ON THE TYPICAL STRUCTURE OF GRAPHS NOT CONTAINING A FIXED VERTEX-CRITICAL SUBGRAPH

OREN ENGELBERG, WOJCIECH SAMOTIJ, AND LUTZ WARNKE

Abstract. This work studies the typical structure of sparse H-free graphs, that is, graphs that do not contain a subgraph isomorphic to a given graph H. Extending the seminal result of Osthus, Prömel, and Taraz that addressed the case where H is an odd cycle, Balogh, Morris, Samotij, and Warnke proved that, for every $r \geqslant 3$, the structure of a random K_{r+1}-free graph with n vertices and m edges undergoes a phase transition when m crosses an explicit (sharp) threshold function $m_r(n)$. They conjectured that a similar threshold phenomenon occurs when K_{r+1} is replaced by any strictly 2-balanced, edge-critical graph H. In this paper, we resolve this conjecture. In fact, we prove that the structure of a typical H-free graph undergoes an analogous phase transition for every H in a family of vertex-critical graphs that includes all edge-critical graphs.

Contents

1. Introduction

1.1. **Background and motivation.** Given a graph H, let $\mathcal{F}_n(H)$ be the family of all graphs with vertex set $[\![n]\!] = \{1, \dots, n\}$ that are *H-free*, that is, graphs which do not contain a (not necessarily induced) subgraph isomorphic to H. A basic question in extremal graph theory is to determine $\mathrm{ex}(n, H)$, the largest number of edges in a graph from $\mathcal{F}_n(H)$. The classical result of Turán [22] determines $\mathrm{ex}(n, K_{r+1})$ for every $r \geqslant 2$ and also characterises the extremal graphs. The works of Erdős, Simonovits, and Stone [9, 10] extend this to an arbitrary non-bipartite graph H, showing that

$$\mathrm{ex}(n,H) = \left(1 - \frac{1}{\chi(H)-1} + o(1)\right)\binom{n}{2}$$

This research was supported in part by the Israel Science Foundation grants 1147/14 and 1145/18 (OE and WS).

and, moreover, that every H-free graph with at least $\mathrm{ex}(n,H) - o(n^2)$ edges may be made $(\chi(H)-1)$-partite by removing from it some $o(n^2)$ edges.

Here, we are interested in the structure of a *typical* H-free graph. This problem was first considered by Erdős, Kleitman, and Rothschild [8], who proved that almost all triangle-free graphs are bipartite. Formally, if F_n is a uniformly chosen random element of $\mathcal{F}_n(K_3)$, then

$$\lim_{n\to\infty} \mathbb{P}(F_n \text{ is bipartite}) = 1.$$

This result was later generalised by Kolaitis, Prömel, and Rothschild [15], who showed that, for every fixed $r \geqslant 2$, almost all graphs in $\mathcal{F}_n(K_{r+1})$ are r-partite. Very recently, Balogh and the second named author [5] proved that this remains true as long as $r \leqslant c\log n/\log\log n$ for some small positive constant c, see also [2, 16].

The result of [15] was further generalised from cliques to the much wider class of edge-critical graphs. We say that a graph H is *edge-critical* if it contains an edge whose removal reduces the chromatic number, that is, if $\chi(H \setminus e) = \chi(H) - 1$ for some $e \in E(H)$; in particular, every clique is edge-critical and so is every odd cycle. Simonovits [21] showed that, for every edge-critical graph H and all large enough n, not only $\mathrm{ex}(n,H) = \mathrm{ex}(n, K_{\chi(H)})$ but also that the only H-free graphs with $\mathrm{ex}(n,H)$ edges are complete $(\chi(H)-1)$-partite graphs, as in the case $H = K_{\chi(H)}$. Prömel and Steger [18] showed that, if H is edge-critical, then almost every H-free graph is $(\chi(H)-1)$-partite.

One drawback of the structural characterisations of typical H-free graphs mentioned above is that they do not say anything about sparse graphs, that is, n-vertex graphs with $o(n^2)$ edges. Indeed, for every non-bipartite H, the family $\mathcal{F}_n(H)$ contains all bipartite graphs and there are at least $2^{\lfloor n^2/4\rfloor}$ of them; this is much more than the number of *all* graphs with n vertices and at most $n^2/20$ edges. In view of this, it is natural to ask the following refined question, first considered by Prömel and Steger [20]: Fix some m with $0 \leqslant m \leqslant \mathrm{ex}(n,H)$. What can be said about the structure of a uniformly selected random element of $\mathcal{F}_n(H)$ with exactly m edges? In particular, for what m does this graph admit a similar description as a uniformly random element of $\mathcal{F}_n(H)$?

Let $\mathcal{G}_{n,m}$ be the family of all graphs with vertex set $[\![n]\!]$ and precisely m edges and let $\mathcal{F}_{n,m}(H) = \mathcal{G}_{n,m} \cap \mathcal{F}_n(H)$ be the subfamily of $\mathcal{G}_{n,m}$ that comprises all H-free graphs. Osthus, Prömel, and Taraz [17] showed that, for every odd integer $\ell \geqslant 3$, there exists an explicit constant c_ℓ such that, letting $m_\ell = m_\ell(n) = c_\ell n^{\frac{\ell}{\ell-1}}(\log n)^{\frac{1}{\ell-1}}$, a uniformly random graph $F_{n,m} \in \mathcal{F}_{n,m}(C_\ell)$ satisfies, for every $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}(F_{n,m} \text{ is bipartite}) = \begin{cases} 0 & \text{if } n/2 \leqslant m \leqslant (1-\varepsilon)m_\ell, \\ 1 & \text{if } m \geqslant (1+\varepsilon)m_\ell. \end{cases}$$

This result was extended by Balogh, Morris, Samotij, and Warnke [4] to the case where H is a clique of an arbitrary order. They showed that, for every $r \geqslant 3$, there is an explicit positive constant c'_r such that, letting $m'_r = m'_r(n) = c'_r n^{2-\frac{2}{r+2}}(\log n)^{\frac{1}{\binom{r+1}{2}-1}}$, a uniformly chosen random graph $F_{n,m} \in \mathcal{F}_{n,m}(K_{r+1})$ satisfies, for every $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}(F_{n,m} \text{ is } r\text{-partite}) = \begin{cases} 0 & \text{if } n \ll m \leqslant (1-\varepsilon)m'_r, \\ 1 & \text{if } m \geqslant (1+\varepsilon)m'_r. \end{cases}$$

1.2. **Our result – edge-critical graphs.** Aiming towards a common generalisation of the results of [4, 17, 18], the authors of [4] made the following conjecture. Recall that the 2-*density* of a graph H is defined by

$$m_2(H) = \max\left\{\frac{e_K - 1}{v_K - 2} : K \subseteq H,\, v_K \geqslant 3\right\}$$

and that H is called *strictly* 2-*balanced* if the maximum above is attained only when $K = H$, that is, if $m_2(K) < m_2(H)$ for every proper subgraph $K \subsetneq H$.

Conjecture 1.1 ([4, Conjecture 1.3])**.** *For every strictly* 2*-balanced, non-bipartite, edge-critical graph* H*, there exists a constant* C *such that the following holds. If*

$$m \geqslant Cn^{2-\frac{1}{m_2(H)}}(\log n)^{\frac{1}{e_H-1}},$$

then almost all graphs in $\mathcal{F}_{n,m}(H)$ *are* $(\chi(H)-1)$*-partite.*

In this paper, we resolve this conjecture and show that the assumption on m is best possible.

Theorem 1.2. *For every strictly* 2*-balanced, non-bipartite, edge-critical graph graph* H*, there exist positive constants* c_H *and* C_H *such that, letting*

$$m_H = m_H(n) = n^{2-\frac{1}{m_2(H)}}(\log n)^{\frac{1}{e_H-1}},$$

the following holds for a uniformly chosen random $F_{n,m} \in \mathcal{F}_{n,m}(H)$*:*

$$\lim_{n\to\infty} \mathbb{P}\big(F_{n,m} \textit{ is } (\chi(H)-1)\textit{-partite}\big) = \begin{cases} 0 & \textit{if } n \ll m \leqslant c_H m_H, \\ 1 & \textit{if } m \geqslant C_H m_H. \end{cases}$$

In fact, Theorem 1.2 is only a special case of a much more general result, Theorem 1.4 below, which we present in the next subsection.

1.3. **Our result – vertex-critical graphs.** We say that a graph H is *vertex-critical* if it contains a vertex whose deletion reduces the chromatic number, that is, if $\chi(H-v) = \chi(H) - 1$ for some $v \in V(H)$; we call every such v a *critical vertex* of H. A star $S \subseteq H$ centred at a critical vertex is called a *critical star* if $\chi(H \setminus S) = \chi(H) - 1$ and if no proper subgraph $S' \subsetneq S$ has this property. For a critical vertex v, we define $\mathrm{crit}(v)$, the *criticality of* v, to be the smallest number of edges incident to v whose removal decreases the chromatic number, that is,

$$\mathrm{crit}(v) = \min\{e_S : S \text{ is a critical star centred at } v\},$$

and define $\mathrm{crit}(H)$, the *criticality of* H, to be the smallest criticality of a vertex, that is,

$$\mathrm{crit}(H) = \min\{\mathrm{crit}(v) : v \text{ is a critical vertex}\}.$$

Note that every edge-critical graph is also vertex-critical. Conversely, a vertex-critical graph is edge-critical precisely when its criticality is equal to one.

The motivation for our investigation of the typical structure of vertex-critical graphs is a result of Hundack, Prömel, and Steger [12] which states that, for every vertex-critical H, almost all H-free graphs are 'almost' $(\chi(H)-1)$-partite in the following precise sense. Given integers $r \geqslant 1$ and $k \geqslant 0$, we will denote by $\mathcal{G}(r,k)$ the class of all graphs G that admit an r-colouring of $V(G)$ for which the subgraph of G induced by each of the r colour classes has maximum degree at most k. In particular, $\mathcal{G}(r,0)$ is the class of all r-colourable graphs and the following theorem generalises the main result of [18].

Theorem 1.3 ([12]). *If H is a vertex-critical graph of criticality $k+1$ and $\chi(H) = r+1 \geqslant 3$, then, for some positive c,*

$$|\mathcal{F}_n(H)| = \big(1 + O(2^{-cn})\big) \cdot |\mathcal{F}_n(H) \cap \mathcal{G}(r,k)|.$$

Remark. A less accurate description of the structure of a typical member of $\mathcal{F}_n(H)$, but valid for *every* non-bipartite H, was given by Balogh, Bollobás, and Simonovits [1].

Our main result is a sparse analogue of Theorem 1.3 that is valid for a subclass of vertex-critical graphs that includes all edge-critical graphs. In order to state it, we need several additional definitions.

Definition. A vertex-critical graph H will be called *simple vertex-critical* if every colouring of H with $\chi(H)-1$ colours admits a monochromatic star with $\mathrm{crit}(H)$ edges or a monochromatic cycle. Further, a vertex-critical graph will be called *plain vertex-critical* if, for every colouring of H with $\chi(H)-1$ colours, the monochromatic graph B satisfies at least one of the following:

(i) B contains a cycle,
(ii) B is the star $K_{1,\mathrm{crit}(H)}$,
(iii) B has a vertex with degree larger than $\mathrm{crit}(H)$, or
(iv) B has two nonadjacent vertices with degree $\mathrm{crit}(H)$.

It is not hard to see that every edge-critical graph is plain vertex-critical and every plain vertex-critical graph is simple vertex-critical.

Remark. Another family of plain vertex-critical graphs are the complete multipartite graphs $K_{1,k_1,\dots,k_r}$ with $1 \leqslant k_1 < k_2 \leqslant \dots \leqslant k_r$. To see this, denote the $r+1$ colour classes of this graph by $V_0, \dots, V_r$, so that $|V_0| = 1$ and $|V_i| = k_i$ for each $i \in [\![r]\!]$. Consider an arbitrary r-colouring $W_1 \cup \dots \cup W_r$ of the vertices of $K_{1,k_1,\dots,k_r}$. If, for some $j \in [\![r]\!]$, we have $|W_j \setminus V_i| \geqslant 2$ for every i, then W_j must contain a cycle; indeed, in this case W_j intersects three different V_i or it intersects some two V_i in at least two vertices each. We may therefore assume that, for every $j \in [\![r]\!]$, there is an $i(j)$ such that $\delta_j = |W_j \setminus V_{i(j)}| \leqslant 1$. This assumption guarantees that each W_j induces a star (if $\delta_j = 1$) or an empty graph (if $\delta_j = 0$). Let $J = \{j \in [\![r]\!] : \delta_j = 1\}$ and observe that

$$\sum_{j \in J} |W_j| = 1 + k_1 + \dots + k_r - \sum_{j \notin J} |W_j| \geqslant 1 + k_1 + \dots + k_{|J|} \geqslant |J| \cdot (k_1 + 1).$$

In particular, either each W_j with $j \in J$ induces a copy of K_{1,k_1} or one of them induces a graph with maximum degree strictly larger than k_1.

For an integer $k \geqslant 2$ and a graph F with $v_F \geqslant k+1$, we let

$$d_k(F) = \frac{e_F - k + 1}{v_F - k},$$

cf. the definition of 2-density given at the start of Section 1.2. Suppose that H is a non-bipartite, vertex-critical graph and let $S_1, \dots, S_t$ be all the critical stars of H. Further, let $k \geqslant 0$ and $r \geqslant 2$ be the integers such that $\mathrm{crit}(H) = k+1$ and $\chi(H) = r+1$. Denote, for each $i \in [\![t]\!]$,

$$\eta_i(H) = \max\big\{d_{k+2}(F) : S_i \subsetneq F \subseteq H\big\} \qquad \text{and} \qquad \eta(H) = \min_{1 \leqslant i \leqslant t} \eta_i(H)$$

and, further,

$$\zeta_i(H) = \min\big\{e_F : S_i \subsetneq F \subseteq H,\ d_{k+2}(F) = \eta_i(H)\big\} \qquad \text{and} \qquad \zeta(H) = \max_{\substack{1 \leqslant i \leqslant t \\ \eta_i(H) = \eta(H)}} \zeta_i(H).$$

We are now ready to define the threshold function:

$$m_H = m_H(n) = \begin{cases} n^{2-\frac{1}{m_2(H)}} & \text{if } m_2(H) > \eta(H), \\ n^{2-\frac{1}{\eta(H)}}(\log n)^{\frac{1}{\zeta(H)-k-1}} & \text{otherwise.} \end{cases} \tag{1}$$

Remark. If H is edge-critical, then $S_1, \dots, S_t$ are the critical edges of H, that is, edges S satisfying $\chi(H \setminus S) = \chi(H) - 1$. If, additionally, H is strictly 2-balanced, then, for every $i \in [\![t]\!]$, the maximum in the definition of $\eta_i(H)$ is uniquely attained at $F = H$ and thus $\eta_i(H) = d_2(H) = m_2(H)$ and $\zeta_i(H) = e_H$; consequently, $\eta(H) = m_2(H)$ and $\zeta(H) = e_H$. This shows that definition (1) extends the definition of m_H given in the statement of Theorem 1.2.

The following generalisation of Theorem 1.2 is the main result of this work.

Theorem 1.4. *Let H be a simple vertex-critical graph with $\chi(H) = r + 1 \geqslant 3$ and criticality $k + 1$ and let $F_{n,m}$ denote the uniformly chosen random element of $\mathcal{F}_{n,m}(H)$. There exists a positive constant C_H such that, for every $m \geqslant C_H m_H$,*

$$\lim_{n\to\infty} \mathbb{P}\big(F_{n,m} \in \mathcal{G}(r, k)\big) = 1.$$

Furthermore, if H is plain vertex-critical, then there exists a positive constant c_H such that, for every $n \ll m \leqslant c_H m_H$,

$$\lim_{n\to\infty} \mathbb{P}\big(F_{n,m} \in \mathcal{G}(r, k)\big) = 0.$$

It may be worth pointing out that there are plain vertex-critical graphs H for which $\eta(H)$ is strictly larger than $m_2(H)$. (One such graph is $H = K_{1,2,3}$, which is plain vertex-critical with criticality two, has exactly one critical star, and satisfies $\eta(H) = 3 > 5/2 = m_2(H)$.) Why is it interesting? Let H be such a graph and let $F_{n,m}$ be a uniformly chosen random element of $\mathcal{F}_{n,m}(H)$. As soon as $m \gg n^{2-\frac{1}{m_2(H)}}$, with probability close to one, $F_{n,m}$ can be made $(\chi(H) - 1)$-partite by removing from it $o(m)$ edges, see Theorem 6.2 below. However, it is only when $m \gg n^{2-\frac{1}{\eta(H)}}(\log n)^{\frac{1}{\zeta(H)-k-1}}$, polynomially above the 2-density threshold, that the 'exact' structure emerges. What can one say about the typical structure of $F_{n,m}$ between these two thresholds?

Unfortunately, our techniques are too weak to extend Theorem 1.4 to arbitrary vertex-critical graphs. Still, they are sufficient to prove an approximate version of the 1-statement. We refrain ourselves from stating this result here; instead, we refer the interested reader to [7]. Having said that, we have no good reason to believe that m_H is the threshold for a general vertex-critical graph H. We believe that it would be extremely interesting to find the threshold for a generic vertex-critical graph H and extend Theorem 1.4 to all such H.

1.4. **Acknowledgements.** We thank József Balogh, Michael Krivelevich, Rob Morris, and Angelika Steger for stimulating discussions about various aspects of this work and Anita Liebenau for bringing [6, Theorem 2] to our attention.

2. Outline of the proof

2.1. **Why is m_H the threshold?** The location of the threshold at which a typical graph in $\mathcal{F}_{n,m}(H)$ 'enters' $\mathcal{G}(r,k)$ can be guessed by comparing the number of graphs in $\mathcal{F}_{n,m}(H) \cap \mathcal{G}(r,k)$ to the number of graphs in $\mathcal{F}_{n,m}(H)$ that are 'one edge away' from $\mathcal{G}(r,k)$. It is relatively straightforward to estimate the former: For a constant proportion of graphs $G \in \mathcal{G}_{n,m} \cap \mathcal{G}(r,k)$, the monochromatic subgraph of G (under its optimal colouring, i.e., the colouring that witnesses $G \in \mathcal{G}(r,k)$) has girth larger than the number of vertices of H (this is true for all m); moreover, the assumption that H is *simple* vertex-critical guarantees that each such graph is H-free. As for the latter quantity, the number of graphs in $\mathcal{G}_{n,m}$ that are 'one edge away' from $\mathcal{G}(r,k)$ is $\Theta(m)$ times larger, but the proportion of them that are H-free is a decreasing function of m. The reason for this is that every such graph G has at least one copy of $K_{1,k+1}$ in its monochromatic subgraph (under every r-colouring) and, since $\chi(H) = r+1$ and $\mathrm{crit}(H) = k+1$, this copy of $K_{1,k+1}$ can be extended to many copies of H in K_n that use only edges that are properly coloured. In particular, if G is H-free, then it must avoid all such copies of $H \setminus K_{1,k+1}$. Furthermore, if G is *plain* vertex-critical, then this implication can be reversed—G is H-free *if and only if* it avoids all such copies of $H \setminus K_{1,k+1}$—under a weak assumption on the monochromatic graph (girth larger than v_H) that is satisfied by a constant proportion of all such graphs.

Finally, if we fix both the optimal r-colouring and the monochromatic graph (which is 'one edge away' from having maximum degree k), the proportion P_m of graphs in $\mathcal{G}_{n,m}$ (among those that contain our fixed monochromatic graph) that avoid all copies of $H \setminus K_{1,k+1}$ of the above type can be bounded using the inequalities of Janson (from above) and Harris (from below) as follows:

$$-\log P_m = \Theta\left(\max_{i\in[\![t]\!]} \min\left\{n^{v_F-k-2}\cdot (m/n^2)^{e_F-k-1} : S_i \subsetneq F \subseteq H\right\}\right),$$

where $S_1,\ldots,S_t$ are the critical stars of H.

The threshold m_H is then the smallest $m \geqslant n^{2-\frac{1}{m_2(H)}}$ for which $-\log P_m \geqslant \log m$, that is, for which the number of H-free graphs that are 'one edge away' from $\mathcal{G}(r,k)$ is of the same order of magnitude as $|\mathcal{F}_{n,m}(H) \cap \mathcal{G}(r,k)|$. We note that the additional requirement $m \geqslant n^{2-\frac{1}{m_2(H)}}$ is needed because $n^{2-\frac{1}{m_2(H)}}$ is the threshold for approximate r-colourability of a random element of $\mathcal{F}_{n,m}(H)$ and comparing $|\mathcal{F}_{n,m}(H) \cap \mathcal{G}(r,k)|$ only to the number of graphs in $\mathcal{F}_{n,m}(H)$ that are 'one edge away' from $\mathcal{G}(r,k)$—as opposed to $|\mathcal{F}_{n,m}(H)|$—cannot be justified below this threshold.

2.2. **The 0-statement.** Our proof of the 0-statement (the second assertion of Theorem 1.4), presented in Section 5, is a formalisation of the above discussion. (Having said that, in the range $n \ll m \ll n^{2-\frac{1}{m_2(H)}}$, we give a separate, elementary counting argument that exploits the fact that a typical graph in $\mathcal{G}_{n,m}$ can be made H-free by removing from it some $o(m)$ edges, see Section 5.1.) One of the key ideas here is to reduce the problem of comparing $|\mathcal{F}_{n,m}(H) \cap \mathcal{G}(r,k)|$ and the number of graphs in $\mathcal{F}_{n,m}(H)$ that are 'one edge away' from $\mathcal{G}(r,k)$ to the analogous problem for a *fixed* r-colouring that is moreover balanced (in the sense that each of its r colour classes has approximately n/r vertices). Various assertions and estimates that justify this reduction are proved in Section 4, where we also show that the number of graphs in $\mathcal{G}_{n,m}$ that are 'one edge away' from $\mathcal{G}(r,k)$ and whose monochromatic graph has girth larger than v_H is indeed $\Theta(m)$ times bigger than $|\mathcal{G}_{n,m} \cap \mathcal{G}(r,k)|$.

2.3. **The 1-statement.** The proof of the 1-statement (the first assertion of Theorem 1.4) is significantly harder. A first, nowadays standard, step is to show that, when $m \gg n^{2-\frac{1}{m_2(H)}}$, almost every graph in $\mathcal{F}_{n,m}(H)$ admits an r-colouring such that:

(i) there are only $o(m)$ monochromatic edges,
(ii) each colour class comprises approximately n/r vertices,
(iii) every vertex has at most as many neighbours in its own colour class as it has in any other colour class.

We derive this approximate version of the 1-statement, stated as Theorem 6.1 below, from [3, Theorem 1.7]. Using several properties of graphs in $\mathcal{G}_{n,m} \cap \mathcal{G}(r,k)$ established in Section 4, we further reduce the 1-statement to showing that, for every *fixed* r-colouring Π[1] satisfying (ii) above, the number of graphs $G \in \mathcal{F}_{n,m}(H)$ that satisfy (i) and (iii) for this particular Π but the maximum degree of $G \setminus \Pi$ (the monochromatic subgraph of G) exceeds k is much smaller than the number of graphs $G \in \mathcal{G}_{n,m}$ such that the maximum degree of $G \setminus \Pi$ is at most k. This reduction is formalised in Section 7.1.

Fix an r-colouring Π satisfying (ii) and let $\mathcal{F}^*$ denote the family of al graphs G that satisfy (i) and (iii) and $\Delta(G \setminus \Pi) > k$. The methods of bounding the number of graphs in $\mathcal{F}^*$ will vary with m and the distribution of the edges of $G \setminus \Pi$. In Section 9, we give a somewhat ad-hoc argument to separately treat the case where $m > \mathrm{ex}(n,H) - \xi n^2$ for some small positive ξ (the *dense case*); we will not discuss it in detail here. We deal with the main, complementary case $m \leqslant \mathrm{ex}(n,H) - \xi n^2$, which we term the *sparse case* in Section 8.

Let $\mathcal{T}$ denote the collection of all possible monochromatic graphs $G \setminus \Pi$ as G ranges over $\mathcal{F}^*$. For every $T \in \mathcal{T}$, we arbitrarily choose a maximal subgraph $B_T \subseteq T$ with $\Delta(B_T) = k$. Since we will separately estimate the number of graphs $G \in \mathcal{F}^*$ such that $B_{G\setminus\Pi} = B$ for every B with $\Delta(B) = k$, we may further fix one such B. The possible monochromatic graphs $T \in \mathcal{T}$ with $B_T = B$ are divided into two classes, denoted $\mathcal{T}_{\mathrm{L}}$ and $\mathcal{T}_{\mathrm{H}}$, depending on what proportion of edges of T are incident to vertices whose degrees are larger than $\rho m/n$, where ρ is a small positive constant, see Section 8.2. We separately enumerate graphs $G \in \mathcal{F}^*$ such that $G \setminus \Pi \in \mathcal{T}_{\mathrm{L}}$ and those satisfying $G \setminus \Pi \in \mathcal{T}_{\mathrm{H}}$. We term these two parts of the argument the *low-degree case* and the *high-degree case*, respectively. We outline these two cases in Sections 8.3 and 8.6, respectively.

In the low-degree case, for each $T \in \mathcal{T}_{\mathrm{L}}$, we give an upper bound on the number of graphs $G \in \mathcal{F}^*$ such that $G \setminus \Pi = T$ and then sum this bound over all T. This upper bound, stated in Proposition 8.1, lies at the heart the low-degree case. We briefly describe the main idea: By construction, every edge of $T \setminus B$ belongs to a copy of $K_{1,k+1}$ in T and, since $\chi(H) = r+1$ and $\mathrm{crit}(H) = k+1$, this copy of $K_{1,k+1}$ can be extended to $\Omega(n^{v_H-k-2})$ copies of H in K_n that use only the edges of Π. Consequently, each $G \in \mathcal{F}^*$ with $G \setminus \Pi = T$ must avoid all such copies of $H \setminus K_{1,k+1}$, for every copy of $K_{1,k+1}$ in T. The number of graphs G with this property is bounded from above with the use of the Hypergeometric Janson Inequality (Lemma 3.1) applied to a carefully chosen subfamily

[1]We identify every r-colouring with a partition of $[\![n]\!]$ into r sets as well as the complete r-partite graph with these partite sets.

of copies of $K_{1,k+1}$ in T with 'nice' properties that enable us to control the correlation term Δ in Lemma 3.1. Finally, the size of the summation over all T is controlled by Lemma 8.2, which in turn utilises bounds on the number of graphs in $\mathcal{T}$ with given values of certain key parameters that are obtained in Section 8.4.

In the high-degree case, we crucially use property (iii) to argue that, for every $T \in \mathcal{T}_{\mathrm{H}}$, all high-degree vertices (vertices whose degree is larger than $\rho m/n$) in every $G \in \mathcal{F}^*$ with $G \setminus \Pi = T$ must have at least $\rho m/n$ neighbours in each colour class of Π. This allows us to enumerate all such G in two steps as follows: First, we specify a graph Z that includes T and the edges of $G \cap \Pi$ incident to a carefully chosen set Y of high-degree vertices. Second, we specify the remaining $m - e(Z)$ edges of $G \cap \Pi$. Since H is vertex-critical, each vertex of Y belongs to many copies of H in $Z \cup \Pi$. This allows us to bound the number of ways to choose the $m-e(Z)$ edges of $G \cap \Pi$ in the second step from above with another application of the Hypergeometric Janson Inequality. For the vast majority of Z, this upper bound will be sufficiently strong to enable a naive union bound over the choice of Z; we shall say that such Z fall into the *regular case*. Unfortunately, there will be a small family of exceptional graphs Z for which this upper bound is too weak (this can happen when there are unusual overlaps between the neighbourhoods of the vertices in Y); we shall term it the *irregular case*. However, we may use Lemma 3.3 to show that the number of such exceptional graphs is so small that even a trivial upper bound of $\binom{e(\Pi)}{m-e(Z)}$ for the number of choices in the second step will be sufficient to show that the number of graphs that fall into the irregular case is tiny.

3. Preliminaries

3.1. **Probabilistic inequalities.** In this section, we present four probabilistic inequalities that will be used in the proof of Theorem 1.4. The first three results presented in this section were proved in [4, Lemmas 3.1, 3.2, 3.6] and the fourth result is a standard bound on the tail probabilities of hypergeometric distributions [14, Theorem 2.10]. We begin with a version of Janson's inequality [13] for the hypergeometric distribution, which is an essential ingredient in the proof of the 1-statement in Theorem 1.4.

Lemma 3.1 (Hypergeometric Janson Inequality). *Suppose that $\{B_i\}_{i\in I}$ is a family of subsets of an n-element set Ω, let $m \in \{0,\dots,n\}$, and let $p = m/n$. Let*

$$\mu = \sum_{i\in I} p^{|B_i|} \qquad \text{and} \qquad \Delta = \sum_{i\sim j} p^{|B_i\cup B_j|},$$

where the second sum is over all ordered pairs $(i,j) \in I^2$ such that $i \neq j$ and $B_i \cap B_j \neq \emptyset$. Let R be the uniformly chosen random m-element subset of Ω and let $\mathcal{B}$ denote the event that $B_i \not\subseteq R$ for all $i \in I$. Then, for every $q \in [0,1]$,

$$\mathbb{P}(\mathcal{B}) \leqslant 2 \cdot \exp\left(-q\mu + q^2\Delta/2\right).$$

The main tool in the proof of the 0-statement in Theorem 1.4 will be the following version of the Harris Inequality [11] for the hypergeometric distribution; it gives a lower bound on the probability $\mathbb{P}(\mathcal{B})$ from the statement of Lemma 3.1.

Lemma 3.2 (Hypergeometric Harris Inequality). *Suppose that $\{B_i\}_{i\in I}$ is a family of subsets of an n-element set Ω. Let $m \in \{0,\dots,\lfloor n/2\rfloor\}$, let R be the uniformly chosen* *random m-element subset of Ω, and let $\mathcal{B}$ denote the event that $B_i \nsubseteq R$ for all $i \in I$. Then, for every $\eta \in (0,1)$,*

$$\mathbb{P}(\mathcal{B}) \geqslant \prod_{i\in I}\left(1-\left(\frac{(1+\eta)m}{n}\right)^{|B_i|}\right) - \exp\left(\eta^2 m/4\right).$$

Another key tool in the proof of the 1-statement in Theorem 1.4 is an estimate of the upper tail of the distribution of the number of edges in a random induced subhypergraph of a sparse z-uniform z-partite hypergraph. It formalizes the following statement: If $\mathcal{M} \subseteq U_1 \times \cdots \times U_z$ contains only a tiny proportion of all the z-tuples in $U_1 \times \cdots \times U_Z$, then the probability that, for a random choice of d-elements sets $W_1 \subseteq U_1, \ldots, W_z \subseteq U_z$, a much larger proportion of $W_1 \times \cdots \times W_z$ falls in $\mathcal{M}$ decays exponentially in d.

Lemma 3.3. *For every integer z and all positive α and λ, there exists a positive τ such that the following holds. Let $U_1, \ldots, U_z$ be finite sets and let d be an integer satisfying $2 \leqslant d \leqslant \min\{|U_1|, \ldots, |U_z|\}$. Suppose that $\mathcal{M} \subseteq U_1 \times \cdots \times U_z$ satisfies*

$$|\mathcal{M}| \leqslant \tau \prod_{i=1}^{z} |U_i|$$

and that $W_1, \ldots, W_z$ are uniformly chosen random d-element subsets of $U_1, \ldots, U_z$, respectively. Then, there are at most $\alpha^d \cdot \prod_{i=1}^{z} \binom{U_i}{d}$ choices of $(W_i)_{i\in[\![z]\!]}$ for which

$$|\mathcal{M} \cap (W_1 \times \cdots \times W_z)| > \lambda d^z.$$

Finally, we will need the following simple bound on lower tails of hypergeometric distributions.

Lemma 3.4. *Let R be the uniformly chosen random m-element subset of an N-element set Ω and let $A \subseteq \Omega$ be a k-element set. Then, for every $t \geqslant 0$,*

$$\mathbb{P}\left(|R \cap A| \leqslant \frac{km}{N} - t\right) \leqslant \exp\left(-\frac{t^2}{2 \cdot km/N}\right).$$

3.2. **Two-density related bounds.** In this short section, we present a useful inequality that will be invoked several times in the proof of the 1-statement of Theorem 1.4.

Lemma 3.5. *Suppose that H is a graph with at least three vertices. If $p \geqslant Cn^{2-\frac{1}{m_2(H)}}$ for some $C \geqslant 0$, then, for every nonempty $F \subseteq H$,*

$$n^{v_F} p^{e_F} \geqslant C^{e_F - 1} n^2 p.$$

Proof. Let F be a nonempty subgraph of H. If F has two vertices, then $F = K_2$ and the assertion is vacuously true. Suppose now that $v_F \geqslant 3$ and observe that

$$n^{v_F-2} p^{e_F-1} \geqslant n^{v_F-2}\left(Cn^{2-\frac{1}{m_2(H)}}\right)^{e_F-1} = C^{e_F-1} n^{v_F-2-(e_F-1)/m_2(H)}.$$

The claimed bound follows as $m_2(H) \geqslant \frac{e_F-1}{v_F-2}$, by the definition of 2-density. □

3.3. **The Turán problem for r-partite graphs.** The case $m \geqslant \mathrm{ex}(n,H) - o(n^2)$ in the proof of Theorem 1.4 will require the following folklore result in extremal graph theory. For integers $r \geqslant 2$ and $n \geqslant 1$, we denote by $K_r(n)$ the balanced complete r-partite graph with $r \cdot n$ vertices.

Lemma 3.6. *For all integers r, s, and n satisfying $r \geqslant 2$ and $n \geqslant s \geqslant 1$,*

$$\mathrm{ex}\big(K_r(n), K_r(s)\big) \leqslant e\big(K_r(n)\big) - n^2/s^2.$$

Proof. Denote the r colour classes of $K_r(n)$ by $V_1, \ldots, V_r$ and, for each $i \in [\![r]\!]$, let R_i be a uniformly chosen random s-element subset of V_i. Suppose that $G \subseteq K_r(n)$ is $K_r(s)$-free and let G' be the subgraph of G induced by $R_1 \cup \cdots \cup R_r$. Since G' may be viewed as a subgraph of $K_r(s)$, we have $e(G') \leqslant e(K_r(s)) - 1$. On the other hand, $\mathbb{E}[e(G')] = e(G) \cdot (s/n)^2$. We conclude that

$$e(G) \leqslant (n/s)^2 \cdot \big(e\big(K_r(s)\big) - 1\big) = e\big(K_r(n)\big) - n^2/s^2,$$

as claimed. □

3.4. **Estimates for binomial coefficients.** We will use the following trivial inequalities that hold for all positive integers $a > b > c$:

$$\binom{a}{b-c} \leqslant \binom{a}{b} \cdot \left(\frac{b}{a-b}\right)^c, \tag{2}$$

$$\binom{b}{c}\binom{a}{c}^{-1} \leqslant \left(\frac{b}{a}\right)^c, \tag{3}$$

$$\binom{a}{c}\binom{b}{c}^{-1} \leqslant \left(\frac{a-c}{b-c}\right)^c, \tag{4}$$

$$\sum_{i=0}^{b}\binom{a}{i} \leqslant \left(\frac{ea}{b}\right)^b. \tag{5}$$

4. On almost r-colourable graphs

In this section, we establish several properties of almost r-colourable graphs, that is, graphs belonging to the family $\mathcal{G}(r,k)$, defined in Section 1.3; these properties will come in handy in our proof of Theorem 1.4. It will be convenient to denote by $\mathcal{G}_{n,m}(r,k) = \mathcal{G}_{n,m} \cap \mathcal{G}(r,k)$ the family of graphs with vertex set $[\![n]\!]$ and precisely m edges that admit an r-colouring whose induced monochromatic graph has maximum degree at most k.

Let $\mathcal{P}_{n,r}$ be the family of all r-colourings of $[\![n]\!]$, that is, all partitions of $[\![n]\!]$ into r parts. For the sake of brevity, we shall often identify a partition $\Pi \in \mathcal{P}_{n,r}$ with the complete r-partite graph with vertex set $[\![n]\!]$ whose colour classes are the r parts of Π. In particular, if G is a graph on the vertex set $[\![n]\!]$, then $G \subseteq \Pi$ means that G is a subgraph of the complete r-partite graph Π or, in other words, that the partition Π is a proper colouring of G. Exploiting this convention, we will also write Π^c to denote the complement of the graph Π, that is, the union of r complete graphs with vertex sets $V_1, \ldots, V_r$.

4.1. **Balanced r-colourings.** We will be interested in balanced r-colourings, that is, partitions of $[\![n]\!]$ whose all parts have approximately n/r elements. More precisely, given a positive γ, we let $\mathcal{P}_{n,r}(\gamma)$ be the family of all partitions of $[\![n]\!]$ into r parts $V_1, \ldots, V_r$ such that

$$\left(\frac{1}{r} - \gamma\right) n \leqslant |V_i| \leqslant \left(\frac{1}{r} + \gamma\right) n \quad \text{for all } i \in [\![r]\!]. \tag{6}$$

That is,

$$\mathcal{P}_{n,r}(\gamma) = \big\{\{V_1, \ldots, V_r\} \in \mathcal{P}_{n,r} : (6) \text{ holds}\big\}.$$

The following easy proposition establishes useful bounds for the number of edges in the complete r-partite graphs defined by balanced and unbalanced r-colourings.

Proposition 4.1. *The following holds for every integer $r \geqslant 2$, every $\gamma > 0$, and all sufficiently large n:*

(i) If $\Pi \in \mathcal{P}_{n,r}(\gamma)$, then

$$e(\Pi) \geqslant (1 - 2r\gamma) \cdot \left(1 - \frac{1}{r}\right)\frac{n^2}{2}.$$

In particular, if $\gamma \leqslant \frac{1}{20r}$, then $e(\Pi) \geqslant n^2/5$.

(ii) If $\Pi \in \mathcal{P}_{n,r} \setminus \mathcal{P}_{n,r}(\gamma)$, then

$$e(\Pi) \leqslant \left(1 - \frac{\gamma^2}{3}\right) \cdot \mathrm{ex}(n, K_{r+1}).$$

Proof. Note that every $\Pi = \{V_1, \ldots, V_r\} \in \mathcal{P}_{n,r}(\gamma)$ satisfies

$$e(\Pi) = \sum_{1 \leqslant i < j \leqslant r} |V_i||V_j| \geqslant \binom{r}{2} \cdot \left[\left(\frac{1}{r} - \gamma\right)n\right]^2 = (1 - r\gamma)^2 \cdot \left(1 - \frac{1}{r}\right)\frac{n^2}{2},$$

proving (i). To see that (ii) holds as well, fix an arbitrary partition Π that does not satisfy (6) and let V and W be two parts of Π with the smallest and the largest size, respectively. Let

$$d = \left\lfloor \frac{|W| - |V|}{2} \right\rfloor,$$

let Π' be a partition obtained from Π by moving some d vertices from W to V, and note that

$$e(\Pi') - e(\Pi) = (|W| - d)(|V| + d) = |V||W| = (|W| - |V|)d - d^2 \geqslant d^2.$$

Since Π does not satisfy (6), it must be that $d \geqslant \lfloor \gamma n/2 \rfloor$ and, since $\mathrm{ex}(n, K_{r+1})$ is the largest number of edges in an r-partite graph with n vertices,

$$e(\Pi) \leqslant e(\Pi') - \frac{\gamma^2 n^2}{5} \leqslant \mathrm{ex}(n, K_{r+1}) - \frac{\gamma^2 n^2}{5} \leqslant \left(1 - \frac{\gamma^2}{3}\right)\mathrm{ex}(n, K_{r+1}), \tag{7}$$

provided that n is sufficiently large. □

4.2. **Monochromatic graphs with small maximum degree.** For every $\Pi \in \mathcal{P}_{n,r}$, define $\mathcal{B}(\Pi, k)$ to be the family of all subgraphs of Π^c with maximum degree at most k. Now, for every $\Pi \in \mathcal{P}_{n,r}$ and $B \in \mathcal{B}(\Pi, k)$, define

$$\mathcal{G}_m(\Pi, B) = \{G \in \mathcal{G}_{n,m} : G \cap \Pi^c = B\},$$

the family of all graphs in $\mathcal{G}_{n,m}$ that, when coloured by Π, have precisely the edges of B monochromatic. Then

$$|\mathcal{G}_m(\Pi, B)| = \binom{e(\Pi)}{m - e(B)}$$

and, since $e(B) \leqslant \Delta(B)n \leqslant kn$ for every $B \in \mathcal{B}(\Pi, k)$,

$$|\mathcal{B}(\Pi, k)| \leqslant \sum_{b=0}^{kn} \binom{e(\Pi^c)}{b} \overset{(5)}{\leqslant} \left(\frac{en^2}{kn}\right)^{kn} \leqslant e^{2kn \log n}, \tag{8}$$

provided that n is sufficiently large. We also have

$$\mathcal{G}_{n,m}(r,k) = \bigcup_{\Pi \in \mathcal{P}_{n,r}} \bigcup_{B \in \mathcal{B}(\Pi,k)} \mathcal{G}_m(\Pi, B).$$

4.3. **The number of graphs with an unbalanced colouring.** The following proposition shows that if $m \gg n \log n$, then almost every graph in $\mathcal{G}_{n,m}(r,k)$ cannot be coloured by an unbalanced partition Π in such a way that the monochromatic graph has maximum degree at most k. In other words, for almost every $G \in \mathcal{G}_{n,m}(r,k)$, all r-colourings of G that yield a monochromatic subgraph with maximum degree k are balanced.

Proposition 4.2. *For all integers $k \geqslant 0$ and $r \geqslant 2$ and every positive γ, there exists a constant $C > 0$ such that, if $m \geqslant Cn \log n$,*

$$\sum_{\Pi \notin \mathcal{P}_{n,r}(\gamma)} \sum_{B \in \mathcal{B}(\Pi,k)} |\mathcal{G}_m(\Pi, B)| \ll \binom{\mathrm{ex}(n, K_{r+1})}{m} \leqslant |\mathcal{G}_{n,m}(r,k)|.$$

Proof. First note that, for an equipartition $\tilde{\Pi}$ of $[\![n]\!]$ into r parts, we have $|\mathcal{G}_m(\tilde{\Pi}, \emptyset)| = \binom{\mathrm{ex}(n,K_{r+1})}{m}$, so it is enough to check the first inequality. Assume that $m \geqslant Cn \log n$ for some sufficiently large constant C. We have

$$\begin{aligned} \binom{\mathrm{ex}(n, K_{r+1})}{m}^{-1} \sum_{B \in \mathcal{B}(\Pi,k)} |\mathcal{G}_m(\Pi, B)| &\overset{(2),\,(8)}{\leqslant} \binom{\mathrm{ex}(n, K_{r+1})}{m}^{-1} \binom{e(\Pi)}{m} e^{2kn \log n} m^{kn} \\ &\overset{(7)}{\leqslant} \left(1 - \frac{\gamma^2}{3}\right)^m e^{4kn \log n} \leqslant e^{-\gamma^2 m/3 + 4kn \log n} \leqslant e^{-\gamma^2 m/4}. \end{aligned} \tag{9}$$

To complete the proof, note that there are at most r^n different r-colourings and that $r^n \leqslant e^{\gamma^2 m/5}$ when n is sufficiently large. Consequently, summing (9) over all $\Pi \notin \mathcal{P}_{n,r}(\gamma)$ one gets the assertion of the proposition. □

4.4. **The number of graphs with many colourings.** Even though the collections $\mathcal{G}_m(\Pi, B)$ are generally not pairwise disjoint, there is not too much overlap between them. In other words, for all $\Pi \in \mathcal{P}_{n,r}(\gamma)$ and $B \in \mathcal{B}(\Pi, k)$, the pair (Π, B) is the unique pair which covers G for almost all $G \in \mathcal{G}_m(\Pi, B)$. More precisely, let $\mathcal{U}_m(\Pi, B)$ be the family of all graphs in $\mathcal{G}_m(\Pi, B)$ for which (Π, B) is the unique pair which covers them. The following result is based on a result implicit in the work of Prömel and Steger [19].

Proposition 4.3. *For all integers $k \geqslant 0$ and $r \geqslant 2$ and real number a, there exists a constant c such that the following holds for all $\Pi \in \mathcal{P}_{n,r}(\frac{1}{2r})$ and $B \in \mathcal{B}(\Pi, k)$. If $m \geqslant cn \log n$, then*

$$|\mathcal{G}_m(\Pi, B) \setminus \mathcal{U}_m(\Pi, B)| \leqslant n^{-a} \cdot |\mathcal{G}_m(\Pi, B)|.$$

Proof. Fix some $\Pi \in \mathcal{P}_{n,r}(\frac{1}{2r})$ and $\Pi' \in \mathcal{P}_{n,r} \setminus \{\Pi\}$. Suppose that $\Pi = \{V_1, \ldots, V_r\}$ and $\Pi' = \{V'_1, \ldots, V'_r\}$ and, for all $i, j \in [\![r]\!]$, let $V_{i,j} = V_i \cap V'_j$. We will say that the vertices in $V_{i,j}$ are moved from V_i to V'_j. For every $i \in [\![r]\!]$, define L_i and S_i as the largest and the second largest subclasses of V_i, respectively. Note that $|V_i| \geqslant \frac{n}{2r}$ implies that $|L_i| \geqslant \frac{n}{2r^2}$. Set $s = \max_{j \in [\![r]\!]} |S_j|$ and let $S = S_j$ for the smallest j for which the maximum in the definition of s is achieved. Note that $1 \leqslant s \leqslant n/2$, as $s = 0$ would imply that $(V'_1, \ldots, V'_r)$ is a permutation of $(V_1, \ldots, V_r)$, and therefore $\Pi = \Pi'$, which will imply also that $B = B'$.

Observe that either some pair $\{L_i, L_j\}$ of largest subclasses, or some largest subclass L_i and S, where $S \nsubseteq V_i$, are moved to the same vertex class V'_z. Denote these sets L_i and L_j or L_i and S by C and D. Since, for every $G \in \mathcal{G}_m(\Pi', B')$, the subgraph of G induced by V'_z has maximum degree at most k, it follows that, for every $G \in \mathcal{G}_m(\Pi, B) \cap \mathcal{G}_m(\Pi', B')$, the bipartite subgraph of G induced between C and D also has maximum degree at most k. In particular,

$$e(C, D) \leqslant k \cdot \min\{|C|, |D|\}.$$

It follows that

$$(\star) = \sum_{B' \in \mathcal{B}(\Pi', k)} |\mathcal{G}_m(\Pi, B) \cap \mathcal{G}_m(\Pi', B')| = |\mathcal{G}_m(\Pi, B) \cap \bigcup_{B' \in \mathcal{B}(\Pi', k)} \mathcal{G}_m(\Pi', B')|$$
$$\leqslant \sum_{t=0}^{k \cdot \min\{|C|,|D|\}} \binom{e(\Pi) - |C| \cdot |D|}{m - e(B) - t} \binom{|C| \cdot |D|}{t},$$

since every $G \in \mathcal{G}_m(\Pi, B)$ contains B and we need to specify its remaining $m - e(B)$ edges (by the definition of C and D, no edge of B connects these two sets). Consequently,

$$\begin{aligned}
(\star) &\overset{(2)}{\leqslant} \binom{e(\Pi) - |C| \cdot |D|}{m - e(B)} \cdot \sum_{t=0}^{k \cdot \min\{|C|,|D|\}} (m - e(B))^t \cdot \binom{|C| \cdot |D|}{t} \\
&\overset{(5)}{\leqslant} \binom{e(\Pi) - |C| \cdot |D|}{m - e(B)} \cdot \left(\frac{(m - e(B)) \cdot e \max\{|C|, |D|\}}{k} \right)^{k \cdot \min\{|C|,|D|\}} \\
&\overset{(3)}{\leqslant} \left(1 - \frac{|C| \cdot |D|}{n^2}\right)^{m - kn} \cdot \binom{e(\Pi)}{m - e(B)} \cdot e^{4k \cdot \min\{|C|,|D|\} \cdot \log n} \\
&\exp\left(-\frac{|C| \cdot |D| \cdot (m - kn)}{n^2} + 4k \cdot \min\{|C|, |D|\} \cdot \log n\right) \cdot \binom{e(\Pi)}{m - e(B)}.
\end{aligned}$$

If $m \geqslant cn \log n$ for a sufficiently large constant $c = c(k, r, a)$, then the simple bounds $\max\{|C|, |D|\} \geqslant \frac{n}{2r^2}$ and $\min\{|C|, |D|\} \geqslant \frac{s}{2r^2}$ imply that

$$\begin{aligned}
(\star) &\leqslant \exp\left(\left(-\frac{m/2}{2r^2 n} + 4k \log n\right) \cdot \min\{|C|, |D|\}\right) \cdot \binom{e(\Pi)}{m - e(B)} \\
&\leqslant n^{-(a+3)sr^2} \cdot \binom{e(\Pi)}{m - e(B)}.
\end{aligned}$$

Finally, observe that, given a Π, we can describe any $\Pi' \neq \Pi$ by first picking the partitions $\{V_{i,j}\}_{j \in [\![r]\!]}$ for every i and then setting $V'_j = \bigcup_{i \in [\![r]\!]} V_{i,j}$. We claim that, for every s, there are at most $n^{r^2} \cdot n^{sr^2}$ ways to choose all $V_{i,j}$ so that $\max_{i \in [\![r]\!]} |S_i| = s$. Indeed, one may first specify the sequence $\big(|V_{i,j}|\big)_{i,j \in [\![r]\!]}$ and then specify, for each $i \in [\![r]\!]$, the elements of each $V_{i,j}$ with $j \in [\![r]\!]$, apart from L_i (which will comprise all the remaining, unspecified elements of V_i). Therefore, by the above computation,

$$\begin{aligned}
|\mathcal{G}_m(\Pi, B) \setminus \mathcal{U}_m(\Pi, B)| &\leqslant \sum_{\substack{\Pi' \in \mathcal{P}_{n,r}(\gamma) \\ \Pi' \neq \Pi}} \sum_{B' \in \mathcal{B}(\Pi', k)} |\mathcal{G}_m(\Pi, B) \cap \mathcal{G}_m(\Pi', B')| \\
&\leqslant \sum_{s \geqslant 1} \left(n^{(s+1)r^2} \cdot n^{-s(a+3)r^2}\right) \cdot |\mathcal{G}_m(\Pi, B)| \leqslant n^{-a} \cdot |\mathcal{G}_m(\Pi, B)|,
\end{aligned}$$

as claimed. $\square$

4.5. **Typical degrees in almost r-colourable graphs.** We shall now show that most vertices of almost every graph in $\mathcal{G}_{n,m}(r,k)$ have degree exactly k in the monochromatic graph. To make this informal statement precise, given a positive number κ, denote by $\mathcal{B}(\Pi, k; \kappa)$ the family of all $B \in \mathcal{B}(\Pi, k)$ such that

$$\left|\left\{v \in [\![n]\!] : \deg_B(v) = k\right\}\right| \geqslant (1-\kappa)n.$$

Proposition 4.4. *For all integers $k \geqslant 0$ and $r \geqslant 2$, every positive κ, all $\Pi \in \mathcal{P}_{n,r}$, and every m satisfying $m \gg n$,*

$$\sum_{B \in \mathcal{B}(\Pi,k) \setminus \mathcal{B}(\Pi,k;\kappa)} |\mathcal{G}_m(\Pi, B)| \ll \sum_{B \in \mathcal{B}(\Pi,k)} |\mathcal{G}_m(\Pi, B)|. \tag{10}$$

Proof. Since $\mathcal{B}(\Pi, 0; \kappa) = \mathcal{B}(\Pi, 0)$, we may assume that $k \geqslant 1$. The left-hand and the right-hand sides of (10) are cardinalities of the (disjoint) unions of families $\mathcal{G}_m(\Pi, B)$ over all $B \in \mathcal{B}(\Pi, k) \setminus \mathcal{B}(\Pi, k; \kappa)$ and all $B \in \mathcal{B}(\Pi, k)$, respectively; denote these two families of graphs by $\mathcal{F}_L$ and $\mathcal{F}_R$. We will compare the sizes of $\mathcal{F}_L$ and $\mathcal{F}_R$ by counting edges in a bipartite graph $\mathcal{H} \subseteq \mathcal{F}_L \times \mathcal{F}_R$ defined as follows: A pair $(G_L, G_R) \in \mathcal{F}_L \times \mathcal{F}_R$ belongs to $\mathcal{H}$ if and only if $G_R \setminus G_L$ is a single edge of $\Pi^c \setminus G_L$ and $G_L \setminus G_R$ is a single edge of $\Pi \cap G_L$.

On the one hand, for every $G_R \in \mathcal{F}_R$,

$$\deg_{\mathcal{H}}(G_R) \leqslant e(\Pi) \cdot e(\Pi^c \cap G_R) \leqslant n^2 \cdot kn.$$

On the other hand, since every $B \in \mathcal{B}(\Pi, k) \setminus \mathcal{B}(\Pi, k; \kappa)$ contains more than κn vertices of degree smaller than k, at least $r \cdot \binom{\kappa n/r}{2}$ pairs of such vertices belong to the same colour class of Π. Consequently, for every $G_L \in \mathcal{F}_L$,

$$\begin{aligned}\deg_{\mathcal{H}}(G_L) &\geqslant \left(r \cdot \binom{\kappa n/r}{2} - e(\Pi^c \cap G_L)\right) \cdot e(\Pi \cap G_L) \\ &\geqslant \left(\frac{\kappa^2 n^2}{3r} - kn\right) \cdot (m - kn) \geqslant \frac{\kappa^2 n^2}{4r} \cdot \frac{m}{2}.\end{aligned}$$

We conclude that

$$|\mathcal{F}_L| \cdot \frac{\kappa^2 n^2 m}{8r} \leqslant e(\mathcal{H}) \leqslant |\mathcal{F}_R| \cdot kn^3,$$

which implies that $|\mathcal{F}_L| \leqslant (8kr/\kappa^2) \cdot (n/m) \cdot |\mathcal{F}_R| \ll |\mathcal{F}_R|$, as claimed. □

4.6. **Almost r-colourable graphs with large monochromatic girth.** We shall now show that in a constant proportion of graphs in $\mathcal{G}_{n,m}(r,k)$, the monochromatic graph has large girth. To make this informal statement precise, given an integer $g \geqslant 3$, denote by $\mathcal{B}_g(\Pi, k)$ the family of all graphs in $\mathcal{B}(\Pi, k)$ that do not contain any cycles of length at most g. The following statement is a key ingredient in our proof of Theorem 1.4.

Proposition 4.5. *For all integers $k \geqslant 0$, $r \geqslant 2$, and $g \geqslant 3$, there exists a positive constant c such that, for all $\Pi \in \mathcal{P}_{n,r}(\frac{1}{2r})$ and every m satisfying $m \gg n$,*

$$\sum_{B \in \mathcal{B}_g(\Pi,k)} |\mathcal{G}_m(\Pi, B)| \geqslant c \cdot \sum_{B \in \mathcal{B}(\Pi,k)} |\mathcal{G}_m(\Pi, B)|.$$

The proof of this proposition is a relatively straightforward corollary of Proposition 4.4 and the following classical result of Bollobás [6] and Wormald [23].

Theorem 4.6 ([6, Theorem 2]). *Suppose that $k \geqslant 2$ and $g \geqslant 3$ are integers and let $0 \leqslant d_1 \leqslant \cdots \leqslant d_n \leqslant k$ be such that $\sum_{i=1}^n d_i =: 2m$ is even and $2m - n \to \infty$ as $n \to \infty$. Let G be a graph chosen uniformly at random from the family of all graphs with vertex set $[\![n]\!]$ such that $\deg_G(i) = d_i$ for every $i \in [\![n]\!]$ and, for each $\ell \geqslant 3$, denote by X_ℓ the number of cycles of length ℓ in G. Denote by $(Z_3, \ldots, Z_g)$ the vector of independent Poisson random variables with*

$$\mathbb{E}[Z_\ell] = \frac{1}{2\ell}\left(\frac{1}{m}\sum_{i=1}^n \binom{d_i}{2}\right)^\ell$$

for each ℓ. Then

$$\lim_{n\to\infty} d_{TV}\big((X_3, \ldots, X_g), (Z_3, \ldots, Z_g)\big) = 0,$$

where d_{TV} is the total variation distance.

Proof of Proposition 4.5. We may assume that $k \geqslant 2$, since otherwise no graph in $\mathcal{B}(\Pi, k)$ can contain a cycle and thus $\mathcal{B}_g(\Pi, k) = \mathcal{B}(\Pi, k)$. Suppose that $\Pi = \{V_1, \ldots, V_r\} \in \mathcal{P}_{n,r}(\frac{1}{2r})$ and let G be a uniformly random element of $\bigcup_{B \in \mathcal{B}(\Pi,k)} \mathcal{G}_m(\Pi, B)$. Conditioned on $G \cap \Pi$ and the degree sequence of $G \cap \Pi^c$, the graphs $G[V_1], \ldots, G[V_r]$ become independent, uniformly chosen random graphs with respective degree sequences. By Proposition 4.4, invoked with $\kappa = 1/(6r)$, with probability $1 - o(1)$,

$$\sum_{v \in V_i} \deg_{G[V_i]}(v) \geqslant (|V_i| - \kappa n) \cdot k \geqslant \frac{2|V_i|}{3} \cdot k \geqslant \frac{4|V_i|}{3} \tag{11}$$

for each $i \in [\![r]\!]$, as $\min_i |V_i| \geqslant n/(2r) = 3\kappa n$. Since, for every $i \in [\![r]\!]$,

$$\frac{1}{e(G[V_i])}\sum_{v \in V_i}\binom{\deg_{G[V_i]}(v)}{2} \leqslant \frac{1}{e(G[V_i])}\sum_{v \in V_i}\frac{\deg_{G[V_i]}(v) \cdot (k-1)}{2} = k - 1,$$

Theorem 4.6 implies that, if the degree sequence of $G \cap \Pi^c$ satisfies (11) for every $i \in [\![r]\!]$, which happens with probability $1 - o(1)$,

$$\mathbb{P}(G \cap \Pi^c \in \mathcal{B}_g(\Pi, k) \mid G \cap \Pi, \text{degree sequence of } G \cap \Pi^c) \geqslant \left(\frac{1}{2} \cdot \inf\left\{\prod_{\ell=3}^g \mathbb{P}\big(\text{Pois}(\lambda) = 0\big) : \lambda \leqslant \frac{(k-1)^\ell}{2\ell}\right\}\right)^r,$$

where $\text{Pois}(\lambda)$ denotes the Poisson random variable with mean λ. The assertion of the proposition follows as $\mathbb{P}(\text{Pois}(\lambda) = 0) = e^{-\lambda}$. □

5. The 0-statement

In this section, we treat the 0-statement of Theorem 1.4. First, using an elementary counting argument, we show that, for every graph H with maximum degree at least two, if $m \ll n^{2-\frac{1}{m_2(H)}}$, then the family $\mathcal{F}_{n,m}(H)$ constitutes an $e^{-o(m)}$-proportion of all graphs with n vertices and m edges. Using a standard estimate on the lower tails of hypergeometric distributions, it will be fairly straightforward to deduce that, when $n \ll m \ll n^{2-\frac{1}{m_2(H)}}$ and both r and k are bounded, the family $\mathcal{G}_{n,m}(r, k)$ is far smaller than $\mathcal{F}_{n,m}(H)$. The details are presented in Section 5.1.

Second, using a much more subtle argument, we show that, for every plain vertex-critical graph H with criticality $k + 1$ and chromatic number $r + 1$, if $\Omega\big(n^{2-\frac{1}{m_2(H)}}\big) \leqslant$

$m \leqslant cn^{2-\frac{1}{\eta(H)}}(\log n)^{\frac{1}{\zeta(H)-k-1}}$ for a sufficiently small positive c, the number of graphs in $\mathcal{F}_{n,m}(H)$ that are 'one edge away' from being in $\mathcal{G}_{n,m}(r,k)$ is far greater than the number of graphs in $\mathcal{G}_{n,m}(r,k)$. Our argument, which relies on the Hypergeometric Harris Inequality as well as several crucial properties of graphs in $\mathcal{G}_{n,m}(r,k)$ that we have established in Section 4, is presented in Section 5.2.

5.1. **Below the 2-density.** We first give a simple lower bound on $|\mathcal{F}_{n,m}(H)|$, valid for every graph H with maximum degree at least two, that exploits the fact that, if $m \ll n^{2-\frac{1}{m_2(H)}}$, a typical graph in $\mathcal{G}_{n,m}$ can be made H-free by removing from it some $o(m)$ edges.

Proposition 5.1. *Let H be an arbitrary graph with maximum degree at least two. For every $\varepsilon > 0$, there is a $\delta > 0$ such that, for all sufficiently large m and every $m \leqslant \delta n^{2-\frac{1}{m_2(H)}}$,*

$$|\mathcal{F}_{n,m}(H)| \geqslant e^{-\varepsilon m} \cdot \binom{\binom{n}{2}}{m}.$$

Proof. Suppose that H is a graph with maximum degree at least two. This means that $K_{1,2} \subseteq H$ and hence $m_2(H) \geqslant m_2(K_{1,2}) \geqslant 1$. Suppose that ε is a positive number. Let F be an arbitrary subgraph of H such that $d_2(F) = m_2(H)$ and note that $e_F \geqslant 2$, as $m_2(H) \geqslant 1$. Finally, let δ be a small positive number satisfying

$$(6\delta)^{e_F-2} \leqslant \frac{1}{72} \qquad \text{and} \qquad \left(\frac{\delta}{e(1+2\delta)}\right)^{2\delta} \geqslant e^{-\varepsilon/2}. \tag{12}$$

Let m be a positive integer satisfying $m \leqslant \delta n^{2-\frac{1}{m_2(H)}}$. If $m \leqslant n^{1/3}$, we let G be a uniformly chosen random graph in $\mathcal{G}_{n,m}$ and note that

$$\mathbb{P}(H \subseteq G) \leqslant \mathbb{P}(K_{1,2} \subseteq G) \leqslant n^3 \cdot \left(\frac{m}{\binom{n}{2}}\right)^2 \leqslant 5n^{-1/3} \leqslant 1 - e^{-\varepsilon},$$

provided that n is sufficiently large. Consequently,

$$|\mathcal{F}_{n,m}(H)| = \mathbb{P}(H \nsubseteq G) \cdot \binom{\binom{n}{2}}{m} \geqslant e^{-\varepsilon m} \cdot \binom{\binom{n}{2}}{m},$$

as desired. We may thus assume from now on that $m > n^{1/3}$.

Set $m' = \lceil (1+\delta)m \rceil$ and note that

$$m' \leqslant (1+\delta)\delta n^{2-\frac{1}{m_2(H)}} + 1 \leqslant 2\delta n^{2-\frac{1}{m_2(H)}} = 2\delta n^{2-\frac{1}{d_2(F)}}$$

provided that n is sufficiently large. Now, let G be a uniformly chosen random graph in $\mathcal{G}_{n,m'}$, and let X denote the number of copies of F in G. Recalling that $d_2(F) = (e_F - 1)/(v_F - 2)$, we have

$$\begin{aligned}\mathbb{E}[X] &\leqslant n^{v_F} \cdot \left(\frac{m'}{\binom{n}{2}}\right)^{e_F} \leqslant n^{v_F} \cdot \left(\frac{3m'}{n^2}\right)^{e_F} \leqslant n^{v_F} \cdot \frac{3m'}{n^2} \cdot \left(6\delta n^{-\frac{1}{d_2(F)}}\right)^{e_F-1} \\ &= (6\delta)^{e_F-1} \cdot 3m' \overset{(12)}{\leqslant} \frac{\delta m'}{4} \leqslant \frac{\delta m}{2}\end{aligned}$$

and consequently, by Markov's inequality,

$$\mathbb{P}(X \geqslant m' - m) = \mathbb{P}(X \geqslant \delta m) \leqslant \frac{1}{2}.$$

We conclude that at least half of the graphs in $\mathcal{G}_{n,m'}$ contain a subgraph with m edges that is F-free and thus also H-free. (Indeed, we may delete an arbitrary edge from each of the at most $m' - m$ copies of F in the original graph). By double counting,

$$|\mathcal{F}_{n,m}(H)| \cdot \binom{\binom{n}{2} - m}{m' - m} \geqslant \frac{1}{2} \cdot \binom{\binom{n}{2}}{m'}.$$

It follows that, denoting $N = \binom{n}{2}$,

$$\frac{|\mathcal{F}_{n,m}(H)|}{\binom{N}{m}} \geqslant \frac{1}{2} \cdot \frac{\binom{N}{m'}}{\binom{N}{m}\binom{N-m}{m'-m}} = \frac{1}{2 \cdot \binom{m'}{m'-m}} \overset{(5)}{\geqslant} \frac{1}{2} \cdot \left(\frac{em'}{m' - m}\right)^{m-m'}.$$

Finally, since $(1 + \delta)m \leqslant m' \leqslant (1 + \delta)m + 1 \leqslant (1 + 2\delta)m$, we conclude that

$$\frac{|\mathcal{F}_{n,m}(H)|}{\binom{\binom{n}{2}}{m}} \geqslant \frac{1}{2} \cdot \left(\frac{\delta}{e(1 + 2\delta)}\right)^{2\delta m} \overset{(12)}{\geqslant} \frac{1}{2} \cdot e^{-\varepsilon m/2} \geqslant e^{-\varepsilon m},$$

provided that n is sufficiently large. □

In order to bound the number of graphs in $\mathcal{G}_{n,m}(r, k)$ from above, we use the simple observation that every graph in $\mathcal{G}_{n,m}(r, k)$ contains a set of at least n/r vertices that induces a graph with average degree at most k, which is much less than the expected average degree of a graph that such a set would induce in a uniformly chosen random graph from $\mathcal{G}_{n,m}$.

Proposition 5.2. *For all positive integers k, r, n, and m satisfying $m \geqslant 6r^2(k + 2)n$, we have*

$$|\mathcal{G}_{n,m}(r, k)| \leqslant \exp\left(-\frac{m}{4r^2}\right) \cdot \binom{\binom{n}{2}}{m},$$

provided that n is sufficiently large.

Proof. Observe that, for every graph $G \in \mathcal{G}_{n,m}(r, k)$, there is a set $W \subseteq [\![n]\!]$ with at least n/r elements such that $e(G[W]) \leqslant k|W|/2$. In particular, if G is a uniformly chosen random graph from $\mathcal{G}_{n,m}$,

$$|\mathcal{G}_{n,m}(r, k)| \leqslant \sum_{\substack{W \subseteq [\![n]\!] \\ |W| \geqslant n/r}} \mathbb{P}\big(e(G[W]) \leqslant k|W|/2\big) \cdot \binom{\binom{n}{2}}{m}.$$

We may bound each term in the above sum using Lemma 3.4, invoked with $\Omega = \binom{[\![n]\!]}{2}$ and $A = \binom{W}{2}$. Indeed, letting

$$t = \frac{m\binom{|W|}{2}}{\binom{n}{2}} - \frac{k|W|}{2},$$

we have

$$\mathbb{P}\big(e(G[W]) \leqslant k|W|/2\big) \leqslant \exp\left(-\frac{t^2}{2m\binom{|W|}{2}/\binom{n}{2}}\right) \leqslant \exp\left(-\frac{m\binom{|W|}{2}}{2\binom{n}{2}} + \frac{k|W|}{2}\right).$$

If n is sufficiently large, then, for every W with $n/r \leqslant |W| \leqslant n$,

$$\frac{m\binom{|W|}{2}}{2\binom{n}{2}} - \frac{k|W|}{2} \geqslant \frac{m}{2} \cdot \frac{n/r \cdot (n/r - 1)}{n \cdot (n - 1)} - \frac{kn}{2} \geqslant \frac{m}{3r^2} - \frac{kn}{2},$$

and, consequently,

$$|\mathcal{G}_{n,m}(r,k)| \leqslant 2^n \cdot \exp\left(-\frac{m}{3r^2} + \frac{kn}{2}\right) \cdot \binom{\binom{n}{2}}{m}.$$

The claimed bound now follows from our assumption that $m \geqslant 6r^2(k+2)n$. □

Propositions 5.1 and 5.2 immediately yield the following corollary.

Corollary 5.3. *Let H be an arbitrary graph with maximum degree at least two and let k and r be positive integers. There exists a positive constant c such that, if $n \ll m \leqslant cn^{2-\frac{1}{m_2(H)}}$,*

$$|\mathcal{F}_{n,m}(H)| \gg |\mathcal{G}_{n,m}(r,k)|.$$

5.2. **Above the 2-density.** In this section, we show that, if H is a plain vertex-critical graph with criticality $k+1$ and chromatic number $r+1 \geqslant 3$, then there exists a positive constant c_H such that $|\mathcal{F}_{n,m}(H)| \gg |\mathcal{G}_{n,m}(r,k)|$ for every m satisfying $\Omega\big(n^{2-\frac{1}{m_2(H)}}\big) \leqslant m \leqslant c_H m_H$. More precisely, we will show that the number of graphs in $\mathcal{F}_{n,m}(H)$ that are 'one edge away' from being in $\mathcal{G}_{n,m}(r,k)$, i.e., graphs $G \in \mathcal{F}_{n,m}(H)$ such that $G \setminus e \in \mathcal{G}_{n,m}(r,k)$ for some $e \in G$, is far greater than the number of graphs in $\mathcal{G}_{n,m}(r,k)$.

Proposition 5.4. *Suppose that H is a plain vertex-critical graph with criticality $k+1$ and chromatic number $r+1 \geqslant 3$. There is a positive constant c_H such that, if m satisfies*

$$\Omega\Big(n^{2-\frac{1}{m_2(H)}}\Big) \leqslant m \leqslant c_H m_H,$$

then $|\mathcal{F}_{n,m}(H)| \gg |\mathcal{G}_{n,m}(r,k)|$.

The main ingredient in our proof of this proposition is the following lower bound on the number of H-free graphs that are 'one edge away' from $\mathcal{G}_m(\Pi, B)$ for given balanced r-colouring Π and $B \in \mathcal{B}(\Pi, k)$ with large girth.

Lemma 5.5. *Suppose that H is a plain vertex-critical graph with criticality $k+1$ and chromatic number $r+1 \geqslant 3$. For every $\varepsilon > 0$, there exists a positive constant c such that the following holds for every m that satisfies*

$$n \log n \ll m \leqslant cn^{2-\frac{1}{\eta(H)}} (\log n)^{\frac{1}{\zeta(H)-k-1}}.$$

For every $\Pi \in \mathcal{P}_{n,r}(\frac{1}{20r})$, all $B \in \mathcal{B}(\Pi, k)$, and each $e \in \Pi^c \setminus B$ such that $B \cup e$ has girth larger than v_H,

$$|\mathcal{U}_m(\Pi, B \cup e) \cap \mathcal{F}_{n,m}(H)| \geqslant \frac{m}{n^{2+\varepsilon}} \cdot |\mathcal{G}_m(\Pi, B)|.$$

Proof. Note first that only the two endpoints of e can have degree larger than k in the graph $B \cup e$ and that, by assumption, the girth of $B \cup e$ is larger than v_H. The definition of plain vertex-critical graphs guarantees that, for every embedding φ of H into $\Pi \cup B \cup e$, there must be a critical star $S \subseteq H$ such that $\varphi(H) \cap (B \cup e) = \varphi(S)$ and $e \in \varphi(S)$; in particular, for every $S \subseteq F \subseteq H$, the map φ, restricted to $V(F)$, is also an embedding of F into $\Pi \cup B \cup e$ that maps S to $B \cup e$ and $F \setminus S$ to Π.

Let $S_1, \ldots, S_t$ be all the critical stars of H and, for each $i \in [\![t]\!]$, let F_i be a subgraph satisfying

$$S_i \subsetneq F_i \subseteq H, \qquad d_{k+2}(F_i) = \eta_i(H), \qquad \text{and} \qquad e_{F_i} = \zeta_i(H).$$

Let G be a uniformly chosen random element of $\mathcal{G}_m(\Pi, B \cup e)$, let $G' = G \cap \Pi$, and observe that G' is a uniformly random subgraph of Π with $m - e(B) - 1$ edges. For every i and every injection $\varphi\colon V(F_i) \to [\![n]\!]$, we let $S_{i,\varphi} = \varphi(S_i)$ and $K_{i,\varphi} = \varphi(F_i \setminus S_i)$ be the labeled graphs that are the images of S_i and $F_i \setminus S_i$ via the embedding φ. Define

$$\Phi_i = \{\varphi : S_{i,\varphi} \subseteq B \cup e \text{ and } K_{i,\varphi} \subseteq \Pi\};$$

in other words, Φ_i comprises all those embeddings of F_i into $\Pi \cup B \cup e$ that embed S_i into $B \cup e$ and map the remaining edges of F_i to Π. Since $B \cup e$ contains at most two copies of S_i, one for each endpoint of e, we have $|\Phi_i| \leqslant 2n^{v_{F_i}-k-2}$. More importantly, the above discussion implies that

$$|\mathcal{G}_m(\Pi, B \cup e) \cap \mathcal{F}_{n,m}(H)| \geqslant \underbrace{\mathbb{P}(G' \not\supseteq K_{i,\varphi} \text{ for all } i \text{ and } \varphi \in \Phi_i)}_{P} \cdot \binom{e(\Pi)}{m - e(B) - 1}.$$

Assume that $m \leqslant cn^{2-\frac{1}{\eta(H)}}(\log n)^{\frac{1}{\zeta(H)-k-1}}$, where

$$c = \frac{\varepsilon}{64t}. \tag{13}$$

We shall bound P from below using the Hypergeometric Harris Inequality (Lemma 3.2). To this end, let

$$p = \frac{3}{2} \cdot \frac{m - e(B) - 1}{e(\Pi)}$$

and note that $p \leqslant \frac{8m}{n^2}$, by part (i) of Proposition 4.1, as $\Pi \in \mathcal{P}_{n,r}(\frac{1}{20r})$. It follows from Lemma 3.2 that

$$\begin{aligned} P + \exp(-m/16) &\geqslant \prod_{i=1}^{t} (1 - p^{e_{F_i \setminus S_i}})^{|\Phi_i|} \geqslant \exp\left(-\sum_{i=1}^{t} |\Phi_i| \cdot 2p^{e_{F_i \setminus S_i}}\right) \\ &\geqslant \exp\left(-\sum_{i=1}^{t} 4n^{v_{F_i}-k-2} p^{e_{F_i}-k-1}\right). \end{aligned}$$

Claim 5.6. For every $i \in [\![t]\!]$,

$$n^{v_{F_i}-k-2} p^{e_{F_i}-k-1} \leqslant 8c \log n.$$

Proof. Since $e_{F_i} > e_{S_i} = k + 1$, we have

$$\begin{aligned} n^{v_{F_i}-k-2} p^{e_{F_i}-k-1} &\leqslant n^{v_{F_i}-k-2} \cdot \left(\frac{8cm}{n^2}\right)^{e_{F_i}-k-1} \\ &\leqslant 8c \cdot n^{v_{F_i}-k-2} \cdot \left(\frac{m}{n^2}\right)^{e_{F_i}-k-1} \\ &\leqslant 8c \cdot \left(n^{\frac{1}{d_{k+2}(F_i)}} \cdot \frac{m}{n^2}\right)^{e_{F_i}-k-1} \\ &\leqslant 8c \cdot \left(n^{\frac{1}{\eta_i(H)} - \frac{1}{\eta(H)}} \cdot (\log n)^{\frac{1}{\zeta(H)-k-1}}\right)^{\zeta_i(H)-k-1}. \end{aligned}$$

The claimed upper bound follows since $\eta_i(H) \geqslant \eta(H)$ and $\zeta_i(H) \leqslant \zeta(H)$ whenever $\eta_i(H) = \eta(H)$. □

In particular, assuming that n is large, we have

$$P \geqslant \exp(-32tc \log n) - \exp(-m/16) \overset{(13)}{\geqslant} n^{-\varepsilon/2} - \exp(-n)$$

Since $B \cup e \in \mathcal{B}(\Pi, k+1)$, we may now invoke Proposition 4.3 with $a = \varepsilon$ to obtain

$$\begin{aligned} |\mathcal{U}_m(\Pi, B \cup e) \cap \mathcal{F}_{n,m}(H)| &\geqslant |\mathcal{G}_m(\Pi, B \cup e) \cap \mathcal{F}_{n,m}(H)| - |\mathcal{G}_m(\Pi, B \cup e) \setminus \mathcal{U}_m(\Pi, B \cup e)| \\ &\geqslant (P - n^{-\varepsilon}) \cdot \binom{e(\Pi)}{m - e(B) - 1} \geqslant n^{-\varepsilon} \cdot \binom{e(\Pi)}{m - e(B) - 1}. \end{aligned}$$

Since $m - e(B) \geqslant m - kn \geqslant m/2$ and $e(\Pi) \leqslant n^2/2$, we may conclude that

$$|\mathcal{U}_m(\Pi, B \cup e) \cap \mathcal{F}_{n,m}(H)| \geqslant \frac{m}{n^{2+\varepsilon}} \cdot \binom{e(\Pi)}{m - e(B)} = \frac{m}{n^{2+\varepsilon}} \cdot |\mathcal{G}_m(\Pi, B)|,$$

as claimed. □

Proof of Proposition 5.4. If $\eta(H) < m_2(H)$, then $m_H = n^{2 - \frac{1}{m_2(H)}}$ and we may may simply invoke Corollary 5.3 and let $c_H = c_{5.3}$. If this is not the case, then $m_H = n^{2 - \frac{1}{\eta(H)}} (\log n)^{\frac{1}{\zeta(H) - k - 1}}$ and we let $c_H = c_{5.5}(\varepsilon)$, where $2\varepsilon = 1 - 1/m_2(H) > 0$.

Since, for all $\Pi \in \mathcal{P}_{n,r}$, every $B' \subseteq \Pi^c$ can be written as $B' = B \cup e$ with $B \in \mathcal{B}(\Pi, k)$ and $e \notin B$ in at most $e(B) + 1 \leqslant kn$ different ways, we have

$$|\mathcal{F}_{n,m}(H)| \geqslant \frac{1}{kn} \sum_{\Pi \in \mathcal{P}_{n,r}} \sum_{B \in \mathcal{B}(\Pi, k)} \sum_{e \in \Pi^c \setminus B} |\mathcal{U}_m(\Pi, B \cup e) \cap \mathcal{F}_{n,m}(H)|. \tag{14}$$

Further, observe that, for every $B \in \mathcal{B}_{v_H}(\Pi, k)$, the number of edges $e \in \Pi^c \setminus B$ such that $B \cup e$ has girth larger than v_H is at least

$$e(\Pi^c) - e(B) - n \cdot \sum_{\ell=2}^{v_H - 1} k(k-1)^{\ell - 1} \geqslant \frac{n^2}{4r}.$$

Let $\gamma = 1/(20r)$. Since $m = \Omega\big(n^{2 - \frac{1}{m_2(H)}}\big) = \Omega\big(n^{1+2\varepsilon}\big)$, we may conclude that

$$\begin{aligned} |\mathcal{F}_{n,m}(H)| &\overset{(14)}{\geqslant} \frac{1}{kn} \sum_{\Pi \in \mathcal{P}_{n,r}(\gamma)} \sum_{B \in \mathcal{B}_{v_H}(\Pi, k)} \frac{n^2}{4r} \cdot \frac{m}{n^{2+\varepsilon}} \cdot |\mathcal{G}_m(\Pi, B)| \\ &\overset{\text{L 5.5}}{\geqslant} \frac{m}{4krn^{1+\varepsilon}} \sum_{\Pi \in \mathcal{P}_{n,r}(\gamma)} \sum_{B \in \mathcal{B}_{v_H}(\Pi, k)} |\mathcal{G}_m(\Pi, B)| \\ &\overset{\text{P 4.5}}{\geqslant} \frac{c_{4.5} m}{4krn^{1+\varepsilon}} \sum_{\Pi \in \mathcal{P}_{n,r}(\gamma)} \sum_{B \in \mathcal{B}(\Pi, k)} |\mathcal{G}_m(\Pi, B)| \\ &\gg \sum_{\Pi \in \mathcal{P}_{n,r}(\gamma)} \sum_{B \in \mathcal{B}(\Pi, k)} |\mathcal{G}_m(\Pi, B)|. \end{aligned}$$

On the other hand,

$$|\mathcal{G}_{n,m}(r,k)| \leqslant \sum_{\Pi \in \mathcal{P}_{n,r}} \sum_{B \in \mathcal{B}(\Pi, k)} |\mathcal{G}_m(\Pi, B)| \overset{\text{P 4.2}}{\leqslant} 2 \sum_{\Pi \in \mathcal{P}_{n,r}(\gamma)} \sum_{B \in \mathcal{B}(\Pi, k)} |\mathcal{G}_m(\Pi, B)|.$$

These two estimates imply the assertion of the proposition. □

6. Approximate 1-statement

In this section, we show that, for every graph H with $\chi(H) = r + 1 \geqslant 3$, then, as soon as $m \gg n^{2 - \frac{1}{m_2(H)}}$, most graphs in $\mathcal{F}_{n,m}(H)$ admit a balanced, unfriendly r-colouring that leaves only $o(m)$ edges monochromatic.

Theorem 6.1. *Suppose that a graph H satisfies $\chi(H) = r + 1 \geqslant 3$. For all positive δ and γ, there is a positive C such that the following holds. If $m \geqslant Cn^{2-\frac{1}{m_2(H)}}$, then almost every graph G in $\mathcal{F}_{n,m}(H)$ admits a partition $\Pi \in \mathcal{P}_{n,r}(\gamma)$ such that*

$$e(G \setminus \Pi) \leqslant \delta m \tag{15}$$

and, letting $\Pi = \{V_1, \ldots, V_r\}$,

$$\deg_G(v, V_i) \leqslant \min_{j \neq i} \deg_G(v, V_j) \quad \text{for all } i \in [\![r]\!] \text{ and } v \in V_i. \tag{16}$$

Our proof of Theorem 6.1 relies on the following result established in [3, Theorem 1.7], which states that, for every H with $\chi(H) = r + 1 \geqslant 3$, when $m \gg n^{2-\frac{1}{m_2(H)}}$, then most graphs $G \in \mathcal{F}_{n,m}(H)$ admit an r-partition $\Pi \in \mathcal{P}_{n,r}$ such that $e(G \setminus \Pi) = o(m)$. With little extra work, we will show that, for most such G, one such partition Π is balanced (i.e., it belongs to $\mathcal{P}_{n,r}(\gamma)$ for some small γ) and unfriendly (i.e., it satisfies (16)).

Theorem 6.2. *For every graph H with $\chi(H) \geqslant 3$ and every positive δ, there exists a positive constant C such that the following holds. If $m \geqslant Cn^{2-\frac{1}{m_2(H)}}$, then almost every graph in $\mathcal{F}_{n,m}(H)$ can be made $(\chi(H) - 1)$-partite by removing from it at most δm edges.*

As a next step towards establishing Theorem 6.1, we now show that there are very few $G \in \mathcal{F}_{n,m}(H)$ admit a non-balanced partition Π that satisfies $e(G \setminus \Pi) \leqslant \delta m$.

Proposition 6.3. *Suppose that a graph H satisfies $\chi(H) = r + 1 \geqslant 3$. For all positive δ and γ and all $m \gg n$, almost every $G \in \mathcal{F}_{n,m}(H)$ does not admit a partition $\Pi \in \mathcal{P}_{n,r} \setminus \mathcal{P}_{n,r}(\gamma)$ that satisfies $e(G \setminus \Pi) \leqslant \delta m$.*

Proof. Since making δ smaller only strengthens the assertion of the theorem, we may assume without loss of generality that $\delta \leqslant \gamma^2 6$ and that

$$\delta \cdot \big(4 - \log(\gamma^2\delta)\big) - (1 - \delta) \cdot \frac{\gamma^2}{3} < -\frac{\gamma^2}{4}; \tag{17}$$

indeed, as $\delta \to 0$, the left-hand side of (17) converges to $-\gamma^2/3$.

Fix an arbitrary partition $\Pi \in \mathcal{P}_{n,r} \setminus \mathcal{P}_{n,r}(\gamma)$, that is, a $\Pi \in \mathcal{P}_{n,r}$ that does not satisfy (6) and recall from the proof of Proposition 4.2, see (7), that

$$e(\Pi) \leqslant \left(1 - \frac{\gamma^2}{3}\right) \cdot \mathrm{ex}(n, K_{r+1}) \leqslant \left(1 - \frac{\gamma^2}{3}\right) \cdot \mathrm{ex}(n, H).$$

Denote $N = \binom{n}{2}$ and $N' = \mathrm{ex}(n, H)$. The number X_Π of graphs $G \in \mathcal{F}_{n,m}(H)$ for which $e(G \setminus \Pi) \leqslant \delta m$ satisfies

$$(\star) = \binom{N'}{m}^{-1} \cdot X_\Pi \leqslant \sum_{t=0}^{\delta m} \frac{\binom{N}{t}\binom{e(\Pi)}{m-t}}{\binom{N'}{m}} = \sum_{t=0}^{\delta m} \frac{\binom{N}{t}\binom{e(\Pi)}{m-t}\binom{m}{t}}{\binom{N'}{m-t}\binom{N'-m+t}{t}}.$$

Note that, for every t, either $m - t \leqslant e(\Pi)$ or the corresponding summand is equal to zero. This observation and the above bound on $e(\Pi)$ imply that

$$\binom{e(\Pi)}{m-t}\binom{N'}{m-t}^{-1} \overset{(3)}{\leqslant} \left(\frac{e(\Pi)}{N'}\right)^{m-t} \leqslant \left(1 - \frac{\gamma^2}{3}\right)^{m-t}$$

and that

$$\binom{N}{t}\binom{N'-m+t}{t}^{-1} \leqslant \binom{N}{t}\binom{N'-e(\Pi)}{t}^{-1} \leqslant \binom{N}{t}\binom{\gamma^2 N'/3}{t}^{-1}$$
$$\overset{(4)}{\leqslant} \left(\frac{N}{\gamma^2 N'/3 - t}\right)^t \leqslant \left(\frac{6N}{\gamma^2 N'}\right)^t \leqslant \left(\frac{12}{\gamma^2}\right)^t,$$

as $t \leqslant \delta m \leqslant \delta N' \leqslant \gamma^2 N'/6$ and $N' \geqslant \mathrm{ex}(n, K_3) \geqslant N/2$. Consequently,

$$(\star) \leqslant \sum_{t=0}^{\delta m} \left(1 - \frac{\gamma^2}{3}\right)^{m-t} \left(\frac{12}{\gamma^2}\right)^t \binom{m}{t} \leqslant \left(1 - \frac{\gamma^2}{3}\right)^{(1-\delta)m} \left(\frac{12}{\gamma^2}\right)^{\delta m} \cdot \sum_{t=0}^{\delta m} \binom{m}{t}.$$

Since $\delta \leqslant \frac{1}{2}$, inequalities (5) and $\log(12e) \leqslant 4$ further imply that

$$(\star) \leqslant \left(1 - \frac{\gamma^2}{3}\right)^{(1-\delta)m} \left(\frac{12}{\gamma^2} \cdot \frac{e}{\delta}\right)^{\delta m} \leqslant \exp\left(\left(\delta \cdot (4 - \log(\gamma^2 \delta)) - (1-\delta) \cdot \frac{\gamma^2}{3}\right) \cdot m\right)$$
$$\overset{(17)}{\leqslant} \exp\left(-\frac{\gamma^2 m}{4}\right).$$

Finally, since there are at most r^n partitions $\Pi \in \mathcal{P}_{n,r}$ and at least $\binom{N'}{m}$ graphs in $\mathcal{F}_{n,m}(H)$ and since $m \gg n$, we have

$$\sum_{\Pi \in \mathcal{P}_{n,r} \setminus \mathcal{P}_{n,r}(\gamma)} X_\Pi \leqslant r^n \cdot e^{-\gamma^2 m/4} \cdot \binom{N'}{m} \leqslant e^{-\gamma^2 m/5} \cdot |\mathcal{F}_{n,m}(H)|,$$

which implies the assertion of the proposition. □

Proof of Theorem 6.1. Let $\mathcal{F}_{n,m}(H; \delta, \gamma)$ be the collection of all graphs $G \in \mathcal{F}_{n,m}(H)$ that satisfy (15) for some $\Pi \in \mathcal{P}_{n,r}(\gamma)$ but no $\Pi \in \mathcal{P}_{n,r} \setminus \mathcal{P}_{n,r}(\gamma)$. Let $C = C_{6.2}(\delta)$ and assume that $m \geqslant Cn^{2 - \frac{1}{m_2(H)}}$. Since Theorem 6.2 and Proposition 6.3 imply that almost all graphs in $\mathcal{F}_{n,m}(H)$ belong to $\mathcal{F}_{n,m}(H; \delta, \gamma)$, it is enough to show that every $G \in \mathcal{F}_{n,m}(H; \delta, \gamma)$ admits a partition $\Pi = \{V_1, \dots, V_r\} \in \mathcal{P}_{n,r}(\gamma)$ that satisfies both (15) and (16).

To see this, given an arbitrary $G \in \mathcal{F}_{n,m}(H; \delta, \gamma)$, let $\Pi \in \mathcal{P}_{n,r}$ be a partition that minimises $e(G \setminus \Pi)$ over all r-partitions of $[\![n]\!]$. Since $e(G \setminus \Pi) \leqslant \delta m$, by the definition of $\mathcal{F}_{n,m}(H; \delta, \gamma)$ and the minimality of Π, then $\Pi \in \mathcal{P}_{n,r}(\gamma)$, again by the definition of $\mathcal{F}_{n,m}(H; \delta, \gamma)$. Suppose that $\Pi = \{V_1, \dots, V_r\}$. If there were $i, j \in [\![r]\!]$ and $v \in V_i$ such that $\deg_G(v, V_i) > \deg_G(v, V_j)$, then the partition Π' obtained from Π by moving the vertex v from V_i to V_j would satisfy $e(G \setminus \Pi') < e(G \setminus \Pi)$, contradicting the minimality of Π. □

7. The 1-statement

In this section, we prepare for the proof of the 1-statement of Theorem 1.4. Our goal is to show that, if H is a simple vertex-critical graph with criticality $k+1$ and chromatic number $r+1 \geqslant 3$, then there is a positive constant C_H such that, if $m \geqslant C_H m_H$, then almost every graph from $\mathcal{F}_{n,m}(H)$ belongs to $\mathcal{G}_{n,m}(r,k)$; recall that m_H is the threshold function defined in (1). Note that it suffices to prove this statement only for graphs H that have no isolated vertices.

7.1. **A sufficient condition.** Given a positive constant δ and a balanced r-partition $\Pi = \{V_1, \ldots, V_r\} \in \mathcal{P}_{n,r}(\gamma)$, let $\mathcal{F}_{n,m}(H;\delta,\Pi)$ be the family of all $G \in \mathcal{F}_{n,m}(H)$ for which Π is an unfriendly partition that leaves at most δm edges of G monochromatic, that is,

$$\mathcal{F}_{n,m}(H;\delta,\Pi) = \big\{G \in \mathcal{F}_{n,m}(H) : (G,\Pi) \text{ satisfy (15) and (16)}\big\}$$

and let

$$\mathcal{F}^*_{n,m}(H;\delta,\Pi) = \big\{G \in \mathcal{F}_{n,m}(H;\delta,\Pi) : G \setminus \Pi \notin \mathcal{B}(\Pi,k)\big\}.$$

In other words, $\mathcal{F}^*_{n,m}(H;\delta,\Pi)$ comprises all those graphs $G \in \mathcal{F}_{n,m}(H;\delta,\Pi)$ for which the monochromatic subgraph of G induced by the r-colouring Π has maximum degree larger than k. The following proposition gives a sufficient condition for the assertion of the 1-statement of Theorem 1.4 to hold true, that is, a sufficient condition for the asymptotic inequality $|\mathcal{F}_{n,m}(H) \setminus \mathcal{G}_{n,m}(r,k)| \ll |\mathcal{F}_{n,m}(H)|$.

Proposition 7.1. *Suppose that H is a simple vertex-critical graph with $\chi(H) = r+1 \geqslant 3$ and criticality $k+1$. For all positive δ and γ, there exists a constant C such that the following holds when $m \geqslant Cn^{2-\frac{1}{m_2(H)}}$: Suppose that there is a function $\omega\colon \mathbb{N} \to (0,\infty)$ satisfying $\omega(n) \to \infty$ as $n \to \infty$ such that, for every $\Pi \in \mathcal{P}_{n,r}(\gamma)$, there exists a map $\mathcal{M}\colon \mathcal{F}^*_{n,m}(H;\delta,\Pi) \to \mathcal{B}(\Pi,k)$ that satisfies*

$$|\mathcal{M}^{-1}(B)| \leqslant \frac{1}{\omega(n)} \cdot \binom{e(\Pi)}{m - e(B)} \tag{18}$$

for every $B \in \mathcal{B}(\Pi,k)$. Then

$$|\mathcal{F}_{n,m}(H) \setminus \mathcal{G}_{n,m}(r,k)| \ll |\mathcal{F}_{n,m}(H)|.$$

Proof. Set $C = C_{6.1}(\delta,\gamma)$ and suppose that $m \geqslant Cn^{2-\frac{1}{m_2(H)}}$. We claim that

$$\mathcal{F}_{n,m}(H) \setminus \mathcal{G}_{n,m}(r,k) \subseteq \big(\mathcal{F}_{n,m}(H) \setminus \mathcal{F}_{n,m}(H;\delta,\gamma)\big) \cup \bigcup_{\Pi \in \mathcal{P}_{n,r}(\gamma)} \mathcal{F}^*_{n,m}(H;\delta,\Pi).$$

Indeed, if $G \in \mathcal{F}_{n,m}(H;\delta,\gamma) \setminus \mathcal{G}_{n,m}(r,k)$, then, on the one hand, $G \in \mathcal{F}_{n,m}(H;\delta,\Pi)$ for some $\Pi \in \mathcal{P}_{n,r}(\gamma)$ but, on the other hand, $G \setminus \Pi \notin \mathcal{B}(\Pi,k)$ and hence $G \in \mathcal{F}^*_{n,m}(H;\delta,\Pi)$. Since Theorem 6.1 states that

$$|\mathcal{F}_{n,m}(H) \setminus \mathcal{F}_{n,m}(H;\delta,\gamma)| \ll |\mathcal{F}_{n,m}(H)|,$$

it suffices if we show that our assumptions imply that

$$\sum_{\Pi \in \mathcal{P}_{n,r}(\gamma)} |\mathcal{F}^*_{n,m}(H;\delta,\Pi)| \ll |\mathcal{F}_{n,m}(H)|. \tag{19}$$

To this end, note first that the assumption that H is *simple* vertex-critical implies that, for all $\Pi \in \mathcal{P}_{n,r}$ and $B \in \mathcal{B}_{v_H}(\Pi,k)$, the graph $B \cup \Pi$ is H-free and, consequently,

$$\mathcal{U}_m(\Pi,B) \subseteq \mathcal{G}_m(\Pi,B) \subseteq \mathcal{F}_{n,m}(H).$$

Since the families $\mathcal{U}_m(\Pi,B)$ are pairwise-disjoint and

$$|\mathcal{U}_m(\Pi,B)| \overset{\text{P 4.3}}{\geqslant} \frac{1}{2} \cdot |\mathcal{G}_m(\Pi,B)| = \frac{1}{2}\binom{e(\Pi)}{m - e(B)},$$

we have

$$\begin{aligned}|\mathcal{F}_{n,m}(H)| &\geqslant \sum_{\Pi\in\mathcal{P}_{n,r}(\gamma)}\sum_{B\in\mathcal{B}_{v_H}(\Pi,k)} |\mathcal{U}_m(\Pi,B)| \\ &\geqslant \frac{1}{2}\sum_{\Pi\in\mathcal{P}_{n,r}(\gamma)}\sum_{B\in\mathcal{B}_{v_H}(\Pi,k)} \binom{e(\Pi)}{m-e(B)} \\ &\overset{\text{P 4.5}}{\geqslant} \frac{c_{4.5}}{2}\sum_{\Pi\in\mathcal{P}_{n,r}(\gamma)}\sum_{B\in\mathcal{B}(\Pi,k)} \binom{e(\Pi)}{m-e(B)}.\end{aligned} \tag{20}$$

Fix an arbitrary $\Pi \in \mathcal{P}_{n,r}(\gamma)$, let $\mathcal{M}$ be the map satisfying (18) for every $B \in \mathcal{B}(\Pi, k)$, and observe that

$$|\mathcal{F}^*_{n,m}(H;\delta,\Pi)| = \sum_{B\in\mathcal{B}(\Pi,k)} |\mathcal{M}^{-1}(B)| \leqslant \frac{1}{\omega(n)} \sum_{B\in\mathcal{B}(\Pi,k)} \binom{e(\Pi)}{m-e(B)}. \tag{21}$$

Summing (21) over all $\Pi \in \mathcal{P}_{n,r}(\gamma)$ and substituting it into (20) yields (19). □

7.2. **Splitting into the sparse and the dense cases.** In the remainder of this paper, we will define, for some sufficiently small positive constants δ and γ and every $\Pi \in \mathcal{P}_{n,r}(\gamma)$, a map $\mathcal{M}\colon \mathcal{F}^*_{n,m}(H;\delta,\Pi) \to \mathcal{B}(\Pi,k)$ and show that these maps satisfy the assumptions of Proposition 7.1. Unfortunately, our main argument, presented in Section 8, will work only under the assumption that $m \leqslant \mathrm{ex}(n,H) - \Omega(n^2)$; the (much easier) complementary case $m \geqslant \mathrm{ex}(n,H) - o(n^2)$ will be treated in Section 9.

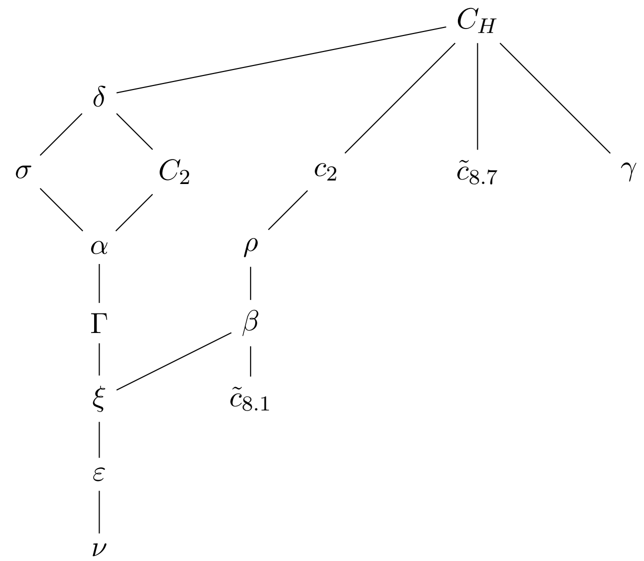

FIGURE 1. The Hasse diagram depicting dependence between the various constants in the proof

In order to formally define the split between the two cases, we need to introduce several additional parameters (cf. Figure 1). First, let γ be any positive constant satisfying

$$\gamma \leqslant \frac{1}{20r}. \tag{22}$$

Second, let ξ be a positive constant that satisfies inequalities (60) and the first inequality (61), which involve absolute constants ε and ν that are defined in (59). Third, let δ

be a small positive constants that also satisfies the first inequality in (61) and, moreover, the inequalities

$$\delta \leqslant \frac{1}{20r} \qquad \text{and} \qquad \delta \leqslant \frac{\xi\rho\tilde{c}_{8.7}}{70} \cdot \min\left\{\sigma, \frac{1}{C_2}\right\}, \tag{23}$$

where $\tilde{c}_{8.7}$ is an absolute positive constant implicit in the statement of Lemma 8.7, ρ is a constant that depends on ξ and on $\tilde{c}_{8.7}$ and is defined at the beginning of Section 8, and σ and C_2 are constant that depend on ξ and the function $(z, \alpha, \lambda) \mapsto \tau$ implicit in the statement of Lemma 3.3 and are defined in Section 8.6. Finally, define

$$C_H = \max\left\{C_{7.1}(\delta, \gamma),\ \frac{1}{\beta},\ \frac{1}{c_2 \cdot \tilde{c}_{8.7}} \cdot \frac{35r}{\xi}\right\}, \tag{24}$$

where $C_{7.1}(\delta, \gamma)$ is a constant that depends on δ and γ and is implicitly defined in the statement of Proposition 7.1, β is a constant that depends on ξ and on $\tilde{c}_{8.1}$ and is defined at the beginning of Section 8, and c_2 is a constant that depends on ρ (see above) and is defined in Section 8.6.

Fix an arbitrary $\Pi \in \mathcal{P}_{n,r}(\gamma)$. Our definition of the map $\mathcal{M}\colon \mathcal{F}^*_{n,m}(H; \delta, \Pi) \to \mathcal{B}(\Pi, k)$ and the arguments we will use to show that $\mathcal{M}$ satisfies the assumptions of Proposition 7.1 with $\omega(n) = 2/n$ will vary depending on whether

$$C_H m_H \leqslant m \leqslant e(\Pi) - \xi n^2 \qquad \text{or} \qquad e(\Pi) - \xi n^2 < m \leqslant \mathrm{ex}(n, H). \tag{25}$$

Our analysis under the assumption that m satisfies the first and the second pair of inequalities in (25) will be referred to as the *sparse case* and the *dense case*, respectively. These two cases will be treated in Sections 8 and 9, respectively.

8. The 1-statement: the sparse case

Fix a partition $\Pi \in \mathcal{P}_{n,r}(\gamma)$. In this section, we verify the assumptions of Proposition 7.1 in the case where

$$C_H m_H \leqslant m \leqslant e(\Pi) - \xi n^2.$$

In order to show that the assumptions of Proposition 7.1 are satisfied, we will first define a natural map $\mathcal{M}\colon \mathcal{F}^*_{n,m}(H; \delta, \Pi) \to \mathcal{B}(\Pi, k)$ by letting $\mathcal{M}(G)$ be an arbitrarily chosen maximal subgraph of $G \setminus \Pi$ with maximum degree k. We will estimate the left-hand side of (18) using two different arguments, depending on the distribution of edges in the monochromatic graph $G \setminus \Pi$: the *low-degree case* and the *high-degree case*.

Let $\mathcal{T}_\Pi$ denote the family of all $T \subseteq \Pi^c$ that are the monochromatic subgraph of some $G \in \mathcal{F}^*_{n,m}(H; \delta, \Pi)$, that is,

$$\mathcal{T}_\Pi = \left\{G \setminus \Pi : G \in \mathcal{F}^*_{n,m}(H; \delta, \Pi)\right\};$$

our definitions imply that every $T \in \mathcal{T}_\Pi$ satisfies $e(T) \leqslant \delta m$ and $\Delta(T) > k$. Define further, for every $T \in \mathcal{T}_\Pi$,

$$\mathcal{F}^*(T) = \left\{G \in \mathcal{F}^*_{n,m}(H; \delta, \Pi) : G \setminus \Pi = T\right\}$$

and observe that

$$|\mathcal{F}^*_{n,m}(H; \delta, \Pi)| = \sum_{T \in \mathcal{T}_\Pi} |\mathcal{F}^*(T)|.$$

In order to describe the split between the low-degree and the high-degree cases, let

$$\beta = \min\left\{\frac{e}{\xi(r-1)}, \frac{\tilde{c}_{8.1}}{22}\right\} \qquad \text{and} \qquad D = \left\lfloor \beta \frac{m}{n \log n} \right\rfloor, \tag{26}$$

where $\tilde{c}_{8.1}$ is an absolute positive constant that is implicit in the statement of Proposition 8.1, and choose a $\rho > 0$ which satisfies

$$\left(\frac{e}{\xi\rho}\right)^{\rho} \leqslant e^{\beta/2} \qquad \text{and} \quad \rho \leqslant \frac{1}{4r}; \tag{27}$$

it is possible to choose such ρ, since the left-hand side of the first inequality in (27) converges to 1 as $\rho \to 0$.

8.1. **Decomposing the monochromatic graphs.** For every $T \in \mathcal{T}_\Pi$, we define the following graphs and sets:

- Let B_T be an arbitrarily chosen maximal subgraph of T with $\Delta(B_T) = k$; note that $B_T \in \mathcal{B}(\Pi, k)$, as defined in Section 4.2.
- Let U_T be an arbitrarily chosen maximal subgraph of T that extends B_T and satisfies $\Delta(U_T) \leqslant D$.
- Let X_T be the set of vertices whose degrees in U_T are exactly D.
- Let H_T the set of all vertices whose degrees in T are larger than $\rho m/n$; note that $|H_T| \leqslant 2\delta n/\rho$.

Finally, for every $B \in \mathcal{B}(\Pi, k)$, let $\mathcal{T}_\Pi(B, t, \ell, h)$ denote the subfamily of $\mathcal{T}_\Pi$ comprising all T with

$$B_T = B, \qquad e(T) = t, \qquad e(U_T) = e(B) + \ell, \qquad \text{and} \qquad |H_T| = h.$$

The map $\mathcal{M}$ that we will supply to Proposition 7.1 is the map defined by $\mathcal{M}(G) = B_T$, where $T = G \setminus \Pi$.

8.2. **The low-degree and the high-degree cases.** We may now define the partition into the low-degree and the high-degree cases. Suppose that $T \in \mathcal{T}_\Pi$. We place T in $\mathcal{T}_{\mathrm{L}}(\Pi)$ when

$$e(U_T \setminus B_T) \log n \geqslant \frac{m|H_T|}{\xi n}; \tag{28}$$

otherwise, we place T in $\mathcal{T}_{\mathrm{H}}(\Pi)$. Since $\mathcal{T}_{\mathrm{L}}(\Pi)$ and $\mathcal{T}_{\mathrm{H}}(\Pi)$ form a partition of $\mathcal{T}_\Pi$, we have, for every $B \in \mathcal{B}(\Pi, k)$,

$$|\mathcal{M}^{-1}(B)| = \sum_{\substack{T \in \mathcal{T}_\Pi \\ B_T = B}} |\mathcal{F}^*(T)| = \sum_{\substack{T \in \mathcal{T}_{\mathrm{L}}(\Pi) \\ B_T = B}} |\mathcal{F}^*(T)| + \sum_{\substack{T \in \mathcal{T}_{\mathrm{H}}(\Pi) \\ B_T = B}} |\mathcal{F}^*(T)|. \tag{29}$$

The low-degree and the high-degree cases are estimates of the first and the second sums in the right hand side of (29), respectively.

8.3. **The low-degree case – summary.** In the low-degree case, we will rely on the following upper bound on $|\mathcal{F}^*(T)|$, which is established in Section 8.5 with the use of the Hypergeometric Janson Inequality (Lemma 3.1).

Proposition 8.1. *There exist positive constants $\tilde{c}$ and $\tilde{C}$ that depend only on H such that the following holds. If $m \geqslant \tilde{C} m_H$ for some $\tilde{C} \geqslant 2$, then, for every $\Pi \in \mathcal{P}_{n,r}(\gamma)$, every $B \in \mathcal{B}(\Pi, k)$, all $t \leqslant m/2$, ℓ, and h, and every $T \in \mathcal{T}_\Pi(B, t, \ell, h)$,*

$$|\mathcal{F}^*(T)| \leqslant \exp\left(-\frac{\tilde{c}}{\beta + \tilde{C}^{-1}} \cdot \ell \log n\right) \cdot \binom{e(\Pi)}{m - t}.$$

This upper bound on $|\mathcal{F}^*(T)|$ will be combined with the following estimate on the size of the sum over all $T \in \mathcal{T}_{\mathrm{L}}(\Pi)$, which is derived in Section 8.4.

Lemma 8.2. *Suppose that $n \log n \ll m \leqslant e(\Pi) - \xi n^2$. For every $B \in \mathcal{B}(\Pi, k)$ and all t, ℓ, and h,*

$$|\mathcal{T}_\Pi(B,t,\ell,h)| \cdot \binom{e(\Pi)}{m-t} \leqslant \exp\left(14\ell \log n + \frac{2mh}{\xi n}\right) \cdot \binom{e(\Pi)}{m-e(B)}.$$

Before we close this section, we show how these two lemmas can be used to estimate the first sum in the right-hand side of (29). Let $\mathcal{L}$ be the family of all triples (t,ℓ,h) that satisfy $t \geqslant \ell \geqslant 1$ and $\ell \log n \geqslant mh/(\xi n)$, cf. (28), and observe that, for every $B \in \mathcal{B}(\Pi, k)$,

$$\sum_{\substack{T \in \mathcal{T}_\mathrm{L}(\Pi) \\ B_T = B}} |\mathcal{F}^*(T)| = \sum_{(t,\ell,h)\in\mathcal{L}} \underbrace{\sum_{T \in \mathcal{T}_\Pi(B,t,\ell,h)} |\mathcal{F}^*(T)|}_{X_{t,\ell,h}}.$$

Since $m \geqslant C_H m_H$, Proposition 8.1 and Lemma 8.2 imply that, for every $(t,\ell,h) \in \mathcal{L}$,

$$\begin{aligned} X_{t,\ell,h} &\leqslant |\mathcal{T}_\Pi(B,t,\ell,h)| \cdot \exp\left(-\frac{\tilde{c}}{\beta + C_H^{-1}} \cdot \ell \log n\right) \cdot \binom{e(\Pi)}{m-t} \\ &\leqslant \exp\left(14\ell\log n + \frac{2mh}{\xi n} - \frac{\tilde{c}}{\beta + C_H^{-1}} \cdot \ell \log n\right) \cdot \binom{e(\Pi)}{m-e(B)} \\ &\overset{(24)}{\leqslant} \exp\left(\left(16 - \frac{\tilde{c}}{2\beta}\right) \cdot \ell \log n\right) \cdot \binom{e(\Pi)}{m-e(B)} \overset{(26)}{\leqslant} n^{-6\ell} \cdot \binom{e(\Pi)}{m-e(B)}. \end{aligned}$$

Since $\ell \geqslant 1$ for every $(t,\ell,h) \in \mathcal{L}$, we may conclude that

$$\sum_{\substack{T \in \mathcal{T}_\mathrm{L}(\Pi) \\ B_T = B}} |\mathcal{F}^*(T)| \leqslant |\mathcal{L}| \cdot n^{-6} \cdot \binom{e(\Pi)}{m-e(B)} \leqslant \frac{1}{n} \cdot \binom{e(\Pi)}{m-e(B)}.$$

8.4. **Enumerating the monochromatic graphs.** In this short section, we enumerate graphs in $\mathcal{T}_\Pi(B,t,\ell,h)$, proving Lemma 8.2.

Proof of Lemma 8.2. We will count the number of ways to construct any $T \in \mathcal{T}_\Pi(B,t,\ell,h)$ in several steps. We first record the following inequality, which holds for all integers $y \leqslant m' \leqslant m$:

$$\frac{\binom{e(\Pi)}{m'-y}}{\binom{e(\Pi)}{m'}} = \frac{\binom{m'}{y}}{\binom{e(\Pi)-m'+y}{y}} \leqslant \frac{\binom{m}{y}}{\binom{\xi n^2 + y}{y}} \overset{(4)}{\leqslant} \left(\frac{m}{\xi n^2}\right)^y. \tag{30}$$

Since $B = B_T \subseteq T$ for every $T \in \mathcal{T}_\Pi(B,t,\ell,h)$, we only need to choose which $t - e(B)$ edges of Π^c form the graph $T \setminus B$. For every $T \in \mathcal{T}_\Pi(B,t,\ell,h)$, let U_T' be the subgraph of $U_T \setminus B$ obtained by removing all edges touching X_T. Since every edge of $U_T \setminus U_T'$ has at least one endpoint in X_T and $\Delta(B) \leqslant k$, we have

$$\ell = e(U_T \setminus B) \geqslant e(U_T') + |X_T| \cdot (D-k)/2 \geqslant e(U_T') + |X_T| \cdot D/3.$$

We choose the edges of $T \setminus B$ in three steps:

(S1) We choose the edges of U_T'.
(S2) We choose the edges of $T \setminus B$ that touch $X_T \setminus H_T$.
(S3) We choose the remaining edges of $T \setminus B$; they all touch H_T.

We count the number of ways to build a graph $T \in \mathcal{T}_\Pi(B,t,\ell,h)$ with u', t_X, and t_H edges chosen in steps (S1), (S2), and (S3), respectively. An upper bound on $|\mathcal{T}_\Pi(B,t,\ell,h)|$ will be obtained by summing over all choices for u', t_X, and t_H. There are at most $\binom{e(\Pi^c)}{u'}$ ways to choose u' edges of U'_T. Since $e(\Pi^c) \leqslant n^2$, we have

$$\binom{e(\Pi^c)}{u'} \cdot \frac{\binom{e(\Pi)}{m-e(B)-u'}}{\binom{e(\Pi)}{m-e(B)}} \overset{(30)}{\leqslant} n^{2u'} \left(\frac{m}{\xi n^2}\right)^{u'} = \left(\frac{m}{\xi}\right)^{u'} \leqslant m^{2u'}.$$

Next, we bound the number of ways to choose the t_X edges that touch $X_T \setminus H_T$. To this end, we arbitrarily order the vertices of $X_T \setminus H_T$ as $v_1, \ldots, v_s$ and then, for each $i \in [\![s]\!]$, we choose the edges incident to v_i and not to any of $v_1, \ldots, v_{i-1}$; denote the number of such edges by d_i. Since we are considering only vertices of $X_T \setminus H_T$, we have $d_i \leqslant \rho m/n$; moreover, $d_1 + \cdots + d_s = t_X$. Let $N_2 = N_2(t_X, s)$ denote the total number of ways to choose the t_X edges when $|X_T \setminus H_T| = s$. We have

$$\begin{aligned} N_2 \cdot \frac{\binom{e(\Pi)}{m-e(B)-u'-t_X}}{\binom{e(\Pi)}{m-e(B)-u'}} &\leqslant \binom{n}{s} \sum_{\substack{d_1,\ldots,d_s \leqslant \rho m/n \\ d_1+\cdots+d_s = t_X}} \prod_{i=1}^{s} \binom{n}{d_i} \cdot \frac{\binom{e(\Pi)}{m-e(B)-u'-(d_1+\cdots+d_i)}}{\binom{e(\Pi)}{m-e(B)-u'-(d_1+\cdots+d_{i-1})}} \\ &\leqslant \left(n \cdot \sum_{d=0}^{\rho m/n} \binom{n}{d} \cdot \max_{m' \leqslant m} \frac{\binom{e(\Pi)}{m'-d}}{\binom{e(\Pi)}{m'}}\right)^s \\ &\overset{(30)}{\leqslant} \left(n + n \cdot \sum_{d=1}^{\rho m/n} \left(\frac{en}{d} \cdot \frac{m}{\xi n^2}\right)^d\right)^s. \end{aligned}$$

Since, for every positive a, the function $x \mapsto (ea/x)^x$ is increasing on the interval $(0,a]$ and $\rho < e/\xi$, we conclude that

$$\left(N_2 \cdot \frac{\binom{e(\Pi)}{m-e(B)-u'-t_X}}{\binom{e(\Pi)}{m-e(B)-u'}}\right)^{1/s} \leqslant n + \rho m \cdot \left(\frac{e}{\xi\rho}\right)^{\rho m/n} \leqslant \left(\frac{e}{\xi\rho}\right)^{2\rho m/n} \overset{(27)}{\leqslant} e^{\beta m/n} \leqslant m^D.$$

Finally, let $N_3 = N_3(t_X, h)$ denote the number of ways to choose the remaining t_H edges of $T \setminus B$. Recalling that $e(B) + u' + t_X + t_H = t$ and arguing similarly as above, we obtain

$$N_3 \cdot \frac{\binom{e(\Pi)}{m-t}}{\binom{e(\Pi)}{m-e(B)-u'-t_X}} \leqslant \left(n + n \cdot \sum_{d=1}^{n-1} \left(\frac{en}{d} \cdot \frac{m}{\xi n^2}\right)^d\right)^h.$$

Using again the fact that $(ea/x)^x \leqslant e^a$ for all $x \in (0,\infty)$, we conclude that

$$\left(N_3 \cdot \frac{\binom{e(\Pi)}{m-t}}{\binom{e(\Pi)}{m-e(B)-u'-t_X}}\right)^{1/h} \leqslant n + n^2 \cdot \exp\left(\frac{m}{\xi n}\right) \leqslant \exp\left(\frac{2m}{\xi n}\right).$$

Combining the above bounds, we obtain

$$|\mathcal{T}_\Pi(B,t,\ell,h)| \cdot \frac{\binom{e(\Pi)}{m-t}}{\binom{e(\Pi)}{m-e(B)}} \leqslant \sum_{\substack{u',t_X,t_H,s \\ u'+t_X+t_H=t \\ u'+sD/3\leqslant \ell}} m^{2u'} \cdot m^{Ds} \cdot \exp\left(\frac{2mh}{\xi n}\right)$$
$$\leqslant nm^3 \cdot m^{3\ell} \cdot \exp\left(\frac{2mh}{\xi n}\right) \leqslant \exp\left(14\ell \log n + \frac{2mh}{\xi n}\right),$$

where the final inequality follows as $nm^3 \leqslant m^4 \leqslant m^{4\ell}$ and $m \leqslant n^2$. □

8.5. **The low-degree case.** In this section, we prove Proposition 8.1, that is, for a given $T \in \mathcal{T}_\Pi(B,t,\ell,h)$, we give an upper bound on the number of graphs in $\mathcal{F}^*(T)$ in terms of t and ℓ. To this end, fix an arbitrary critical star S_{i_0} in H that satisfies

$$\eta_{i_0}(H) = \eta(H) \qquad \text{and} \qquad \zeta_{i_0}(H) = \zeta(H),$$

where $\eta(H)$ and $\zeta(H)$ are the quantities defined above (1). Fix some $T \in \mathcal{T}_\Pi$. For every injection $\varphi\colon V(H) \to [\![n]\!]$, we let $S_\varphi = \varphi(S_{i_0})$ and $K_\varphi = \varphi(H \setminus S_{i_0})$ be the labeled graphs that are the images of S_{i_0} and $H \setminus S_{i_0}$ via the embedding φ. Define

$$\Phi_T = \{\varphi : S_\varphi \subseteq T \text{ and } K_\varphi \subseteq \Pi\};$$

in other words, Φ_T comprises all those embeddings of H into $\Pi \cup T$ that embed S_{i_0} into T and map the remaning edges of H to Π. Since $T \subseteq G$ for every $G \in \mathcal{F}^*(T)$, the graph $G \cap \Pi$ does not contain any of the K_φ with $\varphi \in \Phi_T$. In particular, letting G' be a uniformly chosen random subgraph of Π with $m-t$ edges, we have

$$|\mathcal{F}^*(T)| \leqslant \mathbb{P}\big(K_\varphi \nsubseteq G' \text{ for each } \varphi \in \Phi_T\big) \cdot \binom{e(\Pi)}{m-t}. \tag{31}$$

Proposition 8.1 is derived from (31) and the Hypergeometric Janson Inequality. In order to get a strong bound on the probability in the right-hand side of (31), we will carefully construct a sub-family of Φ_T that satisfies some 'nice' properties and apply Janson's inequality with Φ_T replaced by this sub-family.

Lemma 8.3. *Suppose that* $\Pi = \{V_1,\dots,V_r\}$ *and let* $T \in \mathcal{T}_\Pi(B,t,\ell,h)$. *There are an* $i \in [\![r]\!]$ *and a family* $\mathcal{S}$ *of edge-disjoint copies of* $K_{1,k+1}$ *in* $T[V_i]$ *that satisfy the following properties for some positive constants* c_1 *and* C_1 *that depend only on* r *and* k*:*

- *(GS1) We have* $c_1\ell \leqslant |\mathcal{S}| \leqslant \ell$.
- *(GS2) For every* $v \in [\![n]\!]$, *we have* $|\{S \in \mathcal{S} : v \in V(S)\}| \leqslant D = \left\lfloor \beta \frac{m}{n \log n} \right\rfloor$.
- *(GS3) For every two different vertices* $v, u \in [\![n]\!]$, *let* $A(u,v)$ *be the set of all pairs of stars* $S, S' \in \mathcal{S}$, *each containing both* u *and* v *as leaves. Then,* $\sum_{u,v} |A(u,v)| \leqslant C_1\ell$.

We will first derive Proposition 8.1 from Lemma 8.3 and and then prove the lemma.

Proof of Proposition 8.1. Suppose that $\Pi = \{V_1,\dots,V_r\}$ and let $T \in \mathcal{T}_\Pi(B,t,\ell,h)$ be a graph with at most $m/2$ edges. Let i, $\mathcal{S}$, c_1, and C_1 be the colour class, the family of stars, and the two constants from the statement of Lemma 8.3, respectively. Fix an arbitrary colouring $\psi\colon V(H) \to [\![r]\!]$ that leaves only the edges of S_{i_0} monochromatic and such that the vertices of S_{i_0} are coloured i; such a colouring exists because S_{i_0} is a critical star of H. For every $j \in [\![r]\!]$, randomly choose an equipartition $\{V_{j,w}\}_{w \in V(H)}$ of V_j into v_H parts. We let Φ'_T be the family of all embeddings $\varphi \in \Phi_T$ that satisfy

$$S_\varphi \in \mathcal{S} \qquad \text{and} \qquad \varphi(w) \in V_{\psi(w),w} \text{ for every } w \in V(H).$$

Let $n' = \min\{|V| : V \in \Pi\} \geqslant n/(2r)$. Since there are at least $|\mathcal{S}| \cdot (n' - v_H)^{v_H-(k+2)}$ embeddings $\varphi \in \Phi_T$ such that $S_\varphi \in \mathcal{S}$ and $\varphi(w) \in V_{\psi(w)}$ for every $w \in V(H)$ and, for each such φ, the probability that $\varphi \in \Phi'_T$ is at least $v_H^{-v_H}$, there is a positive constant c that depends only on H such that

$$\mathbb{E}\big[|\Phi'_T|\big] \geqslant c\ell n^{v_H-k-2}.$$

We now fix some partitions $\{V_{j,w}\}_{w \in V(H)}$ for which $|\Phi'_T|$ is at least as large as its expectation and we let

$$\mathcal{S}' = \big\{S_\varphi : \varphi \in \Phi'_T\big\} \qquad \text{and} \qquad \mathcal{K}' = \big\{K_\varphi : \varphi \in \Phi'_T\big\}.$$

We claim that $K_\varphi \neq K_{\varphi'}$ for each pair of distinct $\varphi, \varphi' \in \Phi'_T$. To see this, note first that, since S_{i_0} is a critical star, every vertex in $V(S_{i_0})$ must have a neighbour in $\psi(j)^{-1}$, for each $j \in [\![r]\!] \setminus \{i\}$. Since H has no isolated vertices, this means that each vertex of H is incident to an edge of $H \setminus S_{i_0}$. Therefore, since each $\varphi \in \Phi'_T$ maps every $w \in V(H)$ to its dedicated set $V_{\psi(w),w}$, one can recover φ from the graph K_φ. This means, in particular, that

$$|\mathcal{K}'| = |\Phi'_T| \geqslant c\ell n^{v_H-k-2}. \tag{32}$$

Suppose that $m \geqslant \tilde{C} m_H$ for some $\tilde{C} \geqslant 2$ and let G' be a uniformly chosen random subgraph of Π with $m-t$ edges. The definition of $\mathcal{K}'$ and (31) imply that

$$|\mathcal{F}^*(T)| \leqslant \mathbb{P}(K \nsubseteq G' \text{ for every } K \in \mathcal{K}') \cdot \binom{e(\Pi)}{m-t}.$$

We shall bound this probability from above using the Hypergeometric Janson Inequality. To this end, let $p = \frac{m-e(T)}{e(\Pi)}$ and note that

$$\frac{m}{2n^2} \leqslant p \leqslant \frac{5m}{n^2}, \tag{33}$$

where the first inequality holds because $e(T) \leqslant m/2$ and the last inequality follows from part (i) of Proposition 4.1, as $\Pi \in \mathcal{P}_{n,r}(\gamma)$ and $\gamma \leqslant \frac{1}{20r}$. For any $K, K' \in \mathcal{K}'$, we write $K \sim K'$ if K and K' share an edge but $K \neq K'$. Let μ and Δ be the quantities defined in the statement of the Hypergeometric Janson Inequality (Lemma 3.1), that is,

$$\mu = \sum_{K \in \mathcal{K}'} p^{e_K} \qquad \text{and} \qquad \Delta = \sum_{\substack{K,K' \in \mathcal{K}' \\ K \sim K'}} p^{e_{K \cup K'}}.$$

Since $e_K = e(H \setminus S_{i_0}) = e_H - k - 1$ for every $K \in \mathcal{K}'$, we have, by (32),

$$\mu = |\mathcal{K}'| \cdot p^{e_H-k-1} \geqslant c\ell n^{v_H-k-2} p^{e_H-k-1}. \tag{34}$$

We now bound Δ from above. In order to do this, we shall classify the pairs $(K, K') \in (\mathcal{K}')^2$ with $K \sim K'$ according to their intersection. To this end, for each $J \subseteq V(S_{i_0})$, define $\mathcal{S}'(J)$ to be the set of all pairs of stars from $\mathcal{S}'$ which agree exactly on (the image of) J, that is,

$$\mathcal{S}'(J) = \big\{(S_\varphi, S_{\varphi'}) \in \mathcal{S}' \times \mathcal{S}' : S_\varphi \cap S_{\varphi'} = \varphi(H[J]) = \varphi'(H[J])\big\}.$$

Further, given $J \subseteq V(S_{i_0})$ and $I \subseteq V(H) \setminus V(S_{i_0})$, let $F_{I,J} = H[I \cup J] \setminus S_{i_0}$, that is, $F_{I,J}$ is a graph with vertex set $I \cup J$ that comprises the edges of $H \setminus S_{i_0}$ with both endpoints in $I \cup J$. (Let us note here that $F_{I,J}$ may have some isolated vertices.) Finally, for $J \subseteq V(S_{i_0})$, $I \subseteq V(H) \setminus V(S_{i_0})$, and $S, S' \in \mathcal{S}'(J)$, define $\mathcal{K}(I,J,S,S')$ to be the set of all pairs $K, K' \in \mathcal{K}'$ which extend the stars S, S', respectively, and agree exactly on (the image of) $I \cup J$. In other words,

$$\mathcal{K}(I,J,S,S') = \big\{(K_\varphi, K'_\varphi) \in (\mathcal{K}')^2 : K_\varphi \cap K_{\varphi'} = \varphi(F_{I,J}) = \varphi'(F_{I,J}), S_\varphi = S, S_{\varphi'} = S'\big\}.$$

For brevity, set

$$v' = |V(H) \setminus V(S_{i_0})| = v_H - k - 2 \qquad \text{and} \qquad e' = e(H \setminus S_{i_0}) = e_H - k - 1.$$

These definitions were made in such a way that

$$\begin{aligned}\Delta &= \sum_{J \subseteq V(S_{i_0})} \sum_{(S,S') \in \mathcal{S}'(J)} \sum_{\substack{I \subseteq V(H) \setminus V(S_{i_0}) \\ e(F_{I,J})>0}} \sum_{(K,K') \in \mathcal{K}(I,J,S,S')} p^{2e' - e(F_{I,J})} \\ &\leqslant \sum_{J \subseteq V(S_{i_0})} \sum_{(S,S') \in \mathcal{S}'(J)} \sum_{\substack{I \subseteq V(H) \setminus V(S_{i_0}) \\ e(F_{I,J})>0}} n^{2v' - |I|} p^{2e' - e(F_{I,J})}.\end{aligned} \tag{35}$$

Denote by Δ_0, Δ_1, and Δ_2 the contributions to the sum in the right-hand side of (35) corresponding to $J = \emptyset$, $|J| = 1$, and $|J| \geqslant 2$, respectively, so that $\Delta \leqslant \Delta_0 + \Delta_1 + \Delta_2$. Since $F_{I,\emptyset} = H[I] \subseteq H$, we have

$$\begin{aligned}\frac{\Delta_0}{n^{2v'} p^{2e'}} &= \sum_{(S,S') \in \mathcal{S}'(\emptyset)} \sum_{\substack{I \subseteq V(H) \setminus V(S_{i_0}) \\ e(F_{I,\emptyset})>0}} \frac{1}{n^{|I|} p^{e(F_{I,\emptyset})}} \leqslant |\mathcal{S}'(\emptyset)| \cdot \sum_{\emptyset \neq F \subseteq H} \frac{1}{n^{v_F} p^{e_F}} \\ &\leqslant |\mathcal{S}|^2 \cdot \frac{2^{e_H}}{\min_{\emptyset \neq F \subseteq H} n^{v_F} p^{e_F}} \overset{\text{L. 3.5}}{\leqslant} |\mathcal{S}|^2 \cdot \frac{2^{e_H}}{n^2 p} \leqslant \ell^2 \cdot \frac{2^{e_H}}{n^2 p},\end{aligned}$$

where the last inequality follows from (GS1) in Lemma 8.3. Further, as $v_{F_{I,J}} = |I| + |J|$,

$$\begin{aligned}\frac{\Delta_1}{n^{2v'} p^{2e'}} &= \sum_{\substack{J \subseteq V(S_{i_0}) \\ |J|=1}} \sum_{(S,S') \in \mathcal{S}'(J)} \sum_{\substack{I \subseteq V(H) \setminus V(S_{i_0}) \\ e(F_{I,J})>0}} \frac{n}{n^{v_{F_{I,J}}} p^{e_{F_{I,J}}}} \\ &\leqslant \sum_{S \in \mathcal{S}} \sum_{v \in V(S)} \sum_{\substack{S' \in \mathcal{S} \\ v \in V(S')}} \sum_{\emptyset \neq F \subseteq H} \frac{n}{n^{v_F} p^{e_F}} \\ &\leqslant |\mathcal{S}| \cdot (k+2) \cdot \max_v |\{S' \in \mathcal{S} : v \in V(S')\}| \cdot \frac{2^{e_H} \cdot n}{\min_{\emptyset \neq F \subseteq H} n^{v_F} p^{e_F}} \\ &\leqslant \ell \cdot (k+2) \cdot D \cdot \frac{2^{e_H}}{np},\end{aligned}$$

where the last inequality follows from (GS1) and (GS2) in Lemma 8.3 and from Lemma 3.5. Finally,

$$\begin{aligned}\frac{\Delta_2}{n^{2v'}p^{2e'}} &= \sum_{\substack{J\subseteq V(S_{i_0})\\ |J|\geqslant 2}} \sum_{(S,S')\in\mathcal{S}'(J)} \sum_{\substack{I\subseteq V(H)\setminus V(S_{i_0})\\ e(F_{I,J})>0}} \frac{n^{|J|}}{n^{v_{F_{I,J}}}p^{e_{F_{I,J}}}} \\ &\leqslant \sum_{\substack{u,v\in V_i\\ u\neq v}} \sum_{\substack{S,S'\in\mathcal{S}'\\ u,v\in V(S)\cap V(S')}} \sum_{\substack{J\subseteq V(S_{i_0})\\ |J|\geqslant 2}} \sum_{\substack{I\subseteq V(H)\setminus V(S_{i_0})\\ e(F_{I,J})>0}} \frac{n^{|J|}}{n^{v_{F_{I,J}}}p^{e_{F_{I,J}}}} \\ &\leqslant C_1\ell\cdot 2^{v_H}\cdot\max\left\{\frac{n^{|V(F)\cap V(S_{i_0})|}}{n^{v_F}p^{e_F}} : \emptyset\neq F\subseteq H\setminus S_{i_0}\right\},\end{aligned} \tag{36}$$

where the last inequality follows from (GS3) in Lemma 8.3 (since the stars in $\mathcal{S}\supseteq\mathcal{S}'$ are edge-disjoint, two different $S,S'\in\mathcal{S}'$ that intersect in more than one vertex have to intersect only in leaf vertices). In order to bound the maximum in the right-hand side of (36), given an arbitrary nonempty $F\subseteq H\setminus S_{i_0}$, we let $F'=F\cup S_{i_0}$, so that

$$\frac{n^{|V(F)\cap V(S_{i_0})|}}{n^{v_F}p^{e_F}} = n^{-v_{F'}+k+2}p^{-e_{F'}+k+1}. \tag{37}$$

Claim 8.4. For every F' satisfying $S_{i_0}\subsetneq F'\subseteq H$, we have

$$n^{-v_{F'}+k+2}p^{-e_{F'}+k+1}\leqslant\frac{2}{\tilde{C}\log n}.$$

Proof. Since $e_{F'}>e_{S_{i_0}}=k+1$, we have

$$\begin{aligned}n^{-v_{F'}+k+2}p^{-e_{F'}+k+1} &\overset{(33)}{\leqslant} n^{-v_{F'}+k+2}\cdot\left(\frac{\tilde{C}}{2}\cdot\frac{m_H}{n^2}\right)^{-e_{F'}+k+1} \\ &\leqslant\frac{2}{\tilde{C}}\cdot n^{-v_{F'}+k+2}\cdot\left(\frac{m_H}{n^2}\right)^{-e_{F'}+k+1} \\ &=\frac{2}{\tilde{C}}\cdot\left(n^{\frac{1}{d_{k+2}(F')}}\cdot\frac{m_H}{n^2}\right)^{-e_{F'}+k+1}.\end{aligned}$$

Regardless of which case holds true in the definition of m_H given in (1), we have

$$\frac{m_H}{n^2}\leqslant n^{-\frac{1}{\eta(H)}}(\log n)^{\frac{1}{\zeta(H)-k-1}} = n^{-\frac{1}{\eta_{i_0}(H)}}(\log n)^{\frac{1}{\zeta_{i_0}(H)-k-1}}$$

and, consequently,

$$n^{-v_{F'}+k+2}p^{-e_{F'}+k+1}\leqslant\frac{2}{\tilde{C}}\cdot n^{\left(\frac{1}{d_{k+2}(F')}-\frac{1}{\eta_{i_0}(H)}\right)(-e_{F'}+k+1)}\cdot(\log n)^{-\frac{e_{F'}-k-1}{\zeta_{i_0}(H)-k-1}}.$$

The claimed upper bound follows since $d_{k+2}(F')\leqslant\eta_{i_0}(H)$ and $e_{F'}\geqslant\zeta_{i_0}(H)$ whenever $d_{k+2}(F')=\eta_{i_0}(H)$. □

Substituting (37) into (36) and invoking Claim 8.4 yields

$$\Delta_2\leqslant\frac{C_1\ell\cdot 2^{v_H+1}}{\tilde{C}\log n}\cdot n^{2v'}p^{2e'}.$$

Recalling (34) and the definitions of v' and e', we thus obtain

$$\frac{\Delta}{\mu^2} \leqslant \frac{\Delta_0 + \Delta_1 + \Delta_2}{\mu^2} \leqslant \frac{1}{c^2 \ell} \cdot \left(\frac{2^{e_H} \ell}{n^2 p} + \frac{2^{e_H}(k+2)D}{np} + \frac{C_1 2^{v_H+1}}{\tilde{C} \log n} \right).$$

Since $\ell \leqslant Dn$, or otherwise $\mathcal{T}_\Pi(B, t, \ell, h)$ is empty (see Section 8.1), and

$$\frac{D}{np} \leqslant \frac{\beta m}{n^2 p \log n} \overset{(33)}{\leqslant} \frac{2\beta}{\log n},$$

we conclude that

$$\frac{\Delta}{\mu^2} \leqslant \frac{C'(\beta + \tilde{C}^{-1})}{\ell \log n} \tag{38}$$

where C' is some constant that depends only on H. On the other hand, (34) and Claim 8.4 with $F' = H$ imply that

$$\mu \geqslant \frac{c\tilde{C}\ell \log n}{2}. \tag{39}$$

Finally, we invoke Lemma 3.1 with $q = \frac{\mu}{\mu + \Delta} \leqslant 1$ to conclude that

$$\frac{|\mathcal{F}^*(T)|}{\binom{e(\Pi)}{m-t}} \leqslant \mathbb{P}(K \nsubseteq G' \text{ for every } K \in \mathcal{K}') \leqslant \exp\left(-\frac{\mu^2}{\mu + \Delta} + \frac{\mu^2 \Delta}{2(\mu+\Delta)^2} \right)$$
$$\leqslant \exp\left(-\frac{\mu^2}{2(\mu+\Delta)} \right) \leqslant \exp\left(-\min\left\{ \frac{\mu}{4}, \frac{\mu^2}{4\Delta} \right\} \right).$$

Substituting inequalities (38) and (39) into this bound, we obtain the assertion of the proposition with $\tilde{c} = \min\{1/(4C'), c/8\}$. □

Proof of Lemma 8.3. Suppose that $\Pi = \{V_1, \ldots, V_r\}$ and let $T \in \mathcal{T}_\Pi(B, t, \ell, h)$ for some $B \in \mathcal{B}(\Pi, k)$. Recall from Section 8.1 that U_T is a canonically chosen maximal subgraph of T that extends B and satisfies $\Delta(U_T) \leqslant D$.

Claim 8.5. There are $U' \subseteq U_T$ and an orientation $\vec{U}$ of a subgraph of U' that satisfy

(i) We have $B \subseteq U'$ and $e(U' \setminus B) \geqslant \ell/2$.
(ii) For every $(u, v) \in \vec{U}$, we have $\deg_{U'}(u) \leqslant \max\{\deg_{U'}(v), 4(k+1)\}$.
(iii) For every $v \in [\![n]\!]$, either $\deg^-_{\vec{U}}(v) = 0$ or $\deg^-_{\vec{U}}(v) \geqslant \max\{\deg_{U'}(v)/4, k+1\}$.
(iv) We have $e(\vec{U}) \geqslant \ell/(8k+8)$.

Proof. Let $Q = \{v : \deg_{U_T}(v) \geqslant 4(k+1)\}$. We split the proof into two cases, depending on how many edges of $U_T \setminus B$ have an endpoint in Q.

Case 1. Fewer than half the edges of $U_T \setminus B$ touch Q.

Let U' be the graph obtained from U_T by removing all edges of $U_T \setminus B$ that touch Q. As $\ell = e(U_T \setminus B)$, the graph U' satisfies (i); moreover, as $\Delta(B) \leqslant k$, then $\Delta(U') < 4(k+1)$. Let $W = \{w \in U' : \deg_{U'}(w) \geqslant k+1\}$, let W' be a largest U'-independent subset of W, and let

$$\vec{U} = \big\{(u, v) : \{u, v\} \in U' \text{ and } v \in W'\big\}.$$

Since $\Delta(U') < 4(k+1)$, property (ii) clearly holds. To see that (iii) holds, choose an arbitrary $v \in [\![n]\!]$ and note that $\deg^-_{\vec{U}}(v) = 0$ if $v \notin W'$; if $v \in W' \subseteq W$, then

$$\deg^-_{\vec{U}}(v) = \deg_{U'}(v) \geqslant k + 1 = \max\{\deg_{U'}(v)/4, k+1\}.$$

Finally, we argue that (iv) holds as well. Since B is a maximal subgraph of U_T with maximum degree at most k and $U' \supseteq B$, every edge of $U' \setminus B$ must have an endpoint with degree larger than k. Therefore, by (i),

$$\ell/2 \leqslant e(U' \setminus B) \leqslant \sum_{w \in W} \deg_{U'}(w).$$

As every vertex in $W \setminus W'$ has a U'-neighbour in W' (since W' is a maximal U'-independent subset of W), we further have

$$\sum_{w \in W \setminus W'} \deg_{U'}(w) \leqslant \sum_{v \in W'} \sum_{w \in N_{U'}(v)} \deg_{U'}(w) \leqslant \sum_{v \in W'} \deg_{U'}(v) \cdot \Delta(U')$$

Recalling that $\Delta(U') \leqslant 4k+3$, that $(u,w) \in \vec{U}$ for every $\{u,w\} \in U'$ such that $w \in W'$, and that W' is an independent set in U', we conclude that

$$\ell/2 \leqslant (4k+4) \sum_{w \in W'} \deg_{U'}(w) = (4k+4) e(\vec{U}).$$

Case 2. At least half the edges of $U_T \setminus B$ touch Q.

In this case we just take $U' = U_T$, so that (i) clearly holds. We first let $\vec{U'}$ be an arbitrary orientation of U' such that $\deg_{U'}(u) \leqslant \deg_{U'}(v)$ for all $(u,v) \in \vec{U'}$. We then obtain $\vec{U}$ from $\vec{U'}$ by removing all edges directed to a vertex v that satisfies $\deg^-_{\vec{U'}}(v) < \max\{k+1, \deg_{U'}(v)/4\}$. The construction of $\vec{U}$ guarantees that both (ii) and (iii) are satisfied. Since every edge of U' between Q and Q^c is directed (in $\vec{U'}$) towards its Q-endpoint, we have

$$\sum_{v \in Q} \deg^-_{\vec{U'}}(v) \geqslant \frac{1}{2} \sum_{v \in Q} \deg_{U'}(v).$$

Consequently,

$$\begin{aligned} e(\vec{U}) &\geqslant \sum_{v \in Q} \deg^-_{\vec{U}}(v) = \sum_{v \in Q} \deg^-_{\vec{U'}}(v) - \sum_{\substack{v \in Q \\ \deg^-_{\vec{U}}(v) = 0}} \deg^-_{\vec{U'}}(v) \\ &\geqslant \frac{1}{2} \sum_{v \in Q} \deg_{U'}(v) - \sum_{v \in Q} \max\{k+1, \deg_{U'}(v)/4\} = \frac{1}{4} \sum_{v \in Q} \deg_{U'}(v), \end{aligned}$$

since $\deg_{U'}(v) \geqslant 4(k+1)$ for every $v \in Q$. Finally, as at least half the edges of $U' \setminus B$ touch Q, we have $\sum_{v \in Q} \deg_{U'}(v) \geqslant \ell/2$ and we may conclude that $e(\vec{U}) \geqslant \ell/8$. $\square$

Let $U' \subseteq U_T$ and an orientation $\vec{U}$ of a subgraph of U' be as in Claim 8.5. For each vertex v, denote $\vec{d}_v = \deg^-_{\vec{U}}(v)$ and let $u_1^v, \ldots, u_{\vec{d}_v}^v$ be a uniformly chosen random ordering of the set of the in-neighbours of v in $\vec{U}$. Given $v \in [\![n]\!]$ and $A \subseteq [\![n]\!] \setminus \{v\}$, denote by $S_v(A)$ the $|A|$-star centred at v whose leaves are all elements of A. Define

$$\mathcal{S}' = \big\{ S_v(\{u_i^v, \ldots, u_{i+k}^v\}) : v \in [\![n]\!],\ i \in [\![\vec{d}_v - k]\!],\ \text{and } (k+1) | (i-1) \big\}.$$

In other words, $\mathcal{S}'$ is a (random) collection of $K_{1,k+1}$s in U' created by taking, for every vertex v with positive in-degree in $\vec{U}$, the $\lfloor \vec{d}_v/(k+1) \rfloor$ stars centred at v whose leaves are v's first $k+1$ in-neighbours (in the random ordering defined above), v's next $k+1$ in-neighbours, etc. By construction, the stars in $\mathcal{S}'$ are edge-disjoint and

$|\mathcal{S}'| \leqslant e(U' \setminus B) \leqslant \ell$, as each star must contain an edge of $U \setminus B$ (since $\Delta(B) \leqslant k$). On the other hand, since $\vec{d_v} \geqslant k+1$ for every v such that $\vec{d_v} > 0$,

$$|\mathcal{S}'| = \sum_v \left\lfloor \frac{\vec{d_v}}{k+1} \right\rfloor \geqslant \sum_v \frac{\vec{d_v}}{2(k+1)} = \frac{e(\vec{U})}{2(k+1)} \geqslant \frac{\ell}{16(k+1)^2},$$

by (iv) in Claim 8.5. Finally, since, for every $S \in \mathcal{S}'$, there is an index $i_S \in [\![r]\!]$ such that $S \subseteq U[V_{i_S}]$, by the pigeonhole principle, there must be an $i \in [\![r]\!]$ such that the set

$$\mathcal{S} = \{S \in \mathcal{S}' : S \subseteq T[V_i]\}$$

has size at least $|\mathcal{S}'|/r$. This family satisfies (GS1) with $c_1 = \big(16(k+1)^2 r\big)^{-1}$. To see that (GS2) holds as well, recall that the stars in $\mathcal{S}$ are edge-disjoint, contained in U', and $\Delta(U') \leqslant D$.

In the remainder of the proof we show that, with nonzero probability, our collection $\mathcal{S}$ satisfies also (GS3). To this end, recall that

$$A(u,v) = \big\{(S,S') \in (\mathcal{S}')^2 : u \text{ and } v \text{ are leaves of both } S \text{ and } S'\big\}.$$

Since $\mathcal{S} \subseteq \mathcal{S}'$, it will suffice to show that, with nonzero probability,

$$\sum_{\substack{u,v\in[\![n]\!]\\ u\neq v}} |A(u,v)| \leqslant C_1 \ell. \tag{40}$$

for some C_1 that depends only on r and k. For each pair of distinct $u, v \in [\![n]\!]$, define

$$\mathcal{S}_{u,v} = \{S \in \mathcal{S}' : u \text{ and } v \text{ are leaves of } S\},$$
$$\mathcal{D}_{u,v} = \{w \in [\![n]\!] : u, v \in N^-_{\vec{U}}(w) \text{ and } \vec{d_w} \geqslant k+1\}.$$

Since the stars in $\mathcal{S}'$ are edge-disjoint, for every $w \in \mathcal{D}_{u,v}$, there is at most one $S \in \mathcal{S}_{u,v}$ whose w is the centre. Moreover, if $\mathcal{S}_{u,v}$ contains such a star, then both u and v must fall into one of the $\lfloor \vec{d_w}/(k+1) \rfloor$ intervals of length $k+1$ in the random ordering $u^w_1, \dots, u^w_{\vec{d_w}}$ of $N^-_{\vec{U}}(w)$. In particular, for every $w \in \mathcal{D}_{u,v}$,

$$\mathbb{P}\big(\mathcal{S}_{u,v} \text{ contains a star centred at } w\big) \leqslant \frac{k}{\vec{d_w} - 1} \leqslant \frac{k+1}{\vec{d_w}}, \tag{41}$$

as $\vec{d_w} \geqslant k+1$. Moreover, if $w \in \mathcal{D}_{u,v}$, then $(u,w) \in \vec{U}$ and hence, by (ii) and (iii) in Claim 8.5,

$$\vec{d_w} \geqslant \frac{\max\{\deg_{U'}(w), 4(k+1)\}}{4} \geqslant \frac{\deg_{U'}(u)}{4} \geqslant \frac{|\mathcal{D}_{u,v}|}{4}. \tag{42}$$

We conclude that

$$\sum_{\substack{u,v\in[\![n]\!]\\ u\neq v}} \mathbb{E}\big[|\mathcal{S}_{u,v}|\big] = \mathbb{E}\big[|\mathcal{S}'|\big] + \sum_{u,v} \sum_{\substack{w_1,w_2\in\mathcal{D}_{u,v}\\ w_1\neq w_2}} \prod_{i=1}^{2} \mathbb{P}(S_{u,v} \text{ contains a star centred at } w_i).$$

$$\overset{(41)}{\leqslant} \ell + \sum_{u,v} \left(\sum_{w\in\mathcal{D}_{u,v}} \frac{k+1}{\vec{d_w}} \right)^2 \overset{(42)}{\leqslant} \ell + \sum_{u,v} \sum_{w\in\mathcal{D}_{u,v}} \frac{4(k+1)^2}{\vec{d_w}}$$

$$\leqslant \ell + \sum_{w:\vec{d_w}\geqslant k+1} \frac{4(k+1)^2}{\vec{d_w}} \cdot \binom{\vec{d_w}}{2} \leqslant \ell + 2(k+1)^2 \sum_{w:\vec{d_w}\geqslant k+1} \vec{d_w}.$$

Finally, since $\vec{d}_w \geqslant k+1$ implies that

$$\vec{d}_w \leqslant \deg_{U'}(w) \leqslant \deg_{U'\setminus B}(w) + k \leqslant (k+1)\deg_{U'\setminus B}(w),$$

we have

$$\begin{aligned}\sum_{\substack{u,v\in[\![n]\!]\\ u\neq v}} \mathbb{E}\big[|A(u,v)|\big] &= \sum_{u,v} \mathbb{E}\big[|\mathcal{S}_{u,v}|\big] \leqslant \ell + 2(k+1)^3 \sum_{w} \deg_{U'\setminus B}(w) \\ &= \ell + 2(k+1)^3 \cdot 2e(U'\setminus B) \leqslant \big(4(k+1)^3+1\big)\ell.\end{aligned}$$

In particular, taking $C_1 = 4(k+1)^3+1$, inequality (40) must hold with nonzero probability. □

8.6. **The high-degree case – introduction.** Recall from Section 8.1 that, for $T \in \mathcal{T}_\Pi$, we defined subgraphs B_T and U_T satisfying $B_T \subseteq U_T \subseteq T$ and we denoted by H_T the set of all vertices of T whose degree is larger than $\rho m/n$. Then, $\mathcal{T}_{\mathrm{H}}(\Pi)$ was the family of all $T \in \mathcal{T}_\Pi$ that satisfy (cf. (28))

$$e(U_T \setminus B_T)\log n < \frac{m|H_T|}{\xi n}, \tag{43}$$

Our argument in the high-degree case will analyse the distribution of edges incident to a subset of the set H_T of high-degree vertices that has convenient properties specified by our next lemma.

Lemma 8.6. *Suppose that $\Pi = \{V_1, \ldots, V_r\}$. For every $T \in \mathcal{T}_\Pi$, there exist $i \in [\![r]\!]$ and $Y \subseteq V_i$ with $|Y| \geqslant |H_T|/(2r)$ such that, for every $v \in Y$,*

$$\deg_{T\setminus U_T}(v, V_i \setminus Y) \geqslant \frac{\rho m}{3n}.$$

Proof. By the pigeonhole principle, there is an $i \in [\![r]\!]$ such that $|H_T \cap V_i| \geqslant |H_T|/r$. Fix any such i and let $V_i = V_i' \cup V_i''$ be an arbitrary partition that maximises the number of edges of $T \setminus B_T$ incident to $H_T \cap V_i$ that cross the partition. Then, for every $v \in V_i'$, we have $\deg_T(v, V_i'') \geqslant \deg_T(v, V_i')$ and vice-versa. We let Y be the larger of the two sets $H_T \cap V_i'$ and $H_T \cap V_i''$, so that $|Y| \geqslant |H_T \cap V_i|/2 \geqslant |H_T|/(2r)$. Without loss of generality, $Y = H_T \cap V_i'$. Writing $U = U_T$, we have, for every $v \in Y \subseteq H_T$,

$$\begin{aligned}\deg_{T\setminus U}(v, V_i \setminus Y) \geqslant \deg_{T\setminus U}(v, V_i'') &\geqslant \frac{\deg_{T\setminus U}(v, V_i)}{2} = \frac{\deg_T v - \deg_U v}{2} \\ &\geqslant \frac{\rho m}{2n} - \frac{D}{2} \geqslant \frac{\rho m}{2n} - \frac{\beta m}{2n\log n} \geqslant \frac{\rho m}{3n},\end{aligned}$$

as claimed. □

Fix some $\Pi = \{V_1, \ldots, V_r\}$ and $T \in \mathcal{T}_{\mathrm{H}}(\Pi)$. Let $i_T \in [\![r]\!]$ and $Y_T \subseteq V_{i_T}$ be the index and the set from the statement of Lemma 8.6. Let

$$D_H = \left\lceil \frac{\rho m}{3n} \right\rceil$$

and define $\mathcal{Z}(T)$ to be the family of all graphs that are obtained from T by adding to it edges connecting each $v \in Y$ to some D_H vertices in each V_i with $i \neq i_T$. Note that, for every $Z \in \mathcal{Z}(T)$,

$$e(Z) = e(T) + |Y_T| \cdot (r-1) \cdot D_H.$$

Recall from (16) that, for every $G \in \mathcal{F}^*(T)$ and every $v \in H_T$, we have $\deg_G(v, V_i) \geqslant \rho m/n \geqslant D_H$ for every $i \in [\![r]\!]$. This means, in particular, that for each $G \in \mathcal{F}^*(T)$, there is some $Z \in \mathcal{Z}(T)$ such that $Z \subseteq G$. In other words, defining, for each $Z \in \mathcal{Z}(T)$,

$$\mathcal{F}^*(T; Z) = \{G \in \mathcal{F}^*(T) : Z \subseteq G\},$$

we have

$$\mathcal{F}^*(T) = \bigcup_{Z \in \mathcal{Z}(T)} \mathcal{F}^*(T; Z). \tag{44}$$

We now turn to bounding $|\mathcal{F}^*(T; Z)|$ from above. To this end, fix some $T \in \mathcal{T}_{\mathrm{H}}(\Pi)$ and $Z \in \mathcal{Z}(T)$. For every $v \in Y_T$ and every $i \in [\![r]\!]$, let $N_i(v)$ be an arbitrary subset of $N_{Z \setminus U_T}(v) \cap (V_i \setminus Y)$ with D_H elements (and note that $N_i(v) = N_Z(v) \cap V_i$ when $i \neq i_T$).

Let v_c be the centre of any critical star of H and let H^- be the subgraph of H obtained by removing v_c and all the vertices whose only neighbour in H is v_c. (As H has no isolated vertices, neither does H^-.) Let $W_1 = N_H(v_c) \cap V(H^-)$ and $W_2 = V(H^-) \setminus W_1$; denote $v_1 = |W_1|$ and $v_2 = |W_2|$. Since H^- is obtained from H by removing the critical vertex v_c (and possibly some additional vertices), it is r-colourable; let us fix an arbitrary proper colouring $\psi \colon V(H^-) \to [\![r]\!]$.

Define a v_1-partite v_1-uniform hypergraph $\mathcal{H}_Z$ as follows:

$$V(\mathcal{H}_Z) = \bigsqcup_{w \in W_1} V_{\psi(w)},$$

$$E(\mathcal{H}_Z) = \bigcup_{v \in Y_T} \left\{ (v_w)_{w \in W_1} : v_w \in N_{\psi(w)}(v) \text{ for all } w \in W_1, \text{ all distinct} \right\}.$$

For every injection $\varphi \colon V(H^-) \to [\![n]\!]$, let K_φ be the labeled graph that is the image of H^- via the embedding φ. Define

$$\Phi_Z = \Big\{ \varphi : K_\varphi \subseteq \Pi - Y_T \text{ and } \big(\varphi(w)\big)_{w \in W_1} \in \mathcal{H}_Z \Big\};$$

in other words, Φ_Z comprises all embeddings of H^- into Π that avoid the set Y_T and such that W_1 is mapped into $N_1(v) \cup \dots \cup N_r(v)$ for some $v \in Y_T$, accordingly with the colouring ψ.

Choose an arbitrary $G \in \mathcal{F}^*(T; Z)$. We claim that $G \cap \Pi$ cannot contain any of the K_φ with $\varphi \in \Phi_Z$. Suppose to the contrary that $K_\varphi \subseteq G \cap \Pi$ for some $\varphi \in \Phi_Z$. By the definitions of $\mathcal{H}_Z$ and Φ_Z, there is a vertex $v \in Y_T$ such that $\varphi(w) \in N_{\psi(w)}(v)$ for all $w \in W_1$. Since $N_i(v) \subseteq N_Z(v) \subseteq N_G(v)$ for all $i \in [\![r]\!]$, extending φ to $V(H)$ by first letting $\varphi(v_c) = v$ and then choosing $\varphi(w) \in N_Z(v)$ arbitrarily[2] for all $w \in N_H(v_c) \setminus V(H^-)$ would give an embedding of H into G. In particular, letting G' be a uniformly chosen random subgraph of $\Pi \setminus Z$ with $m - e(Z)$ edges, we have

$$|\mathcal{F}^*(T; Z)| \leqslant \mathbb{P}\big(K_\varphi \nsubseteq G' \text{ for each } \varphi \in \Phi_Z\big) \cdot \binom{e(\Pi)}{m - e(Z)}. \tag{45}$$

The probability in (45) can vary greatly with the distribution of the edges of the associated hypergraph $\mathcal{H}_Z$. For a vast majority of $Z \in \mathcal{Z}(T)$, an upper bound on this probability that we will obtain using the Hypergeometric Janson Inequality will be sufficient to survive a naive union bound argument; we shall refer to this as the *regular case*. There will be, however, a family of exceptional graphs $Z \in \mathcal{Z}(T)$ for which the

[2]One can keep φ injective since $v_H \ll \rho m/n \leqslant \deg_Z(v)$.

distribution of the edges of the associated hypergraph $\mathcal{H}_Z$ precludes obtaining a strong upper bound on the probability in (45). We shall prove (using Lemma 3.3) that the number of such exceptional graphs Z is extremely small; we shall refer to this as the *irregular case*.

To make the above discussion precise, given a hypergraph $\mathcal{H}$ on $\bigsqcup_{w \in W_1} V_{\psi(w)}$, a set $I \subseteq W_1$, and an $L \in \prod_{w \in I} V_{\psi(w)}$, the degree $\deg_{\mathcal{H}}(L)$ of L in $\mathcal{H}$ is defined by

$$\deg_{\mathcal{H}}(L) = |\{K \in \mathcal{H} \colon L \subseteq K\}|,$$

where we write $L \subseteq K$ to mean that K agrees with L on the coordinates indexed by I, and the maximal I-degree of $\mathcal{H}$, denoted by $\Delta_I(\mathcal{H})$, is defined by

$$\Delta_I(\mathcal{H}) = \max\Big\{ \deg_{\mathcal{H}}(L) \colon L \in \prod_{w \in I} V_{\psi(w)} \Big\};$$

in particular $\Delta_\emptyset(\mathcal{H}) = e(\mathcal{H})$.

In order to describe the split between the regular and the irregular cases, we need to introduce several additional parameters. First, let Γ be a constant satisfying

$$\Gamma \geqslant \frac{21r}{\xi} \tag{46}$$

and let α be a positive constant that satisfies

$$\big(3er\alpha^{1/v_H} v_H\big)^\gamma \leqslant \exp(-12\Gamma). \tag{47}$$

Moreover, let

$$c_2 = \frac{1}{2} \cdot \left(\frac{\rho}{2v_H}\right)^{v_H} \tag{48}$$

and let σ and C_2 be positive constants satisfying

$$\max\left\{\sigma \cdot (2r)^{v_1}, \frac{(4r)^{v_1}}{C_2}\right\} \leqslant \min\big\{\tau_{3.3}(z, \alpha, \lambda \leftarrow 2^{-v_1}) : z \in [\![v_H]\!]\big\} \tag{49}$$

Let $\mathcal{Z}_1^R(T)$ be the family of all $Z \in \mathcal{Z}(T)$ such that

$$e(\mathcal{H}_Z) \geqslant \sigma n^{v_1}. \tag{50}$$

Let $\mathcal{Z}_2^R(T)$ be the family of all $Z \in \mathcal{Z}(T) \setminus \mathcal{Z}_1^R(T)$ such that $\mathcal{H}_Z$ contains a subhypergraph $\mathcal{H} \subseteq \mathcal{H}_Z$ which satisfies

$$e(\mathcal{H}) \geqslant c_2 \cdot |Y_T| \cdot \left(\frac{m}{n}\right)^{v_1} \tag{51}$$

and, for every nonempty $I \subseteq W_1$,

$$\Delta_I(\mathcal{H}) \leqslant \max\left\{\left(\frac{m}{n}\right)^{v_1 - |I|}, C_2 \cdot \frac{e(\mathcal{H})}{n^{|I|}}\right\}. \tag{52}$$

Finally, let $\mathcal{Z}^R(T) = \mathcal{Z}_1^R(T) \cup \mathcal{Z}_2^R(T)$ and $\mathcal{Z}^I(T) = \mathcal{Z}(T) \setminus \mathcal{Z}^R(T)$. Since $\mathcal{Z}^R(T)$ and $\mathcal{Z}^I(T)$ form a partition of $\mathcal{Z}(T)$ for every $T \in \mathcal{T}_{\mathrm{H}}(\Pi)$, it follows from (44) that

$$\sum_{\substack{T \in \mathcal{T}_{\mathrm{H}}(\Pi) \\ B_T = B}} |\mathcal{F}^*(T)| \leqslant \sum_{\substack{T \in \mathcal{T}_{\mathrm{H}}(\Pi) \\ B_T = B}} \sum_{Z \in \mathcal{Z}^R(T)} |\mathcal{F}^*(T; Z)| + \sum_{\substack{T \in \mathcal{T}_{\mathrm{H}}(\Pi) \\ B_T = B}} \sum_{Z \in \mathcal{Z}^I(T)} |\mathcal{F}^*(T; Z)|. \tag{53}$$

The regular and the irregular cases are estimates of the first and the second sum in the right-hand side of (53), respectively.

8.7. **The regular case – summary.** In the regular case, we will rely on the following upper bound on the cardinality of $\mathcal{F}^*(T;Z)$, which is established in Section 8.9 with the use of the Hypergeometric Janson Inequality (Lemma 3.1).

Lemma 8.7. *There exists a positive constant $\tilde{c}$ that depends only on H such that the following holds for every $T \in \mathcal{T}_{\mathrm{H}}(\Pi)$ and each $Z \in \mathcal{Z}^R(T)$. If n is sufficiently large and $m \geqslant \tilde{C} n^{2-\frac{1}{m_2(H)}}$ for some $\tilde{C} \geqslant 2$, then*

$$|\mathcal{F}^*(T;Z)| \leqslant \exp\left(-\tilde{c} \cdot \min\left\{\frac{c_2 \cdot \tilde{C} \cdot |Y_T|}{n}, \frac{1}{C_2}, \sigma\right\} \cdot m\right) \cdot \binom{e(\Pi)}{m - e(Z)}.$$

This upper bound on $|\mathcal{F}^*(T;Z)|$ provided by Lemma 8.7 will be combined with the following estimate on the size of the sum over all $Z \in \mathcal{Z}(T)$.

Lemma 8.8. *For every $T \in \mathcal{T}_{\mathrm{H}}(\Pi)$,*

$$|\mathcal{Z}(T)| \cdot \binom{e(\Pi)}{m - e(T) - |Y_T| \cdot (r-1) \cdot D_H} \leqslant \exp\left(\frac{|Y_T| \cdot m}{\xi n}\right) \cdot \binom{e(\Pi)}{m - e(T)}.$$

Proof. Since, for every $Z \in \mathcal{Z}(T)$, the graph $Z \setminus T$ comprises precisely $|Y_T| \cdot (r-1) \cdot D_H$ edges incident to Y_T, we have, letting $b = |Y_T|$,

$$|\mathcal{Z}(T)| \leqslant \binom{n}{(r-1)D_H}^b \overset{(5)}{\leqslant} \left(\frac{en}{(r-1)D_H}\right)^{b(r-1)D_H}.$$

On the other hand, by (30), which holds for all $y \leqslant m' \leqslant m \leqslant e(\Pi) - \xi n^2$, we have

$$\frac{\binom{e(\Pi)}{m-e(T)-b\cdot(r-1)\cdot D_H}}{\binom{e(\Pi)}{m-e(T)}} \leqslant \left(\frac{m}{\xi n^2}\right)^{b(r-1)D_H}.$$

The claimed bound follows after noting that

$$\left(\frac{en}{(r-1)D_H} \cdot \frac{m}{\xi n^2}\right)^{(r-1)D_H} \leqslant \exp\left(\frac{m}{\xi n}\right),$$

as $(ea/x)^x \leqslant e^a$ for all $x \in (0,\infty)$. □

Before we close this section, we show how these two lemmas can be used to estimate the first sum in the right-hand side of (53):

$$\Sigma_B^R = \sum_{\substack{T \in \mathcal{T}_{\mathrm{H}}(\Pi) \\ B_T = B}} \underbrace{\sum_{Z \in \mathcal{Z}^R(T)} |\mathcal{F}^*(T;Z)|}_{\Sigma_T^R}.$$

Since

$$m \geqslant C_H m_H \geqslant C_H n^{2-\frac{1}{m_2(H)}} \overset{(24)}{\geqslant} \frac{1}{c_2 \cdot \tilde{c}_{8.7}} \cdot \frac{35r}{\xi} \cdot n^{2-\frac{1}{m_2(H)}},$$

Lemma 8.7 implies that, for every $T \in \mathcal{T}_{\mathrm{H}}(\Pi)$,

$$\Sigma_T^R \leqslant \sum_{Z \in \mathcal{Z}^R(T)} \exp\left(-\min\left\{\frac{|Y_T|}{n} \cdot \frac{35r}{\xi}, \frac{\tilde{c}_{8.7}}{C_2}, \tilde{c}_{8.7}\sigma\right\} \cdot m\right) \cdot \binom{e(\Pi)}{m - e(Z)}.$$

Since $|Y_T| \leqslant |H_T| \leqslant 2\delta n/\rho$ for every $T \subseteq \Pi^c$ with at most δm edges, we have, for every $T \in \mathcal{T}_{\mathrm{H}}(\Pi)$,

$$\frac{|Y_T|}{n} \leqslant \frac{2\delta}{\rho} \overset{(23)}{\leqslant} \frac{\xi}{35r} \cdot \min\left\{\frac{\tilde{c}_{8.7}}{C_2}, \tilde{c}_{8.7}\sigma\right\}$$

and, consequently,

$$\Sigma_T^R \leqslant \sum_{Z \in \mathcal{Z}^R(T)} \exp\left(-\frac{35rm \cdot |Y_T|}{\xi n}\right) \cdot \binom{e(\Pi)}{m - e(Z)}.$$

Since $e(Z) = e(T) + |Y_T| \cdot (r-1) \cdot D_H$ for every $Z \in \mathcal{Z}(T)$, Lemma 8.8 gives

$$\Sigma_T^R \leqslant \exp\left(-\frac{34rm \cdot |Y_T|}{\xi n}\right) \cdot \binom{e(\Pi)}{m - e(T)} \leqslant \exp\left(-\frac{17m \cdot |H_T|}{\xi n}\right) \cdot \binom{e(\Pi)}{m - e(T)},$$

where the second inequality follows from the inequality $|Y_Y| \geqslant |H_T|/(2r)$, see Lemma 8.6.

Let $\mathcal{L}$ be the family of all triples (t, ℓ, h) that satisfy $t \geqslant \ell \geqslant 1$ and $\ell \log n < mh/(\xi n)$, cf. (43), and observe that

$$\Sigma_B^R \leqslant \sum_{(t,\ell,h) \in \mathcal{L}} \sum_{T \in \mathcal{T}_\Pi(B,t,\ell,h)} \Sigma_T^R \leqslant \sum_{(t,\ell,h) \in \mathcal{L}} |\mathcal{T}_\Pi(B,t,\ell,h)| \cdot \exp\left(-\frac{17mh}{\xi n}\right) \cdot \binom{e(\Pi)}{m-t}.$$

Since, by Lemma 8.8, we have, for every $(t, \ell, h) \in \mathcal{L}$,

$$\begin{aligned} |\mathcal{T}_\Pi(B,t,\ell,h)| \cdot \binom{e(\Pi)}{m-t} &\leqslant \exp\left(14\ell \log n + \frac{2mh}{\xi n}\right) \cdot \binom{e(\Pi)}{m - e(B)} \\ &\leqslant \exp\left(\frac{16mh}{\xi n}\right) \cdot \binom{e(\Pi)}{m - e(B)}, \end{aligned}$$

we may conclude that

$$\Sigma_B^R \cdot \binom{e(\Pi)}{m - e(B)}^{-1} \leqslant \sum_{(t,\ell,h) \in \mathcal{L}} \exp\left(-\frac{mh}{\xi n}\right) \leqslant |\mathcal{L}| \cdot \exp\left(-\frac{m}{\xi n}\right) \leqslant \frac{1}{n},$$

as $h \geqslant 1$ for every $(t, \ell, h) \in \mathcal{L}$.

8.8. **The irregular case – summary.** In the irregular case, we will use Lemma 3.3 to prove upper bounds on the number of graphs Z that fall into $\mathcal{Z}^I(T)$ for some $T \in \mathcal{T}_{\mathrm{H}}(\Pi)$; these upper bounds will be so strong that we will be able to get the desired estimate on the second term in the right-hand side of (53) by combining them with the trivial estimate $\binom{e(\Pi)}{m-e(Z)}$ on the number of completions of Z to a graph in $\mathcal{F}^*(T; Z)$. Since the nature of our argument precludes obtaining a strong bound on $|\mathcal{F}^*(T; Z) \cap \mathcal{Z}^I(T)|$ for every T, we will have to partition the family $\bigcup_{T \in \mathcal{T}_{\mathrm{H}}(\Pi)} \mathcal{Z}^I(T)$ differently. To this end, for every positive integer b, define

$$\mathcal{Z}_\Pi^I(b) = \bigcup_{\substack{T \in \mathcal{T}_{\mathrm{H}}(\Pi) \\ |Y_T| = b}} \mathcal{Z}^I(T)$$

Given some $T \in \mathcal{T}_{\mathrm{H}}(\Pi)$ and a $Z \in \mathcal{Z}(T)$, let $T'_Z \subseteq T$ be the graph obtained from T by removing the $|Y_T| \cdot D_H$ edges vu such that $v \in Y_T$ and $u \in N_{i_T}(v)$. Note that $B_T \subseteq U_T \subseteq T'_Z$, as $N_{i_T}(v)$ was defined to be a subset of $N_{Z \setminus U_T}(v)$, and that T'_Z can be defined in terms of Z only because if $Z \in \mathcal{Z}(T)$, then $T = Z \cap \Pi^c$. Further, for every positive integer b and every $T' \subseteq \Pi^c$, let

$$\mathcal{Z}_\Pi^I(b; T') = \{Z \in \mathcal{Z}_\Pi^I(b) : T'_Z = T'\}.$$

The following upper bound on cardinalities of the families $\mathcal{Z}_\Pi^I(b; T')$ is the main step in the analysis of the irregular case.

Lemma 8.9. *For every $T' \subseteq \Pi^c$ and every $b \geqslant 1$,*

$$|\mathcal{Z}^I_\Pi(b;T')| \cdot \binom{e(\Pi)}{m - e(T') - rbD_H} \leqslant \exp\left(-\frac{\Gamma bm}{n}\right) \cdot \binom{e(\Pi)}{m - e(T')}.$$

This upper bound on $|\mathcal{Z}^I_\Pi(b;T')|$ provided by Lemma 8.9 will be combined with the following estimate on the size of the sum over all T'. For every $B \in \mathcal{B}(\Pi, k)$ and every nonnegative integer t', let $\mathcal{T}'_\Pi(B, t', b)$ comprise all graphs $T' \subseteq \Pi^c$ with t' edges such that $T' = T'_Z$ for some $Z \in \mathcal{Z}(T)$, where $T \in \mathcal{T}_\mathrm{H}(\Pi)$ satisfies $B_T = B$ and $|Y_T| = b$.

Lemma 8.10. *Suppose that $n \log n \ll m \leqslant e(\Pi) - \xi n^2$. For every $B \in \mathcal{B}(\Pi, k)$ and all t' and b,*

$$|\mathcal{T}'_\Pi(B, t', b)| \cdot \binom{e(\Pi)}{m - t'} \leqslant \exp\left(\frac{20rmb}{\xi n}\right) \cdot \binom{e(\Pi)}{m - e(B)}.$$

Proof. We adapt the argument used in the proof of Lemma 8.2. Suppose that $T' \in \mathcal{T}'_\Pi(B, t', b)$. This means that there is a $T \in \mathcal{T}_\mathrm{H}(\Pi)$ such that $|Y_T| = b$, $B = B_T \subseteq U_T \subseteq T' \subseteq T$, and $T \setminus T'$ comprises some bD_H edges incident to $Y_T \subseteq H_T$. Moreover, since $T \in \mathcal{T}_\mathrm{H}(\Pi)$, we have

$$e(U_T \setminus B_T) \overset{(43)}{<} \frac{m|H_T|}{\xi n \log n}.$$

Let U'_T be the subgraph of $U_T \setminus B_T$ obtained by removing all edges touching the set X_T of vertices whose degree in U_T is D. Since every edge of $U_T \setminus U'_T$ has at least one endpoint in X_T and $\Delta(B_T) \leqslant k$, we have

$$e(U_T \setminus B) \geqslant e(U'_T) + |X_T| \cdot (D - k)/2 \geqslant e(U'_T) + |X_T| \cdot D/3.$$

We choose the $t' - e(B)$ edges of $T' \setminus B$ in three steps:

(S1) We choose the edges of U'_T.
(S2) We choose the edges of $T' \setminus B$ that touch $X_T \setminus H_T$.
(S3) We choose the remaining edges of $T' \setminus B$; they all touch H_T.

We count the number of ways to build a graph $T' \in \mathcal{T}'_\Pi(B, t', b)$ with u', t'_X, and t'_H edges chosen in steps (S1), (S2), and (S3), respectively. An upper bound on $|\mathcal{T}'_\Pi(B, t', b)|$ will be obtained by summing over all choices for u', t'_X, and t'_H. There are at most $\binom{e(\Pi^c)}{u'}$ ways to choose u' edges of U'_T and, as in the proof of Lemma 8.2,

$$\binom{e(\Pi^c)}{u'} \cdot \frac{\binom{e(\Pi)}{m - e(B) - u'}}{\binom{e(\Pi)}{m - e(B)}} \leqslant m^{2u'}.$$

Next, let $N_2 = N_2(t_X, s)$ denote the total number of ways to choose the t'_X edges touching $X_T \setminus H_T$ when $|X_T \setminus H_T| = s$. As in the proof of Lemma 8.2, we have

$$N_2 \cdot \frac{\binom{e(\Pi)}{m - e(B) - u' - t'_X}}{\binom{e(\Pi)}{m - e(B) - u'}} \leqslant m^{Ds}$$

Finally, let $N_3 = N_3(t_X, h)$ denote the number of ways to choose the remaining t'_H edges of $T' \setminus B$ when $|H_T| = h$. Recalling that $e(B) + u' + t'_X + t'_H = t'$ and arguing as in the proof of Lemma 8.2, we obtain

$$N_3 \cdot \frac{\binom{e(\Pi)}{m - t'}}{\binom{e(\Pi)}{m - e(B) - u' - t'_X}} \leqslant \exp\left(\frac{2mh}{\xi n}\right).$$

Since $|H_T| \leqslant 2r|Y_T| = 2rb$, by Lemma 8.6, combining the above bounds, we obtain

$$|\mathcal{T}'_\Pi(B,t',b)| \cdot \frac{\binom{e(\Pi)}{m-t'}}{\binom{e(\Pi)}{m-e(B)}} \leqslant \sum_{\substack{u',t'_X,t'_H,s,h \\ u'+t'_X+t'_H=t' \\ u'+sD/3\leqslant mh/(\xi n\log n) \\ h\leqslant 2rb}} m^{2u'} \cdot m^{Ds} \cdot \exp\left(\frac{2mh}{\xi n}\right)$$

$$\leqslant nm^3 \sum_{h\leqslant 2rb} \exp\left(\left(\frac{3\log m}{\log n}+2\right)\cdot\frac{mh}{\xi n}\right) \leqslant \exp\left(\frac{20rmb}{\xi n}\right),$$

as claimed. □

Before we close this section, we show how these two lemmas can be used to estimate the second sum in the right-hand side of (53):

$$\Sigma_B^I = \sum_{\substack{T\in\mathcal{T}_{\mathrm{H}}(\Pi)\\ B_T=B}} \sum_{Z\in\mathcal{Z}^I(T)} |\mathcal{F}^*(T;Z)| \leqslant \sum_{\substack{T\in\mathcal{T}_{\mathrm{H}}(\Pi)\\ B_T=B}} \sum_{Z\in\mathcal{Z}^I(T)} \binom{e(\Pi)}{m-e(Z)}$$

$$= \sum_{t',b} \sum_{T'\in\mathcal{T}'_\Pi(B,t',b)} \underbrace{\sum_{Z\in\mathcal{Z}^I_\Pi(b;T')} \binom{e(\Pi)}{m-e(Z)}}_{\Sigma^I_{T'}}.$$

Since $e(Z) = e(T') + rbD_H$ for every $Z \in \mathcal{Z}^I_\Pi(b;T')$, Lemma 8.9 implies that

$$\Sigma^I_{T'} \leqslant \exp\left(-\frac{\Gamma bm}{n}\right)\cdot\binom{e(\Pi)}{m-e(T')}$$

and, further, Lemma 8.10 implies that

$$\Sigma^I_B \cdot \binom{e(\Pi)}{m-e(B)}^{-1} \leqslant \sum_{t',b} \exp\left(\frac{20rmb}{\xi n} - \frac{\Gamma bm}{n}\right) \overset{(46)}{\leqslant} \sum_{t',b} \exp\left(-\frac{rmb}{\xi n}\right)$$

$$\leqslant mn\cdot\exp\left(-\frac{rm}{\xi n}\right) \leqslant \frac{1}{n}.$$

8.9. **The regular case.** In this section, we prove Lemma 8.7, that is, for given $T \in \mathcal{T}_{\mathrm{H}}(\Pi)$ and $Z \in \mathcal{Z}^R_1(T) \cup \mathcal{Z}^R_2(T)$, we give an upper bound on the number of graphs in $\mathcal{F}^*(T;Z)$.

Proof of Lemma 8.7. Suppose that $\Pi = \{V_1,\ldots,V_r\}$, let $T \in \mathcal{T}_{\mathrm{H}}(\Pi)$, and fix an arbitrary $Z \in \mathcal{Z}^R_1(T)\cup\mathcal{Z}^R_2(T)$. If $Z \in \mathcal{Z}^R_1(T)$, we let $\mathcal{H} = \mathcal{H}_Z$ and recall that $e(\mathcal{H}) \geqslant \sigma n^{v_1}$, see (50). Otherwise, $Z \in \mathcal{Z}^R_2(T)$ and we let $\mathcal{H} \subseteq \mathcal{H}_Z$ be any hypergraph which satisfies both (51) and (52).

Recall the definitions of H^-, $\mathcal{H}_Z$, ψ, and Φ_Z from Section 8.6. For every $j \in [\![r]\!]$, randomly choose an equipartition $\{V_{j,w}\}_{w\in V(H)}$ of $V_j \setminus Y_T$ into v_H parts. We let Φ'_Z be the family of all embeddings $\varphi \in \Phi_Z$ that satisfy

$$\big(\varphi(w)\big)_{w\in W_1} \in \mathcal{H} \quad\text{and}\quad \varphi(w) \in V_{\psi(w),w} \text{ for every } w \in V(H^-).$$

Let $n' = \min\{|V| : V \in \mathcal{P}_{n,r}\} \geqslant n/(2r)$. Since there are at least $e(\mathcal{H})\cdot(n'-|Y_T|-v_H)^{v_2}$ embeddings $\varphi \in \Phi_Z$ such that $\big(\varphi(w)\big)_{w\in W_1} \in \mathcal{H}$ and, for each such φ, the probability that $\varphi \in \Phi'_Z$ is at least $v_H^{-v_H}$, there is a positive constant c that depends only on H such that

$$\mathbb{E}\big[|\Phi'_Z|\big] \geqslant c \cdot e(\mathcal{H}) \cdot n^{v_2}.$$

Now, fix some partitions $\{V_{j,w}\}_{w \in V(H)}$ for which $|\Phi'_Z|$ is at least as large as its expectation and let

$$\mathcal{K}' = \{K_\varphi : \varphi \in \Phi'_Z\}.$$

We claim that $K_\varphi \neq K_{\varphi'}$ for each pair of distinct $\varphi, \varphi' \in \Phi'_Z$. Since H^- has no isolated vertices and each $\varphi \in \Phi'_Z$ maps every $w \in V(H^-)$ to its dedicated set $V_{\psi(w),w}$, one can recover φ from the graph K_φ. This means, in particular, that

$$|\mathcal{K}'| = |\Phi'_Z| \geqslant c \cdot e(\mathcal{H}) \cdot n^{v_2}. \tag{54}$$

Suppose that $m \geqslant \tilde{C} n^{2-\frac{1}{m_2(H)}}$ for some $\tilde{C} \geqslant 2$ and let G' be a uniformly chosen subgraph of $\Pi \setminus Z$ with $m - e(Z)$ edges. The definition of $\mathcal{K}'$ and (45) imply that

$$|\mathcal{F}^*(T;Z)| \leqslant \mathbb{P}(K \nsubseteq G' \text{ for every } K \in \mathcal{K}') \cdot \binom{e(\Pi)}{m - e(Z)}.$$

We shall bound this probability from above using the Hypergeometric Janson Inequality. To this end, let $p = \frac{m - e(Z)}{e(\Pi) - e(Z)}$. Since

$$e(Z) \leqslant e(T) + (r-1) \cdot |H_T| \cdot D_H \leqslant 2r\delta m,$$

as $|H_T| \leqslant 2\delta n/\rho$ and $D_H \leqslant \rho m/n$, we have

$$p \geqslant \frac{m - e(Z)}{n^2} \geqslant (1 - 2r\delta) \cdot \frac{m}{n^2} \geqslant \frac{m}{2n^2} \geqslant \frac{\tilde{C}}{2} \cdot n^{2-\frac{1}{m_2(H)}},$$

as $\delta \leqslant \frac{1}{4r}$, see (23), and

$$p \leqslant \frac{m}{e(\Pi) - e(Z)} \leqslant \frac{m}{n^2/5 - 2r\delta n^2} \leqslant \frac{10m}{n^2},$$

where the second inequality follows from part (i) of Proposition 4.1, as $\Pi \in \mathcal{P}_{n,r}(\gamma)$ and $\gamma \leqslant \frac{1}{20r}$, see (22), and the final inequality holds because $\delta \leqslant \frac{1}{20r}$, see (23). For any $K, K' \in \mathcal{K}'$, we write $K \sim K'$ if K and K' share an edge but $K \neq K'$. Let μ and Δ be the quantities defined in the statement of the Hypergeometric Janson Inequality (Lemma 3.1), that is

$$\mu = \sum_{K \in \mathcal{K}'} p^{e_K} = |\mathcal{K}'| \cdot p^{e_{H^-}} \qquad \text{and} \qquad \Delta = \sum_{\substack{K, K' \in \mathcal{K}' \\ K \sim K'}} p^{e_{K \cup K'}}.$$

Claim 8.11. There is a positive constant c' that depends only on H such that

$$\mu \geqslant c' \cdot \min\left\{\frac{c_2 \cdot \tilde{C} \cdot |Y_T|}{n}, \sigma\right\} \cdot m. \tag{55}$$

Proof. It follows from (54) that

$$\mu = |\mathcal{K}'| \cdot p^{e_{H^-}} \geqslant c \cdot e(\mathcal{H}) \cdot n^{v_2} \cdot p^{e_{H^-}}.$$

Assume first that $Z \in \mathcal{Z}_1^R(T)$. Since $e(\mathcal{H}) \geqslant \sigma n^{v_1}$, we have

$$\mu \geqslant c \cdot \sigma \cdot n^{v_1 + v_2} \cdot p^{e_{H^-}}. \tag{56}$$

Since $v_1 + v_2$ is the number of vertices of H^- and $H^- \subseteq H$, Lemma 3.5 implies that $\mu \geqslant c \cdot \sigma \cdot m$, as $\tilde{C} \geqslant 2$. If, on the other hand, $Z \in \mathcal{Z}_2^R(T)$, then

$$\mu \geqslant c \cdot c_2 \cdot |Y_T| \cdot \left(\frac{m}{n}\right)^{v_1} \cdot n^{v_2} \cdot p^{e_{H^-}} \geqslant c \cdot c_2 \cdot |Y_T| \cdot \left(\frac{pn}{10}\right)^{v_1} \cdot n^{v_2} \cdot p^{e_{H^-}}$$
$$\geqslant c'' \cdot c_2 \cdot |Y_T| \cdot \frac{n^{v_1+v_2+1} p^{e_{H^-}+v_1}}{n}$$

for some c'' that depends only on H. Let H^* be the subgraph of H induced by $\{v_c\} \cup V(H^-)$ and note that $v_{H^*} = v_1 + v_2 + 1$ and $e_{H^*} = e_{H^-} + v_1$. Since v_c is the centre of a critical star of H, it has at least $\chi(H) \geqslant 3$ neighbours and thus $e_{H^*} \geqslant v_1 \geqslant 3$. By Lemma 3.5, with $F = H^*$,

$$n^{v_1+v_2+1} p^{e_{H^-}+v_1} = n^{v_{H^*}} p^{e_{H^*}} \geqslant \frac{\tilde{C}}{2} \cdot n^2 p \geqslant \frac{\tilde{C} m}{4},$$

and we may conclude that $\mu \geqslant c' \cdot c_2 \cdot \tilde{C} \cdot |Y_T| \cdot m/n$. This completes the proof of (55). □

Claim 8.12. There exists a positive constant c' that depends only on H such that

$$\frac{\mu^2}{\Delta} \geqslant c' \cdot \min\left\{\frac{c_2 \cdot \tilde{C} \cdot |Y_T|}{n}, \sigma, \frac{1}{C_2}\right\} \cdot m.$$

Proof. For every $I \subseteq W_1$ and $J \subseteq W_2$ let $H_{I,J}$ be the subgraph of H^- (and thus also of H) induced by $I \cup J$; note that $H_{I,J}$ may have isolated vertices. Further, let $\mathcal{K}(I,J)$ be the set of all pairs $K, K' \in \mathcal{K}'$ that agree exactly on (the image of) $I \cup J$, that is,

$$\mathcal{K}(I,J) = \left\{(K_\varphi, K_{\varphi'}) \in (\mathcal{K}')^2 : K_\varphi \cap K_{\varphi'} = \varphi(H_{I,J}) = \varphi'(H_{I,J})\right\}.$$

These definitions were made in such a way that

$$\begin{aligned}\Delta &= \sum_{K \in \mathcal{K}'} \sum_{\substack{I \subseteq W_1, J \subseteq W_2 \\ \emptyset \neq H_{I,J} \subsetneq H^-}} \sum_{\substack{K' \in \mathcal{K}' \\ (K,K') \in \mathcal{K}'(I,J)}} p^{2e_{H^-} - e(H_{I,J})} \\ &\leqslant \sum_{K \in \mathcal{K}'} p^{e_{H^-}} \sum_{\substack{I \subseteq W_1, J \subseteq W_2 \\ \emptyset \neq H_{I,J} \subsetneq H^-}} |\{K' \in \mathcal{K}' : (K,K') \in \mathcal{K}'(I,J)\}| \cdot p^{e_{H^-} - e_{H_{I,J}}} \\ &\leqslant \mu \sum_{\substack{I \subseteq W_1, J \subseteq W_2 \\ \emptyset \neq H_{I,J} \subsetneq H^-}} \Delta_I(\mathcal{H}) \cdot n^{v_2 - |J|} \cdot p^{e_{H^-} - e_{H_{I,J}}}.\end{aligned} \tag{57}$$

Assume first that $Z \in \mathcal{Z}_1^R(T)$. Using the trivial bound $\Delta_I(\mathcal{H}) \leqslant n^{v_1 - |I|}$, which is valid for all $I \subseteq W_1$, and (57), we obtain

$$\begin{aligned}\frac{\Delta}{\mu} &\leqslant \sum_{\substack{I \subseteq W_1, J \subseteq W_2 \\ \emptyset \neq H_{I,J} \subsetneq H^-}} n^{v_1+v_2-|I|-|J|} \cdot p^{e_{H^-} - e_{H_{I,J}}} = \sum_{\substack{I \subseteq W_1, J \subseteq W_2 \\ \emptyset \neq H_{I,J} \subsetneq H^-}} \frac{n^{v_1+v_2} p^{e_{H^-}}}{n^{v_{H_{I,J}}} p^{e_{H_{I,J}}}} \\ &\leqslant 2^{v_1+v_2} \cdot \frac{n^{v_1+v_2} p^{e_{H^-}}}{\min_{\emptyset \neq F \subseteq H^-} n^{v_F} p^{e_F}} \overset{\text{L. 3.5}}{\leqslant} 2^{v_1+v_2} \cdot \frac{n^{v_1+v_2} p^{e_{H^-}}}{n^2 p}.\end{aligned}$$

Since $n^{v_1+v_2} p^{e_{H^-}} \leqslant \mu/(c \cdot \sigma)$, see (56) and $n^2 p \geqslant m/2$, we may conclude that

$$\frac{\mu^2}{\Delta} \geqslant \frac{c \cdot \sigma}{2^{v_1+v_2+1}} \cdot m.$$

Suppose now that $Z \in \mathcal{Z}_2^R(T)$. In this case, for all nonempty $I \subseteq W_1$,

$$\Delta_I(\mathcal{H}) \leqslant \max\left\{\left(\frac{m}{n}\right)^{v_1-|I|}, C_2 \cdot \frac{e(\mathcal{H})}{n^{|I|}}\right\} \overset{(51)}{\leqslant} \max\left\{\frac{1}{c_2|Y_T|} \cdot \left(\frac{m}{n^2}\right)^{-|I|}, C_2\right\} \cdot \frac{e(\mathcal{H})}{n^{|I|}}.$$
$$\leqslant \max\left\{\frac{1}{c_2|Y_T|} \cdot \left(\frac{10}{p}\right)^{|I|}, C_2\right\} \cdot \frac{e(\mathcal{H})}{n^{|I|}}.$$

Denote by Δ_0 and Δ_1 the contributions to the sum in the right-hand side of (57) corresponding to $I = \emptyset$ and $I \neq \emptyset$, respectively, so that $\Delta \leqslant \Delta_0 + \Delta_1$. Since $H_{\emptyset,J} = H[J] \subseteq H$ and $\Delta_\emptyset(\mathcal{H}) = e(\mathcal{H})$, we have

$$\frac{\Delta_0}{\mu} \leqslant e(\mathcal{H}) \cdot n^{v_2} p^{e_{H^-}} \cdot \sum_{\substack{J \subseteq W_2 \\ H[J] \neq \emptyset}} \frac{1}{n^{|J|} p^{e(H[J])}} \leqslant 2^{v_2} \cdot \frac{e(\mathcal{H}) \cdot n^{v_2} p^{e_{H^-}}}{\min_{\emptyset \neq F \subseteq H} n^{v_F} p^{e_F}}.$$

Recalling that $e(\mathcal{H}) \cdot n^{v_2} p^{e_{H^-}} \leqslant \mu / c$, we conclude, using Lemma 3.5, that

$$\frac{\Delta_0}{\mu} \leqslant \frac{2^{v_2}}{c} \cdot \frac{\mu}{n^2 p} \leqslant \frac{2^{v_2+1} \mu}{cm}.$$

On the other hand,

$$\frac{\Delta_1}{\mu} \leqslant e(\mathcal{H}) \cdot n^{v_2} p^{e_{H^-}} \cdot \sum_{\substack{\emptyset \neq I \subseteq W_1, J \subseteq W_2 \\ \emptyset \neq H_{I,J} \subsetneq H^-}} \max\left\{\frac{1}{c_2|Y_T|} \cdot \frac{10^{|I|}}{p^{|I|}}, C_2\right\} \cdot \frac{1}{n^{|I|+|J|} p^{e_{H_{I,J}}}}$$
$$= e(\mathcal{H}) \cdot n^{v_2} p^{e_{H^-}} \cdot \sum_{\substack{\emptyset \neq I \subseteq W_1, J \subseteq W_2 \\ \emptyset \neq H_{I,J} \subsetneq H^-}} \max\left\{\frac{n}{c_2|Y_T|} \cdot \frac{10^{|I|}}{np^{|I|}}, C_2\right\} \cdot \frac{1}{n^{v_{H_{I,J}}} p^{e_{H_{I,J}}}}.$$

Fix a nonempty $I \subseteq W_1$ and a $J \subseteq W_2$ such that $H_{I,J}$ is nonempty. Since $v_{H_{I,J}} + 1$ and $e_{H_{I,J}} + |I| \geqslant 2$ are the numbers of vertices and edges of the subgraph of H induced by $\{v_c\} \cup I \cup J$, Lemma 3.5 implies that

$$\max\left\{\frac{n}{c_2|Y_T|} \cdot \frac{10^{|I|}}{np^{|I|}}, C_2\right\} \cdot \frac{1}{n^{v_{H_{I,J}}} p^{e_{H_{I,J}}}} \leqslant \max\left\{\frac{n}{c_2|Y_T|} \cdot \frac{2 \cdot 10^{|I|}}{\tilde{C}}, C_2\right\} \cdot \frac{1}{n^2 p}.$$

Recalling again that $e(\mathcal{H}) \cdot n^{v_2} p^{e_{H^-}} \leqslant \mu / c$, we have

$$\frac{\Delta_1}{\mu} \leqslant \frac{\mu}{c} \cdot 2^{v_1+v_2} \cdot \max\left\{\frac{n \cdot 10^{v_1+1}}{c_2|Y_T| \cdot \tilde{C}}, C_2\right\} \cdot \frac{2}{m}.$$

We may conclude that

$$\frac{\mu^2}{\Delta} \geqslant \frac{\mu^2}{\Delta_0 + \Delta_1} \geqslant c' \cdot \min\left\{\frac{c_2 \cdot \tilde{C} \cdot |Y_T|}{n}, \frac{1}{C_2}\right\} \cdot m,$$

where c' is a positive constant that depends only on H. □

Finally, we invoke Lemma 3.1 with $q = \frac{\mu}{\mu+\Delta} \leqslant 1$ to conclude that

$$\frac{|\mathcal{F}^*(T;Z)|}{\binom{e(\Pi)}{m-e(Z)}} \leqslant \mathbb{P}(K \not\subseteq G' \text{ for every } K \in \mathcal{K}') \leqslant \exp\left(-\frac{\mu^2}{\mu+\Delta} + \frac{\mu^2 \Delta}{2(\mu+\Delta)^2}\right)$$
$$\leqslant \exp\left(-\frac{\mu^2}{2(\mu+\Delta)}\right) \leqslant \exp\left(-\min\left\{\frac{\mu}{4}, \frac{\mu^2}{4\Delta}\right\}\right).$$

The assertion of the lemma now follows from Claims 8.11 and 8.12. □

8.10. **The irregular case.** In this section, we prove Lemma 8.9, that is, for given $T' \subseteq \Pi^c$, we give an upper bound on the number of graphs $Z \in \mathcal{Z}^I(T)$, for some $T \in \mathcal{T}_{\mathrm{H}}(\Pi)$ satisfying $|Y_T| = b$, such that $T'_Z = T'$.

Proof of Lemma 8.9. Fix some graph $T' \subseteq \Pi^c$, an integer $b \geqslant 1$, a colour $i \in [\![r]\!]$, and distinct $v_1, \ldots, v_b \in V_i$. We will describe a procedure that constructs, for every graph Z such that $T'_Z = T'$ and $Y_{T_Z} = \{v_1, \ldots, v_b\}$, a hypergraph $\mathcal{H} \subseteq \mathcal{H}_Z$ that satisfies condition (52) for every nonempty $I \subseteq W_1$. Our procedure will examine the neighbourhoods of $v_1, \ldots, v_b$ in the graph $Z \setminus T'$ one-by-one and build $\mathcal{H}$ in an online fashion. If $Z \in \mathcal{Z}^I_{\Pi}(b)$, then the constructed hypergraph $\mathcal{H}$ cannot have too many edges. More precisely, $\mathcal{H}$ has to fail condition (51) and, moreover, $\mathcal{H}_Z$ must not satisfy (50). This means, roughly speaking, that, when $Z \in \mathcal{Z}^I_{\Pi}(b)$, the neighbourhoods of $v_1, \ldots, v_b$ in $Z \setminus T'$ are highly correlated. This will allow us, with the use of Lemma 3.3, to bound the number of choices for these neighbourhoods that result in a graph $Z \in \mathcal{Z}^I_{\Pi}(b)$. Consequently, we will obtain an upper bound on the size of the set $\mathcal{Z}^I_{\Pi}(b; T')$.

Let

$$D_* = \left\lfloor \frac{D_H}{v_1} \right\rfloor \geqslant \left\lfloor \frac{\rho m}{2 v_1 n} \right\rfloor$$

and let $\mathcal{H}_0$ be the empty hypergraph with vertex set $\bigsqcup_{w \in W_1} V_{\psi(w)}$. Do the following for $s = 1, \ldots, b$:

(i) For every nonempty $I \subsetneq W_1$, let

$$M^I_s = \left\{ L \in \prod_{w \in I} V_{\psi(w)} : \deg_{\mathcal{H}_{s-1}}(L) > \frac{C_2}{2} \cdot \frac{e(\mathcal{H}_{s-1})}{n^{|I|}} \right\}.$$

(ii) For each $j \in [\![r]\!]$, choose an arbitrary collection $\{N_{j,w}(v_s)\}_{w \in W_1}$ of v_1 pairwise disjoint subsets of $N_j(v_s)$, each of size D_*, denote $N(v_s) = \prod_{w \in W_1} N_{\psi(w),w}(v_s)$, and let

$$\mathcal{H}_s = \mathcal{H}_{s-1} \cup \left\{ K \in N(v_s) : L \nsubseteq K \text{ for all } L \in \bigcup_{\emptyset \neq I \subsetneq W_1} M^I_s \right\}.$$

Finally, let $\mathcal{H} = \mathcal{H}_b$.

By construction, every $(v_w)_{w \in W_1} \in N(v_s)$ has distinct coordinates and hence $\mathcal{H} \subseteq \mathcal{H}_Z$. Moreover, for every nonempty $I \subsetneq W_1$,

$$\begin{aligned}
\Delta_I(\mathcal{H}) &\leqslant \frac{C_2}{2} \cdot \frac{e(\mathcal{H})}{n^{|I|}} + \Delta_I\big(N(v_s)\big) \leqslant \frac{C_2}{2} \cdot \frac{e(\mathcal{H})}{n^{|I|}} + \prod_{w \in W_1 \setminus I} |N_{\psi(w)}(v_s)| \\
&\leqslant \max\left\{ 2 D_H^{v_1 - |I|}, C_2 \cdot \frac{e(\mathcal{H})}{n^{|I|}} \right\} \leqslant \max\left\{ \left(\frac{m}{n}\right)^{v_1 - |I|}, C_2 \cdot \frac{e(\mathcal{H})}{n^{|I|}} \right\},
\end{aligned}$$

as $|N_j(v_s)| = D_H \leqslant \rho m/n \leqslant m/(2n)$ for all $j \in [\![r]\!]$ and $s \in [\![b]\!]$. Moreover, since $\Delta_{W_1}(\mathcal{H}) \leqslant 1 = (m/n)^{v_1 - |W_1|}$, our $\mathcal{H}$ satisfies (52) for every nonempty $I \subseteq W_1$.

We say that $s \in [\![b]\!]$ is *useful* if

$$e\big(\mathcal{H}_s \setminus \mathcal{H}_{s-1}\big) \geqslant 2^{-v_1} \cdot D_*^{v_1}.$$

If more than half of $s \in [\![b]\!]$ are useful, then

$$e(\mathcal{H}) = \sum_{s=1}^{b} e\big(\mathcal{H}_s \setminus \mathcal{H}_{s-1}\big) \geqslant \frac{b}{2} \cdot 2^{-v_1} \cdot D_*^{v_1} \geqslant 2^{-v_H} \cdot b \cdot \left\lfloor \frac{\rho m}{2v_1 n} \right\rfloor^{v_1} \geqslant c_2 \cdot b \cdot \left(\frac{m}{n}\right)^{v_1},$$

where the last inequality follows from (48); in particular $\mathcal{H}$ satisfies condition (51) and thus $Z \in \mathcal{Z}_2^R(T_Z)$. Therefore, if $Z \in \mathcal{Z}_{\Pi}^I(b)$, then at least half of $s \in [\![b]\!]$ are not useful.

Claim 8.13. Let $s \in [\![b]\!]$ and suppose that $e(\mathcal{H}_{s-1}) < \sigma n^{v_1}$. Then, there are at most

$$\exp\left(-\frac{4\Gamma m}{n}\right) \cdot \binom{n}{rD_H}$$

choices for $N_1(v_s), \ldots, N_r(v_s)$ such that s is not useful.

Proof. For every $I \subseteq W_1$, denote $N_I(v_s) = \prod_{w \in I} N_{\psi(w),w}(v_s)$, where $\{N_{j,w}(v_s)\}$ is the collection defined in step (ii) of the algorithm building $\mathcal{H}$. Letting $M_s^{W_1} = \mathcal{H}_{s-1}$, we have

$$e\big(\mathcal{H}_s \setminus \mathcal{H}_{s-1}\big) \geqslant D_*^{v_1} - \sum_{\emptyset \neq I \subseteq W_1} |N_I(v_s) \cap M_s^I| \cdot D_*^{v_1 - |I|}.$$

In particular, if s is not useful then there must be some nonempty $I \subseteq W_1$ such that

$$\big|N_I(V_s) \cap M_s^I\big| > 2^{-v_1} \cdot D_*^{|I|}. \tag{58}$$

Since $|V_j| \geqslant n/(2r)$ for every $j \in [\![r]\!]$, we have

$$|M_s^{W_1}| = e(\mathcal{H}_{s-1}) < \sigma n^{v_1} \leqslant \sigma \cdot (2r)^{v_1} \cdot \prod_{w \in W_1} |V_{\psi(w)}|.$$

Moreover, for every $\emptyset \neq I \subsetneq W_1$,

$$|M_s^I| \cdot \frac{C_2}{2} \cdot \frac{e(\mathcal{H}_{s-1})}{n^{|I|}} \leqslant \sum_{L \in M_s^I} \deg_{\mathcal{H}_{s-1}}(L) \leqslant \binom{v_1}{|I|} \cdot e(\mathcal{H}_{s-1})$$

and hence

$$|M_s^I| \leqslant \frac{1}{C_2} \cdot \binom{v_1}{|I|} \cdot n^{|I|} \leqslant \frac{2^{v_1}}{C_2} \cdot n^{|I|} \leqslant \frac{(4r)^{v_1}}{C_2} \cdot \prod_{w \in I} |V_{\psi(w)}|.$$

Since we chose σ to be sufficiently small and C_2 to be sufficiently large as a function of α and v_1, see (49), Lemma 3.3 applied $2^{v_1} - 1$ times implies that there are at most

$$(2^{v_1} - 1) \cdot \alpha^{D_*} \cdot \prod_{w \in W_1} \binom{|V_{\psi(w)}|}{D_*}$$

choices of $N(v_s)$ such that (58) holds for some nonempty $I \subseteq W_1$. On the other hand, the number of choices for $N_1(v_s), \ldots, N_r(v_s)$ that can yield a given $N(v_s)$ is at most $\binom{n}{rD_H - v_1 D_*}$. We conclude that the number X of choices for $N_1(v_s), \ldots, N_r(v_s)$ that render s not useful satisfies

$$\begin{aligned} X &\leqslant 2^{v_1} \cdot \alpha^{D_*} \cdot \binom{n}{rD_H - v_1 D_*} \cdot \prod_{w \in W_1} \binom{|V_{\psi(w)}|}{D_*} \\ &= 2^{v_1} \cdot \alpha^{D_*} \cdot \binom{n}{rD_H} \cdot \underbrace{\binom{rD_H}{v_1 D_*} \cdot \binom{n - rD_H + v_1 D_*}{v_1 D_*}^{-1} \cdot \prod_{w \in W_1} \binom{|V_{\psi(w)}|}{D_*}}_{(\star)}. \end{aligned}$$

Since $n - rD_H \geqslant 2n/3 \geqslant |V_j|$ for every $j \in [\![r]\!]$, we have

$$(\star) \leqslant \binom{2n/3 + v_1 D_*}{v_1 D_*}^{-1} \cdot \binom{2n/3}{D_*}^{v_1} \leqslant \binom{2n/3 + v_1 D_*}{v_1 D_*}^{-1} \cdot \binom{v_1 \cdot 2n/3}{v_1 D_*} \overset{(4)}{\leqslant} v_1^{v_1 D_*}.$$

Finally, since $v_1 D_* \geqslant D_H - v_1 \geqslant 2D_H/3 \geqslant \rho m/(3n)$, we conclude that

$$X \cdot \binom{n}{rD_H}^{-1} \leqslant 2^{v_1} \cdot \alpha^{D_*} \cdot \binom{rD_H}{v_1 D_*} \cdot v_1^{v_1 D_*} \overset{(5)}{\leqslant} \left(2 \cdot \alpha^{1/v_1} \cdot \frac{erD_H}{v_1 D_*} \cdot v_1\right)^{v_1 D_*}$$
$$\leqslant \left(3er \cdot \alpha^{1/v_1} v_1\right)^{\frac{\rho m}{3n}} \overset{(47)}{\leqslant} \exp\left(-\frac{4\Gamma m}{n}\right),$$

giving the assertion of the claim. □

We are now ready to prove the claimed upper bound on the size of the family $\mathcal{Z}_\Pi^I(b; T')$. Each graph Z in this family can be constructed by specifying an $i \in [\![r]\!]$, a sequence of distinct vertices $v_1, \ldots, v_b \in V_i$, and a set $S \subseteq [\![b]\!]$ of size at least $b/2$ such that, when we execute the algorithm described above, every $s \in S$ is not useful. Since the number of choices for $N_1(v_s), \ldots, N_r(v_s)$ is at most $\exp(-4\Gamma m/n) \cdot \binom{n}{rD_H}$ when $s \in S$, by Claim 8.13, and at most $\binom{n}{rD_H}$ when $s \in [\![r]\!] \setminus S$, we have

$$|\mathcal{Z}_\Pi^I(b; T')| \leqslant r \cdot n^b \cdot 2^b \cdot \exp\left(-\frac{4\Gamma m}{n} \cdot \frac{b}{2}\right) \binom{n}{rD_H}^b \overset{(5)}{\leqslant} \exp\left(-\frac{3\Gamma m}{2n} \cdot b\right) \left(\frac{en}{rD_H}\right)^{brD_H}.$$

On the other hand, by (30), which holds for all $y \leqslant m' \leqslant m \leqslant e(\Pi) - \xi n^2$, we have

$$\frac{\binom{e(\Pi)}{m - e(T') - brD_H}}{\binom{e(\Pi)}{m - e(T')}} \leqslant \left(\frac{m}{\xi n^2}\right)^{brD_H}.$$

It follows that

$$|\mathcal{Z}_\Pi^I(b; T')| \cdot \binom{e(\Pi)}{m - e(T') - brD_H}$$
$$\leqslant \exp\left(-\frac{3\Gamma m}{2n} \cdot b\right) \cdot \left(\frac{en}{rD_H} \cdot \frac{m}{\xi n^2}\right)^{brD_H} \binom{e(\Pi)}{m - e(T')}.$$

The claimed bound follows after noting that, since $(ea/x)^x \leqslant e^a$ for all $x \in (0, \infty)$,

$$\left(\frac{en}{rD_H} \cdot \frac{m}{\xi n^2}\right)^{rD_H} \leqslant \exp\left(\frac{m}{\xi n}\right) \overset{(46)}{\leqslant} \exp\left(\frac{\Gamma m}{2n}\right). \qquad \square$$

9. The 1-statement: the dense case

Fix a partition $\Pi \in \mathcal{P}_{n,r}(\gamma)$. In this section, we verify the assumptions of Proposition 7.1 in the case where

$$e(\Pi) - \xi n^2 \leqslant m \leqslant \mathrm{ex}(n, H).$$

We start by introducing two additional parameters. Let ε and ν be positive constants satisfying

$$\varepsilon + v_H \nu \leqslant \frac{1}{2r} \qquad \text{and} \qquad \varepsilon \leqslant \nu/8. \tag{59}$$

Earlier on, we chose γ, δ, and ξ sufficiently small so that

$$320\xi \leqslant \nu \qquad \text{and} \qquad \left(\frac{4e}{\nu\varepsilon}\right)^{\varepsilon} \cdot \left(\frac{320\xi}{\nu}\right)^{\nu/4} \leqslant e^{-2} \tag{60}$$

and, additionally,

$$\xi + \delta \leqslant 2\max\{\xi, \delta\} < \min\left\{\frac{\varepsilon^2}{v_H^2}, \frac{\nu}{8r}\right\} \qquad \text{and} \qquad \gamma \leqslant \frac{1}{20r}. \tag{61}$$

In order to show that the assumptions of Proposition 7.1 are satisfied, we will define a natural map $\mathcal{M}\colon \mathcal{F}^*_{n,m}(H;\delta,\Pi) \to \mathcal{B}(\Pi,k)$ by letting $\mathcal{M}(G)$ be the subgraph of $G \setminus \Pi$ obtained by deleting from it all vertices that are non-adjacent to more than νn vertices of a different colour class of Π; we shall show that this graph has maximum degree k. We will then estimate the left-hand side of (18) using ad-hoc, combinatorial arguments.

Suppose that $\Pi = \{V_1, \dots, V_r\}$. For a graph $G \in \mathcal{F}^*_{n,m}(H;\delta,\Pi)$, let X_G denote the set of all vertices of G that have fewer than $|V_j| - \nu n$ neighbours in some colour class V_j other than their own. More precisely,

$$X_G = \bigcup_{i=1}^{r} \big\{v \in V_i : \deg_G(v, V_j) < |V_j| - \nu n \text{ for some } j \neq i\big\}.$$

We first show that the set X_G is rather small and that the graph $(G \setminus \Pi) - X_G$ has maximum degree at most k.

Lemma 9.1. *For every $G \in \mathcal{F}^*_{n,m}(H;\delta,\Pi)$, we have*

$$|X_G| \leqslant \frac{n}{4r}.$$

Proof. Since

$$e(\Pi) - e(G \cap \Pi) = \frac{1}{2} \cdot \sum_{i=1}^{r} \sum_{v \in V_i} \sum_{j \neq i} \big(|V_j| - \deg(v, V_j)\big) \geqslant \frac{1}{2} \cdot |X_G| \cdot \nu n,$$

we have

$$e(\Pi) - \xi n^2 \leqslant e(G) \leqslant e(G \cap \Pi) + \delta n^2 \leqslant e(\Pi) + \delta n^2 - \frac{1}{2} \cdot |X_G| \cdot \nu n.$$

We conclude that

$$|X_G| \leqslant 2 \cdot \frac{\delta + \xi}{\nu} \cdot n \overset{(61)}{\leqslant} \frac{n}{4r},$$

as claimed. □

Lemma 9.2. *For every $G \in \mathcal{F}^*_{n,m}(H;\delta,\Pi)$, the maximum degree of $(G \setminus \Pi) - X_G$ is at most k.*

Proof. Suppose that there were a $G \in \mathcal{F}^*_{n,m}(H;\delta,\Pi)$ such that $(G \setminus \Pi) - X_G$ has a vertex v of degree at least $k+1$. Let $\Pi = \{V_1, \dots, V_r\}$ and suppose that $v \in V_i$. Let $u_1, \dots, u_{k+1} \in V_i \setminus X_G$ be arbitrary neighbours of v and let $u_{k+2}, \dots, u_{v_H - 1}$ be arbitrary vertices of $V_i \setminus X_G$ that are distinct from v and $u_1, \dots, u_{k+1}$; such vertices exist since, by Lemma 9.1, we have $|V_i \setminus X_G| \geqslant n/(2r) - n/(4r) = n/(4r)$. For each $j \in [\![r]\!] \setminus \{i\}$, let

$$N_j = V_j \cap N_G(v) \cap \bigcap_{\ell=1}^{v_H - 1} N_G(u_\ell).$$

and observe that, by the definition of X_G,

$$|N_j| \geqslant |V_j| - v_H \cdot \nu n \geqslant n/(2r) - v_H \cdot \nu n \overset{(59)}{\geqslant} \varepsilon n.$$

Observe further that the subgraph of $G \cap \Pi$ that is induced by $N_1 \cup \dots \cup N_{i-1} \cup N_{i+1} \cup \dots \cup N_r$ is $K_{r-1}(v_H)$-free; indeed, otherwise G would contain every $(r+1)$-colourable vertex-critical graph of criticality $k+1$ with at most v_H vertices, contradicting the fact that G is H-free. This implies, in particular, that

$$e(\Pi) - e(G \cap \Pi) \geqslant e\big(K_{r-1}(\varepsilon n)\big) - \mathrm{ex}\big(K_{r-1}(\varepsilon n), K_{r-1}(v_H)\big) \geqslant (\varepsilon n)^2/v_H^2,$$

where the last inequality follows from Lemma 3.6. Consequently,

$$m = e(G \cap \Pi) + e(G \setminus \Pi) \leqslant e(\Pi) - (\varepsilon n)^2/v_H^2 + \delta n^2 \overset{(61)}{<} e(\Pi) - \xi n^2,$$

a contradiction. □

For every $G \in \mathcal{F}^*_{n,m}(H;\delta,\Pi)$, let $B_G = (G \setminus \Pi) - X_G$ and note, by Lemma 9.2, we have $B_G \in \mathcal{B}(\Pi, k)$. Define, for every $B \in \mathcal{B}(\Pi, k)$,

$$\mathcal{F}^*_B = \{G \in \mathcal{F}^*_{n,m}(H;\delta,\Pi) : B_G = B\}.$$

The following proposition, which is the main result of this section, implies that the map $\mathcal{M} \colon G \mapsto B_G$ satisfies the assumptions of Proposition 7.1.

Proposition 9.3. *For every $B \in \mathcal{B}(\Pi, k)$, we have*

$$|\mathcal{F}^*_B| \leqslant \exp(-n) \cdot \binom{e(\Pi)}{m - e(B)}.$$

The proof of Proposition 9.3 will require one additional lemma, which states that the maximum degree of the graph $G \setminus \Pi$ cannot be very large.

Lemma 9.4. *For every $G \in \mathcal{F}^*_{n,m}(H;\delta,\Pi)$, we have $\Delta(G \setminus \Pi) < \varepsilon n$.*

Proof. Suppose that there was a $G \in \mathcal{F}^*_{n,m}(H;\delta,\Pi)$ such that $\Delta(G \setminus \Pi) \geqslant \varepsilon n$ and pick an arbitrary v with $\deg_{G\setminus\Pi}(v) \geqslant \varepsilon n$. Suppose that $\Pi = \{V_1, \dots, V_r\}$ and recall from (16) that $\deg_G(v, V_i) \geqslant \varepsilon n$ for every $i \in [\![r]\!]$. For each i, let $N_i \subseteq N(v) \cap V_i$ be an arbitrary subset of size exactly εn. Observe that the subgraph of $G \cap \Pi$ that is induced by $N_1 \cup \dots \cup N_r$ is $K_r(v_H)$-free; indeed, otherwise G would contain every vertex-critical $(r+1)$-colourable graph with at most v_H vertices, contradicting the fact that G is H-free. This implies, in particular, that

$$e(\Pi) - e(G \cap \Pi) \geqslant e\big(K_r(\varepsilon n)\big) - \mathrm{ex}\big(K_r(\varepsilon n), K_r(v_H)\big) \geqslant (\varepsilon n)^2/v_H^2,$$

where the last inequality follows from Lemma 3.6. Consequently,

$$m = e(G \cap \Pi) + e(G \setminus \Pi) \leqslant e(\Pi) - (\varepsilon n)^2/v_H^2 + \delta n^2 \overset{(61)}{<} e(\Pi) - \xi n^2,$$

a contradiction. □

Proof of Proposition 9.3. Fix an arbitrary $B \in \mathcal{B}(\Pi, k)$ and choose some $G \in \mathcal{F}^*_B$. Note that X_G cannot be empty as otherwise $G \subseteq \Pi \cup B$, contradicting the fact that $G \in \mathcal{F}^*_{n,m}(H;\delta,\Pi)$. Moreover, by Lemma 9.4, every vertex of X_G has degree at most εn in $G \setminus \Pi$. Denote $\Pi_{X_G} = \Pi \setminus (\Pi - X_G)$; in other words, Π_{X_G} comprises all edges of Π that have an endpoint in X_G. By the definition of X_G,

$$e(\Pi_{X_G}) - e(G \cap \Pi_{X_G}) \geqslant \frac{1}{2} \cdot |X_G| \cdot \nu n.$$

Observe that every graph in $G \in \mathcal{F}^*_B$ may be constructed as follows:

- Choose a nonempty vertex set X with at most $n/(4r)$ elements (to serve as X_G).

- Choose at most $|X| \cdot \varepsilon n$ edges of Π^c, each touching X, to form $(G \setminus \Pi) \setminus B$.
- Choose at most $e(\Pi_X) - |X| \cdot \nu n/2$ edges of Π_X to form $G \cap \Pi_X$.
- Choose the remaining edges of G from $\Pi - X$.

In particular, letting

$$t_X = |X| \cdot \varepsilon n \qquad \text{and} \qquad z_X = e(\Pi_X) - |X| \cdot \nu n/2,$$

we have

$$|\mathcal{F}_B^*| \leqslant \sum_{\substack{X \neq \emptyset \\ |X| \leqslant n/(4r)}} \sum_{t \leqslant t_X} \sum_{z \leqslant z_X} \binom{|X| \cdot n}{t} \binom{e(\Pi_X)}{z} \binom{e(\Pi - X)}{m - z - t - e(B)}.$$

Note that, for all X and all $t \leqslant t_X$ and $z \leqslant z_X$,

$$\binom{e(\Pi_X)}{z} \cdot \binom{e(\Pi_X)}{z+t}^{-1} \overset{(2)}{\leqslant} \left(\frac{z_X + t_X}{e(\Pi_X) - z_X - t_X} \right)^t \leqslant \left(\frac{e(\Pi_X) + |X| \cdot (\varepsilon - \nu/2)n}{|X| \cdot (\nu/2 - \varepsilon)n} \right)^t$$
$$\leqslant \left(\frac{1 + \varepsilon - \nu/2}{\nu/2 - \varepsilon} \right)^t \overset{(59)}{\leqslant} \left(\frac{4}{\nu} \right)^t \leqslant \left(\frac{4}{\nu} \right)^{t_X},$$

so that

$$\binom{|X| \cdot n}{t} \binom{e(\Pi_X)}{z} \cdot \binom{e(\Pi)_X}{z+t}^{-1} \leqslant \binom{|X| \cdot n}{t_X} \cdot \left(\frac{4}{\nu} \right)^{t_X}$$
$$\overset{(5)}{\leqslant} \left(\frac{4e \cdot |X| \cdot n}{\nu t_X} \right)^{t_X} = \left(\frac{4e}{\nu \varepsilon} \right)^{|X| \cdot \varepsilon n}.$$

This gives

$$|\mathcal{F}_B^*| \leqslant \sum_{\substack{X \neq \emptyset \\ |X| \leqslant n/(4r)}} \left(\frac{4e}{\nu \varepsilon} \right)^{|X| \cdot \varepsilon n} \sum_{z \leqslant z_X} \sum_{t \leqslant t_X} \underbrace{\binom{e(\Pi_X)}{z+t} \binom{e(\Pi - X)}{m - z - t - e(B)}}_{N_{X, z+t}}. \tag{62}$$

By Vandermonde's identity, we have, for every y,

$$N_{X,y} \leqslant \binom{e(\Pi_X) + e(\Pi - X)}{m - e(B)} = \binom{e(\Pi)}{m - e(B)}. \tag{63}$$

Moreover, direct calculation shows that

$$\frac{N_{X,y}}{N_{X,y+1}} = \underbrace{\frac{y+1}{e(\Pi_X) - y}}_{\rho_{X,y}} \cdot \underbrace{\frac{e(\Pi - X) - m + y + 1 + e(B)}{m - y - e(B)}}_{\rho'_{X,y}}. \tag{64}$$

Set $y_X = z_X + t_X$ and

$$y'_X = y_X + |X| \cdot \nu n/4 = e(\Pi_X) - |X| \cdot (\nu/4 - \varepsilon)n \overset{(59)}{\leqslant} e(\Pi_X) - |X| \cdot \nu n/8.$$

Assume that $|X| \leqslant n/(4r)$ and $y + 1 \leqslant y'_X$. Using $e(\Pi_X) \leqslant |X| \cdot n$, we have $\rho_y \leqslant 8/\nu$. Moreover,

$$m - y - 1 - e(B) \geqslant e(\Pi) - \xi n^2 - e(\Pi_X) - kn \geqslant e(\Pi - X) - 2\xi n^2$$

and, by part (i) of Proposition 4.1 and our assumption that $\gamma \leqslant \frac{1}{20r}$,

$$m - y - e(B) \geqslant e(\Pi) - 2\xi n^2 - |X| \cdot n \geqslant \frac{n^2}{5} - 2\xi n^2 - \frac{n^2}{4r} \geqslant \frac{n^2}{20}.$$

Consequently, $\rho'_{X,y} \leqslant 40\xi$. Substituting these two estimates into (64) yields

$$\frac{N_{X,y}}{N_{X,y+1}} \leqslant \frac{320\xi}{\nu} \overset{(60)}{\leqslant} 1. \tag{65}$$

We may conclude that, when $|X| \leqslant n/(4r)$ and $y \leqslant y_X$,

$$N_{X,y} = N_{X,y'_X} \cdot \prod_{y'=y}^{y'_X-1} \frac{N_{X,y'}}{N_{X,y'+1}} \overset{(63),\,(65)}{\leqslant} \binom{e(\Pi)}{m-e(B)} \cdot \left(\frac{320\xi}{\nu}\right)^{y'_X-y}$$
$$\overset{(65)}{\leqslant} \binom{e(\Pi)}{m-e(B)} \cdot \left(\frac{320\xi}{\nu}\right)^{y'_X-y_X} = \binom{e(\Pi)}{m-e(B)} \cdot \left(\frac{320\xi}{\nu}\right)^{|X|\cdot\nu n/4}.$$

Finally, substituting this estimate into (62) yields

$$|\mathcal{F}^*_B| \cdot \binom{e(\Pi)}{m-e(B)}^{-1} \leqslant \sum_{\substack{X\neq\emptyset \\ |X|\leqslant n/(4r)}} \left(\frac{4e}{\nu\varepsilon}\right)^{|X|\cdot\varepsilon n} (z_X+1)(t_X+1) \cdot \left(\frac{320\xi}{\nu}\right)^{|X|\cdot\nu n/4}$$
$$\leqslant \sum_{x=1}^{n/(4r)} \binom{n}{x} \cdot (xn)^2 \cdot \left[\left(\frac{4e}{\nu\varepsilon}\right)^{\varepsilon} \cdot \left(\frac{320\xi}{\nu}\right)^{\nu/4}\right]^{xn}$$
$$\overset{(60)}{\leqslant} \sum_{x=1}^{n} n^{5x} \cdot e^{-2xn} \leqslant e^{-n},$$

provided that n is sufficiently large. □

References

1. J. Balogh, B. Bollobás, and M. Simonovits, *The typical structure of graphs without given excluded subgraphs*, Random Structures Algorithms **34** (2009), 305–318.
2. J. Balogh, N. Bushaw, M. Collares, H. Liu, R. Morris, and M. Sharifzadeh, *The typical structure of graphs with no large cliques*, Combinatorica **37** (2017), no. 4, 617–632.
3. J. Balogh, R. Morris, and W. Samotij, *Independent sets in hypergraphs*, J. Amer. Math. Soc. **28** (2015), 669–709.
4. J. Balogh, R. Morris, W. Samotij, and L. Warnke, *The typical structure of sparse K_{r+1}-free graphs*, Trans. Amer. Math. Soc. **368** (2016), no. 9, 6439–6485.
5. J. Balogh and W. Samotij, *An efficient container lemma*, Discrete Anal. (2020), Paper No. 17, 56.
6. B. Bollobás, *A probabilistic proof of an asymptotic formula for the number of labelled regular graphs*, European J. Combin. **1** (1980), no. 4, 311–316.
7. O. Engelberg, *On the typical structure of graphs not containing a fixed vertex-critical subgraph*, Master's thesis, Tel Aviv University, 2017.
8. P. Erdős, D. J. Kleitman, and B. L. Rothschild, *Asymptotic enumeration of K_n-free graphs*, Colloquio Internazionale sulle Teorie Combinatorie (Rome, 1973), Tomo II, Accad. Naz. Lincei, Rome, 1976, pp. 19–27. Atti dei Convegni Lincei, No. 17.
9. P. Erdős and M. Simonovits, *A limit theorem in graph theory*, Studia Sci. Math. Hungar. **1** (1966), 51–57.
10. P. Erdős and A. H. Stone, *On the structure of linear graphs*, Bull. Amer. Math. Soc. **52** (1946), 1087–1091.
11. T. E. Harris, *A lower bound for the critical probability in a certain percolation process*, Proc. Cambridge Philos. Soc. **56** (1960), 13–20.
12. C. Hundack, H. J. Prömel, and A. Steger, *Extremal graph problems for graphs with a color-critical vertex*, Combin. Probab. Comput. **2** (1993), 465–477.
13. S. Janson, *Poisson approximation for large deviations*, Random Structures Algorithms **1** (1990), 221–229.

14. S. Janson, T. Łuczak, and A. Ruciński, *Random graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2000.
15. P. Kolaitis, H. J. Prömel, and B. Rothschild, *K_{l+1}-free graphs: asymptotic structure and a $0-1$ law*, Trans. Amer. Math. Soc. **303** (1987), 637–671.
16. F. Mousset, R. Nenadov, and A. Steger, *On the number of graphs without large cliques*, SIAM J. Discrete Math. **28** (2014), 1980–1986.
17. D. Osthus, H. J. Prömel, and A. Taraz, *For which densities are random triangle-free graphs almost surely bipartite?*, Combinatorica **23** (2003), no. 1, 105–150, Paul Erdős and his mathematics (Budapest, 1999).
18. H. J. Prömel and A. Steger, *The asymptotic number of graphs not containing a fixed color-critical subgraph*, Combinatorica **12** (1992), 463–473.
19. ———, *Random l-colorable graphs*, Random Structures Algorithms **6** (1995), 21–37.
20. ———, *On the asymptotic structure of sparse triangle free graphs*, J. Graph Theory **21** (1996), 137–151.
21. M. Simonovits, *Extermal graph problems with symmetrical extremal graphs. Additional chromatic conditions*, Discrete Math. **7** (1974), 349–376.
22. P. Turán, *Eine Extremalaufgabe aus der Graphentheorie*, Mat. Fiz. Lapok **48** (1941), 436–452.
23. N. C. Wormald, *The asymptotic distribution of short cycles in random regular graphs*, J. Combin. Theory Ser. B **31** (1981), 168–182.

School of Mathematical Sciences, Tel Aviv University, Tel Aviv 6997801, Israel
Email address: oengel48@gmail.com

School of Mathematical Sciences, Tel Aviv University, Tel Aviv 6997801, Israel
Email address: samotij@tauex.tau.ac.il

School of Mathematics, Georgia Institute of Technology, Atlanta GA 30332, US
Email address: warnke@math.gatech.edu