# Assignment-1

Anannya Mathur 2023SIY7565

## Question 1.1

### Question (a):

The set $V_L$ is defined as the vertices $x \in V$ where $N(x, L)$ is contained within $B(x, n^a)$, which means that all nearest $n^a$ neighbours of x might not be at a distance $\leq L$. $V \backslash V_L$ denotes the set of vertices $x \in V$ where $B(x, n^a)$ is contained within $N(x, L)$, that is, all the nearest $n^a$ neighbours of x are at a distance $\leq L$. Let us define $R(v) = \{x \in V | dist(v, x) < d(v, p(v))\}$, where p(v) is the closest vertex to v in R. Thus, R(v) contains vertices that are closer to v that all other vertices of R. If there exists a vertex $r_x \in R \cap B(x)$, all vertices that are closer to x than $r_x$ will be in B(x). Since, p(x) is defined to be the closest vertex to x in R, $p(x) \in B(x)$. Thus, $dist(x, p(x)) \leq L$ as all the vertices in $B(x, n^a)$ are at a distance $\leq L$ from x. Since R hits all the vertices in $B(x, n^a)$ with a high probability of $1 - \frac{1}{n^2}$(as proved in Lecture 2), the problem statement thus stands proved.

### Question (b):

1. Consider shortest path P connecting x and y in G. The next vertex lying on the shortest path P from x to y in G will happen to be one of the closest vertices to x in G and therefore, $N(x, \frac{\epsilon * D}{2} + 1) \cap V(P\ connecting\ x\ and\ y)$ must contain atleast one vertex. Let us consider next vertex, say u $\in N(x, \frac{\epsilon * D}{2} + 1) \cap V(P\ connecting\ x\ and\ y)$ such that d(x,u) is maximum, $d(x, u) \leq \frac{\epsilon * D}{2} + 1$. If we consider the next vertex, say v $\in N(u, \frac{\epsilon * D}{2} + 1) \cap V(P\ connecting\ u\ and\ y)$ such that d(u,v) is maximum, Case 1: $d(u, v) \geq \frac{\epsilon * D}{2}$ , therefore, $d(x, v) = d(x, u) + d(u, v) \geq \frac{\epsilon * D}{2} + 1$, which implies either v is not present in $B(x, n^a)$ or it belongs to $B(x, n^a) \backslash N(x, \frac{\epsilon * D}{2} + 1)$. It can be observed that the path P can be traversed in H in at most $\lceil \frac{2}{\epsilon} \rceil$ hops. Case 2: $d(u, v) \leq \frac{\epsilon * D}{2}$ , therefore, $d(x, v) = d(x, u) + d(u, v) \leq \frac{\epsilon * D}{2} + 1$, which implies v is present in $N(x, L)$ and since, by our definition, we were supposed to pick vertex next to x on path P that maximises d(x,t) for every t present in $N(x, \frac{\epsilon * D}{2} + 1) \cap V(P\ connecting\ x\ and\ y)$, vertex v would have been picked over u. Hence, the consecutive vertices on path P should fall in case 1. Thus, path P can be traversed in H in at most $\lceil \frac{2}{\epsilon} \rceil$ hops. Let us pick the first vertex, say p, that we encounter while traversing the path P where d(x,p) becomes $\geq \frac{D}{2}$. It can be seen that vertex p can be reached from x in $\leq \lceil \frac{1}{\epsilon} \rceil$ hops. d(p,y) $\leq D/2$. As shown earlier in the proof, if a vertex v exists between p and y on path P in H, d(p,v) must be $\geq \frac{\epsilon * D}{2}$, therefore, y can be reached from p in $\leq \lceil \frac{1}{\epsilon} \rceil$ hops.

2. Let vertex p be one such vertex. We assume that d(x,p) $\leq D/2$. As seen in proof of Question b part 1, for any vertex v on path P in H connecting x and p (observe that all vertices lying of path connecting x and p except for p lie in $V_L$ and therefore, the

structure of our proof remains the same), d(x,v) must be $\geq \frac{\epsilon * D}{2}$. Hence, p can reach x in $\leq \lceil \frac{1}{\epsilon} \rceil$ hops. The proof remains the same if d(x,p) were to be $\geq D/2$. Here, we traverse backwards from y to p where d(y,p) is $\leq D/2$.

## Question (c):

Let us consider L= $\frac{\epsilon * D}{2} + 1$. We can keep two sets where one set stores vertices on path P in H that can reach x in $\leq \lceil \frac{1}{\epsilon} \rceil$ hops, while the other set stores vertices on path P in H that can reach y in $\leq \lceil \frac{1}{\epsilon} \rceil$ hops. If we consider that there is a vertex p $\in V(P) \backslash V_L$, where P is the shortest path from x to y in G, the oracle should return minimum d(x,p(v))+d(p(v),t) over v $\in$ union of the two above mentioned sets. $d(x, y, oracle) \leq d(x, p(p)) + d(p(p), y) \leq (d(x, p) + d(p, p(p))) + (d(p(p), p) + d(p, y)) \leq D + (2 * d(p(p), p)) \leq D + 2L = (1+\epsilon)D + 2$. If all the vertices of V(P) lie in $V_L$, the oracle should return minimum d(x,v)+d(v,y) over v $\in$ intersection of the two above mentioned sets. The distance $d(x, y, G) \leq d(x, y, oracle) \leq d(x, v) + d(v, y) = d$, hence, d(x,y,oracle)=d(x,y). The equality holds because a vertex in V(P) exists that is at $\leq \lceil \frac{1}{\epsilon} \rceil$ hops away from both x and y.

# Question 2:

a: $\Delta_1$ should be $n^{1/3}$.

b: $\Delta_2 = n/\Delta_1 = n^{2/3}$

c: $Q_i$ is of size $O(n^{2/3} \log n / 2^i)$

Therefore, number of edges in H $\leq \sum_i |Q_i| * 2^i * \Delta_2 = \sum_i n^{4/3} \log n = \tilde{O}(n^{4/3})$. If all the vertices lying on path P(x,y) have low degree, d(x,y,H)=d(x,y,G). If there is atleast one high degree vertex, then with high probability, a vertex w, lying of P(x,y) in G, has a neighbour r in Q, such that $d(x, y, H) \leq d(x, r, H) + d(r, y, H) \leq d(x, w, G) + 1 + d(w, y, G) + 1 \leq d(x, y, G) + 2$. We now move to the case where highest degree of vertices lying on path P(x,y) is the medium range, say $[2^{i-1}, 2^i]$. Let u be the first such vertex lying on path P, while v be the last such vertex lying on P. Given how $Q_0$ is constructed, $d(x, node \in Q_0) = 1$. Construction of $Q_i$ allows that some vertex q $\in Q_i$ will have a neighbour t on shortest path from x to y in G. Let u' be a neighbour of u in $Q_0$. Construction of $Q_0$ ensures that u,v have neighbours in $Q_0$. Similarly, the construction holds for v. $d(x, y, H) \leq d(x, u) + 1 + (1 + d(u, t) + 1) + (1 + d(t, v) + 1 + 1 + d(v, y)) = d + 6$.

# Question 3

Survey 1: Proving that determing if a t-spanner of G contains $\leq m$ edges is NP-complete.
Authors: David Peleg and Alejandro A Schaffer
Year: 1989 in Journal of Graph Theory, 13(1):99–116
Survey 2: In the year 2018, Arturs Backurs, Liam Roditty, Gilad Segal, Virginia Vassilevska Williams, and Nicole Wein showed that a $O(n^{3/2-\delta})$ time algorithm does not exist to approximate the diameter of sparse weighted graph with a ratio better than 5/3 unless Strong Exponential Time Hypothesis were to fail.

Survey 3: In the year 1999, Donald Aingworth, Chandra Chekuri, Piotr Indyk, and Rajeev Motwani proposed Fast estimation of diameter and shortest paths without making use of matrix multiplication. This work helped in finding extreme bounds for purely additive spanners. It was shown that bound on edges can be improved to $O(n^{3/2})$ for +2 additive spanners.

Open Problems: 1. In the year 2018, author Yusuke Kobayashi left an open question if minimum t– spanner problem on bounded-degree graphs of degree at most D is NP–hard for certain (t,D).

2. Is it possible to design a spanner with high probability through random sampling by creating a probability distribution over the edges of a graph?

3. Find the largest p(n) for which any p node pairs in an undirected unweighted n–node graph, there is a pairwise spanner of p node pairs on O(n) edges with +c additive error.