4. Restricted Isometry Property

We have seen that the NSP implies the guaranteed recovery of sparse solutions of underdetermined linear systems by basis pursuit. However it is somehow difficult to construct matrices satisfying this property. We shall therefore present a sufficient condition called *Restricted Isometry Property*, which was first introduced in [CT05], and which ensures that the NSP is satisfied. We will also see in this section that it implies stable recovery of nearly sparse signals and also robustness to measurement noise.

Definition 4.1. Let $A \in \mathbb{R}^{m\times n}$ and let $s \in [n]$. Then the **restricted isometry constant** $\delta_s = \delta_s(A)$ of order s is the smallest $\delta \geq 0$, such that

$$(1-\delta)\|x\|_2^2 \leq \|Ax\|_2^2 \leq (1+\delta)\|x\|_2^2, \qquad \forall x \in \Sigma_s. \tag{4.1}$$

Furthermore, we say that A satisfies the **restricted isometry property (RIP)** of order s with the constant δ_s if $\delta_s \in [0,1)$.

Remark 4.1. The condition (4.1) states that A acts nearly like an isometry when it is restricted to vectors from Σ_s. The smaller the constant $\delta_s(A)$ is, the closer the matrix A is to an isometry on Σ_s. We will be interested in an later section in constructing matrices with small RIP constants.

Note that if a matrix A satisfies the RIP of order $2s$, the we can interpret (4.1) as saying that A approximately preserves the distance between any pair of s-sparse vectors. To see this, choose $x = y - z$, where $y, z \in \Sigma_s$ and x is at most $2s$-sparse. This will have fundamental implications concerning robustness to noise.

We additional note that the inequality $\delta_1(A) \leq \delta_2(A) \leq \cdots \leq \delta_s(A)$ holds trivially.

Remark 4.2. The restricted isometriy constant δ_s is given as

$$\delta_s = \max_{S\subset[n],|S|\leq s} \|A_S^\top A_S - I\|_2. \tag{4.2}$$

By (4.2) each submatrix A_S, $S \subset [n]$ with $|S| \leq s$ has its singular values in the interval $[1-\delta_s, 1+\delta_s]$. The submatrix A_S will be injective when $\delta_s < 1$. Thus, the relevant situation occurs when $\delta_s < 1$.

It is important to note that while in our definition of the RIP we assume bounds that are symmetric about 1, this is merely a notational convenience. We could also consider arbitrary bounds with

$$\alpha\|x\|_2^2 \leq \|Ax\|_2^2 \leq \beta\|x\|_2^2, \qquad \forall x \in \Sigma_s \tag{4.3}$$

with $0 < \alpha \leq \beta < \infty$. Given such bounds, one can always scale A so that it satisfies the symmetric bound around 1 as in (4.1). Indeed, multiplying A by $\sqrt{2/(\beta+\alpha)}$ will result in an $\tilde{A}$ for which (4.1) hold with $\delta_s = (\beta-\alpha)/(\beta+\alpha)$.

In general, computing RIP constants is of combinatorial nature. For small size matrices these constants can be computed.

Example 4.1. Consider $s = 1$. Let $x \in \Sigma_1$ and denote by x_i the only nonzero entry of x. The RIP property of order 1 becomes

$$(1-\delta_1)x_i^2 \leq \|A^i\|_2^2 x_i^2 \leq (1+\delta_1)x_i^2, \quad i \in [n], \tag{4.4}$$

where A^i denotes the i-th column of A. For $x_i \neq 0$ we get

$$(1-\delta_1) \leq \|A^i\|_2^2 \leq (1+\delta_1), \quad i \in [n]. \tag{4.5}$$

Therefore, a matrix A satisfies the RIP of order 1 if the norm of each of its columns is approximately equal to 1.

Example 4.2. Let $s = 2$ and consider

$$A = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & -1 \\ -\frac{1}{\sqrt{2}} & 1 & 0 \end{pmatrix}. \tag{4.6}$$

We show that A satisfies the RIP of order 2 for $\delta_2 = 1/\sqrt{2}$. By the previous example A satisfies the RIP of order 1 for $\delta_1 = 0$ since the columns are normalized to 1. Next, let $x = (x_1, x_2, x_3)^\top$ be an arbitrary exact

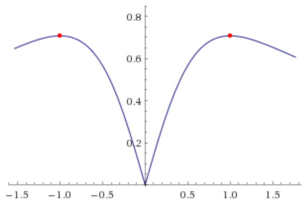

FIGURE 4.1. Plot of function $\alpha \mapsto \frac{\sqrt{2}|\alpha|}{\alpha^2+1}$ from 4.8 that attains its maximum in $\alpha = \pm 1$.

2-sparse vector. Consider first the case where $x_3 = 0$ and define $\alpha := x_2/x_1$ ($x_1 \neq 0$ since otherwise $x \in \Sigma_1$). The RIP property can be written as

$$(1-\delta_2)(1+\alpha^2)x_1^2 \leq (\alpha^2 - \sqrt{2}\alpha + 1)x_1^2 \leq (1+\delta_2)(1+\alpha^2)x_1^2. \tag{4.7}$$

Condition (4.7) is satisfied for every $\alpha \in \mathbb{R}$ if

$$\delta_2 \geq \max_{\alpha \in \mathbb{R}} \frac{\sqrt{2}|\alpha|}{\alpha^2+1} = \frac{1}{\sqrt{2}}, \tag{4.8}$$

see Fig. 4.1. We obtain the same equation when we assume that $x_2 = 0$. Finally, suppose that $x_1 = 0$. In this case, $Ax = (-x_3, x_2)^\top$ and $\|Ax\|_2^2 = \|x\|_2^2$ so that (4.1) holds for $\delta_2 = 0$. We conclude that (4.1) is satisfied for all $x \in \Sigma_2$ with $\delta_2 = 1/\sqrt{2}$.

If A satisfies the RIP of order s with constant δ_s, then for any $s' < s$ we automatically have that A satisfies the RIP of order s' with constant $\delta_{s'} \leq \delta_s$.

It is also straightforward to see that if A satisfies RIP of order $2s$ for any $\delta \in (0,1)$, then $\mathrm{spark}(A) > 2s$. This follows from the lower bound in (4.1). Indeed, let $x \neq 0$ be an arbitrary vector in Σ_{2s}. From the RIP property we have that $\|Ax\| > 0$ so that $x \notin \mathcal{N}(A)$. This in turn implies that every $2s$ columns of A are linearly independent and $\mathrm{spark}(A) > 2s$.

A useful property that is implied by RIP is given in the following proposition.

Proposition 4.1. *If A satisfies the RIP of order $2s$, then for any pair of vectors $x, z \in \Sigma_s$ with disjoint support, we have*

$$|\langle Ax, Az\rangle| \leq \delta_{2s}\|x\|_2\|z\|_2. \tag{4.9}$$

Proof. If $x, z \in \Sigma_s$ are two vectors with disjoint supports and $\|x\|_2 = \|z\|_2 = 1$, then $x \pm z \in \Sigma_{2s}$ and $\|x \pm z\|_2^2 = 2$. If we now combine the RIP of A

$$2\,(1-\delta_{2s}) \leq \|A(x \pm z)\|_2^2 \leq 2(1+\delta_{2s})$$

with the polarization identity[1] we get

$$|\langle Ax, Az\rangle| = \frac{1}{4}\left|\|Ax+Az\|_2^2 - \|Ax-Az\|_2^2\right| \leq \delta_{2s}. \tag{4.10}$$

For arbitrary x, z, define $x' = x/\|x\|$ and $z' = z/\|z\|$. Since x' and z' have equal one norm, we can apply (4.10) to conclude that

$$|\langle Ax, Az\rangle| = \|x\|_2\|z\|_2|\langle Ax', Az'\rangle| \leq \delta_{2s}\|x\|_2\|z\|_2. \tag{4.11}$$

□

The following theorem shows that RIP of sufficiently high order with a constant small enough is indeed a sufficient condition for NSP.

[1]If V is a real vector space, then the inner product is defined by the polarization identity

$$\langle x,\, y\rangle = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right), \quad \forall\, x, y \in V.$$

Theorem 4.2. *Suppose $A \in \mathbb{R}^{m\times n}$ satisfies the RIP of order $2s$ with $s \le n/2$. If $\delta_{2s}(A) < 1/3$, then A has the NSP of order s.*

Proof. Let $v \in \mathcal{N}(A)$ and let $S \subset [n]$ with $|S| \le s$. We will show, that

$$\|v_S\|_2 \le \frac{\delta_{2s}}{1-\delta_s} \cdot \frac{\|v\|_1}{\sqrt{s}}. \tag{4.12}$$

If $\delta_s \le \delta_{2s} < 1/3$, then with Cauchy-Schwarz inequality we obtain $\|v_S\|_1 \le \sqrt{s}\|v_S\|_2 < \|v\|_1/2$ and the NSP of A of order s follows.

To show (4.12), let us assume that $v \in \mathcal{N}(A)$ is fixed. It is enough to consider $S = S_0$ the set of the s largest entries of v taken in the absolute value. Furthermore, we denote by S_1 the set of s largest entries of $v_{S_0^c}$ in the absolute value, by S_2 the set of s largest entries of $v_{(S_0\cup S_1)^c}$ in the absolute value, etc. Using now that $0 = Av = A(v_{S_0} + v_{S_1} + v_{S_2} + \dots)$ and (4.9), we arrive at

$$\begin{aligned}\|v_{S_0}\|_2^2 &\le \frac{1}{1-\delta_s}\|Av_{S_0}\|_2^2 = \frac{1}{1-\delta_s}\langle Av_{S_0}, A(-v_{S_1}) + A(-v_{S_2}) + \dots\rangle \\ &\le \frac{1}{1-\delta_s}\sum_{j\ge 1}|\langle Av_{S_0}, Av_{S_j}\rangle| \le \frac{\delta_{2s}}{1-\delta_s}\sum_{j\ge1}\|v_{S_0}\|_2 \cdot \|v_{S_j}\|_2.\end{aligned}$$

We divide this inequality by $\|v_{S_0}\|_2 \ne 0$ and obtain

$$\|v_{S_0}\|_2 \le \frac{\delta_{2s}}{1-\delta_s}\sum_{j\ge1}\|v_{S_j}\|_2.$$

Taking in account the definition of the sets $S_j, j \in [n]_0$ we obtain

$$\sum_{j\ge1}\|v_{S_j}\|_2 = \sum_{j\ge1}\left(\sum_{k\in S_j}|v_k|^2\right)^{1/2} \le \sum_{j\ge1}\left(s\max_{k\in S_j}|v_k|^2\right)^{1/2} \tag{4.13}$$

$$= \sum_{j\ge1}\sqrt{s}\max_{k\in S_j}|v_k| \le \sum_{j\ge1}\sqrt{s}\min_{k\in S_{j-1}}|v_k| \le \sum_{j\ge1}\sqrt{s}\cdot\frac{\sum_{k\in S_{j-1}}|v_k|}{s} \tag{4.14}$$

$$= \sum_{j\ge1}\frac{\|v_{S_{j-1}}\|_1}{\sqrt{s}} = \frac{\|v\|_1}{\sqrt{s}} \tag{4.15}$$

and the proof is complete. □

Combining Theorems 3.2 and 4.2, we obtain immediately the following corollary.

Corollary 4.3. *Suppose $A \in \mathbb{R}^{m\times n}$ has the RIP of order $2s$ and let $s \in [n]$ with $s \le n/2$. If $\delta_{2s}(A) < 1/3$, then every s-sparse vector x is the unique solution of problem (P_1) from (3.1) with $b = Ax$.*

Remark 4.3. It will turn out that the simplest way of constructing matrices that fulfill the RIP condition with reasonably small constants with high probability (and hence also the NSP) is by taking its entries to be independent standard normal variables, e.g.

$$A = \frac{1}{\sqrt{m}}\begin{pmatrix} a_{1,1} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn}\end{pmatrix},$$

where $a_{ij}, i \in [m], j \in [n]$, are i.i.d. standard normal variables.

4.1. **Stability and Robustness.** We have seen that we can recover sparse solutions of underdetermined linear systems by convex optimization via ℓ_1-minimization. This is surely a very promising result. However, two additional features are necessary to extend this results to real-life applications, namely

Stability: We want to be able to recover (or at least approximate) also vectors $x \in \mathbb{R}^n$, which are not exactly sparse but compressible. Recall that we have characterized compressible vectors by assuming that their best s-term approximation decays rapidly with s. Intuitively, the faster the decay of the best s-term approximation of $x \in \mathbb{R}^n$ is, the better we should be able to approximate x.

Robustness: Equally important, we want to recover sparse or compressible vectors from noisy measurements. The basic model here is the assumptions that the measurement vector b is given by $b = Ax + r$, where r is small (in some sense). Again, the smaller the error r is, the better we should be able to recover an approximation of x.

We have seen in the previous section that the NSP can be extended also to the noisy scenario and we have introduced the stable and robust NSP. As previously mentioned it is difficult to design matrices that satisfy the NSP. The RIP is strictly stronger than the NSP in the sense that if a matrix satisfies the RIP, then it will also satisfy the (stable) NSP as seen in Thm. 4.2 but also the robust NSP.

Moreover, we will see next that the lower bound in RIP is a necessary condition to be able to recover all sparse signal from noisy measurements. We will need the following notion.

Definition 4.2. Let $A \in \mathbb{R}^{m\times n}$ denote a sensing matrix and $\Delta : \mathbb{R}^m \to \mathbb{R}^n$ a decoder, i.e. recovery algorithm. We say that the pair (A, Δ) is C-*robust* if for any $x \in \Sigma_s$ and any $r \in \mathbb{R}^m$ we have that

$$\|\Delta(Ax + r) - x\|_2 \le C\|r\|_2. \tag{4.16}$$

This definition implies that if we add a small amount of noise to the measurements, then the impact on the recovered signal is not arbitrary large. The next result shows that the existence of a robust decoding algorithm requires that A satisfies the lower bound of RIP.

Theorem 4.4. *If the pair* (A, Δ) *is* C-robust*, then*

$$\frac{1}{C}\|x\|_2 \le \|Ax\|_2 \tag{4.17}$$

for all $x \in \Sigma_{2s}$.

Proof. For any $x \in \Sigma_{2s}$ write $x = v - z$ with $v, z \in \Sigma_s$. Define

$$r_v := \frac{A(z-v)}{2} \quad \text{and} \quad r_z := \frac{A(v-z)}{2}, \tag{4.18}$$

and note that

$$Av + r_v = Az + r_z = \frac{A(v+z)}{2}. \tag{4.19}$$

Let $\hat{x} = \Delta(Av + r_v) = \Delta(Az + r_z)$. From the triangle inequatlity and the definition of C-robustness we have that

$$\begin{aligned} \|v - z\|_2 &= \|v - \hat{x} + \hat{x} - z\|_2 \\ &\le \|v - \hat{x}\|_2 + \|\hat{x} - z\|_2 \\ &\le C\|r_v\|_2 + C\|r_z\|_2 \\ &= C\|Av - Az\|_2. \end{aligned}$$

The last equality follows from $Av - Az = r_z - r_v$ and the fact that $r_z = -r_v$. Noting that $x = v - z$ completes the proof. □

We will further concentrate on the recovery properties of the following relaxation of (P_1) and consider the basis pursuit denoising problem $(P_{1,\eta})$ from (3.15)

$$\min_{x\in\mathbb{R}^n} \|x\|_1 \qquad \text{s.t.} \quad \|Ax - b\|_2 \le \eta, \tag{4.20}$$

with $\eta \ge 0$. If $\eta = 0$, $(P_{1,\eta})$ reduces back to (P_1) .

Theorem 4.5. *Suppose $A \in \mathbb{R}^{m\times n}$ satisfies the RIP of order $2s$. Let $\delta_{2s} < \sqrt{2}-1$, $b = Ax^* + r$ and $\|r\|_2 \le \eta$. Then the solution $\hat{x}$ of $(P_{1,\eta})$ satisfies*

$$\|\hat{x} - x^*\|_2 \le \frac{C\sigma_s(x^*)_1}{\sqrt{s}} + D\eta, \tag{4.21}$$

where $C, D > 0$ are two positive constants.

Proof. First, let us recall that by Prop. 4.1, if A has RIP of order $2s$ and $u, v \in \Sigma_s$ are two vectors with disjoint supports, then we have by (4.9)

$$|\langle Au, Av\rangle| \le \delta_{2s}\|u\|_2\|v\|_2. \tag{4.22}$$

Set $h = \hat{x} - x^*$ and define the index set $S_0 \subset [n]$ as the locations of the s largest entries of x^* taken in absolute value. Furthermore, we define $S_1 \subset S_0^c$ to be the indices of the s largest absolute entries of $h_{S_0^c}$, S_2 the indices of the s largest absolute entries of $h_{(S_0\cup S_1)^c}$, etc. Since $\hat{x}$ is a feasible point of $(P_{1,\eta})$, we obtain by using the triangle inequality

$$\|Ah\|_2 = \|A(\hat{x} - x^*)\|_2 \le \|A\hat{x} - b\|_2 + \|b - Ax^*\|_2 \le 2\eta. \tag{4.23}$$

Since $\hat{x}$ is the minimizer of $(P_{1,\eta})$, we get $\|\hat{x}\|_1 = \|x^* + h\|_1 \le \|x^*\|_1$. This we use to show that h must be small outside of S_0. Indeed, we obtain

$$\begin{aligned}
\|h_{S_0^c}\|_1 &= \|(x^* + h)_{S_0^c} - x^*_{S_0^c}\|_1 + \|(x^* + h)_{S_0} - h_{S_0}\|_1 - \|x^*_{S_0}\|_1 \\
&\le \|(x^* + h)_{S_0^c}\|_1 + \|x^*_{S_0^c}\|_1 + \|(x^* + h)_{S_0}\|_1 + \|h_{S_0}\|_1 - \|x^*_{S_0}\|_1 \\
&= \|x^* + h\|_1 + \|x^*_{S_0^c}\|_1 + \|h_{S_0}\|_1 - \|x^*_{S_0}\|_1 \\
&\le \|x^*\|_1 + \|x^*_{S_0^c}\|_1 + \|h_{S_0}\|_1 - \|x^*_{S_0}\|_1 \\
&= \|h_{S_0}\|_1 + 2\|x^*_{S_0^c}\| \le s^{1/2}\|h_{S_0}\|_2 + 2\sigma_s(x^*)_1.
\end{aligned}$$

Using this together with the approach applied already in the proof of Thm. 4.2, see (4.13)–(4.15), we derive

$$\sum_{j\ge 2}\|h_{S_j}\|_2 \le s^{-1/2}\|h_{S_0^c}\|_1 \le \|h_{S_0}\|_2 + 2s^{-1/2}\sigma_s(x^*)_1. \tag{4.24}$$

We use the RIP property of A, (4.22), (4.23), (4.24) and the inequality $\|h_{S_0}\|_2 + \|h_{S_1}\|_2 \le \sqrt{2}\|h_{S_0\cup S_1}\|_2$ and get

$$\begin{aligned}
(1-\delta_{2s})\|h_{S_0\cup S_1}\|_2^2 \le \|Ah_{S_0\cup S_1}\|_2^2 &= \langle Ah_{S_0\cup S_1}, Ah\rangle - \langle Ah_{S_0\cup S_1}, \sum_{j\ge 2} Ah_{S_j}\rangle \\
&\le \|Ah_{S_0\cup S_1}\|_2\|Ah\|_2 + \sum_{j\ge 2}\left|\langle Ah_{S_0}, Ah_{S_j}\rangle\right| + \sum_{j\ge 2}\left|\langle Ah_{S_1}, Ah_{S_j}\rangle\right| \\
&\le 2\eta\sqrt{1+\delta_{2s}}\|h_{S_0\cup S_1}\|_2 + \delta_{2s}(\|h_{S_0}\|_2 + \|h_{S_1}\|_2)\sum_{j\ge 2}\|h_{S_j}\|_2 \\
&\le \|h_{S_0\cup S_1}\|_2(2\eta\sqrt{1+\delta_{2s}} + \sqrt{2}\delta_{2s}\|h_{S_0}\|_2 + 2\sqrt{2}\delta_{2s}s^{-1/2}\sigma_s(x^*)_1)\,.
\end{aligned}$$

We divide the inequality above by $(1-\delta_{2s})\|h_{S_0\cup S_1}\|_2$, replace $\|h_{S_0}\|_2$ with the larger quantity $\|h_{S_0\cup S_1}\|_2$ and subtract $\sqrt{2}\delta_{2s}/(1-\delta_{2s})\|h_{S_0\cup S_1}\|_2$ from both sides to arrive at

$$\|h_{S_0\cup S_1}\|_2 \le (1-\rho)^{-1}(\alpha\eta + 2\rho s^{-1/2}\sigma_s(x^*)_1), \tag{4.25}$$

where

$$\alpha = \frac{2\sqrt{1+\delta_{2s}}}{1-\delta_{2s}} \quad \text{and} \quad \rho = \frac{\sqrt{2}\delta_{2s}}{1-\delta_{2s}}.$$

By using (4.24) and (4.25) we obtain

$$\|h\|_2 \le \|h_{(S_0\cup S_1)^c}\|_2 + \|h_{S_0\cup S_1}\|_2 \le \sum_{j\ge 2}\|h_{S_j}\|_2 + \|h_{S_0\cup S_1}\|_2$$

$$\le 2\|h_{S_0\cup S_1}\|_2 + 2s^{-1/2}\sigma_s(x)_1 \le C\frac{\sigma_s(x)_1}{\sqrt{s}} + D\eta,$$

with $C = 2(1-\rho)^{-1}\alpha$ and $D = 2(1+\rho)(1-\rho)^{-1}$. This shows (4.22). □

Remark 4.4. Some comments are in order. First, note that the reconstruction error is bounded by two terms. The first term is determined by the bound on the measurement noise r. This tells us that if we add a small amount of noise to the measurements, its impact on the recovered signal remains well-controlled. Moreover, as the noise bound approaches zero, we see that the impact of the noise on the reconstruction error will also approach zero. The second term measures the error that occurs by approximating the signal x^* as a s-sparse signal (where the error is measured using the ℓ_1-norm). In the case that x^* is compressible, then the error again remains well-controlled. Note that this term vanishes if x^* is perfectly s-sparse. Moreover, whenever $x^* \in \Sigma_s$ and there is no noise, then we obtain *exact recovery*.

There have been many efforts to improve on the constants in Thm. 4.5 and to weaken the assumption on the constant δ_{2s}, but most of this work results in theorems that are substantially the same.

4.2. **RIP and Measurement Bounds.** It is important to obtain insight into how many measurements are necessary to achieve the RIP. If we ignore the impact of δ_s and only focus on the dimensions of the problem (m, n and s) then we can provide a lower bound. Before proving this lower bound, we need a preliminary lemma.

Lemma 4.6. *Let s and n satisfying $s < n/2$. There exists a set $X \subset \Sigma_s$ such that for any $x \in X$ we have $\|x\|_2 \le \sqrt{s}$ and for any $x, z \in X$ with $x \ne z$*

$$\|x - z\|_2 \ge \sqrt{s/2} \tag{4.26}$$

and

$$\log|X| \ge \frac{s}{2}\log\left(\frac{n}{s}\right) \tag{4.27}$$

holds.

Proof. We consider the set

$$U := \{x \in \{0, \pm 1\}^n \colon \|x\|_0 = s\}. \tag{4.28}$$

By construction, $\|x\|_2^2 = s$ if $x \in U$. Thus we can construct the set X by choosing elements from U. We will then automatically have $\|x\|_2 = \sqrt{s}$ for all elements in U.

Next, we note that $|U| = \binom{n}{s}2^s$. We observe also that for $z, x \in U$ we have $\|x - z\|_0 \le \|x - z\|_2^2$. Thus, if $\|x - z\|_2^2 \le s/2$ we also have $\|x - z\|_0 \le s/2$. This implies that for any fixed $x \in U$

$$\left|\{z \in U \colon \|x - z\|_2^2 \le s/2\}\right| \le \left|\{z \in U \colon \|x - z\|_0 \le s/2\}\right| \le \binom{n}{s/2}3^{s/2}. \tag{4.29}$$

We will construct the set X by iteratively choosing points that satisfy (4.26). After j points are added to the set X, there are at least

$$\binom{n}{s}2^s - j\binom{n}{s/2}3^{s/2} \tag{4.30}$$

points left to choose from. The process stops if this quantity is not positive any more. Hence, we can construct a set of size $|X|$ if

$$|X|\binom{n}{s/2}3^{s/2} \le \binom{n}{s}2^s. \tag{4.31}$$

Further we observe that

$$\frac{\binom{n}{s}}{\binom{n}{s/2}} = \frac{(s/2)!(n-s/2)!}{s!(n-s)!} = \prod_{i=1}^{s/2}\frac{n-s+i}{s/2+i} \ge \left(\frac{n}{s} - \frac{1}{2}\right)^{s/2}, \tag{4.32}$$

where the last inequality follows since $i \mapsto \frac{n-s+i}{s/2+i}$ is decreasing. If now we set $|X| = (n/s)^{s/2}$ then we have

$$|X| \left(\frac{3}{4}\right)^{s/2} = \left(\frac{3n}{4s}\right)^{s/2} \leq \left(\frac{n}{s} - \frac{1}{2}\right)^{s/2} \leq \frac{\binom{n}{s}}{\binom{n}{s/2}}. \tag{4.33}$$

Now (4.33) shows that (4.31) holds for $|X| = (n/s)^{s/2}$. Hence we have constructed a set X with the desired properties. □

Theorem 4.7. *Let A be an $m \times n$ matrix that satisfies the RIP of order $2s$ with constant $\delta_{2s} \in (0, 1/2]$. Then*

$$m \geq Cs \log\left(\frac{n}{s}\right) \tag{4.34}$$

where $C = 1/(2\log(\sqrt{24}+1)) \approx 0.28$.

Proof. Consider the set X from Lem. 4.6. Let A satisfy RIP of order $2s$. Then for any points $x, z \in X$ we have by (4.26)

$$\|Ax - Az\|_2 \geq \sqrt{1-\delta_{2s}}\|x-z\|_2 \geq \sqrt{s/4}, \tag{4.35}$$

since $x - z \in \Sigma_{2s}$ and $\delta_{2s} \leq 1/2$. For all $x \in X$ we also have

$$\|Ax\|_2 \leq \sqrt{1+\delta_{2s}}\|x\|_2 \leq \sqrt{3s/2}, \tag{4.36}$$

in view of $\|x\|_2 \leq \sqrt{s}$. From the lower bound in (4.35) it follows that

$$\left|\mathbb{B}^m_{\sqrt{\frac{s}{16}}}(Ax) \cap \mathbb{B}^m_{\sqrt{\frac{s}{16}}}(Az)\right| \leq 1, \tag{4.37}$$

since $\sqrt{s/4}/2 = \sqrt{s/16}$. Here we denote by $\mathbb{B}^m_r(y) := \{u \in \mathbb{R}^m \colon \|u-y\|_2 \leq r\}$ the ball in $\mathbb{R}^m$ centred at y with radius r w.r.t. the ℓ_2-norm.

On the other hand the entire set of balls is contained within a larger ball of radius $\sqrt{3s/2} + \sqrt{s/16}$ centred at 0 in view of the upper bound (4.36), i.e.

$$\bigcup_{x\in X} \mathbb{B}^m_{\sqrt{\frac{s}{16}}}(Ax) \quad \subset \quad \mathbb{B}^m_{\sqrt{\frac{s}{16}}+\sqrt{\frac{3s}{2}}}(0). \tag{4.38}$$

Now (4.38) implies that

$$\begin{aligned}
&\operatorname{vol}\left(\mathbb{B}^m_{\sqrt{\frac{s}{16}}+\sqrt{\frac{3s}{2}}}(0)\right) \geq |X| \cdot \operatorname{vol}\left(\mathbb{B}^m_{\sqrt{\frac{s}{16}}}(0)\right) \\
\Leftrightarrow\ &\left(\sqrt{\frac{s}{16}} + \sqrt{\frac{3s}{2}}\right)^m \geq |X| \cdot \left(\sqrt{\frac{s}{16}}\right)^m \\
\Leftrightarrow\ &\qquad \left(\sqrt{24}+1\right)^m \geq |X| \\
\Leftrightarrow\ &\qquad\qquad m \geq \frac{\log|X|}{\log\left(\sqrt{24}+1\right)}.
\end{aligned}$$

We can complete the proof by applying the bound in (4.27) for the cardinality of X from Lem. 4.6. □