

A $14k$ -Kernel for Planar Feedback Vertex Set via Region Decomposition

Marthe Bonamy¹ and Łukasz Kowalik²(✉)

¹ LIRMM, Montpellier, France

² University of Warsaw, Warsaw, Poland
Kowalik@mimuw.edu.pl

Abstract. We show a kernel of at most $14k$ vertices for the PLANAR FEEDBACK VERTEX SET problem. This improves over the previous kernel of size bounded by $97k$. Our algorithm has a few new reduction rules. However, our main contribution is an application of the region decomposition technique in the analysis of the kernel size.

1 Introduction

A *feedback vertex set* in a graph $G = (V, E)$ is a set of vertices $S \subseteq V$ such that $G - S$ is a forest. In the FEEDBACK VERTEX SET problem, given a graph G and integer k one has to decide whether G has a feedback vertex set of size k . This is one of the fundamental NP-complete problems, in particular it is among the 21 problems considered by Karp [8]. It has applications e.g. in operating systems (see [9]), VLSI design, synchronous systems and artificial intelligence (see [6]).

In this paper we study kernelization algorithms, i.e., polynomial-time algorithms which, for an input instance (G, k) either conclude that G has no feedback vertex set of size k or return an equivalent instance (G', k') , called *kernel*. In this paper, by the size of the kernel we mean the number of vertices of G' . Burrage et al. [5] showed that FEEDBACK VERTEX SET has a kernel of size $O(k^{11})$, which was next improved to $O(k^3)$ by Bodlaender [3] and to $4k^2$ by Thomassé [10].

In this paper we study PLANAR FEEDBACK VERTEX SET problem, i.e., FEEDBACK VERTEX SET restricted to planar graphs. Then one can get a kernel of size $O(k)$ by general tools based on the bidimensionality theory [7]. However, in this general algorithm the constants hidden in the O notation are very large. Bodlaender and Penninkx [4] gave a simple (to implement) algorithm which outputs a kernel of size at most $112k$. This was improved recently by Abu-Khzam and Khuzam [1] to $97k$. In this paper we improve this bound substantially, to $14k$. More precisely, we show the following result.

Theorem 1.1. *There is an algorithm that, given an instance (G, k) of PLANAR FEEDBACK VERTEX SET, either reports that G has no feedback vertex set of*

Work partially supported by the ANR Grant EGOS (2012–2015) 12 JS02 002 01 (MB) and by the National Science Centre of Poland, grant number UMO-2013/09/B/ST6/03136 (ŁK).

size k or produces an equivalent instance with at most $14k - 24$ vertices. The algorithm runs in $O(kn)$ time, where n is the number of vertices of G .

To obtain Theorem 1.1, we extend the algorithms in the previous works [1, 4] by a few new reduction rules. However, our main contribution is an application of the region decomposition technique in the analysis of the kernel size. Region decomposition was developed for the DOMINATING SET problem by Alber et al. [2]. Roughly, in this method the plane instance is decomposed into $O(k)$ regions (i.e. subsets of the plane) such that every region contains $O(1)$ vertices of the graph. We apply this approach in a slightly relaxed way: the regions are the faces of a k -vertex plane graph and the number of vertices in each region is linear in the length of the corresponding face. In [1, 4] kernel size was bounded using different methods, e.g., using bounds on the number of edges in general and bipartite planar graphs. In our opinion region decomposition gives tighter bounds. In particular, we present a tight example, i.e., an example of a family of graphs which can be returned by our algorithm and have $14k - O(1)$ vertices.

Organization of the paper. In Sect. 2 we present a kernelization algorithm which is obtained from the algorithms in [1, 4] by generalizing a few reduction rules, and adding some completely new rules. In Sect. 3 we present an analysis of the size of the kernel obtained by our algorithm. In the analysis we assume that in the reduced graph, for every induced path with ℓ internal vertices, the internal vertices have at least three neighbors outside the path. Based on this, we get the bound of $(2\ell + 4)k - (4\ell + 6)$ for the number of vertices in the kernel. In Sect. 2 we present reduction rules which guarantee that in the kernel $\ell \leq 6$, resulting in the kernel size bound of $16k - 30$. To get the claimed bound of $14k - 24$ vertices we use a complex set of reduction rules, which allow us to conclude that $\ell \leq 5$. Due to space limitations these additional rules are deferred to the journal version.

Notation. In this paper we deal with multigraphs, though for simplicity we refer to them as graphs. (Even if the input graph is simple, our algorithm may introduce multiple edges.) Recall that by the *degree* of a vertex x in a multigraph G , denoted by $\deg_G(x)$, we mean the number of edges incident to x in G . By $N_G(x)$, or shortly $N(x)$, we denote the set of neighbors of x . Note that in a multigraph $|N_G(x)| \leq \deg_G(x)$, but the equality does not need to hold. The neighborhood of a set of vertices S is defined as $N(S) = (\bigcup_{v \in S} N(v)) \setminus S$. For a face f in a plane graph, a *facial walk* of f is the shortest closed walk induced by all edges incident with f . The *length* of f , denoted by $d(f)$ is the length of its facial walk.

2 Our Kernelization Algorithm

In this section we describe our algorithm which outputs a $16k$ -kernel for PLANAR FEEDBACK VERTEX SET. The algorithm exhaustively applies reduction rules. Each reduction rule is a subroutine which finds in polynomial time a certain structure in the graph and replaces it by another structure, so that the resulting

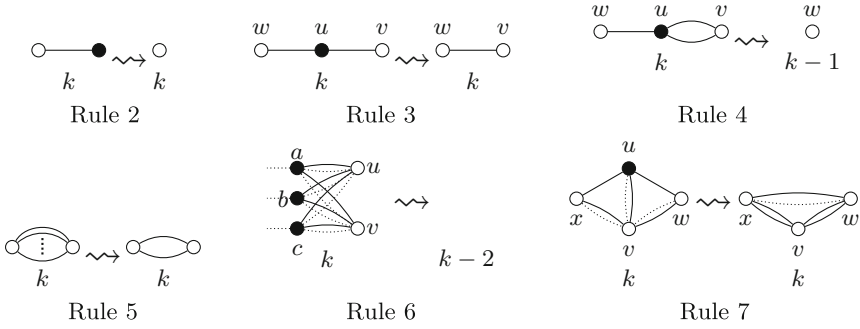


Fig. 1. Reduction rules 1–7. Dashed edges are optional. We draw in black the vertices whose incident edges are all already drawn (as solid or dashed edges), in white the vertices which might be incident to other edges. Regardless of their color, vertices in the figures may not coincide.

instance is equivalent to the original one. More precisely, we say that a reduction rule for parameterized graph problem P is *correct* when for every instance (G, k) of P it returns an instance (G', k') such that:

- (a) (G', k') is an instance of P ,
- (b) (G, k) is a yes-instance of P iff (G', k') is a yes-instance of P , and
- (c) $k' \leq k$.

Below we state the rules we use. The rules are applied in the given order, i.e., in each rule we assume that the earlier rules do not apply. We begin with some rules used in the previous works [1, 4] (Fig. 1).

Rule 1. If there is a loop at a vertex v , remove v and decrease k by one.

Rule 2. Delete vertices of degree at most one.

Rule 3. If a vertex u is of degree two, with incident edges uv and uw , then delete u and add the edge vw . (Note that if $v = w$ then a loop is added.)

Rule 4. If a vertex u has exactly two neighbors v and w , edge uv is double, and edge uw is single, then delete v and u and decrease k by one.

Rule 5. If there are at least three edges between a pair of vertices, remove all but two of the edges.

Rule 6. Assume that there are five vertices a, b, c, v, w such that 1) both v and w are neighbors of each of a, b, c and 2) each vertex $x \in \{a, b, c\}$ is incident with at most one edge xy such that $y \notin \{v, w\}$. Then remove all the five vertices and decrease k by two.

The correctness of the above reduction rules was proven in [1]. (In [1], Rule 6 is formulated in a slightly less general way which forbids multiplicity of some edges, but the correctness proof stays the same.) Now we introduce a few new rules.

Rule 7. If a vertex u has exactly three neighbors v, w and x , v is also adjacent to w and x , and both edges uw and ux are single, then contract uw and add an edge wx (increasing its multiplicity if it already exists). If edge uv was not single, add a loop at v .

Lemma 2.1. *Rule 7 is correct.*

Proof. Let G' be the graph obtained from a graph G by a single application of Rule 7. Let S be a feedback vertex set of size k in G' . We claim S is a feedback vertex set in G too. Assume for a contradiction that there is a cycle C in $G - S$. Then $u \in V(C)$, for otherwise $C \subseteq G'$. If $v \in S$ then $\{wu, ux\} \subseteq C$ and $C - \{wu, ux\} + \{wx\}$ is a cycle in G' , a contradiction. If $v \notin S$, then $w, x \in S$ and hence v is the only neighbor of u in $G - S$, so C is the 2-cycle uvu . But then $G' - S$ contains a loop at v , a contradiction.

Let S be a feedback vertex set of size k in G . If $|\{u, v\} \cap S| = 2$, then $S \setminus \{u\} \cup \{w\}$ is a feedback vertex set of size k in G' . Assume $|\{u, v\} \cap S| = 1$. Then we can assume $v \in S$ for otherwise we replace S by $S \setminus \{u\} \cup \{v\}$, which is also a feedback vertex set in G . If there is a cycle C in $G' - S$, then $wx \in E(C)$, for otherwise $C \subseteq G - S$. But then $C - \{wx\} + \{wu, ux\}$ is a cycle in G , a contradiction. Finally, if $|\{u, v\} \cap S| = 0$ then both w and x are in S , so S is also a feedback vertex set in G' . \square

Rule 8. Let $A \subseteq V(G)$ and let w_1 and w_2 be two vertices in G , $w_1, w_2 \notin A$. If (i) no cycle in $G \setminus \{w_1, w_2\}$ intersects A , and (ii) there is a subgraph $Q \subseteq G[A \cup \{w_1, w_2\}]$ with $|V(Q)| \geq 2$ such that for every vertex $x \in V(Q) \setminus \{w_1\}$, we have $\deg_Q(x) \leq |E(Q)| - |A| - 1$, then remove w_1 and decrease k by 1.

Lemma 2.2. *Rule 8 is correct.*

Proof. Let G' be the graph obtained from a graph G by a single application of Rule 8, i.e., $G' = G - w_1$. Let S be a feedback vertex set of size $k - 1$ in G' . Then every cycle in $G - S$ contains w_1 , so $S \cup \{w_1\}$ is a feedback vertex set of size k in G .

Let S be a feedback vertex set of size k in G . If $w_1 \in S$, then clearly $S \setminus \{w_1\}$ is a solution of the instance $(G', k - 1)$. Hence assume $w_1 \notin S$. We claim that $|S \cap V(Q)| \geq 2$. Assume the contrary, i.e., $|S \cap V(Q)| \leq 1$. Since $Q - S$ is a forest,

$$|E(Q - S)| \leq |V(Q - S)| - 1 = |V(Q)| - |S \cap V(Q)| - 1 \leq |A| + 1 - |S \cap V(Q)|. \quad (1)$$

On the other hand, by the degree bound, and because $w_1 \notin S$ and $|S \cap V(Q)| \leq 1$,

$$|E(Q - S)| \geq |E(Q)| - (|E(Q)| - |A| - 1)|S \cap V(Q)|. \quad (2)$$

By (1) and (2), $|A| + 1 \geq |E(Q)| - (|E(Q)| - |A| - 2)|S \cap V(Q)|$. Since $|S \cap V(Q)| \leq 1$ this implies $|A| + 1 \geq |E(Q)| - (|E(Q)| - |A| - 2) = |A| + 2$, a contradiction. It follows that $|S \cap V(Q)| \geq 2$. Then $S' = S \setminus \{u, v_1, v_2, v\} \cup \{w_1, w_2\}$ is of size at most k . Moreover, S' is a feedback vertex set in G , since S is a feedback vertex set and by (i). Again, this implies that $S' \setminus \{w_1\}$ is a solution of the instance $(G', k - 1)$, as required. \square

Rule 8 is not used directly in our algorithm, because it seems impossible to detect it in $O(n)$ time. However, to get the claimed kernel size we need just a few special cases of Rule 8, which are stated in Lemmas 2.3, 2.4 and 2.5 below (Fig. 2).

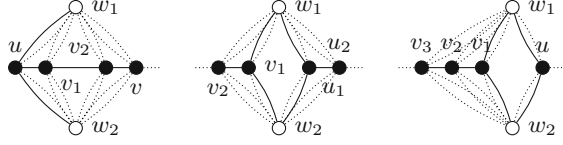


Fig. 2. Configurations in Lemmas 2.3, 2.4, and 2.5.

Lemma 2.3. *Assume there is an induced path $w_1v_1v_2v$ such that for some vertices w_1, w_2 outside the path we have $N(u) = \{v_1, w_1, w_2\}$, $N(\{v_1, v_2\}) \setminus \{u, v\} \subseteq \{w_1, w_2\}$, and there is at most one edge incident to v and a vertex outside $\{w_1, w_2, v_2\}$. Then Rule 8 applies.*

Proof. It is easy to see that condition (i) of Rule 8 is satisfied. We proceed to condition (ii). By symmetry we can assume $|N(w_1) \cap \{v, v_1, v_2\}| \geq |N(w_2) \cap \{v, v_1, v_2\}|$. Let $A = \{u, v_1, v_2, v\}$. We build $E(Q)$ as follows. First, we put the five edges $uw_1, uw_2, w_1v_1, v_1v_2, v_2v$ in Q . Since Rule 3 does not apply, there are no vertices of degree two in G and all of v, v_1 and v_2 are adjacent to w_1 or w_2 (or to both). For every $y \in \{v, v_1, v_2\}$, if $yw_1 \in E$, then we add edge yw_1 to Q , and otherwise we add yw_2 to Q . Thus $|E(Q)| = 8$. Moreover, since $|N(w_1) \cap \{v, v_1, v_2\}| \geq |N(w_2) \cap \{v, v_1, v_2\}|$, in this last step at least two edges added to Q are incident with w_1 , and at most one to w_2 . Hence, for every $x \in V(Q) \setminus \{w_1\}$ we have $\deg_Q(x) \leq 3 = |E(Q)| - |A| - 1$. \square

Lemma 2.4. *Assume there are six vertices $v_1, v_2, u_1, u_2, w_1, w_2$ such that $N(v_1) = \{w_1, w_2, v_2\}$, $N(u_1) = \{w_1, w_2, u_2\}$, there is at most one edge incident to v_2 and a vertex outside $\{w_1, w_2, v_1\}$ and at most one edge incident to u_2 and a vertex outside $\{w_1, w_2, u_1\}$. Moreover, assume that the edges v_1v_2 and u_1u_2 are simple. Then Rule 8 applies.*

Proof. It is easy to see that condition (i) of Rule 8 is satisfied. It is easy to see that condition (ii) of Rule 8 is satisfied. We proceed to condition (ii). By symmetry we can assume $|N(w_1) \cap \{v_2, v_3\}| \geq |N(w_2) \cap \{v_2, v_3\}|$. Let $A = \{u, v_1, v_2, v_3\}$. We build $E(Q)$ as follows. First, we put the six edges $v_1v_2, v_1w_1, v_1w_2, v_2v_3, uw_1$, and uw_2 in Q . Since Rule 3 does not apply, there are no vertices of degree two in G and both v_2 and v_3 are adjacent to w_1 or w_2 (or to both). For every $y \in \{v_2, v_3\}$, if $yw_1 \in E$, then we add edge yw_1 to Q , and otherwise we add yw_2 to Q . Thus $|E(Q)| = 8$. Moreover, since $|N(w_1) \cap \{v_2, v_3\}| \geq |N(w_2) \cap \{v_2, v_3\}|$, in this last step at least one edge added to Q is incident with w_1 , and at most one to w_2 . Hence, for every $x \in V(Q) \setminus \{w_1\}$ we have $\deg_Q(x) \leq 3 = |E(Q)| - |A| - 1$. \square

Lemma 2.5. *Assume there are six vertices $v_1, v_2, v_3, u, w_1, w_2$ such that $N(v_1) = \{w_1, w_2, v_2\}$, $\{v_1, v_3\} \subseteq N(v_2) \subseteq \{w_1, w_2, v_1, v_3\}$, there is at most one edge incident to v_3 and a vertex outside $\{w_1, w_2, v_2\}$ and at most one edge incident to u and a vertex outside $\{w_1, w_2\}$. Moreover, the edges v_1v_2 and v_2v_3 are simple. Then Rule 8 applies.*

Proof. We proceed very similarly as in the proof of Lemma 2.4. \square

Rule 9. Assume there is an induced path with endpoints u and v and with six internal vertices v_1, \dots, v_6 such that for some vertices w_1, w_2 outside the path $N(\{v_1, \dots, v_6\}) \setminus \{u, v\} = \{w_1, w_2\}$. If $|N(w_1) \cap \{v_1, \dots, v_6\}| \geq |N(w_2) \cap \{v_1, \dots, v_6\}|$, then remove w_1 and decrease k by one.

The correctness of Rule 9 is shown in [1]. In [1] it was assumed that when Rule 9 described above is applied, G does not contain an induced path v_1, \dots, v_5 such that for some vertex w , we have $N(v_2, v_3, v_4) \setminus \{v_1, v_5\} = \{w\}$. In our algorithm this is guaranteed by Rule 7 (slightly more general than their Rule 6).

To complete the analysis we need a final rejecting rule which is applied when the resulting graph is too big. In Sect. 3 we prove that Rule 10 is correct.

Rule 10. If the graph has more than $16k - 30$ vertices, return a trivial no-instance (conclude that there is no feedback vertex set of size k in G).

We are able to extend Rule 9 as follows.

Lemma 2.6. *Assume there is an induced path with endpoints u and v and with five internal vertices v_1, \dots, v_5 such that for some vertices w_1, w_2 outside the path $N(\{v_1, \dots, v_5\}) \setminus \{u, v\} = \{w_1, w_2\}$. Then there is an instance (G', k') with $|V(G')| < |V(G)|$ such that (G, k) is a yes-instance iff (G', k') is a yes-instance and $k' \leq k$.*

The proof of Lemma 2.6 contains many cases and is thus deferred to a journal version due to space limitations. We stress here that even without Lemma 2.6, in this paper we give a self-contained kernelization algorithm which returns a kernel of size at most $16k$. If one aims at a $14k$ -kernel, beside adding the reduction rule from Lemma 2.6, the bound in Rule 10 should be replaced by $14k - 26$.

Running time. It is easy to verify that the whole algorithm works in $O(kn)$ time (details deferred to the journal version).

3 The Size Bound

In this section we prove the following theorem.

Theorem 3.1. *Let G be a planar graph such that rules 1–7 do not apply and G does not contain the configurations described in Lemmas 2.3, 2.4 and 2.5. Assume also that for every induced path P with endpoints u and v and with ℓ internal vertices v_1, \dots, v_ℓ the internal vertices have at least three neighbors outside the path, i.e., $|N(\{v_1, \dots, v_\ell\}) \setminus \{u, v\}| \geq 3$. If there is a feedback vertex set of size k in G , then $|V(G)| \leq (2\ell + 4)k - (4\ell + 6)$.*

Let S be a feedback vertex set of size k in G (i.e., a “solution”), and let F be the forest induced by $V(G) \setminus S$. Denote the set of vertices of F by $V_F = V(G) \setminus S$. We call the vertices in S *solution vertices* and the vertices in V_F *forest vertices*.

A partition of V_F . Now we define some subsets of V_F . Let $I_2, I_{3+} \subseteq V_F$ denote the vertices whose degree in F is two or at least three, respectively. The leaves of F are further partitioned into two subsets. Let L_2 and L_{3+} be the leaves of F that have two or at least three solution neighbors, respectively. By rules 2 and 3 all the vertices in G have degree at least 3. Hence, if a leaf of F has fewer than two solution neighbors, Rule 4 or 5 applies. It follows that every leaf of F belongs to $L_2 \cup L_{3+}$. This proves claim (i) of Lemma 3.2 below.

Lemma 3.2. *Graph G satisfies the following properties.*

- (i) *The sets I_2, I_{3+}, L_2, L_{3+} form a partition of V_F .*
- (ii) *For every pair u, v of solution vertices there are at most two vertices $x, y \in L_2$ such that $N(x) \cap S = N(y) \cap S = \{u, v\}$.*
- (iii) *Every vertex of G is of degree at least three.*
- (iv) *Every face of G is of length at least two.*

Claim (ii) follows from the fact that Rule 6 does not apply to G . Claim (iii) follows because rules 2 and 3 do not apply to G and Claim (iv) by Rule 1.

The inner forest. Let F_I be the forest on the vertex set $I_{3+} \cup L_{3+}$ such that $uv \in E(F_I)$ iff for some integer $i \geq 0$, there is a path $ux_1 \cdots x_i v$ in forest F such that $u, v \in I_{3+} \cup L_{3+}$ and for every $j = 1, \dots, i$, vertex x_j belongs to I_2 .

Three sets of short chains. A path in F consisting of vertices from $I_2 \cup L_2$ will be called a *chain*. A chain is maximal if it is not contained in a bigger chain. In what follows we introduce three sets of (not necessarily maximal) chains, denoted by CL_2, C_{2-} and C_{3+} . We will do it so that each vertex in I_2 belongs to *at least one* chain from these sets of chains.

For every vertex $x \in L_2$, we consider the maximal chain (y_1, \dots, y_p) of degree 2 vertices in F such that y_1 is adjacent to x and no y_i has a solution neighbor outside $N_G(x) \cap S$. Then the chain (x, y_1, \dots, y_p) is an element of CL_2 . Note that $L_2 \subseteq V(CL_2)$.

Chains of C_{2-} and C_{3+} are defined using the following algorithm. We consider maximal chains in F , one by one (note that all maximal chains are vertex-disjoint). Let $c = (x_1, x_2, \dots, x_p)$ be a maximal chain. The vertices of c are ordered so that if $\{x_1, x_p\} \cap L_2 \neq \emptyset$, then $x_p \in L_2$. Using vertices of c we form disjoint bounded length chains and put them in the sets C_{2-} and C_{3+} as follows. Assume that for some $i < p$ the vertices of a prefix (x_1, x_2, \dots, x_i) have been already partitioned into such chains (in particular $i = 0$ if we begin to process c). There are three cases to consider.

Consider a shortest chain $c_i = (x_{i+1}, \dots, x_j)$ such that the vertices of c_i have at least three solution neighbors, i.e., $|S \cap N(\{x_{i+1}, \dots, x_j\})| \geq 3$. If the chain c_i exists, we put it in C_{3+} , and we proceed to the next vertices of c . Otherwise we consider the chain $c'_i = (x_{i+1}, \dots, x_p)$. Note that vertices of c'_i have at most two solution neighbors.

If $x_p \in I_2$, then we add the chain c'_i to C_{2-} and we finish processing c . Note that then x_p is adjacent to a vertex $u \in L_{3+} \cup I_{3+}$ (otherwise c is not maximal, as we can extend it by a vertex in L_2). Moreover, because of the order of the

vertices in c , we know that $x_1 \notin L_2$. It follows that x_1 is also adjacent to a vertex $v \in L_{3+} \cup I_{3+}$. Hence, $uv \in E(F_I)$. We assign chain c'_i to edge uv .

If $x_p \in L_2$, then we do not form a new chain and we finish processing c . Note, however, that the vertices $\{x_{i+1}, \dots, x_p\} \cap I_2$ belong to a chain in CL_2 .

Note also that some vertices of the first chain c_0 can belong to two chains, one in C_{3+} and one in CL_2 .

Let us summarize the main properties of the construction.

Lemma 3.3. *The following properties hold:*

- (i) *Every vertex from I_2 belongs to a chain in CL_2, C_{2-} or C_{3+} .*
- (ii) *Every chain in $CL_2 \cup C_{2-}$ has at most two solution neighbors.*
- (iii) *Every chain in C_{3+} has at least three solution neighbors.*
- (iv) *Every chain in C_{2-} is assigned to a different edge of inner forest F_I .*
- (v) *Every chain in $C_{2-} \cup CL_2$ has at most $\ell - 1$ vertices.*
- (vi) *Every chain in C_{3+} has at most ℓ vertices.*

A solution graph H_S . Let us introduce a new plane multigraph $H_S = (S, E_S)$. Since the vertices of H_S are the solution vertices we call it a *solution graph*. From now on, we fix a plane embedding of G . The vertices of H_S are embedded in the plane exactly in the same points as in G . The edge multiset E_S is defined as follows. For every triple (u, x, v) such that $u, v \in S, x \in L_2$ and there is a path uxv in G , we put an edge uv in E_S . Moreover, the edge uv is embedded in the plane exactly as one of the corresponding paths uxv (note that there can be up to four such paths if some edges are double). Note that by Lemma 3.2(ii), every edge of H_S has multiplicity at most two.

The set of faces of H_S is denoted by F_S . By $F_{S,2}$ we denote its subset with the faces of length two, while $F_{S,3+}$ are the remaining faces. Note that there are no faces of length 1 in H_S .

Lemma 3.4. *We have $|V(CL_2)| \leq 3(|E_S| - |F_{S,2}|)$.*

Proof. By the definition, for every vertex $x \in L_2$ there is a corresponding edge $uv \in E_S$, where $N_G(x) \cap S = \{u, v\}$. Also, for every chain c in CL_2 there is a corresponding vertex $x \in L_2$, and thus a corresponding edge $uv \in E_S$. We assign x, c and the vertices of c to the pair $\{u, v\}$.

Consider an arbitrary pair u, v such that $uv \in E_S$. Note that there are exactly $|E_S| - |F_{S,2}|$ such pairs. We claim that there are at most three elements in $V(CL_2)$ assigned to the pair $\{u, v\}$. Indeed, by Lemma 3.2(ii), there are at most two vertices in L_2 assigned to $\{u, v\}$. If there are no such vertices, no chain in CL_2 is assigned to $\{u, v\}$, so the claim holds. If there is exactly one vertex $x \in L_2$ assigned, there is exactly one chain $c \in CL_2$ assigned. By Lemma 2.3, chain c has at most three vertices, so the claim holds. Finally, if there are exactly two vertices $x, y \in L_2$ assigned, there are exactly two chains c_x and c_y assigned. By Lemmas 2.4 and 2.5 we have $|V(c_x)| + |V(c_y)| \leq 3$. This concludes the proof. □

Maximality. In what follows we assume that graph G is *maximal*, meaning that one can add neither an edge to $E(G)$ nor a vertex to L_2 obtaining a graph G' such that S is still a feedback vertex set of G' and all the claims of Lemmas 3.2, 3.3 and 3.4 hold. Note that the number of L_2 -vertices which can be added to G is bounded, since each such vertex corresponds to an edge in H_S , and H_S has at most $6|S|$ edges as a plane multigraph with edge multiplicity at most two. Similarly, once the set of L_2 -vertices is maximal, and hence the vertex set of G is fixed, the number of edges which can be added to G is bounded by $6|V(G)|$. It follows that such a maximal supergraph of G exists. Clearly, it is sufficient to prove Theorem 3.1 only in the case when G is maximal.

Lemma 3.5. *The planar graph H_S is connected.*

Proof. Assume now for contradiction that there is a partition $S = S_1 \cup S_2$ such that there is no edge in H_S between a vertex of S_1 and a vertex of S_2 .

Every face of G is incident to at least one vertex of S , for otherwise the boundary of the face does not contain a cycle, a contradiction. Assume that a face f of G contains a solution vertex u_1 in S_1 and a solution vertex u_2 in S_2 . Then we can add a vertex x , two edges xu_1 and two edges xu_2 . Note that S is still a feedback vertex set in the new graph; in particular now $x \in L_2$. In the new graph there are no more vertices in L_2 adjacent to both u_1 and u_2 because of our assumption that S_1 and S_2 are not connected by an edge in H_S , so Lemma 3.2(ii) holds. Moreover, $|V(CL_2)|$ was increased by one and $|E_S| - |F_{S,2}|$ was also increased by one, so Lemma 3.4 holds. The other claims of Lemmas 3.2 and 3.3 trivially hold, so F is not maximal, a contradiction.

Let \mathcal{F}_1 and \mathcal{F}_2 be the collections of faces of G containing a vertex in S_1 , or in S_2 , respectively. We have shown above that $\mathcal{F}_1 \cup \mathcal{F}_2$ is a partition of the set of all the faces of G . Let V_1 and V_2 denote the sets of vertices incident to a face in \mathcal{F}_1 , or in \mathcal{F}_2 , respectively. Note that $V_1 \cap V_2 \neq \emptyset$, since there must be two neighboring faces, one in \mathcal{F}_1 and the other in \mathcal{F}_2 . Let $x \in V_1 \cap V_2$. Since faces of G are of length at least two, x has in G at least two neighbors in $V_1 \cap V_2$. It follows that $G[V_1 \cap V_2]$ has minimum degree two, so $G[V_1 \cap V_2]$ contains a cycle. However, $(V_1 \cap V_2) \cap S = \emptyset$, since \mathcal{F}_1 and \mathcal{F}_2 are disjoint. Hence $V_1 \cap V_2 \subseteq F$, a contradiction. \square

Bounding the number of forest vertices in a face of H_S . For a face f of H_S and a set of vertices $A \subseteq V(G)$ we define A^f as the subset of A of vertices which are embedded in f or belong to the boundary of f . Note that all vertices of every chain belong to the same face f of H_S . When C is a set of chains, by C^f we denote the subset of chains of C which lie in f , i.e., $C^f = \{c \in C : V(c) \subseteq V(G)^f\}$.

Lemma 3.6. *For every face f of H_S , it holds that $|L_{3+}^f| + |I_{3+}^f| + |C_{3+}^f| \leq d(f) - 2$.*

Proof. First we note that the forest F^f is in fact a tree. Indeed, if F^f has more than one component, we can add an edge between two solution vertices on the boundary of f preserving planarity, what contradicts the assumed maximality.

Consider a plane subgraph A of G induced by $V(G)^f$, i.e., we take the plane embedding of G and we remove the vertices outside $V(G)^f$. Then we can define graph A_S , analogously to H_S . We treat f as a face of A_S . Let $u_1 u_2 \cdots u_{d(f)} u_1$ be the facial walk of f .

Consider an arbitrary vertex x of I_{3+}^f . Let T_1, \dots, T_r be the r trees obtained from the tree T in F containing x after removing r from T . Then $r \geq 3$ since x has at least three neighbors in T . By planarity, there are $2r$ indices $b_1, e_1, b_2, e_2, \dots, b_r, e_r$ such that for every $i = 1, \dots, r$

$$\{u_{b_i}, u_{e_i}\} \subseteq N(V(T_i)) \cap \{u_1, \dots, u_{d(f)}\} \subseteq \{u_{b_i}, u_{b_i+1}, \dots, u_{e_i}\}.$$

Then, for every $j \in \{b_1, b_2, \dots, b_r\}$ there is an edge xu_j , for otherwise we can add it in the current plane embedding, contradicting the maximality of G . This means that every vertex in I_{3+}^f has at least three neighbors in $\{u_1, u_2, \dots, u_{d(f)}\}$.

We further define B as the plane graph obtained from A by (1) replacing every triple (u, x, v) where $x \in L_2$, $u, v \in S$ and uxv forms a path by a single edge uv , (2) removing vertices of $V(CL_2)$, (3) contracting every chain from C_{3+} into a single vertex, and (4) contracting every chain from C_{2-} into a single edge. By (4) we mean that every maximal chain $d = x_1, \dots, x_i$ of I_2 vertices which is contained in a chain from C_{2-} , is replaced by the edge yz where y and z are the forest neighbors (in $L_{3+} \cup I_{3+}$) of x_1 and x_i outside the chain d . Let us call the vertices of B that are not on the boundary of f as *inner vertices*.

Note that the set of inner vertices is in a bijection with $L_{3+}^f \cup I_{3+}^f \cup C_{3+}^f$. Moreover, I forms a tree, since F^f is a tree. Also, each inner vertex has at least three neighbors in $\{u_1, u_2, \dots, u_{d(f)}\}$. We show that $|I| \leq d(f) - 2$ by the induction on $d(f)$. When $d(f) = 2$ the claim follows since each inner vertex has at least three neighbors on the boundary of f . Now assume $d(f) > 2$. Let x be leaf in the tree I . Then the edges from x to the boundary of face f split F into at least three different faces. The subtree $I - x$ lies in one of these faces, say face bounded by the cycle $xu_i u_{i+1} \cdots u_j x$. We remove x and vertices u_{j+1}, \dots, u_{i-1} (there is at least one of them) and we add edge $u_i u_j$. The outer face of the resulting graph is of length at most $d(f) - 1$, so we can apply induction and the claim follows. \square

Lemma 3.7. *For every face f in H_S of length at least three,*

$$|V_F^f \setminus V(CL_2^f)| \leq \ell \cdot (d(f) - 2) - (\ell - 1).$$

Proof. We have

$$|V_F^f \setminus V(CL_2^f)| \leq |L_{3+}^f| + |I_{3+}^f| + |V(C_{3+}^f)| + |V(C_{2-}^f)|.$$

By Lemma 3.3(v) we get

$$|V_F^f \setminus V(CL_2^f)| \leq |L_{3+}^f| + |I_{3+}^f| + \ell |C_{3+}^f| + (\ell - 1) |C_{2-}^f|. \tag{3}$$

By Lemma 3.3(iv), $|C_{2-}^f|$ is bounded by the number of edges of the inner forest F_I . Hence, $|C_{2-}^f| \leq |L_{3+}^f| + |I_{3+}^f| - 1$ when $|L_{3+}^f| + |I_{3+}^f| > 0$ and $|C_{2-}^f| = 0$

otherwise. In the prior case, by (3) we get that

$$|V_F^f \setminus V(CL_2^f)| \leq \ell(|L_{3+}^f| + |I_{3+}^f| + |C_{3+}^f|) - (\ell - 1),$$

and the result then follows from Lemma 3.6. Hence it suffices to prove the claim when $|L_{3+}^f| = |I_{3+}^f| = |C_{2-}^f| = 0$. Then the forest F^f is a non-empty collection of paths, each with both endpoints in L_2 . Let c be such a path on p vertices x_1, \dots, x_p . Then $x_1 \in L_2$ and x_1 has exactly two neighbors u, v in S . Let i be the largest such that $N(\{x_1, \dots, x_i\}) \cap S = \{u, v\}$. By definition, (x_1, \dots, x_i) is a chain in CL_2^f . We infer that if $i = p$ for every such path, then $|V_F^f \setminus V(CL_2^f)| = 0$ and the claim follows. Hence we can assume that $i < p$, i.e., x_{i+1} has a neighbor in $S \setminus \{u, v\}$. Then, by definition, (x_1, \dots, x_{i+1}) is a chain in C_{3+}^f . Since $(x_1, \dots, x_i) \in CL_2^f$, we get $|\{x_1, \dots, x_{i+1}\} \setminus V(CL_2^f)| = 1$. Hence,

$$|V_F^f \setminus V(CL_2^f)| \leq 1 + \ell(|C_{3+}^f| - 1),$$

what, by Lemma 3.6, is bounded by $1 + \ell \cdot (d(f) - 3) = \ell \cdot (d(f) - 2) - (\ell - 1)$, as required. \square

Lemma 3.8. *For every face f in H_S of length two, $V_F^f \subseteq V(CL_2^f)$.*

Proof. Since the boundary of f has only two solution vertices, F^f contains no vertices of L_{3+}^f , $V(C_{3+}^f)$ or I_{3+}^f . Then by Lemma 3.3(iv), C_{2-}^f is also empty. The claim follows. \square

Now we proceed to the bound of Theorem 3.1. By Lemmas 3.7 and 3.8 we have

$$|V_F| \leq |V(CL_2)| + \sum_{f \in F_{S,3+}} (\ell(d(f) - 2) - (\ell - 1))$$

By Lemma 3.4 we get

$$\begin{aligned} |V_F| &\leq 3(|E_S| - |F_{S,2}|) + \sum_{f \in F_{S,3+}} (\ell(d(f) - 2) - (\ell - 1)) \\ &= 3(|E_S| - |F_{S,2}|) + \sum_{f \in F_S} (\ell(d(f) - 2) - (\ell - 1)) + (\ell - 1)|F_{2,S}| \\ &= (2\ell + 3)|E_S| - (3\ell - 1)|F_S| + (\ell - 4)|F_{2,S}| \\ &= (2\ell + 3)|E_S| - (2\ell + 3)|F_S| - (\ell - 4)|F_S| + (\ell - 4)|F_{2,S}| \\ &\leq (2\ell + 3)(|E_S| - |F_S|). \end{aligned}$$

By Lemma 3.5 graph H_S is connected, so we can apply Euler's formula $|S| - |E_S| + |F_S| = 2$. Thus,

$$\begin{aligned} |V(G)| &= |V_F| + |S| \leq (2\ell + 3)(|S| - 2) + |S|, \\ &= (2\ell + 4)k - (4\ell + 6). \end{aligned}$$

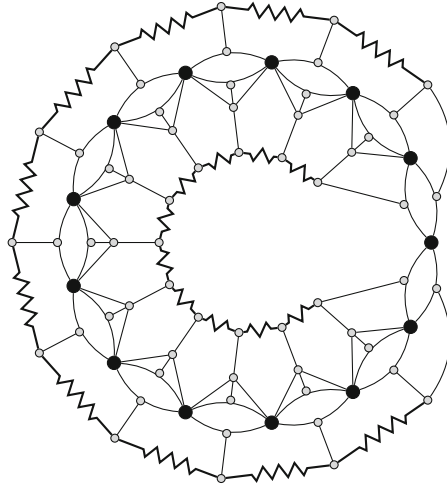


Fig. 3. A tight example. The big black vertices are solution vertices, the small grey ones are forest vertices. The zigzag edges represent paths of $\ell - 1$ forest vertices, each adjacent to the two available solution vertices. Asymptotically for larger cycles, we have $2\ell + 3$ forest vertices for each solution vertex.

This concludes the proof of Theorem 3.1. For $\ell = 6$, we get $|V(G)| \leq 16k - 30$. If we use Lemma 2.6, we can put $\ell = 5$, which results in $|V(G)| \leq 14k - 24$. In Fig. 3 we show an example of a graph, where our reduction rules do not apply and our analysis is tight (up to a constant additive term).

Note. We have learned that very recently Xiao [11] obtained independently a $29k$ -kernel for PLANAR FEEDBACK VERTEX SET.

References

1. Abu-Khzam, F.N., Bou Khuzam, M.: An improved kernel for the undirected planar feedback vertex set problem. In: Thilikos, D.M., Woeginger, G.J. (eds.) IPEC 2012. LNCS, vol. 7535, pp. 264–273. Springer, Heidelberg (2012)
2. Alber, J., Fellows, M.R., Niedermeier, R.: Polynomial-time data reduction for dominating set. *J. ACM* **51**(3), 363–384 (2004)
3. Bodlaender, H.L.: A cubic kernel for feedback vertex set. In: Thomas, W., Weil, P. (eds.) STACS 2007. LNCS, vol. 4393, pp. 320–331. Springer, Heidelberg (2007)
4. Bodlaender, H.L., Penninkx, E.: A linear kernel for planar feedback vertex set. In: Grohe, M., Niedermeier, R. (eds.) IWPEC 2008. LNCS, vol. 5018, pp. 160–171. Springer, Heidelberg (2008)
5. Burrage, K., Estivill-Castro, V., Fellows, M.R., Langston, M.A., Mac, S., Rosamond, F.A.: The undirected feedback vertex set problem has a $\text{poly}(k)$ kernel. In: Bodlaender, H.L., Langston, M.A. (eds.) IWPEC 2006. LNCS, vol. 4169, pp. 192–202. Springer, Heidelberg (2006)

6. Festa, P., Pardalos, P., Resende, M.: Feedback Set Problems. *Encyclopedia of Optimization*, pp. 1005–1016. Springer, New York (2009)
7. Fomin, F.V., Lokshtanov, D., Saurabh, S., Thilikos, D.M.: Bidimensionality and kernels. In: *SODA*, pp. 503–510 (2010)
8. Karp, R.: Reducibility among combinatorial problems. In: Miller, R., Thatcher, J. (eds.) *Complexity of Computer Computations*, pp. 85–103. Plenum Press, New York (1972)
9. Silberschatz, A., Galvin, P.B., Gagne, G.: *Operating System Concepts*, 8th edn. Wiley Publishing, New York (2008)
10. Thomassé, S.: A $4k^2$ kernel for feedback vertex set. *ACM Trans. Algorithms* **6**(2), 1–8 (2010)
11. Xiao, M.: A new linear kernel for undirected planar feedback vertex set: smaller and simpler. In: Gu, Q., Hell, P., Yang, B. (eds.) *AAIM 2014. LNCS*, vol. 8546, pp. 288–298. Springer, Heidelberg (2014)