

Nonrepetitive Colouring via Entropy Compression

Vida Dujmović[§] Gwenaël Joret[†] Jakub Kozik[¶] David R. Wood[‡]

Abstract

A vertex colouring of a graph is *nonrepetitive* if there is no path whose first half receives the same sequence of colours as the second half. A graph is nonrepetitively k -choosable if given lists of at least k colours at each vertex, there is a nonrepetitive colouring such that each vertex is coloured from its own list. It is known that every graph with maximum degree Δ is $c\Delta^2$ -choosable, for some constant c . We prove this result with $c = 1$ (ignoring lower order terms). We then prove that every subdivision of a graph with sufficiently many division vertices per edge is nonrepetitively 5-choosable. The proofs of both these results are based on the Moser-Tardos entropy-compression method, and a recent extension by Grytczuk, Kozik and Micek for the nonrepetitive choosability of paths. Finally, we prove that every graph with pathwidth k is nonrepetitively $O(k^2)$ -colourable.

1 Introduction

A colouring of a graph¹ is *nonrepetitive* if there is no path P such that the first half of P receives the same sequence of colours as the second half of P . More precisely, a k -colouring

Date: 23 Dec 2011; revised 24 Feb 2012. *Mathematics Subject Classification.* 05C15, 05D40.

[§]School of Computer Science, Carleton University, Ottawa, Canada (vida@scs.carleton.ca). Supported by NSERC and an Endeavour Fellowship from the Australian Government.

[†]Département d'Informatique, Université Libre de Bruxelles, Brussels, Belgium (gjoret@ulb.ac.be). Postdoctoral Researcher of the Fonds National de la Recherche Scientifique (F.R.S.-FNRS). Supported by an Endeavour Fellowship from the Australian Government.

[¶]Theoretical Computer Science Department, Faculty of Mathematics and Computer Science, Jagiellonian University, ul. Prof. St. Łojasiewicza 6, 30-348 Krakow, Poland (jkozik@tcs.uj.edu.pl). Supported by the Polish National Science Center (DEC-2011/01/D/ST1/04412).

[‡]Department of Mathematics and Statistics, The University of Melbourne, Melbourne, Australia (woodd@unimelb.edu.au). Supported by a QEII Fellowship and a Discovery Project from the Australian Research Council.

¹We consider simple, finite, undirected graphs G with vertex set $V(G)$, edge set $E(G)$, maximum degree $\Delta(G)$, and maximum clique size $\omega(G)$. The *neighbourhood* of a vertex $v \in V(G)$ is $N_G(v) := \{w \in V(G) : vw \in E(G)\}$. The *neighbourhood* of a set $S \subseteq V(G)$ is $N_G(S) := \bigcup\{N_G(x) : x \in S\} \setminus S$. The degree of a vertex $v \in V(G)$ is $\deg_G(v) := |N_G(v)|$. We use $N(v)$ and $N(S)$ and $\deg(v)$ if the graph G is clear from the context.

of a graph G is a function ψ that assigns one of k colours to each vertex of G . A path is *even* if its order is even. An even path v_1, v_2, \dots, v_{2t} of G is *repetitively* coloured by ψ if $\psi(v_i) = \psi(v_{t+i})$ for all $i \in [1, t] := \{1, 2, \dots, t\}$. A colouring ψ is *nonrepetitive* if no path of G is repetitively coloured by ψ . The *nonrepetitive chromatic number* $\pi(G)$ is the minimum integer k such that G has a nonrepetitive k -colouring.

Observe that every nonrepetitive colouring is *proper*, in the sense that adjacent vertices receive distinct colours. Moreover, a nonrepetitive colouring has no 2-coloured P_4 (a path on four vertices). A proper colouring with no 2-coloured P_4 is called a *star colouring* since each bichromatic subgraph is a star forest; see [1, 7, 15, 18, 36, 41]. The *star chromatic number* $\chi_{\text{st}}(G)$ is the minimum number of colours in a star colouring of G . Thus

$$\chi(G) \leq \chi_{\text{st}}(G) \leq \pi(G) .$$

The seminal result in this field is by Thue [40], who in 1906 proved² that every path is nonrepetitively 3-colourable. Nonrepetitive colourings have recently been widely studied [2–4, 4–6, 8, 9, 12–14, 16, 19, 21, 23, 24, 26–28, 31–33, 37–39]; see the surveys [11, 20–22].

The contributions of this paper concern three different generalisations of the result of Thue: bounded degree graphs, graph subdivisions, and graphs of bounded pathwidth.

1.1 Bounded Degree

In a sweeping generalisation of Thue’s result, Alon et al. [3] proved³ that for some constant c and for every graph G with maximum degree $\Delta \geq 1$,

$$\pi(G) \leq c\Delta^2 . \tag{1}$$

Moreover, the bound in (1) is almost tight—Alon et al. [3] proved that there are graphs with maximum degree Δ that are not nonrepetitively $(c\Delta^2/\log \Delta)$ -colourable for some constant c .

The bound in (1), in fact, holds in the stronger setting of nonrepetitive list colourings. A *list assignment* of a graph G is a function L that assigns a set $L(v)$ of colours to each vertex $v \in V(G)$. Then G is *nonrepetitively L -colourable* if there is a nonrepetitive colouring of G , such that each vertex $v \in V(G)$ is assigned a colour in $L(v)$. And G is *nonrepetitively k -choosable* if for every list assignment L of G such that $|L(v)| \geq k$ for each vertex $v \in V(G)$,

²The nonrepetitive 3-colouring of P_n by Thue [40] is obtained as follows. Given a nonrepetitive sequence over $\{1, 2, 3\}$, replace each 1 by the sequence 12312, replace each 2 by the sequence 131232, and replace each 3 by the sequence 1323132. Thue [40] proved that the new sequence is nonrepetitive. Thus arbitrarily long paths can be nonrepetitively 3-coloured.

³The theorem of Alon et al. [3] was actually for edge colourings, but it is easily seen that the method works in the more general setting of vertex colourings.

there is a nonrepetitive L -colouring of G . The *nonrepetitive choice number* $\pi_{\text{ch}}(G)$ is the minimum integer k such that G is nonrepetitively k -choosable. By definition, $\pi(G) \leq \pi_{\text{ch}}(G)$.

All the known proofs of (1) are based on the Lovász Local Lemma, and thus are easily seen to prove the stronger result that

$$\pi_{\text{ch}}(G) \leq c \Delta(G)^2 . \quad (2)$$

Alon et al. [3] originally proved (2) with $c = 2e^{16}$, which was improved to 36 by Grytczuk [21] and then to 16 again by Grytczuk [20]. The current best bound, assuming $\Delta(G) \geq 2$, is $\pi_{\text{ch}}(G) \leq 12.92(\Delta(G) - 1)^2$ by Haranta and Jendrol' [26]. We improve the constant c to 1.

Theorem 1. *For every graph G with maximum degree Δ ,*

$$\pi_{\text{ch}}(G) \leq (1 + o(1))\Delta^2 .$$

The proof of Theorem 1 is based on the celebrated entropy-compression method of Moser and Tardos [35], and more precisely on an extension by Grytczuk et al. [23] for nonrepetitive sequences (or equivalently, nonrepetitive colourings of paths). The latter authors considered the following variant of the Moser-Tardos algorithm for nonrepetitively colouring paths. Start at the first vertex of the path and repeat the following step until a valid colouring is produced: Randomly colour the current vertex. If doing so creates a repetitively coloured subpath P , then uncolour the second half of P and let the new current vertex be the first uncoloured vertex on the path. Otherwise, go to the next vertex in the path. Grytczuk et al. [23] used this algorithm to obtain a short proof that paths are nonrepetitively 4-choosable, which was first proved by Grytczuk et al. [24] using the Lovász Local Lemma. (It is open whether every path is nonrepetitively 3-choosable.) Our proof of Theorem 1 generalises this method for graphs of bounded degree. While the main conclusion of the Moser-Tardos method was a constructive proof of the Lovász Local Lemma, as Kolipaka and Szegedy [29] write, “variants of the Moser-Tardos algorithm can be useful in existence proofs”. Our result is further evidence of this claim.

We expect that this method can also be used to make constant-factor improvements to other bounds proved using the Lovász Local Lemma, such as the bound of $\chi_{\text{st}}(G) \leq 20\Delta^{3/2}$ by Fertin et al. [15].

1.2 Subdivisions

A *subdivision* of a graph G is a graph obtained from G by replacing the edges of G with internally disjoint paths, where the path replacing vw has endpoints v and w . In a beautiful generalisation of Thue’s theorem, Pezarski and Zmarz [39] proved that every graph has a nonrepetitively 3-colourable subdivision (improving on analogous 5- and 4-colour results by

Grytczuk [20] and Barát and Wood [6] respectively). For each of these theorems, the number of division vertices per edge is $\mathcal{O}(n)$ or $\mathcal{O}(n^2)$ for n -vertex graphs. Improving these bounds, Nešetřil et al. [37] proved that every graph has a nonrepetitively 17-colourable subdivision with $\mathcal{O}(\log n)$ division vertices per edge, and that $\Omega(\log n)$ division vertices are needed on some edge of a nonrepetitively $\mathcal{O}(1)$ -colourable subdivision of K_n . Here we prove that every graph has a nonrepetitively $\mathcal{O}(1)$ -choosable subdivision, which solves an open problem by Grytczuk et al. [24]. All logarithms are binary.

Theorem 2. *Let G be a subdivision of a graph H , such that each edge $vw \in E(H)$ is subdivided at least $\lceil 10^5 \log(\deg(v) + 1) \rceil + \lceil 10^5 \log(\deg(w) + 1) \rceil + 2$ times in G . Then*

$$\pi_{\text{ch}}(G) \leq 5 .$$

Theorem 2 is stronger than the above subdivision results in the following respects: (1) it is for choosability not just colourability; (2) it applies to every subdivision with *at least* a certain number of division vertices per edge, and (3) the required number of division vertices per edge is asymptotically fewer than for the above results. Of course, Theorem 2 is weaker than the results in [6, 39] mentioned above in that the number of colours is 5.

Theorem 2 is also proved using the entropy-compression method mentioned above. An analogous theorem with more colours and $\mathcal{O}(\log \Delta(G))$ division vertices per edges can be proved using the Lovász Local Lemma; see Appendix A.

1.3 Pathwidth

Thue’s result was generalised in a different direction by Brešar et al. [8], who proved that every tree is nonrepetitively 4-colourable⁴. This result was further generalised by considering treewidth⁵, which is a parameter that measures how similar a graph is to a tree. Kündgen and Pelsmajer [31] and Barát and Varjú [4] independently proved that graphs of bounded treewidth have bounded nonrepetitive chromatic number. The best upper bound is due to Kündgen and Pelsmajer [31], who proved that $\pi(G) \leq 4^k$ for every graph G with treewidth k . The best lower bound is due to Al bertson et al. [1], who described graphs G with treewidth k and $\pi(G) \geq \chi_{\text{st}}(G) = \binom{k+2}{2}$. Thus there is a quadratic lower bound on π in

⁴No such result is possible for choosability—Fiorenzi et al. [16] proved that trees have arbitrarily high nonrepetitive choice number. On the other hand, Kozik and Micek [30] proved that $\pi_{\text{ch}}(T) \leq O(\Delta^{1+\epsilon})$ for every tree T with maximum degree Δ .

⁵A *tree decomposition* of a graph G consists of a tree T and a set $\{B_x \subseteq V(G) : x \in V(T)\}$ of subsets of vertices of G , called *bags*, indexed by the vertices of T , such that (1) the endpoints of each edge of G appear in some bag, and (2) for each vertex v of G the set $\{x \in V(T) : v \in B_x\}$ is nonempty and induces a connected subtree of T . The *width* of the tree decomposition is $\max\{|B_x| - 1 : x \in V(T)\}$. The *treewidth* of G is the minimum width of a tree decomposition of G . A *path decomposition* is a tree decomposition whose underlying tree is a path. Thus a path decomposition is simply defined by a sequence of bags B_1, \dots, B_p . The *pathwidth* of G is the minimum width of a path decomposition of G .

terms of treewidth. It is open whether π is bounded from above by a polynomial function of treewidth. We prove the following related result.

Theorem 3. *For every graph G with pathwidth k ,*

$$\pi(G) \leq 2k^2 + 6k + 1 .$$

It is open whether $\pi(G) \in \mathcal{O}(k)$ for every graph G with pathwidth k . For treewidth, a quadratic lower bound on π follows from the quadratic lower bound on χ_{st} , as explained above. However, we show that no such result holds for pathwidth.

Theorem 4. *For every graph G with pathwidth k ,*

$$\chi_{\text{st}}(G) \leq 3k + 1 .$$

2 An Algorithm

This section presents and analyses an algorithm for nonrepetitively list colouring a graph. This machinery will be used to prove Theorems 1 and 2 in the following sections.

If a set X is linearly ordered according to some fixed ordering, then the *index* of an element $e \in X$ in X is the position of e in this ordering of X . Such an ordering induces in a natural way an ordering of each subset Y of X , so that the index of an element $e \in Y$ in Y is well defined.

Let G be a fixed n -vertex graph. Assume that $V(G)$ is ordered according to some arbitrary linear ordering. Let L be a list assignment for G . Assume each list in L has size ℓ . Identify colours with nonnegative integers. Thus the colours in $L(v)$ are ordered in a natural way, for each $v \in V(G)$. Without loss of generality, the colour 0 is in none of these lists. In what follows, we consider an uncoloured vertex to be coloured 0. A *precolouring* of G is a colouring ψ of G such that $\psi(v) \in L(v) \cup \{0\}$ for each $v \in V(G)$. If $\psi(v) \neq 0$ then v is said to be *precoloured* by ψ .

For each path P of G with $2k$ vertices, for each subset $X \subseteq V(G) - V(P)$, and for each vertex $v \in V(P)$, define $s(P, X, v)$ to be the sequence (s_1, \dots, s_{2k}) obtained as follows: Let x, y be the two endpoints of P , with v closer to y than to x in P . Let v_1, \dots, v_p be the vertices of P from $v_1 := v$ to $v_p := x$ defined by P , in order. (Observe that $p \geq 2$ since $v \neq x$.) Let s_1 be the index of v_2 in $N(v_1) - X$, and for each $i \in [2, p - 1]$, let s_i be the index of v_{i+1} in $N(v_i) - (X \cup \{v_{i-1}\})$. Let $s_p := -1$. If $v = y$ then $p = 2k$ and the sequence (s_1, \dots, s_{2k}) is completely defined. If $v \neq y$ then let $q := 2k - p + 1$ and let w_1, \dots, w_q be the vertices of P from $w_1 := v$ to $w_q := y$ defined by P , in order. Let s_{p+1} be the index of w_2 in $N(w_1) - (X \cup \{v_2\})$, and for each $i \in [2, q - 1]$, let s_{p+i} be the index of w_{i+1} in $N(w_i) - (X \cup \{w_{i-1}\})$.

An important feature of the above encoding of the triple (P, X, v) as a sequence $s(P, X, v)$ is that it can be reversed, as we now explain.

Lemma 5. *Suppose $s = s(P, X, v)$ for some even path P of G such that $X \subseteq V(G) - V(P)$, and $v \in V(P)$. Then, given s , X , and v , one can uniquely determine the path P .*

Proof. Let $s = (s_1, \dots, s_{2k})$ and let $p \in [2, 2k]$ be the unique index such that $s_p = -1$. Let $u_p := v$, let u_{p-1} be the s_1 -th vertex in $N(u_p) - X$, and for $i = p - 2, \dots, 1$, let u_i be the s_{p-i} -th vertex in $N(u_{i+1}) - (X \cup \{u_{i+2}\})$. Next, for $j = p + 1, \dots, 2k$, let u_j be the s_j -th vertex in $N(u_{j-1}) - (X \cup \{u_{j-2}\})$. Then the vertices u_1, u_2, \dots, u_{2k} , in this order, determine a path P of G such that $s(P, X, v) = s$.

Observe that if P' is an even path of G such that $X \subseteq V(G) - V(P')$, $v \in V(P')$, and P' is distinct from P , then $s(P', X, v) \neq s(P, X, v)$. Therefore, the path P above is uniquely determined. \square

Let \mathcal{S} be the set of all sequences $s(P, X, v)$ where P is an even path in G , $X \subseteq V(G) - V(P)$, and v is a vertex of P . A *record* is a mapping $R : \mathbb{N} \rightarrow \mathcal{S} \cup \{\emptyset\}$. The *empty record* is the record R such that $R(i) = \emptyset$ for all $i \in \mathbb{N}$.

A *priority function* is a function f that associates to each nonempty subset X of $V(G)$ a vertex $f(X) \in X$. Consider Algorithm 1, which (for a fixed graph G , a list assignment L , a priority function f , and a precolouring ψ) takes as input a positive integer t and a vector $(c_1, \dots, c_t) \in [1, \ell]^t$. Note that precoloured vertices and a specific priority function will only be needed when proving the result on subdivisions. We thus invite the reader to first consider the set Q of precoloured vertices to be empty, and the priority function f to be arbitrary (for instance, $f(X)$ could be the first vertex in X in the fixed ordering of $V(G)$). Also note that the choice of the repetitively coloured path P in the algorithm is assumed to be consistent; that is, according to some (arbitrary) fixed deterministic rule.

Say that the algorithm *succeeds* if it terminates with $X = \emptyset$, and *fails* otherwise. It is easily seen that if the algorithm succeeds, then the produced colouring ϕ is a nonrepetitive L -colouring of G . For a given integer $t \geq 1$, let \mathcal{F}_t be the set of vectors $(c_1, \dots, c_t) \in [1, \ell]^t$ on which the algorithm fails. Let \mathcal{A}_t be the set of distinct pairs (ϕ, R) that are produced by the algorithm on vectors in \mathcal{F}_t . Let \mathcal{R}_t be the set of distinct records R that can be produced by the algorithm on vectors in \mathcal{F}_t . For $R \in \mathcal{R}_t$, let $\mathcal{F}_{t,R}$ be the set of vectors $(c_1, \dots, c_t) \in \mathcal{F}_t$ on which the algorithm produces record R . (Thus $\mathcal{F}_{t,R} \neq \emptyset$.) For a vector $(c_1, \dots, c_t) \in \mathcal{F}_t$, let the *trace* $\text{tr}(c_1, \dots, c_t)$ be the vector (X_1, \dots, X_t) where X_i is the set X at the beginning of the i -th iteration of the while-loop of the algorithm on input (c_1, \dots, c_t) , for each $i \in [1, t]$. (Observe that X_1 always equals $V(G) - Q$.) Finally, let $\mathcal{T}_t := \{\text{tr}(c_1, \dots, c_t) : (c_1, \dots, c_t) \in \mathcal{F}_t\}$.

The next lemma shows that for a fixed record $R \in \mathcal{R}_t$, all the vectors in $\mathcal{F}_{t,R}$ have the same

Algorithm 1 L -colouring the graph G , where f is a priority function, ψ is a precolouring of G , and Q is the set of precoloured vertices under ψ .

Input: $(c_1, \dots, c_t) \in [1, \ell]^t$

Output: a (possibly invalid) colouring ϕ and a record R

```

 $i \leftarrow 1$ 
 $\phi \leftarrow \psi$ 
 $R \leftarrow$  empty record
 $X \leftarrow V(G) - Q$ 
while  $i \leq t$  and  $X \neq \emptyset$  do
     $v \leftarrow f(X)$ 
     $\phi(v) \leftarrow$   $c_i$ -th colour in  $L(v)$ 
    if  $G$  contains a repetitively coloured path  $P$  then
        divide  $P$  into first half  $P_1$  and second half  $P_2$  so that  $v \in V(P_2)$ 
        for  $w \in V(P_2) - Q$  do
             $\phi(w) \leftarrow 0$ 
        end for
         $R(i) \leftarrow s(P, X, v)$ 
         $X \leftarrow X \cup (V(P_2) - Q)$ 
    else
         $R(i) \leftarrow \emptyset$ 
         $X \leftarrow X - \{v\}$ 
    end if
     $i \leftarrow i + 1$ 
end while
return return  $\phi, R$ 

```

trace.

Lemma 6. *For every $t \geq 1$ there exists a function $h_t : \mathcal{R}_t \rightarrow \mathcal{T}_t$ such that for each $R \in \mathcal{R}_t$ and each $(c_1, \dots, c_t) \in \mathcal{F}_{t,R}$ we have $\text{tr}(c_1, \dots, c_t) = h_t(R)$.*

Proof. We construct h_t by induction on t . For $t = 1$ simply let $h_t(R) := (V(G) - Q)$ for each $R \in \mathcal{R}_t$.

Now assume that $t \geq 2$. Let $R \in \mathcal{R}_t$. Let R' be the record obtained from R by setting $R'(i) := R(i)$ for each $i \in \mathbb{N}$ such that $i \neq t$, and $R'(t) := \emptyset$. Then $R' \in \mathcal{R}_{t-1}$, and by induction, $h_{t-1}(R') = (X_1, \dots, X_{t-1})$ for some $(X_1, \dots, X_{t-1}) \in \mathcal{T}_{t-1}$. Let $v_{t-1} := f(X_{t-1})$.

First suppose that $R(t-1) = \emptyset$. Let $X_t := X_{t-1} - \{v_{t-1}\}$ and $h_t(R) := (X_1, \dots, X_{t-1}, X_t)$. Consider a vector $(c_1, \dots, c_t) \in \mathcal{F}_{t,R}$. Then $(c_1, \dots, c_{t-1}) \in \mathcal{F}_{t-1,R'}$, and by induction $\text{tr}(c_1, \dots, c_{t-1}) = (X_1, \dots, X_{t-1})$. Thus at the beginning of the $(t-1)$ -th iteration of the while-loop in the algorithm on input (c_1, \dots, c_t) , the current record is R' , and $v = v_{t-1}$ and

$X = X_{t-1}$. Since $R(t-1) = \emptyset$, the algorithm subsequently coloured v_{t-1} without creating any repetitively coloured path, implying that $X = X_{t-1} - \{v_{t-1}\} = X_t$ at the beginning of the t -th iteration. Hence $\text{tr}(c_1, \dots, c_t) = (X_1, \dots, X_{t-1}, X_t) = h_t(R)$, as desired.

Now assume that $R(t-1) = s$ for some $s \in \mathcal{S}$. Using Lemma 5, let P be the path of G determined by s , X_{t-1} , and v_{t-1} . Let P_1 and P_2 denote the two halves of P , so that $v_{t-1} \in V(P_2)$. Let $X_t := X_{t-1} \cup (V(P_2) - Q)$ and $h_t(R) := (X_1, \dots, X_{t-1}, X_t)$. Consider a vector $(c_1, \dots, c_t) \in \mathcal{F}_{t,R}$. Then $(c_1, \dots, c_{t-1}) \in \mathcal{F}_{t-1,R'}$, and by induction $\text{tr}(c_1, \dots, c_{t-1}) = (X_1, \dots, X_{t-1})$. As before, at the beginning of the $(t-1)$ -th iteration of the while-loop in the algorithm on input (c_1, \dots, c_t) , the current record is R' , and $v = v_{t-1}$ and $X = X_{t-1}$. Then, after colouring v , the path P is repetitively coloured, and all vertices in P_2 are subsequently uncoloured, except for those in Q . Hence we have $X = X_{t-1} \cup (V(P_2) - Q) = X_t$ at the beginning of the t -th iteration. Therefore $\text{tr}(c_1, \dots, c_t) = (X_1, \dots, X_{t-1}, X_t) = h_t(R)$. \square

Lemma 7. *For every $(\phi, R) \in \mathcal{A}_t$ there is a unique vector $(c_1, \dots, c_t) \in \mathcal{F}_t$ such that the algorithm produces (ϕ, R) on input (c_1, \dots, c_t) .*

Proof. The proof is by induction on t . This claim is true for $t = 1$ since in that case the unique vector $(c_1) \in \mathcal{F}_1$ yielding (ϕ, R) is the one where c_1 is the index of colour $\phi(v_1)$ in the list $L(v_1)$, where $v_1 := f(V(G) - Q)$.

Now assume that $t \geq 2$. Let $(X_1, \dots, X_t) := h_t(R)$, where h_t is the function in Lemma 6. Let $v_t := f(X_t)$. (Recall that $X_t \neq \emptyset$.) Let R' be the record obtained from R by setting $R'(i) := R(i)$ for each $i \in \mathbb{N}$ such that $i \neq t$, and $R'(t) := \emptyset$. Then $R' \in \mathcal{R}_{t-1}$, and $h_{t-1}(R') = (X_1, \dots, X_{t-1})$.

First suppose that $R(t) = \emptyset$. Let ϕ' be the colouring obtained from ϕ by setting $\phi'(v_t) := 0$ and $\phi'(w) := \phi(w)$ for each $w \in V(G) - \{v_t\}$. Then $(\phi', R') \in \mathcal{A}_{t-1}$, and by induction there is a unique input vector $(c'_1, \dots, c'_{t-1}) \in \mathcal{F}_{t-1}$ for which the algorithm produces (ϕ', R') . It follows that every vector $(c_1, \dots, c_t) \in \mathcal{F}_t$ resulting in the pair (ϕ, R) satisfies $c_i = c'_i$ for each $i \in [1, t-1]$. But then c_t is also uniquely determined, since it is the index of colour $\phi(v_t)$ in the list $L(v_t)$. Hence there is a unique such vector (c_1, \dots, c_t) .

Now assume that $R(t) = s$ for some $s \in \mathcal{S}$. Using Lemma 5, let P be the path of G determined by s , X_t , and v_t . Let P_1 and P_2 denote the two halves of P , so that $v_t \in V(P_2)$. Let w_1, \dots, w_{2k} denotes the vertices of P , in order, so that $V(P_1) = \{w_1, \dots, w_k\}$ and $V(P_2) = \{w_{k+1}, \dots, w_{2k}\}$. Let $j \in [1, k]$ be the index such that $w_{k+j} = v_t$. Let ϕ' be the colouring obtained from ϕ by setting $\phi'(v_t) := 0$, $\phi'(w_{k+i}) := \phi(w_i)$ for each $i \in [1, k]$ such that $i \neq j$ and $w_{k+i} \notin Q$, and $\phi'(w) := \phi(w)$ for each $w \in V(G) - (V(P_2) - Q)$. Then $(\phi', R') \in \mathcal{A}_{t-1}$, and by induction there is a unique vector $(c'_1, \dots, c'_{t-1}) \in \mathcal{F}_{t-1}$ on the input of which the algorithm produces (ϕ', R') . It follows that every vector $(c_1, \dots, c_t) \in \mathcal{F}_t$ resulting in the pair (ϕ, R) satisfies $c_i = c'_i$ for each $i \in [1, t-1]$. Moreover, c_t is the index of colour $\phi(w_j)$ in the list $L(w_j)$, and therefore is also uniquely determined. \square

Lemma 7 implies that $|\mathcal{A}_t| = |\mathcal{F}_t|$ for all $t \geq 1$.

Recall that \mathcal{S} is the set of all sequences $s(P, X, v)$ where P is an even path in G , $X \subseteq V(G) - V(P)$, and $v \in V(P)$. Once we fix a precolouring ψ of G and a priority function f , as we did above, some triples (P, X, v) will never be considered by the algorithm on any input. (For instance, this is the case if X contains a precoloured vertex.) This leads us to define $\bar{\mathcal{S}}$ as the set of sequences $s \in \mathcal{S}$ such that $R(i) = s$ for some $t \geq 1$, $R \in \mathcal{R}_t$, and $i \in [1, t]$; such sequences are said to be *realisable* (with respect to ψ and f).

For each $k \in [1, \lfloor \frac{n}{2} \rfloor]$, let α_k be the number of sequences of length $2k$ in $\bar{\mathcal{S}}$. Define

$$\beta := \max\{1, \max\{(\alpha_k)^{1/k} : 1 \leq k \leq \lfloor \frac{n}{2} \rfloor\}\} .$$

Thus $\beta \geq 1$ and $\alpha_k \leq \beta^k$ for each $k \in [1, \lfloor \frac{n}{2} \rfloor]$.

Recall that a *Dyck word* of length $2t$ is a binary sequence with t zeroes and t ones such that the number of zeroes is at least the number of ones in every prefix of the sequence.

Let $R \in \mathcal{R}_t$. For each $i \in [1, t]$, let r_i be half the length of the sequence $R(i)$ (in particular $r_i = 0$ if $R(i) = \emptyset$), and let $z(R) := \sum_{i=1}^t r_i$. Then $z(R)$ is exactly the number of coloured vertices at the end of any execution of the algorithm that produces the record R . (Recall that a vertex of colour 0 is interpreted as being uncoloured.) In particular, $z(R) \geq 1$, since there is always at least one coloured vertex, and $z(R) \leq n$. Associate with R the word

$$D(R) := 01^{r_1}01^{r_2} \dots 01^{r_t}1^{z(R)} .$$

Then $D(R)$ is a Dyck word of length $2t$.

Conversely, a Dyck word d is *realisable* if there exist $t \geq 1$ and $R \in \mathcal{R}_t$ such that $D(R) = d$. The set of realisable Dyck words of length $2t$ is denoted \mathcal{D}_t .

3 Bounded Degree Proof

The proof of Theorem 1 makes use of the symbolic approach to combinatorial enumeration via generating functions. We refer the reader to the book by Flajolet and Sedgewick [17] for background on this topic, as well as for undefined terms and notations.

First we introduce a result from [17] that will be used in our proof.

Definition 1 (Definition IV.5 from [17]). Let $B(z)$ be a function analytic at 0. Then B admits a span of d if for some $r \in \mathbb{N}$,

$$\{n \in \mathbb{N} : [z^n]B(z) \neq 0\} \subseteq r + d\mathbb{N} .$$

If the largest span that $B(z)$ admits is 1, then B is *aperiodic*.

Definition 2 (Definition VII.3 from [17]). Let $B(z)$ be a function analytic at 0. It is said to belong to *the smooth inverse-function schema* if there exists a function $\phi(u)$ analytic at 0, such that in a neighbourhood of 0,

$$B(z) = z\phi(B(z)),$$

and $\phi(u)$ satisfies the following conditions:

(H1) The function $\phi(u)$ is such that

$$\phi(0) \neq 0, \quad [u^n]\phi(u) \geq 0, \quad \phi(u) \neq \phi_0 + \phi_1 u.$$

(H2) Within the open disc of convergence of ϕ at 0, $|u| < R$, there exists a (necessarily unique) positive solution $\tau \in (0, R)$ to the characteristic equation $\phi(\tau) - \tau\phi'(\tau) = 0$.

The schema is said to be *aperiodic* if $\phi(u)$ is an aperiodic function of u .

Theorem 8 (Theorem VII.2 from [17]). *Let $B(z)$ belong to the aperiodic smooth inverse-function schema $B(z) = z\phi(B(z))$. Let τ be the positive root of the characteristic equation and let $\rho := \tau/\phi(\tau)$. Then*

$$[z^t]B(z) \sim \sqrt{\frac{\phi(\tau)}{2\phi'(\tau)} \frac{\rho^{-t}}{\sqrt{\pi t^3}}}.$$

A *substring* of some sequence or word is a subsequence of consecutive elements.

The next result is a precise version of Theorem 1. Note that we do not attempt to optimise the lower order terms.

Theorem 9. *For every graph G with maximum degree $\Delta > 1$,*

$$\pi_{\text{ch}}(G) \leq \left\lceil \left(1 + \frac{1}{\Delta^{1/3}-1} + \frac{1}{\Delta^{1/3}}\right) \Delta^2 \right\rceil.$$

Proof. Let G be a graph with maximum degree Δ . Fix an ordering of $V(G)$. Let $n := |V(G)|$ and let L be a list assignment of G . Assume each list in L has size $\ell := \left\lceil \left(1 + \frac{1}{\Delta^{1/3}-1} + \frac{1}{\Delta^{1/3}}\right) \Delta^2 \right\rceil$. Let f be an arbitrary priority function. Consider the algorithm on G , where none of the vertices of G are precoloured (thus $Q = \emptyset$).

We will prove that $|\mathcal{A}_t| = o\left(\left(1 + \frac{1}{\Delta^{1/3}-1} + \frac{1}{\Delta^{1/3}}\right)^t \Delta^{2t}\right)$. It suffices to show that $|\mathcal{R}_t| = o\left(\left(1 + \frac{1}{\Delta^{1/3}-1} + \frac{1}{\Delta^{1/3}}\right)^t \Delta^{2t}\right)$ since the number of distinct colourings that can be produced by the algorithm is at most $(\ell + 1)^n$ (taking into account the extra colour 0).

Let $s \in \bar{\mathcal{S}}$ and suppose $s = (s_1, \dots, s_{2k})$. Observe, that $s_j \neq \Delta$ for each $j \in [2, 2k]$. Also, there is a unique index $p \in [k+1, 2k]$ such that $s_p = -1$. Thus $s_1 \in [1, \Delta]$, $s_p = -1$, and $s_j \in [1, \Delta-1]$ for each $j \in [1, 2k] \setminus \{1, p\}$. Hence there are at most $k\Delta(\Delta-1)^{2(k-1)}$ sequences of length $2k$ in $\bar{\mathcal{S}}$. Hence $\alpha_k < k\Delta^{2k-1}$.

Let $d = (d_1, \dots, d_{2t})$ be a realisable Dyck word of length $2t$. Suppose that d has the form $0^{l_1}1^{k_1}0^{l_2}1^{k_2} \dots 0^{l_q}1^{k_q}1$, for some positive $q, l_1, \dots, l_q, k_1, \dots, k_q$. Note that $\sum_{j=1}^q k_j = t-1$. Associate with the word d a weight $w(d) := k_1 k_2 \dots k_q \Delta^{-q}$. Clearly, for every $i \in [0, k_q]$ the number of distinct records $R \in \mathcal{R}_t$ with $z(R) = i+1$ such that $D(R) = d$ does not exceed

$$\begin{aligned} \alpha_{k_1} \cdots \alpha_{k_{q-1}} \alpha_{k_q-i} &< k_1 \Delta^{2k_1-1} \cdots k_{q-1} \Delta^{2k_{q-1}-1} (k_q - i) \Delta^{2(k_q-i)-1} \\ &\leq k_1 \Delta^{2k_1-1} \cdots k_q \Delta^{2k_q-1} \\ &= w(d) \Delta^{2t} . \end{aligned}$$

Therefore

$$|\mathcal{R}_t| < n \cdot \Delta^{2t} \cdot \sum_{d \in \mathcal{D}_t} w(d) .$$

Claim 10.

$$\sum_{d \in \mathcal{D}_t} w(d) = o\left(\left(1 + \frac{1}{\Delta^{1/3}-1} + \frac{1}{\Delta^{1/3}}\right)^t\right)$$

Proof. Let D' be the set of words on the alphabet $\{0, 1, 2\}$ that

- do not contain substrings 21 and 02,
- in every nonempty prefix the number of nonzero elements is strictly less than the number of zeroes, and
- the number of ones and twos in the whole word is one less than the number of zeroes.

Let γ be the function that, given a word in D' , replaces each 2 by 1 and appends 1. Observe that the image of γ is a Dyck word. Let d be a realisable Dyck word of length $2t$. Then for every proper, nonempty prefix of d , the number of ones is strictly less than the number of zeroes. In particular, d belongs to the image of γ . We are interested in the size of the preimage of d . Suppose that d has the form $0^{l_1}1^{k_1}0^{l_2}1^{k_2} \dots 0^{l_q}1^{k_q}1$, for some positive $q, l_1, \dots, l_q, k_1, \dots, k_q$. By the definition of γ , the elements of the preimage of d are exactly the words of the form $0^{l_1}1^{a_1}2^{b_1}0^{l_2}1^{a_2}2^{b_2} \dots 0^{l_q}1^{a_q}2^{b_q}$ where for every $i \in [1, q]$ we have $a_i + b_i = k_i$ and $a_i > 0$ and $b_i \geq 0$. Hence the size of this preimage is $k_1 k_2 \dots k_q$, which equals $w(d) \Delta^q$. Moreover, every element of the preimage of d has t zeroes and exactly q substrings 01. Let $F_{t,q}$ be the number of words from D' with exactly t zeroes and exactly q substrings 01. It follows from the above observations that

$$\sum_{d \in \mathcal{D}_t} w(d) \leq \sum_{q=0}^{\infty} F_{t,q} \Delta^{-q} .$$

Define the formal power series

$$F(z, y) := \sum_{t=0}^{\infty} \sum_{q=0}^{\infty} F_{t,q} z^t y^q .$$

Then

$$B(z) := F(z, \Delta^{-1}) = \sum_{t=0}^{\infty} z^t \left(\sum_{q=0}^{\infty} F_{t,q} \Delta^{-q} \right) .$$

Hence

$$\sum_{d \in \mathcal{D}_t} w(d) \leq [z^t] B(z) .$$

We now derive a functional equation defining $F(z, y)$ by decomposing elements of D' recursively along the last sequence of nonzero letters. For $(d_1, \dots, d_{2t-1}) \in D'$ we say that *position j visits level k* if the number of zeroes in (d_1, \dots, d_j) exceeds the number of nonzero symbols by k . The sequence (0) is the unique sequence in D' that contains only one zero. Every other sequence $d = (d_1, \dots, d_{2t-1})$ from D' with t zeroes that ends with exactly $p > 0$ nonzero symbols can be uniquely decomposed into p sequences $\delta_1, \dots, \delta_p \in D'$, of join length $2t - p - 2$, and the remaining sequence of the form $01^a 2^b$ with $a + b = p$ and $a > 0$ and $b \geq 0$. Precisely, the sequence δ_i of this decomposition is the substring of (d_1, \dots, d_{2t-p-2}) from the next position after the last position that visits level $i - 1$ (or from the beginning, when $i = 1$) to the last position that visits level i .

Let $\text{SEQ}(D')$ denote the set of finite sequences of sequences from D' . Let $\text{SEQ}_{\geq 1}(D')$ denote the set of nonempty finite sequences of sequences from D' . Let

$$D'' := \{(0)\} \times \text{SEQ}_{\geq 1}(D') \times \text{SEQ}(D') .$$

Let h be the function that maps a sequence $d = (d_1, \dots, d_{2t-1}) \in D' \setminus \{(0)\}$ to the triple $((0), (\delta_1, \dots, \delta_a), (\delta_{a+1}, \dots, \delta_{a+b}))$, where a, b , and the δ_i 's are defined as above. Observe that h is a bijection between $D' \setminus \{(0)\}$ and D'' .

Let $C_{t,q}$ be the number of elements of D'' with t zeroes and q substrings 01 . Define the formal power series

$$C(z, y) := \sum_{t=0}^{\infty} \sum_{q=0}^{\infty} C_{t,q} z^t y^q .$$

Observe that d and $h(d)$ have the same number of zeroes, for every $d \in D' \setminus \{(0)\}$. (Indeed, this is the reason for the leading (0) in the definition of D'' .) Also, the total number of occurrences of substring 01 in $h(d)$ is one less than in d . Thus $F_{t,q} = C_{t,q-1}$ for every $t \geq 1$, and $F(z, y) - z = y \cdot C(z, y)$. On the other hand, it follows from the definition of D'' that

$$C(z, y) = z \left(\sum_{i \geq 1} F(z, y)^i \right) \left(\sum_{i \geq 0} F(z, y)^i \right) = z \left(\frac{F(z, y)}{1 - F(z, y)} \right) \left(\frac{1}{1 - F(z, y)} \right) .$$

This justifies that $F(z, y)$ satisfies the following equation:

$$F(z, y) = z + zy \frac{F(z, y)}{(1 - F(z, y))^2} .$$

In particular,

$$B(z) = z \left(1 + \Delta^{-1} \frac{B(z)}{(1 - B(z))^2} \right) .$$

Hence $B(z)$ belongs to the smooth inverse-function schema with the function $\phi(u) = 1 + \Delta^{-1}u/(1 - u)^2$. It is straightforward to check that ϕ satisfies conditions (H1) and (H2) of Definition 2, and that it is aperiodic. By Theorem 8,

$$[z^t]B(z) = o \left(\left(\frac{\phi(\tau)}{\tau} \right)^t \right) ,$$

where τ is the unique solution of $\phi(\tau) - \tau\phi'(\tau) = 0$ belonging to $(0, 1)$. This also means that τ minimizes $\phi(u)/u$ in the interval $(0, 1)$. Hence for any $\tau' \in (0, 1)$ we have $[z^t]B(z) = o((\phi(\tau')/\tau')^t)$. If we choose $\tau' := 1 - \Delta^{-1/3}$ then

$$\frac{\phi(\tau')}{\tau'} = 1 + \frac{1}{\Delta^{1/3} - 1} + \frac{1}{\Delta^{1/3}}$$

and we obtain

$$[z^t]B(z) = o \left(\left(1 + \frac{1}{\Delta^{1/3} - 1} + \frac{1}{\Delta^{1/3}} \right)^t \right) .$$

This completes the proof of the claim. \square

Returning to the proof of Theorem 9, Claim 10 implies

$$|\mathcal{R}_t| = o \left(\left(1 + \frac{1}{\Delta^{1/3} - 1} + \frac{1}{\Delta^{1/3}} \right)^t \Delta^{2t} \right) .$$

Thus, if t is large enough, then $|\mathcal{A}_t|$ is strictly smaller than ℓ^t , implying that there is at least one vector (c_1, \dots, c_t) among the ℓ^t vectors in $[1, \ell]^t$ on which the algorithm succeeds. Therefore G admits a nonrepetitive L -colouring. \square

4 Subdivision Proof

We now begin the proof of Theorem 2. A sequence (s_1, \dots, s_q) of positive integers is *c-spread* if each entry equal to 1 can be mapped to an entry greater than 1 such that for each $i \in [1, q]$ such that $s_i \geq 2$, there are at least $\lceil c \log s_i \rceil$ entries, either all immediately before s_i or all immediately after s_i , that are equal to 1 and are mapped to s_i .

Lemma 11. Fix $\varepsilon > 0$. Let $w := (1 + \varepsilon)^{-1/2} < 1$. Let $c \in \mathbb{N}$ be such that $2^{2/c} \leq 1 + \varepsilon$ and $w^c \leq \frac{\varepsilon}{2}(1 - w)$. Then for each $q \geq 1$ the number of distinct c -spread sequences of length q is at most $(1 + \varepsilon)^q$.

Proof. The proof is by induction on q . Let $f(q)$ be the number of c -spread sequences of length q . The claim holds when $q \leq c$ since the length- q sequence $(1, \dots, 1)$ is the only c -spread sequence of length q in that case.

Now assume that $q \geq c + 1$. Here are three ways of obtaining c -spread sequences of length q from smaller ones:

1. If (s_1, \dots, s_{q-1}) is c -spread then so is $(1, s_1, \dots, s_{q-1})$.
2. If $r \in \mathbb{N}$ such that $r \geq 2$ and $\lceil c \log r \rceil = q - 1$ then the two length- q sequences $(1, \dots, 1, r)$ and $(r, 1, \dots, 1)$ are c -spread.
3. If $r \in \mathbb{N}$ such that $r \geq 2$ and $z := \lceil c \log r \rceil \leq q - 2$, and if (s_1, \dots, s_{q-z-1}) is a c -spread sequence, then the two length- q sequences $(1, \dots, 1, r, s_1, \dots, s_{q-z-1})$ and $(r, 1, \dots, 1, s_1, \dots, s_{q-z-1})$ are c -spread.

It is not difficult to see that each c -spread sequence of length q can be obtained using the three constructions above. Notice that if $z, r \in \mathbb{N}$ are such that $r \geq 2$ and $z = \lceil c \log r \rceil$, then in particular $z \geq c$ and $r \leq 2^{z/c}$. Letting $f(0) := 1$, we deduce that

$$\begin{aligned}
f(q) &\leq f(q-1) + 2 \sum_{z=c}^{q-1} 2^{z/c} f(q-z-1) \\
&\leq (1 + \varepsilon)^{q-1} + 2 \sum_{z=c}^{q-1} (1 + \varepsilon)^{z/2} (1 + \varepsilon)^{q-z-1} \\
&= (1 + \varepsilon)^{q-1} + 2(1 + \varepsilon)^{q-1} \sum_{z=c}^{q-1} w^z \\
&\leq (1 + \varepsilon)^{q-1} + 2(1 + \varepsilon)^{q-1} \sum_{z=c}^{\infty} w^z \\
&= (1 + \varepsilon)^{q-1} + 2(1 + \varepsilon)^{q-1} \frac{w^c}{1 - w} \\
&\leq (1 + \varepsilon)^{q-1} + \varepsilon(1 + \varepsilon)^{q-1} \\
&= (1 + \varepsilon)^q.
\end{aligned}$$

□

A Dyck word d is said to be *special* if d does not contain 0110110 as a substring. The following crude upper bound on the number of such words will be used in our proof of Theorem 2.

Lemma 12. *The number of special Dyck words of length $2t$ is at most 3.992^{t+1} .*

Proof. For $k \geq 1$, let $g(k)$ be the number of binary words not containing 0110110. Let $\gamma := (2^7 - 1)^{1/7}$. Then $g(k) \leq 2^k \leq \gamma^{k+1}$ for $k \in [1, 7]$, and $g(k) \leq \gamma^7 \cdot g(k-7)$ for $k \geq 8$, since such binary words cannot start with 0110110. Thus $g(k) \leq \gamma^{k+1}$ for all $k \geq 1$. Since every special Dyck word of length $2t$ is a binary word not containing 0110110, it follows that the number of such Dyck words is at most $\gamma^{2t+1} < 3.992^{t+1}$. \square

Theorem 2. Let G be a subdivision of a graph H , such that each edge $vw \in E(H)$ is subdivided at least $\lceil 10^5 \log(\deg(v) + 1) \rceil + \lceil 10^5 \log(\deg(w) + 1) \rceil + 2$ times in G . Then $\pi_{\text{ch}}(G) \leq 5$.

Proof. Let $n := |V(G)|$. Let $L'(v)$ denote a list of available colours for each vertex $v \in V(G)$, and assume all these lists have size 5. Let ψ be an arbitrary precolouring of G with precoloured set $Q := V(H)$ and with $\psi(v) \in L'(v)$ for each $v \in Q$. Fix an ordering of $V(G)$ such that $V(G) - Q$ precedes Q .

Let $c := 10^5$. For each $v \in Q$, let $g(v) := \lceil c \log(\deg(v) + 1) \rceil$ and let $M(v)$ be the set of vertices of G at distance at most $g(v) + 1$ from v . Thus $M(v) \cap M(w) = \emptyset$ for distinct vertices $v, w \in Q$; we say that $u \in V(G) - Q$ belongs to $v \in Q$ if $u \in M(v)$.

For each edge $vw \in E(H)$, let P_{vw} denote the path of G induced by the subdivision vertices introduced on the edge vw in G . Note that $v, w \notin V(P_{vw})$. A set $X \subseteq V(G) - Q$ is *nice* if $X \neq \emptyset$ and, for each edge $vw \in E(H)$, the graph $P_{vw} - X$ is either connected or empty. The *boundary* $\partial(X)$ of a nice set X is the set of vertices $y \in X$ such that $X - \{y\}$ is either nice or empty. Observe that $\partial(X)$ is always nonempty.

Fix an arbitrary ordering of the edges in $E(H)$. For each edge $vw \in E(H)$, orient the path P_{vw} from an arbitrarily chosen endpoint to the other. If Y is a set of consecutive vertices of a path P_{vw} and $x \in V(P_{vw}) - Y$, then x is either *before* Y or *after* Y , depending on the orientation of P_{vw} .

Let f be a priority function defined as follows: For every nice set X , let vw be the first edge in the ordering of $E(H)$ such that $V(P_{vw}) \cap X \neq \emptyset$. If $V(P_{vw}) \subseteq X$, then $V(P_{vw}) \subseteq \partial(X)$, and we let $f(X)$ be an arbitrary vertex in $V(P_{vw})$. If $V(P_{vw}) - X \neq \emptyset$ and there is a vertex $x \in \partial(X) \cap V(P_{vw})$ before $V(P_{vw}) - X$ on P_{vw} , then x is uniquely determined, and we let $f(X) := x$. If $V(P_{vw}) - X \neq \emptyset$ but there is no such vertex x , then we let $f(X)$ be the unique vertex in $\partial(X) \cap V(P_{vw})$ that is after $V(P_{vw}) - X$ on P_{vw} .

For each $u \in V(G) - Q$, let $L(u)$ be the list $L'(u) - \{\phi(v)\}$ if u belongs to $v \in Q$ and $\phi(v) \in L'(u)$, otherwise let $L(u)$ be obtained from $L'(u)$ by removing one arbitrary colour from $L'(u)$. This defines a list $L(u)$ of available colours for each vertex $u \in V(G) - Q$, and all

these lists have size $\ell := 4$. Consider the algorithm on G with the latter lists, with priority function f , and with precolouring ψ . By the definition of the lists $L(u)$, if the algorithm succeeds on some input $(c_1, \dots, c_t) \in [1, \ell]^t$ then it produces a nonrepetitive L' -colouring of G .

Claim 13. *Let $t \geq 1$. Then for each vector $(c_1, \dots, c_t) \in \mathcal{F}_t$, all the sets appearing in the trace $\text{tr}(c_1, \dots, c_t)$ are nice.*

Proof. The proof is by induction on t . The claim is true for $t = 1$ since $X_1 = V(G) - Q$ is nice. Now assume that $t \geq 2$. Let $(c_1, \dots, c_t) \in \mathcal{F}_t$ and let $\text{tr}(c_1, \dots, c_t) = (X_1, \dots, X_t)$. Then $(c_1, \dots, c_{t-1}) \in \mathcal{F}_{t-1}$, and by induction the sets X_1, \dots, X_{t-1} are nice. Let $v_{t-1} := f(X_{t-1})$.

First suppose that $R(t-1) = \emptyset$. Then $X_t = X_{t-1} - \{v_{t-1}\}$, which is a nice set since $v_{t-1} \in \partial(X_{t-1})$ and $X_t \neq \emptyset$.

Now assume that $R(t-1) = s$ for some $s \in \mathcal{S}$. Using Lemma 5, let P be the path of G determined by s , X_{t-1} , and v_{t-1} . Let P_1 and P_2 denote the two halves of P , so that $v_{t-1} \in V(P_2)$. Then $X_t = X_{t-1} \cup (V(P_2) - Q)$. Arguing by contradiction, suppose that X_t is not nice. Then there exists $vw \in E(H)$ such that $P_{vw} - X_t$ has at least two components. Let x, y be two vertices in distinct components of $P_{vw} - X_t$ that are as close as possible on the path P_{vw} . Then the set Z of vertices strictly between x and y on P_{vw} is a subset of X_t . On the other hand, $Z \cap X_{t-1} = \emptyset$ since otherwise x and y would be in distinct components of $P_{vw} - X_{t-1}$, contradicting the fact that X_{t-1} is nice. Thus $Z \subseteq V(P_2) - Q$, and also $\partial(X_{t-1}) \cap Z = \emptyset$. Since P_2 is connected and avoids x and y , we deduce that $Z = V(P_2)$ (and thus $Q \cap V(P_2) = \emptyset$). However, $v_{t-1} \in \partial(X_{t-1})$ and $v_{t-1} \in V(P_2) = Z$, contradicting $\partial(X_{t-1}) \cap Z = \emptyset$. \square

Claim 14. $\beta \leq 1.001$.

Proof. We need to show that $\alpha_k \leq 1.001^k$ for each $k \in [1, \lfloor \frac{n}{2} \rfloor]$. Fix such an integer k . Let \mathcal{W} be the set of triples (P, X, v) that can be considered by the algorithm in the uncolouring step, over all $t \geq 1$ and vectors $(c_1, \dots, c_t) \in \mathcal{F}_t$, such that P has exactly $2k$ vertices.

Observe that if $(P, X, v) \in \mathcal{W}$, then X is a nice subset of $V(G) - (Q \cup V(P))$ by Claim 13; also, $v \in V(P)$ and $X \cup \{v\}$ is again a nice set. By the definition of \mathcal{W} , every sequence $s \in \bar{\mathcal{S}}$ of length $2k$ is ‘produced’ by at least one triple in \mathcal{W} , in the sense that there exists $(P, X, v) \in \mathcal{W}$ such that $s = s(P, X, v)$. We may assume that \mathcal{W} is not empty, since otherwise $\alpha_k = 0$, and we are trivially done.

Let $(P, X, v) \in \mathcal{W}$ and let $s(P, X, v) = (s_1, \dots, s_{2k})$. Let v_1, \dots, v_{2k} be the vertices of P . Note that P may contain vertices of Q . Since v is not in Q , it has degree 2, and thus $s_1 \in \{1, 2\}$. (Note that we could have $s_1 = 2$ if the neighbour of v in P is in Q .) We have $s_p = -1$ for a unique $p \in [2, 2k]$. We claim that the sequence $s' := (s_{p-1}, \dots, s_2, 1, s_{p+1}, \dots, s_{2k})$

obtained from (s_1, \dots, s_{2k}) by removing the p -th entry, reversing the (s_1, \dots, s_{p-1}) prefix, and replacing s_1 by 1, is c -spread.

Case 1. $2k \leq c + 1$: Then P has no vertex u in Q , since otherwise P would have at least $g(u) + 2 \geq \lceil c \log 2 \rceil + 2 > c + 1$ vertices. It follows that there is an edge $xy \in E(H)$ such that P is a subpath of P_{xy} . Then v must be an endpoint of P . Indeed, if not, then the two neighbours of v in P are in distinct components of $P_{xy} - (X \cup \{v\})$, contradicting the fact that $X \cup \{v\}$ is nice. Clearly $s_i = 1$ for each $i \in [2, 2k - 1]$ and $s_{2k} = -1$. If v is an internal vertex of P_{xy} , then one of the two neighbours of v is in X , and it follows that $s_1 = 1$. If v is an endpoint of P_{xy} , then s_1 is the index in the set $N(v)$ of the only neighbour w of v that is in P_{xy} . This index is always 1 by our choice of the ordering of $V(G)$. Hence we again have $s_1 = 1$. Therefore $(s_1, \dots, s_{2k}) = (1, \dots, 1, -1)$ and s' is the sequence of $2k - 1$ one's, which is c -spread.

Case 2. $2k \geq c + 2$: If $s_i > 1$ for some $i \in [2, p]$, then v_i is in Q ; in this case, our goal is to show that s' contains $g(v_i)$ one's immediately before or after s_i that can be mapped to s_i . Similarly, if $s_{p+j} > 1$ for some $j \in [1, q]$, then w_j is in Q ; in this case, our goal is to show that s' contains $g(w_j)$ one's immediately before or after s_{p+j} that can be mapped to s_{p+j} .

Consider a vertex $u \in V(P) \cap Q$. By the definition of L , the colour assigned to u is assigned to no vertex that belongs to u (those in the set $M(u)$) when the algorithm considers the triple (P, X, v) . At that stage, P is repetitively coloured. Let x be the unique vertex at distance k from v in P . Then v and x are assigned the same colour, and x is not in $M(u)$. Walk from u towards x and stop after $g(u) + 1$ steps. This defines a subpath P' of P consisting of exactly $g(u) + 1$ vertices that belong to u , either all immediately before u or all immediately after u in P . Consider the following six possible values of u and P' :

- If $u = v_i$ and $P' = (v_{i+g(u)+1}, v_{i+g(u)}, \dots, v_{i+1})$ then $s_{i+g(u)} = s_{i+g(u)-1} \dots = s_{i+1} = 1$ and s' contains $g(u)$ one's immediately before s_i that can be mapped to s_i .
- If $u = v_i$ and $P' = (v_{i-1}, v_{i-2}, \dots, v_{i-g(u)-1})$ and $i - g(u) - 1 \neq 1$, then $s_{i-1} = s_{i-2} = \dots = s_{i-g(u)} = 1$ and s' contains $g(u)$ one's immediately after s_i that can be mapped to s_i .
- If $u = w_j$ and $P' = (w_{j+1}, w_{j+2}, \dots, w_{j+g(u)+1})$ then $s_{p+j+1} = s_{p+j+2} = \dots = s_{p+j+g(u)} = 1$ and s' contains $g(u)$ one's immediately after s_{p+j} that can be mapped to s_{p+j} .
- If $u = w_j$ and $P' = (w_{j-g(u)-1}, w_{j-g(u)}, \dots, w_{j-1})$ then $s_{p+j-g(u)} = s_{p+j-g(u)+1} = \dots = s_{p+j-1} = 1$ and s' contains $g(u)$ one's immediately before s_{p+j} that can be mapped to s_{p+j} .
- If $u = v_i$ and $P' = (v_{i-1}, v_{i-2}, \dots, v_1 = w_1, w_2, \dots, w_{g(u)-i+3})$ then $s_{i-1} = s_{i-2} = \dots =$

$s_2 = 1$ and $s_{p+1} = s_{p+2} = \dots = s_{p+g(u)-i+2} = 1$, implying that

$$s_{i-1}, s_{i-2}, \dots, s_2, 1, s_{p+1}, s_{p+2}, \dots, s_{p+g(u)-i+1}$$

is a sequence of $g(u)$ one's immediately after s_i in s' that can be mapped to s_i .

- If $u = w_j$ and $P' = (v_{g(u)-j+3}, v_{g(u)-j+2}, \dots, v_1 = w_1, w_2, \dots, w_{j-1})$ then $s_{g(u)-j+2} = s_{g(u)-j+1} = \dots = s_2 = 1$ and $s_{p+1} = s_{p+2} = \dots = s_{p+j-1} = 1$, implying that

$$s_{g(u)-j+2}, s_{g(u)-j+1}, \dots, s_2, 1, s_{p+1}, s_{p+2}, \dots, s_{p+j-1}$$

is a sequence of $g(u)$ one's immediately before s_{p+j} in s' that can be mapped to s_{p+j} .

Hence s' is c -spread, as claimed.

Therefore (s_1, \dots, s_{2k}) is obtained from a c -spread sequence s' of length $2k - 1$ by choosing an index $p \in [1, 2k - 1]$, inserting -1 between the p -th and $(p + 1)$ -th entries, reversing the prefix of s' up to the p -th entry, and possibly changing the first entry to a 2. Hence the number of distinct sequences in $\bar{\mathcal{S}}$ of length $2k$ is at most $2(2k - 1)$ times the number of c -spread sequences of length $2k - 1$. Let $\varepsilon := 0.0002$. Then ε and c satisfy the hypotheses of Lemma 11, and we deduce from that lemma that

$$\alpha_k \leq (4k - 2) \cdot (1 + \varepsilon)^{2k-1} \leq 4k(1 + \varepsilon)^{2k} \leq 1.001^k .$$

(The rightmost inequality holds because $2k \geq c$ and $2c(1 + \varepsilon)^c \leq 1.001^{c/2}$.) □

Next we show that every realisable Dyck word is special. Consider a word $d \in \mathcal{D}_t$ for some $t \geq 1$, and let $R \in \mathcal{R}_t$ be a record such that $D(R) = d$. Suppose that d contains 0110110 as a subsequence. Then there is an index $i \in [1, t - 2]$ such that $|R(i)| = |R(i + 1)| = 4$. Fix an arbitrary vector $(c_1, \dots, c_t) \in \mathcal{F}_{t,R}$, and let (P, X, v) and (P', X', v') be the triples such that $s(P, X, v) = R(i)$ and $s(P', X', v') = R(i + 1)$, respectively, in the execution of the algorithm on input (c_1, \dots, c_t) . Then P contains no vertex from Q , since otherwise P would need to have at least $c + 2 > 4$ vertices, as explained in Case 1 of the proof of Claim 14. Since our ordering of $V(G)$ puts vertices in $V(G) - Q$ before those in Q , and since X is nice, it follows that $s(P, X, v) = R(i) = (1, 1, 1, -1)$. By the same argument $s(P', X', v') = R(i + 1) = (1, 1, 1, -1)$. Let v_1, \dots, v_4 denote the vertices of P , with $v_4 = v$. Then in the i -th iteration of the while-loop of the algorithm, immediately after colouring v_4 , we have $\phi(v_1) = \phi(v_3)$ and $\phi(v_2) = \phi(v_4)$. Vertices v_3 and v_4 are subsequently uncoloured. Thus $X' = X \cup \{v_3\}$.

By our choice of the priority function f , we have $f(X') = v' = v_3$. Indeed, $f(X) = v_4$ and $v_1, v_2 \in V(P_{vw}) - X'$, where $vw \in E(H)$ is the edge such that $P \subseteq P_{vw}$. In particular, $v_3 \in \partial(X')$ and $V(P_{vw}) - X' \neq \emptyset$. Thus either v_4 is before $V(P_{vw}) - X$ on P_{vw} , in which case v_3 is before $V(P_{vw}) - X'$ on P_{vw} , implying $f(X') = v_3$; or v_4 is after $V(P_{vw}) - X$ on P_{vw} , in

which case v_3 is after $V(P_{vw}) - X'$ and there is no vertex in $\partial(X')$ before $V(P_{vw}) - X'$ on P_{vw} , implying again $f(X') = v_3$.

It follows that the vertices of P' are v_0, v_1, v_2, v_3 in order, where $v_0 \in V(P_{vw})$ (and obviously $v_0 \neq v_4$). In the $(i + 1)$ -th iteration of the while-loop, immediately after colouring $v_3 (=v')$, we have $\phi(v_0) = \phi(v_2)$ and $\phi(v_1) = \phi(v_3)$. However, the colours of v_0, v_1, v_2 have not changed since the beginning of the i -th iteration, and $\phi(v_1) = \phi(v_3)$ at that point. This implies that P' was already repetitively coloured at the start of the i -th iteration, a contradiction. Therefore, realisable Dyck words are special, as claimed.

Let $t \geq 1$, let $m := \min\{n, t\}$, let $i \in [1, m]$ and let (d_1, \dots, d_{2t}) be a realisable Dyck word of length $2t$ such that $d_{2t-j} = 1$ for each $j \in [0, i - 1]$. Say there are q maximal subsequences of consecutive ones in (d_1, \dots, d_{2t-i}) . If $q > 0$ then let k_1, \dots, k_q be the lengths of these sequences. If $q \geq 1$, then $\sum_{j=1}^q k_j \leq t - i$, and we deduce that there are at most $\alpha_{k_1} \alpha_{k_2} \dots \alpha_{k_q} \leq \beta^{k_1} \beta^{k_2} \dots \beta^{k_q} \leq \beta^{t-i} \leq \beta^t$ distinct records $R \in \mathcal{R}_t$ with $z(R) = i$ such that $D(R) = (d_1, \dots, d_{2t})$. If $q = 0$, then there is at most $1 \leq \beta^t$ records $R \in \mathcal{R}_t$ with $z(R) = i$ such that $D(R) = (d_1, \dots, d_{2t})$.

Since for each $i \in [1, m]$, there are at most β^t distinct records $R \in \mathcal{R}_t$ with $z(R) = i$ that have the same Dyck word $D(R)$, and since there are exactly $|\mathcal{D}_t|$ distinct realisable special Dyck words of length $2t$, it follows that $|\mathcal{R}_t| \leq m|\mathcal{D}_t|\beta^t \leq n|\mathcal{D}_t|\beta^t$. Using Claim 14 and Lemmas 7 and 12, we obtain

$$|\mathcal{F}_t| = |\mathcal{A}_t| \leq n(\ell + 1)^n |\mathcal{D}_t| \beta^t \leq n(\ell + 1)^n 3.992^{t+1} 1.001^t < n(\ell + 1)^n 3.996^{t+1} = o(\ell^t) .$$

Hence, if t is sufficiently large then there is at least one vector (c_1, \dots, c_t) among the ℓ^t many vectors in $[1, \ell]^t$ on which the algorithm succeeds. Therefore G admits a nonrepetitive L' -colouring. \square

Note that we made no effort to optimise the constant 10^5 in the proof of Theorem 2.

5 Pathwidth Proofs

The proof of Theorem 3 depends on the following lemma of independent interest.

Lemma 15. *Let B_1, \dots, B_m be pairwise disjoint sets of vertices in a graph G , such that no two vertices in distinct B_i are adjacent. Let H be the graph obtained from G by deleting B_i and adding a clique on $N_G(B_i)$ for each $i \in [1, m]$. Then*

$$\pi(G) \leq \pi(H) + \max_i \pi(G[B_i]) .$$

Proof. Nonrepetitively colour $G[B_1 \cup \dots \cup B_m]$ with $\max_i \pi(G[B_i])$ colours. Nonrepetitively colour H with a different set of $\pi(H)$ colours. Suppose on the contrary that G contains

a repetitively coloured path P . Let P' be the set of vertices in P that are in H , ordered according to P . Then $P' \neq \emptyset$, as otherwise P is contained in some B_i , implying B_i contains a repetitively coloured path. Consider a maximal subpath S in P that is not in H . So S was deleted from P in the construction of P' . Since no two vertices in distinct B_i are adjacent, S is contained in a single set B_i . Thus the vertices in P immediately before and after S (if they exist) are in $N_G(B_i)$, and are thus adjacent in H . Hence P' is a path in H . Since the vertices in $B_1 \cup \dots \cup B_m$ receive distinct colours from the vertices in H , the path P' is repetitively coloured. This contradiction proves that G is nonrepetitively coloured. \square

The next lemma provides a useful way to think about graphs of bounded pathwidth. Let $G \cdot K_k$ denote the *lexicographical product* of a graph G and the complete graph K_k . That is, $G \cdot K_k$ is obtained by replacing each vertex of G by a copy of K_k , and replacing each edge of G by a copy of $K_{k,k}$.

Lemma 16. *Every graph G with pathwidth k contains pairwise disjoint sets B_1, \dots, B_m of vertices, such that:*

- *no two vertices in distinct B_i are adjacent,*
- *$G[B_i]$ has pathwidth at most $k - 1$ for each $i \in [1, m]$, and*
- *if H is the graph obtained from G by deleting B_i and adding a clique on $N_G(B_i)$ for each $i \in [1, m]$, then H is a subgraph of $P_m \cdot K_{k+1}$.*

Proof. Consider a path decomposition \mathcal{D} of G with width k . Let X_1, \dots, X_m be the set of bags in \mathcal{D} , such that X_1 is the first bag in \mathcal{D} , and for each $i \geq 2$, the bag X_i is the first bag in \mathcal{D} that is disjoint from X_{i-1} . Thus X_1, \dots, X_m are pairwise disjoint. For $i \in [1, m]$, let B_i be the set of vertices that only appear in bags strictly between X_i and X_{i+1} (or strictly after X_m if $i = m$). By construction, each such bag intersects X_i . Hence $G[B_i]$ has pathwidth at most $k - 1$. Since each X_i separates B_{i-1} and B_{i+1} (for $i \neq m$), no two vertices in distinct B_i are adjacent. Moreover, the neighbourhood of B_i is contained in $X_i \cup X_{i+1}$ (or X_i if $i = m$). Hence the graph H (defined above) has vertex set $X_1 \cup \dots \cup X_m$ where $X_i \cup X_{i+1}$ is a clique for each $i \in [1, m - 1]$. Since $|X_i| \leq k + 1$, the graph H is a subgraph of $P_m \cdot K_{k+1}$. \square

Proof of Theorem 3. We proceed by induction on $k \geq 0$. Every graph with pathwidth 0 is edgeless, and is thus nonrepetitively 1-colourable, as desired. Now assume that G is a graph with pathwidth $k \geq 1$. Let B_1, \dots, B_m be the sets that satisfy Lemma 16. Let H be the graph obtained from G by deleting B_i and adding a clique on $N_G(B_i)$ for each $i \in [1, m]$. Then H is a subgraph of $P_{m+1} \cdot K_{k+1}$, which is nonrepetitively $4(k + 1)$ -colourable by a theorem of Kündgen and Pelsmajer [31]⁶. By induction, $\pi(G[B_i]) \leq 2(k - 1)^2 + 6(k - 1) - 4$.

⁶Say V_1, \dots, V_t is a partition of $V(G)$ such that for all $i \in [1, t]$, we have $N_G(V_i) \subseteq V_{i-1} \cup V_{i+1}$ and $N_G(V_i) \cap V_{i-1}$ is a clique. Kündgen and Pelsmajer [31] proved that $\pi(G) \leq 4 \max_i \pi(G[V_i])$. Clearly $P_m \cdot K_{k+1}$ has such a partition with each V_i a $(k + 1)$ -clique. Thus $\pi(P_m \cdot K_{k+1}) \leq 4(k + 1)$.

By Lemma 15, $\pi(G) \leq \pi(H) + \max_i \pi(G[B_i]) \leq 4(k+1) + 2(k-1)^2 + 6(k-1) - 4 = 2k^2 + 6k - 4$. This completes the proof. \square

Proof of Theorem 4. We proceed by induction on $k \geq 0$. Every graph with pathwidth 0 is edgeless, and is thus star 1-colourable, as desired. Now assume that G is a graph with pathwidth $k \geq 1$. We may assume that G is connected. Let G' be an interval graph such that G' contains G as a spanning subgraph and $\omega(G') = k + 1$. Let $I(v)$ be the interval representing each vertex v . Let X be an inclusion-wise minimal set of vertices in G' such that for every vertex w ,

$$I(w) \subseteq \bigcup \{I(v) : v \in X\} . \quad (3)$$

The set X exists since $X = V(G)$ satisfies (3). It is easily seen that $G[X]$ is an induced path, say (x_1, \dots, x_n) . Colour x_i by $i \bmod 3$ (in $\{0, 1, 2\}$). Observe that $G'[X]$ is star 3-coloured. By (3), the subgraph $G' - X$ is an interval graph with $\omega(G' - X) \leq k$. Thus, by induction, $G' - X$ is star-colourable with colours $\{3, 4, \dots, 3k\}$. Suppose on the contrary that G' contains a 2-coloured path (u, v, w, x) . First suppose that u is in X . Then w is also in X . If v is also in X then so is x , which contradicts the fact that $G'[X]$ is star-coloured. So $v \notin X$. Since u and v receive the same colour, there are at least two vertices p and q between u and v in the path $G'[X]$. Thus replacing p and q by v gives a shorter path that satisfies (3). This contradiction proves that $u \notin X$. By symmetry $x \notin X$. Since X and $G' - X$ are assigned disjoint sets of colours, $v \notin X$ and $w \notin X$. Hence (u, v, w, x) is a 2-coloured path in $G' - X$, which is the desired contradiction. Hence G' is star-coloured with $3k + 1$ colours. \square

6 Open problems

We conclude with a number of open problems:

- Whether there is a relationship between nonrepetitive choosability and pathwidth is an interesting open problem. It is easily seen⁷ that graphs with pathwidth 1 (i.e., caterpillars) are nonrepetitively c -choosable for some constant c . Is every graph (or tree) with pathwidth 2 nonrepetitively c -choosable for some constant c ?
- Except for a finite number of examples, every cycle is nonrepetitively 3-colourable [12]. Every cycle is nonrepetitively 5-choosable. (*Proof.* Precolour one vertex, remove

⁷*Proof Sketch.* A coloured path (v_1, \dots, v_{2p+1}) with $p \geq 1$ is *almost repetitively coloured* if $\psi(v_i) = \psi(v_{p+i+1})$ for all $i \in [1, p]$. By the Lovász Local Lemma, for some constant c , every path is c -choosable such that no subpath is repetitively coloured and no subpath is almost repetitively coloured (see Appendix B). Now, let T be a caterpillar with spine P (that is, T is a tree such that deleting the leaves of T gives a path P). Apply the above result to P , and then colour each leaf of T by a colour in its list distinct from the colour assigned to its neighbour in P . There is no repetitive path, as otherwise P would contain a repetitively coloured or almost repetitively coloured path.

this colour from every other list, and apply the nonrepetitive 4-choosability result for paths.) Is every cycle nonrepetitively 4-choosable? Which cycles are nonrepetitively 3-choosable?

- Does every graph have a nonrepetitively 4-choosable subdivision? Even 3-choosable might be possible.
- Is there a function f such that every graph G has a nonrepetitively $\mathcal{O}(1)$ -colourable subdivision with $f(\pi(G))$ division vertices per edge?
- Is there a function f such that $\pi(G/M) \leq f(\pi(G))$ for every graph G and for every matching M of G , where G/M denotes the graph obtained from G by contracting the edges in M ? This would generalise a result of Nešetřil et al. [37] about subdivisions (when each edge in M has one endpoint of degree 2).
- Is there a polynomial-time Monte Carlo algorithm that nonrepetitively $\mathcal{O}(\Delta^2)$ -colours a graph with maximum degree Δ ? Haeupler et al. [25] show that $\mathcal{O}(\Delta^{2+\varepsilon})$ colours suffice for all fixed $\varepsilon > 0$; also see [10, 29] for related results. Note that testing whether a given colouring of a graph is nonrepetitive is co-NP-complete, even for 4-colourings [33].

References

- [1] MICHAEL O. ALBERTSON, GLENN G. CHAPPELL, HAL A. KIERSTEAD, ANDRÉ KÜNDGEN, AND RADHIKA RAMAMURTHI. Coloring with no 2-colored P_4 's. *Electron. J. Combin.*, 11 #R26, 2004. http://www.combinatorics.org/Volume_11/Abstracts/v11i1r26.html. MR: 2056078.
- [2] NOGA ALON AND JAROSŁAW GRZYTCZUK. Breaking the rhythm on graphs. *Discrete Math.*, 308:1375–1380, 2008. doi: [10.1016/j.disc.2007.07.063](https://doi.org/10.1016/j.disc.2007.07.063). MR: 2392054.
- [3] NOGA ALON, JAROSŁAW GRZYTCZUK, MARIUSZ HALUSZCZAK, AND OLIVER RIORDAN. Nonrepetitive colorings of graphs. *Random Structures Algorithms*, 21(3-4):336–346, 2002. doi: [10.1002/rsa.10057](https://doi.org/10.1002/rsa.10057). MR: 1945373.
- [4] JÁNOS BARÁT AND PÉTER P. VARJÚ. On square-free vertex colorings of graphs. *Studia Sci. Math. Hungar.*, 44(3):411–422, 2007. doi: [10.1556/SScMath.2007.1029](https://doi.org/10.1556/SScMath.2007.1029). MR: 2361685.
- [5] JÁNOS BARÁT AND PÉTER P. VARJÚ. On square-free edge colorings of graphs. *Ars Combin.*, 87:377–383, 2008. MR: 2414029.

- [6] JÁNOS BARÁT AND DAVID R. WOOD. Notes on nonrepetitive graph colouring. *Electron. J. Combin.*, 15:R99, 2008. http://www.combinatorics.org/Volume_15/Abstracts/v15i1r99.html. MR: 2426162.
- [7] OLEG V. BORODIN. On acyclic colorings of planar graphs. *Discrete Math.*, 25(3):211–236, 1979. doi: [10.1016/0012-365X\(79\)90077-3](https://doi.org/10.1016/0012-365X(79)90077-3). MR: 534939.
- [8] BOŠTJAN BREŠAR, JAROSŁAW GRZYTCZUK, SANDI KLAVŽAR, STANISŁAW NIWCZYK, AND IZTOK PETERIN. Nonrepetitive colorings of trees. *Discrete Math.*, 307(2):163–172, 2007. doi: [10.1016/j.disc.2006.06.017](https://doi.org/10.1016/j.disc.2006.06.017). MR: 2285186.
- [9] BOŠTJAN BREŠAR AND SANDI KLAVŽAR. Square-free colorings of graphs. *Ars Combin.*, 70:3–13, 2004. MR: 2023057.
- [10] KARTHEKEYAN CHANDRASEKARAN, NAVIN GOYAL, AND BERNHARD HAEUPLER. Deterministic algorithms for the Lovász local lemma. In *Proc. 21st Annual ACM-SIAM Symposium on Discrete Algorithms (SODA '10)*, pp. 992–1004. SIAM, 2010. <http://dl.acm.org/citation.cfm?id=1873601.1873681>. To appear in *SIAM J. Computing*.
- [11] PANAGIOTIS CHEILARIS, ERNST SPECKER, AND STATHIS ZACHOS. Neochromatica. *Comment. Math. Univ. Carolin.*, 51(3):469–480, 2010. <http://www.dml.cz/dmlcz/140723>. MR: 2741880.
- [12] JAMES D. CURRIE. There are ternary circular square-free words of length n for $n \geq 18$. *Electron. J. Combin.*, 9(1), 2002. http://www.combinatorics.org/Volume_9/Abstracts/v9i1n10.html. MR: 1936865.
- [13] JAMES D. CURRIE. Pattern avoidance: themes and variations. *Theoret. Comput. Sci.*, 339(1):7–18, 2005. doi: [10.1016/j.tcs.2005.01.004](https://doi.org/10.1016/j.tcs.2005.01.004). MR: 2142070.
- [14] SEBASTIAN CZERWIŃSKI AND JAROSŁAW GRZYTCZUK. Nonrepetitive colorings of graphs. *Electron. Notes Discrete Math.*, 28:453–459, 2007. doi: [10.1016/j.endm.2007.01.063](https://doi.org/10.1016/j.endm.2007.01.063). MR: 2324051.
- [15] GUILLAUME FERTIN, ANDRÉ RASPAUD, AND BRUCE REED. Star coloring of graphs. *J. Graph Theory*, 47(3):163–182, 2004. doi: [10.1002/jgt.20029](https://doi.org/10.1002/jgt.20029). MR: 2089462.
- [16] FRANCESCA FIORENZI, PASCAL OCHEM, PATRICE OSSONA DE MENDEZ, AND XUDING ZHU. Thue choosability of trees. *Discrete Applied Math.*, 159(17):2045–2049, 2011. doi: [10.1016/j.dam.2011.07.017](https://doi.org/10.1016/j.dam.2011.07.017). MR: 2832329.
- [17] PHILIPPE FLAJOLET AND ROBERT SEDGEWICK. *Analytic combinatorics*. Cambridge University Press, 2009.

- [18] BRANKO GRÜNBAUM. Acyclic colorings of planar graphs. *Israel J. Math.*, 14:390–408, 1973. doi: [10.1007/BF02764716](https://doi.org/10.1007/BF02764716). MR: 0317982.
- [19] JAROSŁAW GRYTCZUK. Thue-like sequences and rainbow arithmetic progressions. *Electron. J. Combin.*, 9(1):R44, 2002. http://www.combinatorics.org/Volume_9/Abstracts/v9i1r44.html. MR: 1946146.
- [20] JAROSŁAW GRYTCZUK. Nonrepetitive colorings of graphs—a survey. *Int. J. Math. Math. Sci.*, 74639, 2007. doi: [10.1155/2007/74639](https://doi.org/10.1155/2007/74639). MR: 2272338.
- [21] JAROSŁAW GRYTCZUK. Nonrepetitive graph coloring. In *Graph Theory in Paris*, Trends in Mathematics, pp. 209–218. Birkhauser, 2007.
- [22] JAROSŁAW GRYTCZUK. Thue type problems for graphs, points, and numbers. *Discrete Math.*, 308(19):4419–4429, 2008. doi: [10.1016/j.disc.2007.08.039](https://doi.org/10.1016/j.disc.2007.08.039). MR: 2433769.
- [23] JAROSŁAW GRYTCZUK, JAKUB KOZIK, AND PIOTR MICEK. A new approach to non-repetitive sequences. 2011. arXiv: [1103.3809](https://arxiv.org/abs/1103.3809). To appear in *Random Structures Algorithms*.
- [24] JAROSŁAW GRYTCZUK, JAKUB PRZYBYŁO, AND XUDING ZHU. Nonrepetitive list colourings of paths. *Random Structures Algorithms*, 38(1-2):162–173, 2011. doi: [10.1002/rsa.20347](https://doi.org/10.1002/rsa.20347). MR: 2768888.
- [25] BERNHARD HAEUPLER, BARNA SAHA, AND ARAVIND SRINIVASAN. New constructive aspects of the Lovász local lemma. *J. ACM*, 58(6):28, 2011. doi: [10.1145/2049697.2049702](https://doi.org/10.1145/2049697.2049702).
- [26] JOCHEN HARANTA AND STANISLAV JENDROL’. Nonrepetitive vertex colorings of graphs. *Discrete Math.*, 312(2):374–380, 2012. doi: [10.1016/j.disc.2011.09.027](https://doi.org/10.1016/j.disc.2011.09.027).
- [27] FRÉDÉRIC HAVET, STANISLAV JENDROĽ, ROMAN SOTÁK, AND ERIKA ŠKRABUĽÁKOVÁ. Facial non-repetitive edge-coloring of plane graphs. *J. Graph Theory*, 66(1):38–48, 2011. doi: [10.1002/jgt.20488](https://doi.org/10.1002/jgt.20488). MR: 2742187.
- [28] STANISLAV JENDROL AND ERIKA ŠKRABUĽÁKOVÁ. Facial non-repetitive edge colouring of semiregular polyhedra. *Acta Univ. M. Belii Ser. Math.*, 15:37–52, 2009. <http://actamath.savbb.sk/acta1503.shtml>. MR: 2589669.
- [29] KASHYAP KOLIPAKA AND MARIO SZEGEDY. Moser and Tardos meet Lovász. In *Proc. 43rd Annual ACM Symposium on Theory of Computing (STOC ’11)*, pp. 235–244. ACM, 2011. doi: [10.1145/1993636.1993669](https://doi.org/10.1145/1993636.1993669).
- [30] JAKUB KOZIK AND PIOTR MICEK. Nonrepetitive colorings of trees. 2011. Preprint.
- [31] ANDRE KÜNDGEN AND MICHAEL J. PELSMAJER. Nonrepetitive colorings of graphs of bounded tree-width. *Discrete Math.*, 308(19):4473–4478, 2008. doi: [10.1016/j.disc.2007.08.043](https://doi.org/10.1016/j.disc.2007.08.043). MR: 2433774.

- [32] FEDOR MANIN. The complexity of nonrepetitive edge coloring of graphs, 2007. arXiv: [0709.4497](https://arxiv.org/abs/0709.4497).
- [33] DÁNIEL MARX AND MARCUS SCHAEFER. The complexity of nonrepetitive coloring. *Discrete Appl. Math.*, 157(1):13–18, 2009. doi: [10.1016/j.dam.2008.04.015](https://doi.org/10.1016/j.dam.2008.04.015). MR: [2479374](https://dblp.org/mr/2479374).
- [34] MICHAEL MOLLOY AND BRUCE REED. *Graph colouring and the probabilistic method*, vol. 23 of *Algorithms and Combinatorics*. Springer, 2002.
- [35] ROBIN A. MOSER AND GÁBOR TARDOS. A constructive proof of the general Lovász local lemma. *J. ACM*, 57(2), 2010. doi: [10.1145/1667053.1667060](https://doi.org/10.1145/1667053.1667060).
- [36] JAROSLAV NEŠETŘIL AND PATRICE OSSONA DE MENDEZ. Colorings and homomorphisms of minor closed classes. In BORIS ARONOV, SAUGATA BASU, JÁNOS PACH, AND MICHA SHARIR, eds., *Discrete and Computational Geometry, The Goodman-Pollack Festschrift*, vol. 25 of *Algorithms and Combinatorics*, pp. 651–664. Springer, 2003. MR: [2038495](https://dblp.org/mr/2038495).
- [37] JAROSLAV NEŠETŘIL, PATRICE OSSONA DE MENDEZ, AND DAVID R. WOOD. Characterisations and examples of graph classes with bounded expansion. *European J. Combinatorics*, 33(3):350–373, 2011. doi: [10.1016/j.ejc.2011.09.008](https://doi.org/10.1016/j.ejc.2011.09.008).
- [38] WESLEY PEGDEN. Highly nonrepetitive sequences: winning strategies from the local lemma. *Random Structures Algorithms*, 38(1-2):140–161, 2011. doi: [10.1002/rsa.20354](https://doi.org/10.1002/rsa.20354). MR: [2768887](https://dblp.org/mr/2768887)
- [39] ANDRZEJ PEZARSKI AND MICHAŁ ZMARZ. Non-repetitive 3-coloring of subdivided graphs. *Electron. J. Combin.*, 16(1):N15, 2009. http://www.combinatorics.org/Volume_16/Abstracts/v16i1n15.html. MR: [2515755](https://dblp.org/mr/2515755).
- [40] AXEL THUE. Über unendliche Zeichenreihen. *Norske Vid. Selsk. Skr. I. Mat. Nat. Kl. Christiania*, 7:1–22, 1906.
- [41] DAVID R. WOOD. Acyclic, star and oriented colourings of graph subdivisions. *Discrete Math. Theor. Comput. Sci.*, 7(1):37–50, 2005. <http://www.dmtcs.org/dmtcs-ojs/index.php/dmtcs/article/view/60>. MR: [2164057](https://dblp.org/mr/2164057).

A Subdivisions via the Lovász Local Lemma

For the sake of comparing related proof techniques, we now give a proof of a qualitatively similar result to Theorem 2 that uses the Lovász Local Lemma instead of entropy compres-

sion. In particular, we use the following ‘weighted’ version of the Lovász Local Lemma; see [34].

Lemma 17. *Let $\mathcal{E} = \{A_1, \dots, A_n\}$ be a set of “bad” events such that each A_i is mutually independent of $\mathcal{E} \setminus (\mathcal{D}_i \cup \{A_i\})$ for some $\mathcal{D}_i \subseteq \mathcal{E}$. Let p be a real number such that $0 < p \leq \frac{1}{4}$. Let $t_1, \dots, t_n \geq 1$ be integers called weights, such that for all $i \in [1, n]$,*

$$(a) \Pr(A_i) \leq p^{t_i}, \text{ and}$$

$$(b) 2 \sum_{A_j \in \mathcal{D}_i} (2p)^{t_j} \leq t_i .$$

Then with positive probability, no event in \mathcal{E} occurs.

Theorem 18. *For every graph H with maximum degree Δ , every subdivision G of H with at least $3 + 400 \log \Delta$ division vertices per edge is nonrepetitively 23-choosable.*

Proof. We may assume that $\Delta \geq 2$. Let $r := 3 + \lceil 400 \log \Delta \rceil$. Let G be a subdivision of H with at least r division vertices per edge. Arbitrarily colour each original vertex of H from its list. For each edge vw of H , delete the colours chosen for v and w from the list of each division vertex on the edge vw . Now each division vertex has a list of at least 21 colours. Colour each division vertex independently and randomly from its list. Let $p := \frac{1}{21}$.

Suppose that some path P containing exactly one original vertex v is repetitively coloured. Let x be the vertex corresponding to v in the other half of P . Thus x is a division vertex of some edge incident to v , which is a contradiction since the colour of v was removed from the list of colours at x . Thus no path with exactly one original vertex is repetitively coloured. Say an even path with no original vertices is *short*, and an even path with at least two original vertices is *long*. To prove that a colouring of G is nonrepetitive it suffices to prove that no long or short path is repetitively coloured.

Let \mathcal{P} be the set of all short or long paths in G . Say $\mathcal{P} = \{P_1, \dots, P_n\}$ and each P_i has $2\ell_i$ vertices, of which k_i are original vertices. Note that

$$2\ell_i \geq (k_i - 1)(r + 1) + 1 = (r + 1)k_i - r . \quad (4)$$

Orient each path P_i so that the j -th vertex in P_i is well defined. (Edges may be oriented differently in different paths.) Let A_i be the event that P_i is repetitively coloured. Let $\mathcal{E} := \{A_1, \dots, A_n\}$. If $P_i = (v_1, \dots, v_{2\ell_i})$ then v_j and v_{ℓ_i+j} are both division vertices for at least $\ell_i - k_i$ values of $j \in [1, \ell_i]$. Let $t_i := \ell_i - k_i$ be the weight of A_i and of P_i . Hence $\Pr(A_i) \leq (\frac{1}{21})^{t_i}$, and Lemma 17(a) is satisfied.

We claim that $100t_i \geq 99\ell_i$ for each path P_i . This is immediately true if $k_i = 0$. Now assume that $k_i \geq 2$. By (4), we have $2\ell_i \geq (r + 1)(k_i - 1) \geq 400(k_i - 1)$ and $\ell_i \geq 200$. Thus $400k_i \leq 2\ell_i + 400$ and $400t_i = 400\ell_i - 400k_i \geq 398\ell_i - 400 \geq 396\ell_i$, as claimed.

For each $i \in [1, n]$, let \mathcal{D}_i be the set of events A_j such that the corresponding path P_j and P_i have a division vertex in common. Since division vertices are coloured independently, A_i is mutually independent of $\mathcal{E} \setminus (\mathcal{D}_i \cup \{A_i\})$.

Let $P_i \in \mathcal{P}$. Our goal is to bound the number of paths in \mathcal{P} of given weight t that intersect P_i .

First consider the case of short paths of weight t . Such paths have order $2t$. Each vertex is in at most $2t$ short paths of order $2t$. Thus each vertex is in at most $2t$ short paths of weight t . Thus P_i , which has order $2\ell_i$, intersects at most $2\ell_i \cdot 2t \leq \frac{400}{99}t t_i$ short paths of weight t .

Now consider the case of long paths with weight t . Let P_j be such a long path. By (4), we have $(r+1)k_j \leq 2\ell_j + r \leq 4\ell_j = 4t + 4k_j$. Thus $k_j \leq \frac{4t}{r-3}$. Thus for each q , each vertex is the q -th vertex in at most $\Delta^{4t/(r-3)}$ long paths of weight t . Since $r-3 \geq 400 \log \Delta$, each vertex is the q -th vertex in at most $2^{t/100}$ long paths of weight t . Now $|P_i| = 2\ell_i \leq \frac{200}{99}t_i$. Similarly, each path of weight t has at most $\frac{200}{99}t$ vertices. Hence P_i intersects at most $(\frac{200}{99})^2 t_i t 2^{t/100}$ long paths of weight t . The same bound holds for short paths of weight t (since $\frac{400}{99}t t_i \leq (\frac{200}{99})^2 t_i t 2^{t/100}$).

Thus Lemma 17(b) is satisfied if for all i ,

$$2 \sum_{t \geq 1} (\frac{200}{99})^2 t_i t 2^{t/100} (\frac{2}{21})^t \leq t_i ;$$

that is,

$$2(\frac{200}{99})^2 \sum_{t \geq 1} t (\frac{2}{21} 2^{1/100})^t \leq 1 .$$

For $0 < c < 1$, we have $\sum_{t \geq 1} t c^t = \frac{c}{(1-c)^2}$. Let $c := \frac{2}{21} 2^{1/100} \approx 0.0959$. Thus

$$2(\frac{200}{99})^2 \sum_{t \geq 1} t c^t = 2(\frac{200}{99})^2 \frac{c}{(1-c)^2} < 0.958 ,$$

as desired. Hence Lemma 17(b) is satisfied.

Therefore with positive probability, no event in \mathcal{E} occurs. Thus, there exists a choice of colours for the division vertices such that no event in \mathcal{E} occurs. That is, no short or long path is repetitively coloured. Hence G is nonrepetitively colourable from the given lists. Therefore G is nonrepetitively 23-choosable. \square

B Caterpillars

Here we prove that every caterpillar is nonrepetitively 64-choosable. (Note that the constant 64 can be significantly improved using entropy-compression.) A coloured path (v_1, \dots, v_{2p+1}) is *almost repetitively coloured* if $\psi(v_i) = \psi(v_{p+i+1})$ for all $i \in [1, p]$.

Lemma 19. *Every path G is 64-choosable such that no path is repetitively coloured and no path is almost repetitively coloured.*

Proof. Colour each vertex v of G independently and randomly by a colour in the list of v . Let P_1, \dots, P_n be all the subpaths of G of order at least 2. If $|P_i|$ is even, then let A_i be the event that P_i is repetitively coloured. If $|P_i|$ is odd, then let A_i be the event that P_i is almost repetitively coloured. Let $t_i := \lfloor \frac{1}{2}|P_i| \rfloor$ be the weight of A_i and of P_i . Thus $\Pr(A_i) \leq 64^{-t_i}$. Hence Lemma 17(a) is satisfied with $p := 64^{-1}$. Let \mathcal{D}_i be the set of events corresponding to paths that intersect P_i . Thus A_i is mutually independent of $\{A_1, \dots, A_n\} \setminus (\mathcal{D}_i \cup \{A_i\})$. Each vertex is in at most ℓ subpaths of order ℓ . Each path of weight t has order $2t$ or $2t + 1$. So each vertex is in at most $4t + 1$ subpaths of weight t . Hence P_i intersects at most $(2t_i + 1)(4t_i + 1) \leq 15t_i t$ paths of weight t . To apply Lemma 17 we need

$$2 \sum_{t \geq 1} 15t t_i 32^{-t} \leq t_i .$$

That is,

$$30 \sum_{t \geq 1} t 32^{-t} \leq 1 .$$

Now $\sum_{t \geq 1} t c^{-t} = \frac{c}{(c-1)^2}$ for all $c > 1$. Thus

$$30 \sum_{t \geq 1} t 32^{-t} = \frac{30 \cdot 32}{31^2} < 1 ,$$

as desired. Hence Lemma 17(b) is satisfied. Therefore with positive probability, no A_i occurs. Thus, there exists a choice of colours such that no A_i occurs, in which case no path is repetitively coloured and no path is almost repetitively coloured. \square

Theorem 20. *Every caterpillar is nonrepetitively 64-choosable.*

Proof. Let P be the spine of a caterpillar T . By Lemma 19, P is 64-choosable such that no subpath is repetitively coloured and no subpath is almost repetitively coloured. Colour each leaf v of T by an arbitrary colour in the list of v that is distinct from the colour assigned to the neighbour of v (which is in P). There is no repetitive path in T , otherwise there is a repetitively coloured or almost repetitively coloured path in P . \square