



# Upper bounding rainbow connection number by forest number



L. Sunil Chandran<sup>a,1</sup>, Davis Issac<sup>b,\*</sup>, Juho Lauri<sup>c,3</sup>, Erik Jan van Leeuwen<sup>d</sup>

<sup>a</sup> Indian Institute of Science, Bangalore, India

<sup>b</sup> Hasso Plattner Institute, Potsdam, Germany

<sup>c</sup> Bell Labs, Dublin, Ireland

<sup>d</sup> Utrecht University, the Netherlands

## ARTICLE INFO

### Article history:

Received 30 September 2020

Received in revised form 6 January 2022

Accepted 28 January 2022

Available online 1 March 2022

### Keywords:

Rainbow connection

Forest number

Upper bound

## 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.

© 2022 Elsevier B.V. All rights reserved.

## 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**

\* Corresponding author.

E-mail address: davisissac22@gmail.com (D. Issac).

<sup>1</sup> 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 Foundation fellowship.

<sup>2</sup> The major part of the work was done when this author was a PhD student at Max Planck Institute for Informatics, Saarbrücken, Germany.

<sup>3</sup> The major part of the work was done when this author was at Bell Labs, Dublin, Ireland.

**connection number** of  $G$ , denoted as  $\text{src}(G)$ . These measures of rainbow connectivity were introduced by Chartrand et al. [5] in 2008. The concept has gathered significant attention from both combinatorial and algorithmic perspectives. Indeed, the work of Chartrand et al. [5] 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 [3], layered encryption [7], and broadcast scheduling [9].

While introducing the parameters, Chartrand et al. [5] 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  $\text{diam}(G)$ , is the length of a longest shortest path in  $G$ . Now, it is straightforward to verify that  $\text{diam}(G) \leq \text{rc}(G) \leq \text{src}(G) \leq m$ , where  $m$  is the number of edges of  $G$ . In other words, both  $\text{rc}(G)$  and  $\text{src}(G)$  are always sandwiched between one and  $m$ . The extremal cases are not difficult to 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  $\text{rc}(G) \leq \frac{20n}{\delta}$ , 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

Here, we give a summary of the ideas used for proving Theorem 1. We first fix a maximum induced forest  $\mathcal{F}$  of  $G$  and define  $H$  to be the graph obtained from  $G$  by contracting each connected component of  $\mathcal{F}$ , each of which is a tree, 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 *inner skeleton*  $B_1$  is defined to be  $B$  minus the leaves of  $B$  that are non-tree vertices.

We color all the edges of the forest  $\mathcal{F}$  with distinct colors (call them forest colors), then associate with each tree of  $\mathcal{F}$ , an additional color called its *surplus color*, and also keep aside two *global surplus colors*. Note that this makes the total

number of colors  $f(G) + 2$  as required. The idea is to color the edges of  $G$  that appear in the inner skeleton  $B_1$  using the surplus colors in such a way that between any pair of vertices in  $B_1$  there is a rainbow path using edges of  $B_1$  and  $\mathcal{F}$ . However, this turned out to be not always possible; in some cases we had to use some special edges outside  $B_1$  that we call shortcut edges. After rainbow-connecting vertex-pairs in inner skeleton  $B_1$ , we connect the vertices outside of the inner skeleton to the inner skeleton using the two global surplus colors. The hard part of the proof is in making the coloring of  $B_1$  work with one surplus color per tree. For this, 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  $G_1$  of  $G$ , and any  $E' \subseteq E(G)$ , we use  $E'(G_1)$  to denote  $E' \cap E(G_1)$ . 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  $u$  and  $v$  and for a directed edge from  $u$  to  $v$ , we use  $\vec{uv}$ . 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  $\mathcal{F}$ , each connected component is a tree, we will use the phrases “tree of  $\mathcal{F}$ ” and “connected component of  $\mathcal{F}$ ” 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  $v$ . The following is a general easy-to-see observation about trees that we use in the proof.

**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  $e$ .*

## 2. Proof of Theorem 1

Let  $G = (V, E)$  be a connected graph. Our goal is to prove that  $\text{rc}(G) \leq f(G) + 2$ .

Let  $\mathcal{F}$  be a maximum induced forest of  $G$  that has the smallest number of connected components (trees) out of all the maximum induced forests of  $G$ . Let  $F = V(\mathcal{F})$  be the set of vertices in  $\mathcal{F}$ . Let  $\mathcal{T}$  be the set of connected components (trees) of  $\mathcal{F}$  and let  $t = |\mathcal{T}|$ . 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  $\mathcal{T}$ ; otherwise, the edge is called a **non-tree edge**.

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

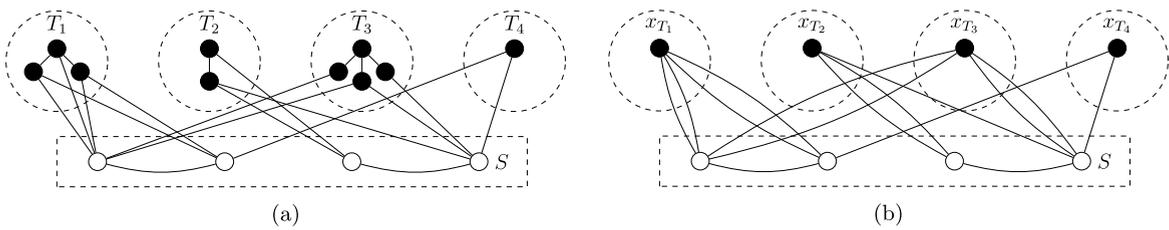
$$\begin{aligned} V(H) &:= V_{\mathcal{T}} \cup S, \text{ where} \\ V_{\mathcal{T}} &:= \{x_T : T \in \mathcal{T}\} \text{ and} \\ E(H) &:= E(G[S]) \cup \{ux_T : u \in S, T \in \mathcal{T}, u \text{ has at least one edge to } V(T) \text{ in } G\}. \end{aligned}$$

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

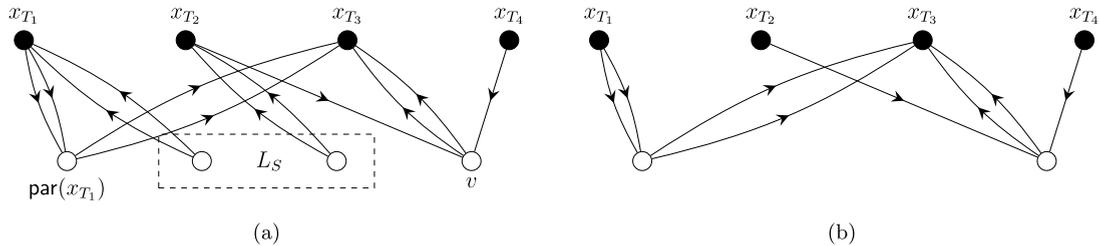
$$\begin{aligned} E_1 &:= E(G[S]) \cup \{ux_T : u \in S, T \in \mathcal{T}, u \text{ has exactly one edge to } V(T) \text{ in } G\} \text{ and} \\ E_2 &:= \{ux_T : u \in S, T \in \mathcal{T}, u \text{ has at least two edges to } V(T) \text{ in } G\}. \end{aligned}$$

The edges in  $E_1$  are called **1-edges** while those in  $E_2$  are **2-edges**. See Fig. 1 for an illustration of the above definitions. We define a function  $f_{\mathcal{T}} : V_{\mathcal{T}} \rightarrow \mathcal{T}$  that maps a tree-vertex to its corresponding tree, i.e.,  $f_{\mathcal{T}}(x_T) = T$ . For each edge in  $H$ , we define its **representatives** in  $G$  as follows. Consider first a 2-edge  $e$  between  $u \in S$  and  $x_T \in V_{\mathcal{T}}$ . 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  $e$  and denote them by  $(e)_1$  and  $(e)_2$ . For a 1-edge  $e$  between  $u \in S$  and  $x_T \in V_{\mathcal{T}}$ , 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 **foots** of  $ux_T$ . The unique path between  $v_1$  and  $v_2$  in  $T$  is called the **foot-path** of  $ux_T$ . For a 1-edge  $ux_T$  (whose one endpoint is a tree-vertex) with representative  $uv$ , we call the vertex  $v$ , the foot of  $ux_T$ .

A **skeleton** is an in-arborescence obtained by taking a spanning tree of  $H$  with an arbitrary node of  $V_{\mathcal{T}}$  fixed as its root with all edges directed towards the root. 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:



**Fig. 1.** (a) A graph  $G$  is partitioned into a maximum induced forest  $\mathcal{F}$  and  $S = V(G) \setminus V(\mathcal{F})$ . The connected components (trees) of  $\mathcal{F}$  are  $T_1, T_2, T_3$  and  $T_4$ . The edges between two black vertices, corresponding to vertices in  $V(\mathcal{F})$ , are tree-edges. (b) The graph  $H$  obtained after contracting the connected components of  $\mathcal{F}$ . We draw a 2-edge with 2 lines and a 1-edge with a single line.



**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  $B_1 := B[V(B) \setminus L_S]$ .

$$\langle |E_2(B)|, n_2, n_3, \dots, n_{|V|} \rangle,$$

where  $n_i$  denotes the number of vertices in level  $i$  in  $B$ .

We now fix a skeleton  $B$  such that it has the lexicographically highest configuration vector out of all possible skeletons. 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  $\vec{uv}$  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  $\tilde{B}_1$  be the underlying undirected tree of  $B_1$ . These concepts are illustrated in Fig. 2.

We now prove a lemma and three corollaries that are useful for our coloring procedure.

**Lemma 1.** 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  $B$ . There exists a  $T \in \mathcal{T}$  such that  $v$  has at least two edges to  $T$  in  $G$ , because otherwise,  $G[F \cup \{v\}]$  is a forest, contradicting the maximality of  $\mathcal{F}$ . 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  $C$ . Note that  $e$  is a 1-edge by assumption. 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  $B$ .  $\square$

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.  $\square$

**Corollary 4.** For each 1-edge  $\vec{uv}$  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 1. 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  $B_1$ .  $\square$

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  $\{h(e) : e \in E' \text{ and } e \text{ is a non-tree edge}\}$ . 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  $\{1, 2, \dots, f + 2\}$ . We call colors  $f + 1$  and  $f + 2$  the **global surplus colors**, and denote them by  $g_1$  and  $g_2$  respectively. 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  $\{1, 2, \dots, f\}$  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  $u_2$  are in some  $T \in \mathcal{T}$ , but  $u_1$  and  $v_2$  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  $v, a \in V(T)$  ( $v, b \in V(T)$  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  $\mathcal{F}$  with distinct colors  $1, 2, \dots, f - t$ .

**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  $i^{\text{th}}$  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  $\vec{uv}$  in  $B$ : if  $u$  is a tree-vertex, then color  $(uv)_1$  with  $s(f_{\mathcal{T}}(u))$ ; otherwise, i.e., if  $u$  is not a tree-vertex, by Corollary 4, 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  $T_{wa}$  to  $P_{ab}$ , where  $w$  is the foot of the edge  $ux_T$  in  $T$ .

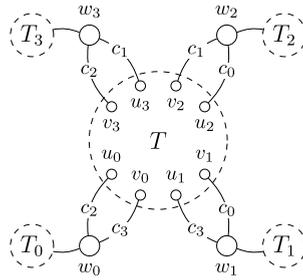


Fig. 3. Illustration of Coloring Rule 3 applied on a tree-vertex  $x_T$  with 2-edge degree 4. Here,  $c_i = s(T_i)$ .

- If for some tree  $T$  it holds that  $b \in V(T)$  and there is a 1-edge  $ux_T$  in  $Q_{ab}$ , then add the path  $T_{wb}$  to  $P_{ab}$ , where  $w$  is the foot of the edge  $ux_T$  in  $T$ .
- If for some tree  $T$  there are two 1-edges  $ux_T$  and  $vx_T$  in  $Q_{ab}$ , add the path  $T_{wz}$  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 2.** 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  $B_1$ . We fix a closest (breaking ties arbitrarily) tree-vertex to  $v$  in  $ST(u, v)$  in  $\tilde{B}_1$  and denote it 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 Fig. 3 for an illustration). Let  $q$  be the 2-edge degree of  $x_T$ . Let  $w_0, w_1, w_2, \dots, w_{q-1}$  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  $c_{((i+2) \bmod q)}$  and the edge  $(x_T w_i)_2$  with  $c_{((i+3) \bmod q)}$ .

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

**Lemma 3.** 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  $(\{c_0, c_1, \dots, c_{q-1}\} \setminus \{c_i, c_j\}) \cup c(T)$ . 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  $(\{c_0, c_1, \dots, c_{q-1}\} \setminus \{c_i\}) \cup c(T)$ .

**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  $i \in \{0, 1, \dots, q - 1\}$ . 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  $w_j$ :  $P := w_i u_i T_{u_i u_j} u_j w_j$ ,  $P' := w_i v_i T_{v_i v_j} v_j w_j$ , and  $P'' := w_i v_i T_{v_i u_j} u_j w_j$ . By our assumption, each of these paths, is either not a rainbow path, or uses a color that is not in  $(\{c_0, c_1, \dots, c_{q-1}\} \setminus \{c_i, c_j\}) \cup c(T)$ . Also, from Coloring Rules 1 and 3, we know that the only colors that are not in  $(\{c_0, c_1, \dots, c_{q-1}\} \setminus \{c_i, c_j\}) \cup c(T)$  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  $T_{u_i u_j}, T_{v_i v_j}$ , and  $T_{v_i u_j}$  are all rainbow paths due to Coloring Rule 1, and moreover the colors used by them are disjoint from  $\{c_0, \dots, c_{q-1}\}$ . For the path  $P$ , this means that either  $c(w_i u_i) = c(u_j w_j)$  or  $\{c(w_i u_i), c(u_j w_j)\} \cap \{c_i, c_j\} \neq \emptyset$ . That is, either  $c_{(i+2) \bmod q} = c_{(j+2) \bmod q}$  or  $\{c_{(i+2) \bmod q}, c_{(j+2) \bmod q}\} \cap \{c_i, c_j\} \neq \emptyset$ . That is, either  $i = j$  or  $\{(i + 2) \bmod q, (j + 2) \bmod q\} \cap \{i, j\} \neq \emptyset$ . But we know that  $(i + 2) \bmod q \neq i$  and that  $(j + 2) \bmod q \neq j$ . Therefore, either  $(i + 2) \bmod q = j$  or  $(j + 2) \bmod q = i$ . Without loss of generality assume that  $(i + 2) \bmod q = j$ .

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

Now consider the third path  $P''$ . We have that either  $c(w_i v_i) = c(u_j w_j)$  or  $\{c(w_i v_i), c(u_j w_j)\} \cap \{c_i, c_j\} \neq \emptyset$ . That is, either  $c_{(i+3) \bmod q} = c_{(j+2) \bmod q}$  or  $\{c_{(i+3) \bmod q}, c_{(j+2) \bmod q}\} \cap \{c_i, c_j\} \neq \emptyset$ . That is, either  $(i + 3) \bmod q = (j + 2) \bmod q$  or  $\{(i + 3) \bmod q, (j + 2) \bmod q\} \cap \{i, j\} \neq \emptyset$ . Substituting that  $(i + 3) \bmod q = ((i + 2) \bmod q + 1) \bmod q = (j + 1) \bmod q$  and that  $i = (j + 3) \bmod q$ , we get that either  $(j + 1) \bmod q = (j + 2) \bmod q$  or  $\{(j + 1) \bmod q, (j + 2) \bmod q\} \cap$

$\{(j + 3) \bmod q, j\} \neq \emptyset$ . Since  $j, (j + 1) \bmod q, (j + 2) \bmod q$ , and  $(j + 3) \bmod q$  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  $P''' := w_i u_i T_{u_i u}$ . We know that the path  $T_{u_i u}$  uses only colors from  $c(T)$  and is rainbow, and that the edge  $w_i u_i$  is colored  $c_{(i+2) \bmod q}$ . Also,  $c_{(i+2) \bmod q} \neq c_i$  as  $(i + 2) \bmod q \neq i$ . Thus  $P$  is a rainbow path and uses only the colors in  $(\{c_0, c_1, \dots, c_{q-1}\} \setminus \{c_i\}) \cup c(T)$ .  $\square$

**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  $Q_{ab}$ .

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 3.

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

Let  $x_T w_i$  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 3. Add this path to  $Q_{ab}$ . (Note that the representative of the 1-edge has been already added to  $P_{ab}$  during Path Rule 2).

**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 3. Add this path to  $P_{ab}$ .

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

**Lemma 4.** 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  $ST(x_T, x_{T_i})$ . Then the color  $c_i$  was not used in Path Rule 3 according to Lemma 3. That means  $e$  was not colored with  $c_i$ , which is a contradiction.  $\square$

**Lemma 5.** Invariant 1 is not violated during Path Rule 3.

**Proof.** Suppose Invariant 1 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  $ST(x_T, x_{T_j})$  by Lemma 4. 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 1. 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  $\text{dist}_{\bar{B}_1}(x_{T'}, x_T) < \text{dist}_{\bar{B}_1}(x_{T'}, x_{T_j})$ . But then during the application of Coloring Rule 3 on  $x_{T'}$ , the color  $s(T_j)$  would never be used as  $x_{T_j} \neq cT(x_{T'}, v)$  for any vertex  $v$ . Thus, the color of  $e'$  is not  $s(T_j)$ . 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  $s(T_j)$ , we have that  $h(e')$  should either be the outgoing edge of  $x_{T_j}$  or the outgoing edge of the parent of  $x_{T_j}$ , 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.  $\square$

**Coloring Rule 4.** For each tree-vertex  $x_T$  with 2-edge degree exactly 3 (see Fig. 4), 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  $x_{T_i} = cT(x_T, w_i)$ , let  $u_i$  and  $v_i$  be the foots of  $x_T w_i$  in  $T$ , let  $P_i := T_{u_i v_i}$ , and let  $c_i := s(T_i)$ .

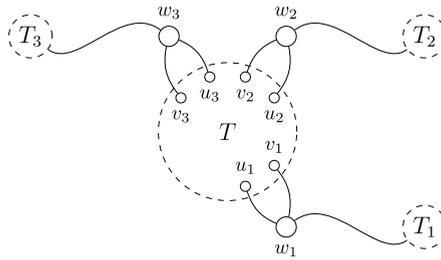


Fig. 4. A scenario in which Coloring Rule 4 is applicable on  $x_T$ .

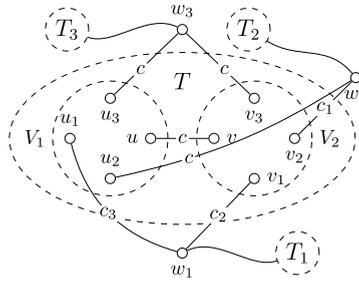


Fig. 5. Case 1 of Coloring Rule 4.

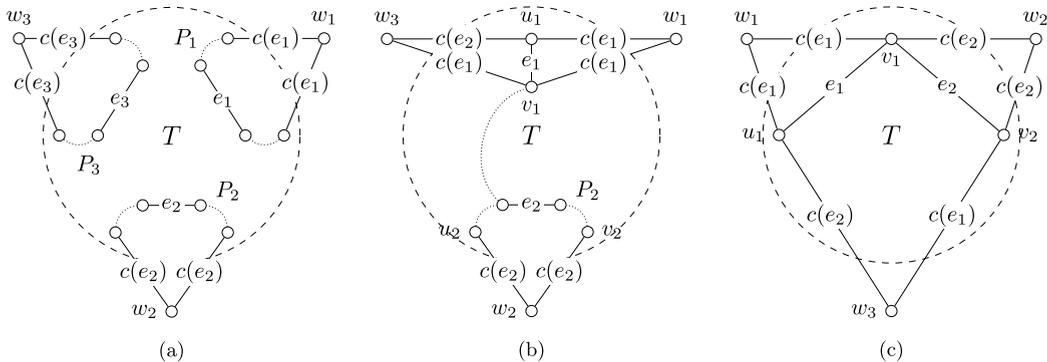


Fig. 6. Cases 2 and 3 of Coloring Rule 4. (a) Case 2. Note that  $P_1$ ,  $P_2$ , and  $P_3$  are not necessarily disjoint. (b) Case 3, scenario 1. Note that  $u_1 = u_3$  and  $v_1 = v_3$ . (c) Case 3, scenario 2. Note that  $u_1 = u_3$ ,  $v_1 = u_2$  and  $v_3 = v_2$ .

**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 Fig. 5).

Without loss of generality, let  $u_i$  and  $v_i$  be the roots of  $x_T w_i$  in  $V_1$  and  $V_2$  respectively for each  $i \in \{1, 2, 3\}$ . Let  $c$  be the color of  $uv$ . Color  $u_1 w_1$  with  $c_3$ ,  $v_1 w_1$  with  $c_2$ ,  $u_2 w_2$  with  $c$ ,  $v_2 w_2$  with  $c_1$ ,  $u_3 w_3$  with  $c$ , and  $v_3 w_3$  with  $c$ , as shown in Fig. 5.

**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 Fig. 6 (a)).

Color both the representatives of  $x_T w_i$  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<sup>4</sup> 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  $E(P_3) \cap E(P_1) \neq \emptyset$ . 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  $e_1 = u_1 v_1$ . 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  $u_1 = u_3$ . Let  $e_2$  be any edge in  $P_2$ . Without loss of generality assume that

<sup>4</sup> We use the following Helly property of trees: if  $T_1, T_2, \dots, T_k$  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  $T_1, T_2, \dots, T_k$ .

$v_1$  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 Fig. 6 (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. The lemma is easy to verify (with the help of Figs. 5 and 6, and using Observation 1) and hence we state it without proof.

**Lemma 6.** For each tree-vertex  $x_T$  on which Coloring 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  $(\{c_1, c_2, c_3\} \setminus \{c_i\}) \cup c(T)$ .

**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 6.

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

Let  $x_T w_i$  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 6.

**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 6. Add this path to  $P_{ab}$ .

The following lemma follows from Lemma 6 and Path Rule 4. The proof is similar to that of Lemma 4 and is omitted.

**Lemma 7.** 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 8.** 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  $w_1x_T$ . 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 4 and 7, 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 4 and 7. 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  $B_1$  (see Fig. 7), let  $q$  be the number of children of  $u$  (note that  $q \geq 2$  as degree of  $u$  is at least 3), let  $u_1, u_2, \dots, u_q$  be the children of  $u$  and let  $x_{T_i}$  be  $CT(u, u_i)$ . Let  $\vec{uv}$  be the outgoing edge from  $u$  in  $B_1$ . If  $uv$  is a 1-edge, due to Coloring Rule 2, we know that there exists an  $i \in [q]$  such that  $u_i$  is a tree-vertex (and hence  $T_i = f_{\mathcal{T}}(u_i)$ ), and  $uv$  is colored with  $s(T_i)$ . Hence, if  $uv$  is a 1-edge, assume without loss of generality that  $u_1$  is a tree-vertex (and hence  $x_{T_1} = u_1$ ) and that  $uv$  is colored with  $s(T_1)$ .

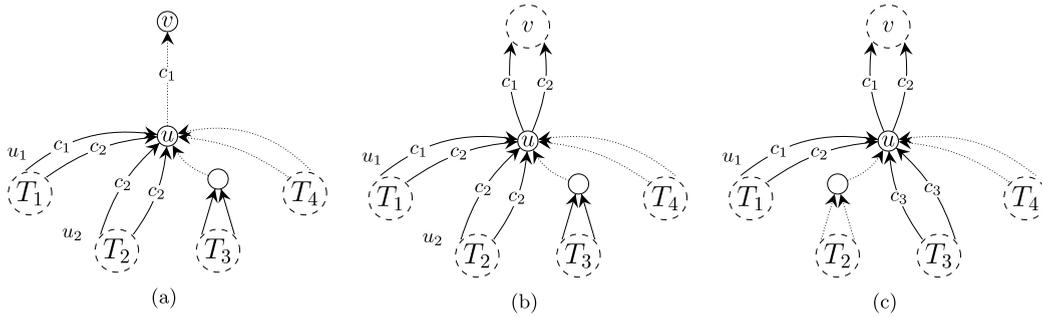


Fig. 7. Three examples of Coloring Rule 5. Here  $c_i = s(T_i)$ . The edges that were colored before the application of the rule are drawn as densely dotted lines.

- If  $u_1u$  is uncolored (then  $u_1u$  is a 2-edge due to Coloring Rule 2, implying that  $u_1$  is a tree-vertex and hence  $T_1 = f_{\mathcal{T}}(u_1)$ ), then color  $(u_1u)_1$  with  $s(T_1)$  and  $(u_1u)_2$  with  $s(T_2)$ .
- For each  $2 \leq i \leq q$ , if  $u_iu$  is uncolored (then  $u_iu$  is a 2-edge due to Coloring Rule 2, implying that  $u_i$  is a tree-vertex and hence  $T_i = f_{\mathcal{T}}(u_i)$ ), then color both its representatives with  $s(T_i)$ .
- If  $uv$  is uncolored (in which case it is a 2-edge due to Coloring Rule 2) then color  $(uv)_1$  with  $s(T_1)$  and  $(uv)_2$  with  $s(T_2)$ .

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

**Part 1**

- If  $Q_{ab}$  contains edge  $u_1u$  and  $u_1u$  is colored during the application of Coloring Rule 5 on  $u$ , do the following. If the other neighbor (if any) of  $u$  in  $Q_{ab}$  is  $u_2$ , then add  $(u_1u)_1$  (which has color  $s(T_1)$ ) to  $P_{ab}$ . Otherwise, add  $(u_1u)_2$  (which has color  $s(T_2)$ ) to  $P_{ab}$ .
- For each  $i \in [2, q]$ , if  $Q_{ab}$  contains edge  $u_iu$  and  $u_iu$  is colored during the application of Coloring Rule 5 on  $u$ , add  $(u_iu)_1$  (which has color  $s(T_i)$ ) to  $P_{ab}$ .
- If  $Q_{ab}$  contains edge  $uv$  and  $uv$  is colored during the application of Coloring Rule 5 on  $u$ : if the other neighbor (if any) of  $u$  in  $Q_{ab}$  is  $u_1$  and  $u_1u$  is a 1-edge, then add  $(uv)_2$  (which has color  $s(T_2)$ ) to  $P_{ab}$ ; otherwise add  $(uv)_1$  (which has color  $s(T_1)$ ) to  $P_{ab}$ .

**Part 2**

- For each tree-vertex  $x_T$  such that the degree of  $x_T$  in  $h(P_{ab})$  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  $P_{ab}$  incident on  $T$ . Add  $T_{xy}$  to  $P_{ab}$ .
- For each tree-vertex  $x_T \in \{h(a), h(b)\}$  such that the degree of  $x_T$  in  $h(P_{ab})$  became 1 during the addition of above edges in Part 1, let  $x$  be the endpoint in  $T$  of the edge of  $P_{ab}$  incident on  $T$ . If  $x_T = h(a)$ , add  $T_{ax}$  to  $P_{ab}$ ; otherwise (i.e., if  $x_T = h(b)$ ), add  $T_{bx}$  to  $P_{ab}$ .

**Lemma 9.** Invariant 1 is not violated during Path Rule 5.

**Proof.** Suppose Invariant 1 is violated during the application of Path Rule 5 on the pair  $(u, P_{ab})$  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 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  $x_T$  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  $c(e') \neq c(e)$ , a contradiction. Thus, the edge  $e$  was not added during Part 2. Then  $e$  was added during Part 1 and hence  $c(e) = c(e') = s(T_i)$  for some  $i \in [q]$ . 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  $(u, P_{ab})$ , we have added at most two edges to  $P_{ab}$ . And, if we have added two edges, they are of different colors. Thus  $e'$  was not added to  $P_{ab}$  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  $u' \neq u$ . 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  $u'$  are ancestors of  $x_T$ . Without loss of generality, assume that  $u'$  is closer than  $u$  to  $x_T$ . Then,  $u$  cannot have any tree vertices as children because otherwise  $x_{T_i} \neq CT(u, u_i)$ . 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**  $e = (uv)_2$ .

We know that  $e = (uv)_2$  is colored with  $s(T_2)$  by Coloring Rule 5. Thus,  $c(e') = c(e) = s(T_2)$  and  $T_i = T_2$ . Since the edge  $(uv)_2$  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  $u' \in Q_{ab}$ , we have that  $u'$  is a descendant of  $u_1$  and not  $u_2$  in  $B_1$ . This implies  $T_i = T_1 \neq T_2$ , a contradiction.

**Case 1.2**  $e = (uv)_1$ .

Since the edge  $(uv)_1$  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  $u' \in Q_{ab}$ , this implies that  $T_i \neq T_1$ , and hence  $c(e) = c(e') = s(T_i) \neq s(T_1)$ . But we know that  $e = (uv)_1$  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  $s(T_i)$  during the application of Coloring Rules 4 or 3 on  $x_{T'}$ . Then  $Q_{ab}$  does not intersect  $ST(T', T_i)$  due to Lemmas 7 and 4. Since  $x_{T_i} = CT(u, u_i)$ , there is no other tree-vertex in the path from  $u$  to  $x_{T_i}$ . Thus,  $u$  is in  $ST(T', T_i)$ . 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  $(uv)_1$  and  $(uu_1)_1$ . Since  $e$  and  $e'$  are distinct we have  $e = (u_1u)_1$ . But since  $uv$  is in  $Q_{ab}$ , we would have added  $(u_1u)_2$  and not  $(u_1u)_1$  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  $x_{T_i}$  and  $u$ . Also,  $x_{T_i}$  is not a child of  $u$ . Then, the only possibility for  $e$  to have color  $s(T_i)$  is if  $i = 2$  and  $e = (u_1u)_2$ . Then  $Q_{ab}$  contains both  $u_1$  and  $u_2$ . In that case, we would have added  $(u_1u)_1$  and not  $(u_1u)_2$  to  $P_{ab}$  during Path Rule 5. Hence,  $e \neq (u_1u)_2$ , 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 1, 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 10.** Invariant 1 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  $c((e)_j) = c(e_1)$  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.  $\square$

**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.

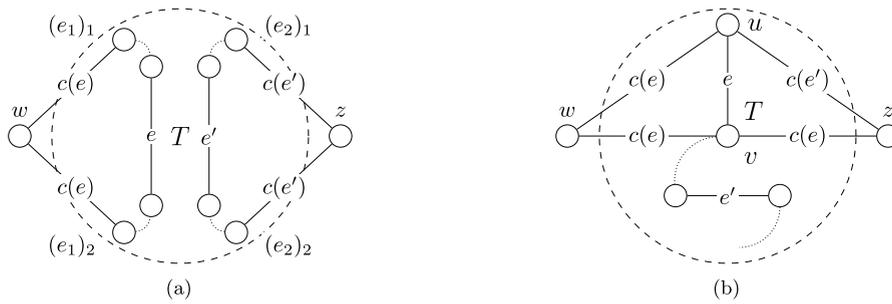


Fig. 8. Coloring Rule 7. (a) Case 1 (b) Case 2.

**Case 1:**  $|E(P_1 \cup P_2)| \geq 2$ . (See Fig. 8 (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 Fig. 8 (b)).

Clearly,  $P_1$  and  $P_2$  both are a single edge and they are the same edge. Let this edge be  $e = uv$ . 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 11.** 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  $T_{xu}$  or  $T_{xv}$  avoids  $e$ . Then at least one of the paths  $T_{xu}$  followed by  $uw$  or  $uz$ , or  $T_{xv}$  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  $T_{w_1z_1}, T_{w_1z_2}, T_{w_2z_1}$  and  $T_{w_2z_2}$  contains neither  $e$  nor  $e'$ . Given this, it is easy to show the necessary rainbow path from  $w$  to  $z$ : if the path  $T_{w_iz_j}$  contains neither  $e$  nor  $e'$  then the path  $ww_iT_{w_iz_j}z_jz$  is the required path. So for the sake of contradiction assume that each  $T_{w_iz_j}$  contains either  $e$  or  $e'$ . Without loss of generality assume that  $T_{w_1z_1}$  contains  $e$ . Let  $y$  be the last vertex on  $T_{w_1z_1}$  that is in  $P_1$  (while going from  $w_1$  to  $z_1$ ). Now,  $T_{yz_1}$  and  $T_{yw_2}$  do not contain  $e$  and hence  $T_{z_1w_2} = T_{yz_1} \cup T_{yw_2}$  does not contain  $e$ . This implies that  $T_{z_1w_2}$  contains  $e'$ . Let  $y'$  be the last vertex on  $T_{z_1w_2}$  that is in  $P_2$  (while going from  $z_1$  to  $w_2$ ). Now,  $T_{y'z_2}$  and  $T_{w_2y'}$  contains neither  $e$  nor  $e'$  and hence  $T_{w_2z_2} = T_{y'z_2} \cup T_{w_2y'}$  contains neither  $e$  nor  $e'$ .

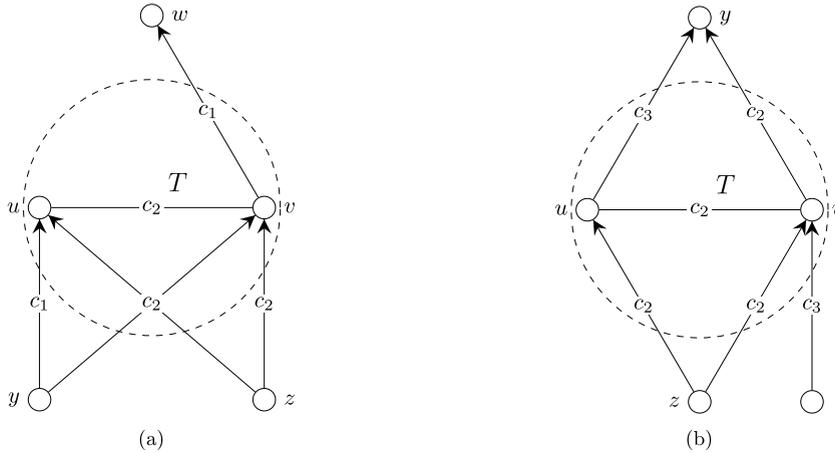
To prove the second statement, observe that there is a rainbow path from  $x$  to either  $w_1$  or  $w_2$  not containing  $e$ , and a rainbow path from  $x$  to either  $z_1$  or  $z_2$  not containing  $e'$ , due to Observation 1.  $\square$

**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 11.

**Lemma 12.** 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 11, 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 4, we have that  $E(T) \cap E_O = \emptyset$ . Thus  $d \notin E(T)$  but has color in  $c(E(T))$  and  $d \in E_O$ . 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.  $\square$



**Fig. 9.** Coloring Rule 8. **(a)** Case 1 where  $c_1 = s(T)$  and  $c_2 = c(uv)$ . **(b)** Case 2 where  $c_2 = c(uv)$  and  $c_3$  is the color of the representative of an arbitrarily chosen 1-edge incident on  $x_T$ .

**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  $x_T$  has at least three 2-edges incident on it, then Coloring Rule 3 or 4 would have been applicable on  $x_T$ . If it has 2-edge degree 1, then Coloring Rule 6 would have been applicable on  $x_T$ . If it has 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 7 becomes applicable on  $x_T$ . Hence, we can assume that the tree  $T$  is 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 Fig. 9 (a)).

Then the outgoing edge of  $x_T$  in  $B_1$  (if any) is a 1-edge, say  $\overrightarrow{x_T w}$ . 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 Fig. 9 (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 8 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  $y_1 y_2$  exists. Add the edge  $y_1 y_2$  to  $P_{ab}$ .

**Lemma 13.** Invariant 1 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 Fig. 9 (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 7 and 4. 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  $x_T w$  is in  $Q_{ab}$ . But then by Path Rule 8, we would have added  $yv$  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  $xx_T$  is the 1-edge incident on  $x_T$  whose representative is colored with  $c_3$ . By Coloring Rule 2, there exists 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  $xx_T$  is the only 1-edge with color  $s(T')$  by Lemma 2, we have that  $e' = xx_T$ . Hence,  $xx_T$  is in  $Q_{ab}$ . But if  $xx_T$  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 7 and 4. 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  $xx_T$ , 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  $ST(x_T, x_{T'}) = ST(x_T, x)$ , 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.  $\square$

The following Lemma follows from the previous coloring rules.

**Lemma 14.** Consider an edge  $e$  in  $B_1$  that remains uncolored after the application of Coloring Rules 1 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  $u \in \{v, x_T\}$  has 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.  $\square$

**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 14 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  $e'$ .

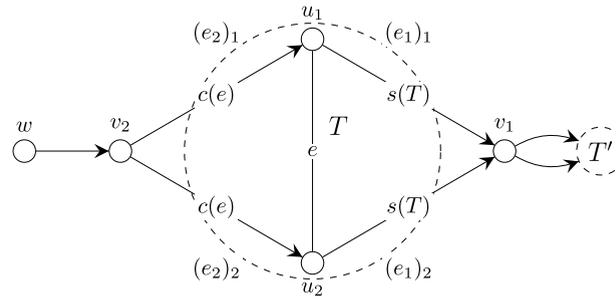


Fig. 10. Coloring Rule 10.

**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 15.** Invariant 1 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.  $\square$

**Lemma 16.** Consider a 2-edge  $e$  incident on tree-vertex  $x_T$  that remains uncolored after the application of Rules 1 to 9. Then,  $x_T$  has 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 14 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.  $\square$

**Coloring Rule 10.** For each incomplete tree-vertex  $x_T$  whose parent's outgoing edge is a 2-edge: (See Fig. 10 for an Illustration). Let  $v_1$  be the parent of  $x_T$ . From Lemma 16, 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 endpoint 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 14, we know that  $v_1$  and  $v_2$  have degree exactly 2. Let  $\overrightarrow{v_1x_T}$  be the outgoing 2-edge from  $v_1$  and let  $\overrightarrow{wv_2}$  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.

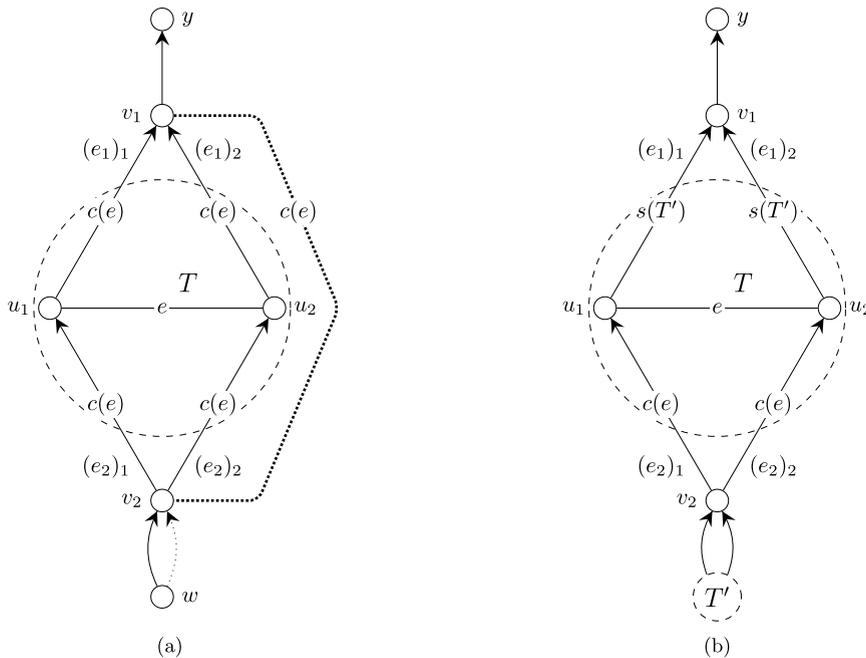
- If  $Q_{ab}$  contains both  $e_1$  and  $e_2$ , add  $v_1u_1$  and  $u_1v_2$  to  $P_{ab}$ .
- If  $Q_{ab}$  contains exactly one edge among  $e_1$  and  $e_2$ , then either  $a$  or  $b$  is in  $V(T)$ . Also both of them cannot be in  $V(T)$ . Let  $z$  be the one among  $a$  or  $b$  that is in  $V(T)$ . If  $Q_{ab}$  contains  $e_1$ , add  $v_1z$  to  $P_{ab}$ ; otherwise, i.e., if  $Q_{ab}$  contains  $e_2$ , add  $v_2z$  to  $P_{ab}$ .

**Lemma 17.** Invariant 1 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:**  $c(d_1) = c(d_2) = c(e)$ .



**Fig. 11.** Coloring Rule 11. **(a)** Case 1; here the edge  $v_1v_2$  is drawn as a thick dotted line to highlight that it is not in  $B_1$ , and the edge  $wv_2$  is drawn with one solid line and one dotted line to denote that it could be a 1-edge or a 2-edge **(b)** Case 2; here  $w = x_T$ .

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  $h(d_1) = v_1x_T$ . 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 4 or 3.

Let  $T''$  be such that  $d_2$  is adjacent on  $x_{T''}$ . Then, by Lemmas 7 and 4, we know that  $Q_{ab}$  does not intersect  $ST(x_{T''}, x_T)$ , in particular  $Q_{ab}$  does not contain  $x_T$ . Since  $h(d_1) = v_1x_T$ , 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 16, 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 endpoint of  $e_1$  and  $v_2$  be the other endpoint of  $e_2$ . Let  $e = u_1u_2$  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 14, we know that  $v_1$  and  $v_2$  have degree exactly 2. Let  $\overrightarrow{v_1 y}$  be the outgoing edge from  $v_1$  and  $\overleftarrow{w v_2}$  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 Fig. 11 (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 Fig. 11 (b) for an illustration).

We will prove in Lemma 24 that  $w v_2$  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 18.** Invariant 1 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  $x_{T'} v_2$  or  $v_2 x_T$ . But since both  $x_{T'} v_2$  and  $v_2 x_T$  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 to  $x_{T''}$ . By Lemmas 7 and 4, we know that  $Q_{ab}$  does not intersect  $ST(x_{T''}, x_{T'})$ . Then  $x_T$  is not in  $ST(x_{T''}, x_{T'})$ . 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  $d_2$  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 application of Coloring Rule 10.

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 25, we will prove that Coloring Rule 10 was not applied on  $x_{T'}$ . Thus, we have a contradiction.



**Fig. 12.** An illustration of the proof of Lemma 19. 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 new skeleton  $B'$  as described in the proof.



**Fig. 13.** An illustration of the transformation in the proof of Lemma 20. 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.

**Case 2.6**  $d_2$  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.  $\square$

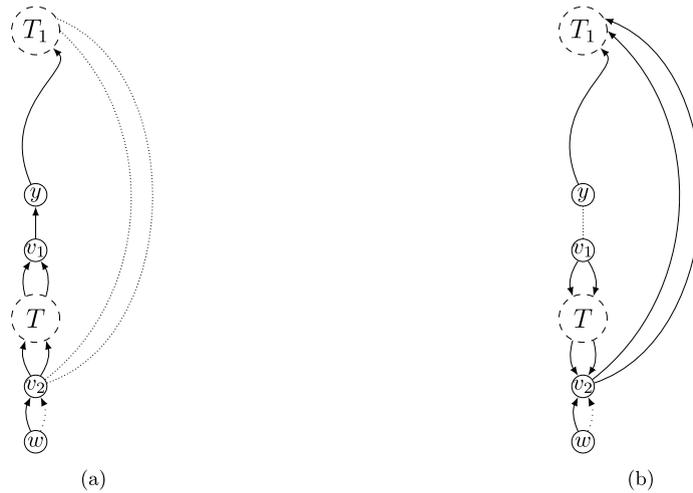
Now we proceed towards proving Lemmas 24 and 25 that were used above. For this we need to prove some auxiliary lemmas first.

**Lemma 19.** *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  $v x_{T_1}$  is a 2-edge in  $H$  then  $x_{T_1}$  is a child of  $v$  in  $B_1$ .*

**Proof.** See Fig. 12 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 level of all vertices in the subtree rooted at  $x_{T_1}$  decrease by 1 and the level of all other vertices remain the same. Thus  $B'$  has a lexicographically higher configuration vector than  $B$ . Thus we have a contradiction to the choice of  $B$ .  $\square$

**Lemma 20.** *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 Fig. 13 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  $v_1 y$  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  $v_1 x_{T_1}$  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 leaves of  $H$ ). Thus, since the edge  $v_1 x_{T_1}$  is not in  $B_1$ , it is not in  $B$  also (recall  $B_1 = B \setminus L_S$ ). Thus,  $v_1 x_{T_1} \in E(H) \setminus E(B)$ . Also  $x_{T_1}$  is not a descendant of  $v_1$  due to Lemma 19. Thus,  $x_{T_1} \in ST(v_1, y)$ . Then, by deleting the 1-edge  $\overrightarrow{v_1 y}$  from  $B$  and adding the 2-edge  $\overrightarrow{v_1 x_{T_1}}$ , 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  $B$ .  $\square$



**Fig. 14.** An illustration of the transformation in the proof of Lemma 21. 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. 15.** An illustration of the transformation in the proof of Lemma 22. The densely dotted edges denote the edges of  $H$  that are not in  $B$ . (a) The scenario before the transformation, and (b) the transformation to the new forest  $\mathcal{F}'$  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.

**Lemma 21.** 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  $V(H) \setminus \{w, x_T\}$ .

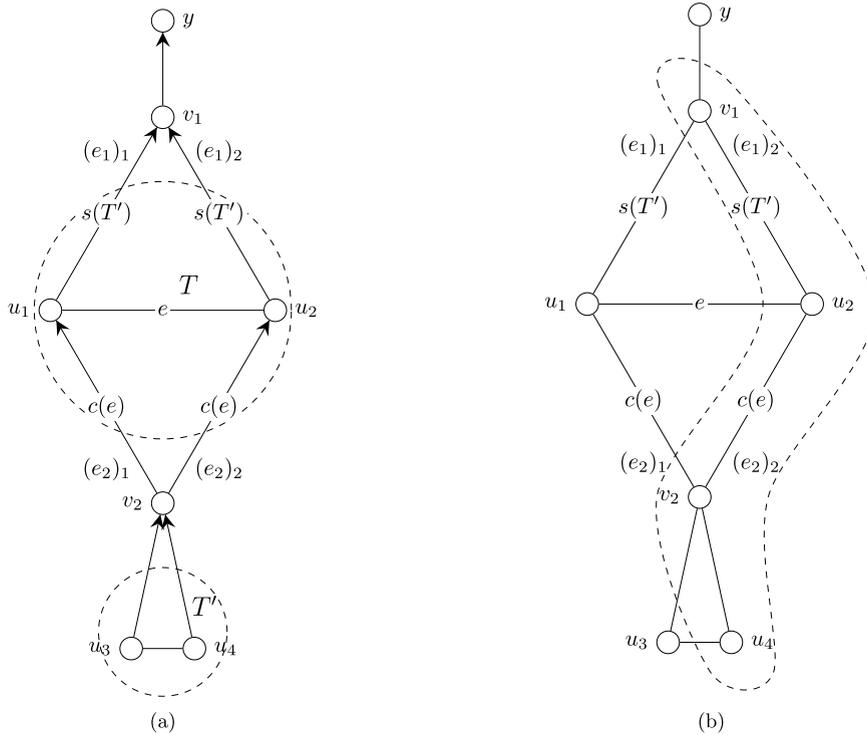
**Proof.** See Fig. 14 for an illustration of the proof. Suppose for the sake of contradiction that  $v_2$  has a 2-edge in  $H$  to a tree-vertex  $x_{T_1} \in V(H) \setminus \{x_T, w\}$ . Then the edge  $v_2x_{T_1}$  is not in  $B_1$  as the only neighbors of  $v_2$  in  $B_1$  are  $x_T$  and  $w$ . Since  $x_{T_1}$  is a tree-vertex, it is not in  $L_S$  (recall that  $L_S$  is the set of non-tree leaves 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 19. Clearly, then  $x_{T_1} \in ST(v_2, x_T)$ . 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  $B$ .  $\square$

**Lemma 22.** 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  $v$  does not have a 2-edge to any tree-vertex except  $x_T$  in  $H$ , then  $v$  does not have edges to any tree-vertex in  $H$  except  $x_T$ .

**Proof.** See Fig. 15 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  $\mathcal{F}'$  as  $F' := (F \setminus \{u_1\}) \cup \{v\}$  and  $\mathcal{F}' := G[F']$ . We claim that  $\mathcal{F}'$  is a forest with fewer trees than  $\mathcal{F}$ , which is a contradiction to the choice of  $\mathcal{F}$ . (Recall that out of all maximum induced forests, we picked  $\mathcal{F}$  to be one having the fewest number of trees). Since  $v$  has at most one edge to any tree in  $G[F \setminus \{u_1\}]$ , it is clear that  $\mathcal{F}'$  is indeed a forest. The number of trees in  $\mathcal{F}'$  is at least one smaller than that of  $\mathcal{F}$  because  $\{u_2, v\} \cup V(T_1)$  induces a single tree in  $\mathcal{F}'$ .  $\square$

**Lemma 23.** 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), and



**Fig. 16.** An illustration of the transformation in the proof of Lemma 25. (a) The initial scenario before transformation, and (b) the transformation to the new forest  $\mathcal{F}'$  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.

- 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 20, and 22 and the fact that  $v_1y$  is a 1-edge, and the second statement follows from Lemmas 21 and 22.  $\square$

**Lemma 24.** 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 Fig. 11). Let  $F' = (F \setminus \{u_1\} \cup \{v_1, v_2\})$ . Let  $\mathcal{F}' = G[F']$ . To see that  $\mathcal{F}'$  is a forest, observe that by Lemma 23, both  $v_1$  and  $v_2$  are not adjacent in  $H$  to any tree-vertex except  $x_T$ . Then since  $|F'| > |F|$ , we have that  $\mathcal{F}$  is not a maximum induced forest, a contradiction.  $\square$

**Lemma 25.** 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 Fig. 11b). 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 Fig. 16a). 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  $\mathcal{F}' := G[F']$ . Note that  $|F'| = |F|$ . We claim that  $\mathcal{F}'$  is a forest with fewer trees than  $\mathcal{F}$ , thereby showing a contradiction to the choice of  $\mathcal{F}$ . (Recall that out of all maximum induced forests, we picked  $\mathcal{F}$  to be one having the fewest number of trees). To see that  $\mathcal{F}'$  is indeed a forest, observe that  $v_1$  does not have an edge to any tree-vertex in  $H$  except  $x_T$  (by Lemma 23),  $v_1$  and  $v_2$  are not adjacent in  $G$  (since the precondition of Case 1 of Rule 11 is not satisfied), and that  $v_2$  does not have an edge in  $H$  to any tree-vertex in  $B_1$  except  $x_T$  and  $x_{T'}$  (by Lemma 21). It is clear from construction that  $\mathcal{F}'$  has at least one tree less than  $\mathcal{F}$ , which concludes the proof.  $\square$

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 26.** 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.  $\square$

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 1. 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 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 26. We define path  $P$  as follows. If  $a_1, a_2 \in L_S$ , then  $P := a_1aP_{ab}ba_2$ . If  $a_1 \in L_S$  but  $a_2 \notin L_S$ , then  $P := a_1aP_{ab}$ . If  $a_2 \in L_S$  but  $a_1 \notin L_S$ , then  $P := P_{ab}ba_2$ . 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.  $\square$

### 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 GraphHedron [16].

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

Another interesting direction is to discover other graph parameters that yield tight bounds on the rainbow connection number of general graphs or of particular graph classes.

#### Declaration of competing interest

None declared.

#### 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, A. Ramaswamy, Rainbow connection number and radius, *Graphs Comb.* 30 (2) (2014) 275–285.
- [2] Y. Caro, A. Lev, Y. Roditty, Z. Tuza, R. Yuster, On rainbow connection, *Electron. J. Comb.* 15 (1) (2008) R57.
- [3] S. Chakraborty, E. Fischer, A. Matsliah, R. Yuster, Hardness and algorithms for rainbow connection, *J. Comb. Optim.* 21 (3) (2009) 330–347.
- [4] L.S. Chandran, A. Das, D. Rajendraprasad, N.M. Varma, Rainbow connection number and connected dominating sets, *J. Graph Theory* 71 (2) (2012) 206–218.
- [5] G. Chartrand, G.L. Johns, K.A. McKeon, P. Zhang, Rainbow connection in graphs, *Math. Bohem.* 133 (1) (2008).
- [6] G. Chartrand, P. Zhang, *Chromatic Graph Theory*, CRC Press, 2008.
- [7] P. Dorbec, I. Schiermeyer, E. Sidorowicz, É. Sopena, Rainbow connection in oriented graphs, *Discrete Appl. Math.* 179 (2014) 69–78.
- [8] J. Ekstein, P. Holub, T. Kaiser, M. Koch, S. Matos Camacho, Z. Ryjáček, I. Schiermeyer, The rainbow connection number of 2-connected graphs, *Discrete Math.* 313 (19) (2013) 1884–1892.
- [9] S.N. Joseph, L.C. DiPippo, *Pseudo-Scheduling: A New Approach to the Broadcast Scheduling Problem*, Springer, Berlin Heidelberg, 2013, pp. 93–104.
- [10] N. Kamčev, M. Krivelevich, B. Sudakov, Some remarks on rainbow connectivity, *J. Graph Theory* 83 (4) (2016) 372–383.
- [11] M. Krivelevich, R. Yuster, The rainbow connection of a graph is (at most) reciprocal to its minimum degree, *J. Graph Theory* 63 (3) (2010) 185–191.
- [12] J. Lauri, *Chasing the Rainbow Connection: Hardness, Algorithms, and Bounds*, PhD thesis, Tampere University of Technology, 2016.
- [13] X. Li, Y. Shi, Y. Sun, Rainbow connections of graphs: a survey, *Graphs Comb.* 29 (1) (2012) 1–38.
- [14] X. Li, Y. Sun, *Rainbow Connections of Graphs*, Springer, 2012.
- [15] X. Li, Y. Sun, An updated survey on rainbow connections of graphs—a dynamic survey, *Theory Appl. Graphs* 1 (2017) 3.
- [16] H. Mélot, Facet defining inequalities among graph invariants: the system GraPHedron, *Discrete Appl. Math.* 156 (10) (2008) 1875–1891.