

Out: 31-Oct

Due: Before class, November 7th

For the following problems, suppose that  $G = (V, E)$  is a connected graph with unit edge weights, and let  $d(u, v)$  denote the shortest-path distance between  $u$  and  $v$  for  $u, v \in V$ .

1. **Hierarchical random partitions.** In class, we showed the existence of random partitions that satisfy the following. For every  $\Delta > 0$ , there is a random partition  $V = C_1 \cup C_2 \cup \dots \cup C_k$  such that  $\text{diam}(C_i) \leq \Delta$  for all  $i = 1, \dots, k$  and so that for every  $u, v \in V$ , we have

$$\mathbb{P}[u \text{ and } v \text{ are in different clusters}] \leq \frac{d(u, v)}{\Delta} \alpha, \quad (0.1)$$

with  $\alpha \leq 8 \ln n$ .

In order to construct tree embeddings, we will need a hierarchy of such random partitions, one for every value  $\Delta = 2^j$  for  $j = 0, 1, \dots, \log_2 n - 1$ . Moreover, we will need these random partitions to be *hierarchical*.

Say that a partition  $P$  is a *refinement* of another partition  $P'$  if for every set  $S \in P$ , there is a set  $S' \in P'$  such that  $S \subseteq S'$ . Given two partitions  $P$  and  $P'$ , we can take their *common refinement*  $P \wedge P'$  which is the partition consisting of sets  $S \cap T$  for all  $S \in P$  and  $T \in P'$ . Note that  $P \wedge P'$  is a refinement of both  $P$  and  $P'$ .

Let  $P_j$  be the random partition discussed above for  $\Delta = 2^j$ . Your goal in this problem is to show that we can obtain another sequence  $\{P'_j\}$  of partitions that satisfy (0.1) with for  $\Delta = 2^j$  and  $\alpha \leq 16 \ln n$ , and such that  $P'_j$  is a refinement of  $P'_{j+1}$  for each  $j$ . Think about using common refinements.

2. **Hierarchical partitions give random trees.** Given the sequence of *hierarchical* partitions  $\{P'_j\}$  from the previous problem, let us define a random tree metric  $T$ . There will be one vertex in the tree for every pair  $(S, j)$  with  $S \in P'_j$  and  $j = 0, 1, \dots, \log_2 n$ . We connect two nodes  $(S, j+1)$  and  $(T, j)$  by an edge of length  $2^j$  if  $T \subseteq S$  and  $S \in P'_{j+1}$  and  $T \in P'_j$ .

Construct also a map  $F : V \rightarrow V(T)$  (here  $V(T)$  is the vertex set of  $T$ ). We put  $F(u) = (\{u\}, 0)$  (noting that  $P'_0$  consists solely of singleton clusters). Prove that  $T$  is a tree, and  $F$  is a random tree embedding (as defined in the lecture). Prove that the expected stretch is at most  $O((\log n)^2)$ .

3. **Bonus question: LP rounding.** Define the Sparsest Cut problem: Given a graph  $G = (V, E)$ , find a subset  $S \subseteq V$  to minimize the quantity

$$\frac{|E(S, \bar{S})|}{|S||\bar{S}|},$$

where  $\bar{S}$  denotes the complement of  $S$  and  $E(S, \bar{S})$  is the set of edge running between  $S$  and  $\bar{S}$ .

Consider the following continuous relaxation with variables  $d_{ij} \in [0, 1]$  for  $i, j \in V$ :

$$\min_{d_{ij}} \frac{\sum_{\{i,j\} \in E} d_{ij}}{\sum_{i,j \in V} d_{ij}},$$

subject to the constraints  $d_{ij} = d_{ji}$ ,  $d_{ij} \in [0, 1]$ , and  $d_{ij} \leq d_{ik} + d_{kj}$  for all  $i, j, k \in V$ . This relaxation can be written as a linear programming problem.

First, show that this is indeed a relaxation (i.e., the optimal value over the choice of  $d_{ij}$  variables is at most the optimal value of the Sparsest Cut problem). Then show that you can use the tree embedding from part (2) to prove that the optimal Sparsest Cut value is at most a factor  $O((\log n)^2)$  times more than the optimal value of the relaxation.

[Hint: The fractional solution is a metric. Embed this into a random tree metric, and then try to find a good cut in the tree. For finding cuts in the tree, it will suffice to look at cuts of the following form: For every edge  $e$  of  $T$ , removal of  $e$  splits  $T$  into two components.]

4. **Bonus question: Locality.** From (0.1) (using the probabilistic method), we can also conclude that there exists a partition  $P$  with diameter bound  $\Delta$ , and such that at least half the pairs  $u, v$  with  $d(u, v) \leq \frac{\Delta}{8 \log n}$  are not separated by the partition.

Suppose we know something more about our space: Every  $\Delta$ -ball is small:  $|B(x, \Delta)| \leq K$ . Can we improve this guarantee, maybe even to something which is independent of  $n$ ? Use the symmetric version of the Lovász Local Lemma (see below) to prove that there exists a partition with diameter bound  $\Delta$  and such that at least half the pairs  $u, v$  with  $d(u, v) \leq \frac{\Delta}{O(\log K)}$  are not separated by the partition.

**Lemma 0.1** (Lovász Local Lemma, symmetric version). *Suppose that  $A_1, A_2, \dots, A_n$  are events in some probability space such that each  $A_i$  is mutually independent of all but  $d$  other events  $A_j$ . If  $\mathbb{P}[A_i] \leq p$  for every  $i = 1, \dots, n$  and if*

$$p \leq \frac{1}{e(d+1)},$$

*then  $\mathbb{P}[\overline{A_1} \wedge \dots \wedge \overline{A_n}] > 0$ .*