ISRAEL JOURNAL OF MATHEMATICS 93 (1996), 333–344

ON THE DISTORTION REQUIRED FOR EMBEDDING FINITE METRIC SPACES INTO NORMED SPACES*

BY

JIŘÍ MATOUŠEK

Department of Applied Mathematics, Charles University
Malostranské nám. 25, 118 00 Praha 1, Czech Republic
e-mail: matousek@kam.mff.cuni.cz

ABSTRACT

We investigate the minimum dimension k such that any n-point metric space M can be D-embedded into some k-dimensional normed space X (possibly depending on M), that is, there exists a mapping $f\colon M \to X$ with

$$\frac{1}{D}\mathrm{dist}_M(x,y) \le \| f(x) - f(y)\| \le \mathrm{dist}_M(x,y) \quad \text{for any } x, y \in M.$$

Extending a technique of Arias-de-Reyna and Rodríguez-Piazza, we prove that, for any fixed $D \ge 1$, $k \ge c(D)\, n^{1/2D}$ for some $c(D) > 0$. For a D-embedding of all n-point metric spaces into the same k-dimensional normed space X we find an upper bound $k \le 12D\, n^{1/\lfloor (D+1)/2 \rfloor} \ln n$ (using the ℓ_∞^k space for X), and a lower bound showing that the exponent of n cannot be decreased at least for $D \in [1,7) \cup [9,11)$, thus the exponent is in fact a jumping function of the (continuously varied) parameter D.

* The research was supported by Charles University grant No. 351, by Czech Republic Grant GAČR 201/93/2167 and by EC Cooperative Action IC-1000 (project ALTEC: *Algorithms for Future Technologies*).
Received September 1, 1994

1. Introduction

Let M be a metric space with metric ρ, let X be a normed space (whose norm will be denoted by $|\cdot|$), and let $f\colon M \to X$ be a mapping. We say that f is a D-**embedding** (or a mapping with **distortion** at most D), $D \geq 0$ a real number, if we have

$$\frac{1}{D}\,\rho(x,y) \leq |\, f(x) - f(y)| \leq \rho(x,y)$$

for any two points $x, y \in M$. We say that M D-**embeds** into X if there exists a D-embedding* $f\colon M \to X$.

The D-embeddability of finite metric spaces into various normed spaces has been studied in several papers. This investigation started in the context of the local Banach space theory, where the general idea was to obtain some analogs for general metric spaces of notions and results dealing with the structure of finite dimensional subspaces of Banach spaces. The distortion of a mapping should play the role of the norm of a linear operator, and the quantity $\log n$, where n is the number of points in a metric space, would serve as an analog of the dimension of a normed space. Parts of this programme have been carried out by Bourgain, Johnson, Lindenstrauss, Milman and others; see, e.g., [JL84], [Bou85], [BMW86], [JLS87], [Bou86].

It appears that D-embeddings can be of considerable interest also in more applied areas. They can serve as a useful representation of graphs and other metric spaces helping to visualize their structure, find clusters, small separators etc.; see Linial *et al.* [LLR 94].

To formulate some important questions in this area and the results, we introduce three functions. We define $\psi_D(n)$ as the minimum k such that for any n-point metric space M there exists a k-dimensional normed space and a D-embedding of M into it. We let $\psi_D^F(n)$ be the minimum k such that there exists a k-dimensional normed space into which any n-point metric space D-embeds, and finally $\psi_D^\infty(n)$ is the smallest k such that any n-point metric space D-embeds into ℓ_∞^k (the vector space $\mathbb{R}k$ with the ℓ_∞-norm). Clearly $\psi_D(n) \leq \psi_D^F(n) \leq \psi_D^\infty(n)$.

---

* We remark that the notion of a D-embedding can be introduced for two general metric spaces (instead of requiring X to be a normed space), with a formally somewhat more complicated definition. Also, a number of various terms besides the mentioned ones are used in the literature in this context; e.g., a D-embedding is also called a D-**isomorphism**, a D-**lipeomorphism**, or the metric spaces (M,ρ) and $f(M)$ with the metric induced by X are said to have **Lipschitz distance** at most D, etc.

Since any n-point metric space can be isometrically embedded into ℓ_∞^{n-1}, we have $\psi_1^\infty(n) \le n-1$. On the other hand, using a volume argument it is easily seen that for any fixed D, $\psi_D(n) \ge c(D)\log n$.

Johnson and Lindenstrauss [JL84] asked whether this lower bound described the actual asymptotic behavior of $\psi_D(n)$. Bourgain [Bou85] disproved this by showing $\psi_D(n) \ge \frac{c}{D^2}(\log n/\log\log n)^2$. In his remarkable paper, he developed a technique for constructing D-embeddings of finite metric spaces into normed spaces, and also a method for proving lower bounds on the distortion and/or dimension required for embedding metric spaces into a fixed normed space (the basic approach in this method, which is the use of random graphs, was originally suggested by Lindenstrauss). He used these techniques for estimating the distortion required to embed any n-point metric space into the n-dimensional Euclidean space. His embedding technique combined with other methods has been used by Johnson *et al.* [JLS87] to prove that for any D,

$$\psi_D(n) \le C(D)\, n^{A/D} \tag{1}$$

(with an absolute constant A). In [Ma91] the author used Bourgain's technique in a much simpler way (and directly for embedding into ℓ_∞^k), showing that

$$\psi_D^\infty(n) \le C\, n^{3/D}\log^2 n$$

(this also improved the value of A in (1)).

Essential progress was achieved by Arias-de-Reyna and Rodríguez-Piazza [AR92], who proved that for any $D < 2$ there exists a constant $c(D) > 0$ such that $\psi_D(n) \ge c(D)n$ (in other words, almost the full dimension n is needed for embeddings with distortion below 2). At the same time they ask whether $\psi_D(n) \le C(D)\log^{\text{const}} n$ is true for distortions $D > 2$. Here we prove that for a fixed D the function $\psi_D(n)$ grows at least as a power of n:

THEOREM 1.1: *For a fixed $D \ge 1$,*

$$\psi_D(n) \ge \begin{cases} c(\lfloor D\rfloor)\, n^{\frac{1}{\lceil D\rceil}} & \text{for } D \in [2,4)\cup[5,6), \\ c(\lfloor D\rfloor)\, n^{\frac{1}{2\lceil D\rceil}} & \text{otherwise,} \end{cases}$$

where $c(\lfloor D\rfloor) > 0$ depends on the integer part of D only.

In particular, for $D < 2$ this improves the result of Arias-de-Reyna and Rodríguez-Piazza, since it shows that for any such D the required dimension is linear in n with an absolute positive constant of proportionality (while in their result the constant tends to 0 with $D \to 2$).

The asymptotics of the functions $\psi_D^\infty(n)$ and $\psi_D^F(n)$ is described by the following result, which also gives the best upper bound for $\psi_D(n)$ known so far.

THEOREM 1.2:

(i) *For any $D \geq 3$ and $n \geq 3$,*

$$\psi_D^\infty(n) \leq 12D\, n^{\frac{1}{\lfloor (D+1)/2 \rfloor}} \ln n\,.$$

(ii)

$$\psi_D^F(n) \geq \begin{cases} c(D)\, n^{\frac{1}{\lfloor (D+1)/2 \rfloor}} & \text{for } D \in [1,7) \cup [9,11), \\ c(D)\, n^{\frac{1}{2\lfloor (D+1)/2 \rfloor}} & \text{otherwise}\ , \end{cases}$$

where $c(D) > 0$ depends on D only.

For the upper bound (embedding) we apply the technique of Bourgain [Bou85] almost in the same way as in [Ma91], and for the lower bound we use Bourgain's counting argument similarly as Johnson *et al.* [JLS87]; we only note that one can use explicit graphs without short cycles as a basis instead of probabilistically constructed ones implicitly applied in the above-mentioned papers. Thus, we mostly repeat previous work. The reason why we include this result nevertheless is that one obtains the surprisingly close upper and lower bounds by looking at the calculations carefully, and it turns out that the actual dependence of the dimension on the required distortion has the strange jumping function $1/\lfloor (d+1)/2 \rfloor$ in the exponent of n. For instance, the required dimension remains nearly linear for distortions below 3, then it jumps suddenly to the order of $\sqrt{n}$ at distortion 3, etc.

In the present proof of Theorem 1.2(ii) the constant $c(D)$ tends to 0 every time D approaches an odd integer (where the value of the exponent jumps) from below. We suspect that a uniform bound for the constant should exist; Proposition 3.1 in Section 3 shows at least that $\psi_D^\infty(n) \geq n/4$ for all $D < 3$.

The lower bounds in Theorems 1.1 and 1.2 become less precise for larger D; this is because of a lack of knowledge about graphs without short cycles. Recall that the **girth** of a graph G is the length of the shortest cycle in G. The quantity relevant for our results is the maximum number of edges a graph of girth g

on n vertices can have. A well known upper bound is $O(n^{1+1/\lfloor(g-1)/2\rfloor})$, and this is conjectured as the correct order of magnitude; see e.g., [Bol78]*. The upper bound is tight for $g = 3, 4$ (using the complete bipartite graph), $g = 5, 6$ (using the incidence graph of a finite projective plane) and $g = 7, 8, 11, 12$ (by algebraic constructions, see [Ben66]). For larger g, there is an old lower bound of $cn^{1+1/2\lfloor(g-1)/2\rfloor}$. An asymptotic improvement follows from a construction of Lubotzky *et al.* [LPS88] which provides a lower bound of roughly $n^{1+4/3g}$ for infinitely many values of g (this gives corresponding improved bounds in the situations of Theorem 1.1 and 1.2(ii); we haven't included these in the theorems, as the exact statements become somewhat complicated).

Now that the order of magnitude of $\psi_D^F(n)$ has been determined reasonably precisely for small D, it would be interesting to find out whether $\psi_D(n)$ is essentially smaller or not, that is, whether the freedom to choose a normed space for a given metric space helps (it definitely does help for some metric spaces, e.g., for the spaces used as lower bound examples in Theorem 1.2). In particular, can every n-point metric space be 2.99-embedded, say, into a normed space of a sublinear dimension?

2. Lower bound for arbitrary normed spaces

In this section we prove Theorem 1.1. Let the distortion D be given, and let g be the nearest even integer strictly larger than $2D$. We start with a $2n$ vertex graph G with girth g and with possibly many edges, according to the lower bounds mentioned in the introduction. By removing at most half of the edges, we may assume that G is bipartite with balanced classes, that is, its edges join vertices of $U = \{u_1, u_2, \ldots, u_n\}$ to vertices of $V = \{v_1, v_2, \ldots, v_n\}$.

For every edge $e = \{u_i, v_j\} \in E(G)$ we choose a suitable sign $\varepsilon_{ij} \in \{\pm 1\}$ (Lemma 2.2 below deals with the choice of these signs). Then we define a new graph G' with vertex set $U^+ \cup U^- \cup V$, where $U^+ = \{u_1^+, \ldots, u_n^+\}$ and $U^- = \{u_1^-, \ldots, u_n^-\}$. For every edge $\{u_i, v_j\} \in E(G)$ we put one edge into $E(G')$; this edge connects v_j to either u_i^+ or u_i^- depending on the sign ε_{ij}. The metric space M requiring a large dimension for a D-embedding is the set of vertices of G' (that is, a $3n$ point space) with the usual graph-theoretic distance (denoted by ρ).

* One can always assume that G is bipartite, so that it has no odd length cycles. This is why the exponent is the same for an odd g and the next even g.

Important facts about this metric are that $\rho(u_i^+, v_j) = 1$ if $\{u_i, v_j\} \in E(G)$ and $\varepsilon_{ij} = 1$, $\rho(u_i^-, v_j) = 1$ if $\{u_i, v_j\} \in E(G)$ and $\varepsilon_{ij} = -1$, and $\rho(u_i^+, u_i^-) \geq g$ for all i. To see the last inequality we note that by identifying U^+ and U^- back to U we never duplicate edges, and thus a path of length ℓ connecting u_i^+ and u_i^- in G' induces a cycle of length $\leq \ell$ in G.

Suppose that $f\colon M \to X$ is a D-embedding into a normed space. For a convenient notation, we denote $f(w)$ by $\tilde{w}$ for any vertex w of G'. As in [AR92], we choose norm 1 linear functionals $h_1, \ldots, h_n \in X^*$ (thus $|h_i(x) - h_i(y)| \leq \|x - y\|$ for all $x, y \in X$) such that $h_i(\tilde{u}_i^+) - h_i(\tilde{u}_i^-) = \|\tilde{u}_i^+ - \tilde{u}_i^-\|$. We form an $n \times n$ matrix $A_0 = (h_i(\tilde{v}_j))_{i,j=1}^n$. It is easy to see that the dimension of X is at least the rank of A_0. Let B be an auxiliary matrix with $b_{ij} = h_i\left(\frac{\tilde{u}_i^+ + \tilde{u}_i^-}{2}\right)$, and set $A = A_0 - B$. The matrix B has rank 1, and thus $\dim(X) \geq \operatorname{rank}(A_0) \geq \operatorname{rank}(A) - \operatorname{rank}(B) \geq \operatorname{rank}(A) - 1$.

To estimate $\operatorname{rank}(A)$, we derive inequalities for its entries corresponding to edges of G.

LEMMA 2.1: *For any pair (i, j) such that $\{u_i, v_j\} \in E(G)$, $\operatorname{sgn}(a_{ij}) = \varepsilon_{ij}$.*

Proof: Since f is a D-embedding, we have $h_i(\tilde{u}_i^+ - \tilde{u}_i^-) = \|\tilde{u}_i^+ - \tilde{u}_i^-\| \geq \rho(u_i^+, u_i^-)/D \geq g/D$. Thus, for the case $\varepsilon_{ij} = 1$ we get

$$a_{ij} = h_i\left(\tilde{v}_j - \frac{\tilde{u}_i^+ + \tilde{u}_i^-}{2}\right) = h_i(\tilde{v}_j - \tilde{u}_i^+) + h_i\left(\frac{\tilde{u}_i^+ - \tilde{u}_i^-}{2}\right) \geq -1 + \frac{g}{2D} > 0\,.$$

The case $\varepsilon_{ij} = -1$ is symmetric. ∎

Now, in order to finish the proof of Theorem 1.1, it suffices to establish the following lemma (the quantitative bounds in Theorem 1.1 follow from the quantitative results about graphs of given girth mentioned in the introduction).

LEMMA 2.2: *There exists a choice of signs $\varepsilon_{ij} \in \{\pm 1\}$, for $\{u_i, v_j\} \in E(G)$, such that any matrix A whose entries satisfy the claim of Lemma 2.1 has rank at least $m/21n$, where $m = |E(G)|$, provided that m is sufficiently large (larger than a suitable constant).*

The proof resembles a proof of a result of a similar type of Alon *et al.* [AFR85], which in turn is based on theorems of Milnor [Mil64] and Thom [Tho65]. We need the following lemma of [AFR85]:

LEMMA 2.3: *Let $p_1, \dots, p_m \in \mathbb{R}[x_1, \dots, x_\ell]$ be polynomials with real coefficients in ℓ variables. Put $J = \{1, 2, \dots, m\}$ and let $J = J_1 \dot\cup J_2 \dot\cup \cdots \dot\cup J_h$ be a partition of J into h pairwise disjoint parts. Define*

$$k = 2 \max_{1 \le i \le h} \left(\sum_{j \in J_i} \deg(p_j) \right).$$

Then the number of sequences $(\sigma_1, \sigma_2, \dots, \sigma_m) \in \{\pm 1\}^m$ such that there exists $x = (x_1, \dots, x_\ell) \in \mathbb{R}^\ell$ with $\operatorname{sgn}(p_j(x)) = \sigma_j$ for all $j = 1, 2, \dots, m$ is at most

$$k(2k-1)^{\ell+h-1}.$$

Proof of Lemma 2.2: It is well known (and easy to check) that any $n \times n$ matrix A of rank at most q can be written as a product UV^T, where U and V are $n \times q$ matrices. We let the entries of U and V be variables, so that each entry a_{ij} of A is a quadratic polynomial in these variables. Then the existence of a matrix A of rank q satisfying the conditions of Lemma 2.2 means that the m polynomials corresponding to entries a_{ij} with $\{u_i, v_j\} \in G$ attain the signs ε_{ij}. There are 2^m possible choices of the ε_{ij}. We divide the m considered polynomials into $h = \lceil m/50 \rceil$ groups by at most 50 and apply Lemma 2.3 with $\ell = 2nq$; we get that the total number of sign patterns attained by these polynomials is at most $100 \times 400^{2nq+m/50}$. For $q \le m/21n$ calculation shows that this number is smaller than 2^m, thus there is a choice of the ε_{ij} corresponding to no rank $\le q$ matrix. ∎

3. Lower bound for a fixed normed space

Here we prove Theorem 1.2(ii). Let $D \ge 1$ be given, let g be the first even integer (strictly) larger than $D+1$, and fix a graph G on vertex set $V = \{v_1, v_2, \dots, v_n\}$ with m edges (as many as possible). Let $\mathcal{G}$ denote the class of all subgraphs of G. For each $H \in \mathcal{G}$ define a metric space M_H on the point set V: the distance of v_i and v_j will be the minimum of $g-1$ and of the graph theoretic distance of v_i and v_j in H.

Let X be a k-dimensional normed space, and suppose that for every $H \in \mathcal{G}$ there exists a D-embedding $f_H \colon M_H \to X$. We may assume that $f_H(v_1) = 0$ for every H, and thus the images of all points are contained in the $(g-1)$-ball B around the origin of X. Set $\beta = \frac{1}{4}[(g-1)/D - 1]$, and let N be a β-net in B

(recall that a β-net is an inclusion maximal subset of B whose any two points have distance at least β). For every $H \in \mathcal{G}$ define a new mapping $\bar{f}_H\colon M_H \to N$, by letting $\bar{f}_H(v_i)$ be the nearest point to $f_H(v_i)$ in N.

We claim that for distinct $H_1, H_2 \in \mathcal{G}$ the mappings $\bar{f}_{H_1}$ and $\bar{f}_{H_2}$ are distinct. The edge sets of H_1 and H_2 differ, so we can choose a pair u, v of vertices which form an edge in one of them, say in H_1, and not in the other one (H_2). The distance of u and v in M_{H_1} is 1, while their distance in M_{H_2} is $g-1$ (otherwise an u-v path in H_2 of length $< g-1$ and the edge $\{u,v\}$ would induce a cycle shorter than g in G). Then we get

$$|\bar{f}_{H_1}(u) - \bar{f}_{H_1}(v)| < |f_{H_1}(u) - f_{H_1}(v)| + 2\beta \le 1 + 2\beta$$

and

$$|\bar{f}_{H_2}(u) - \bar{f}_{H_2}(v)| > |f_{H_2}(u) - f_{H_2}(v)| - 2\beta \ge \frac{g-1}{D} - 2\beta = 1 + 2\beta\,,$$

therefore $\bar{f}_{H_1}(u) \ne \bar{f}_{H_2}(u)$ or $\bar{f}_{H_1}(v) \ne \bar{f}_{H_2}(v)$.

The number of all mappings of V into N is $|N|^n$, and this must be at least $|\mathcal{G}| = 2^m$. By a standard volume argument we find the estimate $|N| \le (2g/\beta)^k$, and calculation gives

$$k \ge \frac{1}{\log_2 \frac{2g}{\beta}} \, \frac{m}{n}\,.$$

Theorem 1.2(ii) is obtained by substituting the known estimates for m. ∎

Remark: For a fixed small D, the upper and lower bounds in Theorem 1.2 differ asymptotically by a logarithmic factor. It seems possible that it is the lower bound which can be improved, even using the current approach. For, say, girth 6 there is no graph G with more than $n^{3/2}$ edges, but it is not clear that there could not exist a class $\mathcal{G}$ having, say, $2^{cn^{3/2}\log n}$ graphs such that for any 2 graphs in $\mathcal{G}$ there is a pair of vertices which is an edge in one of them and has distance at least $g-1$ in the other one.

PROPOSITION 3.1: *There exists an n-point metric space which cannot be D-embedded into $\ell_\infty^{\lfloor n/4\rfloor - 1}$ for any $D < 3$.*

Proof: The metric space will be the complete bipartite graph minus a perfect matching. More formally, suppose that n is a multiple of 4, and for $i = 1, 2, \ldots, n/2$ let a point u_i have distance 1 to all points v_j, $1 \le j \le n$, $i \ne j$

and distance 3 to v_i. Let f be a D-embedding of this metric space to ℓ_∞^k. For every i there must be one coordinate in ℓ_∞^k where $f(u_i)$ and $f(v_i)$ differ by more than 1. We claim that no 3 indices i_1, i_2, i_3 can share the same coordinate. For contradiction, suppose that the indices 1, 2, 3 do, and let $g(w)$ denote the value of this offending coordinate of $f(w)$, $w \in \{u_1, u_2, u_3, v_1, v_2, v_3\}$. For each pair u_j, v_j there are two possibilities for the order relation between $g(u_j)$ and $g(v_j)$, so two indices, say 1 and 2, share the same ordering, say $g(u_1) < g(v_1)$ and $g(u_2) < g(v_2)$; moreover we may assume $g(u_1) \le g(u_2)$. Then we get $g(v_2) \le g(u_1) + 1$ and $g(v_2) > g(u_2) + 1 \ge g(u_1) + 1$, a contradiction. ∎

4. Embeddings into ℓ_∞^k

Here we prove Theorem 1.2(i). We follow [Bou85] and [Ma91]. Let $D = 2q+1 \ge 3$ be an odd integer, let M be an n-point metric space with metric ρ. We show that M can be D-embedded into ℓ_∞^k with $k = (q+1)\lceil 16n^{1/(q+1)} \ln n \rceil$; from this Theorem 1.2(i) follows.

First we describe a mapping $f\colon M \to \ell_\infty^k$. Set an auxiliary parameter $p = n^{-1/(q+1)}$, and for $j = 1, 2, \ldots, q+1$ define probabilities $p_j = \min(\frac{1}{2}, p^j)$. Further, let $m = \lceil 8n^{1/(q+1)} \ln n \rceil$. For $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, q+1$, we choose a random subset $A_{ij} \subseteq M$, by including every point of M into A_{ij} with probability p_j, the choices being mutually independent. We divide the coordinates in ℓ_∞^k into $q+1$ blocks by m coordinates. Then for every $x \in M$ we define the ith coordinate in the jth block for the point $f(x) \in \ell_\infty^k$ as the distance $\rho(x, A_{ij})$. We claim that with a positive probability, f is a D-embedding.

It is easy to see that f is always a contraction. We have the following lemma (due to Bourgain [Bou85], we only refine the formulas a little):

LEMMA 4.1: *Let x, y be two distinct points of M. Then there exists an index $j \in \{1, 2, \ldots, q+1\}$ such that if the set A_{ij} is chosen randomly as above, then the probability of the event*

$$(2) \qquad |\rho(A_{ij}, x) - \rho(A_{ij}, y)| \ge \frac{\rho(x,y)}{D}$$

is at least $p/8$.

First, assuming this lemma, we finish the proof of Theorem 1.2(i). To show that f is a D-embedding, it suffices to show that, with a nonzero probability, for every pair x, y there are i, j such that the event (2) in the above lemma occurs for the set A_{ij}. Consider a fixed pair x, y and select the appropriate index j as in the lemma. The probability that the event (2) does not occur for any of the m indices i is at most $(1-p/8)^m \leq e^{-pm/8} \leq n^{-2}$, and since there are $\binom{n}{2} < n^2$ pairs x, y, the probability that we fail to choose a good set for any of the pairs is smaller than 1. ∎

Proof of Lemma 4.1: Set $r = \rho(x,y)/D$. Let $B_0 = \{x\}$, let B_1 be the (closed) r-ball around y, let B_2 be the (closed) $2r$-ball around x, ..., finishing with B_{q+1} which is a $(q+1)r$-ball around x (if q is odd) or around y (if q is even). The parameters are chosen so that the radii of B_q and B_{q+1} add up to $\rho(x,y)$, i.e. the last 2 balls just touch. Let n_t denote number of points in B_t.

Divide the interval $[1, n]$ into $q+1$ intervals $I_1, I_2, \ldots, I_{q+1}$, where

$$I_j = \left[n^{(j-1)/(q+1)}, n^{j/(q+1)}\right].$$

If we have $n_0 \leq n_1 \leq \cdots \leq n_{q+1}$ then, by the pigeonhole principle, there exists an interval I_j containing some n_t and n_{t+1}, so we have

$$n_t \geq n^{(j-1)/(q+1)} \quad \text{and} \quad n_{t+1} \leq n^{j/(q+1)}. \tag{3}$$

On the other hand, if $n_t > n_{t-1}$ then pick for I_j the interval containing n_t and (3) holds as well.

In this way, we have selected the index j whose existence is claimed in the lemma. We will show that with probability at least $p/8$, the set A_{ij} (randomly selected with point probability p_j) includes a point of B_t (event E_1) and is disjoint from the interior of B_{t+1} (event E_2); such an A_{ij} then satisfies (2). Since B_t and the interior of B_{t+1} are disjoint, E_1 and E_2 are independent.

We calculate

$$\text{Prob}[E_1] = 1 - \text{Prob}\left[A_{ij} \cap B_t = \emptyset\right] = 1 - (1-p_j)^{n_t} \geq 1 - \exp(-p_j n_t) \geq 1 - e^{-p}$$

by (3). It is an elementary calculus exercise to verify that the last quantity is at least $p/2$ for all $p \in [0,1]$. Further

$$\text{Prob}[E_2] \geq (1-p_j)^{n_{t+1}} \geq (1-p_j)^{n^{j/(q+1)}} \geq (1-p_j)^{1/p_j} \geq 1/4$$

(since we always have $p_j \leq 1/2$). Thus $\text{Prob}[E_1 \cap E_2] \geq p/8$, which proves the lemma. ∎

ACKNOWLEDGEMENT: I would like to thank Prof. Joram Lindenstrauss for for a stimulating discussion and for giving me the key reference [AR92]. Thanks also to Prof. János Pach for providing references concerning graphs of given girth.

References

[AFR85] N. Alon, P. Frankl and V. Rödl, *Geometrical realization of set systems and probabilistic communication complexity*, Proc. 26th IEEE Symposium on Foundations of Computer Science, 1985, pp. 277–280.

[AR92] J. Arias-de-Reyna and L. Rodríguez-Piazza, *Finite metric spaces needing high dimension for Lipschitz embeddings in Banach spaces*, Israel Journal of Mathematics **79** (1992), 103–113.

[Ben66] C. T. Benson, *Minimal regular graphs of girth eight and twelve*, Canadian Journal of Mathematics **18** (1966), 1091-1094.

[Big74] N. Biggs, *Algebraic Graph Theory*, Cambridge University Press, 1974.

[Bol78] B. Bollobás, *Extremal Graph Theory*, Academic Press, London, 1978.

[Bou85] J. Bourgain, *On Lipschitz embedding of finite metric spaces in Hilbert space*, Israel Journal of Mathematics **52** (1985), 46–52.

[Bou86] J. Bourgain, *The metrical interpretation of superreflexivity in Banach spaces*, Israel Journal of Mathematics **56** (1986), 222–230.

[BMW86] J. Bourgain, V. Milman and H. Wolfson, *On type of metric spaces*, Transactions of the American Mathematical Society **294** (1986), 295–317.

[Gru67] B. Grünbaum, *Convex Polytopes*, John Wiley & Sons, London, 1967.

[JL84] W. Johnson and J. Lindenstrauss, *Extensions of Lipschitz maps into a Hilbert space*, Contemporary Mathematics **26** (Conference in Modern Analysis and Probability), American Mathematical Society, 1984, pp. 189–206.

[JLS87] W. Johnson, J. Lindenstrauss and G. Schechtman, *On Lipschitz embedding of finite metric spaces in low dimensional normed spaces*, in *Geometrical Aspects of Functional Analysis* (J. Lindenstrauss and V. D. Milman, eds.), Lecture Notes in Mathematics **1267**, Springer-Verlag, Berlin–Heidelberg, 1987.

[LLR94] N. Linial, E. London and Yu. Rabinovich, *The geometry of graphs and some of its algorithmic applications*, Proceedings of the 35th IEEE Symposium on Foundations of Computer Science (1994), pp. 577–591. Journal version: Combinatorica **15** (1995), 215–245.

[LPS88] A. Lubotzky, R. Phillips and P. Sarnak, *Ramanujan graphs*, Combinatorica **8** (1988), 261–277.

[Ma90] J. Matoušek, *Bi-Lipschitz embeddings into low-dimensional Euclidean spaces*, Commentationes Mathematicae Universitatis Carolinae **31** (1990), 589–600.

[Ma91] J. Matoušek, *Note on bi-Lipschitz embeddings into normed spaces*, Commentationes Mathematicae Universitatis Carolinae **33** (1992), 51–55.

[Mil64] J. Milnor, *On the betti numbers of real algebraic varieties*, Proceedings of the American Mathematical Society **15** (1964), 275–280.

[Tho65] R. Thom, *Sur l'homologie des variétés algébriques reélles*, in *Differential and Combinatorial Topology* (S. S. Cairns, ed.), Princeton Univ. Press, 1965.