DEMIXING AND BLIND DECONVOLUTION OF GRAPH-DIFFUSED SPARSE SIGNALS

Fernando J. Iglesias†, Santiago Segarra‡, Samuel Rey-Escudero†, Antonio G. Marques†, David Ramirez*

†Signal Theory & Communications, King Juan Carlos University ‡Institute for Data, Systems and Society, Massachusetts Institute of Technology *Signal Theory & Communications, Carlos III University of Madrid

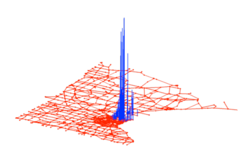
Universidad Rey Juan Carlos

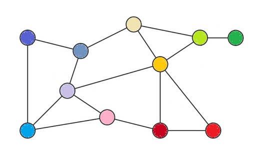
Abstract

We extend the classical joint problem of signal demixing, blind deconvolution, and filter identification to the realm of graphs. The model is that each mixing signal is generated by a sparse input diffused via a graph filter. Then, the sum of diffused signals is observed. We identify and address two problems: 1) each sparse input is diffused in a different graph; and 2) all signals are diffused in the same graph. These tasks amount to finding the collections of sources and filter coefficients producing the observation.

Graph signal processing - 101

- Graph $G = (\mathcal{V}, \mathcal{E})$: encode pairwise relationships
- Interest is not in G itself, but in data associated with nodes in $\mathcal{V}$
- Ex: Opinion profile, buffer congestion, neural activity, epidemic



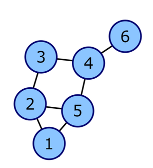


$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_{|\mathcal{V}|} \end{bmatrix} = \begin{bmatrix} 0.6 \\ \vdots \\ 0.7 \end{bmatrix}$$

- Graph SP: broaden SP to graph signals, well suited to netw. process.

Graph signals and graph-shift operator

- Graph signals vector $\mathbf{x} \in \mathbb{R}^N$ (with $|\mathcal{V}| = N$)
- Graph G is endowed with a graph-shift operator $\mathbf{S}$
  - $\Rightarrow$ Matrix $\mathbf{S} \in \mathbb{R}^{N \times N}$ satisfying: $S_{ij} = 0$ for $i \neq j$ and $(i, j) \notin \mathcal{E}$



$$\mathbf{S} = \begin{pmatrix} S_{11} & S_{12} & 0 & 0 & S_{15} & 0 \\ S_{21} & S_{22} & S_{23} & 0 & S_{25} & 0 \\ 0 & S_{32} & S_{33} & S_{34} & 0 & 0 \\ 0 & 0 & S_{43} & S_{44} & S_{45} & S_{46} \\ S_{51} & S_{52} & 0 & S_{54} & S_{55} & 0 \\ 0 & 0 & 0 & S_{64} & 0 & S_{66} \end{pmatrix}$$

S captures local structure in G

- Ex: Adjacency **A**, Degree **D** and Laplacian **L**

Locality of S and frequency-domain representation

- **S** is a local linear operator $\Rightarrow$ If $\mathbf{y} = \mathbf{Sx}$, $y_i = \sum_{j \in \mathcal{N}_i} S_{ij} x_j \Rightarrow$ 1-hop info
- Spectrum of **S** useful to analyze **x** $\Rightarrow$ diagonalizable $\mathbf{S} = \mathbf{V \Lambda V}^{-1}$
- Leverage **S** to define graph Fourier transform (GFT) and iGFT

$$\tilde{\mathbf{x}} = \mathbf{V}^{-1}\mathbf{x}, \qquad \mathbf{x} = \mathbf{V}\tilde{\mathbf{x}} \qquad \text{(Ex: DFT, PCA)}$$

- Key message: the two basic elements of GSP are **x** and **S**

Linear (shift-invariant) graph filter

- With coeff. $\mathbf{h} = [h_0, \ldots, h_L]^T$, then **H** is a graph filter if

$$\mathbf{H} := h_0\mathbf{S}^0 + h_1\mathbf{S}^1 + \ldots + h_L\mathbf{S}^L = \sum_{l=0}^{L} h_l \mathbf{S}^l$$

- Key properties: **H** diagonalized by **V**, distr. (L-hop) implementation
- If $\mathbf{y} = \mathbf{Hx}$, then $\tilde{\mathbf{y}} = diag(\tilde{\mathbf{h}})\tilde{\mathbf{x}}$, with the frequency response being

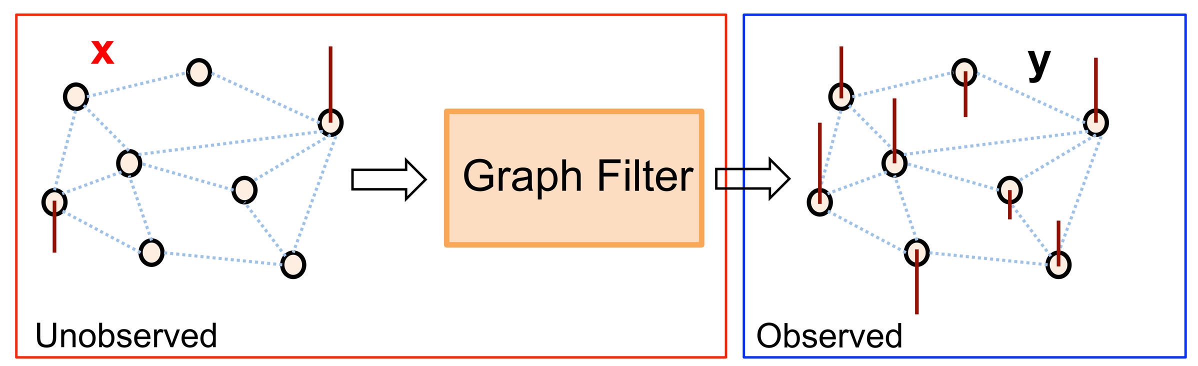
$$\tilde{\mathbf{h}} = \mathbf{\Psi h}, \text{ where } \mathbf{\Psi} := \begin{pmatrix} 1 & \lambda_1 & \ldots & \lambda_1^L \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_N & \ldots & \lambda_N^L \end{pmatrix}$$

Diffused sparse graph signals

- Q: Upon observing a graph signal **y**, how was this signal generated?
- Postulate the following generative model
  - $\Rightarrow$ An originally sparse signal $\mathbf{x} = \mathbf{x}^{(0)}$
  - $\Rightarrow$ Diffused via linear graph dynamics $\mathbf{S} \Rightarrow \mathbf{x}^{(l)} = \mathbf{Sx}^{(l-1)}$
  - $\Rightarrow$ Observed **y** is a linear combination of the diffused signals $\mathbf{x}^{(l)}$

$$\mathbf{y} = \sum_{l=0}^{L} h_l \mathbf{x}^{(l)} = \sum_{l=0}^{L} h_l \mathbf{S}^l \mathbf{x} = \mathbf{Hx}$$

- View few elements in $\text{supp}(\mathbf{x}) =: \{i : x_i \neq 0\}$ as sources or seeds



Classical blind demixing

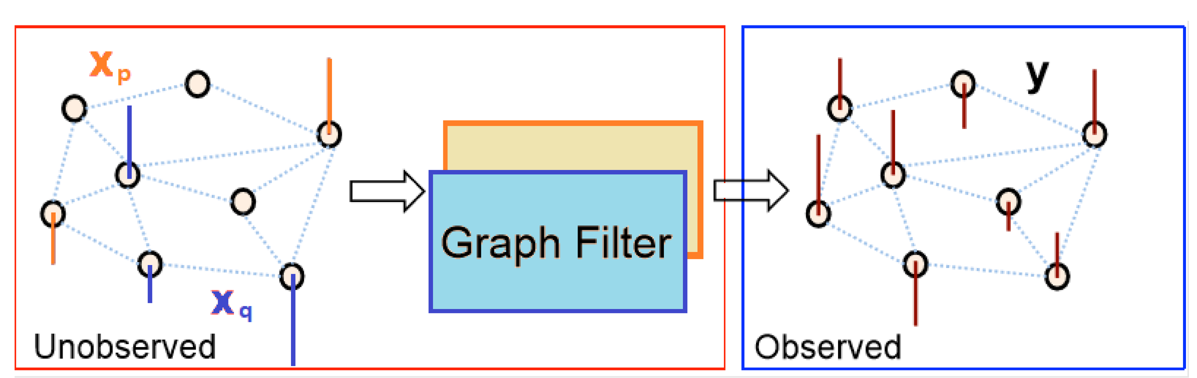
- Unknown signals x_p and filters h_p [Ling-Strohmer17]

$$\mathbf{y} = \sum_{p=1}^{P} \mathbf{x}_p * \mathbf{h}_p$$

  - $\Rightarrow$ Only one observation available
  - $\Rightarrow$ Mixture of blind deconvolutions"
- Undetermined without further assumptions
  - $\Rightarrow$ Probabilistic priors, subspace models
- Natural model of multi-sensor, single-receiver
- Many practical applications
  - $\Rightarrow$ E.g.: neuroscience, spectroscopy, astronomy

Graph blind demixing formulation

- Q: Can we determine the signals $\{\mathbf{x}_p\}_{p=1}^P$ and filters $\{\mathbf{h}_p\}_{p=1}^P$ from $\mathbf{y} = \sum_{p=1}^P \mathbf{H}_p \mathbf{x}_p$?



- Problem: Blind identification of graph filters with sparse inputs
  - $\Rightarrow$ Generalizes classical blind demixing to graphs
- Ill-posed $\Rightarrow N\,P + \sum_{p=1}^{P}(L_p + 1)$ unknowns and N observations
  - $\Rightarrow$ Assume $\mathbf{x}_p$ is Q_p-sparse i.e., $\|\mathbf{x}_p\|_0 := |\text{supp}(\mathbf{x}_p)| = Q_p$
- We address two different setups:
  - $\Rightarrow$ Multi-graph: $\mathbf{H}_p = \sum_{l=0}^{L_p} h_{p,l} \mathbf{S}_p^l$
  - $\Rightarrow$ Single-graph: $\mathbf{H}_p = \sum_{l=0}^{L_p} h_{p,l} \mathbf{S}^l$

"Lifting" the bilinear inverse problem

- Leverage the frequency response of graph filters
  - $\Rightarrow$ Let $\mathbf{U}_p := \mathbf{V}_p^{-1}$ and $\mathbf{y}_p = \mathbf{H}_p \mathbf{x}_p$:

$$\mathbf{y}_p = \mathbf{V}_p \text{diag}(\mathbf{\Psi}_p \mathbf{h}_p) \mathbf{U}_p \mathbf{x}_p$$

  - $\Rightarrow$ $\mathbf{y}_p$ is a bilinear function of $\mathbf{h}_p$ and $\mathbf{x}_p$
- Multi-graph blind demixing $\Rightarrow$ Non-convex feasibility problem

$$\text{find} \quad \{\mathbf{x}_p, \mathbf{h}_p, \mathbf{y}_p\}_{p=1}^P \quad \text{s. to } \mathbf{y}_p = \mathbf{V}_p \text{diag}(\mathbf{\Psi}_p \mathbf{h}_p) \mathbf{U}_p \mathbf{x}_p, \ \mathbf{y} = \sum_{p=1}^{P} \mathbf{y}_p, \ \|\mathbf{x}_p\|_0 \leq Q_p$$

- Key observation: Using the Khatri-Rao product $\odot$, rewrite $\mathbf{y}_p$ [Segarra17]

$$\mathbf{y}_p = \mathbf{V}_p \left(\mathbf{\Psi}_p^T \odot \mathbf{U}_p^T\right)^T \text{vec}(\mathbf{x}_p \mathbf{h}_p^T) = \mathbf{M}_p \text{vec}(\mathbf{x}_p \mathbf{h}_p^T)$$

  - $\Rightarrow$ Reveals $\mathbf{y}_p$ is a linear combination of the entries of $\mathbf{Z}_p := \mathbf{x}_p \mathbf{h}_p^T$
- Matrices $\mathbf{Z}_p$ are rank-1 and row-sparse

$$\text{find } \{\mathbf{Z}_p\}_{p=1}^P \quad \text{s. to } \mathbf{y} = \sum_{p=1}^{P} \mathbf{M}_p \text{vec}(\mathbf{Z}_p), \ \text{rank}(\mathbf{Z}_p) = 1, \ \|\mathbf{Z}_p\|_{2,0} \leq Q_p$$

  - $\Rightarrow$ Pseudo-norm $\|\mathbf{Z}_p\|_{2,0}$ counts the non-zero rows of $\mathbf{Z}_p$

Algorithmic approach via convex relaxation

- Rank minimization s. to row-cardinality constraints is NP-hard. Relax!
  - $\Rightarrow$ Nuclear norm $\|\mathbf{Z}_p\|_* := \sum_k \sigma_k(\mathbf{Z}_p)$ a convex proxy of rank
  - $\Rightarrow$ $\ell_{2,1}$ mixed norm $\|\mathbf{Z}_p\|_{2,1} := \sum_{n=1}^N \|\mathbf{z}_{p,n}^T\|_2$ surrogate of $\|\mathbf{Z}_p\|_{2,0}$
- Convex relaxation

$$\min_{\{\mathbf{Z}_p\}_{p=1}^P} \sum_{p=1}^{P} \eta_p \|\mathbf{Z}_p\|_* + \sum_{p=1}^{P} \beta_p \|\mathbf{Z}_p\|_{2,1} \quad s.to \quad \mathbf{y} = \sum_{p=1}^{P} \mathbf{M}_p \text{vec}(\mathbf{Z}_p)$$

- More sophisticated relaxations [Ramirez17]
  - $\Rightarrow$ Sparse reconstruction with iterative-reweighed optimization
  - $\Rightarrow$ Semidefinite embedding lemma

Single-graph demixing

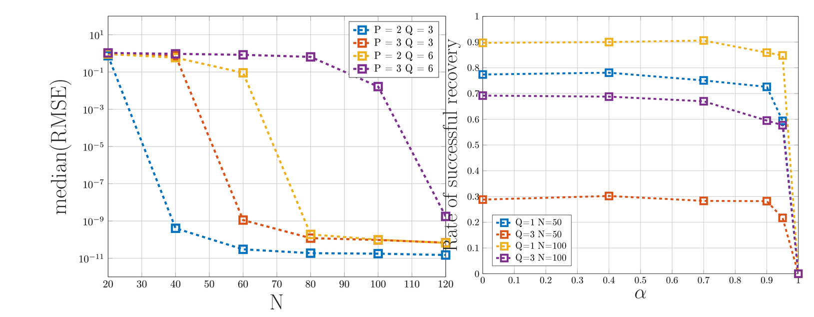
- More challenging case than multi-graph
- Generally non-identifiable, only $\mathbf{Z} = \Sigma_p \mathbf{Z}_p$
  - $\Rightarrow$ Argument in the paper
- We propose a two-step approach
  1. Find $\mathbf{Z}^*$ – rank$(\mathbf{Z}_p) \geq 1$ "multi-graph" demixing with $P = 1$
  2. SVD $\mathbf{Z}^*$ for $\{\mathbf{Z}^*\}_{p=1}^P$ with orthogonal inputs and filters

Summary of additional aspects

- If only few samples of the observation are available, introduce sampling matrices
- Robust demixing: relax the equality constraint in the presence of noise
- Prior information
  - $\Rightarrow$ *Probabilistic* distributions in the objective
  - $\Rightarrow$ *Deterministic* knowledge in terms of constraints. E.g:
    1. $\{\mathbf{h}\}_{p=1}^P$ known removes bilinearity
    2. Known values of $\mathbf{x}_p \Rightarrow$ row-equality constraints

Demixing in random graphs

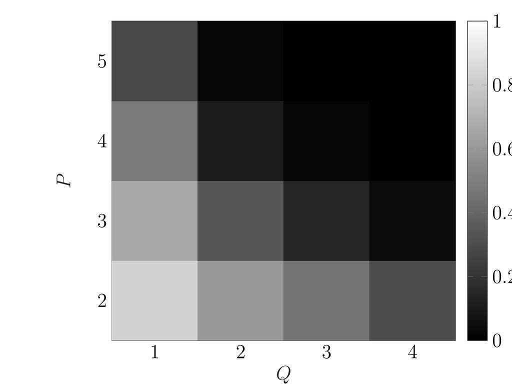
- Recovery rates on Erdős-Rényi graphs ($N = 50$) for varying P and Q ($\{Q_p = Q, P_p = P\}_{p=1}^P$), $L = 2$
- single-graph (left), two coupled graphs (right)



- Left: ($P = 3, Q = 3$) harder than ($P = 2, Q = 6$) $\Rightarrow$ Q is critical
- Right: two coupled graphs ($\alpha = 1$ equal, $\alpha = 0$ random)
  - $\Rightarrow$ Recovery is maintained for large coupling: $\alpha \approx 0.7$
  - $\Rightarrow$ Topology is central!

Demixing in brain graphs

- Graphs ($N = 66$) representing the brain anatomy of several individuals [Hagmann08]



- Feasible demixing even for real-world graphs
  - $\Rightarrow$ Expected performance decay for increasing P and Q

Discussion and road ahead

- Identifiability conditions
  - $\Rightarrow$ Q: When is $\{\mathbf{x}_p, \mathbf{h}_p\}_{p=1}^P$ the unique solution (up to scaling)?
  - $\Rightarrow$ Deterministic or probabilistic model assumptions
- Exact recovery conditions
  - $\Rightarrow$ Q: When does the convex relaxation succeed? Hypotheses:
  - $\Rightarrow$ Lower bound on N to guarantee recovery for given P and Q
  - $\Rightarrow$ Dependence on algebraic features of the graph-shift **S**
  - $\Rightarrow$ Some graph topologies are more amenable
- Envisioned application domains
  - $\Rightarrow$ Opinion formation in social networks
  - $\Rightarrow$ Event-driven information cascades
  - $\Rightarrow$ Identify sources of abnormal brain activity

References

- S. Ling and T. Strohmer, "Blind deconvolution meets blind demixing: algorithms and performance bounds," *IEEE Trans. Information Theory*, vol. 67, no. 7, pp. 4497–4520, 2017.
- S. Segarra, G. Mateos, A. G. Marques and A. Ribeiro, "Blind identification of graph filters," *IEEE Trans. Signal Processing*, vol. 65, no. 5, pp. 1146–1159, 2017.
- D. Ramirez, A. G. Marques and S. Segarra, "Graph-signal reconstruction and blind deconvolution for diffused sparse inputs," *IEEE Int. Conf. Acoustics, Speech and Signal Processing*, Mar. 2017.
- P. Hagmann, et al., "Mapping the structural core of human cerebral cortex," *PLoS Biol*, vol. 6, no. 7, pp. e159, 2008.