PHASE TRANSITION IN THE ONE-BIT JOHNSON-LINDENSTRAUSS LEMMA

AMADOU BAH, BRYSON KAGY, AND EMILY SMITH

Abstract. The Johnson-Lindenstrauss Lemma (J-L Lemma) is a cornerstone of dimension reduction techniques. We study it in the one-bit context, namely we consider the unit sphere $S^{N-1}$, with normalized geodesic metric, and map a finite set $X \subset S^{N-1}$ into the Hamming cube $H_m = \{0,1\}^m$, with normalized Hamming metric. We find that for $0 < \delta < 1$, and $m > \frac{\ln n}{2 \delta^2}$ there is a $\delta$-RIP from $X$ into $H_m$. This is surprising as the value of $m$ is virtually identical to best known bound linear J-L Lemma. In both the linear and one-bit case, the maps are randomly constructed. We show that the probability of $B_m$ being a $\delta$-RIP satisfies a phase transition. It passes from probability of nearly zero to nearly one with a very small change in $m$. Our proof relies on delicate properties of Bernoulli random variables.

1. Introduction

Compressive sensing was first introduced as a practical application of signal processing and has since taken off and proven to be very useful for many aspects of modern life such as MRI scanning, cell phone imaging, electron microscopy, and many more [3,5,8]. It has been previously shown by Johnson-Lindenstrauss [7], that given a very high dimensional data set in $\mathbb{R}^N$, it is possible, with little sacrifice, to map vectors from a subset of this $N$-dimensional space, to a much lower, $m$-dimensional space. Recently, Alon and Klartag [1] studied the minimum number of bits required in order to maintain the Euclidean distance between data points. This differs from our results through the fact that we maintain the geodesic distance between points. The non-linear geodesic metric is basic to our considerations.

Dasgupta-Gupta [4] provide the best quantitative bounds in the J-L Lemma. For any $0 < \delta < 1$, any integer $n$ let

\[ m \geq 4 \frac{\ln n}{\frac{2}{\delta^2} - \frac{3}{\delta^3}}, \]

then for any set of $n$ points in $\mathbb{R}^N$ there exist a map $f : \mathbb{R}^N \rightarrow \mathbb{R}^m$ such that for all $x, y \in \mathbb{R}^N$ we have:

\[ (1 - \delta) \|x - y\|^2 \leq \|f(x) - f(y)\|^2 \leq (1 + \delta) \|x - y\|^2 \]

Research conducted at during an REU sponsored by an NSF MCTP-Grant to the Georgia Institute of Technology.
For comparison below, we remark that $\sqrt{1 + \delta} \simeq 1 + \delta/2$.

We study the one bit context. Consider the unit sphere $S^{N-1}$ with the normalized geodesic metric. We map finite $X \subset S^{N-1}$ into the $m$ dimensional Hamming cube $\mathbb{H}_m$, with normalized Hamming metric. A main result is that for $0 < \delta < 1$, and integer $n$, let $m > 2 \ln n / \delta^2$. For any set $X \subset S^{N-1}$ of cardinality $n$, there is a $\delta$-RIP from $X$ into $\mathbb{H}_m$. The counter-intuitive fact is that our bound for $m$ is virtually identical to the one that holds for the linear J-L Lemma.

We prove the One Bit J-L Lemma in §5. The simplier property of our random one-bit map being one to one is studied in §3. For special choices of $X$, we make a finer analysis of the one-to-one and RIP properties. They satisfy phase transitions that depend only weakly on the number of points we are mapping, see §6 and §4. Some background information is recalled in §2.

2. Background

We formalize below several definitions we will use throughout the paper.

**Hamming Cube.** $\mathbb{H}_m = \{0,1\}^m$ for all $x \in \mathbb{H}_m$ $x = x_1 x_2 ... x_n$ where $x_i \in \{0,1\}$. For all $x, y \in \mathbb{H}_m$ we have the normalized metric

$$d_{\mathbb{H}_m}(x,y) = \frac{1}{m} \#\{i : x_i \neq y_i\}.$$  

**Random $m$-dimensional One-Bit Map.** Given $\{\theta_j : 1 \leq j \leq m\}$ be iid uniformly distributed random vectors in $S^{N-1}$. Define a map $B_m : S^{N-1} \rightarrow \mathbb{H}_m$ by $B_m x = \{\text{sgn}(x \cdot \theta_j)\}_{j=1}^m$. Observe that

$$(2.0.1) \ d_{\mathbb{H}_m}(B_m x, B_m y) = \frac{1}{m} \sum_{i=1}^m \frac{|\text{sgn}(x \cdot \theta_i) - \text{sgn}(y \cdot \theta_i)|}{2}.$$  

**Geodesic Distance.** Fix $x$, $y$ on $S^{N-1}$. The geodesic distance $d_{\text{geo}}(x,y)$, is the shortest distance between the points $x$ and $y$ on the surface. This is given by

$$d_{\text{geo}}(x,y) = \frac{\cos^{-1}(x \cdot y)}{\pi}.$$  

Antipodal points are normalized to be distance one apart. Geodesic distance has this probabilistic interpretation: Let $\text{Wedge}_{xy} = \{\theta \in S^{N-1} : \text{sgn}(x \cdot \theta) \neq \text{sgn}(y \cdot \theta)\}$. These are the $\theta$ which distinguish between $x$ and $y$ under the one-bit map. Selecting $\theta \in S^{N-1}$ at random, the probability of being in $\text{Wedge}_{xy}$ is $d_{\text{geo}}(x,y)$. This is an instance of the Crofton formula.

For the distance in (2.0.1), we then have

$$d_{\mathbb{H}_m}(B_m x, B_m y) = \frac{1}{m} \sum_{i=1}^m 1_{\text{Wedge}_{xy}}(\theta_i).$$
The right hand side is an average of Bernoulli rvs. In particular, the difference between
the Hamming and geodesic metrics is
\[ d_{H_m}(B_m x, B_m y) - d_{geo}(x, y) = \frac{1}{m} \sum_{i=1}^{m} \text{1}_{\text{Wedge}_{xy}(\theta_j)} - d_{geo}(x, y). \]

Standard deviation inequalities for Bernoulli rvs apply to the right hand side above.

The Restricted Isometry Property. \( B_m : X \to \mathbb{H}_m \) has the \( \delta \)-RIP if for all pairs \( x, y \in X \):
\[ |d_{H_m}(B_m x, B_m y) - d_{geo}(x, y)| \leq \delta. \]

Positively Associated Stein-Chen Approximation. For random variables to be
positively associated, their covariance is positive, meaning they increase or decrease
together.
\[ d_{TV}(W, \text{Poi}(\lambda)) \leq \min \left( 1, \frac{1}{\lambda} \right) \left( \text{Var}(W) - \lambda + 2 \sum_{i \in I} P_i^2 \right) \]
where \( W \) is a sum of positively associated Bernoullis with parameter \( P_i \), \( \lambda \) is \( \mathbb{E}(W) \),
and \( d_{TV} \) is total variation distance.

General Form of Stein-Chen Approximation. \[ d_{TV}(W, \text{Poi}(\lambda)) \leq \min \left( 1, \frac{1}{\lambda} \right) \sum_{i \in I} \left( P_i^2 + \sum_{j \in \mathcal{N}_i} (P_i P_j + \mathbb{E}(X_i X_j)) \right) \]
where \( X_i \) are Bernoullis with parameter \( P_i \), \( W \) is a sum of all \( X_i \), \( \lambda = \mathbb{E}(W) \), \( \mathcal{N}_i \) is
the set of random variables that depend on \( X_i \), and \( d_{TV} \) is total variation distance.

3. A One-to-One Mapping From the Unit Sphere to the Hamming Cube

We start with an analysis of a simpler property of \( B_m \) being one-to-one.

**Theorem 3.1.** Let \( 0 < \delta, \epsilon < 1 \), and let \( X \subset \mathbb{S}^{N-1} \) be a subset of \( n \) points with
\( d_{geo}(x, y) > 1 - \delta \), where \( x \neq y \in X \). The random \( m \)-dimensional one-bit map \( B_m : X \to \mathbb{H}_m \)
will be one-to-one with probability at least \( 1 - \epsilon \) provided that
\[ m \geq \frac{\ln \frac{n^2}{2\epsilon}}{\ln \frac{1}{\delta}}. \]

In the special case when the points \( x \) and \( y \) are pairwise orthogonal, \( d_{geo}(x, y) = \frac{1}{2} \),
\[ m \geq 2 \log_2 n + \log_2 \frac{1}{2\epsilon}. \]
By the pigeonhole principle, \( m \) must be at least \( \log_2 n \). Our result shows that if \( m > 2\log_2 n \), then the random \( m \)-dimensional one-bit map is one-to-one with high probability.

**Proof.** By the union bound, we know that:

\[
\mathbb{P}(B_m \text{ is not one-to-one}) \leq \sum_{x \sim y} \mathbb{P}(B_m x = B_m y \text{ with } x \neq y) \leq \binom{n}{2} \delta^m.
\]

In this expression, \( \sum_{x \sim y} \) means the sum over all unordered pairs \( x \sim y \) where \( x \neq y \).

Above, there are \( \binom{n}{2} \) pairs \( x \sim y \in X \). The \( i \)th coordinates of \( B_m x \) and \( B_m y \) are equal with probability at most \( \delta \). The coordinates are independent, hence the inequality above. We require \( \binom{n}{2} \delta^m \leq \epsilon \), which is true if

\[
m \geq \frac{\ln \frac{n^2}{2\epsilon}}{\ln \frac{1}{\delta}}.
\]

This condition is sufficient for \( B_m \) to be one-to-one with probability \( 1 - \epsilon \). In the special case when \( X \) consists of pairwise orthogonal vectors, \( d_{\text{geo}}(x, y) = \frac{1}{2} \), the bound is

\[
m \geq 2\log_2 n + \log_2 \frac{1}{2\epsilon}.
\]

\[\square\]

### 4. A Phase Transition in One-to-One Property

For a special class of \( X \), we analyze the property of \( B_m \) being one-to-one. We show that the probability passes through a phase transition. And the width of the phase transition is essentially independent of the cardinality of \( X \).

**Theorem 4.1.** Fix \( 0 < \epsilon_2 < \epsilon_1 < 1 \). Let \( X \) be \( n \) pairwise orthogonal vectors in \( \mathbb{S}^{N-1} \), and let \( P_{1-1}(m) \) be the probability that \( B_m \) is one-to-one. Then for \( n \geq 10 \),

\[
1 - \epsilon_1 < P_{1-1}(m) \text{ when:}
\]

\[
\log_2 \frac{n(n-1)}{2\ln \frac{1}{1-\epsilon_1}} \leq m
\]

and \( P_{1-1}(m) < 1 - \epsilon_2 \) when

\[
m \leq \log_2 \frac{n(n-1)}{2\ln \frac{1}{1-\epsilon_2}}.
\]

Additionally, the phase transition is bounded as follows:

\[
P_{1-1}(m) = \mathbb{P}(B_m \text{ is one-to-one}) \in \left[ e^{-\frac{(2)^{\frac{n-1}{2}}}{2m}} - \left(\frac{n}{2}\right) \cdot 2^{-m}, e^{-\frac{(2)^{\frac{n-1}{2}}}{2m}} + \left(\frac{n}{2}\right) \cdot 2^{-2m} \right].
\]
We will analyze this from the perspective of the birthday problem. To do this, we will count all pairs of points that $B_m$ map to the same point in the Hamming cube. Namely for $x, y \in X$, let

$$W_{x \sim y} = \begin{cases} 1, & B_m x = B_m y \\ 0, & \text{otherwise.} \end{cases}$$

All $W_{x \sim y}$ are i.i.d with probability $p = \frac{1}{2^m}$ and $W = \sum_{x \sim y} W_{x \sim y}$ is a sum of positively associated Bernoulli random variables. By the Stein-Chen approximation, $W$ is close to a Poisson distribution, in total variation, denoted $d_{TV}$ below. We make this precise below:

$$d_{TV}(W, \text{Poi}(\lambda)) \leq \eta$$

where

$$\eta = \min \left( 1, \frac{1}{\lambda} \right) [\text{Var}(W) - \lambda + 2 \sum_{x \sim y} p^2].$$

**Lemma 4.2.** We claim

$$\text{Var}(W) = \lambda - \sum_{x \sim y} p^2.$$

**Proof.** If the $W_{x \sim y}$ have the property $\mathbb{E}[W_{x \sim y}W_{r \sim s}] = \mathbb{E}[W_{x \sim y}]\mathbb{E}[W_{r \sim s}]$ then the $W_{x \sim y}$ are pairwise independent. Assuming this, variance adds, and $\text{Var}(W)$ can be calculated:

$$\text{Var}(W) = \sum_{x \sim y} \text{Var}(W_{x \sim y})$$

$$= \binom{n}{2} p(1-p)$$

$$= \lambda - \sum_{x \sim y} p^2.$$

It remains to prove that $\mathbb{E}[W_{x \sim y}W_{r \sim s}] = \mathbb{E}[W_{x \sim y}]\mathbb{E}[W_{r \sim s}]$. It will be sufficient to show that $\mathbb{E}[W_{x \sim y}W_{r \sim s}] = p^2$. The only non-trivial case is when $x \sim y$ and $r \sim s$ share exactly one point. We will write this as $\mathbb{E}[W_{x \sim y}W_{y \sim s}]$ where $y$ is the shared point. $\mathbb{E}[W_{x \sim y}W_{y \sim s}]$ is equal to $\mathbb{P}(B_m x = B_m y = B_m s)$ which means we have three distinct points on the sphere mapping to the same point on the Hamming cube. Thus $\mathbb{P}(B_m x = B_m y = B_m s) = p^2$ giving us that $\mathbb{E}[W_{x \sim y}W_{r \sim s}] = p^2$. \qed

**Lemma 4.3.** We claim that $\eta = \binom{n}{2} \cdot 2^{-2m}$ where $\eta$ is (4.1.1)
Using this, we can bound $1 - \mathbb{P}(W \geq 1)$ in the window

$$1 - \mathbb{P}(W \geq 1) \in \left[ e^{-\frac{n^2}{2m}} - \eta, e^{-\frac{n^2}{2m}} + \eta \right].$$

**Proof.** We can find an expression for $\eta$ using the variance of $W$:

$$\eta = \min \left( 1, \frac{1}{\lambda} \right) \left[ \text{Var}(W) - \lambda + 2 \sum_{x \sim y} p^2 \right]$$

$$= \left( \frac{n}{2} \right) \cdot 2^{-2m}.$$  

We can bound $1 - \mathbb{P}(W \geq 1)$ which is equal to $P_{1-1}(m)$:

$$|\mathbb{P}(W \geq 1) - \mathbb{P}(\text{Poi}((\lambda)) \geq 1)| \leq \eta$$

$$1 - \mathbb{P}(W \geq 1) \in \left[ e^{-\frac{n^2}{2m}} - \eta, e^{-\frac{n^2}{2m}} + \eta \right].$$

□

**Solving For m.** Fix $0 < \epsilon_2 < \epsilon_1 < 1$, let $X$ be $n$ pairwise orthogonal vectors in $S^{N-1}$, and let $P_{1-1}(m)$ be the probability that $B_m$ is one-to-one, then $1 - \epsilon_1 P_{1-1}(m)$ when:

$$1 - e^{-\frac{n^2}{2m}} + \eta \leq \epsilon_1.$$  

and $P_{1-1}(m) < 1 - \epsilon_2$ when:

$$1 - e^{-\frac{n^2}{2m}} - \eta \geq \epsilon_2.$$  

In order to ensure that $\eta$ is very small compared to the Poisson distribution, we want $\eta \leq 0.01(1-e^{-\frac{n^2}{2m}}).$ If we fix $n$ and choose $m$ such that $\frac{n^2}{2m} \leq 1$, observe the inequality

$$\frac{n^2}{2} \leq \frac{n^2}{2m} - \frac{n^2}{2m}.$$

It is sufficient to bound $\eta$ as

$$\eta \leq 0.01 \frac{n^2}{2}.$$  

Manipulating this statement we get:

$$\frac{\binom{n}{2} \cdot 2^{-m}}{\binom{n}{2}} \leq 0.005.$$
Since we already assumed that \( \frac{n^2}{m} \leq 1 \), we can rewrite this inequality as \( \frac{1}{(\frac{n^2}{m})} \leq 0.005 \) and solve for \( n \): \( n \geq 10. \) This means that if \( n \geq 10, \) we have \( \eta \leq 0.01(1 - e^{-\frac{n^2}{2m}}) \) which allows us to rewrite our inequalities and gain bounds on \( m \):

\[
\log_2 \frac{n(n-1)}{2 \ln \frac{1 - \eta}{1 - \eta}} \leq m \leq \log_2 \frac{n(n-1)}{2 \ln \frac{1 - \eta}{1 - 0.01}}.
\]

5. A Union Bound for the Restricted Isometry Property

This is the One Bit version of the Johnson-Lindenstrauss Lemma. In particular the quantitative bound on \( m \) below is nearly identical to the best known bound in the linear Johnson-Lindenstrauss Lemma.

**Theorem 5.1.** Fix \( 0 < \delta < \frac{1}{2} \) and \( 0 < \epsilon < 1 \), let \( X \) be \( n \) pairwise orthogonal vectors in \( S^{N-1} \). The random \( m \) dimensional one-bit map, \( B_m : X \rightarrow \mathbb{H}_m \), satisfies the \( \delta \)-RIP with probability at least \( 1 - \epsilon \) when

\[
m \geq \frac{\ln \frac{n^2}{\epsilon}}{2\delta^2}.
\]

In the Restricted Isometry Property (RIP), we want to preserve the pairwise distances between the points so that for all pairs \( x, y \in X \)

\[
|d_{\mathbb{H}_m}(B_m x, B_m y) - d_{geo}(x, y)| \leq \delta.
\]

To ensure that we satisfy the \( \delta \)-RIP with probability at least \( 1 - \epsilon \), we have to be certain that the probability of failure, \( \epsilon \), is small:

\[
P(\exists x, y \in X : |d_{\mathbb{H}_m}(B_m x, B_m y) - d_{geo}(x, y)| \geq \delta) \leq \epsilon.
\]

Using the Union bound,

\[
P(\exists x, y \in X : |d_{\mathbb{H}_m}(B_m x, B_m y) - d_{geo}(x, y)| \geq \delta) \leq \sum_{x \sim y} P(\{|d_{\mathbb{H}_m}(B_m x, B_m y) - d_{geo}(x, y)| \geq \delta\}).
\]

We will first analyze the probability for one pair \( x \sim y \), namely:

\[
P(|d_{\mathbb{H}_m}(B_m x, B_m y) - d_{geo}(x, y)| \geq \delta).
\]

**Lemma 5.2.** We claim that \( P(|d_{\mathbb{H}_m}(B_m x, B_m y) - d_{geo}(x, y)| \geq \delta) \leq 2e^{-2\delta^2 m} \) for all pairs \( x \sim y \).

**Proof.** The difference in the metrics is the average of the random variables:

\[
X_i(x, y) = \frac{|\text{sgn}(x \cdot \theta_i) - \text{sgn}(y \cdot \theta_i)|}{2} - \frac{\cos^{-1}(x \cdot y)}{\pi}.
\]
These are independent centered Bernoulli random variables that satisfy a large deviation inequality which is uniform in the Bernoulli parameter \( \theta \),

\[
P\left( \left| \frac{1}{m} \sum_{i=1}^{m} X_i(x, y) \right| > \delta \right) \leq 2e^{-2\delta^2 m}.
\]

**The Union Bound for the RIP:** The last expression above provides a bound for the probability that one pair \( x \sim y \) fails the \( \delta \)-RIP. Now, summing over all pairs,

\[
\sum_{x \sim y} P\left( \left| \frac{1}{m} \sum_{i=1}^{m} \left[ \text{sgn}(x \cdot \theta_i) - \text{sgn}(y \cdot \theta_i) \right] \right| - \frac{\cos^{-1}(x \cdot y)}{\pi} \right| \geq \delta \right) \leq 2 \left( \frac{n}{2} \right) \cdot e^{-2\delta^2 m}.
\]

We can bound this probability with \( \epsilon \) and solve for \( m \):

\[
m \geq \frac{\ln \frac{n^2}{\pi}}{2\delta^2}.
\]

For all \( m \) greater than this bound, \( m \) must satisfy the \( \delta \)-RIP with probability at least \( 1 - \epsilon \).

### 6. A Phase Transition for the Restricted Isometry Property

For special \( X \), we analyze the property of \( B_m \) is a \( \delta \)-RIP. Again, the size of the window’s dependence on \( |X| \), is very weak. This time the dependence is in terms of \( \ln \ln |X| \).

**Theorem 6.1.** Fix \( 0 < \delta < \frac{1}{2} \), fix \( 0 < \epsilon_2 < \epsilon_1 < 0.99 \). Let \( X \) be \( n \) pairwise orthogonal vectors in \( \mathbb{S}^{N-1} \), and let \( P_{\text{RIP}}(m) \) be the probability that \( B_m \) satisfies \( \delta - \text{RIP} \). If \( n \geq 800 \), then \( 1 - \epsilon_1 < P_{\text{RIP}}(m) \) when:

\[
(6.1.1) \quad m \leq \frac{1}{2} \ln(1 - 4\delta^2) + \delta \ln \frac{1 + 2\delta}{1 - 2\delta} \left[ \ln \frac{n(n - 1)}{2\sqrt{2\pi e} \ln \frac{1}{1 - \frac{\epsilon_1}{\epsilon_2}}} - \ln \frac{n(n - 1)}{2\sqrt{2\pi e} \ln \frac{1}{1 - \frac{\epsilon_1}{\epsilon_2}}} \right]
\]

and \( P_{\text{RIP}}(m) < 1 - \epsilon_2 \) when

\[
(6.1.2) \quad m \geq \frac{1}{2} \ln(1 - 4\delta^2) + \delta \ln \frac{1 + 2\delta}{1 - 2\delta} \left[ \ln \frac{n(n - 1)e^{\frac{1}{\epsilon_2}}}{2\sqrt{2\pi \ln \frac{1}{1 - \frac{\epsilon_1}{\epsilon_2}}}} + \ln \frac{n(n - 1)e^{\frac{1}{\epsilon_2}}}{2\sqrt{2\pi \ln \frac{1}{1 - \frac{\epsilon_1}{\epsilon_2}}}} \right].
\]

Additionally, the phase transition is bounded as follows:

\[
P(B_m \text{ is a } \delta - \text{RIP}) \in \left[ e^{-\lambda_2^\delta - \eta^\delta}, e^{-\lambda_1^\delta + \eta^\delta} \right]
\]
where $\lambda_1^\delta = \frac{n(n-1)}{2} \cdot e^{\frac{1}{2} \ln(1-4\delta^2) + \delta \ln \frac{1-2\delta}{1+2\delta}}$, $\lambda_2^\delta = \frac{n(n-1)}{2} \cdot \frac{1}{\sqrt{2\pi m}} \cdot e^{m \left( \frac{1}{2} \ln(1-4\delta^2) + \delta \ln \frac{1-2\delta}{1+2\delta} \right)}$, and $\eta^\delta = \binom{n}{2} \left[ (p^\delta)^2 + 4(n-2)(p^\delta)^2 \right]$ where $p^\delta \leq \frac{1}{\sqrt{2\pi m}} \cdot e^{m \left( \frac{1}{2} \ln(1-4\delta^2) + \delta \ln \frac{1-2\delta}{1+2\delta} \right)}$.

The graph below shows a simulation of the RIP property with $\delta = 0.2$. The red line is the bound (6.1.1), the green line is (6.1.2). The jagged blue line is the simulated value of the probability of $B_m$ being a 0.2-RIP. The line is jagged, due to the discrete nature of the Hamming metric. The latter fact is of course a complication implicit in our proof.

![Graph showing simulation of RIP property](image)

We will again analyze the phase transition from the perspective of the birthday problem. To do this we will count all $x \sim y \in X$ that fail the RIP property, namely:

$$W^\delta_{x \sim y} = \begin{cases} 1, & B_m x = B_m y \\ 0, & \text{otherwise} \end{cases}.$$  

Then $W^\delta = \sum_{x \sim y} W^\delta_{x \sim y}$ is a sum of Bernoulli random variables. All $W^\delta_{x \sim y}$ are i.i.d with probability $p^\delta = \mathbb{P}(|d_{\text{ham}}(B_m x, B_m y) - d_{\text{geo}}(x, y)| \geq \delta)$. By the general form of the Stein-Chen approximation $W^\delta$ is close to a Poisson distribution in total variation. We make this precise below:

$$d_{TV}(W^\delta, \text{Poi}(\lambda^\delta)) \leq \eta^\delta$$
where

\[
\eta^\delta = \min \left( 1, \frac{1}{\lambda^\delta} \right) \sum_{x \sim y} \left( p^\delta_{x \sim y} \right)^2 + \sum_{r \sim s \in \mathcal{M}_{x \sim y}} \left( p^\delta_{r \sim s} \right)^2 + \mathbb{E}[W^\delta_{x \sim y}, W^\delta_{r \sim s}] \]

(6.1.3)

**Lemma 6.2.** We claim that \( \lambda^\delta = \binom{n}{2} P(|Y - m/2| > m\delta) \) where \( Y \) is Bin\((m, \frac{1}{2})\).

**Proof.** We know that \( \lambda^\delta = \binom{n}{2} p^\delta \). In this special case, the geodesic distances between the points in \( X \) is \( \frac{1}{2} \), which reduces \( p^\delta \) to:

\[
p^\delta = P \left( \left| d_{Hm}(B_m x, B_m y) - \frac{1}{2} \right| \geq \delta \right)
\]

\[
= P \left( \left| \sum_{i=1}^{n} \frac{|\text{sgn}(x \cdot \theta_i) - \text{sgn}(y \cdot \theta_i)|}{2} - \frac{m}{2} \right| \geq m\delta \right)
\]

For each \( i \), \( \frac{|\text{sgn}(x \cdot \theta_i) - \text{sgn}(y \cdot \theta_i)|}{2} \) is Bernoulli with parameter \( d_{\text{geo}}(x, y) = \frac{1}{2} \). The \( \theta_i \) are independently so \( Y = \sum_{i=1}^{n} \frac{|\text{sgn}(x \cdot \theta_i) - \text{sgn}(y \cdot \theta_i)|}{2} \) is Bin\((m, \frac{1}{2})\) and we can rewrite \( p^\delta \) as:

\[
p^\delta = P \left( \left| Y - \frac{m}{2} \right| \geq m\delta \right)
\]

\( \square \)

**Lemma 6.3.** We can bound \( \lambda^\delta \) as such:

\[
e^{-\frac{1}{6}} \sqrt{2\pi} \frac{e^{\frac{1}{2} \ln(1-4\delta^2) + \delta \ln \frac{1-2\delta}{1+2\delta}}}{m} \leq \lambda^\delta \leq e^{\frac{1}{12}} \sqrt{m}
\]

where \( \Lambda = \frac{(n)}{\sqrt{2\pi}} \cdot e^{m|\frac{1}{2} \ln(1-4\delta^2) + \delta \ln \frac{1-2\delta}{1+2\delta}|} \). Additionally for \( \delta < 0.25 \), we can approximate this statement as:

\[
\left( \frac{n}{2} \right) \cdot e^{-\frac{1}{6}} \sqrt{2\pi m} \cdot e^{-2m\delta^2 - 4m\delta^3} \leq \lambda^\delta \leq \left( \frac{n}{2} \right) \cdot e^{\frac{1}{12}} \sqrt{2\pi m} \cdot e^{-2m\delta^2 + 8m\delta^3}
\]

**Proof.** As previously defined,

\[
\lambda^\delta = \left( \frac{n}{2} \right) P(|Y - \frac{m}{2}| > m\delta)
\]

(6.3.1) \[ = 2 \left( \frac{n}{2} \right) \left( \frac{1}{2^m} \right) \left[ \left( \frac{m}{2} + m\delta \right) + \ldots + \left( \frac{m}{m} \right) \right].
\]

Now we will use Sterling’s approximation,

\[
\sqrt{2\pi m} \left( \frac{m}{e} \right)^m \leq m! \leq \sqrt{2\pi m} \left( \frac{m}{e} \right)^m \cdot e^{\frac{1}{12m}},
\]
to obtain bounds for $\lambda^\delta$, but for the upper bound, we will use the fact that $e^{\frac{1}{8m}} \leq e^{\frac{1}{12}}$ for $m \geq 1$.

Let $A = \lceil \frac{m}{2} + m\delta \rceil$ and let us assess only the first term of the $\lambda^\delta$ sum since it is the largest.

$$e^{-\frac{1}{2}} \cdot \sqrt{\frac{m}{2\pi A(m - A)}} \cdot \frac{(m/e)^m}{A^A (m - A)^{m-A}} \leq \frac{m}{A} \leq \sqrt{\frac{m}{2\pi A(m - A)}} \cdot \frac{(m/e)^m}{A^A (m - A)^{m-A}} \cdot e^{\frac{1}{12}}$$

$$\frac{\left(\frac{m}{4} - m^2\delta^2\right)^{\frac{m}{2}} \left(\frac{m + m\delta}{m - m\delta}\right)^{m\delta}}{\left(\frac{1}{4} - \delta^2\right)^{\frac{m}{2}} \left(1 - 2\delta^2\right)^{m\delta}} \leq \frac{\left(\frac{m}{4} - m^2\delta^2\right)^{\frac{m}{2}} \left(\frac{m + m\delta}{m - m\delta}\right)^{m\delta}}{\left(\frac{1}{4} - \delta^2\right)^{\frac{m}{2}} \left(1 + 2\delta\right)^{m\delta}}.$$

We can rewrite these bounds in terms of $\Lambda$:

$$\frac{2^m e^{-\frac{1}{6}}}{\binom{n}{2} \sqrt{m}} \leq \frac{(m/e)^n}{\Lambda} \leq \frac{2^m e^{\frac{1}{12}} \sqrt{m}}{\binom{n}{2}}$$

Using this inequality for the first term in (6.3.1), we gain a lower bound for $\lambda^\delta$:

$$\frac{\Lambda e^{-\frac{1}{6}} \sqrt{m}}{\sqrt{m}} \leq \lambda^\delta.$$  

For the upper bound, we have at most $m$ summands in (6.3.1) so the upper bound is:

$$\lambda^\delta \leq \Lambda e^{\frac{1}{12}} \sqrt{m}$$

For $\delta < 0.25$ we can simply these two statements above using the Taylor approximation for $\ln(1 - x)$, $\frac{-x}{1 - x} \leq \ln(1 - x) \leq -x$

$$\binom{n}{2} \cdot \frac{e^{-\frac{1}{6}}}{\sqrt{2\pi m}} \cdot e^{-2m\delta^2 - 4m\delta^3} \leq \lambda^\delta \leq \binom{n}{2} \cdot \frac{e^{\frac{1}{12}} \sqrt{m}}{\sqrt{2\pi}} \cdot e^{-2m\delta^2 + 4m\delta^3}.$$

\[ \square \]

**Lemma 6.4.** We claim that: $\eta^\delta = \binom{n}{2} \left[ (p^\delta)^2 + 4(n - 2)(p^\delta)^2 \right]$ where $\eta^\delta$ is given in (6.1.3). Using this we can bound $1 - \mathbb{P}(W^\delta \geq 1)$ in the window

$$[e^{-\lambda^\delta_2} - \eta^\delta, e^{-\lambda^\delta_2} + \eta^\delta]$$

where $\lambda^\delta_1 = \binom{n}{2} \cdot \frac{e^{-\frac{1}{6}}}{\sqrt{2\pi m}} \cdot e^{m \left[ \frac{-1}{4} \ln(1 - 4\delta^2) + \delta \ln \frac{1}{4 + 25} \right]}$ and $\lambda^\delta_2 = \binom{n}{2} \cdot \frac{e^{\frac{1}{12}} \sqrt{m}}{\sqrt{2\pi}} \cdot e^{m \left[ \frac{-1}{4} \ln(1 - 4\delta^2) + \delta \ln \frac{1}{4 + 25} \right]}.$
Proof. We recall:

$$
\eta^\delta = \min \left( \frac{1}{\lambda^\delta}, \sum_{x \sim y} \left( (p^\delta)^2 + \sum_{r \sim s \in \mathcal{N}_{x \sim y}} ((p^\delta)^2 + \mathbb{E}[W_{x \sim y}^\delta, W_{r \sim s}^\delta]) \right) \right).
$$

In order to estimate \( \eta^\delta \), we need to estimate \( |\mathcal{N}_{x \sim y}| \). Because the only coordinates that are dependent on \( x \sim y \) are those that share exactly one coordinate with \( x \sim y \), \( |\mathcal{N}_{x \sim y}| \leq 2(n - 2) \). There are two ways this can happen: either \( r \) or \( s \) shares a coordinate with \( x \sim y \). There are \( n - 2 \) ways to choose the remaining coordinates. Assuming pairwise independence, we can estimate the size of \( \eta^\delta \):

$$
\eta^\delta \leq \left( \frac{n}{2} \right) \left( (p^\delta)^2 + 4(n - 2)(p^\delta)^2 \right).
$$

We can now bound \( 1 - \mathbb{P}(W^\delta \geq 1) \) which is equal to \( \mathbb{P}(B_m \text{ is } \delta \text{- RIP}) \):

$$
|\mathbb{P}(W^\delta \geq 1) - \mathbb{P}(\text{Poi}(\lambda^\delta) \geq 1)| < \eta^\delta
$$

$$
1 - \mathbb{P}(W^\delta \geq 1) = \left[ e^{-\lambda^\delta} - \eta^\delta, e^{-\lambda^\delta} + \eta^\delta \right]
$$

where \( \lambda^\delta = \left( \frac{n}{2} \right) \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{n \left[ \frac{1}{2} \ln(1 - 4\delta^2) + \delta \ln \left( \frac{1 - 4\delta^2}{1 + 4\delta^2} \right) \right]} \) and \( \lambda^\delta = \left( \frac{n}{2} \right) \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{n \left[ \frac{1}{2} \ln(1 - 4\delta^2) + \delta \ln \left( \frac{1 - 4\delta^2}{1 + 4\delta^2} \right) \right]} \).

Lemma 6.5. \( W^\delta_{x \sim y} \) are pairwise independent for all pairs, \( x \sim y \).

Proof. To show that \( W^\delta_{x \sim y} \) are pairwise independent, it is sufficient to show that

$$
\mathbb{E}(W^\delta_{x \sim y}, W^\delta_{r \sim s}) = (p^\delta)^2
$$

$$
\mathbb{E}(W^\delta_{x \sim y}, W^\delta_{r \sim s}) = \mathbb{P}(W^\delta_{x \sim y} = W^\delta_{r \sim s} = 1) = \mathbb{P}(W^\delta_{x \sim y} = W^\delta_{y \sim s} = 1).
$$

The only non-trivial case is when \( x \sim y \) and \( r \sim s \) share a common point. We can rewrite this probability as:

$$
\sum_{k \in \mathbb{H}_m} \mathbb{P} \left( |d_{\mathbb{H}_m}(B_m x, k) - \frac{1}{2}| \geq \delta, |d_{\mathbb{H}_m}(B_m s, k) - \frac{1}{2}| \geq \delta, B_m y = k \right)
$$

where \( k \) is an element of \( \mathbb{H}_m \). After an orthogonal transformation, we can take \( x, y, s \) to be the first three coordinates vectors \( e_1, e_2, e_3 \). The distribution of the \( \theta_j \) are unchanged. The signs of the coordinates of the \( \theta_j \) are independent, so the events are independent. Because all of these events are independent, the probability can be written as:

$$
\sum_{k \in \mathbb{H}_m} \mathbb{P} \left( |d_{\mathbb{H}_m}(B_m x, k) - \frac{1}{2}| \geq \delta \right) \mathbb{P} \left( |d_{\mathbb{H}_m}(B_m s, k) - \frac{1}{2}| \geq \delta \right) \mathbb{P}(B_m y = k)
$$

$$
= \sum_{k \in \mathbb{H}_m} (p^\delta)^2 \cdot 2^{-m}.
$$
Because each of these probabilities is identical distributed to $P(|d_{H_m}(B_m x, B_m y) - \frac{1}{2}| \geq \delta)$.

There are are $2^m$ elements in $H_m$, we get that $P(W_{x\sim y} = W_{r\sim s} = 1) = (p^\delta)^2$, as desired.

**Solving For $m$:** Using the previous bounds on $\lambda^\delta$,

$$\lambda^\delta_1 < \lambda^\delta < \lambda^\delta_2$$

where $\lambda^\delta_1 = (\frac{n}{2}) \cdot \frac{e^{-\frac{1}{2} \lambda^\delta_2}}{\sqrt{2\pi m}} \cdot e^{m[\frac{1}{2} \ln(1-4\delta^2) + \delta \ln \frac{\sqrt{2\pi}}{\sqrt{2\pi m}}]}$ and $\lambda^\delta_2 = (\frac{n}{2}) \cdot \frac{e^{-\frac{1}{2} \lambda^\delta_2}}{\sqrt{2\pi}} \cdot e^{m[\frac{1}{2} \ln(1-4\delta^2) + \delta \ln \frac{\sqrt{2\pi}}{\sqrt{m}}]}$.

Fix $0 < \epsilon_2 < \epsilon_1 < 1$, let $X$ be $n$ pairwise orthogonal vectors in $S^{N-1}$, and let $P_{RIP}(m)$ be the probability that $B_m$ is one-to-one, then $1 - \epsilon_1 < P_{RIP}(m)$ when:

$$1 - e^{-\lambda^\delta_1} + \eta^\delta \leq \epsilon_1.$$

and $P_{RIP}(m) \leq 1 - \epsilon_2$ when

$$1 - e^{-\lambda^\delta_2} - \eta^\delta \geq \epsilon_2.$$

In order to ensure that $\eta^\delta$ is very small compared to $1 - e^{-\lambda^\delta_2}$ and $1 - e^{-\lambda^\delta_1}$, we want $\eta^\delta \leq 0.01(1 - e^{-\lambda^\delta_2})$. If we fix $n$ and choose $m$ such that $\lambda^\delta_2 \leq 1$, then using the inequality $\frac{\lambda^\delta_2}{2} \leq \lambda^\delta_2 - \frac{(\lambda^\delta_2)^2}{2}$ it is sufficient to bound $\eta^\delta$ as

$$\eta^\delta \leq 0.01 \frac{\lambda^\delta_2}{2}.$$

Manipulating this statement we gain:

$$(1 + 2(n - 2)) \cdot \frac{e^{\frac{1}{12} \sqrt{m}}}{\sqrt{2\pi}} \cdot e^{m[\frac{1}{2} \ln(1-4\delta^2) + \delta \ln \frac{\sqrt{2\pi}}{\sqrt{m}}]} \leq 0.005.$$

Because $1 + 2(n - 2) \leq 4\frac{n(n-1)}{2}$, we can rewrite this inequality as:

$$\frac{4\lambda^\delta_2}{n} \leq 0.005.$$

Since we assumed that $\lambda^\delta_2 \leq 1$, we can rewrite this inequality as $\frac{4}{n} \leq 0.005$ and solve for $n$:

$$n \geq 800.$$

This means that if $n \geq 800$, we have $\eta^\delta \leq 0.01(1 - e^{-\lambda^\delta_2})$ and $\eta^\delta \leq 0.01(1 - e^{-\lambda^\delta_1})$ which allows us to rewrite our inequalities and get bounds on $m$:
\[ \frac{1}{2} \ln m + m \left( \frac{1}{2} \ln(1 - 4\delta^2) + \delta \ln \frac{1 + 2\delta}{1 - 2\delta} \right) \leq \epsilon_1 \]

\[ \ln \frac{n(n - 1)}{2\sqrt{2\pi} e^{\frac{3}{2}} \ln \frac{1}{1 - \epsilon_1}} \leq \frac{1.01(1 - e^{-2\delta^2})}{\epsilon_1} \]

\[ m \leq \frac{1}{2} \ln(1 - 4\delta^2) + \delta \ln \frac{1 + 2\delta}{1 - 2\delta} \left[ \ln \frac{n(n - 1)}{2\sqrt{2\pi} e^{\frac{3}{2}} \ln \frac{1}{1 - \epsilon_1}} + \ln \ln \frac{n(n - 1)}{2\sqrt{2\pi} e^{\frac{3}{2}} \ln \frac{1}{1 - \epsilon_1}} \right] \]

and

\[ m \geq \frac{1}{2} \ln(1 - 4\delta^2) + \delta \ln \frac{1 + 2\delta}{1 - 2\delta} \left[ \ln \frac{n(n - 1)e^{\frac{1}{12}}}{2\sqrt{2\pi} \ln \frac{1}{1 - \epsilon_2}} + \ln \ln \frac{n(n - 1)e^{\frac{1}{12}}}{2\sqrt{2\pi} \ln \frac{1}{1 - \epsilon_2}} \right]. \]

Let \( q = \frac{1}{2} \ln(1 - 4\delta^2) + \delta \ln \frac{1 + 2\delta}{1 - 2\delta} \). We remark that \( q \) is approximately \( \frac{1}{2\delta} \). Let \( r = \ln \frac{n(n - 1)}{2\sqrt{2\pi}} \) then:

\[ m \leq q \left[ r + \ln \frac{1}{e^{\frac{1}{12}} \ln \frac{1}{1 - \epsilon_1}} - \ln r - \ln \ln \frac{1}{e^{\frac{1}{12}} \ln \frac{1}{1 - \epsilon_1}} \right] \]

and

\[ m \geq q \left[ r + \ln \frac{e^{\frac{1}{12}}}{\ln \frac{1}{1 - \epsilon_2}} + \ln r + \ln \ln \frac{e^{\frac{1}{12}}}{\ln \frac{1}{1 - \epsilon_2}} \right]. \]

This is a statement of the main theorem by inspection.

7. Acknowledgments

We would like to thank Dr. Michael Lacey and Dr. Robert Kesler for their assistance and mentorship. We would also like to thank the Georgia Institute of Technology and the NSF for their funding and support.
References

[1] N. Alon and B. Klartag, *Optimal compression of approximate inner products and dimension reduction*, ArXiv e-prints (October 2016), available at 1610.00239.

[2] Richard Arratia, Larry Goldstein, and Louis Gordon, *Poisson approximation and the chen-stein method*, Statistical Science (1990), 403–424.

[3] Peter Binev, Wolfgang Dahmen, Ronald DeVore, Philipp Lamby, Daniel Savu, and Robert Sharp-ley, *Compressed sensing and electron microscopy*, Modeling nanoscale imaging in electron microscopy, 2012, pp. 73–126.

[4] Sanjoy Dasgupta and Anupam Gupta, *An elementary proof of the johnson-lindenstrauss lemma*, International Computer Science Institute, Technical Report (1999), 99–006.

[5] Massimo Fornasier and Holger Rauhut, *Compressive sensing*, Handbook of mathematical methods in imaging, 2011, pp. 187–228.

[6] Wassily Hoeffding, *Probability inequalities for sums of bounded random variables*, Journal of the American statistical association 58 (1963), no. 301, 13–30.

[7] William B Johnson and Joram Lindenstrauss, *Extensions of lipschitz mappings into a hilbert space*, Contemporary mathematics 26 (1984), no. 189-206, 1.

[8] Michael Lustig, David Donoho, and John M Pauly, *Sparse mri: The application of compressed sensing for rapid mr imaging*, Magnetic resonance in medicine 58 (2007), no. 6, 1182–1195.