Upper Bounding Rainbow Connection Number by Forest Number

L. Sunil Chandran\textsuperscript{1*}, Davis Issac\textsuperscript{2**}, Juho Lauri \textsuperscript{3***}, and Erik Jan van Leeuwen\textsuperscript{3}

\textsuperscript{1} Indian Institute of Science, Bangalore, India.
\textsuperscript{2} Hasso Plattner Institute, Potsdam, Germany.
\textsuperscript{3} Utrecht University, The Netherlands.

Abstract. A path in an edge-colored graph is rainbow if no two edges of it are colored the same, and the graph is rainbow-connected if there is a rainbow path between each pair of its vertices. The minimum number of colors needed to rainbow-connect a graph $G$ is the rainbow connection number of $G$, denoted by $rc(G)$.

A simple way to rainbow-connect a graph $G$ is to color the edges of a spanning tree with distinct colors and then re-use any of these colors to color the remaining edges of $G$. This proves that $rc(G) \leq |V(G)| - 1$.

We ask whether there is a stronger connection between tree-like structures and rainbow coloring than that is implied by the above trivial argument. For instance, is it possible to find an upper bound of $t(G) - 1$ for $rc(G)$, where $t(G)$ is the number of vertices in the largest induced tree of $G$? The answer turns out to be negative, as there are counter-examples that show that even $c \cdot t(G)$ is not an upper bound for $rc(G)$ for any given constant $c$.

In this work we show that if we consider the forest number $f(G)$, the number of vertices in a maximum induced forest of $G$, instead of $t(G)$, then surprisingly we do get an upper bound. More specifically, we prove that $rc(G) \leq f(G) + 2$. Our result indicates a stronger connection between rainbow connection and tree-like structures than that was suggested by the simple spanning tree based upper bound.

Keywords: rainbow connection, forest number, upper bound

1 Introduction

Let $G$ be a connected, simple and finite graph. Consider any edge-coloring of $G$. A path in $G$ is said to be rainbow if no two edges of it are colored the same. The graph $G$ is rainbow-connected if there is a rainbow path between each pair of its vertices. If there is a rainbow shortest path between every pair of its vertices, we say that $G$ is strongly rainbow-connected. The minimum number of colors required to rainbow-connect $G$ is known as the rainbow connection number of $G$, and denoted as $rc(G)$. Similarly, the minimum number of colors needed to strongly rainbow-connect $G$ is the strong rainbow connection number of $G$, denoted as $src(G)$. These measures of rainbow connectivity were introduced by Chartrand et al. \cite{chartrand2008} in 2008. The concept has gathered significant attention from both combinatorial and algorithmic perspectives. Indeed, the work of Chartrand et al. \cite{chartrand2008} has already amassed more than 400 citations. In addition to being a theoretically interesting way of strengthening the usual notion of connectivity, rainbow connectivity has potential applications in networking \cite{chartrand2008}, layered encryption \cite{chartrand2008}, and broadcast scheduling \cite{chartrand2008}.

While introducing the parameters, Chartrand et al. \cite{chartrand2008} established basic bounds along with exact values of the parameters for some structured graphs. To repeat their results, recall that the diameter of $G$, denoted by $diam(G)$, is the length of a longest shortest path in $G$. Now, it is straightforward to verify that $diam(G) \leq rc(G) \leq src(G) \leq m$, where $m$ is the number of edges of $G$. In other words, both $rc(G)$ and $src(G)$ are always sandwiched between one and $m$. The extremal cases are not difficult to

\textsuperscript{*} The major part of the work was done when this author was on a long-term research visit at Max Planck Institute for Informatics, Saarbrücken, Germany. The visit was funded by the Alexander von Humboldt fellowship.

\textsuperscript{**} The major part of the work was done when this author was a PhD student at Max Planck Institute for Informatics, Saarbrücken, Germany.

\textsuperscript{***} The major part of the work was done when this author was at Bell Labs, Dublin, Ireland.
see: \( \text{rc}(G) = \text{src}(G) = 1 \) if and only if \( G \) is complete; \( \text{rc}(G) = \text{src}(G) = m \) if and only if \( G \) is a tree. The authors also determined the exact rainbow connection numbers for cycle graphs, wheel graphs, and complete multipartite graphs.

Much of the research on rainbow connectivity has focused on finding bounds on the parameters, either in terms of the number of vertices \( n \) or some other well-known parameters. It follows that \( \text{rc}(G) \leq n - 1 \) by taking a spanning tree and coloring its edges with distinct colors, and repeating an already used color for the other edges. For 2-connected graphs, Ekstein et al. [8] showed that \( \text{rc}(G) \leq \lceil n/2 \rceil \), and this is tight as witnessed by e.g., odd cycles. Further, it has turned out that domination is a useful concept when deriving upper bounds on \( \text{rc}(G) \) (see e.g., [2,11,4]). Specifically, Krivelevich and Yuster [11] showed that \( \text{rc}(G) \leq \frac{20n}{3} \), later improved by Chandran et al. [4] to \( \text{rc}(G) \leq \frac{3n}{\delta + 1} + 3 \), where \( \delta \) denotes the minimum degree of \( G \). Moreover, the latter authors derived that when \( \delta \geq 2 \), then \( \text{rc}(G) \leq \gamma_c(G) + 2 \), where \( \gamma_c(G) \) is the connected domination number. For some structured graph classes, this leads to upper bounds of the form \( \text{rc}(G) \leq \text{diam}(G) + c \), where \( c \) is a small constant. For instance, it follows that \( \text{rc}(G) \leq \text{diam}(G) + 1 \) when \( G \) is an interval graph and \( \text{rc}(G) \leq \text{diam}(G) + 3 \), when \( G \) is an AT-free graph, both bounds holding when \( \delta \geq 2 \). Basavaraju et al. [1] show that for every bridgeless graph \( G \) with radius \( r \), \( \text{rc}(G) \leq r(r + 2) \), and for a bridgeless graph with radius \( r \) and chordality (length of a largest induced cycle) \( k \), \( \text{rc}(G) \leq rk \).

In addition to domination, various authors (see e.g., [2]) have noted trees to be useful in bounding \( \text{rc}(G) \). As mentioned earlier, \( \text{rc}(G) \leq n - 1 \) follows by coloring the edges of a spanning tree of \( G \) with distinct colors. Moreover, Kamčev et al. [10] proved that \( \text{rc}(G) \leq \text{diam}(G_1) + \text{diam}(G_2) + c \), where \( G_1 = (V, E_1) \) and \( G_2 = (V, E_2) \) are connected spanning subgraphs of \( G \) and \( c \leq |E_1 \cap E_2| \). For a more comprehensive treatment, we refer the curious reader to the books [6,14] and the surveys [13,15] on rainbow connectivity.

In light of the above results, it makes sense to search for bounds on \( \text{rc}(G) \) in terms of other graph parameters, that possibly arise from “tree-related” and “dominating” graph structures. Intuitively, a graph structure that has both characteristics is a maximum induced forest of a graph. Hence, the question arises whether one can bound \( \text{rc}(G) \) in terms of its forest number \( f(G) \), the number of vertices in the largest induced forest in the graph. We answer this in the affirmative by proving the following theorem.

**Theorem 1.** A connected graph \( G \) with forest number \( f(G) \) has \( \text{rc}(G) \leq f(G) + 2 \).

Observe that the bound is tight up to an additive factor of 3 due to trees that have \( \text{rc}(G) = n - 1 = f(G) - 1 \). Our bound improves the upper bound of \( n - 1 \) obtained by coloring the edges of a spanning tree in distinct colors, except when \( f(G) \geq n - 2 \). We leave as an open problem the question of whether the stronger upper bound of \( f(G) - 1 \) is true.

One might be tempted to conjecture a strengthening of our bound, namely that \( \text{rc}(G) \) is at most \( t(G) \), the number of vertices in the largest induced tree in the graph. However, this turns out to be not true. To see this, one can consider a graph \( G \) obtained by taking a \( K_k \) for any \( k \geq 3 \) with a pendant vertex attached to each of its vertices. Then, we have that \( \text{rc}(G) = k \) whereas \( t(G) = 4 \).

Finally, we note that the complement of an induced forest is a feedback vertex set. The feedback vertex set number is the size of the smallest set of vertices in a graph whose removal leaves an induced forest. Hence, Theorem [1] directly implies the following.

**Corollary 1.** A connected graph \( G \) with feedback vertex set number \( \text{fvs}(G) \) has \( \text{rc}(G) \leq |V(G)| - \text{fvs}(G) + 2 \).

1.1 **Overview of our Techniques**

We give the proof of the upper bound in three takes. In Take 1 in Section [2,1], we give a short proof that \( \text{rc}(G) \leq 3f(G) + 1 \). For this, we first observe that given a connected dominating set \( D \) such that each
of the remaining vertices have at least two neighbors in $D$, it is easy to find a rainbow coloring with $|D| + 1$ colors. Then we observe that any maximal induced forest $F$ can be turned into such a connected dominating set by adding at most $2|F|$ more vertices. The bound of $3f(G) + 1$ follows from these two observations.

In Take 2 in Section 2.2, we strengthen our bound from $3f(G) + 1$ to $2f(G) + 2$. In this section, we already introduce the main structures and insights used for our final bound in Take 3. We fix a maximum induced forest $F$ and define $H$ to be the graph obtained from $G$ by contracting each connected component of $G − F$, each of which is a tree, of $G$ into a single vertex. Thus $H$ consists of tree vertices and non-tree vertices. An edge from a non-tree vertex $u$ to a tree-vertex $x_T$ is classified as a 2-edge if $u$ has at least two edges to the tree $T$ (the tree that was contracted into the tree-vertex $x_T$), and as a 1-edge otherwise. We fix a carefully chosen spanning tree of $H$, root it at some (contracted) tree vertex, and direct all the edges towards root. We call this the skeleton $B$. The skeleton is chosen so that the number of 2-edges in it is maximized. The inner skeleton $B_1$ is defined to be $B$ minus the leaves of $B$ that are non-tree vertices.

Our idea is to color all the edges of the forest $F$ with distinct colors and then associate each tree of $F$ with two additional colors called surplus colors, and also keep aside two global surplus colors. Note that this makes the total number of colors $2f(G) + 2$ as required. Then, we show that the 2 surplus colors per tree are sufficient to color the edges of the inner skeleton $B_1$ so that there is a rainbow path between every pair of vertices in $B_1$, also giving a corresponding rainbow path in $G$ between every pair of vertices in $B_1$. We show that if each tree $T$ can take care of the first three edges while following the outward path to the root, then we cover all edges of $B_1$. Although we only have two surplus colors, the three edges can indeed be taken care of. For this, we show that the third edge is not taken care of by any other tree only when the first outgoing edge is a 2-edge. In this case, this outgoing 2-edge, say $x_Tu$, can be colored with the same color as some edge in the path between the 2 neighbors of $u$ in $T$, while still maintaining a rainbow connection in $G$ between every pair of vertices in $B_1$.

After rainbow-connecting $B_1$, we connect the vertices outside of the inner skeleton to the inner skeleton using the two global surplus colors. From the choice of the skeleton, we have that each outer vertex has a 2-edge to at least one tree-vertex. Thus, the outer vertices have at least two edges to one tree, that we color with the two different global surplus colors. Thus to get a rainbow path between two outer vertices $x$ and $y$, we can travel from $x$ into the inner skeleton using global surplus color 1, then use the path inside the inner skeleton, and then go to $y$ using global surplus color 2.

In Take 3 in Section 2.3, we improve the bound from $2f(G) + 2$ to $f(G) + 2$. The improvement comes from the fact that we use only one surplus color per tree instead of two. In order to make the coloring of $B_1$ work with one surplus color per tree, we do a case analysis to color the edges around a vertex in $B_1$. The cases are differentiated mainly on the basis of the number of edges and the number of 2-edges incident on a vertex.

1.2 Preliminaries

For a graph $G$, a subgraph $H$ of $G$, and any $E' \subseteq E(G)$, we use $E'(H)$ to denote $E' \cap E(H)$. For a vertex $v$ of (di)graph $G$, we use $\deg_G(v)$ to denote the degree of $v$ in $G$. We use $\text{dist}_G(u,v)$ to denote number of vertices in any shortest path between $u$ and $v$ in $G$. For graph $G$ and $S \subseteq V(G)$, we define $G \setminus S := G[V(G) \setminus S]$. We use $uv$ for an edge between $u$ to $v$ and for a directed edge from $u$ to $v$, we use $\overrightarrow{uv}$. For the latter, we may omit the arrow, when the direction is not relevant. For a directed graph $G$, we denote by $\tilde{G}$, the underlying undirected graph of it. Since, for a forest $F$, each connected component is a tree, we will use the phrases “tree of $F$” and “connected component of $F$” analogously. An in-arborescence is a directed graph with a special root vertex such that all vertices have a unique
directed path to the root vertex. For a tree $T$ and vertices $u$ and $v$ in $T$, we use $T_{uv}$ to denote the unique path in $T$ between $u$ and $v$.

2 Proof of Theorem

Let $G = (V,E)$ be a connected graph. Our goal is to prove that $\text{rc}(G) \leq f(G) + 2$. In what is to follow, we proceed incrementally by starting from a weaker bound, namely that $\text{rc}(G) \leq 3f(G) - 1$ in Section 2.1 and then improving upon this by proving that $\text{rc}(G) \leq 2f(G) + 2$ in Section 2.2. Finally, in Section 2.3 to prove our final result, we use the ideas presented in these proofs in a more detailed and technically involved manner.

2.1 Take 1: $\text{rc}(G) \leq 3f(G) - 1$

In this subsection, as a starting point, we prove that $\text{rc}(G) \leq 3f(G) - 1$ which is considerably weaker but significantly less involved than our main result. We begin with the following simple construction of a rainbow coloring.

Lemma 1. If there is a set $D \subseteq V$ such that $G[D]$ is connected and every vertex in $V \setminus D$ has at least two neighbors in $D$, then $\text{rc}(G) \leq |D| + 1$.

Proof. Since $G[D]$ is connected, it has at least one spanning tree $T$. Pick any spanning tree $T$ of $G[D]$ and color its edges with distinct colors from 1 to $|D| - 1$. By the second precondition of the lemma, every vertex $v \in V \setminus D$ has at least two neighbors in $D$; call them $d_1(v)$ and $d_2(v)$. Color the edge $vd_1(v)$ with color $|D|$ and $vd_2(v)$ with color $|D| + 1$. We will now prove that there is a rainbow path between any $u, v \in V(G)$.

When $u, v \in V(D)$, the proof is trivial as $G[D]$ is rainbow-connected using only edges from $T$, all of which we colored distinctly. Similarly, when $u \in D$ and $v \in V \setminus D$, a rainbow path between the two uses only edges of $T$ plus either $vd_1(v)$ or $vd_2(v)$, both colored with a color not appearing on any edge in $T$. Finally, suppose both $u$ and $v$ are in $V \setminus D$. Observe that there is a rainbow path between $d_1(u)$ and $d_2(v)$ using only edges from $T$, particularly not using colors $|D|$ or $|D| + 1$. Since $ud_1(u)$ is colored with color $|D|$ and $vd_2(v)$ is colored with color $|D| + 1$, we have a rainbow path between $u$ and $v$, completing the proof.

To proceed, let $\mathcal{F}$ be a maximum induced forest of $G$. Let $F = V(\mathcal{F})$, i.e., the set of vertices in $\mathcal{F}$ and let $\mathcal{T}$ denote the set of connected components of $\mathcal{F}$. The following structural observation will be useful for us.

Lemma 2. For any maximum induced forest $\mathcal{F}$ of $G$ and for any $v \in V \setminus F$, there exists a connected component $T$ of $\mathcal{F}$ such that $v$ has at least two neighbors in $T$.

Proof. Otherwise, $G[F \cup \{v\}]$ is a forest, contradicting the maximality of $\mathcal{F}$.

The proof of the following lemma is folklore after observing that $F$ is a dominating set of $G$ (by the maximality of $\mathcal{F}$), and proved here only for the sake of completeness.

Lemma 3. There exists a vertex set $A \subseteq V \setminus F$ of size at most $2|F| - 2$ such that $G[F \cup A]$ is connected.

Proof. Let $B$ be a smallest subset of $V \setminus F$ such that $G[F \cup B]$ has a smaller number of connected components than $G[F]$. We show that $|B| \leq 2$. From the minimality of $B$, it follows that $B$ induces a
Fig. 1: (a) A graph $G$ is partitioned into a maximum induced forest $F$ and $S = V(G) \setminus V(F)$. The connected components (trees) of $F$ are $T_1, T_2, T_3$ and $T_4$. The edges between two black vertices, corresponding to vertices in $V(F)$, are tree-edges. (b) The graph $H$ obtained after contracting the connected components of $F$. We draw a 2-edge with 2 lines and a 1-edge with a single line.

path in $G$ that connects two connected components of $F$. Let $b_1, b_2, \ldots, b_k$ be the vertices in this path, where $b_1$ and $b_k$ are adjacent to some vertex in distinct connected components $T_1$ and $T_2$ respectively. Due to Lemma 2, $b_2$ has an edge to some connected component $T \in T$. If $T = T_1$, then $B \setminus \{b_1\}$ would also connect $T_1$ and $T_2$. Hence $T \neq T_1$. But then the vertex set $\{b_1, b_2\}$ is sufficient to connect $T$ and $T_1$ and thereby reduce the number of components. This implies $k = 2$ and hence $|B| \leq 2$. Now, we add $B$ to $F$ to reduce the number of connected components by at least one. Then we can repeat the process again until we get a single connected component. At each stage, the currently existing connected components take the role of the $T_i$’s. We only need to repeat the process at most $|F| - 1$ times until we get a single connected component. Since each repetition adds at most 2 vertices, the total number of vertices that we add is at most $2|F| - 2$.

We then arrive at the following conclusion.

**Theorem 2.** Any connected graph $G$ has $rc(G) \leq 3f(G) - 1$.

**Proof.** By Lemma 3, we have a set $A$ such that $G[F \cup A]$ is connected and $|A| \leq 2|F| - 2$. Let $D = F \cup A$. By Lemma 2, we have that each vertex in $V \setminus F$ has at least two neighbors in $F$. This implies that each vertex in $V \setminus D$ has at least two neighbors in $D$. Thus, $D$ satisfies both preconditions of Lemma 1 and hence

$$rc(G) \leq |D| + 1 \leq |F| + |A| + 1 \leq |F| + (2|F| - 2) + 1 = 3f(G) - 1.$$ 

This completes the proof.

**2.2 Take 2: $rc(G) \leq 2f(G) + 2$**

In this subsection we strengthen the previous upper bound to $2f(G) + 2$. Many of the concepts and techniques that we use in the final proof are already introduced here.

As in the previous subsection, let $F$ be a maximum induced forest of $G$. Let $F = V(F)$ be the set of vertices in $F$. Let $T$ be the set of connected components (trees) of $F$ and let $t = |T|$. Let $S := V \setminus F$. Also, let $f = |V(F)| = f(G)$. We call an edge $uv$ of $G$ a **tree-edge** if both $u$ and $v$ belong to the same tree in $T$; otherwise, the edge is called a **non-tree edge**.

Let $H$ be the graph obtained from $G$ by contracting each connected component of $F$ to a single vertex (see Figure 1). Formally, we define $H$ as:

- $V(H) := V_T \cup S$, where
- $V_T := \{x_T : T \in T\}$ and
- $E(H) := E(G[S]) \cup \{ux_T : u \in S, T \in T, u$ has at least one edge to $V(T)$ in $G\}$. 

We call the vertices in \( V_T \) the **tree vertices** and the vertices in \( S \) the **non-tree vertices** of \( H \). Notice that \( V_T \) is an independent set in \( H \), because there are no edges in \( G \) between any two distinct connected components of \( F \). We partition the edges of \( H \) into the following two sets:

\[
E_1 := E(G[S]) \cup \{ux_T : u \in S, T \in \mathcal{T}, \ u \text{ has exactly one edge to } V(T) \text{ in } G\} \quad \text{and} \quad E_2 := \{ux_T : u \in S, T \in \mathcal{T}, \ u \text{ has at least two edges to } V(T) \text{ in } G\}.
\]

The edges in \( E_1 \) are called **1-edges** while those in \( E_2 \) are **2-edges**. See Figure 1 for an illustration of the above definitions. We define a function \( f_T : V_T \to \mathcal{T} \) that maps a tree vertex to its corresponding tree, i.e., \( f_T(x_T) = T \). For each edge in \( H \), we define its **representatives** in \( G \) as follows. Consider first a 2-edge \( e \) between \( u \in S \) and \( x_T \in V_T \). By definition of a 2-edge, \( u \) has at least two edges to \( V(T) \) in \( G \). We arbitrarily choose two of these edges as the representatives in \( G \) of the 2-edge \( e \) in \( G \) and denote them by \((e)_1 \) and \((e)_2 \). For a 1-edge \( e \) between \( u \in S \) and \( x_T \in V_T \), there is a unique edge between \( u \) and \( V(T) \) in \( G \), by the definition of a 1-edge. We call this edge the representative of \( ux_T \) in \( G \), and denote it \((e)_1 \). For a 1-edge \( e \) between \( u \in S \) and \( v \in S \), we call \( uv \) its own representative in \( G \). For simplicity, we might simply say **representatives** instead of representatives in \( G \). Whenever we say a representative, it is implicitly assumed that we are talking about an edge in \( G \). For a 2-edge \( ux_T \) with representatives \( uv_1 \) and \( uv_2 \), we call the vertices \( v_1 \) and \( v_2 \), the **foot**s of \( ux_T \). The unique path between \( v_1 \) and \( v_2 \) in \( T \) is called the **foot-path** of \( ux_T \).

A **skeleton** is an in-arborescence obtained by taking a spanning tree of \( H \) with an arbitrary node of \( V_T \) fixed as its root with all edges directed towards the root. Let \( B \) be a skeleton such that the number of 2-edges in \( B \) is as large as possible (or equivalently, the number of 1-edges is as small as possible, as the total number of edges in a skeleton is always \(|V(H)| - 1\)). The **parent** of a non-root vertex \( v \) in \( B \), denoted by \( \text{par}(v) \), is the unique out-neighbor of \( v \) in \( B \). The **children** of \( v \) are the in-neighbors of \( v \) in \( B \). Whenever we say the parent (or child), we mean the parent (or child) in \( B \), even if \( B \) is not mentioned explicitly. We call a directed edge \( \overrightarrow{uv} \) in \( B \) a 1-edge (or 2-edge respectively), if \( uv \) is a 1-edge (or 2-edge respectively) in \( H \). Let \( L_S \) be the set of vertices of \( S \) that are leaves of \( B \) and let \( B_1 \) be the sub-arborescence of \( B \) defined as \( B_1 := B[V(B) \setminus L_S] \). We call \( B_1 \) the **inner skeleton**. Let \( \bar{B}_1 \) be the underlying undirected tree of \( B_1 \). These concepts are illustrated in Figure 2.

We now prove a lemma that is useful for our coloring procedure.

**Lemma 4.** Every vertex in \( S \) has at least one 2-edge incident on it in \( B \).

**Proof.** Suppose for the sake of contradiction that \( v \) is a vertex in \( S \) that has only 1-edges incident on it in \( B \). By Lemma 2, there exist a \( T \in \mathcal{T} \) such that \( v \) has at least two edges to \( T \) in \( G \). Therefore, \( vx_T \) is a 2-edge in \( H \). Let \( C \) be the connected component of \( B \setminus v \) that contains the vertex \( x_T \). Let \( e \) be the unique edge in \( B \) between \( v \) and \( C \). Removing \( e \) from \( B \) and adding 2-edge \( vx_T \) gives a skeleton with higher number of 2-edges than \( B \). This is a contradiction to the choice of \( B \). \( \square \)

Fig. 2: (a) A skeleton \( B \), where \( x_{T_3} \) is the root. For vertex \( v \), \( x_{T_2} \) and \( x_{T_4} \) are the children. (b) The inner skeleton \( B_1 := B[V(B) \setminus L_S] \).
Fig. 3: The rainbow coloring from Take 2 for the graph $G$ in Figure 1. As an example, the color-giving edge of $T_1$ is the edge colored with 1. Only relevant edges are shown. The direction of edges across $S$ (the white vertices) and $F$ (the black vertices) drawn according to the direction in $B_1$. The surplus colors are $s_1(T_1) = 7$, $s_2(T_1) = 8$, $s_1(T_2) = 9$, $s_2(T_2) = 10$, $s_1(T_3) = 11$, $s_2(T_3) = 12$, $s_1(T_4) = 13$, and $s_2(T_4) = 14$ while the global surplus colors are $g_1 = 15$ and $g_2 = 16$.

The above lemma has the following corollaries.

**Corollary 2.** For every vertex in $L_S$, the unique edge incident on it in $B$ is a 2-edge.

**Corollary 3.** Every leaf of $B_1$ is a tree vertex.

**Proof.** Suppose for the sake of contradiction that there is a leaf $v$ of $B_1$ that is a non-tree vertex. Clearly, $v \notin L_S$ by the definition of $B_1$. Hence, $v$ is not a leaf of $B$. Then, there must be a vertex $u$ in $L_S$ that has an edge to $v$ in $B$. Since both $u$ and $v$ are in $S$, the edge $uv$ is a 1-edge. This is a contradiction to Corollary 2. \hfill \square

**Corollary 4.** For each 1-edge $\overrightarrow{uv}$ in $B_1$, either $u$ is a tree vertex, or a child $u'$ of $u$ in $B_1$ is a tree vertex with $u'u$ being a 2-edge.

**Proof.** Suppose that $u$ is not a tree vertex. Then there is an incoming 2-edge on $u$ in $B$, because its outgoing edge is a 1-edge and there has to be at least one 2-edge incident on it due to Lemma 4. Let the other endpoint of this edge be $u'$. Since at least one of the endpoints of a 2-edge has to be a tree vertex, $u'$ is a tree vertex. Since $u'$ is a tree vertex, it has to be in $B_1$. \hfill \square

**Coloring procedure:** We now give a rainbow coloring of $G$ using $f + t + 2 \leq 2f + 2$ colors (recall that $t = |T|$ is the number of connected components in $F$ and $f = f(G) = |V(F)|$). The coloring procedure

Fig. 4: Coloring of $B_1$ according to the coloring procedure in Take 2.
is illustrated in Figure 3. Since $F$ is a forest with $t$ connected components, the number of edges in $F$ is $f - t$. Color all the edges in $F$ with distinct colors $1, 2, \ldots, f - t$. We call colors $g_1 := f + t + 1$ and $g_2 := f + t + 2$, the \textbf{global surplus colors}. We use the global surplus colors to color the representatives of those edges of $B$ that are incident on the vertices in $L_S$. Each vertex in $L_S$ has only one edge incident on it in $B$, and this edge is a 2-edge due to Corollary 2. Color one of the representatives of this 2-edge with $g_1$ and the other with $g_2$. Now there are $2t$ unused colors, which are the colors $f - t + 1$ to $f + t$.

We allocate each tree in $T$ (i.e., each connected component of $F$), two of these colors as its \textbf{surplus colors}. That is, the $i$th tree $T$ in $T$ is allocated colors $f - t + 2(i - 1) + 1$ and $f - t + 2(i - 1) + 2$ as its surplus colors. We denote the two surplus colors of $T$ by $s_1(T)$ and $s_2(T)$.

Finally, we give a coloring of the edges of $B_1$ and then extend the coloring to their representatives in $G$ (see Figure 4). Whenever we color an edge $e$ in $B_1$ with color $c$, we also color the representatives of $e$ in $G$ also with $c$, though we may not mention this explicitly. Pick each $T \in T$ such that $x_T$ is not the root of $B$ and do the following: Let $\overrightarrow{x_Tv}$ be the outgoing edge of vertex $x_T$ in $B$, i.e., $v = \text{par}(x_T)$; let $w = \text{par}(v)$ (if parent of $v$ exists, i.e., if $v$ is not the root) and let $z = \text{par}(w)$ (if $w$ exists and parent of $w$ exists, i.e., if $w$ is not the root).

\textbf{Case 1.} $\overrightarrow{x_Tv}$ is a 2-edge.

We fix an arbitrary edge in the foot-path of $x_Tv$ as the \textbf{color-giving edge} of $T$. If $x_Tv$ is uncolored, then we color it with the same color as that of the color-giving edge of $T$. Note that the color-giving edge is already colored as we have already colored all the edges inside $F$. If $w$ exists and $\overrightarrow{vw}$ is uncolored, color $\overrightarrow{vw}$ with $s_1(T)$, the first surplus color of $T$. If $z$ exists and $\overrightarrow{wz}$ is uncolored, color $\overrightarrow{wz}$ with $s_2(T)$, the second surplus color of $T$.

\textbf{Case 2.} $\overrightarrow{x_Tv}$ is a 1-edge.

If $\overrightarrow{x_Tv}$ is uncolored, then color $\overrightarrow{x_Tv}$ with $s_1(T)$. If $w$ exists and $\overrightarrow{vw}$ is uncolored, then color $\overrightarrow{vw}$ with $s_2(T)$.

We will prove in Lemma 5 that the above procedure in fact colors all the edges of $B_1$, and moreover does so with distinct colors. We extend this coloring of the edges of $B_1$ to their representatives in $G$ as mentioned before: for each edge $e$ of $B_1$ color their representatives (both representatives in case of 2-edges and the only representative in case of 1-edge) with the color of $e$.

At this point, there might be some edges in $G$ that are still uncolored. We call them \textbf{irrelevant edges}. Indeed, when we exhibit rainbow paths between pairs of vertices later, these edges will not be used. The edges of $G$ that are not irrelevant are called \textbf{relevant edges}. To complete the edge-coloring of $G$, color all the irrelevant edges with an arbitrary color from $[f + t + 2]$.

\textbf{Lemma 5.} All the edges of $B_1$ are colored and they are colored with distinct colors.

\textbf{Proof.} It is easy to see that the colors are distinct as each color is used only once while coloring $B_1$. Therefore, it only remains to prove that all edges are colored. Assume for the sake of contradiction that there is an uncolored edge $\overrightarrow{uw}$ of $B_1$. Observe that in the coloring procedure, for each $x_T \in V_T$, we have colored its outgoing edge and the outgoing edge of its parent (if such an edge exists) irrespective of whether Case 1 or 2 was applied. Hence, neither $u$ or any child of $u$ is a tree vertex. Since $u \notin V_T$, it holds that $u$ is not a leaf of $B_1$ by Corollary 3 and consequently $u$ has at least one child in $B_1$. Let $x$ be one such child. As discussed above, $x$ is not a tree vertex. Now, the edge $\overrightarrow{xu}$ is a 1-edge as both endpoints are in $S$. Then, by Corollary 4 there is a tree vertex $x_T$ that is a child of $x$ with $\overrightarrow{x_Tu}$ being a 2-edge. Then, Case 1 of the coloring procedure was applied on $T$, during which edge $\overrightarrow{uv}$ would have been colored.

We are now ready to prove that the coloring of the edges of $G$ produced above is indeed a rainbow coloring. We will prove that there is a rainbow path between every pair of vertices in $G$. For this, we
first prove in Lemma 6 that there is a rainbow path between any pair of vertices in \( V(G) \setminus L_S \) and then in Lemma 7 we show that there is a rainbow path between any pair of vertices in \( V(G) \). The following observation is helpful in proving these lemmas.

**Observation 1** Let \( v_1, v_2, \) and \( v_3 \) be three vertices in any tree \( T \), and let \( e \) be an edge in \( T_{v_2v_3} \). Then either \( T_{v_1v_2} \) or \( T_{v_1v_3} \) does not contain the edge \( e \).

For a vertex \( v \) in \( V(G) \), if \( v \in S \) define \( h(v) := v \), otherwise (i.e., if \( v \in F \)) define \( h(v) := x_T \), where \( T \in T \) is the tree containing \( v \).

**Lemma 6.** For any pair of vertices \( v_1, v_2 \in V(G) \setminus L_S \), there is a rainbow path between \( v_1 \) and \( v_2 \) in \( G \) that uses only colors in \([f + t]\).

**Proof.** Consider a pair of vertices \( v_1, v_2 \in V(G) \setminus L_S \). We will construct a rainbow path \( P \) from \( v_1 \) to \( v_2 \) using only the edges of \( G \) that have colors in \([f + t]\). From Lemma 5 we know that there is a rainbow path between \( v_1' := h(v_1) \) and \( v_2' := h(v_2) \) in \( B_1 \) that uses only the colors in \([f + t]\). Let this path be \( P' \). We will use the path \( P' \) as a guide to construct our required rainbow path \( P \) in \( G \). First, break \( P' \) into two paths \( P_1' \) and \( P_2' \) as follows. Let \( v_1'' \) be the least common ancestor of \( v_1' \) and \( v_2' \) in \( B_1 \), let \( P_1' \) be the path from \( v_1'' \) to \( v_1' \) and let \( P_2' \) be the path from \( v_2' \) to \( v_2'' \).

We will first construct a path \( P_1 \) in \( G \) starting from \( v_1 \) using the path \( P_1' \) as a guide: we start with \( P_1 \) being just the vertex \( v_1 \). We maintain a current vertex in \( G \), denoted by \( v_G \), which is initialized to \( v_1 \). Let \( e \) be the outgoing edge from \( h(v_G) \) in \( B_1 \). The edge \( e \) is the next edge in \( P_1' \) to be processed. As long as \( h(v_G) \neq v_2' \), repeat the following step.

**Case 1.** \( h(v_G) \in S \). Append \((e)_1\) to \( P_1 \). Also, update \( v_G \) to be the other endpoint of \((e)_1\).

**Case 2.** \( h(v_G) \in V_T \).

Let \( x_T = h(v_G) \). We then branch into two sub-cases.

**Case 2.1** \( e \) is a 1-edge.

Let \( z \) be the endpoint of \((e)_1\) in \( T \). Append the path \( T_{v_Gz} \) to \( P_1 \). Then append \((e)_1\) to \( P_1 \). Update \( v_G \) to be the endpoint of \((e)_1\) outside \( T \).

**Case 2.2** \( e \) is a 2-edge.

Let \( z_1 \) and \( z_2 \) be the foots of \( e \) in \( T \). Observe that \( v_G \in V(T) \) as \( h(v_G) = x_T \). By Observation 1 there exists a \( z \in \{z_1, z_2\} \) such that there is a path from \( v_G \) to \( z \) that excludes the color-giving edge of \( T \). Append this path to \( P_1 \). Now, append to \( P_1 \) the representative of \( e \) having one of the endpoints as \( z \). Update \( v_G \) to be the endpoint outside \( T \) of this representative.

Similarly, we also construct \( P_2 \) starting from \( v_2 \) using \( P_2' \) as a guide. If \( v_2' \in S \), then we take \( P := P_1P_2 \). Otherwise, if \( v_2' \in V_T \), we define \( P \) as follows. Let \( T = f_T(v_2') \). There are vertices \( w_1, w_2 \in V(T) \) that are the endpoints of \( P_1 \) and \( P_2 \), respectively. Let \( P_3 \) be the path \( T_{w_1w_2} \). Take \( P := P_1P_3P_2 \).

By construction, it is clear that \( P \) is indeed a path from \( v_1 \) to \( v_2 \). It only remains to prove that \( P \) is a rainbow path whose colors are a subset of \([f + t]\). Observe that each edge that we have added to \( P \) is either a representative of an edge in \( P' \) (call such edges \( E_{PP} \)) or an edge in \( F \) (call such edges \( E_{PF} \)). Recall that the representatives of an edge \( e \) in \( B_1 \) were colored with the same color as that of \( e \), and that each edge of \( P' \) is colored with a color from \([f + t]\). Thus, each edge of \( E_{PP} \) is a relevant edge that having a color in \([f + t]\). Also recall that each edge in \( F \) was colored with a color from \([f - t]\). Thus, each edge of \( E_{PF} \) is a relevant edge having a color in \([f - t]\). This means that all edges in \( P \) are relevant edges with colors in \([f + t]\). In our coloring of the edges of \( G \), each of the colors \( 1, 2, \ldots, f - t \) except the colors of the color-giving edges have been used only for one relevant edge. Also, the color of
Proof. Consider any pair of vertices $x_1, x_2 \in B$. But when constructing $P$, we have taken care not to include the color-giving edge of $T$ in $P$ if $P$ contains a representative of the outgoing edge in of $T$. Thus, the colors $1, 2, \ldots, f - t$ appear at most once in $P$. Each of the colors from $f - t + 1$ to $f + t$ (surplus colors of the trees) appear on at most two relevant edges. Further, if they appear on two edges, then those two edges are the two representatives of some 2-edge of $B$. Since we constructed $P$ in such a way that at most one representative of any edge in $B$ is included in $P$, each surplus color appears at most once in $P$. Thus $P$ is a rainbow path, concluding the proof. \hfill \Box

Lemma 7. For any pair of vertices $v_1, v_2 \in V(G)$, there is a rainbow path between $v_1$ and $v_2$ in $G$.

Proof. Consider any pair of vertices $v_1, v_2 \in V(G)$. If both $v_1, v_2 \in V(G) \setminus L_S$, then we are done by Lemma 6. So assume without loss of generality that $v_1 \in L_S$. Recall that if $v_1 \in L_S$, then there is 2-edge $e$ in $B$ incident on $v_1$ and that the representatives of $e$ are colored with $g_1$ and $g_2$. Let $e_1$ be the representative of $e$ that is colored $g_1$, and let $a$ be the other endpoint of $e_1$. If $v_2 \in L_S$, let $e_2$ be the edge incident on $v_2$ that is colored $g_2$, and let $b$ be the other endpoint of $e_2$. If $v_2 \notin L_S$, let $b = v_2$. We know that there is a rainbow path $P_{ab}$ from $a$ to $b$ that uses only colors in $[f + t]$ due to Lemma 6. We define the path $P := v_1 P_{ab} v_2$. Since the edge $v_1 a$ is colored with $g_1 = f + t + 1$, the edge $b v_2$ is colored with $g_2 = f + t + 2$, and path $P_{ab}$ uses only colors in $[f + t]$, we have that the path $P$ is indeed a rainbow path between $v_1$ and $v_2$. \hfill \Box

The following result is now immediate from the constructed coloring and Lemma 7.

Theorem 3. Any connected graph $G$ has $rc(G) \leq 2f(G) + 2$.

2.3 Take 3: $rc(G) \leq f(G) + 2$

In this section, we prove our final bound, by further developing the ideas from the previous takes. Recall that in Take 2, we gave two surplus colors to each tree. Here, we give only one surplus color to each tree and thereby reduce the number of colors used. However, our analysis has to be tighter to make the proof work with only one surplus color per tree. This makes the proof much more technical and lengthy.

Similar to Take 2, we fix a maximum induced forest $F$ of $G$, but with an additional property as follows. Let $F$ be the maximum induced forest that has the smallest number of connected components (trees) out of all the maximum induced forests of $G$. Now that $F$ is fixed, we define $T, F, S, H, V_T, E_1, E_2, f_T, f, t$, tree vertices, non-tree vertices, 1-edges, 2-edges, representatives, foots, foot-path and skeleton, the same way as in Take 2. However, the selection of the skeleton $B$ is done in a more involved way here. Given a skeleton $B$ with root $r$, we define the level of each node $v$, denoted by $\ell_B(v)$, as its distance (in terms of number of vertices) to $r$ in $B$. Note that $\ell_B(r) = 1$ per this definition. For a skeleton $B$, we define its configuration vector as the following vector:

$$\langle |E_2(B)|, \Sigma_{v: \ell_B(v) = 1} \deg_B(v), \Sigma_{v: \ell_B(v) = 2} \deg_B(v), \ldots, \Sigma_{v: \ell_B(v) = |V|} \deg_B(v) \rangle,$$

where $\deg_B(v)$ is the total degree (sum of in-degree and out-degree) of $v$ in $B$. We now fix a skeleton $B$ such that it has the lexicographically highest configuration vector out of all possible skeletons. We also define $L_S$ and the inner skeleton $B_1$ the same way as in Take 2, i.e., $L_S$ is the set of vertices in $S$ that are leaves of $B$, and $B_1 = B[V(B) \setminus L_S]$. Let $\tilde{B}_1$ be the underlying undirected tree of $B_1$. We note that since the first element of the configuration vector is the number of 2-edges, $B$ would have been a valid skeleton in Take 2 as well. In particular, Lemma 4 and Corollaries 2, 3 and 4 from Take 2 hold for $B$. 
We define a mapping $h$ from $G$ to $H$ as follows. For a vertex $v \in V(G)$, if $v \in S$ then define $h(v) := v$, otherwise (i.e., if $v \in F$) define $h(v) := x_T$, where $T \in \mathcal{T}$ is the tree containing $v$. For a non-tree edge $e = uv$ in $G$, we define $h(e)$ to be the edge $h(u)h(v)$. For a vertex subset $U$ of $V(G)$, we define $h(U)$ to be $\bigcup_{a \in U} h(a)$. For an edge subset $E'$ of $E(G)$, we define $h(E')$ to be $\{h(e) : e \in E'$ and $e$ is a non-tree edge$\}$. For a subgraph $G'$ of $G$, we define $h(G')$ as the subgraph of $H$ with vertex set $h(V(G'))$ and edge set $h(E(G'))$.

Let the palette of colors be $\{1, 2, \ldots, f + 2\}$. We call colors $f + 1$ and $f + 2$ the **global surplus colors**, and denote them by $g_1$ and $g_2$. We reserve $g_1$ and $g_2$ to color the edges incident on $L_S$. We will first give a coloring of some edges of $G$ using colors $\{1, 2, \ldots, f\}$ such that there is a rainbow path between every pair of vertices in $V(G) \setminus L_S$. Then we will extend the coloring to $L_S$ using the global surplus colors. We give our coloring procedure as a list of **coloring rules**.

For $a, b \in V(G) \setminus L_S$, let $Q_{ab}$ denote the unique path in the inner skeleton $B_1$ between $h(a)$ and $h(b)$. For each such pair of vertices $(a, b)$, we will maintain a subgraph $P_{ab}$ of $G$. Each $P_{ab}$ is initialized to $\emptyset$. After the application of each coloring rule, we will apply a **path rule** for each pair $(a, b)$, which (possibly) adds some newly colored edges to $P_{ab}$. We say that an edge in $B_1$ is colored if its representatives in $G$ are colored (we will make sure that for a 2-edge, either both representatives are colored or both are uncolored at any point of time). Whenever an edge in $B_1$ gets colored by a coloring rule and if it is in $Q_{ab}$, we make sure that we add exactly one of its representatives to $P_{ab}$ in the subsequent path rule. Whenever it happens during a path rule that two edges $u_1v_1$ and $u_2v_2$ are in $P_{ab}$ such that both $v_1$ and $v_2$ are in some $T \in \mathcal{T}$, but $u_1$ and $u_2$ are not in $T$, then we add the path $T_{v_2u_1}$ to $P_{ab}$ (if it is not already included). Similarly, if it happens that there is an edge $uv$ in $P_{ab}$ such that $v, a \in V(T)$ (or $b \in V(T)$ resp.) but $u \notin V(T)$, we add the path $T_{va}$ ($T_{vb}$ resp.) to $P_{ab}$. Also, if both $a$ and $b$ are in the same tree $T$, then we add the path $T_{ab}$ to $P_{ab}$ (during Path Rule [1] below). Thus, when all the coloring rules and path rules have been applied, we will have that for all $a, b \in V(G) \setminus L_S$, it holds that $P_{ab}$ is a path between $a$ and $b$. We will prove that $P_{ab}$ is also a rainbow path. For this, we will maintain the following invariant.

**Invariant 1** For each pair $a, b \in V(G) \setminus L_S$, no two edges in $P_{ab}$ have the same color.

We will prove that the invariant still holds after each path rule. Since new edges are added to $P_{ab}$ only during path rules, this means that the invariant always holds. We also maintain the following three auxiliary invariants. But they are rather straightforward to check from the coloring and path rules and hence we will not explicitly prove them.

**Invariant 2** For any 2-edge in $B$, either both representatives of it are colored or both are uncolored.

A vertex in $B_1$ is said to be **completed** if all the incident edges on it in $B_1$ are colored and is said to be **incomplete** otherwise.

**Invariant 3** For an incomplete tree-vertex $x_T$, the colors of $E(T)$ are disjoint from the colors of the rest of the graph $G$.

**Invariant 4** A nonempty subset of internal edges of a tree $T$ is contained in $P_{ab}$ only if a representative of each edge in $Q_{ab}$ that is incident on $x_T$ (there can be at most two of such edges as $Q_{ab}$ is a path) is in $P_{ab}$.

Now, we start with the coloring and path rules.

**Coloring Rule 1** Color all the edges in $F$ with distinct colors $1, 2, \ldots, f - t$.

**Path Rule 1** For each $a, b \in V(G) \setminus L_S$, if $a$ and $b$ are in the same tree $T$ for some $T \in \mathcal{T}$, then add the path $T_{ab}$ to $P_{ab}$. 
It is easy to see that Invariant 1 is satisfied after the above Path rule as the color of each edge is distinct so far.

For each tree \( T \in \mathcal{T} \), we designate a color in \( [f - t + 1, f] \) as its **surplus color**, denoted by \( s(T) \). More specifically, the surplus color of \( i^{th} \) tree in \( \mathcal{T} \) is defined as the color \( f - t + i \). Also, the colors of the edges of \( T \) (colored by Coloring Rule 1) are called the **internal colors** of \( T \).

**Coloring Rule 2** For each 1-edge \( \bar{uw} \) in \( B \): if \( u \) is a tree vertex, then color \((uv)_1\) with \( s(f_T(u))\); otherwise, i.e., if \( u \) is not a tree vertex, by Corollary 2, there is at least one child of \( u \) in \( B_1 \) that is a tree vertex; pick one such tree vertex \( x_T \) and color \((uv)_1\) with color \( s(T) \).

Note that after Coloring Rule 2 any tree vertex \( x_T \) such that \( T \) is just a single vertex, is completed.

**Path Rule 2** Do the following for each \( a, b \in V(G) \setminus L_S \). For each 1-edge \( e \) in \( Q_{ab} \), add \((e)_1\) to \( P_{ab} \). Next, we add edges inside trees as follows.

- If for some tree \( T \), it holds that \( a \in V(T) \) and there is a 1-edge \( ux_T \) in \( Q_{ab} \), then add the path \( P_{wa} \) to \( P_{ab} \), where \( w \) is the foot of the edge \( ux_T \) in \( T \).
- If for some tree \( T \), it holds that \( b \in V(T) \) and there is a 1-edge \( vx_T \) in \( Q_{ab} \), then add the path \( P_{wb} \) to \( P_{ab} \), where \( w \) is the foot of the edge \( vx_T \) in \( T \).
- If for some tree \( T \), there are two 1-edges \( ux_T \) and \( vx_T \) in \( Q_{ab} \), add the path \( P_{wz} \) to \( P_{ab} \), where \( w \) is the foot of the edge \( ux_T \) in \( T \) and \( z \) the foot of the edge \( vx_T \) in \( T \).

**Lemma 8.** Invariant 1 is satisfied so far. Moreover, each color is used at most for one edge in \( G \).

**Proof.** It is clear that during Coloring Rule 1 all edges are colored distinct. In Coloring Rule 2, we use the surplus colors, which are disjoint from the colors used in Coloring rule 1. It is also not difficult to see that during Coloring Rule 2 the surplus color of a tree vertex is used for only one 1-edge. \( \square \)

For a colored edge \( e \in E(G) \), we define \( c(e) \) to be the color of \( e \). For a subgraph \( G' \) of \( G \), we define \( c(G') \) to be the set of colors used in \( E(G') \). We call the number of 2-edges of \( B_1 \) incident on a vertex, the **2-edge degree** of it. For any two vertices \( u \) and \( v \), the connected component of \( B_1 \setminus u \) containing \( v \) is denoted by \( ST(u, v) \). Note that \( ST(u, v) \) is a subtree of \( B_1 \). The closest (breaking ties arbitrarily) tree vertex to \( v \) in \( ST(u, v) \) in \( B_1 \) is denoted by \( CT(u, v) \). Note that at least one tree vertex exists in \( ST(u, v) \) because all leaves of \( B_1 \) are tree vertices by Corollary 3. Also note that if \( v \) is a tree vertex, then \( CT(u, v) = v \).

**Coloring Rule 3** For each tree vertex \( x_T \) with 2-edge degree at least 4 (see Figure 3 for an illustration). Let \( q \) be the 2-edge degree of \( x_T \). Let \( w_0, w_1, w_2, \ldots, w_{q-1} \) be the other endpoints of the 2-edges incident on \( x_T \). For \( i \in [0, q-1] \), let \( x_T i := CT(x_T, w_i) \) and let \( c_i := s(T_i) \). For each \( i \in [0, q-1] \), color the edge \((x_T w_i)_1\) with \( c_{((i+2) \mod q)} \) and the edge \((x_T w_i)_2\) with \( c_{((i+3) \mod q)} \).

The following lemma follows from the way in which we have colored the edges incident on \( x_T \) in Coloring Rule 3.

**Lemma 9.** For each tree vertex \( x_T \) on which Coloring Rule 3 has been applied as above, for all distinct \( i, j \in [0, q-1] \), there is a rainbow path from \( w_i \) to \( w_j \) in \( G \) that uses only the colors from \((\{c_0, c_1, \ldots, c_{q-1}\} \setminus \{c_i, c_j\}) \cup c(T)\). Moreover, for any \( i \in [q-1] \) and some \( u \in V(T) \), there is a rainbow path in \( G \) from \( u \) to \( w_i \) that uses only colors from \((\{c_0, c_1, \ldots, c_{q-1}\} \setminus \{c_i\}) \cup c(T)\).

**Proof.** Let \( u_i \) and \( v_i \) be the endpoints in \( T \) of \((x_T w_i)_1\) and \((x_T w_i)_2\) respectively, for each \( i \in \{0, 1, \ldots, q - 1\} \). First, we prove that there is a rainbow path from \( w_i \) to \( w_j \) with the required colors as claimed by the
lemma. Suppose for the sake of contradiction that there was no such path. Consider the following three paths between \( w_i \) and \( w_j \): 

\[
P := w_iu_iTu_{u_i}u_jw_j, \quad P' := w_iv_iTv_{v_i}v_jw_j, \quad \text{and} \quad P'' := w_iv_iTv_{v_i}u_jw_j.
\]

By our assumption, each of these paths, is either not a rainbow path, or uses a color that is not in \( \{c_0, c_1, \ldots, c_{q-1}\} \setminus \{c_i, c_j\} \cup c(T) \). Also, from Coloring rules 3 and 4, we know that the only colors that are not in \( \{c_0, c_1, \ldots, c_{q-1}\} \setminus \{c_i, c_j\} \cup c(T) \) that any of these three paths can use are \( c_i \) and \( c_j \). Thus, each of \( P, P' \) and \( P'' \) is either not a rainbow path or uses \( c_i \) or \( c_j \). However, we know that the paths \( Tu_{u_i}, Tv_{v_i}, \) and \( Tu_{u_j} \) are all rainbow paths due to Coloring Rule 1 and moreover the colors used by them are disjoint from \( \{c_0, \ldots, c_{q-1}\} \). For the path \( P \), this means that either \( c(w_iu_i) = c(u_jw_j) \) or \( c(w_iu_i), c(u_jw_j) \cap \{c_i, c_j\} \neq \emptyset \). That is, either \( c(i+2) \mod q = c(j+2) \mod q \) or \( \{c(i+2) \mod q, c(j+2) \mod q\} \cap \{c_i, c_j\} \neq \emptyset \). That is, either \( i = j \) or \( (i + 2) \mod q, (j + 2) \mod q \cap \{i, j\} \neq \emptyset \). But we know that \( (i + 2) \mod q \neq i \) and that \( (j + 2) \mod q \neq j \). Therefore, either \( (i + 2) \mod q = j \) or \( (j + 2) \mod q = i \). Without loss of generality assume that \( (i + 2) \mod q = j \).

By using the same reasoning as above for path \( P' \), we derive that either \( (i + 3) \mod q = j \) or \( (j + 3) \mod q = i \). Since we already have that \( (i + 2) \mod q = j \), it should be the latter case, i.e., \( (j + 3) \mod q = i \).

Now consider the third path \( P'' \). We have that either \( c(w_iv_i) = c(u_jw_j) \) or \( c(w_iv_i), c(u_jw_j) \cap \{c_i, c_j\} \neq \emptyset \). That is, either \( c(i+3) \mod q = c(j+2) \mod q \) or \( \{c(i+3) \mod q, c(j+2) \mod q\} \cap \{c_i, c_j\} \neq \emptyset \). That is, either \( (i + 3) \mod q = (j + 2) \mod q \) or \( \{(i + 3) \mod q, (j + 2) \mod q\} \cap \{i, j\} \neq \emptyset \). Substituting that \( (i + 3) \mod q = ((i + 2) \mod q + 1) \mod q = (j + 1) \mod q \) and that \( i = (j + 3) \mod q \), we get that either \( (j + 1) \mod q = (j + 2) \mod q \) or \( \{(j + 1) \mod q, (j + 2) \mod q\} \cap \{j + 3 \mod q, j \} \neq \emptyset \). Since \( j, (j + 1) \mod q, (j + 2) \mod q, \) and \( (j + 3) \mod q \) are distinct for \( q \geq 4 \), we have a contradiction.

Next, we prove the second part of the lemma, i.e., we prove that there is a rainbow path from \( u \) to \( w_i \) with the colors claimed by the lemma. Suppose for the sake of contradiction that there was no such path. Consider the path \( P''' := w_iu_iTu_{u_i}u_j \). We know that the path \( Tu_{u_i} \) uses only colors from \( c(T) \) and is rainbow, and that the edge \( w_ju_i \) is colored \( c(i+2) \mod q \). Also, \( c(i+2) \mod q \neq c_i \) as \( (i + 2) \mod q \neq i \). Thus \( P \) is a rainbow path and uses only the colors in \( \{c_0, c_1, \ldots, c_{q-1}\} \setminus \{c_i\} \cup c(T) \). \( \square \)

**Path Rule 3** For each \( x_T \) on which Coloring Rule 3 has been applied as above and for each \( a, b \in V(G) \setminus L_S \) such that \( Q_{ab} \) contains \( x_T \) (we say that the path rule is being applied on the pair \( (x_T, P_{ab}) \)), do the following.

**Case 1:** There are two 2-edges incident on \( x_T \) in \( Q_{ab} \).

Fig. 5: Illustration of Coloring Rule 3 applied on a tree vertex \( x_T \) with 2-edge degree 4. Here, \( c_i = s(T_i) \).
Let $w_i$ and $w_j$ be the neighbors of $x_T$ in $Q_{ab}$. Add to $P_{ab}$ the rainbow path from $w_i$ to $w_j$ as given by Lemma 9.

**Case 2:** There is one 2-edge and one 1-edge incident on $x_T$ in $Q_{ab}$.

Let $x_Tw_i$ be the 2-edge. Let $u$ be the endpoint in $T$ of the representative of the 1-edge. There is a rainbow path from $w_i$ to $u$ as given by Lemma 9. Add this path to $Q_{ab}$. (Note that the representative of the 1-edge has already been added to $P_{ab}$ during Path Rule 3.)

**Case 3:** $x_T$ is an endpoint of $Q_{ab}$ and the only edge incident on $x_T$ in $Q_{ab}$ is a 2-edge.

Let $w_i$ be the neighbor of $x_T$ in $Q_{ab}$. We know one of $a$ or $b$ is in $T$. From this vertex ($a$ or $b$ whichever is in $T$) to $w_i$, there is a rainbow path as given by Lemma 9. Add this path to $P_{ab}$.

The following lemma follows from Lemma 9 and Path Rule 3.

**Lemma 10.** Suppose for some $a,b \in V(G) \setminus L_S$ and for some tree $T' \in \mathcal{T}$, $P_{ab}$ contains an edge $e$ that was colored with $s(T')$ during the application of Coloring Rule 3 on some tree vertex $x_{T'}$. Then, $T' \neq T$ and $Q_{ab}$ does not intersect $ST(x_T, x_{T'})$.

**Proof.** Since $s(T')$ was used during the application of Coloring Rule 3 on $x_{T'}$, the vertex $x_{T'}$ should have been taken as $x_T^i$ (in Coloring Rule 3) for some $i$ and $s(T')$ was taken as $c_i$ (in Coloring Rule 3). Since $T_i \neq T$, it is clear that $T' \neq T$. Suppose $Q_{ab}$ intersects $ST(x_T, x_{T'})$ for the sake of contradiction. That is, $Q_{ab}$ intersects $ST(x_T, x_T')$. Then the color $c_i$ was not used in Path Rule 3 according to Lemma 9. That means $e$ was not colored with $c_i$, which is a contradiction. □

**Lemma 11.** Invariant Ⅲ is not violated during Path Rule 3.

**Proof.** Suppose Invariant Ⅲ is violated during the application of Path Rule 3 on the pair $(x_T, P_{ab})$. Then there exist edges $e$ and $e'$ in $P_{ab}$ having the same color after the application of the path rule. We can assume without loss of generality that $e$ was added during the application of Path Rule 3 on $(x_T, P_{ab})$. That means $e$ was colored during the application of Coloring Rule 3 on $x_T$. Then either $e \in E(T)$ or $h(e) = w_i x_T$ for some $i \in [0, q - 1]$. Since each color in $c(T)$ has been used only in one edge in $G$, we have that $h(e) = w_i x_T$ for some $i \in [0, q - 1]$ and hence $c(e) = s(T_j)$ for some $j \in [0, q - 1] \setminus i$. Also $Q_{ab}$ does not intersect $ST(x_T, x_T')$ by Lemma 10. Since the application of Path Rule 3 on $(x_T, P_{ab})$ added a rainbow path to $P_{ab}$, the edge $e'$ was not added during this application. Since each color in $c(F)$ has been used for only one edge in $G$ so far, we know that $e'$ was not added during Path Rule 3. Hence, the following two cases are exhaustive and in both cases we derive a contradiction.

**Case 1:** $e'$ was added during the application of Path Rule 3 on $(x_{T'}, P_{ab})$ for some tree $T' \neq T$.

Since $P_{ab}$ contains $e'$, we have that $Q_{ab}$ contains $h(e')$. Since $e'$ was added during the application of Path Rule 3 on $(x_{T'}, P_{ab})$, either $e' \in E(T')$ or $h(e')$ is incident on $x_{T'}$. In either case, $x_{T'}$ is in $Q_{ab}$. Since $Q_{ab}$ does not intersect $ST(x_T, x_{T'_j})$, we have that $x_{T'}$ is not in $ST(x_T, x_{T'_j})$. This implies that $\text{dist}_{B_3}(x_{T'_j}, x_{T'}) < \text{dist}_{B_3}(x_T, x_{T'_j})$. But then during the application of Coloring Rule 3 on $x_{T'}$, the color $s(T_{j'_i})$ would never be used as $x_{T'_i} \neq ST(x_T, v)$ for any vertex $v$. Thus, the color of $e'$ is not $s(T_{j'_i})$. But we know that $c(e') = c(e) = s(T_j)$, a contradiction.

**Case 2:** $e'$ was added during the application of Path Rule 2 on $P_{ab}$.

This means $e'$ is the representative of a 1-edge and was colored during Coloring Rule 2. Since $e'$ is colored with $s(T_{j'_i})$, we have that $h(e')$ should either be the outgoing edge of $x_{T'_i}$ or the outgoing edge of the parent of $x_{T'_i}$, from Coloring Rule 2. This implies that $h(e')$ is in $ST(x_T, x_{T'_j})$, as the parent of $x_{T'_j}$ is a non-tree vertex. But then $Q_{ab}$ does not contain $h(e')$ as $Q_{ab}$ does not intersect $ST(x_T, x_{T'_j})$. Thus $P_{ab}$ does not contain $e'$, which is a contradiction. □
**Coloring Rule 4** For each tree vertex $x_T$ with 2-edge degree exactly 3 (see Figure 6), let $w_1$, $w_2$, and $w_3$ be the other endpoints of the three 2-edges incident on $x_T$. Further, for $i \in \{1, 2, 3\}$, let $x_{T_i} = \mathcal{C}(x_T, w_i)$, let $u_i$ and $v_i$ be the foots of $x_{T_i}w_i$ in $T$, let $P_i := T_{u_i,v_i}$, and let $c_i := s(T_i)$.

**Case 1:** There exists an edge $uv$ in $T$ such that the cut $(V_1, V_2)$ induced by $uv$ in $T$ is such that for all $i \in \{1, 2, 3\}$, $|V_1 \cap \{u_i, v_i\}| = 1$ and $|V_2 \cap \{u_i, v_i\}| = 1$. (For an illustration, see Figure 7).

Without loss of generality, let $u_i$ and $v_i$ be the foots of $x_{T_i}w_i$ in $V_1$ and $V_2$ respectively for each $i \in \{1, 2, 3\}$. Let $c$ be the color of $uv$. Color $w_1w_1$ with $c_3$, $v_1w_1$ with $c_2$, $w_2w_2$ with $c$, $v_2w_2$ with $c_1$, $w_3w_3$ with $c$, and $v_3w_3$ with $c$, as shown in Figure 7.

**Case 2:** There exist distinct edges $e_1, e_2, e_3$ such that $e_i \in E(P_i)$ for each $i \in \{1, 2, 3\}$. (For an illustration, see Figure 8 (a)).

Color both the representatives of $x_{T_i}w_i$ with the color of $e_i$ for each $i \in \{1, 2, 3\}$.

**Case 3:** Case 1 and 2 do not apply.

Because Case 1 and 2 do not apply, there exist $i, j \in \{1, 2, 3\}$ such that $E(P_i) \cap E(P_j) = \emptyset$, because otherwise $E(P_1) \cap E(P_2) \cap E(P_3) \neq \emptyset$ using the Helly property of trees, and then any edge in this intersection qualifies as uv of Case 1. So, without loss of generality assume that $E(P_1) \cap E(P_2) = \emptyset$. Also, note that $E(P_3) \subseteq E(P_1) \cup E(P_2)$ because otherwise Case 2 applies. So, without loss of generality assume that $E(P_3) \cap E(P_1) = \emptyset$. But then $E(P_3) \cap E(P_1) = E(P_1)$ and $P_1$ consists of a single edge so that Case 2 does not apply. Let this edge be $e_1$. Note that $e_1 = u_1v_1$. Furthermore, at least one of the end-vertices of $P_1$ and $P_3$ coincide so that Case 2 does not apply. Thus, assume without loss of generality that $u_1 = v_3$. Let $e_2$ be any edge in $P_2$. Without loss of generality assume that $v_1$ is the closer vertex among $u_1$ and $v_1$ to path $P_2$ in $T$. The two possible scenarios in this case are shown in Figure 8 (b) and (c). Color $w_1u_1$ and $w_1v_1$ with $c(e_1)$, $w_2u_2$ and $w_2v_2$ with $c(e_2)$, $w_3u_3$ with $c(e_2)$ and $w_3v_3$ with $c(e_1)$.

The following lemma follows from the way in which we have colored the edges incident on $x_T$ in Coloring Rule 4.

**Lemma 12.** For each tree vertex $x_T$ on which Rule 4 has been applied as above, for distinct $i, j \in \{1, 2, 3\}$, there is a rainbow path from $w_i$ to $w_j$ in $G$ that uses only the colors from $\{c_1, c_2, c_3\} \setminus \{c_i, c_j\} \cup c(T)$. Also, for any $i \in \{1, 2, 3\}$, and any $z \in V(T)$, there is a rainbow path from $z$ to $w_i$, that uses only the colors from $\{c_1, c_2, c_3\} \setminus \{c_i\} \cup c(T)$.

---

*Fig. 6: A scenario in which Coloring Rule 4 is applicable on $x_T$. |

---

We use the following Helly property of trees: if $T_1, T_2, \ldots, T_k$ are subtrees of a tree $T$ that pairwise intersect each other on at least one edge, then there is an edge of $T$ that is common to all of $T_1, T_2, \ldots, T_k$. |
Proof. We demonstrate the required paths in each of the three cases of Coloring Rule 4.

Case 1: Between $w_1$ and $w_2$, there is the rainbow path $w_1u_1T_{u_1u_2}u_2w_2$ that uses only the colors in $c(T) \cup \{c_3\}$. Between $w_1$ and $w_3$, there is the rainbow path $w_1v_1T_{v_1v_3}v_3w_3$ that uses only the colors in $c(T) \cup \{c_2\}$. Between $w_2$ and $w_3$, there is the rainbow path $w_2v_2T_{v_2v_3}v_3w_3$ that uses only the colors in $c(T) \cup \{c_1\}$.

Now consider any vertex $z \in V(T)$. Suppose $z \in V_1$. Between $z$ and $w_1$, there is the rainbow path $T_{zu_1}u_1w_1$ that uses only the colors in $c(T) \cup \{c_3\}$. Between $z$ and $w_2$, there is the rainbow path $T_{zu_3}v_2w_2$ that uses only the colors in $c(T) \cup \{c_1\}$. Between $z$ and $w_3$, there is the rainbow path $T_{zu_3}u_3w_3$ that uses only the colors in $c(T)$.

Now, suppose $z \in V_2$. Between $z$ and $w_1$, there is the rainbow path $T_{zu_1}v_1w_1$ that uses only the colors in $c(T) \cup \{c_2\}$. Between $z$ and $w_2$, there is the rainbow path $T_{zu_3}v_2w_2$ that uses only the colors in $c(T) \cup \{c_1\}$. Between $z$ and $w_3$, there is the rainbow path $T_{zu_3}v_3w_3$ that uses only the colors in $c(T)$.

Case 2: First we show the path between $w_1$ and $w_2$. By Observation 4, either $T_{u_1u_2}$ or $T_{u_1v_3}$ does not contain the edge $e_2$. If $T_{u_1u_2}$ does not contain $e_2$, then the path $w_1u_1T_{u_1u_2}u_2w_2$ is a rainbow path and uses only the colors in $c(T)$; otherwise (i.e., if $T_{u_1v_3}$ does not contain $e_2$) then the path $w_1u_1T_{u_1v_3}v_3u_2w_2$...
is a rainbow path and uses only the colors in $c(T)$. The required paths between $w_2$ and $w_3$, and between $w_1$ and $w_2$ can be shown in a similar way.

Now for any vertex $z$ in $T$, we show the required path between $z$ and $w_1$. By Observation 1, either $T_{zu_1}$ or $T_{zu_3}$ does not contain the edge $e_1$. If $T_{zu_1}$ does not contain $e_1$, then the path $T_{zu_1}w_1w_q$ is a rainbow path and uses only the colors in $c(T)$; otherwise (i.e., if $T_{zu_3}$ does not contain $e_1$) then the path $T_{zu_3}v_1w_1$ is a rainbow path and uses only the colors in $c(T)$. The required paths between $w_2$ and $z$, and between $w_3$ and $z$ can be shown in a similar way.

**Case 3:** First we show the required path between $w_1$ and $w_2$. By Observation 1, either $T_{w_1v_2}$ or $T_{w_3v_2}$ does not contain the edge $e_2$. Observe that both $T_{w_1v_2}$ and $T_{w_3v_2}$ does not contain the edge $e_1$ as $v_1$ is closer than $u_1$ to $P_2$, as mentioned in the Coloring Rule. Hence, if $T_{w_1v_2}$ does not contain $e_2$, then the path $w_1v_1T_{v_1w_2}w_2w_2$ is a rainbow path that uses only the colors in $\{c(e_2), c(e_1)\} \cup c(T)$; and otherwise (i.e., if $T_{w_3v_2}$ does not contain $e_2$), the path $w_1v_1T_{v_1w_2}v_3w_3$ is a rainbow path that uses only the colors in $\{c(e_2), c(e_1)\} \cup c(T)$.

Since $u_1 = u_3$, there is the path $w_1u_1w_3$ between $w_1$ and $w_3$ that uses only the colors in $\{c(e_1), c(e_2)\}$. Next, we show the required path between $w_3$ and $w_2$. By Observation 1, either $T_{w_3u_2}$ or $T_{w_3v_2}$ does not contain the edge $e_2$. Let $v'$ be the vertex in $\{u_2, v_2\}$ such that $T_{v_3v'}$ does not contain edge $e_2$. We show that the path $w_3v_3T_{u_3v'}w_2w_2$ is the required path between $w_3$ and $w_2$. We know $w_3v_3$ is colored $c(e_1)$ and $w_2v'$ is colored $c(e_2)$. So, it is sufficient to show that $c(e_1)$ and $c(e_2)$ does not appear in $T_{v_3v'}$. For this, it is sufficient to prove that $e_1$ and $e_2$ is not in $T_{v_3v'}$. Since we picked $v'$ such that $T_{v_3v'}$ does not contain $e_2$, it only remains to prove that $e_1$ is not in $T_{v_3v'}$. Suppose for the sake of contradiction that $e_1$ is in $T_{v_3v'}$. That means both $v_1$ and $u_1$ are in $T_{v_3v'}$. We know that $v_1$ is closer than $u_1$ to $P_3$, as mentioned in the Coloring Rule. Hence, $v_1$ is closer than $u_1$ to $v'$ in $T$. This also implies that $v_1$ is closer than $u_1$ to $v'$ in $T_{v_3v'}$. Then $u_1$ is closer than $v_1$ to $v_3$ in $T_{v_3v'}$ and hence also in $T$. But then $P_3$ contains edges that are not in both $P_1$ and $P_3$, a contradiction.

Now consider any vertex $z \in V(T)$. For each $i \in \{1, 3\}$, let $v_i'$ be the closer vertex among $u_i, v_i$ to $z$, and let $P_i$ be the path $T_{zv_i}v_i'w_i$. We show that $P_i$ is the required path from $z$ to $w_i$ for $i \in \{1, 3\}$. The path from $z$ to $v_i'$ does not contain $e_1$. Also, the edge $w_1v_i'$ is colored with $c(e_1)$. Hence $P_i$ is a rainbow path from $z$ to $w_i$ that uses only colors in $c(T)$. Now consider path $P_3$. First consider the case when $v_3' = u_3 = u_1$. Then $e_2$ is not in $T_{zv_3'}$, because otherwise either $P_3$ contains edges that are not in $P_1 \cup P_2$ or $w_1$ is closer than $v_1$ to $P_2$. Since $w_3u_3$ is colored $c(e_2)$, the path $P_3$ satisfies the requirements. Now consider the case when $v_3' = v_1$. Then $e_1$ is clearly not in $T_{zv_3'}$. Since $v_3u_3$ is colored $c(e_1)$, the path $P_3$ satisfies the requirements.

Now we show the required path from $z$ to $w_2$. By Observation 1, either $T_{w_2v_2}$ or $T_{w_2u_2}$ does not contain the edge $e_2$. Let $v_2'$ be the vertex in $\{u_2, v_2\}$ such that $T_{w_2v_2}$ does not contain edge $e_2$. Then the path $T_{zv_2'}v_2'w_2$ is the required path between $z$ and $w_2$, as $v_2'w_2$ is colored $c(e_2)$.

**Path Rule 4** For each $x_T$ on which Coloring Rule 4 has been applied as above and for each $P_{ab}$ such that $Q_{ab}$ contains $x_T$ (we say that the rule is being applied on the pair $(x_T, P_{ab})$), do the following.

**Case 1:** $x_T$ has two 2-edges incident in $Q_{ab}$.

Let $w_i$ and $w_j$ be the neighbors of $x_T$ in $Q_{ab}$. Add to $P_{ab}$ the rainbow path from $w_i$ to $w_j$ as given by Lemma 12.

**Case 2:** $x_T$ has exactly one 2-edge and exactly one 1-edge incident in $Q_{ab}$.

Let $x_Tw_i$ be the 2-edge and let $z$ be the endpoint in $T$ of the 1-edge. Add to $P_{ab}$ the rainbow path from $w_i$ to $z$ as given by Lemma 12.

**Case 3:** $x_T$ is an endpoint of $Q_{ab}$ and has one 2-edge incident in $Q_{ab}$.

Let $w_i$ be the neighbor of $x_T$ in $Q_{ab}$. We know one of $a$ or $b$ is in $T$. From this vertex (a or $b$, whichever is in $T$) to $w_i$, there is a rainbow path as given by Lemma 12. Add this path to $P_{ab}$. 

\[ \square \]
The following lemma follows from Lemma 12 and Path Rule 4. The proof is similar to that of Lemma 10 and is omitted.

**Lemma 13.** Suppose for some $a,b \in V(G) \setminus L_S$ and for some tree $T' \in \mathcal{T}$, $P_{ab}$ contains an edge $e$ that was colored with $s(T')$ during the application of Coloring Rule 4 on some tree vertex $x_T$. Then, $T' \neq T$ and $Q_{ab}$ does not intersect $ST(x_T, x_{T'})$.

**Lemma 14.** Invariant 1 is not violated during Path Rule 4.

*Proof.* Suppose for the sake of contradiction that Invariant 1 is violated during the application of Path Rule 4 on the pair $(x_T, P_{ab})$ as above. Then there exist edges $e$ and $e'$ in $P_{ab}$ having the same color. We can assume without loss of generality that $e$ was colored during the application of Coloring Rule 4 on $x_T$. This means $e \in E' := E(T) \cup R$, where $R$ is defined as the set of representatives of $w_1x_T, w_2x_T,$ and $w_3x_T$. Since the application of Path Rule 4 on $(x_T, P_{ab})$ added a rainbow path to $P_{ab}$, the edge $e'$ was not added during this application and hence $e' \notin E'$. Each color in $c(T)$ have been used only in $E'$ so far. That means $c(e) = c(e') \notin c(T)$. Hence $e \in E' \setminus E(T) = R$. Without loss of generality assume that $e$ is a representative of $w_i x_T$. Now, $c(e) = s(T_j)$ where $j \in \{2,3\}$. Without loss of generality assume that $c(e) = s(T_2)$. This also means $c(e') = s(T_2)$. That means $e'$ was colored during Coloring Rules 2, 3 or 4. Hence the following two cases are exhaustive and in each case we prove a contradiction.

**Case 1:** $e'$ was colored during the application of Coloring Rules 3 or 4 on $x_{T'}$, for some tree $T' \neq T$.

Since $P_{ab}$ contains $e'$, we have that $Q_{ab}$ contains $h(e')$. Since $e'$ was colored during the application of Coloring Rules 3 or 4 on $x_{T'}$, either $e' \in E(T')$ or $h(e')$ is incident on $x_{T'}$, and hence $x_{T'}$ is in $Q_{ab}$. Since $Q_{ab}$ does not intersect $ST(x_T, x_{T_2})$ by Lemmas 10 and 13, we have that $x_{T'}$ is not in $ST(x_T, x_{T_2})$. Then $\text{dist}_{\tilde{B}_1}(x_{T'}, x_T) < \text{dist}_{\tilde{B}_1}(x_{T'}, x_{T_2})$. But then during the application of Coloring Rule 3 or 4 on $x_{T'}$, the color $s(T_2)$ would never be used as $x_{T_2} \neq CT(x_{T'}, v)$ for any vertex $v$. Thus, the color of $e'$ is not $s(T_2)$. But we know that $c(e') = c(e) = s(T_2)$, a contradiction.

**Case 2:** $e'$ was colored during the application of Coloring Rule 2.

This means $e'$ is the representative of a 1-edge. Since $e'$ is colored with $s(T_2)$, we have that $h(e')$ should either be the outgoing edge of $x_{T_2}$ or the outgoing edge of the parent of $x_{T_2}$, from Coloring Rule 2. This implies that $h(e')$ is in $ST(x_T, x_{T_2})$, as the parent of $x_{T_2}$ is a non-tree edge. But then $Q_{ab}$ does not contain $h(e')$ as $Q_{ab}$ does not intersect $ST(x_T, x_{T_2})$, by Lemmas 10 and 13. Thus $P_{ab}$ does not contain $e'$, which is a contradiction.

\[\square\]
**Coloring Rule 5** For each non-tree vertex u with degree at least 3 in B₁ (see Figure 9), let q be the number of children of u (note that q ≥ 2 as degree of u is at least 3), let u₁, u₂, ..., u₉ be the children of u and let xₜ be the outgoing edge from u in B₁. If uv is a 1-edge, due to Coloring Rule 2, we know that there exist an i ∈ [q] such that uᵢ is a tree vertex (and hence Tᵢ = fₜ(uᵢ)), and uv is colored with s(Tᵢ). Hence, if uv is a 1-edge, assume without loss of generality that u₁ is a tree vertex (and hence xₜ₁ = u₁) and that uv is colored with s(T₁).

- If u₁u is uncolored (then u₁u is a 2-edge due to Coloring Rule 2), implying that u₁ is a tree vertex and hence T₁ = fₜ(u₁), then color (u₁u)₁ with s(T₁) and (u₁u)₂ with s(T₂).
- For each 2 ≤ i ≤ q, if uᵢu is uncolored (then uᵢu is a 2-edge due to Coloring Rule 2), implying that uᵢ is a tree vertex and hence Tᵢ = fₜ(uᵢ), then color both its representatives with s(Tᵢ).
- Let uw be the outgoing edge from u. If uw is uncolored (in which case it is a 2-edge due to Coloring Rule 2), then color (uw)₁ with s(T₁) and (uw)₂ with s(T₂).

**Path Rule 5** For each non-tree vertex u on which Coloring Rule 5 has been applied as above and for each Pᵤₜ, such that Qᵤₜ contains u (we say that the rule is being applied on the pair (u, Pᵤₜ)), execute the following two parts (in the mentioned order).

**Part 1**

- If Qᵤₜ contains edge u₁u and u₁u is colored during the application of Coloring Rule 5 on u, do the following. If the other neighbor (if any) of u in Qᵤₜ is u₂, then add (u₁u)₁ (which has color s(T₁)) to Pᵤₜ. Otherwise, add (u₁u)₂ (which has color s(T₂)) to Pᵤₜ.
- For each i ∈ [2, q], if Qᵤₜ contains edge uᵢu and uᵢu is colored during the application of Coloring Rule 5 on u, add (uᵢu)₁ (which has color s(Tᵢ)) to Pᵤₜ.
- If Qᵤₜ contains edge uw and uw is colored during the application of Coloring Rule 5 on u: if the other neighbor (if any) of u in Qᵤₜ is u₁ and u₁u is a 1-edge, then add (uw)₁ (which has color s(T₁)) to Pᵤₜ; otherwise add (uw)₁ (which has color s(T₁)) to Pᵤₜ.

**Part 2**

- For each tree vertex xₜ such that the degree of xₜ in h(Pᵤₜ) became 2 during the addition of above edges in Part 1, let x and y be the endpoints in T of the two edges of Pᵤₜ incident on T. Add Tₓᵧ to Pᵤₜ.
- For each tree vertex xₜ ∈ {h(a), h(b)} such that the degree of xₜ in h(Pᵤₜ) became 1 during the addition of above edges in Part 1, let x be the endpoint in T of the edge of Pᵤₜ incident on T. If xₜ = h(a), add Tₓₐ to Pᵤₜ; otherwise (i.e., if xₜ = h(b)), add Tₓₜ to Pᵤₜ.

**Lemma 15.** Invariant 7 is not violated during Path Rule 5.

**Proof.** Suppose Invariant 1 is violated during the application of Path Rule 5 on the pair (u, Pᵤₜ) as above. Then there exist edges e and e’ in Pᵤₜ having the same color. We can assume without loss of generality that e was colored during the application of Coloring Rule 5 on u. Suppose e was added during Part 2 of Path Rule 5. Observe that if we add a path inside a tree T in Part 2, then xₜ was incomplete before the application of Coloring Rule 5. By Invariant 3, this implies that the internal colors of T were not used anywhere else so far. Thus, the color of e is unique, in particular c(e’) ≠ c(e), a contradiction. Thus, the edge e was not added during Part 2. Then e was added during Part 1 and hence c(e) = c(e’) = s(Tᵢ) for some i ∈ [q]. Then e’ was colored during one of Coloring Rules 5, 4, 3, or 2

**Case 1:** e’ was colored during the Coloring Rule 5.

Note that during the application of Path Rule 5 on (u, Pᵤₜ), we have added at most two edges to Pᵤₜ. And, if we have added two edges, they are of different color. Thus e’ was not added to Pᵤₜ during
the application of Path Rule 5 on \((u, P_{ab})\) and hence was not colored during the application of Coloring Rule 5 on \(u\). So, \(e'\) was colored during the application of Coloring Rule 5 on some non-tree vertex \(u^1 \neq u\). Notice that for any tree \(T \in \mathcal{T}\), \(s(T)\) is used during the application of Coloring Rule 5 only when the rule is applied to an ancestor of \(x_T\) in \(B_1\). Hence, both \(u\) and \(u^1\) are ancestors of \(x_T\). Without loss of generality, assume that \(u^1\) is closer than \(u\) to \(x_T\). Then, \(u\) cannot have any non-tree vertices as children because otherwise \(x_T \neq \mathcal{C}(u, u_i)\). Then, the only edges colored during the application of Coloring Rule 5 on \(u\), are the representatives of \(uv\). Thus \(h(e) = uv\).

**Case 1.1** \(e = (uv)\).

We know that \(e = (uv)\) is colored with \(s(T_2)\) by Coloring Rule 5. Thus, \(c(e') = c(e) = s(T_2)\) and \(T_1 = T_2\). Since the edge \((uv)\) is added during application of Path Rule 5 on \((u, P_{ab})\), the neighbors of \(u\) in \(Q_{ab}\) are \(v\) and \(u_1\), by Path Rule 5. Since \(u^1 \in Q_{ab}\), we have that \(u^1\) is a descendant of \(u_1\) and not \(u_2\) in \(B_1\). This implies \(T_1 = T_1 \neq T_2\), a contradiction.

**Case 1.2** \(e = (uv)\).

Since the edge \((uv)\) is added during application of Path Rule 5 on \((u, P_{ab})\), either \(u_1\) is not a neighbor of \(u\) in \(Q_{ab}\), or \(uu_1\) is a 2-edge, by Path Rule 5. But \(uu_1\) cannot be a 2-edge as both \(u\) and \(u_1\) are non-tree vertices. (Recall that we said all children of \(u\) are non-tree vertices in Case 1). Hence \(u_1\) is not a neighbor of \(u\) in \(Q_{ab}\). Since \(u^1 \in Q_{ab}\), this implies that \(T_1 \neq T_1\), and hence \(c(e) = c(e') = s(T_1) \neq s(T_1)\). But we know that \(e = (uv)\) is colored with \(s(T_1)\), by Coloring Rule 5. Thus, we have a contradiction.

**Case 2:** \(e'\) was colored during the Coloring Rules 4 or 3.

Let \(T'\) be the tree on which \(e'\) is incident. Then \(e'\) was colored with \(s(T_1)\) during the application of Coloring Rules 4 or 3 on \(x_T\). Then \(Q_{ab}\) does not intersect \(\mathcal{ST}(T', T_1)\) due to Lemmas 13 and 10. Since \(x_T = \mathcal{C}(u, u_i)\), there is no other tree-vertex in the path from \(u\) to \(x_T\). Thus, \(u\) is in \(\mathcal{ST}(T', T_1)\). Hence, we have that \(u\) is not in \(Q_{ab}\). We know that \(e\) is adjacent on \(u\) as every edge colored during the application of Coloring Rule 5 on \(u\) is incident on \(u\). But then \(e \notin P_{ab}\) as \(u\) is not in \(Q_{ab}\). This is a contradiction.

**Case 3:** \(e'\) was colored during the Coloring Rule 2.

This means that \(e'\) is a 1-edge.

**Case 3.1** \(h(e') = uv\).

In the case when \(uv\) is a 1-edge, we selected \(u_1\) during Coloring Rule 5 in such a way that \(c(uv) = s(T_1)\). Thus \(c(e) = c(e' = uv) = s(T_1)\). The only edges that can be potentially colored with \(s(T_1)\) during the application of Coloring Rule 5 on \(u\) are \((uv)\) and \((uu_1)\). Since \(e\) and \(e'\) are distinct we have \(e = (u_1u)\). But since \(uv\) is in \(Q_{ab}\), we would have added \((u_1u)\) and not \((u_1u)\) to \(P_{ab}\) during Path Rule 5. Thus we have a contradiction.

**Case 3.2** \(h(e') \neq uv\).

Then \(e'\) is on the path between \(x_T\) and \(u\). Also, \(x_T\) is not a child of \(u\). Then, the only possibility for \(e\) to have color \(s(T_1)\) is if \(i = 2\) and \(e = (u_1u)\). Then \(Q_{ab}\) contains both \(u_1\) and \(u_2\). In that case, we would have added \((u_1u)\) and not \((u_1u)\) to \(P_{ab}\) during Path Rule 5. Hence, \(e \neq (u_1u)\), a contradiction.

\[\square\]

**Coloring Rule 6** For each incomplete tree vertex \(x_T\) having 2-edge degree exactly 1; let \(e\) be the only 2-edge incident on \(x_T\), pick an edge \(e_1\) in the foot-path of \(e\), color the representatives of \(e\) with the color of \(e_1\).

**Path Rule 6** For each tree vertex \(x_T\) on which Coloring Rule 6 has been applied as above and for each \(P_{ab}\) such that \(Q_{ab}\) contains \(h(e)\) (we say that the path rule is being applied on the pair \((x_T, P_{ab})\)), do the following.
We pick vertex $w$ as follows. If $a \in V(T)$, let $w := a$, and if $b \in V(T)$ let $w := b$. (Note that both $a$ and $b$ cannot be in $T$ as $Q_{ab}$ contains $h(e)$. If $a, b \notin V(T)$ then there is an edge $e_2 \neq e$ of $Q_{ab}$ incident on $x_T$. Furthermore, since $e$ is the only 2-edge incident on $x_T$, the edge $e_2$ is a 1-edge. In this case, let $w$ be the endpoint of $(e_2)_1$ in $T$.

By Observation 7, there is a path in $T$ that excludes $e_1$, from $w$ to one of the foots of $e$. Let this foot be $z$. Add the path in $T$ between $w$ and $z$ to $P_{ab}$. Also add to $P_{ab}$ the representative of $e$ having $z$ as its endpoint in $T$.

**Lemma 16.** Invariant 7 is not violated during Path Rule 6.

Proof. Let $E_N$ be the set of new edges added to $P_{ab}$ during the application of Path Rule 6 to $(x_T, P_{ab})$ and let $E_O$ be the set of already included edges in $P_{ab}$ before this application. Suppose Invariant 1 is violated for the sake of contradiction. Then either there are two edges in $E_N$ with the same color or $c(E_N) \cap c(E_O) \neq \emptyset$. Recall that $E_N$ consists of $E(T_{wz})$ and a representative of $e$, say $(e)_j$. Note that $c(((e)_j)_j) = c(e_1)$ by Coloring Rule 6. So, all the edges in $E_N$ are colored from $c(T)$, the internal colors of $T$. Recall that $e_1$ is not in $T_{wz}$ by our choice of $z$. Thus the edges in $E_N$ all have distinct colors. So, it has to be the case that $c(E_O) \cap c(E_N) \neq \emptyset$. Since $c(E_N) \subseteq c(T)$, this implies that $c(E_O) \cap c(T) \neq \emptyset$. Let $d$ be an edge in $E_O$ with color in $c(T)$. Since the representative of at least one edge of $Q_{ab}$ incident on $x_T$ (namely $e$) was not added to $P_{ab}$ before the application of Path Rule 6, we have that $E(T) \cap E_O = \emptyset$ by Invariant 4. Thus $d \notin E(T)$ but has color in $c(E(T))$ and $d \in E_O$. By Invariant 3, this implies that $x_T$ was complete before the application of Coloring rule 6 making the rule not applicable on $x_T$, a contradiction. □

**Coloring Rule 7** For each tree vertex $x_T$ such that the 2-edge degree of $x_T$ is exactly 2 and $E(T)$ contains at least 2 edges, let $e_1$ and $e_2$ be the 2-edges incident on $x_T$, let $w$ and $z$ be the other endpoints of $e_1$ and $e_2$, respectively, and let $P_1$ and $P_2$ be the foot-paths of $e_1$ and $e_2$, respectively.

**Case 1:** $|E(P_1 \cup P_2)| \geq 2$. (See Figure 10 (a)).

Pick distinct edges $e$ and $e'$ from $P_1$ and $P_2$ respectively. If $(e_1)_1$ and $(e_1)_2$ are uncolored, color them with color of $e$ and if $(e_2)_1$ and $(e_2)_2$ are uncolored, color them with color of $e'$.

**Case 2:** Case 1 does not hold. (See Figure 10 (b)).

Clearly, $P_1$ and $P_2$ both are a single edge and they are the same edge. Let this edge be $e = wz$. Pick any other edge $e'$ in $T$ (such an edge exists because we said that the rule is applicable only if $E(T)$ contains at least two edges). Without loss of generality, assume that $e'$ is closer to $v$ than $u$ in $T$. If $uw$ and $vw$ are uncolored, color them with color of $e$. If $uz$ and $vz$ are uncolored, color them with colors of $e'$ and $e$, respectively.
**Lemma 17.** Consider a tree vertex $x_T$ on which Coloring Rule 7 has been applied as above. There is a rainbow path in $G$ from $w$ to $z$ using only the colors in $c(T)$. Also, from any vertex $x \in V(T)$, there is a rainbow path to both $w$ and $z$ using only the colors in $c(T)$.

**Proof.** If Case 2 of Coloring Rule 7 has been applied then this is rather easy to see as follows. The path $wuz$ is a rainbow path from $w$ to $z$. Also for the second statement, note that at least one of $T_{xw}$ or $T_{xz}$ avoids $e$. Then at least one of the paths $T_{xw}$ followed by $uw$ or $uz$, or $T_{xz}$ followed by $vw$ or $vz$, is a rainbow path (it avoids $e$ and if it contains $e'$, then it is the second case and $c(e')$ is used only once).

So, it only remains to prove the lemma when Case 1 of Coloring Rule 7 is applied. Let $w_1, w_2$ be the foots of $e_1$ and $z_1, z_2$ be the foots of $e_2$. To prove the first statement, it is sufficient to prove that at least one of the four paths $T_{w_1 z_1}, T_{w_1 z_2}, T_{w_2 z_1}$ and $T_{w_2 z_2}$ contains neither $e$ nor $e'$. Given this, it is easy to show the necessary rainbow path from $w$ to $z$: if the path $T_{w_1 z_j}$ contains neither $e$ nor $e'$ then the path $w_1 w T_{w_1 z_j} z_j z$ is the required path. So for the sake of contradiction assume that each $T_{w_1 z_j}$ contains either $e$ or $e'$. Without loss of generality assume that $T_{w_1 z_1}$ contains $e$. Let $y$ be the last vertex on $T_{w_1 z_1}$ that is in $P_1$ (while going from $w_1$ to $z_1$). Now, $T_{y z_1}$ and $T_{yw}$ do not contain $e$ and hence $T_{z_1 w_2} = T_{y z_1} \cup T_{yw}$ does not contain $e$. This implies that $T_{z_1 w_2}$ contains $e'$. Let $y'$ be the last vertex on $T_{z_1 w_2}$ that is in $P_2$ (while going from $z_1$ to $w_2$). Now, $T_{y' z_2}$ and $T_{w_2 y'}$ contains neither $e$ nor $e'$ and hence $T_{w_2 z_2} = T_{y' z_2} \cup T_{w_2 y'}$ contains neither $e$ nor $e'$.

To prove the second statement, observe that there is a rainbow path from $x$ to either $w_1$ or $w_2$ not containing $e$, and a rainbow path from $x$ to either $z_1$ or $z_2$ not containing $e'$, due to Observation 1.

**Path Rule 7** For each tree vertex $x_T$ on which Coloring Rule 7 has been applied as above and for each $P_{ab}$ such that $Q_{ab}$ contains at least one of $e_1$ and $e_2$ (we say that the path rule is being applied on the pair $(x_T, P_{ab})$), do the following.

If $Q_{ab}$ contains both $e_1$ and $e_2$ then let $y_1 := w$ and $y_2 := z$. If $Q_{ab}$ contains only $e_1$ and not $e_2$ then let $y_1 := w$. If $Q_{ab}$ contains only $e_2$ and not $e_1$ then let $y_1 := z$. If $a \in V(T)$, let $y_2 := a$, and if $b \in V(T)$, let $y_2 := b$. (Note that both $a$ and $b$ cannot be in $T$ as $Q_{ab}$ contains $e_1$ or $e_2$). If $a, b \notin V(T)$ and only one of $e_1, e_2$ is in $Q_{ab}$, then there is an edge $e'' \notin \{e_1, e_2\}$ incident on $x_T$ in $Q_{ab}$; and since $e_1$ and $e_2$ are the only 2-edges incident on $x_T$, the edge $e''$ is a 1-edge; let $y_2$ be the endpoint of $(e'')_1$ in $T$. Add to $P_{ab}$ the path between $y_1$ and $y_2$ given by Lemma 17.

**Lemma 18.** Invariant 1 is not violated during Path Rule 7

**Proof.** Let $E_N$ be the set of new edges added to $P_{ab}$ during the application of Path Rule 7 to $(x_T, P_{ab})$ and let $E_O$ be the set of already included edges in $P_{ab}$ before this application. Suppose Invariant 1 is violated for the sake of contradiction. Then either there are two edges in $E_N$ with the same color or $c(E_N) \cap c(E_O) \neq \emptyset$. By Lemma 17, all edges in $E_N$ are colored from $c(T)$ and have distinct colors. So, it has to be the case that $c(E_O) \cap c(E_N) \neq \emptyset$. Since $c(E_N) \subseteq c(T)$, this implies that $c(E_O) \cap c(T) \neq \emptyset$. Let $d$ be an edge in $E_O$ with color in $c(T)$. The representative of at least one edge of $Q_{ab}$ incident on $x_T$ was not in $P_{ab}$ before the application of Path Rule 7 on $(x_T, P_{ab})$, because otherwise the path rule is not applicable on $(x_T, P_{ab})$. Then, by Invariant 1, we have that $E(T) \cap E_O = \emptyset$. Thus $d \notin E(T)$ but has color in $c(E(T))$. But then by Invariant 3, we have that $x_T$ was completed before the application of Coloring Rule 7 thereby making the rule not applicable on $x_T$, which is a contradiction.

**Coloring Rule 8** For each incomplete tree vertex $x_T$ having degree at least 3 in $B_1$: We can assume that Coloring Rules 3, 4, 6, 7 are not applicable on $x_T$ as otherwise $x_T$ would have been completed. If $x_T$ had at least three 2-edges incident on it, then Coloring Rule 3 or 4 would have been applicable on $x_T$. 


If it had 2-edge degree 1, then Coloring Rule 2 would have been applicable on $x_T$. If it had 2-edge degree 0, then it would have been completed after Coloring Rule 2. Hence, we can assume that $x_T$ has 2-edge degree exactly 2. Now, if $|E(T)| \geq 2$, Coloring Rule 3 becomes applicable on $x_T$. Hence, we can assume that the tree $T$ is just an edge. Let $uv$ be this edge. Let the two 2-edges incident on $x_T$ be $yx_T$ and $zx_T$.

**Case 1:** Both $yx_T$ and $zx_T$ are incoming to $x_T$ (see Figure 11 (a)).

Then the outgoing edge of $u$ in $B_1$ is a 1-edge, say $\bar{x_T}w$. Assume without loss of generality that its representative is $vw$. Let $c_1 = s(T)$ and $c_2$ be the color of $uv$. Note that $vw$ is colored with $c_1$ due to Coloring Rule 2. If $yu$ and $yv$ are uncolored, color them with $c_1$ and $c_2$ respectively. If $zu$ and $zv$ are uncolored, color both of them with $c_2$.

**Case 2:** One of the 2-edges, say $yx_T$ is outgoing from $x_T$ (see Figure 11 (b)).

Let $c_2$ be the color of $uv$ and $c_3$ be the color of representative of any 1-edge incoming on $x_T$. Note that at least one such 1-edge exists as the degree of $x_T$ is at least 3. If $yu$ and $yv$ are uncolored, color them with $c_3$ and $c_2$ respectively. If $zu$ and $zv$ are uncolored, color both of them with $c_2$.

**Path Rule 8** For each tree vertex $x_T$ on which Coloring Rule 2 has been applied as above and for each $P_{ab}$ such that $Q_{ab}$ contains at least one of $yx_T$ and $zx_T$ (we say that the path rule is being applied on the pair $(x_T, P_{ab})$), do the following:

- If $Q_{ab}$ contains both $yx_T$ and $zx_T$ then add $yu$ and $uz$ to $P_{ab}$.
- If $Q_{ab}$ contains only $yx_T$ and not $zx_T$ then let $y_1 := y$. If $Q_{ab}$ contains only $zx_T$ and not $yx_T$ then let $y_1 := z$. If $a \in V(T)$, let $y_2 := a$, and if $b \in V(T)$ let $y_2 := b$. (Note that both $a$ and $b$ cannot be in $T$ as $Q_{ab}$ contains $yx_T$ or $zx_T$). If $a, b \notin V(T)$ and only one of $yx_T, zx_T$ is in $Q_{ab}$ then there is an edge $e'' \notin \{yx_T, zx_T\}$ incident on $x_T$ in $Q_{ab}$. Further, since $yx_T$ and $zx_T$ are the only 2-edges incident on $x_T$, the edge $e''$ is a 1-edge. Note that $(e'')_1$ is already added to $P_{ab}$ in Path rule 2. Let $y_2$ be the endpoint of $(e'')_1$ in $T$.

Note that in all cases $y_2 \in \{u, v\}$ and $y_1 \in \{y, z\}$. Hence, the edge $y_1y_2$ exists. Add the edge $y_1y_2$ to $P_{ab}$.

**Lemma 19.** Invariant 2 is not violated during Path Rule 8.
Proof. Suppose the invariant is violated. Then there exist edges $e$ and $e'$ in $P_{ab}$ having the same color. We can assume without loss of generality that $e$ was colored during the application of Coloring Rule 8 on $(x_T, P_{ab})$. We added at most two edges during the application of Path Rule 8 on $x_T$ and if we added two edges we have made sure they have distinct colors. Thus $e'$ was not added during the application of Path Rule 8 on $(x_T, P_{ab})$. The colors that are possible for $e$ are $c_1, c_2$ and $c_3$ according to Coloring Rule 8.

Case 1: $c(e) = c(e') = c_2$.

Recall $c_2 = c(uv)$. If $e' = uv$, then by Invariant 4 the representatives of all the edges of $Q_{ab}$ incident on $x_T$ are in $P_{ab}$ even before the application of Path Rule 8 on $(x_T, P_{ab})$. Then Path Rule 8 is not applicable on $(x_T, P_{ab})$. Thus $e' \notin E(T)$ but $c(e') \in c(E(T))$. Then by Invariant 3 $x_T$ was completed before the application of Coloring Rule 8 making the rule not applicable on $x_T$. Thus, such an $e'$ does not exist.

Case 2: $c(e) = c(e') = c_1 = s(T)$.

This means $e = yu$ and that Case 1 of Coloring Rule 8 (see Figure 11 (a)) was applied on $x_T$. The only coloring rules so far that use surplus colors are Coloring Rules 8, 5, 4, 3, and 2.

Case 2.1 $e'$ was colored during Coloring Rule 2.

Note that this means $e'$ is a 1-edge and the only way $e'$ can have color $s(T)$ is if $e' = vw$. But, in Path Rule 8 we add $yu = e$ to $P_{ab}$ only when $vw$ is not in $Q_{ab}$. Thus we have a contradiction.

Case 2.2 $e'$ was colored during Coloring Rules 4 or 3.

Let $T'$ be the tree on which $e'$ is incident. Since $e'$ was colored with $s(T)$ during Coloring Rules 4 or 3, we know that $Q_{ab}$ does not intersect $ST(x_{T'}, x_T)$ due to Lemmas 13 and 10. Then $Q_{ab}$ does not contain $x_T$ and hence $P_{ab}$ does not contain $e$, which is a contradiction.

Case 2.3 $e'$ was colored during Coloring Rule 5.

Then $h(e')$ is a 2-edge in the path from $x_T$ to root of $B_1$. Since $e'$ is in $Q_{ab}$, this means that $x_Tw$ is in $Q_{ab}$. But then by Path Rule 8 we would have added $yu$ instead of $yu = e$ to $P_{ab}$, a contradiction.

Case 2.4 $e'$ was colored during Coloring Rule 8.

The only application of Coloring Rule 8 that uses $s(T)$ is the application on $x_T$. But since $e'$ was not colored during this application, we have a contradiction.

Case 3: $c(e) = c(e') = c_3$.

This means $e = uy$ and that Case 2 of Coloring Rule 8 was applied on $x_T$. Let $x$ be the neighbor of $x_T$ such that $xxT$ is the 1-edge incident on $x_T$ whose representative is colored with $c_3$. By Coloring Rule 2, there exist a tree $T'$ that is a descendant of $x$ such that $s(T') = c_3$. The only coloring rules so far that use surplus colors are Coloring Rules 8, 5, 4, 3, and 2.

Case 3.1 $e'$ was colored during Coloring Rule 2.

This means $e'$ is a 1-edge. Since $xxT$ is the only 1-edge with color $s(T')$ by Lemma 8, we have that $e' = xxT$. Hence, $xxT$ is in $Q_{ab}$. But if $xxT$ is in $Q_{ab}$, we would have added $vy$ and not $uy$ in Path Rule 8. This is a contradiction to $e = uy$.

Case 3.2 $e'$ was colored during Coloring Rules 4 or 3.

Let $T''$ be such that $e'$ is adjacent on $x_{T''}$. Since $e'$ was colored with $s(T')$ during Coloring Rules 4 or 3, we have that $Q_{ab}$ does not intersect $ST(x_{T''}, x_{T'})$ due to Lemmas 13 and 10. Since $P_{ab}$ contains $e$ that is incident on $x_T$, we have that $Q_{ab}$ contains $x_T$. This implies that $x_T \notin ST(x_{T''}, x_{T'})$ implying that $x_{T''}$ is on the path between $x_T$ and $x_{T'}$. But then $s(T'')$ and not $s(T')$ would have been used to color $xxT$, a contradiction.

Case 3.3 $e'$ was colored during Coloring Rule 5 or 8.
Since $e'$ is colored with $s(T')$, by Coloring Rule 5 and 8 this implies $e'$ is in $\text{ST}(x_T, x_{T'}) = \text{ST}(x_T, x)$, implying that $Q_{ab}$ contains $x$. But then we would have added $vy$ and not $uy = e$ to $P_{ab}$ according to Path Rule 8, a contradiction.

The following Lemma follows from the previous coloring rules.

**Lemma 20.** Consider an edge $e$ in $B_1$ that remains uncolored after the application of Coloring Rules 7 through 8. Let $x_T$ and $v$ be the endpoints of $e$ (Note that due to Coloring Rule 2, $e$ is a 2-edge, and hence one of its endpoints is a tree-vertex). Then, both $x_T$ and $v$ have degree exactly 2 in $B_1$, both edges incident on $x_T$ are 2-edges, and $T$ consists of a single edge.

**Proof.** Suppose $u \in \{v, x_T\}$ had degree not equal to 2 in $B_1$. First, suppose the degree was greater than 2. Then Coloring Rule 8 or 5 would have been applicable on $u$, and hence $u$ would have been completed. Therefore, $u$ has degree 1 in $B_1$. By Corollary 3 every leaf of $B_1$ is a tree vertex. Hence, $u$ is a tree vertex and $u = x_T$. But then Coloring Rule 6 would have been applicable on $x_T$, and $x_T$ would have been completed. Thus, $e$ is already colored, which is a contradiction. Hence, $x_T$ and $v$ have degree 2 in $B_1$.

Now, suppose $x_T$ has only one 2-edge incident in $B_1$. Then, Coloring Rule 6 would have been applied on $x_T$ and $x_T$ would have been completed. Thus, both edges incident on $x_T$ in $B_1$ are 2-edges. If $T$ contained at least two edges, Coloring Rule 7 would have been applied on $x_T$ and $x_T$ would have been completed. Hence, $T$ contains only one edge.

**Coloring Rule 9** For each tree vertex $x_T$ with exactly one uncolored 2-edge $e$ incident on it: note that it follows by Lemma 20 that the tree $T$ comprises of a single edge $e'$. Let $e = vx_T$ and $e' := u_1u_2$. Color $(e)_1 = vu_1$ and $(e)_2 = vu_2$ with the color of $e'$.

**Path Rule 9** For each tree vertex $x_T$ on which Coloring Rule 9 has been applied as above and for each $P_{ab}$ such that $Q_{ab}$ contains $e$ (we say that the path rule is being applied on the pair $(x_T, P_{ab})$), do the following:

First we pick vertex $w \in V(T) = \{u_1, u_2\}$ as follows: if $a \in V(T)$, let $w := a$; if $b \in V(T)$ let $w := b$; (note that both $a$ and $b$ cannot be in $T$ as $Q_{ab}$ contains $e$); if $a, b \notin V(T)$ then there is an edge $e_2 \neq e$ of $Q_{ab}$ incident on $x_T$; furthermore, since $e$ is the only uncolored edge incident on $x_T$, a representative of the edge $e_2$ is already in $P_{ab}; let w$ be the endpoint in $T$ of this representative of $e_2$.

Add $vw$ to $P_{ab}$.

**Lemma 21.** Invariant 4 is not violated during Path Rule 9.

**Proof.** The edge added to $P_{ab}$ during the application of Path Rule 9 to $(x_T, P_{ab})$ has color $c(e') \in c(T)$. If the invariant is violated, then there was an edge $e''$ in $P_{ab}$ already with color $c(e')$. By Invariant 4, $e'$ was not already in $P_{ab}$ as the edge $e$ incident on $x_T$ is in $Q_{ab}$ and the representative of $e$ was not added to $P_{ab}$ before. Thus $e'' \neq e'$ but $c(e'') = c(e')$. Since $x_T$ was incomplete before the application of current coloring rule, by Invariant 3 none of the colors in $c(T)$ were used before anywhere outside of $T$. So, such an $e''$ does not exist, a contradiction.

**Lemma 22.** Consider a 2-edge $e$ incident on tree vertex $x_T$ that remains uncolored after the application of Rules 3 to 9. Then, $x_T$ have degree exactly 2 in $B_1$, $T$ contains only one edge, and the other edge incident on $x_T$ is an uncolored 2-edge.

**Proof.** By Lemma 20 it follows that $x_T$ has degree exactly 2 in $B_1$, $T$ contains only one edge, and the other edge incident on $x_T$ is a 2-edge. If this other 2-edge is colored, then Coloring Rule 9 would have been applied on $x_T$ and $x_T$ would have been completed.
Coloring Rule 10 For each incomplete tree vertex \(x_T\) whose parent’s outgoing edge is a 2-edge: (See Figure 12 for an Illustration). Let \(v_1\) be the parent of \(x_T\). From Lemma 22, it follows that \(x_T\) has degree exactly 2 in \(B_1\), has one incoming and one outgoing 2-edge incident on it, both the 2-edges are uncolored, and the tree \(T\) is just a single edge. Let \(e_1\) and \(e_2\) respectively be the outgoing and incoming 2-edges of \(x_T\). Let \(e\) be the only edge in \(T\). Let \(v_2\) be the other end point of \(e_2\). Let \(u_1\) be the endpoint of \((e_1)_1\) and \((e_2)_1\) in \(T\). Let \(u_2\) be the endpoint of \((e_1)_2\) and \((e_2)_2\) in \(T\). From Lemma 20, we know that \(v_1\) and \(v_2\) have degree exactly 2. Let \(v_1\) or \(v_2\) be the outgoing 2-edge from \(v_1\) and let \(v_2\) be the incoming edge on \(v_2\) in \(B_1\).

Color \((e_2)_1\) and \((e_2)_2\) with the color of \(e\), and color \((e_1)_1\) and \((e_1)_2\) with \(s(T)\).

Path Rule 10 For each tree vertex \(x_T\) on which Coloring Rule 10 has been applied as above and for each \(P_{ab}\) such that \(Q_{ab}\) contains \(e_1\) or \(e_2\) (we say that the path rule is being applied on the pair \((x_T, P_{ab})\)), do the following.

- If \(Q_{ab}\) contains both \(e_1\) and \(e_2\), add \(v_1\) to \(P_{ab}\).
- If \(Q_{ab}\) contains exactly one edge among \(e_1\) and \(e_2\), then either \(a\) or \(b\) is in \(V(T)\). Also both of them cannot be in \(V(T)\). Let \(z\) be the one among \(a\) or \(b\) that is in \(V(T)\). If \(Q_{ab}\) contains \(e_1\), add \(v_1\) to \(P_{ab}\); otherwise, i.e., if \(Q_{ab}\) contains \(e_2\), add \(v_2\) to \(P_{ab}\).

Lemma 23. Invariant 7 is not violated during Path Rule 10.

Proof. Suppose for the sake of contradiction that the invariant is violated. Then there exist distinct edges \(d_1\) and \(d_2\) in \(P_{ab}\) having the same color. We can assume without loss of generality that \(d_1\) was colored during the application of Coloring Rule 10 on \(x_T\). We added at most two edges during the application of Path Rule 10 on \(x_T\), and in the cases where we added two edges, the two edges have distinct colors. Thus, \(d_2\) was not added during the application of Path Rule 10 on \(x_T\) and hence was not colored during the application of Coloring Rule 10 on \(x_T\).

The colors that are possible for \(d_1\) are \(s(T)\) and \(c(e)\).

Case 1: \(c(d_1) = c(d_2) = c(e)\).

This is not possible since the color of \(e\) has not been used to color any other edges so far by Invariant 3 and \(e\) is not in \(P_{ab}\) by Invariant 4.

Case 2: \(c(d_1) = c(d_2) = s(T)\).

This means \(h(d_1) = v_1\) or \(v_2\). The only coloring rules so far that use the surplus colors of trees are Coloring Rules 2, 3, 4, 5, 8, and 10. Hence, \(d_2\) was colored with \(s(T)\) during one of them.

Case 2.1 \(d_2\) was colored during Coloring Rule 2.
This means that \( d_2 \) is a 1-edge. According to Coloring Rule 2, the only 1-edge that can be colored with \( s(T) \) is either the outgoing edge of \( x_T \) or the outgoing edge of the parent of \( x_T \). However, both of them are 2-edges and hence we have a contradiction.

**Case 2.2** \( d_2 \) was colored during Coloring Rules 1 or 3

Let \( T'' \) be such that \( d_2 \) is adjacent on \( x_{T''} \). Then, by Lemmas 13 and 10 we know that \( Q_{ab} \) does not intersect \( ST(x_{T''}, x_T) \), in particular \( Q_{ab} \) does not contain \( x_T \). Since \( h(d_1) = v_1 x_T \), this implies \( P_{ab} \) does not contain \( d_1 \), which is a contradiction.

**Case 2.3** \( d_2 \) was colored during application of Coloring Rule 5

From Coloring Rule 5, this implies that \( d_2 \) was colored during application of Coloring Rule 5 on some ancestor \( v' \) of \( x_T \) such that there are no other tree vertices in the path from \( x_T \) to \( v' \). Then, the only possibility for \( v' \) is \( v_1 \) as the parent of \( v_1 \) is a tree vertex. However, we know that \( v_1 \) has degree 2 in \( B_1 \), and hence Coloring Rule 5 could not have been applied on \( v_1 \). Thus, we have a contradiction.

**Case 2.4** \( d_2 \) was colored during application of Coloring Rule 8

Since \( d_2 \) is colored with \( s(T) \) during Coloring Rule 8. Case 1 of the rule (see Coloring Rule 8) was applied on \( x_T \) and hence the outgoing edge from \( x_T \) is a 1-edge. However, this is a 2-edge and hence we have a contradiction.

**Case 2.5** \( d_2 \) was colored during Coloring Rule 10

Since \( d_2 \) was not colored during the application of Coloring Rule 10 on \( x_T \), we have that \( d_2 \) was colored during the application of Coloring Rule 10 on some \( x_{T''} \neq x_T \). But then \( d_2 \) is not colored with \( s(T) \), a contradiction. \( \square \)

**Coloring Rule 11** For each incomplete tree vertex \( x_T \): from Lemma 24, it follows that \( x_T \) has degree exactly 2 in \( B_1 \), has one incoming and one outgoing 2-edge incident on it, both the 2-edges are uncolored, and the tree \( T \) is just a single edge. Let \( e_1 \) and \( e_2 \) respectively be the outgoing and incoming 2-edges of \( x_T \). Let \( v_1 \) be the other end point of \( e_1 \) and \( v_2 \) be the other end point of \( e_2 \). Let \( e = u_1 u_2 \) be the only edge in \( T \). Without loss of generality, \( u_1 \) be the endpoint of \((e_1)_1 \) and \((e_2)_1 \) in \( T \), and \( u_2 \) be the endpoint of \((e_1)_2 \) and \((e_2)_2 \) in \( T \). From Lemma 20, we know that \( v_1 \) and \( v_2 \) have degree exactly 2. Let \( v_1 y \) be the outgoing edge from \( v_1 \) and \( wv_2 \) be the incoming edge on \( v_2 \) in \( B_1 \). We have that \( v_1 y \) is a 1-edge because otherwise Coloring Rule 10 would have been applicable on \( x_T \), and \( x_T \) would have been already completed.

**Case 1:** There is an edge between \( v_1 \) and \( v_2 \) in \( G \). (See Figure 13 (a) for an illustration).

Color the representatives of \( e_1 \) and \( e_2 \) with \( c(e) \). Color \( v_1 v_2 \) with \( c(e) \). We say that \( v_1 v_2 \) is a shortcut edge. Note that shortcut edges are the only colored edges in \( G \) that are not representatives of edges in \( B \).

**Case 2:** Case 1 does not apply. (See Figure 13 (b) for an illustration).

We will prove in Lemma 30 that \( wv_2 \) is a 2-edge in this case. Let \( T' = f_T(w) \). Color \((e_1)_1 \) and \((e_1)_2 \) with \( s(T') \) and color \((e_2)_1 \) and \((e_2)_2 \) with color of \( e \).

**Path Rule 11** For each tree vertex \( x_T \) on which Coloring Rule 11 has been applied as above and for each \( P_{ab} \) such that \( Q_{ab} \) contains \( e_1 \) or \( e_2 \) (we say that the path rule is being applied on the pair \((x_T, P_{ab})\)), do the following.

- If \( Q_{ab} \) contains both \( e_1 \) and \( e_2 \): if \( v_1 v_2 \in E(G) \), add \( v_1 v_2 \) to \( P_{ab} \); otherwise add \( v_1 u_1 \) and \( v_2 u_1 \) to \( P_{ab} \).
- If \( Q_{ab} \) contains exactly one edge among \( e_1 \) and \( e_2 \), then either \( a \) or \( b \) is in \( V(T) \). Also, both of them cannot be in \( V(T) \). Let \( z \) be the one among \( a \) or \( b \) that is in \( V(T) \). If \( Q_{ab} \) contains \( e_1 \), add \( v_1 z \) to \( P_{ab} \). If \( Q_{ab} \) contains \( e_2 \), add \( v_2 z \) to \( P_{ab} \).
Lemma 24. Invariant [11] is not violated during Path Rule [11].

Proof. Suppose for the sake of contradiction that the invariant is violated. Then there exist distinct edges $d_1$ and $d_2$ in $P_{ab}$ having the same color. We can assume without loss of generality that $d_1$ was colored during the application of Coloring Rule [11] on $x_T$. We added at most two edges during the application of Path Rule [11] on $x_T$, and in the cases where we added two edges, the two edges have distinct colors. Thus $d_2$ was not added during the application of Path Rule [11] on $x_T$ and hence was not colored during Coloring Rule [11] on $x_T$.

The colors that are possible for $d_1$ are $c(e)$ and $s(T')$.

Case 1: $c(d_1) = c(d_2) = c(e)$.

This is not possible since the color of $e$ has not been used to color any other edges so far by Invariant [3] and $e$ is not in $P_{ab}$ by Invariant [4].

Case 2: $c(d_1) = c(d_2) = s(T')$.

This means $h(d_1) = e_1$ and that Case 2 of Coloring Rule [11] was applied on $x_T$. The only coloring rules so far that use the surplus colors of trees are Coloring Rules [2], [3], [4], [5], [8], [10], and [11]. Hence, $d_2$ was colored with $s(T')$ during one of them.

Case 2.1 $d_2$ was colored during Coloring Rule [2].

This means $d_2$ is a 1-edge and $h(d_2)$ is either $x_Tv_2$ or $v_2x_T$. But since both $x_Tv_2$ and $v_2x_T$ are 2-edges (since Case 2 of Coloring Rule [11] was applied on $x_T$), this is not possible.

Case 2.2 $d_2$ was colored during Coloring Rules [4] or [3].
Let $T''$ be the tree such that $d_2$ is adjacent on $x_{T''}$. By Lemmas 13 and 10, we know that $Q_{ab}$ does not intersect $\mathcal{ST}(x_{T''}, x_{T'})$. Then $x_T$ is not in $\mathcal{ST}(x_{T''}, x_{T'})$. This implies $x_{T''}$ is in the path from $x_T$ to $x_{T'}$. But the only vertex in the path from $x_T$ to $x_{T'}$ is $v_2$, a non-tree vertex. Thus, we have a contradiction.

Case 2.3 $d_2$ was colored during application of Coloring Rule 5

From Coloring Rule 5, this implies that $d_2$ was colored during application of Coloring Rule 5 on some ancestor $v'$ of $x_{T'}$ such that there are no other tree vertices in the path from $x_{T'}$ to $v'$. Then, the only possibility for $v'$ is $v_2$ as the parent of $v_2$ is a tree vertex. However, we know that $v_2$ has degree 2 in $B_1$, and hence Coloring Rule 5 could not have been applied on $v_2$. Thus, we have a contradiction.

Case 2.4 $d_2$ was colored during application of Coloring Rule 8

Since $c(d_2) = s(T')$, from Coloring Rule 10 we get that $d_2$ was colored during the application of Coloring Rule 10 on $x_{T'}$. In Lemma 31, we will prove that Coloring Rule 10 was not applied on $x_{T'}$. Thus, we have a contradiction.

Case 2.5 $d_2$ was colored during application of Coloring Rule 11

Since $c(d_2) = s(T')$, from Coloring Rule 11, we get that $d_2$ was colored during the application of Coloring Rule 11 on $x_T$. Moreover, $h(d_2) = e_1$. Recall that we have $h(d_1) = e_1$ too. Since we picked only one representative of $e_1$ into $P_{ab}$ by Path Rule 11 we have that $d_1 = d_2$. This is a contradiction to the fact that $d_1$ and $d_2$ are distinct.

Now we proceed towards proving Lemmas 30 and 31 that were used above. For this we need to prove some auxiliary lemmas first.

Lemma 25. Let $v$ be a non-tree vertex and $x_{T_1}$ be a tree-vertex that is a descendant of $v$ in $B_1$. If $vx_{T_1}$ is a 2-edge in $H$ then $x_{T_1}$ is a child of $v$ in $B_1$.

Proof. See Figure 14 for an illustration of the proof. Suppose $x_{T_1}$ is not a child of $v$ in $B_1$. Let $\overrightarrow{x_{T_1}z}$ be the outgoing edge of $x_{T_1}$ in $B_1$. Let $B'$ be the skeleton obtained by deleting $\overrightarrow{x_{T_1}z}$ from $B$ and adding $\overrightarrow{x_{T_1}v}$. Going from $B$ to $B'$, the number of 2-edges is non-decreasing, the degree of $v$ increases, the degree of $z$ decreases, and the degree of all other vertices remains same. Since $v$ is at a smaller level than $z$ in both
Let $x_T$ be a vertex on which Coloring Rule 11 is being applied. Let $v_1$ be as defined in Coloring Rule 11. The vertex $v_1$ has no 2-edge in $H$ to any vertex except $x_T$.

**Proof.** See Figure 15 for an illustration of the proof. Suppose for the sake of contradiction that $v_1$ has a 2-edge in $H$ to a tree vertex $x_{T_1} \in V(H) \setminus \{x_T\}$. The edge $v_1y$ is a 1-edge as otherwise Coloring Rule 10 would have been applicable on $x_T$ and $x_T$ would have been completed already. Thus $x_{T_1} \neq y$. Then the edge $v_1x_{T_1}$ is not in $B_1$ as $y$ and $x_T$ are the only neighbors of $v_1$ in $B_1$. Since $x_{T_1}$ is a tree vertex, it is not in $L_S$ (recall that $L_S$ is the set of non-tree vertices of $H$). Thus, since the edge $v_1x_{T_1}$ is not in $B_1$, it is not in $B$ also (recall $B_1 = B \setminus L_S$). Thus, $v_1x_{T_1} \in E(H) \setminus E(B)$. Also $x_{T_1}$ is not a descendant of $v_1$ due to Lemma 25. Thus, $x_{T_1} \in \text{ST}(v_1,y)$. Then, by deleting the 1-edge $v_1y$ from $B$ and adding the 2-edge $v_1x_{T_1}$, we get a skeleton $B'$ that has a higher number of 2-edges than $B$ and hence has a lexicographically higher configuration vector. Thus we have a contradiction to the choice of $B$. \qed

**Lemma 26.** Let $x_T$ be a vertex on which Coloring Rule 11 is being applied. Let $v_1$ be as defined in Coloring Rule 11. The vertex $v_1$ has no 2-edge in $H$ to any vertex except $x_T$.

**Proof.** See Figure 15 for an illustration of the proof. Suppose for the sake of contradiction that $v_1$ has a 2-edge in $H$ to a tree vertex $x_{T_1} \in V(H) \setminus \{x_T\}$. The edge $v_1y$ is a 1-edge as otherwise Coloring Rule 10 would have been applicable on $x_T$ and $x_T$ would have been completed already. Thus $x_{T_1} \neq y$. Then the edge $v_1x_{T_1}$ is not in $B_1$ as $y$ and $x_T$ are the only neighbors of $v_1$ in $B_1$. Since $x_{T_1}$ is a tree vertex, it is not in $L_S$ (recall that $L_S$ is the set of non-tree vertices of $H$). Thus, since the edge $v_1x_{T_1}$ is not in $B_1$, it is not in $B$ also (recall $B_1 = B \setminus L_S$). Thus, $v_1x_{T_1} \in E(H) \setminus E(B)$. Also $x_{T_1}$ is not a descendant of $v_1$ due to Lemma 25. Thus, $x_{T_1} \in \text{ST}(v_1,y)$. Then, by deleting the 1-edge $v_1y$ from $B$ and adding the 2-edge $v_1x_{T_1}$, we get a skeleton $B'$ that has a higher number of 2-edges than $B$ and hence has a lexicographically higher configuration vector. Thus we have a contradiction to the choice of $B$. \qed

**Lemma 27.** Let $x_T$ be a vertex on which Coloring Rule 11 is being applied. Let $v_2, w$ be as defined in Coloring Rule 11. Then, $v_2$ has no 2-edge in $H$ to any vertex in $V(H) \setminus \{w, x_T\}$.

**Proof.** See Figure 16 for an illustration of the proof. Suppose for the sake of contradiction that $v_2$ has a 2-edge in $H$ to a tree vertex $x_{T_1} \in V(H) \setminus \{x_T, w\}$. Then the edge $v_2x_{T_1}$ is not in $B_1$ as the only neighbors of $v_2$ in $B_1$ are $x_T$ and $w$. Since $x_{T_1}$ is a tree vertex, it is not in $L_S$ (recall that $L_S$ is the set of non-tree vertices of $H$). Thus, since the edge $v_2x_{T_1}$ is not in $B_1$, it is not in $B$ also (recall $B_1 = B \setminus L_S$). Thus, $v_2x_{T_1} \in E(H) \setminus E(B_1)$. Also $x_{T_1}$ is not a descendant of $v_2$ due to Lemma 25. Clearly, then $x_{T_1} \in \text{ST}(v_2, x_T)$. Since $x_{T_1} \neq x_T$, and the degree of $x_T$ and $v_1$ in $B_1$ is 2, we have that the edge $v_1y$ is on the path from $v_2$ to $x_{T_1}$ in $B_1$. Then, by deleting the 1-edge $v_1y$ from $B$ and adding the 2-edge $v_2x_{T_1}$, we get a skeleton $B'$ that has a higher number of 2-edges and hence has a lexicographically higher configuration vector. Thus we have a contradiction to the choice of $B$. \qed

**Lemma 28.** Let $v$ be a non-tree vertex having degree 2 in $B_1$ such that the neighbors of $v$ in $B_1$ are a tree-vertex $x_T$ and a vertex $x'$, and the tree $T$ consists of a single edge $u_1u_2$. If $v_1$ does not have a 2-edge to any tree-vertex except $x_T$ in $H$, then $v_1$ does not have edges to any tree vertex in $H$ except $x_T$. 

Fig. 15: An illustration of the transformation in the proof of Lemma 26. The densely dotted edges denote the edges of $H$ that are not in $B$. (a) The scenario given by the precondition of the lemma, and (b) the transformation to the new skeleton $B'$ as described in the proof.
Fig. 16: An illustration of the transformation in the proof of Lemma 27. The densely dotted edges denote the edges of $H$ that are not in $B$ and the edge $wv_2$ is drawn with 1 solid line and 1 dotted line to denote that it could be a 1-edge or a 2-edge. (a) The scenario given by the precondition of the lemma, and (b) the transformation to the new skeleton $B'$ as described in the proof.

Fig. 17: An illustration of the transformation in the proof of Lemma 28. The densely dotted edges denote the edges of $H$ that are not in $B$. (a) The scenario given by the precondition of the lemma, and (b) the transformation to the new forest $F'$ as described in the proof where $T$ is removed and a new tree including $v, u_2$ and the vertices of $T_1$ are added.

Proof. See Figure 17 for an illustration. Suppose $v$ has an edge in $H$ to a tree vertex $x_{T_1} \neq x_T$. Note that $vx_{T_1}$ is not a 2-edge by assumption. Thus $vx_{T_1}$ is a 1-edge. We define the forest $F'$ as $F' := (F \setminus \{u_1\}) \cup \{v\}$ and $F' := G[F']$. We prove that $F'$ is a forest with fewer trees than $F$, which is a contradiction to the choice of $F$. (Recall that out of all maximum induced forests, we picked $F$ to be one having the fewest number of trees).

First, we prove that $F'$ is indeed a forest. Suppose for the sake of contradiction that there is a cycle $C$ in $F'$. The cycle $C$ has to contain $v$ because otherwise $C$ is also a cycle in $F$. The cycle $C$ can intersect at most one tree in $T$ as there are no edges across the trees. Let this tree be $T_2$. Then $v$ should have two edges to $T_2$ in order to complete the cycle. We know that $v$ does not have 2-edges to any tree in $T \setminus \{T\}$ by assumption. Hence $T_2 = T$. But, since $|V(T) \cap F'| = 1$, $v$ can only have one edge in $C$ to $T$. Thus, we have a contradiction. Hence, $F'$ is indeed a forest.

Now, we show that the number of trees in $F'$ is smaller than that of $F$. Let $T_1$ be the set of trees in $T \setminus \{T\}$ that have an edge from $v$ in $G$. Clearly, $T_1 \in T_1$ and hence $|T_1| \geq 1$. The vertex set
Lemma 29. Let \( x_T \) be a vertex on which Coloring Rule 11 is being applied. Let \( v_1, v_2, y, w \) be as defined in Coloring Rule 11.

1. \( v_1 \) has no edge in \( H \) to any tree vertex except \( x_T \) (which also implies that \( y \) is a non-tree vertex),
2. if \( wv_2 \) is a 1-edge, vertex \( v_2 \) has no edge in \( H \) to any tree vertex except \( x_T \) (which also implies that \( w \) is a non-tree vertex in this case).

Proof. The first statement follows from Lemmas 26 and 28 and the fact that \( v_1y \) is a 1-edge. The second statement follows from Lemmas 27 and 28.

Lemma 30. Let \( x_T \) be a vertex on which Coloring Rule 11 is being applied and suppose the precondition of Case 1 of the rule is not satisfied. Let \( v_2 \) and \( w \) be as defined in Coloring Rule 11. Then, \( wv_2 \) is a 2-edge.

Proof. Suppose for the sake of contradiction that \( wv_2 \) is a 1-edge. Let \( v_1, u_1, u_2 \) be also as given in Coloring Rule 11 (see Figure 13). Let \( F' = (F \setminus \{u_1\} \cup \{v_1, v_2\}) \). Let \( F' = G[F'] \). We will show that \( F' \) is a forest. Then since \(|F'| > |F|\), we have that \( F \) is not a maximum induced forest, a contradiction.

Suppose for the sake of contradiction that there is a cycle \( C \) in \( F' \).

Case 1: \( C \) intersects more than two trees of \( T \).

Since there are only two vertices in \( F' \) that are not in any tree in \( T \), there have to be two vertices in 2 different trees of \( T \) that are adjacent in \( C \). This is a contradiction as the trees are connected components of \( F \) and therefore have no edges between them in \( G \).

Case 2: \( C \) intersects exactly two trees of \( T \).

Let \( T_1 \) and \( T_2 \) be the trees that \( C \) intersects. Since \( wv_2 \) is a 1-edge, by Lemma 29 we have that both \( v_1 \) and \( v_2 \) are not adjacent in \( H \) to any tree vertex except \( x_T \). Hence, at least one of \( T_1 \) and \( T_2 \) has no edges to \( v_1 \) and \( v_2 \). This means that there should be an edge in \( C \) between \( T_1 \) and \( T_2 \). This is a contradiction as \( T_1 \) and \( T_2 \) are connected components of \( F \) and therefore have no edges between them in \( G \).

Case 3: \( C \) intersects exactly one tree of \( T \).

Let \( T_1 \) be the tree that \( C \) intersects. By Lemma 29 we have that \( v_1 \) and \( v_2 \) do not have an edge to any tree vertex in \( B_1 \) except \( x_T \). Hence, \( T_1 = T \). Since only one vertex of \( T \) is in \( F' \), \( v_1 \) and \( v_2 \) both can have at most one edge to \( T \) in \( C \). This means \( v_1 \) and \( v_2 \) should be adjacent in \( C \) to complete the cycle. Thus, the precondition of Case 1 of Rule 11 is satisfied, a contradiction to our assumption.

Lemma 31. Let \( x_T \) be a vertex on which Coloring Rule 11 is being applied and suppose the precondition of Case 1 of the rule is not satisfied. Let \( T' \) be as defined in Coloring Rule 11 (see Figure 13b). Then, \( x_{T'} \) is not a vertex on which Coloring Rule 10 was applied.

Proof. Suppose for the sake of contradiction that Coloring Rule 10 was applied on \( x_{T'} \). Then \( x_{T'} \) has degree 2 in \( B_1 \) and \( T' \) consists of a single edge. Let this edge be \( u_3u_4 \) (see Figure 18a). Observe that the representatives of the outgoing edge of \( x_{T'} \) are \( u_3v_2 \) and \( u_4v_2 \). Let \( F' := (F \setminus \{u_1, u_3\} \cup \{v_1, v_2\}) \) and \( F' := G[F'] \). Note that \(|F'| = |F|\). We will show that \( F' \) is a forest with fewer trees than \( F \), thereby showing a contradiction to the choice of \( F \). (Recall that out of all maximum induced forests, we picked \( F \) to be one having the fewest number of trees).
Fig. 18: An illustration of the transformation in the proof of Lemma [31]. (a) The initial scenario before transformation, and (b) the transformation to the new forest $F'$ as described in the proof where $T$ and $T'$ are removed and a new tree including $u_4, u_2, v_1, v_2$ is added, thereby reducing the number of trees.
First, we show that $\mathcal{F}'$ is indeed a forest. Suppose for the sake of contradiction that there is a cycle $C$ in $\mathcal{F}'$.

**Case 1:** $C$ intersects more than two trees of $\mathcal{T}$.

Since there are only two vertices in $F'$ that are not in any tree in $\mathcal{T}$, there have to be two vertices in two different trees of $\mathcal{T}$ that are adjacent in $C$. This is a contradiction as the trees are connected components of $\mathcal{F}$ and therefore have no edges between them in $G$.

**Case 2:** $C$ intersects exactly two trees of $\mathcal{T}$.

Let $T_1$ and $T_2$ be the trees that $C$ intersects. Then in order to complete the cycle $C$, there should be at least one edge from each of $v_1$ and $v_2$ to each of the trees $T_1$ and $T_2$. This implies $v_1$ has edges to each of $x_{T_1}$ and $x_{T_2}$ in $H$. But, by Lemma 29 we have that $v_1$ does not have an edge to any tree vertex except $x_T$. Therefore, $x_{T_1} = x_{T_2}$, implying that $T_1 = T_2$, implying that $C$ intersects only one tree, a contradiction.

**Case 3:** $C$ intersects exactly one tree of $\mathcal{T}$.

Let $T_1$ be the tree that $C$ intersects. Since the precondition of Case 1 of Rule 11 is not satisfied, $v_1$ and $v_2$ are not adjacent in $G$. Hence, only one of them is in $C$.

**Case 3.1** $v_1$ is in $C$.

Then, $v_1$ should have two edges in $C$ to $T_1$ in order to complete the cycle. This means that $v_1$ has a 2-edge to $x_{T_1}$ in $H$. By Lemma 29 we have that $v_1$ does not have an edge in $H$ to any tree vertex except $x_T$. Hence, $T_1 = T$. Since only one vertex of $T$ is in $F'$, $v_1$ can have at most one edge to $T$ in $C$. Thus we have a contradiction.

**Case 3.2** $v_2$ is in $C$.

Then, $v_2$ should have two edges in $C$ to $T_1$ in order to complete the cycle. This means that $v_2$ has a 2-edge to $x_{T_1}$ in $H$. By Lemma 29 we have that $v_2$ does not have an edge in $H$ to any tree vertex in $B_1$ except $x_T$ and $x_{T'}$. Hence, $T_1 \in \{T, T'\}$. However, since only one vertex of each $T$ and $T'$ is in $F'$, $v_1$ can have at most one edge to each of $T$ and $T'$ in $C$. Hence, $v_2$ has at most one edge to $T_1$ in $G$. Thus, we have a contradiction.

Thus, we have proved that $\mathcal{F}'$ is indeed a forest. Now, we prove that $\mathcal{F}'$ has fewer trees than $\mathcal{F}$, which concludes the proof. Let $T'_1$ be the set of trees in $\mathcal{T} \setminus \{T, T'\}$ that have an edge from $v_1$ or $v_2$ in $G$. The vertex set \{$v_1, u_2, v_2, u_4$\} $\cup \bigcup_{T'' \in T'_1} V(T'')$ induces a tree in $G$. Hence, the number of trees in $\mathcal{F}'$ is at most $|\mathcal{T} \setminus (T_1 \cup \{T, T'\})| + 1 \leq |\mathcal{T}| - 1$.

Thus we have proved the Lemmas that we used in Coloring Rule 11.

By the end of Coloring Rule 11 we have colored the representatives of all edges in $B_1$. We may have also colored some additional edges of $G$ that are not in $B_1$, namely the shortcut edges (during Coloring Rule 11). We next show that the vertices in $B_1$ are now rainbow connected through these colored edges.

**Lemma 32.** For any pair of vertices $v_1, v_2 \in V(G) \setminus L_S$, $P_{ab}$ is a rainbow path between $v_1$ and $v_2$ in $G$ and uses only colors in $[f]$.

**Proof.** There are no more incomplete tree-vertices because Coloring Rule 11 is applicable on each incomplete tree-vertex and each tree-vertex on which the rule is applied is completed during the rule. This means there are no uncolored 2-edges in $B_1$. Also, Coloring Rule 2 colors all 1-edges in $B_1$. Thus, each edge in $B_1$, and hence their representatives in $G$, have been colored.

Whenever an edge in $B_1$ is colored by a coloring rule and if it is in $Q_{ab}$, we have added exactly one of its representatives to $P_{ab}$ in the proceeding path rule, except possibly Path Rule 11, where we might have added a shortcut edge instead. In the case when a shortcut edge is added, the shortcut edge shortcuts
the two consecutive edges in $Q_{ab}$ whose representatives were not added to $P_{ab}$ and hence the path is not broken.

Also, whenever a tree $T$ has two edges of $P_{ab}$ incident on it, we have added the path between the endpoints of the edges in the tree to $P_{ab}$. And, whenever a tree $T$ with $a \in V(T)$ has one edge of $P_{ab}$ incident on it, we have added the path between the endpoints of the edge and $a$ in the tree to $P_{ab}$. Similarly, whenever a tree $T$ with $b \in V(T)$ has one edge of $P_{ab}$ incident on it, we have added the path between the endpoints of the edge and $b$ in the tree to $P_{ab}$. If there is a tree $T$ with $a, b \in V(T)$ we added the path between the endpoints of $a$ and $b$ in the tree to $P_{ab}$ during Path Rule 1. It follows that $P_{ab}$ is indeed a path between $a$ and $b$ in $G$. Since Invariant 1 holds, we know that $P_{ab}$ is a rainbow path. Since we have used only the colors from 1 to $f$ so far, the lemma follows.

So, now we only need to worry about how to rainbow connect vertices in $L_S$ between themselves and to the other vertices. For this, we give the following coloring rule.

**Coloring Rule 12** For each $v \in L_S$, let $e$ be the unique 2-edge incident on $v$ which exists by Lemma 4. Color $(e)_1$ with $g_1 = f + 1$ and $(e)_2$ with $g_2 = f + 2$. (Recall that $g_1$ and $g_2$ are the global surplus colors).

Now, we complete the proof of the main theorem.

**Proof (Proof of Theorem 1).** Consider any pair of vertices $a_1, a_2 \in V(G)$. If $a_1 \in L_S$, let $e_1$ be the edge incident on $a_1$ that is colored with $g_1$, and let $a$ be the other end of $e_1$. If $a_1 \notin L_S$, let $a = a_1$. If $a_2 \in L_S$, let $e_2$ be the edge incident on $a_2$ that is colored with $g_2$, and let $b$ be the other end of $e_2$. If $a_2 \notin L_S$, let $b = a_2$. We know there is a rainbow path $P_{ab}$ from $a$ to $b$ that uses only colors in $[f]$ due to Lemma 32. We define path $P$ as follows. If $a_1, a_2 \in L_S$, then $P := a_1 aP_{ab} a_2$. If $a_1 \in L_S$ but $a_2 \notin L_S$, then $P := a_1 aP_{ab}$. If $a_2 \in L_S$ but $a_1 \notin L_S$, then $P := P_{ab} a_2$. If $a_1, a_2 \notin L_S$, then $P := P_{ab}$. It is clear from the construction that $P$ is a path between $a_1$ and $a_2$. Since edge $a_1a$ is colored with $g_1 = f + 1$, edge $ba_2$ is colored with $g_2 = f + 2$, and path $P_{ab}$ uses only colors in $[f]$, the path $P$ is indeed a rainbow path.

\[\Box\]

3 Conclusions

We gave an upper bound of $f(G) + 2$ on $rc(G)$, strengthening the intuition that tree-like and dominating structures are helpful in rainbow-connecting graphs. Our bound is tight up to an additive factor of 3 as shown by any tree. The question remains whether the bound can be improved to $f(G) - 1$ so that the bound is tight even with respect to additive factors. Also, then the bound would be a strict improvement over the bound $n - 1$ obtained by coloring the edges of a spanning tree in distinct colors. From our insight developed during the current work, we conjecture such a slightly stronger bound.

**Conjecture 1.** A connected graph $G$ has $rc(G) \leq f(G) - 1$.

We expect that proving this conjecture requires further extensive case analysis. Further, we note that Lauri [12] proposed the following stronger version of the above conjecture, discovered using the automated conjecture-making software GraPHedron [16].

**Conjecture 2 ([12]).** A connected graph $G$ has $src(G) \leq f(G) - 1$.

**Acknowledgement.** We thank Kurt Mehlhorn for funding the research visit of Erik Jan van Leeuwen and Juho Lauri to Max Planck Institute for Informatics, Saarbrücken, which led to this work.
References

1. M. Basavaraju, L. S. Chandran, D. Rajendraprasad, and A. Ramaswamy. Rainbow connection number and radius. *Graphs and Combinatorics*, 30(2):275–285, 2014.

2. Y. Caro, A. Lev, Y. Roditty, Z. Tuza, and R. Yuster. On rainbow connection. *Electron. J. Combin*, 15(1):R57, 2008.

3. S. Chakraborty, E. Fischer, A. Matsliah, and R. Yuster. Hardness and algorithms for rainbow connection. *Journal of Combinatorial Optimization*, 21(3):330–347, 2009.

4. L. S. Chandran, A. Das, D. Rajendraprasad, and N. M. Varma. Rainbow connection number and connected dominating sets. *Journal of Graph Theory*, 71(2):206–218, 2012.

5. G. Chartrand, G. L. Johns, K. A. McKeon, and P. Zhang. Rainbow connection in graphs. *Mathematica Bohemica*, 133(1), 2008.

6. G. Chartrand and P. Zhang. *Chromatic graph theory*. CRC press, 2008.

7. P. Dorbec, I. Schiermeyer, E. Sidorowicz, and É. Sopena. Rainbow connection in oriented graphs. *Discrete Applied Mathematics*, 179:69–78, 2014.

8. J. Ekstein, P. Holub, T. Kaiser, M. Koch, S. Matos Camacho, Z. Ryjáček, and I. Schiermeyer. The rainbow connection number of 2-connected graphs. *Discrete Mathematics*, 313(19):1884–1892, 2013.

9. S. N. Joseph and L. C. DiPippo. *Pseudo-scheduling: A New Approach to the Broadcast Scheduling Problem*, pages 93–104. Springer Berlin Heidelberg, 2013.

10. N. Kamčev, M. Krivelevich, and B. Sudakov. Some remarks on rainbow connectivity. *Journal of Graph Theory*, 83(4):372–383, 2016.

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

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

13. X. Li, Y. Shi, and Y. Sun. Rainbow Connections of Graphs: A Survey. *Graphs and Combinatorics*, 29(1):1–38, 2012.

14. X. Li and Y. Sun. *Rainbow connections of graphs*. Springer, 2012.

15. X. Li and Y. Sun. An updated survey on rainbow connections of graphs-a dynamic survey. *Theory and Applications of Graphs*, (1):3, 2017.

16. H. Mélot. Facet defining inequalities among graph invariants: The system GraPHedron. *Discrete Applied Mathematics*, 156(10):1875–1891, 2008.