6 Bourgain's Theorem

In the last lecture, we studied low-distortion embeddings of general metrics into ℓ_∞ using only a small number of dimensions. In this lecture we will extend that result and give embeddings into ℓ_p for all $1 \le p < \infty$; this fundamental result is essentially due to Bourgain (1985). For completeness, we begin by stating the last lecture.

Theorem 6.1 (Matoušek (1996)) *For any metric space (X, d),*

$$(X, d) \overset{D}{\hookrightarrow} \ell_\infty^{O(Dn^{2/D} \log n)}$$

Theorem 6.1 uses the following embedding:

1. Pick random subsets S_{ij}, with $j = 1, 2, \ldots, \frac{D}{2}$ as follows

   At *level* j, for each $i = 1, 2, \ldots, m$ (where $m = O(n^{2/D} \log n)$), form the set S_{ij} by picking each node independently with probability $(n^{-2/D})^j$.

2. Let $f_{ij}(x) = d(x, S_{ij})$ for all $x \in X$. Finally, set

$$f(x) = \oplus_{j=1}^{D/2} \oplus_{i=1}^{m} f_{ij}(x). \tag{6.1}$$

An easy application of the triangle inequality shows that the above embedding is a contraction. That is, for any two points x and y in X, in every dimension, the distance between x and y is less than $d(x, y)$. This fact will be useful later.

Note that if we use $D = O(\log n)$, we get that $(X, d) \overset{\log n}{\hookrightarrow} \ell_\infty^{O(\log^2 n)}$. Interestingly, we immediately get the fact that f is a low distortion embedding into ℓ_1 as well, albeit with a worse distortion. Indeed, the contraction, when viewing f as a map into ℓ_1, is at most the contraction with respect to ℓ_∞, which is $O(\log n)$. Moreover, since the embedding is a contraction in every dimension, the *expansion* with respect to ℓ_1 is at most $\log^2 n$, the number of dimensions. Thus we get

$$(X, d) \overset{\log^3 n}{\hookrightarrow} \ell_1^{O(\log^2 n)}$$

Using a more careful analysis, we can show that the above embedding has distortion $O(\log^2 n)$ with high probability and a distortion $O(\log^{1.5} n)$ in the ℓ_2-norm (see last lecture). As an aside, note that Theorem 6.1 is meaningful only for $D = O(\log n)$. For larger values of D, the distortion as well as the number of dimensions is larger that those corresponding to $\log n$.

Now, we prove Bourgain's Theorem, which refines the embedding and proof of Theorem 6.1, and obtains an embedding into $\ell_1^{O(\log^2 n)}$ with distortion only $O(\log n)$.

Theorem 6.2 (Bourgain (1985), Linial, London, Rabinovich (1995)) *For any metric space (X, d), and for any p,*

$$(X, d) \overset{O(\log n)}{\hookrightarrow} \ell_p^{O(\log^2 n)}$$

Theorem 6.2 uses the following algorithm to construct an embedding, which is very similar to that used for Theorem 6.1.

Algorithm:

$m := 576 \log n)$
for $j = 1$ **to** $\log n$ **do** /* levels of density */
  for $i = 1$ **to** m **do** /* repeat for high probability */
    choose set S_{ij} by sampling each node in X
    independently with probability 2^{-j}
  end
end
$f_{ij}(x) := d(x, S_{ij})$
$f(x) := \bigoplus_{j=1}^{\log n} \bigoplus_{i=1}^{m} f_{ij}(x)$

Let us first prove Theorem 6.2 for $p = 1$. Formally, let K denote the number of coordinates, i.e., $K = m \log n = 576 \log^2 n$. We will now prove that

$$(X, d) \xhookrightarrow{64 \log n} \ell_1^K .$$

Lemma 6.3 (Expansion) $\|f(x) - f(y)\|_1 \leq K \cdot d(x, y)$

Proof. For any set S and any pair of points $x, y \in X$, $|d(x, S) - d(y, S)| \leq d(x, y)$. Since the embedding has K dimensions, the lemma follows. ■

Theorem 6.4 (Contraction) *With probability* $\frac{1}{2}$*, for all* $x, y \in X$*,*

$$\|f(x) - f(y)\|_1 \geq 9 \log n \ d(x, y) .$$

Proving this theorem will form a major part of this lecture. We will first describe how it implies Theorem 6.2.

Proof of Theorem 6.2. By construction, the embedding uses $K \log^2 n$ dimensions. Theorem 6.4 shows that f has contraction $\frac{1}{9 \log n}$. On the other hand, the expansion of f is at most $K = 576 \log^2 n$ by Lemma 6.3. Multiplying the two, we get that the distortion of the embedding is at most $64 \log n$. ■

Now we begin with a proof of Theorem 6.4. Fix a pair of points $x, y \in X$. Let $r_j(x)$, for $j = 0, 1, \ldots, \log n$, be the smallest radius r such that a ball of radius r around x contains at least 2^j points, that is, $|B(x, r)| \geq 2^j$. Likewise, let $r_j(y)$ be the smallest radius r for which $|B(y, r)| \geq 2^j$. (Remember that the *ball* $B(x, r)$ is just the set $\{z \in X \mid d(x, z) \leq r\}$. Similarly, let the *open ball* $B^o(x, r)$ be defined as $\{z \in X \mid d(x, z) < r\}$. Let $\rho_j = \max\{r_j(x), r_j(y)\}$. By our construction, we have the following observation.

Observation 6.5 $|B(x, \rho_j)| \geq 2^j$ *and* $|B(y, \rho_j)| \geq 2^j$*.*

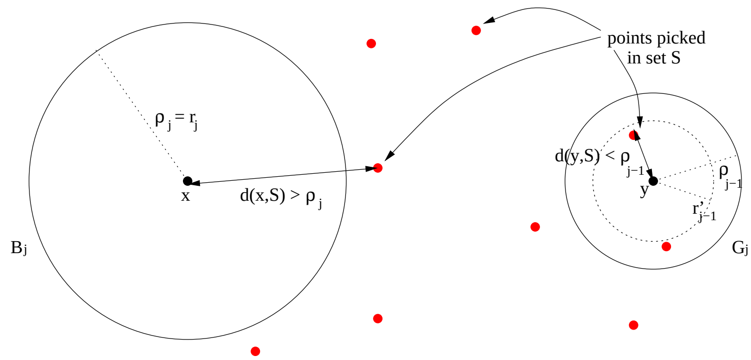

Figure 6.1: Balls B_j and G_j for a fixed pair of points x and y and iteration j

Let t be the smallest value such that $\rho_t + \rho_{t-1} > d(x,y)$; we then redefine $\rho_t = d(x,y) - \rho_{t-1}$. In the following argument, we will consider ρ_j's for values of j at most t.

Remark 6.6 *The value $\rho_0 = 0$, and the ρ_j's are monotonically increasing. Furthermore, since $\rho_{t-1} + \rho_t = d(x,y)$ and $\rho_{t-1} \le \rho_t$, it follows that $\rho_t \ge \frac{1}{2} d(x,y)$.*

For all j, we define a pair of balls B_j ("bad") and G_j ("good") with the following property. One of these balls is centered around x and the other is centered around y. Moreover, at stage j in the construction of the embedding, there is a reasonable probability that we pick at least one point from G_j and none from B_j. This would give us the necessary separation between x and y in the dimensions corresponding to j. (See Figure 6.1). We will formalize this in the rest of the proof.

If $\rho_j = r_j(x)$, let $B_j = B^o(x, \rho_j)$ and $G_j = B(y, \rho_{j-1})$. Otherwise, let $B_j = B^o(y, \rho_j)$ and $G_j = B(x, \rho_{j-1})$. This means we center the bad ball around the node from $\{x, y\}$ that has less nodes in its ρ_j-neighborhood. Note that open ball B_j contains *at most* 2^j points (by definition of $r_j(x)$ and $r_j(y)$), and G_j contains *at least* 2^{j-1} points (by Lemma 6.5). Furthermore the balls B_j and G_j do not intersect.

Observation 6.7 *For all j, $|B_j| \le 2^j$ and $|G_j| \ge 2^{j-1}$.*

Thus both the balls have size about 2^j. So if we are sampling at rate 2^{-j}, there is a reasonable chance that we will hit the "good" ball and miss the "bad" ball.

Claim 6.8 *For any value of j,*

$$\mathbf{Pr}[\text{for at least } 18 \log n \text{ values of } i, |f_{ij}(x) - f_{ij}(y)| \ge (\rho_j - \rho_{j-1})] \ge 1 - \frac{1}{n^3} \enspace .$$

Before we prove the claim, let us prove Theorem 6.4 using this claim.

Proof of Theorem 6.4. Claim 6.8 implies that, for a fixed j, we have

$$\mathbf{Pr}\left[\sum_{i=1}^{m} |f_{ij}(x) - f_{ij}(y)| \geq 18 \log n \cdot (\rho_j - \rho_{j-1})\right] \geq 1 - \frac{1}{n^3} \ .$$

In other words, for a fixed j, the probability of not getting "enough" contribution from the m coordinates corresponding to that j is bounded above by $\frac{1}{n^3}$. Since there are at most $\log n$ values of j, the trivial union bound implies that the probability that *any one* of these values of j did not contribute $18 \log n\,(\rho_j - \rho_{j-1})$ to the distance between x and y is $\frac{\log n}{n^3} \leq 1/n^2$. Hence

$$\mathbf{Pr}\left[\sum_{j,i} |f_{ij}(x) - f_{ij}(y)| \geq (18 \log n) \sum_{j=1}^{t} (\rho_j - \rho_{j-1})\right] \geq 1 - \frac{1}{n^2} \ .$$

However, note that $\sum_{j,i} |f_{ij}(x) - f_{ij}(y)| = \|f(x) - f(y)\|_1$; on the right hand side, the sum $\sum_{j=1}^{t} (\rho_j - \rho_{j-1})$ telescopes to ρ_t, which is at least $\frac{1}{2}\, d(x, y)$. Hence, for any fixed $x, y \in X$,

$$\mathbf{Pr}\left[\|f(x) - f(y)\|_1 \geq 18 \log n \frac{d(x,y)}{2}\right] \geq 1 - \frac{1}{n^2} \ .$$

Since there are $\binom{n}{2}$ pairs $x, y \in X$ that we need to argue about, the trivial union bound again implies

$$\mathbf{Pr}\left[\forall x, y \in X\colon \frac{\|f(x) - f(y)\|_1}{d(x,y)} \geq 9 \log n\right] \geq \frac{1}{2} \ ,$$

which proves the theorem. ■

Proof of Claim 6.8. Consider the coordinate corresponding to i and j. Recall that we form the set S_{ij} by sampling vertices independently at rate 2^{-j}. Let us consider the case when B_j is centered around x, and G_j around y, as in Figure 6.1. (The other case is proved identically.)

We want to compute the probability of the event that we get a good contribution, i.e., of the event

$$\begin{aligned} \mathcal{E} &= \{|f_{ij}(x) - f_{ij}(y)| \geq (\rho_j - \rho_{j-1})\} \\ &= \{|d(S_{ij}, x) - d(S_{ij}, y)| \geq (\rho_j - \rho_{j-1})\} \ . \end{aligned}$$

It will be tricky to calculate the probability of this event directly, so let us define another event $\mathcal{E}'$ thus

$$\begin{aligned} \mathcal{E}' &= \{d(S_{ij}, x) \geq \rho_j \quad \wedge \quad d(S_{ij}, y) \leq \rho_{j-1}\} \\ &= \{\ S_{ij} \text{ misses "bad" ball } B_j \wedge\ S_{ij} \text{ hits "good" ball } G_j\ \} \ . \end{aligned}$$

(Make sure you believe this!) It can be checked that the event $\mathcal{E}'$ implies $\mathcal{E}$, and hence $\mathbf{Pr}[\mathcal{E}] \geq \mathbf{Pr}[\mathcal{E}']$; hence it suffices to lower bound the probability of $\mathcal{E}'$. Finally, let us define

$$\begin{aligned} \mathcal{E}_{\text{hit}} &:= S_{ij} \cap G_j \neq \emptyset \\ \mathcal{E}_{\text{miss}} &:= S_{ij} \cap B_j = \emptyset \end{aligned}$$

and hence $\mathcal{E}' = \mathcal{E}_{\text{hit}} \wedge \mathcal{E}_{\text{miss}}$. Note that $\mathcal{E}_{\text{miss}}$ and $\mathcal{E}_{\text{hit}}$ are independent, since the balls G_j and B_j are disjoint.

$$\begin{aligned}\mathbf{Pr}[\mathcal{E}_{\text{hit}}] &= 1 - \mathbf{Pr}[S_{ij} \cap G_j = \emptyset] \\ &= 1 - (1 - 2^{-j})^{|G_j|} \\ &\geq 1 - (1 - 2^{-j})^{2^{j-1}} \\ &\geq 1 - e^{-1/2} \geq \frac{1}{4} \ .\end{aligned}$$

The third step in the above equations follows from the fact that $|G_j| \geq 2^{j-1}$. Likewise, for B_j we have,

$$\begin{aligned}\mathbf{Pr}[\mathcal{E}_{\text{miss}}] &= (1 - 2^{-j})^{|B_j|} \\ &\geq (1 - 2^{-j})^{2^j} \geq \frac{1}{4} \ .\end{aligned}$$

And finally,

$$\mathbf{Pr}\left[\mathcal{E}'\right] = \mathbf{Pr}[\mathcal{E}_{\text{hit}}] \cdot \mathbf{Pr}[\mathcal{E}_{\text{miss}}] \geq \frac{1}{16} \ .$$

We now use the following fundamental large-deviations bound, which says that the sum of many independent "well-behaved" random variables is closely concentrated around its expectation. (See, e.g., Alon and Spencer (1992) for a proof.)

Lemma 6.9 (Chernoff (1952), Hoeffding (1963)) *Let $X_1, \ldots, X_t$ denote binary random variables, with $E[X_i] \geq \mu$ for all i. Let $X = \sum_{i=1}^{t} X_i$. Then, $E[X] \geq \mu t$. The probability that X deviates significantly from its expectation is bounded as follows:*

$$\mathbf{Pr}[X \leq (1-\epsilon)\mu t] \ \leq \ e^{-\epsilon^2 \mu t/3} \ .$$

To apply the Chernoff bound, let X_i be the indicator variable that the event $\mathcal{E}'$ happens for S_{ij}. Thus we have $\mathbf{E}[X_i] = \mathbf{Pr}[X_i = 1] \geq \frac{1}{16}$. Hence, by Chernoff bound,

$$\mathbf{Pr}\left[\sum_{i=1}^{m} X_i \geq 18 \log n\right] \geq 1 - \frac{1}{n^3} \ .$$

Thus we have shown that the event $\mathcal{E}'$, and hence $\mathcal{E}$ for at least $18 \log n$ values of i, and hence

$$\mathbf{Pr}[|f_{ij}(x) - f_{ij}(y)| \geq \rho_j - \rho_{j-1} \text{ for at least } 18 \log n \text{ values of } i] \geq 1 - \frac{1}{n^3} \ ,$$

proving the claim. ∎

6.1 Embeddings into ℓ_p

In this section, we will show that the same mapping f that we defined for embedding into ℓ_1 works for all $p \geq 1$. It is a fairly surprising result, since we produce a mapping of (X, d) into $\mathbb{R}^{O(\log^2 n)}$ which has a low distortion for all norms ℓ_p *simultaneously.*

Theorem 6.10 *There exists an embedding that maps any arbitrary finite metric space (X, d) into $\ell_p^{O(\log^2 n)}$ with $O(\log n)$ distortion.*

$$(X, d) \xhookrightarrow{O(\log n)} \ell_p^{O(\log^2 n)} \quad .$$

In particular, with probability at least $\frac{1}{2}$, the mapping f defined above has this property.

In order to prove this theorem, we will use the Cauchy-Schwarz Inequality.

Fact 6.11 (Cauchy-Schwarz) *For two vectors $\vec{a}, \vec{b} \in \mathbb{R}^k$, it holds that $\|\vec{a}\|_2 \cdot \|\vec{b}\|_2 \geq \langle \vec{a}, \vec{b} \rangle$.*

Proof Theorem 6.10. We will first prove the theorem for $p = 2$. We give an upper bound on $\|f(x) - f(y)\|_2$. Recall that $K = 576 \log^2 n$ is the number of dimensions of the map f.

$$\begin{aligned}
\|f(x) - f(y)\|_2 &= (\sum_{i,j} |f_{ij}(x) - f_{ij}(y)|^2)^{\frac{1}{2}} \\
&\leq (\sum_{i,j} d(x, y)^2)^{\frac{1}{2}} \\
&= \sqrt{K} \cdot d(x, y) \\
&= O(\log n) \cdot d(x, y)
\end{aligned}$$

To give a lower bound on the distance between x and y in the embedding, we use the Cauchy-Schwarz inequality from Fact 6.11. Set $a_{ij} = |f_{ij}(x) - f_{ij}(y)|$ and $b_{ij} = 1$. Plugging these values in the inequality gives

$$(\textstyle\sum_{i,j} |f_{i,j}(x) - f_{i,j}(y)|^2)^{\frac{1}{2}} \cdot \sqrt{K} \geq \textstyle\sum_{i,j} |f_{ij}(x) - f_{ij}(y)|$$

which means that

$$\begin{aligned}
\sqrt{K} \cdot \|f(x) - f(y)\|_2 &\geq \|f(x) - f(y)\|_1 \\
&\geq O(\log n) \cdot d(x, y) \quad ,
\end{aligned}$$

where the last inequality following from Theorem 6.2. Now using that $\sqrt{K} = \Theta(\log n)$, we get that

$$\|f(x) - f(y)\|_2 \geq \frac{d(x, y)}{O(1)} \quad ,$$

which proves the theorem. ∎

This result can be easily generalized to ℓ_p using the Hölder's inequality (see Homework).

Fact 6.12 (Hölder's Inequality) *For two vectors $\vec{a}, \vec{b} \in \mathbb{R}^k$, and $p, q \in \mathbb{R}^+$ such that $\frac{1}{p} + \frac{1}{q} = 1$, it holds that $\|\vec{a}\|_p \cdot \|\vec{b}\|_q \geq \langle \vec{a}, \vec{b} \rangle$.*

Remark 6.13 *As a sanity check, notice that if $p = q = 2$, then this gives us the Cauchy-Schwarz inequality from Fact 6.11. Furthermore, if $p = 1$ and $q = \infty$, then the inequality implies that $\|\vec{a}\|_1 \cdot b_{\max} \geq \langle \vec{a}, \vec{b} \rangle$, which can also be easily checked.*

A stronger version of Theorem 6.10 proved by Matoušek (1997) gives the following result:

$$(X, d) \xhookrightarrow{O(\log n/p)} \ell_p \ .$$

7 Lower Bounds for Embedding finite Metrics into ℓ_1

In this section we show that the result of Bourgain is tight. We show that for arbitrarily large n there exist metrics on n points such that the minimum distortion for embedding the metric into ℓ_1 is $\Omega(\log n)$. Recall the IP from the last class

$$\begin{aligned} \text{minimize} \quad & \sum_{x,y} \text{cap}(x,y) \cdot d(x,y) \\ \text{subject to} \quad & \sum_{x,y} \text{dem}(x,y) \cdot d(x,y) = 1 \\ & d \text{ is elementary cut metric} \end{aligned} \tag{IP}$$

We looked at the relaxation of this IP where we require d only to be a metric (instead of being a cut-metric). This gives an LP. Let d^* denote the metric that obtains the optimal value for this LP. In the last lecture we have shown that if d^* embeds into ℓ_1 with distortion D, then we have

$$LP^* \leq IP^* \leq D \cdot LP^* \ ,$$

where LP^* and IP^* denote the optimal value for the LP and IP, respectively. From the following theorem we can derive a lower bound on the distortion needed for embedding finite metrics into ℓ_1 .

Theorem 7.14 (Leighton-Rao (1988)) *For infinitely many values of n, there exist instances of the sparsest cut problem on graphs G_n with n vertices such that*

$$IP^*/LP^* \geq \Omega(\log n) \ .$$

Corollary 7.15 *Let the metric $d^*_{G_n}$ be generated by (LP) on the instances in Theorem 7.14. Then*

$$(V_n, d^*_{G_n}) \xhookrightarrow{\Omega(\log n)} \ell_1 \ .$$

Proof. Suppose not, and let $(V_n, d^*_{G_n}) \xhookrightarrow{o(\log n)} \ell_1$. Then by our result from the previous lecture, $IP^* \leq o(\log n) \cdot LP^*$ on these instances, which contradicts Theorem 7.14. ■

Hence, we have shown that there are metrics that require $\Omega(\log n)$ distortion to embed into ℓ_1, showing that Bourgain's embedding is existentially tight. However we still have to prove Theorem 7.14.

The lower bound is achieved by a family of expanders. Let us recall the following definition:

Definition 7.16 *The* expansion *of set $S \subseteq V$ is given by:*

$$\Phi(S) := \frac{\text{cap}(S, V \setminus S)}{\min\{|S|, |V \setminus S|\}}$$

The expansion $\Phi(G)$ *of* G *is defined to be* $\min_S \Phi(S)$. *A graph* G *is an* expander *if the expansion is a constant independent of* n*; i.e.,* $\Phi(G) \geq \Omega(1)$.

Constant-degree expanders are known to exist: this can be shown via a probabilistic argument. Explicit constructions are more difficult to come by, but these are also known:

Theorem 7.17 (Lubotzky, Phillips, and Sarnak (1988)) *For infinitely many n, there exist 3-regular expanders G_n.*

The expansion is closely related to the sparsity of the graph; note that if we define demands $\text{dem}(x, y) = 1$ for all pairs of vertices $x \neq y \in G$, we get:

$$\frac{n}{2} \cdot \text{sparsity}(G) \leq \Phi(G) \leq n \cdot \text{sparsity}(G) \tag{7.3}$$

Proof Theorem 7.14. Let $G_n = (V_n, E_n)$ be a 3-regular expander, and fix any vertex $v \in V_n$. We claim that the number of vertices at distance at least $\frac{1}{2}\log_2 n$ from v is at least $\frac{1}{2}n$.

Indeed, in any Δ-regular graph, the number of vertices at distance no more than $t-1$ is at most $1 + \Delta + (\Delta-1)^2 + \ldots + (\Delta-1)^{t-1}$, which is

$$1 + \frac{(\Delta-1)^t - 1}{\Delta - 1} \quad .$$

Now plugging in $\Delta = 3$ and $t = \frac{1}{2}\log n$, we get that the number of vertices *farther* than $\frac{1}{2}\log_2 n$ from v is at least $n - \sqrt{n} \geq n/2$.

Lemma 7.18 *For the 3-regular expanders of Theorem 7.17,*

$$LP^* \leq O(\frac{1}{n \log n}) \quad .$$

Proof. For each $v \in V_n$, look at all $\frac{1}{2}n$ or more vertices that are at distance at least $\frac{1}{2}\log_2 n$ from v. This gives at least $\frac{1}{2}n \cdot n$ demand pairs, such that sending λ units of flow between any one of these pairs uses up at least $\lambda \cdot \frac{\log_2 n}{2}$ units of volume. Since the total capacity available in the 3-regular graph is only $\frac{3}{2}n$, we must have that

$$\frac{1}{2}n \cdot n \cdot \frac{\lambda \log n}{2} \leq \frac{3}{2}n \ ,$$

and hence $\lambda \leq \frac{6}{n \log n}$. ■

However, since G_n is an expander, $\Phi(G_n) = \Omega(1)$, and hence Equation (7.3) implies that sparsity$(G_n) = IP^* \geq \Theta(\frac{1}{n})$. Therefore, the integrality gap $IP^*/LP^* \geq \Omega(\log n)$, which completes the proof of Theorem 7.14. ■

References

[AR98] Yonatan Aumann and Yuval Rabani. An $O(\log k)$ approximate min-cut max-flow theorem and approximation algorithm. *SIAM Journal on Computing*, 27(1):291–301, 1998.

[BK97] Avrim Blum and David Karger. An $O(n^{3/4})$-coloring algorithm for 3-colorable graphs. *Information Processing Letters*, 61(1):49–53, 1997.

[BL84] Sandeep N. Bhatt and Frank Thomson Leighton. A framework for solving VLSI graph layout problems. *Journal of Computer and System Sciences*, 28(2):300–343, 1984.

[Bou85] Jean Bourgain. On Lipschitz embeddings of finite metric spaces in Hilbert space. *Israel Journal of Mathematics*, 52(1-2):46–52, 1985.

[Che95] Herman Chernoff. A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Annals of Mathematical Statistics*, 23(4):493–507, 1995.

[GLS88] Martin Grötschel, László Lovász, and Alexander Schrijver. *Geometric algorithms and combinatorial optimization*. Springer, New York, 1988.

[HIL98] Johan Håstad, Lars Ivansson, and Jens Lagergren. Fitting points on the real line and its application to RH mapping. In *Proceedings of the 6th European Symposium on Algorithms (ESA)*, pages 465–476, 1998.

[Hoe63] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963.

[LLR95] Nathan Linial, Eran London, and Yuri Rabinovich. The geometry of graphs and some of its algorithmic applications. *Combinatorica*, 15(2):215–245, 1995. Also in *Proc. 35th FOCS*, 1994, pp. 577–591.

[Mat96] Jiří Matoušek. On the distortion required for embedding finite metric spaces into normed spaces. *Israel Journal of Mathematics*, 93:333–344, 1996.

[Mat97] Jiří Matoušek. On embedding expander graphs into ℓ_p spaces. *Israel Journal of Mathematics*, 102:189–197, 1997.