Greedy Signal Recovery and Uniform Uncertainty Principles

SPIE - IE 2008

Deanna Needell
Joint work with Roman Vershynin

UC Davis, January 2008

Outline

- Problem Background
  - Setup
  - L1 Minimization
  - Greedy Algorithms
- ROMP
  - Algorithm
  - Main Theorem
  - Empirical Results
- Future Work

Setup

- Consider $v \in \mathbb{R}^d$, $\|v\|_0 := |\operatorname{supp} v| \leq n \ll d$.

Setup

- Consider $v \in \mathbb{R}^d$, $\|v\|_0 := |\operatorname{supp} v| \leq n \ll d$.
- We call such signals n-*sparse*.

Setup

- Consider $v \in \mathbb{R}^d$, $\|v\|_0 := |\operatorname{supp} v| \leq n \ll d$.
- We call such signals n-*sparse*.
- Given some $N \times d$ measurement matrix Φ, we collect $N \ll d$ nonadaptive linear measurements of v, in the form $x = \Phi v$.

$$\left[\quad \Phi \quad \right] \left[v \right] = \left[x \right]$$

Setup ctd.

- From these measurements we wish to efficiently recover the original signal v.

Setup ctd.

- From these measurements we wish to efficiently recover the original signal v.
- How many measurements $N \ll d$ are needed?

Setup ctd.

- From these measurements we wish to efficiently recover the original signal v.
- How many measurements $N \ll d$ are needed?
- Exact recovery is possible with just $N = 2n$. However, recovery in this regime is not numerically feasible.

Setup ctd.

- Work in Compressed Sensing has shown that the signal v can be efficiently exactly recovered from $x = \Phi v$ with just $N \sim n \,\mathrm{polylog}\, d$.

Setup ctd.

- Work in Compressed Sensing has shown that the signal v can be efficiently exactly recovered from $x = \Phi v$ with just $N \sim n \operatorname{polylog} d$.
- Two major algorithmic approaches:

Setup ctd.

- Work in Compressed Sensing has shown that the signal v can be efficiently exactly recovered from $x = \Phi v$ with just $N \sim n \operatorname{polylog} d$.
- Two major algorithmic approaches:
- L1-Minimization (Donoho et. al.)

Setup ctd.

- Work in Compressed Sensing has shown that the signal v can be efficiently exactly recovered from $x = \Phi v$ with just $N \sim n \operatorname{polylog} d$.
- Two major algorithmic approaches:
- L1-Minimization (Donoho et. al.)
- Iterative methods such as Orthogonal Matching Pursuit (Tropp-Gilbert)

L1-Minimization Methods

- The sparse recovery problem can be stated as solving the optimization problem:

$$\min \|z\|_0 \qquad \text{subject to} \qquad \Phi z = \Phi v$$

L1-Minimization Methods

- The sparse recovery problem can be stated as solving the optimization problem:

$$\min \|z\|_0 \qquad \text{subject to} \qquad \Phi z = \Phi v$$

- For certain measurement matrices Φ this hard problem is equivalent to:

$$\min \|u\|_1 \qquad \text{subject to} \qquad \Phi u = \Phi v$$

(Donoho, Candès-Tao)

Restricted Isometry Condition

- A measurement matrix Φ satisfies the *Restricted Isometry Condition* (RIC) with parameters (m, ε) for $\varepsilon \in (0, 1)$ if we have

$$(1-\varepsilon)\|v\|_2 \leq \|\Phi v\|_2 \leq (1+\varepsilon)\|v\|_2 \quad \forall m\text{-sparse } v.$$

Restricted Isometry Condition

- A measurement matrix Φ satisfies the *Restricted Isometry Condition* (RIC) with parameters (m, ε) for $\varepsilon \in (0,1)$ if we have

$$(1-\varepsilon)\|v\|_2 \leq \|\Phi v\|_2 \leq (1+\varepsilon)\|v\|_2 \quad \forall m\text{-sparse } v.$$

- "Every set of n columns of Φ forms approximately an orthonormal system."

L1 and the RIC

- Assume that the measurement matrix Φ satisfies the Restricted Isometry Condition with parameters $(2n, \sqrt{2} - 1)$.

L1 and the RIC

- Assume that the measurement matrix Φ satisfies the Restricted Isometry Condition with parameters $(2n, \sqrt{2}-1)$.
- Then the L_1 method recovers any n-sparse vector. (Candes-Tao)

L1 and the RIC

- Assume that the measurement matrix Φ satisfies the Restricted Isometry Condition with parameters $(2n, \sqrt{2} - 1)$.
- Then the L_1 method recovers any n-sparse vector. (Candes-Tao)
- What kinds of matrices satisfy the RIC?

L1 and the RIC

- Assume that the measurement matrix Φ satisfies the Restricted Isometry Condition with parameters $(2n, \sqrt{2} - 1)$.
- Then the L_1 method recovers any n-sparse vector. (Candes-Tao)
- What kinds of matrices satisfy the RIC?
- Random Gaussian, Bernoulli, and partial Fourier matrices, with $N \sim n \operatorname{polylog} d$.

Greedy Algorithms: OMP

- Orthogonal Matching Pursuit (Tropp-Gilbert) finds the support of the n-sparse signal v progressively.

Greedy Algorithms: OMP

- Orthogonal Matching Pursuit (Tropp-Gilbert) finds the support of the n-sparse signal v progressively.
- Once $S = \mathrm{supp}(v)$ is found correctly, we can recover the signal: $x = \Phi v$ as $v = (\Phi_S)^{-1} x$.

Greedy Algorithms: OMP ctd.

- At each iteration, OMP finds the largest component of $u = \Phi^* x$ and subtracts off that component's contribution.

Greedy Algorithms: OMP ctd.

- At each iteration, OMP finds the largest component of $u = \Phi^* x$ and subtracts off that component's contribution.
- For every fixed n-sparse $v \in \mathbb{R}^d$, and an $N \times d$ Gaussian measurement matrix Φ, OMP recovers v with high probability, provided $N \sim n \log d$.

Comparing the approaches

- L_1 has uniform guarantees.

Comparing the approaches

- L_1 has uniform guarantees.
- OMP has no such known uniform guarantees.

Comparing the approaches

- L_1 has uniform guarantees.
- OMP has no such known uniform guarantees.
- L_1 is based on linear programming.

Comparing the approaches

- L_1 has uniform guarantees.
- OMP has no such known uniform guarantees.
- L_1 is based on linear programming.
- OMP is quite *fast*, both theoretically and experimentally.

Comparing the approaches

- This gap between the approaches leads us to our new algorithm, Regularized Orthogonal Matching Pursuit.

Comparing the approaches

- This gap between the approaches leads us to our new algorithm, Regularized Orthogonal Matching Pursuit.
- ROMP has polynomial running time.

Comparing the approaches

- This gap between the approaches leads us to our new algorithm, Regularized Orthogonal Matching Pursuit.
- ROMP has polynomial running time.
- ROMP provides uniform guarantees.

ROMP

- INPUT: Measurement vector $x \in \mathbb{R}^N$ and sparsity level n

ROMP

- INPUT: Measurement vector $x \in \mathbb{R}^N$ and sparsity level n
- OUTPUT: Index set $I \subset \{1, \ldots, d\}$

ROMP

- INPUT: Measurement vector $x \in \mathbb{R}^N$ and sparsity level n
- OUTPUT: Index set $I \subset \{1, \ldots, d\}$
- Initialize: Set $I = \emptyset$, $r = x$. Repeat until $r = 0$:

ROMP

- INPUT: Measurement vector $x \in \mathbb{R}^N$ and sparsity level n
- OUTPUT: Index set $I \subset \{1, \ldots, d\}$
- Initialize: Set $I = \emptyset$, $r = x$. Repeat until $r = 0$:
- Identify: Choose a set J of the n biggest coordinates in magnitude of $u = \Phi^* r$.

ROMP ctd.

- Regularize: Among all subsets $J_0 \subset J$ with comparable coordinates:

$$|u(i)| \leq 2|u(j)| \quad \text{for all } i, j \in J_0,$$

choose J_0 with the maximal energy $\|u|_{J_0}\|_2$.

ROMP ctd.

- Regularize: Among all subsets $J_0 \subset J$ with comparable coordinates:

$$|u(i)| \leq 2|u(j)| \quad \text{for all } i, j \in J_0,$$

  choose J_0 with the maximal energy $\|u|_{J_0}\|_2$.
- Update: the index set: $I \leftarrow I \cup J_0$, and the residual:

$$y = \operatorname*{argmin}_{z \in \mathbb{R}^I} \|x - \Phi z\|_2; \qquad r = x - \Phi y.$$

Main Theorem

- Theorem: Stability under measurement perturbations.
  Assume a measurement matrix Φ satisfies the Restricted Isometry Condition with parameters $(4n, .01/\sqrt{\log n})$.

Main Theorem

- Theorem: Stability under measurement perturbations.
  Assume a measurement matrix Φ satisfies the Restricted Isometry Condition with parameters $(4n, .01/\sqrt{\log n})$.
- Let v be an n-sparse vector in $\mathbb{R}^d$.

Main Theorem

- Theorem: Stability under measurement perturbations.
  Assume a measurement matrix Φ satisfies the Restricted Isometry Condition with parameters $(4n, .01/\sqrt{\log n})$.
- Let v be an n-sparse vector in $\mathbb{R}^d$.
- Consider corrupted $x = \Phi v + e$.

Main Theorem

- Theorem: Stability under measurement perturbations.
  Assume a measurement matrix Φ satisfies the Restricted Isometry Condition with parameters $(4n, .01/\sqrt{\log n})$.
- Let v be an n-sparse vector in $\mathbb{R}^d$.
- Consider corrupted $x = \Phi v + e$.
- Then ROMP produces a good approximation to v:

$$\|v - \hat{v}\|_2 \leq C\sqrt{\log n}\|e\|_2.$$

Corollary

- Stability of ROMP under signal perturbations. Assume a measurement matrix Φ satisfies the Restricted Isometry Condition with parameters $(8n, .01/\sqrt{\log n})$.

Corollary

- Stability of ROMP under signal perturbations. Assume a measurement matrix Φ satisfies the Restricted Isometry Condition with parameters $(8n, .01/\sqrt{\log n})$.
- Let v be an arbitrary vector in $\mathbb{R}^d$.

Corollary

- Stability of ROMP under signal perturbations. Assume a measurement matrix Φ satisfies the Restricted Isometry Condition with parameters $(8n, .01/\sqrt{\log n})$.
- Let v be an arbitrary vector in $\mathbb{R}^d$.
- Consider corrupted $x = \Phi v + e$.

Corollary

- Stability of ROMP under signal perturbations. Assume a measurement matrix Φ satisfies the Restricted Isometry Condition with parameters $(8n, .01/\sqrt{\log n})$.
- Let v be an arbitrary vector in $\mathbb{R}^d$.
- Consider corrupted $x = \Phi v + e$.
- Then ROMP produces a good approximation to v_{2n}:

$$\|\hat{v} - v_{2n}\|_2 \leq C' \sqrt{\log n} \Big(\|e\|_2 + \frac{\|v - v_n\|_1}{\sqrt{n}} \Big).$$

Remarks

- In the noiseless case $(e = 0)$, note that the theorem guarantees exact reconstruction.

Remarks

- In the noiseless case $(e = 0)$, note that the theorem guarantees exact reconstruction.
- The runtime is polynomial: In the case of unstructured matrices, the runtime is $O(dNn)$.

Remarks

- In the noiseless case $(e = 0)$, note that the theorem guarantees exact reconstruction.
- The runtime is polynomial: In the case of unstructured matrices, the runtime is $O(dNn)$.
- The theorem gives *uniform guarantees* of sparse recovery.

Remarks

- In the noiseless case $(e = 0)$, note that the theorem guarantees exact reconstruction.
- The runtime is polynomial: In the case of unstructured matrices, the runtime is $O(dNn)$.
- The theorem gives *uniform guarantees* of sparse recovery.
- ROMP succeeds with no prior knowledge about the error vector e.

Empirical Results

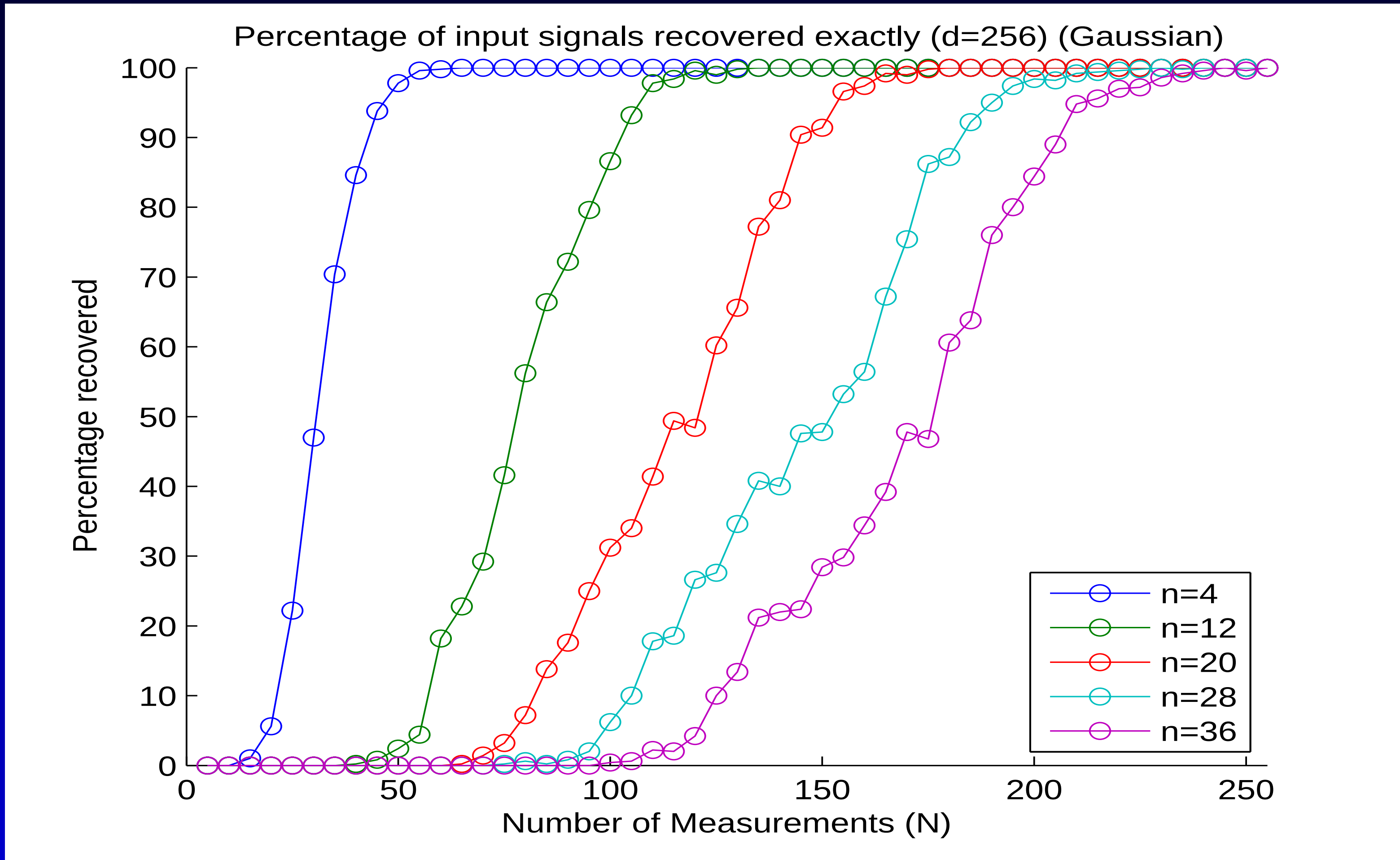

Figure 1: Sparse flat signals with Gaussian matrix.

Empirical Results ctd.

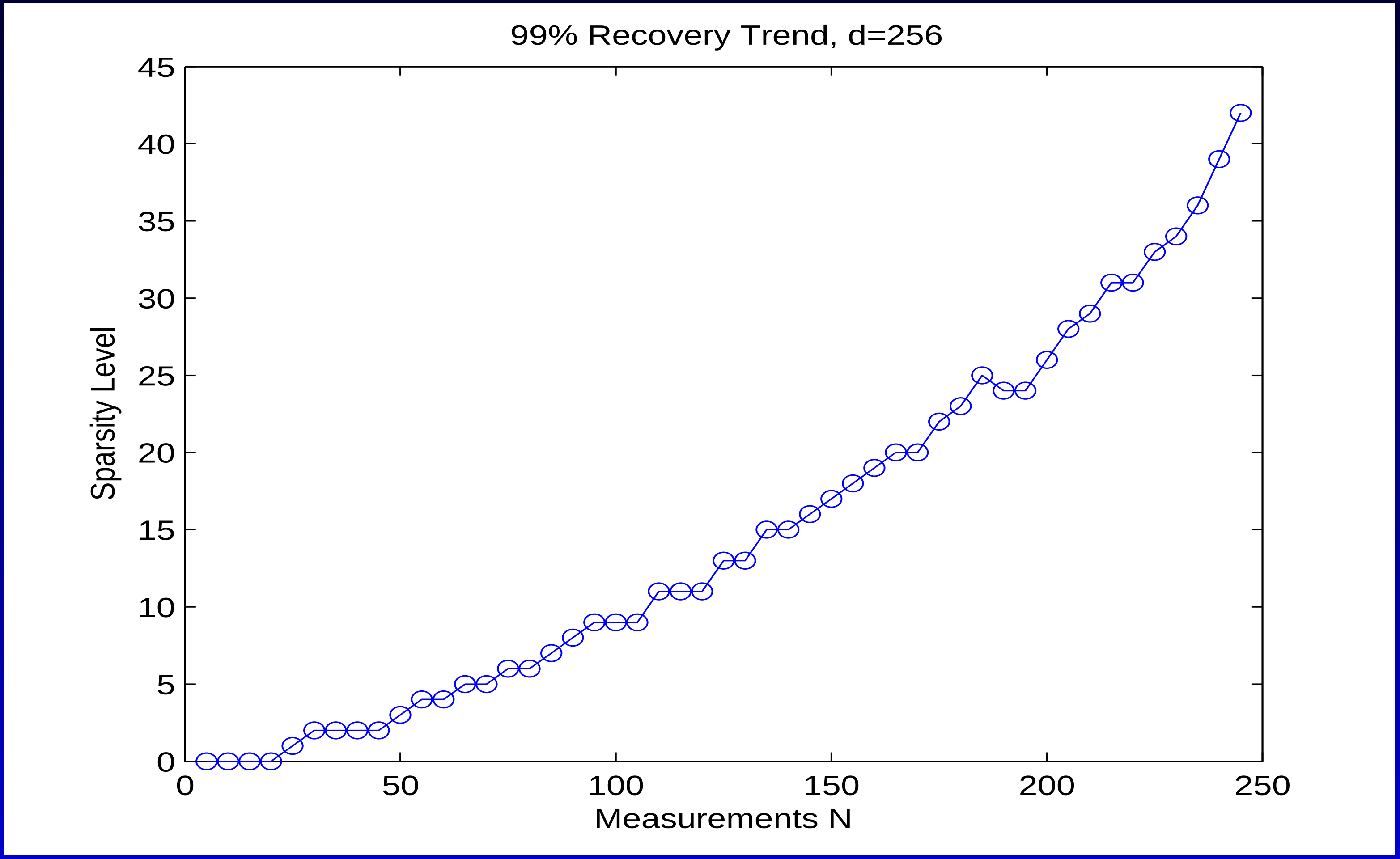

Figure 2: Sparse flat signals, Gaussian.

Empirical Results ctd.

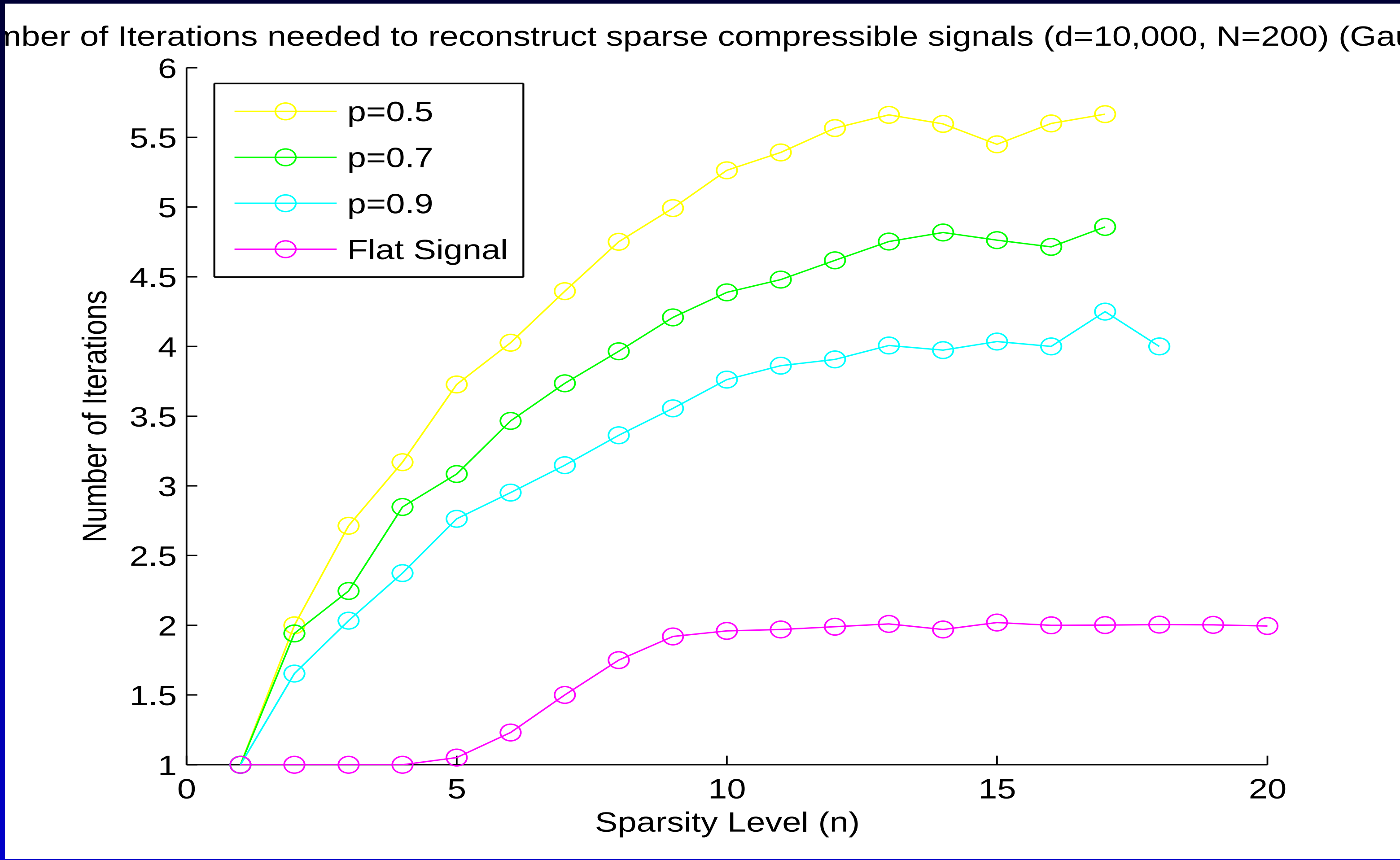

Figure 3: Number of Iterations.

Empirical Results ctd.

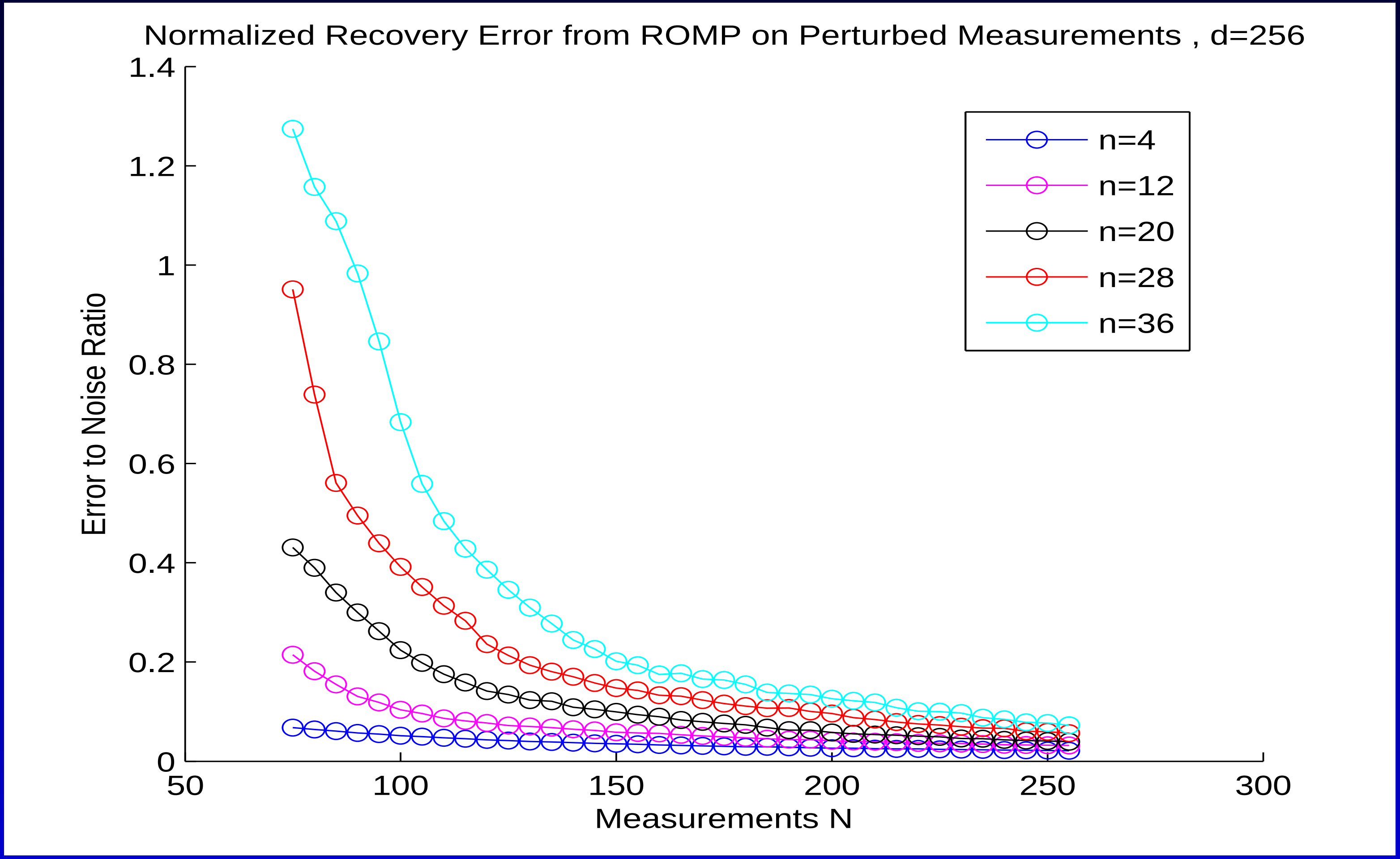

Figure 4: Error to noise ratio $\frac{\|\hat{v}-v\|_2}{\|e\|_2}$.

Empirical Results ctd.

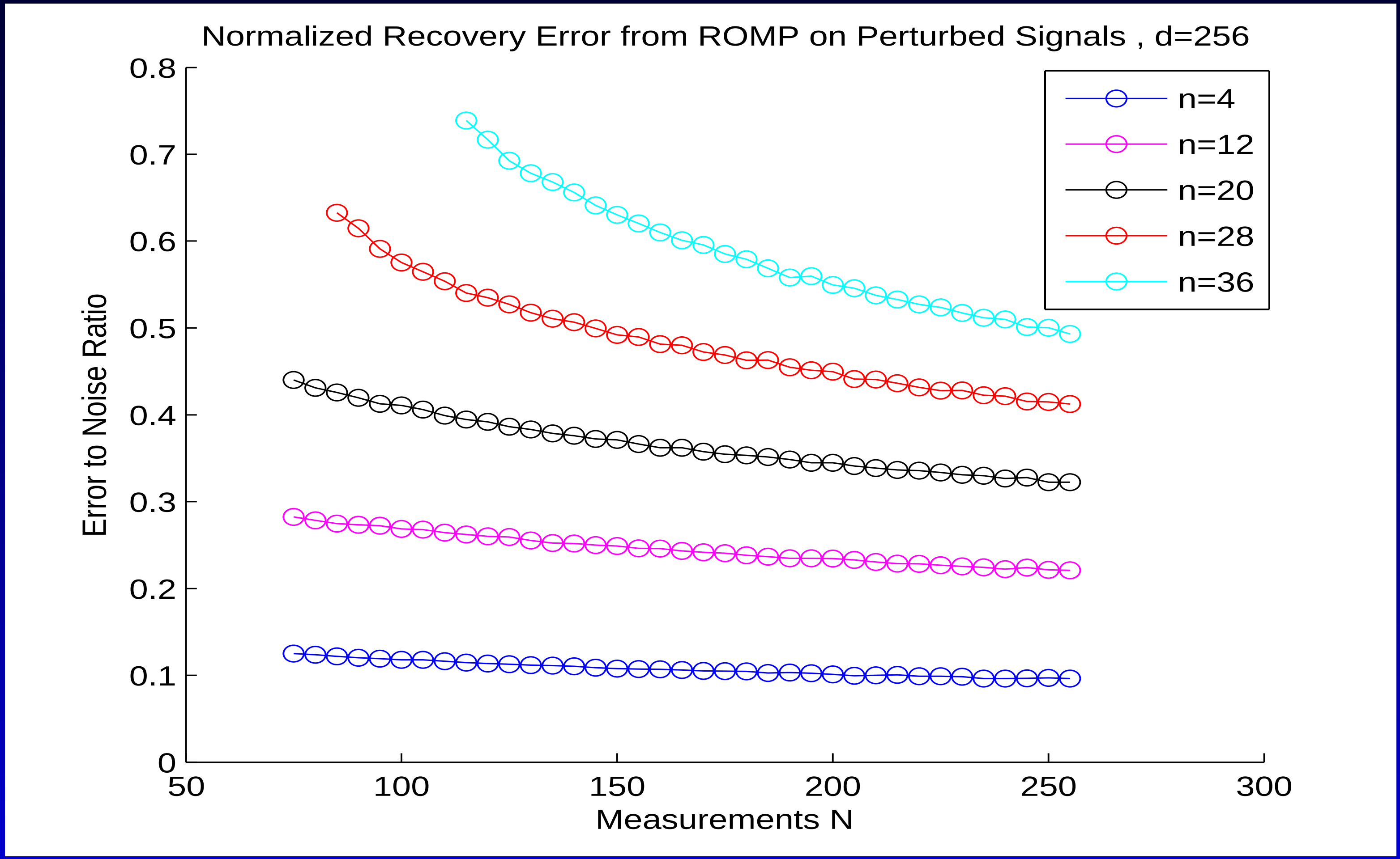

Figure 5: Error to noise ratio $\frac{\|\hat{v}-v_{2n}\|_2}{\|v-v_n\|_1/\sqrt{n}}$.

Future Work

- N-Tropp-Vershynin developing Compressive Sampling Matching Pursuit (CoSaMP)

Future Work

- N-Tropp-Vershynin developing Compressive Sampling Matching Pursuit (CoSaMP)
- Selects $O(n)$ coordinates at each iteration but adds a signal estimation step using least squares. Then prunes this estimation to make sparse.

Future Work

- N-Tropp-Vershynin developing Compressive Sampling Matching Pursuit (CoSaMP)
- Selects $O(n)$ coordinates at each iteration but adds a signal estimation step using least squares. Then prunes this estimation to make sparse.
- Same uniform guarantees as ROMP, but removes the $\sqrt{\log n}$ term in the requirement for ε

Thank you!

- Questions?