The effect of girth on the kernelization complexity of Connected Dominating Set

Neeldhara Misra, Geevarghese Philip, Venkatesh Raman, Saket Saurabh

The Institute of Mathematical Sciences, Chennai, India.
{neeldhara,gphilip,vraman,saket}@imsc.res.in

Abstract

In the Connected Dominating Set problem we are given as input a graph G and a positive integer k, and are asked if there is a set S of at most k vertices of G such that S is a dominating set of G and the subgraph induced by S is connected. This is a basic connectivity problem that is known to be NP-complete, and it has been extensively studied using several algorithmic approaches. In this paper we study the effect of excluding short cycles, as a subgraph, on the *kernelization complexity* of Connected Dominating Set.

Kernelization algorithms are polynomial-time algorithms that take an input and a positive integer k (the *parameter*) and output an equivalent instance where the size of the new instance and the new parameter are both bounded by some function $g(k)$. The new instance is called a $g(k)$ *kernel* for the problem. If $g(k)$ is a polynomial in k then we say that the problem admits polynomial kernels. The girth of a graph G is the length of a shortest cycle in G. It turns out that Connected Dominating Set is "hard" on graphs with small cycles, and becomes progressively easier as the girth increases. More specifically, we obtain the following interesting trichotomy: Connected Dominating Set

- does not have a kernel of *any* size on graphs of girth 3 or 4 (since the problem is W[2]-hard);
- admits a $g(k)$ kernel, where $g(k)$ is $k^{\mathcal{O}(k)}$, on graphs of girth 5 or 6 but has *no* polynomial kernel (unless the Polynomial Hierarchy (PH) collapses to the third level) on these graphs;
- has a cubic ($\mathcal{O}(k^3)$) kernel on graphs of girth at least 7.

While there is a large and growing collection of parameterized complexity results available for problems on graph classes characterized by excluded *minors*, our results add to the very few known in the field for graph classes characterized by excluded *subgraphs*.

Digital Object Identifier 10.4230/LIPIcs.xxx.yyy.p

1 Introduction

In the Dominating Set (DS) problem, we are given a graph G and a non-negative integer k, and the question is whether G contains a set of k vertices whose closed neighborhood contains all the vertices of G. In the connected variant Connected Dominating Set (CDS), we also demand that the subgraph induced by the dominating set be connected. DS and CDS, together with their numerous variants, are two of the most well-studied problems in algorithms and combinatorics [22]. A significant part of the algorithmic study of these NP-complete problems has focused on the design of parameterized algorithms. Informally, a *parameterization* of a problem assigns an integer k to each input instance and a parameterized problem is *fixed-parameter tractable* (FPT) if there is an algorithm that solves the problem in time $f(k) \cdot |I|^{O(1)}$, where $|I|$ is the size of the input and f is an arbitrary computable function that depends only on the parameter k. CDS is W[2]-complete on general graphs and therefore it cannot be solved by a parameterized algorithm, unless an unlikely collapse occurs in the W hierarchy (see [15, 16, 27]). However, there are interesting graph classes where FPT algorithms do exist for the Dominating Set problem. The project of widening the

Conference title on which this volume is based on.
Editors: Billy Editor, Bill Editors; pp. 1–12
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

horizon where such algorithms exist spawned a multitude of ideas that made DS and CDS the testbed for some of the most cutting-edge techniques of parameterized algorithm design. For example, the initial study of parameterized subexponential algorithms for DS on planar graphs [1, 10, 18] resulted in the creation of bidimensionality theory which characterizes a broad range of graph problems that admit efficient approximation schemes and/or fixed-parameter algorithms on a broad range of graphs [11, 12, 14].

Kernelization is a rapidly growing sub-area of parameterized complexity. A parameterized problem is said to admit a *polynomial kernel* if there is a polynomial time algorithm, called a *kernelization* algorithm, that reduces the input instance down to an instance with size bounded by a polynomial $p(k)$ in the parameter k, while preserving the answer. This reduced instance is called a $p(k)$ *kernel* for the problem. If $p(k) = O(k)$, then we call it a *linear kernel*. One of the first results on linear kernels is the celebrated work of Alber, Fellows, and Niedermeier on DS, on planar graphs [2]. This work spurred the interest to prove polynomial (preferably linear) kernels for other parameterized problems. The result from [2] (see also [8]) has been extended to much more general graph classes. More recently, Bodlaender et al. [5] and Fomin et al. [17] obtained algorithmic meta-kernelization results which show that a multitude of problems expressible in a certain logic (or are bidimensional) admit linear kernels on (apex) H-minor free graphs.

Most of the kernelization results mentioned above are on graph classes excluding a fixed graph as a *minor*. While there have been a lot of results obtained in the realm of parameterized algorithms on graph classes excluding some graph as a minor, there have only been a handful of such results on graph classes that are defined by excluding a fixed graph as a *subgraph*. The first result of this kind was obtained by Raman and Saurabh [29] who showed that DS and several of its variants are FPT on any class of graphs that forbids “short” cycles — cycles of length 4. This can equivalently be thought of as excluding a $K_{2,2}$, the complete bipartite graph where each part has size exactly 2. Philip et al. [28] generalized this result and showed that DS remains FPT on $K_{i,j}$-free graphs for any fixed i and j, and in fact has a polynomial kernel of size k^h where h is a constant that depends on i and j. It is a corollary of this result that the DS problem has polynomial kernels on graphs of bounded degeneracy – a class which includes graphs defined by excluding a fixed graph H as a minor. Alon and Gutner had shown previously that DS has a kernel of size $O(k^h)$ on H-minor free graphs, where the constant h depends on the excluded graph H [3, 21]. $K_{i,j}$-free graphs remain the largest class of graphs for which DS is currently known to have a polynomial sized kernel and is fixed-parameter tractable.

In this paper, we study the effect of girth on the kernelization complexity of CDS. Typically the parameterized (or other) complexity of connected variants of a problem tend to be much more than that of the problem itself. For example, Vertex Cover has a $2k$-sized vertex kernel and an efficient fixed-parameter tractable algorithm [27], and its connected variant is known not to have a polynomial sized kernel unless the Polynomial Hierarchy collapses to the third level(which is widely believed to be unlikely) [13]. Similarly, while Feedback Vertex Set has an $O^*(3.83^k)$ FPT algorithm [7], the best known FPT algorithm for its connected variant has an $O^*(c^k)$ running time [25] where c is more than 23.

The parameterized complexity of CDS has been extensively investigated, and many results are known. Thus, it is known that CDS is W[2]-hard on general graphs [15], has a linear kernel on planar, or more generally, on apex-minor-free graphs [17, 20, 24], and is FPT on graphs of bounded degeneracy [19]. CDS is also unlikely to have polynomial sized kernels on graphs of bounded degeneracy [9]. We obtain the complete kernelization complexity landscape for the CDS problem based on the girth of the problem instance. More precisely, we show that

Connected Dominating Set

1. is W[2]-hard on graphs of girth 3 or 4, and hence does not have a kernel of *any* size on these graphs unless $FPT = W[2]$;
2. has an FPT algorithm that runs in time $k^{\mathcal{O}(k)}n^{\mathcal{O}(1)}$ on graphs of girth 5 or 6, and hence has a kernel of size $k^{\mathcal{O}(k)}$ on these graphs;
3. has *no* polynomial kernel (unless PH collapses to the third level) on graphs of girth 5 or 6, and,
4. has a cubic ($\mathcal{O}(k^3)$) kernel on graphs of girth at least 7.

The first result follows directly from a construction in [29], and the second and fourth results are obtained using nontrivial extensions of techniques from [29]. The main technical contribution of this paper is the third result, to obtain which we introduce an intermediate, seemingly unrelated problem (FCC), show that FCC has no polynomial kernels (unless PH collapses to the third level) using the recent kernel lower bound machinery developed by Bodlaender et al. [4], and then provide a parameter-preserving reduction [6] from FCC to CDS.

2 Preliminaries

We use $V(G)$ and to $E(G)$ denote, respectively, the vertex and edge sets of graph G. A graph H is a *subgraph* of G if $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$. The subgraph H is called an *induced subgraph* (induced by the vertex set $V(H)$) of G if $E(H) = \{\{u,v\} \in E(G) \mid u,v \in V(H)\}$. For a subset $S \subseteq V(G)$ the subgraph of G induced by S is denoted by $G[S]$, and we use $G \setminus S$ to denote the subgraph induced by $V(G) \setminus S$. The *open-neighborhood* of a vertex v in G, denoted $N(v)$, is the set of all vertices that are adjacent to v in G. The elements of $N(v)$ are said to be the *neighbors* of v, and $N[v] = N(v) \cup \{v\}$ is called the *closed neighborhood* of v. For a set of vertices $X \subseteq V(G)$, the open and closed neighborhoods of X are defined, respectively, as $N(X) = \bigcup_{u \in X} N(u) \setminus X$ and $N[X] = N(X) \cup X$. A vertex $v \in V(G)$ is said to be a *pendant* vertex of G if $|N(v)| = 1$. The girth of a graph is the size (number of vertices) of the smallest cycle in the graph. We use $\mathbb{G}_r$ to denote the class of all graphs with girth *at least* $r \in \mathbb{N}$.

A *dominating set* of graph G is a vertex-subset $S \subseteq V(G)$ such that for each $u \in V(G) \setminus S$ there exists $v \in S$ such that $\{u,v\} \in E(G)$. Given a graph G and $A, B \subseteq V(G)$, we say that A dominates B if every vertex in $B \setminus A$ is adjacent in G to some vertex in A. A *connected dominating set* of a graph $G = (V,E)$ is a set $S \subseteq V$ of vertices of G such that $G[S]$ is connected and S is a dominating set of G. To describe the running times of algorithms we sometimes use the $\mathcal{O}^*$ notation. The $\mathcal{O}^*$ notation suppresses polynomial factors in the expression.

A *parameterized problem* Π is a subset of $\Gamma^* \times \mathbb{N}$, where Γ is a finite alphabet. An instance of a parameterized problem is a tuple (x,k), where k is called the parameter. A central notion in parameterized complexity is *fixed-parameter tractability (FPT)* which means, for a given instance (x,k), decidability in time $\mathcal{O}(f(k) \cdot p(|x|))$, where f is an arbitrary function of k and p is a polynomial. The notion of *kernelization* is formally defined as follows.

▶ **Definition 1.** **[Kernelization, Kernel]** [16, 27] A kernelization algorithm for a parameterized problem $\Pi \subseteq \Sigma^* \times \mathbb{N}$ is an algorithm that, given $(x,k) \in \Sigma^* \times \mathbb{N}$, outputs, in time polynomial in $|x|+k$, a pair $(x',k') \in \Sigma^* \times \mathbb{N}$ such that (1) $(x,k) \in \Pi$ if and only if $(x',k') \in \Pi$ and (2) $|x'|, k' \leq g(k)$, where g is some computable function. The output instance x' is called the *kernel*, and the function g is called the size of the kernel. If $g(k) = k^{O(1)}$ then we say that Π admits a polynomial kernel.

3 On Graphs of Girth 3 and 4 : W[2]-Hardness

In [29, Theorem 1], it is shown that the closely related DS problem is W[2]-hard on graphs of girth 4. Their construction, reproduced below for completeness, suffices to show that CDS is W[2]-hard on graphs of girth 4:

▶ **Theorem 2.** CDS *is* W[2]-hard *on graphs of girth* 3 *and on graphs of girth* 4.

Proof. Given an instance (G, k) of DS, we construct a bipartite graph H. We take two copies of $V(G)$ call it $V_1 = \{u_1 \mid u \in V(G)\}$ and $V_2 = \{u_2 \mid u \in V(G)\}$. If there is an edge $\{u, v\}$ in E, then we add the edges $\{u_1, v_2\}$ and $\{v_1, u_2\}$ to H. We also include edges of the form $\{u_1, u_2\}$ for each $u \in V(G)$. We create two new vertices $z_1 \in V_1$ and $z_2 \in V_2$, and add an edge from every vertex in V_1 to z_2. This completes the construction of H.

The girth of the reduced instance H is at least 4 because H is bipartite, and H has girth exactly 4 because the reduction takes an edge in the original instance G to a cycle of length 4 in H. If G has a dominating set S of size at most k, then $S_1 = \{s_1 \in V_1 \mid s \in S\}$ and the vertex z_2 together form a connected dominating set of H of size at most $k+1$. For the reverse direction, observe that z_2 is present in any minimal connected dominating set of H. If D' is a connected dominating set of H of size at most $k+1$, then let $D = \{u \mid u \in V(G), u_1 \textit{ or } u_2 \in D'\}$. It can easily be shown that D forms a dominating set of G of size at most k. It follows that the CDS problem restricted to graphs of girth 4 is W[2]-hard.

To see that CDS is W[2]-hard on graphs of girth 3 as well, add a new vertex z_3 and the two edges $\{z_2, z_3\}, \{z_1, z_3\}$ to H to form a triangle so that H has girth 3. The reduced instance is $(H, k+1)$. Essentially the same argument as above shows that this reduction is sound. ◀

4 On Graphs of girth 5 or More: $k^{\mathcal{O}(k)}$ kernel or FPT

We now show that the CDS problem restricted to $\mathbb{G}_5$ is FPT with an algorithm that runs in time $k^{\mathcal{O}(k)} n^{\mathcal{O}(1)}$. A folklore theorem of parameterized complexity [27] states that for any computable function f, if a parameterized problem has an FPT algorithm that runs in time $f(k)n^{O(1)}$ on inputs of size n and parameter k, then the problem has a kernel of size $f(k)$. It follows that CDS restricted to $\mathbb{G}_5$ has a kernel of size $k^{\mathcal{O}(k)}$. We show first that a slightly more general problem is FPT on $\mathbb{G}_5$. Following [29], we define the Connected RWB-Dominating Set (ColCDS) problem as:

Connected RWB-Dominating Set(ColCDS)

Input:	A graph $G = (V, E)$, and a positive integer k. The vertex set of G is partitioned into three sets R, W, B of red, white, and blue vertices, respectively. In addition, G has the following properties: (a) G has girth at least 5; (b) every white vertex is the neighbor of some red vertex; (c) blue vertices have no red neighbors; and (d) $\lvert R\rvert \leq k$.
Parameter:	k
Question:	Does G have a connected dominating set of size at most k that contains all the red vertices?

The semantics of the colors are similar to those in [29]: A red vertex is one which is definitely present in the connected dominating set D that our algorithm is trying to construct. A white vertex is one that is not yet in D but is known to be dominated by some vertex in D. All the remaining vertices are those yet to be dominated and are colored blue.

We note that it is claimed in [29, Corollary 3] that `CDS` restricted to $\mathbb{G}_5$ has a kernel on $\mathcal{O}(k^3)$ vertices, and hence is fixed-parameter tractable. But the argument that they present is incorrect; in fact, as we show later (Theorem 13), `CDS` restricted to $\mathbb{G}_5$ *cannot* have any polynomial-sized kernel unless the Polynomial Hierarchy collapses to the third level. The error in their argument is that they assume that the reduction rules they used for `DS` also work for `CDS` — but rules like deleting a white vertex and edges between white vertices *do not* apply to `CDS`. This is because such vertices and edges may be needed to provide connectivity to a dominating set. However, the fixed-parameter tractability result still holds, as we prove by a different argument in the following lemma.

▶ **Lemma 3.** *`ColCDS` is FPT.*

Observe that once we have Lemma 3, we can solve the `CDS` problem on $\mathbb{G}_5$ by simply coloring all vertices blue and then solving the `ColCDS` problem using Lemma 3. Hence we have

▶ **Theorem 4.** *`CDS` is FPT on graphs of girth at least* 5*.*

Let (G, k) be an instance of `ColCDS`. If a vertex v in G has more than k neighbors, and v is not in a dominating set S of G of size at most k, then there is a vertex $u \in S$ that dominates at least two vertices $x, y \in N(v)$. Then u, v, x, y form a cycle of length at most 4, a contradiction. So we have:

▶ **Lemma 5.** *Let (G, k) be an instance of `ColCDS`. If a vertex v in G has more than k neighbors, then v is present in every dominating set of G of size at most k.*

Proof of Lemma 3. Let (G, k) be an instance of `ColCDS` and S be the set of white and blue vertices in G that have at least $k + 1$ neighbors. By Lemma 5 we know that every vertex of S is part of every dominating set of G size at most k whether connected or otherwise. Thus if $|R \cup S| > k$ then G does not have any connected dominating set of size at most k that contains all the vertices of R and hence we return NO. So we assume that $|R \cup S| \leq k$.

We first obtain an equivalent instance of `ColCDS` by coloring all the vertices of S red and all its blue neighbors white. Now we bound the size of the set B. Observe that in the equivalent instance every blue or white vertex has at most k neighbors and no red vertex has any blue neighbor. Thus the remaining $k' = k - |R|$ white and blue vertices can only dominate at most $k'(k + 1)$ blue vertices and hence $|B| \leq k^2 + k$ if (G, k) is a YES instance of the problem. So if $|B| > k^2 + k$, then we return NO.

Let W' be the set of white vertices that are neighbors to blue vertices. From Lemma 5, $|W'| \leq |B|k \leq k^3 + k^2$. Observe that every *connected dominating set* D of G of size at most k containing all the red vertices contains a *minimal dominating set* D' of size at most k such that $D' \subseteq B \cup W' \cup R$. This is because all the neighbors of B are in W'. We use this property to check whether G has a connected dominating set D of size at most k that contains all the red vertices. We enumerate all the minimal dominating sets D' of G of size at most k such that $R \subseteq D' \subseteq B \cup W' \cup R$. Given such a set D', we only need to check whether we can make it connected by adding at most $k - |D'|$ vertices. To do so we use an algorithm for the Steiner Tree problem. In the Steiner Tree problem we are given a graph G and a subset T of the vertex set called the terminal set, and the objective is to find a smallest set of vertices $N \subseteq V(G) \setminus T$ such that $G[T \cup N]$ is connected. Nederlof [26] gave an algorithm for Steiner Tree that runs in time $2^t n^{\mathcal{O}(1)}$ where $t = |T|$. Given D' we use this algorithm and check whether we can make D' connected by adding at most $k - |D'|$ vertices. If there is at least one D' such that we can connect it by adding at most $k - |D'|$ vertices, then we return YES, else we return NO. Note that $\ell = |B \cup W' \cup R| \leq (k^2 + k) + (k^3 + k^2) + k = \mathcal{O}(k^3)$.

Thus the running time of our algorithm is bounded by $\mathcal{O}^*(\sum_{i=|R|}^{k} \binom{\ell}{i} \cdot 2^i) = \mathcal{O}^*(2^k k^{3k})$. This concludes the proof of theorem. ◀

5 On Graphs of girth 5 and 6: No Polynomial Kernels

In the last section we saw that CDS is FPT on graphs with girth at least 5, with an algorithm of running time $k^{\mathcal{O}(k)} n^{\mathcal{O}(1)}$. This immediately implies that the problem has a kernel of size $k^{\mathcal{O}(k)}$[27]. A natural question to ask is whether CDS has polynomial kernels on these graph classes. We now show that the Connected Dominating Set problem restricted to graphs of girth 5 or 6 does not have a polynomial kernel unless the Polynomial Hierarchy collapses to the third level.

5.1 Known Lower Bound Machinery

To prove our lower bound, we need a few notions and results from the recently developed theory of kernel lower bounds [4, 6, 13]. We use a notion of reductions, similar in spirit to those used in classical complexity to show NP-hardness results, to show this kernelization lower bound. We recall the required definitions and theorems:

▶ **Definition 6.** **[Derived Classical Problem]** [6] Let $\Pi \subseteq \Sigma^* \times \mathbb{N}$ be a parameterized problem, and let $1 \notin \Sigma$ be a new symbol. We define the *derived classical problem* associated with Π to be $\left\{x1^k | \, (x,k) \in \Pi\right\}$.

▶ **Definition 7.** **[Composition Algorithm, Compositional Problem]** [4] A *composition algorithm* for a parameterized problem $\Pi \subseteq \Sigma^* \times \mathbb{N}$ is an algorithm that takes as input a sequence $\langle (x_1, k), (x_2, k), \ldots, (x_t, k) \rangle$ where each $(x_i, k) \in \Sigma^* \times \mathbb{N}$, runs in time polynomial in $\sum_{i=1}^{t} |x_i| + k$, and outputs an instance $(y, k') \in \Sigma^* \times \mathbb{N}$ where $(y, k') \in L \iff (x_i, k) \in L$ for some $1 \leq i \leq t$, and k' is polynomial in k. We say that a parameterized problem is *compositional* if it has a composition algorithm.

▶ **Theorem 8.** [4, Lemmas 1 and 2] *Let L be a compositional parameterized problem whose derived classical problem is NP-complete. If L has a polynomial kernel, then the Polynomial Hierarchy collapses to the third level.*

▶ **Definition 9.** [6] Let P and Q be parameterized problems. We say that P is polynomial parameter reducible to Q, written $P \leq_{ppt} Q$, if there exists a polynomial time computable function $f : \Sigma^* \times \mathbb{N} \to \Sigma^* \times \mathbb{N}$, and a polynomial $p : \mathbb{N} \to \mathbb{N}$, and for all $x \in \Sigma^*$ and $k \in \mathbb{N}$, if $f((x, k)) = (x', k')$, then $(x, k) \in P$ if and only if $(x', k') \in Q$, and $k' \leq p(k)$. We call f a polynomial parameter transformation (or a PPT) from P to Q.

▶ **Theorem 10.** [6, Theorem 3] *Let P and Q be parameterized problems whose derived classical problems are P^c, Q^c, respectively. Let P^c be NP-complete, and $Q^c \in$ NP. Suppose there exists a PPT from P to Q. Then, if Q has a polynomial kernel, then P also has a polynomial kernel.*

5.2 Kernel lower bounds

We begin our reductions by defining the Fair Connected Colors problem, which is a variant of the Connected Colors problem recently introduced by Cygan et al. [9]:

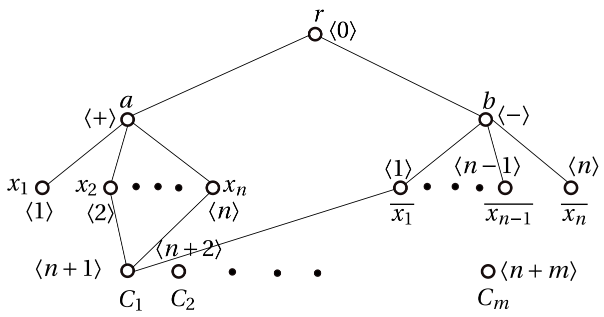

Figure 1 Reduction from CNF SAT to Fair Connected Colors. The color of each vertex is indicated within angled brackets.

Fair Connected Colors

Input:	A graph G, where the vertices $V(G)$ are *properly* colored with k colors in such a way that all neighbors of each vertex have distinct colors.
Parameter:	k
Question:	Does G contain a tree T on k vertices as a subgraph, where each vertex of T has a distinct color?

This problem differs from Connected Colors in that for Connected Colors, the given graph is *arbitrarily* colored with k colors. For Fair Connected Colors we restrict the coloring to be proper and fair (all neighbors of a vertex get different colors) as we need this restriction for the reduction we give in Theorem 13.

▶ **Lemma 11.** *The* Fair Connected Colors *problem is* NP*-complete.*

Proof. A tree on k vertices with all its vertices colored with distinct colors is a polynomial-time verifiable witness to a YES-instance of the problem, and so Fair Connected Colors is in NP. To show hardness, we reduce from the NP-complete CNF SAT problem [23]. Let ϕ be a Boolean formula in CNF on the variables $x_1, \ldots, x_n$ and clauses $C_1, \ldots, C_m$. We assume without loss of generality that there is no clause that contains both a variable and its negation. We construct a graph G on $m + 2n + 3$ vertices colored using $m + n + 3$ colors as follows: We define the vertex set to be $V(G) := \{r, a, b, x_1, \ldots, x_n, \overline{x_1}, \ldots, \overline{x_n}, C_1, \ldots, C_m\}$. We add the edges $\{r, a\}, \{r, b\}$ and $\{a, x_1\}, \{a, x_2\}, \ldots, \{a, x_n\}, \{b, \overline{x_1}\}, \{b, \overline{x_2}\}, \ldots, \{b, \overline{x_n}\}$; and for each vertex C_i, we add an edge from C_i to vertex $y \in \{x_1, \ldots, x_n, \overline{x_1}, \ldots, \overline{x_n}\}$ if and only if the literal y appears in clause C_i in the formula ϕ. This completes the construction of the graph G. We assign the colors $0, +, -$ to vertices r, a, b, respectively. For $1 \leq i \leq n$, we assign color i to vertices x_i and $\overline{x_i}$, and for $1 \leq j \leq m$, we assign color $n + j$ to vertex C_j. This completes the construction; see Figure 1.

Note that the vertices of G are properly colored with $n + m + 3$ colors in such a way that no vertex v is adjacent to two other vertices u, w where u and w are of the same color. The instance of Fair Connected Colors is $(G, n + m + 3)$. It remains to show that ϕ is satisfiable if and only if G contains an $m + n + 3$-vertex tree as a subgraph whose vertices are all colored distinctly.

Suppose ϕ is satisfiable, and let S be the set of literals (negative as well as positive) that are set to true by a satisfying assignment A of ϕ. Notice that A sets at least one literal in each clause of ϕ to true. Also, for each variable x_i, A sets exactly one of $x_i, \overline{x_i}$ to true. Thus each vertex $C_i; 1 \leq i \leq m$ is adjacent to at least one of vertex in S, and S contains exactly one vertex with each of the colors $\{1, 2, \ldots, n\}$. It follows that the subgraph H of G induced on the vertex set $\{r, a, b, C_1, C_2, \ldots, C_m\} \cup S$ is connected and has one vertex from each of the $n+m+3$ colors $\{0,+,-,1,2,\ldots,n+m\}$. Therefore G contains an $m+n+3$-vertex tree as a subgraph whose vertices are all colored distinctly: indeed, any spanning tree of H serves as a witness.

Now suppose G contains an $m+n+3$-vertex tree T as a subgraph whose vertices are all colored distinctly. Then the vertex set $V(T)$ of T must consist of $\{r,a,b,C_1,\ldots,C_m\}$, and exactly n vertices from the set $X=\cup_{i=1}^{n}\{x_i,\overline{x_i}\}$ where exactly one vertex is chosen from $\{x_i,\overline{x_i}\};1\le i\le n$. The unique path from any vertex $C_i;1\le i\le n$ to r in T must use a vertex in $S=X\cap V(T)$. Consider the assignment A of the formula ϕ which sets to true exactly those literals that appear in S. Since $|S\cap\{x_i,\overline{x_i}\}|=1$ for $1\le i\le n$, A is a valid assignment. Since each vertex C_i is adjacent to at least one vertex in S, the assignment satisfies every clause in ϕ, and so ϕ is satisfiable. ◀

The Fair Connected Colors problem is easily seen to be compositional: taking the disjoint union of input graphs suffices for the composition. That is, given k colored graph $G_1,\ldots,G_t$, return $\cup_{i=1}^{t}G_i$ and k. Hence from the Lemma 11 and Theorem 8 we have:

▶ **Lemma 12.** *The* Fair Connected Colors *problem does not have a polynomial kernel unless the Polynomial Hierarchy collapses to the third level.*

We now prove our main result by giving a polynomial parameter transformation (PPT) from Fair Connected Colors to CDS on graphs with graph 5 or 6.

▶ **Theorem 13.** *The CDS problem restricted to graphs of girth* 5 *or* 6 *does not admit a polynomial kernel unless the Polynomial Hierarchy collapses to the third level.*

Proof. Note that by Theorem 10 and Lemma 12 it is sufficient to show that there is a polynomial parameter transformation (PPT) from Fair Connected Colors to each of these problems. We first describe a PPT from Fair Connected Colors to CDS in graphs of girth six. Given an instance (G,k) of Fair Connected Colors, we construct an instance (H,k') of CDS where H has girth six and k' is bounded by a polynomial in k.

We start with a copy of G. For each color class (set of vertices of the same color) $\mathcal{C}_i$ of G, we add a new vertex v_i adjacent to all vertices of $\mathcal{C}_i$, and a new vertex g_i adjacent to v_i. The vertex g_i is essentially a guard vertex that will force v_i to be selected in our solution. We add a new vertex uv for each edge $\{u,v\}$ of G, and replace the edge $\{u,v\}$ by two new edges $\{u,uv\},\{uv,v\}$. That is, we split each edge of G once. For every two color classes $\mathcal{C}_i,\mathcal{C}_j;i<j$ of G,

1. We add two new vertices v_{ij} and g_{ij} and the edge $\{v_{ij},g_{ij}\}$.
2. For each edge $\{u,v\}$ in G where $u\in\mathcal{C}_i,v\in\mathcal{C}_j$, we add the edge $\{uv,v_{ij}\}$ where uv is the new vertex that splits $\{u,v\}$.
3. For each vertex $u\in\mathcal{C}_i$ that has no neighbor in $\mathcal{C}_j$, we add a new vertex u_{ij} and the edges $\{u,u_{ij}\},\{u_{ij},v_{ij}\}$ where v_{ij} is the vertex added in step 1.
4. Symmetrically, for each vertex $u\in\mathcal{C}_j$ that has no neighbor in $\mathcal{C}_i$, we add a new vertex u_{ji} and the edges $\{u,u_{ji}\},\{u_{ji},v_{ij}\}$.

This completes the construction of H; see Figure 2. For later reference, let S be the set of vertices of the form uv introduced in H to split the edges of G, $C=\mathcal{C}_1\cup\cdots\cup\mathcal{C}_k$, $X=\{g_i;1\le i\le k\}$, $Y=\{v_{ij}\in V(H)\}$, $Z=\{v_1,v_2,\ldots,v_k\}$, $W=\{g_{ij};1\le i<j\le_k\}$, and let U be the set of all new vertices added in steps (3) and (4) above.

Observe that H is bipartite, with one part being $A=C\cup X\cup Y$. Hence every cycle in H is of even length, and the smallest cycle has length at least 4. Also, H contains a 4-cycle if and only if there are two vertices in A which have two common neighbors in $V(H)\setminus A$.

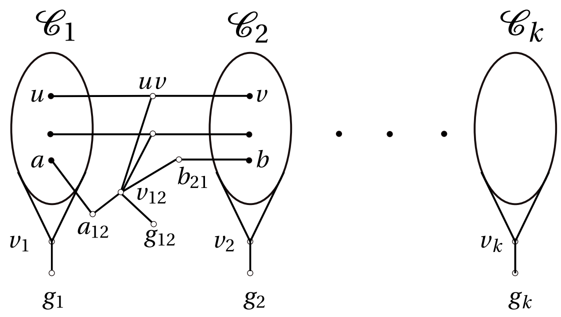

Figure 2 Reduction from Fair Connected Colors to Connected Dominating Set.

But no two vertices in A can have two common neighbors: the vertices in X are all of degree exactly one, and so they are not part of any cycle, and in each of the remaining ways of forming a pair a, b of vertices from A, it is easy to verify that a and b have at most one common neighbor. It follows that H does not contain a 4-cycle, and so the smallest cycle in H has length at least 6. To see that the girth of H is indeed 6, note that we can assume without loss of generality that $\mathcal{C}_1$ contains at least two vertices, say a, b. Observe that there is a path of length two from a to v_{12}, and a path of length two from v_{12} to b. These paths meet only at v_{12}, and together with the two edges $\{b, v_1\}, \{v_1, a\}$ they form a cycle of length 6. Thus let $(H, k^2 + k)$ be the reduced instance. Now we argue that the reduction is indeed sound.

Forward direction. Suppose G contains a tree T on k vertices, where each vertex of T has a distinct color. Let $V(T) = \{t_1, t_2, \ldots, t_k\}$, where $t_i \in \mathcal{C}_i$ for all i. Let T' be the "corresponding" tree in H: the vertex set of T' consists of $V(T)$ and all the new vertices in H that split the edges of T, and the edge set consists of all the new edges formed by splitting the edges of T. Thus T' is a tree on $2k - 1$ vertices. We now add more vertices and edges to T' to obtain a tree on $k^2 + k$ vertices that dominates all of H.

- For $1 \leq i \leq k$, we add the vertex v_i and the edge $\{v_i, t_i\}$ to T'. This adds k vertices.
- For $1 \leq i < j \leq k$, if the vertex $t_i t_j$ is present in T', then we add the vertex v_{ij} and the edge $\{t_i t_j, v_{ij}\}$ to T'. This adds $k - 1$ vertices to T'. Otherwise, let $a = t_i$. We add the vertices a_{ij}, v_{ij} and the edges $\{a, a_{ij}\}, \{a_{ij}, v_{ij}\}$ to T'. This adds two vertices for each "non-edge" in T, for a total of $2(\binom{k}{2} - (k - 1))$ new vertices added to T'.

This completes the construction of T'. Note that T' is a tree on $4k - 2 + 2(\binom{k}{2} - (k - 1)) = k^2 + k$ vertices. In H, (1) the set $\{v_i \mid 1 \leq i \leq k\} \subseteq V(T')$ dominates all the vertices copied over from G, and the new vertices $\{g_1, \ldots, g_k\}$, and (2) the set $\{v_{ij} \mid 1 \leq i < j \leq k\} \subseteq V(T')$ dominates all the other newly added vertices. Thus T' is a connected dominating set of H on $k^2 + k$ vertices.

Reverse direction. Let D be a minimal connected dominating set of H with $1 < |D| \leq k^2 + k$. Observe first that vertices in $X \cup W$ are all pendant vertices, and all of their neighbors have degree at least 2. So $N(X \cup W) = (Y \cup Z) \subseteq D$, and since D is minimal, $D \cap (X \cup W) = \emptyset$. Now since $G[D]$ is connected and $|D| \geq 2$, at least one neighbor of each vertex in D must also be in D. Observe that for any two vertices $u, v \in Y \cup Z$, $N[u] \cap N[v] = \emptyset$, and so each vertex in D can be the neighbor of at most one vertex in $Y \cup Z \subseteq D$. Thus for each vertex $v \in Y \cup Z$, D contains at least one distinct vertex $u \in (N(v) \setminus (Y \cup Z))$, and so $|D| \geq 2|Y \cup Z| = 2(\binom{k}{2} + k) = k^2 + k$. But $|D| \leq k^2 + k$ by assumption, and so $|D| = k^2 + k$. Thus exactly one neighbor of each vertex in $Y \cup Z$ is in D. In particular, D contains exactly one vertex from each set $\mathcal{C}_i; 1 \leq i \leq k$. Further, $D = (Y \cup Z) \cup N(Y \cup Z)$.

Let T_1 be a spanning tree of $H[D]$. From the above arguments we see that all vertices in $Y \cup Z$ are leaves in T_1, and so $T_2 = T_1 \setminus (Y \cup Z)$ is also a tree. Observe that all the vertices in $V(T_2) \cap U$ are leaves in T_2, and so $T_3 = T_2 \setminus U$ is also a tree. Observe that T_3 consists of (1) exactly one vertex from each set $\mathcal{C}_i; 1 \leq i \leq k$, and (2) some vertices from the set S. Let T_4 be the tree obtained from T_3 by removing all those vertices in S that are leaves in T_3. Note that each vertex in $R = S \cap V(T_4)$ has degree exactly two in T_4, and no two vertices in R are adjacent in T_4. So the graph T obtained from T_4 by replacing each vertex $u \in R$ with an edge between the two neighbors of u is also a tree. From the construction, T is (isomorphic to) a subgraph of G. But T is a tree on k vertices where each vertex has a distinct color, and so (G, k) is a YES instance of Fair Connected Colors.

A small modification to the above reduction suffices to show that the Connected Dominating Set problem has no polynomial kernel in graphs of girth 5 as well, unless PH collapses: Add three new vertices a, b, c and the four new edges required to complete the 5-cycle v_1, a, b, c, g_1 so that H has girth 5. The reduced instance is $(H, k^2 + k + 2)$. In the argument to show that this reduction is sound, both the directions go through exactly as before once we observe that exactly one of the sets $\{v_1, a, g_1\}, \{v_1, a, b\}, \{v_1, g_1, c\}$ is contained in any minimal connected dominating set of H. ◀

6 On Graphs of girth 7 or More: A Cubic Kernel

We now show that `CDS` has a cubic kernel on graphs of girth at least 7. As before, our reduction rules color the vertices of G red, white, and blue. Red vertices are those that must necessarily be in any connected dominating set of G of size at most k. White vertices are those non-red vertices that are dominated by the red vertices, and blue vertices are the rest. Initially we color every vertex blue. We have the following four reduction rules.

(R1) Let S be the set of blue vertices in G that have at least $k+1$ blue neighbors. Color all the vertices of S red and all the blue neighbors in $N(S)$ white.

(R2) If $|R| > k$ or $|B| > k^2 + k$, then say NO and stop.

(R3) If G contains an isolated blue vertex, then say NO and stop.

(R4) If G contains a pendant blue or white vertex u adjacent to a vertex v, then remove u from G. If v is not red, then color v red and color all the remaining blue neighbors of v white.

Note that the class $\mathbb{G}_7$ is a subclass of $\mathbb{G}_5$. Hence the correctness of reduction rule (R1) is justified by Lemma 5. The bound obtained on $|B|$ in the proof of Lemma 3 justifies reduction rule (R2). Rule (R3) is justified as we need to include the isolated blue vertex in the dominating set (to dominate that vertex), but as it is isolated the dominating set will not induce a connected graph. Rule (R4) is justified as without loss of generality the vertex v can be in the minimal dominating set we are constructing (as u or v must be in any minimal dominating set to dominate u, and u is a pendant vertex).

From Rule (R2) we have that $|R| \leq k$ and $|B| \leq k^2 + k$. Now using the two additional rules and the fact that G has no cycles of length 5 or 6, we bound $|W|$.

▶ **Lemma 14.** *Let G be reduced with respect to the reduction rules (R1) to (R4) and let (G, k) be a* YES *instance of the* Connected Dominating Set *problem. Then $|W| \leq k^3 + \frac{5}{2}k^2 - \frac{3}{2}k$.*

Proof. We divide W into three parts, $W = W_B \cup W_R \cup W_W$, where

- W_B is the set of all white vertices that have *at least one* blue neighbor,
- W_R is the set of all white vertices in $W \setminus W_B$ that have *only* red neighbors, and
- W_W is the set of all white vertices $W \setminus W_B$ that have *at least one* white neighbor.

We now bound each of these sets.

By rule (R1) we know that any blue vertex v has degree at most k and hence can have at most k white neighbors. Thus $|W_B| \leq k|B| \leq k(k^2 + k)$.

Since G is reduced with respect to rule (R4) each vertex in W_R has at least two red neighbors. From this and the fact that no two vertices have more than one common neighbor, it follows that $|W_R| \leq \binom{|R|}{2} \leq \binom{k}{2}$. Note that we cannot just remove the vertices in W_R from G, since they could be useful in providing connectivity in some smallest connected dominating set.

Let E_W be the set of all edges $e \in E$ where both end vertices of e are white. Each white vertex is adjacent to some red vertex. For any pair of red vertices x, y, there is at most one edge $(u, v) \in E_W$ such that u is adjacent to x and v is adjacent to y. For, if there is another edge $(u', v') \in E_W$ where u' is adjacent to x and v' is adjacent to y, then the vertices x, y, u, v, u', v' form a cycle of length at most 6, a contradiction. It follows that $|E_W| \leq \binom{|R|}{2} \leq \binom{k}{2}$, and so $|W_W| \leq 2|E_W| \leq k^2 - k$.

Putting all the bounds together, if G has a connected dominating set of size at most k, then the number of white vertices in G is at most $k^3 + \frac{5}{2}k^2 - \frac{3}{2}k$. ◀

To obtain an (uncolored) instance of `CDS`, we now attach a new pendant vertex to each red vertex, and remove all colors to obtain an instance (G', k). This last step essentially "forces" all red vertices to be picked in any dominating set of G' of size at most k; it is easy to verify that this step is sound. From Lemma 14 and the bounds $|B| \leq k^2 + k$ and $|R| \leq k$, we get

▶ **Theorem 15.** *The* Connected Dominating Set *problem has a kernel on at most* $k^3 + \frac{7}{2}k^2 + \frac{3k}{2} = \mathcal{O}(k^3)$ *vertices on the class of graphs of girth at least* 7*.*

7 Conclusion

In this paper we studied the effect of excluding short cycles on `CDS` from the kernelization perspective. We obtained a very diverse kernelization landscape. The problem became progressively easier as the size of the girth increased with no kernels to polynomial kernels. It would be interesting to study other problems and excluding some other subgraphs. An interesting problem in this direction is whether `CDS` is FPT on claw-free graphs.

References

1. J. Alber, H. L. Bodlaender, H. Fernau, T. Kloks, and R. Niedermeier. Fixed parameter algorithms for dominating set and related problems on planar graphs. *Algorithmica*, 33(4):461–493, 2002.
2. J. Alber, M. R. Fellows, and R. Niedermeier. Polynomial-time data reduction for Dominating Set. *Journal of the ACM*, 51(3):363–384, 2004.
3. N. Alon and S. Gutner. Kernels for the Dominating Set Problem on Graphs with an Excluded Minor. Technical Report TR08-066, ECCC, 2008.
4. H. L. Bodlaender, R. G. Downey, M. R. Fellows, and D. Hermelin. On problems without polynomial kernels. *Journal of Computer and System Sciences*, 75(8):423–434, 2009.
5. H. L. Bodlaender, F. V. Fomin, D. Lokshtanov, E. Penninkx, S. Saurabh, and D. M. Thilikos. (Meta) Kernelization. In *Proceedings of FOCS 2009*, pages 629–638. IEEE, 2009.
6. H. L. Bodlaender, S. Thomassé, and A. Yeo. Kernel Bounds for Disjoint Cycles and Disjoint Paths. In *Proceedings of ESA 2009*, volume 5757 of *LNCS*, pages 635–646, 2009.

7 Y. Cao, J. Chen, and Y. Liu. On Feedback Vertex Set: New Measure and New Structures. In *Proceedings of SWAT 2010*, volume 6139 of *LNCS*, pages 93–104, 2010.

8 J. Chen, H. Fernau, I. A. Kanj, and G. Xia. Parametric Duality and Kernelization: Lower Bounds and Upper Bounds on Kernel Size. *SIAM Journal on Computing*, 37(4):1077–1106, 2007.

9 M. Cygan, M. Pilipczuk, M. Pilipczuk, and J. Wojtaszczyk. Kernelization hardness of connectivity problems in d-degenerate graphs. Accepted at WG 2010.

10 E. D. Demaine, F. V. Fomin, M. Hajiaghayi, and D. M. Thilikos. Fixed-parameter algorithms for (k, r)-center in planar graphs and map graphs. *ACM Transactions on Algorithms*, 1(1):33–47, 2005.

11 E. D. Demaine, F. V. Fomin, M. Hajiaghayi, and D. M. Thilikos. Subexponential parameterized algorithms on bounded-genus graphs and H-minor-free graphs. *Journal of the ACM*, 52(6):866–893, 2005.

12 E. D. Demaine and M. Hajiaghayi. The Bidimensionality Theory and Its Algorithmic Applications. *The Computer Journal*, 51(3):332–337, 2007.

13 M. Dom, D. Lokshtanov, and S. Saurabh. Incompressibility through Colors and IDs. In *Proceedings of ICALP 2009*, volume 5555 of *LNCS*, pages 378–389. Springer, 2009.

14 F. Dorn, F. V. Fomin, D. Lokshtanov, V. Raman, and S. Saurabh. Beyond Bidimensionality: Parameterized Subexponential Algorithms on Directed Graphs. In *Proceedings of STACS 2010*, volume 5 of *LIPIcs*, pages 251–262, 2010.

15 R. G. Downey and M. R. Fellows. *Parameterized Complexity.* Springer, 1999.

16 J. Flum and M. Grohe. *Parameterized Complexity Theory.* Springer-Verlag, 2006.

17 F. V. Fomin, D. Lokshtanov, S. Saurabh, and D. M. Thilikos. Bidimensionality and Kernels. In *Proceedings of SODA 2010*, pages 503–510, 2010.

18 F. V. Fomin and D. M. Thilikos. Dominating Sets in Planar Graphs: Branch-Width and Exponential Speed-Up. *SIAM Journal on Computing*, 36(2):281–309, 2006.

19 P. A. Golovach and Y. Villanger. Parameterized Complexity for Domination Problems on Degenerate Graphs. In *Proceedings of WG 2008*, volume 5344 of *LNCS*, 2008.

20 Q. Gu and N. Imani. Connectivity Is Not a Limit for Kernelization: Planar Connected Dominating Set. In *Proceedings of LATIN 2010*, volume 6034 of *LNCS*, pages 26–37, 2010.

21 S. Gutner. Polynomial Kernels and Faster Algorithms for the Dominating Set Problem on Graphs with an Excluded Minor. In *Proceedings of IWPEC 2009*, pages 246–257, 2009.

22 T. W. Haynes, S. T. Hedetniemi, and P. J. Slater. *Fundamentals of Domination in Graphs.* CRC Press, 1998.

23 R. M. Karp. Reducibility among combinatorial problems. In *Complexity of Computer Communications*, pages 85–103, 1972.

24 D. Lokshtanov, M. Mnich, and S. Saurabh. Linear Kernel for Planar Connected Dominating Set. In *Proceedings of TAMC 2009*, volume 5532 of *LNCS*, pages 281–290, 2009.

25 N. Misra, G. Philip, V. Raman, S. Saurabh, and S. Sikdar. FPT Algorithms for Connected Feedback Vertex Set. In *Proceedings of WALCOM 2010*, volume 5942 of *LNCS*, pages 269–280, 2010.

26 J. Nederlof. Fast Polynomial-Space Algorithms Using Möbius Inversion: Improving on Steiner Tree and Related Problems. In *Proceedings of ICALP 2009*, 2009.

27 R. Niedermeier. *Invitation to Fixed-Parameter Algorithms.* Oxford University Press, 2006.

28 G. Philip, V. Raman, and S. Sikdar. Solving Dominating Set in Larger Classes of Graphs: FPT Algorithms and Polynomial Kernels. In *Proceedings of ESA 2009*, volume 5757 of *LNCS*, pages 694–705, 2009.

29 V. Raman and S. Saurabh. Short Cycles Make W-hard Problems Hard: FPT Algorithms for W-hard Problems in Graphs with no Short Cycles. *Algorithmica*, 52(2):203–225, 2008.