Note on the nonlinear Dvoretzky theory and the Bartal-Linial-Mendel-Naor theorem, following Naor and Tao

Aziz Ben Nejma

May 15, 2026

Abstract

In this note, we give some context on the linear and nonlinear Dvoretzky theory and prove a sharper version of a theorem of Bartal, Linial, Mendel and Naor, following Naor and Tao.

1. Introduction

Definition 1.1. *Let (X_1, d_1) and (X_2, d_2) be two metric spaces. X_1 is said to embed into X_2 with distortion D if there exists a map $f : X_1 \to X_2$ and $\alpha > 0$ such that*

$$\forall x, y \in X_1 \quad \alpha d_1(x, y) \leq d_2(f(x), f(y)) \leq \alpha D d_1(x, y).$$

Due to the nice (and intuitive) metric properties of Hilbert spaces, embedding metric spaces in Hilbert spaces has been thoroughly investigated and interst in these embeddings dates back to the work of Scheonberg, Bourgain, Johnsson and Lindenstrauss [Sch35, Bou85, JL84] among others. To give a few fundamental examples, it is worth mentioning that every n-point metric space embeds into a Hilbert space with distortion $O(\log(n))$ [Bou85], that n-point planar graphs embed into a Hilbert space with distortion $O(\sqrt{\log(n)})$ [Rao99] and that n-point ultrametric spaces (i.e. metric spaces whose distance d satisfies $d(x, y) \leq \max(d(x, z), d(z, y))$ for all x, y, z) isometrically embed into a Hilbert space.

Another natural question at this point is, given a metric space X, are there large subsets of X that look like subsets of a Hilbert space from the metric point of view ? In other words, how large are subsets of X that embed into a Hilbert space with a "small" distortion ?
In this context, Dvoretzky [Dvo61] proved the following theorem, settling a conjecture of Grothendieck.

Theorem 1.2. *Let B be a Banach space of dimension n and $D > 1$. There exists $k(n, D)$ depending only on n and D and a $k(n, D)$-dimensional linear subspace of B that embeds into a Hilbert space with distortion D. Moreover, $k(n, D) \to +\infty$ as $n \to +\infty$.*

Having in mind this theorem, Bourgain, Figiel and Milman initiated a nonlinear Dvoretzky theory in the seminal work [BFM86] by generalizing Dvoretzky's theorem to metric spaces in the following way :

Theorem 1.3. *Let (X, d) an n-point metric space, and $D > 1$. There exists $k(n, D)$ depending only on n and D and a $k(n, D)$-point subset $S \subseteq X$ such that S embeds into a Hilbert space with distortion D. Moreover, $k(n, D) \geq C(D)\log(n)$ where $C(D)$ is positive constant depending on D.*

This result led to numerous subsequent developments such as the ultrametric skeleton theorem proved by Mendel and Naor [MN13] in which the $\log(n)$ lower bound can be strengthened to a n^α bound for "large" distortions. More precisely,

Theorem 1.4. *Let (X, d) be an n-point metric space and $\varepsilon \in (0, 1)$. There exists a subset $S \subseteq X$ such that $|S| \geq n^{1-\varepsilon}$ and S embeds with distortion $O(\frac{1}{\varepsilon})$ in a Hilbert space.*

The $\log(n)$ versus n^α is a remarkable phenomenon that provides a key distinction between small and large distortions. Indeed, the starting point of Naor and Tao's paper [NT10] is the following dichotomy, proved by Bartal, Linial, Mendel and Naor [BLMN05]

Theorem 1.5. *(a) Let $D \in (1, 2)$, then any n-point metric space X has a subset S such that $|S| \geq c(D)\log(n)$ and S embeds into a Hilbert space with distortion D, where $c(D)$ is a positive constant. Moreover, the $\log(n)$ bound is optimal.*

(b) Let $D > 2$, then any n-point metric space X has a subset S such that $\log(|S|) \geq (1-\alpha(D))\log(n)$ and S embeds into a Hilbert space with distortion D, where $\alpha(D) < 1$ is a positive constant. Moreover, the $\log(n)$ bound is optimal for general metric spaces.

Naor and Tao's paper provides a novel and elementary probabilistic approach in order to obtain explicit constants $\alpha(D)$, using a scale-oblivious fragmentation procedure, that proves the bound of Theorem 1.5(b) with $\alpha(D) \in (0, 1)$ being the unique solution of the equation $\alpha(1-\alpha)^{\frac{1-\alpha}{\alpha}} = \frac{2}{D}$.

2. Main proof

Let (X, d) be an n-point metric space and $D > 2$. We will show that there exists $S \subseteq X$ with $|S| \geq n^{1-\alpha(D)}$ such that S embeds into an ultrametric (hence in a Hilbert) space with distortion D.

In the sequel, we denote by $B(x, r)$ the closed ball centered at x with radius r.

We divide the proof in three steps :

Step 1. *One iteration.* Let $1 = r_0 > r_1 > r_2 > \dots$ a decreasing sequence such that $r_n \to 0$ as $n \to \infty$.

Let ν_0 be any positive measure on X. Let $(x_n)_{n\in\mathbb{N}^*}$ be an i.i.d. sequence of random variables uniformly distributed on X. Define, for $k \in \mathbb{N}^*$,

$$A_k^1 = \{x \in X : d(x, x_k) \leq r_1 \text{ and } d(x, x_i) > r_1 + \frac{2r_0}{D} \text{ for } i < k\}$$

and

$$A^1 = \bigcup_{k\geq 1} A_k^1.$$

Observe that by the triangle inequality, almost-surely, if $m > n$, $a \in A_n^1$ and $b \in A_m^1$, then $d(a,b) > \frac{2r_0}{D}$ (since $d(x_n, a) \leq r_1$ and $d(x_n, b) > r_1 + \frac{2r_0}{D}$) and that $\operatorname{diam}(A_k^1) \leq 2r_1$ for all $k \geq 1$ since $A_k^1 \subseteq B(x_k, r_1)$.

Moreover, for $x \in X$ and $k \in \mathbb{N}^*$, it holds that

$$\mathbb{P}(x \in A_k^1) = \frac{|B(x,r_1)|}{n}(1 - \frac{|B(x, r_1 + \frac{2r_0}{D})|}{n})^{k-1}$$

from which it follows (since the A_k^1's are disjoint) that

$$\mathbb{P}(x \in A^1) = \frac{|B(x,r_1)|}{|B(x, r_1 + \frac{2r_0}{D})|}.$$

Hence,

$$\mathbb{E}\left[\int \mathbf{1}_{A^1}(x)\frac{|B(x, r_1 + \frac{2r_0}{D})|}{|B(x,r_1)|}d\nu_0(x)\right] = \int d\nu_0(x).$$

Therefore, There exists a subset $E^1 \subseteq X$ such that $E^1 = \bigcup_{k\geq 1} E_k^1$ and

$$\int \mathbf{1}_{E^1}(x)\frac{|B(x, r_1 + \frac{2r_0}{D})|}{|B(x,r_1)|}d\nu_0(x) \geq \int d\nu_0(x).$$

Moreover, each E_k^1 has diameter at most $2r_1$ and the distance between any element of E_k^1 and any element E_l^1 for $k \neq l$ is at least $\frac{2r_0}{D}$.

Step 2. *Iterating Step* 1. Now, we argue by induction setting

$$d\nu_n(x) = \mathbf{1}_{E^n}(x)\frac{|B(x, r_n + \frac{2r_{n-1}}{D})|}{|B(x,r_n)|}d\nu_{n-1}(x)$$

and exactly in the same way as in Step 1 replacing r_1 with r_{n+1} and r_0 with r_n in the definition of the random sequence, we hence obtain the existence of $E^{n+1} \subseteq E^n$ (if not, take the intersection $E^{n+1} \cap E^n$ as ν_n is supported in E^n) and the E_k^{n+1}'s that satisfy the same inequalities as in the first step, namely

- E_k^{n+1} has diameter at most $2r_{n+1}$.
- If $a \in E_k^{n+1}$ and $b \in E_l^{n+1}$ with $k \neq l$ then $d(a,b) \geq \frac{2r_n}{D}$.

- $\displaystyle\int \mathbf{1}_{E^{n+1}}(x)\frac{|B(x, r_{n+1}+\frac{2r_n}{D})|}{|B(x,r_{n+1})|}d\nu_n(x) \geq \int d\nu_n(x)$ $\left(\text{i.e. } \int d\nu_{n+1}(x) \geq \int d\nu_n(x)\right)$.

It follows from the last inequality that for all $n \in \mathbb{N}^*$

$$\int d\nu_n(x) \geq \int d\nu_0(x).$$

Setting

$$\nu_0 = \sum_{x\in X}\left(\prod_{n=1}^{+\infty}\frac{|B(x,r_n)|}{|B(x,r_n+\frac{2r_{n-1}}{D})|}\right)\delta_x$$

yields $\dfrac{d\nu_n}{d(\sum_{x\in X}\delta_x)} \leq 1$ for all $n \in \mathbb{N}$. Since ν_n is supported in E_n, it holds that

$$|E^{n+1}| \geq \int d\nu_0(x)$$

for all $n \in \mathbb{N}$. Finally, since the E^n's are nested subsets, we have

$$|E| \geq \int d\nu_0(x)$$

where $E = \bigcap_{n\geq 1} E_n$.

Step 3. *An ultrametric distance on E.*

For $a, b \in E$, let $n(a,b)$ be the largest integer n such that for all $j \leq n$, there exists $k \in \mathbb{N}^*$ such that $a, b \in A_k^j$ and define $\rho(a,b) = 2r_{n(a,b)}$. It is immediate to check (since r is non-increasing and $\lim r_n = 0$) that ρ is an ultrametric distance on E.

Let $a, b \in E$, since $a, b \in A_k^{n(a,b)}$ for some k, we have $d(a,b) \leq 2r_{n(a,b)} = \rho(a,b)$. Moreover, a, b do not belong to the same $A_k^{n(a,b)}$, therefore, $d(a,b) \geq \frac{2r_{n(a,b)}}{D}$. Hence, E embeds in an ultrametric space with distortion D.

Step 4. *Estimating $\int d\nu_0(x)$.*

The quantity $\displaystyle\int d\nu_0(x)$ only depends on the choice of the sequence $(r_n)_n$. Again, we will use the probabilistic method to find a suitable sequence $(r_n)_n$ so that $\displaystyle\int d\nu_0(x) \geq n^{1-\alpha(D)}$. Let U be a random variable uniformly distributed on $[0,1]$. Define the decreasing random sequence $(r_n)_n$ such that $r_n = (1-\alpha)^{\frac{U+n-1}{\alpha}}$ for $n \in \mathbb{N}$. By Jensen's inequality

$$\sum_{x\in X}\mathbb{E}\left[\prod_{n=1}^{+\infty}\frac{|B(x,r_n)|}{|B(x,r_n+\frac{2r_{n-1}}{D})|}\right] \geq \sum_{x\in X}\exp\left(\mathbb{E}\left[\sum_{n=1}^{\infty}\log\left(\frac{|B(x,r_n)|}{|B(x,r_n+\frac{2r_{n-1}}{D})|}\right)\right]\right).$$

Let r be a positive random variable, and $x \in X$. Let $0 = d_0 < d_1 < d_2 < \cdots < d_m$ be the values taken by the function $y \in X \mapsto d(x,y)$ and $d_{m+1} = +\infty$. We have

$$\begin{aligned}\mathbb{E}[\log|B(x,r)|] &= \sum_{j=0}^{m}\mathbb{P}(d_j \leq r < d_{j+1})\log|B(x,d_j)| \\ &= \sum_{j=1}^{m}\mathbb{P}(d_j \leq r)\log\left(|B(x,d_j)\backslash B(x,d_{j-1})|\right)\end{aligned}$$

where the last equality follows from a discrete integration by part, observing that $\log(|B(x,d_0)|) = 0$. Applying this identity to r_n and $r_n + \frac{2r_{n-1}}{D}$ yields

$$\begin{aligned}
\mathbb{E}\left[\sum_{n=1}^{+\infty}\log\left(\frac{|B(x,r_n)|}{|B(x,r_n+\frac{2r_{n-1}}{D})|}\right)\right] &= \sum_{j=1}^{m}\left(\sum_{n=1}^{+\infty}\mathbb{P}\left(r_n \geq d_j\right) - \mathbb{P}\left(r_n + \frac{2r_{n-1}}{D} \geq d_j\right)\right)\log\left(\frac{|B(x,d_j)|}{|B(x,d_{j-1})|}\right)\\
&= \sum_{j=1}^{m}\left(\sum_{n=1}^{+\infty} -\mathbb{P}\left(r_n + \frac{2r_{n-1}}{D} \geq d_j > r_n\right)\right)\log\left(\frac{|B(x,d_j)|}{|B(x,d_{j-1})|}\right)\\
&\geq -\sigma\sum_{j=1}^{m}\log\left(\frac{|B(x,d_j)|}{|B(x,d_{j-1})|}\right)\\
&= -\sigma\log(|X|)\\
&= -\sigma\log(n)
\end{aligned}$$

where $\sigma = \sup_{a>0}\sum_{n=1}^{+\infty}\mathbb{P}(r_n + \frac{2r_{n-1}}{D} \geq a > r_n) \in [0,+\infty]$, from which it follows

$$\int d\nu_0(x) \geq \sum_{x\in X}\mathbb{E}\left[\prod_{n=1}^{+\infty}\frac{|B(x,r_n)|}{|B(x,r_n+\frac{2r_{n-1}}{D})|}\right] \geq \sum_{x\in X} e^{-\sigma\log(n)} = n^{1-\sigma}.$$

Note that we we did not use yet the exact definition of r_n. We will finally use it to show that $\sigma \leq \alpha(D)$.

Let $a > 0$, we have

$$\sum_{n=1}^{+\infty}\mathbb{P}\left(r_n < a \leq r_n + \frac{2r_{n-1}}{D}\right) = \sum_{n=1}^{+\infty}\mathbb{P}\left((1-\alpha)^{\frac{U+n-1}{\alpha}} < a \leq (1-\alpha)^{\frac{U+n-1}{\alpha}} + \frac{2(1-\alpha)^{\frac{U+n-2}{\alpha}}}{D}\right).$$

Moreover,

$$(1-\alpha)^{\frac{U+n-1}{\alpha}} < a \iff U+n-1 > \alpha\frac{\log(a)}{\log(1-\alpha)}$$

and

$$\begin{aligned}
a \leq (1-\alpha)^{\frac{U+n-1}{\alpha}} + \frac{2(1-\alpha)^{\frac{U+n-2}{\alpha}}}{D} &\iff a \leq (1-\alpha)^{\frac{U+n-1}{\alpha}}\left(1+\frac{2}{D}(1-\alpha)^{\frac{-1}{\alpha}}\right)\\
&\iff U+n-1 \leq \alpha\log\left(\frac{a}{1+\frac{2}{D}(1-\alpha)^{\frac{-1}{\alpha}}}\right)/\log(1-\alpha)
\end{aligned}$$

which yields

$$\sum_{n=1}^{+\infty}\mathbb{P}\left(r_n < a \leq r_n + \frac{2r_{n-1}}{D}\right) = \sum_{n=1}^{+\infty}\mathbb{P}\left(U+n-1 \in I\right)$$

where $I = \left(\alpha\frac{\log(a)}{\log(1-\alpha)}, \alpha\log\left(\frac{a}{1+\frac{2}{D}(1-\alpha)^{\frac{-1}{\alpha}}}\right)/\log(1-\alpha)\right]$ is an interval of length $\frac{\alpha}{\log(1-\alpha)}\log\left(\frac{1}{1+\frac{2}{D}(1-\alpha)^{\frac{-1}{\alpha}}}\right) = \alpha$ (since $\alpha(1-\alpha)^{\frac{1-\alpha}{\alpha}} = \frac{2}{D}$).

Since I is of length α and U is uniform on $[0,1]$, we have $\sum_{n\geq 1} \mathbb{P}(U+n-1 \in I) \leq \alpha$, therefore $\sigma \leq \alpha$ and we showed the existence of a subset $E \subseteq X$ of cardinality at least $n^{1-\alpha(D)}$ such that E embeds in an ultrametric space with distortion D. The proof is complete. ■

References

[BFM86] Jean Bourgain, Tadeusz Figiel, and Vitali D. Milman. On Hilbertian subsets of finite metric spaces. *Israel Journal of Mathematics*, 55(2):147–152, 1986.

[BLMN05] Yair Bartal, Nathan Linial, Manor Mendel, and Assaf Naor. On metric Ramsey-type dichotomies. *Journal of the London Mathematical Society*, 71(2):289–303, 2005.

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

[Dvo61] Aryeh Dvoretzky. Some results on convex bodies and Banach spaces. *Proceedings of the International Symposium on Linear Spaces*, pages 123–160, 1961.

[JL84] William B. Johnson and Joram Lindenstrauss. Extensions of Lipschitz mappings into a Hilbert space. *Contemporary Mathematics*, 26:189–206, 1984.

[MN13] Manor Mendel and Assaf Naor. Ultrametric skeletons. *Proceedings of the National Academy of Sciences of the United States of America*, 110(48):19284–19289, 2013.

[NT10] Assaf Naor and Terence Tao. Scale-oblivious metric fragmentation and the non-linear Dvoretzky theorem, 2010.

[Rao99] Satish Rao. Small distortion and volume preserving embeddings for planar and Euclidean metrics. In *Proceedings of the Fifteenth Annual Symposium on Computational Geometry (SoCG)*, pages 300–306. ACM, 1999.

[Sch35] Isaac J. Schoenberg. Remarks to Maurice Fréchet's article "sur la définition axiomatique d'une classe d'espaces vectoriels distanciés applicables vectoriellement sur l'espace de Hilbert". *Annals of Mathematics*, 36(3):724–732, 1935.