

# Extremal properties of flood-filling games

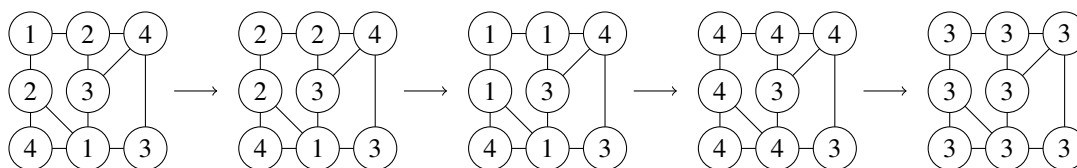
Kitty Meeks<sup>1\*</sup>Dominik K. Vu<sup>2</sup><sup>1</sup> School of Computing Science, University of Glasgow<sup>2</sup> University of Memphisreceived 29<sup>th</sup> Mar. 2018, revised 8<sup>th</sup> Mar. 2019, accepted 3<sup>rd</sup> July 2019.

The problem of determining the number of “flooding operations” required to make a given coloured graph monochromatic in the one-player combinatorial game Flood-It has been studied extensively from an algorithmic point of view, but basic questions about the maximum number of moves that might be required in the worst case remain unanswered. We begin a systematic investigation of such questions, with the goal of determining, for a given graph, the maximum number of moves that may be required, taken over all possible colourings. We give several upper and lower bounds on this quantity for arbitrary graphs and show that all of the bounds are tight for trees; we also investigate how much the upper bounds can be improved if we restrict our attention to graphs with higher edge-density.

**Keywords:** combinatorial games, flood-filling games, Free-Flood-It

## 1 Introduction

Flood-It is a one-player combinatorial game, played on a coloured graph. The goal is to make the entire graph monochromatic (“flood” the graph) with as few moves as possible, where a move involves picking a vertex  $v$  and a colour  $d$ , and giving all vertices in the same monochromatic component as  $v$  colour  $d$ , as illustrated in Figure 1. Implementations of the game played on regular grids are widely available online [5, 4] and as popular smartphone apps. More generally, when played on a planar graph, the game can be regarded as modelling repeated use of the flood-fill tool in Microsoft Paint.



**Fig. 1:** A sequence of moves to flood a graph; the colour of the monochromatic component containing the top left vertex is changed at each move.

Questions arising from this game (and a two-player variant known as the Honey-Bee Game) have received considerable attention from a complexity-theoretic perspective in recent years [9, 12, 15, 18, 19,

\*Supported by a Royal Society of Edinburgh Personal Research Fellowship, funded by the Scottish Government.

21, 23, 24, 26, 30, 32, 33, 34, 37], with such work focussing on questions of the form, “Given a graph  $G$  from a specified class, and a colouring  $\omega$  of the vertices of  $G$ , what is the computational complexity of determining the minimum number of moves required to flood  $G$ ?” The problem is known to be NP-hard in many situations, provided that at least three colours are present in the initial colouring, including in the case that  $G$  is an  $n \times n$  grid (as in the original version of the game) [12] and the case in which  $G$  is a tree [21, 29] (the parameterised complexity of the problem restricted to trees has also been studied [18]). On the other hand, there are polynomial-time algorithms to determine the minimum number of moves required if  $G$  is a path or a cycle, or more generally for any graph if the initial colouring uses only two colours. A more complete description of the complexity landscape for these problems can be found in a recent survey [20].

Two different versions of the game have been considered in the literature, known as the “fixed” and “free” versions. In the fixed version of the game (as in most implementations), players must always change the colour of the monochromatic component containing a single distinguished *pivot vertex*, so the only choice is what colour to assign to this component; in the free version players can choose freely at each move the component whose colour is changed, in addition to the new colour.

In this paper, we initiate a systematic investigation of a different type of question about the game, raised by Meeks and Scott [33] but as yet unanswered in the literature: given a graph  $G$ , what is the maximum number of moves we may need to flood  $G$ , taken over all possible colourings of  $G$  with  $c$  colours? In addition to providing a deeper insight into the behaviour of the game, questions of this form are motivated by recent algorithmic work involving Integer Programming formulations for the optimisation problem, whose running-times might be reduced by better bounds on the worst-case number of moves required.

In Section 2 we obtain a number of straightforward upper and lower bounds on the maximum number of moves that might be required in the worst case in both the fixed and free versions of the game; perhaps surprisingly, we are also able to demonstrate that all of these simple bounds are tight for suitable families of trees. It follows from previous work [34] relating the number of moves required to flood a graph to the number of moves required to flood its spanning trees that any bound on the worst-case number of moves required is tight if and only if it is tight for some family of trees, but we might be able to obtain much better upper bounds on the number of moves required if we know our graph is far from being a tree: intuitively, we expect that a single colouring cannot simultaneously be the worst possible for all spanning trees if the graph has many spanning trees. In Section 3 we investigate this issue, and in the process we determine the worst-case number of moves required to flood a graph that is a blow-up of a long path.

In the remainder of this section, we introduce some key notation and definitions, and mention some results from the existing literature that are of particular relevance to addressing extremal problems.

### 1.1 Notation and definitions

For any graph  $G = (V, E)$ , we denote by  $|G|$  the number of vertices in  $G$  (so  $|G| = |V|$ ). Throughout this paper, we consider only connected graphs: if  $G$  is disconnected then it is impossible to flood  $G$  in the fixed version of the game, and in the free version an optimal strategy for the entire graph is obtained by playing optimal strategy in each connected component. We write  $\mathcal{T}(G)$  for the set of spanning trees of  $G$ .

For  $u, v \in V(G)$ , we let  $\mathcal{P}(u, v)$  be the set of  $u$ - $v$  paths in  $G$ , and the *distance*  $d(u, v)$  from  $u$  to  $v$  is defined to be  $\min_{P \in \mathcal{P}(u, v)} |P| - 1$ . The *eccentricity* of a vertex  $v \in V(G)$  is  $\max_{u \in V(G)} d(u, v)$ , and the *radius* of  $G$  is the minimum eccentricity of any vertex in  $G$ , that is  $\min_{u \in V(G)} \max_{v \in V(G)} d(u, v)$ .

Let  $A \subset V(G)$ . We write  $G[A]$  for the subgraph of  $G$  induced by  $A$ , and  $N(A)$  for the set of vertices in  $V(G) \setminus A$  with at least one neighbour in  $A$ .

A graph  $G = (V_G, E_G)$  is said to be a *blow-up* of a graph  $H = (V_H, E_H)$  if  $V_G$  can be partitioned into sets  $\{V_u : u \in V_H\}$  such that  $v_1 v_2 \in E_G$  if and only if  $v_1 \in V_u$  and  $v_2 \in V_w$  with  $uw \in E_H$ .

Suppose the game is played on a graph  $G$ , equipped with an initial colouring  $\omega : V \rightarrow C$  (not necessarily a proper colouring); we call  $C$  the *colour-set*. We say that  $u$  and  $v$  belong to the same *monochromatic component* of  $G$  with respect to  $\omega$  if there is a path  $u = x_1, x_2, \dots, x_\ell = v$  in  $G$  such that  $\omega(x_1) = \omega(x_2) = \dots = \omega(x_\ell)$ . In the fixed version, a move consists of a single colour  $d$ , and involves assigning colour  $d$  to all vertices in the same monochromatic component as the pivot vertex  $u$ ; in the free version, a move  $m = (v, d)$  involves assigning colour  $d$  to all vertices in the same monochromatic component as  $v$ . Given a colouring  $\omega$  with colour-set  $C$  and any  $d \in C$ , we denote by  $N_d(G, \omega)$  the number of vertices  $v \in V$  such that  $\omega(v) = d$ .

For any graph  $G$  with colouring  $\omega$ , we can obtain a new graph and corresponding colouring by *contracting monochromatic components* of  $G$  with respect to  $\omega$ , that is repeatedly contracting an edge  $e = uv$  such that  $\omega(u) = \omega(v)$ . If  $G'$  (with colouring  $\omega'$ ) is obtained from  $G$  (with initial colouring  $\omega$ ) in this way, it is clear that any sequence of moves that floods  $G$  with initial colouring  $\omega$  will also flood  $G'$  with initial colouring  $\omega'$ , and vice versa (up to possibly changing the vertex at which a move is played to another vertex in the same monochromatic component with respect to  $\omega$ ). We say that the graph  $G_1$  with colouring  $\omega_1$  is *equivalent* to the graph  $G_2$  with colouring  $\omega_2$  if there is a colour-preserving isomorphism (i.e. a bijection between vertices that preserves (non-)adjacency and colours) from the coloured graph obtained by contracting monochromatic components of  $G_1$  with respect to  $\omega_1$  to that obtained by contracting monochromatic components of  $G_2$  with respect to  $\omega_2$ .

We define  $m_G(G, \omega, d)$  to be the minimum number of moves required to give all vertices of  $G$  colour  $d$  in the free version of the game, and  $m_G(G, \omega)$  to be  $\min_{d \in C} m_G(G, \omega, d)$ . Analogously, for the fixed version of the game we write  $m_G^{(v)}(G, \omega, d)$  to be the minimum number of moves required to give all vertices of  $G$  colour  $d$  when all moves are played at the pivot vertex  $v$ , and define  $m_G^{(v)}(G, \omega)$  to be  $\min_{d \in C} m_G^{(v)}(G, \omega, d)$ .

Let  $\Omega(V, C)$  be the set of all surjective functions from  $V$  to  $C$ . We then define

$$M_c(G) = \max_{\omega \in \Omega(V, \{1, \dots, c\})} m_G(G, \omega),$$

so  $M_c(G)$  is the maximum number of moves that might be required in the free version to flood  $G$  in the worst case, taken over all possible initial colourings with  $c$  colours. We define  $M_c^{(v)}(G)$  analogously for the fixed version.

Let  $A$  be any subset of  $V$ . We set  $m_G(A, \omega, d)$  to be the minimum number of moves we must play in  $G$  (with initial colouring  $\omega$ ) in the free version to create a monochromatic component of colour  $d$  that contains every vertex in  $A$ , and  $m_G(A, \omega) = \min_{d \in C} m_G(A, \omega, d)$ . When the ground graph and area to be flooded agree, we may henceforth omit the subscript.

We say a move  $m = (v, d)$  is *played in*  $A$  if  $v \in A$ , and that  $A$  is *linked* if it is contained in a single monochromatic component. Subsets  $A, B \subseteq V$  are *adjacent* if there exists  $ab \in E$  with  $a \in A$  and  $b \in B$ . We will use the same notation when referring to (the vertex-set of) a subgraph  $H$  of  $G$  as for a subset  $A \subseteq V(G)$ .

## 1.2 Background results

One key result which we will exploit throughout this paper gives a characterisation of the number of moves required to flood a graph in terms of the number of moves required to flood its spanning trees.

More precisely we will apply the following result by Meeks and Scott [34].

**Theorem 1.1.** *Let  $G$  be a connected graph with colouring  $\omega$  from colour-set  $C$ . Then, for any  $d \in C$ ,*

$$m(G, \omega, d) = \min_{T \in \mathcal{T}(G)} m(T, \omega, d).$$

A corollary of this result, proved in the same paper, is that the number of moves required to flood a graph  $H$  cannot be increased when moves are played in a larger graph  $G$  which contains  $H$  as a subgraph.

**Corollary 1.2.** *Let  $G$  be a connected graph with colouring  $\omega$  from colour-set  $C$ , and  $H$  a connected subgraph of  $G$ . Then, for any  $d \in C$ ,*

$$m_G(V(H), \omega, d) \leq m_H(H, \omega, d).$$

We will also use a simple monotonicity result for paths, proved by the same authors in a previous paper [32].

**Lemma 1.3.** *Let  $P$  be a path, with colouring  $\omega$  from colour-set  $C$ , and let  $P'$  be a second coloured path with colouring  $\omega'$ , obtained from  $P$  by deleting one vertex and joining its neighbours. Then, for any  $d \in C$ ,  $m(P', \omega', d) \leq m(P, \omega, d)$ . We also have  $m(P', \omega') \leq m(P, \omega)$ .*

Another useful result, proved in an additional paper by Meeks and Scott [33], relates the number of moves required to flood the same graph with different initial colourings.

**Lemma 1.4.** *Let  $G$  be a connected graph, and let  $\omega$  and  $\omega'$  be two colourings of the vertices of  $G$  (from colour-set  $C$ ). Let  $\mathcal{A}$  be the set of all maximal monochromatic components of  $G$  with respect to  $\omega'$ , and for each  $A \in \mathcal{A}$  let  $c_A$  be the colour of  $A$  under  $\omega'$ . Then, for any  $d \in C$ ,*

$$m(G, \omega, d) \leq m(G, \omega', d) + \sum_{A \in \mathcal{A}} m(A, \omega, c_A).$$

This result means that we do not normally need to worry about the possible effect that moves played to flood a particular subgraph might have elsewhere.

## 2 General Bounds

In this section we obtain some general lower and upper bounds on the value of  $M_c(G)$  and  $M_c^{(v)}(G)$ , where  $G$  is an arbitrary connected graph. While these bounds are based on simple observations about the flooding process, we are able to show that all of them are tight for suitably chosen trees.

### 2.1 Lower bounds

We begin with a simple observation about the minimum number of moves required to flood a graph coloured with  $c$  colours.

**Proposition 2.1.** *Let  $G$  be any graph with colouring  $\omega$ , where  $\omega$  uses exactly  $c$  colours on  $G$ . Then*

$$m(G, \omega) \geq c - 1.$$

*Moreover, if every colour appears in at least two distinct monochromatic components with respect to  $\omega$ , then*

$$m(G, \omega) \geq c.$$

**Proof:** To see that the first statement is true, note that any move can reduce the number of colours present in the graph by at most one, and so, in order to reduce the total number of colours present from  $c$  to 1, at least  $c - 1$  moves are required. For the second part of the result, observe that the first move played can only change the colour of one monochromatic component under the initial colouring, and so if every colour initially appears in at least two distinct monochromatic components then the first move cannot reduce the total number of colours present on the graph; thus a total of at least  $c$  moves will be required to reduce the number of colours in the graph to 1.  $\square$

This gives an immediate lower bound on  $M_c(G)$ . A vertex  $v$  is said to be a *dominating vertex* if every other vertex is adjacent to  $v$ .

**Corollary 2.2.** *For any connected graph  $G$ ,  $M_c(G) \geq c - 1$ . If  $G$  has no dominating vertex, then  $M_c(G) \geq c$ .*

The same reasoning allows us to make a slightly stronger statement in the fixed case.

**Proposition 2.3.** *For any connected graph  $G = (V, E)$  and  $v \in V$ ,  $M_c^{(v)}(G) \geq c - 1$ . If  $v$  is not a dominating vertex, then  $M_c^{(v)}(G) \geq c$ .*

To see that this first pair of bounds is tight, consider the complete bipartite graph  $K_{1,n}$  with any colouring using  $c$  colours: we can always flood this graph by playing changing the colour of  $v$ , the unique non-leaf vertex,  $c - 1$  times. Now suppose that  $u$  is a leaf in  $K_{1,n}$ ; we obtain a new graph  $G$  by adding an additional vertex  $x$  adjacent only to  $u$ . Now, for any colouring  $\omega$  of  $G$  with  $c$  colours such that  $\omega(x) \neq \omega(u)$  and  $\omega(x) = \omega(v)$ , we can flood  $G$  by changing the colour of  $v$  precisely  $c$  times.

In order for the lower bounds we have obtained thus far to be tight, we required that some monochromatic component is adjacent to many others. We formalise this observation with a second general lower bound based on the structure maximum number of vertices adjacent to any connected subgraph. Note that this maximum number of neighbours will be bounded by a constant when the graph in question is obtained from some fixed graph by subdividing edges.

**Proposition 2.4.** *Let  $G = (V, E)$  be a connected graph on  $n \geq 1$  vertices, and let  $\omega$  be a proper colouring of  $G$ . Suppose that, for every set  $A \subset V$  such that  $G[A]$  is connected, we have  $|N(A)| \leq s$ . Then*

$$m(G, \omega) \geq \frac{1}{s}(n - 1).$$

**Proof:** We proceed by induction on  $m(G, \omega)$ ; the result is trivially true for  $m(G, \omega) = 0$ , so assume that  $m(G, \omega) > 0$ . Let  $S$  be a sequence which floods  $G$  with some colour  $d$ , where  $|S| = m(G, \omega)$ ; suppose

that the first move of  $S$  is  $\alpha = (v, d')$ , and let  $\omega'$  denote the colouring of  $G$  obtained by playing  $\alpha$  from the initial colouring  $\omega$ . Let  $A$  denote the monochromatic component of  $G$  with respect to  $\omega'$  that contains  $v$ . It is clear that every vertex of  $A$  other than  $v$  must belong to some monochromatic component with respect to  $\omega$  that is adjacent to  $A$ ; since all monochromatic components with respect to  $\omega$  are singletons (as  $\omega$  is a proper colouring) it follows that  $A \subset v \cup N(\{v\})$ . Note that  $\{v\}$  trivially induces a connected subgraph of  $G$ , so by assumption we have  $|N(\{v\})| \leq s$ , implying that  $|A| \leq 1 + s$ .

Now consider the graph  $G'$  obtained from  $G$  by contracting  $A$  to a single vertex; let  $\omega''$  be the colouring of  $G'$  derived from  $\omega'$ , and note that  $\omega''$  is a proper colouring of  $G'$ . It is also clear that contracting edges cannot create a set  $B$  such that  $G[B]$  is connected and  $|N(B)| > s$ , so we may apply the inductive hypothesis to see that

$$m(G', \omega'') \geq \frac{1}{s}(n - |A|) \geq \frac{1}{s}(n - s - 1) = \frac{1}{s}(n - 1) - 1.$$

Thus we can conclude that  $|S| \geq 1 + \frac{1}{s}(n - 1) - 1 = \frac{1}{s}(n - 1)$ , as required.  $\square$

Again, the same argument can be applied to make a slightly stronger statement in the fixed case.

**Proposition 2.5.** *Let  $G = (V, E)$  be a connected graph on  $n \geq 1$  vertices, fix  $v \in V$ , and let  $\omega$  be a proper colouring of  $G$ . Suppose that, for every set  $A \subset V$  such that  $v \in A$  and  $G[A]$  is connected, we have  $|N(A)| \leq s$ . Then*

$$m^{(v)}(G, \omega) \geq \frac{1}{s}(n - 1).$$

We can immediately deduce lower bounds on  $M_c(G)$  and  $M_c^{(v)}(G)$ .

**Corollary 2.6.** *Let  $G = (V, E)$  be a connected graph and suppose that, for every set  $A \subset V$  such that  $G[A]$  is connected, we have  $|N(A)| \leq s$ . Then*

$$M_c(G) \geq \frac{1}{s}(n - 1).$$

*If, for some  $v \in V$ , we have  $|N(A)| \leq s$  for all  $A \subset V$  such that  $v \in A$  and  $G[A]$  is connected, then*

$$M_c^{(v)}(G) \geq \frac{1}{s}(n - 1).$$

To see that these bounds are tight, consider a path on  $2t + 1$  vertices for some  $t \in \mathbb{N}$ , whose vertices are coloured alternately with colours 1 and 2. Here we have  $s = 2$ , and we can flood the path by changing colour of the midpoint exactly  $t$  times.

## 2.2 Upper bounds

We start with a simple bound based on the initial colouring of  $G$ .

**Proposition 2.7.** *Let  $G$  be any connected graph with colouring  $\omega$  from colour-set  $C$ . Then, for any  $d \in C$ ,*

$$m(G, \omega, d) \leq n - N_d(G, \omega).$$

**Proof:** This result follows immediately from the fact that, provided at least one vertex does not yet have colour  $d$ , it is always possible to play a move which increases the number of vertices having colour  $d$  by at least one: changing the colour of a vertex that does not already have colour  $d$  to  $d$  may additionally give some other vertices colour  $d$ , but all vertices that previously had colour  $d$  will be unchanged.  $\square$

This result can also be reformulated in terms of the number of colours used in the initial colouring.

**Proposition 2.8.** *Let  $G$  be any connected graph. Then*

$$M_c(G) \leq n - \left\lceil \frac{n}{c} \right\rceil.$$

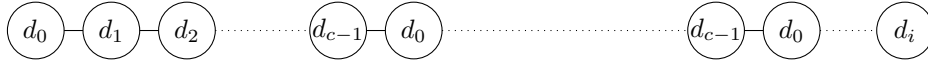
**Proof:** Let  $\omega$  be any colouring of  $G$  with  $c$  colours. There must then be at least one colour  $d$  such that  $N_d(G, \omega) \geq \lceil \frac{n}{c} \rceil$ , implying by Proposition 2.7 that  $m(G, \omega, d) \leq n - \lceil \frac{n}{c} \rceil$ .  $\square$

In the fixed version of the game, we do not necessarily benefit from having many vertices coloured with the same colour initially. Consider a path on  $n$  vertices, whose vertices are coloured alternately with colours 1 and 2, and fix the pivot to be one of the endpoints. By Proposition 2.5, we can see that we will require at least  $n - 1$  moves to flood the graph, even though half the vertices already have the same colour.

Returning our attention to the free version, we can show that the two simple upper bounds above are in fact tight for suitably coloured paths; we start by defining a useful family of colourings for paths.

**Definition.** *Let  $P = v_1 \dots v_n$  be a path with edge-set  $E = \{v_i v_{i+1} : 1 \leq i \leq n - 1\}$ ,  $C = \{d_0, \dots, d_{c-1}\}$  a set of colours, and  $\omega : V(P) \rightarrow C$  a proper colouring of  $P$ . The colouring  $\omega$  is said to be a  $C$ -rainbow colouring of  $P$  if there exists a permutation  $\pi : \{0, \dots, c - 1\} \rightarrow \{0, \dots, c - 1\}$  such that, for  $1 \leq i \leq n$ ,  $\omega(v_i) = d_{\pi(i \bmod c)}$ .*

Note that, up to relabelling of the colours, a  $C$ -rainbow coloured path must be as illustrated in Figure 2. We say that a colouring  $\omega$  of  $P$  is a *rainbow colouring* if it is a  $C$ -rainbow colouring for some colour-set  $C$ .



**Fig. 2:** A  $C$ -rainbow colouring of a path

We now demonstrate that any rainbow colouring of a path attains the upper bound from Proposition 2.7.

**Lemma 2.9.** *Let  $P$  be a path on  $n$  vertices, and  $\omega$  a rainbow colouring of  $P$ . Then, for any colour  $d$ ,*

$$m(P, \omega, d) \geq n - N_d(P, \omega).$$

**Proof:** We proceed by induction on  $m(P, \omega, d)$ . For the base case, suppose that  $m(P, \omega, d) = 0$ , which is only possible if the path is already monochromatic with colour  $d$ ; thus  $N_d(P, \omega) = n$  and so  $n - N_d(P, \omega) = 0 \leq m(P, \omega, d)$ , as required.

Now suppose that  $m(P, \omega, d) > 0$  and that the result holds for any path  $P'$ , colouring  $\omega'$  and colour  $d'$  such that  $m(P', \omega', d') < m(P, \omega, d)$ . Let  $S$  be an optimal sequence of moves to flood  $P$  with colour  $d$  (starting from the initial colouring  $\omega$ ), and let  $\alpha$  be the final move of  $S$ . There are now two cases to consider, depending on whether or not  $P$  is monochromatic immediately before  $\alpha$  is played.

Suppose first that  $P$  is monochromatic in some colour  $d' \neq d$  immediately before  $\alpha$  is played. In this case we know that  $m(P, \omega, d') \leq m(P, \omega, d) - 1$  and so we may apply the inductive hypothesis to see that

$$m(P, \omega, d') \geq n - N_{d'}(P, \omega).$$

Note that the number of vertices with colours  $d$  and  $d'$  in a rainbow colouring can differ by at most one, so we see that

$$\begin{aligned} m(P, \omega, d) &\geq m(P, \omega, d') + 1 \\ &\geq n - N_{d'}(P, \omega) + 1 \\ &\geq n - (N_d(P, \omega) + 1) + 1 \\ &= n - N_d(P, \omega), \end{aligned}$$

as required.

Now suppose that  $P$  is not monochromatic immediately before  $\alpha$  is played. In this case, before the final move, there must be either two or three monochromatic segments; we denote these segments  $P_1, \dots, P_\ell$  (where  $\ell \in \{2, 3\}$ ), and without loss of generality we may assume that  $P_2$  does not have colour  $d$  before  $\alpha$  is played. For each  $i \in \{1, \dots, \ell\}$ , let  $S_i$  be the subsequence of  $S \setminus \alpha$  consisting of moves played in a monochromatic component that intersects  $P_i$ ; note that these subsequences partition  $S \setminus \alpha$ . Moreover, observe that  $S_i$ , played in  $P_i$ , must make  $P_i$  monochromatic; for each  $i \neq 2$  the sequence  $S_i$  must flood  $P_i$  with colour  $d$ , while  $S_2$  must flood  $P_2$  with some colour  $d' \neq d$ . Thus we see that, for  $i \neq 2$ ,

$$m(P_i, \omega, d) \leq |S_i|,$$

and also

$$m(P_2, \omega, d') \leq |S_2|.$$

Observe further that  $\omega$  is a rainbow colouring of  $P_i$  so, as  $|S_i| \leq |S \setminus \alpha| < m(P, \omega, d)$  for all  $i$ , we can then apply the inductive hypothesis to see that, for  $i \neq 2$ ,

$$|S_i| \geq m(P_i, \omega, d) \geq |P_i| - N_d(P_i, \omega),$$

and that

$$|S_2| \geq m(P_2, \omega, d') \geq |P_2| - N_{d'}(P_2, \omega) \geq |P_2| - N_d(P_2, \omega) - 1$$

as the number of vertices that initially have colours  $d$  and  $d'$  can differ by at most one.



This then implies that

$$\begin{aligned}
m(P, \omega, d) &= |S| \\
&= 1 + \sum_{i=1}^{\ell} |S_i| \\
&= 1 + |S_2| + \sum_{i \neq 2} |S_i| \\
&\geq 1 + |P_2| - N_d(P_2, \omega) - 1 + \sum_{i \neq 2} (|P_i| - N_d(P_i, \omega)) \\
&= |P| - N_d(P, \omega),
\end{aligned}$$

as required.  $\square$

As the number of occurrences of each colour in a rainbow colouring of an  $n$ -vertex path with  $c$  colours is either  $\lfloor \frac{n}{c} \rfloor$  or  $\lceil \frac{n}{c} \rceil$ , we see that Proposition 2.8 is also tight.

**Corollary 2.10.** *Let  $P_t$  denote the path on  $t$  vertices, and assume that  $t \geq c$ . Then*

$$M_c(P_t) = t - \left\lceil \frac{t}{c} \right\rceil.$$

*Moreover, this maximum possible number of moves is obtained with a rainbow colouring.*

Since the tightness of these bounds is demonstrated by means of a long path, it is natural to ask how much better we can do if we exclude the presence of such paths; in the next result we show that we can bound the number of moves required in the fixed version in terms of only the total number of colours and the length of the longest induced path starting at the pivot vertex.

**Proposition 2.11.** *Let  $G = (V, E)$  be a connected graph, fix  $v \in V$ , and suppose that  $v$  has eccentricity  $r$ , Then*

$$M_c^{(v)}(G) \leq (c-1)r.$$

**Proof:** For  $1 \leq i \leq r$ , let  $V_i = \{u \in V : d(u, v) = i\}$ , and note that  $V = \{v\} \cup \bigcup_{1 \leq i \leq r} V_i$ . Now fix any colouring  $\omega \in \Omega(V, \{1, \dots, c\})$ . We will argue, by induction on  $r$ , that there is a sequence of moves played at  $v$  which will flood the graph. The base case, for  $r = 0$ , is trivial, so we will assume that  $r > 0$  and that the result holds for all graphs with radius smaller than  $r$ .

Let  $C_1$  be the set of colours, other than  $\omega(v)$ , that occur at vertices of  $V_1$  under  $\omega$ ; note that  $|C_1| \leq c-1$ . Then, cycling  $v$  through all colours in  $C_1$  will create a monochromatic component containing (at least) all of  $\{v\} \cup V_1$ ; we will denote the new colouring of  $G$  resulting from these moves by  $\omega'$ . Note that, in the graph obtained from  $G$  by contracting monochromatic components with respect to  $\omega'$ , the pivot vertex has eccentricity at most  $r-1$ , so by the inductive hypothesis we see that at most  $(c-1)(r-1)$  further moves are required to flood  $G$ . Hence the total number of moves required to flood  $G$  is at most  $(c-1) + (c-1)(r-1) = (c-1)r$ , as required.  $\square$

In the free version, we can choose to play all moves at a vertex whose eccentricity is equal to the radius of the graph; this gives the following immediate corollary.

**Corollary 2.12.** *Let  $G = (V, E)$  be a connected graph with radius  $r$ . Then*

$$M_c(G) \leq (c - 1)r.$$

We now argue that both of these bounds are tight. To do this, we will make use of a specific family of trees: we define  $T_{a,b}$  to be the tree obtained from the star  $K_{1,a}$  by subdividing every edge exactly  $b - 1$  times (so  $T_{a,b}$  is composed of  $a$  paths on  $b + 1$  vertices, all having a common first vertex). We begin with the fixed case.

**Lemma 2.13.** *Let  $v$  be the unique vertex of  $T_{(c-1)^r, r}$  with degree  $(c - 1)^r$ . Then*

$$M_c^{(v)}(G) \geq (c - 1)r.$$

**Proof:** We define a  $c$ -colouring  $\omega$  of  $V(T_{(c-1)^r, r})$  such that  $m^{(v)}(T_{(c-1)^r, r}, \omega) \geq (c - 1)r$ . We first set  $\omega(v) = 1$ . Now let  $\mathcal{S}_{c,r}$  be the set of all sequences of elements from  $\{1, \dots, c\}$  of length  $r$  with the following properties:

1. the first element of the sequence is not 1, and
2. no two consecutive elements of the sequence are the same.

Note that this definition implies that  $|\mathcal{S}_{c,r}| = (c - 1)^r$ . We will set our colouring  $\omega$  to colour one of the paths in our tree with each  $\sigma \in \mathcal{S}_{c,r}$ : to colour a path with  $\sigma$ , we give the vertex adjacent to  $v$  the colour that is the first element of  $\sigma$ , the next vertex the colour that is the second element, and so on. Note that the conditions on elements of  $\mathcal{S}_{c,r}$  ensure that this colouring  $\omega$  is a proper colouring of  $T_{c,r}$ .

Let  $S$  be any sequence of moves played at  $v$  which floods  $G$ . Note that there must be a colour  $c_1$ , other than 1, that is none of the first  $c - 2$  moves of  $S$ . We now define  $c_i$  inductively: set  $S_{i-1}$  to be the shortest initial segment of  $S$  that is a supersequence of  $c_1, \dots, c_{i-1}$ , and choose  $c_i$  to be a colour, not equal to  $c_{i-1}$ , that does not appear in the first  $c - 2$  moves of the sequence  $S$  after  $S_{i-1}$  has been removed. Observe that  $c_1, \dots, c_r$  is an element of  $\mathcal{S}_{c,r}$ , so there is some path in  $T_{c,r}$  whose vertices (starting from the vertex adjacent to  $v$ ) are coloured, in order,  $c_1, \dots, c_r$ . In order to flood this path, we must play  $S_r$ ; but by construction,  $|S_r| \geq (c - 1)r$ , so we must have  $|S| \geq (c - 1)r$ , as claimed.  $\square$

We now generalise this argument to the free version; we use an almost identical construction, but with even more paths incident with the central vertex.

**Lemma 2.14.** *Let  $T_{r(c-1)^{r+1}, r}$  denote the tree obtained from  $K_{1, r(c-1)^{r+1}}$  by subdividing each edge  $r - 1$  times. Then*

$$M_c(T_{r(c-1)^{r+1}, r}) \geq (c - 1)r.$$

**Proof:** Once again, we set  $v$  to be the vertex of  $T_{r(c-1)^{r+1}, r}$  with degree  $r(c - 1)^{r+1}$ . We define a colouring  $\omega$  of  $V(T_{r(c-1)^{r+1}, r})$  by setting  $\omega(v) = 1$  and, for each  $\sigma \in \mathcal{S}_{c,r}$ , we colour  $r(c - 1)$  of the paths in our tree with  $\sigma$ .

If all moves are played at  $v$ , we can use exactly the same reasoning as in the proof of Lemma 2.13, so it remains only to argue that no sequence of fewer than  $r(c-1)$  moves in which not all moves are played at  $v$  can flood the tree. Suppose that some number  $\alpha$  of the moves are not played in monochromatic components containing  $v$ , with  $1 \leq \alpha < (c-1)r$ . Note that any such move cannot change the colour of any vertex outside the path in which it is played. Thus, even if  $\alpha = (c-1)r - 1$ , there must still be at least one path with each colouring from  $\mathcal{S}_{c,r}$  whose colouring is not changed by any of these moves that is not played at  $v$ ; at least  $(c-1)r$  moves played at  $v$  will be required to flood these remaining paths.  $\square$

### 3 Graphs with high edge density

While we have seen that the upper bounds on  $M_c(G)$  that we derived in the previous section are tight, it is natural to ask to whether they are only tight for graphs with few edges: intuitively, increasing the number of edges in the graph should make it easier to flood the graph.

One simple question we might ask is as follows: given  $\delta \in (0, 1)$ , are the upper bounds in Propositions 2.11 and 2.12 tight for any graph  $G = (V, E)$  such that  $|E|/|V|^2 \geq \delta$ ? The answer to this question is yes: we can obtain a graph with arbitrarily high edge-density by adding a large clique to the constructions used in Lemmas 2.13 and 2.14, every vertex of which is adjacent to the central vertex of the tree. The addition of this clique does not change the radius of the graph, and no matter how we colour the clique the same number of moves will still be required to flood the whole graph.

However, this construction seems somewhat artificial: the value of  $M_c(G)$  is determined by some small part of the graph whose edge-density remains small. A more meaningful line of investigation is therefore to ask whether adding many edges (but no new vertices) to these constructions, in such a way that the radius remains unchanged, will significantly reduce  $M_c(G)$ .

In this section, we provide a partial answer to this question. If  $v$  denotes the central vertex of some tree  $T_{a,b}$ , we define  $V_i = \{u \in V(T_{a,b}) : d(v, u) = i - 1\}$  for  $1 \leq i \leq b + 1$ . Then the set of edges we can add without changing the number of vertices in each  $V_i$  is precisely

$$\{uw : u \in V_i, w \in V_j, |i - j| \leq 1\}.$$

Notice that the resulting graph is obtained from a blow-up of a path on  $b + 1$  vertices by making each vertex class (which is an independent set in the blow-up) into a clique. Since adding edges cannot increase the number of moves required (by Corollary 1.2), the following result tells us that  $b + c - \frac{b-1}{c} - 2$  moves suffice to flood this graph in the free version.

**Proposition 3.1.** *Let  $G = (V, E)$  be a blow-up of the path  $P_t$  on  $t$  vertices, and let  $\omega : V(G) \rightarrow C$  be a colouring of  $G$ . Suppose that  $Q$  is a subpath of  $G$  such that every vertex of  $G$  has a neighbour on  $Q$ . Then*

$$m(G, \omega) \leq m(Q, \omega) + (c - 1),$$

and in particular

$$m(G, \omega) \leq t - 2 - \left\lceil \frac{t-2}{c} \right\rceil + (c - 1).$$

**Proof:** Note that, by Corollary 1.2,  $m_G(Q, \omega) \leq m_Q(Q, \omega)$ , so it is possible to play at most  $m_Q(Q, \omega)$  moves in  $G$  to create a monochromatic component  $A$  of some colour  $d \in C$ , where  $A$  contains all of  $Q$ ;

by our assumptions on  $Q$ , every vertex in  $G$  either belongs to  $A$  or has a neighbour in  $A$ . We can therefore flood the remainder of  $G$  with at most  $c - 1$  further moves, changing the colour of  $A$  repeatedly to give it every colour in  $C \setminus \{d\}$ . This implies that

$$m(G, \omega) \leq m(Q, \omega) + (c - 1),$$

as required. The second part of the result then follows immediately from Proposition 2.8, together with the observation that a path containing precisely one vertex from every class except for the two end classes has the required property.  $\square$

The same argument gives an analogous result for the fixed version.

**Proposition 3.2.** *Let  $G = (V, E)$  be a blow-up of the path  $P_t$  on  $t$  vertices, fix  $v \in V$ , and let  $\omega : V(G) \rightarrow C$  be a colouring of  $G$ . If  $v$  belongs to one of the end classes of  $G$ , then*

$$m^{(v)}(G, \omega) \leq (t - 2) + (c - 1) = t + c - 3;$$

*if  $v$  does not belong to either end class, then*

$$m^{(v)}(G, \omega) \leq (t - 3) + (c - 1) = t + c - 4.$$

The main result of this section is that, provided  $b$  is sufficiently large compared with the number of colours, we can in fact improve on the simple bound of Proposition 3.1 in the free version: the number of moves required to flood a blow-up of a long path is in fact the same as the number required, in the worst case, to flood a path of the same length.

**Theorem 3.3.** *Let  $G$  be a blow-up of a path on  $t$  vertices, and suppose that  $c \geq 3$  and  $t \geq 2c^{10}$ . Then*

$$M_c(G) = t - \left\lceil \frac{t}{c} \right\rceil.$$

We have made no attempt to optimise the dependence of  $t$  on  $c$  in the statement of Theorem 3.3, and indeed conjecture that the result is true for much smaller values of  $t$ . However, it is clear that some dependence on  $c$  is necessary, as it follows from Proposition 2.1 that if  $G$  is a blow-up of a path on  $c$  vertices in which every vertex class has size at least two and, for  $1 \leq i \leq c$ ,  $\omega$  assigns colour  $i$  to both vertices corresponding to the  $i^{\text{th}}$  vertex of the path.

We also remark that we cannot improve on the bound of Proposition 3.2 in the fixed case: suppose that the vertex classes are  $V_1, \dots, V_t$ , with the pivot  $v \in V_1$ , and that the classes  $V_1, \dots, V_{t-1}$  are alternately coloured with colours 1 and 2, while  $V_t$  contains a vertex of every colour. It is easy to verify that we require  $t + c - 3$  moves in this case.

The remainder of the section is devoted to the proof of Theorem 3.3. In Section 3.1 we consider a generalisation of rainbow colourings to blow-ups of paths and obtain an upper bound on the number of moves required in this case; in Section 3.2 we generalise our results to colourings in which all vertices ‘‘blown up’’ from a single vertex receive the same colour; finally, in Section 3.3, we deal with arbitrary colourings.

First, we observe that the upper bound in Theorem 3.3 is optimal. Recall from the definition of a blow-up of a graph that, if  $G$  is a blow-up of a path on  $t$  vertices, the vertices of  $G$  can be partitioned into vertex-classes  $V_1, \dots, V_t$  such that each class  $V_i$  is an independent set in  $G$  and  $uw$  is an edge in  $G$  if and only if  $u \in V_i$  and  $w \in V_j$  where  $|i - j| = 1$ .

**Lemma 3.4.** *Let  $G$  be a blow-up of a path on  $t$  vertices. Then*

$$M_c(G) \geq t - \left\lceil \frac{t}{c} \right\rceil.$$

**Proof:** We define a colouring  $\omega : V(G) \rightarrow \{0, \dots, c-1\}$  by setting  $\omega(v) = i \bmod c$ , where  $v \in V_i$ , for  $1 \leq i \leq t$ . Let  $G'$  be the graph obtained from  $G$  by adding all edges within each  $V_i$ ; by Corollary 1.2 this cannot increase the number of moves required to flood the graph. Moreover, it is clear that  $G'$  with colouring  $\omega$  is equivalent (contracting monochromatic components) to a path on  $t$  vertices with a  $C$ -rainbow colouring. We therefore know from Corollary 2.10 that  $m(G', \omega) \geq t - \left\lceil \frac{t}{c} \right\rceil$ , so we also have  $m(G, \omega) \geq t - \left\lceil \frac{t}{c} \right\rceil$ , as required.  $\square$

### 3.1 Rainbow colourings

We begin by defining an important restricted family of colourings for graphs that are blow-ups of paths.

**Definition.** *Let  $G = (V, E)$  be a blow-up of the path  $P_t$  on  $t$  vertices, and let  $C = \{d_0, \dots, d_{c-1}\}$  be a set of colours. We say that the colouring  $\omega : V \rightarrow C$  is a path colouring of  $G$  if there exists a function  $f : \{1, \dots, t\} \rightarrow C$  such that, for each  $1 \leq i \leq t$ , we have  $\omega(v) = f(i)$  for every  $v \in V_i$ .*

Using this definition, we extend our definition of  $C$ -rainbow colourings to blow-ups of paths, to define a subfamily of path colourings.

**Definition.** *Let  $G = (V, E)$  be a blow-up of the path  $P_t$  on  $t$  vertices. Suppose that  $C = \{d_0, \dots, d_{c-1}\}$  is a set of colours, and  $\omega : V \rightarrow C$  is a path colouring of  $V$ . The colouring  $\omega$  is said to be a  $C$ -rainbow colouring of  $G$  if the corresponding colouring of  $P_t$  is a  $C$ -rainbow colouring of the path.*

As for paths, we say that  $\omega$  is a *rainbow* colouring if it is a  $C$ -rainbow colouring for some colour-set  $C$ . Note that the colouring used in the proof of Lemma 3.4 is a rainbow colouring.

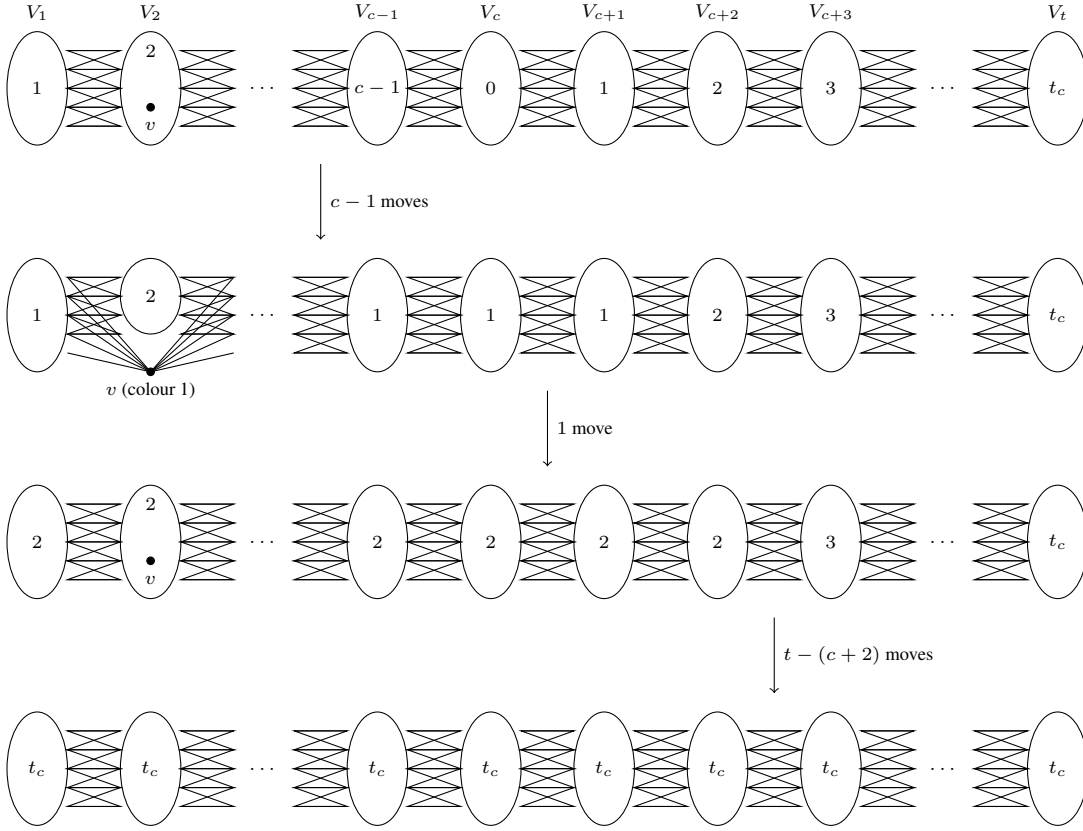
We now prove that Theorem 3.3 holds if we restrict our attention only to rainbow colourings.

**Lemma 3.5.** *Let  $G$  be a blow-up of the path  $P_t$ , and let  $\omega$  be a  $C$ -rainbow colouring of  $G$ . Then, if  $t \geq c + 2$ , we have*

$$m(G, \omega) \leq t - \left\lceil \frac{t}{c} \right\rceil.$$

**Proof:** We prove the result by induction on  $t$ . We begin by considering several base cases, which together cover the situation in which  $c + 2 \leq t \leq 3c + 1$ . As usual, we will denote by  $V_1, \dots, V_t$  the vertex-classes of  $G$ , and we may assume without loss of generality that  $V_i$  receives colour  $i \bmod c$  under  $\omega$ .

For the first of these cases, suppose that  $c + 2 \leq t \leq 2c$ . We describe a strategy to flood  $G$  with  $t - \left\lceil \frac{t}{c} \right\rceil = t - 2$  moves; this strategy is illustrated in Figures 3 and 4. First, we play  $c - 1$  moves at some vertex  $v \in V_2$ , giving this vertex colours  $3, 4, \dots, c - 1, 0, 1$  in turn; this will create a monochromatic

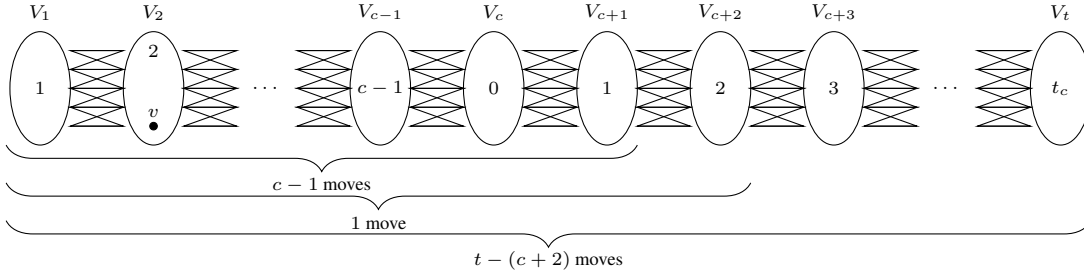


**Fig. 3:** The first base case for Lemma 3.5: detailed version. We write  $t_c$  for  $t \bmod c$ .

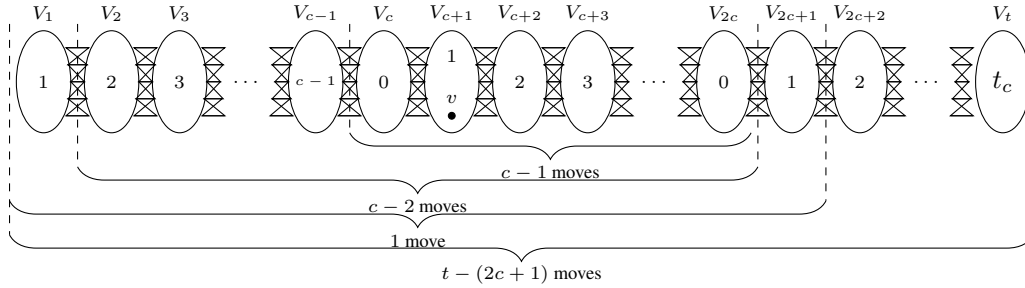
component containing  $v$  and all of  $V_1 \cup V_3 \cup \dots \cup V_{c+1}$ . Note that the only vertices of  $V_1 \cup \dots \cup V_{c+1}$  that do not belong to this monochromatic component are in  $V_2$  and have colour 2. Now we change the colour of this component to take colours  $2, \dots, t \bmod c$  in turn; this will extend our monochromatic component to contain all of  $V_{c+2} \cup \dots \cup V_t$  and, as the component takes colour 2 at the start of this subsequence of moves, all remaining vertices of  $V_2$  will also be flooded. Thus we have described a sequence of  $c - 1 + t - (c + 1) = t - 2$  moves which floods  $G$ , as required.

For the second base case, suppose that  $2c + 1 \leq t \leq 3c$ . In this case we play  $t - \lceil \frac{t}{c} \rceil = t - 3$  moves, all at a vertex  $v \in V_{c+1}$ , as illustrated in Figure 5. We first change the colour of  $v$  to take colours  $2, 3, \dots, 0$  in turn; this creates a monochromatic component containing  $v$  and all of  $V_c \cup V_{c+2} \cup \dots \cup V_{2c}$ . Next we give  $v$  colours  $c - 1, c - 2, \dots, 2$  in turn; at this point there is a monochromatic component containing all of  $V_2 \cup \dots \cup V_{2c}$  except for some vertices in  $V_{c+1}$  of colour 1. Playing one further move to give this component colour 1 therefore creates a monochromatic component containing all of  $V_1 \cup \dots \cup V_{2c+1}$ . To flood the remainder of  $G$ , we give this component colours  $2, \dots, t \bmod c$  in turn. This strategy allows us to flood  $G$  in a total of  $c - 1 + c - 2 + 1 + t - (2c + 1) = t - 3$  moves, as required.

For the final base case, suppose that  $t = 3c + 1$ . In this case we play  $t - 4 = t - \lceil \frac{t}{c} \rceil$  moves, all



**Fig. 4:** A reduced schematic of the first base case of Lemma 3.5. We write  $t_c$  for  $t \bmod c$ .

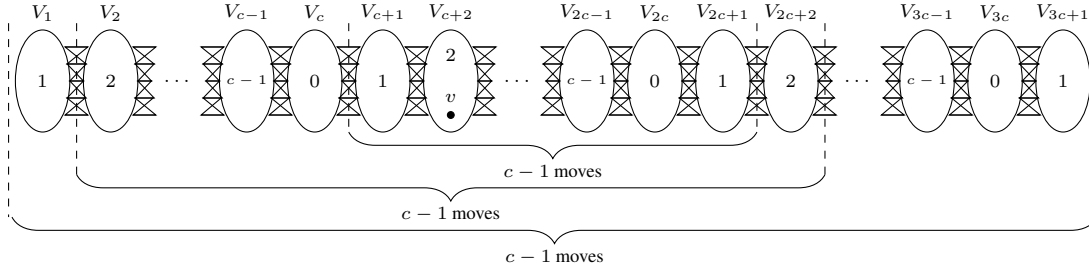


**Fig. 5:** The second base case for Lemma 3.5

at a vertex  $v \in V_{c+2}$ , as illustrated in Figure 6. We begin by giving  $v$  colours  $3, \dots, c-1, 0, 1$  in turn, which creates a monochromatic component containing all of  $V_{c+1} \cup \dots \cup V_{2c+1}$  except for vertices in  $V_{c+2}$  having colour 2. We play a further  $c-1$  moves in this component, giving it colours  $0, c-1, \dots, 2$  in turn, which creates a monochromatic component containing all of  $V_2 \cup \dots \cup V_{2c+2}$ . Finally, we give this component colours  $3, \dots, c-1, 0, 1$  in turn, which floods the remainder of the graph. Thus we can flood  $G$  with a total of  $c-1 + c-1 + c-1 = 3c-3 = t-4$  moves, as required.

From now on, therefore, we may assume that  $t \geq 3c+2$ , and that the result holds for any graph that is a blow-up of  $P_s$  for  $s < t$ . Using a similar strategy to that described in the second base case above, we can play  $2c-2$  moves which create a monochromatic component in  $G$  containing all of  $V_1 \cup \dots \cup V_{2c+1}$ . To achieve this, we play  $2c-1$  moves at a vertex  $v \in V_{c+1}$ : we give this vertex colours  $2, \dots, c-1, 0$ , followed by  $c-1, \dots, 2, 1$ . Note that the resulting monochromatic component ends up with colour 1, and that playing the sequence of moves described above in  $G$  will not change the colour of any vertex outside  $V_1 \cup \dots \cup V_{2c+1}$ . Thus, after playing this sequence, the resulting coloured graph is equivalent to a graph  $G'$  with colouring  $\omega'$ , where  $G'$  is a blow-up of the path  $P_{t-2c}$  and  $\omega'$  is a  $C$ -rainbow colouring of  $G'$ . Note that  $t-2c \geq c+2$ , so we can apply the inductive hypothesis to see that

$$m(G', \omega') \leq t - 2c - \left\lceil \frac{t-2c}{c} \right\rceil = t - 2c + 2 - \left\lfloor \frac{t}{c} \right\rfloor.$$



**Fig. 6:** The third base case for Lemma 3.5

This implies that

$$m(G, \omega) \leq 2c - 2 + t - 2c + 2 - \left\lceil \frac{t}{c} \right\rceil = t - \left\lceil \frac{t}{c} \right\rceil,$$

as required.  $\square$

### 3.2 General path colourings

Before proving that the upper bound is also valid for path blow-ups with any path colouring (provided that the path is sufficiently long compared with the number of colours), we need some auxiliary results. First of all, it is straightforward to verify the following characterisation of path colourings that are *not* rainbow colourings.

**Proposition 3.6.** *Let  $G$  be a blow-up of the path  $P_t$ , let  $f : \{1, \dots, t\} \rightarrow \{0, \dots, c-1\}$  be any function and  $C = \{d_0, \dots, d_{c-1}\}$  a set of colours, and let  $\omega$  be defined by setting  $\omega(u) = d_{f(i)}$  for all  $u \in V_i$  (for  $1 \leq i \leq t$ ). If  $\omega$  is not a  $C$ -rainbow colouring of  $G$ , then there exists  $1 \leq i < j \leq t$  such that  $j - i < c$  and  $f(i) = f(j)$ .*

We also need another result relating the number of moves required to flood a path and a collection of subpaths.

**Lemma 3.7.** *Let  $P$  be a path on  $t$  vertices with colouring  $\omega$  from colour-set  $C$ , where  $|C| = c$ , and let  $Q_1, \dots, Q_r$  be a collection of disjoint subpaths of  $P$ . Then*

$$m(P, \omega) \leq t - \sum_{i=1}^r (|Q_i| - 1) - \left\lceil \frac{t - \sum_{i=1}^r (|Q_i| - 1)}{c} \right\rceil + \sum_{i=1}^r m(Q_i, \omega).$$

**Proof:** For each  $1 \leq i \leq r$ , fix  $d_i \in C$  such that  $m(Q_i, \omega) = m(Q_i, \omega, d)$ . We now define a new colouring  $\omega'$  of  $P$  by setting

$$\omega'(v) = \begin{cases} d_i & \text{if } v \in Q_i \\ \omega(v) & \text{otherwise.} \end{cases}$$



Observe that  $P$  with colouring  $\omega'$  is equivalent to a path on at most  $t - \sum_{i=1}^r (|Q_i| - 1)$  vertices, so by Corollary 2.10 we have

$$m(P, \omega') \leq t - \sum_{i=1}^r (|Q_i| - 1) - \left\lceil \frac{t - \sum_{i=1}^r (|Q_i| - 1)}{c} \right\rceil.$$

Let  $\mathcal{A}$  be the set of maximal monochromatic components of  $P$  with respect to  $\omega'$ , where each  $A \in \mathcal{A}$  has colour  $d_A$  under  $\omega'$ . Then, by Lemma 1.4, we have

$$\begin{aligned} m(P, \omega) &\leq m(P, \omega') + \sum_{A \in \mathcal{A}} m(A, \omega, d_A) \\ &\leq t - \sum_{i=1}^r (|Q_i| - 1) - \left\lceil \frac{t - \sum_{i=1}^r (|Q_i| - 1)}{c} \right\rceil + \sum_{A \in \mathcal{A}} m(A, \omega, d_A). \end{aligned}$$

So it remains to check that  $\sum_{A \in \mathcal{A}} m(A, \omega, d_A) \leq \sum_{i=1}^r m(Q_i, \omega)$ . Note that it is possible that more than one of the subpaths  $Q_1, \dots, Q_r$  belongs to the same maximal monochromatic component with respect to  $\omega'$ ; suppose that  $A_1, \dots, A_s$  are the elements of  $\mathcal{A}$  that contain at least one subpath  $Q_i$ , and observe therefore that  $s \leq r$ . Moreover, it is clear that, for each  $1 \leq i \leq s$ ,

$$m(A_i, \omega, d_{A_i}) \leq \sum_{Q_j \subseteq A_i} m(Q_j, \omega, d_j) = \sum_{Q_j \subseteq A_i} m(Q_j, \omega).$$

Observe also that, for  $A \in \mathcal{A}$  with  $A \notin \{A_1, \dots, A_s\}$ ,  $A$  is also a monochromatic component of  $P$  with respect to  $\omega$ , so we have  $m(A, \omega, d_A) = 0$ . Hence, as no subpath  $Q_j$  belongs to more than one monochromatic component  $A_i$ , we have

$$\sum_{A \in \mathcal{A}} m(A, \omega, d_A) \leq \sum_{i=1}^s \sum_{Q_j \subseteq A_i} m(Q_j, \omega) = \sum_{i=1}^r m(Q_i, \omega),$$

completing the proof.  $\square$

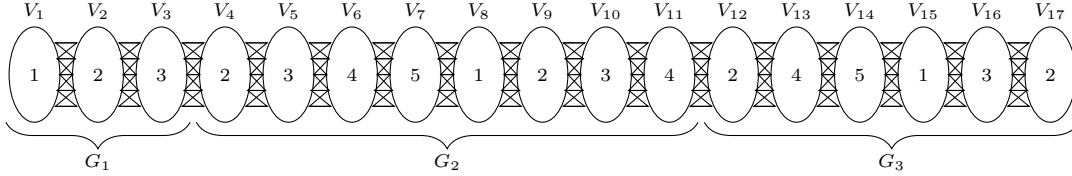
We now use these auxiliary results to extend our upper bound to cover all initial colourings that are path colourings. The key idea of the proof is to define a quantity that captures in a sense how far away the initial colouring is from a rainbow colouring. If the initial colouring is sufficiently different from a rainbow colouring, we can argue that we must be able to create a monochromatic end-to-end path significantly more quickly than in the rainbow case, meaning that we are then able to flood any remaining vertices greedily. In the event that the colouring does not differ so much from a rainbow colouring, we demonstrate how we may perform a sequence of flooding moves that is not too long and which results in a rainbow-coloured path blow-up, allowing us to apply the previous result.

**Lemma 3.8.** *Let  $G$  be a blow-up of the path  $P_t$ , let  $c \geq 3$  and let  $f : \{1, \dots, t\} \rightarrow \{1, \dots, c\}$  be any function, and let  $\omega$  be defined by setting  $\omega(u) = f(i)$  for all  $u \in V_i$  (for  $1 \leq i \leq t$ ). Then, if  $t \geq 2c^2(c-1)^3$ ,*

$$m(G, \omega) \leq t - \left\lceil \frac{t}{c} \right\rceil.$$

**Proof:** Without loss of generality, we may assume that, for every  $1 \leq x < t$ ,  $f(x) \neq f(x+1)$ , as otherwise we could contract all vertices of  $V_x \cup V_{x+1}$  to a single vertex, obtaining an equivalent coloured graph which is a blow-up of a path on  $t-1$  vertices. Now observe that, for any path colouring  $\omega$  of  $G$ , the graph  $G$  can be decomposed into subgraphs  $G_1, \dots, G_r$ , where each  $G_i$  is a blow-up of the path  $P_{t_i}$  and  $\sum_{i=1}^r t_i = t$ , in such a way that  $\omega$  is a  $C$ -rainbow colouring of  $G_i$  for each  $1 \leq i \leq r$ ; it is clear that such a decomposition must exist since setting  $G_i = G[V_i]$  for  $1 \leq i \leq t$  will do.

A more meaningful decomposition with the required properties can be constructed greedily, as illustrated in Figure 7. We set  $s_1 = 1$  and choose  $t_1$  to be the largest integer such that  $\omega$  is a rainbow colouring of  $G[V_1 \cup \dots \cup V_{t_1}]$ ; given  $t_1, \dots, t_i$ , we set  $s_{i+1} = 1 + \sum_{j=1}^i t_j$  and choose  $t_{i+1}$  to be the largest integer such that  $\omega$  is a rainbow colouring of  $G[V_{s_{i+1}} \cup \dots \cup V_{s_{i+1}+t_{i+1}-1}]$ . We call the decomposition constructed in this way the *greedy rainbow decomposition* of  $(G, \omega)$ , and denote by  $\text{grd}(G, \omega)$  the collection of subgraphs in this decomposition. Then  $\text{grd}(G, \omega) = \{G_1, \dots, G_r\}$  (for some  $r \geq 1$ ), where  $G_i = G[V_{s_i} \cup V_{s_i+1} \cup \dots \cup V_{s_{i+1}-1}]$ , and the greedy construction guarantees that, for each  $1 \leq i \leq r = |\text{grd}(G, \omega)|$ , there exists  $x_i$  with  $\max\{s_i, s_{i+1} - c + 1\} \leq x_i \leq s_{i+1} - 1$  such that  $f(x_i) = f(s_{i+1})$ .



**Fig. 7:** An example of a greedy rainbow decomposition

Suppose first that  $|\text{grd}(G, \omega)| > 2c(c-1)$ . We will argue that in this case we can flood  $G$  in at most  $m(G, \omega)$  moves by first creating a monochromatic end-to-end path and then cycling through any remaining colours.

Fix  $Q$  to be any path that contains precisely one vertex from each vertex class  $V_1, \dots, V_t$ . Now, for each  $1 \leq i \leq r-1$ , set  $Q_i$  to be the segment of  $Q$  induced by  $Q \cap (V_{s_i} \cup \dots \cup V_{s_{i+1}})$ . Observe that for each  $Q_i$ , by definition of  $x_i$ , we have  $|Q_i| \leq c$ , and  $m(Q_i, \omega, f(x_i)) \leq |Q_i| - 2$ , by Proposition 2.7. We are not quite able to apply Lemma 3.7, as for any  $i$  it is possible that  $Q_i$  and  $Q_{i+1}$  intersect in one vertex; however, it is clear that  $Q_i \cap Q_{i+2} = \emptyset$  for any  $i$ , so it is certainly the case that  $\{Q_{2i} : 1 \leq i \leq \lfloor \frac{|\text{grd}(G, \omega)|}{2} \rfloor\}$  is a collection of disjoint subpaths of  $Q$ . Setting  $r = |\text{grd}(G, \omega)|$ , Lemma 3.7 now tells us that

$$\begin{aligned} m(Q, \omega) &\leq t - \sum_{i=1}^{\lfloor \frac{r}{2} \rfloor} (|Q_{2i}| - 1) - \left\lceil \frac{t - \sum_{i=1}^{\lfloor \frac{r}{2} \rfloor} (|Q_{2i}| - 1)}{c} \right\rceil + \sum_{i=1}^{\lfloor \frac{r}{2} \rfloor} m(Q_{2i}, \omega) \\ &\leq t - \left\lceil \frac{t}{c} \right\rceil - (c-1), \end{aligned}$$

since  $|Q_i| \leq c$ ,  $m(Q_{2i}, \omega) \leq |Q_{2i}| - 2$ , and  $r = |\text{grd}(G, \omega)| \geq 2c(c-1)$ . Proposition 3.1 then gives  $m(G, \omega) \leq t - \lceil \frac{t}{c} \rceil$ , as required.

So we may assume from now on that  $|\text{grd}(G, \omega)| \leq 2c(c-1)$ . In this case it suffices to prove the following claim, by our initial assumption that  $t \geq 2c^2(c-1)^3$ .

**Claim.** *If  $t \geq c(c-1)^2 |\text{grd}(G, \omega)|$  then*

$$m(G, \omega) \leq t - \left\lfloor \frac{t}{c} \right\rfloor.$$

The base case, for  $|\text{grd}(G, \omega)| = 1$ , follows from Lemma 3.5 (since in this case  $\omega$  must be a rainbow colouring of  $G$ ), so we may assume from now on that  $|\text{grd}(G, \omega)| \geq 2$  and that the claim holds for any such graph  $\tilde{G}$  with colouring  $\tilde{\omega}$  such that  $|\text{grd}(\tilde{G}, \tilde{\omega})| < |\text{grd}(G, \omega)|$ .

By our assumption that  $t \geq c(c-1)^2 |\text{grd}(G, \omega)|$ , there must be some  $1 \leq i \leq |\text{grd}(G, \omega)|$  such that  $t_i \geq c(c-1)^2 > 2c$ . Note that we may assume without loss of generality that  $i \neq |\text{grd}(G, \omega)|$ : if the only such subgraph in the decomposition is  $G_{|\text{grd}(G, \omega)|}$  then we can reverse the ordering of the vertex classes so that we instead have  $t_1 > 2c$  after re-ordering (the subgraphs of the decomposition may not be the same as before, but our longest section can only increase in length in this new setting). We will now consider the subgraph  $H_0 = G[V(G_i) \cup V(G_{i+1})]$ , where  $\omega_0$  is the restriction of  $\omega$  to  $V(H_0)$ .

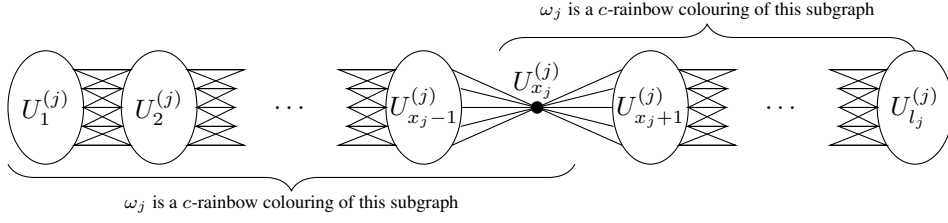
We will describe how to play a sequence of moves in  $H_0$  which results in a rainbow colouring of this subgraph; if this does not decrease the length of the underlying path too much (when monochromatic components are contracted) we then invoke the inductive hypothesis, and otherwise we can complete the proof directly. Specifically, we describe how to obtain a sequence of graphs  $(H_j)_{j=1}^s$  (for some  $s \geq 1$ ), with corresponding colourings  $(\omega_j)_{j=1}^s$ , with three key properties. Note that  $H_0$  itself is not included in this sequence; it does not have the second property listed below.

We denote by  $U_1^{(0)}, \dots, U_{\ell_0}^{(0)}$  the vertex classes of  $H_0$ , where  $\ell_0 = t_i + t_{i+1}$ , and let  $f_0 : \{1, \dots, \ell_0\} \rightarrow C$  be the function such that, for every  $1 \leq z \leq \ell_0$ ,  $\omega_0(u) = f_0(z)$  for each  $u \in U_z^{(0)}$ . For each  $j$ , the coloured graph  $(H_j, \omega_j)$  then has the following properties:

1.  $H_j$  is a blow-up of a path  $P_{\ell_j}$ , where  $\ell_j \geq \ell_0 - (j+1)(c-1) - 1$ , and  $\omega_j$  is a proper path colouring of  $H_j$ ,
2. There exists  $x_j \in \{1, \dots, \ell_j\}$  such that, if the vertex classes of  $H_j$  are  $U_1^{(j)}, \dots, U_{\ell_j}^{(j)}$ , then  $|U_{x_j}^{(j)}| = 1$  and  $\omega_j$  is a  $C$ -rainbow colouring of both  $H_j[U_1^{(j)} \cup \dots \cup U_{x_j}^{(j)}]$  and  $H_j[U_{x_j}^{(j)} \cup \dots \cup U_{\ell_j}^{(j)}]$  (as illustrated in Figure 8), and
3. If the colouring  $\omega'_j$  of  $V(H_0)$  is defined by

$$\omega'_j(v) = \begin{cases} \omega_0(v) & \text{if } v \in U_z^{(0)} \text{ and either } z < x_j \text{ or } z > \ell_0 - (\ell_j - x_j) \\ \omega_j(x_j) & \text{otherwise,} \end{cases}$$

then  $H_0$  with colouring  $\omega'_j$  is equivalent to  $H_j$  with colouring  $\omega_j$ ; moreover, if  $\mathcal{A}_j$  is the set of maximal monochromatic components of  $H_0$  with respect to  $\omega'_j$ , where each  $A \in \mathcal{A}_j$  has colour  $d_A^j$  under  $\omega'_j$ , then  $\sum_{A \in \mathcal{A}_j} m(A, \omega_0, d_A^j) \leq \ell_0 - \ell_j - j - 1$ .

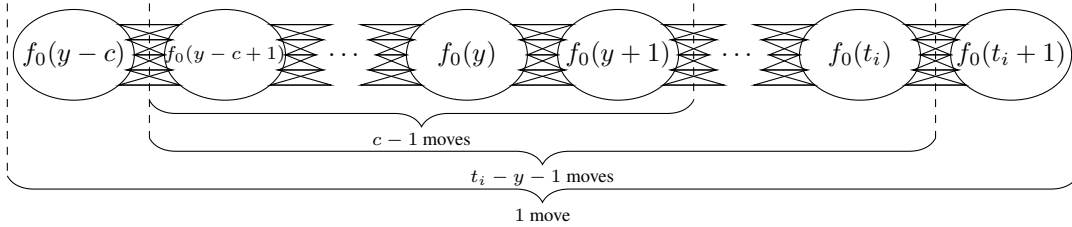


**Fig. 8:** The structure of  $H_j$

Informally, we repeatedly perform moves to create a single monochromatic component at the boundary of the two rainbow-coloured segments (whose colour is consistent with both rainbow colourings), taking care to make sure that we do not play too many moves or shorten the path too much at any stage. The key idea is to exploit the fact that some colour must occur more frequently, as we traverse the subgraph from left to right, than would happen under a rainbow colouring.

We describe in detail how to obtain the first pair  $(H_1, \omega_1)$ ; the method for constructing further pairs is very similar but somewhat simpler. Throughout, the only assumption we require, in addition to the fact that  $(H_j, \omega_j)$  has the three stated properties, is that  $\omega_j$  is not a rainbow colouring of  $H_j$ .

By construction of  $\text{grd}(G, \omega)$ , we know that there is some  $y \in \{t_i - c + 2, \dots, t_i\}$  such that  $f_0(y) = f_0(t_i + 1)$ : if not, then we would have chosen  $G_i$  to include at least one more vertex class. In fact, by our assumption that no two consecutive vertex classes receive the same colour under  $\omega$ , we know that  $y \in \{t_i - c + 2, \dots, t_i - 1\}$ . Since  $\omega_0$  is a  $C$ -rainbow colouring of  $G_i$ , and  $t_i > 2c$ , we also know that  $f_0(y - c) = f_0(y) = f_0(t_i + 1)$ . Now set  $F_0 = H_0[U_{y-c}^{(0)} \cup \dots \cup U_{t_i+1}^{(0)}]$ . We now describe a sequence of moves to flood  $F_0$  with colour  $f(y)$  in at most  $t_i + c - y - 1$  moves, all played at some vertex  $v_0 \in U_y^{(0)}$ ; this sequence of moves is illustrated in Figure 9. We begin by giving  $v_0$  colours  $f_0(y - 1), \dots, f_0(y - c + 1) = f(y + 1)$  in turn; this will create a monochromatic component containing  $v_0$  and all of  $U_{y-c+1}^{(0)} \cup \dots \cup U_{y-1}^{(0)} \cup U_{y+1}^{(0)}$ , and so that the only vertices of  $U_{y-c+1}^{(0)} \cup \dots \cup U_{y+1}^{(0)}$  not linked to this component have colour  $f_0(y)$ . We then give this component colours  $f_0(y + 2), \dots, f_0(t_i + 1) = f_0(y) = f_0(y - c)$  in turn, which will clearly flood all remaining vertices in  $F_0$ . The total number of moves played is therefore  $c - 1 + t_i - y = t_i + c - y - 1$ , implying that  $m(F_0, \omega_0, f_0(y)) \leq t_i + c - y - 1$ .



**Fig. 9:** Flooding the subgraph  $F_0$

Now define  $H_1$  to be the graph obtained from  $H_0$  by contracting all vertices of  $F_0$  to a single vertex  $w_1$ , and  $\omega_1$  to be the colouring of  $H_1$  that agrees with  $\omega_0$  on all vertices of  $H_1$  except  $w_1$ , and gives  $w_1$

colour  $f_0(y)$ . We claim that  $H_1$  with colouring  $\omega_1$  has the three properties listed above. It is clear that  $H_1$  is a blow-up of a path on

$$\ell_1 = \ell_0 - (t_i + 1 - y + c) \geq \ell_0 - 2c + 1 = \ell_0 - (1 + 1)(c - 1) - 1$$

vertices and that  $\omega_1$  is a proper path colouring of  $H_1$ , as required to satisfy the first condition. For the second condition, set  $x_1 = y - c$ , and note that  $U_{x_1}^{(1)} = \{w_1\}$ . Further define  $f_1 : \{1, \dots, \ell_1\} \rightarrow C$  to be the function so that  $\omega_1(u) = f_1(z)$  for every  $u \in U_z^{(1)}$ , for  $1 \leq z \leq \ell_1$ . Observe that  $f_1(z) = f_0(z)$  for  $1 \leq z \leq y - c$ , so it follows from the fact that  $\omega_0$  is a  $C$ -rainbow colouring of  $H_0[U_1^{(0)} \cup \dots \cup U_{t_i}^{(0)}] \supset H_0[U_1^{(0)} \cup \dots \cup U_{y-c}^{(0)}]$  that  $\omega_1$  is a  $C$ -rainbow colouring of  $H_1[U_1^{(1)} \cup \dots \cup U_{x_1}^{(1)}]$ . Moreover, for  $y - c \leq z \leq \ell_1$ , we see that  $f_1(z) = f_0(z + \ell_0 - \ell_1)$ , so the fact that  $\omega_0$  is a  $C$ -rainbow colouring of  $H_0[U_{t_i+1}^{(0)} \cup \dots \cup U_{\ell_0}^{(0)}]$  implies that  $\omega_1$  is a  $C$ -rainbow colouring of

$$H_1[U_{t_i+1-(\ell_0-\ell_1)}^{(0)} \cup \dots \cup U_{\ell_0-(\ell_0-\ell_1)}^{(0)}] = H_1[U_{x_1}^{(1)} \cup \dots \cup U_{\ell_1}^{(1)}].$$

Thus the second condition holds. Finally, for the third condition, it is clear from the definition of  $\omega_1$  that  $H_1$  with colouring  $\omega_1$  is equivalent to the graph  $H_0$  with colouring  $\omega'_1$  (with  $\omega'_1$  defined with respect to  $\omega_1$  as in the statement of the third condition); note also that the only maximal monochromatic component of  $H_0$  with respect to  $\omega'_1$  that is not also a maximal monochromatic component with respect to  $\omega_0$  is  $F_0$  (and it is straightforward to verify that  $F_0$  is indeed a maximal monochromatic component of  $H_0$  with respect to  $\omega'_1$ ). Thus, if  $\mathcal{A}_1$  denotes the set of maximal monochromatic components of  $H_0$  with respect to  $\omega'_1$  and each  $A \in \mathcal{A}_1$  has colour  $d_A$  under  $\omega'_1$ , we see that

$$\begin{aligned} \sum_{A \in \mathcal{A}_1} m(A, \omega_0, d_A) &= m(F_0, \omega_0, \omega_1(x_1)) \\ &= m(F_0, \omega_0, f_0(y)) \\ &\leq t_i + c - y - 1 \\ &= (t_i + 1 - y + c) - 2 \\ &= \ell_0 - \ell_1 - 2, \end{aligned}$$

as required to satisfy the third condition. This completes the definition of  $H_1$  and  $\omega_1$ .

Observe that we can continue applying the same general procedure (omitting the part of the process that makes  $F_0$  monochromatic) to obtain a new pair  $(H_{j+1}, \omega_{j+1})$  with the same three properties so long as  $(H_j, \omega_j)$  has the three stated properties and  $\omega_j$  is not a rainbow colouring of  $H_j$ . Suppose that we construct a sequence of pairs  $(H_j, \omega_j)$  for  $1 \leq j \leq s$  in this way, where  $s$  is as large as possible. By maximality of  $s$ , we may assume that  $\omega_s$  is a rainbow colouring of  $H_s$ , as otherwise we could continue. (Note that if our colouring is not a rainbow colouring, this imposes a minimum condition on the length of the path, so we do not need to consider separately the possibility of our path becoming too short to apply the procedure.) We define  $\omega'$  to be the colouring of  $G$  which agrees with  $\omega$  on all vertices that do not belong to  $H_0$ , and with  $\omega'_s$  on all vertices of  $H_0$ . Note that the maximal monochromatic components of  $G$  with respect to  $\omega'$  that are *not* also maximal monochromatic components with respect to  $\omega$  are precisely the maximal monochromatic components of  $H_0$  with respect to  $\omega' = \omega'_s$  (and recall also that  $\omega_0$  is the

restriction of  $\omega$  to  $H_0$ ). Thus we can apply Lemma 1.4 to see that

$$\begin{aligned} m(G, \omega) &\leq m(G, \omega') + \sum_{A \in \mathcal{A}_s} m(A, \omega_0, d_A^s) \\ &\leq m(G, \omega') + \ell_0 - \ell_s - s - 1 \end{aligned} \quad (3.1)$$

by the third condition on  $(H_s, \omega_s)$ . There are now two cases to consider, depending on the value of  $s$ .

First, suppose that  $s \geq c(c-1)$ . In this case we argue that we can continue by creating a monochromatic end-to-end path and then cycling through any remaining colours. Let  $Q$  be a path in  $G$  which contains precisely one vertex from each class. Note that there will be a segment of  $\ell_0 - \ell_s + 1$  consecutive vertices on  $Q$  which have the same colour under  $\omega'$  so, under this colouring,  $Q$  is equivalent to a path of length  $|Q| - \ell_0 + \ell_s$ ; Proposition 3.1, together with Corollary 2.10, therefore implies that

$$m(G, \omega') \leq t - \ell_0 + \ell_s - \left\lceil \frac{t - \ell_0 + \ell_s}{c} \right\rceil + c - 1.$$

Substituting this in (3.1) and using both the fact that  $\ell_s \geq \ell_0 - (s+1)(c-1) - 1$  (the first condition on  $(H_s, \omega_s)$ ) and the assumption that  $s \geq c(c-1)$ , this gives

$$m(G, \omega) \leq t - \left\lceil \frac{t}{c} \right\rceil,$$

as required.

Now suppose instead that  $s < c(c-1)$ ; in this case we invoke the inductive hypothesis. Since  $\omega_s$  is a  $C$ -rainbow colouring of  $H_0$ , it is clear that  $|\text{grd}(G, \omega')| < |\text{grd}(G, \omega)|$ . Moreover,  $G$  with colouring  $\omega'$  is equivalent to a graph  $\tilde{G}$  with colouring  $\tilde{\omega}$ , where  $\tilde{G}$  is a blow-up of a path on  $r$  vertices with

$$\begin{aligned} r &= t - (\ell_0 - \ell_s) \\ &> c(c-1)^2 |\text{grd}(G, \omega')| \end{aligned}$$

(making use of our assumption on the value of  $t$  and the fact that  $|\text{grd}(G, \omega')| \leq |\text{grd}(G, \omega)| - 1$ ). Thus we can apply the inductive hypothesis to see that the claim holds for  $G$  with colouring  $\omega'$ , implying that

$$m(G, \omega') \leq t - (\ell_0 - \ell_s) - \left\lceil \frac{t - (\ell_0 - \ell_s)}{c} \right\rceil.$$

Substituting into (3.1), then gives the required result, completing the proof of the claim, and hence proving the result.  $\square$

### 3.3 Arbitrary colourings

In this section we show that our upper bound can be extended to all initial colourings. The structure of this proof is in some ways similar to the previous result: we define a notion of the distance of a colouring from a path colouring, and then consider two cases. If the colouring differs sufficiently from a path colouring, we can quickly create an end-to-end path and flood any remaining vertices greedily, whereas if our initial colouring is sufficiently close to a path colouring we demonstrate how to play a sequence of moves that results in a path-coloured graph, allowing us to apply the previous result.

**Lemma 3.9.** *Let  $G$  be a blow-up of the path  $P_t$ , where  $t \geq 2c^{10}$ . Then*

$$M_c(G, \omega) \leq t - \left\lceil \frac{t}{c} \right\rceil.$$

**Proof:** let  $\omega$  be any colouring of  $G$  from colour-set  $C = \{1, \dots, c\}$ . We begin by setting

$$\theta(G, \omega) = |\{i : 1 \leq i \leq t \text{ and } \omega \text{ is not constant on } V_i\}|.$$

Suppose first that  $\theta(G, \omega) \geq c(c-1)$ , and set

$$n_j = |\{i : 1 \leq i \leq t \text{ and } \exists u \in V_i \text{ with } \omega(u) = j\}|.$$

Since there are  $\theta(G, \omega)$  vertex classes that each contain vertices of at least two distinct colours, we see that

$$\sum_{j=1}^c n_j \geq t + \theta(G, \omega) \geq t + c(c-1).$$

Thus there exists some  $j \in \{1, \dots, c\}$  such that  $n_j \geq \left\lceil \frac{t}{c} \right\rceil + (c-1)$ . Observe therefore that there exists a path  $Q$  containing precisely one vertex from each vertex class  $V_1, \dots, V_t$  and so that at least  $\left\lceil \frac{t}{c} \right\rceil + (c-1)$  vertices on  $Q$  have colour  $j$  under  $\omega$ . It then follows from Proposition 2.7 that  $m(Q, \omega) \leq t - \left\lceil \frac{t}{c} \right\rceil - (c-1)$ , so Proposition 3.1 gives

$$m(G, \omega) \leq t - \left\lceil \frac{t}{c} \right\rceil - (c-1) + (c-1) = t - \left\lceil \frac{t}{c} \right\rceil,$$

as required.

Thus from now on we will assume that  $\theta(G, \omega) < c(c-1)$ . In this case it clearly suffices to prove the following claim, since we are assuming that  $t \geq 2c^{10} > 2c^8(\theta(G, \omega) + 1)$ .

**Claim.** *Suppose that  $t > 2c^8(\theta(G, \omega) + 1)$ . Then*

$$m(G, \omega) \leq t - \left\lceil \frac{t}{c} \right\rceil.$$

We prove the claim by induction on  $\theta(G, \omega)$ . In the base case, for  $\theta(G, \omega) = 0$ , we know that  $\omega$  must in fact be a path-colouring of  $G$  and so the result follows immediately from Lemma 3.8. Thus we may assume that  $\theta(G, \omega) \geq 1$  and that the result holds for any graph  $G'$  with colouring  $\omega'$  such that  $\theta(G', \omega') < \theta(G, \omega)$ .

Since  $t \geq 2c^8(\theta(G, \omega) + 1)$ , there exists some vertex class  $V_i$  such that  $\omega$  is not constant on  $V_i$ , but for  $1 \leq j \leq 2c^8$  we either have  $\omega$  constant on every  $V_{i+j}$ , or else  $\omega$  is constant on every  $V_{i-j}$ ; reversing the order of the vertex classes if necessary, we may assume without loss of generality that we have  $\omega$  constant on every  $V_{i+j}$  for  $1 \leq j \leq 2c^8$ .

The first step in our strategy to flood  $G$  is to perform a series of moves in the  $(2c^5 + 1)(c^2(c-1) + 1)$  classes to the right of  $V_i$ , resulting in a colouring of these vertices that makes the subgraph they induce

(after contracting monochromatic components) equivalent to a path. Note that  $\omega$  defines a path colouring on the subgraph induced by these classes.

We split this subgraph into  $c^2(c-1) + 1$  consecutive blocks,  $B_1, \dots, B_{c^2(c-1)+1}$ , where each block consists of the subgraph induced by  $2c^5 + 1$  consecutive vertex classes. As the restriction of  $\omega$  to any block  $B_\ell$  is a path colouring, it follows from Lemma 3.8 that  $B_\ell$  can be made monochromatic in some colour  $d_\ell$  with at most

$$2c^5 + 1 - \left\lceil \frac{2c^5 + 1}{c} \right\rceil = 2c^5 - 2c^4$$

moves. Let  $\bar{\omega}$  be the colouring of  $V(G)$  that assigns  $d_\ell$  to every vertex of  $B_\ell$  (for  $1 \leq \ell \leq c^2(c-1) + 1$ ), and agrees with  $\omega$  elsewhere. Then Lemma 1.4 tells us that

$$m(G, \omega) \leq m(G, \bar{\omega}) + (c^2(c-1) + 1)(2c^5 - 2c^4). \quad (3.2)$$

We now consider the graph  $G'$ , obtained by contracting monochromatic components of  $G$  with respect to  $\bar{\omega}$ , and the corresponding colouring  $\omega'$ . Note that  $G'$  is a blow-up of a path on  $t' \leq t - 2c^5(c^2(c-1) + 1)$  vertices. In the remainder of the proof we will argue that in fact  $m(G', \omega') \leq t' - \left\lceil \frac{t'}{c} \right\rceil$ ; it is straightforward to check that substituting this bound on  $m(G, \bar{\omega})$  in (3.2) gives the required result.

We now prove this bound on  $m(G', \omega')$ . Recall from the construction of  $G'$  that, if the vertex classes of  $G'$  are  $U_1, \dots, U_{t'}$ , we have  $|U_{i+\ell}| = 1$  for  $1 \leq \ell \leq c^2(c-1) + 1$ . We may also assume without loss of generality that the colours assigned to vertices of  $U_i$  by  $\omega'$  are  $\{1, \dots, r\}$  for some  $r \geq 2$ .

If every colour in  $\{1, \dots, r\}$  is assigned to one of the first  $c$  vertex classes to the right of  $U_i$  by  $\omega'$ , then we can create a monochromatic component containing all of  $U_i$  and the  $c$  vertex classes to its right using at most  $c-1$  moves: we give the unique vertex in  $U_{i+1}$  the colours of  $U_{i+2}, U_{i+3}, \dots, U_{i+c}$  in turn. Otherwise, there must be some  $d \in \{1, \dots, r\}$  that is not assigned to any of the first  $c$  vertex-classes to the right of  $U_i$  by  $\omega'$ . Then, by Proposition 2.7, we can perform an ‘‘efficient flooding sequence’’ in which we flood this subpath on  $c$  vertices with at most  $c-2$  moves (as some colour must be repeated), thus reducing the length of an end-to-end path by  $c-1$ . We continue in this way until either we have performed an efficient flooding operation  $c(c-1)$  times, or else we are able, by the method described above, to flood  $U_i$  and the  $c$  vertex classes immediately to the right with  $c-1$  moves.

In the latter case, we have played  $a(c-2) + (c-1)$  moves, for some  $a < c(c-1)$ , to create a coloured graph equivalent to a graph  $\tilde{G}$  with colouring  $\tilde{\omega}$ , where  $\tilde{G}$  is a blow-up of a path on  $t'' = t' - (a(c-1) + c)$  vertices and  $\theta(\tilde{G}, \tilde{\omega}) < \theta(G, \omega)$ .<sup>(i)</sup> It is straightforward to verify that  $t'' \geq 2c^8(\theta(\tilde{G}, \tilde{\omega}) + 1)$  and so it follows from the inductive hypothesis that

$$m(\tilde{G}, \tilde{\omega}) \leq t'' - \left\lceil \frac{t''}{c} \right\rceil = t' - (a(c-1) + c) - \left\lceil \frac{t' - (a(c-1) + c)}{c} \right\rceil;$$

the required bound on  $m(G', \omega')$  then follows easily from the fact that  $m(G', \omega') \leq a(c-2) + (c-1) + m(\tilde{G}, \tilde{\omega})$ .

It remains to consider the case that we terminate after performing  $c(c-1)$  efficient flooding sequences. In this case, we have (after contracting monochromatic components) reduced the length of an end-to-end

---

<sup>(i)</sup> In fact the moves we have played might have flooded a larger component than is described here; but by Lemma 1.4 this can only help us.



path by  $c(c-1)^2$ , with at most  $c(c-1)(c-2)$  moves. Thus, by Proposition 3.1 we can flood the resulting graph with a further

$$t' - c(c-1)^2 - \left\lceil \frac{t' - c(c-1)^2}{c} \right\rceil + c - 1 = t' - \left\lceil \frac{t'}{c} \right\rceil - c(c-1)(c-2)$$

moves, from which the bound on  $m(G', \omega')$  follows immediately, completing the proof.  $\square$

## 4 Conclusions and Open Problems

We have given several upper and lower bounds on the maximum number of moves, taken over all possible colourings, that may be required to flood a given graph  $G$  in both the fixed and free variants of the game, and we have demonstrated that these bounds are tight for suitable families of trees.

Motivated by the intuition that adding a large number of edges to the graph should reduce the number of moves required in the worst case, we also demonstrate that the number of moves required in the worst case (in the free version) to flood a blow-up of a sufficiently long path is the same as the number required to flood a path of the same length. If we add all possible edges to a tree of radius  $r$  that do not reduce the length of a shortest path between any vertex  $u$  and the vertex  $v$  of minimum eccentricity, the resulting graph contains a blow-up of a path of length  $r+1$  as a subgraph, so our result on path blow-ups demonstrates that adding edges in this way gives a dramatic reduction in the worst-case number of moves required to flood the graph. It would be interesting to investigate precisely how many edges must be added to a tree of radius  $r$  to decrease the worst-case number of move required.

Our results provide a partial answer to a question raised by Meeks and Scott in [33], that of determining the maximum number of moves, taken over all possible colourings, that may be required to flood a given graph  $G$ . This general question remains open, and given the emphasis on  $k \times n$  grids in the existing algorithmic analysis of flood-filling games, a natural first direction for further research would be that of determining the exact value of  $M_c(G)$  and  $M_c^{(v)}(G)$  in the case that  $G$  is a  $k \times n$  grid. Using Proposition 2.7 and Lemma 2.9, we only have  $n - \lceil \frac{n}{c} \rceil$  and  $n - \lceil \frac{n}{c} \rceil + (c-1) \lceil \frac{k-1}{2} \rceil$  as lower and upper bounds in the free version, respectively; the bounds for the fixed version are even worse.

The parameterised complexity of determining whether a given coloured graph can be flooded with a specified number of moves has been studied with a wide range of different parameterisations (as in, for example, [18]), but this investigation of extremal properties gives rise to a new natural parameterised problem: given a graph  $G$ , for which we know the value of  $M_c(G)$ , and a colouring  $\omega$  of  $G$ , what is the (parameterised) complexity of determining whether  $G$  with initial colouring  $\omega$  can be flooded in at most  $M_c(G) - k$  moves, where  $k$  is the parameter? It can easily be deduced from the hardness proof for  $2 \times n$  boards in [33] that determining the minimum number of moves required to flood a blow-up of a path is NP-hard, so it is already meaningful to consider this parameterised problem for graphs drawn from the classes considered here.

## References

- [1] Flood it! available at <https://market.android.com>.
- [2] Flood it game. <http://floodit.appspot.com>.

- [3] Flood it! 2. available at <http://itunes.apple.com>.
- [4] Mad virus. <http://www.bubblebox.com/play/puzzle/539.htm>.
- [5] Flood-it. <http://unixpapa.com/floodit/>.
- [6] S. Arora and B. Barak. *Computational complexity: a modern approach*. Cambridge University Press, 2009.
- [7] D. Arthur, R. Clifford, M. Jalsenius, A. Montanaro, and B. Sach. The complexity of flood filling games. arXiv.1001.4420v1 [cs.DS], January 2010.
- [8] D. Arthur, R. Clifford, M. Jalsenius, A. Montanaro, and B. Sach. The complexity of flood filling games. In P. Boldi and L. Gargano, editors, *Fun with Algorithms, 5th International Conference, FUN 2010, Ischia, Italy, June 2-4, 2010. Proceedings*, volume 6099 of *Lecture Notes in Computer Science*, pages 307–318. Springer, 2010.
- [9] R. Belmonte, M. Ghadikolaei, M. Kiyomi, M. Lampis, and Y. Otachi. How bad is the freedom to flood-it? In *9th International Conference on Fun with Algorithms, FUN 2018, June 13-15, 2018, La Maddalena, Italy*, pages 5:1–5:13, 2018. doi: 10.4230/LIPIcs.FUN.2018.5.
- [10] B. Bollobás. *Modern Graph Theory*. Springer, 1998.
- [11] A. Born. Flash application for the computer game “*Biene*” (*Honey-Bee*). <http://www.ursulinen.asn-graz.ac.at/Bugs/html/games/biene.htm>, 2009.
- [12] R. Clifford, M. Jalsenius, A. Montanaro, and B. Sach. The complexity of flood filling games. *Theory of Computing Systems*, 50:72–92, 2012.
- [13] S. A. Cook. The complexity of theorem proving procedures. In *Proc. 3rd Ann. ACM Symp. Theory of Computing*, pages 151–158, 1971.
- [14] T. Cormen, C. Leiserson, and R. Rivest. *Introduction to Algorithms*. MIT Press and McGraw-Hill, 1990.
- [15] A. da Silva, L. Ochi, B. Barros, and R. Pinheiro. Efficient approaches for the Flooding Problem on graphs. *Annals of Operations Research*, pages doi:10.1007/s10479-018-2796-0, 2018.
- [16] U. dos Santos Souza, F. Protti, and M. Dantas da Silva. Parameterized complexity of flood-filling games on trees. In D.-Z. Du and G. Zhang, editors, *Computing and Combinatorics*, volume 7936 of *Lecture Notes in Computer Science*, pages 531–542. 2013.
- [17] R. G. Downey and M. R. Fellows. *Parameterized Complexity*. Springer-Verlag, 1999.
- [18] M. Fellows, U. Souza, F. Protti, and M. D. da Silva. Tractability and hardness of flood-filling games on trees. *Theoretical Computer Science*, 576:102–116, 2015. ISSN 0304-3975. doi: <http://dx.doi.org/10.1016/j.tcs.2015.02.008>. URL <http://www.sciencedirect.com/science/article/pii/S030439751500105X>.

- [19] M. Fellows, F. Protti, F. Rosamond, M. D. da Silva, and U. S. Souza. Algorithms, kernels and lower bounds for the flood-it game parameterized by the vertex cover number. *Discrete Applied Mathematics*, 245:94 – 100, 2018. ISSN 0166-218X. doi: <https://doi.org/10.1016/j.dam.2017.07.004>. URL <http://www.sciencedirect.com/science/article/pii/S0166218X17303116>.
- [20] M. Fellows, F. Rosamond, M. Dantas da Silva, and U. Souza. A survey on the complexity of flood-filling games. In H.-J. Böckenhauer, D. Komm, and W. Unger, editors, *Adventures Between Lower Bounds and Higher Altitudes: Essays Dedicated to Juraj Hromkovič on the Occasion of His 60th Birthday*, pages 357–376. Springer International Publishing, 2018. ISBN 978-3-319-98355-4. doi: 10.1007/978-3-319-98355-4\_20. URL [https://doi.org/10.1007/978-3-319-98355-4\\_20](https://doi.org/10.1007/978-3-319-98355-4_20).
- [21] R. Fleischer and G. Woeginger. An algorithmic analysis of the Honey-Bee Game. *Theoretical Computer Science*, 452:75–87, 2012.
- [22] J. Flum and M. Grohe. *Parameterised Complexity Theory*. Springer, 2006.
- [23] H. Fukui, A. Nakanishi, R. Uehara, T. Uno, and Y. Uno. The complexity of free flooding games. In *Information Processing Society of Japan (IPSJ) SIG Notes*, pages 1–5, August 2011.
- [24] H. Fukui, Y. Otachi, R. Uehara, T. Uno, and Y. Uno. On complexity of flooding games on graphs with interval representations. In J. Akiyama, M. Kano, and T. Sakai, editors, *Computational Geometry and Graphs – TJJCCGG 2012*, pages 73–84. Springer Berlin Heidelberg, 2013.
- [25] M. Garey and D. Johnson. *Computers and intractability: a guide to the theory of NP-completeness*. W. H. Freeman, 1979.
- [26] W.-K. Hon, T. Kloks, F.-H. Liu, H.-H. Liu, and H.-L. Wang. Flood-it on at-free graphs. arXiv:1511.01806 [cs.DM], 2015.
- [27] R. M. Karp. Reducibility among combinatorial problems. In R. E. Miller and J. W. Thatcher, editors, *Complexity of Computer Computations*. Plenum Press, New York, 1972.
- [28] A. Lagoutte. Jeux d’inondation dans les graphes. Technical report, ENS Lyon, HAL: hal-00509488, August 2010.
- [29] A. Lagoutte and S. Tavenas. The complexity of shortest common supersequence for inputs with no identical consecutive letters. arXiv:1309.0422 [cs.DM], 2013.
- [30] A. Lagoutte, M. Noulal, and E. Thierry. Flooding games on graphs. *Discrete Applied Mathematics*, 164:532–538, December 2014.
- [31] L. A. Levin. Universal sequential search problems. *PINFTRANS: Problems of Information Transmission*, 9, 1973.
- [32] K. Meeks and A. Scott. The complexity of flood-filling games on graphs. *Discrete Applied Mathematics*, 160:959–969, 2012.

- [33] K. Meeks and A. Scott. The complexity of Free-Flood-It on  $2 \times n$  boards. *Theoretical Computer Science*, 500:25–43, 2013.
- [34] K. Meeks and A. Scott. Spanning trees and the complexity of flood filling games. *Theory of Computing Systems*, 54:731–753, 2014.
- [35] R. Niedermeier. *Invitation to Fixed Parameter Algorithms*. Oxford University Press, 2006.
- [36] K.-J. Räihä and E. Ukkonen. The shortest common supersequence problem over binary alphabet is NP-complete. *Theoretical Computer Science*, 16:187–198, 1981.
- [37] U. Souza, F. Protti, and M. D. da Silva. An algorithmic analysis of flood-it and free-flood-it on graph powers. *Discrete Mathematics & Theoretical Computer Science*, 16(3):279–290, 2014. URL <http://www.dmtcs.org/dmtcs-ojs/index.php/dmtcs/article/view/2683>.