

ASYMMETRIC RAMSEY PROPERTIES OF RANDOM GRAPHS INVOLVING CLIQUES

MARTIN MARCINISZYN, JOZEF SKOKAN, RETO SPÖHEL, AND ANGELIKA STEGER

CDAM RESEARCH REPORT LSE-CDAM-2007-02

DATE: JANUARY 18, 2007

ABSTRACT. Consider the following problem: For given graphs G and F_1, \dots, F_k , find a coloring of the edges of G with k colors such that G does not contain F_i in color i . Rödl and Ruciński studied this problem for the random graph $G_{n,p}$ in the symmetric case when k is fixed and $F_1 = \dots = F_k = F$. They proved that such a coloring exists asymptotically almost surely (a.a.s.) provided that $p \leq bn^{-\beta}$ for some constants $b = b(F, k)$ and $\beta = \beta(F)$. This result is essentially best possible because for $p \geq Bn^{-\beta}$, where $B = B(F, k)$ is a large constant, such an edge-coloring does not exist. Kohayakawa and Kreuter conjectured a threshold function $n^{-\beta(F_1, \dots, F_k)}$ for arbitrary F_1, \dots, F_k .

In this paper we address the case when F_1, \dots, F_k are cliques of different sizes and propose an algorithm that a.a.s. finds a valid k -edge-coloring of $G_{n,p}$ with $p \leq bn^{-\beta}$ for some constant $b = b(F_1, \dots, F_k)$. This algorithm also extends to the symmetric case. We also show that there exists a constant $B = B(F_1, \dots, F_k)$ such that for $p \geq Bn^{-\beta}$ the random graph $G_{n,p}$ a.a.s. does not have a valid k -edge-coloring provided the so-called KLR-conjecture holds.

CONTENTS

1. Introduction	1
1.1. Our results	3
1.2. Organization of this paper	4
2. Preliminaries and Notation	4
3. An algorithm for computing valid edge colorings	5
3.1. Proof of Claim 15	13
3.2. Proof of Lemma 8	22
4. The 1-statement	27
References	29

1. INTRODUCTION

It follows from Ramsey's celebrated result [16] that *every* k -coloring of the edges of the complete graph on n vertices contains a monochromatic copy of F if n is sufficiently large. While this seems to rely on the fact that K_n is a very dense graph, Folkman [2] and, in a more general

The first author was supported by FAPESP (Proj. Temático-ProNEx Proc. FAPESP 2003/09925-5) while visiting the University of São Paulo.

The second author was partially supported by NSF grants INT-0072064, INT-0305793, NSA grant H98230-04-1-0035, and by FAPESP/CNPq (Proj. Temático-ProNEx Proc. FAPESP 2003/09925-5 and Proc. FAPESP 2004/15397-4). Part of this work was done while this author was visiting ETH Zurich.

The third author was supported by SNF grant 200021-108158.

An extended abstract of the paper appeared in the proceedings of the 10th International Workshop on Randomized Techniques in Computation (RANDOM'06) [14].

setting, Nešetřil and Rödl [15] showed that there also exist locally sparse graphs $G = G(F)$ with the property that every k -coloring of the edges of G contains a monochromatic copy of F . By transferring the problem into a random setting, Rödl and Ruciński [18] showed that in fact such graphs G are quite frequent. More precisely, they proved the following result. Let

$$G \rightarrow (F)_k^e$$

denote the property that *every* edge-coloring of G with k colors contains a monochromatic copy of F . Recall that in the binomial random graph $G_{n,p}$ on n vertices, every edge is present with probability $0 \leq p = p(n) \leq 1$ independently of all other edges.

Theorem 1 ([13], [17], [18]). *Let $k \geq 2$ and F be a non-empty graph that is not a forest. Then there exist constants $b, B > 0$ such that*

$$\lim_{n \rightarrow \infty} \mathbb{P}[G_{n,p} \rightarrow (F)_k^e] = \begin{cases} 0 & \text{if } p \leq bn^{-1/m_2(F)} \\ 1 & \text{if } p \geq Bn^{-1/m_2(F)} \end{cases},$$

where

$$m_2(F) := \max \left\{ \frac{|E(H)| - 1}{|V(H)| - 2} : H \subseteq F \wedge |V(H)| \geq 3 \right\}. \quad (1)$$

Functions like $n^{-1/m_2(F)}$ in Theorem 1 are called thresholds or threshold functions. In Theorem 1, this function can be motivated as follows. For the sake of simplicity, suppose that $m_2(F) = (|E(F)| - 1)/(|V(F)| - 2)$. Then, for $p = cn^{-1/m_2(F)}$, the expected number of copies of F containing a given edge of $G_{n,p}$ is a constant depending on c . If this constant is close to zero, the copies of F in $G_{n,p}$ are loosely scattered and a valid coloring should thus exist. On the other hand, if this constant is large, the copies of F in $G_{n,p}$ highly intersect with each other, and the existence of a valid coloring becomes unlikely.

In Theorem 1 the same graph F is forbidden in every color class. We can generalize this setup by allowing for k different forbidden graphs, one per color. Within classical Ramsey theory the study of these so-called asymmetric Ramsey properties led to many interesting questions (see e.g. [1]) and results, most notably the celebrated paper of Kim [7] where he established an asymptotically sharp bound on the Ramsey number $R(3, t)$.

Within the random setting only very little is known about asymmetric Ramsey properties. Let

$$G \rightarrow (F_1, \dots, F_k)^e$$

denote the property that in *every* edge-coloring of G with k colors, there exists a color i such that F_i is contained in the subgraph of G spanned by the edges which are assigned to i . In [9] Kohayakawa and Kreuter proved the following result for cycles C_ℓ of length ℓ .

Theorem 2 ([9]). *Let $k \geq 2$ and $3 \leq \ell_1 \leq \dots \leq \ell_k$ be integers. Then there exist constants $b, B > 0$ such that*

$$\lim_{n \rightarrow \infty} \mathbb{P}[G_{n,p} \rightarrow (C_{\ell_1}, \dots, C_{\ell_k})^e] = \begin{cases} 0 & \text{if } p \leq bn^{-1/m_2(C_{\ell_2}, C_{\ell_1})} \\ 1 & \text{if } p \geq Bn^{-1/m_2(C_{\ell_2}, C_{\ell_1})} \end{cases},$$

where

$$m_2(C_{\ell_2}, C_{\ell_1}) := \frac{\ell_1}{\ell_1 - 2 + 1/m_2(C_{\ell_2})}.$$

In [9] Kohayakawa and Kreuter also formulated the following conjecture that generalizes their result to general graphs.

Conjecture 3 ([9]). *Let H and F be graphs with $1 < m_2(H) \leq m_2(F)$. Then there exists a constant $b > 0$ such that for all $\varepsilon > 0$, we have*

$$\lim_{n \rightarrow \infty} \mathbb{P}[G_{n,p} \rightarrow (H, F)^e] = \begin{cases} 0 & \text{if } p \leq (1 - \varepsilon)bn^{-1/m_2(H, F)} \\ 1 & \text{if } p \geq (1 + \varepsilon)bn^{-1/m_2(H, F)} \end{cases},$$

where

$$m_2(H, F) := \max \left\{ \frac{|E(J)|}{|V(J)| - 2 + 1/m_2(H)} : J \subseteq F \wedge |V(F)| \geq 2 \right\}. \quad (2)$$

The threshold function in Conjecture 3 can be justified as follows. The expected number of copies of F in $G_{n,p}$ with $p = \Theta(n^{-1/m_2(F,H)})$ is

$$\Theta \left(n^{|V(F)|} p^{|E(F)|} \right) = \Omega \left(n^{2-1/m_2(H)} \right).$$

Since every edge-coloring of $G_{n,p}$ must avoid monochromatic copies of F in color 2, there is at least one edge of color 1 in every subgraph of $G_{n,p}$ isomorphic to F . Select one such edge from each copy of F arbitrarily. It is plausible that these edges span a graph G' with edge density $\Omega(n^{-1/m_2(H)})$ that satisfies certain pseudo-random properties. As it turns out, that seems just about the right density in order to embed a copy of H into G' , no matter which edges were selected from the original graph.

We remark that Ramsey properties of random graphs were first studied with respect to vertex colorings in [13]. In the asymmetric setup with respect to vertex colorings, very general results were obtained in [12]. To the best of our knowledge, besides Theorem 2, no progress has been made towards proving Conjecture 3 since its publication in [9].

1.1. Our results

We study Conjecture 3 in the case when all graphs are cliques and prove a threshold result in the flavor of Theorem 2.

Theorem 4 (Main Result). *Let $k \geq 2$ and $3 \leq \ell_k \leq \dots \leq \ell_1$ be integers. Then there exist constants $b, B > 0$ such that*

$$\lim_{n \rightarrow \infty} \mathbb{P}[G_{n,p} \rightarrow (K_{\ell_1}, \dots, K_{\ell_k})^c] = \begin{cases} 0 & \text{if } p \leq bn^{-1/m_2(K_{\ell_2}, K_{\ell_1})} \\ 1 & \text{if } p \geq Bn^{-1/m_2(K_{\ell_2}, K_{\ell_1})} \end{cases},$$

where

$$m_2(K_{\ell_2}, K_{\ell_1}) := \frac{\binom{\ell_1}{2}}{\ell_1 - 2 + 1/m_2(K_{\ell_2})},$$

and the 1-statement holds provided Conjecture 23 in [10] is true for K_{ℓ_2} .

As is typical for threshold phenomena, the result consists of two separate statements, the so-called 0- and 1-statement, respectively. For the 0-statement one needs to show that for p below the threshold, $G_{n,p}$ allows a k -edge-coloring that avoids every forbidden clique K_{ℓ_i} , $1 \leq i \leq k$, in the corresponding color class i . For the 1-statement one has to show that for p above the threshold, every k -edge-coloring of $G_{n,p}$ contains at least one of the forbidden monochromatic cliques.

The main contribution of this paper is the proof of the 0-statement. The way of proving it that was pursued in [17] and [9] is by contradiction. This approach shows the existence of a coloring, but provides no efficient way of obtaining the coloring from the proof. Our approach is constructive. We provide a (polynomial-time) algorithm that computes a valid coloring for graphs that satisfy certain properties. We employ techniques similar to those in [17] and [9] in order to prove that these properties a.a.s. hold in $G_{n,p}$ with p sufficiently small. Indeed, the results in [17] yield that our algorithm also computes valid colorings of $G_{n,p}$ in the symmetric case, unless the forbidden graph is one of a few special cases, e.g., a triangle. In fact, the symmetric case of triangles was solved in [13] by different methods.

A standard way of attacking the 1-statement, which was also pursued in [9], is via the sparse version of Szemerédi's regularity lemma, which was independently developed by Kohayakawa [8] and Rödl (unpublished, see [11]). Using properties of regularity, one can find a monochromatic copy of a forbidden subgraph in the colored graph $G_{n,p}$. Unfortunately, generalizing this argument from cycles to cliques requires a proof of Conjecture 23 in [10] (cf. Conjecture 30 below)

of Kohayakawa, Łuczak, and Rödl. This so-called KLR-Conjecture formulates a probabilistic version of the classical embedding lemma for dense graphs. It implies many interesting extremal results for random graphs. Despite recent progress [4], the conjecture is, in its full generality, still wide open. As we shall show, a proof of the 1-statement is routinely obtained assuming that the KLR-Conjecture holds. For a comprehensive overview on the status quo of that conjecture, the interested reader is referred to [5].

1.2. Organization of this paper

We start by collecting some definitions and useful facts in Section 2. Section 3 is the main body of our work. We present an algorithm ASYM-EDGE-COL and prove that it a.a.s. computes a valid coloring of $G_{n,p}$ with p as claimed. The proofs of two major technical lemmas are omitted from the main line of argument and deferred to Sections 3.1 and 3.2, respectively. A considerable amount of technicalities in this section is specifically needed to deal with the case $\ell_2 = 3$. For a simpler account of the case $\ell_2 \geq 4$, the reader is referred to [14]. Lastly, in Section 4, we explain how the 1-statement follows from the results in [9] and the KLR-Conjecture.

2. PRELIMINARIES AND NOTATION

Our notation is mostly adopted from [6]. All graphs are simple and undirected. Let $G = (V, E)$ be a graph. We denote the number of vertices in G by $v(G) = v_G := |V(G)|$. Similarly, we denote the number of edges in G by $e(G) = e_G := |E(G)|$. For disjoint sets $U, W \subseteq V(G)$, we denote the set of all edges having one end in U and the other one in W by $E_G(U, W)$ and abbreviate $|E_G(U, W)|$ by $e_G(U, W)$. The neighborhood of a vertex v is denoted by $\Gamma(v)$ and its degree by $\deg(v) := |\Gamma(v)|$.

For any graph H , the most well-known density measure is

$$d(H) := \begin{cases} e_H/v_H & \text{if } v(H) \geq 1 \\ 0 & \text{otherwise} \end{cases},$$

that is, exactly the half of the average degree of H . Maximizing over all subgraphs of H defines the measure

$$m(H) := \max \{d(J) : J \subseteq H\}.$$

A graph H is *balanced w.r.t. d* or simply *balanced* if we have $d(H) = m(H)$. It is *strictly balanced* if for all proper subgraphs $H' \subsetneq H$, we have $d(H') < m(H)$. In [17] the so-called 2-density measure was introduced:

$$d_2(H) := \begin{cases} (e_H - 1)/(v_H - 2) & \text{if } H \text{ is nonempty with } v(H) \geq 3 \\ 1/2 & \text{if } H \cong K_2 \\ 0 & \text{otherwise} \end{cases}.$$

Analogously to the standard density d , the maximization over the 2-density of all subgraphs $J \subseteq H$ is denoted by $m_2(H)$ (cf. (1)). The notion of (*strict*) *2-balancedness* is defined similarly. It is easily verified that every graph satisfies $m(H) \leq m_2(H)$, and that every 2-balanced graph is also balanced.

In the context of asymmetric Ramsey properties, the following generalization of d_2 with two arguments was introduced in [9]. Let H and F be any graphs, and define

$$d_2(H, F) := \begin{cases} e_F/(v_F - 2 + 1/m_2(H)) & \text{if } H \text{ is nonempty and } v(F) \geq 2 \\ 0 & \text{otherwise} \end{cases}.$$

As in (2), it is useful to denote the maximization of this generalized 2-density over all subgraphs $J \subseteq F$ with respect to some graph H by $m_2(H, F)$. The graph F is *balanced w.r.t. $d_2(H, \cdot)$* if we have $d_2(H, F) = m_2(H, F)$. It is readily verified that every clique K_r is balanced w.r.t. $d_2(K_\ell, \cdot)$ for all $r \geq \ell \geq 3$ and, in particular, that every clique is 2-balanced.

The function m_2 is not symmetric in both arguments. Although we shall not use it, the following statement sheds more light on the asymmetric 2-density. It is routinely proved using Proposition 6 below.

Proposition 5. *Suppose H and F are graphs with $m_2(H) \leq m_2(F)$. Then we have*

$$m_2(H) \leq m_2(H, F) \leq m_2(F) ,$$

which implies that $m_2(H) = m_2(H, F) = m_2(F)$ whenever $m_2(H) = m_2(F)$. Moreover, if F is 2-balanced, it is also balanced w.r.t. $d_2(H, \cdot)$, but the converse does not hold in general.

All our logarithms have base e . In our proofs we frequently use the following observation, which we state separately for further reference.

Proposition 6. *For $a, c, C \in \mathbb{R}$ and $b, d > 0$, we have*

$$\frac{a}{b} \leq C \wedge \frac{c}{d} \leq C \Rightarrow \frac{a+c}{b+d} \leq C \quad \text{and} \quad \frac{a}{b} \geq C \wedge \frac{c}{d} \geq C \Rightarrow \frac{a+c}{b+d} \geq C$$

and similarly, if also $b > d$,

$$\frac{a}{b} \leq C \wedge \frac{c}{d} \geq C \Rightarrow \frac{a-c}{b-d} \leq C \quad \text{and} \quad \frac{a}{b} \geq C \wedge \frac{c}{d} \leq C \Rightarrow \frac{a-c}{b-d} \geq C .$$

3. AN ALGORITHM FOR COMPUTING VALID EDGE COLORINGS

Suppose $G = G_{n,p}$ with $p \leq bn^{-1/m_2(K_{\ell_2}, K_{\ell_1})}$ is given. In order to provide a valid coloring of G , it suffices to compute a 2-coloring of $E(G)$ such that there is neither a copy of K_{ℓ_1} in color 1 nor a copy of K_{ℓ_2} in color 2. That implies the 0-statement of Theorem 4 also for k -colorings. Hence, we focus on 2-colorings and abbreviate ℓ_1 by r and ℓ_2 by ℓ in the following. For the sake of completeness, we include the symmetric cases $r = \ell \geq 4$ in our setup. For the rest of this section, suppose $r \geq \ell \geq 3$, $(r, \ell) \neq (3, 3)$ are fixed. As mentioned before, our approach does not carry over to the symmetric triangle case. We also fix the constant $\varepsilon_0 := 0.01$ in what follows.

We describe an algorithm that finds a valid edge-coloring of G a.a.s. The basic idea of the algorithm is to remove edges from the graph successively. An edge e is deleted from G if there are no two cliques of size ℓ and r respectively that intersect exactly on e . When this deletion process stops, the remaining graph G^* a.a.s. satisfies strong structural properties and can be colored easily by a local argument. Any coloring of G^* can then be extended to a coloring of G by inserting the removed edges in the reverse order one by one, always assigning a valid color instantly. The actual algorithm is more complex since sometimes one has to ‘forget’ about the existence of certain cliques in order to remove enough from G . As we shall see, we can easily deal with those cliques later.

In order to simplify notation, for any graph G we define the families

$$\mathcal{L}_G := \{L \subseteq G : L \cong K_\ell\} \quad \text{and} \quad \mathcal{R}_G := \{R \subseteq G : R \cong K_r\}$$

of all ℓ -cliques and r -cliques in G respectively. Furthermore, we introduce the family

$$\mathcal{L}_G^* := \{L \in \mathcal{L}_G : \forall e \in E(L) \exists R \in \mathcal{R}_G \text{ s.t. } E(L) \cap E(R) = \{e\}\} \subseteq \mathcal{L}_G .$$

We define families \mathcal{C} and $\mathcal{A} \subseteq \mathcal{C}$ of graphs as follows. Recall that $\varepsilon_0 = 0.01$.

Definition 7. For $r \geq \ell \geq 3$, let

$$\mathcal{C} = \mathcal{C}(\ell, r) := \{G = (V, E) : \forall e \in E(G) \exists (L, R) \in \mathcal{L}_G \times \mathcal{R}_G \text{ s.t. } E(L) \cap E(R) = \{e\}\}$$

and

$$\mathcal{A} = \mathcal{A}(\ell, r) := \{A \in \mathcal{C}(\ell, r) : m(A) \leq m_2(K_\ell, K_r) + \varepsilon_0 \wedge A \text{ is 2-connected}\} .$$

Intuitively, the graphs in \mathcal{A} are the building blocks of the graphs G^* which may remain after the edge deletion process. We shall see that these families are very small (in fact, they are empty for $r > \ell \geq 4$). Finding their explicit representation is quite technical and deferred to Section 3.2. However, this representation immediately implies the following Lemma, which is all we need for the proof of Theorem 4.

Lemma 8. *For every $r \geq \ell \geq 3$, $(r, \ell) \neq (3, 3)$, the family $\mathcal{A}(\ell, r)$ is finite, and all its members permit a valid edge coloring avoiding K_ℓ in one and K_r in the other color.*

Remark 9. Lemma 8 is false for $r = \ell = 3$. This is the reason why our approach does not carry over to the symmetric triangle case.

For any given graph G , the family \mathcal{S}_G identifies all maximal subgraphs of G isomorphic to a member of \mathcal{A} , i.e.,

$$\mathcal{S}_G := \{S \subseteq G : S \cong A \in \mathcal{A} \wedge \nexists S' \supsetneq S \text{ with } S' \cong A' \in \mathcal{A}\} .$$

Note that there are no two members $S_1, S_2 \in \mathcal{S}_G$ such that $S_1 \subsetneq S_2$. For any edge $e \in E(G)$, let

$$\mathcal{S}_G(e) := \{S \in \mathcal{S}_G : e \in E(S)\} .$$

We call G an \mathcal{A} -graph if, for all $e \in E(G)$, we have

$$|\mathcal{S}_G(e)| = 1 .$$

In particular, an \mathcal{A} -graph is an edge-disjoint union of graphs from \mathcal{A} .

In an \mathcal{A} -graph G , there are two types of triangles: triangles that are subgraphs of an $S \in \mathcal{S}_G$, and triangles that have edges in at least two different graphs \mathcal{S}_G . We call the former *trivial triangles* and define

$$\mathcal{T}_G := \left\{ T \subseteq G : T \cong K_3 \wedge \left| \bigcup_{e \in E(T)} \mathcal{S}_G(e) \right| \geq 2 \right\}$$

as the family of all non-trivial triangles in G . We call an \mathcal{A} -graph G *triangle-sparse* if $\mathcal{T}_G = \emptyset$. The next lemma asserts that triangle-sparse \mathcal{A} -graphs are easily colorable.

Lemma 10. *There exists a procedure A-COLOR that returns for any triangle-sparse \mathcal{A} -graph a valid edge coloring.*

Proof. Lemma 8 asserts that there exists a valid edge coloring of every $A \in \mathcal{A}$. Using this we define the procedure A-COLOR(G) as follows: Assign a valid coloring to every subgraph $S \in \mathcal{S}_G$ locally, that is, regardless of the remaining structure of G . Since G is an \mathcal{A} -graph, we thus assign a color to all edges of G without creating conflicts. As G is triangle-sparse, this procedure cannot produce non-trivial monochromatic triangles, let alone larger monochromatic cliques with edges in different subgraphs $S \in \mathcal{S}_G$. \square

The algorithm ASYM-EDGE-COL is given in Figure 1. Note that edges are removed from and inserted into a working copy $G' = (V, E')$ of G . The local variable \mathcal{L} is identical to the family $\mathcal{L}_{G'}$ until the first execution of line 13. Generally, we have $\mathcal{L} \subseteq \mathcal{L}_{G'}$.

Lemma 11. *Algorithm ASYM-EDGE-COL either terminates with an error in line 14 or finds a valid edge coloring of G .*

Proof. Let G^* denote the argument in the call to A-COLOR in line 15. By Lemma 10, there is an edge coloring of G^* with neither a blue ℓ -clique nor a red r -clique. It remains to show that no forbidden monochromatic cliques are created when this coloring is extended to a coloring of G in lines 16–24.

First, we argue that the algorithm never creates a blue copy of K_ℓ . Observe that *every* copy of K_ℓ that does not lie entirely in G^* is pushed on the stack in the first loop. Therefore, in the execution of the second loop, the algorithm checks the coloring of every such copy. By the order

```

ASYM-EDGE-COL( $G = (V, E)$ )
1   $s \leftarrow \text{EMPTY-STACK}()$ 
2   $E' \leftarrow E$ 
3   $\mathcal{L} \leftarrow \mathcal{L}_G$ 
4  while  $G' = (V, E')$  is no triangle-sparse  $\mathcal{A}$ -graph
5  do if  $\exists e \in E'$  s.t.  $\nexists (L, R) \in \mathcal{L} \times \mathcal{R}_{G'=(V, E')} : E(L) \cap E(R) = \{e\}$ 
6      then for each  $L \in \mathcal{L} : e \in E(L)$ 
7          do  $s.\text{PUSH}(L)$ 
8               $\mathcal{L}.\text{REMOVE}(L)$ 
9               $s.\text{PUSH}(e)$ 
10              $E'.\text{REMOVE}(e)$ 
11         else if  $\exists L \in \mathcal{L} \setminus \mathcal{L}_{G'=(V, E')}^*$ 
12             then  $s.\text{PUSH}(L)$ 
13                  $\mathcal{L}.\text{REMOVE}(L)$ 
14         else error “stuck”
15  A-COLOR( $G' = (V, E')$ )
16  while  $s \neq \emptyset$ 
17  do if  $s.\text{TOP}()$  is an edge
18      then  $e \leftarrow s.\text{POP}()$ 
19           $e.\text{SET-COLOR}(\text{blue})$ 
20           $E'.\text{ADD}(e)$ 
21      else  $L \leftarrow s.\text{POP}()$ 
22          if  $L$  is entirely blue
23              then  $f \leftarrow$  any  $e \in E(L)$  s.t.  $\nexists R \in \mathcal{R}_{G'=(V, E')} : E(L) \cap E(R) = \{e\}$ 
24                   $f.\text{SET-COLOR}(\text{red})$ 

```

FIGURE 1. The implementation of algorithm ASYM-EDGE-COL.

of the elements on the stack, each such test is performed only after all edges of the corresponding clique were inserted and colored. For every blue copy of K_ℓ , one particular edge f (see line 23 of the algorithm) is recolored to red. Since red edges are never flipped back to blue, no blue copy of K_ℓ can occur in the coloring found by the algorithm.

We need to show that the edge f in line 23 always exists. Since the second loop inserts edges into G' in the reverse order in which they were deleted during the first loop, when we select f in line 23, G' has the same structure as at the time when L was pushed on the stack. This happened either in line 7 when there exists no r -clique in G' that intersects with L on some particular edge $e \in E(L)$, or in line 12 when L satisfies the condition of the if-clause in line 11. In both cases we have $L \notin \mathcal{L}_{G'}^*$, and hence there exists an edge $e \in E(L)$ such that all currently existing copies of K_r do not intersect with L exactly in e .

It remains to prove that changing the color of some edges from blue to red by the algorithm never creates an entirely red copy of K_r . By the condition on f in line 23 of the algorithm, at the moment f is recolored there exists no copy of K_r in G' that intersects L exactly in f . So there is either no K_r containing f at all, or every such copy contains also another edge from L . In the latter case, those copies cannot become entirely red since L is entirely blue. \square

It remains to prove the following lemma.

Lemma 12. *There exists a constant $b = b(\ell, r) > 0$ such that for $p \leq bn^{-1/m_2(K_\ell, K_r)}$, a.a.s. algorithm ASYM-EDGE-COL terminates on $G_{n,p}$ without error.*

We prove Lemma 12 by means of an auxiliary algorithm GROW. If ASYM-EDGE-COL fails on some graph G , GROW explicitly computes a subgraph $F \subseteq G$ which is either too large or too

```

GROW( $G' = (V, E)$ )
1  if  $\forall e \in E : |\mathcal{S}_{G'}(e)| = 1$ 
2    then  $T \leftarrow$  any member of  $\mathcal{T}_{G'}$ 
3    return  $\bigcup_{e \in E(T)} \mathcal{S}_{G'}(e)$ 
4  if  $\exists e \in E : |\mathcal{S}_{G'}(e)| \geq 2$ 
5    then  $S_1, S_2 \leftarrow$  any two distinct members of  $\mathcal{S}_{G'}(e)$ 
6    return  $S_1 \cup S_2$ 
7   $e \leftarrow$  any  $e \in E : |\mathcal{S}_{G'}(e)| = 0$ 
8   $F_0 \leftarrow$  any  $R \in \mathcal{R}_{G'} : e \in E(R)$ 
9   $i \leftarrow 0$ 
10 while  $i < \log(n) \wedge \forall \tilde{F} \subseteq F_i : \lambda(\tilde{F}) > -\gamma$ 
11 do if  $\exists R \in \mathcal{R}_{G'} \setminus \mathcal{R}_{F_i}$  s.t.  $|V(R) \cap V(F_i)| \geq 2$ 
12   then  $F_{i+1} \leftarrow F_i \cup R$ 
13   else  $e \leftarrow$  ELIGIBLE-EDGE( $F_i$ )
14      $F_{i+1} \leftarrow$  EXTEND-L( $F_i, e, G'$ )
15    $i \leftarrow i + 1$ 
16 if  $i \geq \log(n)$ 
17   then return  $F_i$ 
18   else return MINIMIZING-SUBGRAPH( $F_i$ )

```

```

EXTEND-L( $F, e, G'$ )
1   $L \leftarrow$  any  $L \in \mathcal{L}_{G'}^* : e \in E(L)$ 
2   $F' \leftarrow F \cup L$ 
3  for each  $e' \in E(L) \setminus E(F)$ 
4  do  $R_{e'} \leftarrow$  any  $R \in \mathcal{R}_{G'} : E(L) \cap E(R) = \{e'\}$ 
5     $F' \leftarrow F' \cup R_{e'}$ 
6  return  $F'$ 

```

FIGURE 2. The implementation of algorithm GROW.

dense to appear in $G_{n,p}$ with p as in the lemma. More precisely, we shall show that for any graph F that GROW may return, the probability that F appears in $G_{n,p}$ is small compared to the size of \mathcal{F} , the class of all graphs that GROW may possibly return. It follows that $G_{n,p}$ a.a.s. does not contain any of these graphs, which implies Lemma 12 by contradiction. Note that we employ algorithm GROW only for proving the lemma. It does not contribute to the running time of algorithm ASYM-EDGE-COL.

In order to formulate algorithm GROW, we need some definitions. Let

$$\gamma = \gamma(\ell, r) := \frac{1}{m_2(K_\ell, K_r)} - \frac{1}{m_2(K_\ell, K_r) + \varepsilon_0} > 0 .$$

For any graph F , let

$$\lambda(F) := v(F) - e(F)/m_2(K_\ell, K_r) .$$

The definition of $\lambda(F)$ is motivated by the fact that the number of copies of F in $G_{n,p}$ with $p = bn^{-1/m_2(K_\ell, K_r)}$ has the order of magnitude

$$n^{v(F)} p^{e(F)} = b^{e(F)} n^{\lambda(F)} .$$

For any graph F , we call an edge $e \in E(F)$ *eligible for extension* if it satisfies

$$\nexists (L, R) \in \mathcal{L}_F \times \mathcal{R}_F \text{ s.t. } E(F) \cap E(R) = \{e\} .$$

Note that F is in \mathcal{C} (as defined in Definition 7) if and only if it contains no edge that is eligible for extension.

The implementation of algorithm GROW is shown in Figure 2. The input is the graph $G' \subseteq G$ after ASYM-EDGE-COL got stuck. It proceeds as follows: Before entering the while-loop, it tests for two simple special cases. If neither of those special cases occurs, it chooses a suitable seed F_0 for the actual growing procedure. In every iteration i of the while-loop, the procedure extends F_i to F_{i+1} by adding new vertices and edges to it. As long as there are copies of K_r in G' that intersect with F_i in at least two vertices but not in all edges, it greedily adds those to F_i . If there are no such copies, it calls a function ELIGIBLE-EDGE that maps F_i to an edge $e \in E(F_i)$ eligible for extension, provided such an edge exists. Crucially, this edge is *unique up to isomorphism of F_i* , i.e., for any two isomorphic graphs F and F' , there exists an isomorphism φ with $\varphi(F) = F'$ such that

$$\text{ELIGIBLE-EDGE}(F') = \varphi(\text{ELIGIBLE-EDGE}(F)) .$$

Note that this implies in particular that e depends only on the graph F_i and not on the surrounding graph G' . One possible way of implementing ELIGIBLE-EDGE is to keep a large table of representatives for all isomorphism classes of graphs with up to n vertices. Since we only want to show the existence of a certain structure in G' and do not care about complexity issues here, the actual implementation of ELIGIBLE-EDGE is irrelevant. Procedure EXTEND-L then adds a graph $L \in \mathcal{L}_{G'}^*$ that contains the edge e returned by ELIGIBLE-EDGE to F_i . To each new edge $e' \in E(L) \setminus E(F_i)$, it glues a graph $R_{e'} \in \mathcal{R}_{G'}$ that intersects L only in e' . (We will argue that this is always possible later.) The algorithm stops as soon as $i \geq \log(n)$ or $\lambda(\tilde{F}) \leq -\gamma$ for some $\tilde{F} \subseteq F_i$. In the former case, it returns F_i , in the latter case, a subgraph $\tilde{F} \subseteq F_i$ that minimizes $\lambda(\tilde{F})$. We assume that, for each graph F , the function MINIMIZING-SUBGRAPH(F) returns such a minimizing subgraph that is *unique up to isomorphism*. Similarly to ELIGIBLE-EDGE, MINIMIZING-SUBGRAPH could be implemented using a huge lookup-table.

We shall argue that GROW terminates without error, i.e., that ELIGIBLE-EDGE always finds an edge eligible for extension, and that all ‘any’-assignments in GROW and EXTEND-L are always successful.

Let us consider the properties of G' when ASYM-EDGE-COL gets stuck. Since the condition in line 5 of ASYM-EDGE-COL fails, G' is in the family \mathcal{C} . In particular, every edge of G' is contained in a copy $L \in \mathcal{L}$, and, since the condition in line 11 fails as well, we know that this L belongs to $\mathcal{L}_{G'}^*$. That is, G' is even in the smaller family

$$\mathcal{C}^* = \mathcal{C}^*(\ell, r) := \{G = (V, E) : \forall e \in E(G) \exists L \in \mathcal{L}_{G'}^* \text{ s.t. } e \in E(L)\} \subseteq \mathcal{C} .$$

Lastly, G' is not a triangle-sparse \mathcal{A} -graph because ASYM-EDGE-COL terminated with an error.

Claim 13. *Algorithm GROW terminates without error on any input graph $G' \in \mathcal{C}^*$ that is no triangle-sparse \mathcal{A} -graph. Moreover, for every iteration i of the while-loop, we have $e(F_{i+1}) > e(F_i)$.*

Proof. We first show that nothing can go wrong if one of the two special cases in lines 1 to 6 occurs. The first case occurs if and only if G' is an \mathcal{A} -graph. By assumption, G' is not triangle-sparse, that is, the family $\mathcal{T}_{G'}$ is not empty. Hence the assignment in line 2 is successful. Clearly, the assignment in line 5 is always successful due to the if-condition in line 4.

It is also easily seen that the assignments in lines 7 and 8 are successful: none of the two special cases occurred, therefore there must be an edge $e \in E(G')$ that is not contained in any $S \in \mathcal{S}_{G'}$. Also, there must be a member of $\mathcal{R}_{G'}$ that contains e because G' is a member of $\mathcal{C}^* \subseteq \mathcal{C}$.

Next, we show that the call to ELIGIBLE-EDGE in line 13 is always successful. Suppose there is no edge in F_i that is eligible for extension for some $i \geq 0$. Then we have $F_i \in \mathcal{C}$ by definition. Moreover, it is easily seen that F_i is 2-connected by construction. However, the choice of F_0 in line 8 guarantees that F_i is not in \mathcal{A} , because the edge e that was selected in line 7 satisfies $|\mathcal{S}_{G'}(e)| = 0$ and is contained in $F_0 \subseteq F_i \subseteq G'$. It follows from the definition of \mathcal{A} that

$m(F_i) > m_2(K_\ell, K_r) + \varepsilon_0$. Thus, there exists a graph $\tilde{F} \subseteq F_i$ with $d(\tilde{F}) = m(F_i)$ that satisfies

$$\begin{aligned} \lambda(\tilde{F}) &\stackrel{\text{def}}{=} v(\tilde{F}) - \frac{e(\tilde{F})}{m_2(K_\ell, K_r)} \\ &= e(\tilde{F}) \left(\frac{1}{m(F_i)} - \frac{1}{m_2(K_\ell, K_r)} \right) \\ &\leq e(\tilde{F}) \left(\frac{1}{m_2(K_\ell, K_r) + \varepsilon_0} - \frac{1}{m_2(K_\ell, K_r)} \right) \\ &\stackrel{\text{def}}{=} -\gamma e(\tilde{F}) \leq -\gamma . \end{aligned}$$

Consequently, GROW terminates in line 10 without calling ELIGIBLE-EDGE. Hence, each call to ELIGIBLE-EDGE in GROW is successful and returns an edge e . The subsequent call to EXTEND-L(F_i, e, G') is successful since we have $G' \in \mathcal{C}^*$ and hence there exist suitable graphs $L \in \mathcal{L}_{G'}^*$ with $e \in E(L)$ and $R_{e'}$ for each $e' \in E(L) \setminus E(F_i)$.

Finally, we show that for every iteration i of the while-loop, we have $e(F_{i+1}) > e(F_i)$. Since an R found in line 12 is a clique in G' but not in F_i , it is clear that $F_{i+1} = F_i \cup R$ has at least one more edge than F_i .

Hence we assume that lines 13 and 14 are called. Let e denote the edge chosen in line 13 and L the subgraph chosen in line 1 of EXTEND-L(F_i, e, G'). By the definition of $\mathcal{L}_{G'}^*$, there exists $R_e \in \mathcal{R}_{G'}$ such that e is the intersection of R_e and L . When EXTEND-L(F_i, e, G') is called, we must have $R_e \in \mathcal{R}_{F_i}$ because the condition on line 11 of GROW failed. Hence, the graph L is not contained in F_i , since otherwise e would not be eligible for extension. Consequently, we have $e(F_{i+1}) \geq e(F_i \cup L) > e(F_i)$. \square

Now, we look at the evolution of F_i in more detail. We say that iteration i of the while-loop in procedure GROW is *non-degenerate* if all of the following assertions hold:

- The condition in line 11 evaluates to false and consequently EXTEND-L is called.
- In line 2 of EXTEND-L, we have $V(F) \cap V(L) = e$.
- In every execution of line 5 of EXTEND-L, we have $V(F') \cap V(R_{e'}) = e'$.

Otherwise, we call iteration i *degenerate*. In non-degenerate iterations, F_{i+1} is uniquely defined up to isomorphism for any given F_i . The structure of F_{i+1} depends solely on the implementation of the subroutine ELIGIBLE-EDGE, which determines the position where to attach the next copy of K_ℓ . Observe that the structures which are added in every step are isomorphic. A graph F_2 that results from two non-degenerate iterations is depicted in Figure 3 for $r = 6$ and $\ell = 4$. The little dashed circle identifies F_0 . The greater dotted circle circumscribes F_1 .

Claim 14. *If iteration i of the while-loop in procedure GROW is non-degenerate, we have*

$$\lambda(F_{i+1}) = \lambda(F_i) .$$

Proof. In a non-degenerate iteration, the graph L added in line 1 of EXTEND-L contributes $\ell - 2$ new vertices and $\binom{\ell}{2} - 1$ new edges to F . Each of these new edges then is replaced by a copy of K_r . Hence, using the definition of $m_2(K_\ell, K_r)$ in Theorem 4, we obtain

$$\begin{aligned} \lambda(F_{i+1}) - \lambda(F_i) &= \ell - 2 + \left(\binom{\ell}{2} - 1 \right) (r - 2) - \left(\binom{\ell}{2} - 1 \right) \frac{\binom{r}{2}}{m_2(K_\ell, K_r)} \\ &= \left(\binom{\ell}{2} - 1 \right) \left(\frac{\ell - 2}{\binom{\ell}{2} - 1} + r - 2 - \left(r - 2 + \frac{1}{m_2(K_\ell)} \right) \right) = 0 . \end{aligned}$$

\square

In a degenerate iteration i , the structure of F_{i+1} is not just a function of F_i , but varies with the structure of G' . Suppose that F_i is extended with an r -clique R in line 5. This R can intersect F_i in virtually every possible way. Moreover, there may be several copies of K_r that satisfy

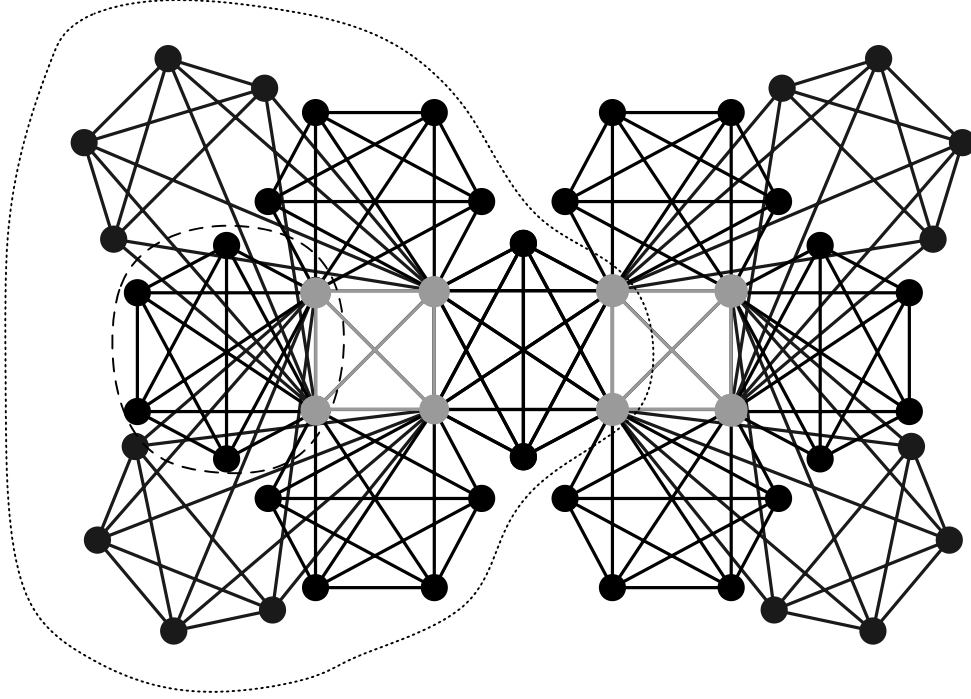


FIGURE 3. A graph F_2 resulting from two non-degenerate iterations for $r = 6$ and $\ell = 4$. The two central copies of K_4 are shaded.

the condition in line 11. The same is true for the graphs added in lines 2 and 5 of EXTEND-L. Thus, degenerate iterations cause difficulties since they enlarge the family of graphs that algorithm GROW can return. However, we will show that at most a *constant* number of degenerate iterations can occur before the algorithm terminates. This allows us to control the number of non-isomorphic graphs that GROW may return. The key for proving this is the next claim.

Claim 15. *There exists a constant $\kappa = \kappa(\ell, r) > 0$ such that if iteration i of the while-loop in procedure GROW is degenerate, we have*

$$\lambda(F_{i+1}) \leq \lambda(F_i) - \kappa .$$

The proof of Claim 15 is the main technical part of this work and postponed to Section 3.1. In combination, Claims 14 and 15 yield the next claim, which in turn leads to a polylogarithmic bound on the number of non-isomorphic graphs that GROW can return.

Claim 16. *There exists a constant $m_0 = m_0(\ell, r)$ such that algorithm GROW performs at most m_0 degenerate iterations before it terminates, regardless of the input instance G' .*

Proof. An easy calculation yields that $\lambda(F_0) = \lambda(K_r) = 2 - 2/(\ell + 1)$. The value of the function λ remains unchanged in every non-degenerate iteration due to Claim 14. However, Claim 15 yields a constant κ , which depends solely on ℓ and r , such that

$$\lambda(F_{i+1}) \leq \lambda(F_i) - \kappa$$

for every degenerate iteration i . Hence, after at most

$$m_0 := \frac{\lambda(F_0) + \gamma}{\kappa}$$

degenerate iterations, we have $\lambda(F_i) \leq -\gamma$, and the algorithm terminates. \square

Let $\tilde{\mathcal{F}} = \tilde{\mathcal{F}}(\ell, r, n)$ denote a family of representatives for the isomorphism classes of all graphs that can be the output of GROW with parameters n and $\gamma(\ell, r)$ on *any* input instance G' for which it enters the while-loop.

Claim 17. *There exists a constant $C = C(\ell, r)$ such that $|\tilde{\mathcal{F}}(\ell, r, n)| \leq \log(n)^C$ for n sufficiently large.*

Proof. Since the output of function MINIMIZING-SUBGRAPH is unique up to isomorphism, it suffices to count the graphs F_i that can be generated in the while-loop of GROW. For $0 \leq d \leq t \leq \lceil \log n \rceil$, let $\mathcal{F}(t, d)$ denote a family of representatives for the isomorphism classes of all graphs F_t that algorithm GROW can generate after exactly t iterations if it performs exactly d degenerate iterations along the way. Let $f(t, d) := |\mathcal{F}(t, d)|$ denote its cardinality.

Observe that in every iteration, we add at most

$$K := \ell - 2 + \binom{\ell}{2}(r - 2)$$

new vertices to F , which is exactly the number of vertices added in a non-degenerate iteration. Hence, we have $v(F_t) \leq r + Kt$. It also follows that in every iteration, the new edges $E(F_{t+1}) \setminus E(F_t)$ span a graph from \mathcal{G}_K , where \mathcal{G}_K denotes the set of all graphs on at most K vertices. F_{t+1} is uniquely defined if one specifies $G \in \mathcal{G}_K$, the number y of vertices in which G intersects F_t , and two ordered lists of vertices from G and F_t respectively of length y , which specify the mapping of the intersection vertices from G into F_t . Thus, the number of ways to extend F_t is bounded from above by

$$\sum_{G \in \mathcal{G}_K} \sum_{y=2}^{v(G)} v(G)^y v(F_t)^y \leq |\mathcal{G}_K| \cdot K \cdot K^K (r + Kt)^K \leq \lceil \log(n) \rceil^{C_0},$$

where the constant C_0 depends only on ℓ and r , and n is sufficiently large.

As the selection of the edge to be extended is unique up to isomorphism of F , the evolution of F is uniquely defined if there are no degenerate iterations along the way, regardless of the input instance G' . This implies in particular that $f(t, 0) = 1$ for all t , and more generally that for $0 \leq d \leq t \leq \lceil \log n \rceil$,

$$f(t, d) \leq \binom{t}{d} (\lceil \log(n) \rceil^{C_0})^d \leq \lceil \log(n) \rceil^{(C_0+1)d}.$$

Here the binomial coefficient corresponds to the choice of the d degenerate iterations. Claim 16 yields a constant $m_0 = m_0(\ell, r) > 0$ such that

$$\begin{aligned} |\tilde{\mathcal{F}}(\ell, r, n)| &\leq \sum_{t=0}^{\lceil \log(n) \rceil} \sum_{d=0}^{\min\{t, m_0\}} f(t, d) \\ &\leq (\lceil \log(n) \rceil + 1)(m_0 + 1) \lceil \log(n) \rceil^{(C_0+1)m_0} \leq \log(n)^C \end{aligned}$$

for an appropriately chosen constant $C = C(\ell, r)$ and n sufficiently large. \square

Let now $\mathcal{F} = \mathcal{F}(\ell, r, n) \supseteq \tilde{\mathcal{F}}$ denote a family of representatives for the isomorphism classes of all graphs that can be the output of GROW, regardless of whether it enters the while-loop or not.

Claim 18. *There exists a constant $b = b(\ell, r) > 0$ such that for $p \leq bn^{-1/m_2(K_\ell, K_r)}$, $G_{n,p}$ does not contain any graph from $\mathcal{F}(\ell, r, n)$ a.a.s.*

Proof. We start by considering the two special cases in lines 1 to 6. Let $\mathcal{F}_0 = \mathcal{F}_0(\ell, r) \subseteq \mathcal{F}$ denote the class of graphs that can be the output of GROW if one of these cases occurs. Clearly, we have $\mathcal{F} = \mathcal{F}_0 \cup \tilde{\mathcal{F}}$. Due to the conditions in lines 1 and 4 of GROW, any $F \in \mathcal{F}_0$ is of the form

$$F = \bigcup_{e \in E(T)} S_{G'}(e)$$

for some triangle $T \in \mathcal{T}_{G'}$, or of the form

$$F = S_1 \cup S_2$$

for some edge-intersecting $S_1, S_2 \in \mathcal{S}_{G'}$. Either way, as all elements of $\mathcal{S}_{G'}$ are 2-connected and in \mathcal{C} , so is the resulting $F \in \mathcal{F}_0$. On the other hand, $F \subseteq G'$ is not in $\mathcal{S}_{G'}$ and thus not isomorphic to a graph in \mathcal{A} , as otherwise the graphs S forming F would not be in $\mathcal{S}_{G'}$ due to the maximality condition in the definition of $\mathcal{S}_{G'}$. It follows that $m(F) > m_2(K_\ell, K_r) + \varepsilon_0$. By Lemma 8, the family \mathcal{F}_0 is finite, hence Markov's inequality yields that $G_{n,p}$ contains no graph from \mathcal{F}_0 a.a.s.

Now, suppose that GROW enters the while-loop. Let \mathcal{F}_1 and \mathcal{F}_2 denote the classes of graphs that algorithm GROW can output in lines 17 and 18, respectively. For all $F \in \mathcal{F}_1$, we have that $e(F) \geq \log(n)$, as F was generated in $\lceil \log(n) \rceil$ iterations, each of which introduces at least one new edge by Claim 13. Moreover, Claims 14 and 15 imply that $\lambda(F_i)$ is non-increasing. Thus, we have that $\lambda(F) \leq \lambda(F_0)$ for all $F \in \mathcal{F}_1$. For all $F \in \mathcal{F}_2$, we have that $\lambda(F) \leq -\gamma$ due to the condition in line 10 of GROW. Since we have chosen $F_0 = K_r$ as the seed of the growing procedure, it follows that for

$$b := e^{-\lambda(K_r) - \gamma} \leq 1 ,$$

the expected number of copies of graphs from $\tilde{\mathcal{F}}$ in $G_{n,p}$ with $p \leq bn^{-1/m_2(K_\ell, K_r)}$ is bounded by

$$\begin{aligned} \sum_{F \in \tilde{\mathcal{F}}} n^{v(F)} p^{e(F)} &= \sum_{F \in \tilde{\mathcal{F}}} b^{e(F)} n^{\lambda(F)} \\ &\leq \sum_{F \in \mathcal{F}_1} e^{(-\lambda(F_0) - \gamma) \log(n)} n^{\lambda(F_0)} + \sum_{F \in \mathcal{F}_2} n^{-\gamma} \\ &= \sum_{F \in \mathcal{F}_1} n^{-\gamma} + \sum_{F \in \mathcal{F}_2} n^{-\gamma} \\ &\leq 2(\log(n))^C n^{-\gamma} = o(1) , \end{aligned}$$

where the last inequality follows from Claim 17. Consequently, Markov's inequality implies that $G_{n,p}$ a.a.s. contains no graph from $\tilde{\mathcal{F}}$.

Combined with the earlier observation that $G_{n,p}$ a.a.s. has no subgraph from \mathcal{F}_0 , this proves that $G_{n,p}$ a.a.s. contains no graph from $\mathcal{F} = \mathcal{F}_0 \cup \tilde{\mathcal{F}}$. \square

Proof of Lemma 12. Suppose that the call to ASYM-EDGE-COL(G) gets stuck for some graph G , and consider $G' \subseteq G$ at this moment. The call to GROW(G') returns a copy of a graph $F \in \mathcal{F}(\ell, r, n)$ that is contained in $G' \subseteq G$. By Claim 18 this event does a.a.s. not occur in $G = G_{n,p}$ with p as claimed. This proves that ASYM-EDGE-COL finds a valid coloring of $G_{n,p}$ with $p \leq bn^{-1/m_2(K_\ell, K_r)}$ a.a.s. \square

Clearly, Lemma 12 implies the 0-statement of Theorem 4. It thus remains to prove Claim 15 and Lemma 8, which we do in Sections 3.1 and 3.2 respectively.

3.1. Proof of Claim 15

We say that Algorithm GROW encounters a *degeneracy of type 1* if the condition on line 11 evaluates to true.

Claim 19. *There exists a constant $\kappa_1 = \kappa_1(\ell, r) > 0$ such that if procedure GROW encounters a degeneracy of type 1 in iteration i of the while-loop, we have*

$$\lambda(F_{i+1}) \leq \lambda(F_i) - \kappa_1 .$$

Proof. The claim is trivial if R overlaps with F_i in r vertices since in that case we have to add at least one edge to F_i but no vertex, and therefore $\kappa_1 \geq 1/m_2(K_\ell, K_r)$. Hence, assume that R extends F_i by x , $1 \leq x \leq r - 2$ vertices. Then it must add at least $\binom{x}{2} + x(r - x)$ edges to F , and thus we have

$$\lambda(F_{i+1}) - \lambda(F_i) \leq x - \frac{\binom{x}{2} + x(r - x)}{m_2(K_\ell, K_r)} =: g(x) .$$

We can rewrite $g(x)$ to

$$g(x) = \frac{x}{r(r-1)(\ell+1)} \left((\ell(r-2) + r)x - r^2 - (r(r-4) + 2)\ell \right) .$$

The function $g(x)$ is a quadratic parabola, which attains zero for $x_1 = 0$ and

$$x_2 = \frac{r^2 + (r(r-4) + 2)\ell}{\ell(r-2) + r} .$$

Since

$$x_2 - (r-2) = \frac{2(r-\ell)}{\ell(r-2) + r} > 0 ,$$

$g(x)$ is negative in the entire interval $[1, r-2]$ and attains its maximum at either end of this interval. In fact, we have

$$g(r-2) - g(1) = \frac{(r-3)(\ell-1)}{(r-1)(\ell+1)} > 0 .$$

Hence, it is safe to choose κ_1 as

$$\kappa_1 = \min \left\{ \frac{1}{m_2(K_\ell, K_r)}, \quad -g(r-2) \right\} .$$

□

Before tackling the degeneracies that can occur when the condition in line 11 evaluates to false and EXTEND-L is called, we state a technical lemma that turns out to be crucial in our framework. In this lemma, we will consider graphs consisting of a copy of some graph J , every edge of which is embedded into an outer r -clique, where these outer r -cliques may overlap. We define the following families of graphs.

Definition 20. For every graph $J = (V, E)$ and every integer $r \geq 3$, let

$$\begin{aligned} \mathcal{H}(J, r) := \left\{ J^r = (V \dot{\cup} U, E \dot{\cup} D) : J^r \text{ is a minimal graph such that} \right. \\ \left. \text{for all } e' \in E, \text{ there are sets } U(e') \subseteq U \text{ and } D(e') \subseteq D \text{ with} \right. \\ \left. (e' \dot{\cup} U(e'), \{e'\} \dot{\cup} D(e')) \cong K_r \right\} . \end{aligned}$$

We refer to V and E as the inner vertices and edges respectively, which form the inner copy of J . Every edge $e' \in E$ forms together with its associated outer vertices $U(e')$ and outer edges $D(e')$ an outer copy of K_r . Hence, $|U(e')| = r-2$ and $|D(e')| = \binom{r}{2} - 1$ for all $e' \in E$. The minimality condition is understood w.r.t. subgraph inclusion, i.e., $J^r \in \mathcal{H}(J, r)$ does not have a subgraph which satisfies the same properties. This ensures in particular that $\mathcal{H}(J, r)$ is finite. Observe that for any nonempty graph J and integer $r \geq 3$ there is a unique graph $J^{r*} \in \mathcal{H}(J, r)$ in which the copies of the outer cliques are pairwise disjoint, i.e., which satisfies $U(e_1) \cap U(e_2) = \emptyset$ and $D(e_1) \cap D(e_2) = \emptyset$ for all pairs $e_1, e_2 \in E$. The following lemma relates the average degree of J^{r*} to that of all other members of the family $\mathcal{H}(J, r)$.

Lemma 21. *Let $r \geq 3$ be a fixed integer and J be a balanced nonempty graph. Let J^{r*} denote the unique member of $\mathcal{H}(J, r)$ with pairwise disjoint outer cliques. Then every member $J^r \in \mathcal{H}(J, r)$ satisfies*

$$\frac{e(J^r)}{v(J^r)} \geq \frac{e(J^{r*})}{v(J^{r*})} .$$

Although the statement of Lemma 21 may seem unsurprising, its proof is quite involved and deferred to Section 3.1.1.

With Lemma 21 at hand, we turn our attention back to the proof of Claim 15. We say that algorithm GROW encounters a *degeneracy of type 2* in iteration i of the while-loop if, during the call to EXTEND-L(F_i, e, G'), the graph L found in line 1 overlaps in more than two vertices

with F_i , or if there exists an edge $e' \in E(L) \setminus E(F)$ such that the graph $R_{e'}$ found in line 4 overlaps in more than two vertices with F' .

Claim 22. *There exists a constant $\kappa_2 = \kappa_2(\ell, r) > 0$ such that if procedure GROW encounters a degeneracy of type 2 in iteration i of the while-loop, we have*

$$\lambda(F_{i+1}) \leq \lambda(F_i) - \kappa_2 .$$

Proof. Consider the graph $F := F_i$ that is passed to EXTEND-L and the output of this procedure $F' := F_{i+1}$. We need to show that there exists a constant $\kappa_2 > 0$ such that

$$\lambda(F) - \lambda(F') = v(F) - v(F') - \frac{e(F) - e(F')}{m_2(K_\ell, K_r)} \geq \kappa_2 . \quad (3)$$

As we have argued before, the structure of F' would be uniquely defined up to isomorphism just by the structure of F if iteration i was non-degenerate. Let F^* denote the output of such a virtual non-degenerate iteration. We transform any degenerated outcome F' into F^* in three steps

$$F' = F^0 \xrightarrow{(i)} F^1 \xrightarrow{(ii)} F^2 \xrightarrow{(iii)} F^3 = F^* ,$$

each time carefully resolving certain kinds of degeneracies. By Claim 14 we have $\lambda(F) - \lambda(F^*) = 0$, and, using a telescoping summation, we may rewrite $\lambda(F) - \lambda(F')$ to

$$\begin{aligned} \lambda(F) - \lambda(F') &= \lambda(F^*) - \lambda(F') = \sum_{j=1}^3 (\lambda(F^j) - \lambda(F^{j-1})) \\ &= \sum_{j=1}^3 \left(v(F^j) - v(F^{j-1}) - \frac{e(F^j) - e(F^{j-1})}{m_2(K_\ell, K_r)} \right) . \end{aligned}$$

We shall show that for all $1 \leq j \leq 3$, we have

$$v(F^j) - v(F^{j-1}) - \frac{e(F^j) - e(F^{j-1})}{m_2(K_\ell, K_r)} \geq \kappa_2 \quad (4)$$

for a suitable $\kappa_2 = \kappa_2(\ell, r) > 0$, provided that F^j and F^{j-1} are not isomorphic. In each step j , $1 \leq j \leq 3$, we consider a different structural property of F' resulting from a degeneracy of type 2. Since we do not know the exact structure of F' , not every step j will necessarily modify the structure of F^{j-1} . However, at least for one j , F^j is not isomorphic to F^{j-1} since F' is not isomorphic to F^* . This suffices to conclude (3) from (4).

Let us carefully analyze the graph that procedure EXTEND-L appends to F when a degeneracy of type 2 occurs. First, it extends F by $L \cong K_\ell$. Let x denote the number of new vertices that L contributes to F , i.e., $x = |V(L) \setminus V(F)|$. Observe that $x \leq \ell - 2$ since L must overlap with F in at least one edge e , which was determined by subroutine ELIGIBLE-EDGE. As all edges of $L \in \mathcal{L}_{G'}^*$ are covered by copies of K_r , and since the condition in line 11 of GROW evaluates to false in iteration i , we have

$$\binom{V(F) \cap V(L)}{2} \subseteq E(F) . \quad (5)$$

This also implies that $x \geq 1$, as otherwise F would not be extended by at least one edge, contradicting Claim 13.

Let $L_x \subseteq L$ denote the graph where all edges in $E(F) \cap E(L)$ were removed from L . Clearly, we have $v(L_x) = \ell$ and

$$e(L_x) = \binom{\ell}{2} - \binom{\ell - x}{2} = \binom{x}{2} + x(\ell - x) .$$

To every edge $e' \in E(L_x)$, EXTEND-L glues a graph $R_{e'} \cong K_r$ in line 5 which intersects with L only in e' . (Recall that such a graph always exists since $L \in \mathcal{L}_{G'}^*$.) Let L_x^R denote the graph

$$L_x \cup \bigcup_{e' \in E(L_x)} R_{e'} .$$

Due to (5), F' can be written as $F \cup L_x^R$, and as the condition in line 11 of GROW fails, we have $E(F') = E(F) \dot{\cup} E(L_x^R)$. Therefore, we have

$$e(F') - e(F) = e(L_x^R) .$$

We may also conclude that $|V(F) \cap V(L_x^R)| = \ell - x$ and thus

$$\begin{aligned} v(F') - v(F) &= v(L_x^R) - |V(F) \cap V(L_x^R)| \\ &= v(L_x^R) - (\ell - x) - |V(F) \cap (V(L_x^R) \setminus V(L_x))| . \end{aligned}$$

Assuming $|V(F) \cap (V(L_x^R) \setminus V(L_x))| > 0$, we apply transformation (i) mapping F^0 to F^1 . The transformation introduces, for every vertex $v \in V(F) \cap (V(L_x^R) \setminus V(L_x))$, an additional vertex v' . All edges incident to v belonging to $E(F)$ remain connected to v and those belonging to $E(L_x^R)$ are redirected to v' . Thus, we disconnect the vertices in $V(L_x^R) \setminus V(L_x)$ from the vertices in $V(F)$. In L_x^R we replace the vertices $V(F) \cap (V(L_x^R) \setminus V(L_x))$ by the new vertices. Since there are no edges in $E(L_x^R) \cap E(F)$, the output of this operation is uniquely defined, and the structure of L_x^R is not affected. Hence, we have

$$v(F^1) - v(F^0) - \frac{e(F^1) - e(F^0)}{m_2(K_\ell, K_r)} = |V(F) \cap (V(L_x^R) \setminus V(L_x))| \geq 1 . \quad (6)$$

By construction, F^1 satisfies $V(F) \cap (V(L_x^R) \setminus V(L_x)) = \emptyset$.

Next, we shall disconnect copies of K_r in L_x^R that mutually intersect. Let K_ℓ^x denote the graph isomorphic to L_x , i.e., a clique on ℓ vertices from which all edges of a clique on $\ell - x$ vertices were removed. Since transformation (i) does not change the structure of L_x^R , that graph consists of an inner copy $L_x \cong K_\ell^x$ with every edge embedded into an r -clique, where these outer cliques may mutually intersect. Using the notation from Definition 20, L_x^R is isomorphic to a graph $(K_\ell^x)^r \in \mathcal{H}(K_\ell^x, r)$.

Transformation (ii) proceeds by replacing L_x^R in F^1 by a copy L_x^{R*} of the unique graph $(K_\ell^x)^{r*} \in \mathcal{H}(K_\ell^x, r)$ in which the outer r -cliques are disjoint. Due to the assumption that $V(F) \cap (V(L_x^R) \setminus V(L_x)) = \emptyset$, L_x^R intersects F in exactly $\ell - x$ anchor points. We obtain the graph F^2 by removing L_x^R and attaching $L_x^{R*} \cong (K_\ell^x)^{r*}$ to the very same anchor points. F^2 is uniquely defined up to isomorphism due to the symmetries in $(K_\ell^x)^{r*}$. Observe that if L_x^R is not isomorphic to $(K_\ell^x)^{r*}$, the minimality condition in the definition of $\mathcal{H}(K_\ell^x, r)$ yields that $v(L_x^R) < v(L_x^{R*})$. Together with Lemma 21, we obtain that

$$\begin{aligned} v(F^2) - v(F^1) &= \frac{e(F^2) - e(F^1)}{m_2(K_\ell, K_r)} \\ &= v(F \cup L_x^{R*}) - v(F \cup L_x^R) - \frac{e(F \cup L_x^{R*}) - e(F \cup L_x^R)}{m_2(K_\ell, K_r)} \\ &= v(L_x^{R*}) - v(L_x^R) - \frac{e(L_x^{R*}) - e(L_x^R)}{m_2(K_\ell, K_r)} \\ &\stackrel{L.21}{\geq} v(L_x^{R*}) - v(L_x^R) - \frac{e(L_x^{R*}) - e(L_x^R) \frac{v(L_x^R)}{v(L_x^{R*})}}{m_2(K_\ell, K_r)} \\ &= (v(L_x^{R*}) - v(L_x^R)) \left(1 - \frac{e(L_x^R)}{v(L_x^{R*}) m_2(K_\ell, K_r)} \right) \\ &\geq 1 - \frac{(e(L_x) \binom{r}{2}) \left(r - 2 + \frac{1}{m_2(K_\ell)} \right)}{(e(L_x)(r-2) + v(L_x)) \binom{r}{2}} \\ &= 1 - \frac{e(L_x) \left(r - 2 + \frac{2}{\ell+1} \right)}{e(L_x)(r-2) + \ell} \end{aligned} \quad (7)$$

$$\begin{aligned}
&= \frac{\ell - e(L_x) \frac{2}{\ell+1}}{e(L_x)(r-2) + \ell} \\
&\stackrel{x \leq \ell-2}{\geq} \frac{2}{\binom{\ell}{2} - 1} \frac{1}{(r-2) + \ell}.
\end{aligned}$$

It remains to prove (4) for the last transformation (iii). We transform F^2 to F^3 by replacing the graph $L_x^{R*} \cong (K_\ell^x)^{r*}$ by a copy $L_{\ell-2}^{R*}$ of $(K_\ell^{\ell-2})^{r*}$ if $x < \ell - 2$. Note that for $\ell = 3$ we always have $x = 1$, so this step never applies. $(K_\ell^{\ell-2})^{r*} \in \mathcal{H}(K_\ell^{\ell-2}, r)$ consists of an inner graph $K_\ell^{\ell-2}$, a clique on ℓ vertices with one missing edge, with mutually disjoint copies of K_r sitting on these inner edges. Recall that the function ELIGIBLE-EDGE determines an edge $e \in E(F)$ which is contained in an ℓ -clique $L \in \mathcal{L}_{G'}^*$ that extends F . In transformation (iii), we detach $L_x^{R*} \cong (K_\ell^x)^{r*}$ from F^2 by removing all but the $\ell - x$ vertices that join L_x^{R*} and F . Then, we glue a copy $L_{\ell-2}^{R*}$ of $(K_\ell^{\ell-2})^{r*}$ to F by identifying the edge e that was returned from ELIGIBLE-EDGE(F_i) and the edge missing in $K_\ell^{\ell-2}$ that complements it to a complete graph.

Since $0 < x < \ell - 2$, we have

$$\begin{aligned}
&v(F^3) - v(F^2) - \frac{e(F^3) - e(F^2)}{m_2(K_\ell, K_r)} \\
&= v(F \cup L_{\ell-2}^{R*}) - v(F \cup L_x^{R*}) - \frac{e(F \cup L_{\ell-2}^{R*}) - e(F \cup L_x^{R*})}{m_2(K_\ell, K_r)} \\
&= (v(L_{\ell-2}^{R*}) - 2) - (v(L_x^{R*}) - (\ell - x)) - \frac{e(L_{\ell-2}^{R*}) - e(L_x^{R*})}{m_2(K_\ell, K_r)} \\
&= (e(L_{\ell-2}) - e(L_x))(r-2) + \ell - 2 - x - \frac{(e(L_{\ell-2}) - e(L_x)) \binom{r}{2}}{m_2(K_\ell, K_r)} \\
&= \ell - 2 - x + \left[\binom{\ell}{2} - 1 - \left(\binom{\ell}{2} - \binom{\ell-x}{2} \right) \right] \left[(r-2) - \left(r - 2 + \frac{1}{m_2(K_\ell)} \right) \right] \\
&= \ell - 2 - x - \left(\binom{\ell-x}{2} - 1 \right) \frac{2}{\ell+1} \\
&= \ell - 2 - x - \frac{(\ell-2-x)(\ell+1-x)}{\ell+1} \\
&= (\ell-2-x) \frac{x}{\ell+1} \stackrel{1 \leq x \leq \ell-2}{\geq} \frac{\ell-3}{\ell+1}.
\end{aligned} \tag{8}$$

Hence, setting κ_2 to

$$\kappa_2 := \min \left\{ 1, \frac{2}{\binom{\ell}{2} - 1} \frac{1}{(r-2) + \ell}, \frac{\ell-3}{\ell+1} \right\}$$

satisfies (4) due to (6), (7), and (8). Note that we can ignore the last term for $\ell = 3$. This completes the proof of Claim 22. \square

Claim 15 now follows directly from Claims 19 and 22. It remains to prove Lemma 21.

3.1.1. *Proof of Lemma 21.* The intuition behind our approach is the following: J^{r*} can be transformed into any given $J^r \in \mathcal{H}(J, r)$ by successively merging outer copies of K_r . We shall do this in $e_J - 1$ steps, fixing a linear ordering on the inner edges E . For every edge $f \in E$, we merge the attached outer copy $\widehat{K}(f) \cong K_r$ to outer copies attached to edges preceding f in that ordering, keeping track of the number of edges $\Delta_e(f)$ and vertices $\Delta_v(f)$ vanishing in the process. One might hope that the density of J^r increases in every step of this process or, slightly stronger, that $\Delta_e(f)/\Delta_v(f) \leq e(J^{r*})/v(J^{r*})$ for all $f \in E$. Unfortunately, this does not hold,

but we shall prove the existence of enough ‘good’ steps in the process to compensate for all ‘bad’ ones which may arise. To find these good steps, we will group the edges $f \in E$ into appropriate ‘phases’ such that the existence of bad steps in a phase implies the existence of some good steps earlier in that phase.

Recall that in every graph $J^r = (V \dot{\cup} U, E \dot{\cup} D) \in \mathcal{H}(J, r)$, the inner copy (V, E) is isomorphic to J . By definition, for each inner edge $f \in E$, we can identify sets of outer vertices $U(f) \subseteq U$ and outer edges $D(f) \subseteq D$ such that $\widehat{K}(f) := (U(f) \dot{\cup} f, D(f) \dot{\cup} \{f\})$ is isomorphic to K_r . While these sets are not necessarily unique, for the rest of the proof, we fix one choice of appropriate sets $U(f)$ and $D(f)$. Note that by the minimality condition in Definition 20, every vertex and edge is included in at least one outer copy. Similarly, let $\widehat{K}_-(f) := (U(f) \dot{\cup} f, D(f))$ denote the subgraph that is obtained by removing f from $\widehat{K}(f)$. For every outer vertex $u \in U$ and for every outer edge $d \in D$ of J^r , the sets

$$E(u) := \{f \in E : u \in U(f)\}$$

and

$$E(d) := \{f \in E : d \in D(f)\}$$

indicate in which outer copies $\widehat{K}(f)$ u and d , respectively, participate.

Note that

$$\sum_{d \in D} |E(d)| = \sum_{d \in D} \sum_{\substack{f \in E: \\ d \in D(f)}} 1 = \sum_{f \in E} \sum_{\substack{d \in D: \\ d \in D(f)}} 1 = \sum_{f \in E} |D(f)| = e_J \left(\binom{r}{2} - 1 \right)$$

and analogously

$$\sum_{u \in U} |E(u)| = e_J(r - 2) .$$

Owing to

$$e(J^r) - |D| = e_J = e(J^{r*}) - e_J \left(\binom{r}{2} - 1 \right) = e(J^{r*}) - \sum_{d \in D} |E(d)| ,$$

we have

$$e(J^r) = e(J^{r*}) - \sum_{d \in D} (|E(d)| - 1) \quad (9)$$

and analogously

$$v(J^r) = v(J^{r*}) - \sum_{u \in U} (|E(u)| - 1) . \quad (10)$$

Next, we impose a linear order on the vertices and edges of J . To simplify the notation we set $\ell := v_J$ for the remainder of the proof. By averaging, it is easy to see that for every balanced graph J there exists an ordering $[v_1, \dots, v_\ell]$ of its vertices such that for all $2 \leq i \leq \ell$, v_i has at most $\lfloor 2d(J) \rfloor$ neighbors among $\{v_1, \dots, v_{i-1}\}$ in J . In fact, we can compute such an ordering by recursively removing a vertex of minimum degree. W.l.o.g. we assume that the inner vertices V are ordered in this way. This ordering induces a mapping $p : E \rightarrow \{2, \dots, \ell\}$, which assigns every inner edge to the greater label of its two vertices. We call $p(f)$ the *phase* of edge f . This mapping induces a partial order on E , which can be extended to a total order \prec by choosing an arbitrary order on edges of the same phase.

For $f \in E$, we define

$$\Delta_E(f) := D(f) \cap \left(\bigcup_{f' \prec f} D(f') \right) , \quad (11)$$

$$\Delta_V(f) := U(f) \cap \left(\bigcup_{f' \prec f} U(f') \right) , \quad (12)$$

and set $\Delta_e(f) := |\Delta_E(f)|$, $\Delta_v(f) := |\Delta_V(f)|$. Intuitively, $\Delta_e(f)$ is the number of edges vanishing from $\widehat{K}_-(f)$ when it is merged with preceding outer copies. Analogously, $\Delta_v(f)$ is the number of vertices vanishing in this merge operation.

$\Delta_E(f)$ contains all edges $d \in D(f)$ that also belong to $D(f')$ for some edge $f' \prec f$. By definition, both f and f' are in $E(d)$. Therefore, we have

$$\sum_{f \in E} \Delta_e(f) = \sum_{f \in E} \sum_{\substack{d \in D: \\ d \in D(f) \\ f \neq \min E(d)}} 1 = \sum_{d \in D} \sum_{\substack{f \in E: \\ d \in D(f) \\ f \neq \min E(d)}} 1 = \sum_{d \in D} (|E(d)| - 1)$$

and, by (9),

$$e(J^r) = e(J^{r*}) - \sum_{f \in E} \Delta_e(f) . \quad (13)$$

Analogously, we obtain from (10) that

$$v(J^r) = v(J^{r*}) - \sum_{f \in E} \Delta_v(f) . \quad (14)$$

In order to calculate the density of J^r , we introduce the following quantities. For every phase i , $2 \leq i \leq \ell$, we define

$$\Delta_e^i := \sum_{\substack{f \in E: \\ p(f)=i}} \Delta_e(f)$$

and

$$\Delta_v^i := \sum_{\substack{f \in E: \\ p(f)=i}} \Delta_v(f) .$$

Owing to (13) and (14), we can express the density of J^r simply as

$$\frac{e(J^r)}{v(J^r)} = \frac{e(J^{r*}) - \Delta_e^2 - \dots - \Delta_e^\ell}{v(J^{r*}) - \Delta_v^2 - \dots - \Delta_v^\ell} . \quad (15)$$

We call phase i *trivial* if $\Delta_v^i = 0$, which implies $\Delta_e^i = 0$. By Proposition 6, to show that

$$\frac{e(J^r)}{v(J^r)} \geq \frac{e(J^{r*})}{v(J^{r*})}$$

it suffices to prove that, in every non-trivial phase i , we have

$$\frac{\Delta_e^i}{\Delta_v^i} \leq \frac{e(J^{r*})}{v(J^{r*})} . \quad (16)$$

To prove this, suppose a non-trivial phase $i \in [\ell] \setminus \{1\}$ is fixed. For every step $f \in E$ with $p(f) = i$, let

$$q(f) := \frac{\Delta_e(f)}{\Delta_v(f)}$$

and

$$T(f) := \left(\Delta_V(f) \dot{\cup} f, \Delta_E(f) \right) \subseteq \widehat{K}_-(f) .$$

Intuitively, the graph $T(f)$ is formed by the edges and vertices that vanish in the merge step f . However, we have to add the two vertices of f to guarantee that the graph is well-defined. We say that an edge $f \in E$, $p(f) = i$, is *bad* if $q(f) > d(J^{r*})$. Note that, by (15) and Proposition 6, phase i trivially satisfies (16) if it does not contain bad edges. We shall demonstrate that this is the case if J is a tree. In all other cases, our strategy is to show that there are sufficiently many good edges in phase i that can compensate for the bad ones.

First, let us rewrite $d(J^{r*}) = e(J^{r*})/v(J^{r*})$. Easy algebraic transformations yield

$$\begin{aligned} d(J^{r*}) &= \frac{e_J \binom{r}{2}}{e_J(r-2) + v_J} \\ &= \frac{r+1}{2} - \frac{\frac{r+1}{2} - \frac{e_J}{v_J}}{\frac{e_J}{v_J}(r-2) + 1} \\ &= d_2(K_r) - \frac{d_2(K_r) - d(J)}{d(J)(r-2) + 1} . \end{aligned} \tag{17}$$

Now, suppose that J is a tree. Then $d(J) \geq 1/2$, and we deduce from (17) that

$$d(J^{r*}) \geq d_2(K_r) - \frac{d_2(K_r) - \frac{1}{2}}{\frac{1}{2}(r-2) + 1} = \frac{r+1}{2} - 1 = \frac{r-1}{2} .$$

On the other hand, since for a tree $[2d(J)] = 1$, phase i consists of a single edge $f = \{v_i, v_j\}$, $j < i$. Furthermore, if some edge incident to v_i vanished in phase i , it would necessarily belong to the outer edges $D(f')$ of some edge $f' \prec f$ (cf. (11)). This, however, is impossible since otherwise phase i would consist of both edges f and f' . Therefore, the vertex v_i remains isolated in $T(f)$, and we have

$$\Delta_E(f) \subseteq \binom{\Delta_V(f) \dot{\cup} \{v_j\}}{2} .$$

From this we obtain

$$q(f) \leq \frac{\binom{\Delta_v(f)+1}{2}}{\Delta_v(f)} = \frac{\Delta_v(f) + 1}{2} \leq \frac{r-1}{2} \leq d(J^{r*})$$

and conclude that in the case of trees, every phase consists of just one good edge.

In the remaining part of the proof, suppose that J contains at least one cycle. As J is balanced, we have $d(J) \geq 1$ and, owing to (17),

$$d(J^{r*}) \geq d_2(K_r) - \frac{d_2(K_r) - 1}{r-1} = \frac{r+1}{2} - \frac{1}{2} = \frac{r}{2} . \tag{18}$$

Let $f = \{v_i, v_j\}$, $j < i$, be any edge in phase i . A trivial bound on $\Delta_e(f)$ is

$$\Delta_e(f) \leq \binom{\Delta_v + 2}{2} - 1 , \tag{19}$$

resulting from

$$\Delta_E(f) \subseteq \binom{\Delta_V(f) \dot{\cup} f}{2} \setminus \{f\} .$$

We will now derive a stronger bound. Owing to (11) and (12), we have

$$T(f) = \left(\Delta_V(f) \dot{\cup} f, \Delta_E(f) \right) = \widehat{K}_-(f) \cap G(f) ,$$

where

$$G(f) := \bigcup_{f' \prec f} \widehat{K}_-(f') .$$

Intuitively, in the merge step f the outer copy $\widehat{K}_-(f)$ is merged from its virtual non-merged position onto $G(f)$, and the graph $T(f)$ describes the vertices and edges vanishing in the process. We decompose $G(f)$ into

$$G_0 := \bigcup_{\substack{f' \in E \\ p(f') < i}} \widehat{K}_-(f') \quad \text{and} \quad G_i(f) := \bigcup_{\substack{f' \prec f \\ p(f') = i}} \widehat{K}_-(f')$$

and let

$$U_0 := \bigcup_{\substack{f' \in E \\ p(f') < i}} U(f') \quad \text{and} \quad U_i(f) := \bigcup_{\substack{f' \prec f \\ p(f') = i}} U(f')$$

denote the corresponding sets of outer vertices.

The crucial observation is that the neighborhood of v_i in $G(f)$ entirely lies in $U_i(f)$ since v_i is not contained in the graph $\widehat{K}_-(f')$ for any edge f' with $p(f') < i$. Analogously, the neighborhood of v_j in $G(f)$ entirely lies in U_0 since v_j is not contained in the graph $\widehat{K}_-(f')$ for any edge $f' \prec f$ with $p(f') = i$.

This yields the same restrictions for the neighborhoods of v_i and v_j in $T(f) \subseteq G(f)$, i.e.,

$$\Delta_E(f) \subseteq \left[\binom{\Delta_V(f) \dot{\cup} f}{2} \setminus \{f\} \right] \setminus \left[E_{F'}(\{v_i\}, \Delta_V(f) \setminus U_i(f)) \dot{\cup} E_{F'}(\{v_j\}, \Delta_V(f) \setminus U_0) \right]$$

and thus

$$\Delta_e(f) \leq \binom{\Delta_v(f) + 2}{2} - 1 - \left| \Delta_V(f) \setminus (U_0 \cap U_i(f)) \right| \quad (20)$$

$$\leq \binom{\Delta_v(f) + 2}{2} - 1 - \Delta_v(f) + |U_0 \cap U_i(f)| . \quad (21)$$

Let $b \in E$, $p(b) = i$, denote the *worst* edge in phase i , i.e.,

$$b = \operatorname{argmax}_{f \in E, p(f)=i} q(f) .$$

As observed before, there is nothing to prove if b is good, i.e., $q(b) \leq d(J^{r*})$. Using (18), it follows from (19) that $\Delta_v(b) = r - 2$ since otherwise we had

$$q(b) = \frac{\Delta_e(b)}{\Delta_v(b)} \leq \frac{\binom{\Delta_v(b)+2}{2} - 1}{\Delta_v(b)} = \frac{\Delta_v(b) + 3}{2} \leq \frac{r}{2} \leq d(J^{r*}) ,$$

and thus b would be good. We obtain from (21) that

$$\Delta_e(b) \leq \binom{r}{2} - 1 - (r - 2 - |U_0 \cap U_i(b)|)$$

and, consequently,

$$\Delta_e(f) \leq q(b)\Delta_v(f) \leq \left(d_2(K_r) - \frac{r - 2 - |U_0 \cap U_i(b)|}{r - 2} \right) \Delta_v(f) \quad (22)$$

for all edges f of phase i .

On the other hand, observing that

$$\binom{\Delta_v(f) + 2}{2} - 1 = \frac{\Delta_v(f) + 3}{2} \Delta_v(f) \leq \frac{r + 1}{2} \Delta_v(f) = d_2(K_r) \Delta_v(f) ,$$

we obtain from (20) that

$$\Delta_e(f) \leq d_2(K_r) \Delta_v(f) - |\Delta_V(f) \setminus U_i(f)| \quad (23)$$

for all edges f of phase i .

Recall that $U_i(f)$ is the union of all sets $U(f')$ for $f' \prec f$ and $p(f') = i$. Hence, we can write

$$U_i(b) = \bigcup_{\substack{f \prec b \\ p(f)=i}} (U(f) \setminus U_i(f)) .$$

As this is a union of pairwise disjoint sets, we get

$$\begin{aligned} |U_0 \cap U_i(b)| &= \sum_{\substack{f \prec b \\ p(f)=i}} |U_0 \cap (U(f) \setminus U_i(f))| = \sum_{\substack{f \prec b \\ p(f)=i}} |(U_0 \cap U(f)) \setminus U_i(f)| \\ &= \sum_{\substack{f \prec b \\ p(f)=i}} \left| \left((U_0 \cup U_i(f)) \cap U(f) \right) \setminus U_i(f) \right| = \sum_{\substack{f \prec b \\ p(f)=i}} |\Delta_V(f) \setminus U_i(f)| . \end{aligned} \quad (24)$$

Using (23) for the edges f in phase i that satisfy $f \prec b$ and (22) for all remaining edges of phase i , we obtain with (24) that

$$\begin{aligned}
\frac{\Delta_e^i}{\Delta_v^i} &= \left(\sum_{\substack{f \prec b \\ p(f)=i}} \Delta_e(f) + \sum_{\substack{f \succ b \\ p(f)=i}} \Delta_e(f) \right) / \Delta_v^i \\
&\leq \left(\sum_{\substack{f \prec b \\ p(f)=i}} (d_2(K_r) \Delta_v(f) - |\Delta_V(f) \setminus U_i(f)|) \right. \\
&\quad \left. + \sum_{\substack{f \succ b \\ p(f)=i}} \left(d_2(K_r) - \frac{r-2 - |U_0 \cap U_i(b)|}{r-2} \right) \Delta_v(f) \right) / \Delta_v^i \\
&= d_2(K_r) - \left(|U_0 \cap U_i(b)| + \frac{r-2 - |U_0 \cap U_i(b)|}{r-2} \sum_{\substack{f \succ b \\ p(f)=i}} \Delta_v(f) \right) / \Delta_v^i \\
&\leq d_2(K_r) - \left(|U_0 \cap U_i(b)| + \frac{r-2 - |U_0 \cap U_i(b)|}{r-2} \Delta_v(b) \right) / \Delta_v^i .
\end{aligned} \tag{25}$$

Recall that we have $\Delta_v(b) = r - 2$ and that each phase consists of at most $\lfloor 2d(J) \rfloor$ merge operations. Hence, we have $\Delta_v^i \leq \lfloor 2d(J) \rfloor (r - 2)$. By plugging those bounds into (25), we obtain

$$\frac{\Delta_e^i}{\Delta_v^i} \leq d_2(K_r) - \frac{1}{\lfloor 2d(J) \rfloor} .$$

Comparing this to (17), we see that (16) follows if

$$\frac{d_2(K_r) - d(J)}{d(J)(r-2) + 1} \leq \frac{1}{\lfloor 2d(J) \rfloor} .$$

Dropping the floor function and expanding yields the quadratic inequality

$$(d(J) - 1)(2d(J) - 1) \geq 0 ,$$

which is obviously satisfied for $d(J) \geq 1$. This proves that (16) holds for all non-trivial phases $i \in [\ell] \setminus \{1\}$, which implies Lemma 21 owing to (15) and Proposition 6 as observed before.

3.2. Proof of Lemma 8

We readily obtain a proof of Lemma 8 from the following explicit representation of the family $\mathcal{A}(\ell, r)$. Recall that, for all $r \geq \ell \geq 3$, we defined

$$\mathcal{A}(\ell, r) := \{A \in \mathcal{C}(\ell, r) : m(A) \leq m_2(K_\ell, K_r) + \varepsilon_0 \wedge A \text{ is 2-connected}\} ,$$

where $\varepsilon_0 := 0.01$ is fixed (cf. Definition 7).

Lemma 23. *We have*

$$\begin{aligned}
\mathcal{A}(\ell, r) &= \emptyset \quad \text{for } r > \ell \geq 4 \quad \text{or } r = \ell \geq 5 , \\
\mathcal{A}(4, 4) &= \{K_6\} , \\
\mathcal{A}(3, r) &= \{K_{r+1}\} \quad \text{for } r \geq 6 , \\
\mathcal{A}(3, 5) &= \{K_6, K_6 \cap_5 K_6\} , \\
\mathcal{A}(3, 4) &= \{K_5, K_5 \cap_4 K_5, K_5 \cap_2 K_5\} ,
\end{aligned}$$

where $K_k \cap_x K_k$ denotes the graph consisting of two k -cliques that overlap in exactly x vertices.

Remark 24. Lemma 8 is false for $r = \ell = 3$, since C_k^{4*} , as introduced in Lemma 21, is a member of $\mathcal{A}(3, 3)$ for all $k \geq 3$.

First we deal with the easy case in which we have $r > \ell \geq 4$ or $r = \ell \geq 5$.

Proof of Lemma 23 (Case $r > \ell \geq 4$ or $r = \ell \geq 5$). Let $A \in \mathcal{A}(\ell, r) \subseteq \mathcal{C}(\ell, r)$ be given. Since every vertex of A is contained in an ℓ -clique and an r -clique that intersect in exactly one edge, the minimum degree in A is at least $(\ell - 1) + (r - 1) - 1 = \ell + r - 3$. It follows that $d(A) \geq (\ell + r - 3)/2$. With $m_2(K_\ell, K_r) = r(r - 1)/2(r - 2 + 2/(\ell + 1))$, we easily obtain that $d(A) > m_2(K_\ell, K_r) + \varepsilon_0$ for $r > \ell \geq 4$ or $r = \ell \geq 5$, contradicting the assumption that $A \in \mathcal{A}(\ell, r)$. \square

We defer the case $r = \ell = 4$ to the end of this section and fix $\ell = 3$ for now. As an intermediate step we define a subfamily $\mathcal{B}(r) \subseteq \mathcal{A}(\ell, r)$ by adding more structure to \mathcal{A} . We show that \mathcal{B} can be explicitly represented in the form that was given in Lemma 23 for \mathcal{A} . Then Lemma 23 follows by showing that we have $\mathcal{A}(3, r) \subseteq \mathcal{B}(r)$.

Definition 25. For $r \geq 4$, let

$$\mathcal{B}(r) := \{B : B \text{ is union of copies of } K_{r+1} \wedge B \text{ is 2-connected} \\ \wedge m(B) \leq m_2(K_3, K_r) + \varepsilon_0\} .$$

Note that, for all $r \geq 4$, we have $\mathcal{B}(r) \subseteq \mathcal{A}(3, r)$ since any union of copies of K_{r+1} is a member of $\mathcal{C}(3, r)$. We can list the members of the family $\mathcal{B}(r)$ explicitly.

Lemma 26. *We have*

$$\begin{aligned} \mathcal{B}(r) &:= \{K_{r+1}\} \quad \text{for } r \geq 6 , \\ \mathcal{B}(5) &:= \{K_6, K_6 \cap_5 K_6\} , \\ \mathcal{B}(4) &:= \{K_5, K_5 \cap_4 K_5, K_5 \cap_2 K_5\} . \end{aligned} \tag{26}$$

Proof. Suppose that r is fixed and $B \in \mathcal{B}(r)$ is given. If $r \geq 6$, the copies of K_{r+1} forming B are edge-disjoint, which we see as follows. Clearly, we have

$$d(K_{r+1} \cap_x K_{r+1}) = \frac{2 \binom{r+1}{2} - \binom{x}{2}}{2(r+1) - x} = \frac{2r^2 + 2r - x^2 + x}{2(2r - x + 2)} .$$

The second derivative

$$\frac{\partial^2}{\partial x^2} d(K_{r+1} \cap_x K_{r+1}) = \frac{2(r+1)^2}{(x - 2(r+1))^3}$$

is negative on the interval $2 \leq x \leq r$. Hence, the minimum of $d(K_{r+1} \cap_x K_{r+1})$ is at either boundary of the interval, i.e., at $x = 2$ or $x = r$. In fact, we have

$$d(K_{r+1} \cap_2 K_{r+1}) = \frac{r^2 + r - 1}{2r} \geq \frac{r(r+3)}{2(r+2)} = d(K_{r+1} \cap_r K_{r+1}) ,$$

and therefore

$$d(K_{r+1} \cap_x K_{r+1}) \geq \frac{r(r+3)}{2(r+2)} = \underbrace{\frac{r(r-1)}{2r-3}}_{=m_2(K_3, K_r)} + \frac{r(r-5)}{2(r+2)(2r-3)} \stackrel{r \geq 6}{>} m_2(K_3, K_r) + \varepsilon_0 .$$

If $r = 5$, one calculates that for $2 \leq x \leq 4$,

$$d(K_6 \cap_x K_6) = \frac{30 - \binom{x}{2}}{12 - x} \stackrel{2 \leq x \leq 4}{>} 20/7 + \varepsilon_0 = m_2(K_3, K_5) + \varepsilon_0 .$$

Similarly, for $r = 4$, one obtains

$$d(K_5 \cap_3 K_5) = 17/7 > 12/5 + \varepsilon_0 = m_2(K_3, K_4) + \varepsilon_0 .$$

This leaves exactly the cases listed in the statement of the lemma. Similar calculations show that there cannot be a third copy of K_{r+1} edge-intersecting with a graph formed by two edge-intersecting copies of K_{r+1} for $r \in \{4, 5\}$.

Hence every $B \in \mathcal{B}(r)$ is an edge-disjoint union of graphs listed in (26). Suppose that B itself is not isomorphic to one of the listed graphs. Since it is 2-connected by definition, it contains a cycle formed by edges from at least two such edge-disjoint graphs. Every edge of this cycle is contained in a copy of K_{r+1} , and we may assume that all these copies are different: if one copy of K_{r+1} contains two edges of the cycle, we can shortcut through this copy to obtain a smaller cycle which still is formed by edges from at least two edge-disjoint graphs as listed. If the cycle that we have just found has length k , then B contains a graph $C_k^{r+1} \in \mathcal{H}(C_k, r+1)$ as in Definition 20. Lemma 21 yields that

$$\begin{aligned} m(B) &\geq d(C_k^{r+1}) \geq d(C_k^{(r+1)*}) \\ &= \frac{k \binom{r+1}{2}}{kr} = \frac{r(r-1)}{2r-3} + \frac{r-3}{2(2r-3)} \\ &\stackrel{r \geq 4}{>} m_2(K_3, K_r) + \varepsilon_0 . \end{aligned}$$

This is a contradiction to the definition of $\mathcal{B}(r)$, so Lemma 26 follows. \square

We need the following technical lemma in order to derive that $\mathcal{A}(3, r)$ is contained in $\mathcal{B}(r)$.

Lemma 27. *Let $A \in \mathcal{A}(3, r)$ be given, and let $H \subseteq A$ denote the union of all copies of K_{r+1} in A . For any vertex $v_0 \in V(H)$ that is contained in not more than one block of H , $\deg_A(v_0) > \deg_H(v_0)$ implies that*

$$\deg_A(v_0) \geq \deg_H(v_0) + r - 1 .$$

Proof. Note that the blocks of H are members of the family $\mathcal{B}(r)$ by Definition 25. For the sake of contradiction, assume that $v_0 \in V(H)$ violates the claim, and let $B \subseteq H \subseteq A$, $B \in \mathcal{B}(r)$, be the unique block of H containing v_0 . Note that we have $\deg_H(v_0) = \deg_B(v_0)$. Let

$$\Delta := \deg_A(v_0) - \deg_B(v_0) .$$

Since there can be no edges from $E(A) \setminus E(B)$ connecting two vertices of B (if $B \cong K_{r+1} \cap_x K_{r+1}$, any such edge would create a structure which is too dense to appear in $A \in \mathcal{A}(r)$), Δ is also the degree of v_0 into $V(A) \setminus V(B)$.

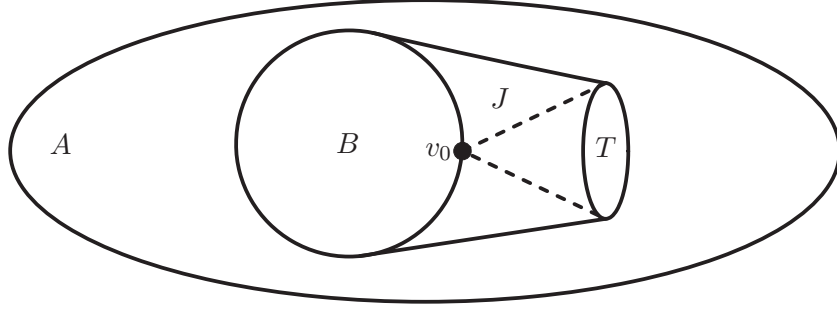
The assumption that v_0 violates the claim is equivalent to $1 \leq \Delta \leq r - 2$. Let $T := A[\Gamma(v_0) \setminus V(B)]$ denote the graph induced by the Δ neighbors of v_0 that are not in B . We shall prove that the graph

$$J := A[V(B) \cup \Gamma(v_0)] = A[V(B) \dot{\cup} V(T)]$$

satisfies $d(J) > m_2(K_3, K_r) + \varepsilon_0$, contradicting the assumption that $A \in \mathcal{A}(3, r)$. Figure 4 shows the formation of graphs A , B , T , and J .

Since A is a member of the family $\mathcal{C}(3, r)$, each of the Δ many edges connecting v_0 to T is covered by at least $(3 - 2) + (r - 2) = r - 1$ triangles. Therefore, each vertex of T has at least $r - 1$ common neighbors with v_0 and thus degree at least r in J . We obtain

$$\begin{aligned} e(J) &= e(B) + e(T) + \sum_{v \in V(T)} (\deg_J(v) - \deg_T(v)) \\ &\geq e(B) + e(T) + \sum_{v \in V(T)} (r - \deg_T(v)) \\ &= e(B) + \Delta r - e(T) \\ &\geq e(B) + \Delta r - \binom{\Delta}{2} \end{aligned}$$

FIGURE 4. The formation of graphs A , B , T , and J .

and, thus,

$$d(J) \geq \frac{e(B) + \Delta(r + 1/2 - \Delta/2)}{v(B) + \Delta} =: f(\Delta) .$$

Since

$$\frac{\partial^2 f(\Delta)}{\partial \Delta^2} = \frac{2e(B) - v(B)^2 - 2rv(B) - v(B)}{(v(B) + \Delta)^3} \leq 0 ,$$

f is concave. Consequently, we have

$$d(J) \geq \min \{ f(1), f(r-2) \} = \min \left\{ \frac{e(B) + r}{v(B) + 1}, \frac{e(B) + (r-2)(r+3)/2}{v(B) + r-2} \right\} . \quad (27)$$

Recall that B is isomorphic to a graph from $\mathcal{B}(r)$. If $B \cong K_{r+1}$, we have

$$f(1) = \frac{\binom{r+1}{2} + r}{r+2} = \frac{r(r-1)}{2r-3} + \frac{r(r-5)}{2(r+2)(2r-3)} \stackrel{r \geq 6}{>} m_2(K_3, K_r) + \varepsilon_0 , \quad (28)$$

and

$$\begin{aligned} f(r-2) &= \frac{\binom{r+1}{2} + (r-2)(r+3)/2}{2r-1} \\ &= \frac{r(r-1)}{2r-3} + \frac{2r^2 - 10r + 9}{(2r-1)(2r-3)} \stackrel{r \geq 4}{>} m_2(K_3, K_r) + \varepsilon_0 . \end{aligned}$$

This proves the claim for $r \geq 6$.

For $r \in \{4, 5\}$ and $B \cong K_{r+1}$, $\Delta = 1$ implies that $J \cong K_{r+1} \cap_r K_{r+1} \in \mathcal{B}(r)$, contradicting the assumption that B is a block of $H \subseteq A$. Hence we can replace (28) with

$$f(2) = \frac{\binom{r+1}{2} + 2r - 1}{r+3} = \frac{r(r-1)}{2r-3} + \frac{3r^2 - 13r + 6}{2(r+3)(2r-3)} \stackrel{r \geq 4}{>} m_2(K_3, K_r) + \varepsilon_0 ,$$

which settles this case.

Lastly, it is easily verified that (27) is also strictly greater than $m_2(K_3, K_r) + \varepsilon_0$ for $B \cong K_{r+1} \cap_x K_{r+1}$, $(r, x) \in \{(4, 2), (4, 4), (5, 5)\}$.

To summarize, we have shown that for $r \geq 4$ and $1 \leq \Delta \leq r-2$, $J \subseteq A$ satisfies $d(J) > m_2(K_3, K_r) + \varepsilon_0$, contradicting the assumption that $A \in \mathcal{A}(3, r)$. \square

Proof of Lemma 23 (Case $r > \ell = 3$). Let $A \in \mathcal{A}(3, r)$ be given, and, as in Lemma 27, let $H \subseteq A$ denote the union of all copies of K_{r+1} in A . Clearly by Lemma 26, we are home if we have $H = A$. Hence, suppose H is a proper subgraph of A for the sake of contradiction. Since we have $A \in \mathcal{C}(3, r)$, any two adjacent vertices have at least $(3-2) + (r-2) = r-1$ common neighbors. Consequently, any vertex of degree exactly r and its neighborhood induce a clique

of size $r + 1$ in A . This implies that the degree of all vertices of A that are not contained in a copy of K_{r+1} is at least $r + 1$. It follows that

$$\frac{\sum_{v \in V(A) \setminus V(H)} \deg_A(v)}{v(A) - v(H)} \geq r + 1 > 2(m_2(K_3, K_r) + \varepsilon_0) . \quad (29)$$

In particular, A contains at least one copy of K_{r+1} , i.e., H is nonempty.

Consider any fixed connected component C of H . If C is 2-connected, it follows from the 2-connectedness of A that C contains two vertices with $\deg_A(v) > \deg_C(v)$ to which Lemma 27 is applicable. Otherwise consider the block structure of C . It is tree-like, and all blocks that correspond to leaves contain some non cut-vertex v satisfying $\deg_A(v) > \deg_C(v)$. (If not, then such a block of C would also be a block of A , contradicting the assumption that A is 2-connected.) As there are at least two leaf blocks, again there are at least two vertices to which Lemma 27 is applicable.

This yields the following lower bound on the average degree of the vertices in C .

$$\begin{aligned} \frac{\sum_{v \in V(C)} \deg_A(v)}{v(C)} &\geq \frac{\sum_{v \in V(C)} \deg_C(v) + 2(r - 1)}{v(C)} \\ &= 2 \frac{\sum_{B \subseteq C} e(B) + r - 1}{\sum_{B \subseteq C} (v(B) - 1) + 1} , \end{aligned}$$

where the sum runs over all blocks B of C . By easy case checking, we see that

$$\frac{e(B)}{v(B) - 1} > m_2(K_3, K_r) + \varepsilon_0$$

for all $r \geq 4$ and $B \in \mathcal{B}(r)$. Moreover, since $r - 1 > m_2(K_3, K_r) + \varepsilon_0$, using Proposition 6 we obtain that

$$\frac{\sum_{v \in V(C)} \deg_A(v)}{v(C)} > 2(m_2(K_3, K_r) + \varepsilon_0) .$$

Summing over all connected components C of H and applying Proposition 6 again yields that

$$\frac{\sum_{v \in V(H)} \deg_A(v)}{v(H)} > 2(m_2(K_3, K_r) + \varepsilon_0) . \quad (30)$$

By Proposition 6, we conclude from equations (29) and (30) that

$$2d(A) = \frac{\sum_{v \in V(A)} \deg_A(v)}{v(A)} > 2(m_2(K_3, K_r) + \varepsilon_0) ,$$

contradicting $A \in \mathcal{A}(3, r)$. Thus, the assumption $H \subsetneq A$ is false. Consequently, $A = H$ is a union of copies of K_{r+1} . By definition, it is also 2-connected and thus in $\mathcal{B}(r)$. Lemma 23 now follows from Lemma 26. \square

Proof of Lemma 23 (Case $r = \ell = 4$). For the calculations, note that $m_2(K_4, K_4) = m_2(K_4) = 5/2$. Let $A \in \mathcal{A}(4, 4)$ be given. Similarly to the case $r > \ell = 3$, let $H \subseteq A$ denote the union of all copies $B \cong K_6$ in A , and assume for the sake of contradiction that $H \subsetneq A$.

Since every vertex of A is contained in two 4-cliques that intersect in exactly one edge, the minimum degree in A is at least $(4 - 1) + (4 - 1) - 1 = 5$. Moreover, any two adjacent vertices have at least $(4 - 2) + (4 - 2) = 4$ common neighbors. Consequently, any vertex of degree exactly 5 and its neighborhood induce a clique of size 6 in A . This implies that the degree of all vertices of A that are not contained in a copy of K_6 is at least 6. Consequently, we have

$$\frac{\sum_{v \in V(A) \setminus V(H)} \deg_A(v)}{v(A) - v(H)} \geq 6 > 2(m_2(K_4) + \varepsilon_0) . \quad (31)$$

On the other hand, trivial calculations yield that the 6-cliques B forming H must be mutually vertex-disjoint. As A is 2-connected, we obtain similarly to the case $r > \ell = 3$ that

$$\begin{aligned} \frac{\sum_{v \in V(H)} \deg_A(v)}{v(H)} &\geq \frac{\sum_{B \subseteq H} \left(\sum_{v \in V(B)} \deg_B(v) + 2 \right)}{v(H)} \\ &= 2 \frac{\sum_{B \subseteq H} (e(B) + 1)}{\sum_{B \subseteq H} v(B)} \\ &= 2 \cdot 16/6 > 2(m_2(K_4) + \varepsilon_0) . \end{aligned} \tag{32}$$

As before, we conclude with Proposition 6 from equations (31) and (32) that

$$2d(A) = \frac{\sum_{v \in V(A)} \deg_A(v)}{v(A)} > 2(m_2(K_4) + \varepsilon_0) ,$$

which implies the claim by contradiction. \square

4. THE 1-STATEMENT

Suppose that $k \geq 2$ and $3 \leq \ell_k \leq \dots \leq \ell_1$ are integers. We shall argue that there exists a constant $B > 0$ such that a.a.s. we have

$$G_{n,p} \rightarrow (K_{\ell_1}, \dots, K_{\ell_k})^e$$

provided $p \geq Bn^{-1/m_2(K_{\ell_2}, K_{\ell_1})}$, and Conjecture 23 in [10] is true for K_{ℓ_2} . For $\ell_2 = \dots = \ell_k = 3$, this statement follows from a theorem of Kohayakawa and Kreuter [9]. Note that we have $K_3 = C_3$, and that every clique is strictly 2-balanced.

Theorem 28 ([9]). *Let $k \geq 2$ and $3 \leq \ell_2 \leq \dots \leq \ell_k$ be integers and suppose H is a strictly 2-balanced graph with $m_2(H) \geq m_2(C_{\ell_2})$. Then there is a constant B such that, for $p = p(n) = Bn^{-1/m_2(C_{\ell_2}, H)}$, we have*

$$\mathbb{P}[G_{n,p} \rightarrow (H, C_{\ell_2}, \dots, C_{\ell_k})^e] = 1 - o(1) .$$

In fact, the result proved in [9] is stronger showing that even subgraphs of $G_{n,p}$ with sufficiently many edges satisfy the same conclusion. A careful inspection of the proof of Theorem 28 given in [9] reveals that only Lemma 17 from there actually depends on the structure of cycles. All other statements and arguments also hold when the cycles $C_{\ell_2}, \dots, C_{\ell_k}$ are replaced by arbitrary graphs H_2, \dots, H_k with $m_2(H) \geq m_2(H_2) \geq \dots \geq m_2(H_k)$. Instead of reproducing the proof, we sketch the necessary alterations for the sake of conciseness. The interested reader is referred to [9], which also contains an excellent outline of the proof in the introductory section explaining the main ideas.

Lemma 17 in [9] states an upper bound on the number of graphs satisfying certain pseudo-random properties. We need to introduce several notions so as to define this family of graphs precisely. Let F be a graph and $V = V(F)$ denote its set of vertices. Suppose $0 < \bar{p} \leq 1$ and $0 < \varepsilon \leq 1$ are real numbers, and U and W are disjoint nonempty subsets of V . Define the \bar{p} -density of F between U and W by

$$d_{F, \bar{p}}(U, W) := \frac{e_F(U, W)}{\bar{p}|U||W|} .$$

We say that the pair (U, W) is $(\varepsilon, F, \bar{p})$ -regular if for all $U' \subseteq U$ and $W' \subseteq W$ with $|U'| \geq \varepsilon|U|$ and $|W'| \geq \varepsilon|W|$, we have

$$|d_{F, \bar{p}}(U', W') - d_{F, \bar{p}}(U, W)| \leq \varepsilon .$$

Let J be a graph of order $v(J) = \ell$ on the vertex set $\{v_1, \dots, v_\ell\}$ and let also $\mathbb{V} = (V_i)_{i=1}^\ell$ be a family of ℓ pairwise disjoint sets, each of cardinality m . Suppose that reals $0 < \gamma \leq 1$, $D \geq 1$, and an integer T are given. We shall be interested in finding the size of the family, denoted by

$$\mathcal{F}(J, \varepsilon, \bar{p}, \gamma, D; \mathbb{V}, T) ,$$

consisting of all ℓ -partite graphs F with partition classes V_i , $1 \leq i \leq \ell$, and with exactly T edges that satisfy the following properties:

- (i) (V_i, V_j) is $(\varepsilon, F, \bar{p})$ -regular and has \bar{p} -density $\gamma \leq d_{F, \bar{p}}(V_i, V_j) \leq D$ whenever $v_i v_j \in E(J)$.
- (ii) F does not contain J as a subgraph.

Let

$$\mathcal{F}'(J, \varepsilon, \bar{p}, \gamma, D; \mathbb{V}, T) \subseteq \mathcal{F}(J, \varepsilon, \bar{p}, \gamma, D; \mathbb{V}, T)$$

be the family of graphs that additionally satisfy that, for all $U \subseteq V_{i-1}$ and $W \subseteq V_i$, $2 \leq i \leq \ell$, with

$$|U| \leq |W| \leq \bar{d}|U| \leq \bar{d}^{\ell-2} ,$$

where $\bar{d} = \bar{p}m$, we have

$$e_F(U, W) \leq C|W| .$$

Then Lemma 17 from [9] reads as follows.

Lemma 29 ([9], Lemma 17). *Let an integer $\ell \geq 3$ be fixed and let constants $0 < \alpha \leq 1$, $0 < \gamma \leq 1$, $C \geq 1$ and $D \geq 1$ be given. Then there are constants $0 < \varepsilon \leq 1$, $\bar{B} \geq 0$, and m_0 that depend only on ℓ , α , γ , C , and D such that, if $\bar{p} = \bar{p}(m) \geq \bar{B}m^{-1/m_2(C)}$, for all integers $m \geq m_0$ and $T \geq 1$, we have*

$$|\mathcal{F}'(C_\ell, \varepsilon, \bar{p}, \gamma, D; \mathbb{V}, T)| \leq \alpha^T \binom{(\ell+2)m^2}{T} .$$

This lemma is a special case of the following conjecture that appeared in [10].

Conjecture 30 ([10], Conjecture 23). *Let J be a graph on $\ell \geq 3$ vertices and let constants $0 < \alpha \leq 1$, $0 < \gamma \leq 1$, and $D \geq 1$ be given. Then there are constants $0 < \varepsilon \leq 1$, $\bar{B} \geq 0$, and m_0 that depend only on J , α , γ , and D such that, if $\bar{p} = \bar{p}(m) \geq \bar{B}m^{-1/m_2(J)}$, for all integers $m \geq m_0$ and $T \geq 1$, we have*

$$|\mathcal{F}(J, \varepsilon, \bar{p}, \gamma, D; \mathbb{V}, T)| \leq \alpha^T \binom{\binom{\ell}{2}m^2}{T} .$$

Note that the factors $(\ell+2)$ and $\binom{\ell}{2}$ respectively in the binomial coefficients are negligible since they contribute only a factor of $O(1)^T$ to the total expression, which may be suppressed by choosing α sufficiently small. Moreover, the constant D was arbitrarily set to 2 in the original formulation of Conjecture 30. Conjecture 30 has been verified for cycles of fixed length in [3], which clearly implies the bound on the size of the family $\mathcal{F}'(C_\ell, \varepsilon, \bar{p}, \gamma, D; \mathbb{V}, T)$ in Lemma 29.

The following generalized version of Theorem 28 can be proved by substituting Lemma 29 by Conjecture 30 and every occurrence of the cycles $C_{\ell_2}, \dots, C_{\ell_k}$ by the graphs H_2, \dots, H_k in the proof of Theorem 28 presented in [9].

Theorem 31. *Let $k \geq 2$ be an integer and H_1, \dots, H_k be graphs such that H_1 is strictly 2-balanced with $m_2(H_1) \geq m_2(H_2) \geq \dots \geq m_2(H_k)$. Suppose Conjecture 30 holds for all H_2, \dots, H_k . Then there is a constant B such that, for $p = p(n) = Bn^{-1/m_2(H_2, H_1)}$, we have*

$$\mathbb{P}[G_{n,p} \rightarrow (H_1, \dots, H_k)^e] = 1 - o(1) .$$

Furthermore, if $H_k \subseteq \dots \subseteq H_2$ holds, the conclusion of Theorem 31 follows even if Conjecture 30 is only known for H_2 . All we need to observe is that

$$G \rightarrow (H_1, H_2, H_2, \dots, H_2)^e$$

implies

$$G \rightarrow (H_1, H_2, H_3, \dots, H_k)^e ,$$

and that the order of magnitude of $p = p(n)$ in Theorem 31 depends only on H_1 and H_2 . Therefore, the 1-statement of Theorem 4 follows from Theorem 31 since $K_{\ell_k} \subseteq \dots \subseteq K_{\ell_2}$.

REFERENCES

- [1] F. Chung and R. Graham. *Erdős on graphs*. A K Peters Ltd., Wellesley, MA, 1998. His legacy of unsolved problems.
- [2] J. Folkman. Graphs with monochromatic complete subgraphs in every edge coloring. *SIAM J. Appl. Math.*, 18:19–24, 1970.
- [3] S. Gerke, Y. Kohayakawa, V. Rödl, and A. Steger. Small subsets inherit sparse ε -regularity. *J. Combin. Theory Ser. B*, 97(1):34–56, 2007.
- [4] S. Gerke, M. Marciniszyn, and A. Steger. A probabilistic counting lemma for complete graphs. In Stefan Felsner, editor, *2005 European Conference on Combinatorics, Graph Theory and Applications (EuroComb '05)*, volume AE of *DMTCS Proceedings*, pages 309–316. Discrete Mathematics and Theoretical Computer Science, 2005.
- [5] S. Gerke and A. Steger. The sparse regularity lemma and its applications. In *Surveys in combinatorics 2005*, volume 327 of *London Math. Soc. Lecture Note Ser.*, pages 227–258. Cambridge Univ. Press, Cambridge, 2005.
- [6] S. Janson, T. Łuczak, and A. Ruciński. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.
- [7] J. H. Kim. The Ramsey number $R(3, t)$ has order of magnitude $t^2/\log t$. *Random Structures Algorithms*, 7(3):173–207, 1995.
- [8] Y. Kohayakawa. Szemerédi’s regularity lemma for sparse graphs. In *Foundations of computational mathematics (Rio de Janeiro, 1997)*, pages 216–230. Springer, Berlin, 1997.
- [9] Y. Kohayakawa and B. Kreuter. Threshold functions for asymmetric Ramsey properties involving cycles. *Random Structures Algorithms*, 11(3):245–276, 1997.
- [10] Y. Kohayakawa, T. Łuczak, and V. Rödl. On K^4 -free subgraphs of random graphs. *Combinatorica*, 17(2):173–213, 1997.
- [11] Y. Kohayakawa and V. Rödl. Regular pairs in sparse random graphs. I. *Random Structures Algorithms*, 22(4):359–434, 2003.
- [12] B. Kreuter. Threshold functions for asymmetric Ramsey properties with respect to vertex colorings. *Random Structures Algorithms*, 9(3):335–348, 1996.
- [13] T. Łuczak, A. Ruciński, and B. Voigt. Ramsey properties of random graphs. *J. Combin. Theory Ser. B*, 56(1):55–68, 1992.
- [14] M. Marciniszyn, J. Skokan, R. Spöhel, and A. Steger. Threshold functions for asymmetric Ramsey properties involving cliques. In *Approximation, randomization and combinatorial optimization*, volume 4110 of *Lecture Notes in Comput. Sci.*, pages 462–474. Springer, Berlin, 2006.
- [15] J. Nešetřil and V. Rödl. The Ramsey property for graphs with forbidden complete subgraphs. *J. Combinatorial Theory Ser. B*, 20(3):243–249, 1976.
- [16] F. P. Ramsey. On a problem of formal logic. *Proceedings of the London Mathematical Society*, 30:264–286, 1930.
- [17] V. Rödl and A. Ruciński. Lower bounds on probability thresholds for Ramsey properties. In *Combinatorics, Paul Erdős is eighty, Vol. 1*, Bolyai Soc. Math. Stud., pages 317–346. János Bolyai Math. Soc., Budapest, 1993.
- [18] V. Rödl and A. Ruciński. Threshold functions for Ramsey properties. *J. Amer. Math. Soc.*, 8(4):917–942, 1995.

INSTITUTO DE MATEMÁTICA E ESTATÍSTICA, UNIVERSIDADE DE SÃO PAULO, 05508-090 SÃO PAULO, SP, BRAZIL

Current address: Department of Mathematics, London School of Economics, Houghton Street, London WC2A 2AE, United Kingdom and Department of Mathematics, MC-382, University of Illinois at Urbana–Champaign, 1409 W. Green Street, Urbana, IL, 61801, USA

E-mail address: J.Skokan@lse.ac.uk

INSTITUTE OF THEORETICAL COMPUTER SCIENCE, ETH ZURICH, 8092 ZURICH, SWITZERLAND

E-mail address: {mmarcini,rspoehel,stegeer}@inf.ethz.ch