Some Mathematical Problems in Graph Signal Processing

Qiyu Sun

University of Central Florida

January 4, 2021

This talk is mainly based on joint works with

- Yang Chen (Hunan Normal Univ.)
- Cheng Cheng (Sun Yat-sen Univ.)
- Nazar Emirov (Univ. Central Florida)
- Qiquan Fang (Zhejiang Univ. of Sci. Tech.)
- Yingchun Jiang (Guilin Univ. Electronic Tech.)
- Chang Eon Shin (Sogang Univ.)
- Guohui Song (Old Dominion Univ.)
- Junzheng Jiang (Guilin Univ. Electronic Tech.)
- Nader Motee (Lehigh Univ.)
- ...

Overview

1 Introduction

- Graph signal processing provides an **innovative framework** to handle data residing on various networks and many irregular domains.
- Graph signal processing is an **emerging interdisciplinary field** and its mathematical foundation is closely related to applied and computational harmonic analysis, algebraic and spectral graph theory, and many more mathematical fields.
- Many important concepts in classical signal processing have been extended to graph setting, including graph Fourier transform, graph wavelet filter banks, and graph filters.
- We have seen many **amazing developments** in last ten years, both in applications and methodology, however there is still a huge research gap between mathematical theory and engineering practice.

- Several review papers:
  - The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains, IEEE Signal Processing Mag. 2013.
  - Graph Signal Processing: Overview, Challenges and Applications, Proc. IEEE, 2018
- Special Issue: *Sampling Signals on Graphs: From Theory to Applications*, IEEE Signal Processing Magazine, November 2020.
- Book chapter: *Introduction to Graph Signal Processing* by Ljubisa Stankovic, Milos Dakovic and Ervin Sejdic, Spring 2018
- Special Issue on *Harmonic Analysis on Graphs*, Journal of Fourier Analysis and Applications (Hans Feichtinger, Isaac Pesenson, Stefan Steinerberger, and S.), 2021.

Motivation and challenge

- Spatially distributed networks (SDNs) have been widely used in (wireless) sensor networks, smart grids, drone fleets and many real world applications.
- Our representative SDNs are sensor networks distributed over a spatial domain with agents communicating with each other via signal broadcasting within a finite range.
- The topological structure of an SDN can be described by some sparse graph with finite Beurling dimension.
- Agents of SDNs have equipped with data processing subsystems with limited capacity and also with communication subsystems for data interchange with neighboring agents.
- Challenges: **Complicated topological structure and distributed implementation of data processing**.

2 Preliminaries

- Graph $\mathcal{G} := (V, E)$, where $V = \{1, 2, \cdots, N\}$ and $E \subset V \times V$.
- Graph provides a flexible model to represent complicated relationships between data on networks.

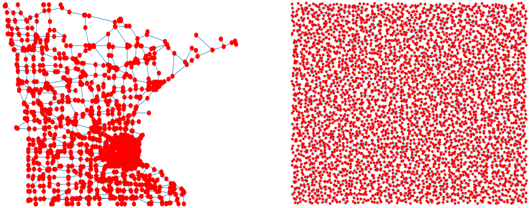


Figure: Minnesota traffic sensor graph and random geometric graph.

Assumption

The graph $\mathcal{G} := (V, E)$ *is "***sparse***" in the sense that*

$$\mu(B(i, r)) \le D_1(\mathcal{G})(1 + r)^d$$

where $\mu(B(i, r))$ *denotes the cardinality of the ball* $B(i, r) = \{j, \rho(i, j) \le r\}$*. (*$D_1$ *density,* d *Beurling dimension)*

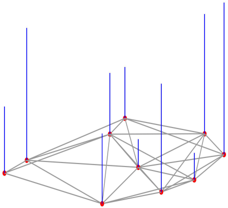
- Maximal degree $\mathrm{Deg}(\mathcal{G}) = \max_i \mu(B(i, 1)) \le D_1 2^d$
- Doubling counting measure

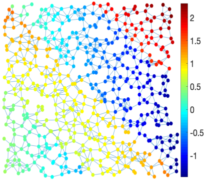
$$\mu(B(i, 2r)) \le D\mu(B(i, r)), r \ge 0$$

implies that $d \le \log_2 D$.

Graph Signals

- A graph signal is a **vector** residing on the graph nodes, denoted by $\mathbf{x} = (x_v)_{v \in V}$, where $x_v, v \in V$ are real-valued, complex-valued or vector-valued.
- graph signals residing on a graph





Original Signal $\mathbf{x}_p(t_1)$

3 Graph Fourier Transform

- Adjacency matrix $\mathbf{A} = (a(i,j))_{i,j\in V}$ of a graph $\mathcal{G} = (V, E)$, where $a(i,j) = 1$ if $(i,j) \in E$ and $a(i,j) = 0$ otherwise.
- Degree matrix $\mathbf{D} = diag\{d(1), d(2), \cdots, d(N)\}$ with $d(i) = \sum_{j\in V} a(i,j)$.
- Laplacian matrix $\mathbf{L} = \mathbf{D} - \mathbf{A}$.
- Write $\mathbf{L} = \mathbf{U}^T \Lambda \mathbf{U} = \sum_{i=1}^N \lambda_i \mathbf{u}_i \mathbf{u}_i^T$ (eigendecomposition).
- **Graph Fourier transform** of a graph signal $\mathbf{x}$ is by $\hat{\mathbf{x}} = \mathbf{U}\mathbf{x}$ and the inverse graph Fourier transform is $\mathbf{x} = \mathbf{U}^T \hat{\mathbf{x}}$.

An illustrative example of graph Fourier transform.

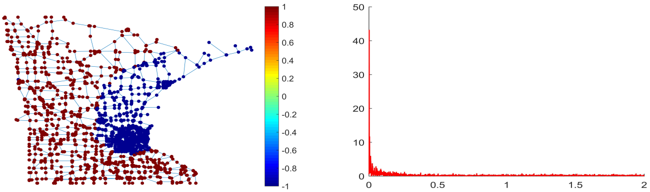


Figure: Piecewise constant signal on Minnesota traffic graph and its Fourier transform (plotted is $|\hat{\mathbf{f}}|$).

- **Phase retrieval:** How to find real/complex/vector-valued graph signals f in some linear space so that they can determined, up to sign $\pm$ in real setting, module z with $|z| = 1$ in complex setting and an orthogonal matrix U in vector-valued setting, from magnitude $|\hat{\mathbf{f}}|$ of their Fourier measurements or $|\langle \psi, f \rangle|$ of their frame measurements?

4 Phase retrieval and vector field on graphs

- **Phase retrieval problem**: How to determine the vector-valued signal $\mathbf{f}_i = (f_{1,i}, f_{2,i}, f_{3,i})^T, i \in V$ from phaseless measurements $|\sum_{j\in V} a_{kj}\mathbf{f_i}|, 1 \le k \le m$?
- How to determine velocity (position) $\mathbf{f}_i = (f_{1,i}, f_{2,i}, f_{3,i})^T, i \in V$ from **absolute speed** (distance) $|f_i|, i \in E$ of each agent and **relative speed** (distance) $|f_i - f_j|, i,j \in E$. In other words, given any vector field $\mathbf{g} = (\mathbf{g}_i)_{i\in V}$ satisfying $\|\mathbf{g}_i\| = \|\mathbf{f}_i\| \ \text{ for all } i \in V$ and $\|\mathbf{g}_i - \mathbf{g}_j\| = \|\mathbf{f}_i - \mathbf{f}_j\| \ \text{ for all } (i,j) \in E$, can we find an orthogonal matrix U of size 3×3 such that $\mathbf{g} = U\mathbf{f}$?

Theorem

Let $\mathcal{G} = (V, E)$ be a complete graph. Then a vector field $\mathbf{f} = (\mathbf{f}_i)_{i\in V}$ can be reconstructed, up to an orthogonal matrix, from its absolute magnitudes $\|\mathbf{f}_i\|, i \in V$ and relative magnitudes $\|\mathbf{f}_i - \mathbf{f}_j\|, i,j \in V$.

- The determination of velocity field on complete graphs from absolute and relative speeds does **not** hold for sparse graphs.

- For a simple graph $\mathcal{G} = (V, E)$ and a 3-dimensional vector field $\mathbf{f} = (\mathbf{f}_i)_{i \in V}$ on $\mathcal{G}$, we define

$$\triangle(\mathbf{f}, \mathcal{G}_c) = \Big\{ \sum_{i \in V_c} t_i \mathbf{f}_i, \ \sum_{i \in V_c} t_i = 1 \text{ and } 0 \le t_i \le 1 \text{ for all } i \in V_c \Big\}, \quad (1)$$

  for any $\mathcal{G}_c = (V_c, E_c)$, a complete subgraph of $\mathcal{G}$ with order 4.
- Let $V_{\mathbf{f}}$ be the set of all complete subgraphs $\mathcal{G}_c$ of order 4 such that $\triangle(\mathbf{f}, \mathcal{G}_c)$ is a 3-simplex (tetrahedron), and $E_{\mathbf{f}}$ be the set of all pairs of complete subgraphs $\mathcal{G}_c, \tilde{\mathcal{G}}_c \in V_{\mathbf{f}}$ of order 4 such that $\triangle(\mathbf{f}, \mathcal{G}_c \cap \tilde{\mathcal{G}}_c)$ is a 2-simplex (triangle) and the hyperplane containing $\triangle(\mathbf{f}, \mathcal{G}_c \cap \tilde{\mathcal{G}}_c)$ does not contain the origin.
- We call the graph $\mathcal{G}_{\mathbf{f}} = (V_{\mathbf{f}}, E_{\mathbf{f}})$ with the vertex set $V_{\mathbf{f}}$ and edge set $E_{\mathbf{f}}$ defined above as the 3-*simplex graph* associated with the vector field $\mathbf{f}$.

Theorem

(Chen, Cheng and S. Arxiv 2019) Let $\mathcal{G} = (V, E)$ *be a simple finite graph,* $\mathbf{f} = (\mathbf{f}_i)_{i \in V}$ *be a vector field on the graph* $\mathcal{G}$ *with its d-simplex graph denoted by* $\mathcal{G}_{\mathbf{f}}$*. If the d-simplex graph* $\mathcal{G}_{\mathbf{f}}$ *is connected and for any vertex* $i \in V$ *there exists a complete subgraph* $\mathcal{G}_c = (V_c, E_c)$ *of order* $d + 1$ *such that* $i \in V_c$ *and* $\triangle(\mathbf{f}, \mathcal{G}_c)$ *is a d-simplex, then the vector field* $\mathbf{f}$ *is determined, up to an orthogonal matrix, from its absolute magnitudes* $\|\mathbf{f}_i\|, i \in V$ *at vertices and relative magnitudes* $\|\mathbf{f}_i - \mathbf{f}_j\|, (i, j) \in E$ *of neighboring vertices.*

- Phase retrieval for real/complex/vector-valued graph signals is a widely open field for further study.

5 Graph signal processing and filter banks

- A graph processing operator on a graph $\mathcal{G}$ is a (linear) map from one graph signal $\mathbf{x}$ to another graph signal $\mathbf{y} = \mathbf{A}(\mathbf{x})$. In the case that the operator is linear, it can be represented by a graph filter/matrix $\mathbf{A} = (a(i;j))_{i,j\in V}$.
- **Convolution**: $\mathbf{x} * \mathbf{y}$ between $\mathbf{x}$ and $\mathbf{y}$: $\widehat{\mathbf{x} * \mathbf{y}} = \hat{\mathbf{x}} \odot \hat{\mathbf{y}}$, where $\hat{\mathbf{x}} = \mathbf{U}\mathbf{x}$ is the Fourier transform of the graph signal $\mathbf{x}$.
- **Polynomial filter and convolution**: For the polynomial filter $\mathbf{H} = h(\mathcal{L}) = \sum_{k=0}^{L} h_k \mathcal{L}^k$, one may verify that $\mathbf{H}\mathbf{x} = \mathbf{y} * \mathbf{x}$ for some graph signal $\mathbf{y}$ with $\hat{y}(\lambda) = h(\lambda)$ on the spectrum set of graph Laplacian. The converse is true if all eigenvalues are distinct.

- **Graph shift** $\mathbf{S} = (S(i,j))$ has bandwidth at most one, i.e., $S(i,j) = 0$ if $\rho(i,j) \geq 2$.
- Illustrative examples of graph shifts: the Laplacian matrix $\mathbf{L} = \mathbf{D} - \mathbf{A}$ and the symmetric normalized Laplacian matrix $\mathbf{L}^{\rm sym} = \mathbf{D}^{-1/2}\mathbf{L}\mathbf{D}$.
- Lowpass spline filter, $\mathbf{H}_0 = \left(\mathbf{I} - \frac{\mathbf{L}^{\rm sym}}{2}\right)^n$; highpass spline filter $\mathbf{H}_1 = \left(\frac{\mathbf{L}^{\rm sym}}{2}\right)^n, n \geq 1$.

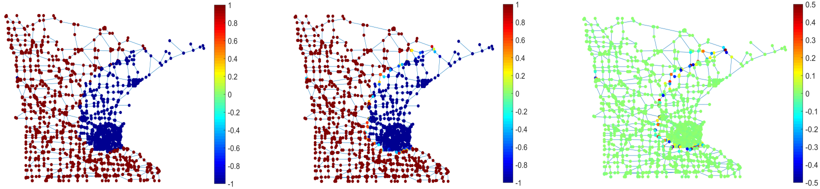


Figure: From left to right: original, lowpass, and highpass filtered signals, where $n = 2$.

- Filters in many engineering problem are designed to be polynomials of single or multiple graph shifts,

$$\mathbf{H} = h(\mathbf{S}_1, \ldots, \mathbf{S}_d) = \sum_{l_1=0}^{L_1} \cdots \sum_{l_d=0}^{L_d} h_{l_1,\ldots,l_d} \mathbf{S}_1^{l_1} \cdots \mathbf{S}_d^{l_d} \tag{2}$$

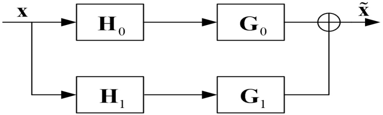
where $h(t_1, \ldots, t_d) = \sum_{l_1=0}^{L_1} \cdots \sum_{l_d=0}^{L_d} h_{l_1,\ldots,l_d} t_1^{l_1} \ldots t_d^{l_d}$, and $\mathbf{S}_1, \ldots, \mathbf{S}_d$ are commutative graph shifts, $\mathbf{S}_k \mathbf{S}_{k'} = \mathbf{S}_{k'} \mathbf{S}_k$.
- Commutative graph shifts play a similar role to the one-order delay $z_1^{-1}, \ldots, z_d^{-1}$ in classical signal processing.
- In practice, multiple graph shifts may have specific features and physical interpretation, such as in time-varying signals, which carry different correlation characteristics for different dimensions/directions.

Theorem

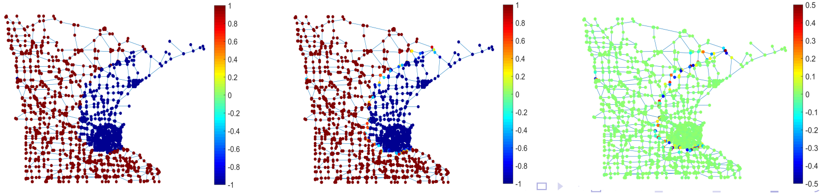
(Emirov, Cheng, Jiang and S. Arxiv 2020) The filtering procedure $\mathbf{x} \longmapsto \mathbf{Hx}$ *can be implemented in a recursive algorithm containing about* $\sum_{m=0}^{d-1} \prod_{k=1}^{m+1} (L_k + 1)$ *steps with the output value at each vertex in each step being updated from some weighted sum of the input values at adjacent vertices of its preceding step.*

6 Nonsubsampled graph filter banks

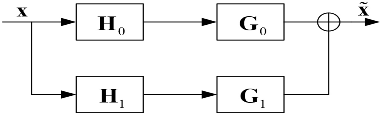
- A *nonsubsampled graph filter bank* (NSGFB) that contains an analysis filter bank $(\mathbf{H}_0, \mathbf{H}_1)$ and a synthesis filter bank $(\mathbf{G}_0, \mathbf{G}_1)$. (Jiang, Cheng, S. IEEE TSP, 2019)



- The analysis filter bank decomposes a graph signal into two components $\mathbf{y}_0 = \mathbf{H}_0\mathbf{x}$ and $\mathbf{y}_1 = \mathbf{H}_1\mathbf{x}$ carrying different frequency information. For instance, spline analysis filter bank, $\mathbf{H}_0 = \left(\mathbf{I} - \mathbf{L}^{\rm sym}/2\right)^n$ and $\mathbf{H}_1 = \left(\mathbf{L}^{\rm sym}/2\right)^n, n \geq 1$.



- An important concept for an NSGFB is the *perfect reconstruction* condition $\tilde{\mathbf{x}} = \mathbf{x}$,



- The nonsubsampled structure in an NSGFB greatly simplifies the design of analysis filter banks for spectral decomposition and synthesis filter banks for signal reconstruction.
- The PR condition can be characterized by the following matrix equation,

$$\mathbf{G}_0\mathbf{H}_0 + \mathbf{G}_1\mathbf{H}_1 = \mathbf{I}. \tag{3}$$

- The existence of synthesis filter banks can be described by the following stability condition

$$A\|\mathbf{x}\|_2^2 \le \|\mathbf{H}_0\mathbf{x}\|_2^2 + \|\mathbf{H}_1\mathbf{x}\|_2^2 \le B\|\mathbf{x}\|_2^2. \tag{4}$$

- If analysis filters $(\mathbf{H}_0, \mathbf{H}_1)$ in NSGFB are polynomials of a graph shift $\mathbf{S}$ that have non common roots, using Chinese reminder theorem, we can construct synthesis filter bank $(\mathbf{G}_0, \mathbf{G}_1)$, being polynomials of $\mathbf{S}$, so that the following Bezout identity holds,

$$\mathbf{G}_0\mathbf{H}_0 + \mathbf{G}_1\mathbf{H}_1 = \mathbf{I}. \quad (5)$$

- Let $(\mathbf{H}_0, \mathbf{H}_1)$ be a stable analysis filter bank. We can construct synthesis filter banks $(\mathbf{G}_0, \mathbf{G}_1)$ of an NSGFB by solving the minimization problem:

$$\underset{\mathbf{G}_0, \mathbf{G}_1}{\text{minimize}} \ \|\mathbf{G}_0\|_F^2 + \|\mathbf{G}_1\|_F^2$$

subject to the perfect reconstruction condition (5). The solution is

$$\mathbf{G}_{0,L} = \mathbf{H}^{-1}\mathbf{H}_0^T \ \text{ and } \ \mathbf{G}_{1,L} = \mathbf{H}^{-1}\mathbf{H}_1^T, \quad (6)$$

where $\mathbf{H} = \mathbf{H}_0^T\mathbf{H}_0 + \mathbf{H}_1^T\mathbf{H}_1$ is positive definite.

- Implementation of the analysis filtering procedure $\mathbf{y}_0 = \mathbf{H}_0\mathbf{x}, \mathbf{y}_1 = \mathbf{H}_1\mathbf{x}$; and the synthesis filtering procedure

$$\tilde{\mathbf{x}} = \mathbf{G}_{0,L}\mathbf{y}_0 + \mathbf{G}_{1,L}\mathbf{y}_1 = \mathbf{H}^{-1}(\mathbf{H}_0^T\mathbf{y}_0 + \mathbf{H}_1^T\mathbf{y}_1).$$

7. Off-diagonal decay for inverse filtering (Wiener's lemma)

- The synthesis filter bank is given by $\mathbf{G}_{0,L} = \mathbf{H}^{-1}\mathbf{H}_0^T$ and $\mathbf{G}_{1,L} = \mathbf{H}^{-1}\mathbf{H}_1^T$, where $\mathbf{H} = \mathbf{H}_0^T\mathbf{H}_0 + \mathbf{H}_1^T\mathbf{H}_1$ is positive definite.
- **Geodesic-width** of a filter $\mathbf{H} = (H(i,j))_{i,j\in V}$ is the smallest nonnegative integer $\sigma(\mathbf{H})$ such that

$$H(i,j) = 0 \text{ if } \rho(i,j) > \sigma(\mathbf{H}). \tag{7}$$

- If $\mathbf{H} = \sum_{l=0}^{L} a_l \mathbf{S}^l$ (polynomial of a graph shift), then $\sigma(\mathbf{H}) \le L$.
- If $\mathbf{H}$ has small geodesic-width, then $\mathbf{H}^{-1}$ has exponential off-diagonal decay.

Theorem

If $\mathbf{H}$ has geodesic-width $\sigma(\mathbf{H}) \le \sigma$ and satisfies $c_1\mathbf{I} \le \mathbf{H} \le c_2\mathbf{I}$, then $\mathbf{H}^{-1}$ has exponential off-diagonal decay,

$$|H^{-1}(i,j)| \le \frac{c_2^2}{c_1}\left(1 - \frac{c_1}{c_2}\right)^{\rho(i,j)/\sigma}.$$

- **Beurling algebra** $\mathbf{B}_{r,\alpha}, 1 \le r \le \infty, \alpha \ge 0$, of filters/matrices $\mathbf{H} = (H(i,j))_{i,j \in V}$:

$$\|\mathbf{H}\|_{\mathcal{B}_{r,\alpha}} = \Big(\sum_{n=0}^{\infty} (h_{\mathbf{H}}(n))^r (n+1)^{\alpha r + d - 1} \Big)^{1/r}$$

where $h_{\mathbf{H}}(n) = \sup_{\rho(i,j) \ge n} |H(i,j)|$ and d is the Beurling dimension of the graph (Beurling (1952) for $r = 1$ and $\alpha = 0$, Jaffard (1990) for $r = \infty$ and $\alpha \ge 0$, S. (2011) for $1 \le r \le \infty$ and $\alpha \ge 0$, Shin and S. (2019))

Theorem

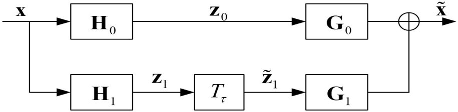
(Shin and S., JFA 2019) If $\mathbf{H} \in \mathcal{B}_{r,\alpha}$ *for some* $1 \le r \le \infty$ *and* $\alpha > d(1 - 1/r) + 1$ *and* $\mathbf{H}$ *is invertible, then*

$$\|\mathbf{H}^{-1}\|_{\mathcal{B}_{r,\alpha}} \le C \|\mathbf{H}^{-1}\|_{\mathcal{B}(\ell^2)} (\|\mathbf{H}^{-1}\|_{\mathcal{B}(\ell^2)} \|\mathbf{H}\|_{\mathcal{B}_{r,\alpha}})^{\alpha + d/r}.$$

- Exponent $\alpha + d/r$ cannot be replaced by $\alpha + d/r - 1 - \epsilon, \epsilon > 0$.
- C^*-algebra $\mathcal{B}(\ell^2)$ replaced by Banach algebra $\mathcal{B}(\ell^q_w), 1 \le q \le \infty$ and Muckenhoupt weight w (Fang, Shin and S., ArXiv 2019).

off-diagonal decay and denoising via NSFBs

- Denoising via NSFBs, where T_τ is a hard-thresholding operator, and the input signal is corrupted by bounded noise, i.e., $\mathbf{x} = \mathbf{x}_0 + \mathbf{n}$.



- Conclusion from off-diagonal decay: $\|\tilde{\mathbf{x}} - \mathbf{x}\|_\infty \le C\tau$.

Table: Denoising performance on the Minnesota traffic graph: SNR over 50 trials.

$\|\mathbf{n}\|_\infty$	1/32	1/16	1/8	1/4	1/2	1
Noisy	34.89	28.85	22.83	16.82	10.81	4.75
graphBior	34.43	28.91	24.06	18.21	12.79	7.39
OSGFB	38.25	32.59	24.44	16.70	12.54	4.69
PRT	35.31	29.41	23.74	18.46	15.45	12.77
NSGFB-L1	38.49	32.44	26.42	19.25	13.82	8.34

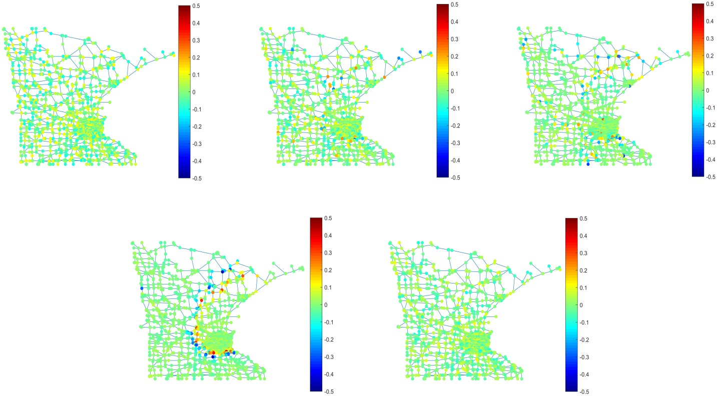


Figure: Top left to right: pure noise signal, residual signals by the Bior, OSGFB. Bottom left to right: residual signals by the Pyramid and proposed algorithms.

8 Stability of Dynamic Systems/Sampling

- Stability of the analysis filter bank $A\|x\|_2^2 \le \|\mathbf{H}_0 x\|_2^2 + \|\mathbf{H}_1 x\|_2^2 \le B\|x\|_2^2$ and invertibility of $\mathbf{H} = \mathbf{H}_0^T \mathbf{H}_0 + \mathbf{H}_1^T \mathbf{H}$.
- **Exponential stability** of a dynamic system $\mathbf{x}'(t) = \mathbf{A}\mathbf{x}(t)$ on a graph:

$$\|\mathbf{x}(t)\|_p \le C \exp(-\alpha t),\ \ t \ge 0,$$

where $1 \le p \le \infty, \alpha > 0$ and $A = (a(i,j))_{i,j\in V}$ has geodesic-width σ.

- Exponential stability $\Longleftrightarrow$ (eigenvalues) all eigenvalues of $\mathbf{A}$ is contained in $\{z|\Re(z) \le -\delta\}$ for some $\delta > 0 \Longleftrightarrow$ (invertibility) $z\mathbf{I} - \mathbf{A}$ is invertible for all $z \in \mathbb{C}_+ \Longleftrightarrow$ (stability) there exists a positive constant A_0 such that

$$\|(z\mathbf{I} - \mathbf{A})\mathbf{c}\|_2 \ge A_0\|\mathbf{c}\|_2 \quad \text{and} \quad \|(z\mathbf{I} - \mathbf{A}^*)\mathbf{c}\|_2 \ge A_0\|\mathbf{c}\|_2 \qquad (8)$$

for all $z \in \mathbb{C}_+$ and $\mathbf{c} \in \ell^2$.

- The linear system $\mathbf{x}'(t) = \mathbf{A}\mathbf{x}(t)$ is exponentially stable if and only if $\|(z\mathbf{I} - \mathbf{A})\mathbf{c}\|_2 \ge A_0\|\mathbf{c}\|_2$ and $\|(z\mathbf{I} - \mathbf{A}^*)\mathbf{c}\|_2 \ge A_0\|\mathbf{c}\|_2$ hold for all $z \in \mathbb{C}_+$ and $\mathbf{c} \in \ell^2$.

Theorem

(Motee and S. 2019) Suppose that the state matrix $\mathbf{A}$ *belongs to* $\mathcal{B}_\tau \cap \mathcal{B}^2$ *for some* $\tau \ge 0$. *If there exists a positive integer* $N_0 \ge \tau$ *and a positive number* $B_{N_0} \ge 4\tau\sqrt{\alpha_2^*/\alpha_1^*}\ \|\mathbf{A}\|_{\mathcal{S}}\ N_0^{-1}$ *such that*

$$\min\left(\left\|(z\mathbf{I} - \mathbf{A})\chi_{i_m}^{N_0}\mathbf{c}\right\|_2, \left\|(z\mathbf{I} - \mathbf{A}^*)\chi_{i_m}^{N_0}\mathbf{c}\right\|_2\right) \ge B_{N_0}\left\|\chi_{i_m}^{N_0}\mathbf{c}\right\|_2 \qquad (9)$$

hold for all $z \in C_+, \mathbf{c} \in \ell^2$ *and* $i_m \in V_{N_0}$, *then the linear system is exponentially stable and its stability threshold* A_0 *satisfies* $A_0 \ge B_{N_0}\sqrt{\alpha_1^*/(4\alpha_2^*)}$, *in which* α_1^* *and* α_2^* *are constants related to the ball covering with centers in* V_{N_0}.

- Applicable to adjust the feedback by increasing/decreasing $a(i,j)$ of the state matrix A at a particular vertex pair to improve exponential stability of a closed-loop network.

Stability of Sampling

- Sampling procedure $f \longmapsto (\langle \psi, f\rangle)_{\psi\in\Psi}$, where $\Psi = \{\psi\}$ is a family of sampling functional. For $f \in \ell^2(V)$, the sampling procedure can be described by a sensing matrix $\mathbf{S} = (s(i,j))_{i\in W, j\in V}$.

Theorem

(Cheng, Jiang and S. ACHA 2019) Let the sensing matrix $\mathbf{S} \in \mathcal{J}_\alpha(\mathcal{W},\mathcal{G})$, Jaffard class for some $\alpha > d$. If there exist a positive constant A_0 and an integer $N_0 \ge 3$ such that $A_0 \ge C N_0^{-\min(\alpha-d,1)}$ and for all $i \in V$, the quasi-restrictions have uniform stability,

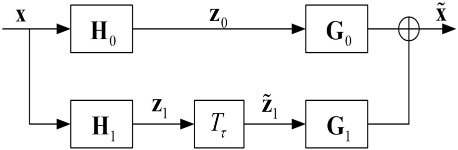
$$\|\chi_{i,W}^{2N_0}\mathbf{S}\chi_{i,V}^{N_0}\mathbf{c}\|_2 \ge A_0\|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{W},\mathcal{G})}\|\chi_{i,V}^{N_0}\mathbf{c}\|_2,\ \mathbf{c}\in\ell^2, \tag{10}$$

then $\mathbf{S}$ has ℓ^2-stability.

- Pivotal for the design of a robust distributed sampling and reconstruction system against supplement, replacement and impairment of agents.
- Sampling on Paley-Wiener spaces and reproducing kernel spaces, ...

9 Distributed implementation of inverse filtering

- Reconstruct graph signals $\tilde{f}$ from their (un)corrupted sampling data $(\langle \psi, f\rangle + \epsilon_\psi), \psi \in \Psi$, where $\Psi = \{\psi\}$ is a family of sampling functional.
- Denoising via NSFBs: $\tilde{\mathbf{x}} = \mathbf{G}_0\mathbf{z}_0 + \mathbf{G}_1\tilde{\mathbf{z}}_1 = \mathbf{H}^{-1}(\mathbf{H}_0^T\mathbf{z}_0 + \mathbf{H}_1^T\tilde{\mathbf{z}}_1)$, where $\mathbf{G}_0 = \mathbf{H}^{-1}\mathbf{H}_0^T$, $\mathbf{G}_1 = \mathbf{H}^{-1}\mathbf{H}_1^T$ and $\mathbf{H} = \mathbf{H}_0^T\mathbf{H}_0 + \mathbf{H}_1^T\mathbf{H}_1$.



- Inverse filtering procedure $\mathbf{y} \longmapsto \mathbf{H}^{-1}\mathbf{y} = \mathbf{x}$, where $\mathbf{H}$ is invertible and has small geodesic width $\sigma(\mathbf{H}) \le \sigma$.
- Iterative distributed algorithm to solve the linear system $\mathbf{Hx} = \mathbf{y}$.
- 1) Iterative approximation algorithm when $\mathbf{H}$ is a polynomial of multiple graph shifts; 2) Preconditioned gradient descent algorithm; 3) Iterative divide-and-conquer algorithm

Iterative approximation algorithm

- $\mathbf{H}$ is a polynomial of multiple graph shifts,
$$\mathbf{H} = h(\mathbf{S}_1, \ldots, \mathbf{S}_d) = \sum_{l_1=0}^{L_1} \cdots \sum_{l_d=0}^{L_d} h_{l_1,\ldots,l_d} \mathbf{S}_1^{l_1} \cdots \mathbf{S}_d^{l_d}.$$
- **Assumption**: Commutative graph shifts $\mathbf{S}_k = \mathbf{U}^* \Lambda_k \mathbf{U}, 1 \le k \le d$ have their joint spectrum Λ contained in the cubic $[\boldsymbol{\mu}, \boldsymbol{\nu}] = [\mu_1, \nu_1] \times \cdots \times [\mu_d, \nu_d]$, and the polynomial $h(t_1, \ldots, t_d) = \sum_{l_1=0}^{L_1} \cdots \sum_{l_d=0}^{L_d} h_{l_1,\ldots,l_d} t_1^{l_1} \ldots t_d^{l_d}$ satisfies
$$h(\mathbf{t}) \ne 0 \ \text{ for all } \ \mathbf{t} \in [\boldsymbol{\mu}, \boldsymbol{\nu}]. \tag{11}$$
- By (11), $1/h$ is an analytic function on $[\boldsymbol{\mu}, \boldsymbol{\nu}]$, and hence it has Fourier expansion in term of shifted Chebyshev polynomials $\bar{T}_{\mathbf{k}}, \mathbf{k} \in \mathbf{Z}_+^d$,
$$1/h(\mathbf{t}) = \sum_{\mathbf{k} \in \mathbf{Z}_+^d} c_{\mathbf{k}} \bar{T}_{\mathbf{k}}(\mathbf{t}), \ \mathbf{t} \in [\boldsymbol{\mu}, \boldsymbol{\nu}].$$
- Partial sums $g_K(\mathbf{t}) = \sum_{|\mathbf{k}| \le K} c_{\mathbf{k}} \bar{T}_{\mathbf{k}}(\mathbf{t})$, converge to $1/h$ exponentially,
$$b_K := \sup_{\mathbf{t} \in [\boldsymbol{\mu}, \boldsymbol{\nu}]} |1 - h(\mathbf{t}) g_K(\mathbf{t})| \le C r_0^K, \ K \ge 0, \tag{12}$$
for some positive constants $C \in (0, \infty)$ and $r_0 \in (0, 1)$.

- Set $\mathbf{G}_K = g_K(\mathbf{S}_1, \ldots, \mathbf{S}_d),\ K \ge 0$. Define the *iterative Chebyshev polynomial approximation algorithm*, or ICPA in abbreviation,

$$\begin{cases} \mathbf{z}^{(m)} = \mathbf{G}_K \mathbf{e}^{(m-1)}; \quad \mathbf{e}^{(m)} = \mathbf{e}^{(m-1)} - \mathbf{H}\mathbf{z}^{(m)}; \\ \mathbf{x}^{(m)} = \mathbf{x}^{(m-1)} + \mathbf{z}^{(m)},\ m \ge 1, \end{cases} \tag{13}$$

with initials $\mathbf{e}^{(0)} = \mathbf{y}$ and $\mathbf{x}^{(0)} = \mathbf{0}$.

Theorem

(Emirov, Cheng, Jiang and S. arXiv 2020) Let degree $K \ge 0$ of Chebyshev polynomial approximation satisfy

$$b_K := \sup_{\mathbf{t} \in [\boldsymbol{\mu}, \boldsymbol{\nu}]} |1 - h(\mathbf{t}) g_K(\mathbf{t})| < 1. \tag{14}$$

Then $\mathbf{x}^{(m)}, m \ge 0$, converges exponentially to $\mathbf{H}^{-1}\mathbf{y}$.

- The ICPA algorithm can be implemented at the vertex level, with each step containing data exchanging among adjacent vertices and weighted linear combination of values at adjacent vertices.

Preconditioned gradient descent algorithm

- An invertible filter $\mathbf{H}$ with geodesic width $\sigma(\mathbf{H})$ is not necessarily a polynomial of some multiple graph shifts.
- Gradient descent algorithm:

$$\mathbf{x}^{(m+1)} = \mathbf{x}^{(m)} - \beta \mathbf{H}^T(\mathbf{H}\mathbf{x}^{(m)} - \mathbf{y})$$

  where the step size β is appropriately chosen. The optimal size is $\beta_{\rm opt} = 2/(\lambda_{\min}(\mathbf{H}^T(\mathbf{H}) + \lambda_{\max}(\mathbf{H}^T(\mathbf{H}))$, expensive to evaluate.
- Solve the right preconditioned linear system

$$\mathbf{H}\mathbf{Q}^{-1}\mathbf{z} = \mathbf{y} \quad \text{and} \quad \mathbf{x} = \mathbf{Q}^{-1}\mathbf{z}, \qquad (15)$$

  via the gradient descent algorithm ($\beta = 1$)

$$\begin{cases} \mathbf{z}^{(m)} = \mathbf{z}^{(m-1)} - \mathbf{Q}^{-1}\mathbf{H}^T\big(\mathbf{H}\mathbf{Q}^{-1}\mathbf{z}^{(m-1)} - \mathbf{y}\big) \\ \mathbf{x}^{(m)} = \mathbf{Q}^{-1}\mathbf{z}^{(m)}, \quad m \ge 1, \end{cases}$$

  with initial $\mathbf{z}^{(0)}$, where $\mathbf{Q}$ is a diagonal matrix such that $\mathbf{H}^T\mathbf{H} \preceq \mathbf{Q}^2$. The above iterative algorithm can be reformulated as a quasi-Newton method,

$$\mathbf{x}^{(m)} = \mathbf{x}^{(m-1)} - \mathbf{Q}^{-2}\mathbf{H}^T(\mathbf{H}\mathbf{x}^{(m-1)} - \mathbf{y}), \quad m \ge 1 \qquad (16)$$

  with initial $\mathbf{x}^{(0)}$.

Theorem

(Cheng, Emirov and Sun, IEEE SPL, 2020) Let $\mathbf{H}$ *be an invertible filter, and* $\mathbf{Q}$ *be a diagonal matrix such that* $\mathbf{H}^T\mathbf{H} \preceq \mathbf{Q}^2$*. Then the preconditioned gradient descent algorithm*

$$\mathbf{x}^{(m)} = \mathbf{x}^{(m-1)} - \mathbf{Q}^{-2}\mathbf{H}^T(\mathbf{H}\mathbf{x}^{(m-1)} - \mathbf{y}), \quad m \geq 1$$

with initial $\mathbf{x}^{(0)}$ *converges exponentially to* $\mathbf{H}^{-1}\mathbf{y}$*.*

- (Emirov, Cheng, Sun and Qu, Arxiv 2020) In the case that $\mathbf{H}$ is not necessary invertible, then the sequence $\mathbf{x}^{(m)}, m \geq 1$ in PGDA with $\mathbf{y}$ being replaced by $\mathbf{0}$ converges to eigenvector of $\mathbf{H}$ associated with eigenvalue zero if it is nonzero. The above conclusion can be used to find principal eigenvectors associated with the minimal/maximal eigenvalue of a Hermitian matrix, and also to locally evaluate principal eigenvectors of the hyperlink matrix and hence identify the local influence of a vertex on its neighborhood.

- Define a diagonal matrix $\mathbf{P_H}$ with diagonal elements $P_{\mathbf{H}}(i,i), i \in V$, given by

$$P_{\mathbf{H}}(i,i) := \max_{k \in B(i,\omega(\mathbf{H}))} \Big\{ \max\Big(\sum_{j \in B(k,\omega(\mathbf{H}))} |H(j,k)|, \sum_{j \in B(k,\omega(\mathbf{H}))} |H(k,j)| \Big) \Big\},$$

 where we denote the set of all s-hop neighbors of a vertex $i \in V$ by $B(i,s) = \{j \in V,\ \rho(j,i) \le s\},\ s \ge 0$.

- The above diagonal matrix $\mathbf{P_H}$ can be evaluated at vertex level and constructed on SDNs with communication range $L \ge \omega(\mathbf{H})$.
  - **Inputs**: Geodesic width $\omega(\mathbf{H})$ of the filter $\mathbf{H}$ and nonzero entries $H(i,j)$ and $H(j,i)$ for $j \in B(i,\omega(\mathbf{H}))$ in the i-th row and column of the filter $\mathbf{H}$.
  - **1)** Calculate $d(i) = \max\Big\{ \sum_{j \in B(i,\omega(\mathbf{H}))} |H(i,j)|, \sum_{j \in B(i,\omega(\mathbf{H}))} |H(j,i)| \Big\}$.
  - **2)** Send $d(i)$ to all neighbors $k \in B(i,\omega(\mathbf{H}))\backslash\{i\}$ and receive $d(k)$ from neighbors $k \in B(i,\omega(\mathbf{H}))\backslash\{i\}$.
  - **3)** Calculate $P_{\mathbf{H}}(i,i) = \max_{k \in B(i,\omega(\mathbf{H}))} d(k)$.
  - **Output**: $P_{\mathbf{H}}(i,i)$.
- Important observation: $\mathbf{H}^T\mathbf{H} \preceq \mathbf{P}_{\mathbf{H}}^2$.

Iterative Divide-and-Conquer Algorithm

- **Problem**: Solve the least-square problem $\|\mathbf{Hx}-\mathbf{y}\|_2^2$ in a distributed manner, where $\mathbf{H}$ is an invertible graph filter with geodesic width σ.
- *Divide*: A graph can be decomposed into a family of **overlapped** subgraphs as $\mathcal{G}=\cup_{i\in V_r}\mathcal{G}_i$, where $\mathcal{G}_i := (B(i,r), E(i,r))$ with $E(i,r)=\{(i,j), j\in B(i,r)\}$.
- *Conquer*: To solve a family of local least squares problem,

$$\mathbf{x}_{k,r} = \arg\min_{\mathbf{x}} \|\mathbf{H}\chi_k^{2r}\mathbf{x}-\mathbf{y}\|_2^2,\ k\in V_r,$$

  where $\chi_k^r : (x(i))_{i\in V} \longmapsto \big(x(i)\chi_{B(k,r)}(i)\big)_{i\in V}$ are truncation operators. The unique solution is given by

$$\mathbf{x}_{k,r} = \chi_k^{2r}\big(\chi_k^{2r}\mathbf{H}^T\mathbf{H}\chi_k^{2r}\big)^{-1}\chi_k^{2r}\mathbf{H}^T\mathbf{y}. \qquad (17)$$

- Patch all local solutions $\chi_k^r x_{k,r}, k\in V_r$ together by simple local average to update in each iteration.

Algorithm 1 Iterative Divide-and-Conquer Algorithm

Initials: $\mathbf{x}^{(0)} = \mathbf{0}$,$\mathbf{z}^{(0)} = \mathbf{y}$, set $m = 0$.

Iteration:

1) Update residue $\mathbf{z}_m = \mathbf{y} - \mathbf{H}\mathbf{x}^{(m)}$.

2) Solve local least squares problem:

$$\mathbf{w}_{k,r} = \chi_k^{2r} \left(\chi_k^{2r} \mathbf{H}^T \mathbf{H} \chi_k^{2r} \right)^{-1} \chi_k^{2r} \mathbf{H}^T \mathbf{z}_m$$

3) Patch all local solutions by simple local average to obtain $\mathbf{w}^{(m)}$, and then update $\mathbf{x}^{(m+1)} = \mathbf{x}^{(m)} + \mathbf{w}^{(m)}$.

4) Evaluate $\|\mathbf{x}^{(m+1)} - \mathbf{x}^{(m)}\|_\infty \le \varepsilon$. If yes, terminate the iteration and output $\mathbf{x}^{(m+1)}$ as the reconstructed signal. Otherwise, set $m = m + 1$.

Theorem

(Cheng, Jiang and S., 2019 ACHA) Let $1 \le p \le \infty$, and $\mathbf{H}$ be an invertible filter with geodesic width σ. If radius r is appropriately selected, then the solution $\mathbf{x}^{(m)}, m \ge 0$, of the iterative divide-and-conquer algorithm converges to $\mathbf{H}^{-1}\mathbf{y}$ exponentially,

$$\|\mathbf{x}^{(m)} - \mathbf{H}^{-1}\mathbf{y}\|_p \le (\delta_{r,\sigma})^m \|\mathbf{H}^{-1}\mathbf{y}\|_p, \ m \ge 0, 1 \le p \le \infty, \qquad (18)$$

where $\delta_{r,\sigma} \in (0,1)$.

- Solutions of local optimization problems provide good local approximations to the solution of the global optimization problem.
- Nazar Emirov, Guohui Song and S. (2021) introduced an iterative divide-and-conquer algorithm for the localized minimization problem

$$\min F(\mathbf{x}) = \sum_{i \in V} f_i(\mathbf{x}),$$

where $f_i, i \in V$, are convex functions depends on variables in the σ-neighborhood of the vertex i only.

10 Summary

- Mathematical signal processing on graphs is an emerging interdisciplinary field. We see many amazing developments in last ten years, however there is still a huge research gap between mathematical theory and engineering practice.
- Distributed algorithms provide indispensable tools for data processing on graphs, and their implementations at the vertex level are crucial for many real-world problems on networks of large size.
- In this talk, I present some of my understanding on filtering and inverse filtering, graph filter banks, local certificate for stability, and distributed algorithms for inverse filtering.


thank you
danke
謝謝
рахмат
ngiyabonga
teşekkür ederim
gracias
Баярлалаа
спасибо
faafetai lava
mersi
barka
welalin
tack
spas
dank je
misaotra
matondo
paldies
grazzi
mahalo
tapadh leat
хвала
asante
manana
obrigada
tenki
vinaka
спасибі
blagodaram
kiitos
dankie
dhanyavad
mauruuru
köszönöm
hvala
grazie
bayarlalaa
bedankt
enkosi
dziękuję
djiere dieuf
mochchakkeram
mamnun
дякую
chnorakaloutioun
gratias ago
gràcies
sulpáy
go raibh maith agat
sobodi
dėkui
sukriya
kop khun krap
taiku
arigatô
takk
dakujem
trugarez
obrigado
mèsi
sağolun
najis tuke
ありがとう
tanemirt
grazie
rahmet
diolch
dhanyavadagalu
shukriya
мерси
kam sah hamnida
terima kasih
감사합니다
xièxie
ευχαριστώ
merci


PEACE
MONEY
2021
GOOD LUCK
Friendship
SATISFACTION
FRIENDSHIP
2021
LOVE
Happiness
Peace
Satisfaction
Success
Health
Money
Wish you
all the best
for 2021
HAPPY NEW YEAR
Love
Success
Health
2021