Kernelization Bounds for Constrained Coloring

Ishay Haviv
The Academic College of Tel Aviv-Yaffo, Israel

Abstract

We study the kernel complexity of constraint satisfaction problems over a finite domain, parameterized by the number of variables, whose constraint language consists of two relations: the non-equality relation and an additional permutation-invariant relation R. We establish a conditional lower bound on the kernel size in terms of the largest arity of an OR relation definable from R. Building on this, we investigate the kernel complexity of uniformly rainbow free coloring problems. In these problems, for fixed positive integers d, ℓ, and $q \geq d$, we are given a graph G on n vertices and a collection $\mathcal{F}$ of ℓ-tuples of d-subsets of its vertex set, and the goal is to decide whether there exists a proper coloring of G with q colors such that no ℓ-tuple in $\mathcal{F}$ is uniformly rainbow, that is, no tuple has all its sets colored with the same d distinct colors. We determine, for all admissible values of d, ℓ, and q, the infimum over all values η for which the problem admits a kernel of size $O(n^{\eta})$, under the assumption $\mathsf{NP} \not\subseteq \mathsf{coNP/poly}$. As applications, we obtain nearly tight bounds on the kernel complexity of various coloring problems under diverse settings and parameterizations. This includes graph coloring problems parameterized by the vertex-deletion distance to a disjoint union of cliques, resolving a question of Schalken (2020), as well as uniform hypergraph coloring problems parameterized by the number of vertices, extending results of Jansen and Pieterse (2019) and Beukers (2021).

2012 ACM Subject Classification Theory of computation → Parameterized complexity and exact algorithms; Mathematics of computing → Graph coloring; Mathematics of computing → Combinatorial algorithms

Keywords and phrases parameterized complexity, kernelization, constraint satisfaction problems, coloring problems

Digital Object Identifier 10.4230/LIPIcs.MFCS.2026.53

Related Version *Full Version*: http://arxiv.org/abs/2604.21531

Acknowledgements We thank the anonymous reviewers for useful feedback.

1 Introduction

Constraint satisfaction problems form a central concept in theoretical computer science, interacting with foundational areas in the field, such as approximation algorithms, hardness of approximation, fine-grained complexity, parameterized complexity, and kernelization. In the *constraint satisfaction problem* associated with a relation $R \subseteq D^r$ of arity r over a finite domain D, the input consists of a set of variables V and a collection $\mathcal{F} \subseteq V^r$ of r-tuples of variables, called *constraints*, and the goal is to decide whether there exists an assignment $c : V \to D$ that satisfies all constraints in $\mathcal{F}$, meaning that for every r-tuple $(x_1, \ldots, x_r) \in \mathcal{F}$, it holds that $(c(x_1), \ldots, c(x_r)) \in R$. More broadly, a constraint satisfaction problem can be defined with respect to a set of relations over D, called the *constraint language*, where each constraint is associated with one of them. This framework naturally encompasses central decision problems studied in theoretical computer science, such as bounded-width satisfiability, graph and hypergraph coloring, and many others. A remarkable result in this context is the dichotomy theorem, which asserts that every constraint satisfaction problem is either solvable in polynomial time or NP-hard. This was proved for the Boolean domain in 1978 by Schaefer [23], and more recently, for all finite domains, by Bulatov [6] and Zhuk [26] independently.

51st International Symposium on Mathematical Foundations of Computer Science (MFCS 2026).
Editors: Michal Koucký and Daniela Petrişan; Article No. 53; pp. 53:1–53:14
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

The present paper studies constraint satisfaction problems from the perspective of *kernelization* (see, e.g., [12]). This theme, situated in the area of parameterized complexity, seeks to determine to what extent instances of a computationally hard problem can be compressed with respect to a chosen parameter of the instances. Specifically, a kernelization algorithm for a parameterized problem is given as input a pair (x, k), where x is the main input and k is the associated parameter, and produces in polynomial time an equivalent instance, whose size is bounded by a function that depends solely on k. A major challenge in the field is to estimate, under plausible complexity assumptions, the smallest possible growth rate of the kernel size of a given parameterized problem. When parameterized by the number of variables n, a constraint satisfaction problem associated with a constraint language of relations of arity at most r admits a trivial kernelization algorithm that produces instances with $O(n^r)$ constraints, simply by removing duplicate constraints. Determining the smallest possible kernel size for such problems has attracted significant attention in recent years, e.g., [11, 18, 20, 10, 19], and is the driving force behind this work.

The kernel complexity of constraint satisfaction problems over the Boolean domain $D = \{0, 1\}$ has been studied extensively in the literature. A remarkable result, proved by Dell and van Melkebeek [11], states that for every integer $k \geq 3$ and for any real $\varepsilon > 0$, the satisfiability problem on CNF formulas with clauses of size k, when parameterized by the number of variables n, does not admit a kernel of size $O(n^{k-\varepsilon})$ unless $\mathsf{NP} \subseteq \mathsf{coNP/poly}$. Since this containment is known to imply the collapse of the polynomial-time hierarchy [25], it follows that the trivial kernel of size $O(n^k)$ is presumably near-optimal. Yet, for other variants of the satisfiability problem, smaller kernels are known to exist. For example, Jansen and Pieterse [16, 18] showed that the Not All Equal satisfiability problem on CNF formulas with clauses of size $k \geq 3$, when parameterized by the number of variables n, admits a kernel with $O(n^{k-1})$ clauses, which can be encoded in $O(n^{k-1} \cdot \log n)$ bits. This was achieved via an algebraic approach they introduced in [16] based on low-degree polynomials expressing the satisfiability constraint. They further provided a nearly matching lower bound for the problem, ruling out any kernel of size $O(n^{k-1-\varepsilon})$ for a constant $\varepsilon > 0$ unless $\mathsf{NP} \subseteq \mathsf{coNP/poly}$ (see also [10]). More generally, Chen, Jansen, and Pieterse [10] provided a criterion for when an intractable Boolean constraint satisfaction problem admits a non-trivial kernel. Letting r denote the maximum arity of the relations in the constraint language at hand, they showed that if an OR relation of arity r is definable from one of them, in the sense that it has exactly one falsifying assignment, then no kernel of size $O(n^{r-\varepsilon})$ exists for $\varepsilon > 0$ unless $\mathsf{NP} \subseteq \mathsf{coNP/poly}$, whereas otherwise, the number of constraints can be efficiently reduced to $O(n^{r-1})$. This result was used in [10] to obtain an exact conditional characterization of the kernel complexity of all Boolean constraint satisfaction problems with relations of arity at most three. The kernel complexity of constraint satisfaction problems over finite domains of size larger than 2 is currently much less understood. To address these cases, Lagerkvist and Wahlström [20] generalized the approach of [10] using tools from universal algebra. Among their results lies a full conditional characterization of the finite-domain constraint satisfaction problems that admit a kernel whose number of constraints grows linearly with the number of variables.

The study of the kernel complexity of non-Boolean constraint satisfaction problems is strongly motivated by graph coloring problems, which provide a rich source of kernelization challenges. For an integer q, the q-COLORING problem asks whether a given graph admits a *proper* coloring with q colors, i.e., an assignment of colors to the vertices such that adjacent vertices receive distinct colors. This problem, which is well known to be NP-hard for every $q \geq 3$, corresponds to the constraint satisfaction problem associated with the *non-equality* relation of arity 2 over the domain $[q] = \{1, \ldots, q\}$. When parameterized by the number of vertices n, the problem admits a trivial kernel of size $O(n^2)$, and a result of Jansen and Pieterse [16, 17] asserts that, for every $q \geq 3$ and for any $\varepsilon > 0$, it admits no kernel of size $O(n^{2-\varepsilon})$ unless $\mathsf{NP} \subseteq \mathsf{coNP/poly}$. This result has inspired natural extensions and generalizations. For example, such a lower bound was obtained for certain (list) H-Coloring problems, which ask whether an input graph admits a homomorphism to a fixed target graph H, possibly subject to lists of allowed vertices in H for each vertex [9, 1]. Another extension, due to Beukers [3], concerns the q-Coloring problem on 3-uniform hypergraphs with n vertices, corresponding to the relation of arity 3 that consists of all triples in $[q]^3$ except those with all equal entries. Under the assumption $\mathsf{NP} \nsubseteq \mathsf{coNP/poly}$, he proved that for every $q \geq 3$ and for any $\varepsilon > 0$, the problem admits no kernel of size $O(n^{3-\varepsilon})$, essentially matching the trivial upper bound.

It is noteworthy that recent research on the kernel complexity of constraint satisfaction problems has revealed intimate connections to additional structural quantities of interest. A prominent example is the *non-redundancy* of such problems, studied, e.g., in [2, 8, 4, 5], which measures the maximum number of constraints in a non-redundant instance on n variables, that is, an instance in which no constraint is implied by the others. At first glance, this quantity may appear unrelated to the optimal kernel size, since a kernelization algorithm is not restricted to removing redundant constraints and may modify the instance in arbitrary ways, even introducing new constraints. Furthermore, the kernelization task is algorithmic, so tractable constraint satisfaction problems admit a kernel of constant size, whereas their non-redundancy can grow polynomially. Also, while combining several tractable relations in the constraint language of a constraint satisfaction problem can render the problem intractable, and thus result in non-constant kernel complexity, the non-redundancy of the problem is dominated by the maximal non-redundancy of the individual relations. Nevertheless, recent works indicate that in many cases, the two quantities are closely tied. Known kernelization algorithms typically operate by eliminating redundant constraints, and upper bounds on non-redundancy can often be transformed into efficient kernelization algorithms. Clarifying the precise relationship between these two measures of constraint satisfaction problems is a compelling question in the area.

1.1 Our Contribution

This paper addresses the kernel complexity of constraint satisfaction problems over a finite domain of arbitrary size q, whose constraint language consists of two relations: the non-equality relation of arity 2 over $[q]$, corresponding to the q-Coloring problem, and an additional relation $R \subseteq [q]^r$ of arity r. Our first contribution establishes a general conditional lower bound on the kernel complexity of such problems. Building on this result, we investigate a family of problems, which we refer to as uniformly rainbow free coloring, and obtain nearly matching upper and lower bounds on their kernel complexity. We then use these bounds to determine the kernel complexity of various coloring problems under diverse settings and parameterizations. This includes graph coloring problems parameterized by the vertex-deletion distance to a disjoint union of cliques, as well as uniform hypergraph coloring problems parameterized by the number of vertices. A detailed description of our contribution is given below.

1.1.1 Constrained Coloring

The constrained coloring problem associated with a relation R is defined as follows.

▶ **Definition 1.** *For positive integers q and r, let $R \subseteq [q]^r$ be a relation of arity r over $[q]$. The input of the R-Constrained Coloring problem consists of a graph $G = (V, E)$ and a collection $\mathcal{F} \subseteq V^r$ of r-tuples of vertices. The goal is to decide whether there exists a proper coloring $c : V \to [q]$ of G, such that for every r-tuple $(x_1, \ldots, x_r) \in \mathcal{F}$, it holds that $(c(x_1), \ldots, c(x_r)) \in R$. The number of constraints in the instance $(G, \mathcal{F})$ is $|E| + |\mathcal{F}|$.*

Our first result establishes a lower bound on the kernel complexity of R-Constrained Coloring problems for permutation-invariant relations R with respect to the number of vertices parameterization. A relation $R \subseteq [q]^r$ is called *permutation-invariant* if every permutation of $[q]$ applied to the entries of a tuple does not affect its membership in R (see Definition 8). For an integer k, we say that *an* OR *relation of arity k is definable from* $R \subseteq [q]^r$ if there exist r subsets $D_1, \ldots, D_r$ of $[q]$, with k of them of size 2 and the others of size 1, so that among the 2^k tuples in the product $D_1 \times \cdots \times D_r$, exactly one does not belong to R (see Definition 9). We prove that for every permutation-invariant relation R that defines an OR relation of arity at least 3, the R-Constrained Coloring problem parameterized by the number of vertices is unlikely to admit a kernel whose size is bounded by a polynomial of degree smaller than that arity. In fact, here and throughout, the lower bounds apply not only to kernels but also to *compressions*, i.e., algorithms that reduce an instance of the problem at hand to an instance of *any* problem, not necessarily itself (see Section 2.3).

▶ **Theorem 2.** *For positive integers q, r, and $k \geq 3$, let $R \subseteq [q]^r$ be a permutation-invariant relation of arity r over $[q]$ from which an* OR *relation of arity k is definable. Then, for any real $\varepsilon > 0$, the R-Constrained Coloring problem parameterized by the number of vertices n does not admit a compression of size $O(n^{k-\varepsilon})$ unless $\mathsf{NP} \subseteq \mathsf{coNP/poly}$.*

Several remarks are in order here. First, for every relation $R \subseteq [q]^r$ of arity $r \geq 3$, the R-Constrained Coloring problem parameterized by the number of vertices n admits a trivial kernel of size $O(n^r)$, obtained by eliminating repeated constraints. Theorem 2 shows that if an OR relation of arity r is definable from R, then this trivial kernel is likely near-optimal. It turns out that this condition is essential for such a lower bound, as a result of Carbonnel [8] implies that if the largest arity of an OR relation definable from R is smaller than r, then the problem admits a smaller kernel with $O(n^{r-\varepsilon})$ constraints for $\varepsilon = 2^{1-r}$ (see Proposition 16). Second, if one is interested in a lower bound on the compression size of a *list-coloring* variant of the R-Constrained Coloring problem, where the input specifies a set of available colors for each vertex, then the assumption of Theorem 2 that the relation R is permutation-invariant can be omitted. In the sequel, we first prove a lower bound for this list-coloring variant for an arbitrary relation R and then apply it to derive Theorem 2 (see Theorem 13 and Lemma 15). Finally, we stress that the lower bound on the compression size of the R-Constrained Coloring problem, given in Theorem 2, is governed by the arity of the OR relations definable from the relation R, even when the constraint satisfaction problem involving only R is solvable in polynomial time. This outcome arises from the interplay between the R-constraints and the non-equality constraints imposed by the proper coloring requirement. Moreover, the theorem sometimes yields a non-trivial lower bound even when the two individual components of the R-Constrained Coloring problem are polynomial-time tractable, as may occur when $q = 2$.

Before turning to our main applications of Theorem 2, let us demonstrate its utility with a quick example concerning the q-Coloring problem on ℓ-uniform hypergraphs for fixed integers $\ell \geq 2$ and $q \geq 3$. In this problem, the input is a hypergraph on n vertices with edges of size exactly ℓ, and the goal is to decide whether it admits a coloring with

q colors so that no edge is monochromatic. As alluded to earlier, for $\ell \in \{2,3\}$, for every $q \geq 3$, and for any $\varepsilon > 0$, the problem does not admit a compression of size $O(n^{\ell-\varepsilon})$ unless $\mathsf{NP} \subseteq \mathsf{coNP/poly}$ [17, 3], nearly matching the trivial upper bound of $O(n^{\ell})$ on the kernel complexity. Theorem 2 enables us to extend the lower bound to all integers $\ell \geq 4$. To do so, consider the relation R that consists of all ℓ-tuples in $[q]^{\ell}$ whose entries are not all equal, and notice that the corresponding R-CONSTRAINED COLORING problem coincides with the q-COLORING problem on hypergraphs with edges of size 2 and ℓ. Observe that an OR relation of arity ℓ is definable from R, as witnessed by restricting it to the product $\{1,2\} \times \{1,3\}^{\ell-1}$, where all 2^{ℓ} tuples belong to R except the all-1 tuple. Since R is permutation-invariant, Theorem 2 implies that the R-CONSTRAINED COLORING problem parameterized by the number of vertices n does not admit a compression of size $O(n^{\ell-\varepsilon})$ unless $\mathsf{NP} \subseteq \mathsf{coNP/poly}$. By a simple reduction to the q-COLORING problem on ℓ-uniform hypergraphs, without edges of size 2, this yields a near-optimal lower bound on its compressibility, as summarized by the following statement.

▶ **Theorem 3.** *For all integers $\ell \geq 2$ and $q \geq 3$, and for any real $\varepsilon > 0$, the q-COLORING problem on ℓ-uniform hypergraphs parameterized by the number of vertices n does not admit a compression of size $O(n^{\ell-\varepsilon})$ unless $\mathsf{NP} \subseteq \mathsf{coNP/poly}$.*

It should be mentioned that the kernel complexity of the q-COLORING problem on ℓ-uniform hypergraphs parameterized by the number of vertices n behaves differently when the number of colors q is 2. Indeed, it was shown in [16] that the NP-hard cases of the 2-COLORING problem on ℓ-uniform hypergraphs, with $\ell \geq 3$, admit a kernel with $O(n^{\ell-1})$ edges, which can be encoded in $O(n^{\ell-1} \cdot \log n)$ bits. We note that Theorem 3 can also be derived from our analysis of the coloring problems described next.

1.1.2 Uniformly Rainbow Free Coloring

The primary family of R-CONSTRAINED COLORING problems considered in this paper is that of uniformly rainbow free colorings. For a graph $G = (V,E)$, a positive integer q, and a coloring $c : V \to [q]$, a set $A \subseteq V$ is called *rainbow* with respect to c if c assigns to the vertices of A pairwise distinct colors. For positive integers d and ℓ, an ℓ-tuple $(F_1,\dots,F_{\ell})$ of d-subsets of V is called *uniformly rainbow* with respect to the coloring c if all of its sets are rainbow and share the same set of colors under c. We define a family of computational problems as follows.

▶ **Definition 4.** *Let d, ℓ, and $q \geq d$ be positive integers. The input of the (d,ℓ,q)-UNIFORMLY RAINBOW FREE COLORING problem consists of a graph $G = (V,E)$ and a collection $\mathcal{F}$ of ℓ-tuples of d-subsets of V. The goal is to decide whether there exists a proper coloring $c : V \to [q]$ of G, such that no tuple in $\mathcal{F}$ is uniformly rainbow with respect to c.*

Note that the (d,ℓ,q)-UNIFORMLY RAINBOW FREE COLORING problem is defined for $q \geq d$, as otherwise no set of size d can be rainbow, and all constraints in $\mathcal{F}$ are trivially satisfied. Note further that for any values of d, ℓ, and q, the problem can be formulated as an R-CONSTRAINED COLORING problem for a suitably defined permutation-invariant relation R.

We provide nearly matching upper and lower bounds on the kernel complexity of the (d,ℓ,q)-UNIFORMLY RAINBOW FREE COLORING problem, parameterized by the number of vertices, for all admissible values of d, ℓ, and q, under the complexity assumption $\mathsf{NP} \not\subseteq \mathsf{coNP/poly}$. We first introduce the following notation.

▶ **Definition 5.** *For positive integers d, ℓ, and $q \geq d$, set*

$$\eta(d,\ell,q) = \begin{cases} d\ell, & \text{if } \ell \geq 2 \text{ and } q \geq d+2, \\ d\ell - 1, & \text{if } \ell \geq 2,\ q = d+1,\ \text{and } (d,\ell) \neq (1,2), \\ (d-1)\ell, & \text{if } (\ell \geq 2 \text{ and } q = d \geq 2) \text{ or } (\ell = 1 \text{ and } d \geq 3), \\ 2, & \text{if } \ell = 1,\ d \leq 2,\ \text{and } q \geq 3, \\ 0, & \text{if } (q = 2 \text{ and } d\ell \leq 2) \text{ or } q = 1. \end{cases}$$

Note that the five cases in Definition 5 are disjoint and cover all triples (d, ℓ, q) with $q \geq d$. With this definition in place, we proceed to the formal statement.

▶ **Theorem 6.** *For positive integers d, ℓ, and $q \geq d$, set $\eta = \eta(d, \ell, q)$ as in Definition 5. The (d, ℓ, q)-Uniformly Rainbow Free Coloring problem parameterized by the number of vertices n admits a kernel with $O(n^{\eta})$ constraints, while for any real $\varepsilon > 0$, it does not admit a compression of size $O(n^{\eta-\varepsilon})$ unless $\mathsf{NP} \subseteq \mathsf{coNP/poly}$.*

Let us unpack Theorem 6 in light of the cases arising from Definition 5. When $\ell \geq 2$ and $q \geq d + 2$, the theorem shows that the (d, ℓ, q)-Uniformly Rainbow Free Coloring problem parameterized by the number of vertices n is unlikely to admit a kernel much smaller than the trivial one, whose number of constraints is bounded by $O(n^{d\ell})$. In contrast, when q is either d or $d + 1$, a polynomially smaller kernel is achievable, apart from some exceptional cases, including those with $\eta(d, \ell, q) = 0$, for which the problem is solvable in polynomial time. Note that, for $d \geq 2$, any coloring that assigns a fixed color to all vertices trivially satisfies all uniformly rainbow free constraints. Nonetheless, when combined with the proper coloring requirement of the graph, these constraints essentially govern the kernel complexity of the problem. It is also worth noting that the bounds on the number of constraints in the kernelized instances in Theorem 6 reflect their encoding size in bits, up to a logarithmic factor.

The lower bounds in Theorem 6 are derived from Theorem 2 through an analysis of the OR relations definable from the relations that characterize the (d, ℓ, q)-Uniformly Rainbow Free Coloring problems. To establish the non-trivial upper bounds stated in Theorem 6, we leverage the linear-algebraic approach of [16]. To adapt this machinery to our setting, we construct low-degree polynomials that capture the uniformly rainbow property, extending ideas from [17, 14, 13]. We also provide an extension of the upper bounds from Theorem 6, needed for some of our applications, where the values of d and ℓ are not unique but are taken from a prescribed set of possibilities. For this setting, we show that the degree of the polynomial kernel complexity is dominated by the maximum value of $\eta(d, \ell, q)$ over all allowed pairs (d, ℓ).

1.1.3 Coloring Graphs Close to Disjoint Union of Cliques

Our results on the uniformly rainbow free coloring problems enable us to settle the kernel complexity of a family of graph coloring problems parameterized by the vertex-deletion distance to a disjoint union of cliques. Following a terminology of Cai [7], for a positive integer q and for a graph family $\mathcal{G}$, let q-Coloring on $\mathcal{G} + k$v graphs be the problem parameterized by k, defined as follows. The input comprises a graph $G = (V, E)$ and a set $X \subseteq V$ of k vertices, whose removal from G gives a graph that belongs to $\mathcal{G}$, and the goal is to decide whether G admits a proper coloring with q colors. We investigate the kernel complexity of the problem in its NP-hard cases, with $q \geq 3$, for graph families $\mathcal{G}$ of disjoint unions of cliques.

Specifically, for a positive integer t, let t-DisjCliques denote the family of all graphs that form a disjoint union of cliques, each of size at most t, and let DisjCliques denote the family of all graphs that form a disjoint union of cliques of any size.

The kernel complexity of the q-Coloring problem on t-DisjCliques $+ k$v graphs was first studied in 2013 by Jansen and Kratsch [15] for the special case of $t = 1$, corresponding to the well-studied vertex cover number parameterization. They showed that in this case, for every integer $q \geq 3$, the problem admits a kernel with $O(k^q)$ vertices and bit-size, while for any $\varepsilon > 0$, it admits no kernel of size $O(k^{q-1-\varepsilon})$ unless $\mathsf{NP} \subseteq \mathsf{coNP/poly}$. In 2017, the upper bound was tightened by Jansen and Pieterse [17], using their linear-algebraic approach [16], to a kernel with $O(k^{q-1})$ vertices and bit-size $O(k^{q-1} \cdot \log k)$, matching the lower bound of [15] up to a $k^{o(1)}$ multiplicative term. In 2020, Schalken [24] studied the kernel complexity of the q-Coloring problem with respect to smaller parameters, with particular attention to the q-Coloring problem on t-DisjCliques $+ k$v graphs with $t = 2$. He proved that the problem admits a kernel with $O(k^{2q-2})$ vertices and bit-size, while for any $\varepsilon > 0$, it admits no kernel of size $O(k^{2q-3-\varepsilon})$ unless $\mathsf{NP} \subseteq \mathsf{coNP/poly}$. For $q \in \{3, 4\}$, he further provided a near-optimal kernel with $O(k^{2q-3})$ vertices and bit-size $O(k^{2q-3} \cdot \log k)$, which was then extended in [13] to all integers $q \geq 3$. Schalken [24] also studied the q-Coloring problem on DisjCliques $+ k$v graphs, which essentially corresponds to the case $t = q$. He showed that the problem admits a kernel with $O(k^r)$ vertices for $r = \max_{\ell \in [q]} (q - \ell + 1)\ell = \lfloor (q+1)^2/4 \rfloor$, and asked whether a smaller kernel exists and whether matching conditional lower bounds could be established.

Based on our bounds on the kernel complexity of uniformly rainbow free coloring problems, we determine the kernel complexity of the q-Coloring problem on t-DisjCliques $+ k$v graphs for all admissible values of q and t, up to a $k^{o(1)}$ multiplicative factor and under the assumption $\mathsf{NP} \not\subseteq \mathsf{coNP/poly}$.

▶ **Theorem 7.** *For positive integers $q \geq 3$ and $t \leq q$, set $r = r(q,t) = \max_{\ell \in [t]} \eta(q - \ell + 1, \ell, q)$, with η as in Definition 5. The q-Coloring problem on t-DisjCliques $+ k$v graphs admits a kernel with $O(k^r)$ vertices, while for any real $\varepsilon > 0$, it does not admit a compression of size $O(k^{r-\varepsilon})$ unless* $\mathsf{NP} \subseteq \mathsf{coNP/poly}$.

As a consequence of Theorem 7, we obtain nearly tight bounds on the kernel complexity of the q-Coloring problem on DisjCliques $+ k$v graphs. Our findings resolve the open question raised in [24], and show that the kernel size achieved there is essentially optimal for $q \geq 4$, but can be improved by almost a linear factor for $q = 3$.

1.2 Outline

The remainder of the paper is structured as follows. In Section 2, we collect definitions and results that will be used throughout the paper. In Section 3, we establish our general lower bound on the compression size of R-Constrained Coloring problems and prove Theorem 2. Due to space limitations, we defer the proofs of the remaining results to the full version of the paper. We end in Section 4 with some concluding remarks.

2 Preliminaries

For a positive integer n, let $[n]$ denote the set of positive integers up to n, and for a set A, let $\mathcal{P}(A)$ denote its power set.

2.1 Graphs

All graphs considered in this paper are finite and simple, having no loops or parallel edges. For a positive integer q, a *coloring* of a graph $G = (V, E)$ with q colors is a function from its vertex set to a set of size q. A coloring $c : V \to [q]$ of G is called *proper* if $c(u) \neq c(v)$ for every edge $\{u, v\} \in E$.

2.2 Relations

We present here two notions that will be used in our study of constraint satisfaction problems. We begin with the notion of a permutation-invariant relation.

▶ **Definition 8.** *For positive integers q and r, let $R \subseteq [q]^r$ be a relation of arity r over $[q]$. The relation R is called* permutation-invariant *if for every permutation $\pi : [q] \to [q]$ and for every r-tuple $x = (x_1, \dots, x_r) \in [q]^r$, it holds that $x \in R$ if and only if $(\pi(x_1), \dots, \pi(x_r)) \in R$.*

Note that permutation-invariance concerns permutations of the domain elements rather than of the tuple positions. Consequently, permuting the entries of a tuple may not preserve membership in a permutation-invariant relation.

We next formalize when an OR relation is definable from another relation. Related and more general notions can be found, e.g., in [8, 21].

▶ **Definition 9.** *Let q, r, and k be positive integers, and let $R \subseteq [q]^r$ be a relation of arity r over $[q]$. We say that* an OR relation of arity k is definable from R *if there exist sets $D_1, \dots, D_r \subseteq [q]$, exactly k of which are of size 2 and the others of size 1, such that all 2^k tuples in the product $D_1 \times \dots \times D_r$ except one belong to R, equivalently, $|(D_1 \times \dots \times D_r) \cap R| = 2^k - 1$.*

2.3 Parameterized Complexity

We now gather fundamental definitions on kernelization from the field of parameterized complexity. For an in-depth introduction to the area, the reader is referred to [12]. A *parameterized problem* is a set $Q \subseteq \Sigma^* \times \mathbb{N}$ for some finite alphabet Σ. A *compression* (also known as a generalized kernel) for a parameterized problem $Q \subseteq \Sigma^* \times \mathbb{N}$ into a parameterized problem $Q' \subseteq \Sigma^* \times \mathbb{N}$ is an algorithm that given an instance $(x, k) \in \Sigma^* \times \mathbb{N}$ returns in time polynomial in $|x| + k$ an instance $(x', k') \in \Sigma^* \times \mathbb{N}$, such that $(x, k) \in Q$ if and only if $(x', k') \in Q'$, and such that $|x'| + k' \leq h(k)$ for some computable function $h : \mathbb{N} \to \mathbb{R}$. The function h is called the *size* of the compression, and when the alphabet Σ is $\{0, 1\}$, it is called the *bit-size* of the compression. A parameterized problem Q is said to admit a compression of size h if there exists a compression of size h for Q into *some* parameterized problem. A compression for a parameterized problem Q into itself is called a *kernelization*, or simply a *kernel*, for Q.

A *transformation* from a parameterized problem $Q \subseteq \Sigma^* \times \mathbb{N}$ into a parameterized problem $Q' \subseteq \Sigma^* \times \mathbb{N}$ is an algorithm that given an instance $(x, k) \in \Sigma^* \times \mathbb{N}$ returns in time polynomial in $|x| + k$ an instance $(x', k') \in \Sigma^* \times \mathbb{N}$, such that $(x, k) \in Q$ if and only if $(x', k') \in Q'$, and such that $k' \leq h(k)$ for some computable function $h : \mathbb{N} \to \mathbb{R}$. Note that in contrast to the notion of compression, the bound $h(k)$ applies here only to k', not to $|x'| + k'$. If h is linear, the transformation is called *linear-parameter*. Such transformations are useful for deriving lower bounds on the compression size of parameterized problems from those of others, as illustrated by the following standard lemma.

▶ **Lemma 10.** *Let Q and Q' be parameterized problems, whose parameters are denoted by k, and suppose that there exists a linear-parameter transformation from Q into Q'. Then, for every positive integer d, if Q' admits a compression of size $O(k^d)$, then so does Q.*

Proof. Composing a compression of size $O(k^d)$ for Q' with a linear-parameter transformation from Q into Q' yields a compression of size $O(k^d)$ for Q. ◀

2.4 Satisfiability Problems

For a positive integer k, a k-CNF formula is a Boolean formula in conjunctive normal form, where each clause contains k literals involving k distinct variables. The classic k-SAT problem asks whether a given k-CNF formula is satisfiable. The following theorem, proved by Dell and van Melkebeek [11], provides a conditional lower bound on the compression size of the k-SAT problem with $k \geq 3$ when parameterized by the number of variables.

▶ **Theorem 11** ([11]). *For every integer $k \geq 3$ and any real $\varepsilon > 0$, the k-SAT problem parameterized by the number of variables n does not admit a compression of size $O(n^{k-\varepsilon})$ unless* $\mathsf{NP} \subseteq \mathsf{coNP/poly}$.

3 A Lower Bound for Constrained Coloring

In this section, we prove a lower bound on the compression size of R-Constrained Coloring problems in terms of the arity of the OR relations definable from R (see Definitions 1 and 9). We first introduce a list-coloring variant of the R-Constrained Coloring problem, then prove a lower bound on its compression size, and finally derive the lower bound for the non-list variant, stated as Theorem 2.

3.1 Constrained List Coloring

For a positive integer q, the input of the q-List Coloring problem consists of a graph $G = (V, E)$ and a function $L : V \to \mathcal{P}([q])$, specifying a set of available colors for each vertex in G. The goal is to decide whether there exists a proper coloring $c : V \to [q]$ of G satisfying $c(v) \in L(v)$ for all $v \in V$. We refer to such a coloring as a *proper list-coloring* of (G, L). We define the R-Constrained List Coloring problem associated with a relation R as follows.

▶ **Definition 12.** *For positive integers q and r, let $R \subseteq [q]^r$ be a relation of arity r over $[q]$. The input of the R-Constrained List Coloring problem consists of a graph $G = (V, E)$, a function $L : V \to \mathcal{P}([q])$, and a collection $\mathcal{F} \subseteq V^r$ of r-tuples of vertices. The goal is to decide whether there exists a proper list-coloring $c : V \to [q]$ of (G, L), such that for every r-tuple $(x_1, \ldots, x_r) \in \mathcal{F}$, it holds that $(c(x_1), \ldots, c(x_r)) \in R$. The number of constraints in the instance $(G, L, \mathcal{F})$ is $|E| + |\mathcal{F}|$.*

The following theorem provides a conditional lower bound on the compression size of R-Constrained List Coloring problems parameterized by the number of vertices.

▶ **Theorem 13.** *For positive integers q, r, and $k \geq 3$, let $R \subseteq [q]^r$ be a relation of arity r over $[q]$ from which an OR relation of arity k is definable. Then, for any real $\varepsilon > 0$, the R-Constrained List Coloring problem parameterized by the number of vertices n does not admit a compression of size $O(n^{k-\varepsilon})$ unless* $\mathsf{NP} \subseteq \mathsf{coNP/poly}$.

The proof of Theorem 13 uses a simple graph gadget, designed to forbid a specific color assignment on two given vertices of a graph, as described in the following lemma. The proof is given in the full version of the paper.

▶ **Lemma 14.** *For every integer $q \geq 3$, there exists a polynomial-time algorithm that, given a graph $G = (V, E)$, a list function $L : V \to \mathcal{P}([q])$, two distinct vertices $u_1, u_2 \in V$, and colors $\alpha_1, \alpha_2 \in [q]$, extends the instance (G, L) of q-List Coloring to an instance (G', L') by adding a constant number of vertices and edges, so that for every proper list-coloring $c : V \to [q]$ of (G, L), c can be extended to a proper list-coloring of (G', L') if and only if $(c(u_1), c(u_2)) \neq (\alpha_1, \alpha_2)$.*

We are ready to prove Theorem 13.

Proof of Theorem 13. Fix positive integers q, r, and $k \geq 3$, and a real $\varepsilon > 0$, and let $R \subseteq [q]^r$ be a relation of arity r over $[q]$ from which an OR relation of arity k is definable. By Definition 9, there exist a set $J = \{j_1, \ldots, j_k\} \subseteq [r]$ of k distinct indices, sets $D_j = \{\alpha_j, \beta_j\} \subseteq [q]$ with $\alpha_j \neq \beta_j$ for $j \in J$, and sets $D_j = \{\alpha_j\} \subseteq [q]$ for $j \in [r] \setminus J$, such that all tuples in the product $D_1 \times \cdots \times D_r$ except the tuple $\alpha = (\alpha_1, \ldots, \alpha_r)$ belong to R.

Since $k \geq 3$, Theorem 11 implies that the k-SAT problem parameterized by the number of variables n does not admit a compression of size $O(n^{k-\varepsilon})$ unless $\mathsf{NP} \subseteq \mathsf{coNP/poly}$. In what follows, we present a linear-parameter transformation from the k-SAT problem parameterized by the number of variables to the R-Constrained List Coloring problem parameterized by the number of vertices. By Lemma 10, it follows that if the R-Constrained List Coloring problem parameterized by the number of vertices n admits a compression of size $O(n^{k-\varepsilon})$, then so does the k-SAT problem parameterized by the number of variables n, implying that $\mathsf{NP} \subseteq \mathsf{coNP/poly}$.

Consider an instance of the k-SAT problem, namely, a k-CNF formula φ on n variables, denoted by $x_1, \ldots, x_n$. We transform such an instance into an instance of R-Constrained List Coloring, consisting of a graph $G = (V, E)$, a list function $L : V \to \mathcal{P}([q])$, and a collection $\mathcal{F} \subseteq V^r$, as follows.

1. The vertex set of G includes $2nk$ vertices, denoted by $t_{i,j}$ and $f_{i,j}$ for all $i \in [n]$ and $j \in J$. For every pair $(i, j) \in [n] \times J$, the vertices $t_{i,j}$ and $f_{i,j}$ are adjacent in G, and their lists are defined by $L(t_{i,j}) = L(f_{i,j}) = D_j$. This ensures that, for every pair $(i, j) \in [n] \times J$, any proper list-coloring of (G, L) assigns color α_j to one of $t_{i,j}$ and $f_{i,j}$ and color β_j to the other.
2. The graph G includes $r - k$ isolated vertices, denoted by v_j for $j \in [r] \setminus J$, whose lists are defined by $L(v_j) = D_j$. This ensures that any proper list-coloring of (G, L) assigns color α_j to v_j for all $j \in [r] \setminus J$.
3. For each $i \in [n]$, we extend the current instance (G, L) as follows. For every $s \in [k - 1]$, we apply the efficient algorithm from Lemma 14 to the pair of vertices $(t_{i,j_s}, t_{i,j_{s+1}})$ twice: once to forbid the colors $(\alpha_{j_s}, \beta_{j_{s+1}})$ and once to forbid the colors $(\beta_{j_s}, \alpha_{j_{s+1}})$. This ensures, for each $i \in [n]$, that any proper list-coloring of (G, L) assigns to all vertices $t_{i,j}$ with $j \in J$ either their corresponding colors α_j or their corresponding colors β_j, while imposing no further constraints on the coloring.
4. We finally define the collection $\mathcal{F} \subseteq V^r$. For each clause $(y_1 \vee \cdots \vee y_k)$ in φ, we add to $\mathcal{F}$ the r-tuple $z = (z_1, \ldots, z_r)$, where for each $j \in [r]$, the vertex z_j is defined as follows.
   - If $j \in J$, let $s \in [k]$ be the index for which $j = j_s$, and let $i \in [n]$ be the index of the variable x_i appearing in the literal y_s.
     - If $y_s = x_i$, then set $z_j = t_{i,j}$.
     - If $y_s = \overline{x_i}$, then set $z_j = f_{i,j}$.
   - If $j \in [r] \setminus J$, then set $z_j = v_j$.

This completes the description of the instance $(G, L, \mathcal{F})$. It is straightforward to verify that it can be constructed in polynomial time. We now analyze the number of vertices in G.

The vertex set includes the $2nk$ vertices $t_{i,j}$ and $f_{i,j}$ with $(i,j) \in [n] \times J$ from Item 1 of the construction, as well as the $r-k$ vertices v_j with $j \in [r] \setminus J$ from Item 2 of the construction. Moreover, for each $i \in [n]$, Item 3 of the construction applies the efficient algorithm from Lemma 14 a constant number of times to the vertices $t_{i,j}$ with $j \in J$, introducing a constant number of additional vertices. Therefore, the total number of vertices in G is $O(n)$, and the transformation is linear-parameter.

For correctness, we shall prove that the formula φ is satisfiable if and only if $(G, L, \mathcal{F})$ is a YES instance of the R-Constrained List Coloring problem. Suppose first that φ is satisfiable, and consider a satisfying assignment for φ. We define a coloring $c : V \to [q]$ of G as follows. First, for each $j \in [r] \setminus J$, define $c(v_j) = \alpha_j$. Next, for each $i \in [n]$, if x_i is set to True, then define $c(t_{i,j}) = \beta_j$ and $c(f_{i,j}) = \alpha_j$ for all $j \in J$, and if x_i is set to False, then define $c(t_{i,j}) = \alpha_j$ and $c(f_{i,j}) = \beta_j$ for all $j \in J$. By definition, the coloring c respects the list function L, and for all $i \in [n]$ and $j \in J$, it assigns distinct colors to the adjacent vertices $t_{i,j}$ and $f_{i,j}$. Since for each $i \in [n]$, the vertices $t_{i,j}$ with $j \in J$ are either all assigned the corresponding colors α_j or all assigned the corresponding colors β_j, by Lemma 14, we can properly extend c to the vertices added in Item 3 of the construction. It follows that c is a proper list-coloring of (G, L).

We now show that for every clause $(y_1 \vee \cdots \vee y_k)$ in φ, its associated r-tuple $z = (z_1, \ldots, z_r)$ in $\mathcal{F}$, defined in Item 4 of the construction, satisfies $c(z) = (c(z_1), \ldots, c(z_r)) \in R$. By construction, for each $j \in [r]$, we have $L(z_j) = D_j$, hence $c(z) \in D_1 \times \cdots \times D_r$. Further, for some $s \in [k]$, the literal y_s is evaluated to True under the given assignment. Let $i \in [n]$ be the index of the variable x_i appearing in the literal y_s. If $y_s = x_i$, then the value of x_i is True, hence $c(z_{j_s}) = c(t_{i,j_s}) = \beta_{j_s}$, and if $y_s = \overline{x_i}$, then the value of x_i is False, hence $c(z_{j_s}) = c(f_{i,j_s}) = \beta_{j_s}$. In both cases, $c(z_{j_s}) = \beta_{j_s} \neq \alpha_{j_s}$, hence $c(z) \neq \alpha$. Since the only tuple in $D_1 \times \cdots \times D_r$ that does not belong to R is α, it follows that $c(z) \in R$, as required.

For the converse direction, suppose that there exists a proper list-coloring $c : V \to [q]$ of (G, L), such that for every r-tuple $z = (z_1, \ldots, z_r) \in \mathcal{F}$, we have $c(z) = (c(z_1), \ldots, c(z_r)) \in R$. We define an assignment for the variables $x_1, \ldots, x_n$ of φ as follows. Fix some $i \in [n]$. For every $j \in J$, the vertices $t_{i,j}$ and $f_{i,j}$ are adjacent in G, and the list function L guarantees that $c(t_{i,j}) \in D_j$ and $c(f_{i,j}) \in D_j$, where $D_j = \{\alpha_j, \beta_j\}$. Moreover, Item 3 of the construction ensures that either $c(t_{i,j}) = \alpha_j$ for all $j \in J$, and thus $c(f_{i,j}) = \beta_j$ for all $j \in J$, or $c(t_{i,j}) = \beta_j$ for all $j \in J$, and thus $c(f_{i,j}) = \alpha_j$ for all $j \in J$. We set the variable x_i to False in the former case, and to True in the latter. We claim that this assignment satisfies φ. To see this, consider a clause $(y_1 \vee \cdots \vee y_k)$ in φ and its associated r-tuple $z = (z_1, \ldots, z_r)$ in $\mathcal{F}$, as defined in Item 4 of the construction. Notice that the list function L ensures that $c(z) = (c(z_1), \ldots, c(z_r)) \in D_1 \times \cdots \times D_r$. Further, by $c(z) \in R$, it follows that $c(z) \neq \alpha$, hence there exists some $j \in J$ such that $c(z_j) = \beta_j$. Let $s \in [k]$ be the index for which $j = j_s$, and let $i \in [n]$ be the index of the variable x_i that appears in the literal y_s. If $y_s = x_i$, then the fact that $c(t_{i,j}) = c(z_j) = \beta_j$ implies that x_i is set to True, and if $y_s = \overline{x_i}$, then the fact that $c(f_{i,j}) = c(z_j) = \beta_j$ implies that x_i is set to False. In both cases, y_s is evaluated to True, hence the clause is satisfied. This completes the proof. ◀

3.2 Constrained Coloring

We now prove Theorem 2, which asserts that if an OR relation of arity $k \geq 3$ is definable from a permutation-invariant relation $R \subseteq [q]^r$, then the R-Constrained Coloring problem parameterized by the number of vertices n does not admit a compression of size $O(n^{k-\varepsilon})$ for any $\varepsilon > 0$ unless NP $\subseteq$ coNP/poly. The proof relies on the following lemma.

▶ **Lemma 15.** *For positive integers q and r, let $R \subseteq [q]^r$ be a permutation-invariant relation of arity r over $[q]$. Then there exists a linear-parameter transformation from the R-Constrained List Coloring problem to the R-Constrained Coloring problem, both parameterized by the number of vertices.*

Proof. Let $R \subseteq [q]^r$ be a permutation-invariant relation. Consider an instance of the R-Constrained List Coloring problem, consisting of a graph $G = (V, E)$, a list function $L : V \to \mathcal{P}([q])$, and a collection $\mathcal{F} \subseteq V^r$. Define a transformation that, given such an instance, returns the pair $(G', \mathcal{F})$, where $G' = (V', E')$ is the graph obtained from G by adding a clique on q vertices, denoted by $z_1, \ldots, z_q$, and connecting each vertex $v \in V$ to all vertices z_i with $i \in [q] \setminus L(v)$. The number of vertices in G' is $|V| + q$, hence the transformation is linear-parameter.

For correctness, suppose first that $(G, L, \mathcal{F})$ is a YES instance of R-Constrained List Coloring, and consider a proper list-coloring $c : V \to [q]$ of (G, L), such that for every r-tuple $(x_1, \ldots, x_r) \in \mathcal{F}$, it holds that $(c(x_1), \ldots, c(x_r)) \in R$. Let c' be the coloring of G' that extends c by assigning to the vertex z_i the color i for all $i \in [q]$. The coloring c' clearly assigns distinct colors to the endpoints of every edge of G and of every edge in the clique $\{z_1, \ldots, z_q\}$. Further, since c respects the list function L, every vertex $v \in V$ satisfies $c'(v) = c(v) \in L(v)$, and thus $c'(v) \neq c'(z_i)$ whenever $i \in [q] \setminus L(v)$. This implies that c' forms a proper coloring of G'. As an extension of c, it also satisfies the constraints in $\mathcal{F}$, so it forms a valid solution for the instance $(G', \mathcal{F})$ of R-Constrained Coloring.

Conversely, suppose that $(G', \mathcal{F})$ is a YES instance of R-Constrained Coloring, and consider a proper coloring $c' : V' \to [q]$ of G', such that for every r-tuple $(x_1, \ldots, x_r) \in \mathcal{F}$, it holds that $(c'(x_1), \ldots, c'(x_r)) \in R$. Since the vertices $z_1, \ldots, z_q$ are pairwise adjacent in G', they receive distinct colors under c', so there exists a permutation $\pi : [q] \to [q]$ such that $\pi(c'(z_i)) = i$ for all $i \in [q]$. Let $c : V \to [q]$ denote the coloring of G defined by $c(v) = \pi(c'(v))$ for all $v \in V$. For every pair of adjacent vertices u and v in G, it holds that $c'(u) \neq c'(v)$, and thus $c(u) \neq c(v)$, so c is a proper coloring of G. Furthermore, c respects the list function L. Indeed, for every vertex v and every $i \in [q] \setminus L(v)$, the vertices v and z_i are adjacent in G', hence $c'(v) \neq c'(z_i)$, which implies that $c(v) = \pi(c'(v)) \neq \pi(c'(z_i)) = i$, yielding that $c(v) \in L(v)$. Finally, for every r-tuple $(x_1, \ldots, x_r) \in \mathcal{F}$, it holds that $(c'(x_1), \ldots, c'(x_r)) \in R$, and since R is permutation-invariant, it follows that $(c(x_1), \ldots, c(x_r)) \in R$. This shows that $(G, L, \mathcal{F})$ is a YES instance of R-Constrained List Coloring, as required. ◀

We are ready to derive Theorem 2.

Proof of Theorem 2. For positive integers q, r, and $k \geq 3$, let $R \subseteq [q]^r$ be a permutation-invariant relation of arity r over $[q]$ from which an OR relation of arity k is definable. By Lemma 15, there exists a linear-parameter transformation from the R-Constrained List Coloring problem to the R-Constrained Coloring problem, both parameterized by the number of vertices n. If for some $\varepsilon > 0$, the latter admits a compression of size $O(n^{k-\varepsilon})$, then by Lemma 10, so does the former. This implies, by Theorem 13, that NP $\subseteq$ coNP/poly, and we are done. ◀

We conclude this section by highlighting the role of OR-definability in the kernel complexity of R-Constrained Coloring problems. Theorem 2 implies that if an OR relation of arity r is definable from a permutation-invariant relation R of arity $r \geq 3$, then the trivial kernel for the R-Constrained Coloring problem parameterized by the number of vertices is nearly optimal under a standard complexity assumption. As mentioned in the introduction, this definability condition is crucial, because if no OR of arity r is definable from R, a smaller kernel is achievable. This stems from the following result of Carbonnel [8].

▶ **Proposition 16** ([8]). *For positive integers q and $r \geq 3$, let $R \subseteq [q]^r$ be a relation of arity r over $[q]$, such that no* OR *relation of arity r is definable from R. Then the R-*CONSTRAINED COLORING *problem parameterized by the number of vertices n admits a kernel with $O(n^{r-\varepsilon})$ constraints for $\varepsilon = 2^{1-r}$.*

4 Concluding Remarks

This paper studies the kernel complexity of R-CONSTRAINED COLORING problems, a class of constraint satisfaction problems that combine the non-inequality relation over a finite domain together with an additional relation R. We provide a general conditional lower bound on the compression size of such problems when parameterized by the number of variables and demonstrate its applicability to several concrete parameterized coloring problems. These include the (d, ℓ, q)-UNIFORMLY RAINBOW FREE COLORING problems and the q-COLORING problems on ℓ-uniform hypergraphs, both parameterized by the number of vertices, as well as the q-COLORING problems on graphs, when parameterized by the vertex-deletion distance to a disjoint union of cliques of size at most t. For all of these problems, and for all admissible values of their relevant constants, we obtain nearly matching upper and lower bounds on the kernel complexity under the assumption $\mathsf{NP} \not\subseteq \mathsf{coNP/poly}$.

It will be interesting to find additional applications of the bounds presented here for determining the kernel complexity of other coloring problems. Natural candidates are the families of H-COLORING and LIST H-COLORING problems, which concern the existence of a homomorphism from a given graph to a fixed target graph H, when parameterized by the number of vertices, by the vertex cover number, or by other structural parameters of interest (see, e.g., [17, 9, 1, 22]). A more ambitious objective would be to characterize, under a plausible complexity assumption, the behavior of the kernel complexity of the R-CONSTRAINED COLORING problem for an arbitrary relation R.

References

1 Yael Berkman and Ishay Haviv. Kernelization for H-coloring. In *Proc. of the 20th International Symposium on Parameterized and Exact Computation (IPEC'25)*, pages 5:1–17, 2025. `doi:10.4230/LIPIcs.IPEC.2025.5`.

2 Christian Bessiere, Clément Carbonnel, and George Katsirelos. Chain length and CSPs learnable with few queries. In *Proc. of the 34th AAAI Conference on Artificial Intelligence (AAAI'20)*, pages 1420–1427, 2020. `doi:10.1609/AAAI.V34I02.5499`.

3 Stijn Beukers. Extending the scope of algebraic kernelization for constraint satisfaction problems. M.Sc. Thesis, Technische Universiteit Eindhoven, 2021.

4 Joshua Brakensiek and Venkatesan Guruswami. Redundancy is all you need. In *Proc. of the 57th Annual ACM Symposium on Theory of Computing (STOC'25)*, pages 1614–1625, 2025. `doi:10.1145/3717823.3718212`.

5 Joshua Brakensiek, Venkatesan Guruswami, Bart M. P. Jansen, Victor Lagerkvist, and Magnus Wahlström. The richness of CSP non-redundancy. *arXiv*, abs/2507.07942, 2025. `doi:10.48550/arXiv.2507.07942`.

6 Andrei A. Bulatov. A dichotomy theorem for nonuniform CSPs. In *Proc. of the 58th IEEE Annual Symposium on Foundations of Computer Science (FOCS'17)*, pages 319–330, 2017. `doi:10.1109/FOCS.2017.37`.

7 Leizhen Cai. Parameterized complexity of vertex colouring. *Discret. Appl. Math.*, 127(3):415–429, 2003. `doi:10.1016/S0166-218X(02)00242-1`.

8 Clément Carbonnel. On redundancy in constraint satisfaction problems. In *Proc. of the 28th International Conference on Principles and Practice of Constraint Programming (CP'22)*, pages 11:1–15, 2022. `doi:10.4230/LIPIcs.CP.2022.11`.

9 Hubie Chen, Bart M. P. Jansen, Karolina Okrasa, Astrid Pieterse, and Paweł Rzążewski. Sparsification lower bounds for list H-coloring. *ACM Trans. Comput. Theory*, 15(3–4):8:1–23, 2023. Preliminary version in ISAAC'20. doi:10.1145/3612938.

10 Hubie Chen, Bart M. P. Jansen, and Astrid Pieterse. Best-case and worst-case sparsifiability of Boolean CSPs. *Algorithmica*, 82(8):2200–2242, 2020. Preliminary version in IPEC'18. doi:10.1007/S00453-019-00660-Y.

11 Holger Dell and Dieter van Melkebeek. Satisfiability allows no nontrivial sparsification unless the polynomial-time hierarchy collapses. *J. ACM*, 61(4):23:1–27, 2014. Preliminary version in STOC'10. doi:10.1145/2629620.

12 Fedor V. Fomin, Daniel Lokshtanov, Saket Saurabh, and Meirav Zehavi. *Kernelization: Theory of Parameterized Preprocessing*. Cambridge University Press, 2019.

13 Ishay Haviv and Dror Rabinovich. A near-optimal kernel for a coloring problem. *Discret. Appl. Math.*, 377:66–73, 2025. doi:10.1016/J.DAM.2025.06.065.

14 Ishay Haviv and Dror Rabinovich. Kernelization for orthogonality dimension. *J. Comput. Syst. Sci.*, 157:103757, 2026. Preliminary version in IPEC'24. doi:10.1016/J.JCSS.2026.103757.

15 Bart M. P. Jansen and Stefan Kratsch. Data reduction for graph coloring problems. *Inf. Comput.*, 231:70–88, 2013. Preliminary version in FCT'11. doi:10.1016/J.IC.2013.08.005.

16 Bart M. P. Jansen and Astrid Pieterse. Sparsification upper and lower bounds for graph problems and not-all-equal SAT. *Algorithmica*, 79(1):3–28, 2017. Preliminary version in IPEC'15. doi:10.1007/S00453-016-0189-9.

17 Bart M. P. Jansen and Astrid Pieterse. Optimal data reduction for graph coloring using low-degree polynomials. *Algorithmica*, 81(10):3865–3889, 2019. Preliminary version in IPEC'17. doi:10.1007/S00453-019-00578-5.

18 Bart M. P. Jansen and Astrid Pieterse. Optimal sparsification for some binary CSPs using low-degree polynomials. *ACM Trans. Comput. Theory*, 11(4):28:1–26, 2019. Preliminary version in MFCS'16. doi:10.1145/3349618.

19 Bart M. P. Jansen and Michał Włodarczyk. Optimal polynomial-time compression for Boolean max CSP. *ACM Trans. Comput. Theory*, 16(1):4:1–20, 2024. doi:10.1145/3624704.

20 Victor Lagerkvist and Magnus Wahlström. Sparsification of SAT and CSP problems via tractable extensions. *ACM Trans. Comput. Theory*, 12(2):13:1–29, 2020. Preliminary version in CP'17. doi:10.1145/3389411.

21 Victor Lagerkvist and Magnus Wahlström. The (coarse) fine-grained structure of NP-hard SAT and CSP problems. *ACM Trans. Comput. Theory*, 14(1):2:1–54, 2022. doi:10.1145/3492336.

22 Marta Piecyk, Astrid Pieterse, Pawel Rzążewski, and Magnus Wahlström. Kernelization for list H-coloring for graphs with small vertex cover. *arXiv*, abs/2507.12005, 2025. doi:10.48550/arXiv.2507.12005.

23 Thomas J. Schaefer. The complexity of satisfiability problems. In *Proc. of the 10th Annual ACM Symposium on Theory of Computing (STOC'78)*, pages 216–226. ACM, 1978. doi:10.1145/800133.804350.

24 Martijn M. R. Schalken. Efficient kernels for q-coloring. M.Sc. Thesis, Technische Universiteit Eindhoven, 2020.

25 Chee-Keng Yap. Some consequences of non-uniform conditions on uniform classes. *Theor. Comput. Sci.*, 26(3):287–300, 1983. doi:10.1016/0304-3975(83)90020-8.

26 Dmitriy Zhuk. A proof of the CSP dichotomy conjecture. *J. ACM*, 67(5):30:1–78, 2020. Preliminary version in FOCS'17. doi:10.1145/3402029.