Rainbow Vertex Coloring Bipartite Graphs and Chordal Graphs

Pinar Heggernes
Department of Informatics, University of Bergen, Norway
pinar.heggernes@uib.no

Davis Issac
Max Planck Institute for Informatics, Saarland Informatics Campus, Germany
dissac@mpi-inf.mpg.de

Juho Lauri
Nokia Bell Labs, Dublin, Ireland
juho.lauri@nokia-bell-labs.com

Paloma T. Lima
Department of Informatics, University of Bergen, Norway
paloma.lima@uib.no

Erik Jan van Leeuwen
Department of Information and Computing Sciences, Utrecht University, The Netherlands
e.j.vanleeuwen@uu.nl

Abstract

Given a graph with colors on its vertices, a path is called a rainbow vertex path if all its internal vertices have distinct colors. We say that the graph is rainbow vertex-connected if there is a rainbow vertex path between every pair of its vertices. We study the problem of deciding whether the vertices of a given graph can be colored with at most k colors so that the graph becomes rainbow vertex-connected. Although edge-colorings have been studied extensively under similar constraints, there are significantly fewer results on the vertex variant that we consider. In particular, its complexity on structured graph classes was explicitly posed as an open question.

We show that the problem remains NP-complete even on bipartite apex graphs and on split graphs. The former can be seen as a first step in the direction of studying the complexity of rainbow coloring on sparse graphs, an open problem which has attracted attention but limited progress. We also give hardness of approximation results for both bipartite and split graphs. To complement the negative results, we show that bipartite permutation graphs, interval graphs, and block graphs can be rainbow vertex-connected optimally in polynomial time.

2012 ACM Subject Classification Mathematics of computing → Graph theory

Keywords and phrases Rainbow coloring, graph classes, polynomial-time algorithms, approximation algorithms

Digital Object Identifier 10.4230/LIPIcs.MFCS.2018.83

1 Introduction

Graph coloring and graph connectivity are two of the most famous topics in graph algorithms. Many different types of colorings and connectivity measures have been considered throughout time. The concept of rainbow coloring brings these two extensively studied topics together, and it was first defined a decade ago by Chartrand et al. [8] using edge-colorings. Let G be a connected, edge-colored graph. A *rainbow path* in G is a path all of whose edges are colored with distinct colors, and G is *rainbow-connected* if there is a rainbow path between every pair of its vertices. The resulting computational problem Rainbow Coloring (RC) takes as input a connected (uncolored) graph G and an integer k, and the task is to decide whether the edges of G can be colored with at most k colors so that G is rainbow-connected. This problem has various applications in telecommunications, data transfer, and encryption [25, 4, 11] and has been studied rather thoroughly from both graph-theoretic and complexity-theoretic viewpoints (see related work below and the surveys [19, 25]).

The intense interest in Rainbow Coloring led Krivelevich and Yuster [18] to define a natural variant on *vertex-colored* graphs. Here, a path in a vertex-colored graph H is a *rainbow vertex path* if all its *internal* vertices have distinct colors. We say that H is *rainbow vertex-connected* if there is a rainbow vertex path between every pair of its vertices. Similarly to the edge variant, Rainbow Vertex Coloring (RVC) is the decision problem in which we are given a connected (uncolored) graph H and an integer k, and the task is to decide whether the vertices of H can be colored with at most k colors such that H is rainbow vertex-connected. The *rainbow vertex connection number* of G, denoted by $\mathbf{rvc}(G)$, is the minimum k such that G has a rainbow vertex coloring with k colors. RVC is NP-complete for every $k \geq 2$ [10, 9], and remains NP-complete for $k = 3$ for bipartite graphs [23]. In addition, it is NP-hard to approximate $\mathbf{rvc}(G)$ within a factor of $2 - \varepsilon$ unless $\mathsf{P} \neq \mathsf{NP}$, for any $\varepsilon > 0$ [13]. It is also known that RVC is linear-time solvable on planar graphs for every fixed k [19]. Finally, assuming the Exponential Time Hypothesis, there is no algorithm for solving RVC in time $2^{o(n^{3/2})}$ for any $k \geq 2$ [19].

A stronger variant of rainbow vertex-colorings was introduced by Li et al. [24]. A vertex-colored graph H is *strongly rainbow vertex-connected* if between every pair of vertices of H, there is a *shortest* path that is also a rainbow vertex path. The Strong Rainbow Vertex Coloring (SRVC) problem takes as input a connected (uncolored) graph H and an integer k, and the task is to decide whether the vertices of H can be colored such that H is strongly rainbow vertex-connected. This definition is the vertex variant of the Strong Rainbow Coloring problem, which was also broadly studied (see related work below and the surveys [19, 25]). The *strong rainbow vertex connection number* of G, denoted by $\mathbf{srvc}(G)$, is the minimum k such that G has a strong rainbow vertex coloring with k colors. SRVC is NP-complete for every $k \geq 2$ [12] and linear-time solvable on planar graphs for every fixed k [19]. In addition, it is NP-hard to approximate $\mathbf{srvc}(G)$ within a factor of $n^{1/2-\varepsilon}$ unless $\mathsf{P} \neq \mathsf{NP}$, for any $\varepsilon > 0$ [13].

While RC has been widely studied in more than 300 published papers, we are unaware of any further complexity results on RVC and SRVC than those mentioned previously. In particular, the complexity of RVC and SRVC on structured graph classes is mostly open. This led Lauri [19, Open problem 6.6] to explicitly ask the following:

For what restricted graph classes do RVC *and* SRVC *remain NP-complete?*

Our Results. In this paper, we make significant progress towards addressing this open problem. In particular, we study bipartite graphs and chordal graphs, and some of their subclasses, and give hardness results and polynomial-time algorithms for RVC and SRVC. Our main result is a hardness result for bipartite apex graphs:

▶ **Theorem 1.** *Let G be a bipartite apex graph of diameter 4. It is NP-complete to decide both whether $\mathbf{rvc}(G) \leq 4$ and whether $\mathbf{srvc}(G) \leq 4$. Moreover, it is NP-hard to approximate $\mathbf{rvc}(G)$ and $\mathbf{srvc}(G)$ within a factor of $5/4 - \varepsilon$, for every $\varepsilon > 0$.*

This result is particularly interesting since no hardness result was known on a sparse graph class (like apex graphs) for any of the variants of rainbow coloring. Moreover, this result can be considered tight in conjunction with the known result that RVC and SRVC are linear-time solvable on planar graphs for every fixed number of colors k [19]. Finally, we observe (like Li et al. [23]) that $\mathbf{rvc}(G)$ and $\mathbf{srvc}(G)$ can be computed in linear time if G is a bipartite graph of diameter 3, providing further evidence that this result is tight.

For general bipartite graphs and for split graphs (a well-known subclass of chordal graphs), we exhibit stronger hardness results:

▶ **Theorem 2.** *Let G be a bipartite graph of diameter 4. It is* NP*-complete to decide both whether $\mathbf{rvc}(G) \leq k$ and whether $\mathbf{srvc}(G) \leq k$, for every $k \geq 3$. Moreover, it is* NP*-hard to approximate both $\mathbf{rvc}(G)$ and $\mathbf{srvc}(G)$ within a factor of $n^{1/3-\varepsilon}$, for every $\varepsilon > 0$.*

We remark that, previously, it was only known that deciding whether $\mathbf{rvc}(G) \leq 3$ for bipartite graphs G is NP-complete by the result of [23]. Our construction, however, is conceptually simpler, gives hardness for every $k \geq 3$, and is easily extended to the strong variant. Moreover, for RVC on general graphs, this result implies a considerable improvement over the previous result of Eiben et al. [13] which only excluded a polynomial-time approximation with a factor of less than 2 assuming $\mathsf{P} \neq \mathsf{NP}$.

▶ **Theorem 3.** *Let G be a split graph of diameter 3. It is* NP*-complete to decide both whether $\mathbf{rvc}(G) \leq k$ and whether $\mathbf{srvc}(G) \leq k$, for every $k \geq 2$. Moreover, it is* NP*-hard to approximate both $\mathbf{rvc}(G)$ and $\mathbf{srvc}(G)$ within a factor of $n^{1/3-\varepsilon}$, for every $\varepsilon > 0$.*

To the best of our knowledge, our results for split graphs give the first non-trivial graph class besides diameter-two graphs for which the complexity of the edge and the vertex variant differ (see e.g, [19, Table 4.2] but note that it contains a typo erroneously claiming that RVC can be solved in polynomial-time for split graphs). In particular, RC can be solved in polynomial time on split graphs when $k \geq 4$ [5, 7]. Moreover, we observe that $\mathbf{rvc}(G)$ and $\mathbf{srvc}(G)$ can be computed in linear time if G is a graph of diameter 2, providing evidence that this result is tight.

To contrast our hardness results, we show that both problems can be solved in polynomial time on several other subclasses of bipartite graphs and chordal graphs.

▶ **Theorem 4.** *If G is a bipartite permutation graph, a block graph, or a unit interval graph, then $\mathbf{rvc}(G)$ and $\mathbf{srvc}(G)$ can be computed in linear time. If G is an interval graph, then $\mathbf{rvc}(G)$ can be computed in linear time.*

Combined, these results paint a much clearer picture of the complexity landscape of RVC and SRVC than was possible previously.

Related Work. We briefly survey the known work for the edge variants of rainbow coloring; we refer to [19, 25] for more detailed surveys. RC is NP-complete for every $k \geq 2$ [4, 2, 22], even on chordal graphs [5]. On split graphs, RC is NP-complete when $k \in \{2, 3\}$, but solvable in polynomial time otherwise [5, 7]. It is also solvable in polynomial time on threshold graphs [5]. On bridgeless chordal graphs, there is a linear-time (3/2)-approximation algorithm for RC, however the problem cannot be approximated with a factor less than 5/4 on this graph class, unless $\mathsf{P} = \mathsf{NP}$ [6]. Some lower bounds on algorithms for solving RC are given by Kowalik et al. [17] and Agrawal [1] under the Exponential Time Hypothesis.

For the strong edge variant, an edge-colored graph is said to be *strongly rainbow-connected* if there is a rainbow shortest path between every pair of its vertices. The problem of deciding whether the edges of a given graph G can be colored in k colors to make G strongly rainbow-connected is referred to as SRC. For $k = 2$, it is not difficult to verify that RC is equivalent to SRC. Not surprisingly, SRC is also NP-complete for $k \geq 2$ [2]. In contrast to RC, SRC remains hard on split graphs for every $k \geq 2$ [19, Theorem 4.1]. Moreover, on n-vertex split graphs, it is NP-hard to approximate SRC within a factor of $n^{1/2-\varepsilon}$ for any $\varepsilon > 0$, while RC admits an additive-1 approximation [5]. The former statement also holds for n-vertex bipartite graphs instead of split graphs [2]. For block graphs, computing SRC can be done in linear time [16], while RC on block graphs is conjectured to be hard (see [19, Conjecture 6.3] or [16]). In general, it appears that despite the interest, there are fewer complexity-theoretic results on SRC. In fact, the same is true when considering combinatorial results (see [25] for a broader discussion).

2 Preliminaries

In this paper, we work on undirected simple graphs. Such a graph is denoted by $G = (V, E)$, where V is the vertex set of G, and E is the edge set. We let n denote the number of vertices of G. For a vertex $x \in V$, $N(x)$ is the set of its *neighbors*, and $\deg(x) = |N(x)|$ is its *degree*. For a $S \subseteq V$, the subgraph of G induced by S is denoted by $G[S]$. A *cut vertex* of G is a vertex whose removal increases the number of connected components of G.

Given a path $P = x_1, x_2, \dots, x_{p-1}, x_p$ in G, the vertices from x_2 to x_{p-1} are called the *internal vertices* of P. The *distance* between two vertices u and v in G, denoted by $\mathrm{dist}(u, v)$, is the length of a shortest path between u and v. The *diameter* of G, denoted by $\mathrm{diam}(G)$, is the maximum distance between any pair of vertices of G.

A *k-coloring of G* is a function $c : V \to \{1, 2, \dots, k\}$. (From now on, we will denote a set of consecutive integers from 1 to k as $[k]$.) A *coloring* is simply a k-coloring for some $k \leq n$. A coloring c is *proper* if $c(u) \neq c(v)$ for every edge $uv \in E$. The *chromatic number* of G, denoted by $\chi(G)$, is the smallest k such that G has a proper k-coloring. A *d-distance coloring* of G is a coloring c of G such that $c(u) \neq c(v)$ whenever $\mathrm{dist}(u, v) \leq d$. The minimum number of colors needed for a d-distance coloring of G is known as the *d-distance chromatic number* of G, and it is denoted by $\chi_d(G)$. Note that $\chi_d(G)$ is equivalent to $\chi(G^d)$, i.e., the chromatic number of the d^{th} power of G.

Since, in this paper, we will only be working on the vertex variant of the rainbow coloring and rainbow connectivity, we might sometimes omit the word "vertex" when there is no confusion. The parameter $\mathbf{srvc}(G)$ was defined by Li et al. [24], and they also verified that $\mathrm{diam}(G) - 1 \leq \mathbf{rvc}(G) \leq \mathbf{srvc}(G) \leq n - 2$. The following upper bound was mentioned in [19] (see the same reference for further discussion and examples).

▶ **Proposition 5** ([19]). *Let G be a connected graph with $\mathrm{diam}(G) = d \geq 3$. Then*

$$d - 1 \leq \mathbf{rvc}(G) \leq \mathbf{srvc}(G) \leq \chi_{d-2}(G).$$

Proof. There are at least two vertices in G connected by a shortest path of length d. Clearly, every coloring must use at least $d - 1$ colors to rainbow-connect this pair. On the other hand, between every pair of vertices u and v, there is a path of length at most d, meaning that it contains at most $d - 1$ internal vertices. As every $(d - 2)$-distance coloring colors these internal vertices distinctly, the statement follows. ◀

A *dominating set* of G is a set $D \subseteq V$ such that every vertex in $V \setminus D$ is adjacent to at least one vertex in D. If $G[D]$ is connected, then D is a *connected dominating set*. The minimum size of a connected dominating set in G, denoted by $\gamma_c(G)$, is known as the *connected domination number* of G. This parameter provides an upper bound on the rainbow vertex connection number of a connected graph, since G becomes rainbow vertex-connected by simply coloring all vertices of the connected dominating set distinctly, and the remaining vertices with any of the already used colors. This observation can be derived from [18].

▶ **Proposition 6** ([18]). *If G is a connected graph, then $\mathbf{rvc}(G) \leq \gamma_c(G)$.*

2.1 Graph classes

As we will be studying the mentioned problems on some graph classes, let us give a brief definition of these classes here. More definitions and properties will be added as needed when we handle these graphs. A detailed background on these graph classes can be found, for example, in the book by Brandstädt, Le, and Spinrad [3].

A graph is an *apex graph* if it contains a vertex (called an *apex*) whose removal results in a planar graph. A graph is *chordal* if all of its induced simple cycles are of length 3. Some well-known subclasses of chordal graphs are interval graphs, split graphs, and block graphs. A graph is an *interval graph* if it is chordal and it contains no triple of non-adjacent vertices, such that there is a path between every two of them that does not contain a neighbor of the third. A graph is a *split graph* if its vertex set can be partitioned into an independent set and a clique. A graph is a *block graph* if every biconnected component (block) of G is a complete graph.

Let σ be a permutation of the integers between 1 and n. We can make a graph G_σ on vertex set $[n]$ in the following way. Vertices i and j are adjacent in G_σ if and only if they appear in σ in the opposite order of their natural order. A graph on n vertices is a *permutation graph* if it is isomorphic to G_σ for some permutation σ of the integers between 1 and n. A graph is a *bipartite permutation* graph if it is both a bipartite graph and a permutation graph.

2.2 Hypergraph coloring

For our hardness reductions we will use a well-known NP-complete problem called Hypergraph Coloring. A *hypergraph* $H = (N, \mathcal{E})$ with vertex set N and hyperedge set $\mathcal{E}$ is a generalization of a graph, in which edges can contain more than two vertices. Thus $\mathcal{E}$ consists of subsets of N of arbitrary size. The definition of a (vertex) coloring of a hypergraph is exactly that same as that of a graph. In a colored hypergraph, an edge is called *monochromatic* if all of its vertices received the same color. A *proper coloring* of a hypergraph generalizes a proper coloring of a graph in a natural way: we require that no hyperedge is monochromatic. To avoid trivial cases, we can assume from now on that every hyperedge contains at least two vertices. Thus a proper coloring must always use at least two colors.

The Hypergraph Coloring problem takes as input a hypergraph H and an integer k and asks whether there is a proper coloring of H with at most k colors. The problem is well-known to be NP-complete for every $k \geq 2$ [26]. The Graph Coloring problem takes as input an undirected graph G and asks to determine the smallest k such that G has a proper k-coloring. This problem is NP-hard to approximate within a factor of $n^{1-\varepsilon}$ for any $\varepsilon > 0$, where n is the number of vertices [30]. Finally, the Planar 3-Coloring problem takes as input a planar graph G and asks whether G has a proper 3-coloring. This problem is NP-complete [14].

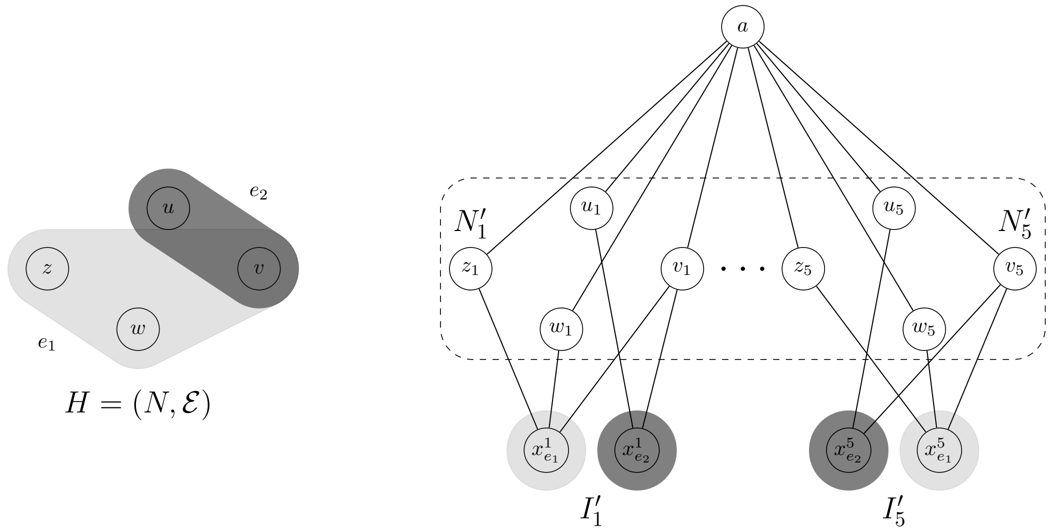

Figure 1 A hypergraph $H = (N, \mathcal{E})$ (left) transformed into a bipartite graph G (right) as described in the proof of Lemma 8. The dashed rectangle with rounded corners contains the sets in N'.

3 Bipartite graphs and their subclasses

In this section, we show that RVC and SRVC are hard on bipartite graphs for $k \geq 3$. We complement these results by showing that both problems can be solved in linear time on bipartite permutation graphs. We first observe that computing $\mathbf{rvc}(G)$ or $\mathbf{srvc}(G)$ is easy on bipartite graphs of diameter 3. The same observation was made by Li et al. [23].

▶ **Proposition 7** ([23]). *If G is a bipartite graph with $\operatorname{diam}(G) = 3$, then $\mathbf{rvc}(G) = \mathbf{srvc}(G) = 2$. Moreover, such a coloring can be found in linear time.*

Proof. The statement follows from Proposition 5 and the fact that every bipartite graph has a proper 2-coloring that can be found in linear time. ◀

It turns out that if $\operatorname{diam}(G) \geq 4$, then $\mathbf{rvc}(G)$ and $\mathbf{srvc}(G)$ of a bipartite graph G become much harder to compute, as claimed in Theorem 2. We prove the following general construction.

▶ **Lemma 8.** *Let H be a hypergraph on n vertices. Then in polynomial time we can construct a bipartite graph G of diameter 4 and with $O(n^3)$ vertices such that for any $k \in [n]$, H has a proper k-coloring if and only if G has a $(k+1)$-coloring under which G is (strongly) rainbow vertex-connected. Moreover, if H is a planar graph, then G is an apex graph.*

Proof. Let $H = (N, \mathcal{E})$ be an arbitrary hypergraph and let $n = |N|$. We construct a bipartite graph $G = (\{a\} \cup N' \cup I',\ E)$ where $N' = N'_1 \cup \cdots \cup N'_{n+1}$, $I' = I'_1 \cup \cdots \cup I'_{n+1}$, $N'_i := \{v_i \mid v \in N\}$, $I'_i := \{x^i_e \mid e \in \mathcal{E}\}$ and $E := \{av_i \mid v \in N, i \in [n+1]\} \cup \{v_i x^i_e \mid v \in N, e \in \mathcal{E}, i \in [n+1], v \in e\}$. Let $V = \{a\} \cup N' \cup I'$. A bipartition of G is given by $(\{a\} \cup I',\ N')$. Observe that $\operatorname{diam}(G) = 4$ and that G has $O(n^3)$ vertices. Moreover, if H is a planar graph, then G consists of vertex a plus $n + 1$ copies of the graph obtained from H by subdividing each edge of H, and thus G is an apex graph. For an illustration of the construction, see Figure 1.

Consider any proper k-coloring $h : N \to [k]$ of H, i.e., no hyperedge of H is monochromatic under h. We construct a coloring $c : V \to [k+1]$ in the following way. First, for every $v \in N$, we give the vertices $v_1, v_2, \ldots, v_n$ of G the same color as v, i.e., $c(v_i) = h(v)$ for all $v \in N$ and $i \in [n+1]$. We give vertex a the color $k+1$, i.e., $c(a) = k+1$. The vertices in I all receive the same color, which is any arbitrary color in $[k+1]$. Now we prove that G is strongly rainbow vertex-connected under c by showing that there is a rainbow vertex shortest path between every pair of vertices. The only non-trivial case is when both vertices of the pair are in I. Consider two distinct vertices $x_e^i, x_f^j \in I$ (it is possible that $e = f$ or $i = j$ but not both). Since e and f are not monochromatic under h, we can pick two distinct vertices $u \in e$ and $v \in f$ such that $h(u) \neq h(v)$. It is clear that the path $x_e^i uav x_f^j$ is a shortest path between x_e^i and x_f^j and that it is a rainbow vertex path. Hence, G is strongly rainbow vertex-connected under c.

Conversely, let c be a $(k+1)$-coloring of G under which G is (strongly) rainbow vertex-connected. For each $i \in [n+1]$, define h_i to be the vertex coloring of H such that $h_i(v) = c(v_i)$ for all $v \in N$. Let M_i be the set of vertices $v \in N$ such that $h_i(v) \neq c(a)$. Let $h_i'(v) = h_i(v)$ if $v \in M_i$ and $h_i'(v) = 1$ otherwise. We claim that there exists an $i \in [n+1]$ such that h_i' is a proper k-coloring of H. For the sake of contradiction, suppose that h_i' is not a proper k-coloring of H for every $i \in [n+1]$. For each $i \in [n+1]$, let $e_i \in \mathcal{E}$ be a monochromatic edge under h_i'. Suppose that, for some $i \in [n+1]$, all vertices in e_i are colored $c(a)$ under c. Then any path from $x_{e_i}^i$ to $x_{e_i}^j$ for some $j \neq i$ uses two vertices having color $c(a)$ under c. Hence, c would not be a rainbow vertex coloring, a contradiction. Therefore, for each $i \in [n+1]$, there is a vertex $v_i \in e_i$ for which $c(v_i) \neq c(a)$. Suppose now that for every $i \in [n+1]$, all vertices in e_i are colored either $c(v_i)$ or $c(a)$ under c. If $c(v_i) = c(v_j)$ for $i \neq j$, then any path from $x_{e_i}^i$ to $x_{e_j}^j$ uses either two vertices having color $c(a)$ or two vertices having color $c(v_i) = c(v_j)$ under c. This would contradict the assumption that G is rainbow vertex-connected under c. Hence, $c(v_i) \neq c(v_j)$ for all distinct $i, j \in [n+1]$. This implies that c uses at least $n+2$ colors, a contradiction to the assumptions that c is a $(k+1)$-coloring of G and that $k \in [n]$. Therefore, for some $i \in [n+1]$, there is a vertex $v_i' \in e_i$ for which $c(v_i') \neq c(a)$ and $c(v_i') \neq c(v_i)$. The latter implies that e_i is not monochromatic under h_i', a contradiction. The claim follows, and thus H has a proper k-coloring. ◀

Proof of Theorem 2. For membership in NP, a certificate that $\mathbf{rvc}(G) \leq k$ ($\mathbf{srvc}(G) \leq k$) consists of a k-coloring and a list of (shortest) paths, one for every pair of non-adjacent vertices, that are rainbow vertex connected. For NP-hardness, we observe that the transformation of Lemma 8 implies a straightforward reduction from Hypergraph Coloring. Since Hypergraph Coloring is NP-complete for each $k \geq 2$, this proves the first part of the theorem.

For the second part of the theorem, we consider an instance of Graph Coloring that consists of a graph on ℓ vertices and apply Lemma 8. Note that the total number of vertices in G is $n = O(\ell^3)$. From the hardness of approximation of Graph Coloring, we know that for all $\varepsilon > 0$, it is NP-hard to distinguish between the case when H is properly colorable with ℓ^ε colors and the case when H is not properly colorable with fewer than $\ell^{1-\varepsilon}$ colors [30]. By Lemma 8, this implies that it is NP-hard to distinguish between the case when G is (strong) rainbow vertex colorable with $\ell^\varepsilon + 1 \leq n^\varepsilon + 1$ colors and the case when G is not (strong) rainbow vertex colorable with fewer than $\ell^{1-\varepsilon} + 1 = \Omega(n^{1/3-\varepsilon})$ colors. The second statement of the theorem follows. ◀

We then proceed to give a proof of Theorem 1. This result can be considered as a first step to understand rainbow coloring on sparse graphs classes.

Proof of Theorem 1. The proof follows along the same lines as the proof of the first part of Theorem 2. Instead of Hypergraph Coloring, however, we reduce from Planar 3-Coloring, the problem of deciding whether a planar graph has a proper 3-coloring. This problem is NP-complete. The statement follows from Lemma 8, because the graph resulting from the construction is a bipartite apex graph of diameter 4.

For the hardness of approximation, we recall that any planar graph has a proper 4-coloring, and thus the graph G constructed in Lemma 8 has a 5-coloring under which G is rainbow vertex-connected. Hence, Lemma 8 combined with the NP-hardness of Planar 3-Coloring makes it NP-hard to decide whether G has a 5-coloring or a 4-coloring under which G is rainbow vertex-connected. ◀

We now complement the above hardness results with a positive result in the case when a bipartite graph is also a permutation graph, as claimed in Theorem 4. Bipartite permutation graphs have a desirable property, related to breadth-first search (BFS), that we will use heavily in our next result. Let us first define a chain graph. A bipartite graph is a *chain graph* if the vertices of the two independent sets A and B can be ordered as $\{a_1, a_2, \ldots, a_k\}$ and $\{b_1, b_2, \ldots, b_\ell\}$, such that $N(a_1) \subseteq N(a_2) \subseteq \cdots \subseteq N(A_k)$, equivalently, $N(b_\ell) \subseteq N(b_{\ell-1}) \subseteq \cdots \subseteq N(b_1)$.

In every bipartite permutation graph G it is possible to find a vertex v such that the levels $L_0, L_1, L_2, \ldots$ of the tree resulting from a BFS starting from v have the following properties. For all i, $L_0 = \{v\}$, L_i is an independent set and $G[L_i \cup L_{i+1}]$ is a chain graph. Moreover, for each level i, there exists a special vertex $a_i \in L_i$ such that $L_{i+1} \subset N(a_i)$. The vertex v can be picked as the first vertex of a *strong ordering*. It has been shown by Spinrad et al. [27] that a bipartite graph is a permutation graph if and only if it has a strong ordering, and such an ordering can be computed in linear time. The properties of the BFS tree above are well-known and easy to deduce from a strong ordering [29].

▶ **Theorem 9.** *If G is a bipartite permutation graph, then* $\mathbf{rvc}(G) = \mathbf{srvc}(G) = \mathrm{diam}(G) - 1$, *and the corresponding (strong) rainbow vertex coloring can be found in time that is linear in the size of G.*

Proof. Let $G = (V, E)$ be a bipartite permutation graph. Let v be a first vertex in a strong ordering for G. We start by doing a BFS on G with v as the root. Let k be the number of levels in the BFS tree in addition to level 0. Hence, L_i is the set of vertices in level i of the BFS tree, $0 \leq i \leq k$, with $L_0 = \{v\}$. Since $\mathrm{dist}(v, y) = k$ for every $y \in L_k$, we conclude that $\mathrm{diam}(G) \geq k$. Furthermore, if $\mathrm{dist}(x, y) > k - 1$ for some $x \in L_1$ and some $y \in L_k$, then we can conclude that $\mathrm{dist}(x, y) = k + 1$, where $x, v, a_1, a_2, \ldots, a_{k-1}, y$ is a shortest path between x and y. In this case, $\mathrm{diam}(G) = k + 1$. We distinguish between these two cases:

Case 1. $\mathrm{diam}(G) = k$.

We construct a strong rainbow vertex coloring $c : V \to [k-1]$ for G in the following way. If $x \in L_i$, we define $c(x) = i$, for $1 \leq i \leq k-1$. We define $c(v) = k - 1$, and we give arbitrary colors between 1 and $k - 1$ to the vertices of L_k. To see that G is indeed rainbow-connected under c, consider any pair $x, y \in V$. If $xy \in E$ or if they are in the same level of the BFS tree, there is nothing to prove, since $\mathrm{dist}(x, y) \leq 2$. Otherwise, we have exactly the following cases:

1. $x = v$ and $y \in L_j$: Then the path $v, a_1, \ldots, a_{j-1}, y$ is shortest and it is rainbow.
2. $x \in L_1$ and $y \in L_k$: In this case, $\mathrm{dist}(x, y) = k - 1$. Otherwise, since each L_i is an independent set, we would have $\mathrm{dist}(x, y) \geq k + 1$, which contradicts our assumption that

diam$(G) = k$. Since dist$(x, y) = k - 1$, every shortest path between x and y is rainbow, as every vertex of such a shortest path has to be in a distinct level of the BFS tree.

3. $x \in L_1$ and $y \in L_j$ with $2 \leq j \leq k - 1$: If dist$(x, y) = j - 1$, then again by the same argument used above, every shortest path between x and y is rainbow. If dist$(x, y) > j - 1$, then dist$(x, y) = j + 1$, and the shortest path $x, v, a_1, \ldots, a_{j-1}, y$ has distinct colors on all its internal vertices. (Note that y might have the same color as v if $j = k - 1$, but this is fine since y is the end of the path.)
4. $x \in L_i$ and $y \in L_j$ with $2 \leq i < j \leq k$: If dist$(x, y) = j - i$, then every shortest path is rainbow. If dist$(x, y) > j - i$, then the path $x, a_{i-1}, a_i, \ldots, a_{j-1}, y$ is rainbow and has length $j - i + 1$, and it is therefore shortest.

Case 2. diam$(G) = k + 1$.

We construct a strong rainbow vertex coloring $c : V \to [k]$ for G in the following way. If $x \in L_i$, we define $c(x) = i$, for $1 \leq i \leq k - 1$. We define $c(v) = k$, and we give arbitrary colors between 1 and k to the vertices of L_k. To see that G is indeed rainbow-connected under c, consider any pair $x, y \in V$. Again, if $xy \in E$ or if they are in the same level of the BFS, there is nothing to prove, since dist$(x, y) \leq 2$. Otherwise, there is only one remaining case:

- $x \in L_i$ and $y \in L_j$, with $0 \leq i < j \leq k$: If dist$(x, y) = j - i$ then every shortest path between x and y is rainbow. Otherwise, the path $x, a_{i-1}, a_i, \ldots, a_{j-1}, y$ is rainbow and has length $j - i + 2$, therefore being shortest.

In both cases, c is a strong rainbow vertex coloring for G with diam$(G) - 1$ colors. By Proposition 5 we can conclude that $\mathbf{rvc}(G) = \mathbf{srvc}(G) = \text{diam}(G) - 1$. ◀

4 Chordal graphs and their subclasses

In this section, we investigate the complexity of RVC and SRVC on chordal graphs and some subclasses of chordal graphs. We start by proving that both problems are NP-complete when the input graph is a split graph, implying that they are also NP-complete on chordal graphs. On the positive side, we show that RVC is polynomial-time solvable on interval graphs, and both RVC and SRVC are polynomial-time solvable on block graphs and on unit interval graphs.

We start by observing that computing $\mathbf{rvc}(G)$ or $\mathbf{srvc}(G)$ is easy on graphs of diameter 2.

▶ **Proposition 10** ([18]). *If G is a graph with* diam$(G) = 2$, *then* $\mathbf{rvc}(G) = \mathbf{srvc}(G) = 1$. *Moreover, such a coloring can be found in linear time.*

Proof. Color each vertex of G with the same color. Since each shortest path between two vertices contains at most one internal vertex, G is strongly rainbow vertex-connected under this coloring. ◀

If G is a split graph of diam$(G) = 3$ (note that split graphs have diameter at most 3), then $\mathbf{rvc}(G)$ and $\mathbf{srvc}(G)$ become much harder to compute, as claimed in Theorem 3. We prove the following general construction, which closely mimics the construction of Lemma 8.

▶ **Lemma 11.** *Let H be a hypergraph on n vertices. Then in polynomial time we can construct a split graph G of diameter 3 and with $O(n^3)$ vertices such that for any $k \in [n]$, H has a proper k-coloring if and only if G has a k-coloring under which G is (strongly) rainbow vertex-connected.*

Proof. Let $H = (N, \mathcal{E})$ be an arbitrary hypergraph and let $n = |N|$. We construct a split graph $G = (N' \cup I',\ E)$ where $N' = N'_1 \cup \cdots \cup N'_{n+1}$, $I' = I'_1 \cup \cdots \cup I'_{n+1}$, $N'_i := \{v_i \mid v \in N\}$, $I'_i := \{x^i_e \mid e \in \mathcal{E}\}$ and $E := \{u_i v_j \mid u, v \in N,\ i, j \in [n+1]\} \cup \{v_i x^i_e \mid v \in N, e \in \mathcal{E}, i \in [n+1], v \in e\}$. Let $V = N' \cup I'$. The constructed graph G is a split graph since $G[I']$ is an independent set and $G[N']$ is a clique. Observe that $\operatorname{diam}(G) = 3$ and that G has $O(n^3)$ vertices. The construction is illustrated in Figure 1: note that since $G[N']$ is a clique, all possible edges now appear between the vertices inside the rectangle with rounded corners.

Consider any proper k-coloring $h : N \to [k]$ of H, i.e., no hyperedge of H is monochromatic under h. We construct a coloring $c : V \to [k]$ in the following way. First, for every $v \in N$, we give the vertices $v_1, v_2, \ldots, v_n$ of G the same color as v, i.e., $c(v_i) = h(v)$ for all $v \in N$ and $i \in [n+1]$. The vertices in I all receive the same color, which is any arbitrary color in $[k]$. Now, we prove that G is strongly rainbow vertex-connected under c by showing that there is a rainbow vertex shortest path between every pair of vertices. The only non-trivial case is when both vertices of the pair are in I. Consider two distinct vertices $x^i_e, x^j_f \in I$ (it is possible that $e = f$ or $i = j$ but not both). Since e and f are not monochromatic under h, we can pick two distinct vertices $u \in e$ and $v \in f$ such that $h(u) \neq h(v)$. It is clear that the path $x^i_e uv x^j_f$ is a shortest path between x^i_e and x^j_f and that it is rainbow vertex path.

Conversely, let c be a k-coloring of G under which G is (strongly) rainbow vertex-connected. For each $i \in [n+1]$, define h_i to be the vertex coloring of H such that $h_i(v) = c(v_i)$ for all $v \in N$. We claim that there exists an $i \in [n+1]$ such that h'_i is a proper k-coloring of H. For the sake of contradiction, suppose that h'_i is not a proper k-coloring of H for every $i \in [n+1]$. For each $i \in [n+1]$, let $e_i \in \mathcal{E}$ be a monochromatic edge under h'_i. Let v_i be an arbitrary vertex in e_i. Suppose now that for every $i \in [n+1]$, all vertices in e_i are colored $c(v_i)$ under c. If $c(v_i) = c(v_j)$ for $i \neq j$, then any path from $x^i_{e_i}$ to $x^j_{e_j}$ uses two vertices having color $c(v_i) = c(v_j)$ under c. This would contradict the assumption that G is rainbow vertex-connected under c. Hence, $c(v_i) \neq c(v_j)$ for all distinct $i, j \in [n+1]$. This implies that c uses at least $n+1$ colors, a contradiction to the assumptions that c is a k-coloring of G and $k \in [n]$. Therefore, for some $i \in [n+1]$, there is a vertex $v'_i \in e_i$ for which $c(v'_i) \neq c(a)$ and $c(v'_i) \neq c(v_i)$. The latter implies that e_i is not monochromatic under h'_i, a contradiction. The claim follows, and thus H has a proper H-coloring. ◀

Proof of Theorem 3. The proof follows in exactly the same way as Theorem 2, except that we apply Lemma 11 instead of Lemma 8. ◀

We now move on to the positive results. As a consequence of the following theorems, we complete the proof of Theorem 4.

▶ **Theorem 12.** *Let G be a block graph, and let ℓ be the number of cut vertices in G. Then $\mathbf{rvc}(G) = \mathbf{srvc}(G) = \ell$. The corresponding (strong) rainbow vertex coloring can be found in time that is linear in the size of G.*

Proof. Let $G = (V, E)$ be a block graph and $\{a_1, a_2, \ldots, a_\ell\}$ be the set of cut vertices of G. We construct a strong rainbow vertex coloring $c : V \to [\ell]$ for G by defining $c(a_i) = i$ for $i \in [\ell]$ and giving the other vertices arbitrary colors between 1 and ℓ. An important property of block graphs is that there is a unique shortest path between every pair of vertices. Moreover, each internal vertex of such a path is a cut vertex. Since all the cut vertices received distinct colors, these shortest paths are all rainbow. The proof follows by observing that $\mathbf{rvc}(G) \geq \mathbf{srvc}(G) \geq \ell$ as well. ◀

For our next result, we need to mention that every interval graph has a representation called an *interval model*. Let $\mathcal{I}$ be a set of n intervals of the real line. Then we can define a graph $G_{\mathcal{I}}$ with a vertex for each interval, such that two vertices are adjacent if and only if their corresponding intervals overlap. A graph G is an interval graph if and only if G is isomorphic to $G_{\mathcal{I}}$ for some set $\mathcal{I}$ of intervals. In this case $\mathcal{I}$ is called an interval model of G.

▶ **Theorem 13.** *If G is an interval graph, then $\mathbf{rvc}(G) = \mathrm{diam}(G) - 1$, and the corresponding rainbow vertex coloring can be found in time that is linear in the size of G.*

Proof. Let $G = (V, E)$ be an interval graph and $\mathcal{I}$ be an interval model for G. The interval corresponding to vertex v is denoted by I_v. For each interval $I \in \mathcal{I}$, we let $r(I)$ be its right endpoint and $\ell(I)$ its right endpoint. Let $I_u \in \mathcal{I}$ be such that $r(I_u) \leq r(I)$ for all $I \in \mathcal{I}$. Let $I_v \in \mathcal{I}$ be such that $\ell(I_v) \geq \ell(I)$ for all $I \in \mathcal{I}$. Let $P = u, x_1, x_2, \ldots, x_k, v$ be a shortest path between u and v in G. Observe that P is a connected dominating set. Furthermore, since P is a shortest path, $k \leq \mathrm{diam}(G) - 1$. By the way we defined u and v, we have that $N(u) \subseteq N(x_1)$ and $N(v) \subseteq N(x_k)$. This implies that the set $\{x_1, x_2, \ldots, x_k\}$ is also a connected dominating set. By Proposition 6, G has a rainbow vertex coloring $c : V \to [k]$ with $c(x_i) = i$, and we can give all the other vertices arbitrary colors. ◀

An interval graph is a *unit interval graph* if it has an interval model in which every interval has the same length (or no interval properly contains another interval). Unit interval graphs have the same BFS tree structure as that of bipartite permutation graphs, with the single difference that every level of the BFS tree is a clique instead of an independent set [15].

▶ **Theorem 14.** *If G is a unit interval graph, then $\mathbf{rvc}(G) = \mathbf{srvc}(G) = \mathrm{diam}(G) - 1$, and the corresponding (strong) rainbow vertex coloring can be found in time that is linear in the size of G.*

Proof. Let $G = (V, E)$ be a unit interval graph. Let v be the vertex corresponding to a first interval in an ordering of the intervals in the unit interval model of G by their right endpoints. Do a BFS on G with v as the root. Let L_i be the set of vertices in level i of the BFS tree, $0 \leq i \leq k$, with $L_0 = \{v\}$. Recall that, for $0 \leq i \leq k-1$, there exists a special vertex $a_i \in L_i$ such that $L_{i+1} \subset N(a_i)$.

Consider a vertex $u \in L_k$. A shortest path between v and u has $k-1$ internal vertices, which implies that $\mathrm{diam}(G) \geq k$. To construct a strong rainbow coloring $c : V \to [k-1]$, we assign, for $1 \leq i \leq k-1$, $c(x) = i$ if $x \in L_i$ and we give arbitrary colors to the vertices of L_k.

To see that G is strongly rainbow vertex-connected under c, consider $x, y \in V$. If both x and y are in the same level of the BFS tree, then they are adjacent. So let us consider the case when $x \in L_i$ and $y \in L_j$, with $1 \leq i < j \leq k$. If there is a shortest path between x and y each of whose vertices is in a distinct level of the BFS tree, then this path is rainbow. If this is not the case, we consider the path $x, a_i, a_{i+1}, \ldots, a_{j-1}, y$. In this case, this path is a shortest path between x and y, and its internal vertices have distinct colors, since only x and a_i belong to the same level of the BFS. This proves that c is indeed a strong rainbow coloring for G with $\mathrm{diam}(G) - 1$ colors. ◀

5 Concluding remarks and related problems

It should be mentioned that other variants of rainbow problems have been studied as well. When a coloring of the edges or the vertices of a graph is already given as input, we can ask whether the graph is rainbow-connected or rainbow vertex-connected. Both of these problems are known to be NP-complete even on highly restricted graphs, like interval graphs, series-parallel graphs, and k-regular graphs for every $k \geq 3$ [21, 20, 28]. However, we stress that these problems are strictly different from RC and RVC. That is, complexity results on one problem are not transferable to the other.

Finally, we end our paper with the following open question.[1] A *diametral path* of a graph G is a shortest path whose length is equal to $\mathrm{diam}(G)$. A graph is a *diametral path* if every connected induced subgraph has a dominating diametral path.

▶ **Conjecture 15.** *Let G be a diametral path graph. Then* $\mathbf{rvc}(G) = \mathrm{diam}(G) - 1$.

References

1 Akanksha Agrawal. Fine-grained complexity of rainbow coloring and its variants. In *Proceedings of the 42nd International Symposium on Mathematical Foundations of Computer Science (MFCS)*, pages 60:1–60:14, 2017.

2 Prabhanjan Ananth, Meghana Nasre, and Kanthi K. Sarpatwar. Rainbow connectivity: Hardness and tractability. In *IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS)*, pages 241–251, 2011.

3 Andreas Brandstädt, Van Bang Le, and Jeremy P. Spinrad. *Graph Classes: A Survey*. SIAM Monographs in Discrete Mathematics and Applications, 1999.

4 Sourav Chakraborty, Eldar Fischer, Arie Matsliah, and Raphael Yuster. Hardness and algorithms for rainbow connection. *Journal of Combinatorial Optimization*, 21(3):330–347, 2011.

5 L. Sunil Chandran and Deepak Rajendraprasad. Rainbow Colouring of Split and Threshold Graphs. In *Proceedings of the 18th Annual International Computing and Combinatorics Conference (COCOON)*, volume 7434 of *Lecture Notes in Computer Science*, pages 181–192. Springer, 2012.

6 L. Sunil Chandran and Deepak Rajendraprasad. Inapproximability of rainbow colouring. In *IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS)*, pages 153–162, 2013.

7 L. Sunil Chandran, Deepak Rajendraprasad, and Marek Tesař. Rainbow colouring of split graphs. *Discrete Applied Mathematics*, 216:98–113, 2017.

8 Gary Chartrand, Garry L Johns, Kathleen A McKeon, and Ping Zhang. Rainbow connection in graphs. *Mathematica Bohemica*, 133(1):85–98, 2008.

9 Lily Chen, Xueliang Li, and Huishu Lian. Further hardness results on the rainbow vertex-connection number of graphs. *Theoretical Computer Science*, 481:18–23, 2013.

10 Lily Chen, Xueliang Li, and Yongtang Shi. The complexity of determining the rainbow vertex-connection of a graph. *Theoretical Computer Science*, 412(35):4531–4535, 2011.

11 Paul Dorbec, Ingo Schiermeyer, Elżbieta Sidorowicz, and Éric Sopena. Rainbow connection in oriented graphs. *Discrete Applied Mathematics*, 179:69–78, 2014.

12 Eduard Eiben, Robert Ganian, and Juho Lauri. On the complexity of rainbow coloring problems. In *Proceedings of the 26th International Workshop on Combinatorial Algorithms (IWOCA)*, volume 9538 of *Lecture Notes in Computer Science*, pages 209–220. Springer, 2015.

13 Eduard Eiben, Robert Ganian, and Juho Lauri. On the complexity of rainbow coloring problems. *Discrete Applied Mathematics*, 246:38–48, 2018.

14 M. R. Garey, David S. Johnson, and Larry J. Stockmeyer. Some Simplified NP-Complete Graph Problems. *Theoretical Compututer Science*, 1(3):237–267, 1976.

[1] The statement is claimed in [19, Proposition 5.2] but its proof contains an error. Essentially, only an upper bound of $\mathrm{diam}(G) + 1$ is known by Proposition 6.

15 Pavol Hell and Jing Huang. Certifying LexBFS Recognition Algorithms for Proper Interval Graphs and Proper Interval Bigraphs. *SIAM Journal on Discrete Mathematics*, 18(3):554–570, 2004.

16 Melissa Keranen and Juho Lauri. Computing minimum rainbow and strong rainbow colorings of block graphs. *arXiv preprint arXiv:1405.6893*, 2014.

17 Łukasz Kowalik, Juho Lauri, and Arkadiusz Socała. On the fine-grained complexity of rainbow coloring. In *Proceedings of the 24th Annual European Symposium on Algorithms (ESA)*, pages 58:1–58:16, 2016.

18 Michael Krivelevich and Raphael Yuster. The rainbow connection of a graph is (at most) reciprocal to its minimum degree. *Journal of Graph Theory*, 63(3):185–191, 2010.

19 Juho Lauri. *Chasing the Rainbow Connection: Hardness, Algorithms, and Bounds*. PhD thesis, Tampere University of Technology, 2016.

20 Juho Lauri. Further hardness results on rainbow and strong rainbow connectivity. *Discrete Applied Mathematics*, 201:191–200, 2016.

21 Juho Lauri. Complexity of rainbow vertex connectivity problems for restricted graph classes. *Discrete Applied Mathematics*, 219:132–146, 2017.

22 Van Bang Le and Zsolt Tuza. Finding optimal rainbow connection is hard. Technical Report CS-03-09, Universität Rostock, 2009.

23 Shasha Li, Xueliang Li, and Yongtang Shi. Note on the complexity of deciding the rainbow (vertex-)connectedness for bipartite graphs. *Applied Mathematics and Computation*, 258:155–161, 2015.

24 Xueliang Li, Yaping Mao, and Yongtang Shi. The strong rainbow vertex-connection of graphs. *Utilitas Mathematica*, 93:213–223, 2014.

25 Xueliang Li, Yongtang Shi, and Yuefang Sun. Rainbow connections of graphs: A survey. *Graphs and Combinatorics*, 29(1):1–38, 2013.

26 László Lovász. Coverings and colorings of hypergraphs. In *Proceedings of the 4th Southeastern Conference on Combinatorics, Graph Theory, and Computing*, pages 3–12, 1973.

27 Jeremy Spinrad, Andreas Brandstädt, and Lorna Stewart. Bipartite permutation graphs. *Discrete Applied Mathematics*, 18(3):279–292, 1987.

28 Kei Uchizawa, Takanori Aoki, Takehiro Ito, Akira Suzuki, and Xiao Zhou. On the Rainbow Connectivity of Graphs: Complexity and FPT Algorithms. *Algorithmica*, 67(2):161–179, 2013.

29 Ryuhei Uehara and Gabriel Valiente. Linear structure of bipartite permutation graphs and the longest path problem. *Information Processing Letters*, 103(2):71–77, 2007.

30 David Zuckerman. Linear degree extractors and the inapproximability of max clique and chromatic number. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing (STOC)*, pages 681–690. ACM, 2006.