Towards a sampling theorem for signals on arbitrary graphs
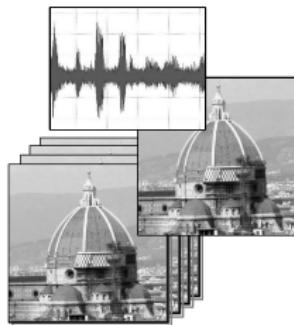
Aamir Anis, Akshay Gadde and Antonio Ortega

University of Southern California
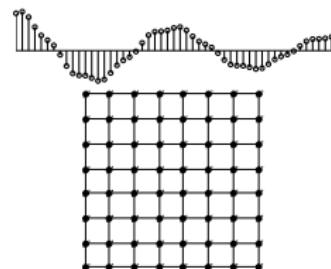
08 May, 2014

USC University of
Southern California

Graph signal processing?

Traditional signal processing applications:
Speech, image, video processing ...
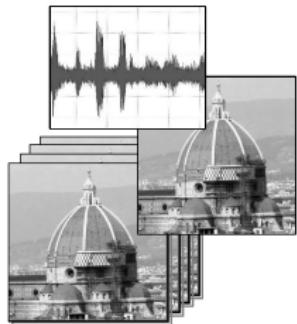
Regular, well-ordered data.
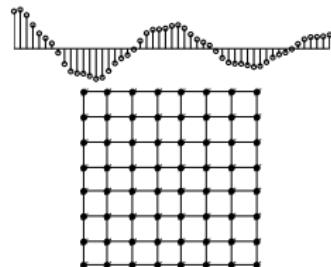
Graph signal processing?

Traditional signal processing applications: Speech, image, video processing ...
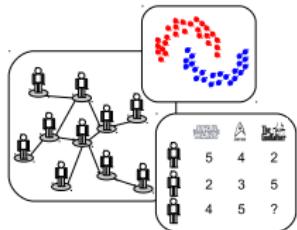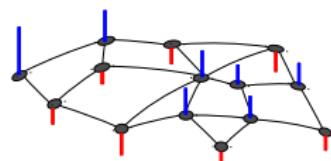
Regular, well-ordered data.

Recent applications: Machine learning, collaborative filtering, social networks ...

Data defined on graphs; Irregular, arbitrarily connected.
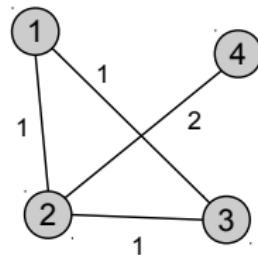
USC University of Southern California

Graph signals

- Graph $G = (\mathcal{V}, \mathcal{E})$: undirected, no self-loops, given by the application.

- Adjacency matrix $\mathbf{A} = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 2 \\ 1 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \end{bmatrix}$.

- Degree matrix $\quad \mathbf{D} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$.

- Laplacian matrix $\mathbf{L} = \mathbf{D} - \mathbf{A}$.

- Symmetric normalized Laplacian
  $\mathcal{L} = \mathbf{D}^{-1/2} \mathbf{L} \mathbf{D}^{-1/2}$.

Graph signals
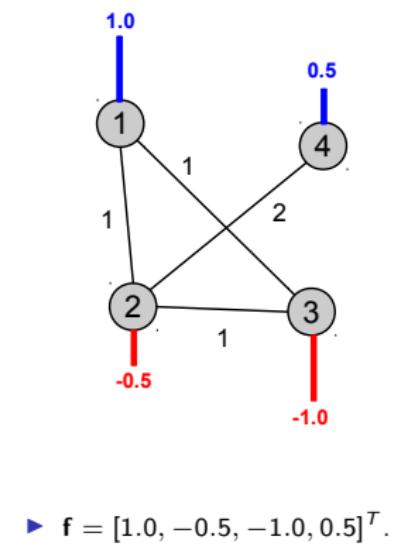
- Graph $G = (\mathcal{V}, \mathcal{E})$: undirected, no self-loops, given by the application.

- Adjacency matrix $\mathbf{A} = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 2 \\ 1 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \end{bmatrix}$.

- Degree matrix $\quad \mathbf{D} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$.

- Laplacian matrix $\mathbf{L} = \mathbf{D} - \mathbf{A}$.

- Symmetric normalized Laplacian $\mathcal{L} = \mathbf{D}^{-1/2} \mathbf{L} \mathbf{D}^{-1/2}$.
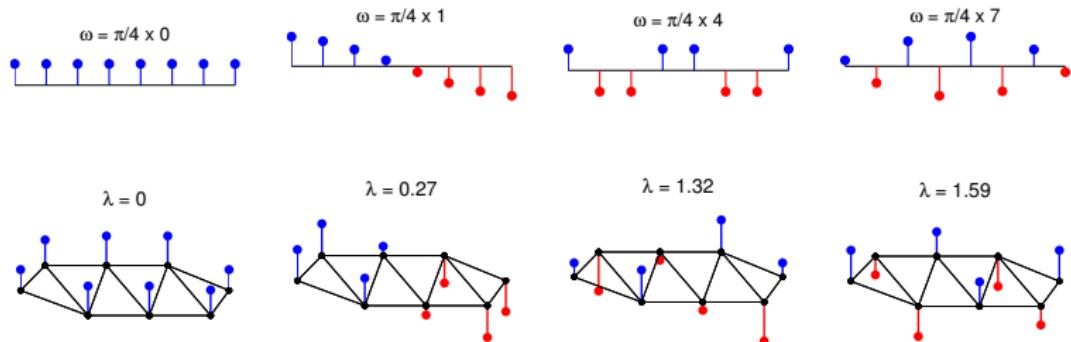
- $\mathbf{f} = [1.0, -0.5, -1.0, 0.5]^T$.

  - Graph signal $f : \mathcal{V} \to \mathbb{R}$, i.e., scalar values on the nodes.
  - Representation: $\mathbf{f} \in \mathbb{R}^N$, where $|\mathcal{V}| = N$.

Graph spectral domain

Spectrum of $\mathcal{L}$ provides frequency interpretation[1]:

- $\lambda_k \in [0, 2]$: *graph frequencies*.
- $\mathbf{u}_k$: *graph Fourier basis*.
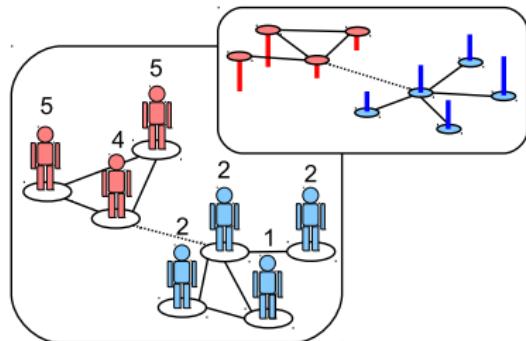
- *Fourier coefficients* of $\mathbf{f}$: $\tilde{\mathbf{f}}(\lambda_i) = \langle \mathbf{f}, \mathbf{u}_i \rangle$.
- *Graph Fourier Transform* (GFT):

$$\tilde{\mathbf{f}} = \mathbf{U}^T \mathbf{f}. \tag{1}$$

[1]Shuman et al., "The Emerging Field of Signal Processing on Graphs: Extending High-Dimensional Data Analysis to Networks and Other Irregular Domains," *IEEE Sig. Proc. Mag.*, 2013.
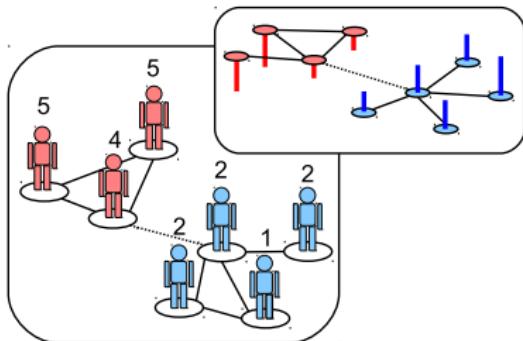
USC University of Southern California

Bandlimited signals

- In many applications, signals of interest are *smooth*.
- Smooth signals $\rightarrow$ lowpass in spectral domain.
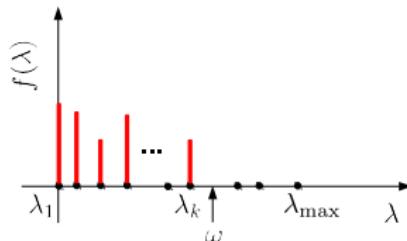
Bandlimited signals

- In many applications, signals of interest are *smooth*.
- Smooth signals → lowpass in spectral domain.

Bandlimited signals in graphs
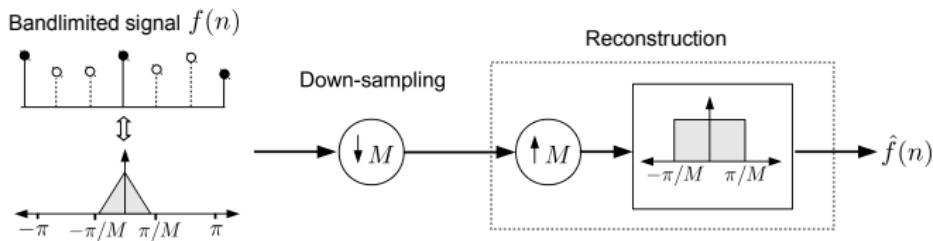
- *ω-bandlimited signal*: GFT has support $[0, \omega]$.
- *Paley-Wiener space $PW_\omega(G)$*: Space of all ω-bandlimited signals.
    - $PW_\omega(G)$ is a subspace of $\mathbb{R}^N$.
    - $\omega_1 \leq \omega_2 \Rightarrow PW_{\omega_1}(G) \subseteq PW_{\omega_2}(G)$.

Sampling

Traditional DSP
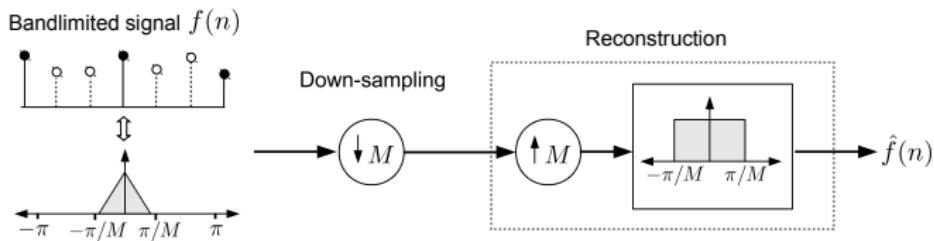
- ▶ Samples dropped in a regular fashion, spectral folding.
- ▶ Cutoff frequency ⇔ sampling rate.

Sampling

Traditional DSP

- ► Samples dropped in a regular fashion, spectral folding.
- ► Cutoff frequency $\Leftrightarrow$ sampling rate.

Bandlimited signal $f(n)$

Down-sampling

Reconstruction

$\downarrow M$

$\uparrow M$

$-\pi/M \quad \pi/M$

$\hat{f}(n)$

$-\pi \quad -\pi/M \quad \pi/M \quad \pi$
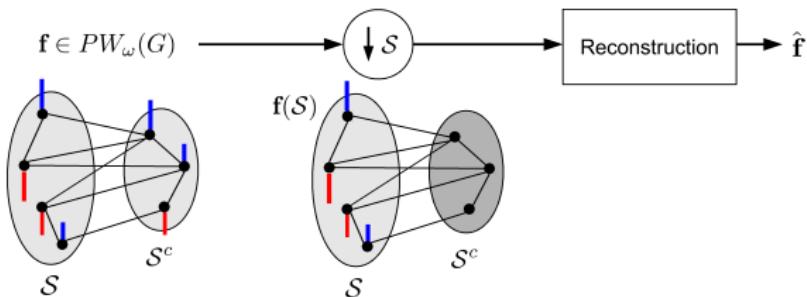
Graph signals

- ► *Irregular*, arbitrarily connected; dropping every *other* sample?
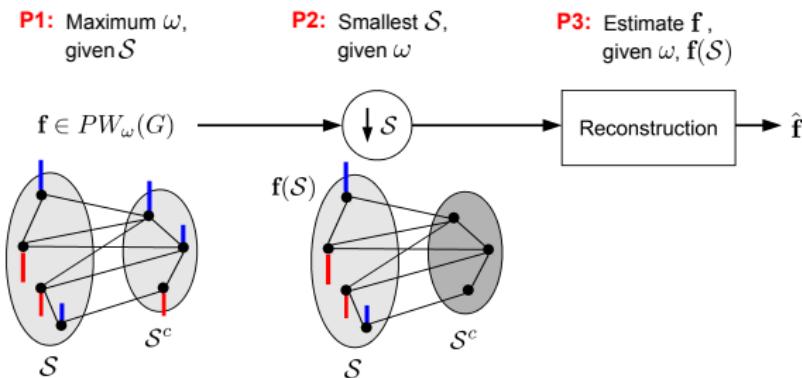
Sampling graph signals

- Input signal: $\mathbf{f} \in PW_\omega(G)$.
- Sampling set: $\mathcal{S} \subset \mathcal{V}$, unknown set: $\mathcal{S}^c$.
- Sampled signal: $\mathbf{f}(\mathcal{S}) \in \mathbb{R}^{|\mathcal{S}|}$.

Sampling graph signals

- Input signal: $\mathbf{f} \in PW_\omega(G)$.
- Sampling set: $\mathcal{S} \subset \mathcal{V}$, unknown set: $\mathcal{S}^c$.
- Sampled signal: $\mathbf{f}(\mathcal{S}) \in \mathbb{R}^{|\mathcal{S}|}$.
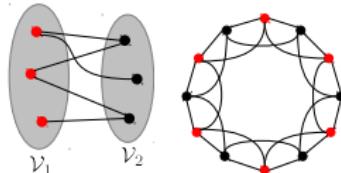
Sampling theorem for graph signals:

P1: Given $\mathcal{S}$, *maximum ω*?

P2: Given ω, *smallest set $\mathcal{S}$*?

P3: Given ω and $\mathbf{f}(\mathcal{S})$, how to recover $\mathbf{f}$?

Contributions

Previous results: Special graphs

- *Bipartite graphs*[2]: Samples on one partition; $\omega < 1$.
- *Circulant graphs*[3]: Regularly-spaced samples.

Previous results: Arbitrary graphs

P1	Pesenson, 2008	*Sufficient* condition for recovery.
	Narang et al., 2013a	Estimate of cutoff frequency, not tight.
P2	Shuman et al., 2013	Heuristic downsampling by 2, no explicit optimization.
P3	Pesenson et al., 2009	Variational splines.
	Narang et al., 2013a	Least-squares approach, requires eigen-decomposition.
	Narang et al., 2013b	Localized iterative graph filtering.

---
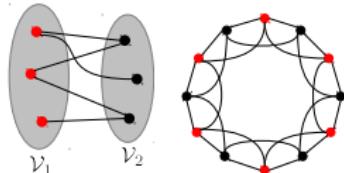
[2] Narang et al., "Perfect reconstruction two-channel wavelet filterbanks for graph structured data," *IEEE Trans. Sig. Proc.*, 2012.

[3] Ekambaram et al., "Multiresolution Graph Signal Processing via Circulant Structures," *IEEE DSP/SPE Workshop*, 2013.

Contributions

Previous results: Special graphs

- *Bipartite graphs*[2]: Samples on one partition; $\omega < 1$.
- *Circulant graphs*[3]: Regularly-spaced samples.

Previous results: Arbitrary graphs

P1	Pesenson, 2008	*Sufficient* condition for recovery.
	Narang et al., 2013a	Estimate of cutoff frequency, not tight.
P2	Shuman et al., 2013	Heuristic downsampling by 2, no explicit optimization.
	Pesenson et al., 2009	Variational splines.
P3	Narang et al., 2013a	Least-squares approach, requires eigen-decomposition.
	Narang et al., 2013b	Localized iterative graph filtering.

Contributions of this work

- Sampling theorem: Necessary and sufficient condition.
- Improved estimate of cutoff in P1.
- Formulation of P2, greedy algorithm.
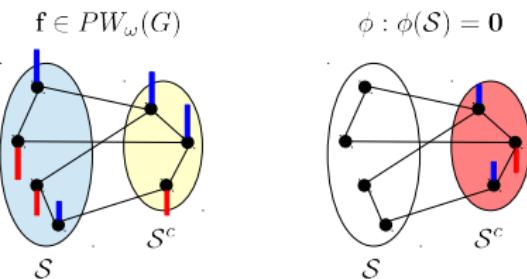
---

[2] Narang et al., "Perfect reconstruction two-channel wavelet filterbanks for graph structured data," *IEEE Trans. Sig. Proc.*, 2012.

[3] Ekambaram et al., "Multiresolution Graph Signal Processing via Circulant Structures," *IEEE DSP/SPE Workshop*, 2013.

USC University of Southern California

Necessary and sufficient condition

$\mathbf{f} \in PW_\omega(G)$

$\mathcal{S}$

$\mathcal{S}^c$

Necessary and sufficient condition

$\mathbf{f} \in PW_\omega(G)$ $\phi : \phi(\mathcal{S}) = \mathbf{0}$

Necessary and sufficient condition

Necessary and sufficient condition

$\mathbf{f} \in PW_\omega(G)$ $\phi : \phi(\mathcal{S}) = \mathbf{0}$ $\mathbf{g} : \mathbf{g}(\mathcal{S}) = \mathbf{f}(\mathcal{S})$

- If $\phi \in PW_\omega(G)$, then $\mathbf{g} = \mathbf{f} + \phi \in PW_\omega(G)$.
- $\mathbf{f} \neq \mathbf{g}$ and $\mathbf{f}(\mathcal{S}) = \mathbf{g}(\mathcal{S}) \Rightarrow$ *trouble!*

Necessary and sufficient condition

$\mathbf{f} \in PW_\omega(G)$ $\phi : \phi(\mathcal{S}) = \mathbf{0}$ $\mathbf{g} : \mathbf{g}(\mathcal{S}) = \mathbf{f}(\mathcal{S})$
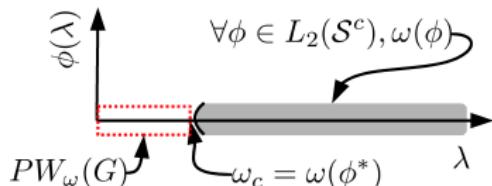
- If $\phi \in PW_\omega(G)$, then $\mathbf{g} = \mathbf{f} + \phi \in PW_\omega(G)$.
- $\mathbf{f} \neq \mathbf{g}$ and $\mathbf{f}(\mathcal{S}) = \mathbf{g}(\mathcal{S}) \Rightarrow$ *trouble!*

Lemma

Let $L_2(\mathcal{S}^c) = \{\phi : \phi(\mathcal{S}) = \mathbf{0}\}$. All signals $\mathbf{f} \in PW_\omega(G)$ can be perfectly recovered from $\mathcal{S}$ if and only if $PW_\omega(G) \cap L_2(\mathcal{S}^c) = \{\mathbf{0}\}$.

USC University of Southern California

Sampling theorem

Choose ω to be less than bandwidth of *smoothest* signal ϕ^* in $L_2(\mathcal{S}^c)$.

Theorem (Sampling theorem)

Let $\omega(\mathbf{f})$ denote the bandwidth of $\mathbf{f}$. For a sampling set $\mathcal{S} \subset \mathcal{V}$, all signals $\mathbf{f} \in PW_\omega(G)$ can be perfectly recovered from their samples $\mathbf{f}(\mathcal{S})$ if and only if

$$\omega < \inf_{\phi \in L_2(\mathcal{S}^c)} \omega(\phi) \triangleq \omega_c(\mathcal{S})$$

We call $\omega_c(\mathcal{S})$ the true cutoff frequency.

The cutoff frequency depends on the size of $\mathcal{S}$ and topologies of G and $\mathcal{S}$.
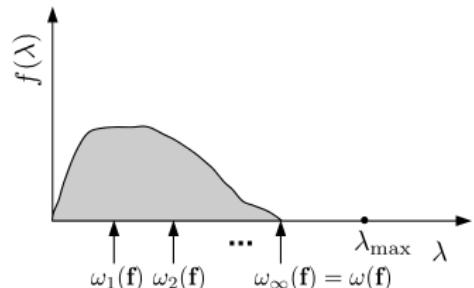
P1: Computing the cutoff frequency

Definition (Spectral moments)

For any signal $\mathbf{f} \neq \mathbf{0}$, $k \in \mathbb{Z}^+$, we define its kth spectral moment as
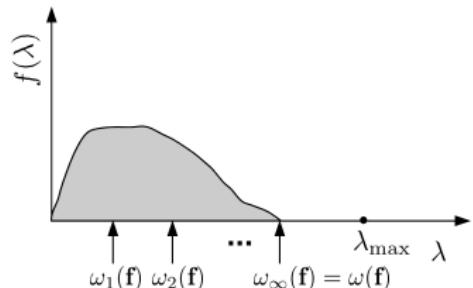
$$\omega_k(\mathbf{f}) \triangleq \left(\frac{\mathbf{f}^T \mathcal{L}^k \mathbf{f}}{\mathbf{f}^T \mathbf{f}} \right)^{1/k}$$

- *Monotonicity:* $\forall \mathbf{f}, k_1 < k_2 \Rightarrow \omega_{k_1}(\mathbf{f}) \leq \omega_{k_2}(\mathbf{f})$.
- *Convergence:* $\lim_{k \to \infty} \omega_k(\mathbf{f}) = \omega(\mathbf{f})$.

P1: Computing the cutoff frequency

Definition (Spectral moments)

For any signal $\mathbf{f} \neq \mathbf{0}$, $k \in \mathbb{Z}^+$, we define its kth spectral moment as

$$\omega_k(\mathbf{f}) \triangleq \left(\frac{\mathbf{f}^T \mathcal{L}^k \mathbf{f}}{\mathbf{f}^T \mathbf{f}} \right)^{1/k}$$

- *Monotonicity:* $\forall \mathbf{f}, k_1 < k_2 \Rightarrow \omega_{k_1}(\mathbf{f}) \leq \omega_{k_2}(\mathbf{f})$.
- *Convergence:* $\lim_{k \to \infty} \omega_k(\mathbf{f}) = \omega(\mathbf{f})$.

For large k, spectral moment $\approx$ bandwidth. Estimate cutoff by
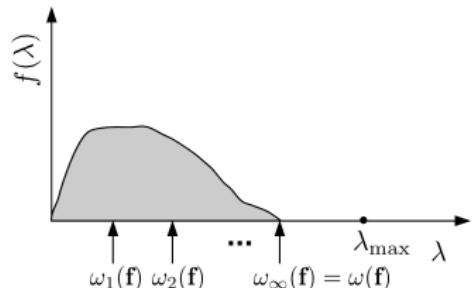
$$\Omega_k(\mathcal{S}) \triangleq \min_{\phi \in L_2(\mathcal{S}^c)} \omega_k(\phi) = \min_{\phi : \phi(\mathcal{S}) = \mathbf{0}} \left(\frac{\phi^T \mathcal{L}^k \phi}{\phi^T \phi} \right)^{1/k}$$

P1: Computing the cutoff frequency

Definition (Spectral moments)

For any signal $\mathbf{f} \neq \mathbf{0}$, $k \in \mathbb{Z}^+$, we define its kth spectral moment as

$$\omega_k(\mathbf{f}) \triangleq \left(\frac{\mathbf{f}^T \mathcal{L}^k \mathbf{f}}{\mathbf{f}^T \mathbf{f}} \right)^{1/k}$$

- *Monotonicity:* $\forall \mathbf{f}, k_1 < k_2 \Rightarrow \omega_{k_1}(\mathbf{f}) \leq \omega_{k_2}(\mathbf{f})$.
- *Convergence:* $\lim_{k \to \infty} \omega_k(\mathbf{f}) = \omega(\mathbf{f})$.

For large k, spectral moment $\approx$ bandwidth. Estimate cutoff by

$$\Omega_k(\mathcal{S}) \triangleq \min_{\phi \in L_2(\mathcal{S}^c)} \omega_k(\phi) = \min_{\phi:\phi(\mathcal{S})=\mathbf{0}} \left(\frac{\phi^T \mathcal{L}^k \phi}{\phi^T \phi} \right)^{1/k} = \left(\min_{\psi} \underbrace{\frac{\psi^T (\mathcal{L}^k)_{\mathcal{S}^c} \psi}{\psi^T \psi}}_{\text{Rayleigh quotient}} \right)^{1/k}$$

P1: Computing the cutoff frequency

Definition (Spectral moments)

For any signal $\mathbf{f} \neq \mathbf{0}$, $k \in \mathbb{Z}^+$, we define its kth spectral moment as

$$\omega_k(\mathbf{f}) \triangleq \left(\frac{\mathbf{f}^T \mathcal{L}^k \mathbf{f}}{\mathbf{f}^T \mathbf{f}} \right)^{1/k}$$

- Monotonicity: $\forall \mathbf{f}$, $k_1 < k_2 \Rightarrow \omega_{k_1}(\mathbf{f}) \leq \omega_{k_2}(\mathbf{f})$.
- Convergence: $\lim_{k \to \infty} \omega_k(\mathbf{f}) = \omega(\mathbf{f})$.

For large k, spectral moment $\approx$ bandwidth. Estimate cutoff by

$$\Omega_k(\mathcal{S}) \triangleq \min_{\phi \in L_2(\mathcal{S}^c)} \omega_k(\phi) = \min_{\phi : \phi(\mathcal{S}) = \mathbf{0}} \left(\frac{\phi^T \mathcal{L}^k \phi}{\phi^T \phi} \right)^{1/k} = \left(\min_{\psi} \underbrace{\frac{\psi^T (\mathcal{L}^k)_{\mathcal{S}^c} \psi}{\psi^T \psi}}_{\text{Rayleigh quotient}} \right)^{1/k}$$

Let $\{\sigma_{1,k}, \psi_{1,k}\} \rightarrow$ smallest eigen-pair of $(\mathcal{L}^k)_{\mathcal{S}^c}$. Cutoff frequency and smoothest signal can be estimated as:

$$\Omega_k(\mathcal{S}) = (\sigma_{1,k})^{1/k},$$
$$\phi_k^* = \psi_{1,k}, \quad \phi_k^* = \mathbf{0}.$$

USC University of Southern California
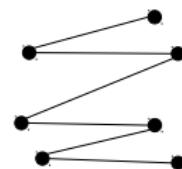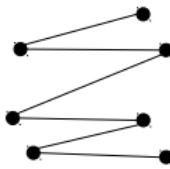
Cutoff frequency $\equiv$ bandwidth of smoothest signal ϕ^* in $L_2(\mathcal{S}^c)$.

---

[4] Shuman et al., " A framework for multiscale transforms on graphs," arXiv:1308.4942, 2013.

USC University of Southern California

Cutoff frequency $\equiv$ bandwidth of smoothest signal ϕ^* in $L_2(\mathcal{S}^c)$.

USC University of
Southern California

From P1 to P2

Cutoff frequency $\equiv$ bandwidth of smoothest signal ϕ^* in $L_2(\mathcal{S}^c)$.

Which choice of $\mathcal{S}$ leads to a higher cutoff frequency?

[4] Shuman et al., " A framework for multiscale transforms on graphs," arXiv:1308.4942, 2013.

From P1 to P2

Cutoff frequency $\equiv$ bandwidth of smoothest signal ϕ^* in $L_2(\mathcal{S}^c)$.

Which choice of $\mathcal{S}$ leads to a higher cutoff frequency?
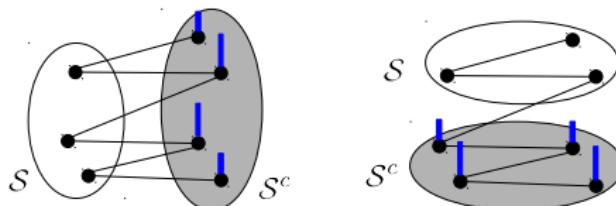
- ▶ Compare $\phi^* \in L_2(\mathcal{S}^c)$ for both graphs.
- ▶ More cross-links $\Rightarrow$ higher variation $\Rightarrow$ higher bandwidth.

Choosing best sampling set for given ω: Max-cut[4]?

[4] Shuman et al., " A framework for multiscale transforms on graphs," arXiv:1308.4942, 2013.

USC University of Southern California

Cutoff frequency $\equiv$ bandwidth of smoothest signal ϕ^* in $L_2(\mathcal{S}^c)$.

Which choice of $\mathcal{S}$ leads to a higher cutoff frequency?

- ▸ Compare $\phi^* \in L_2(\mathcal{S}^c)$ for both graphs.
- ▸ More cross-links $\Rightarrow$ higher variation $\Rightarrow$ higher bandwidth.

Choosing best sampling set for given ω: Max-cut[4]? ... *No.*

- ▸ Roughly divides $\mathcal{V}$ by 2; $\mathcal{S}$ for a given target ω?.

---

[4]Shuman et al., " A framework for multiscale transforms on graphs," arXiv:1308.4942, 2013.

USC University of Southern California

P2: Smallest sampling set

Given ω_c, *smallest set* $\mathcal{S}_{\mathrm{opt}}$ for perfect recovery?

USC University of
Southern California

P2: Smallest sampling set

Given ω_c, *smallest set* $\mathcal{S}_{\mathrm{opt}}$ for perfect recovery?

Formulation

Relax the true cutoff $\omega_c(\mathcal{S})$ by $\Omega_k(\mathcal{S})$, and solve the combinatorial problem:

$$\text{Minimize } |\mathcal{S}| \text{ subject to } \Omega_k(\mathcal{S}) \geq \omega_c$$

USC University of Southern California

P2: Smallest sampling set

Given ω_c, *smallest set $\mathcal{S}_{\mathrm{opt}}$ for perfect recovery?*

Formulation

Relax the true cutoff $\omega_c(\mathcal{S})$ by $\Omega_k(\mathcal{S})$, and solve the combinatorial problem:

$$\text{Minimize } |\mathcal{S}| \text{ subject to } \Omega_k(\mathcal{S}) \geq \omega_c$$

Greedy Approach to get an estimate of $\mathcal{S}_{\mathrm{opt}}$:

- Start with $\mathcal{S} = \{\emptyset\}$.
- Add nodes to $\mathcal{S}$ (from $\mathcal{S}^c$) one-by-one that ensure maximum increase in $\Omega_k(\mathcal{S})$ at each step.

USC University of Southern California

P2: Greedy Approach

Given $\mathcal{S}$, taking which node from $\mathcal{S}^c$ leads to maximum increase in $\Omega_k(\mathcal{S})$?

P2: Greedy Approach

Given $\mathcal{S}$, taking which node from $\mathcal{S}^c$ leads to maximum increase in $\Omega_k(\mathcal{S})$?

▶ *Constrained optimization → regularization*

$$(\Omega_k(\mathcal{S}))^k = \min_{\phi \in L_2(\mathcal{S}^c)} \left(\frac{\phi^T \mathcal{L}^k \phi}{\phi^T \phi} \right) \approx \min_{\mathbf{x}} \left(\frac{\mathbf{x}^T \mathcal{L}^k \mathbf{x}}{\mathbf{x}^T \mathbf{x}} + \alpha \frac{\mathbf{x}^T \text{diag}(\mathbf{1}_\mathcal{S}) \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \right),$$

where, $\mathbf{1}_\mathcal{S} : \mathcal{V} \to \{0, 1\}$ is the indicator function of $\mathcal{S}$.
Minimizer $\approx \phi_k^*$, i.e., estimate of smoothest signal of $L_2(\mathcal{S}^c)$.

USC University of
Southern California

P2: Greedy Approach

Given $\mathcal{S}$, taking which node from $\mathcal{S}^c$ leads to maximum increase in $\Omega_k(\mathcal{S})$?

▶ *Constrained optimization → regularization*

$$(\Omega_k(\mathcal{S}))^k = \min_{\phi \in L_2(\mathcal{S}^c)} \left(\frac{\phi^T \boldsymbol{\mathcal{L}}^k \phi}{\phi^T \phi} \right) \approx \min_{\mathbf{x}} \left(\frac{\mathbf{x}^T \boldsymbol{\mathcal{L}}^k \mathbf{x}}{\mathbf{x}^T \mathbf{x}} + \alpha \frac{\mathbf{x}^T \operatorname{diag}(\mathbf{1}_\mathcal{S}) \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \right),$$

where, $\mathbf{1}_\mathcal{S} : \mathcal{V} \to \{0, 1\}$ is the indicator function of $\mathcal{S}$.
Minimizer $\approx \phi_k^*$, i.e., estimate of smoothest signal of $L_2(\mathcal{S}^c)$.

▶ *Binary relaxation of penalization with* $\mathbf{t} \in \mathbb{R}^N$.

$$(\Omega_k(\mathcal{S}))^k \approx \min_{\mathbf{x}} \left. \left(\frac{\mathbf{x}^T \boldsymbol{\mathcal{L}}^k \mathbf{x}}{\mathbf{x}^T \mathbf{x}} + \alpha \frac{\mathbf{x}^T \operatorname{diag}(\mathbf{t}) \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \right) \right|_{\mathbf{t}=\mathbf{1}_\mathcal{S}} = \lambda_k^\alpha(\mathbf{t})|_{\mathbf{t}=\mathbf{1}_\mathcal{S}}.$$

USC University of
Southern California

P2: Greedy Approach

Given $\mathcal{S}$, taking which node from $\mathcal{S}^c$ leads to maximum increase in $\Omega_k(\mathcal{S})$?

▶ *Constrained optimization* → *regularization*

$$(\Omega_k(\mathcal{S}))^k = \min_{\phi \in L_2(\mathcal{S}^c)} \left(\frac{\phi^T \boldsymbol{\mathcal{L}}^k \phi}{\phi^T \phi} \right) \approx \min_{\mathbf{x}} \left(\frac{\mathbf{x}^T \boldsymbol{\mathcal{L}}^k \mathbf{x}}{\mathbf{x}^T \mathbf{x}} + \alpha \frac{\mathbf{x}^T \mathrm{diag}(\mathbf{1}_{\mathcal{S}}) \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \right),$$

where, $\mathbf{1}_{\mathcal{S}} : \mathcal{V} \to \{0, 1\}$ is the indicator function of $\mathcal{S}$.
Minimizer $\approx \phi_k^*$, i.e., estimate of smoothest signal of $L_2(\mathcal{S}^c)$.

▶ *Binary relaxation of penalization* with $\mathbf{t} \in \mathbb{R}^N$.

$$(\Omega_k(\mathcal{S}))^k \approx \min_{\mathbf{x}} \left(\frac{\mathbf{x}^T \boldsymbol{\mathcal{L}}^k \mathbf{x}}{\mathbf{x}^T \mathbf{x}} + \alpha \frac{\mathbf{x}^T \mathrm{diag}(\mathbf{t}) \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \right) \bigg|_{\mathbf{t} = \mathbf{1}_{\mathcal{S}}} = \lambda_k^\alpha(\mathbf{t}) |_{\mathbf{t} = \mathbf{1}_{\mathcal{S}}} .$$

▶ *Gradient* with respect to $\mathbf{t}$

$$\frac{d\lambda_k^\alpha(\mathbf{t})}{d\mathbf{t}(i)} = \alpha \left(x^*(i) \right)^2 \quad \Rightarrow \quad \frac{d\lambda_k^\alpha(\mathbf{t})}{d\mathbf{t}(i)} \bigg|_{\mathbf{t} = \mathbf{1}_{\mathcal{S}}} \approx \alpha \left(\phi_k^*(i) \right)^2 .$$

USC University of
Southern California

P2: Greedy Approach

Given $\mathcal{S}$, taking which node from $\mathcal{S}^c$ leads to maximum increase in $\Omega_k(\mathcal{S})$?

▶ *Constrained optimization* → *regularization*

$$(\Omega_k(\mathcal{S}))^k = \min_{\phi \in L_2(\mathcal{S}^c)} \left(\frac{\phi^T \mathcal{L}^k \phi}{\phi^T \phi} \right) \approx \min_{\mathbf{x}} \left(\frac{\mathbf{x}^T \mathcal{L}^k \mathbf{x}}{\mathbf{x}^T \mathbf{x}} + \alpha \frac{\mathbf{x}^T \mathrm{diag}(\mathbf{1}_\mathcal{S}) \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \right),$$

where, $\mathbf{1}_\mathcal{S} : \mathcal{V} \to \{0, 1\}$ is the indicator function of $\mathcal{S}$.
Minimizer $\approx \phi_k^*$, i.e., estimate of smoothest signal of $L_2(\mathcal{S}^c)$.

▶ *Binary relaxation of penalization with* $\mathbf{t} \in \mathbb{R}^N$.

$$(\Omega_k(\mathcal{S}))^k \approx \min_{\mathbf{x}} \left(\frac{\mathbf{x}^T \mathcal{L}^k \mathbf{x}}{\mathbf{x}^T \mathbf{x}} + \alpha \frac{\mathbf{x}^T \mathrm{diag}(\mathbf{t}) \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \right)\Bigg|_{\mathbf{t}=\mathbf{1}_\mathcal{S}} = \lambda_k^\alpha(\mathbf{t})|_{\mathbf{t}=\mathbf{1}_\mathcal{S}}.$$

▶ *Gradient* with respect to $\mathbf{t}$

$$\frac{d\lambda_k^\alpha(\mathbf{t})}{d\mathbf{t}(i)} = \alpha \left(x^*(i) \right)^2 \quad \Rightarrow \quad \frac{d\lambda_k^\alpha(\mathbf{t})}{d\mathbf{t}(i)}\Bigg|_{\mathbf{t}=\mathbf{1}_\mathcal{S}} \approx \alpha \left(\phi_k^*(i) \right)^2.$$

▶ *Heuristic*: Given $\mathcal{S}$, pick node $i = \mathrm{argmax}_j (\phi_k^*(j))^2$, i.e., node on which smoothest signal has maximum energy.

USC University of Southern California

P2: Smallest sampling set selection algorithm

Algorithm 1 Greedy heuristic for estimating $\mathcal{S}_{\text{opt}}$

Input: $G = \{\mathcal{V}, E\}$, $\mathcal{L}$, target bandwidth ω_c, some $k \in \mathbb{Z}^+$
Initialize: $\mathcal{S} = \{\emptyset\}$, $\omega = 0$.

1: **while** $\omega \leq \omega_c$ **do**
2:      For $\mathcal{S}$, compute cutoff estimate $\Omega_k(\mathcal{S})$ and corresponding smoothest signal $\phi_k^* \in L_2(\mathcal{S}^c)$.
3:      $\omega \leftarrow \Omega_k(\mathcal{S})$, $v \leftarrow \text{argmax}_i(\phi_k^*(i))^2$.
4:      $\mathcal{S} \leftarrow \mathcal{S} \cup v$.
5: **end while**
6: $\mathcal{S}_{\text{est}} \leftarrow \mathcal{S}$.

USC University of
Southern California

Which node is removed from $\mathcal{S}^c$?

USC University of
Southern California

P2: Intuition

Which node is removed from $\mathcal{S}^c$?

▶ For simplicity, consider $\Omega_k(\mathcal{S})$ for $k = 1$.

$$\min_{\phi \in L_2(\mathcal{S}^c)} \frac{\phi^T \mathcal{L} \phi}{\phi^T \phi} = \min_{\substack{||\psi||=1 \\ \psi(\mathcal{S})=0}} \left(\underbrace{\sum_{j \in \mathcal{S}^c} \overbrace{\frac{\sum_{i \in \mathcal{S}} w_{ij}}{d_j}}^{c_j: \text{"cross degree"}} \psi(j)^2}_{\text{Variation across } \mathcal{S} \text{ and } \mathcal{S}^c} + \underbrace{\sum_{i,j \in \mathcal{S}^c} w_{ij} \left(\frac{\psi(i)}{\sqrt{d_i}} - \frac{\psi(j)}{\sqrt{d_j}} \right)^2}_{\text{Variation within } \mathcal{S}^c} \right)$$

▶ Minimizer ϕ^* expected to have maximum energy on node least connected to $\mathcal{S}$.

▶ Send that node to $\mathcal{S}$.

Nodes in $\mathcal{S}^c$ must be *well-connected* to $\mathcal{S} \rightarrow$ Minimize weak links between $\mathcal{S}$ and $\mathcal{S}^c$.

P2: Intuition

Which node is removed from $\mathcal{S}^c$?

▶ For simplicity, consider $\Omega_k(\mathcal{S})$ for $k = 1$.

$$\min_{\phi \in L_2(\mathcal{S}^c)} \frac{\phi^T \mathcal{L} \phi}{\phi^T \phi} = \min_{\substack{||\psi||=1 \\ \psi(\mathcal{S})=\mathbf{0}}} \left(\underbrace{\sum_{j \in \mathcal{S}^c} \overbrace{\frac{\sum_{i \in \mathcal{S}} w_{ij}}{d_j}}^{c_j : \text{"cross degree"}} \psi(j)^2}_{\text{Variation across } \mathcal{S} \text{ and } \mathcal{S}^c} + \underbrace{\sum_{i,j \in \mathcal{S}^c} w_{ij} \left(\frac{\psi(i)}{\sqrt{d_i}} - \frac{\psi(j)}{\sqrt{d_j}} \right)^2}_{\text{Variation within } \mathcal{S}^c} \right)$$

▶ Minimizer ϕ^* expected to have maximum energy on node least connected to $\mathcal{S}$.

▶ Send that node to $\mathcal{S}$.

Nodes in $\mathcal{S}^c$ must be *well-connected* to $\mathcal{S} \rightarrow$ Minimize weak links between $\mathcal{S}$ and $\mathcal{S}^c$.

▶ Connectivity within $\mathcal{S}^c$ is also taken into account.

▶ For $k > 1$, higher order (*global*) connectivity considered.

USC University of Southern California

Experiments

Four graphs with $N = 100$ used in experiments:

G_1: Bipartite graph
- $\mathcal{V} = \mathcal{V}_1 \cup \mathcal{V}_2$, $|\mathcal{V}_1| = 40$, numbered 1 to 40, $|\mathcal{V}_2| = 60$, numbered 41 to 60.
- Random weights $w_{ij} \sim U(0, 1)$.
- Sparsified using 6 nearest neighbors.

G_2: Regular circulant graph
- Number of neighbors = 8.
- Weights $\propto 1/\text{distance}$ (assuming evenly spaced nodes on a circle).

G_3: *Erdős-Rényi* random graph
- Connection probability $p = 0.2$.
- Weights $w_{ij} \sim U(0, 1)$ if i and j are connected.

G_4: *Watts-Strogatz* 'small-world' model
- Unweighted, underlying circulant graph with 8 neighbors.
- Rewiring probability $\beta = 0.1$.

USC University of Southern California

- Improved cutoff estimate compared to earlier result[5].
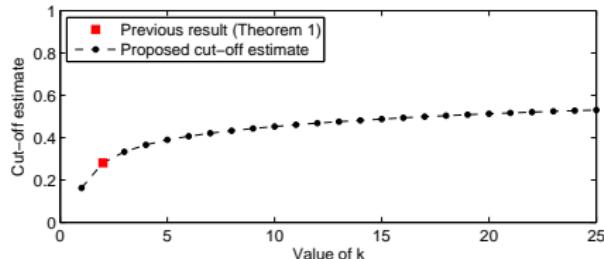
Figure: Behavior of cut-off estimate $\Omega_k(\mathcal{S})$ with k for G_1.

---

[5] Narang et al., "Signal processing techniques for interpolation in graph structured data," *ICASSP*, 2013.

USC University of Southern California

- Improved cutoff estimate compared to earlier result[5].
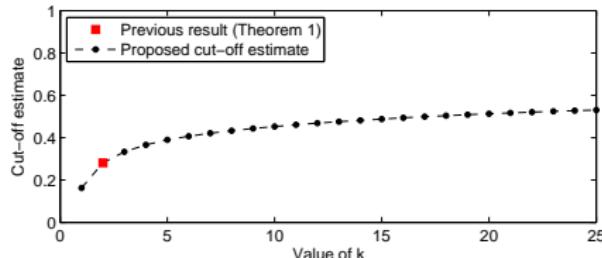
Figure: Behavior of cut-off estimate $\Omega_k(\mathcal{S})$ with k for G_1.

- Increasing trend in estimated cutoff vs. k for all test graphs.
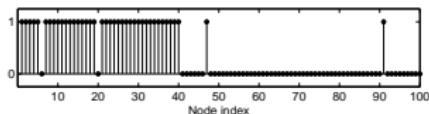
Table: Cut-off frequency estimates for a given subset of nodes.

| Graph	$|\mathcal{S}|$	Earlier result[10] ($k = 2$)	Proposed method		
			$k = 6$	$k = 12$	$k = 18$
G_1	25	0.2815	0.4070	0.4684	0.5040
G_2	40	0.1236	0.3077	0.4696	0.5427
G_3	25	0.4643	0.6292	0.6756	0.7029
G_4	25	0.2313	0.4716	0.6469	0.7106
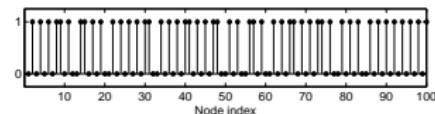
[5] Narang et al., "Signal processing techniques for interpolation in graph structured data," *ICASSP*, 2013.

▶ Sampling set conforms to known results for bipartite and circulant graphs.
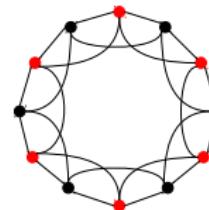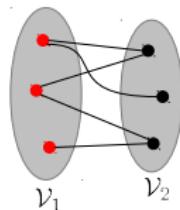
(a)

(b)

Figure: Indicator functions of $\mathcal{S}_{\text{opt}}$ with $k = 18$ for (a) bipartite graph G_1, $|\mathcal{S}_{\text{opt}}| = |\mathcal{V}_1| = 40$ (b) circulant graph G_2, $|\mathcal{S}_{\text{opt}}| = N/2 = 50$.

USC University of Southern California

Experiment: P2

K_c: number of eigenvalues of $\mathcal{L}$ below ω_c.

- Size of estimated sampling set $\mathcal{S}_{\text{est}}$ close to theoretical limit K_c.

Table: Estimated sampling set $\mathcal{S}_{\text{est}}$ for bandwidth $\omega_c = 0.8$.

Graph	K_c	$\vert\mathcal{S}_{\text{est}}\vert$ using Algorithm 1		
		$k = 6$	$k = 12$	$k = 18$
G_1	36	42	41	41
G_2	41	49	46	45
G_3	24	42	35	32
G_4	19	30	24	22

Summary and future work

Summary:

- ▶ Stated a sampling theorem for graph signals.
- ▶ Proposed framework to compute analogs of "Nyquist frequency" and "Nyquist rate".
- ▶ Results provide trade-off between accuracy and complexity.

Future work:

- ▶ Optimal solution for P2.
- ▶ Robustness against noisy samples - stable sampling.
- ▶ Sampling sets for arbitrary bands of frequencies.
- ▶ Wavelets on arbitrary graphs.

USC University of
Southern California