Spanning trees with few non-leaves

Noga Alon *

Abstract

Let $f(n,k)$ denote the smallest number so that every connected graph with n vertices and minimum degree at least k contains a spanning tree in which the number of non-leaves is at most $f(n,k)$. An early result of Linial and Sturtevant asserting that $f(n,3) = 3n/4 + O(1)$ and a related conjecture suggested by Linial led to a significant amount of work studying this function. It is known that for n much larger than k, $f(n,k) \geq \frac{n}{k+1}(1-\varepsilon(k))\ln(k+1)$, where $\varepsilon(k)$ tends to zero as k tends to infinity. Here we prove that $f(n,k) \leq \frac{n}{k+1}(\ln(k+1)+4) - 2$. This improves the error term in the best known upper bound for the function, due to Caro, West and Yuster, which is $f(n,k) \leq \frac{n}{k+1}(\ln(k+1) + 0.5\sqrt{\ln(k+1)} + 145)$. The proof provides an efficient deterministic algorithm for finding such a spanning tree in any given input graph satisfying the assumptions.

1 Introduction

For integers $n > k > 1$ let $f(n,k)$ denote the smallest integer so that every connected graph with n vertices and minimum degree k contains a spanning tree with at most $f(n,k)$ non-leaves. Trivially $f(n,2) = n-2$ for all $n > 2$. An n-vertex graph obtained from $m = \lfloor \frac{n}{k+1} \rfloor$ cliques, each of size $k+1$ or $k+2$, by removing an edge from each clique and by connecting the resulting subgraphs along a cycle by adding edges that ensure the minimum degree is k shows that for every $n > k > 1$

$$f(n,k) \geq 3\frac{n}{k+1} - c(k). \tag{1}$$

An early unpublished result of Linial and Sturtevant [13] (c.f. [7], [5]) is that (1) is tight for $k = 3$ (the regular case has been proved even earlier by Storer [14].) Kleitman and West [11] proved that (1) is tight for $k = 4$, and Griggs and Wu [10] showed it is tight for $k = 5$ as well. Improved results assuming the graph does not contain certain subgraphs appear in [9], [4], [5]. The results above may suggest that for every fixed k, $f(n,k) \leq 3\frac{n}{k+1} - c(k)$, and indeed Linial suggested this conjecture (c.f., [10], [8]). This is, however, false, since as observed in [1] (and more explicitly in [2], [3]) with high probability a random k-regular graph on n vertices does not contain a spanning tree with less than $(1 - o_k(1))\frac{n\ln(k+1)}{k+1}$ non-leaves. Indeed, with high probability any dominating

*Department of Mathematics, Princeton University, Princeton, NJ 08544, USA and Schools of Mathematics and Computer Science, Tel Aviv University, Tel Aviv 69978, Israel. Email: nogaa@tau.ac.il. Research supported in part by NSF grant DMS-1855464 and by BSF grant 2018267.

set in such a graph has at least that many vertices. Despite this fact, this is an illustration of the remarkable ability of Nati Linial to raise interesting questions. Even when the answer to one of his questions is negative, and even when it is clear he has not spent much time thinking about it seriously, it often leads to extensive subsequent research.

For a connected graph $G = (V, E)$, let $\gamma(G)$ denote the minimum size of a dominating set in it, that is, the minimum cardinality of a set of vertices $X \subset V$ so that each $v \in V - X$ has at least one neighbor in X. Let $\gamma_c(G)$ denote the minimum size of a connected dominating set of G, that is, the minimum cardinality of a dominating set of vertices X so that the induced subgraph of G on X is connected. Note that $\gamma_c(G)$ is exactly the minimum possible number of non-leaves in a spanning tree of G. Thus the function $f(n, k)$ discussed above is exactly the maximum possible value of $\gamma_c(G)$, as G ranges over all connected graphs with n vertices and minimum degree at least k. Throughout the rest of this short paper it will be convenient to consider the parameter $\gamma_c(G)$ instead of the function $f(n, k)$.

It is well known that if the minimum degree in G is k and its number of vertices is n, then $\gamma(G) \leq \frac{n(\ln(k+1)+1)}{k+1}$. See [12] or [2], Theorem 1.2.2 for a proof. As mentioned above this is asymptotically tight for large k, see, e.g., [3] for a proof that for any $\varepsilon > 0$ and $k > k_0(\varepsilon)$ a random k-regular graph on n vertices is unlikely to contain a dominating set of size at most $(1-\varepsilon)\frac{n \ln k}{k}$.

Caro, West and Yuster [7] proved that for every connected graph G with n vertices and minimum degree k, $\gamma_c(G)$ is also not much larger than $\frac{n \ln(k+1)}{k+1}$. The precise statement of their result is as follows.

Theorem 1.1 ([7]). *Let G be a connected graph with n vertices and minimum degree at least k. Then*

$$\gamma_c(G) \leq \frac{n(\ln(k+1) + 0.5\sqrt{\ln(k+1)} + 145)}{k+1}$$

Here we prove several results improving the error term in this estimate, and consider the corresponding algorithmic problem.

2 Results

The first result we prove here, obtained together with Michael Krivelevich, is the following slight improvement of Theorem 1.1.

Theorem 2.1. *Let G be a connected graph with n vertices and minimum degree at least k. Then*

$$\gamma_c(G) \leq \frac{n(\ln(k+1) + \ln\lceil\ln(k+1)\rceil + 4)}{k+1}.$$

The main merit here is not the improved estimate, but the proof, which is much simpler than the one in [7]. Like the proof in [7], it provides a simple efficient algorithm for finding a connected dominating set of the required size for a given input graph. As a byproduct of the proof we get an upper bound for the difference between $\gamma(G)$ and $\gamma_c(G)$, as stated in the following theorem.

Define a monotone increasing piecewise linear function $f = f_{n,k}$ mapping $[1, \infty)$ to $[0, \infty)$ (or $[0, \infty)$ to $[-2, \infty)$) as follows. The pieces have x ranging from $(i-1)N$ to iN for $i = 1, 2, 3, \ldots$, where $N = n/(k+1)$, $f(0) = -2$ and the slopes in the pieces are $2, 1, 1/2, 1/3, 1/4, \ldots$ in order. An equivalent formulation is the following. For any real $x \geq 1$, let $x = (y+z)\frac{n}{k+1}$ with $y \geq 0$ an integer and $z \in [0,1]$ a real:

1. If $y = 0$ then $f(x) = \frac{n}{k+1} 2z - 2 = 2x - 2$
2. If $y = 1$ then $f(x) = \frac{n}{k+1}(\frac{z}{y} + 2) - 2 = \frac{n}{k+1}(z+2) - 2$.
3. If $y \geq 2$ then $f(x) = \frac{n}{k+1}(\frac{z}{y} + \frac{1}{y-1} + \cdots + \frac{1}{1} + 2) - 2$.

The function f is piecewise linear and monotone increasing. Its derivative, which exists in all points of $(1, \infty)$ besides the integral multiples of $\frac{n}{k+1}$, is (weakly) decreasing, thus f is concave. In addition it satisfies the following. For every $x = (w+z)\frac{n}{k+1} > \frac{n}{k+1}$ with $w \geq 1$ an integer and $z \in [0,1]$ a real, and for every integer w' satisfying $w \leq w' \leq x - 1$

$$f(x) \geq f(x - w') + 1 \tag{2}$$

Indeed, this follows from the fact that the slope of f is at least $\frac{1}{w}$ for every z in $(x - w', x]$ and thus $f(x) - f(x - w') \geq w' \cdot \frac{1}{w} \geq 1$.

Theorem 2.2. *Let G be a connected graph with n vertices, minimum degree at least k and domination number $\gamma = \gamma(G)$. Then $\gamma_c(G) \leq \gamma + f_{n,k}(\gamma)$. Therefore*

$$\gamma_c(G) < \gamma + \frac{n}{k+1}(\ln\lceil \ln(k+1) \rceil + 3).$$

We also describe an improved argument that provides a better estimate than the ones in Theorems 1.1, 2.1.

Theorem 2.3. *Let G be a connected graph with n vertices and minimum degree at least k. Then*

$$\gamma_c(G) \leq \frac{n}{k+1}(\ln(k+1) + 4) - 2.$$

The proof here too provides an efficient randomized algorithm for finding a connected dominating set with expected size as in the theorem. This algorithm can be derandomized and converted into an efficient deterministic algorithm.

3 Proofs

In the proofs we use the following simple lemma.

Lemma 3.1. *Let $G = (V, E)$ be a connected graph with n vertices and minimum degree at least k. Let $S \subset V$ be a dominating set of G, let H be the induced subgraph of G on S, and suppose the number of its connected components is $x = (y+z)\frac{n}{k+1}$ where y is a nonnegative integer and $0 \leq z \leq 1$ is a real. Then $\gamma_c(G) \leq |S| + f(x)$, where $f = f_{n,k}$ is the function defined in the previous subsection.*

Proof: Starting with the dominating set S we prove, by induction on x, that it is always possible to add to it at most $f(x)$ additional vertices to get a connected dominating set. For $x = 1$ the given set is already connected, and as $f(1) = 0$ the result in this case is trivial. If $1 < x \leq \frac{n}{k+1}$ we note that as long as there are at least two components, each one C can be merged to another one by adding at most two vertices. Indeed, every vertex in the second neighborhood of C is dominated, hence adding the two vertices of a path from C to any such vertex merges C to another component. This means that by adding at most $2(x-1) = f(x)$ vertices to S we get a connected dominating set, as needed.

If $x > \frac{n}{k+1}$ pick arbitrarily one vertex $v = v(C)$ in each of the x connected components of H and let $N(v)$ denote its closed neighborhood consisting of v and all its neighbors in G. This set is of size at least $k+1$. Therefore there is a vertex u of G that belongs to at least $\lceil (k+1)x/n \rceil$ of these closed neighborhoods. (This can in fact be slightly improved as none of the vertices of the dominating set belongs to more than one such closed neighborhood, but we do not use this improvement here). Define $S' = S \cup \{u\}$ and note that adding u merges at least $\lceil (k+1)x/n \rceil$ components. Therefore, if $x > w\frac{n}{k+1}$ for an integer $w \geq 1$, then the number of connected components of the induced subgraph of G on the dominating set S' is $\max\{x - w', 1\}$ for some integer $w' \geq w$. If this maximum is 1 we have added a single vertex to S to get a connected dominating set, and the required result clearly holds as $f(x) \geq 1$. Otherwise, by induction one can add to S' at most $f(x - w')$ additional vertices to get a connected dominating set, and the desired result follows from (2). $\square$

The proof clearly supplies an efficient deterministic algorithm for finding a connected dominating set of the required size, given the initial dominating set S.

Proof of Theorem 2.2: This is an immediate consequence of Lemma 3.1 together with the obvious fact that if $\gamma(G) = \gamma$ then G contains a dominating set S of size γ with at most $|S| = \gamma$ connected components. The known fact that $\gamma \leq \frac{n}{k+1}(\ln(k+1)+1)$ implies that $\gamma \leq \frac{n}{k+1}(y+z)$ with $y = \lceil \ln(k+1) \rceil$ and $z = 1$. The definition of the function $f = f_{n,k}$ thus implies that

$$f_{n,k}(\gamma) \leq \frac{n}{k+1}(\frac{1}{y} + \frac{1}{y-1} + \ldots + \frac{1}{1} + 2) - 2 < \frac{n}{k+1}(\ln y + 3),$$

completing the proof. $\square$

Proof of Theorem 2.1: This follows from Theorem 2.2 together with the fact that $\gamma(G) \leq \frac{n}{k+1}(\ln(k+1)+1)$. $\square$

In order to prove Theorem 2.3 we need two simple lemmas. The first one is a known fact, cf., e.g., [6], Formula (3.2). For completeness we include a short proof.

Lemma 3.2. *For a positive integer k and a real $p \in (0,1)$, let $B(k,p)$ denote the Binomial random variable with parameters k and p. Then the expectation of $\frac{1}{B(k,p)+1}$ satisfies*

$$E[\frac{1}{B(k,p)+1}] = \frac{1}{(k+1)p} - \frac{(1-p)^{k+1}}{(k+1)p}.$$

Proof: By definition

$$E[\frac{1}{B(k,p)+1}] = \sum_{i=0}^{k} \frac{1}{i+1}\binom{k}{i} p^i (1-p)^{k-i} = (1-p)^k \sum_{i=0}^{k} \frac{1}{i+1}\binom{k}{i} (\frac{p}{1-p})^i.$$

By the Binomial formula $(1+x)^k = \sum_{i=0}^{k} \binom{k}{i} x^i$. Integrating we get

$$\frac{(1+x)^{k+1}-1}{k+1} = \sum_{i=0}^{k} \frac{1}{i+1}\binom{k}{i} x^{i+1}.$$

Dividing by x and plugging $x = \frac{p}{1-p}$ the desired result follows. $\square$

Lemma 3.3. *Let $H = (V, E)$ be a graph. For every $v \in V$ let $d_H(v)$ denote the degree of v in H. Then the number of connected components of H is at most $D(H) = \sum_{v \in V} \frac{1}{d_H(v)+1}$.*

Proof: The contribution to $D(H)$ from the vertices in any connected component C of H with m vertices is

$$\sum_{v \in C} \frac{1}{d(v)+1} \geq \sum_{v \in C} \frac{1}{m} = 1.$$

$\square$

Proof of Theorem 2.3: Recall that the function $f = f_{n,k}$ defined in the previous subsection is concave. Therefore, by Jensen's Inequality, for every positive random variable X, $E[f(X)] \leq f(E[X])$.

Let $G = (V, E)$ be a connected graph with n vertices and minimum degree at least k. By Lemma 3.1 if there is a dominating set S of G and the induced subgraph of G on S has x connected components, then

$$\gamma_c(G) \leq |S| + f(x). \tag{3}$$

For a dominating set S, let $H = H(S)$ be the induced subgraph of G on S, and put $D(H) = \sum_{v \in S} \frac{1}{d_H(v)+1}$ where $d_H(v)$ is the degree of v in H. By Lemma 3.3 the number of connected components of H is at most $D(H)$, and since the function $f = f_{n,k}$ defined above is monotone increasing this implies, by (3), that

$$\gamma_c(G) \leq |S| + f(D(H)) = |S| + f(\sum_{v \in S} \frac{1}{d_H(v)+1}). \tag{4}$$

We next describe a random procedure for generating a dominating set S and complete the proof by upper bounding the expectation of the right-hand-side of (4). The procedure is the standard one described in [2], Theorem 1.1.2 for generating a dominating set. Define $p = \frac{\ln(k+1)}{k+1}$ and let T be a random set of vertices of G obtained by picking, randomly and independently, each vertex of G to be a member of T with probability p. Let $Y = Y_T$ be the set of all vertices of G that are not dominated by T, that is, all vertices in $V - T$ that have no neighbors in T. The set S defined by $S = T \cup Y_T$ is clearly dominating. The expected size of T is np. The expected size of Y_T is at most $n(1-p)^{k+1}$, since for any vertex v the probability it lies in Y_T is exactly $(1-p)^{d_G(v)+1} \leq (1-p)^{k+1}$, and the bound for the expectation of $|Y_T|$ follows by linearity of expectation. We proceed to bound the expectation of $f(\sum_{v \in S} \frac{1}{d_H(v)+1})$. By Jensen's Inequality and the convexity of f mentioned above this is at most $f(E[\sum_{v \in S} \frac{1}{d_H(v)+1}])$. Since f is monotone increasing it suffices to bound the expectation $E[\sum_{v \in S} \frac{1}{d_H(v)+1}]$.

Fix a vertex v. The probability it belongs to Y_T (and hence has degree 0 in H) is $(1-p)^{d+1}$, where d is its degree in G. The probability it belongs to T and has degree i in H is $p\binom{d}{i}p^i(1-p)^{d-i}$. Therefore, the expectation of $\frac{1}{d_H(v)+1}$ is, by Lemma 3.2,

$$(1-p)^{d+1} + p\left(\frac{1}{(d+1)p} - \frac{(1-p)^{d+1}}{(d+1)p}\right) < (1-p)^{k+1} + \frac{1}{k+1}.$$

Since $(1-p)^{k+1} \le e^{-p(k+1)} = \frac{1}{k+1}$ this implies, by linearity of expectation, that

$$E[\sum_{v \in S} \frac{1}{d_H(v)+1}] \le \frac{2n}{k+1}.$$

Using, again, linearity of expectation and the fact that $f_{n,k}(\frac{2n}{k+1}) = 3\frac{n}{k+1} - 2$ we conclude that the expectation of the right-hand-side of (4) is at most

$$np + n(1-p)^{k+1} + 3\frac{n}{k+1} - 2 \le \frac{n}{k+1}(\ln(k+1)+4) - 2.$$

Therefore there is a dominating set S for which this expression is at most the above quantity, completing the proof. $\square$

4 Algorithm

The proof of Theorem 2.3 clearly supplies an efficient randomized algorithm generating a connected dominating set of expected size at most as in the theorem in any given connected input graph $G = (V, E)$ with n vertices and minimum degree at least k. This algorithm can be derandomized using the method of conditional expectations, yielding a polynomial time deterministic algorithm for finding such a connected dominating set. Here is the argument. Let $v_1, v_2, \dots, v_n$ be an arbitrary numbering of the vertices of G. The algorithm generates a dominating set S satisfying

$$|S| + f(D(H)) = |T| + |Y_T| + f(\sum_{v \in S} \frac{1}{d_H(v)+1}) \le \frac{n}{k+1}(\ln(k+1)+4) - 2,$$

where $f = f_{n,k}$ is the function defined in the proof of Theorem 2.3, H is the induced subgraph of G on $S = T \cup Y_T$ and $D(H) = \sum_{v \in S} \frac{1}{d_H(v)+1}$. Once such an S is found it is clear that the proof of the theorem provides an efficient way to construct a connected dominating set of the required size using it.

The algorithm produces S as above by going over the vertices v_i in order, where in step i the algorithm decides whether or not to add v_i to S. Let S_i denote $S \cap \{v_1, v_2, \dots, v_i\}$. Thus $S_0 = \emptyset$. For each i, $0 \le i \le n$, define a potential function ψ_i in terms of the conditional expectations of $|S| = |T| + |Y_T|$ given S_i, which is denoted by $E[|S||S_i]$ and the conditional expectation of $\sum_{v\in S} \frac{1}{d_H(v)+1}$ given S_i, denoted by $E[\sum_{v\in S} \frac{1}{d_H(v)+1}|S_i]$. In this notation

$$\psi_i = E[|S||S_i] + f(E[D(H)|S_i] = E[|T||S_i] + E[|Y_T||S_i] + f(E[\sum_{v\in S} \frac{1}{d_H(v)+1}|S_i]).$$

Given the graph G and the set S_i, it is not difficult to compute ψ_i in polynomial time. Indeed, by linearity of expectation, the conditional expectation $E[|T||S_i]$ is computed by adding the contribution of each vertex $v = v_j$ to it. For $j \le i$ this contribution is 1 if $v_j \in T$ and 0 if $v_j \notin T$. For $j > i$ the contribution is p. The contribution of v_j to $E[Y_T|S_i]$ is 0 if v_j is already dominated by a vertex in S_i, and if it is not, then it is $(1-p)^s$, where s is the number of neighbors of v_j (including v_j itself if $j > i$) in the set $V - \{v_1, v_2, \ldots, v_i\}$.

The conditional expectation $E[\sum_{v\in S} \frac{1}{d_H(v)+1}|S_i]$ is also computed using linearity of expectation, where the contribution of each vertex v_j is $E[\frac{1}{d_H(v_j)+1}|S_i]$. This is also simple to compute in all cases. We describe here only one representative example. If $j > i$, q of the neighbors of v_j appear in S_i, and the number of its neighbors in G which lie in $V - \{v_1, v_2, \ldots, v_i\}$ is s, then

$$E[\frac{1}{d_H(v_j)+1}|S_i] = p \cdot \sum_{a=0}^{s} \binom{s}{a} p^a (1-p)^{s-a} \frac{1}{q+1+a}.$$

A similar expression exists in every other possible case.

Put $\psi_i = \psi_i^{(T)} + \psi_i^{(Y)} + \psi_i^{(f)}$, where $\psi_i(T) = E[|T||S_i]$, $\psi_i(Y) = E[|Y_T||S_i]$, and $\psi_i^{(f)} = f[E(D(H)|S_i]$. By the definition of conditional expectation

$$\psi_i^{(T)} = pE[|T| \mid S_{i+1} = S_i \cup v_{i+1}] + (1-p)E[|T| \mid S_{i+1} = S_i] \tag{5}$$

and

$$\psi_i^{(Y)} = pE[|Y_T| \mid S_{i+1} = S_i \cup v_{i+1}] + (1-p)E[|Y_T| \mid S_{i+1} = S_i] \tag{6}$$

Similarly, using the fact that the function f is concave

$$\psi_i^{(f)} = f(pE[\sum_{v\in H} \frac{1}{d_H(v_j)+1}|S_{i+1} = S_i \cup v_{i+1}] + (1-p)E[\sum_{v\in H} \frac{1}{d_H(v_j)+1}|S_{i+1} = S_i])$$

$$\ge pf(E[\sum_{v\in H} \frac{1}{d_H(v_j)+1}|S_{i+1} = S_i \cup v_{i+1}]) + (1-p)f(E[\sum_{v\in H} \frac{1}{d_H(v_j)+1}|S_{i+1} = S_i])$$

$$\ge \min\{f(E[\sum_{v\in H} \frac{1}{d_H(v_j)+1}|S_{i+1} = S_i \cup v_{i+1}])\ ,\ f(E[\sum_{v\in H} \frac{1}{d_H(v_j)+1}|S_{i+1} = S_i])\}.$$

Let ψ_{i+1}^{+} denote the value of ψ_{i+1} with $S_{i+1} = S_i \cup v_{i+1}$ and ψ_{i+1}^{-} denote the value of ψ_{i+1} with $S_{i+1} = S_i$.

By adding the last inequality and (5),(6) we conclude that

$$\psi_i \ge \min\{\psi_{i+1}^{+}, \psi_{i+1}^{-}\}.$$

Therefore, if the algorithm decides in each step $i+1$ whether or not to add v_{i+1} to S_i in order to get S_{i+1} by choosing the option that minimizes the value of ψ_{i+1}, then the potential function ψ_i

is a monotone decreasing function of i. Since ψ_0 is at most $\frac{n}{k+1}(\ln(k+1)+4)-2$ by the proof of Theorem 2.3, so is ψ_n. However, ψ_n is exactly $|S|+f(D(H))$ for the dominating set S constructed by the algorithm. This completes the description of the algorithm and its correctness.

5 Problem

We conclude this short paper with the following problem.

Problem: Determine or estimate the maximum possible value of the difference $\gamma_c(G)-\gamma(G)$, where the maximum is taken over all connected graphs G with n vertices and minimum degree at least k (≥ 3).

By Theorem 2.2 this maximum is at most $\frac{n}{k+1}(\ln\lceil\ln(k+1)\rceil+3)$. It is not difficult to show that it is at least $2\lfloor\frac{n}{k+1}\rfloor-O(1)$. This is shown by the example discussed in the introduction, as we proceed to describe. Assume, for simplicity, that $k+1$ divides n and put $m=\frac{n}{k+1}$. For each $0\leq i<m$ let K_i be the graph obtained from a clique on $k+1$ vertices by deleting a single edge x_iy_i. Let G be the k-regular graph obtained from the vertex disjoint union of the m graphs K_i by adding the edges y_ix_{i+1} for all $0\leq i<m$, where $x_m=x_0$. For this cycle of cliques G, $\gamma(G)=m=\frac{n}{k+1}$ as shown by a dominating set consisting of one vertex in each $K_i-\{x_i,y_i\}$ - this is a minimum dominating set as G is k-regular. On the other hand the induced subgraph on any connected dominating set must contain at least $m-1$ of the edges y_ix_{i+1} and their endpoints, and it is not difficult to check that it must contain at least one additional vertex in each of the cliques besides at most 2, and two additional vertices. Thus $\gamma_c(G)=2(m-1)+m=3\frac{n}{k+1}-2$. It will be interesting to close the $\ln\ln(k+1)$ gap between the upper and lower bounds and decide whether or not the above maximum is $\Theta(\frac{n}{k+1})$.

Acknowledgment I thank Michael Krivelevich for helpful discussions.

References

[1] N. Alon, Transversal numbers of uniform hypergraphs, Graphs and Combinatorics 6 (1990), 1-4.

[2] N. Alon and J. H. Spencer, The Probabilistic Method, Fourth Edition, Wiley, 2016, xiv+375 pp.

[3] N. Alon and N. Wormald, High degree graphs contain large-star factors, in "Fete of Combinatorics", Bolyai Soc. Math. Studies, 20 (Gy. Katona, A. Schrijver and T. Szönyi, eds.), Springer 2010, pp. 9-21.

[4] P. Bonsma, Spanning trees with many leaves in graphs with minimum degree three, SIAM J. Discrete Math. 22(3)(2008), 920–937.

[5] P. Bonsma and F. Zickfeld, Improved bounds for spanning trees with many leaves, Discrete Math. 312 (2012), 1178–1194.

[6] M. T. Chao and W. E. Strawderman, Negative Moments of Positive Random Variables, Journal of the American Statistical Association Vol. 67, No. 338 (1972), pp. 429–431.

[7] Y. Caro, D. B. West and R. Yuster, Connected domination and spanning trees with many leaves, SIAM J. Discrete Math. 13 (2000), no. 2, 202–211.

[8] G. Ding, T. Johnson, P. Seymour, Spanning trees with many leaves, J. Graph Theory 37(4)(2001), 189–197.

[9] J. R. Griggs, D. J. Kleitman, A. Shastri, Spanning trees with many leaves in cubic graphs, J. Graph Theory 13(6)(1989), 669–695.

[10] J. R. Griggs and M. Wu, Spanning trees in graphs of minimum degree 4 or 5, Discrete Math. 104(2)(1992), 167–183.

[11] D. J. Kleitman, D. B. West, Spanning trees with many leaves, SIAM J. Discrete Math. 4(1)(1991), 99–106.

[12] L. Lovász, On the ratio of optimal and integral fractional covers, Disc. Math. 13 (1975), 383–390.

[13] N. Linial and D.G. Sturtevant, Unpublished result, 1987.

[14] J. A. Storer, Constructing full spanning trees for cubic graphs, Inform. Process. Lett. 13 (1981), no. 1, 8–11.