

# Primality of trees

PENNY HAXELL<sup>\*</sup>, OLEG PIKHURKO<sup>†</sup> AND ANUSCH TARAZ<sup>‡</sup>

A graph of order  $n$  is *prime* if one can bijectively label its vertices with integers  $1, \dots, n$  so that any two adjacent vertices get coprime labels. We prove that all bipartite  $d$ -degenerate graphs with separators of size at most  $n^{1-O_d(1/\ln \ln n)}$  are prime. It immediately follows that all large trees are prime, confirming an old conjecture of Entringer and Tout from around 1980. Also, our method allows us to determine the smallest size of a non-prime connected order- $n$  graph for all large  $n$ , proving a conjecture of Rao [*R. C. Bose Centenary Symposium on Discrete Math. and Applications*, Kolkata, 2002] in this range.

AMS 2000 SUBJECT CLASSIFICATIONS: 05C78.

KEYWORDS AND PHRASES: Coprime graph, prime labeling, Entringer–Tout conjecture.

## 1. Introduction

The *coprime graph*  $S_n$  has vertex set  $[n] := \{1, \dots, n\}$  in which two vertices are adjacent if and only if they are coprime (as numbers). For example,  $S_5$  is isomorphic to  $K_5$  minus one edge. Various questions and results about combinatorial properties of  $S_n$  can be found in Erdős [7, 8, 9], Erdős, Sárközy, and Szemerédi [10, 12], Szabó and Tóth [30], Erdős and Sárközy [11, 13], Ahlswede and Khachatrian [1, 2, 3], Sárközy [28], and others.

A graph  $G$  of order  $n$  is called *prime* if it is a subgraph of  $S_n$ , that is, if there is a bijection  $f : V(G) \rightarrow [n]$  such that any two adjacent vertices of  $G$  are assigned coprime numbers. This notion was introduced by Entringer who, according to [15], conjectured around 1980 that every tree is prime. The earliest statement of this conjecture that we could find in the literature comes from the 1982 paper of Tout, Dabbouchy, and Howalla [33], where its

---

<sup>\*</sup>Partially supported by NSERC.

<sup>†</sup>Partially supported by the National Science Foundation, Grant DMS-0758057, and the Alexander von Humboldt Foundation.

<sup>‡</sup>Partially supported by the Deutsche Forschungsgemeinschaft, Grant DFG TA 309/2-2.

formulation is preceded by the names of Entringer and Tout. Therefore we shall refer to it as the *Entringer–Tout Conjecture*.

One popular direction of research was to verify this conjecture for some very special classes of trees (small trees, caterpillars, spiders, complete binary trees, olive trees, palm trees, banana trees, twigs, binomial trees, bistars, etc). We refer the reader to the dynamic survey by Gallian [15, Section 7.2] for references to these and related results.

Here we prove this conjecture for all large  $n$ .

**Theorem 1.** *There exists  $n'$  such that every tree with  $n \geq n'$  vertices is prime.*

In fact, we can show a more general result, extending Theorem 1 to a larger class of graphs. In order to state it, we have to present some definitions first. We say that a graph  $G$  is  $d$ -degenerate if every non-empty subgraph of  $G$  has a vertex of degree at most  $d$ . For example, a graph is 1-degenerate if and only if it is acyclic. Let us call a graph  $G$   $s$ -separable if for every subgraph  $G' \subseteq G$  there is a set  $S \subseteq V(G')$  such that  $|S| \leq s$  and each component of  $G' - S$  has at most  $|V(G')|/2$  vertices. The choice of the constant  $1/2$  is rather arbitrary; we choose it for the convenience of calculations and because of the well-known fact that trees are 1-separable. Also, in order to make our results stronger, we use a weaker version of separability where the upper bound  $s$  depends only on  $|V(G)|$  and not on  $|V(G')|$ .

Lipton and Tarjan [23] showed that every order- $n$  planar graph  $G$  contains a set  $X$  with  $|X| \leq 2\sqrt{2n}$  such that no component of  $G - X$  has more than  $2n/3$  vertices. Clearly, by applying this theorem twice to any given subgraph  $G' \subseteq G$ , we can eliminate all components of order larger than  $|V(G')|/2$ . Thus  $G$  is  $4\sqrt{2n}$ -separable. Likewise, the result of Alon, Seymour, and Thomas [4] implies that any order- $n$  graph without a  $K_h$ -minor is  $2h^{3/2}n^{1/2}$ -separable.

Given an integer  $d \geq 1$ , define a function  $s = s(n)$  by

$$(1) \quad s(n) := n^{1 - \frac{10^6 \cdot d}{\ln \ln n}}.$$

Here is the main result of this paper.

**Theorem 2.** *For every  $d \geq 1$  there exists  $n''$  such that every  $s(n)$ -separable bipartite  $d$ -degenerate graph  $F$  of order  $n \geq n''$  is prime, where  $s(n)$  is defined by (1).*

Mader [24] showed that every  $K_h$ -minor free graph  $F$  has average degree at most  $f(h)$ , with more precise estimates on the function  $f(h)$  given by

Kostochka [20] and Thomason [31, 32]. Since not containing a  $K_h$ -minor is a hereditary property, such a graph  $F$  is necessarily  $f(h)$ -degenerate. This and the above-mentioned result of Alon, Seymour, and Thomas [4] allow us to deduce the following from Theorem 2.

**Corollary 3.** *For every  $h$ , all sufficiently large bipartite graphs without a  $K_h$ -minor are prime.*

Based on an earlier version of this manuscript that had a slightly simpler proof just for trees and using the results of Dusart [6] on the distribution of primes, Spiess [29] estimated that taking  $n' = 10^{10^{100}}$  in Theorem 1 is a suitable choice. Although the Entringer–Tout Conjecture for trees of small order  $n$  seems quite amenable (see [14, 25, 26, 22] with the current record  $n \leq 206$  claimed in a manuscript of Kuo and Fu [21]), closing this gap is beyond any small-order approaches. Therefore, we make no attempt to optimize the constants.

Two main difficulties in proving Theorem 2 are that we have to use every element of  $[n]$  as a label (that is, we look for spanning subgraphs in  $S_n$ ) and that  $S_n$  has a large independent set  $\{2, 4, 6, \dots\}$ . On the other hand, every vertex in the set

$$(2) \quad P_1 := \{p \in [n] : p > n/2 \text{ and } p \text{ is prime}\}$$

is *universal*, that is, it is adjacent to all other vertices of  $S_n$ . Likewise, every vertex  $2p$  in the set

$$(3) \quad P_0 := \{2p \in [n] : p > n/3 \text{ and } p \text{ is prime}\}$$

is adjacent to all vertices in  $S_n$  with odd labels, except  $p$ . The existence of these two sets, each of order  $\Theta(n/\ln n)$ , crucially helps in our proof.

We split the whole proof into three lemmas which are stated in Section 3, deriving Theorem 2 from them in Section 4. In brief, Lemma 5 splits the given graph  $F$  satisfying the assumptions of Theorem 2 into tiny components by removing a small set  $M$  of vertices using the separability property. It also arranges these components into groups in order to balance more evenly the distribution of vertices among groups. Then Lemma 6 specifies where each group is to be mapped inside  $[n]$ . Since we do not have much control over the vertices in  $M$ , they are mapped into  $P_0 \cup P_1$ . As we have already mentioned, one has to be careful to ensure that every group has a sufficiently large independent set to host all even labels that are assigned to it. Apart from the Prime Number Theorem, we use only very basic results about divisibility and primality of integers. Finally, Lemma 7 shows how to embed each group

into its assigned part of  $S_n$ ; this is the point when we need the  $d$ -degeneracy property.

The proof benefits from some of the ideas in [26], where the second author proved that any  $n$ -vertex tree can be embedded into  $S_{(1+\varepsilon)n}$ . The presentation of the proof as three separate lemmas not only improves the readability but also directly states which properties are needed at each step. Also, these lemmas may be useful on their own in proving further results on prime labeling or more general embedding problems.

Here is one example. Rao [27] asked about the value of  $m(n)$ , the smallest size of a connected non-prime graph of order  $n$ . An upper bound on  $m(n)$  can be obtained as follows. Suppose first that  $n = 2k$  is even. Let  $F'$  be the vertex-disjoint union of  $k - 3$  edges and two triangles. Clearly, the independence number  $\alpha(F') \leq k - 1$  because any independent set contains at most one vertex from each clique of  $F'$ . Thus  $F'$  is not prime because there is not enough space to fit all even labels. Finally, make  $F'$  connected by adding  $k - 2$  extra edges. The resulting graph  $F$  shows that  $m(2k) \leq 2k + 1$ . If  $n = 2k + 1$  is odd, then we let  $F'$  be the vertex-disjoint union of  $k - 4$  edges and three triangles. Again,  $F'$  is not prime because  $\alpha(F') \leq k - 1$ . We can make  $F'$  connected by adding  $k - 2$  extra edges, which shows that  $m(2k + 1) \leq 2k + 3$ . Rao [27] conjectured that we have equality here for every  $n$ . Our methods allow us to prove this conjecture for all large  $n$  (see Section 8).

**Theorem 4.** *We have  $m(n) = 2 \lceil n/2 \rceil + 1$  for all large  $n$ .*

Our methods can prove, for all large  $n$ , another conjecture of Rao that every tree on  $n$  vertices admits a prime labeling that assigns the label  $n$  to a leaf. Apparently, his motivation was to prove the Entringer–Tout Conjecture via some induction. In light of Theorem 1, this question is not so interesting now, so we skip the proof.

Unfortunately, we could not make much progress on yet another conjecture of Rao [27] that the value of  $m'(n)$ , the smallest size of a (not necessarily connected) non-prime graph of order  $n$ , is given by the graph  $F'$  above, that is,  $m'(2k) = k + 3$  and  $m'(2k + 1) = k + 5$ .

## 2. Notation

We will use standard graph notation. In particular,  $V(G)$  and  $E(G)$  denote the set of vertices and the set of edges of  $G$  respectively. By  $G[A]$  we mean the graph induced by  $A \subseteq V(G)$  in  $G$ , while for disjoint  $A, B \subseteq V(G)$ ,  $G[A, B]$  denotes the induced bipartite graph with parts  $A$  and  $B$ . Also,

$\Gamma_G(x)$  denotes the neighborhood of  $x$  in  $G$ . When more than one graph is being considered, we sometimes use the term  $G$ -neighbor of  $x$  for an element of  $\Gamma_G(x)$ .

For integers  $m \leq n$ , we denote  $[m, n] := \{m, m + 1, \dots, n\}$ . Thus  $[1, n] = [n]$ . Also, let  $\mathbb{Z}_r := \{0, \dots, r - 1\}$  consist of the residues modulo  $r$ . A *cyclic interval* in  $\mathbb{Z}_r$  is a set of the form  $\{i, i + 1, \dots, i + m\}$  reduced modulo  $r$ . For example, every single-element set or the set  $\{0, r - 1\}$  is a cyclic interval. Let us enumerate all primes as  $p_0 := 2, p_1 := 3, p_2 := 5$ , and so on.

### 3. Statements of the Lemmas

The following definitions apply to each of Lemmas 5–7 as well as to the proof of Theorem 2. Let  $d$  be given. Define

$$\beta := \frac{1}{900d}.$$

Next, we define the following parameters as functions of  $n$ . Let

$$r_0 := n^{150/(\beta \ln \ln n)}.$$

Take the largest integer  $t$  such that

$$(4) \quad r := \prod_{i=1}^t p_i \leq r_0,$$

that is, the product of the first  $t$  odd primes is at most  $r_0$ . Finally, let

$$(5) \quad m := \left\lceil \frac{\beta n}{150r} \right\rceil \quad \text{and} \quad s := \frac{\beta^3 n}{40,000 r^3 \ln n}.$$

Note that, for all large  $n$ , this  $s$  is at least as large as the value  $s(n)$  defined in (1). So it is enough to prove Theorem 2 using the definition of  $s$  in (5).

We now make explicit some inequalities relating these parameters. Since we do not compute an explicit bound on the constant  $n''$  of Theorem 2, we will use the asymptotic version of the Prime Number Theorem that  $p_i = (1 + o(1)) i \ln i$  as  $i \rightarrow \infty$ . In particular, all but at most finitely many integers  $i$  satisfy, for example,  $(i \ln i)/2 \leq p_i \leq 2i \ln i$ . Thus, for every  $\varepsilon > 0$  and all large  $n$ , we have

$$e^{(1-2\varepsilon)t \ln t} \leq \left( \frac{\varepsilon t \ln(\varepsilon t)}{2} \right)^{(1-\varepsilon)t} \leq \prod_{i=\varepsilon t}^t p_i \leq r \leq (2t \ln t)^t \leq e^{(1+\varepsilon)t \ln t},$$

implying that

$$(6) \quad t = (1 + o(1)) \frac{\ln r}{\ln \ln r} \quad \text{and} \quad p_t = (1 + o(1)) \ln r.$$

It follows that  $1 \leq r_0/r \leq 2 \ln n$  for all large  $n$ . We also have that

$$(7) \quad \frac{ns}{m} \leq \frac{\beta^3 n^2}{40,000 r^3 \ln n} \times \frac{160 r}{\beta n} = \frac{\beta^2 n}{250 r^2 \ln n}$$

and

$$(8) \quad \ln n \ll r \ll n, \quad \frac{140 \ln n}{\beta \ln \ln n} < \ln r, \quad \ln \ln r = (1 + o(1)) \ln \ln n.$$

Moreover, for every integer  $d$  we can choose an integer  $n'$  such that for all  $n \geq n'$  the following three lemmas hold.

**Lemma 5** (Preparing the Graph  $F$ ). *Every  $s$ -separable bipartite graph  $F$  of order  $n \geq n'$  admits a partition*

$$(9) \quad V(F) = \cup_{i=0}^1 \left( M_i \cup \left( \cup_{j=0}^{r-1} A_{j,i} \right) \right)$$

such that, for  $m_i := |M_i|$  and  $a_{j,i} := |A_{j,i}|$ , we have

$$(10) \quad m_0 + m_1 = m,$$

$$(11) \quad \left\lfloor \frac{n}{2} \right\rfloor \leq m_0 + \sum_{j=0}^{r-1} a_{j,0} \leq \frac{n}{2} + m,$$

and all the following properties hold.

1.  $F - (M_0 \cup M_1)$  is the vertex-disjoint union of bipartite graphs  $F[A_{j,0}, A_{j,1}]$ ,  $j \in \mathbb{Z}_r$ .
2. Each component of  $F - (M_0 \cup M_1)$  has at most  $3ns/m$  vertices.
3.  $M_0 \cup (\cup_{j=0}^{r-1} A_{j,0})$  is an independent set in  $F$ .
4. For every cyclic interval  $J \subseteq \mathbb{Z}_r$  and  $i = 0, 1$ , we have

$$(12) \quad \left| \sum_{j \in J} a_{j,i} - \frac{n}{2r} |J| \right| \leq \frac{\beta n}{8r}.$$

**Lemma 6** (Preparing the Graph  $S_n$ ). *Let  $n \geq n'$  and suppose that non-negative integers  $m_0$ ,  $m_1$ , and  $a_{j,i}$ , indexed by  $j \in \mathbb{Z}_r$  and  $i = 0, 1$ , are*

given, satisfying (10), (11), (12), and  $\sum_{j=0}^{r-1}(a_{j,0} + a_{j,1}) = n - m$ . Then there exists a partition

$$(13) \quad [n] = \cup_{i=0}^1 \left( Q_i \cup \left( \cup_{j=0}^{r-1} R_{j,i} \right) \right)$$

of the vertex set of  $S_n$  such that the following properties hold for all  $j \in \mathbb{Z}_r$  and  $i = 0, 1$ .

1.  $Q_i \subseteq P_i$ ,  $|Q_i| = m_i$ , and  $|R_{j,i}| = a_{j,i}$ . (Recall that  $P_0$  and  $P_1$  are defined by (3) and (2).)
2. Every vertex in  $R_{j,i}$  has at least  $(1 - \beta)|R_{j,1-i}|$   $S_n$ -neighbors in  $R_{j,1-i}$ .
3. All even elements of  $[n]$  belong to  $Q_0 \cup (\cup_{j=0}^{r-1} R_{j,0})$ .

**Lemma 7** (Embedding Lemma). *Let  $G$  and  $T$  be bipartite graphs with bipartitions  $V(G) = V_0 \cup V_1$  and  $V(T) = W_0 \cup W_1$  such that  $|V_0| = |W_0|$ ,  $|V_1| = |W_1|$ , and  $T$  is  $d$ -degenerate. For  $i = 0, 1$ , let  $n_i := |V_i|$  and suppose that the following assumptions hold.*

1.  $n' \leq n_i \leq 2n_{1-i}$ .
2. Each component of  $T$  has at most  $\beta n_i / \ln n_i$  vertices in  $W_i$ .
3. Every vertex in  $V_i$  has at least  $(1 - \beta)n_{1-i}$   $G$ -neighbors in  $V_{1-i}$ .

Then there is a bijection  $f : V(T) \rightarrow V(G)$  that embeds  $T$  as a subgraph into  $G$  with  $f(V_i) = W_i$  for  $i = 0, 1$ .

Although there are many results in the literature on embedding a spanning graph into a dense bipartite graph (such as e.g. the Blow-up Lemma [19]), we were unable to find one that applies directly to our case and thus we provide a proof of Lemma 7 in Section 7.

### 4. Proof of Theorem 2

Given  $d$ , let  $n'$  be a large constant satisfying each of Lemmas 5–7. By (8), we may pick a sufficiently large  $n''$  so that, in particular, for all  $n \geq n''$  we have  $n/(4r) \geq n'$ , where  $r$  is defined in (4). Let  $n \geq n''$  be arbitrary and let a graph  $F$  satisfy all assumptions of Theorem 2. Apply Lemma 5 to  $F$ . It returns a partition (9). As in Lemma 5, we define  $m_i := |M_i|$  and  $a_{j,i} := |A_{j,i}|$  for  $j \in \mathbb{Z}_r$  and  $i = 0, 1$ . We use these numbers as the input for Lemma 6. This lemma returns a partition (13).

We embed each bipartite graph  $T := F[A_{j,0}, A_{j,1}]$  into  $G := S_n[R_{j,0}, R_{j,1}]$  using Lemma 7. Let us check that all assumptions of Lemma 7 are satisfied. Here  $W_i = A_{j,i}$ ,  $V_i = R_{j,i}$ , and  $n_i = a_{j,i}$  for  $i = 0, 1$ . By taking

$J := \{j\}$  in Conclusion 4 of Lemma 5, we have  $|n_i - n/(2r)| \leq \beta n/(8r)$ . Thus  $n_i \geq n/(4r) \geq n'$  and  $n_i/n_{1-i} \leq (1 + \beta/4)/(1 - \beta/4) < 2$ , that is, Assumption 1 of Lemma 7 holds. Also, Assumptions 2 and 3 follow from respectively Conclusion 2 of Lemma 5 and Conclusion 2 of Lemma 6. (Note that by (7) and the fact that  $n_i \geq n/(4r)$  we find  $3ns/m \leq \beta n_i/\ln n_i$ .) Observe that by Conclusion 3 of Lemma 6 every vertex of  $F - (M_0 \cup M_1)$  that gets an even label is in  $\cup_{j=0}^{r-1} A_{j,0}$ .

Next, for  $i = 0, 1$ , take an arbitrary bijection of  $M_i$  into  $Q_i$ . We have assigned labels to all vertices of  $F$ . By Conclusion 1 of Lemma 5, there are no edges in  $F$  between  $A_{j,0} \cup A_{j,1}$  and  $A_{h,0} \cup A_{h,1}$  for any distinct  $j, h \in \mathbb{Z}_r$ . Thus we can have problems only at edges that intersect  $M_0 \cup M_1$ . But all labels on this set are restricted to  $P_0 \cup P_1$ . Recall that all vertices of  $P_1$  are universal, thus every edge of  $F$  incident to a vertex of  $M_1$  is embedded correctly. Also, all even labels are assigned to a subset of  $M_0 \cup (\cup_{j=0}^{r-1} A_{j,0})$ , which is an independent set in  $F$  by Conclusion 3 of Lemma 5. Thus all edges of  $F$  incident to  $M_0$  are correctly embedded by the property of  $P_0$  (stated after (3)), except possibly some pairs of the form  $(p, 2p)$  where  $2p \in P_0$ , since they may be edges of  $F$  but are not present in  $S_n$ .

Let us show that we can always swap some labels and eliminate one such problematic pair, without creating any new conflicts. Suppose that some adjacent vertices  $x, y \in V(F)$  get respectively labels  $p$  and  $2p$  with  $2p \in P_0$ . Let  $X$  be the set of vertices of  $F$  that get labels from  $P_1$ . Let  $Y$  consist of vertices of  $F$  whose labels are powers of 2 that lie between 2 and  $n$ . Since  $|X| = \Theta(n/\ln n)$  and  $|Y| \geq \lfloor \log_2 n \rfloor$ , we cannot have a complete bipartite graph in  $F$  between  $X$  and  $Y$ . (Otherwise  $F[X \cup Y]$  would have minimum degree at least  $|Y| > d$ , contradicting the  $d$ -degeneracy of  $F$ .) Let  $x' \in X$  and  $y' \in Y$  be non-adjacent. Then we have the following situation:

- $x$  has label  $p$ , which is adjacent to every other vertex of  $S$  except  $2p$ ,
- $y$  has label  $2p$ , which is adjacent to every odd vertex of  $S$  except  $p$ ,
- the label of  $x'$  is adjacent to every other vertex of  $S$ ,
- the label of  $y'$  is adjacent to every odd vertex of  $S$ .

Thus we may swap the labels of  $x$  and  $x'$  and of  $y$  and  $y'$ , and in the new labeling the number of problematic pairs has gone down by at least 1. By repeating this step we eventually get a prime labeling of  $F$ , proving Theorem 2.

### 5. Proof of Lemma 5

Fix a 2-coloring of  $F$  and denote by  $W_0 \cup W_1 = V(F)$  the two color classes, where we assume that  $|W_0| \geq \lfloor n/2 \rfloor$ .

Initially, set  $M := \emptyset$ . As long as  $|M| \leq m - s$ , repeat the following. Take a component  $F'$  of  $F - M$  of the largest order. By the  $s$ -separability find a set  $S \subseteq V(F')$  with  $|S| \leq s$  such that no component of  $F' - S$  has more than  $|V(F')|/2$  vertices. Add  $S$  to  $M$ . When we stop, we have  $m - s < |M| \leq m$ . Finally, add some  $m - |M|$  remaining vertices to  $M$ , ensuring that  $|M| = m$  at the end.

It is not hard to show (see Lemma 8 in [26]) that the maximum order of a component of  $F - M$  is at most

$$\frac{2n}{\lfloor m/s \rfloor} \leq \frac{3ns}{m}.$$

Let  $F - M$  have  $c$  components with vertex sets  $A'_1 \cup \dots \cup A'_c = V(F) \setminus M$ . Let  $A'_{j,i} := A'_j \cap W_i$  for  $j \in [c]$  and  $i = 0, 1$ .

As the next step, we want to define a partition  $M = M_0 \cup M_1$  and perhaps *flip* some components  $F[A'_j] = F[A'_{j,0}, A'_{j,1}]$  (that is, swap the parts  $A'_{j,0}$  and  $A'_{j,1}$ ) so that

$$(14) \quad X := M_0 \cup (\cup_{j=1}^c A'_{j,0})$$

is an independent set in  $F$  and

$$(15) \quad \lfloor n/2 \rfloor \leq |X| \leq n/2 + m.$$

There are two cases that we have to consider. First, suppose that

$$(16) \quad \sum_{j=1}^c \max(|A'_{j,0}|, |A'_{j,1}|) \geq \lfloor n/2 \rfloor.$$

We let  $M_0 := \emptyset$  and  $M_1 := M$ . If initially  $|X| < \lfloor n/2 \rfloor$ , then by (16) there is  $j \in [c]$  such that  $|A'_{j,0}| < |A'_{j,1}|$ . We flip this component  $F[A'_j]$  and repeat. Clearly,  $|X|$  strictly increases each time. Since any component has at most  $3ns/m$  vertices, which by (7) is less than  $m$  even in order of magnitude, the value of  $|X|$  never jumps above  $n/2 + m$  and we eventually achieve (15). Likewise, if initially  $|X| > n/2 + m$ , then  $\sum_{j=1}^c |A'_{j,0}| > \sum_{j=1}^c |A'_{j,1}|$  and there is always  $j \in [c]$  with  $|A'_{j,0}| > |A'_{j,1}|$ . By iteratively flipping such components, we arrive at (15). Since  $M_0 = \emptyset$ , the set  $X$  is independent.

Next, suppose that (16) does not hold. We let  $M_0 := M \cap W_0$  and  $M_1 := M \cap W_1$ . Here we do not flip any components. Thus  $X = W_0$ . By the

definition of  $W_0$ ,  $X$  is independent and

$$\lfloor n/2 \rfloor \leq |X| \leq |M_0| + \sum_{j=1}^c \max(|A'_{j,0}|, |A'_{j,1}|) \leq m + \frac{n}{2},$$

since we assumed that (16) is false. Thus  $X$  has the required properties.

To complete the proof, it remains to collect together the sets  $A'_{j,i}$  into  $2r$  sets  $A_{j,i}$  satisfying Conclusion 4. To this end, we take a random partition  $[c] = R_0 \cup \dots \cup R_{r-1}$ , where each element of  $[c]$  is assigned to one of the  $r$  parts uniformly at random (and independently of all other choices). Next, for  $j \in \mathbb{Z}_r$  and  $i = 0, 1$ , let  $A_{j,i} := \cup_{h \in R_j} A'_{h,i}$ , and set  $a_{j,i} = |A_{j,i}|$ . (The graphs  $F[A_{j,0}, A_{j,1}]$  are what we called groups in the informal sketch in the Introduction.) We now check that at least one  $r$ -partition of  $[c]$  makes Conclusion 4 valid. By the Union Bound, it suffices to show that for every fixed cyclic interval  $J$  and every fixed index  $i = 0, 1$ , the probability that (12) fails is  $o(r^{-2})$ .

Let  $p := |J|/r$  and let  $\Delta := \sum_{j=1}^c |A'_{j,i}|^2$ . Since each  $|A'_{j,i}| \leq 3ns/m$  and  $\sum_{j=1}^c |A'_{j,i}| \leq n$ , we have  $\Delta \leq (m/(3s)) \cdot (3ns/m)^2 = 3n^2s/m$ . The value of  $a := \sum_{j \in J} a_{j,i}$  is the sum of  $c$  independent random variables  $X_1, \dots, X_c$ , where  $X_j$  assumes value  $|A'_{j,i}|$  with probability  $p$  and value 0 with probability  $1 - p$ . By (15), the expectation  $\mathbf{E}[a]$  is within  $2pm$  of  $n|J|/(2r)$  (for both  $i = 0, 1$ ). Thus if (12) fails, then  $a$  deviates from its mean by at least  $\beta n/(8r) - 2pm \geq \beta n/(9r)$ . (This is why the value of  $m$  was chosen as in (5).)

Hoeffding's inequality [17, Theorem 2] states that the probability of this is at most

$$\begin{aligned} 2 \exp\left(-2 \frac{(\beta n/(9r))^2}{\Delta}\right) &\leq 2 \exp\left(-2 \frac{(\beta n/(9r))^2}{3n^2s/m}\right) \\ &\stackrel{(5)}{\leq} 2 \exp\left(\frac{-2\beta^3 n}{243 \cdot 150r^3s}\right) \stackrel{(5)}{=} o(1/n^2) = o(1/r^2). \end{aligned}$$

(Our choice of  $s$  was determined by this calculation.) Thus, almost surely (12) holds for every  $J$  and  $i$ , as desired. This completes the proof of Lemma 5.

### 6. Proof of Lemma 6

For  $j \in [0, 2r - 1]$ , let  $D_j := \{x \in [n] : x \equiv j \pmod{2r}\}$  be the residue class of  $j$  modulo  $2r$ . Recall that  $r$  and  $t$  are as defined in (4).

Let us show that if  $j, h \in [0, 2r - 1]$  are coprime, then every vertex  $x \in D_j$  has at most  $\beta n / (8r)$  non-neighbors in  $D_h$  with respect to the graph  $S := S_n$ . If  $y \in D_h$  is not coprime to  $x$ , then for every prime  $p_i$  dividing both  $x$  and  $y$  we have  $i > t$ . (Otherwise, the same prime  $p_i$  divides both  $j$  and  $h$ , a contradiction.) Trivially, the element  $x \in [n]$  can have at most  $\log_{p_t} n$  distinct primes that exceed  $p_t$  and every such prime  $p_i$  gives at most  $n / (2rp_t) + 1$  possible non-neighbors  $y \in D_h$ . Note that we have by (6) and (8) that, for example,  $p_t \geq \frac{\ln r}{2} \geq \frac{70 \ln n}{\beta \ln \ln n}$ . Hence,  $x$  has at most

$$(17) \quad \log_{p_t} n \left( \frac{n}{2rp_t} + 1 \right) \leq \frac{\ln n}{\ln p_t} \times \frac{n}{rp_t} \stackrel{(8)}{\leq} \frac{\beta n}{8r}$$

non-neighbors in  $D_h$ , as required.

Let  $n_i := m_i + \sum_{j=0}^{r-1} a_{j,i}$  for  $i = 0, 1$ . By (5) and (8),  $m_i \leq m = o(n / \ln n)$ . Since  $|P_0|, |P_1| = \Theta(n / \ln n)$ , we may pick arbitrary disjoint subsets  $Q_0 \subseteq P_0$ , and  $Q_1, Q' \subseteq P_1$  of sizes  $m_0, m_1$ , and  $n_0 - \lfloor n/2 \rfloor$  respectively. (The last number is between zero and  $m$  by our assumption (11).) Set  $Q := Q_0 \cup Q_1 \cup Q'$ , then  $|Q| \leq 2m$ .

As a first guess, we take the following parts. For  $j \in \mathbb{Z}_r$  and  $i = 0, 1$ , let  $R_{j,i} := D_{2j+i} \setminus Q$ , except we let  $R_{0,0} := (D_0 \setminus Q) \cup Q'$ . Thus we have a partition  $[n] = V_0 \cup V_1$ , where  $V_i := Q_i \cup (\cup_{j=0}^{r-1} R_{j,i})$ . Then of the  $\lceil n/2 \rceil$  odd elements of  $S_n$ ,  $n_0 - \lfloor n/2 \rfloor$  of them are in  $V_0$  and the rest make up  $V_1$ . Thus  $|V_1| = \lceil n/2 \rceil - n_0 + \lfloor n/2 \rfloor = n - n_0 = n_1$ , and  $|V_0| = n - n_1 = n_0$ . (This was the whole point of “moving”  $Q'$  to  $V_0$ .)

Let  $i = 0$  or  $1$ . For  $j \in \mathbb{Z}_r$ , let  $d_j := a_{j,i} - |R_{j,i}|$  denote the *discrepancy* at  $R_{j,i}$  relative to the desired value. We are going to correct this by moving vertices between parts. For  $h \in \mathbb{Z}_r$ , define  $\sigma_h := \sum_{k=0}^h d_k$ . Since the sum of all discrepancies  $\sum_{k=0}^{r-1} d_k$  is zero, the relation  $\sigma_h - \sigma_{h-1} = d_h$  also holds for  $h = 0$ . Consider the case  $i = 0$  first. Let  $h \in \mathbb{Z}_r$  be the index such that  $\sigma_h$  is minimum. For  $j \in \mathbb{Z}_r$ , move  $\sigma_{j-1} - \sigma_h \geq 0$  vertices from  $R_{j,0}$  to  $R_{j-1,0}$ . This changes the size of each  $R_{j,0}$  by  $(\sigma_j - \sigma_h) - (\sigma_{j-1} - \sigma_h) = d_j$ , thus making  $|R_{j,0}|$  exactly  $a_{j,0}$ .

How many vertices have been moved into  $R_{j,0}$ ? Let us assume without loss of generality that  $h = 0$ . Then this number is by definition

$$(18) \quad \sigma_j - \sigma_0 = \sum_{k=1}^j (a_{k,0} - |R_{k,0}|).$$

Then using our assumption (12) with  $J := [1, j]$  we get

$$\begin{aligned} \frac{\beta n}{8r} &\stackrel{(12)}{\geq} \left| \left( \sum_{k=1}^j a_{k,0} \right) - \frac{n}{2r} j \right| \\ &\stackrel{(18)}{=} \left| \sigma_j - \sigma_0 + \left( \sum_{k=1}^j |R_{k,0}| \right) - \frac{n}{2r} j \right| \geq \sigma_j - \sigma_0 - 2m. \end{aligned}$$

Therefore  $R_{j,0}$  gets at most  $\beta n/(8r) + 2m$  vertices from  $R_{j+1,0}$ , and in the case of  $j = 0$  at most another  $m$  vertices could have been added from  $Q'$ . Analogously, we get the same upper bound on the number of vertices removed from  $R_{j,0}$ .

When  $i = 1$ , we proceed similarly except we move vertices *upward*. That is, we move  $\sigma_f - \sigma_j \geq 0$  vertices from  $R_{j,1}$  to  $R_{j+1,1}$ , where  $f \in \mathbb{Z}_r$  maximizes  $\sigma_f$ .

The sizes of all parts are as required. It remains to show that Conclusion 2 of the lemma holds. The vertices in  $P_0 \cup P_1 \supseteq Q_0 \cup Q_1 \cup Q'$  satisfy it because all even numbers lie inside  $V_0$  and Conclusion 2 is concerned only with crossing edges. So take any  $x \in R_{j,i}$ . Suppose first that this vertex has not been moved (that is,  $x \in D_{2j+i}$ ). We have to show that the non-neighborhood  $\bar{\Gamma}_{S_n}(x) \cap R_{j,1-i}$  of  $x$  into  $R_{j,1-i}$  is small. This non-neighborhood consists of non-neighbors of  $x$  in  $D_{2j+1-i}$  together with at most  $3m + \beta n/(8r)$  extra vertices as calculated above. We have

$$(19) \quad |\bar{\Gamma}_{S_n}(x) \cap R_{j,1-i}| \leq |\bar{\Gamma}_{S_n}(x) \cap D_{2j+1-i}| + \frac{\beta n}{8r} + 3m \stackrel{(17)}{\leq} \frac{\beta n}{4r} + 3m.$$

Next, suppose that  $x$  was one of the moved vertices. Suppose without loss of generality that  $i = 0$ . Then  $x \in D_{2j+2}$ . We have to estimate the degree of  $x$  into  $R_{j,1}$ . This part comes mainly from the residue class  $D_{2j+1}$ . Since the indices  $2j + 2$  and  $2j + 1$  are coprime (this was the reason why we moved vertices *downward* for  $i = 0$ ), Inequality (17) applies again and gives the same bound as in (19).

Finally we keep the name  $R_{j,i}$  to denote the new sets. Inequality (19) now implies the required Conclusion 2 and thus Lemma 6 is proved.

### 7. Proof of Lemma 7

Let a partition  $V(T) = X_0 \cup X_1 \cup X_2 \cup X_3$  be obtained by assigning each component of  $T$  uniformly at random to one of the four possible parts, with all choices being mutually independent. Let  $X_{j,i} := X_j \cap W_i$  and  $x_{j,i} := |X_{j,i}|$ , for  $j \in [0, 3]$  and  $i = 0, 1$ . With a similar calculation as in Lemma 5,

Assumption 2 of Lemma 7 and Hoeffding’s inequality imply that almost surely each  $x_{j,i}$  is at least  $n_i/5$  (and hence at most  $2n_i/5 \leq 4n_{1-i}/5 \leq 4x_{j,1-i}$ ). Fix one such partition.

For  $i = 0, 1$ , let  $L_i := \{x \in X_{i,i} : d_T(x) \leq 7d\}$  and  $l_i := |L_i|$ . By the  $d$ -degeneracy of  $T$ , we have

$$7d \cdot (x_{i,i} - l_i) \leq |E(T[X_i])| \leq d \cdot (x_{i,0} + x_{i,1}) \leq 5dx_{i,i}.$$

Thus  $l_i \geq 2x_{i,i}/7 \geq n_i/20$ .

Let us roughly sketch how we embed  $T$ . First, we greedily embed  $(X_0 \setminus L_0) \cup (X_1 \setminus L_1)$ . Then we define the bipartite compatibility graph  $C$  with one part being  $L_0 \cup L_1$  and the other consisting of all unused vertices of  $G$ . Then we embed  $X_2 \cup X_3$  making sure that the final graph  $C$  has high minimum degree. Finally, we embed each of  $L_0$  and  $L_1$  by a simple application of Hall’s marriage theorem. Note that there are no edges between  $X_0$  and  $X_1$ ; thus  $L_0 \cup L_1$  is an independent set.

Let us formally describe our embedding procedure, which is divided into *Stages 1–3*. By  $f$  we denote the current partial embedding, that is,  $f$  maps a subset  $\text{Dom}(f)$  of  $V(T)$  injectively into  $V(G)$ . Let  $\text{Im}(f) := \{f(x) : x \in \text{Dom}(f)\}$  be the set of vertices of  $G$  to which we have already assigned a vertex of  $T$ .

Since  $T$  is  $d$ -degenerate, we may fix an ordering  $\prec$  of its vertex set  $W_0 \cup W_1$  such that for every  $x \in W_0 \cup W_1$  we have  $|\Gamma_{\prec}(x)| \leq d$ , where we define  $\Gamma_{\prec}(x)$  to be the set of  $T$ -neighbors of  $x$  that precede  $x$ . As we have already mentioned, it will be the case that our procedure saves  $L_0 \cup L_1$  to the very end. This means that, in any previous stage, at least  $l_i \geq n_i/20$  vertices of  $V_i$  are available. Also, when we define  $f$ , we take the vertices of  $X_i \setminus L_i$  using the ordering  $\prec$ . Thus, when we are about to embed a vertex  $x \in W_i$ , the number of possible places for  $f(x)$  is at least

$$(20) \quad \left| \left( \bigcap_{y \in f(\Gamma_{\prec}(x))} \Gamma_G(y) \right) \setminus \text{Im}(f) \right| \geq |L_i| - d\beta n_i \geq \frac{n_i}{20} - d\beta n_i > 0.$$

Here we used the upper bound on the number of non-neighbors of  $y \in f(\Gamma_{\prec}(x))$  given by Assumption 3. Therefore we can always embed  $x$ .

**Stage 1.** Using this observation and the  $d$ -degeneracy of  $T$ , we embed  $(X_0 \setminus L_0) \cup (X_1 \setminus L_1)$  into  $V(G)$  arbitrarily.

After Stage 1, for each  $x \in L_i$  we know the set  $f(\Gamma_T(x)) \subseteq V_{1-i}$  of the  $f$ -images of the  $T$ -neighborhood of  $x$ . We may therefore define

$$A_x := \left( \bigcap_{y \in f(\Gamma_T(x))} \Gamma_G(y) \right) \setminus \text{Im}(f) \subseteq V_i$$

to be the set of all *x-compatible* vertices. Here  $y \in A_x$  means that all neighbors of  $x$  have already been embedded into the neighborhood of  $y$  in  $G$ , so embedding  $x$  to  $y$  would be valid. Let  $C$  be the bipartite *compatibility graph* with parts  $L_0 \cup L_1$  and  $V(G) \setminus \text{Im}(f)$ , where the neighborhood of  $x \in L_0 \cup L_1$  is defined to be  $A_x$ . When we extend  $f$  to a new vertex in Stage 2, we update  $C$ , which amounts to deleting some vertex from the second part  $V(G) \setminus \text{Im}(f)$  of  $C$ .

By Assumption 3, we have for every  $x \in L_i$  that

$$(21) \quad |A_x| \geq |V_i \setminus \text{Im}(f)| - 7d\beta n_i,$$

that is, each  $A_x$  is large. However, it may be possible that some vertices in  $V_i \setminus \text{Im}(f)$  are covered only by few sets  $A_x$ . Let us call  $y \in V_i \setminus \text{Im}(f)$  *bad* if  $|A'_y| < l_i/2$ , where

$$A'_y := \Gamma_C(y) = \{x \in L_i : y \in A_x\}.$$

For  $i = 0, 1$ , let  $B_i$  consist of all bad vertices in  $V_i \setminus \text{Im}(f)$ . By (21), we have

$$\begin{aligned} l_i (|V_i \setminus \text{Im}(f)| - 7d\beta n_i) &\leq |E(C[L_i, V_i \setminus \text{Im}(f)])| \\ &\leq |B_i| \frac{l_i}{2} + (|V_i \setminus \text{Im}(f)| - |B_i|) l_i. \end{aligned}$$

which implies that

$$(22) \quad |B_i| \leq 14d\beta n_i.$$

**Stage 2.** We embed  $X_2 \cup X_3$  so that  $B_0 \cup B_1$  is completely covered.

To implement Stage 2, we describe a randomized algorithm that embeds  $X_2$  and covers the whole of  $B_0$  with positive probability. It follows that at least one such embedding of  $X_2$  exists. We fix it. Then a similar argument (omitted) shows that we can embed  $X_3$  to include all remaining vertices of  $B_1$  (i.e. those that are not covered by  $f(X_2)$ ).

The algorithm proceeds as follows, using the ordering  $\prec$  on  $X_2$ . Given a vertex  $x \in X_{2,i}$  to be embedded, we define

$$A := \left( \bigcap_{y \in f(\Gamma_{\prec}(x))} \Gamma_G(y) \right) \setminus \text{Im}(f)$$

to be the set of available vertices *at this point*. The size of  $A$  is bounded from below by (20). The embedding rule for  $x$  depends on  $i$ , that is, on which part  $W_i$  contains the current vertex  $x$ . If  $x \in X_{2,1}$ , then we let  $f(x)$  be a vertex of  $A$  chosen uniformly at random (and independently of all previous choices). If  $x \in X_{2,0}$  then we select an arbitrary element of  $A$  for  $f(x)$  with the restriction that  $f(x) \in B_0$  whenever  $A \cap B_0 \neq \emptyset$ .

In order to prove that the whole of  $B_0$  can be covered by  $f(X_2)$ , it is enough to show that for any fixed vertex  $y \in B_0$  the probability that it does not belong to  $f(X_2)$  is  $o(1/n_0)$ . Then by linearity of expectation, the expected size of  $B_0 \setminus f(X_2)$  is  $o(1)$ , so almost surely  $B_0 \subseteq f(X_2)$ .

We now fix  $y \in B_0$  and prove the claim about the probability  $\Pr[y \notin f(X_2)]$  by using the following coupling. Let

$$p := \frac{n_1/20 - (d + 1)\beta n_1}{n_1/20 - d\beta n_1} \leq 1.$$

Since e.g.  $\beta \leq \frac{1}{60d}$ , we have  $p \geq 1 - 1/(2d)$ . Let  $(b_x)_{x \in X_{2,1}}$  be the sequence of i.i.d. Bernoulli 0/1-trials indexed by the vertices of  $X_{2,1}$ , where  $\Pr[b_x = 1] = p$ . When embedding a vertex  $x \in X_{2,1}$  into  $V(G)$ , we select  $f(x) \in A$  as follows. Expose  $b_x$ . If  $b_x = 1$ , then we let  $f(x)$  be a random uniform element of  $Y := \Gamma_G(y) \cap A$ . If  $b_x = 0$ , then with probability  $q := (|Y|/|A| - p)/(1 - p)$  we pick a random element of  $Y$ , and with probability  $1 - q$  we pick a random element of  $A \setminus Y$ . Note that  $p \leq |Y|/|A|$  because  $|A| \geq l_1 - \beta d n_1 \geq n_1/20 - \beta d n_1$  and  $|A \setminus Y| \leq \beta n_1$ . Thus  $0 \leq q \leq 1$ . The probability that we pick an element of  $Y$  is  $p + (1 - p)q = |Y|/|A|$ . It follows that  $f(x)$  is uniformly distributed among all elements of  $A$ . What we have achieved is that if  $b_x = 1$ , then  $y$  and  $f(x)$  are necessarily adjacent in  $G$ .

For  $z \in X_{2,0}$ , let the random variable  $b_z$  be 1 if  $b_x = 1$  for every  $x \in \Gamma_{\prec}(z)$  and let  $b_z = 0$  otherwise. Let  $b := \sum_{z \in X_{2,0}} b_z$ .

Let us show that if  $b \geq |B_0|$ , then the vertex  $y$  necessarily belongs to  $f(X_2)$ . Suppose on the contrary that  $y \notin f(X_2)$  after Stage 2. Our coupling implies that, when we embed any  $z \in X_{2,0}$  with  $b_z = 1$ , then for every  $x \in \Gamma_{\prec}(z)$  we have  $b_x = 1$  and  $f(x) \in \Gamma_G(y)$ . Thus this  $z$  is necessarily mapped into a vertex of  $B_0$  because  $z$  is compatible with at least one unused vertex of  $B_0$ , namely with  $y \in B_0$ . But then  $f(z) \in B_0 \setminus \{y\}$  for at least  $|B_0|$  distinct vertices  $z$ , a contradiction.

The definition of the ordering  $\prec$  implies that the expectation  $\mathbf{E}[b]$  is at least  $|X_{2,0}|p^d \geq |X_{2,0}|/2 \geq 2|B_0|$  by (22). (Recall that  $p^d \geq (1 - 1/(2d))^d \geq 1/2$ .) Also, each  $b_z$  with  $z \in X_{2,0}$  is determined by random bits  $b_x$  with  $x \in \Gamma_{\prec}(z)$ . Hence, we can define a dependency graph by connecting distinct

$z, z' \in X_{2,0}$  if and only if  $\Gamma_{\prec}(z) \cap \Gamma_{\prec}(z') \neq \emptyset$ ; let us denote this by  $z \sim z'$ . Using  $\Delta(T[X_2])$  to denote the maximum degree of  $T[X_2]$ , we let

$$\Delta := \sum_{z \sim z'} \Pr[b_z = b_{z'} = 1] \leq \sum_{z \sim z'} 1 \leq |X_{2,0}| \cdot d \cdot \Delta(T[X_2]) \leq n_0 \cdot d \cdot \frac{\beta n_1}{\ln n_1}.$$

Now, a large deviation result of Janson [18, Theorem 1] (which is Theorem 8.7.2 in [5]) implies that by  $\mathbf{E}[b] \geq |X_{2,0}|/2 \geq n_0/10$  we have

$$\begin{aligned} \Pr[b < |B_0|] &\leq \Pr[b \leq \mathbf{E}[b]/2] \leq e^{-\frac{\mathbf{E}[b]}{4(2+\Delta/\mathbf{E}[b])}} \\ &\leq e^{-\frac{n_0/10}{4(2+10d\beta n_1/\ln n_1)}} \leq e^{-\frac{\ln n_0}{850d\beta}} = o(n_0^{-1}). \end{aligned}$$

(The last inequality used Assumption 1 of Lemma 7, and our choice of  $\beta$  was determined by this computation.) This is the required upper bound on the probability that  $y$  remains uncovered. Thus there is an embedding of  $X_2$  with  $f(X_2) \supseteq B_0$ . Likewise, we can extend  $f$  to  $X_3$  and cover all of  $B_1$ .

**Stage 3.** For  $i = 0, 1$ , we embed  $L_i$ . By (21), each  $x \in L_i$  has  $|L_i| - 7d\beta n_i \geq |L_i|/2$  neighbors in the current compatibility graph  $C$ . Since  $B_0 \cup B_1 \subseteq \text{Im}(f)$ , by definition of  $B_0$  and  $B_1$  we know that each  $y \in V_i \setminus \text{Im}(f)$  has at least  $|L_i|/2$  neighbors in  $C$ . Therefore, when restricted to  $L_i \cup (V_i \setminus f(W_i))$ , the graph  $C$  has minimum degree at least  $|L_i|/2$  on both sides. It easily follows from Hall’s marriage theorem [16] that we have a perfect matching in  $C$ , which shows how to extend  $f$  to  $L_i$ .

This completes the proof of Lemma 7.

### 8. Sketch of Proof of Theorem 4

It is in principle possible to strengthen Lemmas 5–7 so that they formally imply Theorem 4 as well. However, this would make their statements and proofs messier and, we believe, would not justify the gain. Therefore, we give just a high-level sketch of how one has to modify the proofs.

Let us first do the case  $n = 2k$  of Theorem 4, which is slightly easier. Let  $n$  be large and let  $H$  be a connected graph with  $n$  vertices and  $n$  edges. We have to show that  $H$  is prime. Pick an arbitrary edge  $xy$  on a cycle. Let  $W_0 \cup W_1 = V(T)$  be a bipartition of the spanning tree  $T := H - xy$  with  $|W_0| \geq k$ . Apply the proof of Lemma 5 to  $T$ .

Suppose first that  $x$  belongs to  $W_1$ . We can additionally require that  $x \in M_1$  (by, perhaps, increasing  $m$  by 1). Then we follow the proof of Theorem 2 to obtain a prime labeling of  $T$ . Additionally, we make sure that

$x$  gets a label from  $P_1$ ; it is enough to require that we do not change the label at  $x$  when we fix conflicting pairs  $p$  and  $2p \in P_0$  at the end of the proof of Theorem 2.

Likewise, we are done if  $y \in W_1$  or if  $|W_0| = |W_1|$  (when we just swap  $W_0$  and  $W_1$  in the proof of Lemma 5 and proceed as above). Hence, let us assume that  $|W_0| \geq k + 1$  and  $x, y \in W_0$ . Move  $x$  to  $W_1$  and assign  $x$  to  $M_1$ . Proceed with the proof of Theorem 2, making sure again that  $x$  gets a label from  $P_1$ . The edges between  $x$  and  $W_1$  cannot spoil anything. Thus  $H$  is prime, as required.

Let us do the case of odd  $n = 2k + 1$ . Let  $H$  be a connected graph with  $n$  vertices and  $n + 1$  edges. Let  $T := G - wx - yz$  be a spanning tree. Let  $W_0 \cup W_1$  be a bipartition of  $T$  with  $|W_0| > |W_1|$ . As before, we are done if both  $wx$  and  $yz$  intersect  $W_1$ . So suppose that, for example,  $w, x \in W_0$ .

Since  $|W_0| \geq k + 1$ , we can additionally require in Lemma 5 that  $|X| \geq k + 1$ , where  $X$  is defined by (14). (Here we have to change the proof of Lemma 5 slightly: instead of Condition (16) we use  $\sum_{j=1}^c \max(|A'_{j,0}|, |A'_{j,1}|) \geq k + 1$  with the obvious modifications of the rest of the proof.)

Suppose first that  $yz$  intersects  $W_1$ , say  $y \in W_1$ . Make sure that  $y \in M_1$  and that  $y$  gets a label from  $P_1$ . In the proof of Lemma 6, we have  $|Q'| \geq |X| - \lfloor n/2 \rfloor \geq 1$ . We can use a label from  $Q' \subseteq P_1$  on one of  $w$  or  $x$ , thus eliminating any conflict between  $x$  and  $w$ .

Finally, it remains to assume that  $w, x, y, z \in W_0$ . Similarly to above, if  $|Q'| \geq 2$ , then we have at least 2 elements from  $P_1$  which we can use on  $W_0$  and eliminate conflicts on  $wx$  and  $yz$ . Thus we are done unless  $|Q'| = 1$ . In this case  $|W_0| = k + 1$  and  $|W_1| = k$ . Also, note that  $W_1$  is an independent set. Thus we can swap  $W_0$  and  $W_1$  and apply the proof of Lemma 5 to  $T$  (where we do need  $\lfloor n/2 \rfloor$  in the right-hand side of (16)). Everything works as before, provided we make sure that  $w, y$  are assigned to  $M_1$  and get labels from  $P_1$  in the final labeling.

This finishes the sketch of the proof of Theorem 4.

## References

- [1] Ahlswede, R., and Khachatrian, L. H. (1994). On extremal sets without coprimes. *Acta Arith.* **66** 89–99. [MR1262655](#)
- [2] Ahlswede, R., and Khachatrian, L. H. (1995). Maximal sets of numbers not containing  $k + 1$  pairwise coprime integers. *Acta Arith.* **72** 77–100. [MR1346807](#)

- [3] Ahlswede, R., and Khachatrian, L. H. (1996). Sets of integers and quasi-integers with pairwise common divisor. *Acta Arith.* **74** 141–153. [MR1373704](#)
- [4] Alon, N., Seymour, P., and Thomas, R. (1990). A separator theorem for nonplanar graphs. *J. Amer. Math. Soc.* **3** 801–808. [MR1065053](#)
- [5] Alon, N., and Spencer, J. (2008). *The Probabilistic Method*, 3d ed. Wiley Interscience. [MR2437651](#)
- [6] Dusart, P. (1999). The  $k$ th prime is greater than  $k(\ln k + \ln \ln k - 1)$  for  $k \geq 2$ . *Math. Comp.* **68** 411–415. [MR1620223](#)
- [7] Erdős, P. (1962). Remarks on number theory. IV. Extremal problems in number theory. I. *Mat. Lapok* **13** 228–255. [MR0195822](#)
- [8] Erdős, P. (1980). A survey of problems in combinatorial number theory. *Ann. Discrete Math.* **6** 89–115. [MR0593525](#)
- [9] Erdős, P. (1998). Some of my new and almost new problems and results in combinatorial number theory. In: *Number Theory (Eger, 1996)*. de Gruyter, Berlin, pp. 169–180. [MR1628841](#)
- [10] Erdős, P., Sárközy, A., and Szemerédi, E. (1969). On some extremal properties of sequences of integers. *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.* **12** 131–135. [MR0260702](#)
- [11] Erdős, P., and Sárközy, A. (1993). On sets of coprime integers in intervals. *Hardy-Ramanujan J.* **16** 1–20. [MR1216724](#)
- [12] Erdős, P., Sárközy, A., and Szemerédi, E. (1980). On some extremal properties of sequences of integers. II. *Publ. Math. Debrecen* **27** 117–125. [MR0592189](#)
- [13] Erdős, P., and Sárközy, G. N. (1997). On cycles in the coprime graph of integers. *Electronic J. Combin.* **4**, 11 pp. The Wilf Festschrift (Philadelphia, PA, 1996). [MR1444155](#)
- [14] Fu, H.-L., and Huang, K.-C. (1994). On prime labellings. *Discrete Math.* **127** 181–186. [MR1273601](#)
- [15] Gallian, J. A. (2011). A dynamic survey of graph labeling. *Electronic J. Combin.* **DS6**, 256 pp.
- [16] Hall, P. (1935). On representatives of subsets. *J. Lond. Math. Soc.* **10** 26–30.
- [17] Hoeffding, W. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58** 13–30. [MR0144363](#)

- [18] Janson, S. (1990). Poisson approximation for large deviations. *Random Struct. Algorithms* **1** 221–229. [MR1138428](#)
- [19] Komlós, J., Sárközy, G. N., and Szemerédi, E. (1997). Blow-up lemma. *Combinatorica* **17** 109–123. [MR1466579](#)
- [20] Kostochka, A. V. (1982). A lower bound for the Hadwiger number of a graph as a function of the average degree and its vertices (in Russian). *Diskret. Analiz. Novosibirsk* **38** 37–58. [MR0713722](#)
- [21] Kuo, J., and Fu, H.-L. (2010). Modified prime labelling on small trees. Manuscript.
- [22] Lin, S.-H. (1999). *A Study of Prime Labeling*. M.S. Thesis, National Chiao Tung University, Taiwan.
- [23] Lipton, R. J., and Tarjan, R. E. (1979). A separator theorem for planar graphs. *SIAM J. Appl. Math.* **36** 177–189. [MR0524495](#)
- [24] Mader, W. (1968). Homomorphiesätze für Graphen. *Mathematische Annalen* **178** 154–168. [MR0229550](#)
- [25] Pikhurko, O. (2002). Every tree with at most 34 vertices is prime. *Utilitas Math.* **62** 185–190. [MR1941385](#)
- [26] Pikhurko, O. (2007). Trees are almost prime. *Discrete Math.* **307** 1455–1462. [MR2311118](#)
- [27] Rao, S. N. (2002). Prime labelling. In: *R. C. Bose Centenary Symposium on Discrete Math. and Applications, Kolkata*.
- [28] Sárközy, G. N. (1999). Complete tripartite subgraphs in the coprime graph of integers. *Discrete Math.* **202** 227–238. [MR1694477](#)
- [29] Spiess, J. (2010). Embedding of spanning trees into special classes of graphs. Bachelor’s thesis, TU München.
- [30] Szabó, C., and Tóth, G. (1985). Maximal sequences not containing four pairwise coprime integers. *Mat. Lapok* **32** 253–257. [MR0866222](#)
- [31] Thomason, A. G. (1984). An extremal function for contractions of graphs. *Math. Proc. Camb. Phil. Soc.* **95** 261–265. [MR0735367](#)
- [32] Thomason, A. G. (2001). The extremal function for complete minors. *J. Combin. Theory (B)* **81** 318–338. [MR1814910](#)
- [33] Tout, A., Dabboucy, A. N., and Howalla, K. (1982). Prime labeling of graphs. *Nat. Acad. Sci. Letters* **11** 365–368.

PENNY HAXELL  
DEPARTMENT OF COMBINATORICS AND OPTIMIZATION  
UNIVERSITY OF WATERLOO  
WATERLOO, ONTARIO N2L 3G1  
CANADA

OLEG PIKHURKO  
DEPARTMENT OF MATHEMATICAL SCIENCES  
CARNEGIE MELLON UNIVERSITY  
PITTSBURGH, PA 15213  
USA

ANUSCH TARAZ  
ZENTRUM MATHEMATIK  
TECHNISCHE UNIVERSITÄT MÜNCHEN  
D-85747 GARCHING BEI MÜNCHEN  
GERMANY

RECEIVED SEPTEMBER 21, 2010