CoSaMP with Redundant Dictionaries

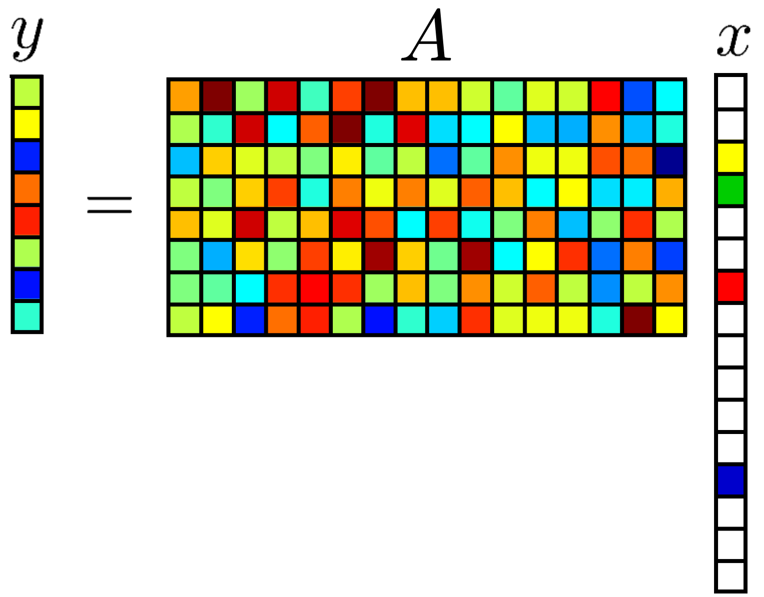
Mark Davenport *Deanna Needell* *Michael Wakin*

Georgia Institute of Technology

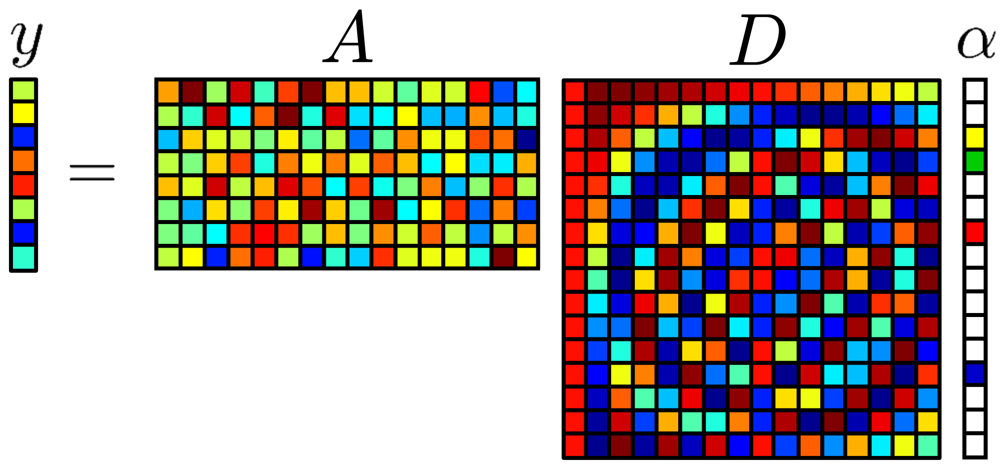
Claremont McKenna College

MINES

Compressive Sensing



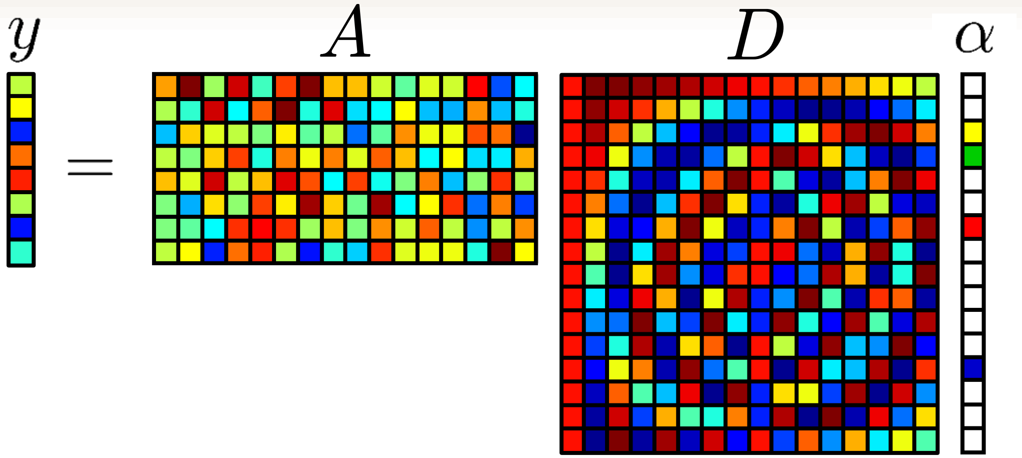
Compressive Sensing



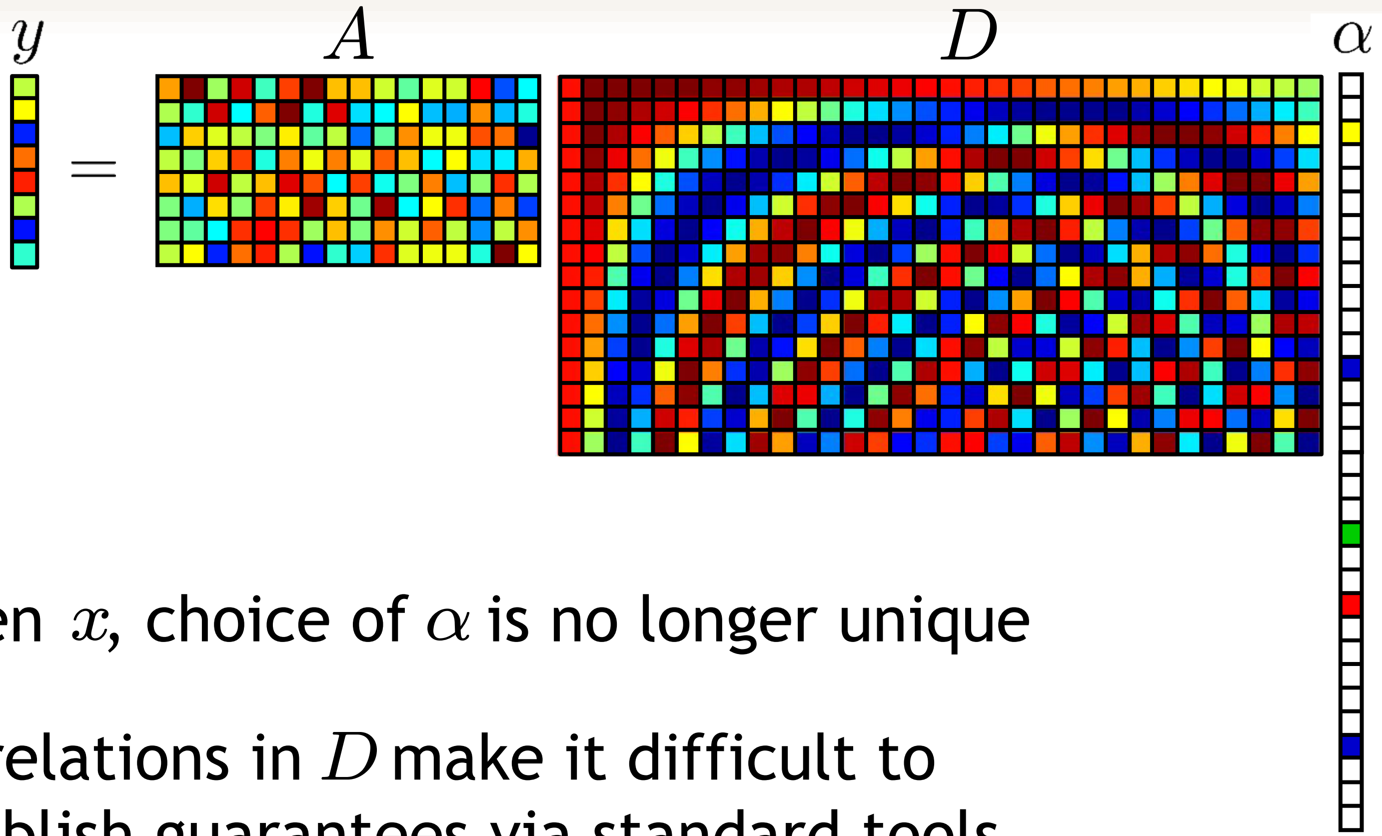
y → $\widehat{\alpha}$ → $\widehat{x} = D\widehat{\alpha}$

The Treachery of Images

The Treachery of α



The Treachery of α



- Given x, choice of α is no longer unique
- Correlations in D make it difficult to establish guarantees via standard tools
- If D is poorly conditioned, we can have $\|D\widehat{\alpha} - D\alpha\|_2 \gg \|\widehat{\alpha} - \alpha\|_2$ or $\|D\widehat{\alpha} - D\alpha\|_2 \ll \|\widehat{\alpha} - \alpha\|_2$

Signal-focused Recovery Strategy

- Focus on x instead of α
- Measure error in terms of $\|\widehat{x} - x\|_2$ instead of $\|\widehat{\alpha} - \alpha\|_2$

$$\sqrt{1-\delta_k}\|\alpha\|_2 \cdot \quad \|AD\alpha\|_2 \cdot \quad \sqrt{1+\delta_k}\|\alpha\|_2$$

$$\Downarrow$$

$$\sqrt{1-\delta_k}\|D\alpha\|_2 \cdot \quad \|AD\alpha\|_2 \cdot \quad \sqrt{1+\delta_k}\|D\alpha\|_2$$

CoSaMP

initialize: $r = y, x^0 = 0, \ell = 0, \Gamma = \emptyset$

until converged:

proxy: $h = A^* r$

identify: $\quad = \{2k$ largest elements of $|h|\}$

merge: $T = \quad \cup \Gamma$

update: $\widetilde{x} = \underset{\mathrm{supp}(z) \subseteq T}{\arg\min} \|y - Az\|_2$

$\Gamma = \{k$ largest elements of $|\widetilde{x}|\}$

$x^{\ell+1} = \widetilde{x}|_{\Gamma}$

$r^{\ell+1} = y - Ax^{\ell+1}$

$\ell = \ell + 1$

output: $\widehat{x} = x^{\ell}$

Key Steps

$$= \{2k \text{ largest elements of } |h|\}$$

$$\widetilde{x} = \underset{\mathrm{supp}(z) \subseteq T}{\arg\min} \|y - Az\|_2$$

$$\Gamma = \{k \text{ largest elements of } |\widetilde{x}|\}$$

$$x^{\ell+1} = \widetilde{x}|_{\Gamma}$$

Key Steps

$$= \{2k \text{ largest elements of } |h|\}$$

$$\Gamma = \{k \text{ largest elements of } |\widetilde{x}|\}$$
$$x^{\ell+1} = \widetilde{x}|_{\Gamma}$$

Given a vector in $\mathbb{R}^n$, use hard thresholding to find best sparse approximation

$\mathcal{P}_{\Lambda}$: orthogonal projector onto $\mathcal{R}(D_{\Lambda})$

$$\Lambda_{\text{opt}}(z, k) = \arg\min_{|\Lambda|=k} \|z - \mathcal{P}_{\Lambda} z\|_2$$

Approximate Projection

$\mathcal{P}_\Lambda$: orthogonal projector onto $\mathcal{R}(D_\Lambda)$

$$\Lambda_{\text{opt}}(z,k) = \arg\min_{|\Lambda|=k} \|z - \mathcal{P}_\Lambda z\|_2$$

$\mathcal{S}(z,k)$: estimate of $\Lambda_{\text{opt}}(z,k)$

$$\|\mathcal{P}_{\Lambda_{\text{opt}}} z - \mathcal{P}_{\mathcal{S}} z\|_2 \cdot \quad \min\left(\epsilon_1 \|\mathcal{P}_{\Lambda_{\text{opt}}} z\|_2,\ \epsilon_2 \|z - \mathcal{P}_{\Lambda_{\text{opt}}} z\|_2\right)$$

measure quality of approximation in "signal space", not "coefficient space"

Signal Space CoSaMP

initialize: $r = y, x^0 = 0, \ell = 0, \Gamma = \emptyset$

until converged:

proxy: $h = A^* r$

identify: $\quad = \mathcal{S}(h, 2k)$

merge: $T = \quad \cup \Gamma$

update: $\widetilde{x} = \underset{z \in \mathcal{R}(D_T)}{\arg\min} \|y - Az\|_2$

$\Gamma = \mathcal{S}(\widetilde{x}, k)$

$x^{\ell+1} = \mathcal{P}_\Gamma(\widetilde{x})$

$r^{\ell+1} = y - Ax^{\ell+1}$

$\ell = \ell + 1$

output: $\widehat{x} = x^\ell$

Recovery Guarantees

Suppose there exists a k-sparse α such that $x = D\alpha$ and that A satisfies the D-RIP of order $4k$.

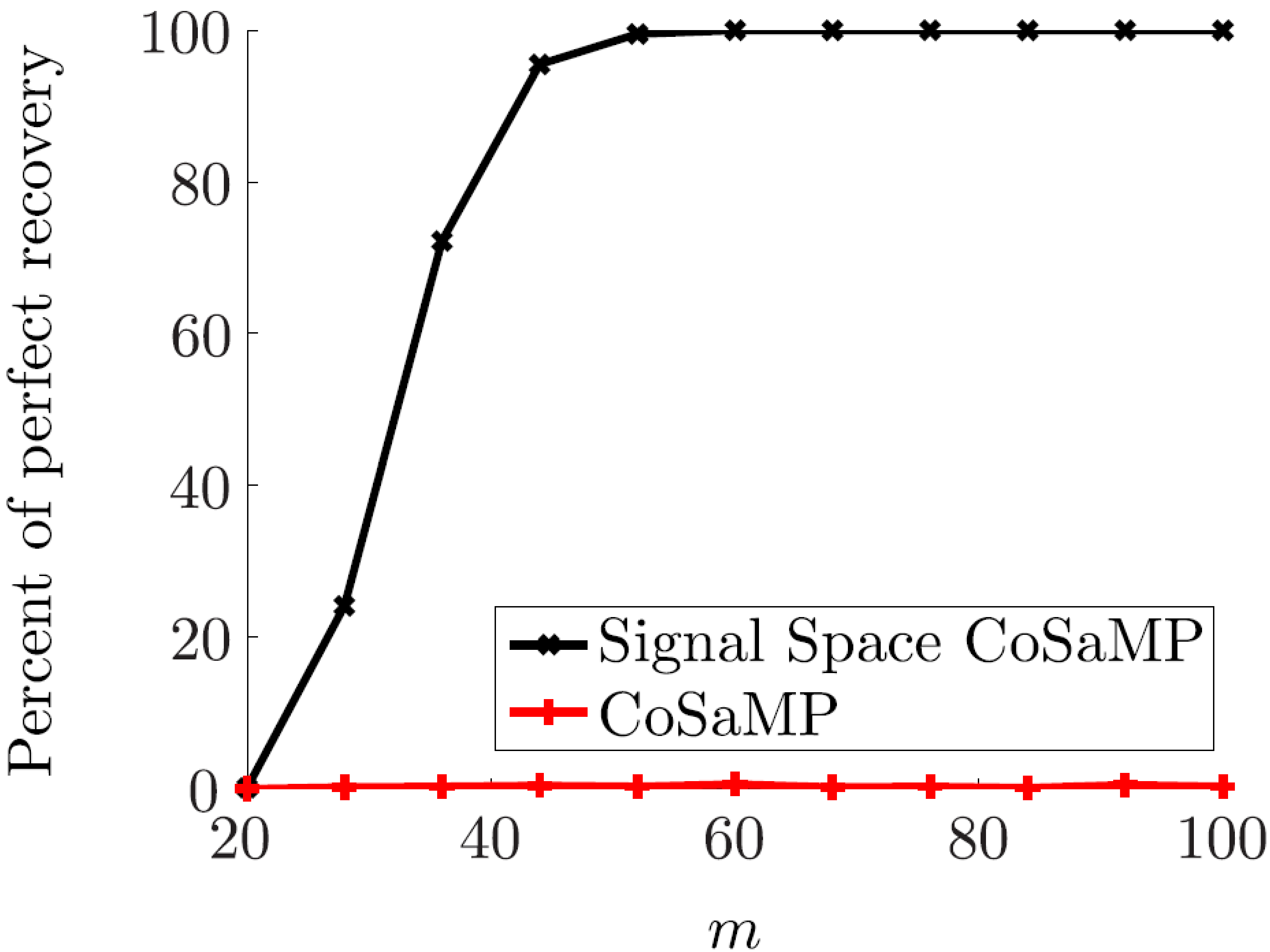
If we observe $y = Ax + e$, then

$$\|x - x^{\ell+1}\|_2 \cdot \; C_1\|x - x^{\ell}\|_2 + C_2\|e\|_2$$

For $\delta_{4k} = 0.029, \epsilon_1 = 0.1, \epsilon_2 = 1,$

$$\|x - x^{\ell}\|_2 \cdot \; 2^{-\ell}\|x\|_2 + 25.4\|e\|_2$$

Renormalized Columns



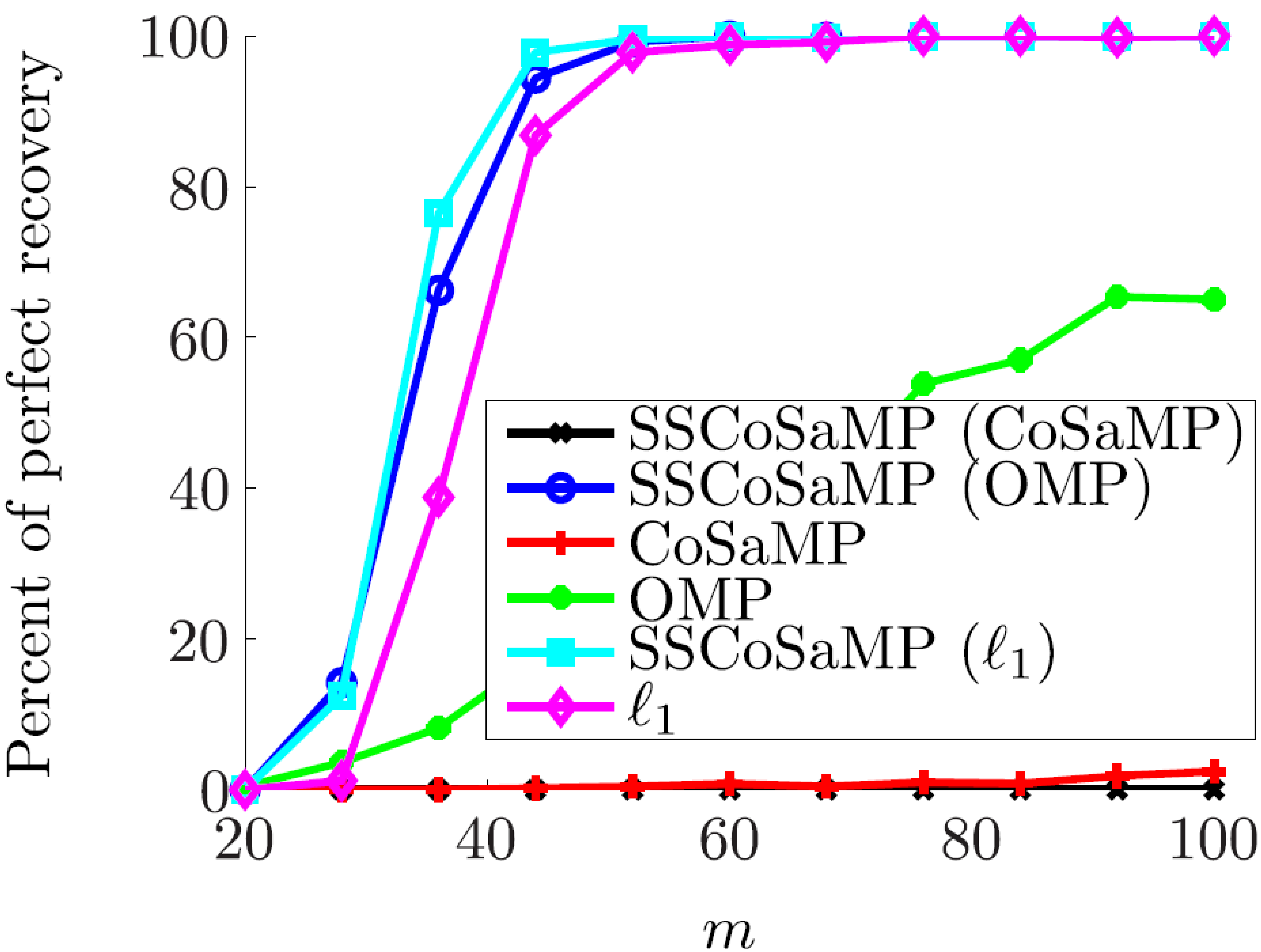
Practical Choices for $\mathcal{S}(z, k)$

- Given z, we want to find a k-sparse α such that $z \approx D\alpha$
- Any sparse recovery algorithm!
- CoSaMP
- Orthogonal Matching Pursuit (OMP)
- ℓ_1-minimization followed by hard-thresholding

$$\mathcal{S}(z, k) = H_k \left(\underset{w : Dw = z}{\arg\min} \|w\|_1 \right)$$

Overcomplete DFT

Separated coefficients



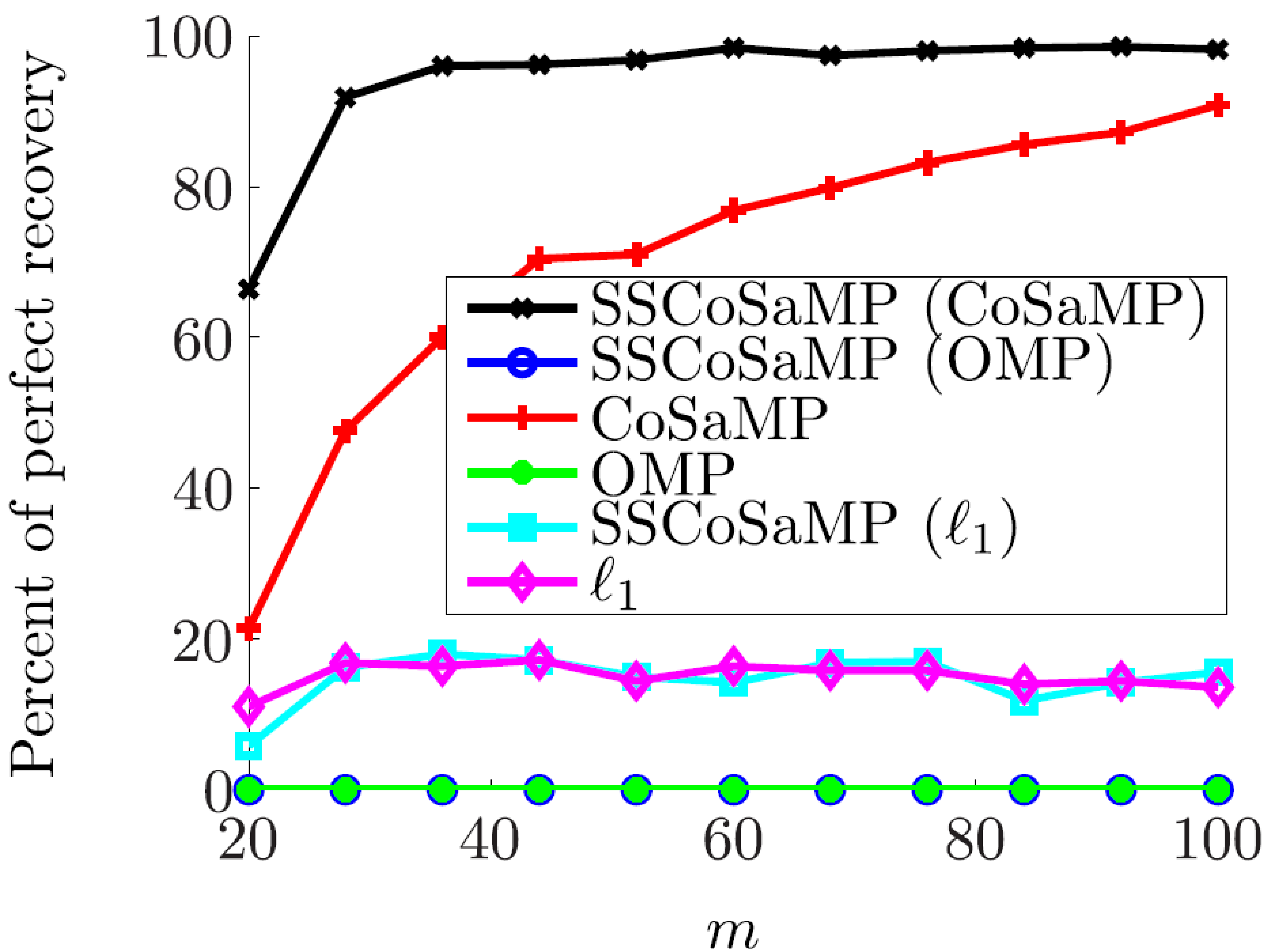
$k = 8$

$n = 256$

$d = 4n$

Overcomplete DFT

Clustered coefficients



$k = 8$

$n = 256$

$d = 4n$

Conclusions

- If we care about x, we should focus on x, not α
- Signal Space CoSaMP provides a theoretical framework for accommodating general dictionaries D
- Choice of $\mathcal{S}(z, k)$ is not obvious
  - existing methods appear to work well in practice
  - future work should focus on theoretical guarantees
- Success in this area has implications far beyond Signal Space CoSaMP
  - "signal space IHT"
  - analysis IHT/CoSaMP
  - ℓ_1- minimization