A Novel Scheme for Support Identification and Iterative Sampling of Bandlimited Graph Signals

Abolfazl Hashemi [1], Rasoul Shafipour [2], Haris Vikalo [1], and Gonzalo Mateos [2]

[1]Dept. of Electrical and Computer Engineering, University of Texas at Austin, Austin, TX, USA
[2]Dept. of Electrical and Computer Engineering, University of Rochester, Rochester, NY, USA

TEXAS The University of Texas at Austin

UNIVERSITY of ROCHESTER

BACKGROUND

- Graph Signal Processing
  - Modeling network processes by exploiting the underlying graph structures
  - Applications: sensor and social networks, transportation systems, gene regulatory networks
- Sampling and Reconstruction
  - Selecting a *small representative* subset of graph nodes
  - Applications: constrained sensing in sensor networks, data summarization
- Notation and model
  - $\mathbf{x} \in \mathbb{R}^N$: signal on graph $\mathcal{G}$ with N nodes, $\mathbf{C} \in \{0,1\}^{k \times N}$ Sampling matrix
  - $\mathbf{A}$: adjacency matrix of $\mathcal{G}$
  - $\mathbf{V}$: basis of the graph signal (here, eigenvectors of the Laplacian matrix $\mathbf{L}$)
  - $\bar{\mathbf{x}} = \mathbf{V}^\top \mathbf{x}$: graph Fourier transform, k-sparse (*bandlimited*), support $\mathcal{K}$
  - $\mathbf{U} \in \mathbb{R}^{N \times k}$: submatrix of $\mathbf{V}$ containing columns indexed by $\mathcal{K}$
  - $\mathbf{y} = \mathbf{x} + \mathbf{n}$: measurement model, $\mathbf{n}$ *bounded* noise with $\mathbb{E}[\mathbf{n}\mathbf{n}^\top] = \mathbf{Q}$
- Goal: finding a *good* sampled signal $\tilde{\mathbf{x}} = \mathbf{C}\mathbf{x}$ for perfect reconstruction:

$$\hat{\mathbf{x}} = \mathbf{U}\bar{\mathbf{x}}_{\mathcal{K}} = \mathbf{U}(\mathbf{C}\mathbf{U})^{-1}\tilde{\mathbf{x}}$$

- Prior work [1] and [2] based on using uniform and leverage score random sampling : nonzero probability of failure, require more than k samples
- Our approach: Iterative scheme based on orthogonal matching pursuit (OMP)
- Support identification from *historical observations* of the graph signal

ITERATIVE SELECTION SAMPLING

- Necessary and sufficient condition for perfect recovery: *invertibility* of $\mathbf{CU}$
- Make it invertible by construction:
  - select a *residual node* $\ell \in [N]$ (to be excluded from the sampling set)
  - suggestion for residual: $\ell = \mathrm{argmin}_{j \in [N]} \|\mathbf{u}_j\|_2$
  - iteratively identify a sampling node and construct sampling set $\mathcal{S}$

$$j_s = \arg\max_{j \in \mathcal{N} \setminus \mathcal{S}} \frac{|\mathbf{r}_{i-1}^\top \mathbf{u}_j|^2}{\|\mathbf{u}_j\|_2^2}$$

  - $\mathbf{r}_i = \mathbf{P}_{\mathcal{S}}^\perp \mathbf{u}_\ell$: the residual vector initialized at $\mathbf{r}_0 = \mathbf{u}_\ell$
  - $\mathbf{P}_{\mathcal{S}}^\perp = \mathbf{I}_N - \mathbf{U}_{\mathcal{S},r}^\top (\mathbf{U}_{\mathcal{S},r}^\top)^\dagger$: the projection operator to complement of the subspace spanned by $\mathbf{U}_{\mathcal{S},r}$ (rows of $\mathbf{U}$ indexed by $\mathcal{S}$)

PERFORMANCE ANALYSIS

- Proposed scheme guarantees perfect recovery in noiseless case for all *connected* graphs with *general structures* and with *normal adjacency*.

Theorem 1:

Let $\mathcal{S}$ be the sampling set constructed by Algorithm 1 and let be $\mathbf{C}$ the corresponding sampling matrix such that $|\mathcal{S}| = k$. Then, the matrix $\mathbf{CU}$ is invertible.

- Proof's remarks:
  - An inductive argument
  - Iterative selection of linearly independent $\mathbf{u}_i$'s
  - Zero residual norm only after the last iteration

Sampling under bounded noise

- Assumption: $\|\mathbf{n}\|_2 \le \epsilon_n$
- Explicit bound on reconstruction error

$$\|\hat{\mathbf{x}} - \mathbf{x}\|_2 \le \sigma_{\max}((\mathbf{U}_{\mathcal{S},r}^\top \mathbf{Q}_{\mathcal{S}}^{-1} \mathbf{U}_{\mathcal{S},r})^{-1} \mathbf{U}_{\mathcal{S},r}^\top \mathbf{Q}_{\mathcal{S}}^{-1}) \epsilon_{\mathbf{n}}$$

- Guaranteed existence of inverse matrices under Algorithm 1
- Preserving *statistical characteristics (e.g. whiteness)* of effective noise

SUPPORT IDENTIFICATION

- Support recovery given P historical templates $\mathbf{X} = [\mathbf{x}^1, \cdots, \mathbf{x}^P] \in \mathbb{R}^{N \times P}$ with shared support from noisy observations $\mathbf{Y} = \mathbf{X} + \mathbf{N}$
- Equivalent task: estimating sparse GFTs $\bar{\mathbf{X}} = [\bar{\mathbf{x}}^1, \cdots, \bar{\mathbf{x}}^P] = \mathbf{V}^T\mathbf{X}$
- Proposed optimization based on *block sparsity* of $\bar{\mathbf{X}}$:

$$\min_{\bar{\mathbf{X}}} \quad \|\bar{\mathbf{X}} - \mathbf{V}^T\mathbf{Y}\|_F^2 \quad \text{s.t.} \quad \|\bar{\mathbf{X}}\|_{2,0} \le k,$$

- Closed-form solution via row-wise l_2 norm thresholding on $\bar{\mathbf{Y}} = \mathbf{V}^T\mathbf{Y}$

$$\bar{\mathbf{X}}(i,:)^\star = \begin{cases} \bar{\mathbf{Y}}(i,:) & \|\bar{\mathbf{Y}}(i,:)\|_2 \ge k^{\text{th}} \text{ largest } \ell_2 \text{ norm} \\ \mathbf{0} & \text{o.w.} \end{cases}$$

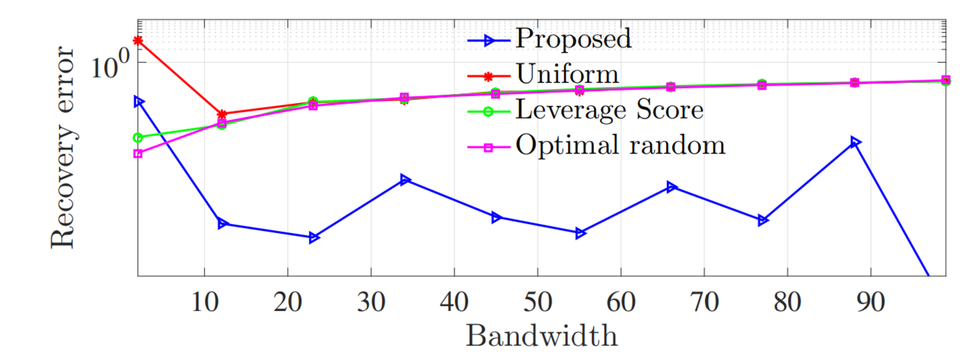
Theorem 2
Assume $\mathbf{V}$ is *orthogonal*. Under bounded noise assumption, the GFTs $\bar{\mathbf{X}}$ and support $\mathcal{K}$ are identifiable if

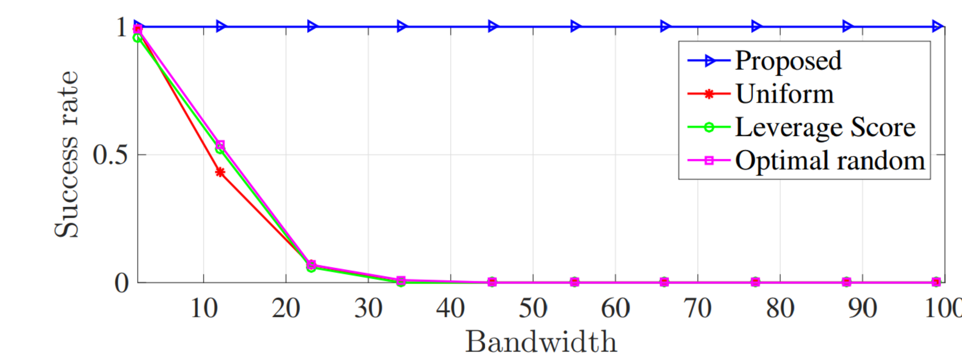
$$\min_{i \in \mathcal{K}} \quad \|\bar{\mathbf{X}}(i,:)\|_2 > 2\epsilon_{\mathbf{n}}\sqrt{P}.$$

- Perfect support identification *in the absence of noise* with *P = 1*
- Easier satisfiability of the established sufficient condition for larger *P*

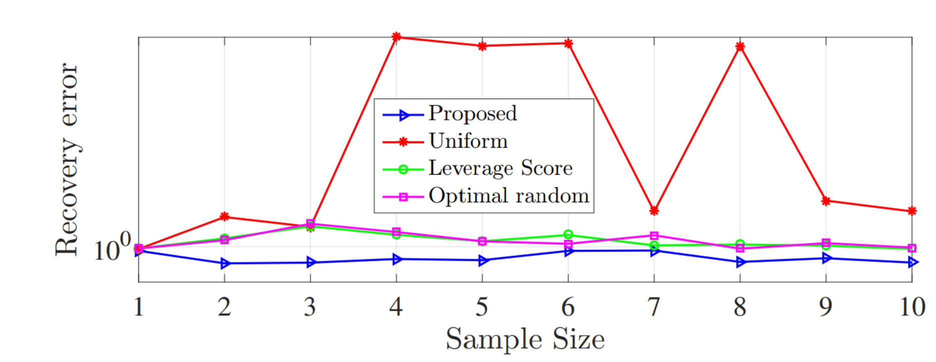
RESULTS

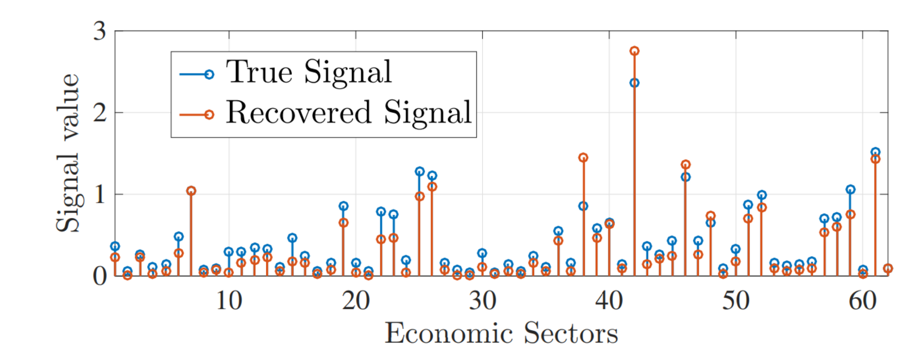
- Simulated Erdos-Renyi graphs with *N = 1000* and *p = 0.1*
- Noisy samples with $\sigma = 0.02$
- Recovery error (error to signal energy)
- Success rate (invertibility of $\mathbf{CU}$)



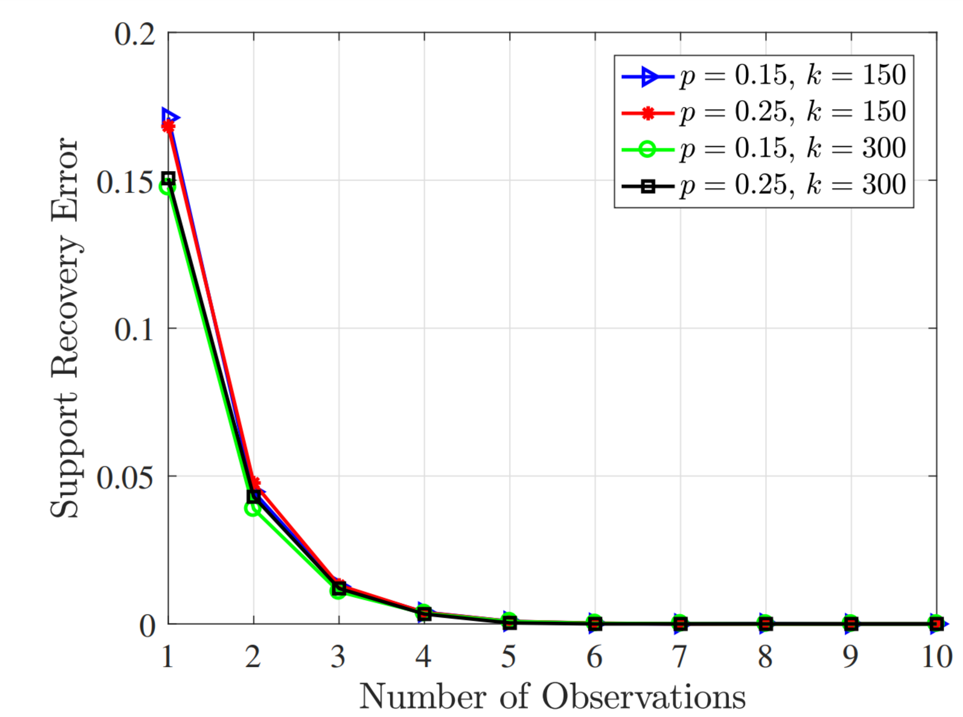


- Real graph of economy sectors with *N = 64*
- Disaggregated GDP is an *approximately* bandlimited graph signal
- Recovery error vs. various sampling set sizes
- Achieved recovery error: *1.32%*





- Simulated Erdos-Renyi graphs with *N = 1000*
- AWGN with *20dB* power
- Various degrees of connectivity *(p)*
- Support recovery error vs. *P*, number of signal templates, (i.e. observations)
- Monotonically better accuracy



CONCLUSION

- Our contributions:
  - Iterative sampling of graph signals in non-Bayesian setting with performance guarantees
  - Extension to unknown support scenario: support identification from historical templates of the graph signal
  - See our extended article "Accelerated sampling of bandlimited graph signals" in arXiv: https://arxiv.org/abs/1807.07222
- Future work: Joint support identification and sampling

[1] S. Chen et al., "Signal recovery on graphs: Fundamental limits of sampling strategies," Dec 2016.
[2] G. Puy et al., "Random sampling of bandlimited signals on graphs," Mar. 2018.