Efficient Signed Graph Sampling via Balancing & Gershgorin Disc Perfect Alignment

Gene Cheung

York University

Toronto, Canada

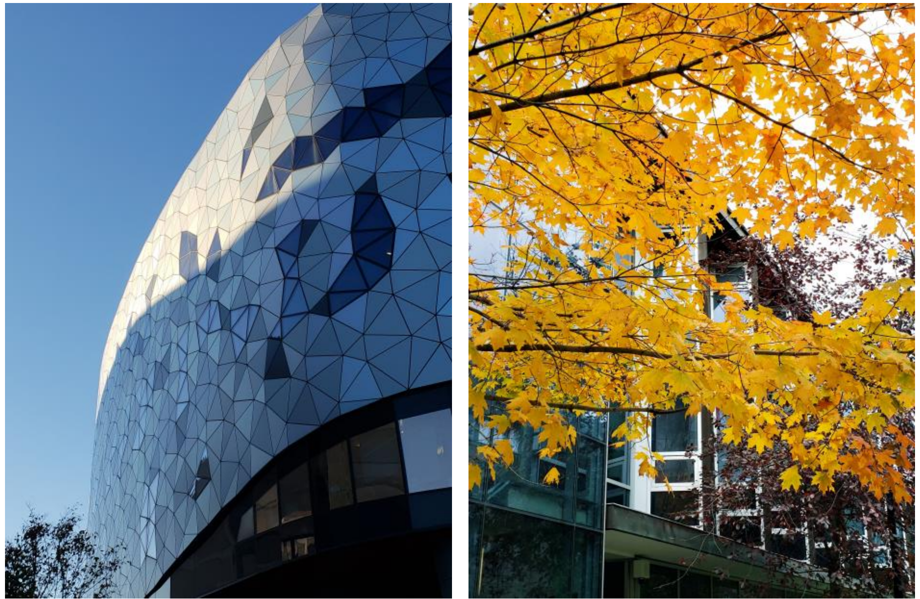
July 28, 2022

Acknowledgement

- **Graph and Image Signal Processing (GISP) Lab** (York University, Toronto, Canada)
  - Post-docs: Cheng Yang*, Xue Zhang, Chinthaka Dinesh
  - Grad students: Saghar Bagheri, Fengbo Lan, Huy Vu, Tam Thuc Do
  - Visiting researchers: Sadid Sahami (NTHU), Fen Wang (Xidian)

- **Collaborators**
  - Michael Brown (York Univ., Canada)
  - Ivan V. Bajic (Simon Fraser Univ., Canada)
  - Antonio Ortega (Univ. of Southern California, USA)
  - Stanley Chan (Purdue Univ., USA)
  - Wai-Tian Tan (Cisco, USA)
  - Jiahao Pang, Dong Tian (InterDigital, USA)
  - Phil Chou (Google, USA)
  - Wei Hu (Peking Univ., China), Jin Zeng (Tongji Univ., China)
  - Chia-Wen Lin (NTHU, Taiwan)
  - Kazuya Kodama (NII, Japan)

interdigital cisco Google GROWERS EDGE



* September 2018 to June 2020.

Gene Cheung (genec@yorku.ca)

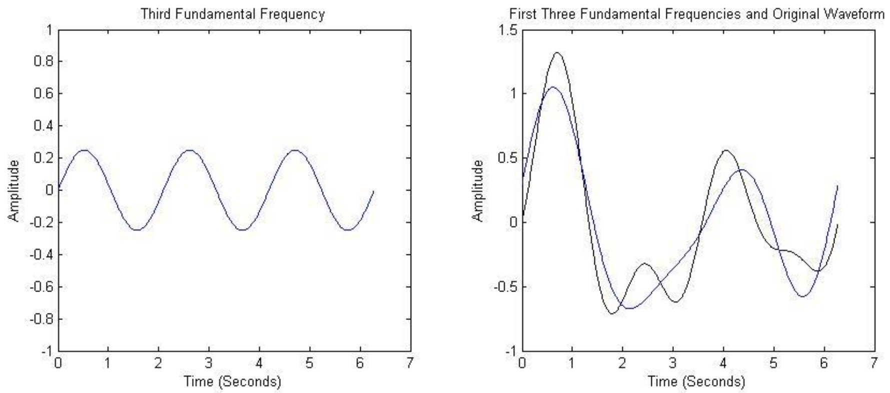
LASSONDE SCHOOL OF ENGINEERING YORK UNIVERSITÉ UNIVERSITY

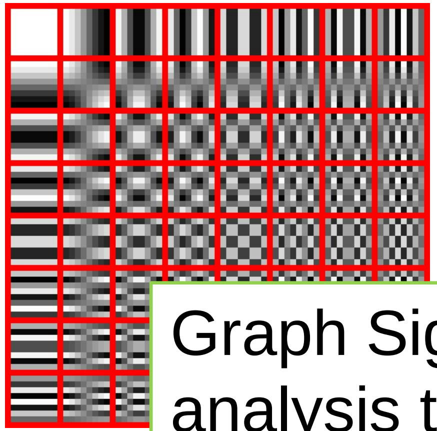
Outline

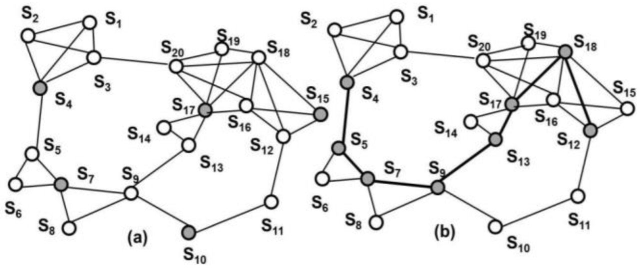
- Overview of Graph Signal Processing (GSP)
- Motivation, related work
- **Key Idea**: Balanced Signed Graph
- Frequencies
  - 1D data kernel (DCT)
  - Positive line graph
  - Balanced signed graph
- Graph Estimation
- Signed Graph Sampling Strategy
  - Sampling objective
  - Gershgorin circle theorem, GDA-based sampling
  - 3-step signed graph sampling recipe
- Results
- Conclusion

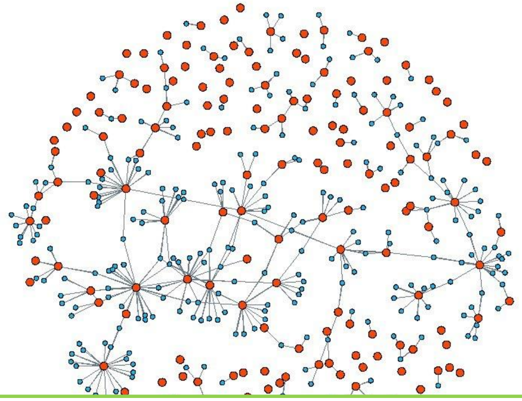
Graph Signal Processing

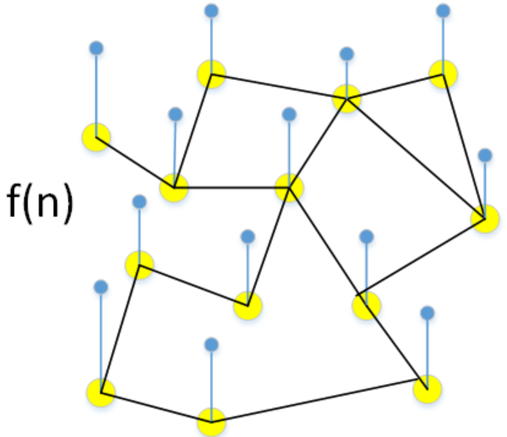
Frequency analysis + **Graph kernels** = **Graph Signal Processing**



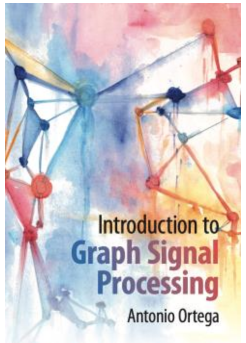




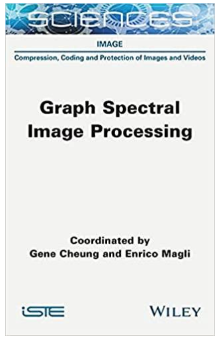




signal on graph kernel



Introduction to
Graph Signal
Processing
Antonio Ortega



SCIENCES
IMAGE
Compression, Coding and Protection of Images and Videos
Graph Spectral
Image Processing
Coordinated by
Gene Cheung and Enrico Magli
ISTE
WILEY

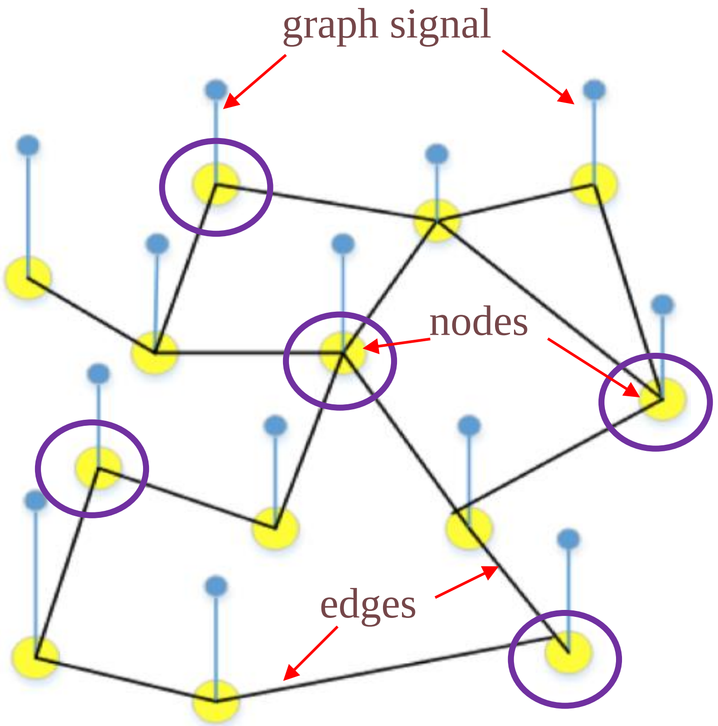
Graph Signal Processing (GSP) studies spectral analysis tools for signals residing on graphs.

[1] A. Ortega, P. Frossard, J. Kovacevic, J. M. F. Moura, and P. Vandergheynst, "**Graph signal processing: Overview, challenges, and applications**," *Proceedings of the IEEE*, vol. 106, no. 5, pp. 808–828, 2018.

Motivation

Graph sampling [1]: Choose a node subset, so that the entire signal can be reconstructed.

- Graph sampling strategies extend **Nyquist sampling** to graph data kernel.
- Bandlimited or smooth signal assumption.



[1] Y. Tanaka et al., “Sampling signals on graphs: From theory to applications,” IEEE Signal Process. Mag., vol. 37, no. 6, pp. 14–30, 2020.

Existing Work

Existing graph sampling methods

Eigen-decomposition-based methods [1,2]

Computational expensive

eigen-decomposition-free methods

Spectral proxies (SP) [3]

Neumann series (NS) [4]

Localization operator (LO) [5]

Gershgoring disc alignment (GDA) [6]

Positive graphs only!

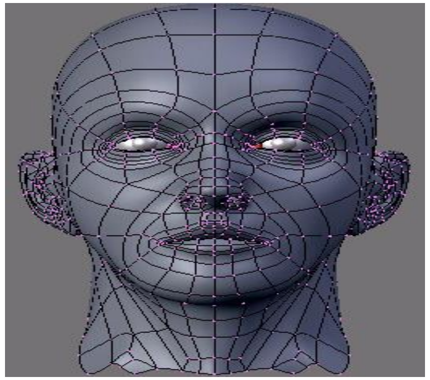
[1] M. Tsitsvero, S. Barbarossa, and P. Di Lorenzo, “Signals on graphs: Uncertainty principle and sampling,” *IEEE TSP*, vol. 64, no. 18, pp. 4845–4860, 2016.
[2] S. Chen, R. Varma, A. Sandryhaila, and J. Kovavcevic, “Discrete signal processing on graphs: Sampling theory,” *IEEE TSP*, vol. 63, no. 24, pp. 6510–6523, 2015.
[3] A. Anis et al., “Efficient sampling set selection for bandlimited graph signals using graph spectral proxies,” *IEEE TSP*, vol. 64, no. 14, pp.3775–3789, 2016.
[4] F. Wang et al., “Low-complexity graph sampling with noise and signal reconstruction via Neumannseries,” *IEEE TSP*, vol. 67, no. 21, pp. 5511–5526, 2019.
[5] A. Sakiyama, Y. Tanaka, T. Tanaka, and A. Ortega, “Eigendecomposition-free sampling set selection for graph signals,” *IEEE TSP*, vol. 67, no. 10, pp. 2679–2692, 2019.
[6] Y. Bai, F. Wang, G. Cheung, Y. Nakatsukasa, and W. Gao, Fast graph sampling set selection using Gershgorin disc alignment,” *IEEE TSP*, vol. 68, pp. 2419–2434, 2020.

Motivation 2

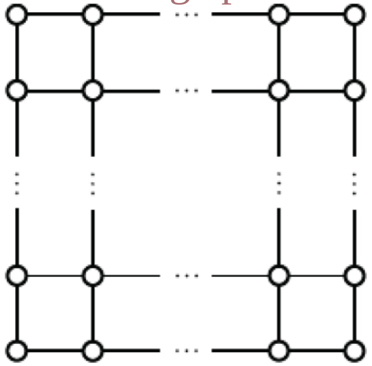
- Previous works designed for **positive graphs**.

Temperature sensor network

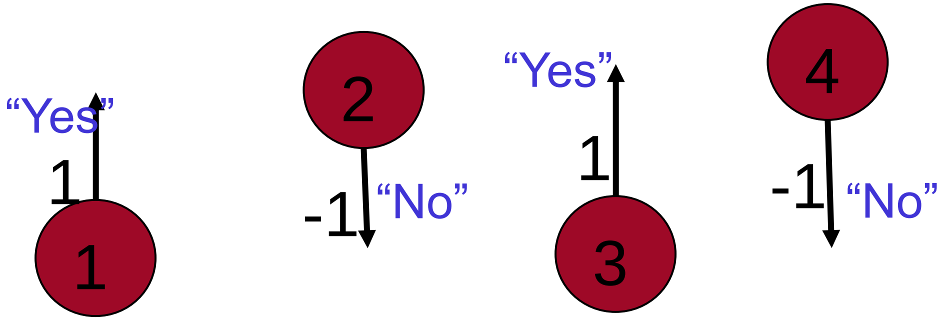
3D geometry

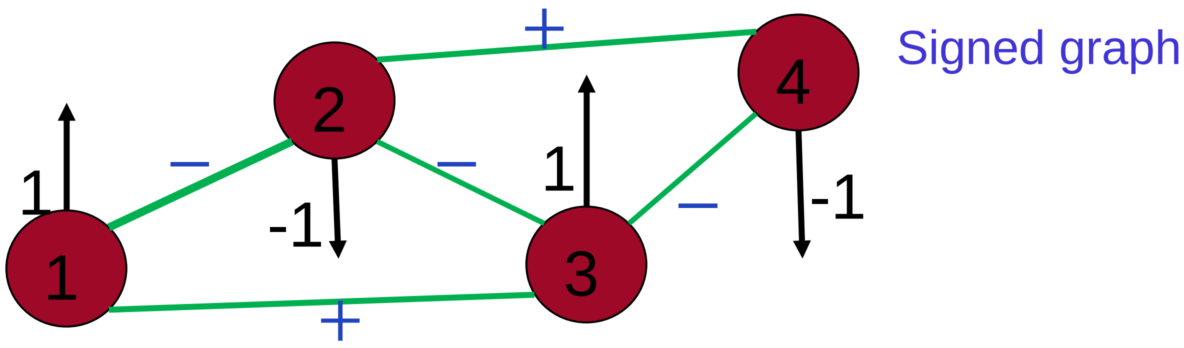


2D image pixels



- Voting records in a Parliament ⟶ ***anti-correlation*** represented as **negative edges** [1, 2].



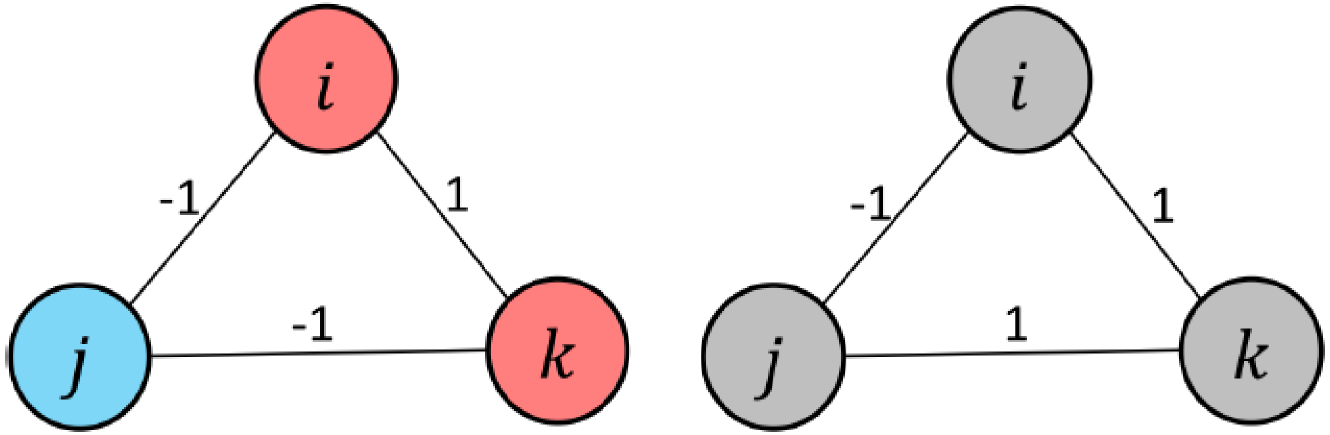


[1] W.-T. Su, G. Cheung, and C.-W. Lin, "Graph Fourier transform with negative edges for depth image coding," *IEEE ICIP*, 2017, pp. 1682–1686.
[2] G. Cheung et al., "Robust semisupervised graph classifier learning with negative edge weights," *IEEE TSIPN*, vol. 4, no. 4, pp. 712–726, 2018.

Key Idea

Balanced Signed Graph [1]: a signed graph with no cycles of odd number of negative edges.

- A natural definition of graph frequencies.
- more amenable to efficient sampling than unbalanced graphs.



[1] D. Easley and J. Kleinberg, “Networks, crowds, and markets: Reasoning about a Highly Connected World”, vol. 8, Cambridge university press, 2010.

Outline

- Overview of Graph Signal Processing (GSP)
- Motivation, related work
- **Key Idea**: Balanced Signed Graph
- Frequencies
    - 1D data kernel (DCT)
    - Positive line graph
    - Balanced signed graph
- Graph Estimation
- Signed Graph Sampling Strategy
    - Sampling objective
    - Gershgorin circle theorem, GDA-based sampling
    - 3-step signed graph sampling recipe
- Results
- Conclusion

Definitions in GSP

- **Adjacency matrix** (W): $\mathbf{W}(i,j) = \mathbf{W}(j,i) = w_{i,j}$

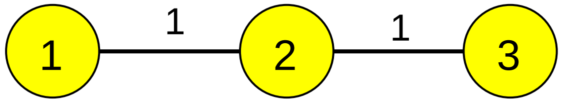
  Edge weight between node *i* and node *j*

- Diagonal **Degree matrix** (D): $\mathbf{D}(i,i) = \sum_j w_{i,j}$

- **Combinatorial graph Laplacian matrix** (L): $\mathbf{L} = \mathbf{D} - \mathbf{W}$

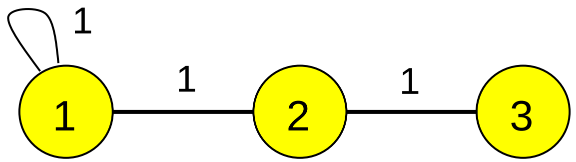
- **Generalized graph Laplacian Matrix**: $\mathcal{L} \triangleq \mathbf{D} - \mathbf{W} + \text{diag}(\mathbf{W})$

To account for self-loops



$$\mathbf{W} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \qquad \mathbf{D} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{L} = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$



$$\mathcal{L} = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

[1] A. Ortega et al., "Graph signal processing: Overview, challenges, and applications," *Proc. IEEE*, vol. 106, no. 5, pp. 808–828, May 2018.

Frequencies for 1D Regular Kernel

Discrete Cosine Transform (DCT)

- ***Eigenvectors*** of 2nd difference matrices:

Minimize Rayleigh quotient

$$\mathbf{T} = \begin{bmatrix} a & b & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & c & d \end{bmatrix}$$

boundary rows

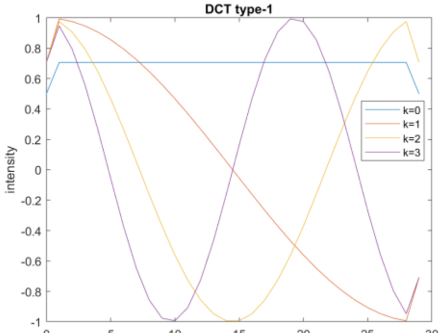
- Two boundary conditions:
  - **Dirichlet** (signal 0 at boundary, **anti-symmetric extension**)
  - **Neumann** (derivative 0 at boundary, **symmetric extension**)

- Two locations to apply conditions:
  - **Midpoint** (half-way bet’n discrete samples)
  - **Meshpoint** (at discrete samples)

[1] G. Strang, “The discrete cosine transform,” *SIAM review*, vol. 41, no.1, pp. 135–147, 1999.

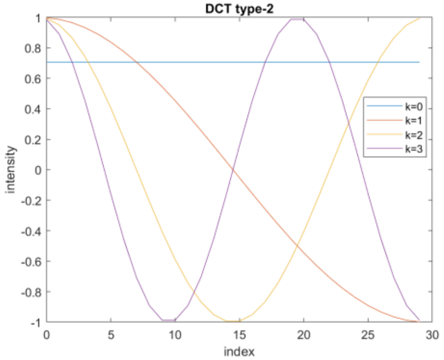
DCT-1

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix}$$



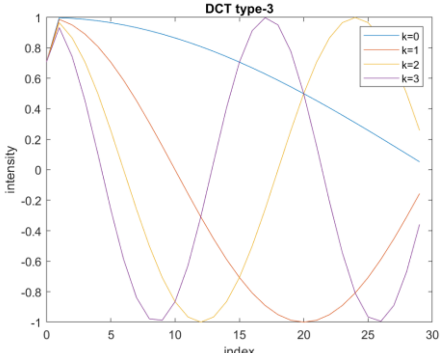
DCT-2

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$



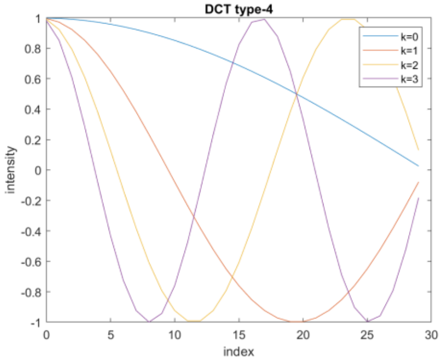
DCT-3

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 2 & -2 \\ -1 & 2 \end{bmatrix}$$



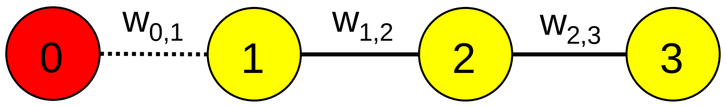
DCT-4

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 3 \end{bmatrix}$$



Frequencies for Positive Line Graph

- Graph Laplacian matrix:



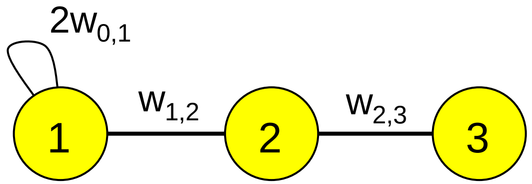
$$\mathcal{L} = \begin{bmatrix} W_{1,2} & -W_{1,2} & 0 & \dots & 0 \\ -W_{1,2} & W_{1,2}+W_{2,3} & -W_{2,3} & \dots & 0 \\ & & \ddots & & \\ 0 & 0 & \dots & -W_{N-1,N} & W_{N-1,N} \end{bmatrix}$$

boundary rows

- Applying interior equation to node 1:

$$(\mathcal{L}\mathbf{x})_1 = -W_{0,1}x_0 + (W_{0,1} + W_{1,2})x_1 - W_{1,2}x_2$$

- **Dirichlet** (signal 0 at boundary) at **midpoint**:



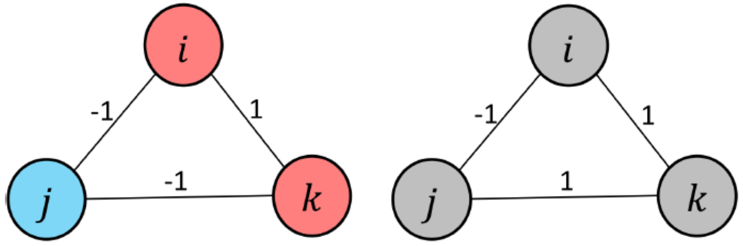
$$(\mathcal{L}\mathbf{x})_1 = (2W_{0,1} + W_{1,2})x_1 - W_{1,2}x_2$$

An external node connected with weight *w* leads to a self-loop of weight *2w* at corresponding boundary node.

Eigenvectors of a generalized graph Laplacian $\mathcal{L}$ for a positive graph *G* with self-loops are graph frequency components.

Frequencies for Balanced Signed Graph

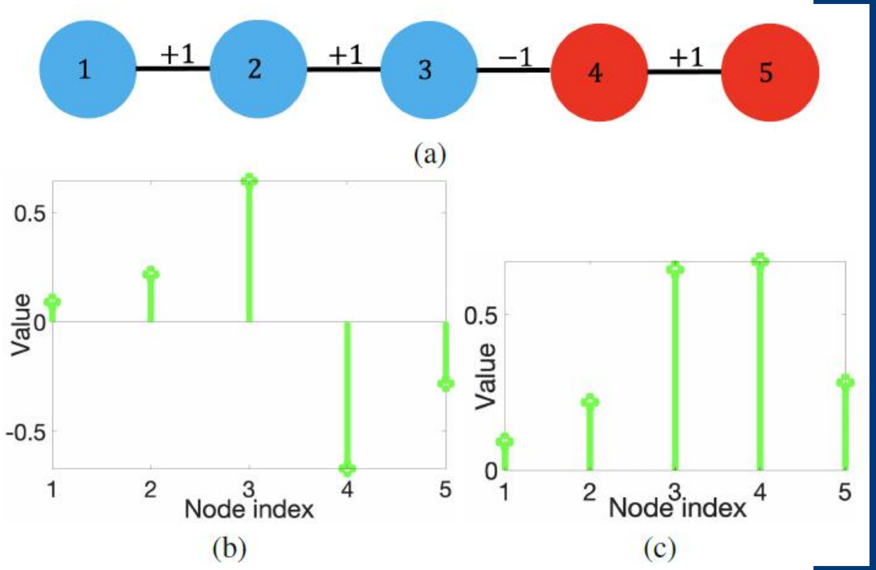
- **Cartwright-Harary Theorem** [2]: A signed graph is balanced iff nodes can be colored to red and blue, s.t. a positive (negative) edge connects nodes of the same (different) color.



- Mapping bet’n balanced signed graph *G* and positive graph *G’*:
  - Graph Laplacians related by *similarity transform*: (same e-val)

$$\mathcal{L}' = \begin{bmatrix}\mathbf{I}_b & \mathbf{0}\\ \mathbf{0} & -\mathbf{I}_r\end{bmatrix}\begin{bmatrix}\mathcal{L}_{11} & \mathcal{L}_{12}\\ \mathcal{L}_{12}^\top & \mathcal{L}_{22}\end{bmatrix}\begin{bmatrix}\mathbf{I}_b & \mathbf{0}\\ \mathbf{0} & -\mathbf{I}_r\end{bmatrix} = \begin{bmatrix}\mathcal{L}_{11} & -\mathcal{L}_{12}\\ -\mathcal{L}_{12}^\top & \mathcal{L}_{22}\end{bmatrix},$$

$$\mathbf{v}_k = \begin{bmatrix}\mathbf{I}_b & \mathbf{0}\\ \mathbf{0} & -\mathbf{I}_r\end{bmatrix}\mathbf{v}'_k$$



(a)

(b)

(c)

Eigenvectors of a generalized graph Laplacian $\mathcal{L}$ for a balanced signed graph *G* with self-loops are graph frequency components.

[1] D. Easley and J. Kleinberg, “Networks, crowds, and markets: Reasoning about a Highly Connected World”, vol. 8, Cambridge university press, 2010.

Outline

- Overview of Graph Signal Processing (GSP)
- Motivation, related work
- **Key Idea**: Balanced Signed Graph
- Frequencies
  - 1D data kernel (DCT)
  - Positive line graph
  - Balanced signed graph
- Graph Estimation
- Signed Graph Sampling Strategy
  - Sampling objective
  - Gershgorin circle theorem, GDA-based sampling
  - 3-step signed graph sampling recipe
- Results
- Conclusion

Graph Laplacian Estimation

- Given *empirical covariance matrix* $\bar{\mathbf{C}}$, we compute $\mathcal{L}$ using GLASSO [1] formulation:

$$\boxed{\min_{\mathcal{L}} \ \mathrm{Tr}(\mathcal{L}\bar{\mathbf{C}}) - \log\det\mathcal{L} + \rho \, ||\mathcal{L}||_1}$$

- Solve via a variant of the *Block Coordinate Descent (BCD)* [2] algorithm.
- In general, solution $\mathcal{L}$ is generalized graph Laplacian for *unbalanced signed graph*.

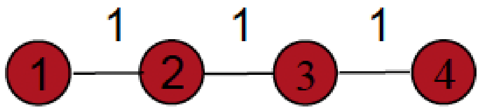
[1] J. Friedman, T. Hastie, and R. Tibshirani, “Sparse inverse covariance estimation with the graphical lasso,” *Biostatistics* (Oxford, England), vol. 9, pp. 432–41, 08, 2008.
[2] S. J. Wright, “Coordinate descent algorithms,” *Math. Program.*, vol. 151, no. 1, pp. 3–34, 2015.

Signal Reconstruction from Samples

desired signal

- Signal Model: $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{e}$ noise

observed signal sampling matrix



Sample set $\{2, 4\}$

$$\mathbf{H} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- Reconstruct signal $\mathbf{x}^* \in \mathbb{R}^N$ from samples $\mathbf{y} \in \mathbb{R}^M,\ M < N$:

$$\boxed{\mathbf{x}^* = \arg\min_{\mathbf{x}} ||\mathbf{H}\mathbf{x} - \mathbf{y}||_2^2 + \mu \mathbf{x}^\top \mathcal{L} \mathbf{x}}$$

Sampling matrix

Graph Laplacian Regularizer (GLR)

- System of Linear Equations for Solution:

$$\boxed{\underbrace{\left(\mathbf{H}^\top \mathbf{H} + \mu \mathcal{L}\right)}_{\mathbf{B}} \mathbf{x}^* = \mathbf{H}^\top \mathbf{y}}$$

Sampling Objective

$$\underbrace{\left(\mathbf{H}^\top \mathbf{H} + \mu \mathcal{L}\right)}_{\mathbf{B}} \mathbf{x}^* = \mathbf{H}^\top \mathbf{y}$$

can be upper-bounded by a small value

- Stability depends on the condition number: $C = \frac{\lambda_{\max}(\mathbf{B})}{\lambda_{\min}(\mathbf{B})}$

can be close to 0 if H is not carefully chosen

- Upper bound of the condition number can be minimized by maximizing $\lambda_{\min}(\mathbf{B})$.
- Select **H** to maximize $\lambda_{\min}(\mathbf{B})$

$$\boxed{\max_{\mathbf{H} \,|\, \mathrm{Tr}(\mathbf{H}^\top \mathbf{H}) \leq M} \lambda_{\min}(\mathbf{B})}$$

sub-sampling budget

- **Theorem:** Maximizing $\lambda_{\min}(\mathbf{B})$ minimizes an MSE upper bound bet’n original and reconstructed signal [1].

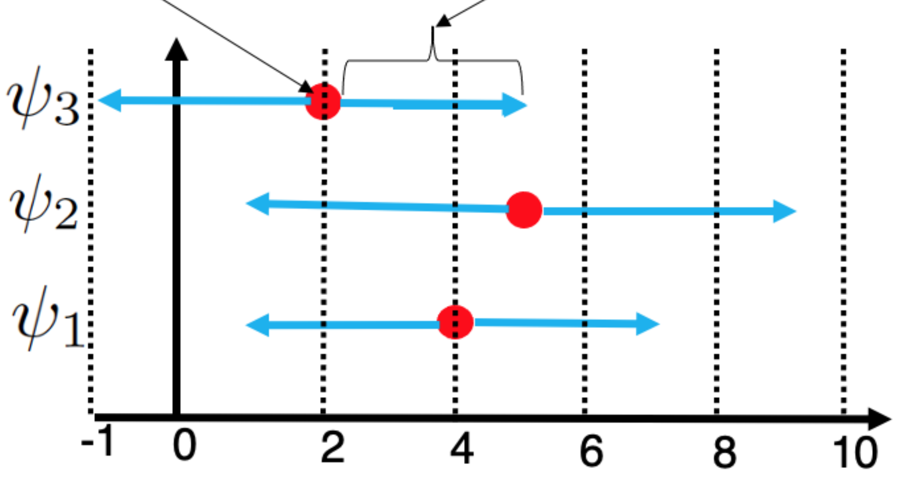
[1] Y. Bai, F.Wang, G. Cheung, Y. Y. Nakatsukasa, and W. Gao, “Fast graph sampling set selection using gershgorin disc alignment," *IEEE Trans. Signal Process*., 2020.

Gershgorin Circle Theorem

- GCT relates matrix entries to bounds of eigenvalues.

Row i of matrix **F** maps to a Gershgorin disc with centre $c_i = \mathbf{F}(i,i)$ and radious $r_i = \sum_{j\neq i} |\mathbf{F}(i,j)|$

$$\mathbf{F} = \begin{bmatrix} 2 & -2 & -1 \\ -2 & 5 & -2 \\ -1 & -2 & 4 \end{bmatrix}$$



$\lambda_{\min}(\mathbf{F})$ is lower-bounded by the *smallest left-end of Gershgorin discs*:

$$\boxed{\lambda^{-}_{\min}(\mathbf{F}) \triangleq \min_i c_i - r_i \leq \lambda_{\min}(\mathbf{F})}$$

[1] [1] R. S. Varga, *Gershgorin and his circles*, Springer, 2004.

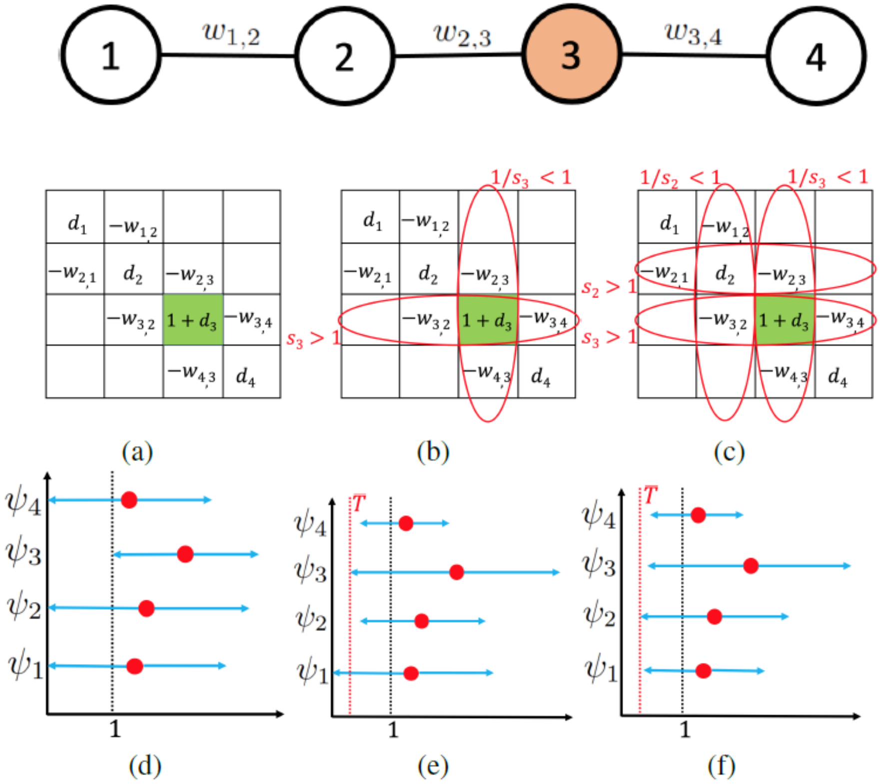
GDA-based Sampling for Positive Graphs

- We focus on maximizing $\lambda_{\min}(\mathbf{B})$:

$$\mathbf{H}^\top\mathbf{H} + \mu\mathcal{L}$$

$$\boxed{\max_{\mathbf{H}\,|\,\mathrm{Tr}(\mathbf{H}^\top\mathbf{H})\leq M} \lambda_{\min}(\mathbf{B})}$$

- Given Gershgorin disc left-ends of $\mathcal{L}$ is at the same exact value, GDA-based graph sampling [1]: Select samples to max smallest disc left-end $\lambda^-_{\min}(\mathbf{B})$ of coefficient matrix **B** via:
  - *Disc shifting* (choosing sample *i*).
  - *Disc scaling* (estimating influence on neighbors given sample *i*).



❑ **Challenge**: Disc left-ends of $\mathcal{L}$ are not at the same exact value.

---

[1] Y. Bai, F.Wang, G. Cheung, Y. Y. Nakatsukasa, and W. Gao, “Fast graph sampling set selection using gershgorin disc alignment," *IEEE Trans. Signal Process.*, 2020.

Signed Graph Sampling

Theorem [1]: Gershgorin disc left-ends of a generalized graph Laplacian matrix $\mathcal{L}_B$ corresponding to a balanced graph $\mathcal{G}_B$ can be aligned exactly to $\lambda_{\min}(\mathcal{L}_B)$ via *similar trasnform* $\hat{\mathbf{S}}\mathcal{L}_B\hat{\mathbf{S}}^{-1}$, where $\hat{s}_i = 1/v_i$, and **v** is the *first eigenvector* of $\mathcal{L}_B$ corresponding to $\lambda_{\min}(\mathcal{L}_B)$.

Sampling Strategy:

1. Approximate $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{W})$ with a balanced graph $\mathcal{G}_B = (\mathcal{V}, \mathcal{E}_B, \mathbf{W}_B)$ while satisfying the following condition:

$$\boxed{\lambda_{\min}(\mathbf{H}^\top\mathbf{H} + \mu\mathcal{L}_B) \leq \lambda_{\min}(\mathbf{H}^\top\mathbf{H} + \mu\mathcal{L})}$$

2. Given balanced graph, perform similarity transform $\mathcal{L}_p = \hat{\mathbf{S}}\mathcal{L}_B\hat{\mathbf{S}}^{-1}$ so that disc left-ends of $\mathcal{L}_p$ are aligned exactly at $\lambda_{\min}(\mathcal{L}_p) = \lambda_{\min}(\mathcal{L}_B)$.
3. Employ GDA sampling method [2] on $\mathcal{L}_p$ to maximize $\lambda^{-}_{\min}(\mathbf{H}^\top\mathbf{H} + \mu\boldsymbol{\mathcal{L}}_p)$.

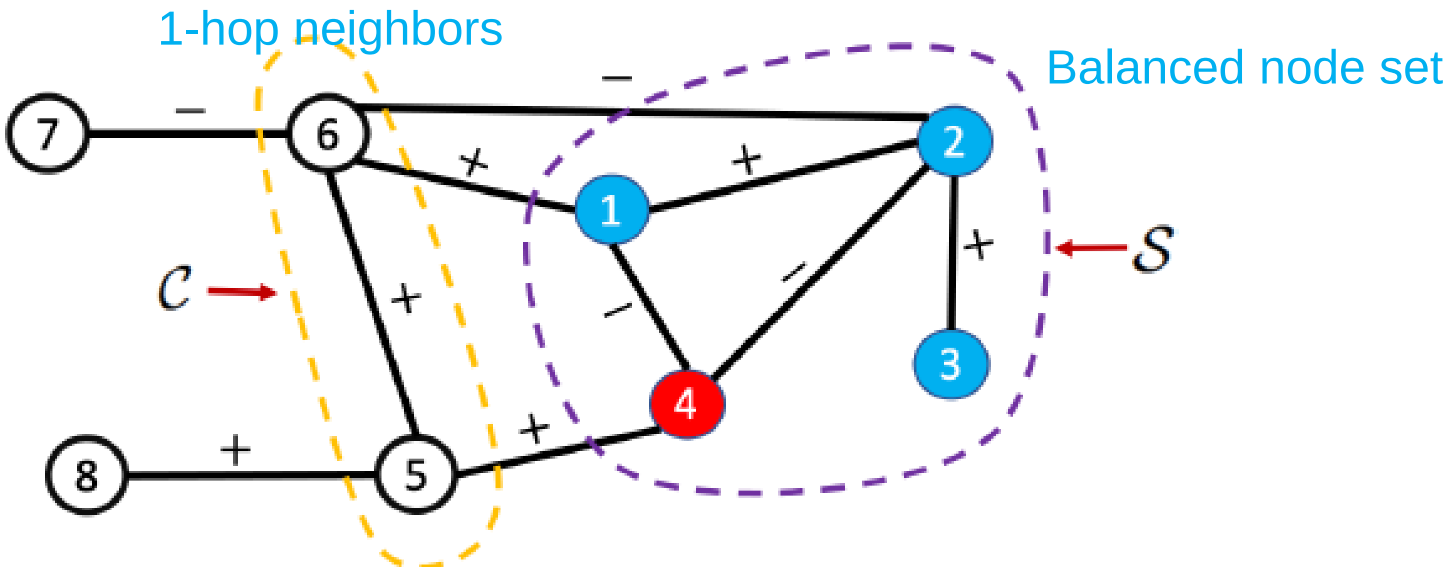
[1] C. Yang, G. Cheung, W. Hu, “Graph Metric Learning via Gershgorin Disc Alignment," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1-15, 2021.

[2] Y. Bai, F.Wang, G. Cheung, Y. Y. Nakatsukasa, and W. Gao, “Fast graph sampling set selection using gershgorin disc alignment," *IEEE Trans. Signal Process.*, 2020

Graph Balancing Algorithm: Overview

- ❑ Construct balanced graph by adding one node at a time to a balanced node set $\mathcal{S}$.
- ❑ At each iteration, choose a most beneficial node $j \in \mathcal{C}$ to add to $\mathcal{S}$, while satisfying constraint

$$\boxed{\lambda_{\min}(\mathbf{H}^\top \mathbf{H} + \mu \mathcal{L}_B) \leq \lambda_{\min}(\mathbf{H}^\top \mathbf{H} + \mu \mathcal{L})}$$



[1] C. Dinesh, S. Bagheri, G. Cheung, I. V. Bajic, "Linear-time Sampling on Signed Graphs via Gershgorin Disc Perfect Alignment," *IEEE ICASSP*, Singapore, May 2022.

Outline

- Overview of Graph Signal Processing (GSP)
- Motivation, related work
- **Key Idea**: Balanced Signed Graph
- Frequencies
  - 1D data kernel (DCT)
  - Positive line graph
  - Balanced signed graph
- Graph Estimation
- Signed Graph Sampling Strategy
  - Sampling objective
  - Gershgorin circle theorem, GDA-based sampling
  - 3-step signed graph sampling recipe
- Results
- Conclusion

Experimental Setup

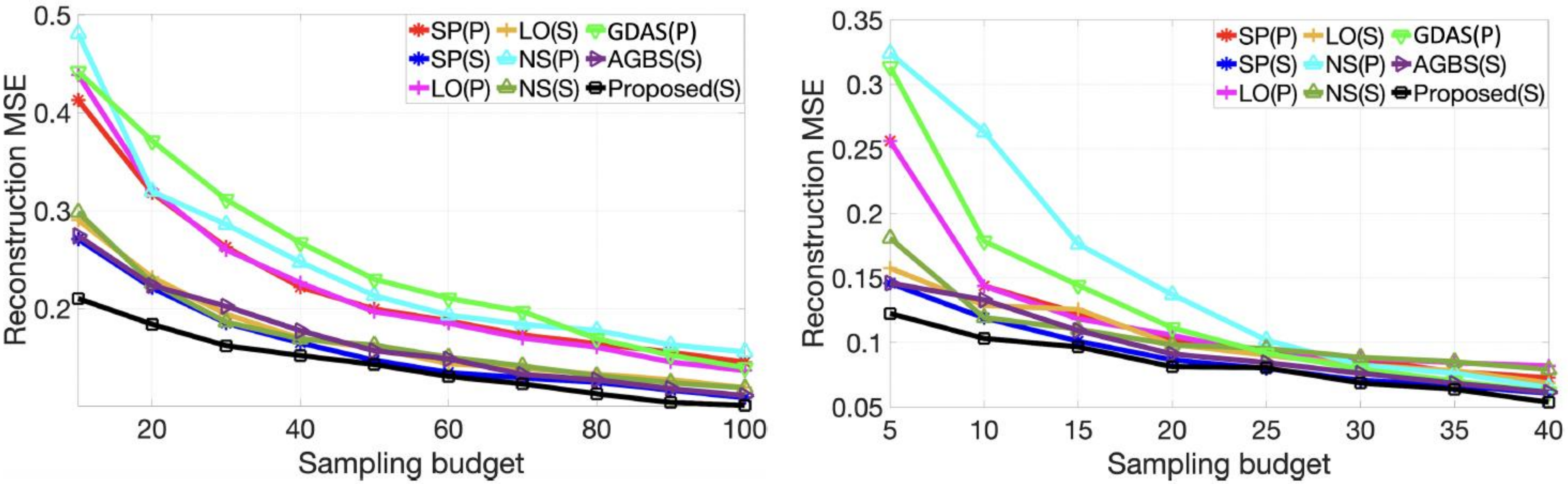
Four datasets:

1. **Canadian parliament voting records data from 2005 to 2021:** voting records of 340 constituencies voted in 3154 elections. The votes are recorded as −1 for “no” and 1 for “yes” and 0 for “abstain / absent”. A signal for a given vote is defined as $\mathbf{x} \in \{1, 0, -1\}^{340}$
2. **US Senate voting records data from 2017 to 2020:** voting records of 100 senators in 1320 elections.
3. **Canadian Car Model Sales Dataset:** Canada vehicle model monthly sales for 2019–2022.
4. **Almanac of Minutely Power Dataset Version 2** (AMPds2): 2 years of ON/OFF status data sampled at 1-minute intervals for 15 residential appliances in a Canadian household.

- Randomly selected 90% of signals from each dataset to learn a signed graph [1] and a positive graph [2].
- Remaining 10% were used to test sampling algorithms.

[1] J. Friedman, T. Hastie, and R. Tibshirani, “Sparse inverse covariance estimation with the graphical lasso,” *Biostatistics* (Oxford, England), vol. 9, pp. 432–41, 08 2008.
[2] H. E. Egilmez, E. Pavez, and A. Ortega, “Graph learning from data under Laplacian and structural constraints,” *IEEE JSTSP*, vol. 11, no. 6, pp. 825–841, Sep. 2017.

Experimental Results

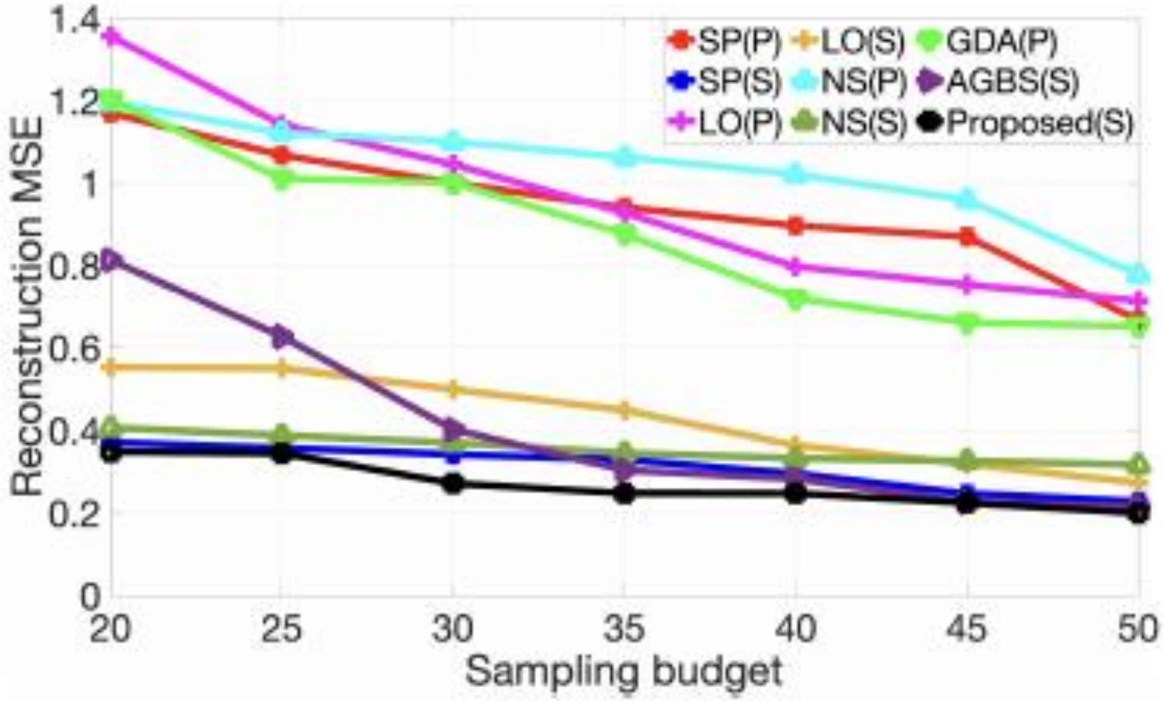


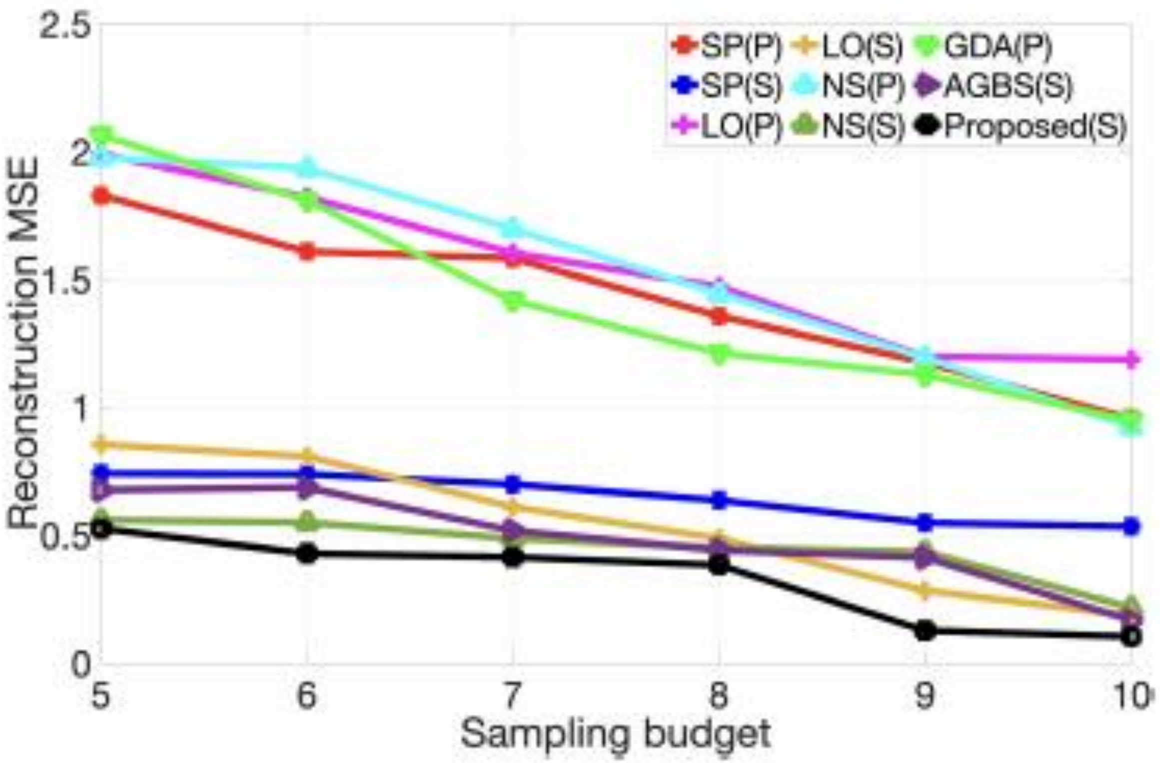
(a) Canadian parliament voting dataset

(b) US senate voting dataset

For the Canadian dataset, our scheme reduced the lowest MSE among competitor schemes by 22.2%, 18.2%, 13.5%, 10.4%, for sampling budget 10, 20, 30, 40.

Experimental Results 2





Similar trends can be observed.

Outline

- Overview of Graph Signal Processing (GSP)
- Motivation, related work
- **Key Idea**: Balanced Signed Graph
- Frequencies
  - 1D data kernel (DCT)
  - Positive line graph
  - Balanced signed graph
- Graph Estimation
- Signed Graph Sampling Strategy
  - Sampling objective
  - Gershgorin circle theorem, GDA-based sampling
  - 3-step signed graph sampling recipe
- Results
- Conclusion

Conclusion

- **Graph Signal Processing (GSP)** studies signals on graphs.

- **Graph sampling** is a fundamental problem in GSP.

- **Balanced signed graph:**
  - Natural frequency interpretation.
  - Amenable to fast graph sampling:
    1. Balance graph.
    2. Align Gershgorin disc left-end via similarity transform [1].
    3. Run GDA-based sampling algorithm [2].

- **Future Work**: Graph sampling for board applications.
  - e.g., video summarization.

Applications: Image coding, denoising, deblurring, interpolation, contrast enhancement, light field image denoising, 3D point cloud denoising, sub-sampling, super-resolution, matrix completion, semi-supervised classifier learning, video summarization, crop yield prediction

[1] C. Yang, G. Cheung, W. Hu, "Graph Metric Learning via Gershgorin Disc Alignment," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1-15, 2021.
[2] Y. Bai, F.Wang, G. Cheung, Y. Y. Nakatsukasa, and W. Gao, "Fast graph sampling set selection using gershgorin disc alignment," *IEEE Trans. Signal Process*., 2020

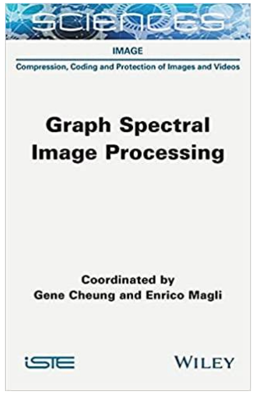
Contact Info

- **Homepage**:

https://www.eecs.yorku.ca/~genec/index.html

- **E-mail**:

genec@yorku.ca



SCIENCES

IMAGE

Compression, Coding and Protection of Images and Videos

Graph Spectral Image Processing

Coordinated by
Gene Cheung and Enrico Magli

ISTE

WILEY

- **New book**:

G. Cheung, E. Magli, (edited) ***Graph Spectral Image Processing***, ISTE/Wiley, August 2021.