

Mathematics of Compressive Sensing: Random matrices are restricted isometries

Justin Romberg

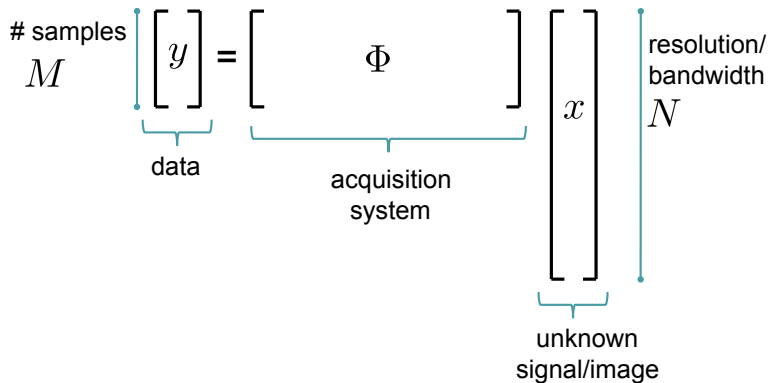
Georgia Tech, School of ECE

Tsinghua Univeristy

October 18, 2013

Beijing, China

Acquisition as linear algebra



- Small number of samples = underdetermined system
Impossible to solve in general
- If x is *sparse* and Φ is *diverse*, then these systems can be “inverted”

Yesterday...

Theorem: Let Φ be an $M \times N$ matrix that is an approximate isometry for $3S$ -sparse vectors. Let x_0 be an S -sparse vector, and suppose we observe $y = \Phi x_0$. Given y , the solution to

$$\min_x \|x\|_1 \quad \text{subject to} \quad \Phi x = y$$

is *exactly* x_0 .

Today...

We will prove (almost from top to bottom):

- That an $M \times N$ iid Gaussian random matrix satisfies

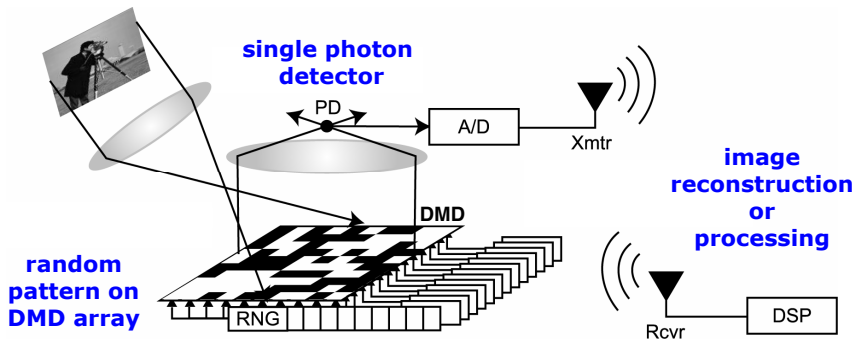
$$(1 - \delta)\|x\|_2^2 \leq \|\Phi x\|_2^2 \leq (1 + \delta)\|x\|_2^2 \quad \forall 2S\text{-sparse } x \quad (1)$$

with (extraordinarily) high probability when

$$M \geq \text{Const} \cdot S \log(N/S)$$

This is easier than it looks at first — we will just put together a bunch of easy results from probability theory.

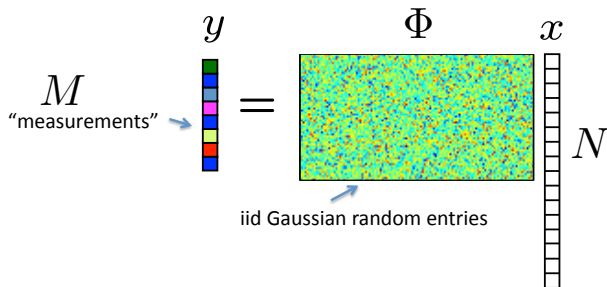
Rice single pixel camera



(Duarte, Davenport, Takhar, Laska, Sun, Kelly, Baraniuk '08)

Gaussian random matrices

- Each entry of Φ is iid $\text{Normal}(0, M^{-1})$

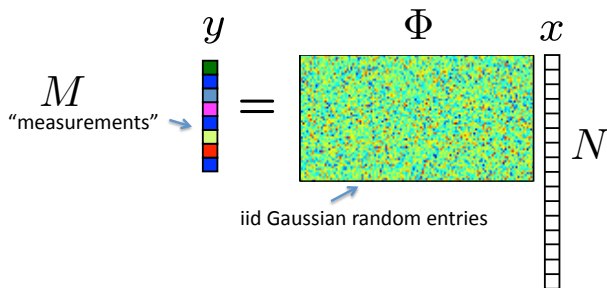


- For *any fixed* $x \in \mathbb{R}^N$, each measurement is

$$y_m \sim \text{Normal}(0, \|x\|_2^2/M)$$

Gaussian random matrices

- Each entry of Φ is iid Normal($0, M^{-1}$)



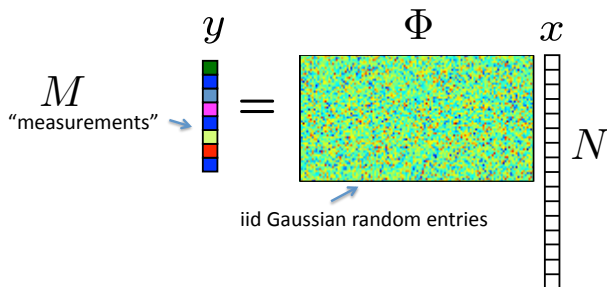
- For *any fixed* $x \in \mathbb{R}^N$, we have

$$\mathbb{E}[\|\Phi x\|_2^2] = \|x\|_2^2$$

the mean of the measurement energy is exactly $\|x\|_2^2$

Gaussian random matrices

- Each entry of Φ is iid $\text{Normal}(0, M^{-1})$

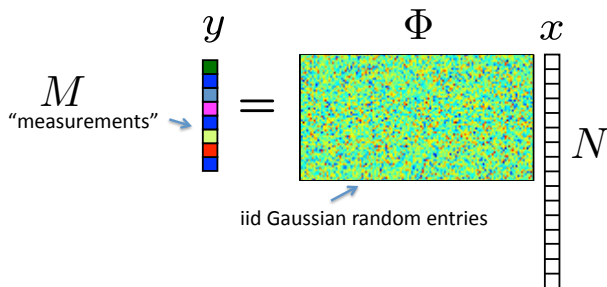


- For *any fixed* $x \in \mathbb{R}^N$, we have

$$\mathbb{P} \left\{ \left| \|\Phi x\|_2^2 - \|x\|_2^2 \right| < \delta \|x\|_2^2 \right\} \geq 1 - 2e^{-M\delta^2/8}$$

Gaussian random matrices

- Each entry of Φ is iid $\text{Normal}(0, M^{-1})$



- For *all* $2S$ -sparse $x \in \mathbb{R}^N$, we have

$$\mathbb{P} \left\{ \max_x \left| \|\Phi x\|_2^2 - \|x\|_2^2 \right| < \delta \|x\|_2^2 \right\} \geq 1 - 2e^{c \cdot S \log(N/S)} e^{-M\delta^2/8}$$

So we can make this probability close to 1 by taking

$$M \geq \text{Const} \cdot S \log(N/S)$$

Random projection of a fixed vector

For Gaussian random Φ operating on a *fixed* $x \in \mathbb{R}^N$

$$\|\Phi x\|_2^2 \approx \|x\|_2^2$$

Theorem: Let Φ be an $M \times N$ matrix whose entries are iid Gaussian

$$\Phi_{i,j} \sim \text{Normal}(0, 1/M).$$

Set $v = \Phi x$. Then

$$\mathbb{E} \|v\|_2^2 = \|x\|_2^2,$$

Random projection of a fixed vector

For Gaussian random Φ operating on a *fixed* $x \in \mathbb{R}^N$

$$\|\Phi x\|_2^2 \approx \|x\|_2^2$$

Theorem: Let Φ be an $M \times N$ matrix whose entries are iid Gaussian

$$\Phi_{i,j} \sim \text{Normal}(0, 1/M).$$

Set $v = \Phi x$. Then

$$\mathbb{E} \|v\|_2^2 = \|x\|_2^2,$$

as

$$\mathbb{E} \left[\sum_{m=1}^M v_m^2 \right] = \sum_{m=1}^M \mathbb{E}[v_m^2] = \sum_{m=1}^M \frac{1}{M} \|x\|_2^2 = \|x\|_2^2,$$

since $v_m = \langle x, \phi_m \rangle \sim \text{Normal}(0, M^{-1} \|x\|_2^2)$

Random projection of a fixed vector

For Gaussian random Φ operating on a *fixed* $x \in \mathbb{R}^N$

$$\|\Phi x\|_2^2 \approx \|x\|_2^2$$

Theorem: Let Φ be an $M \times N$ matrix whose entries are iid Gaussian

$$\Phi_{i,j} \sim \text{Normal}(0, 1/M).$$

Set $v = \Phi x$. Then

$$\mathbb{E} \|v\|_2^2 = \|x\|_2^2,$$

and for any $0 < \delta \leq 1$

$$\begin{aligned} \mathbb{P} \left\{ \left| \|v\|_2^2 - \|x\|_2^2 \right| > \delta \right\} &\leq 2 \exp \left(-\frac{(\delta^2 - \delta^3)M}{4} \right) \\ &\leq 2 \exp(-\delta^2 M/8) \end{aligned}$$

for $\delta \leq 1/2$.

The Markov inequality

Let Y be a positive random variable. Then for all $t > 0$

$$P \{Y \geq t\} \leq \frac{E[Y]}{t}$$

The Markov inequality

Let Y be a positive random variable. Then for all $t > 0$

$$\mathbb{P}\{Y \geq t\} \leq \frac{\mathbb{E}[Y]}{t}$$

Proof:

$$\begin{aligned}\mathbb{E}[Y] &= \int_0^{\infty} y f_Y(y) dy \\ &\geq \int_t^{\infty} y f_Y(y) dy \\ &\geq t \int_t^{\infty} f_Y(y) dy \\ &= t \mathbb{P}\{Y \geq t\}.\end{aligned}$$

The Markov inequality

Let Y be a positive random variable. Then for all $t > 0$

$$\mathbb{P}\{Y \geq t\} \leq \frac{\mathbb{E}[Y]}{t}$$

Also:

$$\mathbb{P}\{Y^2 \geq t^2\} \leq \frac{\mathbb{E}[Y^2]}{t^2}$$

$$\mathbb{P}\{Y^3 \geq t^3\} \leq \frac{\mathbb{E}[Y^3]}{t^3}$$

$$\mathbb{P}\{e^{\lambda Y} \geq e^{\lambda t}\} \leq \frac{\mathbb{E}[e^{\lambda Y}]}{e^{\lambda t}} \quad \lambda > 0$$

\vdots

$$\mathbb{P}\{\phi(Y) \geq \phi(t)\} \leq \frac{\mathbb{E}[\phi(y)]}{\phi(t)}$$

for any strictly monotonic $\phi(\cdot)$.

The Markov inequality

Let Y be a positive random variable. Then for all $t > 0$

$$P\{Y \geq t\} \leq \frac{E[Y]}{t}$$

Chernoff-type bound:

$$P\{Y \geq t\} \leq \frac{E[e^{\lambda Y}]}{e^{\lambda t}} \quad \text{for any } \lambda > 0.$$

A first upper concentration bound ...

For $v = \Phi x$, $\|x\|_2 = 1$, we have that

$$\mathbf{P} \{ \|v\|_2^2 > 1 + \delta \} \leq \frac{\mathbf{E}[e^{\lambda \|v\|_2^2}]}{e^{\lambda(1+\delta)}}$$

A first upper concentration bound ...

For $v = \Phi x$, $\|x\|_2 = 1$, we have that

$$\begin{aligned} \mathbb{P} \{ \|v\|_2^2 > 1 + \delta \} &\leq \frac{\mathbb{E}[e^{\lambda \|v\|_2^2}]}{e^{\lambda(1+\delta)}} \\ &= \frac{\mathbb{E}[e^{\lambda \sum_m v_m^2}]}{e^{\lambda(1+\delta)}} \end{aligned}$$

A first upper concentration bound ...

For $v = \Phi x$, $\|x\|_2 = 1$, we have that

$$\begin{aligned} \mathbb{P} \{ \|v\|_2^2 > 1 + \delta \} &\leq \frac{\mathbb{E}[e^{\lambda \|v\|_2^2}]}{e^{\lambda(1+\delta)}} \\ &= \frac{\mathbb{E}[e^{\lambda \sum_m v_m^2}]}{e^{\lambda(1+\delta)}} \\ &= \frac{\mathbb{E}[e^{\lambda v_1^2} e^{\lambda v_2^2} \dots e^{\lambda v_M^2}]}{e^{\lambda(1+\delta)}} \end{aligned}$$

A first upper concentration bound ...

For $v = \Phi x$, $\|x\|_2 = 1$, we have that

$$\begin{aligned} \mathbb{P} \{ \|v\|_2^2 > 1 + \delta \} &\leq \frac{\mathbb{E}[e^{\lambda \|v\|_2^2}]}{e^{\lambda(1+\delta)}} \\ &= \frac{\mathbb{E}[e^{\lambda \sum_m v_m^2}]}{e^{\lambda(1+\delta)}} \\ &= \frac{\mathbb{E}[e^{\lambda v_1^2} e^{\lambda v_2^2} \dots e^{\lambda v_M^2}]}{e^{\lambda(1+\delta)}} \\ &= \frac{\mathbb{E}[e^{\lambda v_1^2}] \mathbb{E}[e^{\lambda v_2^2}] \dots \mathbb{E}[e^{\lambda v_M^2}]}{e^{\lambda(1+\delta)}} \end{aligned}$$

A first upper concentration bound ...

For $v = \Phi x$, $\|x\|_2 = 1$, we have that

$$\begin{aligned} \mathbb{P} \{ \|v\|_2^2 > 1 + \delta \} &\leq \frac{\mathbb{E}[e^{\lambda \|v\|_2^2}]}{e^{\lambda(1+\delta)}} \\ &= \frac{\mathbb{E}[e^{\lambda \sum_m v_m^2}]}{e^{\lambda(1+\delta)}} \\ &= \frac{\mathbb{E}[e^{\lambda v_1^2} e^{\lambda v_2^2} \dots e^{\lambda v_M^2}]}{e^{\lambda(1+\delta)}} \\ &= \frac{\mathbb{E}[e^{\lambda v_1^2}] \mathbb{E}[e^{\lambda v_2^2}] \dots \mathbb{E}[e^{\lambda v_M^2}]}{e^{\lambda(1+\delta)}} \\ &= \frac{(\mathbb{E}[e^{\lambda v_1^2}])^M}{e^{\lambda(1+\delta)}} \quad (\text{since } v_m \text{ i.i.d.}) \end{aligned}$$

A first upper concentration bound ...

For $v = \Phi x$, $\|x\|_2 = 1$, we have that

$$\mathbb{P} \{ \|v\|_2^2 > 1 + \delta \} \leq \frac{(\mathbb{E}[e^{\lambda v_1^2}])^M}{e^{\lambda(1+\delta)}}, \quad v_1 \sim \text{Normal}(0, M^{-1})$$

A first upper concentration bound ...

For $v = \Phi x$, $\|x\|_2 = 1$, we have that

$$\mathbb{P} \{ \|v\|_2^2 > 1 + \delta \} \leq \frac{(\mathbb{E}[e^{\lambda v_1^2}])^M}{e^{\lambda(1+\delta)}}, \quad v_1 \sim \text{Normal}(0, M^{-1})$$

It is known that

$$\mathbb{E}[e^{\lambda v_1^2}] = \frac{1}{\sqrt{1 - 2\lambda/M}} \quad \text{for } \lambda < M/2.$$

A first upper concentration bound ...

For $v = \Phi x$, $\|x\|_2 = 1$, we have that

$$\mathbb{P} \{ \|v\|_2^2 > 1 + \delta \} \leq \frac{(\mathbb{E}[e^{\lambda v_1^2}])^M}{e^{\lambda(1+\delta)}}, \quad v_1 \sim \text{Normal}(0, M^{-1})$$

And so

$$\mathbb{P} \{ \|v\|_2^2 > 1 + \delta \} \leq \left(\frac{e^{-2\lambda(1+\delta)/M}}{1 - 2\lambda/M} \right)^{M/2} \quad \forall \lambda < M/2$$

A first upper concentration bound ...

We have

$$\mathbb{P} \{ \|v\|_2^2 > 1 + \delta \} \leq \left(\frac{e^{-2\lambda(1+\delta)/M}}{1 - 2\lambda/M} \right)^{M/2} \quad \forall \lambda < M/2$$

Choose

$$\lambda = \frac{M\delta}{2(1+\delta)}$$

(easy to see that in this case $\lambda < M/2$).

A first upper concentration bound ...

We have

$$\mathbb{P} \{ \|v\|_2^2 > 1 + \delta \} \leq \left(\frac{e^{-2\lambda(1+\delta)/M}}{1 - 2\lambda/M} \right)^{M/2} \quad \forall \lambda < M/2$$

Choose

$$\lambda = \frac{M\delta}{2(1+\delta)}$$

(easy to see that in this case $\lambda < M/2$).

And so

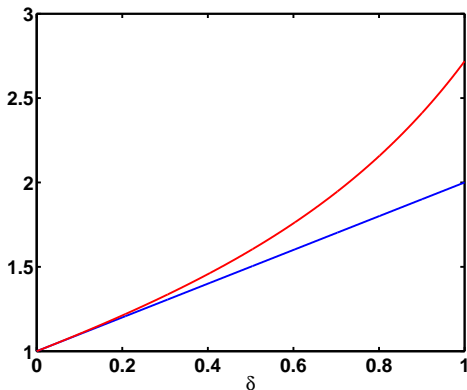
$$\mathbb{P} \{ \|v\|_2^2 > 1 + \delta \} \leq \left((1 + \delta)e^{-\delta} \right)^{M/2}.$$

The upper concentration bound

We have

$$P \{ \|v\|_2^2 > 1 + \delta \} \leq \left((1 + \delta)e^{-\delta} \right)^{M/2}.$$

blue: $1 + \delta$, red: $e^{\delta - (\delta^2 - \delta^3)/2}$



The upper concentration bound

We have

$$\mathbb{P} \{ \|v\|_2^2 > 1 + \delta \} \leq \left((1 + \delta)e^{-\delta} \right)^{M/2}.$$

and so

$$\mathbb{P} \{ \|v\|_2^2 > 1 + \delta \} \leq e^{-(\delta^2 - \delta^3)M/4}$$

The lower concentration bound

The lower bound follows the exact same sequence of steps (work them out at home!):

$$\begin{aligned} \mathbb{P} \{ \|v\|_2^2 < 1 - \delta \} &\leq \left(\frac{e^{2(1-\delta)\lambda/M}}{1 + 2\lambda/M} \right)^{M/2} \\ &\leq \left((1 - \delta)e^\delta \right)^{M/2} \quad \text{by taking } \lambda = \frac{M\delta}{2(1 - \delta)} \\ &\leq e^{-(\delta^2 - \delta^3)M/4} \end{aligned}$$

The Johnson-Lindenstrauss Lemma

We have shown that for any *fixed* $x \in \mathbb{R}^N$

$$(1 - \delta)\|x\|_2^2 \leq \|\Phi x\|_2^2 \leq (1 + \delta)\|x\|_2^2$$

with probability exceeding $1 - 2e^{-c\delta^2 M}$.

(Can take $c = 1/8$.)

A simple application of the union bound means that for any set of K vectors x_1, x_2, \dots, x_K , the above holds with probability exceeding $1 - Ke^{-\delta^2 M/8} \dots$

The Johnson-Lindenstrauss Lemma

Theorem: (J&L, 1984): Let \mathcal{Q} be a arbitrary set of Q vectors in \mathbb{R}^N , and let Φ be an $M \times N$ random linear mapping. Then

$$(1 - \delta)\|x_1 - x_2\|_2^2 \leq \|\Phi(x_1 - x_2)\|_2^2 \leq (1 + \delta)\|x_1 - x_2\|_2^2$$

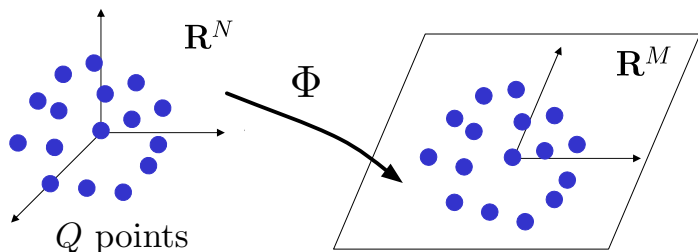
for *all* $x_1, x_2 \in \mathcal{Q}$ with

$$P \{\text{Failure}\} \leq 2Q^2 e^{-\delta^2 M/8} \leq \epsilon$$

when

$$M \geq \frac{8}{\delta^2} \left[2 \log(Q) + \log\left(\frac{1}{\epsilon}\right) + 0.7 \right]$$

The Johnson-Lindenstrauss Lemma



Φ embeds to precision δ with probability ϵ when

$$M \geq \frac{8}{\delta^2} \left[2 \log(Q) + \log \left(\frac{1}{\epsilon} \right) + 0.7 \right]$$

Concentration bound

We have: For any *fixed* $x \in \mathbb{R}^N$

$$(1 - \delta)\|x\|_2^2 \leq \|\Phi x\|_2^2 \leq (1 + \delta)\|x\|_2^2$$

with probability exceeding $1 - 2e^{-c\delta^2 M}$.

(Can take $c = 1/8$.)

We want: this for *all* $2S$ -sparse x simultaneously...

A single $2S$ -dimensional subspace

Theorem: Let V be a $2S$ -dimensional subspace of \mathbb{R}^N . Then

$$\mathbb{P} \left\{ \sup_{x \in V} \left| \|\Phi x\|_2^2 - \|x\|_2^2 \right| > \delta \right\} \leq 2 \cdot 9^{2S} \cdot e^{-c' \delta^2 M}$$

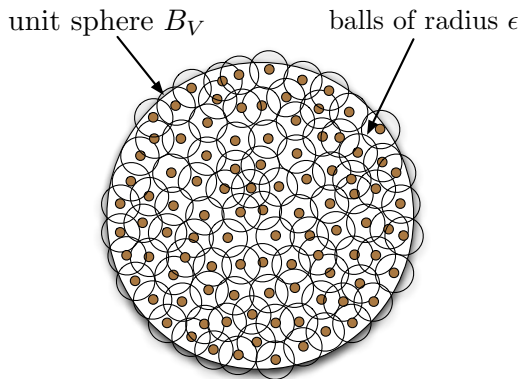
where the constant $c' = c/4$ with c from the previous theorem.

As before, it is enough to prove this for

$$x \in B_V = \{x \in V : \|x\|_2 = 1\}$$

Covering the sphere

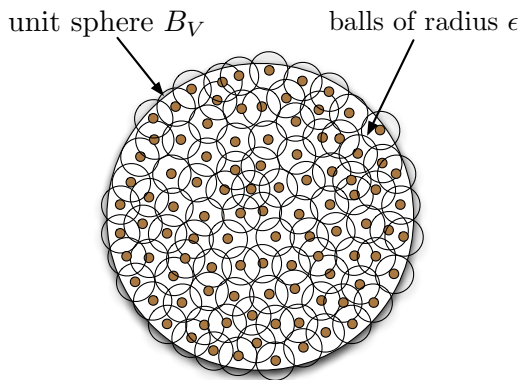
An ϵ -net for B_V :



for every $x \in B_V$, there is a $y \in \text{Net}$ such that $\|x - y\|_2 \leq \epsilon$

$N(B_V, \epsilon)$ is the *size of the smallest ϵ -net*

Covering the sphere



It is a fact that

$$N(B_V, \epsilon) \leq \left(1 + \frac{2}{\epsilon}\right)^{2S}$$

From discrete to continuous

Lemma: Fix $0 \leq \epsilon < 1/2$, and let \mathcal{N}_ϵ be the minimal ϵ -net for B_V . Then

$$\sup_{x \in B_V} \left| \|\Phi x\|_2^2 - \|x\|_2^2 \right| \leq \frac{1}{1 - 2\epsilon} \max_{y \in \mathcal{N}_\epsilon} \left| \|\Phi y\|_2^2 - \|y\|_2^2 \right|$$

A single $2S$ -dimensional subspace

Theorem: Let V be a $2S$ -dimensional subspace of \mathbb{R}^N . Then

$$\mathbb{P} \left\{ \sup_{x \in V} \left| \|\Phi x\|_2^2 - \|x\|_2^2 \right| > \delta \right\} \leq 2 \cdot 9^{2S} \cdot e^{-c' \delta^2 M}$$

where the constant $c' = c/4$ with c from the previous theorem.

A single $2S$ -dimensional subspace

Theorem: Let V be a $2S$ -dimensional subspace of \mathbb{R}^N . Then

$$\mathbb{P} \left\{ \sup_{x \in V} \left| \|\Phi x\|_2^2 - \|x\|_2^2 \right| > \delta \right\} \leq 2 \cdot 9^{2S} \cdot e^{-c' \delta^2 M}$$

where the constant $c' = c/4$ with c from the previous theorem.

So Φ is “well-conditioned” on V when

$$M \geq \text{Const} \cdot S$$

A single $2S$ -dimensional subspace

Theorem: Let V be a $2S$ -dimensional subspace of \mathbb{R}^N . Then

$$\mathbb{P} \left\{ \sup_{x \in V} \left| \|\Phi x\|_2^2 - \|x\|_2^2 \right| > \delta \right\} \leq 2 \cdot 9^{2S} \cdot e^{-c' \delta^2 M}$$

where the constant $c' = c/4$ with c from the previous theorem.

We want this for *all subspaces* in which $2S$ -sparse signals live...

A single $2S$ -dimensional subspace

Theorem: Let V be a $2S$ -dimensional subspace of \mathbb{R}^N . Then

$$\mathbb{P} \left\{ \sup_{x \in V} \left| \|\Phi x\|_2^2 - \|x\|_2^2 \right| > \delta \right\} \leq 2 \cdot 9^{2S} \cdot e^{-c' \delta^2 M}$$

where the constant $c' = c/4$ with c from the previous theorem.

We want this for *all subspaces* in which $2S$ -sparse signals live...

There are $\binom{N}{2S} \leq \left(\frac{Ne}{2S}\right)^{2S}$ such subspaces...

All $2S$ -dimensional subspaces

For $\Gamma \subset \{1, \dots, N\}$, let

$$B_\Gamma = \{x \in \mathbb{R}^N : x_\gamma = 0, \gamma \notin \Gamma, \|x\|_2 = 1\}.$$

Theorem:

$$\mathbb{P} \left\{ \max_{|\Gamma| \leq 2S} \sup_{x \in B_\Gamma} \left| \|\Phi x\|_2^2 - \|x\|_2^2 \right| > \delta \right\} \leq 2 \left(\frac{Ne}{2S} \right)^{2S} 9^{2S} e^{-c'\delta^2 M}$$

SUCCESS!!!

All $2S$ -dimensional subspaces

Theorem:

$$\begin{aligned} \mathbb{P} \left\{ \sup_{\text{all } 2S \text{ sparse } x} \left| \|\Phi x\|_2^2 - \|x\|_2^2 \right| > \delta \right\} &\leq 2 \left(\frac{Ne}{2S} \right)^{2S} 9^{2S} e^{-c'\delta^2 M} \\ &= e^{\log 2 + 2S \log(Ne/2S) + 2S \log 9 - c'\delta^2 M} \end{aligned}$$

Which is to say

$$(1 - \delta)\|x\|_2^2 \leq \|\Phi x\|_2^2 \leq (1 + \delta)\|x\|_2^2 \quad \forall 2S - \text{ sparse } x$$

with high probability when

$$M \geq \frac{\text{Const}}{\delta^2} \cdot S \log(N/S)$$

SUCCESS!!!