and real graphs are presented in Section 3, showing the interesting characteristics of the proposed formulation. Finally, we provide some concluding remarks in Section 4.

## 2. FROM LOCAL SPECTRAL GRAPH THEORY TO A SHORT-GRAPH FOURIER TRANSFORM

We begin by introducing the notation and fundamental formulas used throughout the paper.

Let  $\mathbf{X}$  be a matrix. In the following,  $(\mathbf{X})_{ij}$ ,  $(\mathbf{X})_{:j}$ ,  $(\mathbf{X})_i$  denote the  $(i,j)$ th entry of  $\mathbf{X}$ , the  $j$ th column of  $\mathbf{X}$ , and the  $i$ th row of  $\mathbf{X}$ , respectively.

We consider a graph  $G = (V, \mathbf{A})$ , where  $|V| = n$  and  $\mathbf{A} \in \mathbb{R}_+^{n \times n}$  is the weighted adjacency matrix. The weighted entry  $(\mathbf{A})_{ij}$  represents in most applications a measure of similarity between vertices  $i$  and  $j$ . We assume that  $G$  is connected and undirected, i.e.,  $(\mathbf{A})_{ij} = (\mathbf{A})_{ji}$ . The degree of a node  $i \in V$  is  $d_i \doteq \sum_{j=1}^n (\mathbf{A})_{ij}$ . Let  $\mathbf{D} \in \mathbb{R}^{n \times n}$  be a diagonal matrix with entries  $(\mathbf{D})_{ii} = d_i$ . The Laplacian of  $G$  is defined as  $\mathbf{L} \doteq \mathbf{D} - \mathbf{A}$ , and the normalized Laplacian of  $G$  is defined as  $\mathcal{L} \doteq \mathbf{D}^{-1/2} \mathbf{L} \mathbf{D}^{-1/2}$ . We denote the eigendecompositions of  $\mathcal{L}$  and  $\mathbf{L}$  by  $\mathbf{U} \mathbf{\Lambda} \mathbf{U}^T$  and  $\mathbf{V} \mathbf{\Lambda}_L \mathbf{V}^T$ , respectively. We assume that the eigenvalues, the diagonal entries of  $\mathbf{\Lambda}$  and  $\mathbf{\Lambda}_L$ , are sorted in increasing order. Finally, the volume of a set of vertices  $S \subseteq V$  is  $\text{vol}(S) \doteq \sum_{i \in S} d_i$ .

Let  $\mathbf{f} \in \mathbb{R}^n$  be a signal over the graph vertices, i.e.,  $(\mathbf{f})_i$  is the signal value at vertex  $i$ . The classical Fourier transform can be defined as the transform that diagonalizes the Laplace operator. Similarly, the graph Fourier transform [9] is defined as  $\hat{\mathbf{f}} \doteq \mathbf{V}^T \mathbf{f}$ , where  $\mathbf{V}$  diagonalizes the graph Laplacian. The inverse graph Fourier transform is then simply defined as  $\mathbf{f} \doteq \mathbf{V} \hat{\mathbf{f}}$ .

### 2.1. Local spectral graph theory

The second eigenvalue of the graph Laplacian  $\mathbf{L}$  can be viewed as the solution to

$$\min_{\mathbf{x}} \mathbf{x}^T \mathbf{L} \mathbf{x} \quad \text{s.t.} \quad \mathbf{x}^T \mathbf{D} \mathbf{x} = 1, \quad \mathbf{x}^T \mathbf{D} \mathbf{1} = 0. \quad (3)$$

The optimal solution  $\mathbf{x}^*$  is a generalized eigenvector of  $\mathbf{L}$  with respect to  $\mathbf{D}$  and provides a map from the graph to the real line. This map encodes a measure of similarity (geodesic distance) between graph vertices. This property is exploited for clustering [10] and hashing [11], for example.

In [12] the above problem is modified to incorporate a bias towards a target region (defined by one or more vertices) in the graph. This region is represented as an indicator vector  $\mathbf{s}$ , normalized such that  $\mathbf{s}^T \mathbf{D} \mathbf{1} = 0$  and  $\mathbf{s}^T \mathbf{D} \mathbf{s} = 1$ . More precisely, given a set of nodes  $S \in V$  we define the unit vector  $\mathbf{s} = \text{unit}(S)$  as

$$(\text{unit}(S))_i \doteq \begin{cases} b / \text{vol}(S) & \text{if } i \in S; \\ -b / \text{vol}(V \setminus S) & \text{otherwise,} \end{cases} \quad (4)$$

where  $b = \sqrt{\text{vol}(S) \text{vol}(V \setminus S) / \text{vol}(V)}$ . The modified problem is given by [12]

$$\min_{\mathbf{x}} \mathbf{x}^T \mathbf{L} \mathbf{x} \quad \text{s.t.} \quad \begin{aligned} \mathbf{x}^T \mathbf{D} \mathbf{x} &= 1, \quad \mathbf{x}^T \mathbf{D} \mathbf{1} = 0, \\ \mathbf{x}^T \mathbf{D} \mathbf{s} &\geq \kappa. \end{aligned} \quad (5)$$

The only modification is the addition of the constraint on  $\mathbf{x}^T \mathbf{D} \mathbf{s}$ . This can be interpreted as imposing to the solution a correlation with  $\mathbf{s}$  larger than  $\arccos(\kappa)$ .

Intuitively, as the solution to Problem (3) provides a notion of geodesic distance between graph nodes, the solution to Problem (5) provides a notion of geodesic distance from the seed set to the rest of the vertices. This link will become clear in the following.

**Theorem ([12]).** *Let  $\mathbf{s} \in \mathbb{R}^n$  be a seed vector such that  $\mathbf{s}^T \mathbf{D} \mathbf{1} = 0$  and  $\mathbf{s}^T \mathbf{D} \mathbf{s} = 1$ , and  $\mathbf{s}^T \mathbf{D} \mathbf{v}_2 \neq 0$ , where  $\mathbf{v}_2$  is the second generalized eigenvector of  $\mathbf{L}$  with respect to  $\mathbf{D}$ . In addition, let  $\mathbf{x}^*$  be an optimal solution to Problem (5) with correlation parameter  $\kappa \in [0, 1]$ . Then, there exists some  $\gamma \in (-\infty, (\mathbf{\Lambda})_{2,2})$  and some  $c \in [0, \infty]$  such that*

$$\mathbf{x}^* = c (\mathbf{L} - \gamma \mathbf{D})^+ \mathbf{D} \mathbf{s}. \quad (6)$$

PageRank [13] assigns a numerical weight to each vertex of a graph, assessing its relative importance within the graph; its personalized variant is frequently used to localize the PageRank vector within a subset of the network [8]. The following proposition can be proven using simple algebraic manipulations and the definition of  $\mathbf{L}$ .

**Proposition.** *Let  $c = -\gamma = \frac{1-\alpha}{\alpha}$  in Theorem 6. The vector  $\mathbf{p}$ , defined as  $\mathbf{p} = c (\mathbf{L} - \gamma \mathbf{D})^{-1} \mathbf{D} \mathbf{s}$ , is the solution to the (degree normalized) personalized PageRank (PPR) equation*

$$\mathbf{D} \mathbf{p} = (1 - \alpha)(\mathbf{D} \mathbf{s}) + \alpha \mathbf{A} \mathbf{D}^{-1} (\mathbf{D} \mathbf{p}). \quad (7)$$

Theorem 6 and Equation (7) connect Problem (5) with the personalized PageRank equation [14]. In the field of community detection, PPR vectors are used to find local communities around seed vertices [e.g., 7, 8, 15], where a small but cohesive “seed set” of vertices is expanded to generate its enclosing community (vertices having a stronger relationship to the seed set than to the rest of the graph). In this context, PPR vectors arise as natural units of observation for localized analysis of graphs.

Furthermore, powers of a graph diffusion operator were identified in [6] as natural building blocks to define wavelets on graphs. Notice that the PPR vector is *exactly* equivalent to the recursive application of the diffusion operator  $\mathbf{A} \mathbf{D}^{-1}$  [8], which leads naturally to the notion of geodesic distance.

Given this evidence, we posit that the PPR vector is a fundamental tool to perform a localized spectral analysis of graph signals. This connection is the key observation of this work and drives our definition of a short-graph Fourier transform.

### 2.2. A short-graph Fourier transform

As described in the introduction, we need two elements to define a short-graph Fourier transform: a localization (e.g., classically a translation) and a modulation operators.

**Definition (Localization).** We define the local window at node  $i$  as

$$\mathbf{w}_i \doteq \max(0, \mathbf{x}_i^*) / \|\max(0, \mathbf{x}_i^*)\|_1, \quad (8)$$

where  $\mathbf{x}_i^*$  the solution to Problem (5) with  $\mathbf{s} = \text{unit}(\{i\})$  and the maximum is taken entrywise.

The window  $\mathbf{w}_i$  is defined in terms of its correlation with  $\text{unit}(\{i\})$ , yielding to a simple conceptual interpretation. In this work, we solve Problem (5) using Theorem 6. Given the eigendecomposition of the normalized Laplacian, we have  $\mathbf{L} = \mathbf{D}^{1/2} \mathbf{U} \mathbf{A} \mathbf{U}^T \mathbf{D}^{1/2}$ . Then,

$$\mathbf{x}^* = c \left( \mathbf{D}^{-1/2} \mathbf{U} \right) (\mathbf{A} - \gamma \mathbf{I})^+ \left( \mathbf{D}^{-1/2} \mathbf{U} \right)^T \mathbf{D} \mathbf{s}. \quad (9)$$

Once the eigendecomposition is computed as a preprocessing step, this formula delivers an efficient method for obtaining  $\mathbf{x}^*$ , without any iterations nor matrix inversions (albeit the inversion of the diagonal matrix  $\mathbf{A} - \gamma \mathbf{I}$ ). Interestingly, the spectral localization of  $\mathbf{x}^*$  is determined by the product  $(\mathbf{D}^{-1/2} \mathbf{U})^T \mathbf{D} \mathbf{s} = \mathbf{U}^T \mathbf{D}^{1/2} \mathbf{s}$ , i.e., by the correlation between  $\mathbf{s}$  and each element of the graph Fourier basis.

Since the eigendecomposition of  $\mathcal{L}$  is used in Equation (9), we also use it in our graph modulation operator.

**Definition (Graph modulation).** For  $k \in \{1, 2, \dots, n\}$ , we define the graph modulation operator  $M_k : \mathbb{R}^n \rightarrow \mathbb{R}^n$  by

$$M_k f \doteq \sqrt{\text{vol}(V)} f \circ (\mathbf{D}^{-1/2} \mathbf{U})_{:,k}, \quad (10)$$

where  $\circ$  denotes the entrywise multiplication.

$M_1$  is the identity operator, such as  $\mathcal{T}_0$  is in the classical modulation for temporal signals.

**Definition.** Given the localization and modulation operators, we define the short-graph Fourier transform of a signal  $\mathbf{f} \in \mathbb{R}^n$  at vertex  $i \in V$  and frequency  $k \in \{1, 2, \dots, n\}$  as

$$\text{SGFT}_f(i, k) = \langle \mathbf{f}, M_k \mathbf{w}_i \rangle. \quad (11)$$

The spectrogram of  $\mathbf{f}$  is defined as

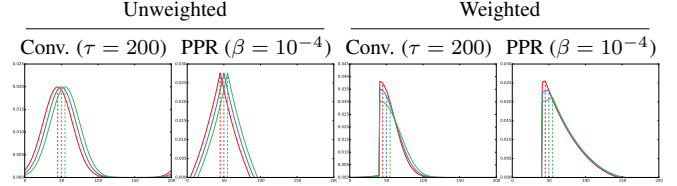
$$\text{spectrogram}_f(i, k) = |\text{SGFT}_f(k, i)|^2. \quad (12)$$

It is not hard to see that Equation (11) reduces to the standard one when the unweighted graph is a Cartesian grid.

**Note.** Shuman et al. provide a different definition for a short-graph Fourier transform [5]. They define the graph modulation operator as  $\widetilde{M}_k f \doteq \sqrt{n} f \circ (\mathbf{V})_{:,k}$ , where  $\circ$  denotes the entrywise multiplication. They also define a convolution operator as  $f * g \doteq \mathbf{V} \left( \widehat{f} \circ \widehat{g} \right)$ , and a translation operator as  $T_i f \doteq \sqrt{n} (f * \delta_i) = \sqrt{n} \mathbf{V} \left( (\mathbf{V}^T)_{:,i} \circ \widehat{f} \right)$ , where  $\delta_i$  is the impulse function at vertex  $i$ . This leads to the (more



(a) The linear graph has 200 vertices, where each vertex is connected to its two neighbors (with periodicity in the edges). In the weighted case, the graph weights for all edges are set to one, excepting edges (40, 41) and (159, 160) (marked in red) which have a weight of  $10^{-3}$ .



(b) We compute localized windows around vertices 45, 50, and 55 (red, blue and green curves, respectively). Note how the convolutional approach [5] fails to properly locate the window in the weighted case, as the window peak does not coincide with the desired vertex.

**Fig. 1. Window localization comparison.** In the unweighted case (in which the graph Fourier and the standard Fourier transforms are equal), convolutional localization [5] works well; however, when the graph weights present a sharp discontinuity, it fails to provide an accurate result. Contrarily, the proposed PPR approach works well in both cases.

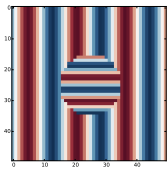
classical) definition  $\text{SGFT}_f(i, k) = \langle \mathbf{f}, \widetilde{M}_k T_i \mathbf{g} \rangle$ . Defining an appropriate window (kernel)  $\mathbf{g}$  is not trivial in the graph space. It is possible to define it in the graph spectral space as  $(\widehat{\mathbf{g}})_k = e^{-\tau \lambda_k}$  and then invert the graph Fourier transform. In this work we do not aim at producing a better method than the one in [5] (although we will exemplify potential localization advantages of our definition). The method here proposed is based on radically different principles, which are of interest by themselves for the spectral study of graph signals.

### 3. EXPERIMENTAL RESULTS

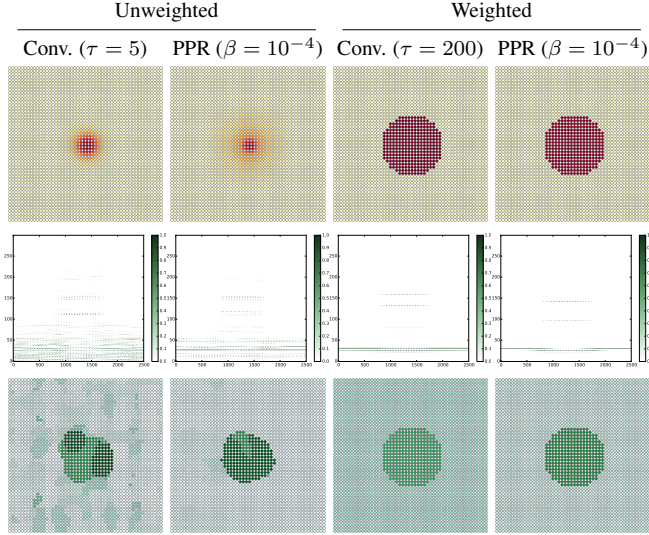
We implemented the proposed short-graph Fourier transform in Python, using the graph-tool library [16]. We make the code publicly available at <https://github.com/marianotepper/sgft>. In all examples we set  $\gamma = (\mathbf{A})_{1,1} - \beta$ , where  $\beta$  takes a particular value in each example.

We begin by examining the localization operator (Equation (8)). For this, we use in Fig. 1 a linear graph, where localization can be easily interpreted and visualized. The main observation is that the window is properly localized when using the proposed approach, while this is not always the case with the convolutional approach.

In the second example, we present results on a 2D grid graph, see Fig. 2. When this graph is unweighted, the graph Fourier transform amounts to the classical 2D Fourier. In the unweighted and weighted cases, the proposed short-graph Fourier transform is able to clearly identify the two different signal regions. Naturally, since the weight discontinuity matches the boundary between both signal regions, the spectrograms of the weighted graph have better spatial and



(a) The graph is a  $50 \times 50$  regular grid, where each vertex is connected to its four neighbors (with periodicity in the edges). The input signal is formed by two sinusoidal waveforms, as shown on the side. In the weighted case, the graph weights for the edges connecting both waveforms are set to  $10^{-5}$  while for the rest of the edges, they are set to one.



(b) **Top row:** Localized windows at the central vertex of the grid. **Center row:** Spectrograms. **Bottom row:** For each vertex, color represents the index of the frequency with maximum magnitude, from light green (low frequencies) to dark green (high frequencies).

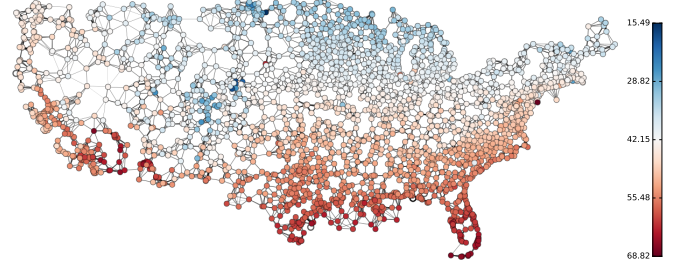
**Fig. 2.** All spectrograms coarsely identify the two sections in the input signal with different patterns (in all cases, we use the first 500 eigenvectors only). In the unweighted case, the proposed method works significantly better than the convolutional approach [5]. In the unweighted and weighted cases, the proposed PPR-based method has better spectral localization, i.e., for each vertex, fewer frequencies are selected.

frequency localizations. The proposed PPR-based spectrogram exhibits better spatial and frequency localizations than the convolutional approach.

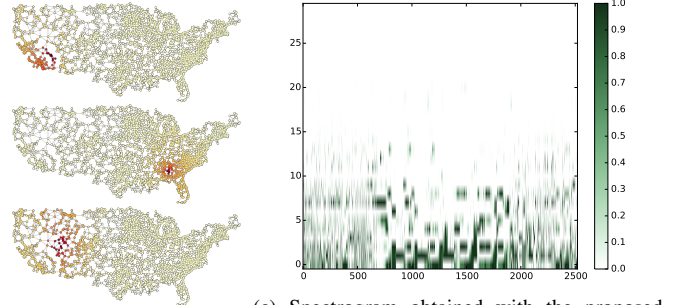
For our last example, we use a real graph comprised of weather stations distributed throughout the US. The input signal is the average temperature in each station in 2014. The localized PPR windows follow nicely the graph topology, being more isotropic or anisotropic, depending on the local graph topology. The spectrogram obtained with the proposed PPR-based method presents clear patterns, which are coherent with the spatial arrangement of the graph vertices.

#### 4. CONCLUSIONS

In this work we presented an extension of the classical short-time Fourier transform (STFT) to signals defined over graphs. We have shown with different examples that this new short-graph Fourier transform can be a valuable tool for extracting information from signals on graphs.

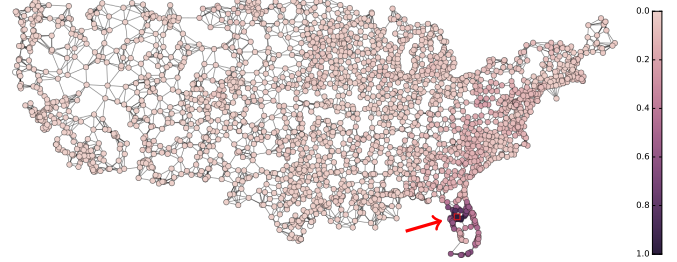


(a) Graph of weather stations in the US, with color representing the average annual temperature in 2014. The graph was built by connecting each station to its 6 spatial nearest neighbors (the corresponding edge weight is the spatial distance between both stations).



(b) Different personalized PageRank windows.

(c) Spectrogram obtained with the proposed technique (for better visualization, we only show the first 30 components).



(d) For the vertex marked with a red square (and a red arrow), we show the correlation between its spectral signature (its column in the spectrogram) and the signatures of the other vertices.

**Fig. 3.** Clear patterns appear in the spectrogram, where nodes with similar spectral signature are localized in spatially coherent areas (e.g., Florida). We use  $\beta = 10^{-3}$  and restrict the computations to the first 500 eigenvectors.

More broadly, we established the connection between local spectral graph theory and localized spectral analysis of graph signals. This is the first work that studies the use of personalized PageRank vectors as fundamental building blocks for local spectral analysis of graph signals.

The STFT becomes ineffective when the signal includes structures having different time-frequency resolution, some very localized in time and others in frequency. Wavelets address this issue by changing the time and frequency resolution. Such as diffusion wavelets extended the definition to the graph domain using powers of a graph diffusion operator, we plan on extending the use of personalized PageRank vectors to produce an alternate definition of wavelets on graphs.

## References

- [1] J. Allen, “Short term spectral analysis, synthesis, and modification by discrete Fourier transform,” *EEE Trans. Acoust.*, vol. 25, no. 3, 1977.
- [2] M. Vetterli and J. Kovačević, *Wavelets and Subband Coding*, Prentice Hall, 1995.
- [3] D. Shuman, S. 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, 2013.
- [4] A. Sandryhaila and J. Moura, “Discrete signal processing on graphs: Frequency analysis,” *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3042–3054, 2014.
- [5] D. Shuman, B. Ricaud, and P. Vandergheynst, “A windowed graph Fourier transform,” in *SSP*, 2012.
- [6] R. Coifman and M. Maggioni, “Diffusion wavelets,” *Appl. Comput. Harmon. Anal.*, vol. 21, no. 1, pp. 53–94, 2006.
- [7] R. Andersen and K. Lang, “Communities from seed sets,” in *WWW*, 2006.
- [8] R. Andersen, F. Chung, and K. Lang, “Local graph partitioning using PageRank vectors,” in *FOCS*, 2006.
- [9] F. Chung, *Spectral Graph Theory*, vol. 92, American Mathematical Soc., 1997.
- [10] J. Shi and J. Malik, “Normalized cuts and image segmentation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 8, pp. 888–905, 2000.
- [11] Y. Weiss, A. Torralba, and R. Fergus, “Spectral hashing,” *NIPS*, 2008.
- [12] M. Mahoney, L. Orecchia, and N. Vishnoi, “A local spectral method for graphs: With applications to improving graph partitions and exploring data graphs locally,” *J. Mach. Learn. Res.*, vol. 13, no. 1, pp. 2339–2365, 2012.
- [13] L. Page, S. Brin, R. Motwani, and T. Winograd, “The PageRank citation ranking: Bringing order to the web,” Tech. Rep. 1999-66, Stanford InfoLab, 1999.
- [14] G. Jeh and J. Widom, “Scaling personalized web search,” in *WWW*, 2003.
- [15] K. Kloster and D. Gleich, “Personalized PageRank solution paths,” *arXiv:1503.00322*, 2015.
- [16] T. Peixoto, “The graph-tool Python library,” *figshare*, 2014.