10.1 The edge cutting procedure, revisited

We will present a result slightly more general than Theorem 9.1.2 from the previous lecture. The structure of the proof will be analogous to that of 9.1.2; in fact we will use the same algorithm. First, we devise a little new notation.

In the following, let $G = (V, E)$ be an undirected graph. Let $u \in V$ and $\rho \in \mathbb{N}$; $\mathsf{Ball}(u, \rho)$ is defined to be the largest $S \subseteq V$ such that every $v \in S$ has $d(u, v) \leq \rho$.

Definition 10.1.1 *Let $E' \subseteq E$. $\mathsf{Ball}(u, \rho)$* is cut by *E' iff there are distinct components C_1 and C_2 in $(V, E - E')$ whereby $\mathsf{Ball}(u, \rho) \cap C_1$ and $\mathsf{Ball}(u, \rho) \cap C_2$ are non-empty. That is, nodes within ρ of u lie in at least two distinct connected components of $(V, E - E')$.*

Theorem 10.1.2 *There is a randomized procedure that, given G and parameter $\delta > 0$, returns $E' \subseteq E$ such that:*

(1) every connected component of $(V, E - E')$ has weak diameter at most δ, and

(2) for all $u \in V$ and $\rho \leq \delta/2$, $Pr[\mathsf{Ball}(u, \rho)$ is cut by $E'] \leq \frac{8\rho}{\delta} \log n$.

Proof: We will use the graph cutting procedure due to Calinescu, Karloff and Rabani (2002) discussed in the previous lecture. As is standard, we will use the notation $x \in_R S$ to say: x is a variable whose value is randomly chosen from the set S.

$\mathsf{CKR}(G, \delta)$:

Let $\{v_1, \ldots, v_n\}$ be an arbitrary indexing of V.

- Select $r \in_R [\delta/4, \delta/2]$ and $\sigma \in_R S_n$ (the group of permutations $\sigma : [n] \to [n]$).
- For each $v_i \in V$ (in order of increasing $\sigma(i)$), set $B_i := \mathsf{Ball}(v_i, r)$ and $\hat{B}_i := B_i \setminus (\bigcup_{j:\sigma(j)<\sigma(i)} B_j)$.
- Return $E' := E \setminus \{\{u, v\} \ : \ (\exists i, j)[i \neq j \ \wedge \ u \in \hat{B}_i \wedge v \in \hat{B}_j]\}$. That is, E' consists of the edges that cross from one ball into another. (Another way of writing this is that $E' := E \setminus (\bigcup_i \partial \hat{B}_i)$.)

By the choice of r, it is easy to see that the weak diameter of any component in $(V, E - E')$ is at most δ, so (1) holds.

For a given $u \in V$ and ρ, let $\mathcal{E}_{\text{cut}}$ be the event that the ball $\mathsf{Ball}(u, \rho)$ is cut by the edges returned in the above procedure. Then part (2) of the theorem consists in showing that $Pr[\mathcal{E}_{\text{cut}}] \leq \frac{8\rho}{\delta} \log n$. To this end, we will say that v_i is *responsible for cutting* $\mathsf{Ball}(u, \rho)$ if "$\mathsf{Ball}(u, \rho)$ is cut by $\hat{B}_i$" (*i.e.* $(\mathsf{Ball}(u, v) \setminus \hat{B}_i) \cup (\hat{B}_i \setminus \mathsf{Ball}(u, v)) \neq \varnothing$), and it is not cut by $\hat{B}_j$, for all j s.t. $\sigma(j) < \sigma(i)$. In other words, v_i is the first vertex we encounter where $\hat{B}_i$ cuts the given ball.

Let $\mathcal{E}(v_i)$ be the event that v_i is responsible for cutting $\mathsf{Ball}(u,\rho)$. Then $\mathcal{E}_{\text{cut}} = \bigcup_i \mathcal{E}(v_i)$, *i.e.* if the ball is cut, there is a v_i whose ball cuts it the *first* time.

Of course, $Pr[\mathcal{E}_{\text{cut}}] \le \sum_i Pr[\mathcal{E}(v_i)]$. We will devote our time now to determining an upper bound for $Pr[\mathcal{E}(v_i)]$. First, define $\mathcal{E}_1(v_i)$ to be the event that $r \in [d(u,v_i)-\rho, d(u,v_i)+\rho]$, and $\mathcal{E}_2$ to be the event that $B(u,\rho) \cap B_j \neq \varnothing$ for all $j : \sigma(j) < \sigma(i)$. Observe that $\mathcal{E}(v_i) = \mathcal{E}_1(v_i) \wedge \mathcal{E}_2(v_i)$; moreover, they are independent since the choice of r is independent from the choice of σ. Hence if we can discern the probabilities of $\mathcal{E}_1$ and $\mathcal{E}_2$, we're done.

Claim 10.1.3 $Pr[\mathcal{E}_1(v_i)] \le \frac{8\rho}{\delta}$.

Let $\mathbf{1}$ be the indicator function from events to $\{0,1\}$, evaluating to 1 if and only if the event is true. Then

$$Pr[\mathcal{E}_1(v_i)] = Pr[\{r \in [d(u,v_i)-\rho, d(u,v_i)+\rho]\}] \le \frac{2\rho}{\delta} \cdot \mathbf{1}[\delta/4 - \rho \le d(u,v_i) \le \delta/2 + \rho].$$

This is because if $d(u,v_i)$ is not within $[\delta/4-\rho, \delta/2+\rho]$, then $r \notin [d(u,v_i)-\rho, d(u,v_i)+\rho]$ for all choices of $r \in [\delta/4, \delta/2]$. The claim follows since δ is the weak diameter of the graph.

In the last class, we showed that if $\rho \le \delta/8$, the theorem holds (almost trivially).

Claim 10.1.4 $Pr[\mathcal{E}_2(v_i)] \le 1/i$.

Assume w.l.o.g. the v_i's were originally indexed so that

$$d(u,v_1) \le d(u,v_2) \le \cdots \le d(u,v_n).$$

Then $\mathcal{E}_2(v_i)$ may be considered as the event that v_i lies before $v_1, \ldots, v_{i-1}$ in σ; *i.e.* $\sigma(v_i) < \sigma(v_j)$, for all $j < i$. This implies $Pr[\mathcal{E}_2(v_i)] \le 1/i$, since σ is chosen uniformly at random.

Combining the two claims, we have that

$$Pr[\mathcal{E}_{\text{cut}}] = \sum_{i=1}^{n} Pr[\mathcal{E}(v_i)] = \sum_{i=1}^{n} 1/i \frac{8\rho}{\delta} = \frac{8\rho}{\delta} H_n < \frac{8\rho}{\delta}(\log n + 0.578)$$

asymptotically.

■

The cutting procedure provides a neat proof of the following.

Theorem 10.1.5 (Fakcharoenphol, Rao and Talwar (2003)) *If this cutting scheme is used to build dominating trees at random, then $E[d_T(u,v)] \le 192 \log n d_G(u,v)$.*

Proof: (Sketch) Use the cutting scheme above to construct cuts E_i' for $\rho = 2^i$, where i ranges from 0 to $\log\delta + 2$. Each cut E_i' induces a set of connected components $\mathcal{C}_i$ (these sets are over V). These sets define a laminar family of sets; as in the proof of Bartal's theorem, we associate the laminar family with a rooted tree T, and define

$d_T(u,v)$ to be the shortest path distance in T from $\{u\}$ to $\{v\}$. Analogous to the proof of Bartal's theorem, we have that:

$$
\begin{aligned}
E[d_T(u,v)] &= (4\delta)\cdot\frac{8d_G(u,v)}{\delta/2}[\log|\mathsf{Ball}(u,\delta/2)| - \log|\mathsf{Ball}(u,\delta/16)|]\\
&+ (4\delta/2)\cdot\frac{8d_G(u,v)}{\delta/4}[\log|\mathsf{Ball}(u,\delta/4)| - \log|\mathsf{Ball}(u,\delta/32)|]\\
&+ (4\delta/4)\cdot\frac{8d_G(u,v)}{\delta/8}[\log|\mathsf{Ball}(u,\delta/8)| - \log|\mathsf{Ball}(u,\delta/64)|]\\
&+ \ldots
\end{aligned}
$$

This sum telescopes, so that only the positive terms involving $\log|\mathsf{Ball}(u,\delta/2)|$, $\log|\mathsf{Ball}(u,\delta/4)|$, and $\log|\mathsf{Ball}(u,\delta/8)|$ remain. The largest of these, $\log|\mathsf{Ball}(u,\delta/2)|$, is at most $\log n$. Therefore

$$
\begin{aligned}
E[d_T(u,v) &\leq [(4\delta)\cdot\frac{8d_G(u,v)}{\delta/2} + (4\delta/2)\cdot\frac{8d_G(u,v)}{\delta/4} + (4\delta/4)\cdot\frac{8d_G(u,v)}{\delta/8}]\log n\\
&\leq (64\cdot 3\log n)d_G(u,v) = 192\log n d_G(u,v).
\end{aligned}
$$

■

10.2 An application to subset embeddings

Next, we will show how to use the cutting scheme to generate good subset embeddings for metrics. In particular, one can prove an approximate version of Bourgain's theorem using the cutting schemes given in class.

We will work with the same assumptions as above, i.e., that all edge lengths are integers, and the diameter of the graph G (which represents the metric (X,d)) is Δ. We will show how to use the partitioning scheme to prove the following theorem:

Theorem 10.2.1 *There exists a subset embedding ϕ_1 such that*

$$
(X,d) \xhookrightarrow{O(\log n\log\Delta)} \ell_1^{O(\log n\log\Delta)} \tag{10.2.1}
$$

In fact, in Homework 2, we will see that the same map gives us a good embedding for all ℓ_p, promising

$$
(X,d) \xhookrightarrow{O(\log n(\log\Delta)^{1/p})} \ell_p^{O(\log n\log\Delta)}
$$

as well. Moreover, we will see that the $\log n$ term in the distortion is completely dependent on our upper bound on the probability of cutting a ball in the procedure. More precisely, if we can prove $Pr[\mathsf{Ball}(u,\rho) \text{ is cut}] \leq \frac{\rho}{\delta}\beta$ for some β, then the $\log n$ term in the distortion can be replaced with β.

Note that if we choose a value $\rho = \gamma\delta/\log n$ for a sufficiently small constant $\gamma > 0$, then the second condition of Theorem 10.1.2 implies that $\mathbf{Pr}[\mathsf{Ball}(u,\rho) \text{ is cut}]$ is at most some constant. Hence, if we wanted to maintain the distance between u and v which are at distance $\delta < d(u,v) \leq 2\delta$, we could do the following: perform a random partitioning with parameter δ and consider the sum of the distances of u and v from the boundary of the components they lie in. Since the partition ensures that the resulting components have diamater at most δ, the vertices u and v lie in different components in the partition; furthermore, the above observation implies that with constant probability, at least one of u and v lie at distance $\rho = \Omega(\delta/\log n)$ from the boundary of the clusters.

Hence, if we look at the $\log \Delta$ partitions done with $\delta = 2^i$ for $i = 1, 2, \ldots, \log \Delta$, for the particular choice of $\delta \approx d(u, v)$, the sum of the distances of u and v from the boundaries of their clusters is a reasonable approximation ($\approx d(u, v)/\log n$) to their real distance. We now have to figure out how to emulate this by embeddings into ℓ_1.

To capture the notion of "boundary" formally, let us define a *dense* version of G, constructed as follows. We begin with the graph G, and for each edge $e \in E$ of length $d(e)$, we subdivide the edge with $d(e) - 1$ fresh vertices placed at unit distance from each other. The resulting graph $G' = (V', E')$ has unit length edges, whereas the distance between the original vertices of G remains unchanged. We can now perform the same cuttings schemes on this graph, with the balls being considered only around the original vertices in V; it can be checked that the probability of an edge being cut is still given by Theorem 10.1.2. However, given a component $C \subseteq V'$ in the resulting partition, let us define the boundary $\mathsf{Bndry}(C)$ as the set of vertices, one of whose neighbors does not lie in C. I.e.,

$$\mathsf{Bndry}(C) = \{\, v \in C \mid \exists u \in V' \setminus C \text{ s.t. } (u, v) \in E' \}.$$

We are now in a position to define the mapping ϕ_1:

1. For each *level* $i = 0, \ldots, \log \Delta$,
2. For $j = 1, \ldots, O(\log n)$,

   Use the above scheme to obtain a random partition Π_{ij} of the graph G' with $\delta_i = 2^i$. For each connected component $C_1, C_2, \ldots, C_{k(i,j)}$ formed in the partition Π_{ij}, choose a random sign $s_{i,j}(C_l) \in_R \{-1, 1\}$.

   Let S_{ij} be the set of boundary vertices $\cup_{1 \le l \le k(i,j)} \mathsf{Bndry}(C_l)$. For $u \in V$, let $s_{i,j}(u)$ be the sign of the connected component in Π_{ij} that contains u. Then let us define

   $$f_{i,j}(u) := s_{i,j}(u) \cdot d(u, S_{ij}).$$

3. Finally, set $\phi_1(u) := \bigoplus_{i,j} f_{i,j}(u)$.

What is the distortion of this embedding? First, let us observe that this subset embedding is a contraction.

Claim 10.2.2 *For all i, j and $\{u, v\} \in E$, $d(u, v) \ge |f_{i,j}(u) - f_{i,j}(v)|$.*

Proof: There are two cases: u and v are either in the same component of $E_{i,j}$ or are in distinct ones. If they're in the same one, then $f_{i,j}(u)$ and $f_{i,j}(v)$ have the same sign, so

$$|f_{i,j}(u) - f_{i,j}(v)| \le |d(u, S_{ij}) - d(v, S_{ij})| \le d(u, v)$$

by the triangle inequality.

If u and v are in distinct components, then they may have different signs. However, in this case, we can infer that

$$|f_{i,j}(u) - f_{i,j}(v)| \leq d(u, S_{ij}) + d(v, S_{ij}).$$

Since a shortest path from u to v must go through the set S_{ij} at some point, the right hand side is at most $d(u,v)$, proving the claim. ■

This claim immediately implies that

$$|\phi_1(u) - \phi_2(v)|_1 = \sum_{i,j} |f_{i,j}(u) - f_{i,j}(v)| \leq O(\log n \log \Delta)\, d(u,v).$$

Next, we will argue that for all u and v, the distances in the embedding is at least a constant fraction of $d(u,v)$.

Claim 10.2.3 $|\phi_1(u) - \phi_1(v)|_1 \geq \Omega(d(u,v))$ *with probability* $1 - n^{-3}$.

Proof: The proof is close in spirit to Question 6 on Homework 1 and uses Chernoff bounds. Let $d(u,v) \in (2^i, 2^{i+1}]$, and consider the partition Π_{ij}. Let X_j be the indicator variable for the event

$$\mathcal{E} = \mathsf{Ball}(u, \gamma \frac{2^i}{\log n}) \text{ is not cut by } \Pi_{ij}.$$

Note that Theorem 10.1.2 implies that $\mathbf{Pr}[X_j = 1] \leq 8\gamma$, and hence choosing $\gamma = 1/32$ implies that this probability is a quarter.

Now if event $\mathcal{E}$ were to happen, then it holds that $|f_{i,j}(u) - f_{i,j}(v)| \geq \gamma \frac{2^i}{\log n}$ with probability $\frac{1}{2}$; indeed, with probability half, the signs $s_{i,j}(u)$ and $s_{i,j}(v)$ would be different, and would cause the distances of u and v from S_{ij} to be added up, and be greater than $\gamma \frac{2^i}{\log n}$. Hence, if Y_j is the indicator for the event that $|f_{i,j}(u) - f_{i,j}(v)| \geq \gamma \frac{2^i}{\log n}$, then $\mathbf{Pr}[Y_j = 1] \geq \frac{1}{2}\mathbf{Pr}[X_j = 1] \geq 1/16$.

Now we can use the Chernoff-Hoeffding bounds and argue that, since each $Y_j = 1$ with constant probability, it holds that for some constant $c > 0$

$$\mathbf{Pr}\left[\sum_{j=1}^{O(\log n)} Y_j \geq c \log n\right] \geq 1 - \frac{1}{n^3}.$$

This immediately implies that

$$\mathbf{Pr}\left[\sum_{j=1}^{O(\log n)} |f_{i,j}(u) - f_{i,j}(v)| \geq \Omega(d(u,v))\right] \geq 1 - \frac{1}{n^3},$$

and hence $|\phi_1(u) - \phi_1(v)|_1 \geq \Omega(d(u,v))$ with probability $1 - 1/n^3$, thus proving the claim. ■

Note that taking a union bound over all pairs $u, v \in V$ implies that the map has a constant contraction (with high probability), which proves Theorem 10.2.1.

References

Călinescu, G., Karloff, H., and Rabani, Y. (2001). Approximation algorithms for the 0-extension problem. In *Proceedings of the twelfth annual ACM-SIAM symposium on Discrete algorithms*, pages 8–16. ACM Press.

Fakcharoenphol, J., Rao, S., and Talwar, K. (2003). A tight bound on approximating arbitrary metrics by tree metrics. In *Proceedings of the thirty-fifth ACM symposium on Theory of computing*, pages 448–455. ACM Press.