CS-67720 Metric Embedding Theory and Its Algorithmic Applications

Lecture 3

Lecturer: Nova Fandina *Fall 2018/19*

We continue the study of embedding into a normed space. Particularly, we discuss the question of dimension reduction.

3.1 Dimension Reduction

The central question of dimensionality reduction is the following. Let X be a normed space of dimension d, $S \subset X$ of size n. Can S be embedded into a normed space Y of dimension k (where $k << d$) and distortion α? Here α depends on d and n, k depends on d,n and α.

3.1.1 Isometric Dimension Reduction

We start with isometric dimension reductions. Let us cite several results of this line of research.

1. Each n-point set of ℓ_2 is isometrically embeddable into ℓ_2^{n-1}.
2. Every n-point subset of ℓ_p, $1 \leq p < \infty$ is isometrically embeddable into $\ell_p^{\binom{n}{2}}$ [3, 6].
   - for $p = \infty$ we have seen that $d \leq n-1$;
   - for $p = 1$ it holds that $d \geq \binom{n-2}{2}$, for $n \geq 4$;
   - for $1 < p < 2$ it holds that $d \geq \binom{n-1}{2}$, for $n \geq 3$.
3. Any n-point set of ℓ_2 is isometrically embeddable into $\ell_p^{O(n^2)}$, for any $p \geq 1$. This can be proven by non-trivial arguments involving Dvoretzky theorem and compactness of the unit ball in ℓ_p.

3.1.2 Dimension Reduction With Distortion

For a given $1 \leq p \leq \infty$, $n \geq 2$, $\epsilon > 0$, denote by $k_p(n, \epsilon)$ the dimension required to embed any n-point subset of ℓ_p into $\ell_p^{k_p(n,\epsilon)}$, with distortion $1 + \epsilon$. What is the tightest bound on $k_p(n, \epsilon)$? For $\epsilon = \Theta(1)$?

- [For $p = 2$] Johnson-Lindenstrauss lemma states: $k_2(n, \epsilon) = O\left(\frac{\log n}{\epsilon^2}\right)$.

  The very recent result of [9] states that this bound is tight: there exists a subset $X \subset \ell_2^d$ on n points such that any embedding $f : X \to \ell_2^k$ with distortion $1 + \epsilon$ requires dimension at least $\Omega\left(\frac{\log n}{\epsilon^2}\right)$, for any given $1 > \epsilon \geq 1/n^{0.499}$.

- [For $p = 1$] Newman and Rabinovich [13]: $k_1(n, \epsilon) = O(\frac{n}{\epsilon^2})$.

  The following lower bounds are known: There exists an n-point set inℓ_1^d, such that every embedding of it into ℓ_1^k, with distortion $\alpha \geq 1$, requires dimension $k = n^{\Omega\left(\frac{1}{\alpha^2}\right)}$ [4, 10]. For the refined distortion of $1 + \epsilon$, Andoni et. al [2] showed the existence of an n-point set in ℓ_1^d, such that every embedding of it into ℓ_1^k, requires dimension $k = n^{1-O(1/\log(\frac{1}{\epsilon}))}$.

- [For $p = \infty$] Matausek [12]: $\Omega(n^{\frac{1}{\alpha}}) = k_\infty(n, \alpha) = O(n^{\frac{1}{\alpha}} \cdot ploylog(n))$.

- [For $p \neq 1, 2, \infty$] The possibility of dimension reduction with guarantees similar to those of the JL lemma is still an open question.

From all those results, we will show the upper bound on $k_2(n, \epsilon)$ (JL lemma), and a (weak) lower bound on $k_p(n, \Theta(1))$, for all $1 \leq p \leq \infty$.

We start with a weak lower bound, which is of independent interest, as it uses a quite standard volume argument. Let E_n denote an n-point equilateral space.

Claim 3.1. *$\forall\ \alpha > 1$, $1 \leq p \leq \infty$, for any $f : E_n \to \ell_p^k$ with dist$(f) = \alpha$, it holds that $k = \Omega(\log_{\max\{\alpha,2\}} n)$.*

Proof. Let $f : E_n \to \mathbb{R}^k$ be an embedding with distortion α. Without loss of generality f is non-expansive, and one of the points of E_n is mapped to 0. Therefore, $\forall x, y \in E_n$, $\frac{1}{\alpha} \leq \|f(x) - f(y)\|_p \leq 1$. Then, the images of E_n are in the unit ball of ℓ_p^k, and the set $\{f(x)|x \in E_n\}$ is $(1/\alpha)$-separated. Hence, by the lemma we proved in the last lecture, $n = |f(E)| \leq (1 + 2\alpha)^k$, for any $\alpha > 1$. Therefore, $k \geq \frac{\log_2 n}{\log_2(1+2\alpha)}$. If $\alpha \geq 2$, then $\log_2(1 + 2\alpha) = O(\log_2 \alpha)$, implying $k = \Omega\left(\frac{\log_2 n}{\log_2 \alpha}\right) = \Omega(\log_\alpha n)$. If $1 < \alpha < 2$, then $\log_2(1 + 2\alpha) = O(1)$, implying $k = \Omega(\log_2 n)$, as required. □

Essentially, the above argument is correct for any k dimensional normed space.

3.2 Dimension Reduction in ℓ_2: the Johnson-Lindenstrauss Lemma

We present the proof of the striking result of dimensionality reduction in ℓ_2.

Theorem 3.2 (J-L Lemma [8]). *For an n-point $X \in \ell_2^d$, for $0 < \epsilon < 1/2$, there exists $f : X \to \ell_2^k$, for $k = O\left(\frac{\log n}{\epsilon^2}\right)$, with dist$(f) = 1 + \epsilon$.*

The idea of the proof is to project a given point set into a randomly chosen k dimensional subspace (from some distribution), and to show that with positive probability, all the pairwise distances are distorted by at most $1+\epsilon$. We present an elegant proof of Indyk and Motwani [7], who showed that it's actually enough to "project" onto k (appropriately normalized) normally distributed directions (i.e., the requirement on the orthogonality can be dropped).

Proof. Let $r^{(1)}, r^{(2)}, \ldots, r^{(k)} \subset \mathbb{R}^d$ be i.i.d., according to d-dimensional Gaussian distribution. Namely, $\forall 1 \leq i \leq k$, $\forall 1 \leq j \leq d$, $r_j^{(i)} \sim N(0, 1)$. For all $x \in X$ set:

$$f(x) = \frac{1}{\sqrt{k}}(< x, r^{(1)} >, < x, r^{(2)} >, \ldots, < x, r^{(k)} >).$$

We will show that for a given $0 < \epsilon < 1$, it is enough to set $k = O\left(\frac{\log n}{\epsilon^2}\right)$, to have for all $x \neq y \in X$: $1/(1 + \epsilon)\|x - y\|_2^2 \leq \|f(x) - f(y)\|_2^2 \leq (1 + \epsilon)\|x - y\|_2^2$, with constant probability.

Note that it is enough to show that the inequality holds for any $x \neq y \in X$, with probability at least $1 - \frac{1}{n^2}$. Since, by the union bound:

$$Pr\left[\exists\, x \neq y \in X, s.t. \frac{\|f(x) - f(y)\|_2^2}{\|x - y\|_2^2} \notin [1/(1 + \epsilon), 1 + \epsilon]\right] \leq$$

$$\sum_{(x,y)\in\binom{X}{2}} Pr\left[(x, y)\ is\ s.t.\ \frac{\|f(x) - f(y)\|_2^2}{\|x - y\|_2^2} \notin [1/(1 + \epsilon), 1 + \epsilon]\right] < n^2/2 \cdot \frac{1}{n^2} = 1/2.$$

Also, since the embedding is linear, it is enough to prove the inequality for any vector of a unit norm. Let us formulate the precise lemma we prove, from which the theorem follows:

Lemma 3.3. *For $0 < \epsilon < 1$, let $k = O\left(\frac{\log n}{\epsilon^2}\right)$. Then, for every $x \in \mathbb{R}^d$, s.t. $\|x\|_2$=1, with probability at least $1 - \frac{1}{n^2}$, $1/(1+\epsilon) \le \|f(x)\|_2^2 \le 1+\epsilon$.*

We use the fact that the normal distribution is 2-stable distribution:

Claim 3.4. *Let $\{r_i\}_{i=1}^t$ be a sequence of independently chosen normal random variables: $r_i \sim N(\mu_i, {\sigma_i}^2)$. Let $\{\gamma_i\}_{i=1}^t \in \mathbb{R}$ be a sequence of scalars. Then, for the random variable defined by $Z = \sum_{i=1}^t \gamma_i r_i$ it holds that $Z \sim N(\sum_{i=1}^t \gamma_i \mu_i, \sum_{i=1}^t {\gamma_i}^2 {\sigma_i}^2)$.*

First, we estimate the expected value of the length of the projected vector. For any $x \in \mathbb{R}^n$, with $\|x\|_2 = 1$: $E[\|f(x)\|_2^2] = \frac{1}{k}\sum_{i=1}^k E\left[(< x, r^{(i)} >)^2\right]$. By Claim 3.4, $< x, r^{(i)} >= y_i \sim N\left(0, \sum_{i=1}^n {x_i}^2\right) \equiv N(0,1)$. Therefore, $E[(y_i)^2] = \sigma^2 + \mu^2 = 1$, implying $E\left[\|f(x)\|_2^2\right] = 1$.

Next, we show that the length of the projected vector is highly concentrated around its expectation. Let $\alpha = 1+\epsilon$, then

$$Pr\left[\|f(x)\|_2^2 > \alpha\right] = Pr\left[\frac{1}{k}\sum_{i=1}^k y_i^2 > \alpha\right] = Pr\left[\sum_{i=1}^k y_i^2 > k\alpha\right] = (\text{ for some } s > 0\ ) =$$

$$Pr\left[e^{s\cdot\sum {y_i}^2} > e^{sk\alpha}\right] \le \frac{E\left[e^{s\cdot\sum {y_i}^2}\right]}{e^{sk\alpha}} = \frac{E\left[\prod e^{s{y_i}^2}\right]}{e^{sk\alpha}} = (y_i \text{ are independent }) = \frac{\prod E\left[e^{s{y_i}^2}\right]}{e^{sk\alpha}} = \left(\frac{E\left[e^{sy^2}\right]}{e^{s\alpha}}\right)^k,$$

where $y \sim N(0,1)$. Next we compute the expected value:

$$E\left[e^{sy^2}\right] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{sy^2} \cdot e^{\frac{-y^2}{2}} dy = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{(s-\frac{1}{2})y^2} dy,$$

changing variables to $u^2 = (1-2s)y^2$, we get

$$E\left[e^{sy^2}\right] = \frac{1}{\sqrt{1-2s}}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{\frac{-u^2}{2}} du = \frac{1}{\sqrt{1-2s}}.$$

So far we have shown that

$$Pr[\|f(x)\|_2^2 > \alpha] \le \left(e^{s\alpha} \cdot \sqrt{1-2s}\right)^{-k},$$

choosing $s = \frac{1}{2} - \frac{1}{2\alpha}$, and recalling that $\alpha = 1+\epsilon$ results in

$$Pr[\|f(x)\|_2^2 > \alpha] \le \left(e^{\frac{\alpha-1}{2}} \cdot \sqrt{\frac{1}{\alpha}}\right)^{-k} = \left(e^{\frac{\epsilon}{2}} \cdot \left(\frac{1}{1+\epsilon}\right)^{\frac{1}{2}}\right)^{-k} = \left(\frac{e^\epsilon}{1+\epsilon}\right)^{\frac{-k}{2}} = e^{\frac{-k}{2}(\epsilon - ln(1+\epsilon))}.$$

For $|\epsilon| < 1$, we have $\ln(1+\epsilon) \le \epsilon - \frac{1}{4}\epsilon^2$, then $Pr[\|f(x)\|^2 > 1+\epsilon] \le e^{-\frac{k}{2}\cdot\frac{\epsilon^2}{4}} = e^{-\frac{k\epsilon^2}{8}}$. Thus, setting $k \ge \frac{8(2\ln n + \ln 2)}{\epsilon^2}$, we obtain $Pr[\|f(x)\|^2 > 1+\epsilon] \le \frac{1}{2n^2}$. In the similar way we show that $Pr[\|f(x)\|_2^2 < \frac{1}{1+\epsilon}] \le \frac{1}{2n^2}$ (with slightly different k). The lemma then follows by the union bound. □

Remark 3.5. *In their proof, Dasgupta and Gupta [5], have used more precise estimation: for any $\epsilon \ge 0$, $\ln(1+\epsilon) \le \epsilon - \frac{\epsilon^2}{2} + \frac{\epsilon^3}{3}$, which results in a slightly better bound on dimension $k = \frac{8\ln n}{\epsilon^2 - 2/3\epsilon^3}$. This is the best currently known bound. In addition, instead of choosing normally distributed variables, [1] showed that one can choose uniformly distributed variables from $\{-1, 1\}$, which will result in the same bound on dimension.*

Remark 3.6. *The exactly same process can be used to embed any n-points in ℓ_2 into ℓ_1. The analysis shows that this implies an embedding into dimension $O(\frac{(\log 1/\epsilon)n}{\epsilon^2})$*

3.3 Dimension Reduction with Bounded ℓ_q-Distortion

So far we have discussed dimension reduction techniques that aim to preserve the interpoint distances of the input metric space in the worst case. However, for various proximity problems with strong dependence on dimension, the worst case guarantee is an excessive requirement. Hence, it is natural to consider techniques that trade worst case distortion for dimension, while preserving a particular parameter of interest. Therefore, the question of embedding into a given number of $k \geq 1$ dimensions seems to be of a particular algorithmic interest.

For the worst case distortion, the following was proved by Matousek[11]: Every n-point set in ℓ_2^d can be embedded into ℓ_2^k, with $O(n^{2/k}\sqrt{\log n/k})$ distortion, for $3 \leq k \leq \ln n$. In the same paper the lower bound was established: For all $k \geq 2$, there is an n-point metric space in ℓ_2^{k+1}, such that any embedding of it into ℓ_2^k requires distortion of $\Omega(n^{2/k})$.

In this lecture we will show that the above negative result does not exclude a positive outcome (while considering the ℓ_q-distortion measure of the quality).

Recall the definitions we have discussed in the first lecture.

Definition 3.1 (ℓ_q-distortion). *Let (X, d_x) be any n-point metric space, and $f : X \to Y$ be an embedding. For any $u \neq v \in X$, define*

$$dist_f(u,v) = \max\{contr_f(u,v), expans_f(u,v)\}.$$

For all $1 \leq q < \infty$, define

$$\ell_q\text{-}dist(f) = \left(\frac{\sum_{u\neq v\in X}(dist_f(u,v))^q}{\binom{|X|}{2}}\right)^{\frac{1}{q}}, \ell_\infty\text{-}dist = \sup_{u\neq v\in X}\{dist_f(u,v)\}.$$

The definition generalizes in the obvious way to the case of any distribution over the pairs of X. Let us also introduce the following related and useful notions.

Definition 3.2 (ℓ_q-contraction, ℓ_q-expansion).

$$\ell_q\text{-}contr(f) = \left(\frac{\sum_{u\neq v\in X}(contr_f(u,v))^q}{\binom{|X|}{2}}\right)^{\frac{1}{q}},$$

$$\ell_q\text{-}expans(f) = \left(\frac{\sum_{u\neq v\in X}(expans_f(u,v))^q}{\binom{|X|}{2}}\right)^{\frac{1}{q}}.$$

It is almost trivial to conclude that $\ell_q\text{-}dist(f) \leq \ell_q\text{-}contr(f) + \ell_q\text{-}expans(f)$.

In what follows we discuss dimension reduction for Euclidean subsets, while preserving ℓ_q-distortion. Particularly, we analyze the ℓ_q-distortion behavior of the JL transform, while embedding into a given number of $k \geq 1$ dimensions. It turns out, that analyzing ℓ_q-contraction and ℓ_q-expansion of the JL transform results in the worst case distortion guarantees (providing the best known bound on required dimension). Let us formalize this in the following subsection.

3.3.1 Dimension Reduction with ℓ_q-distortion

We present several results: first, we analyze the JL transform in terms of its ℓ_q-distortion behavior, next we show that the analysis implies the (w.c.) JL Lemma, and finally, we prove an optimality of the JL dimension reduction for a certain value of q.

The following claim, actually, implies the upper bound results.

Claim 3.7. *Given any n-point subset $X \subset \ell_2^d$, any $k \geq 2$, and any $1 \leq q < k$, the JL transform $f : X \to \ell_2^k$ is such that with constant probability it holds that*

$$\ell_q\text{-}contr(f) = 1 + O\left(\frac{q}{k-q}\right),$$

and

$$\ell_q\text{-}expans(f) = 1 + O\left(\frac{q}{k}\right).$$

Proof. We estimate $E\left[(\ell_q\text{-}contr(f))^q\right]$ and $E\left[(\ell_q\text{-}expans(f))^q\right]$, and then apply Jensen's and Markov's inequalities to obtain the claim.

Since f is a linear map, and by the linearity of expectation it is enough to estimate $E\left[(contr_f(x))^q\right]$, and $E\left[(expans_f(x))^q\right]$ for any unit vector $x \in \mathbb{R}^d$. For such x we have

$$(contr_f(x))^q = \left(\frac{1}{\|f(x)\|_2^2}\right)^{q/2} = \left(\frac{k}{Z}\right)^{q/2}, \text{ where } Z \sim \chi_k^2.$$

The same for the expansion

$$(expans_f(x))^q = \left(\|f(x)\|_2^2\right)^{q/2} = \left(\frac{Z}{k}\right)^{q/2}, \text{ where } Z \sim \chi_k^2.$$

Recall the *pdf* function of χ_k^2: $\forall x \geq 0$, $pdf_{\chi_k^2}(x) = \frac{x^{\frac{k}{2}-1}}{2^{\frac{k}{2}}\Gamma(\frac{k}{2})e^{\frac{x}{2}}}$, where Γ is the Gamma function. (Recall the Gamma function's definition: for $t > 0$, $\Gamma(t) = \int_0^\infty x^{t-1}e^{-x}\,dx$). Denote $l = \lceil q/2 \rceil$, and then $q = l - r$. Therefore,

$$E\left[(\ell_q\text{-}contr(f))^q\right] = \int_0^\infty \left(\frac{k}{x}\right)^{\frac{q}{2}} \frac{x^{\frac{k}{2}-1}}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)e^{\frac{x}{2}}}\,dx = \frac{\left(\frac{k}{2}\right)^{\frac{q}{2}}\Gamma\left(\frac{k}{2}-\frac{q}{2}\right)}{\Gamma\left(\frac{k}{2}\right)} = \frac{\left(\frac{k}{2}\right)^{l}\Gamma\left(\left(\frac{k}{2}+r\right)-l\right)}{\left(\frac{k}{2}\right)^{r}\Gamma\left(\frac{k}{2}\right)}.$$

Applying the rule $\Gamma(x) = (x-1)\Gamma(x-1)$, l times we get

$$\int_0^\infty \left(\frac{k}{x}\right)^{\frac{q}{2}} \frac{x^{\frac{k}{2}-1}}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)e^{\frac{x}{2}}}\,dx = \frac{\left(\frac{k}{2}\right)^{l}}{\prod_{i=0}^{l-1}\left(\frac{k}{2}-\frac{q}{2}+i\right)} \cdot \frac{\Gamma\left(\frac{k}{2}+r\right)}{\left(\frac{k}{2}\right)^{r}\Gamma\left(\frac{k}{2}\right)}.$$

Using the known inequality we obtain that for all $k \geq 1$ and for all $0 \leq r < 1$ it holds

$$\frac{\Gamma(\frac{k}{2}+r)}{\left(\frac{k}{2}\right)^{r}\Gamma(\frac{k}{2})} \leq 1,$$

we conclude

$$\int_0^\infty \left(\frac{k}{x}\right)^{\frac{q}{2}} \frac{x^{\frac{k}{2}-1}}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)e^{\frac{x}{2}}}\,dx \leq \frac{\left(\frac{k}{2}\right)^{l}}{\prod_{i=0}^{l-1}\left(\frac{k}{2}-\frac{q}{2}+i\right)} \leq \left(1+\frac{q}{k-q}\right)^{\lceil\frac{q}{2}\rceil},$$

as required. For the ℓ_q-expansion denote $\frac{q}{2} = l + r$, where $l = \lfloor\frac{q}{2}\rfloor \geq 0$, and $0 \leq r < 1$. Therefore

$$E\left[(\ell_q\text{-}expans(f))^q\right] = \int_0^\infty \left(\frac{x}{k}\right)^{\frac{q}{2}} \frac{x^{\frac{k}{2}-1}}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)e^{\frac{x}{2}}}\,dx = \frac{\Gamma\left(\frac{k}{2}+\frac{q}{2}\right)}{\left(\frac{k}{2}\right)^{\frac{q}{2}}\Gamma\left(\frac{k}{2}\right)} = \frac{\Gamma\left(\left(\frac{k}{2}+r\right)+l\right)}{\left(\frac{k}{2}\right)^{l}\left(\frac{k}{2}\right)^{r}\Gamma\left(\frac{k}{2}\right)},$$

applying similar consideration to the previous case obtain that

$$E\left[(\ell_q\text{-}expans(f))^q\right] \leq \left(1+\frac{q}{k}\right)^{\lfloor\frac{q}{2}\rfloor},$$

as required. This finishes the proof of the claim. ◻

3.3.2 Proving the JL Lemma from ℓ_q-Distortion Components

It can be shown that using the bounds on ℓ_q-$contr(f)$ and ℓ_q-$expans(f)$, it is possible to prove the existence of the embedding with the same guarantees as in the the original JL lemma.

3.3.3 Analyzing ℓ_q-Distortion of the JL Transform

The following theorem suggests that reducing dimension of any finite Euclidean subset, into a constant number of dimensions can be done with a constant average distortion.

Theorem 3.8. *Given any n-point set $X \subset \ell_2^d$, any integer $2 \leq k < n$, and any $1 \leq q < k$ there exists an embedding $f : X \to \ell_2^k$ such that*

$$\ell_q\text{-}dist(f) = 1 + O\left(\frac{q}{k-q}\right) + O\left(\frac{1}{\sqrt{k}}\right)$$

Proof of Theorem 3.8. The embedding $f : X \to \ell_2^k$ is the JL transform. By the observation we made before, it holds:

$$E[(\ell_q\text{-}dist(f))^q] \leq E[(\ell_q\text{-}contr(f))^q] + E[(\ell_q\text{-}expans(f))^q].$$

Then, by Claim 3.7 and by Jensen's inequality we get:

$$E[\ell_q\text{-}dist^{(\Pi)}(f)] \leq 2^{1/q}\left(1 + \frac{q}{k-q}\right)^{1/2} \leq \left(1 + \frac{1}{q}\right)\left(1 + \frac{q}{2(k-q)}\right).$$

For $q \geq \sqrt{k}$, we have that the above is bounded by $1 + O(q/(k-q))$, as required. For $q \leq \sqrt{k}$, we use the following (easy to show) relation: $\forall q < q'$, it holds that $\ell_q\text{-}dist(f) \leq \ell_{q'}\text{-}dist(f)$. Therefore, for all $q \leq \sqrt{k}$, we obtain the result by taking $q' = \sqrt{k}$. $\square$

References

[1] Dimitris Achlioptas. Database-friendly random projections: Johnson-lindenstrauss with binary coins. *Journal of Computer and System Sciences*, 66(4):671 – 687, 2003. Special Issue on {PODS} 2001.

[2] A. Andoni, M.S. Charikar, O. Neiman, and H.L. Nguyen. Near linear lower bound for dimension reduction in l1. In *Foundations of Computer Science (FOCS), 2011 IEEE 52nd Annual Symposium on*, pages 315–323, Oct 2011.

[3] K. Ball. Isometric embedding in ℓ_p-spaces. *European J. Combin.*, 11(4):305–311, 1990.

[4] Bo Brinkman and Moses Charikar. On the impossibility of dimension reduction in l1. *J. ACM*, 52(5):766–788, September 2005.

[5] Sanjoy Dasgupta and Anupam Gupta. An elementary proof of a theorem of johnson and lindenstrauss. *Random Structures and Algorithms*, 22(1):60–65, 2003.

[6] B. Fichet. l_p-spaces in data analysis. *Classification and Related Methods of Data Analysis*, pages 439– 444, 1988.

[7] Piotr Indyk and Rajeev Motwani. Approximate nearest neighbors: Towards removing the curse of dimensionality. In *Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing*, STOC '98, pages 604–613, New York, NY, USA, 1998. ACM.

[8] W. B. Johnson and J. Lindenstrauss. Extensions of lipschitz mappings into a hilbert space. 1982.

[9] Kasper G. Larsen and Jelani Nelson. Optimality of the Johnson-Lindenstrauss Lemma. *arxiv:1609.02094*.

[10] James R. Lee and Assaf Naor. Embedding the diamond graph in lp and dimension reduction in l1, 2004.

[11] Jiří Matoušek. Bi-Lipschitz embeddings into low-dimensional Euclidean spaces. *Commentat. Math. Univ. Carol.*, 31(3):589–600, 1990.

[12] Jiří Matoušek. On the distortion required for embedding finite metric spaces into normed spaces. *Israel J. Math*, 93:333–344, 1996.

[13] I. Newman and Y. Rabinovich. Finite Volume Spaces and Sparsification. *ArXiv e-prints*, February 2010.