On Stochastic Decompositions of Metric Spaces

Scribe of a talk given at the fall school "Metric Embeddings : Constructions and Obstructions". 3-7 November 2014, Paris, France

Ofer Neiman*

Abstract

In this talk we will go over stochastic metric decompositions. These are random partitioning of a metric space into pieces of bounded diameter, such that for each point, a certain ball centered at it has a good chance of being contained in a single cluster. These decompositions play a role in metric embedding, metric Ramsey theory, higher-order Cheeger inequalities for graphs, metric and Lipschitz extension problems and approximation algorithms. We will then see an example of such decompositions for doubling metric spaces.

The second part of the talk will be devoted to embedding finite metrics into normed spaces using these decompositions. We will begin with a basic result due to Rao, and time permitting, the Measure Descent approach.

1 Preliminaries

Let (X, d) be a metric space. For $x \in X$ and $r \geq 0$, denote by $B_X(x, r) = \{z \in X \ : \ d(x, z) \leq r\}$ the closed ball of radius r centered at x (we omit the subscript when clear from context). By $B^\circ(x, r) = \{z \in X \ : \ d(x, z) < r\}$ we mean the open ball. The *diameter* of X is denoted as $\mathrm{diam}(X) = \sup_{x,y \in X} d(x, y)$, and its aspect ratio $\Phi(X) = \frac{\sup_{x,y \in X} d(x,y)}{\inf_{x,y \in X} d(x,y)}$.

Distortion. If (X, d_X) and (Y, d_Y) are metric spaces, a mapping $f : X \to Y$ has distortion at most K if there exists $C > 0$ such that for any $x, y \in X$,

$$\frac{C}{K} \cdot d_X(x, y) \leq d_Y(f(x), f(y)) \leq C \cdot d_X(x, y) \ .$$

The infimum K is called the *distortion* of f. We denote by $c_Y(X)$ the smallest distortion of a mapping from X to Y. In the special case where $Y = \ell_p$ we denote the smallest distortion by $c_p(X)$.

*Ben-Gurion University of the Negev, Israel. Email: `neimano@cs.bgu.ac.il`. Partially funded by the Lynne and William Frankel Center for Computer Sciences.

Stochastic Decompositions. A partition P of X is a pairwise disjoint collection of clusters that covers X. We say that the partition is Δ-bounded if for any cluster $C \in P$, $\mathrm{diam}(C) \leq \Delta$. For $x \in X$, let $P(x)$ denote the unique cluster containing x in P. Denote by $\mathcal{P}$ the collection of all partitions of X. For a distribution $\Pr$ over $\mathcal{P}$, recall that $\mathrm{supp}(\Pr) = \{P \in \mathcal{P} : \Pr[P] > 0\}$.

Definition 1 (Padded-Decomposition)**.** *A stochastic decomposition of a metric space (X, d) is a distribution $\Pr$ over $\mathcal{P}$. Given $\Delta > 0$, the decomposition is called Δ-bounded if for all $P \in \mathrm{supp}(\Pr)$, P is Δ-bounded. For a function $\epsilon : X \to [0, 1]$, a Δ-bounded decomposition is called* ϵ-padded *if the following condition holds:*

- *For all $x \in X$,* $\Pr[B(x, \epsilon(x) \cdot \Delta) \subseteq P(x)] \geq 1/2$.

Definition 2 (Modulus of Decomposability)**.** *A metric space (X, d) is called α-decomposable if for every $\Delta > 0$ there exists a Δ-bounded stochastic decomposition of X with padding parameter $\epsilon(x) = 1/\alpha$, for all $x \in X$. The* modulus of decomposability *of X is defined as*

$$\alpha_X = \inf\{\alpha : X \text{ is } \alpha\text{-decomposable}\} .$$

Let $\mathcal{X}$ be a family of metric spaces. If every member of the family has modulus of decomposability at most β, then we say that the family $\mathcal{X}$ is β-decomposable.

In general, every finite metric space has $\alpha_X \leq O(\log |X|)$ [Bar96], which is quantitatively the best possible, as exhibited by the family of expander graphs. However, there are many families of metric space which are decomposable (that is, $O(1)$-decomposable), as described in the next section.

2 Examples of Decomposable Metric Spaces

There are several families of metric spaces which are known to be decomposable, for instance, metrics arising from shortest path on planar graphs, or bounded tree-width graphs, and in general all graphs excluding some fixed minor. Another example is the family of metrics with bounded Negata-Assouad dimension, which contains doubling metric spaces, subsets of compact Riemannian surfaces, Gromov hyperbolic spaces of bounded local geometry, Euclidean buildings, symmetric spaces, and homogeneous Hadamard manifolds. Here for the sake of simplicity, we show the decomposability of the family of doubling metric spaces. The best quantitative result, which is shown here, is due to [GKL03].

2.1 Doubling Metrics

Let λ be a positive integer. A metric space (X, d) has doubling constant λ if for all $x \in X$ and $r > 0$, the ball $B(x, 2r)$ can be covered by λ balls of radius r. The *doubling dimension* of (X, d) is defined as $\log_2 \lambda$.

Comment: The doubling constant may be defined in terms of diameters of sets, rather than radii of balls, but this which affects the dimension by a factor of 2 at most.

Definition 3. *(Nets) For $r > 0$, an r-*net *of a metric (X, d) is a set $N \subseteq X$ satisfying the following properties:*

1. **Packing:** *For every* $u, v \in N$, $d(u,v) > r$.
2. **Covering:** *For every* $x \in X$ *there exists* $u \in N$ *such that* $d(x,u) \le r$.

Proposition 1. *Let* (X,d) *be a metric with doubling constant* λ*, and* N *be an* r*-net of* X*. If* $S \subseteq X$ *is a set of diameter* t*, then*

$$|N \cap S| \le \lambda^{\lceil \log 4t/r \rceil} .$$

Proof. Note that S is contained in a ball of radius $2t$, and that this ball can be covered by λ^k balls of radius $2t/2^k$. Letting $k = \lceil \log 4t/r \rceil$ we get that these small balls have radius at most $r/2$ and thus cannot contain more than a single point of N. □

Theorem 1. *Let* (X,d) *be a metric space with doubling constant* λ*, then* $\alpha_X \le O(\log \lambda)$.

Proof. Fix any $\Delta > 0$, and take N to be a $\Delta/4$-net of X. We now describe the random partition P. Let σ be a random permutation of N, and choose $r \in [\Delta/4, \Delta/2]$ uniformly at random. For each $u \in N$ define a cluster

$$C_u = \{x \in X \ : \ d(x,u) \le r \text{ and } \sigma(u) < \sigma(v) \text{ for all } v \in N \text{ with } d(x,v) \le r\} .$$

In words, every net point in order of σ collects to its cluster all the unassigned points within distance r from it. Then $P = \{C_u\}_{u \in N} \setminus \{\emptyset\}$. Note that this is indeed a Δ-bounded partition, due to the covering property of nets.

Fix some $x \in X$ and let $t = \Delta/(100 \ln \lambda)$, we need to show that the event $\{B(x,t) \not\subseteq P(x)\}$ happens with probability at most $1/2$. Observe that if $u \in N$ has $d(x,u) \ge \Delta$, then $C_u \cap B(x,t) = \emptyset$ for any choice of r (because $r \le \Delta/2$ and $t < \Delta/2$). Let $S = B(x,\Delta) \cap N$, and note that by Proposition 1, $m := |S| \le \lambda^5$. Arrange the points $s_1, s_2, \ldots, s_m \in S$ in order of increasing distance from x. For $j \in [m]$, let I_j be the interval $[d(x,s_j) - t, d(x,s_j) + t]$. We say that the point s_j *cuts* $B(x,t)$ if s_j is the minimal element (of the permutation σ) for which $r \ge d(x,s_j) - t$, and also $r \in I_j$. Observe that if $B(x,t) \not\subseteq P(x)$ then there must be some s_j which cuts $B(x,t)$.

$$\begin{aligned}
\Pr[B(x,t) \not\subseteq P(x)] &\le \sum_{j=1}^{m} \Pr[s_j \text{ cuts } B(x,t)] \\
&\le \sum_{j=1}^{m} \Pr[r \in I_j \wedge \forall_{i<j} \sigma(s_j) < \sigma(s_i)] \\
&= \sum_{j=1}^{m} \Pr[r \in I_j] \cdot \Pr[\forall_{i<j} \sigma(s_j) < \sigma(s_i) \mid r \in I_j] \\
&\le \sum_{j=1}^{m} \frac{2t}{\Delta/4} \cdot \frac{1}{j} \\
&\le \frac{8t}{\Delta} \cdot (1 + \ln m) .
\end{aligned}$$

The third inequality follows from the independent choices of r and σ. Plugging in the estimates for $t = \Delta/(100 \ln \lambda)$ and $m \le \lambda^5$, gives a bound of $1/2$ on the probability, as required. □

3 Embedding Decomposable Metrics into Normed Spaces

In this section we describe an embedding of finite decomposable metrics into ℓ_p space (for any $p \in [1, \infty)$). This is a simplified version of a result of [Rao99], in which the aspect ratio Φ is replaced by n.

Theorem 2. *Let (X, d) be a finite metric space with modulus of decomposability α and aspect ratio Φ, then $c_p(X) = O(\alpha \cdot \log^{1/p} \Phi)$.*

Proof. Let c be a universal constant to be determined later. Assume w.l.o.g that the minimal distance between two distinct points in X is 1 (by appropriate scaling), and thus $\mathrm{diam}(X) = \Phi$. For each $i \in I = \{0, 1, \ldots, \lceil \log \Phi \rceil\}$ and $j \in J = [c \log n]$ (where $n = |X|$), let P_{ij} be a 2^i-bounded $1/\alpha$-padded partition sampled from the distribution guaranteed to exist by the decomposability of (X, d). For each $i \in I$, $j \in J$, and each $C \in P_{ij}$ let $\tau(C) \in \{0, 1\}$ be a Bernoulli random variable chosen independently and uniformly. Define a random embedding $f_{ij} : X \to \mathbb{R}$ by

$$f_{ij}(x) = \tau(P_{ij}(x)) \cdot d(x, X \setminus P_{ij}(x)) \ ,$$

and let $f : X \to \mathbb{R}^{|I| \cdot |J|}$ by $f = \frac{1}{|J|^{1/p}} \bigoplus_{i \in I, j \in J} f_{ij}$.

Expansion. First we bound the expansion of the map f. Fix any $x, y \in X$, and any $i \in I$, $j \in J$. Next we show that $|f_{ij}(x) - f_{ij}(y)| \leq d(x, y)$. If it is the case that $P_{ij}(x) = P_{ij}(y)$ then by the triangle inequality

$$f_{ij}(x) - f_{ij}(y) = \tau(P_{ij}(x)) \cdot (d(x, X \setminus P_{ij}(x)) - d(y, X \setminus P_{ij}(x))) \leq d(x, y) \ .$$

Otherwise, if $y \notin P_{ij}(x)$, then

$$f_{ij}(x) - f_{ij}(y) \leq f_{ij}(x) \leq d(x, X \setminus P_{ij}(x)) \leq d(x, y) \ ,$$

and the bound on the absolute value follows by symmetry. Finally, we obtain that

$$\|f(x) - f(y)\|_p^p = \frac{1}{|J|} \sum_{i=1}^{|I|} \sum_{j=1}^{|J|} |f_{ij}(x) - f_{ij}(y)|^p \leq O(d(x, y)^p \cdot \log \Phi) \ .$$

Contraction. Now we bound the expected contraction of the embedding. Fix any $x, y \in X$, and let $i \in I$ be the unique value such that $2^i < d(x, y) \leq 2^{i+1}$. Since P_{ij} is 2^i-bounded, it must be that $P_{ij}(x) \neq P_{ij}(y)$, and as τ is chosen independently, there is probability of $1/4$ for the event $\mathcal{C}_j = \{\tau(P_{ij}(x)) = 1 \wedge \tau(P_{ij}(y)) = 0\}$. Also, by the definition of padded decomposition, we have that the event $\mathcal{D}_j = \{B(x, 2^i/\alpha) \subseteq P_{ij}(x)\}$ happens independently with probability at least $1/2$. Define $\mathcal{E}_j = \mathcal{C}_j \cap \mathcal{D}_j$. Thus with probability at least $1/8$ we have that event $\mathcal{E}_j$ holds and so

$$|f_{ij}(x) - f_{ij}(y)| = f_{ij}(x) = d(x, X \setminus P_{ij}(x)) \geq 2^i/\alpha \geq d(x, y)/(2\alpha) \ .$$

Note that events $\{\mathcal{E}_j\}_{j \in J}$ are mutually independent. Let Z_j be an indicator random variable for event $\mathcal{E}_j$, and set $Z = \sum_{j \in J} Z_j$. We have that $\mathbb{E}[Z] \geq |J|/8$, and by standard Chernoff bound

$$\Pr[Z \leq |J|/16] \leq e^{-|J|/128} \leq 1/n^2 \ ,$$

when c is sufficiently large. If indeed $Z \geq |J|/16$ it follows that

$$\|f(x) - f(y)\|_p^p \geq \frac{1}{|J|} \sum_{j \in J} |f_{ij}(x) - f_{ij}(y)|^p \geq \left(\frac{d(x,y)}{32\alpha} \right)^p .$$

By applying a union bound over the $\binom{n}{2}$ pairs, we obtain that with probability at least $1/2$ we have an embedding with distortion $O(\alpha \cdot \log^{1/p} \Phi)$.

□

4 Measured Descent

In this section we enhance the embedding so that the dependence on the aspect ratio is replaced by a dependence on n, and also improve the dependence on the decomposability parameter α. This result was obtained by [KLMN04].

Theorem 3. *For any $1 \leq p \leq \infty$, any finite metric space (X, d) with n points has $c_p(X) = O(\alpha_X^{1-1/p} \cdot \log^{1/p} n)$.*

The distortion guarantee is tight for every possible value of α_X, as shown by [JLM09].

We will need the following lemma, whose proof is similar to that of Theorem 1, which is based on the random partitions of [FRT04, CKR01].

Lemma 2. *For any $\Delta > 0$, any finite metric space (X, d) admits a Δ-bounded ϵ-padded stochastic decomposition, where for each $x \in X$:*

$$\epsilon(x) = \frac{1}{16 + 16 \ln \left(\frac{|B(x,\Delta)|}{|B(x,\Delta/8)|} \right)} .$$

Proof. Fix any $\Delta > 0$, and set $\epsilon : X \to [0, 1]$ as defined in the lemma. We now describe the random partition P. Let σ be a random permutation of X, and choose $r \in [\Delta/4, \Delta/2]$ uniformly at random. For each $u \in X$ define a cluster

$$C_u = \{x \in X \ : \ d(x, u) \leq r \text{ and } \sigma(u) < \sigma(v) \text{ for all } v \in X \text{ with } d(x, v) \leq r\} .$$

In words, every point in order of σ collects to its cluster all the unassigned points within distance r from it. Then $P = \{C_u\}_{u \in X} \setminus \{\emptyset\}$.

Fix some $x \in X$ and let $t = \epsilon(x) \cdot \Delta$, we need to show that the event $\{B(x, t) \nsubseteq P(x)\}$ happens with probability at most $1/2$. Let $a = |B(x, \Delta/8)|$ and $m = |B(x, \Delta)|$. Arrange the points $s_1, s_2, \ldots, s_m \in B(x, \Delta)$ in order of increasing distance from x. For $j \in [m]$, let I_j be the interval $[d(x, s_j) - t, d(x, s_j) + t]$. We say that the point s_j *cuts* $B(x, t)$ if s_j is the minimal element (of the permutation σ) for which $r \geq d(x, s_j) - t$, and also $r \in I_j$. Note that if $d(s_j, x) \leq \Delta/8$ then s_j cannot cut $B(x, t)$, because $d(x, s_j) + t \leq \Delta/8 + t < \Delta/4 \leq r$, so r cannot fall in the interval I_j. Also observe that if $u \notin B(x, \Delta)$ then $C_u \cap B(x, t) = \emptyset$ (for any choice of r).

$$
\begin{aligned}
\Pr[B(x,t) \not\subseteq P(x)] &\leq \sum_{j=1}^{m} \Pr[s_j \text{ cuts } B(x,t)] \\
&= \sum_{j=a+1}^{m} \Pr[s_j \text{ cuts } B(x,t)] \\
&\leq \sum_{j=a+1}^{m} \Pr[r \in I_j] \cdot \Pr[\forall_{i<j} \sigma(s_j) < \sigma(s_i) \mid r \in I_j] \\
&\leq \sum_{j=a+1}^{m} \frac{2t}{\Delta/4} \cdot \frac{1}{j} \\
&\leq \frac{8t}{\Delta} \cdot (1 + \ln(m/a)) \ .
\end{aligned}
$$

Plugging in the estimate for $t = \frac{\Delta}{16\left(1+\ln\left(\frac{|B(x,\Delta)|}{|B(x,\Delta/8)|}\right)\right)}$, gives a bound of $1/2$ on the probability, as required. $\square$

We will use the following definition of local growth-rate. Intuitively, the embedding will have more coordinates in scales for which there is a significant local growth change, and few (even none) when there is little change in the local cardinality of balls. Fix any $r > 0$, and define the *local growth-rate* of x at scale r as

$$
\mathrm{GR}(x,r) = \frac{|B(x,2r)|}{|B(x,r/512)|} \ .
$$

Proof of Theorem 3. For any integer $k \in \mathbb{Z}$ let P_k be a 2^k-bounded random partition sampled from Lemma 2. Denote by ϵ_k the padding function of P_k. For each k and each $C \in P_k$ let $\tau(C)$ be a $\{0,1\}$ Bernoulli uniform random variable chosen independently. For each $x \in X$ and integer $t > 0$ let $k(x,t) = \max\{k \in \mathbb{Z} \ : \ |B(x,2^k)| < 2^t\}$. Let $I = \{-5,-4,-3,-2,-1,0,1,2,3,4,5\}$. For each $t \in T := \{0,1,\ldots,\lceil \log n \rceil\}$ and $i \in I$ define a set

$$
W_t^i = \{x \in X \ : \ \tau(P_{k(x,t)+i}(x)) = 0\} \ .
$$

The embedding $f : X \to \mathbb{R}^{|I|\cdot|T|}$ is defined as $f(x) = (d(x,W_t^i) \ : \ i \in I, t \in T)$. By the triangle inequality, every coordinate of f is non-expansive, so for any $x,y \in X$

$$
\|f(x) - f(y)\|_p^p \leq |I| \cdot |T| \cdot d(x,y)^p = O(\log n) \cdot d(x,y)^p \ .
$$

It remains to show a bound on the contraction. Fix any $x,y \in X$, and let $R = d(x,y)$. It is not hard to verify that

$$
\max\left\{\frac{|B(x,2R)|}{|B(x,R/4)|}, \frac{|B(y,2R)|}{|B(y,R/4)|}\right\} \geq 2 \ . \tag{1}
$$

To see this, note that $B(x,R/4) \cap B(y,R/4) = \emptyset$, while both balls are contained in $B(x,2R)$ and also in $B(y,2R)$. Assume w.l.o.g that $\frac{|B(x,2R)|}{|B(x,R/4)|} \geq 2$. Let t_{lo}, t_{hi} be two integers such that $2^{t_{lo}-1} \leq |B(x,R/512)| < 2^{t_{lo}}$ and $2^{t_{hi}} \leq |B(x,2R)| < 2^{t_{hi}+1}$. Observe that $t_{hi} -$

$t_{lo} > \log \mathrm{GR}(x,R) - 2$, and due to (1) and our assumption on x, $\log \mathrm{GR}(x,R) \geq 1$, so that $t_{hi} - t_{lo} \geq 0$. Fix any integer $t \in [t_{lo}, t_{hi}]$, and let $k = k(x,t)$. Using the maximality of k in the definition of $k(t,x)$ we obtain that $|B(x,2^{k+1})| \geq 2^t$, so that $2^k \geq R/1024$ (otherwise $|B(x,2^{k+1})| \leq |B(x,R/512)| < 2^t$). We also have that $2^k < 2R$ (otherwise $|B(x,2^k)| \geq |B(x,2R)| \geq 2^t$). Let $u \in I$ be such that

$$R/32 \leq 2^{k+u} < R/16 \,.$$

(It can be checked that such $u \in I$ exists.)

For any $z \in B(x, R/2048)$, we claim that

$$k-1 \leq k(t,z) \leq k+2 \,. \tag{2}$$

To see this, note that $d(x,z) \leq R/2048 \leq 2^{k-1}$, and since $|B(x,2^{k+1})| \geq 2^t$ we conclude that $2^t \leq |B(z, 2^{k+1} + d(x,z))| \leq |B(z,2^{k+2})|$, or in other words that $k(t,z) \leq k+2$. Similarly, $2^t > |B(x,2^k)| \geq |B(z, 2^k - d(x,z))| \geq |B(z,2^{k-1})|$, so that $k(t,z) \geq k-1$.

Let $I' = \{-1,0,1,2\}$, and note that for $i \in I'$, by the assertion of Lemma 2,

$$\begin{aligned}
\epsilon_{k+u+i}(x) &= \frac{1}{16\left(1+\ln\left(\frac{|B(x,2^{k+u+i})|}{|B(x,2^{k+u+i-3})|}\right)\right)} \\
&\geq \frac{1}{16\left(1+\ln\left(\frac{|B(x,2^{k+u+2})|}{|B(x,2^{k+u-4})|}\right)\right)} \\
&\geq \frac{1}{16\left(1+\ln\left(\frac{|B(x,2R)|}{|B(x,R/512)|}\right)\right)} \\
&= \frac{1}{16\left(1+\ln \mathrm{GR}(x,R)\right)}
\end{aligned} \tag{3}$$

Let $\delta = \frac{1}{2048(1+\ln \mathrm{GR}(x,R))}$. Consider the set $W_t^u = \{z : \tau(P_{k(t,z)+u}(z)) = 0\}$. Let $\mathcal{E}_{big}$ be the event that $d(y, W_t^u) \geq \delta R/2$, and $\mathcal{E}_{small}$ be the event that $d(y,W_t^u) < \delta R/2$. Observe that these events are independent of the value of $\tau(P_{k+u+i}(x))$ for any $i \in I'$, because P_{k+u+i} is 2^{k+u+i}-bounded and $2^{k+u+i} < R/4$, thus for any $z \in B(y, \delta R/2)$, we have that $d(x,z) > 3R/4 \geq \mathrm{diam}(P_{k+u+i}(x))$ (note that $\mathcal{E}_{big}$ and $\mathcal{E}_{small}$ are indeed independent of values τ gives to points outside $B(y,\delta R/2)$).

If it is the case that $\mathcal{E}_{big}$ holds, then there is probability $1/2$ that $\tau(P_{k+u}(x)) = 0$ (independently as we noted above), in such a case $d(x,W_t^u) = 0$, and we obtain that

$$|d(x,W_t^u) - d(y,W_t^u)| \geq \delta R/2 \,.$$

The other case is that $\mathcal{E}_{small}$ holds. Let $\mathcal{E}$ be the event that for each $i \in I'$, $B(x, \epsilon_{k+u+i}(x) 2^{k+u+i}) \subseteq P_{k+u+i}(x)$ and $\tau(P_{k+u+i}(x)) = 1$. These events are clearly mutually independent, and since $2^{k+u+i} < R/4$ and P_{k+u+i} is 2^{k+u+i} bounded, they are also independent of $\mathcal{E}_{small}$. The probability of $\mathcal{E}$ is at least 2^{-8}. Consider any $z \in B(x, \delta R)$. If $\mathcal{E}$ indeed holds, then for each $i \in I'$: since $2^{k+u+i} \geq 2^{k+u-1} \geq R/64$ and due to (3) we conclude that $\epsilon_{k+u+i}(x) 2^{k+u+i} \geq \delta R$ so that $z \in P_{k+u+i}(x)$. From (2) we recall that $k(t,z) \in k + I'$, and as for any $i \in I'$, $\tau(P_{k+u+i}(x)) = 1$ (assuming event $\mathcal{E}$), it follows that $z \notin W_t^u$. We conclude that $d(x,W_t^u) \geq \delta R$, and as $\mathcal{E}_{small}$ holds:

$$|d(x,W_t^u) - d(y,W_t^u)| \geq \delta R/2 \,.$$

We conclude that for each of the (at least) $\min\{1, \log \mathrm{GR}(x,R) - 1\} \geq \log \mathrm{GR}(x,R)/2$ coordinates $t \in [t_{lo}, t_{hi}]$, with constant probability the contribution from a coordinate corresponding to t (and the appropriate value of u) is at least $\delta R/2$, and thus

$$\mathbb{E}[\|f(x) - f(y)\|_p^p] \geq \Omega\left(\frac{R}{\log \mathrm{GR}(x,R)}\right)^p \cdot \log \mathrm{GR}(x,R) \ .$$

Next, we devise another embedding $g : X \to \mathbb{R}$ using the same procedure, while sampling from the $1/\alpha_X$-padded distribution (guaranteed to exists as (X,d) is α_X-decomposable). The same proof holds (defining $\delta = 1/(128\alpha)$), and we obtain that

$$\mathbb{E}[\|f(x) - f(y)\|_p^p] \geq \Omega\left(\frac{R}{\alpha_X}\right)^p \cdot \log \mathrm{GR}(x,R) \ .$$

Finally, observe that choosing at random between f and g, we obtain in expectation the summation of these estimates (divided by 2), which is at least $\Omega(d(x,y)^p/\alpha_X^{p-1})$. This concludes the proof. □

Remark: In order to achieve an actual bound, not only on the expectation, one can use standard sampling and Chernoff bound as in the proof of Theorem 2, and obtain an embedding into R^D with $D = O(\log^2 n)$.

References

[Bar96] Y. Bartal. Probabilistic approximation of metric spaces and its algorithmic applications. In *Proceedings of the 37th Annual Symposium on Foundations of Computer Science*, pages 184–, Washington, DC, USA, 1996. IEEE Computer Society.

[CKR01] Gruia Calinescu, Howard Karloff, and Yuval Rabani. Approximation algorithms for the 0-extension problem. In *Proceedings of the twelfth annual ACM-SIAM symposium on Discrete algorithms*, SODA '01, pages 8–16, Philadelphia, PA, USA, 2001. Society for Industrial and Applied Mathematics.

[FRT04] Jittat Fakcharoenphol, Satish Rao, and Kunal Talwar. Approximating metrics by tree metrics. *SIGACT News*, 35(2):60–70, 2004.

[GKL03] Anupam Gupta, Robert Krauthgamer, and James R. Lee. Bounded geometries, fractals, and low-distortion embeddings. In *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '03, pages 534–, Washington, DC, USA, 2003. IEEE Computer Society.

[JLM09] Alex Jaffe, James R. Lee, and Mohammad Moharrami. On the optimality of gluing over scales. In *Proceedings of the 12th International Workshop and 13th International Workshop on Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, APPROX '09 / RANDOM '09, pages 190–201, Berlin, Heidelberg, 2009. Springer-Verlag.

[KLMN04] Robert Krauthgamer, James R. Lee, Manor Mendel, and Assaf Naor. Measured descent: A new embedding method for finite metrics. In *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, pages 434–443, Washington, DC, USA, 2004. IEEE Computer Society.

[Rao99] Satish Rao. Small distortion and volume preserving embeddings for planar and euclidean metrics. In *Proceedings of the fifteenth annual symposium on Computational geometry*, SCG '99, pages 300–306, New York, NY, USA, 1999. ACM.