A Short Introduction to Compressive Sensing

Simon Foucart
University of Georgia

Semester Program on “High-Dimensional Approximation”
ICERM
26 September 2014
Compressive Sensing

Keywords:
- Sparsity (Essential)
- Randomness (Nothing better so far)
- Optimization (Merely preferred, but alternatives exist)

Classical theory now found in:
Compressive Sensing

Keywords:

Sparsity (Essential)

Randomness (Nothing better so far)

Optimization (Merely preferred, but alternatives exist)

Classical theory now found in:
Compressive Sensing

Keywords:

- **Sparsity**
  (Essential)
Compressive Sensing

Keywords:

- Sparsity  
  (Essential)
- Randomness  
  (Nothing better so far)
Compressive Sensing

Keywords:

- Sparsity
  (Essential)
- Randomness
  (Nothing better so far)
- Optimization
  (Merely preferred, but alternatives exist)
Compressive Sensing

Keywords:

- Sparsity (Essential)
- Randomness (Nothing better so far)
- Optimization (Merely preferred, but alternatives exist)

Classical theory now found in:

A Mathematical Introduction to Compressive Sensing

Simon Foucart
Holger Rauhut
Birkhäuser
Part I:

The Standard Problem and First Algorithms
Initial Analysis
The Standard Compressive Sensing Problem

$x$: unknown signal of interest in $\mathbb{K}^N$

$y$: measurement vector in $\mathbb{K}^m$ with $m \ll N$

$s$: sparsity of $x = \text{card}\{j \in \{1, \ldots, N\}: x_j \neq 0\}$

Find concrete sensing/recovery protocols, i.e., find

$\begin{align*}
  &\text{measurement matrices } A: x \in \mathbb{K}^N \mapsto y \in \mathbb{K}^m \\
  &\text{reconstruction maps } \Delta: y \in \mathbb{K}^m \mapsto x \in \mathbb{K}^N
\end{align*}$

such that $\Delta(Ax) = x$ for any $s$-sparse vector $x \in \mathbb{K}^N$.

In realistic situations, two issues to consider:

Stability: $x$ not sparse but compressible,

Robustness: measurement error in $y = Ax + e$. 

The Standard Compressive Sensing Problem

\[ \mathbf{x} : \text{unknown signal of interest in } \mathbb{K}^N \]
The Standard Compressive Sensing Problem

\[ x : \text{unknown signal of interest in } \mathbb{K}^N \]
\[ y : \text{measurement vector in } \mathbb{K}^m \]
The Standard Compressive Sensing Problem

\( x \) : unknown signal of interest in \( \mathbb{K}^N \)

\( y \) : measurement vector in \( \mathbb{K}^m \) with \( m \ll N \),
The Standard Compressive Sensing Problem

\(\mathbf{x}\) : unknown signal of interest in \(\mathbb{K}^N\)
\(\mathbf{y}\) : measurement vector in \(\mathbb{K}^m\) with \(m \ll N\),
\(s\) : sparsity of \(\mathbf{x}\)
The Standard Compressive Sensing Problem

\( \mathbf{x} : \) unknown signal of interest in \( \mathbb{K}^N \)

\( \mathbf{y} : \) measurement vector in \( \mathbb{K}^m \) with \( m \ll N \),

\( s : \) sparsity of \( \mathbf{x} = \text{card}\{j \in \{1, \ldots, N\} : x_j \neq 0\} \).
The Standard Compressive Sensing Problem

\( x \): unknown signal of interest in \( \mathbb{K}^N \)
\( y \): measurement vector in \( \mathbb{K}^m \) with \( m \ll N \),
\( s \): sparsity of \( x = \text{card}\{ j \in \{1, \ldots, N\} : x_j \neq 0 \} \).

Find concrete sensing/recovery protocols,
The Standard Compressive Sensing Problem

\[ x \]: unknown signal of interest in \( \mathbb{K}^N \)

\[ y \]: measurement vector in \( \mathbb{K}^m \) with \( m \ll N \),

\[ s \]: sparsity of \( x = \text{card}\{j \in \{1, \ldots, N\} : x_j \neq 0\} \).

Find concrete sensing/recovery protocols, i.e., find

- measurement matrices \( A : x \in \mathbb{K}^N \mapsto y \in \mathbb{K}^m \)
The Standard Compressive Sensing Problem

\[ x : \text{unknown signal of interest in } \mathbb{K}^N \]
\[ y : \text{measurement vector in } \mathbb{K}^m \text{ with } m \ll N, \]
\[ s : \text{sparsity of } x = \text{card}\{j \in \{1, \ldots, N\} : x_j \neq 0\}. \]

Find concrete sensing/recovery protocols, i.e., find

- measurement matrices \( A : x \in \mathbb{K}^N \mapsto y \in \mathbb{K}^m \)
- reconstruction maps \( \Delta : y \in \mathbb{K}^m \mapsto x \in \mathbb{K}^N \)
The Standard Compressive Sensing Problem

\[ x \text{ : unknown signal of interest in } \mathbb{K}^N \]
\[ y \text{ : measurement vector in } \mathbb{K}^m \text{ with } m \ll N, \]
\[ s \text{ : sparsity of } x = \text{card}\{j \in \{1, \ldots, N\} : x_j \neq 0\}. \]

Find concrete sensing/recovery protocols, i.e., find

- measurement matrices \( A : x \in \mathbb{K}^N \mapsto y \in \mathbb{K}^m \)
- reconstruction maps \( \Delta : y \in \mathbb{K}^m \mapsto x \in \mathbb{K}^N \)

such that

\[ \Delta(Ax) = x \text{ for any } s\text{-sparse vector } x \in \mathbb{K}^N. \]
The Standard Compressive Sensing Problem

\[ \mathbf{x} : \text{unknown signal of interest in } \mathbb{K}^N \]
\[ \mathbf{y} : \text{measurement vector in } \mathbb{K}^m \text{ with } m \ll N, \]
\[ s : \text{sparsity of } \mathbf{x} = \text{card}\{j \in \{1, \ldots, N\} : x_j \neq 0\}. \]

Find concrete sensing/recovery protocols, i.e., find

- measurement matrices \( A : \mathbf{x} \in \mathbb{K}^N \mapsto \mathbf{y} \in \mathbb{K}^m \)
- reconstruction maps \( \Delta : \mathbf{y} \in \mathbb{K}^m \mapsto \mathbf{x} \in \mathbb{K}^N \)

such that

\[ \Delta(A\mathbf{x}) = \mathbf{x} \quad \text{for any } s\text{-sparse vector } \mathbf{x} \in \mathbb{K}^N. \]

In realistic situations, two issues to consider:

- Stability: \( \mathbf{x} \) not sparse but compressible,
- Robustness: measurement error in \( \mathbf{y} = A\mathbf{x} + \mathbf{e} \).
The Standard Compressive Sensing Problem

\[ x : \text{unknown signal of interest in } \mathbb{K}^N \]

\[ y : \text{measurement vector in } \mathbb{K}^m \text{ with } m \ll N, \]

\[ s : \text{sparsity of } x = \text{card}\{j \in \{1, \ldots, N\} : x_j \neq 0\}. \]

Find concrete sensing/recovery protocols, i.e., find

- measurement matrices \( A : x \in \mathbb{K}^N \mapsto y \in \mathbb{K}^m \)
- reconstruction maps \( \Delta : y \in \mathbb{K}^m \mapsto x \in \mathbb{K}^N \)

such that

\[ \Delta(Ax) = x \quad \text{for any } s\text{-sparse vector } x \in \mathbb{K}^N. \]

In realistic situations, two issues to consider:

- **Stability:** \( x \) not sparse but compressible,
- **Robustness:** measurement error in \( y = Ax + e \).
A Selection of Applications*

- Magnetic resonance imaging
  
  Figure: Left: traditional MRI reconstruction; Right: compressive sensing reconstruction (courtesy of M. Lustig and S. Vasanawala)

- Sampling theory

- Error correction

- and many more...
A Selection of Applications*

- Magnetic resonance imaging

Figure: Left: traditional MRI reconstruction; Right: compressive sensing reconstruction (courtesy of M. Lustig and S. Vasanawala)
A Selection of Applications*

- Magnetic resonance imaging
- Sampling theory

Figure: Time-domain signal with 16 samples.
A Selection of Applications*

- Magnetic resonance imaging
- Sampling theory
- Error correction
A Selection of Applications*

- Magnetic resonance imaging
- Sampling theory
- Error correction
- and many more...
$\ell_0$-Minimization

Since

$$\|x\|_p^p := \sum_{j=1}^{N} |x_j|^p \overset{p \to 0}{\longrightarrow} \sum_{j=1}^{N} 1\{x_j \neq 0\},$$
\textbf{\(\ell_0\)-Minimization}

Since

\[
\|x\|_p := \sum_{j=1}^{N} |x_j|^p \xrightarrow{p \to 0} \sum_{j=1}^{N} 1\{x_j \neq 0\},
\]

the notation \(\|x\|_0\) \textit{sic} has become usual for

\[
\|x\|_0 := \text{card}(\text{supp}(x)), \quad \text{where } \text{supp}(x) := \{j \in [N] : x_j \neq 0\}.
\]
\[ \ell_0 \text{-Minimization} \]

Since
\[
\|x\|_p^p := \sum_{j=1}^{N} |x_j|^p \xrightarrow{p \to 0} \sum_{j=1}^{N} 1_{\{x_j \neq 0\}},
\]
the notation \( \|x\|_0 \) [sic] has become usual for
\[
\|x\|_0 := \text{card}(\text{supp}(x)), \quad \text{where } \text{supp}(x) := \{j \in [N] : x_j \neq 0\}.
\]

For an \( s \)-sparse \( x \in \mathbb{K}^N \), observe the equivalence of
\(\ell_0\)-Minimization

Since
\[
\|x\|_p^p := \sum_{j=1}^{N} |x_j|^p \xrightarrow{p \to 0} \sum_{j=1}^{N} 1\{x_j \neq 0\},
\]
the notation \(\|x\|_0\) [sic] has become usual for
\[
\|x\|_0 := \text{card}(\text{supp}(x)), \quad \text{where } \text{supp}(x) := \{j \in [N] : x_j \neq 0\}.
\]

For an \(s\)-sparse \(x \in \mathbb{K}^N\), observe the equivalence of

- \(x\) is the unique \(s\)-sparse solution of \(Az = y\) with \(y = Ax\),
\( \ell_0 \)-Minimization

Since
\[
\| x \|_p^p := \sum_{j=1}^{N} |x_j|^p \xrightarrow{p \to 0} \sum_{j=1}^{N} 1\{x_j \neq 0\},
\]
the notation \( \| x \|_0 \) [sic] has become usual for
\[
\| x \|_0 := \text{card}(\text{supp}(x)), \quad \text{where supp}(x) := \{ j \in [N] : x_j \neq 0 \}.
\]
For an \( s \)-sparse \( x \in \mathbb{K}^N \), observe the equivalence of

- \( x \) is the unique \( s \)-sparse solution of \( Az = y \) with \( y = Ax \),
- \( x \) can be reconstructed as the unique solution of

\[
(P_0) \quad \minimize_{z \in \mathbb{K}^N} \| z \|_0 \quad \text{subject to } Az = y.
\]
ℓ₀-Minimization

Since

\[ \|x\|_p^p := \sum_{j=1}^{N} |x_j|^p \xrightarrow{p \to 0} \sum_{j=1}^{N} 1\{x_j \neq 0\}, \]

the notation \( \|x\|_0 \) [sic] has become usual for

\[ \|x\|_0 := \text{card}(\text{supp}(x)), \quad \text{where} \quad \text{supp}(x) := \{ j \in [N] : x_j \neq 0 \}. \]

For an \( s \)-sparse \( x \in \mathbb{K}^N \), observe the equivalence of

- \( x \) is the unique \( s \)-sparse solution of \( Az = y \) with \( y = Ax \),
- \( x \) can be reconstructed as the unique solution of

\[ (P_0) \quad \text{minimize} \quad \|z\|_0 \quad \text{subject to} \quad Az = y. \]

This is a combinatorial problem, NP-hard in general.
Minimal Number of Measurements

Given $A \in \mathbb{K}^{m \times N}$, the following are equivalent:

1. Every $s$-sparse $x$ is the unique $s$-sparse solution of $A z = A x$.

2. $\ker A \cap \{z \in \mathbb{K}^N : \|z\|_0 \leq 2s\} = \{0\}$.

3. For any $S \subset \{1, \ldots, N\}$ with $|S| \leq 2s$, the matrix $A_S$ is injective.

4. Every set of $2s$ columns of $A$ is linearly independent.

As a consequence, exact recovery of every $s$-sparse vector forces $m \geq 2s$.

This can be achieved using partial Vandermonde matrices.
Minimal Number of Measurements

Given $A \in \mathbb{K}^{m \times N}$, the following are equivalent:

1. Every $s$-sparse $x$ is the unique $s$-sparse solution of $A z = A x$,
2. $\ker A \cap \{z \in \mathbb{K}^N : \|z\|_0 \leq 2s\} = \{0\}$,
3. For any $S \subset [N]$ with $|S| \leq 2s$, the matrix $A_S$ is injective,
4. Every set of $2s$ columns of $A$ is linearly independent.

As a consequence, exact recovery of every $s$-sparse vector forces $m \geq 2s$.

This can be achieved using partial Vandermonde matrices.
Minimal Number of Measurements

Given $A \in \mathbb{K}^{m \times N}$, the following are equivalent:

1. Every $s$-sparse $\mathbf{x}$ is the unique $s$-sparse solution of $A \mathbf{z} = A \mathbf{x}$,
Minimal Number of Measurements

Given $A \in \mathbb{K}^{m \times N}$, the following are equivalent:

1. Every $s$-sparse $x$ is the unique $s$-sparse solution of $Az = Ax$,
2. $\ker A \cap \{z \in \mathbb{K}^N : \|z\|_0 \leq 2s\} = \{0\},$
Minimal Number of Measurements

Given $A \in \mathbb{K}^{m \times N}$, the following are equivalent:

1. Every $s$-sparse $x$ is the unique $s$-sparse solution of $Az = Ax$,
2. $\ker A \cap \{z \in \mathbb{K}^N : \|z\|_0 \leq 2s\} = \{0\}$,
3. For any $S \subset [N]$ with $\text{card}(S) \leq 2s$, the matrix $A_S$ is injective,
Minimal Number of Measurements

Given $A \in \mathbb{K}^{m \times N}$, the following are equivalent:

1. Every $s$-sparse $x$ is the unique $s$-sparse solution of $Az = Ax$,
2. $\ker A \cap \{z \in \mathbb{K}^N : \|z\|_0 \leq 2s\} = \{0\}$,
3. For any $S \subset [N]$ with $\text{card}(S) \leq 2s$, the matrix $A_S$ is injective,
4. Every set of $2s$ columns of $A$ is linearly independent.

As a consequence, exact recovery of every $s$-sparse vector forces $m \geq 2s$. This can be achieved using partial Vandermonde matrices.
Minimal Number of Measurements

Given $A \in \mathbb{K}^{m \times N}$, the following are equivalent:

1. Every $s$-sparse $x$ is the unique $s$-sparse solution of $Az = Ax$,
2. $\ker A \cap \{z \in \mathbb{K}^N : \|z\|_0 \leq 2s\} = \{0\}$,
3. For any $S \subset [N]$ with $\text{card}(S) \leq 2s$, the matrix $A_S$ is injective,
4. Every set of $2s$ columns of $A$ is linearly independent.

As a consequence, exact recovery of every $s$-sparse vector forces

$$m \geq 2s.$$
Minimal Number of Measurements

Given $A \in \mathbb{K}^{m \times N}$, the following are equivalent:

1. Every $s$-sparse $x$ is the unique $s$-sparse solution of $Az = Ax$,
2. $\ker A \cap \{z \in \mathbb{K}^N : \|z\|_0 \leq 2s\} = \{0\}$,
3. For any $S \subset [N]$ with $\text{card}(S) \leq 2s$, the matrix $A_S$ is injective,
4. Every set of $2s$ columns of $A$ is linearly independent.

As a consequence, exact recovery of every $s$-sparse vector forces

$$m \geq 2s.$$

This can be achieved using partial Vandermonde matrices.
Exact s-Sparse Recovery from 2s Fourier Measurements*

Identify an s-sparse \( x \in \mathbb{C}^N \) with a function \( x \) on \( \{0, 1, \ldots, N-1\} \) with support \( S \), \( \text{card}(S) = s \).

Consider the 2s Fourier coefficients \( \hat{x}(j) = \frac{1}{N-1} \sum_{k=0}^{N-1} x(k) e^{-i2\pi jk/N}, 0 \leq j \leq 2s-1 \).

Consider a trigonometric polynomial vanishing exactly on \( S \), i.e., \( p(t) := \prod_{k \in S} (1 - e^{-i2\pi k/N} e^{i2\pi t/N}) \).

Since \( p \cdot x \equiv 0 \), discrete convolution gives
\[
0 = (\hat{p} \ast \hat{x})(j) = \frac{1}{N-1} \sum_{k=0}^{N-1} \hat{p}(k) \hat{x}(j-k), 0 \leq j \leq N-1.
\]

Note that \( \hat{p}(0) = 1 \) and that \( \hat{p}(k) = 0 \) for \( k > s \).

The equations \( s, \ldots, 2s-1 \) translate into a Toeplitz system with unknowns \( \hat{p}(1), \ldots, \hat{p}(s) \).

This determines \( \hat{p} \), hence \( p \), then \( S \), and finally \( x \).
Exact $s$-Sparse Recovery from $2s$ Fourier Measurements*

Identify an $s$-sparse $\mathbf{x} \in \mathbb{C}^N$ with a function $x$ on $\{0, 1, \ldots, N - 1\}$ with support $S$, $\text{card}(S) = s$. 
Exact $s$-Sparse Recovery from $2s$ Fourier Measurements*

Identify an $s$-sparse $\mathbf{x} \in \mathbb{C}^N$ with a function $x$ on $\{0, 1, \ldots, N - 1\}$ with support $S$, $\text{card}(S) = s$. Consider the $2s$ Fourier coefficients

$$\hat{x}(j) = \sum_{k=0}^{N-1} x(k)e^{-i2\pi jk/N}, \quad 0 \leq j \leq 2s - 1.$$
Exact $s$-Sparse Recovery from $2s$ Fourier Measurements*

Identify an $s$-sparse $\mathbf{x} \in \mathbb{C}^N$ with a function $x$ on $\{0, 1, \ldots, N - 1\}$ with support $S$, $\text{card}(S) = s$. Consider the $2s$ Fourier coefficients

$$\hat{x}(j) = \sum_{k=0}^{N-1} x(k) e^{-i2\pi j k / N}, \quad 0 \leq j \leq 2s - 1.$$ 

Consider a trigonometric polynomial vanishing exactly on $S$, i.e.,

$$p(t) := \prod_{k \in S} \left( 1 - e^{-i2\pi k / N} e^{i2\pi t / N} \right).$$
Exact $s$-Sparse Recovery from $2s$ Fourier Measurements*

Identify an $s$-sparse $x \in \mathbb{C}^N$ with a function $x$ on $\{0, 1, \ldots, N - 1\}$ with support $S$, $\text{card}(S) = s$. Consider the $2s$ Fourier coefficients

$$\hat{x}(j) = \sum_{k=0}^{N-1} x(k) e^{-i2\pi j k / N}, \quad 0 \leq j \leq 2s - 1.$$ 

Consider a trigonometric polynomial vanishing exactly on $S$, i.e.,

$$p(t) := \prod_{k \in S} \left(1 - e^{-i2\pi k / N} e^{i2\pi t / N}\right).$$

Since $p \cdot x \equiv 0$, discrete convolution gives

$$0 = (\hat{p} \ast \hat{x})(j) = \sum_{k=0}^{N-1} \hat{p}(k) \hat{x}(j - k), \quad 0 \leq j \leq N - 1.$$
Exact $s$-Sparse Recovery from $2s$ Fourier Measurements*

Identify an $s$-sparse $x \in \mathbb{C}^N$ with a function $x$ on $\{0, 1, \ldots, N - 1\}$ with support $S$, $\text{card}(S) = s$. Consider the $2s$ Fourier coefficients

$$\hat{x}(j) = \sum_{k=0}^{N-1} x(k) e^{-i2\pi jk/N}, \quad 0 \leq j \leq 2s - 1.$$ 

Consider a trigonometric polynomial vanishing exactly on $S$, i.e.,

$$p(t) := \prod_{k \in S} \left(1 - e^{-i2\pi k/N} e^{i2\pi t/N}\right).$$

Since $p \cdot x \equiv 0$, discrete convolution gives

$$0 = (\hat{p} \ast \hat{x})(j) = \sum_{k=0}^{N-1} \hat{p}(k) \hat{x}(j - k), \quad 0 \leq j \leq N - 1.$$ 

Note that $\hat{p}(0) = 1$ and that $\hat{p}(k) = 0$ for $k > s$. 

*Traditionally, this section is marked with an asterisk.
Exact $s$-Sparse Recovery from $2s$ Fourier Measurements*

Identify an $s$-sparse $x \in \mathbb{C}^N$ with a function $x$ on $\{0, 1, \ldots, N - 1\}$ with support $S$, $\text{card}(S) = s$. Consider the $2s$ Fourier coefficients

$$\hat{x}(j) = \sum_{k=0}^{N-1} x(k) e^{-i2\pi j k / N}, \quad 0 \leq j \leq 2s - 1.$$ 

Consider a trigonometric polynomial vanishing exactly on $S$, i.e.,

$$p(t) := \prod_{k \in S} (1 - e^{-i2\pi k / N} e^{i2\pi t / N}).$$

Since $p \cdot x \equiv 0$, discrete convolution gives

$$0 = (\hat{p} * \hat{x})(j) = \sum_{k=0}^{N-1} \hat{p}(k) \hat{x}(j - k), \quad 0 \leq j \leq N - 1.$$ 

Note that $\hat{p}(0) = 1$ and that $\hat{p}(k) = 0$ for $k > s$. The equations $s, \ldots, 2s - 1$ translate into a Toeplitz system with unknowns $\hat{p}(1), \ldots, \hat{p}(s)$. 

*This research was supported by the National Science Foundation under Grant Number 0526010.
Exact $s$-Sparse Recovery from $2s$ Fourier Measurements*

Identify an $s$-sparse $x \in \mathbb{C}^N$ with a function $x$ on $\{0, 1, \ldots, N - 1\}$ with support $S$, $\text{card}(S) = s$. Consider the $2s$ Fourier coefficients

$$\hat{x}(j) = \sum_{k=0}^{N-1} x(k) e^{-i2\pi jk/N}, \quad 0 \leq j \leq 2s - 1.$$ 

Consider a trigonometric polynomial vanishing exactly on $S$, i.e.,

$$p(t) := \prod_{k \in S} (1 - e^{-i2\pi k/N} e^{i2\pi t/N}).$$

Since $p \cdot x \equiv 0$, discrete convolution gives

$$0 = (\hat{p} \ast \hat{x})(j) = \sum_{k=0}^{N-1} \hat{p}(k)\hat{x}(j - k), \quad 0 \leq j \leq N - 1.$$ 

Note that $\hat{p}(0) = 1$ and that $\hat{p}(k) = 0$ for $k > s$. The equations $s, \ldots, 2s - 1$ translate into a Toeplitz system with unknowns $\hat{p}(1), \ldots, \hat{p}(s)$. This determines $\hat{p}$, hence $p$, then $S$, and finally $x$. 

*Refers to the asterisk in the title.
Optimization and Greedy Strategies
$\ell_1$-Minimization (Basis Pursuit)
$\ell_1$-Minimization (Basis Pursuit)

Replace $(P_0)$ by

$\mathbf{(P_1)}$ \quad \text{minimize} \quad \|z\|_1 \quad \text{subject to} \quad Az = y.$
\( \ell_1 \)-Minimization (Basis Pursuit)

Replace \((P_0)\) by

\[
(P_1) \quad \text{minimize} \quad \|z\|_1 \quad \text{subject to} \quad Az = y.
\]

- Geometric intuition
Replace \((P_0)\) by \((P_1)\)

\[(P_1) \quad \text{minimize } \|z\|_1 \quad \text{subject to } Az = y.\]

- Geometric intuition
- Unique \(\ell_1\)-minimizers are at most \(m\)-sparse (when \(K = \mathbb{R}\))
ℓ₁-Minimization (Basis Pursuit)

Replace (P₀) by

\[(P₁) \quad \text{minimize } \|z\|₁ \text{ subject to } Az = y.\]

- Geometric intuition
- Unique ℓ₁-minimizers are at most \( m \)-sparse (when \( K = \mathbb{R} \))
- Convex optimization program, hence solvable in practice
\(\ell_1\)-Minimization (Basis Pursuit)

Replace \((P_0)\) by

\[(P_1) \quad \text{minimize} \quad \| z \|_1 \quad \text{subject to} \quad Az = y.\]

- Geometric intuition
- Unique \(\ell_1\)-minimizers are at most \(m\)-sparse (when \(K = \mathbb{R}\))
- Convex optimization program, hence solvable in practice
- In the real setting, recast as the linear optimization program

\[
\text{minimize} \quad \sum_{j=1}^{N} c_j \quad \text{subject to} \quad Az = y \quad \text{and} \quad -c_j \leq z_j \leq c_j.
\]
$\ell_1$-Minimization (Basis Pursuit)

Replace $(P_0)$ by

$$(P_1) \quad \text{minimize } \|z\|_1 \quad \text{subject to } Az = y.$$ 

- Geometric intuition
- Unique $\ell_1$-minimizers are at most $m$-sparse (when $K = \mathbb{R}$)
- Convex optimization program, hence solvable in practice
- In the real setting, recast as the linear optimization program

$$\text{minimize } \sum_{j=1}^{N} c_j \quad \text{subject to } Az = y \text{ and } -c_j \leq z_j \leq c_j.$$ 

- In the complex setting, recast as a second order cone program
Basis Pursuit — Null Space Property

\[ \Delta_1 (Ax) = x \text{ for every vector } x \text{ supported on } S \text{ if and only if } \parallel u_S \parallel_1 < \parallel u_S \parallel_1, \text{ all } u \in \ker \{0\}. \]

For real measurement matrices, real and complex NSPs read

\[ \sum_{j \in S} |u_j| < \sum_{\ell \in S} |u_\ell|, \text{ all } u \in \ker \{0\}, \]

\[ \sum_{j \in S} \sqrt{v_j^2 + w_j^2} < \sum_{\ell \in S} \sqrt{v_\ell^2 + w_\ell^2}, \text{ all } (v, w) \in (\ker \{0\})^2. \]

Real and complex NSPs are in fact equivalent.
Basis Pursuit — Null Space Property

\[ \Delta_1(Ax) = x \text{ for every vector } x \text{ supported on } S \text{ if and only if} \]

Real and complex NSPs are in fact equivalent.
Basis Pursuit — Null Space Property

\[ \Delta_1(Ax) = x \text{ for every vector } x \text{ supported on } S \text{ if and only if} \]

(NSP) \[ \|u_S\|_1 < \|u_{\bar{S}}\|_1, \quad \text{all } u \in \ker A \setminus \{0\}. \]
Basis Pursuit — Null Space Property

\[ \Delta_1(Ax) = x \text{ for every vector } x \text{ supported on } S \text{ if and only if} \]

\[(\text{NSP}) \quad \|u_S\|_1 < \|u_{\bar{S}}\|_1, \quad \text{all } u \in \ker A \setminus \{0\} .\]

For real measurement matrices,
Basis Pursuit — Null Space Property

\[ \Delta_1(Ax) = x \] for every vector \( x \) supported on \( S \) if and only if

\[ (\text{NSP}) \quad \|u_S\|_1 < \|u_{\overline{S}}\|_1, \quad \text{all } u \in \ker A \setminus \{0\}. \]

For real measurement matrices, real and complex NSPs read
Basis Pursuit — Null Space Property

\[ \Delta_1(Ax) = x \text{ for every vector } x \text{ supported on } S \text{ if and only if} \]

\[ (\text{NSP}) \quad \|u_S\|_1 < \|u_{\overline{S}}\|_1, \quad \text{all } u \in \ker A \setminus \{0\}. \]

For real measurement matrices, real and complex NSPs read

\[ \sum_{j \in S} |u_j| < \sum_{\ell \in \overline{S}} |u_\ell|, \quad \text{all } u \in \ker_{\mathbb{R}} A \setminus \{0\}, \]
Basis Pursuit — Null Space Property

\[ \Delta_1(Ax) = x \text{ for every vector } x \text{ supported on } S \text{ if and only if} \]

\[ (\text{NSP}) \quad \|u_S\|_1 < \|u_{\overline{S}}\|_1, \quad \text{all } u \in \ker A \setminus \{0\}. \]

For real measurement matrices, real and complex NSPs read

\[ \sum_{j \in S} |u_j| < \sum_{\ell \in \overline{S}} |u_{\ell}|, \quad \text{all } u \in \ker \mathbb{R} A \setminus \{0\}, \]

\[ \sum_{j \in S} \sqrt{v_j^2 + w_j^2} < \sum_{\ell \in \overline{S}} \sqrt{v_{\ell}^2 + w_{\ell}^2}, \quad \text{all } (v, w) \in (\ker \mathbb{R} A)^2 \setminus \{0\}. \]
Basis Pursuit — Null Space Property

\[ \Delta_1(Ax) = x \text{ for every vector } x \text{ supported on } S \text{ if and only if} \]

\[(\text{NSP}) \quad \|u_S\|_1 < \|u_{\overline{S}}\|_1, \quad \text{all } u \in \ker A \setminus \{0\}.\]

For real measurement matrices, real and complex NSPs read

\[
\sum_{j \in S} |u_j| < \sum_{\ell \in \overline{S}} |u_\ell|, \quad \text{all } u \in \ker_{\mathbb{R}} A \setminus \{0\},
\]

\[
\sum_{j \in S} \sqrt{v_j^2 + w_j^2} < \sum_{\ell \in \overline{S}} \sqrt{v_\ell^2 + w_\ell^2}, \quad \text{all } (v, w) \in (\ker_{\mathbb{R}} A)^2 \setminus \{0\}.
\]

Real and complex NSPs are in fact equivalent.
Orthogonal Matching Pursuit

Starting with $S_0 = \emptyset$ and $x_0 = 0$, iterate

$$S_{n+1} = S_n \cup \{j_{n+1} = \arg\max_{j \in [N]} \{|(A^* (y - A x_n))_j|\}$$

(OMP 1)

$$x_{n+1} = \arg\min_{z \in \mathbb{C}^N} \{\|y - A z\|_2, \text{supp}(z) \subseteq S_{n+1}\}.$$  

(OMP 2)

▶ The norm of the residual decreases according to

$$\|y - A x_{n+1}\|_2^2 \leq \|y - A x_n\|_2^2 - \|(A^* (y - A x_n))_{j_{n+1}}\|_2^2.$$  

▶ Every vector $x \neq 0$ supported on $S$, $\text{card}(S) = s$, is recovered from $y = A x$ after at most $s$ iterations of OMP if and only if $A_S$ is injective and

$$\max_{j \in S} |(A^* r)_j| > \max_{\ell \in S} |(A^* r)_\ell|$$

for all $r \neq 0 \in \{A z, \text{supp}(z) \subseteq S\}$. 


Orthogonal Matching Pursuit

Starting with \( S^0 = \emptyset \) and \( x^0 = 0 \), iterate

\[
\text{OMP}_1 \quad S^{n+1} = S^n \cup \{ j^{n+1} := \arg\max_{j \in [N]} \left| (A^* (y - Ax^n))_j \right| \},
\]

\[
\text{OMP}_2 \quad x^{n+1} = \arg\min_{z \in \mathbb{C}^N} \{ \| y - Az \|_2, \supp(z) \subseteq S^{n+1} \}.
\]
Orthogonal Matching Pursuit

Starting with $S^0 = \emptyset$ and $x^0 = 0$, iterate

\[(OMP_1) \quad S^{n+1} = S^n \cup \{ j^{n+1} := \text{argmax}_{j \in [N]} \{ |(A^*(y - Ax^n))_j| \} \}, \]

\[(OMP_2) \quad x^{n+1} = \text{argmin}_{z \in \mathbb{C}^N} \{ \|y - Az\|_2, \text{supp}(z) \subseteq S^{n+1} \}. \]

- The norm of the residual decreases according to

\[\|y - Ax^{n+1}\|_2^2 \leq \|y - Ax^n\|_2^2 - |(A^*(y - Ax^n))_{j^{n+1}}|^2.\]
Orthogonal Matching Pursuit

Starting with \( S^0 = \emptyset \) and \( x^0 = 0 \), iterate

\[
(\text{OMP}_1) \quad S^{n+1} = S^n \cup \{ j^{n+1} := \arg\max_{j \in [N]} \{(A^*(y - Ax^n))_j\} \},
\]

\[
(\text{OMP}_2) \quad x^{n+1} = \arg\min_{z \in \mathbb{C}^N} \{ \| y - Az \|_2, \text{supp}(z) \subseteq S^{n+1} \}.
\]

- The norm of the residual decreases according to

\[
\| y - Ax^{n+1} \|_2^2 \leq \| y - Ax^n \|_2^2 - \left| (A^*(y - Ax^n))_{j^{n+1}} \right|^2.
\]

- Every vector \( x \neq 0 \) supported on \( S \), \( \text{card}(S) = s \), is recovered from \( y = Ax \) after at most \( s \) iterations of OMP if and only if \( A_S \) is injective and

\[
(\text{ERC}) \quad \max_{j \in S} |(A^*r)_j| > \max_{\ell \in \overline{S}} |(A^*r)_\ell|
\]

for all \( r \neq 0 \in \{ Az, \text{supp}(z) \subseteq S \} \).
First Recovery Guarantees
Coherence

For a matrix with $\ell_2$-normalized columns $a_1, \ldots, a_N$, define

$$\mu := \max_{i \neq j} |\langle a_i, a_j \rangle|.$$

As a rule, the smaller the coherence, the better.

However, the Welch bound reads

$$\mu \geq \sqrt{N - m} \left( \frac{N-1}{m} \right).$$

▶ Welch bound achieved at and only at equiangular tight frames

▶ Deterministic matrices with coherence $\mu \leq c/\sqrt{m}$ exist
For a matrix with $\ell_2$-normalized columns $a_1, \ldots, a_N$, define

$$\mu := \max_{i \neq j} |\langle a_i, a_j \rangle|.$$
Coherence

For a matrix with $\ell_2$-normalized columns $\mathbf{a}_1, \ldots, \mathbf{a}_N$, define

$$\mu := \max_{i \neq j} |\langle \mathbf{a}_i, \mathbf{a}_j \rangle| .$$

As a rule, the smaller the coherence, the better.
Coherence

For a matrix with $\ell_2$-normalized columns $\mathbf{a}_1, \ldots, \mathbf{a}_N$, define

$$
\mu := \max_{i \neq j} |\langle \mathbf{a}_i, \mathbf{a}_j \rangle|.
$$

As a rule, the smaller the coherence, the better.

However, the Welch bound reads

$$
\mu \geq \sqrt{\frac{N - m}{m(N - 1)}}.
$$
Coherence

For a matrix with $\ell_2$-normalized columns $a_1, \ldots, a_N$, define

$$\mu := \max_{i \neq j} |\langle a_i, a_j \rangle|.$$

As a rule, the smaller the coherence, the better.

However, the Welch bound reads

$$\mu \geq \sqrt{\frac{N - m}{m(N - 1)}}.$$

- Welch bound achieved at and only at equiangular tight frames
Coherence

For a matrix with $\ell_2$-normalized columns $a_1, \ldots, a_N$, define

$$\mu := \max_{i \neq j} |\langle a_i, a_j \rangle|.$$ 

As a rule, the smaller the coherence, the better.

However, the Welch bound reads

$$\mu \geq \sqrt{\frac{N - m}{m(N - 1)}}.$$ 

- Welch bound achieved at and only at equiangular tight frames
- Deterministic matrices with coherence $\mu \leq c/\sqrt{m}$ exist
Recovery Conditions using Coherence

▶ Every $s$-sparse $x \in \mathbb{C}^N$ is recovered from $y = Ax \in \mathbb{C}^m$ via at most $s$ iterations of OMP provided $\mu < \frac{1}{2s - 1}$.

▶ Every $s$-sparse $x \in \mathbb{C}^N$ is recovered from $y = Ax \in \mathbb{C}^m$ via Basis Pursuit provided $\mu < \frac{1}{2s - 1}$.

▶ In fact, the Exact Recovery Condition can be rephrased as $\|A^\dagger S A S\|_1 \rightarrow 1 < 1$, and this implies the Null Space Property.
Recovery Conditions using Coherence

- Every $s$-sparse $\mathbf{x} \in \mathbb{C}^N$ is recovered from $\mathbf{y} = A\mathbf{x} \in \mathbb{C}^m$ via at most $s$ iterations of OMP provided
  \[ \mu < \frac{1}{2s - 1}. \]
Recovery Conditions using Coherence

- Every \( s \)-sparse \( \mathbf{x} \in \mathbb{C}^N \) is recovered from \( \mathbf{y} = A\mathbf{x} \in \mathbb{C}^m \) via at most \( s \) iterations of OMP provided

\[
\mu < \frac{1}{2s - 1}.
\]

- Every \( s \)-sparse \( \mathbf{x} \in \mathbb{C}^N \) is recovered from \( \mathbf{y} = A\mathbf{x} \in \mathbb{C}^m \) via Basis Pursuit provided

\[
\mu < \frac{1}{2s - 1}.
\]
Recovery Conditions using Coherence

- Every $s$-sparse $x \in \mathbb{C}^N$ is recovered from $y = Ax \in \mathbb{C}^m$ via at most $s$ iterations of OMP provided
  \[
  \mu < \frac{1}{2s - 1}.
  \]

- Every $s$-sparse $x \in \mathbb{C}^N$ is recovered from $y = Ax \in \mathbb{C}^m$ via Basis Pursuit provided
  \[
  \mu < \frac{1}{2s - 1}.
  \]

- In fact, the Exact Recovery Condition can be rephrased as
  \[
  \|A_S^\dagger A_S\|_{1\rightarrow 1} < 1,
  \]
  and this implies the Null Space Property.
The coherence conditions for $s$-sparse recovery are of the type $\mu \leq c_s$, while the Welch bound (for $N \geq 2^m$, say) reads $\mu \geq c_s \sqrt{m}$. Therefore, arguments based on coherence necessitate $m \geq c_s^2$. This is far from the ideal linear scaling of $m$ in $s$...

Next, we will introduce new tools to break this quadratic barrier (with random matrices).
Summary of Part I

The coherence conditions for $s$-sparse recovery are of the type

$$\mu \leq \frac{c}{s},$$

while the Welch bound (for $N \geq 2^m$, say) reads

$$\mu \geq c \sqrt{m}.$$
Summary of Part I

The coherence conditions for $s$-sparse recovery are of the type

$$\mu \leq \frac{c}{s},$$

while the Welch bound (for $N \geq 2m$, say) reads

$$\mu \geq \frac{c}{\sqrt{m}}.$$
The coherence conditions for $s$-sparse recovery are of the type

$$\mu \leq \frac{c}{s},$$

while the Welch bound (for $N \geq 2m$, say) reads

$$\mu \geq \frac{c}{\sqrt{m}}.$$  

Therefore, arguments based on coherence necessitate

$$m \geq c s^2.$$
The coherence conditions for $s$-sparse recovery are of the type

$$\mu \leq \frac{c}{s},$$

while the Welch bound (for $N \geq 2m$, say) reads

$$\mu \geq \frac{c}{\sqrt{m}}.$$

Therefore, arguments based on coherence necessitate

$$m \geq c s^2.$$

This is far from the ideal linear scaling of $m$ in $s$. ...
Summary of Part I

The coherence conditions for $s$-sparse recovery are of the type

$$\mu \leq \frac{c}{s},$$

while the Welch bound (for $N \geq 2m$, say) reads

$$\mu \geq \frac{c}{\sqrt{m}}.$$  

Therefore, arguments based on coherence necessitate

$$m \geq c s^2.$$  

This is far from the ideal linear scaling of $m$ in $s$...

Next, we will introduce new tools to break this *quadratic barrier* (with random matrices).
Part II:

The Restricted Isometry Property
Restricted Isometry Constants

\[\delta_s = \text{the smallest } \delta > 0 \text{ such that } (1 - \delta) \|z\|_2^2 \leq \|Az\|_2^2 \leq (1 + \delta) \|z\|_2^2 \text{ for all } s\text{-sparse } z \in \mathbb{C}^N.\]

Alternative form:

\[\delta_s = \max_{\text{card}(S) \leq s} \|A^* S A S - \text{Id}\|_2 \rightarrow 2.\]

Suitable CS matrices have small restricted isometry constants.

Typically, \(\delta_s \leq \delta^*\) holds for a number of random measurements \(m \approx c(\delta^*)c\delta^2s \ln(eN/s)\).

In fact, \(\delta_s \leq \delta^*\) imposes \(m \geq c\delta^2s\).
Restricted Isometry Constants

Restricted Isometry Constant $\delta_s = \text{the smallest } \delta > 0 \text{ such that }$

$$(1 - \delta)\|z\|_2^2 \leq \|Az\|_2^2 \leq (1 + \delta)\|z\|_2^2 \text{ for all } s\text{-sparse } z \in \mathbb{C}^N.$$
Restricted Isometry Constants

Restricted Isometry Constant $\delta_s = \text{the smallest } \delta > 0 \text{ such that}$

$$(1 - \delta)\|z\|_2^2 \leq \|Az\|_2^2 \leq (1 + \delta)\|z\|_2^2 \text{ for all } s\text{-sparse } z \in \mathbb{C}^N.$$ 

Alternative form:

$$\delta_s = \max_{\text{card}(S) \leq s} \|A^*_S A_S - \text{Id}\|_{2 \to 2}.$$
Restricted Isometry Constants

**Restricted Isometry Constant** $\delta_s = \text{the smallest } \delta > 0 \text{ such that}

$$(1 - \delta)\|z\|_2^2 \leq \|Az\|_2^2 \leq (1 + \delta)\|z\|_2^2$$

for all $s$-sparse $z \in \mathbb{C}^N$.

Alternative form:

$$\delta_s = \max_{\text{card}(S) \leq s} \|A^*_S A_S - \text{Id}\|_{2 \to 2}.$$

Suitable CS matrices have small restricted isometry constants.
Restricted Isometry Constants

Restricted Isometry Constant $\delta_s = $ the smallest $\delta > 0$ such that

$$(1 - \delta) \|z\|_2^2 \leq \|Az\|_2^2 \leq (1 + \delta) \|z\|_2^2$$

for all $s$-sparse $z \in \mathbb{C}^N$.

Alternative form:

$$\delta_s = \max_{\text{card}(S) \leq s} \|A^*_S A_S - \text{Id}\|_{2 \to 2}.$$ 

Suitable CS matrices have small restricted isometry constants.

Typically, $\delta_s \leq \delta_*$ holds for a number of random measurements

$$m \approx c(\delta_*) s \ln(eN/s).$$
Restricted Isometry Constants

Restricted Isometry Constant $\delta_s = \text{the smallest $\delta > 0$ such that}$

$$(1 - \delta)\|z\|_2^2 \leq \|A z\|_2^2 \leq (1 + \delta)\|z\|_2^2 \quad \text{for all $s$-sparse $z \in \mathbb{C}^N$.}$$

Alternative form:

$$\delta_s = \max_{\text{card}(S) \leq s} \|A^*_S A_S - \text{Id}\|_{2 \rightarrow 2}.$$

Suitable CS matrices have small restricted isometry constants. Typically, $\delta_s \leq \delta_*$ holds for a number of random measurements

$$m \approx \frac{c}{\delta_*^2} s \ln(eN/s).$$
Restricted Isometry Constants

**Restricted Isometry Constant** $\delta_s = \text{the smallest } \delta > 0 \text{ such that }$

$$(1 - \delta)\|z\|_2^2 \leq \|Az\|_2^2 \leq (1 + \delta)\|z\|_2^2 \text{ for all } s\text{-sparse } z \in \mathbb{C}^N.$$

**Alternative form:**

$$\delta_s = \max_{\text{card}(S) \leq s} \|A_S^* A_S - \text{Id}\|_{2 \rightarrow 2}.$$

Suitable CS matrices have small restricted isometry constants. Typically, $\delta_s \leq \delta_* \text{ holds for a number of random measurements}$

$$m \approx \frac{c}{\delta_*^2} s \ln(eN/s).$$

In fact, $\delta_s \leq \delta_*$ imposes $m \geq \frac{c}{\delta_*^2} s.$
RIP-based Recovery Guarantees
Exact Sparse Recovery via $\ell_1$-Minimization when $\delta_{2s} < 1/3$
Exact Sparse Recovery via $\ell_1$-Minimization when $\delta_{2s} < 1/3$

Take $v \in \ker A \setminus \{0\}$.
Exact Sparse Recovery via $\ell_1$-Minimization when $\delta_{2s} < 1/3$

Take $v \in \ker A \setminus \{0\}$. Consider sets $S_0, S_1, \ldots$ of $s$ indices ordered by decreasing magnitude of entries of $v$. 
Exact Sparse Recovery via $\ell_1$-Minimization when $\delta_{2s} < 1/3$

Take $v \in \ker A \setminus \{0\}$. Consider sets $S_0, S_1, \ldots$ of $s$ indices ordered by decreasing magnitude of entries of $v$.

$$\|v_{S_0}\|_2^2 \leq$$
Take $v \in \ker A \setminus \{0\}$. Consider sets $S_0, S_1, \ldots$ of $s$ indices ordered by decreasing magnitude of entries of $v$.

$$\|v_{S_0}\|_2^2 \leq \frac{1}{1 - \delta_{2s}} \|A(v_{S_0})\|_2^2$$
Exact Sparse Recovery via $\ell_1$-Minimization when $\delta_{2s} < 1/3$

Take $v \in \ker A \setminus \{0\}$. Consider sets $S_0, S_1, \ldots$ of $s$ indices ordered by decreasing magnitude of entries of $v$.

$$\|v_{S_0}\|_2^2 \leq \frac{1}{1 - \delta_{2s}} \|A(v_{S_0})\|_2^2 = \frac{1}{1 - \delta_{2s}} \sum_{k \geq 1} \langle A(v_{S_0}), A(-v_{S_k}) \rangle$$
Exact Sparse Recovery via $\ell_1$-Minimization when $\delta_{2s} < 1/3$

Take $v \in \ker A \setminus \{0\}$. Consider sets $S_0, S_1, \ldots$ of $s$ indices ordered by decreasing magnitude of entries of $v$.

$$
\|v_{S_0}\|_2^2 \leq \frac{1}{1 - \delta_{2s}} \|A(v_{S_0})\|_2^2 = \frac{1}{1 - \delta_{2s}} \sum_{k \geq 1} \langle A(v_{S_0}), A(-v_{S_k}) \rangle
$$

$$
\leq \frac{1}{1 - \delta_{2s}} \sum_{k \geq 1} \delta_{2s} \|v_{S_0}\|_2 \|v_{S_k}\|_2.
$$
Exact Sparse Recovery via \( \ell_1 \)-Minimization when \( \delta_{2s} < 1/3 \)

Take \( \mathbf{v} \in \ker A \setminus \{0\} \). Consider sets \( S_0, S_1, \ldots \) of \( s \) indices ordered by decreasing magnitude of entries of \( \mathbf{v} \).

\[
\| \mathbf{v}_{S_0} \|_2^2 \leq \frac{1}{1 - \delta_{2s}} \| A(\mathbf{v}_{S_0}) \|_2^2 = \frac{1}{1 - \delta_{2s}} \sum_{k \geq 1} \langle A(\mathbf{v}_{S_0}), A(-\mathbf{v}_{S_k}) \rangle
\]

\[
\leq \frac{1}{1 - \delta_{2s}} \sum_{k \geq 1} \delta_{2s} \| \mathbf{v}_{S_0} \|_2 \| \mathbf{v}_{S_k} \|_2.
\]

Simplify by \( \| \mathbf{v}_{S_0} \|_2 \)
Exact Sparse Recovery via $\ell_1$-Minimization when $\delta_{2s} < 1/3$

Take $v \in \ker A \setminus \{0\}$. Consider sets $S_0, S_1, \ldots$ of $s$ indices ordered by decreasing magnitude of entries of $v$.

$$\|v_{S_0}\|_2^2 \leq \frac{1}{1 - \delta_{2s}} \|A(v_{S_0})\|_2^2 = \frac{1}{1 - \delta_{2s}} \sum_{k \geq 1} \langle A(v_{S_0}), A(-v_{S_k}) \rangle$$

$$\leq \frac{1}{1 - \delta_{2s}} \sum_{k \geq 1} \delta_{2s} \|v_{S_0}\|_2 \|v_{S_k}\|_2.$$

Simplify by $\|v_{S_0}\|_2$ and observe that

$$\|v_{S_k}\|_2 \leq \frac{1}{\sqrt{s}} \|v_{S_{k-1}}\|_1.$$
Exact Sparse Recovery via $\ell_1$-Minimization when $\delta_{2s} < 1/3$

Take $v \in \ker A \setminus \{0\}$. Consider sets $S_0, S_1, \ldots$ of $s$ indices ordered by decreasing magnitude of entries of $v$.

$$\|v_{S_0}\|_2^2 \leq \frac{1}{1 - \delta_{2s}} \|A(v_{S_0})\|_2^2 = \frac{1}{1 - \delta_{2s}} \sum_{k \geq 1} \langle A(v_{S_0}), A(-v_{S_k}) \rangle$$

$$\leq \frac{1}{1 - \delta_{2s}} \sum_{k \geq 1} \delta_{2s} \|v_{S_0}\|_2 \|v_{S_k}\|_2.$$

Simplify by $\|v_{S_0}\|_2$ and observe that

$$\|v_{S_k}\|_2 \leq \frac{1}{\sqrt{s}} \|v_{S_{k-1}}\|_1$$

to obtain

$$\|v_{S_0}\|_2 \leq \frac{\delta_{2s}}{1 - \delta_{2s}} \frac{1}{\sqrt{s}} \|v\|_1,$$
Exact Sparse Recovery via $\ell_1$-Minimization when $\delta_{2s} < 1/3$

Take $\mathbf{v} \in \ker A \setminus \{\mathbf{0}\}$. Consider sets $S_0, S_1, \ldots$ of $s$ indices ordered by decreasing magnitude of entries of $\mathbf{v}$.

$$\|\mathbf{v}_{S_0}\|_2^2 \leq \frac{1}{1 - \delta_{2s}} \|A(\mathbf{v}_{S_0})\|_2^2 = \frac{1}{1 - \delta_{2s}} \sum_{k \geq 1} \langle A(\mathbf{v}_{S_0}), A(-\mathbf{v}_{S_k}) \rangle$$

$$\leq \frac{1}{1 - \delta_{2s}} \sum_{k \geq 1} \delta_{2s} \|\mathbf{v}_{S_0}\|_2 \|\mathbf{v}_{S_k}\|_2.$$

Simplify by $\|\mathbf{v}_{S_0}\|_2$ and observe that

$$\|\mathbf{v}_{S_k}\|_2 \leq \frac{1}{\sqrt{s}} \|\mathbf{v}_{S_{k-1}}\|_1$$

to obtain

$$\|\mathbf{v}_{S_0}\|_2 \leq \frac{\delta_{2s}}{1 - \delta_{2s}} \frac{1}{\sqrt{s}} \|\mathbf{v}\|_1,$$  hence $\|\mathbf{v}_{S_0}\|_1 \leq \frac{\delta_{2s}}{1 - \delta_{2s}} \|\mathbf{v}\|_1.$
Exact Sparse Recovery via $\ell_1$-Minimization when $\delta_{2s} < 1/3$

Take $\mathbf{v} \in \ker A \setminus \{0\}$. Consider sets $S_0, S_1, \ldots$ of $s$ indices ordered by decreasing magnitude of entries of $\mathbf{v}$.

$$
\|\mathbf{v}_{S_0}\|_2^2 \leq \frac{1}{1 - \delta_{2s}} \|A(\mathbf{v}_{S_0})\|_2^2 = \frac{1}{1 - \delta_{2s}} \sum_{k \geq 1} \langle A(\mathbf{v}_{S_0}), A(-\mathbf{v}_{S_k}) \rangle 
$$

$$
\leq \frac{1}{1 - \delta_{2s}} \sum_{k \geq 1} \delta_{2s} \|\mathbf{v}_{S_0}\|_2 \|\mathbf{v}_{S_k}\|_2.
$$

Simplify by $\|\mathbf{v}_{S_0}\|_2$ and observe that

$$
\|\mathbf{v}_{S_k}\|_2 \leq \frac{1}{\sqrt{s}} \|\mathbf{v}_{S_{k-1}}\|_1
$$

to obtain

$$
\|\mathbf{v}_{S_0}\|_2 \leq \frac{\delta_{2s}}{1 - \delta_{2s}} \frac{1}{\sqrt{s}} \|\mathbf{v}\|_1, \quad \text{hence} \quad \|\mathbf{v}_{S_0}\|_1 \leq \frac{\delta_{2s}}{1 - \delta_{2s}} \|\mathbf{v}\|_1.
$$

Note that $\rho := \delta_{2s}/(1 - \delta_{2s}) < 1/2$ whenever $\delta_{2s} < 1/3$. 
Stable and Robust Sparse Recovery via $\ell_1$-Minimization*

Objective: for $p \in [1, 2]$, for all $x \in \mathbb{C}^N$ and $e \in \mathbb{C}^m$ with $\|e\|_2 \leq \eta$:

$$\|x - \Delta(Ax + e)\|_p \leq \text{stability} \cdot \frac{1}{1/p - 1/2} \|x_{\text{sparse}}\|_1 + \text{robustness} \cdot \frac{1}{1/p - 1/2} \eta,$$

where $\Delta(y) = \Delta_1(\eta)(y) := \arg\min \|z\|_1$ subject to $\|Az - y\|_2 \leq \eta$.

Taking $x = v \in \mathbb{C}^N$, $e = -Av \in \mathbb{C}^m$, and $\eta = \|Av\|_2$ gives

$$\|v\|_p \leq C \cdot \frac{1}{1/p - 1/2} \|v_{\text{sparse}}\|_1 + D \cdot \frac{1}{1/p - 1/2} \|Av\|_2$$

for all $S \subseteq [N]$ with $\text{card}(S) = s$. 

* denotes reference to further technical details or proof.
Stable and Robust Sparse Recovery via $\ell_1$-Minimization*

Objective: for $p \in [1, 2]$, for all $x \in \mathbb{C}^N$ and $e \in \mathbb{C}^m$ with $\|e\|_2 \leq \eta$:

$$\|x - \Delta(Ax + e)\|_p \leq \min_{x_s \text{ s-sparse}} \|x - x_s\|_1 \quad \text{subject to} \quad \|Ax - y\|_2 \leq \eta,$$

where $\Delta(y) = \Delta_{1,\eta}(y) := \arg\min \|z\|_1 \quad \text{subject to} \quad \|Az - y\|_2 \leq \eta$. 

\*Note: This formulation is inspired by the work of Elad and Eldar, and is adapted for the context of sparse recovery. The details of the optimization problem and the implications for stability and robustness are discussed in the next sections.
Stable and Robust Sparse Recovery via $\ell_1$-Minimization*

Objective: for $p \in [1, 2]$, for all $x \in \mathbb{C}^N$ and $e \in \mathbb{C}^m$ with $\|e\|_2 \leq \eta$:

$$\|x - \Delta(Ax + e)\|_p \leq \frac{C}{s^{1-1/p}} \min_{x_s \text{ s-sparse}} \|x - x_s\|_1 + D \frac{s^{1/p - 1/2}}{\eta} \eta,$$

where $\Delta(y) = \Delta_{1,\eta}(y) := \arg\min \|z\|_1$ subject to $\|Az - y\|_2 \leq \eta$.

Taking $x = v \in \mathbb{C}^N$, $e = -Av \in \mathbb{C}^m$, and $\eta = \|Av\|_2$ gives

$$\|v\|_p \leq \frac{C}{s^{1-1/p}} \|v_S\|_1 + D \frac{s^{1/p - 1/2}}{\eta} \|Av\|_2$$

for all $S \subseteq [N]$ with $\text{card}(S) = s$. 
Robust Null Space Property*

For $q \in [1, 2]$, $A \in \mathbb{C}^{m \times N}$ has the $\ell^q$-robust null space property of order $s$ (with respect to $\| \cdot \|$) with constants $0 < \rho < 1$ and $\tau > 0$ if, for any set $S \subset [N]$ with $\text{card}(S) \leq s$,

$$\|v_S\|_q \leq \rho \frac{s}{1 - 1/q} \|v_S\|_1 + \tau \|A v\|$$

for all $v \in \mathbb{C}^N$.

▶ $\ell^q$-RNSP (with respect to $\| \cdot \|$) implies $\ell^p$-RNSP (with respect to $s_1/p - 1/q \| \cdot \|$) whenever $1 \leq p \leq q \leq 2$.

▶ $\ell^q$-RNSP (with respect to $\| \cdot \|$) implies that, for $1 \leq p \leq q$,

$$\|z - x\|_p \leq C \frac{s}{1 - 1/p} (\|z\|_1 - \|x\|_1 + 2 \sigma_s(x)_1) + D \frac{s}{1 - 1/q} \|A(z - x)\|$$

for all $x, z \in \mathbb{C}^N$.

The converse holds when $q = 1$.

▶ $\ell^2$-RNSP (with respect to $\| \cdot \|_2$) holds when $\delta_2^s < 0$.

62.
Robust Null Space Property*

For $q \in [1, 2]$, $A \in \mathbb{C}^{m \times N}$ has the $\ell_q$-robust null space property of order $s$ (wrt $\| \cdot \|$) with constants $0 < \rho < 1$ and $\tau > 0$ if, for any set $S \subset [N]$ with $\text{card}(S) \leq s$,

$$\|v_S\|_q \leq \frac{\rho}{s^{1-1/q}} \|v_{\overline{S}}\|_1 + \tau \|Av\|$$

for all $v \in \mathbb{C}^N$. 

The converse holds when $q = 1$.

$\ell_2$-RNSP (wrt $\| \cdot \|_2$) holds when $\delta_2 < 0$. 62.
Robust Null Space Property*

For $q \in [1, 2]$, $A \in \mathbb{C}^{m \times N}$ has the $\ell_q$-robust null space property of order $s$ (wrto $\| \cdot \|$) with constants $0 < \rho < 1$ and $\tau > 0$ if, for any set $S \subset [N]$ with $\text{card}(S) \leq s$,

$$\|v_S\|_q \leq \frac{\rho}{s^{1-1/q}} \|v_{\overline{S}}\|_1 + \tau \|Av\|$$

for all $v \in \mathbb{C}^N$.

$\ell_q$-RNSP (wrto $\| \cdot \|$) $\Rightarrow$ $\ell_p$-RSNP (wrto $s^{1/p-1/q}\| \cdot \|$) whenever $1 \leq p \leq q \leq 2$. 

$\triangleright$
Robust Null Space Property*

For $q \in [1, 2]$, $A \in \mathbb{C}^{m \times N}$ has the $\ell_q$-robust null space property of order $s$ (wrto $\| \cdot \|$) with constants $0 < \rho < 1$ and $\tau > 0$ if, for any set $S \subset [N]$ with $\text{card}(S) \leq s$,

$$\|v_S\|_q \leq \frac{\rho}{s^{1-1/q}} \|v_{\overline{S}}\|_1 + \tau \|A v\|$$

for all $v \in \mathbb{C}^N$.

- $\ell_q$-RNSP (wrto $\| \cdot \|$) $\Rightarrow$ $\ell_p$-RSNP (wrto $s^{1/p - 1/q}\| \cdot \|$) whenever $1 \leq p \leq q \leq 2$.
- $\ell_q$-RNSP (wrto $\| \cdot \|$) implies that, for $1 \leq p \leq q$,

$$\|z-x\|_p \leq \frac{C}{s^{1-1/p}} \left( \|z\|_1 - \|x\|_1 + 2\sigma_s(x)_1 \right) + D s^{1/p - 1/q} \|A(z-x)\|$$

for all $x, z \in \mathbb{C}^N$. 

The converse holds when $q = 1$. 

- $\ell_2$-RNSP (wrto $\| \cdot \|_2$) holds when $\delta_2 s < 0$. 

- 62.
Robust Null Space Property*

For \( q \in [1, 2] \), \( A \in \mathbb{C}^{m \times N} \) has the \( \ell_q \)-robust null space property of order \( s \) (wrto \( \| \cdot \| \)) with constants \( 0 < \rho < 1 \) and \( \tau > 0 \) if, for any set \( S \subset [N] \) with \( \text{card}(S) \leq s \),

\[
\| v_S \|_q \leq \frac{\rho}{s^{1-1/q}} \| v_S \|_1 + \tau \| Av \| \quad \text{for all } v \in \mathbb{C}^N.
\]

- \( \ell_q \)-RNSP (wrto \( \| \cdot \| \)) \( \Rightarrow \) \( \ell_p \)-RSNP (wrto \( s^{1/p-1/q} \| \cdot \| \)) whenever \( 1 \leq p \leq q \leq 2 \).

- \( \ell_q \)-RNSP (wrto \( \| \cdot \| \)) implies that, for \( 1 \leq p \leq q \),

\[
\| z-x \|_p \leq \frac{C}{s^{1-1/p}} (\| z \|_1-\| x \|_1 + 2\sigma_s(x)_1) + D s^{1/p-1/q} \| A(z-x) \|
\]

for all \( x, z \in \mathbb{C}^N \). The converse holds when \( q = 1 \).
Robust Null Space Property*

For $q \in [1, 2]$, $A \in \mathbb{C}^{m \times N}$ has the $\ell_q$-robust null space property of order $s$ (wrto $\| \cdot \|$) with constants $0 < \rho < 1$ and $\tau > 0$ if, for any set $S \subset [N]$ with $\text{card}(S) \leq s$,

$$\|v_S\|_q \leq \frac{\rho}{s^{1-1/q}} \|v_S\|_1 + \tau \|Av\|$$

for all $v \in \mathbb{C}^N$.

- $\ell_q$-RNSP (wrto $\| \cdot \|$) $\Rightarrow$ $\ell_p$-RSNP (wrto $s^{1/p-1/q} \| \cdot \|$) whenever $1 \leq p \leq q \leq 2$.

- $\ell_q$-RNSP (wrto $\| \cdot \|$) implies that, for $1 \leq p \leq q$,

$$\|z-x\|_p \leq \frac{C}{s^{1-1/p}} \left( \|z\|_1 - \|x\|_1 + 2\sigma_s(x)_1 \right) + D s^{1/p-1/q} \|A(z-x)\|$$

for all $x, z \in \mathbb{C}^N$. The converse holds when $q = 1$.

- $\ell_2$-RNSP (wrto $\| \cdot \|_2$) holds when $\delta_{2s} < 0.62$. 
Iterative Hard Thresholding and Hard Thresholding Pursuit

- Solving the rectangular system $A x = y$ amounts to solving the square system $A^* A x = A^* y$.
- Classical iterative methods suggest the iteration $x_{n+1} = x_n + A^* (y - A x_n)$.
- At each iteration, keep $s$ largest absolute entries and set the other ones to zero.

**IHT:** Start with an $s$-sparse $x_0 \in \mathbb{C}^N$ and iterate:

$$x_{n+1} = H_s \left( x_n + A^* (y - A x_n) \right)$$

until a stopping criterion is met.

**HTP:** Start with an $s$-sparse $x_0 \in \mathbb{C}^N$ and iterate:

$$S_{n+1} = \{s \text{ largest abs. entries of } x_n + A^* (y - A x_n)\},$$

and

$$x_{n+1} = \text{arg min}_{z} \left\{ \| y - A z \|_2, \supp(z) \subseteq S_{n+1} \right\}.$$
Iterative Hard Thresholding and Hard Thresholding Pursuit

- solving the rectangular system $Ax = y$ amounts to solving the square system $A^*Ax = A^*y$,
Iterative Hard Thresholding and Hard Thresholding Pursuit

- solving the rectangular system $Ax = y$ amounts to solving the square system $A^*Ax = A^*y$,
- classical iterative methods suggest the iteration
  $x^{n+1} = x^n + A^*(y - Ax^n)$,
Iterative Hard Thresholding and Hard Thresholding Pursuit

- solving the rectangular system $Ax = y$ amounts to solving the square system $A^*Ax = A^*y$,
- classical iterative methods suggest the iteration $x^{n+1} = x^n + A^*(y - Ax^n)$,
- at each iteration, keep $s$ largest absolute entries and set the other ones to zero.
Iterative Hard Thresholding and Hard Thresholding Pursuit

- solving the rectangular system $Ax = y$ amounts to solving the square system $A^*Ax = A^*y$,
- classical iterative methods suggest the iteration $x^{n+1} = x^n + A^*(y - Ax^n)$,
- at each iteration, keep $s$ largest absolute entries and set the other ones to zero.

**IHT:** Start with an $s$-sparse $x^0 \in \mathbb{C}^N$ and iterate:

$$x^{n+1} = H_s(x^n + A^*(y - Ax^n))$$

until a stopping criterion is met.
Iterative Hard Thresholding and Hard Thresholding Pursuit

- solving the rectangular system \( Ax = y \) amounts to solving the square system \( A^* A x = A^* y \),
- classical iterative methods suggest the iteration \( x^{n+1} = x^n + A^*(y - A x^n) \),
- at each iteration, keep \( s \) largest absolute entries and set the other ones to zero.

IHT: Start with an \( s \)-sparse \( x^0 \in \mathbb{C}^N \) and iterate:

\[
(IHT) \quad x^{n+1} = H_s(x^n + A^*(y - A x^n))
\]

until a stopping criterion is met.

HTP: Start with an \( s \)-sparse \( x^0 \in \mathbb{C}^N \) and iterate:
Iterative Hard Thresholding and Hard Thresholding Pursuit

- solving the rectangular system $Ax = y$ amounts to solving the square system $A^*Ax = A^*y$,
- classical iterative methods suggest the iteration $x^{n+1} = x^n + A^*(y - Ax^n)$,
- at each iteration, keep $s$ largest absolute entries and set the other ones to zero.

IHT: Start with an $s$-sparse $x^0 \in \mathbb{C}^N$ and iterate:

\[
(IHT) \quad x^{n+1} = H_s(x^n + A^*(y - Ax^n))
\]

until a stopping criterion is met.

HTP: Start with an $s$-sparse $x^0 \in \mathbb{C}^N$ and iterate:

\[
(HTP_1)
\]

\[
(HTP_2)
\]
Iterative Hard Thresholding and Hard Thresholding Pursuit

- solving the rectangular system $Ax = y$ amounts to solving the square system $A^*Ax = A^*y$,
- classical iterative methods suggest the iteration $x^{n+1} = x^n + A^*(y - Ax^n)$,
- at each iteration, keep $s$ largest absolute entries and set the other ones to zero.

**IHT:** Start with an $s$-sparse $x^0 \in \mathbb{C}^N$ and iterate:

\[
(IHT) \quad x^{n+1} = H_s(x^n + A^*(y - Ax^n))
\]

until a stopping criterion is met.

**HTP:** Start with an $s$-sparse $x^0 \in \mathbb{C}^N$ and iterate:

\[
(HTP_1) \quad S^{n+1} = \{s \text{ largest abs. entries of } x^n + A^*(y - Ax^n)\},
\]

\[
(HTP_2)
\]
Iterative Hard Thresholding and Hard Thresholding Pursuit

- solving the rectangular system $Ax = y$ amounts to solving the square system $A^*Ax = A^*y$,
- classical iterative methods suggest the iteration $x^{n+1} = x^n + A^*(y - Ax^n)$,
- at each iteration, keep $s$ largest absolute entries and set the other ones to zero.

IHT: Start with an $s$-sparse $x^0 \in \mathbb{C}^N$ and iterate:

$$\text{(IHT)} \quad x^{n+1} = H_s(x^n + A^*(y - Ax^n))$$

until a stopping criterion is met.

HTP: Start with an $s$-sparse $x^0 \in \mathbb{C}^N$ and iterate:

$$\text{(HTP}_1) \quad S^{n+1} = \{ s \text{ largest abs. entries of } x^n + A^*(y - Ax^n) \} ,$$
$$\text{(HTP}_2) \quad x^{n+1} = \text{argmin}\{ \|y - Az\|_2, \text{supp}(z) \subseteq S^{n+1}\} ,$$
Iterative Hard Thresholding and Hard Thresholding Pursuit

- solving the rectangular system $Ax = y$ amounts to solving the square system $A^*Ax = A^*y$.
- classical iterative methods suggest the iteration
  $$x^{n+1} = x^n + A^*(y - Ax^n),$$
  at each iteration, keep $s$ largest absolute entries and set the other ones to zero.

IHT: Start with an $s$-sparse $x^0 \in \mathbb{C}^N$ and iterate:

\begin{align*}
(IHT) \quad x^{n+1} &= H_s(x^n + A^*(y - Ax^n)) \tag{IHT}
\end{align*}

until a stopping criterion is met.

HTP: Start with an $s$-sparse $x^0 \in \mathbb{C}^N$ and iterate:

\begin{align*}
(HTP_1) \quad S^{n+1} &= \{s \text{ largest abs. entries of } x^n + A^*(y - Ax^n)\}, \\
(HTP_2) \quad x^{n+1} &= \text{argmin}\{|y - Az|_2, \text{supp}(z) \subseteq S^{n+1}\}, \tag{HTP}
\end{align*}

until a stopping criterion is met ($S^{n+1} = S^n$ is natural here).
Suppose that $\delta_3 s < 1/\sqrt{3}$. Then, for some $\rho < 1$ and $\tau > 0$, the following facts hold for all $x \in \mathbb{C}^N$ and $e \in \mathbb{C}^m$, where $S$ denotes an index set of $s$ largest absolute entries of $x$:

\begin{align*}
\|x_S - x_n + 1\|_2 &\leq \rho \|x_S - x_n\|_2 + (1 - \rho) \tau \|A(x_S + e)\|_2, \\
\|x_S - x_n\|_2 &\leq \rho \|x_S - x_0\|_2 + \tau \|A(x_S + e)\|_2, \\
\text{The output } x^\# \text{ of the algorithms satisfy} &\\
\|x_S - x^\#\|_2 &\leq \tau \|A(x_S + e)\|_2, \\
\text{With } 2s \text{ in place of } s \text{ in the algorithms, if } &\\
\delta_6 s < 1/\sqrt{3}, \text{ then} &\\
\|x - x^\#\|_p &\leq C s^{1 - 1/p} \sigma_s(x)^{1/p} + D s^{1/p - 1/2} \|e\|_2, \\
\text{For HTP, (pseudo)robustness is achieved in} &\\
\leq c s \text{ iterations.}
\end{align*}
Suppose that \( \delta_{3s} < \frac{1}{\sqrt{3}} \).
Stable and Robust Sparse Recovery via IHT and HTP

Suppose that $\delta_{3s} < 1/\sqrt{3}$. Then, for some $\rho < 1$ and $\tau > 0$, the following facts hold for all $\mathbf{x} \in \mathbb{C}^N$ and $\mathbf{e} \in \mathbb{C}^m$, where $S$ denotes an index set of $s$ largest absolute entries of $\mathbf{x}$:

\begin{align*}
\|\mathbf{x}_S - \mathbf{x}_{n+1}\|_2 & \leq \rho \|\mathbf{x}_S - \mathbf{x}_n\|_2 + (1 - \rho) \tau \|A\mathbf{x}_S + \mathbf{e}\|_2, \\
\|\mathbf{x}_S - \mathbf{x}_0\|_2 & \leq \rho^n \|\mathbf{x}_S - \mathbf{x}_0\|_2 + \tau \|A\mathbf{x}_S + \mathbf{e}\|_2, \\
\text{The output } \mathbf{x}^\# \text{ of the algorithms satisfy } & \|\mathbf{x}_S - \mathbf{x}^\#\|_2 \leq \tau \|A\mathbf{x}_S + \mathbf{e}\|_2, \\
\text{With } 2s \text{ in place of } s \text{ in the algorithms, if } & \delta_{6s} \frac{1}{\sqrt{3}}, \text{ then } \|\mathbf{x} - \mathbf{x}^\#\|_p \leq C s^{1 - 1/p} \sigma_s(\mathbf{x})^{1/p + \frac{1}{p - 1/2}} \|\mathbf{e}\|_2, 1 \leq p \leq 2, \\
\text{For HTP, (pseudo)robustness is achieved in } & \leq c s \text{ iterations.}
\end{align*}
Suppose that $\delta_{3s} < 1/\sqrt{3}$. Then, for some $\rho < 1$ and $\tau > 0$, the following facts hold for all $x \in \mathbb{C}^N$ and $e \in \mathbb{C}^m$, where $S$ denotes an index set of $s$ largest absolute entries of $x$:

1. $\|x_S - x^{n+1}\|_2 \leq \rho \|x_S - x^n\|_2 + (1 - \rho)\tau \|Ax_{\bar{S}} + e\|_2$,
Suppose that $\delta_{3s} < 1/\sqrt{3}$. Then, for some $\rho < 1$ and $\tau > 0$, the following facts hold for all $x \in \mathbb{C}^N$ and $e \in \mathbb{C}^m$, where $S$ denotes an index set of $s$ largest absolute entries of $x$:

- $\|x_S - x^{n+1}\|_2 \leq \rho \|x_S - x^n\|_2 + (1 - \rho)\tau\|Ax_{\overline{S}} + e\|_2$,
- $\|x_S - x^n\|_2 \leq \rho^n\|x_S - x^0\|_2 + \tau\|Ax_{\overline{S}} + e\|_2$, 

For HTP, (pseudo)robustness is achieved in $\leq c_s$ iterations.
Suppose that $\delta_{3s} < 1/\sqrt{3}$. Then, for some $\rho < 1$ and $\tau > 0$, the following facts hold for all $x \in \mathbb{C}^N$ and $e \in \mathbb{C}^m$, where $S$ denotes an index set of $s$ largest absolute entries of $x$:

- $\|x_S - x^{n+1}\|_2 \leq \rho \|x_S - x^n\|_2 + (1 - \rho)\tau \|Ax_\bar{S} + e\|_2$,
- $\|x_S - x^n\|_2 \leq \rho^n \|x_S - x^0\|_2 + \tau \|Ax_\bar{S} + e\|_2$,
- The output $x^\#$ of the algorithms satisfy

$$\|x_S - x^\#\|_2 \leq \tau \|Ax_\bar{S} + e\|_2,$$
Suppose that $\delta_{3s} < 1/\sqrt{3}$. Then, for some $\rho < 1$ and $\tau > 0$, the following facts hold for all $x \in \mathbb{C}^N$ and $e \in \mathbb{C}^m$, where $S$ denotes an index set of $s$ largest absolute entries of $x$:

- $\|x_S - x_{n+1}\|_2 \leq \rho \|x_S - x^n\|_2 + (1 - \rho)\tau \|Ax_\bar{S} + e\|_2$,
- $\|x_S - x^n\|_2 \leq \rho^n \|x_S - x^0\|_2 + \tau \|Ax_\bar{S} + e\|_2$,
- The output $x^\#$ of the algorithms satisfy
  \[ \|x_S - x^\#\|_2 \leq \tau \|Ax_\bar{S} + e\|_2, \]
- With $2s$ in place of $s$ in the algorithms, if $\delta_{6s} < 1/\sqrt{3}$, then
  \[ \|x - x^\#\|_p \leq \frac{C}{s^{1-1/p}} \sigma_s(x)_1 + D s^{1/p - 1/2} \|e\|_2, \quad 1 \leq p \leq 2, \]
Suppose that $\delta_{3s} < 1/\sqrt{3}$. Then, for some $\rho < 1$ and $\tau > 0$, the following facts hold for all $x \in \mathbb{C}^N$ and $e \in \mathbb{C}^m$, where $S$ denotes an index set of $s$ largest absolute entries of $x$:

- $\|x_S - x^{n+1}\|_2 \leq \rho \|x_S - x^n\|_2 + (1 - \rho)\tau \|Ax_S + e\|_2$,
- $\|x_S - x^n\|_2 \leq \rho^n \|x_S - x^0\|_2 + \tau \|Ax_S + e\|_2$,
- The output $x^\#$ of the algorithms satisfy
  \[ \|x_S - x^\#\|_2 \leq \tau \|Ax_S + e\|_2, \]
- With $2s$ in place of $s$ in the algorithms, if $\delta_{6s} < 1/\sqrt{3}$, then
  \[ \|x - x^\#\|_p \leq \frac{C}{s^{1 - 1/p}} \sigma_s(x)_1 + D s^{1/p - 1/2} \|e\|_2, \quad 1 \leq p \leq 2, \]
- For HTP, (pseudo)robustness is achieved in $\leq cs$ iterations.
Recovery by Orthogonal Matching Pursuit

OMP: Start with $S_0 = \emptyset$ and $x_0 = 0$, and iterate:

$S_n = S_{n-1} \cup \{j_n : \arg\max_j \| (A^* (y - A x_{n-1}))_j \| \}$,

$x_n = \arg\min_z \{ \| y - A z \|_2, \text{supp}(z) \subseteq S_n \}$.

If $A$ has $\ell_2$-normalized columns $a_1, \ldots, a_N$, then

$\| y - A x_n \|_2^2 = \| y - A x_{n-1} \|_2^2 - \| A^* (y - A x_{n-1}) \|_2^2 \leq \| y - A x_{n-1} \|_2^2 - \| A^* (y - A x_{n-1}) \|_2^2$.

▶ Suppose that $\delta_{13} < 1/6$. Then $s$-sparse recovery via $12s$ iterations of OMP is stable and robust.

▶ Challenge: natural proof of OMP success in $c s$ iterations?
Recovery by Orthogonal Matching Pursuit

OMP: Start with $S^0 = \emptyset$ and $x^0 = 0$, and iterate:

$$S^n = S^{n-1} \cup \{j^n : = \text{argmax}_j |(A^*(y - Ax^{n-1}))_j| \},$$

$$x^n = \text{argmin}_z \{\|y - Az\|_2, \text{supp}(z) \subseteq S^n \}.$$
Recovery by Orthogonal Matching Pursuit

OMP: Start with $S^0 = \emptyset$ and $x^0 = 0$, and iterate:

\[ S^n = S^{n-1} \cup \{ j^n := \text{argmax}_j |(A^*(y - Ax^{n-1}))_j| \}, \]
\[ x^n = \text{argmin}_z \{ \|y - Az\|_2, \text{supp}(z) \subseteq S^n \}. \]

If $A$ has $\ell_2$-normalized columns $a_1, \ldots, a_N$, then

\[ \|y - Ax^n\|_2^2 = \|y - Ax^{n-1}\|_2^2 - \frac{|(A^*(y - Ax^{n-1}))_{j^n}|^2}{d(a_{j^n}, \text{span}[a_j, j \in S^{n-1}])^2} \]
Recovery by Orthogonal Matching Pursuit

OMP: Start with $S^0 = \emptyset$ and $x^0 = 0$, and iterate:

$$S^n = S^{n-1} \cup \{j^n := \arg\max_j \| (A^*(y - Ax^{n-1}))_j \| \},$$

$$x^n = \arg\min_z \{ \| y - Az \|_2, \text{supp}(z) \subseteq S^n \}.$$

If $A$ has $\ell_2$-normalized columns $a_1, \ldots, a_N$, then

$$\| y - Ax^n \|_2^2 = \| y - Ax^{n-1} \|_2^2 - \frac{|(A^*(y - Ax^{n-1}))_{j^n}|^2}{d(a_{j^n}, \text{span}[a_j, j \in S^{n-1}])^2}$$

$$\leq \| y - Ax^{n-1} \|_2^2 - |(A^*(y - Ax^{n-1}))_{j^n}|^2.$$
Recovery by Orthogonal Matching Pursuit

OMP: Start with $S^0 = \emptyset$ and $x^0 = 0$, and iterate:

$$S^n = S^{n-1} \cup \{j^n := \arg\max_j \left| (A^* (y - Ax^{n-1}))_j \right| \},$$

$$x^n = \arg\min_z \{ \|y - Az\|_2, \text{supp}(z) \subseteq S^n \}.$$ 

If $A$ has $\ell_2$-normalized columns $a_1, \ldots, a_N$, then

$$\|y - Ax^n\|_2^2 = \|y - Ax^{n-1}\|_2^2 - \frac{\left| (A^* (y - Ax^{n-1}))_{j^n} \right|^2}{d(a_{j^n}, \text{span}[a_j, j \in S^{n-1}])^2} \leq \|y - Ax^{n-1}\|_2^2 - \left| (A^* (y - Ax^{n-1}))_{j^n} \right|^2.$$ 

▶ Suppose that $\delta_{13s} < 1/6$. Then $s$-sparse recovery via $12s$ iterations of OMP is stable and robust.
Recovery by Orthogonal Matching Pursuit

OMP: Start with \( S^0 = \emptyset \) and \( x^0 = 0 \), and iterate:

\[
S^n = S^{n-1} \cup \{ j^n := \text{argmax}_j |(A^*(y - Ax^{n-1}))_j| \},
\]

\[
x^n = \text{argmin}_z \{ \|y - Az\|_2, \text{supp}(z) \subseteq S^n \}.
\]

If \( A \) has \( \ell_2 \)-normalized columns \( a_1, \ldots, a_N \), then

\[
\|y - Ax^n\|_2^2 = \|y - Ax^{n-1}\|_2^2 - \frac{|(A^*(y - Ax^{n-1}))_{j^n}|^2}{d(a_{j^n}, \text{span}[a_j, j \in S^{n-1}])^2} \leq \|y - Ax^{n-1}\|_2^2 - |(A^*(y - Ax^{n-1}))_{j^n}|^2.
\]

▶ Suppose that \( \delta_{13s} < 1/6 \). Then \( s \)-sparse recovery via \( 12s \) iterations of OMP is stable and robust.

▶ Challenge: natural proof of OMP success in \( cs \) iterations?
RIP for Random Matrices
Concentration Inequality

Let \( A \in \mathbb{R}^{m \times N} \) be a random matrix with entries \( a_{i,j} = g_{i,j} \sqrt{m} \) where the \( g_{i,j} \) are independent \( \mathcal{N}(0,1) \).

For a fixed \( x \in \mathbb{R}^N \), note that \( (Ax)_i = \sum_{j=1}^N a_{i,j}x_j \), hence
\[
\mathbb{E}((Ax)_i^2) = \sum_{j=1}^N x_j^2 \mathbb{E}(a_{i,j}^2) = \|x\|_2^2 \frac{1}{m},
\]
and
\[
\mathbb{E}(\|Ax\|_2^2) = \|x\|_2^2 \frac{1}{m}.
\]

In fact, \( \|Ax\|_2^2 \) concentrates around its mean: for \( t \in (0,1) \),
\[
\mathbb{P}(\|Ax\|_2^2 - \|x\|_2^2 > t \|x\|_2^2) \leq 2 \exp(-ct^2 m),
\]
Concentration Inequality

Let $A \in \mathbb{R}^{m \times N}$ be a random matrix with entries

$$a_{i,j} = \frac{g_{i,j}}{\sqrt{m}}$$

where the $g_{i,j}$ are independent $\mathcal{N}(0, 1)$.
Concentration Inequality

Let $A \in \mathbb{R}^{m \times N}$ be a random matrix with entries

$$a_{i,j} = \frac{g_{i,j}}{\sqrt{m}}$$

where the $g_{i,j}$ are independent $\mathcal{N}(0, 1)$.

For a fixed $x \in \mathbb{R}^N$, note that $(Ax)_i = \sum_{j=1}^{N} a_{i,j}x_j$, hence

\[
\mathbb{E}((Ax)_i^2) = \mathbb{V}(\sum a_{i,j}x_j) = \sum x_j^2 \mathbb{V}(a_{i,j}) = \frac{\|x\|_2^2}{m},
\]

\[
\mathbb{E}(\|Ax\|_2^2) = \|x\|_2^2.
\]
Concentration Inequality

Let $A \in \mathbb{R}^{m \times N}$ be a random matrix with entries 

$$a_{i,j} = \frac{g_{i,j}}{\sqrt{m}}$$

where the $g_{i,j}$ are independent $\mathcal{N}(0, 1)$.

For a fixed $x \in \mathbb{R}^N$, note that $(Ax)_i = \sum_{j=1}^{N} a_{i,j}x_j$, hence

$$\mathbb{E}((Ax)_i^2) = \mathbb{V}(\sum a_{i,j}x_j) = \sum x_j^2 \mathbb{V}(a_{i,j}) = \frac{\|x\|_2^2}{m},$$

$$\mathbb{E}(\|Ax\|_2^2) = \|x\|_2^2.$$

In fact, $\|Ax\|_2^2$ concentrates around its mean: for $t \in (0, 1)$,

$$(CI) \quad \mathbb{P}(|\|Ax\|_2^2 - \|x\|_2^2| > t\|x\|_2^2) \leq 2 \exp(-ct^2 m).$$
Suppose that the random matrix $A \in \mathbb{R}^{m \times N}$ satisfies (CI).

Let $S \subseteq [N]$ with $\text{card}(S) = s$.

Then $P\left(\|A^* S A S - I_d\|_2 \rightarrow 2 > \delta\right) \leq 2 \exp\left(-c \delta^2 m\right)$ provided $m \geq c' \delta^2 s$.

The argument relies on the following fact:

A subset $U$ of the unit ball of $\mathbb{R}^k$ relative to a norm $\|\cdot\|$ has covering and packing numbers satisfying $N(U, \|\cdot\|, t) \leq P(U, \|\cdot\|, t) \leq (1 + 2t)^k$. 
Covering Arguments

Suppose that the random matrix $A \in \mathbb{R}^{m \times N}$ satisfies (CI).

Then $P(\|A^*SAA - \text{Id}\|_2 \to 2 > \delta) \leq 2 \exp(-c\delta^2 m)$ provided $m \geq c'\delta^2 s$.

The argument relies on the following fact: A subset $U$ of the unit ball of $\mathbb{R}^k$ relative to a norm $\|\cdot\|$ has covering and packing numbers satisfying $N(U,\|\cdot\|,t) \leq P(U,\|\cdot\|,t) \leq (1 + 2t)^k$.
Covering Arguments

Suppose that the random matrix $A \in \mathbb{R}^{m \times N}$ satisfies (CI). Let $S \subseteq [N]$ with $\text{card}(S) = s$. 

$$
\text{P} \left( \|A^* S A S - I\|_2 > \delta \right) \leq 2 \exp\left(-c\delta^2 m\right)
$$

provided $m \geq c'\delta^2 s$. The argument relies on the following fact: A subset $U$ of the unit ball of $\mathbb{R}^k$ relative to a norm $\|\cdot\|$ has covering and packing numbers satisfying 

$$
N(U, \|\cdot\|, t) \leq P(U, \|\cdot\|, t) \leq (1 + 2t)^k
$$
Covering Arguments

Suppose that the random matrix $A \in \mathbb{R}^{m \times N}$ satisfies (CI). Let $S \subseteq [N]$ with $\text{card}(S) = s$. Then

$$\mathbb{P}(\|A^*_S A_S - \text{Id}\|_{2\rightarrow2} > \delta) \leq 2 \exp(-c\delta^2 m)$$
Covering Arguments

Suppose that the random matrix $A \in \mathbb{R}^{m \times N}$ satisfies (CI). Let $S \subseteq [N]$ with $\text{card}(S) = s$. Then

$$\mathbb{P}(\|A^*_S A_S - \text{Id}\|_{2 \to 2} > \delta) \leq 2 \exp(-c\delta^2 m)$$

provided

$$m \geq \frac{c'}{\delta^2} s.$$
Covering Arguments

Suppose that the random matrix $A \in \mathbb{R}^{m \times N}$ satisfies (CI). Let $S \subseteq [N]$ with $\text{card}(S) = s$. Then

$$\mathbb{P}(\| A_S^* A_S - \text{Id} \|_{2 \rightarrow 2} > \delta) \leq 2 \exp(-c\delta^2 m)$$

provided

$$m \geq \frac{c'}{\delta^2} s.$$ 

The argument relies on the following fact:

A subset $U$ of the unit ball of $\mathbb{R}^k$ relative to a norm $\| \cdot \|$ has covering and packing numbers satisfying

$$\mathcal{N}(U, \| \cdot \|, t) \leq \mathcal{P}(U, \| \cdot \|, t) \leq \left(1 + \frac{2}{t}\right)^k.$$
Restricted Isometry Property

Suppose that the random matrix $A \in \mathbb{R}^{m \times N}$ satisfies (CI). Then

$$P(\delta_s > \delta) \leq 2 \exp\left(-c\delta_s^2 m\right)$$

provided $m \geq c' \delta_s^2 \ln(eN/s)$. The arguments are also valid for subgaussian matrices (e.g. Bernoulli matrices), since these satisfy (CI), too. For Gaussian matrices, more powerful techniques can provide an explicit value for $c'$. 
Suppose that the random matrix $A \in \mathbb{R}^{m \times N}$ satisfies (CI). Then

$$\mathbb{P}(\delta_s > \delta) \leq 2 \exp(-c\delta^2 m)$$

provided

$$m \geq \frac{c'}{\delta^2} s \ln(eN/s).$$
Suppose that the random matrix $A \in \mathbb{R}^{m \times N}$ satisfies (CI). Then

$$\mathbb{P}(\delta_s > \delta) \leq 2 \exp(-c\delta^2 m)$$

provided

$$m \geq \frac{c'}{\delta^2} s \ln(eN/s).$$

The arguments are also valid for subgaussian matrices (e.g. Bernoulli matrices), since these satisfy (CI), too.
Suppose that the random matrix $A \in \mathbb{R}^{m \times N}$ satisfies (CI). Then

$$\mathbb{P}(\delta_s > \delta) \leq 2 \exp(-c\delta^2 m)$$

provided

$$m \geq \frac{c'}{\delta^2} s \ln(eN/s).$$

The arguments are also valid for subgaussian matrices (e.g. Bernoulli matrices), since these satisfy (CI), too.

For Gaussian matrices, more powerful techniques can provide an explicit value for $c'$. 
Summary of Part II

The RI conditions for $s$-sparse recovery are of the type $\delta_\kappa s < \delta^*$. They guarantee stable and robust reconstructions in the form, say,

$$\|x - \Delta(Ax + e)\|_2 \leq C\sqrt{s}\sigma(x) + D\|e\|_2$$

for all $x$ and all $e$.

Random matrices fulfill the RI conditions with high probability as soon as

$$m \geq c s \ln(N/s).$$

Next, we will see that this number of measurements is optimal, in the sense that estimates of type (1) require (2) to hold.

We will also examine the gain in replacing $\sigma(x)$ in (1) by $\sigma(x)$ for a fixed $x$. 
Summary of Part II

The RI conditions for $s$-sparse recovery are of the type

$$\delta_{\kappa s} < \delta_*.$$
Summary of Part II

The RI conditions for $s$-sparse recovery are of the type

$$\delta_{\kappa s} < \delta_* .$$

They guarantee stable and robust reconstructions in the form, say,

$$(1) \| x - \Delta(Ax + e) \| _2 \leq \frac{C}{\sqrt{s}} \sigma_s(x)_1 + D \| e \| _2$$

for all $x$ and all $e$. 
Summary of Part II

The RI conditions for \( s \)-sparse recovery are of the type

\[
\delta_{\kappa s} < \delta_*. 
\]

They guarantee stable and robust reconstructions in the form, say,

\[
(1) \quad \|x - \Delta(Ax + e)\|_2 \leq \frac{C}{\sqrt{s}} \sigma_s(x)_1 + D\|e\|_2 \quad \text{for all } x \text{ and all } e.
\]

Random matrices fulfill the RI conditions with high probability as soon as

\[
(2) \quad m \geq cs \ln(N/s).
\]
Summary of Part II

The RI conditions for $s$-sparse recovery are of the type

$$\delta_{\kappa s} < \delta_*.$$  

They guarantee stable and robust reconstructions in the form, say,

\[
\|x - \Delta (Ax + e)\|_2 \leq \frac{C}{\sqrt{s}} \sigma_s(x)_1 + D\|e\|_2 \quad \text{for all } x \text{ and all } e.
\]

Random matrices fulfill the RI conditions with high probability as soon as

\[
m \geq c s \ln(N/s).
\]

Next, we will see that this number of measurement is optimal, in the sense that estimates of type (1) require (2) to hold.
Summary of Part II

The RI conditions for $s$-sparse recovery are of the type

$$\delta_{\kappa s} < \delta_*.$$

They guarantee stable and robust reconstructions in the form, say,

$$\|x - \Delta(Ax + e)\|_2 \leq \frac{C}{\sqrt{s}} \sigma_s(x)_1 + D\|e\|_2$$

for all $x$ and all $e$.

Random matrices fulfill the RI conditions with high probability as soon as

$$m \geq c s \ln(N/s).$$

Next, we will see that this number of measurement is optimal, in the sense that estimates of type (1) require (2) to hold. We will also examine the gain in replacing for all $x$ in (1) by for a fixed $x$. 
Part III:

Optimality, Uniform vs Nonuniform Recovery
Optimality of Uniform Guarantees
Gelfand Widths

For a subset $K$ of a normed space $X$, define $E_m(K, X) := \inf \left\{ \sup_{x \in K} \| x - \Delta(Ax) \|, \text{ } A: X \text{ linear } \rightarrow \mathbb{R}^m, \Delta: \mathbb{R}^m \rightarrow X \right\}$.

The Gelfand $m$-width of $K$ in $X$ is $d_m(K, X) := \inf \left\{ \sup_{x \in K \cap L_m} \| x \|, L_m \text{ subspace of } X, \text{codim}(L_m) \leq m \right\}$.

If $-K = K$, then $d_m(K, X) \leq E_m(K, X)$, and if in addition $K + K \subseteq aK$, then $E_m(K, X) \leq a d_m(K, X)$.
Gelfand Widths

For a subset $K$ of a normed space $X$, define

$$E^m(K, X) := \inf \left\{ \sup_{x \in K} \| x - \Delta(Ax) \|, \quad A : X \xrightarrow{\text{linear}} \mathbb{R}^m, \Delta : \mathbb{R}^m \to X \right\}$$
Gelfand Widths

For a subset $K$ of a normed space $X$, define

$$E^m(K, X) := \inf \left\{ \sup_{x \in K} \|x - \Delta(Ax)\|, \ A : X \to \mathbb{R}^m, \Delta : \mathbb{R}^m \to X \right\}$$

The Gelfand $m$-width of $K$ in $X$ is

$$d^m(K, X) := \inf \left\{ \sup_{x \in K \cap L^m} \|x\|, \ L^m \text{ subspace of } X, \ \text{codim}(L^m) \leq m \right\}$$
Gelfand Widths

For a subset $K$ of a normed space $X$, define

$$E^m(K, X) := \inf \left\{ \sup_{x \in K} \|x - \Delta(Ax)\|, \ A : X \xrightarrow{\text{linear}} \mathbb{R}^m, \Delta : \mathbb{R}^m \to X \right\}$$

The Gelfand $m$-width of $K$ in $X$ is

$$d^m(K, X) := \inf \left\{ \sup_{x \in K \cap L^m} \|x\|, \text{ } L^m \text{ subspace of } X, \text{ codim}(L^m) \leq m \right\}$$

If $-K = K$, then

$$d^m(K, X) \leq E^m(K, X),$$
Gelfand Widths

For a subset $K$ of a normed space $X$, define

$$E^m(K, X) := \inf \left\{ \sup_{x \in K} \| x - \Delta(Ax) \|, \ A : X \to \mathbb{R}^m, \Delta : \mathbb{R}^m \to X \right\}$$

The Gelfand $m$-width of $K$ in $X$ is

$$d^m(K, X) := \inf \left\{ \sup_{x \in K \cap L^m} \| x \|, \ L^m \text{ subspace of } X, \ \text{codim}(L^m) \leq m \right\}$$

If $-K = K$, then

$$d^m(K, X) \leq E^m(K, X),$$

and if in addition $K + K \subseteq aK$, then

$$E^m(K, X) \leq a d^m(K, X).$$
Let $A \in \mathbb{R}^{m \times N}$ with $m \approx c_s \ln(eN/m)$ be such that $\delta_2 < 1/2$.

Let $\Delta_1 : \mathbb{R}^m \rightarrow \mathbb{R}^N$ be the $\ell_1$-minimization map.

Given $1 < p \leq 2$,

$$\|x - \Delta_1(Ax)\|_p \leq C_s 1 - 1/p \sigma_s(x) \leq C_s 1 - 1/p \approx C'_s (m/\ln(eN/m)) \leq 1 - 1/p.$$
Gelfand Widths of $\ell_1$-Balls: Upper Bound

Let $A \in \mathbb{R}^{m \times N}$ with $m \approx cs \ln(eN/m)$ be such that $\delta_{2s} < 1/2$. 
Let $A \in \mathbb{R}^{m \times N}$ with $m \approx c \, s \ln(eN/m)$ be such that $\delta_{2s} < 1/2$. Let $\Delta_1 : \mathbb{R}^m \to \mathbb{R}^N$ be the $\ell_1$-minimization map.
Gelfand Widths of $\ell_1$-Balls: Upper Bound

Let $A \in \mathbb{R}^{m \times N}$ with $m \approx cs \ln(eN/m)$ be such that $\delta_{2s} < 1/2$. Let $\Delta_1 : \mathbb{R}^m \to \mathbb{R}^N$ be the $\ell_1$-minimization map. Given $1 < p \leq 2$, for any $x \in B_1^N$,

$$\|x - \Delta_1(Ax)\|_p \leq \frac{C}{s^{1-1/p}} \sigma_s(x)_1 \leq \frac{C}{s^{1-1/p}} \approx \frac{C'}{(m/\ln(eN/m))^{1-1/p}}.$$
Gelfand Widths of $\ell_1$-Balls: Upper Bound

Let $A \in \mathbb{R}^{m \times N}$ with $m \approx cs \ln(eN/m)$ be such that $\delta_{2s} < 1/2$.
Let $\Delta_1 : \mathbb{R}^m \to \mathbb{R}^N$ be the $\ell_1$-minimization map. Given $1 < p \leq 2$,
for any $x \in B_1^N$,

$$\|x - \Delta_1(Ax)\|_p \leq \frac{C}{s^{1-1/p}} \sigma_s(x)_1 \leq \frac{C}{s^{1-1/p}} \approx \frac{C'}{(m/\ln(eN/m))^{1-1/p}}.$$ 

This gives an upper bound for $E^m(B_1^N, \ell_p^N)$,
Gelfand Widths of $\ell_1$-Balls: Upper Bound

Let $A \in \mathbb{R}^{m \times N}$ with $m \approx c_s \ln(eN/m)$ be such that $\delta_{2s} < 1/2$. Let $\Delta_1 : \mathbb{R}^m \to \mathbb{R}^N$ be the $\ell_1$-minimization map. Given $1 < p \leq 2$, for any $x \in B_1^N$,

$$\|x - \Delta_1(Ax)\|_p \leq \frac{C}{s^{1-1/p}} \sigma_s(x)_1 \leq \frac{C}{s^{1-1/p}} \approx \frac{C'}{(m/\ln(eN/m))^{1-1/p}}.$$

This gives an upper bound for $E^m(B_1^N, \ell_p^N)$, and in turn

$$d^m(B_1^N, \ell_p^N) \leq C \min \left\{ 1, \frac{\ln(eN/m)}{m} \right\}^{1-1/p}.$$
The Gelfand width of $B_N$ in $\ell_p$, $p > 1$, also satisfies
$$d_m(B_N, \ell_p) \geq c_{\min}\left\{1, \frac{\ln(eN/m)}{m}\right\}^{1 - 1/p}.$$ Once established, this will show the optimality of the CS results.

Indeed, suppose the existence of $(A, \Delta)$ such that
$$\|x - \Delta(\text{Ax})\|_p \leq C_s^{1 - 1/p} \sigma_s(x) \|x\|_1$$ for all $x \in \mathbb{R}^N$. Then,
$$c_{\min}\left\{1, \frac{\ln(eN/m)}{m}\right\}^{1 - 1/p} \leq d_m(B_N, \ell_p) \leq E_m(B_N, \ell_p) \leq C_s^{1 - 1/p},$$ in other words
$$c'_{\min}\left\{1, \frac{\ln(eN/m)}{m}\right\} \leq 1.$$ We derive either $s \leq 1/c'$ or $m \geq c's\ln(eN/m)$, in which case
$$m \geq c's\ln(eN/s) + c's\ln(s/m) - c'ms\ln(s/m) \geq \cdots \geq c''s\ln(eN/s).$$
Gelfand Widths of $\ell_1$-Balls: Lower Bound

The Gelfand width of $B_1^N$ in $\ell_p^N$, $p > 1$, also satisfies

$$d^m(B_1^N, \ell_p^N) \geq c \min \left\{ 1, \frac{\ln(eN/m)}{m} \right\}^{1-1/p}.$$
Gelfand Widths of $\ell_1$-Balls: Lower Bound

The Gelfand width of $B_1^N$ in $\ell_p^N$, $p > 1$, also satisfies

\[ d^m(B_1^N, \ell_p^N) \geq c \min \left\{ 1, \frac{\ln(eN/m)}{m} \right\}^{1-1/p}. \]

Once established, this will show the optimality of the CS results.
Gelfand Widths of $\ell_1$-Balls: Lower Bound

The Gelfand width of $B_1^N$ in $\ell_p^N$, $p > 1$, also satisfies

$$d^m(B_1^N, \ell_p^N) \geq c \min \left\{ 1, \frac{\ln(eN/m)}{m} \right\}^{1-1/p}.$$

Once established, this will show the optimality of the CS results. Indeed, suppose the existence of $(A, \Delta)$ such that

$$\|x - \Delta(Ax)\|_p \leq \frac{C}{s^{1-1/p}} \sigma_s(x)_1$$

for all $x \in \mathbb{R}^N$. 

Gelfand Widths of $\ell_1$-Balls: Lower Bound

The Gelfand width of $B_1^N$ in $\ell_p^N$, $p > 1$, also satisfies

$$d^m(B_1^N, \ell_p^N) \geq c \min \left\{ 1, \frac{\ln(eN/m)}{m} \right\}^{1-1/p}.$$ 

Once established, this will show the optimality of the CS results. Indeed, suppose the existence of $(A, \Delta)$ such that

$$\|x - \Delta(Ax)\|_p \leq \frac{C}{s^{1-1/p}} \sigma_s(x)_1 \leq \frac{C}{s^{1-1/p}} \|x\|_1 \quad \text{for all } x \in \mathbb{R}^N.$$
Gelfand Widths of $\ell_1$-Balls: Lower Bound

The Gelfand width of $B^N_1$ in $\ell^N_p$, $p > 1$, also satisfies

$$d^m(B^N_1, \ell^N_p) \geq c \min \left\{ 1, \frac{\ln(\text{e}N/m)}{m} \right\}^{1-1/p}.$$

Once established, this will show the optimality of the CS results. Indeed, suppose the existence of $(A, \Delta)$ such that

$$\|x - \Delta(Ax)\|_p \leq \frac{C}{s^{1-1/p}} \sigma_s(x)_1 \leq \frac{C}{s^{1-1/p}} \|x\|_1 \quad \text{for all } x \in \mathbb{R}^N.$$

Then,

$$c \min \left\{ 1, \frac{\ln(\text{e}N/m)}{m} \right\}^{1-1/p} \leq d^m(B^N_1, \ell^N_p) \leq E^m(B^N_1, \ell^N_p) \leq \frac{C}{s^{1-1/p}},$$
Gelfand Widths of $\ell_1$-Balls: Lower Bound

The Gelfand width of $B^N_1$ in $\ell^N_p$, $p > 1$, also satisfies

$$d^m(B^N_1, \ell^N_p) \geq c \min \left\{ 1, \frac{\ln(eN/m)}{m} \right\}^{1-1/p}.$$

Once established, this will show the optimality of the CS results. Indeed, suppose the existence of $(A, \Delta)$ such that

$$\|x - \Delta(Ax)\|_p \leq \frac{C}{s^{1-1/p}} \sigma_s(x)_1 \leq \frac{C}{s^{1-1/p}} \|x\|_1$$

for all $x \in \mathbb{R}^N$.

Then,

$$c \min \left\{ 1, \frac{\ln(eN/m)}{m} \right\}^{1-1/p} \leq d^m(B^N_1, \ell^N_p) \leq E^m(B^N_1, \ell^N_p) \leq \frac{C}{s^{1-1/p}},$$

in other words

$$c' \min \left\{ 1, \frac{\ln(eN/m)}{m} \right\} \leq \frac{1}{s}.$$
Gelfand Widths of $\ell_1$-Balls: Lower Bound

The Gelfand width of $B_1^N$ in $\ell_p^N$, $p > 1$, also satisfies

$$d^m(B_1^N, \ell_p^N) \geq c \min\left\{1, \frac{\ln(eN/m)}{m}\right\}^{1-1/p}.$$  

Once established, this will show the optimality of the CS results. Indeed, suppose the existence of $(A, \Delta)$ such that

$$\|x - \Delta(Ax)\|_p \leq \frac{C}{s^{1-1/p}} \sigma_s(x)_1 \leq \frac{C}{s^{1-1/p}} \|x\|_1 \quad \text{for all } x \in \mathbb{R}^N.$$  

Then,

$$c \min\left\{1, \frac{\ln(eN/m)}{m}\right\}^{1-1/p} \leq d^m(B_1^N, \ell_p^N) \leq E^m(B_1^N, \ell_p^N) \leq \frac{C}{s^{1-1/p}},$$

in other words

$$c' \min\left\{1, \frac{\ln(eN/m)}{m}\right\} \leq \frac{1}{s}.$$  

We derive either $s \leq 1/c'$ or $m \geq c' s \ln(eN/m)$,
Gelfand Widths of $\ell_1$-Balls: Lower Bound

The Gelfand width of $B_1^N$ in $\ell_p^N$, $p > 1$, also satisfies

$$d_m(B_1^N, \ell_p^N) \geq c \min \left\{ 1, \frac{\ln(eN/m)}{m} \right\}^{1-1/p}.$$ 

Once established, this will show the optimality of the CS results. Indeed, suppose the existence of $(A, \Delta)$ such that

$$\|x - \Delta(Ax)\|_p \leq \frac{C}{s^{1-1/p}} \sigma_s(x)_1 \leq \frac{C}{s^{1-1/p}} \|x\|_1 \quad \text{for all } x \in \mathbb{R}^N.$$ 

Then,

$$c \min \left\{ 1, \frac{\ln(eN/m)}{m} \right\}^{1-1/p} \leq d_m(B_1^N, \ell_p^N) \leq E_m(B_1^N, \ell_p^N) \leq \frac{C}{s^{1-1/p}},$$

in other words

$$c' \min \left\{ 1, \frac{\ln(eN/m)}{m} \right\} \leq \frac{1}{s}.$$

We derive either $s \leq 1/c'$ or $m \geq c' s \ln(eN/m)$, in which case
Gelfand Widths of $\ell_1$-Balls: Lower Bound

The Gelfand width of $B_1^N$ in $\ell_p^N$, $p > 1$, also satisfies

$$d^m(B_1^N, \ell_p^N) \geq c \min \left\{ 1, \frac{\ln(eN/m)}{m} \right\}^{1 - 1/p}.\$$

Once established, this will show the optimality of the CS results. Indeed, suppose the existence of $(A, \Delta)$ such that

$$\|x - \Delta(Ax)\|_p \leq \frac{C}{s^{1 - 1/p}} \sigma_s(x)_1 \leq \frac{C}{s^{1 - 1/p}} \|x\|_1 \quad \text{for all } x \in \mathbb{R}^N.$$

Then,

$$c \min \left\{ 1, \frac{\ln(eN/m)}{m} \right\}^{1 - 1/p} \leq d^m(B_1^N, \ell_p^N) \leq E^m(B_1^N, \ell_p^N) \leq \frac{C}{s^{1 - 1/p}},$$

in other words

$$c' \min \left\{ 1, \frac{\ln(eN/m)}{m} \right\} \leq \frac{1}{s}.$$

We derive either $s \leq 1/c'$ or $m \geq c' s \ln(eN/m)$, in which case

$$m \geq c' s \ln \left( \frac{eN}{s} \right) + c' s \ln \left( \frac{s}{m} \right)$$
Gelfand Widths of $\ell_1$-Balls: Lower Bound

The Gelfand width of $B_1^N$ in $\ell_p^N$, $p > 1$, also satisfies

$$d^m(B_1^N, \ell_p^N) \geq c \min \left\{ 1, \frac{\ln(eN/m)}{m} \right\}^{1-1/p}.$$  

Once established, this will show the optimality of the CS results. Indeed, suppose the existence of $(A, \Delta)$ such that

$$\|x - \Delta(Ax)\|_p \leq \frac{C}{s^{1-1/p}} \sigma_s(x)_1 \leq \frac{C}{s^{1-1/p}} \|x\|_1$$  

for all $x \in \mathbb{R}^N$.

Then,  

$$c \min \left\{ 1, \frac{\ln(eN/m)}{m} \right\}^{1-1/p} \leq d^m(B_1^N, \ell_p^N) \leq E^m(B_1^N, \ell_p^N) \leq \frac{C}{s^{1-1/p}},$$

in other words

$$c' \min \left\{ 1, \frac{\ln(eN/m)}{m} \right\} \leq \frac{1}{s}.$$  

We derive either $s \leq 1/c'$ or $m \geq c' s \ln(eN/m)$, in which case

$$m \geq c' s \ln \left( \frac{eN}{s} \right) + c' m \frac{s}{m} \ln \left( \frac{s}{m} \right).$$
Gelfand Widths of $\ell_1$-Balls: Lower Bound

The Gelfand width of $B_1^N$ in $\ell_p^N$, $p > 1$, also satisfies

$$d^m(B_1^N, \ell_p^N) \geq c \min\left\{1, \frac{\ln(eN/m)}{m}\right\}^{1-1/p}.$$

Once established, this will show the optimality of the CS results. Indeed, suppose the existence of $(A, \Delta)$ such that

$$\|x - \Delta(Ax)\|_p \leq \frac{C}{s^{1-1/p}} \sigma_s(x)_1 \leq \frac{C}{s^{1-1/p}} \|x\|_1$$

for all $x \in \mathbb{R}^N$.

Then,

$$c \min\left\{1, \frac{\ln(eN/m)}{m}\right\}^{1-1/p} \leq d^m(B_1^N, \ell_p^N) \leq E^m(B_1^N, \ell_p^N) \leq \frac{C}{s^{1-1/p}},$$

in other words

$$c' \min\left\{1, \frac{\ln(eN/m)}{m}\right\} \leq \frac{1}{s}.$$

We derive either $s \leq 1/c'$ or $m \geq c' s \ln(eN/m)$, in which case

$$m \geq c' s \ln\left(\frac{eN}{s}\right) + c' m \frac{s}{m} \ln\left(\frac{s}{m}\right) \geq -1/e.$$
Gelfand Widths of $\ell_1$-Balls: Lower Bound

The Gelfand width of $B_1^N$ in $\ell_p^N$, $p > 1$, also satisfies

$$d^m(B_1^N, \ell_p^N) \geq c \min \left\{ 1, \frac{\ln(eN/m)}{m} \right\}^{1-1/p}.$$ 

Once established, this will show the optimality of the CS results. Indeed, suppose the existence of $(A, \Delta)$ such that

$$\|x - \Delta(Ax)\|_p \leq \frac{C}{s^{1-1/p}} \sigma_s(x)_1 \leq \frac{C}{s^{1-1/p}} \|x\|_1 \quad \text{for all } x \in \mathbb{R}^N.$$

Then,

$$c \min \left\{ 1, \frac{\ln(eN/m)}{m} \right\}^{1-1/p} \leq d^m(B_1^N, \ell_p^N) \leq E^m(B_1^N, \ell_p^N) \leq \frac{C}{s^{1-1/p}},$$

in other words

$$c' \min \left\{ 1, \frac{\ln(eN/m)}{m} \right\} \leq \frac{1}{s}.$$

We derive either $s \leq 1/c'$ or $m \geq c' s \ln(eN/m)$, in which case

$$m \geq c' s \ln \left( \frac{eN}{s} \right) + c' m \frac{s}{m} \ln \left( \frac{s}{m} \right) \geq \cdots \geq -1/e.$$
The Gelfand width of $B_1^N$ in $\ell_p^N$, $p > 1$, also satisfies

$$d^m(B_1^N, \ell_p^N) \geq c \min \left\{ 1, \frac{\ln(eN/m)}{m} \right\}^{1 - 1/p}.$$ 

Once established, this will show the optimality of the CS results. Indeed, suppose the existence of $(A, \Delta)$ such that

$$\|x - \Delta(Ax)\|_p \leq \frac{C}{s^{1 - 1/p}} \sigma_s(x)_1 \leq \frac{C}{s^{1 - 1/p}} \|x\|_1$$

for all $x \in \mathbb{R}^N$. Then,

$$c \min \left\{ 1, \frac{\ln(eN/m)}{m} \right\}^{1 - 1/p} \leq d^m(B_1^N, \ell_p^N) \leq E^m(B_1^N, \ell_p^N) \leq \frac{C}{s^{1 - 1/p}},$$

in other words

$$c' \min \left\{ 1, \frac{\ln(eN/m)}{m} \right\} \leq \frac{1}{s}.$$ 

We derive either $s \leq 1/c'$ or $m \geq c' s \ln(eN/m)$, in which case

$$m \geq c' s \ln \left( \frac{eN}{s} \right) + c' m \frac{s}{m} \ln \left( \frac{s}{m} \right) \geq \cdots \geq c'' s \ln \left( \frac{eN}{s} \right).$$
The lower estimate for $d^m(B_1^N, \ell_p^N)$: two key insights
The lower estimate for $d^m(B_1^N, \ell_p^N)$: two key insights

- small width implies $\ell_1$-recovery of $s$-sparse vectors for large $s$. 
The lower estimate for $d^m(B^N_1, \ell^N_p)$: two key insights

- small width implies $\ell_1$-recovery of $s$-sparse vectors for large $s$.

- $\ell_1$-recovery of $s$-sparse vectors only possible for moderate $s$. 
The lower estimate for $d^m(B_1^N, \ell^N_p)$: two key insights

- small width implies $\ell_1$-recovery of $s$-sparse vectors for large $s$.

There is a matrix $A \in \mathbb{R}^{m \times N}$ such that every $s$-sparse $x \in \mathbb{R}^N$ is a minimizer of $\|z\|_1$ subject to $Az = Ax$ for

$$s \approx \left( \frac{1}{2d^m(B_1^N, \ell^N_p)} \right)^{\frac{p}{p-1}}.$$

- $\ell_1$-recovery of $s$-sparse vectors only possible for moderate $s$. 
The lower estimate for $d^m(B^N_1, \ell^N_p)$: two key insights

- small width implies $\ell_1$-recovery of $s$-sparse vectors for large $s$.

There is a matrix $A \in \mathbb{R}^{m \times N}$ such that every $s$-sparse $x \in \mathbb{R}^N$ is a minimizer of $\|z\|_1$ subject to $Az = Ax$ for

$$s \approx \left( \frac{1}{2d^m(B^N_1, \ell^N_p)} \right)^{\frac{p}{p-1}}.$$

- $\ell_1$-recovery of $s$-sparse vectors only possible for moderate $s$.

For $s \geq 2$, if $A \in \mathbb{R}^{m \times N}$ is a matrix such that every $s$-sparse vector $x$ is a minimizer of $\|z\|_1$ subject to $Az = Ax$, then

$$m \geq c_1 s \ln \left( \frac{N}{c_2 s} \right),$$
The lower estimate for $d^m(B_1^N, \ell_p^N)$: two key insights

- small width implies $\ell_1$-recovery of $s$-sparse vectors for large $s$.

There is a matrix $A \in \mathbb{R}^{m \times N}$ such that every $s$-sparse $x \in \mathbb{R}^N$ is a minimizer of $\|z\|_1$ subject to $Az = Ax$ for

$$s \approx \left( \frac{1}{2d^m(B_1^N, \ell_p^N)} \right)^{\frac{p}{p-1}}.$$

- $\ell_1$-recovery of $s$-sparse vectors only possible for moderate $s$.

For $s \geq 2$, if $A \in \mathbb{R}^{m \times N}$ is a matrix such that every $s$-sparse vector $x$ is a minimizer of $\|z\|_1$ subject to $Az = Ax$, then

$$m \geq c_1 s \ln \left( \frac{N}{c_2 s} \right), \quad c_1 \geq 0.45, \quad c_2 = 4.$$
Deriving the lower estimate*

Suppose that

\[ d_m(B_N, \ell_N) < (c\mu)^{1 - 1/p/2}, \]

where

\[ \mu := \min\{1, \ln(eN/m)\} \leq 1. \]

Setting

\[ s \approx 1/(c\mu) \geq 2, \]

there exists

\[ A \in \mathbb{R}^{m \times N} \]

allowing \( \ell_1 \)-recovery of all \( s \)-sparse vectors.

Therefore

\[ m \geq c_{1s} \ln(N c_{2s}) \geq c_{1s} \ln(N c_{2m}) \geq c_{1c} \ln(eN/m) \min\{1, \ln(eN/m)\} \geq c_{1c} \ln(eN/m) m, \]

provided \( c \) is chosen small enough.

This is a contradiction.
Deriving the lower estimate*

Suppose that $d^m(B_1^N, \ell_p^N) < (c\mu)^{1-1/p}/2$, where

$$\mu := \min \left\{ 1, \frac{\ln(eN/m)}{m} \right\} \leq 1.$$
Deriving the lower estimate*

Suppose that \( d^m(B_1^N, \ell_p^N) < (c\mu)^{1-1/p}/2 \), where

\[
\mu := \min \left\{ 1, \frac{\ln(eN/m)}{m} \right\} \leq 1.
\]

Setting \( s \approx 1/(c\mu) \geq 2 \),
Deriving the lower estimate*

Suppose that $d^m(B_1^N, \ell_p^N) < (c\mu)^{1-1/p}/2$, where

$$\mu := \min \left\{ 1, \frac{\ln(eN/m)}{m} \right\} \leq 1.$$

Setting $s \approx 1/(c\mu) \geq 2$, there exists $A \in \mathbb{R}^{m \times N}$ allowing $\ell_1$-recovery of all $s$-sparse vectors.

*This is a contradiction.
Deriving the lower estimate*

Suppose that $d^m(B_1^N, \ell_p^N) < (c\mu)^{1-1/p}/2$, where

$$\mu := \min \left\{ 1, \frac{\ln(eN/m)}{m} \right\} \leq 1.$$

Setting $s \approx 1/(c\mu) \geq 2$, there exists $A \in \mathbb{R}^{m \times N}$ allowing \ell_1-recovery of all $s$-sparse vectors. Therefore

$$m \geq c_1 s \ln \left( \frac{N}{c_2 s} \right)$$
Deriving the lower estimate*

Suppose that $d^m(B_1^N, \ell_p^N) < (c\mu)^{1-1/p}/2$, where

$$\mu := \min \left\{ 1, \frac{\ln(eN/m)}{m} \right\} \leq 1.$$  

Setting $s \approx 1/(c\mu) \geq 2$, there exists $A \in \mathbb{R}^{m \times N}$ allowing $\ell_1$-recovery of all $s$-sparse vectors. Therefore

$$m \geq c_1 s \ln \left( \frac{N}{c_2 s} \right) \geq c_1 s \ln \left( \frac{N}{c'_2 m} \right)$$
Deriving the lower estimate*

Suppose that \(d^m(B_1^N, \ell_p^N) < (c\mu)^{1-1/p}/2\), where

\[
\mu := \min \left\{1, \frac{\ln(eN/m)}{m} \right\} \leq 1.
\]

Setting \(s \approx 1/(c\mu) \geq 2\), there exists \(A \in \mathbb{R}^{m \times N}\) allowing \(\ell_1\)-recovery of all \(s\)-sparse vectors. Therefore

\[
m \geq c_1 s \ln \left(\frac{N}{c_2 s} \right) \geq c_1 s \ln \left(\frac{N}{c'_2 m} \right) \geq c'_1 s \ln \left(\frac{eN}{m} \right)
\]
Deriving the lower estimate*

Suppose that $d^m(B_1^N, \ell_p^N) < (c\mu)^{1-1/p}/2$, where

$$\mu := \min \left\{ 1, \frac{\ln(eN/m)}{m} \right\} \leq 1.$$ 

Setting $s \approx 1/(c\mu) \geq 2$, there exists $A \in \mathbb{R}^{m \times N}$ allowing $\ell_1$-recovery of all $s$-sparse vectors. Therefore

$$m \geq c_1 s \ln\left( \frac{N}{c_2 s} \right) \geq c_1 s \ln\left( \frac{N}{c'_2 m} \right) \geq c'_1 s \ln\left( \frac{eN}{m} \right)$$

$$\geq \frac{c'_1}{c} \min \left\{ 1, \frac{\ln(eN/m)}{m} \right\} \leq 1.$$
Deriving the lower estimate*

Suppose that $d^m(B_1^N, \ell_p^N) < (c\mu)^{1-1/p}/2$, where

$$\mu := \min \left\{ 1, \frac{\ln(eN/m)}{m} \right\} \leq 1.$$  

Setting $s \approx 1/(c\mu) \geq 2$, there exists $A \in \mathbb{R}^{m \times N}$ allowing $\ell_1$-recovery of all $s$-sparse vectors. Therefore

$$m \geq c_1 s \ln \left( \frac{N}{c_2 s} \right) \geq c_1 s \ln \left( \frac{N}{c'_2 m} \right) \geq c'_1 s \ln \left( \frac{eN}{m} \right) \geq c'_1 \frac{\ln(eN/m)}{m} \geq c'_1 m,$$

provided $c$ is chosen small enough. This is a contradiction.
Suppose that $d^m(B^N_1, \ell^N_p) < (c\mu)^{1-1/p}/2$, where

$$\mu := \min \left\{ 1, \frac{\ln(eN/m)}{m} \right\} \leq 1.$$ 

Setting $s \approx 1/(c\mu) \geq 2$, there exists $A \in \mathbb{R}^{m \times N}$ allowing $\ell_1$-recovery of all $s$-sparse vectors. Therefore

$$m \geq c_1 s \ln \left( \frac{N}{c_2 s} \right) \geq c_1 s \ln \left( \frac{N}{c'_2 m} \right) \geq c'_1 s \ln \left( \frac{eN}{m} \right) \geq \frac{c'_1}{c} \frac{\ln(eN/m)}{m} \geq \frac{c'_1}{c} m > m,$$

provided $c$ is chosen small enough.
Deriving the lower estimate

Suppose that $d^m(B_1^N, \ell_p^N) < (c\mu)^{1-1/p}/2$, where

$$\mu := \min \left\{ 1, \frac{\ln(eN/m)}{m} \right\} \leq 1.$$ 

Setting $s \approx 1/(c\mu) \geq 2$, there exists $A \in \mathbb{R}^{m \times N}$ allowing $\ell_1$-recovery of all $s$-sparse vectors. Therefore

$$m \geq c_1 s \ln \left( \frac{N}{c_2 s} \right) \geq c_1 s \ln \left( \frac{N}{c'_2 m} \right) \geq c'_1 s \ln \left( \frac{eN}{m} \right) \geq \frac{c'_1}{c} \ln(eN/m) \min \left\{ 1, \frac{\ln(eN/m)}{m} \right\} \geq \frac{c'_1}{c} m > m,$$

provided $c$ is chosen small enough. This is a contradiction.
Insight 1: width and $\ell_1$-recovery*

Every $s$-sparse $x \in \mathbb{R}^N$ is a minimizer of $\|z\|_1$ subject to $A z = A x$ if and only if the null space property of order $s$ holds, i.e.,

$$\|v_S\|_1 \leq \|v\|_1 / 2$$

for all $v \in \ker A$ and all $S \in \{N\}$ with $|S| \leq s$.

Setting $d := d_m(B_N^{1/p}, \ell_N^p)$, there exists $A \in \mathbb{R}^{m \times N}$ such that $\|v\|_p \leq d \|v\|_1$ for all $v \in \ker A$.

Then, for $v \in \ker A$ and $S \in \{N\}$ with $|S| \leq s$,

$$\|v_S\|_1 \leq s_1 - 1/p \|v_S\|_p \leq d s_1 - 1/p \|v\|_1.$$

Choose $s \approx (\frac{1}{2} d)^{p/(p-1)}$ to derive the null space property of order $s$. 
Insight 1: width and $\ell_1$-recovery*

Every $s$-sparse $x \in \mathbb{R}^N$ is a minimizer of $\|z\|_1$ subject to $Az = Ax$ if and only if the null space property of order $s$ holds,
Insight 1: width and $\ell_1$-recovery*

Every $s$-sparse $x \in \mathbb{R}^N$ is a minimizer of $\|z\|_1$ subject to $Az = Ax$ if and only if the null space property of order $s$ holds, i.e.,

$$\|v_S\|_1 \leq \|v_{\overline{S}}\|_1$$

for all $v \in \ker A$ and all $S \in [N]$ with $|S| \leq s$. 

Choose $s \approx \left(\frac{1}{2}d_m(B_{\ell_1^N}, \ell_1^p)\right)^{p/(p-1)}$ to derive the null space property of order $s$. 
Insight 1: width and $\ell_1$-recovery*

Every $s$-sparse $x \in \mathbb{R}^N$ is a minimizer of $\|z\|_1$ subject to $Az = Ax$ if and only if the null space property of order $s$ holds, i.e.,

$$\|v_S\|_1 \leq \|v\|_1/2 \quad \text{for all } v \in \ker A \text{ and all } S \in [N] \text{ with } |S| \leq s.$$
Insight 1: width and $\ell_1$-recovery*

Every $s$-sparse $x \in \mathbb{R}^N$ is a minimizer of $\|z\|_1$ subject to $Az = Ax$ if and only if the null space property of order $s$ holds, i.e.,

$$\|v_S\|_1 \leq \|v\|_1/2 \quad \text{for all } v \in \text{ker } A \text{ and all } S \in [N] \text{ with } |S| \leq s.$$ 

Setting $d := d^m(B_1^N, \ell_p^N)$, there exists $A \in \mathbb{R}^{m \times N}$ such that

$$\|v\|_p \leq d\|v\|_1 \quad \text{for all } v \in \ker A.$$
Insight 1: width and $\ell_1$-recovery*

Every $s$-sparse $x \in \mathbb{R}^N$ is a minimizer of $\|z\|_1$ subject to $Az = Ax$ if and only if the null space property of order $s$ holds, i.e.,

$$\|v_S\|_1 \leq \|v\|_1/2 \quad \text{for all } v \in \ker A \text{ and all } S \in [N] \text{ with } |S| \leq s.$$  

Setting $d := d_m(B_1^N, \ell_p^N)$, there exists $A \in \mathbb{R}^{m \times N}$ such that

$$\|v\|_p \leq d\|v\|_1 \quad \text{for all } v \in \ker A.$$  

Then, for $v \in \ker A$ and $S \in [N]$ with $|S| \leq s$,
Insight 1: width and $\ell_1$-recovery

Every $s$-sparse $x \in \mathbb{R}^N$ is a minimizer of $\|z\|_1$ subject to $Az = Ax$ if and only if the null space property of order $s$ holds, i.e.,

$$\|v_S\|_1 \leq \|v\|_1/2 \quad \text{for all } v \in \ker A \text{ and all } S \in [N] \text{ with } |S| \leq s.$$  

Setting $d := d_m(B_1^N, \ell_p^N)$, there exists $A \in \mathbb{R}^{m \times N}$ such that

$$\|v\|_p \leq d\|v\|_1 \quad \text{for all } v \in \ker A.$$  

Then, for $v \in \ker A$ and $S \in [N]$ with $|S| \leq s$,

$$\|v_S\|_1 \leq s^{1-1/p}\|v_S\|_p$$
Insight 1: width and $\ell_1$-recovery*

Every $s$-sparse $x \in \mathbb{R}^N$ is a minimizer of $\|z\|_1$ subject to $Az = Ax$ if and only if the null space property of order $s$ holds, i.e.,

$$\|v_S\|_1 \leq \|v\|_1/2 \quad \text{for all } v \in \ker A \text{ and all } S \in [N] \text{ with } |S| \leq s.$$ 

Setting $d := d_m(B_1^N, \ell_p^N)$, there exists $A \in \mathbb{R}^{m \times N}$ such that

$$\|v\|_p \leq d\|v\|_1 \quad \text{for all } v \in \ker A.$$ 

Then, for $v \in \ker A$ and $S \in [N]$ with $|S| \leq s$,

$$\|v_S\|_1 \leq s^{1-1/p}\|v_S\|_p \leq s^{1-1/p}\|v\|_p.$$
Insight 1: width and $\ell_1$-recovery*

Every $s$-sparse $\mathbf{x} \in \mathbb{R}^N$ is a minimizer of $\|\mathbf{z}\|_1$ subject to $A\mathbf{z} = A\mathbf{x}$ if and only if the null space property of order $s$ holds, i.e.,

$$\|\mathbf{v}_S\|_1 \leq \|\mathbf{v}\|_1/2 \quad \text{for all } \mathbf{v} \in \ker A \text{ and all } S \in [N] \text{ with } |S| \leq s.$$  

Setting $d := d_m(B_1^N, \ell_p^N)$, there exists $A \in \mathbb{R}^{m \times N}$ such that

$$\|\mathbf{v}\|_p \leq d\|\mathbf{v}\|_1 \quad \text{for all } \mathbf{v} \in \ker A.$$  

Then, for $\mathbf{v} \in \ker A$ and $S \in [N]$ with $|S| \leq s$,

$$\|\mathbf{v}_S\|_1 \leq s^{1-1/p}\|\mathbf{v}_S\|_p \leq s^{1-1/p}\|\mathbf{v}\|_p \leq ds^{1-1/p}\|\mathbf{v}\|_1.$$
Insight 1: width and $\ell_1$-recovery*

Every $s$-sparse $x \in \mathbb{R}^N$ is a minimizer of $\|z\|_1$ subject to $Az = Ax$ if and only if the null space property of order $s$ holds, i.e.,

$$\|v_S\|_1 \leq \|v\|_1/2 \quad \text{for all } v \in \ker A \text{ and all } S \in [N] \text{ with } |S| \leq s.$$ 

Setting $d := d^m(B_1^N, \ell_p^N)$, there exists $A \in \mathbb{R}^{m \times N}$ such that

$$\|v\|_p \leq d\|v\|_1 \quad \text{for all } v \in \ker A.$$

Then, for $v \in \ker A$ and $S \in [N]$ with $|S| \leq s$,

$$\|v_S\|_1 \leq s^{1-1/p}\|v_S\|_p \leq s^{1-1/p}\|v\|_p \leq ds^{1-1/p}\|v\|_1.$$

Choose $s \approx \left( \frac{1}{2d} \right)^\frac{p}{p-1}$.
Insight 1: width and $\ell_1$-recovery*

Every $s$-sparse $x \in \mathbb{R}^N$ is a minimizer of $\|z\|_1$ subject to $Az = Ax$ if and only if the null space property of order $s$ holds, i.e.,

$$\|v_S\|_1 \leq \|v\|_1/2 \quad \text{for all } v \in \ker A \text{ and all } S \in [N] \text{ with } |S| \leq s.$$ 

Setting $d := d^m(B_1^N, \ell_p^N)$, there exists $A \in \mathbb{R}^{m \times N}$ such that

$$\|v\|_p \leq d\|v\|_1 \quad \text{for all } v \in \ker A.$$ 

Then, for $v \in \ker A$ and $S \in [N]$ with $|S| \leq s$,

$$\|v_S\|_1 \leq s^{1-1/p}\|v_S\|_p \leq s^{1-1/p}\|v\|_p \leq ds^{1-1/p}\|v\|_1.$$ 

Choose $s \approx \left(\frac{1}{2d}\right)^\frac{p}{p-1}$ to derive the null space property of order $s$. 

*Note: The asterisk indicates a special or variant insight or result.
A combinatorial lemma

There exists $n \geq \binom{N}{4}s^2$ subsets $S_1, \ldots, S_n$ of size $s$ such that $|S_j \cap S_k| < s^2$ for all $1 \leq j \neq k \leq n$.

Define $s$-sparse vectors $x_1, \ldots, x_n$ by $(x_j)_i = \begin{cases} 1/s & \text{if } i \in S_j, \\ 0 & \text{if } i \not\in S_j. \end{cases}$

Note that $\|x_j\|_1 = 1$ and $\|x_j - x_k\|_1 > 1$ for all $1 \leq j \neq k \leq n$.  


Lemma

There exists $n \geq \left( \frac{N}{4s} \right)^{\frac{s}{2}}$ subsets $S^1, \ldots, S^n$ of size $s$ such that

$$|S^j \cap S^k| < \frac{s}{2} \quad \text{for all } 1 \leq j \neq k \leq n.$$
A combinatorial lemma

Lemma

There exists $n \geq \left( \frac{N}{4s} \right)^{\frac{s}{2}}$ subsets $S^1, \ldots, S^n$ of size $s$ such that

$$|S^j \cap S^k| < \frac{s}{2} \quad \text{for all } 1 \leq j \neq k \leq n.$$ 

Define $s$-sparse vectors $x^1, \ldots, x^n$ by

$$(x^j)_i = \begin{cases} 1/s & \text{if } i \in S^j, \\ 0 & \text{if } i \notin S^j. \end{cases}$$
Lemma

There exists \( n \geq \left( \frac{N}{4s} \right)^{\frac{s}{2}} \) subsets \( S^1, \ldots, S^n \) of size \( s \) such that

\[
|S^j \cap S^k| < \frac{s}{2} \quad \text{for all } 1 \leq j \neq k \leq n.
\]

Define \( s \)-sparse vectors \( \mathbf{x}^1, \ldots, \mathbf{x}^n \) by

\[
(x^j)_i = \begin{cases} 
1/s & \text{if } i \in S^j, \\
0 & \text{if } i \not\in S^j.
\end{cases}
\]

Note that \( \|\mathbf{x}^j\|_1 = 1 \)
A combinatorial lemma

Lemma

There exists \( n \geq \left( \frac{N}{4s} \right)^{\frac{s}{2}} \) subsets \( S^1, \ldots, S^n \) of size \( s \) such that

\[ |S^j \cap S^k| < \frac{s}{2} \quad \text{for all } 1 \leq j \neq k \leq n. \]

Define \( s \)-sparse vectors \( \mathbf{x}^1, \ldots, \mathbf{x}^n \) by

\[ (\mathbf{x}^j)_i = \begin{cases} 1/s & \text{if } i \in S^j, \\ 0 & \text{if } i \notin S^j. \end{cases} \]

Note that \( \|\mathbf{x}^j\|_1 = 1 \) and \( \|\mathbf{x}^j - \mathbf{x}^k\|_1 > 1 \) for all \( 1 \leq j \neq k \leq n \).
Insight 2: $\ell_1$-recovery and number of measurements*
Insight 2: $\ell_1$-recovery and number of measurements*

For $A \in \mathbb{R}^{m \times N}$, suppose that every $2s$-sparse vector $x \in \mathbb{R}^N$ is a minimizer of $\|z\|_1$ subject to $Az = Ax$. 

In the quotient space $\ell^N_1/\ker A$, this means $\|x\| := \inf_{v \in \ker A} \|x - v\|_1 = \|x\|_1$ for all $2s$-sparse $x \in \mathbb{R}^N$.

In particular, $\|x_j\| = 1$ and $\|x_j - x_k\| > 1$, all $1 \leq j \neq k \leq n$.

The size of this $1$-separating set of the unit sphere satisfies $\left(\frac{N}{4s}\right)^s \leq n \leq \left(1 + 2\frac{1}{s}\right)m = 3m$.

Taking the logarithm yields $m \geq s \ln 9 \ln \left(\frac{N}{4s}\right)$. 
Insight 2: $\ell_1$-recovery and number of measurements

For $A \in \mathbb{R}^{m \times N}$, suppose that every $2s$-sparse vector $x \in \mathbb{R}^N$ is a minimizer of $\|z\|_1$ subject to $Az = Ax$. In the quotient space $\ell_1^N / \ker A$, this means

$$\|[x]\| := \inf_{v \in \ker A} \|x - v\|_1$$
Insight 2: $\ell_1$-recovery and number of measurements*

For $A \in \mathbb{R}^{m \times N}$, suppose that every $2s$-sparse vector $x \in \mathbb{R}^N$ is a minimizer of $\|z\|_1$ subject to $Az = Ax$.
In the quotient space $\ell_1^N / \ker A$, this means

$$\|[x]\| := \inf_{v \in \ker A} \|x - v\|_1 = \|x\|_1 \quad \text{for all 2s-sparse } x \in \mathbb{R}^N.$$
Insight 2: $\ell_1$-recovery and number of measurements*

For $A \in \mathbb{R}^{m \times N}$, suppose that every $2s$-sparse vector $x \in \mathbb{R}^N$ is a minimizer of $\|z\|_1$ subject to $Az = Ax$.

In the quotient space $\ell_1^N / \ker A$, this means

$$\|[x]\| := \inf_{v \in \ker A} \|x - v\|_1 = \|x\|_1$$

for all $2s$-sparse $x \in \mathbb{R}^N$.

In particular,

$$\|[x^j]\| = 1$$
Insight 2: $\ell_1$-recovery and number of measurements

For $A \in \mathbb{R}^{m \times N}$, suppose that every $2s$-sparse vector $x \in \mathbb{R}^N$ is a minimizer of $\|z\|_1$ subject to $Az = Ax$.

In the quotient space $\ell_1^N / \ker A$, this means

$$\|[x]\| := \inf_{v \in \ker A} \|x - v\|_1 = \|x\|_1 \quad \text{for all 2s-sparse } x \in \mathbb{R}^N.$$ 

In particular,

$$\|[x^j]\| = 1 \quad \text{and} \quad \|[x^j] - [x^k]\| > 1, \quad \text{all } 1 \leq j \neq k \leq n.$$
Insight 2: $\ell_1$-recovery and number of measurements*

For $A \in \mathbb{R}^{m \times N}$, suppose that every $2s$-sparse vector $x \in \mathbb{R}^N$ is a minimizer of $\|z\|_1$ subject to $Az = Ax$.

In the quotient space $\ell_1^N / \ker A$, this means

$$\| [x] \| := \inf_{v \in \ker A} \| x - v \|_1 = \| x \|_1 \quad \text{for all } 2s\text{-sparse } x \in \mathbb{R}^N.$$  

In particular,

$$\| [x^j] \| = 1 \quad \text{and} \quad \| [x^j] - [x^k] \| > 1, \quad \text{all } 1 \leq j \neq k \leq n.$$  

The size of this 1-separating set of the unit sphere satisfies

$$n \leq \left( 1 + \frac{2}{1} \right)^m.$$
Insight 2: $\ell_1$-recovery and number of measurements*

For $A \in \mathbb{R}^{m \times N}$, suppose that every $2s$-sparse vector $x \in \mathbb{R}^N$ is a minimizer of $\|z\|_1$ subject to $Az = Ax$. In the quotient space $\ell_1^N / \ker A$, this means

$$\|[x]\| \defeq \inf_{v \in \ker A} \|x - v\|_1 = \|x\|_1 \quad \text{for all } 2s\text{-sparse } x \in \mathbb{R}^N.$$ 

In particular,

$$\| [x^j] \| = 1 \quad \text{and} \quad \| [x^j] - [x^k] \| > 1, \quad \text{all } 1 \leq j \neq k \leq n.$$ 

The size of this 1-separating set of the unit sphere satisfies

$$n \leq \left( 1 + \frac{2}{1} \right)^m = 3^m.$$
Insight 2: $\ell_1$-recovery and number of measurements*

For $A \in \mathbb{R}^{m \times N}$, suppose that every $2s$-sparse vector $x \in \mathbb{R}^N$ is a minimizer of $\|z\|_1$ subject to $Az = Ax$.

In the quotient space $\ell_1^N / \ker A$, this means

$$\| [x] \| := \inf_{v \in \ker A} \| x - v \|_1 = \| x \|_1$$

for all $2s$-sparse $x \in \mathbb{R}^N$.

In particular,

$$\| [x^j] \| = 1 \quad \text{and} \quad \| [x^j] - [x^k] \| > 1, \quad \text{all } 1 \leq j \neq k \leq n.$$

The size of this $1$-separating set of the unit sphere satisfies

$$\left( \frac{N}{4s} \right)^{\frac{s}{2}} \leq n \leq \left( 1 + \frac{2}{1} \right)^m = 3^m.$$
Insight 2: $\ell_1$-recovery and number of measurements*

For $A \in \mathbb{R}^{m \times N}$, suppose that every $2s$-sparse vector $x \in \mathbb{R}^N$ is a minimizer of $\|z\|_1$ subject to $Az = Ax$.

In the quotient space $\ell_1^N / \ker A$, this means

$$\| [x] \| := \inf_{v \in \ker A} \| x - v \|_1 = \| x \|_1 \quad \text{for all } 2s\text{-sparse } x \in \mathbb{R}^N.$$ 

In particular,

$$\| [x^j] \| = 1 \quad \text{and} \quad \| [x^j] - [x^k] \| > 1, \quad \text{all } 1 \leq j \neq k \leq n.$$ 

The size of this 1-separating set of the unit sphere satisfies

$$\left( \frac{N}{4s} \right)^{\frac{s}{2}} \leq n \leq \left( 1 + \frac{2}{1} \right)^m = 3^m.$$ 

Taking the logarithm yields

$$m \geq \frac{s}{\ln 9} \ln \left( \frac{N}{4s} \right).$$
Uniform vs. Nonuniform Guarantees*
Empirical Performances: HTP and $\ell_1$-Minimization
Empirical Performances: HTP and $\ell_1$-Minimization

Gaussian matrices with Rademacher then with Gaussian vectors
Empirical Performances: HTP and $\ell_1$-Minimization

Gaussian matrices with Rademacher then with Gaussian vectors

![Graph 1](image1)

N=1000, s=20

Successful reconstructions (%)

Number of measurements (m)

Basis pursuit
Hard thresholding pursuit

![Graph 2](image2)

N=1000, m=200

Successful reconstructions (%)

Sparsity (s)
Empirical Performances: Phase Transitions*

Figure: L: empirically observed weak threshold
R: strong (dashed) and weak (solid) thresholds (courtesy of J. Tanner)
Uniform vs. Nonuniform Sparse Recovery

When $A$ is a random matrix, uniform results read

$$\Pr(\forall s\text{-sparse } x, \text{recovery of } x \text{ is successful using } A) \geq 1 - \varepsilon,$$

while nonuniform results have the form

$$\forall s\text{-sparse } x, \Pr(\text{recovery of } x \text{ is successful using } A) \geq 1 - \varepsilon.$$

Using Gaussian matrices, nonuniform $s$-sparse recovery results hold for exactly $s$ iterations of OMP in the regime $m \geq cs \ln(N/s)$, but uniform results do not hold in this regime.
Uniform vs. Nonuniform Sparse Recovery

When $A$ is a random matrix, uniform results read

$$\mathbb{P}(\forall s\text{-sparse } \mathbf{x}, \text{ recovery of } \mathbf{x} \text{ is successful using } A) \geq 1 - \varepsilon,$$

while nonuniform results have the form

$$\forall s\text{-sparse } \mathbf{x}, \mathbb{P}(\text{recovery of } \mathbf{x} \text{ is successful using } A) \geq 1 - \varepsilon.$$
When $A$ is a random matrix, uniform results read
\[
\mathbb{P}(\forall s\text{-sparse } x, \text{ recovery of } x \text{ is successful using } A) \geq 1 - \varepsilon,
\]
while nonuniform results have the form
\[
\forall s\text{-sparse } x, \mathbb{P}(\text{recovery of } x \text{ is successful using } A) \geq 1 - \varepsilon.
\]
Uniform vs. Nonuniform Sparse Recovery

When $A$ is a random matrix, uniform results read

$$\mathbb{P}(\forall s\text{-sparse } x, \text{ recovery of } x \text{ is successful using } A) \geq 1 - \varepsilon,$$

while nonuniform results have the form

$$\forall s\text{-sparse } x, \mathbb{P}(\text{recovery of } x \text{ is successful using } A) \geq 1 - \varepsilon.$$

Using Gaussian matrices, nonuniform $s$-sparse recovery results hold for exactly $s$ iterations of OMP in the regime $m \geq cs \ln(N/s)$, but uniform results do not hold in this regime.
Nonuniform Recovery via BP when \( m \gtrsim 2s \ln(N/s) \)

A necessary and sufficient condition for the recovery of a fixed \( x \in \mathbb{C}^N \) supported on \( S \) via \( \ell_1 \)-minimization is

\[ \left| \sum_{j \in S} \text{sgn}(x_j) v_j \right| < \| v_S \|_1 \text{ for all } v \neq 0 \in \ker A. \]

This is implied by (in the real case, equivalent to) the injectivity of \( A_S \) and the existence of \( h \in \mathbb{C}^m \) such that

\[ (A^* h)_j = \text{sgn}(x_j), \quad j \in S, \quad \left\| (A^* h)_\ell \right\| < 1, \quad \ell \in S. \]

The latter is in turn implied by

\[ \left| \langle A^\dagger_S a_\ell, \text{sgn}(x_S) \rangle \right| < 1 \text{ for all } \ell \in S. \]

For subgaussian matrices, this occurs with probability \( \geq 1 - \varepsilon \) provided

\[ m \gtrsim 2s \ln(2^N/s) \]

(which is replaced by

\[ m \gtrsim 2s (\sqrt{\ln(eN/s)} + \sqrt{\ln(1/\varepsilon)/s})^2 \]

for Gaussian matrices using Gordon's escape through the mesh.)
Nonuniform Recovery via BP when $m \gtrsim 2s \ln(N/s)$

- A necessary and sufficient condition for the recovery of a fixed $x \in \mathbb{C}^N$ supported on $S$ via $\ell_1$-minimization is

$$\left| \sum_{j \in S} \text{sgn}(x_j)v_j \right| < \|v_S\|_1$$

for all $v \neq 0 \in \ker A$. 

For subgaussian matrices, this occurs with probability $\geq 1 - \varepsilon$ provided $m \gtrsim 2s \ln(2N/\varepsilon)$ (which is replaced by $m \gtrsim 2s(\sqrt{\ln(eN/s)} + \sqrt{\ln(1/\varepsilon)/s})^2$ for Gaussian matrices using Gordon's escape through the mesh).
Nonuniform Recovery via BP when \( m \gtrsim 2s \ln(N/s) \)*

- A necessary and sufficient condition for the recovery of a fixed \( x \in \mathbb{C}^N \) supported on \( S \) via \( \ell_1 \)-minimization is
  \[
  \left| \sum_{j \in S} \text{sgn}(x_j) v_j \right| < \| v_{\bar{S}} \|_1 \quad \text{for all } v \neq 0 \in \text{ker } A.
  \]

- This is implied by (in the real case, equivalent to) the injectivity of \( A_S \) and the existence of \( h \in \mathbb{C}^m \) such that
  \[
  (A^*h)_j = \text{sgn}(x_j), \quad j \in S, \quad |(A^*h)_\ell| < 1, \quad \ell \in \bar{S}.
  \]
Nonuniform Recovery via BP when $m \gtrsim 2s \ln(N/s)^*$

- A necessary and sufficient condition for the recovery of a fixed $x \in \mathbb{C}^N$ supported on $S$ via $\ell_1$-minimization is
  \[
  \left| \sum_{j \in S} \text{sgn}(x_j) v_j \right| < \|v_S\|_1 \quad \text{for all } v \neq 0 \in \ker A.
  \]

- This is implied by (in the real case, equivalent to) the injectivity of $A_S$ and the existence of $h \in \mathbb{C}^m$ such that
  \[
  (A^* h)_j = \text{sgn}(x_j), \quad j \in S, \quad \left| (A^* h)_\ell \right| < 1, \quad \ell \in \overline{S}.
  \]

- The latter is in turn implied by
  \[
  \left| \langle A_S^\dagger a_\ell, \text{sgn}(x_S) \rangle \right| < 1 \quad \text{for all } \ell \in \overline{S}.
  \]
Nonuniform Recovery via BP when $m \gtrsim 2s \ln(N/s)^*$

- A necessary and sufficient condition for the recovery of a fixed $x \in \mathbb{C}^N$ supported on $S$ via $\ell_1$-minimization is

$$\left| \sum_{j \in S} \text{sgn}(x_j)v_j \right| < \|v_S\|_1 \quad \text{for all } v \neq 0 \in \ker A.$$

- This is implied by (in the real case, equivalent to) the injectivity of $A_S$ and the existence of $h \in \mathbb{C}^m$ such that

$$(A^*h)_j = \text{sgn}(x_j), \ j \in S, \quad |(A^*h)_\ell| < 1, \ \ell \in \overline{S}.$$ 

- The latter is in turn implied by

$$\left| \langle A^*_S a_\ell, \text{sgn}(x_S) \rangle \right| < 1 \quad \text{for all } \ell \in \overline{S}.$$ 

- For subgaussian matrices, this occurs with probability $\geq 1 - \varepsilon$ provided

$$m \gtrsim 2s \ln(2N/\varepsilon).$$
Nonuniform Recovery via BP when $m \gtrsim 2s \ln(N/s)^*$

- A necessary and sufficient condition for the recovery of a fixed $x \in \mathbb{C}^N$ supported on $S$ via $\ell_1$-minimization is
  \[
  \left| \sum_{j \in S} \text{sgn}(x_j)v_j \right| < \|v_S\|_1 \quad \text{for all } v \neq 0 \in \ker A.
  \]

- This is implied by (in the real case, equivalent to) the injectivity of $A_S$ and the existence of $h \in \mathbb{C}^m$ such that
  \[
  (A^*h)_j = \text{sgn}(x_j), \ j \in S, \quad |(A^*h)_\ell| < 1, \ \ell \in \overline{S}.
  \]

- The latter is in turn implied by
  \[
  |\langle A^*_S a_\ell, \text{sgn}(x_S) \rangle| < 1 \quad \text{for all } \ell \in \overline{S}.
  \]

- For subgaussian matrices, this occurs with probability $\geq 1 - \varepsilon$ provided
  \[
  m \gtrsim 2s \ln(2N/\varepsilon)
  \]
  (which is replaced by $m \gtrsim 2s \left( \sqrt{\ln(eN/s)} + \sqrt{\ln(1/\varepsilon)/s} \right)^2$ for Gaussian matrices using Gordon’s escape through the mesh).
Bounded Orthonormal Systems

Let $D \in \mathbb{R}^d$ be endowed with a probability measure $\nu$. A bounded orthonormal system (BOS) with constant $K \geq 1$ is a system $(\phi_1, ..., \phi_N)$ of function of $D$ satisfying

$$\int_D \phi_j(t) \phi_k(t) d\nu(t) = \delta_{j,k},$$

$$\sup_{1 \leq j \leq N} \sup_{t \in D} |\phi_j(t)| \leq K.$$

1. trigonometric polynomials: take $\phi_k(t) = e^{i 2\pi k t}$ for $t \in D = [0, 1]$, $d\nu(t) = dt$,
2. discrete orthonormal systems — let $U \in \mathbb{C}^{N \times N}$ be a unitary matrix with $\sqrt{N} |U_{k,j}| \leq K$ for all $k, j \in \{1, ..., N\}$ (e.g. Fourier or Hadamard matrix): take $\phi_k(t) = \sqrt{NU_{k,j}}$ for $t \in \{1, ..., N\}$, $\nu(B) = \text{card}(B) N$. 


Bounded Orthonormal Systems

Let $\mathcal{D} \in \mathbb{R}^d$ be endowed with a probability measure $\nu$. A bounded orthonormal system (BOS) with constant $K \geq 1$ is a system $(\phi_1, \ldots, \phi_N)$ of functions of $\mathcal{D}$ satisfying

$$\int_{\mathcal{D}} \phi_j(t)\overline{\phi_k(t)}d\nu(t) = \delta_{j,k},$$

$$\sup_{1 \leq j \leq N} \sup_{t \in \mathcal{D}} |\phi_j(t)| \leq K.$$
Bounded Orthonormal Systems

Let $\mathcal{D} \subseteq \mathbb{R}^d$ be endowed with a probability measure $\nu$. A bounded orthonormal system (BOS) with constant $K \geq 1$ is a system $(\phi_1, \ldots, \phi_N)$ of functions of $\mathcal{D}$ satisfying

$$\int_{\mathcal{D}} \phi_j(t)\overline{\phi_k(t)}d\nu(t) = \delta_{j,k},$$

$$\sup_{1 \leq j \leq N} \sup_{t \in \mathcal{D}} |\phi_j(t)| \leq K.$$

1. trigonometric polynomials:
   take $\phi_k(t) = e^{i2\pi kt}$ for $t \in \mathcal{D} = [0, 1]$, $d\nu(t) = dt$, 

2. discrete orthonormal systems — let $U \in \mathbb{C}^{N \times N}$ be a unitary matrix with $\sqrt{N}|U_{k,j}| \leq K$ for all $k, j \in \{1, \ldots, N\}$ (e.g. Fourier or Hadamard matrix): take $\phi_k(t) = \sqrt{N}U_{k,j}$ for $t \in \{1, \ldots, N\}$, $\nu(B) = \text{card}(B)$. 


Bounded Orthonormal Systems

Let $D \in \mathbb{R}^d$ be endowed with a probability measure $\nu$. A bounded orthonormal system (BOS) with constant $K \geq 1$ is a system $(\phi_1, \ldots, \phi_N)$ of function of $D$ satisfying

$$\int_D \phi_j(t) \overline{\phi_k(t)} d\nu(t) = \delta_{j,k},$$

$$\sup_{1 \leq j \leq N} \sup_{t \in D} |\phi_j(t)| \leq K.$$

1. trigonometric polynomials:
   take $\phi_k(t) = e^{i2\pi kt}$ for $t \in D = [0, 1]$, $d\nu(t) = dt$,

2. discrete orthonormal systems — let $U \in \mathbb{C}^{N \times N}$ be a unitary matrix with $\sqrt{N} |U_{k,j}| \leq K$ for all $k, j \in \{1, \ldots, N\}$ (e.g. Fourier or Hadamard matrix):
   take $\phi_k(t) = \sqrt{N} U_{k,t}$ for $t \in \{1, \ldots, N\}$, $\nu(B) = \frac{\text{card}(B)}{N}$. 
For \( f = \sum_{j=1}^{N} x_j \phi_j \), the samples of \( f \) at \( t_1, \ldots, t_m \in D \) are
\[
 f(t_k) = \sum_{j=1}^{N} x_j \phi_j(t_k) = (Ax)_k,
\]
where the sampling matrix \( A \in \mathbb{C}^{m \times N} \) have entries
\[
 A_{k,j} = \phi_j(t_k).
\]
Let an \( s \)-sparse \( x \in \mathbb{C}^N \) be fixed. Let sampling points \( t_1, \ldots, t_m \) be randomly selected according to \( \nu \), and let \( A \in \mathbb{C}^{m \times N} \) be the sampling matrix (3).

If \( m \geq C K^2 s \ln(N) \ln(\frac{1}{\epsilon}) \), then \( x \) is the unique minimizer of \( \| z \|_1 \) subject to \( Az = Ax \) with probability at least \( 1 - \epsilon \). (Proof based on the golfing scheme.)
Nonuniform Recovery for BOS*

For $f = \sum_{j=1}^{N} x_j \phi_j$, the samples of $f$ at $t_1, \ldots, t_m \in D$ are

$$f(t_k) = \sum_{j=1}^{N} x_j \phi_j(t_k) = (Ax)_k,$$

where the sampling matrix $A \in \mathbb{C}^{m \times N}$ have entries

(3)  \hspace{1cm} A_{k,j} = \phi_j(t_k).
Nonuniform Recovery for BOS*

For $f = \sum_{j=1}^{N} x_j \phi_j$, the samples of $f$ at $t_1, \ldots, t_m \in \mathcal{D}$ are

$$f(t_k) = \sum_{j=1}^{N} x_j \phi_j(t_k) = (Ax)_k,$$

where the sampling matrix $A \in \mathbb{C}^{m \times N}$ have entries

(3) \hspace{1cm} A_{k,j} = \phi_j(t_k).

Let an $s$-sparse $x \in \mathbb{C}^{N}$ be fixed.
Nonuniform Recovery for BOS*

For $f = \sum_{j=1}^{N} x_j \phi_j$, the samples of $f$ at $t_1, \ldots, t_m \in D$ are

$$f(t_k) = \sum_{j=1}^{N} x_j \phi_j(t_k) = (A\mathbf{x})_k,$$

where the sampling matrix $A \in \mathbb{C}^{m \times N}$ have entries

(3) $$A_{k,j} = \phi_j(t_k).$$

Let an $s$-sparse $\mathbf{x} \in \mathbb{C}^N$ be fixed. Let sampling points $t_1, \ldots, t_m$ be randomly selected according to $\nu$, and let $A \in \mathbb{C}^{m \times N}$ be the sampling matrix (3).
Nonuniform Recovery for BOS*

For \( f = \sum_{j=1}^{N} x_j \phi_j \), the samples of \( f \) at \( t_1, \ldots, t_m \in D \) are

\[
f(t_k) = \sum_{j=1}^{N} x_j \phi_j(t_k) = (Ax)_k,
\]

where the sampling matrix \( A \in \mathbb{C}^{m \times N} \) have entries

\[
A_{k,j} = \phi_j(t_k).
\]

Let an \( s \)-sparse \( x \in \mathbb{C}^{N} \) be fixed. Let sampling points \( t_1, \ldots, t_m \) be randomly selected according to \( \nu \), and let \( A \in \mathbb{C}^{m \times N} \) be the sampling matrix (3). If

\[
m \geq C K^2 s \ln(N) \ln(\varepsilon^{-1}),
\]

then \( x \) is the unique minimizer of \( \|z\|_1 \) subject to \( Az = Ax \) with probability at least \( 1 - \varepsilon \).
Nonuniform Recovery for BOS*

For \( f = \sum_{j=1}^{N} x_j \phi_j \), the samples of \( f \) at \( t_1, \ldots, t_m \in D \) are

\[
f(t_k) = \sum_{j=1}^{N} x_j \phi_j(t_k) = (Ax)_k,
\]

where the sampling matrix \( A \in \mathbb{C}^{m \times N} \) have entries

(3)

\[
A_{k,j} = \phi_j(t_k).
\]

Let an \( s \)-sparse \( x \in \mathbb{C}^{N} \) be fixed. Let sampling points \( t_1, \ldots, t_m \) be randomly selected according to \( \nu \), and let \( A \in \mathbb{C}^{m \times N} \) be the sampling matrix (3). If

\[
m \geq C K^2 s \ln(N) \ln(\varepsilon^{-1}),
\]

then \( x \) is the unique minimizer of \( \|z\|_1 \) subject to \( Az = Ax \) with probability at least \( 1 - \varepsilon \).

(Proof based on the golfing scheme).
Restricted Isometry Property for BOS*

Let $A \in \mathbb{C}^{m \times N}$ be the random sampling matrix associated to a BOS with constant $K \geq 1$. For $\delta \in (0, 1)$, if $m \geq C K^2 \delta^{-2} \ln 4 (N)$, then, with probability at least $1 - N^{-\gamma} \ln 3 (N)$, the matrix $\frac{1}{\sqrt{m}} A$ has a restricted isometry constant satisfying $\delta s \leq \delta$.

(Proof uses Dudley's inequality, empirical method of Maurey, etc.)
Let $A \in \mathbb{C}^{m \times N}$ be the random sampling matrix associated to a BOS with constant $K \geq 1$. 
Let $A \in \mathbb{C}^{m \times N}$ be the random sampling matrix associated to a BOS with constant $K \geq 1$. For $\delta \in (0, 1)$, if

$$m \geq C K^2 \delta^{-2} s \ln^4(N),$$

then, with probability at least $1 - N^{-\ln^3(N)}$, the matrix $\frac{1}{\sqrt{m}}A$ has a restricted isometry constant satisfying $\delta_s \leq \delta$. 

(Proof uses Dudley's inequality, empirical method of Maurey, etc.)
Let $A \in \mathbb{C}^{m \times N}$ be the random sampling matrix associated to a BOS with constant $K \geq 1$. For $\delta \in (0, 1)$, if

$$m \geq C K^2 \delta^{-2} s \ln^4(N),$$

then, with probability at least $1 - N^{-\ln^3(N)}$, the matrix $\frac{1}{\sqrt{m}} A$ has a restricted isometry constant satisfying $\delta_s \leq \delta$.

(Proof uses Dudley’s inequality, empirical method of Maurey, etc.)
Stability: Uniform Setting*

\[ q \geq p \geq 1, \quad \text{a pair } (A, \Delta) \text{ is mixed } (\ell_q, \ell_p)-\text{instance optimal of order } s \text{ with constant } C > 0 \text{ if } \|x - \Delta(Ax)\|_q \leq C s^{1/p - 1/q} \sigma_s(x)^p \text{ for all } x \in \mathbb{R}^N. \]

\[
\text{Let } A \in \mathbb{R}^{m \times N}. \quad \text{If there exists } \Delta \text{ making } (A, \Delta) \text{ mixed } (\ell_q, \ell_p)-\text{instance optimal of order } s \text{ with constant } C, \text{ then (4)} \]
\[
\|v\|_q \leq C s^{1/p - 1/q} \sigma_2^s(v)^p \text{ for all } v \in \ker A. \]

\[
\text{Conversely, if (4) holds, then there exists } \Delta \text{ making } (A, \Delta) \text{ mixed } (\ell_q, \ell_p)-\text{instance optimal of order } s \text{ with constant } 2C. \]

\[
\text{If there is an } \ell_2-\text{instance optimal pair of order } s \geq 1 \text{ with constant } C, \text{ then } m \geq c N \text{ for some constant } c \text{ depending only on } C. \]
Stability: Uniform Setting*

- Given $q \geq p \geq 1$, a pair $(A, \Delta)$ is mixed $(\ell_q, \ell_p)$-instance optimal of order $s$ with constant $C > 0$ if

$$\|x - \Delta(Ax)\|_q \leq \frac{C}{s^{1/p-1/q}} \sigma_s(x)_p$$

for all $x \in \mathbb{C}^N$. 

*Remark: If there is an $\ell_2$-instance optimal pair of order $s \geq 1$ with constant $C$, then $m \geq cN$ for some constant $c$ depending only on $C$. 
Stability: Uniform Setting*

- Given $q \geq p \geq 1$, a pair $(A, \Delta)$ is mixed $(\ell_q, \ell_p)$-instance optimal of order $s$ with constant $C > 0$ if

$$\|x - \Delta(Ax)\|_q \leq \frac{C}{s^{1/p - 1/q}} \sigma_s(x)_p$$

for all $x \in \mathbb{C}^N$.

- Let $A \in \mathbb{R}^{m \times N}$.
Stability: Uniform Setting*

- Given $q \geq p \geq 1$, a pair $(A, \Delta)$ is mixed $(\ell_q, \ell_p)$-instance optimal of order $s$ with constant $C > 0$ if

$$
\|x - \Delta(Ax)\|_q \leq \frac{C}{s^{1/p-1/q}} \sigma_s(x)_p \quad \text{for all } x \in \mathbb{C}^N.
$$

- Let $A \in \mathbb{R}^{m \times N}$. If there exists $\Delta$ making $(A, \Delta)$ mixed $(\ell_q, \ell_p)$-instance optimal of order $s$ with constant $C$, then

$$(4) \quad \|v\|_q \leq \frac{C}{s^{1/p-1/q}} \sigma_{2s}(v)_p \quad \text{for all } v \in \ker A.$$

Stability: Uniform Setting\(^*\)

- Given \( q \geq p \geq 1 \), a pair \((A, \Delta)\) is mixed \((\ell_q, \ell_p)\)-instance optimal of order \( s \) with constant \( C > 0 \) if

\[
\|x - \Delta(Ax)\|_q \leq \frac{C}{s^{1/p - 1/q}} \sigma_s(x)_p \quad \text{for all } x \in \mathbb{C}^N.
\]

- Let \( A \in \mathbb{R}^{m \times N} \). If there exists \( \Delta \) making \((A, \Delta)\) mixed \((\ell_q, \ell_p)\)-instance optimal of order \( s \) with constant \( C \), then

\[
(4) \quad \|v\|_q \leq \frac{C}{s^{1/p - 1/q}} \sigma_{2s}(v)_p \quad \text{for all } v \in \ker A.
\]

Conversely, if (4) holds, then there exists \( \Delta \) making \((A, \Delta)\) mixed \((\ell_q, \ell_p)\)-instance optimal of order \( s \) with constant \( 2C \).
Stability: Uniform Setting*

- Given \( q \geq p \geq 1 \), a pair \((A, \Delta)\) is mixed \((\ell_q, \ell_p)\)-instance optimal of order \( s \) with constant \( C > 0 \) if

\[
\|x - \Delta(Ax)\|_q \leq \frac{C}{s^{1/p-1/q}} \sigma_s(x)_p \quad \text{for all } x \in \mathbb{C}^N.
\]

- Let \( A \in \mathbb{R}^{m \times N} \). If there exists \( \Delta \) making \((A, \Delta)\) mixed \((\ell_q, \ell_p)\)-instance optimal of order \( s \) with constant \( C \), then

\[
\|v\|_q \leq \frac{C}{s^{1/p-1/q}} \sigma_{2s}(v)_p \quad \text{for all } v \in \ker A.
\]

Conversely, if (4) holds, then there exists \( \Delta \) making \((A, \Delta)\) mixed \((\ell_q, \ell_p)\)-instance optimal of order \( s \) with constant \( 2C \).

- If there is an \( \ell_2 \)-instance optimal pair of order \( s \geq 1 \) with constant \( C \), then

\[
m \geq c N
\]

for some constant \( c \) depending only on \( C \).
Stability: Nonuniform Setting*

Let an $s$-sparse $x \in \mathbb{R}^N$ be fixed. Let $A \in \mathbb{R}^{m \times N}$ be a matrix with $\text{ind}_N(0, m - 1/2)$ entries. If $N \geq c_2 m$ and $m \geq c_3 s \ln(N/m)$, then

$$\|x - \Delta_1(Ax + e)\|_2 \leq C \sigma_s(x)^2 + D \|e\|_2$$

holds for all $e \in \mathbb{R}^m$ with probability at least $1 - 5 \exp(-c_1 m)$. This uses the $\ell_1$-quotient property of $A$ with constant $d$ relative to the $\ell_2$-norm $\|\cdot\|_2$ on $\mathbb{R}^m$, which is expressed in one of the forms:

1. for all $e \in \mathbb{R}^m$, there exists $u \in \mathbb{R}^N$ with $Au = e$ and $\|u\|_1 \leq d \sqrt{s} \|e\|$, where $s^* := m / \ln(N/m)$.

2. $\|\lfloor e \rfloor\|_{\ell_1/\ker A} \leq d \sqrt{s^*} \|e\|$ for all $e \in \mathbb{R}^m$,

3. $\|e\|_* \leq d \sqrt{s^*} \|A^* e\|_\infty$ for all $e \in \mathbb{R}^m$. 
Stability: Nonuniform Setting*

Uniform $\ell_2$-instance optimality may be irrelevant, but nonuniform $\ell_2$-instance optimality is quite relevant...
Stability: Nonuniform Setting*

Uniform $\ell_2$-instance optimality may be irrelevant, but nonuniform $\ell_2$-instance optimality is quite relevant...

Let an $s$-sparse $x \in \mathbb{R}^N$ be fixed.
Stability: Nonuniform Setting*

Uniform $\ell_2$-instance optimality may be irrelevant, but nonuniform $\ell_2$-instance optimality is quite relevant...

Let an $s$-sparse $x \in \mathbb{R}^N$ be fixed. Let $A \in \mathbb{R}^{m \times N}$ be a matrix with $\text{ind } \mathcal{N}(0, m^{-1/2})$ entries.
Stability: Nonuniform Setting*

Uniform $\ell_2$-instance optimality may be irrelevant, but nonuniform $\ell_2$-instance optimality is quite relevant...

Let an $s$-sparse $\mathbf{x} \in \mathbb{R}^N$ be fixed. Let $A \in \mathbb{R}^{m \times N}$ be a matrix with ind $\mathcal{N}(0, m^{-1/2})$ entries. If $N \geq c_2 m$ and $m \geq c_3 s \ln(N/m)$, then

$$\|\mathbf{x} - \Delta_1(A\mathbf{x} + \mathbf{e})\|_2 \leq C \sigma_s(\mathbf{x})_2 + D \|\mathbf{e}\|_2$$

holds for all $\mathbf{e} \in \mathbb{R}^m$ with probability at least $1 - 5 \exp(-c_1 m)$. 

---

*Stability: Nonuniform Setting* indicates a focus on the stability of algorithms or solutions in the context of nonuniform settings, contrasting with uniform settings where optimality may be assessed differently.
Stability: Nonuniform Setting*

Uniform $\ell_2$-instance optimality may be irrelevant, but nonuniform $\ell_2$-instance optimality is quite relevant...

Let an $s$-sparse $x \in \mathbb{R}^N$ be fixed. Let $A \in \mathbb{R}^{m \times N}$ be a matrix with ind $\mathcal{N}(0, m^{-1/2})$ entries. If $N \geq c_2 m$ and $m \geq c_3 s \ln(N/m)$, then

$$\|x - \Delta_1(Ax + e)\|_2 \leq C \sigma_s(x)_2 + D \|e\|_2$$

holds for all $e \in \mathbb{R}^m$ with probability at least $1 - 5 \exp(-c_1 m)$.

This uses the $\ell_1$-quotient property of $A$ with constant $d$ relative to the $\ell_2$-norm $\| \cdot \|$ on $\mathbb{R}^m$, which is expressed in one of the forms:
Stability: Nonuniform Setting*

Uniform $\ell_2$-instance optimality may be irrelevant, but nonuniform $\ell_2$-instance optimality is quite relevant...

Let an $s$-sparse $x \in \mathbb{R}^N$ be fixed. Let $A \in \mathbb{R}^{m \times N}$ be a matrix with ind $\mathcal{N}(0, m^{-1/2})$ entries. If $N \geq c_2 m$ and $m \geq c_3 s \ln(N/m)$, then

$$\|x - \Delta_1(Ax + e)\|_2 \leq C \sigma_s(x)_2 + D \|e\|_2$$

holds for all $e \in \mathbb{R}^m$ with probability at least $1 - 5 \exp(-c_1 m)$.

This uses the $\ell_1$-quotient property of $A$ with constant $d$ relative to the $\ell_2$-norm $\|\cdot\|$ on $\mathbb{R}^m$, which is expressed in one of the forms:

- for all $e \in \mathbb{R}^m$, there exists $u \in \mathbb{R}^N$ with

$$Au = e \quad \text{and} \quad \|u\|_1 \leq d \sqrt{s_*} \|e\|, \quad s_* := m/\ln(N/m),$$
Stability: Nonuniform Setting*

Uniform $\ell_2$-instance optimality may be irrelevant, but nonuniform $\ell_2$-instance optimality is quite relevant...

Let an $s$-sparse $x \in \mathbb{R}^N$ be fixed. Let $A \in \mathbb{R}^{m \times N}$ be a matrix with ind $\mathcal{N}(0, m^{-1/2})$ entries. If $N \geq c_2 m$ and $m \geq c_3 s \ln(N/m)$, then

$$\|x - \Delta_1(Ax + e)\|_2 \leq C \sigma_s(x)_2 + D \|e\|_2$$

holds for all $e \in \mathbb{R}^m$ with probability at least $1 - 5 \exp(-c_1 m)$.

This uses the $\ell_1$-quotient property of $A$ with constant $d$ relative to the $\ell_2$-norm $\|\cdot\|$ on $\mathbb{R}^m$, which is expressed in one of the forms:

- for all $e \in \mathbb{R}^m$, there exists $u \in \mathbb{R}^N$ with
  $$Au = e \quad \text{and} \quad \|u\|_1 \leq d \sqrt{s_*} \|e\|, \quad s_* := m/\ln(N/m),$$
- $\|[e]\|_{\ell_1/\ker A} \leq d \sqrt{s_*} \|e\|$ for all $e \in \mathbb{R}^m$,
Stability: Nonuniform Setting*

Uniform $\ell_2$-instance optimality may be irrelevant, but nonuniform $\ell_2$-instance optimality is quite relevant...

Let an $s$-sparse $x \in \mathbb{R}^N$ be fixed. Let $A \in \mathbb{R}^{m \times N}$ be a matrix with ind $\mathcal{N}(0, m^{-1/2})$ entries. If $N \geq c_2 m$ and $m \geq c_3 s \ln(N/m)$, then

$$\|x - \Delta_1(Ax + e)\|_2 \leq C \sigma_s(x)_2 + D \|e\|_2$$

holds for all $e \in \mathbb{R}^m$ with probability at least $1 - 5 \exp(-c_1 m)$.

This uses the $\ell_1$-quotient property of $A$ with constant $d$ relative to the $\ell_2$-norm $\|\cdot\|$ on $\mathbb{R}^m$, which is expressed in one of the forms:

- for all $e \in \mathbb{R}^m$, there exists $u \in \mathbb{R}^N$ with
  \[ Au = e \quad \text{and} \quad \|u\|_1 \leq d \sqrt{s_*}\|e\|, \quad s_* := m/\ln(N/m), \]

- $\|[e]\|_{\ell_1/\ker A} \leq d \sqrt{s_*}\|e\|$ for all $e \in \mathbb{R}^m$,
- $\|e\|_* \leq d \sqrt{s_*}\|A^*e\|_\infty$ for all $e \in \mathbb{R}^m$. 

* This content is marked as 'Stability: Nonuniform Setting' and is part of a larger discussion on the stability of optimization problems in various settings.