On some network design problems with degree constraints*

Rohit Khandekar
IBM T.J.Watson Research Center
rohitk@us.ibm.com

Guy Kortsarz†
Rutgers University, Camden
guyk@camden.rutgers.edu

Zeev Nutov
The Open University of Israel
nutov@openu.ac.il

Abstract

We study several network design problems with degree constraints. For the minimum-cost Degree-Constrained 2-Node-Connected Subgraph problem, we obtain the first non-trivial bicriteria approximation algorithm, with factor 6 violation for the degrees and a 4-approximation for the cost. This improves upon the logarithmic degree violation and no cost guarantee obtained by Feder, Motwani, and Zhu (2006). Then we consider the problem of finding an arborescence with at least k terminals and with minimum maximum outdegree. We show that the natural LP-relaxation has a gap of $\Omega(\sqrt{k})$ or $\Omega(n^{1/4})$ with respect to the multiplicative degree bound violation. We overcome this hurdle by a *combinatorial* $O(\sqrt{(k \log k)/\Delta^*})$-approximation algorithm, where Δ^* denotes the maximum degree in the optimum solution. We also give an $\Omega(\log n)$ lower bound on approximating this problem. Then we consider the undirected version of this problem, however, with an extra diameter constraint, and give an $\Omega(\log n)$ lower bound on the approximability of this version. We also consider a closely related Prize-Collecting Degree-Constrained Edge-Connectivity Survivable Network problem, and obtain several results in this direction by reducing the prize-collecting variant to the regular one. Finally, we consider a simple variant of the Degree-Constrained Steiner Tree problem when some terminals are required to be leaves. We show that this seemingly simple problem is equivalent to the Group Steiner Tree problem.

1 Introduction

1.1 Problems considered

In network design problems one seeks a cheap subgraph H of a given graph G that satisfies some given properties. In the b-Matching problem H should satisfy prescribed degree constraints, while in the Survivable Network problem H should satisfy prescribed connectivity requirements. The Degree-Constrained Survivable Network problems is a combination of these two fundamental problems, where H should satisfy both degree constraints and connectivity requirements. For most of these problems, even checking whether there exists a feasible solution is NP-hard, hence one considers a bicriteria approximation when the degree constraints are relaxed. Namely, the goal is to compute a cheap solution that satisfies the connectivity requirements and has small degree violation.

Many recent papers considered *edge-connectivity* Degree-Constrained Survivable Network problems, see a recent survey in [20]. Our first problem is the simplest *node-connectivity* problem. A graph H is k-(node-)connected if it contains k internally disjoint paths between every pair of its nodes. In the k-Connected Subgraph problem we are given a graph $G = (V, E)$ with edge-costs and an integer k. The

*A preliminary version of this paper is [15].

†Partially supported by NSF Award Grant number 0829959.

goal is to find a minimum-cost k-connected spanning subgraph H of G. In the Degree-Constrained k-Connected Subgraph problem, we are also given degree bounds $\{b(v) : v \in V\}$. The goal is to find a minimum-cost k-connected spanning subgraph H of G such that in H, the degree of every node v is at most $b(v)$. We consider the case $k = 2$.

Degree-Constrained 2-Connected Subgraph
Instance: An undirected graph $G = (V, E)$ with non-negative edge-costs $\{c_e : e \in E\}$, and degree bounds $\{b(v) : v \in V\}$.
Objective: Find a minimum cost 2-connected spanning subgraph H of G that satisfies the *degree constraints* $\deg_H(v) \leq b(v)$ for all $v \in V$.

In the Steiner k-Tree problem one seeks a minimum-cost tree that contains at least k-terminals (when every node is a terminal we get the k-MST problem). Our next problem is the minimum-degree directed version of this problem. Given a directed graph G, a set S of terminals, and an integer $k \leq |S|$, a *k-arborescence* is an arborescence in G that contains k terminals; in the case of undirected graphs we have a *k-tree*. For a directed/undirected graph or edge-set H let $\Delta(H)$ denote the maximum outdegree/degree of a node in H.

Minimum Degree k-Arborescence
Instance: A directed graph $G = (V, E)$, a root $s \in V$, a subset $S \subseteq V \setminus \{s\}$ of terminals, and an integer $k \leq |S|$.
Objective: Find in G a k-arborescence T rooted at s that minimizes $\Delta(T)$.

The origin of this problem is in peer-to-peer networking, when one wants to bound the maximum load (degree) of a node, while connecting the root to the maximum number of terminals. It is also of interest to bound the height of such a tree, to limit the time for sending messages from the root. This motivates our next problem, for which we only show a lower bound. Hence we show it for the *less* general case of undirected graphs.

Degree and Diameter Bounded k-Tree
Instance: An undirected graph $G = (V, E)$, a subset $S \subseteq V$ of terminals, an integer $k \leq |S|$, and a diameter bound D.
Objective: Find a k-tree T with diameter bounded by D that minimizes $\Delta(T)$.

Let $\lambda_H(u, v)$ denote the the maximum number of edge-disjoint uv-paths in H. In the Edge-Connectivity Survivable Network problem we are given a graph $G = (V, E)$ with edge-costs $c_e \geq 0$, a collection $\mathcal{P} = \{\{u_1, v_1\}, \ldots, \{u_k, v_k\}\}$ of node pairs, and connectivity requirements $\mathcal{R} = \{r_1, \ldots, r_k\}$. The goal is to find a minimum-cost subgraph H of G that satisfies the connectivity requirements $\lambda_H(u_i, v_i) \geq r_i$ for all i.

We consider a combination of the following two generalizations of this problem. In Degree-Constrained Edge-Connectivity Survivable Network, we are given degree bounds $\{b(v) : v \in V\}$. The goal is to find a minimum-cost subgraph H of G that satisfies the connectivity requirements and the degree constraints $\deg_H(v) \leq b(v)$ for all $v \in V$. In the Prize-Collecting Edge-Connectivity Survivable Network we are given a submodular monotone non-decreasing penalty function $\pi : 2^{\{1,\ldots,k\}} \to \mathbb{R}_+$ (π is given by an evaluation oracle). The goal is to find a subgraph H of G that minimizes the *value* $\texttt{val}(H) = c(H) + \pi(\texttt{unsat}(H))$ of H, where $\texttt{unsat}(H) = \{i \mid \lambda_H^S(u_i, v_i) < r_i\}$ is the set of requirements *not* (completely) satisfied by H. Formally, the problem we consider is as follows.

Prize-Collecting Degree-Constrained Edge-Connectivity Survivable Network
Instance: A graph $G = (V, E)$ with non-negative edge-costs $\{c_e : e \in E\}$, a collection $\mathcal{P} = \{\{u_1, v_1\}, \ldots, \{u_k, v_k\}\}$ of node pairs, connectivity requirements $\mathcal{R} = \{r_1, \ldots, r_k\}$, a submodular monotone non-decreasing penalty function $\pi : 2^{\{1,\ldots,k\}} \to \mathbb{R}_+$ given by an evaluation oracle, and degree bounds $\{b(v) : v \in V\}$.
Objective: Find a subgraph H of G that satisfies the *degree constraints* $\deg_H(v) \le b(v)$ for all $v \in V$, and minimizes the *value*

$$\mathtt{val}(H) = c(H) + \pi(\mathtt{unsat}(H))$$

of H, where $\mathtt{unsat}(H) = \{i \mid \lambda^S_H(u_i, v_i) < r_i\}$ is the set of requirements *not* satisfied by H.

The Steiner Tree problem is a particular case of this problem, when we seek a minimum-cost subtree T of G that contains a specified subset S of terminals. In the degree constrained version of Steiner Tree, we are also given degree bounds on nodes in S and need to satisfy the degree constraints. We consider the case of $\{0, 1\}$-constraints, namely, we require that certain nodes in S should be leaves of T, and do not allow to relax this condition, as was done in previous papers [19, 21, 22, 1]. Namely, the degree bounds here are of the "hard capacity" type, and cannot be violated. Formally, our problem can be casted as follows.

Leaf-Constrained Steiner Tree
Instance: A graph $G = (V, E)$ with edge-costs $c_e \ge 0$ and subsets $L \subseteq S \subseteq V$.
Objective: Find a minimum-cost tree T in G that contains S such that every $v \in L$ is a leaf of T.

It was brought to our attention by Magnus Halldorsson, in a personal communication, that this problem has an important application in numerical computation of matrices entries with partial knowledge. The details of the application are rather technical, and thus are omitted.

1.2 Previous and related work

Fürer and Raghavachari [7] were the first to consider degree-constrained connectivity problems. They gave a "plus 1" approximation for the Minimum Degree Steiner Tree problem. Namely, if the lowest maximum degree possible is Δ^*, their algorithm returns a Steiner tree with maximum degree $\Delta^* + 1$. This result is the best possible, as computing an optimal solution is NP-hard even in Minimum Degree Spanning Tree case.

The first result for the min-cost case is due to Ravi et al. [23]; they obtained an $(O(\log n) \cdot b(v), O(\log n))$-approximation for Degree-Constrained MST, namely, the degree of every node v in the output tree is $O(\log n) \cdot b(v)$ while its cost is $O(\log n)$ times the optimal cost. A major breakthrough was obtained by Goemans [9]; his algorithm computes a minimum cost spanning tree with degree at most $\Delta + 2$, with Δ the minimum possible degree.

In [16] and [27] is given an $O(n^\delta)$-approximation algorithm for the Minimum Degree k-Edge-Connected Subgraph problem, for any fixed $\delta > 0$.

It turned out that an extension of the iterative rounding method of Jain [14] may be the leading technique for degree-constrained problems. Singh and Lau [29] were the first to extend this method to achieve the best possible result for Min-Cost Minimum Degree MST; their tree has optimal cost while the maximum degree is at most $\Delta+1$. Lau et al. [19] obtained a $(2b(v)+3, 2)$-approximation for the *edge-connectivity* Degree-Constrained Survivable Network problem, which was recently improved to $(2b(v)+2, 2)$ in [22]. Lau and Singh [21] further obtained a $(b(v)+O(r_{\max}), 2)$-approximation, where $r_{\max}$ denotes the maximum connectivity requirement.

For directed graphs, Bansal et al. [1] gave an $(\lceil \frac{b(v)}{1-\epsilon} \rceil + 4, \frac{1}{\epsilon})$-approximation scheme for the Degree-Constrained k-Edge-Outconnected Subgraph problem; the case $k=1$ is the Degree-Constrained Arborescence problem, for which [1] gave a $b(v)+2$-approximation, without bounding the cost. Some extensions and slight improvements can be found in [25].

Note that all the above results are for *edge-connectivity* Survivable Network problems. The only known result for *node-connectivity* degree-constrained problems is by Feder, Motwani, and Zhu [6] who gave an algorithm that computes in $n^{O(k)}$ time a k-connected spanning subgraph H of G such that $\deg_H(v) = O(\log n) \cdot b(v)$. Their algorithm cannot handle costs.

The special case $k = |S|$ of the Minimum Degree k-Arborescence problem was already studied in [5], where a $\tilde{O}(\sqrt{k})$ *additive* approximation was given. Their technique does not seem to extend to the case $k < |S|$. Even for the easier undirected case, if we ask for a tree containing k nodes and want to minimize the maximum degree (this is the Degree Bounded k-MST problem), the above techniques of [5] seem to fail.

Hajiaghayi and Nasri [11] obtained a constant ratio for a very special case of Degree-Constrained Prize-Collecting Edge-Connectivity Survivable Network problem when the penalty function π is modular.

1.3 Our results and techniques

Recall that for the Degree-Constrained k-Connected subgraph problem, Feder, Motwani, and Zhu [6] gave an algorithm that computes in $n^{O(k)}$ time a k-connected spanning subgraph H of G such that $\deg_H(v) = O(\log n) \cdot b(v)$, and that their algorithm cannot handle costs. Our first result significantly improves their result for $k=2$, from logarithmic factor degree violation to constant factor violation. Furthermore, we are also able to bound the cost.

Theorem 1.1 *The* Degree-Constrained 2-Connected Subgraph *problem admits a bicriteria* $(6b(v)+6, 4)$*-approximation algorithm; namely, a polynomial time algorithm that computes a* 2*-connected spanning subgraph* H *of* G *in which the degree of every node* v *is at most* $6b(v)+6$*, and the cost of* H *is at most* 4 *times the optimal cost.*

To prove Theorem 1.1 we first compute a degree-constrained spanning tree J with $+1$ degree violation using the algorithm of [29]. Then we compute an augmenting edge-set I such that $J \cup I$ is 2-connected, using the iterative rounding method. To apply this method for degree constrained problems, one proves that any basic LP-solution $x > 0$ has an edge e with hight x_e value, or there exists a node $v \in B$ such that $\deg_E(v)$ is close to $b(v)$. Otherwise, one shows a contradiction using the so called "token assignment argument". Here one shows that there exists a laminar family $\mathcal{L}$ of "violated sets" and a set T of nodes, such that x is the unique solution to the equation system defined by cut-constraints of sets in $\mathcal{L}$ and degree constraints of nodes in T. The contradiction is obtained by showing that the number of entries in x is strictly larger than $|\mathcal{L}| + |T|$. All previous "token assignment arguments" associated every node with a *unique* set in the laminar family $\mathcal{L}$. However, even in the simplest node-connectivity problem of augmenting a tree to be 2-connected, this is not possible, as the cut-nodes of the tree are associated with many sets in $\mathcal{L}$. We will allow for a node to be "shared" by many members of $\mathcal{L}$, and still will be able to distribute the tokens to obtained the desired contradiction.

Our second result gives the first approximation algorithm for the Minimum Degree k-Arborescence problem.

Theorem 1.2 *The* Minimum Degree k-Arborescence *problem admits an approximation algorithm with ratio $O(\sqrt{(k \log k)/\Delta^*})$, where Δ^* is the optimal solution value, namely, the minimal maximum out-degree possible. Furthermore, the problem admits no $o(\log n)$-approximation, unless* NP=Quasi(P).

Our algorithm for the Minimum Degree k-Arborescence problem uses a new method, which might be useful for related problems. We show that any k-arborescence with maximum degree Δ^* admits a "balanced partition" into roughly $\sqrt{k} \cdot \sqrt{\Delta^*}$ edge-disjoint arborescences, each containing at most $\sqrt{k} \cdot \sqrt{\Delta^*}$ terminals. We find iteratively, via max-flow computations, trees that contain $\sqrt{k} \cdot \Delta^*$ terminals. This will create many separate trees, that should be connected to the root. Thus, we have to show that there will be not too many such trees. We prove this by using the fact that the flow computation problem is non decreasing and submodular and so admits a $O(\log n)$-approximation [30].

Integrality gap of the natural LP relaxation for Minimum Degree k-MST.

To get some indication that the problem might be hard even on undirected graphs, consider the following natural LP-relaxation for Minimum Degree k-MST. The intended integral solution has $y_v = 1$ for nodes picked in the optimum tree T^*, $x_e = 1$ for $e \in T^*$, and d equal to the maximum degree of T^*.

$$\begin{array}{llcll}
\text{Minimize} & d & & & \\
\text{Subject to} & \sum\limits_{v \neq r} y_v & \geq & k & \\
& \sum\limits_{e \in \delta(S)} x_e & \geq & y_v & \forall v \in V \setminus \{r\} \;\; \forall S \subset V, r \in S, v \notin S \\
& \sum\limits_{e \in \delta(v)} x_e & \leq & d & \forall v \in V \\
& 0 \leq x_e, y_v & \leq & 1 & \forall e \in E \;\; \forall v \in V
\end{array} \tag{1}$$

We show that this LP-relaxation has integrality gap $\Omega(\sqrt{k})$ or $\Omega\left(n^{1/4}\right)$ where $n = |V|$. This holds even for the undirected case. Consider a rooted at r complete Δ-ary tree T of height h and let $k = \lfloor (\Delta + \Delta^2 + \cdots + \Delta^h)/(\Delta + 1) \rfloor$. It is easy to see that giving $x_e = 1/(\Delta+1)$ to all the edges $e \in T$ and $y_v = 1/(\Delta + 1)$ to all nodes $v \neq r$ satisfies all the constraints with fractional objective value $d = 1$. In order to cover k nodes, any integral tree however has to have a maximum degree of at least δ where $\delta + \delta(\delta-1) + \delta(\delta-1)^2 + \cdots + \delta(\delta-1)^{h-1} \geq k$. Such δ satisfies $\delta = \Omega(k^{1/h})$. Thus the optimum integral tree must have maximum degree $\Omega\left(k^{1/h}\right)$ giving an integrality gap of $\Omega\left(k^{1/h}\right)$. If we let $h = 2$, we get that $k = \Delta$ and $n = 1 + \Delta + \Delta^2$ and the integrality gap is $\Omega(\sqrt{\Delta})$ which is $\Omega(\sqrt{k})$ or $\Omega(n^{1/4})$.

Theorem 1.3 *The* Degree and Diameter Bounded k-Tree *problem admits no $o(\log n)$-approximation algorithm, unless* NP=Quasi(P)*. This is so even for the special case in which some optimal solution tree has diameter* 4.

Let $\delta_F(A)$ denote the set of edges in F between A and $V \setminus A$. For $i \in K$ let $A \odot i$ denote that $|A \cap \{u_i, v_i\}| = 1$. Menger's Theorem for edge-connectivity (see [18]) states that for a node pair u_i, v_i of a graph $H = (V, E)$ we have $\lambda_H(u_i, v_i) = \min\limits_{A \odot i} |\delta_E(A)|$. Hence if $\lambda_H(u_i, v_i) \geq r_i$ for a graph $H = (V, E)$, then for any A with $A \odot i$ we must have $|\delta_E(A)| \geq r_i$. A standard "cut-type"

LP-relaxation for Degree-Constrained Edge-Connectivity Survivable Network problem is as follows.

$$\begin{array}{lrcll} \text{Minimize} & \sum_{e \in E} c_e x_e & & & \\ \text{Subject to} & \sum_{e \in \delta_E(A)} x_e & \geq & r_i(A) & \forall i \;\; \forall A \odot i \\ & \sum_{e \in \delta_E(v)} x_e & \leq & b(v) & \forall v \\ & x_e & \in & [0,1] & \forall e \end{array} \tag{2}$$

Theorem 1.4 *Suppose that for an instance of* Prize-Collecting Degree-Constrained Edge-Connectivity Survivable Network *for any $\mathcal{P}' \subseteq \mathcal{P}$ the following holds. For an instance of* Degree-Constrained Edge-Connectivity Survivable Network *defined by $\mathcal{P}'$ there exists a polynomial-time algorithm that computes a subgraph H' of cost at most ρ times the optimal value of LP (2) with requirements restricted to $\mathcal{P}'$ such that $\deg_{H'}(v) \leq \alpha b(v) + \beta$ for all $v \in V$. Then we can compute in polynomial time numbers c^* and π^* with $c^* + \pi^* \leq \mathsf{opt}$, such that for any $\mu \in (0,1)$,* Prize-Collecting Degree-Constrained Edge-Connectivity Survivable Network *admits a polynomial time algorithm that computes a subgraph H such that $\mathtt{val}(H) \leq \frac{\rho}{1-\mu} c^* + \frac{1}{\mu}\pi^*$ and $\deg_H(v) \leq \frac{\alpha}{1-\mu} b(v) + \beta$ for all $v \in V$.*

The above theorem can be used along with the following known results. Louis and Vishnoi [22] obtain $\rho = 2, \alpha = 2, \beta = 2$ for Degree-Constrained Edge-Connectivity Survivable Network. Lau and Singh [21] obtain $\rho = 2, \alpha = 1, \beta = 3$ for Degree-Constrained Steiner Forest and $\rho = 2, \alpha = 1, \beta = 6r_{\max} + 3$ for Degree-Constrained Edge-Connectivity Survivable Network where $r_{\max}$ is the maximum requirement.

In the Group Steiner Tree problem we are given a collection $\mathcal{S}$ of node-subsets (groups), and seek a minimum-cost subtree T of G that contains at least one node from each group. Group Steiner Tree admits ratio $O(\log n \log |\mathcal{S}| \log \mathcal{S}_{\max})$ [8], where $\mathcal{S}_{\max} = \max_{S \in \mathcal{S}} |S|$. Group Steiner Tree with G being a tree admits no $\Omega(\log^{2-\epsilon} n)$ ratio, unless NP has quasi-polynomial Las-Vegas algorithms [12]. Our last result shows that Leaf-Constrained Steiner Tree and Group Steiner Tree are essentially equivalent w.r.t. approximation.

Theorem 1.5

(i) *If* Group Steiner Tree *admits approximation ratio $\rho(|V|, |\mathcal{S}|, \mathcal{S}_{\max})$ then* Leaf-Constrained Steiner Tree *admits ratio $\rho(|L||V|, |S|, |V|)$.*

(ii) *If* Leaf-Constrained Steiner Tree *admits ratio $\rho(|V|, |S|)$ then* Group Steiner Tree *admits ratio $\rho(|V| + |\mathcal{S}|, |\mathcal{S}|)$.*

Consequently, Leaf-Constrained Steiner Tree *admits ratio $O(\log^2 n \log |S|)$, and admits no $\Omega(\log^{2-\epsilon} n)$ ratio, unless NP has quasi-polynomial Las-Vegas algorithms.*

This is the only approximation algorithm for a degree-constrained connectivity problem that produces a solution without violating the degree bounds. It is not surprising that in our problem the degree bounds are 1, since in the case of degree bounds 2, even checking whether the problem admits a feasible solution is NP-hard (by a reduction to the Hamiltonian Path problem).

Theorems 1.1, 1.2, 1.3, 1.4, and 1.5, are proved in Sections 2, 3, 4, 5, and 6, respectively.

2 Degree-Constrained 2-Connected Subgraph (Theorem 1.1)

We start by considering the problem of augmenting a connected graph $J = (V, E_J)$ by a minimum-cost edge-set $I \subseteq E$ such that $\deg_I(v) \leq b(v)$ for all $v \in V$ and such that $J \cup I$ is 2-connected.

Definition 2.1 *For a node v of J let $\mu_J(v)$ be the number of connected components of $J \setminus \{v\}$; v is a* cut-node *of J if $\mu_J(v) \geq 2$.*

Note that $\mu_J(v) \leq \deg_J(v)$ for every $v \in V$. For $S \subseteq V$ let $\Gamma_J(S)$ denote the set of neighbors of S in J. Let s be a non-cut-node of J. A set $S \subseteq V \setminus \{s\}$ is *violated* if $|\Gamma_J(S)| = 1$ and $s \notin \Gamma_J(S)$. Let $\mathcal{S}_J$ denote the set of violated sets of J. Recall that $\delta_F(S)$ denotes the set of edges in F between S and $V \setminus S$. For $S \in \mathcal{S}_J$ let $\zeta_F(S)$ denote the set of edges in F with one endnode in S and the other in $V \setminus (S \cup \Gamma_J(S))$. By Menger's Theorem, $J \cup I$ is 2-connected if, and only if, $|\zeta_I(S)| \geq 1$ for every $S \in \mathcal{S}_J$. Thus a natural LP-relaxation for our augmentation problem is $\tau = \min\{c \cdot x : x \in P(J, b)\}$, where $P(J, b)$ is the polytope defined by the following constraints:

$$
\boxed{
\begin{array}{rcll}
x(\zeta_E(S)) & \geq & 1 & \text{for all } S \in \mathcal{S}_J \\
x(\delta_E(v)) & \leq & b(v) & \text{for all } v \in B \\
0 \leq x_e & \leq & 1 & \text{for all } e \in E
\end{array}
}
$$

Theorem 2.1 *There exists a polynomial time algorithm that given an instance of* Degree-Constrained 2-Connected Subgraph *and a connected spanning subgraph (V, J) of G computes an edge set $I \subseteq E \setminus J$ such that $c(I) \leq 3\tau$ and such that $\deg_I(v) \leq 2\mu_J(v) + 3b(v) + 3$ for all $v \in V$.*

Theorem 2.1 will be proved later. Now we show how to deduce the promised approximation ratio from it. Consider the following two phase algorithm.

Phase 1: With degree bounds $b(v)$, use the $(b(v) + 1, 1)$-approximation algorithm of Singh and Lau [29] for the Degree Constrained Steiner Tree problem to compute a spanning tree J in G.

Phase 2: Use the algorithm from Theorem 2.1 to compute an augmenting edge set I such that $H = J \cup I$ is 2-connected.

We prove the ratio. We have $c(J) \leq \mathsf{opt}$ and $c(I) \leq 3\tau$, hence $c(H) = c(J) + c(I) \leq 4\mathsf{opt}$. We now prove the approximability of the degrees. Let $v \in V$. Note that $\mu_J(v) \leq \deg_J(v) \leq b(v) + 1$. Thus we have

$$\deg_I(v) \leq 2\mu_J(v) + 3b(v) + 3 \leq 2(b(v) + 1) + 3b(v) + 3 = 5b(v) + 5.$$

This implies

$$\deg_H(v) \leq \deg_J(v) + \deg_I(v) \leq b(v) + 1 + 5b(v) + 5 = 6b(v) + 6.$$

In the rest of this section we will prove the following statement, that implies Theorem 2.1.

Lemma 2.2 *Let x be an extreme point of the polytope $P(J, b)$ with $0 < x_e < \frac{1}{3}$ for every $e \in E$. Then there is $v \in B$ such that $\deg_E(v) \leq 2\mu_J(v) + 3b(v) + 3$.*

Lemma 2.2 implies Theorem 2.1 as follows. Given a partial solution I and a parameter $\alpha \geq 1$, the residual degree bounds are $b_I^\alpha(v) = b(v) - \deg_I(v)/\alpha$. The following algorithm starts with $I = \emptyset$ and performs iterations. In every iteration, we work with the residual polytope $P(\mathcal{S}_{J \cup I}, b_I^\alpha)$, and remove some edges from E and/or some nodes from B, until E becomes empty. Let $\alpha = 3$ and

$\beta(v)=2\mu_J(v)+3$ for all $v \in B$. It is easy to see that for any edge-set $I \subseteq E$ we have $\mu_{J\cup I}(v) \leq \mu_J(v)$ for every $v \in V$.

Algorithm as in Theorem 2.1
Input: A connected graph (V,J), an edge-set E on V with costs $\{c_e : e \in E\}$, degree bounds $\{b(v) : v \in V\}$, and non-negative integers $\{\beta(v) : v \in V\}$.
Initialization: $I \leftarrow \emptyset$.
If $P(J,b)=\emptyset$, then return 'UNFEASIBLE' and STOP.
While $E \neq \emptyset$ do:
1. Find a basic solution $x \in P(\mathcal{S}_{J\cup I}, b_I^\alpha)$.
2. Remove from E all edges with $x_e=0$.
3. Add to I and remove from E all edges with $x_e \geq 1/\alpha$.
4. Remove from B every $v \in B$ with $\deg_E(v) \leq \alpha b_I^\alpha(v)+\beta(v)$.

EndWhile
Return I.

It is a routine to prove the following statement.

Lemma 2.3 *The above algorithm computes an edge set I such that $J \cup I$ is 2-connected, $c(I) \leq \alpha\tau$, and $\deg_J(v) \leq \alpha b(v)+\beta(v)$ for all $v \in B$.*

It remains to prove Lemma 2.2. The following statement is well known and can be proved using standard "uncrossing" methods.

Lemma 2.4 *The family $\mathcal{S}_J$ of violated sets is uncrossable, namely, for any $X,Y \in \mathcal{S}_J$ we have $X\cap Y, X\cup Y \in \mathcal{S}_J$ or $X\setminus Y, Y\setminus X \in \mathcal{S}_J$.*

Recall that a set-family $\mathcal{L}$ is *laminar* if for any distinct sets $X,Y \in \mathcal{L}$ either $X \subset Y$, or $Y \subset X$, or $X \cap Y=\emptyset$. Any laminar family $\mathcal{L}$ defines a partial order on its members by inclusion; we carry the usual notion of children, descendants, and leaves of laminar set families. The following statement is proved using a standard "uncrossing" argument.

Lemma 2.5 *For any basic solution $x \in P(J,b)$ with $0<x(e)<1$ for all $e \in E$, there exists a laminar family $\mathcal{L} \subseteq \mathcal{S}$ and $T \subseteq B$, such that x is the unique solution to the linear equation system:*

$$\begin{array}{rcll} x(\zeta_E(S)) & = & 1 & \text{for all } S \in \mathcal{L} \\ x(\delta_E(v)) & = & b(v) & \text{for all } v \in T \end{array}$$

Thus $|\mathcal{L}|+|T|=|E|$ and the characteristic vectors of $\{\zeta_E(S) : S \in \mathcal{L}\}$ are linearly independent.

Let x, $\mathcal{L}$, and T be as in Lemma 2.5. Let I' is the set of edges in E with exactly one endnode in B, I'' is set of edges in E with both end nodes in B, and $F=E\setminus(I'\cup I'')$.

Lemma 2.6 *Let $\{\beta(v) : v \in V\}$ be integers. Then there is $v \in B$ such that $\deg_E(v) \leq \alpha b(v)+\beta(v)$, if the following property holds:*

$$|\mathcal{L}| < \frac{1}{2}(\beta(B)-|B|)+\alpha x(I'')+\frac{1}{2}|I'|+\frac{1}{2}\alpha x(I')+|F| \tag{3}$$

Proof: Note that

$$\begin{aligned}\sum_{v\in B}(\deg_E(v)-\alpha b(v)) &\le \sum_{v\in B}(\deg_E(v)-\alpha x(\delta_E(v)))\\ &= \sum_{v\in B}(\deg_{I'}(v)+\deg_{I''}(v))-\alpha\sum_{v\in B}(x(\delta_{I'}(v))+x(\delta_{I''}(v)))\\ &= |I'|+2|I''|-\alpha x(I')-2\alpha x(I'')\ .\end{aligned}$$

Thus $|I'|+2|I''|-\alpha x(I')-2\alpha x(I'')<\beta(B)+|B|$ implies that $\deg_E(v)\le\alpha b(v)+\beta(v)$ for some $v\in B$. Adding $|I'|+2|F|$ to both sides gives $2(|I'|+|I''|+|F|)<\beta(B)+|B|+\alpha x(I')+2\alpha x(I'')+|I'|+2|F|$. Note that $|I'|+|I''|+|F|=|E|=|\mathcal{L}|+|T|$. Consequently, since $|T|\le|B|$, it is sufficient to prove that

$$2(|\mathcal{L}|+|B|)<\beta(B)+|B|+\alpha x(I')+2\alpha x(I'')+|I'|+2|F|\ .$$

Multiplying both sides by $\frac{1}{2}$ and rearranging terms gives (3). $\square$

Let us say that an edge e covers $S\in\mathcal{L}$ if e has one endnode in S and the other in $V\setminus(S\cup\Gamma_J(S))$. Given $S\in\mathcal{L}$ and an edge set E we will use the following notation.

- $\mathcal{C}$ is the set of children in $\mathcal{L}$ of S,
- E_S is the set of edges in E covering S but not a child of S,
- $E_{\mathcal{C}}$ is the set of edges in E covering some child of S but not S.

To show that there is $v\in B$ with $\deg_E(v)\le\alpha b(v)+\beta(v)$ for $\alpha=3$ and $\beta(v)=2\mu_J(v)+3$, we assign a certain amount of tokens to edges in E and nodes in B, such that the total amount of tokens does not exceed the right hand side of (3). A part of tokens assigned to an edge can be placed at some endnode or at the middle of the edge. A set $S\in\mathcal{L}$ gets the tokens placed at an endnode v of an edge e if $e\in E_S$ and $v\in S$, or if $e\in E_{\mathcal{C}}$ and v does not belong to a child of S. S gets the tokens placed at the middle of e if $e\in E_{\mathcal{C}}$. It is easy to verify that no two sets get the same token part of an edge. S gets also a token from $v\in B$ by the following rule.

Definition 2.2 *We say that $S\in\mathcal{L}$* owns *a node v if $v\in S$ but v is not in a child of S. We say that S* shares *v if $v\in\Gamma_J(S)$.*

Clearly, if $S\in\mathcal{L}$ owns v then no other set in $\mathcal{L}$ owns v. Note that if S shares v, then no ancestor or descendant of S in $\mathcal{L}$ shares v. This implies the following.

Lemma 2.7 *For any $v\in B$, the number of sets in $\mathcal{L}$ sharing v is at most $\mu_J(v)$.*

Thus if $\mu_J(v)+1$ tokens are assigned to v, then every set $S\in\mathcal{L}$ that owns or shares v can get 1 token from v. We will argue by induction that we can redistribute the tokens of S and its descendants in $\mathcal{L}$ such that every proper descendant of S in $\mathcal{L}$ gets at least 1 token and S gets 2 tokens. This differs from the usual token distribution in edge-connectivity problems, where nodes are owned but not shared. In node-connectivity, the cut-nodes may be shared by many members of $\mathcal{L}$, and in our case, by at most $\mu_J(v)$ members.

For $\beta(v)=2\mu_J(v)+3$ and $\alpha=3$, (3) becomes:

$$|\mathcal{L}|<\mu_J(B)+|B|+3x(I'')+\frac{1}{2}|I'|+\frac{3}{2}x(I')+|J|\ . \tag{4}$$

Initial token assignment (total amount of tokens $\leq$ the r.h.s of (4)):

- $\mu_J(v) + 1$ tokens to every $v \in B$,
- x_e tokens to each endnode in B of an edge e,
- $\frac{1}{2}$ token to each endnode in $V \setminus B$ of an edge e.

Lemma 2.8 *We can redistribute the tokens of S and its descendants in $\mathcal{L}$ such that every proper descendant of S in $\mathcal{L}$ gets at least 1 token and S gets 2 tokens.*

Proof: Since $0 < x_e < \frac{1}{3}$ for every $e \in E$, $|\zeta_E(S)| \geq 4$ for every $S \in \mathcal{L}$. Suppose that S is a leaf. If $S \cap B = \emptyset$, then S gets $\frac{1}{2}$ token from an endnode of every edge in $\zeta_E(S)$, and in total at least 2 tokens. If there is $v \in S \cap B$ then S owns v and gets 1 token from v. S gets $x(\zeta_E(S)) = 1$ tokens from edges in $\zeta_E(S)$. Consequently, S gets in total at least 2 tokens, as claimed. Now suppose that S is not a leaf. S gets $|\mathcal{C}|$ tokens from its children, hence if $|\mathcal{C}| \geq 2$ then we are done. Thus we are left with the case that S has a unique child C, and needs 1 token not from C. If S contains or shares some $v \in B$ then we are done. Otherwise, S gets $\frac{1}{2}$ token from the corresponding endnode of each edge in $E_S \cup E_{\mathcal{C}}$. By the linear independence an integrality of cuts $|E_S \cup E_{\mathcal{C}}| \geq 2$, hence S gets the desired token from the edges in $E_S \cup E_{\mathcal{C}}$. $\square$

The proof of Lemma 2.2, and thus also of Theorem 1.1, is now complete.

3 Minimum Degree k-Arborescence (Theorem 1.2)

We may assume that in the input graph G every node is reachable from the root s, that every terminal has indegree 1 and outdegree 0, and that the set of terminal of every arborescence T coincides with the set of leaves of T. Let $U = V \setminus S$. Before describing the algorithm, we need some definitions.

Definition 3.1 *For $W \subseteq U$ and an integer parameter $\alpha \geq 1$ the network $F_\alpha(W)$ with source s' and sink t' is obtained from G as follows.*

1. *Assign infinite capacity to every edge of G and capacity α to every node in U.*
2. *Add a new node s' and add new edges of capacity α each from s' to every node in W.*
3. *Add two new nodes t, t', add an edge of capacity 1 from every terminal to t, and add an edge of capacity k from t to t'.*

Our algorithm runs with an integer parameter α set eventually to

$$\alpha = \left\lceil \sqrt{k \cdot \Delta^* \cdot (\ln k + 1)} \right\rceil \ . \tag{5}$$

Although Δ^* is not known, $\Delta^* \leq k$, and our algorithm applies exhaustive search in the range $1, \ldots, k$.

Recall that a set-function ν defined on subsets of a ground-set U is *submodular* if $\nu(A) + \nu(B) \geq \nu(A \cup B) + \nu(A \cap B)$ for all $A, B \subseteq U$. Consider the following well known generic problem (for our purposes we state only the unweighted version).

Submodular Cover
Instance: A finite set U and a non-decreasing submodular function $\nu : 2^U \mapsto \mathbb{Z}$.
Objective: A minimum-size subset $W \subseteq U$ such that $\nu(W) = \nu(U)$.

The Submodular Cover Greedy Algorithm (for the unweighted version) starts with $W = \emptyset$ and while $\nu(W) < \nu(U)$ repeatedly adds to W an element $u \in U \setminus W$ that maximizes $\nu(W \cup \{u\}) - \nu(W)$. At the end, W is output. It is proved in [30] that the Greedy Algorithm for Submodular Cover has approximation ratio $\ln \max_{u \in U} \nu(\{u\}) + 1$.

A generalization of the following statement is proved in [2].

Lemma 3.1 ([2]) *For $W \subseteq U$ let $\nu_\alpha(W)$ be the maximum st-flow value in the network $F_\alpha(W)$. Then ν_α is non-decreasing and submodular, and $\nu_\alpha(U) \leq k$.*

The algorithm is as follows.

1. Execute the Submodular Cover Greedy Algorithm with $U = V \setminus S$ and with $\nu = \nu_\alpha$; let $W \subseteq U$ be the node-set computed.
2. Let f be a maximum integral flow in $F_\alpha(W)$ and let $J_W = \{e \in E : f(e) > 0\}$ be the set of those edges in E that carry a positive flow in $F_\alpha(W)$.
   Let T_W be an inclusion-minimal arborescence in G rooted at s containing W.
3. Return any k-arborescence contained in the graph $(V, J_W) \cup T_W$.

In the rest of this section we prove that the graph $(V, J_W) \cup T_W$ indeed contains a k-arborescence, and that for any integer $\alpha \geq 1$ it has maximum outdegree at most $\alpha + (\ln k + 1) \cdot k\Delta^*/\alpha$.

This implies the approximation ratio $\alpha/\Delta^* + (\ln k + 1) \cdot k/\alpha = O(\sqrt{(k \log k)/\Delta^*})$ for α given by (5).

Definition 3.2 *A collection $\mathcal{T}$ of sub-arborescence of an arborescence T is an* α-leaf-covering decomposition of T *if the arborescence in $\mathcal{T}$ are pairwise node-disjoint, every leaf of T belongs to exactly one of them, and each of them has at most α leaves.*

Lemma 3.2 *Suppose that G contains a k-arborescence that admits an α-leaf-covering decomposition $\mathcal{T}$. Let R be the set of roots of the arborescence in $\mathcal{T}$. Then $\nu_\alpha(R) = k$, and for the set W computed by the algorithm the following holds:*

(i) *$\nu_\alpha(W) = k$ and thus the graph $J_W \cup T_W$ contains a k-arborescence.*

(ii) *The graph $(V, J_W) \cup T_W$ has maximum outdegree at most $\alpha + |\mathcal{T}| \cdot (\ln k + 1)$.*

Proof: We prove that $\nu_\alpha(R) = k$. For a terminal v in T, let $r_v \in R$ be the root of the (unique) arborescence $T_v \in \mathcal{T}$ that contains v, and let P_v be the path in $F_\alpha(R)$ that consists of: the edge $s'r_v$, the unique path from r_v to v in T_v, and the edges vt' and $t't$. Let f be the flow obtained by sending for every terminal v of T one flow unit along P_v. Then f has value k, since T has k terminals. We verify that f obeys the capacity constraints in $F_\alpha(R)$. For every $r \in R$, the arborescence $T_r \in \mathcal{T}$ which root is r, has at most α terminals; hence the edge $s'r$ carries at most α flow units, which does not exceed its capacity α. This also implies that the capacity α on all nodes in U is met. For every terminal v of T, the edge vt carries one flow unit and has capacity 1. The edge $t't$ carries k flow units and has capacity k. Other edges have infinite capacity.

We prove (i). By Lemma 3.1, ν_α is non-decreasing and $\nu_\alpha(U) \leq k$. As $\nu_\alpha(U) \geq \nu_\alpha(R) = k$ and $\nu_\alpha(W) = \nu_\alpha(U)$, we conclude that $\nu(W) = k$. This implies that in the graph (V, J_W), k terminals are reachable from W, and (i) follows.

We prove (ii). In the graph (V, J_W), the outdegree of any node is at most α. This follows from the capacity α on any node in U. We have $|W| \leq |R| \cdot (\ln k + 1) = |\mathcal{T}| \cdot (\ln k + 1)$, by Lemma 3.2 and the approximation ratio of the Submodular Cover Greedy Algorithm. Since T_W is an arborescence with leaf-set W, the maximum outdegree of T_W is at most $|W| \leq |\mathcal{T}| \cdot (\ln k + 1)$. The statement follows. $\square$

The following lemma implies that the optimal tree T^* admits an α-leaf-covering decomposition $\mathcal{T}$ of size $|\mathcal{T}| \leq k\Delta^*/\alpha$ for any $\alpha \geq 1$. This together with Lemma 3.2 concludes the proof of Theorem 1.2.

Lemma 3.3 *Any arborescence T with k leaves and maximum outdegree Δ admits an α-leaf-covering decomposition $\mathcal{T}$ of size $|\mathcal{T}| \leq \Delta \cdot \lfloor k/(\alpha+1) \rfloor + 1$, for any integer $\alpha \geq 1$.*

Proof: For a node r of an arborescence T with root s let us use the following notation: T_r is the sub-arborescence of T with root r that contains all descendants of r in T, and P_r is the set of internal nodes in the ar-path in T. where a is the closest to r ancestor of r that has outdegree at least 2. Let us say that a node $u \in U$ of T is *α-critical* if T_u has more than α leaves, but no child of u has this property. It is easy to see that T has an α-critical node if, and only if, T has more than α leaves.

Consider the following algorithm. Start with $\mathcal{T} = \emptyset$. While T has an α-critical node u do the following: add T_r to $\mathcal{T}$ for every child r of u, and remove T_u and P_u from T (note that since we remove P_u no new leaves are created). When the while loop ends, if T is nonempty, add the remaining arborescence $T = T_s$ (which now has at most α leaves) to $\mathcal{T}$.

By the definition, the arborescence in $\mathcal{T}$ are pairwise node-disjoint, every leaf of T belongs to exactly one of them, and each of them has at most α leaves. It remains to prove the bound on $\mathcal{T}$. In the loop, when we consider an α-critical node u, at least $\alpha + 1$ leaves are removed from T and at most Δ arborescence are added to $\mathcal{T}$. Hence $|\mathcal{T}| \leq \Delta \cdot \lfloor k/(\alpha+1) \rfloor$ at the end of the loop. At most one additional arborescence is added to $\mathcal{T}$ after the loop. The statement follows. $\square$

Recall that we omit the $\Omega(\log n)$ lower bound for the problem, because the proof of this lower bound is basically identical to the proof of Theorem 1.3.

4 Hardness of Degree and Diameter Bounded k-Tree (Theorem 1.3)

4.1 The Min-Rep graph

For simplicity we mostly avoid the language of one-round two provers and stick to graph terminology which is more basic. The Min-Rep problem was introduced in [17]. This problem is a variant of the Label-Cover problem invented in [13], and it is equivalent to the Symmetric Label-Cover problem (see [4]).

The input to Min-rep consists of a bipartite graph $G = (V_1, V_2, E)$, and partitions $\tilde{V}_1$ and $\tilde{V}_2$ of V_1 and V_2 respectively, into q disjoint sets each, so that $V_1 = \bigcup_{i=1}^q A_i$, and $V_2 = \bigcup_{j=1}^q B_j$. The sets A_i for every i and B_j for every j are called *super-nodes*. If there exists $a_i \in A_i$ and $b_j \in B_j$ so that $(a_i, b_j) \in E$ we say that A_iB_j is a *super-edge*, or a *query*. The super-nodes and super-edges define a super-graph $\tilde{G} = (\tilde{V}_1, \tilde{V}_2, \tilde{E})$ with the A_i, B_j being the super-nodes and there is a super-edge A_iB_j iff A_i and B_j share and edge in $G(V_1, V_2, , E)$. A pair of nodes $a_i \in A_i, b_j \in B_j$ is said to *Min-Rep cover* the super-edge A_iB_j if $(a_i, b_j) \in E$. A subset $C \subseteq V_1 \cup V_2$ of nodes is said to Min-Rep cover a super-edge A_iB_j if there exists $a_i, b_j \in C$ that Min-Rep covers A_iB_j. A subset C that Min-Rep covers all the super-edges is called a *feasible solution* for the Min-Rep instance. A feasible subset $C \subseteq V_1 \cup V_2$ that satisfies $|C \cap A_i| = 1$ and $|C \cap B_j| = 1$ for every A_i and B_j is called a *perfect cover.* Note that a perfect cover has size $2q$.

This size of the Min-Rep graph and all its parameters is related to a central parameter ℓ (called the parallel repetition size) that determines the gap between a "yes" and a "no" instance in Min-Rep. We start with a 3-SAT-5 instance with n variables and m clauses (see [13]). In 3-SAT-5 every clause has 3 literals and every variable (literal) appears in exactly 5 clauses. The instance I of 3-SAT-5 is reduced to a Min-Rep instance so that the following holds.

Theorem 4.1 (The Parallel Repetition Theorem [28]) *For every ℓ there exists a reduction from* 3-SAT-5 *instance of size n to a instance of Min-Rep of size $M = O(n^{\mu \cdot \ell})$ for a universal constant μ, such that:*

1. *A* Min-Rep *instance that corresponds to a "yes" instance of I has a perfect cover.*
2. *Every feasible solution C that corresponds to a "no" instance I contains at least $(2q)2^{c\ell}$ sets (answer nodes) for some universal constant c.*

Note that, as there is $2q$ super-nodes, on average the number of answer nodes that need to be taken from every super-node is exponential in ℓ.

In this paper we set $\ell = \lceil \log \log n \rceil$. (All logs are to the base 2 unless it explicitly said otherwise).

4.2 Partition systems

Let $M' = \sqrt{M}$. For every super-edge A_iB_j we define a ground set M_{ij} of size M' (all pairwise disjoint). In addition we define an *anti-universal* set of partitions [24] of every M_{ij}. The anti-universal partition system is denoted by $C(M_{ij}, 7^\ell)$ and has the following properties [24].

Let each partition be denoted by Z^p_{ij}, $M_{ij} \setminus Z^p_{ij}$ so that $|Z^p_{ij}| = M'/2$ and p goes from $p = 1$ to $p = 7^\ell$. Let $\mathcal{C}$ be the collection of all sets in all partitions, $M_{i,j}$ (both $Z^p_{i,j}$ and $M - Z^p_{i,j}$ belong to $\mathcal{C}$ for every p.) Consider picking a collection of sets that belong to $\mathcal{C}$ under the following rule. For every $1 \le p \le k$ we may pick at most one of Z^k_{ij} or $M_{ij} - Z^k_{ij}$ (but not both). We call the above a legal choice of subsets of M_{ij}. We are concerned with the union of elements of these sets. Consider a particular legal choice $\bigcup_{p=1}^{k} X^p_{ij}$ with $X_{ij} = Z^p_{ij}$ or $X^p_{ij} = M_{ij} - Z^p_{ij}$. We call this particular choice of subsets of $\mathcal{C}$ *a set-cover* of M_{ij} if $\bigcup_{p=1}^{7^\ell} X^p_{ij} = M_{ij}$

Theorem 4.2 ([24]) *There exists a $C(M_{ij}, 7^\ell)$ structure so that for every legal choice of subsets in $\mathcal{C}$ the smallest number of sets in a legal choice that set-covers $M_{i,j}$ is at least $\alpha \log M'$ for some universal constant α.*

The above states that for a legal choice of subsets, you need about $\log M'$ sets, X^p_{ij} to cover $M_{i,j}$. The partition system behaves very much as if every $X^p_{i,j}$ set is a random half of $M_{i,j}$. Indeed in such a random partition, about $\log_2 M'$ $X^p_{i,j}$ sets are required to set-cover $M_{i,j}$.

4.3 A property of the Min-Rep graph for 3-SAT-5

We say that the Min-Rep instance obeys the *projection property* (or the *star property*) if for every super-edge $A_i \cup B_j$, the graph induced by $A_i + B_j$ is a set of edge disjoint stars with heads in A. The head of the star S is denoted $h(S)$. The b_j answer nodes in S are called the *leaves* of the star. The following is known to hold.

Theorem 4.3 [28] *The parallel repetition theorem starting with* 3-SAT-5 *gives a* Min-Rep *instance obeying the projection property.*

The size of every A_i (which is the number of stars of in the graph induced by $A_i \cup B_j$) is $|A_i| = 7^\ell$. See [13].

4.4 A variation on the [13] construction

The set-cover instance of [13] has $\bigcup A_i \cup \bigcup B_j$ as sets. Each super-edge $A_i B_j$ has a *ground set* $M_{i,j}$ of size $\sqrt{M}$ of elements (this is a slight change compared to [13]). Only $A_i \cup B_j$ sets can be connected (or contain) the element of the ground sets M_{ij}.

We explain how are the sets in $A_i \cup B_j$ connected to M_{ij}. Since the number of partitions in $C(M_{ij}, 7^\ell)$ exactly equals the number of stars in the graph induced by $A_i + B_j$ we may assume a bijection between stars in $A_i + B_j$ and anti-universal partitions. Every star S corresponds to a unique partition. A head $a = h(S)$ of a star is connected to $Z^p_{i,J}$ in M_{ij} and every leaf in the star of a is connected to the complement set $M_{i,j} \setminus Z^p_{i,j}$. Here $(Z^p_{i,j}, M_{ij} - Z^p_{ij})$ is the unique partition that corresponds to a.

4.5 The size of the parameters

Assume that we start the reduction to Min-Rep from the usual 3-SAT-5 instance I, $|I| = n$. Therefore, every boolean literal appears in exactly 5 clauses and every clause has exactly 3 literals. The following holds for the parameters of the construction (see more details in [13]).

It can be seen that the size of the graph is $n^{\mu \cdot \ell}$ for a universal constant μ, and that each A_i, B_j contains $O(\log n)$ sets (answer nodes). Therefore, $q = \Omega(n^{\mu \cdot \ell}/\log n)$ for some universal constant μ.

4.6 The reduction and degrees

Now focus on the degree of the nodes. Elements in the ground sets have degree at most $|A_i \cup B_j|$, which is $O(\log n)$.

As the set-cover for a yes instance has size $2q$, its size is at least $M/O(\log n)$. It can be seen by [28] that every a, b participates in at most $polylog(n)$ super-edges, and for every super-edge related to a or b, it is connected to $\sqrt{M}/2$ of the nodes of the ground set. Hence has maximum degree of $\bigcup_i A_i \cup \bigcup_j B_j$ in the set-cover graph is $O\left(\sqrt{M} \cdot poly(logn)\right)$.

The reduction to Degree and Diameter Bounded k-Tree is immediate. Add a node s connected to $\bigcup_i A_i \cup \bigcup_j B_j$. Make $\bigcup_{ij} M_{ij}$ the terminals and select the diameter bound to be 4. The node s is the root.

4.7 The gap

By the diameter bound of 4, the $\bigcup_i A_i \cup \bigcup_j B_j$ sets that are connected to s must set-cover the elements $\bigcup_{i,j} M_{i,j}$. The size of a set cover which is at least $2q = \Omega(M/O(\log n))$. This is much larger than the maximum degree of elements in ground sets. And it is much larger than the degree of any $A_i \cup B_j$. node that was established to be bounded by $O(\sqrt{M} \cdot \text{polylog}(n)) = o(q)$.

This implies that approximating the minimum degree is equivalent to approximation the size of the set-cover. The gap is $\alpha \log M' = \alpha \log M/2$, namely logarithmic in the size $|I|^\ell$ of the instance.

5 Prize-Collecting Degree-Constrained Survivable Network (Theorem 1.4)

Our LP-relaxation for Prize-Collecting Degree-Constrained Edge-Connectivity Survivable Network is:

$$
\begin{array}{llllll}
\text{Minimize} & \sum\limits_{e\in E} c_e x_e + \sum\limits_{I\subseteq K} \pi(I) z_I & & & & \\
\text{Subject to} & \sum\limits_{e\in\delta(T)} f_{i,e} & \geq & (1-\sum\limits_{I:i\in I} z_I) r_i(T) & \forall i \ \ \forall T \odot i & \\
& f_{i,e} & \leq & 1-\sum\limits_{I:i\in I} z_I & \forall i \ \ \forall e & \\
& x_e & \geq & f_{i,e} & \forall i \ \ \forall e & \\
& \sum\limits_{I\subseteq K} z_I & = & 1 & & \\
& \sum\limits_{e\in\delta(v)} x_e & \leq & b(v) & \forall v & \\
& x_e, f_{i,e}, z_I & \in & [0,1] & \forall i \ \ \forall e \ \ \forall I &
\end{array}
\qquad (6)
$$

Without the degree constraints, this LP-relaxation was used in [10] for Prize-Collecting Edge-connectivity Survivable Network. In the intended integral solution H, the variables are supposed to take the following values: $x_e = 1$ if $e \in H$, $f_{i,e} = 1$ if $i \notin \texttt{unsat}(H)$ and e appears on a chosen set of r_i edge-disjoint $\{u_i, v_i\}$-paths in H and $z_I = 1$ if $I = \texttt{unsat}(H)$. We prove the following refinement of Theorem 1.4.

Theorem 5.1 *Suppose that for an instance of* Prize-Collecting Degree-Constrained Edge-Connectivity Survivable Network *for any $\mathcal{P}' \subseteq \mathcal{P}$ the following holds. For an instance of* Degree-Constrained Edge-Connectivity Survivable Network *defined by $\mathcal{P}'$ there exists a polynomial-time algorithm that computes a subgraph H' of cost at most ρ times the optimal value of LP (2) with requirements restricted to $\mathcal{P}'$ such that $\deg_{H'}(v) \leq \alpha b(v) + \beta$ for all $v \in V$. Then for any $\mu \in (0,1)$,* Prize-Collecting Degree-Constrained Edge-Connectivity Survivable Network *admits a polynomial time algorithm that computes a subgraph H such that $c(H) \leq \frac{\rho}{1-\mu} \sum_{e\in E} c_e x^*_e$, $\pi(\texttt{unsat}(H)) \leq \frac{1}{\mu} \sum_{I\subseteq K} \pi(I) z(I)$ and $\deg_H(v) \leq \frac{\alpha}{1-\mu} b(v) + \beta$.*

We prove Theorem 5.1. Let $\{x^*_e, f^*_{i,e}, z^*_I\}$ be a feasible solution to LP (6) and let $\mu \in (0,1)$. We partition the requirements into two classes: we call a requirement r_i *good* if $\sum_{I:i\in I} z^*_I \leq \mu$ and *bad* otherwise. Let $\mathcal{R}_g$ denote the set of good requirements. The following statement shows how to satisfy the good requirements.

Lemma 5.2 *There exists a polynomial-time algorithm that computes a subgraph H of G of cost $c(H) \leq \frac{\rho}{1-\mu} \cdot \sum_e c_e x^*_e$ that satisfies all good requirements such that $\deg_H(v) \leq \frac{\alpha}{1-\mu} b(v) + \beta$ for all $v \in V$.*

Proof: Consider the LP-relaxation (2) of the Degree-Constrained Edge-Connectivity Survivable Network problem with good requirements only, with K replaced by K_g; namely, we seek a minimum cost subgraph H of G that satisfies the set K_g of good requirements and the degree constraints. We claim that $x^{**}_e = \min\{1, x^*_e/(1-\mu)\}$ for each $e \in E$ is a feasible solution to LP (2) with degree bounds $\frac{b(v)}{1-\mu}$. Thus the optimum value of LP (2) is at most $\sum_{e\in E} c_e x^{**}_e$. Consequently, using the algorithm that computes an integral solution to LP (2) of cost at most ρ times the optimal value of LP (2) and with degrees at most $\alpha b(v) + \beta$, we can construct a subgraph H that satisfies all good requirements and has cost at most $c(H) \leq \rho \sum_{e\in E} c_e x^{**}_e \leq \frac{\rho}{1-\mu}\sum_e c_e x^*_e$, and degrees at most $\deg_H(v) \leq \frac{\alpha}{1-\mu} b(v) + \beta$, as desired.

We now show that $\{x^{**}_e\}$ is a feasible solution to LP (2), namely, that $\sum_{e\in\delta(A)} x^{**}_e \geq r_i(A)$ for any $i \in K_g$ and any $A \odot i$. Let $i \in K_g$ and let $\zeta_i = 1 - \sum_{I:i\in I} z^*_I$. Note that $\zeta_i \geq 1-\mu$, by the definition of K_g. By the second and the third sets of constraints in LP (6), for every $e \in E$ we have $\min\{\zeta_i, x^*_e\} \geq f^*_{i,e}$. Thus we obtain: $x^{**}_e = \min\left\{1, \frac{x^*_e}{1-\mu}\right\} = \frac{1}{\zeta_i}\min\left\{\zeta_i, \frac{\zeta_i}{1-\mu}x^*_e\right\} \geq \frac{1}{\zeta_i}\min\{\zeta_i, x^*_e\} \geq \frac{f^*_{i,e}}{\zeta_i} = \frac{f^*_{i,e}}{1-\sum_{I:i\in I} z^*_I}$. Consequently, combining with the first set of constraints in LP (6), for any $A \odot i$ we obtain that $\sum_{e\in\delta(A)} x^{**}_e \geq \frac{\sum_{e\in\delta(A)} f^*_{i,e}}{1-\sum_{I:i\in I} z^*_I} \geq r_i(A)$. □

Let H be as in Lemma 5.2, and recall that $\mathtt{unsat}(H)$ denotes the set of requirements not satisfied by H. Clearly each requirement $i \in \mathtt{unsat}(H)$ is bad. The following lemma bounds the total penalty we pay for $\mathtt{unsat}(H)$.

Lemma 5.3 $\pi(\mathtt{unsat}(H)) \leq \frac{1}{\mu} \cdot \sum_I \pi(I) z^*_I$.

Proof: This lemma was proved in [10] for the case when there are no degree bounds, and the proof of the case with degree bounds is identical. □

The proof of Theorem 5.1 and thus also of Theorem 1.4 is now complete.

6 Leaf-Constrained Steiner Tree (Theorem 1.5)

We start by proving Part (i) of Theorem 1.5, namely, that if Group Steiner Tree admits approximation ratio $\rho(|V|, |\mathcal{S}|, \mathcal{S}_{\max})$ then Leaf-Constrained Steiner Tree admits ratio $\rho(|L|\cdot|V|, |S|, |V|)$. Given an instance $G = (V, E), c, S, L$ of Leaf-Constrained Steiner Tree construct an instance $G' = (V', E'), c', \mathcal{S}'$ of Group Steiner Tree as follows.

- The pair G', c' is obtained from G, c as follows. For every $v \in L$ do the following. For every $u \in \Gamma_G(v) \setminus L$ add a new node v_u, and replace the edge $e = uv$ by the new edge $e' = uv_u$, of the same cost as e. Then remove v and all the edges incident to it from the graph.
- The set of groups is as follows. Every $v \in L$ defines the group $S(v) = \{v_u : u \in \Gamma_G(v) \setminus L\}$. The collection of groups is $\mathcal{S}' = \{\{S(v)\} : v \in L\} \cup \{\{s\} : s \in S \setminus L\}$.

By the construction, $|V'| \leq |V|\cdot|L|$, $|S| = |\mathcal{S}|$, and $\mathcal{S}_{\max} \leq |V|$. Note that to every edge-set $F' \subseteq E'$ corresponds the edge-set $F \subseteq E$, where to every edge $e' = uv_u$ corresponds the edge uv, and the other edges appear in both F and F'. Note that if F corresponds to F', then F, F' have the same cost, namely, $c(F) = c'(F')$, and that if F' is a tree then so is F. Now we prove the following.

Lemma 6.1 *If T' is an inclusion-minimal solution to the obtained* Group Steiner Tree *instance then the edge set T that corresponds to T' is a feasible solution to the* Leaf-Constrained Steiner Tree *instance. Furthermore, to every inclusion-minimal solution T to the* Leaf-Constrained Steiner Tree *instance there exists a feasible solution T' to the* Group Steiner Tree *instance, such that T corresponds to T'.*

Proof: Let T' be an inclusion-minimal solution to the obtained Group Steiner Tree instance. Let $T \subseteq E$ the edge set that corresponds to T'. From the construction it is clear that T satisfies the connectivity requirements. We show that T satisfies the degree constraints. Since T' is an inclusion minimal solution, for every $v \in L$ there is a unique node $v_u \in S(v)$ included in the tree T'. This implies $\deg_T(v) = 1$.

Let T be an inclusion-minimal solution to the Leaf-Constrained Steiner Tree instance. If $|S| = 2$ then the statement is easily verified, so assume that $|S| \geq 3$. There is no edge in T between two nodes in L. Hence every $v \in L$ has its unique neighbor in $V \setminus L$. The tree T' is obtained from T by replacing for every $v \in L$ the unique edge uv incident to v in T by the edge uv_u. Clearly, T corresponds to T', and it is easy to see that T' is a feasible solution to the obtained Group Steiner Tree instance. $\square$

Now we prove Part (ii) of Theorem 1.5, namely that if Leaf-Constrained Steiner Tree admits ratio $\rho(|V|, |S|)$ then Group Steiner Tree admits ratio $\rho(|V| + |\mathcal{S}|, |\mathcal{S}|)$. Given an instance $G = (V, E), c, \mathcal{S}$ of Group Steiner Tree construct an instance $G' = (V', E'), c', S', L'$ of Leaf-Constrained Steiner Tree as follows.

- The pair G', c' is obtained from G, c by adding for every group S_i a new node v_i and connecting v_i to every node in S_i by an edge of cost zero.
- The node sets S', L' are defined by $S' = L' = \{v_1, \ldots, v_{|\mathcal{S}|}\}$.

By the construction, $|V'| = |V| + |\mathcal{S}|$ and $|S'| = |L'| = |\mathcal{S}|$. Now it is easy to see the following.

Lemma 6.2 *If T' is a feasible solution to the obtained* Leaf-Constrained Steiner Tree *instance then the tree $T = T' \setminus L$ is a feasible solution to the original* Group Steiner Tree *instance and $c(T) = c'(T')$. Furthermore, to every feasible solution T to the* Group Steiner Tree *instance there exists a feasible solution T' to the* Group Steiner Tree *instance, such that $T = T' \setminus L$.* $\square$

The proof of Theorem 1.5 is complete.

7 Discussion and open problems

In this paper we gave the first constant degree and cost approximation for the Degree-Constrained 2-Connected Subgraph problem. Recently, in [26], the method here was generalized to obtain constant ratios for several other node-connectivity degree-constrained problems. For the Degree-constrained k-Connected Subgraph problem the algorithm in [26] has ratio $O(2^k)$ for the degrees and $O(\log k)$ for the cost. The main open problem here is to improve the $O(2^k)$ ratio for the degrees.

Now let us focus on the *undirected* Minimum Degree k-MST problem. We do not know a lower bound for this problem beyond the standard APX-hardness result. However, it may be a bad sign that the natural LP already has a large integrality gap. Hence a polylogarithmic approximation for this problem may be hard or not possible to obtain. An easier task might be to get an n^ϵ-approximation scheme, similar to the one that exists for the Directed Steiner Tree problem (see [3]).

Finally, we note that for the Minimum Degree k-MST problem (the undirected variant) it is not hard to design an *"iterative merging"* algorithm with ratio $O(n/k)$ (c.f., [23]). Combined with our result in Theorem 1.2 this implies ratio $O(n^{1/3})$, which in terms of n might be better than the one in Theorem 1.2. We do not know if this holds also for directed graphs. The main open problem here is either to achieve a polylogarithmic ratio, or to give a strong evidence that a polylogarithmic ratio is unlikely.

References

[1] N. Bansal, R. Khandekar, and V. Nagarajan. Additive guarantees for degree bounded directed network design. *SIAM J. Computing*, 39(4):1413–1431, 2009.

[2] J. Bar-Ilan, G. Kortsarz, and D. Peleg. Generalized submodular cover problems and applications. *Theoretical Computer Science*, 250:179–200, 2001.

[3] M. Charikar, C. Chekuri, T. Cheung, Z. Dai, A. Goel, S. Guha, and M. Li. Approximation algorithms for directed Steiner problems. *Journal of Algorithms*, 33:73–91, 1999.

[4] Y. Dodis and S. Khanna. Design networks with bounded pairwise distance. In *STOC*, pages 750–759, 1999.

[5] M. Elkin and G. Kortsarz. An approximation algorithm for the directed telephone multicast problem. *Algorithmica*, 45(4):569–583, 2006.

[6] T. Feder, R. Motwani, and A. Zhu. k-connected spanning subgraphs of low degree. *Electronic Colloquium on Computational Complexity (ECCC)*, 041, 2006.

[7] M. Fürer and B. Raghavachari. Approximating the minimum-degree Steiner tree to within one of optimal. *J. Algorithms*, 17(3):409–423, 1994.

[8] N. Garg, G. Konjevod, and R. Ravi. A polylogarithmic approximation algorithm for the group Steiner tree problem. *J. Algorithms*, 37(1):66–84, 2000.

[9] M. Goemans. Bounded degree minimum spanning trees. In *FOCS*, pages 273–282, 2006.

[10] M. Hajiaghayi, R. Khandekar, G. Kortsarz, and Z. Nutov. Prize-Collecting Steiner Network Problems. In *IPCO*, pages 71–84, 2010.

[11] M. Hajiaghayi and A. Nasri. Prize-collecting steiner networks via iterative rounding. In *LATIN*, pages 515–526, 2010.

[12] E. Halperin and R. Krauthgamer. Polylogarithmic inapproximability. In *STOC*, pages 585–594, 2003.

[13] D. Hochbaum, editor. *Approximation algorithms for NP-hard problems.* PWS Publishing Co., Boston, MA, USA, 1997.

[14] K. Jain. A factor 2 approximation algorithm for the generalized Steiner network problem. *Combinatorica*, 21(1):39–60, 2001.

[15] R. Khandekar, G. Kortsarz, and Z. Nutov. Network-design with degree constraints. In *APPROX-RANDOM*, pages 289–301, 2011.

[16] P. Klein, R. Krishnan, B. Raghavachari, and R. Ravi. Approximation through local optimality: designing networks with small degree. *Networks*, 44:203–215, 2004.

[17] G. Kortsarz. On the hardness of approximating spanners. *Algorithmica*, 30(3):432–450, 2001.

[18] G. Kortsarz and Z. Nutov. *Approximating minimum cost connectivity problems,* in *Approximation Algorithms and Metahueristics,* T. F. Gonzalez ed. Chapman and Hall/CRC, 2007. Chapter 58.

[19] L. C. Lau, J. Naor, M. Salavatipour, and M. Singh. Survivable network design with degree or order constraints. *SIAM J. Computing*, 39(3):1062–1087, 2009.

[20] L. C. Lau, R. Ravi, and M. Singh. *Iterative Methods in Combinatorial Optimization.* Cambridge University Press, 2011.

[21] L. C. Lau and M. Singh. Additive approximation for bounded degree survivable network design. In *STOC*, pages 759–768, 2008.

[22] A. Louis and N. Vishnoi. Improved algorithm for degree bounded survivable network design problem. In *SWAT*, pages 408–419, 2010.

[23] M. V. Marathe, R. Ravi, R. Sundaram, S. S. Ravi, D. J. Rosenkrantz, and H. B. Hunt, III. Bicriteria network design problems. *J. Algorithms*, 28(1):142–171, 1998.

[24] M. Naor, L. Schulman, and A. Srinivasan. Splitters and near-optimal derandomization. In *FOCS*, pages 182–191, 1995.

[25] Z. Nutov. Approximating directed weighted-degree constrained networks. *Theoretical Computer Science*, 412(8-10):901–912, 2011.

[26] Z. Nutov. Degree-constrained node-connectivity, 2011. Manuscript.

[27] R. Ravi, B. Raghavachari, and P. Klein. Approximation through local optimality: Designing networks with small degree. In *FSTTCS*, pages 279–290, 1992.

[28] R. Raz. A parallel repetition theorem. *SIAM J. Comput.*, 27(3):763–803, 1998.

[29] M. Singh and L. C. Lau. Approximating minimum bounded degree spanning trees to within one of optimal. In *STOC*, pages 661–670, 2007.

[30] L. A. Wolsey. An analysis of the greedy algorithm for the submodular set covering problem. *Combinatorica*, 2:385–393, 1982.