Augmenting trees to meet biconnectivity and diameter constraints [1]

VICTOR CHEPOI AND YANN VAXES

Laboratoire d'Informatique de Marseille,
Université de la Méditerranée,
Faculté des Sciences de Luminy,
F-13288 Marseille Cedex 9, France

E-mail: {chepoi,vaxes}@lim.univ-mrs.fr

Abstract. Given a graph $G = (V, E)$ and a positive integer D, we consider the problem of finding a minimum number of new edges E' such that the augmented graph $G' = (V, E \cup E')$ is biconnected and has diameter no greater than D. In this note we show that this problem is NP-hard for all fixed D, by employing a reduction from the DOMINATING SET problem. We prove that the problem remains NP-hard even for forests and trees, but in this case we present approximation algorithms with worst-case bounds 3 (for even D) and 6 (for odd D). A closely related problem of finding a minimum number of edges such that the augmented graph has diameter no greater than D has been shown to be NP-hard by Schoone, Bodlaender, and van Leeuwen, *J. Graph Theory,* **11**(1987) when $D = 3$, and by Li, McCormick, and Simchi–Levi, *Operations Research Letters,* (1991) when $D = 2$.

Key Words: biconnectivity augmentation, diameter, radius, trees, approximation algorithms.

[1] Supported in part by France Telecom Grant CTI 00 1B 090

1. Preliminaries

The problem of augmenting a graph to reach a certain connectivity requirement by adding new edges is one of important problems of network reliability and fault-tolerant computing. In the most basic version, given an undirected graph $G = (V, E)$ one should add a minimum number of edges E' such that the augmented graph $G' = (V, E \cup E')$ is biconnected. Eswaran and Tarjan [9] characterized the minimum number of edges which must be added, leading to a linear time algorithm for this problem [20, 12]; for a survey of related problems and results see Khuller [14]. In this note we consider the biconnectivity augmentation problem with an additional diameter constraint:

Problem BADC (BICONNECTIVITY AUGMENTATION under DIAMETER CONSTRAINTS): *given a graph $G = (V, E)$ and a positive integer D, add a minimum number of new edges E' such that the augmented graph $G' = (V, E \cup E')$ is biconnected and has diameter no greater than D.*

Biconnectivity is a fundamental requirement to the topology of communication networks: a biconnected network survives any single link or single node failure (the probability of two or several simultaneous failures is much smaller). Since the delay of sending a message from node u to node v is roughly proportionnal to the number of nodes (or links) the message has to traverse, it is desirable to route the messages along paths as short as possible. Therefore a network having an underlying graph of small diameter ensures a low communication delay between any two nodes (the all-to-all communication model). The problem BADC can be viewed as a network improvement problem where G is the initial communication network and a minimum number of additional communication links must be added so that the upgraded network G' meets the biconnectivity and the diameter requirements.

In Section 2 we will show that the problem BADC is NP-hard for trees. Moreover, we will establish that this problem and the related

Problem ADC (AUGMENTATION under DIAMETER CONSTRAINTS): *given a graph G and a positive integer D, add a minimum number of edges to obtain a graph of diameter at most D*

are NP-hard for any fixed $D \geq 2$. Our reduction shows that both problems BADC and ADC are at least as difficult as the SET COVER or DOMINATING SET problems. From recent non–approximability results for SET COVER [1, 2, 10, 19] follows that unless P=NP there are no constant approximation polynomial time algorithms for both BADC and ADC and any fixed D. After completing a part of this work we learned about the papers [21, 17]; Schoone, Bodlaender and van Leeuwen [21] show that ADC is NP-hard when $D = 3$, while Li, McCormick, and Simchi–Levi [17] proved that the same problem is NP-hard when $D = 2$.

In Section 3 we present polynomial time approximation algorithms for the problem BADC (and for ADC and even values of D) in trees and forests. Additionally, we show that for trees one can solve in polynomial time the radius version of ADC (this version is motivated by the one-to-all communication model):

Problem ARC (AUGMENTATION under RADIUS CONSTRAINTS): *given a graph G and a positive integer R, add a minimum number of edges such that the augmented graph has radius no greater than R.*

Note that an optimal solution of ARC can be found by solving the following problem for each vertex and selecting a solution with minimum number of edges:

Problem AEC(G, R, b) (AUGMENTATION under ECCENTRICITY CONSTRAINTS): *Given a graph $G = (V, E)$, a positive integer R, and a vertex $b \in V$, find a minimum augmentation E' such that in the graph $G' = (V, E \cup E')$ the eccentricity of b is no greater than R.*

We conclude this introductory section with a few necessary definitions. A polynomial algorithm is called an *α-factor approximation* algorithm for a minimization problem Π if for each instance I of Π, it returns a solution whose value is at most α times the optimal value of I plus a constant not depending on I. For a graph $G = (V, E)$ and two vertices $u, v \in V$, we will denote by $d_G(u, v)$ the *distance* between these vertices (if u and v are in distinct connected components of G we will set $d_G(u, v) = \infty$). The *eccentricity* $e_G(u)$ of a vertex u is the distance to a vertex furthest from u. The *diameter* $\mathrm{diam}(G)$ and the *radius* $\mathrm{rad}(G)$ of G are respectively the largest and the smallest eccentricities of vertices of G. For a positive integer k and a vertex $u \in V$ let

$$B(u, k) = \{v \in V : d_G(u, v) \leq k\}$$

denote the *ball* centered at u of radius k. A *Helly graph* is a graph in which every family of pairwise intersecting balls has a nonempty intersection. Trees are simplest examples of Helly graphs; cf. [3].

2. NP–completeness

The decision variants of four problems formulated in the introduction evidently belong to the class NP. In all subsequent proofs of NP-completeness we will present pseudo-polynomial transformations from some known NP–complete in the strong sense problems to BADC, ADC, and AEC . From Lemma 4.1. of [11] will follow that the decision variants of all four problems are NP–complete as well.

2.1. General graphs

Proposition 1. *The problems BADC, ADC, ARC, and AEC are NP-hard for any fixed integers $D \geq 2$ and $R \geq 2$.*

Proof. We will present pseudo-polynomial transformations from the problems SET COVER and DOMINATING SET. We distinguish two cases.

Case 1. $D = 2$.

To settle this case, we need the following variant of SET COVER which we prove to be NP-hard.

Problem SET COVER$_2$ *A set X of n elements to be covered and a collection of subsets $\mathcal{S}$ of X such that every element of X belongs to at least two sets, while each pair of elements of X belongs to a common set of $\mathcal{S}$. Find a subcollection of sets $\mathcal{S}'$ that forms a cover and $c_{\mathcal{S}} := |\mathcal{S}'|$ is minimized.*

Claim. SET COVER$_2$ is NP-hard.

The given proof of the claim was suggested by the referee and uses the fact that it is NP-hard to find a cover consisting of exactly q sets for a family of sets, each of size 3, over a ground set of $3q$ elements (EXACT COVER BY 3-SETS). Indeed take an instance $\mathcal{S}^0$ of this problem and extend $\mathcal{S}^0$ to the collection $\mathcal{S}$ by adding all pairs of elements of X. Clearly, the instance $\mathcal{S}^0$ has "yes" answer if and only if the instance $\mathcal{S}$ of SET COVER$_2$ has an optimal solution of size q (sets of size 2 are useless).

We polynomially transform SET COVER$_2$ to BADC and ADC with $D = 2$. Let an arbitrary instance of SET COVER$_2$ be given by a set $X = \{x_1, \ldots, x_n\}$ and a collection $\mathcal{S} = \{S_1, \ldots, S_m\}$ of subsets of X. We construct a graph $G = (V, E)$ by creating a vertex for each element of X and each set of $\mathcal{S}$, and adding three new vertices a_1, a_2 and b. Define

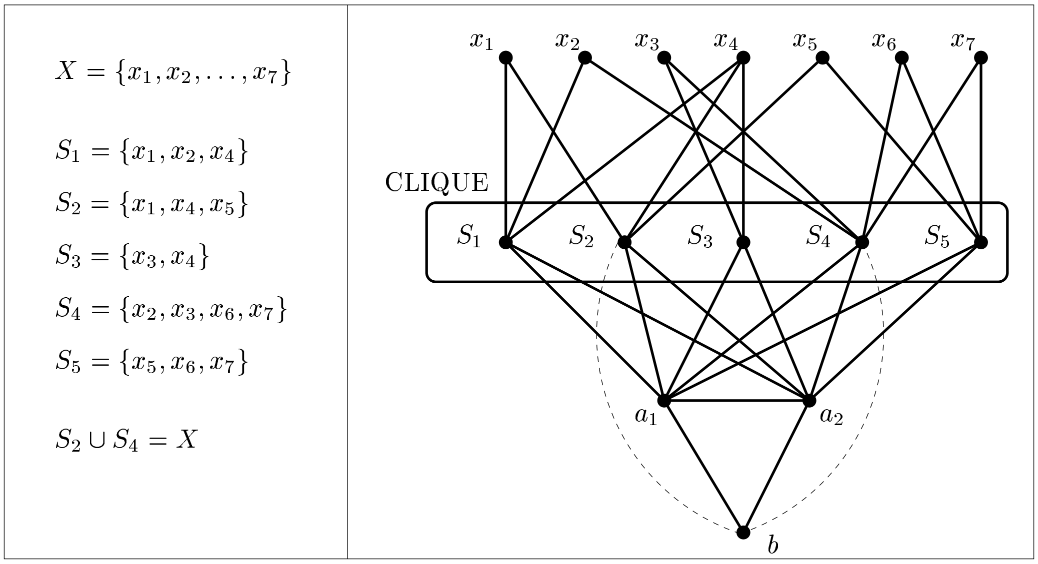

Figure 1:

an edge between each pair S_i, S_j of sets, between the vertices a_1 and a_2 and every S_j. An edge between x_i and S_j exists if and only if $x_i \in S_j$. Additionally assume that b is adjacent only with the vertices a_1 and a_2; see Figure 1 for an illustration. The resulting graph G is biconnected and all pairwise distances except $d_H(b, x_i)$ are at most 2.

We assert that $c_{\mathcal{S}} \leq k$ if and only if there is a solution of BADC (or ADC) with at most k edges. If $\mathcal{S}' \subseteq \mathcal{S}$ is a set cover, then adding an edge between b and each vertex representing a subset of $\mathcal{S}'$ we will get a graph of diameter 2. Conversely, suppose there exists a feasible augmentation of G using at most k edges. One can easily check that it can be replaced by a solution E' of the same size which consists solely of edges connecting b with vertices designing subsets of $\mathcal{S}$. Indeed, an edge x_iS_j can be replaced by an edge bS_l, where $x_i \in S_l$. The edge bx_i and all edges of the type x_ix_j can be replaced by the same number of edges between b and subsets containing the corresponding elements. The same operation can be performed with all edges of the type a_lx_i. Thus all edges of E' have the form bS_j. The collection $\mathcal{S}' = \{S_j : bS_j \in E'\}$ covers X, settling Case 1. This shows also that ARC and AEC are NP-hard for $R = 2$.

Case 2. $D \geq 3$.

Set $D := 2R + 2$ for even D and $D := 2R + 1$ for odd D, where $R > 0$. We will present a pseudo-polynomial transformation from DOMINATING SET. (Recall its formulation: given a graph $G = (V, E)$ and an integer $k > 0$, is there a subset S of vertices with $|S| \leq k$ whose neighbourhoods $B(s, 1), s \in S$, cover V?) To construct an instance of the problem ADC we proceed in the following way. First, we take a copy V' of V, transform it into a clique, and make adjacent every vertex $v' \in V'$ with its twin $v \in V$ and all neighbours of v in G. Then add a vertex a adjacent to all vertices of V' and a vertex b adjacent to a. For each vertex $v \in V$ add a path P_v of length $R - 1$ issuing from v. Finally, add $n := 2|V| + 1$ paths $Q_1, \ldots, Q_n$ with one end in b, and each having length $R + 1$ if $D = 2R + 2$ and length R if $D = 2R + 1$. Denote the resulting graph by H. As an instance of the problem AEC we consider the graph H, the vertex b, and the eccentricity $R + 1$. For an illustration of these constructions see Figure 2. To define an instance of the problem BADC, we take a copy H' of H and make adjacent every vertex v' of H' with its twin v and all its neighbours in H. The resulting graph H^* is biconnected.

We assert that G has a dominating set of size k if and only if the problems ADC and AEC on the graph H or the problem BADC on the graph H^* have a solution consisting of k edges. Notice that $d_H(x, y), d_{H^*}(x, y) \leq D$, unless x is the end-vertex of a Q-path and y is the end-vertex of a P-path or vice versa (further, we will suppose that x and y are vertices of first type). If S is a dominating set of G, then we may define a feasible solution of the problems ADC, AEC, and BADC of size $|S|$ by adding the edges $bs, s \in S$. The eccentricity of b in the augmented graph Γ is $R+1$: indeed, $d_\Gamma(b, y) = 1 + R - 1 = R$ if y is the end-vertex of a path P_s with $s \in S$, and $d_\Gamma(b, y) = 2 + R - 1 = R + 1$ otherwise. This implies that $d_\Gamma(x, y) = d_\Gamma(x, b) + R = D - 1$ in the first case and $d_\Gamma(x, y) = d_\Gamma(x, b) + R + 1 = D$ in the second case, therefore the diameter of Γ is D. Conversely, assume that there is a solution of ADC on H (respectively, BADC on H^*) consisting of $k (\leq |V|)$ edges. From the choice of n one can easily deduce that at least one of the paths $Q_1, \ldots, Q_n$, does not contain vertices incident to added edges. This means that k is at least as large as the number of edges in a solution of the problem AEC. Since the converse also holds, we may find a solution of AEC$(H, R + 1, b)$ consisting of k edges. We need an auxilliary result.

Lemma 1. *Given a graph H, a vertex b, and a positive integer r, there is an optimal solution of the problem AEC(H, r, b) consisting entirely of edges incident to b (i.e., a solution in the form of a star centered at b).*

Proof. Let E' be an arbitrary optimal solution of AEC. Denote by H' the graph obtained

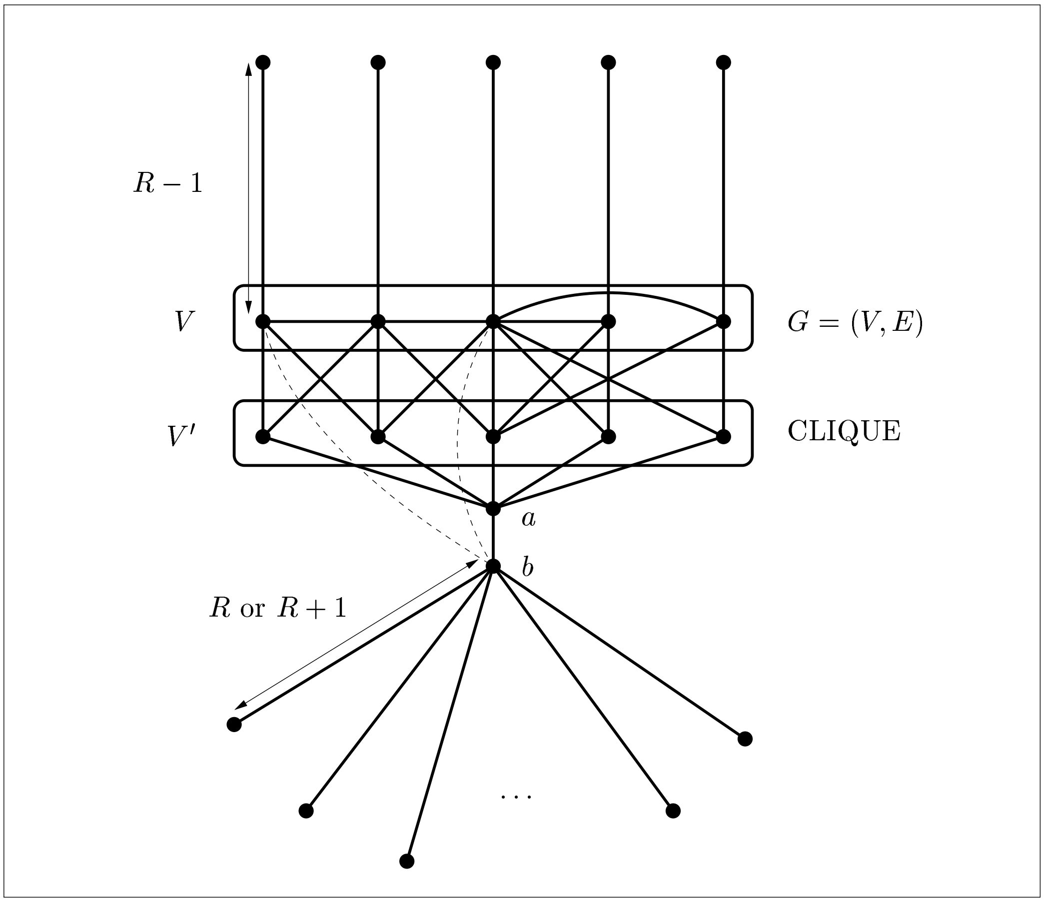

Figure 2:

from H by adding the edges of E'. Every edge $uv \in E'$, $u, v \neq b$, belongs to a shortest path of H' (of length at most r) connecting b with a vertex q (otherwise we can remove the edge uv from E'). This implies that the vertices u and v are not equidistant from b. Suppose $d_{H'}(u,b) < d_{H'}(v,b)$. Replace in E' every edge uv by the edge vb. Since in the new graph H'' still $d_{H''}(b,q) \leq r$, we get an optimal solution of AEC(H,r,b) in the form of a star centered at b. $\square$

By Lemma 1 there is a solution of AEC$(H,R+1,b)$ which induces a star centered at b. Replace in this star every edge bu such that either $u = v' \in V'$ or $u \in P_v$ by the edge bv. Clearly, the new set of edges is a solution of AEC$(H,R+1,b)$ and also defines a star centered at b. The tips of this star constitute a subset S of size at most k of the vertex–set of G. We assert that S is a dominating set of G. Suppose not; then there exists a vertex

$v \in V$ at distance ≥ 2 from every vertex of S. Let p be the end-vertex of the path P_v. Then $d_H(b,p) = d_H(b,v) + d(v,p) \geq 3 + R - 1 = R + 2$, contrary to the fact that we have a solution of AEC$(H, R+1, b)$. Thus S is a dominating set of G, concluding the proof of Proposition 1. $\square$

The proof of Proposition 1 shows that every feasible solution of size k of each of the problems BADC, ADC, ARC, and AEC leads to a feasible solution of the same size for the corresponding problem SET COVER$_2$ ($D = 2$) and DOMINATING SET ($D \geq 3$) and vice versa. Therefore an α-factor approximation algorithm for anyone of the four first problems would lead to an α-factor approximation algorithm for the last two problems. On the other hand, there is a simple polynomial reduction showing that an α-factor approximation algorithm for SET COVER$_2$ would lead to a 2α-factor approximation algorithm for SET COVER. Indeed, take an instance $\mathcal{S}_0$ of SET COVER and extend $\mathcal{S}_0$ to the collection $\mathcal{S}$ as in the proof of the claim. Clearly, $c_{\mathcal{S}} \leq c_{\mathcal{S}_0}$. Let $\mathcal{S}'$ be a solution of the instance $\mathcal{S}$ returned by an α-factor approximation algorithm for SET COVER$_2$. We can derive a feasible solution $\mathcal{S}'_0$ for SET COVER by replacing every 2-element set of $\mathcal{S}'$ which is not in a set of $\mathcal{S}_0$ by one or two sets of $\mathcal{S}_0$ containing the corresponding elements. Since $|\mathcal{S}'_0| \leq 2|\mathcal{S}'| \leq 2\alpha c_{\mathcal{S}} \leq 2\alpha c_{\mathcal{S}_0}$, we are done. Hence, our four augmentation problems are at least as difficult as SET COVER which is $\Omega(\log n)$-hard unless $P = NP$ [10, 1, 19].

On the other hand, one can solve ARC and AEC using SET COVER. As we noticed before, a solution for ARC can be obtained by solving the problem AEC for each vertex of the input graph $G = (V, E)$. To solve AEC(G, R, b) we proceed in the following way. By Lemma 1 there is a solution in the form of a star and we will search a solution of this type. Notice that in the augmented in this way graph every shortest path issuing from b contains at most one new edge. For a potential new edge bu let S_u be a subset of V consisting of all vertices $v \in V$ such that $d_G(b,v) > R$ but $d_{G_u}(b,v) \leq R$, where G_u is the graph obtained from G by adding the edge bu. Denote by $\mathcal{S}$ the resulting collection of subsets $S_u, u \in V$. If $\mathcal{S}'$ is a minimum cover of $\mathcal{S}$, then $E' = \{bu : S_u \in \mathcal{S}'\}$ is an optimal solution of AEC(G, R, b). Similarly, a version of ADC asking for a minimum augmentation of a graph G to obtain a graph in which every pair of vertices can be connected by a path of length at most D which uses at most one new edge can be solved using SET COVER (note that the proof of Proposition 1 can be easily adapted to show that this problem is NP-hard as well). For vertices u, v let G_{uv} be the graph G plus the edge uv, and denote by S_{uv} the set of all pairs

xy, such that $d_{G_{uv}}(x,y) \le D$ and $d_G(x,y) > D$. Then taking a minimum set cover of the collection $\mathcal{S} = \{S_{uv} : u,v \in V\}$ we obtain an optimal solution of this version of ADC.

2.2. Trees and forests

PROPOSITION 2. *The problem BADC remains NP-hard for trees.*

PROOF. We will present a pseudo-polynomial transformation from 3-PARTITION. Recall the formulation of this basic NP-complete problem.

3-PARTITION: *Given a set A of $3m$ elements, a bound $B \in Z^+$, and a size $s(a) \in Z^+$, such that each $s(a)$ satisfies $B/4 < s(a) < B/2$ and such that $\sum_{a\in A} s(a) = mB$, can A be partitioned into m disjoint sets $S_1, S_2, \ldots, S_m$ such that, for $1 \le i \le m$, $\sum_{a\in S_i} s(a) = B$?*

(The above constraints on the $s(a)$'s imply that every S_i must contain exactly three elements from A). First we will present a transformation from 3-PARTITION to BADC on forests (which seems especially simple and elegant), and then we will show how to modify this construction to get a transformation to the problem on trees.

Let $A = \{a_1, a_2, \ldots, a_{3m}\}, B \in Z^+$, and $s(a_1), s(a_2), \ldots, s(a_{3m}) \in Z^+$ be an arbitrary instance of 3-PARTITION. Set $D := B + 6$. Define a forest F consisting of a "bistar" S formed by a path P_0 of length $D+1$ plus m leaves at each end, and $3m$ paths, where the ith path P_i corresponds to the element $a_i \in A$ and has length $s(a_i)$; see left part of Figure 3. Let x_0 and y_0 be the end-vertices of P_0. We assert that 3-PARTITION has answer "yes" if and only if there exists a solution of the problem BADC in F with $D = B + 6$ which has at most $4m$ edges.

The forest F contains $8m$ leaves. Since any leaf needs a new edge in a biconnected augmentation, the smallest biconnectivity augmentation of F consists of exactly $4m$ edges, therefore any feasible solution of the corresponding instance of BADC contains at least $4m$ edges. Every biconnected graph H obtained from F by adding exactly $4m$ edges has a rather precise shape. Namely, H consists of m cycles $C_1, C_2, \ldots, C_m$ glued together along the path P_0. We will call the content of a cycle C_i minus the edges of P_0 an *ear* and denote it by Ear_i. It consists of two edges of S, one incident to x_0 and another to y_0 (after a suitable relabelling, we may assume without loss of generality that these are the edges x_0x_i and y_0y_i), a certain number of paths (which maybe equal to 0 in the degenerate case), and some new edges each connecting either the end–vertices of two paths or an end–vertex of a path with x_i or y_i (in the degenerate case, Ear_i is formed by the edges x_0x_i, y_0y_i and the new edge

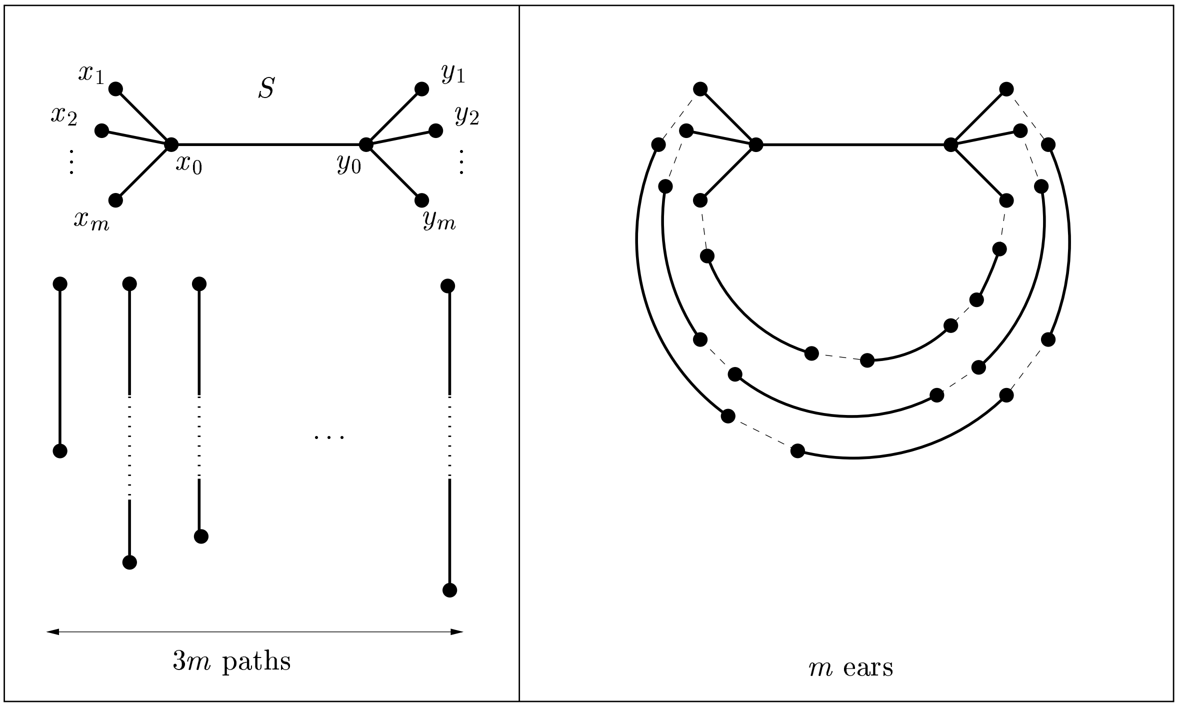

Figure 3:

x_iy_i). For an illustration see the right part of Figure 3. Let l_i be the length of Ear_i. Clearly $\sum_{i=1}^{m} l_i = mB + 6m$. We will call an ear *big, normal,* or *small* if its length is, respectively, larger, equal, or smaller than $B + 6$.

First, suppose that the required 3-partition $S_1, S_2, \ldots, S_m$ of A exists. Then biconnect F by adding the following $4m$ new edges: for every $S_i = \{a_{i_1}, a_{i_2}, a_{i_3}\}$ add an edge between x_i and one end of P_{i_1}, an edge between the second end of P_{i_1} and one end of P_{i_2}, an edge between the second end of P_{i_2} and one end of P_{i_3}, and, finally, an edge between the second end of P_{i_3} and the vertex y_i; see Figure 3. All ears of the resulting graph H are normal, therefore the distance in H between two vertices located on different ears is at most D. Every cycle C_i has length $2B + 13 = 2D + 1$, thus the distance between a vertex on Ear_i and a vertex on P_0 does not exceed D as well. This shows that the diameter of H is D.

Now suppose that we are given a solution of BADC consisting of $4m$ edges. Let H be the augmented graph. To establish that a required 3-partition of A exists, it suffices to show that all ears of H are normal. Assume the contrary, i.e, the graph H contains a big ear, say Ear_1. The length of the cycle C_1 is at least $2B + 14$, therefore the distance in H between the middle vertex z_0 (one of the two middle vertices if $B+7$ is odd) of P_0 and its opposite vertex (or one of its opposite vertices if $|C_1|$ is odd) in C_1 is larger than $D = B+6$, contrary to the choice of H. This establishes that BADC is NP-hard for forests.

To establish the same result for trees we have to modify the previous construction. First, instead of taking paths of length $s(a_i)$ we will take paths of length $l_i := 24s(a_i)$. Set $D := 24B + 6$ and, as in the previous case, let P_0 be a path of length $D+1$ between x_0 and y_0. Let z_i denote the middle vertex of the path P_i for $i = 1, ..., n$ and let z_0 be the middle vertex of P_0 at distance $12B + 4$ from x_0. As an instance of BADC we define a tree T obtained from the forest F (with updated lengths) by adding a star centered at a new vertex c_0 : it consists of a path Q_0 of length $13B + 3$ between c_0 and z_0, and $3m$ paths Q_i of length $8B + 1$ joining the vertex c_0 with the vertex $z_i, i = 1, \ldots, m$; see Figure 4. The paths $Q_0, Q_1, \ldots, Q_{3m}$ pairwise intersect only in the vertex c_0. As before, we assert that 3-PARTITION has answer "yes" if and only if there is a solution of BADC on T with $D = 24B + 6$ using at most $4m$ edges. As in the case of forests, any biconnected augmentation of T needs at least $4m$ edges, and every biconnected graph H obtained from T by adding exactly $4m$ edges has the same form: m ears $\mathrm{Ear}_1, \ldots, \mathrm{Ear}_m$, the path P_0, and, additionally, the paths $Q_0, Q_1, \ldots, Q_{3m}$.

Since every $s(a_i)$ is an integer and $B/4 < s(a_i) < B/2$, one can easily deduce that

$$6B + 6 \leq l_i \leq 12B - 12$$

and

$$3B + 3 \leq l_i/2 \leq 6B - 6.$$

First, suppose that the required 3–partition $\{S_1, S_2, \ldots, S_m\}$ of A exists. Biconnect T as in the case of forests, previously ordering the elements of each triplet $S_i = \{a_{i_1}, a_{i_2}, a_{i_3}\}$ so that $a_{i_1} \leq a_{i_3} \leq a_{i_2}$. Then obviously $l_{i_1} \leq l_{i_3} \leq l_{i_2}$, and, since $l_{i_1} + l_{i_2} + l_{i_3} = 24B$, we conclude that $l_{i_3} \leq (24B - (6B + 6))/2 = 9B - 3$ and $(l_{i_1} + l_{i_3}) \leq 16B$. We assert that the diameter of H is at most D. Pick two vertices u, v of H. If one vertex is located on a path Q_i, $i > 0$, and another vertex is either on some ear or on some path Q_j then

$$d_H(u, v) \leq (8B + 1) + (8B + 1) + (6B - 6) = 22B - 4 < D.$$

In remaining cases, we assert that u and v lie on a common cycle C of length at most $2D+1$, yielding $d_H(u, v) \leq D$. This assertion obviously holds if both vertices u, v are located on ears, or one vertex is on an ear and another one is on the path P_0. If $u \in Q_0$ and $v \in \mathrm{Ear}_i$, then the role of C is played by the cycle consisting of Q_0, Q_{i_2}, and the half of C_i between z_{i_2} and z_0 which passes via v. Since the unique path P_j of Ear_i that belongs entirely to C is not the longest path of this ear, its length is at most $9B-3$. Hence the length of C is at most

$$(12B+4)+(13B+3)+(8B+1)+(6B-6)+(9B-3)+3 = 48B+2 < 2D+1.$$

Finally, suppose that $u \in P_0$ and $v \in Q_j$, $j > 0$. Two subcases can be distinguished. If $j = i_2$, where $S_i = \{a_{i_1}, a_{i_2}, a_{i_3}\}$, then let C be the cycle consisting of Q_0, Q_{i_2}, and the half of C_i between z_0 and z_{i_2} passing via u. Since C is a cycle of the same kind as in the previous case, its length is at most $2D+1$. If $j = i_3$ or i_1, then let C be the cycle consisting of Q_{i_1}, Q_{i_3} and the half of C_i between z_{i_1} and z_{i_3} passing via u. Since, $(l_{i_1}+l_{i_3})/2 \leq 8B$, we conclude that the length of C is at most

$$(24B+7)+(16B+2)+8B+4 = 48B+13 = 2D+1.$$

This shows that the graph H has indeed diameter D.

To establish the converse, let H be a solution of BADC obtained from T by adding exactly $4m$ edges. To show that 3-PARTITION has answer "yes" it suffices to prove that all ears of H are normal. We proceed in two stages: first we prove that every ear consists of three paths and then we show that they are normal. Since the $3m$ paths are distributed over m ears, either all ears have three paths each or there exists an ear containing at least four paths. Suppose by way of contradiction that Ear_i contains at least four paths $P_{i_1}, P_{i_2}, P_{i_3}, P_{i_4}$ ordered as they occur in C_i starting from x_i. Let v be the furthest from x_i end-vertex of the path P_{i_2}. First, note that every (z_0, v)-path passing via c_0 has length at least

$$(13B+3)+(8B+1)+(3B+3) = 24B+7 > D.$$

On the other hand, every (z_0, v)-path passing via x_0 or y_0 has length at least

$$(12B+3)+2+(6B+6)+1+(6B+6) = 24B+18 > D.$$

The obtained contradiction shows that every ear comprises precisely three paths, thus its length is at least $3(6B+6)+6 = 18B+24$. We continue by showing that all ears of H are actually normal. Suppose not, and let Ear_i be big. Pick the vertex $u \in P_0$ located at distance $7B+2$ from z_0 and at distance $5B+2$ from x_0. Let v be the vertex opposite to u in the cycle C_i. From the choice of Ear_i, each of the two (u,v)paths of C_i has length

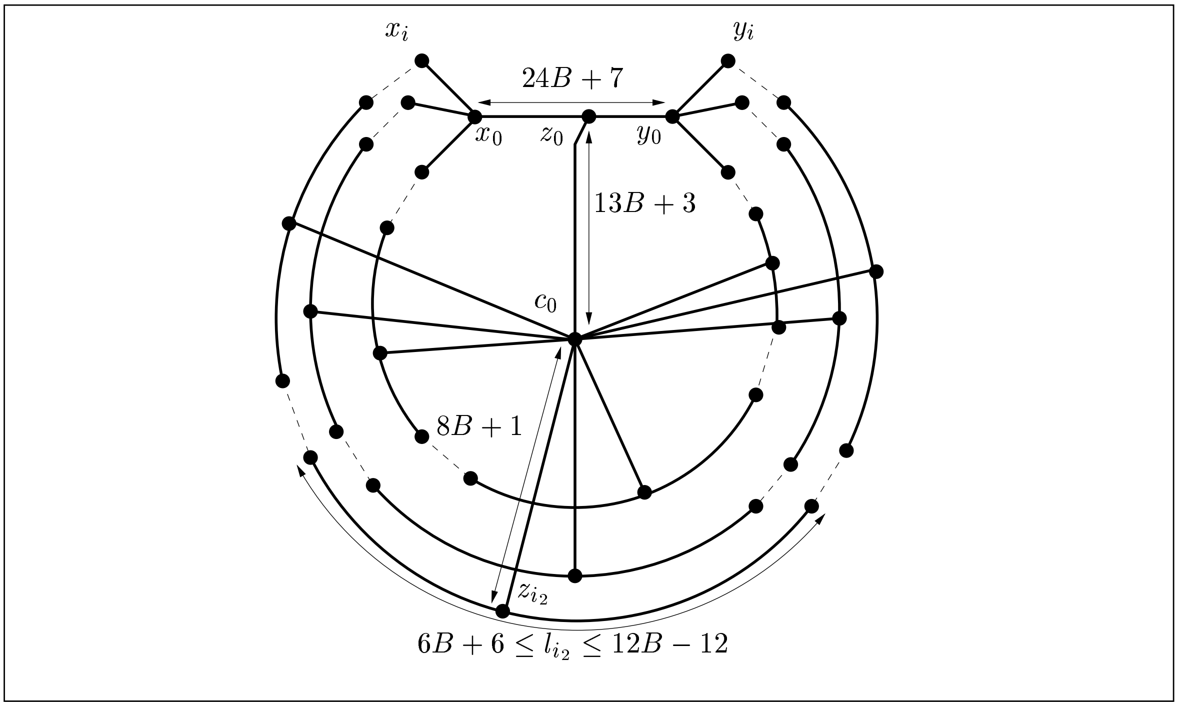

Figure 4:

greater than D. On the other hand, since the length of each ear is at least $18B+24$, one conclude that the subpath of P_0 comprised between u and y_0 is a shortest path between these vertices in H. Hence $d_H(u,y_0) = (7B+2)+(12B+3) = 19B+5$. In order to have $d_H(u,v) \le D$, no shortest (u,v)-path can pass via the vertex y_0. Thus such a path must pass via c_0, whence $d_H(u,v) = d_H(u,c_0) + d_H(c_0,v)$. Notice that $d_H(c_0,v) \ge 8B+1$. On the other hand, we assert that $d_H(u,c_0) \ge 16B+8$. Indeed, the length of every (u,c_0)-path passing via x_0, some vertex x_j, one half of the path P_{j_1}, and the path Q_{j_1} is at least $(5B+2)+2+(3B+3)+(8B+1) = 16B+8$, while the length of the (u,c_0)-path passing via z_0 is $(7B+2)+(13B+3) = 20B+5$. Hence $d_H(u,c_0) \ge 16B+8$, and, as a consequence, $d_H(u,v) \ge 24B+9$, contrary to the choice of H. This shows that all ears of H are normal, whence there exists a feasible 3–partition of A, thus completing the proof of Proposition 2. □

3. Approximation algorithms for trees and forests

In this section we present polynomial time approximation algorithms for solving the problems BADC and ADC for trees, forests, and, more generally, for graphs on which the domination and the k-domination problems can be solved efficiently. Our solution is mainly based on the linear time algorithm for solving k-DOMINATING SET problem on these graphs and the relationship between this problem and the problem AEC. Recall, given a graph $G = (V, E)$ and a set $X \subseteq V$, the k-DOMINATING SET consists in finding a covering of X with a minimum number of balls of radius k. This problem can be solved in linear time for trees [7, 13], sun–free chordal graphs [8], and dually chordal graphs [6] (*dually chordal graphs* are the Helly graphs whose intersection graph of balls is chordal, or, equivalently, graphs whose ball hypergraph is a hypertree), and in polynomial time for all Helly graphs whose intersection graph of balls is perfect.

3.1. Even diameter. We continue by showing how to reduce the problem AEC(G, R, b) to k-DOMINATING SET. For this, set $X := V - B(b, R)$ and $k := R - 1$. For an augmentation E' of G which induces a star centered at b denote $S' = \{s : bs \in E'\}$ and $G' = (V, E \cup E')$ (by Lemma 1 such a star-solution always exists); see Figure 5 for an illustration.

LEMMA 2. *The set of new edges E' is a solution of AEC(G, R, b) if and only if S' is a solution of given instance of the k-DOMINATING SET problem.*

PROOF. If the set X is covered by the balls $B(s, R - 1), s \in S'$, then for any vertex $x \in X$ there is a vertex $s \in S$ such that $d_{G'}(x, b) \leq d_G(x, s) + 1 \leq R$. Notice that s is not adjacent to b in G, otherwise $x \in B(b, R)$. Thus E' is a feasible solution for the eccentricity problem. Conversely, if the eccentricity of b in the augmented graph G' is at most R and $v \in V$, then either $d_G(b, v) \leq R$ and then v belongs to the ball $B(b, R)$, or a shortest path of G' between v and b uses a new edge bs. Since this path does not contain other new edges, we have $d_G(s, v) \leq R - 1$. This shows that S' is a solution for k-DOMINATING SET. $\square$

To solve the problem ARC on a graph G, we must solve AEC(G, R, b) for every vertex b and among all solutions to pick an admissible augmentation with least number of edges. The complexity of this procedure is $O(|V||E|)$ for all graphs G in which the k-DOMINATING SET can be solved in linear $O(|E|)$ time, in particular for trees.

LEMMA 3. *If G is a Helly graph then an optimal solution for the problem ARC constructed by previous method is a 2-factor approximation for the problem ADC with $D = 2R$.*

PROOF. Obviously, if we take a feasible solution for ARC in the form of a star, then the augmented graph has diameter at most $2R$. Conversely, let E' be an optimal solution for

ADC and let G' be the augmented graph. Denote by P the set of end-vertices of the edges from E'. Take around each vertex $p \in P$ the ball $B(p, R-1)$, and let Q be the set of vertices of G not covered by such balls, i.e., $Q := V - \bigcup\{B(p, R-1) : p \in P\}$. We assert that the diameter of Q is at most D, in other words, that $d_G(u,v) \leq D$ for any $u, v \in Q$. Suppose not and let $d_G(u,v) > D$ for some $u, v \in Q$. Since $d_{G'}(u,v) \leq D$, every shortest path in G' between u and v will use at least one new edge. This means that such a path consists of three tiles: a subpath from u until the first vertex p_u from P, followed by a tile consisting of subpaths of G and one or several new edges ending with the last vertex p_v of P, and, finally, a subpath of G from p_v to v. Since $u \notin B(p_u, R-1)$ and $v \notin B(p_v, R-1)$, we deduce that $d_{G'}(u,v) \geq R+1+R > D$, contrary to the choice of E'. Thus the distance in G between any two vertices of Q is at most $2R$, thus the balls $B(q,R), q \in Q$, pairwise intersect. Since G is a Helly graph, these balls have a vertex b in common (which is not necessarily a vertex of Q). Set $E'' = \{bp : p \in P\} - E$. Clearly E'' is an admissible solution for the problem AEC(G, R, b). Since $|E''| \leq 2|E'|$ (the worst case occurs when E' is a matching on P), we obtain the desired inequality. $\square$

PROPOSITION 3. *There is a 3-factor approximation algorithm for the problem BADC with $D = 2R$ on trees and forests (and, more generally, on dually chordal graphs) with complexity $O(|V||E|)$.*

PROOF. The algorithm constructs an admissible augmentation by solving separately the biconnectivity augmentation problem and the problem ARC. Let c' edges were used to biconnect the input graph and other c'' edges were used to solve the augmentation problem under radius constraints. Suppose that an optimal solution for BADC has c edges. Since $c \geq c'$ and $c'' \leq 2c$ by Lemma 3, we are done. $\square$

One can construct examples of trees for which the heuristic employed to solve the problem ADC does not find an optimal solution. If, for example, we take the path on 5 vertices and $D = 2$, then our algorithm will add two edges, while the optimal augmentation uses only one edge and the augmented graph is the 5-cycle. Next, if we consider the path on 10 vertices and $D = 2$, then we will add 7 edges, while the optimal augmentation uses 6 edges and the augmented graph is the Petersen graph. One of the main results of [21] establishes that for $n \geq 11$ and $D = 2$ an optimal augmentation of a path on n vertices has $n-3$ edges, i.e., our augmentation in the form of a star is optimal. Nethertheless, we were not able to find

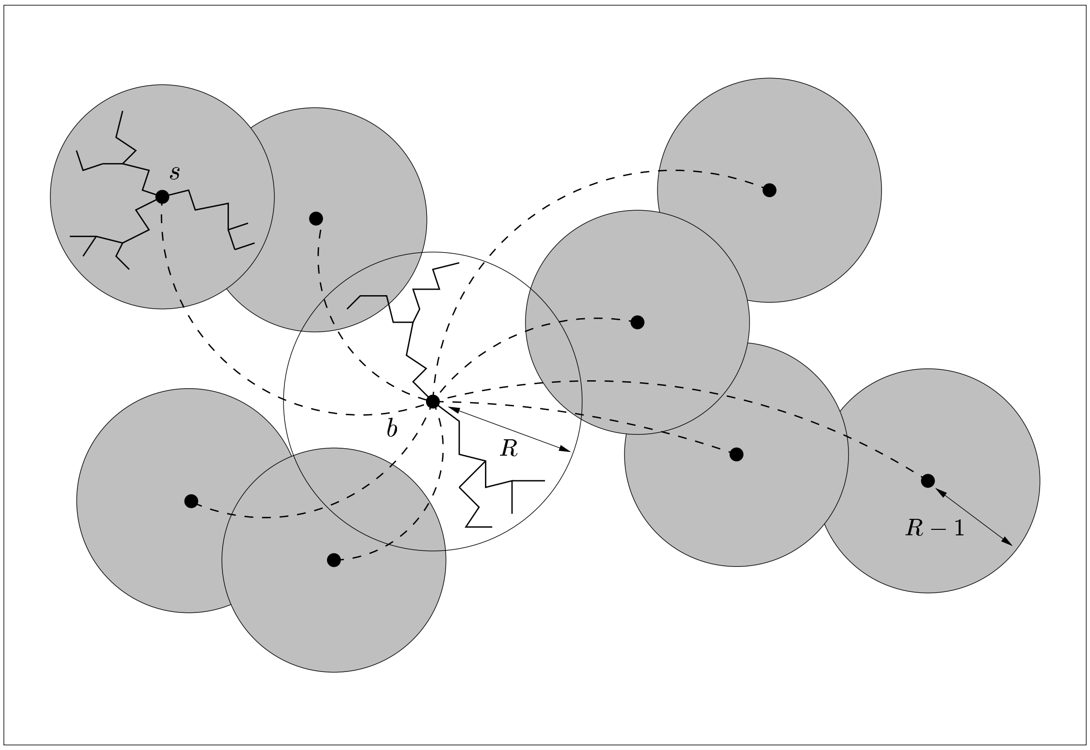

Figure 5:

examples of trees for which the difference between the star–augmentation and the optimal augmentation is larger than one, or to prove that the problem ADC remains NP-hard for trees or forests. We leave this as an *open question* and actually *conjecture* that there exists a constant α not depending of the tree T and of the value of D such that either there is a solution of ADC with at most α edges or the star–solution of the problem ARC is an optimum for ADC (this will result into a polynomial time algorithm for ADC on trees); in the most optimistic version one can suppose that $\alpha \leq 6$.

As to the problem BADC on forests, we can present examples for which the number of edges added by our algorithm is twice the number of edges in an optimal augmentation. For this, consider a forest consisting of D isolated edges. An optimal biconnectivity augmentation has D edges, and the resulting graph is a simple cycle of length $2D$, thus an optimal solution for BADC. On the other hand, our augmentation will use $2D-2$ edges. For the forests occuring in the proof of Proposition 2 the error ratio is $7/4$.

In Appendix A we consider the problem BADC for two particular classes of trees: (i) stars (consisting of a center and k paths of lengths $p_1, \ldots, p_k$) and (ii) trees with a few leaves (trees with n vertices and at most $n/(\mu(2R-1))$ leaves, where $\mu \geq 1$). In both cases, we slightly modify our heuristic in order to get better approximations.

3.2. Odd diameter.

Finally, consider the problem BADC on forests for odd $D = 2R+1$. The simplest approach to design a polynomial approximation algorithm for BADC is to find such an algorithm for the problem ADC, and then to proceed as in the case $D = 2R$. Unfortunately, designing an approximation algorithm for ADC seems to be more difficult than in the case of even diameter. Instead, we address directly the problem BADC and show that a version of the heuristic used for BADC with $D = 2R$ is an approximation algorithm for $D = 2R+1$ as well.

Let T be a tree whose set of leaves is L. Denote by n_{R-1} and n_R the minimum numbers of balls of radius $R-1$ and R, respectively, necessary to cover T. Simple examples of trees show that in general there are no relationships between n_R and n_{R-1}. Nethertheless, the following holds.

Lemma 4. $n_{R-1} \leq \begin{cases} 2n_R + |L|, & \text{if } R \geq 2, \\ 3n_R + |L|, & \text{if } R = 1. \end{cases}$

Moreover, for any covering $\mathcal{C}_R = \{B(s,R) : s \in S\}$ of T with balls of radius R there exists a covering $\mathcal{C}_{R-1} = \{B(s,R-1) : s \in S'\}$ of T with at most $|L \cup S| + |S|$ balls of radius $R-1$ if $R \geq 2$ and with at most $|L \cup S| + 2|S|$ balls of radius $R-1$ if $R = 1$, such that $L \cup S \subseteq S'$.

Proof. It suffices to prove only the second assertion. Define a graph Γ whose vertex set is S and two vertices $s', s'' \in S$ are adjacent in Γ if the path connecting s' and s'' in T does not contain other vertices of S. We assert that Γ is a block-graph, i.e., a graph in which all maximal biconnected subgraphs (alias blocks) are complete subgraphs. First notice that the graph Γ is chordal: existence in Γ of an induced cycle $C = (x_1, x_2, \ldots, x_n, x_1)$ of length $n > 3$ would imply that T has a cycle, because every vertex x_i would lie on the path of T connecting the neighbours of x_i in C. It remains to show that Γ does not contain induced $K_4 - e$ (a clique on four vertices minus one edge). Suppose the contrary, and let all pairs of the quadruple u, v, x, y except the pair xy are edges of Γ. Consider the smallest subtree of T containing these vertices. All four vertices are tips of this subtree, otherwise a non-tip vertex will belong to at least two paths between three other vertices, implying that u, v, x, y do not induce a $K_4 - e$. This also infers that this subtree cannot contain other vertices of S. Therefore our quadruple induces a complete subgraph of Γ, establishing that Γ indeed is a block-graph. With every block-graph one can associate a tree by taking the blocks as vertices and defining an edge between each pair of blocks sharing a common vertex. Therefore, one can speak of pendant blocks of Γ: they correspond to leaves of the associated tree.

Now we will transform the covering $\mathcal{C}_R$ of T into a covering $\mathcal{C}_{R-1}$ consisting of at most $|L \cup S| + |S|$ or $|L \cup S| + 2|S|$ balls of radius $R-1$ obeying $L \cup S \subseteq S'$. Let $m_1 := |L|$, $m_2 := |S|$, and $m_3 := |L \cap S|$. First locate $|L \cup S| = m_1 + m_2 - m_3$ centers at the leaves of T and at the centers from S. We still have to locate m_2 centers if $R \geq 2$ and $2m_2$ centers if $R = 1$. For this we proceed by induction on the number of vertices of T. Pick a pendant block B of Γ, say $B = \{s_1, \ldots, s_k, s_{k+1}\}$, where all vertices of B except possibly s_{k+1} do not belong to other blocks. Suppose without loss of generality that s_1 is the furthest from s_{k+1} vertex of B. On the path between s_1 and s_{k+1} pick the closest to s_{k+1} vertex s'_1 which verifies the condition $d_T(s_1, s'_1) \leq 2R - 1$ (s'_1 can coincide with s_{K+1}). Let x be the neighbour of s'_1 in the path between s'_1 and s_1. Denote by T' and T'' the connected components obtained from T after removing the edge $s'_1 x$, and assume that $s_1 \in T'$. Set $m'_1 := |L \cap T'|$, $m'_2 := |S \cap T'|$ and $m'_3 := |(L \cap S) \cap T'|$. Locate at s'_1 a center of a ball of $\mathcal{C}_{R-1}$.

We assert that the subtree T' is covered by the ball $B(s'_1, R-1)$ and $m'_1 + m'_2 - m'_3$ balls of radius $R-1$ centered at the leaves of T located in T' and at the vertices of S which belong to T'. Suppose by that this is not true, i.e., there exists a vertex $z \in T'$ which is not covered by neither of these balls. In particular, $d_T(z, s'_1) \geq R$. Denote by T_0 the subtree of T (and T') induced by all vertices y such that z lies on the path between y and s'_1. The choice of z implies that $T_0 \neq \{z\}$, thus in T_0 one can pick a leaf z_0 of T. Let $B(s_j, R)$ be the ball of $\mathcal{C}_R$ covering z_0. Since $d_T(z, z_0) \geq R$ by our choice of z, the vertex s_j must be located in the subtree T_0. On the other hand, since $d_T(s_j, z) \geq R$, we conclude that

$$d_T(s_j, s'_1) = d_T(s_j, z) + d_T(z, s'_1) \geq 2R > 2R - 1 = d_T(s_1, s'_1).$$

Therefore $d_T(s_j, s_{k+1}) > d_T(s_1, s_{k+1})$, contrary to the choice of s_1. This establishes our assertion.

Notice also that $B(s_1, R) \subseteq T'$, except possibly the case when $s'_1 = s_{k+1}$ (then obviously $B(s_1, R) \cap T'' \subseteq B(s_{k+1}, R)$). On the other hand, it may happen that for some $s_i \in T'$ the ball $B(s_i, R)$ ($i \neq 1$) intersects the subtree T''. Then, however, the ball $B(s'_1, R-1)$ will contain this intersection (and $m'_2 \geq 2$ holds). Thus, the ball $B(s'_1, R-1)$ covers all vertices of T' not covered by the $(R-1)$-balls centered at the vertices from $(L \cup S) \cap T'$ and, eventually, all vertices of T'' covered by the balls of $\mathcal{C}_R$ whose centers are located in $S \cap T'$.

Now, assume that $R \geq 2$ and apply the induction assumption to the tree T''. Its leaves are the leaves of T which are not in T' plus the vertex s'_1. As a covering of T'' with R-balls we take all balls of $\mathcal{C}_R$ whose centers are in T'', and, additionally, the ball $B(s'_1, R)$, if it is employed to cover some vertices of T'' covered solely by balls of $\mathcal{C}_R$ whose centers are in T' (if this happens, then we already noticed that $m'_2 \geq 2$). By the induction assumption, we can find a required covering of T'' with at most $(m_1 - m'_1 + 1) + 2(m_2 - m'_2) - (m_3 - m'_3)$ $(R-1)$-balls if s'_1 is used only as a leaf and with at most $(m_1 - m'_1 + 1) + 2(m_2 - m'_2 + 1) - (m_3 - m'_3 + 1)$ $(R-1)$-balls if s'_1 is simultaneously a leaf and a center of a ball. The number of balls can be written as

$$(m_1 - m'_1) + 2(m_2 - m'_2) - (m_3 - m'_3) + \mu,$$

where $\mu = 1$ in the first case and $\mu = 2$ in the second case. Recall that in T' we have located other $m'_1 + m'_2 - m'_3$ balls of radius $R-1$. Together with the balls from T'', they constitute a covering of T with at most $m_1 + 2m_2 - m_3 + (\mu - m'_2)$ balls of radius $R-1$. Since in both occuring cases we have $m'_2 \geq \mu$, this is the desired covering.

The case $R = 1$ is similar, but in order to use induction we have to include in S' also the neighbour s''_1 of s'_1 on the path between s'_1 and s_{k+1}. Then we apply all arguments from the previous case to s''_1 instead of s'_1. □

To present an approximation algorithm for the problem BADC on forests, we find a covering with $(R-1)$-balls of the input forest F as in the proof of Lemma 4, pick a leaf b and add all edges between b and all centers of balls in the covering, except the ball centered at b. Let E'' be the set of added edges. Clearly, the complexity of this algorithm is linear.

Proposition 4. *The heuristic is a 6-factor approximation algorithm for the problem BADC with $D = 2R+1$ for $R \geq 2$ and a 8-factor approximation for $R = 1$.*

Proof. Let F be a forest whose set of leaves is L, and let n_R be the minimum number of R-balls necessary to cover F. First, if we add the edges found by our heuristic, then the resulting graph is biconnected and has diameter at most $2R$. Now, let E' be an optimal solution for BADC. Denote by P the set of end-vertices of the edges from E'. Take around each vertex $p \in P$ the ball $B(p, R)$, and let Q be the set of vertices of F not covered by such balls. As in Lemma 3, one can show that the diameter of Q in F is at most D. Therefore Q may be covered by one or two balls of radius R (if we have two balls then their centers are adjacent). Hence $|P| + 2 \geq n_R$. From Lemma 4 we deduce that either $|E''| \leq 2n_R + |L| - 2$

or $|E''| \leq 3n_R + |L| - 2$. Since $|E'| \geq |P|/2$ and $|E'| \geq |L|/2$, all these inequalities imply that $6|E'| \geq |E''| - 2$ for $R \geq 2$ and $8|E'| \geq |E''| - 4$ for $R = 1$, establishing the required conclusion. □

Finally, consider the problem ADC on trees for odd $D = 2R + 1$. For a positive integer k and an edge uv of T we call the set $B(u,k) \cup B(v,k)$ an *edge-ball* centered at uv with radius k. A way to get an admissible augmentation for ADC would be to cover the tree T with a certain number n_1 of edge-balls of radius R and a certain number n_2 of balls of radius $R-1$. Add all edges between the end-vertices of centers of n_1 edge-balls (in the worst case when the edge-centers will be pairwise disjoint, the added edges will create a complete graph K_{2n_1} minus a perfect matching) and then add an edge from the center of each ball of radius $R-1$ to an end-vertex of some center of an edge-ball. In this way we get a graph of diameter D and we inserted at most $n_2 + 2n_1(n_1 - 1)$ new edges. Now, we have to find a covering of T with edge-balls of radius R and balls of radius $R-1$ minimizing the quantity $n_2 + 2n_1(n_1 - 1)$. We call the resulting problem **MIXED COVERING.** Below we establish that an optimal solution for MIXED COVERING provides a 6-factor approximation for the problem ADC. However, we do not know whether MIXED COVERING on trees is polynomially solvable or NP-hard, and we leave this as another *open question.* Obviously, MIXED COVERING can be solved by varying the number n_1 of edge-balls, and for each fixed value of n_1, by locating the edge-balls so that to minimize the number of balls of radius $R-1$ necessary to cover the remaining part of T. This problem is closely related with the following **MIXED CLIQUE COVERING** problem on chordal graphs (note that it is customary to formulate and solve the covering problems of trees by subtrees as clique covering problems of underlying chordal graphs): Given two graphs $G = (V,E)$ and $G_0 = (V,E_0)$ with $E_0 \subseteq E$ (G_0 is a partial subgraph of G) and an integer n_1, locate n_1 cliques of G so that the uncovered vertices can be covered with a minimum number of cliques of the graph G_0. As H. Bodlaender [4] showed, MIXED CLIQUE COVERING is NP-hard if both G and G_0 are chordal graphs. Setting $G := T^{2R+1}$ and $G_0 := T^{2R-2}$, where T^k is the kth power of the tree T, we see that every edge-ball of radius R is a clique of G and every $(R-1)$-ball is a clique of G_0 and vice versa, hence MIXED COVERING on T is reduced to MIXED CLIQUE COVERING for chordal graphs T^{2R+1} and T^{2R-2}.

Proposition 5. *An augmentation provided by an optimal solution of MIXED COVERING is a 6-factor approximation for the problem ADC with $D = 2R + 1$ on trees.*

Proof. Let E' be an optimal solution for ADC and let G' be the augmented graph. Denote by P the set of end-vertices of the edges from E', and let $n_2 := |P|$. As in the proof of Lemma 3, let $Q := V - \cup\{B(p, R-1) : p \in P\}$. Define a graph $T^{2R+1}(Q)$ with Q as a vertex-set and two vertices $u, v \in Q$ are adjacent in $T^{2R+1}(Q)$ if and only if $d_T(u, v) \leq 2R + 1$. Notice that a set $S \subseteq Q$ is a stable set of $T^{2R+1}(Q)$ if and only if $d_T(x, y) > 2R + 1$ for any $x, y \in S$. On the other hand, a clique C of $T^{2R+1}(Q)$ consists of vertices of Q with pairwise distances $\leq 2R + 1$. The least subtree $T(C)$ of T containing the set C has diameter $\leq 2R + 1$, thus its radius is either $\leq R$ or $R + 1$. In the first case $T(C)$ can be covered by a ball of radius R. In the second case, $T(C)$ has two (adjacent) central vertices and $T(C)$ can be covered by two balls of radius R centered at these vertices. Anyway $T(C)$ can be covered with an edge-ball of radius R. As a consequence, a covering of $T^{2R+1}(Q)$ with a minimum number of cliques corresponds to a covering of Q with a minimum number n_1 of edge-balls of radius R. Since the graph $T^{2R+1}(Q)$ is perfect (because it is chordal), the largest stable set Y of $T^{2R+1}(Q)$ has cardinality n_1.

Pick two vertices $u, v \in Y$. Since $d_{G'}(u, v) \leq R$, every shortest path in G' between u and v will use an edge of E'. Such a shortest path consists of a path of T of length R each connecting the vertex u to some vertex $p_u \in P$, the new edge $p_u p_v \in E'$, and the path of T of length R connecting the vertex $p_u \in P$ to v. For a vertex $u \in Y$, let P_u be the collection of all vertices of P which occur as p_u in such a path of G'. Since $d_T(u, v) > 2R + 1$ for distinct vertices $u, v \in Y$, we conclude that $P_u \cap P_v = \emptyset$. Every edge of E' occuring in a shortest (u, v)-path of G' cannot participate in a shortest path of G' between another pair of vertices of Y. Therefore $|E'| \geq n_1(n_1 - 1)/2$. On the other hand, $|E'| \geq n_2/2$. The feasible solution for ADC constructed using an optimal solution for MIXED COVERING contains at most $n_2 + 2n_1(n_1 - 1)$ edges. Hence this number is at most $6|E'|$. □

In Theorem 9.3 from [6] it is shown that DOMINATING SET remains NP-hard for Helly graphs. If in Case 2 of our Proposition 1 as a graph G we will take a Helly graph, then one can easily show that the balls of the resulting graph H still obey the Helly property. Therefore the problems ADC and ARC remain NP-hard on Helly graphs for all $D \geq 3$. We were not able to establish that ADC is NP-hard for even D and any class of graphs where DOMINATING SET or k-DOMINATING SET can be solved in polynomial time. Nethertheless, we can establish this fact for dually chordal graphs and all odd $D \geq 3$.

PROPOSITION 6. *The problem ADC remains NP-hard for dually chordal graphs and all odd values of $D \geq 3$.*

Proof. See Appendix B.

Appendix A: Examples

Consider the problem BADC on some particular instances of trees. To analyze the performance of our heuristics we will use the fact that every biconnected graph of diameter $D = 2R$ with n vertices contains at least $\frac{2R}{2R-1}n - o(1)$ edges. It is conjectured [16] that $o(1)$ is actually 1, but it is established only that $o(1)$ is at most $4(4R-2)^{R-1}$ (see Theorem 2.8 of [5]). If we apply this bound to biconnected graphs of diameter D obtained from a tree T on n vertices, we deduce that every feasible augmentation of T uses at least $n/(2R-1) + 1 - o(1)$ edges.

EXAMPLE 1. (stars) First, suppose that the input tree T is a star with center c and k paths $P_1, \ldots, P_k$ of lengths $p_1, p_2, \ldots, p_k$. Trivially, T has $n := \sum_{i=1}^{k} p_i - k + 1$ vertices. Add $k-1$ edges between the leaf b of the path P_k and all remaining leaves. Next locate at b the center of the ball of radius R and at c a center of a ball of radius $R-1$, and make c adjacent with b. It remains to cover with balls of radius $R-1$ a subpath of length $p_i - 2R$ of each path P_i (a subpath of length $R-1$ of each P_i is covered by the ball centered at b). This can be done with $\lceil \frac{p_i - 2R}{2R-1} \rceil$ balls. Add an edge between b and the center of every such ball, thus getting an admissible solution for BADC. The number of added edges is

$$\sum_{i=1}^{k} \lceil \frac{p_i - 2R}{2R-1} \rceil + k \leq \sum_{i=1}^{k} \frac{p_i - 2R}{2R-1} + 2k.$$

It can be easily seen that this number is smaller or equal than $n/(2R-1) + k$. Comparing this with $n/(2R-1) + 1 - o(1)$, one can conclude that for stars the error is augmenting asymptotically with the number of paths in the star and not with their lengths. Second, for stars with a few branches (say for paths) our heuristic gives an augmentation which (asymptotically) provides not a factor approximation but an approximation within an additive constant to the optimum.

EXAMPLE 2. (Trees with a few leaves). Let $T = (V, E)$ be a tree with n vertices and having at most $n/(\mu(2R-1))$ leaves, where $\mu \geq 1$. Suppose that the diameter of T is larger than $D = 2R$, otherwise we are done. Transform T into a rooted tree whose root is a central vertex b of T. Pick a leaf c as far as possible from b, and let $h := d_T(b,c)$. Add new edges between c and all remaining leaves of T. Let T' be the augmented graph. Next we use an idea from [18] (see also [15]), where it is shown that the tree T has a k-dominating set of cardinality $\le \lfloor n/(k+1) \rfloor$. We will give a better bound for the size of a k-dominating set in the graph T' (we set $k := R-1$).

Divide the vertices of T into levels $T_0, T_1, \ldots, T_h$ according to their height in the tree, assigning all the vertices (except c) of height i to T_i. Clearly $T_0 = \{b\}$. We merge the levels $T_1, \ldots, T_h$ into $2R-1$ sets $D_1, \ldots, D_{2R-1}$ by letting

$$D_i := \bigcup_{j \ge 1} T_{i+j(2R-1)},$$

i.e., T_i and every $(2R-1)$st level thereafter. These sets form a partition of $V - \{b,c\}$, therefore at least one of the sets (say D_i) is of size at most $\lfloor n/(2R-1) \rfloor$. We assert that the set $D_i \cup \{c,b\}$ is a k-dominating set (with $k = R-1$) of the graph T'. Indeed, every vertex $v \in V$ either belongs to a path of T between two vertices of D_i lying in levels $T_{i+j(2R-1)}$ and $T_{i+(j+1)(2R-1)}$, or to a path of T between a vertex of D_i and the vertex b, or to a path of T' between a vertex of D_i and the vertex c. Since every such path has length $\le 2R-1$, we deduce that $\min\{d_{T'}(v,x) : x \in D_i \cup \{b,c\}\} \le R-1$, establishing our assertion. To obtain a feasible solution of BADC for T we augment T' by adding an edge between b and c, and an edge between b and every vertex of D_i. The total number of added edges is at most

$$\frac{n}{\mu(2R-1)} - 1 + \lfloor \frac{n}{2R-1} \rfloor + 1 \le \frac{n}{2R-1}(1 + \frac{1}{\mu}).$$

Comparing this with the lower bound for the number of edges in a biconnected graph of diameter $2R$, we may conclude that asymptotically our heuristic is a $(1+\frac{1}{\mu})$-factor approximation algorithm for the problem BADC.

Appendix B: Proof of Proposition 6

Given a fixed $D = 2R+1$, we will transform the problem CLIQUE to the problem ADC in a dually chordal graph. Let $H = (X,F)$ be an arbitrary instance of the problem CLIQUE, where $X = \{x_1, \ldots, x_n\}$. We construct a dually chordal graph $G = (V,E)$ by taking a copy of H, a vertex b adjacent to all vertices of H, and by adding to each $x_i \in X$ a path $P_i = (x_i, \ldots, y_i)$ of length R; see Figure 6 for an illustration. Obviously, this construction can be accomplished in pseudo-polynomial time and the resulting graph is dually chordal. We

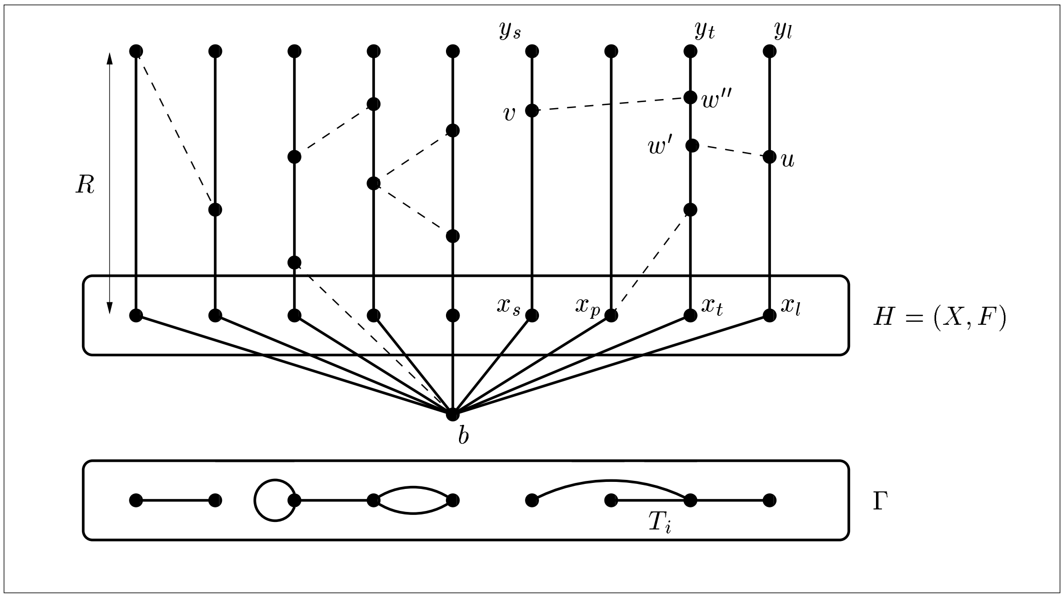

Figure 6:

assert that the graph H has a clique of size φ if and only if there exist a graph $G' = (V, E \cup E')$ of diameter no greater than D and $|E'| \leq n - \varphi$.

First, let Q be a clique of H. To obtain a graph $G' = (V, E \cup E')$ with $|E'| \leq n - |Q|$ and $\mathrm{diam}(G) \leq D$, for each $x_i \notin Q$ we add one edge between the vertex y_i and an arbitrary vertex of Q. Conversely, let $G' = (V, E \cup E')$ be of diameter no greater than D. We assert that H has a clique of size no smaller than $\varphi := n - |E'|$. Define a new graph Γ with $\{1, \ldots, n\}$ as a vertex–set. Two vertices i and j are connected in Γ by s edges if and only if there are exactly s edges in E' with one end in P_i and another end in P_j. Additionally, add a loop at i for each edge $bu \in E'$ with $u \in P_i - \{x_i\}$; see Figure 6.

Consider the connected components of the graph Γ. Since Γ has exactly n vertices and $|E'|$ edges, it must contain at least φ tree–like connected components, i.e., connected components which are trees without loops and multiple edges. Denote them by $T_1, \ldots, T_{\varphi'}$, where $\varphi' \geq \varphi$.

Claim. *For each T_i there is a vertex $j_i \in T_i$ such that $d_{G'}(y_{j_i}, X) = R$ and x_{j_i} is the unique closest to y_{j_i} vertex of X.*

We proceed by induction on the number of vertices of T_i. Pick a leaf $l \in T_i$. Obviously, $d_{G'}(y_l, x_l) = R$. If the vertex y_l obeys the required condition, then set $j_i := l$ and we are done. So, assume that there exists a vertex $x_p \in X$, $p \neq l$, such that $d_{G'}(y_l, x_p) \leq R$. Let t

be the unique neighbour of l in T_i. Denote by uw' the unique edge of E' with one end u in P_l and another one w' in P_t. Clearly,

$$d_{G'}(y_l, x_p) = d_{G'}(y_l, u) + 1 + d_{G'}(w', x_p) \leq R. \tag{1}$$

Deleting l from T_i, we will obtain a smaller tree T_i'. By the induction hypothesis, T_i' contains a vertex s which verifies our assertion in the graph obtained from G' by deleting the edge uw'. We may suppose that $d_{G'}(y_s, x_l) \leq R$, otherwise s is the required vertex of T_i in G'. The shortest path between y_s and x_l will use the edge uw' and some other edge $vw'' \in E'$ with $w'' \in P_t$. In particular, we obtain

$$d_{G'}(y_s, x_l) = d_{G'}(y_s, w'') + d_{G'}(w'', w') + 1 + R - d_{G'}(y_l, u) \leq R. \tag{2}$$

From (2) we conclude that $d_{G'}(y_l, u) \geq d_{G'}(y_s, w'') + d_{G'}(w'', w') + 1$. Together with (1) this imply that

$$d_{G'}(y_s, x_p) < d_{G'}(y_s, w'') + d_{G'}(w'', w') + 2 + d_{G'}(w', x_p) \leq R,$$

contrary to the choice of the vertex $s \in T_i'$. This concludes the proof of the claim.

In each tree–component T_i select a vertex j_i as described in the claim. We assert that the vertices $x_{j_1}, \ldots, x_{j_{\varphi'}}$ induce a complete subgraph in the graph H. To show this, pick two vertices $x_{j_{i'}}$ and $x_{j_{i''}}$. Any shortest path between the vertices $y_{j_{i'}}$ and $y_{j_{i''}}$ will pass via the set X. Since $d_{G'}(y_{j_{i'}}, X) = d_{G'}(y_{j_{i''}}, X) = R$ and $x_{j_{i'}}$ and $x_{j_{i''}}$ are the unique vertices of X closest to $y_{j_{i'}}$ and $y_{j_{i''}}$, respectively, the inequality $d_{G'}(y_{j_{i'}}, y_{j_{i''}}) \leq 2R + 1$ implies that the vertices $x_{j_{i'}}$ and $x_{j_{i''}}$ are adjacent. $\Box$

Acknowledgement. We are grateful to the anonymous referee for several pertinent suggestions.

References

[1] S. Arora and M. Sudan, Improved low degree testing and its applications, in *Proc. 29th ACM Symp. on Theory of Computing,* (1997), pp. 485–495.

[2] M. Bellare, S. Goldwasser, C. Lund, and A. Russell, Efficient multi–prover interactive proofs with applications to approximation problems, in *Proc. 25th ACM Symp. on Theory of Computing,* (1993), pp. 113–131.

[3] C. Berge, *Hypergraphs,* North-Holland, 1989.

[4] H.L. Bodlaender, personal communication.

[5] B. Bollobàs, *Extremal Graph Theory,* Academic Press, 1978.

[6] A. Brandstädt, V. Chepoi, and F. Dragan, The algorithmic use of hypertree structure and maximum neighbourhood orderings, *Discrete Appl. Math.* **82** (1998), 43-77.

[7] R. Chandrasekaran and A. Doughety, Location on tree networks: p-center and q-dispersion problems, *Math. Oper. Res.* **6** (1981), 50-57.

[8] G.J. Chang, Labeling algorithms for domination problems in sun-free chordal graphs, *Discrete Appl. Math.* **22** (1988/1989), 21-34.

[9] K.P. Eswaran and R.E. Tarjan, Augmentation problems, *SIAM J. Computing,* **5**(1976), 653–665.

[10] U. Feige, A threshold of $\ln n$ for approximating set cover, *J. ACM,* **45**(1998), 634–652.

[11] M. R. Garey and D. S. Johnson, *Computers and Intractability: A guide to the theory of NP-completeness,* W.H. Freeman, San Francisco, 1979.

[12] T.S. Hsu and V. Ramachandran, On finding a minimum augmentation to biconnect a graph, *SIAM J. Computing,* **22** (1993), 889–891.

[13] O. Kariv and S.L. Hakimi, An algorithmic approach to network location problems, I: the p-centers, *SIAM J. Appl. Math.* **37(3)** (1979), 513-538.

[14] S. Khuller, Approximation algorithms for finding highly connected subgraphs, in *Approximation Algorithms for NP-hard Problems* (D.S.Hochbaum ed.), PWS Publishing Company, 1997, pp. 236–265.

[15] S. Kutten and D. Peleg, Fast distributed construction of small k-dominating sets and applications, *J. Algorithms,* **28** (1998), 40–66.

[16] A. Laugier, personal communication.

[17] Ch.-L. Li, S.Th. McCormick, and D. Simchi–Levi, On the minimum-cardinality-bounded-diameter and the bounded-cardinality-minimum-diameter edge addition problems, *Operations Research Letters* **11** (1992), 303–308.

[18] D. Peleg and E. Upfal, A tradeoff between size and efficiency of routing tables, *J. ACM* **36** (1989), 510–530.

[19] R. Raz and S. Safra, A sub-constant error-probability low-degree test, and a sub-constant error-probability PCP characterization of NP, in *Proc. 29th ACM Symp. on Theory of Computing,* (1997), pp. 475–484.

[20] A. Rosenthal and A. Goldner, Smallest augmentation to biconnect a graph, *SIAM J. Computing* **6** (1977), 55–66.

[21] A.A. Schoone, H.L. Bodlaender, and J. van Leeuwen, Diameter increase caused by edge deletion, *J. Graph Theory* **11** (1987), 409–427.