Near-Optimality of Greedy Set Selection in the Sampling of Graph Signals

Luiz F. O. Chamon and Alejandro Ribeiro

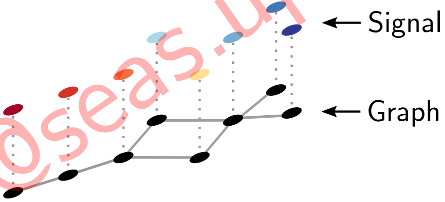
December 7th, 2016

Greedy Sampling of Graph Signals

Greedy Sampling of Graph Signals

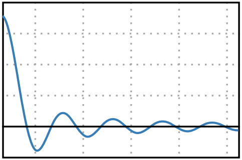
Definition

It's a signal that comes with a graph.



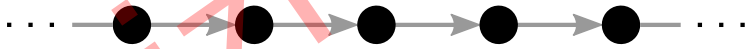
Old dog

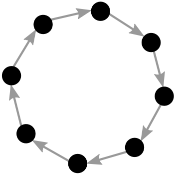
In traditional signal processing, we only look at the values . . .



$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

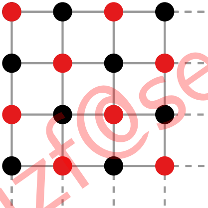
. . . because the structure is implicit

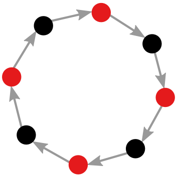




Greedy Sampling of Graph Signals

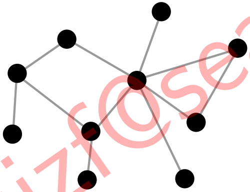
- **Classical signals**: sampling is "easy" on regular domains

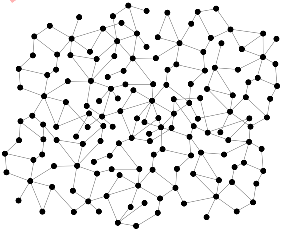




So what? I already know how to do this...

- **Classical signals**: sampling is "easy" on regular domains
- **Graph signals**: not so easy (combinatorial)





Greedy Sampling of Graph Signals

Definition

Pick nodes one at a time by always choosing the one that most improves interpolation at each step.

function $\textsc{GreedySampling}(\ell)$

$\mathcal{G}_0 = \{\}$

for $j = 1, \ldots, \ell$

$u = \operatorname{argmin}_{s \in \mathcal{V} \setminus \mathcal{G}_{j-1}} \mathsf{MSE}\left(\mathcal{G}_{j-1} \cup \{s\}\right)$

$\mathcal{G}_j = \mathcal{G}_{j-1} \cup \{u\}$

end

end

- Pros:
  - Low complexity
  - Sequential
  - Empirically successful
- Cons:
  - Is it guaranteed to be close to optimal?

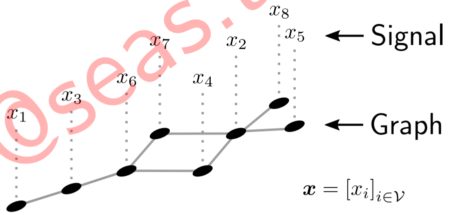
Greedy sampling

- Pros:
  - Low complexity
  - Sequential
  - Empirically successful
- Cons:
  - Is it guaranteed to be close to optimal?

Greedy sampling is guaranteed to do a good job minimizing the interpolation MSE

Graph signal formalism

- A graph signal is a pair $(\mathbb{G}, \boldsymbol{x})$
  - A graph $\mathbb{G} = (\mathcal{V}, \mathcal{E})$
    - $\boldsymbol{A}$ is a matrix representation of $\mathbb{G}$ (e.g., adjacency, Laplacian)
    - **Assumption (Parseval):** $\boldsymbol{A}$ is normal, i.e., $\boldsymbol{A} = \boldsymbol{V}\boldsymbol{\Sigma}\boldsymbol{V}^T$
  - A signal $\boldsymbol{x} \in \mathbb{R}^n$ defined over $\mathcal{V}$



Graph signal formalism

- A graph signal is a pair $(\mathbb{G}, \boldsymbol{x})$
  - A graph $\mathbb{G} = (\mathcal{V}, \mathcal{E})$
    - $\boldsymbol{A}$ is a matrix representation of $\mathbb{G}$ (e.g., adjacency, Laplacian)
    - **Assumption (Parseval):** $\boldsymbol{A}$ is normal, i.e., $\boldsymbol{A} = \boldsymbol{V}\boldsymbol{\Sigma}\boldsymbol{V}^T$
  - A signal $\boldsymbol{x} \in \mathbb{R}^n$ defined over $\mathcal{V}$

- Graph Fourier Transform

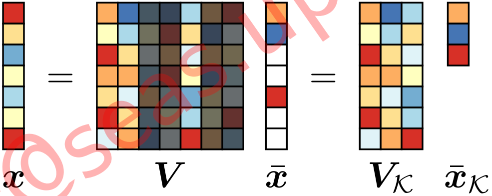
$$\bar{\boldsymbol{x}} = \boldsymbol{V}^T \boldsymbol{x} \quad \longleftrightarrow \quad \boldsymbol{x} = \boldsymbol{V} \bar{\boldsymbol{x}}$$

- A signal is $\mathcal{K}$-bandlimited if $\bar{x}$ is $\mathcal{K}$-sparse: $\bar{\boldsymbol{x}}_{\mathcal{V} \setminus \mathcal{K}} = \boldsymbol{0}$

$$\boldsymbol{x} = \boldsymbol{V}_{\mathcal{K}} \bar{\boldsymbol{x}}_{\mathcal{K}}$$

$\mathcal{K}$-bandlimited graph signal

- $\mathcal{K} = \{1, 2, 5\}$



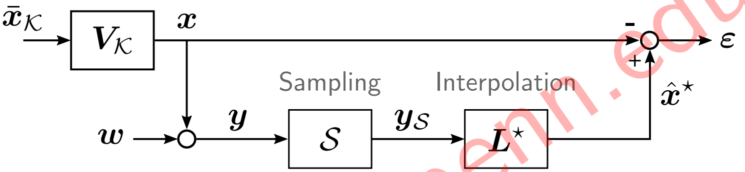
Stochastic graph signal

- **Signal**: $\bar{\boldsymbol{x}}_{\mathcal{K}}$ is a zero-mean RV with covariance $\boldsymbol{\Lambda} = \sigma_x^2 \boldsymbol{I}$

$$\boldsymbol{x} = \boldsymbol{V}_{\mathcal{K}} \bar{\boldsymbol{x}}_{\mathcal{K}}$$

- **Noise**: $\boldsymbol{w}$ is a zero-mean RV with covariance $\boldsymbol{\Lambda}_w = \sigma_w^2 \boldsymbol{I}$

$$\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{w}$$



- Optimal interpolator:

$$L^\star C \left(V_\mathcal{K} \Lambda V_\mathcal{K}^T + \Lambda_w \right) C^T = V_\mathcal{K} \Lambda V_\mathcal{K}^T C^T$$

- Optimal interpolation MSE:

$$\text{MSE}(\mathcal{S}) = \mathbb{E} \left\| x - \hat{x}^\star \right\|^2 = \text{Tr} \left[\left(\sigma_x^{-2} I + \sigma_w^{-2} \sum_{i \in \mathcal{S}} v_i v_i^T \right)^{-1} \right]$$

The sampling set selection problem

$$\begin{aligned} \underset{\mathcal{S} \subseteq \mathcal{V}}{\text{minimize}} \quad & \text{MSE}(\mathcal{S}) \\ \text{subject to} \quad & |\mathcal{S}| = k \end{aligned}$$

- Set function minimization with cardinality constraint

Theorem ([NWF, 1978])

Let $\mathcal{S}^\star$ be the optimal solution of the problem

$$\begin{aligned} \underset{\mathcal{S} \subseteq \mathcal{V}}{\text{minimize}} \quad & f\left(\mathcal{S}\right) \\ \text{subject to} \quad & |\mathcal{S}| = k \end{aligned}$$

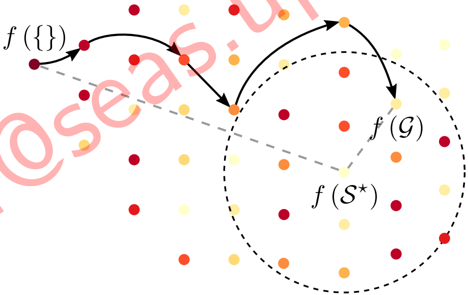
and $\mathcal{G}$ be its greedy solution. If f is (i) monotone decreasing and (ii) supermodular, then

$$\frac{f(\mathcal{G}) - f(\mathcal{S}^\star)}{f(\{\}) - f(\mathcal{S}^\star)} \leq e^{-1} \approx 0.37.$$

Theorem ([NWF, 1978])

If f is (i) monotone decreasing and (ii) supermodular, then

$$\frac{f(\mathcal{G}) - f(\mathcal{S}^\star)}{f(\{\}) - f(\mathcal{S}^\star)} \leq e^{-1} \approx 0.37.$$

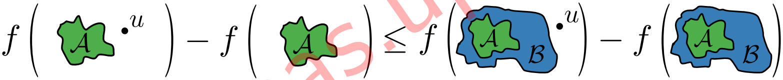


Supermodularity

Definition (Supermodularity)

For $\mathcal{A} \subseteq \mathcal{B}$ and $u \notin \mathcal{B}$,

$$f\left(\mathcal{A} \cup \{u\}\right) - f\left(\mathcal{A}\right) \leq f\left(\mathcal{B} \cup \{u\}\right) - f\left(\mathcal{B}\right)$$



"diminishing returns"

Which f are supermodular?

- $\mathsf{MSE}(\mathcal{S})$ is NOT supermodular
- $\log\det$ of the error covariance matrix is supermodular

Now what?

Definition (Supermodularity)

For $\mathcal{A} \subseteq \mathcal{B}$, $u \notin \mathcal{B}$

$$f\left(\mathcal{A} \cup \{u\}\right) - f\left(\mathcal{A}\right) \leq f\left(\mathcal{B} \cup \{u\}\right) - f\left(\mathcal{B}\right)$$

Now what?

Definition (Approximate supermodularity or α-supermodularity)

For $\mathcal{A} \subseteq \mathcal{B}$ and $u \notin \mathcal{B}$, and $\alpha \in [0, 1]$

$$f(\mathcal{A} \cup \{u\}) - f(\mathcal{A}) \leq \alpha \Big[f(\mathcal{B} \cup \{u\}) - f(\mathcal{B}) \Big]$$

- If $\alpha = 1$, then f is supermodular

Theorem

Let $\mathcal{S}^\star$ be the optimal solution of the problem

$$\begin{aligned} \underset{\mathcal{S} \subseteq \mathcal{V}}{\text{minimize}} \quad & f\left(\mathcal{S}\right) \\ \text{subject to} \quad & \left|\mathcal{S}\right| = k \end{aligned}$$

and $\mathcal{G}_\ell$ be the ℓ-th iteration of a greedy solution. If f is (i) monotone decreasing and (ii) α-supermodular, then

$$\frac{f(\mathcal{G}_\ell) - f(\mathcal{S}^\star)}{f(\{\}) - f(\mathcal{S}^\star)} \leq e^{-\alpha \ell / k}.$$

Greedy α-supermodular minimization

Theorem

If f is (i) monotone decreasing and (ii) α-supermodular, then

$$\frac{f(\mathcal{G}_\ell) - f(\mathcal{S}^\star)}{f(\{\}) - f(\mathcal{S}^\star)} \leq e^{-\alpha\ell/k}.$$

- For $\ell = k$ and $\alpha = 1$, we recover the classical greedy result
- If $\alpha < 1$, then e^{-1} is recovered for $\ell = \alpha^{-1}k$
- Evaluating α is NP-hard

What is α for the MSE?

Theorem

The MSE($\mathcal{S}$) *is* α*-supermodular with*

$$\alpha \geq \frac{1+2\gamma}{(1+\gamma)^4}, \quad \text{for} \quad \gamma = \frac{\sigma_x^2}{\sigma_w^2}.$$

What is α for the MSE?

Theorem

The MSE($\mathcal{S}$) *is α-supermodular with*

$$\alpha \geq \frac{1+2\gamma}{(1+\gamma)^4}, \quad \text{for} \quad \gamma = \frac{\sigma_x^2}{\sigma_w^2}.$$

- $\alpha \to 1$ as $\gamma \to 0$

What is α for the MSE?

Theorem

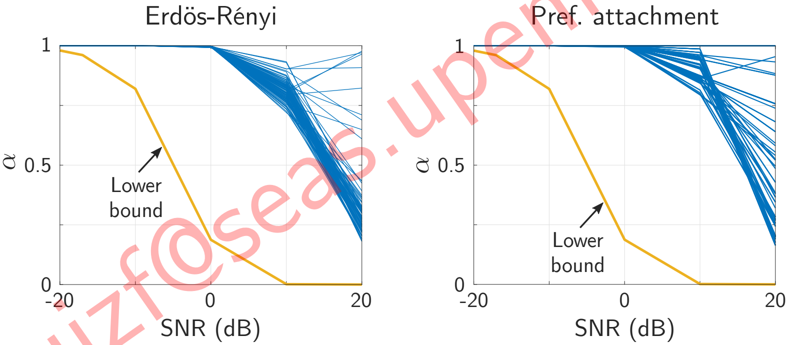
The MSE($\mathcal{S}$) *is α-supermodular with*

$$\alpha \geq \frac{1+2\gamma}{(1+\gamma)^4}, \quad \text{for} \quad \gamma = \frac{\sigma_x^2}{\sigma_w^2}.$$

- $\alpha \to 1$ as $\gamma \to 0$
- $\alpha \to 0$ as $\gamma \to \infty$
  - In the noiseless case, almost every $\mathcal{S}$ with $|\mathcal{S}| \geq |\mathcal{K}|$ yields perfect reconstruction

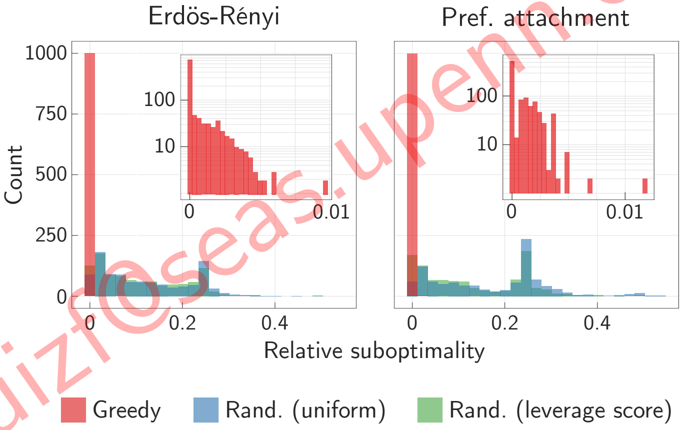
Simulations: α for the MSE

- $n = 10$ nodes, $|\mathcal{K}| = 4$, and 100 realizations

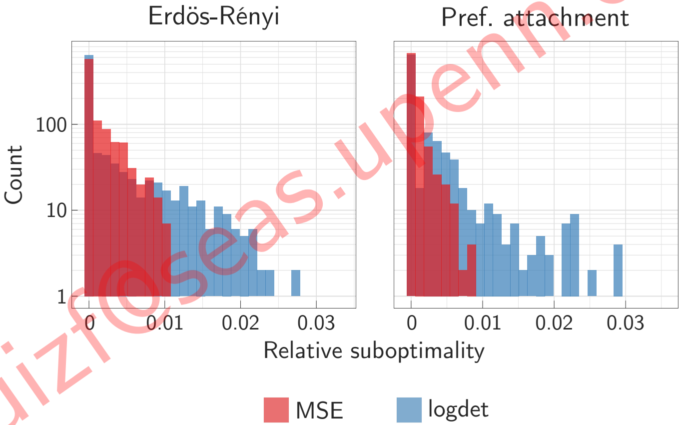


Simulations: greedy sampling

- MSE (10 nodes, $|\mathcal{S}| = |\mathcal{K}| = 4$, and SNR $= 20$ dB)

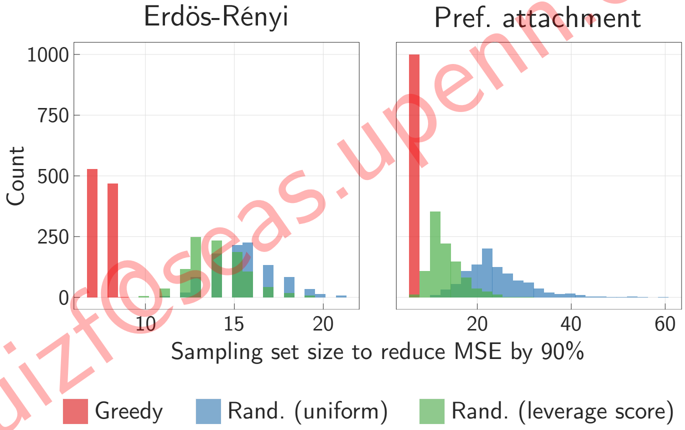


- MSE vs $\log\det$ (10 nodes, $|\mathcal{S}| = |\mathcal{K}| = 4$, and SNR $= 20$ dB)



Simulations: greedy sampling

- 100 nodes, $|\mathcal{K}| = 7$, and SNR $= 20$ dB



Conclusion

- Graph signal sampling is useful, but it's hard
- Interpolation MSE is not supermodular, but almost
- Greedy sampling set selection is efficient and has a guaranteed near-optimal performance

Near-Optimality of Greedy Set Selection in the Sampling of Graph Signals

Luiz F. O. Chamon and Alejandro Ribeiro

More details: `http://www.seas.upenn.edu/~luizf`