A Windowed Digraph Fourier Transform

Rasoul Shafipour[†], Ali Khodabakhsh[‡] and **Gonzalo Mateos**[†]

[†]Dept. of Electrical and Computer Engineering, University of Rochester, Rochester, NY, USA

[‡]Dept. of Electrical and Computer Engineering, University of Texas at Austin, Austin, TX, USA

Objective

Network graph: Weighted digraph $\mathcal{G} = (\mathcal{V}, \mathbf{A})$, where $\mathcal{V}$ is the set of vertices and $\mathbf{A} \in \mathbb{R}^{N\times N}$ is the graph adjacency matrix

Graph signal: A vertex-valued process defined on $\mathcal{G}$, represented by $\mathbf{x} \in \mathbb{R}^N$

Digraph Fourier transform (DGFT): Projection of $\mathbf{x}$ onto a basis $\mathbf{U} \in \mathbb{R}^{N\times N}$ capturing different orthonormal modes of variation with respect to $\mathcal{G}$ $\Rightarrow \tilde{\mathbf{x}} := \mathbf{U}^T\mathbf{x}$

- Fourier transforms obscure the spatial dependency
- **Q:** How to extend classical short-time (aka windowed) Fourier transforms to signals on digraphs?

Preliminiaries

Definition 1 (Total Variation). *For a signal $\mathbf{x}$ on an undirected graph $\mathcal{G} = (\mathcal{V}, \mathbf{A})$ with Laplacian matrix $\mathbf{L} := \text{diag}(\mathbf{A}\mathbf{1}_N) - \mathbf{A}$, the total variation is defined as*

$$TV(\mathbf{x}) = \mathbf{x}^T\mathbf{L}\mathbf{x} = \sum_{i,j=1, j>i}^{N} A_{ij}(x_i - x_j)^2$$

Definition 2 (Directed Variation). *For a signal $\mathbf{x}$ on a digraph $\mathcal{G} = (\mathcal{V}, \mathbf{A})$, we define the directed variation as*

$$DV(\mathbf{x}) = \sum_{i,j=1}^{N} A_{ij}[x_i - x_j]_+^2$$

Spread DGFT design ([1]): Find a set of orthonormal basis signals $\mathbf{U} := [\mathbf{u}_1, \cdots, \mathbf{u}_N] \in \mathbb{R}^{N\times N}$ that

$$\underset{\mathbf{U}}{\text{minimize}} \quad \delta(\mathbf{U}) = \sum_{i=1}^{N-1}(\text{DV}(\mathbf{u}_{i+1}) - \text{DV}(\mathbf{u}_i))^2 \qquad \text{(P1)}$$
$$\text{subject to} \quad \mathbf{U}^T\mathbf{U} = \mathbf{I}, \mathbf{u}_1 = \mathbf{u}_{\min}, \mathbf{u}_N = \mathbf{u}_{\max},$$

where $\mathbf{u}_{\min} = \frac{1}{\sqrt{N}}\mathbf{1}_N$ and $\mathbf{u}_{\max} := \arg\max_{\|\mathbf{u}\|_2=1} \text{DV}(\mathbf{u})$

- $\delta(\mathbf{U})$ is a spectral dispersion measure
- (P1): Nonconvex, but differentiable loss function

  $\Rightarrow$ Find a stationary point via a feasible method

Short-time Fourier transform (STFT):

$$\mathbf{X}(f, i) = \mathcal{F}\{x^{(i)}(t)\} = \mathcal{F}\{x(t)w(t-i)\}$$

- $w(t)$ is an analysis window with a certain band τ ms
- Reveals the time-frequency dependencies of $x(t)$

Problem Statement

- Windowed (localized) graph signal $\mathbf{x}^{(i)}$ for node i:

$$\mathbf{x}^{(i)} := \mathbf{x} \circ \boldsymbol{\phi}_i = \text{diag}(\mathbf{x})\boldsymbol{\phi}_i,$$

  where $\boldsymbol{\phi}_i \in \mathbb{R}^N$ is a windowing signal around $i \in \mathcal{V}$

- Akin to STFT, let the Windowed DGFT (WDGFT) be

$$\tilde{\mathbf{X}} = [\tilde{\mathbf{x}}^{(1)}, \cdots, \tilde{\mathbf{x}}^{(N)}] := [\mathbf{U}^T\mathbf{x}^{(1)}, \cdots, \mathbf{U}^T\mathbf{x}^{(N)}] = \mathbf{U}^T\text{diag}(\mathbf{x})\boldsymbol{\Phi},$$

  where $\boldsymbol{\Phi} = [\boldsymbol{\phi}_1, \cdots, \boldsymbol{\phi}_N] = [\phi_{ji}] \in \mathbb{R}^{N\times N}$

- **Q:** How to learn $\boldsymbol{\Phi} = [\boldsymbol{\phi}_1, \cdots, \boldsymbol{\phi}_N]$ capturing vertex-frequency energy content of graph signals?

Learning windows

- Let $\mathbf{D} \in \mathbb{R}_+^{N\times N}$ store the entries D_{ji} denoting the (directed) proximities or topological structure of the graph from node i to node j

  $\Rightarrow$ Ex: D_{ji} can be the length of the shortest path from node i to node j $\Rightarrow D_{ii} = 0$

- Let $\phi_{ii} = 1$ and ϕ_{ji} be inversely proportional to d_{ji}

  $\Rightarrow$ Ex: An exponential decay $\phi_{ji} = \exp(-\tau_i d_{ji})$

- Learn $\boldsymbol{\tau} =: [\tau_1, \cdots, \tau_N]$ by penalizing the windows for having certain frequency components

  $\Rightarrow$ For $\mathbf{w} = [w_1, \cdots, w_N]^T \in \mathbb{R}_+^N$, solve

$$\min_{\tau \in \mathbb{R}_+^N} \frac{1}{2}\|\text{diag}(\mathbf{w})\mathbf{U}^T\boldsymbol{\Phi}\|_F^2 = \frac{1}{2}\sum_{i=1}^{N}\|\text{diag}(\mathbf{w})\mathbf{U}^T\phi_i\|_2^2,$$

  which decouples into N independent subproblems

$$\min_{\tau_i \geq 0} \quad f_i(\tau_i) = \frac{1}{2}\|\text{diag}(\mathbf{w})\mathbf{U}^T\boldsymbol{\phi}_i\|_2^2 \qquad \text{(P2)}$$

- Solve subproblems in parallel via gradient descent

$$\Rightarrow g = df_i/d\tau_i = -(\mathbf{d}_i \circ \boldsymbol{\phi}_i)^T\mathbf{U}\mathbf{W}\mathbf{U}^T\boldsymbol{\phi}_i,$$

  where $\mathbf{W} = \text{diag}([w_1^2, \cdots, w_N^2])$ and $\mathbf{d}_i = \mathbf{D}(:, i)$

- Follow the update rule $\tau_i^{k+1} := \tau_i^k - \eta^k g(\tau_i^k)$

  $\Rightarrow \{\tau_i\}_i$ converges to a stationary point of $f_i(\tau_i)$

Spectral Vs. Spatial Resolution

Proposition 1. *If $w_1 = 0$, then the optimal value of* (P2) *is zero and is achieved by $\tau_i = 0$, i.e., $f_i(0) = 0$*

$\Rightarrow w_1 = 0$ leads to constant all ones window $\Rightarrow$ WDGFT $\equiv$ DGFT $\Rightarrow$ No resolution in the vertex domain

Proposition 2. *Let $w_1 < w_1'$ be two parameters for penalizing the DC component of the window. If τ is a local minima for* (P2) *with w_1, then the corresponding optimization problem for w_1' has a local (or asymptotic) minima greater than τ*

$\Rightarrow$ Tradeoff between smoothness and locality of windows $\Rightarrow w_1$ can be tuned to achieve a desired resolution

Numerical Results

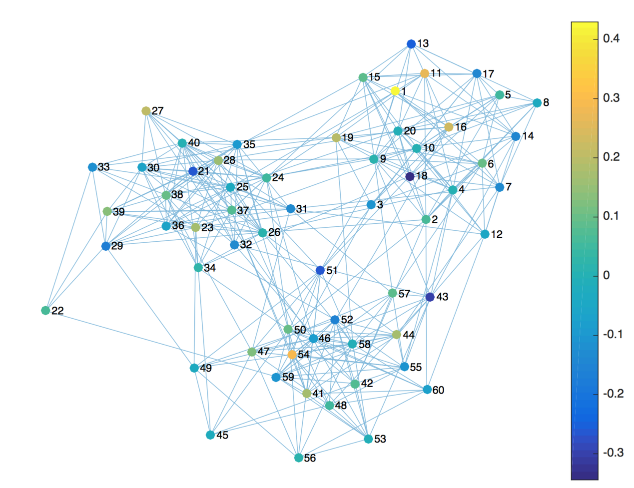

Figure 1: Undirected graph via stochastic block model ($N = 60$, $p_1 = 0.5$, $p_2 = 0.05$, and 3 communities)

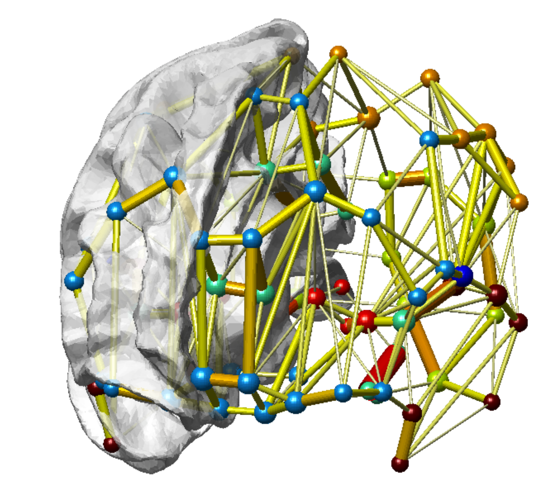

Figure 2: Directed structural brain networks ($N = 47$ and 505 edges, among which 121 links are directed)

- Run Dijkstra algorithm to find a proximity matrix $\mathbf{D}$
- Solve N independent subproblems (P2); see Fig. 3
- Construct signal $\mathbf{x}$ in Fig. 1 by adding $\mathbf{u}_{15}$ restricted to the first 20 nodes, $\mathbf{u}_{30}$ restricted to the middle 20 nodes, and $\mathbf{u}_{45}$ restricted to the last 20 nodes
- Construct $\mathbf{x}$ in the directed case, by adding $\mathbf{u}_{10}$ restricted to 24 highly connected nodes and $\mathbf{u}_{30}$ restricted to the rest
- Obtain spectrograms $|\tilde{\mathbf{X}}| = |\mathbf{U}^T\text{diag}(\mathbf{x})\boldsymbol{\Phi}|$
- For the undirected graph, compare with [2]

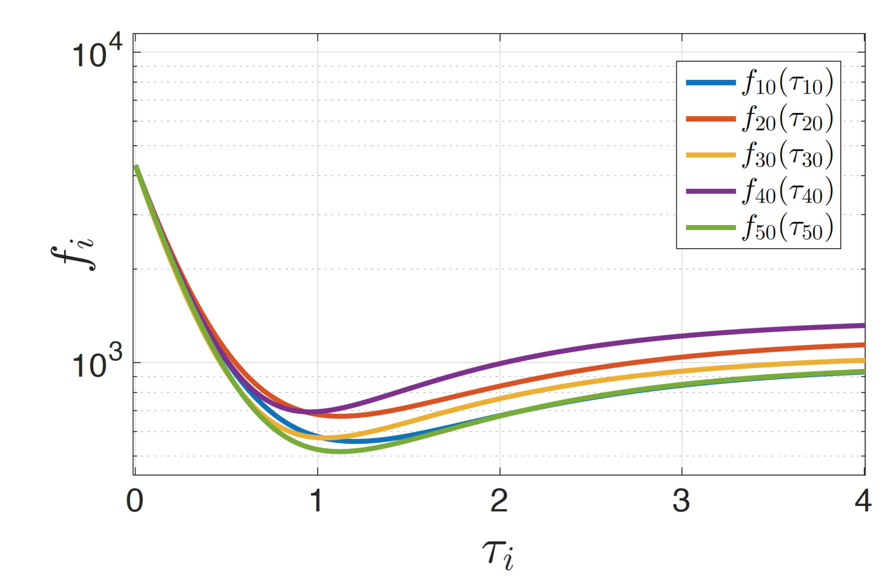

Figure 3: Examples of $f_i(\tau_i)$ in (P2) for different vertices.

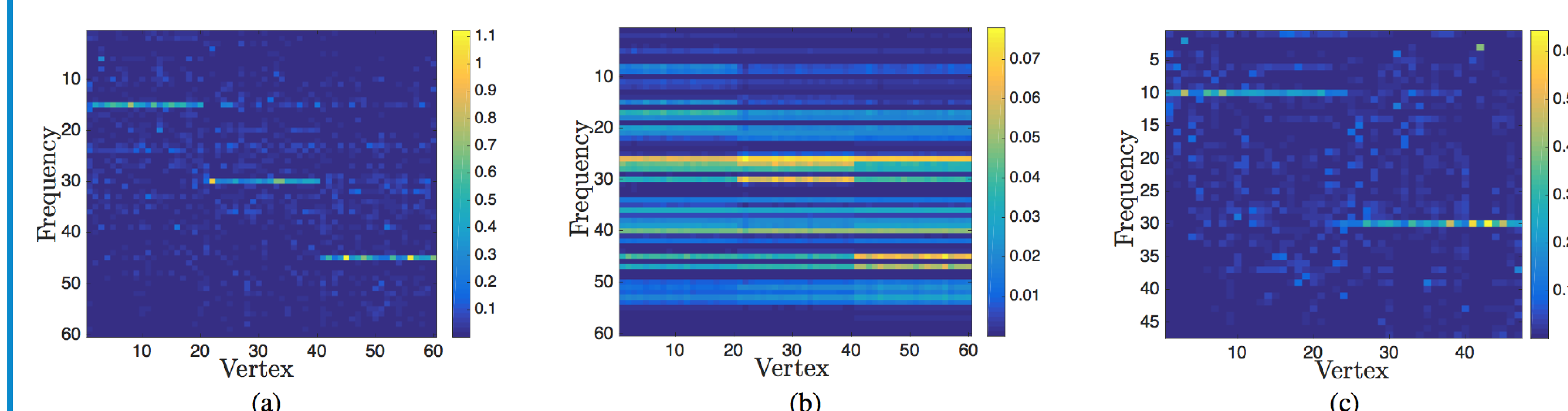

Figure 4: Spectrograms for both undirected and directed examples. (a) Proposed windowed digraph Fourier transform for the graph in Fig. 1 and a signal constructed by three different basis vectors using DGFT (b) Method in [2] for the same graph and a signal constructed by three different eigenvectors of the Laplacian matrix (c) Proposed method for the directed brain graph

References

[1] R. Shafipour, A. Khodabakhsh, G. Mateos, and E. Nikolova. A directed graph Fourier transform with spread frequency components. *IEEE Trans. Signal Process.*, 67(4):946–960, Feb 2019.

[2] D. I. Shuman, B. Ricaud, and P. Vandergheynst. Vertex-frequency analysis on graphs. *Applied and Computational Harmonic Analysis*, 40(2):260–291, 2016.

Contact Information

Web www.ece.rochester.edu/ ~ gmateosb/

Email gmateosb@ece.rochester.edu