Digraph Fourier Transform via Spectral Dispersion Minimization

Gonzalo Mateos
Dept. of Electrical and Computer Engineering
University of Rochester
gmateosb@ece.rochester.edu
http://www.ece.rochester.edu/~gmateosb/

Co-authors: Rasoul Shafipour, Ali Khodabakhsh, and Evdokia Nikolova
Acknowledgment: NSF Award CCF-1750428

Calgary, AB, April 20, 2018
Network as graph $G = (\mathcal{V}, \mathcal{E})$: encode pairwise relationships

Desiderata: Process, analyze and learn from network data [Kolaczyk’09]

Interest here not in $G$ itself, but in data associated with nodes in $\mathcal{V}$

⇒ The object of study is a graph signal

⇒ Ex: Opinion profile, buffer levels, neural activity, epidemic
Graph signal processing and Fourier transform

- Directed graph (digraph) $G$ with adjacency matrix $A$
  $\Rightarrow A_{ij} = \text{Edge weight from node } i \text{ to node } j$

- Define a signal $x \in \mathbb{R}^N$ on top of the graph
  $\Rightarrow x_i = \text{Signal value at node } i$

- Associated with $G$ is the underlying undirected $G^u$
  $\Rightarrow$ Laplacian matrix $L = D - A^u$, eigenvectors $V = [v_1, \cdots, v_N]$

- Graph Signal Processing (GSP): exploit structure in $A$ or $L$ to process $x$

- Graph Fourier Transform (GFT): $\tilde{x} = V^T x$ for undirected graphs
  $\Rightarrow$ Decompose $x$ into different modes of variation
  $\Rightarrow$ Inverse (i)GFT $x = V \tilde{x}$, eigenvectors as frequency atoms
GFT: Motivation and context

- Spectral analysis and filter design [Tremblay et al’17], [Isufi et al’16]

- Promising tool in neuroscience [Huang et al’16]
  ⇒ Graph frequency analyses of fMRI signals

- Noteworthy GFT approaches
  - Eigenvectors of the Laplacian \( L \) [Shuman et al’13]
  - Jordan decomposition of \( A \) [Sandryhaila-Moura’14], [Deri-Moura’17]
  - Lovász extension of the graph cut size [Sardellitti et al’17]
  - Greedy basis selection for spread modes [Shafipour et al’17]
  - Generalized variation operators and inner products [Girault et al’18]

- **Our contribution:** design a novel digraph \((D)\)GFT such that
  - Bases offer notions of frequency and signal variation
  - Frequencies are (approximately) equidistributed in \([0, f_{\text{max}}]\)
  - Bases are orthonormal, so Parseval’s identity holds
Signal variation on digraphs

- **Total variation** of signal \( x \) with respect to \( \mathbf{L} \)

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

\( \Rightarrow \) Smoothness measure on the graph \( \mathcal{G}^u \)

- For Laplacian eigenvectors \( \mathbf{V} = [v_1, \cdots, v_N] \) \( \Rightarrow TV(v_k) = \lambda_k \)

\( \Rightarrow 0 = \lambda_1 < \cdots \leq \lambda_N \) can be viewed as frequencies

- **Def:** Directed variation for signals over digraphs \(([x]_+ = \max(0, x))\)

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

\( \Rightarrow \) Captures signal variation (flow) along directed edges

\( \Rightarrow \) **Consistent**, since \( DV(x) \equiv TV(x) \) for undirected graphs
DGFT with spread frequency components

- **Goal**: find $N$ orthonormal bases capturing different modes of DV on $G$
- Collect the desired bases in a matrix $\mathbf{U} = [\mathbf{u}_1, \cdots, \mathbf{u}_N] \in \mathbb{R}^{N \times N}$
  - $\mathbf{u}_k$ represents the $k$th frequency mode with $f_k := \text{DV}(\mathbf{u}_k)$
- Similar to the DFT, seek $N$ *equidistributed* graph frequencies
  $$f_k = \text{DV}(\mathbf{u}_k) = \frac{k - 1}{N - 1} f_{\max}, \quad k = 1, \ldots, N$$
  - $f_{\max}$ is the maximum DV of a unit-norm graph signal on $G$

- **Q**: Why spread frequencies?
  - Parsimonious representations of slowly-varying signals
  - Interpretability ⇒ better capture low, medium, and high frequencies
  - Aid filter design in the graph spectral domain
Motivation for spread frequencies

**Ex:** Directed variation minimization [Sardellitti et al’17]

\[
\min_U \sum_{i,j=1}^{N} A_{ij} [u_i - u_j] + \\
\text{s.t. } U^T U = I
\]

- **U**\(^*\) is the optimum basis where \( a = \frac{1+\sqrt{5}}{4} \), \( b = \frac{1-\sqrt{5}}{4} \), and \( c = -0.5 \)

- All columns of **U**\(^*\) satisfy \( \text{DV}(u_k^*) = 0 \), \( k = 1, \ldots, 4 \)

  \( \Rightarrow \) Expansion \( x = U^*\tilde{x} \) fails to capture different modes of variation

- **Q:** Can we always find *equidistributed* frequencies in \([0, f_{\text{max}}]\)?
**Challenges: Maximum directed variation**

- **Finding** $f_{\text{max}}$ is in general challenging

  $$u_{\text{max}} = \arg\max_{\|u\|=1} \text{DV}(u) \quad \text{and} \quad f_{\text{max}} := \text{DV}(u_{\text{max}}).$$

- **Let** $v_N$ be the dominant eigenvector of $L$

  $$\Rightarrow \text{Can } 1/2\text{-approximate } f_{\text{max}} \text{ with } \tilde{u}_{\text{max}} = \arg\max_{v \in \{v_N, -v_N\}} \text{DV}(v).$$

- $f_{\text{max}}$ can be obtained analytically for particular classes though

\[ f_{\text{max}} = 2 \max_{i,j} A_{ij} \]

\[ f_{\text{max}} = 2 \max_{i,j} A_{ij} \]

\[ f_{\text{max}} = \lambda_{\text{max}} \]
Equidistributed frequencies may not be feasible. Example: In undirected $G^u$

\[
f_k = \frac{k-1}{N-1} f_{\text{max}}
\]

\[
f_u^{\text{max}} = \lambda_{\text{max}} \quad \text{and} \quad \sum_{k=1}^{N} f_k = \sum_{k=1}^{N} \text{TV}(v_k) = \text{trace}(L)
\]

Idea: Set $u_1 = u_{\text{min}} := \frac{1}{\sqrt{N}} 1_N$ and $u_N = u_{\text{max}}$ and minimize

\[
\delta(U) := \sum_{i=1}^{N-1} [\text{DV}(u_{i+1}) - \text{DV}(u_i)]^2
\]

$\delta(U)$ is the spectral dispersion function

Minimized when the free DV values form an arithmetic sequence
We cast the optimization problem of finding spread frequencies as

$$\min_U \sum_{i=1}^{N-1} [DV(u_{i+1}) - DV(u_i)]^2$$

subject to $U^T U = I$

$u_1 = u_{min}$

$u_N = u_{max}$

- Non-convex, orthogonality-constrained minimization of smooth $\delta(U)$
- Feasible since $u_{max} \perp u_{min}$

Adopt a feasible method in the Stiefel manifold to design the DFGT:

(i) Obtain $f_{max}$ (and $u_{max}$) by minimizing $-DV(u)$ over $\{u \mid u^T u = 1\}$

(ii) Find the orthonormal basis $U$ with minimum spectral dispersion
Feasible method in the Stiefel manifold

- Rewrite the problem of finding orthonormal basis as
  \[
  \min_U \phi(U) := \delta(U) + \frac{\lambda}{2} \left( \|u_1 - u_{\min}\|^2 + \|u_N - u_{\max}\|^2 \right)
  \]
subject to \(U^TU = I_N\)

- Let \(U_k\) be a feasible point at iteration \(k\) and the gradient \(G_k = \nabla \phi(U_k)\)
  \[\Rightarrow\] Skew-symmetric matrix \(B_k := G_k U_k^T - U_k G_k^T\)

- Follow the update rule \(U_{k+1}(\tau) = \left( I + \frac{\tau}{2} B_k \right)^{-1} \left( I - \frac{\tau}{2} B_k \right) U_k\)
  - Cayley transform preserves orthogonality (i.e., \(U_{k+1}^T U_{k+1} = I\))
  - Is a descent path for a proper step size \(\tau\)

**Theorem (Wen et al’13)** The procedure converges to a stationary point of smooth \(\phi(U)\), while generating feasible points at every iteration.
Algorithm

1: **Input:** Adjacency matrix $A$, parameters $\lambda > 0$ and $\epsilon > 0$
2: Find $u_{\text{max}}$ by a similar feasible method and set $u_{\text{min}} = \frac{1}{\sqrt{N}} 1_N$
3: **Initialize** $k = 0$ and orthonormal $U_0 \in \mathbb{R}^{N \times N}$ at random
4: **repeat**
5: Compute gradient $G_k = \nabla \phi(U_k) \in \mathbb{R}^{N \times N}$
6: Form $B_k = G_k U_k^T - U_k G_k^T$
7: Select $\tau_k$ satisfying Armijo-Wolfe conditions
8: Update $U_{k+1}(\tau_k) = (I + \frac{\tau_k}{2} B_k)^{-1}(I - \frac{\tau_k}{2} B_k) U_k$
9: $k \leftarrow k + 1$
10: **until** $\|U_k - U_{k-1}\|_F \leq \epsilon$
11: **Return** $\hat{U} = U_k$

- Overall run-time is $O(N^3)$ per iteration

Additional details in arXiv:1804.03000 [eess.SP]
Compute $\mathbf{U}$ and directed variations using

- Directed Laplacian eigenvectors [Chung’05]
- PAMAL method [Sardellitti et al’17]
- Greedy heuristic [Shafipour et al’17]
- Spectral dispersion minimization

Rescale DV values to $[0, 1]$ and calculate *spectral dispersion* $\delta(\mathbf{U})$

$\Rightarrow 0.256, 0.301, 0.118, \text{ and } 0.076$ respectively

$\Rightarrow$ Confirms the proposed method yields a better frequency spread
Numerical test: US average temperatures

- Consider the graph of the $N = 48$ contiguous United States
  - Connect two states if they share a border
  - Set arc directions from lower to higher latitudes

- Graph signal $x \rightarrow$ Average annual temperature of each state
Numerical test: Denoising US temperatures

- Noisy signal $y = x + n$, with $n \sim \mathcal{N}(0, 10 \times I_N)$
- Define low-pass filter $\tilde{H} = \text{diag}(\tilde{h})$, where $\tilde{h}_i = I \{i \leq 3\}$
- Recover signal via filtering $\hat{x} = U\tilde{H}\tilde{y} = U\tilde{H}U^T y$
  - Compute recovery error $e_f = \frac{\|\hat{x} - x\|}{\|x\|}$
  - Reverse the edge orientations and repeat the experiment

- DGFT basis offers a parsimonious (i.e., bandlimited) signal representation
  - Adequate network model improves the denoising performance
Closing remarks

- Measure of **directed variation** to capture the notion of **frequency** on $\mathcal{G}$

- Find an **orthonormal** set of Fourier bases for signals on digraphs
  - Span a maximal frequency range $[0, f_{\text{max}}]$
  - Frequency modes are as evenly distributed as possible

- Two-step **DGFT** basis design via a feasible method over Stiefel manifold
  i) Find the maximum directed variation $f_{\text{max}}$ over the unit sphere
  ii) Minimize a smooth **spectral dispersion** criterion over $[0, f_{\text{max}}]$
      $\Rightarrow$ Provable convergence guarantees to a stationary point

- **Ongoing work and future directions**
  - Complexity of finding the maximum frequency $f_{\text{max}}$ on a digraph?
    $\Rightarrow$ If NP-hard, what is the best approximation ratio
  - Optimality gap between the local and global optimal dispersions?
Symposium on Graph Signal Processing

Topics of interest

- Graph-signal transforms and filters
- Distributed and non-linear graph SP
- Statistical graph SP
- Prediction and learning for graphs
- Network topology inference
- Recovery of sampled graph signals
- Control of network processes
- Signals in high-order and multiplex graphs
- Neural networks for graph data
- Topological data analysis
- Graph-based image and video processing
- Communications, sensor and power networks
- Neuroscience and other medical fields
- Web, economic and social networks

Paper submission due: June 17, 2018

Organizers:
Gonzalo Mateos (Univ. of Rochester)
Santiago Segarra (MIT)
Sundeep Chepuri (TU Delft)