Almost Optimal Explicit Johnson-Lindenstrauss Transformations

Raghu Meka

UT Austin

`raghu@cs.utexas.edu`

Abstract

The Johnson-Lindenstrauss lemma is a fundamental result in probability with several applications in the design and analysis of algorithms in high dimensional geometry. Most known constructions of linear embeddings that satisfy the Johnson-Lindenstrauss property involve randomness. We address the question of explicitly constructing such embedding families and provide a construction with an almost optimal use of randomness: For $0 < \delta, \varepsilon < 1$, we give an explicit generator $G : \{0,1\}^r \to \mathbb{R}^{s \times n}$ for $s = O(\log(1/\delta)/\varepsilon^2)$ such that for all $w \in \mathbb{R}^n$, $\|w\| = 1$,

$$\Pr_{y \in_u \{0,1\}^r} \left[\, |\, \|G(y)w\|^2 - 1| > \varepsilon \, \right] \leq \delta,$$

and seed-length $r = O\left(\log(n/\delta) \cdot \log\left(\frac{\log(n/\delta)}{\varepsilon} \right) \right)$. In particular, for $\delta = 1/\text{poly}(n)$ and fixed $\varepsilon > 0$ we get seed-length $O(\log n \log \log n)$. Previous constructions required at least $O(\log^2 n)$ random bits to get polynomially small error.

1 Introduction

The celebrated Johnson-Lindenstrauss lemma (JLL) [JL84] has been a fundamental tool in geometric algorithms for the past two decades and is by now a standard technique for handling high dimensional data. Among its many known variants (see [AV99, DG03, IM98, Mat08]), we use the following version originally due to Achlioptas [Ach03] as reference [1].

Theorem 1.1 (Achlioptas). *For all $w \in \mathbb{R}^n$, $\|w\| = 1$, $\varepsilon > 0$, $s = C \log(1/\delta)/\varepsilon^2$,*

$$\Pr_{A \in_u \{1,-1\}^{s \times n}} \left[\, |\, \|(1/\sqrt{s})\, Aw\|^2 - 1 \, | \geq \varepsilon \, \right] \leq \delta.$$

We say a family of random matrices has the JL property (or is a JL family) if a similar condition holds. In typical applications of JLL, δ is taken to be $1/\text{poly}(n)$ and the goal is to embed a given set of $\text{poly}(n)$ points in n dimensions to $O(\log n)$ dimensions with distortion at most ε for a fixed constant ε. This is the setting we concern ourselves with.

Most results on embedding the Euclidean space as above are probabilistic in nature. However, a simple probabilistic argument shows that there exists a fixed collection of $\text{poly}(n, 1/\delta)$ linear mappings satisfying the JL property. Despite much attention, the best known constructions of JL families use at least $O(\log n \log(1/\delta))$ random bits [CW09]. Besides being a natural problem in geometry as well as derandomization, an explicit family of Johnson-Lindenstrauss transformations

---

[1]Throughout, C denotes a universal constant. For a multiset S, $x \in_u S$ denotes a uniformly random element of S.

would likely help derandomize other geometric algorithms and metric embedding constructions. Further, having an explicit construction is of fundamental importance for streaming algorithms as storing the entire matrix (as opposed to the randomness required to generate the matrix) is often too expensive in the streaming context.

Our main result is an explicit generator that takes roughly $O(\log n \log \log n)$ random bits and outputs a matrix $A \in \mathbb{R}^{s \times n}$ satisfying the JL property.

Theorem 1.2. *For every $0 < \varepsilon, \delta < 1$, there exists an explicit generator $G : \{0,1\}^r \to \mathbb{R}^{n \times s}$ for $s = C \log(1/\delta)/\varepsilon^2$, such that for every $w \in \mathbb{R}^n, \|w\| = 1$,*

$$\Pr_{y \in_u \{0,1\}^r} [\, | \, \|G(y)w\|^2 - 1 \, | > \varepsilon \,] \leq \delta.$$

The seed-length of the generator is $r = O\left(\log(n/\delta) \cdot \log\left(\frac{\log(n/\delta)}{\varepsilon}\right)\right)$.

Our construction is elementary in nature using only standard tools in derandomization such as k-wise independence and oblivious samplers [Gol97]. The construction has the additional property that for $\delta = 1/\text{poly}(n)$, the matrix-vector product $G(y)w$ can be computed efficiently in time $O(n \log n)$. The computational efficiency does not follow directly from the dimensions of $G(y)$, as in our construction $G(y)$ is obtained by composing several matrices some of which are of dimension $\tilde{O}(\sqrt{n}) \times n$. Nevertheless, the large matrices are obtained from the discrete Fourier transform matrix facilitating fast matrix-vector product computations.

Further, as one of the motivations for derandomizing JLL is its potential applications in streaming, it is important that the entries of the generated matrices be computable in small space. We observe that for any $i \in [s], j \in [n], y \in \{0,1\}^r$, the entry $G(y)_{ij}$ can be computed in space $O(\log n)$ and time $O(n^{2+o(1)})$ (for fixed $\varepsilon, \delta > 1/\text{poly}(n)$). (See proof of Theorem 1.2 for the exact bound)

1.1 Related Work

Independently and concurrently, Kane and Nelson [KN10a] obtained a construction that is similar in spirit to ours and achieves a slightly better seed-length of $r = O(\log n + \log(1/\delta) \log(\log(1/\delta)/\varepsilon))$. Note that for the important case of δ polynomially small, our seed-length is the same as theirs.

Karnin et al. [KRS09] construct an explicit family $\mathcal{A}_{n,s}$ for embedding $\mathbb{R}^n$ into $\mathbb{R}^s$ with distortion $\varepsilon = 1/s^{-C}$ and error $\delta = 1/s^{-c}$. The seed-length they achieve is $(1 + o(1)) \log n + O(\log^2 s)$.

The works of Diakonikolas et al. [DKN10] and Meka and Zuckerman [MZ10] construct pseudo-random generators for degree 2 threshold functions achieving a seed-length of $\log n \cdot \text{poly}(1/\delta)$ for fooling with error at most δ. As derandomizing the JL lemma is a special case of fooling degree 2 PTFs, these works give a JL family with seed-length $\log n \cdot \text{poly}(1/\delta)$.

The best known explicit JL family is the construction of Clarkson and Woodruff [CW09] who show that a random scaled Bernoulli matrix with $O(\log(1/\delta))$-wise independent entries satisfies the JL lemma. We make use of their result in our construction.

We also note that there are efficient non-black box derandomizations of JLL, [EIO02], [Siv02]. These works, take as input N points in $\mathbb{R}^n$, and deterministically compute an embedding (that depends on the input set) into $\mathbb{R}^{O(\log N)/\varepsilon^2}$ which preserves all pairwise distances between the given set of N points.

Finally, we remark that our goal as well as result is very different from those of the recent works [AC09, AL09, DKS10, AL10, KN10b] on *fast* or *sparse* Johnson-Lindenstrauss transformations as

pioneered by the seminal work of Ailon and Chazelle [AC09]. The goal in these works is to design a family of embedding matrices for which the matrix-vector products Ax can be computed efficiently (usually $O(n \log n)$) and are mainly concerned with the setting where the desired error probability δ is exponentially small. In contrast, we are mainly interested in the case where δ is polynomially small but want to save on randomness.

1.2 Outline of Construction

Our construction is based on a simple iterative scheme: We reduce the dimension from n to $\tilde{O}(\sqrt{n})$ using k-wise independence and *oblivious samplers* [Gol97] and iterate for $O(\log \log n)$ steps.

Fix a vector $w \in \mathbb{R}^n$ with $\|w\| = 1$ and let $\delta = 1/\text{poly}(n)$. We first use an idea of Ailon and Chazelle [AC09] who give a family of unitary transformations $\mathcal{R}$ from $\mathbb{R}^n \to \mathbb{R}^n$ such that for every $w \in \mathbb{R}^n$ and $V \in_u \mathcal{R}$, the vector Vw is *regular* in the sense that $\|Vw\|_\infty = O(\sqrt{\log n/n})$. We derandomize their construction using limited independence to get a bound of $\|Vw\|_\infty = O(n^{-1/4})$.

We next observe that for a vector $w \in \mathbb{R}^n$, with $\|w\|_\infty = O(n^{-1/4}\|w\|_2)$ projecting onto a random set of $O(n^{1/2}\log(1/\delta)/\varepsilon^2)$ coordinates preserves the ℓ_2 norm with distortion at most ε. We then note that the random set of coordinates can be chosen using efficient samplers as in [Gol97]. The idea of using samplers is due to Karnin et al. [KRS09] who use samplers for a similar purpose.

Finally, iterating the above scheme $O(\log \log n)$ times we obtain an embedding of $\mathbb{R}^n$ to $\mathbb{R}^{\text{poly} \log n}$ using roughly $O(\log n \log \log n)$ random bits. We then apply the result of Clarkson and Woodruff and perform the final embedding into $O(\log(1/\delta)/\varepsilon^2)$ dimensions by using a random scaled Bernoulli matrix with $O(\log(1/\delta))$-wise independent entries.

2 Preliminaries

Let $H_n \in \{-1/\sqrt{n}, 1/\sqrt{n}\}^{n \times n}$ be the normalized Hadamard matrix such that $H_n^T H_n = I_n$ (we drop the suffix n when dimension is clear from context). While the Hadamard matrix is known to exist for powers of 2, for clarity, we ignore this minor technicality and assume that it exists for all n.

We make use of Khintchine-Kahane inequalities (cf. [LT91]).

Lemma 2.1 (Khintchine-Kahane). *For every* $w \in \mathbb{R}^n$, $x \in_u \{1, -1\}^n$, $k > 0$,

$$\mathbb{E}[\,|\langle w, x\rangle|^k\,] \leq k^{k/2}\,\mathbb{E}[\,|\langle w, x\rangle|^2\,]^{k/2} = k^{k/2}\|w\|^k.$$

We use efficient oblivious samplers as in [Gol97].

Theorem 2.2. *For every* ε, δ *there exists* $s = C \log(1/\delta)/\varepsilon^2$ *such that the following holds. There exists an explicit collection of subsets of* $[n]$, $\mathcal{S}(n, \varepsilon, \delta)$, *with each* $S \in \mathcal{S}$ *of cardinality* $|S| = s$, *and* $|\mathcal{S}| = n \cdot \text{poly}(1/\varepsilon, 1/\delta)$ *such that for every function* $f : [n] \to [0, 1]$,

$$\Pr_{S \in_u \mathcal{S}}\left[\left|\frac{1}{s}\sum_{i \in S} f(i) - \mathbb{E}_{i \in_u [n]} f(i)\right| > \varepsilon\right] \leq \delta.$$

Corollary 2.3. *For every* $\varepsilon, \delta, B > 0$ *there exists* $s = C \log(1/\delta)B^2/\varepsilon^2$ *such that the following holds. There exists an explicit collection of subsets of* $[n]$, $\mathcal{S}(n, B, \varepsilon, \delta)$, *with each* $S \in \mathcal{S}$ *of cardinality* $|S| = s$, *and* $|\mathcal{S}| = n \cdot \text{poly}(B/\varepsilon, 1/\delta)$ *such that for every function* $f : [n] \to [0, B]$,

$$\Pr_{S \in_u \mathcal{S}}\left[\left|\frac{1}{s}\sum_{i \in S} f(i) - \mathbb{E}_{i \in_u [n]} f(i)\right| > \varepsilon\right] \leq \delta.$$

Proof. Follows by taking $\mathcal{S} = \mathcal{S}(n, \varepsilon/B, \delta)$ as in Theorem 2.2 and using the condition of Theorem 2.2 for $\bar{f} : [n] \to [0, 1]$ defined by $\bar{f}(i) = f(i)/B$. $\qquad\square$

Finally, we use the following result of Clarkson and Woodruff.

Theorem 2.4 (Theorem 2.2, [CW09])**.** *There exist constants c, C such that the following holds. For $0 < \varepsilon, \delta < 1$, $s = c \log(1/\delta)/\varepsilon^2$, let $A \in \mathbb{R}^{s \times n}$ be a random matrix with entries in $\{-1/\sqrt{s}, 1/\sqrt{s}\}$ that are $(C \log(1/\delta))$-wise independent. Then, for every $w \in \mathbb{R}^n$, $\|w\| = 1$,*

$$\Pr_A[\, |\, \| Aw \|^2 - 1 \,| \geq \varepsilon \,] \leq \delta.$$

Note that for all k, m, constructions of k-wise independent spaces over $\{1, -1\}^m$ with seed-length $O(k \log m)$ are known.

3   Main Construction

Suppose that $\delta > 1/n^c$ for a constant $c > 0$. If not we first embed the input vector into $\mathbb{R}^N$ for $N = \lceil 1/\delta \rceil$ by retaining the first n coordinates and setting the other coordinates to be 0. The parameters we get by working over $\mathbb{R}^N$ will be the same as those of Theorem 3.4 and hence of Theorem 1.2. Further, we assume that $\log(1/\delta)/\varepsilon^2 = o(n)$ as else JLL is not interesting.

As outlined in the introduction, we first give a family of rotations to regularize vectors in $\mathbb{R}^n$. For a vector $x \in \mathbb{R}^n$, let $D(x) \in \mathbb{R}^{n \times n}$ be the diagonal matrix with $D(x)_{ii} = x_i$.

Lemma 3.1. *Let $x \in \{1, -1\}^n$ be drawn from a k-wise independent distribution. Then, for every $w \in \mathbb{R}^n$ with $\|w\| = 1$,*

$$\Pr[\, \| HD(x)w \|_\infty > n^{-(1/2-\alpha)} \,] \leq \frac{k^{k/2}}{n^{\alpha k - 1}}.$$

Proof. Let $v = HD(x)w$. Then, for $i \in [n]$, $v_i = \sum_j H_{ij} x_j w_j$ and $\mathbb{E}[v_i^2] = \sum_j H_{ij}^2 w_j^2 = 1/n$. By Markov's inequality and Khintchine-Kahane inequality Lemma 2.1,

$$\Pr[\, |v_i| > n^{-(1/2-\alpha)} \,] \leq \mathbb{E}[v_i^k] \cdot n^{(1/2-\alpha)k} \leq k^{k/2} n^{(1/2-\alpha)k} / n^{k/2} = k^{k/2} n^{-\alpha k}.$$

The claim now follows from a union bound over $i \in [n]$. $\qquad\square$

We now give a family of transformations for reducing n dimensions to $\tilde{O}(n^{1/2})$ dimensions with distortion at most ε. For $S \subseteq [n]$, let $\mathcal{P}_S : \mathbb{R}^n \to \mathbb{R}^{|S|}$ be the projection onto the coordinates in S.

Lemma 3.2. *Let $\mathcal{S} \equiv \mathcal{S}(n, n^{1/4}, \varepsilon, \delta)$, $s = O(n^{1/2} \log(1/\delta)/\varepsilon^2)$ be as in Corollary 2.3 and let $\mathcal{D}$ be a k-wise independent distribution over $\{1, -1\}^n$. For $S \in_u \mathcal{S}$, $x \leftarrow D$, let random linear transformation $A_{S,x} : \mathbb{R}^n \to \mathbb{R}^s$ be defined by $A_{S,x} = \sqrt{n/s} \cdot \mathcal{P}_S \cdot HD(x)$. Then, for every $w \in \mathbb{R}^n$ with $\|w\| = 1$,*

$$\Pr[\, |\|A_{S,x}(w)\|^2 - 1| \geq \varepsilon \,] \leq \delta + k^{k/2}/n^{k/8-1}.$$

Proof. Let $v = HD(x)w$. Then, $\|v\| = 1$ and by Lemma 3.1 applied for $\alpha = 1/8$,

$$\Pr[\, \|v\|_\infty > n^{-3/8} \,] \leq k^{k/2}/n^{k/8-1}.$$

4

Now suppose that $\|v\|_\infty \leq n^{-3/8}$. Define $f : [n] \to \mathbb{R}$ by $f(i) = n \cdot v_i^2 \leq n \cdot n^{-3/4} = n^{1/4} = B$. Then,

$$\|A_{S,x}(w)\|^2 = (n/s)\|\mathcal{P}_S(v)\|^2 = \frac{1}{s}\sum_{i \in S} nv_i^2 = \frac{1}{s}\sum_{i \in S} f(i),$$

and $\mathbb{E}_{i \in_u [n]} f(i) = (1/n)\sum_i n \cdot v_i^2 = 1$. Therefore, by Corollary 2.3,

$$\Pr[\,|\,\|A_{S,x}(w)\|^2 - 1\,| \geq \varepsilon\,] = \Pr_S\left[\left|\frac{1}{s}\sum_{i \in S} f(i) - \mathbb{E}_{i \in_u [n]} f(i)\right| \geq \varepsilon\right] \leq \delta.$$

The claim now follows. $\qquad\qquad\square$

We now recursively apply the above lemma. Fix $\varepsilon, \delta > 0$. Let $\mathcal{A}(n,k) : \mathbb{R}^n \to \mathbb{R}^{s(n)}$ be the collection of transformations $\{A_{S,x} : S \in_u \mathcal{S}, x \leftarrow D\}$ as in the above lemma for $s(n) = s(n, n^{1/4}, \varepsilon, \delta) = Cn^{1/2}\log(1/\delta)/\varepsilon^2$. Note that we can sample from $\mathcal{A}(n,k)$ using $O(k \log n + \log(1/\varepsilon) + \log(1/\delta))$ random bits. Let $n_0 = n$, and let $n_{i+1} = s(n_i)$. Let $k_0 = 16(c+1)$ (recall that $\delta > 1/n^c$) and $k_{i+1} = 2^i k_0$. The parameters n_i, k_i are chosen so that $1/n_i^{k_i}$ is always polynomially small. Fix $t > 0$ to be chosen later so that $k_i < n_i^{1/8}$ for $i < t$.

Lemma 3.3. *For $A_0 \in_u \mathcal{A}(n_0, k_0), A_1 \in_u \mathcal{A}(n_1, k_1), \cdots, A_{t-1} \in_u \mathcal{A}(n_{t-1}, k_{t-1})$ chosen independently, and $w \in \mathbb{R}^n$, $\|w\| = 1$,*

$$\Pr[\,(1-\varepsilon)^t \leq \|A_{t-1}\cdots A_1 A_0(w)\|^2 \leq (1+\varepsilon)^t\,] \geq 1 - t\delta - \sum_{i=0}^{t-1} \frac{k_i^{k_i/2}}{n_i^{k_i/8-1}}.$$

Proof. The proof is by induction on $i = 1, \ldots, t$. For $i = 1$, the claim is same as Lemma 3.2. Suppose the statement is true for $i-1$ and let $v = A_{i-1}\cdots A_0(w)$. Then, $v \in \mathbb{R}^{n_i}$ and by Lemma 3.2 applied to $\mathcal{A}(n_i, k_i)$,

$$\Pr[\,(1-\varepsilon)\|v\|^2 \leq \|A_i(v)\|^2 \leq (1+\varepsilon)\|v\|^2\,] \geq 1 - \delta - \frac{k_i^{k_i/2}}{n_i^{k_i/8-1}}.$$

The claim now follows by induction. $\qquad\qquad\square$

What follows is a series of simple calculations to bound the seed-length and error from the above lemma. Observe that

$$n^{(1/2)^i} \leq n_i = n^{(1/2)^i} \cdot \left(\frac{C\log(1/\delta)}{\varepsilon^2}\right)^{1+(1/2)+\cdots+(1/2)^{i-1}} \leq n^{(1/2)^i}\left(\frac{C\log(1/\delta)}{\varepsilon^2}\right)^2. \qquad (3.1)$$

Let $t = O(\log\log n)$ be such that $2^t = \log n/8\log\log n$. Then, $n_t \leq \log^8 n \cdot (C\log(1/\delta)/\varepsilon^2)^2$, and for $i < t$,

$$k_i < k_t = 16(c+1)2^t = 2(c+1)\log n/\log\log n < \log n = n^{(1/2)^t/8} < n_t^{1/8} < n_i^{1/8}, \qquad (3.2)$$

5

where we assumed that $\log\log n > 2c+2$. Therefore, the error in Lemma 3.3 can be bounded by

$$t\delta + \sum_{i=0}^{t-1}\frac{k_i^{k_i/2}}{n_i^{k_i/8-1}} \le t\delta + n\sum_{i=0}^{t-1}n_i^{-k_i/16} \qquad\text{(Equation 3.2)}$$

$$\le t\delta + n\sum_{i=0}^{t-1}n^{-(1/2)^i\cdot16(c+1)\cdot2^i/16} \qquad\text{(Equation 3.1)}$$

$$\le t\delta + t/n^c \le 2t\delta \qquad\text{(as } \delta > 1/n^c\text{)}.$$

Note that,

$$k_i\log n_i \le 16(c+1)\cdot2^i(\log n/2^i+2C\log\log(1/\delta)+4\log(1/\varepsilon)) = O(\log n+\log n\log(1/\varepsilon)/\log\log n).$$

Therefore, the randomness needed after $t = O(\log\log n)$ iterations is

$$\sum_{i=0}^{t-1}O(k_i\log n_i + \log(1/\varepsilon) + \log(1/\delta)) = O(\log n\cdot\log(\log n/\varepsilon)).$$

Combining the above arguments (and simplifying the resulting expression for seed-length) we get:

Theorem 3.4. *There exists an explicit generator that takes $O(\log(n/\delta)\cdot\log(\log(n/\delta)/\varepsilon))$ random bits and outputs a linear transformation $A : \mathbb{R}^n \to \mathbb{R}^m$ for $m = O((\log^{10}(n/\delta))/\varepsilon^4)$, so that for every $w \in \mathbb{R}^n$, $\|w\| = 1$,*

$$\Pr[\,|\,\|Aw\|^2 - 1\,| > (C\log\log n)\varepsilon\,] \le (C\log\log n)\delta.$$

We can now obtain our final construction of explicit Johnson-Lindenstrauss families by composing the above family with that of Theorem 2.4.

Proof of Theorem 1.2. Follows by composing the transformations of the above theorem for $\varepsilon' = \varepsilon/C\log\log n$, $\delta' = \delta/C\log\log n$ with those of Theorem 2.4 using $O(\log(1/\delta))$-wise independence. The additional randomness required is $O(\log(1/\delta)\log m) = O(\log(1/\delta)(\log\log(n/\delta) + \log(1/\varepsilon))$.

We next bound the time for computing matrix-vector products for the matrices we output. Let $\delta > 1/n^c$. Note that for $i < t$, the matrices A_i of Lemma 3.3 are of the form $\mathcal{P}_S \cdot H_{n_i}D(x)$ for a k-wise independent string $x \in \{1,-1\}^{n_i}$. Thus, for any vector $w_i \in \mathbb{R}^{n_i}$, A_iw_i can be computed in time $O(n_i\log n_i)$ using the discrete Fourier transform. Therefore, for any $w = w_0 \in \mathbb{R}^{n_0}$, the product $A_{t-1}\cdots A_1A_0w_0$ can be computed in time

$$\sum_{i=0}^{t-1}O(n_i\log n_i) \le O(n\log n) + \log n\cdot\sum_{i=1}^{t-1}O\left(n^{1/2^i}(\log(1/\delta)/\varepsilon^2)^2\right) \qquad\text{(Equation 3.1)}$$

$$= O(n\log n + \sqrt{n}\log n\log^2(1/\delta)/\varepsilon^4).$$

It is easy to check that the above bound dominates the time required to perform the final embedding.

A similar calculation shows that for indices $i \in s, j \in [n]$, the entry $G(y)_{ij}$ of the generated matrix can be computed in space $O\left(\sum_i \log n_i\right) = O(\log n + \log(1/\varepsilon)\cdot\log\log n)$ by expanding the product of matrices and enumerating over all intermediary indices. The time required to perform the calculation is $O(s\cdot n_t\cdot n_{t-1}\cdots n_0) = n^2\cdot(\log n/\varepsilon)^{O(\log\log n)}$. $\qquad\square$

References

[AC09]   Nir Ailon and Bernard Chazelle. The fast Johnson-Lindenstrauss transform and approximate nearest neighbors. *SIAM J. Comput.*, 39(1):302–322, 2009.

[Ach03] Dimitris Achlioptas. Database-friendly random projections: Johnson-lindenstrauss with binary coins. *J. Comput. Syst. Sci.*, 66(4):671–687, 2003.

[AL09]   Nir Ailon and Edo Liberty. Fast dimension reduction using rademacher series on dual bch codes. *Discrete & Computational Geometry*, 42(4):615–630, 2009.

[AL10]   Nir Ailon and Edo Liberty. Almost optimal unrestricted fast Johnson-Lindenstrauss transform. *CoRR*, abs/1005.5513, 2010. To appear in SODA 2011.

[AV99]   Rosa I. Arriaga and Santosh Vempala. An algorithmic theory of learning: Robust concepts and random projection. In *FOCS*, pages 616–623, 1999.

[CW09]   Kenneth L. Clarkson and David P. Woodruff. Numerical linear algebra in the streaming model. In *STOC*, pages 205–214, 2009.

[DG03]   Sanjoy Dasgupta and Anupam Gupta. An elementary proof of a theorem of johnson and lindenstrauss. *Random Struct. Algorithms*, 22(1):60–65, 2003.

[DKN10] Ilias Diakonikolas, Daniel Kane, and Jelani Nelson. Bounded independence fools degree 2 threshold functions. In *FOCS*, 2010.

[DKS10] Anirban Dasgupta, Ravi Kumar, and Tamás Sarlós. A sparse Johnson-Lindenstrauss transform. In *STOC*, pages 341–350, 2010.

[EIO02] Lars Engebretsen, Piotr Indyk, and Ryan O'Donnell. Derandomized dimensionality reduction with applications. In *SODA*, pages 705–712, 2002.

[Gol97] Oded Goldreich. A sample of samplers - a computational perspective on sampling (survey). *Electronic Colloquium on Computational Complexity (ECCC)*, 4(20), 1997.

[IM98]   Piotr Indyk and Rajeev Motwani. Approximate nearest neighbors: towards removing the curse of dimensionality. In *STOC*, pages 604–613, 1998.

[JL84]   W B Johnson and J Lindenstrauss. Extensions of Lipschitz mappings into a Hilbert space. *Contemp Math*, 26:189–206, 1984.

[KN10a] Daniel Kane and Jelani Nelson, 2010. Personal Communication.

[KN10b] Daniel Kane and Jelani Nelson. A derandomized sparse Johnson-Lindenstrauss transform, 2010. arxiv:1006.3585.

[KRS09] Zohar Karnin, Yuval Rabani, and Amir Shpilka. Explicit dimension reduction and its applications. *Electronic Colloquium on Computational Complexity (ECCC)*, 16(121), 2009.

[LT91]   Michel Ledoux and Michel Talagrand. *Probability in Banach spaces: isoperimetry and processes.* Springer, 1991.

[Mat08] Jirí Matousek. On variants of the johnson-lindenstrauss lemma. *Random Struct. Algorithms*, 33(2):142–156, 2008.

[MZ10]   Raghu Meka and David Zuckerman. Pseudorandom generators for polynomial threshold functions. In *STOC*, pages 427–436, 2010.

[Siv02] D. Sivakumar. Algorithmic derandomization via complexity theory. In *STOC*, pages 619–626, 2002.