Interpolation of graph signals using shift-invariant graph filters

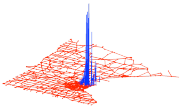
Santiago Segarra[1], Antonio G. Marques[2], Geert Leus[3], & Alejandro Ribeiro[1]

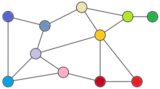
[1]University of Pennsylvania (USA)
[2]King Juan Carlos University (Spain)
[3]Delft University of Technology (Netherlands)

https://www.seas.upenn.edu/~ssegarra/

EUSIPCO, Nice, September 1, 2015

- Emergence of network science and big data
- Networks and graphs: structures that encode pairwise relationships
- Our interest, not in network itself, but in data associated with nodes
  - ⇒ The object of study is a graph signal
- Graph SP: need to extend classical SP results to graph signals
  - ⇒ Modify existing algorithms, gain intuition on concepts preserved/lost



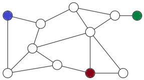


Graph $G = (\mathcal{V}, \mathcal{E}, W)$

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

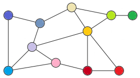
Reconstruction of banlimited graph signals

- Many relevant GSP problems: filter design, sampling, blind deconvolution
- Our focus in this paper: reconstruction of bandlimited graph signals
- Most related problems:
  - ⇒ Estimate the *unknown* signal $\mathbf{y}$ by observing a subset of nodes
- Our problem:
  - ⇒ Reconstruct the *known* signal $\mathbf{y}$ by acting on a subset of nodes
  - ⇒ Injection of a sparse signal followed by a low-pass graph filter



⇒ GRAPH FILTER ⇒

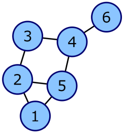
$$\mathbf{y} = \mathbf{H}\mathbf{x}$$



- Not only theoretical merits, also practical relevance
  - ⇒ Graph filters ⇒ percolation of local information
  - ⇒ Distributed nets, opinion formation, biological percolation processes
- Before being more specific: review of graph signals and filters

Graph signals and need for graph topology

- (Node) graph signals are mappings $x : \mathcal{V} \rightarrow \mathbb{R}$
  - $\Rightarrow$ May be represented as a vector $\mathbf{x} \in \mathbb{R}^N$ (with $|\mathcal{V}| = N$)
  - $\Rightarrow$ DSP can be seen as a particular case of GSP $\Rightarrow$ directed cycle graph
- Graph $G = (\mathcal{V}, \mathcal{E}, W)$ is endowed with a graph-shift operator $\mathbf{S}$
  - $\Rightarrow$ Can be represented as a matrix $\mathbf{S} \in \mathbb{R}^{N \times N}$ satisfying:
  - $\Rightarrow$ $S_{ij} = 0$ for $i \neq j$ and $(i, j) \notin \mathcal{E}$ (captures local structure in G)
  - $\Rightarrow$ $\mathbf{S}$ can take nonzero values in the edges of G or in its diagonal



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

- Examples: Adjacency $\mathbf{A}$, Degree $\mathbf{D}$ and Laplacian $\mathbf{L}$

Locality of $\mathbf{S}$ and frequency interpretation

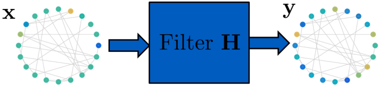
- $\mathbf{S}$ is a local linear operator, i.e., if $\mathbf{y} = \mathbf{S}\mathbf{x}$
  $\Rightarrow y_i = \sum_j S_{ij} x_j = \sum_{j \in \mathcal{N}_i^+} S_{ij} x_j \Rightarrow$ only to 1-hop info
  $\Rightarrow$ if $\mathbf{z} = \mathbf{S}^2\mathbf{x} \Rightarrow \mathbf{z} = \mathbf{S}\mathbf{y} \Rightarrow 2$-hop info
- $\mathbf{S}$ (spectrum) useful to analyze $\mathbf{x}$, here diagonalizable shifts $\mathbf{S} = \mathbf{V}\Lambda\mathbf{V}^{-1}$
  $\Rightarrow \mathbf{V} = [\mathbf{v}_1 \mathbf{v}_2 \ldots \mathbf{v}_N]$ eigenvectors; Λ eigenvalues; if normal, $\mathbf{V}^{-1} = \mathbf{V}^H$
- Good transformations graph signals? Leverage $\mathbf{S}$ to define GFT and iGFT

$$\tilde{\mathbf{x}} = \mathbf{V}^{-1}\mathbf{x} \qquad\qquad \mathbf{x} = \mathbf{V}\tilde{\mathbf{x}}$$

  $\Rightarrow$ Bandlimited signals: $\tilde{\mathbf{x}}$ sparse; particular cases: DFT, PCA
- Key message: the two basic elements of GSP are $\mathbf{x}$ and $\mathbf{S}$

- A graph filter $H : \mathbb{R}^N \to \mathbb{R}^N$ is a map between graph signals

Focus on linear filters
$\Rightarrow$ map represented as an $N \times N$ matrix



- Filter $\mathbf{H}$ is a polynomial on $\mathbf{S}$ with coefficients h_l and degree L

$$\mathbf{H} := h_0\mathbf{S}^0 + h_1\mathbf{S}^1 + h_2\mathbf{S}^2 + \ldots = \sum_{l=0}^{L} h_l\mathbf{S}^l$$

- A graph filter represents a linear transformation that
  - $\Rightarrow$ Accounts for local structure of the graph
  - $\Rightarrow$ Can be implemented distributedly in L steps
  - $\Rightarrow$ Only requires information in the L-neighborhood

- Using $\mathbf{S} = \mathbf{V}\Lambda\mathbf{V}^{-1}$, we may write $\mathbf{H} = \sum_{l=0}^{L} h_l \mathbf{S}^l = \mathbf{V}\left(\sum_{l=0}^{L} h_l \Lambda^l\right)\mathbf{V}^{-1}$
- Since Λ^l are diagonal, the GFT-iGT can be used to write $\mathbf{y} = \mathbf{H}\mathbf{x}$ as

$$\tilde{\mathbf{y}} = diag(\tilde{\mathbf{h}})\tilde{\mathbf{x}}$$

$\Rightarrow$ Output at frequency k depends only on input at frequency k

$\Rightarrow$ $\tilde{\mathbf{h}}$ is the frequency response of the filter $\mathbf{H}$

- Clearly $\tilde{h}_k = \sum_{l=0}^{L} h_l \lambda_k^l$, hence one can obtain $\tilde{\mathbf{h}}$ as $\tilde{\mathbf{h}} = \mathbf{\Psi}\mathbf{h}$, where

$$\mathbf{\Psi} := \begin{pmatrix} 1 & \lambda_1 & \dots & \lambda_1^L \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_N & \dots & \lambda_N^L \end{pmatrix}$$

$\Rightarrow$ Since $\mathbf{\Psi}$ is Vandermonde, invertible if $\lambda_k \neq \lambda_{k'}$ $\Rightarrow$ $\mathbf{h} = \mathbf{\Psi}^{-1}\tilde{\mathbf{h}}$

$\Rightarrow$ To be leveraged when designing (low-pass) graph filters

- Note that GFT for signals and filters is not the same

Revisiting our motivation

- We want to reconstruct a (K-bandlimited) graph signal $\mathbf{y}$
- Most existing problems
  - ⇒ Estimate the unknown signal $\mathbf{y}$ by observing a subset of nodes
- Our problem
  - ⇒ Reconstruct the known signal $\mathbf{y}$ by acting on a subset of nodes
- Approach: design a sparse input that is percolated by a graph filter
  - ⇒ We act on a node by injecting signal values
  - ⇒ Distributed implementation
- Examples include the reconstruction of:
  - ⇒ Global opinion in a social net by influencing a few people
  - ⇒ Brain state by exciting a few brain regions

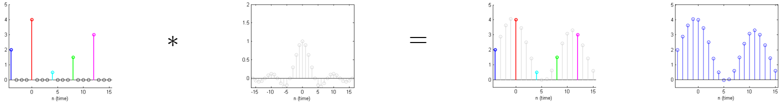
Operation and problem statement

Operation: The reconstruction scheme proceeds in two phases

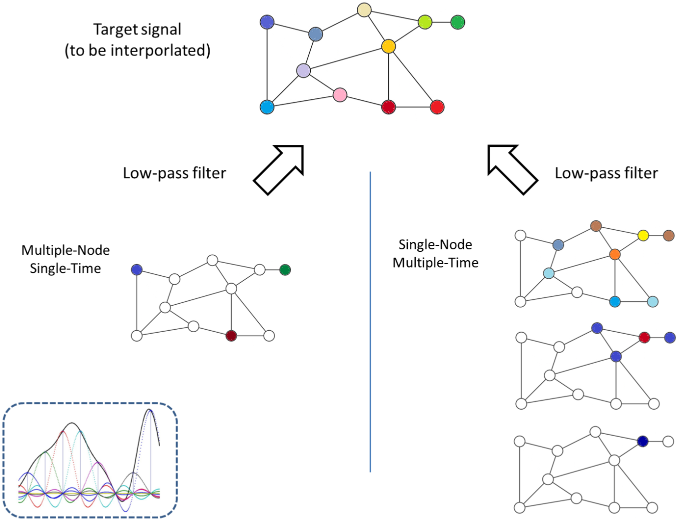
1. **Seeding phase**
   - Output is a sparse signal $\mathbf{x}$
     - $\Rightarrow$ In its simplest form we can act directly on $\{x_i\}_{i \in \mathcal{P}}$
     - $\Rightarrow$ Single seeding node, injects scalars $\{s^t\}_{t \in \mathcal{P}}$ diffused by $\mathbf{S}$ to form $\mathbf{x}$
2. **Filtering phase**
   - Use $\mathbf{x}$ as input
   - Apply a low-pass graph filter $\mathbf{H}$ with freq. response $\tilde{\mathbf{h}} = [\tilde{\mathbf{h}}_K^T \,, \mathbf{0}]^T$
   - Obtain the output signal $\mathbf{z} := \mathbf{H}\mathbf{x}$

Problem statement: How to design $\mathbf{x}$ and $\mathbf{H}$ such that $\mathbf{z} = \mathbf{y}$?

- Resembles (uniform) time interpolation



Illustrating percolation



- A single seeding signal $\mathbf{x}$ with P seeding nodes $\mathbf{x} = [\mathbf{x}_P^T, \mathbf{0}]^T$
- Goal of $\mathbf{y} = \mathbf{H}\mathbf{x}$, rewritten in the frequency domain

$$\tilde{\mathbf{y}} = \mathbf{V}^{-1}\mathbf{H}\mathbf{x} = \mathbf{V}^{-1}\mathbf{V}\text{diag}(\mathbf{\Psi}\mathbf{h})\mathbf{V}^{-1}\mathbf{x} = \text{diag}\big(\mathbf{\Psi}\mathbf{h}\big)\mathbf{V}^{-1}\mathbf{E}_P\mathbf{x}_P$$

$\Rightarrow$ Bilinear problem in $\mathbf{x}_P$ and $\mathbf{h}$ $\qquad$ $\mathbf{E}_P$ first P canon. vecs.

- We split the above system of linear equations into two

$$\tilde{\mathbf{y}}_K = \mathbf{E}_K^T \text{diag}(\mathbf{\Psi}\mathbf{h})\,\mathbf{V}^{-1}\mathbf{E}_P\mathbf{x}_P, \tag{1}$$

$$\mathbf{0}_{N-K} = \bar{\mathbf{E}}_K^T \text{diag}(\mathbf{\Psi}\mathbf{h})\,\mathbf{V}^{-1}\mathbf{E}_P\mathbf{x}_P. \tag{2}$$

- Equation (2) holds for every K-bandlimited signal $\mathbf{y}$

$\Rightarrow$ Design $\mathbf{h}$ to solve (2) and $\mathbf{x}_P$ to solve (1)

- If degree of $\mathbf{H}$ no smaller than distinct eigenvalues in $\{\lambda_i\}_{i=K+1}^{N}$

$\Rightarrow$ $\mathbf{h}^*$ solving (2) can always be found

- Perfect reconstruction if we can solve $\tilde{\mathbf{y}}_K = \text{diag}(\tilde{\mathbf{h}}_K^*)\, \mathbf{E}_K^T \mathbf{V}^{-1} \mathbf{E}_P \mathbf{x}_P$

Perfect reconstruction in MN-ST seeding

Perfect reconstruction of $\mathbf{y}$ is guaranteed via MN-ST seeding if:
i) $\lambda_{k_1} \neq \lambda_{k_2}$ for all $(\lambda_{k_1}, \lambda_{k_2})$ such that $k_1 \leq K$ and $k_2 > K$,
ii) $\text{rank}(\mathbf{E}_K^T \mathbf{V}^{-1} \mathbf{E}_P) \geq K$.

- Condition *i)* ensures that $\mathbf{h}^*$ does not eliminate any active frequency
- Condition *ii)* requires at least $P \geq K$
  - $\Rightarrow$ Additional seeding nodes can reduce the low-pass filter degree
- Seeding $\mathbf{x}_P^* = (\mathbf{E}_K^T \mathbf{V}^{-1} \mathbf{E}_P)^{-1} \text{diag}^{-1}(\tilde{\mathbf{h}}_K^*) \tilde{\mathbf{y}}_K$.
  - $\Rightarrow$ In general, $\mathbf{x}_P^* \neq \mathbf{y}_P$
  - $\Rightarrow$ Contrasts with sinc interpolation after uniform sampling

- A single node (say 1) injects P scalar seeding signals s^t, one per time t
- Collect those P signals in $\mathbf{s}_P := [s^{P-1}, \ldots, s^0]^T$ and define $\mathbf{s}^t = [\mathbf{s}_P]_t \mathbf{e}_1$
- Each seed percolated using $\mathbf{S}$ $\Rightarrow$ output of the seeding phase is

$$\mathbf{x} = \sum_{t=1}^{P} \mathbf{S}^{t-1}\mathbf{s}^t = \sum_{t=1}^{P} [\mathbf{s}_P]_t \mathbf{S}^{t-1}\mathbf{e}_1$$

- Like a filter with input $\mathbf{e}_1$ and coefficients $\mathbf{s}_P$

$$\tilde{\mathbf{x}} = \text{diag}(\mathbf{\Psi}\mathbf{s}_P)\tilde{\mathbf{e}}_1 = \text{diag}(\tilde{\mathbf{e}}_1)\mathbf{\Psi}\mathbf{s}_P$$

- U_1 # zeros in $\{[\tilde{\mathbf{e}}_1]_k\}_{k=1}^K$ and D_1 # of repeated values in $\{\lambda_k\}_{k=1}^K$

Perfect reconstruction in SN-MT seeding

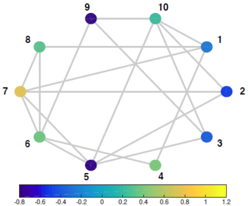
Perfect reconstruction of $\mathbf{y}$ is guaranteed via SN-MT seeding if:
i) $\lambda_{k_1} \neq \lambda_{k_2}$ for all $(\lambda_{k_1}, \lambda_{k_2})$ such that $k_1 \leq K$ and $k_2 > K$,
ii) $U_1 = 0$ and $D_1 = 0$.

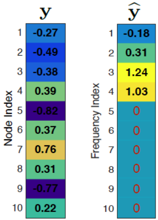
- Seeding $\mathbf{s}_P^* = \mathbf{\Psi}^{-1}\text{diag}^{-1}(\tilde{e}_1)\text{diag}^{-1}(\tilde{\mathbf{h}}_K^*)\tilde{\mathbf{y}}_K$
  - $\Rightarrow$ $U_1 = 0$: seeding node acts on every active frequency
  - $\Rightarrow$ $D_1 = 0$: every active frequency is distinguishable from each other

- SN-MT recovery depends on $\mathbf{\Lambda}$ and rows of $\mathbf{V}^{-1}$ being non-zero
  - $\Rightarrow$ Easy to look for a good seeding node: as in sampling [Marques15]
- MN-ST recovery depends on rank of submatrix of $\mathbf{V}^{-1}$
  - $\Rightarrow$ No clear way to check a priori: as in sampling [Cheng15]
- Extensions to MN-MT seeding developed too
- Approximate (imperfect) reconstruction settings
  - $\Rightarrow$ Insufficient amount of seeding values or filter degree
  - $\Rightarrow$ Noisy seeding value injections
- The set of seeding nodes has a significant impact on robustness
  - $\Rightarrow$ Optimal design to minimize mean or worst-case error

Illustrating perfect recovery of a bandlimited graph signal

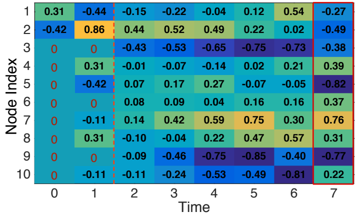
- Erdős-Rényi graph with $p = 2$ and $N = 10$

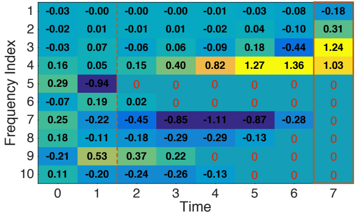




Bandwidth: $K = 4$
$\Rightarrow P \geq 4$ seeds required

- Evolution of the signal (space and frequency) for every shift

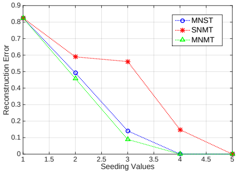




- Perfect recovery for MN-MT seeding

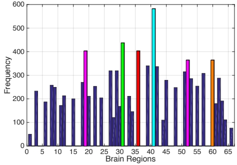
Opinion formation

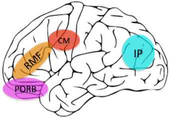
- Induce a desired opinion profile
  - ⇒ Zachary's Karate club graph
- Study robustness of different seeding strategies
  - ⇒ Insufficient seeding nodes
- Better to convince several people once (MN-ST)
  - ⇒ than same person multiple times (SN-MT)

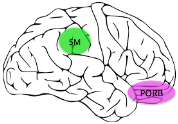


Brain state induction

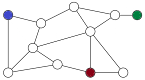
- Take brain to desired neurological state by exciting a few regions
- Study robustness of different seeding sets
  - ⇒ Noise in the signal injection
- Corroborate neurophysiological meaning of the findings



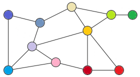




- Specific contribution in this paper
  - ⇒ Reconstruction of bandlimited graph signals using sparse inputs
  - ⇒ Sparse input followed by a graph filter
  - ⇒ Conditions for recovery, differences with time interpolation
  - ⇒ Imperfect reconstruction



⇒ GRAPH FILTER ⇒



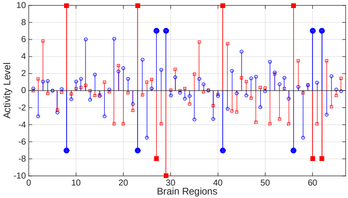
- From a more general point of view
  - ⇒ Decoupling betw. estimating unobserved values and low-pass filtering
  - ⇒ Graph filters can be viewed as linear network operators
  - ⇒ Strong relation between GSP and diffusion/percolation processes

Back-up slides

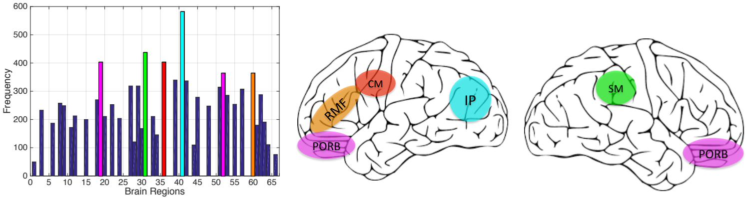
- Brain graph obtained from 66 anatomical regions [Hagmann08]
- Brain signal **y** is a desired brain activity pattern
- $\mathbf{S} = \mathbf{A}$ models a linear evolution of brain activity
- Node injections via transcranial magnetic stimulation (TMS)

- Take the brain from a rest state

  $\Rightarrow$ To one of high-cognitive activity

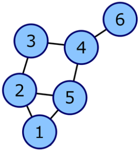


- Fix 6 seeding nodes and inject 11 values in each
  - ⇒ Since N = 66, no need for filtering phase
  - ⇒ Unclear how to apply a filter in a brain system
- Injections are noisy, look for robust combinations of seeding nodes
  - ⇒ Histogram of regions in robust seeding configurations



- Regions with 0 frequency are those inaccessible via TMS
- Either right or left Pars Orbitalis (PORB) in most robust injections

- $\mathbf{x}=[-1,2,0,0,0,0]^T$, $\mathbf{h}=[1,1,0.5]^T$, $\mathbf{y}=(\sum_{l=0}^{L} h_l \mathbf{S})\mathbf{x}=\sum_{l=0}^{L} h_l \mathbf{x}^{(l)}$



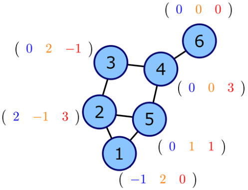
$$\mathbf{S}=\mathbf{A}=\begin{pmatrix} 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

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

$$\Downarrow$$

$$\mathbf{y}=h_0\mathbf{x}^{(0)}+h_1\mathbf{x}^{(1)}+h_2\mathbf{x}^{(2)}$$

Given $\mathbf{x}=[-1,2,0,0,0,0]^T$ and $\mathbf{h}=[1,1,0.5]^T \Rightarrow$ Find $\{\mathbf{x}^{(0)},\mathbf{x}^{(1)},\mathbf{x}^{(2)}\} \Rightarrow$ Find $\mathbf{y}$



$$\mathbf{x}^{(0)}=\mathbf{x}=\begin{pmatrix} -1 \\ 2 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \mathbf{x}^{(1)}=\mathbf{S}\mathbf{x}^{(0)}=\begin{pmatrix} 2 \\ -1 \\ 2 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad \mathbf{x}^{(2)}=\mathbf{S}\mathbf{x}^{(1)}=\begin{pmatrix} 0 \\ 3 \\ -1 \\ 3 \\ 1 \\ 0 \end{pmatrix}$$

$$\mathbf{y}=1\mathbf{x}^{(0)}+1\mathbf{x}^{(1)}+0.5\mathbf{x}^{(2)}=\begin{pmatrix} 1.0 \\ 2.5 \\ 1.5 \\ 1.5 \\ 1.5 \\ 0.0 \end{pmatrix}$$

- P seeding signals $\mathbf{s}^{(t)}$ with a single seeding node $\mathbf{s}^{(t)} = s^{(t)}\mathbf{e}_1$
  - $\Rightarrow$ Define $\mathbf{s}_P := [s^{(P-1)}, \ldots, s^{(0)}]^T$

$$\mathbf{x} = \mathbf{x}^{(P-1)} = \sum_{l=0}^{P-1} \mathbf{S}^l \mathbf{s}^{(P-1-l)} = \sum_{l=0}^{P-1} s^{(P-1-l)} \mathbf{S}^l \mathbf{e}_1$$

  - $\Rightarrow$ Like a filter with input $\mathbf{e}_1$ $\Rightarrow \tilde{\mathbf{x}} = \text{diag}(\mathbf{\Psi}\mathbf{s}_P)\tilde{\mathbf{e}}_1 = \text{diag}(\tilde{\mathbf{e}}_1)\mathbf{\Psi}\mathbf{s}_P$
- U_1 # zeros in $\{[\tilde{\mathbf{e}}_1]_k\}_{k=1}^K$ and D_1 # of repeated values in $\{\lambda_k\}_{k=1}^K$

Single Node - Multiple Time (SN-MT) seeding

- P seeding signals $\mathbf{s}^{(t)}$ with a single seeding node $\mathbf{s}^{(t)} = s^{(t)}\mathbf{e}_1$
  - ⇒ Define $\mathbf{s}_P := [s^{(P-1)}, \ldots, s^{(0)}]^T$

$$\mathbf{x} = \mathbf{x}^{(P-1)} = \sum_{l=0}^{P-1} \mathbf{S}^l \mathbf{s}^{(P-1-l)} = \sum_{l=0}^{P-1} s^{(P-1-l)} \mathbf{S}^l \mathbf{e}_1$$

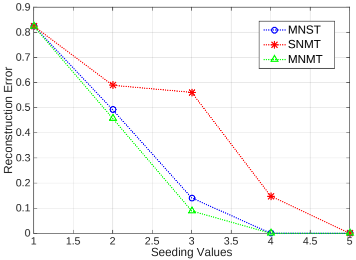
  - ⇒ Like a filter with input $\mathbf{e}_1$ ⇒ $\tilde{\mathbf{x}} = \text{diag}(\mathbf{\Psi}\mathbf{s}_P)\tilde{\mathbf{e}}_1 = \text{diag}(\tilde{\mathbf{e}}_1)\mathbf{\Psi}\mathbf{s}_P$
- U_1 # zeros in $\{[\tilde{\mathbf{e}}_1]_k\}_{k=1}^K$ and D_1 # of repeated values in $\{\lambda_k\}_{k=1}^K$

Perfect reconstruction in SN-MT seeding

Perfect reconstruction of $\mathbf{y}$ is guaranteed via SN-MT seeding if:
i) $\lambda_{k_1} \neq \lambda_{k_2}$ for all $(\lambda_{k_1}, \lambda_{k_2})$ such that $k_1 \leq K$ and $k_2 > K$,
ii) $U_1 = 0$ and $D_1 = 0$.

- Seeding $\mathbf{s}_P^* = \mathbf{\Psi}^{-1}\text{diag}^{-1}(\tilde{e}_1)\text{diag}^{-1}(\tilde{\mathbf{h}}_K^*)\tilde{\mathbf{y}}_K$
  - ⇒ $U_1 = 0$: seeding node acts on every active frequency
  - ⇒ $D_1 = 0$: every active frequency is distinguishable from each other

- Induce opinion in Zachary's karate club [Zachary77]
  - $\Rightarrow$ 34 nodes (members) and 78 undirected edges (friendships)
- Define $\mathbf{S} = \mathbf{I} - \alpha \mathbf{L}$, opinion gets updated influenced by neighbors
- 5-bandlimited target opinion, small discrepancies between neighbors



- Recovery error decreases with the number of seeding values
- Better to convince 3 people at the same time (MN-ST)
  - $\Rightarrow$ than the same person at three different times (SN-MT)