



**HAL**  
open science

## Finding a Small Number of Colourful Components

Laurent Bulteau, Konrad K. Dabrowski, Guillaume Fertin, Matthew Johnson,  
Daniël Paulusma, Stéphane Vialette

► **To cite this version:**

Laurent Bulteau, Konrad K. Dabrowski, Guillaume Fertin, Matthew Johnson, Daniël Paulusma, et al.. Finding a Small Number of Colourful Components. CPM 2019, Jul 2019, Pise, Italy. 10.4230/LIPIcs.CPM.2019.20 . hal-02304030

**HAL Id: hal-02304030**

**<https://hal.science/hal-02304030>**

Submitted on 19 Nov 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Finding a Small Number of Colourful Components\*

**Laurent Bulteau**

Université Paris-Est, LIGM (UMR 8049), CNRS, ENPC, UPEM, ESIEE Paris, France  
laurent.bulteau@u-pem.fr

**Konrad K. Dabrowski**<sup>1</sup>

Department of Computer Science, Durham University, Durham, UK  
konrad.dabrowski@durham.ac.uk

**Guillaume Fertin**

Université de Nantes, LS2N (UMR 6004), CNRS, Nantes, France  
guillaume.fertin@univ-nantes.fr

**Matthew Johnson**

Department of Computer Science, Durham University, Durham, UK  
matthew.johnson2@durham.ac.uk

**Daniël Paulusma**<sup>2</sup>

Department of Computer Science, Durham University, Durham, UK  
daniel.paulusma@durham.ac.uk

**Stéphane Vialette**

Université Paris-Est, LIGM (UMR 8049), CNRS, ENPC, UPEM, ESIEE Paris, France  
vialette@u-pem.fr

---

## Abstract

A partition  $(V_1, \dots, V_k)$  of the vertex set of a graph  $G$  with a (not necessarily proper) colouring  $c$  is colourful if no two vertices in any  $V_i$  have the same colour and every set  $V_i$  induces a connected graph. The COLOURFUL PARTITION problem is to decide whether a coloured graph  $(G, c)$  has a colourful partition of size at most  $k$ . This problem is closely related to the COLOURFUL COMPONENTS problem, which is to decide whether a graph can be modified into a graph whose connected components form a colourful partition by deleting at most  $p$  edges. Nevertheless we show that COLOURFUL PARTITION and COLOURFUL COMPONENTS may have different complexities for restricted instances. We tighten known NP-hardness results for both problems and in addition we prove new hardness and tractability results for COLOURFUL PARTITION. Using these results we complete our paper with a thorough parameterized study of COLOURFUL PARTITION.

**2012 ACM Subject Classification** Mathematics of computing → Graph theory

**Keywords and phrases** colourful component, colourful partition, tree, vertex cover

## 1 Introduction

Research in comparative genomics, which studies the structure and evolution of genomes from different species, has motivated a number of interesting graph colouring problems. In this paper we focus on the multiple genome alignment problem, where one takes a set of sequenced genomes, lets the genes be the vertex set of a graph  $G$ , and joins by an edge any pair of genes whose *similarity* (determined by their nucleotide sequences) exceeds a given

---

\* This paper was supported by Research Project Grant RPG-2016-258 of the Leverhulme Trust.

<sup>1</sup> Also supported by EPSRC (EP/K025090/1)

<sup>2</sup> Also supported by EPSRC (EP/K025090/1)

threshold. The vertices are also coloured to indicate the species to which each gene belongs. This leads to a *coloured graph*  $(G, c)$ , where  $c : V(G) \rightarrow \{1, 2, \dots\}$  denotes the colouring of  $G$ . We emphasize that  $c$  is not necessarily proper (adjacent vertices may have the same colour).

One seeks to better understand the evolutionary processes affecting these genomes by attempting to partition the genes into *orthologous* sets (that is, collections of genes that originated from the same ancestral species but diverged following a speciation event) that are sufficiently similar. This translates into partitioning  $V(G)$  such that each part

- (i) contains no more than one vertex of each colour, and
- (ii) induces a connected component.

These two conditions ensure that each part contains vertices representing an orthologous set of similar genes. In addition, one seeks to find a partition that is in some sense optimal.

When Zheng et al. [17] considered this model, one approach they followed was to try to delete as few edges as possible such that the connected components of the resulting graph  $G'$  are *colourful*, that is, contain no more than one vertex of any colour; in this case the connected components give the partition of  $V(G)$ . This led them to the ORTHOGONAL PARTITION problem, introduced in [8] and also known as the COLOURFUL COMPONENTS problem [1, 3, 7]. (Note that a graph is *colourful* if each of its components is colourful.)

**COLOURFUL COMPONENTS**

*Instance:* A coloured graph  $(G, c)$  and an integer  $p \geq 1$

*Question:* Is it possible to modify  $G$  into a colourful graph  $G'$  by deleting at most  $p$  edges?

The central problem of this paper is a natural variant, COLOURFUL PARTITION, which was introduced by Adamaszek and Popa [1]. A partition is *colourful* if every partition class induces a connected colourful graph. The *size* of a partition is its number of partition classes.

**COLOURFUL PARTITION**

*Instance:* A coloured graph  $(G, c)$  and an integer  $k \geq 1$

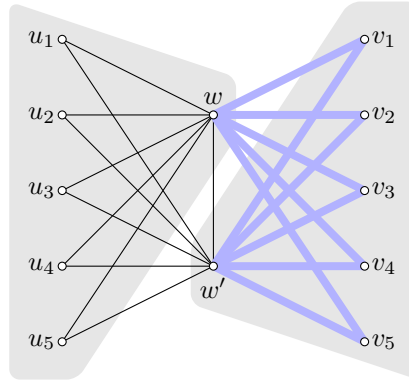
*Question:* Does  $(G, c)$  have a colourful partition of size at most  $k$ ?

The following example demonstrates that COLOURFUL COMPONENTS and COLOURFUL PARTITION are not the same: a colourful partition with the fewest parts might require the deletion of more edges than the minimum needed to obtain a colourful graph. Later we find a family of instances on which the problems have different complexities (see Corollary 11).

► **Example 1.** Let  $G$  have vertices  $u_1, \dots, u_k, v_1, \dots, v_k, w, w'$  and edges  $ww'$  and  $u_iw, u_iw', v_iw, v_iw'$  for  $i \in \{1, \dots, k\}$ . Let  $c$  assign colour  $i$  to each  $u_i$  and  $v_i$ , colour  $k + 1$  to  $w$  and colour  $k + 2$  to  $w'$ . Then  $(\{u_1, \dots, u_k, w\}, \{v_1, \dots, v_k, w'\})$  is a colourful partition for  $(G, c)$  of size 2, which we obtain by deleting  $2k + 1$  edges. However, deleting the  $2k$  edges  $v_iw$  and  $v_iw'$  for  $i \in \{1, \dots, k\}$  also yields a colourful graph (with a colourful partition of size  $k + 1$ ). See Figure 1 for an illustration.

## 1.1 Known Results

Adamaszek and Popa [1] proved, among other results, that COLOURFUL PARTITION does not admit a polynomial-time approximation within a factor of  $n^{\frac{1}{14} - \epsilon}$ , for any  $\epsilon > 0$  (assuming  $P \neq NP$ ). A coloured graph  $(G, c)$  is  *$\ell$ -coloured* if  $1 \leq c(u) \leq \ell$  for all  $u \in V(G)$ . Bruckner



■ **Figure 1** The graph  $G$  of Example 1 with  $k = 5$  (the colouring is not indicated, but recall that  $u_i$  and  $v_i$  are coloured alike and otherwise colours are distinct). The shaded areas represent a colourful partition of minimum size; there are 11 edges between the two parts. The 10 highlighted edges are a smallest set whose deletion leaves a colourful graph.

et al. [3] proved the following two results for COLOURFUL COMPONENTS. The first result follows from observing that for  $\ell = 2$ , the problem becomes a maximum matching problem in a bipartite graph after removing all edges between vertices coloured alike. This observation can also be used for COLOURFUL PARTITION.

► **Theorem 2** ([3]). COLOURFUL PARTITION and COLOURFUL COMPONENTS are polynomial-time solvable for 2-coloured graphs.

► **Theorem 3** ([3]). COLOURFUL COMPONENTS is NP-complete for 3-coloured graphs of maximum degree 6.

The situation for trees is different than for general graphs (see Example 1). A tree  $T$  has a colourful partition of size  $k$  if and only if it can be modified into a colourful graph by at most  $k - 1$  edge deletions. Hence, the problems COLOURFUL PARTITION and COLOURFUL COMPONENTS are equivalent for trees. The following hardness and FPT results are due to Bruckner et al. [3] and Dondi and Sikora [7]. Note that trees of diameter at most 3 are stars and double stars (the graph obtained from two stars by adding an edge between their central vertices), for which both problems are readily seen to be polynomial-time solvable. A *subdivided star* is the graph obtained by subdividing the edges of a star.

► **Theorem 4** ([3]). COLOURFUL PARTITION and COLOURFUL COMPONENTS are polynomial-time solvable for coloured trees of diameter at most 3, but NP-complete for coloured trees of diameter 4.

► **Theorem 5** ([7]). COLOURFUL PARTITION and COLOURFUL COMPONENTS are polynomial-time solvable for coloured paths (which have path-width at most 1), but NP-complete for coloured subdivided stars (which are trees of path-width at most 2).

► **Theorem 6** ([3]). COLOURFUL PARTITION and COLOURFUL COMPONENTS are FPT for coloured trees, when parameterized by the number of colours.

► **Theorem 7** ([7]). COLOURFUL PARTITION and COLOURFUL COMPONENTS are FPT for coloured trees, when parameterized by the number of colourful components (or equivalently, by the size of a colourful partition).

In fact, Dondi and Sikora [7] found a cubic kernel for coloured trees when parameterized by  $k$  and also gave an  $O^*(1.554^k)$ -time exact algorithm for coloured trees. In addition to Theorem 6, Bruckner et al. [3] showed that COLOURFUL COMPONENTS is FPT for general coloured graphs when parameterized by the number of colours  $\ell$  and the number of edge deletions  $p$ .

Misra [13] considered the size of a vertex cover as the parameter.

► **Theorem 8** ([13]). COLOURFUL COMPONENTS is FPT when parameterized by vertex cover number.

## 1.2 Our Aims and Results

Our main focus is the COLOURFUL PARTITION problem, and our research aims are:

1. to tighten known NP-hardness results and generalize tractability results for COLOURFUL PARTITION (Section 3);
2. to prove new results analogous to those known for COLOURFUL COMPONENTS (Section 3) and to show that the two problems are different (Sections 3 and 4); and
3. to use the new results to determine suitable parameters for obtaining FPT results (Section 4).

Along the way we also prove some new results for COLOURFUL COMPONENTS.

First, we show an analogue of Theorem 3 by proving that COLOURFUL PARTITION is NP-complete even for 3-coloured 2-connected planar graphs of maximum degree 3. This result is best possible, as COLOURFUL PARTITION is polynomial-time solvable for 2-coloured graphs (Theorem 2) and for graphs of maximum degree 2 (trivial). We show that it also gives us a family of instances on which COLOURFUL COMPONENTS and COLOURFUL PARTITION have different complexities.

Second, we focus on coloured trees. Due to Theorem 6, COLOURFUL PARTITION is polynomial-time solvable for  $\ell$ -coloured trees for any fixed  $\ell$ . Hence the coloured trees in the hardness proofs of Theorems 4 and 5 use an arbitrarily large number of colours. They also have arbitrarily large degree. Moreover, they have arbitrarily large degree. We define the *colour-multiplicity* of a coloured graph  $(G, c)$  as the maximum number of vertices in  $G$  with the same colour. We prove that COLOURFUL PARTITION is NP-complete even for coloured trees of maximum degree 6 and colour-multiplicity 2. As both problems are equivalent on trees, we obtain the same result for COLOURFUL COMPONENTS (note that the graphs in the proof of Theorem 3 are not trees).

Third, we fix the number of colourful components. Note that  $k$ -COLOURFUL PARTITION is polynomial-time solvable for coloured trees (due to Theorem 7) and for  $\ell$ -coloured graphs for any fixed  $\ell$  (such graphs have at most  $k\ell$  vertices). We prove that 2-COLOURFUL PARTITION is NP-complete for split graphs and for coloured planar bipartite graphs of maximum degree 3 and path-width 3, but polynomial-time solvable for coloured graphs of treewidth at most 2. The latter two results also complement Theorem 5, which implies NP-completeness of COLOURFUL PARTITION for path-width 2.

In Section 4 we show that COLOURFUL PARTITION and COLOURFUL COMPONENTS are FPT when parameterized by the treewidth and the number of colours, which generalizes Theorem 6. Our choice for this combination of parameters was guided by our results from Section 3. As we will argue at the start of Section 4, for the six natural parameters: number of colourful components; maximum degree; number of colours; colour-multiplicity; path-width; treewidth; and all their possible combinations, we now obtained either para-NP-completeness or an FPT algorithm for COLOURFUL PARTITION. This motivates the search

for other parameters for this problem. As the vertex cover number of the subdivided stars in the proof of Theorem 5 can be arbitrarily large, it is natural to consider this parameter. We provide an analogue to Theorem 8 by proving that COLOURFUL PARTITION is FPT when parameterized by vertex cover number. A vertex  $u$  of a coloured graph  $(G, c)$  is *uniquely coloured* if its colour  $c(u)$  is not used on any other vertex of  $G$ . It is easy to show NP-hardness for COLOURFUL PARTITION and COLOURFUL COMPONENTS for instances with no uniquely coloured vertices (see also Theorem 12). We also prove that COLOURFUL COMPONENTS is para-NP-hard when parameterized by the number of non-uniquely coloured vertices, whereas COLOURFUL PARTITION is FPT in this case. Thus there are families of instances on which the two problems have different parameterized complexities.

## 2 Preliminaries

All graphs in this paper are simple, with no loops or multiple edges. Let  $G = (V, E)$  be a graph. A subset  $U \subseteq V$  is *connected* if it induces a connected subgraph of  $G$ . For a vertex  $u \in V$ ,  $N(u) = \{v \mid uv \in E\}$  is the *neighbourhood* of  $u$ , and  $\deg(u) = |N(u)|$  is the *degree* of  $u$ . A tree is *binary* if every vertex in it has degree at most 3. A graph is *cubic* if every vertex has degree exactly 3. A connected graph on at least three vertices is *2-connected* if it does not contain a vertex whose removal disconnects the graph. We say that a graph  $G = (V, E)$  is *split* if  $V$  can be partitioned into two (possibly empty) sets  $K$  and  $I$ , where  $K$  is a clique and  $I$  is an independent set. A mapping  $c : E \rightarrow \{1, 2, 3\}$  is a *proper 3-edge colouring* of  $G$  if  $c(e) \neq c(f)$  for any two distinct edges  $e$  and  $f$  that share a common end-vertex. A set  $S \subseteq V$  is a *vertex cover* of a graph  $G = (V, E)$  if  $G - S$  is an independent set. The VERTEX COVER problem asks if a given graph has a vertex cover of size at most  $s$  for a given integer  $s$ . The *vertex cover number*  $\text{vc}(G)$  of a graph  $G$  is the minimum size of a vertex cover in  $G$ . We will use the following lemma.

► **Lemma 9** ([4]). VERTEX COVER is NP-complete for 2-connected cubic planar graphs with a proper 3-edge colouring given as input.

A *tree decomposition* of a graph  $G$  is a pair  $(T, \mathcal{X})$  where  $T$  is a tree and  $\mathcal{X} = \{X_i \mid i \in V(T)\}$  is a collection of subsets of  $V(G)$ , called *bags*, such that

1.  $\bigcup_{i \in V(T)} X_i = V(G)$ ;
2. for every edge  $xy \in E(G)$ , there is an  $i \in V(T)$  such that  $x, y \in X_i$ ; and
3. for every  $x \in V(G)$ , the set  $\{i \in V(T) \mid x \in X_i\}$  induces a connected subtree of  $T$ .

The *width* of  $(T, \mathcal{X})$  is  $\max\{|X_i| - 1 \mid i \in V(T)\}$ , and the *treewidth* of  $G$  is the minimum width over all tree decompositions of  $G$ . If  $T$  is a path, then  $(T, \mathcal{X})$  is a *path decomposition* of  $G$ . The *path-width* of  $G$  is the minimum width over all path decompositions of  $G$ .

A tree decomposition  $(T, \mathcal{X})$  of a graph  $G$  is *nice* if  $T$  is a rooted binary tree such that the nodes of  $T$  are of four types:

1. a *leaf node*  $i$  is a leaf of  $T$  with  $X_i = \emptyset$ ;
2. an *introduce node*  $i$  has one child  $i'$  with  $X_i = X_{i'} \cup \{v\}$  for some vertex  $v \in V(G)$ ;
3. a *forget node*  $i$  has one child  $i'$  with  $X_i = X_{i'} \setminus \{v\}$  for some vertex  $v \in V(G)$ ; and
4. a *join node*  $i$  has two children  $i'$  and  $i''$  with  $X_i = X_{i'} = X_{i''}$ ,

and, moreover, the root  $r$  is a forget node with  $X_r = \emptyset$ . Kloks [11] proved that every tree decomposition of a graph can be converted in linear time to a nice tree decomposition of the same width such that the size of the obtained tree is linear in the size of the original tree.

A *literal* is a (propositional) variable  $x$  (*positive literal*) or a negated variable  $\bar{x}$  (*negative literal*). A set  $S$  of literals is *tautological* if  $S \cap \bar{S} \neq \emptyset$ , where  $\bar{S} = \{\bar{x} \mid x \in S\}$ . A *clause* is a finite non-tautological set of literals. A (*CNF*) *formula* is a finite set of clauses. For  $k \geq 1$ , a *k-formula* is a formula in which every clause contains exactly  $k$  pairwise distinct variables. A *k-formula* is *positive* if every clause in it contains only positive literals. A *truth assignment* for a set of variables  $X$  is a mapping  $\tau : X \rightarrow \{\mathbf{true}, \mathbf{false}\}$ . The *k-SATISFIABILITY* problem asks whether there exists a truth assignment for a given *k-formula*  $F$  such that every clause of  $F$  contains at least one true literal. The NOT-ALL-EQUAL POSITIVE 3-SATISFIABILITY problem asks whether there exists a truth assignment for a given positive 3-formula  $F$  such that every clause of  $F$  contains at least one true literal and at least one false literal. For both problems we say that such a desired truth assignment *satisfies*  $F$ . Both 3-SATISFIABILITY [10] and NOT-ALL-EQUAL POSITIVE 3-SATISFIABILITY [15] are NP-complete. In contrast, it is well known and readily seen that 2-SATISFIABILITY can be solved in polynomial time [12].

Let  $X$  be a set. Then  $\mathcal{P}(X) = \{Y \mid Y \subseteq X\}$  is the *power set* of  $X$ , that is, the set of subsets of  $X$ . A *partition*  $P$  of  $X$  is a subset of  $\mathcal{P}(X)$  such that  $\emptyset \notin P$  and every element of  $X$  is in exactly one set in  $P$ . A *block* is an element of a partition  $P$  and the *size* of  $P$  is  $|P|$ . Given two partitions  $P$  and  $Q$ , we say that  $P$  is *coarsening* of  $Q$  if for every  $Y \in Q$  there is a  $Y' \in P$  such that  $Y \subseteq Y'$ . Let  $\emptyset : \emptyset \rightarrow \emptyset$  denote the trivial function.

### 3 Classical Complexity

We will prove four hardness results and one polynomial-time result on COLOURFUL PARTITION. Note that COLOURFUL PARTITION belongs to NP. We start by proving the following result.

► **Theorem 10.** COLOURFUL PARTITION is NP-complete for 3-coloured 2-connected planar graphs of maximum degree 3.

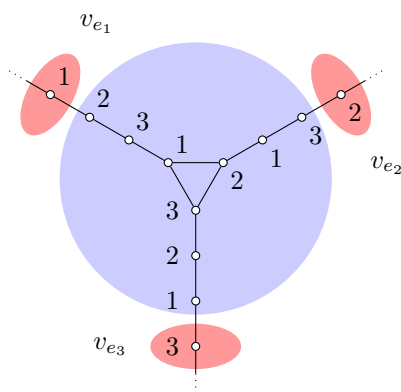
**Proof.** We use a reduction from VERTEX COVER. By Lemma 9 we may assume that we are given a 2-connected cubic planar graph  $G$  with a proper 3-edge colouring  $c$ . From  $G$  and  $c$  we construct a coloured graph  $(G', c')$  as follows.

- For each  $e \in E(G)$  with  $c(e) = i$ , create a vertex  $v_e$  with colour  $c'(v_e) = i$  in  $G'$  (a *red* vertex).
- For every vertex  $v \in V(G)$  with incident edges  $e_1, e_2, e_3$  (with colours 1, 2, 3, respectively), create a copy of the 3-coloured 9-vertex gadget shown in Figure 2 (a *blue* set), and connect it to the red vertices  $v_{e_1}, v_{e_2}$  and  $v_{e_3}$ , as shown in the same figure.

Note that  $G'$  is 3-coloured, 2-connected, planar and has maximum degree 3. We claim that  $G$  has a vertex cover of size at most  $s$  if and only if  $G'$  has a colourful partition of size at most  $3n + s$ .

First suppose that  $G$  has a vertex cover  $S$  of size at most  $s$ . We create the colourful partition of  $G'$  as described in Figure 3. More precisely, for every vertex  $v \in S$ , we pick four colourful components covering the whole vertex gadget plus all neighbouring red vertices that have not yet been picked. Since  $S$  is a vertex cover of  $G$ , all red vertices of  $G'$  are now part of some colourful component. For every vertex  $v \in V(G) \setminus S$ , we pick three colourful components covering the whole vertex gadget. This yields a colourful partition of  $G'$  of size at most  $4|S| + 3(n - |S|) = 3n + |S| \leq 3n + s$ .

Now suppose that  $G$  has a colourful partition  $(V_1, \dots, V_k)$  for some  $k \leq 3n + s$ . Observe that no  $V_i$  contains vertices from two distinct vertex gadgets; otherwise  $V_i$  would contain



■ **Figure 2** The blue 9-vertex gadget for a vertex  $v$  incident to edges  $e_1, e_2, e_3$  in  $G$ , connected to the three red vertices  $v_{e_1}, v_{e_2}, v_{e_3}$  in  $G'$  in the proof of Theorem 10. A crucial property is that the three red vertices have distinct colours, as  $c$  is a 3-edge colouring of  $G$ , and these three colours fix the colours of the blue vertex-gadget.

a red vertex and its two neighbours, which have the same colour by construction. Hence, every  $V_i$  can be mapped to some vertex  $v_j$  of  $G$  (where we map every colourful component that consists of only a single red vertex to an arbitrary adjacent vertex gadget). Observe that  $|V_i| \leq 3$  for all  $i \in \{1, \dots, k\}$ . As every vertex gadget has size 9, at least three colourful components are mapped to it. Moreover, there cannot be two neighbouring gadgets with only three components, as the two of them plus the red vertex connecting them have 19 vertices in total. So the vertices of  $G$  corresponding to vertex gadgets with at least four components form a vertex cover  $S$  in  $G$ . Since there are at most  $k - 3n \leq s$  such vertex gadgets,  $S$  has size at most  $s$ . ◀

The coloured graph  $(G', c')$  constructed in the proof of Theorem 10 can be modified into a colourful graph by omitting exactly one edge adjacent to each red vertex and exactly three edges inside each blue component. It can be readily checked that this is the minimum number of edges required. Hence COLOURFUL COMPONENTS is polynomial-time solvable on these coloured graphs  $(G', c')$ , whereas COLOURFUL PARTITION is NP-complete. Thus we have:

► **Corollary 11.** *There exists a family of instances on which COLOURFUL COMPONENTS and COLOURFUL PARTITION have different complexities.*

For our next result we need to introduce the following problem.

**MULTICUT**

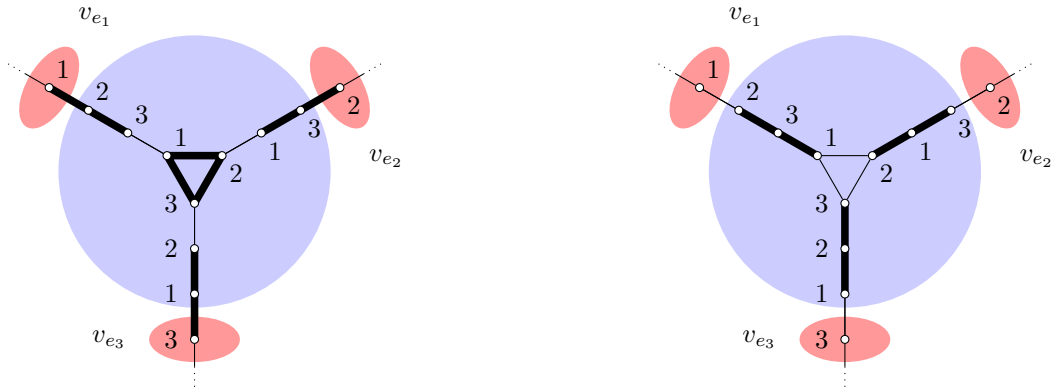
*Instance:* A graph  $G$ , a positive integer  $r$  and a set  $S$  of pairs of distinct vertices of  $G$ .

*Question:* Is there a set of at most  $r$  edges whose removal from  $G$  results in a graph in which  $s$  and  $t$  belong to different connected components for every pair  $(s, t) \in S$ ?

If  $(s, t)$  is a pair in  $S$ , we say that  $s$  and  $t$  are *mates*. Moreover, every vertex that belongs to a pair in  $S$  is a *terminal*. The MULTICUT problem is known to be NP-complete even for binary trees [5].

We are now ready to prove the following theorem. In particular, the last condition in this theorem shows that the number of uniquely coloured vertices is not a useful parameter.





■ **Figure 3** The colourful components (in bold) corresponding to a vertex in the vertex cover  $S$  of  $G$  (left figure) and to a vertex not in  $S$  (right figure) in the proof of Theorem 10.

► **Theorem 12.** COLOURFUL PARTITION and COLOURFUL COMPONENTS are NP-complete for coloured trees with maximum degree at most 6, colour-multiplicity 2, and no uniquely coloured vertices.

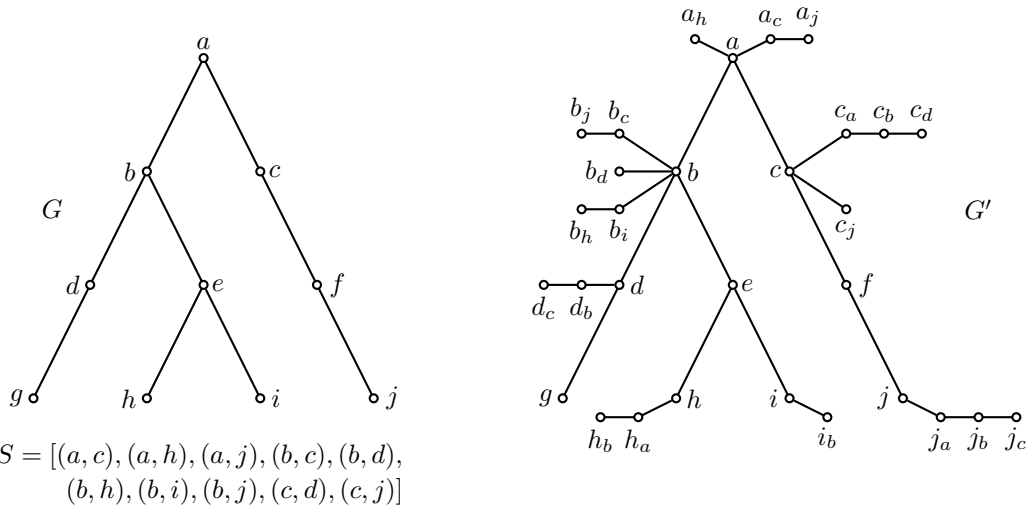
**Proof.** As the two problems are equivalent on trees, it suffices to prove the result for COLOURFUL PARTITION. We first show the result for the case where we allow uniquely coloured vertices. Afterwards we show how to modify our construction to get rid of such vertices.

As mentioned, we reduce from MULTICUT on binary trees. Let  $(G, S, r)$  be an instance of MULTICUT, where  $G$  is a binary tree,  $S$  is a set of terminal pairs and  $r$  is positive integer. We construct a graph  $G'$  that has  $G$  as a subgraph and some additional vertices and edges that we now describe. For a vertex  $v \in V(G)$ , let  $e_v^1, \dots, e_v^{\deg(v)}$  be the edges incident with  $v$  and let  $A_v^1, \dots, A_v^{\deg(v)}$  be a collection of sets whose contents we now define. For each mate  $u$  of  $v$  in the pairs of  $S$ , create a vertex  $v_u$  and place it in the set  $A_v^i$  such that the unique path in  $G$  from  $v$  to  $u$  includes the edge  $e_v^i$ . For each  $A_v^i$ , add edges so that the vertices in  $A_v^i$  induce a path, and add an edge from one end-vertex of this path to  $v$ . Let  $P_v^i$  be the path that includes  $v$  and the vertices of  $A_v^i$ . This completes the construction of  $G'$ . As  $G$  is binary,  $G'$  is a tree of maximum degree at most 6. We now construct a colouring  $c$  on the vertices of  $G'$ . For each pair  $(u, v)$  in  $S$ , we let  $c(v_u) = c(u_v)$  be a colour that is not used on any other vertex of  $G'$ . All other vertices of  $G'$  are assigned a unique colour. Hence  $(G', c)$  has colour-multiplicity 2. See Figure 4 for an illustration.

We claim that it is possible to separate all mates from each other by removing at most  $r$  edges from  $G$  if and only if  $(G', c)$  has a colourful partition of size at most  $r + 1$ .

First suppose that it is possible to separate all mates of  $S$  from each other by removing at most  $r$  edges from  $G$ . We remove the same  $r$  edges from  $G'$  (note that all edges in  $G$  are also in  $G'$ ). The resulting graph consists of at most  $r + 1$  components. Let  $V_1, \dots, V_k$ ,  $k \leq r + 1$ , be the vertex sets of these components. Note that vertices  $u_v$  and  $v_u$  belong to different sets  $V_i$  and  $V_j$ , as  $u$  and  $v$ , being mates, must belong to different components, while  $v_u$  belongs to the same component as  $v$ , and  $u_v$  belongs to the same component as  $u$ . Hence,  $(V_1, \dots, V_k)$  is a colourful partition of  $G'$  that has size at most  $r + 1$ .

Now suppose that  $G'$  has a colourful partition  $(V_1, \dots, V_k)$ ,  $k \leq r + 1$ . As  $G'$  is a tree, we can obtain this partition by removing a set  $F$  of at most  $r$  edges from  $G'$ . If  $F$  contains edges not in  $E(G)$ , then we show that we can modify  $F$  to obtain a set  $F^*$  of at most  $r$  edges, all of which belong to  $E(G)$ , such that removing the edges in  $F^*$  from  $G'$  leaves a graph whose



■ **Figure 4** A graph  $G$  and a set of terminals  $S$  of an instance of MULTICUT and the instance  $G'$  of COLOURFUL PARTITION obtained using the reduction in the proof of Theorem 12. The colouring of  $G'$  is not indicated and the value of  $r$ , which does not affect the construction, is not stated.

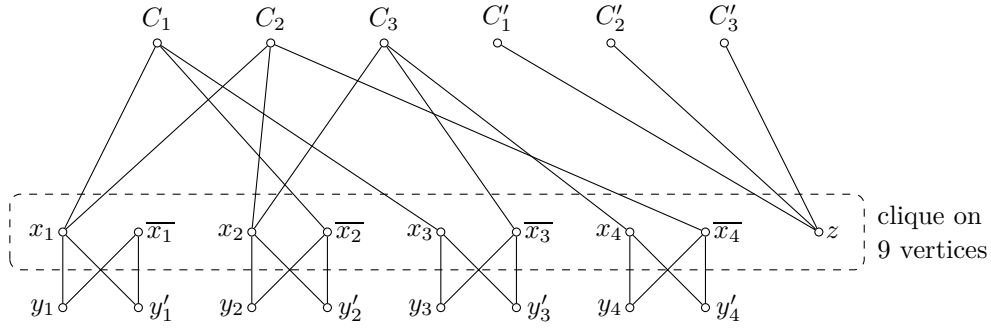
connected components form a colourful partition. To show this, suppose  $F$  contains an edge that belongs to some  $P_w^i$ . In  $F$  we replace *all* the edges of  $P_w^i$  with the edge  $e_w^i$  to obtain a set  $F'$  with  $|F'| \leq |F|$ . We claim that if a pair of vertices  $v_u$  and  $v_v$  are separated by  $F$ , then they are separated by  $F'$  and so removing the edges in  $F'$  from  $G'$  also leaves a colourful partition. This immediately holds if neither  $v_u$  nor  $v_v$  are in  $P_w^i$ . If one of  $v_u, v_v$  does belong to  $P_w^i$ , then  $w \in \{u, v\}$  and the unique path from  $u$  to  $v$  includes  $e_w^i$ , which therefore separates  $v_u$  and  $v_v$ . Thus deleting the vertices of  $F'$  also gives a colourful partition of size at most  $r + 1$ . Furthermore,  $F'$  has a larger intersection with  $E(G)$  than  $F$  does. Hence, by repetition of this modification, we obtain  $F^*$ . Finally we demonstrate that  $F^*$  separates all mates of  $S$ . For each pair  $(u, v) \in S$ , we know that  $v$  and  $v_u$  belong to the same component of  $G - F^*$ , and that  $u$  and  $u_v$  belong to the same component of  $G - F^*$  (as  $F^*$  contains only edges of  $G$ ). Thus as  $F^*$  separates the like-coloured pair  $u_v$  and  $v_u$ , it also separates the mates  $u$  and  $v$ .

We now prove the same result with the addition that there are no uniquely coloured vertices. Let  $(G', c)$  be an instance of COLOURFUL PARTITION constructed above and note that each colour is used at most twice in  $c$ . We take two copies  $G_1$  and  $G_2$  of the graph  $G'$  and for  $v \in V(G)$  let  $v_1$  and  $v_2$  be the corresponding vertices in  $G_1$  and  $G_2$ . We define a new colouring  $c'$  on the disjoint union  $G_1 + G_2$  as follows:

- If  $c$  assigns a colour to a unique vertex  $v$  of  $G$ , then let  $c'(v_1) = c'(v_2) = c(v)$ .
- If  $c$  assigns the same colour to two distinct vertices  $u, v$  in  $G$ , then let  $c'(u_1) = c'(v_1) = c(u)$  and let  $c'$  assign a new colour to  $u_2$  and  $v_2$  that is not used anywhere else.

It is easy to verify that  $(G_1 + G_2, c')$  has a colourful partition of size at most  $2(r + 1)$  if and only if  $(G', c)$  has a colourful partition of size at most  $r + 1$ .

Finally, let  $x$  be a vertex in  $G'$  that has degree 1 (such a vertex must exist if  $G'$  is a tree with at least one edge). We let  $G''$  be the graph obtained from  $G_1 + G_2$  by adding vertices  $y_1$  and  $y_2$  and edges  $x_1y_1, y_1y_2$  and  $y_2x_2$ . Let  $c''$  be the colouring that colours  $y_1$  and  $y_2$  with a new colour that is not used anywhere else and colours all other vertices of  $G''$  in the same way as  $c'$  does. Note that  $G''$  is a tree of maximum degree at most 6 and  $c''$  colours the



■ **Figure 5** The split graph  $G$  constructed using the reduction of the proof of Theorem 13 from the 3-SATISFIABILITY instance  $C_1 = (x_1 \vee \overline{x_2} \vee x_3)$ ,  $C_2 = (x_1 \vee x_2 \vee \overline{x_4})$ ,  $C_3 = (x_2 \vee \overline{x_3} \vee x_4)$ . The colouring is not indicated.

vertices of  $G''$  such that every colour is used exactly twice. Furthermore, in every colourful partition of  $(G'', c'')$ , the vertices  $y_1$  and  $y_2$  must be in different parts of the partition. Since the colours on  $y_1$  and  $y_2$  are not used anywhere else, in any colourful partition of minimum size  $y_1$  must be in the same colourful component as  $x_1$  and  $y_2$  must be in the same colourful component as  $x_2$ . Thus  $(G'', c'')$  has a colourful partition of size at most  $2(r+1)$  if and only if  $(G', c)$  has a colourful partition of size at most  $r+1$ . ◀

We now present our third NP-hardness result.

► **Theorem 13.** 2-COLOURFUL PARTITION is NP-complete for coloured split graphs.

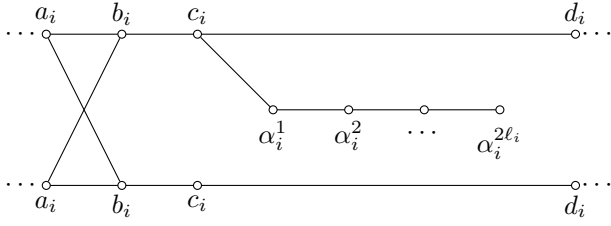
**Proof.** We use a reduction from the NP-complete problem 3-SATISFIABILITY [10]. Let  $(X, \mathcal{C})$  be a 3-formula, where  $X = \{x_1, x_2, \dots, x_n\}$  and  $\mathcal{C} = \{C_1, C_2, \dots, C_m\}$ . From  $(X, \mathcal{C})$  we construct a coloured graph  $(G, c)$  as follows.

- For every  $x_i$  create two vertices  $x_i, \overline{x_i}$ ; we say that these vertices are of  $x$ -type.
- Add an edge between every pair of  $x$ -type vertices.
- For every  $x_i$  create two vertex  $y_i$  and  $y_i'$  and add the edges  $x_i y_i, x_i y_i', \overline{x_i} y_i, \overline{x_i} y_i'$  (the vertices  $x_i$  and  $\overline{x_i}$  will be the only two neighbours of  $y_i$  and  $y_i'$ ).
- Create a vertex  $z$  and add an edge between  $z$  and every  $x$ -type vertex.
- For every clause  $C_j$  create a vertex  $C_j$  and a vertex  $C_j'$ .
- Add an edge between  $z$  and every  $C_j'$  (every  $C_j'$  vertex will have  $z$  as its only neighbour).
- If a clause  $C_j$  consists of literals  $x_g, x_h, x_i$ , then add the edges  $C_j x_g, C_j x_h$  and  $C_j x_i$ .
- Assign a distinct colour to every  $x$ -type vertex, every  $y_i$  vertex, every clause vertex  $C_j$  and  $z$ .
- Set  $c(y_i') = c(y_i)$  for  $i \in \{1, \dots, n\}$  and  $c(C_j') = c(C_j)$  for  $j \in \{1, \dots, m\}$ .

See Figure 5 for an illustration. We note that  $G$  is a split graph, as the set of all  $x$ -type vertices and  $z$  is a clique, whereas the other vertices form an independent set.

We claim that  $\mathcal{C}$  has a satisfying truth assignment if and only if  $(G, c)$  has a colourful partition of size 2.

First suppose that  $\mathcal{C}$  has a satisfying truth assignment  $\tau$ . We put an  $x$ -type vertex in  $V_1$  if and only if  $\tau$  assigns **true** to it. We also put every  $C_j$  and every  $y_i$  in  $V_1$ . As  $\tau$  is a satisfying truth assignment,  $V_1$  is connected. We let  $V_2$  consist of all other vertices and note that by construction  $V_2$  is also connected. Moreover, both  $V_1$  and  $V_2$  are colourful.



■ **Figure 6** The variable gadget for Theorem 14.

Now suppose that  $(G, c)$  has a colourful partition  $(V_1, V_2)$ . We assume without loss of generality that  $z \in V_2$ . Suppose there exists a vertex  $C'_j$  that belongs to  $V_1$ . Then, as  $C'_j$  has  $z$  as its only neighbour,  $V_1$  only contains  $C'_j$ , which means that  $V_2$  contains  $y_1$  and  $y'_1$ , which have the same colour, a contradiction. Hence, every  $C'_j$  must belong to  $V_2$ . As for every  $j \in \{1, \dots, m\}$  the two vertices  $C_j$  and  $C'_j$  have the same colour, this implies that every  $C_j$  belongs to  $V_1$ . Suppose that there exist two vertices  $x_i$  and  $\bar{x}_i$  that both belong to  $V_1$ . As  $y_i$  and  $y'_i$  have the same colour, one of them must belong to  $V_2$ . However, this is not possible, as  $V_2$  is connected and contains  $z$  (and every  $C'_j$ ), but  $x_i$  and  $\bar{x}_i$  are the only neighbours of  $y_i$  and  $y'_i$ . Suppose that there exist two vertices  $x_i$  and  $\bar{x}_i$  that both belong to  $V_2$ . Then, as at least one of  $y_i, y'_i$  belongs to  $V_1$ , we find that  $V_1$ , which contains every  $C_j$ , is not connected. Hence this is also not possible. We conclude that  $x_i$  and  $\bar{x}_i$  belong to different sets  $V_1, V_2$ . We can define a truth assignment  $\tau$  with  $\tau(x) = \mathbf{true}$  if  $x$  belongs to  $V_1$  and  $\tau(x) = \mathbf{false}$  if  $x$  belongs to  $V_2$ . As  $V_1$  is connected, each clause  $C_j$  is adjacent to at least one  $x$ -type vertex in  $V_1$ . So every clause  $C_j$  contains at least one literal to which  $\tau$  assigns  $\mathbf{true}$ . ◀

We now prove our final NP-hardness result.

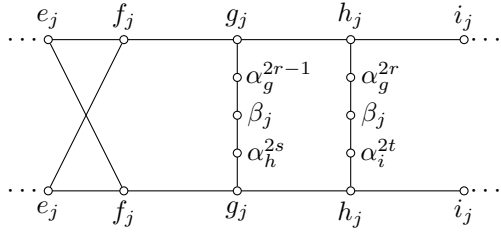
► **Theorem 14.** 2-COLOURFUL PARTITION is NP-complete for coloured planar bipartite graphs of maximum degree 3 and path-width 3.

**Proof.** We reduce from the NOT-ALL-EQUAL POSITIVE 3-SATISFIABILITY problem, which is NP-complete [15]. Let  $(X, \mathcal{C})$  be a positive 3-formula, where  $X = \{x_1, x_2, \dots, x_n\}$  and  $\mathcal{C} = \{C_1, C_2, \dots, C_m\}$ . From  $(X, \mathcal{C})$  we construct a coloured graph  $(G, c)$  as follows. The idea is to build the graph  $G$  as a sequence of gadgets, where two consecutive gadgets are joined by two edges.

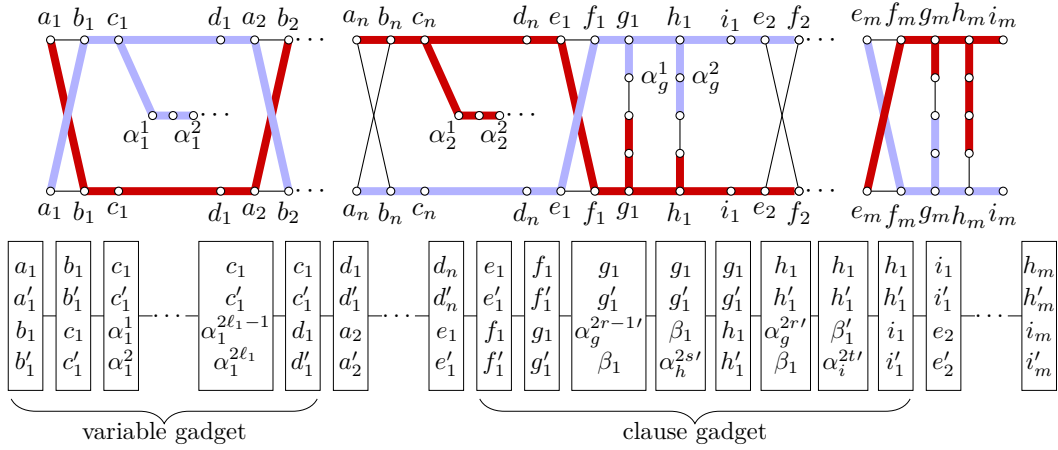
We first construct the variable gadgets. For every  $x_i$ , let  $\ell_i$  be the number of clauses in which  $x_i$  appears. We construct the gadget displayed in Figure 6. In particular, we introduce the new colours  $a_i, b_i, c_i, d_i, \alpha_i^1, \dots, \alpha_i^{2\ell_i}$ .

We now construct the clause gadgets. Let  $C_j = \{x_g, x_h, x_i\}$ . Suppose that  $C_j$  is the  $r$ th clause in which  $x_g$  occurs, the  $s$ th clause in which  $x_h$  occurs, and the  $t$ th clause in which  $x_i$  occurs. We construct the gadget displayed in Figure 7. In particular, we select the already introduced colours  $\alpha_g^{2r-1}, \alpha_g^{2r}, \alpha_h^{2s}$  and  $\alpha_i^{2t}$ , and also introduce six new colours  $e_j, f_j, g_j, h_j, i_j, \beta_j$ .

We obtain  $G$  by first adding for  $l \in \{1, \dots, n-1\}$ , an edge between the top vertices coloured  $d_l$  and  $a_{l+1}$  and an edge between the bottom vertices coloured  $d_l$  and  $a_{l+1}$ . In this way we obtained a sequence of variable gadgets. We now add an edge between the top vertices coloured  $d_n$  and  $e_1$  and the bottom vertices coloured  $d_n$  and  $e_1$ . This is followed by edges between the top vertices coloured  $i_l$  and  $e_{l+1}$  and an edge between the bottom vertices coloured  $i_l$  and  $e_{l+1}$  for  $l \in \{1, \dots, m-1\}$ . This reduction is illustrated in Figure 8.



■ **Figure 7** The clause gadget for Theorem 14.



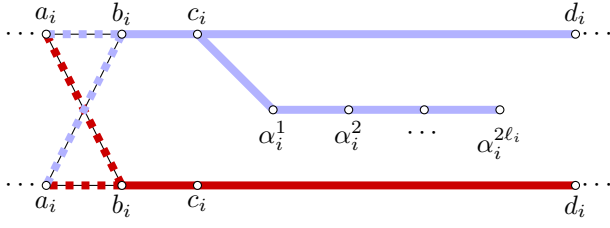
■ **Figure 8** Top: The graph  $G$  obtained by the reduction from NOT-ALL-EQUAL POSITIVE 3-SATISFIABILITY, with a partition into red and blue components, corresponding to a truth assignment. The graph  $G$  can be seen to be planar by letting one of the crossing edges  $\{a_i, b_i\}$  in each variable gadget and one of the crossing edges  $\{e_i, f_i\}$  in each clause gadget go round the outside of the rest of  $G$ . Bottom: A path-decomposition of  $G$  of width 3 (vertices are identified by their colour, we use a ' for the bottom (right) vertex in order to distinguish it from the top (left) vertex of the same colour).

Note that  $G$  is planar and bipartite, and that it admits a path decomposition of width 3, as shown in Figure 8.

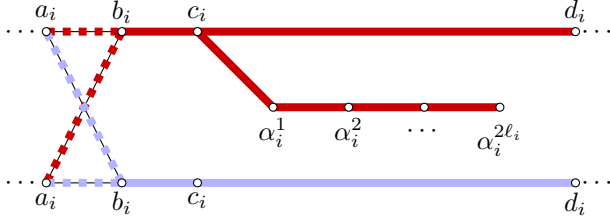
We claim that  $(X, \mathcal{C})$  has a satisfying truth assignment if and only if  $(G, c)$  has a colourful partition of size at most 2.

First suppose that  $(X, \mathcal{C})$  has a satisfying truth assignment  $\tau$ . We will define a colourful partition of  $(G, c)$  with two colourful sets, which we call *red* and *blue*, respectively. Vertices of variable gadgets will be allocated to exactly one of these two sets based on the truth value of the variable. Similarly, vertices of clause gadgets will be allocated to exactly one of these sets based on the truth values of their literals. Below we describe all possibilities. It is readily seen that both the blue set and the red set are colourful. Dotted lines indicate two options depending on the previous gadget in the sequence. We choose the option that ensures the red and blue sets remain connected.

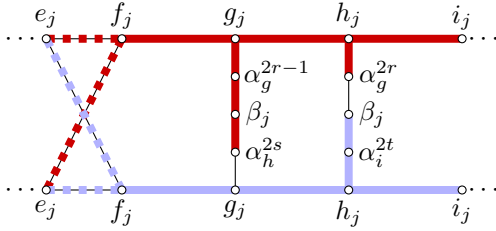
■  $\tau(x_i) = \text{true}$ :



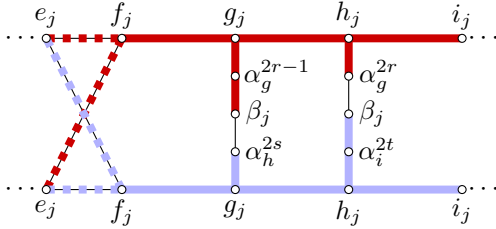
■  $\tau(x_i) = \text{false}$ :



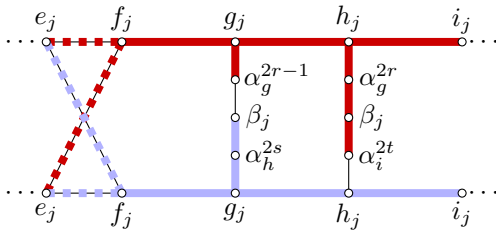
■ The clause  $C_j = \{x_g, x_h, x_i\}$ , with  $\tau(x_g) = \tau(x_h) = \text{true}$  and  $\tau(x_i) = \text{false}$  (exchange the colours blue and red for opposite truth values):



■ The clause  $C_j = \{x_g, x_h, x_i\}$ , with  $\tau(x_g) = \text{true}$  and  $\tau(x_h) = \tau(x_i) = \text{false}$  (exchange the colours blue and red for opposite truth values).



■ The clause  $C_j = \{x_g, x_h, x_i\}$ , with  $\tau(x_g) = \tau(x_i) = \text{true}$  and  $\tau(x_h) = \text{false}$  (exchange the colours blue and red for opposite truth values).



Now suppose that  $(G, c)$  has a colourful partition  $(V_1, V_2)$ . We define a truth assignment  $\tau$  as follows. In the gadget for the variable  $x_i$ , the vertices with colours  $\alpha_i^1, \dots, \alpha_i^{2\ell_i}$  belong to the same component (otherwise, by connectedness, one component consists of only a subset of these vertices, so the other component contains two vertices with the same colour, a contradiction). We define  $\tau(x_i) = \text{true}$  if they belong to  $V_1$  and  $\tau(x_i) = \text{false}$  if they belong to  $V_2$ . In the gadget for the clause  $C_j = \{x_g, x_h, x_i\}$ , there are two vertices with

the same colour  $\beta_j^1$ , so they must belong to different sets  $V_1$  and  $V_2$ . These two vertices are only connected to vertices with colours  $\alpha_g^{2r-1}$ ,  $\alpha_g^{2r}$ ,  $\alpha_h^{2s}$  and  $\alpha_i^{2t}$ . Hence, as  $V_1$  and  $V_2$  are connected, one of these four vertices must be in  $V_1$  and another one must be in  $V_2$ . This means that some variable in  $C_j$  was set to **true** and another one was set to **false**. We conclude that  $\tau$  is a satisfying truth assignment.  $\blacktriangleleft$

The idea behind our next result is that the sets  $V_1$  and  $V_2$  of a colourful partition of size 2 form connected subtrees in a tree decomposition. Branching over all options, we “guess” two vertices  $a$  and  $b$  of one bag to belong to different sets  $V_i$ . By exploiting the treewidth-2 assumption we can translate the instance into an equivalent instance of 2-SATISFIABILITY.

► **Theorem 15.** 2-COLOURFUL PARTITION *is polynomial-time solvable for graphs of treewidth at most 2.*

**Proof.** Let  $(G, c)$  be a coloured graph on  $n$  vertices such that  $G$  has treewidth at most 2. Without loss of generality we may assume that  $G$  is connected. We may assume that  $G$  is not colourful, otherwise we are trivially done. If  $G$  has treewidth 1, then it is a tree and we apply Theorem 7. Hence we may assume that  $G$  has treewidth 2. Let  $(T, \mathcal{X})$  be a tree decomposition of  $G$  of width 2 (so all bags of  $\mathcal{X}$  have size at most 3). We can obtain this tree decomposition in linear time [16]. Let us state and then explain two properties that we may assume hold for  $(T, \mathcal{X})$ . These properties, along with the main definitions necessary for our algorithm, are illustrated in Figure 9.

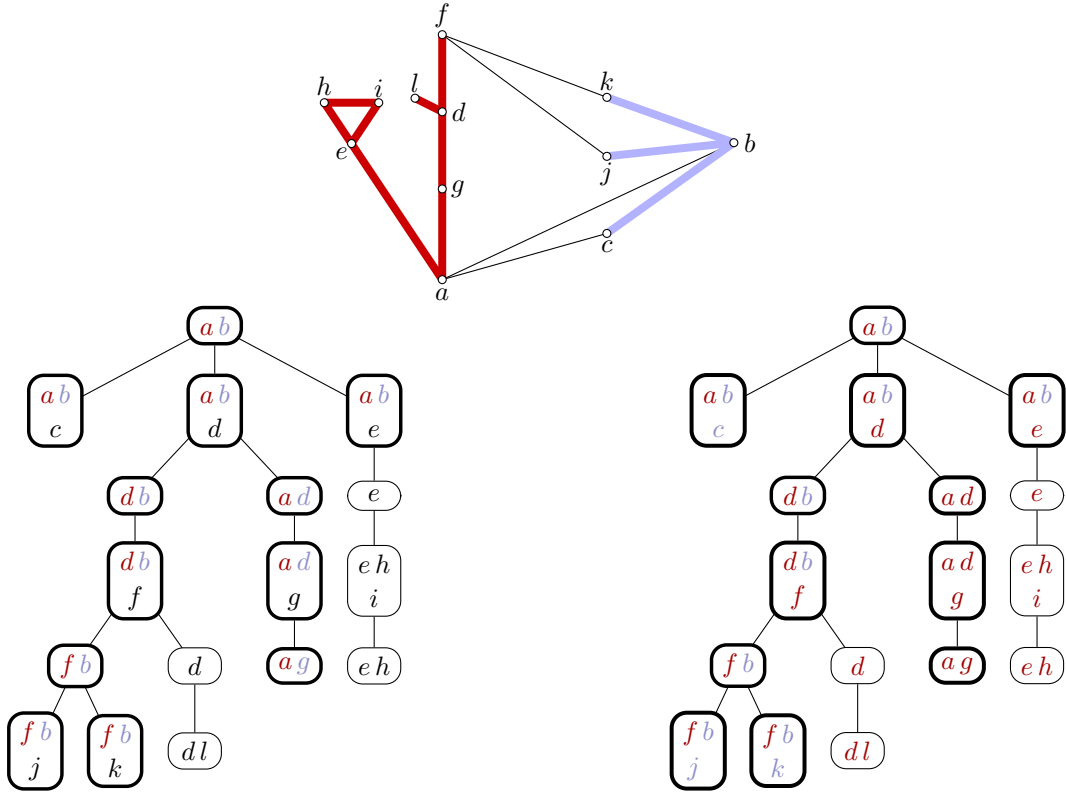
1. For any two adjacent nodes  $i, j$  in  $T$ , one of  $X_i$  and  $X_j$  strictly contains the other.
2. All bags are pairwise distinct.

In fact, what we will show is that if we find that  $(T, \mathcal{X})$  does not have these two properties, then we can make simple changes to obtain a tree decomposition that does. First, if there are two adjacent nodes  $i, j$  in  $T$  such that neither  $X_i$  nor  $X_j$  contains the other, we remove the edge from  $ij$  from  $T$ , create a new node  $k$  that is adjacent to both  $i$  and  $j$  and let  $X_k = X_i \cap X_j$ . Note that  $X_k \neq \emptyset$ , as  $G$  is connected. We now have a tree decomposition that has Property 1 unless  $X_i = X_j$  for a pair of adjacent nodes  $i$  and  $j$ . But now for *every* pair of identical bags  $X_i$  and  $X_j$ , we delete  $j$  and make each of its neighbours adjacent to  $i$  and so obtain a tree decomposition with Properties 1 and 2.

Let  $a$  and  $b$  be a fixed pair of adjacent vertices in  $G$ . Almost all of the remainder of this proof is concerned with describing an algorithm that decides whether or not  $G$  has a colourful partition of size 2 in which  $a$  and  $b$  belong to different parts. Clearly such an algorithm suffices: we can apply it to each of the  $O(n^2)$  pairs of adjacent vertices in  $G$  to determine whether  $G$  has *any* colourful partition of size 2 (since such a partition must separate at least one pair of adjacent vertices).

As  $a$  and  $b$  are adjacent, there exists at least one bag that contains both of them. We may assume that this bag is  $X_0 = \{a, b\}$ ; otherwise we simply add a new node 0 to  $T$  with  $X_0 = \{a, b\}$  and make 0 adjacent to  $i$  such that  $X_i$  is a (larger) bag that contains both  $a$  and  $b$ .

We now orient all edges of  $T$  away from 0 and think of  $T$  as being rooted at 0. We write  $i \rightarrow j$  to denote an edge oriented from  $i$  to  $j$ . If  $i \rightarrow j$  is present in  $T$ , then  $i$  is the *parent* of  $j$  and  $j$  is a *child* of  $i$ . The *head subtree* of  $T$  is the subtree obtained by removing all 1-vertex bags along with all their descendants. For any oriented edge  $i \rightarrow j$  in  $T$ , we have either  $X_i \subsetneq X_j$  or  $X_j \subsetneq X_i$  by Property 1. The tree  $T[i]$  denotes the subtree of  $T$  rooted at  $i$ ; in particular  $T = T[0]$ . The set  $V[i]$  denotes  $\bigcup_{j \in V(T[i])} X_j$ . Finally, as we shall explain, we may assume the following property.



■ **Figure 9** Illustration of the polynomial-time algorithm for treewidth 2. Top: an input graph (the vertex colouring is not represented), with a partition cutting  $(a, b)$ , i.e.  $a$  is in one part,  $b$  is in the other. Bottom-left: a tree-decomposition of the graph satisfying Properties 1, 2 and 3. The bags in the head subtree are in bold. Precuts computed as in Claim 4 are shown using colours on the top line of bags that have precuts. For example, the bag  $\{d, b, f\}$  is precut on  $(d, b)$ , as can be verified directly in the graph (if  $d, b$  and  $f$  are not all in the same part of an  $(a, b)$ -partition, then  $d$  must be in the same part as  $a$ ). Here  $f$  is attached to  $b$ , and  $e$  is attached to  $a$  but not to  $b$ . Bottom-right: the tree decomposition showing a possible output of the 2-SATISFIABILITY formula, where the colours represent the values of each  $x_u$ .

3. For any node  $i$  of  $T$ , the subgraph  $G[V[i]]$  induced by  $V[i]$  is connected.

It is possible that we must again modify the tree decomposition to obtain this property. Suppose that it does not hold for some node  $i$ . That is, the vertices of  $V[i]$  can be divided into two sets  $U$  and  $W$  such that there is no edge from  $U$  to  $W$  in  $G$ . We create two trees  $T_{i,U}$  and  $T_{i,W}$  that are isomorphic to  $T[i]$ : for each vertex  $j$  in  $T[i]$ , we let  $j_U$  and  $j_W$  be the corresponding nodes in  $T_{i,U}$  and  $T_{i,W}$  respectively and let  $X_{j_U} = X_j \cap U$ ,  $X_{j_W} = X_j \cap W$ . Then the tree decomposition is modified by replacing  $T[i]$  by  $T_{i,U}$  and  $T_{i,W}$  and making each of  $i_U$  and  $i_W$  adjacent to the parent of  $i$ . (The vertex  $i$  certainly has a parent since  $i = 0$  would imply that  $G$  is not connected.) If at any point we create a node whose associated bag is empty or identical to that of its parent, we delete it and make its children (if it has any) adjacent to its parent. In this way we obtain a decomposition that now satisfies each of Properties 1, 2 and 3.

We say that a colourful partition  $P = (V_1, V_2)$  of  $G$  is an  $(a, b)$ -partition if  $a \in V_1$  and  $b \in V_2$  and we say that  $P$  cuts a pair  $(u_1, u_2)$  if  $u_1 \in V_1$  and  $u_2 \in V_2$ . We emphasize that the order



is important. A colourful partition  $P$  respects a bag  $X$  if  $X \subseteq V_1$  or  $X \subseteq V_2$ . Note that every colourful partition respects all 1-vertex bags. Let  $X$  be a bag that contains vertices  $u$  and  $v$ . Then  $X$  is *precut* on  $(u, v)$  if every  $(a, b)$ -partition either cuts  $(u, v)$  or respects  $X$ . Note that  $X_0$  is precut on  $(a, b)$  by definition. We make three structural claims.

► **Claim 1.** *If an  $(a, b)$ -partition  $P$  respects a bag  $X_i$ , then it respects every bag  $X_j$  such that  $i \rightarrow j$ .*

We prove Claim 1 as follows. Consider the set  $\mathcal{C}$  of bags that are not respected by  $P$ . Then  $\mathcal{C}$  is the intersection of the set of bags containing at least one vertex from  $V_1$  and the set of bags containing at least one vertex from  $V_2$ . Since both  $V_1$  and  $V_2$  are connected in  $G$ , the set of nodes whose bags contain at least one vertex from  $V_1$  and the set of nodes whose bags contain at least one vertex from  $V_2$  induce subtrees of  $T$ . Hence their intersection,  $\mathcal{C}$ , is a set of nodes that also induce a subtree of  $T$ . Since  $(a, b)$  is not respected by  $P$ , we know  $0 \in \mathcal{C}$ . This means that if  $i$  is not in  $\mathcal{C}$ , then for every other vertex  $j \in T[i]$ ,  $j$  cannot be in  $\mathcal{C}$ . ◊

► **Claim 2.** *Let  $i \rightarrow j$  be an oriented edge of  $T$  with  $X_i = \{u, v\}$  and  $|X_j| = 3$ . If  $X_i$  is precut on  $(u, v)$ , then  $X_j$  is precut on  $(u, v)$ .*

We prove Claim 2 as follows. By Property 1, we know that  $X_i \subsetneq X_j$ , so  $X_j$  contains  $u$  and  $v$ . Suppose  $(u, v)$  is a precut on  $X_i$  and let  $P$  be an  $(a, b)$ -partition. If  $P$  cuts  $(u, v)$  then we are done. Otherwise,  $P$  respects  $X_i = \{u, v\}$ , so by Claim 1,  $P$  also respects  $X_j$ . Thus  $(u, v)$  is a precut on  $X_j$ . ◊

► **Claim 3.** *Let  $i \rightarrow j$  be an oriented edge of  $T$  with  $X_i = \{u, v, w\}$  and  $|X_j| = 2$  such that  $X_i$  is precut on  $(u, v)$ . If  $X_j = \{v, w\}$ , then  $X_j$  is precut on  $(w, v)$ . If  $X_j = \{u, w\}$ , then  $X_j$  is precut on  $(u, w)$ .*

We prove Claim 3 as follows. Let  $P = (V_1, V_2)$  be an  $(a, b)$ -partition of  $G$ . Suppose  $P$  does not respect  $X_j$ . Then if  $Y$  contains  $w$ , we must have that  $w$  is not in the same part of the partition as the other vertex — either  $u$  or  $v$  — of  $X_j$ . By Claim 1,  $P$  does not respect  $X_i$  either, so we know  $u \in V_1$  and  $v \in V_2$ . Thus if  $X_j = \{v, w\}$ , then  $w \in V_1$ , and if  $X_j = \{u, w\}$ , then  $w \in V_2$ . ◊

By Claim 1 and the fact that 1-vertex bags are respected, for every node  $i$  not in the head subtree of  $T$ ,  $X_i$  is respected by every  $(a, b)$ -partition of  $G$ . We now show the following claim.

► **Claim 4.** *For every node  $i$  in the head subtree,  $X_i$  is precut on some pair of its vertices, and these precuts can be computed in linear time.*

We prove that Claim 4 holds by proving a slightly stronger statement: for all  $d$ , for each node  $j$  in the head subtree at distance  $d$  from 0,  $X_j$  is precut on a pair of its vertices, and if  $X_j$  contains three vertices, then it is precut on some pair  $(u, v)$  such that  $X_i = \{u, v\}$  where  $i$  is the parent of  $j$ . We prove this by induction on  $d$ . The base case holds as  $X_0$  is precut on  $(a, b)$ . For the inductive case, suppose that  $j$  is some node at distance  $d > 0$  with parent  $i$ . As  $i$  and  $j$  are in the head subtree, each of  $X_i$  and  $X_j$  has either two or three vertices. Suppose first that  $X_j = \{u, v, w\}$ , and then we can assume that  $X_i$  is, say,  $\{u, v\}$ . (Recall that  $X_i \subsetneq X_j$  or  $X_i \subsetneq X_j$  by Property 1.) By the induction hypothesis, we know that  $X_i$  is precut on  $(u, v)$  or on  $(v, u)$  and so, by Claim 2,  $X_j$  is precut in the same way. Now suppose that  $X_j$  has two vertices and  $X_i = \{u, v, w\}$ . Since  $X_0$  has two vertices, it follows that  $X_i \neq X_0$ , so  $i$  has a parent  $h$ , and we can suppose that  $X_h = \{u, v\}$ . By the induction hypothesis, without loss of generality we may assume  $X_h$  is precut on  $(u, v)$ , so  $X_i$  is precut on  $(u, v)$  by the same argument as above. Then, as  $\mathcal{X}$  has no identical bags

by Property 2, we find that  $X_j = \{u, w\}$  or  $X_j = \{v, w\}$ . It follows from Claim 3 that  $X_j$  is precut.  $\diamond$

We say that two vertices  $u$  and  $v$  of  $G$  are *attached* if they are adjacent or  $\mathcal{X}$  contains the bag  $\{u, v\}$ . If  $u$  and  $v$  are not attached, they are *detached*.

► **Claim 5.** *Let  $X_i = \{u, v, w\}$  be a 3-vertex bag precut on  $(u, v)$ . If  $w$  and  $u$  (or  $w$  and  $v$ ) are detached, then in every  $(a, b)$ -partition of  $G$  the vertices  $w$  and  $v$  (respectively  $w$  and  $u$ ) are in the same colourful set.*

We prove Claim 5 as follows. Let  $P = (V_1, V_2)$  be an  $(a, b)$ -partition. If  $P$  respects  $X_i$ , then  $w$  is in the same colourful set as both  $u$  and  $v$ . Thus we may assume that  $P$  does not respect  $X_i$  and so  $u \in V_1$  and  $v \in V_2$ , since  $X_i$  is precut on  $(u, v)$ . We may assume without loss of generality that  $w$  is detached from  $u$  and so we must show that  $w \in V_2$ .

Let us assume instead that  $w \in V_1$  and derive a contradiction. By the connectivity of  $G[V_1]$ , there exists an induced path  $Q$  on  $\ell$  edges in  $G[V_1]$  that connects  $u$  to  $w$ . As  $u$  and  $w$  are detached, they are not adjacent, so  $\ell$  is at least 2. No internal vertex of  $Q$  is in  $X_i$  since  $u$  and  $w$  are its end-vertices and  $v$  is not in  $V_1$ . Let  $X_{i_1}, \dots, X_{i_\ell}$  be bags of  $\mathcal{X}$  such that  $X_{i_j}$  is a bag that contains the pair of vertices joined by the  $j$ th edge of  $Q$ . We take a walk in  $T$  from  $i_1$  to  $i_\ell$  by stitching together paths from  $i_j$  to  $i_{j+1}$ ,  $1 \leq j \leq \ell - 1$ . As  $X_{i_j}$  and  $X_{i_{j+1}}$  correspond to incident edges of  $Q$ , and both  $X_{i_j}$  and  $X_{i_{j+1}}$  contain the internal vertex of  $Q$  incident with both these edges, every bag of the nodes along the path between them also contains this vertex. Thus every bag of the nodes along our walk from  $i_1$  to  $i_\ell$  contains an internal vertex of  $Q$  and so none of these bags is  $X_i$ . Therefore our walk must be contained within a connected component of  $T \setminus i$ . Let  $j$  be the node of this component adjacent to  $i$  in  $T$ . We note that  $X_i$  contains  $u$  and  $w$ ,  $X_{i_1}$  contains  $u$ ,  $X_{i_\ell}$  contains  $w$ , and the paths from  $i$  to each of  $i_1$  and  $i_\ell$  go through  $j$ . Thus  $X_j$  must contain  $u$  and  $w$ , and so, by Property 1, it follows that  $X_j = \{u, w\}$ . As  $u$  and  $w$  are detached, we have our contradiction.  $\diamond$

We now build  $\phi$ , an instance of 2-SATISFIABILITY, that will help us to find an  $(a, b)$ -partition of  $G$  (if one exists). For each vertex  $u$  of  $G$ , we create a variable  $x_u$  (understood as “ $u \in V_1$ ”). We add clauses (or pairs of clauses) on two variables that are equivalent to the following statements:

$$x_a = \mathbf{true} \tag{1}$$

$$x_b = \mathbf{false} \tag{2}$$

$$x_u \neq x_v \text{ if } u \text{ and } v \text{ have the same colour} \tag{3}$$

For any bag  $X_i$  precut on some pair  $(u, v)$ ,

$$x_u \vee \neg x_v \tag{4}$$

$$\neg x_u \Rightarrow \neg x_w \quad \forall w \in V[i] \setminus \{u\} \tag{5}$$

$$x_v \Rightarrow x_w \quad \forall w \in V[i] \setminus \{v\} \tag{6}$$

For each bag  $X_i = \{u\}$  of size 1,

$$x_w = x_u \quad \forall w \in V[i] \setminus \{u\} \tag{7}$$

For each bag  $\{u, v, w\}$  of size 3 precut on  $(u, v)$

$$x_w = x_u \text{ if } w \text{ is detached from } v \tag{8}$$

$$x_w = x_v \text{ if } w \text{ is detached from } u \tag{9}$$

We claim that  $G$  has an  $(a, b)$ -partition if and only if  $\phi$  is satisfiable.

First suppose that  $G$  has an  $(a, b)$ -partition  $P$ . For each vertex  $u$  in  $G$ , set  $x_u = \mathbf{true}$  if  $u \in V_1$ , and  $x_u = \mathbf{false}$  otherwise. As  $P$  cuts  $(a, b)$ , Statements (1) and (2) are satisfied. As  $P$  is colourful, Statement (3) is satisfied.

Consider a bag  $X_i$  precut on  $(u, v)$ . Two cases are possible: either  $P$  cuts  $(u, v)$ , or  $X$  is respected by  $P$ . In the first case we have  $x_u = \mathbf{true}$  and  $x_v = \mathbf{false}$ , which is enough to satisfy Statements (4), (5) and (6). In the second case,  $x_u = x_v$  (which satisfies Statement (4)), and by Claim 1, for each  $j$  in  $T[i]$ ,  $X_j$  is respected by  $P$ . Thus all vertices in  $V[i]$  are in the same subset  $V_1$  or  $V_2$ , hence they have the same value of  $x_w$ , which satisfies Statements (5) and (6).

For Statement (7), note that each bag  $X_i$  of size 1, and, by Claim 1, every bag  $X_j$  such that  $j$  is a descendent of  $i$  is always respected, so again all values of  $x_w$  for  $w \in V[i]$  are identical. Finally, Statements (8) and (9) follow from Claim 5.

Now suppose that  $\phi$  is satisfiable. We construct two disjoint vertex sets  $V_1$  and  $V_2$  by putting every vertex  $u$  with  $x_u = \mathbf{true}$  in  $V_1$  and every vertex with  $x_u = \mathbf{false}$  in  $V_2$ . We claim that  $(V_1, V_2)$  is an  $(a, b)$ -partition of  $G$ . By Statements (1) and (2),  $a \in V_1$  and  $b \in V_2$ . By Statement (3), two vertices of the same colour cannot belong to the same part, hence it remains to prove that both sets are connected.

In fact, we shall prove by induction on  $d$ , that for the set of vertices belonging to bags of nodes at distance at most  $d$  from 0 in  $T$ , the two subsets found by dividing the set according to membership of  $V_1$  or  $V_2$  each induce a connected subgraph of  $G$ . For the base case, we consider  $X_0 = \{a, b\}$  and the two subsets contain a single vertex so we are done. For the inductive case, we consider a node  $j$  at distance  $d > 0$  from 0. Let  $i$  be the parent of  $j$  in  $T$ . It is enough to show that any vertex in  $X_j \setminus X_i$  is in the same component of  $G[V_1]$  or  $G[V_2]$  as a vertex of  $X_i$  assigned to the same set ( $V_1$  or  $V_2$ ). Let us first assume that  $X_j$  is in the head subtree. If  $|X_j| = 2$ , then  $X_j \subsetneq X_i$  so we may assume that  $X_j = \{u, v, w\}$  and also that  $X_i = \{u, v\}$ , and  $X_i$  and  $X_j$  are both precut on  $(u, v)$  (using Claims 2 and 4). We distinguish three cases.

► **Case 1.**  $x_u = x_v = \mathbf{true}$ .

Considering Statement (6) for bag  $X_i$ , we have  $x_w = \mathbf{true}$ , as well as  $x_{w'} = \mathbf{true}$  for every vertex  $w' \in V[i] \setminus X_i$ . Hence  $V[i] \subseteq V_1$ . By Property 3,  $G[V[X]]$  is connected, so  $w$  is in the same component of  $G[V_1]$  as  $v$ .

► **Case 2.**  $x_u = x_v = \mathbf{false}$ .

Symmetrically to Case 1, we can show that  $w$  is in the same component of  $G[V_2]$  as  $u$ .

► **Case 3.**  $x_u \neq x_v$ .

By Statement (4), in this case we have  $x_u = \mathbf{true}$  and  $x_v = \mathbf{false}$ . Suppose that  $x_w = \mathbf{true}$  (the case for  $x_w = \mathbf{false}$  again follows symmetrically). Since  $x_w \neq x_v$ , by Statement (9),  $w$  must be attached to  $u$ . If there is an edge from  $u$  to  $w$ , then we are done. If there is no such edge, then  $X$  has a bag  $X_k = \{u, w\}$ , and  $k$  must be a child of  $j$ . As  $k$  is in the head subtree,  $X_k$  is precut on  $(u, w)$  or  $(w, u)$  by Claim 4. By Statements (5) and (6), since  $x_u = x_w = \mathbf{true}$ ,  $x_{w'} = \mathbf{true}$  for every  $w' \in V[k] \setminus \{u, w\}$  and thus  $V[k] \subseteq V_1$ . Since  $G[V[k]]$  is connected and contains both  $u$  and  $w$ , we are done.

Now we consider the case that  $X_j$  is not in the head subtree. Thus  $X_j$  is in a subtree that has at its root a bag containing a single vertex  $u$  and, by Statement (7), every vertex in the bags of the subtree are in the same subset  $V_1$  or  $V_2$  as  $u$  and these vertices together induce

a connected subgraph. Thus, as  $u$  is also in the parent of the root of the subtree (since  $G$  is connected, the parent is not empty), we are done.

Thus we have a polynomial-time algorithm to decide 2-COLOURFUL PARTITION on instance  $(G, c)$ . First we compute in polynomial time a tree-decomposition  $(T, \mathcal{X})$  of  $G$  with Properties 1, 2 and 3. Then for every pair of adjacent vertices  $a$  and  $b$  in  $G$ , we check whether there is a colourful  $(a, b)$ -partition of size 2 in polynomial time using the corresponding 2-SATISFIABILITY formula  $\phi$ . ◀

## 4 Parameterized Complexity

The number of colourful components, maximum degree, number of colours, colour-multiplicity, path-width, and treewidth are all natural parameters for COLOURFUL PARTITION. However, as a consequence of our results of the previous section, the only sensible combination of these to consider as a parameter is the treewidth (or path-width) plus the number of colours. This can be seen as follows.

We first observe that if an instance of  $k$ -COLOURFUL PARTITION has colour-multiplicity more than  $k$ , then it is a no-instance. Therefore, if 2-COLOURFUL PARTITION is NP-complete for some class  $\mathcal{C}$  of coloured graphs, then it must be NP-complete for the subclass consisting of all coloured graphs in  $\mathcal{C}$  with colour-multiplicity at most 2. Now let  $X$  be a combination of the above six parameters. Theorem 14, combined with the above observation, implies that COLOURFUL PARTITION is para-NP-complete when parameterized by  $X$  if  $X$  is any subset (including the whole set) of the number of colourful components, maximum degree, colour-multiplicity, path-width and treewidth. Hence, we may assume that  $X$  contains the number of colours. If  $X$  also contains the number of colourful components or the colour-multiplicity, then COLOURFUL PARTITION is trivially FPT, as the size of the input is bounded. Assume that  $X$  contains neither the number of colourful components nor the colour-multiplicity. If  $X$  contains neither the path-width nor the treewidth, then COLOURFUL PARTITION is para-NP-complete due to Theorem 10. Hence, we are indeed left to consider the case where  $X$  contains the number of colours and the path-width or treewidth. For this case we are able to prove the following result.

► **Theorem 16.** *COLOURFUL PARTITION and COLOURFUL COMPONENTS are FPT when parameterized by the number of colours plus the treewidth.*

**Proof.** Let  $(G, c)$  be a coloured graph and let  $C$  be the set of colours used on  $G$ . We may use the algorithm of [2] to obtain a tree decomposition of the input graph  $G$  in FPT time. We then convert it to a nice tree decomposition  $T$  using the aforementioned result of Kloks [11].

Let  $r$  denote the root of  $T$ . For any node  $i \in V(T)$ , let  $T_i$  denote the subtree of  $T$  induced by  $i$  and its descendants and let  $G_i = G[\bigcup_{j \in V(T_i)} X_j]$ . For both problems, we apply a dynamic programming algorithm over  $(T, \mathcal{X})$ .

First, we describe the tables that are constructed for the nodes of  $T$  when solving the COLOURFUL PARTITION problem. Let  $i \in V(T)$ . We define **table** $_i$  as a partial function whose inputs are ordered pairs  $(P, \rho)$  where

- $P$  is a partition of  $X_i$ , and
- $\rho : P \rightarrow \mathcal{P}(C)$  is a function assigning a set of colours to each set in  $P$ .

Let  $Q$  be a partition of the vertices of  $G_i$  into colourful components. We say that  $Q$  induces the partition  $P$  on  $X_i$  and the function  $\rho : P \rightarrow \mathcal{P}(C)$  if both the following conditions hold:

- Two elements of  $X_i$  are in the same set in  $P$  if and only if they are in the same set in  $Q$ .

- For  $D \in P$ , let  $D' \in Q$  be the colourful component that contains the vertices of  $D$ . Then  $\rho(D) = \{c(v) \mid v \in D' \setminus D\}$ .

For a pair  $(P, \rho)$ , the value of  $\mathbf{table}_i(P, \rho)$  will be the minimum possible number of colourful components in a colourful partition  $Q$  of  $G_i$  among all such partitions  $Q$  that induce  $(P, \rho)$ ; if no such partition exists then  $\mathbf{table}_i(P, \rho)$  is void.

Recall that  $X_r = \emptyset$ . Therefore  $\mathbf{table}_r(\emptyset, \emptyset)$  is the minimum number of colourful components for which  $G$  has a colourful partition.

Now we explain how we construct  $\mathbf{table}_i$  for each  $i \in V(T)$ . In what follows we write  $\mathbf{table}_i(P, \rho) \leftarrow t$  to refer to the following procedure: If  $\mathbf{table}_i(P, \rho)$  is undefined, set it to be equal to  $t$ . If  $\mathbf{table}_i(P, \rho) = t' > t$  then set  $\mathbf{table}_i(P, \rho) = t$ , otherwise, leave  $\mathbf{table}_i(P, \rho)$  unchanged.

If  $i$  is a *leaf* node,  $\mathbf{table}_i$  is constructed in a straightforward way because  $X_i = \emptyset$ , so we set  $\mathbf{table}_i(\emptyset, \emptyset) \leftarrow 0$  and all other entries in  $\mathbf{table}_i$  remain void.

If  $i$  is an *introduce* node, let  $j$  be the unique child node of  $i$  and let  $\{v\} = X_i \setminus X_j$ . For every pair  $(P, \rho)$  such that  $\mathbf{table}_j(P, \rho)$  is not void, we do the following. Consider every possible subset of sets  $R \subseteq P$  such that  $v$  has at least one neighbour in every set of  $R$  and  $Y_v = \{v\} \cup \bigcup R$  is colourful (note that  $R$  may be empty). If the sets in  $\{\rho(Y) \mid Y \in R\} \cup \{c(u) \mid u \in Y_v\}$  are not pairwise disjoint, then disregard this choice of  $R$ . Otherwise, let  $P' = P \setminus R \cup \{Y_v\}$  and  $\rho' : P' \rightarrow \mathcal{P}(C)$  such that  $\rho'(Y) = \rho(Y)$  for  $Y \in P \setminus R$  and  $\rho'(Y_v) = \bigcup_{Y \in R} \rho(Y)$ . Apply  $\mathbf{table}_i(P', \rho') \leftarrow \mathbf{table}_j(P, \rho) + 1 - |R|$ .

If  $i$  is a *forget* node, let  $j$  be the unique child node of  $i$  and let  $\{v\} = X_j \setminus X_i$ . For every pair  $(P, \rho)$  such that  $\mathbf{table}_j(P, \rho)$  is not void, let  $P'$  be the partition  $P$  restricted to  $X_i$  (that is, if  $\{v\} \in P$  then delete this set from  $P$  and otherwise, remove  $v$  from the set in  $P$  that contains it) and let  $\rho' : P' \rightarrow \mathcal{P}(C)$  be the function that takes the value  $\rho'(Y) = \rho(Y)$  if  $Y \in P \cap P'$  and  $\rho'(Y) = \rho(Y \cup \{c(v)\})$  otherwise. Apply  $\mathbf{table}_i(P', \rho') \leftarrow \mathbf{table}_j(P, \rho)$ .

If  $i$  is a *join* node, let  $j$  and  $k$  be the two child nodes of  $i$ . For every pair of pairs  $(P, \rho)$ ,  $(P', \rho')$  such that  $\mathbf{table}_j(P, \rho)$  and  $\mathbf{table}_k(P', \rho')$  are not void, we do the following. Recall that  $P$  and  $P'$  are partitions of  $X_i = X_j = X_k$ . We construct a partition  $P''$  and a function  $\rho''$  as follows: Start by setting  $P'' = P$ . If two sets  $P_1, P_2 \in P''$  contain vertices  $v_1, v_2$ , respectively, such that  $v_1$  and  $v_2$  are in the same set of  $P'$ , then replace  $P_1$  and  $P_2$  by  $P_1 \cup P_2$  in  $P''$ . Repeat this process exhaustively and note that the resulting partition  $P''$  is a coarsening of both  $P$  and  $P'$ . Consider an element  $Y \in P''$ . If  $Y$  is not colourful, then we discard this pair of pairs  $(P, \rho)$ ,  $(P', \rho')$ , and consider the next pair. Since  $P''$  is a coarsening of  $P$  and  $P'$ , there must exist sets  $P_1, \dots, P_a \in P$  and  $P'_1, \dots, P'_b \in P'$  such that  $P_1 \cup \dots \cup P_a = P'_1 \cup \dots \cup P'_b = Y$ . Let  $c(Y)$  denote the set of colours used on  $Y$ . If the sets  $c(Y), \rho(P_1), \dots, \rho(P_a), \rho'(P'_1), \dots, \rho'(P'_b)$  are not pairwise disjoint, then we discard this pair of pairs  $(P, \rho)$ ,  $(P', \rho')$ , and consider the next pair. Otherwise, set  $\rho''(Y) = \rho(P_1) \cup \dots \cup \rho(P_a) \cup \rho'(P'_1) \cup \dots \cup \rho'(P'_b)$ . Apply  $\mathbf{table}_i(P'', \rho'') \leftarrow \mathbf{table}_j(P, \rho) + \mathbf{table}_k(P', \rho') - |P| - |P'| + |P''|$ .

It is easy to verify that the above procedure will complete  $\mathbf{table}_i(P, \rho)$  correctly and we can thus obtain the size of an optimal solution to COLOURFUL PARTITION. Note that for  $i \in T$ , the set  $X_i$  contains at most  $\text{tw}(G) + 1$  elements and so the number of partitions  $P$  of  $X_i$  that we need to consider is bounded above by  $(\text{tw}(G) + 1)^{\text{tw}(G) + 1}$ , that is, a function of the treewidth of  $G$ . For each such partition  $P$ , there are at most  $2^{|C||P|}$  functions  $\rho : P \rightarrow \mathcal{P}(C)$  that we need to consider. Thus, for each  $i$ , the number of pairs  $(P, \rho)$  that need to be considered is bounded by a function of  $\text{tw}(G)$  and  $|C|$ . Thus, for every  $i$ ,  $\mathbf{table}_i$  can be completed in FPT-time. Since the number of nodes in  $T$  is  $O(n)$ , it follows that COLOURFUL PARTITION can be solved in FPT-time parameterized by  $\text{tw}(G) + |C|$ . This completes the

proof for the COLOURFUL PARTITION problem.

We now describe the tables that are constructed for the nodes of  $T$  when solving the COLOURFUL COMPONENTS problem. Let  $i \in V(T)$ . We define  $\mathbf{table}'_i$  as a partial function whose inputs are ordered pairs  $(P, \rho)$  where

- $P$  is a partition of  $X_i$ , and
- $\rho : P \rightarrow \mathcal{P}(C)$  is a function assigning a set of colours to each set in  $P$ .

Note that the COLOURFUL COMPONENTS problem is equivalent to finding the minimum number of edges that must be deleted from the input graph  $G$  such that the vertex set of the resulting graph can be partitioned into sets  $V_1, \dots, V_k$  (for arbitrary  $k$ ) such that each set  $V_i$  is colourful. Note that in this case the number of sets  $V_1, \dots, V_k$  does not matter for our purposes and we do not insist that  $G[V_i]$  is connected. However, we do insist that for distinct  $i, j \in \{1, \dots, k\}$ , all edges with one end-vertex in  $V_i$  and the other in  $V_j$  must be in the set of deleted edges. Furthermore, note that in any optimal solution no edge will be deleted with both end-vertices in the same set  $V_i$ . Thus the COLOURFUL COMPONENTS problem is equivalent to finding a partition of  $V(G)$  into colourful sets  $V_1, \dots, V_k$  that minimizes the number of edges whose end-vertices are in different sets of the partition. We call a partition of  $V(G)$  into colourful sets  $V_1, \dots, V_k$  a *colourful set partition* (note that this is different from the definition of colourful partition, for which we insist that each partition set induces a connected graph in  $G$ ).

Let  $Q$  be a colourful set partition of the vertices of  $G_i$ . Similarly to the COLOURFUL PARTITION case, we say that  $Q$  *induces* the partition  $P$  on  $X_i$  and the function  $\rho : P \rightarrow \mathcal{P}(C)$  if both the following conditions hold:

- Two elements of  $X_i$  are in the same set in  $P$  if and only if they are in the same set in  $Q$ .
- For  $D \in P$ , let  $D' \in Q$  be the set that contains the vertices of  $D$ . Then  $\rho(D) = \{c(v) \mid v \in D' \setminus D\}$ .

For a pair  $(P, \rho)$ , the value of  $\mathbf{table}'_i(P, \rho)$  will be the minimum possible number of edges that need to be deleted in a colourful set partition  $Q$  of  $G_i$  among all such partitions  $Q$  that induce  $(P, \rho)$ ; if no such partition exists then  $\mathbf{table}'_i(P, \rho)$  is void.

Recall that  $X_r = \emptyset$ . Therefore  $\mathbf{table}'_r(\emptyset, \emptyset)$  is the minimum number edges that need to be deleted from  $G$  to obtain a colourful set partition.

Now we explain how we construct  $\mathbf{table}'_i$  for each  $i \in V(T)$ . Similarly to the COLOURFUL PARTITION case, in what follows we write  $\mathbf{table}'_i(P, \rho) \leftarrow t$  to refer to the following procedure: If  $\mathbf{table}'_i(P, \rho)$  is undefined, set it to be equal to  $t$ . If  $\mathbf{table}'_i(P, \rho) = t' > t$  then set  $\mathbf{table}'_i(P, \rho) = t$ , otherwise, leave  $\mathbf{table}'_i(P, \rho)$  unchanged.

If  $i$  is a *leaf* node,  $\mathbf{table}'_i$  is constructed in a straightforward way because  $X_i = \emptyset$ , so we set  $\mathbf{table}'_i(\emptyset, \emptyset) \leftarrow 0$  and all other entries in  $\mathbf{table}'_i$  remain void.

If  $i$  is an *introduce* node, let  $j$  be the unique child node of  $i$  and let  $\{v\} = X_i \setminus X_j$ . For every pair  $(P, \rho)$  such that  $\mathbf{table}'_j(P, \rho)$  is not void, we do the following. We choose each  $R \in P$  in turn and let  $n_R$  be the number of neighbours  $v$  has in  $X_i \setminus R$ . If  $c(v) \in \rho(R)$  or there is a vertex in  $R$  with the same colour as  $v$ , then we discard this choice of  $R$  and move on to the next one. Otherwise, let  $P' = P \setminus \{R\} \cup \{R \cup \{v\}\}$  and  $\rho' : P' \rightarrow \mathcal{P}(C)$  such that  $\rho'(Y) = \rho(Y)$  for  $Y \in P \setminus R$  and  $\rho'(R \cup \{v\}) = \rho(R)$ . Apply  $\mathbf{table}'_i(P', \rho') \leftarrow \mathbf{table}'_j(P, \rho) + n_R$ , then move onto the next choice of  $R$ . Finally, let  $n_\emptyset$  be the number of neighbours that  $v$  has in  $X_i \setminus \{v\}$ , let  $P' = P \cup \{\{v\}\}$  and let  $\rho' : P' \rightarrow \mathcal{P}(C)$  such that  $\rho'(Y) = \rho(Y)$  for  $Y \in P$  and  $\rho'(\{v\}) = \emptyset$ . Apply  $\mathbf{table}'_i(P', \rho') \leftarrow \mathbf{table}'_j(P, \rho) + n_\emptyset$ .

If  $i$  is a *forget* node, let  $j$  be the unique child node of  $i$  and let  $\{v\} = X_j \setminus X_i$ . For every pair  $(P, \rho)$  such that  $\mathbf{table}'_j(P, \rho)$  is not void, let  $P'$  be the partition  $P$  restricted to  $X_i$  (that

is, if  $\{v\} \in P$  then delete this set from  $P$  and otherwise, remove  $v$  from the set in  $P$  that contains it) and let  $\rho' : P' \rightarrow \mathcal{P}(C)$  be the function that takes the value  $\rho'(Y) = \rho(Y)$  if  $Y \in P \cap P'$  and  $\rho'(Y) = \rho(Y \cup \{c(v)\})$  otherwise. Apply  $\mathbf{table}'_i(P', \rho') \leftarrow \mathbf{table}'_j(P, \rho)$ .

If  $i$  is a *join* node, let  $j$  and  $k$  be the two child nodes of  $i$ . For every pair of pairs  $(P, \rho)$ ,  $(P, \rho')$  such that  $\mathbf{table}'_j(P, \rho)$  and  $\mathbf{table}'_k(P, \rho')$  are not void, we do the following (note that the partition  $P$  in each pair is the same). If  $\rho(R) \cap \rho'(R) = \emptyset$  for all  $R \in P$  then let  $e_P$  be the number of edges in  $G[X_i]$  whose end-vertices are in distinct sets of  $P$  and apply  $\mathbf{table}'_i(P, \rho) = \mathbf{table}'_j(P, \rho) + \mathbf{table}'_k(P, \rho') - e_P$ .

Similarly to the case for COLOURFUL PARTITION, it is easy to verify that the above procedure will complete  $\mathbf{table}'_i(P, \rho)$  correctly and we can thus obtain the size of an optimal solution to COLOURFUL COMPONENTS. Furthermore, it is easy to verify that the procedure also runs in FPT-time. This completes the proof.  $\blacktriangleleft$

We now prove two FPT results for two different parameters. Our proof for the next result uses similar arguments to the proof sketch of Theorem 8 given in [13]. However, the details of both proofs are different, as optimal solutions for COLOURFUL PARTITION do not necessarily translate into optimal solutions for COLOURFUL COMPONENTS. This holds even if the instance has a vertex cover of size 2, as we showed in Example 1 (the vertices  $w, w'$  form a vertex cover in  $G$ ).

► **Theorem 17.** COLOURFUL PARTITION is FPT when parameterized by vertex cover number.

**Proof.** In fact we will show the result for the optimization version of COLOURFUL PARTITION. Let  $(G, c)$  be a coloured graph. We will prove that we can find the size of a *minimum* colourful partition (one with smallest size) in FPT time. By a simple greedy argument, we can find a vertex cover  $S$  of  $G$  that contains at most  $2 \text{vc}(G)$  vertices. It is therefore sufficient to show that COLOURFUL PARTITION is FPT when parameterized by  $|S|$ . If two vertices of  $G$  have the same colour then they will always be in different colourful components of  $G$ . Thus if two vertices with the same colour are adjacent, we can delete the edge that joins them, that is, we may assume that  $c$  is a proper colouring of  $G$  (note that deleting edges from  $G$  maintains the property that  $S$  is a vertex cover). Since  $S$  is a vertex cover,  $T = V(G) \setminus S$  is an independent set. Let  $C$  be the set of colours used on  $G$ . We let  $s = |S|$  and for a set  $S' \subseteq S$ , we let  $T_i(S')$  be the set of vertices with colour  $i$  whose neighbourhood in  $S$  is  $S'$ , that is, for all  $u \in T_i(S')$ , we have that  $N_S(u) = S'$  and  $c(u) = i$ .

► **Rule 1.** If there is a colour  $i \in C$  and a set  $S' \subseteq S$  such that  $|T_i(S')| \geq s + 1$ , then delete  $|T_i(S')| - s$  (arbitrary) vertices of  $T_i(S')$  from  $G$ .

We claim that we can safely apply Rule 1. In order to see this, consider any colourful partition  $(V_1, \dots, V_k)$  of  $(G, c)$ . The number of colourful components in the solution containing at least one vertex of  $S$  is at most  $s$ . Hence at most  $s$  vertices of  $T_i(S')$  can occur in these components. All other vertices of  $T_i(S')$  will be in 1-vertex components of the solution. Since the vertices of  $T_i(S')$  have the same neighbourhood, we may choose the latter set of vertices arbitrarily. Then, given a minimum colourful partition for the resulting coloured graph  $(G', c')$ , we obtain a minimum colourful partition for  $(G, c)$  by restoring the deleted vertices into 1-vertex components. This proves the claim.  $\diamond$

We apply Rule 1 exhaustively. For convenience we again denote the resulting instance by  $(G, c)$  and let  $T = V(G) \setminus S$ . Note that now  $|T_i(S')| \leq s$  for every  $i \in C$  and every  $S' \subseteq S$ . Consequently, the number of vertices of  $T$  with colour  $i$  is at most  $s2^s$ . Note that this means that  $|T| \leq |C|s2^s$  and thus the total number of vertices in  $G$  is at most  $s + |C|s2^s$ , which

gives us an FPT-algorithm in  $s + |C|$ . Hence in order to prove our result it remains to bound the size of  $C$  by a function of  $s$ .

We let  $C_T$  denote the set of colours that appear on vertices of  $T$  but not on vertices of  $S$ . For two colours  $i, j \in C_T$ , if  $T_i(S') = T_j(S')$  holds for all  $S' \subseteq S$ , then we say that  $i$  and  $j$  are *clones* and note that these colours are interchangeable. For a colourful partition  $P = (V_1, \dots, V_k)$  for  $G$ , we let  $P_S = (V_1 \cap S, \dots, V_k \cap S)$  be the partition of  $S$  induced by  $P$  (note that in this case we allow some of the blocks to be empty).

We consider each partition  $Q$  of  $S$ . If a block of  $Q$  is not colourful, then we discard  $Q$ . Otherwise we determine a minimum colourful partition  $P$  for  $G$  with  $P_S = Q$ ; note that such a partition  $P$  may not exist, as it may not be possible to make the blocks of  $Q$  connected (by using vertices  $T$  in addition to edges both of whose endpoints lie in  $S$ ). Finally, we will choose the colourful partition for  $G$  that has minimum size overall.

Let  $Q$  be a partition of  $S$  in which each block is colourful.

► **Rule 2.** *If there are  $s$  distinct colours  $i_1, \dots, i_s$  in  $C_T$  that are pairwise clones, then delete all vertices with colour  $i_s$  from  $G$ .*

We claim that can safely apply Rule 2 if we only consider colourful partitions  $P$  that induce  $Q$  on  $S$ . We say that a colour  $i \in C_T$  is *redundant* for a colourful partition  $P$  if after deleting all vertices of  $T$  with colour  $i$ , the resulting colourful partition  $P'$  induces the same partition on  $S$  as  $P$  does, that is  $P'_S = P_S$ . Suppose we have  $s$  distinct colours  $i_1, \dots, i_s \in C_T$  that are pairwise clones. Any induced partition  $Q$  on  $S$  requires using at most  $s - 1$  vertices of  $T$  to connect the vertices in the same block of  $Q$ . Hence, there must be at least one colour that is redundant for  $P$ . As the colours  $i_1, \dots, i_s$  are indistinguishable in  $(G, c)$  it does not matter which colour we choose to delete, so we may assume that  $i_s$  is redundant.

Let  $(G', c')$  be the resulting coloured graph after removing all vertices with colour  $i_s$  from  $G$ . Let  $P'$  be a minimum colourful partition for  $(G', c')$  with  $P'_S = Q$ . Then we compute a minimum colourful partition  $P$  of  $G$  with  $P_S = Q$  as follows. We construct an auxiliary bipartite graph  $F$  with partition classes  $X$  and  $Y$ . The vertices of  $X$  represent the vertices of  $T$  with colour  $i_s$ , and the vertices of  $Y$  represent colourful components of  $P$  that contain a vertex of  $S$ . We add an edge between two vertices  $x \in X$  and  $y \in Y$  if and only if the vertex of  $T$  corresponding to  $x$  is adjacent to at least one vertex of the component of  $P$  corresponding to  $y$ . We then compute a maximum matching  $M$  in  $F$ , which prescribes how the vertices with colour  $i_s$  must be added to  $P'$  to obtain  $P$ . Note that  $P_S = P'_S = Q$ . As  $M$  is a matching, at most one vertex with colour  $i_s$  is added to the components of  $P'$  that contain a vertex of  $S$ . Hence,  $P$  is colourful. As  $M$  is maximum, the number of components that consist of isolated vertices coloured  $i_s$  is minimized. Hence,  $P$  is a minimum colourful partition of  $(G, c)$  with  $P_S = Q$ . ◊

We now apply Rule 2 exhaustively, and again call the resulting graph  $G$  and define  $S$  and  $T$  as before.

► **Claim 1.**  $|C_T| \leq (s - 1)(s + 1)^{2^s}$ .

We prove Claim 1 as follows. First note that  $0 \leq |T_i(S')| \leq s$  for every  $i \in C$  and every  $S' \subseteq S$  by Rule 1. Thus for every  $S' \subseteq S$ , there are  $s + 1$  possible values of  $|T_i(S')|$ . Since  $S$  has  $2^s$  subsets  $S'$ , for any colour  $i$  there are at most  $(s + 1)^{2^s}$  possible mappings  $|T_i(S')| : \mathcal{P}(S) \rightarrow \{0, \dots, s\}$ . If two colours  $i, j$  have the same mapping, they are clones. By Rule 2, at most  $s - 1$  colours can be pairwise clones. Therefore  $|C_T| \leq (s - 1)(s + 1)^{2^s}$ . ◊

We continue as follows. The number of different colours used on vertices of  $S$  is at most  $s$ . Hence  $C \setminus C_T$  has size at most  $s$ . Recall that for every colour  $i \in C$ , the number of vertices



of  $T$  with colour  $i$  is at most  $s2^s$ . We combine these two facts with Claim 1. Then

$$\begin{aligned}
|V| &= |S| + |T| \\
&\leq s + |C|s2^s \\
&= s + |C \setminus C_T|s2^s + |C_T|s2^s \\
&\leq s + s^22^s + (s-1)(s+1)^{2^s}s2^s,
\end{aligned}$$

which means that by brute force we can compute a minimum colourful partition  $P$  for  $G$  subject to the restriction that  $P_S = Q$  in  $f(s)$  time for some function  $f$  that only depends on  $s$ .

The correctness of our FPT-algorithm follows from the above description. It remains to analyze the running time. Applying Rule 1 exhaustively takes  $O(2^s|C|) = O(2^sn)$  time, as the number of different subsets  $S' \subseteq S$  is  $2^s$ . We then branch into at most  $s^s$  directions by considering every partition of  $S$ . Applying Rule 2 exhaustively takes  $O(2^sn^2)$  time, as for each colour  $i \in C_T$  we first calculate the values of  $|T_i(S')|$  for every  $S' \subseteq S$ , which can be done in  $O(2^sn)$  time. Doing this for every colour takes a total of  $O(2^sn^2)$  time and partitioning the colours into sets that are clones can be done in  $O(2^{2s}n^2)$  time, and deleting colours can be done in  $O(sn)$  time. As every auxiliary graph  $F$  has at most  $n$  vertices, we can compute a maximum matching in  $F$  in  $O(n^{\frac{5}{2}})$  time by using the Hopcroft-Karp algorithm [9]. Finally, translating a minimum solution into a minimum solution for the graph in which we restored the vertices we removed due to exhaustive application of Rules 1 and 2 takes  $O(n)$  time. This means that the total running time is  $O(2^sn) + s^s(O(2^{2s}n^2) + O(s) + O(n^{\frac{5}{2}})) + f(s) + O(n) = f'(s)O(n^{\frac{5}{2}})$  for some function  $f'$  that only depends on  $s$ , as desired. ◀

For our final result we need to introduce a problem used by Robertson and Seymour [14] in their graph minor project to prove their algorithmic result on recognizing graphs that contain some fixed graph  $H$  as a minor. They proved that this problem is cubic-time solvable as long as  $Z_1 \cup \dots \cup Z_r$  has constant size.

**DISJOINT CONNECTED SUBGRAPHS**

*Instance:* a graph  $G$  and  $r$  pairwise disjoint subsets  $Z_1, \dots, Z_r$  of  $V(G)$  for some  $r \geq 1$ .  
*Question:* Is it possible to partition  $V(G) \setminus (Z_1 \cup \dots \cup Z_r)$  into sets  $S_1, \dots, S_r$  such that every  $S_i \cup Z_i$  induces a connected subgraph of  $G$ ?

We are now ready to prove our final result.

► **Theorem 18.** *When parameterized by the number of non-uniquely coloured vertices, COLOURFUL COMPONENTS is para-NP-complete, but COLOURFUL PARTITION is FPT.*

**Proof.** The MULTITERMINAL CUT problem is to test for a graph  $G$ , integer  $p$  and terminal set  $S$ , if there is a set  $E'$  with  $|E'| \leq p$  such that every terminal in  $S$  is in a different component of  $G - E'$ . This problem is NP-complete even if  $|S| = 3$  [6]. To prove the first part, give each of the three vertices in  $S$  colour 1 and the vertices in  $G - S$  colours  $2, \dots, |V| - 2$ .

To prove the second part, let  $(G, c)$  be a coloured graph and  $k$  be an integer be an instance of COLOURFUL PARTITION. We may assume without loss of generality that  $G$  is connected. Let  $Q$  with  $|Q| = q$  be the set of non-uniquely coloured vertices. If  $k \geq q$ , then we place each of the  $q$  vertices of  $Q$  in a separate component and assign the uniquely coloured vertices to components in an arbitrary way subject to maintaining connectivity of the  $q$  components. This yields a colourful partition of  $(G, c)$  of size at most  $k$ . Now assume that  $k \leq q - 1$ . We consider every possible partition of  $Q$  into  $k$  sets  $Z_1, \dots, Z_k$ , where some

of the sets  $Z_i$  may be empty. It remains to solve DISJOINT CONNECTED SUBGRAPHS on the input  $(G, Z_1, \dots, Z_k)$ . Note that  $|Z_1| + \dots + |Z_k|$  has size  $q$ . Hence, by the aforementioned result of Robertson and Seymour [14], solving DISJOINT CONNECTED SUBGRAPHS takes cubic time. As there are  $O(q^q)$  partitions to consider, the result follows. ◀

As an immediate consequence of Theorem 18 we obtain the following result.

► **Corollary 19.** *There exists a family of instances on which COLOURFUL COMPONENTS and COLOURFUL PARTITION have different parameterized complexities.*

## 5 Conclusions

We conclude our paper with some open problems. We showed that COLOURFUL PARTITION and COLOURFUL COMPONENTS are NP-complete for coloured trees of maximum degree at most 6 (and colour-multiplicity 2). What is their complexity for coloured trees of maximum degree  $d$  for  $3 \leq d \leq 5$ ? COLOURFUL COMPONENTS is known to be NP-complete for 3-coloured graphs of maximum degree 6 (Theorem 3); we also ask if one can prove a result analogous to Theorem 10: what is its complexity for coloured graphs of maximum degree 3?

We also proved that 2-COLOURFUL PARTITION is NP-complete for coloured (planar bipartite) graphs of path-width 3 (and maximum degree 3), but polynomial-time solvable for coloured graphs of treewidth 2. We believe that the latter result can be extended to  $k$ -COLOURFUL PARTITION ( $k \geq 3$ ), but leave this for future research. A more interesting question is whether the problem is FPT on graphs of treewidth 2 when parameterized by  $k$ .

---

## References

- 1 Anna Adamaszek and Alexandru Popa. Algorithmic and hardness results for the colorful components problems. *Algorithmica*, 73(2):371–388, 2015.
- 2 Hans L. Bodlaender. A linear-time algorithm for finding tree-decompositions of small treewidth. *SIAM Journal on Computing*, 25(6):1305–1317, 1996.
- 3 Sharon Bruckner, Falk Hüffner, Christian Komusiewicz, Rolf Niedermeier, Sven Thiel, and Johannes Uhlmann. Partitioning into colorful components by minimum edge deletions. *Proc. CPM 2012, LNCS*, 7354:56–69, 2012.
- 4 Laurent Bulteau, Guillaume Fertin, and Irena Rusu. Maximal strip recovery problem with gaps: Hardness and approximation algorithms. *Journal of Discrete Algorithms*, 19:1–22, 2013.
- 5 Gruia Călinescu, Cristina G. Fernandes, and Bruce Reed. Multicuts in unweighted graphs and digraphs with bounded degree and bounded tree-width. *Journal of Algorithms*, 48(2):333–359, 2003.
- 6 Elias Dahlhaus, David S. Johnson, Christos H. Papadimitriou, Paul D. Seymour, and Mihalis Yannakakis. The complexity of multiterminal cuts. *SIAM Journal on Computing*, 23(4):864–894, 1994.
- 7 Riccardo Dondi and Florian Sikora. Parameterized complexity and approximation issues for the colorful components problems. *Theoretical Computer Science*, 739:1–12, 2018.
- 8 George He, Jiping Liu, and Cheng Zhao. Approximation algorithms for some graph partitioning problems. *Journal of Graph Algorithms and Applications*, 4(2):1–11, 2000.
- 9 John E. Hopcroft and Richard M. Karp. An  $n^{5/2}$  algorithm for maximum matchings in bipartite graphs. *SIAM Journal on Computing*, 2(4):225–231, 1973.
- 10 Richard M. Karp. Reducibility among combinatorial problems. *Complexity of Computer Computations*, pages 85–103, 1972.

- 11 Ton Kloks. *Treewidth: Computations and Approximations*, volume 842 of *LNCS*. Springer, 1994.
- 12 Melven R. Krom. The decision problem for a class of first-order formulas in which all disjunctions are binary. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 13(1–2):15–20, 1967.
- 13 Neeldhara Misra. On the parameterized complexity of colorful components and related problems. *Proc. IWOCOA 2018, LNCS*, 10979:237–249, 2018.
- 14 Neil Robertson and Paul D. Seymour. Graph minors. XIII. The disjoint paths problem. *Journal of Combinatorial Theory, Series B*, 63(1):65–110, 1995.
- 15 Thomas J. Schaefer. The complexity of satisfiability problems. *Proc. STOC 1978*, pages 216–226, 1978.
- 16 Joseph A. Wald and Charles J. Colbourn. Steiner trees, partial 2-trees, and minimum IFI networks. *Networks*, 13(2):159–167, 1983.
- 17 Chunfang Zheng, Krister M. Swenson, Eric Lyons, and David Sankoff. OMG! Orthologs in multiple genomes – competing graph-theoretical formulations. *Proc WABI 2011, LNBI*, 6833:364–375, 2011.