Approximating k-node connected subgraphs via critical graphs

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

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

Abstract

We present two new approximation algorithms for the problem of finding a k-node connected spanning subgraph (directed or undirected) of minimum cost. The best known approximation guarantees for this problem were $O(\min\{k, \frac{n}{\sqrt{n-k}}\})$ for both directed and undirected graphs, and $O(\ln k)$ for undirected graphs with $n \geq 6k^2$, where n is the number of nodes in the input graph. Our first algorithm has approximation ratio $O(\frac{n}{n-k}\ln^2 k)$, which is $O(\ln^2 k)$ except for very large values of k, namely, $k = n - o(n)$. This algorithm is based on a new result on ℓ-connected p-critical graphs, which is of independent interest in the context of graph theory. Our second algorithm uses the primal-dual method and has approximation ratio $O(\sqrt{n}\ln k)$ for all values of n, k. Combining these two gives an algorithm with approximation ratio $O(\ln k \cdot \min\{\sqrt{k}, \frac{n}{n-k}\ln k\})$, which asymptotically improves the best known approximation guarantee for directed graphs for all values of n, k, and for undirected graphs for $k > \sqrt{n/6}$. Moreover, this is the first algorithm that has an approximation guarantee better than $\Theta(k)$ for all values of n, k. Our approximation ratio also provides an upper bound on the integrality gap of the standard LP-relaxation to the problem.

1 Introduction and preliminaries

A basic problem in network design is given a graph to find its minimum cost k-connected spanning subgraph; a graph is *k(-node) connected* if it is simple and there are at least k internally disjoint paths from every node to the other. This problem is NP-hard for undirected graphs with $k = 2$, and for directed graphs with $k = 1$. The best known approximation guarantees for this problem were $O(\min\{k, \frac{n}{\sqrt{n-k}}\})$ for both directed and undirected graphs [16, 4], and $O(\ln k)$ for undirected graphs with $n \geq 6k^2$ [4], where n is the number of nodes in the input graph. [1] Better approximation guarantees are known for restricted edge costs, as follows. For metric costs: $2 + 2(k-1)/n$ for undirected graphs [17] (for a slight improvement to $2 + (k-1)/n$ see [16]) and $2 + k/n$ for directed graphs [16]. For uniform costs there is a $(1 + 1/k)$-approximation algorithm for both directed and undirected graphs [3]. The case when the input graph is complete and the cost of every edge is in $\{0, 1\}$ (so called "vertex-connectivity augmentation problem") is polynomially solvable for directed graphs [7]; polynomial algorithms that compute a near optimal solution for undirected graphs are given in [11, 13]. But in this paper we consider the case of general costs only.

The main result of this paper is the following theorem:

Theorem 1.1 *There exists an algorithm for the minimum-cost k-connected subgraph problem with approximation ratio $O(\ln k \cdot \min\{\sqrt{k}, \frac{n}{n-k} \ln k\})$ and running time $O(k^2 nm^2)$.*

This gives the first algorithm that has an approximation guarantee better than $\Theta(k)$ for all values of n, k, and improves the previously best known approximation guarantees for directed graphs for all values of n, k, and for undirected graphs for $k > \sqrt{n/6}$. Note that our approximation ratio is $O(\ln^2 k)$ except for very large values of k (namely, $k = n - o(n)$). In particular, for instances with $n > kc$ where $c > 1$ is a fixed constant, the approximation ratio is $O(\ln^2 k)$. For example, the previously best approximation ratios for $k = \sqrt{n}$ and $k = n/2$ were $O(k)$ and $O(\sqrt{k})$, respectively; our approximation ratio for both these cases is $O(\ln^2 k)$. For $k = n - o(n)$ the improvement is from $O(k)$ to $O(\sqrt{k} \ln k)$. Our algorithm is combinatorial, and runs significantly faster than the $O(\frac{n}{\sqrt{n-k}})$-approximation algorithm of [4] which solves linear programs.

Remark: A generalization of the min-cost k-connected subgraph problem is the *Survivable Network Design Problem (SNDP)*: find a cheapest spanning subgraph such that for every pair (u, v) of nodes there are at least k_{uv} pairwise internally disjoint paths from u to v. It is interesting to compare Theorem 1.1 with results in [15, 27] which show that SNDP with $k_{uv} \in \{0, k\}$, $k = \Theta(n)$, and costs in $0, 1$, is unlikely to have a polylogarithmic approximation guarantee. On the other hand, Jain [14] showed that the version of SNDP where the paths are only required to be pairwise *edge* disjoint admits a 2-approximation algorithm.

Our $O(\frac{n}{n-k} \ln^2 k)$-approximation algorithm is based on a new result on ℓ-connected p-critical graphs which is of independent interest in graph theory. Namely, we will prove

[1] For undirected graphs, Ravi and Williamson [28] claimed an $O(\ln k)$-approximation algorithm, but the proof was found to contain an error, see [29].

that any ℓ-connected graph (directed or not) on n nodes has a subset U of nodes with $|U| = O(\frac{n}{n-\ell}\ln \ell)$ such that no node-cut of cardinality ℓ contains U; we call such U an ℓ-cover (since U covers the complements of node-cuts of cardinality ℓ), and denote by $\tau_\ell(G)$ the minimum cardinality of an ℓ-cover in G. An ℓ-connected graph G is p-critical if $p < \tau_\ell(G)$ (this definition is shown to be equivalent to the one used in the papers on the topic, see Section 1.2). For undirected graphs, our result partly bridges the gap between two main bounds: the obvious fact that $\tau_\ell(G) \leq \ell + 1$ and a result of Mader [21] which states that $\tau_\ell(G) \leq 3$ for $n \geq 6\ell^2$. Other previous bounds were for undirected graphs only, and of the type $\tau_\ell(G) = \Theta(\ell)$ (e.g., [18]), or of the type $\tau_\ell(G) = \Theta(1)$ for $n = \Omega(\ell^2)$ (e.g., [22]). Our result gives the first nontrivial bound in the intermediate range for undirected graphs, and overall the first nontrivial bound for directed graphs. Moreover, our proof provides a polynomial algorithm that computes an ℓ-cover within the stated bound.

Throughout the paper, let $\mathcal{G} = (V, \mathcal{E})$ denote the input graph with nonnegative costs on the edges; n denotes the number of nodes in $\mathcal{G}$, and m the number of edges in $\mathcal{G}$. Unless stated otherwise, "graph" stands for both directed and undirected graph.

This paper is organized as follows. In the rest of this section we introduce a standard LP-relaxation to the min-cost k-connected subgraph problem, and state some simple facts about ℓ-outconnected graphs and p-critical graphs. In Sections 2 and 3 we give our algorithm for the min-cost k-connected subgraph problem: in Section 2 we establish existence of an ℓ cover of size $O(\frac{n}{n-\ell}\ln \ell)$ and show how to compute it, which implies an $O(\frac{n}{n-k}\ln^2 k)$-approximation algorithm, while Section 3 present our primal-dual $O(\sqrt{n}\ln k)$-approximation algorithm.

1.1 LP-relaxation and ℓ-outconnected graphs

For an edge set or a graph J on a node set V and $S, T \subseteq V$ let $\delta_J(S,T)$ denote the set of edges in J going from S to T. By Menger's Theorem, there are k internally disjoint paths from a node s to a node t in a graph $G = (V, E)$ if, and only if, $|\delta_E(S,T)| \geq k - (n - |S \cup T|)$ for all disjoint $S, T \subset V$ with $s \in S$ and $t \in T$. We will compare the cost of our solutions to the optima opt_k of the following LP-relaxation for the minimum cost k-node connected spanning subgraph that has been introduced in [7] and used in [4]:

$$\begin{array}{llll} opt_k = & \min & \sum_{e \in \mathcal{E}} c_e x_e & \\ & \text{s.t.} & \sum_{e \in \delta_{\mathcal{E}}(S,T)} x_e \geq k - (n - |S \cup T|) & \forall\, \emptyset \neq S, T \subset V, S \cap T = \emptyset \\ & & 0 \leq x_e \leq 1 & \forall e \in \mathcal{E}. \end{array}$$

A graph is *ℓ-outconnected from a node r* if it contains ℓ internally disjoint paths from r to any other node; a graph is *ℓ-inconnected to r* if its reverse graph is ℓ-outconnected from r (for undirected graphs these two concepts mean the same). Frank and Tardos [8] showed that for directed graphs, the problem of finding an ℓ-outconnected spanning subgraph of minimum cost is solvable in polynomial time; a faster algorithm with time complexity $O(\ell^2 n^2 m) = O(m^3)$ is due to Gabow [9] (observe that $n\ell = O(m)$ in an ℓ-outconnected graph).

Let $G = (V, E)$ be an ℓ-connected spanning subgraph of cost zero of a directed graph $\mathcal{G}$, and suppose that G has a subset U of nodes such that no node-cut of cardinality ℓ contains all of them. Then using the algorithm of [8] it is easy to get a $2|U|$-approximation algorithm for the problem of augmenting G to be $(\ell+1)$-connected by adding an edge set of minimum cost: for each node $r \in U$ we compute an $(\ell+1)$-outconnected spanning subgraph from r and an $(\ell+1)$-inconnected spanning subgraph to r and take the union of these $2|U|$ subgraphs. In fact, the following lemma, which can be easily deduced from [5, Theorem 7] (e.g, see [4, Lemma 3.4]) implies that the augmenting edge set produced has cost at most $\frac{2|U|}{k-\ell} opt_k$.

Lemma 1.2 *Let G be an ℓ-outconnected from r subgraph of cost zero of a directed graph $\mathcal{G}$, and for an integer p let α^p be the minimum cost of an $(\ell+p)$-outconnected spanning subgraph of $\mathcal{G}$. Then $\alpha^1 \leq \alpha^p/p$. In particular, for $\ell < k$ the minimum cost of an $(\ell+1)$-outconnected spanning subgraph of $\mathcal{G}$ is at most $\frac{1}{k-\ell} opt_k$.*

Khuller and Raghavachari [17] observed that the algorithm of [8] implies a 2-approximation algorithm for the problem of finding an optimal ℓ-outconnected subgraph of an undirected graph, as follows. First, replace every undirected edge e of $\mathcal{G}$ by the two antiparallel directed edges with the same ends and of the same cost as e. Then compute an optimal ℓ-outconnected subdigraph from r and output its underlying (undirected) simple graph. Several papers used this observation for designing approximation algorithms for node connectivity problems, e.g., see [1, 2, 16, 4].

1.2 p-critical graphs and k-connected subgraphs

Let $\kappa(G)$ denote the *connectivity of G*, that is the maximum integer ℓ for which G is ℓ-connected. A (directed or undirected) graph $G = (V, E)$ is *p-critical* if $\kappa(G-U) = \kappa(G) - |U|$ for any $U \subset V$ with $|U| \leq p$. One can characterize p-critical graphs in terms of covers of set families, as follows. Let $G = (V, E)$ be an ℓ-connected graph. Let $X^* = X^*_G = \{v \in V - X : \delta_G(X, v) = \emptyset\}$ denote the "node complement" of X in G. We say that $X \subset V$ is an *ℓ-fragment* if $X^* \neq \emptyset$ and $|V - (X \cup X^*)| = \ell$. It is well known that if G is ℓ-connected, then $|V| \geq \ell + 1$, and if $|V| = \ell + 1$ then G must be a complete graph. Note that Menger's Theorem implies the following well known statement:

Proposition 1.3 *An ℓ-connected graph G (on at least $\ell + 2$ nodes) is $(\ell + 1)$-connected if, and only if, G has no ℓ-fragments.*

Given a family $\mathcal{F}$ of subsets of a groundset V we say that $U \subseteq V$ *covers* $\mathcal{F}$ if U intersects every set in $\mathcal{F}$. Let $\mathcal{F}_\ell(G)$ be the family of all ℓ-fragments of G. We say that $U \subseteq V$ is an *ℓ-cover of G* if U covers $\{X \cup X^* : X \in \mathcal{F}_\ell(G)\}$; let $\tau_\ell(G)$ denote the minimum cardinality of an ℓ-cover of G. From Proposition 1.3 we have:

Proposition 1.4 *Let $G = (V, E)$ be a graph with $\kappa(G) = \ell$ and $|V| \geq \ell + 2$. Then:*

(i) G is p-critical if, and only if, for any $U \subseteq V$ with $|U| \leq p$ there exists an ℓ-fragment X with $U \cap (X \cup X^*) = \emptyset$. Thus if G is p-critical then G is p'-critical for any $p' \leq p$, and $\tau_\ell(G) - 1$ is the maximum p for which G is p-critical.

(ii) U is an ℓ-cover of G if, and only if, there exists an edge set F incident to U (that is, every edge in F has at least one endpoint in U) such that $G + F$ is $(\ell + 1)$-connected.

Combining Proposition 1.4(ii) with Lemma 1.2 and the discussion before it, we get the following statement, which was implicitly proved in [4] for undirected graphs.

Proposition 1.5 ([4]) *Suppose that there is a polynomial algorithm that finds in any ℓ-connected graph G on n nodes an ℓ-cover of G of size at most $t(\ell, n)$. Then there exists a polynomial algorithm that for instances of the minimum k-connected subgraph problem on n nodes finds a feasible solution of cost at most $opt_k \cdot 2\sum_{\ell=0}^{k-1} \frac{t(\ell,n)}{k-\ell} = opt_k \cdot O(\ln k \cdot \max_{0 \leq \ell \leq k-1} t(\ell, n))$.*

For undirected graphs with $n \geq 6k^2$ Cheriyan et al. [4] gave a $6H(k)$-approximation algorithm for the undirected min-cost k-connected subgraph problem combining Proposition 1.5 with the following theorem due to Mader:

Theorem 1.6 ([21]) *Any undirected 3-critical graph G has less than $6\kappa(G)^2$ nodes.*

In a recent paper [22] Mader improved his bound for 3-critical graphs to $n \leq \kappa(G)(2\kappa(G) - 1)$; hence via Proposition 1.5 the $6H(k)$-approximation algorithm of [4] is valid for $n \geq k(2k - 1)$ as well.

On the other hand, it is easy to see that there are no $\kappa(G)$-critical non-complete graphs. But for undirected graphs, a stronger result was conjectured in [26], and answered by Su:

Theorem 1.7 ([30]) *If a noncomplete graph G is p-critical, then $p \leq \lfloor \kappa(G)/2 \rfloor$.*

For a survey on p-critical graphs see [24, 25]; for some recent results see [22, 23] and [18].

2 Computing logarithmic covers

Note that in terms of covers of set families Theorem 1.6 states that $\tau_\ell(G) \leq 3$ for any undirected graph G with $\kappa(G) = \ell$ and $n \geq 6\ell^2$, and Theorem 1.7 states that $\tau_\ell(G) \leq \lfloor \ell/2 \rfloor + 1$ (if $n \geq \ell + 2$). Our result on p-critical graphs partly bridges the gap between these two bounds, and also gives the first nontrivial bound on $\tau_\ell(G)$ for directed graphs. Let $\theta = \theta(\ell, n) = \frac{2n}{n+\ell}$.

Theorem 2.1 *There exists a polynomial algorithm that given an ℓ-connected graph G on $n \geq \ell + 2$ nodes finds an ℓ-cover of G of size at most*

$$t(\ell, n) = 2 + \frac{3n}{n-\ell} + \frac{1}{\ln \theta} \ln \frac{1}{2}(\ell - 1 - \ell^2/n) = O\left(\frac{n}{n-\ell} \ln \ell\right)$$

if G is undirected, and of size at most $2t(\ell, n)$ if G is directed.

Combining with Proposition 1.5 we get:

Theorem 2.2 *For the minimum cost k-connected subgraph problem there exists a polynomial algorithm that finds a feasible solution of cost at most $opt_k \cdot O(\frac{n}{n-k} \ln^2 k)$.*

Remark: Note that $\tau_\ell(G)$ is the minimum cardinality of a cover (transversal) of the $(n-\ell)$-uniform hypergraph $\{X \cup X^* : X \in \mathcal{F}_\ell(G)\}$. Several general bounds on covers of uniform hypergraphs are known, e.g., see [10]. But, as far as we can see, none of them implies the bound given in Theorem 2.1.

We need several definitions and simple facts to prove Theorem 2.1. In the rest of this section, let ℓ be a fixed integer, and let G be a graph with $\kappa(G) \geq \ell$. An ℓ-fragment X of G is *small* if $|X| \leq |X^*|$, that is if $|X| \leq \lfloor (n-\ell)/2 \rfloor$. Note that by Proposition 1.3, G is $(\ell+1)$-connected if, and only if, G (and the reverse graph of G, if G is directed) has no small ℓ-fragments. Let $\mathcal{S}_\ell(G)$ denote the family of all small ℓ-fragments of G. The following Lemma is well known, e.g., see [11, Lemma 1.2], where it was stated for undirected graphs.

Lemma 2.3 *Let X, Y be two intersecting ℓ-fragments in an ℓ-connected (directed or undirected) graph G on n nodes. If $n - |X \cup Y| \geq \ell$ then $X \cap Y$ is an ℓ-fragment, and if a strict inequality holds, then also $X \cup Y$ is an ℓ-fragment. In particular, the intersection of two intersecting small ℓ-fragments is also a small ℓ-fragment.*

A *core* of G is an inclusion minimal small ℓ-fragment. By Lemma 2.3 the cores of G are pairwise disjoint, and let $\nu(G) = \nu_\ell(G)$ denote their number; note that if $\kappa(G) > \ell$ then $\nu_\ell(G) = 0$. For a core C_i of G, let A_i be the union of all small ℓ-fragments that contain a unique core which is C_i. Let $\mathcal{A}_\ell(G) = \{A_1, \ldots, A_{\nu(G)}\}$. The properties of the sets in $\mathcal{A}_\ell(G)$ that we use are summarized in the following statement:

Corollary 2.4 *The sets in $\mathcal{A}_\ell(G)$ are pairwise disjoint. Moreover, for every $A \in \mathcal{A}_\ell(G)$ holds: either A is an ℓ-fragment, or $|A| \geq n - \ell$ (and $A^* = \emptyset$); thus $|A \cup A^*| \geq n - \ell$.*

Proof: Suppose to the contrary that A_i and A_j intersect for some $1 \leq i \neq j \leq \nu(G)$. Then, by the definition of A_i, A_j, there are two small ℓ-fragments D_i, D_j such that: D_i contains a unique core which is C_i, D_j contains a unique core which is C_j, and such that D_i, D_j intersect. By Lemma 2.3 $D_i \cap D_j$ is a small ℓ-fragment, and thus contains a core C. This implies that D_i contains the two cores C_i and C, which gives a contradiction.

To prove the second statement, let us fix some set $A \in \mathcal{A}_\ell(G)$. Since the sets in $\mathcal{A}_\ell(G)$ are disjoint, A contains a unique core, say C. Consider the family $\mathcal{D}$ of all small ℓ-fragments that contain a unique core which is C, so A is the union of the sets in $\mathcal{D}$. If $n - |A| \leq \ell$, then clearly $|A \cup A^*| \geq |A| \geq n - \ell$ (in fact, in this case $A^* = \emptyset$, and thus $|A \cup A^*| = |A|$). Otherwise, $n - |A| \geq \ell + 1$; then by Lemma 2.3, the union of the sets in $\mathcal{D}$ is an ℓ-fragment, and thus $|A \cup A^*| = n - \ell$. In both cases, the statement is valid. $\square$

Note that Corollary 2.4 does *not* imply that the sets in $\mathcal{A}_\ell(G)$ are small, or that they are ℓ-fragments; it might be that A is large and that $A^* = \emptyset$ for some $A \in \mathcal{A}_\ell(G)$, but in any case, $|A \cup A^*| \geq n - \ell$ holds.

Lemma 2.5 *Let $\mathcal{A}$ be a family of sets on a groundset V such that $|A| \geq n - \ell$ holds for every $A \in \mathcal{A}$, where $n = |V|$ and ℓ is an integer. Then there exists an element $r \in V$ that covers (that is, intersects) at least $\left(1 - \frac{\ell}{n}\right) |\mathcal{A}|$ sets in $\mathcal{A}$.*

Proof: For $r \in V$, let $\mathcal{A}_r = \{A \in \mathcal{A} : r \in A\}$ be the sets in $\mathcal{A}$ covered by r. The claim follows since we have

$$\sum_{r \in V} |\mathcal{A}_r| = \sum_{A \in \mathcal{A}} |A| \geq |\mathcal{A}|(n - \ell).$$

$\square$

For $r \in V$ let $F_r = \{vr : v \in V - r\}$, and let $G + F_r$ be the graph obtained by adding an edge from every node $v \in V$ to r, if such does not exist in G. We say that $r \in V$ *outercovers* $A \in \mathcal{A}_\ell(G)$ if $r \in A^*$.

Lemma 2.6 *Let C_i be a core of an ℓ-connected graph G. If r outercovers A_i then any small ℓ-fragment X of $G + F_r$ that contains C_i must contain a core of G distinct from C_i.*

Proof: Let X be a small ℓ-fragment of $G + F_r$ that contains C_i. Assume by contradiction that this is the unique core of G that X contains. Note that X is a small ℓ-fragment of G. Since A_i is defined as the union of all small ℓ-fragments of G containing C_i as their unique core, we obtain that $X \subseteq A_i$. This gives a contradiction, since then $r \in A_i^* \subseteq X^*$, which implies that X cannot be an ℓ-fragment of $G + F_r$. □

Lemma 2.7 *Let r be a node that outercovers q sets in $\mathcal{A}_\ell(G)$. Then $\nu_\ell(G+F_r) \leq \nu_\ell(G) - q/2$.*

Proof: If $\kappa(G + F_r) > \ell$ then $\nu_\ell(G + F_r) = 0$ and the statement is obvious, so assume that $\kappa(G + F_r) = \ell$. By Lemma 2.3, the cores of $G + F_r$ are pairwise disjoint. Clearly, every core of $G + F_r$ is a small ℓ-fragment of G, and thus contains at least one core of G. Let t be the number of cores of $G + F_r$ containing exactly one core of G. By Lemma 2.6, any core C of $G + F_r$ that contains some core C_i of G with $r \in A_i^*$ must contain another core of G distinct from C_i, so such C contains at least two cores of G. Thus $t \leq \nu_\ell(G) - q$. From this we get that $\nu_\ell(G + F_r) \leq t + (\nu_\ell(G) - t)/2 \leq \nu_\ell(G) - q/2$, as required. □

Since the sets in $\mathcal{A}_\ell(G)$ are pairwise disjoint, a node can belong to at most one of them. Thus, if $\mathcal{A}' \subseteq \mathcal{A}_\ell(G)$ and r covers $\{A \cup A^* : A \in \mathcal{A}'\}$, then there is most one set $A' \in \mathcal{A}'$ such that $r \in A'$; for any other $A \in \mathcal{A} - A'$ we must have $r \in A^*$, hence r outercovers at least $|\mathcal{A}'| - 1$ sets in $\mathcal{A}'$. Combining this with Corollary 2.4 and Lemmas 2.5 and 2.7 we get:

Corollary 2.8 *Any ℓ-connected graph G has a node r that outercovers at least $\nu_\ell(G)(1 - \ell/n) - 1$ sets in $\mathcal{A}_\ell(G)$, and $\nu_\ell(G + F_r) \leq \frac{n+\ell}{2n}\nu_\ell(G) + 1/2$.*

Let us apply the following algorithm on an ℓ-connected graph G starting with $U = \emptyset$.

While $\nu_\ell(G) > 0$ do:
1. Find a node r for which $\nu_\ell(G + F_r)$ is minimal;
2. $U \leftarrow U + r$, $G \leftarrow G + F_r$;

End While
Output U.

By Proposition 1.4(ii), at the end of the algorithm U is an ℓ-cover, and let us estimate its size. Let t_j be the number of cores in G after j iterations of the main loop, and set $\theta = \frac{2n}{n+\ell}$ and $\alpha = 1/\theta$. Corollary 2.8 gives the recursive bound

$$t_{j+1} \leq \alpha t_j + 1/2.$$

We will prove later that $t_1 \le \ell$ (see Corollary 2.11 below) which implies:

$$\begin{aligned} t_j &\le \alpha^{j-1}\ell + \frac{1}{2}(1+\alpha+\cdots+\alpha^{j-2}) = \alpha^{j-1}\ell + \frac{1-\alpha^{j-1}}{2(1-\alpha)} = \\ &= \alpha^{j-1}\left(\ell - \frac{1}{2(1-\alpha)}\right) + \frac{1}{2(1-\alpha)} = \frac{1}{\theta^{j-1}}\left(\ell - \frac{n}{n-\ell}\right) + \frac{n}{n-\ell}. \end{aligned}$$

The inequality can be easily proved by induction on j. Let $\beta = 3n/(n-\ell)$.

Claim 2.9

$$t_j \le \beta \quad \textit{for} \quad j \ge j(\beta) \equiv \frac{1}{\ln\theta}\ln\frac{1}{2}(\ell - 1 - \ell^2/n) + 1.$$

Proof: We solve for j the inequality

$$\frac{1}{\theta^{j-1}}\left(\ell - \frac{n}{n-\ell}\right) + \frac{n}{n-\ell} \le \frac{3n}{n-\ell} = \beta.$$

That is

$$\theta^{j-1} \ge \frac{\ell(n-\ell)-n}{2n} = \frac{1}{2}(\ell - 1 - \ell^2/n).$$

The claim follows by taking logarithm base θ from both sides, and then changing the logarithm base:

$$j - 1 \ge \log_\theta \frac{1}{2}(\ell - 1 - \ell^2/n) = \frac{1}{\ln\theta}\ln\frac{1}{2}(\ell - 1 - \ell^2/n).$$

$\square$

On the other hand, if $t_j > 0$ then $t_{j+1} \le t_j - 1$ since $\nu_\ell(G + F_r) \le \nu_\ell(G) - 1$ for any node r belonging to a core C_i of G (indeed, every core of $G + F_r$ must contain some core of G, but cannot contain C_i). Thus the number of iterations in the algorithm (which equals to the size of the cover found) is bounded by $\lceil j(\beta)\rceil + \lfloor\beta\rfloor \le t(\ell, n)$. This proves Theorem 2.1 for undirected graphs. In the case of a directed graph G, at the end of the algorithm G has no small ℓ-fragments, but G may not be $(\ell+1)$-connected since the reverse graph of G might have small ℓ-fragments. Thus we apply the above procedure twice: on G and on the reverse graph of G, and take the union of the resulting two node sets.

Let us now show that $t_1 \le \ell$, and discuss some consequences from our approach. The following statement is obvious.

Lemma 2.10 *Let r be a node of a (directed or undirected) noncomplete graph $G = (V, E)$ with $\kappa(G) = \ell$, and let $N_r = \{v \in V : vr \in E\}$ be the nodes in G having r as their neighbor. Then N_r covers all ℓ-fragments of $G + F_r$.*

An ℓ-connected graph J is *minimally ℓ-connected* if $J - e$ is not ℓ-connected for every edge e of J. Mader [19, 20] showed that any minimally ℓ-connected graph J on n nodes has at least $\frac{(\ell-1)n+2}{2\ell-1}$ nodes of degree (indegree, if J is directed) ℓ each. Since $\mathcal{F}_\ell(G) \subseteq \mathcal{F}_\ell(J)$ for any ℓ-connected spanning subgraph J of an ℓ-connected graph G, Lemma 2.10 implies the following corollary, which also proves that $t_1 \leq \ell$.

Corollary 2.11 *Let $G = (V, E)$ be an ℓ-connected graph. Then there is $R \subseteq V$ with $|R| \geq \frac{(\ell-1)n+2}{2\ell-1}$ such that for any $r \in R$ the following holds: r and at most ℓ nodes having r as their neighbor cover all ℓ-fragments of G, and in particular $\nu_\ell(G + F_r) \leq \ell$.*

We note that for a directed graph G, Corollary 2.11 implies only the trivial bound $\tau_\ell(G) \leq \ell + 1$; however for undirected G, the following theorem provides an easy proof of Theorem 1.7, which is similar to the one given by Jordán in [12]; recall that in terms of covers, Theorem 1.7 states that $\tau_\ell(G) \leq \lfloor \ell/2 \rfloor + 1$.

Theorem 2.12 *Let G be an undirected ℓ-connected graph and let W be a cover of $\mathcal{F}_\ell(G)$. Then there exists an ℓ-cover $U \subseteq W$ of size at most $\lfloor |W|/2 \rfloor$.*

Proof: In [19], Mader implicitly proved (e.g., see [11] and [16, Corollary 2.2]) that if W covers all the ℓ-fragments of an undirected ℓ-connected graph G then there exists a forest F on W such that $G+F$ is $(\ell+1)$-connected. Since F is a forest, there exists $U \subseteq W$ such that $|U| \leq \lfloor |W|/2 \rfloor$ and every edge in F is incident to a node in U. Thus, by Proposition 1.4 (ii), U is a cover of $\mathcal{F}_\ell(G)$ as required. $\square$

Let us now analyze the time complexity of our approximation algorithm for k-connected spanning subgraphs. Using max-flow techniques an ℓ-cover as in Theorem 2.1 can be found in $O(\ell m^2)$ time, as follows. For the first iteration, we find a minimally ℓ-connected spanning subgraph J of G, and choose a node s of degree (indegree, if G is directed) ℓ in J; such J can be found in $O(\ell m^2)$ time by repeatedly checking every edge for deletion. By Lemma 2.10 the set $N = N_s$ of nodes having s as their neighbor in J covers all ℓ-fragments of $J + F_s$, and thus also of $G + F_s$. Now we set $G \leftarrow G + F_s$. We compute for every $u \in N_s$ and $v \in V$ a set of ℓ internally disjoint paths. This can be done in $O(\ell m)$ time per pair, thus in $O(\ell^2 nm)$ total time, using the Ford-Fulkerson algorithm (the node-capacitated version) and flow decomposition. For each pair uv we check whether v is reachable from u in the corresponding residual network. If so, then the pair uv is discarded; otherwise, a minimal ℓ-fragment containing v is found, and if its size is $\leq (n-l)/2$, it is the minimal core containing v. At each iteration, for every $r \in V$, we can recompute the cores of $G + F_r$ in $O(lm)$ time. Thus each iteration can be implemented in $O(\ell nm)$ time, and since the number of iteration is at most ℓ, an ℓ-cover as in Theorem 2.1 can be found in $O(\ell^2 nm) = O(\ell m^2)$ time, as claimed.

We also need to find a minimum-cost edge set to increase the outconnectivity from ℓ to $\ell+1$ from each node in the cover found. Frank [6] showed that a generalization of this problem can be solved in $O(n^2m)$ time, but with some care Frank's algorithm can be implemented in $O(m^2)$ time. As the size of the cover found is $O(\ell)$, we get that the overall time complexity for increasing connectivity from ℓ to $\ell+1$ is $O(\ell m^2)$, where $m \geq n\ell$ is the number of edges in $\mathcal{G}$. Consequently, the overall running time of the algorithm is $O(k^2m^2)$.

3 A primal-dual algorithm

In this section we prove the following theorem:

Theorem 3.1 *For the problem of making a k_0-connected graph (directed or undirected) k-connected by adding a minimum-cost edge set there exists an approximation algorithm with approximation ratio $O(\sqrt{n}H(k-k_0)) = O(\sqrt{n}\ln k)$ and time complexity $O(km(k^2n^2 + \sqrt{n}m)) = O(n^5m)$, where $H(j)$ denotes the jth Harmonic number.*

We start by giving an algorithm for increasing the connectivity of a *directed graph* by one. We use as a subroutine the primal-dual algorithm of Ravi and Williamson [28], which we adapt to directed graphs. Given an ℓ-connected graph G the algorithm of [28] uses the primal-dual method to find an edge set F so that $G+F$ is $(\ell+1)$-connected. We use the same approach, but unlike the algorithm in [28], the primal-dual procedure terminates when we find an edge set F^+ so that $\nu_\ell(G+F^+) = \sqrt{2n}$; we will show that $c(F^+) = O(\frac{\sqrt{n}}{k-\ell})opt_k$. We then find in $G+F^+$ a node set U of size $O(\sqrt{n})$ by picking one node from every core; for every $r \in U$ we find an $(\ell+1)$-inconnected subgraph to r subgraph. The cost of each subgraph found is $O(\frac{1}{k-\ell})opt_k$ and their number is $|U| = O(\sqrt{n})$. Thus the cost of the edge set found during this step is also $O(\frac{\sqrt{n}}{k-\ell})opt_k$. We apply this procedure twice: on G and on the reversed graph of G. Consequently, the total cost of the edge set found is $O(\frac{\sqrt{n}}{k-\ell})opt_k$.

Let $G = (V, E)$ be an ℓ-connected spanning subgraph of a directed graph $\mathcal{G} = (V, \mathcal{E})$, such that all the edges in E have cost zero, and let $I = \mathcal{E} - E$. Let $\mathcal{S} = \mathcal{S}_\ell(G)$ denote the set of small ℓ-fragments of G. For an edge set F and $S \in \mathcal{S}$, let $d_F(S) = |\delta_F(S, S^*)|$ be the number of edges in F going from S to S^*. Recall that $G+F$ is $(\ell+1)$-connected if, and only if, the graph $G+F$ and its reverse graph have no small ℓ-fragments. Note that for $F \subseteq I = \mathcal{E} - E$, $G+F$ has no small ℓ-fragments if, and only if, $d_F(S, S^*) \geq 1$ for any $S \in \mathcal{S}$. Consider the following linear program (P) and its dual program (D), where (P) is a linear relaxation for the problem of finding a minimum cost augmenting edge set F such that $G+F$ contains no small ℓ-fragments:

$$
\begin{array}{llll}
 & \min & \sum_{e\in I} c_e x_e & \\
\text{(P)} & \text{s.t.} & \sum_{e\in\delta_I(S,S^*)} x_e \geq 1 & \forall S\in\mathcal{S} \\
 & & x_e \geq 0 & \forall e\in I
\end{array}
\qquad
\begin{array}{llll}
 & \max & \sum_{S\in\mathcal{S}} y_S & \\
\text{(D)} & \text{s.t.} & \sum_{S\in\mathcal{S}:e\in\delta_I(S,S^*)} y_S \leq c_e & \forall e\in I \\
 & & y_S \geq 0 & \forall S\in\mathcal{S}.
\end{array}
$$

Lemma 3.2 *Let x be an optimal solution to (P). Then $\sum_{e\in I} c_e x_e \leq \frac{1}{k-\ell} opt_k$.*

Proof: Let x be an optimal solution to the LP-relaxation for the min-cost k-connected spanning subgraph problem (given in Section 1.1). Define $x'_e = 1$ if $e \in E$ and $x'_e = \frac{1}{k-l} x_e$ otherwise. Then x' is a feasible solution to (P). Since all the edges in E have cost zero, the claim follows. □

Given a feasible solution y to (D), an edge $e \in I$ is *tight* if the corresponding inequality in the dual program (D) holds with equality. If $F^+ \subseteq I$ is a set of tight edges, then

$$
c(F^+) = \sum_{e\in F^+} c_e = \sum_{e\in F} \sum_{S\in\mathcal{S}:e\in\delta(S,S^*)} y_S = \sum_{S\in\mathcal{S}} d_F(S) y_S. \tag{1}
$$

Recall that by Lemma 2.3 the cores of G are disjoint. Let us fix the threshold $\beta = \sqrt{2n}$ and apply the following procedure:

Procedure 1:
While $\nu_\ell(G) \geq \beta$, raise dual variables corresponding to cores of G uniformly until some edge $e \in I$ becomes tight, and add this edge to G.

Let $\tilde{F}^+$ be the set of edges added to the input graph G by Procedure 1.

Lemma 3.3 *Let $F^+ \subseteq \tilde{F}^+$. Then $c(F^+) \leq \dfrac{|F^+|}{\beta}\dfrac{opt_k}{k-\ell}$.*

Proof: Let y be the dual solution produced by Procedure 1. Since the edges in F^+ are tight, we have $c(F^+) = \sum_{S\in\mathcal{S}} d_{F^+}(S) y_S$, by (1). Let $\mathcal{C}_i$ be the family of cores of G at iteration i, and let ε_i be the amount at which they were raised at iteration i, $i = 1, \ldots, q$. Note that $y_S = \sum_{\{i:S\in\mathcal{C}_i\}} \varepsilon_i$ for any set $S \in \mathcal{S}$. Using this, together with the fact that the sets in $\mathcal{C}_i$ are disjoint and that $|\mathcal{C}_i| \geq \beta$ we get:

$$
\sum_{S\in\mathcal{S}} d_{F^+}(S) y_S = \sum_{i=1}^{q} \varepsilon_i \sum_{S\in\mathcal{C}_i} d_{F^+}(S) \leq \sum_{i=1}^{q} \varepsilon_i |F^+| \frac{|\mathcal{C}_i|}{\beta} = \frac{|F^+|}{\beta} \sum_{i=1}^{q} \varepsilon_i |\mathcal{C}_i| = \frac{|F^+|}{\beta} \sum_{S\in\mathcal{S}} y_S \leq \frac{|F^+|}{\beta}\frac{opt_k}{k-\ell}.
$$

The first inequality follows by upper bounding 1 by $|\mathcal{C}_i|/\beta$, and noting that in $\sum_{S\in\mathcal{C}_i} d_{F^+}(S)$ every edge is counted exactly once (since the graph we consider are directed). The last inequality follows from Lemma 3.2 and the Weak Duality Theorem. □

After executing Procedure 1, let $C_1, \ldots, C_{\nu^+}$ be the cores of $G + \tilde{F}^+$.

Procedure 2:
For $j = 1, \ldots, \nu^+$, choose $r_j \in C_j$, and compute an optimal edge set $\tilde{F}_j^+$ such that $G + \tilde{F}_j^+$ is $(\ell+1)$-outconnected from r_j.

Note that by Lemma 1.2, $c(\tilde{F}_j^+) \leq \frac{1}{k-\ell} opt_k$, $j = 1, \ldots, \nu^+$.

We then apply Procedures 1 and 2 on the reverse graph of $G+I$ to find appropriate edge sets $\tilde{F}^-$ and $\tilde{F}_1^-, \ldots, \tilde{F}_{\nu^-}^-$. Let $\tilde{F}$ be the union of all the edge sets found. Then $G + \tilde{F}$ is $(\ell+1)$-connected. The last step in our algorithm is finding an inclusion minimal edge set $F \subseteq \tilde{F}$ such that $G+F$ is $(\ell+1)$-connected. Note that $|\tilde{F}|$ might be large, but the following statement shows that $|F| = O(n)$.

Theorem 3.4 ([20]) *Let G be an ℓ-connected directed graph, and let F be an inclusion minimal augmenting edge set such that $G + F$ is $(\ell+1)$-connected. Then $|F| \leq 2n - 1$.*

Lemma 3.5 *The algorithm produces a feasible solution of cost at most $\frac{4\sqrt{2n}}{k-\ell} opt_k$.*

Proof: By Theorem 3.4, $|F| \leq 2n$. Set $F^+ = \tilde{F}^+ \cap F$, $F^- = \tilde{F}^- \cap F$, $F_j^+ = \tilde{F}_j^+ \cap F$ for $j = 1, \ldots, \nu^+$, and $F_j^- = \tilde{F}_j^- \cap F$ for $j = 1, \ldots, \nu^-$. Applying Lemmas 3.3 and 1.2, Theorem 3.4, and recalling that $\nu^+, \nu^- \leq \beta \leq \sqrt{2n}$, we get:

$$c(F) \leq c(F^+) + c(F^-) + \sum_{j=1}^{\nu^+} c(F_j^+) + \sum_{j=1}^{\nu^-} c(F_j^-) \leq \frac{2opt_k}{k-l}\left(\frac{2n}{\beta} + \beta\right) \leq \frac{4\sqrt{2n}}{k-\ell} opt_k.$$

□

Suppose now that the input graph $\mathcal{G}$ contains a k_0-connected spanning subgraph of cost zero. We can repeatedly apply the above algorithm starting with $\ell = k_0$ until $\ell = k-1$, to compute a k-connected spanning subgraph of $\mathcal{G}$; the overall cost of the subgraph found will be at most $4\sqrt{2n} H(k-k_0) opt_k = O(\sqrt{n} \ln k) opt_k$.

For undirected graphs, an $8\sqrt{2n}H(k-k_0)$-approximation algorithm easily follows using the reduction due to Khuller and Raghavachari [17] described at the end of Section 1.1.

To finish the proof of Theorem 3.1, let us discuss the implementation and the time complexity of the algorithm. As was mentioned, Procedure 1 in our algorithm is similar to the one used in [28], and we can adapt the implementation of [28] as well. We omit the details, but note that for implementing all Procedures 1 in the algorithm, as well as finding minimal edge sets $F \subseteq \tilde{F}$ such that $G + F$ is $(\ell+1)$-connected, $\ell = 0, \ldots, k-1$, can be done in $O(k^3 mn^2) = O(k^2 nm^2)$ total time, see Section 5 in [28]. Using the algorithm of [9], the overall time required for Procedure 2 implementations is $O(k^3 mn^2 \sqrt{n})$. Note however, that Procedure 2 requires finding a minimum-cost edge set to increase the outconnectivity from ℓ to $\ell+1$. As was mentioned, this problem can be solved in $O(m^2)$ time using Frank's algorithm [6]. Thus the total time required for Procedure 2 executions is $O(k\sqrt{n}m^2)$.

Acknowledgment The authors thank Joseph Cheriyan and an anonymous referee for their comments.

References

[1] V. Auletta, Y. Dinitz, Z. Nutov, and D. Parente, A 2-approximation algorithm for finding an optimum 3-vertex connected spanning subgraph, *Journal of Algorithms* 32, 1999, 21–30.

[2] J. Cheriyan, T. Jordán, and Z. Nutov, On rooted node-connectivity problems, *Algorithmica* 30 (special issue on APPROX'98), 2001, 353-375.

[3] J. Cheriyan and R. Thurimella, Approximating minimum-size k-connected spanning subgraphs via matching, *SIAM J. Comput.* 30, no. 2, 2000, 528–560.

[4] J. Cheriyan, S. Vempala, and A. Vetta, Approximation algorithms for minimum-cost k-vertex connected subgraphs, in *Proc. 34th ACM Symp. on Theory of Computing* (STOC 2002), 306–312.

[5] A. Frank, Connectivity augmentation problems in network design, *Mathematical Programming*, State of the Art, J. R. Birge and K. G. Murty eds., 1994, 34–63.

[6] A. Frank, Increasing the rooted connectivity of a digraph by one, *Math. Programming*, Vol. 84, Number 3 (1999), 565-576.

[7] A. Frank and T. Jordán, Minimal edge-coverings of pairs of sets, *J. Comb. Theory B* 65, 1995, 73–110.

[8] A. Frank and É. Tardos, An application of submodular flows, *Linear Algebra and its Applications* 114/115, 1989, 329–348.

[9] H. N. Gabow, A representation for crossing set families with application to submodular flow problems, *Proc. 4th Annual ACM-SIAM Symp. on Discrete Algorithms* (SODA 1993), 202–211.

[10] *Handbook of Combinatorics*, Vol. 1 Ch. 7, L. Graham, M. Grötschel, and L.Lovász Eds., 1995.

[11] T. Jordán, On the optimal vertex-connectivity augmentation, *J. Combinatorial Theory Series B* 63, 1995, 8–20.

[12] T. Jordán, On the existence of (k, ℓ)-critical graphs, *Discrete Math.* **179**, 1–3, 1998, 273–275.

[13] B. Jackson and T. Jordán, A near optimal algorithm for vertex connectivity augmentation, in Proc. ISAAC 2000, LNCS 1969 (2000), 212-325.

[14] K. Jain, A Factor 2 Approximation Algorithm for the Generalized Steiner Network Problem , *Combinatorica*, Vol 21-1, 2001, 39-60.

[15] G. Kortsarz, R. Krauthgamer, J. Lee, On the hardness of approximating vertex connectivity network design problems. *Approx 2002*, 185-199.

[16] G. Kortsarz and Z. Nutov, Approximating node connectivity problems via set covers, *Algorithmica* 37, 2003, 75–92.

[17] S. Khuller and B. Raghavachari, Improved approximation algorithms for uniform connectivity problems, *J. of Algorithms* 21, 1996, 434–450.

[18] M. Kriesell, Upper bounds to the number of vertices in a k-critically n-connected graph, *Graphs and Combinatorics* 18, 2002, no. 1, 133–146.

[19] W. Mader, Ecken vom Grad n in minimalen n-fach zusammenhängenden Graphen, *Archive der Mathematik* 23, 1972, 219–224.

[20] W. Mader, Minimal n-fach in minimalen n-fach zusammenhängenden Digraphen, *J. Combinatorial Theory Series B*, 38, 1985, 102–117.

[21] W. Mader, Endlichkeitsätze für k-kritische Graphen (German), *Math. Ann.* 229, 1977, 143–153.

[22] W. Mader, On k-con-critically n-connected graphs, *J. Combin. Theory Ser. B* 86, 2002, no. 2, 296–314.

[23] W. Mader, High connectivity keeping sets in n-connected graphs, to appear in *Combinatorica.*

[24] W. Mader, Connectivity and edge-connectivity in finite graphs, *Surveys in Combinatorics (Proc. Seventh British Combinatorial Conf., Cambridge, 1979)*, 66–95, London Math. Soc. Lecture Notes Ser., 38, Cambridge Univ. Press, 1979.

[25] W. Mader, On k-critically n-connected graphs, *Progress in Graph Theory* (Waterloo, Ont., 1982), 389–398, Academic Press, Toronto, On, 1984.

[26] B. Maurer and P. Slater, On k-critical, n-connected graphs, *Discrete Mathematics* 20, 1988, 255-262.

[27] Z. Nutov, Approximating connectivity augmentation problems, SODA 2005, to appear.

[28] R. Ravi and D. P. Williamson, An approximation algorithm for minimum-cost vertex-connectivity problems, *Algorithmica* 18, 1997, 21–43.

[29] R. Ravi and D. P. Williamson, Erratum: An approximation algorithm for minimum-cost vertex-connectivity problems, *Algorithmica* **34**(1), 2002, 98–107.

[30] J. Su, Proof of Slater's conjecture on k-critical n-connected graphs, *Kexue Tongbao* (English Ed.) 33, 1988, no. 20, 1675–1678.