UNIVERSITÉ CÔTE D'AZUR

ÉCOLE DOCTORALE

SCIENCES ET TECHNOLOGIES DE L'INFORMATION ET DE LA COMMUNICATION

THÈSE DE DOCTORAT

Sous-structures dans les graphes dirigés

William LOCHET

Laboratoire d'informatique, Signaux et Système de Sophia Antipolis

Présentée en vue de l'obtention
du grade de docteur en Informatique
d'Université Côte d'Azur
Dirigée par : Fréderic Havet / Stéphan Thomassé
Soutenue le : 19 Juillet 2018

Devant le jury, composé de :
Victor Chepoi, Professeur, Aix-Marseille Université
Daniel Gonçalves, Chargé de Recherches, LIRMM
Fréderic Havet, Directeur de Recherches, I3S
Bojan Mohar, Professor, Simon Fraser University and University of Ljubljana
András Sebő, Directeur de Recherches, G-scop
Stéphan Thomassé, Professeur, ENS Lyon
Ioan Todinca, Professeur, Université d'Orléans

Sous-structures dans les graphes dirigés

Jury:

Directeurs

Frédéric Havet, Directeur de Recherches, I3S
Stéphan Thomassé, Professeur, ENS de Lyon

Rapporteurs

Bojan Mohar, Professor, Simon Fraser University and University of Ljubljana
András Sebő, Directeur de Recherches, G-scop

Examinateurs

Victor Chepoi, Professeur, Aix-Marseille Université
Daniel Gonçalves, Chargé de Recherches, LIRMM
Ioan Todinca, Professeur, Université d'Orléans

Acknowledgements

Firstly, I would like to express my sincere gratitude to András and Bojan for their careful reading and comments that helped improve the presentation of this thesis. I would also like to thank Daniel, Ioan and Victor for accepting to be part of my committee.

Je voudrais ensuite remercier Fred et Stéphan pour ces trois années à travailler a vos côtés. Que ce soit humainement, ou scientifiquement, j'ai apprecié chaque discussion que j'ai pu avoir avec vous deux. Par vos questions, vous m'avez donné l'occasion de réfléchir à de superbes problèmes, et c'est la, à mon sens, le but principal d'une thèse.

Because a thesis consists mostly of hours spent on the board discussing maths with other people, I would like to thank all the people I was lucky enough to work with since I entered this wonderful world of graph theory: Anders, Cyril, Gwen, Marthe, Nicolas B., Nicolas N., Peter, Pierre A., Pierre C., Phablo, Raul, Stéphane B., Stéphane P. and Tash. A ce sujet, je remercie tout particulierement Nathann pour ces heures passées à réfléchir et me raconter des théorèmes.

Je remercie toute l'equipe COATI pour sa bonne humeur, sa gentillesse, ses discussions foot, ses discussions vélo, le tout enrobé en permanence d'un leger ésprit taquin. Une mention particulière pour l'équipe du 230 de m'avoir supporté tous ces trajets.

Je remercie aussi l'equipe MC2 pour sa passion de l'informatique, sa convivialité, ses potins arrachés à Nicolas, les bières partagées au foyer avec Michael et ses petits, ainsi que les nombreux restos ou Rosi, Tash et Nam ont bien voulu m'accompagner.

I would also like to say thank you to all the PhD students who had to live with my grumpiness for these three years: Alex, Andrea, Fionn, Guillaume, Nicolas, Phablo, Sebastian, Tim. Thank you Matthieu, Nam and Romain for being such great friends.

Finalement je tiens à remercier mes parents, mon frère, ainsi que toute ma famille pour m'avoir soutenu depuis toutes ces années.

Resumé

Le but principal de cette thèse est de présenter des conditions suffisantes pour garantir l'existence de subdivisions dans les graphes dirigés. Bien que ce genre de questions soit assez bien maitrisé dans le cas des graphes non orientés, très peu de résultats sont connus sur le sujet des graphes dirigés. La conjecture la plus célèbre du domaine est sans doute celle attribuée à Mader en 1985 qui dit qu'il existe une fonction f tel que tout graphe dirigé de degré sortant minimal supérieur à $f(k)$ contient le tournoi transitif sur k sommets comme subdivision. Cette question est toujours ouverte pour $k = 5$. Cette thèse présente quelques résultats intermédiaires tendant vers cette conjecture. Il y est d'abords question de montrer l'existence de subdivisions de graphes dirigés autre que les tournois, en particulier les arborescences entrantes. Il y a aussi la preuve que les graphes dirigés de grand degré sortant contiennent des immersions de grand tournois transitifs, question qui avait été posée en 2011 par DeVos et al. En regardant un autre paramètre, on montre aussi qu'un grand nombre chromatique permet de forcer des subdivisions de certains cycles orientés, ainsi que d'autre structures, pour des graphes dirigés fortement connexes.

Cette thèse présente également la preuve de la conjecture de Erdős-Sands-Sauer-Woodrow qui dit que les tournois dont les arcs peuvent être partitionnés en k graphes dirigés transitifs peuvent être dominé par un ensemble de sommet dont la taille dépend uniquement de k.

Pour finir, cette thèse présente la preuve de deux résultats, un sur l'orientation des hypergraphes et l'autre sur la coloration AVD, utilisant la technique de compression d'entropie.

Mots-clés: Graphes dirigés; Coloration; Subdivisions; Compression d'entropie

Abstract

The main purpose of the thesis was to exhibit sufficient conditions on digraphs to find subdivisions of complex structures. While this type of question is pretty well understood in the case of (undirected) graphs, few things are known for the case of directed graphs (also called digraphs). The most notorious conjecture is probably the one due to Mader in 1985. He asked if there exists a function f such that every digraph with minimum outdegree at least $f(k)$ contains a subdivision of the transitive tournament on k vertices. The conjecture is still wide open as even the existence of $f(5)$ remains open. This thesis presents some weakening of this conjecture. Among other results, we prove that digraphs with large minimum outdegree contain large in-arborescences. We also prove that digraphs with large minimum outdegree contain large transitive tournaments as immersions, which was conjectured by DeVos et al. in 2011. Changing the parameter, we also prove that large chromatic number can force subdivision of cycles and other structures in strongly connected digraphs.

This thesis also presents the proof of the Erdős-Sands-Sauer-Woodrow conjecture that states that the domination number of tournaments whose arc set can be partitioned into k transitive digraphs only depends on k. The conjecture, asked in 1982, was still open for $k = 3$.

Finally this thesis presents proofs for two results, one about orientation of hypergraphs and the other about AVD colouring using the recently developed probabilistic technique of entropy compression.

Keywords: Directed graphs; Colouring; Subdivisions; Entropy compression

Contents

Chapter 1

Introduction

This chapter serves as introduction to this thesis. After giving some standard definitions and notations in Section 1.1, in Sections 1.2, 1.3, 1.4, 1.5 we motivate and present the results obtained on subdivisions and other related structures in certain classes of digraphs. In Section 1.6, we discuss domination in tournaments and the proof of the Erdős-Sands-Sauer-Woodrow conjecture. Finally, Section 1.7 introduces the method of entropy compression and presents two results obtained using this method.

1.1 Definitions

We denote by $[j, k]$ the set of integers $\{j, \dots, k\}$ and $[k]$ the set $[1, k]$.

A *graph* is a pair $G = (V, E)$ of finite sets such that E is a set of unordered pairs of elements of V. A *multigraph* is a graph where we allow an edge to appear multiple times. We call these edges *parallel edges*, and the *multiplicity* of an edge is the number of times it appears in $E(G)$. The elements of V are the *vertices* and the elements of E are the *edges*. For a graph G, we write $V(G)$ for its vertex set and $E(G)$ for its edge set. Let u and v be vertices of V. If $\{u, v\}$ belongs to $E(G)$, we say that u and v are *adjacent*, *neighbours* or that u *sees* v, and we will use uv to denote the edge. Let $G = (V, E)$ be a graph and v a vertex of G. The *neighbourhood* of v, denoted by $N(v)$ is the set of vertices adjacent to v, and the *degree* of v is $d(v) = |N(v)|$. For a subset of vertices A, we write $N_A(v) = N(v) \cap A$ and set $d_A(v) = |N_A(v)|$. We also write $N(A) = \cup_{x \in A} N(x)$. We say that G has *minimum degree* at least k if every vertex v is such that $d(v) \geq k$. The *average degree* of G is the number $\frac{|E|}{|V|}$. Let $A \subseteq V$, we call *subgraph of G induced by A*, the graph $G[A] = (A, E \cap A^2)$. We denote by $G - A = G[V \setminus A]$. If $B \subseteq E$, then $G[B] = (V, B)$ and $G \setminus B = G[E \setminus B]$. The *complete graph* on k vertices, denoted by K_k is the graph on k vertices with every possible edge. A *stable set* S of G is a set of vertices such that $G[S]$ has no edge. A *clique* K of G is a set of vertices such that $G[K]$ is a complete graph. The *clique number*, denoted by $\omega(G)$, is the size of the largest clique in G and the *stability number*, denoted by $\alpha(G)$, is the size of its largest stable set.

A *path* P is a sequence of distinct vertices $x_1 \dots x_l$ such that $x_i x_{i+1}$ is an edge for every

$0 < i < l$. The vertices x_1 and x_l are called the *endvertices* of P and P is referred to as an x_1x_l-path, or a path between x_1 and x_l. The vertices x_i for $i \in [2, l-1]$ are referred to as the *inner vertices* of P. A set of paths $P_1, \ldots, P_k$ are said to be *internally disjoint* if the sets of inner vertices of these paths are pairwise disjoint. If S and T are two sets of vertices, a (S,T)-path is a st-path for some $s \in S$ and $t \in T$. A *cycle* C is a sequence of distinct vertices $x_1 \ldots x_l$ such that x_ix_{i+1} is an edge for every $0 < i < l$ and x_lx_1 is an edge. The *length* of a cycle is the number of vertices in this cycle. The *girth* of G is the length of its smallest cycle.

The *grid* of size k is the graph on $[k] \times [k]$ where vertex (i,j) is adjacent to vertices $(i-1,j)$, $(i+1,j)$, $(i,j-1)$ and $(i,j+1)$ when these vertices exist.

In a graph G, we say that vertices x and y have *connectivity* k if there exists k internally disjoint paths between x and y. A graph is *k-connected* if every pair of vertices has connectivity k and $|V| \geq k+1$. We write *connected* for 1-connected. Let X and Y be sets of vertices. A set of vertices S is an *(X,Y)-vertex-cut* if $G - S$ does not contain any $(X \setminus S, Y \setminus S)$-path. The following theorem due to Menger [52] relates the connectivity between vertices and vertex-cuts.

Theorem 1 (Menger's theorem [52])**.** *Let G be a graph, and let $S, T \subseteq V(G)$. The maximum number of vertex-disjoint (S,T)-paths is equal to the minimum size of an (S,T)-vertex-cut.*

A *proper k-colouring* of G is a function $c : V(G) \to [k]$, such that the vertices in $c^{-1}(i)$ form a stable set for every i. A graph G is said to be *k-colourable* if there exists a proper k-colouring of G. The *chromatic number* of G, denoted by $\chi(G)$ is the least k such that a proper k-colouring exists.

Let G and H be two graphs. We say that G admits a *subdivision* of H if there exist functions $\pi_1 : V(H) \to V(G)$ and π_2 mapping the edges of H to paths of G satisfying the following conditions:

- the map π_1 is an injection;
- for every edge $e \in E(H)$ between u and v, $\pi_2(e)$ is a $\pi_1(u)\pi_1(v)$-path; and
- for all distinct $e, e' \in E(H)$, $\pi_2(e)$ and $\pi_2(e')$ are internally disjoint.

We say that G admits an *immersion* of H if there exist functions $\pi_1 : V(H) \to V(G)$ and π_2 mapping the edges of H to paths of G satisfying the following conditions:

- the map π_1 is an injection;
- for every edge $e \in E(H)$ between u and v, $\pi_2(e)$ is a $\pi_1(u)\pi_1(v)$-path; and
- for all distinct $e, e' \in E(H)$, $\pi_2(e)$ and $\pi_2(e')$ have no edge in common.

In a subdivision or in an immersion, the vertices of $\pi_1(V(H))$ are called the *branching vertices* of H.

Let $e = xy$ be an edge of a graph $G = (V, E)$. By G/e we denote the graph obtained from G by *contracting* the edge e into a new vertex v_e, which is adjacent to all former neighbours of x and y. We say that G admits a *minor* of H if there exists a graph G' such that, G' can obtained from G by repeated contractions, and H is a subgraph of G'. These notions are strongly related as minor and immersion are both weaker versions of subdivision.

A *digraph* is a pair $D = (V, A)$ of finite sets such that A is a set of ordered pairs of elements of V. The elements of V are the *vertices* and the elements of A are the *arcs*. For a digraph D, we write $V(D)$ for its vertex set and $A(D)$ for its arc set. A *multidigraph* is a digraph where we allow an arc to appear multiple times. We call these arcs *parallel arcs*, and the *multiplicity* of an arc is the number of times it appears in $A(D)$. We say that a multidigraph has *multiplicity* k if all its arcs have multiplicity smaller than or equal to k. Let u and v be vertices of V. If (u, v) belongs to $A(D)$, we say that u *sees* v, u *dominates* v, v is *seen by* u or v is *dominated* by u. We also say that v is an *outneighbour* of u and u is an *inneighbour* of v. For an arc $a = (u, v)$, u is said to be the *tail* of a and v its *head*. Let $D = (V, A)$ be a digraph and X and Y be two subsets of V. Then $A(X, Y)$ represents the set of arcs of A with tail in X and head in Y. Let v be a vertex of some digraph D. The *inneighbourhood* of v, denoted by $N^-(v)$, is the set of inneighbours of v and the *indegree* of v, $d^-(v) = |N^-(v)|$ is its size. For a subset of vertices A, let $N^-_A(v) = N^-(v) \cap A$ and denote by $d^-_A(v) = |N^-_A(v)|$. We also write $N^-(A) = \cup_{x \in A} N^-(x)$. The *minimum indegree* (resp. *maximum indegree*) of a digraph D, denoted by $\delta^-(D)$ (resp. $\Delta^-(D)$), is the minimum (resp. maximum) value of $d^-(v)$ over all vertices of D. The *closed inneighbourhood* of v, denoted by $N^-[v]$ is defined as $N^-[v] = N^-(v) \cup \{v\}$. The *outneighbourhood* of v, denoted by $N^+(v)$, is the set of outneighbours of v and the *outdegree* of v, $d^+(v) = |N(v)|$ is its size. For a subset of vertices A, let $N^+_A(v) = N^+(v) \cap A$ and $d^+_A(v) = |N^+_A(v)|$. We also write $N^+(A) = \cup_{x \in A} N^+(x)$. The *minimum outdegree* (resp. *maximum outdegree*) of a digraph D, denoted by $\delta^+(D)$ (resp. $\Delta^+(D)$), is the minimum (resp. maximum) value of $d^+(v)$ over all vertices of D. The *closed outneighbourhood* of a v, denoted by $N^+[v]$, is defined as $N^+[v] = N^+(v) \cup \{v\}$. The *underlying graph* of a digraph D is the graph obtained from D by forgetting the orientation. The *degree* of a vertex v in a digraph D is the degree of v in the underlying graph of D. The *chromatic number* $\chi(D)$ of a digraph D is the chromatic number of its underlying graph. The *chromatic number* of a class of digraphs $\mathcal{D}$, denoted by $\chi(\mathcal{D})$, is the smallest k such that $\chi(D) \leq k$ for all $D \in \mathcal{D}$.

A digraph D is said to be *Eulerian* if $d^+(v) = d^-(v)$ for every vertex v.

A *tournament* on k vertices is an orientation of the complete graph on k vertices. The *transitive tournament* on k vertices, denoted by TT_k, is the acyclic tournament on k vertices. The *complete digraph* on k vertices, denoted by $\overleftrightarrow{K_k}$ is the digraph on k vertices with all possible arcs. A multidigraph is *complete* if the underlying graph is complete. Let $S \subseteq V$, we call *subdigraph of G induced by S*, the digraph $D[S] = (S, A \cap S^2)$. We denote by $D - S = D[V \setminus S]$. If $B \subseteq A$, then $D[B] = (V, B)$ and $D \setminus B = D[A \setminus B]$.

An *oriented path* P is a sequence of vertices together with arcs such that the underlying graph is a path on the same sequence. Let $P = (x_1, x_2, \cdots, x_n)$ be an oriented path. We say that P is an *(x_1, x_n)-path*. The vertex x_1 is the *initial vertex* of P and x_n its *terminal vertex*.

The vertices x_i for $i \in [2, l-1]$ are referred to as the *inner vertices* of P. A set of oriented paths $P_1, \ldots, P_k$ is said to be *internally disjoint* if the sets of inner vertices of these paths are pairwise disjoint. P is a *directed path* or simply a *dipath*, if $x_i x_{i+1} \in A(D)$ for all $i \in [n-1]$. Let P be a dipath, we denote by $s(P)$ its initial vertex and by $t(P)$ its terminal vertex. For any two vertices u and v, a (u,v)*-dipath* or *dipath from* u *to* v is a dipath P with $s(P) = u$ and $t(P) = v$. For two sets X, Y of vertices, an (X,Y)*-dipath* or *dipath from* X *to* Y is a dipath P such that $s(P) \in X$, $t(P) \in Y$, and no internal vertex is in $X \cup Y$. Let x and y be two vertices of D, an (x,y)*-walk* is a non-empty sequence of vertices $v_0, \ldots, v_l$ such that, $x = v_0$, $y = v_l$ and $(v_i, v_{i+1}) \in A(D)$ for all $i \in [l-1]$. Note that the same vertex can appear multiple times. An *oriented cycle* C is a set of vertices such that the underlying graph is a cycle. An oriented cycle $x_1, \ldots x_l$ of a digraph D is said to be *directed* if $(x_i, x_{i+1} \in A(D)$ for $0 < i < l$ and $(x_l, x_1) \in A(D)$. For convenience, we will sometimes use path or cycle to talk about oriented path or oriented cycles. A digraph D is said to be *acyclic* if it doesn't contain a directed cycle as a subdigraph. In a digraph D, a directed cycle C of a dipath P is said to be *Hamiltonian* if it contains all vertices of D.

A *block* of a path P or a cycle C is a maximal directed subpath of P or of C. A path is entirely determined by the sequence $(b_1, \ldots, b_p)$ of the lengths of its blocks and the sign $+$ or $-$ indicating if the first arc is forward or backward respectively. Therefore we denote by $P^+(b_1, \ldots, b_p)$ (resp. $P^-(b_1, \ldots, b_p)$) the oriented path whose first arc is forward (resp. backward) with p blocks, such that the ith block along it has length b_i.

An *out-arborescence* is a rooted tree where all the arcs are oriented away from the root and a k*-out-arborescence* is an out-arborescence in which every vertex apart form the leaves has outdegree k. An *in-arborescence* is a tree where all the arcs are oriented towards the root and a k*-in-arborescence* is an in-arborescence in which every vertex apart form the leaves has indegree k. We use *arborescence* to talk about an in- or out-arborescence. The *depth* of an arborscence is the length of its longest dipath. Let x be a vertex of an arborescence T. The *ancestors* of x is the set of vertices on the path from the root to x and the *father* of x is the last vertex (towards x) on this path.

Let D and H be two digraphs, we say that D admits a *subdivision* of H if there exist functions $\pi_1 : V(H) \to V(G)$ and π_2 mapping the arcs of H to dipaths of D satisfying the following conditions:

- the map π_1 is an injection;
- for every arc $e \in A(H)$ from u to v, $\pi_2(e)$ is a $(\pi_1(u), \pi_1(v))$- dipath; and
- for all distinct $e, e' \in A(H)$, $\pi_2(e)$ and $\pi_2(e')$ are internally disjoint.

We say that D admits an *immersion* of H if there exist functions $\pi_1 : V(H) \to V(G)$ and π_2 mapping the arcs of H to dipaths of D satisfying the following conditions:

- the map π_1 is an injection;
- for every arc $e \in A(H)$ from u to v, $\pi_2(e)$ is a $\pi_1(u)\pi_1(v)$- dipath; and

- for all distinct $e, e' \in A(H)$, $\pi_2(e)$ and $\pi_2(e')$ are arc-disjoint.

In a digraph D, we say that vertices x and y have *connectivity* k if there exists k vertex-disjoint dipaths betweeen x and y. We say that vertices x and y have *arc-connectivity* k if there exists k arc-disjoint paths betweeen x and y. A digraph is *k-strongly-connected* if every pair of vertices has connectivity k. We write *strongly connected* or *strong* for 1-strongly-connected. Let X and Y be sets of vertices. A set S of vertices is an *(X, Y)-vertex-cut* if $D - S$ does not contain any $(X \setminus S, Y \setminus S)$-dipath. A set A of arcs is an *(X, Y)-arc-cut* if $D \setminus A$ does not contain any (X, Y)-dipath.

Theorem 2 (Menger's theorem [52]). *Let D be a digraph, and let $S, T \subseteq V(D)$. The maximum number of vertex-disjoint (S, T)-dipaths is equal to the minimum size of an (S, T)-vertex-cut.*

The same result holds for arc-connectivity

Theorem 3 (Menger's theorem [52]). *Let D be a digraph, and let $S, T \subseteq V(D)$. The maximum number of arc-disjoint (S, T)-dipaths is equal to the minimum size of an (S, T)-arc-cut.*

A *hypergraph* is a pair $\mathcal{H} = (S, E)$ of disjoint sets, where the elements of E are non-empty subsets of S. The set S is called the *ground set* of $\mathcal{H}$. The elements of E are called the *hyperedges* of $\mathcal{H}$. A hypergraph $\mathcal{H}$ is *k-uniform* if every hyperedge has cardinality k.

Let $\mathcal{H}$ be a hypergraph on a ground set S and S' be a subset of S. The *projection* of a hyperedge E onto S' is defined as $E \cap S'$. Let $\mathcal{H} \cap S'$ denotes the hypergraph defined on S' where the edges are the projections of the edges of $\mathcal{H}$ onto S'.

Let $D = (V, A)$ be a digraph, the *inneighbourhood hypergraph* of D is the hypergraph $\mathcal{H} = (V, E)$, where $E = \{N^-[x]; x \in V\}$.

1.2 Structures in graphs

What are the properties of a graph forbidding a certain structure? This is a very natural setting in graph theory under which many interesting conjectures and results can be rephrased. One of the classical results in Graph Theory is Kuratowski's Theorem, saying that planar graphs are precisely the graphs avoiding K_5 and $K_{3,3}$ as minors. Generalising this result, the celebrated Robertson-Seymour Theorem states:

Theorem 4 (Robertson-Seymour [60]). *Every class of graphs closed under the minor relation can be defined by a finite set of forbidden minors.*

This theorem means that in order to understand a minor-closed class, it is sufficient to understand graphs avoiding a certain set of obstructions as minor. A very important consequence of this theorem and its proof is the development of the graph decomposition technique which led to important results in both structural and algorithmic graph theory. The main parameter in this field is the tree-width, where graphs with small tree-width behave like trees. On the other hand, if the graph has large tree-width, the following theorem allows us to understand some of its structure:

Theorem 5 (Grid Minor Theorem, Robertson-Seymour [58]). *For every k, there exists a function $f(k)$ such that any graph with tree-width at least $f(k)$ contains a grid of size k as a minor.*

This gives a template for solving many problems: either the graph has small tree-width and the problem can be solved easily (most of the time using a dynamic programming argument), or it contains a large grid as a minor and then we can solve the problem using this specific structure. Probably the most important example of a proof using this technique is the k-linkage problem. An instance of this problem is a graph G and a set of k pairs of vertices (s_i, t_i) and the goal is to decide if there are k internally disjoint paths of G joining every possible pair (s_i, t_i). In [59], Robertson and Seymour gave a cubic-time algorithm for this problem using the above mentioned strategy, this was later improved to quadratic time.

Even outside this setting of graph decomposition, the question of what parameter can force a certain structure in a graph is natural on its own. The most famous conjecture concerning properties of graphs with a forbidden minor is Hadwiger's Conjecture, which tries to generalise the 4-Colour Theorem:

Conjecture 6 (Hadwiger, 1943). For every $t \geq 1$, every graph without K_{t+1} as a minor is t-colourable.

The case $t = 5$ was proved by Robertson, Seymour and Thomas [61] but it remains open for $t > 5$. On the other hand, it was independently proved in 1984 by Kostochka and Thomason [44, 68] that a graph without K_t as a minor is $O(t(logt)^{1/2})$-colourable. In fact they showed the following stronger theorem:

Theorem 7 (Kostochka [44] and Thomason [68]). *For every $t \geq 1$, every graph with minimum degree greater than $\Theta(t(logt)^{1/2})$ contains K_t as a minor.*

Understanding graphs with a large minimum degree has also been studied beyond this link with Hadwiger's Conjecture. The following beautiful result due to Mader [47] in 1967 is probably the most famous example.

Theorem 8 (Mader [47]). *For every $k \geq 1$, there exists an integer $f(k)$ such that every graph with minimum degree at least $f(k)$ contains a subdivision of the complete graph on k vertices.*

Mader, Erdős and Hajnal, conjectured that the right value for $f(k)$ should be in $O(k^2)$. Bollobás and Thomason [8] as well as Komlós and Szemerédi [43] proved this conjecture 30 years later.

Another related result was obtained recently by DeVos et al. [19]. Studying a variant of Hadwiger's Conjecture that deals with immersions instead of minors, they showed that graphs with minimum degree greater than $200t$ contain an immersion of the complete graph on t vertices, proving the first linear bound for these kind of questions. Proving a linear bound for Hadwiger's Conjecture would be a major breakthrough.

Developing a decomposition theory for digraphs is a major challenge in computer science because, in many applications, the most natural model is a digraph. However, apart from a recent breakthrough due to Kawarabayashi and Kreutzer [38] who proved in a very technical paper the directed version of the Grid Minor Theorem, not much is known. There are many reasons for that, one being that the directed versions of the problems are inherently more difficult than their undirected version. This difficulty is particularly clear for the linkage problem. An instance of the directed k-linkage problem is a digraph D and a set of k pairs of vertices (s_i, t_i) and the goal is to decide if there are k internally disjoint dipaths of D joining every possible pair (s_i, t_i). While the k-linkage problem is solvable in quadratic time for every k, the directed version was proved to be NP-hard for $k = 2$ in [29]. In fact, Thomassen proved in [71], that the problem is NP-hard for $k = 2$ even in the case of digraphs with arbitrarily high connectivity. When introducing the directed tree-width, Reed [57] and Johnson, Robertson, Seymour and Thomas [36] chose a definition such that the k-linkage is polynomial-time solvable on digraphs with bounded directed-tree-width, but the notion is hard to handle. Another reason is our lack of understanding concerning the structural aspect of the problem. The biggest example is probably the following conjecture due to Mader [49] in 1985, which is still open for $k \geq 5$.

Conjecture 9 (Mader [49])**.** For every $k \geq 1$, there exists an integer $f(k)$ such that every digraph with minimum outdegree at least $f(k)$ contains a subdivision of TT_k, the transitive tournament on k vertices.

A large part of this thesis is dedicated to understand some structural questions related to Mader's Conjecture.

1.3 Subdivisions of cliques and local connectivity

Before discussing Conjecture 9, we will talk about the proof of the statement in the undirected case. The first thing to note, is that in any graph with average degree greater than k, there exists a subgraph with minimum degree greater than $k/2$. In fact Mader in [48] proved that large average degree forces the existence of a subgraph with large connectivity.

Theorem 10 (Mader [48])**.** *Any graph G with average degree greater than $4k$ contains a k-connected subgraph.*

Proof. For inductive purposes we will prove the following stronger result. Every graph G on $n \geq 2k-1$ vertices and $m > (2k-3)(n-k+1)$ edges has a k-connected subgraph. Note that the condition $m > (2k-3)(n-k+1)$ implies for large value of n a average degree of roughly $4k$. The proof is by induction on n. If $n = 2k-1$, then G must be the complete graph, so we can assume $n \geq 2k$. If a vertex v satisfies $d(v) < 2k-3$, we can apply induction on $G-v$ and we are done. Thus $\delta(G) > 2k-2$. Assume G is not k-connected. By Theorem 1, there exists a vertex-cut X of less than k vertices. Call V_1 and V_2 the two components of $G-X$

and $G_i = G[V_i \cup X]$, for $i \in [2]$. If the G_i are not k-connected, then by induction hypothesis on each G_i: $e(G_i) < (2k-3)(|G_i| - k + 1)$. However, since $|G_1 \cap G_2| = |X| < k$, we get that $e(G) \le e(G_1) + e(G_2) \le (2k-3)(n-k+1)$, a contradiction. □

For a pair of vertices with high connectivity, the following tight result was proved by Mader [47].

Theorem 11 (Mader [47]). *Any graph G with minimum degree k contains a pair of vertices with connectivity k.*

The proofs that Mader gave of Theorem 11 and Theorem 8 both use the same idea of induction where the graph is partitioned into a clique and the rest and the induction will go on by contracting vertices onto the clique, while ensuring that the rest of the graph keeps the right properties. We now present theses proofs.

Let $K = \{x_1, \dots, x_k\}$ be a clique in a graph G. For every $v \in N(x_1) \setminus \cap_{i=1}^{k} N(x_k)$, let $\phi(v)$ be the smallest integer such that $x_{\phi(v)} v \notin E(G)$. Let G_K be the graph obtained from G by removing x_1 and adding for each vertex $v \in N(x_1) \setminus \cap_{i=1}^{k} N(x_k)$ the edge $vx_{\phi(v)}$. The following lemma is the key part of the proof of Theorem 11. The idea is that, if K is a maximal clique, then doing the transformation described just above, we can remove one vertex of the clique, while preserving the connectivity between vertices outside of K.

Lemma 12 (Mader [47]). *For every $k \ge 1$, if $K = \{x_1, \dots, x_k\}$ is a maximal clique, then G_K is such that:*

- *the vertices in $V(G) \setminus K$ have the same degree in G and G_K;*
- *the connectivity between two vertices in $V(G) \setminus K$ is the same in G and G_K.*

Proof. The first property is by construction of $\phi(v)$, since $\cap_{i=1}^{k} N(x_i)$ is empty because K is a maximal clique.

Let u and v be vertices in $V(G) \setminus K$ and suppose there exist l disjoint paths $P_1, \dots, P_l$ in G_K between u and v. Without loss of generality, we can assume that all these paths use vertices of K. This means there exist l paths $Q_1, \dots, Q_l$ of G_K from u to K intersecting only on u, and l paths $S_1, \dots S_l$ of G_K from v to K intersecting only on v. Moreover, for every $i \in [l]$ and $j \in [l]$ (possibly the same), the path Q_i and the path S_j can only intersect on K. Note that we can assume that $Q_i \cap K$ and $S_i \cap K$ are reduced to a single vertex for every $i \in [l]$. So let $u_i = K \cap Q_i$ and $v_i = K \cap S_i$ denote these vertices. Without loss of generality, we can assume that the u_i and the v_i are ordered on K, meaning that if $i < j$ and $u_i = x_{l(i)}$ and $u_j = x_{l(j)}$, then $l(i) < l(j)$. Finding two sets of paths $Q'_1, \dots, Q'_l$ and $S'_1, \dots, S'_l$ of G with the exact same properties would prove the lemma as one can easily use edges in K to close the paths between u and v.

We will obtain each path Q'_i of G from u to u'_i by changing only the last vertex of Q_i as follows: If the last edge of Q_i belongs to G then $u'_i = u_i$ and $Q'_i = Q_i$. If this edge doesn't belong to G, it means by construction of G_K that this edge is an edge $tx_{\phi(t)}$ for some $t \in V(G) \setminus K$. However, by definition of $\phi(t)$, it means that t is adjacent in G to all the vertices x_r with $r < \phi(t)$ (including x_1). In particular, if there exists u_j with $j < i$ such that the last edge of Q_j is not in G, then $tu_j \in E(G)$. So in this case u'_i is the biggest u_j with $j < i$ such that the last edge of Q_i is not in G, if it exists and x_0 if not. By doing the same with the S_i between v and K and using edges of K to connect them, we find l disjoint paths between u and v in G. □

To prove Theorem 11 Mader proved the following stronger result.

Theorem 13 (Mader [47]). *For every $k \geq 1$, let G be a graph with a clique K such that $|V(G) \setminus K| \geq 2$ and every vertex of $V(G) \setminus K$ has degree at least k in G. Then there exists a pair of vertices of $V(G) \setminus K$ with connectivity k in G.*

Proof. The proof is by induction on the order of G. Note that, no matter the order of G, if $|V(G) \setminus K| = 2$, then the result is trivially true. This makes the case $|G| = k + 2$ trivial. Suppose now that $|G|$ is larger than $k + 2$. If K is maximal, then using the previous lemma and the induction hypothesis on G_K gives the result. If K is not maximal, then we can add vertices to K to make it maximal. If during this procedure, $|V(G) \setminus K| \leq 2$, then we can conclude. If not we can apply induction on G_K. □

The proof of Theorem 8 is a bit different, as we will sometimes need to remove vertices of a clique which is not maximal, thus we cannot guarantee the degree of every vertex in $V(G) \setminus K$. However by doing an induction on k, we will be able to guarantee a large average degree.

The following lemma, that we do not prove here, is quite similar to Lemma 12:

Lemma 14 (Mader [47]). *For every $l \geq 1$, let G be a graph and K a clique in G. Suppose there exists a subdivision of K_l in G_K with all branching vertices $\{x_1, \ldots, x_l\}$ in $V(G) \setminus K$ and such that paths of the subdivision between vertices in $\{x_2, \ldots, x_l\}$ only use vertices in $V(G) \setminus K$. Then there exists a subdivision of K_l in G with branching points $\{x_1, \ldots, x_l\}$ in $V(G) \setminus K$ and such that paths of the subdivision between vertices in $\{x_2, \ldots, x_l\}$ only use vertices in $V(G) \setminus K$.*

Proof of Theorem 8. We will prove by induction on p that a graph with average degree 3^p contains K_p as a subdivision. The case $p = 2$ is trivial. Suppose now that the result is true for some p. We will prove it for $p + 1$. We will prove by induction on the size of G the following:

Let G be a graph of average degree greater than 3^{p+1} and $K = \{x_1, \ldots, x_l\}$ a clique of size smaller than $p + 2$. Then G contains a subdivision of K_{p+1} with all the branching vertices $\{x_1, \ldots, x_{p+1}\}$ in $V(G) \setminus K$ and such that that paths of the subdivision between vertices in $\{x_2, \ldots, x_{p+1}\}$ only use vertices in $V(G) \setminus K$.

Let $n = |G|$. Again, the case where $n = 3^{p+1} + 1$ is trivial. If K is maximal, then applying the induction hypothesis on the size of G to G_K works as the number of edges of G_K is equal to

$$E(G) - (p + 1) \geq \frac{3^{p+1}}{2} \times n - (p + 1) > \frac{3^{p+1}}{2} \times (n - 1)$$

and thus the average degree is greater than 3^{p+1}. This allows us to conclude using Lemma 14. Suppose now that K is not maximal. By adding some vertices to it we can assume that $|K| = p + 1$. Consider the set of vertices $X = \cap_{i=1}^{k} N(x_l)$. If the subgraph $G[X]$ has minimum degree $3^p + 1$, then we can apply the induction hypothesis on p to $G[X]$ to find a subdivision of K_p in $X \setminus \{x\}$ for some vertex $x \in X$ and extend it to a subdivision of K_{p+1} using x and edges with K. Suppose $G[X]$ has minimum degree smaller than $3^p + 1$ and let $s \in X$ be a vertex of degree at most 3^p in $G[X]$. If we add s to K, then the difference in the number of edges between G and G_K is precisely $|N_{X \cup K}(s)| \leq p + 1 + 3^p$. Since

$$\frac{3^{p+1}}{2} \times n - (p+1) - 3^p > \frac{3^{p+1}}{2} \times (n-1),$$

G_K has average degree greater than 3^{p+1} and we can conclude by applying the induction hypothesis on the size of G to G_K and using Lemma 14. □

1.4 Subdivisions and immersions in digraphs with large minimum outdegree

When one tries to generalise Theorem 8 to digraphs, the first natural question is to ask if there exists a function $f(k)$ such that every digraph with minimum outdegree at least $f(k)$ contains a subdivision of the complete digraph on k vertices. However the following construction, due to DeVos et al. [20](slightly adapted), shows the existence of digraphs with arbitrarily large outdegree and without two arc-disjoint directed cycles between the same pair of vertices. In particular it cannot contain $\overleftrightarrow{K_3}$ as a subdivision. Let D_k be the digraph obtained from a k-out-arborescence of depth $k+1$ by adding to each leaf the set of its ancestors as outneighbours. This is indeed a digraph with minimum outdegree at least k. Let x and y be two vertices of D_k. Without loss of generality, we can assume that x is deeper than y in the arborescence. Hence every cycle containing x and y must use the arc from the father of y to y, proving the fact that there cannot exist two arc-disjoint directed cycles going through both vertices.

However, for transitive tournaments, the question remains open. In [49] Mader made the following conjecture.

Conjecture 9 (Mader [49])**.** For every $k \geq 1$, there exists an integer $f(k)$ such that every digraph with minimum outdegree at least $f(k)$ contains a subdivision of TT_k, the transitive tournament on k vertices.

The question turned out to be way more difficult than the non oriented case, as the existence of $f(5)$ remains unknown. Weakening the statement, DeVos, McDonald, Mohar and Scheide [20] made the following conjecture replacing subdivision with immersion and proved it for the case of Eulerian digraphs.

Conjecture 15 (DeVos et al. [20])**.** For every $k \geq 1$, there exists an integer $h(k)$ such that every digraph with minimum outdegree at least $h(k)$ contains an immersion of TT_k.

Note that the construction D_k given above also prove that, even in the case of immersions, $\overleftrightarrow{K_3}$ cannot be forced through minimum outdegree.

An important part of the proof of Mader and the one of Bollobás and Thomason about subdivision of the complete graph is the fact that one can easily find a subgraph with large connectivity inside a graph with large average degree. In digaphs however, the picture is completely different, as the complete bipartite digraph with all arcs oriented in the same direction is an example of a digraph with large average outdegree and without a subdigraph with large minimum outdegree. An interesting remark is that for Eulerian digraphs, a similar result was obtained in [20], namely that every Eulerian digraph with large outdegree contains an immersion of an Eulerian digraph with large arc-connectivity.

Another useful tool for finding subdivisions of the complete graph is to find a subgraph with good linkage property. A graph or a digraph is said to be *k-linked* if every instance of the k-linkage problem admits a solution. Thomas and Wollan in [67] proved the following.

Theorem 16 (Thomas and Wollan [67]). *Every $10k$-connected graph is k-linked.*

However Thomassen in [71] proved the existence of digraphs with arbitrarily large connectivity which are not 2-linked.

Another example about the difficulties of understanding digraphs with large minimum degree is the following conjecture due to Alon [4] and still open for $k = 2$.

Conjecture 17 (Alon [4]). There exists a function $f(k)$ such that every digraph with minimum degree at least $f(k)$ can be partitioned into two digraphs of minimum outdegree at least k.

The existence of $f(1)$ means that one can find two disjoint directed cycles in digraphs with large minimum outdegree. This was proved by Thomassen in [69]. Even showing that large minimum outdegree digraphs can be partitioned into a directed cycles and a digraph with minimum outdegree 2 is open. Note that the existence of $f(2)$ would imply the existence of $f(k)$ for every k. Indeed, suppose $f(2)$ is known and let D be a digraph with minimum outdegree $f(2)^{k+1}$. For each vertex v, we partition $N^+(v)$ into $f(2)^k$ sets $S_1(v), \dots, S_{f(2)^k}(v)$, each of size $f(2)$. Consider the digraph H obtained from D by replacing every vertex v by an $f(2)$-out-arborescence of depth k, $T(v)$, and where the *ith* leaf of $T(v)$ is linked to the roots of the arborescence associated to the vertices in $S_i(v)$. Note that contracting each $f(k)$-out-arborescence into one vertex yields exactly the digraph D. So H can be seen as the digraph D where every vertex is blown-up into an $f(k)$-out-arborescence and the outdegree of each vertex is distributed to the leaves of the out-arborescence replacing it. H is a digraph with minimum outdegree $f(2)$, so there exists a partition of it into two digraphs of outdegree 2. It is easy to see that, by looking at the roots of the out-arborescences, this corresponds to a partition of D into two digraphs of minimum outdegree 2^{k+1}.

A very nice result about digraphs with large minimum outdegree is the existence of vertices with large connectivity. Remember that in graphs with minimum degree d there exists a pair of vertices of connectivity d and the proof of this fact uses ideas similar to the one of Theorem 8. Such a tight result is not possible in digraphs, as Mader showed in [49] the existence of digraphs with minimum outdegree $12m$ and maximum connectivity $11m$. However he showed in [50] the existence of a function $f(k)$ such that every digraph with minimum outdegree greater than $f(k)$ contains a pair of vertices with connectivity k. Despite being relatively short, the proof is quite difficult to understand, in part because of the optimisation of the function $f(k)$. We will present our understanding of the proof for arc-connectivity. First because, despite being simpler, it uses most of the very nice ideas of Mader for the verxtex-connectivity and because it served us as base to solve Conjecture 15 in [45].

Theorem 18. *For every $k \geq 1$, there exists an integer $f(k)$ such that every digraph with minimum outdegree at least $f(k)$ contains a pair of vertices with arc-connectivity at least k.*

Proof. We will prove by contradiction that every digraph D where every vertex but at most k^2 has outdegree greater than k^4 contains a pair of vertices with arc-connectivity k. Let D be the smallest, in term of arcs and vertices, counterexample. By minimality of D, all the vertices have outdegree exactly k^4 except the k^2 with outdegree 0. Let T be the set of vertices with outdegree 0. Pick a vertex v with outdegree k^4. For every vertex $y \in D$ there does not exist a set of k arc-disjoint directed paths from v to y. Hence, by Theorem 3, there exists a set E_y of less than k arcs such that there is no directed path from v to y in $D \setminus E_y$. For every $y \in D - v$, define C_y as the set of vertices which can reach y in $D \setminus E_y$. Now take Y a minimal set such that $\cup_{y \in Y} C_y$ covers $D - y$. Y consists of at least k^3 elements as $\cup_{y \in Y} E_y$ contains all the arcs of D with tail v.

For each $y \in Y$, define S_y as the set of vertices which belong to C_y and no other $C_{y'}$ for $y' \in Y$, $y' \neq y$. Since Y is minimal, every S_y is non-empty. Note that for $u \in S_y$, if there exists $y' \in Y \setminus y$ and $v \in C_{y'}$ such that $(u, v) \in A(D)$, then $(u, v) \in E_{y'}$. Note that $T \subseteq Y$ as vertices in T have outdegree 0 and if $y \in Y \setminus T$ then S_y consists only of vertices of outdegree k^4 in D.

Let R be the digraph with vertex set Y and arcs the pairs (y, y') such that there is an arc from S_y to $C_{y'}$. As noted before, $d^-_R(y) \leq |E_y| \leq k$. The average outdegree of the vertices of $Y \setminus T$ in R is then at most $\frac{k \times k^3}{k^3 - k^2} < k^2 - 1$. Let y be a vertex of $R \setminus T$ with outdegree at most this average. Consider the digraph induced by $S_y \cup v$. As noted before the arcs of $A(S_y, \bar{S}_y)$ are precisely the arcs contributing to the outdegree of y in R. Therefore, $D[S_y \cup v]$ is a digraph strictly smaller than D, without two vertices with arc-connectivity k and such that apart from k^2 vertices, every vertex has outdegree k^4, a contradiction. □

Let $F(k, l)$ be the multidigraph consisting of k vertices $x_1, \ldots, x_k$ and l arcs from x_i to x_{i+1} for every $1 \leq i \leq k - 1$. It is clear that $F(k, \binom{k}{2})$ contains an immersion of TT_k, so extending the previous proof to $F(k, l)$ would prove Conjecture 15. The main difficulty to do that, if one removes the arcs of a subdivision of $F(k-1, l)$ and tries to extend it using the same proof as above, is that there is no control on the number of vertices that lose some outdegree. The way around this problem is to remark that, using the notations of last proof, you can add new arcs in $D[S_y \cup v]$ between vertices which belonged to same directed path of the copy of the subdivision of $F(k-1, l)$, so that each of the $(k-1)l$ directed paths of the subdivision of $F(k-1, l)$ contributes to the loss of outdegree of at most one vertex. With this, we are able to control the number of vertices that lose some outdegree and proceed with the induction. The complete proof is presented in Chapter 3.

Trying to extend this technique to obtain the result for subdivisions seems hard, as it looks like the removal of the vertices on the directed paths disturbs the structure of the digraph too much.

Overall not much is known about Conjecture 9. In [41], Kühn et al. proved that a digraph of order n whose minimum outdegree is at least d contains a subdivision of a complete digraph of order $\frac{d^2}{8n^{3/2}}$. In [1], with several co-authors, we looked at subdivisions of digraphs different from the transitive tournament. The first thing we did was to give a tight bound for orientations of paths.

Theorem 19. *Let $(k_1, k_2, \ldots, k_\ell)$ be a sequence of positive integers, and let D be a digraph with $\delta^+(D) \geq \sum_{i=1}^{\ell} k_i$. For every $v \in V(D)$, D contains a path $P^+(k'_1, k'_2, \ldots, k'_\ell)$ with initial vertex v such that $k'_i \geq k_i$ if i is odd, and $k'_i = k_i$ otherwise.*

The bound is indeed thight as the complete digraph on k vertices has minimum outdegree $k-1$ and contains no path on more than k vertices.

The next natural digraphs to look at are trees. It is easy to see that, by a greedy algorithm, finding a large out-arborescence in a digraph with large minimum outdegree is trivial. The question of in-arborescence is a however interesting as one needs to control the vertices with large indegree. The main result of paper [1] was to prove the existence of large in-arborescences in digraphs with sufficiently large minimum outdegree.

Theorem 20. *Let F be an in-arborescence. There exists a constant $C(F)$ such that every digraph with minimum outdegree at least $C(F)$ contains a subdivision of F.*

The proofs of the two previous theorems are presented in Chapter 3 together with some other results. Finding subdivisions remains however open for oriented trees in general. We believe it would be an interesting step towards Conjecture 9.

Problem 21. Let T be an oriented tree. Does there exist a constant $a(T)$ such that any digraph with minimum outdegree at least $a(T)$ contains T as a subdivision.

One interesting problem while attacking Conjecture 9 is that it is hard to find a strengthening of the conjecture which makes sense. For example, requiring that D has large in and outdegree, despite making the question trivial for trees does not add anything for the transitive tournament. Indeed suppose that there exists a function $g(k)$ such every digraph with minimum in and outdegree greater than $g(k)$ contains a subdivision of TT_k. Let D be a digraph with outdegree greater than $g(k)$, and $\bar{D}$ be the digraph obtained from D by reversing all the arcs. Consider H obtained by taking a disjoint copy of both D and $\bar{D}$ and adding all arcs from $\bar{D}$ to D. H is a digraph with minimum in and outdegree greater than $g(k)$ and thus contains a subdivision of TT_k. Without loss of generality, we can assume half of the branching vertices of this subdivision belong to D and thus D contains a subdivision of $TT_{k/2}$.

1.5 Subdivisions in digraphs with large (di)chromatic number

What can we say about the subgraphs of a graph with large chromatic number? Of course, one way for a graph to have large chromatic number is to contain a large complete subgraph. However, if we consider graphs with large chromatic number and small clique number, then we can ask what other subgraphs must occur. We can avoid any graph that contains a cycle because, as proved by Erdős [22], there are graphs with arbitrarily high girth and chromatic number. Reciprocally, one can easily show that every graph with chromatic number n contains every tree of order n as a subgraph.

The following more general question attracted lots of attention.

Problem 22. Which are the graph classes $\mathcal{G}$ such that every graph with sufficiently large chromatic number contains an element of $\mathcal{G}$?

If such a class is finite, then it must contain a tree, by the above-mentioned result of Erdős. If it is infinite however, it does not necessary contain a tree. For example, every graph with chromatic number at least 3 contains an odd cycle. This was strengthened by Erdős and Hajnal [23] who proved that every graph with chromatic number at least k contains an odd cycle of length at least k. A counterpart of this theorem for even length was settled by Mihók and Schiermeyer [53]: every graph with chromatic number at least k contains an even cycle of length at least k. Further results on graphs with prescribed lengths of cycles have been obtained [32, 53, 73, 46].

During this thesis, we looked at similar problems for directed graphs: which are the digraph classes $\mathcal{D}$ such that every digraph with sufficiently large chromatic number contains an element of $\mathcal{D}$? Let us denote by $\text{Forb}(H)$ (resp. $\text{Forb}(\mathcal{H})$) the class of digraphs that do not contain H (resp. any element of $\mathcal{H}$) as a subdigraph. The previous question can be reformulated as follows:

Problem 23. Which are the classes of digraphs $\mathcal{D}$ such that $\chi(\text{Forb}(\mathcal{D}))$ is finite?

If $\mathcal{D}$ is a simple digraph, then $\chi(\text{Forb}(\mathcal{D})) < +\infty$ only if D is an oriented tree. Burr [13] proved that every $(k-1)^2$-chromatic digraph contains every oriented tree of order k. This was slightly improved by Addario-Berry et al. [2] who proved the following.

Theorem 24 (Addario-Berry et al. [2])**.** *Every $(k^2/2 - k/2 + 1)$-chromatic digraph contains every oriented tree of order k. In other words, for every oriented tree T of order k, $\chi(\text{Forb}(T)) \leq k^2/2 - k/2$.*

Burr conjectured that the right bound should be linear:

Conjecture 25 (Burr [13])**.** Every $(2k-2)$-chromatic digraph D contains a copy of any oriented tree T of order k.

This question appears to be extremely difficult, as even the case of tournaments which can be seen as the easiest case of n-chromatic digraphs, was only recently proved by Kühn et al. [40].

For special oriented trees T, better bounds on the chromatic number of $\mathrm{Forb}(T)$ are known. The most famous one, known as the Gallai-Roy Theorem, can be seen as a variant of Dilworth's Theorem on orders:

Theorem 26 (Gallai [30], Hasse [34], Roy [62], Vitaver [72])**.** $\chi(\mathrm{Forb}(P^+(k))) = k$.

The chromatic number of the class of digraphs not containing a prescribed oriented path with two blocks (blocks are maximal directed subpaths) has been determined by Addario-Berry et al. [3].

Theorem 27 (Addario-Berry et al. [3])**.** *Let P be an oriented path with two blocks on k vertices.*

- *If $k = 3$, then $\chi(\mathrm{Forb}(P)) = 3$.*
- *If $k \geq 4$, then $\chi(\mathrm{Forb}(P)) = k - 1$.*

One interesting case is the case when $\mathcal{D}$ is the set of all subdvisions of a certain digraph. Let us denote by S-Forb(D) (resp. S-Forb($\mathcal{D}$)) the class of digraphs that contain no subdivision of D (resp. any element of $\mathcal{D}$) as a subdigraph. Note that in the case of undirected graphs, as every graph with chromatic number greater than k contains a graph with minimum degree greater than k, Theorem 8 implies the following.

Theorem 28. *For every graph H, there exists a constant $C(H)$ such that any graph with chromatic number greater than $C(H)$ contains a subdivison of H.*

Consider the case where $\mathcal{D}$ is the set of all subdivisions of a fixed cycle. Let us denote by $\vec{C}_k$ the directed cycle of length k. For all k, transitive tournaments provide examples of digraphs with arbitrarily large chromatic number and without $\vec{C}_k$ as a subdivision. More generally, the following construction communicated to us by Nešetřil generalise this result to any oriented cycle.

Theorem 29. *For any positive integers b, c, there exists an acyclic digraph D with $\chi(D) \geq c$ in which all oriented cycles have more than b blocks.*

Proof. By [22], there exist graphs with chromatic number c and girth greater than cb. Let G be such a graph and consider a proper c-colouring ϕ of it. Let D be the acyclic orientation of G in which an edge uv of G is oriented from u to v if and only if $\phi(u) < \phi(v)$. By construction, the length of all directed paths in D is less than c and since each oriented cycle of D has length more than cb, they all have more than b blocks. □

This directly implies the following theorem.

Theorem 30. *For any finite family $\mathcal{C}$ of oriented cycles,*

$$\chi(\text{S-Forb}(\mathcal{C})) = \infty.$$

On the other hand, considering strongly connected digraphs may lead to dramatically different results. An example is provided by the following celebrated result due to Bondy [9] : *every strong digraph of chromatic number at least k contains a directed cycle of length at least k.* Denoting the class of strong digraphs by $\mathcal{S}$, this result can be rephrased as follows.

Theorem 31 (Bondy [9])**.** $\chi(\text{S-Forb}(\vec{C}_k) \cap \mathcal{S}) = k - 1$.

The proof of this result is particularly elegant as it starts with the longest cycle where each vertex gets one different colour and extend this colouring to the rest of the digraph.

One natural question is to extend Bondy's result to other oriented cycles.

Problem 32. Let C be an oriented cycle. Is $\chi(\text{S-Forb}(C) \cap \mathcal{S})$ bounded?

The first class of oriented cycles to consider is probably the case of cycles with two blocks. Addario-Berry et al. [3] posed the following problem.

Problem 33 (Addario-Berry et al. [3])**.** Let k and ℓ be two positive integers. Does S-Forb($C(k, \ell) \cap \mathcal{S}$) have bounded chromatic number?

In [17] we proved the following theorem:

Theorem 34. S-Forb($C(k, \ell) \cap \mathcal{S}$) *has chromatic number $O((k + l)^4)$.*

The proof of this theorem, as well as other related results will be presented in Chapter 2. The function has been recently improved to $O((k + l)^2)$ by Kim et al. in [42].

A *p-spindle* is the union of k internally disjoint (x, y)-dipaths for some vertices x and y. Vertex x is said to be the *tail* of the spindle and y its *head.* A $(k_1 + k_2)$*-bispindle* is the union of k_1 (x, y)-dipaths and k_2 (y, x)-dipaths, all these dipaths being internally disjoint. One possible interpretation of Theorems 34 and 31 is to view a cycle on two blocks as a 2-spindle (two internally disjoint dipaths between the same pair of vertices) and a directed cycle as a (1+1)-spindle. A natural question is to try to extend Theorems 31 and 34 to larger spindles. First, we give a construction of digraphs with arbitrarily large chromatic number that contains no 3-spindle and no $(2 + 2)$-bispindle. Therefore, the most we can expect in all strongly connected digraphs with large chromatic number are $(2 + 1)$-bispindles.

Theorem 35. *For every integer k, there exists a strongly connected digraph D with $\chi(D) > k$ that contains no 3-spindle and no $(2 + 2)$-bispindle.*

Proof. Let $D_{k,4}$ be an acyclic digraph with chromatic number greater than k in which every cycle has at least four blocks. The existence of such a digraph is given by Theorem 29. Let $S = \{s_1, \ldots, s_l\}$ be the set of vertices of $D_{k,4}$ with outdegree 0 and $T = \{t_1, \ldots, t_m\}$ the set of vertices with indegree 0.

Consider the digraph D obtained from $D_{k,4}$ as follows. Add a dipath $P = (x_1, x_2, \dots, x_l, z, y_1, y_2, \dots, y_m)$ and the arcs (s_i, x_i) for all $i \in [l]$ and (y_j, t_j) for all $j \in [m]$. It is easy to see that D is strong. Moreover, in D, every directed cycle uses the arc (x_l, z). Therefore D does not contain a $(2+2)$-bispindle, because it contains two arc-disjoint directed cycles.

Suppose now that D has a 3-spindle with tail u and head v, and let Q_1, Q_2, Q_3 be its three (u, v)-dipaths. Observe that u and v are not vertices of P, because all vertices of this dipath have either indegree 2 or outdegree 2. In D, each oriented cycle with two blocks between vertices outside P must use the arc (x_l, z). The union of Q_1 and Q_2 form a cycle on two blocks, which means that one of the two paths, say Q_1, contains (x_l, z). But Q_2 and Q_3 also form a cycle on two blocks, but they cannot contain (x_l, z), a contradiction. □

Let $B(k_1, k_2; k_3)$ denote the digraph formed by three internally disjoint paths between two vertices x, y, two (x, y)-directed paths, one of size k_1, the other of size k_2, and one (y, x)-directed path of size k_3. Note that $B(k_1, k_2; k_3)$ is a (2+1)-bispindle where the length of the dipaths are specified. In [16], we conjectured the following:

Conjecture 36. There is a function $g : \mathbb{N}^3 \to \mathbb{N}$ such that every strong digraph with chromatic number at least $g(k_1, k_2, k_3)$ contains a subdivision of $B(k_1, k_2; k_3)$.

As an evidence, we proved the following result:

Theorem 37. *For every $k \geq 1$, there is a constant γ_k such that if D is a strong digraph with $\chi(D) > \gamma_k$, then D contains a subdivision of $B(k, 1; k)$.*

The proof is quite technical. However the main idea is to use the fact that strong digraphs with a Hamiltonian cycle and without a subdivision of $B(k, 1; k)$ easily have bounded chromatic number. To colour a digraph without a subdivision of $B(k, 1; k)$, we will then contract the long directed cycles one after the other until we reach a digraph without any long cycle. By Theorem 31, the resulting digraph has bounded chromatic number, and by carefully analysing how the contracted cycles can interact with one another, we deduce a colouring of the original digraph. The full proof, as well as the simpler case of $B(k, 1; 1)$ is presented in Chapter 2.

The following conjecture, probably very hard and way stronger than the results we are looking at here, would be an interesting tool for Conjecture 36.

Conjecture 38. For every $k \geq 1$, there exists an integer $f(k)$ such that every strong digraph with chromatic number greater than $f(k)$ contains a subdigraph H with chromatic number at least k and such that H contains a Hamiltonian cycle.

Recently, another notion of colouring for digraphs has received a lot of attention, mainly for its link to the celebrated Erdős-Hajnal conjecture (see [7] for more details), the directed colouring, or simply dicolouring. A *k-dicolouring* is a k-partition $(V_1, \dots, V_k)$ of $V(D)$ such that $D[V_i]$ is acyclic for every $i \in [k]$, and the *dichromatic number* of D is the minimum k such that D admits a k-dicolouring.

To analyse digraphs with large dichromatic number, one can consider only strong digraphs, as shown by the next lemma:

Lemma 39. *The dichromatic number of a digraph is the maximum of the dichromatic numbers of its strong components.*

Remember that Theorem 8 implies the following:

Theorem 28. *For every graph H, there exists a constant $C(H)$ such that any graph with chromatic number greater than $C(H)$ contains a subdivison of H.*

In [1] we proved the following theorem, which can be seen as a generalisation of the previous theorem:

Theorem 40. *Let F be a digraph on n vertices and m arcs. Every digraph D with $\vec{\chi}(D) > 4^m(n-1)+1$ contains a subdivision of F.*

The proof is presented in Chapter 2. An interesting question would be to see how much the bound $4^m(n-1)+1$ can be improved.

1.6 The Erdős-Sands-Sauer-Woodrow conjecture

We interpret a quasi-order on a set S as a digraph, where the vertices are the elements of S and the arcs are the ordered pairs (x, y) such that $x \leq y$. The *transitive closure* of a digraph $D = (V, A)$ is the digraph defined on V, with arc set the set of ordered pairs (x, y) such that there exists a dipath from x to y in D. The transitive closure of a digraph is a quasi-order. In a multidigraph D, a set S is *dominating* if for every $u \in D - S$, there exists $s \in S$ such that su is an arc. Let $\gamma(D)$ be the size of the smallest dominating set in D. In 1986, Sands, Sauer and Woodrow proved the following result:

Theorem 41 (Sands et al. [63])**.** *Let D be a digraph whose arcs are coloured with two colours. Then there exists a stable set S such that for every x not in S, there is a monochromatic dipath from x to a vertex of S.*

By considering the transitive closure of each colour class, the result can be stated as follows:

Theorem 42 (Sands et al. [63])**.** *Every multidigraph D whose arc set is the union of the arc sets of two quasi-orders contains a stable set dominating D.*

This statement can be seen as a generalisation of the Stable Marriage theorem. This theorem states that, given a complete bipartite graph B between n men and n women, where each person has ranked all members of the opposite gender in order of preference, there exists a perfect matching (set of marriages) such that there are no two people of opposite gender who would both rather marry each other than their current partners. To prove that there always exists such a matching, consider the following 2-arc-coloured (blue and red) digraph D: each vertex of D corresponds to an edge of B, the blue arcs correspond to the preferences of the men: if u prefers v_1 over v_2 then there is a blue arc from uv_1 to

uv_2 in D. Likewise the red arcs represent the preferences of the women. Applying Theorem 42 to D gives the desired matching, as a stable set in D is precisely a matching in B and two people of opposite gender who would both rather have each other than their current partners corresponds to a non dominated vertex in D.

In the same paper, Sands et al. conjectured the following generalisation:

Conjecture 43 (Sands et al. [63])**.** For every $k \geq 1$ there exists an integer $f(k)$ such that, for any multidigraph D whose arc set is the union of the arc sets of k quasi-orders, there exists $f(k)$ stable sets such that the union is dominating.

A weaker conjecture concerning tournaments was also asked in the same paper (and attributed to Erdős).

Conjecture 44 (Erdős and Sands et al. [63])**.** For every $k \geq 1$ there exists an integer $f(k)$ such that for any complete multidigraph D whose arc set is the union of the arc sets of k quasi-orders, $\gamma(D) \leq f(k)$.

This conjecture turned out to be very difficult, as despite attracting a lot of attention, the case $k = 3$ remained unresolved. In fact, before our proof of Conjecture 44 that will be presented in this thesis, even the following weaker conjecture due to Gyárfás was still open for $k = 3$.

Conjecture 45 (Gyárfás)**.** For every $k \geq 1$ there exists an integer $f(k)$ such that for any tournament T whose arc set is the union of the arc sets of k quasi orders, $\gamma(T) \leq f(k)$.

We will call the tournaments whose arc set is the union of the arc sets of k quasi orders *k-transitive* tournaments.

A nice application of Conjecture 45 was proved in [56]. Given two points $\mathbf{p} = (p_1, p_2, \ldots, p_d)$ and $\mathbf{q} = (q_1, q_2, \ldots, q_d)$ in $\mathbb{R}^d$, define $\text{box}(\mathbf{p}, \mathbf{q})$ as the smallest box in $\mathbb{R}^d$ containing $\mathbf{p}$ and $\mathbf{q}$. In other words, $\text{box}(\mathbf{p}, \mathbf{q}) = \{x \in \mathbb{R}^d \mid \forall i \min(p_i, q_i) \leq x_i \leq \max(p_i, q_i)\}$. In 1987, Bárány and Lehel [11] proved that every finite subset X of $\mathbb{R}^d$ can be covered by $h(d)$ X-boxes (boxes between points of X). To derive this statement from Conjecture 45, we consider a coloured clique on X. Take $\mathbf{p} = (p_1, \ldots, p_d)$ and $\mathbf{q} = (q_1, \ldots, q_d)$ and assume $p_1 < q_1$ (we can assume that all the coordinates are different). The colour of the edge $\mathbf{pq}$ is given according to the 2^{d-1} possible relations of the other $d - 1$ coordinates. This gives a clique with 2^{d-1} colours. Note that if $\mathbf{pq}$ is a red edge, then $\text{box}(\mathbf{p}, \mathbf{q})$ contains all the vertices $\mathbf{r}$ such that $(\mathbf{p}, \mathbf{r}, \mathbf{q})$ is a red dipath. Now for each colour class, there are two possible orientations giving a transitive order: if $p_1 < q_1$ or if $q_1 < p_1$. Consider all the tournaments obtained by taking any of these orientations. This gives $2^{2^{d-1}}$ k-transitive tournaments. If Conjecture 45 is true, then each one has a dominating set of size depending only on d. Consider the union S of the dominating sets of all these tournaments, and take all the boxes between pairs of points of S. We claim that this is a covering of X. Suppose it is not, and let $\mathbf{x}$ be a point not covered by the boxes of S. This means that for any colour class, $\mathbf{x}$ is not in the middle of a path on three vertices between two vertices of S, so there exists an orientation of this colour class such that $\mathbf{x}$ is not dominated by S in this colour. By doing this for all the classes, we obtain a tournament where S is not dominating $\mathbf{x}$, a contradiction.

A useful result for this problem is the following result due to Fisher [26]:

Theorem 46 (Fisher [26])**.** *For every digraph D, there exists a probability distribution on the vertices w such that for every vertex x, $w(N^-(x)) \geq w(N^+(x))$.*

Fisher uses this result in [26] to prove the tournament version of Seymour's second neighbourhood conjecture, which states that any oriented graph has a vertex whose outdegree is at most its second outdegree. The proof of this theorem uses Farkas's Lemma.

Lemma 47 (Farkas's Lemma)**.** *Let $\mathbf{M} \in \mathbb{R}^{n \times m}$ and $\mathbf{b} \in \mathbb{R}^m$. Then exactly one of the following statements is true:*

1. *There exists $\mathbf{x} \in \mathbb{R}^n$ such that $\mathbf{Mx} = \mathbf{b}$ and $\mathbf{x} \geq 0$.*

2. *There exists $\mathbf{y} \in \mathbb{R}^m$ such that $\mathbf{M}^\intercal \mathbf{y} \geq 0$ and $\mathbf{b}^\intercal \mathbf{y} < 0$.*

Let D be a digraph on vertices $\{v_1, \dots, v_n\}$. The *adjacency matrix* $\mathbf{M(D)} = (m_{ij})$ of D is the $n \times n$ matrix where $m_{ij} = 1$ if $(v_i, v_j) \in A(D)$, $m_{ij} = -1$ if $(v_j, v_i) \in A(D)$ and else $m_{ij} = 0$. Note that a probability distribution w satisfying Theorem 46 can be seen as a vector $\mathbf{w} \in \mathbb{R}^n$ such that $\mathbf{w} \geq \mathbf{0}$, $\mathbf{w}\mathbf{1}^\intercal = 1$ and $\mathbf{M(D)w} \leq \mathbf{0}$.

Proof of Theorem 46. Suppose that w doesn't exist. This means that the following system has no solutions ($\mathbf{I}$ is the identity matrix):

$$C = \begin{bmatrix} \mathbf{M(D)} & \mathbf{I} \\ \mathbf{1}^\intercal & \mathbf{0}^\intercal \end{bmatrix} \begin{pmatrix} \mathbf{w} \\ \mathbf{z} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} \text{ with } \begin{pmatrix} \mathbf{w} \\ \mathbf{z} \end{pmatrix} \geq \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}$$

Farkas's Lemma implies that the following system has a solution:

$$C = \begin{bmatrix} \mathbf{M(D)}^\intercal & \mathbf{1} \\ \mathbf{I} & \mathbf{0}^\intercal \end{bmatrix} \begin{pmatrix} \mathbf{u} \\ v \end{pmatrix} \geq \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix} \text{ with } \begin{pmatrix} \mathbf{0}^\intercal & 1 \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ v \end{pmatrix} < 0$$

As $\mathbf{M(D)}^\intercal = -\mathbf{M(D)}$ we get that $\mathbf{M(D)u} \leq v\mathbf{1}$ with $v < 0$, so $\mathbf{M(D)u} < 0$. Moreover, as $\mathbf{Iu} \geq \mathbf{0}$, then $\mathbf{u} \geq 0$ and by normalising $\mathbf{u}$, we get the required probability distribution on the vertices. $\square$

If we apply this result to a tournament, we get that $w(N^-[x]) \geq 1/2$. To get the intuition, it can be useful to forget this weight function and consider that the indegree of every vertex is at least half the vertices.

A powerful tool to bound the domination on certain class of tournaments is the VC-dimension of the inneighbourhood hypergraph. Let $\mathcal{H} = (V, E)$ be a hypergraph. We say that a subset S' of V is *shattered* if $\mathcal{H} \cap S'$ contains all subsets of S'. The *VC-dimension* of a hypergraph is the size of the biggest shattered set of $\mathcal{H}$. A fundamental result concerning VC-dimension is the following lemma, due to Sauer [64] (see also Lemma 10.2.5 of [51]).

Lemma 48 (Sauer [64])**.** *Let $\mathcal{H}$ be a hypergraph of VC-dimension at most k. Then for every subset A of the ground set, $|E(\mathcal{H} \cap A)| < (\frac{e|A|}{k})^k$.*

The *VC-dimension* of a tournament is the VC-dimension of its inneighbourhood hypergraph. It was first observed by Alon et al.[5] that any tournament T with VC-dimension bounded by t has $\gamma(T) \leq O(t\log(t))$, because the fractional domination of a tournament is bounded by 2 and thus we can apply the ϵ-net Theorem (see chapter 10.2 of [51] for a nice presentation). Rephrasing the proof of the ϵ-net Theorem in the case of tournaments with large indegree gives something along these lines:

For simplicity, we assume here that the indegree of every vertex is at least $|T|/2$. The probability distribution given by Theorem 46 will be the tool to deal with the general case in a similar fashion. Let c be some constant only depending in the VC-dimension t. We want to prove that there exists a dominating set of size c by contradiction. Take two random subsets of size c, S_1 and S_2, uniformly in T and let A be the following event: "There exists a vertex v that dominates S_1 and $|N^-(v) \cap S_2| \geq 1/4$." We will compute the probability of A in two different ways and reach a contradiction. First, since we assumed that no set of size c is dominating, there exists an element v dominating S_1, and the probability that $|N^-(v) \cap S_2| \geq 1/4$ is a lower bound on the probability of A. Since c is a constant, we can view roughly the random variable S_2 as picking c vertices uniformly and independently among vertices of T. Since $|N^-(v)| \geq 1/2$ the random variable $|N^-(v) \cap S_2|$ follows a binomial distribution with parameter c and $p \geq 1/2$. Using a classical bound on the binomial distribution (see Lemma 10.2.6 of [51]), we have that $Pr[A] \geq 1/2$. For the second way of computing $Pr[A]$, instead of taking S_1 and S_2 uniformly at random, we select a set A of size $2c$ uniformly at random and then partition it uniformly at random into (S_1, S_2). The resulting distribution is the same as above. To bound $Pr[A]$, we fix $S \subseteq A$ and compute the probability that a vertex v dominates S_1 and $|N^-(v) \cap S_2| \geq 1/4$ conditioned by the fact that $N^-(v) \cap A = S$. If $N^-(v) \cap A < c/4$, then the probability is 0; if $N^-(v) \cap A \geq c/4$, then the probability is bounded by the probability that a random sample of size c in $2c$ elements avoids at least $c/4$ positions. This is at most

$$\frac{\binom{2c-c/4}{c}}{\binom{2c}{c}} \leq (1-1/8)^c \leq \exp(-c/8).$$

Now remark that for a vertex v, the probability that v dominates S_1 and $|N^-(v) \cap S_2| \geq 1/4$ depends only on $A \cap N^-(v)$. However, we proved that, for each possible intersection, the probability is bounded by $\exp(-c/8)$. Since the VC-dimension of the inneighbourhood of T is bounded from above by t, by Lemma 48 there are at most $(ec/k)^k$ possible intersections and we get,

$$1/2 \leq Pr[A] \leq \exp(-c/8)(ec/k)^k$$

which for c big enough compared to k is a contradiction.

When we tried to attack Conjecture 44, we first looked at the case $k = 3$ of Conjecture 45. We proved it using a sampling argument inspired by the proof of the ϵ-net Theorem presented above, despite the fact that 3-transitive tournaments do not have bounded VC-dimension. Trying to extend the proof to $k > 3$, we found in [10] a very simple proof of Conjecture 44 which does not use the VC-dimension at all.

Theorem 49. *For every $k \geq 1$, if T is a complete multidigraph whose arc set is the union of the arc sets of k quasi-orders, then $\gamma(T) = O(ln(2k) \cdot k^{k+2})$.*

The two proofs are presented in Chapter 4.

1.7 Entropy Compression

The probabilistic method is a surprisingly powerful method in combinatorics and one of its central tool is the Lovász Local lemma (LLL in short).

Lemma 50 (LLL)**.** *Let $A_1, A_2, \ldots, A_n$ be random events of an arbitrary probability space. Suppose that each event A_i is mutually independent of all the other event A_j but at most d, and that $Pr[A_i] \leq p$ for all i. If $ep(d+1) \leq 1$, then $Pr[\cap_i \bar{A}_i] > 0$.*

One very common setting is to view the probability space as a set of independent variables $X_1, X_2, \ldots, X_l$ and each event A_i depends on a finite set $D(A_i)$ of these variables. The LLL tells us that under the right conditions on probability and dependency, there exists an assignment of the variables such that none of the A_i happens. Note that the A_i are usually called the "bad events" because the goal is to avoid them. A very natural question that has attracted a lot of attention over the years was to find a way to compute these assignments. In 2009, Moser [55] gave the answer to this problem by showing that the most natural algorithm works:

Algorithm 1: Moser Fix it Algorithm

> Pick $X_1, \ldots, X_n$ uniformly at random;
> **while** *There exists i such that A_i is true* **do**
>   Repick all the variables in $D(A_i)$ uniformly at random;
> **return** the X_i;

What is especially interesting is how he managed to prove that the algorithm terminates. Suppose for example that the X_i are random bits and assume that the algorithm never finishes after k steps, where a step is one random choice. In this case, a way to see this algorithm is to see it as a deterministic algorithm which takes the random choices as a k-bits vector input and produces some assignments of the X_i, which correspond to the state of the X_i at the end of the algorithm run on the vector. Moser's idea was that he could modify the algorithm to produce some sequence of integers such that, given this sequence and the value of the X_i at the end of the algorithm, one can deduce the value of the k-bit vector taken as input. The sequence produced is called a *log*. Usually the logs produced by the algorithm have very specific properties and this allow us to show that the set of possible logs after k steps is strictly smaller than 2^k. Since 2^k is the number of k-bit vectors, this gives a contradiction. The intuitive way of seeing the log is to see it being built step by step during the algorithm: each time the algorithm erases a set of random choices, it encodes it in the log using less information. For example, if A_i is a bad event, and the algorithm erased the values of $D(A_i)$, then the log need to encode two things: how to know that A_i was the bad event and then the value of the variable in $D(A_i)$. Usually, describing which A_i is a bad event is the hard part, and this is where we use the local aspect of the problem. Indeed, since $Pr[A_i]$ is small, then describing the assignment of $D(A_i)$ such that A_i happens doesn't cost much compared to the set of all possible choices. Writing the proof usually requires more efforts, as one needs to define the log, show that we can deduce the random choices from it, and compute the number of possible logs (sometimes using analytic combinatorics). However the point of view of "information deleted versus information needed to recover it" is the one people use to get the intuition.

This method of showing the existence of a combinatorial object by proving that a random algorithm terminates by encoding its random choices into a set of smaller size is called "the entropy compression method". Since then, it has been used a lot, sometimes to improve on bounds obtained by the usual LLL but not always. During this thesis, we use this setting for two different results, which we shortly describe here and that one can find in Chapter 5.

1.7.1 Orientations of Hypergraphs

In [15], Caro, West, and Yuster presented a generalization to hypergraphs of the notion of *orientation* defined for graphs. Their acknowledged purpose was to study how hypergraphs can be oriented in such a way that minimum and maximum degree are close to each other, knowing that the following theorem is true for graphs.

Theorem 51 (Folklore). *If G is a graph, then there exists an orientation D of G such that $|d^+(u) - d^-(u)| \leq 1$ for every vertex u.*

Identifying an *orientation* of a hyperedge with a total ordering of its elements, they defined a notion of degree on oriented r-uniform hypergraphs.

Definition 52. Let $\mathcal{H}$ be an r-uniform hypergraph, and let every $S \in E(\mathcal{H})$ define a total order on its elements as a bijection $\sigma_S : S \mapsto [r]$. The *degree* $d_P(U)$ of a set of vertices $U \subseteq V(\mathcal{H})$ with respect to a set of positions $P \subseteq [r]$ (where $|P| = |U|$) is equal to:

$$d_P(U) = |\{S \in \mathcal{H} : U \subseteq S \text{ and } \sigma_S(U) = P\}|$$

From there they defined *equitable* orientations:

Definition 53. The orientation of a r-uniform hypergraph $\mathcal{H}$ is said to be *p-equitable* if $|d_P(U) - d_{P'}(U)| \leq 1$ for any choice of $U \subseteq V(\mathcal{H})$ and $P, P' \subseteq [r]$ of cardinality p. It is said to be *nearly p-equitable* if the looser requirement $|d_P(U) - d_{P'}(U)| \leq 2$ holds.

Caro, West, and Yuster [15] proved that all hypergraphs admit a 1-equitable as well as a $(r-1)$-equitable orientation, and also proved that some hypergraphs do not admit any p-equitable orientation for all other values of p. Additionally, they parameterized the notion of *maximum degree* in order to focus on hypergraphs which are *sparse* with respect to the problem at hand:

$$\Delta_p(\mathcal{H}) = \max_{\substack{U \subseteq V(\mathcal{H}) \\ |U|=p}} |\{S \in \mathcal{H} : U \subseteq S\}|$$

Thus, they proved that for any fixed values of p and k, and for every sufficiently large integer r, every r-uniform hypergraph $\mathcal{H}$ with $\Delta_p(\mathcal{H}) \leq k$ admits a nearly p-equitable orientation. They conjectured that this setting actually ensured the existence of a p-equitable orientation, which we proved in [18] with Nathann Cohen.

Theorem 54. *Let p, k be fixed integers. There exists r_0 such that for every $r \geq r_0$, every r-uniform hypergraph with $\Delta_p(\mathcal{H}) \leq k$ admits a p-equitable orientation.*

Note that, as r is big compared to $\Delta_p(\mathcal{H})$, a p-equitable orientation means that $d_P(U)$ is equal to 0 or 1 for every choice of set of positions P and set of vertices U.

In order to prove the existence of nearly p-equitable orientations, Caro, West, and Yuster [15] used the Lovász Local Lemma. We use the entropy compression technique to improve their result. The algorithm we propose will orient the hyperedges of H one by one at random such that, at any time during the algorithm, the number of times a set U of p vertices is sent to a set of positions P by the hyperedges which have been oriented is 0 or 1. This means that if at some point in the algorithm all the hyperedges are oriented, then H admits a p-equitable orientation. Suppose now we have a partial orientation of H and we try to orient an hyperedge X. For any subset U of p vertices of X, the orientation we give on X cannot send U to a set of position P which it has already been sent to before. Call *valid* such a permutation. So for any subset of p vertices of X, there is a list of forbidden positions, and a valid orientation is one which avoids all of those. This can be seen as a generalization of derangements. Remember that a permutation σ of $[n]$ is a *derangement* if $\sigma(i) \neq i$ for all i.

The main idea of the algorithm is to ensure that, at any time, the number of valid orientations available for every hyperedge is large. To do so we need the following lemma, which can be seen as a generalization of the classical result saying that the number of derangement of size n tends to $n!/e$. The proof will make use of another variant of the LLL, namely the Lopsided Lovász Local lemma.

Lemma 55. *Let $p, k \in \mathbb{N}$ and $\alpha < 1$ be fixed. Let X be a set of cardinality r and let $\mathcal{L}_S$ be, for every $S \in \binom{X}{p}$, a collection of p-subsets of X with $|\mathcal{L}_S| \leq k$. Then, if no p-subset occurs in more than r^α of the $\mathcal{L}_S$, a random permutation σ of X satisfies $\sigma(S) \notin \mathcal{L}_S$ for every S with probability $\geq (1 - 2k/\binom{r}{p})^{\binom{r}{p}} = e^{-2k} + o(1)$ when r grows large.*

The collection $\mathcal{L}_S$ corresponds to the list of forbidden positions for S. What the previous lemma says is that if a position is not forbidden too many times, i.e it appears in too many different $\mathcal{L}_S$, then there will be a constant fraction of the permutations available as valid orientations.

So the algorithm will be the following: Chose the permutations one by one among the valid ones, and if at some point there exists an edge X of H and a set P of positions such that P is forbidden too many times for X, then erase the permutations that forbid P for X. The algorithm stops once every permutation is chosen.

The idea behind this algorithm is to make sure that at each step, the number of good permutations is high enough to make the entropy compression argument work. In this sense it is a bit different from a normal LLL argument and reminds us more of a semi-random method, where we pick part of the solution while ensuring that the rest has nice properties.

1.7.2 AVD-colouring

An *edge colouring* of a graph G is defined as a mapping $c : E(G) \to \mathbb{N}$ associating integers (colours) to the edges of G. An edge colouring is said to be *proper* if no two adjacent edges receive the same colour. Vizing Theorem says that every graph has a proper edge colouring with $\Delta(G) + 1$ colours. In 2002 Zhang, Liu, and Wang [74], introduced a new variant of edge colouring. Given a proper edge colouring c of G and a vertex $u \in V(G)$, we let $S_c(u)$ denote the set of colours appearing on edges incident to u. A proper edge colouring c is *adjacent vertex distinguishing* (*AVD* for short) if $S_c(u) \neq S_c(v)$ for every edge $uv \in E(G)$. They conjectured:

Conjecture 56 (Zhang et al. [74])**.** Every connected graph G with maximum degree Δ which is distinct from K_2 and C_5 has an AVD-colouring with at most $\Delta + 2$ colours.

This conjecture captured the attention of several researchers over the years. It is known to be true e.g. if G is bipartite [6] or if $\Delta = 3$ [6]. It also holds asymptotically almost surely for random 4-regular graphs [31]. Hatami [35] in 2005 proved that every graph with maximum degree $\Delta > 10^{20}$ and no isolated edge has an AVD-colouring with at most $\Delta + 300$ colours. Since then no further progress has been made on Conjecture 56.

With Gwenaël Joret in [37], we improved this result to $\Delta + 19$ for large value of Δ.

Theorem 57. *For Δ large enough, every digraph with maximum degree Δ and no isolated edge has an AVD-colouring with at most $\Delta + 19$ colours.*

Algorithm

We will now describe the proof in the case of d-regular graphs. The presentation uses a simpler argument than the one presented in Chapter 5. It leads to a worse constant but contains the general ideas. The method of the proof is the following one, already present in the proof of Hatami. Start with a proper $(d + 1)$-edge-colouring c given by Vizing's Theorem. Let c' be a partial edge colouring, then $S_{c'}(u)$ is the set of colours incident to u in this colouring. The goal is to uncolour the edges such that the partial colouring c' obtained has the following properties:

(i) The graph of uncoloured edges has maximum degree bounded by $q+3$ for some constant q; and

(ii) c' is AVD, meaning that if uv is an edge of G, then $S_{c'}(u) \neq S_{c'}(v)$.

Now pick c_s a proper edge colouring of the uncoloured graph using a set of new colours (different from those of c): the colouring c_f obtained by taking the union of c_s and c' is an AVD-colouring with $d + q + 5$ colours. Indeed, it is easy to see that this produces a proper edge colouring of the graph. Moreover if u and v are two adjacent vertices, then by definition of c', there exists a colour i adjacent to v and not to u in c'. But since c_s uses new colours, the colour i is still incident to v and not to u in c_f.

To find c', we describe an algorithm that chooses, for every v, a set $U(v)$ of three edges adjacent to v such that $\bigcup_{u \in V(G)} U(v)$ will be the set of uncoloured edges. At the beginning of the algorithm, all the sets $U(v)$ are set to empty. Now at each step, the algorithm will pick deterministically a vertex v such that $U(v) = \emptyset$ and pick $U(v)$ among the $\binom{d}{3}$ possible choices. The goal of the algorithm is to make sure that, if it ends, then property (i) and (ii) are satisfied. We will define two types of bads events, for which the algorithm will reset some of the $U(v)$ to empty when they occur. The first type is defined to ensure that property (i) is satisfied. Note that, the degree of a vertex v in the uncoloured graph is equal to three plus the number of times v has been chosen in $U(u)$ for $u \in N(v)$. So the first bad event will be:

Bad event of type 1. *After choosing $U(x)$, $v \in U(x)$ belongs to q sets $U(u)$.*

In this case, the algorithm erases the choices of these $U(u)$.

To deal with property (ii), we define the following notion: We say that an edge uv is *finished* if $U(w) \neq \emptyset$ for each vertex w in $N(u) \cup N(v)$ (note that this set includes u and v). When an edge is finished, it means that all the random choices relevant to the set of colours adjacent to u and v at the end of the algorithm have been made. In particular if these sets are the same, then the algorithm triggers a bad event:

Bad event of type 2. *After choosing $U(x)$, there exists $v \in N(x)$ and $u \in N(v)$ such that, uv is finished and $S_{c'}(u) = S_{c'}(v)$.*

In this case, the algorithm erases the choice for $U(v)$.

Encoding

It is clear that, if the algorithm terminates, it produces the desired partial colouring c'. In order to complete the proof, we only need to show that, there exists a set of random choices such that the algorithm terminates. Using the entropy compression template, we need to show that, any time we erase a random choice, we can encode it with less information. To do this encoding, we suppose that we know the state of the $U(v)$ for every vertex in the graph after treating the bad event. If we are able to recover the value of the variable that we just erased, it means that, by looking at the log and the state of the sets $U(v)$ when the algorithm stops, we can recover the values of $U(v)$ at any moment of the algorithm.

When a Bad event of type 1 is triggered, the algorithm erases q choices of $U(u)$, so $\binom{d}{3}^q$ possibilities. To encode this we need to give two information: the vertex v and neighbours of v for which we deleted the choices of U. Because the bad event occurs when a vertex x of $N(v)$ chooses v in $U(x)$, describing v is just describing for which neighbour of x the bad event happens and thus can be encoded using an integer in $[d]$. To encode the $q-1$ neighbours of v outside of u we can use an integer in $\binom{d}{q-1}$. And finally for these vertices we need to encode the two other choices which makes U : $\binom{d}{2}$. Overall we encode the $\binom{d}{3}^q$ possibilities using an integer in $d \times \binom{d}{q-1} \times \binom{d}{2}^q$. And for q big enough (10 for example works here) $\binom{d}{3}^q < d \times \binom{d}{q-1} \times \binom{d}{2}^q$, so we managed to compress some information.

When a Bad event of type 2 happens, the algorithm erases choices for $U(v)$, so $\binom{d}{3}$ possibilities. Note that to recover the information, we just need to know the vertices u and v. Indeed since u and v have the same set of adjacent colours, $U(v)$ is precisely the set of colours adjacent to v and not to v. In this case encoding the $\binom{d}{3}$ possible choices only requires an integer in $[d^2]$ which is needed to say which neighbour of x is v and which neighbour of v is u.

Overall this proves that, if t is big enough, then we can encode all the of t random choices such that the algorithm doesn't terminate in a set of logs of size $o\left(\binom{d}{3}^t\right)$. Thus there exists a set of t random choices such that the algorithm terminates, which gives the desired partial colouring.

The general proof is presented in Chapter 5.

Chapter 2

Subdivisions in digraphs with large chromatic number

In this chapter, we present some results on subdivisions in digraphs with large chromatic and dichromatic number.

After some definitions we present some proofs from [17] obtained with Nathann Cohen, Frédéric Havet and Nicolas Nisse about oriented cycles in digraphs with large chromatic number. Then we present the proof obtained in [16] with Nathann Cohen, Frédéric Havet and Raul Lopes about subdivisions of spindles (so directed paths between the same pair of vertices). Finally we present a proof of a theorem obtained in [1] with Pierre Aboulker, Nathann Cohen, Frédéric Havet, Phablo Moura and Stéphan Thomassé saying that any digraph H is contained in a digraph with dichromatic number greater than a function of H.

2.1 Definitions

Let F be a digraph. An *F-subdivision* is a subdivision of F. A digraph D is said to be *F-subdivision-free*, if it contains no F-subdivision.

The *union* of two digraphs D_1 and D_2 is the digraph $D_1 \cup D_2$ defined by $V(D_1 \cup D_2) = V(D_1) \cup V(D_2)$ and $A(D_1 \cup D_2) = A(D_1) \cup A(D_2)$. If $\mathcal{D}$ is a set of digraphs, we denote by $\bigcup \mathcal{D}$ the union of the digraphs in $\mathcal{D}$, i.e. $V(\bigcup \mathcal{D}) = \bigcup_{D \in \mathcal{D}} V(D)$ and $A(\bigcup \mathcal{D}) = \bigcup_{D \in \mathcal{D}} A(D)$.

If D is a dipath or a directed cycle, then we denote by $D[a,b]$ the subdipath of D with initial vertex a and terminal vertex b. We denote by $D[a,b[$ the dipath $D[a,b] - b$, by $D]a,b]$ the dipath $D[a,b] - a$, and by $D]a,b[$ the dipath $D[a,b] - \{a,b\}$. If P and Q are two dipaths such that $V(P) \cap V(Q) = \{t(P)\} = \{s(Q)\}$, the *concatenation* of P and Q, denoted by $P \odot Q$, is the dipath $P \cup Q$.

In a digraph D, the *distance* from a vertex x to another y, denoted by $\mathrm{dist}_D(x,y)$ or simply $\mathrm{dist}(x,y)$ when D is clear from the context, is the minimum length of an (x,y)-dipath or $+\infty$ if no such dipath exists. For a set $X \subseteq V(D)$ and vertex $y \in V(D)$, we define $\mathrm{dist}(X,y) = \min\{\mathrm{dist}(x,y) \mid x \in X\}$ and $\mathrm{dist}(y,X) = \min\{\mathrm{dist}(y,x) \mid x \in X\}$, and for two sets $X, Y \subseteq V(D)$, $\mathrm{dist}(X,Y) = \min\{\mathrm{dist}(x,y) \mid x \in X, y \in Y\}$.

An important notion in this chapter is the notion of levelling. An *out-generator* in D is a vertex u such that, for every $x \in V(D)$, there exists an (u, x)-dipath in D. Analogously, an *in-generator* in D is a vertex u such that, for every $x \in V(D)$, there exists an (x, u)-dipath in D. For simplicity, we call a vertex *generator* if it is an in- or out-generator. Observe that every vertex in a strong digraph is an in- and out-generator. Let D be a digraph. Let w, u be in- and out-generators of D, respectively. We remark that w and u are not necessarily different. For every nonnegative integer i, the *ith out-level from* u in D is the set $L_i^{u,+} = \{v \in V(D) \mid \text{dist}_D(u, v) = i\}$, and the *ith in-level from* w in D is the set $L_i^{w,-} = \{v \in V(D) \mid \text{dist}_D(v, w) = i\}$. Note that $\bigcup_i L_i^{u,+} = \bigcup_i L_i^{w,-} = V(D)$.

An *out-Breadth-First-Search Tree* or *out-BFS-tree* T^+ with *root* u, is a subdigraph of D spanning $V(D)$ such that T^+ is an oriented tree and, for every $v \in V(D)$, $dist_{T^+}(u, v) = \text{dist}_D(u, v)$. Similarly, an *in-Breadth-First-Search Tree* or *in-BFS-tree* T^- with *root* w, is a subdigraph of D spanning $V(D)$ such that T^- is an oriented tree and, for every $v \in V(D)$, $dist_{T^-}(v, w) = \text{dist}_D(v, w)$.

It is well-known that if D has an out-generator, then there exists an out-BFS-tree rooted at this vertex. Likewise, if D has an in-generator, then there exists an in-BFS-tree rooted at this generator. Let T denote an in- or out-BFS-tree rooted at u. For any vertex x of D, there is a single (u, x)-dipath in T if T is an out-BFS-tree, and a single (x, u)-dipath in T if T is an in-BFS-tree. The *ancestors* or *successors* of x in T are naturally defined.

If y is an ancestor of x, we denote by $T[y, x]$ the (y, x)-dipath in T. If y is a successor of x, we denote by $T[x, y]$ the (x, y)-dipath in T.

Lemma 58. *Let D be a strong digraph and let T be an in- or out-BFS-tree in D. There is a level L such that $\vec{\chi}(D[L]) \geq \vec{\chi}(D)/2$.*

Proof. First, let us suppose, without loss of generality, that T is an out-BFS-tree in D. The proof when T is an in-BFS-tree is analogous.

Let D_1 and D_2 be the subdigraphs of D induced by the vertices of odd and even levels, respectively. Since there is no arc from L_i to L_j for every $j \geq i + 2$, the strong components of D_1 and D_2 are contained in the levels. Hence, by Lemma 39, $\vec{\chi}(D_1) = \max\{\vec{\chi}(D[L_i]) \mid i \text{ is odd}\}$ and $\vec{\chi}(D_2) = \max\{\vec{\chi}(D[L_i]) \mid i \text{ is even}\}$. Moreover, note that $V(D_1) \cup V(D_2) = V(D)$ because D is strong. Therefore, $\vec{\chi}(D) \leq \vec{\chi}(D_1) + \vec{\chi}(D_2) \leq 2 \cdot \max\{\vec{\chi}(D[L_i]) \mid i \in \mathbb{N}\}$. □

The following lemma is well-known.

Lemma 59. *Let D_1 and D_2 be two digraphs. $\chi(D_1 \cup D_2) \leq \chi(D_1) \times \chi(D_2)$.*

Proof. Let $D = D_1 \cup D_2$. For $i \in \{1, 2\}$, let c_i be a proper colouring of D_i with $\{1, \ldots, \chi(D_i)\}$. Extend c_i to $(V(D), A(D_i))$ by assigning the colour 1 to all vertices in V_{3-i}. Now the function c defined by $c(v) = (c_1(v), c_2(v))$ for all $v \in V(D)$ is a proper colouring of D with colour set $\{1, \ldots, \chi(D_1)\} \times \{1, \ldots, \chi(D_2)\}$. □

2.2 Oriented cycles

In this section we first prove that S-Forb$(C(k,\ell)) \cap \mathcal{S}$ has bounded chromatic number for every k, ℓ. We use the notion of levelling, but only for out-generators. For simplicity, we note $L_i^u = L_i^{u,+}$. Let v be a vertex of D, we set $\mathrm{lvl}^u(v) = \mathrm{dist}_D(u,v)$, hence $v \in L^u_{\mathrm{lvl}(v)}$. In the following, the vertex u is always clear from the context. Therefore, for sake of clarity, we drop the superscript u.

We write *BFS-tree* for out-BFS-tree. For an ancestor y of x, we note $y \geq_T x$. For any two vertices v_1 and v_2, the *least common ancestor* of v_1 and v_2 is the common ancestor x of v_1 and v_2 for which $\mathrm{lvl}(x)$ is maximal. (This is well-defined since u is an ancestor of all vertices.)

The following proposition directly follows from the definitions

Proposition 60. *Let D be a digraph having an out-generator u. If x and y are two vertices of D with* $\mathrm{lvl}(y) > \mathrm{lvl}(x)$, *then every (x,y)-dipath has length at least* $\mathrm{lvl}(y) - \mathrm{lvl}(x)$.

2.2.1 Cycles with 2 blocks

Theorem 61. *Let k and ℓ be positive integers such that $k \geq \max\{\ell, 3\}$ and $\ell \geq 2$, and let D be a digraph in* S-Forb$(C(k,\ell)) \cap \mathcal{S}$. *Then, $\chi(D) \leq (k+\ell-2)(k+\ell-3)(2\ell+2)(k+\ell+1)$.*

Proof. Since D is strongly connected, it has an out-generator u. Let T be a BFS-tree with root u. We define the following sets of arcs.

$$\begin{aligned}
A_0 &= \{(x,y) \in A(D) \mid \mathrm{lvl}(x) = \mathrm{lvl}(y)\};\\
A_1 &= \{(x,y) \in A(D) \mid 0 < |\mathrm{lvl}(x) - \mathrm{lvl}(y)| < k+\ell-3\};\\
A' &= \{(x,y) \in A(D) \mid \mathrm{lvl}(x) - \mathrm{lvl}(y) \geq k+\ell-3\}.
\end{aligned}$$

Since $k+\ell-3 > 0$ and there is no arc (x,y) with $\mathrm{lvl}(y) > \mathrm{lvl}(x)+1$, (A_0, A_1, A') is a partition of $A(D)$. Observe moreover that $A(T) \subseteq A_1$. We further partition A' into two sets A_2 and A_3, where $A_2 = \{(x,y) \in A' \mid y \text{ is an ancestor of } x \text{ in } T\}$ and $A_3 = A' \setminus A_2$. Then (A_0, A_1, A_2, A_3) is a partition of $A(D)$. Let $D_j = (V(D), A_j)$ for all $j \in \{0,1,2,3\}$.

Claim 61.1. $\chi(D_0) \leq k+\ell-2$.

Subproof. Observe that D_0 is the disjoint union of the $D[L_i]$ where $L_i = \{v \mid \mathrm{dist}_D(u,v) = i\}$. Therefore it suffices to prove that $\chi(D[L_i]) \leq k+\ell-2$ for every non-negative integer i.

$L_0 = \{u\}$ so the result holds trivially for $i = 0$.

Assume now $i \geq 1$. Suppose for a contradiction $\chi(D[L_i]) \geq k+\ell-1$. Since $k \geq 3$, by Theorem 27, $D[L_i]$ contains a copy Q of $P^+(k-1,\ell-1)$. Let v_1 and v_2 be the initial and terminal vertices of Q, and let x be the least common ancestor of v_1 and v_2. By definition, for $j \in \{1,2\}$, there exists an (x,v_j)-dipath P_j in T. By definition of least common ancestor, $V(P_1) \cap V(P_2) = \{x\}$, $V(P_j) \cap L_i = \{v_j\}$, $j = 1,2$, and both P_1 and P_2 have length at least 1. Consequently, $P_1 \cup P_2 \cup Q$ is a subdivision of $C(k,\ell)$, a contradiction. ◊

Claim 61.2. $\chi(D_1) \leq k + \ell - 3$.

Subproof. Let ϕ_1 be the colouring of D_1 defined by $\phi_1(x) = \mathrm{lvl}(x) \pmod{k + \ell - 3}$. By definition of D_1, this is clearly a proper colouring of D_1. $\Diamond$

Claim 61.3. $\chi(D_2) \leq 2\ell + 2$.

Subproof. Suppose for a contradiction that $\chi(D_2) \geq 2\ell + 3$. By Theorem 27, D_2 contains a copy Q of $P^-(\ell + 1, \ell + 1)$, which is the union of two dipaths which are disjoint except in there initial vertex y, say $Q_1 = (y_0, y_1, y_2, \ldots, y_{\ell+1})$ and $Q_2 = (z_0, z_1, z_2, \ldots, z_{\ell+1})$ with $y_0 = z_0 = y$. Since Q is in D_2, all vertices of Q belong to $T[u, y]$. Without loss of generality, we can assume $z_1 \geq_T y_1$.

If $z_{\ell+1} \geq_T y_{\ell+1}$, then let j be the smallest integer such that $z_j \geq_T y_{\ell+1}$. Then the union of $T[y_1, y] \odot Q_2[y, z_j] \odot T[z_j, y_{\ell+1}]$ and $Q_1[y_1, y_{\ell+1}]$ is a subdivision of $C(k, \ell)$, because $T[y_1, y]$ has length at least $k - 2$ as $\mathrm{lvl}(y) \geq \mathrm{lvl}(y_1) + k + \ell - 3$. This is a contradiction.

Henceforth $y_{\ell+1} \geq_T z_{\ell+1}$. Observe that all the z_j, $1 \leq j \leq \ell + 1$ are in $T[y_{\ell+1}, y_1]$. Thus, by the Pigeonhole principle, there exists $i, j \geq 1$ such that $y_{i+1} \geq_T z_{j+1} \geq_T z_j \geq_T y_i \geq_T z_{j-1}$.

If $\mathrm{lvl}(z_{j-1}) \geq \mathrm{lvl}(y_i) + \ell - 1$, then $T[y_i, z_{j-1}] \odot (z_{j-1}, z_j)$ has length at least ℓ. Hence its union with $(y_i, y_{i+1}) \odot T[y_{i+1}, z_j]$, which has length greater than k, is a subdivision of $C(k, \ell)$, a contradiction.

Thus $\mathrm{lvl}(z_{j-1}) < \mathrm{lvl}(y_i) + \ell - 1$ (in particular, in this case, $j > 1$ and $i > 2$). Therefore, by definition of A', $\mathrm{lvl}(y_i) \geq \mathrm{lvl}(z_j) + k - 1$ and $\mathrm{lvl}(y_{i-1}) \geq \mathrm{lvl}(z_{j-1}) + k - 1$. Hence both $T[z_{j-1}, y_{i-1}]$ and $T[z_j, y_i]$ have length at least $k - 1$. So the union of $T[z_{j-1}, y_{i-1}] \odot (y_{i-1}, y_i)$ and $(z_{j-1}, z_j) \odot T[z_j, y_i]$ is a subdivision of $C(k, k)$ (and thus of $C(k, \ell)$), a contradiction. $\Diamond$

Claim 61.4. $\chi(D_3) \leq k + \ell + 1$.

Subproof. In this claim, it is important to note that $k + \ell - 3 \geq k - 1$ because $\ell \geq 2$. We use the fact that $\mathrm{lvl}(x) - \mathrm{lvl}(y) \geq k - 1$ if (x, y) is an arc in A_3.

Suppose for a contradiction that $\chi(D_3) \geq k + \ell + 1$. By Theorem 27, D_3 contains a copy Q of $P^-(k, \ell)$ which is the union of two dipaths which are disjoint except in there initial vertex y, say $Q_1 = (y_0, y_1, y_2, \ldots, y_k)$ and $Q_2 = (z_0, z_1, z_2, \ldots, z_\ell)$ with $y_0 = z_0 = y$.

Assume that a vertex of $Q_1 - y$ is an ancestor of y. Let i be the smallest index such that y_i is an ancestor of y. If it exists, by definition of A_3, $i \geq 2$. Let x be the common ancestor of y_i and y_{i-1} in T. By definition of A_3, y_i is not an ancestor of y_{i-1}, so x is different from y_i and y_{i-1}. Moreover by definition of A', $\mathrm{lvl}(y) - k \geq \mathrm{lvl}(y_{i-1}) - k \geq \mathrm{lvl}(y_i) - 1 \geq \mathrm{lvl}(x)$. Hence $T[x, y_{i-1}]$ and $T[x, y]$ have length at least k. Moreover these two dipaths are disjoint except in x. Therefore, the union of $T[x, y_{i-1}]$ and $T[x, y] \odot Q_1[y, y_{i-1}]$ is a subdivision of $C(k, k)$ (and thus of $C(k, \ell)$), a contradiction.

Similarly, we get a contradiction if a vertex of $Q_2 - y$ is an ancestor of y. Henceforth, no vertex of $V(Q_1) \cup V(Q_2) \setminus \{y\}$ is an ancestor of y.

Let x_1 be the least common ancestor of y and y_1. Note that $|T[x_1, y]| \geq k$ so $|T[x_1, y_1]| < k$, for otherwise G would contain a subdivision of $C(k, k)$. Therefore $\mathrm{lvl}(y_1) - \mathrm{lvl}(x_1) < k$. We define inductively $x_2, \ldots, x_k$ as follows: x_{i+1} is the least common ancestor of x_i and y_{i+1}.

As above $|T[x_i, y_{i-1}]| \geq k$ so $\mathrm{lvl}(y_i) - \mathrm{lvl}(x_i) < k$. Symmetrically, let t_1 be the least common ancestor of y and z_1 and for $1 \leq i \leq \ell - 1$, let t_{i+1} be the least common ancestor of t_i and z_{i+1}. For $1 \leq i \leq \ell$, we have $\mathrm{lvl}(z_i) - \mathrm{lvl}(t_i) < k$. Moreover, by definition all x_i and t_j are ancestors of y, so they all are on $T[u, y]$.

Let P_y (resp. P_z) be a shortest dipath in D from y_k (resp. z_ℓ) to $T[u, y] \cup Q_1[y_1, y_{k-1}] \cup Q_2[z_1, z_{\ell-1}]$. Note that P_y and P_z exist since D is strongly connected. Let y' (resp. z') be the terminal vertex of P_y (resp. P_z). Let w_y be the last vertex of $T[x_k, y_k]$ in P_y (possibly, $w_y = y_k$.) Similarly, let w_z be the last vertex of $T[t_\ell, z_\ell]$ in P_z (possibly, $w_z = z_\ell$.) Note that $P_y[w_y, y']$ is a shortest dipath from w_y to y' and $P_z[w_z, z']$ is a shortest dipath from w_z to z'.

If $y' = y_j$ for $0 \leq j \leq k-1$, consider $R = T[x_k, w_y] \odot P_y[w_y, y_j]$ as an (x_k, y_j)-dipath. By Proposition 60, R has length at least k because $\mathrm{lvl}(y_j) - \mathrm{lvl}(x_k) \geq \mathrm{lvl}(y_j) - \mathrm{lvl}(y_k) + 1 \geq k$. Therefore the union of R and $T[x_k, y] \cup Q_1[y, y_j]$ is a subdivision of $C(k, k)$, a contradiction.

Similarly, we get a contradiction if z' is in $\{z_1, \dots, z_{\ell-1}\}$. Consequently, P_y is disjoint from $Q_1[y, y_{k-1}]$ and P_z is disjoint from $Q_2[y, z_{\ell-1}]$.

Suppose P_y and P_z intersect in a vertex s. By the above statement, $s \notin V(Q) \setminus \{y_k, z_\ell\}$. Therefore the union of $Q_1 \odot P_y[y_k, s]$ and $Q_2 \odot P_z[z_\ell, s]$ is a subdivision of $C(k, \ell)$, a contradiction. Henceforth P_y and P_z are disjoint.

Assume both y' and z' are in $T[u, y]$. By symmetry, we can assume $y' \geq_T z'$ and then the union of $Q_1 \odot P_y \odot T[y', z']$ and $Q_2 \odot P_z$ form a subdivision of $C(k, \ell)$. This is a contradiction.

Henceforth a vertex among y' and z' is not in $T[u, y]$. Let us assume that y' is not in $T[u, y]$ (the case $z' \notin T[u, y]$ is similar), and so $y' = z_i$ for some $1 \leq i \leq \ell - 1$. If $\mathrm{lvl}(y') \geq \mathrm{lvl}(x_k) + k$, then both $T[x_k, w_y] \odot P_y[w_y, y']$ and $T[x_k, y] \odot Q_2[y, z_i]$ have length at least k by Proposition 60, so their union is a subdivision of $C(k, k)$, a contradiction. Hence $\mathrm{lvl}(x_k) \geq \mathrm{lvl}(z_i) - k + 1 \geq \mathrm{lvl}(z_\ell) \geq \mathrm{lvl}(t_\ell)$.

If $z' = y_j$ for some j, then necessarily $\mathrm{lvl}(z') \geq \mathrm{lvl}(x_k) + k \geq \mathrm{lvl}(t_\ell) + k$ and both $T[t_\ell, w_z] \odot P_z[w_z, z']$ and $T[t_\ell, y] \odot Q_1[y, y_j]$ have length at least k, so their union is a subdivision of $C(k, k)$, a contradiction.

Therefore $z' \in T[u, y]$. The union of $T[t_\ell, z']$ and $T[t_\ell, w_z] \odot P_z[w_z, z']$ is not a subdivision of $C(k, k)$ so by Proposition 60, $\mathrm{lvl}(z') \leq \mathrm{lvl}(t_\ell) + k - 1 \leq \mathrm{lvl}(z_\ell) + k - 1 \leq \mathrm{lvl}(z_{\ell-1})$.

If $\mathrm{lvl}(z') \leq \mathrm{lvl}(x_k)$, then the union of Q_1 and $Q_2 \odot P_z \odot T[z', y_k]$ is a subdivision of $C(k, \ell)$, a contradiction. Hence $\mathrm{lvl}(z') > \mathrm{lvl}(x_k)$. Therefore $\mathrm{lvl}(y') = \mathrm{lvl}(z_i) \leq \mathrm{lvl}(x_k) + k - 1 \leq \mathrm{lvl}(z') + k - 2 \leq \mathrm{lvl}(z_\ell) + 2k - 3$, which implies that $i = \ell - 1$ that is $y' = z_i = z_{\ell-1}$. Now the union of $T[x_1, y_1] \odot Q_1[y_1, y_k] \odot P_y$ and $T[x_1, y] \odot Q_2[y, z_{\ell-1}]$ is a subdivision of $C(k, \ell)$, a contradiction. $\Diamond$

Claims 61.1, 61.2, 61.3, and 61.4, together with Lemma 59 yield the result. $\square$

2.2.2 Cycles with four blocks in strong digraphs

Let $\hat{C}_4$ be the orientation of the 4-cycle with 4 blocks.

Theorem 62. *Let D be a digraph in* $\text{S-Forb}(\hat{C}_4)$*. If D admits an out-generator, then* $\chi(D) \leq 24$.

Proof. The general idea is the same as in the proof of Theorem 61.

Suppose that D admits an out-generator u and let T be an BFS-tree with root u We partition $A(D)$ into three sets according to the levels of u.

$$\begin{aligned}
A_0 &= \{(x,y) \in A(D) \mid \text{lvl}(x) = \text{lvl}(y)\};\\
A_1 &= \{(x,y) \in A(D) \mid |\text{lvl}(x) - \text{lvl}(y)| = 1\};\\
A_2 &= \{(x,y) \in A(D) \mid \text{lvl}(y) \leq \text{lvl}(x) - 2\}.
\end{aligned}$$

For $i = 0, 1, 2$, let $D_i = (V(D), A_i)$.

Claim 62.1. *$\chi(D_0) \leq 3$.*

Subproof. Suppose for a contradiction that $\chi(D) \geq 4$. By Theorem 27, it contains a $P^-(1,1)$ (y_1, y, y_2), that is (y, y_1) and (y, y_2) are in $A(D_0)$. Let x be the least common ancestor of y_1 and y_2 in T. The union of $T[x, y_1]$, (y, y_1), (y, y_2), and $T[x, y_2]$ is a subdivision of $\hat{C}_4$, a contradiction. $\Diamond$

Claim 62.2. *$\chi(D_1) \leq 2$.*

Subproof. Since the arcs are between consecutive levels, then the colouring ϕ_1 defined by $\phi_1(x) = lvl(x) \mod 2$ is a proper 2-colouring of D_1. $\Diamond$

Claim 62.3. *$\chi(D_2) \leq 4$.*

Subproof. Let x be a vertex of $V(D)$. If y and z are distinct outneighbours of x in D_2, then their least common ancestor w is either y or z, for otherwise the union of $T[w, y]$, (x, y), (x, z), and $T[w, z]$ is a subdivision of $\hat{C}_4$. Consequently, there is an ordering $y_1, \ldots, y_p$ of $N^+_{D_2}(x)$ such that the y_i appear in this order on $T[u, x]$.

Let us prove that, in D_2, $N^+(y_i) = \emptyset$ for $2 \leq i \leq p - 1$. Suppose for a contradiction that y_i has an outneighbour z in D_2. Let t be the least common ancestor of y_1 and z. If $t = z$, then the union of $(y_i, z) \odot T[z, y_1]$, (x, y_1), (x, y_p), and $T[y_i, y_p]$ is a subdivision of $\hat{C}_4$; if $t = y_1 \neq z$, then the union of (y_i, z), $(x, y_1) \odot T[y_1, z]$, (x, y_p), and $T[y_i, y_p]$ is a subdivision of $\hat{C}_4$. Otherwise, if $t \notin \{y_1, z\}$, $T[t, y_1]$, $T[t, z]$, $(x, y_i) \odot (y_i, z)$ and (x, y_1) is a subdivision of $\hat{C}_4$.

Henceforth, in D_2, every vertex has at most two outneighbours that are not sinks (vertices with outdegree 0). Let V_0 be the set of sinks in D_2. It is a stable set in D_2. Furthermore $\Delta^+(D_2 - V_0) \leq 2$, and since $D_2 - V_0$ is acyclic, it is 2-degenerate and thus 3-colourable. Therefore $\chi(D_2) \leq 4$. $\Diamond$

Claims 62.1, 62.2, 62.3, and Lemma 59 implies $\chi(D) \leq 24$. $\square$

2.3 Spindles

Remember that $B(k_1, k_2; k_3)$ denotes the digraph formed by three internally disjoint paths between two vertices x, y, two (x, y)-directed paths, one of order k_1, the other of order k_2, and one (y, x)-directed path of order k_3. In [16] we conjectured the following.

Conjecture 36. There is a function $g : \mathbb{N}^3 \to \mathbb{N}$ such that every strong digraph with chromatic number at least $g(k_1, k_2, k_3)$ contains a subdivision of $B(k_1, k_2; k_3)$.

As an evidence, we proved this conjecture for $k_2 = 1$ and arbitrary k_1 and k_3.

Theorem 37. *For every $k \geq 1$, there is a constant γ_k such that if D is a strong digraph with $\chi(D) > \gamma_k$, then D contains a subdivision of $B(k, 1; k)$.*

The goal of this section is to present the proof of this result.

2.3.1 Definitions and preliminaries

A (directed) graph G is *k-degenerate* if every subgraph H of G has a vertex of degree at most k. The following statements are well-known.

Proposition 63 (Folklore)**.** *Every k-degenerate (directed) graph is $(k+1)$-colourable.*

Theorem 64 (Brooks [12])**.** *Let G be a connected graph. Then $\chi(G) \leq \Delta(G)$ unless G is a complete graph or an odd cycle.*

Lemma 65. *Let D be a digraph, $D_1, \dots, D_l$ be disjoint subdigraphs of D and D' the digraph obtained by contracting each D_i into one vertex d_i. Then $\chi(D) \leq \chi(D') \cdot \max\{\chi(D_i) \mid i \in [l]\}$.*

Proof. Set $k_1 = \max\{\chi(D_i) \mid i \in [l]\}$ and $k_2 = \chi(D')$. For each i, let ϕ_i be a proper colouring of D_i using colours in $[k_1]$ and let ϕ' be a proper colouring of D' using colours in $[k_2]$. Define $\phi : V(D) \to [k_1] \times [k_2]$ as follows. If x is a vertex belonging to some D_i, then $\phi(x) = (\phi_i(x), \phi'(d_i))$, else $\phi(x) = (1, \phi'(x))$. Let x and y be adjacent vertices of D. If they belong to the same subdigraph D_i, then $\phi_i(x) \neq \phi_i(y)$ and so $\phi(x) \neq \phi(y)$. If they do not belong to the same component, then the vertices corresponding to these vertices in D' are adjacent and so $\phi(x) \neq \phi(y)$. Thus ϕ is a proper colouring of D using $k_1 \cdot k_2$ colours. □

The *rotative tournament on $2k-1$ vertices*, denoted by R_{2k-1}, is the tournament with vertex set $\{v_1, \dots, v_{2k-1}\}$ in which v_i dominates v_j if and only if $j - i$ modulo $2k-1$ belongs to $\{1, 2, \dots, k-1\}$.

Proposition 66. *Every strong tournament of order $2k-1$ contains a $B(k, 1; 1)$-subdivision.*

Proof. Let T be a strong tournament of order $2k-1$. By Camion's Theorem, it has a Hamiltonian directed cycle $C = (v_1, v_2, \dots, v_{2k-1}, v_1)$. If there exists an arc (v_i, v_j) with $j - i \geq k$ (indices are modulo $2k-1$), then the union of $C[v_i, v_j]$, (v_i, v_j) and $C[v_j, v_i]$ is a $B(k, 1; 1)$-subdivision. Henceforth, we may assume that $T = R_{2k-1}$. Then the union of $C[v_1, v_{k-1}] \odot (v_{k-1}, v_{k+1}, v_{k+2})$, (v_1, v_k, v_{k+2}), and $C[v_{k+2}, v_1]$ is a $B(k, 1; 1)$-subdivision. □

We will need the following lemmas:

Lemma 67. *Let $\sigma = (u_1, \dots, u_p)$ be a sequence of integers in $[k]$, and let l be a positive integer. If $p \geq l^k$, then there exists a set L of l indices such that for any $i, j \in L$ with $i < j$ the following holds : $u_i = u_j$ and $u_t > u_i$, for all $i < t < j$.*

Proof. By induction on k. The result holds trivially when $k = 1$. Assume now that $k > 1$. Let L_1 be the elements of the sequence with value 1. If L_1 has at least l elements, we are done. If not, then there is a subsequence σ' of $\left\lceil \frac{l^k-(l-1)}{l} \right\rceil = l^{k-1}$ consecutive elements in $\{2, \dots, k-1\}$. Applying the induction hypothesis to σ' yields the result. □

Lemma 68. *Let $\sigma = (u_1, \dots, u_p)$ be a sequence of integers in $[k]$. If $p > k(m-1)$, then there exists a subsequence of m consecutive integers such that the last one is the largest.*

Proof. By induction on k. The result holds trivially when $k = 1$. Let i be the smallest integer such that $u_t \leq k-1$ for all $t \geq i$. If $i > m$, then $u_{i-1} = k$, and the subsequence of the $i-1$ first elements of σ is the desired sequence. If $i \leq m$, apply the induction on $\sigma' = (u_t)_{i \leq t \leq p}$ which is a sequence of more than $(k-1)(m-1)$ integers in $[k-1]$, to get the result. □

2.3.2 $B(k, 1; 1)$

Before proving the main result for spindles, we present a proof of a weaker result in order to introduce some of the ideas.

Theorem 69. *Let $k \geq 3$ be an integer and let D be a strong digraph. If $\chi(D) > (2k-2)(2k-3)$, then D contains a subdivision of $B(k, 1; 1)$.*

Let $\mathcal{C}$ be a collection of directed cycles. It is *nice* if all cycles of $\mathcal{C}$ have length at least $2k-2$, and any two distinct cycles of $\mathcal{C}$ intersect on at most one vertex. A *component* of $\mathcal{C}$ is a connected component in the adjacency graph of $\mathcal{C}$, where vertices correspond to cycles in $\mathcal{C}$ and two vertices are adjacent if the corresponding cycles intersect. Note that if $\mathcal{S}$ is a component of $\mathcal{C}$, then $\bigcup \mathcal{S}$ is both a connected component and a strong component of $\bigcup \mathcal{C}$. For sake of simplicity, we denote by $D[\mathcal{S}]$ the digraph $D[\bigcup \mathcal{S}]$. Observe that this digraph contains $\bigcup \mathcal{S}$ but has more arcs.

We will prove that every $B(k, 1; 1)$-subdivision-free strong digraph D has bounded chromatic number in the following way: We take a maximal nice collection $\mathcal{C}$ of directed cycles. We will prove that for every component $\mathcal{S}$ of $\mathcal{C}$, the digraph $D[\mathcal{S}]$ has bounded chromatic number. Then we will prove that, since it contains no long directed cycle and it is strong, the digraph $D_{\mathcal{C}}$ obtained from D by contracting each component of $\mathcal{C}$ into one vertex has bounded chromatic number. Those two results allow us to conclude by Lemma 65.

We will need the following lemma:

Lemma 70. *Let $\mathcal{C}$ be a nice collection of directed cycles in a $B(k, 1; 1)$-subdivision-free digraph D and let C, C' be two cycles of the same component $\mathcal{S}$ of $\mathcal{C}$. There is no dipath P from C to C' whose arcs are not in $A(\bigcup \mathcal{S})$.*

Proof. By the contrapositive. We suppose that there exists such a dipath P and show that there is a $B(k,1;1)$-subdivision in D.

By definition of $\mathcal{S}$, there exists a dipath Q from C' to C in $\bigcup \mathcal{S}$. By choosing C and C' such that Q is as small as possible, then $s(Q) \neq t(P)$ and $t(Q) \neq s(P)$ (note that $s(Q)$ and $t(Q)$ can be the same vertex).

Since C has length at least $2k-2$, either $C[t(Q), s(P)]$ has length at least $k-1$ or $C[s(P), t(Q)]$ has length at least k.

- If $C[t(Q), s(P)]$ has length at least $k-1$, then the union of $Q \odot C[t(Q), s(P)] \odot P$, $C'[s(Q), t(P)]$ and $C'[t(P), s(Q)]$ is a $B(k,1;1)$-subdivision between $s(Q)$ and $t(P)$.
- If $C[s(P), t(Q)]$ has length at least k, then the union of $C[s(P), t(Q)]$, $P \odot C'[t(P), s(Q)] \odot Q$ and $C[t(Q), s(P)]$ is a $B(k,1;1)$-subdivision between $s(P)$ and $t(Q)$.

□

Lemma 71. *Let $k \geq 3$ be an integer, and let $\mathcal{C}$ be a nice collection of directed cycles in a $B(k,1;1)$-subdivision-free digraph D and $\mathcal{S}$ a component of $\mathcal{C}$. Then $\chi(D[S]) \leq 2k-2$.*

Proof. By induction on the number of directed cycles in $\mathcal{S}$. Let C be a cycle of $\mathcal{S}$. There is no chord (x,y) of C such that $C[x,y]$ has length at least k, for otherwise there would be a $B(k,1;1)$-subdivision. Hence $D[C]$ has maximum degree at most $2k-2$. Moreover, by Proposition 66, $D[C]$ is not a tournament of order $2k-1$. Thus, by Brooks' Theorem (64), $\chi(D[C]) \leq 2k-2$. Let c be a proper colouring of C with $2k-2$ colours. Let $\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_r$ be the components of $\mathcal{S} \setminus \{C\}$. Since $\mathcal{S}$ is the union of the $\mathcal{S}_l$ for $l \in [r]$, and $\{C\}$, each $\mathcal{S}_l$ has less cycles than $\mathcal{S}$. By the induction hypothesis, there exists a proper colouring c_l using $2k-2$ colours for each $D[\mathcal{S}_l]$.

Now, we claim that each $D[\mathcal{S}_l]$ intersects C in exactly one vertex. It is easy to see that C must intersect at least one cycle of each $\mathcal{S}_l$. Now suppose there exist two vertices of C, x and y, in $D[\mathcal{S}_l]$. By definition of a nice collection, they cannot belong to the same cycle of $\mathcal{S}_l$, so there exist two cycles C_i and C_j of S_l such that $x \in C_i$ and $y \in C_j$. Now $C[x,y]$ is a dipath from C_i to C_j whose arcs are not in $A(\bigcup \mathcal{S}_l)$. This contradicts Lemma 70.

Consequently, free to permute the colours of c_l, we may assume that each vertex of C receives the same colour in c and in c_l. In addition, by Lemma 70, there is no arc between different $D[\mathcal{S}_l]$ nor between $D[\mathcal{S}_l]$ and C. Hence the union of c_l and c is a proper colouring of $D[\mathcal{S}]$ using $2k-2$ colours. □

Lemma 72. *Let $\mathcal{C}$ be a maximal nice collection of directed cycles in a $B(k,1;1)$-subdivision-free strong digraph D and $D_{\mathcal{C}}$ the digraph obtained from D by contracting each component of $\mathcal{C}$ into one vertex. Then $\chi(D_{\mathcal{C}}) \leq 2k-3$.*

Proof. First note that since D is strong, then so is $D_{\mathcal{C}}$. Suppose $\chi(D_{\mathcal{C}}) \geq 2k-2$. By Bondy's Theorem (31), there exists a directed cycle $C = (x_1, \ldots, x_l, x_1)$ of length at least $2k-2$ in $D_{\mathcal{C}}$. We derive a directed cycle C' in D the following way: Suppose the vertex x_i corresponds to a component $\mathcal{S}_i$ of $\mathcal{C}$: the arc (x_{i-1}, x_i) corresponds in D to an arc whose head is a vertex

p_i of $\bigcup \mathcal{S}_i$, and the arc (x_i, x_{i+1}) corresponds to an arc whose tail is a vertex l_i of $\bigcup \mathcal{S}_i$. Let P_i be a dipath from p_i to l_i in $D[\mathcal{S}_i]$. Note that P_i intersects each cycle of S_i on a, possibly empty, subdipath of P_i. Then C' is the directed cycle obtained from C by replacing the vertices x_i by the path P_i.

C' is a directed cycle of D of length at least $2k-2$ because it is no shorter than C. Let C_1 be a cycle of $\mathcal{C}$. By construction of C' and $D_{\mathcal{C}}$, C' and C_1 can intersect only along a subdipath of one P_i. Suppose this dipath is more than just one vertex. Let x and y be the initial and terminal vertex, respectively, of this dipath. Then the union of $C'[x,y]$, $C_1[x,y]$ and $C_1[y,x]$ is a $B(k,1;1)$-subdivision, a contradiction.

So C' is a directed cycle of length at least $2k-2$, intersecting each cycle of $\mathcal{C}$ on at most one vertex, and which does not belong to $\mathcal{C}$, for otherwise it would be reduced to one vertex in $D_{\mathcal{C}}$. This contradicts the fact that $\mathcal{C}$ is maximal. □

We can finally prove Theorem 69.

Proof of Theorem 69. Let $\mathcal{C}$ be a maximal nice collection of directed cycles in D. Lemmas 71, 72 and 65 give the result. □

2.3.3 $B(k,1;k)$

We will now present a proof of Theorem 37.

We prove the result by the contrapositive. We consider a $B(k,1;k)$-subdivision-free digraph D. We shall prove that $\chi(D) \leq \gamma_k = 8k^2(4k^2+2)(2\cdot(4k)^{4k}+1)(2\cdot(6k^2)^{3k}+14k)$.

Our proof heavily uses the notion of k-suitable collection of directed cycles, which can be seen as a generalization of the notion of nice collection of directed cycles used to prove Theorem 69.

A collection $\mathcal{C} = \{C_1, C_2, \ldots, C_N\}$ of directed cycles is *k-suitable* if all cycles of $\mathcal{C}$ have length at least $8k$, and for any two distinct directed cycles $C_i, C_j \in \mathcal{C}$, the intersection $C_i \cap C_j$ is either empty or a dipath of order at most k, denoted by $P_{i,j}$. We denote by $s_{i,j}$ (resp. $t_{i,j}$) the initial (resp. terminal) vertex of $P_{i,j}$.

The proof of Theorem 37 uses the same general idea as Theorem 69: take a maximal k-suitable collection of directed cycles $\mathcal{C}$; show that the digraph $D_{\mathcal{C}}$ obtained by contracting the components of $\mathcal{C}$ has bounded chromatic number, and that each component also has bounded chromatic number; conclude using Lemma 65. However, because the intersection of cycles in this collection are more complicated and because there might be arcs between directed cycles of the same component, bounding the chromatic number of the components is way more challenging. The next part is devoted to this.

k-suitable collections of directed cycles

Let ϕ be a colouring of a graph G. A subset of vertices or a subgraph S of G is *rainbow-coloured* by ϕ if all vertices of S have distinct colours.

Set $\alpha_k = 2\cdot(6k^2)^{3k} + 14k$. The first step of the proof is the following lemma.

Lemma 73. *Let $\mathcal{C}$ be a k-suitable collection of directed cycles in a $B(k,1;k)$-subdivision-free digraph. There exists a proper colouring ϕ of $\bigcup \mathcal{C}$ with α_k colours, such that, each subdipath of length $7k$ of each directed cycle of $\mathcal{C}$ is rainbow-coloured.*

In order to prove this lemma, we need some definitions and preliminary results.

Lemma 74. *Let $\mathcal{C}$ be a k-suitable collection of directed cycles in a $B(k,1;k)$-subdivision-free digraph. Let C_1, C_2, C_3 be three pairwise-intersecting directed cycles of $\mathcal{C}$, and let v be a vertex in $V(C_2) \cap V(C_3) \setminus V(C_1)$. Then exactly one of the following holds:*

(i) $C_2[t_{1,2}, v]$ and $C_3[t_{1,3}, v]$ both have length less than $3k$;

(ii) $C_2[v, s_{1,2}]$ and $C_3[v, s_{1,3}]$ both have length less than $3k$.

Proof. Observe first that since C_2 has length at least $8k$ and $P_{1,2}$ has length at most $k-1$, the sum of the lengths of $C_2[t_{1,2}, v]$ and $C[v, s_{1,2}]$ is at least $7k+1$. Similarly, the sum of the lengths of $C_2[t_{1,3}, v]$ and $C[v, s_{1,3}]$ is at least $7k+1$. In particular, if (i) holds, then (ii) does not hold and vice-versa.

Suppose for a contradiction that none of (i) and (ii) holds. By symmetry and the above inequalities, we may assume that both $C_2[t_{1,2}, v]$ and $C_3[v, s_{1,3}]$ have length more than $3k$. But $v \notin V(C_1)$, so $v \notin V(P_{1,3})$. Thus $C_3[v, t_{1,3}]$ has also length at least $3k$ as well.

If there is a vertex in $V(C_1) \cap V(C_2) \cap V(C_3)$, then $C_3[v, t_{1,3}]$ would have length less than $2k$ (since it would be contained in $P_{2,3} \cup P_{1,3}$ and each of those paths has length less than k), a contradiction. Hence $V(C_1) \cap V(C_2) \cap V(C_3) = \emptyset$. In particular, $P_{1,2}$, $P_{1,3}$, and $P_{2,3}$ are disjoint.

The dipath $C_2[s_{1,2}, t_{2,3}]$ has length at least $3k$ because it contains $C_2[t_{1,2}, v]$. Moreover, the dipath $C_3[t_{2,3}, s_{1,3}]$ has length at least $2k$ because $C_3[v, s_{1,3}]$ has length at least $3k$ and $C_3[v, t_{2,3}]$ has length less than k. Thus $C_3[t_{2,3}, s_{1,3}] \odot C_1[s_{1,3}, s_{1,2}]$ has length at least $2k$. Consequently, the union of $C_2[s_{1,2}, t_{2,3}]$, $C_2[t_{2,3}, s_{1,2}]$, and $C_3[t_{2,3}, s_{1,3}] \odot C_1[s_{1,3}, s_{1,2}]$ is a $B(k,1;k)$-subdivision, a contradiction. $\square$

Let $\mathcal{C}$ be a k-suitable collection of directed cycles. For every set of vertices or digraph S, we denote by $\mathcal{C} \cap S$ the set of directed cycles of $\mathcal{C}$ that intersect S.

Let $C_1 \in \mathcal{C}$. For each $C_j \in \mathcal{C} \cap C_1$ such that $C_j \neq C_1$, let Q_j be the subdipath of C_j containing all the vertices that are at distance at most $3k$ from $P_{1,j}$ in the cycle underlying C_j. Then the dipaths $C_j[s(Q_j), s_{1,j}]$ and $C_j[t_{1,j}, t(Q_j)]$ have length $3k$. Set $Q_j^- = C[s(Q_j), s_{1,j}[$ and $Q_j^+ = C]t_{1,j}, t(Q_j)]$.

Set $I(C_1) = C_1 \cup \bigcup_{C_j \in \mathcal{C} \cap C_1} Q_j$, $I^+(C_1) = \bigcup_{C_j \in \mathcal{C} \cap C_1} Q_j^+$ and $I^-(C_1) = \bigcup_{C_j \in \mathcal{C} \cap C_1} Q_j^-$. Observe that Lemma 74 implies directly the following.

Corollary 75. *Let $\mathcal{C}$ be a k-suitable collection of directed cycles and let $C_1 \in \mathcal{C}$.*

(i) $I^+(C_1)$ and $I^-(C_1)$ are vertex-disjoint digraphs.

(ii) $I^-(C_1) \cap C_j = Q_j^-$ and $I^+(C_1) \cap C_j = Q_j^+$, for all $C_j \in \mathcal{C} \cap C_1$.

Lemma 76. *Let $\mathcal{C}$ be a k-suitable collection of directed cycles in a $B(k,1;k)$-subdivision-free digraph D. Let $C_1 \in \mathcal{C}$ and let A be a connected component of $\bigcup \mathcal{C} - I(C_1)$. Then all vertices of $\bigcup(\mathcal{C} \cap A) - A$ belong to a unique directed cycle C_A in $\mathcal{C}$.*

Proof. Suppose it is not the case. Then there are two distinct directed cycles C_2, C_3 of $\mathcal{C} \cap A$ that intersect with C_1. Observe that there is a sequence of distinct directed cycles $C_2 = C_1^*, C_2^*, \ldots, C_q^* = C_3$ of $\mathcal{C} \cap A$ such that $C_j^* \cap C_{j+1}^* \neq \emptyset$ because A is a connected component of $\bigcup \mathcal{C} - I(C_1)$. Free to consider the first $C_j^* \neq C_2$ in this sequence such that $V(C_j^*) \not\subseteq A$ in place of C_3, we may assume that all C_j^*, $2 \leq j \leq q-1$, have all their vertices in A. In particular, there exists a (C_3, C_2)-dipath Q_A in $D[A]$.

Let $R_3 = C_1[t_{1,2}, t_{1,3}] \odot Q_3^+$. Clearly, R_3 has length at least $3k$. Let v be the last vertex in $Q_2 \cap R_3$ along Q_2. (This vertex exists since $t_{1,2} \in Q_2 \cap R_3$.) Since there is a (C_3, C_2)-dipath in $D[A]$, by Corollary 75, $C_3[t(Q_3), s(Q_A)]$ is in $D[A]$. Thus there exists a $(t(Q_3), C_2)$-dipath R_A in $D[A]$. Let w be its terminal vertex. By definition of A, w is in $C_2[t(Q_2), s(Q_2)]$, therefore $C_2[w, v]$ has length at least $3k$ since it contains $C_2[s(Q_2), s_{1,2}]$. Consequently, both $C_2[v, t(Q_2)]$ and $R_3[v, t(Q_3)]$ have length less than k for otherwise the union of $C_2[w, v]$, $C_2[v, w]$ and $R_3[v, t(Q_3)] \odot R_A$ would be a $B(k,1;k)$-subdivision. In particular, $v \neq t(Q_2)$. This implies that $s_{2,3} \in V(Q_2 \cap R_3)$. Moreover, $Q_2[s_{2,3}, t(Q_2)]$ has length less than $2k$ because $Q_2[s_{2,3}, v]$ is a subdipath of $P_{2,3}$ and so has length less than k. Therefore $C_2[t_{1,2}, s_{2,3}] = Q_2[t_{1,2}, s_{2,3}]$ has length at least k because Q_2 has length at least $3k$. It follows that the union of $C_2[s_{2,3}, t_{1,2}]$, $C_2[t_{1,2}, s_{2,3}]$ and $R_3[t_{1,2}, s_{2,3}]$ is a $B(k,1;k)$-subdivision, a contradiction. $\square$

Lemma 77. *Let $\mathcal{C}$ be a k-suitable collection of directed cycles in a $B(k,1;k)$-subdivision-free digraph. For any directed cycle $C_1 \in \mathcal{C}$, the digraph $I^+(C_1)$ has no directed cycle.*

Proof. Suppose for a contradiction that $I^+(C_1)$ contains a directed cycle C'. Clearly, it must contain arcs from at least two Q_j^+.

Assume that C' contains several vertices of Q_j^+. Necessarily, there must be two vertices x, y of $Q_j^+ \cap C'$ such that no vertex of $C']x, y[$ is in C_j and y is before x in Q_j^+. Therefore $C'[x, y] \odot Q^+[y, x]$ is also a directed cycle in $I^+(C_1)$. Free to consider this cycle, we may assume that $C' \cap Q_j^+$ is a dipath.

Doing so, for all j, we may assume that $C' \cap Q_j^+$ is a dipath for every $C_j \in \mathcal{C} \cap C_1$. Without loss of generality, we may assume that there are directed cycles $C_2, \ldots, C_p$ such that

- C' is in $Q_2^+ \cup \cdots \cup Q_p^+$;
- for all $2 \leq j \leq p$, $C' \cap Q_j^+$ is a dipath P_j^+ with initial vertex a_j and terminal vertex b_j;
- the a_j and the b_j appear according to the following order around C': $(a_2, b_p, a_3, b_2, \ldots, a_p, b_{p-1}, a_2)$ with possibly $a_{j+1} = b_j$ for some $1 \leq j \leq p$ where $a_{p+1} = a_2$.

For $2 \leq j \leq p$, set $B_j = C_j[b_j, a_j]$. Note that B_j has length at least $4k$, because Q_2^+ has length less than $3k$.

Consider the closed directed walk

$$W = C_p[a_2, b_p] \odot B_p \odot C_{p-1}[a_p, b_{p-1}] \odot \cdots \odot B_3 \odot C_2[a_3, b_2] \odot B_2.$$

W contains a directed cycle C_W. Wihtout loss of generality, we may assume that this cycle is of the form

$$C_W = B_q[v, a_q] \odot C_{q-1}[a_q, b_{q-1}] \odot \cdots \odot B_3 \odot C_2[a_3, b_2] \odot B_2[b_2, v]$$

for some vertex $v \in B_2 \cap B_q$. (The case when W is a directed cycle corresponds to $q = p+1$ and $B_2 = B_{p+1}$.)

Note that necessarily, $q \geq 4$, for B_3 does not intersect B_2, for otherwise $b_3 = b_2$ since the intersection of C_2 and C_3 is a dipath.

Observe that $C_W[b_2, v] = C_2[b_2, v]$ or $C_W[v, a_4]$ has length at least k. Indeed, if $q = p+1$, then it follows from the fact that B_2 has length as least $4k$; if $5 \leq q \leq p$, then it comes from the fact that B_4 is a subdipath of $C_W[v, a_r]$; if $q = 4$, then it follows from Lemma 74 applied to C_3, C_2, C_4 in the role of C_1, C_2, C_3 respectively. In both cases, $C_W[b_2, a_4]$ has length at least k.

Furthermore, $C_W[a_4, b_2]$ has length at least k because it contains B_3. Therefore the union of $C_W[b_2, a_4]$, $C_W[a_4, b_2]$ and $C'[b_2, a_4] = C_3[b_3, a_4]$ is a $B(k, 1; k)$-subdivision, a contradiction. $\square$

Lemma 78. *Let $\mathcal{C}$ be a k-suitable collection of directed cycles in a $B(k, 1; k)$-subdivision-free digraph.*

Let ϕ be a partial colouring of a directed cycle $C_1 \in \mathcal{C}$ such that only a path of length at most $7k$ is coloured and this path is rainbow-coloured. Then ϕ can be extended into a colouring of $I(C_1)$ using α_k colours, such that every subdipath of length at most $7k$ of C_1 is rainbow-coloured and Q_j is rainbow-coloured, for every $C_j \in \mathcal{C} \cap C_1$.

Proof. We can easily extend ϕ to C_1 using $14k$ colours (including the at most $7k$ already used colours) so that every subdipath of C_1 of length $7k$ is rainbow-coloured.

We shall now prove that there exists a colouring ϕ^+ of $I^+(C_1)$ with $(6k^2)^{3k}$ (new) colours so that Q_j^+ is rainbow-coloured for every $C_j \in \mathcal{C} \cap C_1$, and a colouring ϕ^- of $I^-(C_1)$ with $(6k^2)^{3k}$ (other new) colours so that Q_j^- is rainbow-coloured for every $C_j \in \mathcal{C} \cap C_1$. The union of the three colourings ϕ, ϕ^+, and ϕ^- is clearly the desired colouring of $I(C_1)$. (Observe that a vertex of $I(C_1)$ is coloured only once because C_1, $I^+(C_1)$ and $I^-(C_1)$ are disjoint by Corollary 75.)

It remains to prove the existence of ϕ^+ and ϕ^-. By symmetry, it suffices to prove the existence of ϕ^+. To do so, we consider an auxiliary digraph D_1^+. For each $C_j \in \mathcal{C} \cap C_1$, let T_j^+ be the transitive tournament whose Hamiltonian dipath is Q_j^+. Let $D_1^+ = \bigcup_{C_j \in \mathcal{C} \cap C_1} T_j^+$. The arcs of $A(T_j^+) \setminus A(Q_j^+)$ are called *fake arcs*. Clearly, ϕ^+ exists if and only if D_1^+ admits a proper $(6k^2)^{3k}$-colouring. Henceforth it remains to prove the following claim.

Claim 78.1. $\chi(D_1^+) \leq (6k^2)^{3k}$.

Subproof. To each vertex v in $I^+(C_1)$ we associate the set $\mathrm{Dis}(v)$ of the lengths of the $C_j[t_{1,j}, v]$ for all directed cycles $C_j \in \mathcal{C} \cap C_1$ containing v such that $C_j[t_{1,j}, v]$ has length at most $3k$.

Suppose for a contradiction that $\chi(D_1^+) \leq (6k^2)^{3k}$. By Theorem 26, D_1^+ admits a dipath of length $(6k^2)^{3k}$. Replacing all fake arcs (u, v) in some $A(T_j^+)$, by $Q_j^+[u, v]$ we obtain a directed walk P in $I^+(C_1)$ of length at least $(6k^2)^{3k}$. By Lemma 77, P is necessarily a dipath. Set $P = (v_1, \ldots, v_p)$. We have $p \geq (6k^2)^{3k}$.

For $1 \leq i \leq p$, let $m_i = \min \mathrm{Dis}(v_i)$. Lemma 67 applied to $(m_i)_{1\leq i\leq p}$ yields a set L of $6k^2$ indices such that for any $i < j \in L$, $m_i = m_j$ and $m_k > m_i$, for all $i < k < j$. Let $l_1 < l_2 < \cdots < l_{6k^2}$ be the elements of L and let $m = m_{l_1} = \cdots = m_{l_{6k^2}}$.

For $1 \leq j \leq 6k^2 - 1$, let $M_j = \max \bigcup_{l_j \leq i < l_{j+1}} \mathrm{Dis}(v_i)$. By definition $M_j \leq 3k$. Applying Lemma 68 to $(M_j)_{1\leq j\leq 6k^2}$, we get a sequence of size $2k$ $M_{j_0+1}, \ldots, M_{j_0+2k}$ such that M_{j_0+2k} is the greatest. For sake of simplicity, we set $\ell_i = j_0 + i$ for $1 \leq i \leq 2k$. Let f be the smallest index not smaller than ℓ_{2k} for which $M_{\ell_{2k}} \in \mathrm{Dis}(v_f)$.

Let j_1 be an index such that $C_{j_1}[t_{1,j_1}, v_{\ell_1}]$ has length m and set $P_1 = C_{j_1}[t_{1,j_1}, v_{\ell_1}]$. Let j_2 be an index such that $C_{j_2}[t_{1,j_2}, v_{\ell_k}]$ has length m and set $P_2 = C_{j_2}[t_{1,j_2}, v_{\ell_k}]$. Let j_3 be an index such that $C_{j_3}[t_{1,j_3}, v_f]$ has length $M_{\ell_{2k}}$ and set $P_3 = C_{j_3}[v_f, s_{1,j_3}]$ (some vertices of P_3 are not in $I^+(C_1)$).

Note that any internal vertex x of P_1 or P_2 has an integer in $\mathrm{Dis}(x)$ which is smaller than m and every internal vertex y of P_3 has an integer in $\mathrm{Dis}(y)$ which is greater than $M_{\ell_{2k}}$, or does not belong to $I^+(C_1)$. Hence, P_1, P_2 and P_3 are internally disjoint from $P[v_{\ell_1}, v_f]$.

We distinguish between the intersections of P_1, P_2 and P_3:

- Suppose P_3 does not intersect $P_1 \cup P_2$.
  - Assume first that P_1 and P_2 are disjoint. If $s(P_1)$ is in $C_1[t(P_3), s(P_2)]$, then the union of $P_1 \odot P[v_{\ell_1}, v_{\ell_k}]$, $P[v_{\ell_k}, v_f] \odot P_3 \odot C_1[t(P_3), s(P_1)]$ and $C_1[s(P_1), s(P_2)] \odot P_2$ is a $B(k, 1; k)$-subdivision, a contradiction. If $s(P_1)$ is in $C_1[s(P_2), t(P_3)]$, then the union of $C_1[s(P_2), s(P_1)] \odot P_1 \odot P[v_{\ell_1}, v_{\ell_k}]$, $P[v_{\ell_k}, v_f] \odot P_3 \odot C_1[t(P_3), s(P_2)]$, and P_2 is a $B(k, 1; k)$-subdivision, a contradiction.
  - Assume now P_1 and P_2 intersect. Let u be the last vertex along P_2 on which they intersect. The union of $P_1[u, v_{\ell_1}] \odot P[v_{\ell_1}, v_{\ell_k}]$, $P[v_{\ell_k}, v_f] \odot P_3 \odot C[t(P_3), s(P_1)] \odot P_1[s(P_1), u]$, and $P_2[u, v_{\ell_k}]$ is a $B(k, 1; k)$-subdivision, a contradiction.
- Assume P_3 intersects $P_1 \cap P_2$. Let v be the first vertex along P_3 in $P_1 \cap P_2$ and let u be the last vertex of $P_1 \cap P_2$ along P_2. The union of $P_1[u, v_{\ell_1}] \odot P[v_{\ell_1}, v_{\ell_k}]$, $P[v_{\ell_k}, v_f] \odot P_3[v_f, v] \odot P_1[v, u]$, and $P_2[u, v_{\ell_k}]$ is a $B(k, 1; k)$-subdivision, a contradiction.
- Assume now that P_3 intersects $P_1 \cup P_2$ but not $P_1 \cap P_2$. Let v be the first vertex along P_3 in $P_1 \cup P_2$.
  - If $v \in P_2$, let u be the last vertex on $P_2 \cap P_3$ along P_3. Observe that $P_3[v, u]$ is also a subdipath of P_2 and therefore contains no vertex of P_1. Furthermore, there is a dipath Q from u to v_{ℓ_1} in $P_3[u, t(P_3)] \cup C_1 \cup P_1$. Hence, the union of

$P[v_{\ell_k}, v_f] \odot P_3[v_f, v]$, $Q \odot P[v_{\ell_1}, v_{\ell_k}]$, and $P_2[u, v_{\ell_k}]$ is a $B(k,1;k)$-subdivision, a contradiction.

– If $v \in P_1$, let u be the last vertex on $P_1 \cap P_3$ along P_3. Observe that $P_3[v,u]$ is also a subdipath of P_1 and therefore contains no vertex of P_2. Furthermore, there is a dipath Q from u to v_{ℓ_k} in $P_3[u, t(P_3)] \cup C_1 \cup P_2$. The union of $P[v_{\ell_k}, v_f] \odot P_3[v_f, u]$, $P_1[u, v_{\ell_1}] \odot P[v_{\ell_1}, v_{\ell_k}]$ and Q is a $B(k,1;k)$-subdivision, a contradiction.

$\Diamond$

Claim 78.1 shows the existence of ϕ^+ and completes the proof of Lemma 78. $\square$

We are now ready to prove Lemma 73. In fact, we prove the following stronger statement.

Lemma 79. *If there exists a partial colouring ϕ such that one of the directed cycle C_1 has a path of length less than $7k$ which is rainbow-coloured, then we can extend this colouring to all $D[\mathcal{C}]$ using less than α_k colours such that, on each directed cycle, every subdipath of length $7k$ is rainbow-coloured.*

Proof. By induction on the number of directed cycles in $\mathcal{C}$. Consider a rainbow-colouring of a subdipath of length less than $7k$ of a directed cycle $C_1 \in \mathcal{C}$. By Lemma 78, we can extend this colouring to a colouring ϕ_1 of $I(C_1)$ at most α_k colours. Note that the non-coloured vertices of $\bigcup \mathcal{C}$ are in one of the connected components of $\bigcup \mathcal{C} - I(C_1)$. Let A be a connected component of $\bigcup \mathcal{C} - I(C_1)$. The coloured (by ϕ_1) vertices of $\mathcal{C} \cap A$ are those of $(\mathcal{C} \cap A) - A$. Hence, by Lemma 76, they all belong to some directed cycle C_j and so to the dipath Q_j which has length at most $7k$. Hence, by the induction hypothesis, we can extend ϕ_1 to A. Doing this for each component, we extend ϕ_1 to the whole $\bigcup \mathcal{C}$. $\square$

Set $\beta_k = k(4k^2+2)(2\cdot(4k)^{4k}+1)\alpha_k$. The second step of the proof is the following lemma.

Lemma 80. *Let $\mathcal{C}$ be a k-suitable collection of directed cycles in a $B(k,1;k)$-subdivision-free digraph D. For every component $\mathcal{S}$ of $\mathcal{C}$, we have $\chi(D[\mathcal{S}]) \le \beta_k$.*

Proof. We define a sort of Breadth-First-Search for $\mathcal{S}$. Let C_0 be a directed cycle of $\mathcal{S}$ and set $L_0 = \{C_0\}$. For every directed cycle C_s of $\mathcal{S} \cap C_0$, we put C_s in level L_1 and say that C_0 is the *father* of C_s. We build the levels L_i inductively until all directed cycles of $\mathcal{S}$ are put in a level : L_{i+1} consists of every directed cycle C_l not in $\bigcup_{j\le i} L_j$ such that there exists a directed cycle in L_i intersecting C_l. For every $C_l \in L_{i+1}$, we choose one of the directed cycles in L_i intersecting it to be its *father*. Henceforth every directed cycle in L_{i+1} has a unique father even though it might intersect many directed cycles of L_i. A directed cycle C is an *ancestor* of C' if there is a sequence $C = C_1, \dots, C_q = C'$ such that C_i is the father of C_{i+1} for all $i \in [q-1]$.

For a vertex x in $\bigcup \mathcal{S}$, we say that x *belongs to* level L_i if i is the smallest integer such that there exists a directed cycle in L_i containing x. Observe that the vertices of each directed cycle C_l of $\mathcal{S}$ belong to consecutive levels, that is there exists i such that $V(C_l) \subseteq L_i \cup L_{i+1}$.

To bound the chromatic number of $D[\mathcal{S}]$, we partition its arc set in (A_0, A_1, A_2), where

- A_0 is the set of arcs of $D[\mathcal{S}]$ whose ends belong to the same level, and
- A_1 is the set of arcs of $D[\mathcal{S}]$ whose ends belong to different levels i and j with $|i-j| < k$.
- A_2 is the set of arcs of $D[\mathcal{S}]$ whose ends belong to different levels i and j with $|i-j| \geq k$.

For $i \in \{0, 1, 2\}$, let D_i be the spanning subdigraph of $D[\mathcal{S}]$ with arc set A_i. We shall now bound the chromatic numbers of D_0, D_1 and D_2.

Claim 80.1. $\chi(D_1) \leq k$.

Subproof. Let ϕ_1 be the colouring that assigns to all vertices of level L_i the colour i modulo k, it is easy to see that ϕ_1 is a proper colouring of D_1. $\Diamond$

Let C_l be a directed cycle of L_i, $i \geq 1$ and $C_{l'}$ its father. Let p_l^+ and r_l^+ be the vertices such that $C_l[t_{l,l'}, p_l^+]$ and $C_l[p_l^+, r_l^+]$ have length k. Let p_l^- and r_l^- be the vertices such that $C_l[p_l^-, s_{l,l'}]$ and $C_l[r_l^-, p_l^-]$ have length k. Let R_l^- be the set of vertices of $C_l]r_l^-, s_{l,l'}[$, P_l^- the set of vertices of $C_l]p_l^-, s_{l,l'}[$, R_l^+ the set of vertices of $C_l]t_{l,l'}, r_l[$, P_l^+ the set of vertices of $C_l]t_{l,l'}, p_l^+[$, and finally let R_l' be the set of vertices belonging to L_i in $C_l \setminus \{R_l^+ \cup R_l^-\}$.

Claim 80.2. *Let x be a vertex in L_i with $i \geq 1$. Let C_l and C_m be two directed cycles of L_i containing x. Then either $x \in P_l^+$ and $x \in P_m^+$, or $x \in P_l^-$ and $x \in P_m^-$.*

Subproof. Suppose for a contradiction that $x \in P_l^+$ and $x \notin P_m^+$. Let $C_{l'}$ and $C_{m'}$ be the fathers of C_l and C_m respectively (they can be the same directed cycle). By definition of the L_j's, there exists a dipath P from $t_{l,l'}$ to $s_{m,m'}$ only going through $C_{l'}$, $C_{s'}$ and their ancestors. In particular P is disjoint from $C_l - C_{l'}$ and $C_s - C_{s'}$. Observe that $C_l[s_{l,l'}, t_{l,m}]$ has length at most $3k$ because it is contained in the union of $P_{l,l'}$, $P_{l,m}$, and $C_l[t_{l,l'}, x]$ which has length at most k because $x \in P_l^+$. Hence $C_l[t_{l,m}, s_{l,l'}]$ has length at least k. Moreover $C_m[s_{m,m'}, t_{l,m}]$ contains $C_m[t_{m,m'}, x]$ which has length at least k because $x \notin P_m^+$. Thus the union of $C_l[t_{l,m}, s_{l,l'}] \odot P$, $C_m[t_{l,m}, s_{m,m'}]$, and $C_m[s_{m,m'}, t_{l,m}]$ is a $B(k, 1; k)$-subdivision, a contradiction. The case where $x \in P_l^-$ and $x \notin P_m^-$ is symmetrical and the case where x does not belong to $P_l^- \cup P_l^+ \cup P_m^- \cup P_m^+$ is identical. $\Diamond$

Claim 80.2 implies that each level L_i may be partitioned into sets X_i^+, X_i^- and X_i', where X_i^+ (resp. X_i^-) is the set of vertices x of L_i such that every $x \in R_l^+$ (resp. $x \in R_l^-$) for every directed cycle C_l of L_i containing x and X_i' is set of vertices in L_i but not in $X_i^+ \cup X_i^-$. Set $X^+ = V(C_0) \cup \bigcup_{i \geq 1} X_i^+$, $X^- = \bigcup_{i \geq 1} X_i^-$ and $X' = \bigcup_{i \geq 1} X_i'$. Clearly (X^+, X^-, X') is a partition of $V(D[\mathcal{S}])$.

Claim 80.3. $\chi(D_2) \leq 4k^2 + 2$.

Subproof. Since $X^+ \cup X^- \cup X' = V(D_2)$, we have $\chi(D_2) \leq \chi(D_2[X^+ \cup X']) + \chi(D_2[X^- \cup X'])$. We shall prove that $\chi(D_2[X^+ \cup X']) \leq 2k^2 + 1$ and $\chi(D_2[X^- \cup X']) \leq 2k^2 + 1$, which imply the result.

Let x and y be two adjacent vertices of $D_2[X^+ \cup X']$. Let L_i be the level of x and L_j be the level of y. Without loss of generality, we may assume that $j \geq i + k$. Let C_x be the directed cycle of L_i such that $x \in C_x$ and C_y the directed cycle of L_j such that $y \in C_y$. By considering ancestors of C_x and C_y, there is a shortest sequence of directed cycles $C_1, \dots, C_p$ such that $C_1 = C_x$ and $C_p = C_y$ and for all $l \in [p-1]$, either C_l is the father of C_{l+1} or C_{l+1} is the father of C_l. In particular C_{p-1} is the father of C_p. Since $y \in X^+ \cup X'$, then $C[y, t_{p-1,p}]$ has length at least k.

Assume that (x, y) is an arc. In $\bigcup_{l=1}^{p-1} C_l$, there is a dipath P from $t_{p-1,p}$ to x. This dipath has length at least $k-1$ because it must go through all levels $L_{i'}$, $i \leq i' \leq j-1$ because the vertices of any directed cycle of $\mathcal{S}$ are in two consecutive levels. Hence the union of $P \odot (x, y)$, $C_p[t_{p-1,p}, y]$, and $C_p[y, t_{p-1,p}]$ is a $B(k, 1; k)$-subdivision, a contradiction. Hence (y, x) is an arc.

Suppose that C_x is not an ancestor of C_y. In particular, C_2 is the father of C_1 and there exists a path P from $t_{1,2}$ to y in $\bigcup_{l=2}^{p-1} C_l$ of length at least $k-1$ and internally disjoint from C_1. Hence the union of $P \odot yx$, $C_1[x, t_{1,2}]$ and $C_1[t_{1,2}, x]$ is a subdivsion of $B(k, 1; k)$. Hence C_x is an ancestor of C_y.

In particular, C_l is the father of C_{l+1} for all $l \in [p-1]$. Let P be the dipath from $t_{1,2}$ to y in $\bigcup_{l=2}^{p} C_l$. It has length at least $k-1$ because it must go through all levels L_i, $1 \leq i \leq p-1$. $C_1[x, t_{1,2}]$ has length less than k, for otherwise the union of $P \odot yx$, $C_1[x, t_{1,2}]$ and $C_1[t_{1,2}, x]$ would be a subdivision of $B(k, 1; k)$.

To summarize, the only arcs of $D_2[X^+ \cup X']$ are arcs (y, x) such that C_x is an ancestor of C_y and $C_1[x, t_{1,2}]$ has length less than k with $C_1 \dots C_p$ the sequence of directed cycles such that $C_1 = C_x$ to $C_p = C_y$ and C_l is the father of C_{l+1} for all $l \in [p-1]$. In particular, $D_2[X^+ \cup X']$ is acyclic.

Let y be a vertex of $D_2[X^+ \cup X']$. Let L_p be the level of y and let $C_0, \dots, C_p$ be the sequence of directed cycles such that C_{l-1} is the father of C_l for all $l \in [p]$. For $0 \leq l \leq p-1$, let R_l be the subdipath of C_l of length $k-1$ terminating at $t_{l,l+1}$. By the above property, the outneighbbours of y are in $\bigcup_{l=0}^{p-1} R_l$. Suppose for a contradiction that y has outdegree at least $2k^2 + 1$. Then there are $2k+1$ distinct indices $l_1 < \cdots < l_{2k+1}$ such that for all $i \in [2k+1]$, C_{l_i} contains an outneighbour X_i of y. Let P be the shortest dipath from x_1 to y in $\bigcup_{l=l_1}^{p} C_l$. This dipath intersects all directed cycles C_l $l_1 \leq l \leq p$. Let z be the first vertex of P along $C_{l_{k+1}}[x_{k+1}, t_{l_{k+1}, l_{k+2}}]$. Vertex z belongs to either $L_{l_{k+1}-1}$ or $L_{l_{k+1}}$. Thus $P[x_1, z]$ and $P[z, y]$ have length at least $k-1$ and k respectively since P goes through all levels from L_{l_1} to L_p. Hence the union of $(y, x_1) \odot P[x_1, z]$, $(y, x_{k+1}) \odot C_{l_{k+1}}[x_{k+1}, z]$, and $P[z, y]$ is a $B(k, 1; k)$-subdivision, a contradiction. Therefore $D_2[X^+ \cup X']$ has maximum outdegree at most $2k^2$.

$D_2[X^+ \cup X']$ is acyclic and has maximum outdegree at most $2k^2$. Therefore it is $2k^2$-degenerate, and so $\chi(D_2[X^+ \cup X']) \leq 2k^2 + 1$. By symmetry, we have $\chi(D_2[X^- \cup X']) \leq 2k^2 + 1$. $\Diamond$

To bound $\chi(D_0)$ we partition the vertex set according to a colouring ϕ of $\bigcup \mathcal{S}$ given by Lemma 73. For every colour $c \in [\alpha_k]$, let $X^+(c)$ be the set $X^+ \cap \phi^{-1}(c)$ of vertices of X^+ coloured c, and $X^-(c)$ the set $X^- \cap \phi^{-1}(c)$ of vertices of X^- coloured c. Similarly, let $X_i^+(c) = X_i^+ \cap \phi^{-1}(c)$ and $X_i^-(c) = X_i^- \cap \phi^{-1}(c)$. We denote by $D_0^+(c)$ (resp. $D_0^-(c)$, $D_0'(c)$) the subdigraph of D_0 induced by the vertices of $X^+(c)$, (resp. $X^-(c)$, $X'(c)$).

Claim 80.4. $\chi(D'_0(c)) = 1$ *for every* $c \in [\alpha_k]$.

Subproof. We need to prove that $D'_0(c)$ has no arc. Suppose for a contradiction that (x, y) is an arc of $D'_0(c)$. By definition of D_0, the vertices x and y are in a same level L_i. Let C_l and C_m be two directed cycles of L_i such that $x \in C_l$ and $y \in C_m$.

If $C_l = C_m$, then both $C_l[x, y]$ and $C_l[y, x]$ have length at least $7k$ because the subdipaths of length $7k$ of C_l are rainbow-coloured by ϕ. Hence the union of those paths and (x, y) is a $B(k, 1; k)$-subdivision, a contradiction. Henceforth, C_l and C_m are distinct directed cycles.

Suppose first that C_l and C_m intersect. By Claim 80.2, $s_{l,m}$ belongs to P_l^-, P_l^+ or L_{i-1}, and by construction of R'_l, $C_l[x, s_{l,m}]$ and $C_l[s_{l,m}, x]$ are both longer than k. Therefore they form with $(x, y) \odot C_m[y, s_{l,m}]$ a $B(k, 1; k)$-subdivision, a contradiction.

Suppose now that C_l and C_m do not intersect. Let C'_l and C'_m be the fathers of C_l and C_m respectively. Let P be the dipath from $s_{m,m'}$ to $s_{l,l'}$ in $\bigcup_{j<i} L_j$. Then the union of $C_l[s_{l,l'}, x]$, $(x, y) \odot C_m[y, s_{m,m'}] \odot P$, and $C_l[x, s_{l,l'}]$ is a $B(k, 1; k)$-subdivision, a contradiction. $\Diamond$

Claim 80.5. $\chi(D_0^+(c)) \leq (4k)^{4k}$ *for every* $c \in [\alpha_k]$.

Subproof. Set $p = (4k)^{4k}$. Suppose for a contradiction that there exists c such that $\chi(D_0^+(c)) > p$. Observe that $D_0^+(c)$ is the disjoint union of the $D[X_i^+(c)]$. Thus there exists a level L_{i_0} such that $\chi(D[X_i^+(c)]) > p$. Moreover $i_0 > 0$, because the vertices of C_0 coloured c form a stable set. By Theorem 26, there exists a dipath $P = (v_0, \dots, v_p)$ of length p in $D[X_i^+(c)]$.

Suppose that P contains two vertices x and y of a same directed cycle C of $\mathcal{S}$. Without loss of generality, we may assume that $P]x, y[$ contains no vertices of C. Now both $C[x, y]$ and $C[y, x]$ have length at least $7k$ because the subdipaths of length $7k$ of C are rainbow-coloured by ϕ. Thus the union of $C[x, y]$, $P[x, y]$ and $C[y, x]$ is a $B(k, 1; k)$-subdivision, a contradiction. Hence P intersects every directed cycle of $\mathcal{S}$ at most once.

For every $v \in V(P)$, let $\mathrm{Len}(v)$ be the set of lengths of $C_l[t_{l,l'}, v]$ for all directed cycles $C_l \in L_{i_0}$ containing v and whose father is $C_{l'}$.

For $1 \leq i \leq p$, let $m_i = \min \mathrm{Len}(v_i)$. By Claim 80.2, $\mathrm{Len}(v_i) \subseteq [2k]$. Lemma 67 applied to $(m_i)_{1 \leq i \leq p}$ yields a set L of $4k^2$ indices such that for any $i < j \in L$, $m_i = m_j$ and $m_k > m_i$, for all $i < k < j$. Let $l_1 < l_2 < \cdots < l_{4k^2}$ be the elements of L and let $m = m_{l_1} = \cdots = m_{l_{4k^2}}$.

For $1 \leq j \leq 4k^2 - 1$, let $M_j = \max \bigcup_{l_j \leq i < l_{j+1}} \mathrm{Len}(v_i)$. By definition $M_j \leq 2k$. Applying Lemma 68 to $(M_j)_{1 \leq j \leq 4k^2}$, we get a sequence of size $2k$ $M_{j_0+1}, \dots, M_{j_0+2k}$ such that M_{j_0+2k} is the greatest. For sake of simplicity, we set $\ell_i = j_0 + i$ for $1 \leq i \leq 2k$. Let f be the smallest index not smaller than ℓ_{2k} for which $M_{\ell_{2k}} \in \mathrm{Len}(v_f)$.

Let j_1 and j'_1 be indices such that $v_{\ell_1} \in C_{j_1}$, C_{j_1} is in L_{i_0}, $C_{j'_1}$ is the father of C_{j_1} and $C_{j_1}[t_{j'_1,j_1}, v_{\ell_1}]$ has length m. Set $P_1 = C_{j_1}[t_{j'_1,j_1}, v_{\ell_1}]$. Let j_2 and j'_2 be indices such that $v_{\ell_k} \in C_{j_2}$, C_{j_2} is in L_{i_0}, $C_{j'_2}$ is the father of C_{j_2} and $C_{j_2}[t_{j'_2,j_2}, v_{\ell_k}]$ has length m. Set $P_2 = C_{j_2}[t_{j'_2,j_2}, v_{\ell_k}]$. Let j_3 and j'_3 be indices such that $v_f \in C_{j_3}$, C_{j_3} is in L_i, $C_{j'_3}$ is the father of C_{j_3} and $C_{j_3}[t_{j'_3,j_3}, v_f]$ has length $M_{\ell_{2k}}$. Set $P_3 = C_{j_3}[v_f, s_{j'_3,j_3}]$. Note that any internal vertex x of P_1 or P_2 has an integer in $\mathrm{Len}(x)$ which is smaller than m and every internal vertex y of P_3 either has an integer in $\mathrm{Len}(y)$ which is greater than $M_{\ell_{2k}}$, or does not belong to $X^+(c)$. Hence, P_1, P_2 and P_3 are disjoint from $P[v_{\ell_1}, v_f]$.

We distinguish cases according to the intersection between P_1, P_2 and P_3: Let P_4 be a shortest dipath in $\cup_{i<i_0} L_i$ from $t_{j'_1,j_1}$ to $t_{j'_2,j_2}$ and P_5 be a shortest dipath in $\cup_{i<i_0} L_i$ from $s_{j'_3,j_3}$ to $t_{j'_2,j_2}$

- Suppose P_3 does not intersect $P_1 \cup P_2$.
  - Suppose P_1 and P_2 are disjoint. Let v be the last vertex of P_4 in $P_4 \cap P_5$. The union of $P_5[v, t_{j'_1,j_1}] \odot P_1 \odot P[v_{\ell_1}, v_{\ell_k}]$, $P_4[v, t_{j'_2,j_2}] \odot P_2$, and $P[v_{\ell_k}, v_f] \odot P_3 \odot P_5[s_{j'_3,j_3}, v]$ is a $B(k, 1; k)$-subdivision, a contradiction.
  - Assume now P_1 and P_2 intersect. Let u be the last vertex along P_2 on which they intersect. The union of $P_1[u, v_{\ell_1}] \odot P[v_{\ell_1}, v_{\ell_k}]$, $P_2[u, v_{\ell_k}]$, and $P[v_{\ell_k}, v_f] \odot P_3 \odot P_5 \odot P_1[t_{j'_1,j_1}, u]$ is a $B(k, 1; k)$-subdivision, a contradiction.
- Assume P_3 intersects $P_1 \cap P_2$. Let v be the first vertex along P_3 in $P_1 \cap P_2$ and let u be the last vertex of $P_1 \cap P_2$ along P_2. The union of $P_1[u, v_{\ell_1}] \odot P[v_{\ell_1}, v_{\ell_k}]$, $P_2[u, v_{\ell_k}]$, and $P[v_{\ell_k}, v_f] \odot P_3[v_f, v] \odot P_1[v, u]$ is a $B(k, 1; k)$-subdivision, a contradiction.
- Assume now that P_3 intersects $P_1 \cup P_2$ but not $P_1 \cap P_2$. Let v be the first vertex along P_3 in $P_1 \cup P_2$.
  - If $v \in P_2$, let u be the last vertex of $P_2 \cap P_3$ along P_3. Observe that $P_3[v, u]$ is also a subdipath of P_2 and therefore contains no vertex of P_1. Hence, the union of $P_3[u, s_{j'_3,j_3}] \odot P_5 \odot P_1 \odot P[v_{\ell_1}, v_{\ell_k}]$, $P_2[u, v_{\ell_k}]$, and $P[v_{\ell_k}, v_f] \odot P_3[v_f, v]$ is a $B(k, 1; k)$-subdivision, a contradiction.
  - If $v \in P_1$, let u be the last vertex of $P_1 \cap P_3$ along P_3. Observe that $P_3[v, u]$ is also a subdipath of P_1 and therefore contains no vertex of P_2. Hence the union of $P_1[u, v_{\ell_1}] \odot P[v_{\ell_1}, v_{\ell_k}]$, $P_3[u, s_{j'_3,j_3}] \odot P_6 \odot P_2$, and $P[v_{\ell_k}, v_f] \odot P_3[v_f, u]$, is a $B(k, 1; k)$-subdivision, a contradiction.

$\Diamond$

Similarly to Claim 80.5, one proves that $\chi(D_0^-(c)) \leq (4k)^{4k}$ for every $c \in [\alpha_k]$. Hence, $\chi(D_0(c)) \leq \chi(D_0^+(c)) + \chi(D_0^-(c)) + \chi(D'_0(c)) \leq 2 \cdot (4k)^{4k} + 1$. Thus

$$\chi(D_0) \leq (2 \cdot (4k)^{4k} + 1)\alpha_k.$$

Via Lemma 59, this equation and Claims 80.1 and 80.3 yield

$$\chi(D) \leq \chi(D_0) \times \chi(D_1) \times \chi(D_2) \leq k(4k^2 + 2)(2 \cdot (4k)^{4k} + 1)\alpha_k = \beta_k.$$

$\square$

Proof of Theorem 37

Consider a maximal k-suitable collection $\mathcal{C}$ of directed cycles in D. Recall that $D_{\mathcal{C}}$ is the digraph obtained by contracting every component of $\mathcal{C}$ into one vertex. For each connected component $\mathcal{S}_i$ of $\mathcal{C}$, we denote by s_i the new vertex created.

Claim 80.6. *$\chi(D_{\mathcal{C}}) \leq 8k$.*

Proof. First note that since D is strong so is $D_{\mathcal{C}}$.

Suppose for a contradiction that $\chi(D_{\mathcal{C}}) > 8k$. By Theorem 31, there exists a directed cycle $C = (x_1, x_2, \dots, x_l, x_1)$ of length at least $8k$. For each vertex x_j that corresponds to an s_i in D, the arc (x_{j-1}, x_j) corresponds in D to an arc whose head is a vertex p_i of $\mathcal{S}_i$ and the arc (x_j, x_{j+1}) corresponds to an arc whose tail is a vertex l_i of $\mathcal{S}_i$. Let P_j be the dipath from p_i to l_i in $\bigcup \mathcal{C}$. Note that this dipath intersects the elements of $\mathcal{S}_i$ only along a subdipath. Let C' be the directed cycle obtained from C where we replace all contracted vertices x_j by the dipath P_j. First note that C' has length at least $8k$. Moreover, a directed cycle of $\mathcal{C}$ can intersect C' only along one P_j, because they all correspond to different strong components of $\bigcup \mathcal{C}$. Thus C' intersects each directed cycle of $\mathcal{C}$ on a subdipath. Moreover this subdipath has length less than k for otherwise D would contain a $B(k, 1; k)$-subdivision. So C' is a directed cycle of length at least $8k$ which intersects every directed cycle of $\mathcal{C}$ along a subdipath of length less than k. This contradicts the maximality of $\mathcal{C}$. $\square$

Using Lemma 65 with Claim 80.6 and Lemma 80, we get that $\chi(D) \leq 8k \cdot \beta_k$. This proves Theorem 37 for $\gamma_k = 8k \cdot \beta_k = 8k^2(4k^2 + 2)(2 \cdot (4k)^{4k} + 1)(2 \cdot (6k^2)^{3k} + 14k)$.

2.4 Subdivisions in digraphs with large dichromatic number

Recall that a *k-dicolouring* is a k-partition $\{V_1, \dots, V_k\}$ of $V(D)$ such that $D[V_i]$ is acyclic for every $i \in [k]$, and that the *dichromatic number* of D, noted $\vec{\chi}(D)$, is the minimum k such that D admits a k-dicolouring.

We present here a proof of the following Theorem, obtained in [1].

Theorem 40. *Let F be a digraph on n vertices and m arcs. Every digraph D with $\vec{\chi}(D) > 4^m(n-1) + 1$ contains a subdivision of F.*

The proof uses induction on m with the following lemma providing the induction step.

Lemma 81. *Let F be a digraph and let $a = (x, y)$ be an arc in $A(F)$. Suppose there exists a constant c such that any digraph D with $\vec{\chi}(D) \geq c$ contains a subdivision of $F - a$. Then any digraph D with $\vec{\chi}(D) \geq 4c - 3$ contains a subdivision of F.*

Proof. Let D be a digraph with $\vec{\chi}(D) \geq 4c - 3$. We shall prove that D contains a subdivision of F.

By Lemma 39, we may assume that D is strong. Let u be a vertex in D and T_u an out-BFS-tree with root u. By Lemma 58, there is a level L^u such that $\vec{\chi}(D[L^u]) \geq 2c - 1$. By Lemma 39, there is a strong component C of $D[L^u]$ such that $\vec{\chi}(C) = \vec{\chi}(D[L^u]) \geq 2c-1$. Since D is strong, there is a shortest (v, u)-dipath P in D such that $V(P) \cap V(C) = \{v\}$. Let T_v be an in-BFS-tree in C rooted at v. By Lemma 58, there is a level L^v of T_v such that $\vec{\chi}(D[L^v]) \geq c$. By definition of c, $D[L^v]$ contains a subdivision S of $F - a$. Let x' and y' be the vertices in S corresponding to the vertices x and y of F. Now $T_v[x', v] \cup P \cup T_u[u, y']$ is a directed (x', y')-walk with no internal vertex in L^v. Hence it contains an (x', y')-dipath Q whose internal vertices are not in S. Therefore, $S \cup Q$ is a subdivision of F in D. □

We can now prove Theorem 40

Proof of Theorem 40. We prove the result by induction on m. If $m = 0$, then F is an empty digraph and the result is trivial. If $m > 0$, then consider an arc $a \in A(F)$. By the induction hypothesis on $F - a$, we obtain that any digraph D with $\vec{\chi}(D) \geq 4^{m-1}(n-1)+1$ contains a subdivision of $F - a$. By Lemma 81, we obtain that any digraph with $\vec{\chi}(D) \geq 4(4^{m-1}(n-1)+1) - 3$ contains a subdivision of F, which proves the result. □

Chapter 3

Digraphs with large minimum outdegree

In 1985, Mader made the following beautiful conjecture:

Conjecture 9 (Mader [49])**.** For every $k \geq 1$, there exists an integer $f(k)$ such that every digraph with minimum outdegree at least $f(k)$ contains a subdivision of TT_k, the transitive tournament on k vertices.

The conjecture remains open for $k \geq 5$ which shows our lack of understanding of what it means for a digraph to have large minimum degree. In this chapter we present some work on weakenings of the conjecture. First we present some results obtained in [1] with Pierre Aboulker, Nathann Cohen, Frédéric Havet, Phablo Moura and Stéphan Thomassé concerning subdivisions of digraphs which are simpler than the transitive tournaments. Then we present the proof of Conjecture 15 concerning immersions of transitive tournaments obtained in [45].

3.1 Subdivisions in digraphs with large minimum outdegree

3.1.1 Subdivisions of oriented paths

The first case to consider when one tries to weaken Mader's conjecture is probably the case of oriented paths.

Theorem 19. *Let $(k_1, k_2, \dots, k_\ell)$ be a sequence of positive integers, and let D be a digraph with $\delta^+(D) \geq \sum_{i=1}^{\ell} k_i$. For every $v \in V(D)$, D contains a path $P^+(k'_1, k'_2, \dots, k'_\ell)$ with initial vertex v such that $k'_i \geq k_i$ if i is odd, and $k'_i = k_i$ otherwise.*

Proof. By induction on ℓ. If $\ell = 1$, then the result holds trivially. Assume now $\ell \geq 2$, and suppose that, for every path $P^+(x_1, x_2, \dots, x_t)$ with $t < \ell$ and every digraph G with $\delta^+(G) \geq \sum_{i=1}^{t} x_i$, G contains a path $P^+(x'_1, x'_2, \dots, x'_t)$ starting at any vertex of G such that $x'_i \geq x_i$ if i is odd, and $x'_i = x_i$ otherwise.

Let v be a vertex of D. Since $\delta^+(D) \geq \sum_{i=1}^{\ell} k_i$, there exists a (v, u)-dipath $P_{v,u}$ in D of length exactly k_1, for some vertex $u \in V(D)$. Let $D' = D - (P_{v,u} - u)$, let C be the connected component of D' containing u, and let H be a *sink* strong component of C (i.e. a strong component without arcs leaving it) that is reachable by a directed path in C starting at u. We denote by $P_{u,x}$ a (u, x)-dipath in C such that $V(P_{u,x}) \cap V(H) = \{x\}$.

Note that no vertex of H dominates a vertex in $V(P_{u,x}) \setminus \{x\}$ since H is a sink strong component. Thus, $\delta^+(H) \geq \delta^+(D) - k_1 \geq k_2$. As a consequence, H contains a directed cycle of length at least k_2. Using this and the fact that H is strongly connected, we conclude that there exists a dipath $P_{y,x}$ in H from a vertex $y \in V(H) \setminus \{x\}$ to x of length exactly k_2. Let $G = H - (P_{y,x} - y)$. One may easily verify that $\delta^+(G)$ is at least $\delta^+(D) - k_1 - k_2 \geq \sum_{i=3}^{\ell} k_i$.

Let $Q_{v,y} = P_{v,u}P_{u,x}P_{y,x}$. Note that $Q_{v,y}$ is a path $P^+(k_1', k_2)$ starting at v with $k_1' \geq k_1$. Therefore, the result follows immediately if $\ell = 2$. Suppose now that $\ell \geq 3$. By the induction hypothesis, G contains a path $W_y := P^+(k_3', \ldots, k_\ell')$ with initial vertex y such that $k_i' \geq k_i$ if i is odd, and $k_i' = k_i$ otherwise. Therefore, $Q_{v,y}W_y$ is the desired path $P^+(k_1', k_2', \ldots, k_\ell')$ with initial vertex v. $\square$

The result is tight, as $\sum_{i=1}^{\ell} k_i = |V(P^+(k_1, k_2, \ldots, k_\ell))| - 1$, and the complete digraph $\vec{K}_k$ on k vertices has minimum outdegree $k - 1$ and contains no path on more than k vertices.

3.1.2 Subdivisions of in-arborescences

The aim of this subsection is to prove that in-arborescences are contained in digraphs with large minimum outdegree as subdivision. We need the following result about flows.

Lemma 82. *Let D be a digraph with $|A(D)| \geq 1$, let $S \subseteq V(D)$ be a nonempty subset of vertices of indegree 0 in D, and let $T \subseteq V(D)$ such that $T \cap S = \emptyset$. If $d^+(v) \geq \Delta^-(D)$ for all $v \in V(D) \setminus T$, then there exist $|S|$ vertex-disjoint (S, T)-dipaths in D.*

Proof. Suppose to the contrary that there do not exist $|S|$ vertex-disjoint (S, T)-dipaths in D. By Theorem 2, there exists a set of vertices $X \subseteq V(D)$ with cardinality $|X| < |S|$ which is an (S, T)-vertex-cut. Let C be the set of vertices in $D - X$ that are reachable in D by a dipath with initial vertex in $S \setminus X$. Set $k = |X \cap S|$. Observe that $k < |S|$.

Let us count the number $a(C, X)$ of arcs with tail in C and head in X. Since the vertices in S have indegree 0 and every vertex in C has outdegree at least $\Delta^-(D)$,

$$a(C, X) \geq |C| \cdot \Delta^-(D) - [|C| - (|S| - k)] \cdot \Delta^-(D) = (|S| - k) \cdot \Delta^-(D).$$

Moreover, $a(C, X)$ is at most the number of arcs with head in X which is at most $(|X| - k) \cdot \Delta^-(D)$, because the vertices in $S \cap X$ have indegree 0. Hence $(|S| - k) \cdot \Delta^-(D) \leq a(C, X) \leq (|X| - k) \cdot \Delta^-(D)$. This is a contradiction to $|X| < |S|$. $\square$

Let k and ℓ be positive integers. The ℓ-in-arborescence of depth k is denoted by $B(k, \ell)$. The number of vertices of $B(k, \ell)$ is denoted by $b(k, \ell)$; so $b(k, \ell) = \sum_{i=0}^{k} \ell^i = \frac{\ell^{k+1} - 1}{\ell - 1}$.

Observe that every in-arborescence T is a subdigraph of $B(k,\ell)$ with $\ell = \Delta^-(T)$ and k the maximum length of a dipath in T. Therefore proving the result for $B(k,\ell)$ for all k and ℓ implies the result for in-arborescences.

We define a recursive function $f\colon \mathbb{N} \to \mathbb{N}$ as follows. For all positive integers k and ℓ such that $\ell \geq 2$, $f(1,\ell) = \ell$ and, for $k \geq 2$, we define $f(k,\ell) = t(k,\ell)\cdot(\ell-1)\cdot k + t(k,\ell)$, where $t(k,\ell) := f(k-1,\, b(k-1,\ell)\cdot(\ell-1)+1)\cdot b(k-1,\ell)$.

If $\mathcal{D}$ is a family of digraphs, a *packing* of elements of $\mathcal{D}$ is the disjoint union of copies of elements of $\mathcal{D}$.

Theorem 83. *Let $k \geq 1$ and $\ell \geq 2$ be integers, and let D be a digraph with $\delta^+(D) \geq f(k,\ell)$. Then D contains a subdivision of $B(k,\ell)$, the ℓ-in-arborescence of depth k.*

Proof. We prove the result by induction on k and ℓ. If $k = 1$, then $\delta^+(D) \geq \ell$. Thus, $\Delta^-(D)\cdot|V(D)| \geq \sum_{v\in V(D)} d^-(v) = \sum_{v\in V(D)} d^+(v) \geq \ell\cdot|V(D)|$. Hence there is a vertex with indegree at least ℓ in D and, consequently, the result follows when $k = 1$. Assume now $k \geq 2$, and suppose that, for every positive integers $k' < k$ and ℓ', and every digraph H with $\delta^+(H) \geq f(k',\ell')$, H contains a subdivision of the ℓ'-branching in-arborescence of depth k'.

Let $\mathcal{F}$ be a packing of ℓ-branching in-arborescences subdigraphs of any non-zero depth in D that covers the maximum number of vertices.

We denote by $U \subseteq V(D)$ the set of vertices not covered by $\mathcal{F}$, that is, $U = V(D) \setminus \bigcup_{A\in\mathcal{F}} V(A)$. Let r_A denote the root of the in-arborescence A, for each $A \in \mathcal{F}$, and let $R = \{r_A \in V(D)\colon A \in \mathcal{F}\}$ be the set of the roots of the arborescences in $\mathcal{F}$.

We now construct the digraph H with vertex set R such that there exists an arc (r_A, r_B) in H if and only if r_A dominates some vertex of $V(B)$ in D.

Claim 83.1. *If $\delta^+(H) \geq t(k,\ell)/b(k-1,\ell)$, then D contains a subdivision of $B(k,\ell)$.*

Subproof. Let $p = b(k-1,\ell)\cdot(\ell-1)+1$. By the induction hypothesis, H contains a subdivision T of $B(k-1,p)$. Let R' be the set of branching vertices of T, that is, $R' = \{r \in T\colon d^-_T(r) = p\}$. We assume that each in-arborescence of $\mathcal{F}$ has at most $b(k-1,\ell)$ vertices, as any larger arborescence would yield the theorem. Thus, for each $r \in R'$, there exists a vertex h_r in the in-arborescence rooted at r such that h_r is dominated in D by ℓ vertices of $V(T)$. Similarly, for each $r \in V(T)$ with indegree 1, there exists a vertex h_r in the in-arborescence rooted at r such that h_r is dominated in D by a vertex of $V(T)$. Using these remarks, we next define a procedure to obtain a subdigraph of T that is a subdivision of $B(k-1,\ell)$.

For each $r \in R'$, we remove from T all arcs with head r but exactly ℓ arcs from vertices in $V(T)$ that dominate h_r in D. We denote by T' the component of the subdigraph of T obtained by applying the above-described procedure and that contains the root of T. One may easily verify that T' is a subdivision of $B(k-1,\ell)$. Let P_r be the path from h_r to r in the in-arborescence corresponding to r, for every $r \in V(T')$ such that either $r \in R'$ or r has indegree 1.

Let Q be the in-arborescence obtained from T' in the following way. For each $r \in V(T')$ such that either $r \in R'$ or r has indegree 1, we add h_r to T', and we add an arc from every inneighbour of r in T' to h_r. Additionally, we remove all arcs with head r in T', and link h_r to r by using the dipath P_r. Finally, for each $r \in V(T')$ that is a leaf, we replace r by its corresponding in-arborescence belonging to $\mathcal{F}$.

By this construction, we have that Q is a subdigraph of D such that every internal vertex has either indegree ℓ or 1. Furthermore, it has depth at least k. Therefore, by possibly pruning some levels of Q, we obtain a subdivision of $B(k, \ell)$. $\Diamond$

Suppose now $\delta^+(H) < t(k,\ell)/b(k-1,\ell)$. Observe that, for every $v \in R$ such that $d^+_H(v) < t(k,\ell)/b(k-1,\ell)$, we have, in the digraph D, that $d^+_U(v) \geq t(k,\ell) \cdot (\ell-1) \cdot k$ since $\delta^+(D) \geq t(k,\ell) \cdot (\ell-1) \cdot k + t(k,\ell)$. We define $X = \{v \in R\colon d^+_U(v) \geq t(k,\ell) \cdot (\ell-1) \cdot k\}$.

Let D' be the digraph obtained from $D[U \cup X]$ by removing all arcs with head in X. From D', we construct a digraph G by replacing every vertex $v \in X$ by $t(k,\ell)$ new vertices $v_1, \ldots, v_{t(k,\ell)}$, and adding, for each $i \in [t(k,\ell)]$, at least $(\ell-1) \cdot k$ arcs from v_i to $N^+_{D'}(v)$ in such a way that $d^-_{D'}(u) = d^-_G(u)$, for all $u \in N^+_{D'}(v)$. In other words, we "redistribute" the outneighbours of v in D' among its $t(k,\ell)$ copies in G so that every copy has outdegree at least $(\ell-1) \cdot k$, and the indegrees of vertices belonging to U are not changed. Let $S \subseteq V(G)$ be the set of vertices that replaced those of X, that is, $S = \bigcup_{v \in X} \{v_1, \ldots, v_{t(k,\ell)}\}$. Let T be the set of vertices in U that have large outdegree outside U in the digraph D, more formally, $T = \{v \in U\colon d^+_{V(D)\setminus U}(v) \geq t(k,\ell) + 1\}$.

For every $i \in [k-1]$, let $\mathcal{F}_i = \{A \in \mathcal{F}\colon A \text{ has depth exactly } i\}$. Note that $\{\mathcal{F}_i\}_{i \in [k-1]}$ forms a partition of the packing $\mathcal{F}$. Additionally, observe that, due to the maximality of $\mathcal{F}$, every vertex in U is dominated by at most $\ell-1$ vertices belonging to U, and by at most $\ell-1$ roots of in-arborescences in $\mathcal{F}_i$, for each $i \in [k-1]$. Thus, the indegree in G of every vertex belonging to U is at most $(\ell-1) + (\ell-1) \cdot (k-1) = (\ell-1) \cdot k$. Therefore, we have $\Delta^-(G) \leq (\ell-1) \cdot k$. Moreover, since $\delta^+(D) \geq t(k,\ell) \cdot (\ell-1) \cdot k + t(k,\ell)$, we have $d^+_G(v) \geq t(k,\ell) \cdot (\ell-1) \cdot k$ for every $v \in U \setminus T$. Hence, $d^+_G(v) \geq (\ell-1) \cdot k$, for every $v \in V(G) \setminus T$. By Lemma 82, there exists a set $\mathcal{P}$ of $|S|$ vertex-disjoint paths from S to T in G.

Note that, in D, every vertex belonging to T has at least $t(k,\ell) + 1$ outneighbours in $V(D) \setminus U$. Therefore one can greedily extend each path of $\mathcal{P}$ with an outneighbour of its terminal vertex in $V(D) \setminus U$ in order to obtain a set $\mathcal{P}'$ of $|S|$ vertex-disjoint paths from S to $V(D) \setminus U$ such that for any $v \in X$ all the paths in $\mathcal{P}'$ with initial vertex v have distinct terminal vertices (and different from v).

We now construct the digraph M on the vertex set R where there exists an arc from $v = r_A$ to r_B in M whenever

either r_A dominates some vertex of $V(B)$ in D,

or there is a dipath from some v_i to $V(B)$ in $\mathcal{P}'$.

Since, for each $v \in X$, all vertices in $\{v_i\}_{i \in [t(k,\ell)]}$ are the initial vertices of vertex-disjoint dipaths in $\mathcal{P}'$, we obtain $\delta^+(M) \geq t(k,\ell)/b(k-1,\ell)$. Therefore, the result follows by Claim 83.1 with M playing the role of H. $\square$

3.1.3 Cycles with two blocks

Recall that $C(k_1, k_2)$ is the oriented cycle with two blocks, one of length k_1 and one of length k_2. It can also be seen as the union of two internally disjoint dipaths, one of length k_1 and one of length k_2 with the same initial vertex and same terminal vertex.

Theorem 84. *Let D be a digraph with $\delta^+(D) \geq 2(k_1+k_2)-1$. Then D contains a subdivision of $C(k_1, k_2)$ as a subdigraph.*

Proof. Let us assume, without loss of generality, that $k_1 \geq k_2$. Let ℓ be a positive integer. An (ℓ, k_1, k_2)-*fork* is a digraph obtained from the union of three disjoint dipaths $A = (a_0, a_1, \cdots, a_\ell)$, $B^1 = (b^1_1, \cdots, b^1_{k_1-1})$ and $B^2 = (b^2_1, \cdots, b^2_{k_2-1})$ by adding the arcs (a_ℓ, b^1_1) and (a_ℓ, b^2_1).

Since a $(1, k_1, k_2)$-fork has $k_1 + k_2$ vertices and $\delta^+(D) \geq k_1 + k_2 + 1$, then D contains a $(1, k_1, k_2)$-fork as a subdigraph by a greedy argument. Let $\ell \geq 1$ be the largest integer such that D contains an (ℓ, k_1, k_2)-fork as a subdigraph. Let F be such a fork. For convenience, we denote its subpaths and vertices by their labels in the above definition.

If there exist $i, j \in [\ell - 1] \cup \{0\}$, where $i \leq j$ (resp. $j \leq i$), such that $a_i \in N^+(b^1_{k_1-1})$ and $a_j \in N^+(b^2_{k_2-1})$, then the union of the dipaths $(a_\ell, B^1, a_i, \cdots, a_j)$ and (a_ℓ, B^2, a_j) (resp. (a_ℓ, B^1, a_i) and $(a_\ell, B^2, a_j, \cdots, a_i)$) is a subdivision of $C(k_1, k_2)$.

Suppose now that $b^1_{k_1-1}$ has no outneighbour in $\{a_0, \cdots, a_{\ell-1}\}$, that is, $N^+(b^1_{k_1-1}) \cap (A \setminus \{a_\ell\}) = \emptyset$ (the case $N^+(b^2_{k_2-1}) \cap (A \setminus \{a_\ell\}) = \emptyset$ is similar). Since $|B^1 \cup B^2 \cup \{a_\ell\}| = k_1 + k_2 - 1$ and $\delta^+(D) \geq 2(k_1 + k_2 - 1) + 1$, $b^1_{k_1-1}$ has two distinct outneighbours, say c^1_1 and c^2_1, not in F.

Let $i_1 \geq 1$ be the largest integer such that there exist two disjoint dipaths C^1 and C^2 in $D - F$ with initial vertex c^1_1 and c^2_1, respectively, and length i_1 and $i_2 = \min\{k_2, i_1\}$. Set $C^1 = (c^1_1, \cdots, c^1_{i_1})$ and $C^2 = (c^2_1, \cdots, c^2_{i_2})$. By maximality of ℓ, if $i_1 \geq k_2$, then $i_1 < k_1 - 1$. Otherwise, the union of $A \cup B^1$, C^1, C^2, $(b^1_{k_1-1}, c^1_1)$ and $(b^1_{k_1-1}, c^2_1)$ would contain an $(\ell + k_1 - 1, k_1, k_2)$-fork, contradicting the maximality of ℓ.

Suppose to the contrary that none of $c^1_{i_1}$ and $c^2_{i_2}$ have outneighbour in $A \setminus \{a_\ell\}$. Since $|V(B^1) \cup V(B^2) \cup \{a_\ell\} \cup V(C^1) \cup V(C^2)| = k_1 + k_2 + i_1 + i_2 - 1 < 2(k_1 + k_2) - 2$ (since $i_1 + i_2 < k_1 + k_2 - 1$) and $\delta^+(D) \geq 2(k_1 + k_2 - 1) + 1$, then there exist $c^1_{i_1+1}, c^2_{i_2+1} \in V(D - (F \cup C^1 \cup C^2))$ such that $(c^1_{i_1}, c^1_{i_1+1}), (c^2_{i_2}, c^2_{i_2+1}) \in A(D)$ and $c^1_{i_1+1} \neq c^2_{i_2+1}$. This contradicts the maximality of i_1. Henceforth, we assume that $c^1_{i_1}$ has an outneighbour $a_j \in A \setminus \{a_\ell\}$ for some $0 \leq j < \ell$. The case in which $c^2_{i_2}$ has an outneighbour in $A \setminus \{a_\ell\}$ is similar.

If also $b^2_{k_2-1}$ has an outneighbour $a_m \in A \setminus \{a_\ell\}$, then the union of the dipaths $(a_\ell, B^1, C^1, a_j, \ldots, a_m)$ and (a_ℓ, B^2, a_m) (if $m \geq j$) or of the dipaths (a_ℓ, B^1, C^1, a_j) and $(a_\ell, B^2, a_m, \ldots, a_j)$ (if $m < j$) is a subdivision of $C(k_1, k_2)$.

If $b^2_{k_2-1}$ has an outneighbour $z \in V(C^1 \cup C^2)$, say $z = c^1_h$ for some $h \leq i_1$ (the case in which $z \in V(C^2)$ is similar), then the union of the dipaths (a_ℓ, B^2, c^1_h) and $(a_\ell, B, c^1_1, \cdots, c^1_h)$ is a subdivision of $C(k_1, k_2)$.

So, we may assume that $b^2_{k_2-1}$ has no outneighbour in $A \setminus \{a_\ell\} \cup C^1 \cup C^2$. Hence, $b^2_{k_2-1}$ has two distinct outneighbours, say c^3_1 and c^4_1, not in $F \cup C^1 \cup C^2$. Let $i_3 \geq 1$ be the largest integer such that there exist two disjoint dipaths C^3 and C^4 in $D - (F \cup C^1 \cup C^2)$ with initial vertex c_1^3 and c_1^4, respectively, and length i_3 and $i_4 = \min\{k_2, i_3\}$. By the maximality of ℓ, if $i_4 \geq k_2$ then $i_3 < k_1 - 1$ since otherwise the union of $A \cup B^2$, C^3, C^3, $(b^2_{k_2-1}, c^3_1)$ and $(b^2_{k_2-1}, c^4_1)$ would contain an $(\ell + k_1 - 1, k_1, k_2)$-fork, contradicting the maximality of ℓ.

For sake of contradiction, assume that both $c^3_{i_3}$ and $c^4_{i_4}$ have no outneighbour in $(A\setminus\{a_\ell\})\cup C^1\cup C^2$. Because $|V(B^1)\cup V(B^2)\cup V(C^3)\cup V(C^4)\cup\{a_\ell\}| = k_1+k_2+i_3+i_4-1 < 2(k_1+k_2)-2$ (since $i_3 + i_4 < k_1 + k_2 - 1$) and $\delta^+(D) \geq 2(k_1 + k_2 - 1) + 1$, then there exist distinct vertices $c^3_{i_3+1}, c^4_{i_4+1} \in V(D - (F \cup C^1 \cup C^2 \cup C^3 \cup C^4))$ (if $i_3 \geq k_2$, we only define c_{i_3+1}) such that $(c^3_{i_3}, c^3_{i_3+1}), (c^4_{i_4}, c^4_{i_4+1}) \in A(D)$. This contradicts the maximality of i_3.

So one of $c^3_{i_3}, c^4_{i_4}$ has an outneighbour in $(A \setminus \{a_\ell\}) \cup C^1 \cup C^2$. We assume that it is $c^3_{i_3}$; the case when it is $c^4_{i_4}$ is similar.

If $c^3_{i_3}$ has an outneighbour $a_q \in A \setminus \{a_\ell\}$ (for some $q < \ell$), then the union of either the dipaths $(a_\ell, B^1, C^1, a_j, \ldots, a_q)$ and (a_ℓ, B^2, C^3, a_q) (if $q \geq j$), or the dipaths (a_ℓ, B^1, C^1, a_j) and $(a_\ell, B^2, C^3, a_q, \cdots, a_j)$ (if $q < j$), is a subdivision of $C(k_1, k_2)$.

If $c^3_{i_3}$ has an outneighbour $c^1_h \in V(C^1)$ for some $1 \leq h \leq i_1$, then the union of the dipaths $(a_\ell, B^1, c^1_1, \cdots, c^1_h)$ and $(a_\ell, B^2, C^3, c^1_h)$ is a subdivision of $C(k_1, k_2)$. Similarly, we find a subdivision of $C(k_1, k_2)$ if $c^3_{i_3}$ has an outneighbour in C^2. □

3.1.4 Three dipaths between two vertices

A slight adaptation of the proof of Theorem 84 leads to a stronger result. Let k_1, k_2, k_3 be positive integers. Remember that $B(k_1, k_2; k_3)$ is the digraph formed by three internally disjoint paths between two vertices x, y, two (x, y)-dipaths, one of size k_1, the other of size k_2, and one (y, x)-dipath of size k_3. We denote by $B_{uv}(k_1, k_2; k_3)$ a subdivision of $B(k_1, k_2; k_3)$ where the vertex u plays the role of x and the vertex v plays the role of y.

Theorem 85. *Let k_1, k_2, k_3 be positive integers with $k_1 \geq k_2$. Let D be a digraph with $\delta^+(D) \geq 3k_1 + 2k_2 + k_3 - 5$. Then D contains $B(k_1, k_2; k_3)$ as a subdivision.*

Proof. Let ℓ be an integer. An $(\ell, k_3; k_1, k_2)$*-fork* is a digraph obtained from the union of four disjoint directed paths $P = (p_1, \ldots, p_\ell)$, $A = (a_1, \cdots, a_{k_3-1})$, $B^1 = (b^1_1, \cdots, b^1_{k_1-1})$ and $B^2 = (b^2_1, \cdots, b^2_{k_2-1})$ by adding the arcs (p_ℓ, a_1), (a_{k_3-1}, b^1_1) and (a_{k_3-1}, b^2_1).

Since a $(1, k_3; k_1, k_2)$-fork has $k_1 + k_2 + k_3 - 2$ vertices and $\delta^+(D) \geq k_1 + k_2 + k_3 - 2$, then D contains a $(1, k_3; k_1, k_2)$-fork as a subdigraph. So, let $\ell \geq 1$ be the largest integer such that D contains an $(\ell, k_3; k_1, k_2)$-fork as a subdigraph. Let F be such a fork. For convenience, we denote its subpaths and vertices by their labels in the above definition.

If there exist $i, j \in [\ell]$, with $i \leq j$, such that $p_i \in N^+(b^1_{k_1-1})$ and $p_j \in N^+(b^2_{k_2-1})$ or $p_i \in N^+(b^2_{k_2-1})$ and $p_j \in N^+(b^1_{k_1-1})$, then F contains a $B_{a_{k_3-1}p_j}(k_1, k_2; k_3)$.

So, let us assume that $b^1_{k_1-1}$ has no outneighbour in P (the case where $b^2_{k_2-1}$ has no outneighbour in P is similar). Since $|A\cup B^1\cup B^2| = k_1+k_2+k_3-3$ and $\delta^+(D) \geq k_1+k_2+k_3-1$, $b^1_{k_1-1}$ has two distinct outneighbours, say c^1_1 and c^2_1, not in F.

Let $i_1 \geq 1$ be the largest integer such that there exist two disjoint directed paths $C^1 = (c^1_1, \ldots c^1_{i_1})$ and $C^2 = (c^2_1, \ldots, c^2_{i_2})$ in $D-F$ with initial vertex c^1_1 and c^2_1 respectively and length i_1 and $i_2 = \min\{k_2-1, i_1\}$. If $i_1 \geq k_1-1$, then $i_2 \geq k_2-1$, and thus $P\cup A\cup B^1\cup C_1\cup C_2$ would contain a fork that contradicts the maximality of ℓ. Hence we may assume that $i_1 \leq k_1 - 2$ (and in particular $|V(C_1) \cup V(C_2)| \leq k_1 + k_2 - 3$).

For sake of contradiction, assume that both $c^1_{i_1}$ and $c^2_{i_2}$ have no outneighbour in P. Since $|V(A) \cup V(B^1) \cup V(B^2) \cup V(C^1) \cup V(C^2)| \leq 2k_1 + 2k_2 + k_3 - 6 < \delta^+(D) - 2$, then there exist $c^1_{i_1+1}, c^2_{i_2+1} \in V(D - (F \cup C^1 \cup C^2))$ such that $(c^1_{i_1}, c^1_{i_1+1}), (c^2_{i_2}, c^2_{i_2+1}) \in A(D)$ and $c^1_{i_1+1} \neq c^2_{i_2+1}$. This contradicts the maximality of i_1. Henceforth, we assume that $c^1_{i_1}$ has an outneighbour $p_i \in P$ (the case in which $c^2_{i_2}$ has an outneighbour in P is similar).

If $b^2_{k_2-1}$ has also an outneighbour $p_j \in P$, then $F \cup C_1$ contains a $B_{a_{k_3}p_j}(k_1, k_2; k_3)$ if $i \leq j$, and a $B_{a_{k_3}p_i}(k_1, k_2; k_3)$ if $j \leq i$.

So, we may assume that $b^2_{k_2-1}$ has no outneighbour in P. Hence, $b^2_{k_2-1}$ has two distinct outneighbours, say c^3_1 and c^4_1, not in $F \cup C^1$. Let $i_3 \geq 1$ be the largest integer such that there exist two disjoint dipaths C^3 and C^4 in $D - (F \cup C^1)$ with initial vertex c^3_1 and c^4_1 respectively and length i_3 and $i_4 = \min\{k_2 - 1, i_3\}$. If $i_3 \geq k_1$, then $i_4 \geq k_2 - 1$ and thus $P \cup A \cup B^2 \cup C_3 \cup C_4$ contains a fork that contradicts the maximality of F. Thus, we may assume that $i_3 \leq k_1 - 2$. In particular $|V(C_3) \cup V(C_4)| \leq k_1 + k_2 - 3$.

Suppose to the contrary that both $c^3_{i_3}$ and $c^4_{i_4}$ have no outneighbour in P, where $c^3_{i_3}$ and $c^4_{i_4}$ are the last vertices of C^3 and C^4. Note that $|V(A) \cup V(B^1) \cup V(B^2) \cup V(C^1) \cup V(C^3) \cup V(C^4)| \leq 3k_1 + 2k_2 + k_3 - 7 \leq \delta^+(D) - 2$. Hence, there exist distinct vertices $c^3_{i_3+1}, c^4_{i_4+1} \in V(D - (F \cup C^1 \cup C^3 \cup C^4))$ such that $(c^3_{i_3}, c^3_{i_3+1}), (c^4_{i_4}, c^4_{i_4+1}) \in A(D)$. This contradicts the maximality of i_3.

Therefore, one of $c^3_{i_3}, c^4_{i_4}$ has an outneighbour in p_j in P. We assume that it is $c^3_{i_3}$; the case when it is $c^4_{i_4}$ is similar. We conclude that $F \cup C_1 \cup C_3$ contains a $B_{a_{k_3}p_j}(k_1, k_2; k_3)$ if $i < j$, and a $B_{a_{k_3}p_i}(k_1, k_2; k_3)$ if $j < i$. □

3.2 Immersions of transitive tournaments

An interesting weakening of Mader's conjecture has been made by DeVos et al.[20]:

Conjecture 15 (DeVos et al. [20]). For every $k \geq 1$, there exists an integer $h(k)$ such that every digraph with minimum outdegree at least $h(k)$ contains an immersion of TT_k.

In the same paper, they proved the result for Eulerian digraphs. In fact they showed that, in the case of Eulerian digraphs, large minimum outdegree is enough to force immersions of large complete digraphs.

Remember that $F(k, l)$ is the multidigraph consisting of k vertices $x_1, \ldots, x_k$ and l arcs from x_i to x_{i+1} for every $1 \leq i \leq k - 1$. It is clear that $F(k, \binom{k}{2})$ contains an immersion of TT_k, so the following theorem proved in [45] implies Conjecture 15.

Theorem 86. *For every $k \geq 1$ and $l \geq 1$, there exists a function $h(k, l)$ such that every multidigraph with minimum outdegree at least $h(k, l)$ and multiplicity at most kl contains an immersion of $F(k, l)$.*

Proof. Note that $F(k, 1)$ is a dipath on k vertices and thus $h(k, 1) = k$. We prove the result for $h(k, l) = 2k^3l^2$ and $l \geq 2$. We proceed by induction on k. For $k = 1$ this is trivial because $F(1,l)$ is one vertex. Suppose now that the result holds for k and assume for a contradiction that it does not hold for $k+1$. Let D be the multidigraph with the smallest number of arcs and vertices such that D has multiplicity at most $(k+1)l$, all but at most $c_1 = k + (k+1)l$ vertices have outdegree at least $h(k+1,l)$ and without an immersion of $F(k+1,l)$. By minimality of D, every vertex has outdegree exactly $h(k+1,l)$, except c_1 of them with outdegree 0. Call T the set of vertices of outdegree 0. Suppose we want to remove arcs from D such that the multiplicity of the remaining digraph is at most kl, while keeping the minimum outdegree as large as possible. For a vertex v, the worst case is when, for every vertex $y \in N^+(v)$, the multiplicity of (v,y) is equal to $(k+1)l$. In this case we have to remove at most l arcs for each of the $\frac{h(k+1,l)}{(k+1)l}$ vertices of $N^+(v)$. Therefore, by removing T and these parallel arcs, we obtain a multidigraph of outdegree greater than $d' = h(k+1,l) - c_1(k+1)l - \frac{h(k+1,l)}{k+1}$ with multiplicity kl. Because $h(k+1,l) - h(k,l) = 2(3k^2+3k+1)l^2$ and $c_1(k+1)l + \frac{h(k+1,l)}{(k+1)} = k(k+1)l + 3(k+1)^2l^2$, we get that $d' \geq h(k,l)$ and by induction there exists an immersion of $F(k,l)$ in $D - T$. Call $X = \{x_1, \cdots, x_k\}$ the set of vertices of the immersion and $P_{i,j}$ the jth directed path of this immersion from x_i to x_{i+1}. We can assume this immersion is of minimum size, so that every vertex in $P_{i,j}$ has exactly one outgoing arc in $P_{i,j}$. Let D' be the multidigraph obtained from D by removing all the arcs of the $P_{i,j}$ and the vertices $x_1, \dots, x_{k-1}$. By the previous remark, the outdegree of each vertex in D' is either 0 if this vertex belongs to T or at least $h(k+1,l) - (k-1)l - (k-1)(k+1)l$.

For every vertex $y \in D' - x_k$, there do not exist l arc-disjoint directed paths from x_k to y in D', for otherwise there would be an immersion of $F(k+1,l)$ in D. Hence, by Menger's Theorem there exists a set E_y of less than l arcs such that there is no directed path from x_k to y in $D' \setminus E_y$. Define C_y for every vertex $y \in D' - x_k$ as the set of vertices which can reach y in $D' \setminus E_y$. Now take Y a minimal set such that $\cup_{y \in Y} C_y$ covers $D' - x_k$. We claim that Y consists of at least $c_2 \geq \frac{h(k+1,l)-(k-1)l-(k-1)(k+1)l}{l} \geq 2c_1$ elements, as $\cup_{y \in Y} E_y$ must contain all the arcs of D' with x_k as tail.

For each $y \in Y$, define S_y as the set of vertices which belong to C_y and no other $C_{y'}$ for $y' \in Y$. Since Y is minimal, every S_y is non-empty. Note that for $u \in S_y$, if there exists $y' \in Y \setminus y$ and $v \in C_{y'}$ such that $uv \in A(D)$, then $uv \in E_{y'}$. Note that $T \subset Y$ as vertices in T have outdegree 0 and if $y \in Y \setminus T$ then S_y consists only of vertices of outdegree $h(k+1,l)$ in D.

Let R be the digraph with vertex set Y and arcs from y to y' if there is an arc from S_y to $C_{y'}$. As noted before, $d^-_R(y) \leq |E_y| \leq l$. The average outdegree of the vertices of $Y \setminus T$ in R is then at most $\frac{c_1 l + (c_2 - c_1)l}{c_2 - c_1} \leq 2l$. Let y be a vertex of $R \setminus T$ with outdegree at most this average. Let H be the digraph induced on D' by the vertices in S_y to which we add X, all the arcs that existed in D (with multiplicity) from vertices of S_y to vertices of X and the following arcs: For each $P_{i,j}$, let $z_1, z_2, \dots, z_l = P_{i,j} \cap S_y$, where z_i appears before z_{i+1} on $P_{i,j}$ and add all the arcs (z_i, z_{i+1}) to H. Note that, if (x,y) is an arc of D', then by minimality of the immersion of $F(k,l)$, every time x appears before y on some $P_{i,j}$, then $P_{i,j}$ uses one of the arcs (x,y). Thus for each pair of vertices x and y in H, either $(x,y) \in A(D)$ and the number of (x,y) arcs in H is equal to the one in D, or $(x,y) \notin A(D)$ and the number of

(x, y) arcs in H is bounded by $(k-1)l$. This implies that H has multiplicity at most $(k+1)l$.

Claim 86.1. *H is a multidigraph with multiplicity at most $(k+1)l$, such that all but at most c_1 vertices have outdegree greater than $h(k+1, l)$ and H does not contain an immersion of $F(k+1, l)$.*

Proof of the claim. Suppose H contains an immersion of $F(k+1, l)$, then by replacing the new arcs by the corresponding directed paths along the $P_{i,j}$ we get an immersion of $F(k+1, l)$ in D. Moreover, we claim that the number of vertices in H with outdegree smaller than $h(k+1, l)$ is at most $k + 2l + (k-1)l = c_1$. Indeed, the vertices of H that can have outdegree smaller in H than in D are the x_i, or the vertices with outgoing arcs in $E_{y'}$ for some $y' \in Y \setminus y$, or the vertices along the $P_{i,j}$. But with the additions of the new arcs, we know that there is at most one vertex per path $P_{i,j}$ that loses some outdegree in H. $\diamond$

However, since H is strictly smaller than D, we reach a contradiction.

$\square$

Chapter 4

The Erdős-Sands-Sauer-Woodrow conjecture

All the proofs in this chapter are joint work with Nicolas Bousquet and Stéphan Thomassé.

In this chapter we will present a proof of conjecture 44:

Theorem 49. *For every $k \geq 1$, if T is a complete multidigraph whose arc set is the union of the arc sets of k quasi-orders, then $\gamma(T) = O(ln(2k) \cdot k^{k+2})$.*

The first part of this chapter is dedicated to the proof of the case $k = 3$ of Conjecture 45:

Proposition 87. *There exists an integer C such that every 3-transitive tournament T has $\gamma(T) \leq C$.*

While Proposition 87 is much weaker than Theorem 49, we decided to present it for two reasons. First it uses a new variation of the VC-dimension technique that we find interesting and could be applied to prove other results, and second it provides a context onto how we arrived to the short and simple proof of Theorem 49.

4.1 Domination in 3-transitive tournaments

In this section, T is a 3-transitive tournament. We will prove that T has bounded domination number.

By definition, $A(T)$ decomposes into three sets, R, G and B such that $T[R]$, $T[G]$ and $T[B]$ are quasi-orders. We will use the colours *red*, *blue* and *green* to speak about the sets R, G and B, respectively. A *chain* is a monochromatic subtournament of T. As a chain is a transitive tournament, we will refer to a chain as its directed Hamiltonian path $x_1, \ldots, x_k$.

Remember that the VC-dimension of a tournament is the VC-dimension of the inneighbourhood hypergraph. It is not very difficult to see that a 3-transitive tournament does not necessarily have bounded VC-dimension. For example, if you take two red chains such that all the arcs from one chain to the other are blue and the arcs in the other direction are green, then no matter the orientations of the arcs between the red chains, the resulting tournament will be a 3-transitive tournament. In particular, a random orientation will give arbitrarily large VC-dimension.

Let T be a tournament and S a set of vertices of T. The *VC-dimension* of S is the VC-dimension of S in the inneighbourhood hypergraph. S is said to be *shattered* if it is shattered in the inneighbourhood hypergraph. In a tournament, the VC-dimension argument tells us that the vertices which are difficult to dominate are the vertices with large in and outdegree inside some shattered set. However, in the example we just provided, the structure of a shattered set makes the domination of these vertices very easy.

The goal of our approach was to understand the structure of shattered sets in order to deal with them during the sampling part of the proof of the ϵ-net Theorem. In the first part of the proof, we analyse shattered sets in a 3-transitive tournament and prove that they have a long monochromatic chain, meaning that we only have to consider shattered chains.

4.1.1 Structure of shattered sets

Let $C = x_0, \dots, x_k$ be a chain of order $k+1$, we denote by $I(C)$ the set of vertices x_i for $0 < i < k$.

Lemma 88. *Let $C = x_0, \dots, x_k$ be a chain of a 3-transitive tournament T. If v is a vertex of T such that (v, x_0) and (x_k, v) belong to $A(T)$, then none of the arcs between v and $I(C)$ have the same colour as C.*

Proof. Suppose C is a red chain and let x_i be a vertex of $I(C)$. If (x_i, v) belongs to T and is coloured red, then by transitivity (x_0, v) belongs to T and is coloured red. Likewise, if (v_i, x) is coloured red then (v, x_k) belongs to T. □

Let $C = x_0, \dots, x_k$ be a chain of a 3-transitive tournament T and let $G(C)$ be the set of vertices v of T such that (v, x_0) and (x_k, v) belong to $A(T)$. The previous lemma implies that, for a vertex v of $G(C)$, it is not possible to have one arc from $I(C)$ to v and one arc from v to $I(C)$ of the same colour. Indeed since this colour cannot be the colour of C by the previous lemma, it would force an arc of another colour between vertices in C. Thus, the vertices of $G(C)$ divide in two *types*, depending on the colour of $A(v, I(C))$ and $A(I(C), v)$. For example, if C is a red chain, then the two types are the following:

- The vertices v such that the arcs in $A(v, I(C))$ are green and the arcs in $A(I(C), v)$ are blue.

- The vertices v such that the arcs in $A(v, I(C))$ are blue and the arcs in $A(I(C), v)$ are green.

Lemma 89. *Let $C = x_0, \dots, x_k$ be a chain of a 3-transitive tournament T. Let u and v be two vertices of $G(C)$ with the same type such that there exist two indices $0 < i < k$ and $0 < j < k$ where ux_ivx_j is a directed cycle, then the arc between u and v is of the same colour as C.*

Proof. Suppose C is a red chain. Without loss of generality, we can assume that (u, x_i) and (v, x_j) are blue and (x_i, v) and (x_j, u) are green and moreover that (u, v) belongs to T. If (u, v) is blue, then by transitivity (u, x_j) belongs to T. Likewise, if (u, v) is green then (x_j, v) belongs to T. This means that (u, v) is red. $\square$

Lemma 90. *Let T be a 3-transitive tournament and S a set of vertices of T such that S is a shattered set of the inneighbourhood hypergraph of T. Then $T[S]$ cannot have two chains of length 6 of different colours.*

Proof. Suppose it does and let $C_1 = (c_0^1, \dots, c_5^1)$ and $C_2 = (c_0^2, \dots, c_5^2)$ be those two chains. For each $i \in [2]$, we define 5 sets of vertices of $G(C_i)$:

- Set A_i consists of vertices x such that xc_1^i, xc_2^i, $c_3^i x$ and $c_4^i x$ belong to $A(T)$.
- Set B_i consists of vertices x such that xc_1^i, xc_3^i, $c_2^i x$ and $c_4^i x$ belong to $A(T)$.
- Set L_i consists of vertices x such that xc_1^i, xc_4^i, $c_3^i x$ and $c_2^i x$ belong to $A(T)$.
- Set D_i consists of vertices x such that xc_3^i, xc_2^i, $c_1^i x$ and $c_4^i x$ belong to $A(T)$.
- Set E_i consists of vertices x such that xc_4^i, xc_2^i, $c_3^i x$ and $c_1^i x$ belong to $A(T)$.

The purpose of these sets is that, if there exist two vertices, say a and b, such that $a \in A_1 \cap A_2$ and $b \in B_1 \cap B_2$ and these vertices have the same type towards C_1 and C_2, then we reach a contradiction. Indeed $ac_2^1 bc_3^1$ and $ac_2^2 bc_3^2$ are directed cycles, but since C_1 and C_2 are two chains of different colours, then Lemma 89 implies that the arc between a and b has two different colours, which is impossible.

To find two vertices like this, we use the fact that the set is shattered. Because the set is shattered, it means that for every possible subset S' of S, there exists a vertex x in T such that the inneighbourhood of x in S is precisely S' (and thus the outneighbourhood is precisely $S \setminus S'$). This means that we can find five vertices: a (resp. b, l, d, e) that belongs to $A_1 \cap A_2$ (resp. $B_1 \cap B_2$, $L_1 \cap L_2$, $D_1 \cap D_2$, $E_1 \cap E_2$). Now each of these vertices has one of two specific types regarding C_1 and C_2, this means there are 4 combinations of types. By the Pigeonhole Principle, we can find two vertices, say a and b such that they both have the same type towards C_1 and the same type towards C_2. This ends the proof.

$\square$

Lemma 91. *Let T be a 3-transitive tournament and S a set of vertices of T such that S is a shattered set of the inneighbourhood hypergraph of T. Then $T[S]$ contains a chain of length $|S|/25$.*

Proof. By the previous lemma and without loss of generality, all the chains of S in blue or red have size smaller than 5. Consider the digraph induced by the arcs with colour blue. By Theorem 26, there exists a stable set of size $|S|/5$. Now consider the digraph induced by the arcs of colour red on these $|S|/5$ vertices. Again, by Theorem 26, there exists a stable set of size $|S|/25$. This set is a green chain of length $|S|/25$. $\square$

4.1.2 Dominating shattered sets

Let u and v be two vertices of a 3-transitive tournament T, define $S(u,v)$ as the set of vertices x such that there exist two disjoint vertices $u_1 = u_1(x), u_2 = u_2(x)$, such that u, u_1, v, u_2 is a monochromatic chain (in that order), x dominates u and v and x is dominated by the u_i.

Lemma 92. *Let T be a 3-transitive tournament. For every pair of vertices u and v of T, there exists a set $\phi(u,v)$ of four vertices which dominates $S(u,v)$.*

Proof. Without loss of generality we can assume that (u,v) is a blue arc. Let x be a vertex of $S(u,v)$ and u_1, u_2 the vertices completing the blue chain, then one of the following is true:

(i) The arcs (u_i, x) are green and the arcs (x,u) and (x,v) are red; or

(ii) the arcs (u_i, x) are red and the arcs (x,u) and (x,v) are green.

To prove this, first note that none of the arcs (u_i,x), (x,u) or (x,v) can be blue. Indeed, if (x,u) or (x,v) is blue, then (x,u_2) must be an arc, and if (u_1,x) or (u_2,x) is blue, then (u,x) must be an arc. Finally, since uu_1vu_2 is a blue chain, (x,u) and (x,v) must have different colour than (u_2,x) and (u_1,x).

Consider S_1 the set of vertices of $S(u,v)$ such that (i) is true and let $U_1 = \{u_1(x); x \in S_1\}$ and $U_2 = \{u_2(x); x \in S_1\}$. First note that if $a \in U_1$ and $b \in U_2$ then (a,b) is a blue arc by transitivity with v. Now let $x \in S_1$, we claim that for every y in U_1, either the arc (y,x) exists and is green or the arc (x,y) exists and is red. Indeed if (y,x) is blue, then (u,x) must be blue, if (x,y) is blue then (x,v) must be blue, if (x,y) is green then $(u_2(x),y)$ must be green and if (y,x) is red then (y,u) must be red. In all cases, this is a contradiction. The same can be said with U_2, so we have the following property: For every vertex $x \in S_1$, the set of arcs $A(x, U_1 \cup U_2)$ is red and the set of arcs $A(U_1 \cup U_2, x)$ is green. This means that for any x_1 and x_2 two different vertices of S_1, if the arc (x_1,x_2) is red then by transitivity $N^-_{U_1\cup U_2}(x_1) \subseteq N^-_{U_1\cup U_2}(x_2)$ and if the arc (x_1,x_2) is green then $N^+_{U_1\cup U_2}(x_1) \subseteq N^+_{U_1\cup U_2}(x_2)$ which implies $N^-_{U_1\cup U_2}(x_2) \subseteq N^-_{U_1\cup U_2}(x_1)$.

Now define the following bicoloured complete multidigraph T' on S_1: (x,y) is a grey arc of T' if $N^-_{U_1\cup U_2}(x) \subseteq N^-_{U_1\cup U_2}(y)$ and (x,y) is a blue arc of T' if it is a blue arc of T. T' satisfies the hypothesis of Theorem 42 and therefore has a dominating vertex x, and thus, in T, $\{x\} \cup u_1(x)$ dominates S_1.

By doing the same thing with the set of vertices such that (ii) is true, we obtain a set of four vertices dominating $S(u,v)$.

□

For a set X, let $\phi(X)$ be the union of all the $\phi(x,y)$ for x and y in X.

4.1.3 Sampling argument

We are now ready to prove Proposition 87. Basically we will show that, if we take a set of vertices S of a certain size at random, then with positive probability $S \cup \phi(S)$ is a dominating set of T.

We need the following (not tight) lemma:

Lemma 93. *Let S be a set of size $2s$ and $S = (A, B)$ a random bipartition such that $|A| = |B| = s$ and G a subset of S of size t. Then $Pr[|G \cap A| < t/4$ or $|G \cap B| < t/4] < \frac{t}{4}(2/3)^{t/4}$.*

Proof. By considering the set between A and B with the smallest intersection with G, we get the following bound:

$$\frac{\sum_{i=0}^{t/4} \binom{t}{i}\binom{2s-t}{s-i}}{\sum_{i=0}^{t/2} \binom{t}{i}\binom{2s-t}{s-i}} \leq \frac{\sum_{i=0}^{t/4} \binom{t}{i}\binom{2s-t}{s-i}}{\sum_{i=t/4}^{t/2} \binom{t}{i}\binom{2s-t}{s-i}} \leq \frac{\sum_{i=0}^{t/4} \binom{t}{i}}{\sum_{i=t/4}^{t/2} \binom{t}{i}} \leq \frac{t}{4}\frac{\binom{t}{t/4}}{\binom{t}{t/2}} \leq \frac{t}{4}\frac{t/4 \dots t/2}{t/2 \dots 3t/4}.$$

□

We first show the case where every vertex has large indegree.

Proposition 94. *There exists a constant C such that any 3-transitive tournament T with minimum indegree greater than $|T|/6$ has domination number bounded by C.*

Proof. Let t be an integer and let A be a random set of vertices picked by t^2 independent random draws, where each element is drawn from $V(T)$ uniformly. The goal is to show that $A \cup \phi(A)$ is a dominating set with positive probability when t is big enough (but still a constant). Suppose $A \cup \phi(A)$ is not a dominating set and let B be another random set obtained by picking t^2 vertices randomly. Formally, we regard A and B as sequences of elements of T, with possible repetitions. So $A = (a_1, \dots, a_{t^2})$ and $B = (b_1, \dots, b_{t^2})$. For a sequence $S = (s_1, \dots, s_l)$, $|S|$ will denote the number of elements in the sequence. Abusing the notation slightly, $d_S^-(u)$ will denote the number of indices $i \in [l]$ such that $s_i \in N^-(u)$, so one vertex could count multiple times. Let E_0 be the event "there exists a vertex u such that u dominates $A \cup \phi(A)$ and $d_B^-(u) \geq t^2/12$". Since we assumed that $A \cup \phi(A)$ is not a dominating set, there exists a vertex u dominating this set. Hence $Pr[E_0] \geq Pr[d_B^-(u) \geq t^2/12]$. Since $d^-(u) \geq n/6$, the following claim can be proved with Chebyshev's inequality (see Lemma 10.2.6 of [51] for the complete proof)

Claim 94.1. *$Pr[E_0] \geq 1/2$.*

Now we are going to bound $Pr[E_0]$ differently. First let $N = (z_1, \dots, z_{2t^2})$ be a random sequence of $2t^2$ elements chosen uniformly. In a second stage, we randomly chose t^2 positions in N and put the elements at these positions in A, and the remaining elements in B. The resulting distribution of A and B is the same as the one described before. We now prove that the probability of E_0 is small. To do so, consider the set N' corresponding to the elements of N (so removing repetitions). If this set has VC-dimension bounded by $25t$, it means we can do the same proof as the ϵ-net Theorem and prove that $Pr[E_0]$ is smaller that $1/2$. If the VC-dimension is not bounded by $25t$, then by Lemma 91, there exists a shattered chain of length t. Let $T_1, \dots, T_l$ be a maximal collection of disjoint subsequences of N, such that the elements of each T_i forms a shattered chain of length t in N', without repetition. This means that one element can belong to different chains if it appeared multiple times in N. Note that $l \leq 2t$. Let $R = N \setminus \cup_i T_i$ be the remaining subsequence of N. Note that, if we note R' the set of elements contained in R, then R' has VC-dimension bounded by $25t$.

For each $i \in [l]$, let L_i be the event: "there exists a vertex x with $d^-_{T_i}(x) \geq t/48$ and the random partition of N is such that x is not dominated by $A \cap T_i \cup \phi(A \cap T_i)$". Let C be the event:"$|R| > \frac{t^2}{23}$, there exists a vertex x with $N^-_R(x) > |R|/12$ and the random cut is such that x is not dominated by $A \cap R$". Remember that E_0 is the event that there exists a vertex u such that u dominates $A \cup \phi(A)$ and $d^-_B(u) \geq t^2/12$. In order for such a vertex u to exists, this vertex u must satisfy $d^-_N(u) \geq t^2/12$. Suppose this vertex is such that $d^-_{T_i}(u) < t/48$ for all $i \in [l]$. This means that $\frac{2t^2-|R|}{48} + R \geq \frac{t^2}{12}$, which implies that $|R| \geq \frac{2t^2}{47} \geq \frac{t^2}{23}$ and $N^-_R(x) > |R|/12$. This means that, if E_0 is satisfied, then either one of the L_i is satisfied, or C is satisfied. Next we bound the probabilities of all these events, which will finish the proof.

Claim 94.2. $Pr[L_i] \leq 2(3/4)^{t/100}t^3 + \frac{t}{4}(2/3)^{t/4}$.

Subproof. We will condition on the event S_i : "$|T_i \cap A| > \frac{t}{4}$". By Lemma 93, $Pr[\bar{S}_i] < \frac{t}{4}(2/3)^{t/4}$. To compute $Pr[L_i|S_i]$, we condition on the position (relatively to T_i) of the first, middle and last inneighbour of a vertex x with $d^-_{T_i}(x) \geq t/48$ on T_i. Remember that the T_i are without repetition, so these vertices are defined without ambiguity. If A contains one vertex of T_i between the first and the middle vertex and one vertex between the middle and the last, then by Lemma 92, x is dominated by $(A \cap T_i) \cup \phi(A \cap T_i)$. Because x has indegree greater than $\frac{t}{48}$ in T_i, this is bounded by the probability that a random sample of $t/4$ elements out of t elements in T avoid one of the two sets of $s = \lfloor t/96 \rfloor$ elements. For each of the two sets, the probability that it is avoided is bounded by

$$\frac{\binom{t-s}{t/4}}{\binom{t}{t/4}} = \frac{(t-s)!(t-t/4)!}{(t-s-t/4)!t!} = \frac{(t-s-t/4+1)\dots(t-t/4)}{(t-s+1)\dots t} \leq (3/4)^s.$$

We can suppose that t is big enough so that this value is bounded by $(3/4)^{t/100}$, and the probability that one of the two sets is avoided is then smaller than $2(3/4)^{t/100}$. Overall, since there are at most $\binom{t}{3}$ choices for the positions of the first, middle and last inneighbour on T_i, we get that

$$Pr[L_i] \leq 2(3/4)^{t/100}t^3 + \frac{t}{4}(2/3)^{t/4}$$

. $\Diamond$

Claim 94.3. $Pr[C] \leq (276/t^2)^{t^2/276}(\frac{et}{575})^{25t} + \frac{t^2}{100}(2/3)^{t^2/100}$

Subproof. We condition on the event R_1 that $|R \cap A| > r/4$. By Lemma 93, $P[\bar{R}_1] < \frac{t^2}{100}(2/3)^{t^2/100}$. If we look at the trace of the inneighbourhood hypergraph on R, it has VC-dimension bounded by $25t$ by Lemma 91. Fix the inneighbourhood S of x in R, then the probability $P_S = Pr[S \cap A = 0|R_1]$ is at most the probability that a random sample of $r/4$ elements avoids the $r/12$ elements of S so:

$$P_S \leq \frac{\binom{r-r/12}{r/4}}{\binom{r}{r/4}} \leq \frac{(r-r/12)!(r-r/4)!}{(r-r/3)!r!} < \frac{(r-r/4)!}{(r-r/3)!} \leq (12/r)^{r/12}$$

Because the VC-dimension of R is bounded by $25t$, by Lemma 48 there are at most $(\frac{er}{25t})^{25t}$ possible inneighbourhoods S and by the union bound, $Pr[C] \leq (276/t^2)^{t^2/276}(\frac{et}{575})^{25t} + \frac{t^2}{100}(2/3)^{t^2/100}$. $\Diamond$

Finally, because $l \leq 2t$, we have

$$Pr[E_0] \leq (276/t^2)^{t^2/276}(\frac{et}{575})^{25t} + \frac{t^2}{100}(2/3)^{t^2/100} + 2t(2(3/4)^{t/100}t^3 + \frac{t}{4}(2/3)^{t/4})$$

And when t is big enough, this is smaller than $1/2$, so we reach a contradiction. $\square$

Now we show how to prove the general case, using the probability distribution given by Lemma 46 (actually we use a slightly more precise version found in [27]).

Proof of Proposition 87. Let C be as Proposition 94. If T has a dominating vertex we are done so we can assume it does not. By Thereom 1.5 of [27], there exists a rational number $p(v)$ for every vertex v such that p is a probability distribution on the vertices, $p(v) \leq 1/3$ for all v and $p(N^-[v]) \geq 1/2$. This implies that $p(N^-(v)) \geq 1/6$. By blowing each vertex v of T with a red transitive tournament of size $Mf(v)$ for some integer M, we get a tournament T' satisfying the conditions of Proposition 94, so with a dominating set of size C. Finally, by taking for each vertex in the dominating set of T', its original vertex in T we obtain a dominating set of T of cardinality at most C. $\square$

4.2 General proof

In this section we prove Theorem 49.

The important part in the last proof is that, if A is a shattered set and B is the set of vertices with large in and outdegree in A, then the structure of B is simple. The question we had was to know if this was still true if T is the union of more than three quasi-orders. One way for B to be simple is to have bounded VC-dimension and a natural question was to ask what happens when B doesn't have bounded VC-dimension? Suppose there exists a set C in B such that C has large in and outdegree in B, what can we say about that set? The answer to this question, when all the arcs involved belong to the same transitive digraph is the following lemma which is the essential part of the proof of Theorem 49.

Let P be a quasi-order on S. We say that $A \subseteq P$ is *ϵ-dense* in P if there exists a probability distribution w on the vertices of P such that $w(N^-[x]) \geq \epsilon$ for every element x of A. Let T be a complete multidigraph whose arcs are the union of the arcs of k quasi orders $P_1, \dots, P_k$, we define $N_i^-[x]$ (resp $N_i^+[x]$) as the closed inneighbourhood (resp. outneighbourhood) of the digraph induced by P_i.

Lemma 95. *Let ϵ be a real in $[0, 1]$. There exists an integer $g(\epsilon)$ such that for every quasi-order P on a set A and two subsets $C \subseteq B$ of A such that B is ϵ-dense in P and C is ϵ-dense in B, there exists a set of $g(\epsilon)$ elements in A dominating C.*

Proof. Let $w_A : A \to [0,1]$ and $w_B : B \to [0,1]$ be the probability distributions such that $w_A(N^-[x]) \geq \epsilon$ for every $x \in B$ and $w_B(N^-[x]) \geq \epsilon$ for every $x \in C$.

Let $g(\epsilon) = \lfloor \frac{ln(\epsilon)}{ln(1-\epsilon)} \rfloor + 1$ and pick independently at random according to the distribution w_A a (multi)set S of $g(\epsilon)$ elements of A. For every vertex $x \in B$, $Pr(x \in N^+[S]) \geq 1-(1-\epsilon)^{g(\epsilon)} > 1-\epsilon$. Thus, by linearity of w_B and the expectation, $\mathbb{E}(w_B(N^+[S])) > 1-\epsilon$. Therefore, there exists a choice of S such that $w_B(N^+[S]) > 1-\epsilon$. Since $w_B(N^-[y]) \geq \epsilon$ for every $y \in C$, the set $N^-[y]$ intersects $N^+[S]$. In particular, by transitivity, S dominates y. □

The next lemma is a direct consequence of Theorem 46.

Lemma 96. *Let T be a complete multidigraph whose arc set is the union of the arc sets of k quasi-orders. There exists a probability distribution w on $V(T)$ and a partition of $V(T)$ into sets $T_1, T_2, \ldots, T_k$ such that for every i and $x \in T_i$, $w(N_i^-[x]) \geq 1/2k$.*

Proof. By Theorem 46 there exists a weight function $w : V(T) \to [0,1]$ such that $w(N^-[x]) \geq 1/2$ for all $x \in T$. For every i in $[k]$, let T_i' be the subset of vertices such that $w(N_i^-[x]) \geq 1/2k$. The sets T_i' cover the vertices, so we can extract a partition with the required properties. □

We are now ready to prove the main theorem:

Proof of Theorem 49. Consider $P_1 = T_1, T_2, \ldots, T_k$ together with w the partition given by Lemma 96 applied to T. Each of the T_i is a complete multidigraph whose arc set is the union of the arc sets of k quasi-orders, this means we can apply Lemma 96 and obtain $T_{i,1}$, $T_{i,2}$ $\ldots$, $T_{i,k}$ together with a probability distribution w_i on T_i such that $w_i(N_j^-[x]) \geq 1/2k$ for every $x \in T_{i,j}$. By repeating this process k times, we obtain a sequence of $k+1$ partitions $P_1, \ldots, P_{k+1}$ with $P_i = \cup_{j_1,j_2,\ldots,j_i \leq k} T_{j_1,j_2,\ldots,j_i}$ such that for every $l \leq k+1$ and each $j_1, \ldots, j_l$ in $[k]^l$, $T_{j_1,j_2,\ldots,j_l}$ is a subset of $T_{j_1,j_2,\ldots,j_{l-1}}$ and the probability distribution $w_{j_1\ldots,j_{l-1}}$ is such that $w_{j_1\ldots,j_{l-1}}(N_{j_l}^-[x]) \geq 1/2k$ for every x in $T_{j_1,\ldots,j_l}$.

Fix $j_1, \ldots, j_{k+1}$, in $[k]^{k+1}$, by the Pigeonhole Principle there exist two indices, $i < l$ such that $j_i = j_l$, then by applying Lemma 95 where $T_{j_1,\ldots,j_{i-1}}$, $T_{j_1,\ldots,j_i}$ and $T_{j_1,\ldots,j_l}$ play the roles of, respectively, A, B and C there exists a set of size $g(1/2k)$ that dominates $T_{j_1,\ldots,j_l}$ and thus $T_{j_1,\ldots,j_{k+1}}$. This means that $\gamma(T) \leq k^{k+1} \cdot g(1/2k)$. Moreover since $g(1/2k) \leq ln(2k) \times (2k - 1/2 + o(1))$, we have $\gamma(T) = O(ln(2k) \cdot k^{k+2})$. □

4.3 The Sands-Sauer-Woodrow conjecture

In a recent paper, Harutyunyan et al. [33] proved that the fractional domination of a digraph D is bounded by $2\alpha(D)$. This means that there exists a probability distribution on the vertices, where every vertex x has $d^-[x] > \frac{1}{2\alpha(D)}$. With this result, we can adapt the previous proof to show that there exists a function $f(\alpha, k)$ such that any multidigraph whose arcs are the union of k quasi-orders and with stability number smaller than α is dominated by $f(\alpha, k)$ vertices. However Conjecture 43 remains open.

Chapter 5

Entropy compression

5.1 Orientations of Hypergraphs

We present here the proof of the following theorem, obtained with Nathann Cohen in [18].

Theorem 54. *Let p, k be fixed integers. There exists r_0 such that for every $r \geq r_0$, every r-uniform hypergraph with $\Delta_p(\mathcal{H}) \leq k$ admits a p-equitable orientation.*

We first describe a random algorithm that produces such orientation. To analyse the behaviour of this algorithm we will need the following lemma on permutations, which we will prove in the last subsection.

Lemma 55. *Let $p, k \in \mathbb{N}$ and $\alpha < 1$ be fixed. Let X be a set of cardinality r and let $\mathcal{L}_S$ be, for every $S \in \binom{X}{p}$, a collection of p-subsets of X with $|\mathcal{L}_S| \leq k$. Then, if no p-subset occurs in more than r^α of the $\mathcal{L}_S$, a random permutation σ of X satisfies $\sigma(S) \notin \mathcal{L}_S$ for every S with probability $\geq (1 - 2k/\binom{r}{p})^{\binom{r}{p}} = e^{-2k} + o(1)$ when r grows large.*

5.1.1 Algorithm

In what follows, we assume that every finite set S has an implicit enumeration on its elements, and in particular that the hyperedges of a hypergraph $\mathcal{H}$ are implicitly ordered. We will say that i *represents* an element $s \in S$ when s is the i-th element of S in this implicit ordering.

We will orient the hyperedges of $\mathcal{H}$ one by one as a (partial) equitable orientation of $\mathcal{H}$, i.e. in such a way that no p-subset of $V(\mathcal{H})$ appears more than once at the same position among the oriented hyperedges. To do so, we require the partial orientation to enforce an additional property.

Definition 97. Let $\mathcal{H}$ be a partially oriented r-uniform hypergraph. We say that an hyperedge $S \in \mathcal{H}$ is *pressured* by a family $\{S_1, \ldots, S_l\}$ of hyperedges (oriented by $\sigma_{S_1}, \ldots, \sigma_{S_l}$) if there exists $P \in \binom{[r]}{p}$ such that $\sigma_{S_i}^{-1}(P) \subseteq S$ for every i.

Note that Lemma 55 ensures that a partial orientation of $\mathcal{H}$ can be extended to an unoriented hyperedge S, provided that no family of more than r^α oriented hyperedges pressures S. It asserts, for $c < e^{-2k}$ and r sufficiently large, that at least $cr!$ orientations of S are admissible for this extension: we name them *good* permutations of S. Algorithm 2 selects an ordering randomly among them, while ensuring that no other hyperedge is pressured by a family of hyperedges larger than $r_1 = \lfloor r^\alpha \rfloor$.

Algorithm 2: A non-deterministic algorithm

Data: A r-uniform hypergraph $\mathcal{H}$ with $\Delta_p(\mathcal{H}) \leq k$
Result: A p-equitable orientation of $\mathcal{H}$

```
while not all hyperedges are oriented do
    S_1 ← unoriented hyperedge of smallest index
    Pick for S_1 the orientation indexed v_i (among ≥ cr! available)
    if some hyperedge S of H is pressured by a family {S_1, ..., S_{r_1}} then
        Cancel the orientation of all hyperedges S_i.
Return the oriented H
```

Algorithm 2 starts with every hyperedge being unoriented. At each step it orients the unoriented hyperedge of smallest index by choosing a random permutation amongst the $cr!$ first good permutations. We call *bad event* the event that a hyperedge $S \in \mathcal{H}$ is pressured by a family $\{S_1, \ldots, S_{r_1}\}$ of cardinality r_1. If a bad event occurs after orienting S_1, then the algorithm erases the orientation of the $S_1, \ldots, S_{r_1}$.

It is trivial to see that Algorithm 2 only returns p-equitable orientations of $\mathcal{H}$. Moreover, every time the algorithm chooses a random permutation, it does so among at least $cr!$ good ones by Lemma 55. Note that we need to consider large families pressuring already oriented hyperedges: indeed, we might have to cancel the orientation of such a hyperedge to redefine it again later.

Theorem 98. *Let $p, k \in \mathbb{N}$, $\alpha, c \in \mathbb{R}_{>0}$ with $\alpha < 1$ and $c < e^{-2k}$. For every sufficiently large r, there is a set of random choices for which Algorithm 2 terminates.*

In order to prove this result we will analyse the possible executions of the M first steps of Algorithm 2. To this end we make it deterministic by defining a log (following the idea of [55]) and obtain Algorithm 3, in the following way:

- Take as input a vector $v \in [cr!]^M$ which simulates the random choices.
- Output a log when it is not able to orient all hyperedges.

We define a *log of order M* to be a triple (R, X, F) where:

- R is a binary word whose length lies between M and $2M$.
- X is a sequence of h 7-tuples of integers $(x_1, x_2, x_3, x_4, x_5, x_6, x_7)$ where:

$$x_1 \leq \binom{r}{p}; \quad x_2 \leq k; \quad x_3 \leq \binom{\binom{r}{p}}{r_1-1}; \quad x_4 \leq k^{r_1-1};$$
$$x_5 \leq p!^{r_1-1}; \quad x_6 \leq (r-p)!^{r_1-1}; \quad x_7 \leq r!;$$

- F is an integer smaller than $(r!+1)^{|\mathcal{H}|}$ representing a partial orientation of $\mathcal{H}$.

The log of order M (or just log) is actually a trace of the deterministic algorithm's execution after M steps. Its objective is to encode which orientations get canceled during the algorithm's execution. We will show later that Algorithm 3 cannot produce the same log from two different input vectors $v, v' \in [cr!]^M$, and that, for M big enough, that the set of possible logs is smaller than $(cr!)^M$. We now describe the log and how Algorithm 3 produces it.

- R is initialized to the empty word. We append 1 to R whenever Algorithm 3 adds a new orientation; we append 0 whenever it cancels one.

- Consider the following bad event: after orienting S_1, a hyperedge $S \in \mathcal{H}$ is pressured by a family $\{S_1, \ldots, S_{r_1}\}$ of cardinality r_1. We note s_i the set of vertices that S_i maps to P. We associate the following 7-tuple which identifies the sets S_i as well as their orientations:

  - $x_1 < \binom{r}{p}$ represents the set s_1 among the $\binom{r}{p}$ possible subsets of size p of S_1.
  - $x_2 < k$ identifies S as one of the (at most k) hyperedges containing s_1.
  - $x_3 < \binom{\binom{r}{p}}{r_1-1}$ is an integer representing the set of subsets $s_2, \ldots s_{r_1}$ amongst the $\binom{r}{p}$ subsets of size p of S.
  - $x_4 < k^{r_1-1}$ is an integer representing the sequence $(y_2, \ldots, y_{r_1}) \in [k]^{r_1-1}$ such that the y_l-th edge containing s_l is S_l.
  - $x_5 < p!^{r_1-1}$ is an integer representing the sequence $(p_1, \ldots, p_{r_1})$, where $p_i \in [p!]$ represents the subpermutation of S_i onto s_i (we know it's a permutation of P).
  - $x_6 < (r-p)!^{r_1-1}$ is the integer representing the sequence $[p_2, \ldots, p_{r_1}]$, where $p_i \in [(r-p)!]$ represent the subpermutation of S_i onto $[r] \setminus s_i$.
  - $x_7 < r!$ is the integer representing the permutation chosen for S_1.

  X is the list of the 7-tuples describing the bad events, in the order in which they happen.

- F is the integer representing the partial orientation of $\mathcal{H}$ (i.e. a choice among $r!+1$ per hyperedge of $\mathcal{H}$) after M steps.

This gives the following algorithm:

Algorithm 3: A deterministic algorithm

Data:

1. A r-uniform hypergraph $\mathcal{H}$ with $\Delta_p(\mathcal{H}) \leq k$,
2. A vector $v \in [cr!]^M$

Result: A p-equitable orientation of $\mathcal{H}$, or a log of order M

$R \leftarrow \emptyset, X \leftarrow \emptyset$
for $1 \leq i \leq M$ **do**
  $S_1 \leftarrow$ unoriented hyperedge of smallest index
  Pick for S the orientation indexed v_i among $\geq cr!$ available
  if *some hyperedge of* $\mathcal{H}$ *is pressured by* $\{S_1, \ldots, S_{r_1}\}$ **then**
    Append 1 to the end of R
    Append to X a 7-tuple describing the conflict
    Cancel the orientation of all $r_1 + 1$ hyperedges involved in the conflict
  else if *all hyperedges are oriented* **then**
    Return the oriented $\mathcal{H}$
  else
    Append 0 to the end of R
$F \leftarrow$ the integer representing the partial orientation of $\mathcal{H}$.
Return (R, X, F)

We will show the following claim.

Claim 98.1. *Let e be a vector in $[cr!]^M$ from which Algorithm 3 cannot produce a p-equitable orientation of $\mathcal{H}$ and outputs a log (R, X, F). We can reconstruct e from (R, X, F).*

Proof of the claim. First we show that we can find, for every $z \leq M$, the set $C(z)$ of hyperedges which are oriented after z steps. We proceed by induction on z, starting from $C(0) = \emptyset$. At step $z+1$, Algorithm 3 chooses an orientation for the smallest index i not in $C(z)$. If, in R, the $(z+1)$-th 1 is not followed by a 0, then there is no bad event triggered by this step. In this case the set $C(z+1)$ is the set $C(z) \cup i$. Suppose now that the $(z+1)$-th 1 is followed by a sequence of 0: this means that the algorithm encountered a bad event. By looking at the number of sequences of 0 in R before the $z+1$-th 1 we can deduce the number of bad events before this one. This mean we can find, in X, the 7-tuple $(x_1, x_2, x_3, x_4, x_5, x_6, x_7)$ associated to this bad event. We take the following notations for the bad event : After orienting S_1, a hyperedge S of H is pressured by a family $\{S_1, \ldots, S_{r_1}\}$ of cardinality r_1. We note s_i the subset of S_i that is sent to P. S_1 is the last hyperedge we oriented (known by induction), x_1 indicates s_1 amongst the subset of S, x_2 indicates S amongst the set of hyperedges containing s_1, x_3 indicates the s_d for $d \in [2, r_1]$, and x_4 indicates the S_d for $d \in [2, r_1]$. In this case the set $C(z+1)$ is the set $C(z)$ for which we removed all the ES_d for $d \in [2, r_1]$.

We can now deduce the set $S(z)$ of all chosen orientations after z steps. We also proceed by induction, this time starting from step M. By construction, F is exactly the integer representing the partial orientation of $\mathcal{H}$ at step M. If the last letter of R is a 1, this means the last step of the algorithm consisted only of choice of an orientation. We just showed that we know which orientation was chosen after $M-1$ steps, so we can deduce the state of all orientations after $M-1$ steps. If the last letter is a 0, Algorithm 3 encountered a bad event. Keeping the notation of the bad event, let $(x_1, x_2, x_3, x_4, x_5, x_6, x_7)$ be the 7-tuple associated to this bad event. Like before x_1, x_2, x_3, x_4 and the knowledge of $C(M-1)$ allow us to know which permutations Algorithm 3 erased at this step. Moreover x_7 tells us the random choice made by Algorithm 3 and from x_7 and x_1 we can deduce P. For each s_i we know the orientation chosen for S_i at the step $M-1$ sends P onto s_i, from x_5 we deduce exactly in which order and from x_6 we get the rest of the orientation. Therefore we can deduce the set of chosen orientations before the bad event occurred. With the sets $S(z)$ and $C(z)$ known for all $z \le M$ we can easily deduce e. $\diamond$

The previous claim has the following corollary:

Corollary 99. *If $\mathcal{H}$ admits no p-equitable orientation, then Algorithm 3 defines an injection from the set of vectors $[cr!]^M$ into L^M.*

Let L_M be the set of all possible logs after M steps of Algorithm 3. To show Theorem 98 it suffices to show that, for M big enough, $|L_M|$ is strictly smaller than $(cr!)^M$.

Lemma 100. *For M big enough, $|L_M| < (cr!)^M$.*

Proof. We will compute a bound for $|L_M|$. R is a binary word of size $\le 2M$, and there are at most 4^M such words. X is a list of 7-tuples. As Algorithm 3 made M choices and each bad event removes r_1 of those, there exist at most $\frac{M}{r_1}$ bad events. Moreover, for each 7-tuple, $(x_1, x_2, x_3, x_4, x_5, x_6, x_7)$ we have $x_1 \le \binom{r}{p}$, $x_2 \le k$, $x_3 \le \binom{\binom{r}{p}}{r_1-1}$, $x_4 \le k^{r_1-1}$, $x_5 \le p!^{r_1-1}$, $x_6 \le (r-p)!^{r_1-1}$, $x_7 \le r!$. Using the bounds $\binom{n}{k} \le (\frac{n \cdot e}{k})^k$ or $\binom{n}{k} \le n^k$ we get the following bound.

$$
\begin{aligned}
|X| &\le \left(r^p \cdot k \cdot \left(\frac{r^p \cdot e}{r_1 - 1} \right)^{r_1 - 1} \cdot (k \cdot p! \cdot (r-p)!)^{r_1-1} \cdot r! \right)^{M/r_1} \\
&\le \frac{(r! \cdot (r^p)^{r_1} \cdot (r-p)!^{r_1-1})^{M/r_1} \cdot (k \cdot e \cdot p!)^M}{(r_1-1)^{M(r_1-1)/r_1}} \\
&\le \left[r^p \cdot r!^{r_1} \cdot \left(\frac{r^p}{r(r-1)\dots(r-p+1)} \right)^{r_1-1} \right]^{M/r_1} \cdot \left(\frac{k \cdot e \cdot p!}{(r_1-1)^{(r_1-1)/r_1}} \right)^M
\end{aligned}
$$

We can assume $r > 2p$, and so $\frac{r}{r-p+1} < 2$:

$$
|X| \le r!^M \cdot \left(r^{p/r_1} \cdot 2^p \cdot \frac{k \cdot e \cdot p!}{(r_1-1)^{(r_1-1)/r_1}} \right)^M
$$

As $|F| \le (r!+1)^{|\mathcal{H}|}$ and $|L_M| \le |F||X||R|$ we get the following bound on $|L_M|$:

$$|L_M| \le r!^M \cdot \left(4 \cdot r^{p/r_1} \cdot 2^p \cdot \frac{k \cdot e \cdot p!}{(r_1-1)^{(r_1-1)/r_1}}\right)^M \cdot (r!+1)^{|\mathcal{H}|}$$

□

5.1.2 Derangements

This subsection is devoted to the proof of Lemma 55. The main ingredient is the following lemma from Erdős and Spencer [24]:

Lemma 101 (Lopsided Lovász Local Lemma)**.** *Let $A_1, \ldots, A_m$ be events in a probability space, each with probability at most p. Let G be a graph defined on those events such that for every A_i, and for every set S avoiding both A_i and its neighbours, the following relation holds:*

$$Pr[A_i | \bigwedge_{A_j \in S} \bar{A}_j] \le P[A_i]$$

Then if $4dp \le 1$, all the events can be avoided simultaneously:

$$Pr[\bar{A}_1 \wedge \cdots \wedge \bar{A}_m] \ge (1-2p)^m > 0$$

Thanks to this result we can prove the following, which can be seen as a generalization of the fact that a random permutation of n points is a derangement with asymptotic probability $1/e$.

Lemma 55. *Let $p, k \in \mathbb{N}$ and $\alpha < 1$ be fixed. Let X be a set of cardinality r and let $\mathcal{L}_S$ be, for every $S \in \binom{X}{p}$, a collection of p-subsets of X with $|\mathcal{L}_S| \le k$. Then, if no p-subset occurs in more than r^α of the $\mathcal{L}_S$, a random permutation σ of X satisfies $\sigma(S) \notin \mathcal{L}_S$ for every S with probability $\ge (1 - 2k/\binom{r}{p})^{\binom{r}{p}} = e^{-2k} + o(1)$ when r grows large.*

Proof. For every $S \in \binom{X}{p}$, we define the *bad event* B_S with:

$$B_S = \bigvee_{S' \in \mathcal{L}_S} [\sigma(S) = S']$$

Each B_S has a probability $Pr[B_S] \le k/\binom{r}{p}$. On these bad events we define a lopsidependency graph (see [24]) G_B with the following adjacencies:

$$\left\{(B_{S_1}, B_{S_2}) : S_1, S_2 \in \binom{X}{p} \text{ s.t. } \left[S_1 \bigcup \mathcal{L}_{S_1}\right] \bigcap \left[S_2 \bigcup \mathcal{L}_{S_2}\right] \neq \emptyset\right\}$$

As a p-subset of X intersects at most $O(r^{p-1})$ others, and noting that every p-subset can occur at most r^α times, we have that:

$$\Delta(G_B) \leq (k+1)r^\alpha \times O(r^{p-1}) = o(r^p)$$

In order to apply the Lopsided Lovász Local Lemma to the events B_S and graph G_B, we must ensure for every $S \in \binom{X}{p}$ and $S_B \subseteq V(G_B)\backslash N_{G_B}[B_S]$ that:

$$Pr[B_S | \bigwedge_{B_{S'} \in S_B} \bar{B}_{S'}] \leq Pr[B_S] \tag{5.1}$$

Indeed, if we denote by T (for *trace*) the number of elements of $\bigcup_{B_{S'} \in S_B} S'$ sent by the random permutation σ into $\bigcup \mathcal{L}_S$:

$$Pr[B_S] = \sum_t Pr[B_S \mid T = t] Pr[T = t]$$

$$Pr[B_S | \bigwedge_{B_{S'} \in S_B} \bar{B}_{S'}] = \sum_t Pr[B_S \mid T = t, \bigwedge_{B_{S'} \in S_B} \bar{B}_{S'}] Pr[T = t \mid \bigwedge_{B_{S'} \in S_B} \bar{B}_{S'}]$$

As $\bigcup \mathcal{L}_S$ is disjoint from the $\bigcup \mathcal{L}_{S'}, \forall B_{S'} \in S_B$, we have:

$$Pr[B_S \mid T = t, \bigwedge_{B_{S'} \in S_B} \bar{B}_{S'}] = Pr[B_S \mid T = t]$$

And thus:

$$Pr[B_S | \bigwedge_{B_{S'} \in S_B} \bar{B}_{S'}] = \sum_t Pr[B_S \mid T = t] Pr[T = t \mid \bigwedge_{B_{S'} \in S_B} \bar{B}_{S'}]$$

In order to prove (5.1), we will first need the following observation:

Claim 101.1. *$Pr[B_S \mid T = t]$ is a decreasing function of t.*

Proof of the claim. We compute the value of $Pr[B_S \mid T = t]$ exactly, denoting by $r' \leq r$ the cardinality of $\bigcup_{B_{S'} \in S_B} S'$. It is equal to 0 when $t > r' - p$, and is otherwise equal to:

$$\begin{aligned}
Pr[B_S \mid T = t] &= \sum_{S' \in \mathcal{L}_S} Pr[\sigma(S) = S' \mid T = t] \\
&= \frac{|\mathcal{L}_S|}{\binom{r-t}{p}} \frac{\binom{r'-p}{t}}{\binom{r'}{t}} \\
&= \left(|\mathcal{L}_S| \frac{(r'-p)!p!}{r'!} \right) \left(\frac{(r-p-t)!(r'-t)!}{(r'-p-t)!(r-t)!} \right) \\
&= Pr[B_S \mid T = t-1] \left(\frac{(r'-p-t+1)}{(r-p-t+1)} \frac{(r-t+1)}{(r'-t+1)} \right) \\
&\leq Pr[B_S \mid T = t-1]
\end{aligned}$$

◇

Additionally, we will prove a relationship on the members of $\sum_t Pr[T = t]$ and on those of $\sum_t Pr[T = t \mid \bigwedge_{B_{S'} \in S_B} \bar{B}_{S'}]$, which both sum to 1.

Claim 101.2. *If $Pr[T = t \mid \bigwedge_{B_{S'} \in S_B} \bar{B}_{S'}]$ is nonzero, then*

$$\frac{Pr[T = t + 1]}{Pr[T = t]} \leq \frac{Pr[T = t + 1 | \bigwedge_{B_{S'} \in S_B} \bar{B}_{S'}]}{Pr[T = t \mid \bigwedge_{B_{S'} \in S_B} \bar{B}_{S'}]}$$

Proof of the claim. According to Bayes' Theorem applied to the right side of the equation,

$$\frac{Pr[T = t + 1 | \bigwedge_{B_{S'} \in S_B} \bar{B}_{S'}]}{Pr[T = t \mid \bigwedge_{B_{S'} \in S_B} \bar{B}_{S'}]} = \frac{Pr[\bigwedge_{B_{S'} \in S_B} \bar{B}_{S'} \mid T = t + 1] Pr[T = t + 1]}{Pr[\bigwedge_{B_{S'} \in S_B} \bar{B}_{S'} \mid T = t] Pr[T = t]}$$

We thus only need to ensure the following, which is a consequence of Lemma 102:

$$Pr[\bigwedge_{B_{S'} \in S_B} \bar{B}_{S'} \mid T = t + 1] \geq Pr[\bigwedge_{B_{S'} \in S_B} \bar{B}_{S'} \mid T = t]$$

$\diamond$

We are now ready to prove (5.1), and we define d_t for every t where $Pr[T = t]$ is nonzero:

$$d_t = Pr[T = t] - Pr[T = t \mid \bigwedge_{B_{S'} \in S_B} \bar{B}_{S'}]$$

Because d_t is a difference of probability distributions the sum $\sum_t d_t$ is null, and we can rewrite (5.1) using d_t:

$$\begin{aligned} 0 &\leq Pr[B_S] - Pr[B_S | \bigwedge_{B_{S'} \in S_B} \bar{B}_{S'}] \\ &\leq \sum_t Pr[B_S \mid T = t] Pr[T = t] - \sum_t Pr[B_S \mid T = t] Pr[T = t \mid \bigwedge_{B_{S'} \in S_B} \bar{B}_{S'}] \\ &\leq \sum_t d_t Pr[B_S \mid T = t] \end{aligned}$$

We will thus prove that the sum $\sum_t d_t Pr[B_S \mid T = t]$ is nonnegative. It is a consequence of Claim 101.2 that all nonnegative values of d_t appear before all nonpositive ones, and so that there is a t_0 such that $d_t \geq 0$ if and only if $t \leq t_0$. As a result, $|\sum_{t \leq t_0} d_t| = |\sum_{t > t_0} d_t| = \frac{1}{2}$ and we can write:

$$\begin{aligned} \sum_t d_t Pr[B_S \mid T = t] &= \sum_{t \leq t_0} d_t Pr[B_S \mid T = t] + \sum_{t > t_0} d_t Pr[B_S \mid T = t] \\ &\geq \frac{1}{2} Pr[B_S \mid T = t_0] - \frac{1}{2} Pr[B_S \mid T = t_0 + 1] \geq 0 \quad \text{(by Claim 101.1)} \end{aligned}$$

The second hypothesis of Lemma 101 is that $4pd \leq 1$, which translates in our case to $4\frac{k}{\binom{r}{p}} o(r^p) = o(1)$ and is thus satisfied when r grows large. Hence, we have that:

$$Pr[\bigwedge_S \bar{B}_S] \geq \left[1 - 2k / \binom{r}{p}\right]^{\binom{r}{p}} = e^{-2k} + o(1)$$

$\square$

Lemma 102. *Let A, B be two sets of size r, and let $\sigma : A \mapsto B$ be a random bijection. For every $A_1, \ldots, A_k \subseteq A' \subseteq A$ and $B_1, \ldots, B_k \subseteq B' \subseteq B$, the following function increases with t.*

$$Pr\left[\bigwedge_i [\sigma(A_i) \neq B_i] \;\middle|\; \sigma(A')\backslash B' \text{ has cardinality } t\right] \tag{5.2}$$

Proof. We implicitly assume in this proof that the conditionning event has a nonzero probability for t and $t+1$. Let S_1, S_2 be two sets of cardinality $|A'|$ with symmetric difference $S_1 \Delta S_2 = \{x, y\}$ where $x \in S_2$ is an element of $B\backslash B'$. Let σ_{xy} be the permutation transposing x and y. Then,

$$\begin{aligned} Pr\left[\bigwedge_i [\sigma(A_i) \neq B_i] \;\middle|\; \sigma(A') = S_1\right] &\leq Pr\left[\bigwedge_i [\sigma_{xy}\sigma(A_i) \neq B_i] \;\middle|\; \sigma(A') = S_1\right] \\ &= Pr\left[\bigwedge_i [\sigma(A_i) \neq B_i] \;\middle|\; \sigma(A') = S_2\right] \end{aligned}$$

We are now ready to derive the result:

$$(5.2) = \frac{1}{\binom{|B\backslash B'|}{t}\binom{|B'|}{|A'|-t}} \sum_{\substack{S \subseteq B \\ |S|=|A'| \\ |S\backslash B'|=t}} Pr\left[\bigwedge_i [\sigma(A_i) \neq B_i] \;\middle|\; \sigma(A') = S\right]$$

Using our previous remark, we find an upper bound on the last term of the equation by averaging it over sets S' obtained from S by the exchange of two elements:

$$\begin{aligned} (5.2) &\leq \frac{1}{\binom{|B\backslash B'|}{t}\binom{|B'|}{|A'|-t}} \sum_{\substack{S \subseteq B \\ |S|=|A'| \\ |S\backslash B'|=t}} \frac{1}{(|B\backslash B'| - t)(|A'| - t)} \sum_{\substack{S' \subseteq B \\ |S'|=|A'| \\ |S'\backslash B'|=t+1 \\ |S \Delta S'|=2}} Pr\left[\bigwedge_i [\sigma(A_i) \neq B_i] \;\middle|\; \sigma(A') = S'\right] \\ &= \frac{1}{\binom{|B\backslash B'|}{t}\binom{|B'|}{|A'|-t}} \frac{(t+1)(|B'| - |A'| + t + 1)}{(|B\backslash B'| - t)(|A'| - t)} \sum_{\substack{S' \subseteq B \\ |S'|=|A'| \\ |S'\backslash B'|=t+1}} Pr\left[\bigwedge_i [\sigma(A_i) \neq B_i] \;\middle|\; \sigma(A') = S'\right] \\ &= \frac{\binom{|B\backslash B'|}{t+1}\binom{|B'|}{|A'|-t-1}}{\binom{|B\backslash B'|}{t}\binom{|B'|}{|A'|-t}} \frac{(t+1)(|B'| - |A'| + t + 1)}{(|B\backslash B'| - t)(|A'| - t)} Pr\left[\bigwedge_i [\sigma(A_i) \neq B_i] \;\middle|\; \sigma(A')\backslash B' \text{ has cardinality } t+1\right] \\ &= Pr\left[\bigwedge_i [\sigma(A_i) \neq B_i] \;\middle|\; \sigma(A')\backslash B' \text{ has cardinality } t+1\right] \end{aligned}$$

□

5.2 AVD-colouring

In this section, we present the proof of Theorem 57 obtained in [37] with Gwenaël Joret.

In Subsection 5.2.1 we set up the plan for the proof of Theorem 57. In particular, vertices with big degrees (at least $\Delta/2$, roughly) and those with small degrees are treated independently, and differently. Then, Subsection 5.2.2 and Subsection 5.2.3 are devoted to handling big and small degree vertices, respectively.

5.2.1 Initial colouring

Fix some ϵ with $0 < \epsilon < 1/2$, which will be assumed small enough later in the proof. Let G be a graph with maximum degree Δ and no isolated edge. We assume Δ is large enough for various inequalities appearing in the proof to hold.

The beginning of our proof of Theorem 57 follows closely that of Hatami [35]. In particular, we reuse his approach of treating differently vertices with 'small' degrees and those with 'big' degrees, except we use $(1/2 - \epsilon)\Delta$ as the threshold instead of $\Delta/3$ in [35]. This larger threshold helps a little bit in reducing the additive constant in the main theorem; however, the bulk of the reduction from 300 to 19 comes from treating big degree vertices differently.

Let $d := \lceil (1/2 - \epsilon)\Delta \rceil$. Taking Δ large enough, we may assume that $d < \Delta/2$. We begin as in [35] by modifying the graph G as follows. Let G' be the *multigraph* obtained from G by contracting each edge $uv \in E(G)$ such that $d_G(u) < d$ and $d_G(v) < d$ but neither u nor v has any other neighbour w with $d_G(w) < d$. Then G' has maximum degree Δ and maximum edge multiplicity at most 2. Every proper edge colouring c' of G' can be extended to a proper edge colouring c of G with the same set of colours as follows: For each edge $e \in E(G)$ appearing in G', set $c(e) := c'(e)$. For each edge $uv \in E(G)$ that was contracted, we know that $d_G(u) + d_G(v) < \Delta$. Thus some colour α of c' is not used on any of the edges incident to u and v, set then $c(uv) := \alpha$.

In [35], the author points out that if moreover c' is an AVD-colouring of G' then c is an AVD-colouring of G. Using this observation, the proof in [35] then focuses on finding an AVD-colouring of G'. This is done by starting with a proper edge colouring c' with $\Delta + 2$ colours, which exists by Vizing's theorem, and then recolouring some edges of G' with new colours to obtain an AVD-colouring of G'. The advantage of working in G' instead of G is that the subgraph of G' induced by the vertices with degree strictly less than d has no isolated edge, which is important in that proof.

In our proof, we follow a similar approach but we keep the focus on G: We start with a proper edge colouring c' of G' with $\Delta + 2$ colours obtained from Vizing's theorem and extend it to an edge colouring c of G as in the above remark. Thus, the edge colouring c uses $\Delta + 2$ colours and satisfies the following property:

$$S_c(u) \cap S_c(v) = \{c(uv)\} \quad \forall uv \in E(G) \text{ s.t.} \{w \in N(u) \cup N(v) : d_G(w) < d\} = \{u, v\}. \quad (5.3)$$

Then, we modify c to obtain an AVD-colouring of G. Thus G' is only used to produce the initial edge colouring c of G. One advantage of working in G is that we avoid having to deal with parallel edges, which would introduce (trivial but annoying) technicalities in our approach. On the other hand, a small price to pay compared to [35] is that we will have to watch out for these edges uv such that $\{w \in N(u) \cup N(v) : d_G(w) < d\} = \{u, v\}$ in our proof.

Say that a vertex $u \in V(G)$ is *small* if $d_G(u) < d$, and *big* otherwise. Let A and B be the sets of small and big vertices of G, respectively. Our goal is to transform the edge colouring c into an AVD-colouring of G. The plan for doing so is roughly as follows. First we show that we can uncolour a bounded number of edges per big vertex in such a way that edges uv with $S_c(u) = S_c(v)$ and $u, v \in B$ that remain form a matching satisfying some specific properties. Then we show how we can recolour these uncoloured edges, plus a few other edges of G, to obtain an edge colouring where every edge uv with $u, v \in B$ satisfies $S_c(u) \neq S_c(v)$. Finally, we recolour edges with both endpoints in A in such a way that the resulting edge colouring is an AVD-colouring of G.

5.2.2 Big vertices

For each vertex $u \in B$, choose an arbitrary subset $N^+(u)$ of $N(u)$ of size d. We use a randomised algorithm, Algorithm 4, to select a subset $U^+(u) \subseteq N^+(u)$ of size 2 for each vertex $u \in B$. For each vertex $v \in V(G)$, we let $U^-(v) := \{u \in B : v \in U^+(u)\}$. Algorithm 4 chooses the subsets $U^+(u)$ iteratively, one big vertex u at a time. Hence, we see the sets $U^+(u)$ as variables, and the sets $U^-(v)$ ($v \in V(G)$) as being determined by these variables. (For definiteness, we set $U^+(u) := \emptyset$ for every small vertex u.) Just after choosing the subset $U^+(u)$ of a big vertex u, the algorithm checks whether this choice triggered any 'bad event'. If so, the bad event is handled, which involves *resetting* the variable $U^+(u)$, which means setting $U^+(u) := \emptyset$, and possibly resetting other variables $U^+(v)$ for some well-chosen big vertices v close to u in G.

Thanks to these bad events, the selected subsets satisfy a number of properties. A key property is that $|U^-(v)| \leq q$ for every $v \in V(G)$, with $q := 13$ being the constant that is optimised in this proof.

At any time during the execution of the algorithm, we say that an edge $uv \in E(G)$ is *selected* if $v \in U^+(u)$ or $u \in U^+(v)$. In the algorithm, we will make sure that if $v \in U^+(u)$ then $u \notin U^+(v)$ (that is, an edge can be selected 'at most once').

After the algorithm terminates, selected edges will be used to fix locally the edge colouring c for big vertices: The plan is to recolour them using $q + 3$ new colours, and then recolour a well-chosen matching of G with yet another new colour, in such a way that at the end $S_c(u) \neq S_c(v)$ holds for all edges $uv \in E(G)$ with $u, v \in B$. The resulting edge colouring of G will use $\Delta + q + 6$ colours in total.

Let us explain the conventions and terminology used in Algorithm 4. First, we assume that the vertices of G are ordered according to some fixed arbitrary ordering. This naturally induces an ordering of each subset of $V(G)$ as well, of each set of pairs of vertices (say using lexicographic ordering), and more generally of any set of structures built using vertices of G. This is used implicitly in what follows.

We say that an edge uv linking two big vertices is *finished* if $U^+(w) \neq \emptyset$ for each big vertex w in $N(u) \cup N(v)$ (note that this set includes u and v). For $u \in V(G)$, define $S'(u)$ as the set $S_c(u)$ minus colours of edges incident to u that are selected. That is,

$$S'(u) := \{c(uv) : v \in N(u), v \notin U^+(u) \cup U^-(u)\}.$$

In the algorithm, we check whether $S'(u) = S'(v)$ for two big vertices u, v with $d_G(u) = d_G(v)$ only when uv is finished, the idea being that if there are still big vertices adjacent to u or v waiting to be treated then this could potentially impact the sets $S'(u)$ and $S'(v)$. We say that an edge $uv \in E(G)$ is *bad* if $u, v \in B$, uv is finished, $d_G(u) = d_G(v)$, and $S'(u) = S'(v)$.

As mentioned earlier, we need to watch out for edges uv such that $d_G(u) = d_G(v) < d$ and $d_G(w) \geq d$ for all $w \in (N(u) \cup N(v)) - \{u, v\}$. This is because every edge incident to u or v distinct from uv is incident to a big vertex, and all these edges will have a fixed colour when we are done dealing with big vertices. Indeed, in the next section we only recolour edges in the subgraph induced by small vertices. Thus, if u and v were to see the same set of colours at the end of this step, we would have no way to fix this later. Note that at the beginning u and v see disjoint sets of colours in the edge colouring c except for colour $c(uv)$. Once the algorithm terminates, we will recolour at most $q + 2$ edges incident to each big vertex w (with new colours). Thus, if $d_G(u) = d_G(v) \geq q + 4$, we know that there will be at least one edge e incident to u distinct from uv that kept its original colour $c(e)$. Since v does not see the colour $c(e)$, we are then assured that u and v see different sets of colours after the recolouring step. Therefore, it is only when $d_G(u) = d_G(v) \leq q+3$ that we need to be careful when selecting edges incident to u or v to recolour. Let us call such edges *fragile* edges, i.e. uv is fragile if $d_G(u) = d_G(v) \leq q + 3$ and $d_G(w) \geq d$ for all $w \in (N(u) \cup N(v)) - \{u, v\}$. Fragile edges will be carefully handled in the algorithm.

As mentioned, the algorithm considers remaining big vertices u with $U^+(u) = \emptyset$ one by one, selects the subset $U^+(u)$ randomly each time, and deals with any bad event that may occur. Let us explain how the random choices are made. Given a big vertex u with $U^+(u) = \emptyset$, an unordered pair $\{v, w\} \subseteq N^+(u)$ is *admissible* for u if the following three conditions are satisfied:

- $v, w \notin U^-(u)$;
- if $vw \in E(G)$ then vw is not fragile, and
- setting $U^+(u) := \{v, w\}$ does not create any bad edge incident to u.

At the beginning of the `while` loop, the algorithm chooses an admissible pair for the vertex $u \in B$ under consideration uniformly at random among the first $s := \binom{d-q}{2} - 3d$ admissible pairs. Lemma 103 below shows that there are always at least s such pairs, thus this random choice can always be made.

Five types of bad events are considered in the algorithm. They correspond to the five conditions tested by the `if / else if` statements; we refer to them as Bad Event 1, Bad Event 2, etc. in order. These events state the existence of certain structures in the graph. We remark that there could be more than one instance of the structure under consideration

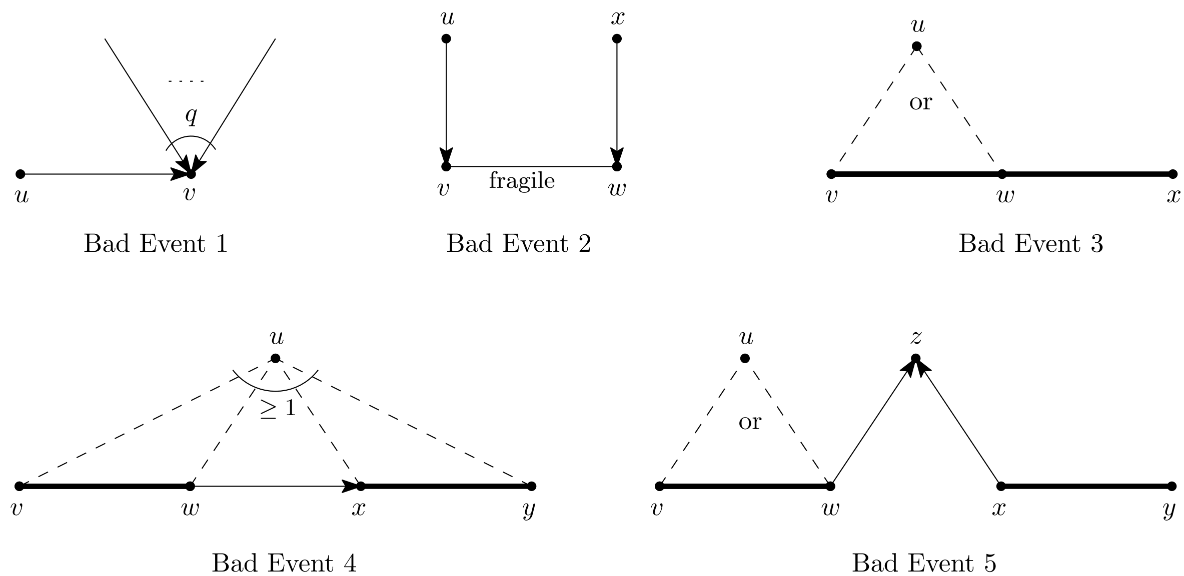

Figure 5.1: The five types of bad events in Algorithm 4. Bad edges are drawn in bold. (Note that possibly $z = u$ in Bad Event 5.)

in the graph. (For instance, there could be two vertices $v \in N(u)$ with $|U^-(v)| = q + 1$ in Bad Event 1.) In this case, we assume that the algorithm chooses one according to some deterministic rule. For the convenience of the reader, the five types of bad events considered are illustrated in Figure 5.1. Let us emphasise that if any bad event is triggered, then the current vertex u is always reset (i.e. the algorithm sets $U^+(u) := \emptyset$). This will ensures that no other bad event remains in the graph after dealing with the bad event under consideration.

The following lemma establishes some key properties of Algorithm 4. Note that by an *invariant* of the `while` loop, we mean a property that is true every time the condition of the loop is being tested. Thus, such a property holds when a new iteration of the loop starts, and also when the loop (and thus the algorithm) stops.

Lemma 103. *The following properties are invariants of the* `while` *loop in Algorithm 4:*

1. $|U^-(v)| \leq q$ *for every* $v \in V(G)$.
2. *At least one of* $U^-(v), U^-(w)$ *is empty for each fragile edge* vw.
3. *Bad edges form a matching.*
4. *If* $w, x \in B$ *belong to distinct bad edges, then* $(\{w\} \cup U^+(w)) \cap (\{x\} \cup U^+(x)) = \emptyset$.
5. *Every* $u \in B$ *with* $U^+(u) = \emptyset$ *has at least* s *admissible pairs.*

Proof. Let us start with property (1). Clearly, $|U^-(v)| \leq q$ for every $v \in V(G)$ the first time the condition of the `while` loop is being tested. This remains true for every subsequent test

Algorithm 4: Uncolouring some edges incident to big vertices.

```
U⁺(u) ← ∅   ∀u ∈ B
while ∃v ∈ B with U⁺(v) = ∅ do
    u ← first such vertex
    U⁺(u) ← admissible pair chosen uniformly at random among first s ones
    if ∃v ∈ N(u) with |U⁻(v)| = q + 1 then
        U⁺(w) ← ∅   ∀w ∈ U⁻(v)
    else if ∃ fragile edge vw and x ∈ V(G) − {u, v, w} s.t. u ∈ U⁻(v) and
      x ∈ U⁻(w) then
        U⁺(a) ← ∅   ∀a ∈ {u, x}
    else if ∃ distinct bad edges vw, wx with u ∈ N(v) ∪ N(w) then
        U⁺(a) ← ∅   ∀a ∈ {u, v, x}
    else if ∃ two independent bad edges vw, xy with u ∈ N(v) ∪ N(w) ∪ N(x) ∪ N(y)
      and x ∈ U⁺(w) then
        U⁺(a) ← ∅   ∀a ∈ {u, v, w, y}
    else if ∃ two independent bad edges vw, xy with u ∈ N(v) ∪ N(w), and
      ∃z ∈ V(G) − {v, w, x, y} with w, x ∈ U⁻(z) then
        U⁺(a) ← ∅   ∀a ∈ {u, v, w, x, y}
```

of the condition, thanks to Bad Event 1: Selecting the subset $U^+(u)$ for a vertex $u \in B$ could create up to two vertices v with $|U^-(v)| = q+1$ but these get fixed immediately when u is reset. Hence, (1) is an invariant of the loop.

Similarly, it is clear that property (2) is an invariant of the `while` loop, thanks to Bad Event 2.

Let us consider property (3). The property is true at the beginning of the algorithm, since there are no bad edges. Next, suppose that property (3) held true at the beginning of the loop but that there are two incident bad edges e, f just after selecting the admissible pair $U^+(u)$ for a big vertex u. Then at least one of e, f, say e, became bad just after treating u. Note that e cannot be incident to u, by definition of admissible pairs. Thus e is at distance 1 from u. It suffices to show that some bad event is triggered, since then u is reset and e is no longer bad (since e is not finished). This is clearly true, since either Bad Event 1 or Bad Event 2 is triggered, and if not then Bad Event 3 is triggered for sure (because of the existence of the pair e, f). Thus we see that property (3) is an invariant of the loop.

The proof for property (4) is similar. The property clearly holds at the beginning of the algorithm. Next, suppose that it held true at the beginning of the loop but that just after selecting the subset $U^+(u)$ for a big vertex u, there are two independent bad edges vw, xy s.t. $\{w\} \cup U^+(w)$ and $\{x\} \cup U^+(x)$ intersect. Then at least one of the two edges, say vw, is at distance exactly 1 from u. (Recall that u cannot be incident to either of the two edges, by definition of admissible pairs.) As before, it suffices to show that some bad event occurs, since then u is reset and vw is no longer bad. Let $z \in (\{w\} \cup U^+(w)) \cap (\{x\} \cup U^+(x))$ and say none of the first four bad events happens. Then $z \notin \{v, w, x, y\}$, since otherwise Bad

Event 4 would have been triggered. But this shows that Bad Event 5 occurs. We deduce that property (4) is maintained.

Finally, it remains to show that property (5) is an invariant of the loop. Consider thus any vertex $u \in B$ with $U^+(u) = \emptyset$ when the condition of the loop is being tested. (Thus, a new iteration of the loop is about to start.) From invariant (1), we know that there are at least $\binom{d-q}{2}$ unordered pairs of distinct vertices in $N^+(u) - U^-(u)$. Next, a key observation is that for every $x \in N(u)$, if there exists $\{v, w\} \subseteq N^+(u) - U^-(u)$ s.t. setting $U^+(u) := \{v, w\}$ makes the edge ux bad, then the set $\{v, w\}$ is uniquely determined. Hence, potential bad edges forbid at most $|N(u)| \leq \Delta$ pairs of vertices in $N^+(u) - U^-(u)$. Finally, among the remaining pairs $\{v, w\}$, at most $\lfloor d/2 \rfloor$ of them are s.t. vw is a fragile edge. Therefore, we conclude that there are at least $\binom{d-q}{2} - \Delta - \lfloor d/2 \rfloor \geq \binom{d-q}{2} - 3d = s$ admissible pairs for u. □

The properties listed in Lemma 103 hold in particular when Algorithm 4 stops. However, it is not clear at first sight that the algorithm should ever stop. Our next result shows that it does so with high probability. For simplicity, we sometimes call one iteration of the `while` loop a *step*.

Theorem 104. *The probability that Algorithm 4 stops in at most t steps tends to* 1 *as* $t \to \infty$.

We use an 'entropy compression' argument to prove this theorem, a proof method introduced by Moser and Tardos [54] in their celebrated algorithmic proof of the Lovász Local Lemma. In a nutshell, the main idea of the proof is to look at sequences of t random choices such that the algorithm does not stop in at most t steps. Exploiting the fact that the algorithm did not stop, we show how one can get an implicit lossless encoding of these sequences, by writing down a concise log of the execution of the algorithm. Then, looking at the structure of the algorithm, we prove that there are only $o(s^t)$ such logs. Since in total there are s^t random sequences of length t, we deduce that only a $o(1)$-fraction of these make the algorithm run for at least t steps. Theorem 104 follows.

To describe the log of an execution of the algorithm, we need the following definitions. First, recall that a *Dyck word* of semilength k is a binary word $w_1 w_2 \dots w_{2k}$ with exactly k 0s and k 1s such that the number of 0s is at least the number of 1s in every prefix of the word. A *descent* in a Dyck word is a maximal sequence of consecutive 1s, its *length* is the number of 1s.

For our purposes, it will be more convenient to drop the requirement that a Dyck word has the same number of 0s and 1s. Let us define a *partial Dyck word* of semilength k as a binary word $w_1 w_2 \dots w_p$ with exactly k 0s and *at most* k 1s such that the number of 0s is at least the number of 1s in every prefix of the word. Descents are defined in the same way as for normal Dyck words.

Let us consider a sequence $(r_1, \dots, r_t)$ of t random choices such that Algorithm 4 does not stop in at most t steps when run with these random choices. In other words, the algorithm is about to start its $(t+1)$-th iteration of the `while` loop, at which point we freeze its execution.

Each random choice r_i consisted in choosing an admissible pair for some big vertex u among its first s admissible pairs, thus we see r_i as a number in $[s]$.

For each $i \in [t+1]$, let U_i^+ and U_i^- denote the functions U^+ and U^-, respectively, at the beginning of the i-th iteration, and let B_i denote the subset of vertices $u \in B$ with $U_i^+(u) = \emptyset$. We associate to the sequence $(r_1, \dots, r_t)$ a corresponding *log* $(W, \gamma, \delta, U_{t+1}^+)$, where W is a partial Dyck word of semilength t such that the length of each descent is in the set $\{2, 3, 4, 5, q+1\}$, and $\gamma = (\gamma_1, \dots, \gamma_t)$ and $\delta = (\delta_1, \dots, \delta_t)$ are two sequences of integers.

The partial Dyck word W is built as follows during the execution of the algorithm: Starting with the empty word, we add a 0 at the end of the word each time a big vertex is treated. If the corresponding random choice triggers a bad event, we moreover add ℓ 1s at the end of the word, where ℓ is the number of big vertices that are reset (so $\ell = q+1, 2, 3, 4, 5$ for bad events of types $1, 2, 3, 4, 5$, respectively). Thus descents in W are in bijection with bad events treated during the execution, and the length of a descent tells us the type of the corresponding bad event.

The two sequences γ and δ are defined as follows. For $i \in [t]$, the integers γ_i and δ_i encode information about the bad event handled during iteration i. If there was none, we simply set $\gamma_i := \delta_i := -1$. Otherwise, γ_i is a nonnegative integer encoding the set of big vertices that are reset when the bad event is handled, and δ_i is a nonnegative integer encoding extra information which will help us recover the random choice r_i from the log. The precise definitions of γ_i and δ_i depend on the type of the bad event (see the list below); however, before giving these definitions we must explain the assumptions we make.

The definition of γ_i assumes that the set B_i is known. In turn, γ_i will encode enough information to determine completely B_{i+1} from B_i. Since $B_1 = B$, it then follows that we can read off all the sets $B_1, B_2, \dots, B_{t+1}$ from the sequence γ: For $i = 1, \dots, t$, either $\gamma_i \geq 0$, in which case B_{i+1} is determined by B_i and γ_i. Or $\gamma_i = -1$, in which case no bad event occurred during iteration i, and thus $B_{i+1} := B_i - \{u\}$ where u is the first vertex in B_i.

As already mentioned, the purpose of the log is to encode all t random choices $r_1, \dots, r_t$ that have been made during the execution. To encode r_i ($i \in [t]$), we work *backwards*: We assume that the function U_{i+1}^+ is known, and we show that one can then deduce r_i and U_i^+ using the log. Since U_{t+1}^+ is part of the log, this implies that the log uniquely determines $r_t, r_{t-1}, \dots, r_1$, as desired. Let us remark that if no bad event occurred during the i-th iteration, then we can already deduce r_i and U_i^+ from U_{i+1}^+ using the sets B_i and B_{i+1}. Indeed, in this case $B_i = B_{i+1} \cup \{u\}$ where u is the vertex treated during the i-th iteration. Thus, for $v \in B$,

$$U_i^+(v) = \begin{cases} U_{i+1}^+(v) & \text{if } v \neq u \\ \emptyset & \text{if } v = u \end{cases}$$

Furthermore, $U_{i+1}^+(u)$ tells us what was the random choice r_i that was made for u during iteration i. Indeed, using U_i^+ we can deduce what was the set of admissible pairs for u at the beginning of iteration i. Then, r_i is the position of the pair $U_{i+1}^+(u)$ in the ordering of these admissible pairs. Therefore, it is only when a bad event happens that we need extra information to determine r_i and U_i^+. This is precisely the role of δ_i.

Definitions of γ and δ. Let $i \in [t]$. If no bad event occurred during iteration i, set

$\gamma_i := -1$ and $\delta_i := -1$. Otherwise, say that a bad event β of type j was handled. The definition of γ_i assumes that B_i is known, while that of δ_i assumes that B_i and U^+_{i+1} are both known. In particular, we know the vertex u treated at the beginning of the iteration, since it is the first vertex in B_i. With these remarks in mind, γ_i and δ_i are defined as follows:

$j = 1$ The bad event β was triggered because the admissible pair chosen for u contained a vertex v with $|U^-_i(v)| = q$. Vertex u and the q vertices in $U^-_i(v)$ were subsequently reset. There are at most d choices for vertex v and at most $\binom{\Delta}{q}$ choices for $U^-_i(v)$. We may thus encode v and $U^-_i(v)$ with a number $\gamma_i \in \left[d\binom{\Delta}{q}\right]$. Observe that $B_{i+1} = B_i \cup U^-_i(v)$.

Now that v and $U^-_i(v)$ are identified, we want to encode the admissible pair $\{v, x\}$ that was chosen for u at the beginning of the iteration, and the sets $U^+_i(w)$ for each vertex $w \in U^-_i(v)$. There are at most d choices for x, and similarly for each $w \in U^-_i(v)$ there are at most d choices for the vertex in $U^+_i(w)$ which is distinct from v. We let $\delta_i \in [d^{q+1}]$ encode these choices. Since U^+_i only differs from U^+_{i+1} on vertices $w \in U^-_i(v)$, with the encoded information we can deduce U^+_i from U^+_{i+1}. Note also that r_i is determined by the admissible pair $\{v, x\}$ that was chosen for u.

$j = 2$ The bad event β was triggered because the admissible pair chosen for u contained a vertex v incident to a fragile edge vw with $U^-_i(w) \neq \emptyset$. Then two vertices were reset, namely u and some vertex x in $U^-_i(w)$. There are at most d choices for vertex v. Once v is identified, we know vertex w since fragile edges form a matching. Finally, there are at most $q + 2$ choices for x, since $d_G(w) \leq q + 3$ and $x \neq v$. We let $\gamma_i \in [(q + 2)d]$ encode v, w, and x. Observe that $B_{i+1} = B_i \cup \{x\}$.

Next, to encode r_i we only need to specify the vertex in the admissible pair chosen for u that is distinct from v (d choices). Similarly, there are at most d possibilities for the set $U^+_i(x)$ since we know that it includes w. Hence, we can encode this information with a number $\delta_i \in [d^2]$. Note that, knowing x and $U^+_i(x)$, we can directly infer U^+_i from U^+_{i+1}, since $U^+_i(y) = U^+_{i+1}(y)$ for all $y \in B - \{x\}$.

$j = 3$ After selecting the admissible pair for u, we had $S'(v) = S'(w) = S'(x)$ for three distinct vertices $v, w, x \in B - \{u\}$ with $vw, wx \in E(G)$ and $u \in N(v) \cup N(w)$. Then u, v, x were reset. There are at most $2\Delta^3$ choices for the triple v, w, x (the factor 2 is due to the fact that u can be adjacent to v or w). We let $\gamma_i \in [2\Delta^3]$ encode v, w, x. Observe that $B_{i+1} = B_i \cup \{v, x\}$.

Knowing v, w, x and U^+_{i+1}, our next aim is to encode U^+_i and r_i using δ_i. First, we simply encode the admissible pair $\{u_1, u_2\}$ that was chosen for u during the i-th iteration explicitly, thus there are $\binom{d}{2}$ possibilities. Next, we observe that $U^+_i(y) = U^+_{i+1}(y)$

for every $y \in B - \{u, v, x\}$, and $U_i^+(u) = \emptyset$. Thus it only remains to encode $U_i^+(v)$ and $U_i^+(x)$. Here the idea is that, since at this point we know the set $U_i^+(w)$ and the admissible pair $\{u_1, u_2\}$ chosen for u, there are only $O(1)$ possibilities for the sets $U_i^+(v)$ and $U_i^+(x)$ in order to have that $S'(v) = S'(w) = S'(x)$ just before u, v, x were reset.

Let us focus on the set $U_i^+(v)$, the argument for $U_i^+(x)$ will be symmetric. First, let us write down the following local information: (1) Is $w \in U_i^+(v)$? (2) Is $w \in U_i^+(x)$? (3) Is $v \in U_i^+(x)$? Thus there are 8 possibilities. (1)–(2) gives enough information to reconstruct the set $S'(w)$ just before the resets, since we already know $U_i^+(w)$ and whether $w \in \{u_1, u_2\}$ or not. From (3) we also know the set $S''(v) := S'(v) - \{c(vv') : v' \in U_i^+(v)\}$ just before the resets, since we know whether $v \in \{u_1, u_2\}$ or not, and whether $v \in U_i^+(z)$ or not for every $z \in N(v) - \{u\}$. Now, it only remains to observe that $U_i^+(v)$ is determined by the two sets $S''(v)$ and $S'(w)$, namely $U_i^+(v) = \{v' \in N(v) : c(vv') \in S''(v) - S'(w)\}$.

Proceeding similary for the set $U_i^+(x)$ (8 possibilities again), this fully determines U_i^+. Now, given U_i^+ we know exactly the set of admissible pairs for u at the beginning of the i-th iteration. Since we know that the pair $\{u_1, u_2\}$ was chosen, we can deduce the value of r_i. Hence, this shows that U_i^+ and r_i can be encoded using a number $\delta_i \in [64\binom{d}{2}]$. (The constant 64 could be reduced with a more careful analysis but this would not make a difference later on.)

$j = 4$ Here we let $\gamma_i \in [4\Delta^4]$ encode the four vertices v, w, x, y as seen from u (the factor 4 comes from the fact that u is adjacent to at least one of them but we do not know which one). Since u, v, w, y are reset during this iteration, we have $B_{i+1} = B_i \cup \{v, w, y\}$.

Next, we set up δ_i to encode U_i^+ and r_i knowing U_{i+1}^+. As in the previous case, we encode the admissible pair $\{u_1, u_2\}$ that was chosen for u during the i-th iteration explicitly ($\binom{d}{2}$ choices). Once we know U_i^+, we know which are the admissible pairs for u at the beginning of the i-th iteration, and thus we can determine r_i, exactly as before. Thus, it only remains to encode $U_i^+(v), U_i^+(w)$, and $U_i^+(y)$.

Let us start with $U_i^+(w)$. We already know that $x \in U_i^+(w)$, and we encode the other vertex in $U_i^+(w)$ explicitly (d choices).

Next, consider $U_i^+(v)$. Here, the idea is the same as for Bad Event 3, namely once $U_i^+(w)$ is known there are only $O(1)$ possibilities for $U_i^+(v)$ to have that $S'(v) = S'(w)$ just before the resets. To be precise, we write down the following local information: (1) Is $w \in U_i^+(v)$? (2) Is $w \in U_i^+(x)$? (3) Is $w \in U_i^+(y)$? (4) Is $v \in U_i^+(x)$?

(5) Is $v \in U_i^+(y)$? Thus there are 32 possibilities. (1)–(3) gives us enough information to reconstruct the set $S'(w)$ just before the resets, since we already know $U_i^+(w)$ and whether $w \in \{u_1, u_2\}$ or not. Similarly, (4)–(5) allow us to determine the set $S''(v) := S'(v) - \{c(vv') : v' \in U_i^+(v)\}$ just before the resets, which in turn determines $U_i^+(v)$ since $U_i^+(v) = \{v' \in N(v) : c(vv') \in S''(v) - S'(w)\}$.

For $U_i^+(y)$, we proceed exactly as for $U_i^+(v)$, exchanging v with y and w with x. The only difference here is that x is not reset, thus $U_i^+(x) = U_{i+1}^+(x)$. We similarly conclude that there are at most 32 possibilities for the set $U_i^+(y)$. In summary, we may encode all the necessary information with a number $\delta_i \in \left[2^{10} d \binom{d}{2}\right]$.

$j = 5$: We let $\gamma_i \in [2\Delta^5]$ encode the vertices v, w, x, y, z. (Recall that possibly $z = u$.) Since u, v, w, x, y are reset during this iteration, we have $B_{i+1} = B_i \cup \{v, w, x, y\}$.

Next, we encode U_i^+ and r_i based on U_{i+1}^+. Again, we encode the admissible pair $\{u_1, u_2\}$ chosen for u explicitly ($\binom{d}{2}$ choices), which will determine r_i once we know U_i^+. It only remains to encode $U_i^+(v), U_i^+(w), U_i^+(x)$, $U_i^+(y)$.

Similarly as for Bad Event 4, there are most d possibilities for the set $U_i^+(w)$, since we already know that $z \in U_i^+(w)$. The same is true $U_i^+(x)$.

For $U_i^+(v)$ we proceed exactly as in the previous case, exploiting the fact that $U_i^+(w)$ is already encoded: Writing down which sets among $U_i^+(v), U_i^+(x), U_i^+(y)$ include vertex w, and similarly which of $U_i^+(x), U_i^+(y)$ include v, is enough to determine $U_i^+(v)$. Thus there are 32 choices. This is also true for $U_i^+(y)$ since the situation is completely symmetric (swapping v, w with y, x, respectively). Hence, we can record the desired information with a number $\delta_i \in \left[2^{10} d^2 \binom{d}{2}\right]$.

Let $\mathcal{R}_t$ denote the set of sequences $(r_1, \ldots, r_t)$ with each $r_i \in [s]$ such that Algorithm 4 does not stop in at most t steps when using $r_1, \ldots, r_t$ for the random choices. Also, let $\mathcal{L}_t$ denote the set of logs defined by the algorithm on these sequences. The following lemma follows from the discussion above.

Lemma 105. *For each $t \geq 1$, there is a bijection between the two sets $\mathcal{R}_t$ and $\mathcal{L}_t$.*

Next, we bound $|\mathcal{L}_t|$ from above when t is large. To do so we need to count some specific Dyck words where each descent is weighted with some integer: Given a set $E = \{(l_1, w_1), \ldots, (l_k, w_k)\}$ of couples of positive integers with all l_j's distinct, we let $C_{t,E}$ be the number of Dyck words of semilength t where each descent has length in the set $\{l_1, \ldots, l_k\}$, and each descent of length l_j is weighted with an integer in $[w_j]$.

For our purposes, we take $E := \{(l_1, w_1), \ldots, (l_5, w_5)\}$, where (l_j, w_j) is determined by the characteristics of Bad Event j: l_j is the number of vertices that are reset, and w_j is an upper bound on the number of values the corresponding pair (γ_i, δ_i) can take in the log during the corresponding i-th iteration of the algorithm. Thus, following the discussion of bad events above, we take:

- $l_1 = q+1$ and $w_1 = \binom{\Delta}{q} d^{q+2}$
- $l_2 = 2$ and $w_2 = (q+2)d^3$
- $l_3 = 3$ and $w_3 = 2^7 \Delta^3 \binom{d}{2}$
- $l_4 = 4$ and $w_4 = 2^{12} \Delta^4 d \binom{d}{2}$
- $l_5 = 5$ and $w_5 = 2^{11} \Delta^4 d^2 \binom{d}{2}$

In our logs we deal with *partial* Dyck words that are weighted as above. The difference between the number of 0s and 1s in the partial Dyck word corresponds to the number of big vertices $u \in B$ for which $U^+(u)$ is currently not set; we call this quantity its *defect*. Observe that partial weighted Dyck words of semilength t and defect k can be mapped injectively to weighted Dyck words of semilength $t+k$ by adding k occurrences of 011 at the end, where each of the k new descents of length 2 are weighted with, say, the number 1. Since $k \leq n = |V(G)|$, we obtain the following lemma.

Lemma 106. $|\mathcal{L}_t| \leq \sum_{k=0}^{n} C_{t+k,E}$.

In our setting, n and s are fixed while t varies; thus, to prove that $|\mathcal{L}_t| \in o\left(s^t\right)$, it is enough by the above lemma to show that $C_{t,E} \in o\left(s^t\right)$. In order to bound $C_{t,E}$ from above, we follow [25] and use a bijection between Dyck words and rooted plane trees.

Lemma 107. *The number $C_{t,E}$ is equal to the number of weighted rooted plane trees on $t+1$ vertices, where each vertex has a number of children in $E \cup \{0\}$, and for each $i \in [5]$ each vertex with l_i children is weighted with an integer in $[w_i]$ (leaves are not weighted).*

The proof of this lemma is essentially that of Lemma 7 in [25].

Now we use generating functions and the analytic method described e.g. in [21, Section 1.2]. Let

$$y(x) := \sum_{t \geq 1} C_{t,E} x^t$$

denote the generating function associated to our objects, and let

$$\phi(x) := 1 + \sum_{i=1}^{5} w_i x^{l_i}.$$

Then $y(x)$ satisfies $y(x) = x\phi(y(x))$. As noted in [21, Theorem 5] (see also [28, p.278, Proposition IV.5]), the following asymptotic bound holds for $C_{t,E}$.

Theorem 108. *Let R denote the radius of convergence of ϕ and suppose that $\lim_{x\to R^-} \frac{x\phi'(x)}{\phi(x)} > 1$. Then there exists a unique solution $\tau \in (0,R)$ of the equation $\tau\phi'(\tau) = \phi(\tau)$, and $C_{t,E} = O(\gamma^t)$, where $\gamma := \phi(\tau)/\tau$.*

The radius of convergence of our function ϕ is $R = \infty$, and $\lim_{x\to\infty} \frac{x\phi'(x)}{\phi(x)} > 1$, thus the theorem applies. For our purposes, it is not necessary to compute exactly τ, a good upper bound on $\gamma = \phi(\tau)/\tau$ will be enough. To obtain such an upper bound we use the following lemma.

Lemma 109. *For every $x \in (0,R)$, if $x\phi'(x)/\phi(x) < 1$ then $\phi(\tau)/\tau < \phi(x)/x$.*

Proof. As noted in [28, Note IV.46] the function $x\phi'(x)/\phi(x)$ is increasing on $(0,R)$. Thus, $x\phi'(x)/\phi(x) < 1$ if and only if $x < \tau$. Consider the function $x\phi'(x)/\phi(x)$ on $(0,\tau)$. Since $x\phi'(x)/\phi(x) < 1$, we have $x\phi'(x) - \phi(x) < 0$. Moreover, since $\frac{\partial}{\partial x}(\frac{\phi(x)}{x}) = \frac{x\phi'(x)-\phi(x)}{x^2}$, we see that $\frac{\phi(x)}{x}$ is decreasing on $(0,\tau)$. Hence, $\frac{\phi(x)}{x} > \frac{\phi(\tau)}{\tau}$. $\square$

Using these tools we can bound γ from above.

Lemma 110. *$\gamma < s$ when d is large enough.*

Proof. We will use Lemma 109. Let $\epsilon_1 > 0$ be fixed (at the end of the proof ϵ_1 will be taken small enough as a function of $q = 13$). Let

$$x := \left(\frac{1}{q(1+\epsilon_1)w_1}\right)^{1/(q+1)}.$$

We claim that $x\phi'(x)/\phi(x) < 1$ when d is large enough. First, let us give some intuition: If we counted only the subset of weighted Dyck words where each descent is of length $l_1 = q+1$ and is weighted with an integer in $[w_1]$, then the corresponding function ϕ would be $\phi(x) = 1 + w_1x^{q+1}$, and one would get $\tau = \left(\frac{1}{qw_1}\right)^{1/(q+1)}$. As it turns out, the value of τ for our function of ϕ tends to that one (from below) as $d \to \infty$, hence our choice of $\left(\frac{1}{qw_1}\right)^{1/(q+1)}$, slightly scaled down, for x.

To show $x\phi'(x)/\phi(x) < 1$, we make the following observations, each of which is self evident:

- $x\phi'(x) = \sum_{i=1}^{5} l_i w_i x^{l_i}$
- $\phi(x) \geq 1 + w_1 x^{q+1}$
- $x = O\left(\frac{1}{d^2}\right)$
- $l_i w_i = O(d^{2l_i - 1})$ for each $i \in [2,5]$.

It follows that

$$\frac{\sum_{i=2}^{5} l_i w_i x^{l_i}}{\phi(x)} = O\left(\frac{1}{d}\right)$$

and

$$\frac{x\phi'(x)}{\phi(x)} \leq \frac{(q+1)w_1 x^{q+1}}{1+w_1 x^{q+1}} + O\left(\frac{1}{d}\right) = \frac{\frac{1}{1+\epsilon_1}\cdot\frac{q+1}{q}}{1+\frac{1}{(1+\epsilon_1)q}} + O\left(\frac{1}{d}\right) = \frac{\frac{1}{1+\epsilon_1}\cdot\frac{q+1}{q}}{\frac{1}{1+\epsilon_1}\cdot\frac{q+1}{q}+\frac{\epsilon_1}{1+\epsilon_1}} + O\left(\frac{1}{d}\right).$$

Thus $x\phi'(x)/\phi(x) < 1$ when d is large enough, as claimed. Hence, to prove the lemma it is enough to show that $\phi(x)/x < s$ for d large enough, by Lemma 109.

Observe that

$$\frac{\phi(x)}{x} = \frac{1}{x} + w_1 x^q + O(d).$$

Since $s = \binom{d-q}{2} - 3d = \Theta(d^2)$, to prove that $\phi(x)/x < s$ for d large enough it is enough to show that $1/x + w_1 x^q < (1-\delta)s$ for some fixed $\delta > 0$. Let

$$c_{q,\epsilon_1} := (q(1+\epsilon_1))^{1/(q+1)} + \left(\frac{1}{q(1+\epsilon_1)}\right)^{q/(q+1)}.$$

Using that $\binom{a}{b} \leq \frac{a^b}{b!}$ and $d \leq \Delta/2$, we obtain the following bound:

$$\frac{1}{x} + w_1 x^q = c_{q,\epsilon_1}\left(\binom{\Delta}{q} d^{q+2}\right)^{1/(q+1)} \leq c_{q,\epsilon_1}\left(\frac{\Delta^{2q+2}}{2^{q+2}q!}\right)^{1/(q+1)} = c_{q,\epsilon_1}\left(\frac{1}{2^{q+2}q!}\right)^{1/(q+1)}\Delta^2.$$

Since $s = \binom{d-q}{2} - 3d$, for any fixed $\epsilon' > 0$ we have $s \geq \frac{1-\epsilon'}{2}d^2 \geq \frac{(1-\epsilon')(1/2-\epsilon)^2}{2}\Delta^2$ if d is large enough. Hence, to conclude the proof it suffices to show that the following inequality holds if ϵ, ϵ' and ϵ_1 are chosen small enough:

$$c_{q,\epsilon_1}\left(\frac{1}{2^{q+2}q!}\right)^{1/(q+1)} < \frac{(1-\epsilon')(1/2-\epsilon)^2}{2}. \tag{5.4}$$

This is true, since $c_{q,0}\left(\frac{1}{2^{q+2}q!}\right)^{1/(q+1)} \simeq 0.12292 < 1/8$ for $q = 13$. □

It follows from Theorem 108 and the above lemma that $C_{t,E} \in o(s^t)$, and hence $|\mathcal{L}_t| \in o(s^t)$, when d is large enough. (To avoid any confusion, let us emphasise that here the $o(\cdot)$ notation is w.r.t. the variable t, that is, we first assume that d is large enough for Lemma 110 to hold, and then when the graph is fixed we let t vary.) Since there are s^t random sequences of length t, Theorem 104 follows from Lemma 105 and Lemma 110.

It follows from Theorem 104 that Algorithm 4 stops on some random sequence, and thus a function U^+ satisfying the properties of Lemma 103 exists. Consider such a function U^+ and the corresponding set of selected edges. Recall that each vertex of G is incident to at most $q+2$ selected edges, as follows from property (1) of Lemma 103. Using Vizing's theorem we recolour the set of selected edges properly using $q+3$ new colours, say from the set $[\Delta+3, \Delta+q+5]$. Let c' denote the resulting edge colouring of G. That is, $c'(e) := c(e)$ if e was not selected, and $c'(e)$ denotes the new colour of e if e was selected.

For each bad edge $uv \in M$, choose one of its two endpoints, say u, and mark one edge uw for some vertex $w \in U^+(u)$, with $w \neq v$ in case $v \in U^+(u)$. It follows from property (4) of Lemma 103 that marked edges form a matching, and that each bad edge is incident to exactly one marked edge. Recolouring all marked edges with a new colour, say colour $\Delta+q+6$, we obtain a proper edge colouring c'' of G with $\Delta+q+6$ colours such that $S_{c''}(u) \neq S_{c''}(v)$ for all edges $uv \in E(G)$ with $u, v \in B$ and $d_G(u) = d_G(v)$. Indeed, if uv is a bad edge this holds because there is exactly one marked edge incident to uv, and it is distinct from uv itself. If uv is not a bad edge, then by definition u and v see distinct sets of colours in the edge colouring c when considering only non-selected edges. Since marked edges form a subset of selected edges, we see that $S_{c''}(u) \neq S_{c''}(v)$ as desired.

Finally, consider edges $uv \in E(G)$ with $u, v \in A$ that are isolated in $G[A]$ (i.e. such that all neighbours of u, v outside $\{u, v\}$ are big) with $d_G(u) = d_G(v)$. Recall that $d_G(u) = d_G(v) \geq 2$, since uv is not isolated in G. Recall also that in the initial colouring c of G we had $S_c(u) \cap S_c(v) = \{c(uv)\}$, that is, u and v saw no common colour except for that of uv. If uv is fragile, then at least one of u, v is such that no incident edge was selected, by property 2 of Lemma 103, and hence $S_{c''}(u) \neq S_{c''}(v)$ (since marked edges form a subset of selected edges). If uv is not fragile, then $d_G(u) = d_G(v) \geq q+4$ by definition. Since at most $q+2$ edges incident to u were selected, and same for v, we see that u and v are each incident to a non-selected edge distinct from uv. It follows that $S_{c''}(u) \neq S_{c''}(v)$.

5.2.3 Small vertices

At this point, we know that $S_{c''}(u) \neq S_{c''}(v)$ for every edge $uv \in E(G)$ with $u, v \in B$, and for every edge $uv \in E(G)$ with $u, v \in A$ which is isolated in $G[A]$. However, we could have $S_{c''}(u) = S_{c''}(v)$ for some non-isolated edges uv of $G[A]$. Let A' be the subset of vertices of A that are not incident to an isolated edge of $G[A]$. In this section we modify the colouring c'' on the graph $G[A']$ only, and make sure that $S_{c''}(u) \neq S_{c''}(v)$ for every $uv \in E(G)$ with $u, v \in A'$. Since this has no effect on the sets $S_{c''}(u)$ for $u \in B \cup (A - A')$, the resulting colouring will be an AVD-colouring of G.

First, uncolour every edge of $G[A']$ and fix an arbitrary ordering of these edges. We colour these edges one by one using the following iterative algorithm; at all times, we let c''' denote the current partial edge colouring of G. Consider the first uncoloured edge uv in the ordering. Let $s := \lceil 2\epsilon\Delta \rceil$. Since $(d_G(u)-1)+(d_G(v)-1) \leq 2(d-1) \leq \Delta - 2\epsilon\Delta$, there are at least $s+q+6$ available choices for the edge uv in order to maintain a proper (partial) edge colouring. In case all other edges around u are already coloured, we possibly remove one colour from the set of available choices as follows: Say that a neighbour w of u in $A' - \{v\}$ is *dangerous* for u if $d_G(u) = d_G(w)$, all edges incident to w are already coloured, and $S_{c'''}(w) = S_{c'''}(u) \cup \{i\}$ for some colour $i \in [\Delta+q+6]$; the colour i is a *dangerous colour* for u. Dangerous neighbours and colours for v are defined similarly. If u has exactly one dangerous neighbour, remove the corresponding dangerous colour from the set of available choices. Do the same for v. Thus, there are at least $s+q+4 \geq s$ available choices remaining for the edge uv. Colour uv with a colour chosen at random among the first s colours available.

As with the algorithm from the previous section we define some bad events that could happen after colouring the edge uv. Here, we only need to consider one type of bad event:

The edge uv received a colour that was dangerous for u or v.

If such an event happens, consider a corresponding dangerous neighbour w, say it was dangerous for u. Let F denote the set of edges incident to u in $G[A']$ that are distinct from uv. Observe that $|F| \geq 2$, since otherwise we would have removed the dangerous colour for u from the available choices. Our ordering of the edges of $G[A']$ induces an ordering of the edges in F; it will be convenient to see this ordering as a *cyclic* ordering. With these notations, the bad event is handled as follows:

Uncolour uv and the edge just after uw in the cyclic ordering of F.

After possibly handling one such bad event, the algorithm proceeds with the next uncoloured edge in this fashion, until every edge is coloured.

Lemma 111. *If the algorithm terminates, then the resulting edge colouring c''' is an AVD-colouring of G.*

Proof. Consider an edge $uv \in E(G)$ with $d_G(u) = d_G(v)$. We already know that $S_{c'''}(u) \neq S_{c'''}(v)$ if $u, v \in B$, so let us assume that $u, v \in A$. Arguing by contradiction, suppose that $S_{c'''}(u) = S_{c'''}(v)$. Recall that $G[A']$ has no isolated edges, thus there is at least one edge incident to u or v which is distinct from uv in $G[A']$. Let e be the last edge coloured by the algorithm among all such edges. Suppose w.l.o.g. that e is incident to v, say $e = vw$. Then, just before the edge vw was coloured for the last time, vertex u was dangerous for v, with dangerous colour $c'''(vw)$. Hence, a bad event has been triggered after colouring vw. The bad event that was handled by the algorithm could have been the one with edge uv, or another one corresponding to another edge incident to v or w. In any case, the edge vw got uncoloured, a contradiction. $\square$

Thanks to the above lemma, to conclude the proof it only remains to show that the algorithm terminates with nonzero probability, which we do now.

Theorem 112. *The algorithm terminates with high probability.*

Proof. The proof is very similar to the corresponding proof in the previous section (but simpler). Let us encode the first t steps (iterations) of an execution of the algorithm with a corresponding *log* $(W, \gamma, \delta, c''')$, where

- W is a partial Dyck word of semilength t, obtained by adding a 0 (a 1) each time an edge is coloured (uncoloured, respectively);
- $\gamma = (\gamma_1, \dots, \gamma_t)$;
- $\delta = (\delta_1, \dots, \delta_t)$;

- c''' is the current edge colouring at the end of the t-th iteration.

For each $i \in [t]$, we let $\gamma_i := -1$ and $\delta_i := -1$ in case no bad event was triggered during the i-th iteration. Otherwise, if a bad event occurred, say involving a vertex w that was dangerous for one of the two endpoints of the edge uv coloured during the iteration, we let $\gamma_i \in [2d]$ identify vertex w knowing uv (recall that u and v have degree at most d). Observe that this identifies also the extra edge that is uncoloured (besides the edge uv).

Then, we let $\delta_i \in [2]$ identify the colours of the two edges that got uncoloured, assuming we know these two edges and the edge colouring c''' at the end of iteration i. Observe that we already know the *set* of colours that was used for these two edges, these are the two colours appearing around w but not around the vertex (u or v) that triggered the bad event. Thus it only remains to specify the mapping of these two colours to the two edges (2 possibilities).

Reading W and γ from the beginning, one can deduce which subset of the edges of $G[A']$ was coloured at any time during the execution. Then using the edge colouring c''' at the end of the t-th iteration and working backwards, we can reconstruct the edge colouring c''' at any time during the execution using γ and δ, and deduce in particular which random choice was made for the edge under consideration during the i-th iteration. Hence, the log $(W, \gamma, \delta, c''')$ uniquely determines the t random choices that were made.

As before, we see a random choice as a number in $[s]$. Let $\mathcal{R}_t$ denote the set of random vectors $(r_1, \ldots, r_t)$ of length t, where each entry is a number in $[s]$. Let $\mathcal{L}_t$ denote the set of logs after t steps resulting from executions of the algorithm that last for at least t steps. By the discussion above, there is an injective mapping from $\mathcal{L}_t$ to $\mathcal{R}_t$. Since $|\mathcal{R}_t| = s^t$, to prove Theorem 112 it only remains to show that $|\mathcal{L}_t| = o(s^t)$.

Here, a rather crude counting will do. First, we count the partial Dyck words W of semilength t that can appear in our logs. Each such word has only descents of length 2. They can be mapped to Dyck words of semilength t simply by adding the missing 1s at the end. Notice that each Dyck word of semilength t is the image of at most two such partial Dyck words. (Two of our partial Dyck words have the same image iff they are the same except one ends with 0 and the other ends with 011.) Hence, the number of our partial Dyck words of semilength t is at most twice the number of Dyck words of semilength t, and thus is at most $2 \cdot 4^t$.

Next, given a log $(W, \gamma, \delta, c''')$, the indices $i \in [t]$ such that $\gamma_i \neq -1$ and $\delta_i \neq -1$ correspond to descents of W. Thus there are at most $t/2$ such indices, and we see that the number of possible pairs (γ, δ) for a given W is at most $(2d)^{t/2} \cdot 2^{t/2} = (4d)^{t/2} \leq (2\Delta)^{t/2}$.

Finally, the number of partial edge colourings c''' of G is at most $|E(G)|^{\Delta+q+7}$, and is in particular independent of t.

Assuming that Δ is large enough so that $s = \lceil 2\epsilon\Delta \rceil > (32\Delta)^{1/2}$, we conclude that

$$|\mathcal{L}_t| \leq 2 \cdot 4^t \cdot (2\Delta)^{t/2} \cdot |E(G)|^{\Delta+q+7} = O\left((32\Delta)^{t/2}\right) = o(s^t),$$

as desired. $\square$

Chapter 6

Conclusions and further work

In this thesis, we investigated different problems regarding structures in digraphs, in particular subdivisions.

The first topic was to study the relation between subdivisions and chromatic number of digraphs. In Theorem 61, we generalised Bondy's Theorem (Theorem 31), by showing that large cycles with two blocks can be found in strong digraphs with large chromatic number. An interesting result would be to generalise this result to any oriented cycle.

Problem 32. Let C be an oriented cycle. Is $\chi(\text{S-Forb}(C) \cap \mathcal{S})$ bounded?

Let us remind our conjecture, which would be an interesting tool towards this question.

Conjecture 38. For every $k \geq 1$, there exists an integer $f(k)$ such that every strong digraph with chromatic number greater than $f(k)$ contains a subdigraph H with chromatic number at least k and such that H contains a Hamiltonian cycle.

Another interesting generalisation would be the following.

Conjecture 36. There is a function $g : \mathbb{N}^3 \to \mathbb{N}$ such that every strong digraph with chromatic number at least $g(k_1, k_2, k_3)$ contains a subdivision of $B(k_1, k_2; k_3)$.

In Theorem 37, we proved a weaker result for $B(k, 1; k)$. The proof is quite technical, and we believe some new ideas are necessary to prove the general statement. In this case as well, being able to prove Conjecture 38 would be an interesting step.

A famous open question concerning the chromatic number of digraphs is Burr's conjecture, which states that digraphs with chromatic number greater than $2k - 2$ contain all the oriented trees on k vertices as subdigraphs. The current best upper bound is in $\frac{k^2}{2} + o(k^2)$ and getting it down to linear seems to be very challenging. An interesting weaker problem is the following, where we replace subdigraph with subdivision.

Conjecture 113. There exists a constant a such that for every k, every digraph D with $\chi(D) \geq ak$ contains every oriented tree of size k as a subdivision.

The most important question concerning subdivisions in digraphs is probably Mader's conjecture about subdivisions of transitive tournaments.

Conjecture 9 (Mader [49])**.** For every $k \geq 1$, there exists an integer $f(k)$ such that every digraph with minimum outdegree at least $f(k)$ contains a subdivision of TT_k, the transitive tournament on k vertices.

A major part of this thesis was devoted to the study of this question, which remains widely open and deserves further research. In Theorem 19 and Theorem 20 we proved the existence of large oriented paths and in-arborescences in digraphs with large minimum outdegree. It might be interesting to generalise these results to other acyclic digraphs, in particular to oriented trees:

Problem 21. Let T be an oriented tree. Does there exist a constant $a(T)$ such that any digraph with minimum outdegree at least $a(T)$ contains T as a subdivision.

Another interesting case is the one for oriented cycles:

Problem 114. Let C be a oriented cycle. Does there exist a constant $f(C)$ such that any digraph with minimum outdegree at least $f(C)$ contains a subdivision of C.

Mader in [50] proved the existence of a pair of vertices with connectivity k in digraphs with large minimum outdegree. This means finding k vertex-disjoint dipaths between some pair of vertice x and y. However, forcing the length of these paths seems difficult. We proved in Theorem 84 that digraph with large minimum outdegree contains subdivisions of $C(k,k)$ (the cycle with two blocks of length k). A natural generalisation is the following problem:

Problem 115. Let $k_1, k_2, \ldots, k_l$ be positive integers. Does there exist an integer $f(k_1, k_2, \ldots, k_l)$ such that in any digraph with minimum outdegree at least $f(k_1, k_2, \ldots, k_l)$ there exist two vertices x and y and a collection of l internally disjoint (x, y)-dipaths $P_1, \ldots, P_l$, such that $|P_i| \geq k_i$ for every $i \in [l]$.

We do not know about the existence of $f(2, 2, 3)$.

One of the difficulties in solving Mader's conjecture is the effect of removing directed paths for connectivity. In the case of undirected graphs, the famous Lovász' path removal conjecture (see [39]) states the following:

Conjecture 116 (Lovász' path removal conjecture)**.** There is an integer-valued function $f(k)$ such that if G is any $f(k)$-connected graph and x and y are any two vertices of G, then there exists a path P with ends x and y such that $G - V(P)$ is k-connected.

A weaker version of the conjecture which removes the edges of the path instead of the vertices has been proved by Kawarabayashi et al. in [39]. A similar statement for digraphs has been proved false by Thomassen in [71]. This probably indicates why strengthening the requirement from large minimum outdegree to large connectivity doesn't seem to be helpful for finding a subdivision of TT_k. Understanding, which parameters stronger than the minimum outdegree condition can help us in finding transitive tournaments, is an interesting question. Remember that proving the result for large out and indegree would imply Mader's conjecture. On the other hands, it makes the version of the conjecture for oriented trees trivial. It is likely that the key parameter isn't something as symmetrical as k-connectivity. In this regard, Mader's proof of the existence of vertices with high connectivity in digraphs with large minimum outdegree, is interesting to analyse. In the undirected case, by considering the dense part of an (x, y) k-vertex-cut, we can progress towards a k-connected subgraph. In the directed case however, we obtain a digraph with some vertices of degree 0 and for every vertex x with $d^+(x) > 0$, there exists a vertex y such that the connectivity between x and y is greater than k. Another interesting observation is that, despite the fact that we mainly focused on acyclic digraphs, they are not the only possible candidates to be found in all digraphs with large minimum outdegree. For example, by considering a miximal dipath, one can easily show that every digraph with minimum degree at least k contains a directed cycle of length at least k. With constructions given in [20] and in [70], we have a pretty rough idea of the probable candidates. However the fact that we are not able to give a precise and simple characterisation shows, once again, our lack of understanding. Investigating this question might help us find a key notion for Mader's conjecture.

Another obstacle for Mader's conjecture, and also part of the reason why we believe it is a crucial one, is that the class of digraphs with large minimum outdegree remains a mystery. Let us remind the following conjecture due to Alon in [4] and whose difficulty is particularly intriguing:

Conjecture 17 (Alon [4])**.** There exists a function $f(k)$ such that every digraph with minimum degree at least $f(k)$ can be partitioned into two digraphs of minimum outdegree at least k.

We can also cite the following infamous conjecture on the class of digraphs with large minimum degree:

Conjecture 117 (Cacceta-Häggkvist [14])**.** Every simple digraph of order n with minimum outdegree at least r has a directed cycle with length at most $\lceil n/r \rceil$.

The case $r = \frac{n}{3}$ of the conjecture is still open, and is probably the most studied question in structural digraph theory (see [66] for a survey on this conjecture). Thus finding any structure for this class of digraphs is interesting.

In Theorem 86, we proved a weakening of Mader's conjecture, due to DeVos et al., about immersions instead of subdivisions. Proving the conjecture for butterfly-minors (see [38] for more details on butterfly-minors), which is another weakening of subdivisions, seems to be an interesting middle step towards Mader's conjecture. Especially since finding a complete graph (or just a dense one) as a minor is an important part of both the proof of Thomas and Wollan of Theorem 16 and the proof of Bollobás and Thomason in [8] that a minimum degree of $O(k^2)$ is enough to force a subdivision of the complete graph on k vertices.

Conjecture 118. There exists a function f such that every digraph with minimum outdegree at least $f(k)$ contains a TT_k as a butterfly-minor.

Proving Alon's conjecture (Conjecture 17) would also be a major step in our understanding of digraphs with large minimum outdegree. A tempting approach is to use the entropy compression technique. The major obstacle for this approach are the vertices of large indegree as they are the ones creating large dependencies in a random partition. It could be interesting to look at the case of undirected graphs. In [65] Stiebitz proved that graphs with minimum degree $s+t+1$ can be partitioned into two graphs, one of minimum degree s and one of minimum degree t, which is optimal. However, as far as we know, no probabilistic proof of the existence of a partition (even not the optimal one) exists, and once again dealing with the vertices of large degree seems to be the main difficulty.

In Theorem 49, we provided a proof that the domination of k-transitive tournaments is bounded by a function of k. The generalisation of this theorem to digraphs remains open:

Conjecture 43 (Sands et al. [63])**.** For every $k \geq 1$ there exists an integer $f(k)$ such that, for any multidigraph D whose arc set is the union of the arc sets of k quasi-orders, there exists $f(k)$ stable sets such that the union is dominating.

Our proof can be extended to the class of digraphs with bounded stability number. An interesting result would be to find classes of digraphs with unbounded stability number and where the conjecture holds.

Bibliography

[1] P. Aboulker, N. Cohen, F. Havet, W. Lochet, P. Moura, and S. Thomassé. Subdivisions in digraphs of large out-degree or large dichromatic number. *arXiv preprint arXiv:1610.00876*, 2016.

[2] L. Addario-Berry, F. Havet, C. Linhares Sales, B. Reed, and S. Thomassé. Oriented trees in digraphs. *Discrete Mathematics*, 313(8):967 – 974, 2013.

[3] L. Addario-Berry, F. Havet, and S. Thomassé. Paths with two blocks in n-chromatic digraphs. *Journal of Combinatorial Theory, Series B*, 97(4):620 – 626, 2007.

[4] N. Alon. Disjoint directed cycles. *Journal of Combinatorial Theory, Series B*, 68(2):167 – 178, 1996.

[5] N. Alon, G. Brightwell, H.A. Kierstead, A.V. Kostochka, and P.R Winkler. Dominating sets in k-majority tournaments. *Journal of Combinatorial Theory, Series B*, 96(3):374 – 387, 2006.

[6] P. N. Balister, E. Győri, J. Lehel, and R. H. Schelp. Adjacent vertex distinguishing edge-colorings. *SIAM J. Discrete Math.*, 21(1):237–250, 2007.

[7] E. Berger, K. Choromanski, M. Chudnovsky, J. Fox, M. Loebl, A. Scott, P.D. Seymour, and S. Thomassé. Tournaments and colouring. *Journal of Combinatorial Theory, Series B*, 103(1):1 – 20, 2013.

[8] B. Bollobás and A. Thomason. Proof of a conjecture of Mader, Erdős and Hajnal on topological complete subgraphs. *Eur. J. Comb.*, 19:883–887, 1998.

[9] J. A. Bondy. Disconnected orientations and a conjecture of Las Vergnas. *J. London Math. Soc.*, 14, 1976.

[10] N. Bousquet, W. Lochet, and S. Thomassé. A proof of the Erdős-Sands-Sauer-Woodrow conjecture. *arXiv preprint arXiv:1703.08123*, 2017.

[11] I. Bárány and J. Lehel. Covering with euclidean boxes. *European Journal of Combinatorics*, 8(2):113 – 119, 1987.

[12] R. Brooks. On colouring the nodes of a network. *Mathematical Proceedings of the Cambridge Philosophical Society*, 37(2):194–197, 1941.

[13] S.A. Burr. Subtrees of directed graphs and hypergraphs. In *Proceedings of the Eleventh Southeastern Conference on Combinatorics, Graph Theory and Computing (Florida Atlantic Univ., Boca Raton, Fla., 1980), Vol. I*, volume 28, pages 227–239, 1980.

[14] L. Caccetta and R. Häggkvist. *On minimal digraphs with given girth.* Department of Combinatorics and Optimization, University of Waterloo, 1978.

[15] Y. Caro, D. West, and R. Yuster. Equitable hypergraph orientations. *Electron. J. Combin.*, 18(1), 2011.

[16] N. Cohen, F. Havet, W. Lochet, and R. Lopes. Bispindles in strongly connected digraphs with large chromatic number. *The Electronic Journal of Combinatorics*, 25(2):2–39, 2018.

[17] N. Cohen, F. Havet, W. Lochet, and N. Nisse. Subdivisions of oriented cycles in digraphs with large chromatic number. *arXiv preprint arXiv:1605.07762*, 2016.

[18] N. Cohen and W. Lochet. Equitable orientations of sparse uniform hypergraphs. *The Electronic Journal of Combinatorics*, 23(4):P4–31, 2016.

[19] M. DeVos, Z. Dvořák, J. Fox, J. McDonald, B. Mohar, and D. Scheide. A minimum degree condition forcing complete graph immersion. *Combinatorica*, 34(3):279–298, Jun 2014.

[20] M. DeVos, J. McDonald, B. Mohar, and D. Scheide. Immersing complete digraphs. *European Journal of Combinatorics*, 33(6):1294 – 1302, 2012.

[21] M. Drmota. Combinatorics and asymptotics on trees. *Cubo Journal*, 6(2), 2004.

[22] P. Erdős. Graph theory and probability. *Canad. J. Math*, 11:34–38, 1959.

[23] P. Erdős and A. Hajnal. On chromatic number of graphs and set-systems. *Acta Mathematica Hungarica*, 17(1-2):61–99, 3 1966.

[24] P. Erdős and J. Spencer. Lopsided Lovász Local lemma and latin transversals. *Discrete Applied Mathematics*, 30(2):151 – 154, 1991.

[25] L. Esperet and A. Parreau. Acyclic edge-coloring using entropy compression. *European Journal of Combinatorics*, 34(6):1019–1027, 2013.

[26] D.C. Fisher. Squaring a tournament: A proof of Dean's conjecture. *Journal of Graph Theory*, 23(1):43–48, 1996.

[27] D.C. Fisher and J. Ryan. Probabilities within optimal strategies for tournament games. *Discrete Applied Mathematics*, 56(1):87 – 91, 1995.

[28] P. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge University Press, New York, NY, USA, 2009.

[29] S. Fortune, J. Hopcroft, and J. Wyllie. The directed subgraph homeomorphism problem. *Theoretical Computer Science*, 10(2):111 – 121, 1980.

[30] T. Gallai. On directed paths and circuits. *In: Erdös, P. and Katona, G., Eds., Theory of Graphs, Academic Press, New York*, pages 115–118, 1968.

[31] C. Greenhill and A. Ruciński. Neighbour-distinguishing edge colourings of random regular graphs. *Electronic Journal of Combinatorics*, 13:#R77, 2006.

[32] A. Gyárfás. Graphs with k odd cycle lengths. *Discrete Mathematics*, 103(1):41 – 48, 1992.

[33] A. Harutyunyan, T. Le, A. Newman, and S. Thomassé. Domination and fractional domination in digraphs. *arXiv preprint arXiv:1708.00423*.

[34] M. Hasse. Zur algebraischen bergrund der graphentheorie i. *Math. Nachr.*, 28:275–290, 1964.

[35] H. Hatami. $\Delta + 300$ is a bound on the adjacent vertex distinguishing edge chromatic number. *J. Combin. Theory Ser. B*, 95(2):246–256, 2005.

[36] T. Johnson, N. Robertson, P.D. Seymour, and R. Thomas. Directed tree-width. *Journal of Combinatorial Theory, Series B*, 82(1):138 – 154, 2001.

[37] G. Joret and W. Lochet. Progress on the adjacent vertex distinguishing edge colouring conjecture. *arXiv preprint arXiv:1804.06104*, 2018.

[38] K. Kawarabayashi and S. Kreutzer. The directed grid theorem. In *Proceedings of the Forty-seventh Annual ACM Symposium on Theory of Computing*, STOC '15, pages 655–664, New York, NY, USA, 2015. ACM.

[39] K. Kawarabayashi, O. Lee, B. Reed, and P. Wollan. A weaker version of Lovász' path removal conjecture. *Journal of Combinatorial Theory, Series B*, 98(5):972 – 979, 2008.

[40] D. Kühn, R. Mycroft, and D. Osthus. A proof of Sumner's universal tournament conjecture for large tournaments. *Proceedings of the London Mathematical Society*, 102(4):731–766, 2011.

[41] D. Kühn, D. Osthus, and A. Young. A note on complete subdivisions in digraphs of large outdegree. *Journal of Graph Theory*, 57(1):1–6, 2008.

[42] R. Kim, S.-J. Kim, J. Ma, and B. Park. Cycles with two blocks in k-chromatic digraphs. *Journal of Graph Theory*, 0(0).

[43] J. Komlós and E. Szemerédi. Topological cliques in graphs ii. *Combinatorics, Probability and Computing*, 5(1):79–90, 1996.

[44] A. V. Kostochka. Lower bound of the hadwiger number of graphs by their average degree. *Combinatorica*, 4(4):307–316, Dec 1984.

[45] W. Lochet. Immersion of transitive tournaments in digraphs with large minimum outdegree. *Journal of Combinatorial Theory, Series B*, 2018.

[46] C. Löwenstein, D. Rautenbach, and I. Schiermeyer. Cycle length parities and the chromatic number. *Journal of Graph Theory*, 64(3):210–218, 2010.

[47] W. Mader. Homomorphieeigenschaften und mittlere kantendichte von graphen. *Mathematische Annalen*, 174(4):265–268, Dec 1967.

[48] W. Mader. Existenz n-fach zusammenhängender teilgraphen in graphen genügend grosser kantendichte. *Abh. Math. Sem. Univ. Hamburg*, 37, 1972.

[49] W. Mader. Degree and local connectivity in digraphs. *Combinatorica*, 5(2):161–165, Jun 1985.

[50] W. Mader. Existence of vertices of local connectivity k in digraphs of large outdegree. *Combinatorica*, 15(4):533–539, Dec 1995.

[51] J. Matousek. *Lectures on Discrete Geometry*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2002.

[52] K. Menger. Zur allgemeinen kurventheorie. *Fundamenta Mathematicae*, 10(1):96–115, 1927.

[53] P. Mihok and I. Schiermeyer. Cycle lengths and chromatic number of graphs. *Discrete Math*, 286(1-2):147–149, 2004.

[54] R. Moser and G. Tardos. A constructive proof of the general Lovász Local Lemma. *J. ACM*, 57(2):Art. 11, 2010.

[55] R.A. Moser. A constructive proof of the Lovász Local Lemma. In *Proceedings of the Forty-first Annual ACM Symposium on Theory of Computing*, STOC '09, pages 343–350, New York, NY, USA, 2009. ACM.

[56] D. Pálvölgyi and A. Gyárfás. Domination in transitive colorings of tournaments. *Journal of Combinatorial Theory, Series B*, 107:1 – 11, 2014.

[57] B. Reed. Introducing directed tree width. *Electronic Notes in Discrete Mathematics*, 3:222 – 229, 1999. 6th Twente Workshop on Graphs and Combinatorial Optimization.

[58] N. Robertson and P.D. Seymour. Graph minors. V. Excluding a planar graph. *Journal of Combinatorial Theory, Series B*, 41(1):92 – 114, 1986.

[59] N. Robertson and P.D Seymour. Graph minors .XIII. The disjoint paths problem. *Journal of Combinatorial Theory, Series B*, 63(1):65 – 110, 1995.

[60] N. Robertson and P.D. Seymour. Graph minors. XX. Wagner's conjecture. *Journal of Combinatorial Theory, Series B*, 92(2):325 – 357, 2004. Special Issue Dedicated to Professor W.T. Tutte.

[61] N. Robertson, P.D. Seymour, and R. Thomas. Hadwiger's conjecture for k_6-free graphs. *Combinatorica*, 13(3):279–361, 1993.

[62] B. Roy. Nombre chromatique et plus longs chemins d'un graphe. *Rev. Francaise Informat. Recherche Operationnelle*, 1:129–132, 1967.

[63] B. Sands, N. Sauer, and R. Woodrow. On monochromatic paths in edge-coloured digraphs. *Journal of Combinatorial Theory, Series B*, 33(3):271 – 275, 1982.

[64] N. Sauer. On the density of families of sets. *Journal of Combinatorial Theory, Series A*, 13(1):145 – 147, 1972.

[65] M. Stiebitz. Decomposing graphs under degree constraints. *Journal of Graph Theory*, 23(3):321–324.

[66] B. Sullivan. A summary of problems and results related to the Caccetta-Häggkvist conjecture. *arXiv preprint math/0605646*, 2006.

[67] R. Thomas and P. Wollan. An improved linear edge bound for graph linkages. *European Journal of Combinatorics*, 26(3):309 – 324, 2005. Topological Graph Theory and Graph Minors, second issue.

[68] A. Thomason. An extremal function for contractions of graphs. *Mathematical Proceedings of the Cambridge Philosophical Society*, 95(2):261–265, 1984.

[69] C. Thomassen. Disjoint cycles in digraphs. *Combinatorica*, 3(3):393–396, Sep 1983.

[70] C. Thomassen. Even cycles in directed graphs. *European Journal of Combinatorics*, 6(1):85 – 89, 1985.

[71] C. Thomassen. Highly connected non-2-linked digraphs. *Combinatorica*, 11(4):393–395, Dec 1991.

[72] L. M. Vitaver. Determination of minimal coloring of vertices of a graph by means of boolean powers of the incidence matrix. *Doklady Akademii Nauk SSSR*, 147:758–759, 1962.

[73] S.S. Wang. Structure and coloring of graphs with only small odd cycles. *SIAM Journal on Discrete Mathematics*, 22(3):1040–1072, 2008.

[74] Z. Zhang, L. Liu, and J. Wang. Adjacent strong edge coloring of graphs. *Appl. Math. Lett.*, 15(5):623–626, 2002.