# Guiding Network Analysis using Graph Slepians: An Illustration for the C. Elegans Connectome

**Article** · August 2017

**3 authors:**

**Some of the authors of this publication are also working on these related projects:**

Vim radiosurgery View project

Partial least squares View project

# Guiding Network Analysis using Graph Slepians: An Illustration for the C. Elegans Connectome

Dimitri Van De Ville[1,2], Robin Demesmaeker[1], Maria Giulia Preti[1,2]

[1] École Polytechnique Fédérale de Lausanne (EPFL), Switzerland
[2] University of Geneva, Switzerland

## ABSTRACT

Spectral approaches of network analysis heavily rely upon the eigendecomposition of the graph Laplacian. For instance, in graph signal processing, the Laplacian eigendecomposition is used to define the graph Fourier transform and then transpose signal processing operations to graphs by implementing them in the spectral domain. Here, we build on recent work that generalized Slepian functions to the graph setting. In particular, graph Slepians are band-limited graph signals with maximal energy concentration in a given subgraph. We show how this approach can be used to guide network analysis; i.e., we propose a visualization that reveals network organization of a subgraph, but while striking a balance with global network structure. These developments are illustrated for the structural connectome of the C. Elegans.

**Keywords:** Graph signal processing, Graph Laplacian, Slepian functions, Graph cut, Laplacian embedding

## 1. INTRODUCTION

In many applications, data can be advantageously modeled and analyzed using graphs, in particular, to reveal emergent, global network properties. In addition, increased availability of large and rich datasets, combined with better computational resources, have been driving the growing interest in network science and graph signal processing, respectively by different communities. One fundamental concept in network analysis is graph partitioning and embedding; e.g., how nodes can be grouped in subgraphs or how they can be represented in a low-dimensional space based on the graph topology.[1] Spectral graph theory relates these problems to the graph Laplacian and its eigendecomposition.[2] For instance, the convex relaxation of the graph cut can be solved by thresholding the Laplacian eigenvector with the smallest non-zero eigenvalue—known as the Fiedler vector.[3] Moreover, the Laplacian eigenvectors also provide a meaningful embedding by mapping nodes onto a line (or higher-dimensional space) that minimizes distance between connected nodes.[4] For graph signal processing, the eigenvectors of the graph Laplacian can be used to define a graph Fourier transform (GFT) and then to define processing steps in the spectral domain;[5] e.g., the spectral graph wavelet transform[6] is such a notable example.

Recently, we have introduced the graph equivalent of Slepian functions[7] that were proposed fifty years ago on regular domains to find a trade-off between temporal and spectral energy concentration.[8,9] The design of graph Slepians is based on an optimization criterion that expresses energy concentration of a band-limited graph signal in a predefined subgraph. In our previous work,[7] we proposed to generalize this criterion such that it reflects a modified Laplacian embedding distance, which basically gives a meaning to eigenvalues as "localized frequencies", instead of as energy concentrations.

Here, we apply the graph Slepian framework to guide network analysis. After a brief introduction on the theory and design of graph Slepians, we show how the information revealed by graph Slepians can be summarized to provide insightful visualizations. We then focus on the structural connectome of the *C. Elegans* organism as an example to guide network analysis.

Further author information: (dimitri.vandeville, robin.demesmaeker, maria.preti)@epfl.ch

## 2. GRAPH SLEPIANS IN A NUTSHELL

### 2.1 Graph Fourier Transform

Let us consider an undirected weighted graph with $N$ nodes. Its connectivity is characterized by an $N \times N$ symmetric adjacency matrix $\mathbf{A}$, where the elements $A_{i,j}$ indicate the edge weights between nodes $i, j = 1, \ldots, N$. We assume the graph is connected and the edge weights are positive. Next to the graph and its topology, we can also consider graph signals that define a mapping from the nodes to a length-$N$ vector associating a value with every node; e.g., the graph topology models the presence of communication links between nodes, while the graph signal represents measures taken at the nodes.

The graph Laplacian is defined as $\mathbf{L}' = \mathbf{D} - \mathbf{A}$, where $\mathbf{D}$ is the diagonal degree matrix with elements $D_{i,i} = \sum_{j=1}^{N} A_{i,j}$. We further consider the normalized graph Laplacian $\mathbf{L} = \mathbf{D}^{-1/2}\mathbf{L}'\mathbf{D}^{-1/2}$ that factors out differences in degree and thus is only reflecting relative connectivity.

Spectral methods for graphs are based on the insight that the eigendecomposition of the graph Laplacian gives access to a GFT.[5] The eigenvectors of $\mathbf{L}$ minimize

$$\lambda = \frac{\mathbf{u}^T\mathbf{L}\mathbf{u}}{\mathbf{u}^T\mathbf{u}} \tag{1}$$

and can be ordered according to increasing eigenvalues $\lambda_1 = 0 \leq \lambda_2 \leq \ldots \leq \lambda_N$. The eigenvectors play the role of basis vectors of the graph spectrum, and the associated eigenvalues the one of frequencies.[2] By stacking the vectors $\mathbf{u}_k$ as the columns of $\mathbf{U}$, which is an $N \times N$ unitary matrix, we obtain the GFT that links a graph signal $\mathbf{s}$ and its spectral coefficient vector $\hat{\mathbf{s}}$ as

$$\mathbf{s} = \mathbf{U}\hat{\mathbf{s}}, \text{ and } \hat{\mathbf{s}} = \mathbf{U}^T\mathbf{s}.$$

### 2.2 Graph Slepians

To generalize Slepians to graphs, we introduce selectivity and bandwidth for the graph signals.[7, 10] First, selectivity can be specified by a subset $\mathcal{S}$ that contains the $N_S$ nodes in which we want the energy concentration to be optimal. We introduce the $N \times N$ diagonal selection matrix $\mathbf{S}$ with elements $S_{k,k} = 0/1$ that indicate the presence of index $k$ in $\mathcal{S}$; so the number of selected indices is trace($\mathbf{S}$) $= N_S$. Second, for the notion of "bandwidth", we propose to restrict the spectrum to the eigenvectors with the $N_W < N$ smallest eigenvalues $\lambda$. We then consider the truncated graph spectrum matrix $\mathbf{U}_W$ of size $N \times N_W$; i.e., it only contains the first $N_W$ columns of $\mathbf{U}$. Consequently, any band-limited graph signal can be represented by

$$\mathbf{g} = \mathbf{U}_W\hat{\mathbf{g}}.$$

Finding the linear combination of eigenvectors within the bandlimit $N_W$ and with maximal energy in $\mathcal{S}$ reverts to optimizing the Rayleigh quotient

$$\mu = \frac{\hat{\mathbf{g}}^T\mathbf{U}_W^T\mathbf{S}\mathbf{U}_W\hat{\mathbf{g}}}{\hat{\mathbf{g}}^T\hat{\mathbf{g}}}, \tag{2}$$

where $\mathbf{C} = \mathbf{U}_W^T\mathbf{S}\mathbf{U}_W$ is the concentration matrix. Since $\mathbf{L}$ is real and symmetric, $\mathbf{U}$ is real as well and we revert to the regular transpose $\cdot^T$. The Slepian vectors are orthonormal over the entire graph as well as orthogonal over the subset $\mathcal{S}$; i.e., we have $\mathbf{g}_k^T\mathbf{g}_l = \delta_{k-l}$ as well as $\mathbf{g}_k^T\mathbf{S}\mathbf{g}_l = \mu_k\delta_{k-l}$, where $\delta$ is the Kronecker delta.

By convention, the eigenvalues of the Slepian eigendecomposition are sorted according to decreasing energy concentration $1 > \mu_1 \geq \mu_2 \geq \ldots > 0$ in $\mathcal{S}$. The eigenvalue spectrum presents a sudden transition between well-localized and poorly-localized eigenvectors.

## 2.3 Graph Slepians Revisited: The Link with Laplacian Embedding

The GFT plays a central role in Laplacian embedding and graph clustering. Specifically, in Laplacian embedding, the aim is to find a mapping from the nodes onto a line such that strongly connected nodes stay as close as possible, which can be expressed as:[4]

$$\arg\min_{\mathbf{g}} \sum_{i,j=1}^{N} A_{i,j}(g_i - g_j)^2 = \arg\min_{\mathbf{g}} \mathbf{g}^T \mathbf{L} \mathbf{g}, \qquad (3)$$

where $\mathbf{g}^T\mathbf{g} = 1$ and $\mathbf{g}^T\mathbf{1} = 0$. The solution is the eigenvector of the Laplacian with the smallest non-zero eigenvalue; i.e., the Fiedler vector.[3] The optimization (3) is also related to the classical graph cut problem that consists of partitioning the graph into clusters of nodes such that the cut size is minimal. In this setting, $g_i = \pm 1$ indicate the labels of the nodes, which can be relaxed to take any real values.

We now use the Slepian design to generalize the criterion (3) as to find a bandlimited solution that is optimized for a selected set of nodes. To reveal this link, we first rewrite the quadratic form (3) as

$$\mathbf{g}^T \mathbf{L} \mathbf{g} = \mathbf{g}^T \mathbf{U} \mathbf{\Lambda} \mathbf{U}^T \mathbf{g} = \hat{\mathbf{g}}^T \mathbf{\Lambda} \hat{\mathbf{g}} = \hat{\mathbf{g}}^T \mathbf{\Lambda}^{1/2} \mathbf{U}^T \mathbf{U} \mathbf{\Lambda}^{1/2} \hat{\mathbf{g}}.$$

Therefore, we propose to obtain a generalized Slepian criterion by introducing in this quadratic form bandwidth and node selection in the spectral and graph domains, respectively:

$$\xi = \frac{\hat{\mathbf{g}}^T \mathbf{\Lambda}_W^{1/2} \overbrace{\mathbf{U}_W^T \mathbf{S} \mathbf{U}_W}^{\mathbf{C}} \mathbf{\Lambda}_W^{1/2} \hat{\mathbf{g}}}{\hat{\mathbf{g}}^T \hat{\mathbf{g}}}, \qquad (4)$$

where $\mathbf{\Lambda}_W$ is the band-limited $N_W \times N_W$ diagonal matrix. Consequently, we propose to solve the eigendecomposition of the modified concentration matrix $\mathbf{C}_{\mathrm{emb}} = \mathbf{\Lambda}_W^{1/2} \mathbf{C} \mathbf{\Lambda}_W^{1/2}$. As before, the Slepian vectors $\mathbf{g}_k = \mathbf{U}_W \hat{\mathbf{g}}_k$ are double orthogonal as we have $\mathbf{g}_k^T \mathbf{g}_l = \delta_{k-l}$ and $\mathbf{g}_k^T \mathbf{S} \mathbf{g}_l = \xi_k \delta_{k-l}$. We store the graph Slepians as columns of the matrix $\mathbf{G}$ of size $N \times N_W$. The concentration value of each graph Slepian can still be recovered using the quadratic form $\mu = \hat{\mathbf{g}}^T \mathbf{C} \hat{\mathbf{g}}$, or, vice versa, the modified embedded distance can be found for the previous design as $\xi = \hat{\mathbf{g}}^{\mathbf{C}}_{\mathrm{emb}} \hat{\mathbf{g}}$.

This demonstrates that Laplacian embedding can be generalized as a Slepian problem with the additional weighting of the Laplacian eigenvalues. It is interesting to note that the eigenvalues $\xi$ of $\mathbf{C}_{\mathrm{emb}}$ represent the modified embedded distance, or, equivalently, a "frequency" that is localized in the subgraph $\mathcal{S}$, whereas the eigenvalues $\mu$ of $\mathbf{C}$ represent the energy concentration in the subgraph.

# 3. AN ILLUSTRATION FOR THE WORM CONNECTOME

The C. Elegans is an intensely studied organism in biology. In particular, the wiring diagram of its 302 neurons has been carefully mapped during a long and effortful study.[11] Here, we use the adjacency matrix that considered 279 somatic neurons and combined connectivity from chemical synapses and gap junctions.[12] We retrieved the type of each neuron (including polymodal status) from the WormAtlas*.

## 3.1 Graph Laplacian Embedding

In their modeling work, Varshney et al[13] studied network properties of the worm connectome using different approaches, including Laplacian embedding. In particular, the topological view generated by mapping on the first two eigenvectors with smallest non-zero eigenvalues already reveals an interesting organization, see Fig. 1A. The horizontal dimension ($\mathbf{u}_2$) reveals two lobes of neurons mostly according to ventral and neck neurons, respectively. The vertical dimension ($\mathbf{u}_3$) show mostly a sensory-inter-motor neuron gradient (from bottom to top), which is very similar to previous results.[13]
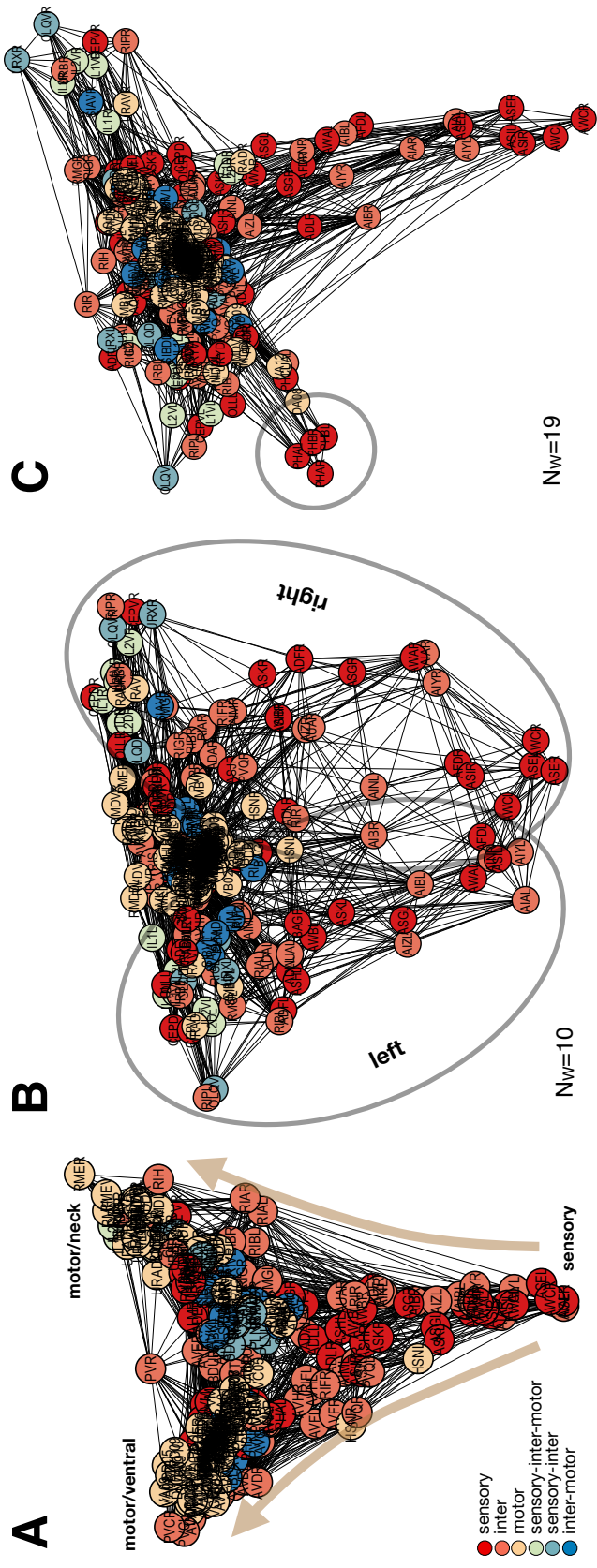
---

*http://www.wormatlas.org/

Figure 1. (A) Laplacian embedding on the two eigenvectors with smallest non-zero eigenvalues. (B) Slepian embedding when selecting sensory neurons and bandwidth $N_W = 10$. (C) Slepian embedding when selecting sensory neurons and bandwidth $N_W = 19$.

## 3.2 Graph Slepian Analysis

To provide a 2-D visualization, we want to maximize the variance captured by the Slepians in the selected subgraph and avoid the selection of the "most important" Slepian vectors. Therefore, we construct the restricted Gram matrix that contains the inner products of the graph Slepians computed on the subgraph $\mathcal{S}$; i.e., we have $\mathbf{K} = \mathbf{G}^T \mathbf{S} \mathbf{G}$. We then diagonalize this matrix using the SVD as $\mathbf{K} = \mathbf{V} \mathbf{\Lambda}^{(\mathbf{K})} \mathbf{V}^T$ and project $\mathbf{G}$ onto this basis to obtain coordinates as $\mathbf{G}\mathbf{V}$, from which we keep the two dimensions with largest eigenvalues in $\mathbf{\Lambda}^{(\mathbf{K})}$. The columns of $\mathbf{V}$ can be considered as a data-driven spectral window that allows maximizing the variance in the visualization for nodes that are part of $\mathcal{S}$.

We specifically pick the subgraph $\mathcal{S}$ as the sensory neurons, or polymodal neurons with a sensory component. A total of 83 neurons are selected, which represents 29.75% of all neurons. For low bandwidth up to $N_W = 5$, the Slepian view on the network is basically identical to the Laplacian embedding. Then, for higher bandwidth such as $N_W = 10$, the energy concentration in the subset of sensory neurons is able to increase and a different structure is revealed, as shown in Fig. 1B. First, many of the motorneurons have moved towards the center of the representation, which means that many of them are not so well connected to sensory neurons. Second, the two lobes of ventral/neck neurons are replaced by two lobes of mainly left/right neurons, including some remaining inter- and motorneurons. This structure is preserved as the bandwidth is further increased, although some additional organization within $\mathcal{S}$ is uncovered, such as the branch of pharynx neurons for $N_W = 19$, see Fig. 1C.

For these results, we used the graph Slepian design based on the modified embedded distance, however, the graph Slepian design based on the energy concentration leads to almost identical results. The reason is that the summarizing method described before relies on the subspace spanned by the Slepians and not on their eigenvalues, and these spaces are very similar. However, for any spectral graph operations (e.g., filtering[5] or wavelet construction[6]), the Slepian design based on the modified embedded distance is advantageously as the eigenvalues can be used as frequencies; for an exemple see.[7]

One well known phenomenon for classical Slepian functions is the phase transition between concentrated and non-concentrated functions, which is given by the Shannon number or time-bandwidth product. For the graph Slepians, the graph-bandwidth product corresponds to $K = N_W N_S / N$. However, as opposed to the conventional case, this is not necessarily equal to the sum of concentration values $\sum_{k=1}^{N_W} \mu_k$ due to the irregularity of the domain. In Fig. 2A, we plot the graph-bandwidth product and the sum of concentration values for the two Slepian designs when selecting the sensory neurons. We notice that for the concentration design, the sum is almost identical to $K$ as the bandwidth varies. For the embedded-distance design, we notice a gap between both measures, although the sum of concentration values still increases approximately linearly. In Fig. 2B, we show a similar behavior when selecting the motor neurons (140 in total).

## 4. CONCLUSION

We have shown how graph Slepians can be designed and used to guide network analysis. The proposed method is flexible as it allows to easily recompute the Slepian vectors for different choices of the subgraph and bandwidth, once the (truncated) eigendecomposition of the graph Laplacian is available. For large(r) graphs, one can rely on efficient large-scale solvers[14] implemented in widely available software libraries such as ARPACK. Future extensions of this approach could include the relationship with graph uncertainty principles,[10,15,16] statistical resampling for graphs,[17] or discovery of hierarchical graph structure[18,19] by gradual refinement of the subgraph. The design could also be extended for directed graphs using recent extensions of spectral decompositions for directed graphs.[20,21]

The Matlab implementation of the graph Slepians is publicly available[†].

## ACKNOWLEDGMENTS

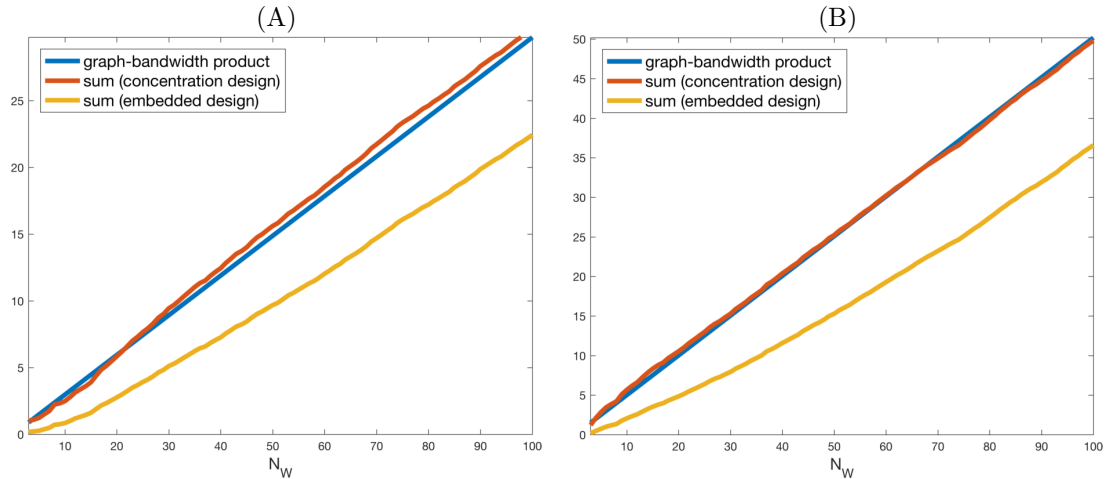[†]https://codeocean.com/2017/05/10/graphslepians/details

Figure 2.   Plots of the graph-bandwidth product and the sum of concentration values for the two Slepian designs and varying bandwidth $N_W$ when selecting (A) sensory or (B) motor neurons.

# REFERENCES

1. U. Luxburg, "A tutorial on spectral clustering," *Statist. Comput.* **17**(4), pp. 395–416, 2007.

2. F. R. K. Chung, *Spectral Graph Theory*, Amer. Math. Soc., Providence, 1997.

3. M. Fiedler, "Laplacian of graphs and algebraic connectivity," *Combinatorics and Graph Theory* **25**, pp. 57–70, 1989.

4. M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural Computation* **15**(6), pp. 1373–1396, 2003.

5. 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* **30**(3), pp. 83–98, 2013.

6. D. K. Hammond, P. Vandergheynst, and R. Gribonval, "Wavelets on graphs via spectral graph theory," *Appl. and Comp. Harm. Anal.* **30**(2), pp. 129–150, 2011.

7. D. Van De Ville, R. Demesmaeker, and M. G. Preti, "When Slepian meets Fiedler: Putting a focus on the graph spectrum," *IEEE Signal Processing Letters* **24**, pp. 1001–1004, 2017.

8. D. Slepian and H. O. Pollak, "Prolate spheroidal wave functions, fourier analysis and uncertainty," *Bell System Technical Journal* **40**(1), pp. 43–63, 1961.

9. D. Slepian, "Prolate spheroidal wave-functions, fourier-analysis, and uncertainty. Discrete case," *Bell System Technical Journal* **57**(5), pp. 1371–1430, 1978.

10. M. Tsitsvero, S. Barbarossa, and P. Di Lorenzo, "Signals on graphs: Uncertainty principle and sampling," *IEEE Transactions on Signal Processing* **64**(18), pp. 4845–4860, 2016.

11. J. G. White, E. Southgate, J. N. Thomson, and S. Brenner, "The structure of the nervous system of the nematode Caenorhabditis elegans," *Phil Trans R Soc Lond B* **314**, pp. 1–340, 1986.

12. B. L. Chen, D. H. Hall, and D. B. Chklovskii, "Wiring optimization can relate neuronal structure and function," *Proceedings of the National Academy of Sciences of the United States of America* **103**, pp. 4723–4728, Mar. 2006.

13. L. R. Varshney, B. L. Chen, E. Paniagua, D. H. Hall, and D. B. Chklovskii, "Structural properties of the Caenorhabditis elegans neuronal network," *PLoS computational biology* **7**(2), p. e1001066, 2011.

14. R. B. Lehoucq and D. C. Sorensen, "Deflation techniques for an implicitly restarted Arnoldi iteration," *SIAM Journal on Matrix Analysis and Applications* **17**(4), pp. 789–821, 1996.

15. A. Agaskar and Y. M. Lu, "A spectral graph uncertainty principle," *IEEE Transactions on Information Theory* **59**(7), pp. 4338–4356, 2013.

16. O. Teke and P. P. Vaidyanathan, "Uncertainty principles and sparse eigenvectors of graphs," *IEEE Transactions on Signal Processing* , in press.

17. E. Pirondini, A. Vybornova, M. Coscia, and D. Van De Ville, "Spectral method for generating surrogate graph signals," *IEEE Signal Processing Letters* **23**, pp. 1275–1278, September 2016.

18. A. Arenas, A. Fernández, and S. Gómez, "Analysis of the structure of complex networks at different resolution levels," *New Journal of Physics* **10**, p. 053039, May 2008.

19. J. Irion and N. Saito, "Hierarchical graph Laplacian eigen transforms," *JSIAM Letters* **6**, pp. 21–24, 2014.

20. A. Sandryhaila and J. M. F. Moura, "Discrete signal processing on graphs," *IEEE Transactions on Signal Processing* **61**, pp. 1644–1656, Mar. 2013.

21. H. N. Mhaskar, "A unified framework for harmonic analysis of functions on directed graphs and changing data," *Applied and Computational Harmonic Analysis* , in press.