Polynomial Integrality Gap of Flow LP for Directed Steiner Tree

Shi Li *
Dept. of Comp. Sci. and Eng.
University at Buffalo, USA
shil@buffalo.edu

Bundit Laekhanukit †
ITCS,
SUFE, China.
lbundit+sufe@gmail.com

Abstract

In the Directed Steiner Tree (DST) problem, we are given a directed graph $G = (V, E)$ on n vertices with edge-costs $c \in \mathbb{R}^E_{\geq 0}$, a root vertex r, and a set K of k terminals. The goal is to find a minimum-cost subgraph of G that contains a path from r to every terminal $t \in K$. DST has been a notorious problem for decades as there is a large gap between the best-known polynomial-time approximation ratio of $O(k^\epsilon)$ for any constant $\epsilon > 0$, and the best quasi-polynomial-time approximation ratio of $O\left(\frac{\log^2 k}{\log\log k}\right)$.

Towards understanding this gap, we study the integrality gap of the standard flow LP relaxation for the problem. We show that the LP has an integrality gap polynomial in n. Previously, the integrality gap LP is only known to be $\Omega\left(\frac{\log^2 n}{\log\log n}\right)$ [Halperin et al., SODA'03 & SIAM J. Comput.] and $\Omega(\sqrt{k})$ [Zosin-Khuller, SODA'02] in some instance with $\sqrt{k} = O\left(\frac{\log n}{\log\log n}\right)$. Our result gives the first known lower bound on the integrality gap of this standard LP that is polynomial in n, the number of vertices. Consequently, we rule out the possibility of developing a poly-logarithmic approximation algorithm for the problem based on the flow LP relaxation.

1 Introduction

Network design problems play an important role in the area of combinatorial optimization both in theory and practice. Two most famous problems in the area are the *Minimum Spanning Tree* problem, and its analogue in directed graphs, the *Minimum-Cost Arborescence* problem. They play a fundamental role in illustrating the powerful concept of greedy algorithms in standard textbooks. While the two textbook problems ask to find a tree or an arborescence that spans all the vertices, in most applications, it is more natural to connect only a subset of vertices, called *terminals*, which models clients in the network, while using some non-terminals or less-important nodes, called *Steiner vertices*, as relay. This motivates the *Steiner tree* problems on both undirected and directed graphs, which have become the central focus in the area of network design for several decades. Since the problems are known to be NP-hard, the algorithmic development has mainly been on finding good approximation algorithms.

In contrast to the Minimum Steiner Tree problem (on undirected graphs), the complexity status of the *Directed Steiner Tree* (DST) problem is much less-understood. Formally, in the Directed Steiner Tree problem, we are given a directed graph $G = (V, E)$ on $n = |V|$ vertices, with edge costs $c \in \mathbb{R}^E_{\geq 0}$, a root $r \in V$ and a set $K \subseteq V$ of $k = |K|$ terminals. The goal of the problem is to find a minimum cost subgraph T of G that contains a path from r to t, for every terminal $t \in K$. By minimality, we may assume WLOG that the solution subgraph T is a tree (more precisely, an out-arborescence) rooted at r.

*The work is supported in part by NSF grant CCF-1844890.

†This work is supported by Science and Technology Innovation 2030 –"New Generation of Artificial Intelligence" Major Project No.(2018AAA0100903), NSFC grant 61932002, Program for Innovative Research Team of Shanghai University of Finance and Economics (IRTSHUFE) and the Fundamental Research Funds for the Central Universities. Bundit Laekhanukit is also supported by the national 1000-talent award by the Chinese government.

The best known polynomial-time approximation ratio for the problem is only $O(k^\epsilon)$ with a running time of $n^{O(1/\epsilon)}$ for any constant $\epsilon > 0$ [Zel97, CCC+99]. With quasi-polynomial time algorithms, one can achieve an approximation ratio of $O\left(\frac{\log^2 k}{\log\log k}\right)$ [GLL19, GN20]. It is known from the work of Halperin and Krauthgamer [HK03] that unless NP $\subseteq$ ZPTIME($n^{\text{poly}\log(n)}$), there is no polynomial time $O(\log^{2-\epsilon} k)$-approximation for the problem for any constant $\epsilon > 0$. The long-standing open question on DST is whether a poly-logarithmic approximation for the problem can be achieved in polynomial time.

In this paper, we study the integrality gap of the standard flow-based relaxation of DST, in the hope to shed some light on the approximability of DST under polynomial-time algorithms. The interesting part on the flow LP is that while its integrality gap is known to be lower bounded by $\Omega(\sqrt{k})$ [ZK02], the parameter k (i.e., the number of terminals) in the construction is only $O\left(\frac{\log^2 n}{\log^2\log n}\right)$. Thus, the lower bound given by the instance of [ZK02] is only $\Omega\left(\frac{\log n}{\log\log n}\right)$. Indeed, the hardness result of [HKK+07] implies an integrality gap lower bound of $\Omega\left(\frac{\log^2 n}{\log\log n}\right)$, which is better than that of [ZK02] when we concern the dependence on n.

1.1 Our Results

In this paper, we show that the integrality gap of the flow LP is indeed lower bounded by $\Omega(n^\delta)$ for some absolute constant $\delta \in (0,1)$. Our construction resemblances the previous one by Zosin and Khuller [ZK02], in the sense that our instance also has a 5-level structure. In the instance of [ZK02], each vertex in the graph corresponds to some subset of K (the set of terminals) of size roughly $\sqrt{k}$. This leads to an exponential dependence of n on k. In our construction, both the terminals and Steiner vertices correspond to subsets of some ground set, so n and k are polynomially related. Crucially, we can still prove the useful properties that are needed to show the integrality gap. The formal statement of our result is stated in Theorem 1, after we formally defined the flow LP relaxation (FLP). As a consequence of our result, we rule out poly-logarithmic approximation algorithms for DST based on the flow LP.

1.2 Related Work

Compared to DST, the Minimum Steiner Tree problem (in undirected graphs) is better-understood. Simply computing the minimum spanning tree over the metric closure of the terminals leads to a 2-approximation algorithm for the problem. The current best approximation ratio for the problem is $\ln(4) + \epsilon < 1.39$ due to Byrka et al. [BGRS10]. On the negative side, the problem is known to be APX-hard [BP89, CC08].

The DST problem has been a subject of studies for decades. The first non-trivial approximation result on this problem is due to Zelikovsky [Zel97], which gives an $O(k^\epsilon)$-approximation in $n^{O(1/\epsilon)}$-time for any constant $\epsilon > 0$, which is designed for directed acyclic graphs. The result is later extended in [CCC+99] to general graphs. Moreover, the ϵ in [CCC+99] can be set to $1/\log k$, leading to an $O(\log^3 k)$-approximation algorithm for DST in $n^{O(\log k)}$ time; this is the first poly-logarithmic approximation algorithm for the problem in quasi-polynomial time. A similar result was obtained by Kortsarz and Peleg [KP99]. The approximation ratio in the quasi-polynomial-time regime has been improved to $O\left(\frac{\log^2 k}{\log\log k}\right)$ in [GLL19, GN20]. Grandoni et al. [GLL19] also showed that this is the best approximation guarantee for any quasi-polynomial-time algorithm unless NP $\subseteq \bigcup_{c>0}$ BPTIME(2^{n^c}) or the *Projection Game Conjecture* [Mos15] is false.

There have been quite a few studies devoted to understand the power of LP/SDP hierarchy for the problem. It was shown in [Rot11] and [FKK+14] that the integrality gap of some basic LP relaxations can be brought down to $O(L\log^2 k)$ for DST on *L-level* graphs, if we lift them by $O(L)$ levels using some well-known LP/SDP hierarchies such as Sherali-Adams, Lovász–Schrijver and Lasserre hierarchies. The quasi-polynomial time $O\left(\frac{\log^2 k}{\log\log k}\right)$-approximation of [GLL19] is also based on the Sherali-Adams hierarchy.

2 Flow LP Relaxation for Directed Steiner Tree

The standard flow LP relaxation is given in (FLP). In the correspondent integer program, for every $e \in E$, we have a variable x_e indicating whether e is in the output tree T. For every terminal $t \in K$, we let $\mathcal{P}_t$ be the set of simple paths in G from r to t. Then for every $t \in K, P \in \mathcal{P}_t$, we have a variable f^t_P indicating if the $r \to t$ path in T is P. In the LP, we relax the 0/1-constraints on variables to non-negativity constraints (3) and (4).

$$\min \quad \sum_{e \in E} c_e x_e \tag{FLP}$$

s.t.

$$\sum_{P \in \mathcal{P}_t} f^t_P = 1 \qquad \forall t \in K \tag{1}$$

$$\sum_{P \in \mathcal{P}_t : e \in P} f^t_P \le x_e \qquad \forall e \in E, t \in K \tag{2}$$

$$x_e \ge 0 \qquad \forall e \in E \tag{3}$$

$$f^t_P \ge 0 \qquad \forall t \in K, P \in \mathcal{P}_t \tag{4}$$

Equation (1) says that exactly one $r \to t$ path should exist in T, for every $t \in K$. Equation (2) says that if the $r \to t$ path P in the output tree T uses an edge e, then e must be in T. So, the constraints are valid. In the LP, we can view f^t_P as the amount of flow sending from r to t using a path P. Then Equation (1) and Equation (2) together are equivalent to saying that the maximum flow from r to t in the graph G with capacities $(x_e)_{e \in E}$ has value at least 1, for every terminal $t \in K$. For this reason, we call (FLP) the flow LP. It can be solved efficiently despite the fact that it has exponential number of variables since the requirement for every t is just a maximum-flow problem, which is equivalent to a polynomial-sized LP.

Zosin and Khuller [ZK02] showed that (FLP) has an integrality gap of $\Omega(\sqrt{k})$, for a family of instances with $\sqrt{k} = O\left(\frac{\log n}{\log \log n}\right)$. Therefore, the dependence of the integrality gap on k is only logarithmic. In this paper, we show the following theorem:

Theorem 1. *The integrality gap of (FLP) is $\Omega(n^\delta)$ for some absolute constant $\delta \in (0,1)$.*

Organization. The remainder of this paper is organized as follows. We first introduce the 5-level instance of Zosin and Khuller [ZK02] in Section 3, where we give some necessary properties to prove an integrality gap. We give our construction satisfying the properties in Section 4, which finishes the proof of Theorem 1.

3 Zosin-Khuller Type Gap Instance

In this section, we define the Zosin-Khuller type gap instance for the flow LP of the directed Steiner tree problem. To define such an instance, we need to specify the following objects.

(P1) $H = (\mathcal{A} \uplus \mathcal{B}, E_H)$ is a bipartite graph with $|\mathcal{A}| \le |\mathcal{B}|$, where all vertices in $\mathcal{A}$ have the same degree d and all vertices in $\mathcal{B}$ have the same degree d'.

(P2) K is a set of k terminals, and $\{M_t\}_{t \in K}$ is a partition of E_H into k matchings of size s each, one for each terminal $t \in K$. Therefore, we have $s \le |\mathcal{A}| \le |\mathcal{B}|, k \ge d \ge d'$ and $sk = d|\mathcal{A}| = d'|\mathcal{B}| = |E_H|$. For convenience, we also treat each terminal $t \in K$ as a color and say all the edges in M_t have color t. So, the partition $\{M_t\}_{t \in K}$ gives a k-coloring of E_H.

(P3) For every vertex $v \in \mathcal{B}$, $K_v := \{t \in K : v \text{ is matched in } M_t\}$ is the set of colors of the incident edges of v. Notice that, for every $v \in \mathcal{B}$, we have $|K_v| = d'$, and that K_v's are determined by (P1) and (P2).

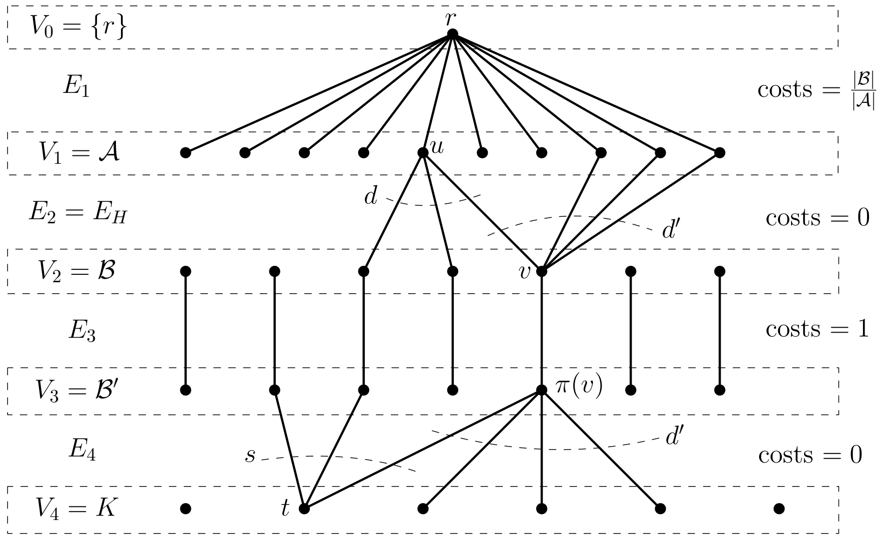

Figure 1: An illustration of the Zosin-Khuller type gap instance. All the edges go downwards. We only show some representative edges in E_2 and E_4. The set of colors of incoming edges of v is precisely the set out-neighbors of $\pi(v)$.

Given the above objects, the Zosin-Khuller type gap instance is defined over a 5-level directed graph $G = (V = V_0 \uplus V_1 \uplus V_2 \uplus V_3 \uplus V_4, E = E_1 \uplus E_2 \uplus E_3 \uplus E_4)$, where edges in E_i, $i \in \{1,2,3,4\}$ go from V_{i-1} to V_i. The vertices and edges in G are defined as follows. (See Figure 1 for an illustration of the instance.)

- $V_0 = \{r\}, V_1 = \mathcal{A}, V_2 = \mathcal{B}, V_3 = \mathcal{B}'$ and $V_4 = K$, where $\mathcal{B}'$ is a new set with the same cardinality as $\mathcal{B}$. Let $\pi : \mathcal{B} \to \mathcal{B}'$ be a bijection from $\mathcal{B}$ to $\mathcal{B}'$. We say that the vertex $\pi(v) \in \mathcal{B}'$ is the copy of the vertex $v \in \mathcal{B}$, and thus, $\mathcal{B}'$ is the copy of $\mathcal{B}$.
- $E_1 = \{(r,u) : u \in \mathcal{A}\}$, $E_2 = E_H$ (except that edges in E_2 are directed), $E_3 = \{(v, \pi(v)) : v \in \mathcal{B}\}$, and $E_4 = \{(\pi(v), t) : v \in \mathcal{B}, t \in K_v\}$.

So, in the graph G, we have edges from r to all vertices in $V_1 = \mathcal{A}$, the graph $(V_1 \cup V_2, E_2)$ is just the graph H, and $(V_2 \cup V_3, E_3)$ is the matching between $\mathcal{B}$ and its copies. The only non-straightforward element is E_4: If the vertex $v \in \mathcal{B}$ is incident to an edge with color $t \in K$ in H, then we have an edge from $\pi(v)$ to t in E_4. Therefore, the set of d' colors incident to a vertex $v \in \mathcal{B}$ in H is the same as the set of d' out-neighbors of $\pi(v)$ in G. The in-degree of a terminal $t \in K$ is s.

The root of the DST instance is r, and the terminals are $V_4 = K$. To complete the DST instance, it remains to define the costs of edges in G: Edges in E_1, E_2, E_3 and E_4 have costs $\frac{|\mathcal{B}|}{|\mathcal{A}|}, 0, 1$ and 0 respectively.

Given the Zosin-Khuller type instance, we can naturally define an LP solution $x \in [0,1]^E$, where every edge $e \in E$ has $x_e = \frac{1}{s}$. The cost of the LP solution is $\frac{1}{s} \cdot \Big(|E_1| \cdot \frac{|\mathcal{B}|}{|\mathcal{A}|} + |E_3| \cdot 1\Big) = \frac{1}{s} \cdot \Big(|\mathcal{A}| \cdot \frac{|\mathcal{B}|}{|\mathcal{A}|} + |\mathcal{B}| \cdot 1\Big) = \frac{2|\mathcal{B}|}{s}$. For every terminal $t \in K$, we can find s disjoint paths from r to t in G: For each $(u,v) \in M_t$, we take the path $r \to u \to v \to \pi(v) \to t$. Notice that the edge $(\pi(v), t)$ exists since $t \in K_v$. Therefore, the x induces a valid solution to (FLP).

The following lemma says that if the objects specified in (P1) to (P3) have a good property, then the DST instance has a large integrality gap.

Lemma 2. *Let $\alpha \geq 1$ be a real number. Suppose for every $u \in \mathcal{A}$, there exists some $J_u \subseteq K$ of terminals with $|J_u| \leq \frac{d}{\alpha}$, such that for every $(u,v) \in E_H$, we have $|K_v \setminus J_u| \leq \frac{d'}{\alpha}$. Then the optimum solution to the DST instance has cost at least $\frac{\alpha|\mathcal{B}|}{s}$.*

Proof. We can break the optimum directed Steiner tree T^* of G into many sub-trees, each containing exactly one edge in E_1. We show that any such sub-tree T' has a small *density*, which is defined as the number of terminals in T' divided by the cost of the edges in T'. In particular, we show its density is at most $\frac{d'}{\alpha}$. If this holds, then the optimum tree has cost at least $k/\frac{d'}{\alpha} = \alpha \cdot \frac{k}{d'} = \alpha \cdot \frac{|\mathcal{B}|}{s}$ as $ks = d'|\mathcal{B}| = |E_H|$.

So, we fix a sub-tree T' that contains exactly one edge in E_1, and it remains to show that the density of T' is at most $\frac{d'}{\alpha}$. Let u be the unique vertex in $V_1 = \mathcal{A}$ in the tree. Let V' be the set of vertices in V_2 in the tree; without loss of generality we assume that, for every $v \in V'$, we have $(v, \pi(v)) \in T'$ since, otherwise, we can remove (u, v) from the solution. So, each vertex in V' is a neighbor of u in H. The set of terminals that can be reached from V' is $\bigcup_{v\in V'} K_v$. Therefore, the number of terminals in the tree T' is at most

$$\left|\bigcup_{v\in V'} K_v\right| = \left|J_u \cup \bigcup_{v\in V'} (K_v \setminus J_u)\right| \le \frac{d}{\alpha} + \frac{d'}{\alpha}|V'|.$$

The cost of T' is exactly $\frac{|\mathcal{B}|}{|\mathcal{A}|} + |V'| = \frac{d}{d'} + |V'|$, and the density of the tree is at most $\dfrac{\frac{d}{\alpha} + \frac{d'}{\alpha}|V'|}{\frac{d}{d'} + |V'|} = \dfrac{d'}{\alpha}$. □

Therefore, if the condition in Lemma 2 holds, then the DST instance has integrality gap at least $\alpha/2$.

In the Zosin-Khuller gap instance in [ZK02], the objects from (P1) to (P3) are defined as follows. Let K be the set of k terminals, and assume $\sqrt{k}$ is an integer. We have $\mathcal{A} = \binom{K}{\sqrt{k}}$, $\mathcal{B} = \binom{K}{\sqrt{k}+1}$, and there is an edge from $A \in \mathcal{A}$ to $B \in \mathcal{B}$ in H if and only if $A \subseteq B$. The color of (A, B) is the unique terminal in $B \setminus A$. Notice that all the required properties are satisfied. $d = k - \sqrt{k}, d' = \sqrt{k} + 1$, and for a vertex $B \in \mathcal{B}$, we have $K_B = B$.

To satisfy the condition in Lemma 2, we define $J_A = A$ for every $A \in \mathcal{A}$. So, $|J_A| = \sqrt{k}$. Then, for every $B \in \mathcal{B}$ that is adjacent to A, we have $K_B \setminus J_A = B \setminus A$, which has size 1. Therefore, we can define $\alpha = \min\{\frac{d}{\sqrt{k}}, \frac{d'}{1}\} = \min\{\sqrt{k} - 1, \sqrt{k} + 1\} = \sqrt{k} - 1$ to make the condition of Lemma 2 holds. The instance gives a gap of $\Omega(\sqrt{k})$. However, it has size exponential in $\sqrt{k}$.

4 Instance with Polynomial Integrality Gap

The construction of our gap instance is similar to that of [ZK02] in the sense that our $\mathcal{A}$ and $\mathcal{B}$ correspond to subsets of a ground set. However, in our construction, a terminal is also correspondent to a subset of the ground set (as opposed to a single element). In the graph H, we have an edge from some element $A \in \mathcal{A}$ to some element $B \in \mathcal{B}$ if and only if $A \subseteq B$, and the color of the edge is the set $B \setminus A$, which will be an element in K. Using this construction, we make the number of terminals exponential in the size of the ground set. This can lead to a polynomial integrality gap. We carefully design the sizes of the subsets so that the conditions in Lemma 2 hold with a large α.

We formally define the objects specified in (P1) to (P3). Let $\rho = \frac{1}{16}$ and $\theta = \frac{1}{64}$. So $\theta = 4\rho^2$, and $\theta = \frac{\rho}{4}$. Let m be an integer multiply of 64 so that ρm and θm are integers. $[m]$ will be the ground set. Though ρ and θ are fixed, most of the time it is more convenient for us to keep the two notions, without replacing them with their values. The objects are defined as follows:

- $\mathcal{A} = K = \binom{[m]}{\rho m}, \mathcal{B} = \binom{[m]}{2\rho m}$, there is an edge from some $A \in \mathcal{A}$ to some $B \in \mathcal{B}$ in H if and only if $A \subseteq B$. The edge has color $B \setminus A \in K$.
- Therefore, $d = \binom{(1-\rho)m}{\rho m}, d' = \binom{2\rho m}{\rho m}, k = \binom{m}{\rho m}$ and $K_B = \{C \subseteq B : |C| = \rho m\}$ for every $B \in \mathcal{B}$.

It is easy to check that all the required properties are satisfied: $|\mathcal{A}| \le |\mathcal{B}|$, the graph H is bi-regular and all matchings have the same size, as H is highly symmetric.

Now we shall define the set J_A for any $A \in \mathcal{A}$ to satisfy the condition of Lemma 2. This is defined as follows:

$$J_A = \{C \in K : |C \cap A| > \theta m\}.$$

We then show that this definition satisfies the condition of Lemma 2 for some $\alpha = e^{\Omega(m)}$. From now on, we fix any $A \in \mathcal{A} = \binom{[m]}{\rho m}$.

Lemma 3. *There is an absolute constant $c > 1$ such that $|J_A| \leq \frac{d}{c^m}$.*

Proof. Let C be a set that is uniformly at random chosen from $K = \binom{[m]}{\rho m}$. Then, $\frac{|J_A|}{k} = \Pr[|C \cap A| > \theta m]$. Define $\mu := \mathbb{E}[|C \cap A|] = \rho m \cdot \frac{\rho m}{m} = \rho^2 m < \theta m$. Notice that the events $\{(i \in C)\}_{i \in A}$ are negatively correlated. Then, applying Chernoff bound for negatively correlated random variables (see Theorem 5), we have

$$\begin{aligned}
\Pr[|C \cap A| > \theta m] &= \Pr\left[|C \cap A| > \left(1 + \frac{\theta m}{\mu} - 1\right)\mu\right] \\
&\leq \exp\left(-\frac{(\theta m/\mu - 1)^2 \mu}{2 + \theta m/\mu - 1}\right) = \exp\left(-\frac{\theta m/\mu - 1}{\theta m/\mu + 1} \cdot (\theta m - \mu)\right) \\
&= \exp\left(-\frac{\theta - \rho^2}{\theta + \rho^2} \cdot (\theta - \rho^2) m\right) = \exp\left(-\frac{9\rho^2 m}{5}\right).
\end{aligned}$$

The last equality used that $\theta = 4\rho^2$. On the other hand, we have

$$\begin{aligned}
\frac{k}{d} = \binom{m}{\rho m} \Big/ \binom{(1-\rho)m}{\rho m} &= \prod_{i=0}^{\rho m - 1} \frac{m - i}{(1-\rho)m - i} \leq \left(\frac{1-\rho}{1-2\rho}\right)^{\rho m} \\
&\leq \exp\left(\frac{\rho}{1-2\rho} \cdot \rho m\right) = \exp\left(\frac{\rho^2 m}{1-2\rho}\right) = \exp\left(\frac{8\rho^2 m}{7}\right).
\end{aligned}$$

Therefore,

$$\frac{|J_A|}{d} = \frac{k}{d} \cdot \Pr[|C \cap A| > \theta m] \leq \exp\left(-\frac{9\rho^2 m}{5}\right) \cdot \exp\left(\frac{8\rho^2 m}{7}\right) = \exp\left(-\frac{23}{35}\rho^2 m\right).$$

The lemma then follows by letting $c = \exp\left(\frac{23}{35}\rho^2\right) > 1$. □

Lemma 4. *For any $B \in \mathcal{B}$ such that $A \subseteq B$, we have $|K_B \setminus J_A| \leq \frac{d'}{c'^m}$ for some absolute constant $c' > 1$.*

Proof. Notice that $\frac{|K_B \setminus J_A|}{d'} = \frac{|K_B \setminus J_A|}{|K_B|}$, which is exactly the probability that a randomly chosen element in K_B is not in J_A. Let C be a random subset of B of size ρm. So, the quantity is the probability that $|C \cap A| \leq \theta m$. Notice that we have $\mu := \mathbb{E}[|C \cap A|] = \rho m \cdot \frac{\rho m}{2\rho m} = \frac{\rho m}{2} > \theta m$. Applying the lower tail Chernoff bound (see Theorem 5) on negatively correlated variables, we obtain

$$\Pr[|C \cap A| \leq \theta m] = \Pr\left[|C \cap A| \leq (1 - 1/2)\mu\right] \leq \exp\left(-\frac{(1/2)^2}{2}\mu\right) = \exp\left(-\frac{\mu}{8}\right) = \exp\left(-\frac{\rho m}{16}\right) = e^{-\frac{m}{256}}.$$

The first equality follows because $\mu = \frac{\rho m}{2} = 2\theta m$. Thus, the lemma holds with $c' = e^{\frac{1}{256}}$. □

Therefore, the condition of Lemma 2 holds with $\alpha = \min\{c^m, c'^m\} \geq e^{\Omega(m)}$. Let n be the number of vertices in the constructed DST instance. Then we have $n \leq e^{O(m)}$. Therefore, we have $\alpha = n^{\Omega(1)}$. This finishes the proof of Theorem 1.

5 Future Directions

Our result rules out a polynomial-time poly-logarithmic approximation for DST using the flow LP relaxation. Our instance is on a 5-level graph, and thus it admits a poly-logarithmic approximation via recursive greedy [CCGG98] or LP hierarchies [Rot11, FKK+14]. One interesting future direction is to understand the power of $O(1)$-level lift of the flow LP using some hierarchy. That is, whether the lifted LP still has a polynomial integrality gap. To show a lower bound, we may need to lift our gap instance for the basic flow LP in some manner.

Acknowledgment We thank Jittat Fakcharoenphol for the useful discussions.

References

[BGRS10] Jaroslaw Byrka, Fabrizio Grandoni, Thomas Rothvoß, and Laura Sanità. An improved lp-based approximation for steiner tree. In *Proceedings of the Forty-Second ACM Symposium on Theory of Computing*, STOC '10, page 583–592, New York, NY, USA, 2010. Association for Computing Machinery.

[BP89] M. Bern and P. Plassmann. The steiner problem with edge lengths 1 and 2. *Inf. Process. Lett.*, 32:171–176, 1989.

[CC08] Miroslav Chlebík and J. Chlebíková. The steiner tree problem on graphs: Inapproximability results. *Theor. Comput. Sci.*, 406:207–214, 2008.

[CCC+99] Moses Charikar, Chandra Chekuri, To-Yat Cheung, Zuo Dai, Ashish Goel, Sudipto Guha, and Ming Li. Approximation algorithms for directed steiner problems. *J. Algorithms*, 33(1):73–91, 1999.

[CCGG98] Moses Charikar, Chandra Chekuri, Ashish Goel, and Sudipto Guha. Rounding via trees: Deterministic approximation algorithms for group steiner trees and k-median. In *Proceedings of the Thirtieth Annual ACM Symposium on the Theory of Computing, Dallas, Texas, USA, May 23-26, 1998*, pages 114–123, 1998.

[FKK+14] Zachary Friggstad, Jochen Könemann, Young Kun-Ko, Anand Louis, Mohammad Shadravan, and Madhur Tulsiani. Linear programming hierarchies suffice for directed steiner tree. In *Integer Programming and Combinatorial Optimization - 17th International Conference, IPCO 2014, Bonn, Germany, June 23-25, 2014. Proceedings*, pages 285–296, 2014.

[GLL19] Fabrizio Grandoni, Bundit Laekhanukit, and Shi Li. $O(\log^2 k / \log\log k)$-approximation algorithm for directed steiner tree: a tight quasi-polynomial-time algorithm. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing, STOC 2019, Phoenix, AZ, USA, June 23-26, 2019*, pages 253–264, 2019.

[GN20] Rohan Ghuge and Viswanath Nagarajan. Quasi-polynomial algorithms for submodular tree orienteering and other directed network design problems. In *To appear in SODA'20.*, 2020.

[HK03] Eran Halperin and Robert Krauthgamer. Polylogarithmic inapproximability. In *Proceedings of the 35th Annual ACM Symposium on Theory of Computing, June 9-11, 2003, San Diego, CA, USA*, pages 585–594, 2003.

[HKK+07] Eran Halperin, Guy Kortsarz, Robert Krauthgamer, Aravind Srinivasan, and Nan Wang. Integrality ratio for group steiner trees and directed steiner trees. *SIAM J. Comput.*, 36(5):1494–1511, 2007.

[KP99] Guy Kortsarz and David Peleg. Approximating the weight of shallow steiner trees. *Discrete Applied Mathematics*, 93(2-3):265–285, 1999.

[Mos15] Dana Moshkovitz. The projection games conjecture and the np-hardness of ln n-approximating set-cover. *Theory of Computing*, 11:221–235, 2015.

[PS97] Alessandro Panconesi and Aravind Srinivasan. Randomized distributed edge coloring via an extension of the chernoff-hoeffding bounds. *SIAM J. Comput.*, 26(2):350–368, 1997.

[Rot11] Thomas Rothvoß. Directed steiner tree and the lasserre hierarchy. *CoRR*, abs/1111.5473, 2011.

[Zel97] Alexander Zelikovsky. A series of approximation algorithms for the acyclic directed steiner tree problem. *Algorithmica*, 18(1):99–110, 1997.

[ZK02] Leonid Zosin and Samir Khuller. On directed steiner trees. In *Proceedings of the Thirteenth Annual ACM-SIAM Symposium on Discrete Algorithms, January 6-8, 2002, San Francisco, CA, USA.*, pages 59–63, 2002.

A Chernoff Bounds for Negatively Correlated Random Variables

A set of 0/1 random variables $\{X_1, X_2, \cdots, X_n\}$ are said to be negatively correlated if for any subset $S \subseteq [n]$ and any $b \in \{0, 1\}$, it holds that

$$\Pr\left[\bigwedge_{i\in S}(X_i = b)\right] \le \prod_{i\in S}\Pr[X_i = b].$$

Theorem 5 (Theorem 3.2 in [PS97])**.** *Let $X_1, X_2, \cdots, X_n$ be a set of negatively correlated 0/1 random variables. Let $\mu = \mathbb{E}\left[\sum_{i=1}^n X_i\right]$. Then for any $\delta \in (0, 1)$, we have*

$$\Pr\left[\sum_{i=1}^n X_i \ge (1+\delta)\mu\right] \le e^{-\frac{\delta^2\mu}{2+\delta}}, \text{ and } \Pr\left[\sum_{i=1}^n X_i \le (1-\delta)\mu\right] \le e^{-\frac{\delta^2\mu}{2}}.$$