Mathematical Foundation for Compressed Sensing

Jan-Olov Strömberg

Royal Institute of Technology, Stockholm, Sweden

Lecture 5, March 5, 2012

An outline for today

- A short summary from last lecture:

An outline for today

- A short summary from last lecture:
  - Best known for: RIP implying MP_1

An outline for today

- A short summary from last lecture:
  - Best known for: RIP implying MP_1
  - A lower estimate for RIP.

An outline for today

- A short summary from last lecture:
  - Best known for: RIP implying MP_1
  - A lower estimate for RIP.
  - RIP of Random matrices .

An outline for today

- A short summary from last lecture:
  - Best known for: RIP implying MP_1
  - A lower estimate for RIP.
  - RIP of Random matrices .
- RIP of Random matrices (continued).

A short summary from last lecture:

Theorem 3.3: If $\mathbf{A}$ fulfils the RIP property with constant $\delta_{2s} < \sqrt{2} - 1 \approx 0.412\ldots$, then $\mathbf{A} \in \mathrm{MP}_1(s)$.

A short summary from last lecture:

Theorem 3.3: If $\mathbf{A}$ fulfils the RIP property with constant $\delta_{2s} < \sqrt{2} - 1 \approx 0.412\ldots$, then $\mathbf{A} \in \mathrm{MP}_1(s)$.

Best known: enough with $\frac{\delta_{2s}}{\sqrt{1-\delta_{2s}^2}} < \frac{1}{\sqrt{2}}$. Thus it is enough with $\delta_{2s} < \frac{1}{\sqrt{3}} \approx 0.577\ldots$.

A short summary from last lecture:

One essential estimate for the proof of Theorem 3.3 is :

A short summary from last lecture:

One essential estimate for the proof of Theorem 3.3 is :
Lemma 3.5: If **A** satisfies RIP properties and $\mathbf{A}\mathbf{x} = 0$, then

$$\|\mathbf{x}_{S_1}\|_2 \leq \frac{\delta_{2s}}{1-\delta_s} \sum_{k=2} \|\mathbf{x}_{S_k}\|_2.$$

A short summary from last lecture:

One essential estimate for the proof of Theorem 3.3 is :
Lemma 3.5: If $\mathbf{A}$ satisfies RIP properties and $\mathbf{Ax} = 0$, then

$$\|\mathbf{x}_{S_1}\|_2 \leq \frac{\delta_{2s}}{1-\delta_s} \sum_{k=2} \|\mathbf{x}_{S_k}\|_2.$$

A short summary from last lecture:

One essential estimate for the proof of Theorem 3.3 is :
Lemma 3.5: If $\mathbf{A}$ satisfies RIP properties and $\mathbf{Ax} = 0$, then

$$\|\mathbf{x}_{S_1}\|_2 \leq \frac{\delta_{2s}}{1-\delta_s} \sum_{k=2} \|\mathbf{x}_{S_k}\|_2.$$

Remark: With some effort the constant $\frac{\delta_{2s}}{1-\delta_s}$ in Lemma 3.5 can be replaced by to $\frac{\delta_{2s}}{\sqrt{1-\delta_{2s}^2}}$.

A lower estimate for RIP.

A $m \times N$ matrix cannot satisfy a RIP with $\delta_s < 1$ unless m is large enough, depending on s an N.

A lower estimate for RIP.

A $m \times N$ matrix cannot satisfy a RIP with $\delta_s < 1$ unless m is large enough, depending on s an N.

Theorem 4.1 If $\delta_s < 1$ there is a constant $C > 0$ such that if any $m \times N$ matrix **A** har RIP with constant δ_s, then

$$m > Cs \log(Ne/s).$$

A lower estimate for RIP.

A $m \times N$ matrix cannot satisfy a RIP with $\delta_s < 1$ unless m is large enough, depending on s an N.

Theorem 4.1 If $\delta_s < 1$ there is a constant $C > 0$ such that if any $m \times N$ matrix $\mathbf{A}$ har RIP with constant δ_s, then

$$m > Cs\log(Ne/s).$$

More precisely we get

$$m \geq \frac{s\log(2Ne/s) - \log 4}{2(\log(\sqrt{1-\delta_s} + 4\sqrt{1+\delta_s}) - \log(1-\delta_s))}.$$

RIP for Gaussian Matrices

Let random matrix $\mathbf{A} = (A_{ij})_{1 \leq i \leq m, 1 \leq j \leq N}$ where each A_{ij} is an independent Gaussian random variable, i.e each $A_{ij} \in N(0,1)$ with distribution function $\phi(t)$ where

$$\phi(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}.$$

RIP for Gaussian Matrices

Let random matrix $\mathbf{A} = (A_{ij})_{1\le i\le m, 1\le j\le N}$ where each A_{ij} is an independent Gaussian random variable, i.e each $A_{ij} \in N(0,1)$ with distribution function $\phi(t)$ where

$$\phi(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}.$$

If m is large enough, depending on s, δ and N is will with large probability have the RIP property with constant $\delta_s = \delta$.

Theorem 4.5 There is a constant $C > 0$ such that for any $\epsilon > 0$, and $0 < \delta \le 1$, the following holds: If

$$m > C\frac{s}{\delta^2}\log(N/(s\epsilon)) + \text{ lower order terms},$$

then the Gaussian $m \times N$ matrix $\mathbf{A}$ (as above) satisfy RIP

RIP for Gaussian Matrices

Let random matrix $\mathbf{A} = (A_{ij})_{1 \le i \le m, 1 \le j \le N}$ where each A_{ij} is an independent Gaussian random variable, i.e each $A_{ij} \in N(0, 1)$ with distribution function $\phi(t)$ where

$$\phi(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}.$$

If m is large enough, depending on s, δ and N is will with large probability have the RIP property with constant $\delta_s = \delta$.

Theorem 4.5 There is a constant $C > 0$ such that for any $\epsilon > 0$, and $0 < \delta \le 1$, the following holds: If

$$m > C \frac{s}{\delta^2} \log(N/(s\epsilon) + \text{ lower order terms},$$

More precisely

$$m > \frac{2s \log(N/s) + 4s \log\log(N/s)}{\delta_s - \log(1 + \delta_s)} + lowerorderterms.$$

then the Gaussian $m \times N$ matrix $\mathbf{A}$ (as above) satisfy RIP

RIP for Gaussian Matrices

Beginning of the proof of Theorem 4.5:

Let $\mathbf{x}$ be a vector in R^N with length $\|\mathbf{x}\|_2 = 1$ and define the random variable $T = \|\mathbf{A}\mathbf{x}\|_2^2/m$.

RIP for Gaussian Matrices

Beginning of the proof of Theorem 4.5:

Let $\mathbf{x}$ be a vector in R^N with length $\|\mathbf{x}\|_2 = 1$ and define the random variable $T = \|\mathbf{Ax}\|_2^2/m$. The distribution function $\psi(\tau)$ can be determined exactly:

RIP for Gaussian Matrices

Beginning of the proof of Theorem 4.5:

Let $\mathbf{x}$ be a vector in R^N with length $\|\mathbf{x}\|_2 = 1$ and define the random variable $T = \|\mathbf{Ax}\|_2^2/m$. The distribution function $\psi(\tau)$ can be determined exactly:

- Let $Z = \mathbf{Ax}$ be the random columnvector with elements $Z_j, 1 \leq j \leq m$.

RIP for Gaussian Matrices

Beginning of the proof of Theorem 4.5:

Let $\mathbf{x}$ be a vector in R^N with length $\|\mathbf{x}\|_2 = 1$ and define the random variable $T = \|\mathbf{Ax}\|_2^2/m$. The distribution function $\psi(\tau)$ can be determined exactly:

- Let $Z = \mathbf{Ax}$ be the random columnvector with elements $Z_j, 1 \leq j \leq m$.
- Observe that Z_j are independent randomvariables with distribution function

$$\phi(t) = \frac{1}{\sqrt{\pi}} e^{-t^2/2}.$$

RIP for Gaussian Matrices

Beginning of the proof of Theorem 4.5:

Let $\mathbf{x}$ be a vector in R^N with length $\|\mathbf{x}\|_2 = 1$ and define the random variable $T = \|\mathbf{Ax}\|_2^2/m$. The distribution function $\psi(\tau)$ can be determined exactly:

- Let $Z = \mathbf{Ax}$ be the random columnvector with elements $Z_j, 1 \leq j \leq m$.
- Observe that Z_j are independent randomvariables with distribution function

$$\phi(t) = \frac{1}{\sqrt{\pi}} e^{-t^2/2}.$$

- Write a vector-distribution function of $(Z_1, \ldots, \mathbf{z}_m)$

$$\phi((t_1 \ldots, t_m) = (2\pi)^{-m/2} e^{-\sum_i t_i^2/2}.$$

- Set $\tau = \sum_i t_i^2/m$. It follows that T has distribution function

$$\Psi(\tau) = c_m \tau^{\frac{m}{2}-1} \mathrm{e}^{-m\tau/2},$$

for some normalization constant c_m.

RIP for Gaussian Matrices

We take the logarithm of $\Psi(\tau)$ analyse for maximimum and estimate the second derivative.

RIP for Gaussian Matrices

We take the logarithm of $\Psi(\tau)$ analyse for maximimum and estimate the second derivative.

- $$\log \Psi(\tau) = \log c_m + (\frac{m}{2} - 1) \log \tau - \frac{m}{2}\tau,$$

RIP for Gaussian Matrices

We take the logarithm of $\Psi(\tau)$ analyse for maximimum and estimate the second derivative.

- $$\log \Psi(\tau) = \log c_m + (\frac{m}{2} - 1) \log \tau - \frac{m}{2}\tau,$$

- derivative $\Psi(\tau)$

$$\frac{d}{dx} \log \Psi(\tau) = (\frac{m}{2} - 1)\frac{1}{\tau} - \frac{m}{2},$$

RIP for Gaussian Matrices

We take the logarithm of $\Psi(\tau)$ analyse for maximimum and estimate the second derivative.

- $$\log \Psi(\tau) = \log c_m + (\frac{m}{2} - 1) \log \tau - \frac{m}{2}\tau,$$
- derivative $\Psi(\tau)$
$$\frac{d}{dx} \log \Psi(\tau) = (\frac{m}{2} - 1)\frac{1}{\tau} - \frac{m}{2},$$
- second derivative of $\Psi(\tau)$
$$\frac{d^2}{dx^2} \log \Psi(\tau) = -(\frac{m}{2} - 1)\frac{1}{\tau^2}.$$

We conclude
Lemma 5.1 $\Psi(\tau)$ has maximum value

$$\Psi_{max} = \Psi_{max,m} \sim \sqrt{m}/\sqrt{2\pi} \text{ at } \tau_m = 1 - \frac{2}{m}.$$

Furthermore

$$\int_{|\tau-1|>\delta} \Psi(\tau) d\tau \sim \frac{\sqrt{2}e^{-\frac{m}{2}(\delta - \log(1+\delta))}}{\sqrt{m\pi}\delta}.$$

$$\int_{\tau > 1+\delta} \Psi(\tau) d\tau$$

$$
\begin{aligned}
&\int_{\tau > 1+\delta} \Psi(\tau)\, d\tau \\
\sim \Psi(1+\delta) &\int_{\tau > 1+\delta} e^{-\left|\frac{d \log \Psi(1+\delta)}{d\tau}\right| \tau}\, d\tau
\end{aligned}
$$

$$\int_{\tau>1+\delta} \Psi(\tau) d\tau$$

$$\sim \Psi(1+\delta) \int_{\tau>1+\delta} e^{-\left|\frac{d \log \Psi(1+\delta)}{d\tau}\right| \tau} d\tau$$

$$= \frac{\Psi(1+\delta)}{\left|\frac{d \log \Psi(1+\delta)}{d\tau}\right|}$$

$$
\begin{aligned}
&\int_{\tau > 1+\delta} \Psi(\tau) d\tau \\
\sim \Psi(1+\delta) &\int_{\tau > 1+\delta} e^{-\left|\frac{d \log \Psi(1+\delta)}{d\tau}\right|\tau} d\tau \\
&= \frac{\Psi(1+\delta)}{\left|\frac{d \log \Psi(1+\delta)}{d\tau}\right|} \\
\sim \frac{\Psi(1)}{\Psi_{max}} \Psi_{max} &\frac{\Psi(1+\delta)}{\Psi(1)} \frac{1}{\left|\frac{d \log \Psi(1+\delta)}{d\tau}\right|}
\end{aligned}
$$

$$
\begin{aligned}
& \int_{\tau > 1+\delta} \Psi(\tau) d\tau \\
\sim \Psi(1+\delta) & \int_{\tau > 1+\delta} e^{-\left|\frac{d \log \Psi(1+\delta)}{d\tau}\right| \tau} d\tau \\
& = \frac{\Psi(1+\delta)}{\left|\frac{d \log \Psi(1+\delta)}{d\tau}\right|} \\
\sim \frac{\Psi(1)}{\Psi_{max}} \Psi_{max} & \frac{\Psi(1+\delta)}{\Psi(1)} \frac{1}{\left|\frac{d \log \Psi(1+\delta)}{d\tau}\right|} \\
\sim 1 \times \frac{\sqrt{m}}{\sqrt{2\pi}(1+\delta)} & \frac{\left((1+\delta) e^{-\delta}\right)^{\frac{m}{2}}}{\left(\frac{m}{2}\left(1 - \frac{1}{1+\delta}\right) + \frac{1}{1+\delta}\right)}
\end{aligned}
$$

$$
\begin{aligned}
&\int_{\tau>1+\delta} \Psi(\tau) d\tau \\
\sim \Psi(1+\delta) &\int_{\tau>1+\delta} e^{-\left|\frac{d \log \Psi(1+\delta)}{d\tau}\right|\tau} d\tau \\
&= \frac{\Psi(1+\delta)}{\left|\frac{d \log \Psi(1+\delta)}{d\tau}\right|} \\
&\sim \frac{\Psi(1)}{\Psi_{max}} \Psi_{max} \frac{\Psi(1+\delta)}{\Psi(1)} \frac{1}{\left|\frac{d \log \Psi(1+\delta)}{d\tau}\right|} \\
&\sim 1 \times \frac{\sqrt{m}}{\sqrt{2\pi}(1+\delta)} \frac{\left((1+\delta)e^{-\delta}\right)^{\frac{m}{2}}}{\left(\frac{m}{2}(1-\frac{1}{1+\delta}) + \frac{1}{1+\delta}\right)} \\
&\sim \frac{\sqrt{2} e^{-\frac{m}{2}(\delta - \log(1+\delta))}}{\sqrt{m\pi}\delta + \frac{2}{\sqrt{m}}}.
\end{aligned}
$$

The remaining part the integral over $0 < \tau < 1 - \delta$ will be much smaller:

The remaining part the integral over $0 < \tau < 1 - \delta$ will be much smaller:

$$\int_{0<\tau<1-\delta} \Psi(\tau) d\tau$$

The remaining part the integral over $0 < \tau < 1 - \delta$ will be much smaller:

$$\int_{0<\tau<1-\delta} \Psi(\tau) d\tau$$

$$\sim 1 \times \frac{\sqrt{m}}{\sqrt{2\pi}(1-\delta)} \frac{\left((1-\delta)e^{\delta}\right)^{\frac{m}{2}}}{\left(\frac{m}{2}\left(\frac{1}{1-\delta}-1\right)-\frac{1}{1-\delta}\right)}$$

The remaining part the integral over $0 < \tau < 1 - \delta$ will be much smaller:

$$\int_{0<\tau<1-\delta} \Psi(\tau) d\tau$$

$$\sim 1 \times \frac{\sqrt{m}}{\sqrt{2\pi}(1-\delta)} \frac{\left((1-\delta)e^{\delta}\right)^{\frac{m}{2}}}{\left(\frac{m}{2}\left(\frac{1}{1-\delta} - 1\right) - \frac{1}{1-\delta}\right)}$$

$$\sim \frac{\sqrt{2}e^{-\frac{m}{2}(-\delta-\log(1-\delta))}}{\sqrt{m\pi}\delta + \frac{2}{\sqrt{m}}}.$$

The remaining part the integral over $0 < \tau < 1 - \delta$ will be much smaller:

$$\int_{0<\tau<1-\delta} \Psi(\tau) d\tau$$

$$\sim 1 \times \frac{\sqrt{m}}{\sqrt{2\pi}(1-\delta)} \frac{\left((1-\delta)e^{\delta}\right)^{\frac{m}{2}}}{\left(\frac{m}{2}\left(\frac{1}{1-\delta}-1\right)-\frac{1}{1-\delta}\right)}$$

$$\sim \frac{\sqrt{2}e^{-\frac{m}{2}(-\delta-\log(1-\delta))}}{\sqrt{m\pi}\delta+\frac{2}{\sqrt{m}}}.$$

And we have

$$\delta - \log(1+\delta) < -\delta - \log(1-\delta).$$

To see this set $f(t) = t - \log(1+t)$, then

$$(f(t) - f(-t))' = 2 - \frac{2}{1-t^2} < 0$$

To see this set $f(t) = t - \log(1 + t)$, then

$$(f(t) - f(-t))' = 2 - \frac{2}{1 - t^2} < 0$$

Thus for $\delta > 0$:

$$\delta - \log(1 + \delta) < -\delta - \log(1 - \delta).$$

To see this set $f(t) = t - \log(1+t)$, then

$$(f(t) - f(-t))' = 2 - \frac{2}{1-t^2} < 0$$

Thus for $\delta > 0$:

$$\delta - \log(1+\delta) < -\delta - \log(1-\delta).$$

This implies the integral over $0 < \tau < 1 - \delta$ neglectible for large m.

Thus for one fixed point probability for the RIP estimate does not hold, is of magnitude

$$\frac{\sqrt{2}e^{-\frac{m}{2}(\delta-\log(1+\delta))}}{\sqrt{m\pi}\delta},$$

for large m.

We summarise in the

Lemma 5.1 Let $\mathbf{x}$ any fixed vector in $\mathbb{R}^N$ an let $\mathbf{A}$ be a Gaussian radom matrix then we have the RiP estimate

$$|\|\mathbf{A}\mathbf{x}\|_2^2 - \|\mathbf{x}\|_2^2| \leq \delta \|\mathbf{x}\|_2^2$$

with probablility $1 - \epsilon$ where

$$\epsilon = \frac{\sqrt{2} e^{-\frac{m}{2}(\delta - \log(1+\delta))}}{\sqrt{m\pi}\delta}$$

We summarise in the

Lemma 5.1 Let $\mathbf{x}$ any fixed vector in $\mathbb{R}^N$ an let $\mathbf{A}$ be a Gaussian radom matrix then we have the RiP estimate

$$\left| \|\mathbf{A}\mathbf{x}\|_2^2 - \|\mathbf{x}\|_2^2 \right| \leq \delta \|\mathbf{x}\|_2^2$$

with probablility $1 - \epsilon$ where

$$\epsilon = \frac{\sqrt{2} e^{-\frac{m}{2}(\delta - \log(1+\delta))}}{\sqrt{m\pi}\delta}$$

What has this to do with RIP for s-sparse vectors?

We summarise in the

Lemma 5.1 Let $\mathbf{x}$ any fixed vector in $\mathbb{R}^N$ an let $\mathbf{A}$ be a Gaussian radom matrix then we have the RiP estimate

$$|\|\mathbf{A}\mathbf{x}\|_2^2 - \|\mathbf{x}\|_2^2| \leq \delta \|\mathbf{x}\|_2^2$$

with probablility $1 - \epsilon$ where

$$\epsilon = \frac{\sqrt{2}e^{-\frac{m}{2}(\delta - \log(1+\delta))}}{\sqrt{m\pi}\delta}$$

What has this to do with RIP for s-sparse vectors?

- Firs let $s = 1$.

We summarise in the

Lemma 5.1 Let $\mathbf{x}$ any fixed vector in $\mathbb{R}^N$ an let $\mathbf{A}$ be a Gaussian radom matrix then we have the RiP estimate

$$|\|\mathbf{Ax}\|_2^2 - \|\mathbf{x}\|_2^2| \leq \delta \|\mathbf{x}\|_2^2$$

with probablility $1 - \epsilon$ where

$$\epsilon = \frac{\sqrt{2} e^{-\frac{m}{2}(\delta - \log(1+\delta))}}{\sqrt{m\pi}\delta}$$

What has this to do with RIP for s-sparse vectors?

- Firs let $s = 1$.

We summarise in the

Lemma 5.1 Let $\mathbf{x}$ any fixed vector in $\mathbb{R}^N$ an let $\mathbf{A}$ be a Gaussian radom matrix then we have the RiP estimate

$$|\|\mathbf{A}\mathbf{x}\|_2^2 - \|\mathbf{x}\|_2^2| \leq \delta \|\mathbf{x}\|_2^2$$

with probablility $1 - \epsilon$ where

$$\epsilon = \frac{\sqrt{2} e^{-\frac{m}{2}(\delta - \log(1+\delta))}}{\sqrt{m\pi}\delta}$$

What has this to do with RIP for s-sparse vectors?

- Firs let $s = 1$.
  Then $X(s)$ is the unioun of N one dimensional subspaces each spanned by a singel point in $\{\mathbf{x}_i\}_1^N$ on the unit sphere.

We summarise in the

Lemma 5.1 Let $\mathbf{x}$ any fixed vector in $\mathbb{R}^N$ an let $\mathbf{A}$ be a Gaussian radom matrix then we have the RiP estimate

$$|\|\mathbf{A}\mathbf{x}\|_2^2 - \|\mathbf{x}\|_2^2| \leq \delta\|\mathbf{x}\|_2^2$$

with probablility $1-\epsilon$ where

$$\epsilon = \frac{\sqrt{2}e^{-\frac{m}{2}(\delta - \log(1+\delta))}}{\sqrt{m\pi}\delta}$$

What has this to do with RIP for s-sparse vectors?

- Firs let $s = 1$.
  Then $X(s)$ is the unioun of N one dimensional subspaces each spanned by a singel point in $\{\mathbf{x}_i\}_1^N$ on the unit sphere.
  The probability that the RIP holds for all of them with $\delta_1 = \delta$ is larger than $1 - N\epsilon$, where ϵ and δ is as above.

We summarise in the

Lemma 5.1 Let **x** any fixed vector in $\mathbb{R}^N$ an let **A** be a Gaussian radom matrix then we have the RiP estimate

$$|\|\mathbf{A}\mathbf{x}\|_2^2 - \|\mathbf{x}\|_2^2| \leq \delta \|\mathbf{x}\|_2^2$$

with probablility $1 - \epsilon$ where

$$\epsilon = \frac{\sqrt{2} e^{-\frac{m}{2}(\delta - \log(1+\delta))}}{\sqrt{m\pi}\delta}$$

What has this to do with RIP for s-sparse vectors?

- Firs let $s = 1$.
  Then $X(s)$ is the unioun of N one dimensional subspaces each spanned by a singel point in $\{\mathbf{x}_i\}_1^N$ on the unit sphere.
  The probability that the RIP holds for all of them with $\delta_1 = \delta$ is larger than $1 - N\epsilon$, where ϵ and δ is as above.

- Let $s = 2$.

- Let $s = 2$.

- Let $s = 2$.
  Then $X(s)$ is the union of $N(N-1)/2$ two dimensional subspaces.

- Let $s = 2$.
  Then $X(s)$ is the union of $N(N-1)/2$ two dimensional subspaces.
  If the have the RIP estimate for all points: $\mathbf{x}_i, 1 \leq i \leq N$, $\mathbf{x}_i + \mathbf{x}_j, i \neq j$ and $\mathbf{x}_i - \mathbf{x}_j, i \neq j$, all together N^2 points .

- Let $s = 2$.
  Then $X(s)$ is the union of $N(N-1)/2$ two dimensional subspaces.
  If the have the RIP estimate for all points: $\mathbf{x}_i, 1 \leq i \leq N$, $\mathbf{x}_i + \mathbf{x}_j, i \neq j$ and $\mathbf{x}_i - \mathbf{x}_j, i \neq j$, all together N^2 points . Then we can use polarisation argument and get

Lemma 5.3 Let $\mathbf{A}$ be a $m \times N$ Gaussian random matrix

Lemma 5.3 Let $\mathbf{A}$ be a $m \times N$ Gaussian random matrix

Let

$$\epsilon = N^2 \frac{\sqrt{2} e^{-\frac{m}{2}(\delta - \log(1+\delta))}}{\sqrt{m\pi}\delta}.$$

Lemma 5.3 Let **A** be a $m \times N$ Gaussian random matrix
Let

$$\epsilon = N^2 \frac{\sqrt{2} e^{-\frac{m}{2}(\delta - \log(1+\delta))}}{\sqrt{m\pi}\delta}.$$

Then with probability larger than $1 - \epsilon$ will the (realisation of) the matrix **A** satisfy the estimate

Lemma 5.3 Let $\mathbf{A}$ be a $m \times N$ Gaussian random matrix
Let

$$\epsilon = N^2 \frac{\sqrt{2} e^{-\frac{m}{2}(\delta - \log(1+\delta))}}{\sqrt{m\pi}\delta}.$$

Then with probability larger than $1 - \epsilon$ will the (realisation of) the matrix $\mathbf{A}$ satisfy the estimate

$$|\|\mathbf{Ax}\|_2^2 - \|\mathbf{x}\|_2^2| \leq \delta \|\mathbf{x}\|_1^2,$$

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

Lemma 5.3 Let $\mathbf{A}$ be a $m \times N$ Gaussian random matrix
Let

$$\epsilon = N^2 \frac{\sqrt{2} e^{-\frac{m}{2}(\delta - \log(1+\delta))}}{\sqrt{m\pi}\delta}.$$

Then with probability larger than $1-\epsilon$ will the (realisation of) the matrix $\mathbf{A}$ satisfy the estimate

$$|\|\mathbf{A}\mathbf{x}\|_2^2 - \|\mathbf{x}\|_2^2| \leq \delta \|\mathbf{x}\|_1^2,$$

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

Corollary: This gives the RIP estimate with constant $\delta_s = s\delta$ with probability larger than $1-\epsilon$ provided

$$m > \frac{2}{\delta/s - \log(1-\delta/s)} (\log(N^2/\epsilon) + c) \sim \frac{4s^2}{\delta^2} (\log(N^2/\epsilon) + c).$$

We have used that $X(s)$ is containt in the ball $B_{(\sqrt{}s)} = \{\mathbf{x} : \|\mathbf{x}\|_1 \leq \sqrt{s}\}$ which is the convex hull of the points $\mathbf{x}_{\pm i} = \pm\sqrt{s}\mathbf{e}_i$.

We have used that $X(s)$ is containt in the ball $B_(\sqrt{s}) = \{\mathbf{x} : \|\mathbf{x}\|_1 \leq \sqrt{s}\}$ which is the convex hull of the points $\mathbf{x}_{\pm i} = \pm\sqrt{s}\mathbf{e}_i$.

How to get a better estimate for Gaussian random matrices?

We have used that $X(s)$ is containt in the ball $B_(\sqrt{s}) = \{\mathbf{x} : \|\mathbf{x}\|_1 \leq \sqrt{s}\}$ which is the convex hull of the points $\mathbf{x}_{\pm i} = \pm\sqrt{s}\mathbf{e}_i$.
How to get a better estimate for Gaussian random matrices?
When taking the convex hull we lost a factor $\sqrt{s}$
void this by using more points.

Recall the Definitions

We have used that $X(s)$ is containt in the ball $B_{(}\sqrt{s}) = \{\mathbf{x} : \|\mathbf{x}\|_1 \leq \sqrt{s}\}$ which is the convex hull of the points $\mathbf{x}_{\pm i} = \pm\sqrt{s}\mathbf{e}_i$.

How to get a better estimate for Gaussian random matrices?

When taking the convex hull we lost a factor $\sqrt{s}$

Avoid this by using more points.

Recall the Definitions

Let $0 < r < 1$, and let $N_s(r)$ be the maximal number of points $\{\mathbf{x}_i\}$ in a set in the unit ball $B_s = \{\mathbf{x} \in \mathbb{R}^m : \|\mathbf{x}\|_2 \leq 1$ in $\mathbb{R}^s$ such that

$$\|\mathbf{x}_i - \mathbf{x}_j\|_2 > r \text{ for all } i \neq j.$$

We will show

We will show

Lemma 5.4: Let $\{\mathbf{x}_i\}_1^M$ a set with maximal number $M = N_s(r)$ of points in the unit ball B_s.

Then the closed convex hull W of $\{\mathbf{x}_i\}$ contains $(1-r)B_s = \{x : \|\mathbf{x}\| \leq 1-r\}$.

Proof: Let $\mathbf{z}$ be a point in the unit ball B_s that is not in W.

We will show

Lemma 5.4: Let $\{\mathbf{x}_i\}_1^M$ a set with maximal number $M = N_s(r)$ of points in the unit ball B_s.

Then the closed convex hull W of $\{\mathbf{x}_i\}$ contains $(1-r)B_s = \{x : \|\mathbf{x}\| \le 1-r\}$.

Proof: Let $\mathbf{z}$ be a point in the unit ball B_s that is not in W. Then there exist a hyperplan dimension $s-1$ that separates z from W,

We will show

Lemma 5.4: Let $\{\mathbf{x}_i\}_1^M$ a set with maximal number $M = N_s(r)$ of points in the unit ball B_s.

Then the closed convex hull W of $\{\mathbf{x}_i\}$ contains $(1-r)B_s = \{x : \|\mathbf{x}\| \leq 1-r\}$.

Proof: Let $\mathbf{z}$ be a point in the unit ball B_s that is not in W.
Then there exist a hyperplan dimension $s-1$ that separates z from W,
stated otherwise there: is a linear functional $l(x)$ on R^N such that $l(z) > 1$ but $l(x) \leq 1$ for all $x \in W$.

Proof of Lemma 5.4 continued

Let $\mathbf{z}_0$ be point on the hyperplan with minimal distance to the origin and let $\mathbf{x}_{M+1} = \mathbf{z}_0 / \|\mathbf{z}\|_2$.

Proof of Lemma 5.4 continued

Let $\mathbf{z}_0$ be point on the hyperplan with minimal distance to the origin and let $\mathbf{x}_{M+1} = \mathbf{z}_0 / \|\mathbf{z}\|_2$.
Then for any $\mathbf{x}_i, 1 \leq i \leq M$

$$\|\mathbf{x}_i - \mathbf{x}_{M+1}\|_2 \geq \|\mathbf{z}_0 - \mathbf{x}_{M+1}\|_2 = 1 - \|\mathbf{z}_0\|_2 \geq 1 - \|\mathbf{z}\|_2.$$

Proof of Lemma 5.4 continued

Let $\mathbf{z}_0$ be point on the hyperplan with minimal distance to the origin and let $\mathbf{x}_{M+1} = \mathbf{z}_0 / \|\mathbf{z}\|_2$.
Then for any $\mathbf{x}_i, 1 \le i \le M$

$$\|\mathbf{x}_i - \mathbf{x}_{M+1}\|_2 \ge \|\mathbf{z}_0 - \mathbf{x}_{M+1}\|_2 = 1 - \|\mathbf{z}_0\|_2 \ge 1 - \|\mathbf{z}\|_2.$$

Since $M = N_s(r)$ was the maximal number of points with distanse $> r$ we have $\|\mathbf{x}_i - \mathbf{x}_{M+1}\|_2 \le r$.

Proof of Lemma 5.4 continued

Let $\mathbf{z}_0$ be point on the hyperplan with minimal distance to the origin and let $\mathbf{x}_{M+1} = \mathbf{z}_0/\|\mathbf{z}\|_2$.
Then for any $\mathbf{x}_i, 1 \leq i \leq M$

$$\|\mathbf{x}_i - \mathbf{x}_{M+1}\|_2 \geq \|\mathbf{z}_0 - \mathbf{x}_{M+1}\|_2 = 1 - \|\mathbf{z}_0\|_2 \geq 1 - \|\mathbf{z}\|_2.$$

Since $M = N_s(r)$ was the maximal number of points with distanse $> r$ we have $\|\mathbf{x}_i - \mathbf{x}_{M+1}\|_2 \leq r$. We conclude that $1 - \|\mathbf{z}\|_2 \leq r$ which is the same as $\|\mathbf{z}\|_2 \geq 1 - r$.

Proof of Lemma 5.4 continued

Let $\mathbf{z}_0$ be point on the hyperplan with minimal distance to the origin and let $\mathbf{x}_{M+1} = \mathbf{z}_0 / \|\mathbf{z}\|_2$.
Then for any $\mathbf{x}_i, 1 \leq i \leq M$

$$\|\mathbf{x}_i - \mathbf{x}_{M+1}\|_2 \geq \|\mathbf{z}_0 - \mathbf{x}_{M+1}\|_2 = 1 - \|\mathbf{z}_0\|_2 \geq 1 - \|\mathbf{z}\|_2.$$

Since $M = N_s(r)$ was the maximal number of points with distanse $> r$ we have $\|\mathbf{x}_i - \mathbf{x}_{M+1}\|_2 \leq r$. We conclude that $1 - \|\mathbf{z}\|_2 \leq r$ which is the same as $\|\mathbf{z}\|_2 \geq 1 - r$.
This completes the proof of Lemma 5.4

Finishing the proof of Theorem 4.5:

Recall Theorem 4.5: There is a constant $C > 0$ such for any $\epsilon > 0, \delta > 0$ if we let

$$m > C\frac{s}{\delta^2}\log(Ne/(s\epsilon)).$$

Then the Gaussian $m \times N$ matrix $\mathbf{A}$ (as above) satisfy RIP constant $\delta_s = \delta$ with a probability not less than $1 - \epsilon$

Finishing the proof:

Finishing the proof of Theorem 4.5:
Recall Theorem 4.5: There is a constant $C > 0$ such for any $\epsilon > 0, \delta > 0$ if we let

$$m > C \frac{s}{\delta^2} \log(Ne/(s\epsilon)).$$

Then the Gaussian $m \times N$ matrix $\mathbf{A}$ (as above) satisfy RIP constant $\delta_s = \delta$ with a probability not less than $1 - \epsilon$

Finishing the proof:
Let $0 < r < 1$, to be optimized later.

Finishing the proof of Theorem 4.5:

Recall Theorem 4.5: There is a constant $C > 0$ such for any $\epsilon > 0, \delta > 0$ if we let

$$m > C\frac{s}{\delta^2}\log(Ne/(s\epsilon)).$$

Then the Gaussian $m \times N$ matrix $\mathbf{A}$ (as above) satisfy RIP constant $\delta_s = \delta$ with a probability not less than $1 - \epsilon$

Finishing the proof:

Let $0 < r < 1$, to be optimized later.

For each index subset S in $\{1, \ldots, N\}$ of length s we find the $M = N_s(r)$ points $\{\mathbf{x}_{S,i}\}_1^M$ supported on S, with norm $\|\mathbf{x}_{S,i}\|_2 = 1$ and such that

Finishing the proof of Theorem 4.5:

Recall Theorem 4.5: There is a constant $C > 0$ such for any $\epsilon > 0, \delta > 0$ if we let

$$m > C\frac{s}{\delta^2}\log(Ne/(s\epsilon)).$$

Then the Gaussian $m \times N$ matrix $\mathbf{A}$ (as above) satisfy RIP constant $\delta_s = \delta$ with a probability not less than $1 - \epsilon$

Finishing the proof:

Let $0 < r < 1$, to be optimized later.

For each index subset S in $\{1, \ldots, N\}$ of length s we find the $M = N_s(r)$ points $\{\mathbf{x}_{S,i}\}_1^M$ supported on S, with norm $\|\mathbf{x}_{S,i}\|_2 = 1$ and such that the convex hull of those points contains the ball of radus $1 - r$ around 0 of points supported in S,

i.e points x supported on S an with norm $\|\mathbf{x}\|_2 = 1 - r$.

We will use the RIP estimate for all the points: $\mathbf{x}_{S,i}, 1 \leq i \leq M$, $\mathbf{x}_{S,i} + \mathbf{x}_{S,j}, i \neq j$ and $\mathbf{x}_{S,i} - \mathbf{x}_{S,j}, i \neq j$ according to Lemma 5.1.

We will use the RIP estimate for all the points: $\mathbf{x}_{S,i}, 1 \leq i \leq M$, $\mathbf{x}_{S,i} + \mathbf{x}_{S,j}, i \neq j$ and $\mathbf{x}_{S,i} - \mathbf{x}_{S,j}, i \neq j$ according to Lemma 5.1. Alltogether $(N_s(r))^2$ points for each subset S.

We will use the RIP estimate for all the points: $\mathbf{x}_{S,i}, 1 \leq i \leq M$, $\mathbf{x}_{S,i} + \mathbf{x}_{S,j}, i \neq j$ and $\mathbf{x}_{S,i} - \mathbf{x}_{S,j}, i \neq j$ according to Lemma 5.1. Alltogether $(N_s(r))^2$ points for each subset S.
Recall that number of such subsets is

$$\binom{N}{s} \sim (N\mathrm{e}/s)^s/\sqrt{2\pi s}.$$

Totally we will use the RIP esitimate for the following number of points

We will use the RIP estimate for all the points: $\mathbf{x}_{S,i}, 1 \leq i \leq M$, $\mathbf{x}_{S,i} + \mathbf{x}_{S,j}, i \neq j$ and $\mathbf{x}_{S,i} - \mathbf{x}_{S,j}, i \neq j$ according to Lemma 5.1. Alltogether $(N_s(r))^2$ points for each subset S.

Recall that number of such subsets is

$$\begin{pmatrix} N \\ s \end{pmatrix} \sim (N\mathrm{e}/s)^s/\sqrt{2\pi s}.$$

Totally we will use the RIP esitimate for the following number of points

$$M_{tot} \begin{pmatrix} N \\ s \end{pmatrix} (N_s(r))^2$$
$$\sim \frac{(N\mathrm{e}/s)^s}{\sqrt{2\pi s}}(1 + \frac{2}{r})^{2s},$$

According to Lemma 5.3 we have we have the RIP estimate for all these points with constant δ with a probability larger than $1 - \epsilon$ where

We will use the RIP estimate for all the points: $\mathbf{x}_{S,i}, 1 \leq i \leq M$, $\mathbf{x}_{S,i} + \mathbf{x}_{S,j}, i \neq j$ and $\mathbf{x}_{S,i} - \mathbf{x}_{S,j}, i \neq j$ according to Lemma 5.1. Alltogether $(N_s(r))^2$ points for each subset S.
Recall that number of such subsets is

$$\left(\begin{array}{c} N \\ s \end{array} \right) \sim (N\mathrm{e}/s)^s/\sqrt{2\pi s}.$$

Totally we will use the RIP esitimate for the following number of points

$$M_{tot} \left(\begin{array}{c} N \\ s \end{array} \right) (N_s(r))^2$$

$$\sim \frac{(N\mathrm{e}/s)^s}{\sqrt{2\pi s}}(1+\frac{2}{r})^{2s},$$

According to Lemma 5.3 we have we have the RIP estimate for all these points with constant δ with a probability larger than $1-\epsilon$ where

We will use the RIP estimate for all the points: $\mathbf{x}_{S,i}, 1 \leq i \leq M$, $\mathbf{x}_{S,i} + \mathbf{x}_{S,j}, i \neq j$ and $\mathbf{x}_{S,i} - \mathbf{x}_{S,j}, i \neq j$ according to Lemma 5.1. Alltogether $(N_s(r))^2$ points for each subset S.
Recall that number of such subsets is

$$\left(\begin{array}{c} N \\ s \end{array} \right) \sim (N\mathrm{e}/s)^s/\sqrt{2\pi s}.$$

Totally we will use the RIP esitimate for the following number of points

$$M_{tot} \left(\begin{array}{c} N \\ s \end{array} \right) (N_s(r))^2$$
$$\sim \frac{(N\mathrm{e}/s)^s}{\sqrt{2\pi s}} (1 + \frac{2}{r})^{2s},$$

According to Lemma 5.3 we have we have the RIP estimate for all these points with constant δ with a probability larger than $1 - \epsilon$ where

This completes the proof of

Theorem 4.5 There is a constant $C > 0$ such that for any $\epsilon > 0$, and $0 < \delta \leq 1$, the following holds: If

$$m > \frac{2s\log(N/(s\epsilon) + 4s\log\log(N/s)}{\delta_s - \log(1+\delta_s)} + \text{ lower order terms}$$

then the Gaussian $m \times N$ matrix $\mathbf{A}$ (as above) satisfy RIP constant $\delta_s = \delta$ with a probability not less than $1 - \epsilon$.

Corollary There exit $m \times N$ matrices RIP constand δ_s if

$$m > \frac{2s\log(N/s) + 4s\log\log(N/s)}{\delta_s - \log(1+\delta_s)} + \text{ lower order terms }.$$

Construction of matrices satisfyin RIP

Construction of RIP matrices We want do a construction of an $m \times N$ matrix $\mathbf{A}$ statistfying the RIP property
Theorem 5.5 There is a constant $C > 0$ such that if $\delta_s > 0$ and

$$m > C\frac{s^2}{\delta_s^2}\log N.$$

then it is possible to construct an $m \times N$ matrix $\mathbf{A}$ statistfying the RIP property with constant δ_s.

Theorem 5.5 is a consequence of the following:

Theorem 5.5 is a consequence of the following:
Lemma 5.6 : Let $s << m$ and $N > 0$ very large and $0 < r < 1$ chosen later. Then there exists matrices $\mathbf{A}$ with column vectors $A_n, \|A_j\|_2 = 1, 1 \leq n \leq N$, such that every subset $S = A_{n_j}$ of s is a set a lineary independent vectors in $\mathbb{C}^N$, such

Theorem 5.5 is a consequence of the following:
Lemma 5.6 : Let $s << m$ and $N > 0$ very large and $0 < r < 1$ chosen later. Then there exists matrices $\mathbf{A}$ with column vectors $A_n, \|A_j\|_2 = 1, 1 \leq n \leq N$, such that every subset $S = A_{n_j}$ of s is a set a lineary independent vectors in $\mathbb{C}^N$, such
that the l_2 distance from A_{n_k} to the $span\{A_{n_j}, j < k\}$ is larger than r, for all $1 < k \leq s$. Provided m is large enough depending on s,N and r.

Proof is done by construction:
The algorithm for construction of the columns vectors $A_n, 0 \leq n \leq N$ is as follows:

1. Choose m independent columns $A_1, \ldots A_m$ in $\mathbb{C}^m$.

Proof is done by construction:
The algorithm for construction of the columns vectors $A_n, 0 \leq n \leq N$ is as follows:

1. Choose m independent columns $A_1, \dots A_m$ in $\mathbb{C}^m$.
2. Set $n =: m$.

Proof is done by construction:

The algorithm for construction of the columns vectors $A_n, 0 \leq n \leq N$ is as follows:

1. Choose m independent columns $A_1, \ldots A_m$ in $\mathbb{C}^m$.
2. Set $n =: m$.
3. Any $s-1$ - tiple of chosen column vectors span a subspace of codimension $m-s+1$.

Proof is done by construction:

The algorithm for construction of the columns vectors $A_n, 0 \leq n \leq N$ is as follows:

1. Choose m independent columns $A_1, \ldots A_m$ in $\mathbb{C}^m$.
2. Set $n =: m$.
3. Any $s-1$ - tiple of chosen column vectors span a subspace of codimension $m-s+1$.
4. Build an m-dimensional plate by intersecting each such $s-1$ - dimensional subspace of the unit ball and take direct sum with the perpendicular ball of radius r and dimension $m-s+1$.

Proof is done by construction:

The algorithm for construction of the columns vectors $A_n, 0 \leq n \leq N$ is as follows:

1. Choose m independent columns $A_1, \ldots A_m$ in $\mathbb{C}^m$.
2. Set $n =: m$.
3. Any $s-1$ - tiple of chosen column vectors span a subspace of codimension $m-s+1$.
4. Build an m-dimensional plate by intersecting each such $s-1$ - dimensional subspace of the unit ball and take direct sum with the perpendicular ball of radius r and dimension $m-s+1$.
5. The volume of such a plate can be estimated from above..

Proof is done by construction:

The algorithm for construction of the columns vectors $A_n, 0 \leq n \leq N$ is as follows:

1. Choose m independent columns $A_1, \dots A_m$ in $\mathbb{C}^m$.
2. Set $n =: m$.
3. Any $s-1$ - tiple of chosen column vectors span a subspace of codimension $m-s+1$.
4. Build an m-dimensional plate by intersecting each such $s-1$ - dimensional subspace of the unit ball and take direct sum with the perpendicular ball of radius r and dimension $m-s+1$.
5. The volume of such a plate can be estimated from above..
6. Choose the new column A_n on the unite sphere, but outside each such plate.

- Any $s-1$ - tiple of previously chosen column vectors together with A_n will span the whole space.

- Any $s-1$ - tiple of previously chosen column vectors together with A_n will span the whole space.
- If $n < N$ and the union of the all plates constructed so far does not cover the unit ball in C^N: set $n =: n+1$ and return to step nr 3.

- Any $s-1$ - tiple of previously chosen column vectors together with A_n will span the whole space.
- If $n < N$ and the union of the all plates constructed so far does not cover the unit ball in C^N: set $n =: n+1$ and return to step nr 3.
- The construction is finished! But –

- Any $s-1$ - tiple of previously chosen column vectors together with A_n will span the whole space.
- If $n < N$ and the union of the all plates constructed so far does not cover the unit ball in C^N: set $n =: n+1$ and return to step nr 3.
- The construction is finished! But –

- Any $s-1$ - tiple of previously chosen column vectors together with A_n will span the whole space.
- If $n < N$ and the union of the all plates constructed so far does not cover the unit ball in C^N: set $n =: n+1$ and return to step nr 3.
- The construction is finished! But –
  r and m is not yet selected

Proof of RIP property of constructed matrix **A**:

- The number r has to be chosen close enougth to 1. Also, m has to be choosen so large that the union of the plates not can cover the unit ball for any $n < N$.

Proof of RIP property of constructed matrix $\mathbf{A}$:

- The number r has to be chosen close enougth to 1.
  Also, m has to be choosen so large that the union of the plates not can cover the unit ball for any $n < N$.
  This is ensured as long as

  (total number plates)$\times$(plate volume) $<$ (volume of the m- dimensi

Proof of RIP property of constructed matrix **A**:

- The number r has to be chosen close enougth to 1.
  Also, m has to be choosen so large that the union of the plates not can cover the unit ball for any $n < N$.
  This is ensured as long as

  (total number plates)$\times$(plate volume) $<$ (volume of the m- dimensi

- the parameter r has to be chosen close enough to 1 to ensure that the RIP property to hold:

Proof of RIP property of constructed matrix **A**:

- The number r has to be chosen close enougth to 1.
  Also, m has to be choosen so large that the union of the plates not can cover the unit ball for any $n < N$.
  This is ensured as long as

  (total number plates)$\times$(plate volume) $<$ (volume of the m- dimensi

- the parameter r has to be chosen close enough to 1 to ensure that the RIP property to hold:
- let $W = span\{A_{n_j}, j < k\}$ then $\mathbf{u} = A_{n_k}$ is uniquely decomposed

$$\mathbf{u} = \mathbf{u}_W + \mathbf{u}_{W^\perp}$$

.

Proof of RIP property of constructed matrix **A**:

- The number r has to be chosen close enougth to 1.
  Also, m has to be choosen so large that the union of the plates not can cover the unit ball for any $n < N$.
  This is ensured as long as

  (total number plates)$\times$(plate volume) $<$ (volume of the m- dimensi

- the parameter r has to be chosen close enough to 1 to ensure that the RIP property to hold:
- let $W = span\{A_{n_j}, j < k\}$ then $\mathbf{u} = A_{n_k}$ is uniquely decomposed

$$\mathbf{u} = \mathbf{u}_W + \mathbf{u}_{W^\perp}$$

  .
- set

$$\mathbf{n}_{j_1,\ldots,j_k} = \frac{\mathbf{u}_{W^\perp}}{\|\mathbf{u}_{W^\perp}\|_2} \text{ and } \mathbf{v}_{j_1,\ldots,j_k} = \mathbf{u} - \mathbf{n}_{j_1,\ldots,j_k}$$

- $$\mathbf{U} = (\mathbf{n}_{j_1,\dots,j_k})_{k=1}^{s}$$

  will be an orthonormal matrix with orthonormal columns and

- $$\mathbf{U} = (\mathbf{n}_{j_1,\ldots,j_k})_{k=1}^{s}$$

  will be an orthonormal matrix with orthonormal columns and

- $$\mathbf{V} = (\mathbf{v}_{j_1,\ldots,j_k})_{k=1}^{s}$$

  will be a matrix whose columns has norms

- $$\mathbf{U} = (\mathbf{n}_{j_1,\ldots,j_k})_{k=1}^{s}$$

  will be an orthonormal matrix with orthonormal columns and

- $$\mathbf{V} = (\mathbf{v}_{j_1,\ldots,j_k})_{k=1}^{s}$$

  will be a matrix whose columns has norms

  $$\|\mathbf{v}_{j_1,\ldots,\jmath_k}\|_2 \leq \sqrt{2(1-r)}.$$

- $$\mathbf{U} = (\mathbf{n}_{j_1,\dots,j_k})_{k=1}^{s}$$

  will be an orthonormal matrix with orthonormal columns and

- $$\mathbf{V} = (\mathbf{v}_{j_1,\dots,j_k})_{k=1}^{s}$$

  will be a matrix whose columns has norms

  $$\|\mathbf{v}_{j_1,\dots,j_k}\|_2 \leq \sqrt{2(1-r)}.$$

- if $\mathbf{x}$ is supported on $\{j_1,\dots,j_s\}$
  will we get $\mathbf{Ax} = \mathbf{Ux} + \mathbf{Vx}$ and

  $$\|\mathbf{Ux}\|_2 - \|\mathbf{Vx}\|_2 \leq \|\mathbf{Ax}\|_2 \leq \|\mathbf{Ux}\|_2 + \|\mathbf{Vx}\|_2$$

- Since $\|\mathbf{Ux}\|_2 = \|\mathbf{x}\|_2$ will we get

$$\left|\|\mathbf{Ax}\|_2^2 - \|\mathbf{x}\|_2^2\right| \leq 2\|\mathbf{x}\|_2\|\mathbf{Vx}\|_2 + \|\mathbf{Vx}\|_2^2.$$

- Since $\|\mathbf{Ux}\|_2 = \|\mathbf{x}\|_2$ will we get

$$\left|\|\mathbf{Ax}\|_2^2 - \|\mathbf{x}\|_2^2\right| \leq 2\|\mathbf{x}\|_2\|\mathbf{Vx}\|_2 + \|\mathbf{Vx}\|_2^2.$$

By the triangle inequality we get

$$\|\mathbf{Vx}\|_2 \leq \sqrt{2(1-r)}\|\mathbf{x}\|_1 \leq \sqrt{2s(1-r)}\|\mathbf{x}\|_2.$$

- Since $\|\mathbf{Ux}\|_2 = \|\mathbf{x}\|_2$ will we get

$$|\|\mathbf{Ax}\|_2^2 - \|\mathbf{x}\|_2^2| \leq 2\|\mathbf{x}\|_2\|\mathbf{Vx}\|_2 + \|\mathbf{Vx}\|_2^2.$$

  By the triangle inequality we get

$$\|\mathbf{Vx}\|_2 \leq \sqrt{2(1-r)}\|\mathbf{x}\|_1 \leq \sqrt{2s(1-r)}\|\mathbf{x}\|_2.$$

  Thus

$$|\|\mathbf{Ax}\|_2^2 - \|\mathbf{x}\|_2^2| \leq (2\sqrt{2s(1-r} + 2s(1-r))\|\mathbf{x}\|_2^2.$$

- Finally we choose r such that $2\sqrt{2s(1-r)} + 2s(1-r) \leq \delta_s$

Careful estimate of the covering and the conditions for m in Lemma 5.6 will give us the estimate how large m needs to be in Theorem 5.5