Weighted Superimposed Codes and Constrained Compressed Sensing

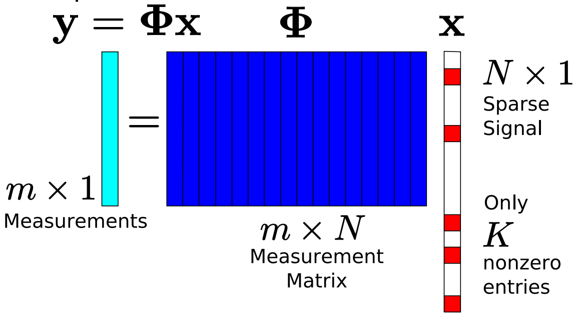
Wei Dai (ECE UIUC)
Joint work with Olgica Milenkovic (ECE UIUC)

University of Illinois at Urbana-Champaign

DIMACS 2009

Compressed Sensing

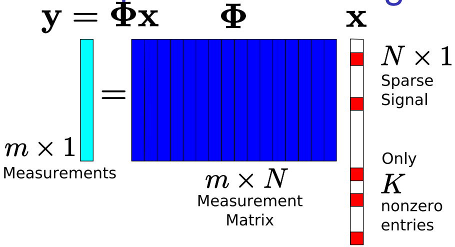
- Classic setup



Kashin, 1977; Bresler et al., 1999; Donoho et al., 2004; Candés et al., 2005; $\cdots$

- Only one constraint
  - $\mathbf{x} \in \mathbb{R}^N$ is K-sparse

Constrained Compressed Sensing



- Constraints on $\mathbf{x}$
  - x_i's are correlated (Dai & Milenkovic; Baraniuk, et al.; $\cdots$).
  - x_i are bounded integers.
  - May improve performance.
- Constraints on $\mathbf{\Phi}$
  - Sparse/structured (Dai & Milenkovic; Indyk, et al.; Do, et al.; Strauss, et al.).
  - l_p-norm + nonnegativity.
  - May introduce performance loss.
- Performance requirement on noise tolerance.

Application 1: CS DNA Microarrays

DNA Microarray: measures the concentration of certain molecules (such as mRNA) for tens of thousands of genes simultaneously.
Major issue: each sequence has a unique identifier $\Rightarrow$ high cost.
CS DNA Microarray (Dai, Sheikh, Milenkovic and Baraniuk; Hassibi)

Constraints:

- $\mathbf{x}$: x_i =the # of certain molecules.
  $|x_i| \leq t$: Bounded integer.
- $\mathbf{\Phi}$: $\Phi_{i,j}$ =the affinity (the probability) between the probe and target.
  $\|\mathbf{\Phi}_i\|_{l_1} = 1$, $\Phi_{i,j} \geq 0$.

The same model works for low light imaging, drug screening$\cdots$

Application 2: Multiuser Communications

A multi-access channel with K users

$$\mathbf{y} = \sum_{i=1}^{K} h_i \sqrt{P_i} \mathbf{t}_i + \mathbf{e}.$$

$\mathbf{t}_i \in \mathcal{C}_i$

$\mathcal{C}_i$: i^{th} user's codebook $\quad |\mathcal{C}_i| = n_i$

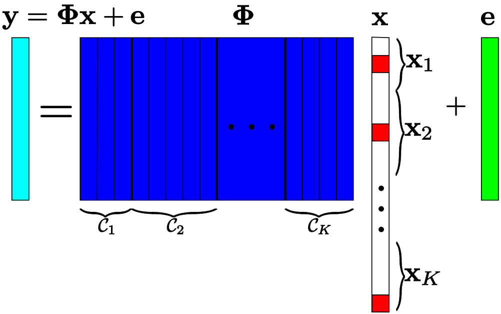
Application 2: Multiuser Communications

A multi-access channel with K users

$$\mathbf{y} = \sum_{i=1}^{K} h_i \sqrt{P_i}\mathbf{t}_i + \mathbf{e}.$$

$\mathbf{t}_i \in \mathcal{C}_i$

$\mathcal{C}_i$: i^{th} user's codebook $\qquad |\mathcal{C}_i| = n_i$



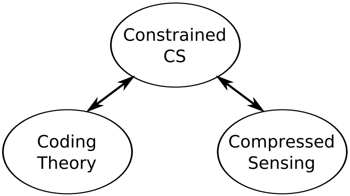
Questions regarding to Constrained CS (CCS)

- How to analyze the gain/loss for a given set of constraints?
- How do the constraints affect the reconstruction algorithms?

Questions regarding to Constrained CS (CCS)

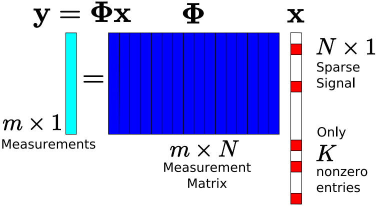
- How to analyze the gain/loss for a given set of constraints?
- How do the constraints affect the reconstruction algorithms?

Our Observation: coding theoretic techniques help.



Superimposed Codes

- Euclidean Superimposed Codes (Ericson and Györfi, 1988)
  - $x_i = 0/1$.
  - $\|\mathbf{v}_i\|_2 = 1$.
  - Distance requirement
    $\Rightarrow$ deterministic noise tolerance.
    $\|\boldsymbol{\Phi}(\mathbf{x}_1 - \mathbf{x}_2)\|_2 \geq d \quad \forall \mathbf{x}_1 \neq \mathbf{x}_2$

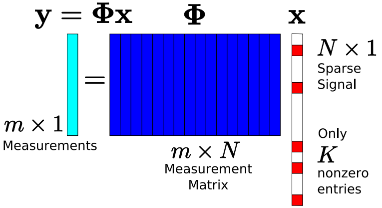


- Applications $\Rightarrow$ Weighted superimposed codes (WSC) (D. and Milenkovic, 2008)
  - $|x_i| \leq t$ is an integer.
  - $\|\mathbf{v}_i\|_p = 1$.
  - Distance requirement
    $\|\boldsymbol{\Phi}(\mathbf{x}_1 - \mathbf{x}_2)\|_p \geq d \quad \forall \mathbf{x}_1 \neq \mathbf{x}_2$.

- A hybrid of CS and Euclidean superimposed codes

Superimposed Codes

- Euclidean Superimposed Codes (Ericson and Györfi, 1988)
  - $x_i = 0/1$.
  - $\|\mathbf{v}_i\|_2 = 1$.
  - Distance requirement
    $\Rightarrow$ deterministic noise tolerance.
    $\|\mathbf{\Phi}(\mathbf{x}_1 - \mathbf{x}_2)\|_2 \geq d \quad \forall \mathbf{x}_1 \neq \mathbf{x}_2$



- Applications $\Rightarrow$ Weighted superimposed codes (WSC) (D. and Milenkovic, 2008)
  - $|x_i| \leq t$ is an integer.
  - $\|\mathbf{v}_i\|_p = 1$.
  - Distance requirement
    $\|\mathbf{\Phi}(\mathbf{x}_1 - \mathbf{x}_2)\|_p \geq d \quad \forall \mathbf{x}_1 \neq \mathbf{x}_2$.

- A hybrid of CS and Euclidean superimposed codes

Rate Bounds for WSCs

Definition: Let

$$N(m, K, d, t) = \max\{N : \exists \mathcal{C}\}.$$

The asymptotic code rate is defined as

$$R(K, d, t) = \lim_{m\to\infty} \sup \frac{\log N(m,K,d,t)}{m}.$$

Theorem:

- For Euclidean norm,
$$\frac{\log K}{4K}(1 + o(1)) \le R(K, d, t) \le \frac{\log K}{2K}(1 + o_{t,d}(1)).$$
- For l_1-WSC and nonnegative l_1-WSC
$$\frac{\log K}{4K}(1 + o(1)) \le R(K, d, t) \le \frac{\log K}{K}(1 + o_{t,d}(1)).$$

Rate Bounds for WSCs

Definition: Let

$$N(m, K, d, t) = \max\{N : \exists \mathcal{C}\}.$$

The asymptotic code rate is defined as

$$R(K, d, t) = \lim_{m\to\infty} \sup \frac{\log N(m,K,d,t)}{m}.$$

Theorem:

- For Euclidean norm,
$$\frac{\log K}{4K}(1 + o(1)) \leq R(K, d, t) \leq \frac{\log K}{2K}(1 + o_{t,d}(1)).$$
- For l_1-WSC and nonnegative l_1-WSC
$$\frac{\log K}{4K}(1 + o(1)) \leq R(K, d, t) \leq \frac{\log K}{K}(1 + o_{t,d}(1)).$$

Interpretation

- For WSCs,

$$\frac{K \log N}{\log K} \leq m \leq \frac{4K \log N}{\log K}.$$

The bounds are not independent of d
$\Rightarrow$ can make the distance arbitrarily close to one.

- For classic CS,

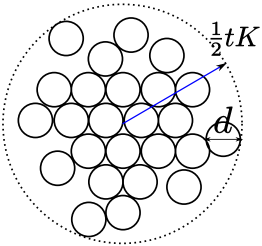
$$m \geq O\left(K \log\left(\frac{N}{K}\right)\right).$$

No performance garantee under noise.

The Proof of the Upper Bound

Low-hanging fruit: sphere-packing bound:
Minimum distance $d \Rightarrow$ Balls $B\left(\mathbf{\Phi x}, \frac{d}{2}\right)$ are disjoint

$$\sum_{k=1}^{K} \binom{N}{k} (2t)^k \leq \left(\frac{tK + \frac{d}{2}}{\frac{d}{2}} \right)^m \quad \Rightarrow \quad \frac{\log N}{m} \leq \frac{\log K}{K}.$$



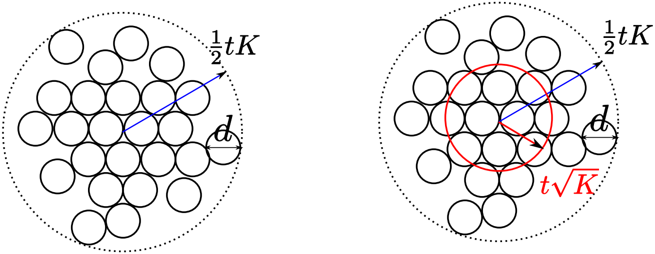
High-hanging fruit: a large fraction of balls lie in the sphere of a smaller radius.

$$\frac{\log N}{m} \leq \frac{\log \sqrt{K}}{K} = \frac{\log K}{2K}.$$

The Proof of the Upper Bound

Low-hanging fruit: sphere-packing bound:
Minimum distance $d \Rightarrow$ Balls $B\left(\mathbf{\Phi x}, \frac{d}{2}\right)$ are disjoint

$$\sum_{k=1}^{K} \binom{N}{k} (2t)^k \leq \left(\frac{tK + \frac{d}{2}}{\frac{d}{2}}\right)^m \quad \Rightarrow \quad \frac{\log N}{m} \leq \frac{\log K}{K}.$$



High-hanging fruit: a large fraction of balls lie in the sphere of a smaller radius.

$$\frac{\log N}{m} \leq \frac{\log \sqrt{K}}{K} = \frac{\log K}{2K}.$$

Proof of the Lower Bounds: Random Coding

Random codes:

$\mathbf{H} \in \mathbb{R}^{m \times N}$ =a Gaussian random matrix ($H_{i,j} \sim N\left(0, \frac{1}{m}\right)$).

$\boldsymbol{\Phi}$: $\quad \boldsymbol{v}_i = \boldsymbol{h}_i / \left\| \boldsymbol{h}_i \right\|_p$.

$d \leq \left\| \Delta \boldsymbol{y} \right\|_p = \left\| \boldsymbol{\Phi} \cdot (\boldsymbol{x}_1 - \boldsymbol{x}_2) \right\|_p$.

$(\Delta \boldsymbol{y})_i \approx$Linear combination of Gaussian rvs.

l_p-norm of a Gaussian vector: large deviations.

$$R(K, d, t) = \lim_{(m,N)\to\infty} \sup \frac{\log N}{m} \geq \frac{\log K}{4K} (1 + o(1)).$$

Difficulty with nonnegativity.

- Gaussian approximation.
- The Berry-Esseen theorem for bounding the approx. error.

Proof of the Lower Bounds: Random Coding

Random codes:

$\mathbf{H} \in \mathbb{R}^{m \times N}$ =a Gaussian random matrix ($H_{i,j} \sim N\left(0, \frac{1}{m}\right)$).

$\mathbf{\Phi}$: $\boldsymbol{v}_i = \boldsymbol{h}_i / \left\|\boldsymbol{h}_i\right\|_p$.

$$d \leq \left\|\Delta \boldsymbol{y}\right\|_p = \left\|\mathbf{\Phi} \cdot (\boldsymbol{x}_1 - \boldsymbol{x}_2)\right\|_p.$$

$(\Delta \boldsymbol{y})_i \approx$Linear combination of Gaussian rvs.

l_p-norm of a Gaussian vector: large deviations.

$$R(K, d, t) = \lim_{(m,N)\to\infty} \sup \frac{\log N}{m} \geq \frac{\log K}{4K} (1 + o(1)).$$

Difficulty with nonnegativity.

- Gaussian approximation.
- The Berry-Esseen theorem for bounding the approx. error.

Proof of the Lower Bounds: Random Coding

Random codes:
$\mathbf{H} \in \mathbb{R}^{m \times N}$ =a Gaussian random matrix ($H_{i,j} \sim N\left(0, \frac{1}{m}\right)$).
$\mathbf{\Phi}$: $\boldsymbol{v}_i = \boldsymbol{h}_i / \left\|\boldsymbol{h}_i\right\|_p$.

$d \leq \left\|\Delta \boldsymbol{y}\right\|_p = \left\|\mathbf{\Phi} \cdot (\boldsymbol{x}_1 - \boldsymbol{x}_2)\right\|_p$.
$(\Delta \boldsymbol{y})_i$ ≈Linear combination of Gaussian rvs.
l_p-norm of a Gaussian vector: large deviations.

$$R(K, d, t) = \limsup_{(m,N)\to\infty} \frac{\log N}{m} \geq \frac{\log K}{4K}(1 + o(1)).$$

Difficulty with nonnegativity.

- Gaussian approximation.
- The Berry-Esseen theorem for bounding the approx. error.

Proof of the Lower Bounds: Random Coding

Random codes:
$\mathbf{H} \in \mathbb{R}^{m \times N}$ =a Gaussian random matrix ($H_{i,j} \sim N\left(0, \frac{1}{m}\right)$).
$\mathbf{\Phi}$: $\quad \boldsymbol{v}_i = \boldsymbol{h}_i / \left\| \boldsymbol{h}_i \right\|_p$.

$$d \leq \left\| \Delta \boldsymbol{y} \right\|_p = \left\| \mathbf{\Phi} \cdot (\boldsymbol{x}_1 - \boldsymbol{x}_2) \right\|_p .$$

$(\Delta \boldsymbol{y})_i$ ≈Linear combination of Gaussian rvs.
l_p-norm of a Gaussian vector: large deviations.

$$R(K, d, t) = \limsup_{(m,N)\to\infty} \frac{\log N}{m} \geq \frac{\log K}{4K} \left(1 + o(1)\right).$$

Difficulty with nonnegativity.

- Gaussian approximation.
- The Berry-Esseen theorem for bounding the approx. error.

Code Construction and Decoding Algorithms

- Coding theory:
  - Offers myriad of construction techniques.
  - No efficient decoding methods for WSC codes were known before.
- CS:
  - Offers decoding algorithmic solutions
    l_1-minimization, OMP, SP, CoSaMP ...
- Combination?

Decoding

The WESC decoder:

$$\hat{x}_i = \text{round}\left(\mathbf{v}_i^* \mathbf{y}\right).$$

no iteration.

OMP: K iterations.

Discrete input $\Rightarrow$ complexity reduction

The WESC decoder: $O(mN)$

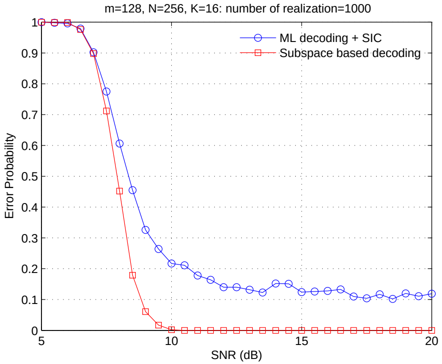
OMP: $O(KmN)$

Code Rate for both WESC decoder and OMP:

$$R \leq \frac{1}{8K^2t^2} \quad \Rightarrow \quad m = O\left(K^2 \log N\right).$$

Multiuser Interference Cancellation and Decoding

- High mobility $\Rightarrow$ No channel information at transmitters.
- Coding and decoding motivated by CS.



Conclusion

WSCs for constrained CS:

- Quantified the code rate
- Noise tolerance
- Efficient decoding