Treedepth Bounds in Linear Colorings

Jeremy Kun[1], Michael P. O'Brien[2(✉)], and Blair D. Sullivan[2]

[1] Google, Raleigh, USA
[2] North Carolina State University, Raleigh, USA
mpobrie3@ncsu.edu

Abstract. *Low-treedepth* colorings are an important tool for algorithms that exploit structure in classes of bounded expansion; they guarantee subgraphs that use few colors have *bounded treedepth*. These colorings have an implicit tradeoff between the total number of colors used and the treedepth bound, and prior empirical work suggests that the former dominates the run time of existing algorithms in practice. We introduce *p-linear colorings* as an alternative to the commonly used *p*-centered colorings. They can be efficiently computed in bounded expansion classes and use at most as many colors as *p*-centered colorings. Although a set of $k < p$ colors from a *p*-centered coloring induces a subgraph of treedepth at most k, the same number of colors from a *p*-linear coloring may induce subgraphs of larger treedepth. A simple analysis of this treedepth bound shows it cannot exceed 2^k, but no graph class is known to have treedepth more than $2k$. We establish polynomial upper bounds via constructive coloring algorithms in trees and intervals graphs, and conjecture that a polynomial relationship is in fact the worst case in general graphs. We also give a co-NP-completeness reduction for recognizing *p*-linear colorings and discuss ways to overcome this limitation in practice.

1 Introduction

Algorithms for graph classes that exhibit *bounded expansion* structure [2,9–11] offer a promising framework for efficiently solving many NP-hard problems on real-world networks. The structural restrictions of bounded expansion, which allow for pockets of localized density in globally sparse graphs, are compatible with properties of many real-world networks such as clustering and heavy-tailed degree distributions. Moreover, multiple random graph models designed to mimic these properties have been proven to asymptotically almost surely belong to classes of bounded expansion [3]. From a theoretical perspective, graphs belonging to classes of bounded expansion can be characterized by *low-treedepth colorings* of bounded size, i.e. using only a small number of colors. Roughly speaking, a low-treedepth coloring is one in which the subgraphs induced on each small set of colors have small *treedepth*, a structural property stronger than treewidth. This definition naturally implies an algorithmic pipeline [3,4,10] for classes of bounded expansion involving four stages: computing a low-treedepth coloring, using the coloring to decompose the graph into subgraphs of small treedepth,

solving the problem efficiently on each such subgraph, and combining the subsolutions to construct a global solution. The complexities of algorithms using this paradigm often are of the form $O(\binom{k}{p}2^{d\log d} \cdot n^c)$ where k is the coloring size and d is the treedepth of the subgraphs.

A recent implementation [12] and experimental evaluation [13] of this pipeline has identified that the coloring size has a much larger effect on the run time than the treedepth in practice. Although graphs in classes of bounded expansion are guaranteed to admit colorings of constant size with respect to the number of vertices, the only known polynomial-time algorithms for computing these colorings are approximations [2]. Consequently it is unclear to what extent our current coloring algorithms can be altered to reduce the coloring size. A more viable approach to improving the performance of the algorithmic pipeline without significant high-level changes would be to develop a new type of low-treedepth coloring that uses fewer colors but potentially has weaker guarantees about the treedepth of the subgraphs.

The traditional low-treedepth colorings for classes of bounded expansion are known as *p-centered colorings*. This name stems from the property that on any subgraph H, a p-centered coloring either uses at least p colors or is a *centered coloring*, which restricts the multiplicity of colors in induced subgraphs. In this paper we introduce an alternative that closely mirrors this paradigm but only extends the color multiplicity guarantees to path subgraphs. For this reason we refer to them as *p-linear colorings* and *linear colorings*. We identify that p-linear colorings share three important properties with p-centered colorings that allow them to be used in the bounded expansion algorithmic pipeline.

1. The minimum coloring size is constant in graphs of bounded expansion.
2. A coloring of bounded size can be computed in polynomial time.
3. Small sets of colors induce graphs of small treedepth.

The third of these properties is of particular interest, since understanding the tradeoffs between coloring size and treedepth in switching between p-centered and p-linear colorings fundamentally depends on bounding the maximum treedepth of a graph that admits a linear coloring with k colors. Equivalently, we frame this problem as determining the gap between the minimum number of colors needed for a linear versus a centered coloring in any given graph. A naïve analysis does not exclude the possibility that this gap is exponentially large, despite the fact that the largest known difference is a constant multiplicative factor. We conjecture that our proven constant factor lower bound is also the upper bound; as evidence, we prove that in trees and interval graphs the difference is polynomially bounded (in the coloring size) and give polynomial time algorithms (with respect to the graph size) to certify this difference. Surprisingly, we also prove that some p-linear colorings cannot be verified in polynomial time unless P = co-NP and discuss the practical implications of these findings. In the interest of space, the proofs of all lemmas are omitted from the main text and can be found in [7].

2   Definitions and Background

In this section we detail the background and terminology necessary to understand p-linear colorings.

2.1   Graph Terminology

We denote the vertices and edges of a graph G as $V(G)$ and $E(G)$, respectively, and assume all graphs are simple and undirected except where specifically noted otherwise. The *open neighborhood* of a vertex v, denoted $N(v)$, is the set of vertices u such that $uv \in E(G)$, while the *closed neighborhood*, $N[v]$ is defined as $N(v) \cup \{v\}$. Vertex a is an *apex* with respect to a subgraph H if $V(H) \subseteq N(a)$.

We say P is a $v_1 v_\ell$-*path* if $V(P) = \{v_1, \ldots, v_\ell\}$ and $E(P) = \{v_i v_{i+1} : 1 \leq i \leq \ell - 1\}$; we will notate this as $P = v_1, \ldots, v_\ell$. Given disjoint paths $P = v_1, \ldots, v_\ell$ and $Q = u_1, \ldots, u_{\ell'}$, the path $P \cdot Q = v_1, \ldots, v_\ell, u_1, \ldots, u_{\ell'}$ is the *concatenation* of P and Q if v_ℓ is adjacent to u_1. A path is *Hamiltonian* with respect to subgraph H if $V(P) = V(H)$.

In a rooted tree T, we let T_v be the subtree of T rooted at v and the *leaf paths* of T_v be the set of paths from a leaf of T_v to v. Vertices u and v are *unrelated* in T if u is neither an ancestor nor a descendant of v.

A *coloring* ϕ of a graph G is a mapping of the vertices of G to *colors* $1, \ldots, k$ and has *size* $|\phi| = k$. A coloring is *proper* if no pair of adjacent vertices have the same color. For any subgraph H and color c, if there is exactly one vertex $v \in H$ such that $\phi(v) = c$ we say c *appears uniquely* in H and v is a *center* of H. A subgraph with no unique color is said to be *non-centered*.

2.2 p-Centered Colorings and Bounded Expansion

Definition 1. *A p-centered coloring ϕ of graph G is a coloring such that for every connected subgraph H, H has a center or $\phi|_H$ uses at least p colors.*

Nešetřil and Ossana de Mendez established that bounding the minimum size of a p-centered coloring is a necessary and sufficient condition for a graph class to have bounded expansion.

Proposition 1 ([9])**.** *A class of graphs $\mathcal{C}$ has bounded expansion iff there exists a function f such that for all $G \in \mathcal{C}$ and all $p \geq 1$, G admits a p-centered coloring with $f(p)$ colors.*

There are varying methods to compute p-centered colorings, such as transitive-fraternal augmentations [5,9] and generalized coloring numbers [17], we focus here on *distance-truncated transitive-fraternal augmentations* (DTFAs) [14], which iteratively augment the graph with additional edges to impose constraints on proper colorings. This linear time algorithm guarantees that after $(2 \log p)^p$ DTFA iterations, any proper coloring of the augmented graph is a p-centered coloring whose size is bounded in classes of bounded expansion.

2.3   Centered Colorings and Treedepth

Note that if ϕ is a p-centered coloring of G and H is a subgraph of G whose vertices use at most $p-1$ colors in ϕ, H must have a center. This relates p-centered colorings to a more restricted class of graphs defined by *centered colorings*.

Definition 2. *A* centered coloring *ϕ of graph G is a coloring such that every connected subgraph has a center. The minimum size of a centered coloring of G is denoted $\chi_{cen}(G)$.*

Note that a centered coloring is also proper, or else there would be a connected subgraph of size two with no center. Observe that if X is the set of all centers of G, then $G \backslash X$ must either be empty or disconnected. This implies that if $|G| \gg \chi_{\mathrm{cen}}(G)$, then G breaks into many components after only a few vertex deletions. This property is captured by *treedepth decompositions*.

Definition 3. *A* treedepth decomposition *$\mathcal{T}$ of graph G is a rooted forest with the same vertex set as G such that $uv \in E(G)$ implies u is an ancestor of v in $\mathcal{T}$ or vice versa. The* depth *of $\mathcal{T}$ is the length of the longest path from a leaf in $\mathcal{T}$ to a root in its component. The* treedepth *of G,* $\mathrm{td}(G)$*, is the minimum depth of a treedepth decomposition of G.*

Given a centered coloring of size k, we can generate a treedepth decomposition of depth at most k by choosing any center v to be the root and setting the children of v to be the roots of the treedepth decompositions of the components of $G \backslash \{v\}$. Likewise, given a treedepth decomposition of depth k, we can generate a centered coloring using k colors by bijectively assigning the colors to levels of the tree and coloring vertices according to their level. We refer to the colorings and decompositions resulting from these procedures as *canonical*; together they imply that the treedepth and centered coloring numbers are equal for all graphs.

3 p-Linear and Linear Colorings

We introduce p-*linear colorings* as an alternative to p-centered colorings.

Definition 4. *A p-linear coloring is a coloring ψ of a graph G such that for every path[1] P, either P has a center or $\psi|_P$ uses at least p colors.*

It is proven in [14] that after performing 2^p DTFA iterations, any proper coloring of the augmented graph is a p-linear coloring. This implies that p-linear colorings indeed have constant size in bounded expansion classes and can be constructed in polynomial time (like p-centered colorings).

In the interest of maintaining consistency with prior terminology, we define *linear colorings* analogously to centered colorings.

---

[1] This includes non-induced paths.

Definition 5. *A* linear coloring *is a coloring ψ of a graph G such that every path has a center. The* linear coloring number *is the minimum number of colors needed for a linear coloring and is denoted $\chi_{lin}(G)$.*

Note that linear colorings must also be proper. A simple recursive argument shows that every path of length d requires at least $\log_2(d+1)$ colors in a linear coloring; thus a graph of linear coloring number k has no path of length 2^k. Because every depth-first search tree is a treedepth decomposition, $\mathrm{td}(G) \leq 2^{\chi_{\mathrm{lin}}(G)}$, proving that small numbers of colors in p-linear colorings induce graphs of bounded treedepth[2].

Our study of the divergence between linear and centered coloring numbers will naturally focus on linear colorings that are not also centered colorings. We say ψ is a *non-centered linear coloring* (NCLC) of graph G if G contains a connected induced subgraph with no center. For NCLC ψ, we say a connected induced subgraph H is a *witness* to ψ if H is non-centered but every proper connected subgraph of H has a center. For the sake of completeness, we prove in Lemma 1 that many simple graph classes do not admit NCLCs.

Lemma 1. *If G is a path, star, cycle, complete graph, or complete bipartite graph, any linear coloring of G is also a centered coloring.*

4   Treedepth Lower Bounds

To understand the tradeoff between the number of colors and treedepth of small color sets when using p-linear colorings in lieu of p-centered colorings, it is important to know the maximum treedepth of a graph of fixed linear coloring number k, $t_{\max}(k)$. In Lemmas 3 and 4, we prove lower bounds on $t_{\max}(k)$ through explicit constructions of graph families. In order to show that these graphs have large treedepth, we first establish assumptions about the structure of treedepth decompositions that can be made without loss of generality.

Lemma 2. *Let G be a graph and $S \subset V(G)$ such that $G[S]$ is connected and with respect to some component $C \in G \backslash S$, every vertex in S is an apex of C. Then for any treedepth decomposition $\mathcal{T}$ of depth k, we can construct a treedepth decomposition $\mathcal{T}'$ such that:*

1. $\mathrm{depth}(\mathcal{T}') \leq k$
2. *Each vertex in S is an ancestor of every vertex in C in $\mathcal{T}'$*
3. *For each pair of vertices $\{u, w\} \subseteq V(C)$ or $\{u, w\} \subseteq V(G \setminus C)$, u is an ancestor of w in $\mathcal{T}'$ iff it is an ancestor of w in $\mathcal{T}$.*

Using Lemma 2, we now show that $t_{\max}(k) \geq 2k$.

Lemma 3. *There exists an infinite sequence of graphs $R_1, R_2, \dots$ such that*

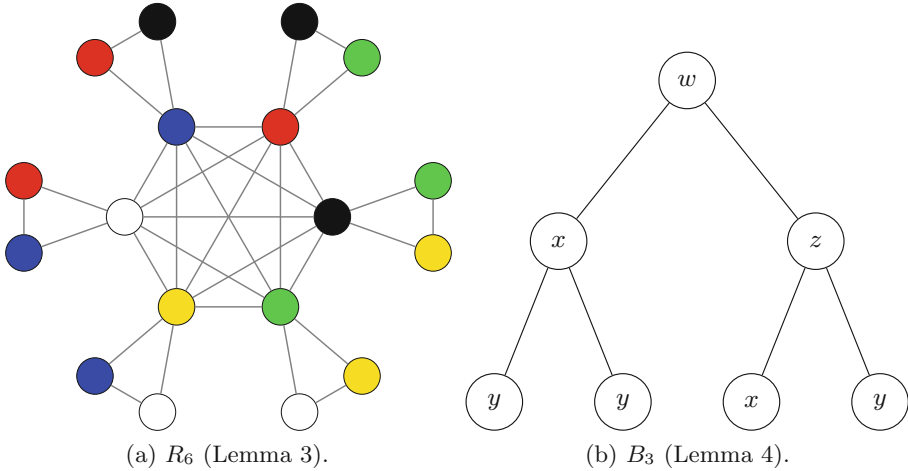$$\lim_{i \to \infty} \frac{\chi_{cen}(R_i)}{\chi_{lin}(R_i)} = 2.$$

---

[2] This tightens a bound in [14] from double to single exponential.

The graphs in Lemma 3 contain large cliques. We now show that this is not a necessary condition for the linear and centered coloring numbers to diverge (Fig. 1).

Lemma 4. *Let B_ℓ be the complete binary tree with ℓ levels. Then*

$$\lim_{\ell \to \infty} \frac{\chi_{cen}(B_\ell)}{\chi_{lin}(B_\ell)} \geq \frac{3}{2}$$

(a) R_6 (Lemma 3).   (b) B_3 (Lemma 4).

Fig. 1. Linear colorings of graphs in Lemmas 3–4. (Color figure online)

We conjecture that the construction in Lemma 3 is optimal.

Conjecture 1. For any graph G, $\chi_{\mathrm{cen}}(G) \leq 2\chi_{\mathrm{lin}}(G)$.

While the exclusion of a path of length 2^k indicates $t_{\max}(k) \leq 2^k$, this nonetheless leaves a large gap between the upper and lower bounds on $t_{\max}(k)$. To move towards a proof of Conjecture 1, we consider two restricted graph classes—namely, trees and interval graphs—in the next two sections and establish polynomial upper bounds on $t_{\max}(k)$ for graphs in these classes.

5   Treedepth Upper Bounds on Trees

Schäffer proved that there is a linear time algorithm for finding a minimum-sized centered coloring of a tree T [16]. In this section we prove the following theorem by showing a correspondence between the centered coloring from Schäffer's algorithm and colors on paths in any linear coloring of T.

Theorem 1. *There exists a polynomial time algorithm that takes as input a tree T and a linear coloring ψ of T with size k and outputs a centered coloring of T whose size is at most $O(k^3)$.*

Schäffer's algorithm finds a particular centered coloring whose colors are ordered in a way that reflects their roles as centers. For this reason, the coloring is called a *vertex ranking* and the colors are referred to as *ranks*; it guarantees that in each subgraph, the vertex of maximum rank is also a center. We will use this terminology in this section to clearly distinguish between the ranks in the vertex ranking and colors in the linear coloring. Note that the canonical centered coloring of a treedepth decomposition is a vertex ranking if the colors are ranked decreasing from the root downwards, which implies that every centered coloring can be converted to a vertex ranking of the same size. Of central importance to Schäffer's algorithm are what we will refer to as *rank lists*.

Definition 6. *For a vertex ranking r of tree T, the* rank list *of T, denoted $L(T)$, can be defined recursively as $L(T) = L(T \backslash T_v) \cup \{r(v)\}$ where v is the vertex of maximum rank in T.*

Schäffer's algorithm arbitrarily roots T and builds the ranking from the leaves to the root of T, computing the rank of each vertex from the rank lists of each of its children.

Proposition 2 ([16])**.** *Let r be a vertex ranking of T produced by Schäffer's algorithm and let $v \in T$ be a vertex with children $u_1, \ldots, u_\ell$. If x is the largest integer appearing on rank lists of at least two children of v (or 0 if all such rank lists are pairwise disjoint) then $r(v)$ is the smallest integer satisfying $r(v) > x$ and $r(v) \notin \bigcup_{i=1}^{\ell} L(u_i)$.*

Our proof of Theorem 1 is based on tracking sets of colors of ψ on leaf paths as Schäffer's algorithm moves up the rooted tree. Define the *color vector* of a path P with respect to linear coloring ψ to be the set of colors from ψ appearing on P. Let $S(v)$ be the set of all color vectors of all leaf paths in T_v. Let $\kappa(v)$ be the maximum cardinality of any color vector in $S(v)$ and $S_\kappa(v) = \{C \in S(v) : |C| = \kappa(v)\}$. We show below that every vertex v has a corresponding vertex u that is "similar" in rank but "dissimilar" in values of κ and/or S_κ.

Lemma 5. *Let v be a vertex of rank $i > k$. There exists a vertex $u \in T_v$ such that*

- $i - \kappa(v) - 1 < r(u) < i$ *and*
- *Either $\kappa(u) < \kappa(v)$ or $|S_\kappa(u)| \leq \lfloor \frac{1}{2}|S_\kappa(v)| \rfloor$.*

Using Lemma 5 as a recursive step, we prove Theorem 1 by tracing the values of functions κ and S down towards the leaves.

Proof (Theorem 1). Let u_i be the vertex of maximum rank in T. There is a maximal sequence of vertices $u_2, \ldots, u_q$ such that u_{i+1} satisfies the properties of

Lemma 5 with respect to T_{u_i}. Note that the ranks of $u_1, \ldots, u_q$ are monotonically decreasing and $r(u_i) - r(u_{i+1}) \leq \kappa(u_i)$. Moreover, every vertex in T satisfies $1 \leq |S_\kappa(v)| \leq \binom{k}{\kappa(v)}$ and $1 \leq \kappa(v) \leq k$. Since the only vertices with $\kappa(v) = 1$ are the leaves,

$$r(u_i) \leq \sum_{i=1}^{k} i \left(\log_2 \binom{k}{i} + 1 \right) \leq O(k^3).$$

Consequently, r is a centered coloring of size at most $O(k^3)$ that can be computed in linear time.   □

6 Treedepth Upper Bounds on Interval Graphs

Because linear colorings are equivalent to centered colorings when restricted to paths, we turn our attention to the linear coloring numbers of "pathlike" graphs. We investigate a particular class of "pathlike" graphs in this section and prove a polynomial relationship between their centered and linear coloring numbers.

Definition 7. *A graph G is an* interval graph *if there is an injective mapping f from $V(G)$ to intervals on the real line such that $uv \in E(G)$ iff $f(u)$ and $f(v)$ overlap.*

We refer to the mapping f as the *interval representation* of G. Since the overlap between intervals $f(u)$ and $f(v)$ is independent of the interval representations of the other vertices, every subgraph of an interval graph is also an interval graph. The interval representation of G implies a natural "left-to-right" layout that gives it the "pathlike" qualities, which are manifested in restrictions on the length of induced cycles (*chordal*) and paths between vertex triples (*AT-free*).

Definition 8. *A graph is* chordal *if it contains no induced cycles of length ≥ 4.*

Definition 9. *Vertices u, v, w are an* asteroidal triple *(AT) if there exist uv-, vw-, and wu-paths P_{uv}, P_{vw}, and P_{wu}, respectively, such that $N[w] \cap P_{uv} = N[u] \cap P_{vw} = N[v] \cap P_{uv} = \emptyset$. A graph with no AT is called* AT-free.

Proposition 3 ([8])**.** *Graph G is an interval graph iff G is chordal and AT-free.*

Intuitively, Definition 9 is a set of three vertices such that every pair is connected by a path that avoids the neighbors of the third. Roughly speaking, in the context of linear colorings, Proposition 3 indicates that if w is a center of a "long" uv-path P in G, any vertex w' such that $\psi(w) = \psi(w')$ must have a neighbor on P. We devote the rest of this section proving Theorem 2.

Theorem 2. *There exists a polynomial time algorithm that takes as input an interval graph G and a linear coloring of G with size k and outputs a centered coloring of G with size at most k^2.*

Our algorithm makes extensive use of the following well-known property of maximal cliques in interval graphs.

Proposition 4 ([8])**.** *If G is an interval graph, its maximal cliques can be linearly ordered in polynomial time such that for each vertex v, the cliques containing v appear consecutively.*

In particular, we identify a *prevailing path* in G whose vertices "span" the maximal cliques and a *prevailing subgraph* that consists of the prevailing path as well as vertices in maximal cliques "between" consecutive vertices on the prevailing path. We will show that any linear coloring is a centered coloring when restricted to the prevailing subgraph and that after removing the prevailing subgraph, the remaining components each use fewer colors.

Let $C_1, \ldots C_m$ be an ordering of the maximal cliques of G that satisfies Proposition 4. We say vertex v is *introduced* in C_i if $v \in C_i$ but $v \notin C_{i-1}$, and denote this as $I(v) = i$. Likewise, v is *forgotten* in C_j if $v \in C_j$ but $v \notin C_{j+1}$, and denote this as $F(v) = j$. The procedure for constructing a prevailing subgraph and prevailing path is described in Algorithm 1. This algorithm selects the vertex v from the current maximal clique that is forgotten "last" and adds v to the prevailing path and $C_{F(v)}$ to the prevailing subgraph. We prove in Lemma 6 that if P, Q are a prevailing path and subgraph, the vertices in $Q \backslash P$ can be inserted between vertices of P to form a Hamiltonian path of Q.

Algorithm 1. Construction of a prevailing path and subgraph.

Input: interval graph G
Output: prevailing path P and prevailing subgraph Q
1: $C_1, \ldots, C_m \leftarrow$ maximal cliques of G labeled in accordance with Proposition 4
2: $P \leftarrow \emptyset$
3: $V_Q \leftarrow \emptyset$
4: $i \leftarrow 1$
5: $j \leftarrow 1$
6: **while** $i < m$ **do**
7:    $v_j \leftarrow \arg \max_{u \in C_i} F(u)$
8:    $P \leftarrow P \cdot \{v_j\}$
9:    $i \leftarrow F(v)$
10:    $V_Q \leftarrow V_Q \cup V(C_i)$
11:    $j \leftarrow j + 1$
12: **end while**
13: $Q \leftarrow G[V_Q]$
14: **return** P, Q

Lemma 6. *Every prevailing subgraph has a Hamiltonian path.*

Although the fact that the prevailing subgraph Q has a Hamiltonian path implies Q has a center with respect to ψ, we must ensure that the proper subgraphs of Q also have a center. In Lemma 7, we prove $\psi|_Q$ is centered by showing every proper connected subgraph of Q also has a Hamiltonian path.

Lemma 7. *If Q is a prevailing subgraph of an interval graph G and ψ a linear coloring of G, $\psi|_Q$ is a centered coloring.*

Since any linear coloring ψ of the prevailing subgraph Q must also be a centered coloring, $\mathrm{td}(Q) \leq |\psi|$. To get a bound on the treedepth of G, we focus on the relationship between Q and $G \backslash Q$. In particular, we show that the components of $G \backslash Q$ use fewer than $|\psi|$ colors by proving that each such component has an apex in the prevailing path.

Lemma 8. *Let P, Q be a prevailing path and subgraph of an interval graph G. For each component X of $G \backslash Q$, there is a vertex $a \in P$ such that $X \subseteq N(a)$.*

We can now establish a polynomial upper bound on the treedepth of interval graphs, proving Theorem 2.

Proof (Theorem 2). Let $\mathcal{A}$ be the algorithm that constructs a treedepth decomposition $\mathcal{T}$ of G by finding a prevailing subgraph Q (Algorithm 1), using $\psi|_Q$ to create a treedepth decomposition of Q, and recursively constructing treedepth decompositions of $G \backslash Q$. If $\mathrm{depth}(\mathcal{T}) \leq k^2$ and $\mathcal{A}$ runs in polynomial time, then the canonical centered coloring of $\mathcal{T}$ is a centered coloring of G of size at most k^2. We prove $\mathcal{A}$ satisfies these requirements by induction on $k = |\psi|$. At $k = 1$, the graph consists of isolated vertices and $\mathcal{A}$ trivially constructs a treedepth decomposition of G of depth 1 in polynomial time.

Assume $\mathcal{A}$ has the desired properties for linear colorings of size at most $k-1$. Because the maximal cliques of an interval graph can be enumerated and ordered in polynomial time (Proposition 4), identifying Q via Algorithm 1 can be done in polynomial time. By Lemma 7, the canonical treedepth decomposition of Q has depth at most k. Since every component X of $G \backslash Q$ has an apex a in P (Lemma 8), we can assume a is an ancestor in $\mathcal{T}$ of each vertex in X (Lemma 2). Because ψ is proper, $\psi(a)$ does not appear in $\psi|_X$ and since induced subgraphs of interval graphs are themselves interval graphs, $\mathcal{A}$ finds a treedepth decomposition of X whose depth is at most $(k-1)^2$. Thus $\mathcal{T}$ has depth $k + (k-1)^2 \leq k^2$. The recursion only lasts $k \leq n$ steps, so $\mathcal{A}$ runs in polynomial time.   □

7   Hardness of Recognizing Linear Colorings

Based on the similarity in definition between linear and centered colorings, one might assume that computing them should be roughly equally difficult. Finding a centered coloring of a fixed size is NP-hard [1], but given a coloring of a graph, we can *recognize* whether it is centered in polynomial time by attempting to create the canonical treedepth decomposition; this procedure will identify a non-centered subgraph if the coloring is not centered. To the contrary, we will prove that LINEAR COLORING RECOGNITION, the problem of recognizing whether a coloring is linear, is co-NP-complete. In order to prove the hardness of LINEAR COLORING RECOGNITION, we first define a dual problem. The NON-CENTERED PATH problem takes a graph G and coloring ψ as input and decides whether G
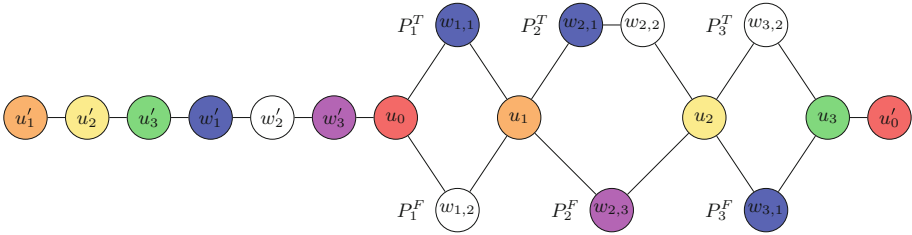
Fig. 2. The graph G and coloring ψ for $\Phi = (x_1 \vee x_2 \vee \neg x_3) \wedge (\neg x_1 \vee x_2 \vee x_3) \wedge (\neg x_2)$. (Color figure online)

has a non-centered path P. We focus on proving the hardness of NON-CENTERED PATH because a certificate to that problem is easily definable: a path where every color appears at least twice (Fig. 2).

Theorem 3. NON-CENTERED PATH *is NP-complete.*

Corollary 1. LINEAR COLORING RECOGNITION *is co-NP-complete.*

The co-NP-hardness of recognizing linear colorings is compounded by three stronger hardness implications. First, the coloring ψ given in Theorem 3 has size $m+n+1$, which means that unless the exponential time hypothesis [6] fails, there is no $2^{o(k)}$ algorithm to recognize a linear coloring of size k. Second, the graph G is outerplanar with pathwidth two, which implies that neither treewidth-style dynamic programming nor a Baker-style layering approach is likely to solve this problem efficiently. Finally, by subdividing each edge and coloring all subdivision vertices with a (single) new color, we obtain a bipartite graph with degeneracy two, proving hardness for each of those classes. Nonetheless, the fact that $\chi_{\mathrm{cen}}(G) = O(\log m + \log n)$ while $|\psi| = m + n + 1$ leaves open the possibility that LINEAR COLORING RECOGNITION becomes easier for colorings of minimum size.

8   Conclusion

We have introduced p-linear and linear colorings as an alternative to p-centered and centered colorings for use in algorithms for classes of bounded expansion. The p-linear colorings are computable in polynomial time and require a constant number of colors in classes of bounded expansion, while inducing graphs of bounded treedepth for all small sets of colors, allowing direct substitution in existing algorithmic pipelines. A major direction for future work is to bring the upper bound on $t_{\max}(k)$ of 2^k closer to the lower bound of $2k$. In particular, it appears our current toolkit for analyzing linear colorings must be expanded in order to prove (or disprove) Conjecture 1. We also believe it is worth studying whether recognizing linear colorings can be done in polynomial time if we assume the coloring is of size $\chi_{\mathrm{lin}}(G)$. Finally, using p-linear colorings

in practice will require an efficient method for translating a linear coloring into a treedepth decomposition. Although there exist general-purpose algorithms to find treedepth decompositions efficiently in graphs of bounded linear coloring number (e.g. [15]), a more specialized algorithm that avoids "heavy machinery" is likely necessary to be practically useful.

References

1. Bodlaender, H.L.: Rankings of graphs. SIAM J. Discrete Math. **11**(1), 168–181 (1998)
2. Nešetřil, J., Ossona de Mendez, P.: Sparsity. Algorithms and Combinatorics, vol. 28. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-27875-4
3. Demaine, E.,D., et al.: Structural sparsity of complex networks: random graph models and linear algorithms. CoRR, abs/1406.2587 (2015)
4. Dvořák, Z., Král, D., Thomas, R.: Testing first-order properties for subclasses of sparse graphs. JACM **60**(5), 36:1–36:24 (2013)
5. Grohe, M., Kreutzer, S., Siebertz, S.: Deciding first-order properties of nowhere dense graphs. J. ACM (JACM) **64**(3), 17 (2017)
6. Impagliazzo, R., Paturi, R.: On the complexity of k-SAT. J. Comput. Syst. Sci. **62**(2), 367–375 (2001)
7. Kun, J., O'Brien, M.P., Sullivan, B.D.: Treedepth bounds in linear colorings. CoRR, abs/1802.09665, May 2018
8. Lekkeikerker, C., Boland, J.: Representation of a finite graph by a set of intervals on the real line. Fundam. Math. **51**(1), 45–64 (1962)
9. Nešettil, J., de Mendez, P.O.: Grad and classes with bounded expansion I. Decompositions. Eur. J. Comb. **29**(3), 760–776 (2008)
10. Nešettil, J., de Mendez, P.O.: Grad and classes with bounded expansion II. Algorithmic aspects. Eur. J. Comb. **29**(3), 777–791 (2008)
11. Nešettil, J., de Mendez, P.O.: Grad and classes with bounded expansion III. Restricted graph homomorphism dualities. Eur. J. Comb. **29**(4), 1012–1024 (2008). Homomorphisms: structure and highlights
12. O'Brien, M.P., Hobbs, C.G., Jasnick, K., Reidl, F., Rodrigues, N.G., Sullivan, B.D.: CONCUSS, v2.0., June 2016. https://doi.org/10.5281/zenodo.30281
13. O'Brien, M.P., Sullivan, B.D.: An experimental evaluation of a bounded expansion algorithmic pipeline. CoRR, abs/1712.06690, December 2017
14. Reidl, F.: Structural sparseness and complex networks. Dr. Aachen, Techn. Hochsch., Aachen, 2015. Aachen, Techn. Hochsch., Dissertation (2015)
15. Reidl, F., Rossmanith, P., Villaamil, F.S., Sikdar, S.: A faster parameterized algorithm for treedepth. In: Esparza, J., Fraigniaud, P., Husfeldt, T., Koutsoupias, E. (eds.) ICALP 2014. LNCS, vol. 8572, pp. 931–942. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-662-43948-7_77

16. Schäffer, A.A.: Optimal node ranking of trees in linear time. Inf. Process. Lett. **33**(2), 91–96 (1989)
17. Zhu, X.: Colouring graphs with bounded generalized colouring number. Discrete Math. **309**(18), 5562–5568 (2009)