Tight lower bounds for the asymmetric k-center problem

Eran Halperin* Guy Kortsarz† Robert Krauthgamer*

Abstract

In the k-CENTER problem, the input is a bound k and n points with the distance between every two of them, such that the distances obey the triangle inequality. The goal is to choose a set of k points to serve as centers, so that the maximum distance from the centers C to any point is as small as possible. This fundamental facility location problem is NP-hard. The symmetric case is well-understood from the viewpoint of approximation; it admits a 2–approximation, but not better.

We address the approximability of the asymmetric k-center problem. Our first result shows that the linear program used by Archer [Arc01] to devise $O(\log^* k)$–approximation has integrality ratio that is at least $(1 - o(1)) \log^* n$; this improves on the previous bound 3 of [Arc01]. Using a similar construction, we then prove that the problem cannot be approximated within a ratio of $\frac{1}{4} \log^* n$, unless $\mathsf{NP} \subseteq \mathsf{DTIME}(n^{\log \log \log n})$. These are the first lower bounds for this problem that are tight, up to constant factors, with the $O(\log^* n)$–approximation due to [PV98, Arc01].

1 Introduction

The k-center problem is one of the most fundamental facility location problems. The input to this problem is a set V of n points and the distance $\mathsf{dist}(u, v)$ from every point u to every point v, together with a bound k. These distances are assumed to obey the triangle inequality, i.e., $\mathsf{dist}(u, w) \leq \mathsf{dist}(u, v) + \mathsf{dist}(v, w)$ for all $u, v, w \in V$. The goal is to choose k of the n points to serve as centers, so as to minimize the maximum distance from the centers to any point. Formally, the goal is:

$$\min_{C \subseteq V} \max_{v \in V} \mathsf{dist}(C, v), \quad \text{where } \mathsf{dist}(C, v) := \min_{c \in C} \{\mathsf{dist}(c, v)\}.$$

The problem is NP-hard [GJ79] and therefore it is natural to seek approximation algorithms with small approximation ratio for the problem. The symmetric case, where $\mathsf{dist}(u, v) = \mathsf{dist}(v, u)$ for all $u, v \in V$ is well-understood from the viewpoint of approximation. It admits a 2–approximation [Gon85, HS85, DF85], but not better (by an easy reduction from SET-COVER [HN79, Ple80]).

For the asymmetric case (where the distances obey the triangle inequality but $\mathsf{dist}(u, v)$ and $\mathsf{dist}(v, u)$ need not be equal), Panigrahy and Vishwanathan [PV98] devise an $O(\log^* n)$–approximation algorithm, where $\log^* n$ is the number of times one has to take the logarithm of n until we get a constant. Some time later, Archer [Arc01] achieved an $O(\log^* k)$–approximation, which is based on a linear programming approach. Archer also claims that a construction found by a computer shows an integrality ratio of 3 to this linear program.

*International Computer Science Institute, Berkeley, CA 94704, USA and Computer Science Division, University of California, Berkeley, CA 94720, USA. Supported in part by NSF grants CCR-9820951 and CCR-0121555 and DARPA cooperative agreement F30602-00-2-0601. Email: {eran,robi}@cs.berkeley.edu

†Department of Computer Sciences, Rutgers University, Camden, NJ 08102, USA. Email: guyk@camden.rutgers.edu

1.1 Our results

Our main result is an $\Omega(\log^* n)$ hardness of approximation for Asymmetric k-Center. This result is tight, up to constant factors, with the approximation algorithms of [PV98, Arc01]. Our results hold even for relatively simple inputs (of Asymmetric k-Center), namely, distances that arise from a layered digraph by taking for every two vertices their shortest path distance (in the digraph). Furthermore, our proof shows that approximating Asymmetric k-Center is hard even for bicriteria (i.e., pseudo-approximation) algorithm that are allowed to use more than k centers, namely $k \log^* n$ centers. We prove the following theorem

Theorem 1.1. Asymmetric k-Center *cannot be approximated within ratio* $\frac{1}{4}\log^* n$, *unless* $\mathsf{NP} \subseteq \mathsf{DTIME}(n^{\log\log\log n})$. *This holds even for layered digraphs of constant in-degree and out-degree.*

We actually prove the following stronger theorem, which excludes, for example, the possibility of a bicriteria algorithm that is allowed to use $k \log^* n$ centers when approximating Asymmetric k-Center within ratio $\frac{1}{3}\log^* n$.

Theorem 1.2. *For any* $0 < \beta(n) \leq \alpha(n) \leq n$, Asymmetric k-Center *cannot be approximated within ratio* $\frac{\log^* \alpha - \log^*(\beta \log^* n)}{4}$, *unless* $\mathsf{NP} \subseteq \mathsf{DTIME}(n^{\alpha})$, *This holds even for layered digraphs of constant in-degree and out-degree, and even for bicriteria algorithms that are allowed to use* βk *centers.*

Techniques. Our approach here is inspired by [HK03], in which an integrality ratio (for the Group-Steiner-Tree problem) due to [HKK+03] is used to prove a hardness of approximation (for the same problem).

Therefore, we first construct in Section 3 an instance for Asymmetric k-Center in which the ratio between the integral and fractional solutions to the linear program of [Arc01] is as large as $(1 - o(1)) \log^* n$.

We then turn this integrality ratio into a hardness of approximation result in Section 4. In essence, we turn the aforementioned integrality ratio into a reduction by carefully composing several reductions that were designed by Dinur, Guruswami, Khot and Regev [DGKR03] to prove a hardness result for vertex cover in uniform hypergraphs.

Remark. During the writing of these results, it was communicated to us that Chuzoy, Guha, Khanna and Naor independently obtained an $\Omega(\log^* n)$ hardness for Asymmetric k-Center.

2 Preliminaries

Let $G = (V, A)$ be a directed unweighted graph. For any two vertices, $v, u \in V$, if there is a path from u to v of length at most r, then we say that u r-covers v and that v is r-covered by u. Similarly, for sets $U_1, U_2 \subseteq V$, we say that U_1 r-covers U_2 if for every $u_2 \in U_2$, there is a $u_1 \in U_1$ which r-covers u_2. For every $v \in V$, let $N_{out}(v) := \{u \mid (v, u) \in A\}$, and $N_{in}(v) := \{u \mid (u, v) \in A\}$ be the outgoing and incoming immediate neighbors of v.

We say that $G = (V, A)$ is a layered graph if the set of vertices V can be partitioned into t sets $L_1, \ldots, L_t$, called *layers*, such that for each arc $(u, v) \in A$ there is $i \in \{1, \ldots, t-1\}$, such that $u \in L_i$ and $v \in L_{i+1}$.

The directed graph $G = (V, A)$ induces a distance function dist, where the distance $\mathsf{dist}(u, v)$ from a vertex u to v is simply the length of the shortest path from u to v. We will slightly abuse the definitions and consider a directed graph G as a possible input to ASYMMETRIC k-CENTER, in which case the actual input will be the distance function $\mathsf{dist}(\cdot,\cdot)$.

For every number $x > 0$, let $\mathrm{tower}^{(1)}\{x\} = 2^x$. For every $i > 1$, we define recursively $\mathrm{tower}^{(i)}\{x\} = 2^{\mathrm{tower}^{(i-1)}\{x\}}$. Hence, $\mathrm{tower}^{(i)}\{x\} = 2^{2^{\cdot^{\cdot^{2^x}}}}$, where the number of 2's in the tower is exactly i. By this definition, $\log^* n$ is the maximal i such that $\mathrm{tower}^{(i)}\{2\} \le n$.

Similary, for every number $x > 0$, let $\log^{(1)} x = \log(x)$. For $i > 1$, we define recursively $\log^{(i)} x = \log(\log^{(i-1)} x$. We call this the iterated logarithm of order i.

3 Integrality ratio

Let the digraph $G = (V, A)$ be an instance of ASYMMETRIC k-CENTER, and let $n = |V|$. The algorithms of both [PV98] and [Arc01] consider the following promise problem as a subroutine to their algorithm: If the optimal solution equals 1, return a solution whose value is smaller than $\alpha(n)$; otherwise, return either a solution smaller than $\alpha(n)$ or do not return any solution. Using this subroutine, [PV98] show how to find an $\alpha(n)$-approximation to ASYMMETRIC k-CENTER. In [PV98] they find a solution to the promise problem with $\alpha(n) = O(\log^* n)$. Archer [Arc01] uses the following linear relaxation to the problem and gets a solution with $\alpha = O(\log^* k)$.

$$\boxed{\begin{array}{lll} \sum_{v\in V} x_v \le k & & \\ \sum_{\{u|(u,v)\in A\}} x_u \ge 1 & , & v \in V \\ 0 \le x_v \le 1 & , & v \in V \end{array}} \tag{1}$$

In this section we show that this linear program admits a $(1 - o(1)) \log^* n$ integrality ratio; that is, there are graphs with n vertices, in which the linear program is feasible, while no set of k centers d-covers all the vertices with $d < (1 - o(1)) \log^* n$. This can be thought of as if the fractional solution is 1, and the integral solution is at least $(1 - o(1)) \log^* n$. Clearly, $\log^* n \ge \log^* k$.

3.1 The Graph Construction

We describe a probabilistic construction of a directed graph $G = (V, A)$, and show that with high probability, the integral solution is at least $\log^* n$, while the linear program corresponding to the graph is feasible.

The graph G is a layered graph. It consists of $t + 1$ layers $L_0, L_1, \dots, L_t$, where t will be determined later (t will be approximately $\log^* n$). The arcs in the graph are always directed from a vertex of L_i to a vertex of L_{i+1} for some $i \in \{0, \dots, t-1\}$. The first layer L_0 consists of a single vertex which will be connected to all the vertices of L_1. For $i \ge 1$, the arcs from layer L_i to layer L_{i+1} are randomly chosen in the following way. For each vertex $v \in L_{i+1}$, we randomly pick d_i distinct vertices in L_i to be the set $N_{in}(v)$, i.e. the set of vertices that have an arc directed to v. The different choices of $N_{in}(v)$ for different vertices are independent. The numbers d_i are defined recursively by $d_{i+1} = 2^{(d_i)^3}$, where $d_1 = t^2$. We further choose the size of the set L_i to be $n_i = d_i d_t$. Finally, we set the number of centers k to be $t \cdot d_t$.

3.2 The fractional solution

We now show that a fractional solution to the linear program (1) is feasible. For every $i \in \{1, \dots, t-1\}$ and every vertex $v \in L_i$ we set $x_v = \frac{1}{d_i}$. Furthermore, we set $x_v = 1$ for the single vertex $v \in L_0$.

Clearly, the set of fractional centers fractionally cover each vertex in the graph, since each vertex $u \in L_{i+1}$ has incoming arcs from d_i distinct vertices, each contributing $\frac{1}{d_i}$ to its cover. The number of centers used in level L_i is $n_i/d_i = d_t$, and in level L_0 we use exactly one center. Thus, the total size of the fractional center set is $(t-1)d_t + 1 < k$, and therefore the constraints of the linear program are satisfied.

3.3 The integral solution

We now show that with high probability, any choice of k centers will not t-cover the whole set of vertices. We actually prove a stronger claim. We prove that with high probability, no choice of k centers t-covers the set L_t. In order to prove this, we show the following.

Lemma 3.1. *With high probability, for any choice of a set $C \subseteq V \setminus L_0$ of k centers, there is a vertex $v \in L_t$, such that C cannot reach v.*

Clearly, if we prove Lemma 3.1, then with high probability, in order to cover L_t, we must use L_0, and thus, the value of the integral solution is $t+1$. In the remaining of this section we prove Lemma 3.1.

Fix a set of centers $C \subseteq L_1$, such that $|C| = k$. Without loss of generality we can assume that $C \subseteq L_1$, since we can always replace each center $c \in L_i$ by a single vertex $v \in L_1$ where v can reach c, and then all the vertices of L_t that are reachable from c are also reachable from v. Let Y_i be the set of vertices in L_i which cannot be reached from C. By definition, $|Y_1| = n_1 - k \geq 2d_t$. In the next lemma we show that with high probability this is also true for Y_i for $i > 1$.

Lemma 3.2. *If $|Y_i| \geq 2d_t$, then the probability that $|Y_{i+1}| < 2d_t$ is at most $e^{-d_i d_t}$.*

Proof. Consider a vertex $v \in L_{i+1}$. We bound the probability that v is in Y_{i+1}. Since $N_{in}(v)$ is a random set of size d_i, the probability that $v \in Y_{i+1}$ is

$$\frac{\binom{|Y_i|}{d_i}}{\binom{n_i}{d_i}} \geq (\frac{|Y_i| - d_i}{n_i - d_i})^{d_i} \geq (\frac{1}{d_i})^{d_i} \geq \frac{1}{2^{(d_i)^2}},$$

where the second inequality follows from the assumption that $|Y_i| \geq 2d_t$ in conjunction with $n_i = d_i d_t$. Since the neighbors of each of the vertices of L_{i+1} are chosen independently, $|Y_{i+1}|$ is dominating a binomial variable $Z \sim B(n_{i+1}, \frac{1}{2^{(d_i)^2}})$. Let $\mu = E[Z] = \frac{n_{i+1}}{2^{(d_i)^2}} \geq 16 d_1 d_t$ (assuming $d_1 > 2$). By Chernoff,

$$\begin{aligned} \Pr(|Y_{i+1}| < 2d_t) &\leq \Pr(Z < 2d_t) \\ &\leq \Pr(Z \leq \frac{\mu}{2}) \\ &\leq e^{-\frac{\mu}{8}} \leq e^{-2d_1 d_t} \end{aligned}$$

□

We now use the union bound to show that with high probability *for any* choice of a set of centers $C \subseteq L_1$ of size k we cannot reach L_t.

Lemma 3.3. *The probability that there exists a set $C \subseteq L_1$ of k centers that t-covers L_t is at most e^{-n_1}.*

Proof. As noted above, $|Y_1| \geq 2d_t$. By applying Lemma 3.2, we get that for any fixed set of k centers $C \subseteq L_1$, the probability that $|Y_t| < 2d_t$ is bounded by $(t-1)e^{-2d_1 d_t} = (t-1)e^{-2n_1}$. Since the number of such sets C is bounded by the total number of possible sets of L_1, that is 2^{n_1}, we get by the union bound that the probability that there exist a set C for which $|Y_t| < 2d_t$ is at most $2^{n_1}(t-1)e^{-2n_1} < e^{-n_1}$. $\square$

Since $n_i = d_i d_t$, the total number of vertices n is $1 + n_1 + \ldots + n_t \leq 2d_t^2$, and thus, $n_1 > \sqrt{n}$. Hence, the probability to cover L_t by centers that do not contain L_0 is at most $e^{-\sqrt{n}}$, and this completes the proof of Lemma 3.1.

By the definition of d_i, one can verify that that $\log^* d_t = t + \Theta(1) + \log^* d_1 = t + \log^* t + \Theta(1)$. Since $n \leq 2d_t^2$, we get that $t \geq \log^* n - \log^*(\log^* n) - \Theta(1)$.

4 Hardness of approximation

In this section we prove Theorem 1.1, i.e., that Asymmetric k-Center cannot be approximated within ratio $\frac{1}{4}\log^* n$, unless $\mathsf{NP} \subseteq \mathsf{DTIME}(n^{\log\log\log n})$. In essence, we turn the integrality ratio presented in Section 3 into a reduction, by carefully composing several Max-k-Cover reductions; the Max-k-Cover reductions that we use follow almost immediately from the hardness result for vertex cover in uniform hypergraphs (a special case of Set-Cover) due to Dinur, Guruswami, Khot and Regev [DGKR03].

Terminology. Set-Cover is the following problem: Given a bipartite graph $H = (A, B, E_H)$, the goal is to find a subset $S \subseteq A$ of minimum size, such that all the vertices of B are covered by S. (A vertex is *covered* by S if it is adjacent to at least one vertex of S.) Let Set-Cover(H) denote the size of a set cover of B, measured as fraction of the vertices of A.

The maximization variant of Set-Cover called Max-k-Cover, is the following problem: Given a bipartite graph $H = (A, B, E_H)$ and an integer k, the goal is to find a subset $S \subseteq A$ of size k that maximizes the number of vertices in B that are covered by S. Let k-Cover(H) denote the largest fraction of vertices in B that can be covered by k vertices from A.

4.1 The Max-k-Cover reduction

We now state the hardness result of [DGKR03] and some of its parameters (such as gap location and degree). We will then show that this result yields a hardness result for Max-k-Cover with parameters (such as gap location and size) that are very useful for us.

Theorem 4.1 ([DGKR03]). *Let L be any* NP *language. Then for every fixed $k \geq 3$ and $\varepsilon > 0$ there is a polynomial time algorithm (i.e., reduction) that, given an instance x for L, computes an instance $H = (A, B, E_H)$ for* Set-Cover*, such that if $x \in L$ then* Set-Cover$(H) \leq 1/(k-1-\varepsilon)$ *and if $x \notin L$ then* Set-Cover$(H) \geq 1 - \epsilon$.

Additional properties. The reduction of [DGKR03] has a few additional properties that will be important for us. First, the parameters k, ε need not be fixed; they can part of the input to the reduction and may depend on $n := |x|$, and then the reduction's running time is polynomial in the output size. To this end, an inspection of the reduction gives the crude estimate $|A| \leq 2^{k^{O(1/\varepsilon^3)}} n^{O(1/\varepsilon^2)}$. Furthermore, the degree of vertices in A is at most $\Delta_{k,\varepsilon} := \mathrm{tower}^{(2)}\{O(k^2/\varepsilon^3)\}$ and hence $|E_H| \leq \Delta_{k,\varepsilon}|A|$. It is also easy to verify that the degree of vertices in B is exactly k (and in particular, at least 1), and thus $|B| \leq |E_H|$.

Using these properties, we now obtain a hardness for Max-k-Cover, which guarantees that for a NO instance, using even almost all the vertices of A, one must miss (i.e., not cover) a non-negligible fraction of B.

Corollary 4.1. *Let L be any* NP *language. Then there is an algorithm (i.e., reduction) that, given an instance x for L together with $r \geq 3$, computes in time that is polynomial in the output, an instance $H = (A, B, E_H)$ for* Set-Cover *such that if $x \in L$ then $\frac{1}{r}|A|$-Cover$(H) = 1$ and if $x \notin L$ then $(1 - \frac{1}{r})|A|$-Cover$(H) \leq 1 - 1/\operatorname{tower}^{(2)}\{r^{O(1)}\}$.*

Proof. Given an instance x for L, denote $n := |x|$ and apply Theorem 4.1 with $k := 2r$ and $\varepsilon := 1/3r$. If $x \in L$ then all of B can be covered using a subset of A whose size is at most $|A|/(k-1-\varepsilon) \leq |A|/r$, and thus $\frac{1}{r}|A|$-Cover$(H) = 1$.

Suppose now that $x \notin L$, and assume for contradiction that there exists a subset $S \subseteq A$ of size $(1 - \frac{1}{r})|A| = (1 - 3\varepsilon)|A|$ that covers at least a $1 - \varepsilon/\Delta_{k,\varepsilon}$ fraction of B. It then follows that S can be extended into a set cover of B by greedily adding to it at most $\frac{\varepsilon}{\Delta_{k,\varepsilon}}|B| \leq \varepsilon|A|$ vertices (of A), and thus Set-Cover$(H) \leq (1 - 3\varepsilon) + \varepsilon = 1 - 2\varepsilon$, which contradicts Theorem 4.1. We conclude that $(1 - \frac{1}{r})|A|$-Cover$(H) \leq 1 - \varepsilon/\Delta_{k,\varepsilon} \leq 1 - 1/\operatorname{tower}^{(2)}\{r^{O(1)}\}$. $\square$

Remark. In the resulting graph H, we have $|A| \leq 2^{2^{O(r^4)}} n^{O(r^2)}$ and $|B|/|A| \leq \Delta_{2r,1/3r} \leq \operatorname{tower}^{(2)}\{r^{O(1)}\}$. Our intended application requires a Set-Cover instance similar to H, but with a predetermined size s for $|A|$. Assuming that $s \geq 3r|A|$, we achieve this by essentially taking a disjoint union of a suitable number of copies of H, as follows. Given s, we construct a graph $H' = (A', B', E_{H'})$ by taking $\lfloor s/|A| \rfloor$ disjoint copies of H and letting A' be the union of all the copies of A, and similarly letting B' be the union of all the copies of B; if $|A|$ does not divide s then we add to A' new vertices so that its size will be exactly s. Notice that the fraction of these added vertices in A' is less than $\frac{|A|}{s} \leq \frac{1}{3r}$. In addition, the degree of every vertex in A' is at most $\Delta_{k,\varepsilon}$, and therefore $|B'|/|A'| \leq \Delta_{2r,1/3r} \leq \operatorname{tower}^{(2)}\{r^{O(1)}\}$. It is easy to verify that if $x \in \mathcal{L}$ then

$$\text{Set-Cover}(H') \leq |A'| \frac{1 + 1/3r}{k - 1 - \epsilon} \leq \frac{1}{r}|A'|,$$

and thus $\frac{1}{r}|A'|$-Cover$(H') = 1$. If $x \notin \mathcal{L}$ then

$$\text{Set-Cover}(H') \geq \frac{1 - \varepsilon}{1 + 1/3r} > 1 - 2\varepsilon,$$

and similarly to the proof of Corollary 4.1, we get that $(1 - \frac{1}{r})|A'|$-Cover$(H') \leq 1 - \varepsilon/\Delta_{k,\varepsilon} \leq 1 - 1/\operatorname{tower}^{(2)}\{r^{O(1)}\}$. Finally, we can assume without loss of generality that $|B'| \geq |A'|$, simply by duplicating $|B'|$ sufficiently many times. (In fact, this property already holds in the reduction of [DGKR03].)

4.2 Our reduction

Our reduction starts from an arbitrary NP language L. Given an instance x for L, let us assume throughout that $n := |x|$ is sufficiently large. Similarly to the integrality ratio construction (Section 3) we construct a layered digraph $G = (V, E)$ with t layers $L_0, L_1, \ldots, L_t$ (for t that we will later choose to be $\Theta(\log^* n)$). The arcs in this digraph are always directed from a vertex of L_i to a vertex of L_{i+1} for some $i \in \{0, \ldots, t-1\}$. The first layer L_0 consists of a single vertex v_0 that has arcs connecting it to all the vertices of L_1. As described below, every two successive layers L_i, L_{i+1} for $i \geq 1$ essentially form a copy of the graph H from Corollary 4.1, constructed with certain parameter r_i.

The motivation for this construction is that if $x \in \mathcal{L}$ then each layer L_{i+1} is covered by $|L_i|/r_i$ centers from layer L_i, while if $x \notin \mathcal{L}$ then, the phenomenon that occurs in the integrality ratio instance occurs here too, namely, for every choice of k centers in L_1 there is some non-negligible fraction of L_2 that is not reached, and continuing iteratively we get that some vertices of L_t cannot be reached. Let L_1 consist of $N := n^{\log^{(4)} n}$ vertices; we will see later that the size of every layer (and thus of V) will be at most $O(N \log\log N)$. Set $r_1 := 8(\log^* N)^2$ and $r_{i+1} := \text{tower}^{(3)}\{r_i\}$ for $i = 1, \dots, t-1$. Hence, $r_i = \text{tower}^{(3i-3)}\{r_1\}$.

Now, for each $i = 1, 2, \dots, t-1$ (iteratively) construct the digraph induced on layers L_i, L_{i+1}, as follows. Apply Corollary 4.1 (and the remark following it) with parameter r_i and with $s = |L_i|$ to create a graph $H'_i = (A'_i, B'_i, E_{H'_i})$. To see that this is possible, recall that $|L_1| = N$, and that for $i \geq 2$, the size of L_i is determined by the previoues iteration; we will also verify later that $s = |L_i| \geq 3r_i|A_i|$.

Finally, let us choose $t = \lfloor 0.3 \log^* N \rfloor$ and set the number of centers to be $k := tN/r_1$. Let us now upper bound $|V|$ and show that $t \geq \frac{1}{4}\log^* |V|$. Since $|L_{i+1}| \leq \Delta_{2r_i, 1/3r_i}|L_i| \leq \text{tower}^{(3)}\{r_i\}|L_i| = \text{tower}^{(3i)}\{r_1\}|L_i|$, we get that $|V| \leq \sum_{i=0}^{t} |L_i| \leq \text{tower}^{(3t-2)}\{r_1\}N$. By our choice of t we get that $|V| \leq N \log^{(4)} N$, and thus $t \geq \frac{1}{4}\log^* N$.

It remains to verify that $s = |L_i| \geq 3r_i|A_i|$. To see this, notice that $|A_i| \leq \text{tower}^{(3)}\{r_i\} \cdot n^{O(r_i^2)}$ while $\text{tower}^{(3)}\{r_i\} \leq \log^{(4)} N$, and thus $|A_i| \leq n^{\log^{(6)} n}$. The claim now follows since $|L_i| \geq |L_1| = N$.

4.3 YES instance

We now show that if $x \in \mathcal{L}$ then the resulting graph G has a set of $k = tN/r_1$ centers that 1-covers G. By the assumption $x \in L$, for every two consecutive levels L_i, L_{i+1} with $i \geq 1$ there is a subset of L_i of size at most $|L_i|/r_i$ that 1-covers L_{i+1}. In addition, all of L_1 can be covered by the single vertex $v_0 \in L_0$. Clearly, taking all these vertices as centers yields a 1-cover of all of G. Now notice that $\frac{|L_i|}{r_i} \leq \frac{|L_{i-1}|}{r_{i-1}}$, since $r_i \frac{|L_{i+1}|}{|L_i|} \leq r_i \cdot \text{tower}^{(2)}\{r_i^{O(1)}\} \leq r_{i+1}$. It follows that the number of centers at each layer L_i is at most $\frac{|L_i|}{r_i} \leq \frac{|L_1|}{r_1} = \frac{N}{r_1}$, and thus the total number of centers is at most $1 + (t-1)N/r_1 \leq k$.

4.4 NO instance

We now show that if $x \notin \mathcal{L}$ then G has no set of k centers that $(t-1)$-covers all of G. It would then follows that the value of this Asymmetric k-Center instance is at least $t \geq \frac{1}{4}\log^* |V|$). We will actually prove a stronger claim, stated as follows.

Lemma 4.2. *If $x \notin \mathcal{L}$ then for every choice of k centers in G, there is a vertex of L_t that is not $(t-1)$-covered.*

Proof. Consider a set of k centers and assume for contradiction that L_t is $(t-1)$-covered. Since G is a layered graph with non-zero in-degree, we can replace each center c with a center $v \in L_1$ that $(t-1)$-covers all the vertices of L_t that are $(t-1)$-covered by c. (Notice that the vertex of L_0 cannot be used to $(t-1)$-cover L_t.)

By the assumption $x \notin \mathcal{L}$, and since the graph induced on $L_1 \cup L_2$ is just H'_1, we know that these $k \leq tN/r_1 \leq (1 - \frac{1}{r_1})|L_1|$ centers 1-cover at most $1 - 1/\,\text{tower}^{(2)}\{r_1^{O(1)}\} \leq 1 - 1/r_2$ fraction of the vertices of L_2. Continuing inductively, for every $i = 2, \ldots, t$, since the graph induced on $L_i \cup L_{i+1}$ is similar to the graph H'_i, the fact that at most $1 - 1/r_i$ fraction of L_i is $(i-1)$-covered by these centers implies that at most $1 - 1/\operatorname{tower}^{(2)}\{r_i^{O(1)}\} \leq 1 - 1/r_{i+1}$ fraction of the vertices of L_{i+1} are i-covered. We therefore conclude that at most $1 - 1/r_t$ fraction of L_t are $(t-1)$-covered by these centers. Since $1/r_t > 0$, we get that L_t is not $(t-1)$-covered by these centers, as desired. $\square$

The proof of Theorem 1.1 now follows immediately from Section 4.3 and Lemma 4.2.

We note that the proof given here for Theorem 1.1 can be easily modified to achieve hardness of $(\frac{1}{2} - o(1)) \log^* n$. The current choice of parameters is made so as to simplify the exposition.

Theorem 1.2 follows from the observation that the size of the reduction is dominated by the last layer, and if we are given the option of using βk centers, then r_1 must increase proportionally to β. Therefore, β and α determine the parameters of the first and last layers in the graph, while the hardness result is proportional to the number of layers. A full description of this extension is omitted from this version.

References

[Arc01] A. F. Archer. Two $O(\log^* k)$-approximation algorithms for the asymmetric k-center problem. In *Proceedings of the 8th Conference on Integer Programming and Combinatorial Optimization*, pages 1–14, 2001.

[DF85] M. E. Dyer and A. M. Frieze. A simple heuristic for the p-centre problem. *Oper. Res. Lett.*, 3(6):285–288, 1985.

[DGKR03] I. Dinur, V. Guruswami, S. Khot, and O. Regev. A new multilayered PCP and the hardness of hypergraph vertex cover. To appear in *35th Annual ACM Symposium on Theory of Computing*, 2003.

[GJ79] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-completeness*. W.H. Freeman and Company, 1979.

[Gon85] T. F. Gonzalez. Clustering to minimize the maximum intercluster distance. *Theoret. Comput. Sci.*, 38(2-3):293–306, 1985.

[HK03] E. Halperin and R. Krauthgamer. Polylogarithmic inapproximability. To appear in *35th Symposium on Theory of Computing*, June 2003.

[HKK+03] E. Halperin, G. Kortsarz, R. Krauthgamer, A. Srinivasan, and N. Wang. Integrality ratio for group Steiner trees and directed Steiner trees. In *14th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 275–284, January 2003.

[HN79] W. L. Hsu and G. L. Nemhauser. Easy and hard bottleneck location problems. *Discrete Appl. Math.*, 1(3):209–215, 1979.

[HS85] D. S. Hochbaum and D. B. Shmoys. A best possible heuristic for the k-center problem. *Math. Oper. Res.*, 10(2):180–184, 1985.

[Ple80] J. Plesník. On the computational complexity of centers locating in a graph. *Apl. Mat.*, 25(6):445–452, 1980.

[PV98] R. Panigrahy and S. Vishwanathan. An $O(\log^* n)$ approximation algorithm for the asymmetric p-center problem. *J. Algorithms*, 27(2):259–268, 1998.