Metric Embedding Theory and its Algorithms (67720) – Exercise 3

Mike Greenbaum 211747639

November 2021

1. (a) wlog the distance between two neigbors in the graph is 1, the result can be easily scaled to any positive scalar. Define the tree vertexes to be $v_1, \ldots, v_n$ where $\forall i \in [n-1], (v_i, v_{i+1}) \in E_T$.
Let $\frac{1}{2\cdot\binom{n}{2}} < \varepsilon < 1$.

Define $g(\varepsilon) = \frac{1}{\sqrt{2\frac{n}{n-1}\cdot\frac{\sqrt{\varepsilon}}{1-\varepsilon}} - \sqrt{\varepsilon}}, h(n) = g\left(\frac{1}{2\binom{n}{2}}\right)$.

Notice that $g(\varepsilon)$ is decreasing with ε and therefore the maximum for $\frac{1}{2\cdot\binom{n}{2}} < \varepsilon < 1$ is $h(n) = g\left(\frac{1}{2\binom{n}{2}}\right) = \Theta\left(\frac{1}{n}\right)$, therefore $g(\varepsilon) > 0$ and $h(n)$ is bounded from above.
Let c be the upper bound of $h(n)$ and $k = \frac{c\cdot n\sqrt{\varepsilon}}{c\cdot\sqrt{\varepsilon}+1}$.
Define $B = \{\{v_i, v_j\} \mid d_T(v_i, v_j) \leq k\}$.
Define $C_l = |\{(v_i, v_j) \mid d_T(v_i, v_j) = l\}|$. Notice that $C_l = 2\cdot(n-l)$. (v_1 with v_{l+1} ... v_{n-l} with v_n).
Notice that $|B| = \sum_{l=1}^{k}\frac{C_l}{2} = \sum_{l=1}^{k}(n-l) \geq \sum_{l=1}^{k}(n-k) = kn - k^2$
therefore

$$\begin{aligned}
kn - k^2 \geq (1-\varepsilon)\cdot\binom{n}{2} &\iff \frac{c\cdot n^2\sqrt{\varepsilon}}{c\cdot\sqrt{\varepsilon}+1} - \left(\frac{c\cdot n\sqrt{\varepsilon}}{c\cdot\sqrt{\varepsilon}+1}\right)^2 \geq \frac{1-\varepsilon}{2}\cdot n\cdot(n-1)\\
&\iff \frac{c\cdot n\sqrt{\varepsilon}}{c\cdot\sqrt{\varepsilon}+1} - \frac{c^2\cdot n\cdot\varepsilon}{(c\cdot\sqrt{\varepsilon}+1)^2} \geq \frac{1-\varepsilon}{2}\cdot(n-1)\\
&\iff \frac{c^2\cdot n\cdot\varepsilon + c\cdot n\sqrt{\varepsilon} - c^2\cdot n\cdot\varepsilon}{(c\cdot\sqrt{\varepsilon}+1)^2} \geq \frac{1-\varepsilon}{2}\cdot(n-1)\\
&\iff \frac{c\cdot n\sqrt{\varepsilon}}{(c\cdot\sqrt{\varepsilon}+1)^2} \geq \frac{1-\varepsilon}{2}\cdot(n-1)\\
&\iff \frac{n\sqrt{\varepsilon}}{\left(\sqrt{\varepsilon}+\frac{1}{c}\right)^2} \geq \frac{1-\varepsilon}{2}\cdot(n-1) \iff \frac{2n\sqrt{\varepsilon}}{(1-\varepsilon)\cdot(n-1)} \geq \left(\sqrt{\varepsilon}+\frac{1}{c}\right)^2\\
&\iff \sqrt{\frac{2n\sqrt{\varepsilon}}{(1-\varepsilon)\cdot(n-1)}} \geq \left(\sqrt{\varepsilon}+\frac{1}{c}\right) \iff \sqrt{2\frac{n}{n-1}\cdot\frac{\sqrt{\varepsilon}}{1-\varepsilon}} \geq \left(\sqrt{\varepsilon}+\frac{1}{c}\right)\\
&\iff \sqrt{2\frac{n}{n-1}\cdot\frac{\sqrt{\varepsilon}}{1-\varepsilon}} - \sqrt{\varepsilon} \geq \frac{1}{c} \iff c \geq \frac{1}{\sqrt{2\frac{n}{n-1}\cdot\frac{\sqrt{\varepsilon}}{1-\varepsilon}} - \sqrt{\varepsilon}} = \frac{1}{g(\varepsilon)}\\
&\Leftarrow c \geq \frac{1}{g\left(\frac{1}{2\binom{n}{2}}\right)} = h(n)
\end{aligned}$$

Notice that we chose c to be the upper bound of $h(n) = \frac{1}{g\left(\frac{1}{2\binom{n}{2}}\right)}$ and therefore $kn - k^2 \geq (1-\varepsilon)\cdot\binom{n}{2}$.

Notice that $\forall(v_i, v_j) \in B$ s.t. $n - |i-j| = \min\{|i-j|, n-|i-j|\}$:

$$dist_f(v_i, v_j) = expans_f(v_i, v_j) = \frac{|i-j|}{n-|i-j|} \leq \frac{k}{n-k} = \frac{\frac{c\cdot n\sqrt{\varepsilon}}{c\cdot\sqrt{\varepsilon}+1}}{n - \frac{c\cdot n\sqrt{\varepsilon}}{c\cdot\sqrt{\varepsilon}+1}} = \frac{\frac{c\cdot n\sqrt{\varepsilon}}{c\cdot\sqrt{\varepsilon}+1}}{\frac{n}{c\cdot\sqrt{\varepsilon}+1}} = \frac{c\cdot n\cdot\sqrt{\varepsilon}}{n} = c\cdot\sqrt{\varepsilon} = O\left(\frac{1}{\varepsilon}\right)$$

Notice that $\forall(v_i, v_j) \in B$ s.t. $|i-j| = \min\{|i-j|, n-|i-j|\}$, then $dist_f(v_i, v_j) = expans_f(v_i, v_j) = \frac{|i-j|}{|i-j|} = 1 = O\left(\frac{1}{\varepsilon}\right)$.

So we proved that there exists $B \subseteq \binom{n}{2}$ s.t. $|B| \geq (1-\varepsilon)\cdot\binom{n}{2}$ and that $\forall (x,y) \in B, dist_f(x,y) = O\left(\frac{1}{\sqrt{\varepsilon}}\right)$, therefore for $\alpha(\varepsilon) = O\left(\frac{1}{\sqrt{\varepsilon}}\right)$, we proved that that the embedding is α-scaling distortion. (We proved in class that it is the tight bound).

We will use the fact that we proved in class that $l_q - dist(f) \leq \left(4\int_{\frac{1}{2\binom{n}{2}}}^{\frac{1}{2}} (\alpha(x))^q dx\right)^{\frac{1}{q}}$.

Notice that

$$l_q - dist(f) \leq \left(4\int_{\frac{1}{2\binom{n}{2}}}^{\frac{1}{2}} (\alpha(x))^q dx\right)^{\frac{1}{q}} = \left(c\cdot 4\int_{\frac{1}{2\binom{n}{2}}}^{\frac{1}{2}} \left(\frac{1}{\sqrt{x}}\right)^q dx\right)^{\frac{1}{q}} = \left(4c\int_{\frac{1}{2\binom{n}{2}}}^{\frac{1}{2}} x^{\frac{-q}{2}} dx\right)^{\frac{1}{q}}$$

$$= \begin{cases} \left(4c\frac{x^{\frac{-q}{2}+1}}{\frac{-q}{2}+1} \Big|_{\frac{1}{2\binom{n}{2}}}^{\frac{1}{2}}\right)^{\frac{1}{q}} & q \neq 2 \\ \left(4c\ln(x) \Big|_{\frac{1}{2\binom{n}{2}}}^{\frac{1}{2}}\right)^{\frac{1}{q}} & q = 2 \end{cases} = \begin{cases} \left(2c\frac{x^{\frac{2-q}{2}}}{2-q} \Big|_{\frac{1}{2\binom{n}{2}}}^{\frac{1}{2}}\right)^{\frac{1}{q}} & q \neq 2 \\ \Theta\left(\log(n)\right)^{0.5} & q = 2 \end{cases}$$

$$\leq \begin{cases} \left(2c\frac{x^{\frac{2-q}{2}}}{q-2} \Big|_{\infty}^{\frac{1}{2\binom{n}{2}}}\right)^{\frac{1}{q}} & q > 2 \\ \left(2c\frac{x}{2-q} \Big|_{0}^{\frac{1}{2}}\right)^{\frac{1}{q}} & q < 2 \\ \Theta\left(\sqrt{\log(n)}\right) & q = 2 \end{cases} = \begin{cases} \left(\Theta(\frac{1}{q-2}\cdot n^{-\frac{2-q}{2}}\right)^{\frac{1}{q}} & q > 2 \\ (\Theta(1))^{\frac{1}{q}} & q < 2 \\ \Theta\left(\sqrt{\log(n)}\right) & q = 2 \end{cases}$$

$$= \begin{cases} \Theta\left(n^{1-\frac{2}{q}}\cdot\left(\frac{1}{q-2}\right)^{\frac{1}{q}}\right) = \Theta\left(n^{1-\frac{2}{q}}\right) & q > 2 \\ \Theta(1) & q < 2 \\ \Theta\left(\sqrt{\log(n)}\right) & q = 2 \end{cases}$$

We saw the bounds are tight, we will prove this claim algebraically again without the use of scaling distortion because I am not sure we are allowed to use the equation $l_q - dist(f) \leq \left(4\int_{\frac{1}{2\binom{n}{2}}}^{\frac{1}{2}} (\alpha(x))^q dx\right)^{\frac{1}{q}}$.

Define $C_k = |\{(v_i,v_j) \mid d_T(v_i,v_j) = k\}|$. Notice that $C_k = 2\cdot(n-k)$. (v_1 with v_{k+1} ... v_{n-k} with v_n). Notice that:

$$l_q - dist(f) = \left(\frac{\sum_{u\neq v\in X} dist_f(u,v)^q}{\binom{n}{2}}\right)^{\frac{1}{q}} = \left(\frac{2\cdot\sum_{i=1}^n\sum_{j=i+1}^n dist_f(v_i,v_j)^q}{\binom{n}{2}}\right)^{\frac{1}{q}}$$

$$= \left(\frac{2\cdot\sum_{i=1}^n\sum_{j=i+1}^n \max\left\{\frac{d_Y(v_i,v_j)}{d_X(v_i,v_j)}, \frac{d_X(v_i,v_j)}{d_Y(v_i,v_j)}\right\}^q}{\binom{n}{2}}\right)^{\frac{1}{q}}$$

$$= \left(\frac{2\cdot\sum_{i=1}^n\sum_{j=i+1}^n \max\left\{\frac{j-i}{\min\{j-i,n-(j-i)\}}, \frac{\min\{j-i,n-(j-i)\}}{j-i}\right\}^q}{\binom{n}{2}}\right)^{\frac{1}{q}}$$

$$= \left(\frac{2\cdot\sum_{i=1}^n\sum_{j=i+1}^n \left(\frac{j-i}{\min\{j-i,n-(j-i)\}}\right)^q}{\binom{n}{2}}\right)^{\frac{1}{q}} \leq \left(\frac{\sum_{k=1}^{n-1} C_k\cdot\left(\frac{k}{\min\{k,n-k\}}\right)^q}{\binom{n}{2}}\right)^{\frac{1}{q}}$$

$$= \left(\frac{\sum_{k=1}^{n-1} 2\cdot(n-k)\cdot\frac{k^q}{\min\{k,n-k\}^q}}{\binom{n}{2}}\right)^{\frac{1}{q}} \leq \left(\frac{\sum_{k=1}^{\frac{n}{2}} 2\cdot(n-k)\cdot\frac{k^q}{k^q} + \sum_{k=\frac{n}{2}}^{n-1} 2\cdot(n-k)\cdot\frac{k^q}{(n-k)^q}}{\binom{n}{2}}\right)^{\frac{1}{q}}$$

$$= \left(\frac{\sum_{k=1}^{\frac{n}{2}} 2\cdot(n-k) + 2\cdot\sum_{k=\frac{n}{2}}^{n-1}\frac{k^q}{(n-k)^{q-1}}}{\binom{n}{2}}\right)^{\frac{1}{q}}$$

Notice that $\forall q \geq 1$. $\frac{k^q}{(n-k)^{q-1}}$ is a monotonic function between $\frac{n}{2}, n$. For $q \neq 2$, we can bound it via the integral and get that $\int_{\frac{n}{2}}^{n}\frac{x^q}{(n-k)^{q-1}}dx \leq \int_{\frac{n}{2}}^{n}\frac{n^q}{(n-k)^{q-1}}dx = \frac{n^q\cdot(n-x)^{2-q}}{q-2}\Big|_{\frac{n}{2}}^{n} = O\left(\max\left\{n^2, n^q\right\}\right)$.

Therefore, for $1 \le q < 2$, we get that $l_q - dist(f) \le \left(\frac{O(n^2)+O(\max\{n^2,n^q\})}{\binom{n}{2}}\right)^{\frac{1}{q}} = O(1)^{\frac{1}{q}} = O(1)$

Therefore, for $q > 2$, we get that $l_q - dist(f) \le \left(\frac{O(n^2)+O(\max\{n^2,n^q\})}{\binom{n}{2}}\right)^{\frac{1}{q}} = O\left(n^{q-2}\right)^{\frac{1}{q}} = O\left(n^{1-\frac{2}{q}}\right)$

For $q = 2$, therefore we can bound it via the integral of $\int_{\frac{n}{2}}^{n} \frac{x^2}{(n-k)^{2-1}}dx \le \int_{\frac{n}{2}}^{n} \frac{n^2}{(n-k)^{2-1}}dx = -n^2 \cdot \ln(n-x) \mid_{\frac{n}{2}}^{n} =$ $\Theta\left(n^2 \cdot \log(n)\right)$. plugging it back we get that $l_2 - dist(f) \le \left(\frac{O(n^2)+O(n^2\log(n))}{\binom{n}{2}}\right)^{\frac{1}{2}} = O(\log(n))^{\frac{1}{2}} = O\left(\sqrt{\log(n)}\right)$.

Therefore, we got exactly what the we saw in the class in 9.8.

(b) The claim is not true. Let $G = K_n$, Denote $V_G = \{v_1, \ldots, v_n\}$ then $\forall i \in [n-1], (v_i, v_{i+1} \in E_T)$. Define $\boxed{a = 1, \varepsilon = 1}$ Let the weights $w(e) = a + \varepsilon \cdot \mathbb{1}_{e \notin E_T}$. Notice that the only MST in G if $\varepsilon > 0$ is $T = (V, E_T)$. (And that it is indeed a tree). Define $C_k = |\{(v_i, v_j) \mid d_T(v_i, v_j) = a \cdot k\}|$. Notice that $C_k = 2 \cdot (n-k)$. (v_1 with v_{k+1} ... v_{n-k} with v_n).

Notice that:

$$
\begin{aligned}
l_1 - dist(f) &= \left(\frac{\sum_{u \ne v \in X} dist_f(u,v)^1}{\binom{n}{2}}\right)^{\frac{1}{1}} = \left(\frac{2 \cdot \sum_{i=1}^{n} \sum_{j=i+1}^{n} dist_f(v_i, v_j)}{\binom{n}{2}}\right) \\
&= \left(\frac{2 \cdot \sum_{i=1}^{n} \sum_{j=i+1}^{n} \max\left\{\frac{d_Y(v_i,v_j)}{d_X(v_i,v_j)}, \frac{d_X(v_i,v_j)}{d_Y(v_i,v_j)}\right\}^q}{\binom{n}{2}}\right)^{\frac{1}{q}} \\
&= \left(\frac{2 \cdot \sum_{i=1}^{n} \sum_{j=i+1}^{n} \max\left\{\frac{d_X(v_i,v_j)}{a \cdot \min\{j-i, n-(j-i)\}}, \frac{a \cdot \min\{j-i, n-(j-i)\}}{d_X(v_i,v_j)}\right\}}{\binom{n}{2}}\right) \\
&\ge \left(\frac{2 \cdot \sum_{i=1}^{n} \sum_{j=i+1}^{n} \left(\frac{a \cdot \min\{j-i, n-(j-i)\}}{d_X(v_i,v_j)}\right)}{\binom{n}{2}}\right) \\
&\ge \left(\frac{2 \cdot \sum_{i=1}^{n} \sum_{j=i+1}^{n} \left(\frac{a \min\{j-i, n-(j-i)\}}{a+\varepsilon}\right)}{\binom{n}{2}}\right) = \frac{1}{a+\varepsilon}\left(\frac{\sum_{k=1}^{n-1} C_k \cdot a \cdot k}{\binom{n}{2}}\right) \\
&= \frac{a}{a+\varepsilon}\left(\frac{\sum_{k=1}^{n-1} 2 \cdot (n-k) \cdot k}{\binom{n}{2}}\right) = \frac{a}{a+\varepsilon} \frac{\Theta(n^3)}{\binom{n}{2}} = \Theta\left(\frac{n}{1+\frac{\varepsilon}{a}}\right) = \Theta(n)
\end{aligned}
$$

Therefore, we get that $l_1 - dist(f) = \Omega(n) \ne O(1)$, as required.

Now we will prove that not all MST have same average distortion. Let $G = K_n$, and $w(e) = 1, \forall e \in E$. First notice that that it is the same as what we proved before for $a = 1, \varepsilon = 0$, and we proved that for $T = (V, E_T)$ we have $l_1 - dist(f) = \Omega\left(\frac{n}{1+\varepsilon}\right) = \Omega(n)$.

Notice that if we define $T' = \langle V, E = \{\{v_1, v_i\} \mid 2 \le i \le n\}\rangle$. Notice that T' is also a MST.

Notice that f is non contracting and $\forall x, y \in V, dist_f(x,y) = \frac{d_Y(x,y)}{d_X(x,y)} \le \frac{2}{1} = 2$.

Therefore $dist(f) \le 2$ and therefore, $l_1 - dist(f) \le dist(f) = 2 = O(1)$.

So the best tree and the worst MST tree are not behaving the same for all graphs.

(c) We will use the same construction as in the previous question for $\boxed{a = c+2, \varepsilon = c+2}$. Notice that only tree that is of weight at most $w(MST) + c$ is $T = (V, E_T)$ (choosing a different edge will add at least $c+1$). Notice that we proved that $l_1 - dist(f) = \Omega\left(\frac{n}{1+\frac{\varepsilon}{a}}\right) = \Omega(n)$, therefore, $l_1 - dist(f) \ne O(1)$. Therefore, the claim is incorrect for tree of weight $w(MST) + c$.

We will use the same construction as in the previous question for $\boxed{a = 1, \varepsilon = 2 + n \cdot (\gamma)}$. Notice that only tree that is of weight at most $(1+\gamma) \cdot w(MST)$ is $T = (V, E_T)$ (the weight of MST is $n-1$ and if we choose a different edge we will get at least $(1+\gamma)w(MST) + 1$).

Notice that we proved that $l_1 - dist(f) = \Omega\left(\frac{n}{1+\frac{\gamma \cdot n+1}{1}}\right) = \Omega\left(\frac{1}{\gamma}\right)$, therefore, $l_1 - dist(f) \ne O(1)$. Therefore, the claim is incorrect for tree of weight $(1+\gamma) \cdot w(MST)$.

2. (a) Let $c = 4$, wlog assume that $\log_2(s) \in \mathbb{N}$, define $v_1, \ldots, v_{2 \cdot s} = \{0,1\}^{\log_2(s)}$. Define $height(U)$ to be the height of the tree.

We will follow the direction and prove via induction on $height(U)$ that $\exists f : U \to l_\infty^{O(\log(s))}$ s.t. $\forall x \neq y \in Leaves(U)$, $\left(1-\frac{c}{k}\right) d_U(x,y) \le \|f(x)-f(y)\|_\infty \le \left(1+\frac{c}{k}\right) d_U(x,y)$ and s.t. $\|f(x)\|_\infty \le Diam(U)$.

Base case: $height(U)=0$, notice that $|U|=1$ and we will define $f(x)=0$, it holds because there are no pairs.

Base case 2: $height(U)=1$, first notice that $|U|\le s$ because the maximal degree is s, notice that the distance between all the points is equal and therefore we have at most s-point equilateral space which we proved in exercise 1 question 2.a can be embedded isometrically into $l_\infty^{O(\log(s))}$.

Induction: Assume the theorem is true forall $height(U') < height(U)$ and we will prove that it is true for $diam(U)$.

Denote by v the root of the tree, it has at most s children, which we will denote as $u_1,\dots,u_l$. $\forall i \in [l]$, define U_i to be the sub tree of u_i. Notice that $height(U_i) < height(U)$ therefore, $\exists f_i : U_i \to l_\infty^{O(\log(s))}$ s.t. $\forall x \neq y \in Leaves(U_i)$, $\left(1-\frac{c}{k}\right) d_U(x,y) \le \|f(x)-f(y)\|_\infty \le \left(1+\frac{c}{k}\right) d_U(x,y)$ and s.t. $\|f_i(x)\|_\infty \le Diam(U_i)$.

Define $\boxed{\alpha = \frac{(k-1)\,Diam(U)}{k}}$

We will define f in the following way: $f(u) = \sum_{i=1}^{l} \mathbb{1}_{u\in U_i}\cdot [v_i\cdot\alpha + f_i(u)]$.

Due to $U_1,\dots,U_l$ having no intersection we can deduce that $\forall i\in[l], \forall x \in U_i, f(x) = f_i(x)+\alpha\cdot v_i$ and $f(v)=0$.

Notice that $\Delta(v) = Diam(U)$ and therefore $\forall i \in [l], \forall x \in Leaves(U_i)$:

$$\begin{aligned}\|f(x)\|_\infty &= \|f_i(x)+v_i\cdot\alpha\|_\infty \le \|f_i(x)\|_\infty + \|v_i\cdot\alpha\|_\infty \le Diam(U_i) + 1\cdot\alpha \\ &\overset{k-HST}{\le} \frac{Diam(U)}{k} + \alpha = \frac{1+(k-1)}{k}\cdot Diam(U) = Diam(U)\end{aligned}$$

therefore we proved that $\forall x \in Leaves(U), \|f(x)\|_\infty \le Diam(U)$.

Notice that $\forall i \in [l], \forall x,y \in Leaves(U_i)$:

$$\|f(x)-f(y)\|_\infty = \|(f_i(x)+\alpha\cdot v_i) - (f_i(y)+\alpha\cdot v_i)\|_\infty = \|f_i(x)-f_i(y)\|_\infty$$

.

Therefore, due to induction assumption, we can conclude that $\forall i \in [l], \forall x,y \in Leaves(U_i)$:

$$\left(1-\frac{c}{k}\right) d_U(x,y) = \left(1-\frac{c}{k}\right) d_{U_i}(x,y) \le \|f_i(x)-f_i(y)\|_\infty = \|f(x)-f(y)\|_\infty \le \left(1+\frac{c}{k}\right) d_{U_i}(x,y) = \left(1+\frac{c}{k}\right) d_U(x,y)$$

Therefore, due to induction assumption, we can conclude that $\forall i \in [l], \forall j \neq i \in [l], \forall x \in Leaves(U_i), y \in Leaves(U_j)$:

$$\begin{aligned}\|f(x)-f(y)\|_\infty &= \|(f_i(x)+\alpha\cdot v_i) - (f_j(y)+\alpha\cdot v_j)\|_\infty \\ &= \|(f_i(x)-f_j(y)) + \alpha\cdot(v_i-v_j)\|_\infty \le \|f_i(x)\|_\infty + \|f_j(y)\|_\infty + \alpha\cdot\|(v_i-v_j)\|_\infty \\ &\le Diam(U_i) + Diam(U_j) + \alpha \\ &\overset{k-HST}{\le} \frac{Diam(U)}{k} + \frac{Diam(U)}{k} + \alpha = Diam(U) + \frac{2+(k-1)}{k} \le \left(1+\frac{c}{k}\right)\cdot d_U(x,y)\end{aligned}$$

Therefore, due to induction assumption, we can conclude that $\forall i \in [l], \forall j \neq i \in [l], \forall x \in Leaves(U_i), y \in Leaves(U_j)$:

$$\begin{aligned}\|f(x)-f(y)\|_\infty &= \|(f_i(x)+\alpha\cdot v_i) - (f_j(y)+\alpha\cdot v_j)\|_\infty \\ &= \|(f_i(x)-f_j(y)) + \alpha\cdot(v_i-v_j)\|_\infty \ge \alpha\cdot\|(v_i-v_j)\|_\infty - \|f_i(x)-f_j(y)\|_\infty \\ &\ge \alpha\cdot\|(v_i-v_j)\|_\infty - \|f_i(x)\|_\infty - \|f_j(y)\|_\infty \\ &\ge \alpha - Diam(U_i) - Diam(U_j) \\ &\overset{k-HST}{\ge} \alpha - \frac{Diam(U)}{k} - \frac{Diam(U)}{k} \\ &= Diam(U)\frac{(k-1)-1-1}{k} \ge \left(1-\frac{c}{k}\right)\cdot d_U(x,y)\end{aligned}$$

Therefore, we proved that $\forall i \in [l], \forall j \neq i \in [l], \forall x \in Leaves(U_i), y \in Leaves(U_j)$, $\left(1-\frac{c}{k}\right)\cdot d_U(x,y) \le \|f(x)-f(y)\|_\infty \le \left(1+\frac{c}{k}\right)\cdot d_U(x,y)$.

We proved for all cases and therefore, $\forall x,y \in Leaves(U)$, $\left(1-\frac{c}{k}\right)\cdot d_U(x,y) \le \|f(x)-f(y)\|_\infty \le \left(1+\frac{c}{k}\right)\cdot d_U(x,y)$.

Notice we proved all the induction hypothesis and therefore the theorem is true. Therefore we know that $\forall(U,d)$ on a $k-HST$ for $k>1$ there exists an embedding $f : U \to l_\infty^{O\log(s))}$ s.t. $\forall x,y \in Leaves(U)$, $\left(1-\frac{c}{k}\right)\cdot d_U(x,y) \le \|f(x)-f(y)\|_\infty \le \left(1+\frac{c}{k}\right)\cdot d_U(x,y)$.

Notice that $expans_f(x,y) = \frac{\|f(x)-f(y)\|_\infty}{d_U(x,y)} \leq \frac{\left(1+\frac{c}{k}\right)\cdot d_U(x,y)}{d_U(x,y)} = 1 + \frac{c}{k}$. Therefore $expans(f) \leq 1 + \frac{c}{k}$.

Notice that $contr_f(x,y) = \frac{d_U(x,y)}{\|f(x)-f(y)\|_\infty} \leq \frac{d_U(x,y)}{\left(1-\frac{c}{k}\right)\cdot d_U(x,y)} = \frac{1}{1-\frac{c}{k}}$. Therefore $contr(f) \leq \frac{1}{1-\frac{c}{k}}$.

Therefore $dist(f) = contr(f) \cdot expans(f) \leq \frac{1+\frac{c}{k}}{1-\frac{c}{k}} = 1 + \frac{2\frac{c}{k}}{1-\frac{c}{k}} = 1 + \frac{2c}{k-c} = 1 + O\left(\frac{1}{k}\right)$.

Therefore we found an embedding $f : U \to l_\infty^{O(\log(s))}$ s.t. $dist(f) = 1 + O\left(\frac{1}{k}\right)$ as required (note that the metric is only defined om leaves).

Note: we can prove the same claim for all nodes and not only the leaves if we choose $l_\infty^{\log(s)+1}$ which is a stronger claim than k-HST distortion if we embed v as a vertex instead of 0.

(b) First, as stated in class, exists an embedding f from ultrametric to $k-HST$ with distortion k, denote the new metric space Y when $k > 2$. As we proved in the first exercise (3.a), $\dim(Y) = \dim(f(X)) = \dim(X) \cdot O(\log(k))$. denote s the maximal degree in the tree in Y. we will prove that $\log(s) = \Theta(\dim(Y))$.
Let $v \in Leaves(Y)$ and $r > 0$, define $B_{v,r} = \left\{u \mid \Delta(u) = \max_{u' \text{ is parent of } v \wedge \Delta(u') \leq r} \Delta(u')\right\}$. Assume by contradiction that $|B_v| > 1$, therefore $\exists u_1 \neq u_2 \in B_v$. Due to the definition of $B_{v,r}$, it follows that $\Delta(u_1) = \Delta(u_2)$. wlog u_1 is a parent of u_2, therefore $\Delta(u_2) \leq \frac{\Delta(u_1)}{k} < \Delta(u_1) = \Delta(u_2)$. We got a contradiction and therefore $|B_{v,r}| = 1$. Therefore, $|B_v| = 1$, i.e. the argmax is well defined. Define $u = \operatorname{argmax}_{u' \text{ is parent of } v \wedge \Delta(u') \leq r} \Delta(u')$.
Firstly, $\forall u' \in Leaves(Y)$, if $LCA(u,u') = u$, then $d(v,u') = \Delta(LCA(v,u')) \leq \Delta(LCA(u,u')) = \Delta(u) \leq r$, therefore $u' \in B(v,r)$. Therefore $\boxed{\{u' \in Leaves(Y) \mid LCA(u,u') = u\} \subseteq B(v,r)}$.
Now, if $u' \notin B(v,r)$, Assume by contradiction that $\Delta(LCA(u,u')) \neq u$, Then $LCA(u,u')$ is a parent of u and we get a contradiction to the maximality of u because $\Delta(u) \overset{k-HST}{<} \Delta(LCA(u,u')) \leq r$. Notice that $B(v,r)$ contains only leaves, therefore $u' \in B(v,r) \implies u' \in \{' \in Leaves(Y) \mid LCA(u,u') = u\}$. Therefore $B(v,r) \subseteq \{u' \in Leaves(Y) \mid LCA(u,u') = u\}$, which means $\boxed{\{u' \in Leaves(Y) \mid LCA(u,u') = u\} = B(v,r)}$.
Notice that if u is a leaf then $B(v,r) = \{v\} = \{u\}$ and it can be covered by 1 balls.
Else denote $u_1, \ldots, u_l$ the direct children of u, and denote $u'_1, \ldots, u'_l$ s.t. u'_i is a leaf of the sub tree created by u_i. Let $i \in [l]$ and u' be a leaf of u_i, therefore

$$d(u', u'_i) = \Delta(LCA(u', u'_i)) \overset{k-HST}{\leq} \Delta(LCA(u', u_i)) = \Delta(u_i) \overset{k-HST}{\leq} \frac{\Delta(u)}{k} \leq \frac{r}{k} \leq \frac{r}{2}$$

Therefore $u' \in B(u'_i, \frac{r}{2})$. i.e. $\{u' \in Leaves(Y) \mid u' \text{ is a child of } u_i\} \subseteq B(u'_i, \frac{r}{2})$. Therefore

$$\begin{aligned} B(v,r) &= \{u' \in Leaves(Y) \mid LCA(u,u') = u\} = \{u' \in Leaves(Y) \mid \exists 1 \leq i \leq l \to LCA(u,u'_i) = u_i\} \\ &= \bigcup_{i=1}^{l} \{u' \in Leaves(Y) \mid LCA(u,u'_i) = u_i\} \subseteq \bigcup_{i=1}^{l} B\left(u'_i, \frac{r}{2}\right) \end{aligned}$$

Therefore, we showed that the leaves of the tree can be covered by at most l balls of radius $\frac{r}{2}$ and $l \leq s$ by definition and therefore $\boxed{\dim(f(X)) = O(\log(s))}$.

Due to what we proved it the first part, exists $g : k - HST \to l_\infty^{O(\log(s))}$ with distortion $1 + O(\frac{1}{k})$. we choose $\boxed{k = 7}$.
Notice that $\log(s) = O(\dim(f(X))) = O(\dim(X) \cdot O(\log(k))) = O(\dim(X) \cdot O(1)) = O(\dim(X))$, therefore exists $g : 7 - HST \to l_\infty^{O(\dim(X))}$ with distortion $1 + O(\frac{1}{7})$.
Notice that distortion is bounded by multiplication and therefore, if we define $h(x) = g(f(x))$ we will get an embedding to $l_\infty^{O(\dim(X))}$ with distortion at most $7 \cdot (1 + O(\frac{1}{7})) = O(1)$, as required.

3. (a) We will prove by induction on n that we can embed L_n with distortion α into $k - HST$ with $k > 2$ core $\geq n^{g(\alpha)}$.
Base case: $n = 1$, embed it as the root and it holds by definition.
Base case: $n = 2$, it can be embedded to a tree with root and a child and the core it n with distortion 1.
Induction: Assume that the theorem holds for $j < n$ and we will prove for n.

Denote $h(\alpha) = \frac{1}{2^{\frac{1}{g(\alpha)}}} = 2^{-\frac{1}{g(\alpha)}}$. Define $\boxed{l = \left(\frac{n^{g(\alpha)}}{2}\right)^{\frac{1}{g(\alpha)}} = \frac{n}{2^{\frac{1}{g(\alpha)}}} = n \cdot h(\alpha)}$.

Notice that

$$\frac{n}{n-2l} \le \alpha \iff 1+\frac{2l}{n-2l} \le \alpha \iff 2l \le (n-2l)\cdot(\alpha-1)$$
$$\iff 0 \le n-2l-\frac{2l}{\alpha-1} \iff 0 \le n-2l\left(1+\frac{1}{\alpha-1}\right) \iff 0 \le n-2n\cdot h(\alpha)\left(\frac{\alpha}{\alpha-1}\right)$$
$$\iff 0 \le 1-2h(\alpha)\left(\frac{\alpha}{\alpha-1}\right) \iff 2h(\alpha)\left(\frac{\alpha}{\alpha-1}\right) \le 1$$
$$\iff 2h(\alpha) \le \frac{\alpha-1}{\alpha} \iff 2\cdot 2^{-\frac{1}{g(\alpha)}} \le \frac{\alpha-1}{\alpha} \iff 2^{1-\frac{1}{g(\alpha)}} \le \frac{\alpha-1}{\alpha}$$
$$\iff 2^{\frac{g(\alpha)-1}{g(\alpha)}} \le \frac{\alpha-1}{\alpha} \iff \frac{g(\alpha)-1}{g(\alpha)} \le \log_2\left(\frac{\alpha-1}{\alpha}\right)$$
$$\iff 1-\frac{1}{g(\alpha)} \le \log_2\left(\frac{\alpha-1}{\alpha}\right) \iff 1-\log_2\left(\frac{\alpha-1}{\alpha}\right) \le \frac{1}{g(\alpha)} \iff 1-\log_2\left(1-\frac{1}{\alpha}\right) \le \frac{1}{g(\alpha)}$$
$$\iff g(\alpha) \le \frac{1}{1-\log_2\left(1-\frac{1}{\alpha}\right)} \iff g(\alpha) \le 1+\frac{\log_2\left(1-\frac{1}{\alpha}\right)}{1-\log_2\left(1-\frac{1}{\alpha}\right)}$$

Notice that $\lim_{x\to\infty} x\cdot\log_2\left(1-\frac{1}{x}\right) = \lim_{x\to\infty}\frac{\log_2\left(1-\frac{1}{x}\right)}{\frac{1}{x}} = \lim_{x\to 0^+}\frac{\log_2(1-x)}{x} \overset{Lupital}{=} \lim_{x\to 0^+}\frac{\frac{-1}{1-x}\ln(2)}{1} = -\ln(2)$

Therefore $\lim_{\alpha\to\infty}\frac{\frac{\log_2\left(1-\frac{1}{\alpha}\right)}{1-\log_2\left(1-\frac{1}{\alpha}\right)}}{\frac{1}{\alpha}} = \frac{-\ln(2)}{1}$

Therefore, $-\frac{\log_2\left(1-\frac{1}{\alpha}\right)}{1-\log_2\left(1-\frac{1}{\alpha}\right)} = \Theta\left(\frac{1}{\alpha}\right)$. (similarly we can show for $\alpha \to 1$).

Therefore, we got that $\frac{n}{n-2l} \le \alpha \iff g(\alpha) \le 1+\frac{\log_2\left(1-\frac{1}{\alpha}\right)}{1-\log_2\left(1-\frac{1}{\alpha}\right)} \iff g(\alpha) \le 1-\Theta\left(\frac{1}{\alpha}\right)$.

We know that $g(\alpha) = 1-O\left(\frac{1}{\alpha}\right)$ and therefore $\frac{n}{n-2l} \le \alpha$.

Notice that $l < \frac{n}{2}$ which means that the following definition doesn't include any element twice.

As the proof in the lecture, we will define $P = \{v_1,\ldots,v_l\}, \overline{Q} = \{v_{n-l},\ldots,v_n\}$. By the hypothesis lemma, it holds that $\exists f_1 : P \to k-HST$ with distortion α and core $|P|^{g(\alpha)}$ and $\exists f_2 : \overline{Q} \to k-HST$ with distortion α and core $\left|\overline{Q}\right|^{g(\alpha)}$. Add a node r which will be the root, its children will be $T_P, T_{\overline{Q}}$ and $\boxed{\Delta(r) = n}$. The new tree is the embedding f (and add leafs for $L_n \setminus (P\cup\overline{Q})$.

Notice that $\forall x,y \in S(P)$, $d_T(x,y) = d_{T_P}(x,y)$ and therefore $dist_f(x,y) = dist_{f_1}(x,y) \le \alpha$ and $|S(P)| \ge |P|^{g(\alpha)}$

Notice that $\forall x,y \in S(\overline{Q})$, $d_T(x,y) = d_{T_{\overline{Q}}}(x,y)$ and therefore $dist_f(x,y) = dist_{f_2}(x,y) \le \alpha$ and $\left|S(\overline{Q})\right| \ge \left|\overline{Q}\right|^{g(\alpha)}$.

Notice that

$$|S(P)| + \left|S(\overline{Q})\right| \ge |P|^{g(\alpha)} + \left|\overline{Q}\right|^{g(\alpha)} = l^{g(\alpha)} + l^{g(\alpha)} = 2\cdot l^{g(\alpha)} = 2\cdot\frac{n^{g(\alpha)}}{2} = n^{g(\alpha)}$$

Now notice that $\forall x \in P, y \in \overline{Q}$, $d_X(x,y) \le n$. Notice that $\forall x \in P, y \in \overline{Q}, contr_f(x,y) = \frac{d_X(x,y)}{\Delta(r)} \le \frac{n}{\frac{n}{2}} = 2$.

Denote $g(\alpha) \le 1-c\cdot\frac{1}{\alpha}$ for some $c > 0$

Now notice that $\forall x \in P, y \in \overline{Q}$,

$$d_X(x,y) \ge (n-l)-l = n-2\cdot l$$

Notice that $\forall x \in P, y \in \overline{Q}$

$$expans_f(x,y) = \frac{\Delta(r)}{d_X(x,y)} \le \frac{n}{n-2l} \le \alpha$$

Notice that

$$contr_f(x,y) = \frac{d_X(x,y)}{\Delta(r)} \le \frac{n}{n} \le 1$$

Define $S(X) = S(P)\cup S(\overline{Q})$. We proved before that $|S(X)| = n^{g(\alpha)}$ and we can conclude that $\forall x,y \in S(P)\cup S(\overline{Q})$, $dist_f(x,y) \le \alpha$. (showed all cases). Therefore, the induction is done and we proved what we needed to proved.

Note: we embedded into $k-HST$ with $k = 2\cdot\frac{1}{h(\alpha)} \ge 2$ because the root is at distance n and by induction, the previous roots are at distance $\frac{n}{2}\cdot h(\alpha)$.

(b) Define $t = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{24\ln(2)}$. Let $Z \subseteq \binom{X}{2}$ s.t. $|Z| = n^{t\cdot\varepsilon^2}$ and $\forall x, y \in X, \mathbb{P}(x \in Z) = \mathbb{P}(y \in Z)$. i.e. Z is a random group of size $n^{t\cdot\varepsilon^2}$.

Let $x, y \in Z$, notice that $x, y \sim Uniform(X) = Uniform(H_d)$. Therefore we can write it as $\forall i \in [d], [x]_i, [y]_i \in Uniform(\{0,1\})$.

Denote $X_{i,x,y} = |[x]_i - [y]_i|, X_{x,y} = \sum_{i=1}^{d} X_i$. Notice that $X_{i,x,y} = |[x]_i - [y]_i| \sim Uniform(\{0,1\})$. Therefore

$$\mathbb{E}[X_{x,y}] = \mathbb{E}\left[\sum_{i=1}^{d} X_{i,x,y}\right] = \mathbb{E}\left[\sum_{i=1}^{d} |[x]_i - [y]_i|\right] = \sum_{i=1}^{d} \mathbb{E}[|[x]_i - [y]_i|] = \sum_{i=1}^{d} \frac{1}{2} = \frac{d}{2}$$

Notice that $\forall i \neq j, X_{i,x,y}, X_{j,x,y}$ are independent variables and $X_{i,x,y} \sim Ber\left(\frac{1}{2}\right)$ and therefore the Chernoff inequality holds for $X_{x,y}$, therefore

$$\mathbb{P}\left(|X_{x,y} - \mathbb{E}[X_{x,y}]| \geq \frac{\varepsilon}{2} \cdot \mathbb{E}[X_{x,y}]\right) \leq 2 \cdot \exp\left\{-\frac{\left(\frac{\varepsilon}{2}\right)^2}{3} \cdot \mathbb{E}[X_{x,y}]\right\} = 2 \cdot \exp\left\{-\frac{\varepsilon^2}{12} \cdot \frac{d}{2}\right\} = 2 \cdot \exp\left\{-\frac{\varepsilon^2 \cdot d}{24}\right\}$$

$$= 2 \cdot \exp\left\{-\frac{\varepsilon^2 \cdot \log_2(n)}{24}\right\} = 2 \cdot \exp\left\{-\frac{\varepsilon^2 \cdot \frac{\ln(n)}{\ln(2)}}{24}\right\} = 2 \cdot n^{-\frac{\varepsilon^2}{24\cdot\ln(2)}}$$

Now notice that

$$\begin{aligned}
\mathbb{P}\left[\exists x' \neq y' \in Z, \left|\|x' - y'\|_1 - \frac{d}{2}\right| \geq \frac{\varepsilon}{2} \cdot \frac{d}{2}\right] &\leq \sum_{x' \neq y' \in Z} \mathbb{P}\left[\left|\|x' - y'\|_1 - \frac{d}{2}\right| \geq \frac{\varepsilon}{2} \cdot \frac{d}{2}\right] \\
&= \sum_{x' \neq y' \in Z} \mathbb{P}\left[\left|\sum_{i=1}^{d} |[x']_i - [y']_i| - \frac{d}{2}\right| \geq \frac{\varepsilon}{2} \cdot \frac{d}{2}\right] \\
&= \sum_{x' \neq y' \in Z} \mathbb{P}\left[\left|\sum_{i=1}^{d} X_{i,x',y'} - \frac{d}{2}\right| \geq \frac{\varepsilon}{2} \cdot \frac{d}{2}\right] \\
&= \sum_{x' \neq y' \in Z} \mathbb{P}\left[|X_{x',y'} - \mathbb{E}[X_{x',y'}]| \geq \frac{\varepsilon}{2} \cdot \mathbb{E}[X_{x',y'}]\right] \\
&\leq \sum_{x' \neq y' \in Z} 2 \cdot n^{\left\{\frac{\varepsilon^2}{24\cdot\ln(2)}\right\}} = \binom{|Z|}{2} 2 \cdot n^{-\frac{\varepsilon^2}{24\cdot\ln(2)}} \\
&\leq |Z|^2\, 2 \cdot n^{-\frac{\varepsilon^2}{24\cdot\ln(2)}} \leq n^{2t\cdot\varepsilon^2} \cdot 2n^{-\frac{\varepsilon^2}{24\cdot\ln(2)}} = 2 \cdot n^{-\frac{1}{2}\frac{\varepsilon^2}{24\cdot\ln(2)}}
\end{aligned}$$

Therefore, $\lim_{n\to\infty} \mathbb{P}\left[\exists x' \neq y' \in Z, \left|\|x' - y'\|_1 - \frac{d}{2}\right| \geq \frac{\varepsilon}{2} \cdot \frac{d}{2}\right] = 0$,

Therefore, for big enough n, $\mathbb{P}\left[\exists x' \neq y' \in Z, \left|\|x' - y'\|_1 - \frac{d}{2}\right| \geq \frac{\varepsilon}{2} \cdot \frac{d}{2}\right] < 1$.

Therefore,

$$\mathbb{P}\left[\forall x' \neq y' \in Z, \left|\|x' - y'\|_1 - \frac{d}{2}\right| < \frac{\varepsilon}{2} \cdot \frac{d}{2}\right] = 1 - \mathbb{P}\left[\exists x' \neq y' \in Z, \left|\|x' - y'\|_1 - \frac{d}{2}\right| \geq \frac{\varepsilon}{2} \cdot \frac{d}{2}\right] > 0$$

Therefore, we showed that exists a group Z s.t. $\forall x, y \in Z, \frac{d}{2} - \frac{\varepsilon}{2} \cdot \frac{d}{2} \leq \|x' - y'\|_1 \leq \frac{d}{2}(1 + \frac{\varepsilon}{2})$ and $|Z| = n^{t\cdot\varepsilon^2} = n^{\Omega(\varepsilon^2)}$ forall $0 < \varepsilon < 1$.

Let $\delta = \frac{\varepsilon}{3}$, from what we proved we know that there exists a group Z_δ s.t. $\forall x, y \in Z, \frac{d}{2} - \frac{\delta}{2} \cdot \frac{d}{2} \leq \|x' - y'\|_1 \leq \frac{d}{2}(1 + \frac{\delta}{2})$ and $|Z| = n^{\Omega(\delta^2)} = n^{\Omega(\varepsilon^2)}$.

Denote $X = \{x_1, \ldots, x_n\}$ Now, we will embed in $\mathbb{R}^n$ with l_1 loss and the embedding will be $\forall i \in [n], f(x_i) = \left(\frac{d}{2} - \frac{\delta}{2}\right) \cdot e_i$. Notice that the space is an equilateral metric space and that $\forall x \neq y \in Z_\delta$ it holds that

$$contr_f(x, y) = \frac{d_X(x, y)}{d_Y(x, y)} = \frac{\|x - y\|_1}{\frac{d}{2}\left(1 - \frac{\delta}{2}\right)} \leq \frac{\frac{d}{2}(1 + \frac{\delta}{2})}{\frac{d}{2}\left(1 - \frac{\delta}{2}\right)} = 1 + \frac{\delta}{1 - \frac{\delta}{2}}$$

Notice that

$$\frac{\delta}{1 - \frac{\delta}{2}} \leq \varepsilon \iff \delta \leq \varepsilon - \frac{\varepsilon}{2}\delta \iff \delta \cdot \left(1 + \frac{\varepsilon}{2}\right) \leq \varepsilon \iff \delta \leq \frac{\varepsilon}{1 + \frac{\varepsilon}{2}}$$

$$\iff \frac{\varepsilon}{3} \leq \frac{\varepsilon}{1 + \frac{\varepsilon}{2}} \iff 1 + \frac{\varepsilon}{2} \leq 3$$

We know that $\varepsilon < 1$ and therefore $1 + \frac{\varepsilon}{2} \leq 3$, and therefore $\frac{\delta}{1-\frac{\delta}{2}} \leq \varepsilon$. Therefore $\forall x, y \in Z_\delta, contr_f(x,y) \leq 1 + \frac{\delta}{1-\frac{\delta}{2}} \leq 1 + \varepsilon$ and

$$expans_f(x,y) = \frac{d_Y(x,y)}{d_X(x,y)} = \frac{\frac{d}{2} - \frac{\delta}{2}}{\|x-y\|_1} \leq \frac{\frac{d}{2} - \frac{\delta}{2}}{\frac{d}{2} - \frac{\delta}{2}} = 1$$

Therefore, $\forall x, y \in Z, expans_f(x,y) \leq 1$.
Therefore, f is non expanding, therefore $\forall x, y \in Z_\delta, dist_f(x,y) = contr_f(x,y) \leq 1 + \varepsilon$.
Therefore, (f, Z_δ) is a partial embedding with distortion $1 + \varepsilon$.
Therefore, f is a Ramsey embedding with distortion $1 + \varepsilon$ and core size $|Z_\delta| = n^{\Omega(\varepsilon^2)}$, as required.