(Meta) Kernelization

Hans L. Bodlaender

Fedor V. Fomin

Daniel Lokshtanov

Eelko Penninkx

Saket Saurabh

Dimitrios M. Thilikos

Technical Report UU-CS-2009-012
May 2009

Department of Information and Computing Sciences

Utrecht University, Utrecht, The Netherlands
www.cs.uu.nl

ISSN: 0924-3275

Department of Information and Computing Sciences
Utrecht University
P.O. Box 80.089
3508 TB Utrecht
The Netherlands

(Meta) Kernelization

Hans L. Bodlaender* Fedor V. Fomin† Daniel Lokshtanov†
Eelko Penninkx* Saket Saurabh† Dimitrios M. Thilikos‡

Abstract

Polynomial time preprocessing to reduce instance size is one of the most commonly deployed heuristics to tackle computationally hard problems. In a parameterized problem, every instance I comes with a positive integer k. The problem is said to admit a *polynomial kernel* if, in polynomial time, we can reduce the size of the instance I to a polynomial in k, while preserving the answer. In this paper, we show that all problems expressible in Counting Monadic Second Order Logic and satisfying a compactness property admit a polynomial kernel on graphs of bounded genus. Our second result is that all problems that have finite integer index and satisfy a weaker compactness condition admit a linear kernel on graphs of bounded genus. The study of kernels on planar graphs was initiated by a seminal paper of Alber, Fellows, and Niedermeier [*J. ACM, 2004*] who showed that Planar Dominating Set admits a linear kernel. Following this result, a multitude of problems have been shown to admit linear kernels on planar graphs by combining the ideas of Alber et al. with problem specific reduction rules. Our theorems unify and extend *all* previously known kernelization results for planar graph problems. Combining our theorems with the Erdős-Pósa property we obtain various new results on linear kernels for a number of packing and covering problems.

1 Introduction

Preprocessing (data reduction or kernelization) as a strategy of coping with hard problems is universally used in almost every implementation. The history of preprocessing, like applying reduction rules to simplify truth functions, can be traced back to the 1950's [49]. A natural question in this regard is how to measure the quality of preprocessing rules proposed for a specific problem. For a long time the mathematical analysis of polynomial time preprocessing algorithms was neglected. The basic reason for this anomaly was that

*Department of Information and Computing Sciences, Utrecht University, P.O. Box 80.089, 3508 TB Utrecht, The Netherlands {hansb|penninkx}@cs.uu.nl

†Department of Informatics, University of Bergen, N-5020 Bergen, Norway.
{fedor.fomin|daniello|saket.saurabh}@ii.uib.no

‡Department of Mathematics, National & Kapodistrian University of Athens, Panepistimioupolis, GR-15784, Athens, Greece. sedthilk@math.uoa.gr

if we start with an instance I of an NP-hard problem and can show that in polynomial time we can replace this with an equivalent instance I' with $|I'| < |I|$ then that would imply P=NP in classical complexity. The situation changed drastically with advent of parameterized complexity. Combining tools from parameterized complexity and classical complexity it has become possible to derive upper and lower bounds on the sizes of reduced instances, or so called *kernels*. In fact some times we can solve a computationally hard problem on some special inputs by just using data reduction. In this regard, we can trace back to the work of Arnborg et al. [5], where they show that each problem in monadic second order logic on graphs of bounded treewidth can be solved in linear time, by using only reductions, that is, no tree decomposition is built. There are similar ideas even in the earlier papers by Fellows and Langston, see for an example [30, 31, 32].

In parameterized complexity each problem instance comes with a parameter k and the parameterized problem is said to admit a *polynomial kernel* if there is a polynomial time algorithm (the degree of polynomial is independent of k), called a *kernelization* algorithm, that reduces the input instance down to an instance with size bounded by a polynomial $p(k)$ in k, while preserving the answer. This reduced instance is called a $p(k)$ *kernel* for the problem. If $p(k) = O(k)$, then we call it a *linear kernel*. Kernelization has been extensively studied in the realm of parameterized complexity, resulting in polynomial kernels for a variety of problems. Notable examples include a $2k$-sized vertex kernel for Vertex Cover [17], a $355k$ kernel for Dominating Set on planar graphs [3], which later was improved to a $67k$ kernel [15], and an $O(k^2)$ kernel for Feedback Vertex Set [52] parameterized by the solution size.

One of the most important results in the area of kernelization is given by Alber et al. [3]. They gave the first linear sized kernel for the Dominating Set problem on planar graphs. The work of Alber et al. [3] triggered an explosion of papers on kernelization, and in particular on kernelization of problems on planar graphs. Combining the ideas of Alber et al. [3] with problem specific data reduction rules, kernels of linear sizes were obtained for a variety of parameterized problems on planar graphs including Connected Vertex Cover, Minimum Edge Dominating Set, Maximum Triangle Packing, Efficient Edge Dominating Set, Induced Matching, Full-Degree Spanning Tree, Feedback Vertex Set, Cycle Packing, and Connected Dominating Set [3, 11, 12, 16, 38, 39, 42, 46, 47]. Dominating Set has received special attention from kernelization view point, leading to a linear kernel on graphs of bounded genus [34] and polynomial kernel on graphs excluding a fixed graph H as a minor and on d-degenerated graphs [4, 48]. We refer to the survey of Guo and Niedermeier [37] for a detailed treatment of the area of kernelization.

Most of the papers on linear kernels on planar graphs have the following idea in common: find an appropriate region decomposition of the input planar graph based on the problem in question, and then perform *problem specific* rules to reduce the part of the graph inside each region. The first step towards the general abstraction of all these algorithms was initiated by Guo and Niedermeier [38], who introduced the notion of "problems with distance property". However, to prove that some problem posses a linear kernel on planar graphs, the approach of Guo and Niedermeier still requires constructing reduction rules which are problem dependent. Thus all previous work on kernelization was strongly based on the design of problem specific reduction rules. In this paper we step aside and find properties of problems, *like expressibility in certain logic*, which allows these reduction rules to be automated.

Our results can be seen as what Grohe and Kreutzer call *algorithmic meta theorems* [36, 44]. Meta theorems expose the deep relations between logic and combinatorial structures, which is a fundamental issue of computational complexity. Also such theorems yield a better understanding of the scope of general algorithmic techniques and the limits of tractability. The very typical example is the celebrated Courcelle's theorem [20] which states that all graph properties definable in Monadic Second Order Logic (MSO) can be decided in linear time on graphs of bounded treewidth. More recent examples of such meta theorems state that all first-order definable properties of planar graphs can be decided in linear time [35] and that all first-order definable optimization problems on classes of graphs with excluded minors can be approximated in polynomial time to any given approximation ratio [23]. Our meta theorems not only give a uniform and natural explanation for a large family of known kernelization results but also provide a variety of new results. In what follows we build up towards our theorems. We first give necessary definitions needed to formulate our results.

Let $\mathcal{G}_g$ be the family of all graphs that can be embedded into a surface Σ of Euler-genus at most g. Given a graph G embedded on a surface Σ of Euler-genus g, and a set S, we define $\mathbf{R}_G^r(S)$ to be the set of all vertices of G whose radial distance from some vertex of S is at most r. The radial distance between two vertices x, y is the minimum length of an alternating sequence of vertices and faces starting from x and ending in y, such that every two consecutive elements of this sequence are incident to each other. We say that a parameterized problem $\Pi \subseteq \mathcal{G}_g \times \mathbb{N}$ is *compact* if there exist an integer r such that for all $(G = (V, E), k) \in \Pi$, there is an embedding of G into a surface Σ of Euler-genus at most g and a set $S \subseteq V$ such that $|S| \leq r \cdot k$, $\mathbf{R}_G^r(S) = V$ and $k \leq |V|^r$. Similarly, Π is *quasi-compact* if there exists an integer r such that for every $(G, k) \in \Pi$, there is an embedding of G into a surface Σ of Euler-genus at most g and a set $S \subseteq V$ such that $|S| \leq r \cdot k$, $\mathbf{tw}(G \setminus \mathbf{R}_G^r(S)) \leq r$ and $k \leq |V|^r$ (by $\mathbf{tw}(G)$ we denote the treewidth of G). Notice that if a problem is compact then it is also quasi-compact.

We use Counting Monadic Second Order Logic (CMSO) [6, 19, 21], an extension of MSO, as a basic tool to express properties of vertex/edge sets in graphs. In fact, it is known that every set $\mathcal{F}$ of graphs of bounded treewidth is CMSO-definable if and only if $\mathcal{F}$ is finite state [45]. Our first result concerns a parameterized analogue of graph optimization problems where the objective is to find a maximum or minimum sized vertex or edge set satisfying a CMSO-expressible property. In particular, the problems considered are defined as follows. In a p-MIN-CMSO graph problem $\Pi \subseteq \mathcal{G}_g \times \mathbb{N}$, we are given a graph $G = (V, E)$ and an integer k as input. The objective is to decide whether there is a vertex/edge set S of size at most k such that the CMSO-expressible predicate $P_\Pi(G, S)$ is satisfied. In a p-EQ-CMSO problem the size of S is required to be exactly k and in a

p-MAX-CMSO problem the size of S is required to be at least k. The *annotated* version Π^α of a p-MIN/EQ/MAX-CMSO problem Π is defined as follows. The input is a triple $(G = (V, E), Y, k)$ where G is a graph, $Y \subseteq V$ is a set of black vertices, and k is a non-negative integer. In the *annotated version* of a p-MIN/EQ-CMSO graph problem, S is additionally required to be a subset of Y. For the annotated version of a p-MAX-CMSO graph problem S is not required to be a subset of Y, but instead of $|S| \geq k$ we demand that $|S \cap Y| \geq k$.

Our results. For a parameterized problem $\Pi \subseteq \mathcal{G}_g \times \mathbb{N}$, let $\overline{\Pi} \subseteq \mathcal{G}_g \times \mathbb{N}$ denote the set of all no instances of Π. Our first result is the following theorem.

Theorem 1. *Let $\Pi \subseteq \mathcal{G}_g \times \mathbb{N}$ be a p-MIN/EQ/MAX-CMSO problem and either Π or $\overline{\Pi}$ is compact. Then the annotated version Π^α admits a cubic kernel if Π is a p-EQ-CMSO problem and a quadratic kernel if Π is a p-MIN/MAX-CMSO problem.*

We remark that a polynomial kernel for an annotated graph problem Π^α, is a polynomial time algorithm that given an input $(G = (V, E), Y, k)$ of Π^α, computes an equivalent instance $(G' = (V', E'), Y', k')$ of Π^α such that $|V'|$ and $k' \leq k^{O(1)}$. Theorem 1 has the following corollary.

Corollary 1. *Let $\Pi \subseteq \mathcal{G}_g \times \mathbb{N}$ be an NP-complete p-MIN/EQ/MAX-CMSO problem such that either Π or $\overline{\Pi}$ is compact and Π^α is in NP. Then Π admits a polynomial kernel.*

Theorem 1 and its corollary give polynomial kernels for *all* graph problems on surfaces for which such kernels are known. However, many problems in the literature are known to admit linear kernels on planar graphs. Our next theorem unifies and generalizes *all* known linear kernels for graph problems on surfaces. To this end we utilize the notion of *finite integer index*. This term first appeared in the work by Bodlaender and van Antwerpen-de Fluiter [13, 24] and is similar to the notion of *finite state* [1, 14, 19] (see Subsection 2.6 for the formal definitions).

Theorem 2. *Let $\Pi \subseteq \mathcal{G}_g \times \mathbb{N}$ has finite integer index and either Π or $\overline{\Pi}$ is quasi-compact. Then Π admits a linear kernel.*

Our two theorems are similar in spirit, yet they have a few differences. In particular, not every p-MIN/EQ/MAX-CMSO graph problem has finite integer index. For an example the INDEPENDENT DOMINATING SET problem is a p-MIN-CMSO problem, but it does not have finite integer index. Also, the class of problems that have finite integer index does not have a syntactic characterization and hence it takes some more work to apply Theorem 2 than Theorem 1. On the other hand, Theorem 2 yields linear kernels, applies to quasi-compact problems and unifies and generalizes results presented in [2, 3, 11, 12, 16, 34, 38, 39, 42, 46, 47] as a corollary.

To demonstrate the applicability of our results we show in Section 6 how our theorems lead to polynomial or linear kernels for a variety of problems. For ease of reference we provide in the Appendix A, a compendium of parameterized problems for which polynomial or linear kernels are derived.

2 Preliminaries

In this section we give various definitions which we make use of in the paper. Let $G = (V, E)$ be a graph. A graph $G' = (V', E')$ is a *subgraph* of G if $V' \subseteq V$ and $E' \subseteq E$. The subgraph G' is called an *induced subgraph* of G if $E' = \{\{u, v\} \in E \mid u, v \in V'\}$, in this case, G' is also called the subgraph *induced by* V' and denoted with $G[V']$.

2.1 Parameterized algorithms and Kernels

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

Definition 1. [**Kernelization**] *A kernelization algorithm, or in short, a kernel for a parameterized problem* $\Pi \subseteq \Gamma^* \times \mathbb{N}$ *is an algorithm that given* $(x, k) \in \Gamma^* \times \mathbb{N}$ *outputs in time polynomial in* $|x| + k$ *a pair* $(x', k') \in \Gamma^* \times \mathbb{N}$ *such that (a)* $(x, k) \in \Pi$ *if and only if* $(x', k') \in \Pi$ *and (b)* $|x'|, k' \leq g(k)$*, where* g *is some computable function. The function* g *is referred to as the size of the kernel. If* $g(k) = k^{O(1)}$ *or* $g(k) = O(k)$ *then we say that* Π *admits a polynomial kernel and linear kernel respectively.*

2.2 Surfaces and Distances

In this paper we consider graphs embeddable into surfaces. Let $\mathcal{G}_g$ be the class of all graphs that can be embedded into a surface Σ of Euler-genus at most g. We say that a graph G is Σ-embedded if it is accompanied with an embedding of the graph into Σ. We define the *normal distance* between two vertices x and y to be the shortest path distance between them. The *radial distance* between x and y is defined to be one less than the minimum length of a sequence starting from x and ending at y such that vertices and faces alternate in the sequence. Given an Σ-embedded graph $G = (V, E)$ and a set $S \subseteq V$, we denote by $\mathbf{R}^r_G(S)$ and $\mathbf{B}^r_G(S)$ the set of all vertices that are in radial distance at most r and normal distance at most r from some vertex in S respectively. Notice that for every set $S \subseteq V$ and every $r \geq 0$, it holds that $\mathbf{B}^r_G(S) \subseteq \mathbf{R}^{2r+1}_G(S)$ for any embedding of G into a surface Σ. An alternative way of viewing radial distance is to consider the *radial graph*, R_G: an embedded multigraph whose vertices are the vertices and the faces of G (each face f of G is represented by a point v_f in it). An edge between a vertex v and a vertex v_f is drawn if and only if v is incident to f. Thus R_G is a bipartite multigraph, embedded in the same surface as G. Hence, if $G \in \mathcal{G}_g$ then $R_G \in \mathcal{G}_g$. Also the radial distance of a pair of vertices in G corresponds to the normal distance in R_G.

Let $G = (V, E)$ be a graph. A *tree decomposition* of G is a pair $(T = (V_T, E_T)), \mathcal{X} = \{X_t\}_{t \in V_T})$ such that: $\cup_{u \in V_T} = V$, $\forall e = (u, v) \in E, \exists t \in V_T : \{u, v\} \subseteq X_t$ and $\forall v \in V$, $T[\{t \mid v \in X_t\}]$ is connected. The *width* of a tree decomposition is $\max_{t \in V_T} |X_t| - 1$

and the *treewidth* of $G = (V, E)$ is the minimum width over all tree decompositions of G. A tree decomposition is called a *nice tree decomposition* if the following conditions are satisfied: Every node of the tree T has at most two children; if a node t has two children t_1 and t_2, then $X_t = X_{t_1} = X_{t_2}$; and if a node t has one child t_1, then either $|X_t| = |X_{t_1}| + 1$ and $X_{t_1} \subset X_t$ or $|X_t| = |X_{t_1}| - 1$ and $X_t \subset X_{t_1}$. It is possible to transform a given tree decomposition into a nice tree decomposition in time $O(|V| + |E|)$ [8].

Given an edge $e = (u, v)$ of a graph $G = (V, E)$, by contracting an edge (u, v) we mean identifying the vertices u and v and removing all the loops and duplicate edges. A *minor* of a graph G is a graph H that can be obtained from a subgraph of G by contracting edges. A graph class $\mathcal{C}$ is *minor closed* if any minor of any graph in $\mathcal{C}$ is also an element of $\mathcal{C}$. A minor closed graph class $\mathcal{C}$ is *H-minor-free* or simply *H-free* if $H \notin \mathcal{C}$.

2.3 Compactness and Protrusions

Let $\mathcal{G}_g$ be the family of all graphs that can be embedded into a surface Σ of Euler-genus at most g. We say that a parameterized problem $\Pi \subseteq \mathcal{G}_g \times \mathbb{N}$ is *compact* if there exist an integer r such that for all $(G = (V, E), k) \in \Pi$, there is an embedding of G into a surface Σ of Euler-genus at most g and a set $S \subseteq V$ such that $|S| \leq r \cdot k$, $\mathbf{R}^r_G(S) = V$ and $k \leq |V|^r$. Similarly, Π is *quasi-compact* if there exists an integer r such that for every $(G, k) \in \Pi$, there is an embedding of G into a surface Σ of Euler-genus at most g and a set $S \subseteq V$ such that $|S| \leq r \cdot k$, $\mathbf{tw}(G \setminus \mathbf{R}^r_G(S)) \leq r$ and $k \leq |V|^r$.

Given a graph $G = (V, E)$ and $S \subseteq V$, we define $\partial_G(S)$ as the set of vertices in S that have a neighbor in $V \setminus S$. For a set $S \subseteq V$ the neighbourhood of S is $N_G(S) = \partial_G(V \setminus S)$. When it is clear from the context, we omit the subscripts. We now define the notion of a *protrusion*.

Definition 1. [r-**protrusion**] *Given a graph $G = (V, E)$, we say that a set $X' \subseteq V$ is an r-*protrusion *of G if $|N(X')| \leq r$ and $\mathbf{tw}(G[X' \cup N(X')]) \leq r$.*

For an r-protrusion X', the vertex set $X = X' \cup N(X')$ is an *extended r-protrusion*. The set X is the extended protrusion of X' and X' is the protrusion of X.

2.4 Counting Monadic Second Order Logic

The syntax of MSO of graphs includes the logical connectives $\vee$, $\wedge$, $\neg$, $\Leftrightarrow$, $\Rightarrow$, variables for vertices, edges, set of vertices and set of edges, the quantifiers $\forall$, $\exists$ that can be applied to these variables, and the following five binary relations: (1) $u \in U$ where u is a vertex variable and U is a vertex set variable; (2) $d \in D$ where d is an edge variable and D is an edge set variable; (3) $\mathbf{inc}(d, u)$, where d is an edge variable, u is a vertex variable, and the interpretation is that the edge d is incident on the vertex u; (4) $\mathbf{adj}(u, v)$, where u and v are vertex variables u, and the interpretation is that u and v are adjacent; (5) equality of variables representing vertices, edges, set of vertices and set of edges. In addition to the usual features of monadic second-order logic, if we have atomic formulas testing whether the cardinality of a set is equal to n modulo p, where n and p are integers such that

$0 \leq n < p$ and $p \geq 2$, then this extension of the MSO is called the *counting monadic second-order logic*. So essentially CMSO is MSO with following atomic formula: If U denotes a set X, then $\mathbf{card}_{n,p}(U) = \mathbf{true}$ if and only if $|X|$ is n **mod** p. We refer to [6, 19, 21] for a detailed introduction on CMSO.

2.5 t-Boundaried Graphs

Here we define the notion of *t-boundaried graphs* and various operations on them.

Definition 2. [**t-Boundaried Graphs**] *A t-boundaried graph is a graph $G = (V, E)$ with t distinguished vertices, uniquely labeled from 1 to t. The set $\partial(G)$ of labeled vertices is called the boundary of G. The vertices in $\partial(G)$ are referred to as boundary vertices or terminals.*

For a graph $G = (V, E)$ and a vertex set $S \subseteq V$, we will sometimes consider the graph $G[S]$ as the $|\partial(S)|$-boundaried graph with $\partial(S)$ being the boundary.

Definition 3. [**Gluing by $\oplus$**] *Let G_1 and G_2 be two t-boundaried graphs. We denote by $G_1 \oplus G_2$ the t-boundaried graph obtained by taking the disjoint union of G_1 and G_2 and identifying each vertex of $\partial(G_1)$ with the vertex of $\partial(G_2)$ with the same label; that is, we glue them together on the boundaries. In $G_1 \oplus G_2$ there is an edge between two labeled vertices if there is an edge between them in G_1 or in G_2.*

Definition 4. [**Legality**] *Let $\mathcal{G}$ be a graph class, G_1 and G_2 be two t-boundaried graphs, and $G_1, G_2 \in \mathcal{G}$. We say that $G_1 \oplus G_2$ is legal with respect to $\mathcal{G}$ if the unified graph $G_1 \oplus G_2 \in \mathcal{G}$. If the class $\mathcal{G}$ is clear from the context we do not say with respect to which graph class the operation is legal.*

Definition 5. [**Replacement**] *Let $G = (V, E)$ be a graph containing an extended r-protrusion X. Let X' be the restricted protrusion of X and let G_1 be an r-boundaried graph. The act of replacing X' with G_1 corresponds to changing G into $G[V \setminus X'] \oplus G_1$. Replacing $G[X]$ with G_1 corresponds to replacing X' with G_1.*

2.6 Finite Integer Index

Definition 6. *For a parameterized problem, Π on a graph class $\mathcal{G}$ and two t-boundaried graphs G_1 and G_2, we say that $G_1 \equiv_\Pi G_2$ if there exists a constant c such that for all t-boundaried graphs G_3 and for all k,*

- *$G_1 \oplus G_3$ is legal if and only if $G_2 \oplus G_3$ is legal.*
- *$(G_1 \oplus G_3, k) \in \Pi$ if and only if $(G_2 \oplus G_3, k + c) \in \Pi$.*

Definition 7. [**Finite Integer Index**] *Π has finite integer index in $\mathcal{G}$ if for every t there exists a finite set $\mathcal{S}$ of t-boundaried graphs such that $\mathcal{S} \subseteq \mathcal{G}$ and for any t-boundaried graph G_1 there exists a $G_2 \in \mathcal{S}$ such that $G_2 \equiv_\Pi G_1$. Such a set $\mathcal{S}$ is called a set of representatives for (Π, t).*

Note that for every t, the relation $\equiv_\Pi$ on t-boundaried graphs is an equivalence relation. A problem Π is finite integer index, if and only if for every t, $\equiv_\Pi$ is of finite index, that is, has a finite number of equivalence classes.

3 Reduction Rules

In this section we give reduction rules for compact annotated p-MIN/EQ/MAX-CMSO graph problems and quasi-compact parameterized problems having finite integer index. Our reduction rules have the following form:

> *If there is a constant size separator such that after its removal we obtain a connected component of unbounded size and of constant treewidth, then we replace this component with something of constant size.*

The implementation of this rule depends on whether we are dealing with an annotated p-MIN-CMSO, p-EQ-CMSO or p-MAX-CMSO problems, or whether the problem in question has finite integer index. Our reduction rules for annotated p-MIN/EQ/MAX-CMSO problems have three parts. In the first two parts we zero in on an area to reduce, in the last part we perform the reduction. In all cases, we assume that we are given a sufficiently large r-protrusion. In the following discussion we only treat annotated p-MIN/EQ/MAX-CMSO problems where the set S being searched for is a set of vertices. The case where S is a set of edges can be dealt in an identical manner.

3.1 Reduction for Annotated p-MIN-CMSO Problems

We now describe the reduction rule that we apply to annotated p-MIN-CMSO problems. The technique employed in this section will act as a template for how we handle the annotated p-EQ/MAX-CMSO problems. Recall that in an annotated p-MIN-CMSO problem Π^α we are given a graph $G = (V, E)$ where a subset Y of the vertices of G is colored *black* and an integer k. The objective is to find a set $S \subseteq Y$ of size at most k such that a fixed CMSO-definable property $P_\Pi(G, S)$ holds. For our reduction rule, we are also given a sufficiently large r-protrusion X'. In the first step of the reduction, we show that the set $Y \cap X'$ can be reduced to $O(k)$ vertices without changing whether (G, k) is a yes instance to Π^α or not. In the second step we show that the r-protrusion X' can be covered by $O(k)$ r'-protrusions such that each r'-protrusion contains at most a constant number of vertices from Y. In the third and final step of the reduction rule, we replace the largest r'-protrusion with an equivalent, but smaller r'-boundaried graph. We now provide the reduction rule for annotated p-MIN-CMSO problems.

Lemma 1. *Let Π^α be an annotated p-MIN-CMSO problem. Let $G = (V, E)$ be a graph, $Y \subseteq V$ be the set of black vertices and k be an integer. Let X be an extended r-protrusion of G. Then there is an integer c, and an $O(|X|)$ time algorithm, that computes a set of vertices $Z \subseteq X \cap Y$ with $|Z| \le ck$, such that if there exists a $W \subseteq Y$ with $P_\Pi(G, W)$ and $|W| \le k$, then there exists a $W' \subseteq Y$ with $P_\Pi(G, W')$, $|W'| \le k'$, and $W' \cap X \subseteq Z$.*

Proof. The algorithm starts by making a tree decomposition of $G[X]$ of width at most r, using the linear time algorithm to compute treewidth by Bodlaender [8]. Now we add $\partial(X)$ to each bag, and add one bag containing only the vertices in $\partial(X)$. The tree decomposition has width at most $2r$ as the bag size is at most $r+1+|\partial(X)| \leq 2r+1$.

Consider the following equivalence relation on subsets $Q \subseteq X \cap Y$. We say that $Q \sim Q'$, if and only if for all $R \subseteq V - X$:

$$P_\Pi(G, Q \cup R) \Leftrightarrow P_\Pi(G, Q' \cup R). \tag{1}$$

The number of equivalence classes is bounded by a function of the treewidth [14, 22] of $G[X]$, and thus for fixed r, can be assumed to be bounded by a constant, say c. We would like to find a minimum sized representative of each of the equivalence classes. We describe an algorithm running in time $O(|X|)$ to find the desired set in each equivalence class. Let us consider an algorithm that solves an optimization version of the problem

$$\min\{|W| \mid W \subseteq Y \wedge P_\Pi(G, W)\}$$

on graphs of bounded treewidth, using a dynamic programming approach. For an example, we can use the algorithm described by Borie et al. [14]. The algorithm of Borie et al. [14] computes for each equivalence class in Relation 1 (or, possibly, a refinement of the Relation 1) the minimum size of a set in the class. This is done in a dynamic programming fashion, computing each value given a table of these values for the children of the bag in the tree decomposition. The running time is linear for fixed c. It is not hard to observe that we can also compute for each equivalence class a minimum size set $Q \subseteq X \cap Y$ that belongs to the class. This can be done in linear time. If there are more than one minimum size sets in a class, then the algorithm just picks one.

Let $\mathcal{Q}$ denote that set of equivalence classes of Relation 1 and suppose that for each class $q \in \mathcal{Q}$ that is non-empty, we have a minimum size representative Q_q. By the above argument we can find a minimum size representative Q_q for each class in $O(|X|)$ time. Now, set

$$Z = \bigcup_{q \in \mathcal{Q}, |Q_q| \leq k} Q_q \ .$$

Suppose now that there exists $W \subseteq Y$ with $P_\Pi(G, W)$ and $|W| \leq k$. Consider the equivalence class q that contains $X \cap W$. Let Q_q be the selected minimum size representative of q. Consider the set $W' = (W \setminus X) \cup Q_q$. As $P_\Pi(G, (W \setminus X) \cup (X \cap W))$, we have that $P_\Pi(G, (W \setminus X) \cup Q_c) = P_\Pi(G, W')$. Since, Q_q is a minimum size representative from q, we have that $|Q_q| \leq |W \setminus X|$, and that $|W'| \leq |W|$. Finally, since the number of equivalence classes in $\mathcal{Q}$ is a function of r and each representative Q_q has size at most k, we have that $|Z| = O(k)$. This proves the lemma. □

Using Lemma 1, we change the set Y to $(Y \setminus X) \cup Z$. We now show how to exploit the fact that Z contains $O(k)$ vertices.

Partitioning Protrusions: In the second step of the reduction rule we partition the extended r-protrusion X into smaller r'-protrusions.

Lemma 2. *Let $G = (V, E)$ be a graph, $Y \subseteq V$ be the set of black vertices of G and k be an integer. Furthermore, let X be an extended r-protrusion and $Z = X \cap Y$. There is a $O(|X|)$ time algorithm that finds $O(|Z|)$ extended r'-protrusions $X_1, X_2, \ldots, X_\ell$ such that $X = X_1 \cup X_2 \cup \ldots \cup X_\ell$ and for every $i \leq \ell$, $Z \cap X_i \subseteq \partial(X_i)$.*

Proof. We start by making a nice tree decomposition of $G[X]$, and adding $\partial(X)$ to all bags. Now, we mark a number of bags. First, we mark the root bag and each forget node where a vertex in Z is forgotten. As each vertex is forgotten at most once in a nice tree decomposition, so far we have $O(|Z|)$ marked bags. Now, mark each bag that is the lowest common ancestor of two marked bags, until we cannot marks in this fashion. Standard counting arguments for trees show that this operation at most doubles the number of marks. Hence, there are at most $O(|Z|)$ marked bags.

We now split X into $X_1, X_2, \ldots, X_\ell$ as follows: we take parts of the tree decomposition, with internally no marked bags, plus the marked bags at the border. Note that each such part has at most two marked bags, each of size at most $2r$ and thus the size of $\partial(X_i)$ is at most $4r$ for every $i \leq \ell$. Note that by the way we constructed the sets X_i, $Z \cap X_i \subseteq \partial(X_i)$ for every i. □

Reducing Protrusions: In the third phase of our reduction rule, we find a protrusion to replace, and perform the replacement. Notice that this part of the reduction works both for annotated p-MIN-CMSO and for annotated p-EQ-CMSO problems.

Lemma 3. *Let Π^α be an annotated p-MIN-CMSO or p-EQ-CMSO problem. There is a fixed constant c depending only on Π^α such that there is an algorithm that given a graph $G = (V, E) \in \mathcal{G}_g$, a set $Y \subseteq V$ of black vertices, an integer k and an extended r-protrusion X with $|X| > c$ such that $Y \cap X \subseteq \partial(X)$, runs in time $O(|X|)$, and produces a graph $G^* = (V^*, E^*) \in \mathcal{G}_g$ such that $|V^*| < |V|$ and $(G^*, k) \in \Pi^\alpha$ if and only if $(G, k) \in \Pi^\alpha$.*

Proof. For two t-boundaried graphs G_1 and G_2, we say that they are equivalent with respect to a subset S of $\partial(G_1) = \partial(G_2)$ if for every $G_3 = (V_3, E_3)$ and set $S' \subseteq V_3$ we have that $P_\Pi(G_1 \oplus G_3, S \cup S')$ if and only if $P_\Pi(G_2 \oplus G_3, S \cup S')$. If G_1 and G_2 are equivalent with respect to S we say that $G_1 \sim_S G_2$. The canonical equivalence relation for CMSO properties with free set variables has finite index [14, 22] and hence the number of equivalence classes of $\sim_S$ depends only on t for every fixed $S \subseteq \partial(G_1)$. We make a new equivalence relation defined on the set of t-boundaried graphs belonging to a graph class $\mathcal{G}$. Two t-boundaried graphs $G_1, G_2 \in \mathcal{G}_g$ are equivalent if

- for every t-boundaried graph G_3, $G_1 \oplus G_3$ is legal if and only if $G_2 \oplus G_3$ is legal; and
- for every $S \subseteq \partial(G_1) = \partial(G_2)$, $G_1 \sim_S G_2$.

Now, $\sim_S$ has a finite number of equivalence classes for every $S \subseteq \partial(G_1)$ and the class $\mathcal{G}_g$ is characterized by a finite set of forbidden minors. Hence the number of equivalence classes in the equivalence relation defined above is a function of t. Let $\mathcal{S}$ be a set of r-boundaried graphs containing one smallest representative for each equivalence class of the relation above. Let c the size of the largest graph in $\mathcal{S}$. We have that X is a r-boundaried graph with boundary $\partial(X)$. Let G_1 be a graph in $\mathcal{S}$ such that G_1 and $G[X]$ are equivalent.

Now we replace $G[X]$ with $G_1 = (V_1, E_1)$ in the graph G, and let the resulting graph be $G^* = (V^*, E^*)$. Let Y_1 be the set of black vertices in $\partial(G_1)$. We let $Y^* = Y \setminus X \cup Y_1$ be the set of black vertices in G^*. Since $|X| > c$, we have that $|V_1| < |X|$ and hence $|V^*| < |V|$. It remains to prove that $(G, k) \in \Pi^\alpha$ if and only if $(G^*, k) \in \Pi^\alpha$. In one direction, suppose there is a set $S \subseteq Y$ such that $P_\Pi(G, S)$ holds. Then $S \cap X \subseteq \partial(G[X])$ and since $G[X]$ and G_1 are equivalent with respect to the relation above we have that $P_\Pi(G^*, S)$ holds. In the other direction, suppose $P_\Pi(G^*, S)$ holds. Since $Y^* \cap V_1 \subseteq \partial(G_1)$ and $G[X]$ and G_1 are equivalent with respect to the relation above, we have that $P_\Pi(G, S)$ holds. This concludes the proof. □

Lemmata 1, 2 and 3 together yield a reduction rule for all annotated p-MIN-CMSO problems.

Lemma 4. *Let* Π^α *be an annotated* p-MIN-CMSO *problem. There is a fixed constant* c *depending only on* Π^α *such that there is an algorithm that given a graph* $G = (V, E) \in \mathcal{G}_g$*, a set* $Y \subseteq V$ *of black vertices, an integer* k *and an extended* r*-protrusion* X *with* $|X| > ck$*, runs in time* $O(|X|)$*, and produces a graph* $G^* = (V^*, E^*) \in \mathcal{G}_g$ *such that* $|V^*| < |V|$ *and* $(G^*, k) \in \Pi^\alpha$ *if and only if* $(G, k) \in \Pi^\alpha$*.*

Proof. The algorithm starts by applying Lemma 1 to X, thus making all but at most ak black vertices uncolored for some fixed constant a. By Lemma 2, $X = X_1 \cup X_2 \ldots \cup X_{bk}$ for some fixed constant b, where for every i, X_i is an extended $4r$-protrusion such that $Y \cap X_i \subseteq \partial(X_i)$. By the pigeon-hole principle some X_i has size at least $|X|/bk > c/b$. Choose c such that c/b is sufficiently large to apply the algorithm in Lemma 3, and then apply Lemma 3 on the extended protrusion X_i. This concludes the proof. □

3.2 Reduction for Annotated p-EQ-CMSO Problems

In this section we give a reduction rule for annotated p-EQ-CMSO problems. The rule is very similar to the one for the p-MIN-CMSO problems described in the previous section. Therefore we only highlight the differences between the two rules in our arguments. The main difference between the two problem variants is that we need to keep track of solutions of every fixed size between 0 and k, instead of just the smallest one in each class. Because of this we require the protrusion to contain at least ck^2 vertices instead of ck vertices, in order to be able to reduce it.

Lemma 5. *Let* Π^α *be an annotated* p-EQ-CMSO *problem. There is a fixed constant* c *depending only on* Π^α *such that there is an algorithm that given a graph* $G = (V, E) \in \mathcal{G}_g$*, a set* $Y \subseteq V$ *of black vertices, an integer* k *and an extended* r*-protrusion* X *with* $|X| > ck^2$*,* *runs in time $O(k|X|)$, and produces a graph $G^* = (V^*, E^*) \in \mathcal{G}_g$ such that $|V^*| < |V|$ and $(G^*, k) \in \Pi^\alpha$ if and only if $(G, k) \in \Pi^\alpha$.*

Proof. We show that if $Y \cap X \geq ak^2$ for some fixed constant k, then we can remove some vertices from Y preserving the answer. The proof proceeds almost as the proof of Lemma 1. The main difference is that now, instead of taking a minimum size representative from each equivalence class we consider all possible sizes for S between 0 and k, for each equivalence class. That is, for each ℓ, $0 \leq \ell \leq k$, and each equivalence class q, we make a set $Q_{q,\ell} \subseteq X$ from the class with $|Q_{q,\ell}| = \ell$. Now, set

$$Z = \bigcup_{q \in \mathcal{Q}, 0 \leq \ell \leq k, |Q_{q,\ell}| = \ell} Q_{q,\ell} .$$

In the dynamic programming algorithm, we must also have one table entry for each class and each size. This gives a running time of $O(k|X|)$ for the first part of the reduction rule.

Next, we remove all vertices in $X \setminus Z$ from Y and apply Lemma 2. This gives us $X = X_1 \cup X_2 \ldots \cup X_{bk^2}$ for some constant b, where for every i, X_i is an extended $4r$-protrusion with $Z \cap X_i \subseteq \partial(X_i)$.

By the pigeon-hole principle some X_i has size at least $|X|/bk^2 > c/b$. Choose c such that c/b is sufficiently large to apply the algorithm in Lemma 3, and then we apply Lemma 3 on the extended protrusion X_i. This concludes the proof. □

3.3 Reduction for Annotated p-MAX-CMSO Problems

We now give a reduction rule for annotated p-MAX-CMSO problems. The rule is still similar to the ones described in the two previous sections, but differs more from the p-MIN-CMSO problems than p-EQ-CMSO did.

Lemma 6. *Let Π^α be an annotated p-MAX-CMSO problem. There is a fixed constant c depending only on Π^α such that there is an algorithm that given a graph $G = (V, E) \in \mathcal{G}_g$, a set $Y \subseteq V$ of black vertices, an integer k and an extended r-protrusion X with $|X| > ck$, runs in time $O(|X|)$, and produces a graph $G^* = (V^*, E^*) \in \mathcal{G}_g$ such that $|V^*| < |V|$ and $(G^*, k) \in \Pi^\alpha$ if and only if $(G, k) \in \Pi^\alpha$.*

Proof. We again begin in a manner similar to the proof of Lemma 1. The main ingredient in the proof of Lemma 1 is that for a given extended r-protrusion X, we consider the equivalence relation $\sim$ on subsets $Q \subseteq X$, where we demand that $Q \sim Q'$ if and only if for all $R \subseteq V - X$: $P_\Pi(Q \cup R) \Leftrightarrow P_\Pi(Q' \cup R)$. The number of equivalence classes of $\sim$ is bounded, but for maximization problems we need to keep the *largest* representative of each class. Hence we can not guarantee that the union of all the representatives we store is bounded by a function of k. To overcome this difficulty we use the expressive power provided by annotation. We compute a largest representative $Q_q \subseteq X$ for each equivalence class q of $\sim$. Then, we build a vertex set $Z \subseteq X \cap Y$ as follows. For each non-empty equivalence class q, if $|Q_q \cap Y| \leq k$, then we add $Q_q \cap Y$ to Z. If $|Q_q \cap Y| > k$, then we select arbitrarily a k-sized subset of $Q_q \cap Y$ and add it to Z. Thus $|Z| \leq ak$ for some constant a. We remove all vertices in $X \setminus Z$ from Y, without changing the membership of (G, k) in Π^α.

Next we apply Lemma 2 on X. This gives us $X = X_1 \cup X_2 \ldots \cup X_{bk}$ for some constant b, where for every i, X_i is an extended $4r$-protrusion with $Z \cap X_i \subseteq \partial(X_i)$.

We now describe how to modify Lemma 3 so that it can also be applied to annotated p-MAX-CMSO problems. For a set $S \subseteq Y$ we define $P'_\Pi(G, S)$ as there exists a set S' containing S such that $P_\Pi(G, S')$ holds. Clearly $(G, k) \in \Pi^\alpha$ if and only if there is a set $S \subseteq Y$ of size k such that $P'_\Pi(G, S)$. If the relation $\sim_S$ in Lemma 3 is defined using $P'_\Pi(G, S)$ instead of $P_\Pi(G, S)$, the proof goes through also for annotated p-MAX-CMSO problems.

By the pigeon-hole principle some X_i has size at least $|X|/bk > c/b$. Choose c such that c/b is sufficiently large to apply the algorithm in the modified version of Lemma 3, and then apply the modified version of Lemma 3 to p-MAX-CMSO problems on the extended protrusion X_i. This concludes the proof. □

3.4 Reductions Based on Finite Integer Index

In the previous sections we gave reduction rules for annotated p-MIN/EQ/MAX-CMSO problems. These reduction rules, together with the results proved later in this article will give quadratic or cubic kernels for the problems in question. However, for many problems we can in fact show that they admit a linear kernel. In this section we provide reduction rules for graph problems that have finite integer index. These reduction rules will yield linear kernels for the problems they apply to. We are now ready to prove the reduction lemma for problems that have finite integer index.

Lemma 7. *Let $\Pi \subseteq \mathcal{G}_g \times \mathbb{N}$ has finite integer index in $\mathcal{G}_g$ and either Π or $\overline{\Pi}$ is quasi-compact. There exists a constant c and an algorithm that given a graph $G = (V, E) \in \mathcal{G}$, an integer k and an extended r-protrusion X in G with $|X| > c$, runs in time $O(|X|)$ and returns a graph $G^* = (V^*, E^*) \in \mathcal{G}_g$ and an integer k^* such that $|V^*| < |V|$, $k^* \leq k$ and $(G^*, k^*) \in \Pi$ if and only if $(G, k) \in \Pi$.*

Proof. Let $\mathcal{S}$ be a set of representatives for (Π, r) and let $c = \max_{Y \in \mathcal{S}} |Y|$. Similarly let $\mathcal{S}'$ be a set of representatives for $(\Pi, 2r)$ and let $c' = \max_{Y \in \mathcal{S}'} |Y|$. If $|X| > 3c'$ we find an extended $2r$-protrusion $X' \subseteq X$ such that $c' < |X'| \leq 3c'$ and work on X' instead of X. This can be done in time $O(|X|)$ since $G[X]$ has treewidth at most r. From now on, we assume that $|X| \leq 3c'$. This initial step is the only step of the algorithm that does not work with constant size structures, and hence the running time of the algorithm is upper bounded by $O(|X|)$. The algorithm proceeds as follows.

Because Π has finite integer index there is a graph $H = (V_H, E_H) \in \mathcal{S}$ such that $H \equiv_\Pi G[X]$. We show how to compute H from X. Since either Π or $\overline{\Pi}$ is quasi-compact there exists an integer p such that $k \leq |V|^p$. Let $k_{max} = (6c')^p$. For every $G_1 = (V_1, E_1) \in \mathcal{S}$, $G_2 = (V_2, E_2) \in \mathcal{S}$ and $k' \leq k_{max}$ we compute whether $(G_1 \oplus G_2, k') \in \Pi$. For each such triple the computation can be done in time $O((|V_1| + |V_2|)^p)$ since Π has finite integer index [13, 24]. Now, for every $G_1 \in \mathcal{S}$ and $k' \leq (|X| + |V_1|)^p$ we compute whether $(G[X] \oplus G_1, k') \in \Pi$. When all these computations are done, the results are stored in a table.

It is not hard to see that $H \equiv_\Pi G[X]$ if and only if there exists a constant c such that for all $G_2 \in \mathcal{S}$ and $k' \leq k_{max}$, $(H \oplus G_2, k') \in \Pi \Leftrightarrow (G[X] \oplus G_2, k' + c) \in \Pi$. Also, c is the constant such that for all r-boundaried graphs G_2 and and integers k', $(H \oplus G_2, k') \in \Pi \Leftrightarrow (G[X] \oplus G_2, k' + c) \in \Pi$. For each $H \in \mathcal{C}$ we can check whether $H \equiv_\Pi G[X]$ using this condition and the pre-computed table, and if $H \equiv_\Pi G[X]$, find the constant c.

After we have found a $H \in \mathcal{S}$ and the corresponding constant c, such that $H \equiv_\Pi G[X]$, we make G^* from G by replacing the extended r-protrusion X with H. Also, we set $k^* = k - c$. Since $|X| > c$ and H has at most c vertices, $|V^*| < |V|$. By the choice of H and c, $(G^*, k^*) \in \Pi$ if and only if $(G, k) \in \Pi$. This concludes the proof. □

4 Decomposition Theorems

Definition 8. *We say that a graph $G = (V, E)$ is (α, β, γ)-*structured around *S if $|S| \leq \alpha$ and V can be partitioned into $S, C_1, C_2, \ldots, C_\gamma$ such that $N_G(C_i) \subseteq S$, and $\max\{|N_G(C_i)|, \mathbf{tw}(G[C_i \cup N_G(C_i)])\} \leq \beta$, for every $i = 1, \ldots, \gamma$.*

Let $G = (V, E)$ be a graph embedded in some surface Σ. A *noose* in G is a closed curve N of Σ meeting only the vertices of G, we denote these vertices by $V_N = V \cap N$. We also define the *length* of N as the number of vertices it meets and we denote it by $|N|$, that is, $|N| = |V_N|$. The *face-width* of G is the minimum length of a non-contractible noose in G. We denote the Euler genus of a surface Σ by $\mathbf{eg}(\Sigma)$.

Lemma 8. *Let $G = (V, E)$ be a graph in $\mathcal{G}_g$ and let $S \subseteq V$ such that $\mathbf{B}^r_G(S) \leq q$. Then, $\mathbf{tw}(G) \leq 4(2r + 1)\sqrt{q + 2g} + 8r + 12g$.*

Proof. The result follows closely the argument of the proof of [25, Theorem 3.2] that, in turn, is based on [25, Lemma 3.1] about the distribution of every such set S in the interior of a $(\rho \times \rho)$-grid. As now we have graphs of higher genus we have to apply [27, theorem 4.7] and find a lower bound on the size of S in the graph obtained by a $(\rho \times \rho)$-grid after adding $O(g)$ edges. □

From now on, we set $f(r, g) = 4(2r + 1)\sqrt{2 + 2g} + 8r + 12g$.

Lemma 9. *Let $G = (V, E)$ be a graph, embedded in a surface Σ of Euler genus g such that either $g = 0$ or the face-width of G is strictly greater than $4r + 2$. Let $S \subseteq V$ be a set containing at least 3 vertices where $\mathbf{B}^r_G(S) = V$. Then there exists a set $S' \subseteq V$ such that $S \subseteq S'$ and G is $(r \cdot (4r + 2) \cdot (3|S| - 6 + 6g), f(r, g), r \cdot (3|S| - 6 + 6g))$-structured around S'.*

Proof. We first need the following claim.

Claim: Let $G = (V, E)$ be a graph, embedded in a surface Σ of Euler genus g such that either $g = 0$ or the face-width of G is more than $4r + 2$. Let S be a set of at least 3 vertices such that $\mathbf{B}_G^r(S) = V$. Then there is a collection $\mathcal{R}$ of closed subsets of Σ, also known as *regions*, such that

- If two sets in $\mathcal{R}$ have common points, then these points lay on their boundaries.
- $\bigcup_{R\in\mathcal{R}} R \cap V = V$.
- The boundary of each $R \in \mathcal{R}$ is the union of two paths of length $\leq 2r + 1$ between two vertices of S called the *anchors* of R. We denote the set of vertices on the boundary of R by $\mathbf{bor}(R)$.
- For each $R \in \mathcal{R}$ with u and v as anchors it holds that $R \cap V \subseteq \mathbf{B}_G^r(\{u, v\})$.
- For each $R \in \mathcal{R}$ with u and v as anchors it holds that $S \cap R = \{u, v\}$.
- $|\mathcal{R}| \leq r \cdot (3|S| + 6g - 6)$.

The above claim follows from [38, Lemma 1] for the case where $g = 0$. For the sake of completeness, we briefly present this proof together with its natural extension for embeddings of higher genus provided that the face-width of G is $> 4r + 2$. A collection $\mathcal{R}$ of closed subsets of Σ is constructed by a greedy algorithm as follows: *Start with an empty $\mathcal{R}$ and, as long as there are vertices not contained in some region R in $\mathcal{R}$, find a area-maximal region R defined by two paths of length $\leq 2r + 1$ between two vertices in S (its anchors) that do not contain any vertex in S and add it to $\mathcal{R}$.* Notice that such a region R always exists even for non-planar embeddings, provided that the face-width of G is bigger than $4r + 2$, as this region is always inside a big enough disk of the surface. To prove the claimed upper bound on the size of $\mathcal{R}$, consider a multigraph $G_{\mathcal{R}}$ (again embedded in Σ) with vertex set S and there is an edge between $u, v \in S$ whenever u and v are the anchors of some region $R \in \mathcal{R}$. Notice that $G_{\mathcal{R}}$ is also embedded in Σ and has face-width bigger than 1. This in turn implies that if two vertices u, v are joined by many copies of the same edge in $G_{\mathcal{R}}$, then a pair of these edges will define a disk Δ in Σ containing in its interior all other copies. Moreover, we may assume that $\Sigma \setminus \Delta$ contains a vertex w of S (this is necessary in case G has no planar embedding). We claim that the following *thinness property* holds: if there are $x \geq 2r$ copies of the edge $\{u, v\}$, then there exists a vertex $w' \in S$ laying in the interior of one of the $x - 1 \geq 2r - 1$ area minimal disks $\Delta_1, \ldots, \Delta_x$ defined by the copies of $\{u, v\}$ inside Δ (the order is chosen such that consecutive disks have common edges). We prove this using the argument of the proof of [38, Lemma 1]. Assume to the contrary and consider the disk Δ_r. Indeed, by area-maximally of the choice of the regions in the above greedy procedure, it follows that Δ_r contains a vertex $z \in V$ whose distance in G from u and v and every other vertex in S is bigger than r, a contradiction. This implies that there exists a vertex $w' \in S$ laying in the interior of one of the the $x - 1 \geq 2r - 1$ area minimal disks $\Delta_1, \ldots, \Delta_x$ defined by the copies of $\{u, v\}$ inside Δ. Using the thinness property of $G_{\mathcal{R}}$ (see also [3, Lemma 5]) along with the Euler formula for graphs embedded in higher genus surfaces, we derive the claimed bound for $|\mathcal{R}|$ and the claim follows.

To complete the proof of the lemma, we define $S' = \bigcup_{R\in\mathcal{R}} \mathbf{bor}(R) \cap V$, that is, S' contains all the vertices belonging to the boundary of the sets in $\mathcal{R}$. As each such boundary is the union of two (u, v)-paths of length $\leq 2r + 1$ where $u, v \in S$, we have that such a boundary can have at most $4r + 2$ vertices. As $|\mathcal{R}| \leq r \cdot (3|S| + 6g - 6)$, we obtain that $|S'| \leq r \cdot (4r+2) \cdot (3|S| - 6 + 6g)$. For each $R_i \in \mathcal{R}$ assign a set $\mathcal{C}_i$ containing each connected component $C = (V_C, E_C)$ of $G \setminus S'$ for which $N_G(V_C) \subseteq \mathbf{bor}(R)$. If a connected component of $G\setminus S'$ can be assigned to more than one R_i, break ties arbitrarily so that $\{\mathcal{C}_i, i = 1, \ldots, |\mathcal{R}|\}$ forms a partition of the set of the connected components of $G \setminus S'$. Notice that for each $C \in \mathcal{C}_i$, $N_G(V_C) \subseteq R_i \subseteq S'$ and if we set $C_i = \cup_{C\in\mathcal{C}_i} V_C$ we have that $N_G(C_i) \subseteq R_i \subseteq S'$ and thus $|N_G(C_i)| \leq 4r + 2 \leq f(r, g)$ (for $i = 1, \ldots, |\mathcal{R}|$). It remains to prove that if $J_i = G[C_i \cup N_G(C_i)]$ then $\mathbf{tw}(J_i) \leq f(r, g)$. For this, recall that $R \cap V \subseteq \mathbf{B}^r_G(\{u, v\})$ and from Lemma 8, we obtain that $\mathbf{tw}(J_i) \leq f(r, g)$, (for $j = 1, \ldots, |\mathcal{R}|$). □

We are now in position to prove the following.

Lemma 10. *Let $G = (V, E)$ be a graph, embedded in a surface Σ of Euler genus g and let $S \subseteq V$, $|S| \geq 3$ such that $\mathbf{R}^r_G(S) = V$ for some $r \geq 0$. Then there exists a set S' such that $S \subseteq S'$ and G is $(h(r, g) \cdot |S|, h(r, g), h(r, g) \cdot |S|)$-structured around S', where $h(r, g) = O(rg)$.*

Proof. We use induction on g. In particular, we prove that there exists a set $S' \subseteq V$ such that G is $(\alpha_{g,|S|}, \beta_g, \gamma_{g,|S|})$-structured around S' where, for $g = 0$, we set $\alpha_{0,x} = r \cdot (4r + 2) \cdot (3x - 6)$, $\beta_0 = f(r, 0)$ and $\gamma_{0,x} = r \cdot (3x - 6)$ and for $g \geq 1$ we have $\alpha_{g,x} = (2g - 1) \cdot r \cdot 6(4r + 2)^2 + r \cdot (4r + 2) \cdot (3x + 6g - 6)$, $\beta_g = f(r, g)$ and $\gamma_{g,x} = (2g - 1) \cdot r \cdot 6(4r + 2) + r \cdot (3x + 6g - 6)$. Once we have proved this, the lemma follows by suitably choosing the function h. For our proof we distinguish two cases:

Case 1. $g = 0$ or the embedding of G has face-width $> 4r + 2$. Then we draw G together with its radial graph $R_G = (V_R, E_R)$ and denote it by $\overline{G} = G \cup R_G = (\overline{V}, \overline{E})$, the superposition of the two drawings. By the definition of the radial graph, it follows that $\mathbf{B}^r_{\overline{G}}(S) = \overline{V}$. Clearly, as G is a subgraph of $\overline{G}$, the result will follow if we prove that $\overline{G}$ is $(\alpha_{g,|S|}, \beta_g, \gamma_{g,|S|})$-structured around some set $S' \subseteq V_R$. If G is embeddable in the sphere, then $\overline{G}$ is also embeddable in the sphere and if G is embeddable in a surface Σ of Euler genus $g > 4r + 2$ then the face-width of $\overline{G}$ is also greater than $4r + 2$. Therefore, in either case we have all the conditions required to apply Lemma 9 and hence the claim follows by applying Lemma 9.

Case 2. There is a non-contractible noose N in Σ of length at most $4r + 2$. Then we split the graph along the vertices V_N of the noose and we distinguish two subcases depending if N is a surface separating or not.

Subcase 2.1. If N is a surface separating noose, then the splitting of the vertices of N creates two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ embedded in surfaces Σ_1 and Σ_2 respectively such that if $\mathbf{eg}(\Sigma_i) = g_i, i = 1, 2$ then $g_1 + g_2 \leq g$ and $g_1 \cdot g_2 > 0$. Let S_i consists of the vertices in $S \cap V_i$ and all the (duplicated) vertices met by N. Notice that $\mathbf{R}^r_{G_i}(S_i) = V_i$ and that $|S_1| + |S_2| \leq 2|N| + |S| \leq 2(4r + 2) + |S|$.

By the induction hypothesis, G_i is $(\alpha_{g_i,|S_i|}, \beta_{g_i}, \gamma_{g_i,|S_i|})$-structured around some set S_i' where $S_i \subseteq S_i'$ and $g_i > 1, i = 1, 2$. Let G^+ be the disjoint union of G_1 and G_2 and observe that G^+ is $(\alpha_{g_1,|S_1|} + \alpha_{g_2,|S_2|}, \max\{\beta_{g_1}, \beta_{g_2}\}, \gamma_{g_1,|S_1|} + \gamma_{g_2,|S_2|})$-structured around $S_1' \cup S_2'$. Notice that

$$\begin{aligned}
\alpha_{g_1,|S_1|} + \alpha_{g_2,|S_2|} &\leq (2g_1 - 1 + 2g_2 - 1) \cdot r \cdot 6(4r+2)^2 + \\
&\quad r \cdot (4r+2)(3|S_1| + 3|S_2| + 6g_1 + 6g_2 - 12) \\
&\leq (2g_1 + 2g_2 - 2) \cdot r \cdot 6(4r+2)^2 + \\
&\quad r \cdot (4r+2)(3(2(4r+2) + |S|) + 6g - 6) \\
&\leq (2g_1 + 2g_2 - 1) \cdot r \cdot 6(4r+2)^2 - r \cdot 6(4r+2)^2 + \\
&\quad r \cdot 6(4r+2)^2 + r \cdot (4r+2) \cdot (3|S| + 6g - 6) \\
&\leq (2g - 1) \cdot r \cdot 6(4r+2)^2 + r \cdot (4r+2) \cdot (3|S| + 6g - 6) \\
&= \alpha_{g,|S|}.
\end{aligned}$$

Similarly, we can show that $\gamma_{g_1,|S_1|} + \gamma_{g_2,|S_2|} \leq \gamma_{g,|S|}$ and, as $\max\{\beta_{g_1}, \beta_{g_2}\} \leq \beta_g$, we conclude that G^+ is $(\alpha_{g,|S|}, \beta_g, \gamma_{g,|S|})$-structured around $S_1' \cup S_2'$. As all the duplicated vertices of V_N are in $S_1 \cup S_2$, we can identify back these duplicated vertices occurring in S_1 and S_2 and obtain that G is $(\alpha_{g,|S|}, \beta_g, \gamma_{g,|S|})$-structured around some set S'.

Subcase 2.2. If N is not a surface separating noose, then the splitting of the vertices of N creates a new graph $G_0 = (V_0, E_0)$ embedded in a surface Σ_0 of Euler genus g_0 where $g_0 < g$. Notice that if S_0 is the set of vertices in G_0 consisting of the non-duplicated vertices of S and all the duplicated vertices, then $\mathbf{R}^r_{G_0}(S_0) = V_0$ and $|S_0| \leq 2|N| + |S| \leq 2(4r+2)+|S|$. From the induction hypothesis G_0 is $(\alpha_{g_0,|S_0|}, \beta_{g_0}, \gamma_{g_0,|S_0|})$-structured around some set S_0' where $S_0 \subseteq S_0'$. Notice that

$$\begin{aligned}
\alpha_{g_0,|S_0|} &\leq (2g_0 - 1) \cdot r \cdot 6(4r+2)^2 + r \cdot (4r+2) \cdot (3|S_0| + 6g_0 - 6) \\
&\leq (2(g-1) - 1) \cdot r \cdot 6(4r+2)^2 + r \cdot (4r+2) \cdot (3(2(4r+2) + |S|) + 6g - 6) \\
&\leq (2g - 1) \cdot r \cdot 6(4r+2)^2 - 2 \cdot r \cdot 6(4r+2)^2 + \\
&\quad r \cdot 6(4r+2)^2 + r \cdot (4r+2) \cdot (3|S| + 6g - 6) \\
&\quad (2g - 1) \cdot r \cdot 6(4r+2)^2 + r \cdot (4r+2) \cdot (3|S| + 6g - 6) \\
&= \alpha_{g,|S|}.
\end{aligned}$$

and similarly $\gamma_{g_0,|S_0|} \leq \gamma_{g,|S|}$. As $\beta_{g_0} \leq \beta_g$, we conclude that G_0 is $(\alpha_{g,|S|}, \beta_g, \gamma_{g,|S|})$-structured around S_0'. As all the duplicated vertices of V_N are in S_0, we can identify them back and deduce that G is $(\alpha_{g,|S|}, \beta_g, \gamma_{g,|S|})$-structured around some set S'. This concludes the proof. □

5 Kernels

In this section we prove Theorems 1 and 2. We say that an instance (G', k') of a parameterized problem Π is *reduced with respect to a set $\mathcal{Q}$ of reduction rules* if none of the reduction rules in $\mathcal{Q}$ can be applied to (G', k').

5.1 Proof of Theorem 1

Proof. We first give a proof for the case when Π is compact and Π^α is an annotated p-MIN-CMSO problem. A proof for the case when Π is compact and Π^α is an annotated p-EQ/MAX-CMSO problem is identical.

We know that Π is compact and hence there exists an integer r such that for all $(G = (V, E), k) \in \Pi^\alpha$, there is an embedding of G into a surface of genus at most g, and a set $S \subseteq V$ such that $\mathbf{R}^r_G(S) = V$ and $|S| \leq r \cdot k$. We show that for all $(G, k) \in \Pi^\alpha$, the equivalent instance $(G' = (V', E'), k)$, reduced with respect to the reduction rule given by Lemma 4, has $|V'| = O(k^2)$. Since $(G' = (V', E'), k) \in \Pi^\alpha$ and Π is compact, there exists an embedding of G' into a surface of genus at most g and a set $S \subseteq V'$ such that $\mathbf{R}^r_{G'}(S) = V'$. Hence by applying Lemma 10 we obtain a set S' such that G' is (α, β, γ)-structured around S', where $\alpha, \gamma = O(rg|S|)$ and $\beta = O(rg)$. This implies that V' can be partitioned into $S', C_1, C_2, \ldots, C_\gamma$ such that $N_{G'}(C_i) \subseteq S'$, $|N_{G'}(C_i)| \leq \beta$ and $\mathbf{tw}(G'[C_i \cup N_G(C_i)]) \leq \beta$ for every $i \leq \gamma$. Observe that each C_i is a β-protrusion in G' and $|C_i \cup N_G(C_i)| \leq ck$, where c is a constant of Lemma 4, otherwise we could have applied Lemma 4. This implies that

$$|V'| \leq |S'| + \sum_{i=1}^{\gamma} |C_i| = O\left(rg|S| + \sum_{i=1}^{\gamma} ck\right) = O(k^2),$$

for some fixed g and r. Here constants hidden in big-Oh depend only on r and g.

So given an input (G, k), if the size of the reduced graph is more than $c^* k^2$ for some constant c^* then we return NO else we have G' as the desired annotated kernel for G.

Now we give a proof for the case when $\overline{\Pi}$ is compact and Π^α is an annotated p-MAX-CMSO problem. A proof for the case when $\overline{\Pi}$ is compact and Π^α is an annotated p-MIN/EQ-CMSO problem is similar. Towards this end, we claim that for all $(G, k) \in \overline{\Pi}^\alpha$, the equivalent instance $(G' = (V', E'), k)$, reduced with respect to the reduction rule given by Lemma 6, has $|V'| = O(k^2)$. The proof for the claim is identical to the one we gave above to bound all the YES instance for an annotated compact p-MIN-CMSO problem. So given an input (G, k), if the equivalent instance $(G' = (V', E'), k)$, reduced with respect to the reduction rule given by Lemma 6, is more than $c^* k^2$ for some constant c^* then we return YES else we have G' as the desired annotated kernel for G. The reason we return YES is that if (G, k) would have a NO instance then the size of $|V'| \leq c^* k^2$ as $(G, k) \in \overline{\Pi}^\alpha$. ☐

5.2 Proof of Corollary 1

Proof. We know that Π is NP-complete and the annotated version Π^α is in NP. So given an instance $(G = (V, E), k)$, we apply Theorem 1 on the annotated instance $(G = (V, E), V, k)$, that is we take V as Y, the set of black vertices. If we get YES or NO as an answer then we return the same. Else for $(G' = (V', E'), k)$ of size polynomial in k, we apply polynomial time many to one reduction from Π^α to Π on G' and obtain a graph $G'' = (V'', E'') \in \mathcal{G}_g$ and an integer k' such that $|V''|, k' \leq k^{O(1)}$ and $(G', k) \in \Pi^\alpha$ if and only if $(G'', k') \in \Pi$. This implies that in this case we have polynomial kernel for Π. ☐

5.3 Proof of Theorem 2

In this section we prove Theorem 2. The idea of the proof is that if a problem has finite integer index then an instance reduced with respect to the reduction rule given by Lemma 7 has bounded radial distance. We first prove a lemma which will assist us in proving Theorem 2.

Lemma 11. *Let $G = (V, E)$ be a graph embedded into a surface of genus at most g and $S \subseteq V$ such that $\mathbf{tw}(G \setminus \mathbf{R}_G^r(S)) \leq r$ and all r-protrusions of G have size at most m. Then there exists a $S' \subseteq V$ such that $|S'| \leq |S| + g \cdot (m + 2r + 1)$ and $V = \mathbf{R}_G^{m+3r+1}(S')$.*

Proof. We prove the lemma using induction on Euler genus g of the graph $G \setminus \mathbf{R}_G^r(S)$. We distinguish two cases:

Case 1. $g = 0$ or the embedding of $G \setminus \mathbf{R}_G^r(S)$ has face-width $> m + 2r + 1$. We claim that $V = \mathbf{R}_G^{m+3r+1}(S)$, that is, $S' = S$. For this assume, towards a contradiction, that $x \in V \setminus \mathbf{R}_G^{m+3r+1}(S)$ and consider the subgraph J of G induced by the set $\mathbf{R}_G^{m+2r+1}(x)$, that is, the vertices of G that are within radial distance at most $m+2r+1$ from x. Notice that J is a subgraph of $G \setminus \mathbf{R}_G^r(S)$, therefore $\mathbf{tw}(J) \leq r$. As either $g = 0$ or because the face-width of G is $> m + 2r + 1$, all the vertices and edges of J are embedded inside a closed disk in Σ. Moreover, there exist $m + 2r + 1$ nested disjoint cycles $C_1, \ldots, C_{m+2r+1}$ with vertex sets $V_1, \ldots, V_{m+2r+1}$ respectively in R_G such that, if Δ_i is the closed disk with C_i as its border and contains x then $i < j$ implies that $\Delta_i \subset \Delta_j$. For $i = 1, \ldots, m+2r+1$, V_i contains vertices and faces whose radial distance from x, in G, is exactly i. We need the following claim.

Claim. Each cycle of the radial graph R_G that is entirely in $(\Delta_{m+r+1} \setminus \Delta_m)$ and separates S and x, has length $> 2r$.

Proof. Indeed, if this is not the case for some cycle C with vertex set V_C, then $L = V \cap V_C$ is a separator of G where $|L| \leq r$ and such that the connected component, say F, of $G \setminus L$ that contains x is a subgraph of J. Then $\mathbf{tw}(F) \leq \mathbf{tw}(J) \leq r$ and F is an r-protrusion of G. As F contains x and has more than m vertices, it is a contradiction to the assumption that all r-protrusions of G have size at most m. □

Applying the above claim to the cycles $C_{m+1}, \ldots, C_{m+2r+1}$ we obtain that they all have length $> 2r$. We now construct an auxiliary graph R^* by taking $R_G \cap (\Delta_{m+r+1} \setminus \Delta_m)$, adding a vertex s adjacent to all vertices of C_{m+1} and adding a vertex t adjacent to all vertices in C_{m+2r+1}. We claim that there is no (s, t)-separator in R^* of size $\leq 2r$. Indeed, such a separator would imply the existence of a cycle C in R_G of size $\leq 2r$, a contradiction to the above claim. By Menger's theorem, it follows that there are $> 2r$ disjoint (s, t)-paths in R^* that correspond to $> 2r$ disjoint paths from the vertices of C_{m+1} to the vertices of C_{m+2r+1}. The intersection of these paths with cycles $C_{m+1}, \ldots, C_{m+2r+1}$, imply the existence of a $(2r + 1) \times (2r + 1)$-grid as a minor of R_J, where R_J is the radial graph corresponding to J. Therefore, $\mathbf{tw}(R_J) > 2r$ (from [9, Lemma 88]), which, using [43, Lemma 3], implies that $\mathbf{tw}(J) > r$, a contradiction.

Case 2. There is a non-contractible noose N, meeting the vertices of V in V_N, in $G \setminus \mathbf{R}_G^r(S)$ of length at most $r' = m + 2r + 1$ (assume that $G \setminus \mathbf{R}_G^r(S)$ is embedded in a surface Σ).

Let $S'' = S \cup V_N$. Observe that $\mathbf{tw}(G \setminus \mathbf{R}^r_G(S'')) \leq r$ and all r-protrusions of G have size at most m. Furthermore let Σ' be the surface in which $G \setminus \mathbf{R}^r_G(S'')$ can be embedded. Then $\mathbf{eg}(\Sigma') < \mathbf{eg}(\Sigma) \leq g$. Hence by induction hypothesis, there exists a set S' such that

$$|S'| \leq |S''| + (g-1)(m+2r+1) \leq |S| + |V_N| + (g-1)(m+2r+1) \leq |S| + g(m+2r+1),$$

and $V = \mathbf{R}^{m+3r+1}_G(S')$. This concludes the proof. □

Proof of Theorem 2. Let us assume that Π is quasi-compact. Fix $t = c^* rg$ where c^* is a constant to be defined later. Let $(G' = (V', E'), k')$, $k' \leq k$, be a reduced instance with respect to reduction rule given by Lemma 7. Hence there is no extended t-protrusion of size more than c, where c is a constant appearing in the statement of Lemma 7 and $(G, k) \in \Pi$ if and only if $(G', k') \in \Pi$ and $G' \in \mathcal{G}_g$. Hence, what remains to show is that $|V'| = O(k)$. The proof for this is similar to the one given for Theorem 1.

Now we show that if $(G' = (V', E'), k') \in \Pi$ then $|V'| \leq O(k)$. Since Π is quasi-compact and $(G', k') \in \Pi$, there is an embedding of G into a surface of genus at most g and a set $S \subseteq V'$ such that $|S| \leq r \cdot k'$ and $\mathbf{tw}(G' \setminus \mathbf{R}^r_{G'}(S)) \leq r$. Since all t-protrusions of G' are of size most c we have that all r-protrusions of G' are of size at most c. Hence by Lemma 11 there exists a set S' such that $|S'| \leq |S| + g \cdot (c + 2r + 1)$ and $V' = \mathbf{R}^{3r+c+1}_{G'}(S')$. Given S', we apply Lemma 10 and obtain a set S'' such that G' is (α, β, γ)-structured around S'', where $\alpha, \gamma = O(rg|S'|)$ and $\beta = O(rg)$. This implies that V' can be partitioned into $S'', C_1, C_2, \ldots, C_\gamma$ such that $N_{G'}(C_i) \subseteq S''$, $|N_{G'}(C_i)| \leq \beta$ and $\mathbf{tw}(G'[C_i \cup N_{G'}(C_i)]) \leq \beta$ for every $i \leq \gamma$. Fix c^* such that $t \geq \beta$. This implies that $G'[C_i \cup N_{G'}(C_i)]$ is a t-protrusion of G' and hence its size is bounded by c. Now we are ready to bound the size of V'.

$$|V'| \leq |S''| + \sum_{i=1}^{\gamma} |C_i| \leq |S''| + \gamma \cdot \beta \cdot c = O(r^3 g^3 c^2 k) = O(k),$$

for fixed r, g and c. So given an input (G, k), if the size of the reduced graph is more than $\tilde{c} \cdot k$ for some $\tilde{c}$, we return NO else we have G' as the desired kernel.

The proof for the case when $\overline{\Pi}$ is quasi-compact is similar to the proof we gave for the case when $\overline{\Pi}$ was compact and Π^α was an annotated p-MAX-CMSO problem in Theorem 1. This concludes the proof. □

6 Implications of Our Results

In this section we mention a few parameterized problems for which we can obtain either polynomial or linear kernel using either Theorem 1 or Theorem 2. Various other problems for which we can obtain either polynomial or linear kernels using our results are mentioned in appendix.

6.1 A Sufficient Condition for Finite Integer Index

We first give a sufficient condition which implies that a large class of p-MIN/MAX-CMSO problems has finite integer index. We prove it here for vertex versions of p-MIN/MAX-CMSO problems that is if Π is a p-MIN/MAX-CMSO problem then P_Π is a property of vertex sets. The edge version can be dealt in a similar manner.

Let Π be a p-MIN-CMSO problem and $\mathcal{F}_t$ be the set of pairs $(G = (V, E), S)$ where G is a t-boundaried graph and $S \subseteq V$. For a t-boundaried graph $G = (V, E)$ we define the function $\zeta_G : \mathcal{F}_t \rightarrow \mathbb{N} \cup \{\infty\}$ as follows. For a pair $(G' = (V', E'), S') \in \mathcal{F}_t$, if there is no set $S \subseteq V$ ($S \subseteq E$) such that $P_\Pi(G \oplus G', S \cup S')$ holds, then $\zeta_G((G', S')) = \infty$. Otherwise $\zeta_G((G', S'))$ is the size of the smallest $S \subseteq V$ ($S \subseteq E$) such that $P_\Pi(G \oplus G', S \cup S')$ holds. If Π is a p-MAX-CMSO problem then we define $\zeta_G((G', S'))$ to be the size of the largest $S \subseteq V$ ($S \subseteq E$) such that $P_\Pi(G \oplus G', S \cup S')$ holds. If there is no set $S \subseteq V$ ($S \subseteq E$) such that $P_\Pi(G \oplus G', S \cup S')$ holds then $\zeta_G((G', S')) = \infty$.

Definition 9. *A p-MIN-CMSO problem Π is said to be* strongly monotone *if there exists a function $f : \mathbb{N} \rightarrow \mathbb{N}$ such that the following condition is satisfied. For every t-boundaried graph $G = (V, E)$, there is a subset $S \subseteq V$ such that for every $(G' = (V', E'), S') \in \mathcal{F}_t$ such that $\zeta_G((G', S'))$ is finite, $P_\Pi(G \oplus G', S \cup S')$ holds and $|S| \leq \zeta_G((G', S')) + f(t)$.*

Definition 10. *A p-MAX-CMSO problem Π is said to be* strongly monotone *if there exists a function $f : \mathbb{N} \rightarrow \mathbb{N}$ such that the following condition is satisfied. For every t-boundaried graph $G = (V, E)$, there is a subset $S \subseteq V$ such that for every $(G' = (V', E'), S') \in \mathcal{F}_t$ such that $\zeta_G((G', S'))$ is finite, $P_\Pi(G \oplus G', S \cup S')$ holds and $|S| \geq \zeta_G((G', S')) - f(t)$.*

Lemma 12. *Every strongly monotone p-MIN-CMSO and p-MAX-CMSO problem has finite integer index.*

Proof. We prove for p-MIN-CMSO problems, the proof for p-MAX-CMSO is similar. Let Π be a monotone p-MIN-CMSO problem. Then P_Π is a finite state property of t-boundaried graphs with a distinguished vertex set S [14, 22]. In particular for every t, there exists a finite set $\mathcal{S}$ of pairs (G, S) such that $G = (V, E)$ is a t-boundaried graph and $S \subseteq V$ such that the set $\mathcal{S}$ satisfies the following properties. For any t-boundaried graph $G_1 = (V_1, E_1)$ and set $S_1 \subseteq V_1$ there is a pair $(G_2 = (V_2, E_2), S_2) \in \mathcal{S}$ such that for every t-boundaried graph $G_3 = (V_3, E_3)$ and set $S_3 \subseteq V_3$ we have that $P_\Pi(G_1 \oplus G_3, S_1 \cup S_3) \iff P_\Pi(G_2 \oplus G_3, S_2 \cup S_3)$. We fix such a set $\mathcal{S}$.

For a t-boundaried graph $G = (V, E)$ we define the *signature* $\zeta_G^{\mathcal{S}} : \mathcal{S} \rightarrow \mathbb{N} \cup \{\infty\}$ of G to be ζ_G with domain restricted to $\mathcal{S}$. We now argue that for any t-boundaried graph G, the maximum finite value of $\zeta_G^{\mathcal{S}}$ is at most $f(t)$ larger than the minimum finite value taken by $\zeta_G^{\mathcal{S}}$. Since Π is strongly monotone there exists a subset S of V that satisfies the conditions of Definition 9. Let $(G' = (V', E'), S') \in \mathcal{S}$ such that $\zeta_G^{\mathcal{S}}((G', S'))$ is finite. Then $P_\Pi(G \oplus G', S \cup S')$ holds and hence $\zeta_G^{\mathcal{S}}((G', S')) \leq |S|$. Furthermore the conditions of Definition 9 imply that $|S| - f(t) \leq \zeta_G^{\mathcal{S}}((G', S')) \leq |S|$. Hence the minimum and the maximum finite values of $\zeta_G^{\mathcal{S}}$ can differ by at most $f(t)$. By the pigeon hole principle there is a finite set $\mathcal{R}$ of t-boundaried graphs such that for any t-boundaried graph G there is a $G_R \in \mathcal{R}$ and a constant c_R depending only on the size of G and G_R, such that for all $(G', S') \in \mathcal{S}$ we have $\zeta_{G_R}((G', S')) + c_R = \zeta_G((G', S'))$.

We now argue that $\mathcal{R}$ forms a set of representatives for (Π, t). Let $G = (V, E)$ be a t-boundaried graph and let $G_R = (V_R, E_R) \in \mathcal{R}$ and c_R be a constant such that for all $(G', S') \in \mathcal{S}$ we have $\zeta_{G_R}((G', S')) + c_R = \zeta_G((G', S'))$. Let $G' = (V', E')$ be a t-boundaried graph and k be an integer such that $(G \oplus G', k) \in \Pi$. We argue that $(G_R \oplus G', k - c_R) \in \Pi$. Let $Z \subseteq V \cup V'$ be a minimum size set such that $P_\Pi(G \oplus G', Z)$ is satisfied, $Z' = Z \cap V'$ and $Z_G = Z \setminus Z'$. Since $(G \oplus G', k) \in \Pi$ we have that $|Z| \leq k$ and hence $\zeta_G(G', Z \cap V')$ is finite. Let $(G_\mathcal{S}, Z_\mathcal{S}) \in \mathcal{S}$ be the representative of (G', Z'). Then $P_\Pi(G \oplus G_\mathcal{S}, Z_G \cup Z_\mathcal{S})$ holds and $|Z_G| = \zeta_G(G_\mathcal{S}, Z_\mathcal{S})$. Now we have that $\zeta_{G_R}((G_\mathcal{S}, Z_\mathcal{S})) + c_R = \zeta_G((G_\mathcal{S}, Z_\mathcal{S}))$. Hence, there is a set $S_R \subseteq V_R$ of size $|Z_G| - c_R$ such that $P_\Pi(G_R \oplus G_\mathcal{S}, S_R \cup Z_\mathcal{S})$ holds. Then, since $(G_\mathcal{S}, Z_\mathcal{S})$ is the representative of (G', Z') we have that $P_\Pi(G_R \oplus G', S_R \cup Z')$ holds. Now we have that $|S_R \cup Z'| \leq |S_R| + |Z'| = |Z_G| - c_R + |Z'| = |Z| - c_R \leq k - c_R$. This implies that $(G_R \oplus G', k - c_R) \in \Pi$. The proof for the other direction that if $(G_R \oplus G', k - c_R) \in \Pi$ then $(G \oplus G', k) \in \Pi$ is symmetric. This concludes the lemma. □

6.2 Covering and Packing Problems

Minor Covering and Packing: We give below a few generic problems which subsumes many problems in itself and fit in our kernelization framework. Let $\mathcal{H}$ be a finite set of connected planar graphs.

VERTEX-$\mathcal{H}$-COVERING *Input:* A graph $G = (V, E) \in \mathcal{G}_g$ and a non-negative integer k. *Question:* Is there an $S \subseteq V$ such that $\|S\| \leq k$ and $G[V \setminus S]$ does not contain any of the graphs in $\mathcal{H}$ as a minor?

VERTEX-$\mathcal{H}$-PACKING *Input:* A graph $G \in \mathcal{G}_g$ and a non-negative integer k. *Question:* Does there exist k vertex disjoint subgraphs $G_1, \ldots, G_k$ of G such that each of them contain some graph in $\mathcal{H}$ as a minor.

Lemma 13. *Let $\mathcal{H}$ be a finite set of connected planar graphs and let Π_1 denote* VERTEX-$\mathcal{H}$-PACKING. *Then* VERTEX-$\mathcal{H}$-COVERING *has finite integer index and is quasi-compact and* VERTEX-$\mathcal{H}$-PACKING *has finite integer index and $\overline{\Pi}_1$ is quasi-compact.*

Proof. Let Π_v denote VERTEX-$\mathcal{H}$-COVERING. We first show that Π_v is quasi compact. If $(G = (V, E), k) \in \Pi_v$ then we know that there exists a set $S \subseteq V$ such that $G[V \setminus S]$ does not contain any graph in $\mathcal{H}$ as a minor. Now we show that the treewidth of $G[V \setminus S]$ is at most a constant. To show this we need following results.

- Every planar graph $H = (V_H, E_H)$ is a minor of the $r \times r$ grid, where $r = 14|V_H| - 24$ [51].
- For any fixed graph H, every H-minor free graph that does not contain a $w \times w$ grid as a minor has treewidth $O(w)$ [26].

Let $q = \max\{|H| \mid H \in \mathcal{H}\}$ and $w = 14q - 24$. Then observe that $G[V \setminus S]$ does not contain $w \times w$ grid as a minor and hence $\mathbf{tw}(G[V \setminus S]) \leq O(w)$. This implies that Π_v is quasi-compact.

Next we show that $\overline{\Pi}_1$ is quasi-compact. We first introduce some definitions. Given a graph $G = (V, E)$, we define the *covering number of G with respect to the class $\mathcal{H}$*, $\mathbf{cov}_{\mathcal{H}}(G)$, as the minimum k such that there exists $S \subseteq V$ of size k such that $G[V \setminus S]$ does not contain any of the graphs in $\mathcal{H}$ as a minor. The *packing number of G with respect to the class $\mathcal{H}$*, is defined as,

$$\begin{aligned} \mathbf{pack}_{\mathcal{H}}(G) \;\; = \;\; & \max\{k \mid \exists \text{ a partition } V_1, \ldots, V_k \text{ of } V \text{ such that} \\ & \qquad \forall_{i \in \{1,\ldots,k\}} G[V_i] \text{ contains a graph in } \mathcal{H} \text{ as a minor}\}. \end{aligned}$$

Less formally, $\mathbf{pack}_{\mathcal{H}}(G) \geq k$ if G contains k vertex-disjoint minors in $\mathcal{H}$. We need the following Erdős-Pósa type of result shown in [33] for our purpose.

Claim 1. *Let $\mathcal{H}$ be a finite set of connected planar graphs, $q = \max\{|H| \mid H \in \mathcal{H}\}$ and $\mathcal{G}$ be a non-trivial minor-closed graph class. Then there is a constant $\sigma_{\mathcal{G},q}$ depending only on $\mathcal{G}$ and q such that for every graph $G \in \mathcal{G}$, it holds that*

$$\mathbf{pack}_{\mathcal{H}}(G) \leq \mathbf{cover}_{\mathcal{H}}(G) \leq \sigma_{\mathcal{G},q} \cdot \mathbf{pack}_{\mathcal{H}}(G).$$

Using Claim 1 we show that $\overline{\Pi}_1$ is quasi-compact. Observe that if $(G, k) \in \Pi_1$ and $(G, k+1) \notin \Pi_1$, then by Claim 1, $\mathbf{cover}_{\mathcal{H}}(G) = O(k)$. Hence by an argument, similar to the one used for showing that Π_v is quasi-compact, we have that $\overline{\Pi}_1$ is quasi-compact.

It is known that Π_v is a p-MIN-CMSO problem and Π_1 is a p-MAX-CMSO problem. Now using Lemma 12 we show that Π_v is finite integer index. We show that Π_v is strongly monotone. Given a t-boundaried graph $G = (V, E)$, with $\partial(G)$ as its boundary, let $S'' \subseteq V$ be a minimum set of vertices in G such that $G[V \setminus S'']$ does not contain any graph in $\mathcal{H}$ as a minor. Take $S = S'' \cup \partial(G)$. Now for any $(G' = (V', E'), S') \in \mathcal{F}_t$ such that $\zeta_G((G', S'))$ is finite we have that $G \oplus G'[(V \cup V') \setminus (S \cup S')]$ does not contain any graph in $\mathcal{H}$ as a minor and $|S| \leq \zeta_G((G', S')) + t$. This proves that Π_v is strongly monotone.

Next we show that Π_1 is finite integer index. Given a t-boundaried graph $G = (V, E)$, we define its signature, ζ_G, as follows. Let $\mathcal{F}_t^{qt} \subseteq \mathcal{G}_g$ be the set of all t-boundaried graph of size at most qt and $\zeta_G : \mathcal{F}_t^{qt} \to \mathbb{N}$. For every $G' \in \mathcal{F}_t^{qt}$ we define $\zeta_G(G') = \mathbf{pack}_{\mathcal{H}}(G \oplus G')$. For any $G', G'' \in \mathcal{F}_t^{qt}$, $|\zeta_G(G') - \zeta_G(G'')| \leq t + qt$, as $\zeta_G(\emptyset) \leq \zeta_G(G') \leq \zeta_G(\emptyset) + qt + t$ for all $G' \in \mathcal{F}_t^{qt}$. Hence, by the pigeon hole principle there is a finite set $\mathcal{R}$ of t-boundaried graphs such that for any t-boundaried graph G there is a $G_R \in \mathcal{R}$ and a constant c_R depending only on the size of G and G_R, such that $\zeta_{G_R} + c_R = \zeta_G$.

We now argue that $\mathcal{R}$ forms a set of representatives for (Π_1, t). Let $G = (V, E)$ be a t-boundaried graph and let $G_R = (V_R, E_R) \in \mathcal{R}$ and c_R be a constant such that $\zeta_{G_R} + c_R = \zeta_G$. Let $G' = (V', E')$ be a t-boundaried graph and k be an integer such that $(G \oplus G', k) \in \Pi_1$. We argue that $(G_R \oplus G', k - c_R) \in \Pi_1$. Let $\mathcal{S}$ be a set of k vertex disjoint minors in $G \oplus G'$ and H be a minor in $\mathcal{S}$ which goes across or touch the boundary $\partial(G)$. Now contract the part of this minor belonging to the side of G' as much close to the boundary $\partial(G)$ as possible. Since the size of largest graph in $\mathcal{H}$ is at most q, we have that the part of the H belonging to the side of G' can be contracted to the border except for some q vertices, which could possibly be hanging out of the boundary. We do this for every minor in $\mathcal{S}$ which is going across. Let $\mathcal{S}'$ be the set of minors resulting after the contraction operation has been performed. Now we take the boundary $\partial(G)$ and all the minors of $\mathcal{S}'$ hanging out of it, and call this resulting graph $\tilde{G} = (\tilde{V}, \tilde{E})$. Observe that we can have at most t minors in $\mathcal{S}'$ which can hang out of the border as these are vertex disjoint minors. Hence $|\tilde{V}| \leq qt$ and $\tilde{G}$ is a t-boundaried graph. Now we know that $\zeta_{G_R} + c_R = \zeta_G$ and hence $\zeta_{G_R}(\tilde{G}) + c_R = \zeta_G(\tilde{G})$. By definition $\zeta_{G_R}(\tilde{G}) = \mathbf{pack}_{\mathcal{H}}(G \oplus \tilde{G}) = \mathbf{pack}_{\mathcal{H}}(G \oplus G')$, and hence $(G_R \oplus G', k - c_R) \in \Pi_1$. The proof for the other direction that if $(G_R \oplus G', k - c_R) \in \Pi_1$, then $(G \oplus G', k) \in \Pi_1$ is symmetric. This concludes that Π_1 is finite integer index. □

Vertex-$\mathcal{H}$-Covering contains various problems as a special case, for example: (a) Feedback Vertex Set by taking $\mathcal{H} = \{\triangle\}$ where $\triangle$ is a cycle of length at most 3; deleting at most k vertices to obtain a graph of fixed treewidth t; deleting at most k vertices to obtain a graph into a graph class $\mathcal{M}$, where $\mathcal{M}$ is characterized by a finite set of minors. Similarly Vertex-$\mathcal{H}$-Packing contains problem like Cycle Packing as a special case.

Subgraph Covering and Packing: Let $\mathcal{S}$ be a finite set of connected graphs.

> Vertex-$\mathcal{S}$-Covering
> *Input:* A graph $G \in \mathcal{G}_g$ and a non-negative integer k.
> *Question:* Is there a $S \subseteq V$ such that $|S| \leq k$ and $G[V \setminus S]$ does not contain
> any of the graphs in $\mathcal{S}$ as a subgraph?

We similarly define Edge-$\mathcal{S}$-Covering by demanding $S \subseteq E$ in the above definition.

> Vertex-$\mathcal{S}$-Packing
> *Input:* A graph $G \in \mathcal{G}_g$ and a non-negative integer k.
> *Question:* Does there exists k vertex disjoint subgraphs $G_1, G_2, ..., G_k$ in G
> such that, for all i G_i is isomorphic to a graph in $\mathcal{S}$.

We define Edge-$\mathcal{S}$-Packing by demanding that $G_1, \ldots, G_k$ be *edge disjoint* subgraphs of G.

We can not show that Vertex/Edge-$\mathcal{S}$-Covering or Vertex/Edge-$\mathcal{S}$-Packing or there no instances are compact unless we do the following simple preprocessing.
Redundant Vertex and Edge Rule: Given an input $(G = (V, E), k)$ to Vertex/Edge-$\mathcal{S}$-Covering or Vertex/Edge-$\mathcal{S}$-Packing remove all edges and vertices that are not part of any subgraph isomorphic to any graph in $\mathcal{S}$.

We can perform the Redundant Vertex and Edge Rule in $O(|V| \cdot |\mathcal{S}|)$ time by looking at a small ball around an edge e or a vertex v and check whether the ball contains a subgraph isomorphic to a graph in $\mathcal{S}$ and contains the edge e or the vertex v. This algorithm to check a subgraph isomorphic to a given graph containing a particular vertex or edge appears in a paper by Eppstein [29].

Lemma 14. *Let $\mathcal{S}$ be a finite set of connected graphs and Π_1 and Π_2 correspond to* Vertex-$\mathcal{S}$-Packing *and* Edge-$\mathcal{S}$-Packing *respectively. Then the following hold: (a)* Vertex-$\mathcal{S}$-Covering *has finite integer index and is compact; (b)* Vertex-$\mathcal{S}$-Packing *has finite integer index and $\overline{\Pi}_1$ is compact; (c)* Edge-$\mathcal{S}$-Covering *is p-*min*-CMSO problem and is compact; and (d)* Edge-$\mathcal{S}$-Packing *is p-*max*-CMSO problem and $\overline{\Pi}_2$ is compact.*

Proof. Let Π_v and Π_e denote Vertex-$\mathcal{S}$-Covering and Edge-$\mathcal{S}$-Covering respectively. Let $s = \max\{|V^*| \mid G^* = (V^*, E^*) \in \mathcal{S}\}$. Without loss of generality we assume that an input (G, k) to all these problems are *reduced* with respect to Redundant Vertex and Edge Rule.

We first show that these problems or their no instances are compact. Let $(G, k) \in \Pi_v$ then we know that there is a $S \subseteq V$ such that $|S| \leq k$ and $G[V \setminus S]$ does not contain any of the graphs in $\mathcal{S}$ as a subgraph and every vertex and edge is in some subgraph in G which is isomorphic to a subgraph in $\mathcal{S}$. This implies that every vertex in $u \in V \setminus S$ is in at most $r = O(s)$ distance away from a vertex in S and hence $\mathbf{B}_G^r(S) = V$. We can similarly show that Π_e is compact. Next we show that $\overline{\Pi}_1$ is compact. Observe that if $(G, k) \in \Pi_1$ and $(G, k+1) \notin \Pi_1$ then we know that we have a set $\mathcal{Z}$ of k vertex disjoint subgraphs in G where each of them is isomorphic to a subgraph in $\mathcal{S}$. Take S as the union of all the vertices appearing in any of the subgraph in $\mathcal{Z}$. Note that $|S| \leq s \cdot k$. Observe that S hits all the subgraphs isomorphic to a subgraph in $\mathcal{S}$ and hence $\mathbf{B}_G^r(S) = V$, where $r = O(s)$. This implies that $\overline{\Pi}_1$ is compact. We can similarly show that $\overline{\Pi}_2$ is compact.

It is well known that Π_v and Π_e are p-min-CMSO problems while Π_1 and Π_2 are p-max-CMSO problems. The proof that Π_v and Π_1 are finite integer index is similar to the proof given for Vertex-$\mathcal{H}$-Covering and Vertex-$\mathcal{H}$-Packing, where $\mathcal{H}$ is a finite set of connected planar graphs, have finite integer index in Lemma 13. The only thing we need to replace is minors by subgraphs in that proof. □

6.3 Domination and its Variants

In the r-Dominating Set problem, we are given a graph $G = (V, E)$, and a positive integer k, and the objective is to find a subset $S \subseteq V$ such that $B_G^r(S) = V$ and $|S| \leq k$. For $r = 1$, if we demand that $G[S]$ is connected then we get Connected Dominating

SET. A problem is called q-THRESHOLD DOMINATING SET if we demand that $B^1_G(S) = V$ and for all $v \in (V \setminus S)$, $|N(v) \cap S| \geq q$. An independent set C of vertices in a graph $G = (V, E)$ is an efficient dominating set (or perfect code) when each vertex not in C is adjacent to exactly one vertex in C. In EFFICIENT DOMINATING SET problem we are given a graph $G = (V, E)$ and a positive integer k and the objective is to find an efficient dominating set of size at most k.

Lemma 15. *r-DOMINATING SET, CONNECTED DOMINATING SET, q-THRESHOLD DOMINATING SET and EFFICIENT DOMINATING SET are compact and have finite integer index.*

Proof. All these problems are compact by definition. To show that these problems have finite integer index, we make use of Lemma 12. Clearly, they are p-MIN-CMSO problems. We now show that these problems are strongly monotone. We first show it for r-DOMINATING SET. Given a t-boundaried graph $G = (V, E)$, with $\partial(G)$ as its boundary, let $S'' \subseteq V$ be a minimum r-dominating set of G. Take $S = S'' \cup \partial(G)$. Now for any $(G' = (V', E'), S') \in \mathcal{F}_t$ such that $\zeta_G((G', S'))$ is finite we have that $S \cup S'$ is a r-dominating set and $|S| \leq \zeta_G((G', S')) + t$. This proves that r-DOMINATING SET is strongly monotone. Similarly, we can show that q-THRESHOLD DOMINATING SET is strongly monotone by taking $S = S'' \cup \partial(G)$ where $S'' \subseteq V$ is a minimum q-threshold dominating set of G. To show that CONNECTED DOMINATING SET is strongly monotone we take $S = S'' \cup \partial(G)$ where $S'' \subseteq V$ is a union of minimum connected dominating set for each connected components of G.

We now prove that EFFICIENT DOMINATING SET has finite integer index. Let Π be the DOMINATING SET problem and Π' be the EFFICIENT DOMINATING SET problem. The property $P(G)$ that G has an efficient dominating set (of any size) is expressible in CMSO and hence this property is finite state in t-boundaried graphs. As argued in the above paragraph, DOMINATING SET has finite integer index. Furthermore by a theorem of Bange et al. [7], if a graph G has an efficient dominating set, then the size of the minimum efficient dominating set is equal to the size of the minimum dominating set of the graph. Hence for two t-boundaried graphs G_1, G_2 if G_1 and G_2 are in the same equivalence class of P and $G_1 \equiv_{\Pi} G_2$ then $G_1 \equiv_{\Pi'} G_2$. Hence Π' has a finite set of representatives. □

6.4 Problems on Directed Graphs

Our results also apply to problems on directed graphs of bounded genus. In this direction we mention three problems considered in the literature. In DIRECTED DOMINATION [2] we are given a directed graph $D = (V, A)$ and a positive integer k and the objective is to find a subset $S \subseteq V$ of size at most k such that for very vertex $u \in V \setminus S$ there is a vertex $v \in S$ such that $(u, v) \in A$. INDEPENDENT DIRECTED DOMINATION[1] [41] takes as input a directed graph $D = (V, A)$ and a positive integer k and the objective

[1]In literature it is known as "KERNELS". We call it differently here to avoid confusion with problem kernels.

is to find a subset $S \subseteq V$ of size at most k such that S is an independent set and for every vertex $u \in V \setminus S$ there is a vertex $v \in S$ such that $(u, v) \in A$. In the Minimum Leaf Out-branching [40] problem we are given a directed graph $D = (V, A)$ and a positive integer k and the objective is to find a rooted directed spanning tree (with all arcs directed outwards from the vertices) with at least k internal vertices.

Lemma 16. Independent Directed Domination *is a* p-min-CMSO *compact problem,* Directed Domination *is compact and has finite integer index and* Π=Minimum Leaf Out-branching *is a* p-max-CMSO *problem and* $\overline{\Pi}$ *is compact.*

Proof. Independent Directed Domination and Directed Domination can easily be seen to be p-min-CMSO problems and by their definition they are compact. Directed Domination can be shown to have finite integer index as follows: given a t-boundaried graph $G = (V, E)$, with $\partial(G)$ as its boundary, let $S'' \subseteq V$ be a minimum directed dominating set of G. Take $S = S'' \cup \partial(G)$. Now for any $(G' = (V', E'), S') \in \mathcal{F}_t$ such that $\zeta_G((G', S'))$ is finite we have that $S \cup S'$ is a directed dominating set and $|S| \leq \zeta_G((G', S')) + t$. This proves that Directed Dominating Set is strongly monotone.

Let Π be the Minimum Leaf Out-branching problem. Observe that Π is a p-max-CMSO problem. The set of no instances can be seen to be compact by observing the fact that if $(G, k) \in \Pi$ and $(G, k+1) \notin \Pi$ then G has an out-branching with *exactly* k internal vertices with all other vertices being its leaves. This implies that $\overline{\Pi}$ is compact. □

Finally by applying Theorem 2 together with Lemmata 12, 13, 14, 15, and 16 we get the following corollary.

Corollary 2. *For* $g \geq 0$, Feedback Vertex Set, Edge Dominating Set, Vertex Cover, Dominating Set, r-Dominating Set, q-Threshold Dominating Set, Connected Dominating Set, Directed Domination, r-Scattered Set, Connected Vertex Cover, Minimum-Vertex Feedback Edge Set, Minimum Maximal Matching, Efficient Dominating Set, Independent Set, Induced d-Degree Subgraph, Min Leaf Spanning Tree, Induced Matching, Triangle Packing, Cycle Packing, Maximum Full-Degree Spanning Tree, Cycle Domination, Vertex-$\mathcal{H}$-Packing, Vertex-$\mathcal{H}$-Covering, Vertex-$\mathcal{S}$-Covering *and* Vertex-$\mathcal{S}$-Packing *admit a linear kernel on graph of genus at most* g.

Corollary 2 unifies and generalizes results presented in [2, 3, 11, 12, 16, 34, 38, 39, 42, 46, 47]. By applying Theorem 1, Corollary 1, and Lemmata 13, 14, 15, and 16, we get the following corollary for problems which are not finite integer index.

Corollary 3. *For* $g \geq 0$, Independent Dominating Set, Independent Directed Domination, Minimum Leaf Out-branching, Odd Set, Edge-$\mathcal{S}$-Covering *and* Edge-$\mathcal{S}$-Packing *admit a polynomial kernels on graphs of genus at most* g.

7 Open Problems and Further Directions

This paper gives for the first time meta theorems, as called by Grohe and Kreutzer [36, 44], where logical and combinatorial conditions on problems lead to kernels of polynomial or linear sizes. Our results are very general in the sense that they can be applied to a large number of combinatorial problems on graphs on fixed surfaces and generalize a large collection of published results. Still, there are several directions in which our results could possibly be extended. We conclude the paper in this section with some new problems and further research directions opened by our results.

Further extensions: The first natural question for further research is if our logical and combinatorial conditions can be extended to larger classes of problems. The condition that problems should satisfy some kind of compactness or quasi-compactness cannot be omitted. For instance, even though the problem of finding a path of length k is expressible in first order logic, it does not admit a polynomial kernel on planar graphs, unless polynomial hierarchy collapses to the third level, a collapse which is deemed unlikely [10]. Two interesting questions for further research are:

- Can we replace the compactness condition with the weaker notion of *quasi-compactness* in Theorem 1?
- Do all quasi-compact CMSO problems admit a linear kernel on graphs of bounded genus?

It is very natural to ask whether our results can be extended to more general classes of graphs. The most natural candidates for such extensions are graphs of bounded local-treewidth and graphs excluding some fixed graph as a minor.

Practical considerations: Our meta theorems provide simple criteria to decide whether a problem admits a polynomial or linear kernel. Our proofs are constructive and essentially provide *meta-algorithms* that construct kernels for problems in an automated way. Of course, it is expected that for concrete problems, tailor-made kernels will have much smaller constant factors, than what would follow from a direct application of our results. However, our approach might be useful for computer aided design of kernelization algorithms: with a Myhill-Nerode approach, a computer program can output a set of rules that transform each region to a minimum size representative and estimate the obtained kernel size. This seems an interesting and far from trivial algorithm-engineering problem. In general, finding linear kernels *with small constant factors* for concrete problems on planar graphs or graphs with small genus remains a worthy topic of further research.

Some concrete open problems: We conclude with some concrete problems that cannot be resolved by our approach. These include Directed Feedback Vertex Set [18], and Odd Cycle Transversal [50] to name a few. All these problems are expressible in CMSO but none of them are known to be quasi-compact. For each of these problems we do not know if it has a polynomial kernel even on planar graphs, and we leave them open.

Acknowledgments.

We thank Jiong Guo for sending us the complete version of [38].

References

[1] K. R. Abrahamson and M. R. Fellows. Finite automata, bounded treewidth and well-quasiordering. In N. Robertson and P. Seymour, editors, *Proceedings of the AMS Summer Workshop on Graph Minors, Graph Structure Theory, Contemporary Mathematics vol. 147*, pages 539–564. American Mathematical Society, 1993.

[2] J. Alber, B. Dorn, and R. Niedermeier. A general data reduction scheme for domination in graphs. In *SOFSEM '06:*, volume 3831 of *Lecture Notes in Computer Science*, pages 137–147, Berlin, 2006. Springer.

[3] J. Alber, M. R. Fellows, and R. Niedermeier. Polynomial-time data reduction for dominating sets. *J. ACM*, 51:363–384, 2004.

[4] N. Alon and S. Gutner. Kernels for the dominating set problem on graphs with an excluded minor. Technical Report TR08-066, Electronic Colloquium on Computational Complexity (ECCC), 2008.

[5] S. Arnborg, B. Courcelle, A. Proskurowski, and D. Seese. An algebraic theory of graph reduction. *J. ACM*, 40:1134–1164, 1993.

[6] S. Arnborg, J. Lagergren, and D. Seese. Easy problems for tree-decomposable graphs. *J. Algorithms*, 12:308–340, 1991.

[7] D. W. Bange, A. E. Barkauskas, and P. J. Slater. Efficient dominating sets in graphs. In *Proceedings of 3rd Conference on Discrete Mathematics*, pages 189–199. SIAM, 1988.

[8] H. L. Bodlaender. A linear time algorithm for finding tree-decompositions of small treewidth. *SIAM J. Comput.*, 25:1305–1317, 1996.

[9] H. L. Bodlaender. A partial k-arboretum of graphs with bounded treewidth. *Theor. Comp. Sc.*, 209:1–45, 1998.

[10] H. L. Bodlaender, R. G. Downey, M. R. Fellows, and D. Hermelin. On problems without polynomial kernels. In *Proc. 35th ICALP*, volume 5125 of *LNCS*, pages 563–574. Springer, 2008.

[11] H. L. Bodlaender and E. Penninkx. A linear kernel for planar feedback vertex set. In M. Grohe and R. Niedermeier, editors, *Proceedings Third International Workshop on Parameterized and Exact Computation, IWPEC 2008*, pages 160–171. Springer Verlag, Lecture Notes in Computer Science, vol. 5018, 2008.

[12] H. L. Bodlaender, E. Penninkx, and R. B. Tan. A linear kernel for the k-disjoint cycle problem on planar graphs. In *Proceedings of the 19th International Symposium on Algorithms and Computation (ISAAC)*, volume 5369 of *LNCS*, pages 306–317. Springer, 2008.

[13] H. L. Bodlaender and B. van Antwerpen-de Fluiter. Reduction algorithms for graphs of small treewidth. *Information and Computation*, 167:86–119, 2001.

[14] R. B. Borie, R. G. Parker, and C. A. Tovey. Automatic generation of linear-time algorithms from predicate calculus descriptions of problems on recursively constructed graph families. *Algorithmica*, 7:555–581, 1992.

[15] J. Chen, H. Fernau, I. A. Kanj, and G. Xia. Parametric duality and kernelization: Lower bounds and upper bounds on kernel size. In *Proc. 22nd STACS*, volume 3404 of *LNCS*, pages 269–280. Springer, 2005.

[16] J. Chen, H. Fernau, I. A. Kanj, and G. Xia. Parametric duality and kernelization: Lower bounds and upper bounds on kernel size. *SIAM J. Comput.*, 37:1077–1106, 2007.

[17] J. Chen, I. A. Kanj, and W. Jia. Vertex Cover: Further observations and further improvements. *J. Algorithms*, 41(2):280–301, 2001.

[18] J. Chen, Y. Liu, S. Lu, B. O'Sullivan, and I. Razgon. Fixed-parameter algorithm for the directed feedback vertex set problem. *J. ACM*, 55(5): article 21, 2008.

[19] B. Courcelle. The monadic second-order logic of graphs I: Recognizable sets of finite graphs. *Information and Computation*, 85:12–75, 1990.

[20] B. Courcelle. The monadic second-order logic of graphs. III. Tree-decompositions, minors and complexity issues. *RAIRO Inform. Théor. Appl.*, 26(3):257–286, 1992.

[21] B. Courcelle. The expression of graph properties and graph transformations in monadic second-order logic. *Handbook of Graph Grammars*, pages 313–400, 1997.

[22] B. Courcelle and M. Mosbah. Monadic second-order evaluations on tree-decomposable graphs. *Theor. Comp. Sc.*, 109:49–82, 1993.

[23] A. Dawar, M. Grohe, and S. Kreutzer. Locally excluding a minor. In *Proceedings of the 22nd IEEE Symposium on Logic in Computer Science (LICS 2007)*, pages 270–279. IEEE Computer Society, 2007.

[24] B. de Fluiter. *Algorithms for Graphs of Small Treewidth.* PhD thesis, Utrecht University, 1997.

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

[26] E. D. Demaine and M. Hajiaghayi. Graphs excluding a fixed minor have grids as large as treewidth, with combinatorial and algorithmic applications through bidimensionality. In *Proceedings of the 16th Annual ACM-SIAM Symposium on Discrete Algorithms, SODA2005*, pages 682–689, 2005.

[27] E. D. Demaine, M. Hajiaghayi, and D. M. Thilikos. The bidimensional theory of bounded-genus graphs. *SIAM J. Discrete Math.*, 20(2):357–371 (electronic), 2006.

[28] R. G. Downey and M. R. Fellows. *Parameterized Complexity*. Springer, 1998.

[29] D. Eppstein. Subgraph isomorphism in planar graphs and related problems. *J. Graph Algorithms and Applications*, 3:1–27, 1999.

[30] M. R. Fellows and M. A. Langston. An analogue of the Myhill-Nerode theorem and its use in computing finite-basis characterizations. In *Proceedings of the 30th Annual Symposium on Foundations of Computer Science*, pages 520–525, 1989.

[31] M. R. Fellows and M. A. Langston. On well-partial-order theory and its application to combinatorial problems of VLSI design. *SIAM J. Disc. Math.*, 5:117–126, 1992.

[32] M. R. Fellows and M. A. Langston. On search, decision and the efficiency of polynomial-time algorithms. *J. Comp. Syst. Sc.*, 49:769–779, 1994.

[33] F. V. Fomin, S. Saurabh, and D. M. Thilikos. Improving the gap of Erdös-Pósa property for minor-closed graph classes. In *Proceedings of 7th Cologne-Twente Workshop on Graphs and Combinatorial Optimization*, pages 2–6, 2008.

[34] F. V. Fomin and D. M. Thilikos. Fast parameterized algorithms for graphs on surfaces: Linear kernel and exponential speed-up. In *Proc. 31st ICALP*, volume 3142 of *LNCS*, pages 581–592. Springer, 2004.

[35] M. Frick and M. Grohe. Deciding first-order properties of locally tree-decomposable structures. *J. ACM*, 48(6):1184–1206, 2001.

[36] M. Grohe. Logic, graphs, and algorithms. In J. Flum, E. Grädel, and T. Wilke, editors, *Logic and Automata???–History and Perspectives*. Amsterdam University Press, 2007.

[37] J. Guo and R. Niedermeier. Invitation to data reduction and problem kernelization. *SIGACT News*, 38(1):31–45, 2007.

[38] J. Guo and R. Niedermeier. Linear problem kernels for NP-hard problems on planar graphs. In *Proc. 34th ICALP*, volume 4596 of *LNCS*, pages 375–386. Springer, 2007.

[39] J. Guo, R. Niedermeier, and S. Wernicke. Fixed-parameter tractability results for full-degree spanning tree and its dual. In *Proceedings of the Second International Workshop on Parameterized and Exact Computation (IWPEC)*, volume 4169 of *LNCS*, pages 203–214. Springer, 2006.

[40] G. Gutin, E. J. Kim, and I. Razgon. Minimum leaf out-branching problems. *arxiv.org/abs/0801.1979*, 2008.

[41] G. Gutin, T. Kloks, C. M. Lee, and A. Yeo. Kernels in planar digraphs. *J. Comp. Syst. Sc.*, 71:174–184, 2005.

[42] I. A. Kanj, M. J. Pelsmajer, G. Xia, and M. Schaefer. On the induced matching problem. In *Proceedings of the 25th Annual Symposium on Theoretical Aspects of Computer Science (STACS 2008)*, volume 08001, pages 397–408. Internationales Begegnungs- und Forschungszentrum fuer Informatik (IBFI), Schloss Dagstuhl, Germany, Berlin, 2008.

[43] A. Koutsonas and D. M. Thilikos. Planar feedback vertex set and face cover: Combinatorial bounds and subexponential algorithms. In *34th Workshop on Graph-Theoretic Concepts in Computer Science (WG '08)*, volume 5344 of *Lecture Notes in Computer Science*, pages 264–274. Springer Verlag, 2008.

[44] S. Kreutzer. Algorithmic meta-theorems. *CoRR*, abs/0902.3616, 2009.

[45] D. Lapoire. Recognizability equals definability, for every set of graphs of bounded tree-width. In *Proceedings 15th Annual Symposium on Theoretical Aspects of Computer Science*, pages 618–628. Springer Verlag, Lecture Notes in Computer Science, vol. 1373, 1998.

[46] D. Lokshtanov, M. Mnich, and Saurabh. Linear kernel for planar connected dominating set. In *Proceedings of Theory and Applications of Models of Computation, (TAMC 2009)*, Lecture Notes in Comput. Sci. Springer, 2009.

[47] H. Moser and S. Sikdar. The parameterized complexity of the induced matching problem in planar graphs. In *Proceedings First Annual International WorkshopFrontiers in Algorithmics (FAW)*, volume 4613 of *LNCS*, pages 325–336. Springer, 2007.

[48] G. Philip, V. Raman, and S. Sikdar. Polynomial kernels for dominating set in $K_{i,j}$ -free and d-degenerate graphs. Manuscript, 2009.

[49] W. V. Quine. The problem of simplifying truth functions. *Amer. Math. Monthly*, 59:521–531, 1952.

[50] B. Reed, K. Smith, and A. Vetta. Finding odd cycle transversals. *Oper. Res. Lett.*, 32(4):299–301, 2004.

[51] N. Robertson, P. D. Seymour, and R. Thomas. Quickly excluding a planar graph. *J. Comb. Theory Series B*, 62:323–348, 1994.

[52] S. Thomassé. A quadratic kernel for feedback vertex set. In *Proc. 20th SODA*, pages 115–119. ACM/SIAM, 2009.

A Problem Compendium

We refer to problem compendium given in [28] or the compendium of parameterized problems provided at http://bravo.ce.uniroma2.it/home/cesati/research/compendium/ for the definitions of problems given below.

A.1 Problems that have Finite Integer Index and Quasi-Compactness Property – *Linear Kernels*

Dominating Set, r-Dominating Set, q-Threshold Dominating Set, Efficient Dominating Set, Vertex Cover, Connected r-Dominating Set, Connected Vertex Cover, Minimum-Vertex Feedback Edge Set, Vertex-$\mathcal{H}$-Covering, Minimum Maximal Matching, Connected Dominating Set, Vertex-$\mathcal{S}$-Covering, Clique-Transversal, Almost-Outerplanar, Feedback Vertex Set, Cycle Domination, Edge Dominating Set.

A.2 Problems that have Finite Integer Index and No Instances have Quasi-Compactness Property – *Linear Kernels*

Independent Set, Induced d-Degree Subgraph, r-Scattered Set, Min Leaf Spanning Tree, Induced Matching, Triangle Packing, Cycle Packing, Maximum Full-Degree Spanning Tree, Vertex-$\mathcal{H}$-Packing, Vertex-$\mathcal{S}$-Packing.

A.3 Problems that are not covered by A.1 and A.2 and are p-min/eq/max-CMSO problems having Compactness Property – *Polynomial Kernels*

Independent Dominating Set, Independent Directed Domination, Minimum Leaf Out-branching, Edge-$\mathcal{S}$-Covering, Edge-$\mathcal{S}$-Packing, Odd Set.

A.4 Problems that have Finite Integer Index but neither the problem nor its no instances satisfy Quasi-Compactness Property

Minimum Partition Into Cliques, Hamiltonian Path Completion.

A.5 Problems that do not have Finite Integer Index

Longest Path, Longest Cycle, Maximum Cut, Minimum Covering by Cliques, Independent Dominating Set, Minimum Leaf Out-branching, Odd Set.

The following lemma is shown in [24].

Lemma 17. [24] Longest Path, Longest Cycle, Maximum Cut *and* Minimum Covering by Cliques *are not of finite integer index.*