Information Theoretic Limits for Standard and One-bit Compressed Sensing with Graph-structured Sparsity

ICASSP 2022

Adarsh Barik, Dr. Jean Honorio

Purdue University

Sparsity

Nowadays we deal with high dimensional data. Fortunately, it can often be represented as sparse vectors.

Can there be structure within sparsity?

- Sparsity appears as clusters/blocks
- Non-zero entries form a rooted sub-tree

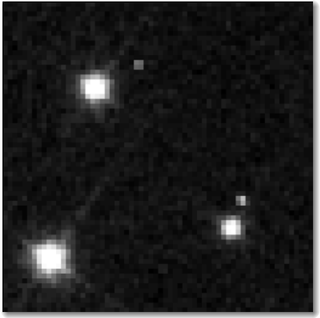


Figure: Hubble image (cropped)[Indyk, 2015]

Weighted Graph Model

Weighted Graph Model [Hegde et al., 2015]

- Graph G, total nodes d, sparsity s, no. of connected components g, weight budget B
- Set of subgraphs with s nodes that are clustered in g connected components with total edge weights $\leqslant B$

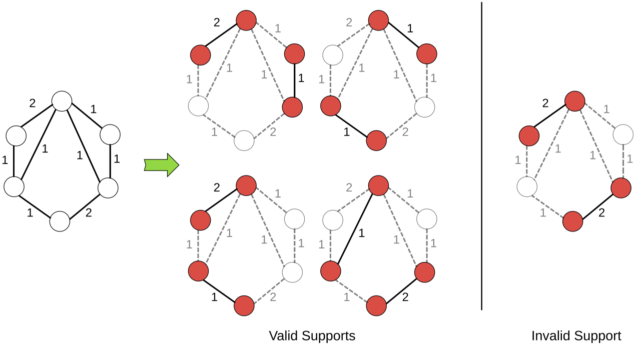


Figure: $d = 6, s = 4, g = 2, B = 3$

Weighted Graph Model

- Weight degree $\rho(v)$: Largest number of adjacent nodes of v connected by edges with the same weight
- Weight-degree of G, ρ is the maximum weight-degree across v

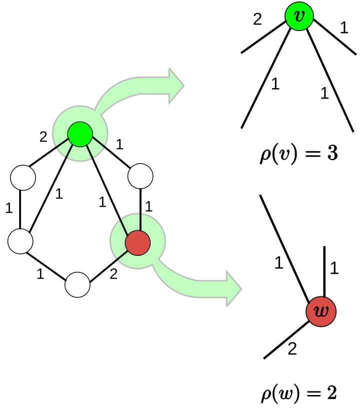


Figure: Weight degree of graph is $\rho = 3$

Weighted Graph Model: Examples

Many sparsity models can be described using Weighted Graph Model.

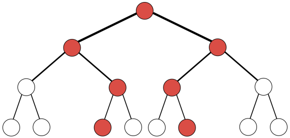


Figure: Tree sparsity with $d = 15, s = 7, g = 1, B = 6, \rho = 3$

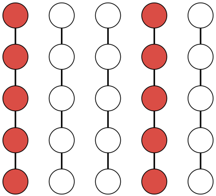
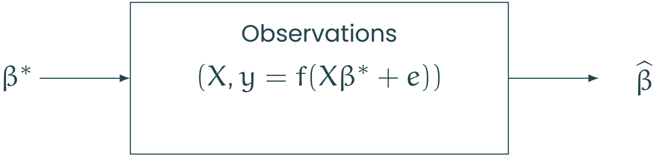


Figure: Block sparsity with $d = 25, s = 10, g = 2, B = 8, \rho = 2$

Problem Statement

Compressed Sensing

- Generative model: $y = f(X\beta^* + e)$ where $y \in \mathbb{R}^n, X \in \mathbb{R}^{n \times d}, \beta^* \in \mathbb{R}^d$ and noise $e \in \mathbb{R}^n$
- Infer a high dimensional sparse signal β^* from low dimensional noisy observations (X, y)
- How many observations are necessary?



Our Results: Standard Compressed Sensing

$\exists$ a (G, s, g, B)—WGM for Standard Compressed Sensing: $y = X\beta^* + e$ such that

Noiseless Case

- if

$$n \in \tilde{o}((s-g)(\log \rho(G) + \log \frac{B}{s-g}) + g \log \frac{d}{g} + (s-g) \log \frac{g}{s-g} + s \log 2),$$

- then $\inf_{\widehat{\beta}} \sup_D \mathbb{P}_{\substack{\beta^* \sim \mathrm{Unif}(\mathcal{F}) \\ (X,y) \sim D(\beta^*)^n}} (\beta^* \neq \widehat{\beta}) \geq \frac{1}{2}$

Noisy Case

- if

$$n \in \tilde{o}((s-g)(\log \rho(G) + \log \frac{B}{s-g}) + g \log \frac{d}{g} + (s-g) \log \frac{g}{s-g} + s \log 2),$$

- then $\inf_{\widehat{\beta}} \sup_D \mathbb{P}_{\substack{\beta^* \sim \mathrm{Unif}(\mathcal{F}) \\ (X,y) \sim D(\beta^*)^n}} (\|\beta^* - \widehat{\beta}\| \geq C\|e\|) \geq \frac{1}{10}$

Our Results: One-bit Compressed Sensing

$\exists$ a (G, s, g, B)—WGM for One-bit Compressed Sensing: $y = \text{sign}(X\beta^* + e)$ such that

Exact Recovery

- if

$$n \in o((s-g)(\log \rho(G) + \log \frac{B}{s-g}) + g \log \frac{d}{g} + (s-g)\log \frac{g}{s-g} + s \log 2),$$

- then $\inf_{\widehat{\beta}} \sup_{D} \mathbb{P}_{\substack{\beta^* \sim \text{Unif}(\mathcal{F}) \\ (X,y) \sim D(\beta^*)^n}} (\beta^* \neq \widehat{\beta}) \geqslant \frac{1}{2}$

Approximate Recovery

- if

$$n \in o((s-g)(\log \rho(G) + \log \frac{B}{s-g}) + g \log \frac{d}{g} + (s-g)\log \frac{g}{s-g} + s \log 2),$$

- then $\inf_{\widehat{\beta}} \sup_{D} \mathbb{P}_{\substack{\beta^* \sim \text{Unif}(\mathcal{F}) \\ (X,y) \sim D(\beta^*)^n}} (\| \frac{\beta^*}{\|\beta^*\|} - \frac{\widehat{\beta}}{\|\widehat{\beta}\|} \| \geqslant \epsilon) \geqslant \frac{1}{2}$

Proof Outline

- Construction of a restricted ensemble
  - Constructing Weighted Graph Model
  - Choosing coefficients of β
- Establishing bounds
  - Lower bound on number of possible signals from restricted ensemble
  - Upper bound mutual information between signal and observation
- Using Fano's inequality

Optimality of the Results

	Standard Compressed Sensing	
Sparsity Structure	Our Lower Bound	Upper Bound
Weighted Graph Model	$\widetilde{\Omega}(s(\log \rho(G)\frac{B}{s}) + g \log \frac{d}{g})$	$\mathcal{O}(s(\log \rho(G)\frac{B}{s}) + g \log \frac{d}{g})$
Tree Structured	$\widetilde{\Omega}(s)$	$\mathcal{O}(s)$
Block Structured	$\widetilde{\Omega}(s + g \log \frac{d}{s})$	$\mathcal{O}(s + g \log \frac{d}{s})$
Regular s-sparsity	$\widetilde{\Omega}(s \log \frac{d}{s})$	$\mathcal{O}(s \log \frac{d}{s})$

	One-bit Compressed Sensing	
Sparsity Structure	Our Lower Bound	Upper Bound
Weighted Graph Model	$\Omega(s(\log \rho(G)\frac{B}{s}) + g \log \frac{d}{g})$	NA
Tree Structured	$\Omega(s)$	NA
Block Structured	$\Omega(s + g \log \frac{d}{s})$	NA
Regular s-sparsity	$\Omega(s \log \frac{d}{s})$	$\mathcal{O}(s \log \frac{d}{s})$

Thank You!