# A Digraph Fourier Transform with Spread Frequency Components


Gonzalo Mateos

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



GlobalSIP, November 14, 2017

# Co-authors

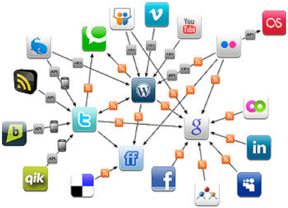
Rasoul Shafipour
University of Rochester

Ali Khodabakhsh
University of Texas at Austin

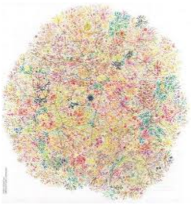
Evdokia Nikolova
University of Texas at Austin

# Network Science analytics

Online social media



Internet



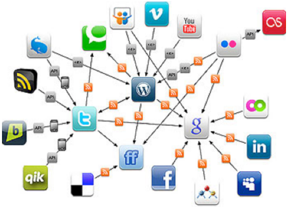
Clean energy and grid analytics

- Network as graph $G = (\mathcal{V}, \mathcal{E})$: encode pairwise relationships
- Desiderata: Process, analyze and learn from network data [Kolaczyk'09]

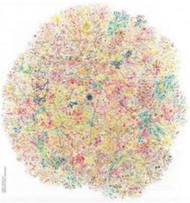
# Network Science analytics



Online social media



Internet



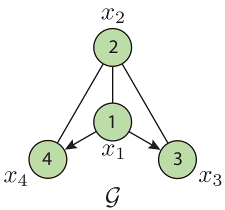
Clean energy and grid analytics

- 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 $\mathcal{V}$

  $\Rightarrow$ The object of study is a graph signal
- Ex: Opinion profile, buffer congestion levels, neural activity, epidemic

# Graph signal processing and Fourier transform

- Directed graph (digraph) $\mathcal{G}$ with adjacency matrix $\mathbf{A}$
  - $\Rightarrow A_{ij} =$ Edge weight from node $i$ to node $j$
- Define a signal $\mathbf{x} \in \mathbb{R}^N$ on top of the graph
  - $\Rightarrow x_i =$ Signal value at node $i$



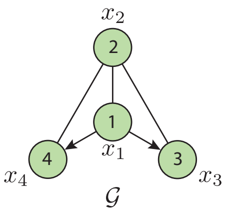
# Graph signal processing and Fourier transform

- Directed graph (digraph) $\mathcal{G}$ with adjacency matrix $\mathbf{A}$

  $\Rightarrow A_{ij} =$ Edge weight from node $i$ to node $j$

- Define a signal $\mathbf{x} \in \mathbb{R}^N$ on top of the graph

  $\Rightarrow x_i =$ Signal value at node $i$




- Associated with $\mathcal{G}$ is the underlying undirected $\mathcal{G}^u$

  $\Rightarrow$ Laplacian marix $\mathbf{L} = \mathbf{D} - \mathbf{A}^u$, eigenvectors $\mathbf{V} = [\mathbf{v}_1, \cdots, \mathbf{v}_N]$

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

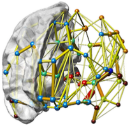
- Graph Fourier Transform (GFT): $\tilde{\mathbf{x}} = \mathbf{V}^T\mathbf{x}$ for undirected graphs

  $\Rightarrow$ Decompose $\mathbf{x}$ into different modes of variation

  $\Rightarrow$ Inverse (i)GFT $\mathbf{x} = \mathbf{V}\tilde{\mathbf{x}}$, eigenvectors as frequency atoms

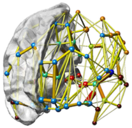
# GFT: Motivation and context

- 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



# GFT: Motivation and context

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



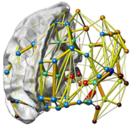
- Promising tool in neuroscience [Huang et al'16]

  $\Rightarrow$ Graph frequency analyses of fMRI signals

- Noteworthy GFT approaches
  - Eigenvectors of the Laplacian $\mathbf{L}$ [Shuman et al'13]
  - Jordan decomposition of $\mathbf{A}$ [Sandryhaila-Moura'14], [Deri-Moura'17]
  - Lovaśz extension of the graph cut size [Sardellitti et al'17]

# 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]
  - $\Rightarrow$ Graph frequency analyses of fMRI signals



- Noteworthy GFT approaches
  - Eigenvectors of the Laplacian $\mathbf{L}$ [Shuman et al'13]
  - Jordan decomposition of $\mathbf{A}$ [Sandryhaila-Moura'14], [Deri-Moura'17]
  - Lovaśz extension of the graph cut size [Sardellitti et al'17]
- **Our contribution:** design a novel digraph (D)GFT such that
  - Bases offer notions of frequency and signal variation
  - Frequencies are (approximately) equidistributed in $[0, f_{\max}]$
  - Bases are orthonormal, so Parseval's identity holds

# Signal variation on digraphs

- Total variation of signal $\mathbf{x}$ with respect to $\mathbf{L}$

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

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

- For Laplacian eigenvectors $\mathbf{V} = [\mathbf{v}_1, \cdots, \mathbf{v}_N] \Rightarrow \mathrm{TV}(\mathbf{v}_k) = \lambda_k$

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

# Signal variation on digraphs

- Total variation of signal $\mathbf{x}$ with respect to $\mathbf{L}$

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

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

- For Laplacian eigenvectors $\mathbf{V} = [\mathbf{v}_1, \cdots, \mathbf{v}_N]$ $\Rightarrow$ $\text{TV}(\mathbf{v}_k) = \lambda_k$

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

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

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

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

$\Rightarrow$ Consistent, since $\text{DV}(\mathbf{x}) \equiv \text{TV}(\mathbf{x})$ for undirected graphs

# DGFT with spread frequeny components

- Goal: find $N$ orthonormal bases capturing different modes of DV on $\mathcal{G}$
- Collect the desired bases in a matrix $\mathbf{U} = [\mathbf{u}_1, \cdots, \mathbf{u}_N] \in \mathbb{R}^{N \times N}$

  $\Rightarrow \mathbf{u}_k$ represents the $k$th frequency component with $f_k := \text{DV}(\mathbf{u}_k)$

# DGFT with spread frequeny components

- Goal: find $N$ orthonormal bases capturing different modes of DV on $\mathcal{G}$
- Collect the desired bases in a matrix $\mathbf{U} = [\mathbf{u}_1, \cdots, \mathbf{u}_N] \in \mathbb{R}^{N \times N}$

  $\Rightarrow$ $\mathbf{u}_k$ represents the $k$th frequency component with $f_k := \mathrm{DV}(\mathbf{u}_k)$
- Similar to the DFT, seek $N$ *equidistributed* graph frequencies

$$f_k = \mathrm{DV}(\mathbf{u}_k) = \frac{k-1}{N-1} f_{\max}, \quad k = 1, \ldots, N$$

  $\Rightarrow$ $f_{\max}$ is the maximum DV of a unit-norm graph signal on $\mathcal{G}$
- **Q**: Why spread frequencies?

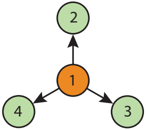
  $\Rightarrow$ To better capture low, medium, and high *frequencies*

  $\Rightarrow$ Aid filter design in the graph spectral domain

# Motivation for spread frequencies

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

$$\min_{\mathbf{U}} \quad \sum_{i,j=1}^{N} A_{ij}[\mathbf{u}_i - \mathbf{u}_j]_+$$
$$\text{s.t.} \quad \mathbf{U}^T\mathbf{U} = \mathbf{I}$$

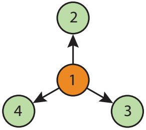



$$\mathbf{U}^{\star} = \begin{bmatrix} 0.5 & c & c & c \\ 0.5 & a & 0 & b \\ 0.5 & b & a & 0 \\ 0.5 & 0 & b & a \end{bmatrix}$$

# Motivation for spread frequencies

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

$$\min_{\mathbf{U}} \quad \sum_{i,j=1}^{N} A_{ij}[\mathbf{u}_i - \mathbf{u}_j]_+$$
$$\text{s.t.} \quad \mathbf{U}^T\mathbf{U} = \mathbf{I}$$




$$\mathbf{U}^{\star} = \begin{bmatrix} 0.5 & c & c & c \\ 0.5 & a & 0 & b \\ 0.5 & b & a & 0 \\ 0.5 & 0 & b & a \end{bmatrix}$$

- $\mathbf{U}^*$ is the optimum basis where $a = \frac{1+\sqrt{5}}{4}$, $b = \frac{1-\sqrt{5}}{4}$, and $c = -0.5$
- All columns of $\mathbf{U}^*$ satisfy $\text{DV}(\mathbf{u}_k^*) = 0,\ k = 1, \ldots, 4$
  - $\Rightarrow$ Expansion $\mathbf{x} = \mathbf{U}^*\tilde{\mathbf{x}}$ fails to capture *different* modes of variation
- **Q:** Can we always find *equidistributed* frequencies?

# Challenges: Maximum directed variation

- Finding $f_{\text{max}}$ is in general challenging
  - Solve the (non-convex) spherically-constrained problem

$$\mathbf{u}_{\text{max}} = \underset{\|\mathbf{u}\|=1}{\text{argmax}}\ \text{DV}(\mathbf{u}) \quad \text{and} \quad f_{\text{max}} := \text{DV}(\mathbf{u}_{\text{max}}).$$

- **Q**: Can we find a basis $\tilde{\mathbf{u}}_{\max}$ with approximate $\tilde{f}_{\max} \approx f_{\text{max}}$?

# Challenges: Maximum directed variation

- Finding $f_{\max}$ is in general challenging
  - Solve the (non-convex) spherically-constrained problem

$$\mathbf{u}_{\max} = \underset{\|\mathbf{u}\|=1}{\operatorname{argmax}}\ \mathrm{DV}(\mathbf{u}) \quad \text{and} \quad f_{\max} := \mathrm{DV}(\mathbf{u}_{\max}).$$

- **Q**: Can we find a basis $\tilde{\mathbf{u}}_{\max}$ with approximate $\tilde{f}_{\max} \approx f_{\max}$?

**Proposition:** For a digraph $\mathcal{G}$, recall $\mathcal{G}^u$ and its Laplacian $\mathbf{L}$. Let $\mathbf{v}_N$ be the dominant eigenvector of $\mathbf{L}$. Then,

$$\tilde{f}_{\max} := \max\left\{\mathrm{DV}(\mathbf{v}_N), \mathrm{DV}(-\mathbf{v}_N)\right\} \geq \frac{f_{\max}}{2}$$

- We can 1/2-approximate $f_{\max}$ with $\tilde{\mathbf{u}}_{\max} = \underset{\mathbf{v}\in\{\mathbf{v}_N, -\mathbf{v}_N\}}{\operatorname{argmax}}\ \mathrm{DV}(\mathbf{v})$

# Challenges: Equidistributed frequencies

- *Equidistributed* $f_k = \frac{k-1}{N-1} f_{\max}$ may not be feasible. Ex: In undirected $\mathcal{G}^u$

$$f^u_{\max} = \lambda_{\max} \quad \& \quad \sum_{k=1}^{N} f_k = \sum_{k=1}^{N} \mathrm{TV}(\mathbf{v}_k) = \mathrm{trace}(\mathbf{L})$$

# Challenges: Equidistributed frequencies

- *Equidistributed* $f_k = \frac{k-1}{N-1} f_{\max}$ may not be feasible. Ex: In undirected $\mathcal{G}^u$

$$f^u_{\max} = \lambda_{\max} \quad \& \quad \sum_{k=1}^{N} f_k = \sum_{k=1}^{N} \text{TV}(\mathbf{v}_k) = \text{trace}(\mathbf{L})$$

- **Idea:** Set $\mathbf{u}_1 = \mathbf{u}_{\min} := \frac{1}{\sqrt{N}} \mathbf{1}_N$ and $\mathbf{u}_N = \tilde{\mathbf{u}}_{\max}$ and minimize

$$\delta(\mathbf{U}) := \sum_{i=1}^{N-1} \left[\text{DV}(\mathbf{u}_{i+1}) - \text{DV}(\mathbf{u}_i)\right]^2$$

$\Rightarrow$ $\delta(\mathbf{U})$ is the *spectral dispersion function*

$\Rightarrow$ $\delta(\mathbf{U})$ is minimized if the *free* DV values form an arithmetic sequence

$\Rightarrow$ Consistent with our design criteria

# Spectral dispersion minimization

- We cast the optimization problem of finding spread frequencies as

$$\begin{aligned} \min_{\mathbf{U}} \quad & \sum_{i=1}^{N-1} \left[\mathrm{DV}(\mathbf{u}_{i+1}) - \mathrm{DV}(\mathbf{u}_i)\right]^2 \\ \text{subject to} \quad & \mathbf{U}^T\mathbf{U} = \mathbf{I} \\ & \mathbf{u}_1 = \mathbf{u}_{\min} \\ & \mathbf{u}_N = \tilde{\mathbf{u}}_{\max} \end{aligned}$$

$\Rightarrow$ Tackle via feasible optimization method in the Stiefel manifold

- Here instead we resort to a simple yet efficient heuristic

# A DGFT construction heuristic

- Use eigenvectors of $\mathbf{L}$, the Laplacian of $\mathcal{G}^u$, to construct $\mathbf{U}$
- Fix $f_1 = 0$ ($\mathbf{u}_1 = \mathbf{u}_{\min}$) and $f_N = \tilde{f}_{\max}$ ($\mathbf{u}_N = \tilde{\mathbf{u}}_{\max}$)
- Let $f_i := \text{DV}(\mathbf{v}_i)$ and $\overline{f}_i := \text{DV}(-\mathbf{v}_i)$, where $\mathbf{v}_i$ is the $i$th eigenvector of $\mathbf{L}$

# A DGFT construction heuristic

- Use eigenvectors of $\mathbf{L}$, the Laplacian of $\mathcal{G}^u$, to construct $\mathbf{U}$
- Fix $f_1 = 0$ ($\mathbf{u}_1 = \mathbf{u}_{\min}$) and $f_N = \tilde{f}_{\max}$ ($\mathbf{u}_N = \tilde{\mathbf{u}}_{\max}$)
- Let $f_i := \text{DV}(\mathbf{v}_i)$ and $\overline{f}_i := \text{DV}(-\mathbf{v}_i)$, where $\mathbf{v}_i$ is the $i$th eigenvector of $\mathbf{L}$
- Define the set of all candidate frequencies as $F := \{f_i, \overline{f}_i : 1 < i < N\}$
  - $\Rightarrow$ Enforce orthonormality: opt exactly one from each pair $\{f_i, \overline{f}_i\}$
- Goal: find the most spread frequency set among the $2^{N-2}$ choices
  - $\Rightarrow$ Exhaustive search intractable even for small graphs
  - $\Rightarrow$ **Q**: Near-optimal solution in polynomial time?

# Frequency selection via supermodular minimization

- For frequency subset $S \subseteq F$, let $s_1 \leq s_2 \leq \ldots \leq s_m$ be the elements of $S$
- Spectral dispersion for $S$ takes the form

$$\delta(S) = \sum_{i=0}^{m} (s_{i+1} - s_i)^2, \quad \text{where } s_0 = 0 \text{ and } s_{m+1} = \tilde{f}_{\max}$$

# Frequency selection via supermodular minimization

- For frequency subset $S \subseteq F$, let $s_1 \leq s_2 \leq \ldots \leq s_m$ be the elements of $S$
- Spectral dispersion for $S$ takes the form

$$\delta(S) = \sum_{i=0}^{m} (s_{i+1} - s_i)^2, \quad \text{where } s_0 = 0 \text{ and } s_{m+1} = \tilde{f}_{\max}$$

- Let $\mathcal{B}$ be the set of all subsets $S \subseteq F$ satisfying $|S \cap \{f_i, \overline{f}_i\}| = 1$, $1 < i < N$

# Frequency selection via supermodular minimization

- For frequency subset $S \subseteq F$, let $s_1 \leq s_2 \leq \ldots \leq s_m$ be the elements of $S$
- Spectral dispersion for $S$ takes the form

$$\delta(S) = \sum_{i=0}^{m} (s_{i+1} - s_i)^2, \quad \text{where } s_0 = 0 \text{ and } s_{m+1} = \tilde{f}_{\max}$$

- Let $\mathcal{B}$ be the set of all subsets $S \subseteq F$ satisfying $|S \cap \{f_i, \overline{f}_i\}| = 1$, $1 < i < N$
- Frequency selection from $F$ boils down to

$$\min_{S} \quad \delta(S), \quad \text{s. t. } S \in \mathcal{B}$$

$\Rightarrow$ Supermodular minimization subject to a matroid basis constraint

$\Rightarrow$ NP-hard and hard to approximate to any factor

# Greedy DGFT bases selection: Algorithm

- Form a non-negative increasing submodular function to be maximized

$$\tilde{\delta}(S) := \tilde{f}_{\max}^2 - \delta(S)$$

- Maximize a monotone submodular function under matroid constraints

  $\Rightarrow$ Can adopt a simple greedy algorithm [Fisher et al'78]

# Greedy DGFT bases selection: Algorithm

- Form a non-negative increasing submodular function to be maximized

$$\tilde{\delta}(S) := \tilde{f}_{\max}^2 - \delta(S)$$

- Maximize a monotone submodular function under matroid constraints

  $\Rightarrow$ Can adopt a simple greedy algorithm [Fisher et al'78]

1: **Input:** Set of candidate frequencies $F$
2: **Initialize** $S = \emptyset$
3: **repeat**
4: $e \leftarrow \text{argmax}_{f \in F} \{\delta(S) - \delta(S \cup \{f\})\}$
5: $S \leftarrow S \cup \{e\}$
6: Delete from $F$ the pair $\{f_i, \overline{f}_i\}$ that $e$ belongs to
7: **until** $F = \emptyset$

# Greedy DGFT bases selection: Guarantees

- **Q:** What about worst-case guarantees for the approximate solution?

**Theorem (Fisher et al'78)** Let $S^*$ be the solution of

$$\min_{S} \quad \delta(S), \quad \text{s. t. } S \in \mathcal{B}$$

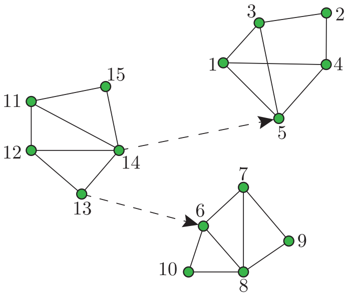
and $S^{\mathrm{g}}$ be the output of the greedy algorithm. Then,

$$\tilde{\delta}(S^{\mathrm{g}}) \geq \frac{1}{2} \times \tilde{\delta}(S^*) \quad \text{or equivalently} \quad \delta(S^{\mathrm{g}}) \leq \frac{1}{2}(\tilde{f}_{\max}^2 + \delta(S^*))$$

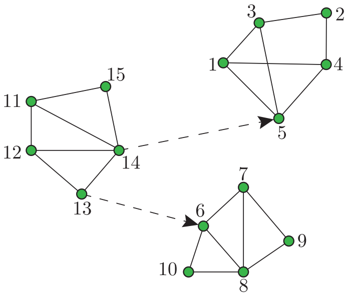
- Usually performs significantly better in practice

# Numerical test: Synthetic graph

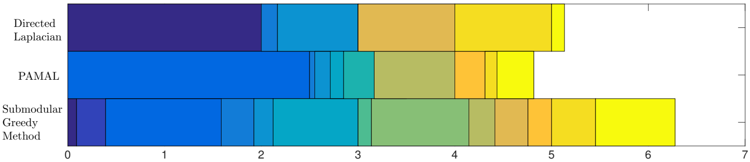
- Digraph studied in [Sardellitti et al'17]
- Compute directed variations using
  - Directed Laplacian eigenvectors [Chung'05]
  - PAMAL method [Sardellitti et'al 17]
  - Proposed submodular greedy algorithm



# Numerical test: Synthetic graph




- Digraph studied in [Sardellitti et al'17]
- Compute directed variations using
  - Directed Laplacian eigenvectors [Chung'05]
  - PAMAL method [Sardellitti et'al 17]
  - Proposed submodular greedy algorithm




- Rescale DV values to the $[0, 1]$ interval and calculate *spectral dispersion*

  $\Rightarrow$ 0.256, 0.301, and **0.118**, respectively

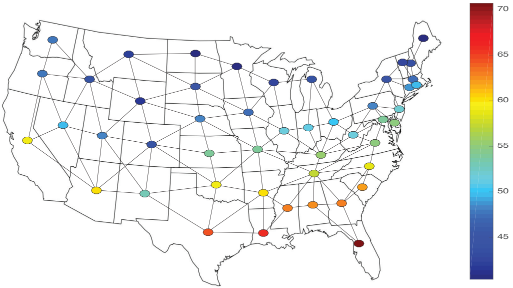
  $\Rightarrow$ Confirms the proposed method yields a better frequency spread

# Numerical test: US average temperatures

- Consider the graph of the contiguous 48 states of the United States
  - $\Rightarrow$ Connect two states if they share a border
  - $\Rightarrow$ Set arc directions from higher to lower latitudes
- Graph signal $\mathbf{x} \rightarrow$ Average annual *temperature* of each state

# Numerical test: US average temperatures

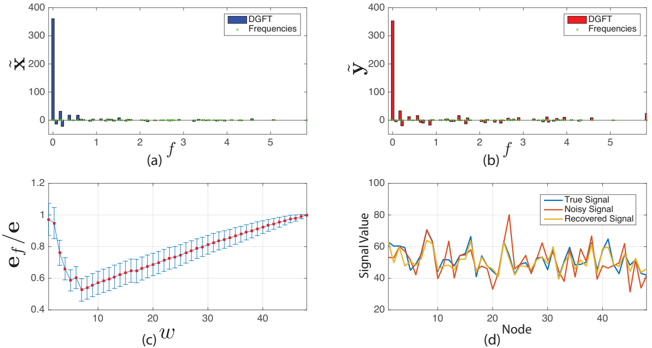
- Consider the graph of the contiguous 48 states of the United States
  - $\Rightarrow$ Connect two states if they share a border
  - $\Rightarrow$ Set arc directions from higher to lower latitudes
- Graph signal $\mathbf{x}$ $\rightarrow$ Average annual *temperature* of each state



# Numerical test: Denoising US average temperatures

- Noisy signal $\mathbf{y} = \mathbf{x} + \mathbf{n}$, with $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, 10\, \mathbf{I}_N)$
- Define low-pass filter $\tilde{\mathbf{H}} = \text{diag}(\tilde{\mathbf{h}})$, where $\tilde{h}_i = \mathbb{I}\{i \leq w\}$
- Recover signal via filtering $\hat{\mathbf{x}} = \mathbf{U}\tilde{\mathbf{H}}\tilde{\mathbf{y}} = \mathbf{U}\tilde{\mathbf{H}}\mathbf{U}^T\mathbf{y}$

  $\Rightarrow$ Compute recovery error $e_f = \frac{\|\hat{\mathbf{x}} - \mathbf{x}\|}{\|\mathbf{x}\|}$




- DGFT basis $\mathbf{U}$ offers parsimonious (i.e., bandlimited) signal representation

# Closing remarks

- Measure of directed variation to capture the notion of frequency on $\mathcal{G}$
- Find an orthonormal set of graph Fourier bases for digraphs
  - Spans a maximal frequency range $[0, f_{\max}]$
  - Frequency components are as evenly distributed as possible
- Two-step DGFT basis construction approach using eigenvectors $\mathbf{V}$ of $\mathbf{L}$

  i) 1/2-approximate $f_{\max}$ with $\max\{\text{DV}(\mathbf{v}_N), \text{DV}(-\mathbf{v}_N)\}$

  ii) Minimize spectral dispersion via a greedy algorithm

# Closing remarks

- Measure of directed variation to capture the notion of frequency on $\mathcal{G}$
- Find an orthonormal set of graph Fourier bases for digraphs
  - Spans a maximal frequency range $[0, f_{\max}]$
  - Frequency components are as evenly distributed as possible
- Two-step DGFT basis construction approach using eigenvectors $\mathbf{V}$ of $\mathbf{L}$

  i) 1/2-approximate $f_{\max}$ with $\max\{\mathrm{DV}(\mathbf{v}_N), \mathrm{DV}(-\mathbf{v}_N)\}$

  ii) Minimize spectral dispersion via a greedy algorithm
- Ongoing work and future directions
  - Complexity of finding the maximum frequency $f_{\max}$ on a digraph?

    $\Rightarrow$ If NP-hard, what is the best approximation ratio
  - Optimality gap between the local and global optimal dispersions?

    $\Rightarrow$ Generalize guarantees to any orthonormal basis