Universal Bounds for the Sampling of Graph Signals

Luiz F. O. Chamon and Alejandro Ribeiro

March 7th, 2017
Face recognition

Jane Doe
Face recognition
Face recognition
What is a graph signal?

Definition

- It’s a signal that comes with a graph
- The graph represents *proximity* between the signal samples
Face recognition: which graph?
Face recognition: which graph?
Face recognition: which graph?
Why universal bounds?

- Understand which factors influence the reconstruction performance
- Gauge the interpolation quality a sampling set or heuristic
- Performance benchmark
Universal sampling bounds

Greedy sampling evaluation

Sampled kernel PCA
Graph signal formalism

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

Diagram:

- **Signal**: \(x_i = [x]_i\)
- **Graph**: \(u_i \in V\)
Graph signal formalism

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

- Graph Fourier Transform
  \[
  \bar{\mathbf{x}} = V^T \mathbf{x} \quad \leftrightarrow \quad \mathbf{x} = V \bar{\mathbf{x}}
  \]

- A signal is \(\mathcal{K}\)-bandlimited if \(\bar{\mathbf{x}}\) is \(\mathcal{K}\)-sparse: \(\bar{\mathbf{x}}_{\mathcal{V}\setminus\mathcal{K}} = 0\)
  \[
  \mathbf{x} = V_{\mathcal{K}} \bar{\mathbf{x}}_{\mathcal{K}}
  \]
Signal: $\tilde{x}_K$ is a zero-mean RV with covariance $\Lambda = \sigma_x^2 I$

$$x = V_K \tilde{x}_K$$

Noise: $w$ is a zero-mean RV with covariance $\Lambda_w = \sigma_w^2 I$

$$y = x + w$$

SNR:

$$\gamma = \frac{\sigma_x^2}{\sigma_w^2}$$
Optimal interpolator:

\[ L^* C \left( V_K \Lambda V_K^T + \Lambda_w \right) C^T = V_K \Lambda V_K^T C^T \]

Optimal interpolation MSE:

\[
\text{MSE}(S) = \mathbb{E} \| \mathbf{x} - \hat{\mathbf{x}}^* \|^2 = \sigma_x^2 \text{Tr} \left[ \left( I + \gamma \sum_{i \in S} \mathbf{v}_i \mathbf{v}_i^T \right)^{-1} \right]
\]
Theorem

For any sampling set $S$, it holds that

$$\text{MSE}(S) \geq \frac{|\mathcal{K}|^2 \sigma_x^2}{|\mathcal{K}| + \bar{\ell}|S|}$$

where $\bar{\ell}_m$ is the sum of the $m$ largest structural SNRs:

$$\bar{\ell}_m = \max_{\mathcal{X}:|\mathcal{X}|=m} \sum_{j \in \mathcal{X}} \gamma \| \mathbf{v}_j \|_2^2,$$

with $\mathbf{v}_j^T$ the $j$-th row of $\mathbf{V}_\mathcal{K}$. 
Theorem

For any sampling set $S$, it holds that

$$\text{MSE}(S) \geq \frac{|\mathcal{K}|^2 \sigma_x^2}{|\mathcal{K}| + \ell |S|}$$

- Depends on graph signal statistics and structural properties
- Universal: holds for ANY sampling set
- Increases with the signal bandwidth
- Decreases with $|S|$ with rate dependent on the structural SNR

$$\ell_j = \gamma \|v_j\|_2^2$$
Corollary

Any sampling set $S$ for which $\text{MSE}(S) = \eta$ satisfies

$$\bar{\ell}|_{S|} \geq \frac{|\mathcal{K}|^2 \sigma_x^2 - \eta |\mathcal{K}|}{\eta}.$$ 

Since $\bar{\ell}|_{S|} \leq |S| \ell_{\text{max}}$, it also holds that

$$|S| \geq \frac{|\mathcal{K}|^2 \sigma_x^2 - \eta |\mathcal{K}|}{\eta \ell_{\text{max}}}.$$
$n = 20$ nodes, $|\mathcal{K}| = 5$, and SNR = 20 dB
function $\text{GREEDY_SAMPLING}(k)$

$G_0 = \emptyset$

for $j = 1, \ldots, k$

$u = \arg\min_{s \in \mathcal{V} \setminus G_{j-1}} \text{MSE} \left( G_{j-1} \cup \{s\} \right)$

$G_j = G_{j-1} \cup \{u\}$

end

end
Universal performance bound

\[ n = 1000 \text{ nodes}, \ |\mathcal{K}| = 70, \text{ and } \text{SNR} = 20 \ \text{dB} \]
Kernel PCA is a nonlinear version of PCA

**PCA**

- $u_i \in \mathbb{R}^m$
- $\sum_i u_i u_i^T$
- $C_x = V \Lambda V^T$

**Kernel PCA**

- $\varphi (u_i, u_j)$
- $\kappa (u_i, u_j)$
- $\langle f_i, f_j \rangle_F$
- $\Phi = V \Lambda V^T$
Kernel PCA

Subspace identification

\[ u_i \in \mathbb{R}^m \xrightarrow{\varphi} f_i \in \mathcal{F} \]

\[ \kappa(u_i, u_j) \]

\[ \langle f_i, f_j \rangle_{\mathcal{F}} \]

Projection

\[ \Phi = V \Lambda V^T \]

\( \mathcal{K} = [k] \)

\[ \tilde{y} = V_{\mathcal{K}}^T \tilde{y} \quad [\Theta(kn)] \]

\[ \tilde{y} = [\kappa(u_i, y)]_{i=1,\ldots,n} \quad [n \text{ KEs}] \]

\[ y \in \mathbb{R}^m \]
Sampled kernel PCA

Subspace identification

\[ u_i \in \mathbb{R}^m \xrightarrow{\varphi} f_i \in \mathcal{F} \]

\[ \kappa(u_i, u_j) \quad \langle f_i, f_j \rangle_{\mathcal{F}} \]

Projection

\[ \Phi = V \Lambda V^T \]

\[ \mathcal{K} = [k] \]

\[ \hat{y} = V_{\mathcal{K}}^T L^* \hat{y}_S [\Theta(k|S|)] \]

\[ \hat{y}_S = [\kappa(u_i, y)]_{i \in S} \]

\[ ||S|| \text{ KEs} \]

\[ y \in \mathbb{R}^m \]
Kernel PCA (polynomial kernel, $|\mathcal{K}| = 25$)

- Direct kPCA: 280 KEs and $\mathcal{O}(7000)$ operations
- Sampled kPCA: 40 KEs and $\mathcal{O}(1000)$ operations
Conclusion

- Sampling graph signals is not straightforward due to the irregularity of the domain.
- There are universal performance bounds that can be used to benchmark sampling sets.
- Sampling is an effective complexity reduction technique (e.g., sampled kPCA).
Universal Bounds for the Sampling of Graph Signals

Luiz F. O. Chamon and Alejandro Ribeiro

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