

Towards the linear arboricity conjecture

Asaf Ferber *

Jacob Fox †

Vishesh Jain‡

Abstract

The linear arboricity of a graph G , denoted by $\text{la}(G)$, is the minimum number of edge-disjoint linear forests (i.e. forests in which every connected component is a path) in G whose union covers all the edges of G . A famous conjecture due to Akiyama, Exoo, and Harary from 1980 asserts that $\text{la}(G) \leq \lceil (\Delta(G) + 1)/2 \rceil$, where $\Delta(G)$ denotes the maximum degree of G . This conjectured upper bound would be best possible, as is easily seen by taking G to be a regular graph. In this paper, we show that for every graph G , $\text{la}(G) \leq \frac{\Delta}{2} + O(\Delta^{2/3-\alpha})$ for some $\alpha > 0$, thereby improving the previously best known bound due to Alon and Spencer from 1992. For graphs which are sufficiently good spectral expanders, we give even better bounds. Our proofs of these results further give probabilistic polynomial time algorithms for finding such decompositions into linear forests.

1 Introduction

A *linear forest* is a forest in which every connected component is a path. Given a graph G , we define its *linear arboricity*, denoted by $\text{la}(G)$, to be the minimum number of edge-disjoint linear forests in G whose union is $E(G)$. This notion was introduced by Harary [15] in 1970 as one of the covering invariants of graphs, and has been studied quite extensively since then.

It is immediate that $\text{la}(G) \leq e(G)$ as every edge uv (along with the isolated vertices $V(G) \setminus \{u, v\}$) forms a linear forest. A less trivial upper bound can be obtained as follows: by a classical theorem due to Vizing, $E(G)$ can be partitioned into at most $\Delta + 1$ matchings, where $\Delta := \Delta(G)$ denotes the maximum degree of G ; observe that each matching is a linear forest, and therefore we get that $\text{la}(G) \leq \Delta + 1$. For a lower bound, note that every linear forest has at most $n - 1$ edges (and equality holds if and only if the linear forest is a Hamiltonian path). Therefore, if G is a Δ -regular graph, then

$$\text{la}(G) \geq \frac{e(G)}{(n-1)} \geq \frac{n\Delta}{2(n-1)} > \frac{\Delta}{2},$$

which implies (recall that $\text{la}(G)$ is an integer) that $\text{la}(G) \geq \lceil (\Delta + 1)/2 \rceil$. The following conjecture, known as the *linear arboricity conjecture*, of Akiyama, Exoo and Harary [1] asserts that this bound is the best possible:

*Massachusetts Institute of Technology. Department of Mathematics. Email: ferbera@mit.edu. Research is partially supported by NSF 6935855.

†Stanford University. Department of Mathematics. Email: jacobfox@stanford.edu. Research supported by a Packard Fellowship and by NSF Career Award DMS-1352121.

‡Massachusetts Institute of Technology. Department of Mathematics. Email: visheshj@mit.edu. Research is partially supported by NSF CCF 1665252, NSF DMS-1737944 and ONR N00014-17-1-2598.

Conjecture 1.1 (The linear arboricity conjecture). *Let G be a graph of maximum degree Δ . Then,*

$$\text{la}(G) \leq \left\lceil \frac{\Delta + 1}{2} \right\rceil.$$

Remark 1.2. It is easy to see that every graph G , with maximum degree $\Delta(G)$, can be embedded into a $\Delta(G)$ -regular graph (perhaps on a greater number of vertices). Therefore, the above conjecture is equivalent to the statement that for a Δ -regular graph G we have $\text{la}(G) = \lceil (\Delta(G) + 1)/2 \rceil$.

The linear arboricity conjecture was shown to be asymptotically correct as $\Delta \rightarrow \infty$ by Alon in 1988 [3]. He showed that for every Δ -regular graph G ,

$$\text{la}(G) \leq \frac{\Delta}{2} + O\left(\frac{\Delta \log \log \Delta}{\log \Delta}\right);$$

in the same paper, he also proved that the linear arboricity conjecture holds for graphs G with girth $\Omega(\Delta(G))$. The bound for general graphs was subsequently improved by Alon and Spencer in 1992 (see [4]) to:

$$\text{la}(G) \leq \frac{\Delta}{2} + O\left(\Delta^{2/3}(\log \Delta)^{1/3}\right). \quad (1)$$

Even though this conjecture has received a considerable amount of attention over the years, and has been proven (i) in special cases (see, e.g., [1, 2, 3, 8, 14, 25, 26]) (ii) for almost all d -regular graphs of constant degree by McDiarmid and Reed [18], and (iii) for a typical Erdős-Rényi graph with edge-density either $\log^{117} n/n \leq p = o(1)$ or p a fixed constant by Glock, Kühn and Osthus [12], there have been no asymptotic improvements in the error term (that is, the second summand in the bound (1) of Alon and Spencer) for general graphs. Our first main result improves this term by a *polynomial* factor:

Theorem 1.3. *There exist absolute constants $\alpha > 0$ and $C > 0$ for which the following holds. For any Δ -regular graph G ,*

$$\text{la}(G) \leq \frac{\Delta}{2} + C\Delta^{\frac{2}{3}-\alpha}.$$

Remark 1.4. In the proof of Theorem 1.3, we make use of Lemma 2.13, the proof of which relies on a ‘nibbling’ argument. As this argument is well-known but quite lengthy, we have used the results from [7] as a black box, and we get a bound of (say) $\alpha = 1/100$. While a more careful analysis of the nibbling process tailored to our argument may very well give a better bound on α , we have made no attempt to do so, since we believe that a ‘natural barrier’ for our argument should be $\alpha = 1/6$ i.e. $\sqrt{\Delta}$ (which is anyway far from Conjecture 1.1), and any further progress towards the conjecture should require new ideas.

It was shown by Peroche [22] that computing the linear arboricity of a graph is NP -complete; this is to be contrasted with variants like the arboricity of a graph (i.e. the minimum number of edge-disjoint forests in G whose union is $E(G)$) for which polynomial time algorithms are available [11]. Our proof of Theorem 1.3 gives an algorithm for computing a decomposition of $E(G)$ into at most $\frac{\Delta}{2} + C\Delta^{\frac{2}{3}-\alpha}$ edge-disjoint linear forests, which runs in time polynomial in $|V(G)|$ with high probability. Since the linear arboricity of a Δ -regular graph is at least $\frac{\Delta}{2}$, we thereby get an approximation algorithm providing the best-known approximation guarantee (to our knowledge) for efficiently approximating the linear arboricity of a regular graph.

Corollary 1.5. *There exist absolute constants $\alpha > 0$ and $C > 0$ for which the following holds. Let G be a Δ -regular graph. Then, there is a probabilistic polynomial time algorithm for approximating $\text{la}(G)$ to within $\left(1 + \frac{C}{\Delta^{1/3+\alpha}}\right)$ -multiplicative error.*

Our second main result deals with (n, Δ, λ) -graphs, which we now define. A Δ -regular graph G is said to be an (n, Δ, λ) -graph if $|V(G)| = n$ and the second largest (in absolute value) eigenvalue of the adjacency matrix of G is at most λ . For all such graphs with λ not too large compared to Δ , we are able to obtain better bounds on the error than the one coming from [Theorem 1.3](#).

Theorem 1.6. *There exist absolute constants $\beta > 0$ and $C > 0$ for which the following holds. For every (n, Δ, λ) -graph G with $\lambda \leq \Delta^{2/3}$,*

$$\text{la}(G) \leq \frac{\Delta}{2} + C(\lambda\Delta)^{\frac{2}{5}-\beta}.$$

Just like for [Theorem 1.3](#), our proof of [Theorem 1.6](#) also leads to an algorithm for computing such a decomposition of $E(G)$ in time which is polynomial in $|V(G)|$ with high probability.

1.1 The general proof scheme

Our proof outlines follow and extend ideas from [\[5\]](#). Let G be a d -regular graph on n vertices. Consider the following procedure to upper bound $\text{la}(G)$: First, find a vertex partitioning $V(G) = V_1 \cup V_2 \cup \dots \cup V_t$, where t is an even positive integer to be specified later, with the following properties:

1. $||V_i| - |V_j|| \leq 1$ for all $i, j \in [t]$, and
2. $d(v, V_i) \in \frac{d}{t} \pm 100 \left(\frac{d \log d}{t} \right)^{1/2}$ for all $v \in V(G)$ and all $i \in [t]$.

The existence of such a partitioning is guaranteed by [Lemma 2.8](#), which is proved by a standard application of Chernoff's bounds ([Lemma 2.1](#)) followed by the Lovász Local Lemma ([Lemma 2.4](#)).

Second, for all $i \neq j$, let B_{ij} be the induced bipartite graph $G[V_i, V_j]$. By Property 2 and Vizing's theorem ([Theorem 2.7](#)), one can decompose $E(B_{ij})$ into at most

$$\Delta(B_{ij}) + 1 \leq \left\lfloor \frac{d}{t} + 102 \left(\frac{d \log d}{t} \right)^{1/2} \right\rfloor := s$$

matchings. Let $\mathcal{M}_{ij}$ be any such decomposition into s matchings (it might be the case that a few of them are empty), and let $\mathfrak{M} := \{\mathcal{M}_{ij}\}_{i < j}$ be the collection of all such decompositions (that is, one decomposition for every B_{ij}).

Third, let $\mathcal{P} := \{P_1, \dots, P_{t/2}\}$ be a Hamiltonian path decomposition of K_t ; the existence of such a decomposition is ensured by the fact that t is even and a classical result of Walecki from the 1890s which can be found in [\[17\]](#) and provides an explicit such decomposition. It is easy to see that using our collection of decompositions $\mathfrak{M}$, one can find a collection $\mathcal{F}_{\mathcal{P}}$ of forests, one for every such Hamiltonian path $P = v_{i_1}, \dots, v_{i_t}$ in $\mathcal{P}$, satisfying the following two properties:

- $\mathcal{F}_{\mathcal{P}}$ consists of at most s edge-disjoint linear forests;
- $\mathcal{F}_{\mathcal{P}}$ contains all the edges $\bigcup_{i < j \in E(P)} E(B_{ij})$.

Indeed, let P be such a Hamiltonian path; after possibly relabeling the vertices, we may assume that $P = 123 \dots t$. Observe that by taking one matching from each decomposition $\mathcal{M}_{i,i+1}$ we obtain a linear forest. Therefore, by repeating this procedure s times, since each $\mathcal{M}_{i,i+1}$ consists of at most s matchings, one can build a collection of at most s linear forests for every P . Clearly, such a collection contains edge-disjoint linear forests whose union consists of all the edges of all the bipartite graphs $\{B_{i,i+1}\}_{i \in [t-1]}$.

As there are $\frac{t}{2}$ Hamiltonian paths in $\mathcal{P}$, the above construction gives us at most

$$\frac{st}{2} \leq \frac{d}{2} + 51 (dt \log d)^{1/2}$$

linear forests which cover all the edges in all the bipartite graphs $\{B_{ij}\}_{i \neq j}$. Let $L := \bigcup_{i=1}^t E(G[V_i])$ be the set of all the edges which have not been covered by these linear forests (we will also identify L with the graph on $V(G)$ whose edges are L , in which case we will refer to L as the *leave graph*). Since $\Delta(L) \leq \frac{d}{t} + 100 \left(\frac{d \log d}{t}\right)^{1/2}$ by Property 2 of the partitioning, Vizing's theorem shows that L can be decomposed into at most $\Delta(L) + 1 \leq \frac{d}{t} + 101 \left(\frac{d \log d}{t}\right)^{1/2}$ matchings. Since any matching is manifestly a linear forest, we have thus obtained a decomposition of the edges of G into at most

$$\frac{st}{2} + \Delta(L) + 1 \leq \frac{d}{2} + \frac{d}{t} + 152 (dt \log d)^{1/2} \quad (2)$$

linear forests. In order to optimize the error term $d/t + 152(dt \log d)^{1/2}$, we would like to pick t so that the two summands in the error term are the same. This is achieved by choosing $t^3 = \Theta(d/\log d)$, in which case

$$\text{la}(G) \leq \frac{d}{2} + \Theta(d^{2/3} \log^{1/3} d).$$

This is the strategy used in [5] to recover the bound of Alon and Spencer.

Let us now discuss the weak points in the construction and the analysis that we have presented, along with ideas for improving them. The formal details will be given in subsequent sections.

- (i) In the above construction, we decompose the leave graph L into matchings and treat each matching as a linear forest by itself. This gives us the $\Theta(d/t)$ error term in the above analysis. Note, however, that adding a matching contained in some $G[V_i]$ to any of the linear forests obtained from a path P which has i as an endpoint still results in a linear forest. Therefore, it makes sense to try to ‘swallow’ all the edges of L in our current linear forests. We discuss this in more detail in [Section 2.6](#), where we also present the key technical lemma ([Lemma 2.13](#)) needed to make this idea work. The upshot of [Lemma 2.13](#) is that it allows us to replace the $\Theta(d/t)$ term in the error by $\Theta((d/t)^{1-\gamma})$ for some $\gamma > 0$. Optimizing the error term now results in the choice $t^{3-2\gamma} = \Theta(d^{1-2\gamma}/\log d)$, which gives an error of $O(d^{2/3-\alpha})$ for some $\alpha > 0$, as desired in [Theorem 1.3](#).
- (ii) In the above construction, we take $\Delta(B_{ij}) + 1$ matchings in each B_{ij} , whereas ideally, we would like to take only ‘average degree’ many matchings. This error, summed up for all B_{ij} , gives us the $\Theta(\sqrt{dt \log d})$ term in (2). In the proof of [Theorem 1.6](#), we will show ([Lemma 2.10](#)) that if G

satisfies some expansion properties, then we can approximately decompose each B_{ij} into edge-disjoint perfect (up to divisibility) matchings. If we remove the linear forests generated by these matchings using the above procedure, then we remove the “correct” number of linear forests, and the “leave graph” L has much smaller maximum degree. Now, we apply [Theorem 1.3](#) to L .

2 Auxiliary lemmas

In this section, we gather various preliminaries, as well as state and prove the key lemmas needed for our proofs.

2.1 Probabilistic estimates

Throughout this paper, we will make extensive use of the following well-known bound on the upper and lower tails of a sum of independent indicators, due to Chernoff (see, e.g., Appendix A in [\[4\]](#)).

Lemma 2.1 (Chernoff’s inequality). *Let $X_1, \dots, X_n$ be independent random variables with $\Pr[X_i = 1] = p_i$ and $\Pr[X_i = 0] = 1 - p_i$ for all i . Let $X = \sum_i X_i$, and let $\mathbb{E}(X) = \mu$. Then*

- $\Pr[X < (1 - a)\mu] < e^{-a^2\mu/2}$ for every $a > 0$;
- $\Pr[X > (1 + a)\mu] < e^{-a^2\mu/3}$ for every $0 < a < 1$.

Remark 2.2. If all the p_i ’s are the same, the obtained bounds are also valid if, instead of taking X as the sum of i.i.d random variables, we take it to be hypergeometrically distributed with mean μ [\[16\]](#).

Before introducing the next tool to be used, we need the following definition.

Definition 2.3. Let $(A_i)_{i=1}^n$ be a collection of events in some probability space. A graph Γ on the vertex set $[n]$ is called a *dependency graph* for $(A_i)_i$ if A_i is mutually independent of all the events $\{A_j : ij \notin E(\Gamma)\}$.

The following is the so-called Lovász local lemma in its symmetric version (see, e.g., [\[4\]](#)).

Lemma 2.4 (Lovász local lemma). *Let $(A_i)_{i=1}^n$ be a sequence of events in some probability space, and let Γ be a dependency graph for $(A_i)_i$. Suppose that $\Pr[A_i] \leq q$ for every $i \in [n]$ and $eq(\Delta(\Gamma) + 1) < 1$. Then, $\Pr[\bigcap_{i=1}^n \overline{A_i}] > 0$.*

2.2 Algorithmic Lovász local lemma

The original proof of the Lovász local lemma in [\[9\]](#) is non-constructive in that it does not provide any way of finding a point in the probability space avoiding the ‘bad’ events. However, in the case when the ‘bad’ events $(A_i)_{i=1}^n$ are determined by a finite collection of mutually independent random variables $(X_j)_{j=1}^m$, the breakthrough work of Moser and Tardos [\[21\]](#) shows that the following simple randomized algorithm efficiently computes an assignment to the random variables $(X_j)_{j=1}^m$ which avoids all the ‘bad’ events – start with a random assignment to the variables $\mathcal{P}$, and check whether some event in $\mathcal{A}$ is violated. If so, arbitrarily pick such a violated event, and sample another random assignment for the values of the variables on which this event depends (this step is called a *resampling* of the event). Continue this process until there are no violated events.

Theorem 2.5 ([21]). *Let $\mathcal{P}$ be a finite set of mutually independent random variables in a probability space. Let $\mathcal{A}$ be a finite set of events determined by these variables. Consider the dependency graph Γ on these events given by adding an edge connecting two events $A, B \in \mathcal{A}$ if and only if A and B depend on some common random variable in $\mathcal{P}$. Suppose that $\Pr[A] \leq q$ for every $A \in \mathcal{A}$ and $\text{eq}(\Delta(\Gamma) + 1) < 1$. Then, there exists an assignment of values to the variables $\mathcal{P}$ not violating any of the events in $\mathcal{A}$. Moreover the randomized algorithm described above resamples an event $A \in \mathcal{A}$ at most an expected $O(1/\Delta(\Gamma))$ times before it finds such an evaluation. Thus, the expected total number of resampling steps is at most $|\mathcal{A}|/\Delta(\Gamma)$.*

Remark 2.6. All the applications of the local lemma in this paper fit the general framework of the above theorem and seek to avoid at most $p_1(n)$ events, where $p_1(n)$ is some polynomial in the number of vertices $n := V(G)$. Moreover, every event in each of our applications can be sampled in time $p_2(n)$, for some polynomial $p_2(n)$. It follows that all of our applications of the local lemma can be performed algorithmically in expected time $O(p_1(n)p_2(n))$. Thus, by Markov's inequality, it follows that the probability of the algorithm taking more than $O(p_1(n)p_2(n)p_3(n))$ time is at most $1/p_3(n)$.

2.3 Vizing's theorem

The *chromatic index* of a graph G , denoted by $\chi'(G)$, is the minimum number of colors needed to color $E(G)$ in such a way that each color class is a matching. It follows immediately from this definition that $\chi'(G) \geq \Delta(G)$; perhaps surprisingly, Vizing [24] proved that this trivial lower bound is nearly optimal:

Theorem 2.7 (Vizing's Theorem). *Every graph G satisfies*

$$\chi'(G) \in \{\Delta(G), \Delta(G) + 1\}.$$

Moreover, the strategy in Vizing's original proof can be used to obtain a polynomial time algorithm to edge color any graph G with $\Delta(G) + 1$ colors ([20]). Note that, as mentioned in the introduction, Vizing's theorem immediately gives the bound $\text{la}(G) \leq \Delta(G) + 1$.

2.4 Random vertex partitioning

Given a d -regular graph with d sufficiently large, the following lemma gives a partition $V(G) = V_1 \cup \dots \cup V_t$ for which 'all the degrees are correct'.

Lemma 2.8. *There exists an absolute constant d_0 for which the following holds. For all $d \geq d_0$, all d -regular graphs G , and all integers $1 \leq t \leq d/100$, there exists a partition $V(G) = V_1 \cup \dots \cup V_t$ satisfying the following two properties:*

1. *For all $1 \leq i, j \leq t$, $||V_i| - |V_j|| \leq 1$.*
2. *For all $v \in V(G)$ and for all $i \in [t]$, the number of edges from v into V_i , denoted by $d_G(v, V_i)$, satisfies $d_G(v, V_i) \in \frac{d}{t} \pm 100 \left(\frac{d \log d}{t} \right)^{1/2}$.*

Proof. Note that for (say) $d \geq \log^2 n$, the lemma follows easily by Chernoff's inequality for the hypergeometric distribution and the union bound. Since we are also interested in graphs with smaller

degree, we need a slightly more complicated proof where the union bound is replaced by a standard application of the local lemma (Lemma 2.4).

Let $s := \lceil n/t \rceil$ and let $F_1, \dots, F_s$ be an arbitrary partition of $V(G)$ such that $F_1, \dots, F_{s-1}$ are of size t each. Let $f : V(G) \rightarrow \{1, \dots, t\}$ be a random function chosen as follows: for each $k \in [s]$, the restriction $f|_{F_k}$ is a permutation of $[|F_k|]$ chosen uniformly at random. Given such an f , define $V_i := \{v \in V(G) : f(v) = i\}$. Observe that for each $i \in [t]$, $|V_i|$ is either $s-1$ or s , so that the desired property 1 of the lemma holds. We wish to show that, with positive probability, there exists an f such that the corresponding partition $V(G) = V_1 \cup \dots \cup V_t$ satisfies property 2 of the lemma.

To this end, fix a vertex $v \in V(G)$ and for each $k \in [s]$, let $p_k := |N_G(v) \cap F_k|/|F_k|$. Since each $f|_{F_k}$ is chosen uniformly at random from among all permutations of $[|F_k|]$, it follows that for all $i \in [t]$,

$$\frac{d}{t} - 1 \leq \sum_{i=1}^{s-1} p_i \leq \mathbb{E}[d_G(v, V_i)] = \sum_{i=1}^s p_i \leq \frac{d}{t} + 1.$$

Therefore, by Chernoff's bounds (Lemma 2.1),

$$\Pr \left[d_G(v, V_i) \notin \frac{d}{t} \pm 100\sqrt{\frac{d \log d}{t}} \right] \leq \exp(-2000 \log d) = \frac{1}{d^{2000}}.$$

Let $\mathcal{E}_{i,v}$ denote the event ' $d(v, V_i) \notin \frac{d}{t} \pm 100\sqrt{\frac{d \log d}{t}}$ ', and note that for all $i \in [t]$ and $v \in V(G)$, $\mathcal{E}_{i,v}$ may depend on an event $\mathcal{E}_{j,u}$ only if at least one of the following two conditions hold: $u = v$; or u and v have neighbors to the same F_k for some k . In particular, each event $\mathcal{E}_{i,v}$ depends on at most $t + d^2 t \leq d^4$ events. Finally, since

$$\frac{e(d^4 + 1)}{d^{2000}} < 1,$$

the local lemma guarantees the existence of an f as desired. $\square$

2.5 Finding dense, regular spanning subgraphs in 'nice' bipartite graphs

The next lemma shows that almost-regular balanced bipartite graphs induced by large disjoint subsets of a good expander contain a spanning regular graph covering almost all the edges. The proof is similar to the proof of Lemma 2.12 in [10], and is based on the following generalization of the Gale-Ryser theorem due to Mirsky [19].

Theorem 2.9 ([19]). *Let $G = (A \cup B, E)$ be a balanced bipartite graph with $|A| = |B| = m$, and let r be an integer. Then, G contains an r -factor if and only if for all $X \subseteq A$ and $Y \subseteq B$*

$$e_G(X, Y) \geq r(|X| + |Y| - m).$$

Lemma 2.10. *Let G be an (n, d, λ) graph. Let $1 \leq t \leq d/100$ be some integer such that $t \mid n$. Let A and B be disjoint subsets of $V(G)$ of sizes $|A| = |B| = \frac{n}{t}$ and consider the bipartite subgraph of G induced by these sets, denoted by $G' := G[A, B]$. Assume further that $\frac{d}{t} - 100\sqrt{\frac{d \log d}{t}} \leq \delta(G') \leq \Delta(G') \leq \frac{d}{t} + 100\sqrt{\frac{d \log d}{t}}$. Then, G' contains an $r := \lfloor \frac{d}{t} - \gamma \rfloor$ -factor (i.e. an r -regular spanning subgraph) for $\gamma = 104 \max\{\lambda, \sqrt{\frac{d \log d}{t}}\}$, provided that $\gamma < r/2$.*

Proof. Since $\gamma \geq 0$, the statement is vacuously true whenever $r \leq 0$. Hence, we may assume that $r > 0$. By Mirsky's criterion, it suffices to verify that for all $X \subseteq A$ and $Y \subseteq B$, we have

$$e_G(X, Y) \geq \left(\frac{d}{t} - \gamma \right) \left(|X| + |Y| - \frac{n}{t} \right).$$

We divide the analysis into five cases:

Case 1: $|X| + |Y| \leq \frac{n}{t}$. Since $e_G(X, Y) \geq 0$, there is nothing to prove in this case.

Case 2: $|X| + |Y| > \frac{n}{t}$, $|Y| \geq |X|$ and $|X||Y^c| \leq \frac{\lambda^2 n^2}{d^2}$, where $Y^c := B \setminus Y$. Suppose for contradiction that $e_G(X, Y) < r(|X| + |Y| - \frac{n}{t})$. Then, it must be the case that

$$e_G(X, Y^c) \geq \delta(G')|X| - e_G(X, Y) > (\delta(G') - r)|X| + r|Y^c|.$$

On the other hand, we know by the expander mixing lemma that

$$e_G(X, Y^c) \leq \frac{d}{n}|X||Y^c| + \lambda\sqrt{|X||Y^c|} \leq 2\lambda\sqrt{|X||Y^c|},$$

where the second inequality holds since $|X||Y^c| \leq \frac{\lambda^2 n^2}{d^2}$. Hence, we must have

$$2\lambda\sqrt{|X||Y^c|} \geq (\delta(G') - r)|X| + r|Y^c|.$$

Since both terms on the right hand side are nonnegative, $2\lambda\sqrt{|X||Y^c|}$ should also be greater than either of them, for which we must have

$$\frac{r^2}{4\lambda^2}|Y^c| \leq |X| \leq \frac{4\lambda^2}{(\delta(G') - r)^2}|Y^c|.$$

In particular, we must have $r(\delta(G') - r) \leq 4\lambda^2$, which implies $\lambda \geq r$, which violates our assumption about λ .

Case 3: $|X| + |Y| > \frac{n}{t}$, $|Y| \geq |X|$, $|X||Y^c| > \frac{\lambda^2 n^2}{d^2}$ and $|X| < \frac{nr}{2d}$. If $e_G(X, Y) < r(|X| + |Y| - \frac{n}{t})$, then by the same argument as above, we must have

$$2\frac{d}{n}|X||Y^c| \geq (\delta(G') - r)|X| + r|Y^c|.$$

In particular, we must have $|X| \geq \frac{nr}{2d}$, which violates our assumption about $|X|$.

Case 4: $|X| + |Y| > \frac{n}{t}$, $|Y| \geq |X|$ and $|X| \geq \frac{nr}{2d}$. By assumption, we have $2|Y| > n/t$, so that $|Y| > n/2t$. Moreover, since $\gamma < r/2$, we have that $r > d/2t$. Therefore, $\sqrt{|X||Y|} \geq \sqrt{n^2/4t^2} \geq n/2t$. On the other hand, we also have $(2n\lambda)/(t\gamma) \leq (2n\gamma)/(104t\gamma) \leq n/50t$. Combining these two inequalities, we see that $\frac{\gamma}{2}|X||Y| \geq \frac{n\lambda}{t}\sqrt{|X||Y|}$. Therefore, by the expander mixing lemma, it suffices to verify that

$$\frac{t}{n} \left(\frac{d}{t} - \frac{\gamma}{2} \right) |X||Y| \geq \left(\frac{d}{t} - \gamma \right) \left(|X| + |Y| - \frac{n}{t} \right).$$

Dividing both sides by n/t , we see that this is implied by the inequality

$$xy - \beta(x + y - 1) \geq 0,$$

where $\beta = \frac{d/t-\gamma}{d/t-\gamma/2}$, $x = t|X|/n$, $y = t|Y|/n$, $x + y \geq 1$, $0 \leq x \leq 1$, and $0 \leq y \leq 1$. Observe that the objective function on the left hand side of the desired inequality is bilinear in x and y , and

therefore the minimum will be obtained on the triangular boundary of the region. On this boundary, the inequality reduces to one of the following: $xy \geq 0$; $x \geq \beta x$; $y \geq \beta y$, and is readily verified since $\beta \leq 1$.

Case 5: $|X| + |Y| > \frac{n}{t}$ and $|Y| \leq |X|$. This is exactly the same as cases (2)-(4) with the roles of X and Y interchanged. $\square$

Remark 2.11. Under the conditions of the above lemma, an r -factor in G' can be found efficiently using algorithmic versions of Mirsky's criterion based on standard network flow algorithms (see, e.g., [6]).

Remark 2.12. In the application of this lemma to [Theorem 1.6](#), we will have to deal with bipartite graphs as above, except that we are allowed to have $|A| = |B| + 1$. In this scenario, it is impossible to find an r -factor. However, by adding a “fake” vertex to B with suitable edge connections to A , finding an r -factor in this new graph using the above lemma, decomposing this r -factor into r edge-disjoint perfect matchings using repeated applications of Hall's theorem, and finally removing all edges incident to the “fake” vertex, we see that $G' := G[A, B]$ contains r -edge disjoint matchings such that every vertex in $A \cup B$ is matched in at least $r - 1$ such matchings.

2.6 Avoiding short cycles

In this section, we introduce our key technical lemma for proving [Theorem 1.3](#). Since the usefulness of this lemma may not be apparent at first glance, we encourage the reader to refer to this section only after encountering its application in the proof of [Theorem 1.3](#).

Lemma 2.13. *There exist universal constants $B, D > 20$ for which the following holds. Let G be a graph with maximum degree Δ and minimum degree δ such that $\Delta - \delta \leq \Delta^{5/6}$ and $\Delta \geq D$. Let $M_1, \dots, M_{\Delta+1}$ be a fixed collection of matchings in the complete graph on $V(G)$. Then, there exists a collection of matchings $M'_1, \dots, M'_{\Delta+1}$ in G , where some of them may possibly be empty, such that:*

1. *the graph G' , which is obtained from G by deleting all the edges $\bigcup_{i \in [\Delta+1]} E(M'_i)$, has maximum degree at most $\Delta^{1-1/B}$;*
2. *for all $v \in V(G)$, there are at most Δ^b indices $i \in [\Delta+1]$ for which v lies on a cycle in $M'_i \cup M_i$ of length at most $\Delta^{1/40}$, where $b := \frac{6}{B} + \frac{1}{20}$.*

Moreover, such a collection of matchings may be obtained in $\text{poly}(V(G))$ time with high probability.

The proof of this lemma builds on the proof of the main result in the work of Dubhashi, Grable, and Panconesi [7]. Since the details are somewhat involved, we defer them to [Appendix A](#).

3 Proofs of main results

In this section, we conclude the proofs of our main results. Since these proofs build on the general strategy discussed earlier, we encourage the reader to review the construction in [Section 1.1](#) before proceeding. We start by proving [Theorem 1.6](#) as a warm up since its proof is simpler.

3.1 Proof of Theorem 1.6

Let G be an (n, d, λ) -graph and set $\gamma = 104 \max\{\lambda, \sqrt{\frac{d \log d}{t}}\}$. As in the general proof scheme presented in Section 1.1, we start with a vertex decomposition $V(G) = V_1 \cup \dots \cup V_t$ satisfying the conclusions of Lemma 2.8, where $t \leq d/100$ is a positive even integer which will be specified below. For all $i \neq j$, let $\mathcal{M}_{ij}$ be a collection of $r := \lfloor \frac{d}{t} - \gamma \rfloor$ edge-disjoint matchings of the bipartite graph $B_{ij} := G[V_i, V_j]$ as in Remark 2.12 – such a decomposition exists for all sufficiently large d since $\gamma < \frac{r}{2}$ holds by our choice of t below, and our assumption that $\lambda \leq d^{2/3}$.

Let $\mathcal{P} := \{P_1, \dots, P_{t/2}\}$ be a Hamiltonian path decomposition of K_t , and for each $P \in \mathcal{P}$, let $\mathcal{F}_P$ be the collection of r edge-disjoint linear forests obtained as in Section 1.1. This gives us a set of $\frac{rt}{2}$ edge-disjoint linear forests of G . The key observation here is that the graph L induced by all edges of G which are not in any such linear forest has maximum degree $\Delta(L) \leq d - (t-1)(r-1) \leq (\gamma+1)t + r$ since each vertex in $V_i \cup V_j$ is in at least $r-1$ of the r edge-disjoint matchings selected in B_{ij} . Our goal now is to find a decomposition of the edges of L into as few linear forests as possible. The bound (1) ensures that we can find a decomposition into at most $\Delta(L)/2 + C(\Delta(L))^{2/3} \log^{1/3}(\Delta(L))$ linear forests. Together with the collection of $rt/2$ edge-disjoint linear forests that we built earlier, this shows that

$$\begin{aligned} \text{la}(G) &\leq \frac{rt}{2} + \frac{\Delta(L)}{2} + C(\Delta(L))^{2/3} \log^{1/3}(\Delta(L)) \\ &\leq \frac{rt}{2} + \frac{\gamma t + t + r}{2} + C(\Delta(L))^{2/3} \log^{1/3}(\Delta(L)) \\ &\leq \frac{(\lfloor \frac{d}{t} - \gamma \rfloor + \gamma)t}{2} + \frac{t + r}{2} + C(\Delta(L))^{2/3} \log^{1/3}(\Delta(L)) \\ &\leq \frac{d}{t} + \frac{t + r}{2} + C(\Delta(L))^{2/3} \log^{1/3}(\Delta(L)). \end{aligned}$$

Setting $t = \tilde{\Theta}\left(\frac{d^3}{\gamma^2}\right)^{1/5}$ to optimize the error term (in which case $\gamma = 104\lambda$) shows that $\text{la}(G) \leq \frac{d}{2} + \tilde{O}((d\lambda)^{2/5})$, where the tilde hides logarithmic dependence on d . If instead of (1), we use Theorem 1.3 to handle the linear arboricity of L , then we get that

$$\text{la}(G) \leq \frac{d}{2} + O\left((d\lambda)^{\frac{2}{5}-\beta}\right)$$

for some $\beta > 0$, as desired.

3.2 Proof of Theorem 1.3

Let G be a d -regular graph on n vertices with d sufficiently large. Let $V(G) = V_1 \cup \dots \cup V_t$ be a vertex-partition satisfying the conclusions of Lemma 2.8, where $t \leq d/100$ is a positive even integer which will be specified below. As before, let $\mathcal{M}_{ij}$ denote a decomposition of the bipartite graph $B_{ij} := G[V_i, V_j]$ into at most $\Delta(B_{ij}) + 1$ matchings, and let $\mathfrak{M} = \{\mathcal{M}_{ij}\}_{i \neq j}$ denote the collection of such decompositions.

Let $\mathcal{P} := \{P_1, \dots, P_{t/2}\}$ be a Hamiltonian path decomposition of K_t , and for each $P \in \mathcal{P}$, let $\mathcal{F}_P$ be the collection of at most s edge-disjoint linear forests obtained as in Section 1.1. Fix an arbitrary labeling $\mathcal{F}_P = \{F_{P,1}, \dots, F_{P,s}\}$ of these forests. Moreover, for each $P \in \mathcal{P}$, let s_P and t_P denote its endpoints, and observe that all the pairs $\{s_P, t_P\}_{P \in \mathcal{P}}$ are disjoint.

Next, for each $P \in \mathcal{P}$, let $\mathcal{M}_{s_P} := \{M''_{P,1}, \dots, M''_{P,s}\}$ be a decomposition of the edges of $G[V_{s_P}]$ into s matchings; the existence of such a decomposition is guaranteed by Vizing's theorem. For each $i \in [s]$, let $F''_{P,i} := F_{P,i} \cup M''_{P,i}$, and observe that $\mathcal{F}_P := \{F''_{P,1}, \dots, F''_{P,s}\}$ is a collection of edge-disjoint linear forests which covers all the edges $\bigcup_{ij \in E(P)} E(B_{ij}) \cup E(G[V_{s_P}])$. For each $i \in [s]$, let $M_{P,i}$ be the set of all pairs $\{x, y\} \subseteq V_{t_P}$ for which there exists a path in $F''_{P,i}$ of length exactly $2t - 1$ with x and y as its endpoints. Note that such paths correspond precisely to two 'full paths' of length $t - 1$ in $F_{P,i}$ whose endpoints in V_{s_P} are an edge of $M''_{P,i}$. Since each $M''_{P,i}$ is a matching, it follows immediately that each $M_{P,i}$ is a matching of the complete graph on the vertex set V_{t_P} .

For each $P \in \mathcal{P}$, consider the graph $G_P^* := G[V_{t_P}]$. By Lemma 2.8, we have

$$\Delta_P - k \leq \delta(G_P^*) \leq \Delta(G_P^*) =: \Delta_P,$$

where $k = 200((d \log d)/t)^{1/2}$. Below, we will choose t to be less than $\sqrt{d}$. Therefore, for d sufficiently large, $\Delta_P \geq D$ and $k \leq \Delta_P^{5/6}$, so that by applying Lemma 2.13 to G_P^* we obtain a collection of $\Delta_P + 1$ matchings $M'_1, \dots, M'_{\Delta_P+1}$ in G_P^* , where some of them are possibly empty, such that:

1. the graph G'_P , which is obtained from G_P^* by deleting all the edges $\bigcup_{i \in [\Delta_P+1]} E(M'_i)$, has maximum degree at most $\Delta_P^{1-1/B}$, and
2. for all $v \in V(G)$, there are at most Δ_P^b indices $i \in [\Delta_P+1]$ for which v lies on a cycle in $M_i \cup M'_i$ of length at most $\Delta_P^{1/40}$.

With this in hand, let $F'_{P,i} := F''_{P,i} \cup M'_{P,i}$ for all $i \in [\Delta_P+1]$, and let $\mathcal{F}'_P := \{F'_{P,1}, \dots, F'_{P,\Delta_P+1}\}$. Since each $F'_{P,i}$ is a graph of maximum degree at most 2, it is a disjoint union of cycles, paths and isolated vertices. We wish to remove one edge from each cycle in each $F'_{P,i}$. For the analysis, it will be convenient to do it in the following manner: for any cycle C in any $F'_{P,i}$ of length at most $t\Delta_P^{1/40}$, remove an edge arbitrarily from $M'_{P,i}$; on the other hand, for a cycle C in some $F'_{P,i}$ of length at least $t\Delta_P^{1/40}$, delete an edge chosen uniformly at random from among the first (with respect to a fixed, but otherwise arbitrary ordering of the edges) $\lfloor \Delta_P^{1/40}/2 \rfloor$ edges of M'_i appearing in this cycle. Let $F^*_{P,i}$ denote the (random) linear forest resulting from $F_{P,i}$ after this deletion, and let $\mathcal{F}_P^* := \{F^*_{P,1}, \dots, F^*_{P,\Delta_P+1}\}$ be the collection of edge-disjoint linear forests obtained from the Hamiltonian path P in this manner.

For each $v \in V_{t_P}$, let $X(v)$ denote the (random) number of edges in $\bigcup_{i \in [\Delta_P+1]} F^*_{P,i} \setminus F'_{P,i}$ which are incident to v . We claim that there is a choice of $\mathcal{F}_P^*$ for which $X(v) \leq 9\Delta_P^{39/40}$ for all $v \in V_{t_P}$. For this, fix $v \in V_{t_P}$ and observe that since v is part of at most Δ_P^b cycles of length $\leq \Delta_P^{1/40}$ in $M_i \cup M'_i$, it follows that v can be a part of at most Δ_P^b cycles in $F'_{P,i}$ of length $\leq t\Delta_P^{1/40}$. Hence, the contribution to $X(v)$ from such cycles is at most Δ_P^b . Moreover, the probability that any cycle of length at least $t\Delta_P^{1/40}$ contributes to $X(v)$ is bounded above by $1/\lfloor \Delta_P^{1/40}/2 \rfloor \leq 3/\Delta_P^{1/40}$, since such a cycle contributes to $X(v)$ only when the edge deleted from it is incident to v , where the edge to be deleted is chosen uniformly at random from among $\lfloor \Delta_P^{1/40}/2 \rfloor$ edges, of which at most one is incident to v . Since there are at most $\Delta_P + 1$ cycles containing v to start with, and since deletions from long cycles are made independently, it follows from Chernoff's bounds that with probability at least $1 - \exp(-\Delta_P^{38/40})$, $X(v) \leq \Delta_P^b + 8\Delta_P^{39/40} \leq 9\Delta_P^{39/40}$. Let $\mathcal{E}_v$ denote the event that this does not happen. Note that $\mathcal{E}_v$ can depend on $\mathcal{E}_u$ only if v and u are both incident to the first $\lfloor \Delta_P^{1/40}/2 \rfloor$ edges

of M'_i in a long cycle. Again, since there are at most $\Delta_P + 1$ cycles to start with, it follows that any $\mathcal{E}_v$ can depend on at most Δ_P^2 other $\mathcal{E}_u$'s. Therefore, since $\Delta_P^2 \exp(-\Delta_P^{39/40}) \ll 1$, it follows from the local lemma that $\Pr[\cap_{v \in V_{t_P}} \overline{\mathcal{E}_v}] > 0$, which proves the desired claim.

Finally, repeat the above construction for each $P \in \mathcal{P}$ to obtain a collection of edge disjoint linear forests $\mathfrak{F} := \{\mathcal{F}_P^*\}_{P \in \mathcal{P}}$, and let L denote the leave graph obtained by deleting from G any edge which appears in $\mathfrak{F}$. Observe that L consists of edges of the following two types:

- edges within V_{t_P} that are not contained in $\cup_{i \in [\Delta_P+1]} E(M'_{P,i})$ i.e. edges in the graph G'_P ;
- edges in $\cup_{P \in \mathcal{P}} F'_P$ that are not contained in $\mathfrak{F}$ i.e. edges removed during the deletion process described above.

Recall from [Lemma 2.8](#) that $\Delta_P \leq s$ for all $P \in \mathcal{P}$. Since the V'_i 's are disjoint, it follows from the above discussion that $\Delta(L) \leq 9\Delta_P^{39/40} + \Delta_P^{1-1/B} \leq 10s^{1-\gamma}$, where $\gamma := \min\{\frac{1}{40}, \frac{1}{B}\}$. Therefore, by Vizing's theorem, one can decompose L into at most $10s^{1-\gamma} + 1$ edge-disjoint matchings. These matchings, together with $\mathfrak{F}$, give a decomposition of $E(G)$ into a number of linear forests which is at most

$$\frac{st}{2} + 10s^{1-\gamma} + 1 \leq \frac{d}{2} + 200 \left(\sqrt{dt \log d} + \left(\frac{d}{t} \right)^{1-\gamma} \right).$$

Optimizing the error term by setting the two summands in the parentheses to be equal gives $t = \left(\frac{d^{1-2\gamma}}{\log d} \right)^{\frac{1}{3-2\gamma}}$, in which case, we get that

$$\text{la}(G) \leq \frac{d}{2} + d^{2/3-\alpha},$$

for some $\alpha > 0$, as desired.

References

- [1] J. Akiyama, G. Exoo, and F. Harary. Covering and packing in graphs. III. Cyclic and acyclic invariants. *Math. Slovaca*, 30(4):405–417, 1980.
- [2] J. Akiyama, G. Exoo, and F. Harary. Covering and packing in graphs. IV. Linear arboricity. *Networks*, 11(1):69–72, 1981.
- [3] N. Alon. The linear arboricity of graphs. *Israel J. Math.*, 62(3):311–325, 1988.
- [4] N. Alon and J. H. Spencer. *The probabilistic method*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., New York, 1992. With an appendix by Paul Erdős, A Wiley-Interscience Publication.
- [5] N. Alon, V. J. Teague, and N. C. Wormald. Linear arboricity and linear k -arboricity of regular graphs. *Graphs Combin.*, 17(1):11–16, 2001.
- [6] R. P. Anstee. The network flows approach for matrices with given row and column sums. *Discrete Math.*, 44(2):125–138, 1983.

- [7] D. Dubhashi, D. A. Grable, and A. Panconesi. Near-optimal, distributed edge colouring via the nibble method. *Theoret. Comput. Sci.*, 203(2):225–251, 1998.
- [8] H. Enomoto and B. Péroche. The linear arboricity of some regular graphs. *J. Graph Theory*, 8(2):309–324, 1984.
- [9] P. Erdős and L. Lovász. Problems and results on 3-chromatic hypergraphs and some related questions. *Colloq. Math. Soc. János Bolyai*, 10:609–627, 1975.
- [10] A. Ferber and V. Jain. 1-factorizations of pseudorandom graphs. *Preprint, arXiv:1803.10361*, 2018.
- [11] H. N. Gabow and H. H. Westermann. Forests, frames, and games: algorithms for matroid sums and applications. *Algorithmica*, 7(5-6):465–497, 1992.
- [12] S. Glock, D. Kühn, and D. Osthus. Optimal path and cycle decompositions of dense quasirandom graphs. *J. Combin. Theory Ser. B*, 118:88–108, 2016.
- [13] D. A. Grable. A large deviation inequality for functions of independent, multi-way choices. *Combin. Probab. Comput.*, 7(1):57–63, 1998.
- [14] F. Guldan. The linear arboricity of 10-regular graphs. *Math. Slovaca*, 36(3):225–228, 1986.
- [15] F. Harary. Covering and packing in graphs. I. *Ann. New York Acad. Sci.*, 175:198–205, 1970.
- [16] W. Hoeffding. Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.*, 58:13–30, 1963.
- [17] E. Lucas. *Récréations mathématiques*. 2ième éd., nouveau tirage. Librairie Scientifique et Technique Albert Blanchard, Paris, 1960.
- [18] C. McDiarmid and B. Reed. Linear arboricity of random regular graphs. *Random Structures Algorithms*, 1(4):443–445, 1990.
- [19] L. Mirsky. Combinatorial theorems and integral matrices. *J. Combinatorial Theory*, 5:30–44, 1968.
- [20] J. Misra and D. Gries. A constructive proof of Vizing’s theorem. *Inform. Process. Lett.*, 41(3):131–133, 1992.
- [21] R. A. Moser and G. Tardos. A constructive proof of the general Lovász local lemma. *J. ACM*, 57(2):Art. 11, 15, 2010.
- [22] B. Péroche. Complexité de l’arboricité linéaire d’un graphe. II. *RAIRO Rech. Opér.*, 19(3):293–300, 1985.
- [23] V. Rödl. On a packing and covering problem. *European J. Combin.*, 6(1):69–78, 1985.
- [24] V. G. Vizing. On an estimate of the chromatic class of a p -graph. *Diskret. Analiz No.*, 3:25–30, 1964.
- [25] J.-L. Wu. On the linear arboricity of planar graphs. *J. Graph Theory*, 31(2):129–134, 1999.
- [26] J.-L. Wu and Y.-W. Wu. The linear arboricity of planar graphs of maximum degree seven is four. *J. Graph Theory*, 58(3):210–220, 2008.

A Proof of Lemma 2.13

In this appendix, we show how the proof of the main result in [7], which is based on the celebrated Rödl nibble [23], implies Lemma 2.13. The organization of this appendix is as follows: Algorithm 1 records the nibbling algorithm used in [7]; Theorem A.1 and Theorem A.4 record the conclusion of the analysis in [7]; Corollary A.5 adapts the analysis in [7] for our choice of parameters; Lemma A.6 and Lemma A.7 show that Algorithm 1 produces only a small number of short cycles with respect to any fixed collection of matchings, and finally, Proposition A.8 proves Lemma 2.13.

Before proceeding to formal details, let us provide a high level overview of what follows. The goal in [7] is to produce a proper edge-coloring of a Δ -regular graph G using $(1 + \epsilon)\Delta$ colors (here, ϵ is allowed to depend on Δ). Their algorithm runs in two phases – the first phase, which is based on the semi-random ‘nibble’ method of Rödl, is the one relevant to our paper; the second phase actually uses a trivial algorithm. In the first phase, the algorithm seeks to color ‘most’ of the edges using a *palette* of Δ colors. Starting with the input graph $G_0 := G$, the algorithm generates a sequence $G_0, G_1, \dots, G_{t_\epsilon}$ of graphs, where G_i is the graph induced by the edges which are still uncolored at the end of stage i . In each stage i , each edge has a palette of all ‘available’ colors, where initially, the palette of each edge is the set $\{1, \dots, \Delta\}$. Each vertex selects an $\epsilon/2$ -fraction of uncolored edges incident to it, and each selected edge picks a *tentative* color from its palette independently and uniformly at random. If a *selected* edge has no ‘color-conflicts’ with any neighboring edge, then the corresponding color becomes the final color of the edge. All the palettes of the remaining edges are updated by deleting all the final colors of neighboring colored edges. This process is then repeated in the next stage. The algorithm continues for a number of rounds by the end of which (with high probability) each vertex has no more than $\epsilon\Delta$ uncolored edges incident to it.

As in all nibbling-based arguments, the key idea is that in each stage, the number of edges which experience color conflicts is only a small fraction of the number of edges selected to be colored at this stage. The main effort in [7] is spent in showing that this holds true with high probability throughout the process. They do this by showing inductively – and this is what we will use in our analysis – that the graphs G_i and the color palettes of each edge behave almost like ‘random’ subgraphs and subsets of the original ones. We now give a formal description of the algorithm and analysis in [7]. Following this, we will show how to tailor it to our application.

Algorithm 1 is the first phase of the algorithm used in [7] as described above. The analysis of this algorithm is based on controlling the following three quantities:

- $|A_i(u)|$, the size of the implicit palette of vertex u at the end of stage i , where the implicit palette $A_i(u)$ denotes the set of colors not yet successfully used by any edge incident to u .
- $|A_i(e)|$, the size of the palette $A_i(e)$ of edge e at the end of stage i . Note that $A_i(uv) = A_i(u) \cap A_i(v)$.
- $\deg_{i,\gamma}(u)$, the number of neighbors of u which, at the end of stage i , have color γ in their palettes.

Before we discuss their analysis of this algorithm, we need some notation. Define d_i and a_i as follows: first, define initial values

$$d_0, a_0 := \Delta$$

Algorithm 1 The Nibble Algorithm

The initial graph $G_0 := G$, the input graph. Each edge $e = uv$ is initially given the palette $A_0(e) = \{1, \dots, \Delta\}$. For $i = 0, \dots, t_\epsilon - 1$ stages, repeat the following:

- (*Select nibble*) Each vertex u randomly selects an $\epsilon/2$ fraction of the uncolored edges incident to itself. An edge is considered selected if either or both of its endpoints selects it.
- (*Choose tentative color*) Each selected edge e chooses independently at random a tentative color $t(e)$ from its palette $A_i(e)$ of currently available colors.
- (*Check color conflicts*) Color $t(e)$ becomes the final color of e unless some edge incident to e has chosen the same tentative color.
- (*Update graph and palettes*) The graph and the palettes are updated by setting

$$G_{i+1} = G_i - \{e | e \text{ got a final color}\}$$

and, for each edge e , setting

$$A_{i+1}(e) = A_i(e) - \{t(f) | f \text{ incident to } e, t(f) \text{ is the final color of } f\}.$$

and then, recursively define

$$\begin{aligned} d_i &:= (1 - p_\epsilon) d_{i-1} = (1 - p_\epsilon)^i \Delta; \\ a_i &:= (1 - p_\epsilon)^2 a_{i-1} = (1 - p_\epsilon)^{2i} \Delta = d_i^2 / \Delta, \end{aligned}$$

where

$$p_\epsilon := \epsilon \left(1 - \frac{\epsilon}{4}\right) e^{-2\epsilon(1-\epsilon/4)}.$$

In particular, note that setting

$$t_\epsilon := \frac{1}{p_\epsilon} \log \frac{4}{\epsilon},$$

we have $d_{t_\epsilon} \leq \epsilon \Delta / 4$. Also, provided that $\epsilon < 1/100$, we have $d_{t_\epsilon} = (1 - p_\epsilon)^{t_\epsilon} \Delta \geq e^{-2p_\epsilon t_\epsilon} \Delta = \epsilon^2 \Delta / 16$.

Theorem A.1 ([7], Lemmas 9, 12 and 15, and the discussion in Section 5.5). *There exist constants $K, D > 0$ such that if $\epsilon < 1/100$, $\epsilon^2 \Delta \geq D$, and at the end of stage i of [Algorithm 1](#) the following holds for all vertices u , edges e and colors γ with $e_i \leq 1/2$:*

$$\begin{aligned} |A_i(u)| &= (1 \pm e_i) d_i \\ |A_i(e)| &= (1 \pm e_i) a_i \\ \deg_{i,\gamma}(u) &= (1 \pm e_i) a_i, \end{aligned}$$

then the following holds for all vertices u , edges e , and colors γ :

$$\begin{aligned} \mathbb{E}[|A_{i+1}(u)|] &= (1 \pm (1 + K\epsilon)e_i) d_{i+1} \\ \mathbb{E}[|A_{i+1}(e)|] &= (1 \pm (1 + K\epsilon)e_i) a_{i+1} \\ \mathbb{E}[\deg_{i+1,\gamma}(u)] &= (1 \pm (1 + K\epsilon)e_i) a_{i+1}. \end{aligned}$$

Remark A.2. In our case (Lemma 2.13), we have

$$|A_0(u)| = |A_0(e)| = \Delta,$$

and

$$\deg_{0,\gamma} = \deg(u) = \Delta \left(1 \pm \frac{\Delta - \delta}{\Delta}\right)$$

for all vertices u , edges e , and colors γ . Therefore, we can take

$$e_0 = \frac{\Delta - \delta}{\Delta} \leq \Delta^{-1/6}.$$

In order to show that the above random variables concentrate around their expectation, we will (as in [7]) use the following concentration inequality due to Grable [13]. The statement of this inequality uses the notion of the ‘variance of a strategy for determining a random variable’ of the form $Y = f(X_1, \dots, X_n)$, whose definition we reproduce verbatim from [7] for the reader’s convenience. A *querying strategy* for Y is a decision tree whose internal nodes designate queries to be made. Each node of the tree represents a query of the type “what was the random choice of X_i ?”. A node has as many children as there are random choices for X_i . Every path from the root to a node which goes through vertices corresponding to $X_{i_1}, \dots, X_{i_k}$ defines an assignment $a_1, \dots, a_k$ to these random variables. We can think of each node as storing the value $\mathbb{E}[Y | X_{i_1} = a_1, \dots, X_{i_k} = a_k]$. In particular, the leaves store the possible values of Y , since by then all relevant random choices have been determined. Define the *variance of a query* (internal node) q concerning choice X_i to be

$$v_q = \sum_{a \in A_i} \Pr[X_i = a] \mu_{q,a}^2,$$

where

$$\mu_{q,a} = \mathbb{E}[Y | X_i = a \text{ and all previous queries}] - \mathbb{E}[Y | \text{all previous queries}].$$

By “all previous queries”, we mean the condition imposed by the queried choices and exposed values determined by the path from the root of the strategy down to the node q . In words, $\mu_{q,a}$ measures the amount which our expectation changes when the answer to query q is revealed to be a . Also define the maximum effect of query q as

$$c_q = \max_{a,b \in A_i} |\mu_{q,a} - \mu_{q,b}|.$$

A way to think about c_q is the following. Consider the children of node q ; c_q is the maximum difference between any values $\mathbb{E}[Y | \text{all previous queries}]$ stored at the children. A *line of questioning* ℓ is a path in the decision tree from the root to a leaf and the variance of a line of questioning is the sum of the variances of the queries along it. Finally, the *variance of a strategy* $\mathcal{S}$ is the maximum variance over all lines of questioning

$$V(\mathcal{S}) = \max_{\ell} \sum_{q \in \ell} v_q.$$

We are now ready to state Grable’s concentration inequality.

Theorem A.3 ([13]). *Let $\mathcal{S}$ be a strategy for determining a random variable Y , and suppose the variance of $\mathcal{S}$ is at most V . Then, for every $0 \leq \varphi \leq V / \max c_q^2$,*

$$\Pr \left[|Y - \mathbb{E}[Y]| > 2\sqrt{\varphi V} \right] \leq 2\exp(-\varphi).$$

We note that $\max c_q^2$ will always be at most 16 in all the applications of [Theorem A.3](#) that we will need ([\[7\]](#)).

The next theorem records the bounds on the variance (in the above sense) of various random variables which we are interested in.

Theorem A.4 ([\[7\]](#), Lemmas 10, 13 and 16, and the discussion in Section 5.5.). *Fix $0 \leq i \leq t_\epsilon - 1$. Let $X_u := |A_i(u)| - |A_{i+1}(u)|$, $Y_e := |A_i(e)| - |A_{i+1}(e)|$, and $Z_{\gamma,u} := \deg_{i+1,\gamma}(u)$. Suppose the assumptions of [Theorem A.1](#) are satisfied. Then, for any fixed u, e, γ , with probability at least $1 - \exp(-\epsilon d_i/10)$, there exist strategies $\mathcal{S}_{X,u}$, $\mathcal{S}_{Y,e}$ and $\mathcal{S}_{Z,\gamma,u}$ for determining X_u , Y_e and $Z_{\gamma,u}$ with the following variance bounds:*

$$\begin{aligned} V(\mathcal{S}_{X,u}) &\leq 100\epsilon d_i \\ V(\mathcal{S}_{Y,e}) &\leq 10000a_i \\ V(\mathcal{S}_{Z,\gamma,u}) &\leq 100\epsilon a_i. \end{aligned}$$

As an immediate application of [Theorem A.3](#) along with these bounds, we obtain the following corollary, which shows that random variables in the conclusion of [Theorem A.1](#) are indeed sufficiently well-concentrated around their expectation.

Corollary A.5. *Fix $0 \leq i \leq t_\epsilon - 1$. Suppose that the assumptions of [Theorem A.1](#) are satisfied. For any u, e, γ , let $\mathcal{X}_u$ be the event that $||A_{i+1}(u)| - \mathbb{E}[A_{i+1}(u)]| \geq \sqrt{\epsilon d_i^{5/3}}$, $\mathcal{Y}_e$ be the event that $||A_{i+1}(e)| - \mathbb{E}[A_{i+1}(e)]| \geq \sqrt{a_i^{5/3}}$, and $\mathcal{Z}_{\gamma,u}$ be the event that $|\deg_{i+1,\gamma}(u) - \mathbb{E}[\deg_{i+1,\gamma}(u)]| \geq \sqrt{\epsilon a_i^{5/3}}$. Then,*

$$\begin{aligned} \Pr[\mathcal{X}_u] &\leq 2\exp(-d_i^{1/3}) + \exp(-\epsilon d_i/10) \\ \Pr[\mathcal{Y}_e] &\leq 2\exp(-a_i^{1/3}) + \exp(-\epsilon d_i/10) \\ \Pr[\mathcal{Z}_{\gamma,u}] &\leq 2\exp(-a_i^{1/3}) + \exp(-\epsilon d_i/10). \end{aligned}$$

Proof. The variance bounds from [Theorem A.4](#) hold for any fixed u, e, γ except with probability at most $\exp(-\epsilon d_i/10)$. Whenever these bounds hold, we apply [Theorem A.3](#) with $\varphi = d_i^{1/3}$ (in the case of $\mathcal{X}_u$) or $\varphi = a_i^{1/3}$ (in the case of $\mathcal{Y}_e$ and $\mathcal{Z}_{\gamma,u}$). Finally, we use that $2000a_i^{4/3} \leq a_i^{5/3}$ and $2000d_i^{4/3} \leq d_i^{5/3}$ since $\epsilon^2 \Delta \geq D$ by assumption. $\square$

The next lemma is tailored for our application. Roughly speaking, we are given matchings $M_1, \dots, M_s$ in the complete graph on $V(G)$. We wish to design a random procedure to properly color (most of) the edges of G using s colors in such a way that by considering the matchings $M'_1, \dots, M'_s$ induced by each color class, the probability of any vertex u becoming part of too many short cycles in any of the graphs $M_i \cup M'_i$ is sufficiently small.

Lemma A.6. *Let $M_1, \dots, M_s$ be a fixed collection of matchings in the complete graph on $V(G)$, where $s := \Delta + 1$. Let $0 < \beta < 1/10$. For any vertex u , let $C_i(u)$ denote the number of indices γ for which u lies on a cycle of length at most $\Delta^{\beta/2}$ in $M_\gamma \cup M'_\gamma$ by the end of round i of the algorithm. Let $\mathcal{C}_{i,u}$ denote the event that $C_{i+1}(u) - C_i(u) \geq 1000\Delta^\beta/\epsilon^4$. Suppose that the assumptions of [Theorem A.1](#) are satisfied. Then,*

$$\Pr[\mathcal{C}_{i,u}] \leq \exp(-\Delta^\beta/\epsilon^4).$$

Proof. Fix $u \in V(G)$ as in the statement of the lemma. For each $\gamma \in [s]$ such that u is not already lying on a cycle in $M_\gamma \cup M'_\gamma$ by the end of round i of the algorithm, let P_γ denote the unique maximal path (with a fixed, but otherwise arbitrary, orientation) in $M_\gamma \cup M'_\gamma$ containing u at the end of round i . Observe that if the first and last edges of P_γ do not belong to M_γ , then P_γ cannot be extended to a cycle in $M_\gamma \cup M'_\gamma$. Let p_γ denote the first vertex of P_γ and let q_γ denote its last vertex. For any vertex $v \in V(G)$, let v^γ denote the unique (if it exists) vertex such that $\{v, v^\gamma\} \in M_\gamma$, and let $\bar{v}^\gamma$ denote the unique (if it exists) vertex such that the edge $\{v, \bar{v}^\gamma\}$ is colored γ during the execution of the algorithm by the end of round $i + 1$. Finally, for each $\gamma \in [s]$, consider the following sequence of vertices defined inductively: $w_{0,\gamma} := p_\gamma$, $w_{2i+1,\gamma} := \bar{w}_{2i,\gamma}^\gamma$ for $i \geq 0$, and $w_{2i,\gamma} := w_{2i-1,\gamma}^\gamma$ for $i \geq 1$.

Note that P_γ closes into a cycle of length at most $\ell = 2\lceil \Delta^{\beta/2}/2 \rceil + 1$ in $M_\gamma \cup M'_\gamma$ during the $(i + 1)^{st}$ round only if one of the vertices $w_{1,\gamma}, w_{3,\gamma}, w_{5,\gamma}, \dots, w_{\ell,\gamma}$ is q_γ . In particular, at least one of the edges $\{w_{0,\gamma}, q_\gamma\}, \{w_{2,\gamma}, q_\gamma\}, \dots, \{w_{\ell-1,\gamma}, q_\gamma\}$ must be *tentatively* colored by γ during the $(i + 1)^{st}$ round. Letting E_γ denote the random variable recording the number of such edges, it follows that $C_{i+1}(u) - C_i(u) \leq \sum_{\gamma \in [s]} E_\gamma$. Moreover, since a given edge e is tentatively colored by a given color γ during the $(i + 1)^{st}$ round with probability at most $1/|A_i(e)| \leq 2/a_i \leq 2\Delta/d_{t_\epsilon}^2 \leq 512/\Delta\epsilon^4$, and since the tentative colors for different edges are chosen independently, we see that conditioning on any choice for the collection of vertices $\{w_{2,\gamma} \notin P_\gamma, \dots, w_{\ell-1,\gamma} \notin P_\gamma\}_{\gamma \in [s]}$, the random variable $\sum_{\gamma \in [s]} E_\gamma$ is stochastically dominated by the random variable $\text{Bin}(\Delta^{1+(\beta/2)}, 512/\Delta\epsilon^4)$. Therefore, by Chernoff's bound for the binomial distribution followed by the law of total probability to remove the conditioning, it follows that

$$\Pr[C_{i,u}] \leq \Pr[\text{Bin}(\Delta^{1+(\beta/2)}, 512/\Delta\epsilon^4) \geq 1000\Delta^\beta/\epsilon^4] \leq \exp(-\Delta^\beta/\epsilon^4),$$

which completes the proof. $\square$

The following lemma combines [Corollary A.5](#) and [Lemma A.6](#) to prove the existence of a ‘good’ outcome of a given round of the algorithm.

Lemma A.7. *Fix $0 \leq i \leq t_\epsilon - 1$. Suppose that the assumptions of [Lemma A.6](#) are satisfied and $\Delta^{-1/B} \leq \epsilon < 1/100$ for some $B > 20$. Then with positive probability, the following holds at the end of stage $i + 1$ for all vertices u , edges e and colors γ simultaneously:*

$$\begin{aligned} |A_{i+1}(u)| &= (1 \pm e_{i+1})d_{i+1} \\ |A_{i+1}(e)| &= (1 \pm e_{i+1})a_{i+1} \\ \deg_{i+1,\gamma}(u) &= (1 \pm e_{i+1})a_{i+1} \\ C_{i+1}(u) - C_i(u) &\leq 1000\Delta^\beta/\epsilon^4, \end{aligned}$$

where $0 < \beta < \frac{1}{10} - \frac{1}{B}$ is fixed, and

$$e_{i+1} = \kappa \left(e_i + \sqrt{a_i^{-1/3}} \right) \leq \kappa \left(e_i + (1 - p_\epsilon)^{-i} \Delta^{-1/6} \right), \quad (3)$$

with $\kappa = 1 + K\epsilon$.

Proof. Let $\mathcal{X}_u, \mathcal{Y}_e, \mathcal{Z}_{\gamma,u}$ be the events defined in [Corollary A.5](#), and let $\mathcal{C}_u := \mathcal{C}_{i,u}$ be the event defined in [Lemma A.6](#). It suffices to show that

$$\Pr \left[\left(\bigcap_{u \in V(G)} \mathcal{X}_u^c \right) \cap \left(\bigcap_{e \in E(G)} \mathcal{Y}_e^c \right) \cap \left(\bigcap_{\gamma \in [s], u \in V(G)} \mathcal{Z}_{\gamma,u}^c \right) \cap \left(\bigcap_{u \in V(G)} \mathcal{C}_u^c \right) \right] > 0.$$

We will show this using the symmetric local lemma. To this end, we note that two events of the form $\mathcal{E}(u, e, \gamma)$ and $\mathcal{E}(u', e', \gamma')$ can depend on each other only if at least one of u or e is within distance at most (say) $4\Delta^{\beta/2}$ from one of u' or e' . Since the maximum degree of G is Δ , it follows that the dependency graph of the events listed above has maximum degree at most (say) $\text{poly}(\Delta)\Delta^{4\Delta^{\beta/2}} \leq \exp(\Delta^{3\beta/4})$, where the last inequality holds for all Δ sufficiently large. Also, by [Lemma A.6](#), events of the form $\mathcal{C}_u$ hold with probability at most $\exp(-\Delta^\beta)$, whereas by [Corollary A.5](#), the other events hold with probability at most $\exp(-\epsilon d_{t_\epsilon}/10) + 2\exp(-a_{t_\epsilon}^{1/3})$. Hence, if this latter quantity were much less than $\exp(\Delta^{3\beta/4})$, we would be done. This is indeed true provided that $\beta < \frac{1}{10} - \frac{1}{B}$ and Δ is sufficiently large. $\square$

Finally, we iterate [Lemma A.7](#) to prove the main result of this appendix.

Proposition A.8. *There exist constants $B, D > 20$ for which the following holds. Let G be a graph with maximum degree Δ and minimum degree δ such that $\Delta - \delta \leq \Delta^{5/6}$ and $\Delta \geq D$. Let $M_1, \dots, M_s$ be a fixed collection of matchings in the complete graph on $V(G)$, where $s := \Delta + 1$. Then, for any fixed $\Delta^{-1/B} \leq \epsilon < 10^{-4}$, the following holds with positive probability for the execution of [Algorithm 1](#) on G with parameter ϵ for t_ϵ stages: for fixed $0 < \beta < \frac{1}{10} - \frac{1}{B}$, for all $0 \leq i \leq t_\epsilon - 1$, and for all vertices u , all edges e , and all colors γ ,*

- $|A_i(u)| = (1 \pm \epsilon^3)d_i$
- $|A_i(e)| = (1 \pm \epsilon^3)a_i$
- $\deg_{i,\gamma}(u) = (1 \pm \epsilon^3)a_i$
- $C_{i+1}(u) - C_i(u) \leq 1000\Delta^\beta/\epsilon^4$.

In particular, for every vertex u , $C_{t_\epsilon}(u) \leq 1000t_\epsilon\Delta^\beta/\epsilon^4 \leq \Delta^\beta/\epsilon^6$. Further, the number of uncolored edges incident to u is at most $|A_{t_\epsilon}(u)| \leq 2d_{t_\epsilon} \leq \epsilon\Delta/2$.

Remark A.9. By taking the matchings $M'_1, \dots, M'_\Delta$ to be the edges colored (at the end of stage t_ϵ) by $1, \dots, \Delta$ respectively, it is immediately seen that the above proposition, with $\epsilon = \Delta^{-1/B}$ and $\beta = 1/20$, implies [Lemma 2.13](#).

Proof. We view the execution of [Algorithm 1](#) as a branching process, where in each round of the algorithm, we branch out according to which edges are assigned final colors, and which final colors are assigned to these edges. Generate this tree for t_ϵ levels, and consider any root to leaf path such that for each intermediate ‘branch’, the endpoint further from the root satisfies the conclusions of [Lemma A.7](#) given the parameters at its parent. Such a root-to-leaf path is guaranteed to exist by [Lemma A.7](#). To complete the proof, we track the error introduced by the iterative application of [Lemma A.7](#), and show that it is no more than what is stated in the proposition.

Setting $A := \Delta^{-1/6}$ and $P := (1 - p_\epsilon)^{-1}$, we get from [Equation \(3\)](#) that

$$e_\ell \leq \kappa^\ell e_0 + A[\kappa^\ell + \kappa^{\ell-1}P + \dots + \kappa P^{\ell-1}]$$

for all $0 \leq \ell \leq t_\epsilon$. Since $P = (1 - p_\epsilon)^{-1}$ is also of the form $1 + K'\epsilon$ for some constant $K' > 0$, it follows that

$$e_\ell \leq (1 + L\epsilon)^\ell e_0 + \ell(1 + L\epsilon)^\ell \Delta^{-1/6},$$

where $L = \max\{K, K'\}$. By [Remark A.2](#), $e_0 \leq (\Delta - \delta)/\Delta \leq \Delta^{-1/6}$. Therefore,

$$e_\ell \leq 2\ell \exp(L\epsilon\ell)\Delta^{-1/6}.$$

The right hand side is maximized when $\ell = t_\epsilon$, in which case it is at most

$$\left(\frac{1}{\epsilon}\right)^{3L} \Delta^{-1/6} \leq \epsilon^3,$$

where the last inequality holds provided we take $B \geq 18(L + 1)$. □