

# Lossy Kernels for Connected Dominating Set on Sparse Graphs

**Eduard Eiben**

Algorithms and Complexity Group, TU Wien, Austria and  
Department of Informatics, University of Bergen, Norway  
eiben@ac.tuwien.ac.at

**Mithilesh Kumar**

Department of Informatics, University of Bergen, Norway  
mithilesh.kumar@ii.uib.no

**Amer E. Mouawad**

Department of Informatics, University of Bergen, Norway  
a.mouawad@ii.uib.no

**Fahad Panolan**

Department of Informatics, University of Bergen, Norway  
fahad.panolan@ii.uib.no

**Sebastian Siebertz**

Faculty of Mathematics, Informatics and Mechanics, University of Warsaw, Poland  
siebertz@mimuw.edu.pl

---

## Abstract

For  $\alpha > 1$ , an  $\alpha$ -approximate (bi-)kernel for a problem  $\mathcal{Q}$  is a polynomial-time algorithm that takes as input an instance  $(I, k)$  of  $\mathcal{Q}$  and outputs an instance  $(I', k')$  (of a problem  $\mathcal{Q}'$ ) of size bounded by a function of  $k$  such that, for every  $c \geq 1$ , a  $c$ -approximate solution for the new instance can be turned into a  $(c \cdot \alpha)$ -approximate solution of the original instance in polynomial time. This framework of *lossy kernelization* was recently introduced by Lokshtanov et al. We study CONNECTED DOMINATING SET (and its distance- $r$  variant) parameterized by solution size on sparse graph classes like biclique-free graphs, classes of bounded expansion, and nowhere dense classes. We prove that for every  $\alpha > 1$ , CONNECTED DOMINATING SET admits a polynomial-size  $\alpha$ -approximate (bi-)kernel on all the aforementioned classes. Our results are in sharp contrast to the kernelization complexity of CONNECTED DOMINATING SET, which is known to not admit a polynomial kernel even on 2-degenerate graphs and graphs of bounded expansion, unless  $\text{NP} \subseteq \text{coNP/poly}$ . We complement our results by the following conditional lower bound. We show that if a class  $\mathcal{C}$  is somewhere dense and closed under taking subgraphs, then for some value of  $r \in \mathbb{N}$  there cannot exist an  $\alpha$ -approximate bi-kernel for the (CONNECTED) DISTANCE- $r$  DOMINATING SET problem on  $\mathcal{C}$  for any  $\alpha > 1$  (assuming the Gap Exponential Time Hypothesis).

**2012 ACM Subject Classification** Theory of computation  $\rightarrow$  Graph algorithms analysis, Theory of computation  $\rightarrow$  Parameterized complexity and exact algorithms, Theory of computation  $\rightarrow$  Fixed parameter tractability, Theory of computation  $\rightarrow$  Approximation algorithms analysis, Theory of computation  $\rightarrow$  W hierarchy

**Keywords and phrases** Lossy Kernelization, Connected Dominating Set, Sparse Graph Classes

**Digital Object Identifier** 10.4230/LIPIcs.STACS.2018.29

**Funding** Eduard Eiben was supported by Pareto-Optimal Parameterized Algorithms (ERC Starting Grant 715744) and by the Austrian Science Fund (FWF, projects P26696 and W1255-N23).



© Eduard Eiben, Mithilesh Kumar, Amer E. Mouawad,  
Fahad Panolan, and Sebastian Siebertz;  
licensed under Creative Commons License CC-BY

35th Symposium on Theoretical Aspects of Computer Science (STACS 2018).

Editors: Rolf Niedermeier and Brigitte Vallée; Article No. 29; pp. 29:1–29:15

Leibniz International Proceedings in Informatics



Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany



SYMPOSIUM  
ON THEORETICAL  
ASPECTS  
OF COMPUTER  
SCIENCE

The work of Sebastian Siebertz is supported by the National Science Centre of Poland via POLONEZ grant agreement UMO-2015/19/P/ST6/03998, which has received funding from the European Union's Horizon 2020 research and innovation programme (Marie Skłodowska-Curie grant agreement No. 665778).

## 1 Introduction

**Lossy kernelization.** A powerful method in parameterized complexity theory is to compute on input  $(I, k)$  a problem *kernel* in a polynomial-time pre-processing step, that is, to reduce the input instance in polynomial time to an equivalent instance  $(I', k')$  of size  $g(k)$  for some function  $g$  bounded in the parameter only. If the reduced instance  $(I', k')$  belongs to a different problem than  $(I, k)$ , we speak of a *bi-kernel*. It is well known that a problem is fixed-parameter tractable if and only if it admits a kernel, however, in general the function  $g$  can grow arbitrarily fast. For practical applications we are mainly interested in linear or at worst polynomial kernels. We refer to the textbooks [6, 11, 12] for extensive background on parameterized complexity and kernelization.

One shortcoming of the above notion of kernelization is that it does not combine well with approximation algorithms or heuristics. An approximate solution on the reduced instance provides no insight whatsoever about the original instance, the only statement we can derive from the definition of a kernel is that the reduced instance  $(I', k')$  is a positive instance if and only if the original instance  $(I, k)$  is a positive instance. This issue was recently addressed by Lokstanov et al. [23], who introduced the framework of *lossy kernelization*. Intuitively, the framework combines notions from approximation and kernelization algorithms to allow for approximation preserving kernels.

Formally, a *parameterized optimization (minimization or maximization) problem*  $\Pi$  over finite vocabulary  $\Sigma$  is a computable function  $\Pi: \Sigma^* \times \mathbb{N} \times \Sigma^* \rightarrow \mathbb{R} \cup \{\pm\infty\}$ . A *solution* for an instance  $(I, k) \in \Sigma^* \times \mathbb{N}$  is a string  $s \in \Sigma^*$ , such that  $|s| \leq |I| + k$ . The *value* of the solution  $s$  is  $\Pi(I, k, s)$ . For a minimization problem, the *optimum value* of an instance  $(I, k)$  is  $\text{OPT}_\Pi(I, k) = \min_{s \in \Sigma^*, |s| \leq |I| + k} \Pi(I, k, s)$ , for a maximization problem it is  $\text{OPT}_\Pi(I, k) = \max_{s \in \Sigma^*, |s| \leq |I| + k} \Pi(I, k, s)$ . An *optimal solution* is a solution  $s$  with  $\Pi(I, k, s) = \text{OPT}_\Pi(I, k)$ . If  $\Pi$  is clear from the context, we simply write  $\text{OPT}(I, k)$ .

A vertex-subset graph problem  $\mathcal{Q}$  defines which subsets of the vertices of an input graph are feasible solutions. We consider the following parameterized minimization problem associated with  $\mathcal{Q}$ :

$$\mathcal{Q}(G, k, S) = \begin{cases} \infty & \text{if } S \text{ is not a valid solution for } G \text{ as determined by } \mathcal{Q} \\ \min\{|S|, k + 1\} & \text{otherwise.} \end{cases}$$

Note that this bounding of the objective function at  $k + 1$  does not make sense for approximation algorithms if one insists on  $k$  being the unknown optimum solution of the instance  $I$ . The parameterization above is by the value of the solution that we want our algorithms to output.

► **Definition 1.1.** Let  $\alpha > 1$  and let  $\Pi$  be a parameterized minimization problem. An  $\alpha$ -*approximate polynomial time pre-processing algorithm*  $\mathcal{A}$  for  $\Pi$  is a pair of polynomial time algorithms. The first algorithm is called the *reduction algorithm*, and computes a map  $R_{\mathcal{A}}: \Sigma^* \times \mathbb{N} \rightarrow \Sigma^* \times \mathbb{N}$ . Given as input an instance  $(I, k)$  of  $\Pi$ , the reduction algorithm outputs another instance  $(I', k') = R_{\mathcal{A}}(I, k)$ . The second algorithm is called the *solution lifting algorithm*. It takes as input an instance  $(I, k) \in \Sigma^* \times \mathbb{N}$ , the output instance

$(I', k') = R_{\mathcal{A}}(I, k)$ , and a solution  $s'$  to the instance  $(I', k')$ . The solution lifting algorithm works in time polynomial in  $|I|, k, |I'|, k'$  and  $s'$ , and outputs a solution  $s$  to  $(I, k)$  such that

$$\frac{\Pi(I, k, s)}{\text{OPT}(I, k)} \leq \alpha \cdot \frac{\Pi(I', k', s')}{\text{OPT}(I', k')}.$$

► **Definition 1.2.** An  $\alpha$ -approximate kernelization algorithm is an  $\alpha$ -approximate polynomial time pre-processing algorithm for which we can prove an upper bound on the size of the output instances in terms of the parameter of the instance to be pre-processed. We speak of a linear or polynomial kernel, if the size bound is linear or polynomial, respectively. If we allow the reduced instance to be an instance of another problem, we speak of an  $\alpha$ -approximate bi-kernel.

We refer to the work of Lokshitanov et al. [23] for an extensive discussion of related work and examples of problems that admit lossy kernels.

**Sparse graphs and domination.** We consider finite, undirected and simple graphs and refer to the textbook [10] for all undefined notation. We write  $K_{i,j}$  for the complete bipartite graph with partitions of size  $i$  and  $j$ , respectively. We call a class  $\mathcal{C}$  of graphs biclique-free if there are  $i, j \in \mathbb{N}$  such that  $K_{i,j}$  is not a subgraph of  $G$  for all  $G \in \mathcal{C}$ .

The notion of nowhere denseness was introduced by Nešetřil and Ossona de Mendez [28, 29] as a general model of *uniform sparseness* of graphs. Many familiar classes of sparse graphs, like planar graphs, graphs of bounded tree-width, graphs of bounded degree, and all classes that exclude a fixed (topological) minor, are nowhere dense. An important and related concept is the notion of a graph class of *bounded expansion*, which was also introduced by Nešetřil and Ossona de Mendez [25, 26, 27].

► **Definition 1.3.** Let  $H$  be a graph and let  $r \in \mathbb{N}$ . An  $r$ -subdivision of  $H$  is obtained by replacing all edges of  $H$  by internally vertex disjoint paths of length at most  $r$ .

► **Definition 1.4.** A class  $\mathcal{C}$  of graphs is *nowhere dense* if there exists a function  $t: \mathbb{N} \rightarrow \mathbb{N}$  such that for all  $r \in \mathbb{N}$  and for all  $G \in \mathcal{C}$  we do not find an  $r$ -subdivision of the complete graph  $K_{t(r)}$  as a subgraph of  $G$ . Otherwise,  $\mathcal{C}$  is called *somewhere dense*.

► **Definition 1.5.** A class  $\mathcal{C}$  of graphs has *bounded expansion* if there exists a function  $d: \mathbb{N} \rightarrow \mathbb{N}$  such that for all  $r \in \mathbb{N}$  and all graphs  $H$ , such that an  $r$ -subdivision of  $H$  is a subgraph of  $G$  for some  $G \in \mathcal{C}$ , satisfy  $|E(H)|/|V(H)| \leq d(r)$ .

Every class of bounded expansion is nowhere dense, which in turn excludes some biclique as a subgraph and hence is biclique-free. For extensive background on bounded expansion and nowhere dense graphs we refer to the textbook of Nešetřil and Ossona de Mendez [30].

► **Definition 1.6.** In the parameterized DOMINATING SET (DS) problem we are given a graph  $G$  and  $k \in \mathbb{N}$ , and the objective is to determine the existence of a subset  $D \subseteq V(G)$  of size at most  $k$  such that every vertex  $u$  of  $G$  is *dominated* by  $D$ , that is, either  $u$  belongs to  $D$  or has a neighbor in  $D$ . More generally, for fixed  $r \in \mathbb{N}$ , in the DISTANCE- $r$  DOMINATING SET ( $r$ -DS) problem we are asked to determine the existence of a subset  $D \subseteq V(G)$  of size at most  $k$  such that every vertex  $u \in V(G)$  is within distance at most  $r$  from a vertex of  $D$ . In the CONNECTED (DISTANCE- $r$ ) DOMINATING SET (CDS/ $r$ -CDS) problem we additionally demand that the (distance- $r$ ) dominating set shall be connected.

DOMINATING SET plays a central role in the theory of parameterized complexity, as it is a prime example of a  $W[2]$ -complete problem with the size of the optimal solution as the parameter, thus considered intractable in full generality. For this reason, the (CONNECTED) DOMINATING SET problem and (CONNECTED) DISTANCE- $r$  DOMINATING SET problem have been extensively studied on restricted graph classes. A particularly fruitful line of research in this area concerns kernelization algorithms for the aforementioned problems [1, 3, 15, 16, 17, 31]. Philip et al. [31] obtained a kernel of size  $\mathcal{O}(k^{(d+1)^2})$  on  $d$ -degenerate graphs, for constant  $d$ , and more generally a kernel of size  $\mathcal{O}(k^{\max(i^2, j^2)})$  on graphs excluding the biclique  $K_{i,j}$  as a subgraph. On the lower bounds side, Cygan et al. [7] have shown that the existence of a size  $\mathcal{O}(k^{(d-1)(d-3)-\epsilon})$  kernel,  $\epsilon > 0$ , for DOMINATING SET on  $d$ -degenerate graphs would imply  $\text{NP} \subseteq \text{coNP/poly}$ . For the CONNECTED DOMINATING SET problem linear kernels are only known for planar [22] and  $H$ -topological-minor-free graphs [17]. Polynomial kernels are excluded already for graphs of bounded degeneracy [8], assuming  $\text{NP} \not\subseteq \text{coNP/poly}$ .

For the more general DISTANCE- $r$  DOMINATING SET problem we know the following results. Dawar and Kreutzer [9] showed that for every  $r \in \mathbb{N}$  and every nowhere dense class  $\mathcal{C}$ , the DISTANCE- $r$  DOMINATING SET problem is fixed-parameter tractable on  $\mathcal{C}$ . Drange et al. [13] gave a linear bi-kernel for DISTANCE- $r$  DOMINATING SET on any graph class of bounded expansion for every  $r \in \mathbb{N}$ , and a pseudo-linear kernel for DOMINATING SET on any nowhere dense graph class; that is, a kernel of size  $f(\epsilon) \cdot k^{1+\epsilon}$ , for some function  $f$  and any  $\epsilon > 0$ . Precisely, the kernelization algorithm of Drange et al. [13] outputs an instance of an annotated problem where some vertices are not required to be dominated; this will be the case in the present paper as well (except for the case of biclique-free graphs). Kreutzer et al. [21] provided a polynomial bi-kernel for the DISTANCE- $r$  DOMINATING SET problem on every nowhere dense class for every fixed  $r \in \mathbb{N}$  and finally, Eickmeyer et al. [14] could prove the existence of pseudo-linear bi-kernels of size  $f(r, \epsilon) \cdot k^{1+\epsilon}$ , for some function  $f$ .

It is known that bounded expansion classes of graphs are the limit for the existence of polynomial kernels for the CONNECTED DOMINATING SET problem. Drange et al. [13] gave an example of a subgraph-closed class of bounded expansion which does not admit a polynomial kernel for CONNECTED DOMINATING SET unless  $\text{NP} \subseteq \text{coNP/Poly}$ . They also showed that nowhere dense classes are the limit for the fixed-parameter tractability of the DISTANCE- $r$  DOMINATING SET problem if we assume closure under taking subgraphs (classes which are closed under taking subgraphs will be called *monotone classes*).

**Our results.** In this paper we prove the following results.

- For any  $\alpha > 1$ , CDS admits an  $\alpha$ -approximate kernel on  $K_{d,d}$ -free graphs of size  $k^{\mathcal{O}(\frac{d^2}{\alpha-1})}$ .
- For any  $\alpha > 1$ , CDS admits an  $\alpha$ -approximate bi-kernel on graphs of bounded expansion of size  $\mathcal{O}(f(\alpha) \cdot k)$  (i.e., linear in  $k$ ), where  $f(\alpha)$  is a function depending only on  $\alpha$ .
- For every  $\alpha > 1$  and every  $r \in \mathbb{N}$ , CONNECTED DISTANCE- $r$  DOMINATING SET admits an  $\alpha$ -approximate kernel of size polynomial in the parameter on classes of nowhere dense graphs, where the degree of the polynomial depends only on  $r$  and not on  $\alpha$ .

Observe that our results are in sharp contrast with the above mentioned lower bounds on kernel size of traditional kernels on these classes. We complement our results by the following conditional lower bound. We show that if a class  $\mathcal{C}$  is somewhere dense and closed under taking subgraphs, then for some value of  $r \in \mathbb{N}$  there cannot exist an  $\alpha$ -approximate bi-kernel for the (CONNECTED) DISTANCE- $r$  DOMINATING SET problem on  $\mathcal{C}$  for any  $\alpha > 1$  (assuming the Gap Exponential Time Hypothesis).

**Organization.** We explain our methods and provide a general framework for computing lossy kernels for CONNECTED DOMINATING SET in Section 2. Using this framework we can directly derive the claimed bounds for  $\alpha$ -approximate kernels on biclique-free graphs, while for bounded expansion and nowhere dense classes we obtain  $\alpha$ -approximate bi-kernels for CONNECTED DISTANCE- $r$  DOMINATING SET of polynomial size with the degree of the polynomial depending both on  $r$  and on  $\alpha$ . The rest of the paper is devoted to the far more technical part of improving the bounds for bounded expansion and nowhere dense classes of graphs. Due to space constraints, the conference version of this latter part contains only a proof outline.

## 2 A general framework

Although the technical details for dealing with biclique-free graphs and with bounded expansion and nowhere dense classes are quite different, the high-level approach is identical. The kernelization algorithms follow the same two-step strategy. First, our goal is to compute a “small” set of vertices whose domination is sufficient, i.e. the set of *dominatees* or the so-called domination core.

► **Definition 2.1** (*k*-domination core). Let  $G$  be a graph and  $Z \subseteq V(G)$ . We say that  $Z$  is a *k*-domination core if every set  $D$  of size at most  $k$  that dominates  $Z$  also dominates  $V(G)$ .

Having found a domination core  $Z$  of appropriate size, the next step is to reduce the number of *dominators*, i.e. vertices whose role is to dominate other vertices, and the number of *connectors*, i.e. vertices whose role is to connect the solution.

► **Definition 2.2.** Let  $G$  be a graph and let  $D, Z \subseteq V(G)$ . We say that  $D$  is a *Z*-dominator if  $D$  dominates  $Z$  in  $G$ , i.e. every vertex  $z \in Z \setminus D$  is at distance at most one from some vertex in  $D$ . We denote by  $\mathbf{ds}(G, Z)$  ( $\mathbf{cds}(G, Z)$ ) the size of a smallest (connected) *Z*-dominator in  $G$ . By  $\mathbf{ds}(G)$  ( $\mathbf{cds}(G)$ ) we mean  $\mathbf{ds}(G, V(G))$  ( $\mathbf{cds}(G, V(G))$ ).

We classify all vertices outside the core according to their domination properties.

► **Definition 2.3.** Let  $G$  be a graph and  $Z \subseteq V(G)$ . We define an equivalence relation  $\sim_Z$  on  $V(G) \setminus Z$  by  $u \sim_Z v \Leftrightarrow N(u) \cap Z = N(v) \cap Z$ .

Clearly, to find a kernel for DOMINATING SET it is now sufficient to construct the graph  $G'$  which contains the *k*-domination core  $Z$  and one representative of each equivalence class of  $\sim_Z$ . Then  $G$  admits a dominating set of size at most  $k$  if and only if  $G'$  contains a *Z*-dominator of size at most  $k$ . This simple two-step approach of computing a small domination core  $Z$  and then bounding the number of equivalence classes of the relation  $\sim_Z$  forms the basis of the kernelization algorithms for DOMINATING SET in [13, 14, 21]. To control the number of classes of  $\sim_Z$ , we give the following definition. The *index* of an equivalence relation is the number of equivalence classes.

► **Definition 2.4.** Let  $G$  be a graph. We define the *neighborhood complexity function* of  $G$  as the function  $\mu: \mathbb{N} \rightarrow \mathbb{N}$  with  $\mu(z) = \max_{Z \subseteq V(G), |Z|=z} \text{index}(\sim_Z)$ .

For example all classes of bounded VC-dimension have polynomially bounded neighborhood complexity functions [33, 34], while bounded expansion classes have linear and nowhere dense classes have almost linear neighborhood complexity [18].

► **Proposition 2.5.** *Let  $\mathcal{C}$  be a class of graphs such that the neighborhood complexity function for all  $G \in \mathcal{C}$  is bounded by a fixed polynomial and such that on input  $(G, k)$ , for  $G \in \mathcal{C}$ , we can decide in polynomial time whether  $\mathbf{ds}(G) > k$  or otherwise compute a  $k$ -domination core  $Z \subseteq V(G)$  of size polynomial in  $k$ . Then DOMINATING SET parameterized by  $k$  admits a polynomial-size kernel on  $\mathcal{C}$ .*

The above proposition can be applied, e.g., to obtain a polynomial kernel on biclique-free graphs (we will prove the existence of a polynomial  $k$ -domination core below). However, as the hardness results even for degenerate graphs show, this approach does not extend to connected dominating sets. We may have to include more vertices in the kernel to ensure connectivity of the dominating sets. This turns out to be a major problem for the construction of polynomial size kernels for CONNECTED DOMINATING SET.

When reducing the number of vertices outside the domination core, we borrow approximation techniques that are closely related to the STEINER TREE problem.

► **Definition 2.6.** Let  $G$  be a graph and let  $Y \subseteq V(G)$  be a set of *terminals*. A *Steiner tree* for  $Y$  is a subtree of  $G$  spanning  $Y$ . We write  $\mathbf{st}_G(Y)$  for the order of (i.e. the number of vertices of) the smallest Steiner tree for  $Y$  in  $G$  (including the vertices of  $Y$ ). If  $\mathcal{Y} = \{V_1, \dots, V_\ell\}$  is a family of vertex disjoint subsets of  $G$ , a *group Steiner tree* for  $\mathcal{Y}$  is a subtree of  $G$  that contains (at least) one vertex of each group  $V_i$ . We write  $\mathbf{st}_G(\mathcal{Y})$  for the order of the smallest group Steiner tree for  $\mathcal{Y}$ .

The GROUP STEINER TREE problem on  $t$  groups can be solved in  $\mathcal{O}(2^t \cdot n^{\mathcal{O}(1)})$ -time [24].

The following definition and proposition form the key to our approach to handle connectivity in the lossy kernelization framework.

► **Definition 2.7.** Let  $D$  be a connected graph and  $t \in \mathbb{N}$ . A  $(D, t)$ -*covering family* is a family  $\mathcal{F}(D, t)$  of connected subgraphs of  $D$  such that (i) for each  $T \in \mathcal{F}(D, t)$ ,  $|V(T)| \leq 2t$  and (ii)  $\bigcup_{T \in \mathcal{F}(D, t)} V(T) = V(D)$ .

► **Proposition 2.8.** *Let  $D$  be a connected graph and  $t \in \mathbb{N}$ . Then there is a  $(D, t)$ -covering family  $\mathcal{F}(D, t)$  such that  $|\mathcal{F}(D, t)| \leq \frac{|V(D)|}{t} + 1$ , and  $\sum_{T \in \mathcal{F}(D, t)} |V(T)| \leq (1 + \frac{1}{t})|V(D)| + 1$ .*

**Proof Sketch.** Let  $T_D$  be a spanning tree of  $D$ . We create a  $(D, t)$ -covering family  $\mathcal{F}(D, t) = \{T_1, T_2, \dots, T_\ell\}$ , which is a set of subtrees of  $T_D$  constructed as follows. We root  $T_D$  at an arbitrary vertex  $r \in V(T_D)$ . For any pair of vertices  $u, v \in V(T_D)$ ,  $u$  is called a child of  $v$  if  $uv \in E(T_D)$  and  $v$  lies on the path from  $u$  to  $r$ . For each vertex  $v \in V(T_D)$ , we let  $\mathbf{weight}(v) = 1 + \sum_{u \in \mathbf{child}(v)} \mathbf{weight}(u)$ , where  $\mathbf{child}(v)$  denotes the set of children of  $v$  in  $T_D$ . In other words,  $\mathbf{weight}(v)$  is the number of vertices in the subtree rooted at  $v$ . Leaves have weight one. We use  $T_v$  to denote the subtree rooted at  $v$ . We construct  $\mathcal{F}(D, t) = \{T_1, T_2, \dots\}$  from  $T_D$  as follows:

1. If  $T_D$  is empty, terminate. Otherwise, compute the weights of all vertices in  $T_D$ , then sort the vertices in increasing order of weight.
2. If there exists a vertex whose weight is between  $t$  and  $2t$  (inclusive), pick the vertex with the smallest such weight, add  $T_v$  to  $\mathcal{F}(D, t)$ , delete  $T_v$  from  $T_D$ , then go back to step (1).
3. If there exists a vertex whose weight is strictly greater than  $2t$ , pick the vertex with the smallest such weight, greedily compute a subset  $S \subseteq \mathbf{child}(v)$  such that  $t < \sum_{u \in S} \mathbf{weight}(u) < 2t$ , let  $R = \mathbf{child}(v) \setminus S$ , add  $T_v - \bigcup_{w \in R} V(T_w)$  to  $\mathcal{F}(D, t)$ , delete  $T_u$  from  $T_D$ , for every  $u \in S$ , then go back to step (1). Note that by our choice of  $v$ , all children of  $v$  must have weight at most  $t - 1$  as otherwise case (2) would apply.

4. Otherwise, every vertex in  $T_D$  has weight strictly less than  $t$  and hence  $T_D$  has at most  $t$  vertices (by the definition of the weight function). In this case, simply add  $T_D$  to  $\mathcal{F}(D, t)$  and terminate.

The correctness proof is deferred to the full version of the paper due to the paucity of space.  $\blacktriangleleft$

While the previous proposition allows us to “take apart” connected dominating sets, the next proposition explains how to “put them back together”.

► **Proposition 2.9.** *Let  $G$  be a graph,  $X \subseteq V(G)$ , such that  $G[X]$  is connected, and let  $D$  be an  $X$ -dominator such that  $G[D]$  has at most  $p$  connected components. Then a set  $Q \subseteq X$  of size at most  $2p$  such that  $G[D \cup Q]$  is connected, can be computed in polynomial time.*

We have now collected all the tools required to control the number of vertices which have to be added to ensure connectivity.

► **Theorem 2.10.** *Let  $\mathcal{C}$  be a class of graphs such that the neighborhood complexity function for all  $G \in \mathcal{C}$  is bounded by a fixed polynomial of degree  $d$  and such that on input  $(G, k)$ , for  $G \in \mathcal{C}$ , we can decide in polynomial time whether  $\mathbf{c}ds(G) > k$  or otherwise compute a  $k$ -domination core  $Z \subseteq V(G)$ . Then for every  $\epsilon > 0$ , CONNECTED DOMINATING SET parameterized by  $k$  admits a  $(1 + \epsilon)$ -approximate kernel with  $|Z|^{\mathcal{O}(d/\epsilon)}$  vertices on  $\mathcal{C}$ .*

**The reduction algorithm.** Let  $(G, k)$  be the input instance, where  $G \in \mathcal{C}$  is connected and  $k$  is a positive integer. We first describe the reduction algorithm  $\mathcal{R}_A$ . As a first step we run the polynomial time algorithm (which exists by assumption of the theorem) to decide whether  $\mathbf{c}ds(G) > k$  and otherwise compute a  $k$ -domination core  $Z \subseteq V(G)$ . In the first case, we output a trivial negative instance  $((\{v\}, \emptyset), 0)$ . In the second case, we proceed as follows.

We partition the graph into two sets  $Z$  and  $R = V(G) \setminus Z$ . We compute the equivalence relation  $\sim_Z$  on  $R$  (see Definition 2.3), that is, we partition vertices in  $R$  according to their neighborhoods in  $Z$ . This is clearly possible in polynomial time. Let  $\mathcal{R}$  be the set of equivalence classes defined by  $\sim_Z$ . As a direct implication of our assumption, we can bound the size of  $\mathcal{R}$  by  $\mathcal{O}(|Z|^d)$ .

► **Proposition 2.11.** *The equivalence relation  $\sim_Z$  has  $\mathcal{O}(|Z|^d)$  classes, i.e.  $|\mathcal{R}| \in \mathcal{O}(|Z|^d)$ .*

As  $Z$  is a  $k$ -domination core, to find a dominating set of size at most  $k$  it is enough to find a set which dominates  $Z$ . Hence for the purpose of domination, it is redundant to pick more than one vertex from an equivalence class in  $\mathcal{R}$ . The following construction finds a small set of relevant vertices which “approximately” preserves the connectivity requirements.

Let  $t \geq 1$  be a constant, which we fix later. Let  $\mathcal{Z}$  be the family of groups  $\{\{z\} \mid z \in Z\}$  and let  $\mathcal{R}$  be the set of equivalence classes defined by  $\sim_Z$ . The set  $\mathcal{R} \cup \mathcal{Z}$  forms a family of groups of vertices in  $V(G)$ . For every subset  $\mathcal{Q} = \{Q_1, \dots, Q_\ell\} \subseteq \mathcal{R} \cup \mathcal{Z}$  of size at most  $2t$  of groups in  $\mathcal{R} \cup \mathcal{Z}$ , construct a GROUP STEINER TREE instance on the graph  $G$  with groups  $Q_1, \dots, Q_\ell$ . Note that since  $t$  is a constant each instance can be solved in polynomial time using the algorithm of Misra et al. [24]. For each subset  $\mathcal{Q}$  denote by  $T_{\mathcal{Q}}$  the corresponding solution. For every instance that we solve, if the size of  $T_{\mathcal{Q}}$  is at most  $2t$  then we mark the vertices of  $T_{\mathcal{Q}}$  in  $G$ . We denote the set of all marked vertices by  $\bigcup T_{\mathcal{Q}}$ . If  $\bigcup T_{\mathcal{Q}}$  is not a dominating set in  $G$ , then we may declare that  $\mathbf{c}ds(G) > k$ . Otherwise, since  $G$  is assumed to be connected, we can run the polynomial-time algorithm of Proposition 2.9 (with parameter  $X = V(G)$ ) to obtain a set  $W \subseteq V(G)$  such that  $\bigcup T_{\mathcal{Q}} \cup W$  is a connected dominating set in  $G$  and  $|\bigcup T_{\mathcal{Q}} \cup W| \leq 3|\bigcup T_{\mathcal{Q}}|$ . Let  $Y = \bigcup T_{\mathcal{Q}} \cup W$ . We output the instance  $(G[Y], k)$ .

**Approximation guarantee.** Now we prove that  $\text{OPT}(G[Y], k) \leq (1 + \epsilon)\text{OPT}(G, k)$ . Let  $D^*$  be a connected dominating set of  $G$  of minimum cardinality. If  $|D^*| > k$ , then  $\text{OPT}(G[Y], k) \leq (1 + \epsilon)\text{OPT}(G, k)$  holds trivially. So we assume that  $|D^*| \leq k$ . We let  $\mathcal{F}(D^*, t) = \{T_1, T_2, \dots, T_m\}$  denote a  $(D^*, t)$ -covering family. Proposition 2.8 implies that there exists such a family for which  $|\mathcal{F}(D^*, t)| \leq \frac{|V(D^*)|}{t} + 1$  and  $\sum_{T \in \mathcal{F}(D^*, t)} |V(T)| \leq (1 + \frac{1}{t})|V(D^*)| + 1$ . Moreover, the size of each connected subgraph  $T$  (in this case also subtree) is at most  $2t$ . We construct a new family  $\mathcal{F}'$  from  $\mathcal{F}(D^*, t)$  as follows. For each  $T \in \mathcal{F}(D^*, t)$ , we replace  $T$  by  $T_{\mathcal{Q}}$ , where  $\mathcal{Q}$  is the set of groups from  $\mathcal{R} \cup \mathcal{Z}$  such that  $Q \in \mathcal{Q}$  if and only if  $V(T) \cap Q \neq \emptyset$  and  $T_{\mathcal{Q}}$  is the set of marked vertices in an optimal Steiner tree connecting vertices from the groups in  $\mathcal{Q}$ . Note that the fact that  $T$  is of size at most  $2t$  guarantees the existence of  $T_{\mathcal{Q}}$  (by construction). Moreover, the size of  $T_{\mathcal{Q}}$  is at most the size of  $T$ , since  $T$  is also a solution for GROUP STEINER TREE for  $\mathcal{Q}$ . Let  $D_{\mathcal{F}'}$  denote the union of all vertices in  $\mathcal{F}'$ .

Let  $D'$  be a subset of  $D_{\mathcal{F}'}$ , of cardinality at most  $|D^*|$ , such that for any  $w \in D^*$ , there is a vertex  $w' \in D'$  with the property that  $\{w, w'\} \subseteq Q \in \mathcal{R} \cup \mathcal{Z}$  and  $w' \in D_{\mathcal{F}'}$ . That is, if  $D^*$  has a vertex from a group  $Q$  in  $\mathcal{R} \cup \mathcal{Z}$ , then  $D'$  also has a vertex from group  $Q$ . We claim that  $D'$  is a dominating set in  $G$ . Notice that  $Z \cap V(D) = Z \cap D'$  and if any vertex in  $Z$  is adjacent to a vertex in a group  $Q$ , then it is adjacent to all vertices in group  $Q$ . This implies that  $D'$  also dominates  $Z$  and since  $|D'| \leq |D| \leq k$ , by the definition of a  $k$ -domination core,  $D'$  is a dominating set in  $G$ .

This implies that  $D_{\mathcal{F}'} \supseteq D'$  is also a dominating set in  $G$ . Applying Proposition 2.9 in  $G[Y]$  (with  $D_{\mathcal{F}'}$  as dominator and since  $G[Y]$  is connected), we obtain a connected dominating set of size at most  $2|\mathcal{F}(D^*, t)| + |D_{\mathcal{F}'}| \leq \frac{2|V(D^*)|}{t} + 2 + (1 + \frac{1}{t})|V(D^*)| + 1 = (1 + \frac{3}{t})|V(D^*)| + 3$ . Now we can fix the constant  $t$  appropriately (as roughly  $\frac{3}{\epsilon}$ ) and we get that  $\text{OPT}(G[Y], k) \leq (1 + \epsilon)\text{OPT}(G, k)$ .

**Size of the kernel.** Now we show that  $|Y| \in |Z|^{\mathcal{O}(d)/\epsilon}$ . By Proposition 2.11, we have that  $|\mathcal{R} \cup \mathcal{Z}| = \mathcal{O}(Z^d)$ . From the construction, it follows that  $|\bigcup T_{\mathcal{Q}}| = \mathcal{O}(2t|\mathcal{R} \cup \mathcal{Z}|^{\mathcal{O}(t)}) = |Z|^{\mathcal{O}(d/\epsilon)}$ . Notice that  $Y = \bigcup T_{\mathcal{Q}} \cup W$ , where  $W$  is obtained by applying Proposition 2.9 and hence we have that  $Y = |\bigcup T_{\mathcal{Q}} \cup W| \leq 3|\bigcup T_{\mathcal{Q}}| = |Z|^{\mathcal{O}(d/\epsilon)}$ .

**The solution lifting algorithm.** The solution lifting algorithm works as follows. Given a solution  $D'$  to the reduced instance  $(G', k')$ , if  $D'$  is not a connected dominating set of  $G'$ , then the solution lifting algorithm will output  $\emptyset$ . If  $D'$  is a connected dominating set, then the algorithm returns  $D'$  if  $|D'| \leq k$  and  $V(G)$  otherwise. Let  $D$  be the output of the solution lifting algorithm.

**The final step.** We prove that the above reduction algorithm together with the solution lifting algorithm constitute a  $(1 + \epsilon)$ -approximate kernel. Note that if  $D'$  is not a valid solution of  $G'$ , then  $\emptyset$  is not a valid solution for  $G$  and  $\text{CDS}(G', k', D') = \text{CDS}(G, k, D) = \infty$ . Hence we can restrict ourselves to the case when  $D'$  is a connected dominating set of  $G'$ . First, consider the case where the reduction algorithm outputs  $Y \subseteq V(G)$  and the reduced instance is hence  $(G', k') = (G[Y], k)$ . From our above observation, we have that  $\text{OPT}(G[Y], k) \leq (1 + \epsilon)\text{OPT}(G, k)$ . We show that in this case  $\text{CDS}(G, k, D) = \text{CDS}(G', k', D')$ . If  $|D'| > k$ , then  $\text{CDS}(G, k, D) = \text{CDS}(G, k, V(G)) = k + 1 = \text{CDS}(G', k', D')$ . So assume that  $|D'| \leq k$ , which implies  $D = D'$ . Since  $D'$  is a connected dominating set of  $G[Y]$  and  $Y$  contains a  $k$ -domination core of  $G$ , it follows that  $D'$  dominates  $G$  and  $\text{CDS}(G, k, D) = \text{CDS}(G', k', D')$ . Combining  $\text{CDS}(G, k, D) = \text{CDS}(G', k', D')$  and  $\text{OPT}(G[Y], k) \leq (1 + \epsilon)\text{OPT}(G, k)$  we get  $\frac{\text{CDS}(G, k, D)}{\text{OPT}(G, k)} \leq (1 + \epsilon) \frac{\text{CDS}(G', k', D')}{\text{OPT}(G', k')}$ .



When  $(G', k') = ((\{v\}, \emptyset), 0)$ , we can easily verify that the above mentioned approximation guarantee holds.  $\square$

The remainder of the paper is concerned with proving the existence of small domination cores for concrete sparse classes of graphs. For example, to prove the following theorem for biclique-free graphs we prove the existence of a polynomial  $k$ -domination core and simply use the fact that such classes have polynomially bounded neighborhood complexity.

► **Theorem 2.12.** *For every  $\epsilon > 0$ , CONNECTED DOMINATING SET, parameterized by solution size, admits a  $(1 + \epsilon)$ -approximate kernel with  $k^{\mathcal{O}(d^2/\epsilon)}$  vertices on  $K_{d,d}$ -free graphs.*

The most technical part will be to prove the existence of a linear domination core for bounded expansion classes (the definition of a domination core is slightly changed to obtain such good bounds, see Section 3). Most surprisingly, the general framework summarized in Theorem 2.10 does not produce a bi-kernel of size  $\mathcal{O}(|Z|^{1/\epsilon})$  but rather of size  $f(\epsilon) \cdot |Z|$  for some function  $f$  on bounded expansion classes and of polynomial size on nowhere dense classes.

### 3 Graphs of bounded expansion

In this section we show that CONNECTED DOMINATING SET, parameterized by solution size  $k$ , admits a  $(1 + \epsilon)$ -approximate bi-kernel on at most  $\mathcal{O}(f(\epsilon) \cdot k)$  vertices on graphs of bounded expansion. The reduced instance will be an instance of SUBSET CONNECTED DOMINATING SET (SCDS), defined as follows:

$$\text{SCDS}((G, Z), k, D) = \begin{cases} \infty & \text{if } D \text{ is not a connected } Z\text{-dominator in } G \\ \min\{|D|, k + 1\} & \text{otherwise} \end{cases}$$

We first formally define the class of graphs of bounded expansion. Towards that we need the definition of shallow minors.

► **Definition 3.1.** A graph  $M$  is an  $r$ -shallow minor of  $G$ , for some  $r \in \mathbb{N}$ , if there is a family of disjoint subsets  $V_1, \dots, V_{|M|}$  of  $V(G)$  such that (i) each graph  $G[V_i]$  is connected and has radius at most  $r$ , and (ii) there is a bijection  $\omega: V(M) \rightarrow \{V_1, \dots, V_{|M|}\}$  such that for any  $uv \in E(M)$  there is an edge in  $G$  with one endpoint in  $\omega(u)$  and another in  $\omega(v)$ .

The set of all  $r$ -shallow minors of a graph  $G$  is denoted by  $G \nabla r$ . The set of all  $r$ -shallow minors of all members of a graph class  $\mathcal{G}$  is denoted by  $\mathcal{G} \nabla r = \bigcup_{G \in \mathcal{G}} (G \nabla r)$ . It will be convenient to work with the following equivalent definition of bounded expansion classes.

► **Definition 3.2** (Grad and bounded expansion [25]). For a graph  $G$  and an integer  $r \geq 0$ , the *greatest reduced average density (grad) at depth  $r$*  is,  $\nabla_r(G) = \max_{M \in \mathcal{G} \nabla r} \text{density}(M) = \max_{M \in \mathcal{G} \nabla r} |E(M)|/|V(M)|$ . For a graph class  $\mathcal{G}$ ,  $\nabla_r(\mathcal{G}) = \sup_{G \in \mathcal{G}} \nabla_r(G)$ . A graph class  $\mathcal{G}$  has *bounded expansion* if there is a function  $f: \mathbb{N} \rightarrow \mathbb{R}$  such that for all  $r$  we have  $\nabla_r(\mathcal{G}) \leq f(r)$ .

The first phase of our algorithm, i.e. finding a domination core, closely follows the work of Drange et al. [13] but requires subtle changes.

#### 3.1 Finding the domination core

In the following we fix a graph class  $\mathcal{G}$  that has bounded expansion and let  $(G, k)$  be the input instance of CONNECTED DOMINATING SET, where  $G \in \mathcal{G}$  and  $G$  is connected.

To obtain a linear domination core for classes of bounded expansion we have to invest a considerable amount of work. The following construction shows that we cannot work with  $k$ -domination cores.

► **Lemma 3.3.** *There exists a class  $\mathcal{C}$  of bounded expansion such that for all  $k \in \mathbb{N}$  there is  $G \in \mathcal{C}$  such that every  $k$ -domination core for  $G$  has  $\Omega(k^2)$  vertices.*

Instead, we introduce the notion of a  $c$ -exchange domination core, which is different from the definition used in the previous section and from the one considered in [13]. Here,  $c$  is a fixed constant which we set later.

► **Definition 3.4** ( $c$ -exchange domination core). Let  $G$  be a graph and  $Z \subseteq V(G)$ . We say that  $Z$  is a  $c$ -exchange domination core if for every set  $X$  that dominates  $Z$  one of the following conditions holds: (1.)  $X$  dominates  $G$ , or (2.) there exist  $A \subseteq X$  and  $B \subseteq V(G)$  such that  $|B| < |A| \leq c$  and  $(X \setminus A) \cup B$  dominates  $Z$ . Moreover the number of connected components of  $(X \setminus A) \cup B$  is at most the number of connected components of  $X$ . In particular, if  $X$  is a connected set then  $(X \setminus A) \cup B$  is also connected.

Clearly,  $V(G)$  is a  $c$ -exchange domination core, for any  $c$ , but we look for a  $c$ -exchange domination core that is linear in  $k$ . Hence, we start with  $Z = V(G)$  and gradually reduce  $|Z|$  by removing one vertex at a time, while maintaining the invariant that  $Z$  is a  $c$ -exchange domination core. To this end, we need to prove Lemma 3.5. Note that we only remove vertices from  $Z$  at this stage (no vertex deletions), and hence the graph remains intact.

► **Lemma 3.5.** *There exists a constant  $C_{\text{core}} > 0$  depending only on a fixed (finite) number of grads of  $\mathcal{G}$  and a polynomial-time algorithm that, given a graph  $G \in \mathcal{G}$  and a  $c$ -exchange domination core  $Z \subseteq V(G)$  with  $|Z| > C_{\text{core}} \cdot k$ , either correctly concludes that  $\text{cds}(G) > k$  or finds a vertex  $z \in Z$  such that  $Z \setminus \{z\}$  is still a  $c$ -exchange domination core. Here  $c$  is a non-zero positive constant.*

We remark that we do not know how to prove the existence of a *small*  $c$ -exchange distance- $r$  domination core (a generalization needed for  $r$ -CDS) on graphs of bounded expansion. In fact the generalization of our proof will not guarantee that the number of connected components after exchange is at most the number of connected component before exchange. This obstacle prevents us from finding a linear or almost linear bi-kernel for the case of  $r$ -CDS.

## 3.2 Reducing connectors and dominators

Armed with a  $c$ -exchange domination core  $Z$  whose size is linear in  $k$ , our next goal is to reduce the number of connectors and dominators (the number of vertices in  $V(G) \setminus Z$ ). To that end, we need the following lemma which is a generalized version of Lemma 2.11 in [13].

► **Lemma 3.6** (Trees closure lemma). *Let  $\mathcal{G}$  be a class of bounded expansion and let  $q$  and  $r$  be positive integers. Let  $G \in \mathcal{G}$  be a graph and  $X \subseteq V(G)$ . Then a superset of vertices  $X' \supseteq X$  can be computed in polynomial time, with the following properties: (1) For every  $Y \subseteq X$  of size at most  $q$ , if  $\text{st}_G(Y) \leq rq$  then  $\text{st}_{G[X']}(Y) = \text{st}_G(Y)$ . (2)  $|X'| \leq C_{\text{tc}} \cdot |X|$ , where  $C_{\text{tc}}$  is a constant depending only on  $r$ ,  $q$ , and the class  $\mathcal{G}$ .*

Lemma 3.6 ensures that the reduction algorithm of Theorem 2.10 produces a set  $\bigcup T_Q$  of linear size. We additionally use the linear neighborhood complexity of classes of bounded expansion to conclude as in the proof of Theorem 2.10.

The rest of section is devoted to prove Lemma 3.6. Towards that we need some more definitions and known results which we state first. Given two graphs  $G$  and  $H$ , the *lexicographic product*  $G \odot H$  is defined as the graph on the vertex set  $V(G) \times V(H)$  where vertices  $(u, a)$  and  $(v, b)$  are adjacent if  $uv \in E(G)$  or if  $u = v$  and  $ab \in E(H)$ .

► **Lemma 3.7** ([19, 20]). *For a graph  $G$  and non-negative integers  $t \geq 1$  and  $r$  we have  $\nabla_r(G \odot K_t) \leq 5t^2(r+1)^2 \cdot \nabla_r(G)$ .*

Let  $G$  be a graph and  $X$  be a subset of its vertices. For  $u \in V(G) \setminus X$ , we define the  *$r$ -projection* of  $u$  onto  $X$  as follows:  $M_r^G(u, X)$  is the set of all vertices  $w \in X$  for which there exists a path in  $G$  that starts in  $u$ , ends in  $w$ , has length at most  $r$ , and whose internal vertices do not belong to  $X$ . Note that  $M_1^G(u, X) = N_X(u)$ .

► **Lemma 3.8** ([13]). *Let  $\mathcal{G}$  be a class of graphs of bounded expansion. There exists a polynomial-time algorithm that, given a graph  $G \in \mathcal{G}$ ,  $X \subseteq V(G)$ , and an integer  $r \geq 1$ , computes the  $r$ -closure of  $X$ , denoted by  $cl_r(X)$ , with the following properties:*

- (i)  $X \subseteq cl_r(X) \subseteq V(G)$ ,
- (ii)  $|cl_r(X)| \leq C_{cl1} \cdot |X|$ , and
- (iii)  $|M_r^G(u, cl_r(X))| \leq C_{cl2}$  for each  $u \in V(G) \setminus cl_r(X)$ , where  $C_{cl1}$  and  $C_{cl2}$  are constants depending only on  $r$  and a fixed (finite) number of grads of  $\mathcal{G}$ .

**Proof of Lemma 3.6.** First, using Lemma 3.8 we compute  $X_0 = cl_{rq}(X)$ . Then,  $|X_0| \leq C_{cl1} \cdot |X|$  and for each vertex  $u \notin X_0$  we have  $|M_{rq}^G(u, X_0)| \leq C_{cl2}$ . Now, for each set  $Y \subseteq X_0$  of at most  $q$  vertices, compute an optimal Steiner tree  $T_Y$  whose edges do not belong to  $G[X_0]$ ; in case there is no such tree, set  $T_Y = \emptyset$ . Note that  $T_Y$  can be computed in polynomial time for any fixed  $q$  [2]. Define  $X'$  to be  $X_0$  plus the vertex sets of all trees  $T_Y$  that have size at most  $rq$ .

► **Claim 3.9.**  $|X'| \leq C_{tc} \cdot |X_0|$ , where  $C_{tc}$  is a constant depending only on  $r, q$ , and a finite number of grads of  $\mathcal{G}$ .

**Proof of the Claim.** Let  $H$  be a graph on vertex set  $X_0$ , where  $uv \in E(H)$  if and only if there exists  $Y$  such that  $\{u, v\} \subseteq Y$ ,  $T_Y \neq \emptyset$  and has size at most  $rq$ , and hence its vertex set was added to  $X$ . Note that we do not add multiedges. For every such set  $Y$ ,  $H[Y]$  induces a clique in  $H$ . Let  $\omega(H)$  denote the number of cliques in  $H$ . Clearly  $|X'| \leq |X_0| + rq \cdot \omega(H)$ , so it suffices to prove an upper bound on  $\omega(H)$ .

Consider an edge  $uv \in E(H)$ . The existence of this edge implies that  $u$  and  $v$  appear together in some tree  $T_Y$  of size at most  $rq$ . Since  $T_Y$  does not contain any edges from  $G[X_0]$  (by construction), there must exist a path  $P_{u,v}$  of length at most  $rq$  connecting  $u$  and  $v$ . The internal vertices of  $P_{u,v}$  do not belong to  $X_0$ . Take any  $w \in X' \setminus X_0$ , and consider for how many pairs  $\{u, v\} \subseteq X_0$  it can hold that  $w \in P_{u,v}$ . If  $\{u, v\}$  is such a pair, then in particular  $u, v \in M_{rq}^G(w, X_0)$ . But we know that  $|M_{rq}^G(w, X_0)| \leq C_{cl2}$ , so the number of such pairs is at most  $\tau \leq (C_{cl2})^2$ . Consequently, we observe that graph  $H$  is an  $(rq-1)$ -shallow minor of  $G \odot K_\tau$ : when each vertex  $w \in X' \setminus X_0$  is replaced with  $\tau$  copies, then we can realize all the paths between  $u$  and  $v$ , in  $G \odot K_\tau$ , so that they are internally vertex-disjoint. From Lemma 3.7, we know that  $\nabla_{rq-1}(G \odot K_\tau)$  is bounded polynomially in  $\nabla_{rq-1}(G)$  and  $\tau$ , which in turn is also bounded polynomially in  $\nabla_{rq-1}(\mathcal{G})$ . Hence  $\nabla_{rq-1}(G \odot K_\tau)$  is bounded polynomially in  $\nabla_{rq-1}(\mathcal{G})$ . The number of cliques in graph of bounded expansion is linear in the number of vertices [4]. Combining the fact that  $H$  has bounded expansion with  $|X'| \leq |X_0| + rq \cdot \omega(H)$ , the claim follows. ◀

► **Claim 3.10.** *If  $Y \subseteq X_0$  has size at most  $q$  and  $st_G(Y) \leq rq$  then  $st_{G[X']}(Y) = st_G(Y)$ .*

**Proof of the Claim.** Let  $T_Y$  be an optimal Steiner tree for  $Y$  in  $G$ , and let  $T_1, T_2, \dots, T_p$  be the subtrees of size greater than one obtained after deleting all edges of  $T_Y$  for which both endpoints are in  $X_0$ . Note that deleting such edges can only create either singleton vertices or subtrees of size greater than one. Moreover, let  $Y_i$ ,  $1 \leq i \leq p$ , denote the set  $Y \cap V(T_i)$ . The existence of  $T_i$  certifies that some tree of size at most  $|T_i|$  was added when constructing  $X'$  from  $X_0$ , and hence  $\text{st}_{G[X']}(Y_i) \leq |T_i|$ . Consequently, we infer that

$$\text{st}_{G[X']}(Y) \leq \sum_{i=1}^p \text{st}_{G[X']}(Y_i) + |Y \setminus \bigcup_{i=1}^p Y_i| \leq \sum_{i=1}^p |T_i| + |Y \setminus \bigcup_{i=1}^p Y_i| \leq |T_Y| = \text{st}_G(Y).$$

The opposite inequality  $\text{st}_{G[X']}(Y) \geq \text{st}_G(Y)$  follows directly from the fact that  $G[X']$  is an induced subgraph of  $G$ .  $\blacktriangleleft$

Claim 3.9 and the fact that  $|X_0| \leq C_{\text{c11}}|X|$  prove property (2). Claim 3.10 and the fact that  $X \subseteq X_0$  prove property (1).  $\blacktriangleleft$

#### 4 Nowhere dense graphs

As we solve the more general CONNECTED DISTANCE- $r$  DOMINATING SET on nowhere dense classes, we work with the following definition of a domination core.

► **Definition 4.1.** Let  $G$  be a graph. A set  $Z \subseteq V(G)$  is a  $(k, r)$ -*domination core* for  $G$  if every set  $D$  of size at most  $k$  that  $r$ -dominates  $Z$  also  $r$ -dominates  $G$

Domination cores of polynomial size exist for nowhere dense classes, as the following lemma shows.

► **Lemma 4.2** (Kreutzer et al. [21]). *There exists a polynomial  $q$  (of degree depending only on  $r$ ) and a polynomial-time algorithm that, given a graph  $G \in \mathcal{C}$  and  $k \in \mathbb{N}$  either correctly concludes that  $G$  cannot be  $r$ -dominated by a set of at most  $k$  vertices, or finds a  $(k, r)$ -domination core  $Z \subseteq V(G)$  of  $G$  of size at most  $q(k)$ .*

We remark that the non-constructive polynomial bounds that follow from [21] can be replaced by the much improved constructive bounds from [32].

It now remains to prove a lemma analogous to Lemma 3.6 for nowhere dense classes to conclude again similarly as in the proof of Theorem 2.10.

#### 5 Lower bounds

Our lower bound is based on Proposition 3.2 of [23] which establishes equivalence between FPT-approximation algorithms and approximate kernelization. More precisely, the Proposition states that for every function  $\alpha$  and decidable parameterized optimization problem  $\Pi$ ,  $\Pi$  admits a fixed parameter tractable  $\alpha$ -approximation algorithm if and only if  $\Pi$  has an  $\alpha$ -approximate kernel.

We use a reduction from SET COVER to the DISTANCE- $r$  DOMINATING SET problem, which under the assumption that the Gap Exponential Time Hypothesis (gap-ETH) holds does not admit a fixed-parameter tractable  $\alpha$ -approximation algorithm for any function  $\alpha$  [5].

For every monotone somewhere dense graph class  $\mathcal{C}$ , there exists  $r \in \mathbb{N}$  such that the exact  $r$ -subdivision of every graph can be found as a member of  $\mathcal{C}$  [29]. Now it is straight forward to adapt the  $W[2]$ -hardness proof for DISTANCE- $r$  DOMINATING SET for monotone somewhere dense graph classes [13] to our setting.

► **Theorem 5.1.** *If the Gap Exponential Time Hypothesis holds, then for every monotone somewhere dense class of graphs  $\mathcal{C}$  there is no  $\alpha(k)$ -approximate kernel for the DISTANCE- $r$  DOMINATING SET problem on  $\mathcal{C}$  for any function  $\alpha: \mathbb{N} \rightarrow \mathbb{N}$ .*

## 6 Conclusion

The study of computationally hard problems on restricted classes of inputs is a very fruitful line of research in algorithmic graph structure theory and in particular in parameterized complexity theory. This research is based on the observation that many problems such as DOMINATING SET, which are considered intractable in general, can be solved efficiently on restricted graph classes. In this work we were able to provide lossy kernels for graphs of bounded expansion whose size matches the size of the best known kernel for DOMINATING SET. We were furthermore able to identify the exact limit for the existence of lossy kernels for  $r$ -CDS. One interesting open question is whether our polynomial bounds on the size of the lossy kernel on nowhere dense classes can be improved to pseudo-linear bounds. For  $K_{d,d}$ -free graphs we have an additional  $\frac{1}{\alpha-1}$  multiplicative factor in the exponent. This leads to the question whether it is possible to reduce the size of our kernel on  $K_{d,d}$ -free graphs to  $f(\alpha)k^{\mathcal{O}(d^2)}$  for some function  $f$ . And, in light of the  $\mathcal{O}(k^{(d-1)(d-3)-\epsilon})$  lower bound for DOMINATING SET, is it possible to obtain a lossy kernel for DOMINATING SET on biclique-free graphs that beats this bound? Our hope is that such a “fine-grained” analysis of the kernelization complexity of domination problems will lead to a better understanding of the boundary between “hard” and “easy” instances.

---

## References

- 1 Jochen Alber, Michael R. Fellows, and Rolf Niedermeier. Polynomial-time data reduction for dominating set. *J. ACM*, 51(3):363–384, 2004.
- 2 Andreas Björklund, Thore Husfeldt, Petteri Kaski, and Mikko Koivisto. Fourier meets möbius: Fast subset convolution. In *Proceedings of the Thirty-ninth Annual ACM Symposium on Theory of Computing*, pages 67–74, New York, NY, USA, 2007.
- 3 Hans L. Bodlaender, Fedor V. Fomin, Daniel Lokshtanov, Eelko Penninkx, Saket Saurabh, and Dimitrios M. Thilikos. (Meta) Kernelization. In *50th Annual IEEE Symposium on Foundations of Computer Science, Atlanta, Georgia, USA*, pages 629–638, 2009.
- 4 Andreas Brandstädt, Van Bang Le, and Jeremy P. Spinrad. *Graph Classes: A Survey*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1999.
- 5 Parinya Chalermsook, Marek Cygan, Guy Kortsarz, Bundit Laekhanukit, Pasin Manurangsi, Danupon Nanongkai, and Luca Trevisan. From gap-eth to fpt-inapproximability: Clique, dominating set, and more. *arXiv preprint arXiv:1708.04218*, 2017.
- 6 Marek Cygan, Fedor V Fomin, Łukasz Kowalik, Daniel Lokshtanov, Dániel Marx, Marcin Pilipczuk, Michał Pilipczuk, and Saket Saurabh. *Parameterized algorithms*, volume 3. Springer, 2015.
- 7 Marek Cygan, Fabrizio Grandoni, and Danny Hermelin. Tight kernel bounds for problems on graphs with small degeneracy - (extended abstract). In *Algorithms - ESA 2013 - 21st Annual European Symposium, Sophia Antipolis, France, September 2-4, 2013. Proceedings*, pages 361–372, 2013.
- 8 Marek Cygan, Marcin Pilipczuk, Michał Pilipczuk, and Jakub Onufry Wojtaszczyk. Kernelization hardness of connectivity problems in  $d$ -degenerate graphs. *Discrete Applied Mathematics*, 160(15):2131–2141, 2012.

- 9 Anuj Dawar and Stephan Kreutzer. Domination problems in nowhere-dense classes. In *FSTTCS 2009*, volume 4 of *LIPICs*, pages 157–168. Schloss Dagstuhl — Leibniz-Zentrum für Informatik, 2009.
- 10 Reinhard Diestel. *Graph Theory, 4th Edition*, volume 173 of *Graduate texts in mathematics*. Springer, 2012.
- 11 Rodney G Downey and Michael R Fellows. *Fundamentals of parameterized complexity*, volume 4. Springer, 2013.
- 12 Rodney G Downey and Michael Ralph Fellows. *Parameterized complexity*. Springer, 1999.
- 13 Pål Grønås Drange, Markus Sortland Dregi, Fedor V. Fomin, Stephan Kreutzer, Daniel Lokshtanov, Marcin Pilipczuk, Michal Pilipczuk, Felix Reidl, Fernando Sánchez Villaamil, Saket Saurabh, Sebastian Siebertz, and Somnath Sikdar. Kernelization and sparseness: the case of dominating set. In *33rd Symposium on Theoretical Aspects of Computer Science, STACS 2016*, pages 31:1–31:14, 2016.
- 14 Kord Eickmeyer, Archontia C. Giannopoulou, Stephan Kreutzer, O-joung Kwon, Michal Pilipczuk, Roman Rabinovich, and Sebastian Siebertz. Neighborhood complexity and kernelization for nowhere dense classes of graphs. In *44th International Colloquium on Automata, Languages, and Programming, ICALP 2017*, pages 63:1–63:14, 2017.
- 15 Fedor V. Fomin, Daniel Lokshtanov, Saket Saurabh, and Dimitrios M. Thilikos. Bidimensionality and kernels. In *Proceedings of the Twenty-First Annual ACM-SIAM Symposium on Discrete Algorithms, Austin, Texas, USA*, pages 503–510, 2010.
- 16 Fedor V. Fomin, Daniel Lokshtanov, Saket Saurabh, and Dimitrios M. Thilikos. Linear kernels for (connected) dominating set on  $H$ -minor-free graphs. In *Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2012, Kyoto, Japan, January 17-19, 2012*, pages 82–93, 2012.
- 17 Fedor V. Fomin, Daniel Lokshtanov, Saket Saurabh, and Dimitrios M. Thilikos. Linear kernels for (connected) dominating set on graphs with excluded topological subgraphs. In *30th International Symposium on Theoretical Aspects of Computer Science, STACS 2013, February 27 - March 2, 2013, Kiel, Germany*, pages 92–103, 2013.
- 18 Jakub Gajarský, Petr Hliněný, Jan Obdržálek, Sebastian Ordyniak, Felix Reidl, Peter Rossmanith, Fernando Sánchez Villaamil, and Somnath Sikdar. Kernelization using structural parameters on sparse graph classes. *J. Comput. Syst. Sci.*, 84:219–242, 2017. doi:10.1016/j.jcss.2016.09.002.
- 19 Sarel Har-Peled and Kent Quanrud. Approximation algorithms for polynomial-expansion and low-density graphs, 2015. (CoRR/abs/1501.00721). URL: <http://arxiv.org/abs/1501.00721>.
- 20 Sarel Har-Peled and Kent Quanrud. Approximation algorithms for polynomial-expansion and low-density graphs. In Nikhil Bansal and Irene Finocchi, editors, *Algorithms - ESA 2015 - 23rd Annual European Symposium, Patras, Greece, September 14-16, 2015, Proceedings*, volume 9294 of *Lecture Notes in Computer Science*, pages 717–728. Springer, 2015.
- 21 Stephan Kreutzer, Roman Rabinovich, and Sebastian Siebertz. Polynomial kernels and wideness properties of nowhere dense graph classes. In *SODA*, pages 1533–1545. SIAM, 2017.
- 22 Daniel Lokshtanov, Matthias Mnich, and Saket Saurabh. A linear kernel for a planar connected dominating set. *Theor. Comput. Sci.*, 412(23):2536–2543, 2011.
- 23 Daniel Lokshtanov, Fahad Panolan, M. S. Ramanujan, and Saket Saurabh. Lossy kernelization. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017, Montreal, QC, Canada, June 19-23, 2017*, pages 224–237, 2017.
- 24 Neeldhara Misra, Geevarghese Philip, Venkatesh Raman, Saket Saurabh, and Somnath Sikdar. FPT algorithms for connected feedback vertex set. In *WALCOM: Algorithms and*

- Computation, 4th International Workshop, WALCOM 2010, Dhaka, Bangladesh, February 10-12, 2010. Proceedings*, pages 269–280, 2010.
- 25 Jaroslav Nešetřil and Patrice Ossona de Mendez. Grad and classes with bounded expansion I. Decompositions. *European Journal of Combinatorics*, 29(3):760–776, 2008.
  - 26 Jaroslav Nešetřil and Patrice Ossona de Mendez. Grad and classes with bounded expansion II. Algorithmic aspects. *European Journal of Combinatorics*, 29(3):777–791, 2008.
  - 27 Jaroslav Nešetřil and Patrice Ossona de Mendez. Grad and classes with bounded expansion III. Restricted graph homomorphism dualities. *European Journal of Combinatorics*, 29(4):1012–1024, 2008.
  - 28 Jaroslav Nešetřil and Patrice Ossona de Mendez. First order properties on nowhere dense structures. *The Journal of Symbolic Logic*, 75(03):868–887, 2010.
  - 29 Jaroslav Nešetřil and Patrice Ossona de Mendez. On nowhere dense graphs. *European Journal of Combinatorics*, 32(4):600–617, 2011.
  - 30 Jaroslav Nešetřil and Patrice Ossona de Mendez. *Sparsity — Graphs, Structures, and Algorithms*, volume 28 of *Algorithms and combinatorics*. Springer, 2012.
  - 31 Geevarghese Philip, Venkatesh Raman, and Somnath Sikdar. Polynomial kernels for dominating set in graphs of bounded degeneracy and beyond. *ACM Trans. Algorithms*, 9(1):11, 2012.
  - 32 Michał Pilipczuk, Sebastian Siebertz, and Szymon Toruńczyk. On wideness and stability. *arXiv preprint arXiv:1705.09336*, 2017.
  - 33 Norbert Sauer. On the density of families of sets. *Journal of Combinatorial Theory, Series A*, 13(1):145–147, 1972.
  - 34 Saharon Shelah. A combinatorial problem; stability and order for models and theories in infinitary languages. *Pacific Journal of Mathematics*, 41(1):247–261, 1972.