Discrete Mathematics and Theoretical Computer Science DMTCS vol. **10:3**, 2008, 135–148

On the k-restricted structure ratio in planar and outerplanar graphs

Gruia Călinescu[1†] and Cristina G. Fernandes[2]

[1] *Department of Computer Science, Illinois Institute of Technology, Chicago, IL 60616, USA*
[2] *Department of Computer Science, University of São Paulo, Rua do Matão, 1010, 05508-900 São Paulo - Brazil*

received March 24, 2008, *revised Oct 18, 2008*, *accepted Oct 24, 2008*.

A planar k-restricted structure is a simple graph whose blocks are planar and each has at most k vertices. Planar k-restricted structures are used by approximation algorithms for Maximum Weight Planar Subgraph, which motivates this work.

The planar k-restricted ratio is the infimum, over simple planar graphs H, of the ratio of the number of edges in a maximum k-restricted structure subgraph of H to the number edges of H. We prove that, as k tends to infinity, the planar k-restricted ratio tends to $1/2$. The same result holds for the weighted version.

Our results are based on analyzing the analogous ratios for outerplanar and weighted outerplanar graphs. Here both ratios tend to 1 as k goes to infinity, and we provide good estimates of the rates of convergence, showing that they differ in the weighted from the unweighted case.

Keywords: planar graphs, outerplanar graphs, approximation algorithms

1 Introduction

We provide insight on approximation algorithms for MAXIMUM WEIGHT PLANAR SUBGRAPH and related problems. MAXIMUM WEIGHT PLANAR SUBGRAPH is the NP-hard problem Liu and Geldmacher (1977); Yannakakis (1978) of finding a heaviest planar subgraph in an edge-weighted simple graph $G = (V, E, w)$, where $w : E \to \mathbb{Q}^+$. That is, the objective is to find a set of edges $F \subseteq E$ such that (V, F) is a planar graph, and which maximizes $w(F) = \sum_{e \in F} w(e)$. In the unweighted version, called MAXIMUM PLANAR SUBGRAPH, the goal is to find a planar subgraph of G with maximum number of edges — that is, all edges of G have weight 1.

A *planar k-restricted structure* is a graph whose blocks (2-connected components) are planar and each has at most k vertices. As we are only interested in planar graphs, from now on we omit the word “planar” — all our k-restricted structures are planar. The best known approximation algorithms for MAXIMUM PLANAR SUBGRAPH Călinescu et al. (1998) and MAXIMUM WEIGHT PLANAR SUBGRAPH Călinescu et al. (2003) are based on 3-restricted structures.

[†]Research supported in part by NSF grant CCF-0515088.

Note that a 2-restricted structure is a forest, and finding a maximum weight forest is a classical problem with polynomial-time algorithms (see, for example, Cormen et al. (2001)). It is known how to find a maximum 3-restricted structure Lovász and Plummer (1986); Gabow and Stallmann (1985); Szigeti (2003) and an ϵ-approximation for a maximum weight 3-restricted structure (implicit in Camerini et al. (1992)). For $k > 3$, several approximation algorithms, originally devised for the Steiner tree problem, produce approximate maximum weight planar k-restricted structures Berman and Ramaiyer (1994); Zelikovsky (1996). When applied to MAXIMUM WEIGHT PLANAR SUBGRAPH, the approximation ratio of these algorithms depends on the planar k-restricted structure ratio ρ_k, defined below. The running time of these algorithms is a polynomial in n^k, where $n = |V|$.

For a simple graph H, denote by $\mathrm{ms}_k(H)$ the maximum number of edges in a subgraph of H whose blocks have size at most k. Define ρ_k as the infimum, over planar simple graphs H, of $\mathrm{ms}_k(H)/|E(H)|$. When H is edge-weighted, denote by $\mathrm{ms}'_k(H)$ the maximum weight of edges in a subgraph of H whose blocks have size at most k. Define ρ'_k as the infimum, over edge-weighted planar graphs H, of $\mathrm{ms}'_k(H)/w(E(H))$. In this paper we prove that $\lim_{k\to\infty} \rho_k = \lim_{k\to\infty} \rho'_k = 1/2$. Similar ratios have been extensively studied for the Steiner tree problem Du et al. (1991); Promel and Steger (2000); Robins and Zelikovsky (2000); Zelikovsky (1993), and are used in the analysis of most approximation algorithms for Steiner tree. The ratios ρ_k and ρ'_k play a similar role with respect to MAXIMUM PLANAR SUBGRAPH.

Our results are based on analyzing β_k and β'_k, defined as ρ_k, but for outerplanar and edge-weighted outerplanar simple graphs. Precisely, we prove that there exist positive constants c_1 and c_2 such that $1 - c_1 \frac{\lg k}{k} \leq \beta_k \leq 1 - c_2 \frac{\lg k}{k}$, and that there exist positive constants c'_1 and c'_2 such that $1 - c'_1 \frac{1}{\lg k} \leq \beta'_k \leq 1 - c'_2 \frac{1}{\lg k}$. Hence $\lim_{k\to\infty} \beta_k = \lim_{k\to\infty} \beta'_k = 1$. We use only base 2 logarithms in this paper. Our proofs are constructive, that is, they consist of explicit examples and deterministic algorithms.

Previous works show that $\rho_2 = \rho'_2 = 1/3$ Dyer et al. (1985), $\rho_3 = 4/9$ Călinescu et al. (1998), $1/3 + 1/72 \leq \rho'_3 \leq 5/12$ Călinescu et al. (2003), that $\beta_2 = \beta'_2 = 1/2$, and $\beta_3 = \beta'_3 = 2/3$ Călinescu et al. (1998, 2003). It is also known that a planar simple graph can have its edge set partitioned into two outerplanar graphs Gonçalves (2005). Thus $\lim_{k\to\infty} \rho_k \geq \frac{1}{2} \lim_{k\to\infty} \beta_k = \frac{1}{2}$ and $\lim_{k\to\infty} \rho'_k \geq \frac{1}{2} \lim_{k\to\infty} \beta'_k = \frac{1}{2}$.

Note that it is trivial that $\rho_k \geq \rho'_k$ and $\beta_k \geq \beta'_k$, and also $\rho_k \geq \rho_{k-1}$, $\rho'_k \geq \rho'_{k-1}$, $\beta_k \geq \beta_{k-1}$, and $\beta'_k \geq \beta'_{k-1}$. Our results indicate that for $\beta_k > \beta'_k$ for large values of k.

As in Călinescu et al. (2003), it is possible to prove that the approximation ratio of Berman and Ramaiyer's algorithm Berman and Ramaiyer (1994) with blocks of size at most k, when applied to MAXIMUM WEIGHT PLANAR SUBGRAPH, is at least

$$\rho'_2 - \frac{\rho'_2 - \rho'_3}{2} - \frac{\rho'_3 - \rho'_4}{3} - \cdots - \frac{\rho'_{k-1} - \rho'_k}{k-1} = \tag{1}$$

$$\rho'_2 \frac{1}{2} + \rho'_3 \left(\frac{1}{2} - \frac{1}{3}\right) + \rho'_4 \left(\frac{1}{3} - \frac{1}{4}\right) + \cdots + \rho'_{k-1} \left(\frac{1}{k-2} - \frac{1}{k-1}\right) + \rho'_k \frac{1}{k-1}. \tag{2}$$

This is a convex combination in ρ'_j, for $2 \leq j \leq k$. Our results do not improve the approximation ratio for MAXIMUM PLANAR SUBGRAPH or MAXIMUM WEIGHT PLANAR SUBGRAPH. They do lead immediately to an alternative proof of the main result of Călinescu et al. (2003): an approximation algorithm with ratio strictly greater than $1/3$, by using a large k in Berman-Ramaiyer's formula of the previous paragraph, the fact that $\rho'_j \geq \rho'_{j-1} \geq \frac{1}{3}$ for all $j \geq 3$, and that $\lim_{k\to\infty} \rho'_k = 1/2$.

Note that the restriction of working on simple graphs is natural. Indeed dealing with multigraphs in the unweighted case is similar to dealing with weighted simple graphs. Therefore, from now on, our graphs are simple unless otherwise specified.

We continue this note as follows. In the next section we give examples for the upper bounds of the limits, and construct the machinery we use for outerplanar graphs. Then we prove the lower bounds for the limits for both weighted and unweighted outerplanar graphs. For the sake of simplicity we do not present here our best c_1 and c_1'; they do not equal our c_2 and c_2'. We leave open finding the exact formulas for β_k and β_k' for all k; such formulas are known for Steiner trees in graphs Borchers and Du (1997).

2 Upper bounds

Theorem 1 *For any positive integer k, $\rho_k \leq 1/2$.*

Proof: Consider the following planar graph G. It has $p+2$ vertices $x, y, v_1, v_2, \ldots, v_p$, and $2p$ edges: $\{x, v_i\}$ and $\{y, v_i\}$ for each $1 \leq i \leq p$. Let S be an arbitrary k-restricted structure of G. Any block in S with t vertices, where $t > 2$, must contain x and y (and $t-2$ vertices v_i). Two blocks cannot have two common vertices, and thus S has at most one block with more than two vertices; this block can have at most $2(k-2)$ edges. Thus the maximum number of edges in a k-restricted structure of G is $2(k-2) + (p - (k-2)) = p + k - 2$. Making p very large compared to k shows that ρ_k cannot exceed $1/2$. $\square$

2.1 Outerplanar graphs

We continue with notions useful when working with outerplanar graphs, notions previously used in Călinescu et al. (2003). Let H be a maximal outerplanar graph (no edge can be added without violating outerplanarity); then it is immediate that the boundary of the exterior face of H is a Hamiltonian cycle C and each interior face is triangular (Harary, 1972, p. 106).

Let E be the edge set of H, b be the exterior face of H, F be the set of faces of H, and $F' = F \setminus \{b\}$. Let D be the dual undirected multigraph of H. Let us call the vertices of D (which are faces of H) *nodes*, and the edges of D, *arcs*. All nodes of D but b have degree three. Also, the edges in the Hamiltonian cycle C correspond to the arcs incident to b in D. Let D' be the graph obtained from D by subdividing each arc incident to b, and then removing b. See Figure 2.1 for an example.

The "dual" construction is also possible: given any undirected tree whose all internal vertices have degree three, one can obtain a maximal outerplanar graph by embedding the tree in the plane, collapsing all the leaves into a single vertex, obtaining a plane multigraph, and taking the planar dual of this multigraph (Bondy and Murty, 1976, Sec. 9.2).

The next lemma appears in Călinescu et al. (2003), but we include its proof for the sake of completeness.

Lemma 2 *D' is a tree all of whose internal nodes have degree three.*

Proof: First, let us prove that D' has no cycle. It is enough to show that any cycle in D contains b. A cycle in D corresponds to a cut in H (Bondy and Murty, 1976, p. 143, ex. 9.2.3). Because C is a Hamiltonian cycle in H, any cut in H contains at least two edges in C, which correspond to arcs incident to b in D. Therefore, any cycle in D contains at least two arcs incident to b, so it contains b.

Second, let us prove that D' is connected. If D' were not connected, there would be two nodes u and v in different components of D'. Let us argue that we can choose u and v to be nodes in $V(D)$. If u were

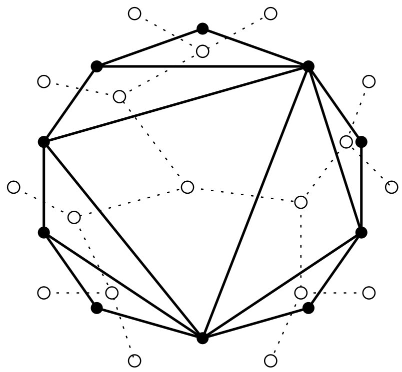

Fig. 1: An outerplanar graph (solid lines) and the tree D' (white vertices and dotted lines) obtained from its dual.

not a node in $V(D)$, then it would be a node that originated from the subdivision of an arc incident to b, and thus it would have degree one in D'. Change u to its unique neighbor in D'. Do the same for v. Note that the new u and v must still be in different components of D', since they are in the same component as the initial u and v, respectively. So we can assume u and v are nodes of D. And because D is connected, there is a path in D between u and v. For this path not to exist in D', it has to go through node b. But this implies that b would be a cut node in D. If b were a cut node in D, then there would be a *minimal* cut in D containing exactly a proper subset of the arcs incident to b (consider the set of arcs going from b to one of the components of D after the removal of b). A minimal cut in D corresponds to a cycle in H (Bondy and Murty, 1976, p. 143, ex. 9.2.3). This implies that there would be a cycle in H whose edges are a proper subset of the edges of C, a contradiction (a proper subset of the edges of any cycle induces an acyclic graph, since it is enough to remove one edge of a cycle to be left with a path, which is acyclic).

Therefore D' is indeed a tree. Recall that all nodes of D but b have degree three. Before removing b, we subdivided all of the arcs incident to b. The new nodes have degree one in D', and all others have the same degree as in D, i.e., three. $\square$

A k-restricted structure K, subgraph of H, is also outerplanar, so its blocks with more than two vertices have each a Hamiltonian cycle. Thus having blocks of size at most k is equivalent, for outerplanar graphs, to having cycles of length at most k.

Let e be an edge of H. We use e to denote also the corresponding arc in D or D'. A cycle C in H (seen as a set of arcs) is a minimal cut in D (Bondy and Murty, 1976, p. 143, ex. 9.2.3). So C disconnects D. See Figure 2.1 for an example. Recall that F' is the set of nodes of D other than b. Since C is minimal, the nodes of D are partitioned into exactly two connected components, given by the set of nodes $F_1^C \ni b$ (outside C) and $F_0^C \subseteq F'$ (inside C), so that C is exactly the set of arcs of D with one endpoint in F_0^C and one in F_1^C. We omit the superscript when C is clear from the context.

Note that F_0 induces a connected subtree of the tree D', and F_0 does not contain leaves of D'. Thus we can relate the number of nodes of F_0 to $|C|$ as follows. Root D' at an arbitrary leaf. If $|F_0| = 1$, then $|C| = 3$, since all internal nodes of D' have degree 3. If $|F_0| > 1$, let u be a node of F_0 as low in the tree as possible; the parent of u also belongs to F_0, while the two children of u do not (u cannot be a leaf). Then $F_0 \setminus \{u\}$ induces a connected component in D' (or D). The number of arcs of D' (or D)

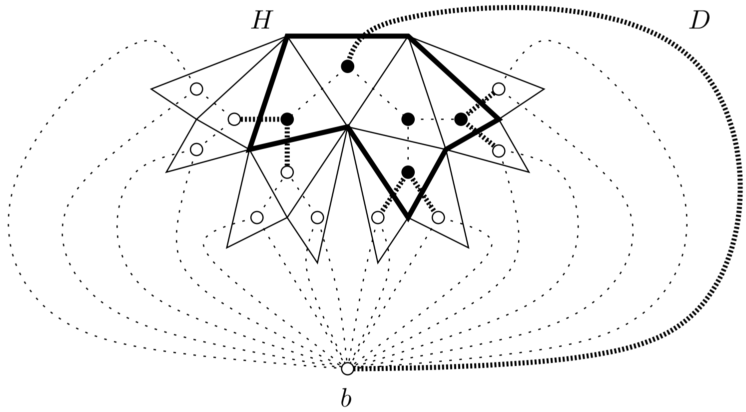

Fig. 2: The dark edges show a cycle C in H and the gray edges show the corresponding cut in D. The black nodes indicate F_0 and the white nodes, F_1.

with exactly one endpoint in F_0 is exactly one more than the number of arcs of D' (or D) with exactly one endpoint in $F_0 \setminus \{u\}$: we have two new arcs, from u to its children, and we miss the arc from u to its parent. Immediate induction gives that $|C| = |F_0| + 2$.

The statement of the next lemma is a bit technical, so we give some intuition before stating it. Let S be a k-restricted structure on an outerplanar graph H. Consider the embedding of S induced by H. If we add to S an edge of H that lies on a non-external face of S, then the resulting graph is still outerplanar. See Figure 2.1 for an illustration.

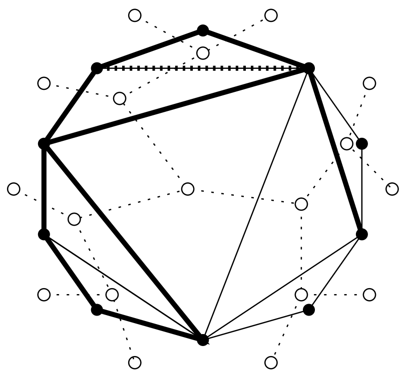

Fig. 3: The dark edges indicate a 4-restricted structure S. The dashed dark edge is inside some non-external face of S, and could be added to S with the result still being a 4-restricted structure.

Lemma 3 *Let $H = (V, E)$ be a maximal outerplanar graph and $K \subseteq E$ be such that (V, K) is a k-restricted structure. Let $Q = E \setminus K$ and $e \in Q$. Consider the duals D and D' described above and* *consider* e *as an arc of* D *and* D'. *If there is no path in* D' *with only arcs of* Q *starting from an endpoint of* e *to a leaf, then* $(V, K \cup \{e\})$ *is a* k*-restricted structure.*

Proof: For a contradiction, suppose that (V, K) is a k-restricted structure, $(V, K \cup \{e\})$ is not a k-restricted structure, and there is no path in D' with only arcs of Q starting from an endpoint of e to a leaf. So there is also no path in D with only arcs of Q from an endpoint of e to b.

Let C be a cycle in $(V, K \cup \{e\})$ of length greater than k. As (V, K) is a k-restricted structure, it has no cycle of length greater than k. So $e \in C$. Let $F_0 = F_0^C$ and $F_1 = F_1^C$ be defined as above. The arcs in D with exactly one endpoint in F_0 correspond exactly to the edges in C, and thus are in K, except for e.

Let $v \in F$ be the endpoint of e in F_1. Let R be the set of nodes reachable in D from v using only arcs of $Q \setminus \{e\}$. By assumption, b is not in R. Also, all arcs in D, except for e, with exactly one endpoint in R are in K, as they are not in Q. Besides, there is at least one such arc f.

Let M be the set of arcs in D with exactly one endpoint in $F_0 \cup R$. From the two paragraphs above, $M \subseteq K$. Note that $F_0 \cup R$ induces a connected subgraph of D'. Indeed, the subgraph of D' induced by F_0 is connected and all vertices in R are connected to v in (the subgraph of D' induced by) R, and v is adjacent through e to a vertex of F_0. Also, $F_0 \cup R$ contains no leaves of D', as $b \notin F_0 \cup R$. Therefore M is a minimal cut in D and thus, according to (Bondy and Murty, 1976, p. 143, ex. 9.2.3), is also a cycle in H.

As D without b is a tree, e is the only edge in D between F_0 and R. So f is not in C, and is in M. Also, M contains all edges in C but e. Therefore the cycle M, which is in (V, K), has more than k arcs, a contradiction, because (V, K) is a k-structure. □

Theorem 4 *There exists a constant* $c_2 > 0$ *such that* $\beta_k \leq 1 - c_2 \frac{\lg k}{k}$ *for all* $k \geq 4$.

Proof: Let us describe a triangulated outerplanar graph H_k, simply denoted H here. We describe the tree D'. Then D can be deduced from D' by collapsing leaves, while H is constructed from D by planar duality.

Consider balanced binary trees T_1, T_2, T_3 (all interior nodes have two children, and all the leaves are on the same level), each with between k and $2k$ nodes, add one more node r and make it adjacent to the roots of T_1, T_2, and T_3. This is the tree D', and when convenient, we assume D' is rooted at r. Also, let $D = (F, E)$ and $H = (V, E)$ be constructed as above. See Figure 2.1 for an illustration. We have that $3k \leq |E| \leq 6k$.

Let Q be a **minimal** set of edges such that $(V, E \setminus Q)$ is a k-restricted structure. It is enough to show that $|Q| \geq 6c_2 \lg k$, for a constant c_2 which we choose later.

We know that $(V, E \setminus Q)$ is a k-restricted structure, and therefore it cannot have cycles longer than k. The edges in Q correspond to arcs in D'.

Assume first that one arc e incident to r is in Q. By the minimality of Q, we have that $(V, (E \setminus Q) \cup \{e\})$ is not a k-restricted structure. Thus, by Lemma 3, Q contains a path from r to a leaf of D', and since T_1, T_2, and T_3 are balanced with at least k nodes, this path must have at least $\lg(k-1)$ nodes.

Consider now the second case, when the three arcs incident to r are not in Q. Let the set R of nodes of F' (recall that these are the interior faces of H) be constructed recursively as follows: $r \in R$, and if a non-leaf node u has a parent in R and all its (three) incident arcs are not in Q, then add u to R. Observe that all arcs that have exactly one endpoint in R are not in Q. Hence one cycle of $(V, E \setminus Q)$ is given by the minimal cut of D which has R on one side. This cycle has $|R| + 2$ edges/arcs. As $(V, E \setminus Q)$ is a

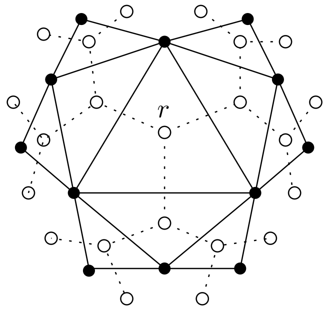

Fig. 4: The example for $k = 4$ in Theorem 4.

k-restricted structure, $|R| < k$. If Q were empty, $R = F'$. But $|F'| > 3\lfloor k/2 \rfloor$ (the interior nodes of T_1, T_2, and T_3), so we get a contradiction. Thus, we may assume Q is non-empty.

As Q is minimal, Lemma 3 assures that there is a path in Q from each arc in Q to a leaf in D'. Let P be such a path with, say, j arcs. Note that r is not in P, as in this second case no arc incident to R is in Q. If a node u is the highest node of F' touched by P, then only the nodes in the balanced subtree rooted at u (including u) are not included in R because of P.

We decompose Q in ℓ such paths $P_1, P_2, \ldots, P_\ell$ for some $\ell \geq 1$. Thus the number of nodes of F' not in R cannot exceed $2^{|P_1|} + 2^{|P_2|} + \cdots + 2^{|P_\ell|}$. This number, given $|Q|$, is maximized when $\ell = 1$ (one single long path), and it must exceed $k/2$ since $|F'| > 3\lfloor k/2 \rfloor$ and $|R| < k$. We conclude that $|Q| \geq \lg(k/2)$, in the second case.

In both cases $|Q| \geq \lg(k/2) \geq \frac{1}{2}\lg k$, so we can pick $c_2 = 1/12$, finishing the proof. □

Theorem 5 *There exists a constant $c_2' > 0$ such that $\beta_k' \leq 1 - c_2' \frac{1}{\lg k}$ for all $k \geq 4$.*

Proof: The example graph $H = H_k$ is a triangulated outerplanar graph. We describe D', and D can be deduced from D' by collapsing leaves, while H is constructed from D by planar duality.

Consider a balanced binary tree T_1 with between $4k+1$ and $8k$ nodes; add one more node r and make it adjacent to the root of T_1. This is D', and we root it at r for convenience. The graphs $D = (F, E)$ (r is also considered to be a leaf here) and $H = (V, E)$ are constructed as above. See Figure 5 for an illustration. Let n be the number of nodes in D'. Note that $n = 2^h$ for some integer h. Also, $4k + 2 \leq n \leq 8k + 1$, so $\lg k < h \leq \lg k + 4 \leq 3 \lg k$. Assign to each arc e of D' a weight as follows: if the higher node of e is at level i (with r at level 0), then $w(e) = 1/2^i$. Then $w(H) = h$.

Let Q be a set of edges such that $(V, E \setminus Q)$ is a k-restricted structure. We proceed to show that $w(Q) \geq 1/8$.

We know that $(V, E \setminus Q)$ is a k-restricted structure, and therefore it cannot have cycles longer than k. Again, the edges in Q correspond to arcs in D'. Let r' be the neighbor of r in D'. If some arc incident to r' is in Q, then this arc has weight 1 or $1/2$, and we are done proving that $w(Q) \geq 1/8$. Assume in the following that Q does not contain any arc incident to r'.

Let the set R of nodes of F' (the interior faces of H) be constructed recursively as follows: $r' \in R$, and if a non-leaf node u has a parent in R and all its (three) incident arcs are not in Q, then add u to R. One cycle of $(V, E \setminus Q)$ is given by the minimal cut of D which has R on one side; this cycle has $|R| + 2$ edges.

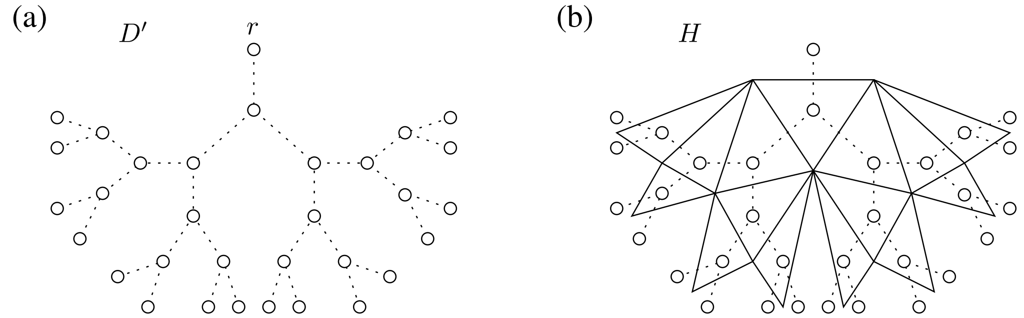

Fig. 5: An example for Theorem 5. (a) A tree D'. (b) The corresponding graph H.

Thus we have that $|R| < k$. If Q was empty, $R = F'$ and $|F'| \geq 2k$, since T_1 has at least $4k + 1$ nodes and thus at least $2k$ interior nodes. But this is a contradiction, so we may assume that Q is non-empty.

An interior node u of D' at level i (with r at level 0) has, including itself, $2^{h-i} - 1$ descendants which are interior nodes of D'. The lightest edge incident to u has weight $1/2^i$. Thus at a cost (weight) of $1/2^i$ one can prevent at most 2^{h-i} vertices from F' to enter R. The cost per vertex removed is at least $1/2^h = 1/n \geq 1/(8k + 1)$. The edges of Q must prevent at least $k + 1$ nodes from F' to enter R (since $|R| < k$ and $|F'| \geq 2k$) and therefore $w(Q) \geq \frac{k+1}{8k+1} \geq \frac{1}{8}$.

Thus we proved that $w(Q) \geq 1/8$ for any set Q of edges such that $(V, E \setminus Q)$ is a k-restricted structure. It follows that a maximum weight k-restricted structure in H has weight at most $h - 1/8$. Thus

$$\beta'_k \leq \frac{(h - 1/8)}{w(E(H))} = \frac{h - 1/8}{h} = \left(1 - \frac{1}{8}\frac{1}{h}\right) \leq \left(1 - \frac{1}{24}\frac{1}{\lg k}\right), \tag{3}$$

as $h \leq 3 \lg k$. Putting $c'_2 = 1/24$ completes the proof. □

3 Lower bounds

Theorem 6 *There exists a constant $c_1 > 0$ such that $\beta_k \geq 1 - c_1 \frac{\lg k}{k}$ for all $k \geq 4$.*

Proof: Let $H = (V, E)$ be an outerplanar graph, and add an edge set E' such that $H' = (V, E \cup E')$ is a maximal outerplanar graph. Construct D and D' for H'; both D and D' have arc set $E \cup E'$.

We give an algorithm to construct a set of edges $Q \subseteq E$ such that $(V, E \setminus Q)$ is a k-restricted structure, and $|Q| \leq c_1(|E|/k) \lg k$, for a constant c_1 to be chosen later. For a tree T and a non-root node u in T, denote by T_u the subtree of T rooted at u plus the parent of u and the arc from u to its parent in T.

To construct Q, the algorithm iteratively builds a sequence of sets of edges $Q_0, Q_1, \ldots, Q_\ell$ and a sequence of rooted trees $D_0, D_1, \ldots, D_\ell$. It starts with $j = 0$, $Q_j = \emptyset$, and $D_j = D'$, rooted at an arbitrary leaf. If D_j has at most k arcs of E, the algorithm outputs $Q = Q_0 \cup Q_1 \cup \cdots \cup Q_j$. Else, increment j and find in D_{j-1} the lowest node u_j such that T_{u_j}, for $T = D_{j-1}$, has more than k arcs of E; such a u_j must exist since the unique child of the root of D_{j-1} is a valid candidate. Construct in T_{u_j} a path P_j as follows: set $u = u_j$ and do the following until u becomes a leaf of D_{j-1}. Find u', a child of u such that $T_{u'}$ has at most as many edges of E as $T_{u''}$, where u'' is the other child of u. Put in P_j the edge

$\{u, u'\}$, make $u = u'$, and resume constructing P_j. Of course P_j may contain edges/arcs from E'. Define $Q_j = P_j \cap E$ and $D_j = D_{j-1} - Z_j$, where Z_j is the set of proper descendants of u_j in D_{j-1} (that is, without u_j itself). Note that u_j is a leaf in D_j. Go back to the checking of the number of arcs of E in D_j. Denote by ℓ the final value of j.

Immediate induction gives that, for $j = 1, \ldots, \ell$, each D_j is a rooted almost-binary tree (the root has only one child, while any other non-leaf node has two children). By the choice of u_j, the subtrees rooted at both children of u_j have at most k edges of E. Therefore P_j contains at most $1 + \lg k$ arcs of E and thus $|Q_j| \leq 1 + \lg k \leq 2\lg k$. If we let E_j be $E \cap E(D_j)$, then $|E_{j+1}| \leq |E_j| - k$, as at most one arc of T_{u_j} remains in E_{j+1}: the arc from u_j to its parent in D'. Hence it is immediate by induction on ℓ that $|E| \geq \ell k$ and $|Q| \leq \sum_{i=1}^{\ell} |Q_i| \leq 2\ell \lg k$. We conclude that indeed $|Q| \leq 2(|E|/k)\lg k$, and we now set $c_1 = 2$.

It remains to prove that $K = (V, E \setminus Q)$ is a k-restricted structure. Define $Q' = \left(\cup_{i=1}^{\ell} P_j\right) \setminus E$. If $\ell = 0$, then $|E| \leq k$, and no cycle of K can have more than k edges. Thus we assume in the following that $\ell > 0$. Note that u_1 is connected to a leaf of D' by the path P_1. For $2 \leq i \leq \ell$, each u_i is connected by the path P_i to either a leaf of D' or a previous $u_{i'}$ ($1 \leq i' < i$). Therefore by immediate induction each u_i, with $1 \leq i \leq \ell$, is connected to a leaf of D' by a path contained in $Q \cup Q'$.

Assume for a contradiction that K is not a k-restricted structure. Let C be a cycle of K with more than k edges. Recall that F_0^C, or simply F_0, is a set of nodes of D and $b \notin F_0$. Consider F_0 as a set of nodes in D' and recall also that F_0 induces a connected subgraph of D'. We claim that $F_0 \cap \{u_1, u_2, \ldots, u_\ell\} \neq \emptyset$. Indeed, next we prove this by induction on ℓ, by looking at the algorithm above, that picked $u_1, u_2, \ldots, u_\ell$.

When $\ell = 1$, we reason as follows: as F_0 induces a connected subgraph of D', we have that either $u_1 \in F_0$, or $F_0 \subseteq A_1$, the set of proper descendants of u_1, or $F_0 \subseteq V(D') \setminus (A_1 \cup \{u_1\})$. If $u_1 \in F_0$, we are done. So we analyze the two other possibilities. Assume that $F_0 \subseteq V(D') \setminus (A_1 \cup \{u_1\})$. Since $\ell = 1$, the whole tree induced by $V(D') \setminus A_1$ has at most k edges from E, and all the edges with exactly one endpoint in F_0 are in this induced tree (as in this case $u_1 \notin F_0$), contradicting $|C| > k$. So it is not possible that $F_0 \subseteq V(D') \setminus (A_1 \cup \{u_1\})$. Assume now that $F_0 \subseteq A_1$. Let u' and u'' be the children of u_1. Then F_0 does not contain u_1 and, as it is connected, F_0 must lay entirely in the subtree rooted at either u' or u''. Say, by symmetry, it is contained in the subtree rooted at u'. By our choice of u_1 (with lowest level in the tree $D' = D_0$), the subtree $T_{u'}$ has at most k edges/arcs of E. Noting that F_0 does not contain leaves of D', we conclude that all the arcs of D' with exactly one endpoint in F_0 must be in $T_{u'}$, and therefore $|C| \leq k$, a contradiction.

Now assume $\ell > 1$. The argument above gives that F_0 cannot be a subset of A_1. Unless $u_1 \in F_0$, in which case we are done, we must have that F_0 is a subset of $V(D_1)$ that does not contain leaves of D_1. Then we can apply the same arguments to u_2 as we did to u_1, with D_1 instead of D_0. This continues for u_3 and so forth until we conclude that, unless $F_0 \cap \{u_1, u_2, \ldots, u_{\ell-1}\} \neq \emptyset$, we must have $u_\ell \in F_0$.

We proved that, assuming $|C| > k$, we have $F_0 \cap \{u_1, u_2, \ldots, u_\ell\} \neq \emptyset$. So let i, with $1 \leq i \leq \ell$, be such that $u_i \in F_0$. We know that u_i is connected by a path P contained in $Q \cup Q'$ to a leaf of D'; this path must have an edge e with exactly one endpoint in F_0, since no leaf of D' is in F_0. By the way F_0 is defined, $e \in C$. However, $e \in Q \cup Q' \subseteq Q \cup E'$, and $C \subseteq E(K) = E \setminus Q$. That is, $e \in E \cap E' = \emptyset$, a contradiction. We conclude that K is indeed a k-restricted structure, completing the proof. □

Theorem 7 *There exists a constant* $c_1' > 0$ *such that* $\beta_k' \geq 1 - c_1' \frac{1}{\lg k}$ *for all* $k \geq 4$.

Proof: The idea of the proof comes from Lemma 25 of Călinescu et al. (2003), where the case $k = 3$ is considered. Let $H = (V, E, w)$ be a weighted outerplanar graph, and add an edge set E' such that $H' = (V, E \cup E')$ is a maximal outerplanar graph. Set $w(e) = 0$ for all e in E'. We have that $w(E \cup E') = w(E)$ and thus without loss of generality we may construct a k-restricted structure in H' instead of H. Construct D and D' for H'; both D and D' have arc set $E \cup E'$.

The proof idea consists of picking a value $j \geq \frac{\lg k}{c_1'}$ (with $c_1' > 0$ to be picked later) and color each arc of D' with one color from the set $\{1, 2, \ldots, j\}$ so that, if E_i is the set of edges/arcs colored i, then the graph $G_i = (V, (E \cup E') \setminus E_i)$ is a k-restricted structure. It then follows that

$$\sum_{i=1}^{j} w(E(G_i)) = j \cdot w(E \cup E') - \sum_{i=1}^{j} w(E_i) = j \cdot w(E \cup E')(1 - 1/j) = j \cdot w(E)(1 - 1/j). \quad (4)$$

It is enough to then pick a G_i with maximum $w(E(G_i))$ to obtain a k-restricted structure of weight at least $(1 - 1/j)w(E) \geq \left(1 - c_1' \frac{1}{\lg k}\right) w(E)$, which is what we want to prove.

Consider $j \geq 3$. Root D' at an arbitrary leaf. The resulting rooted tree is almost binary — every node but the root has two children. For convenience, assume the root has a right child only. Color the arc incident to the root 1. Traverse the tree in preorder, going right first, processing every non-leaf node u as follows. Let v be the parent of u, u' be the left child of u and u'' be the right child of u. Also, assume the arc $\{v, u\}$ has color i, where $1 \leq i \leq j$. Color the arc $\{u, u'\}$ with color i. If u is the right child of v, color the arc $\{u, u''\}$ with color $(i \bmod j) + 1$. (That is, if $i = j$, then color 1, else color $i + 1$.) If u is the left child of v, and the arc from v to its other child is colored i', then use for $\{u, u''\}$ color $(i' \bmod j) + 1$ if $(i' \bmod j) + 1 \neq i$, otherwise use for $\{u, u''\}$ color $(i \bmod j) + 1$. See the example in Figure 3.

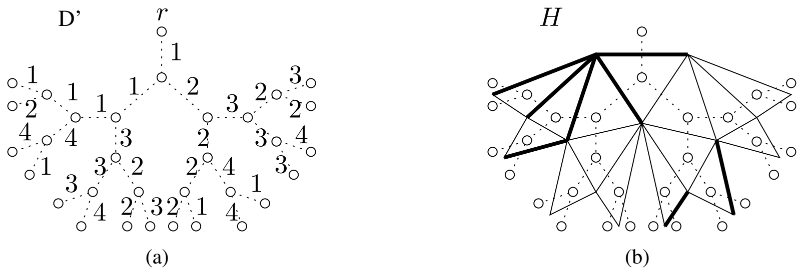

Fig. 6: (a) An example of the coloring of D' with $j = 4$ for the proof in Theorem 7. (b) The darker edges are the edges in G_1.

Now we analyze the size of some arbitrary cycle C in G_i. As before, C is given by the arcs of D' with exactly one endpoint in $F_0 = F_0^C$, where F_0 induces a connected component of D' and does not contain leaves of D'.

We claim that F_0 cannot contain a node adjacent to an arc in E_i. Indeed, if u is a node of F_0 adjacent to an arc in E_i, then it is immediate from the way E_i is constructed that there is a path P from u to a leaf of D' with all of its arcs in E_i: just follow the arcs downwards in D' going left only. Since F_0 does not contain leaves in D', by following P from u downwards we find an arc of D' with only one endpoint in F_0. Thus there is an arc of E_i which is an edge in C; however G_i does not contain edges from E_i. Thus no node of F_0 is adjacent to E_i.

Let u be the highest node of F_0 in D'. Then u is not a leaf of D' and thus is not the root of the rooted D'. All the nodes of F_0 are in the subtree of D' rooted at u. Thus we may assume without loss of generality that $i = j$. Let the arc from u to its parent be colored i', and the arc from u to its right child u'' be colored i''. By construction, $i' \neq i''$. Also, as no node in F_0 is adjacent to an arc in E_j, we have that $i' \neq j$ and $i'' \neq j$.

We claim that the number of nodes of F_0 is at most $2^{j-i''} - 1$ if $i'' > i'$, and at most $2^{j-i''} - 2^{j-i'-1} - 1$ if $i'' < i'$. The proof is by induction on $2j - 3 - i'' - i'$. Note that our coloring ensures that the edge from u to its left child u' is colored i'.

The induction has two base cases: $i'' = j - 1$ and $i' = j - 2$, or $i'' = j - 2$ and $i' = j - 1$. In the first case, both u' and u'', unless they are leaves (in which case they cannot be in F_0), have one incident arc colored j and they cannot be in F_0. Thus $|F_0| = 1$ as desired. In the second case, u'', unless a leaf, has edges incident colored as follows: $j - 2$ to the parent, and $j - 1$ to the right child. This is the first case, and therefore none of the children of u'' can be in F_0. Unless u' is a leaf, the edge going from u' to its right child has color j, and therefore $u' \notin F_0$. We conclude that $F_0 \subseteq \{u, u''\}$ and therefore $|F_0| \leq 2$, as desired.

Consider now the general induction step. We have three case. In the first case, $i' < i''$. Then u', unless a leaf, has the arc going to its right child colored $i'' + 1$. If $i'' = j - 1$, then $u' \notin F_0$; else by induction at most $2^{j-(i''+1)} - 1$ nodes in the subtree rooted at u' can be in F_0. Unless it is a leaf, u'' has arcs incident to it colored i'', to its left child, and $i'' + 1$, to its right child. Thus if $i'' = j - 1$, we have that $u'' \notin F_0$; otherwise, by induction, at most $2^{j-(i''+1)} - 1$ nodes in the subtree rooted at u'' can be in F_0. In all subcases ($i'' = j - 1$ or not), the number of nodes of F_0 in the subtrees rooted at u' or at u'' is at most $2^{j-i''-1} - 1$. Then the number of nodes of F_0 in the subtree rooted at u cannot exceed one (for u) plus the number of nodes of F_0 in the subtrees rooted at u' and u''. This number is $2^{j-i''-1} - 1 + 2^{j-i''-1} - 1 + 1 = 2^{j-i''} - 1$, completing this induction case.

The second case is when $i'' = i' - 1$. Unless it is a leaf, u'' has the arc going to its right child colored $i'' + 1 = i'$, and to its left child colored i''. Induction gives $2^{j-i'} - 1$ as the maximum number of nodes of F_0 in the subtree rooted at u''. Unless it is a leaf, u' has the arc going to its right child colored $i' + 1$ and the arc going to its left child colored i'. Induction gives $2^{j-(i'+1)} - 1$ as the maximum number of nodes of F_0 in the subtree rooted at u'. Then the total number of nodes of F_0 in the subtree rooted at u cannot exceed

$$\begin{aligned} 1 + 2^{j-i'} - 1 + 2^{j-i'-1} - 1 &= 2^{j-i'} + 2^{j-i'} - 2^{j-i'-1} - 1 && (5) \\ &= 2^{j-i'+1} - 2^{j-i'-1} - 1 && (6) \\ &= 2^{j-i''} - 2^{j-i'-1} - 1, && (7) \end{aligned}$$

as desired.

In the third case, $i'' < i' - 1$. Unless it is a leaf, u'' has the arc going to its right child colored $i'' + 1$, and to its left child colored i''. Induction gives $2^{j-(i''+1)} - 1$ as the maximum number of nodes of F_0 in the subtree rooted at u''. Unless it is a leaf, u' has the arc going to its right child colored $i'' + 1$ and the arc going to its left child colored i'. Note that $i'' + 1 < i'$. Induction gives that the subtree rooted at u' has at most $2^{j-(i''+1)} - 2^{j-i'-1} - 1$ nodes of F_0. Combining and adding u, we obtain that the total number of nodes in the subtree rooted at u cannot exceed

$$1 + 2^{j-i''-1} - 1 + 2^{j-i''-1} - 2^{j-i'-1} - 1 = 2^{j-i''} - 2^{j-i'-1} - 1, \quad (8)$$

as desired.

Thus the maximum number of nodes of F_0 is obtained when $i'' = 1$ and $i' = 2$, and is $2^{j-1} - 2^{j-3} - 1$. Hence the cycle C has size $|F_0| + 2 \leq 2^{j-1}$. To ensure that $|C| \leq k$, we pick $j = \lceil \lg k \rceil \geq \lg k$. With this value of j, each G_i, for $i = 1, 2, \ldots, j$, is a k-restricted structure.

We pick $c'_1 = 1$ at this moment. Thus we showed that H has a k-restricted structure G_i with

$$w(E(G_i)) \geq (1 - 1/j)w(E) \geq \left(1 - \frac{1}{\lg k}\right) w(E), \tag{9}$$

finishing the proof. □

The proofs of Theorems 6 and 7 give the more precise bounds of $\beta_k \geq 1 - \frac{\lfloor \lg(k+1) \rfloor}{k}$ and $\beta'_k \geq 1 - \frac{1}{\lfloor 3 - \lg 3 + \lg(k-1) \rfloor}$. We conjecture that $\beta_k = 1 - \gamma \frac{d_k + \lg k}{k}$ and $\beta'_k = 1 - \gamma' \frac{1}{d'_k + \lg k}$, with γ and γ' positive constants, $d_k = o(\lg k)$ and $d'_k = o(\lg k)$. Besides, it maybe the case that $\gamma = \gamma' = 1$ and d_k and d'_k are bounded by a constant. However we do not believe that d_k or d'_k are zero.

References

P. Berman and V. Ramaiyer. Improved approximations for the Steiner tree problem. *JALG*, 17:381–408, 1994.

J. Bondy and U. Murty. *Graph Theory with Applications*. Macmillan/Elsevier, 1976.

A. Borchers and D.-Z. Du. The k-Steiner ratio in graphs. *SIAM Journal on Computing*, 26(3):857–869, 1997.

G. Călinescu, C. Fernandes, U. Finkler, and H. Karloff. A better approximation algorithm for finding planar subgraphs. *Journal of Algorithms*, 27(2):269–302, 1998.

G. Călinescu, C. Fernandes, H. Karloff, and A. Zelikovski. A new approximation algorithm for finding heavy planar subgraphs. *Algorithmica*, 36(2):179–205, 2003.

P. Camerini, G. Galbiati, and F. Maffioli. Random pseudo-polynomial algorithms for exact matroid problems. *JALG*, 13:258–273, 1992.

T. Cormen, C. Leiserson, R. Rivest, and C. Stein. *Introduction to Algorithms*. MIT Press, 2. edition, 2001.

D.-Z. Du, Y.-J. Zhang, and Q. Feng. On better heuristic for Euclidean Steiner minimum trees. In *Proc. 32nd Annual IEEE Symposium on Foundations of Computer Science*, pages 431–439, 1991.

M. Dyer, L. Foulds, and A. Frieze. Analysis of heuristics for finding a maximum weight planar subgraph. *European Journal of Operations Research*, 20:102–114, 1985.

H. Gabow and M. Stallmann. Efficient algorithms for graphic matroid intersection and parity. In *12th Colloq. on Automata, Language and Programming*, pages 210–220, 1985.

D. Gonçalves. Edge partition of planar graphs into two outerplanar graphs. In *37th Annual ACM Symposium on Theory of Computing*, pages 504–512, 2005.

F. Harary. *Graph theory*. Addison-Wesley Publishing Company, 1972.

P. Liu and R. Geldmacher. On the deletion of nonplanar edges of a graph. In *10th Southeastern Conference on Combinatorics, Graph Theory, and Computing*, pages 727–738, 1977.

L. Lovász and M. Plummer. *Matching Theory*. Elsevier, 1986.

H. Promel and A. Steger. A new approximation algorithm for the Steiner tree problem with performance ratio 5/3. *Journal of Algorithms*, 36:89–101, 2000.

G. Robins and A. Zelikovsky. Improved Steiner tree approximation in graphs. In *ACM-SIAM Symposium on Discrete Algorithms*, pages 770–779, 2000.

Z. Szigeti. On the graphic matroid parity problem. *J. Combin. Theory Ser. B*, 88:247–260, 2003.

M. Yannakakis. Node- and edge-deletion NP-complete problems. In *ACM Symposium on Computational Geometry*, pages 253–264, 1978.

A. Zelikovsky. An 11/6-approximation algorithm for the network Steiner problem. *Algorithmica*, 9: 463–470, 1993.

A. Zelikovsky. Better approximation bounds for the network and Euclidean Steiner tree problems. Technical Report CS-96-06, Department of Computer Science, University of Virginia, 1996.