New bounds for linear arboricity and related problems

Micha Christoph[†] Nemanja Draganić[‡] António Girão[‡] Eoin Hurley[‡]
Lukas Michel[‡] Alp Müyesser[‡]

28 July 2025

Abstract

A linear forest is a collection of vertex-disjoint paths. The Linear Arboricity Conjecture states that every graph of maximum degree Δ can be decomposed into at most $\lceil(\Delta+1)/2\rceil$ linear forests. We prove that $\Delta/2+\mathcal{O}(\log n)$ linear forests suffice, where n is the number of vertices of the graph. If $\Delta = \Omega(n^{\varepsilon})$, this is an exponential improvement over the previous best error term. We achieve this by generalising Pósa rotations from rotations of one endpoint of a path to simultaneous rotations of multiple endpoints of a linear forest. This method has further applications, including the resolution of a conjecture of Feige and Fuchs on spanning linear forests with few paths and the existence of optimally short tours in connected regular graphs.

1 Introduction

In 1850, Thomas Kirkman posed the following problem in *The Lady's and Gentleman's Diary*.

> Fifteen young ladies in a school walk out three abreast for seven days in succession: it is required to arrange them daily so that no two shall walk twice abreast.

Put differently, Kirkman was asking for a decomposition[1] of the edges of the complete graph on fifteen vertices into copies of a triangle. Investigating a similar problem in 1882, Walecki proved that for all odd n, the complete graph K_n can be decomposed into Hamilton cycles, which are cycles that contain every vertex. What other graphs can K_n be decomposed into? The answer to this question is perhaps counter-intuitive: if basic divisibility conditions are met, then K_n can be decomposed into a wide variety of graphs. Examples include large trees in Ringel's conjecture [MPS21, KS25], complete graphs in the study of designs [Kee14, GKLO23], cycles of various lengths in the Oberwolfach problem [KS22], and many others. Notably, all of these central conjectures have been proved in the past few decades as the machinery associated with absorption and the probabilistic method has become more sophisticated and versatile [AS16, KKK+23, GKLO23, Mon24, KSSS24].

The complete graph, however, is a very special and highly symmetric object, and so it is unclear exactly which of its properties are necessary for these decomposition results. One way to investigate this is to consider dense graphs as a slight weakening of complete graphs. It turns out that here too, a surprising variety of decompositions are possible, and here too there has been

[†]Department of Mathematics, ETH Zürich, Switzerland (micha.christoph@math.ethz.ch). Research supported by SNSF Ambizione Grant No. 216071.

[‡]Mathematical Institute, University of Oxford, United Kingdom ({nemanja.draganic,antonio.girao,eoin.hurley, lukas.michel,alp.muyesser}@maths.ox.ac.uk). Research of Nemanja Draganić supported by SNSF project 217926. Research of Eoin Hurley supported by ERC Advanced Grant 883810.

[1]A *decomposition* of a graph is a partition of its edge set.

significant progress in the past few decades. Perhaps the largest body of work in this direction concerns the decomposition of graphs that satisfy minimum degree conditions. Prominent examples of this include progress on the Nash-Williams conjecture on triangle decompositions [BKLO16, DP21], generalising Kirkman's problem, and the Hamilton decomposition theorem for regular Dirac graphs [CKL+16], generalising Walecki's result. Another exciting recent direction concerns the decomposition of graphs that are dense and quasirandom. Here, results include generalisations of Ringel's conjecture [KS25] and of the Oberwolfach problem [KS22], as well as *approximate* decompositions into arbitrary graphs of bounded degree [KKOT19]. This work shares some machinery with the case of complete graphs and makes frequent use of Szemerédi's famous regularity lemma. Overall, for the decomposition of dense graphs into graphs of fixed size or very sparse spanning structures, the following perspective has emerged.

> If there is no obvious barrier to the existence of a particular graph decomposition, then the decomposition likely exists.

But what if we go beyond complete and dense graphs? After all, most graphs encountered in areas like statistical physics, computer science, or group theory are graphs of bounded degree. Is it still true that beyond obvious barriers, decompositions exist? Of course, in general we can no longer hope for decompositions into complete graphs or Hamilton cycles since sparse graphs may not even contain these structures. So, we have to ask for different decompositions.

Two classical and elegant results on the decomposition of sparse graphs are Vizing's theorem [Viz64] and the Nash-Williams theorem [NW61, NW64]. Vizing's theorem, inspired by work of Shannon [Sha49], states that any graph of maximum degree Δ can be decomposed into $\Delta + 1$ matchings. The Nash-Williams theorem states that a graph can be decomposed into k forests if and only if every set of ℓ vertices spans at most $k \cdot (\ell - 1)$ edges. Thus, in both cases, decompositions exist beyond the obvious barriers. In recent decades, however, progress on sparse decomposition problems has lagged far behind its dense counterparts. This is largely because various tools—such as Szemerédi's regularity lemma—offer little leverage in sparse graphs[2]. Among the most famous open problems in this area are the List Edge Colouring Conjecture [Viz76, ERT80, Kah00], the Erdős-–Gallai Conjecture on decomposing graphs into cycles and edges [EGP66, BM24], and the Linear Arboricity Conjecture.

1.1 The Linear Arboricity Conjecture

A *linear forest* is a collection of vertex-disjoint paths, and the *linear arboricity* $\mathrm{la}(G)$ of a graph G is the minimum number of linear forests needed to decompose the edges of G [Har70]. Matchings are a special case of linear forests, and linear forests are a special case of forests, so the linear arboricity of a graph lies between the aforementioned results of Vizing and Nash-Williams. Since Hamilton paths are linear forests with the maximum possible number of edges, linear arboricity can also be viewed as a relaxation of a decomposition into Hamilton paths.

How large is the linear arboricity of a graph? Since a linear forest covers at most two edges incident to any vertex, every graph G with maximum degree Δ satisfies $\mathrm{la}(G) \geq \lceil \Delta/2 \rceil$. Moreover, a Δ-regular graph G on n vertices has $n \cdot \Delta/2$ edges while any linear forest of G has strictly less than n edges, and so $\mathrm{la}(G) \geq \lceil (\Delta+1)/2 \rceil$. The celebrated Linear Arboricity Conjecture of Akiyama, Exoo, and Harary [AEH80] asserts that, beyond these obvious barriers, there should always be a decomposition into linear forests.

Conjecture 1.1 (Linear Arboricity Conjecture)**.** *The linear arboricity of a graph with maximum degree Δ is at most $\lceil (\Delta+1)/2 \rceil$.*

[2] Although there exist sparse variants of the regularity lemma [GS05, Koh97, Sco11], they only provide information for sparse graphs whose edges are sufficiently well distributed, that is, pseudorandom graphs.

This conjecture immediately received significant attention [AEH81, Eno81, Pér82, Tom82, EP84, Gul86a, Gul86b]. By 1988, it was known to hold in the special cases $\Delta \in \{3, 4, 5, 6, 8, 10\}$, and the best general upper bound was $\mathrm{la}(G) \leq \lceil (3\Delta + 2)/4 \rceil$. In an influential paper, Alon [Alo88] then confirmed the Linear Arboricity Conjecture asymptotically, showing that

$$\mathrm{la}(G) \leq \frac{\Delta}{2} + \mathcal{O}\left(\frac{\Delta \log \log \Delta}{\log \Delta}\right) = (1 + o(1)) \cdot \frac{\Delta}{2}.$$

This was one of the early applications of the Lovász Local Lemma, giving the Linear Arboricity Conjecture a prominent place in the history of the probabilistic method [AS16].

Since then, some additional special cases of the conjecture have been solved, such as graphs with minimum degree at least $(1 + o(1)) \cdot n/2$ [GKO16, GS24], planar graphs [Wu99, WW08], random or dense quasirandom graphs [DGCS25, GKO16], and random regular graphs of constant degree [MR90]. The error term of Alon's upper bound was subsequently improved to $\mathcal{O}(\Delta^{2/3} \log^{1/3} \Delta)$ by Alon and Spencer [AS16], to $\mathcal{O}(\Delta^{2/3-\varepsilon})$ by Ferber, Fox, and Jain [FFJ20], and to $\mathcal{O}(\sqrt{\Delta} \log^4 \Delta)$ by Lang and Postle [LP23]. A central ingredient in all of these upper bounds is a certain degree of randomisation, either through a random partition of the vertex set or through nibble-based strategies. For both approaches, $\Omega(\sqrt{\Delta})$ is a natural bottleneck for the error term, caused by concentration inequalities used throughout these proofs. This is a reflection of the fact that the standard deviation of a binomial random variable with Δ trials is $\Theta(\sqrt{\Delta})$. Our main contribution is the following upper bound on the linear arboricity.

Theorem 1.2. *Every n-vertex graph G with maximum degree Δ satisfies* $\mathrm{la}(G) \leq \Delta/2 + \mathcal{O}(\log n)$.

This result breaks the square-root barrier for the Linear Arboricity Conjecture if $\Delta = \Omega(\log^2 n)$ and represents an exponential improvement in the error term for $\Delta = \Omega(n^\varepsilon)$. Before our work, there were no results that improved on the upper bound of Lang and Postle [LP23] even for very dense graphs with $\Delta = \Omega(n)$.

For the fractional relaxation of linear arboricity, the previous best error term was also stuck at the same square-root barrier. Formally, the *fractional linear arboricity* $\mathrm{fla}(G)$ of a graph G is the smallest real number k for which we can choose a random linear forest of G that contains every fixed edge of G with probability at least $1/k$. Clearly, $\Delta/2 \leq \mathrm{fla}(G) \leq \mathrm{la}(G)$. Feige, Ravi, and Singh [FRS14] showed that $\mathrm{fla}(G) \leq \Delta/2 + \mathcal{O}(\sqrt{\Delta})$. We improve the error term in this result to $\mathcal{O}(\log \Delta)$, providing an exponential improvement for all ranges of Δ.

Theorem 1.3. *Every graph G with maximum degree Δ satisfies* $\mathrm{fla}(G) \leq \Delta/2 + \mathcal{O}(\log \Delta)$.

Our proofs use a novel variant of Pósa rotations for linear forests. We discuss our proof techniques in more detail in Section 1.3. In particular, this section outlines the key lemma that we use to prove our results, Lemma 2.3. In addition to being crucial for our upper bounds on the linear arboricity, this lemma also has further applications, which we discuss in the following.

1.2 Spanning linear forests with few paths and short tours

If the Linear Arboricity Conjecture holds, then, by averaging, it follows that every d-regular graph[3] on n vertices has a spanning linear forest with as few as $2 \cdot n/(d+2)$ paths[4]. This consequence of the Linear Arboricity Conjecture is a challenging and well-studied problem in its own right. Magnant and Martin [MM09] conjectured that every d-regular graph has a spanning linear forest with $n/(d+1)$ paths. Feige and Fuchs later conjectured that the weaker

[3] Recall that any graph with maximum degree Δ is contained in a Δ-regular graph, and so we may, without loss of generality, assume that a given graph is regular whilst obtaining bounds for the Linear Arboricity Conjecture.

[4] See [FF22, Section 1.1] for a detailed discussion of these calculations.

bound $\mathcal{O}(n/d)$ should hold. These conjectures have been reiterated several times in recent work [MMPS24, LMS25].

The conjecture of Magnant and Martin would be tight, as witnessed by a disjoint union of cliques. However, this conjecture is only known to hold for $d \leq 6$ [MM09, FF22], $d = \Omega(n)$ [Han18, GL21], and for partitioning almost all vertices of a graph into paths [MMPS24, LMS25]. The upper bound on the fractional linear arboricity of Feige, Ravi, and Singh [FRS14] implies that there is a spanning linear forest with $\mathcal{O}(n/\sqrt{d})$ paths. Very recently, the authors of the current paper [CDG+25] broke the square-root barrier for this problem, improving the upper bound to $\mathcal{O}((n \log d)/d)$.

These results employed a variety of different approaches. In addition to strategies also used for the Linear Arboricity Conjecture, such as random partitioning arguments [MMPS24], previous techniques included Szemerédi's regularity lemma [Han18], robust expander decompositions [GL21, LMS25], and entropy-based methods [CDG+25]. Using our new methods based on rotations of linear forests, we establish the conjecture of Magnant and Martin up to a factor of two. In particular, this proves the conjecture of Feige and Fuchs.

Theorem 1.4. *Every d-regular n-vertex graph has a spanning linear forest with at most $2 \cdot n/(d+1)$ paths.*

For regular bipartite graphs, we even confirm the conjecture of Magnant and Martin, see Theorem 2.7. We prove these results in Section 2 as an immediate consequence of the key lemma that we use to make progress on the Linear Arboricity Conjecture.

Finally, our results imply that connected d-regular graphs have short tours, where a *tour* is a sequence of adjacent vertices that ends at the starting vertex and visits every vertex at least once. The problem of finding such tours was introduced by Vishnoi [Vis12] as a restriction of the Travelling Salesman Problem to regular graphs. Feige, Ravi, and Singh [FRS14] showed that every connected d-regular graph on n vertices has a tour of length $(1 + \mathcal{O}(1/\sqrt{d})) \cdot n$ and observed that one cannot do better than $(1 + \mathcal{O}(1/d)) \cdot n$. The authors of the current paper [CDG+25] very recently showed that tours of length $(1 + \mathcal{O}((\log d)/d)) \cdot n$ exist. Here, we resolve this problem. This is a direct consequence of Theorem 1.4 in combination with a result of Feige, Ravi, and Singh [FRS14, Corollary 1].

Corollary 1.5. *Every connected d-regular n-vertex graph has a tour of length at most $(1 + \mathcal{O}(1/d)) \cdot n$.*

It is straightforward to construct regular connected graphs that are not Hamiltonian. However, our results show that, at least in some sense, regular graphs are close to being Hamiltonian.

1.3 Proof techniques

A rotation of a path or a linear forest is a simple operation that modifies the path or linear forest in such a way that exactly one of its endpoints moves to a different vertex of the graph (see also Figure 1). This idea can be traced back to Dirac's theorem [Dir52] stating that graphs on n vertices with minimum degree at least $n/2$ are Hamiltonian. It was also used by Pósa [Pós76] to determine the threshold for the Hamiltonicity of random graphs. Pósa's proof is notable for being the first to consider *the family of all possible sequences of rotations* that can be performed to a given path of maximal length. This allows one to show that many vertices can become an endpoint of such a path, which in turn is crucial for showing that sufficiently dense random graphs are Hamiltonian. Since this was a very influential idea, rotations are often referred to as *Pósa rotations* in the literature. Although the idea of rotations has been around for a long time, several breakthroughs in recent years were only made possible by introducing novel twists to this method [GCS24, DCS24, DMC+24].

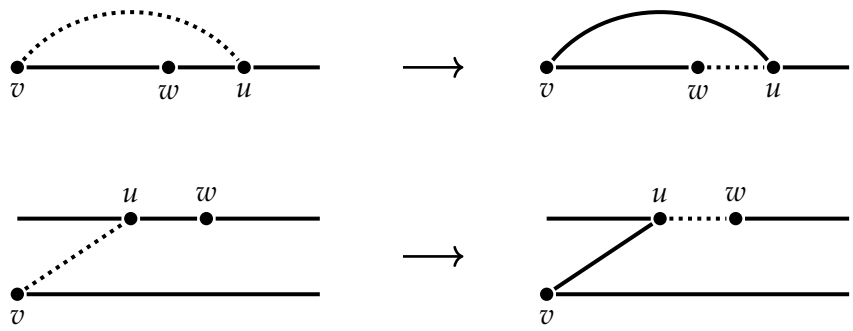

Figure 1: The two different types of rotations of a linear forest. Dotted edges represent edges of the graph that are not contained in the linear forest.

A common feature of these applications is some notion of expansion in the underlying graph. In contrast, we apply rotations to sparse regular graphs without any expansion conditions. Moreover, while Pósa rotations are only about rotations of *a single endpoint* of a path, our main novelty is to consider simultaneous rotations of *a set of endpoints* of a linear forest. Concretely, our key lemma, Lemma 2.3, proves the following.

> Suppose we are given a minimum spanning linear forest of a d-regular graph and a set of ℓ endpoints. If we allow any sequence of rotations that only involves this set of endpoints, then at least $\Omega(\ell \cdot d)$ vertices can become an endpoint.

As we make no assumptions on the structure of the underlying regular graph, this requires a highly delicate argument. We prove this lemma in Section 2 where we also concisely explain the key idea of its proof. For now, note that if there are at least $\Omega(\ell \cdot d)$ distinct potential endpoints, it immediately follows that $\ell = \mathcal{O}(n/d)$, which implies Theorem 1.4 and the conjecture of Feige and Fuchs. Moreover, in Section 3, we use this key lemma to create a distribution on the spanning linear forests of the graph in which every vertex has a probability of at most $\mathcal{O}(1/d)$ of being an endpoint. To prove our bounds on the (fractional) linear arboricity, Theorems 1.2 and 1.3, we then iteratively remove $\Delta/2$ such random linear forests from the graph and decompose the remaining edges into matchings using Vizing's theorem. Note that a vertex has a high degree in the remaining graph only if it was an endpoint in many of the removed linear forests. Since every vertex is unlikely to be an endpoint, we can apply standard concentration inequalities and a union bound to show that the maximum degree of the remaining graph is $\mathcal{O}(\log n)$. This gives the error term of $\mathcal{O}(\log n)$ in our result. We refer the reader to Section 4 for a discussion on the difficulties of improving this error term further.

Throughout the remainder of the paper, every linear forest is a spanning linear forest.

2 Small linear forests in regular graphs

In this section, we will show that any d-regular graph G on n vertices has a spanning linear forest with at most $2 \cdot n/(d+1)$ paths. Our main tool to achieve this are rotations of linear forests. A rotation modifies a linear forest by adding an edge incident to some endpoint of the linear forest and removing a different edge so that the resulting subgraph is still a linear forest. Intuitively, a rotation moves one endpoint of the linear forest to a different vertex of the graph. Figure 1 illustrates the two different types of possible rotations of a linear forest.

Definition 2.1. Let F be a linear forest of a graph G. Let v be an endpoint of F, and let u be a neighbour of v in G. If u and v are on the same path in F, let w be the neighbour of u that is closer to v on that path. Otherwise, let w be any neighbour of u on its path in F. Then, the

linear forest F' obtained from F by removing the edge uw and adding the edge uv is called a *rotation* of F in G.

We say that such a rotation rotates the *old endpoint* v to the *new endpoint* w, using the *pivot* u. The restriction on w in the case where u and v are on the same path is necessary to ensure that the resulting subgraph is still a linear forest. Note that w is allowed to be the same vertex as v, which could happen if u is the unique neighbour of v on its path in F. Also, observe that if F' is a rotation of F, then F' has the same number of paths as F.

If we can use rotations to reach a linear forest where two endpoints of *distinct* paths are at adjacent vertices of the graph, then we can reduce the number of paths of the linear forest by adding the edge between these two endpoints. Suppose that F is a linear forest where we can no longer do this, for example because F has the minimum number of paths possible. To show that F has few paths, we will consider the set C of all vertices of the graph that could become an endpoint of F through some sequence of rotations. We claim that if F has ℓ endpoints, then the size of C is at least $\ell \cdot (d+1)/4$. Since the size of C is at most the total number of vertices of the graph, this claim implies that F has at most $4 \cdot n/(d+1)$ endpoints and therefore at most $2 \cdot n/(d+1)$ paths, as required.

We now sketch our proof of this claim. First, suppose for simplicity that G is bipartite. In this case, by only rotating endpoints in one part of the bipartition of G, we may assume that C is an independent set. Then, we will analyse the edges between C and its *boundary* $B := N_G(C) \setminus C$[5].

In particular, we exploit the tension between the following two facts. Firstly, each vertex in C has d neighbours in B while each vertex in B has at most d neighbours in C. Secondly, each vertex in C has at most two path-neighbours[6] in B while each vertex in B has at least one path-neighbour in C[7]. Using a double counting argument on these edges and combining this with the fact that C contains at least one endpoint, we can find a vertex $u \in B$ with $(d+1)/2$ neighbours in C, but which has only one path-neighbour in C.

Now, rotate one endpoint of F to an arbitrary neighbour v of u and *fix it there*. If we never allow v to be rotated again, we claim that no other endpoint can be rotated to a neighbour v' of u. Indeed, if that were possible, a short case analysis shows that either one of the two endpoints v or v' could be rotated to a vertex outside of C, contradicting the definition of C, or one of the two edges uv or uv' could be added to F, contradicting the minimality of F. This case analysis uses the fact that u has only one path-neighbour in C.

Therefore, fixing an endpoint of F in the neighbourhood of u reduces the number of endpoints that we rotate by one while it reduces the number of vertices that could become an endpoint through a sequence of rotations by at least $(d+1)/2$. Iterating this argument shows that C has size at least $\ell \cdot (d+1)/2$, proving the result for bipartite graphs.

For non-bipartite graphs, we divide the argument into two cases based on the number of edges spanned by C. If C spans few edges, we can proceed exactly as for bipartite graphs. If instead C spans many edges, we iteratively fix endpoints at vertices that have a high degree within C and show that this reduces the size of C sufficiently. In both cases, we then iterate the argument to show that C has size at least $\ell \cdot (d+1)/4$, as required.

We now formalise this strategy. We begin by defining the set of all vertices of the graph that could become an endpoint of a linear forest through a sequence of rotations. In the following, we will use $\mathrm{End}(F)$ to denote the multiset of endpoints of a linear forest F, where an isolated

[5] If G is a graph and $X \subseteq V(G)$, the *neighbourhood* of X in G is $N_G(X) := \{v \in V(G) : uv \in E(G) \text{ for some } u \in X\}$.

[6] A *path-neighbour* of a vertex v is a neighbour of v on its path in F.

[7] The fact that every vertex in B has at least one path-neighbour in C follows from a standard argument for Pósa rotations.

vertex of F is contained twice in $\mathrm{End}(F)$. In particular, if F' is a rotation of F that rotates v to w, then $\mathrm{End}(F') = (\mathrm{End}(F) \setminus \{v\}) \cup \{w\}$.

Definition 2.2. Let F be a linear forest of a graph G, and let $X \subseteq \mathrm{End}(F)$ be a multiset of *fixed endpoints*. Then, $\mathcal{F}(F, X)$ denotes the family of linear forests obtained by a sequence of rotations starting at F that never rotates any endpoint of X. Define the set

$$C(F, X) := \bigcup_{F' \in \mathcal{F}(F,X)} \mathrm{End}(F') \setminus X.$$

We call $C(F, X)$ a *rotation-component*. Intuitively, $C(F, X)$ consists of all vertices that could become an endpoint of F through a sequence of rotations that leaves all endpoints of X fixed. Note that if an endpoint v is contained once in X but twice in $\mathrm{End}(F)$, then we allow one of the endpoints at v to be rotated, but not both. In particular, $C(F, X)$ might not be disjoint from X. Since rotations are reversible, we have $\mathcal{F}(F, X) = \mathcal{F}(F', X)$ for all $F' \in \mathcal{F}(F, X)$, and therefore also $C(F, X) = C(F', X)$. Moreover, if $X \subseteq Y \subseteq \mathrm{End}(F)$, then $C(F, X) \supseteq C(F, Y)$.

A *minimum linear forest* F of G is a linear forest with the minimum number of paths among all linear forests of G. We now state the key lemma of this section which asserts that for any minimum linear forest F, the size of the rotation-component $C(F, X)$ is always at least $(d+1)/4$ times the number of endpoints that are not fixed.

Lemma 2.3. *Let F be a minimum linear forest of a d-regular graph, and let $X \subseteq \mathrm{End}(F)$. Then,*

$$|C(F, X)| \geq \frac{d+1}{4} \cdot |\mathrm{End}(F) \setminus X|.$$

This result suffices to prove that every d-regular graph on n vertices has a linear forest with at most $2 \cdot n/(d+1)$ paths.

Proof of Theorem 1.4. Let G be a d-regular graph on n vertices, and let F be a minimum linear forest of G. By Lemma 2.3, we have

$$|\mathrm{End}(F)| \leq \frac{4}{d+1} \cdot |C(F, \varnothing)| \leq \frac{4}{d+1} \cdot n.$$

Since the number of endpoints of a linear forest is twice the number of its paths, this implies that F has at most $2 \cdot n/(d+1)$ paths. □

To prove Lemma 2.3, we first make the following observations about rotations that are used extensively throughout our proofs. These observations follow directly from the definition of rotations together with a small amount of case analysis as indicated in Figure 2.

(R1) Suppose that v and u are endpoints of F and neighbours in G.
  (a) If w is a neighbour of u on its path in F, then we can rotate v to w using the pivot u.
  (b) Alternatively, if u and v are on different paths in F, we can reduce the number of paths of F by adding the edge uv to F. In particular, this applies if u is an isolated vertex of F.

(R2) Suppose that v is an endpoint of F which is a neighbour of u in G. If w and w' are neighbours of u on its path in F, then we can rotate v to one of the two vertices w or w' using the pivot u.

(R3) Suppose that v and v' are endpoints of F which are both neighbours of u in G. Note that we allow $v = v'$ if v is an isolated vertex of F.

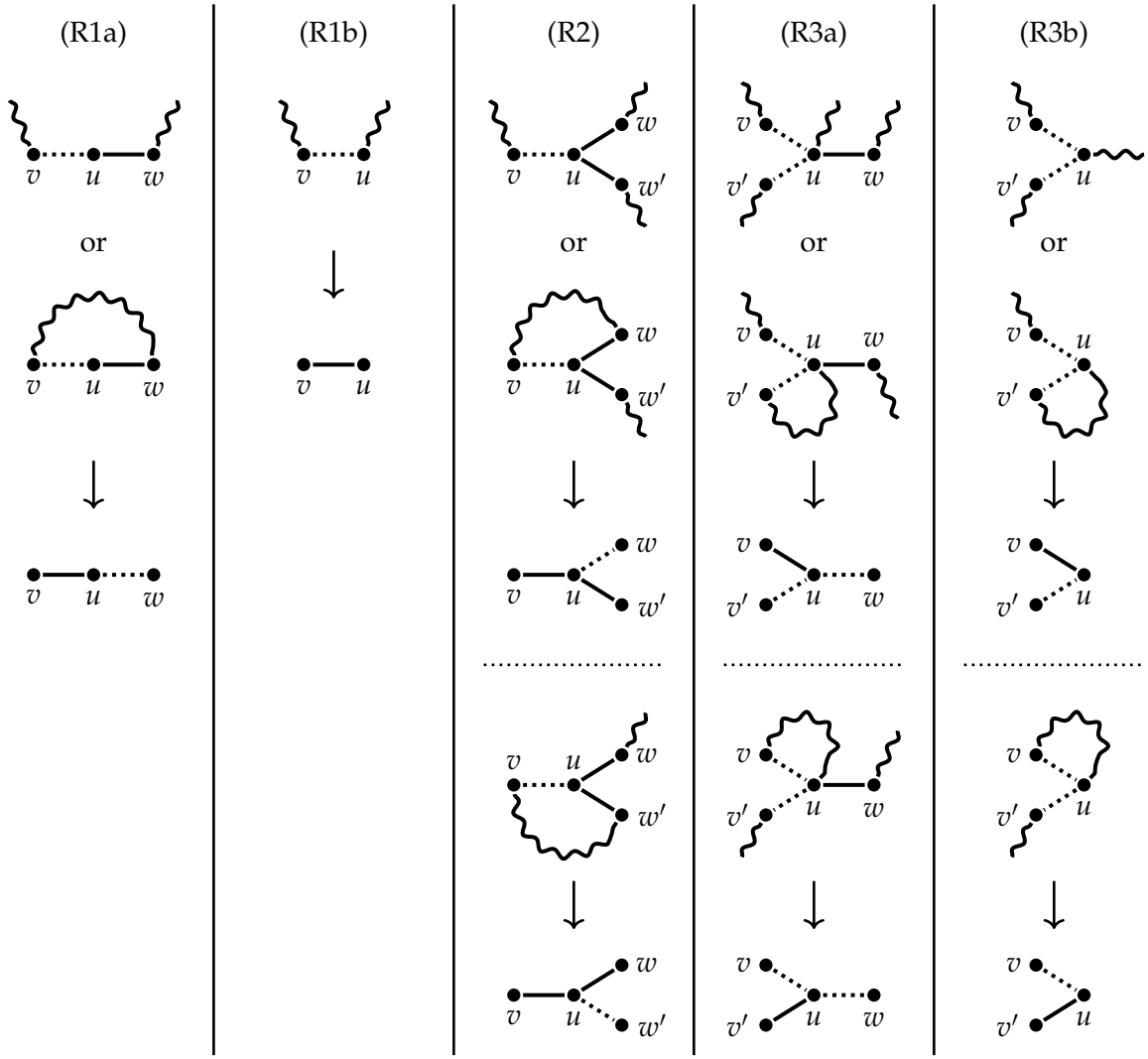

Figure 2: The case analysis for the observations about possible rotations of linear forests.

(a) If w is a neighbour of u on its path in F, then we can rotate one of the two endpoints v or v' to w using the pivot u.

(b) Alternatively, if u is an endpoint of F, we can reduce the number of paths of F by adding one of the two edges uv or uv' to F.

Now, consider the *boundary* $B(F,X) := N_G(C(F,X)) \setminus C(F,X)$ of $C(F,X)$. Since $C(F,X) = C(F',X)$ for all $F' \in \mathcal{F}(F,X)$, we also have $B(F,X) = B(F',X)$ for all $F' \in \mathcal{F}(F,X)$. It is useful to classify each vertex of $C(F,X)$ and $B(F,X)$ based on how many neighbours on its path in F are contained in the opposite set. So, define $C_i(F,X) := \{v \in C(F,X) : |N_F(v) \cap B(F,X)| = i\}$ and $B_i(F,X) := \{v \in B(F,X) : |N_F(v) \cap C(F,X)| = i\}$. Since every vertex is incident to at most two edges of F, we may assume that $i \in \{0,1,2\}$.

We start by proving some basic properties about these sets. The second of these properties, namely that $B_0(F,X) = \varnothing$, is a well-known fact for Pósa rotations.

Lemma 2.4. *Let F be a linear forest and let $X \subseteq \mathrm{End}(F)$. Then, for all $F' \in \mathcal{F}(F,X)$ and $i \in \{0,1,2\}$ we have $B_i(F,X) = B_i(F',X)$. Moreover, if F is a minimum linear forest, then $B_0(F,X) = \varnothing$.*

Proof. Note that F' is obtained by a sequence of rotations starting at F that never rotates any endpoint of X. It suffices to consider a single rotation of this sequence, so we may assume that

F' is a rotation of F that rotates an old endpoint v to a new endpoint w, using a pivot u. Note that $v, w \in C(F, X) = C(F', X)$. In particular, $|N_F(u) \cap C(F, X)| = |N_{F'}(u) \cap C(F', X)|$, and for each vertex in $B(F, X)$ apart from u, the edges of F and F' incident to that vertex are exactly the same. This shows that $B_i(F, X) = B_i(F', X)$.

Now, let F be a minimum linear forest. We want to show that $B_0(F, X) = \varnothing$, so suppose for a contradiction that there exists a vertex $u \in B_0(F, X)$. Let $v \in N_G(u) \cap C(F, X)$ be any neighbour of u in $C(F, X)$. Since $v \in C(F, X)$, there exists a linear forest $F' \in \mathcal{F}(F, X)$ with $X \cup \{v\} \subseteq \mathrm{End}(F')$. Because F' has the same number of paths as F, F' is a minimum linear forest, and we know that $u \in B_0(F, X) = B_0(F', X)$.

If u is an isolated vertex of F', then by (R1b) we could reduce the number of paths of F' by adding the edge uv, contradicting that F' is a minimum linear forest. If u is an endpoint of F' but has a neighbour $w \in N_{F'}(u)$ on its path in F', then by (R1a) we could rotate v to w using the pivot u. But this would imply that $w \in C(F', X)$, contradicting $u \in B_0(F', X)$. Finally, if u has two neighbours on its path in F', then by (R2) there exists at least one neighbour $w \in N_{F'}(u)$ such that we could rotate v to w using the pivot u, again contradicting $w \notin C(F', X)$. In all cases we get a contradiction, and so $B_0(F, X) = \varnothing$. □

We now show that we can reduce the size of $C(F, X)$ by fixing an endpoint at a vertex in $C(F, X)$ that has a neighbour in $B_1(F, X)$ which itself has many neighbours in $C(F, X)$.

Lemma 2.5. *Let F be a minimum linear forest and let $X \subseteq \mathrm{End}(F)$. Let $v \in \mathrm{End}(F) \setminus X$ and suppose that $u \in N_G(v) \cap B_1(F, X)$. Then, $|C(F, X \cup \{v\})| \leq |C(F, X)| - |N_G(u) \cap C(F, X)|$.*

Proof. We claim that $N_G(u) \cap C(F, X \cup \{v\}) = \varnothing$. Suppose for a contradiction that there exists a vertex $v' \in N_G(u) \cap C(F, X \cup \{v\})$. Since $v' \in C(F, X \cup \{v\})$, there exists a linear forest $F' \in \mathcal{F}(F, X \cup \{v\})$ with $X \cup \{v, v'\} \subseteq \mathrm{End}(F')$. Note that F' is a minimum linear forest. By Lemma 2.4 we know that $u \in B_1(F, X) = B_1(F', X)$.

If u is an endpoint of F', then by (R3b) we could reduce the number of paths of F' by adding one of the two edges uv or uv', contradicting that F' is a minimum linear forest. Otherwise, u has two neighbours on its path in F'. Since $u \in B_1(F', X)$, there must be a neighbour $w \in N_{F'}(u)$ that is not contained in $C(F', X)$. But then by (R3a) we could rotate one of the two endpoints v or v' to w using the pivot u, contradicting $w \notin C(F', X)$. In both cases we get a contradiction, and so $N_G(u) \cap C(F, X \cup \{v\}) = \varnothing$, as claimed. Since $C(F, X \cup \{v\}) \subseteq C(F, X)$, this implies that $|C(F, X \cup \{v\})| \leq |C(F, X)| - |N_G(u) \cap C(F, X)|$. □

To use this result to prove a good lower bound on $|C(F, X)|$, we need to find a vertex in $B_1(F, X)$ with many neighbours in $C(F, X)$. This is easy if there are many edges between $C(F, X)$ and $B_1(F, X)$. We will use the following double counting argument to provide a lower bound on the number of such edges. Here, for $A, B \subseteq V(G)$, we denote the number of edges of G connecting a vertex in A to a vertex in B by $e(A, B)$, and we write $e(A) := e(A, A)$.

Lemma 2.6. *Let F be a minimum linear forest of a d-regular graph G, and let $X \subseteq \mathrm{End}(F)$. Then,*

$$e(C(F, X), B_1(F, X)) \geq \frac{d}{2} \cdot |B_1(F, X)| + \frac{d}{2} \cdot |C_1(F, X)| + d \cdot |C_0(F, X)| - 2 \cdot e(C(F, X)).$$

Proof. By double counting the number of edges of F between $C(F, X)$ and $B(F, X)$, we get

$$|C_1(F, X)| + 2 \cdot |C_2(F, X)| = |B_1(F, X)| + 2 \cdot |B_2(F, X)|.$$

Note that the left side of this equation is the same as $2 \cdot |C(F, X)| - |C_1(F, X)| - 2 \cdot |C_0(F, X)|$. The total number of edges of G between $C(F, X)$ and $B(F, X)$ is $d \cdot |C(F, X)| - 2 \cdot e(C(F, X))$.

Because at most $d \cdot |B_2(F,X)|$ of these edges are incident with $B_2(F,X)$ and we know that $B_0(F,X) = \varnothing$ by Lemma 2.4, it follows that

$$\begin{aligned} e(C(F,X), B_1(F,X)) &\geq d \cdot |C(F,X)| - 2 \cdot e(C(F,X)) - d \cdot |B_2(F,X)| \\ &= \frac{d}{2} \cdot |B_1(F,X)| + \frac{d}{2} \cdot |C_1(F,X)| + d \cdot |C_0(F,X)| - 2 \cdot e(C(F,X)). \end{aligned}$$ □

We now first prove that bipartite graphs have small linear forests. In that setting, observe that if we fix all endpoints that are contained in one part of the bipartition, then $C(F,X)$ will be contained in the other part of the bipartition. This implies that $e(C(F,X)) = 0$. As a result, Lemma 2.6 shows that there exists a vertex in $B_1(F,X)$ with at least $(d+1)/2$ neighbours in $C(F,X)$. This allows us to apply Lemma 2.5 and use induction to complete the proof.

Theorem 2.7. *Every d-regular n-vertex bipartite graph has a spanning linear forest with at most $n/(d+1)$ paths.*

Proof. Let G be a d-regular bipartite graph on n vertices, and let U be one part of its bipartition. We claim that if F is a minimum linear forest of G and $X \subseteq \mathrm{End}(F)$ satisfies $\mathrm{End}(F) \setminus X \subseteq U$, then $C(F,X) \subseteq U$ and $|C(F,X)| \geq |\mathrm{End}(F) \setminus X| \cdot (d+1)/2$. We prove the claim by induction on $\ell := |\mathrm{End}(F) \setminus X|$. The claim clearly holds for $\ell = 0$, so suppose that $\ell > 0$.

Observe that if a rotation of F rotates an old endpoint $v \in U$ to a new endpoint w, then w is at distance two from v. Since G is bipartite, it follows that $w \in U$. This shows that $C(F,X) \subseteq U$. In particular, $e(C(F,X)) = 0$. Since $\varnothing \neq \mathrm{End}(F) \setminus X \subseteq C_1(F,X) \cup C_0(F,X)$, this implies by Lemma 2.6 that $e(C(F,X), B_1(F,X)) > |B_1(F,X)| \cdot d/2$. In particular, there exists a vertex $u \in B_1(F,X)$ with $|N_G(u) \cap C(F,X)| \geq (d+1)/2$.

Let $v \in N_G(u) \cap C(F,X)$ be arbitrary, let $X' := X \cup \{v\}$, and let $F' \in \mathcal{F}(F,X)$ be any linear forest with $X' \subseteq \mathrm{End}(F')$. Using the fact that $C(F,X) = C(F',X)$, Lemma 2.5 implies that

$$|C(F,X)| = |C(F',X)| \geq |C(F',X')| + |N_G(u) \cap C(F',X)| \geq |C(F',X')| + \frac{d+1}{2}.$$

By the induction hypothesis, it then follows that

$$|C(F,X)| \geq |C(F',X')| + \frac{d+1}{2} \geq \frac{d+1}{2} \cdot |\mathrm{End}(F') \setminus X'| + \frac{d+1}{2} = \frac{d+1}{2} \cdot |\mathrm{End}(F) \setminus X|.$$

This proves the claim. Finally, if $V(G) = V \cup W$ is the bipartition of G, the claim implies that

$$\begin{aligned} |\mathrm{End}(F)| &= |\mathrm{End}(F) \setminus (\mathrm{End}(F) \cap V)| + |\mathrm{End}(F) \setminus (\mathrm{End}(F) \cap W)| \\ &\leq \frac{2}{d+1} \cdot |C(F, \mathrm{End}(F) \cap V)| + \frac{2}{d+1} \cdot |C(F, \mathrm{End}(F) \cap W)| \leq \frac{2}{d+1} \cdot n. \end{aligned}$$

Since the number of endpoints of a linear forest is twice the number of its paths, this implies that F has at most $n/(d+1)$ paths. □

In non-bipartite graphs, we will define an iterative process that fixes endpoints of F one by one at vertices that have a high degree within $C(F,X)$. If this reduces the size of $C(F,X)$ sufficiently, we are done by induction. Otherwise, we show that $C(F,X)$ spans only few edges, and so we can finish the proof as in the bipartite case.

Proof of Lemma 2.3. Let G be a d-regular graph, let F be a minimum linear forest of G, and let $X \subseteq \mathrm{End}(F)$. We proceed by induction on $\ell := |\mathrm{End}(F) \setminus X|$. The result clearly holds for $\ell = 0$, so suppose that $\ell > 0$.

Consider the following iterative process. Let $X_1 := X$ and let $F_1 := F$. Then, given X_i and F_i, let $\Delta_i := \Delta(C(F_i, X_i))$ be the maximum degree of the subgraph induced by $C(F_i, X_i)$, and choose any vertex $v_i \in C(F_i, X_i)$ with degree Δ_i in $C(F_i, X_i)$. Define $X_{i+1} := X_i \cup \{v_i\}$ and let $F_{i+1} \in \mathcal{F}(F_i, X_i)$ be a linear forest with $X_{i+1} \subseteq \mathrm{End}(F_{i+1})$. This process gives a chain of rotation-components

$$C(F, X) = C(F_1, X_1) \supseteq C(F_2, X_2) \supseteq C(F_3, X_3) \supseteq \cdots \supseteq C(F_{\ell+1}, X_{\ell+1}) = \varnothing.$$

Since the maximum degree of the subgraph induced by $C(F_i, X_i)$ is Δ_i, this implies that

$$e(C(F, X)) \leq \sum_{i=1}^{\ell} |C(F_i, X_i) \setminus C(F_{i+1}, X_{i+1})| \cdot \Delta_i. \tag{1}$$

If $|C(F_{i+1}, X_{i+1})| \leq |C(F, X)| - i \cdot (d+1)/4$, then we know by the induction hypothesis that

$$\begin{aligned} |C(F, X)| &\geq \frac{d+1}{4} \cdot i + |C(F_{i+1}, X_{i+1})| \\ &\geq \frac{d+1}{4} \cdot i + \frac{d+1}{4} \cdot |\mathrm{End}(F_{i+1}) \setminus X_{i+1}| = \frac{d+1}{4} \cdot |\mathrm{End}(F) \setminus X|. \end{aligned}$$

So, we may assume that $|C(F_{i+1}, X_{i+1})| > |C(F, X)| - i \cdot (d+1)/4$. Since $\Delta_1 \geq \cdots \geq \Delta_k$, this implies that

$$\sum_{i=1}^{\ell} |C(F_i, X_i) \setminus C(F_{i+1}, X_{i+1})| \cdot \Delta_i < \sum_{i=1}^{\ell} \frac{d+1}{4} \cdot \Delta_i = \frac{d+1}{4} \sum_{i=1}^{\ell} |N_G(v_i) \cap C(F_i, X_i)|. \tag{2}$$

We claim that

$$|\mathrm{End}(F_{\ell+1}) \setminus X| + \sum_{i=1}^{\ell} |N_G(v_i) \cap C(F_i, X_i)| \leq |C_1(F_{\ell+1}, X)| + 2 \cdot |C_0(F_{\ell+1}, X)|. \tag{3}$$

To see this, observe that every vertex on the left side of the inequality is contained in $C(F, X) = C(F_{\ell+1}, X)$, and consider the following cases:

(a) Suppose that $u \in C_2(F_{\ell+1}, X)$. Clearly $u \notin \mathrm{End}(F_{\ell+1}) \setminus X$ since u has two neighbours on its path in $F_{\ell+1}$. Moreover, we claim that $u \notin N_G(v_i) \cap C(F_i, X_i)$ for all i. Indeed, if $u \in N_G(v_i)$ for some i, then by (R2) there exists at least one neighbour $w \in N_{F_{\ell+1}}(u)$ such that we could rotate v_i to w using the pivot u. This would imply $w \in C(F_{\ell+1}, X)$, contradicting $u \in C_2(F_{\ell+1}, X)$.

(b) If $u \in C_1(F_{\ell+1}, X)$, we claim that either $u \in \mathrm{End}(F_{\ell+1}) \setminus X$ and $u \notin N_G(v_i) \cap C(F_i, X_i)$ for all i, or $u \notin \mathrm{End}(F_{\ell+1}) \setminus X$ and $u \in N_G(v_i) \cap C(F_i, X_i)$ for up to one value of i. Indeed, let $w \in N_{F_{\ell+1}}(u)$ be the unique neighbour of u in $B(F_{\ell+1}, X)$. If $u \in \mathrm{End}(F_{\ell+1}) \setminus X$ and $u \in N_G(v_i)$ for some i, then by (R1a) we could rotate v_i to w using the pivot u, contradicting $w \notin C(F_{\ell+1}, X)$. Moreover, if $u \in N_G(v_i) \cap N_G(v_j)$ for some $i \neq j$, then by (R3a) we could rotate one of the two endpoints v_i or v_j to w using the pivot u, again contradicting $w \notin C(F_{\ell+1}, X)$. This proves the claim.

(c) If $u \in C_0(F_{\ell+1}, X)$, we claim that either $u \in \mathrm{End}(F_{\ell+1}) \setminus X$ and $u \in N_G(v_i) \cap C(F_i, X_i)$ for up to one value of i, or $u \notin \mathrm{End}(F_{\ell+1}) \setminus X$ and $u \in N_G(v_i) \cap C(F_i, X_i)$ for up to two values of i. Indeed, if $u \in \mathrm{End}(F_{\ell+1}) \setminus X$ and $u \in N_G(v_i) \cap N_G(v_j)$ for some $i \neq j$, then by (R3b) we could reduce the number of paths of $F_{\ell+1}$ by adding one of the two edges uv_i or uv_j, contradicting that $F_{\ell+1}$ is a minimum linear forest. Moreover, if $u \in N_G(v_i) \cap N_G(v_j)$ for some $i < j$, then $u \notin C(F_k, X_k)$ for any $k > j$. Otherwise, we could find a linear forest $F' \in \mathcal{F}(F_k, X_k)$ with $\{u, v_i, v_j\} \subseteq \mathrm{End}(F')$, but then by (R3b) we could reduce the number of paths of F' by adding one of the two edges uv_i or uv_j, contradicting that F' is a minimum linear forest. This proves the claim.

In all cases, the contribution of a vertex to the left side of inequality (3) is at most as large as its contribution to the right side, which proves this inequality.

We may assume that $|\mathrm{End}(F_{\ell+1}) \setminus X| \cdot (d+1)/4 = |\mathrm{End}(F) \setminus X| \cdot (d+1)/4 \geq |C(F,X)|$, as we would be done otherwise. Since $C(F,X) = C(F_{\ell+1},X)$, we therefore have

$$\frac{d+1}{2} \cdot |\mathrm{End}(F_{\ell+1}) \setminus X| \geq |C(F,X)| = |C(F_{\ell+1},X)| \geq \frac{1}{2} \cdot |C_1(F_{\ell+1},X)| + |C_0(F_{\ell+1},X)|. \quad (4)$$

By combining the inequalities (1) to (4) and using again that $C(F,X) = C(F_{\ell+1},X)$, we get that

$$\begin{aligned} 2 \cdot e(C(F_{\ell+1},X)) &< \frac{d+1}{2} \sum_{i=1}^{\ell} |N_G(v_i) \cap C(F_i,X_i)| \\ &\leq \frac{d+1}{2} \cdot |C_1(F_{\ell+1},X)| + (d+1) \cdot |C_0(F_{\ell+1},X)| - \frac{d+1}{2} \cdot |\mathrm{End}(F_{\ell+1}) \setminus X| \\ &\leq \frac{d}{2} \cdot |C_1(F_{\ell+1},X)| + d \cdot |C_0(F_{\ell+1},X)|. \end{aligned}$$

By Lemma 2.6, it follows that $e(C(F_{\ell+1},X), B_1(F_{\ell+1},X)) > |B_1(F_{\ell+1},X)| \cdot d/2$. Since we know that $F_{\ell+1} \in \mathcal{F}(F,X)$, by Lemma 2.4 this is equivalent to $e(C(F,X), B_1(F,X)) > |B_1(F,X)| \cdot d/2$. As in the proof of Theorem 2.7, we can then find a vertex $v \in C(F,X)$ such that if $X' := X \cup \{v\}$ and $F' \in \mathcal{F}(F,X)$ is a linear forest with $X' \subseteq \mathrm{End}(F')$, then $|C(F,X)| \geq |C(F',X')| + (d+1)/2$. By the induction hypothesis, it then follows that

$$|C(F,X)| \geq |C(F',X')| + \frac{d+1}{2} \geq \frac{d+1}{4} \cdot |\mathrm{End}(F') \setminus X'| + \frac{d+1}{2} \geq \frac{d+1}{4} \cdot |\mathrm{End}(F) \setminus X|. \quad \square$$

3 Linear arboricity up to logarithmic error

We now use the results from the previous section to show that the linear arboricity of any graph G on n vertices with maximum degree Δ is at most $\Delta/2 + \mathcal{O}(\log n)$. We achieve this by repeatedly choosing linear forests with few paths and removing them from the graph. If the maximum degree of the remaining graph after $\Delta/2$ iterations is $\mathcal{O}(\log n)$, we can decompose the remaining edges into $\mathcal{O}(\log n)$ linear forests, say with Vizing's theorem, to obtain the desired upper bound on the linear arboricity.

Note that the degree of a vertex after the first $\Delta/2$ iterations can only be larger than $\mathcal{O}(\log n)$ if the vertex appeared as an endpoint in more than $\mathcal{O}(\log n)$ linear forests during those iterations. To avoid this, we will choose the linear forests in each iteration randomly in such a way that every vertex has a low probability of being an endpoint. Then, we can simply use standard concentration inequalities and a union bound to bound the probability that any vertex is an endpoint in more than $\mathcal{O}(\log n)$ linear forests, which completes the proof.

We start by showing that for every d-regular graph G, there exists a distribution on the minimum linear forests of G such that every vertex has a low probability of being an endpoint. Recall that $\mathrm{End}(F)$ denotes the multiset of endpoints of a linear forest F. In the proof, we will make extensive use of the notation introduced in Section 2.

Lemma 3.1. *Let G be a d-regular graph. Then, we can choose a random minimum linear forest F of G in such a way that for all vertices $v \in V(G)$ it holds that $\mathbb{P}(v \in \mathrm{End}(F)) \leq 16/(d+1)$.*

Proof. We claim that for all minimum linear forests F of G and all $X \subseteq \mathrm{End}(F)$, we can choose a random linear forest $F' \in \mathcal{F}(F,X)$ in such a way that for all vertices $v \in V(G)$ it holds that $\mathbb{P}(v \in \mathrm{End}(F') \setminus X) \leq 16/(d+1)$. For $X = \varnothing$, this implies the result. We prove the claim by induction on $\ell := |\mathrm{End}(F) \setminus X|$. The claim clearly holds for $\ell = 0$, so suppose that $\ell > 0$.

For all $u \in C(F,X)$, let $F_u \in \mathcal{F}(F,X)$ be a linear forest with $X \cup \{u\} \subseteq \mathrm{End}(F_u)$. By the induction hypothesis, we know that we can choose a random linear forest $F'_u \in \mathcal{F}(F_u, X \cup \{u\})$ such that for all vertices $v \in V(G)$ it holds that $\mathbb{P}(v \in \mathrm{End}(F'_u) \setminus (X \cup \{u\})) \leq 16/(d+1)$. Choose a random linear forest $F'' \in \mathcal{F}(F,X)$ by first choosing $u \in C(F,X)$ uniformly at random and then choosing the random linear forest F'_u. By the linearity of expectation, we know that

$$\sum_{v \in C(F,X)} \mathbb{P}(v \in \mathrm{End}(F'') \setminus X) = \mathbb{E}(|\mathrm{End}(F'') \setminus X|) = |\mathrm{End}(F) \setminus X|.$$

By Lemma 2.3, we also have $|C(F,X)| \geq |\mathrm{End}(F) \setminus X| \cdot (d+1)/4$. This implies that the set $U := \{v \in C(F,X) : \mathbb{P}(v \in \mathrm{End}(F'') \setminus X) \leq 8/(d+1)\}$ must contain at least half of the vertices of $C(F,X)$.

Now, choose the random linear forest $F' \in \mathcal{F}(F,X)$ by first choosing $u \in U$ uniformly at random and then choosing the random linear forest F'_u. For all vertices $v \in V(G) \setminus U$, we have

$$\begin{aligned}\mathbb{P}(v \in \mathrm{End}(F') \setminus X) &= \frac{1}{|U|} \sum_{u \in U} \mathbb{P}(v \in \mathrm{End}(F'_u) \setminus X) \\ &= \frac{1}{|U|} \sum_{u \in U} \mathbb{P}(v \in \mathrm{End}(F'_u) \setminus (X \cup \{u\})) \leq \frac{16}{d+1}.\end{aligned}$$

For all vertices $v \in U$, we have

$$\begin{aligned}\mathbb{P}(v \in \mathrm{End}(F') \setminus X) &= \frac{1}{|U|} \sum_{u \in U} \mathbb{P}(v \in \mathrm{End}(F'_u) \setminus X) \\ &\leq \frac{1}{|U|} \sum_{u \in C(F,X)} \mathbb{P}(v \in \mathrm{End}(F'_u) \setminus X) \\ &= \frac{|C(F,X)|}{|U|} \cdot \mathbb{P}(v \in \mathrm{End}(F'') \setminus X) \leq \frac{16}{d+1}.\end{aligned}$$

□

We can now prove Theorem 1.2. As described earlier in this section, our strategy is to repeatedly remove linear forests from the graph G using Lemma 3.1. Since that result only applies to regular graphs, we need to regularise G after every step, but the remainder of the proof works exactly as outlined before.

Proof of Theorem 1.2. Let G be a graph on n vertices with maximum degree Δ. Let $d := \lceil \Delta/2 \rceil$. By adding vertices and edges to G, we may extend G to a $2d$-regular graph G_d.

Consider the following iterative process. Start with the $2d$-regular graph G_d. Then, given a $2i$-regular graph G_i, let $C_i \subseteq G_i$ be a 2-regular spanning subgraph of G_i. This subgraph exists because G_i is regular with even degree. Choose a random minimum linear forest F_i of G_i according to Lemma 3.1 so that for all vertices $v \in V(G_i)$ it holds that $\mathbb{P}(v \in \mathrm{End}(F_i)) \leq 16/(2i+1)$. Then, let $E_i \subseteq E(C_i)$ be the subset of all edges of C_i that are incident to some endpoint of F_i. Note that every vertex of G_i is incident to at least two edges of $F_i \cup E_i$. This implies that the subgraph $G_i \setminus (F_i \cup E_i)$ has maximum degree at most $2(i-1)$ and can therefore be extended to a $2(i-1)$-regular graph G_{i-1}.

We iterate this process until we reach the empty graph G_0. Note that the restriction of each linear forest F_i to the initial graph G is a linear forest. We also have that every edge of G is contained either in one of the linear forests F_i or in one of the sets E_i. So, the degree of a vertex $v \in V(G)$ in $G' := G \setminus (\bigcup_{i=1}^d F_i)$ is at most the number of edges of $\bigcup_{i=1}^d E_i$ that are incident to v.

Let D_v^i be the indicator random variable of the event that v is incident to an edge of E_i. Since v is incident to at most two edges of E_i, the degree of v in G' is at most $D_v := \sum_{i=1}^d 2D_v^i$. Moreover,

v can only be incident to an edge of E_i if v or one of its two neighbours in C_i is an endpoint of F_i. By construction, the probability that a vertex is an endpoint of F_i is at most $16/(2i+1) \leq 8/i$, regardless of the choices of $F_{i+1}, \ldots, F_d$. It follows that $\mathbb{P}(D_v^i = 1) \leq 24/i$, even conditioned on the values of $D_v^{i+1}, \ldots, D_v^d$. So, the random variables $D_v^1, \ldots, D_v^d$ are stochastically dominated by a collection of independent Bernoulli random variables $X_1, \ldots, X_d$ with $\mathbb{P}(X_i = 1) = 24/i$. In particular, $X := \sum_{i=1}^d 2X_i$ satisfies $D_v \leq X$. Since $\mu := \mathbb{E}(X) \leq 48(\log d + 1)$, a standard Chernoff bound with $\delta := 48(\log n + 1)/\mu \geq 1$ implies that

$$\mathbb{P}(D_v > (1+\delta)\mu) \leq \mathbb{P}(X > (1+\delta)\mu) \leq e^{-\frac{\delta^2\mu}{2+\delta}} \leq e^{-\frac{\delta\mu}{3}} = e^{-16(\log n+1)} < \frac{1}{n}.$$

So, a union bound shows that with positive probability, $D_v \leq (1+\delta)\mu = \mathcal{O}(\log n)$ for all $v \in V(G)$. In that case, the maximum degree of G' is at most $\mathcal{O}(\log n)$ which implies that G' can be decomposed into at most $\mathcal{O}(\log n)$ linear forests. Together with $F_1, \ldots, F_d$, this provides a decomposition of G into at most $d + \mathcal{O}(\log n) = \Delta/2 + \mathcal{O}(\log n)$ linear forests. □

The proof for fractional linear arboricity follows almost the same strategy, except that it ignores vertices with degree larger than $\mathcal{O}(\log \Delta)$ in the remaining graph. This produces a family of $\Delta/2 + \mathcal{O}(\log \Delta)$ linear forests such that each edge is contained in that family with high probability. If we then choose a uniformly random linear forest from that family, this proves the upper bound on the fractional linear arboricity.

Proof of Theorem 1.3. Let G be a graph on n vertices with maximum degree Δ. Define d, F_i, G', D_v, and X as in the proof of Theorem 1.2. Since $\mu := \mathbb{E}(X) \leq 48(\log d + 1)$, a standard Chernoff bound with $\delta := 48(\log d + 1)/\mu \geq 1$ implies that

$$\mathbb{P}(D_v > (1+\delta)\mu) \leq \mathbb{P}(X > (1+\delta)\mu) \leq e^{-\frac{\delta^2\mu}{2+\delta}} \leq e^{-\frac{\delta\mu}{3}} = e^{-16(\log d+1)} < \frac{1}{d^2}.$$

So, the probability that v has degree larger than $c := (1+\delta)\mu = \mathcal{O}(\log d)$ in G' is less than $1/d^2$. Let $E \subseteq E(G')$ be the subset of all edges of G' that are incident to a vertex of degree larger than c in G'. In particular, every edge of G is contained in E with probability at most $2/d^2$. Moreover, the graph $G'' := G' \setminus E$ has maximum degree c which implies that G'' can be decomposed into at most $c+1$ linear forests $F_1', \ldots, F_{c+1}'$. Choose a linear forest $F \in \{F_1, \ldots, F_d, F_1', \ldots, F_{c+1}'\}$ uniformly at random. Then, for every edge $e \in E(G)$ of G, we have

$$\mathbb{P}(e \in F) = \mathbb{P}(e \notin E) \cdot \mathbb{P}(e \in F \mid e \notin E) \geq \left(1 - \frac{2}{d^2}\right) \frac{1}{d+c+1} \geq \frac{1}{d + \mathcal{O}(\log d)}.$$

So, the fractional linear arboricity of G is at most $d + \mathcal{O}(\log d) = \Delta/2 + \mathcal{O}(\log \Delta)$. □

4 Open problems

By proving the conjecture of Feige and Fuchs, we removed a major obstacle towards resolving the Linear Arboricity Conjecture. However, the bound on the linear arboricity that we obtained in Theorem 1.2 has an error term of $\mathcal{O}(\log n)$, which depends on n. A natural next step is to replace this error term with a term polynomial in $\log \Delta$. To achieve this, one would like to replace the union bound that we use with a variant of the Lovász Local Lemma or a similar technique. Unfortunately, the dependency structure of the corresponding bad events is extremely intricate, and so it is not obvious how to make such a strategy work. We believe that it would already be interesting to provide such an improvement for bipartite graphs.

Problem 4.1. *Do all bipartite graphs G with maximum degree Δ satisfy* $\mathrm{la}(G) \leq \Delta/2 + (\log \Delta)^{\mathcal{O}(1)}$*?*

In bipartite graphs, we can sample a random linear forest more easily than for non-bipartite graphs. Indeed, as shown in the proof of Theorem 2.7, we can always find a vertex $u \in B_1(F, X)$ with at least $(d+1)/2$ neighbours in $C(F, X)$. If we then fix an endpoint of the linear forest at a uniformly random neighbour of u in $C(F, X)$, all neighbours of u will afterwards be removed from $C(F, X)$. Repeating this process results in a random linear forest where every vertex is an endpoint with probability at most $2/(d+1)$.

When we iteratively remove linear forests from a graph, in later iterations we would like to fix endpoints only at vertices that have not been endpoints in too many previous iterations. This would be possible if we never fix too many endpoints in the neighbourhood of a vertex since we could then simply avoid these vertices when generating a new linear forest. Note that the above random linear forest has a very low probability of fixing many endpoints in the neighbourhood of a given vertex. However, we cannot apply the Lovász Local Lemma because the vertex u as well as the subset of its neighbours that are contained in $C(F, X)$ can depend on previously fixed endpoints, creating large dependencies. Even just the fact that we only have a subset of the neighbours of u available when we fix an endpoint seems to make this problem difficult. More precisely, even if we knew in advance the sequence of vertices u obtained while fixing endpoints, but not the subset of available neighbours, we do not know how to ensure that we fix at most $(\log \Delta)^{\mathcal{O}(1)}$ endpoints in the neighbourhood of any vertex.

We believe that the following online matching game exhibits the same difficulty and could be of independent interest. Suppose that G is a bipartite graph with bipartition $U \cup W$ such that every vertex in U has degree $c \cdot d$ while every vertex in W has degree at most d. Let $U = \{u_1, \ldots, u_n\}$. Now, for $i = 1, \ldots, n$, we receive a subset $U_i \subseteq N(u_i)$ of size $c \cdot d/2$, and we immediately have to match u_i to an unmatched vertex in U_i. If the sets U_i were determined in advance and $c \geq 2$, Hall's theorem would imply that we could find a perfect matching.

Problem 4.2. *If c is a sufficiently large constant, is there a way to construct a perfect matching in this online matching game where an adversary chooses the subsets U_i based on the matching of $u_1, \ldots, u_{i-1}$?*

An alternative approach to replacing the error term of $\mathcal{O}(\log n)$ with $(\log \Delta)^{\mathcal{O}(1)}$ could be to prove stronger properties about how $C(F, X)$ evolves when fixing endpoints. This might be enough to recover the limited dependencies necessary to apply the Lovász Local Lemma.

In terms of the existence of linear forests with few paths, the conjecture of Magnant and Martin remains open up to a factor of two, but we believe that our techniques could be sufficient to eventually resolve this conjecture completely. Note that even for regular bipartite graphs where we proved this conjecture, our result still differs by a factor of roughly two from the best lower bound of $n/(2d)$ that comes from a disjoint union of complete bipartite graphs.

There are also several potential generalisations of the conjecture of Feige and Fuchs that remain open. For instance, this conjecture might generalise to cycles[8] and to directed regular graphs.

Conjecture 4.3. *Every directed d-regular n-vertex graph has a collection of $\mathcal{O}(n/d)$ vertex-disjoint cycles that cover all vertices of the graph.*

The authors of the current paper [CDG+25] very recently proved that there exists a collection of $\mathcal{O}((n \log d)/d)$ cycles that satisfy this, which is the current best upper bound. In addition, they provided a randomised polynomial-time algorithm that computes such a collection. As a corollary, they also obtained an efficient algorithm for computing a tour of length $(1 + \mathcal{O}((\log d)/d)) \cdot n$ in a connected d-regular n-vertex graph.

In contrast, all of the results in this paper are non-constructive, and so we do not obtain efficient algorithms. This raises the question whether there are efficient algorithms for finding spanning

[8] Here, an individual edge is considered to be a cycle of length two.

linear forests with $\mathcal{O}(n/d)$ paths in d-regular n-vertex graphs. This would also imply the existence of efficient algorithms for computing a tour of length $(1 + \mathcal{O}(1/d)) \cdot n$ in connected regular graphs. An efficient algorithm would likely improve our understanding of small linear forests significantly, which may lead to further progress on the Linear Arboricity Conjecture.

Problem 4.4. *Is there an algorithm that can efficiently compute a spanning linear forest of a d-regular n-vertex graph with at most $\mathcal{O}(n/d)$ paths?*

In particular, what is the computational complexity of finding a rotation-component $C(F, X)$? This could plausibly be NP-hard. In extremal combinatorics it is unusual for a proof to rely on hard-to-compute objects, but this does occur with algebraic methods such as the method of interlacing polynomials [MSS15a, MSS15b] and topological methods such as applications of the Borsuk-Ulam Theorem [Mat03]. Note that for Hamiltonian cubic graphs where minimum linear forests are Hamilton paths, it is known that the number of rotations required to rotate one endpoint to some fixed vertex of the graph may be exponential in n. This follows, for example, from a construction in [BS22].

References

[AEH80] Jin Akiyama, Geoffrey Exoo, and Frank Harary (1980). Covering and packing in graphs. III. Cyclic and acyclic invariants. *Mathematica Slovaca* **30**(4), 405–417. ↑2

[AEH81] Jin Akiyama, Geoffrey Exoo, and Frank Harary (1981). Covering and packing in graphs. IV. Linear arboricity. *Networks* **11**(1), 69–72. ↑3

[Alo88] Noga Alon (1988). The linear arboricity of graphs. *Israel Journal of Mathematics* **62**(3), 311–325. ↑3

[AS16] Noga Alon and Joel H. Spencer (2016). The probabilistic method. Fourth edn. (John Wiley & Sons). ↑1, 3

[BKLO16] Ben Barber, Daniela Kühn, Allan Lo, and Deryk Osthus (2016). Edge-decompositions of graphs with high minimum degree. *Advances in Mathematics* **288**, 337–385. ↑2

[BM24] Matija Bucić and Richard Montgomery (2024). Towards the Erdős-Gallai cycle decomposition conjecture. *Advances in Mathematics* **437**, Paper No. 109434. ↑2

[BS22] Marcin Briański and Adam Szady (2022). A short note on graphs with long Thomason chains. *Discrete Mathematics* **345**(1), Paper No. 112624. ↑16

[CDG+25] Micha Christoph, Nemanja Draganić, António Girão, Eoin Hurley, Lukas Michel, and Alp Müyesser (2025). Cycle-factors of regular graphs via entropy. *2025 IEEE 66th Annual Symposium on Foundations of Computer Science*. To appear. ↑4, 15

[CKL+16] Béla Csaba, Daniela Kühn, Allan Lo, Deryk Osthus, and Andrew Treglown (2016). Proof of the 1-factorization and Hamilton decomposition conjectures. *Memoirs of the American Mathematical Society* **244**(1154), v+164. ↑2

[DCS24] Nemanja Draganić, David Munhá Correia, and Benny Sudakov (2024). Pancyclicity of hamiltonian graphs. *Journal of the European Mathematical Society* . ↑4

[DGCS25] Nemanja Draganić, Stefan Glock, David Munhá Correia, and Benny Sudakov (2025). Optimal Hamilton covers and linear arboricity for random graphs. *Proceedings of the American Mathematical Society* **153**(3), 921–935. ↑3

[Dir52] Gabriel Andrew Dirac (1952). Some theorems on abstract graphs. *Proceedings of the London Mathematical Society (3)* **2**, 69–81. ↑4

[DMC+24] Nemanja Draganić, Richard Montgomery, David Munhá Correia, Alexey Pokrovskiy, and Benny Sudakov (2024). Hamiltonicity of expanders: optimal bounds and applications. arXiv:2402.06603. ↑4

[DP21] Michelle Delcourt and Luke Postle (2021). Progress towards Nash-Williams' conjecture on triangle decompositions. *Journal of Combinatorial Theory, Series B* **146**, 382–416. ↑2

[EGP66] Paul Erdős, A. W. Goodman, and Lajos Pósa (1966). The representation of a graph by set intersections. *Canadian Journal of Mathematics* **18**, 106–112. ↑2

[Eno81] Hikoe Enomoto (1981). The linear arboricity of 5-regular graphs. Technical report, Department of Information Science, University of Tokyo. ↑3

[EP84] Hikoe Enomoto and Bernard Péroche (1984). The linear arboricity of some regular graphs. *Journal of Graph Theory* **8**(2), 309–324. ↑3

[ERT80] Paul Erdős, Arthur L. Rubin, and Herbert Taylor (1980). Choosability in graphs. *Congressus Numerantium* **26**, 125–157. ↑2

[FF22] Uriel Feige and Ella Fuchs (2022). On the path partition number of 6-regular graphs. *Journal of Graph Theory* **101**(3), 345–378. ↑3, 4

[FFJ20] Asaf Ferber, Jacob Fox, and Vishesh Jain (2020). Towards the linear arboricity conjecture. *Journal of Combinatorial Theory, Series B* **142**, 56–79. ↑3

[FRS14] Uriel Feige, R. Ravi, and Mohit Singh (2014). Short tours through large linear forests. *Integer Programming and Combinatorial Optimization, IPCO 2014, Lecture Notes in Computer Science*, vol. 8494, 273–284. ↑3, 4

[GCS24] Stefan Glock, David Munhá Correia, and Benny Sudakov (2024). Hamilton cycles in pseudorandom graphs. *Advances in Mathematics* **458**, Paper No. 109984. ↑4

[GKLO23] Stefan Glock, Daniela Kühn, Allan Lo, and Deryk Osthus (2023). The existence of designs via iterative absorption: hypergraph F-designs for arbitrary F. *Memoirs of the American Mathematical Society* **284**(1406), v+131. ↑1

[GKO16] Stefan Glock, Daniela Kühn, and Deryk Osthus (2016). Optimal path and cycle decompositions of dense quasirandom graphs. *Journal of Combinatorial Theory, Series B* **118**, 88–108. ↑3

[GL21] Vytautas Gruslys and Shoham Letzter (2021). Cycle partitions of regular graphs. *Combinatorics, Probability and Computing* **30**(4), 526–549. ↑4

[GS05] Stefanie Gerke and Angelika Steger (2005). The sparse regularity lemma and its applications. *Surveys in combinatorics 2005, London Mathematical Society Lecture Note Series*, vol. 327, 227–258. ↑2

[GS24] Yuping Gao and Songling Shan (2024). Linear arboricity of graphs with large minimum degree. arXiv:2405.18494. ↑3

[Gul86a] FILIP GULDAN (1986). The linear arboricity of 10-regular graphs. *Mathematica Slovaca* **36**(3), 225–228. ↑3

[Gul86b] FILIP GULDAN (1986). Some results on linear arboricity. *Journal of Graph Theory* **10**(4), 505–509. ↑3

[Han18] JIE HAN (2018). On vertex-disjoint paths in regular graphs. *Electronic Journal of Combinatorics* **25**(2), Paper No. 2.12. ↑4

[Har70] FRANK HARARY (1970). Covering and packing in graphs. I. *Annals of the New York Academy of Sciences* **175**(1), 198–205. ↑2

[Kah00] JEFF KAHN (2000). Asymptotics of the list-chromatic index for multigraphs. *Random Structures & Algorithms* **17**(2), 117–156. ↑2

[Kee14] PETER KEEVASH (2014). The existence of designs. arXiv:1401.3665. ↑1

[KKK+23] DONG YEAP KANG, TOM KELLY, DANIELA KÜHN, ABHISHEK METHUKU, and DERYK OSTHUS (2023). Graph and hypergraph colouring via nibble methods: A survey. *European Congress of Mathematics*, vol. 21, 771–823. ↑1

[KKOT19] JAEHOON KIM, DANIELA KÜHN, DERYK OSTHUS, and MYKHAYLO TYOMKYN (2019). A blow-up lemma for approximate decompositions. *Transactions of the American Mathematical Society* **371**(7), 4655–4742. ↑2

[Koh97] Y. KOHAYAKAWA (1997). Szemerédi's regularity lemma for sparse graphs. *Foundations of Computational Mathematics (Rio de Janeiro, 1997)*, 216–230. ↑2

[KS22] PETER KEEVASH and KATHERINE STADEN (2022). The generalised Oberwolfach problem. *Journal of Combinatorial Theory, Series B* **152**, 281–318. ↑1, 2

[KS25] PETER KEEVASH and KATHERINE STADEN (2025). Ringel's tree packing conjecture in quasirandom graphs. *Journal of the European Mathematical Society (JEMS)* **27**(5), 1769–1826. ↑1, 2

[KSSS24] MATTHEW KWAN, ASHWIN SAH, MEHTAAB SAWHNEY, and MICHAEL SIMKIN (2024). High-girth Steiner triple systems. *Annals of Mathematics* **200**(3), 1059–1156. ↑1

[LMS25] SHOHAM LETZTER, ABHISHEK METHUKU, and BENNY SUDAKOV (2025). Nearly Hamilton cycles in sublinear expanders, and applications. arXiv:2503.07147. ↑4

[LP23] RICHARD LANG and LUKE POSTLE (2023). An improved bound for the linear arboricity conjecture. *Combinatorica* **43**(3), 547–569. ↑3

[Mat03] JIŘÍ MATOUŠEK (2003). Using the Borsuk-Ulam theorem. Lectures on topological methods in combinatorics and geometry. Universitext (Springer, Berlin). ↑16

[MM09] COLTON MAGNANT and DANIEL M. MARTIN (2009). A note on the path cover number of regular graphs. *The Australasian Journal of Combinatorics* **43**, 211–217. ↑3, 4

[MMPS24] RICHARD MONTGOMERY, ALP MÜYESSER, ALEXEY POKROVSKIY, and BENNY SUDAKOV (2024). Approximate path decompositions of regular graphs. *Journal of the London Mathematical Society.* To appear. ↑4

[Mon24] RICHARD MONTGOMERY (2024). Transversals in Latin squares. *Surveys in combinatorics 2024, London Mathematical Society Lecture Note Series*, vol. 493, 131–158. ↑1

[MPS21] RICHARD MONTGOMERY, ALEXEY POKROVSKIY, and BENNY SUDAKOV (2021). A proof of Ringel's conjecture. *Geometric and Functional Analysis* **31**(3), 663–720. ↑1

[MR90] COLIN MCDIARMID and BRUCE REED (1990). Linear arboricity of random regular graphs. *Random Structures & Algorithms* **1**(4), 443–445. ↑3

[MSS15a] ADAM W. MARCUS, DANIEL A. SPIELMAN, and NIKHIL SRIVASTAVA (2015). Interlacing families I: Bipartite Ramanujan graphs of all degrees. *Annals of Mathematics* **182**(1), 307–325. ↑16

[MSS15b] ADAM W. MARCUS, DANIEL A. SPIELMAN, and NIKHIL SRIVASTAVA (2015). Interlacing families II: Mixed characteristic polynomials and the Kadison–Singer problem. *Annals of Mathematics* **182**(1), 327–350. ↑16

[NW61] C. ST. J. A. NASH-WILLIAMS (1961). Edge-disjoint spanning trees of finite graphs. *The Journal of the London Mathematical Society* **36**(1), 445–450. ↑2

[NW64] C. ST. J. A. NASH-WILLIAMS (1964). Decomposition of finite graphs into forests. *The Journal of the London Mathematical Society* **39**(1), 12–12. ↑2

[Pér82] BERNARD PÉROCHE (1982). On partition of graphs into linear forests and dissections. Preprint. ↑3

[Pós76] LAJOS PÓSA (1976). Hamiltonian circuits in random graphs. *Discrete Mathematics* **14**(4), 359–364. ↑4

[Sco11] ALEXANDER SCOTT (2011). Szemerédi's regularity lemma for matrices and sparse graphs. *Combinatorics, Probability and Computing* **20**(3), 455–466. ↑2

[Sha49] CLAUDE E. SHANNON (1949). A theorem on coloring the lines of a network. *Journal of Mathematics and Physics* **28**(1-4), 148–151. ↑2

[Tom82] PAVEL TOMASTA (1982). Note on linear arboricity. *Mathematica Slovaca* **32**(3), 239–242. ↑3

[Vis12] NISHEETH K. VISHNOI (2012). A permanent approach to the traveling salesman problem. *2012 IEEE 53rd Annual Symposium on Foundations of Computer Science*, 76–80. ↑4

[Viz64] VADIM G. VIZING (1964). On an estimate of the chromatic class of a p-graph. *Diskret. Analiz* **3**, 25–30. ↑2

[Viz76] VADIM G. VIZING (1976). Coloring the vertices of a graph in prescribed colors. *Diskret. Analiz* **29**, 3–10. ↑2

[Wu99] JIAN-LIANG WU (1999). On the linear arboricity of planar graphs. *Journal of Graph Theory* **31**(2), 129–134. ↑3

[WW08] JIAN-LIANG WU and YU-WEN WU (2008). The linear arboricity of planar graphs of maximum degree seven is four. *Journal of Graph Theory* **58**(3), 210–220. ↑3