On the Structure of Clique-free Graphs

Hans Jürgen Prömel*
Institut für Informatik, Humboldt-Universität zu Berlin, 10099 Berlin, Germany

Thomas Schickinger†

Angelika Steger
Institut für Informatik, Technische Universität München, 80290 München, Germany

ABSTRACT

Using a clever inductive counting argument Erdős, Kleitman and Rothschild showed in 1976 that almost all triangle-free graphs are bipartite, i.e., that the cardinality of the two graph classes is asymptotically equal. In this paper we investigate the structure of the 'few' triangle-free graphs which are not bipartite. As it turns out, with high probability these graphs are bipartite up to a few vertices. More precisely, almost all of them can be made bipartite by removing just one vertex. Almost all others can be made bipartite by removing two vertices, and then three vertices and so on. We also show that similar results hold if we replace triangle-free by $K_{\ell+1}$-free and bipartite by ℓ-partite.

1. INTRODUCTION

What is a random triangle-free graph? – In 1976 Erdős, Kleitman and Rothschild [2] showed that if one conditions a random graph $G_{n,\frac{1}{2}}$ with edge probability 1/2 on being triangle-free, then such a graph is almost surely bipartite, i.e.,

$$\Pr[\; G \text{ bipartite } \mid G = G_{n,\frac{1}{2}} \text{ triangle-free }] = 1 - o(1) \;.$$

* Supported by DFG-grant Pr 296/4-2.
† Supported by DFG-grant Ste 464/3-1.

Random Structures & Algorithms Vol. (Volume Number), ((Year))

While this result surely answers the question of what the structure of a random triangle-free graph is, it is also in some sense unsatisfactory, as, clearly, a bipartite graph does not capture many interesting properties a triangle-free graph might have.

In this paper we thus investigate the structure of those triangle-free graphs which are not bipartite. As it turns out, this class is surprisingly well-structured. Let us call a graph $G = (V, E)$ i-quasibipartite iff there exist i vertices $v_1, \ldots, v_i \in V$ such that $G[V \setminus \{v_1, \ldots, v_i\}]$ is bipartite. Then we prove

- almost all triangle-free graphs which are not bipartite are 1-quasibipartite,
- almost all triangle-free graphs which are not 1-quasibipartite are 2-quasibipartite,

and in general for all $i \geq 1$

- almost all triangle-free graphs which are not i-quasibipartite are $i+1$-quasibipartite.

We also show that similar results hold for the class of all graphs which do not contain a clique K_{l+1} of size $l+1$. This extends a result of Kolaitis, Prömel and Rothschild [3].

2. MAIN RESULT

In order to state our result precisely we introduce some notation. Throughout the rest of the paper we assume that ℓ is an arbitrary, but fixed integer with $\ell \geq 2$.

We consider graphs with labeled vertices. With $\mathcal{F}_0^\ell(n)$ we denote the set of all $K_{\ell+1}$-free graphs on n vertices and the set $\mathcal{P}_0^\ell(n)$ contains all ℓ-partite graphs. Using these abbreviations the result from Kolaitis, Prömel and Rothschild [3] can be stated as follows.

Theorem 2.1 [3]. *For $\ell \geq 2$ almost all $K_{\ell+1}$-free graphs are ℓ-partite, i.e.,*

$$|\mathcal{F}_0^\ell(n)| = (1 + o(1))|\mathcal{P}_0^\ell(n)| .$$

Note that the special case $\ell = 2$ corresponds to the result of Erdős, Kleitman and Rothschild [2].

We extend this result as follows: Let $\mathcal{P}_i^\ell(n)$ denote the set of all $K_{\ell+1}$-free graphs $G = (V, E)$ on n vertices which can be made ℓ-partite by removing i vertices but not by removing only $i - 1$ vertices, i.e., $\exists v_1, \ldots, v_i \in V$ such that $G[V \setminus \{v_1. \ldots, v_i\}] \in \mathcal{P}_0^\ell(n)$, whereas $\forall v'_1, \ldots, v'_{i-1} \in V$ it holds that $G[V \setminus \{v'_1, \ldots, v'_{i-1}\}] \notin \mathcal{P}_0^\ell(n)$. We call these $K_{\ell+1}$-free graphs *i-quasi-ℓ-partite* .

Additionally, we define the set $\mathcal{F}_i^\ell(n) := \mathcal{F}_0^\ell(n) \setminus [\mathcal{P}_0^\ell(n) \cup \ldots \cup \mathcal{P}_{i-1}^\ell(n)]$, i.e. $\mathcal{F}_i^\ell(n)$ contains all $K_{\ell+1}$-free graphs which cannot be made ℓ-partite by removing at most $i-1$ vertices. For these graphs the following extension to Theorem 2.1 can be shown.

Theorem 2.2. *For all $i \geq 0$ and $\ell \geq 2$ it holds that*

$$|\mathcal{F}_i^\ell(n)| = (1 + o(1))|\mathcal{P}_i^\ell(n)| .$$

The special case $i = 1$ and $\ell = 2$ was already shown in [5].

3. THE PROOF

For the proof of Theorem 2.2 we use an inductive counting method which was introduced by Kleitman and Rothschild in [4]. For some further applications of this method we refer the interested reader to the survey [6]. The key idea of this method is to define appropriate 'bad sets' of graphs, i.e., sets of graphs with 'strange' properties. On the one hand, these bad sets should cover the whole set $\mathcal{F}_i^\ell(n) \setminus \mathcal{P}_i^\ell(n)$. On the other hand, their sizes should be "small" compared to the size of the graph class $\mathcal{P}_i^\ell(n)$.

In Section A. we collect a few lemmas on the growth rate of $|\mathcal{P}_i^\ell(n-1)|$. In Section B. the bad sets are defined and their sizes are examined. The results of these two sections are then combined in Section C. and Theorem 2.2 is proved.

Notation. In this paper we will use the following notational conventions. $\Gamma(v)$ denotes the *neighborhood* of vertex v. With $\log n = \log_2 n$ we denote the logarithm to base 2. For $i \geq 1$ the *iterated logarithm* $\log^{(i)}(n)$ is recursively defined by $\log^{(1)}(n) = \log(n)$ and $\log^{(i+1)}(n) = \log(\log^{(i)}(n))$.

$H(x)$ denotes the *entropy function* and is defined as $H(x) := -x\log x - (1-x)\log(1-x)$. This function is used in the well known bound

$$\binom{n}{\lambda n} \leq 2^{nH(\lambda)} .$$

For our asymptotic estimates we will often assume that n is sufficiently large. This will not always be explicitly mentioned.

A. Growth Rates

In this section we show some estimates on the growth rate of $|\mathcal{P}_i^\ell(n)|$ considered as a function of n and i. Later we will compare these growth rates to the growth rates of the bad sets and it will turn out that these sets grow asymptotically slower than $\mathcal{P}_i^\ell(n)$. Hence, for n sufficiently large the cardinality of the bad sets is tiny in comparison to $|\mathcal{P}_i^\ell(n)|$.

Lemma 3.1. *For all $i \geq 0$ and n sufficiently large it holds that*

$$\log \frac{|\mathcal{P}_i^\ell(n-1)|}{|\mathcal{P}_i^\ell(n)|} \leq -\frac{\ell-1}{\ell}n + i + 1 .$$

Proof. We construct pairwise different graphs G in $\mathcal{P}_i^\ell(n)$ as follows: We choose a graph $G' = (V', E') \in \mathcal{P}_i^\ell(n-1)$ for the first $n-1$ vertices. Consider i vertices $v_1, \ldots, v_i \in V'$ such that $G'' := G'[V' \setminus \{v_1, \ldots, v_i\}]$ is ℓ-partite and fix an arbitrary ℓ-coloring of G''. Clearly, this coloring contains a color class S with at most $\frac{n-1-i}{\ell}$ vertices. If we don't connect the nth vertex v to $S \cup \{v_1, \ldots, v_i\}$ one easily verifies that the resulting graph G cannot contain a $K_{\ell+1}$.

Hence, there are at least $2^{n-1-i-\frac{n-1-i}{\ell}}$ possibilities to connect v to G'. Therefore,

$$|\mathcal{P}_i^\ell(n)| \geq |\mathcal{P}_i^\ell(n-1)| \cdot 2^{(1-\frac{1}{\ell})n-i-1}$$

and the claim follows. ∎

From this lemma we easily deduce a corollary which will later come in handy.

Corollary 3.2. *For $i \geq 0$ and all integral positive functions $g(n) \in \mathcal{O}(\log n)$ it holds that*

$$\log \frac{|\mathcal{P}_i^\ell(n-g(n))|}{|\mathcal{P}_i^\ell(n)|} = -\frac{\ell-1}{\ell} n \cdot g(n) + \mathcal{O}(g(n)^2) .$$

Proof. Multiple application of Lemma 3.1. ∎

Finally, we will need a lemma on the relation between the number of ℓ-partite graphs and the number of i-quasi-ℓ-partite graphs.

Lemma 3.3. *For all $i > 0$ and n sufficiently large it holds that*

$$\frac{|\mathcal{P}_0^\ell(n-i(2\ell+1))|}{|\mathcal{P}_i^\ell(n)|} \leq 1 .$$

Proof. For the proof of this lemma we need a small $K_{\ell+1}$-free graph H with chromatic number $\ell+1$. H will serve as a witness that a graph G which contains H as a subgraph is not ℓ-colorable.

Let $H = (S, F)$ be a graph on $2\ell+1$ vertices with

$$S = \{v, v_1, \ldots, v_\ell, v'_1, \ldots, v'_\ell\} .$$

F is assumed to be a minimal set of edges such that

1. v is connected to all vertices $v_1, \ldots, v_\ell$.

2. $G[\{v'_1, \ldots, v'_\ell\}]$ is a complete graph K_ℓ.

3. For every $i \in \{1, \ldots, \ell\}$: $G[\{v_i, v'_1, \ldots, v'_{i-1}, v'_{i+1}, \ldots, v'_\ell\}]$ is a complete graph K_ℓ.

It is easy to see that such a graph H is $K_{\ell+1}$-free. Moreover, at least $\ell+1$ colors are necessary to color it: Assume that there is a legal ℓ-coloring. Then $v'_1, \ldots, v'_\ell$ must be colored with all ℓ different colors. Hence, v_i and v'_i must receive the same color for all $i \in \{1, \ldots, \ell\}$. Since v is connected to all vertices $v_1, \ldots, v_\ell$ we obtain a contradiction.

We construct pairwise different graphs $G = (V, E)$ in $\mathcal{P}_i^\ell(n)$ as follows: On the first $i(2\ell+1)$ vertices i disjoint copies of H are fixed. Then we choose a graph $G' \in \mathcal{P}_0^\ell(n-i(2\ell+1))$ for the remaining vertices. ∎

B. Size of the Bad Sets

In order to exploit the fact that we consider only $K_{\ell+1}$-free graphs we want to identify dense subgraphs and parts of the graph with many edges incident to such a dense subgraph. If we have found a dense subgraph we can conclude that it may be connected to the rest of the graph only in a quite restricted way, since otherwise a $K_{\ell+1}$ would be constructed.

We assume that

$$\varepsilon_1, \varepsilon_{\mathcal{A}}, \varepsilon_{\mathcal{D}}, \varepsilon'_{\mathcal{D}}, \varepsilon_{\mathcal{E}}, \text{ and } \varepsilon_{\mathcal{R}}$$

are positive constants. Later we will explain how these constants have to be chosen.

Definition 3.4. *A $q(k)$-set $Q = \bigcup_{j=1}^{k} Q_j$ in a graph $G = (V, E)$ on $n = |V|$ vertices consists of k pairwise disjoint vertex sets $Q_1, ..., Q_k \subseteq V$ such that $|Q_1| = \ldots = |Q_k| = f_k(n) := \lceil \log^{(k^2-k+2)}(n) \rceil$ and $\{x, y\} \in E$ for all $x \in Q_h, y \in Q_j$ with $1 \leq h < j \leq k$.*

Given a $q(k)$-set $Q = \bigcup_{j=1}^{k} Q_j$ we denote by $R(Q)$ the set

$$R(Q) := \{v \in V \setminus Q \mid |\Gamma(v) \cap Q_j| \geq \varepsilon_{\mathcal{R}} |Q_j| \text{ for all } j = 1, \ldots, k\} .$$

For ease of notation we assume that a $q(0)$-set denotes the empty set and that $R(\emptyset) = V$. We also use the term q-set as short hand for a $q(\ell - 1)$-set.

In the sequel we assume without loss of generality that the constants ε_1, $\varepsilon_{\mathcal{A}}$, $\varepsilon_{\mathcal{D}}$, $\varepsilon'_{\mathcal{D}}$, $\varepsilon_{\mathcal{E}}$, and $\varepsilon_{\mathcal{R}}$ are chosen such that they satisfy the following inequalities for all sufficiently large n and arbitrary k and w with $1 \leq k < \ell$ and $\frac{n}{4\ell} \leq w \leq n$:

1.	$k\varepsilon_{\mathcal{R}} f_k(n) \binom{f_k(n)}{\varepsilon_{\mathcal{R}} f_k(n)} \leq 2^{\frac{1}{4}\varepsilon_{\mathcal{A}} f_k(n)}$	**6.**	$\varepsilon'_{\mathcal{D}} \leq \frac{1}{32\ell}$
2.	$3\varepsilon_{\mathcal{A}} \leq \varepsilon_1 \frac{1}{16\ell}$	**7.**	$2\ell\varepsilon_{\mathcal{A}} \leq \frac{1}{64\ell}$
3.	$\binom{w}{\frac{w}{2} - \varepsilon_{\mathcal{D}} n} \leq 2^{w - \varepsilon_1 n}$	**8.**	$2\ell\varepsilon_{\mathcal{A}} \leq \frac{\varepsilon_{\mathcal{E}}}{2^{\ell+1}}$
4.	$\frac{1}{2}\varepsilon_{\mathcal{D}} \leq \varepsilon'_{\mathcal{D}}$	**9.**	$\varepsilon_{\mathcal{E}} \leq \frac{1}{\ell} - (\ell - 1)\varepsilon_{\mathcal{A}}$
5.	$H\left(\frac{\frac{1}{2} - \varepsilon'_{\mathcal{D}}}{1 - \varepsilon_{\mathcal{D}}}\right) \leq 1 - \varepsilon_1$	**10.**	$2\varepsilon_{\mathcal{A}} + 4\ell\varepsilon_{\mathcal{E}} + \ell^2\varepsilon_{\mathcal{E}} \leq \frac{1}{10\ell}$

In order to see that this set of inequalities has a solution consider the ordering

$$\sigma := (\varepsilon'_{\mathcal{D}}, \varepsilon_{\mathcal{D}}, \varepsilon_1, \varepsilon_{\mathcal{E}}, \varepsilon_{\mathcal{A}}, \varepsilon_{\mathcal{R}}) .$$

If the constants are fixed according to σ it is quite straightforward to check that the inequalities can be satisfied by choosing values that are sufficiently small.

Using these constants we will now define a couple of bad sets. For every bad set two things we will done: First, the cardinality of the bad set is compared to the cardinality of (quasi-)ℓ-partite graphs. In Section C. it will turn out that these bounds are strong enough. Additionally, structural results are shown for the graphs in $\mathcal{F}_i^\ell(n) \setminus \mathcal{P}_i^\ell(n)$ that don't belong to the bad sets defined so far. As we define more and more bad sets we obtain more and more knowledge about the structure of the remaining graphs. In the end we will derive contradictory results on this structure and, thus, no such graphs remain. This means that we have covered the set of graphs in $\mathcal{F}_i^\ell(n) \setminus \mathcal{P}_i^\ell(n)$ with the bad sets.

Before going into details we will explain the intuition behind the bad sets: Assume that Theorem 2.2 holds. Hence, most graphs in $\mathcal{F}_i^\ell(n)$ look like a typical graph in $\mathcal{P}_i^\ell(n)$. Without a formal proof we claim that the structure of a typical i-quasi-ℓ-partite graph is similar to the structure of a random ℓ-partite graph with i additional vertices (connected to the rest of the graphs in a suitable way). Note that the partitions of random ℓ-partite graphs have approximately equal cardinality, i.e., they contain about $\frac{n}{\ell}$ vertices, and edges between the partitions occur independently with probability about $\frac{1}{2}$.

The bad sets are characterized by properties which violate this intuition. Hence, we expect that these graphs occur very 'unlikely'. The lemmas on the cardinality of the bad sets verify this intuitive argument.

Now we define the first two types of bad sets: Consider a typical graph in $\mathcal{F}_i^\ell(n)$ and a set X of vertices that belongs to k color classes. Assume that $|X|$ is not too small and that there are many vertices of each of the k colors. We expect for every vertex $v \in V \setminus X$ which does not belong to one of the k color classes of X that $|\Gamma(v) \cap X| \approx \frac{1}{2}|X|$, because this is the typical situation we would encounter in a random graph. We call this intuitive argument the *neighborhood assumption.*

A $q(k)$-set Q belongs to k color classes. Hence, most of the about $\frac{\ell-k}{\ell}n$ vertices in the other $\ell - k$ color classes should have approximately $\frac{1}{2}|Q|$ neighbors in Q and, thus, $|R(Q)| \approx \frac{\ell-k}{\ell}n$, i.e. $R(Q)$ contains almost all the other $\ell - k$ color classes. This intuition is violated by the definition of the sets $\mathcal{A}_i^\ell(n,k)$. The intuition behind $\mathcal{B}_i^\ell(n,k)$ is similar.

Definition 3.5. *For $k \in \{1,\ldots,\ell-1\}$ and $i \geq 0$ we define the following sets of graphs*

$$\begin{aligned}
\mathcal{A}_i^\ell(n,k) &:= \{G \in \mathcal{F}_i^\ell(n) \setminus \mathcal{P}_i^\ell(n) \mid \exists\, q(k)\text{-set } Q : |R(Q)| \leq (\tfrac{\ell-k}{\ell} - \varepsilon_{\mathcal{A}})n\}\ , \\
&\text{and for } k \in \{0,\ldots,\ell-2\} \text{ let} \\
\mathcal{B}_i^\ell(n,k) &:= \{G \in \mathcal{F}_i^\ell(n) \setminus \mathcal{P}_i^\ell(n) \mid \exists\, v \in V,\ \exists q(k)\text{-set } Q\ :\ Q \subseteq \Gamma(v), \\
&\qquad |R(Q)| \geq (\tfrac{\ell-k}{\ell} - \varepsilon_{\mathcal{A}})n, |R(Q) \cap \Gamma(v)| \leq f_k(n)\}\ .
\end{aligned}$$

Lemma 3.6. *For $k \in \{1,\ldots,\ell-1\}$, $i \geq 0$ and n sufficiently large it holds that*

$$\log \frac{|\mathcal{A}_i^\ell(n,k)|}{|\mathcal{F}_0^\ell(n - kf_k(n))|} \leq \left(\frac{\ell-1}{\ell} - \frac{\varepsilon_{\mathcal{A}}}{2\ell}\right) \cdot kf_k(n) \cdot n\ .$$

Proof. We construct all graphs $G = (V,E) \in \mathcal{A}_i^\ell(n,k)$ as follows: First we choose k sets $Q_1,...,Q_k \subseteq V$ which form a $q(k)$-set $Q = Q_1 \cup \ldots \cup Q_k$ (at most $\binom{n}{f_k(n)}^k \leq n^{kf_k(n)}$ possibilities) and the edges inside Q (at most $2^{(kf_k(n))^2}$). Then we fix a graph on the remaining $n - kf_k(n)$ vertices in $V \setminus Q$ (at most $|\mathcal{F}_0^\ell(n - kf_k(n))|$ possibilities) and a suitable set $R(Q) \subseteq V \setminus Q$ (at most 2^n possibilities). Finally, we connect Q to $R(Q)$ (at most $2^{kf_k(n)\cdot|R(Q)|}$ possibilities) and Q to $V \setminus [Q \cup R(Q)]$: To this aim we choose for every vertex $x \in V \setminus [Q \cup R(Q)]$ a set Q_{i_0} with $|Q_{i_0} \cap \Gamma(x)| \leq \varepsilon_{\mathcal{R}}|Q_{i_0}|$. Then we fix these neighbors and connect x in an arbitrary way to all other sets Q_j.

Hence, for the edges from Q to $V \setminus [Q \cup R(Q)]$ there are at most

$$\left[k \cdot \sum_{j=0}^{\varepsilon_{\mathcal{R}} f_k(n)} \binom{f_k(n)}{j} \cdot 2^{(k-1)f_k(n))}\right]^{(n-|R(Q)|-kf_k(n))} .$$

possibilities. By Inequality 1 this expression is less than

$$2^{(\frac{1}{4}\varepsilon_{\mathcal{A}} f_k(n)+(k-1)f_k(n))(n-|R(Q)|)} .$$

Thus, we deduce that

$$\begin{aligned}
\log \frac{|\mathcal{A}_i^\ell(n,k)|}{|\mathcal{F}_0^\ell(n-kf_k(n))|} &\leq kf_k(n)\log n + (kf_k(n))^2 + n + kf_k(n) \cdot |R(Q)| \\
&\quad + \left(\frac{1}{4}\varepsilon_{\mathcal{A}} f_k(n) + (k-1)f_k(n)\right)(n - |R(Q)|) \\
&\leq \left((k-1)f_k(n) + \frac{1}{2}\varepsilon_{\mathcal{A}} f_k(n)\right) \cdot n + |R(Q)| \cdot f_k(n) \\
&\leq \left(k - 1 + \frac{1}{2}\varepsilon_{\mathcal{A}} + \frac{\ell - k}{\ell} - \varepsilon_{\mathcal{A}}\right) \cdot f_k(n) \cdot n \\
&\leq \left(k - \frac{k}{\ell} - \frac{1}{2}\varepsilon_{\mathcal{A}}\right) \cdot f_k(n) \cdot n .
\end{aligned}$$

■

In the next proofs we will repeatedly use the following observation: Let $G = (V,E)$ be a graph with $G \in \mathcal{F}_i^\ell(n) \setminus \mathcal{P}_i^\ell(n)$. Now delete an arbitrary vertex $v \in V$ and consider the graph $G' := G[V \setminus \{v\}]$. Then $G' \in \mathcal{F}_i^\ell(n-1)$. This follows directly from the fact that $G \notin \mathcal{P}_0^\ell(n) \cup \ldots \cup \mathcal{P}_{i-1}^\ell(n) \cup \mathcal{P}_i^\ell(n)$ by definition and, thus, $G' = G[V \setminus \{v\}] \notin \mathcal{P}_0^\ell(n-1) \cup \ldots \cup \mathcal{P}_{i-1}^\ell(n-1)$.

Lemma 3.7. *For $k \in \{0, \ldots, \ell-2\}$, $i \geq 0$ and n sufficiently large it holds that*

$$\log \frac{|\mathcal{B}_i^\ell(n,k)|}{|\mathcal{F}_i^\ell(n-1)|} \leq \left(\frac{\ell-1}{\ell} - \varepsilon_{\mathcal{A}}\right) \cdot n .$$

Proof. We construct all graphs $G = (V,E) \in \mathcal{B}_i^\ell(n,k)$ as follows: First we choose a vertex $v \in V$ and a graph $G' \in \mathcal{F}_i^\ell(n-1)$ on $V \setminus \{v\}$ (at most $n|\mathcal{F}_i^\ell(n-1)|$ possibilities). Then we fix a suitable $q(k)$-set $Q \subseteq V \setminus \{v\}$ (at most $n^{kf_k(n)}$ possibilities) and connect v to $R(Q)$ (at most $f_k(n) \cdot \binom{|R(Q)|}{f_k(n)} \leq n^{f_k(n)}$ possibilities). Finally, we connect v to $V \setminus [\{v\} \cup R(Q)]$ (at most $2^{n-|R(Q)|}$ possibilities).

All in all, we obtain

$$\log \frac{|\mathcal{B}_i^\ell(n,k)|}{|\mathcal{F}_i^\ell(n-1)|} \leq \log n + (k+1)f_k(n) \cdot \log n + n - |R(Q)|$$

$$\leq \left(\frac{k}{\ell} + 2\varepsilon_{\mathcal{A}}\right) \cdot n \leq \left(\frac{\ell-1}{\ell} - \varepsilon_{\mathcal{A}}\right) \cdot n\ .$$

■

The exclusion of the sets $\mathcal{A}_i^\ell(n,k)$ and $\mathcal{B}_i^\ell(n,k)$ already enables us to obtain our first structural corollary.

Corollary 3.8. *Let $\mathcal{A}_i^\ell(n) = \bigcup_{k=1}^{\ell-1} \mathcal{A}_i^\ell(n,k)$ and $\mathcal{B}_i^\ell(n) = \bigcup_{k=0}^{\ell-2} \mathcal{B}_i^\ell(n,k)$. For n sufficiently large, all graphs $G = (V,E) \in \mathcal{F}_i^\ell(n) \setminus [\mathcal{P}_i^\ell(n) \cup \mathcal{A}_i^\ell(n) \cup \mathcal{B}_i^\ell(n)]$ have the property that for every vertex $v \in V$ there exists a q-set Q_v with $Q_v \subseteq \Gamma(v)$.*

The proof of Corollary 3.8 is completely analogous to the corresponding result (Lemma 1.7) in [3]. In particular, we also employ the following lemma from [3] which generalizes a result of Bollobás and Erdős [1]:

Lemma 3.9 [3]. *Let $0 < \varepsilon < 1$ and $k \geq 0$. Then for all $N \geq N_0(\varepsilon,k)$ the following holds. If G is a graph with vertex set $A_0 \cup \ldots \cup A_k$, where A_j are pairwise disjoint sets of size $|A_j| = N$ and $|\Gamma(u) \cap A_j| \geq \varepsilon N$ for all $u \in A_0$ and all $j = 1,\ldots,k$, then there exist subsets $A_0' \subset A_0, \ldots, A_k' \subset A_k$ such that $|A_j'| = \lceil \log^{(2k)}(N) \rceil$ for all $j = 0,\ldots,k$ and such that A_0' is completely connected to all A_j', i.e., we have $\Gamma(u) \supseteq \bigcup_{j=1}^k A_j'$ for all $u \in A_0'$.*

We include a proof of Corollary 3.8 for sake of completeness and as we will also reuse the proof idea later on in the proof of Corollary 3.14.

Proof of Corollary 3.8. Consider an arbitrary graph $G = (V,E) \in \mathcal{F}_i^\ell(n) \setminus [\mathcal{P}_i^\ell(n) \cup \mathcal{A}_i^\ell(n) \cup \mathcal{B}_i^\ell(n))]$ and an arbitrary vertex $v \in V$. We will show that $\Gamma(v)$ contains a q-set.

Because of $G \notin \mathcal{B}_i^\ell(n,0)$ we conclude that $|\Gamma(v)| \geq f_0(n) \geq f_1(n)$ and, thus, $\Gamma(v)$ contains a $q(1)$-set Q of size $f_1(n)$. We finish the proof by induction on k. Hence, assume that we have already found a $q(k)$-set $Q = \bigcup_{j=1}^k Q_j$ in $\Gamma(v)$ for a certain k with $1 \leq k \leq \ell - 2$.

Since $G \notin \mathcal{A}_i^\ell(n,k)$ it holds that $|R(Q)| \geq (\frac{\ell-k}{\ell} - \varepsilon_{\mathcal{A}})n$ and due to $G \notin \mathcal{B}_i^\ell(n,k)$ we deduce that $|\Gamma(v) \cap R(Q)| > f_k(n)$. Thus, let A_0 denote an arbitrary subset of $\Gamma(v) \cap R(Q)$ with $|A_0| = f_k(n)$. By applying Lemma 3.9 on A_0 and $A_j = Q_j, 1 \leq i \leq k$ we obtain sets $Q_1', \ldots, Q_{k+1}' \subseteq \Gamma(v)$ which form a $q(k+1)$-set since

$$\begin{aligned} \lceil \log^{(2k)}(f_k(n)) \rceil &= \lceil \log^{(2k)}(\lceil \log^{(k^2-k+2)}(n) \rceil) \rceil \geq \lceil \log^{((k+1)^2-(k+1)+2)}(n) \rceil \\ &= f_{k+1}(n)\ . \end{aligned}$$

■

Up to now we have only excluded graphs which contain a q-set Q for which $R(Q)$ is too small. The next definition also excludes graphs for which $R(Q)$ is too large.

Definition 3.10. *Let $\mathcal{C}_i^\ell(n)$ denote the set of all graphs $G = (V,E)$ in $\mathcal{F}_i^\ell(n) \setminus [\mathcal{P}_i^\ell(n) \cup \mathcal{A}_i^\ell(n) \cup \mathcal{B}_i^\ell(n)]$ that contain a vertex v and a q-set $Q = \bigcup_{j=1}^{\ell-1} Q_j \subseteq \Gamma(v)$, such that $|R(Q)| \geq (\frac{1}{\ell} + \varepsilon_{\mathcal{A}})n$.*

Lemma 3.11. *For all $i > 0$ and n sufficiently large it holds that*

$$\log \frac{|\mathcal{C}_i^\ell(n)|}{|\mathcal{F}_i^\ell(n-1)|} \leq \left(\frac{\ell-1}{\ell} - \frac{1}{2}\varepsilon_{\mathcal{A}}\right) \cdot n \,.$$

Proof. We construct all graphs $G = (V, E) \in \mathcal{C}_i^\ell(n)$ as follows: First we choose a vertex $v \in V$ and a graph $G' \in \mathcal{F}_i^\ell(n-1)$ on $V \setminus \{v\}$ (at most $n|\mathcal{F}_i^\ell(n-1)|$ possibilities). Then we choose an appropriate set Q in $V \setminus \{v\}$ (at most $n^{(\ell-1)f_{\ell-1}(n)}$ possibilities).

Note that v cannot be connected to any vertex in $R(Q)$, as otherwise G would contain a $K_{\ell+1}$. Therefore, if $|R(Q)| \geq (\frac{1}{\ell} + \varepsilon_{\mathcal{A}})n$ then there are at most $2^{(\frac{\ell-1}{\ell} - \varepsilon_{\mathcal{A}})n}$ possibilities to connect v to $V \setminus \{v\}$. It follows that

$$\begin{aligned} \log \frac{|\mathcal{C}_i^\ell(n)|}{|\mathcal{F}_i^\ell(n-1)|} &\leq \log n + (\ell-1) f_{\ell-1}(n) \cdot \log n + \frac{\ell-1}{\ell} n - \varepsilon_{\mathcal{A}} n \\ &\leq \left(\frac{\ell-1}{\ell} - \frac{1}{2}\varepsilon_{\mathcal{A}}\right) \cdot n \,. \end{aligned}$$

∎

Before we derive our next structural corollary we first introduce another bad set. The bad set $\mathcal{D}_i^\ell(n)$ is once again based on the neighborhood assumption.

Definition 3.12. *Let $\mathcal{D}_i^\ell(n)$ denote the set of all graphs $G = (V, E)$ in $\mathcal{F}_i^\ell(n) \setminus [\mathcal{P}_i^\ell(n) \cup \mathcal{A}_i^\ell(n) \cup \mathcal{B}_i^\ell(n) \cup \mathcal{C}_i^\ell(n)]$ that contain three vertices u, v, w and q-sets $Q_u \subseteq \Gamma(u)$, $Q_v \subseteq \Gamma(v)$, $Q_w \subseteq \Gamma(w)$ such that the following conditions are satisfied for $W := R(Q_w) \setminus [R(Q_u) \cup R(Q_v)]$. $|W| \geq \frac{1}{4\ell} n$ and*

- *(i)* $|\Gamma(u) \cap W| \leq \frac{|W|}{2} - \varepsilon_{\mathcal{D}} n$, *or*
- *(ii)* $|\Gamma(u) \cap \Gamma(v) \cap W| \leq \frac{|W|}{4} - \varepsilon'_{\mathcal{D}} n$.

Lemma 3.13. *For all $i \geq 0$ and n sufficiently large it holds that*

$$\log \frac{|\mathcal{D}_i^\ell(n)|}{|\mathcal{F}_i^\ell(n-1)|} \leq \left(\frac{\ell-1}{\ell} - \varepsilon_{\mathcal{A}}\right) \cdot n \,.$$

Proof. We construct all graphs in $\mathcal{D}_i^\ell(n)$ as follows: First we choose the vertex u and a graph $G' \in \mathcal{F}_i^\ell(n-1)$ for the remaining vertices. Then we fix the vertices v and w and the q-sets Q_u, Q_v and Q_w. There are at most

$$n^3 \cdot \binom{n}{(\ell-1)f_{\ell-1}(n)}^3 \cdot |\mathcal{F}_i^\ell(n-1)| \leq n^{4\log n} \cdot |\mathcal{F}_i^\ell(n-1)|$$

possibilities for that.

Assume that Condition (i) of Definition 3.12 holds. Note that $|R(Q_u)| \geq (\frac{1}{\ell} - \varepsilon_{\mathcal{A}})n$ since $G \notin \mathcal{A}_i^\ell(n)$. This implies that the number of choices for the edges from

u to $V \setminus [R(Q_u) \cup W]$ is bounded from above by

$$2^{n-|R(Q_u)|-|W|} \leq 2^{(1-\frac{1}{\ell}+\varepsilon_{\mathcal{A}})n-|W|} .$$

Furthermore, there are at most

$$\sum_{j=0}^{\frac{1}{2}|W|-\varepsilon_{\mathcal{D}}n} \binom{|W|}{j} \leq |W| \cdot \binom{|W|}{\frac{1}{2}|W|-\varepsilon_{\mathcal{D}}n} \leq n \cdot 2^{|W|-\varepsilon_1 n}$$

possibilities for the edges from u to W, where the last estimate follows by Inequality 3. All in all, we obtain the following upper bound on the number of choices:

$$n^{4\log n} \cdot |\mathcal{F}_i^\ell(n-1)| \cdot 2^{(1-\frac{1}{\ell}+\varepsilon_{\mathcal{A}}-\varepsilon_1)n} .$$

If, on the other hand, Condition (ii) of Definition 3.12 holds we may safely assume that Condition (i) does not hold for any choice of the vertices u, v and w and their q-sets. Hence, it follows that $|\Gamma(v) \cap W| =: \xi \geq \frac{1}{2}|W| - \varepsilon_{\mathcal{D}}n \geq (\frac{1}{8\ell} - \varepsilon_{\mathcal{D}})n$. First we fix the edges from u to $V \setminus [R(Q_u) \cup (\Gamma(v) \cap W)]$. The number of choices is bounded from above by

$$2^{n-|R(Q_u)|-|\Gamma(v)\cap W|} \leq 2^{(1-\frac{1}{\ell}+\varepsilon_{\mathcal{A}})n-\xi} .$$

For the edges from u to $\Gamma(v) \cap W$ there are at most

$$\sum_{j=0}^{\frac{1}{4}|W|-\varepsilon'_{\mathcal{D}}n} \binom{\xi}{j} \leq n \cdot \binom{\xi}{\frac{1}{4}|W|-\varepsilon'_{\mathcal{D}}n} \leq n \cdot 2^{\xi \cdot H((\frac{1}{4}|W|-\varepsilon'_{\mathcal{D}}n)/\xi)}$$

possibilities. For $\varepsilon'_{\mathcal{D}} \geq \frac{1}{2}\varepsilon_{\mathcal{D}}$ (cf. Inequality 4) one easily checks that

$$\frac{\frac{1}{4}|W|-\varepsilon'_{\mathcal{D}}n}{\xi} \leq \frac{\frac{1}{4}|W|-\varepsilon'_{\mathcal{D}}n}{\frac{1}{2}|W|-\varepsilon_{\mathcal{D}}n} \leq \frac{(\frac{1}{2}-\varepsilon'_{\mathcal{D}})n}{(1-\varepsilon_{\mathcal{D}})n} \leq \frac{1}{2} .$$

Thus,

$$H\left(\frac{\frac{1}{4}|W|-\varepsilon'_{\mathcal{D}}n}{\xi}\right) \leq H\left(\frac{(\frac{1}{2}-\varepsilon'_{\mathcal{D}})n}{(1-\varepsilon_{\mathcal{D}})n}\right) \leq 1-\varepsilon_1$$

by Inequality 5. The total number of choices is therefore bounded from above by

$$\begin{aligned}
&n^{4\log n} \cdot |\mathcal{F}_i^\ell(n-1)| \cdot 2^{(1-\frac{1}{\ell}+\varepsilon_{\mathcal{A}})n-\varepsilon_1\xi} \\
&\leq\ n^{4\log n} \cdot |\mathcal{F}_i^\ell(n-1)| \cdot 2^{(1-\frac{1}{\ell}+\varepsilon_{\mathcal{A}}-\varepsilon_1(\frac{1}{8\ell}-\varepsilon_{\mathcal{D}}))n} \\
&\leq\ n^{4\log n} \cdot |\mathcal{F}_i^\ell(n-1)| \cdot 2^{(1-\frac{1}{\ell}+\varepsilon_{\mathcal{A}}-\varepsilon_1\frac{1}{16\ell})n} ,
\end{aligned}$$

applying Inequality 6. Using Inequality 2 we can combine the two cases and obtain

$$\log\frac{|\mathcal{D}_i^\ell(n)|}{|\mathcal{F}_i^\ell(n-1)|} \ \leq\ 1+4(\log n)^2+n\cdot\left(1-\frac{1}{\ell}+\varepsilon_{\mathcal{A}}-\varepsilon_1\frac{1}{16\ell}\right)$$

$$\leq\ 5(\log n)^2 + n \cdot \left(1 - \frac{1}{\ell} + \varepsilon_{\mathcal{A}} - 3\varepsilon_{\mathcal{A}}\right) \leq n \cdot \left(1 - \frac{1}{\ell} - \varepsilon_{\mathcal{A}}\right) .$$

■

Now we are in a position to prove our next structural corollary. It shows that all graphs which have not yet been excluded remain i-quasi-ℓ-partite after the deletion of at most ℓ vertices.

Corollary 3.14. *For all graphs $G \in \mathcal{F}_i^\ell(n) \setminus [\mathcal{P}_i^\ell(n) \cup \mathcal{A}_i^\ell(n) \cup \mathcal{B}_i^\ell(n) \cup \mathcal{C}_i^\ell(n) \cup \mathcal{D}_i^\ell(n)]$ and arbitrary vertices $V' \subseteq V$ with $|V'| =: k \leq \ell$ it holds that $G' := G[V \setminus V'] \in \mathcal{F}_i^\ell(n-k)$.*

Proof. Consider a graph $G \in \mathcal{F}_i^\ell(n) \setminus [\mathcal{P}_i^\ell(n) \cup \mathcal{A}_i^\ell(n) \cup \mathcal{B}_i^\ell(n) \cup \mathcal{C}_i^\ell(n) \cup \mathcal{D}_i^\ell(n)]$. It suffices to show that $G \notin \bigcup_{j=i+1}^{\ell+i} \mathcal{P}_j^\ell(n)$ because if this condition holds at least $\ell + i + 1$ vertices must be deleted in order to make the graph ℓ-partite.

Assume that $G = (V, E) \in \mathcal{P}_j^\ell(n)$ for some $j \in \{i+1, \ldots, \ell+i\}$. Then we can find vertices $v_1, \ldots, v_j$ such that $G' = (V', E') := G[V \setminus \{v_1, \ldots, v_j\}] \in \mathcal{P}_0^\ell(n)$. Let $S_1, \ldots, S_\ell$ be a valid ℓ-coloring of G'.

Consider an arbitrary vertex $x \in S_1$ with q-set $Q_x \subseteq \Gamma(x)$. It is easy to see that $R(Q_x) \subseteq S_1 \cup \{v_1, \ldots, v_j\}$. Hence, $|S_1| \geq (\frac{1}{\ell} - 2\varepsilon_{\mathcal{A}})n$ since $G \notin \mathcal{A}_i^\ell(n)$. By symmetry this also holds for $S_2, \ldots, S_\ell$.

For brevity let $v_1 =: v$. By Q_v we denote a q-set with $Q_v \subseteq \Gamma(v)$ which exists due to Corollary 3.8. Without loss of generality we assume that S_ℓ is a color class where the cardinality of the intersection with $R(Q_v)$ is maximum. By $G \notin \mathcal{C}_i^\ell(n)$ it follows that $|R(Q_v) \cap S_k| \leq \frac{1}{2}(\frac{1}{\ell} + \varepsilon_{\mathcal{A}})n$ for $k = 1, \ldots, \ell - 1$.

v must have a neighbor u in S_ℓ. Otherwise, G could be made ℓ-partite by removing just $v_2, \ldots, v_j$. Let $Q_u \subseteq \Gamma(u)$ be a q-set. As we have already seen, $R(Q_u) \subseteq S_\ell \cup \{v_1, \ldots, v_j\}$.

For the color classes $S_1, \ldots, S_{\ell-1}$ we fix vertices $z_1, \ldots, z_{\ell-1}$ with $z_k \in S_k$ and q-sets $Q_{z_k} \subseteq \Gamma(z_k)$ for $k = 1, \ldots, \ell - 1$. Note that $R(Q_{z_k}) \subseteq S_k \cup \{v_1, \ldots, v_j\}$. Furthermore, we define $W_k := R(Q_{z_k}) \setminus [R(Q_u) \cup R(Q_v)]$. Since $R(Q_{z_k}) \cap R(Q_u) \subseteq \{v_1, \ldots, v_j\}$ and $R(Q_v) \cap R(Q_{z_k}) \subseteq (R(Q_v) \cap S_k) \cup \{v_1, \ldots, v_j\}$ it holds that

$$\begin{aligned} |W_k| &\geq |R(Q_{z_k})| - j - |R(Q_v) \cap S_k| \geq \left(\frac{1}{\ell} - \varepsilon_{\mathcal{A}}\right) n - j - \frac{1}{2}\left(\frac{1}{\ell} + \varepsilon_{\mathcal{A}}\right) n \\ &\geq \left(\frac{1}{2\ell} - 2\varepsilon_{\mathcal{A}}\right) n \geq \frac{1}{4\ell} \cdot n , \end{aligned}$$

where the last inequality is a consequence of Inequality 7. Since $G \notin \mathcal{D}_i^\ell(n)$ and because of Inequality 6 it follows that

$$|\Gamma(u) \cap \Gamma(v) \cap W_k| \geq \left(\frac{1}{16\ell} - \varepsilon'_{\mathcal{D}}\right) n \geq \frac{1}{32\ell} \cdot n$$

Using the same approach as in the proof of Corollary 3.8 we construct a $K_{\ell-1}$ in $\Gamma(u) \cap \Gamma(v)$: By induction on k we show that $\Gamma(u) \cap \Gamma(v) \cap (S_1 \cup \ldots \cup S_k) =: N_k$ contains a $q(k)$-set. The base case $k = 0$ is trivial.

Now assume that we have already found a $q(k)$-set Q with $Q \subseteq N_k$. Consider the set $L := [\Gamma(u) \cap \Gamma(v) \cap S_{k+1}] \setminus R(Q)$. $|L|$ is small because, clearly, $R(Q) \cap (S_1 \cup \ldots \cup S_k) = \emptyset$ and, therefore,

$$\begin{aligned} |V| = n &\geq |S_1| + \ldots + |S_k| + |R(Q)| + |L| \\ &\geq \left(\frac{1}{\ell} - 2\varepsilon_{\mathcal{A}}\right) \cdot k \cdot n + \left(\frac{\ell - k}{\ell} - \varepsilon_{\mathcal{A}}\right) n + |L| \\ &= (1 - 2k\varepsilon_{\mathcal{A}} - \varepsilon_{\mathcal{A}})n + |L| \end{aligned}$$

and, thus, $|L| \leq 2(k+1)\varepsilon_{\mathcal{A}} n \leq 2\ell\varepsilon_{\mathcal{A}} n \leq \frac{1}{64\ell} \cdot n$ by Inequality 7. This implies that $|\Gamma(u) \cap \Gamma(v) \cap S_{k+1} \cap R(Q)| = |\Gamma(u) \cap \Gamma(v) \cap S_{k+1}| - |L| \geq \frac{1}{64\ell} \cdot n$. Hence, as in the proof of Corollary 3.8 we can construct a $q(k)$-set within N_k.

This yields a contradiction because we can easily find a $K_{\ell+1}$ composed of the vertices u, v and an arbitrary $K_{\ell-1}$ taken from the $q(\ell-1)$-set within $N_{\ell-1} \subseteq \Gamma(u) \cap \Gamma(v)$. ∎

Finally, we define our last type of bad sets and derive a bound on their cardinality.

Definition 3.15. *For $k \in \{1, \ldots, l\}$ let $\mathcal{E}_i(n,k)$ denote the set of all graphs $G = (V,E) \in \mathcal{F}_i^\ell(n) \setminus [\mathcal{P}_i^\ell(n) \cup \mathcal{A}_i^\ell(n) \cup \mathcal{B}_i^\ell(n) \cup \mathcal{C}_i^\ell(n) \cup \mathcal{D}_i^\ell(n)]$ that contain a k-clique $S_k = \{v_1, \ldots, v_k\}$ such that $G' := G[V \setminus S_k] \in \mathcal{F}_i^\ell(n-k)$ with*

$$S_k \text{ is not contained in a } k+1\text{-clique}\,, \tag{3.1}$$

and for each vertex v_j there exists a q-set $Q_j \subseteq \Gamma(v_j)$ such that $|R(Q_j)| \geq \frac{n}{l} - \varepsilon_{\mathcal{A}} n$ and

$$U := V \setminus \left(S_k \cup \bigcup_{j=1}^{k} R(Q_j) \right)$$

satisfies

$$|U| \geq \varepsilon_{\mathcal{E}} n\,. \tag{3.2}$$

Lemma 3.16. *For $k \in \{1, \ldots, l\}$, $i \geq 0$ and n sufficiently large it holds that*

$$\log \frac{|\mathcal{E}_i(n,k)|}{|\mathcal{F}_i^\ell(n-k)|} \leq \frac{\ell-1}{\ell} nk - \frac{\varepsilon_{\mathcal{E}}}{2^{\ell+1}} n\,.$$

Proof. We construct all graphs $G = (V,E)$ in $\mathcal{E}_i(n,k)$ as follows: First we choose $S_k = \{v_1, \ldots, v_k\}$ and a suitable graph $G' \in \mathcal{F}_i^\ell(n-k)$ on the vertices $V \setminus S_k$ (at most $n^k \cdot |\mathcal{F}_i^\ell(n-k)|$ possibilities).

Now we choose the q-sets $Q_1, \ldots, Q_k \subseteq V$ such that the corresponding sets $R(Q_j)$ satisfy the conditions of Definition 3.15 (at most $n^{(\ell-1)kf_{\ell-1}(n)}$ possibilities).

Finally, we connect S_k to $V \setminus S_k$. For every $1 \leq h \leq k$ and every h-tupel $1 \leq i_1 < \ldots < i_h \leq k$ we define the set $T_{i_1,\ldots,i_h} = (R(Q_{i_1}) \cap \ldots \cap R(Q_{i_h})) \setminus \bigcup_{j \notin \{i_1,\ldots,i_h\}} R(Q_j)$.

Then the following holds for the set U by definition:

$$U = V \setminus \left(\bigcup_{h=1}^{k} \bigcup_{i_1,\dots,i_h} T_{i_1,\dots,i_h} \cup S_k \right) ,$$

The number of possibilities to connect S_k to $T_{i_1,\dots,i_h}$ is at most $2^{(k-h)|T_{i_1,\dots,i_h}|}$. To see this, observe that v_j and $R(Q_j)$ cannot be connected.

By assumption (3.1) of Definition 3.15 there are at most $(2^k-1)^{|U|} \le 2^{(k-2^{-k})\cdot|U|}$ possibilities to connect S_k to U. Putting all this together we obtain at most

$$2^{(k-2^{-k})\cdot|U|} \cdot 2^{\sum_{h=1}^{k}\sum_{i_1,\dots,i_h}|T_{i_1,\dots,i_h}|(k-h)}$$

possible choices for the connections from S_k to $V \setminus S_k$. It holds that

$$\begin{aligned}
\sum_{h=1}^{k}\sum_{i_1,\dots,i_h}|T_{i_1,\dots,i_h}|(k-h) &= k\left(\sum_{h=1}^{k}\sum_{i_1,\dots,i_h}|T_{i_1,\dots,i_h}|\right) - \sum_{h=1}^{k}\sum_{i_1,\dots,i_h}|T_{i_1,\dots,i_h}|h \\
&= k|V\setminus U| - \sum_{j=1}^{k}|R(Q_j)| \\
&\le kn - k|U| - k\left(\frac{n}{l} - \varepsilon_{\mathcal{A}} n\right) \\
&= \frac{l-1}{l}nk - k|U| + k\varepsilon_{\mathcal{A}} n .
\end{aligned}$$

Now the lemma follows by Inequality 8:

$$\begin{aligned}
\log\frac{|\mathcal{E}_k(n)|}{|\mathcal{F}_i^\ell(n-k)|} &\le k\log n + (l-1)kf_{\ell-1}(n)\log n + (k-2^{-k})|U| \\
&\quad + \frac{l-1}{l}nk - k|U| + k\varepsilon_{\mathcal{A}} n \\
&\le \frac{l-1}{l}nk - 2^{-k}|U| + 2k\varepsilon_{\mathcal{A}} n \le \frac{l-1}{l}nk - \frac{\varepsilon_{\mathcal{E}}}{2^k}n + 2k\varepsilon_{\mathcal{A}} n \\
&\le \frac{l-1}{l}nk - \frac{1}{2^{l+1}}\varepsilon_{\mathcal{E}} n .
\end{aligned}$$

∎

Corollary 3.17. *Let $\mathcal{E}_i^\ell(n) := \bigcup_{k=1}^{l}\mathcal{E}_i^\ell(n,k)$. For all sufficiently large n, the graphs $G \in \mathcal{F}_i^\ell(n)\setminus[\mathcal{P}_i^\ell(n)\cup\mathcal{A}_i^\ell(n)\cup\mathcal{B}_i^\ell(n)\cup\mathcal{C}_i^\ell(n)\cup\mathcal{D}_i^\ell(n)\cup\mathcal{E}_i^\ell(n)]$ have the following property:*

For $1 \le k < \ell$ every clique of size k is contained in a clique of size ℓ.

Proof. Consider an arbitrary graph $G \in \mathcal{F}_i^\ell(n)\setminus[\mathcal{P}_i^\ell(n)\cup\mathcal{A}_i^\ell(n)\cup\mathcal{B}_i^\ell(n)\cup\mathcal{C}_i^\ell(n)\cup\mathcal{D}_i^\ell(n)\cup\mathcal{E}_i^\ell(n)]$. Assume that $S = \{v_1,\dots,v_k\}$ is a maximal clique in G of size

$1 \le k < \ell$. For every vertex $v_j \in S$ there is a q-set $Q_j \subseteq \Gamma(v_j)$. Since $G \notin \mathcal{C}_i^\ell(n)$ it holds that $\sum_{j=1}^k |R(Q_j)| \le k(\frac{1}{\ell} + \varepsilon_{\mathcal{A}})n$, i.e., we obtain for the set U of Lemma 3.16 that

$$|U| \ge \left(1 - \frac{k}{\ell} - k\varepsilon_{\mathcal{A}}\right) n \ge \left(\frac{1}{\ell} - (\ell-1)\varepsilon_{\mathcal{A}}\right) n \ge \varepsilon_{\mathcal{E}} n \ ,$$

where the last inequality is a consequence of Inequality 9. This contradicts the assumption that $G \notin \mathcal{E}_k^\ell(n)$. ∎

Now we have defined enough bad sets in order to cover all graphs in $\mathcal{F}_i^\ell(n) \setminus \mathcal{P}_i^\ell(n)$. This is shown in the following lemma.

Lemma 3.18. *For all sufficiently large n, $\mathcal{F}_i^\ell(n) \setminus [\mathcal{P}_i^\ell(n) \cup \mathcal{A}_i^\ell(n) \cup \mathcal{B}_i^\ell(n) \cup \mathcal{C}_i^\ell(n) \cup \mathcal{D}_i^\ell(n) \cup \mathcal{E}_i^\ell(n)]$ is empty.*

Proof. Assume there exists a graph $G = (V, E) \in \mathcal{F}_i^\ell(n) \setminus [\mathcal{P}_i^\ell(n) \cup \mathcal{A}_i^\ell(n) \cup \mathcal{B}_i^\ell(n) \cup \mathcal{C}_i^\ell(n) \cup \mathcal{D}_i^\ell(n) \cup \mathcal{E}_i^\ell(n)]$. We will show that G has to be ℓ-partite, which contradicts the definition of the set $\mathcal{F}_i^\ell(n)$.

Note that Corollary 3.8 implies that for every vertex $v \in V$ there is a q-set $Q_v \subseteq \Gamma(v)$. The corresponding sets $R(Q_v)$ are denoted by $R(v)$ for brevity. In the sequel we collect some auxiliary results on the structure of the graph G:

1. For all $v \in V$ it holds that $(\frac{1}{\ell} - \varepsilon_{\mathcal{A}}) \cdot n \le |R(v)| \le (\frac{1}{\ell} + \varepsilon_{\mathcal{A}}) \cdot n$.

This is an immediate consequence of the fact that $G \notin \mathcal{A}_i^\ell(n, \ell-1) \cup \mathcal{C}_i^\ell(n)$.

2. For $v, v' \in V$ with $\{v, v'\} \in E$ it holds that $|R(v) \cap R(v')| \le 2\varepsilon_{\mathcal{E}} n$.

This can be seen as follows: By Corollary 3.17 we deduce that v and v' belong to an ℓ-clique $S = \{v_1, \ldots, v_\ell\}$. Since this clique cannot be contained in an $(\ell+1)$-clique it follows by Definition 3.15 that

$$\left| \bigcup_{j=1}^{l} R(v_j) \right| \ge n - \varepsilon_{\mathcal{E}} n \ .$$

As $|R(v_j)| \le (\frac{1}{\ell} + \varepsilon_{\mathcal{A}})n$ for all $1 \le i \le \ell$ by (1) we conclude that

$$|R(v) \cap R(v')| \le (\ell\varepsilon_{\mathcal{A}} + \varepsilon_{\mathcal{E}})n \le 2\varepsilon_{\mathcal{E}} n \ .$$

3. For all $v \in V$ and $x \in R(v)$ it holds that $|R(v) \cap R(x)| \ge \frac{9}{10\ell} n$.

Since $x \in R(v)$ we conclude by the definition of $R(v)$ that there are vertices $v_1, \ldots, v_{\ell-1}$ in Q_v such that $v_1, \ldots, v_{\ell-1}, x$ as well as $v_1, \ldots, v_{\ell-1}, v$ form an ℓ-clique. Set $T := \bigcup_{j=1}^{\ell-1} R(v_j)$. By (1) and (2) we deduce using Inequality 8 that

$$\begin{aligned} |T| \ &\ge \ \left((\ell-1)\left(\frac{1}{\ell} - \varepsilon_{\mathcal{A}}\right) - \binom{\ell-1}{2} \cdot 2\varepsilon_{\mathcal{E}}\right) n \\ &\ge \ \left(\frac{\ell-1}{\ell} - (\ell-1)\varepsilon_{\mathcal{A}} - (\ell-1)\ell\varepsilon_{\mathcal{E}}\right) n \ge \left(\frac{\ell-1}{\ell} - \ell^2\varepsilon_{\mathcal{E}}\right) n \end{aligned}$$

and that

$$|R(x) \cap T| \leq 2\ell\varepsilon_{\mathcal{E}} n, \qquad |R(v) \cap T| \leq 2\ell\varepsilon_{\mathcal{E}} n \ .$$

This implies that

$$|R(x) \cup R(v)| \leq |R(x) \cap T| + |R(v) \cap T| + (V \setminus T) \leq \left(4\ell\varepsilon_{\mathcal{E}} + \frac{1}{\ell} + \ell^2\varepsilon_{\mathcal{E}}\right) n \ .$$

Due to $|R(v)|, |R(x)| \geq (\frac{1}{\ell} - \varepsilon_{\mathcal{A}})n$ it follows that

$$\begin{aligned} |R(x) \cap R(v)| &= |R(x)| + |R(v)| - |R(x) \cup R(v)| \\ &\geq \left(\frac{2}{\ell} - 2\varepsilon_{\mathcal{A}}\right) n - \left(4\ell\varepsilon_{\mathcal{E}} + \frac{1}{\ell} + \ell^2\varepsilon_{\mathcal{E}}\right) n \\ &\geq \left(\frac{1}{\ell} - 2\varepsilon_{\mathcal{A}} - 4\ell\varepsilon_{\mathcal{E}} - \ell^2\varepsilon_{\mathcal{E}}\right) n \geq \frac{9}{10\ell} n \end{aligned}$$

by Inequality 10.

After these preliminary observations we are in a position to prove the lemma: Let $S = \{v_1, \ldots, v_\ell\}$ be an ℓ-clique which exists due to Corollary 3.17. For $j = 1, \ldots, \ell$ we set

$$S(v_j) := \left\{ v \in V \;\middle|\; |R(v) \cap R(v_j)| \geq \frac{2}{3}\frac{n}{\ell} \right\} \ .$$

Consider two vertices x, y that belong both to $S(v_j)$ and which are connected by an edge. As $|R(v_j)| \leq (\frac{1}{\ell} + \varepsilon_{\mathcal{A}})n$ by (1) it follows that $|R(x) \cap R(y)|$ is at least, say, $\frac{1}{4\ell}n$. This contradicts (2) (cf. Inequality 7) and we conclude that the sets $S(v_j)$ are stable.

On the other hand, observe that (1) together with $|\bigcup_{j=1}^{\ell} R(v_j)| \geq n - \varepsilon_{\mathcal{E}} n$ implies that for every vertex $x \in V$ there must be an index $j \in \{1, \ldots, \ell\}$ such that $R(x) \cap R(v_j) \neq \emptyset$. Consider an arbitrary vertex $y \in R(x) \cap R(v_j)$. By (3) it follows that $|R(x) \cap R(y)| \geq \frac{9}{10\ell}n$ and $|R(v_j) \cap R(y)| \geq \frac{9}{10\ell}n$. Since $|R(y)| \leq (\frac{1}{\ell} + \varepsilon_{\mathcal{A}})n$ it we conclude that $|R(x) \cap R(v_j)| \geq \frac{2}{3}\frac{n}{\ell}$. From that we deduce that $\bigcup_{j=1}^{\ell} S(v_j) = V$, i.e., G is ℓ-colorable. ∎

C. Putting it all together

Finally, we are in a position to show our main result.

Proof of Theorem 2.2. We will show the slightly stronger statement that

$$|\mathcal{F}_i^\ell(n)| = (1 + c_i \cdot 2^{-\gamma n})|\mathcal{P}_i^\ell(n)| \ .$$

for appropriate constants $c_i \geq 1$ (where $c_0 \leq c_i$ for $i \geq 1$) and $\gamma > 0$.

By Lemma 3.18 it follows that

$$|\mathcal{F}_i^\ell(n)| \leq |\mathcal{P}_i^\ell(n)| + |\mathcal{A}_i^\ell(n)| + |\mathcal{B}_i^\ell(n)| + |\mathcal{C}_i^\ell(n)| + |\mathcal{D}_i^\ell(n)| + |\mathcal{E}_i^\ell(n)| \ .$$

Thus, it suffices to show that $|\mathcal{X}(n)|/|\mathcal{P}_i^\ell(n)| = \frac{1}{5\ell}2^{-\gamma n}$ for all bad sets $\mathcal{X}(n)$, i.e.,

$$\begin{aligned}\mathcal{X}(n) \in\ &\{\mathcal{A}_i^\ell(n,1),\ldots,\mathcal{A}_i^\ell(n,\ell-1),\mathcal{B}_i^\ell(n,0),\ldots,\mathcal{B}_i^\ell(n,\ell-2),\\ &\mathcal{C}_i^\ell(n),\mathcal{D}_i^\ell(n),\mathcal{E}_i^\ell(n,1),\ldots,\mathcal{E}_i^\ell(n,\ell)\}\ .\end{aligned}$$

We proceed by induction on n. Note that we may assume without loss of generality that the claim is true for all $n \le n_0$ for an arbitrarily large constant n_0. This is due to the fact that we can simply choose $c_i \ge 1$ sufficiently large. For the induction step we note that for all bad sets $\mathcal{X}(n)$ apart from $\mathcal{A}_i^\ell(n,1),\ldots,\ \mathcal{A}_i^\ell(n,\ell-1)$ one easily checks by Lemma 3.7, Lemma 3.11, Lemma 3.13 and Lemma 3.16 that for $\gamma > 0$ sufficiently small we have

$$\log\frac{|\mathcal{X}(n)|}{|\mathcal{F}_i^\ell(n-k)|} \le \left(\frac{\ell-1}{\ell} - 3\gamma\right) nk$$

for a suitable constant $k \ge 1$ and n sufficiently large.

Furthermore, it holds that

$$\frac{|\mathcal{X}(n)|}{|\mathcal{P}_i^\ell(n)|} = \frac{|\mathcal{X}(n)|}{|\mathcal{F}_i^\ell(n-k)|}\cdot\frac{|\mathcal{F}_i^\ell(n-k)|}{|\mathcal{P}_i^\ell(n-k)|}\cdot\frac{|\mathcal{P}_i^\ell(n-k)|}{|\mathcal{P}_i^\ell(n)|}\ .$$

By the induction hypothesis and Corollary 3.2 we deduce that

$$\begin{aligned}\log\frac{|\mathcal{X}(n)|}{|\mathcal{P}_i^\ell(n)|} &\le \left(\frac{\ell-1}{\ell} - 3\gamma\right) nk + \log(1+c_i) - \frac{\ell-1}{\ell}nk + \mathcal{O}(1)\\ &\le \log(c_i) - 2\gamma n\end{aligned}$$

and, consequently, $|\mathcal{X}(n)|/|\mathcal{P}_i^\ell(n)| \le \frac{c_i}{5\ell}2^{-\gamma n}$.

For the bad sets $\mathcal{X}(n) \in \{\mathcal{A}_i^\ell(n,1),\ldots,\mathcal{A}_i^\ell(n,\ell-1)\}$ we proceed similarly by observing that

$$\log\frac{|\mathcal{X}(n)|}{|\mathcal{F}_0^\ell(n-g(n))|} \le \left(\frac{\ell-1}{\ell} - 3\gamma\right) ng(n)$$

for $g(n) = kf_k(n) = \mathcal{O}(\log n)$ due to Lemma 3.6. Moreover, it holds that

$$\begin{aligned}\frac{|\mathcal{X}(n)|}{|\mathcal{P}_i^\ell(n)|} =\ &\frac{|\mathcal{X}(n)|}{|\mathcal{F}_0^\ell(n-g(n))|}\cdot\frac{|\mathcal{F}_0^\ell(n-g(n))|}{|\mathcal{P}_0^\ell(n-g(n))|}\cdot\frac{|\mathcal{P}_0^\ell(n-g(n))|}{|\mathcal{P}_i^\ell(n-g(n)+i(2\ell+1))|}\cdot\\ &\frac{|\mathcal{P}_i^\ell(n-g(n)+i(2\ell+1))|}{|\mathcal{P}_i^\ell(n)|}\ .\end{aligned}$$

Applying the induction hypothesis, Corollary 3.2 and Lemma 3.3 we conclude that

$$\begin{aligned}\log\frac{|\mathcal{X}(n)|}{|\mathcal{P}_i^\ell(n)|} \le\ &\left(\frac{\ell-1}{\ell} - 3\gamma\right) n\,g(n) + \log(1+c_0) + 0\\ &-\frac{\ell-1}{\ell}n\,g(n) + \mathcal{O}(g(n)^2)\\ \le\ &\log(c_i) - 2\gamma n\ .\end{aligned}$$

Once again, it follows that $|\mathcal{X}(n)|/|\mathcal{P}_i^\ell(n)| \leq \frac{c_i}{5\ell}2^{-\gamma n}$ and the induction step is completed. ∎

REFERENCES

[1] B. Bollobás B. and P. Erdős, On the structure of edge graphs, *Bull. London Math. Soc.*, 5, 317–321 (1973).

[2] P. Erdős, D. J. Kleitman, and B. L. Rothschild, Asymptotic enumeration of K_n-free graphs, In *International Colloquium on Combinatorial Theory*, volume 2 of *Atti dei Convegni Lincei*, Roma, 19–27 (1976).

[3] Ph. G. Kolaitis, H. J. Prömel, and B. L. Rothschild, K_{l+1}-free graphs: Asymptotic structure and a 0–1 law, *Trans. Amer. Math. Soc.*, 303, 637–671 (1987).

[4] D. J. Kleitman and B. L. Rothschild, Asymptotic enumeration of partial orders on a finite set, *Trans. Amer. Math. Soc.*, 205, 205–220 (1975).

[5] H. J. Prömel, T. Schickinger, and A. Steger, A note on triangle-free and bipartite graphs, Special issue of *Discrete Mathematics* in honor of the 65th birthday of Daniel J. Kleitman (to appear).

[6] H. J. Prömel, A. Steger, and A. Taraz, Asymptotic enumeration, global structure and constrained evolution, *Discrete Mathematics*, 229, 213–233 (2001).