

Linear Kernel for Planar Connected Dominating Set

Daniel Lokshtanov* Matthias Mnich† Saket Saurabh ‡

Abstract

We provide polynomial time data reduction rules for CONNECTED DOMINATING SET on planar graphs and analyze these to obtain a linear kernel for the planar CONNECTED DOMINATING SET problem. To obtain the desired kernel we introduce a method that we call *reduce or refine*. Our kernelization algorithm analyzes the input graph and either finds an appropriate reduction rule that can be applied, or zooms in on a region of the graph which is more amenable to reduction. We find this method of independent interest and believe that it will be useful to obtain linear kernels for other problems on planar graphs.

1 Introduction

Preprocessing of data is one of the oldest and widely used methods in practical algorithms. Parameterized Complexity provides a natural way to measure the quality of preprocessing. In parameterized complexity a problem Π consists of a pair (I, k) where I is the input and k is a parameter (which typically is the solution size). A problem Π is said to have a *kernelization algorithm* if there exists a preprocessing algorithm, which given a parameterized instance (I, k) of Π , runs in time polynomial in $|I|$ and k and outputs a simpler instance (I', k') of Π , such that (I, k) is a yes-instance if and only if (I', k') is a yes-instance and the size of (I', k') is bounded by a function of k alone. The reduced instance I' is called the kernel for the problem. The problem Π is said to have a polynomial (linear) kernel if the reduced instance is bounded by a polynomial (linear) function of k .

Kernelization has been extensively studied, resulting in polynomial kernels for a variety of problems. Notable examples include a $2k$ -vertex kernel for VERTEX COVER [7], a $355k$ -vertex kernel for DOMINATING SET on planar graphs [1] which later was improved to a $67k$ -vertex kernel [6], and a $O(k^2)$ kernel for FEEDBACK VERTEX SET [19] parameterized by the solution size. A significant amount of research has gone into providing linear kernels for NP-hard problems on planar graphs. A foundation for linear kernelization on planar graphs was built by Alber et al. [1] who gave a $335k$ -sized vertex kernel for planar DOMINATING SET. The main ingredient in the analysis of the reduced instance was the notion of region decomposition for the input planar graph where the number of regions depended linearly on the size of the parameter. These ideas were later abstracted by Guo and Niedermeier [15] who gave a framework to obtain linear kernels for planar graph problems possessing a certain “locality property”. This framework has been successfully applied to yield linear kernels for the CONNECTED VERTEX COVER, MINIMUM EDGE DOMINATING SET, MAXIMUM TRIANGLE PACKING, EFFICIENT EDGE DOMINATING SET, INDUCED MATCHING and FULL-DEGREE SPANNING TREE problems [15, 16, 18]. However, the framework proposed by Guo and Niedermeier [15] in its current form is not able to handle problems like FEEDBACK VERTEX SET and ODD CYCLE TRANSVERSAL because these do not admit the “locality property” required by the framework. However, Bodlaender

*University of California, San Diego, USA. daniello@ii.uib.no

†Technische Universiteit Eindhoven, Eindhoven, The Netherlands. m.mnich@tue.nl

‡The Institute of Mathematical Sciences, Chennai, India. saket@imsc.res

and Penninkx [4] and Bodlaender et al. [5] have obtained linear kernels for FEEDBACK VERTEX SET and CYCLE PACKING on planar graphs respectively. These linear kernels for FEEDBACK VERTEX SET and CYCLE PACKING are obtained by using a generalized notion of distance, which enables the problems to become “local in certain sense”.

The list of problems for which linear kernels are known for planar graphs excludes problems which demand the solution to be connected. The mere exception is the the CONNECTED VERTEX COVER problem for which the reduction rules for planar VERTEX COVER apply [15]. In this article we try to fill this void by studying the CONNECTED DOMINATING SET problem for planar graphs from the viewpoint of linear kernelization. The problem is defined as follows.

CONNECTED DOMINATING SET: Given a graph $G = (V, E)$ and a positive integer k , the objective is to find a subset $D \subseteq V$ of size at most k such that $G[D]$ is connected and for every vertex $v \in V$ either $v \in D$ or one of its neighbors is in D .

CONNECTED DOMINATING SET is a well-studied NP-hard problem that finds applications in various network design problems. It remains NP-hard when restricted to the class of planar graphs [11], and has a $O(\log n)$ -approximation algorithm [13]. The parameterized version of the problem is known to be W[2]-complete for general graphs and admits a sub-exponential time parameterized algorithm for planar graphs [8]. In general graphs the problem has also been studied in the realm of moderately exponential time algorithms leading to an algorithm with running time $O(1.9407^n)$ [9]. Here we provide polynomial time data reduction rules for CONNECTED DOMINATING SET on planar graphs which lead to a linear kernel for the problem. In particular, we prove the following theorem.

Theorem 1 *The CONNECTED DOMINATING SET problem parameterized by solution size k has a linear kernel on planar graphs.*

This answers a question asked by Guo [14] during the visit of the third author to Jena in 2007. Our results are based on the *reduce-or-refine* technique, that we introduce here. Until now the notion of *region decomposition* was used only in the analysis of the kernel size, and not explicitly applied in the reduction rules. We utilize the fact that a region decomposition can be obtained in polynomial time given a solution set S . In particular, we compute S using the known polynomial time approximation scheme for CONNECTED DOMINATION SET and compute the decomposition from S using algorithms described by Alber et al. [1] and Guo and Niedermeier [15]. The main technical part of our proofs is devoted to showing that if a region contains more vertices than a fixed constant, we can in polynomial time find a vertex in this region whose removal will not affect the size of an optimal solution. The idea is to check whether the region contains more than a fixed constant number of copies of a particular structure. If so then we can *reduce* the graph by removing a vertex in such a structure. If there are few or no copies of the structure in this region then we can zoom in on, or *refine*, to a smaller region that still contains many vertices but completely excludes the structure. The process is then repeated for a different “bad” structure until the region we have zoomed in on looks so simple that it is easy to identify a vertex to remove.

Since the number of regions in our computed region decomposition is $O(k)$, if the graph has too many vertices then we can identify a region in which a useless vertex can be found. Combining this with the fact that the number of edges in a planar graph is upper bounded by a linear function of the number of the vertices we obtain the desired linear upper bound on the size of the kernel.

2 Preliminaries

In this section we collect necessary definitions and results required to obtain a linear kernel for the problem. We also give a few basic reduction rules for CONNECTED DOMINATING SET.

Let $G = (V, E)$ be a connected planar graph, without loops or multiple edges. For each vertex $v \in V$, define the *open neighbourhood* of v as $N(v) = \{u \in V \mid \{u, v\} \in E\}$ and the *closed neighbourhood* of v as $N[v] = N(v) \cup \{v\}$. A vertex v is *universal* for G if $N[v] = V$. A path in G between distinct vertices v, w is called $[v, w]$ -*path*. Let $G[U]$ denote the induced graph on U , for any vertex set $U \subseteq V$. For a graph $G = (V, E)$ a subset $D \subseteq V$ is called a *dominating set* if for every vertex $v \in V$ either $v \in D$ or $N(v) \cap D \neq \emptyset$. For a graph G , the size of a minimum dominating set is denoted by $\gamma(G)$ and the size of a minimum connected dominating set is denoted by $\gamma_c(G)$. A graph that can be drawn in the plane without edge crossing is called *planar graph*. A *plane graph* is a planar graph with a fixed embedding in the plane. Throughout the paper, we assume that we are working with an arbitrary but fixed embedding of G in the plane.

With respect to connected dominating sets, the following reductions rules will frequently help to simplify the input graph. If G has a universal vertex v then $\{v\}$ is a minimum connected dominating set for G . Henceforth we assume that G has no universal vertex.

Lemma 1 *Let G be a graph and let v be a vertex of G contained in some minimum connected dominating set S of G . Let G_v be the graph obtained from G by removing the edges of $G[N(v)]$. Then $\gamma_c(G) = \gamma_c(G_v)$.*

Proof: Let S be a minimum connected dominating set of G containing v . In G_v , any vertex $u \in N(v) \setminus S$ is still dominated by v ; and any pair of vertices in $N(v) \cap S$ are connected via v . Hence S is a connected dominating set of G_v .

Let S' be a minimum connected dominating set of G_v . Since G_v is a subgraph of G on the same vertex set as G , it follows that S' is a connected dominating set of G . \square

Whenever possible, we remove “twin vertices” from the graph.

Lemma 2 *Let $G = (V, E)$ be a graph and let u, u' be distinct vertices such that $N[u] = N[u']$. Then $\gamma_c(G) = \gamma_c(G - u')$.*

Proof: Let S be a minimum connected dominating set of G . Since $N[u] = N[u']$, vertices u and u' are adjacent. By minimality of S , at most one of u, u' belongs to S . Without loss of generality $u' \notin S$; thus S is a connected dominating set of $G - u'$.

Let S' be a minimum connected dominating set of $G - u'$. Then either $u \in S'$ or $u \notin S'$. If $u \in S'$ then u' is dominated by u and S' is a connected dominating set of G . And if $u \notin S'$ then some neighbor w of u belongs to S' , which by $N[u] = N[u']$ also dominates u' in G . In both cases we see that S' is a connected dominating set of G , which concludes the proof. \square

This reduction rule is only used to reduce the graph in practice; we do not use it in the analysis of the kernel size. Now we introduce the notions of a region and a region decomposition that were first introduced by Alber et al. [1].

Definition 1 *Let G be a plane graph and let v, w be distinct vertices of G . A region $R(v, w)$ between v and w is a closed subset of the plane such that*

- *the boundary of $R(v, w)$ is formed by two simple $[v, w]$ -paths each of length at most three, and*
- *all vertices strictly inside region $R(v, w)$ belong to $N(v) \cup N(w)$, and are called inner vertices of $R(v, w)$.*

If $R(v, w), R'(v, w)$ are regions between v and w then $R(v, w) \cup R'(v, w)$ denotes the region that is defined by the union of the closed subsets of the plane defined by $R(v, w)$ and $R'(v, w)$. We use $V(R(v, w))$ to denote the set of inner vertices of the region $R(v, w)$.

Definition 2 Let $G = (V, E)$ be a plane graph and let $S \subseteq V$. An S -region decomposition of G is a set \mathcal{R} of regions $R(v, w)$ between distinct vertices $v, w \in S$ such that

- each region $R(v, w)$ contains no vertex from $S \setminus \{v, w\}$, and
- any two distinct regions can only intersect in their boundaries.

For an S -region decomposition \mathcal{R} , let $V(\mathcal{R}) = \cup_{R(v,w) \in \mathcal{R}} V(R(v, w))$. An S -region decomposition \mathcal{R} of G is maximal if there is no region $R(v, w)$ such that $\mathcal{R} \cup \{R(v, w)\}$ is an S -region decomposition of G satisfying $V(\mathcal{R}) \subsetneq V(\mathcal{R} \cup \{R(v, w)\})$.

We now state two known results about maximal region decompositions. The results say that given a plane graph G and a dominating set S , one can obtain an S -region decomposition of G with $O(|S|)$ regions that together cover all but $O(|S|)$ vertices of G .

Proposition 1 (Guo and Niedermeier [15]) Let G be a plane graph and let S be a dominating set of G . There exists a maximal S -region decomposition of G containing at most $3|S|$ regions. Furthermore given a set S a maximal S -region decomposition can be found in polynomial time.

Proposition 2 (Alber et al. [1]) Let G be a plane graph and let S be a dominating set of G . If \mathcal{R} is a maximal S -region decomposition of G then at most $170|S|$ vertices of G do not belong to \mathcal{R} .

Since any connected dominating set of a graph is also a dominating set, Propositions 1 and 2 together imply that a planar graph G has a maximal S -region decomposition for a connected dominating set S with $O(|S|)$ regions covering all but $O(|S|)$ vertices of G .

3 A Reduce-or-Refine Scheme

In this section, we provide a polynomial time algorithm to bound the number of vertices in each region by some constant C . As long as there exists a region with more than C vertices, this region will either be “refined” into multiple regions or some vertices will be removed from it. We show that in polynomial time the algorithm produces an instance where the number of vertices in each region is bounded by a constant and the total number of regions is $O(k)$.

First, we give a detailed outline of how we will obtain a linear kernel for the connected dominating set problem on planar graphs. Given a planar graph G and a positive integer k , whenever possible, we reduce it with respect to a few reduction rules, which can be applied in polynomial time and simplify G . Then, for a given connected dominating set D of G (obtained by applying approximation algorithm) and a corresponding maximal D -region decomposition of G , we focus on a region with a maximum number of vertices. If that number is “large”, then we will color the vertices of that region in a specific way. This coloring will either guide us to a new reduction rule, an application of which will simplify the graph, or to a large homogeneous subregion in which we can remove some redundant vertices or edges without changing the size of a minimum connected dominating set. On the other hand, if the maximum number of vertices in each region is “small”, then we will be able to establish a kernel of linear size, based on a linear bound on the number of regions in a maximal D -region decomposition of G .

We now begin to establish the linear kernel, by following the steps as outlined above. We start with a following observation.

Lemma 3 Let $G = (V, E)$ be a plane graph and let $R(v, w)$ be a region between $v, w \in V$. There exists a minimum connected dominating set S of G such that S contains at most two inner vertices from $R(v, w)$.

Proof: Let S be a minimum connected dominating set of G with more than two inner vertices from $R(v, w)$. Let P_1 and P_2 be the two $[v, w]$ -paths forming the boundary of $R(v, w)$. Then either P_1 or P_2 shares some vertex with S , as the graph $G[S]$ contains a path between vertices inside $R(v, w)$ and vertices outside $R(v, w)$. Without loss of generality, let P_1 have some vertex in common with S . Let S' be the vertex set obtained from S by removing all inner vertices of $R(v, w)$ from it and adding all vertices from $P_1 \setminus S$. Clearly, $|S'| \leq |S|$, since the number of vertices being added is at most three (recall that P_1 has at most four vertices including w and v), and we remove at least three vertices from S . The set S' is also a connected dominating set since every vertex in $R(v, w)$ (including boundary vertices) is adjacent to one of v or w , both of which are in S' . \square

Lemma 3 allows us to assume that a minimum sized connected dominating set we are looking for contains at most two internal vertices in any region.

Let $N_R(v, w)$ denote the common neighbourhood of v and w in the region $R(v, w)$, that is,

$$N_R(v, w) = \{u \in R(v, w) \mid u \in N(v) \cap N(w)\}.$$

Our case analysis is based on the number of common vertices v and w have in the region $R(v, w)$. We have primarily three cases: (a) $N_R(v, w)$ has at least 106 vertices; (b) $N_R(v, w)$ is non-empty has at most 105 vertices; and (c) $N_R(v, w)$ is empty. The number 106 is just an appropriately chosen constant so that the proof of various lemmas go through.

Case 1: $N_R(v, w)$ contains at least 106 vertices.

Let x_1, \dots, x_ℓ be a labeling of the vertices in $N_R(v, w)$ such that for all $i = 1, \dots, \ell-1$, there is a region $r_i(v, w)$ between v and w with clockwise ordering (v, x_i, w, x_{i+1}) of boundary vertices. We define a coloring c on the set $\{r_1(v, w), \dots, r_{\ell-1}(v, w)\}$. We color the region $r_i(v, w)$ black, white, black-and-white, or transparent according to the scheme outlined below. We refer to Figure 1 for an example.

- black, if $r_i(v, w)$ contains some inner vertices adjacent to v and no inner vertices adjacent to w ,
- white, if $r_i(v, w)$ contains some inner vertices adjacent to w and no inner vertices adjacent to v ,
- black-and-white, if $r_i(v, w)$ contains some inner vertices adjacent to v and some inner vertices adjacent to w ,
- transparent, if $r_i(v, w)$ contains no inner vertices.

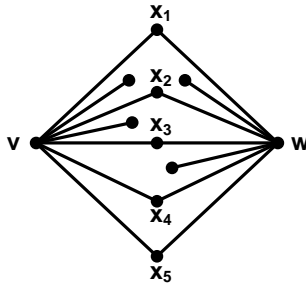


Figure 1: Regions $r_i = r_i(v, w)$ for $i = 1, 2, 3, 4$. The coloring c colors r_1 black-and-white, r_2 black, r_3 white and r_4 transparent.

The *black (white) weight* of c is the number of regions that are colored black or black-and-white (white or black-and-white).

Observation 1 Let $G = (V, E)$ be a plane graph and let $R(v, w)$ be a region between $v, w \in V$. If the coloring c has black weight at least 7 then any minimum connected dominating set of G containing at most two inner vertices of $R(v, w)$ contains v . Similarly, if the coloring c has white weight at least 7 then any minimum connected dominating set of G containing at most two inner vertices of $R(v, w)$ contains w .

Proof: Let S be a minimum connected dominating set of G containing at most two inner vertices of $R(v, w)$, and excluding v . Suppose that c has black weight 7 or more. We exhibit a subset that cannot be dominated by S , which gives a contradiction. Let B be a set of five regions from $r_2(v, w), \dots, r_{\ell-2}(v, w)$ that are colored black or black-and-white. Let $Z(B)$ be a set of vertices containing exactly one vertex $z(r)$ from each region $r \in B$ such that $z(r)$ is a neighbor of v but not a neighbor of w . Note that any inner vertex of $R(v, w)$ is adjacent to at most two vertices of $Z(B)$. Thus at most four of the five chosen vertices can be dominated by the inner vertices of S . Furthermore, observe that the vertices of $Z(B)$ are not adjacent to any boundary vertices of $R(v, w)$ except v . However, S does not contain v and therefore, it follows that at least one vertex in $Z(B)$ is not dominated by S . The proof for the case when c has white weight at least 7 is similar. \square

Case 1.1: The coloring c has black weight at least 8 and white weight at least 8.

Let S be a minimum connected dominating set. Note that S must contain v and w , by Observation 1. Now apply Lemma 1 to turn the induced subgraphs $G[N(v)]$ and $G[N(w)]$ into independent sets.

Lemma 4 (reduce) Let $G = (V, E)$ be a plane graph and let $R(v, w)$ be a region between $v, w \in V$. Suppose that coloring c has black weight at least 8 and white weight at least 8. Let y be an inner vertex of a black or black-and-white region such that y is a neighbor of v but not a neighbor of w . Then $\gamma_c(G) = \gamma_c(G - y)$.

Proof: Let S be a minimum connected dominating set of G . By Observation 1, we can choose S such that $v, w \in S$. If $y \in S$ then let $S' = S \setminus \{y\} \cup \{x_1\}$. Note that S' has size at most that of S , and using the fact that $v, w \in S'$ it follows that S' is a dominating set of G . Moreover, the graph induced by S' is connected because $v, w \in S'$ are connected by their common neighbor $x_1 \in S'$. Thus S' is a connected dominating set of G not containing y , and so is a connected dominating set of $G - y$.

Let S' be a minimum connected dominating set of $G' = G - y$. Observe that the coloring c' of G' has black weight at least 7 and white weight at least 7. Thus, by Observation 1, both $v, w \in S'$. Hence the vertex y is dominated in G by $v \in S'$, and it follows that S' is a connected dominating set of G . \square

An analogous reduction rule applies for inner vertices y inside some white or black-and-white region.

Case 1.2: The coloring c has black weight at least 8 and white weight at most 7.

In this case, there exists a region $r(v, w)$ that is colored black.

Lemma 5 (reduce) Let G be a plane graph and let $R(v, w)$ be a region between $v, w \in V$. Suppose that coloring c has black weight at least 8 and white weight at most 7. Let y be an inner vertex of a black region such that y is a neighbor of v but not a neighbor of w . Then $\gamma_c(G) = \gamma_c(G - y)$.

Proof: Let S be a minimum connected dominating set of G . By Observation 1, we can choose S such that $v \in S$. Notice that each neighbor y' of y in G is also a neighbor of v , and thus any such neighbor y' is dominated by v . Hence by minimality of S , it holds $y \notin S$ and thus S is a connected dominating set of $G - y$.

Let S' be a minimum connected dominating set of $G' = G - y$. The coloring c' of G' has black weight at least 7, and thus Observation 1 ensures $v \in S'$. Now y is a neighbor of v in G , and it follows that S' is a connected dominating set of G . \square

The case of coloring c having large white weight and small black weight is similar.

Case 1.3: The coloring c has black weight at most 7 and white weight at most 7.

Lemma 6 (*refine*) *Let $G = (V, E)$ be a plane graph and let $R(v, w)$ be a region between $v, w \in V$ such that $|N_R(v, w)| \geq 106$. Suppose that coloring c has black weight at most 7 and white weight at most 7. Then there exists a region $R'(v, w)$ such that $|N_{R'}(v, w)| \geq 8$ and containing only transparent regions from $\{r_1(v, w), \dots, r_{\ell-1}(v, w)\}$.*

Proof: For all $i = 1, \dots, \ell - 1$, let $m(i)$ be the largest integer such that $r^{(i)} = (r_i(v, w), r_{i+1}(v, w), \dots, r_{i+m(i)}(v, w))$ is a sequence of consecutive regions that are all colored transparent. Let m^* be the maximum over all $m(i)$. Now $|N_R(v, w)| \geq 106$ means that there are at least 105 regions of type $r_i(v, w)$, of which at most 14 are colored either black, white or black-and-white. Hence at least 91 regions are colored transparently; thus $m^* \geq 7$. Let $j \in \{1, \dots, \ell - 1\}$ be such that $m(j) = m^*$. Then $R'(v, w) = r_j(v, w) \cup r_{j+1}(v, w) \cup \dots \cup r_{j+m(j)}(v, w)$ is a region such that $|N_{R'}(v, w)| \geq 8$. \square

Case 1.4: The coloring c colors all regions transparent.

Observation 2 *Let $G = (V, E)$ be a plane graph and let $R(v, w)$ be a region between $v, w \in V$ such that $|N_R(v, w)| \geq 7$. If the coloring c has black weight equal to zero and white weight equal to zero then any minimum connected dominating set of G containing at most two inner vertices of $R(v, w)$ contains at least one of v and w .*

Proof: Suppose for contradiction that S is a minimum connected dominating set of G containing at most two inner vertices of $R(v, w)$ and containing neither v nor w . Notice that for each $i \in \{2, \dots, \ell - 1\}$, vertex x_i can be adjacent to vertices x_{i-1} and x_{i+1} but to no other vertices of $N_R(v, w)$. Also, observe that $x_2, \dots, x_{\ell-1}$ are inner vertices of $R(v, w)$. Since $G[\{x_1, \dots, x_\ell\}]$ is a subgraph of a path and $G[S]$ is connected, it follows that if $x_i \in S$ then either $x_j \in S$ for all $j \in \{2, \dots, i - 1\}$ or $x_j \in S$ for all $j \in \{i + 1, \dots, \ell - 1\}$. Thus, if $x_4 \in S$ then either $\{x_2, x_3, x_4\} \subseteq S$ or $\{x_4, x_5, x_6\} \subseteq S$, a contradiction. Hence either $x_3 \in S$ or $x_5 \in S$. If $x_3 \in S$ then $x_2 \in S$, since $x_4 \notin S$. But then $S \cap \{x_2, \dots, x_6\} = \{x_2, x_3\}$ and x_5 is not dominated by S , giving the desired contradiction. Symmetrically, if $x_5 \in S$ then $x_6 \in S$ and hence x_3 is not dominated by S . \square

Lemma 7 (*reduce*) *Let $G = (V, E)$ be a plane graph and let $R(v, w)$ be a region between $v, w \in V$ such that $|N_R(v, w)| \geq 8$. Suppose that the coloring c has black weight equal to zero and white weight equal to zero. Let y be an inner vertex of $R(v, w)$. Then $\gamma_c(G) = \gamma_c(G - y)$.*

Proof: First, by Lemma 3 there exists a minimum connected dominating set S of G that contains at most two inner vertices of $R(v, w)$. By Observation 2, the set S contains at least one of v and w . It then follows by Lemma 1 that S is a connected dominating set of the graph G' that is obtained from G by removing all edges between vertices of $N_R(v, w)$. Since $R(v, w)$ has at least 6 inner vertices and S contains at most two of them, there exists an inner vertex y' of $R(v, w)$ that is not contained in S . If S does not contain y then S is a connected dominating set of $G - y$. If S contains y then $S \setminus \{y\} \cup \{y'\}$ is a connected dominating set of $G' - y$ and hence also of $G - y$.

Second, Lemma 3 yields that there exists a minimum connected dominating set S' of $G' = G - y$ that contains at most two inner vertices of $R(v, w)$. By Observation 2 it follows that S' contains at least one of v or w . Thus S' dominates y and is a connected dominating set of G . \square

We summarize Case 1.

Lemma 8 (*reduce*) *There is an algorithm that, given a plane graph $G = (V, E)$ and a region $R(v, w)$ between vertices $v, w \in V$ such that $|N(v, w)| \geq 106$, in polynomial time computes a subgraph G' of G with fewer vertices than G such that $\gamma_c(G') = \gamma_c(G)$.*

Proof: The algorithm proceeds as follows. First, it constructs the coloring c of the regions $r_1(v, w), \dots, r_{\ell-1}(v, w)$.

If c has black weight at least 8 and white weight at least 8 then let y be an inner vertex of a black or black-and-white region that is a neighbor of v but not a neighbor of w . Now by Lemma 4, $G' = G - y$ is a subgraph of G with the desired properties.

If c has black weight at least 8 and white weight at most 7 then let y be an inner vertex of a black region and let $G' = G - y$. By Lemma 5 it holds $\gamma_c(G') = \gamma_c(G)$. Proceed similarly if c has black weight at most 7 and white weight at least 8, in which case we let $G' = G - y'$ for some inner vertex y' of a white region.

If c has black weight at most 7 and white weight at most 7 then by Lemma 6 there exists a region $R'(v, w)$ entirely contained in $R(v, w)$ such that any inner vertex of $R'(v, w)$ is a common neighbor of v and w . Thus by Lemma 7, letting y be an inner vertex of $R'(v, w)$ makes $G' = G - y$ a subgraph of G with the desired properties. \square

Case 2: $N_R(v, w)$ contains at most 105 vertices.

Lemma 9 (*refine*) *Let $G = (V, E)$ be a plane graph and let $R(v, w)$ be a region between $v, w \in V$ such that $|N_R(v, w)| \leq 105$. Let m be the number of vertices in $R(v, w)$ other than v and w . Then there is a region $R'(v, w)$ entirely contained in $R(v, w)$ that contains at least $m/104 + 2$ vertices and such that $N_{R'}(v, w) = \emptyset$.*

Proof: Since $|N_R(v, w)| \leq 105$ there are at most 104 regions of type $r_i(v, w)$. Any region of type $r_i(v, w)$ has no inner vertices that are adjacent to both v and w ; thus $N_{r_i}(v, w) = \emptyset$. By the Pigeonhole Principle, one of the at most 104 regions of type $r_i(v, w)$ contains at least $m/104$ vertices other than v and w . \square

Case 3: $N_R(v, w)$ contains no inner vertices of $R(v, w)$.

Let $\mathcal{P} = \{P_1, \dots, P_m\}$ be a maximum-size set of internally vertex-disjoint induced $[v, w]$ -paths of length three entirely contained in $R(v, w)$, such that for all $i = 1, \dots, m - 1$ the vertices of $P_i \cup P_{i+1}$ form the boundary of a region $s_i(v, w)$ not containing vertices from P_j for any $j \notin \{i, i + 1\}$. For each $i = 1, \dots, m$, let v_i be the internal vertex of P_i that is adjacent to v , and let w_i be the internal vertex of P_i that is adjacent to w .

Case 3.1: There are at least 12 internally vertex-disjoint $[v, w]$ -paths of length three.

Observation 3 *Let $G = (V, E)$ be a plane graph and let $R(v, w)$ be a region between $v, w \in V$. If $|\mathcal{P}| \geq 12$ then every minimum connected dominating set of G containing at most two inner vertices of $R(v, w)$ contains both v and w .*

Proof: Observe that v is the only vertex with any neighbors among $\{v_3, \dots, v_9\}$ that is not an inner vertex of $R(v, w)$. Furthermore, every inner vertex of $R(v, w)$ is adjacent to at most 3 vertices among $\{v_3, \dots, v_9\}$. Hence if S is a minimum connected dominating set of G containing at most two inner vertices of $R(v, w)$ and not containing v then S cannot dominate the set $\{v_3, \dots, v_9\}$. Thus S contains v and, by a symmetric argument, S also contains w . \square

Lemma 10 (*reduce*) Let $G = (V, E)$ be a plane graph and let $R(v, w)$ be a region between $v, w \in V$. If $|\mathcal{P}| \geq 12$ then $\gamma_c(G) = \gamma_c(G - v_3)$.

Proof: First, by Lemma 3 there exists a minimum connected dominating set S of G that contains at most two inner vertices of $R(v, w)$. By Observation 3, the set S contains both v and w . If S does not contain v_3 then S is a connected dominating set of $G - v_3$. If S contains v_3 and no other inner vertices of $R(v, w)$ then v_3 is a leaf of $G[S]$, and hence $S \setminus \{v_3\}$ is a connected dominating set of $G - v_3$. If S contains v_3 and some other inner vertex u of $R(v, w)$ then let $S' = S \setminus \{v_3, u\} \cup \{v_2, w_2\}$. Since every vertex in $R(v, w)$ is adjacent to either v or w and since v and w are connected via v_2 and w_2 it follows that S' is a connected dominating set of $G - v_3$.

Second, Lemma 3 implies that there exists a minimum connected dominating set S of $G - v_3$ that contains at most two inner vertices of $R(v, w)$. By Observation 3, the set S contains both v and w . Hence S dominates v_3 and is a connected dominating set of G . \square

Case 3.2: There are at most 11 internally vertex-disjoint $[v, w]$ -paths of length three.

Lemma 11 (*refine*) Let $G = (V, E)$ be a plane graph and let $R(v, w)$ be a region between $v, w \in V$ such that $N_R(v, w)$ contains no inner vertices of $R(v, w)$. Let m be the number of vertices in $R(v, w)$ other than v and w . If $|\mathcal{P}| \leq 11$ then there is a region $R'(v, w)$ entirely contained in $R(v, w)$ that contains at least $m/10 + 2$ vertices such that every $[v, w]$ -path in $R(v, w)$ contains at least one border vertex of $R(v, w)$ except v and w .

Proof: There are at most 10 regions of type $s_i(v, w)$. No region of type $s_i(v, w)$ contains a pair (v_j, w_j) of inner vertices that are internal to some path P_j , for $j \neq i$. Thus, if $v' \in N(v)$ is an inner vertex of $s_i(v, w)$ that is adjacent to w_i or w_{i+1} then $v' \notin \{v_i, v_{i+1}\}$. Symmetrically, if $w' \in N(w)$ is an inner vertex of $s_i(v, w)$ that is adjacent to v_i or v_{i+1} then $w' \notin \{w_i, w_{i+1}\}$. Hence by the Pigeonhole Principle, one of the at most 10 regions of type $s_i(v, w)$ contains at least $m/10$ vertices other than v and w . \square

Case 3.3: There are no internally vertex-disjoint $[v, w]$ -paths of length three.

Lemma 12 (*reduce*) Let $G = (V, E)$ be a plane graph and $R(v, w)$ be a region between $v, w \in V$ with at least 1274 vertices such that every $[v, w]$ -path in $R(v, w)$ contains at least one border vertex of $R(v, w)$ other than v and w . There is a polynomial time algorithm that given G and $R(v, w)$ computes a subgraph G' of G with $|V(G')| < |V(G)|$ and $\gamma_c(G') = \gamma_c(G)$.

Proof: Since the boundary of $R(v, w)$ contains at most 6 vertices, either v or w has at least 634 neighbors that are inner vertices of $R(v, w)$. Without loss of generality, assume that $N(v)$ contains at least 632 inner vertices of $R(v, w)$. If any vertex b in $R(v, w)$ other than v has at least 106 neighbors in common with v inside $R(v, w)$ then there is a region $R(b, v)$ such that $|N_R(b, v)| \geq 106$. A subgraph G' of G with $|V(G')| < |V(G)|$ and $\gamma_c(G') = \gamma_c(G)$ is then obtained by applying Lemma 8 with G and $R(b, v)$.

Suppose now that no such vertex b exists. Then there is a vertex $x \in N(v)$ that is an inner vertex of $R(v, w)$ with no neighbors outside $N(v)$. We claim that $\gamma_c(G - x) = \gamma_c(G)$. In one direction let S be a minimum connected dominating set of G containing at most 2 inner vertices of $R(v, w)$. Every boundary vertex of $R(v, w) \setminus \{v, w\}$ in S dominates at most 105 inner vertices of $R(v, w)$ in $N(v)$. Every inner vertex of $R(v, w)$ in S dominates at most 106 inner vertices of $R(v, w)$ in $N(v)$, since a vertex dominates itself. Thus, if S does not contain v then S dominates at most $4 \cdot 105 + 2 \cdot 106 = 632 < 634$ vertices of $N(v)$ that are inner vertices of $R(v, w)$ yielding a contradiction. Hence, S contains v . Since $N(x) \subseteq N(v)$ and S is a minimum connected dominating set of G , S does not contain x . Thus, S is a connected dominating set of $G - x$.

In the other direction, let S be a minimum connected dominating set of G containing at most 2 inner vertices of $R(v, w)$. By a discussion identical to the one in the previous paragraph, S must contain v . Hence S is a connected dominating set of G . \square

We summarize the case analysis.

Lemma 13 *There is a polynomial time algorithm that, given a plane graph G and a region $R(v, w)$ between vertices $v, w \in V$ with at least 1322672 vertices, computes a subgraph G' of G such that G' has fewer vertices than G and $\gamma_c(G') = \gamma_c(G)$.*

Proof: If $|N_R(v, w)| \geq 106$ then Lemma 8 applied to G and $R(v, w)$ yields the desired subgraph G' . Otherwise, by Lemma 9 there exists a region $R'(v, w)$ with at least 12720 vertices entirely contained in $R(v, w)$ and such that no common neighbours of v and w are inner vertices of $R'(v, w)$. If $R'(v, w)$ contains at least 12 internally vertex-disjoint induced $[v, w]$ -paths of length three then by Lemma 10 then $G' = G - v_3$ is the desired subgraph of G . If $R'(v, w)$ contains at most 11 internally vertex-disjoint induced $[v, w]$ -paths of length three then by Lemma 11 there is a region $R''(v, w)$ such that $R''(v, w)$ contains at least 1274 vertices and there is no $[v, w]$ -path such that all its internal vertices are inner vertices of $R''(v, w)$. In this case Lemma 12 implies that a subgraph G' of G with $V(G') < V(G)$ and $\gamma_c(G') = \gamma_c(G)$ can be computed in polynomial time. \square

We are now in position to give a proof of our main result.

Proof: [of Theorem 1] We prove that the planar CONNECTED DOMINATING SET problem with parameter k admits a kernel of size $3968187k$. We give an algorithm that given an integer k and a plane graph $G = (V, E)$ with at least $3968187k$ vertices, in polynomial time either concludes that $\gamma_c(G) > k$ or computes a subgraph G' of G such that G' has fewer vertices than G and $\gamma_c(G') = \gamma_c(G)$. The algorithm proceeds as follows. First, let $\epsilon = 1/3968186$ and compute a $(1 + \epsilon)$ -approximation S_1 of a minimum connected dominating set for G using the PTAS of Demaine and Hajiaghayi [8]. If S_1 has strictly more than $(1 + \epsilon)k$ vertices then answer $\gamma_c(G) > k$ and stop.

Otherwise, compute a maximal S_1 -region decomposition \mathcal{R} of G via the algorithm by Guo and Niedermeier [15]. By Proposition 1, there are at most $3(1 + \epsilon)k$ regions in \mathcal{R} , and by Proposition 2 at most $170(1 + \epsilon)k$ vertices do not belong to any region in \mathcal{R} . By the pigeonhole principle, there exists a region $R(v, w) \in \mathcal{R}$ between vertices $v, w \in V$ that contains at least 1322672 vertices. Hence by Lemma 13, a subgraph G' of G with fewer vertices than G and $\gamma_c(G') = \gamma_c(G)$ can be computed in polynomial time. \square

4 Conclusion and Consequent Research

In this paper we showed that CONNECTED DOMINATING SET on planar graphs admits a linear kernel. This is the first linear kernel for a “connectivity” problem on planar graphs that does not follow directly from the framework of Guo and Niedermeier [15].

Our algorithm is impractical for two reasons. The first is the huge constant in the kernel size, and the second is the choice of $\epsilon = 1/3968186$ in the PTAS for DOMINATING SET that yields an unmanageable running time. We think that both these problems can be remedied; choosing $\epsilon = 1$ yields a 2-approximation for DOMINATING SET on planar graphs that runs quite quickly, at the cost of a factor 2 in the kernel size. Also, the constant in our kernelization can be improved significantly. In this paper we focused only on showing the existence of a linear kernel and in many places we deliberately picked a proof that yielded a higher constant but was more readable and understandable. It would be interesting to see how far the kernel size can

be reduced. A possible way to attack this problem would be to eliminate the “refine” steps and re-analyzing the cases, taking into account the noise that the “refine” steps removed.

The linear kernel for the DOMINATING SET problem has been extended from planar graphs to graphs excluding a complete bipartite graph $K_{3,h}$ as a minor [2, 17]. A natural question to ask is whether the same can be done for the minimum connected dominating set problem. Finally, it would be interesting to see whether the reduce or refine technique could be applied to achieve this, or to give kernels for other problems.

There have been a lot of new developments in the area of kernelization for problems on planar graphs after the announcement of results in this paper. Gu and Imani [12] significantly reduced the size of our kernel and obtained a kernel of size $413k$ for CONNECTED DOMINATING SET on planar graphs. In another development Bodlaender et al. [3] and Fomin et al. [10] have developed meta theorems for kernelization on sparse graphs and using these theorems we can conclude that CONNECTED DOMINATING SET has a linear kernel on graphs excluding a fixed apex graph H as a minor.

References

- [1] J. Alber, M. R. Fellows, and R. Niedermeier. Polynomial-time data reduction for dominating set. *J. ACM*, 51(3):363–384 (electronic), 2004.
- [2] N. Alon and S. Gutner. Kernels for the dominating set problem on graphs with an excluded minor. Technical report, ECCO Report TR08-066, 2008.
- [3] H. L. Bodlaender, F. V. Fomin, D. Lokshtanov, E. Penninkx, S. Saurabh, and D. M. Thilikos. (Meta) Kernelization. In *FOCS*, pages 629–638, 2009.
- [4] H. L. Bodlaender and E. Penninkx. A linear kernel for planar feedback vertex set. In *IWPEC*, volume 5018 of *Lecture Notes in Comput. Sci.*, pages 160–171. Springer, Berlin, 2008.
- [5] H. L. Bodlaender, E. Penninkx, and R. B. Tan. A linear kernel for the k -disjoint cycle problem on planar graphs. In *ISAAC*, volume 5369 of *Lecture Notes in Comput. Sci.*, pages 306–317. Springer, Berlin, 2008.
- [6] J. Chen, H. Fernau, I. A. Kanj, and G. Xia. Parametric duality and kernelization: lower bounds and upper bounds on kernel size. *SIAM J. Comput.*, 37(4):1077–1106, 2007.
- [7] J. Chen, I. A. Kanj, and W. Jia. Vertex Cover: Further observations and further improvements. *Journal of Algorithms*, 41(2):280–301, 2001.
- [8] E. D. Demaine and M. Hajiaghayi. Bidimensionality: new connections between FPT algorithms and PTASs. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 590–601 (electronic), New York, 2005. ACM.
- [9] F. V. Fomin, F. Grandoni, and D. Kratsch. Solving connected dominating set faster than 2^n . *Algorithmica*, 52(2):153–166, 2008.
- [10] F. V. Fomin, D. Lokshtanov, S. Saurabh, and D. M. Thilikos. Bidimensionality and Kernels. In *SODA*, pages 503–510, 2010.
- [11] M. R. Garey and D. S. Johnson. *Computers and intractability*. W. H. Freeman and Co., San Francisco, Calif., 1979. A guide to the theory of NP-completeness, A Series of Books in the Mathematical Sciences.

- [12] Q. Gu and N. Imani. Connectivity is not a limit for kernelization: Planar connected dominating set. In *LATIN*, volume 6034 of *Lecture Notes in Computer Science*, pages 26–37, 2010.
- [13] S. Guha and S. Khuller. Approximation algorithms for connected dominating sets. *Algorithmica*, 20(4):374–387, 1998.
- [14] J. Guo. Private communication, 2007.
- [15] J. Guo and R. Niedermeier. Linear problem kernels for NP-hard problems on planar graphs. In *ICALP*, volume 4596 of *Lecture Notes in Comput. Sci.*, pages 375–386. Springer, Berlin, 2007.
- [16] J. Guo, R. Niedermeier, and S. Wernicke. Fixed-parameter tractability results for full-degree spanning tree and its dual. In *IWPEC*, volume 4169 of *Lecture Notes in Comput. Sci.*, pages 203–214. Springer, Berlin, 2006.
- [17] S. Gutner. Polynomial kernels and faster algorithms for the dominating set problem on graphs with an excluded minor. In *IWPEC*, volume 5917 of *Lecture Notes in Computer Science*, pages 246–257, 2009.
- [18] H. Moser and S. Sikdar. The parameterized complexity of the induced matching problem in planar graphs. In *FAW*, volume 4613 of *Lecture Notes in Comput. Sci.*, pages 325–336. Springer, Berlin, 2007.
- [19] S. Thomassé. A quadratic kernel for feedback vertex set. In *Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 115–119 (electronic), New York, 2009. ACM.