


On Kernelization for Edge Dominating Set under Structural Parameters

Eva-Maria C. Hols 

Department of Computer Science, Humboldt-Universität zu Berlin, Germany
hols@informatik.hu-berlin.de

Stefan Kratsch 

Department of Computer Science, Humboldt-Universität zu Berlin, Germany
kratsch@informatik.hu-berlin.de

Abstract

In the NP-hard EDGE DOMINATING SET problem (EDS) we are given a graph $G = (V, E)$ and an integer k , and need to determine whether there is a set $F \subseteq E$ of at most k edges that are incident with all (other) edges of G . It is known that this problem is fixed-parameter tractable and admits a polynomial kernelization when parameterized by k . A caveat for this parameter is that it needs to be large, i.e., at least equal to half the size of a maximum matching of G , for instances not to be trivially negative. Motivated by this, we study the existence of polynomial kernelizations for EDS when parameterized by *structural* parameters that may be much smaller than k .

Unfortunately, at first glance this looks rather hopeless: Even when parameterized by the deletion distance to a disjoint union of paths P_3 of length two there is no polynomial kernelization (under standard assumptions), ruling out polynomial kernelizations for many smaller parameters like the feedback vertex set size. In contrast, somewhat surprisingly, there is a polynomial kernelization for deletion distance to a disjoint union of paths P_5 of length four. As our main result, we fully classify for all finite sets \mathcal{H} of graphs, whether a kernel size polynomial in $|X|$ is possible when given X such that each connected component of $G - X$ is isomorphic to a graph in \mathcal{H} .

2012 ACM Subject Classification Mathematics of computing → Graph algorithms

Keywords and phrases Edge dominating set, kernelization, structural parameters

Digital Object Identifier 10.4230/LIPIcs.STACS.2019.36

Related Version A full version is available at [16], <https://arxiv.org/abs/1901.03582>.

Funding *Eva-Maria C. Hols*: Supported by DFG Emmy Noether-grant (KR 4286/1).

1 Introduction

In the EDGE DOMINATING SET problem (EDS) we are given a graph $G = (V, E)$ and an integer k , and need to determine whether there is a set $F \subseteq E$ of at most k edges that are incident with all (other) edges of G . It is known that this is equivalent to the existence of a maximal matching of size at most k . The EDGE DOMINATING SET problem is NP-hard but admits a simple 2-approximation by taking any maximal matching of G . It can be solved in time $\mathcal{O}^*(2.2351^k)^1$ [18], making it fixed-parameter tractable for parameter k . Additionally, for EDS any given instance (G, k) can be efficiently reduced to an equivalent one (G', k') with only $\mathcal{O}(k^2)$ vertices and $\mathcal{O}(k^3)$ edges [33] (this is called a *kernelization*).

The drawback of choosing the solution size k as the parameter is that k is large on many types of easy instances. This has been addressed for many other problems by turning to so called *structural parameters* that are independent of the solution size. Two lines of research in this direction have yielded polynomial kernelizations for several other NP-hard problems.

¹ \mathcal{O}^* -notation hides factors that are polynomial in the input size.



One possibility is to choose the parameter as the size of a set X such that $G - X$ belongs to some class \mathcal{C} where the problem in question can be efficiently solved; such sets X are called *modulators*. The other possibility is to parameterize above some lower bound for the solution, i.e., the parameter is the difference between the solution size k and the lower bound.

The VERTEX COVER problem, where, given a graph G and an integer k , we are asked whether there are k vertices that are incident with all edges, has been successfully studied under different structural parameters. It had been observed that VERTEX COVER is FPT parameterized by the size of a modulator to a class \mathcal{C} when one can solve vertex cover on graphs that belong to \mathcal{C} in polynomial time; e.g. if \mathcal{C} is the graph class of forests or, more generally, of bipartite or König graphs. Furthermore, there also exist kernelizations for VERTEX COVER parameterized by modulators to some graph classes \mathcal{C} . The first of a number of such results is due to Jansen and Bodlaender [19] who gave a kernelization with $\mathcal{O}(\ell^3)$ vertices where ℓ is the size of a (minimum) feedback vertex set of the input graph. Clearly, the solution size k cannot be bounded in terms of ℓ alone because forests already have arbitrarily large minimum vertex covers. This result has been generalized, e.g., for parameterization by the size of an odd cycle transversal [24].

There are also parameterized algorithms for VERTEX COVER above lower bounds that address the specific complaint about the seemingly unnecessarily large parameter value k in many graph classes. It was first shown that VERTEX COVER parameterized by $\ell = k - MM$ where MM stands for the size of a maximum matching is FPT [27]. In other words, the parameter value ℓ is the difference between k and the obvious lower bound. This has been improved to work also for parameterization by $\ell = k - LP$ where LP stands for the minimum *fractional vertex cover* (as determined by the LP relaxation) [5, 25] and, recently, even for parameter $\ell = k - (2LP - MM)$ [13]. All of these above lower bound parameterizations of VERTEX COVER also have randomized polynomial kernelizations [24, 23].

Motivated by the number of positive results for VERTEX COVER parameterized by structural parameters we would like to know whether some of these results carry over to the related but somewhat more involved EDGE DOMINATING SET problem.

Our results. For kernelization subject to the size of a modulator to some tractable class \mathcal{C} there is bad news: Even if \mathcal{C} contains only the disjoint unions of paths of length two (consisting of three vertices each) we show that there is no polynomial kernelization for parameterization by $|X|$ with $G - X \in \mathcal{C}$ unless $\text{NP} \subseteq \text{coNP}/\text{poly}$ (and the polynomial hierarchy collapses). The same is true when \mathcal{C} contains at least all disjoint unions of triangles. Thus, for the usual program of studying modulators to well-known *hereditary* graph classes \mathcal{C} there is essentially nothing left to do because the only permissible connected components would have one or two vertices.² That said, as the next result shows, this perspective would ignore an interesting landscape of positive and negative results that can be obtained by permitting certain forms of connected components in $G - X$ but not necessarily all induced subgraphs thereof, i.e., by dropping the requirement that \mathcal{C} needs to be hereditary (closed under induced subgraphs).

Indeed, there is, e.g., a polynomial kernelization for parameter $|X|$ when all connected components of $G - X$ are paths of length *four*. This indicates that the structure even of constant-sized components permitted in $G - X$ determines in a nontrivial way whether or not there is a polynomial kernelization. Note the contrast with VERTEX COVER where a modulator to component size d admits a kernelization with $\mathcal{O}(k^d)$ vertices for each fixed d . Naturally, we are interested in finding out exactly which cases admit polynomial kernelizations.

² This very modest case actually admits a polynomial kernelization.

This brings us to our main result. For \mathcal{H} a set of graphs, say that G is an \mathcal{H} -component graph if each connected component of G is isomorphic to some graph in \mathcal{H} . We fully classify the existence of polynomial kernelizations for parameterization by the size of a modulator to the class of \mathcal{H} -component graphs for all finite sets \mathcal{H} . To clarify, the input consists of (G, k, X) such that $G - X$ is an \mathcal{H} -component graph and the task is to determine whether G has an edge dominating set of size at most k ; the parameter is $|X|$. Note that these problems are fixed-parameter tractable for all finite sets \mathcal{H} because G has treewidth at most $|X| + \mathcal{O}(1)$.

► **Theorem 1.** *For every finite set \mathcal{H} of graphs, the EDGE DOMINATING SET problem parameterized by the size of a given modulator X to the class of \mathcal{H} -component graphs falls into one of the following two cases:*

1. *It has a kernelization with $\mathcal{O}(|X|^d)$ vertices, $\mathcal{O}(|X|^{d+1})$ edges, and size $\mathcal{O}(|X|^{d+1} \log |X|)$. Moreover, unless $\text{NP} \subseteq \text{coNP/poly}$, there is no kernelization to size $\mathcal{O}(|X|^{d-\varepsilon})$ for any $\varepsilon > 0$. Here $d = d(\mathcal{H})$ is a constant depending only on the set \mathcal{H} .*
2. *It has no polynomial kernelization unless $\text{NP} \subseteq \text{coNP/poly}$.*

To obtain the classification one needs to understand how connected components of $G - X$ that are isomorphic to some graph $H \in \mathcal{H}$ can interact with a solution for G , and to derive properties of H that can be leveraged for kernelizations or lower bounds for kernelization. Crucially, edge dominating sets for G may contain edges between X and components of $G - X$. From the perspective of such a component (isomorphic to H) this is equivalent to first covering edges incident with some vertex set $B \subseteq V(H)$ (the endpoints of chosen edges to X) and then covering the remaining edges by a minimum edge dominating set for $H - B$. Depending on the size of a minimum edge dominating set of $H - B$ and further properties of H , such a set B may be used to rule out any polynomial kernelizations or to give a lower bound of $\mathcal{O}(|X|^{d-\varepsilon})$ for the kernel size, where $d = |B|$. Conversely, absence of such sets or an upper bound for their size can be leveraged for kernelizations. Some sets B may make others redundant, further complicating both upper and lower bounds.

For a given finite set \mathcal{H} of graphs, the lower bound obtained from the classification is simply the strongest one over all $H \in \mathcal{H}$. If this does not already rule out a polynomial kernelization then, for each $H \in \mathcal{H}$, we can reduce the number of components isomorphic to H to $\mathcal{O}(|X|^{d(H)})$ where $d(H)$ depends only on H . Moreover, we also have the almost matching lower bound of $\mathcal{O}(|X|^{d(H)-\varepsilon})$, assuming $\text{NP} \not\subseteq \text{coNP/poly}$. The value $d(\mathcal{H})$ is the maximum over all $d(H)$ for $H \in \mathcal{H}$ that yield such a polynomial lower bound; it can be computed in time depending only on \mathcal{H} , i.e., in constant time for each fixed \mathcal{H} .

Regarding parameterization above lower bounds, we prove that it is NP-hard to determine whether a graph G has an edge dominating set of size equal to the lower bound of half the size of a maximum matching. This rules out any positive results for parameter $\ell = k - \frac{1}{2}MM$.

Related work. The parameterized complexity of EDGE DOMINATING SET has been studied in a number of papers [9, 10, 31, 32, 33, 34, 8, 18]. Structural parameters were studied, e.g., by Escoffier et al. [8] who obtained an $\mathcal{O}^*(1.821^\ell)$ time algorithm where ℓ is the vertex cover size of the input graph, and by Kobler and Rotics [22] who gave a polynomial-time algorithm for graphs of bounded clique-width. It is easy to see that EDS is fixed-parameter tractable with respect to the treewidth of the input graph. Prieto [26] was the first to find a kernelization to $\mathcal{O}(k^2)$ vertices for the standard parameterization by k ; this was improved to $\mathcal{O}(k^2)$ vertices and $\mathcal{O}(k^3)$ edges by Xiao et al. [33] and further tweaked by Hagerup [15]. Our work appears to be the first to study the existence of polynomial kernelizations for EDS subject to structural parameters, though some lower bounds, e.g., for parameter treewidth are obvious.

Classically, EDGE DOMINATING SET remains NP-hard on planar cubic graphs, bipartite graphs with maximum degree three [36]. This implies NP-hardness already for $|X| = 0$ when considering parameterization by a modulator to any graph class containing this special case. EDGE DOMINATING SET has also been studied from the perspective of approximation [12, 4, 3, 29, 8], enumeration [20, 14, 21], and exact exponential-time algorithms [28, 32, 30, 35].

Organization. We begin with some preliminaries in Section 2. Section 3 provides some intuition for the main result by proving the lower bound for EDGE DOMINATING SET parameterized by the size of a modulator to a P_3 -component graph as well as the polynomial kernelization for parameterization by the size of a modulator to a P_5 -component graph. Section 4 gives a detailed statement of the main result including the required definitions to determine which result applies for any given set \mathcal{H} . Due to space restrictions, the proof of the main result and the hardness proof for parameter $\ell = k - \frac{1}{2}MM$ are deferred to the full version of this work. We conclude in Section 5.

2 Preliminaries

We use standard graph notation as given by Diestel [7]. In particular, for a graph $G = (V, E)$ we let $N(v) = \{u \in V \mid \{u, v\} \in E\}$ and $N[v] = N(v) \cup \{v\}$; similarly, $N[X] = \bigcup_{x \in X} N[x]$ and $N(X) = N[X] \setminus X$. We let $E(X, Y) = \{\{x, y\} \mid x \in X, y \in Y\}$ and we let $\delta(v) = \{\{u, v\} \mid u \in V, \{u, v\} \in E\}$. By $G[X]$ we denote the induced subgraph of G on vertex set X and by $G - X$ the induced subgraph on vertex set $V \setminus X$; we let $G - v = G - \{v\}$. We denote the size of a minimum edge dominating set of a graph G by $\text{EDS}(G)$.

Let \mathcal{H} be a set of graphs. We say that a graph G is an \mathcal{H} -component graph if each connected component of G is isomorphic to some graph in \mathcal{H} . Clearly, disconnected graphs in \mathcal{H} do not affect which graphs G are \mathcal{H} -component graphs and, thus, our proofs need only consider the connected graphs $H \in \mathcal{H}$. We write H -component graph rather than $\{H\}$ -component graph for single (connected) graphs H .

Let $[n]$ denote the set $\{1, 2, \dots, n\}$.

Parameterized complexity. A *parameterized problem* \mathcal{Q} is a subset of $\Sigma^* \times \mathbb{N}$ where Σ is any finite set. The second component k of instances (x, k) is called the *parameter*. A parameterized problem \mathcal{Q} is *fixed-parameter tractable* if there is an algorithm that correctly solves all instances (x, k) in time $f(k)|x|^c$ where f is a computable function and c is a constant independent of k . A *kernelization* for \mathcal{Q} is an efficient algorithm that, given an instance (x, k) , takes time polynomial in $|x| + k$ and returns an instance (x', k') of size at most $f(k)$ such that $(x, k) \in \mathcal{Q}$ if and only if $(x', k') \in \mathcal{Q}$ where f is a computable function. The function f is also called the *size* of the kernelization and a kernelization is polynomial (resp. linear) if $f(k)$ is polynomially (resp. linearly) bounded in k .

We use the notion of a cross-composition [2], which is a convenient front-end for the seminal kernel lower bound framework of Bodlaender et al. [1] and Fortnow and Santhanam [11]. A relation $\mathcal{R} \subseteq \Sigma^* \times \Sigma^*$ is a *polynomial equivalence relation* if equivalence of two strings $x, y \in \Sigma^*$ can be tested in time polynomial in $|x| + |y|$ and if \mathcal{R} partitions any finite set $S \subseteq \Sigma^*$ into a number of classes that is polynomially bounded in the largest element of S .

► **Definition 2** ((OR-)cross-composition [2]). *Let $L \subseteq \Sigma^*$ be a language, let \mathcal{R} be a polynomial equivalence relation on Σ^* , and let $\mathcal{Q} \subseteq \Sigma^* \times \mathbb{N}$ be a parameterized problem. An (OR-)cross-composition of L into \mathcal{Q} (with respect to \mathcal{R}) is an algorithm that, given t instances $x_1, \dots, x_t \in$*

Σ^* of L belonging to the same equivalence class of \mathcal{R} , takes time polynomial in $\sum_{i=1}^t |x_i|$ and outputs an instance $(y, k) \in \Sigma^* \times \mathbb{N}$ such that the following hold:

- “PB”: The parameter value k is polynomially bounded in $\max_{i=1}^t |x_i| + \log t$.
- “OR”: The instance (y, k) is yes for \mathcal{Q} if and only if at least one instance x_i is yes for L . An (OR-)cross-composition of L into \mathcal{Q} of cost $f(t)$ instead satisfies “OR” and “CB”:
- “CB”: The parameter value k is bounded by $\mathcal{O}(f(t) \cdot (\max_{i=1}^t |x_i|)^c)$, where c is some constant independent of t .

If L is NP-hard then both forms of cross-compositions are known to imply lower bounds for kernelizations for \mathcal{Q} . Theorem 4 additionally builds on Dell and van Melkebeek [6].

► **Theorem 3** ([2, Corollary 3.6.]). *If an NP-hard language L has a cross-composition to \mathcal{Q} then \mathcal{Q} admits no polynomial kernelization or polynomial compression unless $\text{NP} \subseteq \text{coNP/poly}$.*

► **Theorem 4** ([2, Theorem 3.8.]). *Let $d, \varepsilon > 0$. If an NP-hard language L has a cross-composition into \mathcal{Q} of cost $f(t) = t^{1/d+o(1)}$, where t is the number of instances, then \mathcal{Q} has no polynomial kernelization or polynomial compression of size $\mathcal{O}(k^{d-\varepsilon})$ unless $\text{NP} \subseteq \text{coNP/poly}$.*

All our composition-based proofs use for L the NP-hard MULTICOLORED CLIQUE problem. Therein we are given a graph $G = (V, E)$, an integer k , and a partition of V into k sets V_1, \dots, V_k of equal size; we need to determine whether there is a clique of size k in G that contains exactly one vertex from each set V_i . Such a set X is called a *multicolored k -clique*.

3 EDS parameterized by the size of a modulator to a P_3 - resp. P_5 -component graph

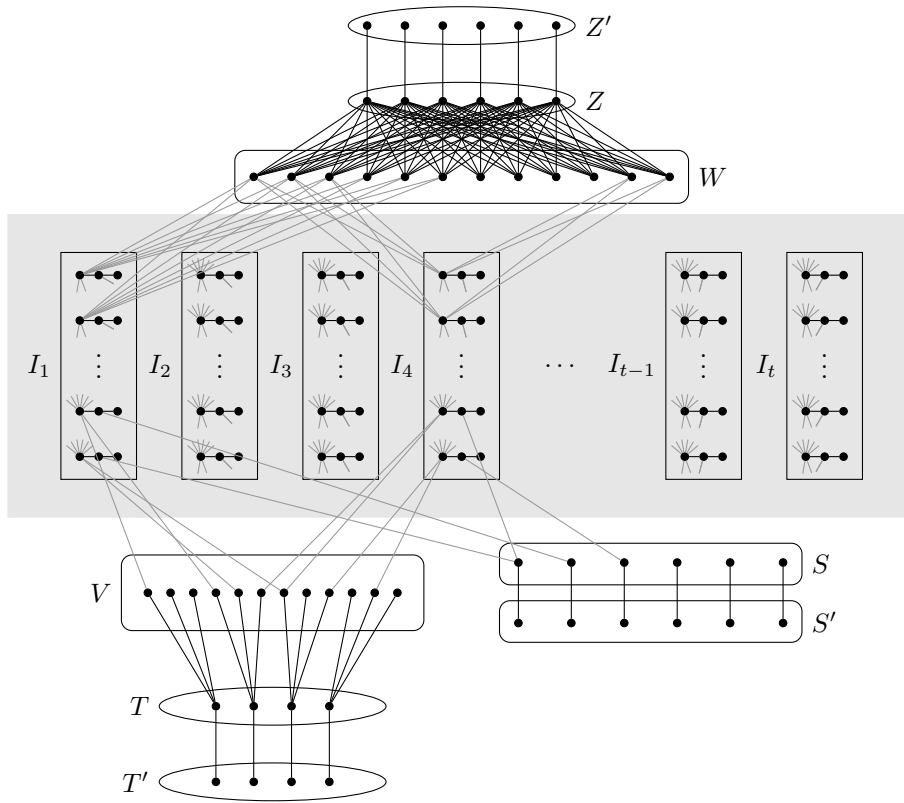
In this section we study the difference of EDGE DOMINATING SET parameterized by the size of a modulator to a P_3 -component graph and EDGE DOMINATING SET parameterized by the size of a modulator to a P_5 -component graph, which are both more restrictive than parameterization by size of a feedback vertex set (modulator to a forest). Note that the latter is FPT, because the treewidth is at most the size of the feedback vertex set plus one and EDGE DOMINATING SET parameterized by the treewidth is FPT. Hence, EDGE DOMINATING SET parameterized by the above modulators is FPT too.

First, we show that EDGE DOMINATING SET parameterized by the size of a modulator to a P_3 -component graph has no polynomial kernelization unless $\text{NP} \subseteq \text{coNP/poly}$. This rules out polynomial kernelizations for a large number of interesting parameters like feedback vertex set size or size of a modulator to a linear forest. Somewhat surprisingly, we then show that when parameterized by the modulator to a P_5 -component graph we do get a polynomial kernelization.

3.1 Lower bound for EDS parameterized by the size of a modulator to a P_3 -component graph

We give a kernelization lower bound for EDGE DOMINATING SET parameterized by the size of a modulator X , such that deleting X results in a disjoint union of P_3 's. To prove this we give a cross-composition from MULTICOLORED CLIQUE.

► **Theorem 5.** *EDGE DOMINATING SET parameterized by the size of a modulator to a P_3 -component graph (and thus also parameterized by the size of a modulator to a linear forest) does not admit a polynomial kernelization unless $\text{NP} \subseteq \text{coNP/poly}$.*



■ **Figure 1** Construction of the graph G' with $k = 4$, where $X' = W \cup Z \cup Z' \cup V \cup T \cup T' \cup S \cup S'$.

Proof. To prove the theorem we give a cross-composition from the NP-hard MULTICOLORED CLIQUE problem to EDGE DOMINATING SET parameterized by the size of a modulator to a P_3 -component graph. Input instances are of the form (G_i, k_i) where G_i comes with a partition of the vertex set into k color classes. (Since the color classes are of equal size it holds that $k \leq |V(G_i)|$.) For the polynomial equivalence relation \mathcal{R} we take the relation that puts two instances $(G_1, k_1), (G_2, k_2)$ of MULTICOLORED CLIQUE in the same equivalence class if $k_1 = k_2$ and $|V(G_1)| = |V(G_2)|$. It is easy to check that \mathcal{R} is a polynomial equivalence relation. (Instances with size at most N have at most N vertices. Thus, we get at most N^2 classes for instances of size at most N .)

Let a sequence of instances $I_i = (G_i, k)_{i=1}^t$ of MULTICOLORED CLIQUE be given that are equivalent under \mathcal{R} . We identify the color classes of the input graphs so that all graphs have the same vertex set V and the same color classes V_1, V_2, \dots, V_k . Let $n := |V_i|$ be the number of vertices of each color class; thus, each instance has $|V| = n \cdot k$ vertices. We assume w.l.o.g. that every instance has at least one edge in $E(V_p, V_q)$ for all $1 \leq p < q \leq k$; otherwise, this instance would be a trivial no instance and we can delete it. Furthermore, we can assume w.l.o.g. that $t = 2^s$ for an integer s , since we may copy some instances if needed (while at most doubling the number of instances and increasing $\log t$ by less than one).

Now, we construct an instance (G', k', X') of EDGE DOMINATING SET parameterized by the size of a modulator to a P_3 -component graph, where the size of X' is polynomially bounded in $n + k + s$ (see Figure 1 for an illustration). We add a set V consisting of $k \cdot n$ vertices to graph G' which represents the vertices of the t instances. The set V is partitioned into the k color classes V_1, V_2, \dots, V_k . To choose which vertices are contained in a clique of

size k , we add a set $T = \{t_1, t_2, \dots, t_k\}$ and a set $T' = \{t'_1, t'_2, \dots, t'_k\}$, each of size k , to G' . We make $t_j \in T$, with $j \in [k]$, adjacent to all vertices in V_j and to vertex $t'_j \in T'$. Next, we add two sets Z, Z' , each of size s , and a set W of size $2s$ to G' and add edges to G' such that each vertex in Z has exactly one private neighbor in Z' and is adjacent to all vertices in W . The set W contains $\binom{2s}{s} \geq 2^s$ different subsets of size s . For each instance (G_i, k) , with $i \in [t]$, we pick a different subset of size s of W and denote it by $W(i)$. For all $1 \leq p < q \leq k$ we add a vertex $s_{p,q}$ and a vertex $s'_{p,q}$ to G' ; these will correspond to the edge sets $E(V_p, V_q)$. Let $S = \{s_{p,q} \mid 1 \leq p < q \leq k\}$ and $S' = \{s'_{p,q} \mid 1 \leq p < q \leq k\}$. We make vertex $s_{p,q}$ adjacent to vertex $s'_{p,q}$ for all $1 \leq p < q \leq k$. For each graph G_i , for $i \in [t]$, we add $|E(G_i)|$ paths of length two to the graph G' ; every P_3 represents exactly one edge of the graph G_i . Let $P_i^e = u_{i,1}^e u_i^e u_{i,2}^e$ denote the path of instance $i \in [t]$ that represents edge $e \in E(G_i)$. Finally, we make vertices in P_i^e , with $i \in [t]$ and $e \in E(G_i)$, adjacent to vertices in the sets W, V , and S as follows: We make vertex $u_{i,1}^e$ of path P_i^e , with $i \in [t]$, which represents edge $e = \{x, y\} \in E(G_i)$ adjacent to the vertices x, y in V and to all vertices in the set $W(i) \subseteq W$. Additionally, we make vertex u_i^e adjacent to vertex $s_{p,q}$ where $1 \leq p < q \leq k$ such that $e \in E(V_p, V_q)$.

The set X' is defined to contain all vertices that do not participate in the paths P_i^e , i.e., $X' = W \cup Z \cup Z' \cup V \cup T \cup T' \cup S \cup S'$. Clearly, $G - X'$ is a P_3 -component graph and $|X'| = 4s + k \cdot n + 2k + 2 \cdot \binom{k}{2}$. Let $k' = k + s + \sum_{i=1}^t |E(G_i)|$. Note that the size of k' can depend linearly on the number of instances, because our parameter is the size of X' , which is polynomially bounded in $n + s$, as $k \leq n$. We return the instance (G', k', X') ; clearly, this instance can be generated in polynomial time.

Now, we have to show that (G', k', X') is a YES-instance of EDS if and only if there exists an $i^* \in [t]$ such that (G_{i^*}, k) is a YES-instance of MULTICOLORED CLIQUE.

(\Rightarrow): Assume first that (G', k', X') is yes for EDS and that there exists an edge dominating set F of size at most k' in G' . We can always pick F such that it fulfills the following properties (most hold for all solutions of size at most k'):

1. The vertex sets S, T , and Z must be subsets of $V(F)$: E.g., for each edge $\{z, z'\}$ with $z \in Z$ and $z' \in Z'$ the set $V(F)$ must contain z or z' ; if it contains z' then $\{z, z'\} \in F$ as it is the only edge incident with z' ; either way we get $z \in V(F)$. The same applies for S and S' , and for T and T' .
2. Because $S, T, Z \subseteq V(F)$ but $S \cup T \cup Z$ is an independent set, the set F must contain at least $|S|$ edges incident with S , $|T|$ edges incident with T , and $|Z|$ edges incident with Z . By straightforward replacement arguments we may assume that F contains exactly the following edges incident with $S \cup T \cup Z$: $|T|$ edges between T and V , $|Z|$ edges between Z and W , and $|S|$ edges between S and middle vertices u_i^e of P_3 's in $G' - X'$. Furthermore, we can assume that these edges are a matching, because no color class is empty, no edge set $E(V_p, V_q)$ is empty, and Z is adjacent to all vertices in W .
3. For each $P_i^e = u_{i,1}^e u_i^e u_{i,2}^e$, which represents the edge e of instance (G_i, k) , at least vertex u_i^e must be an endpoint of an edge in F : Indeed, to cover the edge $\{u_i^e, u_{i,2}^e\}$ one of its two vertices must be in $V(F)$. Similar to Property 1 above, if $u_{i,2}^e \in V(F)$ then F must contain its sole incident edge $\{u_i^e, u_{i,2}^e\}$ and, hence, $u_i^e \in V(F)$.
4. An edge in F cannot have its endpoints in two different P_3 's of $G' - X'$ because no such edges exist.

Let $F_T = F \cap E(T, V)$, let $F_Z = F \cap E(Z, W)$, let $F_S = F \cap E(S, \{u_i^e \mid i \in [t], e \in E(G_i)\})$, and let $F_R = F \setminus (F_T \cup F_Z \cup F_S)$. Hence, due to Properties 1 and 2, we have

$$|F_R| \leq k' - |F_T| - |F_Z| - |F_S| \leq \sum_{i=1}^t |E(G_i)| - \binom{k}{2}.$$

By Property 3, all vertices u_i^e are endpoints of edges in F . Among $F_T \cup F_Z \cup F_S$ this can only be true for the $|S| = \binom{k}{2}$ edges in F_S . Since there are exactly $\sum_{i=1}^t |E(G_i)|$ vertices u_i^e , which is (greater or) equal to $|F_R| + |F_S|$, and there are no edges connecting different such vertices, each edge in $F_R \cup F_S$ is incident with a private vertex u_i^e . This also implies that all edges in F_R have no endpoints in $V \cup W$ as those sets are not adjacent to any vertex u_i^e . Thus, in W exactly the $|Z| = s$ endpoints of F_Z are endpoints of F . Similarly, in V exactly the $|T| = k$ endpoints of F_T are endpoints of F ; let $X \subseteq V$ denote this set of k vertices. Observe that by construction of G' the set X contains exactly one vertex from each color class, because $t_j \in T$, for $j \in [k]$, is only adjacent to vertices of V_j .

Now, consider any path $P_i^e = u_{i,1}^e u_i^e u_{i,2}^e$ where u_i^e is an endpoint of an edge $f \in F_S$. Clearly, the other endpoint of f lies in S , and, by the above accounting, no other edge of F is incident with $u_{i,1}^e$ or $u_{i,2}^e$. In particular, this implies that all neighbors of $u_{i,1}^e$ in W and V must be endpoints of edges in F . If $e = \{x, y\}$ then these neighbors of $u_{i,1}^e$ are the set $W(i) \subseteq W$ and the vertices $x, y \in V$, and, by construction of G' , the edge $\{x, y\}$ must exist in G_i . Thus, $W(i) \cup \{x, y\} \subseteq V(F)$ which implies that $x, y \in X$.

Repeating this argument for all $|S| = \binom{k}{2}$ paths of this type, we can conclude the following: (1) All paths correspond to the same instance $i^* \in [t]$ because we require $W(i) \subseteq V(F)$, but exactly $|Z| = |W(i^*)| = s$ such vertices are in $V(F)$. (Different values of i would require different sets $W(i)$, exceeding size s .) (2) There are $\binom{k}{2}$ edges of G_{i^*} represented by the paths and all their endpoints must be in $X = V \cap V(F)$. Since $|X| = k$, the edges must form a clique of size k on vertex set X in G_{i^*} . We already observed above that X contains exactly one vertex per color class, hence, instance (G_{i^*}, k) is yes, as claimed.

(\Leftarrow): For the other direction, assume that for some $i^* \in [t]$ the MULTICOLORED CLIQUE instance (G_{i^*}, k) is a YES-instance. Let $X = \{x_1, x_2, \dots, x_k\} \subseteq V$ be a multicolored clique of size k in G_{i^*} with $x_j \in V_j$ for $j \in [k]$, let E' be the set of edges of the clique X , and let $e_{p,q} = \{x_p, x_q\}$ for $1 \leq p < q \leq k$. We construct an edge dominating set F of G' of size at most k' as follows: First we add the k edges $\{t_j, x_j\}$ for $j \in [k]$ between T and $X \subseteq V$; thus, $T \cup X \subseteq V(F)$. We then add a maximum matching (of size s) between $W(i^*) \subseteq W$ and Z to the set F . This matching saturates $W(i^*)$ and Z because $|Z| = |W(i^*)| = s$; thus, $W(i^*) \cup Z \subseteq V(F)$. Next, we add the edges $\{u_{i^*,1}^{e_{p,q}}, s_{p,q}\}$ for all edges $e_{p,q} \in E'$, with $1 \leq p < q \leq k$, to the set F ; hence $S \subseteq V(F)$. Finally, for all other paths P_i^e , with $i \in [t]$, $e \in E(G_i)$, and $i \neq i^*$ or $e \notin E'$, we add the edge $\{u_{i,1}^e, u_i^e\}$ to F . (We have thus selected exactly one edge incident with each path of $G' - X'$.) By construction, it holds that $|F| = k + s + \sum_{i=1}^t |E(G_i)| = k'$.

It remains to show that F is indeed an edge dominating set of G' . To prove this, it suffices to show that $V(G') - V(F)$ is an independent set in G' . We already know that $S \cup T \cup W(i^*) \cup X \cup Z \subseteq V(F)$. Moreover, $V(F)$ contains the middle vertex u_i^e for all P_3 's in $G' - X'$ and it contains $u_{i,1}^e$ for all P_3 's that do not correspond to an edge of the clique X (i.e., with $i \neq i^*$ or with $i = i^*$ but $e \neq e_{p,q}$ for any $1 \leq p < q \leq k$). The sets S' , T' , and Z' are independent sets whose neighborhoods S , T , and Z are subsets of $V(F)$. Similarly, all vertices $u_{i,2}^e$ have their single neighbor u_i^e in $V(F)$. Thus, only vertices in $W \setminus W(i^*)$ and $V \setminus X$ could possibly be adjacent to vertices $u_{i^*,1}^{e_{p,q}}$, which correspond to the edges of $G_{i^*}[X]$, in $G' - V(F)$, but this can be easily refuted: Indeed, each $u_{i^*,1}^{e_{p,q}}$ is adjacent only to x_p and x_q in V , which are both in $X \subseteq V(F)$, and to the vertices in $W(i^*)$ in W , but $W(i^*) \subseteq V(F)$ as well. Thus $V(G') - V(F)$ is an independent set in G' and hence F is an edge dominating set for G' of size at most k' . Thus, (G', k', X') is yes, which completes the cross-composition.

By Theorem 3 the cross-composition from MULTICOLORED CLIQUE implies the claimed lower bound for kernelization. \blacktriangleleft

We proved that EDGE DOMINATING SET parameterized by the size of a modulator to a P_3 -component graph has no polynomial kernelization unless $\text{NP} \subseteq \text{coNP}/\text{poly}$. A similar proof establishes the same lower bound for modulators to K_3 -component graphs. As mentioned in the introduction this rules out polynomial kernels using modulators to essentially all interesting hereditary graph classes.³

3.2 Polynomial kernelization for EDS parameterized by the size of a modulator to a P_5 -component graph

To illustrate why other, non-hereditary, sets \mathcal{H} may well allow polynomial kernelizations for parameterization by the size of a modulator X to an \mathcal{H} -component graph, we sketch a simple kernelization for the case of $\mathcal{H} = \{P_5\}$, i.e., when components of $G - X$ are isomorphic to the path of length four. This does not use the full generality of the kernelization obtained in Section 4 because P_5 does not have any (later called) uncovered vertices or (later called) strongly beneficial sets (which are the main source of complication).

For the kernelization we need the following theorem which is due to Hopcroft and Karp [17]. The second claim of the theorem is not standard (but well known).

► **Theorem 6** ([17]). *Let G be an undirected bipartite graph with partition R and S , on n vertices and m edges. Then we can find a maximum matching of G in time $\mathcal{O}(m\sqrt{n})$. Furthermore, in time $\mathcal{O}(m\sqrt{n})$ we can find either a maximum matching that saturates R or a set $Y \subseteq R$ such that $|N_G(Y)| < |Y|$ and such that there exists a maximum matching M in $G - N_G[Y]$ that saturates $R \setminus Y$.*

► **Theorem 7.** *EDGE DOMINATING SET parameterized by the size of a given modulator X to a P_5 -component graph admits a kernelization with $\mathcal{O}(|X|)$ vertices.*

Proof. Let (G, k, X) be an instance of EDGE DOMINATING SET parameterized by the size of a modulator to a P_5 -component graph, and let \mathcal{C} be the set of connected components of $G - X$. We construct a bipartite graph G_B where one part is the set X , the other part consists of one vertex s_P for every connected component P in \mathcal{C} , and where there is an edge between $x \in X$ and s_P with $P = w_1w_2w_3w_4w_5 \in \mathcal{C}$ if and only if x is adjacent to a vertex of P that is not the middle vertex w_3 . Now, we apply Theorem 6 to obtain either a maximum matching in G_B that saturates X or a set $Y \subseteq X$ such that $|N_{G_B}(Y)| < |Y|$ and such that there exists a maximum matching in $G_B - N_{G_B}[Y]$ that saturates $X \setminus Y$. If there exists a maximum matching in G_B that saturates X then let $X_1 = X$ and $X_2 = \emptyset$. Otherwise, if there exists a set Y with the above properties then let $X_1 = X \setminus Y$ and $X_2 = Y$. Observe that X_2 also contains the vertices in X that are only adjacent to middle vertices of components in \mathcal{C} , and the vertices in X that are not adjacent to any component in \mathcal{C} . Let M be a maximum matching in $G_B - N_{G_B}[X_2]$ that saturates X_1 . The partition $X_1 \dot{\cup} X_2$ of X fulfills the following properties:

- Let \mathcal{C}_2 be the set of connected components P in \mathcal{C} where s_P is a vertex in $N_{G_B}(X_2)$, i.e., $\mathcal{C}_2 = \{P = w_1w_2w_3w_4w_5 \in \mathcal{C} \mid N_G(\{w_1, w_2, w_4, w_5\}) \cap X_2 \neq \emptyset\}$. It holds either that \mathcal{C}_2 is the empty set (when $X_2 = \emptyset$) or that it contains less than $|X_2|$ connected components of \mathcal{C} , i.e., $|\mathcal{C}_2| < |X_2|$ (when $Y = X_2 \neq \emptyset$).

³ It certainly does completely settle the question for modulators to \mathcal{H} -component graphs for all hereditary classes \mathcal{H} . If \mathcal{H} contains any connected graph with at least three vertices then we get a lower bound; else all connected components have one or two vertices and there is a polynomial kernelization.

36:10 Edge Dominating Set

- For every vertex $x \in X_1$, let $P_x = w_1^x w_2^x w_3^x w_4^x w_5^x$ be the connected component in $\mathcal{C}_1 := \mathcal{C} \setminus \mathcal{C}_2$ that is paired to x by M , i.e., $\{x, s_{P_x}\} \in M$. It holds that there exists a vertex $w^x \in \{w_1^x, w_2^x, w_4^x, w_5^x\}$ such that $\{w^x, x\} \in E(G)$ (definition of G_B). Note that \mathcal{C}_1 also contains all connected components that are not adjacent to any vertex in X or where only the middle vertex of a path in \mathcal{C} is adjacent to a vertex in X .

Using the above partition, one can show that there exists an optimum solution S that contains for each path P_x with $x \in X_1$ the locally optimal solution $\{\{x, w^x\}, \{w_3^x, w_2^x\}\}$ resp. $\{\{x, w^x\}, \{w_3^x, w_4^x\}\}$ depending on whether $w^x \in \{w_4^x, w_5^x\}$ or $w^x \in \{w_1^x, w_2^x\}$. More generally, for every vertex w of a path $P \in \mathcal{C}$, except the middle vertex, and every vertex $x \in X$ that is adjacent to w there exists a local optimum solution to P that uses edge $\{w, x\}$ and has the middle vertex of P as an endpoint of the second solution edge. This is the crucial difference to a path $P' = v_1 v_2 v_3$ of length two. Here, the only locally optimal solution that dominates P' and contains an edge between P' and X is $\{v_2, x\}$ with $x \in X$, but this local solution does not contain the vertices v_1 and v_3 . We used this in our lower bound construction to control which P_3 's may be used to “buy” vertices in X .

► **Reduction Rule 1.** Delete X_1 from G , i.e., let $G' = G - X_1$, $X' = X \setminus X_1 = X_2$, and $k' = k$.

▷ **Claim 8.** Reduction Rule 1 is safe.

Proof. Let F be an edge dominating set of size at most k in G . We construct an edge dominating set F' of size at most $k' = k$ in G' by deleting every edge $e = \{x, y\} \in F$ if both endpoints of e are contained in X_1 , or if exactly one endpoint is contained in X_1 and the other endpoint is isolated in G' ; and by replacing every edge $e = \{x, y\} \in F$ with $x \in X_1$ and $y \notin X_1$ by exactly one edge in $\delta_{G'}(y)$ if $\delta_{G'}(y) \neq \emptyset$. It holds that F' has size at most $k = k'$ because we either delete edges in F or replace them one for one by a new edge. Since every vertex in $V(G') \cap V(F)$ is either contained in $V(F')$ or isolated in G' it holds that F' is an edge dominating set in G' .

For the other direction, let F' be an edge dominating set of size at most k' in G' . Consider the path $P_x = w_1^x w_2^x w_3^x w_4^x w_5^x$ for some vertex $x \in X_1$. It holds that the only vertex in P_x that can be adjacent to a vertex in $X' = X \setminus X_1 = X_2$ is vertex w_3^x ; otherwise P_x would be a component in \mathcal{C}_2 and not in \mathcal{C}_1 (by definition of \mathcal{C}_1 and \mathcal{C}_2). Furthermore, the edge dominating set F' must dominate the two non-adjacent edges $\{w_1^x, w_2^x\}$ and $\{w_4^x, w_5^x\}$. Since w_1^x, w_2^x, w_4^x , and w_5^x are only adjacent to vertices in P_x the set F' must contain one of the two edges $e_{1,2}^x = \{w_1^x, w_2^x\}$, $e_{2,3}^x = \{w_2^x, w_3^x\}$ and one of the two edges $e_{3,4}^x = \{w_3^x, w_4^x\}$, $e_{4,5}^x = \{w_4^x, w_5^x\}$. To obtain an edge dominating set of size at most k in G we replace for each vertex $x \in X_1$ these edges with the local optimum solution $\{\{x, w^x\}, \{w_3^x, w_2^x\}\}$ resp. $\{\{x, w^x\}, \{w_3^x, w_4^x\}\}$ depending whether $w^x \in \{w_4^x, w_5^x\}$ or $w^x \in \{w_1^x, w_2^x\}$. It holds that $|F| \leq |F'|$ because for every vertex $x \in X_1$ we replace the at least two edges in $F' \cap \{e_{1,2}^x, e_{2,3}^x, e_{3,4}^x, e_{4,5}^x\}$ by the two edges of the locally optimal solution $\{\{x, w^x\}, \{w_3^x, w_2^x\}\}$ resp. $\{\{x, w^x\}, \{w_3^x, w_4^x\}\}$.

It remains to show that F is indeed an edge dominating set in G . The set $V(F)$ contains all vertices in $V(F')$, except some vertices in the connected components P_x with $x \in X_1$ where we change the edge dominating set F' . Furthermore, $V(F)$ contains all vertices in X_1 because for every vertex $x \in X_1$ the edge $\{w^x, x\}$ is contained in F . Thus, the only edges that are possibly not dominated by F have one endpoint in a path P_x with $x \in X_1$. Since w_3^x is contained in $V(F)$ (by construction), since every edge in P_x is dominated by F (by construction), and since the vertices in $\{w_1^x, w_2^x, w_4^x, w_5^x\}$ are only adjacent to vertices in $P_x \cup X_1$, it follows that F is an edge dominating set in G . \square

After applying Reduction Rule 1 it holds that for each path $P = w_1w_2w_3w_4w_5 \in \mathcal{C}_1$ only the vertex w_3 can be adjacent to a vertex in X , and we can assume that every (optimum) solution contains the edges $\{w_2, w_3\}$ and $\{w_3, w_4\}$. Additionally, one can show that there exists an optimum solution that does not contain any edge between \mathcal{C}_1 and X because we can replace any such edge $e = \{x, v\}$ with $v \in V(\mathcal{C}_1)$ by the edge $\{x, u\}$ with $u \in N_G(x) \setminus V(\mathcal{C}_1)$ (or delete this edge when $N_G(x) \setminus V(\mathcal{C}_1) = \emptyset$). This allows us to delete \mathcal{C}_1 from G .

► **Reduction Rule 2.** Delete all connected components in \mathcal{C}_1 and decrease k by the size of a minimum edge dominating set in \mathcal{C}_1 , i.e., let $G' = G - \mathcal{C}_1$, $X' = X$, and $k' = k - \text{EDS}(\mathcal{C}_1)$.

▷ **Claim 9.** Reduction Rule 2 is safe.

Proof. First, we will show that there exists an edge dominating set F of size at most k in G such that no edge in F has one endpoint in a connected component of \mathcal{C}_1 and the other endpoint in X . Let F be an edge dominating set of size at most k in G with $F \cap E(\mathcal{C}_1, X)$ minimal, and let $P = w_1w_2w_3w_4w_5$ be a path in \mathcal{C}_1 . We can assume, w.l.o.g., that F contains the edges $\{w_2, w_3\}$ and $\{w_3, w_4\}$ because F must dominate the non-adjacent edges $\{w_1, w_2\}$, $\{w_4, w_5\}$, and the vertices w_1, w_2, w_4, w_5 are only adjacent to vertices in P ; otherwise, P is contained in \mathcal{C}_2 and not \mathcal{C}_1 . Now, assume for contradiction that there exists an edge $e = \{x, y\} \in F \cap E(\mathcal{C}_1, X)$ with $x \in X$ and $y \in P$ where $P = w_1w_2w_3w_4w_5$ is a path in \mathcal{C}_1 . It holds that $y = w_3$ because w_3 is the only vertex in P that is adjacent to a vertex in X . If every vertex $u \in N_G(x)$ is contained in $V(F)$ then let $\tilde{F} = F \setminus \{e\}$. Otherwise, let $\tilde{F} = F \setminus \{e\} \cup \{\{x, u\}\}$, where $u \in N_G(x) \setminus V(F)$. It holds that \tilde{F} is an edge dominating set in G because $y = w_3$ is still a vertex in $V(\tilde{F})$ which implies $V(F) \subseteq V(\tilde{F})$. Furthermore, u is not contained in a connected component of \mathcal{C}_1 because for every path $P = w_1w_2w_3w_4w_5$ in \mathcal{C}_1 the vertex w_3 is contained in $V(F)$ and no other vertex is adjacent to a vertex in X . Now, the set \tilde{F} is an edge dominating set of size at most k in G with $\tilde{F} \cap E(\mathcal{C}_1, X) \subsetneq F \cap E(\mathcal{C}_1, X)$ which contradicts the minimality of $F \cap E(\mathcal{C}_1, X)$ and proves that there exists an edge dominating set F of size at most k in G with $F \cap E(\mathcal{C}_1, X) = \emptyset$. This implies that $F' = F \setminus E(\mathcal{C}_1)$ is an edge dominating set of size at most k' in G' when F is a solution to (G, k, X) with $F \cap E(\mathcal{C}_1, X) = \emptyset$.

For the other direction, let F' be an edge dominating set of size at most k' in G' . To obtain an edge dominating set F of size at most k in G we add for every path $P = w_1w_2w_3w_4w_5$ in \mathcal{C}_1 the two edges $\{w_2, w_3\}$ and $\{w_3, w_4\}$, which are a minimum edge dominating set of P , to F' . It follows that F has size $|F'| + \text{EDS}(\mathcal{C}_1) \leq k$. The set F dominates all edges in $G - X$ as well as all edges between \mathcal{C}_2 and X because $F' \subseteq F$, and because F contains an edge dominating set of \mathcal{C}_1 . Additionally, F dominates all edges between \mathcal{C}_1 and X because F dominates all middle vertices of the paths in \mathcal{C}_1 which are the only vertices in \mathcal{C}_1 that are adjacent to X . Hence, F is an edge dominating set of size at most k in G . ◀

Let (G', k', X') be the reduced instance. It holds that the set of connected components in $G' - X'$ is \mathcal{C}_2 because we delete all other connected components during Reduction Rule 2. Since $|\mathcal{C}_2| \leq |X_2| = |X'|$ it follows that G' has at most $5 \cdot |\mathcal{C}_2| + |X'| \leq 6|X'|$ vertices. It remains to show that we can perform the reduction in polynomial time. We apply each Reduction Rule at most once. Furthermore, we can apply the Reduction Rules in polynomial time because we can compute the partition of X as well as the sets \mathcal{C}_1 and \mathcal{C}_2 in polynomial time, and because we can delete sets of vertices from G and X in polynomial time. ◀

While this is not the full story about the classification in the following section, it hopefully shows the spirit of how upper and lower bounds for kernelization can arise. Solution edges between components of $G - X$ and X play a crucial role and they affect the solutions for components in nontrivial ways, e.g., apart from control opportunities, it depends on how much budget is needed for $H - B$ when edges between B and X are in the solution.

4 EDS parameterized by the size of a modulator to an \mathcal{H} -component graph

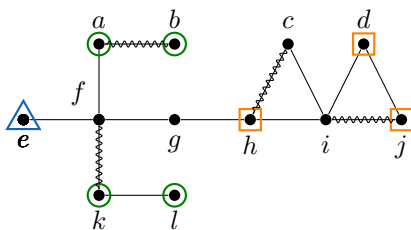
In this section, we develop a complete classification of EDGE DOMINATING SET parameterized by the size of a modulator to an \mathcal{H} -component graph regarding existence of polynomial kernelizations for all finite sets \mathcal{H} . This is motivated by the observed difference between modulating to P_3 -component graphs (no polynomial kernelization unless $\text{NP} \subseteq \text{coNP/poly}$) vs. modulating to P_5 -component graphs (polynomial kernelization). To this end, we will study which properties graphs $H \in \mathcal{H}$ must have, such that EDGE DOMINATING SET parameterized by the size of a modulator to an \mathcal{H} -component graph has resp. does not have a polynomial kernel. To recall, the input of our problem is a tuple (G, k, X) where $G - X$ is an \mathcal{H} -component graph and we ask whether G has an edge dominating set of size at most k ; the parameter is $|X|$.

In contrast to VERTEX COVER, where we can delete a vertex in the modulator if we know that this vertex must be in a solution of certain size, this is not the case for EDGE DOMINATING SET because we do not necessarily know which incident edge should be chosen. Of course, we can check for a vertex x in the modulator X how not having this vertex as an endpoint of a solution edge influences the size of a minimum edge dominating set of $G - X$. But, even if we find out that a vertex x in the modulator X must be an endpoint of a solution edge, we do not know if the other endpoint of the solution edge incident with x is in X or in a connected component of $G - X$. If there would be a connected component C in $G - X$ with the property that there exists a vertex $v \in N(x) \cap V(C)$ with $\text{EDS}(C) = \text{EDS}(C - v) + 1$, then it could be possible to have x as an endpoint of a solution edge without paying more than the cost of a minimum edge dominating set in C . Thus, instead of finding vertices in the modulator that must be endpoint of a solution edge, we want to find vertices in the modulator that can be endpoints of a solution edge without spending more budget than the size of a minimum edge dominating set in $G - X$. Similarly, getting edges to r vertices in X while increasing the cost in C by less than r is of interest (cost equal to r can always be had). The following definition classifies relevant vertices and vertex sets in a graph H , which may occur as a component of $G - X$.

► **Definition 10.** Let $H = (V, E)$ be a connected graph.

- We call a vertex $v \in V$ extendable if $\text{EDS}(H - v) + 1 = \text{EDS}(H)$. We denote the set of extendable vertices of H by $Q(H)$. (Intuitively, these vertices allow a local solution for an H -component in $G - X$ that includes an edge $\{v, x\}$ with $x \in X$ and $v \in V(H)$.)
- We call a set $Y \subseteq Q(H)$ free if for all vertices $v \in Y$ and for all minimum edge dominating sets F in H there exists a minimum edge dominating set F' in $H - v$ of size $|F| - 1$ and with $V(F) \setminus Y \subseteq V(F')$. By $W(H)$ we denote the unique maximum free set of H . We call a vertex $w \in W(H)$ free.⁴ (Intuitively, vertices in Y can be used for solution edges between components and X , while covering the same vertices of $H - Y$ as any local optimum solution; thus, they cannot be used for lower bounds like for P_3 -components.)
- We call a vertex $v \in V$ uncovered if no minimum edge dominating set F of H contains an edge incident with v , i.e. $v \notin V(F)$. We denote the set of uncovered vertices by $U(H)$. (Intuitively, H -components with any $v \in U(H)$ adjacent to $x \in X$ are easy to handle because $x \notin V(F)$ would imply that the local cost for H increases above $\text{EDS}(H)$.)

⁴ We show in the full version that $W(H)$ is unique.



■ **Figure 2** Example of an H -component with $\text{EDS}(H) = 4$. The wavy edges are a possible minimum edge dominating set of H .

- For any $Y \subseteq V$ define $\text{cost}(Y) := |Y| + \text{EDS}(H - Y) - \text{EDS}(H)$.
(Intuitively, $\text{cost}(Y)$ is equal to the additional budget that is needed for an H -component of $G - X$ when exactly the vertices in Y have solution edges to X . Note that $\text{cost}(\{v\}) = 0$ for all extendable vertices v .)
- We call a set $B \subseteq V \setminus W(H)$ beneficial if for all $\tilde{B} \subsetneq B$ we have $|B| - \text{cost}(B) > |\tilde{B}| - \text{cost}(\tilde{B})$ or, equivalently, $\text{EDS}(H - B) < \text{EDS}(H - \tilde{B})$. Note that this must also hold for $\tilde{B} = \emptyset$ which implies that for all beneficial sets we have $|B| - \text{cost}(B) > 0$ or, equivalently, $\text{EDS}(H - B) < \text{EDS}(H)$.
(Intuitively, the solution may include $|B|$ edges between B and some $X' \subseteq X$ while increasing the cost for the H -component by exactly $\text{cost}(B)$; this saves $|B| - \text{cost}(B) > 0$ over taking any $|B|$ edges incident with X' . The condition for all $\tilde{B} \subsetneq B$ ensures that the savings of getting $|B|$ edges at cost $\text{cost}(B)$ is greater than for any proper subset.)
- We call a beneficial set B strongly beneficial if $\text{cost}(B) < \sum_{i=1}^h \text{cost}(B_i)$ holds for all covers $B_1, B_2, \dots, B_h \subsetneq B$ of B . (Intuitively, for a strongly beneficial set B we cannot get the same number of edges to X by using sets B_i in several different H -components.)

► **Example 11** (Illustration of Definition 10). Figure 2 shows a connected graph H . The size of an edge dominating set in H is at least four because a solution has to dominate the four pairwise non-adjacent edges $\{a, b\}$, $\{k, l\}$, $\{j, d\}$ and $\{g, h\}$. Thus, $\text{EDS}(H) = 4$ because the wavy edges are an edge dominating set of H .

The vertices $\{a, b, k, l\}$, marked with a green cycle, as well as the vertices $\{d, h, j\}$, marked with an orange rectangle, are extendable. But only the green marked vertices $\{a, b, k, l\}$ are free: Let F be any minimum edge dominating set in H . The set F must contain exactly one of the two edges $e_1 = \{a, b\}$ and $e_2 = \{a, f\}$, and exactly one of the two edges $e_3 = \{k, l\}$ and $e_4 = \{k, f\}$. Now, $F' = F \setminus \{e_1, e_2, e_3, e_4\} \cup \{f, k\}$ is an edge dominating set in $H - a$ and $H - b$ of size $|F| - 1$, and $F' = F \setminus \{e_1, e_2, e_3, e_4\} \cup \{a, f\}$ is an edge dominating set in $H - k$ and $H - l$ of size $|F| - 1$ which implies that the vertices $\{a, b, k, l\}$ are free. The vertices $\{d, h, j\}$ are not free because no minimum edge dominating set F' in $H - d$, resp. $H - h$, resp. $H - j$ has vertex c , which is not extendable, as an endpoint of a solution edge, but the graph H has a minimum edge dominating set that has c as an endpoint, namely the one containing the wavy edges $\{a, b\}$, $\{h, c\}$, $\{d, j\}$. The vertex e , marked with a blue triangle, is uncovered.

The set $\{c, g\}$ is strongly beneficial, whereas the set $\{c, g, i, j\}$ is only beneficial, but not strongly beneficial: The set $\{c, g\}$ is beneficial because $\text{EDS}(H - \{c, g\}) = 3$ and $\text{EDS}(H - c) = \text{EDS}(H - g) = \text{EDS}(H) = 4$, and strongly beneficial because the only possible non-trivial cover of $\{c, g\}$ is $\{c\}, \{g\}$ and $\text{cost}(\{c, g\}) = 1 < 2 = \text{cost}(\{c\}) + \text{cost}(\{g\})$. The set $\{c, g, i, j\}$ is beneficial because $\text{EDS}(H - \{c, g, i, j\}) = 2$ and $\text{EDS}(H - B) \geq 3$ for all $B \subsetneq \{c, g, i, j\}$. But $\{c, g, i, j\}$ is not strongly beneficial because $\text{cost}(\{c, g, i, j\}) = 2 = 1 + 1 + 0 = \text{cost}(\{c, g\}) + \text{cost}(\{i\}) + \text{cost}(\{j\})$. Observe that the set $\{c, g, i\}$ is not beneficial even though

36:14 Edge Dominating Set

$\text{EDS}(H - \{c, g, i\}) = 3 < 4 = \text{EDS}(H)$, because $\{c, g\} \subsetneq \{c, g, i\}$ and $\text{EDS}(H - \{c, g, i\}) = 3 = \text{EDS}(H - \{c, g\})$.

We are now able to give a more detailed version of Theorem 1, which specifies for each finite set \mathcal{H} of connected graphs the kernelization complexity of EDGE DOMINATING SET parameterized by the size of a modulator to \mathcal{H} -component graphs.

► **Theorem 12.** *Let \mathcal{H} be any finite set of connected graphs. The EDGE DOMINATING SET problem parameterized by the size of a modulator to \mathcal{H} -component graphs behaves as follows:*

1. *If \mathcal{H} contains any graph H fulfilling one of the following items then there is no polynomial kernelization unless $\text{NP} \subseteq \text{coNP/poly}$:*
 - a. *There is an extendable vertex in H that is not free, i.e., $Q(H) \setminus W(H) \neq \emptyset$.*
 - b. *There is a strongly beneficial set B in H that contains an uncovered vertex, i.e., $B \cap U(H) \neq \emptyset$.*
 - c. *There is a vertex in H that is neither uncovered, free, nor neighbor of a free vertex, i.e., $V(H) \setminus (N[W(H)] \cup U(H)) \neq \emptyset$.*
 - d. *There is a strongly beneficial set $B \subseteq N(W(H))$ in H such that no minimum edge dominating set F_B of $H - B$ covers all vertices of $N(W(H)) \setminus B$.*
2. *Else, if \mathcal{H} contains at least one graph that has a strongly beneficial set, then there is a kernelization to $\mathcal{O}(|X|^d)$ vertices, $\mathcal{O}(|X|^{d+1})$ edges, and size $\mathcal{O}(|X|^{d+1} \log |X|)$, and there is no kernelization to size $\mathcal{O}(|X|^{d-\varepsilon})$, for any $\varepsilon > 0$, unless $\text{NP} \subseteq \text{coNP/poly}$ where d is the size of the largest strongly beneficial set in any $H \in \mathcal{H}$.*
3. *Else, there is a kernelization to $\mathcal{O}(|X|^2)$ vertices, $\mathcal{O}(|X|^3)$ edges, and size $\mathcal{O}(|X|^3 \log |X|)$, and there is no kernelization to size $\mathcal{O}(|X|^{2-\varepsilon})$, for any $\varepsilon > 0$, unless $\text{NP} \subseteq \text{coNP/poly}$.*

Observe, Theorem 1 directly follows from Theorem 12 because disconnected graphs in \mathcal{H} do not affect the resulting class of \mathcal{H} -component graphs, i.e., given any finite set \mathcal{H} of graphs we can take the subset \mathcal{H}' of connected graphs in \mathcal{H} and apply Theorem 12 to \mathcal{H}' . As an example for applying the theorem, for $\mathcal{H} = \{P_3\}$ we get Item 1a, for $\mathcal{H} = \{P_4\}$ we get Item 1b, for $\mathcal{H} = \{K_3\}$ and $\mathcal{H} = \{K_5\}$ we get Item 1c, and for $\mathcal{H} = \{P_2\} = \{K_2\}$, $\mathcal{H} = \{K_4\}$, $\mathcal{H} = \{P_5\}$, as well as $\mathcal{H} = \{E = \Xi\}$ we get Item 3.

► **Remark.** We showed that EDGE DOMINATING SET parameterized by the size of a given modulator X to a P_5 -component graph admits a kernelization with $\mathcal{O}(|X|)$ vertices (see Theorem 7). The reason why the kernelization procedure of Item 3 only reduces to $\mathcal{O}(|X|^2)$ vertices instead of $\mathcal{O}(|X|)$ vertices is that H -components can have uncovered vertices. This leads to a different marking argument similar to the case for EDGE DOMINATING SET parameterized by solution size. Note that EDS parameterized by solution size is covered by Item 3.

Proof outline for Theorem 12. We begin by establishing a number of useful properties of the terms introduced in Definition 10, e.g., that each graph H containing a beneficial set B also contains a strongly beneficial set $B' \subseteq B$ (Proposition 13 (11)).

The kernelization lower bound of Item 1 is proved by generalizing the lower bound obtained for P_3 -component graphs in Theorem 5. We define so-called *control pairs* by abstracting properties of P_3 -components used in the proof and show that there is no polynomial kernelization when any graph $H \in \mathcal{H}$ has a control pair (Theorem 15). For a connected graph $H = (V, E)$ the pair (C, B) is called a *control pair*, if $B \subseteq V$ is strongly beneficial, if $C \subseteq V \setminus (Q(H) \cup B)$ and no vertex $c \in C$ is extendable in $H - B$, if there exists a minimum edge dominating set F in H such that $C \subseteq V(F)$, and if for all minimum edge dominating

sets F_B in $H - B$ it holds that $C \not\subseteq V(F_B)$. Observe that for $H = P_3 = v_1v_2v_3$ the set B is the vertex v_2 and the set C is the vertex v_1 (or v_2). Afterwards, we show that graphs H fulfilling Items 1a, 1b, 1c, or 1d have control pairs (Lemmas 17, 18, 19, and 20).

In Item 1d, and in the items below, we (may) use that no graph in \mathcal{H} fulfills Items 1a, 1b, or 1c. Accordingly, each graph $H \in \mathcal{H}$ has $V(H) = N[W(H)] \cup U(H)$, i.e., each vertex of H is uncovered, free, or neighbor of a free vertex. Moreover, every extendable vertex is also free, i.e., $Q(H) = W(H)$, and strongly beneficial sets contain no (uncovered) vertices of $U(H)$. This implies that all strongly beneficial sets are subsets of $N(W(H))$, the neighborhood of the free vertices, as neither uncovered nor free vertices can be contained and no further vertices except those in $N(W(H))$ exist in H (in this case).

For Item 2 we have that no graph in \mathcal{H} fulfills any of the Items 1a through 1d and that at least one graph in \mathcal{H} has a strongly beneficial set. Thus, in addition to the above restrictions on $H \in \mathcal{H}$, we know that for each strongly beneficial set B , which here must be a subset of $N(W(H))$, there is a minimum edge dominating set F_B of $H - B$ that covers all vertices in $N(W(H)) \setminus B$. We give a general kernelization procedure that reduces the number of components in $G - X$ to $\mathcal{O}(|X|^d)$ where d is the size of the largest strongly beneficial set among graphs $H \in \mathcal{H}$ (Lemma 27). We then rule out kernelizations of size $\mathcal{O}(|X|^{d-\varepsilon})$ using only H -components, where H is any graph in \mathcal{H} that exhibits the largest size d of strongly beneficial sets (Lemma 31). Note that in the present item d is always at least two because having a strongly beneficial set B of size one would mean that $v \in B$ is an extendable vertex that is not free (because beneficial sets are disjoint from the set $W(H)$ of free vertices), which is handled by Item 1a.

Finally, for Item 3, it remains to consider the case that no graph $H \in \mathcal{H}$ fulfills any of the Items 1a through 1d and that no graph in H has a strongly beneficial set. It follows that no graph in H has any beneficial sets (Proposition 13 (11)) and, as before, we have $V(H) = N[W(H)] \cup U(H)$. We obtain a kernelization to $\mathcal{O}(|X|^2)$ vertices, $\mathcal{O}(|X|^3)$ edges, and size $\mathcal{O}(|X|^3 \log |X|)$ (Lemma 23). The lower bound ruling out kernelizations of size $\mathcal{O}(|X|^{2-\varepsilon})$ for any $\varepsilon > 0$, and in fact for any set \mathcal{H} , follows easily by a simple reduction from VERTEX COVER for which a lower bound ruling out size $\mathcal{O}(n^{2-\varepsilon})$ is known [6] (Lemma 35). ◀

5 Conclusion

As our main result, we have given a complete classification for EDGE DOMINATING SET parameterized by the size of a modulator to \mathcal{H} -component graphs for all finite sets \mathcal{H} . An obvious follow-up question is to extend this result to infinite sets \mathcal{H} . Our lower bounds of course continue to work in this setting, and the upper bounds still permit us to reduce the number of connected components (under the same conditions as before, e.g., that relevant beneficial sets have bounded size). However, for infinite \mathcal{H} , polynomial kernels also require us to shrink connected components of $G - X$, and to derive general rules for this. Moreover, even determining beneficial sets etc. for graphs $H \in \mathcal{H}$ could no longer be dismissed as being constant time. It is conceivable that such a classification is doable whenever graphs in \mathcal{H} have bounded treewidth, as this simplifies the required additional steps. Since most known tractable graph classes for EDGE DOMINATING SET have bounded treewidth (and tractability for $G - X$ is required, or else NP-hardness for $|X| = 0$ rules out kernels and fixed-parameter tractability), this seems like a reasonable goal. Apart from this, it would be nice to close the gap between size $\mathcal{O}(|X|^{d+1} \log |X|)$ and the lower bound of $\mathcal{O}(|X|^{d-\varepsilon})$.

References

- 1 Hans L. Bodlaender, Rodney G. Downey, Michael R. Fellows, and Danny Hermelin. On problems without polynomial kernels. *J. Comput. Syst. Sci.*, 75(8):423–434, 2009. doi:10.1016/j.jcss.2009.04.001.
- 2 Hans L. Bodlaender, Bart M. P. Jansen, and Stefan Kratsch. Kernelization Lower Bounds by Cross-Composition. *SIAM J. Discrete Math.*, 28(1):277–305, 2014. doi:10.1137/120880240.
- 3 Jean Cardinal, Stefan Langerman, and Eythan Levy. Improved approximation bounds for edge dominating set in dense graphs. *Theor. Comput. Sci.*, 410(8-10):949–957, 2009. doi:10.1016/j.tcs.2008.12.036.
- 4 Miroslav Chlebík and Janka Chlebíková. Approximation hardness of edge dominating set problems. *J. Comb. Optim.*, 11(3):279–290, 2006. doi:10.1007/s10878-006-7908-0.
- 5 Marek Cygan, Marcin Pilipczuk, Michal Pilipczuk, and Jakub Onufry Wojtaszczyk. On multiway cut parameterized above lower bounds. *TOCT*, 5(1):3:1–3:11, 2013. doi:10.1145/2462896.2462899.
- 6 Holger Dell and Dieter van Melkebeek. Satisfiability Allows No Nontrivial Sparsification unless the Polynomial-Time Hierarchy Collapses. *J. ACM*, 61(4):23:1–23:27, 2014. doi:10.1145/2629620.
- 7 Reinhard Diestel. *Graph Theory, 4th Edition*, volume 173 of *Graduate texts in mathematics*. Springer, 2012.
- 8 Bruno Escoffier, Jérôme Monnot, Vangelis Th. Paschos, and Mingyu Xiao. New Results on Polynomial Inapproximability and Fixed Parameter Approximability of Edge Dominating Set. *Theory Comput. Syst.*, 56(2):330–346, 2015. doi:10.1007/s00224-014-9549-5.
- 9 Henning Fernau. Edge dominating set: Efficient Enumeration-Based Exact Algorithms. In Hans L. Bodlaender and Michael A. Langston, editors, *Parameterized and Exact Computation, Second International Workshop, IWPEC 2006, Zürich, Switzerland, September 13-15, 2006, Proceedings*, volume 4169 of *Lecture Notes in Computer Science*, pages 142–153. Springer, 2006. doi:10.1007/11847250_13.
- 10 Fedor V. Fomin, Serge Gaspers, Saket Saurabh, and Alexey A. Stepanov. On Two Techniques of Combining Branching and Treewidth. *Algorithmica*, 54(2):181–207, 2009. doi:10.1007/s00453-007-9133-3.
- 11 Lance Fortnow and Rahul Santhanam. Infeasibility of instance compression and succinct PCPs for NP. *J. Comput. Syst. Sci.*, 77(1):91–106, 2011. doi:10.1016/j.jcss.2010.06.007.
- 12 Toshihiro Fujito and Hiroshi Nagamochi. A 2-approximation algorithm for the minimum weight edge dominating set problem. *Discrete Applied Mathematics*, 118(3):199–207, 2002. doi:10.1016/S0166-218X(00)00383-8.
- 13 Shivam Garg and Geevarghese Philip. Raising The Bar For Vertex Cover: Fixed-parameter Tractability Above a Higher Guarantee. In Robert Krauthgamer, editor, *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2016, Arlington, VA, USA, January 10-12, 2016*, pages 1152–1166. SIAM, 2016. doi:10.1137/1.9781611974331.ch80.
- 14 Petr A. Golovach, Pinar Heggernes, Dieter Kratsch, and Yngve Villanger. An Incremental Polynomial Time Algorithm to Enumerate All Minimal Edge Dominating Sets. *Algorithmica*, 72(3):836–859, 2015. doi:10.1007/s00453-014-9875-7.
- 15 Torben Hagerup. Kernels for Edge Dominating Set: Simpler or Smaller. In Branislav Rován, Vladimiro Sassone, and Peter Widmayer, editors, *Mathematical Foundations of Computer Science 2012 - 37th International Symposium, MFCS 2012, Bratislava, Slovakia, August 27-31, 2012. Proceedings*, volume 7464 of *Lecture Notes in Computer Science*, pages 491–502. Springer, 2012. doi:10.1007/978-3-642-32589-2_44.
- 16 Eva-Maria C. Hols and Stefan Kratsch. On Kernelization for Edge Dominating Set under Structural Parameters. *CoRR*, abs/1901.03582, 2019. arXiv:1901.03582.
- 17 John E. Hopcroft and Richard M. Karp. An $n^{5/2}$ Algorithm for Maximum Matchings in Bipartite Graphs. *SIAM J. Comput.*, 2(4):225–231, 1973. doi:10.1137/0202019.

- 18 Ken Iwaida and Hiroshi Nagamochi. An Improved Algorithm for Parameterized Edge Dominating Set Problem. *J. Graph Algorithms Appl.*, 20(1):23–58, 2016. doi:10.7155/jgaa.00383.
- 19 Bart M. P. Jansen and Hans L. Bodlaender. Vertex Cover Kernelization Revisited - Upper and Lower Bounds for a Refined Parameter. *Theory Comput. Syst.*, 53(2):263–299, 2013. doi:10.1007/s00224-012-9393-4.
- 20 Mamadou Moustapha Kanté, Vincent Limouzy, Arnaud Mary, and Lhouari Nourine. On the Neighbourhood Helly of Some Graph Classes and Applications to the Enumeration of Minimal Dominating Sets. In Kun-Mao Chao, Tsan-sheng Hsu, and Der-Tsai Lee, editors, *Algorithms and Computation - 23rd International Symposium, ISAAC 2012, Taipei, Taiwan, December 19-21, 2012. Proceedings*, volume 7676 of *Lecture Notes in Computer Science*, pages 289–298. Springer, 2012. doi:10.1007/978-3-642-35261-4_32.
- 21 Mamadou Moustapha Kanté, Vincent Limouzy, Arnaud Mary, Lhouari Nourine, and Takeaki Uno. Polynomial Delay Algorithm for Listing Minimal Edge Dominating Sets in Graphs. In Frank Dehne, Jörg-Rüdiger Sack, and Ulrike Stege, editors, *Algorithms and Data Structures - 14th International Symposium, WADS 2015, Victoria, BC, Canada, August 5-7, 2015. Proceedings*, volume 9214 of *Lecture Notes in Computer Science*, pages 446–457. Springer, 2015. doi:10.1007/978-3-319-21840-3_37.
- 22 Daniel Kobler and Udi Rotics. Edge dominating set and colorings on graphs with fixed clique-width. *Discrete Applied Mathematics*, 126(2-3):197–221, 2003. doi:10.1016/S0166-218X(02)00198-1.
- 23 Stefan Kratsch. A Randomized Polynomial Kernelization for Vertex Cover with a Smaller Parameter. In Piotr Sankowski and Christos D. Zaroliagis, editors, *24th Annual European Symposium on Algorithms, ESA 2016, August 22-24, 2016, Aarhus, Denmark, volume 57 of LIPIcs*, pages 59:1–59:17. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2016. doi:10.4230/LIPIcs.ESA.2016.59.
- 24 Stefan Kratsch and Magnus Wahlström. Representative Sets and Irrelevant Vertices: New Tools for Kernelization. In *53rd Annual IEEE Symposium on Foundations of Computer Science, FOCS 2012, New Brunswick, NJ, USA, October 20-23, 2012*, pages 450–459. IEEE Computer Society, 2012. doi:10.1109/FOCS.2012.46.
- 25 Daniel Lokshantov, N. S. Narayanaswamy, Venkatesh Raman, M. S. Ramanujan, and Saket Saurabh. Faster Parameterized Algorithms Using Linear Programming. *ACM Trans. Algorithms*, 11(2):15:1–15:31, 2014. doi:10.1145/2566616.
- 26 Elena Prieto. *Systematic kernelization in FPT algorithm design*. PhD thesis, The University of Newcastle, Australia, 2005.
- 27 Venkatesh Raman, M. S. Ramanujan, and Saket Saurabh. Paths, Flowers and Vertex Cover. In Camil Demetrescu and Magnús M. Halldórsson, editors, *Algorithms - ESA 2011 - 19th Annual European Symposium, Saarbrücken, Germany, September 5-9, 2011. Proceedings*, volume 6942 of *Lecture Notes in Computer Science*, pages 382–393. Springer, 2011. doi:10.1007/978-3-642-23719-5_33.
- 28 Venkatesh Raman, Saket Saurabh, and Somnath Sikdar. Efficient Exact Algorithms through Enumerating Maximal Independent Sets and Other Techniques. *Theory Comput. Syst.*, 41(3):563–587, 2007. doi:10.1007/s00224-007-1334-2.
- 29 Richard Schmieid and Claus Viehmann. Approximating edge dominating set in dense graphs. *Theor. Comput. Sci.*, 414(1):92–99, 2012. doi:10.1016/j.tcs.2011.10.001.
- 30 Johan M. M. van Rooij and Hans L. Bodlaender. Exact Algorithms for Edge Domination. *Algorithmica*, 64(4):535–563, 2012. doi:10.1007/s00453-011-9546-x.
- 31 Jianxin Wang, Beiwei Chen, Qilong Feng, and Jianer Chen. An Efficient Fixed-Parameter Enumeration Algorithm for Weighted Edge Dominating Set. In Xiaotie Deng, John E. Hopcroft, and Jinyun Xue, editors, *Frontiers in Algorithmics, Third International Workshop, FAW 2009, Hefei, China, June 20-23, 2009. Proceedings*, volume 5598 of *Lecture Notes in Computer Science*, pages 237–250. Springer, 2009. doi:10.1007/978-3-642-02270-8_25.

- 32 Mingyu Xiao. Exact and Parameterized Algorithms for Edge Dominating Set in 3-Degree Graphs. In Weili Wu and Ovidiu Daescu, editors, *Combinatorial Optimization and Applications - 4th International Conference, COCOA 2010, Kailua-Kona, HI, USA, December 18-20, 2010, Proceedings, Part II*, volume 6509 of *Lecture Notes in Computer Science*, pages 387–400. Springer, 2010. doi:10.1007/978-3-642-17461-2_31.
- 33 Mingyu Xiao, Ton Kloks, and Sheung-Hung Poon. New parameterized algorithms for the edge dominating set problem. *Theor. Comput. Sci.*, 511:147–158, 2013. doi:10.1016/j.tcs.2012.06.022.
- 34 Mingyu Xiao and Hiroshi Nagamochi. Parameterized edge dominating set in graphs with degree bounded by 3. *Theor. Comput. Sci.*, 508:2–15, 2013. doi:10.1016/j.tcs.2012.08.015.
- 35 Mingyu Xiao and Hiroshi Nagamochi. A refined exact algorithm for Edge Dominating Set. *Theor. Comput. Sci.*, 560:207–216, 2014. doi:10.1016/j.tcs.2014.07.019.
- 36 Mihalis Yannakakis and Fanica Gavril. Edge dominating sets in graphs. *SIAM Journal on Applied Mathematics*, 38(3):364–372, 1980.