

8.1 Partial and Scaling Distortion Embeddings

Recall the following definitions we have discussed in the first lecture.

**Definition 8.1.** Given an embedding \( f : X \to Y, \forall x \neq y \in X, \) \( \text{dist}_f(x,y) = \max \left\{ \frac{d_y(f(x),f(y))}{d_x(x,y)}, \frac{d_x(x,y)}{d_y(f(x),f(y))} \right\}. \)

**Definition 8.2 (Partial Embedding).** Let \((X,d_x)\) and \((Y,d_y)\) be any metric spaces, and \( G \subseteq \binom{X}{2} \). We say that \( (f,G) \) is partial embedding with distortion \( \alpha \geq 1 \), if \( \forall (x,y) \in G \) it holds that \( \text{dist}_f(x,y) \leq \alpha. \) Partial embedding \( (f,G) \) is called \((1-\epsilon)\)-partial, if \( |G| \geq (1-\epsilon)^{(\binom{X}{2})}. \)

**Definition 8.3 (Scaling Distortion).** Let \( \alpha : [0,1] \to \mathbb{R}^+ \) be a non-increasing function. We say that embedding \( f : X \to Y \) has an \( \alpha \)-scaling distortion if for all \( 0 \leq \epsilon \leq 1 \) there exists a set \( G_\epsilon \subseteq \binom{X}{2} \) such that \( (f,G_\epsilon) \) is a \((1-\epsilon)\)-partial embedding with distortion \( \alpha(\epsilon). \)

Note that for \( \epsilon < \left( \frac{n}{2} \right)^{-1} \) the above definition captures the notion of the worst case distortion of a non-contractive (or non-expansive) embedding \( f. \)

**Scaling with Average Distortion.** We discuss the strong relationship between scaling and \( \ell_q \) distortions. Recall the definition of the \( \ell_q \)-distortion: \( \ell_q \text{-dist}(f) = \left( \sum_{x \neq y \in X} (\text{dist}_f(x,y))^q \right)^{\frac{1}{q}}, \forall q \geq 1. \)

**Claim 8.1 (Exercise).** If an embedding \( f \) has an \( \alpha \)-scaling distortion, then \( \forall 1 \leq q < \infty: \)

1. \( \ell_q \text{-dist}(f) \leq \left( \frac{n}{2} \right)^{-\frac{1}{2}} \left( \sum_{i=1}^{\binom{n}{2}} \alpha \left( \frac{1}{2^q} \right) \right)^{\frac{1}{q}} + \alpha \left( \frac{1}{2^q} \right)^{\frac{1}{q}}. \)

2. \( \ell_q \text{-dist}(f) \leq \left( 2 \int_{\frac{1}{2^q}}^{\frac{1}{2}} (\alpha(x))^q dx \right)^{\frac{1}{q}}. \)

For another direction we make the following observation.

**Claim 8.2.** Let \( 1 \leq q < \infty, \alpha \geq 1, \) and \( f \) be an embedding. If \( \ell_q \text{-dist}(f) \leq \alpha, \) then for \( \epsilon = \frac{1}{2q}, \) \( f \) is an \((1-\epsilon)\)-partial embedding with distortion at most \( 2\alpha. \)

**Proof.** We have to show that there exists a set \( G \subseteq \binom{X}{2}, \) such that \( |G| \geq (1-\epsilon)^{(\binom{n}{2})}, \) and the distortion of \( f \) on that set is at most \( 2\alpha. \) Namely, we have to show that there are at most \( \epsilon^{\binom{S}{2}} \) pairs that can be distorted by more than \( 2\alpha. \) Assume by contradiction there is \( S \subseteq \binom{X}{2} \), \( |S| > \epsilon^{\binom{n}{2}} \) and every pair from \( S \) distorted by more than \( 2\alpha. \) Therefore, \( (\ell_q \text{-dist}(f))^q > \frac{\epsilon^{\binom{S}{2}}}{\epsilon^{\binom{n}{2}}} = \alpha^q, \) a contradiction. \( \square \)

The above claim means that for all \( \gamma > 1, f \) is \((1-1/\gamma^q)\)-partial embedding, with distortion \( \gamma \alpha. \) In other words, for any \( 0 < \epsilon < 1 (\epsilon = 1/\gamma^q), f \) is \((1-\epsilon)\)-partial embedding with distortion \( \alpha/\epsilon^{\frac{1}{q}}. \) Namely, \( f \) has \( \ell_q \text{-dist}(f)/\epsilon^{\frac{1}{q}}\)-scaling distortion.


8.1.1 Embedding into Trees with Scaling Distortion

Now we are ready to state the main result of this section.

**Theorem 8.3** ([1], [2]). The following holds.

1. ([1]) Any finite metric space is embeddable into ultra-metric, with scaling distortion \( O\left(\frac{1}{\sqrt[3]{\epsilon}}\right) \).

2. ([1]) Any weighted graph contains a spanning tree, with scaling distortion \( O\left(\frac{1}{\sqrt[3]{\rho}}\right) \).

3. ([2]) For any \( 0 < \rho < 1 \), any weighted graph \( G \) contains a spanning tree, with scaling distortion \( \tilde{O}\left(\sqrt{1/\epsilon/\rho}\right) \), and weight bounded by \((1 + \rho)\text{MST}(G)\). This result is tight with respect to \( \rho \).

For the first two items we conclude: for \( 1 \leq q < 2 \), \( \ell_q\text{-dist}(f) = O(1) \); \( \ell_2\text{-dist}(f) = O(\sqrt{\log n}) \); for \( q > 2 \), \( \ell_q\text{-dist}(f) = O(n^{1-2/q}) \).

**Proof of Theorem 8.3(1).** Let \((X, d)\) be an \( n\)-point metric space. We construct the embedding by induction on \( n \). The idea is to partition (in a smart way) \( X \) by combining the trees obtained by inductive steps on \( X_1 \) and \( X_2 \). Let \( f_1 : X_1 \rightarrow U_1 \) and \( f_2 : X_2 \rightarrow U_2 \) be embeddings obtained by induction. The ultra-metric tree for \( X \) is be obtained by composing \( U_1 \) and \( U_2 \), on the new root \( r \), with label \( \Delta(r) = \text{diam}(X) = \Delta \). Thus, we have to show how to decompose \( X \). Next we discuss what properties should such decomposition satisfy.

We have to show that there exists \( c > 0 \), such that for any \( 0 < \epsilon < 1 \) there is \( G_\epsilon \subseteq \left(\frac{2}{3}\right) \) with \( |G_\epsilon| \geq (1 - \epsilon)^{|X|/2} \), such that \( \forall x, y \in G_\epsilon, \text{dist}_f(x, y) \leq \frac{1}{c\sqrt[3]{\epsilon}} \). In other words, we have to show that there are at most \( \epsilon(|X|/2) \) pairs of points of \( X \) with distortion larger than \( \frac{1}{c\sqrt[3]{\epsilon}} \).

By the induction’s assumption, there are at most \( \epsilon(|X_1|/2) \) pairs of \( X_1 \), and at most \( \epsilon(|X_2|/2) \) pairs of \( X_2 \) with distortion larger than \( \frac{1}{c\sqrt[3]{\epsilon}} \). Note that if \( x \in X_1 \) and \( y \in X_2 \) such that \( d(x, y) \geq c\sqrt[3]{\epsilon}\Delta \), then \( \text{dist}_f(x, y) \leq \frac{1}{c\sqrt[3]{\epsilon}} \). Consider the set \( B_\epsilon = \{(x, y) | x \in X_1, y \in X_2, d(x, y) < c\sqrt[3]{\epsilon}\Delta\} \). Thus, we want to partition \( X \) in such a way, that for every \( \epsilon \), the number of pairs with large distortion is bounded by:

\[
\epsilon\left\lfloor \frac{|X_1|}{2} \right\rfloor + \epsilon\left\lfloor \frac{|X_2|}{2} \right\rfloor + |B_\epsilon| \leq \epsilon\left\lfloor \frac{|X|}{2} \right\rfloor \iff |B_\epsilon| \leq \epsilon|X_1| \cdot |X_2|.
\]

Thus, we show how to partition \( X \) such that above inequality holds for every \( \epsilon \).

Let \( u \in X \) be a point such that \( |\bar{B}(u, \frac{\Delta}{2})| \leq \frac{n}{2} \). Note that there is such \( x \), since if \( x, y \in X \) such that \( \Delta = d(x, y) \), then open balls of radius \( \frac{\Delta}{2} \) around \( x \) and \( y \) are disjoint and at least one of them contains at most \( \frac{n}{2} \) points. Let \( r > 0 \) be a radius (we will choose the value of \( r \) later), and let \( X_1(r) = \bar{B}(u, r) \) and \( X_2(r) = X \setminus X_1 \) (note that \( X_1 \) and \( X_2 \) are dependent on \( r \)). Define the following subsets of \( X \):

\[
S_1^{(r, \epsilon)} = \{w \in X_1(r) | d(w, u) > r - c\sqrt[3]{\epsilon}\Delta\}, \quad S_2^{(r, \epsilon)} = \{w \in X_2(r) | d(w, u) < r + c\sqrt[3]{\epsilon}\Delta\}.
\]

Note that \( B_\epsilon \subseteq S_1^{(r, \epsilon)} \times S_2^{(r, \epsilon)} \), implying \( |B_\epsilon| \leq |S_1^{(r, \epsilon)}| \cdot |S_2^{(r, \epsilon)}| \). Thus, we will prove that there exists \( r \), such that for all \( 0 < \epsilon < 1 \), \( |S_1^{(r, \epsilon)}| \cdot |S_2^{(r, \epsilon)}| \leq \epsilon|X_1(r)| \cdot |X_2(r)| \).

Let \( \bar{\epsilon} = \max\{\epsilon | |B(u, \frac{\sqrt[3]{\epsilon}\Delta}{4})| \geq en\} \). Note that this set is not empty, as at least \( \epsilon = \frac{1}{n} \) belongs to it. Also note that \( \bar{\epsilon} \leq 1/2 \). Thus, for any \( \epsilon > \bar{\epsilon} \), \( B\left(u, \frac{\sqrt[3]{\epsilon}\Delta}{4}\right) < en \). We will choose \( r \) in \( \left[\frac{\sqrt[3]{\epsilon}\Delta}{4}, \frac{\sqrt[3]{\epsilon}\Delta}{2}\right] \).

**Lemma 8.4.** If \( \epsilon > 32\bar{\epsilon} \), then (every \( r \) is good) \( \forall r \in \left[\frac{\sqrt[3]{\epsilon}\Delta}{4}, \frac{\sqrt[3]{\epsilon}\Delta}{2}\right] \), \( |S_1^{(r, \epsilon)}| \cdot |S_2^{(r, \epsilon)}| \leq \epsilon \cdot |X_1(r)| \cdot |X_2(r)| \).
Proof. Fix some \( r \in [\sqrt{\frac{\Delta}{4}}, \sqrt{\frac{\Delta}{2}}] \) and \( \epsilon > 32\bar{\epsilon} \). Note that \( |S_1^{(r, \epsilon)}| \leq |X_1^{(r)}| \), and \( |S_2^{(r, \epsilon)}| \leq |B(u, r + c\sqrt{\Delta})| \).

Also, it holds that \( r + c\sqrt{\Delta} \leq \frac{\sqrt{\Delta}}{2} \leq \frac{\sqrt{\Delta}}{2} + c\sqrt{\Delta} \leq \frac{\epsilon}{\bar{\epsilon}} \sqrt{\Delta} \leq \frac{\sqrt{\Delta}}{2}\left(\frac{1}{2\sqrt{32}} + c\right) \leq \frac{\sqrt{\Delta}}{2} \), where the last inequality holds if we choose \( c = \frac{1}{32\sqrt{2}} \), which will work for all inductive steps. Therefore,

\[
|S_2^{(r, \epsilon)}| \leq |B(u, r + c\sqrt{\Delta})| \leq \left|B(u, \frac{\sqrt{\Delta}}{4})\right| \leq \epsilon \frac{\sqrt{\Delta}}{2n}.
\]

Therefore, \( |S_1^{(r, \epsilon)}| \cdot |S_2^{(r, \epsilon)}| \leq \epsilon \cdot |X_1^{(r)}| \cdot \frac{n}{2} \leq (|X_2^{(r)}|\geq \frac{\epsilon}{2}) \leq \epsilon \cdot |X_1^{(r)}| \cdot |X_2^{(r)}| \).

\( \square \)

Lemma 8.5. There exists \( r \in [\sqrt{\frac{\Delta}{4}}, \sqrt{\frac{\Delta}{2}}] = I, \) such that for all \( \epsilon \leq 32\bar{\epsilon} \), \( |S_1^{(r, \epsilon)}| \cdot |S_2^{(r, \epsilon)}| \leq \epsilon \cdot |X_1^{(r)}| \cdot |X_2^{(r)}| \).

We first prove a small lemma. Let \( 0 \leq n < 2 \) be any real numbers, and let \( A(r_1, r_2) \) denote the size of the strip \( B(u, r_2) \setminus B(u, r_1) \).

Lemma 8.6. \( A(\sqrt{\frac{\Delta}{4}}, \sqrt{\frac{\Delta}{2}}) \leq 4\bar{\epsilon}n. \)

\( \square \)

Proof of Lemma 8.5. We say that \( r \) is a “bad” radius for some \( \epsilon \leq 32\bar{\epsilon} \) if \( |S_1^{(r, \epsilon)}| \cdot |S_2^{(r, \epsilon)}| \geq \epsilon |X_1^{(r)}| \cdot |X_2^{(r)}| \).

Denote by \( J \) the union of the intervals that constitute all bad values of \( r \) from \( I \). We will show that \( |J| < |I| \). We build \( J \) iteratively. At the beginning \( J = \emptyset \). At some step of the construction, consider all the values of \( r \in I \setminus J \) and all the values of \( \epsilon \leq 32\bar{\epsilon} \) such that the pair \((r, \epsilon)\) is a “bad” pair: \( r \) is bad for \( \epsilon \).

From all these pairs we choose one with the maximum \( \epsilon \) (we say maximum as we consider \( \epsilon \in [1/n, 32\bar{\epsilon}] \)). Denote this pair by \((\hat{r}, \hat{\epsilon})\). We add to \( J \) segment of length \( 2c\sqrt{\Delta} \) with center in \( \hat{r} \): \( \hat{r} - c\sqrt{\Delta}, \hat{r} + c\sqrt{\Delta} \).

Note that the length of the segment we add does not increase from step to step. We have to prove that \( |J| < |I| \) on the termination of the algorithm.

Consider the chosen pair \((\hat{r}, \hat{\epsilon})\). Then, \( A(\hat{r} - c\sqrt{\Delta}, \hat{r} + c\sqrt{\Delta}) \geq |S_1^{(\hat{r}, \hat{\epsilon})} \cup S_2^{(\hat{r}, \hat{\epsilon})}| = |S_1^{(\hat{r}, \hat{\epsilon})}| + |S_2^{(\hat{r}, \hat{\epsilon})}| \).

Note that \( |X_1^{(\hat{r})}| \geq |B(u, \sqrt{\frac{\Delta}{4}})| \geq \hat{\epsilon}n \). Therefore, \( |S_1^{(\hat{r}, \hat{\epsilon})}| \cdot |S_2^{(\hat{r}, \hat{\epsilon})}| \geq (\text{\( \hat{r}, \hat{\epsilon} \) is bad}) \geq \hat{\epsilon}X_1^{(\hat{r})} \cdot |X_2^{(\hat{r})}| \geq \frac{\hat{\epsilon}n^2}{2}. \)

Therefore, by the inequality of arithmetic and geometric means

\[
A(\hat{r} - c\sqrt{\Delta}, \hat{r} + c\sqrt{\Delta}) > 2\sqrt{\frac{\hat{\epsilon}n}{2}}.
\]

Therefore, we conclude that

\[
|J| \leq \sum_{i=1}^{t} |J_i| = \sum_{i=1}^{t} 2c\sqrt{\epsilon_i} \Delta = 2c\Delta \sum_{i} \sqrt{\epsilon_i} < (\text{have to prove}) < |I| = \frac{\sqrt{\Delta}}{4}.
\]

Recall that \( c = \frac{1}{32\sqrt{2}} \), therefore, we have to show that \( \sum_i \sqrt{\epsilon_i} < 4\sqrt{2}\sqrt{\epsilon} \). Note that each point of \( I \) belongs to at most 2 segments of \( J \), because the radius \( \hat{r} \) is always chosen outside the segments of \( J \), and the lengths of the segments do not increase from step to step. Therefore,

\[
\sum_i 2\sqrt{\frac{\epsilon_i}{2}} n < \sum_i A(\hat{r}_i - c\sqrt{\epsilon_i} \Delta, \hat{r}_i + c\sqrt{\epsilon_i} \Delta) \leq (\text{each point belongs to at most 2 segments}) \leq 2A(\sqrt{\frac{\Delta}{4}}, \sqrt{\frac{\Delta}{2}}) \leq 2\cdot 4\cdot \hat{\epsilon}n.
\]

Therefore \( \sum_i \sqrt{\epsilon_i} < 4\sqrt{2}\sqrt{\epsilon} \), which completes the proof. Note that this process can be computed in polynomial time, by discretization of values \( \epsilon \) and \( r \).

This completes the proof of the theorem.

\( \square \)
References
