INSTITUTE OF MATHEMATICS

THE CZECH ACADEMY OF SCIENCES

Almost all trees are almost graceful

Anna Adamaszek
Peter Allen
Codruţ Grosu
Jan Hladký

Preprint No. 39-2016

PRAHA 2016

ALMOST ALL TREES ARE ALMOST GRACEFUL

ANNA ADAMASZEK, PETER ALLEN, CODRUŢ GROSU, AND JAN HLADKÝ

Abstract. The Graceful Tree Conjecture of Rosa from 1967 asserts that the vertices of each tree T of order n can be injectively labelled by using the numbers $\{1, 2, \ldots, n\}$ in such a way that the absolute differences induced on the edges are pairwise distinct.

We prove the following relaxation of the conjecture for each $\gamma > 0$ and for all $n > n_0(\gamma)$. Suppose that *(i)* the maximum degree of T is bounded by $O_\gamma(n/\log n)$, and *(ii)* the vertex labels are chosen from the set $\{1, 2, \ldots, \lceil(1+\gamma)n\rceil\}$. Then there is an injective labelling of $V(T)$ such that the absolute differences on the edges are pairwise distinct. In particular, asymptotically almost all trees on n vertices admit such a labelling.

As a consequence, for any such tree T we can pack $\lceil(2+2\gamma)n\rceil - 1$ copies of T into $K_{\lceil(2+2\gamma)n\rceil-1}$ cyclically. This proves an approximate version of the Ringel–Kotzig conjecture (which asserts the existence of a cyclic packing of $2n-1$ copies of any T into K_{2n-1}) for these trees.

The proof proceeds by showing that a certain very natural randomized algorithm produces a desired labelling with high probability.

1. Introduction

1.1. **Graceful labelling.** Let G be a graph with n vertices and q edges. A *vertex labelling* of G is an assignment of natural numbers to the vertices of G, subject to certain conditions. A vertex labelling f is called *graceful* if f is an injection from $V(G)$ to the set $\{1, \ldots, q+1\}$ such that, if each edge $xy \in E(G)$ is assigned the *induced* label $|f(x) - f(y)|$, then the resulting edge labels are distinct. G is called *graceful* if it admits a graceful labelling.

Graceful labellings were first introduced by Rosa [18] under the name of *β-valuations*. It was Golomb [9] who used the term *graceful* for the first time.

A natural problem associated with graceful labellings is to determine which graphs are graceful. According to an unpublished result of Erdős (a version of which for so-called harmonious labellings[1] was later recorded by Graham and Sloane [10]) almost all graphs are not graceful. It was shown by Rosa [18] that a graph with every vertex of even degree and

AA: Most of the work was done while the author was at Max-Planck-Institut für Informatik, Saarbrücken, Germany supported by the Humboldt Foundation. *CG:* This research was supported by the Deutsche Forschungsgemeinschaft within the research training group 'Methods for Discrete Structures' (GRK 1408). *JH:* The research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement number 628974. The Institute of Mathematics is supported by RVO:67985840.

[1] see Section 5.3

number of edges congruent to 1 or 2 (mod 4) is not graceful. Nevertheless, it appears that many graphs that exhibit some regularity in structure are graceful. For example, paths P_n and wheels W_n are graceful ([18], [7]). A comprehensive survey on the current status of knowledge on graceful graphs can be found in [8].

The most important problem in the area is to determine whether every tree is graceful. Conjectured first by Rosa in 1967, the problem remains wide open.

Conjecture 1 (Graceful Tree Conjecture)**.** *For any n-vertex tree T there exists an injective labelling $\psi : V(T) \to \{1, \ldots, n\}$ that induces pairwise distinct labels on the edges of T.*

While Conjecture 1 has attracted a lot of attention, it was proved only for some special classes of trees (paths and caterpillars [18], firecrackers [5], banana trees [19], olive trees [16], trees of diameter at most 7 [20], and several others, see Table 4 in [8]). The related conjecture of Bermond [1] that all lobsters are graceful is still open, and has also been the subject of many papers.

Van Bussel [4] introduced the following relaxation of gracefulness. A map $\psi : V(G) \to [m]$ from a vertex set of a graph is *m-graceful* if ψ is injective, and the map ψ_* induced on the edges, $\psi_* : E(G) \to [m-1]$, $\psi_*(xy) := |\psi(x) - \psi(y)|$, is injective as well. If the codomain $[m]$ is clear from the context, we simply call ψ *graceful.* Van Bussel proved that every tree T admits a $\big(2v(T) - 2\big)$-graceful labelling.

1.2. **Tree packings.** The motivation for considering graceful labellings comes from the area of graph packings. A collection $G_1, \ldots, G_t$ of graphs *pack* into a graph H if there are embeddings $\psi_1, \ldots, \psi_t$ of $G_1 \ldots, G_t$ into H such that each edge of H is used in at most one embedding. If in addition each edge of H is used in some embedding, we say that $G_1, \ldots, G_t$ *decompose* H.

There are several open problems in the area of graph packings. These are special cases of the following general meta-conjecture: if $G_1, \ldots, G_t$ are drawn from a family of 'sparse' graphs, and H is 'dense', then provided some 'simple' necessary conditions are satisfied, $G_1, \ldots, G_t$ pack into H.

In particular, we can consider the (sparse) family of trees and the (dense) complete graph. Even here there are several incarnations of the meta-conjecture. The one which will concern us is Ringel's conjecture from 1963.

Conjecture 2 (Ringel's conjecture, [17])**.** *For any $(n+1)$-vertex tree T the complete graph K_{2n+1} can be decomposed into $2n+1$ edge-disjoint subgraphs isomorphic to T.*

This conjecture can be strengthened by requiring the embeddings to have a special structure. Specifically, suppose that the vertices of K_{2n+1} are the integers $0, 1, \ldots, 2n$. For any subgraph G of K_{2n+1} we may define the *cyclic shift* of G as the subgraph $S(G)$ with

$$S(G) = (\{x+1 \,:\, x \in V(G)\}, \{(x+1, y+1) \,:\, (x,y) \in E(G)\})$$

where all addition is performed modulo $2n+1$.

If G is any graph with n edges, we say that K_{2n+1} can be *cyclically decomposed* into copies of G if there is a subgraph $G' \simeq G$ of K_{2n+1} such that the cyclic shifts $G', S(G'), \dots, S^{2n}(G')$ are edge-disjoint (and thus form a decomposition of K_{2n+1}).

As reported by Rosa [18], the following conjecture is due to Kotzig.

Conjecture 3 (Ringel–Kotzig conjecture)**.** *For any $(n+1)$-vertex tree T the complete graph K_{2n+1} can be cyclically decomposed into copies of T.*

Finally, we can connect this to graceful labellings. If T has a graceful labelling, then T satisfies the Ringel–Kotzig conjecture. Furthermore, if T has an m-graceful labelling, then we can cyclically pack $2m-1$ copies of T into K_{2m-1}. Specifically, let $\psi : V(T) \to [m]$ be an m-graceful labelling of T. Then ψ is also an embedding of T into K_{2m-1}, since ψ is injective. We claim the cyclic shifts of ψ form a packing of $2m-1$ copies of T into K_{2m-1}. Indeed, by symmetry we only have to check that if uv is an edge of T, then $\psi(u)\psi(v)$ is not the image of any $u'v' \in E(T)$ under any non-trivial power of the cyclic shift ψ' of ψ. Without loss of generality we may assume $1 \le \psi(u) < \psi(v) \le m$. If the range of ψ' contains both $\psi(u)$ and $\psi(v)$ then it must contain the interval from $\psi(u)$ to $\psi(v)$. If $u'v' \in E(T)$ satisfies $\psi'(u')\psi'(v') = \psi(u)\psi(v)$, then we have

$$\big|\psi(u) - \psi(v)\big| = \big|\psi'(u') - \psi'(v')\big| = \big|\psi(u') - \psi(v')\big| \,,$$

where the final step uses that the range of ψ' contains the interval from $\psi(u)$ to $\psi(v)$. Now since ψ is graceful, we have $u'v' = uv$, and hence $\psi' = \psi$.

Conjecture 2 was wide open until very recently when in [3, 2] an approximate version for bounded-degree trees was proven.[2]

The main result of [3, 2] was later reproven and extended first by Messuti, Rödl, and Schacht [14] and later (in even greater generality) by Ferber, Lee, and Mousset [6], Kim, Kühn, Osthus, and Tyomkyn [13], and Joos, Kim, Kühn, Osthus [12] using quite different methods. In particular, these methods seem not to hint at any approaches for graceful tree labellings.

1.3. **Our result.** In this paper we prove an approximate version of Conjecture 1 for trees with maximum degree $o\big(\frac{n}{\log n}\big)$. This implies approximate versions of the Ringel–Kotzig and Ringel conjectures.

Theorem 4. *For every $\gamma > 0$ there exist $\eta > 0$ and $n_0 \in \mathbb{N}$ such that the following holds for every $n > n_0$. Suppose that T is an n-vertex tree and $\Delta(T) \le \frac{\eta n}{\log n}$. Then there exists a graceful labelling $\psi : V(T) \to [\lceil(1+\gamma)n\rceil]$.*

---

[2]The main result in [3, 2] is more general and allows almost perfect tilings of a complete graph with an arbitrary family of bounded degree graphs, addressing also the Tree Packing Conjecture of Gyarfás.

This theorem applies to random trees. More precisely, we work with the set $\mathcal{T}_n$ of all labelled trees on the vertex set $\{1, \dots, n\}$. By a classical result of Moon, [15], a tree selected uniformly at random from the set $\mathcal{T}_n$ has maximum degree $o(\log n)$ with probability tending to one as n tends to infinity. In particular Theorem 4 applies to almost all trees.

Our proof of Theorem 4 is an application of the Probabilistic Method, inspired by [3]. More precisely, the proof is an application of the Differential Equations Method (DEM). That is, we run a certain randomized algorithm which sequentially labels the vertices of the n-vertex tree T, and we prove that with high probability this process produces a $[\lceil(1+\gamma)n\rceil]$-graceful labelling of T. As the algorithm progresses with the labelling, the sets of (edge- and vertex-) labels available for future steps keep getting sparser. The key for the analysis of the algorithm is to introduce suitable measures of quasirandomness for these sets, and prove that the sets of available labels evolve in a quasirandom way.

Let us note that DEM has been used extensively in discrete mathematics in particular thanks to the tools developed by Wormald [21]. We do not use Wormald's machinery and it is not clear to us whether that formalism applies in our setting at all. To get a handle on various parameters during the run of the process — this handle provided in other scenarios by DEM itself — we introduce in Section 2 a variant of Hoeffding's bound.

After introducing notation and some preliminary facts in Section 2, we outline the proof of Theorem 4 in Section 3. In Section 3, we also give a detailed description of our labelling algorithm and introduce the key quasirandomness concepts. The missing bits in our proof that the algorithm will produce a labelling required in Theorem 4 are given in Section 4. In Section 5 we suggest various strengthenings of our main result.

2. Notation and auxiliary results

For a graph G, the *order* of G is the number of vertices of G. We write $\Delta(G)$ for the maximum degree of G.

We write $a = b \pm \varepsilon$ when we have $a \in [b-\varepsilon, b+\varepsilon]$. Extending this, and in a slight abuse of notation, we write $a \pm \delta = b \pm \varepsilon$ for the inclusion $[a-\delta, a+\delta] \subseteq [b-\varepsilon, b+\varepsilon]$. We write log for the natural logarithm.

We use $\mathbf{P}[\cdot]$ and $\mathbf{E}[\cdot]$ to denote the probability and the expectation, respectively.

2.1. **Hoeffding's bound.**

The following theorem of Hoeffding [11, Theorem 2] gives bounds on sums of independent real-valued bounded random variables.

Theorem 5. *Let $Y_1, \ldots, Y_n$ be independent random variables with $0 \leq Y_i \leq a_i$ for each i. Let $X = Y_1 + \cdots + Y_n$, and let $\mu = \mathbf{E}X$. Then we have*

$$\mathbf{P}[X - \mu \geq t] \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^n a_i^2}\right) \quad \textit{and}$$
$$\mathbf{P}[X - \mu \leq -t] \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^n a_i^2}\right).$$

Theorem 5 is one of the most commonly used concentration bounds in probabilistic combinatorics. We shall need an extension of this theorem to non-independent random variables. The following lemma is more or less a martingale inequality, but we did not find it in this form in the literature, and so give a proof from first principles.

Lemma 6. *Let Ω be a finite probability space, and $(\mathcal{F}_0, \ldots, \mathcal{F}_n)$ a filtration. Suppose that for each $1 \leq i \leq n$ we have a nonnegative real number a_i, an $\mathcal{F}_i$-measurable random variable U_i satisfying $0 \leq U_i \leq a_i$, and an $\mathcal{F}_{i-1}$-measurable real-valued random variable p_i. Let γ be a real number. Let $X = \sum_{i=1}^n U_i$.*

If $\mathbf{E}[U_i|\mathcal{F}_{i-1}] \leq p_i$ holds almost surely for each $1 \leq i \leq n$, and

$$\gamma \geq \sum_{i=1}^n p_i \tag{1}$$

holds almost surely, then for each $t > 0$ we have

$$\mathbf{P}[X - \gamma \geq t] \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^n a_i^2}\right).$$

Proof. We need the following standard bound. Let Z be a random variable with $0 \leq Z \leq a$ and $\mathbf{E}[Z] = p$. Then for each $u > 0$ we have (see [11, Equation (4.16)])

$$\mathbf{E}\big[\exp(uZ)\big] \leq \exp\left(\tfrac{1}{8}u^2a^2 + up\right). \tag{2}$$

We write $\min(\mathcal{F})$ for the set of minimal non-empty elements in a σ-algebra $\mathcal{F}$. By definition of a σ-algebra, $\min(\mathcal{F})$ forms a partition of Ω. We write $\min^*(\mathcal{F})$ for the sets of $\min(\mathcal{F})$ of positive measure. We use $\min^*(\mathcal{F})$ instead of $\min(\mathcal{F})$ in the calculations below to circumvent conditioning on measure-zero sets.

Our first step is to prove, by induction on n, that under the conditions of the lemma (with that given n), for each $u > 0$ we have

$$\mathbf{E}\big[\exp(uX)\big] \leq \exp\left(\tfrac{1}{8}u^2\sum_{i=1}^n a_i^2 + u\gamma\right). \tag{3}$$

For the base case of $n = 1$, we have $X = U_1$. Then,

$$\begin{aligned}\mathbf{E}\big[\exp(uX)\big] &= \mathbf{E}\big[\exp(uU_1)\big] \\ &= \sum_{F\in\min^*(\mathcal{F}_0)} \mathbf{P}[F]\mathbf{E}\big[\exp(uU_1)\big|F\big] \\ &\overset{(2)}{\leq} \sum_{F\in\min^*(\mathcal{F}_0)} \mathbf{P}[F]\exp\left(\tfrac{1}{8}u^2a_1^2 + up_1(F)\right) \\ &\leq \sum_{F\in\min^*(\mathcal{F}_0)} \mathbf{P}[F]\exp\left(\tfrac{1}{8}u^2a_1^2 + u\gamma\right) = \exp\left(\tfrac{1}{8}u^2a_1^2 + u\gamma\right),\end{aligned}$$

as desired. Now suppose $n \geq 2$. We have

$$\begin{aligned}\mathbf{E}\big[\exp(uX)\big] &= \mathbf{E}\left[\exp(uU_1)\exp\left(u\textstyle\sum_{i=2}^n U_i\right)\right] \\ &= \sum_{F_0\in\min^*(\mathcal{F}_0)}\ \sum_{F\in\min^*(\mathcal{F}_1):F\subseteq F_0} \mathbf{P}[F]\exp\left(uU_1|F\right)\mathbf{E}\left[\exp\left(u\textstyle\sum_{i=2}^n U_i|F\right)\right]. \qquad (4)\end{aligned}$$

Note that since U_1 is $\mathcal{F}_0$-measurable, $U_1|F$ is a constant in the above. Given $F \in \min^*(\mathcal{F}_1)$ contained in $F_0 \in \min^*(\mathcal{F}_0)$, consider the $n-1$ random variables $U_2|F, U_3|F, \ldots, U_n|F$ and the $n-1$ random variables $p_2|F, p_3|F, \ldots, p_n|F$. Let $\gamma' := \gamma - p_1|F$; note that since p_1 is $\mathcal{F}_0$-measurable, $p_1|F = p_1|F_0$ is a constant. By definition, for each $2 \leq i \leq n$ we have $\mathbf{E}[U_i|F, \mathcal{F}_{i-1}] \leq p_i|F$, and because $\gamma \geq \sum_{i=1}^n p_i$ holds almost surely, and $\mathbf{P}[F] > 0$, we have $\gamma' \geq \sum_{i=2}^n p_i|F$ almost surely. Thus the conditions of the induction hypothesis are satisfied. We have

$$\mathbf{E}\left[\exp\left(u\textstyle\sum_{i=2}^n U_i|F\right)\right] \leq \exp\left(\frac{1}{8}u^2\sum_{i=2}^n a_i^2 + u\gamma'\right) = \exp\left(\frac{1}{8}u^2\sum_{i=2}^n a_i^2 + u(\gamma - p_1|F_0)\right).$$

This last is independent of the choice of $F \in \min^*(\mathcal{F}_1)$ with $F \subseteq F_0$, and so substituting into (4) we get

$$\begin{aligned}\mathbf{E}\left[\exp(uX)\right] &\leq \sum_{F_0\in\min^*(\mathcal{F}_0)} \mathbf{P}[F_0]\mathbf{E}\left[\exp(uU_1|F_0)\right]\exp\left(\frac{1}{8}u^2\sum_{i=2}^n a_i^2 + u(\gamma - p_1|F_0)\right) \\ &\overset{(2)}{\leq} \sum_{F_0\in\min^*(\mathcal{F}_0)} \mathbf{P}[F_0]\exp\left(\tfrac{1}{8}u^2a_1^2 + up_1|F_0\right)\exp\left(\frac{1}{8}u^2\sum_{i=2}^n a_i^2 + u(\gamma - p_1|F_0)\right) \\ &= \exp\left(\frac{1}{8}u^2\sum_{i=1}^n a_i^2 + u\gamma\right),\end{aligned}$$

as desired.

Using (3) we can now complete the proof of the lemma. For each $u, t > 0$ we have,

$$\begin{aligned}
\mathbf{P}\big[X - \mu \geq t\big] &= \mathbf{P}\big[\exp(uX) \geq \exp(u(\gamma + t))\big] \\
\boxed{\text{Markov's Inequality}} \quad &\leq \mathbf{E}\left[\exp(uX)\exp(-u(\gamma + t))\right] \\
\boxed{\text{by (3)}} \quad &\leq \exp\Big(\tfrac{1}{8}u^2 \sum_{i=1}^{n} a_i^2 + u\gamma - u\gamma - ut\Big),
\end{aligned}$$

and putting $u = 4t/\sum_{i=1}^{n} a_i^2$ we obtain

$$\mathbf{P}\big[X - \gamma \geq t\big] \leq \exp\left(\frac{2t^2}{\sum_{i=1}^{n} a_i^2} - \frac{4t^2}{\sum_{i=1}^{n} a_i^2}\right),$$

as desired. □

Finally, we are in position to deduce the probabilistic lemma we will need.

Lemma 7. *Let Ω be a finite probability space, and $(\mathcal{F}_0, \ldots, \mathcal{F}_n)$ a filtration. Suppose that for each $1 \leq i \leq n$ we have a nonnegative real number a_i, an $\mathcal{F}_i$-measurable random variable Y_i satisfying $0 \leq Y_i \leq a_i$, nonnegative real numbers μ and ν, and an event $\mathcal{E}$ with $\mathbf{P}[\mathcal{E}] > 0$. Suppose that almost surely, either $\mathcal{E}$ does not occur or $\sum_{i=1}^{n} \mathbf{E}\big[Y_i \big| \mathcal{F}_{i-1}\big] = \mu \pm \nu$. Then for each $t > 0$ we have*

$$\mathbf{P}\Big[\mathcal{E} \text{ and } \Big|\sum_{i=1}^{n} Y_i - \mu\Big| \geq \nu + t\Big] \leq 2\exp\Big(-\frac{2t^2}{\sum_{i=1}^{n} a_i^2}\Big).$$

Proof. Given $Y_1, \ldots, Y_n$, we define random variables $U_1, \ldots, U_n$ by $U_i = Y_i$ if $\mathbf{P}[\mathcal{E}|\mathcal{F}_{i-1}] > 0$, and $U_i = 0$ otherwise. Note that U_i is $\mathcal{F}_i$-measurable for each i. Furthermore, we claim that for each $1 \leq t \leq n$ we have almost surely

$$\sum_{i=1}^{t} \mathbf{E}[U_i|\mathcal{F}_{i-1}] \leq \mu + \nu\,. \tag{5}$$

Indeed, suppose t is minimal such that this statement fails, and let F be a set in $\mathcal{F}_{t-1}$ with $\mathbf{P}[F] > 0$ witnessing its failure. By minimality of t, we have $U_t > 0$, so $\mathbf{P}[\mathcal{E}|F] > 0$. Thus with positive probability, $\mathcal{E}$ occurs and $\sum_{i=1}^{n} \mathbf{E}[U_i|\mathcal{F}_{i-1}] \geq \sum_{i=1}^{t} \mathbf{E}[U_i|\mathcal{F}_{i-1}] > \mu + \nu$, contradicting the assumption of the lemma.

Observe that the random variables $p_i := \mathbf{E}[U_i|\mathcal{F}_{i-1}]$ satisfy (1) for $\gamma = \mu + \nu$ by (5). By Lemma 6, we have

$$\mathbf{P}\Big[\sum_{i=1}^{n} U_i \geq \mu + \nu + t\Big] \leq \exp\Big(-\frac{2t^2}{\sum_{i=1}^{n} a_i^2}\Big).$$

If $\mathcal{E}$ occurs then almost surely $Y_i = U_i$ for each $1 \leq i \leq n$. Therefore, we have

$$\mathbf{P}\Big[\mathcal{E} \text{ and } \sum_{i=1}^{n} Y_i \geq \mu + \nu + t\Big] \leq \exp\Big(-\frac{2t^2}{\sum_{i=1}^{n} a_i^2}\Big).$$

The same argument applied to the random variables $a_i - Y_i$ gives

$$\mathbf{P}\Big[\mathcal{E} \text{ and } \sum_{i=1}^{n} Y_i \le \mu - \nu - t\Big] \le \exp\Big(- \frac{2t^2}{\sum_{i=1}^{n} a_i^2}\Big)\,,$$

and so the lemma statement holds by the union bound. □

2.2. **Cutting a tree.** The following lemma (variants of which are well-known) tells us that trees can be easily separated into small components.

Lemma 8. *For any $\varepsilon \in (0,1)$, any n such that $\varepsilon n \ge 2\log n$, and any tree T with $\Delta(T) \le \frac{\varepsilon^2 n}{4\log n}$ and $v(T) \le n$, there exists a set $\mathcal{R}$ of edges of T with $|\mathcal{R}| \le \varepsilon v(T)$ such that the components of $T - \mathcal{R}$ have order at most $\frac{\varepsilon n}{\log n}$.*

Proof. We prove the lemma by induction on $v(T)$. The statement is trivially true for $v(T) \le \frac{\varepsilon n}{\log n}$, so we may assume $v(T) > \frac{\varepsilon n}{\log n}$. It is enough to show that there exists one edge $uv \in T$ such that one of the two components of $T - uv$ has order between $2\varepsilon^{-1}$ and $\frac{\varepsilon n}{\log n}$, since then the statement follows by applying the induction hypothesis to the other component.

We find the edge uv by the following 'walk' procedure. We start with a leaf vertex u_1 and its neighbour v_1. Now for each $t \ge 1$ in succession we do the following. If the component of $T - u_t v_t$ containing v_t has order between $2\varepsilon^{-1}$ and $\frac{\varepsilon n}{\log n}$, we stop and return $uv = u_t v_t$. If not, we set $u_{t+1} = v_t$ and v_{t+1} to be the neighbour of u_{t+1} not equal to u_t which maximises the order of the component of $T - u_{t+1}v_{t+1}$ containing v_{t+1}.

If at time t the component of $T - u_t v_t$ containing v_t has more than $\frac{\varepsilon n}{\log n}$ vertices, then by averaging, the component of $T - u_{t+1}v_{t+1}$ containing v_{t+1} has at least

$$\tfrac{1}{\Delta(T)}\big(\tfrac{\varepsilon n}{\log n} - 1\big) \ge \tfrac{1}{\Delta(T)} \cdot \tfrac{\varepsilon n}{2\log n} \ge 2\varepsilon^{-1}$$

vertices, where the first inequality is by choice of n and the second by the bound on $\Delta(T)$. Thus the algorithm finds the desired uv. □

3. Setup

Before we embark on the proof of Theorem 4, we outline the main ideas. We write $\tilde{n} = \lceil (1+\gamma)n \rceil$. We wish to label our tree in a random process, at each step labelling one new vertex which has one labelled neighbour (so that the set of labelled vertices is always a subtree), and choosing the new vertex label to be admissible, that is, to avoid re-using either the vertex label or the induced edge label. We keep track of the sets of vertex and induced edge labels which remain available at the t^{th} step, for which we will write A_t and C_t respectively. We will show that A_t looks like a uniform random subset of $[\tilde{n}]$ of cardinality $\tilde{n} - t$ and that C_t looks like a uniform random subset of $[\tilde{n} - 1]$ of cardinality $\tilde{n} - t - 1$. The reason why we do this is that if A_t and C_t were really uniformly distributed it would be easy to show that at every step there is (with probability very close to 1) an admissible choice for the new label.

Choosing the new labels uniformly from all admissible labels will not lead to random-looking sets A_t and C_t. Let us illustrate this in the initial situation when, starting with $A_0 = [\tilde{n}]$ and $C_0 = [\tilde{n}-1]$, we label first two neighboring vertices v_1 and v_2. Then the first created edge label on v_1v_2 will be assigned the smallest possible value of 1 with probability $\frac{2}{\tilde{n}}$, but the largest possible value of $\tilde{n}-1$ with probability only $\frac{2}{\tilde{n}(\tilde{n}-1)}$, and the probabilities of labels with intermediate values interpolate these two extremes. This effect will persist, so that C_t will rapidly become non-uniform, being much sparser for small labels than large. That will in turn cause A_t to evolve in a non-uniform way as well.

But observe that if uv is an edge of T, if we choose $\psi(u)$ a uniformly random vertex label, and then choose $\psi(v)$ randomly 'close to' $\tilde{n} - \psi(u)$ then the distributions of each of $\psi(u)$, $\psi(v)$ and the induced edge label $\big|\psi(u) - \psi(v)\big|$ are close to the uniform distribution. Thus we would like to label our tree such that most edges and vertices are labelled in about this way.

We will label the vertices of T in the order $v_1, \ldots, v_n$ (which we will determine), and in each case we choose an admissible label in a certain short interval $J(v_i)$ within $[\tilde{n}]$. We choose the $J(v_i)$ to be uniformly distributed[3] over $[\tilde{n}]$, and such that for most v_i the intervals $J(v_i)$ and $J(v_j)$, where v_j is the parent of v_i in the ordering, are equally far from and on opposite sides of $\frac{1}{2}(1+\gamma)n$.

In reality, we certainly do not choose vertex labels uniformly at random in the intervals; there is a great deal of dependency in the label choices. However, that dependency is sequential: conditioning on the labelling history of the first $t-1$ vertices, we know the distribution of the t^{th} vertex, and we will be able to show that the average of these distributions is close to uniform (in a sense which we will make precise later). This enables us to apply Lemma 7, which gives us concentration results that imply our quasirandomness properties are maintained. A similar statement applies to the edge labels.

3.1. **Preparation.** We now describe the general setup which we use for proving Theorem 4. We first reduce to the following special case.

Theorem 9. *For every $\gamma > 0$ such that γ^{-1} is an integer, there exist $\eta > 0$ and $n_0 \in \mathbb{N}$ such that the following holds. Let $\mu = \big\lceil \exp\big(10^8\gamma^{-4}\big)\big\rceil^{-1}$, let $\delta_0 = \mu^{2/\mu}$, and let $\delta = \delta_0^{10}$. For every $n > n_0$ divisible by $2\delta^{-1}\gamma^{-1}$, if T is an n-vertex tree with $\Delta(T) \leq \frac{\eta n}{\log n}$, there exists a graceful labelling $\psi : V(T) \to [(1+\gamma)n]$.*

Before proving this, we show why it implies Theorem 4.

Proof of Theorem 4. Given $\gamma > 0$, let $\gamma' = 1/\lceil 2\gamma^{-1}\rceil$. Let $\eta > 0$ and $n_0' \in \mathbb{N}$ be returned by Theorem 9 for input γ'. Let $\mu = \big\lceil \exp\big(10^8\gamma'^{-4}\big)\big\rceil^{-1}$, let $\delta_0 = \mu^{2/\mu}$, and let $\delta = \delta_0^{10}$. Let $n_0 \geq n_0'$ be the smallest integer such that

$$(1+\gamma')\big(n_0 + 2\delta^{-1}\gamma'^{-1}\big) \leq (1+\gamma)n_0$$

[3] A technical complication arises in that the extremes of the interval $[\tilde{n}]$ are not covered as much as the rest; we will ignore this for now.

and such that $\frac{\eta n}{\log n} \geq 2$. Given $n \geq n_0$, and an n-vertex tree T with $\Delta(T) \leq \frac{\eta n}{\log n}$, let n' be the smallest integer which is at least n and divisible by $2\delta^{-1}\gamma'^{-1}$. Let T' be an n'-vertex tree obtained by attaching a path with $n'-n$ vertices to a leaf of T. Theorem 9 applies, so there is a graceful labelling of T' with $(1+\gamma')n' \leq (1+\gamma)n$ labels. The induced labelling of T is also graceful, as desired. □

We now give the setup we will use to prove Theorem 9.

μ, δ_0, δ, ε, η

Setup 10. *Given $\gamma > 0$ such that γ^{-1} is an integer, we choose $\mu = \lceil \exp\left(10^8\gamma^{-4}\right)\rceil^{-1}$. We set $\delta_0 = \mu^{2/\mu}$, $\delta = \delta_0^{10}$, $\varepsilon = \delta^{10}$, and $\eta = \varepsilon^{10}$. Set*

$$n_0 = 2^{10/\eta^2}\,. \tag{6}$$

Suppose now that $n > n_0$ divisible by $2\delta_0^{-10}\gamma^{-1}$ is given. Note that since μ^{-1} is an integer, δ_0^{-1} is also an integer, so this is possible.

δ_i

For each $0 \leq i \leq n$ set $\delta_i = \mu^{(2n-i)/(\mu n)}$. Observe that for $i = 0$, this definition is consistent with the previous definition of δ_0.

ℓ, m

Let $\ell = \delta_0^2 n$, and $m = \delta_0^2\ell$. Because n is divisible by δ_0^{-4}, and δ_0^{-1} is an integer, these two quantities are integers and m divides ℓ.

$\tilde{n}$

Let $\tilde{n} = (1+\gamma)n$. Because $2\delta_0^{-4}\gamma^{-1}$ divides n, in particular $\tilde{n}$ is an integer multiple of $2m$.

$\mathbb{A}, \mathbb{C}$

Let $\mathbb{A} := [\tilde{n}]$, and $\mathbb{C} := [\tilde{n}-1]$. We will choose vertex labels from $\mathbb{A}$, and edge labels from $\mathbb{C}$.

$\mathcal{I}_{\mathsf{V}}$

Let $\mathcal{I}_{\mathsf{V}}$ be the collection of intervals of length $m-1$ (i.e., size m) in $\mathbb{A}$ whose first elements are in the set

$$\{1, m+1, 2m+1, \ldots, \tilde{n}-m+1\}\,,$$

$\mathcal{I}_{\mathsf{E}}$

and let $\mathcal{I}_{\mathsf{E}}$ be the collection of intervals of length $m-1$ in $\mathbb{C}\cup\{0\}$ whose first elements are in the set

$$\{0, m, \ldots, |\mathbb{C}|-m+1\}\,.$$

$\mathcal{J}$

Finally, let $\mathcal{J}$ be the set of intervals of length $\ell-1$ in $\mathbb{A}$ whose first elements are in the set

$$\left\{1, m{+}1, 2m{+}1, \ldots, \tfrac{\tilde{n}}{2}{-}m{-}\ell{+}1, \tfrac{\tilde{n}}{2}{-}\ell{+}1, \tfrac{\tilde{n}}{2}{+}1, \tfrac{\tilde{n}}{2}{+}m{+}1, \tfrac{\tilde{n}}{2}{+}2m{+}1, \ldots, \tilde{n}{-}m{-}\ell{+}1, \tilde{n}{-}\ell{+}1\right\}.$$

(See Figure 3.1.)

$\overline{J}$

For each $J \in \mathcal{J}$, we define the complementary interval *$\overline{J} \in \mathcal{J}$ to be the (unique) interval such that the sum of the elements of $J \cup \overline{J}$ is $\ell(\tilde{n}+1)$. By definition, J and $\overline{J}$ are disjoint, one having only elements less than or equal to $\frac{1}{2}\tilde{n}$ and the other having only elements greater than $\frac{1}{2}\tilde{n}$.*

Note that

$$|\mathcal{I}_{\mathsf{V}}| = |\mathcal{I}_{\mathsf{E}}| = \frac{\tilde{n}}{m} \quad \text{and} \quad |\mathcal{J}| = \frac{\tilde{n}}{m} - 2\left(\frac{\ell}{m}-1\right)\,. \tag{7}$$

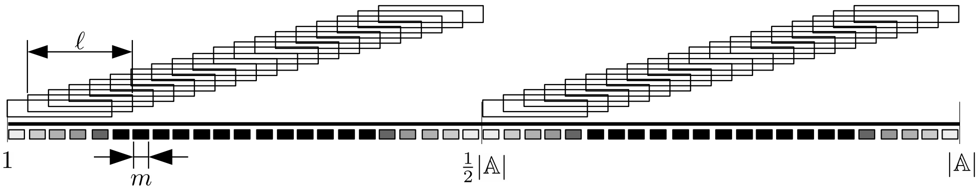


Figure 3.1. The intervals $\mathcal{I}_{\mathsf{V}}$ (below the line) and $\mathcal{J}$ (above the line). The shade of grey used for an interval $I \in \mathcal{I}_{\mathsf{V}}$ corresponds to the number of intervals of $\mathcal{J}$ containing I.

In analysing our labelling algorithm, the values $(\delta_i)_{i=1}^n$ will quantify errors in our quasirandomness properties in steps $i = 1, \dots, n$.[4] Our choice of δ_i ensures that the following holds. For each $1 \le t \le n$, we have

$$\sum_{i=1}^{\lceil \frac{t}{\delta n} \rceil} \delta \mu^{-1} \delta_{i\delta n} < \tfrac{1}{100} \delta_t \,. \tag{8}$$

This inequality will be important in showing that our error terms do not grow too fast.

3.2. **Assigning intervals of labels to vertices.** Our next step is to give an order on $V(T)$ and for each vertex the promised 'small interval' in which we will eventually choose its label. As mentioned in the outline of the proof, for most edges we will label the two endpoints from complementary small intervals. Specifically, we will do this for all edges not in the set $\mathcal{R}$ given by the following lemma.

Lemma 11. *Assume Setup 10. Given an n-vertex tree T with $\Delta(T) \le \frac{\eta n}{\log n}$, there exists a set $\mathcal{R} \subseteq E(T)$, an ordering $V(T) = \{v_1, \dots, v_n\}$ of the vertices of T, and a collection of intervals $J(v_i) \in \mathcal{J}$ for each $1 \le i \le n$ with the following properties.* $\mathcal{R}$ $v_1, \dots, v_n$

(PRE1) $|\mathcal{R}| \le \varepsilon n$.

(PRE2) For each $i \ge 2$, the vertex v_i has exactly one neighbour in the set $\{v_1, \dots, v_{i-1}\}$. We call this vertex the parent *of v_i, and denote it by v_i^+.* v^+

(PRE3) If $v_i v_j \in E(T) \setminus \mathcal{R}$ then $|i - j| \le \frac{\varepsilon n}{\log n}$.

(PRE4) For each interval $S \subseteq [n]$ and each $J \in \mathcal{J}$, we have $\sum_{i \in S} \mathbb{1}_{J(v_i)=J} = \frac{|S|}{|\mathcal{J}|} \pm \delta^2 n$.

(PRE5) For each $v_i v_j \in E(T) \setminus \mathcal{R}$ we have $J(v_i) = \overline{J(v_j)}$.

Proof. Given a tree T, let $\mathcal{R}$ be the set of edges of T returned by Lemma 8 with input ε. Then we have $|\mathcal{R}| \le \varepsilon n$, giving (PRE1). Let $T_1, \dots, T_s$ be the components of $T - \mathcal{R}$.

Let v_1 be an arbitrary vertex of $V(T)$. Now for each $i \ge 2$ in turn, we choose v_i to be a vertex of $V(T) \setminus \{v_1, \dots, v_{i-1}\}$ which has a neighbour in $\{v_1, \dots, v_{i-1}\}$, if possible choosing a vertex in the same component of $T - \mathcal{R}$ as v_{i-1}. (PRE2) follows from this construction.

Suppose now that in the above construction we cannot choose v_i in the same component of $T - \mathcal{R}$ as v_{i-1}. Then this is because we already chose all vertices of that component. It

[4] see Claim 2

follows that each component forms an interval in our ordering on the vertices. In particular, if $v_iv_j \in E(T) \setminus \mathcal{R}$ then v_iv_j is in one component T_k of $T - \mathcal{R}$. Then Lemma 8 tells us that $|i-j| \le v(T_k) \le \frac{\varepsilon n}{\log n}$, giving (PRE3).

We now take an arbitrary proper colouring of T with two colours, red and blue. For each component T_k of $T-\mathcal{R}$ we choose independently an interval $J(T_k)$ uniformly at random from $\mathcal{J}$. For each red $v_j \in T_k$ we set $J(v_j) = J(T_k)$, and for each blue $v_j \in T_k$ we set $J(v_j) = \overline{J(T_k)}$. This gives (PRE5).

It remains to show that with positive probability we obtain (PRE4). To that end, let $S \subseteq [n]$ be an interval, and J be an element of $\mathcal{J}$. We view S as an interval in $(v_1, \dots, v_n)$. Consider the intersections of various components T_k with S. Since these components form intervals in $(v_1, \dots, v_n)$, at most two components have non-empty intersection with S and are not contained in S. These at most two components contain at most $2\frac{\varepsilon n}{\log n} < \delta^5 n$ vertices. Let $\mathcal{K}$ index the components which are contained in S, i.e., $T_k \subseteq S$ for each $k \in \mathcal{K}$. For each $k \in \mathcal{K}$ define Y_k as follows. If $J(T_k) = J$, let Y_k be equal to the number of red vertices in T_k; if $\overline{J(T_k)} = J$ let Y_k be equal to the number of blue vertices in T_k, and otherwise let $Y_k = 0$.

Because the sets $J(T_k)$ are chosen independently, we get that Y_k are independent random variables, with $0 \le Y_k \le v(T_k)$ for each $k \in \mathcal{K}$. We have $\mathbf{E}[Y_k] = \frac{v(T_k)}{|\mathcal{J}|}$ for each $k \in \mathcal{K}$. Because $v(T_k) \le \frac{\varepsilon n}{\log n}$ for each $k \in \mathcal{K}$, and $\sum_{k\in\mathcal{K}} v(T_k) \le n$, we have

$$\sum_{k\in\mathcal{K}} v(T_k)^2 \le \frac{n}{\frac{\varepsilon n}{\log n}} \cdot \left(\frac{\varepsilon n}{\log n}\right)^2 = \frac{\varepsilon n^2}{\log n}\,.$$

Let $X = \sum_{k\in\mathcal{K}} Y_k$. By Theorem 5 we have

$$\mathbf{P}\left[|X - \mathbf{E}[X]| > \tfrac{1}{4}\delta^4 n\right] \le 2\exp\left(-\tfrac{1}{8}\cdot\delta^8 n^2 \cdot \tfrac{\log n}{\varepsilon n^2}\right) = 2\exp\left(-\tfrac{1}{8\delta^2}\log n\right) \le n^{-10}\,,$$

where we used the relation between ε and δ from Setup 10.

We see that with probability at least $1 - n^{-10}$ we have

$$X = \frac{|S| \pm \delta^5 n}{|\mathcal{J}|} \pm \tfrac{1}{4}\delta^4 n = \frac{|S|}{|\mathcal{J}|} \pm \tfrac{1}{2}\delta^4 n\,.$$

Since the number of vertices $v_i \in S$ with $J(v_i) = J$ is X, we conclude that for the chosen S and J the probability that (PRE4) fails is at most n^{-10}. Taking the union bound over $|\mathcal{J}| \le n$ choices of J and $\binom{n}{2}$ choices of S we get that with probability at least $1 - n^{-7}$ we have (PRE4), as desired. $\square$

3.3. **Quasirandom properties.** We give two quasirandomness properties of a pair (A, C) of subsets of $\mathbb{A}$ and $\mathbb{C}$ respectively. The reader should think of A and C as being respectively the vertex labels which have not yet been used, and the edge labels which have not yet been used, at some stage in the labelling process. The first property, (QUASI1), states that the density of C on each interval $I_{\mathsf{E}} \in \mathcal{I}_{\mathsf{E}}$ is approximately $\frac{|C|}{|\mathbb{C}|}$. The second, (QUASI2), states that certain 'structures' involving both A and C occur about as often as one would expect in random sets of the same densities. To make this precise, we need to define 'structure'.

Definition 12 (structure, pattern, chosen label)**.** *A* structure X *is a graph whose vertices are labelled with either pairwise distinct elements of* $\mathbb{A}$ *(we call such vertices* fixed*) or pairwise distinct intervals in* $\mathbb{A}$ *(we call such vertices* free*), and whose edges are labelled with either pairwise distinct elements of* $\mathbb{C}$ *(*fixed*) or with distinct choices from* $\mathfrak{e}_1$ *or* $\mathfrak{e}_2$ *(*free*). When dealing with structures, we identify vertices and edges with their labels. So, vertices in* X *are numbers or intervals and edges in* X *are numbers or special free symbols* $\mathfrak{e}_1$ *and* $\mathfrak{e}_2$*.*

In any structure, we require that if u *and* v *are fixed vertices and* uv *is a fixed edge, we have* $uv = |u - v|$*, and such that each free edge has one endpoint fixed and the other free. We write* $\mathrm{Diff}(\mathfrak{e}_i; X) := |a - \min(I)|$ *for such a free edge* $\mathfrak{e}_i$ *with fixed endpoint labelled* a *and free endpoint labelled* I*. We write* $\mathrm{Free}_{\mathsf{V}}(X)$ *for the set of free vertex labels in* X*, and* $\mathrm{Free}_{\mathsf{E}}(X)$ *for the set of free edge labels in* X*. We write* $\mathrm{free}(X)$ *for the total number of free (vertex- or edge-) labels in* X*.* $\mathrm{Diff}(\mathfrak{e}_i; X)$ $\mathrm{Free}_{\mathsf{V}}(X)$ $\mathrm{Free}_{\mathsf{E}}(X)$ $\mathrm{free}(X)$

Suppose that we are given structures X *and* X'*. We say that* X' follows the pattern X *if* $\mathrm{free}(X') = 0$ *and there is a graph isomorphism* ϱ *from* X' *to* X *such that for each vertex* $u \in X'$ *the labels satisfy either* $u = \varrho(u)$ *or* $u \in \varrho(u)$*, and for each edge* $uv \in X'$ *the labels satisfy either* $uv = \varrho(uv)$ *or* $uv \in \varrho(uv)$*. We call the labels of vertices* $u \in X'$ *such that* $u \in \varrho(u)$*, and of edges* $uv \in X'$ *such that* $uv \in \varrho(uv)$*,* chosen labels*. We write* $\mathrm{Ch}_{\mathsf{V}}(X'; X)$ *for the set of chosen vertex labels in* X' *and* $\mathrm{Ch}_{\mathsf{E}}(X'; X)$ *for the set of chosen edge labels in* X'*.* $\mathrm{Ch}_{\mathsf{V}}(X'; X)$ $\mathrm{Ch}_{\mathsf{E}}(X'; X)$

Given $A \subseteq \mathbb{A}$ *and* $C \subseteq \mathbb{C}$*, and a structure* X*, we write*

$$X(A, C) \tag{9}$$

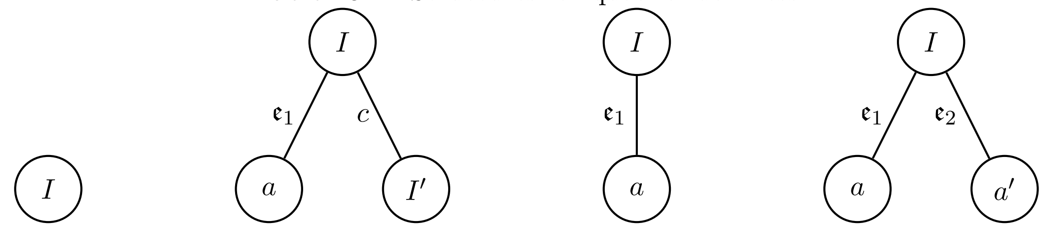
for the set of all structures following the pattern X *whose chosen vertex labels are in* A *and whose chosen edge labels are in* C*.*

The patterns we are interested in are $X_1 [\![I]\!]$, $X_2 [\![a, I, c, I']\!]$, $X_3 [\![a, I]\!]$ and $X_4 [\![a, a', I]\!]$, where we have $a, a' \in \mathbb{A}$, $c \in \mathbb{C}$ and $I, I' \in \mathcal{I}_{\mathsf{V}} \cup \mathcal{J}$ are distinct (see Figure 3.2). The reason for using the double square brackets is to distinguish the parameters from the notion introduced in (9); we can for example write $X_2 [\![a, I, c, I']\!] (A, C)$ for the set of all structures following the pattern $X_2 [\![a, I, c, I']\!]$ whose chosen vertex labels are in A and whose chosen edge label is in C. $X_1 [\![I]\!]$, $X_2 [\![a, I, c, I']\!]$, $X_3 [\![a, I]\!]$, $X_4 [\![a, a', I]\!]$,

We let $\mathcal{X}$ be the set of structures of the form $X_1 [\![I]\!]$, $X_2 [\![a, I, c, I']\!]$, $X_3 [\![a, I]\!]$ and $X_4 [\![a_1, a_2, I]\!]$ where a, a' are distinct elements of $\mathbb{A}$, where c is an element of $\mathbb{C}$, and where $I, I' \in \mathcal{I}_{\mathsf{V}}$ are distinct. $\mathcal{X}$

Observe that the density $\frac{|A \cap I|}{|I|}$ of A on $I \in \mathcal{I}_{\mathsf{V}}$ is equal to $\frac{|X_1 [\![I]\!](A,C)|}{|X_1 [\![I]\!](\mathbb{A},\mathbb{C})|}$. For the other three structures, one should think of $a \in \mathbb{A}$ as the chosen label for some vertex v_i of T. Then $|X_3 [\![a, I]\!] (A, C)|$ is the number of ways to give a neighbour v_j of v_i a label in I which has not yet been used and which induces a label that has not yet been used (which is obviously a useful quantity to control). If one thinks of a as being the chosen label of a vertex v_i, then $X_2 [\![a, J(v_j), c, J(v_k)]\!]$ is the number of ways to label a child v_j and grandchild v_k within their

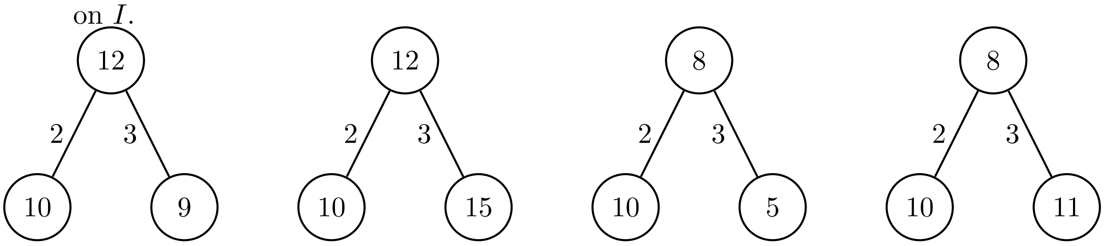
FIGURE 3.2. Structures for quasirandomness



$X_1 [\![I]\!]$ $X_2 [\![a, I, c, I']\!]$ $X_3 [\![a, I]\!]$ $X_4 [\![a, a', I]\!]$

chosen intervals, not re-using vertex or edge labels used previously, and using the edge label c on $v_j v_k$. The quantity $X_4 [\![a, a', J(v_j)]\!]$ plays a similar rôle, except that we fix the vertex label used for v_k to be a' rather than the edge label for $v_j v_k$.

FIGURE 3.3. The four structures following $X_2 [\![10, I, 3, I']\!]$ (for suitable $I \ni 8, 12$ and $I' \ni 5, 9, 11, 15$) which induce label 2 on $\mathfrak{e}_1$. The two leftmost structures are also the only structures following $X_2 [\![10, I, 3, I']\!]$ that have label 12 on I.



For the analysis, two points will be important. First, we have

$$|X(A, C)| \leq \max_{I \in \mathrm{Free}_{\mathsf{V}}(X)} |I| = m \,. \tag{10}$$

for any structure $X \in \mathcal{X}$ and sets A and C. Second, for each $X \in \mathcal{X}$, given any free vertex and a label in $\mathbb{A}$ there is at most one structure in $X(A, C)$ in which the corresponding vertex has that label, and similarly for any free edge and label in $\mathbb{C}$ there are at most two structures in $X(A, C)$ in which the corresponding edge has that label. Example of structures with a specific label on a specific free vertex or edge are given in Figure 3.3. That is, in Figure 3.3 we show the only four structures following $X_2 [\![10, I, 3, I']\!]$ (for suitable $I \ni 8, 12$ and $I' \ni 5, 9, 11, 15$) which induce label 2 on $\mathfrak{e}_1$. The two leftmost structures are also the only structures following $X_2 [\![10, I, 3, I']\!]$ that have label 12 on I. Since I and I' overlap, they cannot be two distinct sets of $\mathcal{I}_{\mathsf{V}}$, and thus $X_2 [\![10, I, 3, I']\!] \notin \mathcal{X}$. When requiring that I and I' are such that $X_2 [\![10, I, 3, I']\!] \in \mathcal{X}$, there would be at most two solutions to the edge problem, and at most one solutions to the vertex problem.

Our quasirandomness conditions are as follows.

Definition 13. *(Quasirandomness conditions) A pair of sets $A \subseteq \mathbb{A}$ and $C \subseteq \mathbb{C}$ is α-* quasirandom *if*

(QUASI1) For all $I_{\mathsf{E}} \in \mathcal{I}_{\mathsf{E}}$ we have $|I_{\mathsf{E}} \cap C| = m \cdot \frac{|A|}{\tilde{n}} \pm \alpha m$.

(QUASI2) For all $X \in \mathcal{X}$ we have $|X(A,C)| = |X(\mathbb{A},\mathbb{C})|\big(|A|/\tilde{n}\big)^{\mathrm{free}(X)} \pm \alpha m$.

It may seem strange that we only have $|A|$ in the estimates, and not $|C|$, but we will prove that throughout our labelling process, $|A|$ and $|C|$ are close enough for the difference to be immaterial. Note that we only insist on counts of structures in $\mathcal{X}$ being preserved, so that $I, I' \in \mathcal{I}_{\mathsf{V}}$, even though we defined our structures allowing $I, I' \in \mathcal{I}_{\mathsf{V}} \cup \mathcal{J}$. The following claim lets us deduce the latter (in the cases we need it) from the former.

Lemma 14. *Suppose that (A,C) is α-quasirandom, that a and a' are distinct elements of $\mathbb{A}$, that $c \in \mathbb{C}$, and that $J \in \mathcal{J}$. If $\ell \geq 3\alpha^{-1}$, then we have*

$$\big|X_2\left[\!\left[a, J, c, \overline{J}\right]\!\right](A,C)\big| = \big(|A|/\tilde{n}\big)^3\big|\{(b,b') \in J \times \overline{J} : |b-b'| = c\}\big| \pm 3\alpha\ell\,, \tag{11}$$

$$\big|X_3\left[\!\left[a, J\right]\!\right](A,C)\big| = \big(|A|/\tilde{n}\big)^2\ell \pm 2\alpha\ell \quad \textit{and} \tag{12}$$

$$\big|X_4\left[\!\left[a, a', J\right]\!\right](A,C)\big| = \big(|A|/\tilde{n}\big)^3\ell \pm 2\alpha\ell\,. \tag{13}$$

Proof. For (12), observe that summing (QUASI2) with $X_3\left[\!\left[a, I\right]\!\right]$ over $I \in \mathcal{I}_{\mathsf{V}}$ with $I \subseteq J$, we have

$$\big|X_3\left[\!\left[a, J\right]\!\right](A,C)\big| = \big(|A|/\tilde{n}\big)^2\big|X_3\left[\!\left[a, J\right]\!\right](\mathbb{A},\mathbb{C})\big| \pm \alpha\ell$$

and $\big|X_3\left[\!\left[a, J\right]\!\right](\mathbb{A},\mathbb{C})\big| = \ell \pm 1 = \ell \pm \alpha\ell$, where the error term comes from the possibility that $a \in J$. This gives (12). Now $\big|X_4\left[\!\left[a, a', J\right]\!\right](\mathbb{A},\mathbb{C})\big| = \ell \pm 3 = \ell \pm \alpha\ell$, where the error term comes from the possibility that some of a, a', and $\frac{a+a'}{2}$ are in J, and by the same logic we obtain (13).

For (11), observe that

$$\big|X_2\left[\!\left[a, J, c, \overline{J}\right]\!\right](A,C)\big| = \sum_{\substack{I, I' \in \mathcal{I}_{\mathsf{V}} \\ I \subseteq J, I' \subseteq \overline{J}}} \big|X_2\left[\!\left[a, I, c, I'\right]\!\right](A,C)\big|$$

and furthermore in the sum at most $2\ell/m$ terms are non-zero. Summing (QUASI2), and observing that we only need to sum the αm error term $2\ell/m$ times, we obtain

$$\big|X_2\left[\!\left[a, J, c, \overline{J}\right]\!\right](A,C)\big| = \big(|A|/\tilde{n}\big)^3\big|X_2\left[\!\left[a, J, c, \overline{J}\right]\!\right](\mathbb{A},\mathbb{C})\big| \pm 2\alpha\ell\,.$$

Now $\big|X_2\left[\!\left[a, J, c, \overline{J}\right]\!\right](\mathbb{A},\mathbb{C})\big| = \big|\{(b,b') \in J \times \overline{J} : |b-b'| = c\}\big| \pm 3$, where again the error term comes from the possibility $a \in J \cup \overline{J}$ and from the at most two labels in J at distance c from a. By choice of ℓ, (11) follows. ☐

3.4. **The algorithm.** The idea of the labelling algorithm is now straightforward. We will label the vertices in order, choosing at each time t to give v_t a vertex label in $J(v_t)$ which has not previously been used, and which induces an edge label on $v_t v_t^+$ which has not previously been used. Unfortunately, this simple version of the algorithm does not quite maintain the quasirandom properties mentioned above, because different intervals $I \in \mathcal{I}_{\mathsf{V}}$ are contained in different numbers of intervals $J \in \mathcal{J}$; this is the 'technical complication' mentioned in Footnote 3. Indeed, in Figure 3.1 we saw that the labels around the extremes and the centre of $\mathbb{A}$ are used less frequently than those in the intermediate ranges. To correct this we introduce a distribution $\mathrm{Cor}_{\mathsf{V}}$ on $\mathcal{I}_{\mathsf{V}} \cup \{*\}$. At each time we in addition randomly sample from $\mathrm{Cor}_{\mathsf{V}}$, and either do nothing (if $\mathrm{Cor}_{\mathsf{V}}$ returns $*$) or remove a randomly chosen so far unused vertex label from I (if $\mathrm{Cor}_{\mathsf{V}}$ returns I).

Most of the mass of $\mathrm{Cor}_{\mathsf{V}}$ is on $*$. That means that the total number of labels removed during the run will be tiny compared to the 'extra' γn labels we are given by Theorem 9. On the other hand, near the extremes and the centre of $\mathbb{A}$ a substantial proportion of the vertex labels will be removed by $\mathrm{Cor}_{\mathsf{V}}$ without serving as vertex labels. These two contrasting properties are consistent since only a small number of vertex labels are near the extremes and the centre of $\mathbb{A}$.

An analogous complication arises when dealing with edge labels $\mathbb{C}$. Indeed, we can see, for example, that for $\tilde{n}-1 \in \mathbb{C}$ to appear as an edge label on uv, we must have chosen $J(u)$ and $J(v)$ to be the two extreme intervals of $\mathcal{J}$, while for $\frac{n}{2} \in \mathbb{C}$ (or any other edge label not close to 0 or $\tilde{n}$) there are $\frac{2\ell^2}{m^2}$ possible choices of the pair $J(u)$, $J(v)$. To deal with these discrepancies, we introduce a suitable distribution $\mathrm{Cor}_{\mathsf{E}}$ on $\mathcal{I}_{\mathsf{E}} \cup \{*\}$. In analogy with $\mathrm{Cor}_{\mathsf{V}}$, we remove random vertex labels from intervals I_{E} chosen according to $\mathrm{Cor}_{\mathsf{E}}$.

We will see that our labelling algorithm labels v_i more or less uniformly in $J = J(v_i)$. If $v_i v_i^+$ is not in $\mathcal{R}$, then the induced edge label on $v_i v_i^+$ is chosen (approximately) from the
$\mathfrak{Q}(J,c)$ distribution in which the probability of choosing c is

$$\mathfrak{Q}(J,c) := \frac{\left|\left\{(a,a') \in J \times \overline{J} : |a-a'| = c\right\}\right|}{\ell^2} .$$

For convenience, we define $\mathfrak{Q}(J,0)$ according to the above formula (even though 0 is not an edge label).

Observe that for any edge label c,

$$\text{there are at most } {}^2\!/\!\delta_0^2 \text{ sets } J \in \mathcal{J} \text{ such that } \mathfrak{Q}(J,c) > 0 . \tag{14}$$

Suppose now that $|c-c'| \le 2m$. Then for any $J \in \mathcal{J}$ we have

$$\left|\mathfrak{Q}(J,c) - \mathfrak{Q}(J,c')\right| \le \frac{2m}{\ell^2} , \tag{15}$$

since $J \cap \overline{J} = \emptyset$ by definition of $\mathcal{J}$. We consider this a 'small error', and use the approximation $\sum_{c\in I_{\mathsf{E}}} \mathfrak{Q}(J(v_i),c) \approx m\mathfrak{Q}(J(v_i),\min(I_{\mathsf{E}}))$ for each $I_{\mathsf{E}} \in \mathcal{I}_{\mathsf{E}}$ in order to simplify the definition
$\mathrm{Cor}_{\mathsf{V}}$ of $\mathrm{Cor}_{\mathsf{E}}$ below. Our two correction distributions are then defined by the following formulas,
$\mathrm{Cor}_{\mathsf{E}}$

$$\mathbf{P}[\mathrm{Cor}_{\mathsf{V}} = I] = \frac{1 - \frac{m}{\ell}|\{J \in \mathcal{J} : I \subseteq J\}|}{|\mathcal{J}|} \qquad \text{for each } I \in \mathcal{I}_{\mathsf{V}}, \tag{16}$$

$$\mathbf{P}[\mathrm{Cor}_{\mathsf{E}} = I_{\mathsf{E}}] = \frac{1 - m\sum_{J\in\mathcal{J}} \mathfrak{Q}(J,\min(I_{\mathsf{E}}))}{|\mathcal{J}|} \qquad \text{for each } I_{\mathsf{E}} \in \mathcal{I}_{\mathsf{E}}, \tag{17}$$

and $\mathbf{P}[\mathrm{Cor}_{\mathsf{V}} = *] = 1 - \sum_{I \in \mathcal{I}_{\mathsf{V}}} \mathbf{P}[\mathrm{Cor}_{\mathsf{V}} = I]$, and $\mathbf{P}[\mathrm{Cor}_{\mathsf{E}} = *] = 1 - \sum_{I_{\mathsf{E}} \in \mathcal{I}_{\mathsf{E}}} \mathbf{P}[\mathrm{Cor}_{\mathsf{E}} = I_{\mathsf{E}}]$. Let us briefly justify that these are really probability distributions, that is, that these formulae are all non-negative. By construction of $\mathcal{I}_{\mathsf{V}}$ and $\mathcal{J}$, each interval of $\mathcal{I}_{\mathsf{V}}$ is in at most $\frac{\ell}{m}$ intervals of $\mathcal{J}$, so that the expression in (16) is nonnegative. Similarly, by construction $\sum_{J \in \mathcal{J}} \mathfrak{Q}(J, \min(I_{\mathsf{E}}))$ is at most $\frac{1}{m}$ for each $I_{\mathsf{E}} \in \mathcal{I}_{\mathsf{E}}$, so that (17) is nonnegative. Note that for 'most' I and I_{E}, actually (16) and (17) evaluate to zero. Now, we have

$$\mathbf{P}[\mathrm{Cor}_{\mathsf{V}} = *] = 1 - \frac{|\mathcal{I}_{\mathsf{V}}| - \frac{m}{\ell}|\{(I,J) \in \mathcal{I}_{\mathsf{V}} \times \mathcal{J} : I \subseteq J\}|}{|\mathcal{J}|}.$$

Any $J \in \mathcal{J}$ contains exactly $\frac{\ell}{m}$ intervals I from $\mathcal{I}_{\mathsf{V}}$. Hence

$$\mathbf{P}[\mathrm{Cor}_{\mathsf{V}} = *] = \frac{|\mathcal{J}| - |\mathcal{I}_{\mathsf{V}}| + |\mathcal{J}|}{|\mathcal{J}|} \tag{18}$$

which is nonnegative (and in fact very close to 1) by (7) and (6). Finally, we have

$$\mathbf{P}[\mathrm{Cor}_{\mathsf{E}} = *] = 1 - \frac{|\mathcal{I}_{\mathsf{E}}| - m \sum_{I_{\mathsf{E}} \in \mathcal{I}_{\mathsf{E}}} \sum_{J \in \mathcal{J}} \mathfrak{Q}(J, \min(I_{\mathsf{E}}))}{|\mathcal{J}|}.$$

Fix $J \in \mathcal{J}$ and consider the set of differences $\{km : km = |a - a'|, (a, a') \in J \times \overline{J}\}$. This is a set of size $2(\frac{\ell}{m} - 1) + 1$. The largest and the smallest labels in this set can be written as a difference of elements from $J \times \overline{J}$ in exactly m ways each, the second largest and second smallest labels can be written as a difference in $2m$ ways each, and so on. For any $I_{\mathsf{E}} \in \mathcal{I}_{\mathsf{E}}$, $\min(I_{\mathsf{E}})$ is a multiple of m, hence

$$\mathbf{P}[\mathrm{Cor}_{\mathsf{E}} = *] = \frac{|\mathcal{J}| - |\mathcal{I}_{\mathsf{E}}| + m|\mathcal{J}| \sum_{j=-\ell/m}^{\ell/m} \frac{\ell - |jm|}{\ell^2}}{|\mathcal{J}|} = \frac{2|\mathcal{J}| - |\mathcal{I}_{\mathsf{E}}|}{|\mathcal{J}|}, \tag{19}$$

which again by (7) and (6) is very close to 1, and in particular nonnegative.

Given sets $A \subseteq \mathbb{A}$ and $C \subseteq \mathbb{C}$, an interval $I \subseteq \mathbb{A}$ and a label a, we say that $a' \in \mathbb{A}$ is *admissible for a and I with respect to A and C* if $a' \in A \cap I$ and $|a' - a| \in C$. We let the set of admissible vertices for a and I with respect to A and C be $\mathrm{Adm}(a, I; A, C)$. Observe that we have admissible label

$$|X_3 [\![a, I]\!] (A, C)| = |\mathrm{Adm}(a, I; A, C)| . \tag{20}$$

We generate a labelling of $V(T)$ by Algorithm 1. For $t = 1$, at lines 3 and 6 of Algorithm 1 we define

$$\mathrm{Adm}(\psi_0(v_1^+), J(v_1); A_1, C_1) := J(v_1) \quad \text{and} \quad \{|\psi_1(v_1^+) - a|\} := \emptyset . \tag{21}$$

Algorithm 1: Labeling of T.

1 Let $\psi_0 := \emptyset$, let $A_1 := \mathbb{A}$ and let $C_1 := \mathbb{C}$;
2 **foreach** $t = 1, \ldots, n$ **do**
3 Choose $a \in \mathrm{Adm}\big(\psi_{t-1}(v_t^+), J(v_t); A_t, C_t\big)$ uniformly at random ▷ may fail, see (21) for $t = 1$;
4 $\psi_t := \psi_{t-1} \cup \{v_t \mapsto a\}$ ▷ enhance the partial labelling;
5 $A_t^r := A_t \setminus \{a\}$ ▷ remove corresponding vertex label;
6 $C_t^r := C_t \setminus \{|\psi_t(v_t^+) - a|\}$ ▷ remove corresponding edge label, see (21) for $t = 1$;
7 Sample x from $\mathrm{Cor}_{\mathsf{V}}$ ▷ correction on vertex labels;
8 **if** $x = I \subseteq \mathcal{I}_{\mathsf{V}}$ **then**
9 Choose $r_t^v \in I \cap A_t^r$ uniformly at random ▷ may fail;
10 $A_{t+1} := A_t^r \setminus \{r_t^v\}$;
11 **end**
12 **else**
13 $A_{t+1} := A_t^r$, $r_t^v := *$;
14 **end**
15 Sample y from $\mathrm{Cor}_{\mathsf{E}}$ ▷ correction on edge labels;
16 **if** $y = I_{\mathsf{E}} \in \mathcal{I}_{\mathsf{E}}$ **then**
17 Choose $r_t^e \in I_{\mathsf{E}} \cap C_t^r$ uniformly at random ▷ may fail;
18 $C_{t+1} := C_t^r \setminus \{r_t^e\}$;
19 **end**
20 **else**
21 $C_{t+1} := C_t^r$, $r_t^e := *$;
22 **end**
23 **end**
24 **return** ψ_n ;

3.5. **Probabilistic formalities.** To apply Lemma 7, as we will want to do, we need a probability space Ω and a filtration $\mathcal{F}_0, \ldots$. Let Ω be the set of sequences of length $3n$ over the alphabet $\mathbb{A} \cup \mathbb{C} \cup \{*, \mathrm{fail}\}$. We generate a sequence in Ω from a run of Algorithm 1 by recording, for each time $t = 1, \ldots, n$, the vertex label chosen at line 3, the choice of r_t^v , and the choice of r_t^e. In the event that the algorithm fails — which occurs when it requests to sample a uniform element from an empty set in lines 3, 9, or 17 — we record fail at the point when the algorithm fails and in all remaining places of the sequence. We obtain a probability measure on Ω as the probability that running Algorithm 1 generates a given sequence.

When we use Lemma 7, we will have random variables $Y_1, \ldots, Y_n$ tracking sequential contributions to some parameter. Each Y_i is determined by some initial segment of $(\mathbb{A} \cup \mathbb{C} \cup \{*, \mathrm{fail}\})^{3n}$, called a *history*, $\mathscr{H}_i(\omega)$ of $\omega \in \Omega$, we have an estimate for $\sum_i \mathbf{E}[Y_i|\mathscr{H}_{i-1}]$, and the lengths of these initial segments are monotone increasing. Since the lengths of the histories are increasing, they generate in the natural way a filtration on Ω, as required for Lemma 7.

In the rest of the paper, we will not need the details of this construction of Ω, but simply the observation that conditioning on some history is equivalent to conditioning on the behaviour of Algorithm 1 up to a given point, and that Lemma 7 applies to random variables of the above type.

4. Proof of Theorem 9

4.1. **Technical overview.** Before starting the proof, we give a brief overview of the structure. Ultimately, all we need to do is show that Algorithm 1 runs successfully with positive probability. We will show something rather stronger, namely that in fact with high probability, at each time t in the running of Algorithm 1 the pair (A_t, C_t) is δ_t-quasirandom. This is a stronger claim because this quasirandomness in particular asserts that the sets from which labels are chosen at lines 3, 9 and 17 are non-empty.

In turn, to prove δ_t-quasirandomness of (A_t, C_t), we consider separately each $I_{\mathsf{E}} \in \mathcal{I}_{\mathsf{E}}$ for (QUASI1) and each $X \in \mathcal{X}$ for (QUASI2). We describe our approach for a given $X \in \mathcal{X}$; that for $I_{\mathsf{E}} \in \mathcal{I}_{\mathsf{E}}$ is analogous. We can write $\big|X(A_t, C_t)\big|$ as $\big|X(\mathbb{A}, \mathbb{C})\big|$ minus the (random) change caused by the (random) choices of $\psi_1(v_1)$, r_1^v, r_1^e, $\psi_2(v_2)$, and so on up to r_{t-1}^e. Thus what we want to do is estimate the sum of a collection of random variables. These random variables are sequentially dependent, so that we can use Lemma 7 to provide such an estimate. We will see that the probability bounds coming from Lemma 7 are strong enough to simply use the union bound over all choices of X and t, completing the proof.

The difficulty in this programme is that in order to apply Lemma 7 we need estimates for the expected changes at each step, conditioned on the history up to that step. In order to obtain these estimates, we need to know that (A_i, C_i) is δ_i-quasirandom for earlier i. This may seem suspiciously circular: but it is not. To see this, consider the first time t at which quasirandomness fails. This is the first time at which some sum of changes deviates excessively from its expected value. The probability of this event is bounded by Lemma 7 in terms of the sum of the conditional expectations of changes, and those conditional expectations are calculated assuming δ_i-quasirandomness of (A_i, C_i) for some values $i < t$, in other words for times i when, because $i < t$ and t is the first time at which quasirandomness fails, we do have this quasirandomness. Lemma 7 then tells us that the event of quasirandomness first failing at time t is unlikely, and sufficiently unlikely that taking a union bound over t we conclude that quasirandomness failing at any time is unlikely.

Let us now discuss how we obtain these sums of conditional expectations. The *removal term* change caused by r_i^v depends only on the sets A_i^r and C_i^r, and similarly the removal term change caused by r_i^e depends only on A_{i+1} and C_i^r. It is thus quite easy to estimate the sums of conditional expectations of these changes, which we do in Claim 2. It is rather harder to estimate the change caused by $\psi_i(v_i)$, which in addition to the sets A_i and C_i depends also on the choice of v_i^+, which in turn depends on earlier labellings, and so on; ultimately there is some dependence on the entire labelling. Analysing this seems at first hopeless. But in fact there is only significant dependence on v_i^+; we will show that, assuming quasirandomness, any choice of $(v_i^+)^+$ leads to the conditional expectation of change when labelling v_i being approximately a quantity, $p_{X,i}$, which we call the *crude estimate*. This quantity $p_{X,i}$ does not depend at all on the labelling process, and thus we can quite easily estimate the sum over i of the $p_{X,i}$'s. Putting this estimate together with the estimated sum of removal terms, which we do in Claim 2, yields the 'correct' value of $\left|X(\mathbb{A},\mathbb{C})\right| - \left|X(A_t,C_t)\right|$. In other words, it is enough to show that the sum of crude estimates corresponds to the actual changes caused by labelling.

This is still not an easy task. We perform it in two steps. First, we argue in Claim 3 that $p_{X,i}$ is approximately the expectation of change caused by labelling v_i, conditioned on the history up to immediately before labelling v_i^+. We define a *fine estimate* $q_{X,i}$ which corresponds to the expectation of change caused by labelling v_i, conditioned on the history up to immediately after labelling v_i^+. An application of Lemma 7 then tells us that with high probability the sum of the $p_{X,i}$'s is approximately the sum of the $q_{X,i}$'s. Now, $q_{X,i}$ is still not the conditional expectation we would like to find: some vertices may be labelled in between labelling v_i^+ and labelling v_i, and these labellings, together with removal terms in the same interval, cause $q_{X,i}$ and the expectation of change caused by labelling v_i, conditioned on the history up to immediately before labelling v_i, to be different. But provided there are only few such intervening vertices, the difference is small. Our choice of order, using (PRE3), guarantees that for most i there are indeed few such intervening vertices, and we conclude that (deterministically) the sum of the $q_{X,i}$ is close to the sum over i of the expectation of change caused by labelling v_i, conditioned on the history up to immediately before labelling v_i. This last sum is what we need in order to apply Lemma 7 to estimate the sum of the actual changes caused by labelling the v_i, which completes the proof.

In total, then, since the $p_{X,i}$ are quantities independent of the labelling process we do not need to assume anything to estimate their sum in Claim 2. To show that their sum approximates the sum of the $q_{X,i}$'s with high probability, which we do in Claim 3, and to show that the sum of the $q_{X,i}$'s is with high probability close to the sum of the actual changes, we need to assume quasirandomness before the time when we label v_i. As this is before time i, as discussed this assumption is valid.

4.2. **Proof of Theorem 9.** Given $\gamma > 0$, let constants and sets be as defined in Setup 10. Given an n-vertex tree T with $\Delta(T) \leq \frac{\eta n}{\log n}$, Lemma 11 produces an edge set $\mathcal{R}$, an ordering $V(T) = \{v_1, \ldots, v_n\}$ and intervals $J(v_i)$.

In order to apply Lemma 7, we will often need to use the following observation. Since $\Delta(T) \leq \frac{\eta n}{\log n}$, we have

$$\sum_{v \in V(T)} \deg(v)^2 \leq \frac{\log n}{\eta} \cdot \left(\frac{\eta n}{\log n}\right)^2 = \frac{\eta n^2}{\log n}\,. \tag{22}$$

We run Algorithm 1. We say the algorithm *fails* if at any time a step is not possible: in other words, if the sets from which we should choose uniformly $\psi_t(v_t)$, r_t^v, or r_t^e, are empty. Then Theorem 9 holds if with positive probability Algorithm 1 does not fail. We will show that with high probability Algorithm 1 maintains the property that (A_t, C_t) is δ_t-quasirandom for each $1 \leq t \leq n$, and furthermore the following claim holds.

Claim 1. *With probability at least* $1 - n^{-1}$, *at each time* $1 \leq \tau \leq n$, *if Algorithm 1 has not failed before time* τ, *we have*

$$|A_\tau|, |C_\tau| = \tilde{n} - \tau \pm 10\ell\,. \tag{23}$$

Rephrasing the claim, with probability at least $1 - n^{-1}$, at each time $1 \leq \tau \leq n$, it holds that Algorithm 1 has failed before time τ or we have (23). So, we emphasize that the 'if' does not stand for conditioning. We shall have several further 'if–then' statements in later stages of the analysis.

Proof. Observe that, unless the algorithm fails, in each step after the first (in which no edge label is given to any edge) one vertex label and one edge label is used in the labelling, so $|A_\tau| \leq \tilde{n} - \tau + 1$ and $|C_\tau| \leq |\mathbb{C}| - t + 2 = \tilde{n} - \tau + 1$. By (7), (18) and (19), we have

$$\mathbf{P}[\mathrm{Cor}_{\mathsf{V}} \neq *] = \mathbf{P}[\mathrm{Cor}_{\mathsf{E}} \neq *] = \frac{2(\ell - m)}{\tilde{n} - 2(\ell - m)} < \frac{2\ell}{n}\,.$$

By Theorem 5, the probability that in the n samples (i.e. all samples) of either $\mathrm{Cor}_{\mathsf{V}}$ or $\mathrm{Cor}_{\mathsf{E}}$, the result is not $*$ more than 10ℓ times is at least $1 - n^{-2}$. That is, with probability is at least $1 - 2n^{-3} > 1 - n^{-1}$ we have $|A_n|, |C_n| \geq \tilde{n} - n - 10\ell$. Observe that in the case of this good event, we also get $|A_\tau|, |C_\tau| \geq \tilde{n} - n - 10\ell$ for each $\tau \leq n$, no matter at which times the potential non-$*$ samples were sampled.. □

Suppose that (A_t, C_t) is δ_t-quasirandom and (23) holds. By choice of δ_t, for each $a \in \mathbb{A}$ and $I \in \mathcal{I}_{\mathsf{V}}$, using (QUASI2) we have $\left|X_3 [\![a, I]\!] (A_t, C_t)\right| > 0$, so that Algorithm 1 does not fail at line 3. Similarly, since $\left|X_1 [\![I]\!] (A_t, C_t)\right| > 0$, Algorithm 1 does not fail at line 9. It also does not fail at line 17, since (QUASI1) tells us that

$$|I_{\mathsf{E}} \cap C_t^r| = |I_{\mathsf{E}} \cap C_t| \pm 1 = m \cdot \frac{|A_t|}{\tilde{n}} \pm \delta_t m \pm 1 \overset{(23)}{\geq} m \cdot \frac{\tilde{n} - n - 10\ell}{\tilde{n}} - \delta_t m - 1 > 0\,.$$

Thus, in order to prove Theorem 9, it is enough to show that with high probability (A_t, C_t) is δ_t-quasirandom for each $1 \leq t \leq n$. We now embark upon proving this.

Since the vertex labels r_t^v and r_t^e are chosen uniformly at random within intervals of respectively $\mathcal{I}_{\mathsf{V}}$ and $\mathcal{I}_{\mathsf{E}}$, it is quite easy to analyse their effect on (QUASI1) and (QUASI2). It is rather harder to analyse the effect of the edge and vertex labels used at step t, since these are not chosen uniformly. However, the idea one should have in mind is that this choice is 'close to uniform' in a sense we will make precise later, and thus it is useful to write down 'crude estimates' for the effect of the vertex and edge labels used at step t in the labelling which pretends these choices are really uniform. Specifically, the following estimates correspond (more or less) to the expected effect if $\phi_t(v_t)$ were chosen uniformly from the unused vertex labels in $J(v_t)$, if the edge label $\left|\phi_t(v_t)-\phi_t(v_t^+)\right|$ were chosen to be $c\in C_t$ with probability proportional to $\mathfrak{Q}(J(v_t),c)$, independently, and if (QUASI1) and (QUASI2) held with zero error at time t. Of course all these assumptions are false, but we will see that 'on average' they hold, which is enough for our proof.

$p_{I_\mathsf{E},t}$ $r_{I_\mathsf{E},t}$ For $I_\mathsf{E}\in\mathcal{I}_\mathsf{E}$ and $1\le t\le n$, we define

$$\begin{aligned} &\textit{the crude estimate} \qquad p_{I_\mathsf{E},t} := m\mathfrak{Q}(J(v_t),\min(I_\mathsf{E})) \quad \text{and} \\ &\textit{the removal term} \qquad r_{I_\mathsf{E},t} := \mathbb{1}_{r_t^e\in I_\mathsf{E}}\,. \end{aligned} \tag{24}$$

The crude estimate is an idealised version of the expected number of edge labels in I_E that are used at time t (equivalently, to the probability that at time t we use an edge label in the interval I_E). The removal term is the indicator of the event that at time t we remove an edge label from I_E. For $X\in\mathcal{X}$ we define similar terms, with the same intent. Again, the crude estimate $p_{X,t}$ is an estimate for the expected change $\left|X(A_t,C_t)\right|-\left|X(A_t^r,C_t^r)\right|$, and the removal term $r_{X,t}$ is the actual change $\left|X(A_t^r,C_t^r)\right|-\left|X(A_{t+1},C_{t+1})\right|$. For the latter, recall that A_t^r and C_t^r are the available vertex and edge labels, respectively, at time t after removing the vertex label and edge label used in labelling T.

$p_{X,t}$ $r_{X,t}$

$$\begin{aligned} p_{X,t} &:= \tfrac{|X(\mathbb{A},\mathbb{C})|(\tilde n-t)^{\mathrm{free}(X)-1}}{\tilde n^{\mathrm{free}(X)-1}}\Bigg(\sum_{I\in\mathrm{Free}_\mathsf{V}(X)}\tfrac{\mathbb{1}_{I\subseteq J(v_t)}}{\ell}+\sum_{\mathfrak{e}\in\mathrm{Free}_\mathsf{E}(X)}\mathfrak{Q}(J(v_t),\mathrm{Diff}(\mathfrak{e};X))\Bigg), \text{and} \\ r_{X,t} &:= \left|\left\{X'\in X(A_t^r,C_t^r) : r_t^v\in\mathrm{Ch}_\mathsf{V}(X';X) \text{ or } r_t^e\in\mathrm{Ch}_\mathsf{E}(X';X)\right\}\right|\,. \end{aligned} \tag{25}$$

Note that the crude estimates $p_{I_\mathsf{E},t}$ and $p_{X,t}$ are determined before the algorithm starts, and so are their corresponding partial sums $\sum_{i=1}^t p_{I_\mathsf{E},i}$ and $\sum_{i=1}^t p_{X,i}$. By contrast, the partial sums of the removal terms, $\sum_{i=1}^t r_{I_\mathsf{E},i}$ and $\sum_{i=1}^t r_{X,i}$, are sums of random variables which in principle depend upon all of the random choices we make throughout the labelling. However, if we assume that (A_i,C_i) is δ_i-quasirandom for each $1\le i\le t-1$ then we can obtain good bounds on these partial sums which hold with high probability by considering only the choice of the r_i^e and r_i^v.

We estimate the partial sums $\sum_{i=1}^t r_{I_\mathsf{E},i}$ and $\sum_{i=1}^t r_{X,i}$ together. The reason is that eventually we will be able to show that (for example) $\sum_{i=1}^t p_{I_\mathsf{E},i}$ is with high probability a good estimate for the number of edge labels in I_E used in the labelling up to time t, and it follows that $\sum_{i=1}^t(p_{I_\mathsf{E},i}+r_{I_\mathsf{E},i})$ is a good estimate for $|\mathbb{C}\cap I_\mathsf{E}|-|C_t\cap I_\mathsf{E}|$, which is what we want to know in order to verify (QUASI1). Recall that in the introduction we mentioned that our proof can be seen as an application of the Differential Equations Method. This claim is where we (implicitly) verify that the crude estimates we chose actually correspond to solutions to some differential equations: one should understand the right hand sides of (26) and (27) as (what we expect for) the differences $|\mathbb{C}\cap I_{\mathsf{E}}|-|C_t\cap I_{\mathsf{E}}|$ and $|X(\mathbb{A},\mathbb{C})|-|X(A_t,C_t)|$, respectively.

Claim 2. *With probability at least $1-2n^{-1}$, for each $I_{\mathsf{E}}\in\mathcal{I}_{\mathsf{E}}$, each $X\in\mathcal{X}$, and each $1\le t\le n$, if (A_i,C_i) is δ_i-quasirandom for each $1\le i<t$ we have*

$$\sum_{i=1}^{t}(p_{I_{\mathsf{E}},i}+r_{I_{\mathsf{E}},i})=\frac{t}{|\mathcal{J}|}\pm\tfrac14\delta_t m\ ,\ \textit{and}\tag{26}$$

$$\sum_{i=1}^{t}(p_{X,i}+r_{X,i})=|X(\mathbb{A},\mathbb{C})|\frac{\tilde{n}^{\mathrm{free}(X)}-(\tilde{n}-t)^{\mathrm{free}(X)}}{m|\mathcal{J}|\tilde{n}^{\mathrm{free}(X)-1}}\pm\tfrac14\delta_t m\,.\tag{27}$$

Proof. By (PRE4), with $S=[t]$, we have

$$\sum_{i=1}^{t}p_{I_{\mathsf{E}},i}=\big(\tfrac{t}{|\mathcal{J}|}\pm\delta^2 n\big)\sum_{J\in\mathcal{J}}m\mathfrak{Q}(J,\min(I_{\mathsf{E}}))\,,$$

and by (17) and Theorem 5, we have, with probability at least $1-n^{-10}$,

$$\sum_{i=1}^{t}r_{I_{\mathsf{E}},i}=t\frac{1-\sum_{J\in\mathcal{J}}m\mathfrak{Q}(J,\min(I_{\mathsf{E}}))}{|\mathcal{J}|}\pm\delta^2 n\,,$$

so (26) follows with probability at least $1-n^{-10}$. Taking the union bound over all choices of t and $I_{\mathsf{E}}\in\mathcal{I}_{\mathsf{E}}$, we see that with probability at least $1-n^{-2}$, (26) holds for all t and I_{E} as desired.

For (27) we need to be a little more careful: the quantity $p_{X,i}$ depends on i as well as $J(v_i)$, and (with similar effect) the quantity $r_{X,i}$ depends on $|X(A_i^r,C_i^r)|$ and $|A_i^r\cap I|$ and $|C_i^r\cap I_{\mathsf{E}}|$ for each $I\in\mathcal{I}_{\mathsf{V}}$ and $I_{\mathsf{E}}\in\mathcal{I}_{\mathsf{E}}$ as well as the outcomes of $\mathrm{Cor}_{\mathsf{V}}$ and $\mathrm{Cor}_{\mathsf{E}}$. We divide the interval $[t]$ into intervals $S_1,\dots,S_{\lceil t/\delta n\rceil}$, all except possibly the last consisting of δn elements. Note S_j
that by assumption δn is an integer which divides n. The point of doing this is that any time $i\in S_j$, $|X(A_i^r,C_i^r)|$, $|A_i^r\cap I|$ and $|C_i^r\cap I_{\mathsf{E}}|$ are up to a small error constant on any S_j.

For $i\in S_j$, since $\tilde{n}-j\delta n\ge\gamma n$, we have

$$\tilde{n}-i\le\tilde{n}-j\delta n+\delta n\le(\tilde{n}-j\delta n)\big(1+\tfrac{\delta}{\gamma}\big)\ .$$

Thus, for any $1\le s\le 3$,

$$(\tilde{n}-i)^s\le(\tilde{n}-j\delta n)^s\big(1+\tfrac{\delta}{\gamma}\big)^s\le\big(1+8\delta\gamma^{-1}\big)(\tilde{n}-j\delta n)^s\,,$$

and hence $(\tilde{n}-i)^s = \big(1\pm 8\delta\gamma^{-1}\big)(\tilde{n}-j\delta n)^s$. Using this and (PRE4), for each $X\in\mathcal{X}$ we have

$$\sum_{i\in S_j} p_{X,i} = \big(1\pm \tfrac{8\delta}{\gamma}\big)\frac{|X(\mathbb{A},\mathbb{C})|(\tilde{n}-j\delta n)^{\mathrm{free}(X)-1}}{\tilde{n}^{\mathrm{free}(X)-1}}\sum_{I\in \mathrm{Free}_{\mathsf{V}}(X)}\big(\tfrac{|S_j|}{|\mathcal{J}|}\pm\delta^2 n\big)\frac{|\{J\in\mathcal{J}: I\subseteq J\}|}{\ell}$$
$$+\big(1\pm \tfrac{8\delta}{\gamma}\big)\frac{|X(\mathbb{A},\mathbb{C})|(\tilde{n}-j\delta n)^{\mathrm{free}(X)-1}}{\tilde{n}^{\mathrm{free}(X)-1}}\sum_{\mathfrak{e}\in \mathrm{Free}_{\mathsf{E}}(X)}\sum_{J\in\mathcal{J}}\big(\tfrac{|S_j|}{|\mathcal{J}|}\pm\delta^2 n\big)\mathfrak{Q}(J,\mathrm{Diff}(\mathfrak{e};X)) \tag{28}$$

$$= \frac{|X(\mathbb{A},\mathbb{C})|(\tilde{n}-j\delta n)^{\mathrm{free}(X)-1}}{\tilde{n}^{\mathrm{free}(X)-1}}\sum_{I\in \mathrm{Free}_{\mathsf{V}}(X)}\frac{|S_j|\cdot|\{J\in\mathcal{J}: I\subseteq J\}|}{|\mathcal{J}|\ell}$$
$$+\frac{|X(\mathbb{A},\mathbb{C})|(\tilde{n}-j\delta n)^{\mathrm{free}(X)-1}}{\tilde{n}^{\mathrm{free}(X)-1}}\sum_{\mathfrak{e}\in \mathrm{Free}_{\mathsf{E}}(X)}\sum_{J\in\mathcal{J}}\tfrac{|S_j|}{|\mathcal{J}|}\mathfrak{Q}(J,\mathrm{Diff}(\mathfrak{e};X))\pm 100\delta^2 n\gamma^{-1}\,. \tag{29}$$

Let us hint where the value of the final error term $\pm 100\delta^2 n\gamma^{-1}$ in (29) comes from. There are two error terms in (28). To bound the error introduced by the term $\pm\frac{8\delta}{\gamma}$, we use that

$$\frac{|X(\mathbb{A},\mathbb{C})|(\tilde{n}-j\delta n)^{\mathrm{free}(X)-1}}{\tilde{n}^{\mathrm{free}(X)-1}}\overset{(10)}{\le} m\,,$$
$$\sum_{I\in \mathrm{Free}_{\mathsf{V}}(X)}\big(\tfrac{|S_j|}{|\mathcal{J}|}\pm\delta^2 n\big)\cdot\frac{|\{J\in\mathcal{J}: I\subseteq J\}|}{\ell}\le 2\big(\tfrac{\delta n}{2\frac{n}{m}}+\delta^2 n\big)\cdot\frac{\delta_0^{-2}}{\ell}\le 5\delta\,,\text{ and}$$
$$\sum_{\mathfrak{e}\in \mathrm{Free}_{\mathsf{E}}(X)}\sum_{J\in\mathcal{J}}\big(\tfrac{|S_j|}{|\mathcal{J}|}\pm\delta^2 n\big)\cdot\mathfrak{Q}(J,\mathrm{Diff}(\mathfrak{e};X))\overset{(14)}{\le} 2\cdot {}^{2}\!/\delta_0^2\cdot\big(\tfrac{\delta n}{2\frac{n}{m}}+\delta^2 n\big)\cdot\frac{1}{\ell}\le 5\delta\,.$$

The error coming from the term $\pm\delta^2 n$ can be bounded similarly.

We now estimate $\sum_{i\in S_j} r_{X,i}$. We will use Lemma 7 to do this. To that end, for each $i=0,\ldots,n-1$, let $\mathscr{H}_i$ be the history up to and including the choice of $\psi_{i+1}(v_{i+1})$, and let $\mathscr{H}_n$ be the complete history. Hence, the difference between $\mathscr{H}_{n-1}$ and $\mathscr{H}_n$ is only in the information about the choice of r_n^v and r_n^e. Let $\mathcal{E}_t$ be the intersection of the event that (A_i,C_i) is δ_i-quasirandom for each $1\le i<t$ and the good event of Claim 1.[5] Suppose now that $\mathcal{E}_t$ occurs. That is in the calculations below, we shall work with an arbitrary conditional subspace $\mathscr{H}_i$, for some $i<t$, but only with such that $\mathscr{H}_i\cap\mathcal{E}_t$ has positive probability.

Since (A_i^r,C_i^r) differs by one vertex and one edge label[6] from (A_i,C_i), and for any given $X\in\mathcal{X}$ these two labels meet at most three $X'\in X(A_i,C_i)$, we have $\big|X(A_i^r,C_i^r)\big|=\big|X(A_i,C_i)\big|\pm 3$. Furthermore, (A_{i+1},C_i^r) differs by at most one vertex label from (A_i^r,C_i^r), and this vertex label meets at most one $X'\in X(A_i^r,C_i^r)$. Thus, using (QUASI2) to estimate $\big|X(A_i,C_i)\big|$, for

[5]We emphasize that the "good event of Claim 1" involves conditions even on times $\tau>t$.

[6]with the only exception $i=1$ when we have $C_i^r=C_i$

all $i \in S_j$ we have

$$\begin{aligned} \big|X(A_{i+1}, C_i^r)\big|, \big|X(A_i^r, C_i^r)\big| &= \frac{|X(\mathbb{A},\mathbb{C}) \cdot |A_i|^{\mathrm{free}(X)}}{\tilde{n}^{\mathrm{free}(X)}} \pm \delta_i m + \pm 4 \\ &\overset{(23)}{=} \frac{|X(\mathbb{A},\mathbb{C})|(\tilde{n} - i \pm 10\ell)^{\mathrm{free}(X)}}{\tilde{n}^{\mathrm{free}(X)}} \pm \delta_i m \pm 4 \\ &= \frac{|X(\mathbb{A},\mathbb{C})|(\tilde{n} - j\delta n)^{\mathrm{free}(X)}}{\tilde{n}^{\mathrm{free}(X)}} \pm 2\delta_i m \,. \end{aligned} \tag{30}$$

Given $I \in \mathcal{I}_{\mathsf{V}}$, by (QUASI2) with the structure $X_1 [\![I]\!]$, we have

$$|A_i^r \cap I| = m \cdot \frac{\tilde{n} - i \pm 10\ell}{\tilde{n}} \pm \delta_i m = m \cdot \frac{\tilde{n} - j\delta n}{\tilde{n}} \pm 2\delta_i m \,. \tag{31}$$

Using this together with (30), since r_i^v is chosen uniformly in $A_i^r \cap I$ for an interval $I \in \mathcal{I}_{\mathsf{V}}$ drawn from $\mathrm{Cor}_{\mathsf{V}}$, we have

$$\begin{aligned} &\mathbf{E}\Big[\big|\big\{X' \in X(A_i^r, C_i^r) : r_i^v \in \mathrm{Ch}_{\mathsf{V}}(X'; X)\big\}\big| \Big| \mathscr{H}_{i-1}\Big] \\ \boxed{\text{by (30) and (31)}} \quad &= \sum_{I \in \mathrm{Free}_{\mathsf{V}}(X)} \mathbf{P}[\mathrm{Cor}_{\mathsf{V}} = I] \frac{\frac{|X(\mathbb{A},\mathbb{C})|(\tilde{n}-j\delta n)^{\mathrm{free}(X)}}{\tilde{n}^{\mathrm{free}(X)}} \pm 2\delta_i m}{m\frac{\tilde{n}-j\delta n}{\tilde{n}} \pm 2\delta_i m} \\ \boxed{\text{by (16)}} \quad &= \Big(\sum_{I \in \mathrm{Free}_{\mathsf{V}}(X)} \frac{1 - \frac{m}{\ell}|\{J \in \mathcal{J} : I \subseteq J\}|}{|\mathcal{J}|}\Big) \cdot \frac{|X(\mathbb{A},\mathbb{C})|(\tilde{n}-j\delta n)^{\mathrm{free}(X)-1}}{m\tilde{n}^{\mathrm{free}(X)-1}} \pm \frac{10\delta_i}{\gamma|\mathcal{J}|} \,. \end{aligned}$$

Observe that, since the free vertex labels of $X \in \mathcal{X}$ are distinct members of $\mathcal{I}_{\mathsf{V}}$, they are disjoint and hence

$$\text{any given vertex label is in } \mathrm{Ch}_{\mathsf{V}}(X'; X) \text{ for at most one } X' \in X(\mathbb{A},\mathbb{C}). \tag{32}$$

Let us now fix an interval S_j. We apply Lemma 7, with $|S_j|$ many random variables $\big|\big\{X' \in X(A_i^r, C_i^r) : r_i^v \in \mathrm{Ch}_{\mathsf{V}}(X'; X)\big\}\big|$ for $i \in S_j$ and with the event $\mathcal{E}_t$. Observe that these random variables are upper-bounded by 1 by (32). Then Lemma 7 states that if $\mathcal{E}_t$ occurs, then with probability at least $1 - \exp\big(-\frac{2\delta^4 n^2}{n}\big) \geq 1 - n^{-10}$, we have

$$\begin{aligned} &\sum_{i \in S_j} \big|\big\{X' \in X(A_i^r, C_i^r) : r_i^v \in \mathrm{Ch}_{\mathsf{V}}(X'; X)\big\}\big| \\ =&|S_j|\Big(\sum_{I \in \mathrm{Free}_{\mathsf{V}}(X)} \frac{1 - \frac{m}{\ell}|\{J \in \mathcal{J} : I \subseteq J\}|}{|\mathcal{J}|}\Big) \cdot \frac{|X(\mathbb{A},\mathbb{C})|(\tilde{n}-j\delta n)^{\mathrm{free}(X)-1}}{m\tilde{n}^{\mathrm{free}(X)-1}} \pm \frac{10\delta n \delta_{j\delta n}}{\gamma|\mathcal{J}|} \pm \delta^2 n \\ =&|S_j|\Big(\sum_{I \in \mathrm{Free}_{\mathsf{V}}(X)} \frac{1 - \frac{m}{\ell}|\{J \in \mathcal{J} : I \subseteq J\}|}{|\mathcal{J}|}\Big) \cdot \frac{|X(\mathbb{A},\mathbb{C})|(\tilde{n}-j\delta n)^{\mathrm{free}(X)-1}}{m\tilde{n}^{\mathrm{free}(X)-1}} \pm \frac{20\delta n \delta_{j\delta n}}{\gamma|\mathcal{J}|} \,. \end{aligned} \tag{33}$$

We now turn to estimating the effects of the r_i^e. Let us fix an arbitrary history $\mathscr{H}_{i-1} \subseteq \mathcal{E}_t$ which leads to a given set C_i^r. By (QUASI1), for each $I_{\mathsf{E}} \in \mathcal{I}_{\mathsf{E}}$ we have $|C_i^r \cap I_{\mathsf{E}}| = m\frac{\tilde{n}-j\delta n}{\tilde{n}} \pm 2\delta_i m$. Given any set $L \subseteq C_i^r$ of edge labels with $\max(L) - \min(L) \leq m$, the set L is contained in two consecutive intervals of $\mathcal{I}_{\mathsf{E}}$. Let these be $I_{\mathsf{E}}^{(1)}$ and $I_{\mathsf{E}}^{(2)}$, and let L_1 and L_2 be the corresponding subsets of L. Suppose that $c \in I_{\mathsf{E}}^{(1)} \cup I_{\mathsf{E}}^{(2)}$. Then (14) tells us that

$\mathfrak{Q}(J, \min(I^{(1)}_{\mathsf{E}})) = \mathfrak{Q}(J, c) = \mathfrak{Q}(J, \min(I^{(2)}_{\mathsf{E}})) = 0$ for all but at most $2\frac{\ell}{m}$ choices of $J \in \mathcal{J}$. By (15) the three quantities never differ by more than $\frac{2m}{\ell^2}$. Thus we have

$$\sum_{J\in\mathcal{J}} m\mathfrak{Q}(J, \min(I^{(g)}_{\mathsf{E}})) = \sum_{J\in\mathcal{J}} m\mathfrak{Q}(J, c) \pm 8\tfrac{m}{\ell}$$

for each $g = 1, 2$. Using this, we have for an arbitrary history $\mathscr{H}_{i-1} \subseteq \mathcal{E}_t$ which leads to the given set C^r_i, that

$$\begin{aligned}
\mathbf{P}[r^e_i \in L|\mathscr{H}_{i-1}] &= \frac{1 - \sum_{J\in\mathcal{J}} m\mathfrak{Q}(J, \min(I^{(1)}_{\mathsf{E}}))}{|\mathcal{J}|} \cdot \frac{|L_1|}{m\frac{\tilde{n}-j\delta n}{\tilde{n}} \pm 2\delta_i m} \\
&\quad + \frac{1 - \sum_{J\in\mathcal{J}} m\mathfrak{Q}(J, \min(I^{(2)}_{\mathsf{E}}))}{|\mathcal{J}|} \cdot \frac{|L_2|}{m\frac{\tilde{n}-j\delta n}{\tilde{n}} \pm 2\delta_i m} \\
&= \frac{1 \pm 8\frac{m}{\ell} - \sum_{J\in\mathcal{J}} m\mathfrak{Q}(J, c)}{|\mathcal{J}|} \cdot \frac{|L|}{m\frac{\tilde{n}-j\delta n}{\tilde{n}} \pm 2\delta_i m} \\
&= \frac{\big(1 - \sum_{J\in\mathcal{J}} m\mathfrak{Q}(J, c)\big)|L|}{|\mathcal{J}|m\frac{\tilde{n}-j\delta n}{\tilde{n}}}\big(1 \pm 10\gamma^{-1}\delta_i\big) \pm \frac{16|L|}{\ell|\mathcal{J}|\gamma},
\end{aligned}$$

where $c \in I^{(1)}_{\mathsf{E}} \cup I^{(2)}_{\mathsf{E}}$ is arbitrary. Recall that $A_{i+1} = A^r_i \setminus \{r^v_i\}$. Using the above calculation and (30), we obtain that when $i \in S_j$,

$$\begin{aligned}
&\mathbf{E}\Big[\big|\big\{X' \in X(A_{i+1}, C^r_i) : r^e_i \in \mathrm{Ch}_{\mathsf{E}}(X'; X)\big\}\big|\Big|\mathscr{H}_i, r^v_i\Big] \\
&= \Big(\tfrac{|X(\mathbb{A},\mathbb{C})|(\tilde{n}-j\delta n)^{\mathrm{free}(X)}}{\tilde{n}^{\mathrm{free}(X)}} \pm 2\delta_i m\Big) \sum_{\mathfrak{e}\in\mathrm{Free}_{\mathsf{E}}(X)} \Big(\tfrac{(1-\sum_{J\in\mathcal{J}} m\mathfrak{Q}(J,\mathrm{Diff}(\mathfrak{e};X))}{|\mathcal{J}|m\frac{\tilde{n}-j\delta n}{\tilde{n}}}\big(1 \pm 10\gamma^{-1}\delta_i\big) \pm \tfrac{16}{\ell|\mathcal{J}|\gamma}\Big) \\
&= \tfrac{|X(\mathbb{A},\mathbb{C})|(\tilde{n}-j\delta n)^{\mathrm{free}(X)}}{\tilde{n}^{\mathrm{free}(X)}} \sum_{\mathfrak{e}\in\mathrm{Free}_{\mathsf{E}}(X)} \tfrac{(1-\sum_{J\in\mathcal{J}} m\mathfrak{Q}(J,\mathrm{Diff}(\mathfrak{e};X))}{|\mathcal{J}|m\frac{\tilde{n}-j\delta n}{\tilde{n}}} \pm \tfrac{40m}{\ell|\mathcal{J}|\gamma} \pm \tfrac{50\delta_i}{\gamma|\mathcal{J}|} \\
&= \tfrac{|X(\mathbb{A},\mathbb{C})|(\tilde{n}-j\delta n)^{\mathrm{free}(X)-1}}{\tilde{n}^{\mathrm{free}(X)-1}} \sum_{\mathfrak{e}\in\mathrm{Free}_{\mathsf{E}}(X)} \tfrac{(1-\sum_{J\in\mathcal{J}} m\mathfrak{Q}(J,\mathrm{Diff}(\mathfrak{e};X))}{|\mathcal{J}|m} \pm \tfrac{60\delta_i}{\gamma|\mathcal{J}|}.
\end{aligned}$$

Now any given edge label is in at most four $X' \in X(A_{i+1}, C^r_i)$, and $\mathscr{H}_i, r^v_i$ is a history, so by Lemma 7, if $\mathcal{E}_t$ occurs then with probability at least $1 - \exp\big(-\frac{2\delta^4 n^2}{16n}\big) \geq 1 - n^{-10}$ we have

$$\begin{aligned}
&\sum_{i\in S_j} \big|\big\{X' \in X(A^r_i, C^r_i) : r^e_i \in \mathrm{Ch}_{\mathsf{E}}(X'; X)\big\}\big| \\
&= |S_j|\tfrac{|X(\mathbb{A},\mathbb{C})|(\tilde{n}-j\delta n)^{\mathrm{free}(X)-1}}{\tilde{n}^{\mathrm{free}(X)-1}}\Big(\sum_{\mathfrak{e}\in\mathrm{Free}_{\mathsf{E}}(X)} \tfrac{(1-\sum_{J\in\mathcal{J}} m\mathfrak{Q}(J,\mathrm{Diff}(\mathfrak{e};X)))}{|\mathcal{J}|m}\Big) \pm \tfrac{60\delta n\delta_{j\delta n}}{\gamma|\mathcal{J}|} \pm \delta^2 n \qquad (34) \\
&= |S_j|\tfrac{|X(\mathbb{A},\mathbb{C})|(\tilde{n}-j\delta n)^{\mathrm{free}(X)-1}}{\tilde{n}^{\mathrm{free}(X)-1}}\Big(\sum_{\mathfrak{e}\in\mathrm{Free}_{\mathsf{E}}(X)} \tfrac{(1-\sum_{J\in\mathcal{J}} m\mathfrak{Q}(J,\mathrm{Diff}(\mathfrak{e};X)))}{|\mathcal{J}|m}\Big) \pm \tfrac{70\delta n\delta_{j\delta n}}{\gamma|\mathcal{J}|}.
\end{aligned}$$

Putting together (29), (33) and (34), with probability at least $1-n^{-9}$ we have that if $\mathcal{E}_t$ occurs then

$$\sum_{i\in S_j}(p_{X,i}+r_{X,i}) = \frac{|X(\mathbb{A},\mathbb{C})|(\tilde{n}-j\delta n)^{\mathrm{free}(X)-1}}{\tilde{n}^{\mathrm{free}(X)-1}}|S_j|\cdot\frac{\mathrm{free}(X)}{m|\mathcal{J}|}\pm\frac{200\delta n\delta_{j\delta n}}{\gamma|\mathcal{J}|}\,. \tag{35}$$

By the union bound over all $X\in\mathcal{X}$, all $1\le t\le n$ and all j, we see that with probability at least $1-n^{-2}$, if $\mathcal{E}_t$ occurs then the equation (35) holds for all $X\in\mathcal{X}$, all times t and all sets S_j. Now one part of $\mathcal{E}_t$ is the assumption that (A_i,C_i) is δ_i-quasirandom for each $1\le i<t$, and the other part is the good event of Claim 1. The latter event occurs with probability at least $1-n^{-1}$ by Claim 1, so that with probability at least $1-2n^{-1}$ the following holds. Whenever t is such that (A_i,C_i) is δ_i-quasirandom for $1\le i<t$, we have (35) for each $X\in\mathcal{X}$ and j.

Suppose now that, for some t, we have (35) for each $X\in\mathcal{X}$ and j. To complete the proof of the claim, we need to show that putting together these partial sums on short intervals, we obtain the desired (27). Here we are implicitly verifying that we have a solution to a certain first-order differential equation (which we do not write down as we do not need to know it), and consequently an integral naturally appears.

We have

$$\int_{x=0}^{t}(\tilde{n}-x-\delta n)^{\mathrm{free}(X)-1}dx\le\sum_{j=1}^{\lceil\frac{t}{\delta n}\rceil}(\tilde{n}-j\delta n)^{\mathrm{free}(X)-1}|S_j|\le\int_{x=0}^{t}(\tilde{n}-x)^{\mathrm{free}(X)-1}dx\,.$$

Plugging this into (35), we get

$$\begin{aligned}\sum_{i=1}^{t}(p_{X,i}+r_{X,i}) &= \int_{x=0}^{t}\frac{|X(\mathbb{A},\mathbb{C})|(\tilde{n}-x)^{\mathrm{free}(X)-1}\mathrm{free}(X)}{\tilde{n}^{\mathrm{free}(X)-1}m|\mathcal{J}|}dx\pm\frac{\mathrm{free}(X)\delta n}{|\mathcal{J}|}\pm\sum_{j=1}^{\lceil\frac{t}{\delta n}\rceil}\frac{200\delta n\delta_{j\delta n}}{\gamma|\mathcal{J}|}\\ &= |X(\mathbb{A},\mathbb{C})|\frac{\tilde{n}^{\mathrm{free}(X)}-(\tilde{n}-t)^{\mathrm{free}(X)}}{\tilde{n}^{\mathrm{free}(X)-1}m|\mathcal{J}|}\pm\tfrac{1}{4}\delta_t m\,,\end{aligned}$$

as desired, where the final line follows from the choice of δ, from (8), and since $\frac{n}{|\mathcal{J}|}<2m$. □

We would now like to argue that $\sum_{i=1}^{t}p_{I_{\mathsf{E}},i}$ is a good estimate for the number of edge labels in I_{E} used in the labelling process up to time t. However, we are not able to do this in one step. Instead, we define a *fine estimate* $q_{I_{\mathsf{E}},i}$, which plays the same rôle as $p_{I_{\mathsf{E}},i}$ except that we condition on the behaviour of Algorithm 1 up to and including the time h at which we label $v_h=v_i^+$. We will see that at least $\sum_{i=1}^{t}p_{I_{\mathsf{E}},i}$ is a good estimate for $\sum_{i=1}^{t}q_{I_{\mathsf{E}},i}$. We define similarly a fine estimate $q_{X,i}$ corresponding to $p_{X,i}$. We write down formulae valid for $i\ge2$, when $v_h=v_i^+$ exists. $q_{I_{\mathsf{E}},i}$

$q_{X,i}$

$$q_{I_{\mathsf{E}},i} := \frac{\left|\left\{a \in \mathrm{Adm}(\psi_h(v_h), J(v_i); A_h, C_h) : |a - \psi_h(v_h)| \in I_{\mathsf{E}}\right\}\right| \tilde{n}^2}{|A_h|^2 \ell}$$
$$q_{X,i} := \sum_{a \in \mathrm{Adm}(\psi_h(v_h), J(v_i); A_h, C_h)} \frac{\left|\left\{X' \in X(A_h, C_h) : a \in \mathrm{Ch}_{\mathsf{V}}(X'; X) \text{ or } |a - \psi_h(v_h)| \in \mathrm{Ch}_{\mathsf{E}}(X'; X)\right\}\right|}{|A_h|^2 \ell \tilde{n}^{-2}} \tag{36}$$

Additionally, we define $q_{I_{\mathsf{E}},1} = q_{X,1} = 0$.

Observe that $\frac{|A_h||C_h|\ell}{\tilde{n}|\mathbb{C}|}$ is, by (QUASI2) with the structure $X_3 \llbracket \psi_h(v_h), J(v_i) \rrbracket$, a good estimate for the size of the set of vertex labels from which we will label v_i, ignoring any changes that may occur between time j and the time i when we label v_i. Thus $q_{I_{\mathsf{E}},i}$ is a good estimate for the expectation of labelling v_i in such a way as to use an edge label in I_{E}, if we ignore any changes that might occur between times j and i. We will see that it is usually reasonable to ignore such changes. Similarly, $q_{X,i}$ is a good estimate for the expected number of structures following the pattern X whose chosen labels contain either the vertex or edge label used at time i, ignoring any changes between times j and i, and estimating the size of the set of vertex labels from which we label v_i by $\frac{|A_h||C_h|\ell}{\tilde{n}|\mathbb{C}|}$.

We now show that the partial sums of the crude estimates are, with high probability, good estimates for the partial sums of the fine estimates. There are two parts to this. First, we will argue that if $v_i v_i^+ \notin \mathcal{R}$, then letting $v_h = v_i^+$, we have $\mathbf{E}[q_{I_{\mathsf{E}},i} | \psi_{h-1}] \approx p_{I_{\mathsf{E}},i}$, and similarly for the $q_{X,i}$. In other words, $p_{I_{\mathsf{E}},i}$ is a good estimate for the expectation of $q_{I_{\mathsf{E}},i}$ conditioned on the labelling history up to the time immediately before labelling v_i^+, whatever that history might be (as long as it maintains quasirandomness). This is the main combinatorial work in our proof. Second, we observe that the effect of the remaining terms where $v_i v_i^+ \in \mathcal{R}$ is small, simply because $\mathcal{R}$ is small, and apply Lemma 7 to argue that the sum of conditional expectations is with high probability close to the partial sum $\sum_{i=1}^{t} q_{I_{\mathsf{E}},i}$, and similarly for the $q_{X,i}$.

Claim 3. *With probability at least $1 - 4n^{-1}$ the following holds. For each $I_{\mathsf{E}} \in \mathcal{I}_{\mathsf{E}}$, each $X \in \mathcal{X}$, each $1 \le t \le n$, and each $1 \le k \le \lceil \frac{t}{\delta n} \rceil$, if (A_i, C_i) is δ_i-quasirandom for each $1 \le i < t$, we have*

$$\sum_{i=(k-1)\delta n+1}^{\max(t,k\delta n)} q_{I_{\mathsf{E}},i} = \sum_{i=(k-1)\delta n+1}^{\max(t,k\delta n)} p_{I_{\mathsf{E}},i} \pm \frac{2000\delta_{k\delta n}\delta m}{\gamma^4}\ , \text{ and} \tag{37}$$
$$\sum_{i=(k-1)\delta n+1}^{\max(t,k\delta n)} q_{X,i} = \sum_{i=(k-1)\delta n+1}^{\max(t,k\delta n)} p_{X,i} \pm \tfrac{10^6 \delta m \delta_{k\delta n}}{\gamma^7}\ . \tag{38}$$

Proof. For this proof, for each $0 \le i \le n$, let $\mathscr{H}_i$ denote the history up to, but not including, the labelling of v_{i+1}. Thus $\mathscr{H}_0$ is the empty history.

We begin with (37). Given I_{E} and $1 \le t \le n$, and $1 \le k \le \lceil \frac{t}{\delta n} \rceil$, we define random variables

Y_h for each $1 \le h < t$ by

$$Y_h = \sum_{h<i\le t} q_{I_{\mathsf{E}},i} \cdot \mathbb{1}_{v_iv_h\in T} \cdot \mathbb{1}_{(k-1)\delta n+1\le i\le \max(t,k\delta n)} \,.$$

We have $q_{I_{\mathsf{E}},1}=0$, and hence

$$\sum_{h=1}^{t-1} Y_h = \sum_{i=(k-1)\delta n+1}^{\max(t,k\delta n)} q_{I_{\mathsf{E}},i} \,. \tag{39}$$

We will apply Lemma 7 to estimate the sum on the left-hand side. As before, we let $\mathcal{E}_t$ be the event that (A_i, C_i) is δ_i-quasirandom for each $1 \le i < t$, and the good event of Claim 1 holds. In particular, for each h we have $|A_h|, |C_h| \ge \frac{\gamma}{2}\tilde{n}$.

We need to show that, assuming $\mathcal{E}_t$, we can give good bounds on $\sum_{h=1}^{t-1} \mathbf{E}[Y_h|\mathscr{H}_{h-1}]$. In turn, to obtain such bounds it is enough to show that $\mathbf{E}[q_{I_{\mathsf{E}},i}|\mathscr{H}_{h-1}] \approx p_{I_{\mathsf{E}},i}$ for each $1 < h < i \le t$ with $v_iv_h \in E(T)\setminus\mathcal{R}$. The terms with $h=1$ or $v_iv_h \in \mathcal{R}$ contribute at most $\frac{\eta n}{\log n} + \varepsilon n$ to the sum by assumption on $\Delta(T)$ and by (PRE1), which is small enough to ignore.

Suppose we have $1 < h < i \le t$, with $v_h = v_i^+$ and $v_iv_h \in E(T)\setminus\mathcal{R}$. We say that (a,a') is an *admissible pair* if $a \in J(v_h)\cap A_h$ and $a' \in J(v_i)\cap A_h$, and $\big|\psi_{h-1}(v_h^+)-a\big|, |a'-a| \in C_h$ are distinct. Note that since $J(v_h)=\overline{J(v_i)}$ by (PRE5), a and a' are distinct. It follows that $\psi_{h-1}\cup\{v_h\mapsto a, v_i\mapsto a'\}$ is a graceful labelling of $T[v_1,\dots,v_h,v_i]$. admissible pair

Note that

$$\mathbf{E}[q_{I_{\mathsf{E}},i}|\mathscr{H}_{h-1}] = \mathbf{E}_{a\sim\mathsf{UNIFORM}(\mathrm{Adm}(\psi_{h-1}(v_h^+),J(v_h);A_h,C_h))}\left[\mathbf{E}\left[q_{I_{\mathsf{E}},i}|\mathscr{H}_{h-1},\psi_h(v_h)=a\right]\right],$$

because we choose $\psi_h(v_h)$ uniformly. Thus, by definition of $q_{I_{\mathsf{E}},i}$, we have

$$\begin{aligned}
\tfrac{|A_h|^2\ell}{\tilde{n}^2}\mathbf{E}[q_{I_{\mathsf{E}},i}|\mathscr{H}_{h-1}] &= \frac{\sum_{(a,a')\text{ admissible}} \mathbb{1}_{|a-a'|\in I_{\mathsf{E}}\cap C_h}}{\big|\mathrm{Adm}(\psi_{h-1}(v_h^+), J(v_h); A_h, C_h)\big|} \\
&\overset{(20)}{=} \frac{\sum_{(a,a')\text{ admissible}} \mathbb{1}_{|a-a'|\in I_{\mathsf{E}}\cap C_h}}{\big|X_3\left[\!\left[\psi_{h-1}(v_h^+), J(v_h)\right]\!\right](A_h,C_h)\big|} \\
&= \frac{\sum_{c\in I_{\mathsf{E}}\cap C_h}\Big|X_2\left[\!\left[\psi_{h-1}(v_h^+), \overline{J(v_i)}, c, J(v_i)\right]\!\right](A_h,C_h)\Big|}{\big|X_3\left[\!\left[\psi_{h-1}(v_h^+), J(v_h)\right]\!\right](A_h,C_h)\big|}\,.
\end{aligned}$$

Therefore,

$$\mathbf{E}[q_{I_{\mathsf{E}},i}|\mathscr{H}_{h-1}] = \frac{\tilde{n}^2\sum_{c\in I_{\mathsf{E}}\cap C_h}\left|X_2\left[\!\left[\psi_{h-1}(v_h^+), \overline{J(v_i)}, c, J(v_i)\right]\!\right](A_h,C_h)\right|}{|A_h|^2\ell\big|X_3\left[\!\left[\psi_{h-1}(v_h^+), J(v_h)\right]\!\right](A_h,C_h)\big|}\,.$$

Because we assume $\mathcal{E}_t$ and $h<t$, so (A_h,C_h) is δ_h-quasirandom and we can use Lemma 14 to estimate both the X_2-term in the numerator and the X_3-term in the denumerator. We

obtain

$$\begin{aligned}
\mathbf{E}[q_{I_{\mathsf{E}},i}|\mathscr{H}_{h-1}] &\overset{(11),(12)}{=} \frac{\tilde{n}^2 \sum_{c\in I_{\mathsf{E}}\cap C_h}\Big(\ell^2\mathfrak{Q}(J(v_i),c)|A_h|^3\tilde{n}^{-3}\pm 3\delta_h\ell\Big)}{|A_h|^2\ell\big(\ell|A_h|^2\tilde{n}^{-2}\pm 2\delta_h\ell\big)} \\
&= \frac{\tilde{n}\sum_{c\in I_{\mathsf{E}}\cap C_h}\mathfrak{Q}(J(v_i),c)}{|A_h|}\pm\frac{50\delta_h m}{\ell\gamma^4} \\
&\overset{(15)}{=} \frac{\tilde{n}|I_{\mathsf{E}}\cap C_h|\mathfrak{Q}(J(v_i),\min(I_{\mathsf{E}}))}{|A_h|}\pm\frac{60\delta_h m}{\ell\gamma^4}\,.
\end{aligned}$$

By (QUASI1) we have $|I_{\mathsf{E}}\cap C_h| = m\frac{|A_h|}{\tilde{n}}\pm\delta_h m$, so that

$$\mathbf{E}[q_{I_{\mathsf{E}},i}|\mathscr{H}_{h-1}] = m\mathfrak{Q}(J(v_i),\min(I_{\mathsf{E}}))\pm\tfrac{70\delta_h m}{\gamma^4\ell} = p_{I_{\mathsf{E}},i}\pm\tfrac{70\delta_i m}{\gamma^4\ell}\,, \tag{40}$$

where we use $\delta_h<\delta_i$ since $h<i$.

For most values of i we actually have a stronger estimate than (40). If $\mathfrak{Q}(J(v_i),c)=0$ for all $c\in I_{\mathsf{E}}$, then $p_{I_{\mathsf{E}},i}=q_{I_{\mathsf{E}},i}=0$ by (24) and (36). That is, in this situation we have

$$\mathbf{E}[q_{I_{\mathsf{E}},i}|\mathscr{H}_{h-1}] = p_{I_{\mathsf{E}},i}\,. \tag{41}$$

For any given $I_{\mathsf{E}}\in\mathcal{I}_{\mathsf{E}}$, by (14) there are at most $2\frac{\ell}{m}$ sets $J\in\mathcal{J}$ such that $\mathfrak{Q}(J,c)\neq 0$ for some $c\in I_{\mathsf{E}}$, so by (PRE4), and choice of δ, for each $1\le k\le\delta^{-1}$, the number of $i\in\{(k-1)\delta n+1,\dots,k\delta n\}$ such that $\mathfrak{Q}(J(v_i),c)\neq 0$ for some $c\in I_{\mathsf{E}}$ is at most $2\frac{\ell}{m}\cdot\frac{2\delta n}{|\mathcal{J}|}$.

$\mathcal{D}_h$ We want to sum up (40) and (41). To this end, for $h=2,\dots,t-1$, set

$$\mathcal{D}_h = \{i\in\mathbb{N}\,:\,(k-1)\delta n+1\le i\le\max(t,k\delta n), v_i^+=v_h, v_iv_h\notin\mathcal{R}\}\,. \tag{42}$$

Observe that $\dot{\bigcup}_{h=2}^{t-1}\mathcal{D}_h=\{i\in\mathbb{N}\,:\,(k-1)\delta n+1\le i\le\max(t,k\delta n), v_i^+\neq v_1, v_iv_i^+\notin\mathcal{R}\}$. Thus, we have

$$\sum_{h=2}^{t-1}\sum_{i\in\mathcal{D}_h}\mathbf{E}[q_{I_{\mathsf{E}},i}|\mathscr{H}_{h-1}] \overset{(40),(41)}{=} \Bigg(\sum_{\substack{i=(k-1)\delta n+1\\ v_i^+\neq v_1\,,\,v_iv_i^+\notin\mathcal{R}}}^{\max(t,k\delta n)} p_{I_{\mathsf{E}},i}\Bigg)\pm\frac{4\ell\delta n}{m|\mathcal{J}|}\cdot\frac{70\delta_{k\delta n}m}{\ell\gamma^4}\,, \tag{43}$$

since $i\le k\delta n$ and thus $\delta_i\le\delta_{k\delta n}$. Taking into account the at most εn values of i with $v_iv_i^+\in\mathcal{R}$ and at most $\frac{\eta n}{\log n}$ values of i with $v_i^+=v_1$, each of which terms contributes an error of at most 1, and using (7), we have

$$\begin{aligned}
\sum_{h=1}^{t-1}\mathbf{E}[Y_h|\mathscr{H}_{h-1}] &= \sum_{h=1}^{t-1}\sum_{i=(k-1)\delta n+1\,,\,v_i^+=v_h}^{\max(t,k\delta n)}\mathbf{E}[q_{I_{\mathsf{E}},i}|\mathscr{H}_{h-1}] \\
&= \Bigg(\sum_{i=(k-1)\delta n+1}^{\max(t,k\delta n)} p_{I_{\mathsf{E}},i}\Bigg)\pm\frac{600\delta_{k\delta n}\delta m}{\gamma^4}\pm\varepsilon n\pm\tfrac{\eta n}{\log n} \\
&= \Bigg(\sum_{i=(k-1)\delta n+1}^{\max(t,k\delta n)} p_{I_{\mathsf{E}},i}\Bigg)\pm\frac{1000\delta_{k\delta n}\delta m}{\gamma^4}\,.
\end{aligned}$$

We are now in a position to apply Lemma 7, with the random variables $(Y_h)_{h=1}^{t-1}$ satisfying $0 \le Y_h \le \deg_T(v_h)$ for each h, and with the event $\mathcal{E}_t$. By (22) we have $\sum_{h=1}^{t-1} \deg_T(v_h)^2 \le \frac{\eta n^2}{\log n}$, so applying Lemma 7 we conclude that with probability at least $1-\exp\big(-\frac{2\delta^2 n^2 \log n}{\eta n^2}\big) > 1-n^{-10}$, if $\mathcal{E}_t$ holds, we have

$$\sum_{h=1}^{t-1} Y_h \overset{(39)}{=} \sum_{i=(k-1)\delta n+1}^{\max(t,k\delta n)} q_{I_{\mathsf{E}},i} = \Big(\sum_{i=(k-1)\delta n+1}^{\max(t,k\delta n)} p_{I_{\mathsf{E}},i}\Big) \pm \frac{1000\delta_{k\delta n}\delta m}{\gamma^4} \pm \delta n$$
$$= \Big(\sum_{i=(k-1)\delta n+1}^{\max(t,k\delta n)} p_{I_{\mathsf{E}},i}\Big) \pm \frac{2000\delta_{k\delta n}\delta m}{\gamma^4}\,.$$

Taking the union bound over $1 \le t \le n$ and $I_{\mathsf{E}} \in \mathcal{I}_{\mathsf{E}}$ and k, we conclude that with probability at least $1-n^{-2}$, if $\mathcal{E}_t$ holds then (37) holds for each $1 \le t \le n$, each $I_{\mathsf{E}} \in \mathcal{I}_{\mathsf{E}}$, and each k. Recall that the good event of Claim 1 holds with probability at least $1-n^{-1}$. Thus the following event holds with probability at least $1-2n^{-1}$. For each $1 \le t \le n$, if (A_i, C_i) is δ_i-quasirandom for each $1 \le i < t$, then (37) holds for each $I_{\mathsf{E}} \in \mathcal{I}_{\mathsf{E}}$ and each k.

The proof that with high probability (38) holds follows the same idea, although the combinatorial manipulations are a little more involved. Now, given $X \in \mathcal{X}$ and $1 \le t \le n$, we define Y_h

$$Y_h = \sum_{h<i\le t} q_{X,i} \cdot \mathbb{1}_{v_i v_h \in T} \cdot \mathbb{1}_{(k-1)\delta n+1 \le i \le \max(t,k\delta n)}\,.$$

and again the critical point is to show that, assuming $\mathcal{E}_t$, for each $1 < h < i \le t$ such that $v_i v_h \in E(T) \setminus \mathcal{R}$, we have $\mathbf{E}[q_{X,i}|\mathscr{H}_{h-1}] \approx p_{X,i}$.

As before, given a choice of $1 < h < i \le t$ with $v_h v_i^+$ such that $v_i v_h \notin \mathcal{R}$, we say that (a, a') is an *admissible pair* if $a \in J(v_h) \cap A_h$ and $a' \in J(v_i) \cap A_h$, and $\big|\psi_{h-1}(v_h^+) - a\big|, |a'-a| \in C_h$ are distinct. admissible pair

By (36), and since $\psi_h(v_h)$ is chosen uniformly from $\mathrm{Adm}\big(\psi_{h-1}(v_h^+), J(v_h); A_h, C_h\big)$, we have

$$\begin{aligned}
&\frac{|A_h|^2 \ell \mathbf{E}[q_{X,i}|\mathscr{H}_{h-1}] \cdot \big|\mathrm{Adm}\big(\psi_{h-1}(v_h^+), J(v_h); A_h, C_h\big)\big|}{\tilde{n}^2} \\
&= \sum_{(a,a')\text{ admissible}} \;\sum_{X' \in X(A_h, C_h)} \mathbb{1}_{a' \in \mathrm{Ch}_{\mathsf{V}}(X';X) \text{ or } |a-a'| \in \mathrm{Ch}_{\mathsf{E}}(X';X)} \qquad (44)\\
&= \sum_{X' \in X(A_h, C_h)} \;\sum_{(a,a')\text{ admissible}} \mathbb{1}_{a' \in \mathrm{Ch}_{\mathsf{V}}(X';X) \text{ or } |a-a'| \in \mathrm{Ch}_{\mathsf{E}}(X';X)} \\
&= \sum_{X' \in X(A_h, C_h)} \;\sum_{(a,a')\text{ admissible}} \mathbb{1}_{a' \in \mathrm{Ch}_{\mathsf{V}}(X';X)} + \mathbb{1}_{|a-a'| \in \mathrm{Ch}_{\mathsf{E}}(X';X)} - \mathbb{1}_{a' \in \mathrm{Ch}_{\mathsf{V}}(X';X)} \mathbb{1}_{|a-a'| \in \mathrm{Ch}_{\mathsf{E}}(X';X)}\,,
\end{aligned}$$

where the last equality holds since vertex labels of X' are by definition pairwise distinct, and edge labels of X' are by definition pairwise distinct.

For a given $X' \in X(A_h, C_h)$, we have

$$\begin{aligned} \sum_{(a,a')\text{ admissible}} \mathbb{1}_{a' \in \mathrm{Ch}_{\mathsf{V}}(X';X)} &= \sum_{a' \in \mathrm{Ch}_{\mathsf{V}}(X';X) \cap J(v_i)} \Big| X_4 \left[\!\left[\psi_{h-1}(v_h^+), a', J(v_h) \right]\!\right] (A_h, C_h) \Big| \\ &\overset{(13)}{=} \sum_{a' \in \mathrm{Ch}_{\mathsf{V}}(X';X) \cap J(v_i)} \big(\tfrac{|A_h|^3 \ell}{\tilde{n}^3} \pm 2\delta_h \ell \big) \\ &= \sum_{I \in \mathrm{Free}_{\mathsf{V}} X} \mathbb{1}_{I \subseteq J(v_i)} \big(\tfrac{|A_h|^3 \ell}{\tilde{n}^3} \pm 2\delta_h \ell \big) \,, \end{aligned} \tag{45}$$

where we can apply Lemma 14 since (A_h, C_h) is by assumption δ_h-quasirandom. Similarly, we have

$$\begin{aligned} \sum_{(a,a')\text{ admissible}} \mathbb{1}_{|a-a'| \in \mathrm{Ch}_{\mathsf{E}}(X';X)} &= \sum_{c \in \mathrm{Ch}_{\mathsf{E}}(X';X)} \Big| X_2 \left[\!\left[\psi_{h-1}(v_h^+), J(v_h), c, J(v_i) \right]\!\right] (A_h, C_h) \Big| \\ &\overset{(11)}{=} \sum_{c \in \mathrm{Ch}_{\mathsf{E}}(X';X)} \big(\tfrac{|A_h|^3 \ell^2 \mathfrak{Q}(J(v_h), c)}{\tilde{n}^3} \pm .\delta_h \ell \big) \\ &\overset{(15)}{=} \sum_{\mathfrak{e} \in \mathrm{Free}_{\mathsf{E}}(X)} \big(\tfrac{|A_h|^3 \ell^2 \mathfrak{Q}(J(v_h), \mathrm{Diff}(\mathfrak{e};X))}{\tilde{n}^3} \pm 4\delta_h \ell \big) \,, \end{aligned} \tag{46}$$

where the final line follows since $\mathrm{Diff}(\mathfrak{e}; X)$ is within m of the edge label chosen for $\mathfrak{e}$ in any X' following the pattern X. Finally, since $\mathrm{Ch}_{\mathsf{E}}(X'; X)$ is a set of size at most 2, for each $a' \in \mathrm{Ch}_{\mathsf{V}}(X'; X)$ there are at most 4 choices of a such that $|a - a'| \in \mathrm{Ch}_{\mathsf{E}}(X'; X)$. Since $\mathrm{Ch}_{\mathsf{V}}(X'; X)$ is a set of size at most 2, in total there are at most 8 pairs (a, a') such that $a' \in \mathrm{Ch}_{\mathsf{V}}(X'; X)$ and $|a - a'| \in \mathrm{Ch}_{\mathsf{E}}(X'; X)$. We therefore have

$$\sum_{(a,a')\text{ admissible}} \mathbb{1}_{a' \in \mathrm{Ch}_{\mathsf{V}}(X';X)} \mathbb{1}_{|a-a'| \in \mathrm{Ch}_{\mathsf{E}}(X';X)} \le 8 \,. \tag{47}$$

Observe that the final expression in each of (45), (46) and (47) is independent of X'. Thus, substituting into (44), we have

$$\begin{aligned}
&\frac{|A_h|^2\ell\mathbf{E}[q_{X,i}|\mathscr{H}_{h-1}]\cdot\big|\mathrm{Adm}\big(\psi_{h-1}(v_h^+),J(v_h);A_h,C_h\big)\big|}{\tilde{n}^2}\\
=&\big|X(A_h,C_h)\big|\sum_{I\in\mathrm{Free}_{\mathsf{V}}X}\mathbb{1}_{I\subseteq J(v_i)}\big(\tfrac{|A_h|^3\ell}{\tilde{n}^3}\pm2\delta_h\ell\big)\\
+&\big|X(A_h,C_h)\big|\sum_{\mathfrak{e}\in\mathrm{Free}_{\mathsf{E}}(X)}\big(\tfrac{|A_h|^3\ell^2\mathfrak{Q}(J(v_h),\mathrm{Diff}(\mathfrak{e};X))}{\tilde{n}^3}\pm4\delta_h\ell\big)\pm8m\\
\overset{(10)}{=}&\frac{|A_h|^3\ell^2\big|X(A_h,C_h)\big|}{\tilde{n}^3}\Big(\sum_{I\in\mathrm{Free}_{\mathsf{V}}(X)}\frac{\mathbb{1}_{I\subseteq J(v_i)}}{\ell}+\sum_{\mathfrak{e}\in\mathrm{Free}_{\mathsf{E}}(X)}\mathfrak{Q}(J(v_i),\mathrm{Diff}(\mathfrak{e};X))\Big)\pm20\delta_h\ell m\\
\overset{(\mathrm{QUASI2})}{=}&\frac{|A_h|^{3+\mathrm{free}(X)}\ell^2\big|X(\mathbb{A},\mathbb{C})\big|}{\tilde{n}^{3+\mathrm{free}(X)}}\Big(\sum_{I\in\mathrm{Free}_{\mathsf{V}}(X)}\frac{\mathbb{1}_{I\subseteq J(v_i)}}{\ell}+\sum_{\mathfrak{e}\in\mathrm{Free}_{\mathsf{E}}(X)}\mathfrak{Q}(J(v_i),\mathrm{Diff}(\mathfrak{e};X))\Big)\pm\frac{40\delta_h\ell m}{\gamma^3}\\
\overset{(25)}{=}&\frac{|A_h|^{3+\mathrm{free}(X)}\ell^2\big|X(\mathbb{A},\mathbb{C})\big|}{\tilde{n}^{3+\mathrm{free}(X)}}\Big(\frac{\tilde{n}^{\mathrm{free}(X)-1}p_{X,i}}{|X(\mathbb{A},\mathbb{C})|(\tilde{n}-i)^{\mathrm{free}(X)-1}}\Big)\pm\frac{40\delta_h\ell m}{\gamma^3}\\
=&\frac{|A_h|^{3+\mathrm{free}(X)}\ell^2p_{X,i}}{\tilde{n}^4(\tilde{n}-h)^{\mathrm{free}(X)-1}}\pm\frac{50\delta_h\ell m}{\gamma^3}\,.
\end{aligned}\tag{48}$$

where for the second equality we use the fact that $\mathfrak{Q}(J(v_i),\mathrm{Diff}(\mathfrak{e};X))=\mathfrak{Q}(J(v_h),\mathrm{Diff}(\mathfrak{e};X))$ since $J(v_i)=\overline{J(v_h)}$ by (PRE5), for the third we use the assumption that (A_h,C_h) is δ_h-quasirandom, and for the last line we use the fact $i=h\pm\frac{\varepsilon n}{\log n}$, which holds by (PRE3) since $v_iv_h\notin\mathcal{R}$.

Now, since (A_h,C_h) is δ_h-quasirandom, by (12) and since $|A_h|\geq\frac{\gamma}{2}\tilde{n}$ by (23), we have

$$\big|\mathrm{Adm}\big(\psi_{h-1}(v_h^+),J(v_h);A_h,C_h\big)\big|=\big(|A_h|^2/\tilde{n}^2\big)\ell\pm2\delta_h\ell=\big(1\pm\tfrac{8\delta_h}{\gamma^2}\big)|A_h|^2\tilde{n}^{-2}\ell\,.\tag{49}$$

We can rewrite (48) as

$$\begin{aligned}
\mathbf{E}[q_{X,i}|\mathscr{H}_{h-1}]&=\frac{\frac{|A_h|^{1+\mathrm{free}(X)}\ell p_{X,i}}{\tilde{n}^2(\tilde{n}-h)^{\mathrm{free}(X)-1}}\pm\frac{50\delta_hm\tilde{n}^2}{\gamma^3|A_h|^2}}{\big|\mathrm{Adm}\big(\psi_{h-1}(v_h^+),J(v_h);A_h,C_h\big)\big|}\\
\boxed{\text{substituting (49) and using that } |A_h|\geq\tfrac{\gamma}{2}\tilde{n}}\quad&=\frac{|A_h|^{\mathrm{free}(X)-1}p_{X,i}}{(\tilde{n}-h)^{\mathrm{free}(X)-1}}\big(1\pm\tfrac{16\delta_h}{\gamma^2}\big)\pm\frac{2000\delta_hm}{\gamma^7\ell}\,.
\end{aligned}$$

By (23), we have $|A_h| = \tilde{n} - h \pm 10\ell$. Furthermore, by (25) we have $p_{X,i} \leq 4m/\ell$. Finally, we have $\mathrm{free}(X) \leq 3$. We thus get

$$\begin{aligned}
\mathbf{E}[q_{X,i}|\mathscr{H}_{h-1}] &= \frac{(\tilde{n}-h\pm 10\ell)^{\mathrm{free}(X)-1}p_{X,i}}{(\tilde{n}-h)^{\mathrm{free}(X)-1}}\big(1\pm\tfrac{16\delta_h}{\gamma^2}\big) \pm \frac{2000\delta_h m}{\gamma^7\ell}\\
&= p_{X,i}\big(1\pm\tfrac{16\delta_h}{\gamma^2}\big)\big(1\pm\tfrac{20\ell}{\gamma\tilde{n}}\big)^2 \pm \frac{2000\delta_h m}{\gamma^7\ell}\\
&= p_{X,i} \pm \frac{100\delta_h m}{\gamma^2\ell} \pm \frac{800m}{\gamma\tilde{n}} \pm \frac{2000\delta_h m}{\gamma^7\ell} = p_{X,i} \pm \tfrac{3000\delta_i m}{\gamma^7\ell}\,.
\end{aligned} \tag{50}$$

As before, for most values of i we obtain a stronger estimate. If $\mathfrak{Q}(J(v_i),c) = 0$ for each $c \in \mathrm{Ch}_{\mathsf{E}}(X';X)$ and each $X' \in X(\mathbb{A},\mathbb{C})$, and $I \cap J(v_i) = \emptyset$ for each $I \in \mathrm{Free}_{\mathsf{V}}(X)$, then $p_{X,i} = q_{X,i} = 0$ by (25) and (36). In this situation, we can write

$$\mathbf{E}[q_{X,i}|\mathscr{H}_{h-1}] = p_{X,i}\,. \tag{51}$$

Since X has at most two free edge labels, for each of which an edge label can be chosen in an interval of length m, there are at most $4\frac{\ell}{m}$ sets $J \in \mathcal{J}$ such that $\mathfrak{Q}(J,c) \neq 0$ for some $c \in \mathrm{Ch}_{\mathsf{E}}(X';X)$ and $X' \in X(\mathbb{A},\mathbb{C})$. Since X has at most two free vertex labels, there are at most $2\frac{\ell}{m}$ sets $J \in \mathcal{J}$ such that $I \subseteq J$ for some $I \in \mathrm{Free}_{\mathsf{V}}(X)$. Putting this together, for all but at most $6\frac{\ell}{m}$ sets $J \in \mathcal{J}$, if $J(v_i) = J$ we have (51). For each $k \geq 1$, and each interval $i \in \{(k-1)\delta n+1,\dots,k\delta n\}$, by (PRE4), for all but at most $\frac{2\delta n}{|\mathcal{J}|}\cdot 6\frac{\ell}{m}$ choices of i we have (51). Thus, using the notation from (42), after summing up (50) and (51) we obtain

$$\sum_{h=2}^{t-1}\sum_{i\in\mathcal{D}_h}\mathbf{E}[q_{X,i}|\mathscr{H}_{h-1}] = \Big(\sum_{\substack{i=(k-1)\delta n+1\\ v_i^+\neq v_1, v_iv_i^+\notin\mathcal{R}}}^{\max(t,k\delta n)} p_{X,i}\Big) \pm \frac{12\ell\delta n}{m|\mathcal{J}|}\cdot\frac{3000\delta_{k\delta n}m}{\ell\gamma^7}\,,$$

where (as in (43)) since $i \leq k\delta n$ we have $\delta_i \leq \delta_{k\delta n}$. Since $0 \leq q_{X,i}, p_{X,i} \leq 6$ for each i, taking into account the at most εn values of i with $v_iv_i^+ \in \mathcal{R}$ and at most $\frac{\eta n}{\log n}$ values with $v_i^+ = v_1$, we have

$$\begin{aligned}
\sum_{h=1}^{t-1}\mathbf{E}[Y_h|\mathscr{H}_{h-1}] &= \sum_{h=1}^{t-1}\sum_{i=(k-1)\delta n+1\,,\,v_i^+=v_h}^{\max(t,k\delta n)}\mathbf{E}[q_{X,i}|\mathscr{H}_{h-1}]\\
&\overset{(7)}{=} \sum_{i=(k-1)\delta n+1}^{\max(t,k\delta n)} p_{X,i} \pm \tfrac{10^5\delta m\delta_{k\delta n}}{\gamma^7} \pm 6\varepsilon n \pm 6\tfrac{\eta n}{\log n}\,.
\end{aligned}$$

Finally, since $0 \leq Y_h \leq 6\deg_T(v_h)$, by Lemma 7 and (22), with probability at least $1 - n^{-10}$, if $\mathcal{E}_t$ holds then we have

$$\sum_{i=(k-1)\delta n+1}^{\max(t,k\delta n)} q_{X,i} = \sum_{h=1}^{t-1} Y_h = \sum_{i=(k-1)\delta n+1}^{\max(t,k\delta n)} p_{X,i} \pm \tfrac{2\cdot 10^5 \delta m \delta_{k\delta n}}{\gamma^7} \pm \delta^2 n$$
$$= \sum_{i=(k-1)\delta n+1}^{\max(t,k\delta n)} p_{X,i} \pm \tfrac{10^6 \delta m \delta_{k\delta n}}{\gamma^7} \,.$$

Taking the union bound over the choices of t and X and k, we see that with probability at least $1 - n^{-8}$, for each t such that $\mathcal{E}_t$ holds, we have (38) for all $X \in \mathcal{X}$ and k. Since the good event of Claim 1 holds with probability at least $1 - n^{-1}$, we conclude that the statement holds with probability at least $1 - 4n^{-1}$, as desired. □

We are now in a position to complete the proof of Theorem 9 by showing that with high probability (A_t, C_t) is δ_t-quasirandom for each $0 \leq t \leq n-1$. Suppose that the good events of Claims 1, 2 and 3 hold; this event $\mathcal{E}$ has probability at least $1 - 7n^{-1}$. Let $\mathscr{H}_0$ be the empty history, and $\mathscr{H}_i$ denote the history up to and including the labelling of v_i for each $1 \leq i \leq n$.

The proof that (A_t, C_t) is δ_t-quasirandom goes by induction on t, with the base case $t = 0$ being trivial. Observe that $(A_1, C_1) = (\mathbb{A}, \mathbb{C})$ is by definition δ_1-quasirandom. Now suppose that (A_i, C_i) are δ_i-quasirandom for each $1 \leq i \leq t$.

Now suppose $1 < i \leq t$ is such that $v_i v_i^+ \notin \mathcal{R}$. By (PRE3), $v_h = v_i^+$ satisfies $h \geq i - \frac{\varepsilon n}{\log n}$, and thus (A_h, C_h) and (A_i, C_i) differ by at most $\frac{2\varepsilon n}{\log n}$ vertex labels and at most $\frac{2\varepsilon n}{\log n}$ edge labels. Thus for any $I_{\mathsf{E}} \in \mathcal{I}_{\mathsf{E}}$, we have

$$\begin{aligned}&\big|\big\{a \in \mathrm{Adm}(\psi_h(v_h), J(v_i); A_i, C_i) : |a - \psi_h(v_h)| \in I_{\mathsf{E}}\big\}\big|\\ =&\big|\big\{a \in \mathrm{Adm}(\psi_h(v_h), J(v_i); A_h, C_h) : |a - \psi_h(v_h)| \in I_{\mathsf{E}}\big\}\big| \pm \tfrac{4\varepsilon n}{\log n} \,.\end{aligned}$$

Furthermore, we have

$$\begin{aligned}\big|\mathrm{Adm}\big(\psi_h(v_h), J(v_i); A_i, C_i\big)\big| &= \big|\mathrm{Adm}\big(\psi_h(v_h), J(v_i); A_h, C_h\big)\big| \pm \tfrac{4\varepsilon n}{\log n}\\ &\overset{(12)}{=} \ell\big(|A_h|/\tilde{n}\big)^2 \pm 2\delta_h \ell \pm \tfrac{4\varepsilon n}{\log n} = \ell |A_h|^2 \tilde{n}^{-2} \pm 3\delta_i \ell \,,\end{aligned} \tag{52}$$

where in the final equality we use $\delta_h < \delta_i$ since $h < i$. Thus we have

$$\begin{aligned}&\mathbf{E}\big[|\psi_i(v_i) - \psi_i(v_i^+)| \in I_{\mathsf{E}} \big| \mathscr{H}_{i-1}\big]\\ =&\frac{\big|\big\{a \in \mathrm{Adm}(\psi_h(v_h), J(v_i); A_i, C_i) : |a - \psi_h(v_h)| \in I_{\mathsf{E}}\big\}\big|}{\big|\mathrm{Adm}\big(\psi_h(v_h), J(v_i); A_i, C_i\big)\big|}\\ =&\frac{\big|\big\{a \in \mathrm{Adm}(\psi_h(v_h), J(v_i); A_h, C_h) : |a - \psi_h(v_h)| \in I_{\mathsf{E}}\big\}\big| \pm \tfrac{4\varepsilon n}{\log n}}{\ell |A_h|^2 \tilde{n}^{-2} \pm 2\delta_i \ell}\\ \overset{(36)}{=}& q_{I_{\mathsf{E}},i}\big(1 \pm \tfrac{50\delta_i}{\gamma^2}\big) \pm \frac{50\varepsilon n}{\gamma^2 \ell \log n} \,.\end{aligned} \tag{53}$$

We now argue that a similar equation for $q_{X,i}$ holds. Given $X \in \mathcal{X}$, we have

$$\begin{aligned}&\sum_{a\in\mathrm{Adm}(\psi_h(v_h),J(v_i);A_i,C_i)} \big|\big\{X'\in X(A_i,C_i) : a\in \mathrm{Ch}_{\mathsf{V}}(X';X) \text{ or } |a-\psi_h(v_h)|\in \mathrm{Ch}_{\mathsf{E}}(X';X)\big\}\big|\\ &= \sum_{a\in\mathrm{Adm}(\psi_h(v_h),J(v_i);A_h,C_h)} \big|\big\{X'\in X(A_h,C_h) : a\in \mathrm{Ch}_{\mathsf{V}}(X';X) \text{ or } |a-\psi_h(v_h)|\in \mathrm{Ch}_{\mathsf{E}}(X';X)\big\}\big| \pm 24\cdot\tfrac{2\varepsilon n}{\log n}\end{aligned}$$

and putting this together with (52) we obtain

$$\begin{aligned}&\mathbf{E}\Big[\big|\big\{X'\in X(A_i,C_i) : \psi_i(v_i)\in \mathrm{Ch}_{\mathsf{V}}(X';X) \text{ or } |\psi_i(v_i)-\psi_h(v_h)|\in \mathrm{Ch}_{\mathsf{E}}(X';X)\big\}\big|\Big|\mathscr{H}_{i-1}\Big]\\ &=\frac{\sum_{a\in\mathrm{Adm}(\psi_h(v_h),J(v_i);A_i,C_i)} \big|\big\{X'\in X(A_i,C_i) : a\in \mathrm{Ch}_{\mathsf{V}}(X';X) \text{ or } |a-\psi_h(v_h)|\in \mathrm{Ch}_{\mathsf{E}}(X';X)\big\}\big|}{\big|\mathrm{Adm}\big(\psi_h(v_h),J(v_i);A_i,C_i\big)\big|}\\ &=\frac{\sum_{a\in\mathrm{Adm}(\psi_h(v_h),J(v_i);A_h,C_h)} \big|\big\{X'\in X(A_h,C_h) : a\in \mathrm{Ch}_{\mathsf{V}}(X';X) \text{ or } |a-\psi_h(v_h)|\in \mathrm{Ch}_{\mathsf{E}}(X';X)\big\}\big| \pm \frac{50\varepsilon n}{\log n}}{\ell|A_h|^2\tilde{n}^{-2}\pm 3\delta_i\ell}\\ &\overset{(36)}{=} q_{X,i}\big(1\pm\tfrac{50\delta_i}{\gamma^2}\big)\pm\tfrac{400\varepsilon n}{\gamma^2\ell\log n}\,.\end{aligned} \tag{54}$$

We are finally in position to estimate $\sum_{i=2}^{t}\mathbf{E}\big[|\psi_i(v_i)-\psi_i(v_i^+)|\in I_{\mathsf{E}}\big|\mathscr{H}_{i-1}\big]$, which is a key quantity in order to verify (QUASI1). Putting together Claim 3 and (53), we have

$$\begin{aligned}\sum_{i=2}^{t}\mathbf{E}\big[|\psi_i(v_i)-\psi_i(v_i^+)|\in I_{\mathsf{E}}\big|\mathscr{H}_{i-1}\big] &= \sum_{i=1}^{t}\Big(q_{I_{\mathsf{E}},i}\big(1\pm\tfrac{50\delta_i}{\gamma^2}\big)\pm\tfrac{50\varepsilon n}{\gamma^2\ell\log n}\Big)\\ &= \sum_{k=1}^{\lceil\frac{t}{\delta n}\rceil}\Big(\big(1\pm\tfrac{50\delta_i}{\gamma^2}\big)\sum_{i=(k-1)\delta n+1}^{\max(t,k\delta n)} q_{I_{\mathsf{E}},i}\Big)\pm\tfrac{50\varepsilon n^2}{\gamma^2\ell\log n}\\ &= \sum_{k=1}^{\lceil\frac{t}{\delta n}\rceil}\Big(\big(1\pm\tfrac{50\delta_i}{\gamma^2}\big)\sum_{i=(k-1)\delta n+1}^{\max(t,k\delta n)} p_{I_{\mathsf{E}},i}\pm\tfrac{4000\delta_{k\delta n}\delta m}{\gamma^4}\Big)\pm 50\delta m\\ &= \sum_{i=1}^{t}p_{I_{\mathsf{E}},i}\pm\sum_{k=1}^{\lceil\frac{t}{\delta n}\rceil}\Big(\tfrac{50\delta_{k\delta n}}{\gamma^2}\cdot\tfrac{2\delta n}{|\mathcal{J}|}+\tfrac{4000\delta_{k\delta n}\delta m}{\gamma^4}\Big)\pm 50\delta m\\ &\overset{(8),(7)}{=}\sum_{i=1}^{t}p_{I_{\mathsf{E}},i}\pm\tfrac14\delta_t m\,.\end{aligned}$$

Next, by Lemma 7, with probability at least $1-n^{-10}$, if $\mathcal{E}$ occurs we have

$$\sum_{i=2}^{t}\mathbb{1}_{|\psi_i(v_i)-\psi_i(v_i^+)|\in I_{\mathsf{E}}}=\sum_{i=1}^{t}p_{I_{\mathsf{E}},i}\pm\tfrac14\delta_t m\pm\delta n\,,$$

and, using Claim 2, we conclude that

$$|C_{t+1} \cap I_{\mathsf{E}}| = |I_{\mathsf{E}}| - \sum_{i=2}^{t} \mathbb{1}_{|\psi_i(v_i)-\psi_i(v_i^+)| \in I_{\mathsf{E}}} - \sum_{i=1}^{t} r_{I_{\mathsf{E}},i} = |I_{\mathsf{E}}| - \tfrac{t}{|\mathcal{J}|} \pm \tfrac{3}{4}\delta_t m \overset{(23)}{=} \frac{m|A_{t+1}|}{\tilde{n}} \pm \delta_{t+1} m$$

as required for (QUASI1). Taking the union bound over I_{E}, with probability at least $1-n^{-9}$, if $\mathcal{E}$ occurs we have (QUASI1) for (A_{t+1}, C_{t+1}).

By a similar argument, but using (54), with probability at least $1-n^{-5}$, if $\mathcal{E}$ occurs we have (QUASI2) for (A_{t+1}, C_{t+1}). Putting these together, with probability at least $1-n^{-4}$, if $\mathcal{E}$ occurs then (A_{t+1}, C_{t+1}) is δ_{t+1}-quasirandom, as desired, and therefore Algorithm 1 does not fail at time $t+1$. Thus, if $\mathcal{E}$ occurs, taking the union bound over $1 \le t \le n$ we conclude by induction that, with probability at least $1-n^{-3}$, (A_{t+1}, C_{t+1}) is δ_{t+1}-quasirandom for each $1 \le t \le n$, and Algorithm 1 does not fail at any time.

Finally, $\mathcal{E}$ occurs with probability at least $1-7n^{-1}$, so we conclude that Algorithm 1 succeeds with probability at least $1-8n^{-1} > 0$. When it succeeds, the resulting ψ_n is the desired graceful labelling of T with $(1+\gamma)n$ labels. □

5. Concluding remarks

5.1. **Improvements on Theorem 4.** It would be desirable to remove the degree constraint of Theorem 4 and prove that all sufficiently large trees have approximate graceful labellings. Inspection of our proof reveals that the log-factor in our degree bound is required only in order to have polynomially small probabilities in various places, which in turn we need to apply the union bound. If we could somehow do without this, we would otherwise require only $\Delta(T) \le \eta n$. In our proof, we take the union bound over all times t and intervals $S \subseteq [t]$, and over all structures $\mathcal{X}$ (and over a bounded number of other choices). The former can be avoided: it suffices to establish quasirandomness for t a multiple of εn, and similarly to discretise the choice of S in (PRE4). The latter cannot so easily be avoided: we need to ensure that (for example) at each time t there are admissible vertices for labelling each v_t, which requires $X_3 \llbracket \psi_{t-1}(v_t^+), J(v_t) \rrbracket (A_t, C_t)$ to be non-empty. We cannot afford occasional failures here, which is what we would expect if the probability of its being empty were a small constant, rather than polynomially small. Nevertheless, it is possible that with more care, following a strategy similar to that presented here one can handle all trees with maximum degree εn, and perhaps even all trees.

It would be very interesting to obtain genuine graceful labellings for a large class of trees. Perhaps the approach here can be put together with the Absorbing Method for this purpose. We are currently investigating this possibility.

5.2. **Bipartite graceful labelling.** We believe our result can be extended to the notion introduced as α-valuations by Rosa [18], and now commonly referred to as *bipartite graceful labellings*. Namely, we say that a labelling ψ of a bipartite graph is *bipartite graceful* if it is graceful and all vertex labels in one colour class V_1 are smaller than in the other V_2.

However, already Rosa found out that the concept is too restrictive in that the bipartite version of Conjecture 1 does not hold (for example, a complete ternary tree of depth 2 and with 13 vertices does not have a bipartite graceful labelling). However, it seems likely that small modifications to our proof can be used to show that the trees as in Theorem 4 have approximate bipartite graceful labellings. That is, the labelling ψ in Theorem 4 can be taken such that $\psi(V_1) \subseteq \{1, \ldots, |V_1| + \frac{\gamma n}{2}\}$ and $\psi(V_2) \subseteq \{|V_1| + \frac{\gamma n}{2} + 1, \ldots, n + \gamma n\}$.

Let us briefly sketch the required modifications. By a similar reduction as that to Theorem 9, we can assume that $|V_1|, |V_2| \geq \gamma n$ in addition to any divisibility properties we need. We need to redefine J and 'complementary interval' to replace the 'midpoint' $\frac{1+\gamma}{2}n$ in our proof with $|V_1| + \frac{\gamma n}{2}$. We then need to alter our procedure of picking the intervals $J(v_i)$, insisting that $J(v_i)$ is always below the midpoint for $v_i \in V_1$, and always above for $v_i \in V_2$, but otherwise following the same random procedure. We will also need to alter the distributions $\mathrm{Cor}_{\mathsf{V}}$ and $\mathrm{Cor}_{\mathsf{E}}$ appropriately. Finally, we apply Algorithm 1, and claim that it succeeds with high probability; doing so it returns an approximate bipartite graceful labelling.

In order to prove this, it is necessary to change (PRE4), taking into account the ratio of vertices in V_1 and in V_2 on the interval S. Similarly, it is necessary to change the quasirandomness definition (QUASI2), again taking into account the ratio of vertices in V_1 and V_2 up to time t. We expect that these modifications to our proof suffice, but these promise significant extra technical complexity, and we have not checked the details.

In light of this sketch, we do not think Theorem 4 should be seen as strong evidence in favour of the Graceful Tree Conjecture. The above sketch would be similarly strong evidence in favour of the statement that all trees admit a bipartite graceful labelling: which is false.

5.3. **Harmonious labellings.** The concept of *harmonious labellings* is the same as of graceful labellings, except that the formula $|f(x) - f(y)|$ defining the label induced on the edge xy is replaced by $(f(x) + f(y)) \mod q$, where q is the number of edges of the graph in question. This concept was introduced by Graham and Sloane [10] who also put forward the counterpart to Conjecture 1.

Conjecture 15 (Harmonious Tree Conjecture)**.** *For any n-vertex tree T there exists an injective labelling $\psi : V(T) \to [n]$ such that the values*

$$\Big(\psi(u) + \psi(v) \mod (n-1)\Big)_{uv \in E(T)} \tag{55}$$

are pairwise distinct.

This conjecture is, too, wide open.

We believe that the tools we introduced in the paper may be used to obtain a counterpart of Theorem 4 for harmonious labellings. The bound on the maximum degree of the tree T would stay $\frac{\eta n}{\log n}$. The relaxation compared to Conjecture 15 would amount to ψ mapping to $[\tilde{n}]$, $\tilde{n} := \lceil (1+\gamma)n \rceil$, and modulus in (55) being $\tilde{n} - 1$.

In order for our analysis of Algorithm 1 to work, we need the property that the marginal distributions of vertex labels and edge labels are close to uniform throughout the whole process. We obtained this by our careful choice of the sets $J(v)$ for $v \in V(T)$. Once we have this, the remaining analysis does not essentially require gracefulness.

In order to modify our method to work for harmonious labellings we would choose $\mathcal{J}$ as follows. We consider $[\tilde{n}]$ with the natural cyclic order, and let $\mathcal{J}$ be the collection of intervals of length $\ell - 1$ in this order starting at 1, $m+1, \ldots, \tilde{n} - m + 1$. We would not need to define 'complementary interval'. Then, in Lemma 11, we would simply choose $J(v)$ independently and uniformly at random from $\mathcal{J}$ for each $v \in V(T)$. It is now obvious that if each v were labelled uniformly in $J(v)$ then the result is a uniform distribution of vertex labels, and easy to check that the distribution of edge labels is also uniform. We expect that from this point one can simply follow the algorithm and analysis given, making the obvious small changes to obtain a harmonious rather than graceful labelling. However we have not checked the details.

Note that in our approach we cut the tree T into small subtrees by removing the edges $\mathcal{R}$. This is not used only to assign the intervals $J(v)$, but also to guarantee that most vertices are labelled shortly after their parent is labelled. The former property is not required for harmonious labelling, but the latter property is still required for the analysis.

6. Acknowledgment

Part of the project was done while CG and JH visited Max Planck Institute for Informatics. We thank Michal Adamaszek for his contribution to this project in its initial stages.

The contents of this publication reflects only the authors' views and not necessarily the views of the European Commission of the European Union.

References

1. J. C. Bermond, *Graceful graphs, radio antennae and French windmills*, Graph Theory and Combinatorics, Pitman, London, 1979, pp. 18–37.
2. J. Böttcher, J. Hladký, D. Piguet, and A. Taraz, *An approximate version of the tree packing conjecture via random embeddings*, APPROX/RANDOM 2014 (Germany) (K. Jansen et al., ed.), Leibniz International Proceedings in Informatics, vol. 28, Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2014, pp. 490–499.
3. J. Böttcher, Jan Hladký, D. Piguet, and A. Taraz, *An approximate version of the tree packing conjecture*, Israel J. Math. **211** (2016), no. 1, 391–446.
4. F. Van Bussel, *Relaxed graceful labellings of trees*, Electronic J. Combin. **9** (2002), 12p.
5. W. C. Chen, H. I. Lü, and Y. N. Yeh, *Operations of interlaced trees and graceful trees*, Southeast Asian Bull. Math. **21** (1997), 337–348.
6. A. Ferber, C. Lee, and F. Mousset, *Packing spanning graphs from separable families*, arXiv:1512.08701.
7. R. Frucht, *Graceful numbering of wheels and related graphs*, Annals New York Academy of Sciences **319** (1979), 219–229.
8. J.A. Gallian, *A dynamic survey of graph labeling*, Electronic J. Combin. (2014), 384p.
9. S. W. Golomb, *How to number a graph*, Graph Theory and Computing (R. C. Read, ed.), Academic Press, 1972, pp. 23–37.

10. R. L. Graham and N. J. A. Sloane, *On additive bases and harmonious graphs*, SIAM Journal on Algebraic and Discrete Methods **1** (1980), no. 4, 382–404.
11. W. Hoeffding, *Probability inequalities for sums of bounded random variables*, J. Amer. Statist. Assoc. **58** (1963), 13–30.
12. F. Joos, J. Kim, D. Kühn, and D. Osthus, *Optimal packings of bounded degree trees*, arXiv:1606.03953.
13. J. Kim, D. Kühn, D. Osthus, and M. Tyomkyn, *A blow-up lemma for approximate decompositions*, arXiv:1604.07282.
14. S. Messuti, V. Rödl, and M.. Schacht, *Packing minor-closed families of graphs into complete graphs*, arXiv:1602.06780, to appear in Journal of Combinatorial Theory, Series B.
15. J. W. Moon, *On the maximum degree in a random tree*, Michigan Math. J. **15** (1968), no. 4, 429–432.
16. A. M. Pastel and H. Raynaud, *Numerotation gracieuse des oliviers*, Colloq. Grenoble, Publications Université de Grenoble, 1978, pp. 218–223.
17. G. Ringel, *Problem 25*, Theory of Graphs and its Applications (Proc. Sympos. Smolenice 1963), Nakl. CSAV, Praha, 1964, p. 162.
18. A. Rosa, *On certain valuations of the vertices of a graph*, Theory of Graphs (Internat. Symposium, Rome, July 1966), Gordon and Breach, N. Y. and Dunod Paris, 1967, pp. 349–355.
19. G. Sethuraman and J. Jesintha, *All banana trees are graceful*, Advances Appl. Disc. Math. **4** (2009), 53–64.
20. T.-M. Wang, C.-C. Yang, L.-H. Hsu, and E. Cheng, *Infinitely many equivalent versions of the graceful tree conjecture*, Appl. Anal. Discrete Math. **9** (2015), to appear.
21. N. C. Wormald, *Differential equations for random processes and random graphs*, Ann. Appl. Probab. **5** (1995), no. 4, 1217–1235.

Department of Computer Science, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen, Denmark.
E-mail address: `anna@mpi-inf.mpg.de`

Department of Mathematics, London School of Economics, Houghton Street, London, WC2A 2AE, UK.
E-mail address: `p.d.allen@lse.ac.uk`

Freie Universität Berlin, Germany.
E-mail address: `grosu.codrut@gmail.com`

Institute of Mathematics of the Czech Academy of Sciences, Žitná 25, Praha.
E-mail address: `honzahladky@gmail.com`