Towards optimal kernel for connected vertex cover in planar graphs[*]

Łukasz Kowalik[1], Marcin Pilipczuk[1], and Karol Suchan[2,3]

[1] Institute of Informatics, University of Warsaw, Poland
{kowalik@,malcin@}mimuw.edu.pl
[2] Universidad Adolfo Ibáñez, Chile
karol.suchan@uai.cl
[3] WMS, AGH - University of Science and Technology, Poland.

Abstract. We study the parameterized complexity of the connected version of the vertex cover problem, where the solution set has to induce a connected subgraph. Although this problem does not admit a polynomial kernel for general graphs (unless $NP \subseteq coNP/poly$), for planar graphs Guo and Niedermeier [ICALP'08] showed a kernel with at most $14k$ vertices, subsequently improved by Wang et al. [MFCS'11] to $4k$. The constant 4 here is so small that a natural question arises: could it be already an optimal value for this problem? In this paper we answer this quesion in negative: we show a $\frac{11}{3}k$-vertex kernel for Connected Vertex Cover in planar graphs. We believe that this result will motivate further study in search for an optimal kernel.

1 Introduction

Many NP-complete problems, while most likely cannot be fully solved efficiently, admit kernelization algorithms, i.e. efficient algorithms which replace input instances with an equivalent, but often much smaller ones. More precisely, a *kernelization algorithm* takes an instance I of size n and a parameter $k \in \mathbb{N}$, and after a time polynomial in n it outputs an instance I' (called a *kernel*) with a parameter k' such that I is a yes-instance iff I' is a yes instance, $k' \leq k$, and $|I'| \leq f(k)$ for some function f depending only on k. The most desired case is when the function f is polynomial, or even linear (then we say that the problem admits a polynomial or linear kernel). In such a case, when the parameter k is relatively small, the input instance, possibly very large, is "reduced" to a small one. Intuitively, kernelization aims at finding the core difficulty in the input instance. The output instance can be then processed in many ways, including approximation algorithms or heuristics. For small values of k and small kernels, one can often even use an exact (exponential-time) algorithm.

A typical example of the above phenomenon is the well-known VERTEX COVER problem, which admits a kernel with at most $2k$ vertices [4], where k is the size of an optimum vertex cover in the input instance. However, for some problems, reducing to a linear number of vertices seems a hard task, e.g. for the FEEDBACK VERTEX SET problem, the best known result is the $4k^2$-vertex kernel of Thomassé [16]. Even worse, there are many natural problems (examples include DOMINATING SET or STEINER TREE) for which it is proved that they do not admit a polynomial kernel, unless some widely believed complexity hypothesis fails ($FPT \neq W[2]$ in the first case and NP $\subsetneq$ coNP/poly in the second one). This motivates investigation of polynomial kernels in natural restrictions of graph classes. Note that it is of particular interest to guarantee the output instance belongs to the same class as the input instance.

A classic example is the $335k$-vertex kernel for the DOMINATING SET problem in planar graphs due to Alber et al. [1] (note that in planar graphs the number of edges is linear in the number of vertices, thus kernels with linear number of vertices are in fact linear kernels). Later, this work was substantially generalized (see e.g. [9]), and researchers obtained a number of linear kernels for planar graphs. Let us mention the $112k$-vertex kernel for FEEDBACK VERTEX SET by Bodlaender and Penninkx [2], the linear kernel for INDUCED MATCHING by Moser and Sikdar [14] or the $624k$-vertex kernel for the MAXIMUM TRIANGLE PACKING by Guo and Niedermeier [11].

Observe that the constants in the linear functions above are crucial: since we deal with NP-complete problems, in order to find an exact solution in the reduced instance, most likely we need exponential time (or at least superpolynomial, because for planar graphs $2^{O(\sqrt{k})}$-time algorithms are often possible), and these constants appear in the exponents. Motivated by this, researchers seek for linear kernels with constants as small as possible. For example, by now there is known a $67k$-vertex kernel for DOMINATING SET [3], a $28k$-vertex kernel for INDUCED MATCHING [8] or a $75k$-vertex kernel for MAXIMUM TRIANGLE PACKING [17].

In this work, we study the CONNECTED VERTEX COVER problem, a variant of the classical VERTEX COVER: we are given a planar graph $G = (V, E)$ and a parameter k, and we ask whether there exists a vertex cover S (i.e. a set $S \subseteq V$ such that every edge of G has an endpoint in S) of size k which induces a connected subgraph of G. This problem is NP-complete also in planar graphs [10], and, contrary to its simpler relative, it does not admit a polynomial kernel in arbitrary graphs [6]. However, Guo and Niedermeier [11] showed a $14k$-vertex kernel for planar graphs. Very recently, it was improved to $4k$ by Wang et al. [17]. The constant 4 here is already so small that a natural question arises: could it be the optimal value for this problem? In this paper we answer this question in negative: we show a $\frac{11}{3}k$-vertex kernel for CONNECTED VERTEX COVER in planar graphs.

Let us recall that in the analysis of the $4k$-vertex kernel by Wang et al. [17], the vertices of the graph are partitioned into three parts: vertices of degree one, the solution, and the rest of the graph, and it is proven that, after applying a few reduction rules, the sizes of these parts can be bounded by k, k and $2k$,

respectively. We present (in Lemma 2) a deeper analysis of these bounds and we show that an instance where all the bounds are close to being tight is somewhat special. This analysis is the main technical contribution of this paper. We believe that this result will motivate further study in search for an optimal kernel.

Organization of the paper The paper is organized as follows. In Section 2 we present our kernelization algorithm along with a proof of its correctness. In Section 3 we show that our algorithm outputs a kernel with the number of vertices bounded by $\frac{11}{3}k$. Finally, in Section 4 we describe an example which shows that our analysis is tight (and hence improving on the kernel size would require adding new reduction rules to the algorithm).

Terminology and notation We use standard fixed parameter complexity and graph theory terminology, see e.g. [7, 5]. For brevity, we call a vertex of degree d a *d-vertex* and if a vertex v has a d-vertex w as a neighbor we call w a d-neighbor. By $N_G(v)$ we denote the set of neighbors of v and we omit the subscript when it is clear from the context. For a graph G and a subset of its vertices S, by $G[S]$ we denote the subgraph of G induced by S.

2 Algorithm

In this section we present our kernelization algorithm. It works in three phases. In what follows, (G_0, k_0) denotes the input instance.

2.1 Phase 1

Phase 1 is a typical kernelization algorithm. A set of rules is specified and in each rule the algorithm searches the graph for a certain configuration. If the configuration is found, the algorithm performs a modification of the graph (sometimes, also of the parameter k), typically decreasing the size of the graph. Each rule has to be *correct*, which means that the new graph is planar and the graph before application of the rule has a connected vertex cover of size k if and only if the new graph has a connected vertex cover of size equal to the new value of k. We apply the rules in order, i.e. Rule i can be applied only if, for every $j < i$, Rule j does not apply. The first three rules come from [17].

Rule 1. If a vertex v has more than one 1-neighbors, then remove all these neighbors except for one.

Rule 2. For a 2-vertex v with $N(v) = \{u, w\}$ and $uw \in E(G)$, contract the edge uw and decrease the parameter k by one.

Rule 3. For a 2-vertex v with $N(v) = \{u, w\}$ and $uw \notin E(G)$, if v is not a cut-vertex, then remove v and add a new 1-neighbor to each of u and w; otherwise, contract the edge uv and decrease the parameter k by one.

Rule 4. If there is an edge uv such that both u and v have a 1-neighbor, then remove the 1-neighbor of u, contract uv and decrease the parameter k by one.

Rule 5. If there is a 3-vertex v with $N(v) = \{x, y, z\}$ and such that $\deg(z) = 1$, then remove v and z, add an edge xy if it was not present before, and decrease the parameter k by one.

Rule 6. If there are two 3-vertices a and b with a common neighborhood $N(a) = N(b) = \{x, v, y\}$, and such that removing any two vertices from $\{x, v, y\}$ makes the graph disconnected, then remove a and add three 1-vertices x', v' and y', adjacent to x, v and y, respectively.

Rule 7. If there is a 3-vertex a with $N(a) = \{x, v, y\}$ and such that v is a 4-vertex with $N(v) = \{x, a, y, q\}$, where q is a 1-vertex, then remove vertices a, v and q, add an edge xy as well as two 1-vertices x' and y', connected to x and y respectively, and decrease the parameter k by one.

Lemma 1. *Rules 1-7 are correct.*

Proof. The correctness of Rules 1-3 were proven in [17]. Rule 4 is obvious, as both u and v need to be included in any connected vertex cover of the graph. In the proofs for Rules 5-7, by G and G' we denote the graph respectively before and after the currently considered rule was applied.
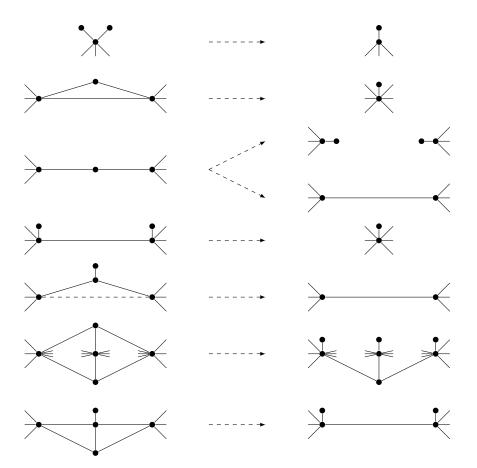
<u>Rule 5.</u> Let S' be a connected vertex cover in G'. We claim that $S = S' \cup \{v\}$ is a connected vertex cover in G. As $v \in S$, S is a vertex cover of G. Moreover, as $xy \in E(G')$, either x or y belongs to S', and if both x and y belong to S', then they remain connected in $G[S]$ via the vertex v. We infer that $G[S]$ is connected.

In the other direction, let S be a minimum connected vertex cover of G. Clearly $v \in S$ and $z \notin S$. As Rules 1-3 are not applicable, the vertices x and y have degrees at least 3. Therefore, as $G[S]$ is connected, x or y belongs to S; without loss of generality assume that $x \in S$. It follows that $S' = S \setminus \{v\}$ is a vertex cover of G'. Moreover, if $y \in S$, then x and y remain connected in $G[S']$ via the edge xy. Thus $G'[S']$ is connected and the proof of the correctness of Rule 5 is finished.

<u>Rule 6.</u> We first note that any connected vertex cover in G' is also a connected vertex cover in G, since it needs to contain x, v, y due to new vertices x', v', y'.

In the other direction, let S be a minimum connected vertex cover in G. Note that S needs to contain at least two vertices out of $\{x, v, y\}$, since removing $\{x, v, y\} \setminus S$ cannot disconnect G. If $a \in S$ while $b \notin S$, then $(S \setminus \{a\}) \cup \{b\}$ is also a minimum connected vertex cover of G. Thus we can assume that if $a \in S$ then $b \in S$ as well. We infer that if $\{x, v, y\} \subseteq S$, then $S \setminus \{a\}$ is a connected vertex cover in G' too. Otherwise, without loss of generality let $v \notin S$. As S is a vertex cover, $a, b \in S$. As $N_G(a) = N_G(b)$, we have that $G[S \setminus \{a\}]$ is connected. Thus $(S \setminus \{a\}) \cup \{v\}$ is a connected vertex cover of both G and G', and Rule 6 is correct.

<u>Rule 7.</u> Let S' be a minimum connected vertex cover in G'. The vertices x' and y' ensure that $x, y \in S'$ and $x', y' \notin S'$. Then clearly $S = S' \cup \{v\}$ is a connected vertex cover in G, as x and y remain connected in $G[S]$ via the vertex v.

Fig. 1. Rules used in Phase 1.

In the other direction, let S be a minimum connected vertex cover in G. Clearly $v \in S$ and $q \notin S$. Since Rule 3 is not applicable, the degrees of x and y are at least 3. As $G[S]$ is connected, we infer that x or y belongs to S; without loss of generality we can assume that $x \in S$. We claim that $S' = (S \setminus \{v, a\}) \cup \{y\}$ is a connected vertex cover of G' of size at most $|S| - 1$. If $y \in S$, the statement is obvious. Otherwise, since S is a vertex cover of G, we have that $a \in S$ and $|S'| = |S| - 1$. This finishes the proof of the lemma.

2.2 Phase 2

In what follows, the graph obtained after Phase 1 is denoted by G_1. Graph G_1, similarly as the reduced graph in [17], does not contain 2-vertices and every vertex has at most one 1-neighbor. The goal of Phase 2 is to decrease the number of vertices in the graph by replacing some pairs of 1-vertices by 2-vertices using the following rule (a kind of inverse of 3):

Rule 8. If there are two vertices u and v, both with 1-neigbors, say x_u and x_v, and u and v are incident with the same face then identify x_u and x_v.

In the above rule we assume that we have a fixed plane embedding of G_1 (if the input graph is not given as a plane embedding, it can be found in linear time from the set of edges by an algorithm of Hopcroft and Tarjan [12]). Note that Rule 8 preserves planarity. Observe also that since the graph before application of this rule is connected, then the in new graph the vertex which appears after identifying x_u and x_v is not a cut-vertex. Moreover, Rule 4 guarantees that $uv \notin E(G_1)$. It follows that the correctness of Rule 3 implies the corectness of Rule 8.

In order to get our bound on the kernel we do not apply Rule 8 greedily, but we maximize the number of times it is applied. To this end, an auxiliary graph G_M is built. Let S_1 be the set of vertices of graph G_1 that have a 1-neighbor. The vertex set of G_M is equal to S_1. Two vertices u and v of S_1 are adjacent in G_M if and only if u and v are incident to the same face in G_1. (Note that G_M does not need to be planar.) Our algorithm finds a maximum matching M_0^* in G_M in polynomial time. Next we modify M_0^* to get another matching M^* of the same size. We start with $M^* = \emptyset$. Then, for each face f of G_1 we find the set M_f of all edges uv of M_0^* such that both u and v are incident to f. Let $v_1, \ldots, v_{2|M_f|}$ be the vertices of $V(M_f)$ in the clockwise order around f. We add the set $\{v_1 v_2, \ldots, v_{2|M_f|-1} v_{2|M_f|}\}$ to M^*. It is clear that after applying this procedure to all the faces of G_1 we have $|M^*| = |M_0^*|$. Moreover, the graph $G_1 \cup M^*$ is planar, and we can extend the plane embedding of G_1 to a plane embedding of $G_1 \cup M^*$. It follows that Rule 8 can be applied $|M^*|$ times according to the matching M^*. The time needed to perform Phase 2 is dominated by finding the matching M_0^* which can be done in $O(\sqrt{|V(G_M)|} \cdot |E(G_M)|) = O(n^{2.5})$ using the Micali-Vazirani algorithm [13].

2.3 Phase 3

Let (G_2, k_2) be the instance obtained after Phase 2. In the next section we show that if G_2 contains a connected vertex cover of size k_2, then $|V(G_2)| \leq \frac{11}{3} k_2$. Together with the correctness of Rules 1-8 this justifies the correctness of the last step of our kernelization algorithm: if $|V(G_2)| > \frac{11}{3} k_2$ the algorithm reports that in the input graph G_0 does not contain a connected vertex cover of size k_0.

3 Analysis

Let S be a minimum connected vertex cover in G_1. Clearly, S is a connected vertex cover in G_2, as S does not contain any 1-vertices. Moreover, the correctness of Rule 3 ensures that S is also a minimum one. The goal of this section is to show that $|V(G_2)| \leq \frac{11}{3} |S|$. However, most of the time we fill focus on the graph G_1.

Observe that we can assume that every vertex in S has degree at least 3: this is not the case only in the trivial case when G_1 is a single edge. Note that $V(G_1) \setminus S$ is an independent set. We denote it by I. The set I is further partitioned into three subsets: I_1, I_3 and $I_{\geq 4}$ which contain vertices of I of degree 1, 3 and at least 4 respectively. Note that each neighbor of a vertex in I_1 belongs to S. We denote by S_1 the set vertices in S which have a neigbor in I_1 and let $S_{\geq 3} = S \setminus S_1$.

Roughly, we want to show that S is a big part of $V(G_1)$. Following [17], we can bound $|I|$ as follows. Consider the bipartite planar graph B which consists of the edges of G_1 between S and $I \setminus I_1$. Then $3|I \setminus I_1| \leq |E(B)| < 2(|S| + |I \setminus I_1|)$, where the second inequality follows from the well-known fact that in a simple bipartite planar graph the number of edges is smaller than twice the number of vertices. This implies that $|I \setminus I_1| < 2|S|$. Since in G_1 every vertex has at most one 1-neighbor, so $|I_1| = |S_1| \leq |S|$. Hence $|I| < 3|S|$. However, with our additional rules, this inequality is not tight. There are three events which make $|I|$ even smaller than $3|S|$. Obviously, this happens when the matching M^* is large. Second good event is when $S_{\geq 3}$ is large, because this means that $|S_1|$ is much smaller than $|S|$, so we can improve our bound on $|I_1|$. Finally, it is also good when the set $I_{\geq 4}$ is large, because then we get a better bound on $|I \setminus I_1|$. We will show that at least one of these three situations happen in G_1. This is guarnteed by the following lemma.

Lemma 2. $|S_{\geq 3}| + |I_{\geq 4}| + |M^*| \geq |S|/3$.

In order to prove Lemma 2 we need the following auxiliary result.

Lemma 3. *In any simple planar graph which contains $n_{\geq 3}$ vertices of degree at least 3 there is a matching of cardinality at least $n_{\geq 3}/3$.*

Proof. Let G be an arbitrary planar graph. Let $n_{\leq 2}$ denote the number of vertices of G of degree at most 2 and $V_{\geq 3}(G)$ — the set of vertices of degree at least 3 in G. We use induction on $n_{\leq 2}$. If $n_{\leq 2} = 0$, the lemma follows by a result

of Nishizeki and Baybars [15] who proved that any n-vertex planar graph of minimum degree 3 contains a matching of size at least $\frac{1}{3}(n+2)$. Now assume $n_{\leq 2} > 0$. Let v be an arbitrary vertex of degree at most 2. There are three cases to consider.

Case 1: $\deg(v) = 0$. Then we remove v and we apply the induction hypothesis.

Case 2: $\deg(v) = 1$. Let w be its only neighbor. If $\deg(w) \neq 3$ we can just remove v and use the induction hypothesis. If $\deg(w) = 3$ then let G' be the graph obtained from G by removing v and w. G' has at least $n_{\geq 3} - 3$ vertices of degree at least 3 ($V_{\geq 3}(G) \setminus V_{\geq 3}(G')$ may contain only w and its two neighbours different than v), so by the induction hypothesis G' has a matching M_0 of size at least $n_{\geq 3}/3 - 1$. Then $M_0 \cup \{vw\}$ is the desired matching in G.

Case 3: $\deg(v) = 2$. Let $N(v) = \{x, y\}$. If $xy \notin E(G)$ then we obtain G' from G by removing v and adding an edge xy. Note that G' is simple, planar and $V_{\geq 3}(G) = V_{\geq 3}(G')$. By the induction hypothesis G' has a matching M_0 of size at least $n_{\geq 3}/3$. If $xy \notin M_0$ then M_0 is the desired matching in G, otherwise we just use $M_0 \setminus \{xy\} \cup \{xv\}$. Hence we are left with the case when $xy \in E(G)$. If $\deg(x) \neq 3$ and $\deg(y) \neq 3$ then we can just remove v and use the matching from the induction hypothesis. Hence, w.l.o.g. $\deg(x) = 3$. Then let G' be the graph obtained from G by removing v and x. G' has at least $n_{\geq 3} - 3$ vertices of degree at least 3 ($V_{\geq 3}(G) \setminus V_{\geq 3}(G')$ may contain only x, y and the third neighbour of x different than y and v), so by the induction hypothesis G' has a matching M_0 of size at least $n_{\geq 3}/3 - 1$. Then $M_0 \cup \{vx\}$ is the desired matching in G. $\square$

Proof (of Lemma 2). Let us consider an auxiliary graph W. Its vertex set consists of three types of vertices: elements of S, $I_{\geq 4}$ and additionally, for each $v \in S_{\geq 3}$ we add three vertices v_1, v_2 and v_3. The edge set can be contructed as follows. First, for every $v, w \in S \cup I_{\geq 4}$, we add vw to W whenever $vw \in E(G_1)$. Second, for each $v \in S_{\geq 3}$ we add three edges vv_1, vv_2 and vv_3. Finally, we consider faces of G_1, one by one. For each such face f we do the following. Let $u_1, \ldots, u_\ell$ be the vertices of $V(W)$ incident to f, in clockwise order. Then, if $\ell > 1$, we add edges $u_1 u_2, u_2 u_3, \ldots u_{\ell-1} u_\ell, u_\ell u_1$. In this process we do not create double edges, i.e. if an edge is already present in the graph, we do not add another copy of it. It is clear that W is a planar graph. Let us prove the following claim.
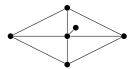

Fig. 2. A 6-vertex graph with minimum connected vertex cover of size 3.

Claim: Every vertex $v \in S_1$ has degree at least 3 in graph W, unless G_1 is the graph from Fig. 2.

By Rule 1 there is exactly one 1-neighbor of v in G_1, let us call it q. By Rule 5, $\deg_{G_1}(v) \geq 4$, so we have $|N_{G_1}(v) \setminus \{q\}| \geq 3$. Note that if v is adjacent in G_1 with a vertex $x \in V(W)$, then $vx \in E(W)$. Hence we can assume that $N_{G_1}(v) \setminus \{q\}$ contains at least one vertex outside $V(W)$, for otherwise the claim holds. Let a denote such a neigbor.

Since $a \notin I_1 \cup I_{\geq 4} \cup S$, we infer that $\deg_{G_1}(a) = 3$. Let x and y be the two other neighbors of a except from v. Since $a \notin S$, we know that $x, y \in S \subseteq V(W)$. By our construction $vx, vy \in E(W)$, so $\deg_W(v) \geq 2$.

Now assume that $(N_{G_1}(v) \setminus \{q\}) \setminus V(W) = \{a\}$. It follows that $|N_{G_1}(v) \cap V(W)| \geq \deg_{G_1}(v) - 2 \geq 2$. If $N_{G_1}(v) \cap V(W)$ contains a vertex $z \notin \{x, y\}$, then $vz \in E(W)$ and the claim holds. It follows that $N_{G_1}(v) \cap V(W) = \{x, y\}$ and $N_{G_1}(v) = \{a, x, y, q\}$. Hence the vertices v, x, y, a, q induce the configuration from Rule 7, a contradiction.

Finally assume that $|(N_{G_1}(v) \setminus \{q\}) \setminus V(W)| \geq 2$. Consider an arbitary vertex $b \in (N_{G_1}(v) \setminus \{q\}) \setminus V(W)$, $b \neq a$. As shown above, $\deg_{G_1}(b) = 3$ and if x' and y' denote the two other neighbors of b except from v, then $x', y' \in V(W)$ and $vx', vy' \in E(W)$. It follows that $\{x, y\} = \{x', y'\}$, for otherwise the claim holds. It implies that when $|(N_{G_1}(v) \setminus \{q\}) \setminus V(W)| \geq 3$, then G_2 contains $K_{3,3}$ as a subgraph, a contradiction with planarity. Hence, $\{q, a, b\} \subseteq N_{G_1}(v) \subseteq \{q, a, b, x, y\}$. Note that a removal of v disconnects q from the rest of the graph G_1. Thus, as Rule 6 is not applicable for vertices v, x, y, a, b, $G_1 \setminus \{x, y\}$ is connected. However, $N_{G_1}(\{a, b, v, q\}) = \{x, y\}$ and y is not of degree 2, as otherwise Rule 3 would be applicable. We infer that G_1 is isomorphic to the graph from Figure 2. This finishes the proof of the claim.

Now we return to the proof of Lemma 2. If G_1 is the graph from Fig. 2 we see that $|S_{\geq 3}| = 2$ and $|S| = 3$ so the lemma holds. Hence by the above claim we can assume that $\deg_W(v) \geq 3$ for any $v \in S_1$. Since also each vertex v in $S_{\geq 3}$ has at least three neighbors v_1, v_2, v_3 in W, we conclude that for any $v \in S$ we have $\deg_W(v) \geq 3$. By Lemma 3, graph W contains a matching M of size

$$|M| \geq |S|/3. \tag{1}$$

The edges of M are of two kinds:

- edges incident with a vertex in $S_{\geq 3} \cup I_{\geq 4}$ (there are at most $|S_{\geq 3}| + |I_{\geq 4}|$ of such edges),
- edges with both endpoints in S_1 (there are at most $|M^*|$ of such edges).

Hence,

$$|M| \leq |S_{\geq 3}| + |I_{\geq 4}| + |M^*|. \tag{2}$$

By combining (1) with (2) we get the claim of the lemma. $\square$

Theorem 4. *Let G_2 be the graph obtained after the kernelization algorithm and let S be any minimum connected vertex cover of G_2. Then $|V(G_2)| \leq \frac{11}{3}|S|$.*

Proof. Consider the bipartite planar graph B which consists of the edges of G_1 between S and $I_3 \cup I_{\geq 4}$. Then $3|I_3| + 4|I_{\geq 4}| \leq |E(B)| < 2(|S| + |I_3| + |I_{\geq 4}|)$. This implies that

$$|I_3| + |I_{\geq 4}| < 2|S| - |I_{\geq 4}|. \tag{3}$$

By Lemma 2, $|S_1| \le 2|S_{\ge 3}| + 3|I_{\ge 4}| + 3|M^*|$. Using this we get

$$|I_1| = |S_1| = \frac{2}{3}|S_1| + \frac{1}{3}|S_1| \le \frac{2}{3}|S_1| + \frac{2}{3}|S_{\ge 3}| + |I_{\ge 4}| + |M^*|. \qquad (4)$$

Now we are ready to bound the number of vertices in G_2:

$$
\begin{aligned}
|V(G_2)| &= |S| + |I_1| - |M^*| + |I_3| + |I_{\ge 4}| \\
&\le^{(3),(4)} |S| + \frac{2}{3}|S_1| + \frac{2}{3}|S_{\ge 3}| + 2|S| \\
&= \frac{11}{3}|S|.
\end{aligned}
$$

$\square$

4   An example with tight analysis

In this section we show an example of a planar graph where the analysis from the previous section is tight. That is, we construct a graph G with the following properties: no reduction of Phase 1 is applicable, G admits a connected vertex cover of size roughly $|V(G)|/4$, and Rule 8 may be used at most $|V(G)|/12$ times.

Consider a gadget graph H depicted on Figure 3. For any integer $\ell \ge 3$, the graph G_ℓ is constructed by taking ℓ copies of H and connecting them in the following manner:

1. In all copies of H, all vertices s are identified into a single vertex; similarly, all vertices t are identified into a single vertex.
2. The vertex v from the i-th copy of H is identified with the vertex u from the $(i + 1)$-th copy of H and the vertex v from the last copy of H is identified with the vertex u from the first copy of H; moreover, the 1-neighbours of the aforementioned pairs vertices are also identified.
3. Any multiple edges, resulting in the above operations, are removed.

It is easy to see that the graph G_ℓ is planar and a direct check ensures us that Rules 1-7 are not applicable to G_ℓ. Now note that the set S consisting of vertices s, t and all vertices x, y, u and v is a minimum connected vertex cover of G_ℓ. Moreover, $|V(G_\ell)| = 12\ell + 2$ and $|S| = 3\ell + 2$. Let us analyze sets S_1, $S_{\ge 3}$, I_1, I_3 and $I_{\ge 4}$ (defined as in the previous section) in the graph G.

1. $S_{\ge 3} = \{s, t\}$, thus $|I_1| = |S_1| = 3\ell$, i.e., almost every vertex in S has a 1-neighbour;
2. $I_{\ge 4} = \emptyset$ and $|I_3| = 6\ell$, i.e., $|I_3| + |I_{\ge 4}| = 2|S| - 4$;

Finally, let us analyze how many times Rule 8 can be applied to G_ℓ. Note that no pair of vertices x and y lie on the same face of the graph G_ℓ, thus any edge in the matching M^* (constructed in Phase 2) needs to have an endpoint in a vertex u or v. There are ℓ such vertices, thus $|M^*| \le \ell$ (in fact it is easy to see that $|M^*| = \ell$, as we can match u to x in every copy of H). We conclude that

$$|S_{\ge 3}| + |I_{\ge 4}| + |M^*| = 2 + 0 + \ell = \frac{|S| + 4}{3},$$

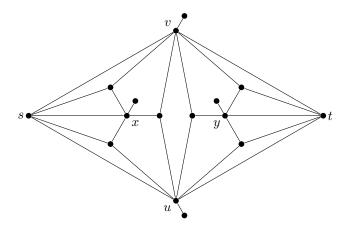and the bound in Lemma 2 is tight up to an additive constant.

Fig. 3. A gadget graph H used in the construction of an example with tight analysis.

References

1. Alber, J., Fellows, M.R., Niedermeier, R.: Polynomial-time data reduction for dominating set. J. ACM 51(3), 363–384 (2004)
2. Bodlaender, H.L., Penninkx, E.: A linear kernel for planar feedback vertex set. In: Grohe, M., Niedermeier, R. (eds.) IWPEC. Lecture Notes in Computer Science, vol. 5018, pp. 160–171. Springer (2008)
3. Chen, J., Fernau, H., Kanj, I.A., Xia, G.: Parametric duality and kernelization: Lower bounds and upper bounds on kernel size. SIAM J. Comput. 37(4), 1077–1106 (2007)
4. Chen, J., Kanj, I.A., Jia, W.: Vertex cover: Further observations and further improvements. J. Algorithms 41(2), 280–301 (2001)
5. Diestel, R.: Graph Theory. Springer (2005)
6. Dom, M., Lokshtanov, D., Saurabh, S.: Incompressibility through colors and IDs. In: Albers, S., Marchetti-Spaccamela, A., Matias, Y., Nikoletseas, S.E., Thomas, W. (eds.) ICALP (1). Lecture Notes in Computer Science, vol. 5555, pp. 378–389. Springer (2009)
7. Downey, R.G., Fellows, M.R.: Parameterized Complexity. Springer (1999), citeseer.ist.psu.edu/downey98parameterized.html
8. Erman, R., Kowalik, L., Krnc, M., Walen, T.: Improved induced matchings in sparse graphs. Discrete Applied Mathematics 158(18), 1994–2003 (2010)
9. Fomin, F.V., Lokshtanov, D., Saurabh, S., Thilikos, D.M.: Bidimensionality and kernels. In: Charikar, M. (ed.) SODA. pp. 503–510. SIAM (2010)
10. Garey, M.R., Johnson, D.S.: The rectilinear steiner tree problem in np complete. SIAM Journal of Applied Mathematics 32, 826–834 (1977)
11. Guo, J., Niedermeier, R.: Linear problem kernels for np-hard problems on planar graphs. In: Proc. ICALP'07. Lecture Notes in Computer Science, vol. 4596, pp. 375–386 (2007)
12. Hopcroft, J., Tarjan, R.: Efficient planarity testing. J. ACM 21, 549–568 (October 1974)

13. Micali, S., Vazirani, V.V.: An $O(\sqrt{|V|}|E|)$ algorithm for finding maximum matching in general graphs. In: FOCS. pp. 17–27. IEEE (1980)
14. Moser, H., Sikdar, S.: The parameterized complexity of the induced matching problem. Discrete Applied Mathematics 157(4), 715–727 (2009)
15. Nishizeki, T., Baybars, I.: Lower bounds on the cardinality of the maximum matchings of planar graphs. Discrete Mathematics 28(3), 255–267 (1979)
16. Thomassé, S.: A $4k^2$ kernel for feedback vertex set. ACM Transactions on Algorithms 6(2) (2010)
17. Wang, J., Yang, Y., Guo, J., Chen, J.: Linear problem kernels for planar graph problems with small distance property. In: Proc. MFCS'11. Lecture Notes in Computer Science, vol. 6907, pp. 592–603 (2011)