

Kernelization for H -Coloring

Yael Berkman

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

Ishay Haviv 

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

Abstract

For a fixed graph H , the H -COLORING problem asks whether a given graph admits an edge-preserving function from its vertex set to that of H . A seminal theorem of Hell and Nešetřil asserts that the H -COLORING problem is NP-hard whenever H is loopless and non-bipartite. A result of Jansen and Pieterse implies that for every graph H , the H -COLORING problem parameterized by the vertex cover number k admits a kernel with $O(k^{\Delta(H)})$ vertices and bit-size bounded by $O(k^{\Delta(H)} \cdot \log k)$, where $\Delta(H)$ denotes the maximum degree in H . For the case where H is a complete graph on at least three vertices, this kernel size nearly matches conditional lower bounds established by Jansen and Kratsch and by Jansen and Pieterse.

This paper presents new upper and lower bounds on the kernel size of H -COLORING problems parameterized by the vertex cover number. The upper bounds arise from two kernelization algorithms. The first is purely combinatorial, and its size is governed by a structural quantity of the graph H , called the non-adjacency witness number. As applications, we obtain kernels whose size is bounded by a fixed polynomial for natural classes of graphs H with unbounded maximum degree, such as planar graphs and, more broadly, graphs with bounded degeneracy. More strikingly, we show that for almost every graph H , the degree of the polynomial that bounds the size of our combinatorial kernel grows only logarithmically in $\Delta(H)$. Our second kernel leverages linear-algebraic tools and involves the notion of faithful independent representations of graphs. It strengthens the general bound from prior work and, among other applications, yields near-optimal kernels for problems concerning the dimension of orthogonal graph representations over finite fields. We complement our kernelization results with conditional lower bounds, thereby nearly settling the kernel complexity of the problem for various target graphs H .

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

Keywords and phrases Kernelization, Graph coloring, Graph homomorphism

Digital Object Identifier 10.4230/LIPIcs.IPEC.2025.5

Related Version Full Version: <http://arxiv.org/abs/2507.13129>

Funding Yael Berkman: Research supported by the Israel Science Foundation (grant No. 1218/20). Ishay Haviv: Research supported by the Israel Science Foundation (grant No. 1218/20).

Acknowledgements We thank the anonymous reviewers for their valuable feedback.

1 Introduction

A central concept in graph theory is that of *graph homomorphisms*, namely, functions from the vertex set of one graph G to the vertex set of another graph H that map adjacent vertices in G to adjacent vertices in H . If such a function exists, the graph G is said to be H -colorable. For a fixed graph H , the computational H -COLORING problem asks whether a given input graph is H -colorable. One may always assume that the target graph H is a core, i.e., a graph admitting no homomorphism to a proper subgraph, as replacing H with a minimal core subgraph does not change the problem. This class of problems occupies a fundamental place



© Yael Berkman and Ishay Haviv;
licensed under Creative Commons License CC-BY 4.0

20th International Symposium on Parameterized and Exact Computation (IPEC 2025).

Editors: Akanksha Agrawal and Erik Jan van Leeuwen; Article No. 5; pp. 5:1–5:17

Leibniz International Proceedings in Informatics



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

in computational graph theory and has been studied extensively for decades (see, e.g., [21]). A milestone in the area is the dichotomy theorem of Hell and Nešetřil [20], which asserts that the problem is solvable in polynomial time whenever H has a loop or is bipartite, and is NP-hard otherwise. In recent years, H -COLORING problems have received attention from multiple perspectives, such as algorithmic design [8, 41, 39], computational lower bounds [5], and fine-grained complexity [35, 37, 10].

This paper investigates the family of H -COLORING problems from the perspective of parameterized complexity, more specifically, from the viewpoint of kernelization. Parameterized complexity is a framework for analyzing decision problems whose instances are equipped with a quantitative parameter, with the goal of determining the effect of the parameter's value on the problem's computational complexity. A central theme in this field, known as kernelization or data reduction, seeks to design efficient preprocessing algorithms that substantially reduce the input size. These algorithms, called *kernels*, take an input instance (x, k) , where x is the main input and k is the parameter, and transform it in polynomial time into an equivalent instance (x', k') whose bit-size is bounded by $f(k)$ for a computable function $f : \mathbb{N} \rightarrow \mathbb{N}$ depending only on k . The slowest achievable growth rate of such a function f reflects the problem's compressibility limits with respect to the parameter k . In the context of graph problems, a widely-studied and ubiquitous parameter is the size of a vertex cover, i.e., a set of vertices that includes at least one endpoint of every edge in the graph. When H -COLORING is parameterized by the vertex cover number k , the input consists of a graph G along with a vertex cover of G of size k .

An archetypal problem within the H -COLORING framework is the q -COLORING problem for a fixed integer q , which has attracted persistent interest in algorithmic and complexity-theoretic research. This problem asks whether the vertices of a given graph can be colored with q colors, so that no two adjacent vertices share a color. It precisely coincides with the H -COLORING problem where H is the complete graph on q vertices, and is well known to be NP-hard for $q \geq 3$. The kernelization complexity of the q -COLORING problem parameterized by the vertex cover number k was first comprehensively studied by Jansen and Kratsch in [23] (see also [22]). They showed that for every integer $q \geq 3$, the problem admits a kernel producing graphs with $O(k^q)$ vertices which can be encoded in $O(k^q)$ bits. This result was subsequently refined by Jansen and Pieterse [24], who obtained a kernel with $O(k^{q-1})$ vertices and bit-size $O(k^{q-1} \cdot \log k)$. Remarkably, this bound is nearly optimal, as it was proved in [23, 24] that for every integer $q \geq 3$ and any real $\varepsilon > 0$, the problem does not admit a kernel of size $O(k^{q-1-\varepsilon})$ unless $\text{NP} \subseteq \text{coNP/poly}$, a containment that would imply the collapse of the polynomial-time hierarchy to its third level [42].

Another captivating case of the H -COLORING problem arises from a geometric lens on graph theory, namely, through the concept of *orthogonal representations* introduced by Lovász [28] in 1979. For a positive integer d and a field \mathbb{F} , a d -dimensional orthogonal representation of a graph $G = (V, E)$ over \mathbb{F} assigns to each vertex $v \in V$ a non-self-orthogonal vector $x_v \in \mathbb{F}^d$, such that for every edge $\{u, v\} \in E$, the vectors x_u and x_v are orthogonal. Here, two vectors $x, y \in \mathbb{F}^d$ are called orthogonal if their standard inner product $\langle x, y \rangle = \sum_{i=1}^d x_i y_i$ equals zero, and a vector $x \in \mathbb{F}^d$ is self-orthogonal if $\langle x, x \rangle = 0$. The orthogonal representation is termed *faithful* if two vertices u and v are adjacent in G if and only if their vectors x_u and x_v are orthogonal. The minimum dimension d for which a graph admits a (faithful) orthogonal representation over \mathbb{F} captures an intriguing structural property, and has found applications in combinatorics, theoretical computer science, and information theory (see [29, Chapter 10] and, e.g., [30, 4, 2, 12, 17, 11]). This motivates the study of the d -ORTHO-DIM $_{\mathbb{F}}$ problem, which, for fixed d and \mathbb{F} , asks whether a given graph

admits a d -dimensional orthogonal representation over \mathbb{F} . This problem naturally fits within the H -COLORING framework by letting $H = H(\mathbb{F}, d)$ be the (possibly infinite) graph whose vertices are all non-self-orthogonal vectors in \mathbb{F}^d , with adjacency defined by orthogonality over \mathbb{F} . Note that the problem is known to be NP-hard for every integer $d \geq 3$ and every field \mathbb{F} [36]. It was recently shown [18] that for every integer $d \geq 3$ and every field \mathbb{F} , the d -ORTHO-DIM $_{\mathbb{F}}$ problem parameterized by the vertex cover number k admits a kernel with $O(k^d)$ vertices and bit-size $O(k^d)$, whereas for any $\varepsilon > 0$, it admits no kernel of size $O(k^{d-1-\varepsilon})$ unless $\text{NP} \subseteq \text{coNP}/\text{poly}$. A sharper upper bound was obtained in [18] for the real field \mathbb{R} , yielding a near-optimal kernel with $O(k^{d-1})$ vertices and bit-size $O(k^{d-1} \cdot \log k)$. For finite fields, however, the precise kernel complexity remained unresolved, with a multiplicative gap of roughly k separating the upper and lower bounds on the kernel size.

The near-optimal kernels for the q -COLORING problems were extended by Jansen and Pieterse in [24] to the H -COLORING problem for arbitrary graphs H . They showed that for every graph H , the H -COLORING problem parameterized by the vertex cover number k admits a kernel with $O(k^{\Delta(H)})$ vertices and bit-size $O(k^{\Delta(H)} \cdot \log k)$, where $\Delta(H)$ denotes the maximum degree of a vertex in H . This result was actually established in a stronger form, with respect to the less restrictive parameter known as the twin cover number. While the bound on the kernel size from [24] is essentially optimal when H is a complete graph, it is natural to ask to what extent the dependence on the maximum degree of H captures the kernel complexity of the problem for general graphs H . This question serves as the driving force behind our work, which explores the polynomial behavior of the kernel complexity of H -COLORING problems parameterized by the vertex cover number.

1.1 Our Contribution

The present paper provides upper and lower bounds on the kernel size of H -COLORING problems parameterized by the vertex cover number. Our upper bounds include two kernelization algorithms: the first is purely combinatorial, and the second leverages linear-algebraic tools. Below, we elaborate on each of them. Table 1 provides an overview of their applications.

The Combinatorial Kernel

Our first kernel is simple and combinatorial, inspired by the strategy developed for coloring problems by Jansen and Kratsch in [23]. Consider an instance of the H -COLORING problem parameterized by the vertex cover number k , namely, a graph $G = (V, E)$ and a vertex cover $X \subseteq V$ of G of size k . For a fixed integer q , the kernel starts with the subgraph of G induced by X , and then, for every set $S \subseteq X$ of size at most q whose vertices share a neighbor from $V \setminus X$ in G , introduces a new vertex whose neighborhood is exactly S . While the number of vertices in the constructed graph is clearly bounded by $O(k^q)$, we prove that the correctness of the kernelization for H -COLORING is governed by a structural quantity of the graph H , which we denote by $q(H)$ and refer to as the *non-adjacency witness number* of H .

Given a set of vertices in H that have no common neighbor, one may ask whether this can be certified by a succinct witness, namely, a small subset that also has no common neighbor. The non-adjacency witness number $q(H)$ is defined as the smallest size of such a subset, whose existence is guaranteed for every set of vertices in H with no common neighbor (see Definition 6). This notion has previously been studied in the literature, with particular attention given to graphs H with $q(H) = 2$ (see, e.g., [6, 14, 13]). We note that $q(H)$ may also be formulated as the smallest integer q for which the collection of open neighborhoods in H satisfies what is known as the Helly property of order q . Our analysis reveals that

the size of the aforementioned kernel for the H -COLORING problem parameterized by the vertex cover number is bounded by a polynomial, whose degree is given by the non-adjacency witness number of H .

► **Theorem 1.** *For every graph H , the H -COLORING problem parameterized by the vertex cover number k admits a kernel with $O(k^{q(H)})$ vertices and bit-size $O(k^{q(H)})$.*

We remark that when applying Theorem 1, it is sometimes useful to replace H with a core subgraph of H . As noted earlier, this substitution does not change the problem, however, it may reduce the non-adjacency witness number and thereby lead to a smaller kernel.

In an attempt to exploit Theorem 1 and to compare it with the kernel of $O(k^{\Delta(H)})$ vertices obtained in [24], we undertake a systematic study of the non-adjacency witness number of graphs. Our analysis illuminates connections to fundamental graph parameters, such as maximum degree, clique number, and degeneracy, as well as structural properties related to the presence of specific subgraphs. By combining these insights with Theorem 1, we obtain economical kernels for a broad range of graphs H . In particular, as demonstrated below, we achieve kernels whose size is bounded by a fixed polynomial for natural classes of graphs H with unbounded maximum degree.

For a fixed integer d , consider the class of d -degenerate graphs, namely, those graphs in which every subgraph has a vertex of degree at most d . We show that for every such graph H , its non-adjacency witness number satisfies $q(H) \leq d + 1$, so Theorem 1 implies that the corresponding H -COLORING problem parameterized by the vertex cover number k admits a kernel whose size is bounded by $O(k^{d+1})$. For planar graphs, which are known to be 5-degenerate, this yields a bound of $O(k^6)$ on the kernel size, and via a refined analysis, we further reduce it to $O(k^4)$. Perhaps most strikingly, we apply Theorem 1 to show that for *almost every graph* H , there is a kernel for the H -COLORING problem parameterized by the vertex cover number, where the degree of the polynomial bounding its size grows only logarithmically with the maximum degree of H . To this end, we provide a tight estimate for the typical non-adjacency witness number of random graphs.

The Algebraic Kernel

Our second kernel for the H -COLORING problem parameterized by the vertex cover number is more algebraic in nature. For certain graphs H , it allows us to sharpen the kernel size from Theorem 1 by a multiplicative linear factor. Somewhat surprisingly, the condition that makes this improvement applicable is related to the concept of faithful orthogonal representations of graphs (which, as described earlier, assign a non-self-orthogonal vector to each vertex, such that two vertices are adjacent if and only if their vectors are orthogonal). Concretely, we show that if a graph H admits a faithful d -dimensional orthogonal representation over some computationally efficient field, then the H -COLORING problem parameterized by the vertex cover number k admits a kernel with $O(k^{d-1})$ vertices and bit-size $O(k^{d-1} \cdot \log k)$.

In fact, we obtain this kernel size under a weaker assumption on the graph H , using the concept of *independent representations* of graphs. This notion refers to an assignment of a d -dimensional vector to each vertex of H , so that the vector associated with each vertex does not lie in the vector space spanned by the vectors of its neighborhood. Note that every orthogonal representation is, in particular, an independent representation. The minimum dimension of an independent representation of a graph over a given field characterizes a well-studied graph quantity, called *minrank*, which plays a pivotal role in the study of fundamental problems in information theory, such as Shannon capacity [15], index coding [1, 40, 16], storage capacity [33], and hat guessing games [38]. For our purposes, we introduce a *faithful*

analogue of this notion, demanding that two vertices u and v in the graph are adjacent if and only if the vector associated with u lies in the linear subspace spanned by the vectors of the neighborhood of v (see Definition 9). With this faithful variant in hand, we prove the following theorem (see also Theorem 14).

► **Theorem 2 (Simplified).** *If a graph H has a faithful d -dimensional independent representation over either a finite field or the real field \mathbb{R} , then the H -COLORING problem parameterized by the vertex cover number k admits a kernel with $O(k^{d-1})$ vertices and bit-size $O(k^{d-1} \cdot \log k)$.*

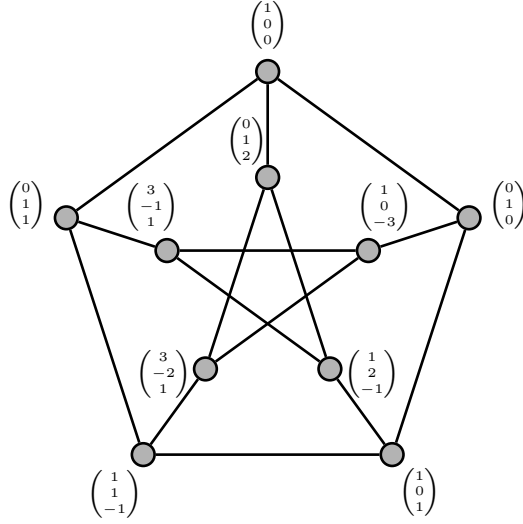
A few remarks are in order here. First, it is not difficult to show that every graph H has a faithful independent representation of dimension $\Delta(H) + 1$ over any sufficiently large field (see Lemma 13). By Theorem 2, this implies that for every graph H , the H -COLORING problem parameterized by the vertex cover number k admits a kernel with $O(k^{\Delta(H)})$ vertices and bit-size $O(k^{\Delta(H)} \cdot \log k)$. Therefore, Theorem 2 strengthens the kernel bound implied by [24]. It is worth noting that a result of Maehara and Rödl [31] asserts that every graph H has a faithful orthogonal representation over \mathbb{R} of dimension $2 \cdot \Delta(H)$, and it is an open question whether this dimension can be reduced to $\Delta(H) + 1$ (see, e.g., [29, Chapter 10.4]). Thus, for Theorem 2 to fully subsume the kernel bound of [24], it is somewhat crucial to adopt the more flexible notion of faithful independent representations (rather than orthogonal).

Second, we observe that for every graph H that possesses a faithful d -dimensional independent representation over some field, the non-adjacency witness number satisfies $q(H) \leq d$ (see Lemma 11). This shows that Theorem 2 improves the kernel size obtained via Theorem 1 only when the minimum dimension of a faithful independent representation of H matches the non-adjacency witness number of H (in fact, of a minimal core subgraph of H). This situation arises, for example, for the aforementioned graphs $H = H(\mathbb{F}, d)$, which correspond to the existence of a d -dimensional orthogonal representation over a field \mathbb{F} . As an application of Theorem 2, we unleash a near-optimal kernel for the d -ORTHO-DIM $_{\mathbb{F}}$ problems over finite fields \mathbb{F} , thereby settling a question posed in [18].

► **Theorem 3.** *For every integer $d \geq 3$ and every finite field \mathbb{F} , the d -ORTHO-DIM $_{\mathbb{F}}$ problem parameterized by the vertex cover number k admits a kernel with $O(k^{d-1})$ vertices and bit-size $O(k^{d-1} \cdot \log k)$.*

We further demonstrate the applicability of Theorem 2 for the celebrated family of *Kneser graphs*. For positive integers m and r with $m \geq 2r$, the Kneser graph $K(m, r)$ has all r -subsets of an m -element set as vertices, where two are adjacent if they are disjoint. It turns out that the non-adjacency witness number of $K(m, r)$ is $m - 2r + 2$, coinciding with its chromatic number that was determined in an influential paper of Lovász [27]. By Theorem 1, this shows that the $K(m, r)$ -COLORING problem parameterized by the vertex cover number k admits a kernel with $O(k^{m-2r+2})$ vertices and bit-size. To further reduce this kernel size, we prove that $K(m, r)$ admits a faithful $(m - 2r + 2)$ -dimensional independent representation over any sufficiently large field, which yields, via Theorem 2, a kernel with $O(k^{m-2r+1})$ vertices and bit-size $O(k^{m-2r+1} \cdot \log k)$. In particular, for the 3-chromatic Kneser graphs $K(2r + 1, r)$, which include the iconic Petersen graph $K(5, 2)$, the resulting kernel has near-quadratic size (see Figure 1).

The proof of Theorem 2 builds on a powerful sparsification technique that was introduced by Jansen and Pieterse in [25] and applied in their work [24]. To illustrate this technique, consider an instance of the H -COLORING problem parameterized by the vertex cover number, consisting of a graph $G = (V, E)$ and a vertex cover $X \subseteq V$ of G . The approach begins by introducing variables over some field \mathbb{F} that represent an assignment of vertices from H to



■ **Figure 1** The Petersen graph – A faithful 3-dimensional orthogonal representation over \mathbb{R} .

the vertices in X . Then, for every vertex $v \in V \setminus X$, one constructs one or more low-degree polynomials in the variables corresponding to its neighbors in X , which evaluate to zero if and only if v can be assigned a compatible vertex from H . The crux is that it suffices to retain in the graph G only a subset of the vertices in $V \setminus X$, whose corresponding polynomials linearly span all the others. The degree of the used polynomials determines the dimension of the entire space of polynomials, and consequently, the resulting kernel size. This method was applied in [24], where the assignment of each vertex of X is encoded by a set of variables representing an indicator vector over the binary field \mathbb{F}_2 .

The way we apply the machinery of [24, 25] to prove Theorem 2 is substantially different. We rely on the existence of a faithful d -dimensional independent representation of the graph H over some field \mathbb{F} , interpreting its vectors as encodings of the vertices of H . Accordingly, we associate each vertex in the vertex cover X with a d -dimensional variable vector over \mathbb{F} to represent its assignment. The kernelization algorithm first applies our combinatorial kernel to ensure that every vertex outside X has degree at most d . Then, it aims to reduce the number of vertices outside X that have degree exactly d . By the definition of an independent representation, a vertex v outside X can be assigned a compatible vertex from H only if every set of d vectors of its neighbors in X is linearly dependent over \mathbb{F} . This condition can be naturally captured by requiring the determinant polynomial over the variables corresponding to such d neighbors of v to be zero. Once the vectors of the neighbors of v span a subspace of dimension smaller than d , the assignment of a compatible vertex to v is essentially realized through some other vertex outside X of degree at most $d - 1$.

However, the degree of the determinant polynomial of a $d \times d$ matrix is d . In order to obtain polynomials of degree $d - 1$, which in turn yield a kernel with $O(k^{d-1})$ vertices, we use an idea from [18, 19] and require all vectors in the faithful independent representation of H to share a fixed value, say 1, in their first entry. When the field \mathbb{F} is sufficiently large relative to the number of vertices in H , this can be achieved by applying an invertible linear transformation to the vectors in the representation to ensure that all first entries are nonzero, and then scaling them appropriately. Over a smaller field, though, such a representation may not exist. We overcome this obstacle by showing that if a graph has a faithful d -dimensional independent representation over some field \mathbb{F} , then it also has one in which all vectors have 1

as their first entry, over any sufficiently large *extension* field of \mathbb{F} (see Lemma 12). This allows us to apply the approach described above over such an extension field and obtain a kernel of the desired size.

■ **Table 1** Kernel size for the H -COLORING problem parameterized by the vertex cover number k for various target graphs H .

Graph H	Kernel bit-size	Kernel type
d -degenerate	$O(k^{d+1})$	combinatorial
Planar	$O(k^4)$	combinatorial
Odd cycle C_{2m+1} , $m \geq 2$	$O(k^2)$	combinatorial
$H(\mathbb{F}, d)$, \mathbb{F} finite, $d \geq 3$	$\tilde{O}(k^{d-1})$	algebraic
Kneser graph $K(m, r)$	$\tilde{O}(k^{m-2r+1})$	algebraic

Lower Bounds

Given our combinatorial and algebraic kernels for the H -COLORING problem parameterized by the vertex cover number k , it is natural to ask how close the obtained kernel sizes come to being optimal. As already discussed, and shown in [23, 24], when H is the complete graph on q vertices with $q \geq 3$, the problem admits no kernel of size $O(k^{q-1-\varepsilon})$ for any $\varepsilon > 0$, unless $\text{NP} \subseteq \text{coNP/poly}$. This firmly settles the kernel complexity for all cases where H , or its minimal core subgraph, is a complete graph. Turning to other fundamental graphs, the first non-trivial case that arises is that of odd cycles of length at least five. For every integer $m \geq 2$, it is easy to verify that the cycle C_{2m+1} on $2m+1$ vertices has non-adjacency witness number $q(C_{2m+1}) = 2$. Consequently, Theorem 1 yields a kernel for the C_{2m+1} -COLORING problem parameterized by the vertex cover number k with $O(k^2)$ vertices and bit-size. Furthermore, neither Theorem 2 nor the kernel from [24] improves upon this bound. Another notable case is the family of 3-chromatic Kneser graphs $K(2r+1, r)$, for which Theorem 2 yields a kernel with $O(k^2)$ vertices and bit-size $O(k^2 \cdot \log k)$. The following theorem indicates that, in both cases, the obtained kernel sizes are nearly optimal.

► **Theorem 4.** *For each of the following cases, for any real number $\varepsilon > 0$, the H -COLORING problem parameterized by the vertex cover number k does not admit a kernel of size $O(k^{2-\varepsilon})$ unless $\text{NP} \subseteq \text{coNP/poly}$.*

1. $H = C_{2m+1}$ for an integer $m \geq 1$.
2. $H = K(2r+1, r)$ for an integer $r \geq 1$.

We prove Theorem 4 in a stronger form in several respects. First, we show that the near-quadratic lower bound persists even when the H -COLORING problem is parameterized by the total number of vertices in the input graph. Since the entire vertex set of a graph trivially forms a vertex cover, this directly implies the statement above. Second, the proof encompasses not only the graphs H stated in the theorem but also extends to various other graphs H . Specifically, the result is proved for all non-bipartite core graphs that are *projective*, a class that is known to cover nearly all graphs, including prominent and well-studied graph families (see, e.g., [26]). Third, as is standard in such results, the lower bound applies to the compressibility of the H -COLORING problem into any target problem, not necessarily itself. Our generalized result is stated as follows.

► **Theorem 5.** *For every non-bipartite projective core graph H and for any real number $\varepsilon > 0$, the H -COLORING problem parameterized by the number of vertices n does not admit a compression of size $O(n^{2-\varepsilon})$ unless $\text{NP} \subseteq \text{coNP/poly}$.*

The proof of Theorem 5 is based on a compression lower bound, due to Chen, Jansen, Okrasa, Pieterse, and Rzażewski [3], for a variant of the H -COLORING problem called LIST- H -COLORING. In this variant, the input consists of a graph G along with a list of allowed vertices of H for each vertex of G , and the goal is to decide whether there exists a homomorphism from G to H that respects these lists. This problem generalizes the standard H -COLORING problem, which corresponds to the special case where all lists are equal to the full vertex set of H . Feder, Hell, and Huang [7] introduced a class of geometric intersection graphs, called *bi-arc graphs*, and used this class to fully characterize the complexity of the LIST- H -COLORING problem. They showed that the problem is solvable in polynomial time if H is bi-arc, and NP-hard otherwise. Chen et al. [3] proved that, unless $\text{NP} \subseteq \text{coNP}/\text{poly}$, the LIST- H -COLORING problem parameterized by the number of vertices admits no sub-quadratic compression for any H that is not bi-arc. They also posed the question of whether such a lower bound holds for the (non-list) H -COLORING problem, when H is loopless and non-bipartite. As a step toward this challenge, Theorem 5 confirms the lower bound for the substantial class of non-bipartite projective core graphs H . To this end, we show that for such graphs H , instances of LIST- H -COLORING can be efficiently reduced into instances of H -COLORING, with the number of vertices preserved up to a multiplicative constant. The proof borrows a gadget construction due to Okrasa and Rzażewski [35] and echoes early NP-hardness proofs for H -COLORING problems [34, 32], as well as recent lower bounds in fine-grained complexity [35, 37].

When the H -COLORING problems are parameterized by the number of vertices, they admit a trivial kernel of quadratic size. Therefore, the lower bounds implied by Theorem 5 for the vertex cover parameterization are limited to the near-quadratic regime. To surpass this barrier, we establish a lower bound on the kernel size of the H -COLORING problem parameterized by the vertex cover number for all non-bipartite projective core graphs H with $q(H) \geq 4$. We show that, for each such graph H , the degree of any polynomial that bounds the kernel size is unlikely to be smaller than the non-adjacency witness number of H minus one. The precise statement and proof appear in the full version of the paper. Note that this lower bound is reasonably close to our upper bounds. Indeed, for every non-bipartite projective core graph H with $q(H) \geq 4$, the upper bound given in Theorem 1 exceeds this lower bound by only a near-linear multiplicative factor. Since almost all graphs are non-bipartite projective cores, our results determine the kernel complexity of the problem up to a near-linear factor for almost every graph H . Furthermore, if a non-bipartite projective core graph H admits a faithful independent representation of dimension $q(H) \geq 4$ over a computationally efficient field, then the kernel given in Theorem 2 essentially meets the lower bound (see also Theorem 14). As an illustrative consequence, we derive that the kernel sizes we achieve for all non-bipartite Kneser graphs are nearly tight (extending Item 2 of Theorem 4).

1.2 Outline

The rest of the paper is organized as follows. In Section 2, we collect several definitions and facts that will be used throughout the paper. In Section 3, we present and analyze our combinatorial and algebraic kernels for the H -COLORING problem parameterized by the vertex cover number, thereby proving Theorems 1 and 2. Further results, including lower bounds on compression size, the study of the non-adjacency witness number, and detailed applications to specific graphs H , appear in the full version of the paper. Concluding remarks and open questions appear in the final Section 4.

2 Preliminaries

Throughout the paper, we omit floor and ceiling signs when they are not crucial, and all logarithms are in base 2. For a positive integer n , we write $[n]$ to denote the set $\{1, 2, \dots, n\}$.

2.1 Graphs

The graphs considered in this paper are, unless stated otherwise, finite and simple, meaning they have no loops or parallel edges. For a graph $G = (V, E)$ and a set $X \subseteq V$, we let $G[X]$ denote the subgraph of G induced by X . The set X is called a *vertex cover* when the graph $G[V \setminus X]$ is edgeless. For a vertex $v \in V$, we let $N_G(v)$ denote the set of neighbors of v in G . As is customary, we denote by $\Delta(G)$ the maximum degree of a vertex in G . For two graphs $G = (V_G, E_G)$ and $H = (V_H, E_H)$, a *homomorphism* from G to H is a mapping $f : V_G \rightarrow V_H$ such that for all vertices $u, v \in V_G$ with $\{u, v\} \in E_G$, it holds that $\{f(u), f(v)\} \in E_H$. If there exists a homomorphism from G to H , the graph G is said to be *H -colorable*.

2.2 Linear Algebra

For a field \mathbb{F} and an integer d , two vectors $x, y \in \mathbb{F}^d$ are said to be *orthogonal* if $\langle x, y \rangle = 0$, with respect to the standard inner product defined by $\langle x, y \rangle = \sum_{i=1}^d x_i y_i$, where all operations are performed over \mathbb{F} . A vector $x \in \mathbb{F}^d$ is *self-orthogonal* if $\langle x, x \rangle = 0$, and *non-self-orthogonal* otherwise.

The *order* of a finite field is the number of its elements. It is well known that the order of any finite field is a prime power, and that every prime power is the order of some finite field. For two fields \mathbb{F} and \mathbb{K} , we say that \mathbb{K} is an *extension field* of \mathbb{F} if $\mathbb{F} \subseteq \mathbb{K}$ and the operations of \mathbb{F} agree with those of \mathbb{K} when restricted to elements in \mathbb{F} . A finite field of order p^m , for a prime p and a positive integer m , has an extension field of order $p^{m'}$ for every integer $m' \geq m$.

A multivariate polynomial over a field \mathbb{F} is called *homogeneous of degree d* if each of its monomials has degree d , and is called *multilinear* if each of them forms a product of distinct variables. The set of all multilinear homogeneous polynomials of degree d in n variables over a field \mathbb{F} forms a vector space over \mathbb{F} of dimension $\binom{n}{d}$.

2.3 Parameterized Complexity

We collect here basic definitions on kernelization from the field of parameterized complexity. For an in-depth introduction to the topic, the reader is referred to, e.g., [9]. A *parameterized problem* is a set $Q \subseteq \Sigma^* \times \mathbb{N}$ for some finite alphabet Σ . A *compression* (also known as 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 in addition, $|x'| + k' \leq h(k)$ for some computable function h . The function h is called the *size* of the compression. If $|\Sigma| = 2$, the function h is called the *bit-size* of the compression. When we say that a parameterized problem Q admits a compression of size h , we mean that 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 .

For a fixed graph H , the computational H -COLORING problem asks to decide whether an input graph G is H -colorable. When parameterized by the vertex cover number, the input further includes a vertex cover X of G , whose size is the parameter of the problem. While

this definition is standard and convenient, we note that including the vertex cover X in the input is not essential, as X can be replaced by a vertex cover of G computed via an efficient 2-approximation algorithm.

3 Kernels for H -COLORING

In this section, we present and analyze our kernelization algorithms for the H -COLORING problem parameterized by the vertex cover number. We begin with a combinatorial kernel, whose size is governed by the non-adjacency witness number of H , and then proceed to an algebraic kernel, whose size is bounded in terms of the dimension of a faithful independent representation of H .

3.1 The Combinatorial Kernel

The non-adjacency witness number of a graph measures the smallest number of vertices needed to certify that a given set of vertices has no common neighbor. We formally define it below and study it in detail in the full version of the paper.

► **Definition 6.** *The non-adjacency witness number of a graph $G = (V, E)$, denoted $q(G)$, is the smallest positive integer q such that for every set $T \subseteq V$ of vertices with no common neighbor in G , there exists a set $T' \subseteq T$ of size $|T'| \leq q$ with no common neighbor in G .*

A main ingredient of our combinatorial kernel is the following lemma.

► **Lemma 7.** *Let $H = (V_H, E_H)$ be a graph, and let q be an integer satisfying $q \geq q(H)$. Consider a graph $G = (V, E)$ and a vertex cover $X \subseteq V$ of G of size k . Define $G' = (V', E')$ as the graph obtained from $G[X]$ by adding, for every non-empty set $S \subseteq X$ of size at most q such that $S \subseteq N_G(v)$ for some $v \in V \setminus X$, a vertex v_S adjacent to the vertices of S . Then the following holds.*

1. *The set X forms a vertex cover of G' .*
2. *The number of vertices in G' is at most $k + \sum_{i=1}^q \binom{k}{i}$.*
3. *The graph G' can be encoded in $\binom{k}{2} + \sum_{i=1}^q \binom{k}{i}$ bits.*
4. *The graph G is H -colorable if and only if the graph G' is H -colorable.*

Proof. By the definition of the graph G' , every edge of G' connects either two vertices of X or a vertex of X to a vertex v_S , hence X forms a vertex cover of G' . It also follows from the definition that the number of vertices in G' does not exceed $k + \sum_{i=1}^q \binom{k}{i}$. Additionally, G' can be represented by a binary string that expresses the adjacencies inside $G[X]$ as well as whether the vertex v_S is included in G' or not, for each non-empty set $S \subseteq X$ of size $|S| \leq q$. Therefore, the graph G' can be encoded in $\binom{k}{2} + \sum_{i=1}^q \binom{k}{i}$ bits.

It remains to prove the fourth item of the lemma, namely, that the graph G is H -colorable if and only if the graph G' is H -colorable. Suppose first that G is H -colorable, and let $f : V \rightarrow V_H$ be a homomorphism from G to H . Consider the function $f' : V' \rightarrow V_H$ defined as follows. First, for every vertex $u \in X$, let $f'(u) = f(u)$. Since every edge of $G'[X]$ is also an edge of $G[X]$, its endpoints are mapped by f' to adjacent vertices in H . Now, every vertex in $V' \setminus X$ has the form v_S for a set $S \subseteq X$ of size $|S| \leq q$, where there exists a vertex $v \in V \setminus X$ with $S \subseteq N_G(v)$. We define $f'(v_S) = f(v)$ for such a vertex v . Since f and f' agree on X , it follows that $f'(v_S)$ is adjacent in H to $f'(u)$ for all vertices $u \in S = N_{G'}(v_S)$. We thus obtain that f' is a homomorphism from G' to H , so G' is H -colorable.

For the converse direction, suppose that G' is H -colorable, and let $f' : V' \rightarrow V_H$ be a homomorphism from G' to H . Consider the function $f : V \rightarrow V_H$ defined as follows. First, for every vertex $u \in X$, let $f(u) = f'(u)$. Since every edge of $G[X]$ is also an edge of

$G'[X]$, its endpoints are mapped by f to adjacent vertices in H . Next, consider any vertex $u \in V \setminus X$. Since X is a vertex cover of G , all the neighbors of u lie in X . We claim that there exists a vertex in H that is adjacent to all vertices $f(v)$ with $v \in N_G(u)$. To see this, let $T = \{f(v) \mid v \in N_G(u)\}$, and suppose for contradiction that the vertices of T have no common neighbor in H . By the definition of the non-adjacency witness number, there exists a non-empty set $T' \subseteq T$ of size $|T'| \leq q(H) \leq q$ with no common neighbor in H . Notice that there also exists a non-empty set $S \subseteq N_G(u)$ of size $|S| \leq q$ such that $T' = \{f(v) \mid v \in S\}$. However, the graph G' includes the corresponding vertex v_S , which is mapped by f' to a vertex in H that is adjacent to all vertices of T' , a contradiction. This shows that the set T must admit a common neighbor in H , allowing us to define $f(u)$ as such a neighbor. It follows that f is a homomorphism from G to H , so G is H -colorable, as desired. ◀

We are ready to derive Theorem 1, which states that for every graph H , the H -COLORING problem parameterized by the vertex cover number k admits a kernel with $O(k^{q(H)})$ vertices and bit-size $O(k^{q(H)})$. Consequences for concrete graph classes appear in the full version of the paper.

Proof of Theorem 1. Let $H = (V_H, E_H)$ be a fixed graph, and set $q = q(H)$. If H is edgeless, then the H -COLORING problem is equivalent to determining whether an input graph is edgeless, and is thus solvable in polynomial time. Therefore, we may and will assume that H has at least one edge, which easily implies that $q \geq 2$.

The input of the H -COLORING problem parameterized by the vertex cover number k consists of a graph $G = (V, E)$ and a vertex cover $X \subseteq V$ of G of size $|X| = k$. Consider the kernelization algorithm that, given such an input, produces the graph $G' = (V', E')$ defined as in Lemma 7, that is, the graph obtained from $G[X]$ by adding, for every non-empty set $S \subseteq X$ of size $|S| \leq q$ such that $S \subseteq N_G(v)$ for some $v \in V \setminus X$, a vertex v_S adjacent to the vertices of S . Then, the algorithm returns the graph G' along with the set X .

The analysis of the algorithm is based on Lemma 7. By Item 1 of the lemma, the set X forms a vertex cover of G' , so the pair (G', X) produced by the algorithm is a valid output. By Item 2, the number of vertices in G' satisfies $|V'| \leq k + \sum_{i=1}^q \binom{k}{i} \leq O(k^q)$, and by Item 3, the number of bits needed to encode G' is at most $\binom{k}{2} + \sum_{i=1}^q \binom{k}{i}$, which is $O(k^q)$ since $q \geq 2$. Item 4 ensures the correctness of the kernelization algorithm. Finally, since q is a fixed constant, one can enumerate all subsets of X of size at most q and construct the graph G' in polynomial time. The proof is now complete. ◀

3.2 The Algebraic Kernel

Our algebraic kernel for the H -COLORING problem parameterized by the vertex cover number uses the notion of independent representations, a generalization of orthogonal representations of graphs.

3.2.1 Orthogonal and Independent Graph Representations

We begin with formal definitions of these two representation types.

► **Definition 8.** A d -dimensional orthogonal representation of a graph $G = (V, E)$ over a field \mathbb{F} is an assignment of a vector $x_v \in \mathbb{F}^d$ with $\langle x_v, x_v \rangle \neq 0$ to each vertex $v \in V$, such that for all vertices $u, v \in V$, if $\{u, v\} \in E$ then $\langle x_u, x_v \rangle = 0$. The orthogonal representation is called faithful if for all $u, v \in V$, it holds that $\{u, v\} \in E$ if and only if $\langle x_u, x_v \rangle = 0$.

► **Definition 9.** A d -dimensional independent representation of a graph $G = (V, E)$ over a field \mathbb{F} is an assignment of a vector $x_v \in \mathbb{F}^d$ to each vertex $v \in V$, such that for every vertex $v \in V$, it holds that $x_v \notin \text{span}(\{x_w \mid w \in N_G(v)\})$. The independent representation is called faithful if for all $u, v \in V$, it holds that $\{u, v\} \in E$ if and only if $x_u \in \text{span}(\{x_w \mid w \in N_G(v)\})$. Notice that the forward implication holds trivially.

We note that all vectors in an independent representation of a graph are nonzero. It can be verified that a graph has a 2-dimensional independent representation over a field if and only if it is bipartite.

► **Remark 10.** Every (faithful) orthogonal representation of a graph $G = (V, E)$ over a field \mathbb{F} is also a (faithful) independent representation of G over \mathbb{F} . To see this, consider an orthogonal representation $(x_v)_{v \in V}$ of G over \mathbb{F} . For each vertex $v \in V$, the vector x_v is orthogonal to the vectors x_w with $w \in N_G(v)$, and thus to the subspace they span, but not to itself, so $x_v \notin \text{span}(\{x_w \mid w \in N_G(v)\})$. This shows that $(x_v)_{v \in V}$ forms an independent representation of G . If the orthogonal representation $(x_v)_{v \in V}$ of G is faithful, then for all vertices $u, v \in V$ with $\{u, v\} \notin E$, the vector x_v is not orthogonal to x_u but is orthogonal to the vectors x_w with $w \in N_G(v)$, implying that $x_u \notin \text{span}(\{x_w \mid w \in N_G(v)\})$. This confirms that $(x_v)_{v \in V}$ forms a faithful independent representation of G in this case.

We proceed by presenting several lemmas that will be used in the analysis of the algebraic kernel. Their proofs are deferred to the full version of the paper. The first lemma shows that the dimension of any faithful independent representation of a graph forms an upper bound on its non-adjacency witness number (recall Definition 6).

► **Lemma 11.** For a graph G , a positive integer d , and a field \mathbb{F} , if G has a faithful d -dimensional independent representation over \mathbb{F} , then $q(G) \leq d$.

The following lemma states that any faithful independent representation over a field can be transformed into one of the same dimension in which every vector starts with 1, over any sufficiently large extension field.

► **Lemma 12.** Let $G = (V, E)$ be a graph, let d be a positive integer, and let $\mathbb{F} \subseteq \mathbb{K}$ be fields such that \mathbb{K} is an extension field of \mathbb{F} with $|\mathbb{K}| > |V|$. If G has a faithful d -dimensional independent representation over \mathbb{F} , then it also has a faithful d -dimensional independent representation over \mathbb{K} , in which all vectors have 1 as their first entry.

We finally observe that every graph admits a faithful independent representation of dimension one greater than its maximum degree over any sufficiently large field.

► **Lemma 13.** Let $G = (V, E)$ be a graph, let \mathbb{F} be a field with $|\mathbb{F}| \geq |V|$, and set $d = \Delta(G) + 1$. Then G has a faithful d -dimensional independent representation over \mathbb{F} .

3.2.2 The Algorithm

We are ready to present our algebraic kernel and confirm Theorem 2. To state the result in full generality, we borrow the following terminology from [25]. A field is said to be *efficient* if field operations and Gaussian elimination can be performed in polynomial time in the size of a reasonable input encoding. Note that all finite fields and the real field \mathbb{R} (restricted to rational numbers) are efficient.

► **Theorem 14.** For a graph H , a positive integer d , and an efficient field \mathbb{F} , suppose that there exists a faithful d -dimensional independent representation of H over \mathbb{F} . Then the H -COLORING problem parameterized by the vertex cover number k admits a kernel with $O(k^{d-1})$ vertices and bit-size $O(k^{d-1} \cdot \log k)$.

Proof. Consider a graph $H = (V_H, E_H)$, a positive integer d , and an efficient field \mathbb{F} , such that H has a faithful d -dimensional independent representation over \mathbb{F} . We may assume that $d \geq 3$, as otherwise H is bipartite, and the H -COLORING problem can be solved in polynomial time. We define a field \mathbb{K} as follows. If \mathbb{F} is finite, then we choose \mathbb{K} to be some finite extension field of \mathbb{F} whose order exceeds $|V_H|$ (e.g., a finite field of order $|\mathbb{F}|^\ell$ for the smallest positive integer ℓ satisfying $|\mathbb{F}|^\ell > |V_H|$). If \mathbb{F} is infinite, then \mathbb{K} is simply taken to be \mathbb{F} . In both cases, the field \mathbb{K} is efficient and satisfies $|\mathbb{K}| > |V_H|$.

The input of the H -COLORING problem parameterized by the vertex cover number k consists of a graph $G = (V, E)$ and a vertex cover $X \subseteq V$ of G of size $|X| = k$. Consider the kernelization algorithm that, given such an input, performs the following steps.

1. Produce the graph $G' = (V', E')$ defined as in Lemma 7 for $q = d$. Namely, G' is the graph obtained from $G[X]$ by adding, for every non-empty set $S \subseteq X$ of size $|S| \leq d$ such that $S \subseteq N_G(v)$ for some $v \in V \setminus X$, a vertex v_S adjacent to the vertices of S .
2. Associate with each vertex $u \in X$ a d -dimensional symbolic vector y_u over the field \mathbb{K} , where the first entry of y_u is fixed to 1 and the remaining $d - 1$ entries are free variables. Note that the total number of variables is $k \cdot (d - 1)$.
3. Consider the collection Y of all sets $S \subseteq X$ of size precisely d for which the vertex v_S is included in the graph G' , i.e., $Y = \{S \subseteq X \mid v_S \in V', |S| = d\}$. For each such set $S \in Y$, define a polynomial p_S over the variables of $(y_u)_{u \in X}$ as the determinant over the field \mathbb{K} of a $d \times d$ matrix whose columns are the vectors y_u with $u \in S$ (ordered arbitrarily).
4. Consider the subspace of polynomials $W = \text{span}(\{p_S \mid S \in Y\})$, and compute a set $Y' \subseteq Y$ such that the polynomials in $\{p_S \mid S \in Y'\}$ form a basis of W .
5. Let $G'' = (V'', E'')$ be the graph obtained from G' by removing all vertices v_S with $S \in Y \setminus Y'$, and return it along with the set X .

Let us emphasize that the given independent representation of H is not used by the algorithm itself, but it plays a crucial role in its correctness proof. Yet, the algorithm uses the field \mathbb{K} , over which it computes the polynomials p_S for $S \in Y$ and a basis of the subspace W .

We now analyze the algorithm and show that it satisfies the assertion of the theorem. First, by Item 1 of Lemma 7, the set X forms a vertex cover of G' , and so also forms a vertex cover of its subgraph G'' . Therefore, the pair (G'', X) produced by the algorithm is a valid output. The vertex set of G'' consists of the k vertices of X , vertices v_S with $|S| \leq d - 1$, whose number is at most $\sum_{i=1}^{d-1} \binom{k}{i}$, and the vertices v_S with $S \in Y'$. To bound the size of Y' , we observe that each polynomial p_S with $S \in Y$ is multilinear and homogeneous of degree $d - 1$. Indeed, the determinant polynomial of a $d \times d$ matrix is a linear combination of monomials of degree d , where each monomial is a product of exactly d entries, one from each row. Since p_S is defined as the determinant of a $d \times d$ matrix, whose first-row entries are all ones and all other entries are distinct variables, it follows that p_S has the required form. This implies that W is a subspace of the space of multilinear homogeneous polynomials of degree $d - 1$ in $k \cdot (d - 1)$ variables, whose dimension is $\binom{k \cdot (d-1)}{d-1}$. Since the size of a basis of such a subspace cannot exceed the dimension of the ambient space, it follows that $|Y'| \leq \binom{k \cdot (d-1)}{d-1}$. We obtain that the number of vertices in G'' satisfies

$$|V''| \leq k + \sum_{i=1}^{d-1} \binom{k}{i} + \binom{k \cdot (d-1)}{d-1} \leq O(k^{d-1}),$$

where the last inequality holds because d is a fixed constant. The graph G'' can be expressed by a binary string that represents the adjacencies in $G''[X]$ and the sets $S \subseteq X$ with $|S| \leq d$ for which the vertex v_S is included in G'' . Since each such set S can be encoded in $d \cdot \lceil \log k \rceil$ bits, the total number of bits needed to encode G'' does not exceed $\binom{k}{2} + |V'' \setminus X| \cdot d \cdot \lceil \log k \rceil$,

which is bounded by $O(k^{d-1} \cdot \log k)$, given that $d \geq 3$. We further notice that G'' can be constructed in polynomial time. Indeed, the construction of G' involves enumerating all subsets of X of size at most d . Then, to determine the vertices retained in G'' , we compute the set Y' by applying Gaussian elimination over the efficient field \mathbb{K} on a system with $\binom{k \cdot (d-1)}{d-1}$ variables, which can be performed in polynomial time.

We turn to proving the correctness of the algorithm, namely, that G is H -colorable if and only if G'' is H -colorable. By assumption, the graph H has a faithful d -dimensional independent representation over \mathbb{F} . Since \mathbb{K} is an extension field of \mathbb{F} with $|\mathbb{K}| > |V_H|$, Lemma 12 implies that H also has a faithful d -dimensional independent representation $(x_v)_{v \in V_H}$ over \mathbb{K} , in which all vectors have 1 as their first entry. By Lemma 11, we have $q(H) \leq d$, which allows us to apply Item 4 of Lemma 7 and conclude that the graph G is H -colorable if and only if the graph G' is H -colorable. It is thus sufficient to prove that G' is H -colorable if and only if G'' is H -colorable. Since G'' is a subgraph of G' , it is clear that if G' is H -colorable then so is G'' . It remains to show that if G'' is H -colorable, then so is G' .

Suppose that the graph G'' is H -colorable, and let $f'' : V'' \rightarrow V_H$ be a homomorphism from G'' to H . We define an assignment ρ of elements from \mathbb{K} to the variables of $(y_u)_{u \in X}$ by assigning to each variable vector y_u the vector $x_{f''(u)}$ from the given independent representation over \mathbb{K} . Note that this is possible because the first entry in y_u is 1, just as in $x_{f''(u)}$, while the other entries in y_u are free variables. We claim that for every set $S \in Y$, the polynomial p_S defined in Item 3 of the algorithm vanishes at the assignment ρ . It suffices to verify this only for sets $S \in Y'$, since the set $\{p_S \mid S \in Y'\}$ forms a basis of the subspace $W = \text{span}(\{p_S \mid S \in Y'\})$. Consider a set $S \in Y'$, and recall that the vertex v_S is adjacent in G'' to the vertices of S , hence $f''(v_S)$ is adjacent in H to all vertices $f''(u)$ with $u \in S$. By the definition of an independent representation, the vector $x_{f''(v_S)}$ does not lie in the linear subspace spanned by the vectors $x_{f''(u)}$ with $u \in S$. In particular, the dimension of this subspace is strictly smaller than d , so the determinant of a $d \times d$ matrix whose columns are the vectors $x_{f''(u)}$ with $u \in S$ is zero. This implies that p_S evaluates to zero at the assignment ρ , as desired.

We now show that there exists a homomorphism $f' : V' \rightarrow V_H$ from G' to H . We start by defining $f'(u) = f''(u)$ for each $u \in X$. Since every edge of $G'[X]$ is also an edge of $G''[X]$, its endpoints are mapped by f' to adjacent vertices in H . Next, for every vertex v_S in G' with $|S| \leq d-1$, set $f'(v_S) = f''(v_S)$. Since the neighborhood of such a vertex v_S is contained in X and is identical to its neighborhood in G'' , the vertex $f'(v_S)$ is adjacent in H to all the images under f' of the neighbors of v_S in G' . It remains to show that for every set $S \in Y$, the vertex v_S in G' can be assigned a vertex in H that is adjacent to all vertices $f'(u)$ with $u \in S$.

To see this, consider a set $S \in Y$, and recall that $|S| = d$ and that the polynomial p_S vanishes at the assignment ρ described above. Therefore, the vectors of $\{x_{f'(w)} \mid w \in S\}$ span a subspace of dimension smaller than d , so there exists a non-empty set $S' \subseteq S$ of size $|S'| \leq d-1$ such that the vectors of $\{x_{f'(w)} \mid w \in S'\}$ span the exact same subspace. Since the vertex v_S is included in G' , it follows that $S \subseteq N_{G'}(v)$ for some $v \in V \setminus X$, so by $S' \subseteq S$, the vertex $v_{S'}$ is included in G' as well as in G'' . We define $f'(v_S) = f''(v_{S'})$, and show that $f'(v_S)$ is adjacent in H to all vertices $f'(u)$ with $u \in S$. Indeed, for every vertex $u \in S$, it holds that

$$\begin{aligned} x_{f'(u)} \in \text{span}(\{x_{f'(w)} \mid w \in S\}) &= \text{span}(\{x_{f'(w)} \mid w \in S'\}) \\ &\subseteq \text{span}(\{x_v \mid v \in N_H(f''(v_{S'}))\}), \end{aligned} \quad (1)$$

where the containment holds because the vertices of S' are adjacent in G'' to $v_{S'}$, so they are mapped by f'' , and thus by f' , to neighbors of $f''(v_{S'})$ in H . By the faithfulness of the given independent representation of H , it follows from (1) that the vertices $f'(u)$ and $f'(v_S)$ are

adjacent in H . We conclude that $f''(v_{S'})$ is adjacent in H to all vertices $f'(u)$ with $u \in S$, as desired. This gives us a homomorphism from G' to H , implying that G' is H -colorable, as required. \blacktriangleleft

We end this section with the observation that Theorem 14 strengthens the kernelization bounds implied by the prior work [24]. Indeed, by Lemma 13, every graph H has a faithful independent representation of dimension $\Delta(H)+1$ over, for instance, the real field \mathbb{R} . Invoking Theorem 14 yields a kernel for the H -COLORING problem parameterized by the vertex cover number k with $O(k^{\Delta(H)})$ vertices and bit-size $O(k^{\Delta(H)} \cdot \log k)$. As illustrated in the full version of the paper, the kernel obtained via Theorem 14 is sometimes markedly more compact.

4 Concluding Remarks

This paper studies the kernelization complexity of H -COLORING problems parameterized by the vertex cover number and presents two kernels for them. The first is purely combinatorial, with size governed by the non-adjacency witness number of H , while the second is more algebraic in nature, relying on the existence of a low-dimensional faithful independent representation of H over some field. We have shown that the combinatorial kernel, despite its simplicity, achieves significant improvements over prior work for certain graphs H , and that the algebraic kernel can sometimes save a further linear factor in the kernel size. Moreover, we have provided lower bounds on the kernel size for projective graphs H and have demonstrated that in certain cases our kernels attain near-optimal bounds.

The natural challenge raised by this work is to determine, for general graphs H , the exact polynomial behavior of the kernel complexity of the H -COLORING problem parameterized by the vertex cover number. In particular, it would be valuable to obtain lower bounds on the kernel complexity of the problem for non-bipartite core graphs H that are not projective. For such graphs H , it would also be interesting to decide whether the H -COLORING problem admits no sub-quadratic compression when parameterized by the number of vertices. These are the remaining unresolved instances of a question raised in [3]. Finally, it would be of interest to explore the kernel complexity of H -COLORING problems under alternative parameterizations, a direction that has already seen several developments, e.g., in [23, 24, 18].

References

- 1 Ziv Bar-Yossef, Yitzhak Birk, T. S. Jayram, and Tomer Kol. Index coding with side information. *IEEE Trans. Inform. Theory*, 57(3):1479–1494, 2011. Preliminary version in FOCS'06. doi:10.1109/TIT.2010.2103753.
- 2 Peter J. Cameron, Ashley Montanaro, Michael W. Newman, Simone Severini, and Andreas J. Winter. On the quantum chromatic number of a graph. *Electron. J. Comb.*, 14(1):R81, 2007. doi:10.37236/999.
- 3 Hubie Chen, Bart M. P. Jansen, Karolina Okrasa, Astrid Pieterse, and Pawel Rzażewski. 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.
- 4 Bruno Codenotti, Pavel Pudlák, and Giovanni Resta. Some structural properties of low-rank matrices related to computational complexity. *Theor. Comput. Sci.*, 235(1):89–107, 2000. doi:10.1016/S0304-3975(99)00185-1.
- 5 Marek Cygan, Fedor V. Fomin, Alexander Golovnev, Alexander S. Kulikov, Ivan Mihajlin, Jakub Pachocki, and Arkadiusz Socala. Tight lower bounds on graph embedding problems. *J. ACM*, 64(3):18:1–22, 2017. Preliminary versions in ICALP'15 and SODA'16. doi:10.1145/3051094.

- 6 Mitre Costa Dourado, Fábio Protti, and Jayme Luiz Szwarcfiter. Computational aspects of the Helly property: A survey. *J. Braz. Comput. Soc.*, 12(1):7–33, 2006. doi:10.1007/BF03192385.
- 7 Tomás Feder, Pavol Hell, and Jing Huang. List homomorphisms and circular arc graphs. *Combinatorica*, 19(4):487–505, 1999. doi:10.1007/S004939970003.
- 8 Fedor V. Fomin, Pinar Heggenes, and Dieter Kratsch. Exact algorithms for graph homomorphisms. *Theory Comput. Syst.*, 41(2):381–393, 2007. Preliminary version in FCT’05. doi:10.1007/S00224-007-2007-X.
- 9 Fedor V. Fomin, Daniel Lokshtanov, Saket Saurabh, and Meirav Zehavi. *Kernelization: Theory of Parameterized Preprocessing*. Cambridge University Press, 2019.
- 10 Robert Ganian, Thekla Hamm, Viktoriia Korchemna, Karolina Okrasa, and Kirill Simonov. The fine-grained complexity of graph homomorphism parameterized by clique-width. *ACM Trans. Algorithms*, 20(3):19, 2024. Preliminary version in ICALP’22. doi:10.1145/3652514.
- 11 Alexander Golovnev and Ishay Haviv. The (generalized) orthogonality dimension of (generalized) Kneser graphs: Bounds and applications. *Theory of Computing*, 18(22):1–22, 2022. Preliminary version in CCC’21. doi:10.4086/TOC.2022.V018A022.
- 12 Alexander Golovnev, Oded Regev, and Omri Weinstein. The minrank of random graphs. *IEEE Trans. Inform. Theory*, 64(11):6990–6995, 2018. Preliminary version in RANDOM’17. doi:10.1109/TIT.2018.2810384.
- 13 Marina Groshaus, Min Chih Lin, and Jayme Luiz Szwarcfiter. On neighborhood-Helly graphs. *Discret. Appl. Math.*, 216(1):191–202, 2017. doi:10.1016/J.DAM.2016.04.029.
- 14 Marina Groshaus and Jayme Luiz Szwarcfiter. Biclique-Helly graphs. *Graphs Comb.*, 23(6):633–645, 2007. doi:10.1007/S00373-007-0756-6.
- 15 Willem H. Haemers. On some problems of Lovász concerning the Shannon capacity of a graph. *IEEE Trans. Inform. Theory*, 25(2):231–232, 1979. doi:10.1109/TIT.1979.1056027.
- 16 Ishay Haviv. On minrank and the Lovász theta function. In *International Conference on Approximation Algorithms for Combinatorial Optimization Problems (APPROX’18)*, pages 13:1–15, 2018. doi:10.4230/LIPICS.APPROX-RANDOM.2018.13.
- 17 Ishay Haviv. On minrank and forbidden subgraphs. *ACM Transactions on Computation Theory (TOCT)*, 11(4):20:1–13, 2019. Preliminary version in RANDOM’18. doi:10.1145/3322817.
- 18 Ishay Haviv and Dror Rabinovich. Kernelization for orthogonality dimension. In *Proc. of the 19th International Symposium on Parameterized and Exact Computation (IPEC’24)*, pages 8:1–17, 2024. doi:10.4230/LIPICS.IPEC.2024.8.
- 19 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.
- 20 Pavol Hell and Jaroslav Nešetřil. On the complexity of H -coloring. *J. Comb. Theory, Ser. B*, 48(1):92–110, 1990. doi:10.1016/0095-8956(90)90132-J.
- 21 Pavol Hell and Jaroslav Nešetřil. *Graphs and homomorphisms*, volume 28 of *Oxford lecture series in mathematics and its applications*. Oxford University Press, 2004.
- 22 Bart M. P. Jansen. The Power of Data Reduction: Kernels for Fundamental Graph Problems. Ph.D. Thesis, Utrecht University, 2013.
- 23 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.
- 24 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.
- 25 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.
- 26 Benoît Larose. Families of strongly projective graphs. *Discuss. Math. Graph Theory*, 22(2):271–292, 2002. doi:10.7151/DMGT.1175.
- 27 László Lovász. Kneser’s conjecture, chromatic number, and homotopy. *J. Comb. Theory, Ser. A*, 25(3):319–324, 1978. doi:10.1016/0097-3165(78)90022-5.

- 28 László Lovász. On the Shannon capacity of a graph. *IEEE Trans. Inform. Theory*, 25(1):1–7, 1979. doi:10.1109/TIT.1979.1055985.
- 29 László Lovász. *Graphs and Geometry*, volume 65. Colloquium Publications, 2019.
- 30 László Lovász, Michael Saks, and Alexander Schrijver. Orthogonal representations and connectivity of graphs. *Linear Algebra Appl.*, 114–115:439–454, 1989.
- 31 Hiroshi Maehara and Vojtech Rödl. On the dimension to represent a graph by a unit distance graph. *Graphs Comb.*, 6(4):365–367, 1990. doi:10.1007/BF01787703.
- 32 Hermann A. Maurer, Ivan Hal Sudborough, and Emo Welzl. On the complexity of the general coloring problem. *Inf. Control.*, 51(2):128–145, 1981. doi:10.1016/S0019-9958(81)90226-6.
- 33 Arya Mazumdar. Storage capacity of repairable networks. *IEEE Trans. Inform. Theory*, 61(11):5810–5821, 2015. Preliminary versions in ISIT’14 and Allerton’14. doi:10.1109/TIT.2015.2472521.
- 34 Jaroslav Nešetřil. Representations of graphs by means of products and their complexity. In *Proc. of the Mathematical Foundations of Computer Science (MFCS’81)*, pages 94–102, 1981.
- 35 Karolina Okrasa and Pawel Rzażewski. Fine-grained complexity of the graph homomorphism problem for bounded-treewidth graphs. *SIAM J. Comput.*, 50(2):487–508, 2021. Preliminary version in SODA’20. doi:10.1137/20M1320146.
- 36 René Peeters. Orthogonal representations over finite fields and the chromatic number of graphs. *Combinatorica*, 16(3):417–431, 1996. doi:10.1007/BF01261326.
- 37 Marta Piecyk and Pawel Rzażewski. Fine-grained complexity of the list homomorphism problem: Feedback vertex set and cutwidth. In *Proc. of the 38th International Symposium on Theoretical Aspects of Computer Science (STACS’21)*, pages 56:1–17, 2021.
- 38 Søren Riis. Information flows, graphs and their guessing numbers. *Electron. J. Comb.*, 14(1):R44, 2007. Preliminary version in NETCOD’06. doi:10.37236/962.
- 39 Pawel Rzażewski. Exact algorithm for graph homomorphism and locally injective graph homomorphism. *Inf. Process. Lett.*, 114(7):387–391, 2014. doi:10.1016/J.IPL.2014.02.012.
- 40 Karthikeyan Shanmugam, Alexandros G. Dimakis, and Michael Langberg. Local graph coloring and index coding. In *Proc. of the IEEE International Symposium on Information Theory (ISIT’13)*, pages 1152–1156, 2013. doi:10.1109/ISIT.2013.6620407.
- 41 Magnus Wahlström. New plain-exponential time classes for graph homomorphism. *Theory Comput. Syst.*, 49(2):273–282, 2011. Preliminary version in CSR’09. doi:10.1007/S00224-010-9261-Z.
- 42 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.