Rainbow Connectivity: Hardness and Tractability

Prabhanjan Ananth[1], Meghana Nasre[1], and Kanthi K Sarpatwar[2]

1 **Indian Institute of Science**
Bangalore
`meghana,prabhanjan@csa.iisc.ernet.in`

2 **University of Maryland**
College Park, Maryland, U.S.A.
`kanthik@gmail.com`

Abstract

A path in an edge colored graph is said to be a rainbow path if no two edges on the path have the same color. An edge colored graph is (strongly) rainbow connected if there exists a (geodesic) rainbow path between every pair of vertices. The (strong) rainbow connectivity of a graph G, denoted by ($src(G)$, respectively) $rc(G)$ is the smallest number of colors required to edge color the graph such that G is (strongly) rainbow connected. In this paper we study the rainbow connectivity problem and the strong rainbow connectivity problem from a computational point of view. Our main results can be summarised as below:

- For every fixed $k \geq 3$, it is NP-Complete to decide whether $src(G) \leq k$ even when the graph G is bipartite.
- For every fixed odd $k \geq 3$, it is NP-Complete to decide whether $rc(G) \leq k$. This resolves one of the open problems posed by Chakraborty et al. (J. Comb. Opt., 2011) where they prove the hardness for the even case.
- The following problem is *fixed parameter tractable*: Given a graph G, determine the maximum number of pairs of vertices that can be rainbow connected using two colors.
- For a directed graph G, it is NP-Complete to decide whether $rc(G) \leq 2$.

1998 ACM Subject Classification G.2 Discrete Mathematics

Keywords and phrases Computational Complexity, Rainbow Connectivity, Graph Theory, Fixed Parameter Tractable Algorithms

Digital Object Identifier 10.4230/LIPIcs.xxx.yyy.p

1 Introduction

This paper deals with the notion of *rainbow connectivity* and *strong rainbow connectivity* of a graph. Unless mentioned otherwise, all the graphs are assumed to be connected and undirected. Consider an edge coloring (not necessarily proper) of a graph $G = (V, E)$. A path between a pair of vertices is said to be a *rainbow path*, if no two edges on the path have the same color. If the edges of G can be colored using k colors such that, between every pair of vertices there exists a rainbow path then G is said to be k-rainbow connected. Further, if the k-coloring ensures that between every pair of vertices one of its geodesic *i.e.,* one of the shortest paths is a rainbow path, then G is said to be k-strongly rainbow connected. The minimum number of colors required to (strongly) rainbow connect a graph G is called the (strong) rainbow connection number denoted by ($src(G)$, respectively) $rc(G)$.

The concept of rainbow connectivity was recently introduced by Chartrand et al. in [6] as a measure of strengthening connectivity. The rainbow connection problem, apart from being an interesting combinatorial property, also finds an application in routing messages on cellular networks [4]. In their original paper[6], Chartrand et al. determined $rc(G)$ and $src(G)$, in special cases where G is a complete bipartite or multipartite graph. Rainbow connectivity from a computational point of view was first studied by Caro et al. [3] who conjectured that computing the rainbow connection number of a given graph is NP-hard. This conjecture was confirmed by Chakraborty et al. [4], who proved that even deciding whether rainbow connection number of a graph equals 2 is NP-Complete. They further showed that the problem of deciding whether rainbow connection of a graph is at most k is NP-hard where k is an even integer. The status of the k-rainbow connectivity problem was left open for the case when k is odd. One of our results is to resolve this problem.

Our Results. We present the following new results in this paper:

1. For every fixed $k \geq 3$, deciding whether $src(G) \leq k$, is NP-Complete even when G is bipartite. As a consequence of our reduction, we show that it is NP-hard to approximate the problem of finding the strong connectivity of a graph by a factor of $n^{\frac{1}{2}-\epsilon}$, where n is the number of vertices in G.
2. For every fixed odd $k \geq 3$, deciding whether $rc(G) \leq k$ is NP-Complete.
3. We consider the following natural extension of the 2-rainbow connectivity problem: Given a graph G, determine the maximum number of pairs of vertices that can be rainbow connected with two colors. We show that the above problem is *fixed parameter tractable* when the number of pairs to be rainbow connected is a parameter.
4. We extend the notion of rainbow connectivity for directed graphs and show that for a directed graph G it is NP-Complete to decide whether $rc(G) \leq 2$.

In [4], Chakraborty et al. introduced the problem of *subset rainbow connectivity*, where in addition to the graph $G = (V, E)$ we are given a set P containing pairs of vertices. The goal is to answer whether there exists an edge coloring of G with k colors such that every pair in P has a rainbow path. We also use the subset rainbow connectivity problem and analogously define the subset strong rainbow connectivity problem to prove our hardness results.

Related Work. The problem of rainbow connectivity has received considerable attention after it was introduced by Chartrand et al. in [6]. Caro et al. [3], Krivelevich et al. [9], Chandran et al. [5] gave lower bounds for rainbow connection number of graphs as a function of the number of vertices and the minimum degree of the graph. Upper bounds were also given by Chandran et al. [5] for special graphs like interval graphs and AT-free graphs. In [2], Basavaraju et al. gave a constructive argument to show that any graph G can be colored with $r(r+2)$ colors in polynomial time where r is the radius of the graph. The threshold function for random graph to have $rc(G) = 2$ was studied by Caro et al. [3]. In case of strong rainbow connection number, Li et al. [10] and Li and Sun [11] gave upper bounds on some special graphs. Interestingly, no good upper bounds are known for the strong rainbow connection number in the general case.

2 Strong rainbow connectivity

In this section, we prove the hardness result for the k-strong rainbow connectivity problem: given a graph G and an integer $k \geq 3$, decide whether $src(G) \leq k$. In order to show the hardness of this problem, we first consider an intermediate problem called the *k-subset strong rainbow connectivity* problem. The input to the k-subset strong rainbow connectivity problem is a graph $G = (V, E)$ along with a set of pairs $P = \{(u, v) : (u, v) \subseteq V \times V\}$ and an integer k. Our goal is to answer whether there exists an edge coloring of G with at most k colors such that every pair $(u, v) \in P$ has a geodesic rainbow path.

The overall plan is to prove that k-subset strong rainbow connectivity is NP-hard by showing a reduction from the vertex coloring problem. We then establish the polynomial time equivalence of the k-subset strong rainbow connectivity problem and the k-strong rainbow connectivity problem.

2.1 k-subset strong rainbow connectivity

Let $G = (V, E)$ be an instance of the k-vertex coloring problem. The problem is to decide whether G can be vertex colored using k colors if there exists an assignment of at most k colors to the vertices of G such that no pair of adjacent vertices are colored using the same color. This is one of the most well-studied problems in computer science and is known to be NP-hard for $k \geq 3$. Given an instance $G = (V, E)$ of the k-vertex coloring problem, we construct an instance $\langle G' = (V', E'), P\rangle$ of the k-subset strong rainbow connectivity problem.

The graph G' that we construct is a star, with one leaf vertex corresponding to every vertex $v \in V$ and an additional central vertex a. The set of pairs P captures the edges in E, that is, for every edge $(u, v) \in E$ we have a pair (u, v) in the set P. The goal is to color the edges of G' using at most k colors such that every pair in the set P has a geodesic rainbow path. More formally, we define the parameters $\langle G' = (V', E'), P\rangle$ of the k-subset strong rainbow connectivity problem below:

$$V' = \{a\} \cup V; \qquad E' = \{(a, v) : v \in V\}$$
$$P = \{(u, v) : (u, v) \in E\};$$

We now prove the following lemma which establishes the hardness of the k-subset strong rainbow connectivity problem.

▶ **Lemma 1.** *The graph $G = (V, E)$ is vertex colorable using $k(\geq 3)$ colors iff the graph $G' = (V', E')$ can be edge colored using k colors such that for every pair $(u, v) \in P$ there is a geodesic rainbow path between u and v in G'.*

Proof. Assume that G can be vertex colored using k colors; we show an assignment of colors to the edges of the graph G'. Let c be the color assigned to a vertex $v \in V$; we assign the color c to the edge $(a, v) \in E'$. Now consider any pair $(u, v) \in P$. Recall that $(u, v) \in P$ because there exists an edge $(u, v) \in E$. Since the coloring was a proper vertex coloring of G, the edges (a, u) and (a, v) in G' are assigned different colors by our coloring. Thus, the path $u - a - v$ is a rainbow path; further since that is the only path between u and v it is also a geodesic rainbow path.

To prove the other direction, assume that there exists an edge coloring of G' using k colors such that between every pair of vertices in P there is a geodesic rainbow path. It is easy to see that if we assign the color c of the edge $(a, v) \in E'$ to the vertex $v \in V$, we get a coloring that is a proper vertex coloring for G. ◀

Recall the problem of subset rainbow connectivity where we are content with any rainbow path between every pair in P. Note that our graph G' constructed in the above reduction is a tree, in fact a star and hence between every pair of vertices in P there is exactly one path. Thus, all the above arguments apply for the k-subset rainbow connectivity problem as well. As a consequence we can conclude the following:

▶ **Lemma 2.** *For every $k \geq 3$, both the problems k-subset strong rainbow connectivity and k-subset rainbow connectivity are NP-hard even when the input graph G is a star.*

2.2 k-strong rainbow connectivity

In this section, we establish the hardness of deciding whether a given graph can be strongly rainbow connected using just k colors. Formally, we have the following.

▶ **Theorem 3.** *For every $k \geq 3$, deciding whether a given graph G can be strongly rainbow colored using k colors is NP-hard. Further, the hardness holds even when the graph G is bipartite.*

Proof. We reduce the k-subset strong rainbow connectivity problem to this problem. Let $\langle G = (V, E), P \rangle$ be an instance of the k-subset strong rainbow connectivity problem. Using Lemma 2, we know that k-subset strong rainbow connectivity is NP-hard even when G is a star as well as the pairs $(v_i, v_j) \in P$ are such that both v_i and v_j are leaf nodes of the star. We assume both these properties on the input $\langle G, P \rangle$ and use them crucially in our reduction. Let us denote the central vertex of the star G by a and the leaf vertices by $L = \{v_1, \ldots, v_n\}$, that is, $V = \{a\} \cup L$. Using the graph G and the pairs P, we construct the new graph $G' = (V', E')$ as follows: for every leaf node $v_i \in L$, we introduce two new vertices u_i and u_i'. For every pair of leaf nodes $(v_i, v_j) \in (L \times L) \setminus P$, we introduce two new vertices $w_{i,j}$ and $w_{i,j}'$.

$$
\begin{aligned}
V' &= V \cup V_1 \cup V_2 \\
V_1 &= \{u_i : i \in \{1, \ldots, n\}\} \cup \{w_{i,j} : (v_i, v_j) \in (L \times L) \setminus P\} \\
V_2 &= \{u_i' : i \in \{1, \ldots, n\}\} \cup \{w_{i,j}' : (v_i, v_j) \in (L \times L) \setminus P\}
\end{aligned}
$$

The edge set E' is be defined as follows:

$$
\begin{aligned}
E' &= E \cup E_1 \cup E_2 \cup E_3 \\
E_1 &= \{(v_i, u_i) : v_i \in L, u_i \in V_1\} \cup \\
&\quad \{(v_i, w_{i,j}), (v_j, w_{i,j}) : (v_i, v_j) \in (L \times L) \setminus P\} \\
E_2 &= \{(x, x') : x \in V_1, x' \in V_2\} \\
E_3 &= \{(a, x') : x' \in V_2\}
\end{aligned}
$$

We now prove that G' is k-strong rainbow connected iff $\langle G, P \rangle$ is k-subset strong rainbow connected. To prove one direction, we first note that, for all pairs $(v_i, v_j) \in P$, there is a two length path $v_i - a - v_j$ in G and this path is also present in G'. Further, this path is the only two length path in G' between v_i and v_j; hence any strong rainbow coloring of G' using k colors must make this path a rainbow path. This implies that if G cannot be edge colored with k colors such that every pair in P is strongly rainbow connected, the graph G' cannot be strongly rainbow colored using k colors.

To prove the other direction, assume that there is an edge coloring $\chi : E \to \{c_1, c_2, \ldots, c_k\}$ of G such that all pairs in P are strongly rainbow connected. We extend this edge coloring of G to an edge coloring of G' such that G' is strong rainbow connected:

- We retain the color on the edges of G, *i.e.* $\chi'(e) = \chi(e) : e \in E$.
- For each edge $(v_i, u_i) \in E_1$, we set $\chi'(v_i, u_i) = c_3$.
- For each pair of edges $\{(v_i, w_{i,j}), (v_j, w_{i,j})\} \in E_1$, we set $\chi'(v_i, w_{i,j}) = c_1$, $\chi'(v_i, w_{i,j}) = c_2$ (Assume without loss of generality that $i < j$).
- The edges in E_2 form a complete bipartite graph between the vertices in V_1 and V_2. To color these edges, we pick a perfect matching M of size $|V_1|$ and assign $\chi'(e) = c_1, \forall e \in E_2 \cap M$ and $\chi'(e) = c_2, \forall e \in E_2 \setminus M$.

- Finally, for each edge $(a, x') \in E_3$, we set $\chi'(a, x') = c_3$.

It is straightforward to verify that this coloring makes G' strong rainbow connected. We note that the graph $G' = (V', E')$ constructed above is in fact bipartite. The vertex set V' can be partitioned into two sets A and B, where $A = \{a\} \cup V_1$ and $B = L \cup V_2$ such that there are no edges between vertices in the same partition. By this and lemma 2, we conclude the statement of the theorem. ◀

From the same construction when $k = 3$, it follows that deciding whether a given graph G can be rainbow colored using at most 3 colors is NP-hard. To see this, note that between any pair of vertices $(v_i, v_j) \in P$, a path in G' that is not contained in G is of length at least 4 and the shortest path between v_i and v_j is in G. Further, we always color the edges $E' \setminus E$ using 3 colors; hence none of these paths can be rainbow path. Thus, we conclude the following corollary.

▶ **Corollary 4.** *Deciding whether $rc(G) \leq 3$ is NP-hard even when the graph G is bipartite.*

As a consequence of the reduction from the k-subset strong rainbow connectivity to the k-strong rainbow connectivity, we have the following result (the proof is in full version [1]):

▶ **Theorem 5.** *There is no polynomial time algorithm that approximates strong rainbow connection number of a graph $G = (V, E)$ within a factor of $n^{1/2-\epsilon}$, unless $NP = ZPP$. Here n denotes the number of vertices of G.*

3 Rainbow connectivity

In this section, we investigate the complexity of deciding whether the rainbow connection number of a graph G, $rc(G) \leq k$. We prove the NP-hardness of the k-rainbow connectivity problem i.e., deciding if $rc(G) \leq k$, when k is odd. We recall from the lemma 2 that the k-subset rainbow connectivity problem is NP-hard. In the following theorem, we give a reduction of the k-subset rainbow connectivity problem to the rainbow connectivity problem.

▶ **Theorem 6.** *For every odd integer $k \geq 3$, deciding whether $rc(G) \leq k$ is NP-Complete.*

Proof. We reduce the k-subset rainbow connectivity problem to the present problem. Let $\langle G = (V, E), P \rangle$ be an instance of the k-subset rainbow connectivity problem. Since k is assumed to be odd, let $k = 2m + 1$ where $m \in \mathbb{N}$. Let us denote the vertices of G as $V = \{v_1, \ldots, v_n\}$. Given the graph G and a set of pairs of vertices P, we construct an instance $G' = (V', E')$ of the k-rainbow connectivity problem as follows: For each vertex $v_i \in V$, we add $2m$ new vertices denoted by $u_{i,j}$ where $j \in \{1, \ldots, 2m\}$. Further, we add the following two paths: $v_i - u_{i,1} - u_{i,2} \cdots - u_{i,m}$ and $v_i - u_{i,m+1} - u_{i,m+2} \cdots - u_{i,2m}$. We also add edges $(u_{i,m}, u_{i,2m})$ and $(u_{i,1}, u_{i,m+1})$ (if $m = 1$, we only add one edge). For every pair of vertices $(v_i, v_j) \in (V \times V) \setminus P$: we add the edges $(u_{i,m}, u_{j,2m})$ and $(u_{i,2m}, u_{j,m})$. We add two more new vertices x, y and for every $v_i \in V$ we add the following edges: $(x, u_{i,m})$, $(x, u_{i,2m})$, $(y, u_{i,m})$ and $(y, u_{i,2m})$. Figure 1 shows a subgraph of the graph G'. The construction shows extra vertices added corresponding to v_i and v_j such that the pair $(v_i, v_j) \in (V \times V) \setminus P$. More formally, the vertex set V' can be defined as:

$$
\begin{aligned}
V' &= V \cup V_{1,m} \cup V_{m+1,2m} \cup V_{x,y} \\
V_{1,m} &= \{u_{i,j} : v_i \in V, j \in \{1, \ldots, m\}\} \\
V_{m+1,2m} &= \{u_{i,j} : v_i \in V, j \in \{m+1, \ldots, 2m\}\} \\
V_{x,y} &= \{x, y\}
\end{aligned}
$$

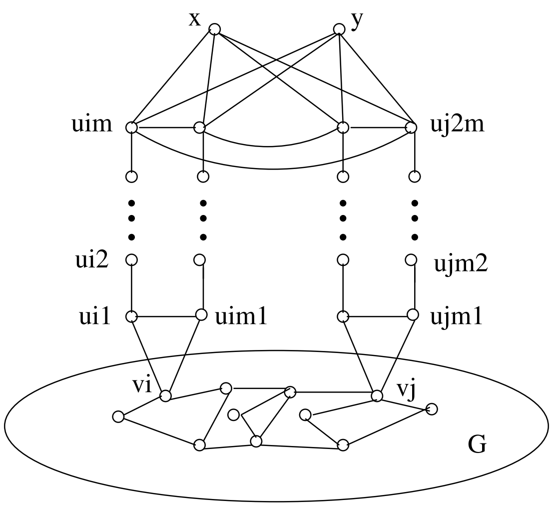

Figure 1 A subgraph of the graph G'. The construction shows extra vertices added corresponding to vertices v_i and v_j belonging to G. The pair $(v_i, v_j) \in (V \times V) \setminus P$.

The edge set E' can be defined as:

$$\begin{aligned}
E' &= E \cup E_1 \cup E_2 \cup E_{x,y} \\
E_1 &= \{(u_{i,j}, u_{i,j+1}) \ : \ v_i \in V, \ j \in \{1, \ldots, m\}, \ j \mod m \neq 0\} \cup \\
& \{(u_{i,1}, u_{i,m+1}), (u_{i,m}, u_{i,2m}) : v_i \in V\} \\
E_2 &= \{(u_{i,m}, u_{j,2m}), (u_{i,2m}, u_{j,m}) \ : \ (v_i, v_j) \in (V \times V) \setminus P\} \cup \\
&= \{(v_i, u_{i,1}), (v_i, u_{i,m+1}) : v_i \in V\} \\
E_{x,y} &= \{(x, u_{i,m}), (x, u_{i,2m}), (y, u_{i,m}), (y, u_{i,2m}) \ : \ i \in \{1, \ldots, n\}\}
\end{aligned}$$

We claim that G can be edge colored using k colors such that every pair belonging to P is rainbow connected if and only if $rc(G') \leq k$. Assume that G can be edge colored using k colors such that all pairs in P are rainbow connected. Let $\chi : E \to \{c_1, \ldots, c_{2m+1}\}$ be such a coloring. We obtain a coloring $\chi' : E' \to \{c_1, \ldots, c_{2m+1}\}$ as follows:

- For every $v_i \in V$:
  $\chi'(v_i, u_{i,1}) = c_1$; $\chi'(v_i, u_{i,m+1}) = c_{m+1}$;
  $\chi'(u_{i,j}, u_{i,j+1}) = c_{j+1}$ where $j \in \{1, \ldots, 2m-1\}$ and $j \mod m \neq 0$;
  $\chi'(x, u_{i,m}) = c_{m+1}$; $\chi'(x, u_{i,2m}) = c_{2m+1}$;
  $\chi'(y, u_{i,m}) = c_{2m+1}$; $\chi'(y, u_{i,2m}) = c_1$.
  If $m \neq 1$, $\chi'(u_{i,1}, u_{i,m+1}) = c_{m+1}$; $\chi'(u_{i,m}, u_{i,2m}) = c_1$
  else $\chi'(u_{i,1}, u_{i,m+1}) = c_1$.
- For every $(v_i, v_j) \in (V \times V) \setminus P$: $\chi'(u_{i,m}, u_{j,2m}) = c_{2m+1}$ and $\chi'(u_{i,2m}, u_{i,m}) = c_{2m+1}$.
- For every edge $(v_i, v_j) \in E$: $\chi'(v_i, v_j) = \chi(v_i, v_j)$.

We claim that if χ makes G k-subset rainbow connected then χ' makes the graph G' k-rainbow connected. This requires a case-wise analysis, which we defer to the full version [1].

To prove the other direction, assume that $rc(G') \leq k$. Let $\chi : E' \to \{c_1, \ldots, c_k\}$ be an edge coloring of G' such that χ makes G' rainbow connected. We will translate this edge coloring of G' to an edge coloring of G as follows: color the edge (v_i, v_j) in G with the color $\chi(v_i, v_j)$. We claim that all pairs in P are rainbow connected in G. This follows from the observation that for a pair $(v_i, v_j) \in P$, any path between v_i and v_j which is of length at most

$2m+1$ in G' has to be completely contained in G. Since χ makes G' rainbow connected, the rainbow path between v_i and v_j in G' has to lie completely inside G itself. Correspondingly, there is a rainbow path between v_i and v_j in G. Hence, all pairs in P are rainbow connected in G. It is clear that given an edge k-coloring, for $k \in \mathbb{N}$, we can check in polynomial time, that the edge coloring rainbow connects every pair of vertices. Hence the problem of deciding if $rc(G) \leq k$ is in NP. Hence we have our theorem. ◀

3.1 Parameterized complexity

In this section, we study the parameterized complexity [7, 8] of a maximization version of the rainbow connection problem. More precisely, given a graph $G = (V, E)$, color the edges of G using 2 colors such that maximum number of pairs are rainbow connected. Since deciding whether $rc(G) \leq 2$ is NP-Complete [4], it follows that the above maximization problem is NP-hard. Any edge coloring of a graph $G = (V, E)$ with 2 colors, trivially satisfies $|E|$ pairs. Hence, we are interested in deciding whether G can be 2-colored such that at least $|E| + k$ pairs of vertices are rainbow connected. We show that the problem is *fixed parameter tractable*, when k is the parameter. A problem is said to be fixed parameter tractable (FPT) if given an instance with size $|x|$ whose parameter size is k there exists an algorithm with running time of form $f(k).p(|x|)$ where $p(.)$ is a polynomial function and $f(.)$ is any arbitrary function. One way of showing that a problem is fixed parameter tractable is to exhibit polynomial time reductions to obtain a "kernel" which is basically an equivalent instance whose size is purely a function of the parameter. For formal definitions, we refer the reader to [7, 8].

We first state a useful lemma (proof in full version [1]). Let us call a *non-edge* in G as an anti-edge; formally we call $e = (u, v)$ an anti-edge of a graph $G = (V, E)$ if $e \notin E$.

▶ **Lemma 7.** *Let $G = (V, E)$ be a connected graph with at least k anti-edges and a clique of size $\geq k$. The edges of G can be colored using 2 colors such that at least $|E| + k$ pair of vertices are rainbow-connected.*

Using the above lemma 7 we now show that the problem is fixed parameter tractable.

▶ **Theorem 8.** *Given a graph $G = (V, E)$, decide whether the edges of G can be colored using 2 colors such that at least $|E| + k$ pair of vertices are rainbow connected. The above problem has a kernel with at most $4k$ vertices and hence is fixed parameter tractable.*

Proof. Let v be any arbitrary vertex and let O_v be the set of vertices which are not adjacent to v. We claim that there is a coloring which rainbow connects at least $|O_v|$ pair of non-adjacent vertices. Consider a breadth first search (bfs) tree rooted at v. Denote the set of vertices in each level of the bfs tree by L_i, $i \geq 1$. Then, $L_1 = \{v\}$, $L_2 = N(v)$ and $O_v = \cup_{i>2} L_i$. We now color the edges from L_{i-1} to L_i by red if i is odd and by blue if i is even. For $i > 2$, every vertex of L_i is rainbow connected to some vertex of L_{i-2}. Thus we have $|O_v|$ pairs of non-adjacent vertices rainbow connected by this coloring. Hence if $|O_v| \geq k$ for any vertex $v \in V$, we have a trivial *yes* instance at hand. Otherwise, $|O_v| < k$, for all $v \in V$.

Recall that our goal is to color the graph using 2 colors such that at least $|E| + k$ pair of vertices are rainbow connected. If G has less than k anti-edges, clearly this is not possible and we have a *no* instance. Assume that this is not the case. Now consider a vertex v and let $N(v)$ denote the neighbors of v in G. Let H denote the complement of the graph induced by the neighbourhood of v, ie the complement of $G[N(v)]$[1] Further, let $C_1, C_2, \ldots, C_r$ denote the components of H. If there are more than k isolated vertices in H, we have a clique of size $\geq k$ in G. Further, since there are at least k anti-edges, using lemma 7, we have a coloring which rainbow connects at least $|E| + k$ pairs of G. Thus we have a *yes* instance.

It remains to deal with the case when the number of isolated vertices in H is less than k. Let C_i be some non-trivial component of H, that is C_i contains at least two vertices. (If no non-trivial component exists, we are already done, since we can bound the number of vertices of G from above by $2k$). We now show a coloring of edges of G such that at least $|C_i| - 1$ non-adjacent vertices are rainbow connected. For this, consider a spanning tree of C_i and color the vertices of the spanning tree level by level using alternate colors. That is, color the root as red, the vertices at the next level in the spanning tree as blue and so on. We map the colors on the vertices of C_i back to the edges of G as follows. If a vertex $u_1 \in C_i$ got the color red, we color the edge $(v, u_1) \in G$ as red. Thus for every edge (u_1, u_2) in C_i that got distinct colors on its end points, we ensure that one pair got rainbow connected via the path u_1, v, u_2. Further, since (u_1, u_2) is an edge in H, it is an anti-edge in G. Thus for every non-trivial component C_i we can rainbow connect $|C_i| - 1$ anti-edges of G. Counting this across all the components we have $\sum_{i=1}^{r} |C_i| - r$ pairs of anti-edges in G rainbow connected. If $\sum_{i=1}^{r} |C_i| - r \geq k$ we have a *yes* instance, otherwise we have:

$$\Sigma_{i=1}^{r} |C_i| - r < k. \tag{1}$$

Let the number of non-trivial components of H be s. Each of these non-trivial components have at least 2 vertices. Hence we have the following:

$$\Sigma_{i=1}^{r} |C_i| \geq 2 * s + (r - s) = r + s \tag{2}$$

Since the number of isolated vertices in H is strictly less than k, we have $r < s + k$. Further, from equations (1) and (2) we get $s < k$. Combining these we have $r < 2k$. Thus we can bound the number of vertices in H as:

$$|H| = \Sigma_{i=1}^{r} |C_i| < r + k < 3k \tag{3}$$

Therefore we have:

$$|G| = |H| + 1 + |O_v| < 3k + 1 + k \implies |G| \leq 4k. \tag{4}$$

Hence, we have a $4k$ vertex kernel. ◀

4 Rainbow connectivity on directed graphs

In this section, we consider the rainbow connectivity problem for directed graphs. All the directed graphs considered in this section are assumed to be connected *i.e.*, between any two vertices u, v in the directed graph there is either a directed path from u to v or from v to u. Consider an edge-coloring of a directed graph $G = (V, E)$. We say that there exists a rainbow path between a pair of vertices (u, v) if there exists a directed path from u to v or from v to u with distinct edge colors. A edge coloring of the edges in a directed graph is said to make the graph rainbow connected if between every pair of vertices there is a

[1] $G[H]$ denotes the induced subgraph of G on vertices of H

rainbow path. Analogous to the undirected version, the minimum colors needed to rainbow color a directed graph G is called the rainbow connection number of the directed graph. The rainbow connection number of a directed graph is at least the rainbow connection number of the underlying undirected graph; however, there are examples where the directed graph requires many more colors than the underlying undirected graph. Consider the directed graph $G = (V, E)$ with $V = \{v_1, \ldots, v_n\}$ and $E = \{(v_i, v_{i+1}) : i = 1, \ldots, n-1\} \cup \{(v_1, v_n)\}$. The rainbow connection number of G is $n-2$ while the rainbow connection number of its underlying undirected graph, which is a cycle, is $\lceil \frac{n}{2} \rceil$.

We study the computational complexity of the problem of computing rainbow connection number for a directed graph. We prove that the problem of deciding whether the rainbow connection of a simple directed graph is at most 2 is NP-hard. As in the case of undirected graphs, we define the problem of subset rainbow connectivity on directed graphs. Given a directed graph $G = (V, E)$ and a set of pairs $P \subseteq V \times V$ decide whether the edges of G can be colored using 2 colors such that every pair in P is rainbow connected (in the directed sense). Throughout this section we will use the term rainbow connected to mean that it is rainbow connected in the directed sense. Our plan, as in the previous cases, is to show that the 2-subset rainbow rainbow connectivity is NP-hard by a showing a reduction from the 3SAT problem. We then establish the polynomial time equivalence of 2-subset rainbow connectivity and 2-rainbow connectivity for a directed graph G.

Let $\mathcal{I}$ be an instance of the 3SAT problem with $X = \{x_1, \ldots, x_n\}$ as the set of variables and $C_1, \ldots, C_m$ being the clauses. We construct from $\mathcal{I}$ a directed graph $G = (V, E)$ and a set of pairs $P \subseteq V \times V$ which is an instance of the 2-subset rainbow connectivity problem. For readability sake, we reuse the symbols C_i, x_i to represent the vertices.

$$\begin{aligned} V &= \{C_i : i \in \{1, \ldots, m\}\} \cup X \cup \bar{X} \cup \{T, R, B\} \\ \bar{X} &= \{\bar{x_i} : x_i \in X\} \end{aligned}$$

The edge set E is defined as below. We say that $x_i \in C_j$ to imply that the clause C_j contains the positive occurrence of the variable x_i. If x_i appears negated in the clause C_j we denote it as $\bar{x}_i \in C_j$.

$$\begin{aligned} E = &\{(R, T), (T, B)\} \cup \\ &\{(x_i, T), (T, \bar{x_i}), (x_i, \bar{x_i}) : x_i \in X\} \cup \\ &\{(C_j, x_i) : x_i \in C_j\} \cup \\ &\{(\bar{x_i}, C_j) : \bar{x_i} \in C_j\} \end{aligned}$$

The set of pairs P is defined as follows:

$$\begin{aligned} P = &\{(C_i, T) : i \in \{1, \ldots, m\}\} \cup \\ &\{(x_i, C_j), (\bar{x_i}, C_j) : x_i \in C_j\} \cup \\ &\{(x_i, C_j), (\bar{x_i}, C_j) : \bar{x_i} \in C_j\} \cup \\ &\{(R, B)\} \cup \{(R, \bar{x_i}), (B, x_i) : x_i \in X\} \end{aligned}$$

We now state the following lemma (proof in full version [1]) which establishes the correctness of our reduction.

▶ **Lemma 9.** *There exists a satisfying assignment for $\mathcal{I}$ if and only if there is an edge coloring of $G = (V, E)$ with 2 colors such that all the pairs in P are rainbow connected.*

We now prove the equivalence of the following two problems.

▶ **Lemma 10.** *The following two problems are polynomial time equivalent:*
(1) Given a directed graph $G = (V, E)$ decide whether G is 2-rainbow connected.
(2) Given a directed graph $G = (V, E)$, and a set of pairs $P \subseteq V \times V$, decide whether $\langle G, P \rangle$ is 2-subset rainbow connected.

Proof. It suffices to prove that problem (2) reduces to problem (1). Given $\langle G = (V, E), P \rangle$ we construct an instance $G' = (V', E')$ as follows:

$$\begin{aligned} V' &= V \cup V_1 \cup \{v_{ex}\} \\ V_1 &= \{w_{i,j} : (v_i, v_j) \in (V \times V) \setminus P, v_i \neq v_j\} \end{aligned}$$

The edge set E' is defined as:

$$\begin{aligned} E' = \; & E \cup \{(v_i, w_{i,j}), (w_{i,j}, v_j) : (v_i, v_j) \in (V \times V) \setminus P, v_i \neq v_j\} \cup \\ & \{(v, v_{ex}), (v_{ex}, x) : v \in V, x \in V_1\} \cup E_1 \end{aligned}$$

The set of edges in E_1 are amongst the vertices in V_1 such that the induced subgraph $T = (V_1, E_1)$ is a tournament.

Assume that G has an edge coloring χ using two colors, say *red* and *blue* such that every pair of vertices in P is rainbow connected. We give a coloring χ' the edges of G' as follows:
- Set $\{\chi'(v, v_{ex}) = red : v \in V\}$ and set $\{\chi'(v_{ex}, x) = blue : x \in V_1\}$.
- For every pair $(v_i, v_j) \in (V \times V) \setminus P$, we set $\chi'(v_i, w_{i,j}) = red$ and $\chi'(w_{i,j}, v_j) = blue$.
- Color the edges of the graph induced by V_1 arbitrarily.
- Set $\{\chi'(v_i, v_j) = \chi(v_i, v_j) : v_i \in V, v_j \in V\}$.

It is easy to verify that the above coloring makes G' rainbow connected.

In the other direction, we note that no pair of vertices in P has a directed 2 length path in G' which is not contained entirely in G. Hence if G' has an edge coloring using 2 colors such that every pair has a rainbow path, then the coloring of the induced subgraph G of G' rainbow connects every pair of vertices in P. This completes the proof of the lemma. ◀

Using lemma 9 and lemma 10 we can conclude the following theorem.

▶ **Theorem 11.** *Given a directed graph $G = (V, E)$, it is NP-hard to decide whether G can be colored using two colors such that between every pair of vertices there is a rainbow path.*

This result can be extended to k colors by using induction, however due to space limitations we do not present the result here.

5 Conclusion

In this paper, we present several hardness results related to the rainbow connectivity problem. The hardness results for the strong rainbow connectivity and rainbow connectivity problem are due to a series of reductions starting from the vertex coloring problem. The nature of our reduction helps us derive an inapproximability result in the case of the strong rainbow connectivity problem. Surprisingly, no good upper bounds are known for the strong rainbow connection number of graphs for the general case. Our study on parameterized version of the rainbow connectivity problem shows a linear kernel when we want to maximize the number of pairs which are rainbow connected using two colors. We initiate the study of rainbow connectivity in directed graphs. Further, we show that the hardness of deciding whether a directed graph can be rainbow connected using at most 2 colors.

6 Acknowledgements

The first two authors would like to thank Deepak Rajendraprasad and Dr. L. Sunil Chandran for the useful discussions on the topic.

References

1 Prabhanjan Ananth, Meghana Nasre, and Kanthi K. Sarpatwar. Hardness and parameterized algorithms on rainbow connectivity problem. *CoRR*, abs/1105.0979, 2011.

2 M. Basavaraju, L.S. Chandran, D. Rajendraprasad, and A. Ramaswamy. Rainbow connection number and radius. *Arxiv preprint arXiv:1011.0620*, 2010.

3 Y. Caro, A. Lev, Y. Roditty, Z. Tuza, and R. Yuster. On rainbow connection. *The Electronic Journal of Combinatorics*, 15(R57):1, 2008.

4 S. Chakraborty, E. Fischer, A. Matsliah, and R. Yuster. Hardness and algorithms for rainbow connection. *Journal of Combinatorial Optimization*, pages 1–18, 2009.

5 L.S. Chandran, A. Das, D. Rajendraprasad, and N.M. Varma. Rainbow connection number and connected dominating sets. *To appear in EUROCOMB*, 2011.

6 G. Chartrand, G.L. Johns, K.A. McKeon, and P. Zhang. Rainbow connection in graphs. *Math. Bohem*, 133(1):85–98, 2008.

7 Rod G. Downey and M.R. Fellows. *Parameterized complexity*. Monographs in computer science. Springer, 1999.

8 J. Flum and M. Grohe. *Parameterized complexity theory*. Texts in theoretical computer science. Springer, 2006.

9 M. Krivelevich and R. Yuster. The rainbow connection of a graph is (at most) reciprocal to its minimum degree. *J. Graph Theory*, 63:185–191, March 2010.

10 H. Li, X. Li, and S. Liu. The (strong) rainbow connection numbers of cayley graphs of abelian groups. *Arxiv preprint arXiv:1011.0827*, 2010.

11 X. Li and Y. Sun. On strong rainbow connection number. *Arxiv preprint arXiv:1010.6139*, 2010.