Last Time

Last time we introduced the following Restricted Isometry Property (RIP) of order $k$. 
Last Time

- Last time we introduced the following Restricted Isometry Property (RIP) of order $k$.
- Let $\Phi$ be an $n \times N$ matrix.
Last Time

- Last time we introduced the following Restricted Isometry Property (RIP) of order $k$
- Let $\Phi$ be an $n \times N$ matrix
- There exists $0 < \delta = \delta_k < 1$ such that

\[(1 - \delta)\|x\|_{\ell_2^N}^2 \leq \|\Phi(x)\|_{\ell_2^n}^2 \leq (1 + \delta)\|x\|_{\ell_2^N}^2, \quad x \in \Sigma_k\]
Last Time

- Last time we introduced the following Restricted Isometry Property (RIP) of order $k$

- Let $\Phi$ be an $n \times N$ matrix

- There exists $0 < \delta = \delta_k < 1$ such that

\[(1 - \delta) \|x\|_{\ell^2_N}^2 \leq \|\Phi(x)\|_{\ell^2_N}^2 \leq (1 + \delta) \|x\|_{\ell^2_N}^2, \quad x \in \Sigma_k\]

- We showed that random matrices will with high probability have the RIP for the range of

\[k \leq c(\delta)n / \log(N/n)\]
Last Time

- Last time we introduced the following **Restricted Isometry Property (RIP)** of order $k$
- Let $\Phi$ be an $n \times N$ matrix
- There exists $0 < \delta = \delta_k < 1$ such that
  \[
  (1 - \delta)\|x\|_{\ell_2}^2 \leq \|\Phi(x)\|_{\ell_2}^2 \leq (1 + \delta)\|x\|_{\ell_2}^2, \quad x \in \Sigma_k
  \]
- We showed that random matrices will with high probability have the **RIP** for the range of
  $k \leq c(\delta)n / \log(N/n)$
- These matrices then gave optimal performance for encoding compact classes such as finite balls in the $\ell_p^N$ spaces.
Matrices $\Phi$ satisfying RIP are closely related to the following Johnson-Lindenstrauss Lemma.
Johnson-Lindenstrauss Lemma

Matrices $\Phi$ satisfying RIP are closely related to the following Johnson-Lindenstrauss Lemma.

Given a set of points $Q$ in $\mathbb{R}^N$, then for any $n \geq c\epsilon^{-2}\log[\#(Q)]$, there is a linear mapping $\Phi$ from $\mathbb{R}^N$ into $\mathbb{R}^n$ such that for all $x, y \in Q$

$$(1 - \epsilon)\text{dist}(x, y) \leq \text{dist}(\Phi(x), \Phi(y)) \leq (1 + \epsilon)\text{dist}(x, y)$$
Johnson-Lindenstrauss Lemma

Matrices $\Phi$ satisfying RIP are closely related to the following Johnson-Lindenstrauss Lemma

Given a set of points $Q$ in $\mathbb{R}^N$, then for any $n \geq c\epsilon^{-2}\log[\#(Q)]$, there is a linear mapping $\Phi$ from $\mathbb{R}^N$ into $\mathbb{R}^n$ such that for all $x, y \in Q$

$$(1 - \epsilon)\text{dist}(x, y) \leq \text{dist}(\Phi(x), \Phi(y)) \leq (1 + \epsilon)\text{dist}(x, y)$$

This lemma is easily proved from the Concentration of Measure Inequalities
**Johnson-Lindenstrauss Lemma**

- Matrices $\Phi$ satisfying RIP are closely related to the following Johnson-Lindenstrauss Lemma

- Given a set of points $Q$ in $\mathbb{R}^N$, then for any $n \geq c\epsilon^{-2}\log[\#(Q)]$, there is a linear mapping $\Phi$ from $\mathbb{R}^N$ into $\mathbb{R}^n$ such that for all $x, y \in Q$

  $$(1 - \epsilon) \text{dist}(x, y) \leq \text{dist}(\Phi(x), \Phi(y)) \leq (1 + \epsilon) \text{dist}(x, y)$$

- This lemma is easily proved from the Concentration of Measure Inequalities

- A Random draw of a matrix satisfying CMI will satisfy JL
**Johnson-Lindenstrauss Lemma**

- Matrices $\Phi$ satisfying RIP are closely related to the following Johnson-Lindenstrauss Lemma

- Given a set of points $Q$ in $\mathbb{R}^N$, then for any $n \geq c\epsilon^{-2} \log[#(Q)]$, there is a linear mapping $\Phi$ from $\mathbb{R}^N$ into $\mathbb{R}^n$ such that for all $x, y \in Q$

$$ (1 - \epsilon) \text{dist}(x, y) \leq \text{dist}(\Phi(x), \Phi(y)) \leq (1 + \epsilon) \text{dist}(x, y) $$

- This lemma is easily proved from the Concentration of Measure Inequalities

- A Random draw of a matrix satisfying CMI will satisfy JL

- It is also easy to prove RIP from JL
Johnson-Lindenstrauss Lemma

Matrices $\Phi$ satisfying RIP are closely related to the following Johnson-Lindenstrauss Lemma

Given a set of points $Q$ in $\mathbb{R}^N$, then for any $n \geq c\epsilon^{-2} \log[\#(Q)]$, there is a linear mapping $\Phi$ from $\mathbb{R}^N$ into $\mathbb{R}^n$ such that for all $x, y \in Q$

$$(1 - \epsilon)\text{dist}(x, y) \leq \text{dist}(\Phi(x), \Phi(y)) \leq (1 + \epsilon)\text{dist}(x, y)$$

This lemma is easily proved from the Concentration of Measure Inequalities

A Random draw of a matrix satisfying CMI will satisfy JL

It is also easy to prove RIP from JL

$\text{CMI} \rightarrow \text{JL} \rightarrow \text{RIP}$
Last time we introduced two measures of performance in compressed sensing - neither handles adequately the performance on general signals.
Finer Measure of Performance

- Last time we introduced two measures of performance in compressed sensing - neither handles adequately the performance on general signals.
- We want performance for each $x \in \mathbb{R}^N$. 
Last time we introduced two measures of performance in compressed sensing - neither handles adequately the performance on general signals.

We want performance for each $x \in \mathbb{R}^N$.

Recall $\Sigma_k := \{x \in \mathbb{R}^N : \#\text{supp}(x) \leq k\}$

$$\sigma_k(x) := \inf_{z \in \Sigma_k} \| x - z \|$$
Finer Measure of Performance

- Last time we introduced two measures of performance in compressed sensing - neither handles adequately the performance on general signals.

- We want performance for each $x \in \mathbb{R}^N$.

- Recall $\Sigma_k := \{x \in \mathbb{R}^N : \#\text{supp}(x) \leq k\}$.

$$\sigma_k(x)_X := \inf_{z \in \Sigma_k} \|x - z\|_X$$

- Given an encoding - decoding pair $(\Phi, \Delta)$, we say that this pair is Instance-Optimal of order $k$ for $X$ if for an absolute constant $C > 0$

$$\|x - \Delta(\Phi(x))\|_X \leq C\sigma_k(x)_X$$
Last time we introduced two measures of performance in compressed sensing - neither handles adequately the performance on general signals.

We want performance for each \( x \in \mathbb{R}^N \).

Recall \( \Sigma_k := \{ x \in \mathbb{R}^N : \#\text{supp}(x) \leq k \} \)

\[
\sigma_k(x)_X := \inf_{z \in \Sigma_k} \| x - z \|_X
\]

Given an encoding - decoding pair \((\Phi, \Delta)\), we say that this pair is \textbf{Instance-Optimal} of order \( k \) for \( X \) if for an absolute constant \( C > 0 \)

\[
\| x - \Delta(\Phi(x)) \|_X \leq C \sigma_k(x)_X
\]

Given \( n, N \), the best encoding - decoding pairs are those which have the largest \( k \).
Optimal Matrices

What properties of a matrix $\Phi$ determine the range of instance optimality?
Optimal Matrices

- What properties of a matrix $\Phi$ determine the range of instance optimality?
- The secret lies in the null space $\mathcal{N}$ of $\Phi$
What properties of a matrix $\Phi$ determine the range of instance optimality?

The secret lies in the null space $\mathcal{N}$ of $\Phi$

We say $\Phi$ has the null space property in $X$ if

$$\|\eta\|_X \leq C_0 \sigma_k(\eta)_X, \quad \eta \in \mathcal{N}$$
Optimal Matrices

- What properties of a matrix $\Phi$ determine the range of instance optimality?
- The secret lies in the null space $\mathcal{N}$ of $\Phi$
- We say $\Phi$ has the null space property in $X$ if

$$\|\eta\|_X \leq C_0\sigma_k(\eta)_X, \quad \eta \in \mathcal{N}$$

- When $X = \ell_q^N$ for some $q$ then an equivalent formulation is

$$\|\eta_T\|_X \leq C_1\|\eta_{T^c}\|_X, \quad \forall \eta \in \mathcal{N}, \#(T) \leq k$$
Optimal Matrices

- What properties of a matrix $\Phi$ determine the range of instance optimality?
- The secret lies in the null space $\mathcal{N}$ of $\Phi$
- We say $\Phi$ has the null space property in $X$ if
  \[ \|\eta\|_X \leq C_0 \sigma_k(\eta)_X, \quad \eta \in \mathcal{N} \]

- When $X = \ell^N_q$ for some $q$ then an equivalent formulation is
  \[ \|\eta_T\|_X \leq C_1 \|\eta_{Te}\|_X, \quad \forall \eta \in \mathcal{N}, \#(T) \leq k \]

- Elements in the null space should have no structure - look like noise
Main Result

Theorem (Cohen-Dahmen-DeVore) Given an $n \times N$ matrix $\Phi$, a norm $\| \cdot \|_X$ and a value of $k$, then to have instance optimality in $X$ with a constant $C_0$ a necessary and sufficient is that $\Phi$ has the null space of order $2k$ with a constant $C_1$ where $C_1 = C_0/2$ in the sufficient part and $C_1 = C_0$ in the necessary part.
Proof of Sufficiency

- define a decoder \( \Delta \) for \( \Phi \)

\[
\Delta(y) := \text{Argmin}_{z \in \mathcal{F}(y)} \sigma_k(z) X
\]
Proof of Sufficiency

- define a decoder \( \Delta \) for \( \Phi \)

\[
\Delta(y) := \text{Argmin}_{z \in \mathcal{F}(y)} \sigma_k(z)x
\]

- \( \eta := x - \Delta(\Phi(x)) \) is in \( \mathcal{N} \)
Proof of Sufficiency

- define a decoder \( \Delta \) for \( \Phi \)

\[
\Delta(y) := \operatorname{Argmin}_{z \in \mathcal{F}(y)} \sigma_k(z)_X
\]

- \( \eta := x - \Delta(\Phi(x)) \) is in \( \mathcal{N} \)

- \[
\|x - \Delta(\Phi(x))\|_X \leq (C_0/2) \sigma_{2k}(x - \Delta(\Phi(x)))_X \\
\leq (C_0/2)(\sigma_k(x)_X + \sigma_k(\Delta(\Phi(x)))_X) \leq C_0 \sigma_k(x)_X
\]
Proof of Sufficiency

- Define a decoder $\Delta$ for $\Phi$

$$
\Delta(y) := \text{Argmin}_{z \in \mathcal{F}(y)} \sigma_k(z)_X
$$

- $\eta := x - \Delta(\Phi(x))$ is in $\mathcal{N}$

- $\|x - \Delta(\Phi(x))\|_X \leq (C_0/2)\sigma_{2k}(x - \Delta(\Phi(x))_X$

$$
\leq (C_0/2)(\sigma_k(x)_X + \sigma_k(\Delta(\Phi(x)))_X) \leq C_0\sigma_k(x)_X
$$

- We used $\sigma_{2k}(x + z)_X \leq \sigma_k(x)_X + \sigma_k(z)_X$
Proof of Sufficiency

- define a decoder $\Delta$ for $\Phi$

$$\Delta(y) := \operatorname{Argmin}_{z \in \mathcal{F}(y)} \sigma_k(z)_X$$

- $\eta := x - \Delta(\Phi(x))$ is in $\mathcal{N}$

$$\|x - \Delta(\Phi(x))\|_X \leq (C_0/2)\sigma_{2k}(x - \Delta(\Phi(x)))_X$$
$$\leq (C_0/2)(\sigma_k(x)_X + \sigma_k(\Delta(\Phi(x)))_X) \leq C_0\sigma_k(x)_X$$

- We used $\sigma_{2k}(x + z)_X \leq \sigma_k(x)_X + \sigma_k(z)_X$

- The last inequality uses the fact that $\Delta(\Phi(x))$ minimizes $\sigma_k(z)$ over $\mathcal{F}(y)$
Verifying Null Space Property

The usual way to verify NSP is through RIP.
Verifying Null Space Property

- The usual way to verify NSP is through RIP
- We shall illustrate this for $X = \ell_1$
Verifying Null Space Property

- The usual way to verify NSP is through RIP
- We shall illustrate this for $X = \ell_1$
- **Cohen-Dahmen-DeVore** If $\Phi$ has RIP for $2k$ and some $\delta < 1/2$ then $\Phi$ is instance-optimal of order $k$ in $\ell_1$:

$$\|x - \Delta(\Phi(x))\|_{\ell_1} \leq C\sigma_k(x)_{\ell_1}$$
Verifying Null Space Property

- The usual way to verify NSP is through RIP.
- We shall illustrate this for $X = \ell_1$.
- **Cohen-Dahmen-DeVore** If $\Phi$ has RIP for $2k$ and some $\delta < 1/2$ then $\Phi$ is instance-optimal of order $k$ in $\ell_1$:

  $$\|x - \Delta(\Phi(x))\|_{\ell_1} \leq C\sigma_k(x)_{\ell_1}$$

- This means that there is some decoder - not necessarily practical (we shall discuss practical decoders in the next lecture).
Verifying Null Space Property

The usual way to verify NSP is through RIP

We shall illustrate this for $X = \ell_1$

Cohen-Dahmen-DeVore If $\Phi$ has RIP for $2k$ and some $\delta < 1/2$ then $\Phi$ is instance-optimal of order $k$ in $\ell_1$:

$$\|x - \Delta(\Phi(x))\|_{\ell_1} \leq C\sigma_k(x)_{\ell_1}$$

This means that there is some decoder - not necessarily practical (we shall discuss practical decoders in the next lecture)

Hence we know there are matrices $\Phi$ which have Instance Optimality in $\ell_1$ for $k \leq c_0n/\log(N/n)$
Verifying Null Space Property

- The usual way to verify NSP is through RIP.
- We shall illustrate this for $X = \ell_1$.
- **Cohen-Dahmen-DeVore** If $\Phi$ has RIP for $2k$ and some $\delta < 1/2$ then $\Phi$ is instance-optimal of order $k$ in $\ell_1$:

  $$\|x - \Delta(\Phi(x))\|_{\ell_1} \leq C\sigma_k(x)_{\ell_1}$$

- This means that there is some decoder - not necessarily practical (we shall discuss practical decoders in the next lecture).
- Hence we know there are matrices $\Phi$ which have Instance Optimality in $\ell_1$ for $k \leq c_0n/\log(N/n)$.
- This range of $k$ cannot be improved.
Proof of NSP for $X = \ell_1$

- Assume $\Phi$ has RIP of order $2k$
Proof of NSP for $X = \ell_1$

- Assume $\Phi$ has RIP of order $2k$
- Fix $\eta \in \mathcal{N}$. Let $T = T_0$ be the set of indices of its $k$ largest coordinates, let $T_1$ be the set of indices its next $k$ largest coordinates and so on. Let $S = T_0 \cup T_1$. 
Proof of NSP for $X = \ell_1$

- Assume $\Phi$ has RIP of order $2k$.

- Fix $\eta \in \mathcal{N}$. Let $T = T_0$ be the set of indices of its $k$ largest coordinates, let $T_1$ be the set of indices its next $k$ largest coordinates and so on. Let $S = T_0 \cup T_1$.

- Since $\Phi(\eta_S) = -\Phi(\eta_{S^c})$ and $\|\eta_T\|_{\ell_1} \leq \|\eta_S\|_{\ell_1}$

\[
[2k]^{-1/2} \|\eta_T\|_{\ell_1} \leq [2k]^{-1/2} \|\eta_S\|_{\ell_1} \leq \|\eta_S\|_{\ell_2} \leq C_0 \|\Phi(\eta_S)\|_{\ell_2} = C_0 \|\Phi(\eta_{S^c})\|_{\ell_2} \leq C_0 \sum_{j=2}^{s} \|\Phi(\eta_{T_j})\|_{\ell_2} \leq C'_0 \sum_{j=2}^{s} \|\eta_{T_j}\|_{\ell_2}
\]
Proof of NSP for $X = \ell_1$

- Assume $\Phi$ has RIP of order $2k$
- Fix $\eta \in \mathcal{N}$. Let $T = T_0$ be the set of indices of its $k$ largest coordinates, let $T_1$ be the set of indices its next $k$ largest coordinates and so on. Let $S = T_0 \cup T_1$.
- Since $\Phi(\eta_S) = -\Phi(\eta_{S^c})$ and $\|\eta_T\|_{\ell_1} \leq \|\eta_S\|_{\ell_1}$

$$[2k]^{-1/2}\|\eta_T\|_{\ell_2} \leq [2k]^{-1/2}\|\eta_S\|_{\ell_1} \leq \|\eta_S\|_{\ell_2} \leq C_0\|\Phi(\eta_S)\|_{\ell_2} = C_0\|\Phi(\eta_{S^c})\|_{\ell_2} \leq C_0\sum_{j=2}^{s} \|\Phi(\eta_{T_j})\|_{\ell_2} \leq C_0'\sum_{j=2}^{s} \|\eta_{T_j}\|_{\ell_2}$$

- Now $|\eta_i| \leq \frac{1}{k} \sum_{\nu \in T_{j-1}} |\eta_\nu|$ when $i \in T_j$ and so

$$\|\eta_{T_j}\|_{\ell_2} \leq k^{-1/2} \|\eta_{T_{j-1}}\|_{\ell_1}$$
Proof of NSP for $X = \ell_1$

- Assume $\Phi$ has RIP of order $2k$
- Fix $\eta \in \mathcal{N}$. Let $T = T_0$ be the set of indices of its $k$ largest coordinates, let $T_1$ be the set of indices its next $k$ largest coordinates and so on. Let $S = T_0 \cup T_1$.

Since $\Phi(\eta_S) = -\Phi(\eta_{S^c})$ and $\|\eta_T\|_1 \leq \|\eta_S\|_1$

\[
[2k]^{-1/2}\|\eta_T\|_1 \leq [2k]^{-1/2}\|\eta_S\|_1 \leq \|\eta_S\|_2 \leq C_0\|\Phi(\eta_S)\|_2 = C_0\|\Phi(\eta_{S^c})\|_2 \leq C_0\sum_{j=2}^{s} \|\Phi(\eta_{T_j})\|_2 \leq C'_0\sum_{j=2}^{s} \|\eta_{T_j}\|_2
\]

Now $|\eta_i| \leq \frac{1}{k} \sum_{\nu \in T_{j-1}} |\eta_{\nu}|$ when $i \in T_j$ and so

$\|\eta_{T_j}\|_2 \leq k^{-1/2}\|\eta_{T_{j-1}}\|_1$

- Hence,

\[
\|\eta_T\|_1 \leq [2k]^{1/2}C'_0 k^{-1/2} \sum_{j=1}^{s-1} \|\eta_{T_j}\|_1 \leq \sqrt{2}C'_0\|\eta_{T^c}\|_1
\]
Instance Optimality in $\ell_p$

Given $n, N$ and a constant $C_0$ then we have instance optimality in $\ell_2$ for $k$ and this $C_0$ only if $k \leq \frac{C_0 n}{N}$.
Instance Optimality in $\ell_p$

- Given $n, N$ and a constant $C_0$ then we have instance optimality in $\ell_2$ for $k$ and this $C_0$ only if $k \leq \frac{C_0 n}{N}$

- This shows that instance-optimal is not a viable concept for $\ell_2$
Instance Optimality in $\ell_p$

- Given $n, N$ and a constant $C_0$ then we have instance optimality in $\ell_2$ for $k$ and this $C_0$ only if $k \leq \frac{C_0 n}{N}$

- This shows that instance-optimal is not a viable concept for $\ell_2$

- We have to make $cN$ measurement to even get instance optimality for $k = 1$
Instance Optimality in $\ell_p$

- Given $n, N$ and a constant $C_0$ then we have instance optimality in $\ell_2$ for $k$ and this $C_0$ only if $k \leq \frac{C_0 n}{N}$
- This shows that instance-optimal is not a viable concept for $\ell_2$
- We have to make $cN$ measurement to even get instance optimality for $k = 1$
- This bound cannot be improved
Instance Optimality in $\ell_p$

- Given $n$, $N$ and a constant $C_0$ then we have instance optimality in $\ell_2$ for $k$ and this $C_0$ only if $k \leq \frac{C_0 n}{N}$.
- This shows that instance-optimal is not a viable concept for $\ell_2$.
- We have to make $cN$ measurement to even get instance optimality for $k = 1$.
- This bound cannot be improved.
- For $1 < p < 2$ the range of $k$ is
  \[ k \leq c_0 N^{\frac{2-2/p}{1-2/p}} [n / \log(N/n)]^{\frac{p}{2-p}} \]
Instance-Optimality in Probability

We saw that Instance-Optimality for $\ell_2^N$ is not viable.
Instance-Optimality in Probability

- We saw that Instance-Optimality for $\ell_2^N$ is not viable.
- We shall next show that it is possible to have Instance-Optimality in $\ell_2^N$ if we are willing to accept some small probability of failure.
We saw that Instance-Optimality for $\ell_2^N$ is not viable.

We shall next show that it is possible to have Instance-Optimality in $\ell_2^N$ if we are willing to accept some small probability of failure.

Let $\Phi(\omega)$ be a collection of random matrices.
Instance-Optimality in Probability

- We saw that Instance-Optimality for $\ell_2^N$ is not viable.
- We shall next show that it is possible to have Instance-Optimality in $\ell_2^N$ if we are willing to accept some small probability of failure.
- Let $\Phi(\omega)$ be a collection of random matrices.
- **Property P1:** We say this family satisfies RIP of order $k$ with probability $1 - \epsilon$ if a random draw from $\{\Phi(\omega)\}$ will satisfy RIP of order $k$ with probability $1 - \epsilon$: denote by $\Omega_1(k, \epsilon)$ the favorable draws.
Instance-Optimality in Probability

- We saw that Instance-Optimality for $\ell_2^N$ is not viable.
- We shall next show that it is possible to have Instance-Optimality in $\ell_2^N$ if we are willing to accept some small probability of failure.
- Let $\Phi(\omega)$ be a collection of random matrices.
  - **Property P1**: We say this family satisfies RIP of order $k$ with probability $1 - \epsilon$ if a random draw from $\{\Phi(\omega)\}$ will satisfy RIP of order $k$ with probability $1 - \epsilon$: denote by $\Omega_1(k, \epsilon)$ the favorable draws.
  - **Property P0**: We say $\{\Phi(\omega)\}$ is bounded with probability $1 - \epsilon$ if given any $x \in \mathbb{R}^N$ with probability $1 - \epsilon$ a random draw from $\{\Phi(\omega)\}$ will satisfy $\|\Phi(\omega)(x)\|_{\ell_2^N} \leq C_0\|x\|_{\ell_2^N}$ with $C_0$ an absolute constant: denote by $\Omega_0(x, \epsilon)$ the favorable draws.
Theorem: Cohen-Dahmen-DeVore

If $\{\Phi(\omega)\}$ satisfies RIP of order $3k$ and boundedness each with probability $1 - \epsilon$ then there are decoders $\Delta(\omega)$ such that given any $x \in \ell_2^N$ we have with probability $1 - 2\epsilon$

$$\|x - \Delta(\omega)\Phi(\omega)(x)\|_{\ell_2^N} \leq C_0\sigma_k(x)_{\ell_2^N}$$
Theorem: Cohen-Dahmen-DeVore

- If \( \{ \Phi(\omega) \} \) satisfies RIP of order \( 3k \) and boundedness each with probability \( 1 - \epsilon \) then there are decoders \( \Delta(\omega) \) such that given any \( x \in \ell_2^N \) we have with probability \( 1 - 2\epsilon \)

\[
\| x - \Delta(\omega)\Phi(\omega)(x) \|_{\ell_2^N} \leq C_0\sigma_k(x)_{\ell_2^N}
\]

- Poor man’s decoder
Theorem: Cohen-Dahmen-DeVore

- If \( \{ \Phi(\omega) \} \) satisfies RIP of order \( 3k \) and boundedness each with probability \( 1 - \epsilon \) then there are decoders \( \Delta(\omega) \) such that given any \( x \in \ell_2^N \) we have with probability \( 1 - 2\epsilon \)
  \[
  \| x - \Delta(\omega)\Phi(\omega)(x) \|_{\ell_2^N} \leq C_0 \sigma_k(x)_{\ell_2^N}
  \]

- Poor man's decoder

- \( \Delta(y) := \operatorname{Argmin}_{z \in \Sigma_k} \| y - \Phi(z) \|_{\ell_2} \)
Theorem: Cohen-Dahmen-DeVore

- If \( \{\Phi(\omega)\} \) satisfies RIP of order \( 3k \) and boundedness each with probability \( 1 - \epsilon \) then there are decoders \( \Delta(\omega) \) such that given any \( x \in \ell_2^N \) we have with probability \( 1 - 2\epsilon \)
  \[
  \|x - \Delta(\omega)\Phi(\omega)(x)\|_{\ell_2^N} \leq C_0\sigma_k(x)_{\ell_2^N}
  \]

- Poor man’s decoder
  \[
  \Delta(y) := \text{Argmin}_{z \in \Sigma_k} \|y - \Phi(z)\|_{\ell_2}
  \]

- We shall discuss other decoders in the next lecture
Proof of theorem

Let $x \in \mathbb{R}^N$ and $\Phi = \Phi(\omega)$ be the draw of the matrix $\Phi$. 
Proof of theorem

- Let $x \in \mathbb{R}^N$ and $\Phi = \Phi(\omega)$ be the draw of the matrix $\Phi$
- $T$ the set of indices of $k$ largest coefficients of $x$
Proof of theorem

- Let $x \in \mathbb{R}^N$ and $\Phi = \Phi(\omega)$ be the draw of the matrix $\Phi$
- $T$ the set of indices of $k$ largest coefficients of $x$
- $\Omega' := \Omega_1(k, \epsilon) \cap \Omega_0(x - x_T, \epsilon)$ and so $P(\Omega') \geq 1 - 2\epsilon$
Proof of theorem

Let \( x \in \mathbb{R}^N \) and \( \Phi = \Phi(\omega) \) be the draw of the matrix \( \Phi \)

\( T \) the set of indices of \( k \) largest coefficients of \( x \)

\( \Omega' := \Omega_1(k, \epsilon) \cap \Omega_0(x - x_T, \epsilon) \) and so \( P(\Omega') \geq 1 - 2\epsilon \)

For any \( \omega \in \Omega' \), we have

\[
\|x - x^*\|_{\ell_2} \leq \|x - x_T\|_{\ell_2} + \|x_T - x^*\|_{\ell_2} \leq \sigma_k(x)_{\ell_2} + \|x_T - x^*\|_{\ell_2}
\]
Estimate of Second Term

\[ \|x_T - x^*\|_2 \leq (1 - \delta)^{-1} \|\Phi(x_T - x^*)\|_2 \]

\[ \leq (1 - \delta)^{-1}(\|y - \Phi(x_T)\|_2 + \|y - \Phi(x^*)\|_2) \]

\[ \leq 2(1 - \delta)^{-1}\|y - \Phi(x_T)\|_2 = 2(1 - \delta)^{-1}\|\Phi(x - x_T)\|_2 \]

\[ \leq 2C(1 - \delta)^{-1}\|x - x_T\|_2 = 2C(1 - \delta)^{-1}\sigma_k(x)\|_2. \]
Applications of the Theorem

- Applies to Bernouli, Gaussian, and many other random families
Applications of the Theorem

- Applies to Bernouli, Gaussian, and many other random families
- If $k \leq cn/\log(N/n)$ then probability of failure is $\epsilon = e^{-cn}$
Applications of the Theorem

- Applies to Bernouli, Gaussian, and many other random families
- If $k \leq cn/\log(N/n)$ then probability of failure is $\epsilon = e^{-cn}$
- Since we can only construct matrices satisfying RIP for the large range of $k$ through probability, the above theorems seem quite satisfactory
Applications of the Theorem

- Applies to Bernoulli, Gaussian, and many other random families
- If $k \leq cn/\log(N/n)$ then probability of failure is $\epsilon = e^{-cn}$
- Since we can only construct matrices satisfying RIP for the large range of $k$ through probability, the above theorems seem quite satisfactory
- Notice that the probability is on the draw of $\Phi$ and not on $x$
Applications of the Theorem

- Applies to Bernouli, Gaussian, and many other random families
- If $k \leq cn/\log(N/n)$ then probability of failure is $\epsilon = e^{-cn}$
- Since we can only construct matrices satisfying RIP for the large range of $k$ through probability, the above theorems seem quite satisfactory
- Notice that the probability is on the draw of $\Phi$ and not on $x$
- Results resting on the choice of $x$ would not be satisfactory