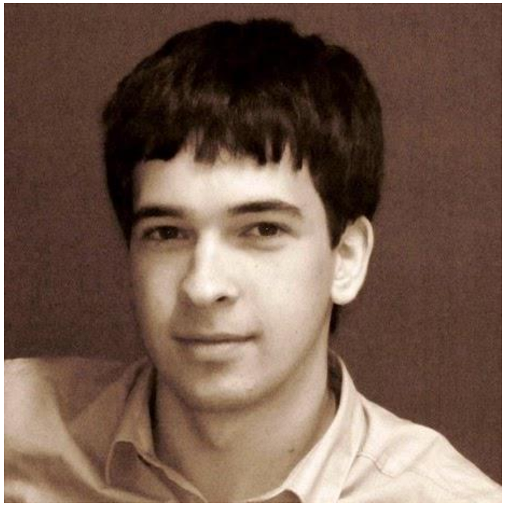
Constrained adaptive sensing

Deanna Needell

Claremont McKenna College
Dept. of Mathematics

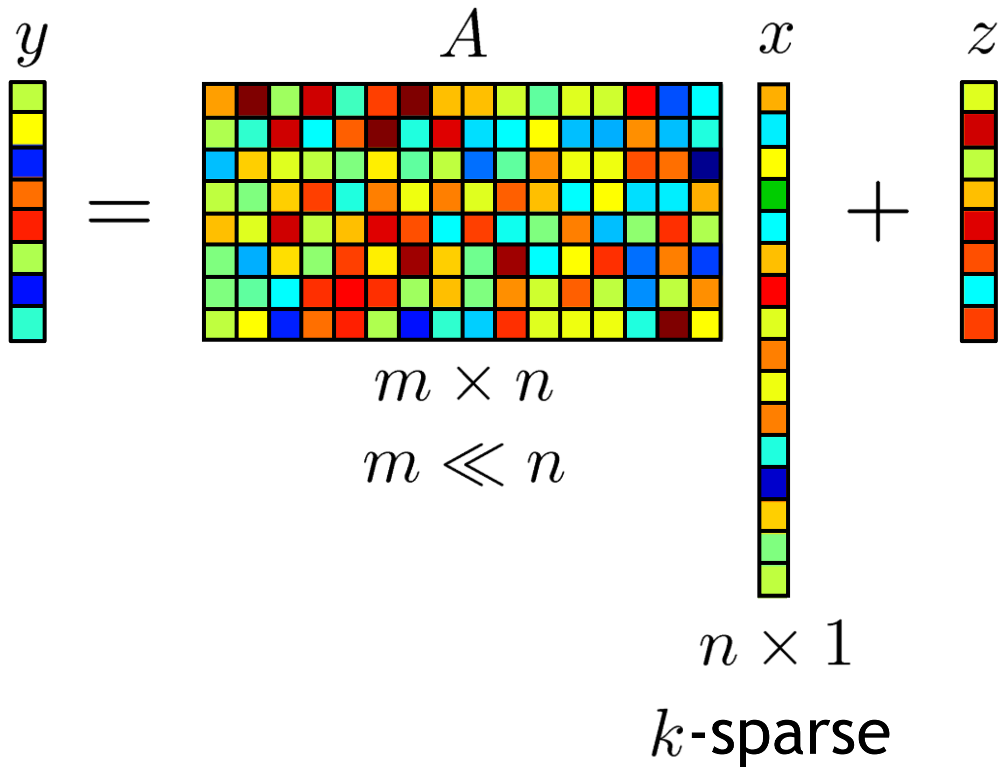
Mark
Davenport



Andrew
Massimino

Tina
Woolf

Sensing sparse signals



When (and how well) can we estimate x from the measurements y?

Nonadaptive sensing

Prototypical sensing model:

$$y = Ax + z \qquad z \sim \mathcal{N}(0, \sigma^2 I)$$

There exist matrices A and recovery algorithms that produce an estimate $\widehat{x}$ such that for *any* x with $\|x\|_0 \leq k$ we have

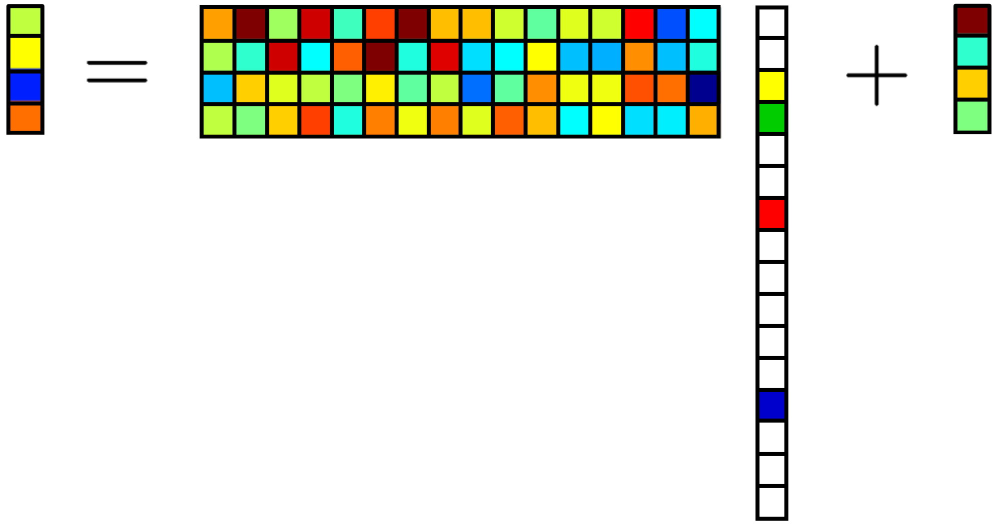
$$\mathbb{E}\|\widehat{x} - x\|_2^2 \leq C \frac{n \log n}{\|A\|_F^2} k\sigma^2.$$

For ***any*** matrix A and ***any*** recovery algorithm $\widehat{x}$, there exist x with $\|x\|_0 \leq k$ such that

$$\mathbb{E}\|\widehat{x} - x\|_2^2 \leq C' \frac{n \log(n/k)}{\|A\|_F^2} k\sigma^2.$$

Adaptive sensing

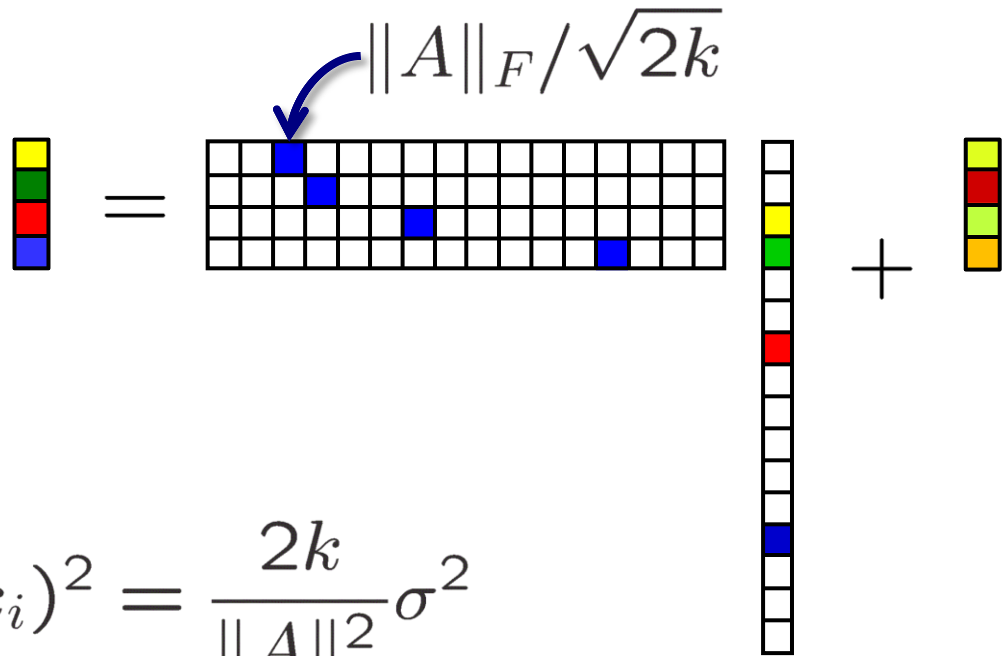
Think of sensing as a game of 20 questions



Simple strategy: Use half of our sensing energy to find the support, and the remainder to estimate the values.

Thought experiment

Suppose that after the first stage we have perfectly estimated the support



$$\mathbb{E}(\widehat{x}_i - x_i)^2 = \frac{2k}{\|A\|_F^2}\sigma^2$$

$$\mathbb{E}\|\widehat{x}_i - x_i\|^2 = \frac{2k}{\|A\|_F^2}k\sigma^2 \ll \frac{n \log n}{\|A\|_F^2}k\sigma^2$$

Benefits of adaptivity

Adaptivity offers the ***potential*** for tremendous benefits

Suppose we wish to estimate a 1-sparse vector whose nonzero has amplitude μ:

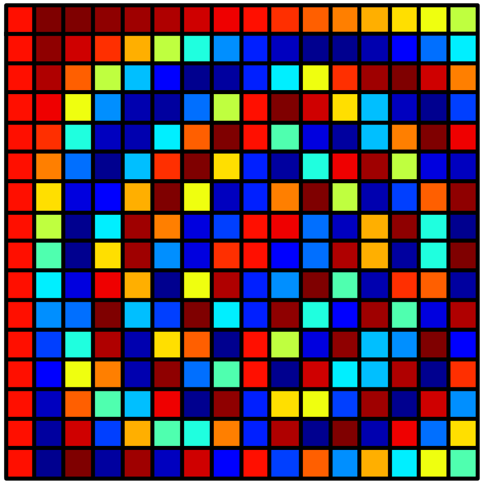
- No method can find the nonzero when $\frac{\mu^2}{\sigma^2} \approx n/\|A\|_F^2$
- A simple binary search procedure will succeed in finding the location of the nonzero with probability $1-\delta$ when $\frac{\mu^2}{\sigma^2} > 16n\log(\frac{1}{2\delta}+1)/\|A\|_F^2$
- Not hard to extend to k-sparse vectors
- See Arias-Castro, Candès, Davenport; Castro; Malloy, Nowak

Provided that the SNR is sufficiently large, adaptivity can reduce our error by a factor of n/k!

Sensing with constraints

Existing approaches to adaptivity require the ability to acquire ***arbitrary*** linear measurements, but in many (most?) real-world systems, our measurements are ***highly constrained***

Suppose we are limited to using measurement vectors chosen from some fixed (finite) ensemble $\mathcal{M} = \{a_1, a_2, \ldots, a_M\}$



- How much room for improvement do we have in this case?
- How should we actually go about adaptively selecting our measurements?

Room for improvement?

It depends!

If x is k-sparse and the a_i are chosen (potentially adaptively) by selecting up to m rows from the DFT matrix, then for ***any*** adaptive scheme we will have

$$\mathbb{E}\|\widehat{x} - x\|_2^2 \geq \frac{n}{m} k \sigma^2$$

On the other hand, if $\mathcal{M} = \{a_1, a_2, \ldots, a_M\}$ contains vectors which are better aligned with our class of signals (*or if x is sparse in an alternative basis/dictionary*), then dramatic improvements may still be possible

How to adapt?

Suppose we knew the locations of the nonzeros

$$\Lambda = \mathrm{supp}(x)$$

One can show that the error in this case is given by

$$\mathbb{E}\|\widehat{x} - x\|_2^2 = \|A_\Lambda^\dagger\|_F^2 \sigma^2 = \mathrm{tr}\left((A_\Lambda^* A_\Lambda)^{-1} \right) \sigma^2$$

Ideally, we would like to choose a sequence $\{a_i\}_{i=1}^m$ according to

$$\{\widehat{a}_i\}_{i=1}^m = \underset{\{a_i\}_{i=1}^m : a_i \in \mathcal{M}}{\mathrm{argmin}} \ \mathrm{tr}\left((A_\Lambda^* A_\Lambda)^{-1} \right)$$

where here A denotes the matrix with rows given by the sequence $\{a_i\}_{i=1}^m$

A toy problem

- Suppose our signal is 1-sparse (in Haar wavelets) and after m/2 measurements we know the location Λ.
- Which Fourier measurements, and how is error?
- We want to minimize $\|(\mathbf{F}'\mathbf{H}_\Lambda^*)^\dagger\|_F^2 = \sum_{i=1}^{s} \frac{1}{\sigma_i^2}$
- Thus we want to maximize $\|\mathbf{F}'\mathbf{H}_\Lambda^*\|_F^2 = \sigma_1^2$
- One easily optimizes by repeatedly sampling with row f_j

A toy problem

- The MSE can be computed as

$$\mathbb{E}\|\widehat{\mathbf{x}}-\mathbf{x}\|_2^2 = \frac{2/m}{|\langle \mathbf{f}_j, \mathbf{H}_\Lambda^* \rangle|^2}\sigma^2$$

- And is bounded b $\displaystyle \frac{2\sigma^2}{m} \le \mathbb{E}\|\widehat{\mathbf{x}}-\mathbf{x}\|_2^2 \le \frac{n\sigma^2}{m}$

matches lower bound

Room for improvement?

It depends!

If x is k-sparse and the a_i are chosen (potentially adaptively) by selecting up to m rows from the DFT matrix, then for ***any*** adaptive scheme we will have

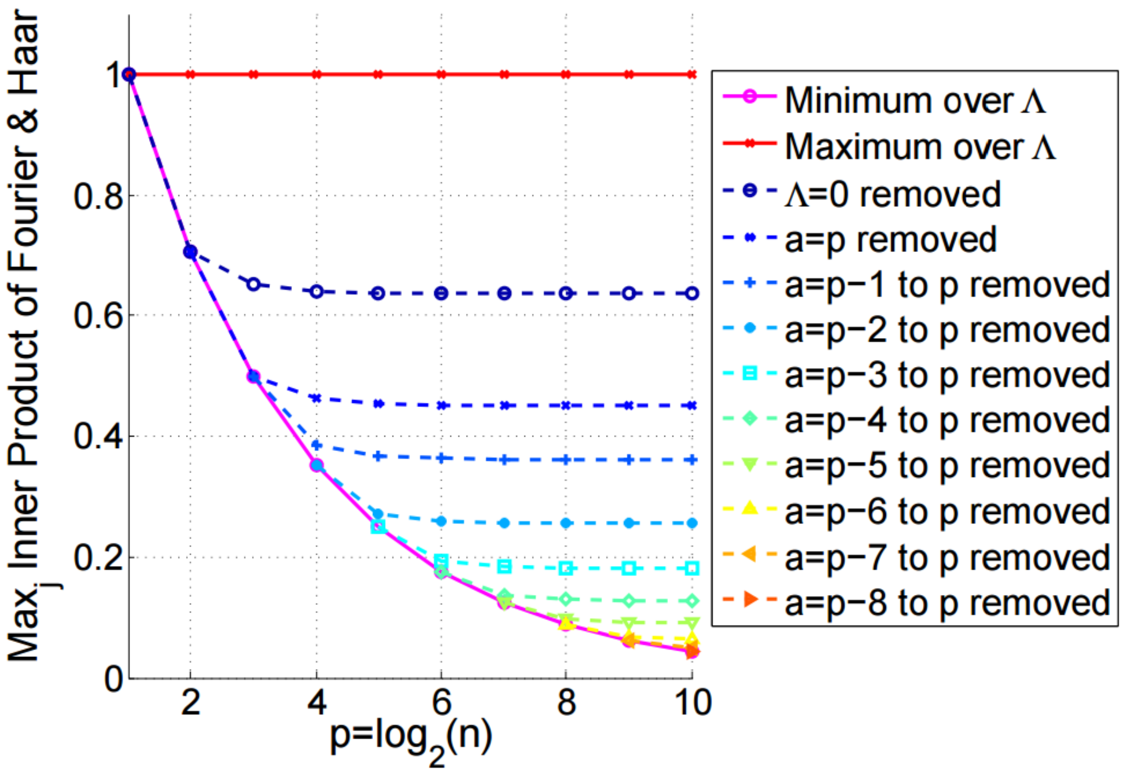
$$\mathbb{E}\|\widehat{x}-x\|_2^2 \ge \frac{n}{m}k\sigma^2$$

On the other hand, if $\mathcal{M} = \{a_1, a_2, \ldots, a_M\}$ contains vectors which are better aligned with our class of signals (*or if x is sparse in an alternative basis/dictionary*), then dramatic improvements may still be possible

A toy problem

- "How many" signals actually benefit?

$$\frac{2\sigma^2}{m} \leq \mathbb{E}\|\widehat{\mathbf{x}} - \mathbf{x}\|_2^2 \leq \frac{n\sigma^2}{m} \qquad \mathbb{E}\|\widehat{\mathbf{x}} - \mathbf{x}\|_2^2 = \frac{2/m}{|\langle \mathbf{f}_j, \mathbf{H}_\Lambda^* \rangle|^2}\sigma^2$$



$$H_8 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & -1 & -1 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}$$

How to adapt?

Suppose we knew the locations of the nonzeros

$$\Lambda = \mathrm{supp}(x)$$

One can show that the error in this case is given by

$$\mathbb{E}\|\widehat{x} - x\|_2^2 = \|A_\Lambda^\dagger\|_F^2 \sigma^2 = \mathrm{tr}\left((A_\Lambda^* A_\Lambda)^{-1}\right)\sigma^2$$

Ideally, we would like to choose a sequence $\{a_i\}_{i=1}^m$ according to

$$\{\widehat{a}_i\}_{i=1}^m = \operatorname*{argmin}_{\{a_i\}_{i=1}^m : a_i \in \mathcal{M}} \mathrm{tr}\left((A_\Lambda^* A_\Lambda)^{-1}\right)$$

where here A denotes the matrix with rows given by the sequence $\{a_i\}_{i=1}^m$

Convex relaxation

We would like to solve $\{\widehat{a}_i\}_{i=1}^m = \underset{\{a_i\}_{i=1}^m : a_i \in \mathcal{M}}{\text{argmin}} \ \text{tr}\left((A_\Lambda^* A_\Lambda)^{-1}\right)$

Instead we consider the relaxation

$$\widehat{S} = \underset{\text{diagonal matrices } S \succeq 0}{\text{argmin}} \ \text{tr}\left((A_\Lambda^* S A_\Lambda)^{-1}\right)$$

$$\text{subject to } \text{tr}(S) \leq \mathcal{E}$$

The diagonal entries of $\widehat{S}$ tell us "how much" of each sensing vector we should use, and the constraint $\text{tr}(S) \leq \mathcal{E}$ ensures that $\|\sqrt{S}A\|_F^2 \leq \mathcal{E}$ (assuming A has unit-norm rows)

Equivalent to notion of "A-optimality" criterion in optimal experimental design

Generating the sensing matrix

In practice, S tends to be somewhat sparse, placing high weight on a small number of measurements and low weights on many others

Where "sensing energy" is the operative constraint (as opposed to number of measurements) we can use $\sqrt{S}A$ directly to sense

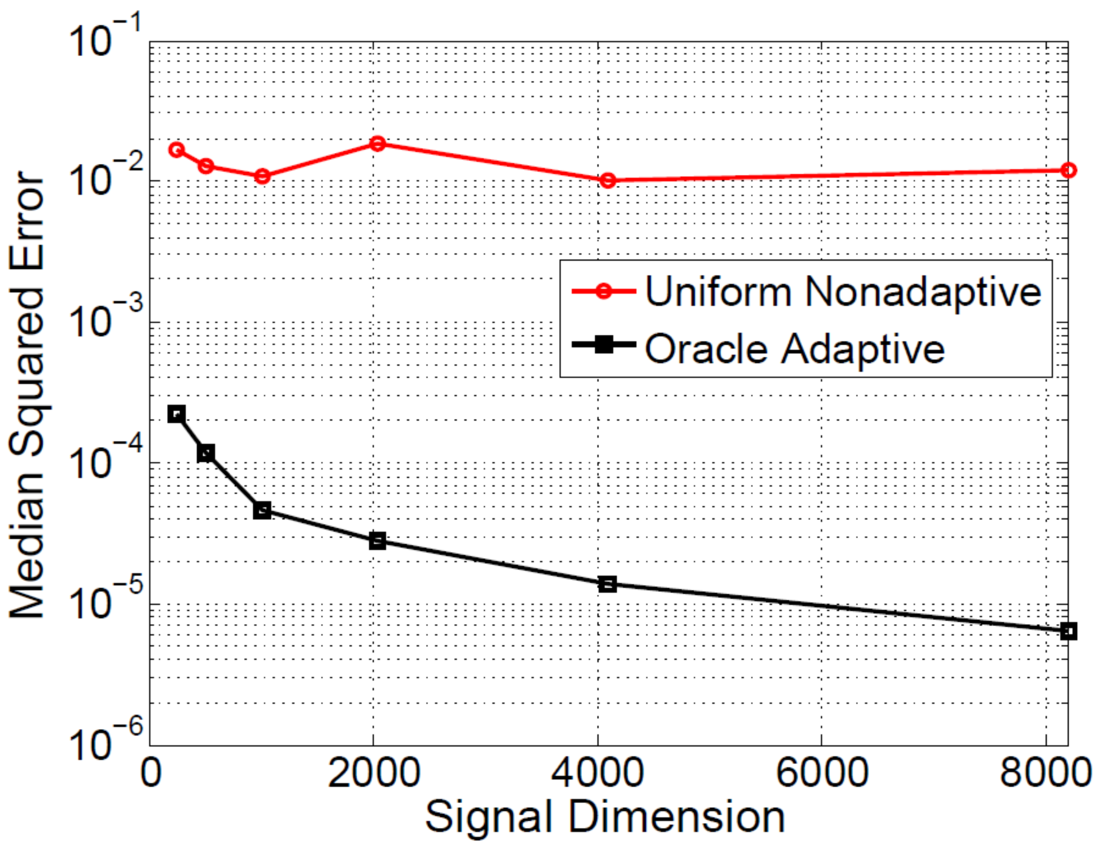
If we wish to take exactly m measurements, one option is to draw m measurement vectors by sampling with replacement according to the probability mass function

$$p_i = \frac{\widehat{s}_{ii}}{\mathcal{E}}$$

Example

DFT measurements of signal with sparse Haar wavelet transform (supported on connected tree)

Recovery performed using CoSaMP



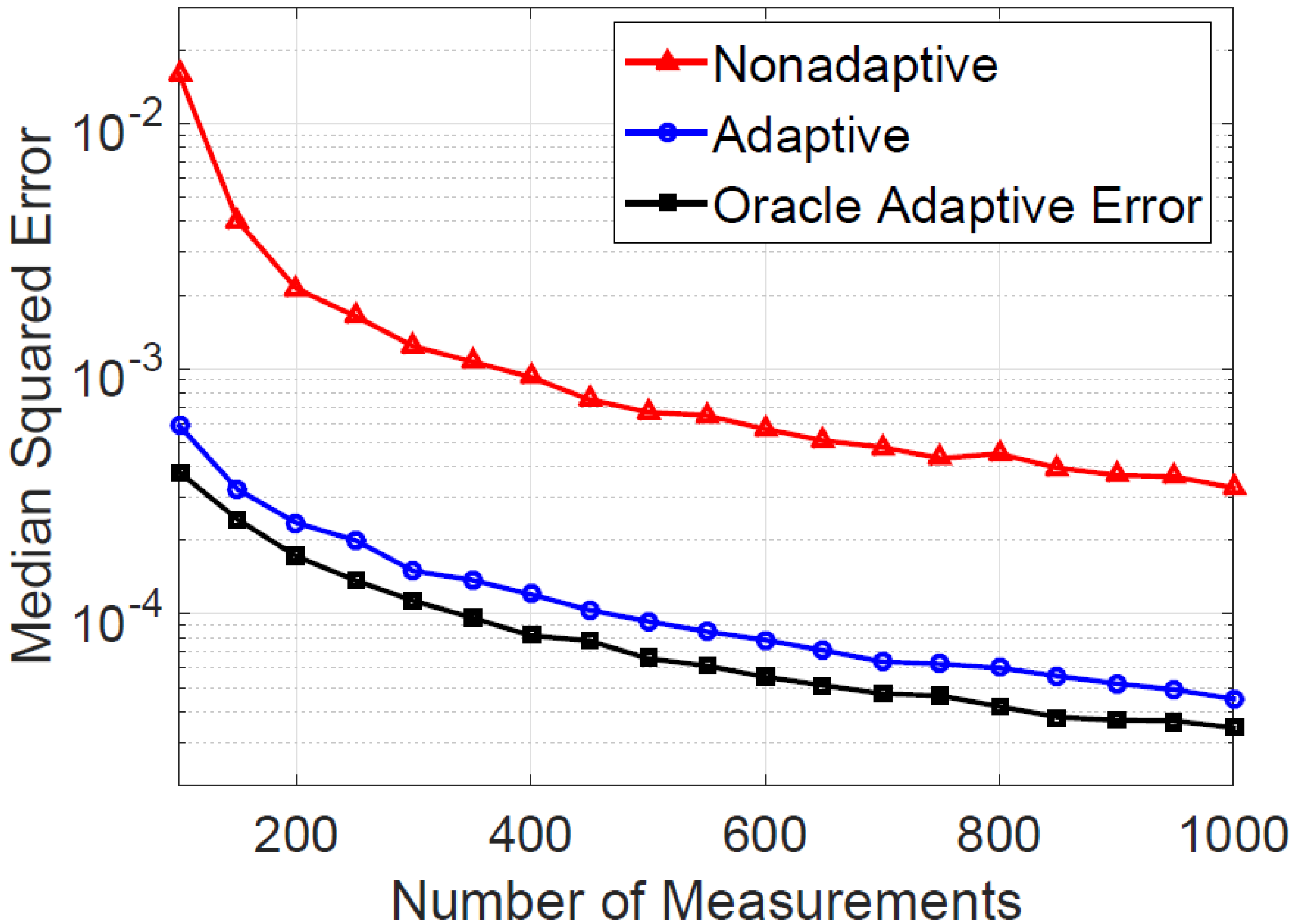
Constrained sensing in practice

The "oracle adaptive" approach can be used as a building block for a practical algorithm

Simple approach:

- Divide sensing energy / measurements in half
- Use first half by randomly selecting measurement vectors and using a conventional sparse recovery algorithm to estimate the support
- Use this support estimate to choose second half of measurements

Simulation results



Summary

- Adaptivity (sometimes) allows tremendous improvements
- Not always easy to realize these improvements in the constrained setting
  - existing algorithms not applicable
  - room for improvement may not be quite as large
- Simple strategies for adaptively selecting the measurements based on convex optimization can be surprisingly effective

Thank You!

arXiv:1506.05889

http://www.cmc.edu/pages/faculty/DNeedell/