In the previous lecture we have shown that if there exists a (special) bundle of probabilistic partitions with padding parameter $\gamma(\mathcal{H})$, then we can construct a probabilistic embedding into a distribution of ultrametrics with expected distortion $O(k \cdot \gamma(\mathcal{H}))$.

In today's lecture we show how to build such a special bundle, for a given n-point metric space, with $\gamma(\mathcal{H}) = O(\log n)$.

11.1 ABN Partitions and Improved Distortion of Prob. Embedding into Trees

In this section we develop another probabilistic partitions of the finite metric spaces. As we will see later these partitions allow to improve distortion guarantee of the probabilistic embedding into trees, and to construct a coarse scaling version of Bourgain's embedding.

11.1.1 Refined Probabilistic Partitions

We start with the necessary definitions:

Definition 11.1 (Local Growth Rate)**.** *Let $x \in X$ and $r > 0$. The local growth rate of x at radius r with parameter $\Gamma > 1$ is defined by $\rho_\Gamma(x, r) = \frac{|B(x,\Gamma r)|}{|B(x,r/\Gamma)|}$.*

In this lecture we will take $\Gamma = 64$, and we will omit the subscript Γ from the notation.

Following is the main partition result. The padding parameter of this partition is a function of $x \in X$, and of the cluster to which it belongs. More precisely, all the points that share the same cluster will have the same padding parameter - such function is called a uniform function over clusters, reflecting the name of the partition.

For a finite space X, $\Delta \leq diam(X)$, and any Δ-bounded probabilistic partition $\mathcal{P}$ of X, $\forall x \in X$, $\forall\ 0 < \delta \leq 1/2$, let $P(x)$ denote the cluster of the partition $\mathcal{P}$ that contains x. We define $\hat{\rho}(v, \Delta) = \max\{\rho(v, \Delta), 2\}$, and $v_{P(x)} = argmin_{u \in P(x)}\{\rho(u, \Delta)\}$. Then we define $\eta^{(\delta)}_{P(x)} = \frac{\delta}{128 \cdot \log(\hat{\rho}(v_{P(x)}, \Delta))}$.

Now we are ready to formulate the main partition lemma.

Theorem 11.1 (Uniform Partition Lemma[1])**.** *For any finite space X, for any Δ, there exists a Δ-bounded probabilistic partition $\mathcal{P}$ of X such that for all $x \in X$, for all $0 < \delta \leq 1/2$,*

$$Pr_{\mathcal{P}}\left[B\left(x, \eta^{(\delta)}_{P(x)} \cdot \Delta\right) \not\subseteq P(x)\right] \leq \delta.$$

Note the difference from the partition lemma we have seen in the last lecture, in which the balls we considered were all of the same radius $\eta^{(\delta)}\Delta = \frac{\delta}{c \log n}\Delta$, while here we consider balls of radii defined by the partition $P \in \mathcal{P}$ (i.e. these radii are the random variables, thus the property of the uniform partition does not immediately imply the padding parameter. To deduce a padding parameter we will work a bit harder right after the proof).

Proof. The partition will be a refined version of the probabilistic partition we developed in the last lecture. Let us remind the conceptual steps of this partition.

At each step j of the construction, the center $v_j \in X$ of the cluster C_j was chosen arbitrarily, from all the points of the space that are still not covered. The radius r_j of the cluster C_j was chosen randomly from the following distribution: the line of the real numbers is divided into intervals I_l of the length $\frac{\Delta}{c \log n}$, the interval is chosen with probability $1/2^l$, and the radius is chosen from the obtained interval with uniform probability. The proof of the required padding parameter, i.e. showing that: $Pr[B(x, \eta_n^{(\delta)}\Delta) \nsubseteq P(x)] = Pr[\exists\ bad\ C_j\ for\ B(x, \eta_n^{(\delta)}\Delta)] \leq \delta$, where $\eta_n^{\delta} = \frac{\delta}{c \log n}$, follows from the following lemma:

$\forall x \in X, \forall 0 < \delta < 1, \forall j \geq 1$

$$Pr[C_j\ is\ bad\ for B(x, \eta_n^{\delta}\Delta) \mid C_1 \ldots C_{j-1}\ neutral] \leq$$

$$\leq \delta \cdot Pr[C_j\ is\ good\ for B(x, \eta_n^{\delta}\Delta) | C_1 \ldots C_{j-1}\ neutral].$$

As a first step, replace n with some parameter $\chi \geq 2$, and slightly change the distribution to be the following: divide the interval $[\Delta/4, \Delta/2]$ by intervals I_l of lengths $\frac{\Delta}{8 \log \chi}$ (this is why we want $\chi \geq 2$), where we choose the interval I_l again with probability $1/2^l$, for all l, except the last interval, $I_{last} = I_{2\lceil \log \chi \rceil}$, which we choose with probability $2/\chi^2$.

Note that, if for some cluster C_j, $d(v_j, x) < \Delta/4 - \eta_n^{(\delta)}\Delta$, then, since the radius of every cluster is at least $\Delta/4$ (by the construction), this cluster will be good for the ball $B(x, \eta_n^{\delta}\Delta)$. Similarly, if for some cluster C_j, $d(v_j, x) > \Delta/2 + \eta_n^{(\delta)}$, then, since the radius of every cluster is at most $\Delta/2$, this cluster will be neutral for the ball $B(x, \eta_n^{\delta}\Delta)$. Namely, for such clusters the padding property trivially holds. Then, we assume without loss of generality, there are no such clusters in the partition. Divide all the indexes $1 \leq j \leq n$ of the clusters C_j, into two groups: $A = \{j \mid d(v_j, x) \notin I_{last}\}$, and $B = \{j \mid d(v_j, x) \in I_{last},\ or\ \Delta/2 \leq d(x, v_j) \leq \Delta/2 + \eta_n^{\delta}\Delta\}$. Define $\eta_\chi^{(\delta)} = \frac{\delta}{128 \cdot \log \chi}$. Therefore,

$$Pr[B\left(x, \eta_\chi^{(\delta)}\Delta\right) \nsubseteq P(x)] \leq Pr[\exists\ j\ s.t.\ C_j\ is\ bad\ for\ B\left(x, \eta_\chi^{(\delta)}\Delta\right)]$$
$$= Pr[\exists\ j \in A\ s.t.\ C_j\ is\ bad\ for\ B\left(x, \eta_\chi^{(\delta)}\Delta\right)] + Pr[\exists\ j \in B\ s.t.\ C_j\ is\ bad\ for\ B\left(x, \eta_\chi^{(\delta)}\Delta\right)].$$

Following the lines of the proof of Lemma 10.6, we conclude that the first probability is at most $\delta/2$ (you should go over the proof of the lemma again and see that it indeed implies the stated). To bound the second probability, we note that there are at most n clusters in a partition and thus by the union bound:

$$Pr[\exists\ j \in B\ s.t.\ C_j\ is\ bad\ for\ B\left(x, \eta_\chi^{(\delta)}\Delta\right)] \leq n \cdot Pr[C_j\ is\ bad\ for\ B\left(x, \eta_\chi^{(\delta)}\Delta\right),\ j \in B] \leq$$

$$\leq n \cdot \left(\frac{2}{\chi^2} + \frac{2}{\chi^2}\right) \delta/8 = \frac{n}{\chi^2}\delta/2.$$

Note that for $\chi = n$, we get the required, which implies a probabilistic partition with padding parameter $O(\log n)$.

Next, we change the process of the construction as follows: at each step $j \geq 1$, we choose the center v_j of the cluster $\hat{C}_j$ as the point that brings to the minimum $\hat{\rho}(u, \Delta)$ (while computing this value according to all the metric space X), among all the points $u \in X$ that are still uncovered (by the clusters $\hat{C}_1, \ldots, \hat{C}_{j-1}$). Denote this quantity by χ_j, and note that $\chi_j \geq 2$. The radius $\hat{r}_j$ for the cluster $\hat{C}_j$ is chosen according to the same process as before, except that the interval $[\Delta/4, \Delta/2]$ is divided into sub-intervals of lengths $\frac{\Delta}{8 \log \chi_j}$.

Let $\mathcal{P}$ be a random partition generated according to the above process, and let $x \in X$ be any point. Let $\hat{C}_l$ denote that the cluster of $\mathcal{P}$ that contains x. Let $\hat{r}_l = \eta_{C_l}^{(\delta)}\Delta = \frac{\delta}{128 \log(\hat{\rho}(v_l, \Delta))}\Delta$, then we have to show that $Pr_{\mathcal{P}}[B(x, \hat{r}_l) \nsubseteq P(x)] \leq \delta$. Note that

$$Pr_{\mathcal{P}}[B(x, \hat{r}_l) \nsubseteq P(x)] \leq Pr[\exists\ bad\ C_j\ , 1 \leq j \leq l,\ for\ B(x, \hat{r}_l)] \leq Pr[\exists\ bad\ C_j\ for\ B(x, \hat{r_j})],$$

where $\hat{r_j} = \eta_{C_j}^{(\delta)}\Delta$; the last inequity is true due to the monotonicity of the radii $\hat{r_j}$ (the radii decrease as j increases).

Therefore, applying the same considerations as before, we conclude that the above probability for the clusters $\hat{C}_j$, such that $j \in A$, is bounded by $\delta/2$. For the clusters $\hat{C}_j$ such that $j \in B$:

$$Pr\left[\exists\ j \in B\ s.t.\ C_j\ is\ bad\ for\ B\left(x, \eta_{C_j}^{(\delta)}\Delta\right)\right] \leq \sum_{j \in B} \frac{1}{{\chi_j}^2} \cdot \delta/2 \leq \sum_{j \in B} \frac{1}{\chi_j} \cdot \delta/2.$$

Thus, we have to show that $\sum_{j \in B} \frac{1}{\chi_j} \leq 1$. By the definition of χ_j: $\sum_{j \in B} \frac{1}{\chi_j} \leq \sum_{j \in B} \frac{|B\left(v_j, \frac{\Delta}{64}\right)|}{|B(v_j, 64\Delta)|}$. Note that for all $j \in B$ it holds that $B(x, 63\Delta) \subseteq B(v_j, 64\Delta)$, since $d(x, v_j) \leq \Delta/2 + \hat{r}_j \leq \Delta$. Therefore,

$$\sum_{j \in B} \frac{|B\left(v_j, \frac{\Delta}{64}\right)|}{|B(v_j, 64\Delta)|} \leq \sum_{j \in B} \frac{|B(v_j, \Delta/64)|}{|B(x, 63\Delta)|}.$$

Also, each $B(v_j, \Delta/64) \subseteq B(x, 63\Delta)$ In addition, for any $j_1 \neq j_2 \in B$, the balls $B(v_{j_1}, \Delta/64)$ and $B(v_{j_2}, \Delta/64)$ are disjoint, since $r_j \geq \Delta/4$. Therefore, $\sum_{j \in B} \frac{|B(v_j, \Delta/64)|}{|B(x, 63\Delta)|} \leq 1$, as required. This completes the proof of the theorem. □

Next we show that this lemma implies a probabilistic embedding of any n-point metric space into k-HST's, with expected distortion $O(k \cdot \log n)$. Recall that we just have to explain how to build a special bundle of probabilistic partitions $\mathcal{H}$, with $\gamma(\mathcal{H}) = O(\log n)$.

Claim 11.2. *Given any n-point metric space there exists a special bundle $\mathcal{H}$ of probabilistic partitions, such that $\gamma(H) = O(\log n)$.*

Proof. For all $i \geq 1$, $\Delta_i = \Delta/k^i$ bounded partition $\mathcal{P}_i$ is defined to be the Δ_i-bounded uniform partition of Theorem 11.1. We show that for all $i \geq 1$, the function defined by: for $x \in X$,

$$\gamma_i(x) = \begin{cases} 128 \log(\rho(x, \Delta_i)), & if\ \ \rho(x, \Delta_i) \geq 2 \\ 0, & otherwise. \end{cases}$$

is the padding parameter of the probabilistic partition $\mathcal{P}_i$. Namely, we show that $\gamma_i(\cdot)$ satisfies the following: $\forall x \in X$, $\forall \delta < 1$, if $\gamma_i(x) > 0$, then $\Pr\left[B\left(x, \frac{\delta}{\gamma_i(x)}\Delta_i\right) \nsubseteq P(X)\right] \leq \delta$; if $\gamma_i(x) = 0$, then $\forall\, 0 < r < \Delta_i/16,\ Pr[B(x,r) \nsubseteq P(x)] = 0$.

If so, if $\gamma_i(x) > 0$, then by the definition $\rho(x, \Delta_i) \geq 2$, and thus:

$$Pr\left[B\left(x, \frac{\delta}{128log(\rho(x,\Delta_i))}\right) \nsubseteq P(x)\right] \leq Pr\left[B\left(x, \frac{\delta}{128log(\hat{\rho}(v_{P(x)},\Delta_i))}\right) \nsubseteq P(x)\right] \leq \delta,$$

since by the construction of the partition, $\rho(x, \Delta_i) \geq \hat{\rho}(v_{P(x)}, \Delta_i)$.

If $\gamma_i(x) = 0$, then by the definition $\rho(x, \Delta_i) < 2$. We first prove the following small lemma:

Lemma 11.3. *Let $x \neq y \in X$ and $\Gamma > 2$. If $(2/\Gamma)\Delta < d(x,y) < (\Gamma - 1/\Gamma)\Delta$, then $\max\{\rho_\Gamma(x,\Delta), \rho_\Gamma(y,\Delta)\} \geq 2$.*

Proof. Let $a = |B(x, \Delta/\Gamma)|$, and $b = |B(y, \Delta/\Gamma)|$. Note that two this balls are disjoint since $\frac{2}{\Gamma}\Delta < d(x,y)$. Assume w.l.o.g. that $b = \min\{a,b\}$. It is easy to see that $B(x,\Delta/\Gamma) \subseteq B(y, \Gamma\Delta)$, since $d(x,y) \leq \left(\Gamma - \frac{1}{\Gamma}\right)\Delta$. Therefore $\rho(y,\Delta) \geq (a+b)/b \geq 2$. □

Now we continue the proof of Claim 11.2. If $\rho(x, \Delta_i) < 2$, then for all $0 < r < \Delta_i/16$, it holds that: $Pr[B(x,r) \nsubseteq P(x)] \leq Pr[B(x, \Delta_i/16) \nsubseteq P(x)]$. Consider the clusters $C_j = B(v_j, r_j)$, in order of the construction, starting from the first and considering the next only if the previous clusters are neutral for $B(x, \Delta_i/16)$. We prove that each cluster is either neutral or good.

1. If $d(v_j, x) > \Delta_i/8$, assume by contradiction that $B(x, \Delta_i/16) \cap C_j \neq \emptyset$. Since by construction the maximum radius of a cluster is $\Delta_i/2$, we have that $d(v_j, x) \leq \Delta_i/2 + \Delta_i/16 \leq \Delta_i$. It follows that $1/32\Delta_i < d(x, v_j) < (64 - 1/64)\Delta_i$, and thus by Lemma 11.3: $\max\{\rho(v_j, \Delta_i), \rho(x, \Delta_i)\} = \rho(x, \Delta_i) \geq 2$ (since by construction we choose the centers of the clusters to be the points with the minimum growth rate, among the uncovered points), which is a contradiction.

2. If $d(v_j, x) \leq \Delta_i/8$, then the cluster C_j is a good cluster for $B(x, \Delta_i/16)$, since by the construction, every radius $r_j \geq \Delta_i/4$, and therefore with probability 1 we have $B(x, \Delta_i/16) \subseteq C_j$, i.e. C_j is a good cluster.

Now, it remains to compute the padding parameter of the bundle. Given any $x \in X$ we have:

$$\sum_{i=1}^{\log\Phi} \gamma_i(x) \leq 128 \cdot \sum_{i=1}^{\log\Phi} \log(\rho(x, \Delta_i)) = 128 \sum_{i=1}^{\log\Phi} \log\left(\frac{|B(x, 64\Delta_i)|}{|B(x, \Delta_i/64)|}\right) \leq$$

since we assumed that $k \geq 16$, then for all $i \geq 4$ it holds that radius $64\Delta_i \leq \Delta_{i-3}/64$, therefore,

$$\leq 128 \log\left(\frac{|B(x, 64\Delta_1)| \cdot |B(x, 64\Delta_2)| \cdot |B(x, 64\Delta_3)|}{1}\right) \leq O(\log n),$$

which completes the proof. □

Coarse Scaling Probabilistic Embedding into k-HST's.

We show that the Uniform Partition Lemma implies probabilistic embedding with coarse scaling distortion of $O(\log(1/\epsilon))$. Given any $x \neq y \in X$ such that $d(x,y) \geq R(x, \epsilon/2)$ (recall that $R(x,\epsilon)$ is a minimal radius such that the ball $B(x, R(x,\epsilon))$ has at least $\epsilon \cdot n$ points.) we have to show that

$$E\left[\frac{d_T(x,y)}{d(x,y)}\right] \leq O\left(\log\left(1/\epsilon\right)\right).$$

Let l be a maximal index such that $\Delta_l = \Delta_0/k^l > 16d(x,y)$. Recalling the algorithm that builds a k-HST from a special bundle $\mathcal{H}$, we have

$$E\left[\frac{d_T(x,y)}{d(x,y)}\right] \leq O\left(k \cdot \left(\sum_{1\leq j<l} \gamma_i(x) + 1\right)\right).$$

Therefore,

$$\sum_{1\leq j<l} \gamma_i(x) \leq c \sum_{1\leq j<l} \log\left(\frac{|B(x, 64\Delta_i)|}{|B(x, \Delta_i/64)|}\right) = c\log\left(\prod_{1\leq j<l} \frac{|B(x, 64\Delta_i)|}{|B(x, \Delta_i/64)|}\right) \leq^{k\geq 16} \leq c\log\left(\frac{n^3}{|B(x, \Delta_{l-1}/64)|^3}\right).$$

Note that $\Delta_{l-1} = k\Delta_l$, and $k \geq 16$ therefore $\Delta_{l-1}/64 \geq 16\Delta_l/64 \geq 16^2 d(x,y)/64 \geq d(x,y)$. Therefore $B(x, R(x,\epsilon/2)) \subseteq B(x, \Delta_{l-1}/64)$, meaning that $|B(x, \Delta_{l-1}/64)| \geq \epsilon n/2$. Therefore, continuing the estimation we have $\sum_{1\leq j<l} \gamma_i(x) \leq O\left(\log(1/\epsilon)\right)$, as required.

References

[1] Ittai Abraham, Yair Bartal, and Ofer Neiman. Advances in metric embedding theory. STOC '06, pages 271–286, 2006.