AN ASYMMETRIC CONTAINER LEMMA AND
THE STRUCTURE OF GRAPHS WITH NO INDUCED 4-CYCLE

ROBERT MORRIS, WOJCIECH SAMOTIJ, AND DAVID SAXTON

ABSTRACT. The method of hypergraph containers, introduced recently by Balogh, Morris, and Samotij, and independently by Saxton and Thomason, has proved to be an extremely useful tool in the study of various monotone graph properties. In particular, a fairly straightforward application of this technique allows one to locate, for each non-bipartite graph H, the threshold at which the distribution of edges in a typical H-free graph with a given number of edges undergoes a transition from 'random-like' to 'structured'. On the other hand, for non-monotone hereditary graph properties the standard version of this method does not allow one to establish even the existence of such a threshold.

In this paper we introduce a refinement of the container method that takes into account the asymmetry between edges and non-edges in a sparse member of a hereditary graph property. As an application, we determine the approximate structure of a typical graph with n vertices, m edges, and no induced copy of the 4-cycle, for each function $m = m(n)$ satisfying $n^{4/3}(\log n)^4 \leqslant m \ll n^2$. We show that almost all such graphs G have the following property: the vertex set of G can be partitioned into an 'almost-independent' set (a set with $o(m)$ edges) and an 'almost-clique' (a set inducing a subgraph with density $1 - o(1)$). The lower bound on m is optimal up to a polylogarithmic factor, as standard arguments show that if $n \ll m \ll n^{4/3}$, then almost all such graphs are 'random-like'. As a further consequence, we deduce that the random graph $G(n, p)$ conditioned to contain no induced 4-cycles undergoes phase transitions at $p = n^{-2/3+o(1)}$ and $p = n^{-1/3+o(1)}$.

1. INTRODUCTION

Two of the central objects of study in combinatorics are the family of H-*free* graphs, that is, the collection of graphs that do not contain H as a subgraph, and the family of *induced-H-free* graphs, that is, graphs without an induced subgraph isomorphic to H. An extremely well-studied problem (see, e.g., [25] and references therein) is to determine the largest number of edges in an H-free graph with a given number of vertices. This line of research dates back to the seminal works of Turán [47] and of Erdős and Stone [23], which are considered to be the cornerstones of the field of extremal graph theory.

Another natural and well-studied problem, which also makes sense in the setting of induced-H-free graphs, can be informally phrased as follows:

> What does a typical H-free (or induced-H-free) graph look like?

The first to address this problem were Erdős, Kleitman, and Rothschild [22], who proved that almost all triangle-free graphs are bipartite. That is, the proportion of triangle-free graphs on a given set of n vertices that are bipartite (among all triangle-free graphs) tends to one as n tends to infinity. This result was generalised by Kolaitis, Prömel, and Rothschild [34], who showed that for every $r \geqslant 2$, almost all K_{r+1}-free graphs are

r-partite, and later by Prömel and Steger [39], who showed that the same remains true if one replaces K_{r+1} with any $(r+1)$-colourable edge-critical[1] graph. Further results in this direction were obtained by Hundack, Prömel, and Steger [28] and by Balogh, Bollobás, and Simonovits [5, 6, 7]. Since the problem of describing the typical structure of an H-free graph is essentially a counting problem in disguise, we should also mention here the closely-related work of Erdős, Frankl, and Rödl [21], who estimated the number of H-free graphs for every non-bipartite H, observing a close connection between this counting problem and the extremal question mentioned above.

The problem of understanding the typical structure of induced-H-free graphs seems to be significantly harder and, as a result, much less is known. The pioneers of this line of research were Prömel and Steger, who described the typical structure of induced-C_4-free graphs [38] and induced-C_5-free graphs [42]. They also proved an analogue of the Erdős–Frankl–Rödl theorem for induced-H-free graphs [40] after finding the correct generalisation of the extremal question in this setting [41], which involves the notion of a *colouring number*. (This notion was later extended to the more general context of hereditary graph properties by Alekseev [2] and by Bollobás and Thomason [12].) Much later, Alon, Balogh, Bollobás, and Morris [3] gave a rough structural description of a typical induced-H-free graph for an arbitrary H (in fact, their result applies to all hereditary properties of graphs). Soon afterwards, Balogh and Butterfield [8] gave a precise structural description of a typical induced-H-free graph for all H that are critical[2]. Finally, let us mention two recent works of Kim, Kühn, Osthus, and Townsend [31] and of Keevash and Lochet [30] on the typical structure of induced-$C_{2\ell}$-free graphs and induced-$(K_{a+b} \setminus K_a)$-free graphs, respectively.

Even though many of the theorems above describe very precisely the structure of a typical H-free (or induced-H-free) graph, they say nothing about sparse H-free graphs. This is because the number of H-free graphs with n vertices is in each case much greater than the number of all n-vertex graphs with $o(n^2)$ edges; for example, there are more than $2^{n^2/4}$ bipartite (and hence H-free for any non-bipartite H) graphs with n vertices. This fact naturally leads one to consider the following refined question:

Question 1.1. Given a graph H and a function $m = m(n)$, what does a typical H-free (or induced-H-free) graph with n vertices and m edges look like?

The first to address this question were Prömel and Steger [43], who proved that almost every triangle-free graph with n vertices and m edges is bipartite whenever $m \geqslant C n^{7/4} \log n$ and it is not bipartite if $n \ll m \ll n^{3/2}$. A few years later, Łuczak [35] showed that this latter bound is (in some sense) sharp, by proving that if $m \gg n^{3/2}$, then almost every triangle-free graph with n vertices and m edges can be made bipartite by removing from it only $o(m)$ edges. More generally, it is not very hard to verify that if $n \ll m \ll n^{2-1/m_2(H)}$, where

$$m_2(H) = \max \left\{ \frac{e(F) - 1}{v(F) - 2} : F \subseteq H \text{ and } e(F) \geqslant 2 \right\}$$

---

[1] A graph H is *edge-critical* if it contains an edge e such that $\chi(H \setminus e) < \chi(H)$.

[2] The definition of criticality in the context of induced-H-free graphs is rather complicated, so we will only note here that it is a natural analogue of the notion of edge-criticality and refer the interested reader to [8] for the details.

is the so-called 2-*density* of H, then almost all H-free graphs with n vertices and m edges are *quasirandom*, in the sense that all sets of vertices of size $\Omega(n)$ induce subgraphs of (asymptotically) the same density. Łuczak [35] proved that, for every non-bipartite H, if a certain probabilistic version of the embedding lemma for regular partitions of sparse graphs (conjectured a few years earlier by Kohayakawa, Łuczak, and Rödl [33]) holds, then above the threshold, if $m \gg n^{2-1/m_2(H)}$, almost every H-free graph with n vertices and m edges can be made $(\chi(H) - 1)$-partite by removing only $o(m)$ edges.

The existence of this phase transition was confirmed several years ago by Balogh, Morris, and Samotij [10] and by Saxton and Thomason [44], using (what is now known as) the *method of hypergraph containers*. This method (see Section 1.2 or the recent survey [9]) allows one to prove the conjecture of Kohayakawa, Łuczak, and Rödl mentioned above, but also provides a more direct way of determining the rough structural description of a typical H-free graph above the 2-density threshold $n^{2-1/m_2(H)}$. We should also mention here the earlier works of Conlon and Gowers [14] and Schacht [45] on the closely related problem of determining the size and structure of the largest H-free subgraph of a random graph, since these breakthroughs had a significant impact on [10, 44].

The exact analogue of the Erdős–Kleitman–Rothschild theorem in the setting of sparse graphs was obtained by Osthus, Prömel, and Taraz [37], who proved that in fact $m = \frac{\sqrt{3}}{4} n^{3/2} (\log n)^{1/2}$ is a sharp threshold at which a typical triangle-free graph with n vertices and m edges becomes bipartite. A generalization of this result from triangle-free to K_{r+1}-free graphs, the sparse analogue of the Kolaitis–Prömel–Rothschild theorem, was obtained recently by Balogh, Morris, Samotij, and Warnke [11]. The exact analogue of Turán's theorem in $G(n, p)$, which sharpens the results of Conlon–Gowers and Schacht in the case $H = K_{r+1}$, was obtained by DeMarco and Kahn [17, 18].

Despite the significant developments described above on the problem of determining the typical structure of a sparse H-free graph, there has been (as far as we are aware) essentially no progress on the corresponding problem for induced-H-free graphs. One reason for this is that, in contrast to the case of H-free graphs, the hypergraph container method does not (in general) provide the correct threshold for the appearance of structure in a typical induced-H-free graph. From the point of view of the container theorems, an induced-H-free graph is a two-edge-coloured graph that does not contain a (two-edge-coloured) clique with the same number of vertices. Since the container method does not take into account the asymmetry between the two colours, it cannot distinguish between an induced copy of H and a clique.

In this paper we introduce a new 'asymmetric' version of the method of hypergraph containers that can distinguish between these two settings and provides the correct threshold for the emergence of structure in typical induced-H-free graphs (at least for non-bipartite graphs H, see Theorem 6.2). As an illustrative example, we use it to determine the structure of a typical induced-C_4-free graph with n vertices and m edges whenever $n^{4/3}(\log n)^4 \leqslant m \ll n^2$. The lower bound on m is best possible up to a polylogarithmic factor, as we shall also show that if $n \ll m \ll n^{4/3}(\log n)^{1/3}$, then a typical such graph does not exhibit a similar structure, and if $n \ll m \ll n^{4/3}$, then it is actually quasirandom (in the precise sense described above). We expect that the ideas contained in this work will allow analogous thresholds to be determined for families of graphs containing no induced copy of an arbitrary graph H, see Section 6.

1.1. **The structure of graphs with no induced** 4**-cycle.** Given a graph H and $n \in \mathbb{N}$, let $\mathcal{F}_n^{\mathrm{ind}}(H)$ denote the family of all graphs with vertex set $\{1, \dots, n\}$ that contain no induced copy of H and let $\mathcal{F}_{n,m}^{\mathrm{ind}}(H)$ denote the family of graphs in $\mathcal{F}_n^{\mathrm{ind}}(H)$ with precisely m edges. A *split graph* is a graph whose vertex set can be partitioned into a clique and an independent set. It is easy to check that a split graph cannot contain an induced copy of C_4; indeed, the property of being a split graph is hereditary and C_4 itself is not a split graph. Conversely, as mentioned above, it was proved by Prömel and Steger [38] over 25 years ago that almost all graphs in $\mathcal{F}_n^{\mathrm{ind}}(C_4)$ are split graphs. However, since almost all n-vertex split graphs admit a partition into a clique and an independent set of roughly equal sizes and have approximately $n^2/4$ edges, this result says nothing about a typical member of $\mathcal{F}_{n,m}^{\mathrm{ind}}(C_4)$ when m is not approximately $n^2/4$. It is worth mentioning that Gishboliner and Shapira [26] recently described the structure of *all* induced-C_4-free graphs; their description is much coarser, however.

We will prove that if $n^{4/3}(\log n)^4 \leqslant m \ll n^2$, then a typical member of $\mathcal{F}_{n,m}^{\mathrm{ind}}(C_4)$ is 'almost' a split graph. We will write a.a.s. (shorthand for asymptotically almost surely) as an abbreviation of "with probability tending to 1 as $n \to \infty$" and say that a graph G with n vertices and $p\binom{n}{2}$ edges is *ε-quasirandom* if every subset of more than εn vertices of G induces a subgraph with density between $(1-\varepsilon)p$ and $(1+\varepsilon)p$. We will say that a graph G is *ε-close to a split graph* if there exists a partition $V(G) = A \cup B$ such that $e_G(A) \geqslant (1-\varepsilon)\binom{|A|}{2}$ and $e_G(B) \leqslant \varepsilon e(G)$. Our first main result is the following structural description of a typical graph in $\mathcal{F}_{n,m}^{\mathrm{ind}}(C_4)$.

Theorem 1.2. *For every $\varepsilon > 0$, there exists $\delta > 0$ such that the following holds. Let G be a uniformly chosen random graph in $\mathcal{F}_{n,m}^{\mathrm{ind}}(C_4)$.*

    (a) *If $n \ll m \leqslant \delta n^{4/3}$, then a.a.s. G is ε-quasirandom.*

    (b) *If $n \ll m \leqslant \delta n^{4/3}(\log n)^{1/3}$, then a.a.s. G is not $1/4$-close to a split graph.*

    (c) *If $n^{4/3}(\log n)^4 \leqslant m \leqslant \delta n^2$, then a.a.s. G is ε-close to a split graph.*

The following result is a relatively straightforward consequence of Theorem 1.2. It determines the number of edges in (and therefore, by Theorem 1.2, the typical structure) of the random graph $G(n, p)$ conditioned on not containing an induced copy of C_4. We write $G_{n,p}^{\mathrm{ind}}(C_4)$ to denote the random graph chosen according to this conditional distribution.

Corollary 1.3. *The following bounds hold asymptotically almost surely as $n \to \infty$:*

$$
e\big(G_{n,p}^{\mathrm{ind}}(C_4)\big) = \begin{cases} \big(1 + o(1)\big)p\binom{n}{2} & \text{if } n^{-1} \ll p \ll n^{-2/3}, \\ n^{4/3}(\log n)^{O(1)} & \text{if } n^{-2/3} \leqslant p \leqslant n^{-1/3}(\log n)^4, \\ \Theta\big(p^2 n^2/\log(1/p)\big) & \text{if } p \geqslant n^{-1/3}(\log n)^4. \end{cases}
$$

Note that it follows immediately from Theorem 1.2 that $G_{n,p}^{\mathrm{ind}}(C_4)$ is a.a.s. ε-quasirandom if $n^{-1} \ll p \ll n^{-2/3}$ and a.a.s. ε-close to a split graph if $p \geqslant n^{-1/3}(\log n)^4$. We remark that we have not attempted to optimize the exponents of $\log n$, since (we believe that) our technique cannot give the correct power.

We would like to draw the reader's attention to the (somewhat surprising) fact that in the middle range $n^{-2/3+o(1)} \leqslant p \leqslant n^{-1/3+o(1)}$, the typical value of $e\big(G_{n,p}^{\mathrm{ind}}(C_4)\big)$ stays essentially constant. This is because the proportion of n-vertex graphs with m edges that are induced-C_4-free drops very sharply from $e^{-o(m)}$ to $e^{-\Omega(m \log n)}$ as m crosses a

very narrow interval around $n^{4/3}$, as shown by Theorem 1.2. A similar phenomenon has been observed in several random Turán problems for forbidden bipartite graphs (even cycles [32, 36] and complete bipartite graphs [36]) as well as Turán-type problems in additive combinatorics [15, 16]. It would be very interesting to determine whether a similar 'long flat segment' appears in the graphs of $p \mapsto e\big(G_{n,p}^{\mathrm{ind}}(H)\big)$ and $p \mapsto \mathrm{ex}\big(G(n,p), H\big)$ for every bipartite H.

Our proof of Theorem 1.2 relies on two new results: (i) an asymmetric container lemma, which generalises the main results of [10, 44], and (ii) a new robust stability theorem for induced copies of C_4 in 'pregraphs' (see below). We discuss these two ingredients in the remainder of this section.

1.2. **The asymmetric container lemma.**

The hypergraph container theorems, proved independently by Balogh, Morris, and Samotij [10] and by Saxton and Thomason [44], state (roughly speaking) that the family of independent sets of a uniform hypergraph whose edges are distributed somewhat evenly can be covered with a small number of sets, called *containers*, each of which is 'almost independent' in the sense that it contains only few edges of the hypergraph. This fact has proved to be a very convenient and useful tool in the study of the families of H-free graphs, as well as other monotone properties of graphs, hypergraphs, sets of integers, etc. There are several reasons for this. First, there is a natural correspondence between H-free graphs with a given number n of vertices and independent sets in the e_H-uniform hypergraph $\mathcal{H}$ whose vertex set is $E(K_n)$, the edge set of the complete graph with n vertices, and whose edges are the edge sets of all copies of H found in K_n. Second, classical results in extremal graph theory provide very precise and explicit descriptions of graphs with few copies of H, which correspond to the containers for independent sets of $\mathcal{H}$. Third, the bounds for the number of containers given by [10, 44] are essentially optimal, which allows one to deduce many best-possible estimates on the number of H-free graphs with given numbers of vertices and edges and describe their typical structure.

The container theorems can also be used to enumerate graphs with no *induced* copy of H. In fact, this was already done by Saxton and Thomason in their original paper [44], where they obtained (implicitly) upper bounds on $|\mathcal{F}_n^{\mathrm{ind}}(H)|$ for all H. One way to phrase this problem in the language of independent sets is to consider the hypergraph $\mathcal{H}$ whose vertex set is $E(K_n) \times \{0,1\}$ and whose edges are

- (i) all the $\binom{v_H}{2}$-element sets of the form $E \times \{1\} \cup \big(\binom{W}{2} \setminus E\big) \times \{0\}$, where W ranges over all v_H-element sets of vertices of K_n and E is the subset of $\binom{W}{2}$ covered by $E(H)$ in one of the $v_H!/|\mathrm{Aut}(H)|$ non-isomorphic embeddings of H into W and
- (ii) all the $\binom{n}{2}$ pairs $\{(e,0),(e,1)\}$, where e ranges over all edges of K_n.

One can see that n-vertex graphs with no induced copy of H are in a natural one-to-one correspondence with the independent sets of $\mathcal{H}$ with $\binom{n}{2}$ elements. Even though the container theorems may be applied only to uniform hypergraphs, since one is usually interested in upper bounds, one may disregard the 2-uniform edges of type (ii) and construct containers for independent sets of the resulting smaller $\binom{v_H}{2}$-uniform hypergraph, which clearly include all independent sets of the original hypergraph.

One soon realises that the above approach is somewhat flawed when one is interested in the family $\mathcal{F}_{n,m}^{\mathrm{ind}}(H)$ whenever m is either very small or very close to $\binom{n}{2}$ and H is neither complete nor empty. This is because the original container theorems completely

disregard the obvious asymmetry between the edges and the non-edges of H in each of the $\binom{v_H}{2}$-uniform edges of $\mathcal{H}$. As a result, one cannot expect to deduce optimal bounds on $|\mathcal{F}_{n,m}^{\mathrm{ind}}(H)|$ for all m using this approach. Our main motivation for this work is to address this issue.

Departing somewhat from the language of independent sets, we shall regard a graph $G \subseteq K_n$ as the characteristic function $h_G \colon E(K_n) \to \{0,1\}$ of its edge set; that is, $h_G(e) = 1$ if $e \in E(G)$ and $h_G(e) = 0$ otherwise. The family $\mathcal{F}_n^{\mathrm{ind}}(H)$, viewed as a set of functions $h \colon E(K_n) \to \{0,1\}$, may be described by a set of *constraints* of the form

$$\neg\left(h|_E \equiv 1 \wedge h|_{\binom{W}{2}\setminus E} \equiv 0\right).$$

In other words, a function $h \in \mathcal{F}_n^{\mathrm{ind}}(H)$ cannot simultaneously map all elements of E to 1 and all elements of $\binom{W}{2} \setminus E$ to 0, for any $W \subset V(K_n)$ with $|W| = v_H$ and any E that is the edge set of an embedding of H into W.

There is nothing special here about the family $\mathcal{F}_n^{\mathrm{ind}}(H)$ or the set $E(K_n)$. Therefore, for the remainder of this discussion, we shall replace $E(K_n)$ with an arbitrary finite set V, let $\mathcal{H}$ be an arbitrary family of pairs of disjoint subsets of V, and let

$$\mathcal{F}(\mathcal{H}) = \left\{h \in \{0,1\}^V : \neg(h|_{A_0} \equiv 0 \wedge h|_{A_1} \equiv 1) \text{ for all } (A_0, A_1) \in \mathcal{H}\right\}.$$

In other words, one obtains the family $\mathcal{F}(\mathcal{H})$ from $\{0,1\}^V$ by discarding all $h \colon V \to \{0,1\}$ that map each element of A_0 to 0 and each element of A_1 to 1 for some pair $(A_0, A_1) \in \mathcal{H}$. We shall informally refer to these pairs of sets as *constraints* and say that h *violates* (resp. *satisfies*) a constraint (A_0, A_1) if h maps (resp. does not map) each element of A_0 to 0 and each element of A_1 to 1. Finally, let us note here for future reference that according to the above definition, $\mathcal{F}(\mathcal{H})$ is empty whenever $\mathcal{H}$ contains the pair $(\emptyset, \emptyset)$; in other words, every function violates the 'empty' constraint $(\emptyset, \emptyset)$.

The container theorems imply that if such a family $\mathcal{H}$ contains only pairs (A_0, A_1) with a given value of $|A_0| + |A_1|$ and the sets $A_0 \cup A_1$ are distributed somewhat uniformly, then there is a small family $\mathcal{C}$ of partitions $V = V_0 \cup V_1 \cup V_*$ such that

$$\mathcal{F}(\mathcal{H}) \subseteq \bigcup_{(V_0, V_1, V_*) \in \mathcal{C}} \{0\}^{V_0} \times \{1\}^{V_1} \times \{0,1\}^{V_*}$$

and, importantly, every function in each of the cylinders $\{0\}^{V_0} \times \{1\}^{V_1} \times \{0,1\}^{V_*}$ violates only few constraints in $\mathcal{H}$. In particular, one does not allow a trivial covering of $\mathcal{F}(\mathcal{H})$ with $\{0,1\}^V$, which corresponds to $V_* = V$. Roughly speaking, we might say that $\mathcal{F}(\mathcal{H})$ may be 'tightly' covered by a small family of cylinders.

In this work, we take a refined approach to this covering problem. We shall build families of containers that are tailored to the subfamily of all $h \in \mathcal{F}(\mathcal{H})$ that attain the values 0 and 1 given numbers of times, unlike in previous works. More precisely, for each integer m with $0 \leqslant m \leqslant |V|$, we shall consider the subfamily $\mathcal{F}_m(\mathcal{H}) \subseteq \mathcal{F}(\mathcal{H})$ defined by

$$\mathcal{F}_m(\mathcal{H}) = \left\{h \in \mathcal{F}(\mathcal{H}) : |h^{-1}(1)| = m\right\}$$

and build a family of containers for the elements of $\mathcal{F}_m(\mathcal{H})$ only.

We shall focus our attention on families $\mathcal{F}(\mathcal{H})$ determined by collections $\mathcal{H}$ of constraints that are uniform in the sense that each $(A_0, A_1) \in \mathcal{H}$ satisfies $|A_0| = k_0$ and $|A_1| = k_1$ for some fixed integers k_0 and k_1. We shall refer to such collections $\mathcal{H}$ as (k_0, k_1)-*uniform hypergraphs*. In standard applications of the container method this should not be a huge restriction, provided that we are only interested in constraints of bounded size,

that is, pairs (A_0, A_1) where $|A_0| + |A_1|$ is bounded from above by a constant. Indeed, given a non-uniform family of constraints of bounded size, we may restrict our attention to the 'densest' (k_0, k_1)-uniform hypergraph that is contained in the family, losing only some constant factors. In fact, this is precisely what we are going to do in our proof of Theorem 1.2.

For a (k_0, k_1)-uniform hypergraph $\mathcal{H}$ and two disjoint sets T_0 and T_1, we define

$$\deg_{\mathcal{H}}(T_0, T_1) = |\{(A_0, A_1) \in \mathcal{H} : T_0 \subseteq A_0 \text{ and } T_1 \subseteq A_1\}|.$$

Furthermore, for each pair of integers (ℓ_0, ℓ_1), we let

$$\Delta_{(\ell_0, \ell_1)}(\mathcal{H}) = \max\left\{\deg_{\mathcal{H}}(T_0, T_1) : T_0, T_1 \subseteq V \text{ with } |T_0| = \ell_0 \text{ and } |T_1| = \ell_1\right\}.$$

Abusing notation somewhat, we shall identify a partition $V = V_0 \cup V_1 \cup V_*$ with the cylinder $\{0\}^{V_0} \times \{1\}^{V_1} \times \{0,1\}^{V_*}$ and the function $a \colon V(\mathcal{H}) \to \{0, 1, *\}$ defined by $a^{-1}(x) = V_x$ for each $x \in \{0, 1, *\}$. In particular, a function $h \colon V(\mathcal{H}) \to \{0, 1\}$ belongs to the cylinder $a \colon V(\mathcal{H}) \to \{0, 1, *\}$ if $h(v) = a(v)$ for all $v \in V(\mathcal{H})$ such that $a(v) \neq *$. In other words, $h(v)$ is forced to equal $a(v)$ unless $a(v) = *$, in which case $h(v)$ can be either 0 or 1.

We are now ready to state the main result of this section, an asymmetric container theorem. In the statement of the theorem, $\mathcal{F}_{\leqslant m}(\mathcal{H})$ is a shorthand for $\bigcup_{m'=0}^{m} \mathcal{F}_{m'}(\mathcal{H})$.

Theorem 1.4. *For all integers $k_0, k_1 \geqslant 0$, not both zero, and each $K > 0$, the following holds. Suppose that $\mathcal{H}$ is a non-empty (k_0, k_1)-uniform hypergraph and b, m, and r are integers satisfying*

$$\Delta_{(\ell_0, \ell_1)}(\mathcal{H}) \leqslant K \cdot \frac{b^{\ell_0 + \ell_1 - 1}}{m^{\ell_0} \cdot v(\mathcal{H})^{\ell_1}} \cdot e(\mathcal{H}) \cdot \left(\frac{m}{r}\right)^{\mathbb{1}[\ell_0 > 0]} \tag{1}$$

for every pair $(\ell_0, \ell_1) \in \{0, \dots, k_0\} \times \{0, \dots, k_1\}$ with $(\ell_0, \ell_1) \neq (0, 0)$. Then there exist a family $\mathcal{S} \subseteq \binom{V(\mathcal{H})}{\leqslant k_0 b} \times \binom{V(\mathcal{H})}{\leqslant k_1 b}$ and functions $f \colon \mathcal{S} \to \{0, 1, *\}^{V(\mathcal{H})}$ and $g \colon \mathcal{F}_{\leqslant m}(\mathcal{H}) \to \mathcal{S}$ such that, letting $\delta = 2^{-(k_0 + k_1)(k_0 + k_1 + 1)} K^{-1}$:

(a) *Every $h \in \mathcal{F}_{\leqslant m}(\mathcal{H})$ belongs to the cylinder $f(g(h))$.*

(b) *Either $|f(S)^{-1}(0)| \geqslant \delta v(\mathcal{H})$ or $|f(S)^{-1}(1)| \geqslant \delta r$ for every $S \in \mathcal{S}$; moreover, the former can hold only if $k_1 > 0$ and the latter can hold only if $k_0 > 0$.*

(c) *If $g(h) = (S_0, S_1)$ for some $h \in \mathcal{F}_{\leqslant m}(\mathcal{H})$, then $S_0 \subseteq h^{-1}(0)$ and $S_1 \subseteq h^{-1}(1)$.*

Yet another rephrasing of condition (a) is that whenever $g(h) = S$, then h is forced to take the value 0 on $f(S)^{-1}(0)$ and it is forced to take the value 1 on $f(S)^{-1}(1)$. Note the asymmetry between the guaranteed lower bounds on the cardinalities of the sets $f(S)^{-1}(0)$ and $f(S)^{-1}(1)$ in (b). Roughly speaking, we are equally satisfied with (i) containers forcing our function h to take the value 0 on a positive proportion of $V(\mathcal{H})$ and (ii) containers forcing our function to take the value 1 only on some δr elements of $V(\mathcal{H})$. Condition (c) states that for every $h \in \mathcal{F}_{\leqslant m}(\mathcal{H})$, the value of $g(h)$ is 'consistent' with h. This additional property of the function g will not be used in our application of the theorem to enumerating $\mathcal{F}_{n,m}^{\text{ind}}(C_4)$. However, we state it here as the analogous property in the original container theorems was crucial in avoiding superfluous logarithmic factors in many applications of the container method. Finally, let us point out here that we shall be allowing all of our hypergraphs to contain edges with multiplicities greater than one. In particular, both $e(\cdot)$ and $\deg_{\mathcal{H}}(\cdot, \cdot)$ count edges with their multiplicities.

A reader who is familiar with the container method might notice that by setting $r = m = v(\mathcal{H})$ in Theorem 1.4, one recovers the statement of the original container theorem [10, Proposition 3.1] in the somewhat more general context of (k_0, k_1)-uniform hypergraphs. To illustrate the 'asymmetry' in Theorem 1.4, we need to assume that $m \ll v(\mathcal{H})$. For brevity, let $N = v(\mathcal{H})$ and consider two cylinders, described by the following two partitions of $V(\mathcal{H})$:

(i) $V(\mathcal{H}) = V_0 \cup V_1 \cup V_*$, where $|V_0| = \delta N$ and $V_1 = \emptyset$,
(ii) $V(\mathcal{H}) = V_0' \cup V_1' \cup V_*'$, where $V_0' = \emptyset$ and $|V_1'| = \delta r$.

Observe that the cylinder described in (i) contains at most $\binom{(1-\delta)N}{m}$ functions from $\mathcal{F}_m(\mathcal{H})$, whereas the cylinder described in (ii) contains at most $\binom{N-\delta r}{m-\delta r}$ functions from $\mathcal{F}_m(\mathcal{H})$. Assume that $r \ll m \ll N$. Since

$$\binom{(1-\delta)N}{m} \approx (1-\delta)^m \cdot \binom{N}{m} \qquad \text{and} \qquad \binom{N-\delta r}{m-\delta r} \approx \left(\frac{m}{N}\right)^{\delta r} \binom{N}{m},$$

then both cylinders will have equal $volume^3$ when $r \approx m/\log(N/m) \ll m$. On the other hand, when $r \ll m \ll N$, then the assumptions on the maximum degrees of $\mathcal{H}$ stated in (1) are weaker by a factor of $\left(\frac{N}{m}\right)^{\ell_0} \cdot \left(\frac{m}{r}\right)^{\mathbb{1}[\ell_0 > 0]}$ when compared to the original container theorems, see [10, Proposition 3.1]. This allows one to choose a smaller b, which results in a smaller family of containers.

As we believe that having a trade-off between the upper bound on the size of containers for independent sets in a hypergraph $\mathcal{H}$ and the upper bounds on maximum degrees $\Delta_\ell(\mathcal{H})$ can be useful in other applications of the container method, we conclude this section with a sharpening of the original container theorems, [10, Proposition 3.1] and [44, Theorem 3.4], that follows easily from Theorem 1.4. We write $\mathcal{I}(\mathcal{H})$ for the family of independent sets of $\mathcal{H}$ and $\Delta_\ell(\mathcal{H})$ for the largest number of edges of $\mathcal{H}$ that contain a particular ℓ-element subset of $V(\mathcal{H})$.

Theorem 1.5. *Suppose that positive integers b, k, and r and a non-empty k-uniform hypergraph $\mathcal{H}$ satisfy*

$$\Delta_\ell(\mathcal{H}) \leqslant \left(\frac{b}{v(\mathcal{H})}\right)^{\ell-1} \frac{e(\mathcal{H})}{r} \tag{2}$$

for every $\ell \in \{1, \ldots, k\}$. Then there exist a family $\mathcal{S} \subseteq \binom{V(\mathcal{H})}{\leqslant kb}$ and functions $f \colon \mathcal{S} \to \mathcal{P}(V(\mathcal{H}))$ and $g \colon \mathcal{I}(\mathcal{H}) \to \mathcal{S}$ such that for every $I \in \mathcal{I}(\mathcal{H})$,

$$g(I) \subseteq I \subseteq f(g(I)) \qquad and \qquad |f(g(I))| \leqslant v(\mathcal{H}) - \delta r,$$

where $\delta = 2^{-k(k+1)}$.

To obtain Theorem 1.5, we simply apply Theorem 1.4 to the $(0, k)$-uniform hypergraph with the same vertex set as $\mathcal{H}$ whose edges are all pairs $(\emptyset, A)$ such that A is an edge of $\mathcal{H}$. We shall spell out a few more details at the end of Section 2.

---

^{3}By the volume of a cylinder $\{0\}^{V_0} \times \{1\}^{V_1} \times \{0, 1\}^{V_*}$, we will mean the number of functions $h \colon V(\mathcal{H}) \to \{0, 1\}$, with $|h^{-1}(1)| = m$, that are contained in the cylinder, that is, $\binom{|V_*|}{m-|V_1|}$.

1.3. **Robust balanced stability for induced C_4s.** In order to determine the structure of a typical graph in $\mathcal{F}_{n,m}^{\mathrm{ind}}(C_4)$ using the container method, we ought to characterise all containers whose volume is (close to) the largest possible. Our containers for $\mathcal{F}_{n,m}^{\mathrm{ind}}(C_4)$ will be cylinders in $\{0,1\}^{E(K_n)}$ that correspond to partitions $E(K_n) = E_0 \cup E_1 \cup E_*$ with the following property: There are only few 4-vertex subsets $\{v_1, v_2, v_3, v_4\}$ such that $v_1 v_2, v_3 v_4 \in E_0 \cup E_*$ and $v_1 v_3, v_1 v_4, v_2 v_3, v_2 v_4 \in E_1 \cup E_*$. Each such set $\{v_1, v_2, v_3, v_4\}$ induces a copy of C_4 in some graph described by the partition $E(K_n) = E_0 \cup E_1 \cup E_*$.[4] Since we are interested only in graphs with exactly m edges, the volume of a container is simply the number of graphs with m edges that this cylinder contains, that is, $\binom{|E_*|}{m-|E_1|}$. The precise statements of our results are rather technical, but roughly speaking we show that each container whose volume is close to largest possible has the following structure: the graph E_1 contains an 'almost-complete' graph with vertex set W, and most edges in E_* have an endpoint in W.

To avoid excessive use of indices, we shall view partitions of $E(K_n)$ of the above type as partial two-colourings of the edges of K_n that we shall call pregraphs. More precisely, by a *pregraph* $\mathcal{P}$ of $E(K_n)$ we will mean a pair (M, E) of disjoint subsets of $E(K_n)$. We shall refer to the elements of the set E as *edges* and the elements of the set M as *mixed edges*.[5] A *good copy* of C_4 in $\mathcal{P}$ is a copy of C_4 in M whose vertex set is independent in E. Note that each good copy of C_4 corresponds to a set $\{v_1, v_2, v_3, v_4\}$ described in the previous paragraph (but not vice-versa). This means, in particular, that the pregraph corresponding to each container contains only few good copies of C_4. We shall therefore restrict our attention to characterising pregraphs with few good copies of C_4. As we will later see, a sufficiently precise and useful characterisation of containers can be derived from a robust stability theorem for pregraphs, which we state here in an abbreviated form; for the full statement, we refer the reader to Section 3. We will say that a graph G is ε-close to K_ℓ if one can transform G into K_ℓ by adding or deleting at most $\varepsilon\binom{\ell}{2}$ edges.

Theorem 1.6. *For every $\varepsilon > 0$ there exist positive constants C, δ, and β such that the following holds for all integers ℓ and n with $\ell \geqslant C\sqrt{n}$. Let $\mathcal{P} = (M, E)$ be a pregraph on n vertices with*

$$|E| \leqslant \binom{\ell}{2} \qquad \text{and} \qquad |M| \geqslant (1-\delta)\ell n.$$

Then either E is ε-close to K_ℓ or $\mathcal{P}$ contains at least $\beta\ell^4$ good copies of C_4.

Observe that Theorem 1.6 provides a structural characterisation of all those pregraphs (M, E) on n vertices with $|E| \leqslant \binom{\ell}{2}$ and $|M| \geqslant (1 - o(1))\ell n$ for some $\ell \gg \sqrt{n}$ that contain only $o(\ell^4)$ good copies of C_4. For each such pregraph (M, E), there is a set U of ℓ vertices on which E is almost complete. Moreover, all but $o(\ell n)$ mixed edges have an endpoint in U. Indeed, if some $\Omega(\ell n)$ mixed edges did not have an endpoint in U, then Theorem 1.6 applied to the pregraph induced by the complement of U would produce $\Omega(\ell^4)$ good copies of C_4.

1.4. **Organisation of the paper.** The rest of the paper is organised as follows. In Section 2 we prove the asymmetric container lemma, in Section 3 we prove Theorem 1.6, in Section 4 we prove the lower bounds in Theorem 1.2, and in Section 5 we complete

---

[4]These are all G such that $E_1 \subseteq E(G) \subseteq E_1 \cup E_*$ and $E_0 \subseteq E(K_n) \setminus E(G) \subseteq E_0 \cup E_*$.

[5]The sets E and M correspond to the sets E_1 and E_* above, respectively.

the proof of Theorem 1.2 and Corollary 1.3. Finally, in Section 6 we discuss some open questions and further applications of the asymmetric container lemma.

2. The proof of the asymmetric container lemma

2.1. **Proof outline.** Our proof of Theorem 1.4 follows the general strategy of [10]. Namely, we construct a function $f^*: \mathcal{F}(\mathcal{H}) \to \{0, 1, *\}^{V(\mathcal{H})}$ that satisfies the following two conditions for every $h \in \mathcal{F}(\mathcal{H})$. Writing f_h^* as a shorthand for $f^*(h)$, the two conditions are:

  (a) h belongs to the cylinder f_h^*,
  (b) $|(f_h^*)^{-1}(0)| \geqslant \delta v(\mathcal{H})$ or $|(f_h^*)^{-1}(1)| \geqslant \delta r$,

cf. (a) and (b) in the statement of Theorem 1.4. Crucially, the function f^* takes only at most $\binom{v(\mathcal{H})}{\leqslant k_0 b} \cdot \binom{v(\mathcal{H})}{\leqslant k_1 b}$ different values. This last property is a simple consequence of the fact that the algorithmic construction of f^* can be encoded as a sequence of decisions that naturally correspond to a pair of subsets of $V(\mathcal{H})$ containing at most $k_0 b$ and $k_1 b$ elements, respectively. In particular, we shall obtain an implicit decomposition $f^* = f \circ g$ promised in Theorem 1.4.

The function f^* is constructed by an algorithm that operates in a sequence of at most $k_0 + k_1 - 1$ rounds. At the beginning of each round, we are given an (i_0, i_1)-uniform hypergraph $\mathcal{G}$ with the same vertex set as $\mathcal{H}$ and such that $h \in \mathcal{F}(\mathcal{G})$; at the beginning of the first round, $(i_0, i_1) = (k_0, k_1)$ and $\mathcal{G} = \mathcal{H}$. We let $(i_0', i_1') = (i_0, i_1 - 1)$ if $i_1 > 0$ and let $(i_0', i_1') = (i_0 - 1, i_1) = (i_0 - 1, 0)$ otherwise. By the end of the round, we will have either (i) defined a function $f_h^*: V(\mathcal{H}) \to \{0, 1, *\}$ satisfying both (a) and (b) above, or (ii) constructed an (i_0', i_1')-uniform hypergraph $\mathcal{G}^*$ with $V(\mathcal{G}^*) = V(\mathcal{H})$ and such that $h \in \mathcal{F}(\mathcal{G}^*)$ whose maximum degrees satisfy conditions akin to the conditions on the maximum degrees of $\mathcal{H}$ given by (1). This is achieved in the following way.

We start with $\mathcal{G}^*$ empty and $f_h^* \equiv *$. We set $c = 1$ if $i_1 > 0$ and $c = 0$ otherwise, so $i_c' = i_c - 1$. Our algorithm considers a sequence of questions of the form "Is $h(v) = c$?" for some carefully chosen (sequence of) vertices $v \in V(\mathcal{H})$. If the answer is YES, then we set $f_h^*(v) = c$ and, more importantly, we add new (i_0', i_1')-uniform constraints to $\mathcal{G}^*$ in the following way. As $h(v) = c$, if h satisfies a constraint[6] (A_0, A_1) with $v \in A_c$, then it also satisfies the constraint (A_0', A_1') defined by $A_c' = A_c \setminus \{v\}$ and $A_{1-c}' = A_{1-c}$. In view of this, for each $(A_0, A_1) \in \mathcal{G}$ with $v \in A_c$, we add to $\mathcal{G}^*$ the corresponding (A_0', A_1'). If the answer is NO, then we only set $f_h^*(v) = 1 - c$. (We thus choose to ignore all the constraints $(A_0, A_1) \in \mathcal{G}$ such that $v \in A_{1-c}$.) The round ends when either the number of YES answers reaches b or if no constraints remain involving only the vertices that we have not yet asked about. Our assumptions on the maximum degrees of the hypergraph $\mathcal{G}$ imply that in the latter case, the number of NO answers will be sufficiently large to deduce that $|(f_h^*)^{-1}(1-c)|$ is sufficiently large (that is, at least $\delta v(\mathcal{H})$ if $c = 1$ and at least δr if $c = 0$). If this does not happen (and hence the number of YES answers reaches b), then we shall be able to show that the hypergraph $\mathcal{G}^*$, which we have created based on the YES answers, contains a subhypergraph with sufficiently many edges, whose maximum

---

[6]Recall from Section 1.2 that h satisfies the constraint (A_0, A_1) if an only if f does *not* simultaneously take only the value 0 on A_0 and only the value 1 on A_1; equivalently, f either takes the value 1 on some element of A_0 or the value 0 on some element of A_1.

degrees satisfy the required conditions. In this case, we let $\mathcal{G} \leftarrow \mathcal{G}^*$ and $(i_0, i_1) \leftarrow (i_0', i_1')$ and proceed to the next round.

Since, as noted before, no function satisfies the empty constraint $(\emptyset, \emptyset)$, it follows that in the round when $i_0 + i_1 = 1$, no YES answers can be given. (Otherwise, a non-empty $(0,0)$-uniform hypergraph $\mathcal{G}^*$ with $h \in \mathcal{F}(\mathcal{G}^*)$ would be constructed.) In particular, the function f_h^* will have to be defined in this round, provided that the algorithm reaches it.

Even though the sequence of values of c that we choose (i.e., we let $c = 1$ as long as i_1 is not yet zero) may seem somewhat arbitrary, it has a very important consequence. Namely, if $\mathcal{G}$ is an $(i_0, 0)$-uniform hypergraph with $V(\mathcal{G}) = V(\mathcal{H})$ and $h \in \mathcal{F}_{\leqslant m}(\mathcal{G})$, then there must be a vertex $v \in V(\mathcal{G})$ such that $\deg_{\mathcal{G}}(\{v\}, \emptyset) \geqslant e(\mathcal{G})/m$. Indeed, the set $h^{-1}(1)$ has at most m elements and it has to intersect A_0 for each $(A_0, \emptyset) \in \mathcal{G}$. Note that if $m \ll v(\mathcal{H})$, then $e(\mathcal{G})/m$ is much larger than the average degree of $\mathcal{G}$. This simple observation is the reason why restricting to the family $\mathcal{F}_{\leqslant m}(\mathcal{H})$ allows us to create a smaller family of containers.

Finally, since each of the questions asked by the algorithm is a YES/NO question, we may encode the execution of the algorithm, and thus also the function f^*, as a set of at most $(k_0 + k_1 - 1) \cdot b$ vertices for which the answer was YES.

We conclude this outline with an important technical remark. Throughout this section we allow all of our hypergraphs to contain edges with multiplicities greater than one. Moreover, when computing various degrees $\deg(\cdot, \cdot)$ or cardinalities $e(\cdot)$ of the edge sets of various hypergraphs, we shall always count edges with multiplicities. As first discovered by Saxton and Thomason in [44] and later reiterated in [10], this seemingly insignificant detail has far-reaching consequences in both the statement and the proof of the container theorems.

2.2. **Setup.** Let k_0 and k_1 be nonnegative integers and let K be a positive real. Let b, m, and r be positive integers and suppose that $\mathcal{H}$ is a (k_0, k_1)-uniform hypergraph satisfying (1) for every pair (ℓ_0, ℓ_1) as in the statement of Theorem 1.4. We claim that without loss of generality we may assume that $b \leqslant m \leqslant v(\mathcal{H})$. Indeed, if $m > v(\mathcal{H})$, then we may replace m with $v(\mathcal{H})$ as $\mathcal{F}_{\leqslant m} \subseteq \mathcal{F}(\mathcal{H}) = \mathcal{F}_{\leqslant v(\mathcal{H})}(\mathcal{H})$ and the right-hand side of (1) is a non-increasing function of m. If $b > v(\mathcal{H}) \geqslant m$, then we may replace b with $v(\mathcal{H})$. This is because $\binom{V(\mathcal{H})}{\leqslant k_i b} = \binom{V(\mathcal{H})}{\leqslant k_i v(\mathcal{H})}$ and the assumed upper bounds on the maximum degrees of $\mathcal{H}$ remain true even after we replace b with $v(\mathcal{H})$. Indeed, if $\ell_0 > 0$, then for every $\ell_1 \in \{0, \ldots, k_1\}$,

$$\Delta_{(\ell_0, \ell_1)}(\mathcal{H}) \leqslant \Delta_{(1,0)}(\mathcal{H}) \leqslant K \cdot \frac{e(\mathcal{H})}{r} \leqslant K \cdot \frac{v(\mathcal{H})^{\ell_0 + \ell_1 - 1}}{r \cdot m^{\ell_0 - 1} \cdot v(\mathcal{H})^{\ell_1}} \cdot e(\mathcal{H}),$$

as $v(\mathcal{H}) \geqslant m$, and if $\ell_0 = 0$, then for every $\ell_1 \in \{1, \ldots, k_1\}$,

$$\Delta_{(0, \ell_1)}(\mathcal{H}) \leqslant \Delta_{(0,1)}(\mathcal{H}) \leqslant K \cdot \frac{e(\mathcal{H})}{v(\mathcal{H})} = K \cdot \frac{v(\mathcal{H})^{\ell_1 - 1}}{v(\mathcal{H})^{\ell_1}} \cdot e(\mathcal{H}).$$

Finally, if $v(\mathcal{H}) \geqslant b > m$, then we may replace m with b, since $\mathcal{F}_{\leqslant m}(\mathcal{H}) \subseteq \mathcal{F}_{\leqslant b}(\mathcal{H})$, the bound on $\Delta_{(0, \ell_1)}(\mathcal{H})$ in (1) does not depend on m, and if $\ell_0 > 0$, then

$$\Delta_{(\ell_0, \ell_1)}(\mathcal{H}) \leqslant \Delta_{(1, \ell_1)}(\mathcal{H}) \leqslant K \cdot \frac{b^{\ell_1}}{r \cdot v(\mathcal{H})^{\ell_1}} \cdot e(\mathcal{H}) = K \cdot \frac{b^{\ell_0 + \ell_1 - 1}}{r \cdot b^{\ell_0 - 1} \cdot v(\mathcal{H})^{\ell_1}} \cdot e(\mathcal{H})$$

for every $\ell_1 \in \{0, \ldots, k_1\}$.

11

We shall be working only with hypergraphs whose uniformities come from the set

$$\mathcal{U} := \big\{(1,0), (2,0), \ldots, (k_0,0), (k_0,1), \ldots, (k_0, k_1)\big\}.$$

We now define a collection of numbers that will be upper bounds on the maximum degrees of the hypergraphs constructed by our algorithm. To be more precise, for each $(i_0, i_1) \in \mathcal{U}$ and all (ℓ_0, ℓ_1), we shall force the maximum (ℓ_0, ℓ_1)-degree of the (i_0, i_1)-uniform hypergraph not to exceed the quantity $\Delta^{(i_0,i_1)}_{(\ell_0,\ell_1)}$, defined as follows.

Definition 2.1. For every $(i_0, i_1) \in \mathcal{U}$ and every $(\ell_0, \ell_1) \in \{0, \ldots, i_0\} \times \{0, \ldots, i_1\}$ with $(\ell_0, \ell_1) \neq (0, 0)$, we define the number $\Delta^{(i_0,i_1)}_{(\ell_0,\ell_1)}$ using the following recursion:

(1) Set $\Delta^{(k_0,k_1)}_{(\ell_0,\ell_1)} := \Delta_{(\ell_0,\ell_1)}(\mathcal{H})$ for all $(\ell_0, \ell_1) \in \{0, \ldots, k_0\} \times \{0, \ldots, k_1\} \setminus \{(0,0)\}$.

(2) If $i_0 = k_0$ and $0 \leqslant i_1 < k_1$, then

$$\Delta^{(i_0,i_1)}_{(\ell_0,\ell_1)} := \max\left\{2 \cdot \Delta^{(i_0,i_1+1)}_{(\ell_0,\ell_1+1)},\ \frac{b}{v(\mathcal{H})} \cdot \Delta^{(i_0,i_1+1)}_{(\ell_0,\ell_1)}\right\}.$$

(3) If $0 < i_0 < k_0$ and $i_1 = 0$, then

$$\Delta^{(i_0,i_1)}_{(\ell_0,\ell_1)} := \max\left\{2 \cdot \Delta^{(i_0+1,i_1)}_{(\ell_0+1,\ell_1)},\ \frac{b}{m} \cdot \Delta^{(i_0+1,i_1)}_{(\ell_0,\ell_1)}\right\}.$$

The above recursive definition will be convenient in some parts of our analysis. In other parts, we shall require the following explicit formula for $\Delta^{(i_0,i_1)}_{(\ell_0,\ell_1)}$, which one easily derives from Definition 2.1 using a straightforward induction on $k_0 + k_1 - i_0 - i_1$.

Observation 2.2. *For all i_0, i_1, ℓ_0, and ℓ_1 as in Definition 2.1,*

$$\Delta^{(i_0,i_1)}_{(\ell_0,\ell_1)} = \max\left\{2^{d_0+d_1}\left(\frac{b}{v(\mathcal{H})}\right)^{k_1-i_1-d_1}\left(\frac{b}{m}\right)^{k_0-i_0-d_0} \Delta_{(\ell_0+d_0,\ell_1+d_1)}(\mathcal{H}) : 0 \leqslant d_j \leqslant k_j - i_j\right\}.$$

For future reference, we note the following two simple corollaries of Observation 2.2 and our assumptions on the maximum degrees of $\mathcal{H}$, see (1). Suppose that $(i_0, i_1) \in \mathcal{U}$. If $i_1 > 0$, then necessarily $i_0 = k_0$ and hence,

$$
\begin{aligned}
\Delta^{(i_0,i_1)}_{(0,1)} &\leqslant \max\left\{2^{d_1}\left(\frac{b}{v(\mathcal{H})}\right)^{k_1-i_1-d_1} K \cdot \frac{b^{d_1}}{v(\mathcal{H})^{d_1+1}} \cdot e(\mathcal{H}) : 0 \leqslant d_1 \leqslant k_1 - i_1\right\} \\
&\leqslant 2^{k_1} K \left(\frac{b}{v(\mathcal{H})}\right)^{k_1-i_1} \frac{e(\mathcal{H})}{v(\mathcal{H})} = 2^{k_1} K \left(\frac{b}{v(\mathcal{H})}\right)^{k_1-i_1}\left(\frac{b}{m}\right)^{k_0-i_0} \frac{e(\mathcal{H})}{v(\mathcal{H})}.
\end{aligned}
\tag{3}
$$

Moreover, if $i_0 > 0$ then

$$
\begin{aligned}
\Delta^{(i_0,i_1)}_{(1,0)} &\leqslant \max\left\{2^{d_0+d_1}\left(\frac{b}{v(\mathcal{H})}\right)^{k_1-i_1-d_1}\left(\frac{b}{m}\right)^{k_0-i_0-d_0} K \cdot \frac{b^{d_0+d_1}}{m^{d_0} \cdot v(\mathcal{H})^{d_1}} \cdot \frac{e(\mathcal{H})}{r}\right\} \\
&\leqslant 2^{k_0+k_1} K \left(\frac{b}{v(\mathcal{H})}\right)^{k_1-i_1}\left(\frac{b}{m}\right)^{k_0-i_0} \frac{e(\mathcal{H})}{r},
\end{aligned}
\tag{4}
$$

where the maximum is over all pairs (d_0, d_1) of integers satisfying $0 \leqslant d_j \leqslant k_j - i_j$.

Definition 2.3. Given $(i_0, i_1) \in \mathcal{U}$, $(\ell_0, \ell_1) \in \{0, \ldots, i_0\} \times \{0, \ldots, i_1\}$ with $(\ell_0, \ell_1) \neq (0, 0)$, and an (i_0, i_1)-uniform hypergraph $\mathcal{G}$, we define

$$M^{(i_0,i_1)}_{(\ell_0,\ell_1)}(\mathcal{G}) = \left\{(T_0, T_1) \in \binom{V(\mathcal{G})}{\ell_0} \times \binom{V(\mathcal{G})}{\ell_1} : \deg_{\mathcal{G}}(T_0, T_1) \geqslant \frac{1}{2} \cdot \Delta^{(i_0,i_1)}_{(\ell_0,\ell_1)}\right\}.$$

Finally, let us say that $c \in \{0,1\}$ is *compatible* with $(i_0, i_1) \in \mathcal{U}$ if the unique pair $(i_0', i_1') \in \mathcal{U} \cup \{(0,0)\}$ with $i_0' + i_1' = i_0 + i_1 - 1$ satisfies $i_c' = i_c - 1$ (and $i_{1-c}' = i_{1-c}$). By the definition of $\mathcal{U}$, it follows that 1 is compatible with $(i_0, i_1) \in \mathcal{U}$ if and only if $i_1 > 0$.

2.3. **The algorithm.** We shall now define precisely a single round of the algorithm that we described informally in Section 2.1. To this end, fix some $(i_0, i_1) \in \mathcal{U}$ and a compatible $c \in \{0,1\}$ and (as in the definition of a compatible c) set

$$i_c' = i_c - 1 \qquad \text{and} \qquad i_{1-c}' = i_{1-c}. \tag{5}$$

Suppose that $\mathcal{G}$ is an (i_0, i_1)-uniform hypergraph with $V(\mathcal{G}) = V(\mathcal{H})$. A single round of the algorithm takes as input an arbitrary $h \in \mathcal{F}(\mathcal{G})$ and outputs an (i_0', i_1')-uniform hypergraph $\mathcal{G}^*$ satisfying $V(\mathcal{G}^*) = V(\mathcal{G})$ and $h \in \mathcal{F}(\mathcal{G}^*)$ as well as a set of vertices of $\mathcal{G}$ on which h takes the value c at most b times. Crucially, the number of possible outputs of the algorithm (over all possible input functions $h \in \mathcal{F}(\mathcal{G})$) is at most $\binom{v(\mathcal{H})}{\leqslant b}$.

Assume that there is an implicit linear order $\preccurlyeq$ on $V(\mathcal{G})$. The *c-maximum vertex* of a hypergraph $\mathcal{A}$ with $V(\mathcal{A}) = V(\mathcal{G})$ is the $\preccurlyeq$-smallest vertex among those v that maximise $|\{(A_0, A_1) \in \mathcal{A} : v \in A_c\}|$.

The algorithm. Set $\mathcal{A}^{(0)} := \mathcal{G}$, let S be the empty set, and let $\mathcal{G}_*^{(0)}$ be the empty (i_0', i_1')-uniform hypergraph on $V(\mathcal{G})$. Do the following for each integer $j \geqslant 0$ in turn:

(S1) If $|S| = b$ or $\mathcal{A}^{(j)}$ is empty, then set $J := j$ and STOP.

(S2) Let v_j be the c-maximum vertex of $\mathcal{A}^{(j)}$.

(S3) If $h(v_j) = c$, then add j to the set S and let

$$\mathcal{G}_*^{(j+1)} := \mathcal{G}_*^{(j)} \cup \left\{ \big(A_0 \setminus \{v_j\}, A_1 \setminus \{v_j\}\big) : (A_0, A_1) \in \mathcal{A}^{(j)} \text{ and } v_j \in A_c \right\}.$$

(S4) Let $\mathcal{A}^{(j+1)}$ be the hypergraph obtained from $\mathcal{A}^{(j)}$ by removing from it all pairs (A_0, A_1) such that either of the following hold:

(a) $v_j \in A_c$;

(b) there exist $T_0 \subseteq A_0$ and $T_1 \subseteq A_1$, not both empty, such that

$$(T_0, T_1) \in M_{(\ell_0, \ell_1)}^{(i_0', i_1')}\big(\mathcal{G}_*^{(j+1)}\big)$$

for some $\ell_0 \in \{0, \ldots, i_0'\}$ and $\ell_1 \in \{0, \ldots, i_1'\}$.

Finally, set $\mathcal{A} := \mathcal{A}^{(J)}$ and $\mathcal{G}_* := \mathcal{G}_*^{(J)}$. Moreover, set

$$W := \{0, \ldots, J-1\} \setminus S = \Big\{ j \in \{0, \ldots, J-1\} : h(v_j) \neq c \Big\}.$$

Observe that the algorithm always stops after at most $v(\mathcal{G})$ iterations of the main loop. Indeed, since all constraints (A_0, A_1) with $v_j \in A_c$ are removed from $\mathcal{A}^{(j+1)}$ in part (a) of step (S4), the vertex v_j cannot be the c-maximum vertex of any $\mathcal{A}^{(j')}$ with $j' > j$ and hence the map $\{0, \ldots, J-1\} \ni j \mapsto v_j \in V(\mathcal{G})$ is injective.

2.4. **The analysis.** We shall now establish some basic properties of the algorithm described in the previous subsection. To this end, let us fix some $(i_0, i_1) \in \mathcal{U}$ and a compatible $c \in \{0,1\}$ and let i_0' and i_1' be the numbers defined in (5). Moreover, suppose that $\mathcal{G}$ is an (i_0, i_1)-uniform hypergraph and that we have run the algorithm with input $h \in \mathcal{F}(\mathcal{G})$ and obtained the (i_0', i_1')-uniform hypergraph $\mathcal{G}_*$, the integer J, the injective map $\{0, \ldots, J-1\} \ni j \mapsto v_j \in V(\mathcal{G})$, and the partition of $\{0, \ldots, J-1\}$ into S and

13

W such that $h(v_j) = c$ if and only if $j \in S$. We first state two straightforward, but fundamental, properties of the algorithm.

Observation 2.4. *If $h \in \mathcal{F}(\mathcal{G})$, then $h \in \mathcal{F}(\mathcal{G}_*)$.*

Proof. Observe that $\mathcal{G}_*$ contains only constraints of the form:

    (*i*) $(A_0 \setminus \{v\}, A_1)$, where $v \in A_0$ and $h(v) = 0$, or

    (*ii*) $(A_0, A_1 \setminus \{v\})$, where $v \in A_1$ and $h(v) = 1$,

where $(A_0, A_1) \in \mathcal{G}$, see (S3). Hence, if h violated a constraint of type (*i*) (resp. (*ii*)) then h would also violate the constraint (A_0, A_1), as $h(v) = 0$ (resp. $h(v) = 1$). $\qquad\square$

The next observation says that if the algorithm applied to two functions h and h' outputs the same set $\{v_j : j \in S\}$, then the rest of the output is also the same.

Observation 2.5. *Suppose that the algorithm applied to $h' \in \mathcal{F}(\mathcal{G})$ outputs a hypergraph $\mathcal{G}'_*$, an integer J', a map $j \mapsto v'_j$, and a partition of $\{0, \ldots, J' - 1\}$ into S' and W'. If $\{v_j : j \in S\} = \{v'_j : j \in S'\}$, then $\mathcal{G}_* = \mathcal{G}'_*$, $J = J'$, $v_j = v'_j$ for all j, and $W = W'$.*

Proof. The only step of the algorithm that depends on the input function h is (S3). There, an index j is added to the set S if and only if $h(v_j) = c$. Therefore, the execution of the algorithm depends solely on the set $\{v_j : j \in S\}$. $\qquad\square$

The next two lemmas will allow us to maintain suitable upper and lower bounds on the degrees and densities of the hypergraphs obtained by applying the algorithm iteratively. The first lemma, which is the easier of the two, states that if all the maximum degrees of $\mathcal{G}$ are appropriately bounded, then all the maximum degrees of $\mathcal{G}_*$ are also appropriately bounded.

Lemma 2.6. *Given $(\ell_0, \ell_1) \in \{0, \ldots, i_0\} \times \{0, \ldots, i_1\}$ with $\ell_0 + \ell_1 \geqslant 2$ and $\ell_c > 0$, set $\ell'_c = \ell_c - 1$ and $\ell'_{1-c} = \ell_{1-c}$. If $\Delta_{(\ell_0, \ell_1)}(\mathcal{G}) \leqslant \Delta_{(\ell_0, \ell_1)}^{(i_0, i_1)}$, then $\Delta_{(\ell'_0, \ell'_1)}(\mathcal{G}_*) \leqslant \Delta_{(\ell'_0, \ell'_1)}^{(i'_0, i'_1)}$.*

Proof. Suppose (for a contradiction) that there exist sets T'_0 and T'_1, with $|T'_0| = \ell'_0$ and $|T'_1| = \ell'_1$, such that $\deg_{\mathcal{G}_*}(T'_0, T'_1) > \Delta_{(\ell'_0, \ell'_1)}^{(i'_0, i'_1)}$. Let j be the smallest integer satisfying

$$\deg_{\mathcal{G}_*^{(j+1)}}(T'_0, T'_1) > \frac{1}{2} \cdot \Delta_{(\ell'_0, \ell'_1)}^{(i'_0, i'_1)}$$

and note that $j \geqslant 0$, since $\mathcal{G}_*^{(0)}$ is empty. We claim first that

$$\deg_{\mathcal{G}_*}(T'_0, T'_1) = \deg_{\mathcal{G}_*^{(j+1)}}(T'_0, T'_1). \tag{6}$$

Indeed, observe that $(T'_0, T'_1) \in M_{(\ell'_0, \ell'_1)}^{(i'_0, i'_1)}(\mathcal{G}_*^{(j+1)})$, and therefore the algorithm removes from $\mathcal{A}^{(j)}$ (when forming $\mathcal{A}^{(j+1)}$ in step (S4)) all pairs (A_0, A_1) such that $T'_0 \subseteq A_0$ and $T'_1 \subseteq A_1$. As a consequence, no further pairs (A'_0, A'_1) with $T'_0 \subseteq A'_0$ and $T'_1 \subseteq A'_1$ are added to $\mathcal{G}_*$ in step (S3).

We next claim that

$$\deg_{\mathcal{G}_*^{(j+1)}}(T'_0, T'_1) - \deg_{\mathcal{G}_*^{(j)}}(T'_0, T'_1) \leqslant \Delta_{(\ell_0, \ell_1)}^{(i_0, i_1)}. \tag{7}$$

To see this, recall that when we extend $\mathcal{G}_*^{(j)}$ to $\mathcal{G}_*^{(j+1)}$ in step (S3), we only add pairs $(A_0 \setminus \{v_j\}, A_1 \setminus \{v_j\})$ such that $(A_0, A_1) \in \mathcal{A}^{(j)} \subseteq \mathcal{G}$ and $v_j \in A_c$. Therefore, setting $T_c = T'_c \cup \{v_j\}$ and $T_{1-c} = T'_{1-c}$, we have

$$\deg_{\mathcal{G}_*^{(j+1)}}(T'_0, T'_1) - \deg_{\mathcal{G}_*^{(j)}}(T'_0, T'_1) \leqslant \deg_{\mathcal{G}}(T_0, T_1) \leqslant \Delta_{(\ell_0, \ell_1)}(\mathcal{G}) \leqslant \Delta_{(\ell_0, \ell_1)}^{(i_0, i_1)},$$

14

where the last inequality is by our assumption, as claimed.

Combining (6) and (7), it follows immediately that

$$\deg_{\mathcal{G}_*}(T_0', T_1') \leqslant \frac{1}{2} \cdot \Delta_{(\ell_0', \ell_1')}^{(i_0', i_1')} + \Delta_{(\ell_0, \ell_1)}^{(i_0, i_1)} \leqslant \Delta_{(\ell_0', \ell_1')}^{(i_0', i_1')},$$

where the final inequality holds by Definition 2.1. This contradicts our choice of (T_0', T_1') and therefore the lemma follows. □

We are now ready for our final lemma, which is really the heart of the matter. We will show that if $\mathcal{G}$ has sufficiently many edges and all of the maximum degrees of $\mathcal{G}$ are appropriately bounded, then either the output hypergraph $\mathcal{G}_*$ has sufficiently many edges or the value of $h(v)$ will be determined for sufficiently many vertices v. We remark that here we shall use the assumption that h takes the value 1 at most m times.

Lemma 2.7. *Suppose that $|h^{-1}(1)| \leqslant m$ and let $\alpha > 0$. If*

(A1) $e(\mathcal{G}) \geqslant \alpha \cdot \left(\frac{b}{v(\mathcal{H})}\right)^{k_1 - i_1} \left(\frac{b}{m}\right)^{k_0 - i_0} e(\mathcal{H})$ *and*

(A2) $\Delta_{(\ell_0, \ell_1)}(\mathcal{G}) \leqslant \Delta_{(\ell_0, \ell_1)}^{(i_0, i_1)}$ *for every* $(0, 0) \neq (\ell_0, \ell_1) \in \{0, \ldots, i_0\} \times \{0, \ldots, i_1\}$,

then at least one of the following statements is true:

(P1) $e(\mathcal{G}_*) \geqslant 2^{-i_0 - i_1 - 1} \alpha \cdot \left(\frac{b}{v(\mathcal{H})}\right)^{k_1 - i_1'} \left(\frac{b}{m}\right)^{k_0 - i_0'} e(\mathcal{H})$.

(P2) $c = 1$ *and* $|W| \geqslant 2^{-k_1 - 1} K^{-1} \alpha \cdot v(\mathcal{H})$.

(P3) $c = 0$ *and* $|W| \geqslant 2^{-k_0 - k_1 - 1} K^{-1} \alpha \cdot r$.

Proof. Suppose first that $c = 0$ and observe that[7]

$$e(\mathcal{G}_*) = \sum_{j \in S} \left(e(\mathcal{G}_*^{(j+1)}) - e(\mathcal{G}_*^{(j)})\right) = \sum_{j \in S} \deg_{\mathcal{A}^{(j)}}(\{v_j\}, \emptyset), \tag{8}$$

since $e(\mathcal{G}_*^{(j+1)}) - e(\mathcal{G}_*^{(j)}) = \deg_{\mathcal{A}^{(j)}}(\{v_j\}, \emptyset)$ for each $j \in S$ and $\mathcal{G}_*^{(j+1)} = \mathcal{G}_*^{(j)}$ for each $j \notin S$. To bound the right-hand side of (8), we count the edges removed from $\mathcal{A}^{(j)}$ in (a) and (b) of step (S4), which gives

$$e(\mathcal{A}^{(j)}) - e(\mathcal{A}^{(j+1)}) \leqslant \deg_{\mathcal{A}^{(j)}}(\{v_j\}, \emptyset) + \sum_{(\ell_0, \ell_1)} \left| M_{(\ell_0, \ell_1)}^{(i_0', i_1')}(\mathcal{G}_*^{(j+1)}) \setminus M_{(\ell_0, \ell_1)}^{(i_0', i_1')}(\mathcal{G}_*^{(j)}) \right| \cdot \Delta_{(\ell_0, \ell_1)}(\mathcal{G}).$$

Summing over $j \in \{0, \ldots, J-1\}$, it follows (using (8)) that

$$e(\mathcal{G}) - e(\mathcal{A}) \leqslant e(\mathcal{G}_*) + |W| \cdot \Delta_{(1,0)}(\mathcal{G}) + \sum_{(\ell_0, \ell_1)} \left| M_{(\ell_0, \ell_1)}^{(i_0', i_1')}(\mathcal{G}_*) \right| \cdot \Delta_{(\ell_0, \ell_1)}^{(i_0, i_1)},$$

since $\mathcal{A} = \mathcal{A}^{(J)} \subseteq \ldots \subseteq \mathcal{A}^{(0)} = \mathcal{G}$ and $\Delta_{(\ell_0, \ell_1)}(\mathcal{G}) \leqslant \Delta_{(\ell_0, \ell_1)}^{(i_0, i_1)}$ by (A2). Observe also that if $c = 1$, then we obtain an identical bound, with $\Delta_{(1,0)}(\mathcal{G})$ replaced by $\Delta_{(0,1)}(\mathcal{G})$.

In order to discuss both cases simultaneously, we set $\chi(0) = (1, 0)$ and $\chi(1) = (0, 1)$. Observe that

$$\Delta_{\chi(c)}(\mathcal{A}) \leqslant \Delta_{\chi(c)}(\mathcal{A}^{(j)}) \leqslant \Delta_{\chi(c)}(\mathcal{G}) \leqslant \Delta_{\chi(c)}^{(i_0, i_1)}, \tag{9}$$

since $\mathcal{A} \subseteq \mathcal{A}^{(j)} \subseteq \mathcal{G}$ and $\mathcal{G}$ satisfies (A2). It follows that, for both $c \in \{0, 1\}$,

$$e(\mathcal{G}) - e(\mathcal{A}) \leqslant e(\mathcal{G}_*) + |W| \cdot \Delta_{\chi(c)}^{(i_0, i_1)} + \sum_{(\ell_0, \ell_1)} \left| M_{(\ell_0, \ell_1)}^{(i_0', i_1')}(\mathcal{G}_*) \right| \cdot \Delta_{(\ell_0, \ell_1)}^{(i_0, i_1)}. \tag{10}$$

---

[7]Recall that $\mathcal{G}_*$ (and $\mathcal{G}_*^{(j)}$ etc.) are multi-hypergraphs and that edges are counted with multiplicity.

Now, recall that v_j is the c-maximum vertex of $\mathcal{A}^{(j)}$ and observe that therefore, by (8) and (9),

$$e(\mathcal{G}_*) = \sum_{j \in S} \Delta_{\chi(c)}\big(\mathcal{A}^{(j)}\big) \geqslant |S| \cdot \Delta_{\chi(c)}(\mathcal{A}) = b \cdot \Delta_{\chi(c)}(\mathcal{A}), \tag{11}$$

where the equality is due to the fact that $|S| \neq b$ only when $\mathcal{A}$ is empty, see step (S1).

Next, to bound the sum in (10), observe that, by Definition 2.3, we have

$$\big|M^{(i_0', i_1')}_{(\ell_0, \ell_1)}(\mathcal{G}_*)\big| \cdot \frac{1}{2} \cdot \Delta^{(i_0', i_1')}_{(\ell_0, \ell_1)} \leqslant \sum_{(T_0, T_1) \in \binom{V(\mathcal{G})}{\ell_0} \times \binom{V(\mathcal{G})}{\ell_1}} \deg_{\mathcal{G}_*}(T_0, T_1) = \binom{i_0'}{\ell_0}\binom{i_1'}{\ell_1} \cdot e(\mathcal{G}_*)$$

for each (ℓ_0, ℓ_1) and therefore

$$\sum_{(\ell_0, \ell_1)} \big|M^{(i_0', i_1')}_{(\ell_0, \ell_1)}(\mathcal{G}_*)\big| \cdot \Delta^{(i_0, i_1)}_{(\ell_0, \ell_1)} \leqslant 2 \cdot \sum_{(\ell_0, \ell_1)} \binom{i_0'}{\ell_0}\binom{i_1'}{\ell_1} \cdot e(\mathcal{G}_*) \cdot \Big(\Delta^{(i_0, i_1)}_{(\ell_0, \ell_1)} / \Delta^{(i_0', i_1')}_{(\ell_0, \ell_1)}\Big)$$
$$\leqslant 2 \cdot \big(2^{i_0' + i_1'} - 1\big) \cdot e(\mathcal{G}_*) \cdot \max_{(\ell_0, \ell_1)}\Big\{\Delta^{(i_0, i_1)}_{(\ell_0, \ell_1)} / \Delta^{(i_0', i_1')}_{(\ell_0, \ell_1)}\Big\}. \tag{12}$$

We claim that $\Delta^{(i_0, i_1)}_{(\ell_0, \ell_1)} / \Delta^{(i_0', i_1')}_{(\ell_0, \ell_1)} \leqslant m/b$ if $c = 0$ and $\Delta^{(i_0, i_1)}_{(\ell_0, \ell_1)} / \Delta^{(i_0', i_1')}_{(\ell_0, \ell_1)} \leqslant v(\mathcal{H})/b$ if $c = 1$. Indeed, both inequalities following directly from Definition 2.1, since if $c = 0$, then $(i_0', i_1') = (i_0 - 1, i_1)$, and if $c = 1$, then $(i_0', i_1') = (i_0, i_1 - 1)$. We split the remainder of the proof into two cases, depending on the value of c.

Suppose first that $c = 1$ and observe that substituting (12) into (10) yields, using the bound $\Delta^{(i_0, i_1)}_{(\ell_0, \ell_1)} / \Delta^{(i_0', i_1')}_{(\ell_0, \ell_1)} \leqslant v(\mathcal{H})/b$,

$$e(\mathcal{G}) - e(\mathcal{A}) \leqslant e(\mathcal{G}_*) + |W| \cdot \Delta^{(i_0, i_1)}_{(0, 1)} + 2 \cdot \big(2^{i_0' + i_1'} - 1\big) \cdot e(\mathcal{G}_*) \cdot \frac{v(\mathcal{H})}{b}. \tag{13}$$

Moreover, by (11), and since $i_1 \geqslant 1$ when $c = 1$, we have

$$\frac{e(\mathcal{G}_*)}{b} \geqslant \Delta_{(0,1)}(\mathcal{A}) \geqslant \frac{i_1 \cdot e(\mathcal{A})}{v(\mathcal{A})} \geqslant \frac{e(\mathcal{A})}{v(\mathcal{H})}, \tag{14}$$

since the maximum degree of a hypergraph is at least as large as its average degree. Combining (13) and (14), we obtain

$$e(\mathcal{G}) \leqslant e(\mathcal{G}_*) \cdot \frac{v(\mathcal{H})}{b} \cdot \left(\frac{b}{v(\mathcal{H})} + 1 + 2^{i_0' + i_1' + 1} - 2\right) + |W| \cdot \Delta^{(i_0, i_1)}_{(0, 1)}$$
$$\leqslant e(\mathcal{G}_*) \cdot \frac{v(\mathcal{H})}{b} \cdot 2^{i_0 + i_1} + |W| \cdot \Delta^{(i_0, i_1)}_{(0, 1)}, \tag{15}$$

since $b \leqslant v(\mathcal{H})$. Now, if the first summand on the right-hand side of (15) exceeds $e(\mathcal{G})/2$, then (A1) implies (P1), since $(i_0', i_1') = (i_0, i_1 - 1)$. Otherwise, the second summand is at least $e(\mathcal{G})/2$ and by (A1) and (3),

$$|W| \geqslant \frac{e(\mathcal{G})}{2 \cdot \Delta^{(i_0, i_1)}_{(0, 1)}} \geqslant \frac{\alpha}{2^{k_1 + 1}K} \cdot v(\mathcal{H}),$$

which is (P2).

The case $c = 0$ is slightly more delicate; in particular, we will finally use our assumption that $|h^{-1}(1)| \leqslant m$. Observe first that if $c = 0$, then substituting (12) into (10) yields,

using the bound $\Delta^{(i_0,i_1)}_{(\ell_0,\ell_1)}/\Delta^{(i'_0,i'_1)}_{(\ell_0,\ell_1)} \leqslant m/b$,

$$e(\mathcal{G}) - e(\mathcal{A}) \leqslant e(\mathcal{G}_*) + |W| \cdot \Delta^{(i_0,i_1)}_{(1,0)} + \left(2^{i_0+i_1} - 2\right) \cdot e(\mathcal{G}_*) \cdot \frac{m}{b}, \tag{16}$$

cf. (13). We claim that

$$\frac{e(\mathcal{G}_*)}{b} \geqslant \Delta_{(1,0)}(\mathcal{A}) \geqslant \frac{e(\mathcal{A})}{m}. \tag{17}$$

The first inequality follows from (11), so we only need to prove the second inequality. To do so, observe that $\mathcal{G}$ is an $(i_0, 0)$-uniform hypergraph (since $c = 0$) and therefore each function in $\mathcal{F}(\mathcal{G})$ must take the value 1 on at least one element of each set A_0 such that $(A_0, \emptyset) \in \mathcal{G}$. Now, recall that $h \in \mathcal{F}(\mathcal{G})$, that $\mathcal{A} \subseteq \mathcal{G}$, and that h takes the value 1 at most m times. It follows that $e(\mathcal{A}) \leqslant m \cdot \Delta_{(1,0)}(\mathcal{A})$, as claimed.

Combining (16) and (17), we obtain (cf. (15))

$$\begin{aligned}
e(\mathcal{G}) &\leqslant e(\mathcal{G}_*) \cdot \frac{m}{b} \cdot \left(\frac{b}{m} + 1 + 2^{i_0+i_1} - 2\right) + |W| \cdot \Delta^{(i_0,i_1)}_{(1,0)} \\
&\leqslant e(\mathcal{G}_*) \cdot \frac{m}{b} \cdot 2^{i_0+i_1} + |W| \cdot \Delta^{(i_0,i_1)}_{(1,0)},
\end{aligned} \tag{18}$$

since $b \leqslant m$. Now, if the first summand on the right-hand side of (15) exceeds $e(\mathcal{G})/2$, then (A1) implies (P1), since $(i'_0, i'_1) = (i_0 - 1, i_1)$. Otherwise, the second summand is at least $e(\mathcal{G})/2$ and by (A1) and (4),

$$|W| \geqslant \frac{e(\mathcal{G})}{2 \cdot \Delta^{(i_0,i_1)}_{(1,0)}} \geqslant \frac{\alpha}{2^{k_0+k_1+1}K} \cdot r,$$

which is (P3). $\qquad\square$

2.5. **Construction of the container.** In this section, we present the construction of containers for functions in $\mathcal{F}_{\leqslant m}(\mathcal{H})$ and analyse their properties, thus proving Theorem 1.4. For each $s \in \{0, \ldots, k_0 + k_1\}$, define

$$\alpha_s = 2^{-s(k_0+k_1+1)} \qquad \text{and} \qquad \beta_s = \alpha_s \cdot \left(\frac{b}{v(\mathcal{H})}\right)^{\min\{k_1,s\}} \left(\frac{b}{m}\right)^{\max\{0,s-k_1\}}.$$

Given an $h \in \mathcal{F}_{\leqslant m}(\mathcal{H})$, we construct the container f_h^* for h using the following procedure.

Construction of the container. Let $\mathcal{H}^{(k_0,k_1)} = \mathcal{H}$, let $S_0 = S_1 = \emptyset$, and let $(i_0, i_1) = (k_0, k_1)$. Do the following for $s = 0, \ldots, k_0 + k_1 - 1$:

(C1) Let $c \in \{0, 1\}$ be the number that is compatible with (i_0, i_1) and let (i'_0, i'_1) be the pair defined by $i'_c = i_c - 1$ and $i'_{1-c} = i_{1-c}$.

(C2) Run the algorithm with $\mathcal{G} \leftarrow \mathcal{H}^{(i_0,i_1)}$ to obtain the (i'_0, i'_1)-uniform hypergraph $\mathcal{G}_*$, the sequence $v_0, \ldots, v_{J-1} \in V(\mathcal{H})$, and the partition $\{0, 1, \ldots, J-1\} = S \cup W$.

(C3) Let $S_c \leftarrow S_c \cup \{v_j : j \in S\}$.

(C4) If $e(\mathcal{G}_*) < \beta_{s+1} \cdot e(\mathcal{H})$, then define $f_h^* : V(\mathcal{H}) \to \{0, 1, *\}$, the container for h, by

$$f_h^*(v) = \begin{cases} 1 - c & \text{if } v = v_j \text{ for some } j \in W, \\ * & \text{otherwise,} \end{cases}$$

and STOP.

(C5) Otherwise, let $\mathcal{H}^{(i'_0,i'_1)} \leftarrow \mathcal{G}_*$ and $(i_0, i_1) \leftarrow (i'_0, i'_1)$ and CONTINUE.

We will show that the above procedure indeed constructs containers for $\mathcal{F}_{\leqslant m}(\mathcal{H})$ that have the desired properties. To this end, we first claim that for each pair $(i_0, i_1) \in \mathcal{U} \cup \{(0,0)\}$, the hypergraph $\mathcal{H}^{(i_0,i_1)}$, if it was defined, satisfies:

(i) $h \in \mathcal{F}(\mathcal{H}^{(i_0,i_1)})$ and
(ii) $\Delta_{(\ell_0,\ell_1)}(\mathcal{H}^{(i_0,i_1)}) \leqslant \Delta_{(\ell_0,\ell_1)}^{(i_0,i_1)}$ for every $(0,0) \neq (\ell_0, \ell_1) \in \{0, \ldots, i_0\} \times \{0, \ldots, i_1\}$.

Indeed, one may easily prove (i) and (ii) by induction on $(k_0 + k_1) - (i_0 + i_1)$. The basis of the induction is trivial as $\mathcal{H}^{(k_0,k_1)} = \mathcal{H}$, see Definition 2.1. The inductive step follows immediately from Observation 2.4 and Lemma 2.6.

Second, we claim that for each input $h \in \mathcal{F}_{\leqslant m}(\mathcal{H})$, step (C4) is called for some s and hence the function $f_h^* \colon V(\mathcal{H}) \to \{0, 1, *\}$ is defined. If this were not true, the condition in step (C5) would be met $k_0 + k_1$ times and, consequently, we would finish with a non-empty $(0,0)$-uniform hypergraph $\mathcal{H}^{(0,0)}$, i.e., we would have $(\emptyset, \emptyset) \in \mathcal{H}^{(0,0)}$. But this contradicts (i), since no function satisfies the empty constraint and thus $h \notin \mathcal{F}(\mathcal{H}^{(0,0)})$.

Suppose, therefore, that step (C4) is executed when $\mathcal{G} = \mathcal{H}^{(i_0,i_1)}$ for some $(i_0, i_1) \in \mathcal{U}$, and note that $s = (k_0 + k_1) - (i_0 + i_1)$. We claim that $e(\mathcal{H}^{(i_0,i_1)}) \geqslant \beta_s e(\mathcal{H})$. Indeed, this is trivial if $s = 0$, whereas if $s > 0$ and this were not true, then we would have executed step (C4) at the previous step. We therefore have

$$e(\mathcal{G}) = e(\mathcal{H}^{(i_0,i_1)}) \geqslant \beta_s \cdot e(\mathcal{H}) \qquad \text{and} \qquad e(\mathcal{G}_*) < \beta_{s+1} \cdot e(\mathcal{H}),$$

which, by Lemma 2.7 and (ii), implies that either (P2) or (P3) of Lemma 2.7 holds. Note that if $c = 1$, then $k_1 \geqslant i_1 > 0$ and we have

$$|(f_h^*)^{-1}(0)| \geqslant 2^{-k_1-1}K^{-1}\alpha_s \cdot v(\mathcal{H}) \geqslant \alpha_{k_0+k_1}K^{-1}v(\mathcal{H}) = \delta v(\mathcal{H}),$$

where $\delta = 2^{-(k_0+k_1)(k_0+k_1+1)}K^{-1}$. On the other hand, if $c = 0$, then $k_0 \geqslant i_0 > 0$ and

$$|(f_h^*)^{-1}(1)| \geqslant 2^{-k_0-k_1-1}K^{-1}\alpha_s \cdot r \geqslant \alpha_{k_0+k_1}K^{-1}r = \delta r.$$

This verifies that f_h^* satisfies property (b) from the statement of Theorem 1.4.

To complete the proof, we need to show that f^* decomposes as $f^* = f \circ g$ for some $g \colon \mathcal{F}_{\leqslant m}(\mathcal{H}) \to \binom{V(\mathcal{H})}{\leqslant k_0 b} \times \binom{V(\mathcal{H})}{\leqslant k_1 b}$ and to verify that properties (a) and (c) from the statement of the theorem hold. We claim that one may take $g(h) = (S_0, S_1)$, where S_0 and S_1 are the sets constructed by the above procedure, see (C3). To this end, it suffices to show that if for some $h, h' \in \mathcal{F}(\mathcal{H})$ the above procedure produces the same pair (S_0, S_1), then $f_h^* = f_{h'}^*$. To see this, observe first that the set S defined in step (C2) is precisely the set of all indices $j \in \{0, \ldots, J - 1\}$ that satisfy $v_j \in S_c$. Indeed, the former set is contained in the latter by construction, see (C3). The reverse inclusion holds because

$$S = \big\{ j \in \{0, \ldots, J-1\} : h(v_j) = c \big\}$$

and $h(v) = c$ for every $v \in S_c$. By Observation 2.5, it follows that the output of the algorithm depends only on the pair (S_0, S_1) and hence $f_h^* = f_{h'}^*$, as claimed.

Finally, observe that $S_0 \subseteq h^{-1}(0)$ and $S_1 \subseteq h^{-1}(1)$, by construction, and that h belongs to the cylinder $f(g(h)) = f_h^*$, since $h(v) = 1 - c$ for every $v = v_j$ with $j \in W$, by the definition of $W = \big\{ j \in \{0, \ldots, J-1\} : h(v_j) \neq c \big\}$. This verifies properties (a) and (c) and hence completes the proof of Theorem 1.4. $\qquad \square$

2.6. **Derivation of Theorem 1.5.** We conclude this part of the paper with the easy derivation of Theorem 1.5 from Theorem 1.4. Given a k-uniform hypergraph $\mathcal{H}$ satisfying the assumptions of Theorem 1.5 for some b and r, one may invoke Theorem 1.4 with $\mathcal{H}_{1.4}$ being the $(0, k)$-uniform hypergraph with the same vertex set as $\mathcal{H}$ whose edges are all pairs $(\emptyset, A)$ such that A is an edge of $\mathcal{H}$. Since $k_0 = 0$ and $\Delta_{(0,\ell)}(\mathcal{H}_{1.4}) = \Delta_\ell(\mathcal{H})$ for every $\ell \in \{1, \dots, k\}$, one can easily check that $\mathcal{H}_{1.4}$ satisfies the assumptions of Theorem 1.4 with the same b, $m_{1.4} \leftarrow v(\mathcal{H})$, and $K_{1.4} \leftarrow v(\mathcal{H})/r$.

Now, observe that the family $\mathcal{F}(\mathcal{H}_{1.4})$ comprises precisely the characteristic functions of all independent sets of $\mathcal{H}$ and that $\mathcal{F}_{\leqslant m}(\mathcal{H}_{1.4}) = \mathcal{F}_{\leqslant v(\mathcal{H})}(\mathcal{H}_{1.4}) = \mathcal{F}(\mathcal{H}_{1.4})$. Given an independent set $I \in \mathcal{I}(\mathcal{H})$, let $h \in \mathcal{F}(\mathcal{H}_{1.4})$ be its characteristic function. Let $S := (\emptyset, S_1) = g_{1.4}(h)$ and $X := f_{1.4}(S)^{-1}(0)$ and set $g(I) := S_1$ and $f(g(I)) := V(\mathcal{H}) \setminus X$. Recalling again that $k_0 = 0$, it is straightforward to verify that properties (a), (b), and (c) from the statement of Theorem 1.4 imply the assertion of Theorem 1.5. □

3. Robust balanced stability for induced C_4s

Recall from Section 1.3 that a pregraph is a pair (M, E) of disjoint subsets of $E(K_n)$. The elements of E are called edges whereas the elements of M are called mixed edges. A good copy of C_4 in a pregraph (M, E) is a copy of C_4 in M whose vertex set is independent in E. In particular, the vertex set of each good copy of C_4 induces four, five, or six edges of M, four of which play the roles of edges of C_4.[8]

Given a pregraph $\mathcal{P} = (M, E)$, we define three hypergraphs with vertex set M, denoted $\mathcal{H}_0^{\mathcal{P}}$, $\mathcal{H}_1^{\mathcal{P}}$, and $\mathcal{H}_2^{\mathcal{P}}$. The $(i, 4)$-uniform hypergraph $\mathcal{H}_i^{\mathcal{P}}$ comprises all pairs (A, B) such that B is a good copy of C_4 and A is the set of the remaining i mixed edges induced by the vertex set of this copy (which induces exactly $4 + i$ edges of M). Recall that we say that a graph G is ε-close to K_ℓ if one can transform G into K_ℓ by adding or deleting at most $\varepsilon \binom{\ell}{2}$ edges. The following theorem, a robust stability statement for good copies of the 4-cycle in a pregraph, is the main result of this section.

Theorem 3.1. *For every $\varepsilon > 0$, there exist positive constants β, δ, λ, and C such that the following holds for all ℓ and n satisfying $\ell \geqslant C\sqrt{n}$. Suppose that $\mathcal{P} = (M, E)$ is a pregraph on n vertices with $e(E) \leqslant \binom{\ell}{2}$ and either*

(M1) $e(M) \geqslant 4\ell n$, or
(M2) $e(M) \geqslant (1 - \delta)\ell n$, E is not ε-close to K_ℓ, and $\ell \leqslant \lambda n$, or
(M3) there exists $U \subseteq V(K_n)$ with $|U| = \ell$, $e_E(U) \geqslant (1 - \varepsilon)\binom{\ell}{2}$, and $e_M(U^c) \geqslant 7\sqrt{\varepsilon}\ell n$.

Then there exist $i \in \{0, 1, 2\}$ and $\mathcal{H}_i \subseteq \mathcal{H}_i^{\mathcal{P}}$ such that

$$e(\mathcal{H}_i) \geqslant \beta\ell^4, \qquad \Delta_{(0,1)}(\mathcal{H}_i) \leqslant \frac{\ell^3}{n}, \qquad and \qquad \Delta_{(0,2)}(\mathcal{H}_i) \leqslant \ell$$

and, if $i > 0$, then also $\Delta_{(1,0)}(\mathcal{H}_i) \leqslant \ell^2$.

Let us say that an $(i, 4)$-uniform hypergraph $\mathcal{H}_i$ is *permissible* if it satisfies both (all three, if $i > 0$) maximum degree conditions stated in Theorem 3.1. We shall thus be looking for a permissible subhypergraph $\mathcal{H}_i \subseteq \mathcal{H}_i^{\mathcal{P}}$, for some $i \in \{0, 1, 2\}$, that has $\Omega(\ell^4)$ edges. We shall build the $\mathcal{H}_0$, $\mathcal{H}_1$, and $\mathcal{H}_2$ by adding to them one edge at a time, making

---

[8]If the vertex set of a good copy of C_4 induces six mixed edges, then there are three choices for these four edges, each corresponding to a different embedding of C_4 into K_4.

sure that we stay within the class of permissible hypergraphs, until one of them has sufficiently many edges. (Trivially, an empty hypergraph is permissible.)

It will be convenient to use the following nomenclature. A pair (S, T) of disjoint sets of edges of K_n is *saturated* in a hypergraph $\mathcal{H}$ if $\deg_{\mathcal{H}}(S, T)$ attains or exceeds its maximum permitted value. That is, if

(i) $(|S|, |T|) = (0, 1)$ and $\deg_{\mathcal{H}}(S, T) \geqslant \lfloor \ell^3 / n \rfloor$, or
(ii) $(|S|, |T|) = (1, 0)$ and $\deg_{\mathcal{H}}(S, T) \geqslant \ell^2$, or
(iii) $(|S|, |T|) = (0, 2)$ and $\deg_{\mathcal{H}}(S, T) \geqslant \ell$.

Thus, in the setting of Theorem 3.1, we shall be looking for an $i \in \{0, 1, 2\}$ and an edge of $\mathcal{H}_i^{\mathcal{P}} \setminus \mathcal{H}_i$ which does not contain any saturated pair. We first show how to deduce Theorem 3.1 from the following, seemingly weaker, statement by performing an appropriate preprocessing of the pregraph $\mathcal{P}$. This preprocessing of $\mathcal{P}$ will 'disable' all saturated pairs of types (i) and (ii), so that we will only have to worry about pairs of type (iii).

Theorem 3.2. *For every $0 < \varepsilon \leqslant 1/2$, there exist positive constants β, δ, λ, and C such that the following holds for all ℓ and n satisfying $\ell \geqslant C\sqrt{n}$. Suppose that $\mathcal{P} = (M, E)$ is a pregraph on n vertices with $e(E) \leqslant \binom{\ell}{2}$ and either*

(M1*) $e(M) \geqslant 3\ell n$ or
(M2*) $e(M) \geqslant (1 - \delta)\ell n$, E is not ε-close to K_ℓ, and $\ell \leqslant \lambda n$.
Then for any collection $\mathcal{C}$ of at most $12\beta\ell^3$ pairs of elements of M, there exist at least $3\beta\ell^4$ good copies of C_4 in $\mathcal{P}$ that contain no pair from $\mathcal{C}$.

Derivation of Theorem 3.1 from Theorem 3.2. Given $0 < \varepsilon \leqslant 2$,[9] let $\beta_{3.2}$, $\delta_{3.2}$, $\lambda_{3.2}$, and $C_{3.2}$ be the constants whose existence is asserted by Theorem 3.2 with $\varepsilon_{3.2} \leftarrow \varepsilon/4$ and let

$$\delta = \min\left\{\frac{\delta_{3.2}}{3}, \frac{\varepsilon}{10}\right\}, \quad \beta = \min\left\{\frac{\varepsilon^2\beta_{3.2}}{4}, \frac{\delta}{20}\right\}, \quad \lambda = \frac{\lambda_{3.2}}{2}, \quad \text{and} \quad C = \frac{C_{3.2}}{\sqrt{\varepsilon}}.$$

Suppose that a pregraph $\mathcal{P} = (M, E)$ satisfies the assumptions of Theorem 3.1. We shall build the (initially empty) hypergraphs $\mathcal{H}_0$, $\mathcal{H}_1$, and $\mathcal{H}_2$ edge by edge, making sure that we stay within the class of permissible hypergraphs, until one of them has sufficiently many edges. To this end, suppose that we have succeeded in constructing some permissible $\mathcal{H}_0$, $\mathcal{H}_1$, and $\mathcal{H}_2$, but each of them has fewer than $\beta\ell^4$ edges. We shall modify the pregraph $\mathcal{P}$ by removing from M all mixed edges f for which there exists $i \in \{0, 1, 2\}$ such that either $(\emptyset, \{f\})$ or $(\{f\}, \emptyset)$ (or both) is saturated in $\mathcal{H}_i$. This will ensure that every good copy of C_4 that we will later find in this modified colouring will not contain any saturated pair (S, T) of type (i) or (ii). To achieve this, we first move all mixed edges f for which $(\{f\}, \emptyset)$ is saturated in either $\mathcal{H}_1$ or $\mathcal{H}_2$ from M to E and then move all f for which $(\emptyset, \{f\})$ is saturated in any of the $\mathcal{H}_i$ from M to an initially empty set N. Denote the modified pregraph by $\mathcal{P}' = (M', E')$. Observe, crucially, that each good copy of C_4 in $\mathcal{P}'$ is also good in $\mathcal{P}$, as $E' \supseteq E$ and $M' \subseteq M$. Moreover, each such copy yields an edge of one of the $\mathcal{H}_i^{\mathcal{P}}$ with no saturated pair of type (i) or (ii), where $4 + i$ is the number of edges of $M' \cup N$ induced by the vertex set of this 4-cycle.[10]

---

[9]Note that the result for $\varepsilon > 2$ is implied by the statement for $\varepsilon = 2$, since condition (M3) is then stronger than condition (M1), and every graph with at most $\binom{\ell}{2}$ edges is 2-close to K_ℓ.

[10]The four edges forming a good copy of C_4 in $\mathcal{P}'$ belong to M', but the remaining two edges induced by the vertex set of this cycle could belong to $N \subseteq M \setminus M'$.

Let $\ell' = \lfloor (1+\delta)\ell \rfloor$. As each of the $\mathcal{H}_i$ has fewer than $\beta\ell^4$ edges, then

$$e(E' \setminus E) \leqslant \sum_{i=1}^{2} \frac{ie(\mathcal{H}_i)}{\ell^2} < 3\beta\ell^2 \leqslant \frac{\delta\ell^2}{2}$$

and

$$e(M \setminus M') \leqslant e(E' \setminus E) + \sum_{i=0}^{2} \frac{4e(\mathcal{H}_i)}{\lfloor \ell^3/n \rfloor} < 3\beta\ell^2 + 13\beta\ell n < 20\beta\ell n \leqslant \delta\ell n.$$

In particular,

$$e(E') \leqslant \binom{\ell}{2} + \frac{\delta\ell^2}{2} \leqslant \binom{\ell'}{2}.$$

Moreover, if $e(M) \geqslant 4\ell n$, then $e(M') \geqslant 3\ell' n$, and if $e(M) \geqslant (1-\delta)\ell n$, then $e(M') \geqslant (1-3\delta)\ell' n \geqslant (1-\delta_{3.2})\ell' n$. Finally, if E' is $(\varepsilon/4)$-close to $K_{\ell'}$, then E is ε-close to K_ℓ, as $e(K_{\ell'}) - e(K_\ell) \leqslant 2\delta\ell^2$, and $\delta \leqslant \varepsilon/10$. Therefore, if $\mathcal{P}$ satisfies the assumptions of Theorem 3.1 with either (M1) or (M2), then $\mathcal{P}'$ satisfies the assumptions of Theorem 3.2 with $\varepsilon_{3.2} \leftarrow \varepsilon/4$ and $\ell_{3.2} \leftarrow \ell'$, see (M1*) and (M2*).

Now, let $\mathcal{C}$ be the collection of all T such that $(\emptyset, T)$ is a saturated pair of type (iii) in one of the $\mathcal{H}_i$ and observe that

$$|\mathcal{C}| \leqslant \sum_{i=0}^{2} \frac{4e(\mathcal{H}_i)}{\ell} < 12\beta\ell^3 \leqslant 12\beta_{3.2}(\ell')^3,$$

as each edge of $\mathcal{H}_i$ contains at most four such saturated pairs (if $f_1, f_2 \in M$ do not share a vertex, then $\deg_{\mathcal{H}_i^\mathcal{P}}(\emptyset, \{f_1, f_2\}) \leqslant 2$). Therefore, if $\mathcal{P}$ satisfies the assumptions of Theorem 3.1 with either (M1) or (M2), then we may invoke Theorem 3.2 to find at least $3\beta_{3.2}(\ell')^4 \geqslant 3\beta\ell^4$ good copies of C_4 in $\mathcal{P}'$, none of which contains a pair from $\mathcal{C}$.

On the other hand, if $\mathcal{P}$ satisfies the assumptions of Theorem 3.1 with (M3), then $\mathcal{P}'$ restricted to the set U^c satisfies the assumptions of Theorem 3.2 with $\ell_{3.2} \leftarrow 2\sqrt{\varepsilon}\ell$, as

$$e_{E'}(U^c) \leqslant e_E(U^c) + e(E' \setminus E) \leqslant \varepsilon\binom{\ell}{2} + \frac{\delta\ell^2}{2} \leqslant \varepsilon\ell^2 \leqslant \binom{2\sqrt{\varepsilon}\ell}{2}$$

and

$$e_{M'}(U^c) \geqslant e_M(U^c) - e(M \setminus M') \geqslant 7\sqrt{\varepsilon}\ell n - \delta\ell n \geqslant 6\sqrt{\varepsilon}\ell n \geqslant 3(2\sqrt{\varepsilon}\ell)|U^c|,$$

see (M1*). Since

$$2\sqrt{\varepsilon}\ell \geqslant 2\sqrt{\varepsilon} \cdot C\sqrt{n} \geqslant C_{3.2}\sqrt{|U^c|} \quad \text{and} \quad |\mathcal{C}| < 12\beta\ell^3 \leqslant 12\beta_{3.2} \cdot \left(2\sqrt{\varepsilon}\ell\right)^3,$$

we may again invoke Theorem 3.2 to find at least $3\beta_{3.2}(2\sqrt{\varepsilon}\ell)^4 \geqslant 3\beta\ell^4$ good copies of C_4 in $\mathcal{P}'$, none of which contains a pair from $\mathcal{C}$.

Finally, it follows from our construction that each good copy of C_4 in $\mathcal{P}'$ corresponds to an edge of $\mathcal{H}_i^\mathcal{P}$ for some $i \in \{0, 1, 2\}$ that additionally does not contain any saturated pairs of type (i) or (ii). Moreover, by our definition of $\mathcal{C}$, none of the at least $3\beta\ell^4$ copies we have found above contains a saturated pair of type (iii) either. Recalling that $e(\mathcal{H}_0) + e(\mathcal{H}_1) + e(\mathcal{H}_2) < 3\beta\ell^4$, it follows that one of these good C_4s yields a pair $(A, B) \in \mathcal{H}_i^\mathcal{P} \setminus \mathcal{H}_i$ such that $\mathcal{H}_i \cup \{(A, B)\}$ is permissible. Iterating this process, we must eventually arrive at a permissible hypergraph $\mathcal{H}_i$ (for some $i \in \{0, 1, 2\}$) with at least $\beta\ell^4$ edges, as required. $\qquad\square$

The remainder of the this section is dedicated to the proof of Theorem 3.2. We begin by proving the following proposition, which proves Theorem 3.2 when the condition (M1*) holds and will moreover serve as a helpful warm-up for the proof of the theorem. It will also be a step in the proof of the theorem under the assumption (M2*).

Proposition 3.3. *Suppose that integers ℓ and n satisfy $\ell \geqslant \sqrt{n}$ and that $\mathcal{P} = (M, E)$ is a pregraph on n vertices with $e(E) \leqslant \binom{\ell}{2}$ and $e(M) \geqslant 3\ell n$. Then for any collection $\mathcal{C}$ of at most $\ell^3/40$ pairs of elements of M, there exist at least $\ell^4/40$ good copies of C_4 in $\mathcal{P}$ that contain no pair from $\mathcal{C}$.*

Our proofs will use the following two auxiliary statements. The first is a well-known result of Caro [13] and Wei [48]. We remark that, in this section, if G is a graph (such as M or E), we will write $d_G(v)$ and $d_G(v, S)$ to denote the number of neighbours of v and the number of neighbours of v in S, respectively.

Lemma 3.4. *For every graph G,*

$$\alpha(G) \geqslant \sum_{v \in V(G)} \frac{1}{1 + d_G(v)}.$$

The second is an easy consequence of Jensen's inequality applied to the convex function $[0, \infty) \ni x \mapsto 1/(1 + x)$. Given a nonnegative integer d and a real number $q \in [0, 1]$, we shall denote by $\mathrm{Bin}(d, q)$ the binomial random variable with parameters d and q.

Fact 3.5. *For every $d \in \mathbb{N}$ and $q \in [0, 1]$,*

$$\mathbb{E}\left[\frac{1}{1 + \mathrm{Bin}(d, q)}\right] \geqslant \frac{1}{1 + qd}.$$

Proof of Proposition 3.3. Fix a pregraph $\mathcal{P} = (M, E)$ on n vertices and a collection $\mathcal{C}$ satisfying the assumptions of the proposition. We first remove all vertices whose degree in M is less than 2ℓ. As this way we lose at most $2\ell n$ edges of M, we arrive at an m-vertex subset $W \subseteq V(K_n)$, for some $2\ell \leqslant m \leqslant n$, such that $\delta(M[W]) \geqslant 2\ell$. Clearly, it is sufficient to find $\ell^4/40$ good copies of C_4 in $\mathcal{P}$ restricted to W, none of which contains a pair from $\mathcal{C}$. Therefore, shall replace the original M, E, and $\mathcal{P}$ with their restrictions to the set W.

Set $q = m/\ell^2 \leqslant n/\ell^2 \leqslant 1$ and form a random subset $R \subseteq W$ by retaining each element of W independently with probability q. We apply Lemma 3.4 to the graph $E[R]$ to find an independent set $I \subseteq R$ with

$$|I| \geqslant \sum_{v \in R} \frac{1}{1 + d_{E[R]}(v)} = \sum_{v \in W} \frac{\mathbb{1}[v \in R]}{1 + d_E(v, R)}.$$

By Fact 3.5, we have

$$\mathbb{E}[|I|] \geqslant \mathbb{E}\left[\sum_{v \in W} \frac{\mathbb{1}[v \in R]}{1 + d_E(v, R)}\right] = \sum_{v \in W} q \cdot \mathbb{E}\left[\frac{1}{1 + \mathrm{Bin}(d_E(v), q)}\right] \geqslant \sum_{v \in W} \frac{q}{1 + qd_E(v)}.$$

As the function $[0, \infty) \ni x \mapsto q/(1 + qx)$ is convex, the sum in the right-hand side above is minimised when $d_E(v) = 2e(E)/m$ for every $v \in W$. As $e(E) \leqslant \ell^2/2$, then

$$\mathbb{E}[|I|] \geqslant \frac{qm}{1 + 2qe(E)/m} \geqslant \frac{qm}{1 + q \cdot \ell^2/m} = \frac{m^2}{2\ell^2}. \tag{19}$$

Next, let us choose, for each vertex $v \in W$, an arbitrary set M_v of 2ℓ edges of M that are incident to v. We shall say that a copy of $K_{1,2}$ is *good* if its centre v lies in I, both of its edges are in M_v, and the pair comprising its two non-centre vertices does not belong to E. The number X_g of such good $K_{1,2}$s satisfies

$$X_g \geqslant \sum_{v \in I} \left(\binom{2\ell}{2} - e(E) \right) \geqslant \left(\binom{2\ell}{2} - \binom{\ell}{2} \right) \cdot |I| \geqslant \frac{4\ell^2}{3} \cdot |I|. \tag{20}$$

We shall say that a copy of $K_{1,2}$ in M is *saturated* if (the set consisting of) its two edges belong to $\mathcal{C}$. Let X_s be the number of saturated $K_{1,2}$s in M whose centre vertex belongs to the (random) set $I \subseteq R$. Writing X for the number of good $K_{1,2}$s that are not saturated, we have $X \geqslant X_g - X_s$ and hence, recalling that $|\mathcal{C}| \leqslant \ell^3/40$,

$$\mathbb{E}[X] \geqslant \mathbb{E}[X_g] - \mathbb{E}[X_s] \geqslant \frac{4\ell^2}{3} \cdot \mathbb{E}[|I|] - q \cdot |\mathcal{C}| \geqslant \frac{2m^2}{3} - \frac{\ell m}{40} \geqslant \frac{3m^2}{5}, \tag{21}$$

where we have used (19), (20), and the inequality $m \geqslant 2\ell$.

Since I is an independent set in E, it follows that any pair of good $K_{1,2}$s with the same non-centre vertices form a good C_4 and therefore we have at least $X - \binom{m}{2}$ such C_4s. However, we must disregard those C_4s that contain a saturated $K_{1,2}$ whose two non-centre vertices lie in I, since the two edges of such a saturated $K_{1,2}$ could come from two different good non-saturated $K_{1,2}$s whose centre vertices lie in I. The expected number of saturated $K_{1,2}$s of this type is at most $q^2 \cdot |\mathcal{C}|$ and each of them lies in at most 2ℓ of our good C_4s, since the edges of our good C_4s came only from the sets M_v. We must therefore discard (in expectation) at most $2\ell q^2 |\mathcal{C}|$ of the (at least) $X - \binom{m}{2}$ good C_4s found using pairs of good $K_{1,2}$s.

To summarise, let Z be the number of good C_4s that contain no saturated $K_{1,2}$ and at least two vertices of I. By (21) and the argument above, we have

$$\mathbb{E}[Z] \geqslant \mathbb{E}[X] - \binom{m}{2} - 2\ell q^2 |\mathcal{C}| \geqslant \left(\frac{3}{5} - \frac{1}{2} - \frac{2|\mathcal{C}|}{\ell^3} \right) m^2 \geqslant \frac{m^2}{20}.$$

Finally, observe that each good copy of C_4 containing no saturated $K_{1,2}$ has probability at most $2q^2$ of being counted by Z. It therefore follows that the total number of such copies of C_4 must be at least $m^2/(40q^2) = \ell^4/40$, as required. $\qquad\square$

We next consider pregraphs $\mathcal{P} = (M, E)$ for which one can find a small set A of vertices of K_n that contains only a tiny proportion of the edges of E, but still a large proportion of mixed edges have an endpoint in A. The following proposition will be invoked in the proof of Theorem 3.2.

Proposition 3.6. *Suppose that integers ℓ and n satisfy $\ell \geqslant 4\sqrt{n}$ and set $\alpha := 1/640$. Let $\mathcal{P} = (M, E)$ be a pregraph on n vertices with $e(E) \leqslant \binom{\ell}{2}$ and suppose that there exists a set $A \subseteq V(K_n)$, with $|A| \leqslant \alpha n$ and $e_E(A) \leqslant \alpha \ell^2$, such that*

$$\sum_{w \in A} d_M(w) \geqslant \frac{\ell n}{2}.$$

Then for any collection $\mathcal{C}$ of at most $\alpha \ell^3$ pairs of elements of M, there exist at least $\alpha \ell^4$ good copies of C_4 in $\mathcal{P}$ that contain no pair from $\mathcal{C}$.

Proof. The proof of Proposition 3.6 follows the general strategy of the proof of Proposition 3.3, but there are some key differences. In particular, we will find the independent

set I inside the set A alone and we shall select vertices of R with different probabilities. Rather than invoking Lemma 3.4 and Fact 3.5, we shall give a somewhat finer argument to produce a large independent set $I \subseteq R$ and use it to construct good copies of C_4.

We start by iteratively removing from A all vertices v that do not satisfy

$$d_M(v) \geqslant \max \left\{ 2\ell, \frac{n}{16\alpha\ell} \cdot d_E(v, A) \right\}. \tag{22}$$

Observe that the set A' of vertices remaining after this deletion satisfies

$$\sum_{v \in A'} d_M(v) \geqslant \sum_{v \in A} d_M(v) - 2 \cdot \left(|A| \cdot 2\ell + e_E(A) \cdot \frac{n}{16\alpha\ell} \right)$$

$$\geqslant \frac{\ell n}{2} - 4\alpha\ell n - \alpha\ell^2 \cdot \frac{n}{8\alpha\ell} \geqslant \frac{\ell n}{3}. \tag{23}$$

Let $a = |A'|$ and order the elements of A' as $v_1, \ldots, v_a$ so that $d_M(v_i) \leqslant d_M(v_j)$ whenever $1 \leqslant i \leqslant j \leqslant a$. For each $i \in [a]$, let

$$q_i = \frac{8n}{\ell \cdot d_M(v_i)} \leqslant \frac{4n}{\ell^2} \leqslant \frac{1}{4}$$

and form a random set $R \subseteq A'$ by keeping each v_i independently with probability q_i. Define

$$I = \left\{ v_i \in R : v_j \notin R \text{ for every } j > i \text{ such that } v_i v_j \in E \right\}$$

and observe that I is an independent set in the graph E.[11]

Similarly to before, we shall say that a copy of $K_{1,2}$ in M is good if its centre lies in I and the pair comprising its two non-centre vertices does not belong to E. Observe that the number X_g of good $K_{1,2}$s satisfies

$$X_g \geqslant \sum_{v \in I} \left(\binom{d_M(v)}{2} - e(E) \right) \geqslant \sum_{v \in I} \left(\binom{d_M(v)}{2} - \binom{\ell}{2} \right) \geqslant \sum_{v \in I} \frac{d_M(v)^2}{3},$$

as $d_M(v) \geqslant 2\ell$ for each $v \in I$. We shall now estimate the probability that a given vertex $v \in A'$ belongs to the random set I. To this end, suppose that $v = v_i$ for some $i \in [a]$ and note that, by (22), there are at most $d_M(v_i) \cdot 16\alpha\ell/n$ indices j such that $v_i v_j \in E$. Moreover, by our choice of the ordering, $q_j \leqslant q_i$ whenever $j > i$. Letting $d = d_M(v) = d_M(v_i)$, and recalling that $8n/(\ell d) \leqslant 1/4$, it follows that

$$\mathbb{P}\big(v \in I\big) \geqslant q_i \cdot (1 - q_i)^{16\alpha\ell d/n} = \frac{8n}{\ell d} \cdot \left(1 - \frac{8n}{\ell d} \right)^{16\alpha\ell d/n} \geqslant e^{-160\alpha} \cdot \frac{8n}{\ell d} \geqslant \frac{6n}{\ell d},$$

where we used the bounds $1 - x \geqslant e^{-5x/4}$ when $0 \leqslant x \leqslant 1/4$ and $e^{-1/4} > 3/4$.

We will need to disregard the saturated $K_{1,2}$s, that is, all those whose pair of edges belongs to $\mathcal{C}$. Let X_s be the number of those saturated $K_{1,2}$s whose centre vertex belongs to the set I. Writing X for the number of good $K_{1,2}$s that are not saturated, we have $X \geqslant X_g - X_s$, and hence

$$\mathbb{E}[X] \geqslant \mathbb{E}[X_g] - \mathbb{E}[X_s] \geqslant \sum_{v \in A'} \mathbb{P}(v \in I) \cdot \frac{d_M(v)^2}{3} - \max_i q_i \cdot |\mathcal{C}|$$

$$\geqslant \sum_{v \in A'} \frac{2n d_M(v)}{\ell} - \frac{4n}{\ell^2} \cdot |\mathcal{C}| \geqslant \frac{2n^2}{3} - \frac{2|\mathcal{C}|n^2}{\ell^3} \geqslant \frac{3n^2}{5},$$

---

[11]The idea of forming a large independent set this way is taken from the proof of Lemma 3.4 given in [4].

where we have used (23) and the inequality $n \geqslant 2\ell$ (which holds since A' is non-empty).

Since I is an independent set in E, it follows that any pair of good $K_{1,2}$s with the same non-centre vertices forms a good C_4. Thus we have at least $X - \binom{n}{2}$ such C_4s. However, we must still disregard those C_4s that contain a saturated $K_{1,2}$ with two non-centre vertices in I. Fix some $K_{1,2}$ from $\mathcal{C}$ and suppose that its non-centre vertices are v_i and v_j. Observe that it can lie in at most $d_M(v_i)$ of our good copies of C_4. Therefore, the expected number of good C_4s that we are forced to disregard because of this single $K_{1,2}$ is at most

$$q_i \cdot q_j \cdot d_M(v_i) = \frac{64n^2}{\ell^2 d_M(v_j)} \leqslant \frac{32n^2}{\ell^3}.$$

Consequently, the expected number of good copies of C_4 that we have to disregard because of one of the saturated $K_{1,2}$s from $\mathcal{C}$ is at most $32|\mathcal{C}|n^2/\ell^3$.

To summarise, let Z be the number of good C_4s that contain no saturated $K_{1,2}$ and at least two vertices of I. We have shown that

$$\mathbb{E}[Z] \geqslant \mathbb{E}[X] - \binom{n}{2} - \frac{32|\mathcal{C}| \cdot n^2}{\ell^3} \geqslant \left(\frac{3}{5} - \frac{1}{2} - \frac{32|\mathcal{C}|}{\ell^3}\right)n^2 \geqslant \frac{n^2}{20}.$$

But as each good copy of C_4 containing no saturated $K_{1,2}$ has chance at most $2q_1^2$ to be counted by Z, the number of them is at least $n^2/(40q_1^2) \geqslant \ell^4/640$. This completes the proof of the proposition. $\square$

Proof of Theorem 3.2. We begin by defining the constants whose existence is claimed in the statement of the theorem. Given $0 < \varepsilon \leqslant 1/2$, set $\alpha = 2^{-16}$ and define

$$C = \frac{4}{\alpha^3}, \qquad \delta = \min\left\{\frac{\alpha^3}{2^7}, \frac{\varepsilon}{16}\right\}, \qquad \beta = \frac{\delta^4}{2^{100}}, \qquad \text{and} \qquad \lambda = \frac{\delta^7}{2^{10}}.$$

Suppose that $\ell \geqslant C\sqrt{n}$ and let $\mathcal{P} = (M, E)$ be a pregraph on n vertices with $e(E) \leqslant \binom{\ell}{2}$. If $\mathcal{P}$ satisfies (M1*), then we may immediately invoke Proposition 3.3, noting that $|\mathcal{C}| \leqslant 12\beta\ell^3 \leqslant \ell^3/40$, to find find $\ell^4/40$ good copies of C_4 that contain no pair from $\mathcal{C}$.

We may therefore assume from now on that $\mathcal{P}$ satisfies (M2*), that is,

$$e(M) \geqslant (1-\delta)\ell n, \qquad E \text{ is not } \varepsilon\text{-close to } K_\ell, \qquad \text{and} \qquad \ell \leqslant \lambda n.$$

We begin by iteratively removing all vertices v whose degree in M is smaller than $(1-2\delta)\ell$. As this way we can remove at most $(1-2\delta)\ell n$ edges of M, we will eventually arrive at a set $W \subseteq V(K_n)$ with $\delta(M[W]) \geqslant (1-2\delta)\ell$. Set $m = |W|$, and note that, since we removed at most $(1-2\delta)\ell(n-m)$ edges of M, we have $e_M(W) \geqslant \max\{(1-\delta)\ell m, \delta\ell n\}$, and therefore

$$m \geqslant \sqrt{\delta\ell n} \geqslant \sqrt{\frac{\delta}{\lambda}} \cdot \ell \geqslant \frac{32}{\delta^3} \cdot \ell. \tag{24}$$

Observe that the subgraph of E induced by W is also not $(\varepsilon/2)$-close to K_ℓ. Indeed, otherwise there would be an ℓ-element set $U \subseteq W$ with $e_E(U) \geqslant (1 - \varepsilon/2)\binom{\ell}{2}$, which would imply that E itself is ε-close to K_ℓ, as $e(E) \leqslant \binom{\ell}{2}$. We may thus work with the restrictions of M, E, and $\mathcal{P}$ to the set W. We shall surpress W from the notation and write M, E, and $\mathcal{P}$ in place of $M[W]$, $E[W]$ and $(M[W], E[W])$. In particular,

$$e(M) \geqslant (1-\delta)\ell m, \qquad e(E) \leqslant \binom{\ell}{2}, \qquad E \text{ is not } (\varepsilon/2)\text{-close to } K_\ell,$$

and moreover $\delta(M) \geqslant (1-2\delta)\ell$.

We split the proof into two cases, depending on the shape of the degree sequence of E.

Case 1. There is a set $L \subseteq W$ of αm vertices v satisfying $d_E(v) \leqslant (1-\alpha)\ell^2/m$.

Set $q = Cm/\ell^2 \leqslant 1/C$ and form a random subset $R \subseteq W$ by keeping each element of W independently with probability q. We apply Lemma 3.4 to the graph $E[R]$ (cf. the proof of Proposition 3.3) to find an independent set $I \subseteq R$ with

$$|I| \geqslant \sum_{v \in R} \frac{1}{1 + d_{E[R]}(v)} = \sum_{v \in W} \frac{\mathbb{1}[v \in R]}{1 + d_E(v, R)}.$$

By Fact 3.5, we have

$$\mathbb{E}[|I|] \geqslant \mathbb{E}\left[\sum_{v \in W} \frac{\mathbb{1}[v \in R]}{1 + d_E(v, R)}\right] = \sum_{v \in W} q \cdot \mathbb{E}\left[\frac{1}{1 + \mathrm{Bin}(d_E(v), q)}\right] \geqslant \sum_{v \in W} \frac{q}{1 + q d_E(v)}.$$

As the function $[0, \infty) \ni x \mapsto q/(1 + qx)$ is convex, the sum in the right-hand side above is minimised when $d_E(v) = 2e(E)/m$ for every $v \in W$. However, we assumed that $d_E(v) \leqslant (1-\alpha)\ell^2/m$ for every $v \in L$, so a slightly stronger bound holds. Indeed, since

$$2e(E) \leqslant \ell^2 = \alpha m \cdot (1-\alpha)\frac{\ell^2}{m} + (1-\alpha)m \cdot \left(\frac{1}{1-\alpha} - \alpha\right)\frac{\ell^2}{m},$$

then it follows that

$$\mathbb{E}[|I|] \geqslant \frac{\alpha m \cdot q}{1 + q \cdot (1-\alpha)\ell^2/m} + \frac{(1-\alpha)m \cdot q}{1 + q \cdot (1/(1-\alpha) - \alpha)\ell^2/m} \geqslant \left(1 + \frac{\alpha^3}{2}\right)\frac{m^2}{\ell^2}. \qquad (25)$$

One may verify the the last inequality in (25) by multiplying the numerators and the denominators in the left-hand side by $m/(\ell^2 q) = 1/C = \alpha^3/4$ and observing that

$$\frac{\alpha}{1-\alpha} + \frac{1-\alpha}{1/(1-\alpha) - \alpha} = 1 + \frac{\alpha^3}{(1-\alpha)(1-\alpha+\alpha^2)} \geqslant 1 + \alpha^3.$$

Set $d = \lfloor(1-2\delta)\ell\rfloor$ and choose, for each vertex $v \in W$, an arbitrary set M_v of d edges of M that are incident to v. As before, we shall say that a copy of $K_{1,2}$ is good if its centre v lies in I, both of its edges are in M_v, and the pair of its non-centre vertices is not in E. As E is not $(\varepsilon/2)$-close to K_ℓ, then for every $v \in W$, the set $\widehat{M_v}$ of the d other endpoints of the edges in M_v contains at least $\binom{d}{2} - (1-\varepsilon/2)\binom{\ell}{2}$ pairs that do not belong to E. In particular, as $\delta \leqslant \varepsilon/16$, each vertex of I is the centre of at least $\varepsilon\ell^2/8$ good $K_{1,2}$s. Unfortunately, this lower bound is not sufficiently strong for the naive argument given in the proof of Proposition 3.3 to work, as $\mathbb{E}[|I|]$ is too small. Instead, we shall exploit the rough structure of E.

To this end, we partition the set W into sets W_L and W_H of low and high degree vertices, which are defined as follows:

$$W_H := \left\{v \in W : d_E(v) \geqslant \delta\ell/2\right\} \qquad \text{and} \qquad W_L := W \setminus W_H.$$

Given an independent set I, we split it into I_L and I_H, which are defined as follows:

$$I_H := \left\{v \in I : |\widehat{M_v} \cap W_H| \geqslant \delta\ell\right\} \qquad \text{and} \qquad I_L := I \setminus I_H.$$

Observe that if $v \in I_L$, then $\widehat{M_v}$ contains at least $\binom{d-\delta\ell}{2} - \delta\ell^2/2 \geqslant (1-7\delta)\ell^2/2$ pairs that do not belong to E. We shall argue differently for different I, depending on the relative sizes of the sets I_L and I_H.

In both cases, we will find a (random) collection of at least $\delta^2 m^2/16$ good C_4s (in expectation) each of which is the union of two $K_{1,2}$s centred at some $v, w \in I$ and such that neither of (the pairs of edges of) these $K_{1,2}$s belongs to $\mathcal{C}$. We first argue that this is sufficient. Indeed, even though we will still have to disregard those copies of C_4 that contain a $K_{1,2}$ with two non-centre vertices in I whose edges belong to $\mathcal{C}$, the expected number of such saturated $K_{1,2}$s is at most $q^2 \cdot |\mathcal{C}|$ and each of them lies in at most $d \leqslant \ell$ of our good copies of C_4, as the edges of these good C_4s came only from the sets M_v. Hence, letting Z be the (random) number of good C_4s that contain at least two vertices of I and no $K_{1,2}$ whose edges belong to $\mathcal{C}$, we will have

$$\mathbb{E}[Z] \geqslant \frac{\delta^2 m^2}{16} - q^2 \cdot |\mathcal{C}| \cdot \ell \geqslant \frac{\delta^2 m^2}{16} - 12C^2 \beta m^2 \geqslant \frac{\delta^2 m^2}{32}.$$

But as each good copy of C_4 containing no saturated $K_{1,2}$ has chance at most $2q^2$ to be counted by Z, the number of them is at least

$$\frac{\delta^2 m^2}{64q^2} = \frac{\delta^2 \ell^4}{64C^2} \geqslant 3\beta \ell^4.$$

Therefore, in order to complete the proof of the theorem in Case 1, it suffices to prove the existence of (a random collection of) $\delta^2 m^2/16$ good copies of C_4 (in expectation) of the less restrictive type described above.

Subcase 1A. $|I_H| \leqslant \delta|I|$.

Recall that if $v \in I_L$, then $\widehat{M}_v$ contains at least $(1 - 7\delta)\ell^2/2$ pairs that do not belong to E. It follows that the number X_g of good $K_{1,2}$s satisfies

$$X_g \geqslant (1 - 7\delta)\frac{\ell^2}{2} \cdot |I_L| \geqslant (1 - 8\delta)\frac{\ell^2}{2} \cdot |I|.$$

Writing again X_s for the number of saturated $K_{1,2}$s (those whose edges belong to $\mathcal{C}$) whose centre vertex belongs to I and X for the number of good $K_{1,2}$s that are not saturated, we have $X \geqslant X_g - X_s$ and consequently,

$$\mathbb{E}[X] \geqslant \mathbb{E}[X_g] - \mathbb{E}[X_s] \geqslant (1 - 8\delta)\frac{\ell^2}{2} \cdot \mathbb{E}[|I|] - q \cdot |\mathcal{C}|$$

$$\geqslant (1 - 8\delta)\left(1 + \frac{\alpha^3}{2}\right)\frac{m^2}{2} - 12C\beta\ell m \geqslant \left(1 + \frac{\alpha^3}{4}\right)\frac{m^2}{2},$$

where we used (25), the facts that $\delta < \alpha^3/2^7$ and $\beta < \alpha^3/(8 \cdot 24C)$, and the trivial inequality $m \geqslant (1 - 2\delta)\ell \geqslant \ell/2$. Since I is an independent set in E, any pair of good $K_{1,2}$s with the same non-centre vertices forms a good C_4. Thus we have at least $X - \binom{m}{2}$ such C_4 and

$$\mathbb{E}[X] - \binom{m}{2} \geqslant \frac{\alpha^3 m^2}{8} \geqslant \frac{\delta^2 m^2}{16},$$

as required.

Subcase 1B. $|I_H| > \delta|I|$.

Let us write X_g for the number of good $K_{1,2}$s with at least one non-centre vertex in W_H. We will show in this case that

$$X_g \geqslant \frac{\delta^2 \ell^2}{4} \cdot |I| \qquad \text{and} \qquad |W_H| \leqslant \frac{\delta^2 m}{16},$$

from which it will be straightforward (as in Subcase 1A) to deduce the existence of the required collection of good C_4s.

To prove the lower bound on X_g, recall first that each vertex $v \in I_H$ is the centre of at least $\varepsilon\ell^2/8$ good $K_{1,2}$s; we claim that at least $\delta\ell^2/4$ of these have at least one non-centre vertex in W_H. To prove this, set $w = |\widehat{M}_v \cap W_L|$ and suppose first that $w \leqslant \varepsilon\ell/2$. Then at most $\varepsilon^2\ell^2/8$ good $K_{1,2}$s centred at v have both non-centre vertices in W_L and since $\varepsilon/8 - \varepsilon^2/8 \geqslant \varepsilon/16 \geqslant \delta/4$, the claim follows in this case. On the other hand, if $w > \varepsilon\ell/2$, then there are at least

$$\min\left\{ w\left(d - w - \frac{\delta\ell}{2}\right) : \frac{\varepsilon\ell}{2} < w \leqslant d - \delta\ell \right\} \geqslant \frac{\delta\ell^2}{4}$$

good $K_{1,2}$s centred at v with at least one non-centre vertex in W_H. Indeed, since $|\widehat{M}_v| = d$ and each $u \in \widehat{M}_v \cap W_L$ has degree at most $\delta\ell/2$ in E, there are at least $d - w - \delta\ell/2$ good $K_{1,2}$s centred at v that contain u and a third vertex from W_H. Thus

$$X_g \geqslant \frac{\delta\ell^2}{4} \cdot |I_H| \geqslant \frac{\delta^2\ell^2}{4} \cdot |I|,$$

as claimed. To prove the claimed upper bound on $|W_H|$, observe that

$$\ell^2 \geqslant 2e(E) \geqslant \sum_{v \in W_H} d_E(v) \geqslant |W_H| \cdot \frac{\delta\ell}{2}$$

which implies, by (24), that

$$|W_H| \leqslant \frac{2\ell}{\delta} \leqslant \frac{\delta^2 m}{16},$$

as required. Now, writing X for the number of good $K_{1,2}$s with a non-centre vertex in W_H that are moreover not saturated and X_s for the number of saturated $K_{1,2}$s (that is, $K_{1,2}$s whose pair of edges belongs to $\mathcal{C}$) whose centre vertex belongs to I, we have $X \geqslant X_g - X_s$ and hence,

$$\mathbb{E}[X] \geqslant \mathbb{E}[X_g] - \mathbb{E}[X_s] \geqslant \frac{\delta^2\ell^2}{4} \cdot \mathbb{E}[|I|] - q \cdot |\mathcal{C}| \geqslant \frac{\delta^2 m^2}{4} - 12C\beta\ell m \geqslant \frac{\delta^2 m^2}{8},$$

where we again used the bounds $\beta < \delta^2/(8 \cdot 24C)$ and $m \geqslant (1 - 2\delta)\ell \geqslant \ell/2$.

Finally, since I is an independent set in E, it follows that there are at least $X - |W_H|m$ good C_4s formed by pairs of $K_{1,2}$s that are counted by X and hence the expected number of good copies of C_4 that are formed by two $K_{1,2}$s centred at vertices in I, neither of which belongs to $\mathcal{C}$, is at least

$$\mathbb{E}[X] - |W_H|m \geqslant \frac{\delta^2 m^2}{16},$$

as required. This completes the proof in Case 1.

Case 2. There are fewer than αm vertices v satisfying $d_E(v) \leqslant (1 - \alpha)\ell^2/m$.

In this case, we shall find our good C_4s in various ways, depending on the distribution of degrees (in both the graphs M and E) on the set A of vertices whose degree in E is somewhat larger than average. To be precise, set $\gamma = 1/32$ and define

$$A := \left\{ v \in W : d_E(v) \geqslant (1 + \gamma)\ell^2/m \right\} \qquad \text{and} \qquad B := W \setminus A. \qquad (26)$$

We claim that $e_E(A) \leqslant \gamma^2 \ell^2$. To prove this, observe first that

$$2e(E) = \sum_{v \in W} d_E(v) \geqslant |A| \cdot (1 + \gamma) \frac{\ell^2}{m} + \big((1 - \alpha)m - |A|\big)\big(1 - \alpha\big)\frac{\ell^2}{m}$$

$$= \left((1 - \alpha)^2 + (\gamma + \alpha) \cdot \frac{|A|}{m}\right)\ell^2.$$

Noting that $\alpha = \gamma^3/2$, and recalling that $e(E) \leqslant \binom{\ell}{2}$, it follows that

$$\frac{|A|}{m} \leqslant \frac{\alpha(2 - \alpha)}{\gamma + \alpha} \leqslant \gamma^2,$$

and therefore

$$\sum_{v \in A} d_E(v) = 2e(E) - \sum_{v \in B} d_E(v) \leqslant \ell^2 - \big((1 - \alpha)m - |A|\big)\big(1 - \alpha\big)\frac{\ell^2}{m}$$

$$\leqslant \big(2\alpha + \gamma^2\big)\ell^2 \leqslant 2\gamma^2 \ell^2,$$

so in particular $e_E(A) \leqslant \gamma^2 \ell^2$, as claimed.

We next use Propositions 3.3 and 3.6 to show that we may assume that $e_M(A) < 9\gamma^3 \ell m$ and $e_M(A, B) < \ell m/2$. Indeed, if $e_M(A) \geqslant 9\gamma^3 \ell m$, then let us fix an arbitrary superset A' of A with exactly $\gamma^2 m$ elements and apply Proposition 3.3 to the pregraph $\mathcal{P}$ restricted to the set A', with $\ell_{3.3} \leftarrow 3\gamma\ell$ and $n_{3.3} \leftarrow \gamma^2 m$. To see that the conditions of the proposition are satisfied, note that

$$e_E(A') \leqslant \gamma^2 \ell^2 + \gamma^2 m \cdot (1 + \gamma)\frac{\ell^2}{m} \leqslant \binom{3\gamma\ell}{2} \quad \text{and} \quad e_M(A') \geqslant 9\gamma^3 \ell m = 3(3\gamma\ell)(\gamma^2 m)$$

and that $(3\gamma\ell)^2 \geqslant 9\gamma^2 n \geqslant \gamma^2 m$ and $|\mathcal{C}| \leqslant 12\beta\ell^3 \leqslant (3\gamma\ell)^3/40$. The proposition provides $(3\gamma\ell)^4/40$ good copies of C_4 that contain no pair from $\mathcal{C}$, and so in this case we are done. Similarly, if $e_M(A, B) \geqslant \ell m/2$, then, noting that

$$e_E(A) \leqslant \gamma^2 \ell^2 < \frac{\ell^2}{640}, \qquad |A| \leqslant \gamma^2 m < \frac{m}{640} \qquad \text{and} \qquad |\mathcal{C}| \leqslant 12\beta\ell^3 < \frac{\ell^3}{640},$$

we may invoke Proposition 3.6 to find $\ell^4/640$ good copies of C_4, none of which contains a pair from $\mathcal{C}$, and so in this case we are also done. We may therefore assume from now on that $e_M(A) < 9\gamma^3 \ell m$ and $e_M(A, B) < \ell m/2$, and hence that

$$\sum_{v \in B} d_M(v) = 2e(M) - 2e_M(A) - e_M(A, B) \geqslant \left(2(1 - \delta) - 18\gamma^3 - \frac{1}{2}\right)\ell m \geqslant \frac{4\ell m}{3}. \quad (27)$$

For the rest of the proof, we will search for good C_4s formed by two $K_{1,2}$s whose centre vertices belong to B. Let us say that a copy of $K_{1,2}$ in M is good if its centre lies in B and the pair of its non-centre vertices does not belong to E. Observe that for each $v \in B$, letting $N_M(v)$ denote the M-neighbourhood of v, we have

$$e_E\big(N_M(v)\big) \leqslant e_E(A) + \sum_{w \in N_M(v) \cap B} d_E(w)$$

$$\leqslant \gamma^2 \ell^2 + d_M(v) \cdot (1 + \gamma)\frac{\ell^2}{m} \leqslant \frac{\gamma}{2} \cdot d_M(v)^2, \quad (28)$$

since $d_M(v) \geqslant \delta(M) \geqslant \ell/2$ and $\ell/m \leqslant \gamma/16$ by (24). We therefore have at least $(1/2 - \gamma)d_M(v)^2$ good $K_{1,2}$s centred at v, for each $v \in B$.

It only remains to bound the number of good C_4s composed of two good $K_{1,2}$s, and remove those that contain a pair from $\mathcal{C}$. Our strategy will be similar to that used above, but there are two additional problems to overcome in this case: the set B is not an independent set and we do not have an upper bound on the degrees $d_M(v)$. To deal with the first problem, we will use our upper bound on $d_E(v)$ for $v \in B$, together with a slightly more careful application of convexity than was needed earlier in the proof. To deal with the second issue, we will partition B according to the approximate size of $d_M(v)$ and restrict our search to one of the parts.

We first partition B into two parts, depending (roughly speaking) on whether or not $d_M(v) = O(\ell)$. Define

$$B_L := \left\{ v \in B : d_M(v) \leqslant 2^{20}\ell \right\} \qquad \text{and} \qquad B_H := B \setminus B_L.$$

We first consider the case in which sufficiently many of the mixed edges incident to B have an endpoint in B_L.

Subcase 2A.
$$\sum_{v \in B_L} d_M(v) \geqslant \frac{5\ell m}{4}. \tag{29}$$

Let X denote the number of good $K_{1,2}$s whose centre vertex lies in B_L and whose pair of edges does not belong to the family $\mathcal{C}$. By (28), we have

$$X \geqslant \sum_{v \in B_L} \left(\binom{d_M(v)}{2} - \frac{\gamma}{2} d_M(v)^2 \right) - |\mathcal{C}|$$

$$\geqslant \left(\frac{1}{2} - \gamma \right) \sum_{v \in B_L} d_M(v)^2 - 12\beta\ell^3 \geqslant \frac{1}{3} \sum_{v \in B_L} d_M(v)^2,$$

since by the Cauchy–Schwarz inequality and (29),

$$\sum_{v \in B_L} d_M(v)^2 \geqslant \frac{1}{|B_L|} \left(\sum_{v \in B_L} d_M(v) \right)^2 \geqslant \frac{25}{16} \cdot \ell^2 m. \tag{30}$$

Let Y denote the number of (ordered) pairs of $K_{1,2}$s that are counted by X and have the same non-centre vertices. By the convexity of the function $x \mapsto x(x-1)$ and by (30), we have

$$Y \geqslant X \left(\frac{X}{\binom{m}{2}} - 1 \right) \geqslant \frac{1}{3} \sum_{v \in B_L} d_M(v)^2 \cdot \left(\frac{25\ell^2}{8m} - 1 \right) \geqslant \frac{\ell^2}{m} \cdot \sum_{v \in B_L} d_M(v)^2,$$

since $\ell^2 \geqslant Cn \geqslant Cm$. Now, let us denote by Y_b the number of (ordered) pairs of $K_{1,2}$s counted by Y that do not correspond to good C_4s (that is, pairs of good $K_{1,2}$s with the same non-centre vertices, whose centre vertices are adjacent in E). By the definition (26) of B, this number satisfies

$$Y_b \leqslant \sum_{v \in B_L} d_E(v) \binom{d_M(v)}{2} \leqslant \left(\frac{1+\gamma}{2} \right) \frac{\ell^2}{m} \cdot \sum_{v \in B_L} d_M(v)^2.$$

Thus, writing Z_g for the number of good C_4s consisting of pairs of $K_{1,2}$s counted by Y and combining the last three displayed equations, we obtain

$$Z_g \geqslant \frac{Y - Y_b}{4} \geqslant \frac{1}{12} \cdot \frac{\ell^2}{m} \cdot \sum_{v \in B_L} d_M(v)^2 \geqslant \frac{\ell^4}{8}.$$

Finally, we must disregard those good C_4s, counted in Z_g, that contain a $K_{1,2}$ of mixed edges that belongs to the family $\mathcal{C}$. The edges of such a $K_{1,2}$ must come from different good $K_{1,2}$s counted by X and therefore (by the definition of B_L) there are at most $2^{20}\ell\cdot|\mathcal{C}|$ such C_4s. It follows that the number Z of good C_4s that contain no $K_{1,2}$s whose edges belong to $\mathcal{C}$ satisfies

$$Z \geqslant Z_g - 2^{20}\ell \cdot |\mathcal{C}| \geqslant \left(\frac{1}{8} - 2^{24}\beta\right)\ell^4 \geqslant 3\beta\ell^4,$$

as required.

Note that if (29) fails to hold, then $\sum_{v\in B_H} d_M(v) \geqslant (4/3 - 5/4)\ell m = \ell m/12$, by (27). In this case we will choose a subset of B_H on which the M-degrees are roughly constant and apply the same argument as in Subcase 2A.

Subcase 2B.
$$\sum_{v\in B_H} d_M(v) \geqslant \frac{\ell m}{12}. \tag{31}$$

For each integer $t \geqslant 0$, set

$$b_t = 2^{-4t-28}m \qquad \text{and} \qquad d_t = 2^{3t+20}\ell$$

and define

$$B_t = \left\{v \in B_H : d_t < d_M(v) \leqslant d_{t+1}\right\}.$$

We claim that there exists t such that $|B_t| \geqslant b_t$. Indeed, since $B_H = \bigcup_{t\geqslant 0} B_t$, if there were no such t, then we would have

$$\sum_{v\in B_H} d_M(v) < \sum_{t\geqslant 0} b_t d_{t+1} = \sum_{t\geqslant 0} \frac{\ell m}{2^{t+5}} < \frac{\ell m}{12},$$

contradicting (31). Fix any such t and let X denote the number of $K_{1,2}$s whose centre vertex lies in B_t, whose pair of non-centre vertices is not in E, and whose pair of edges does not belong to the family $\mathcal{C}$. Observe that

$$X \geqslant \sum_{v\in B_t} \left(\binom{d_M(v)}{2} - \frac{\gamma}{2}d_M(v)^2\right) - |\mathcal{C}| \geqslant b_t \cdot \left(\frac{1}{2} - \gamma\right)d_t^2 - |\mathcal{C}|$$

$$\geqslant \left(\frac{1}{2} - \gamma\right)2^{2t+12}\ell^2 m - 12\beta\ell^3 \geqslant 2^{2t+10}\ell^2 m.$$

As before, let Y denote the number of (ordered) pairs of $K_{1,2}$s that are counted by X and have the same non-centre vertices. By the convexity of the function $x \mapsto x(x-1)$, we have

$$Y \geqslant X\left(\frac{X}{\binom{m}{2}} - 1\right) \geqslant 2^{4t+20}\ell^4,$$

where we again used the assumption that $\ell^2 \geqslant Cm$. The number Y_b of (ordered) pairs counted by Y that do not correspond to good C_4s (that is, pairs of good $K_{1,2}$s whose centre vertices are adjacent in E) satisfies

$$Y_b \leqslant \sum_{v\in B_t} d_E(v)\binom{d_M(v)}{2} \leqslant b_t \cdot \left(\frac{1+\gamma}{2}\right)\frac{\ell^2}{m} \cdot d_{t+1}^2 \leqslant 2^{2t+18}\ell^4.$$

31

Thus, the number Z_g of good C_4s counted by Y satisfies

$$Z_g \geqslant \frac{Y - Y_b}{4} \geqslant \left(2^{4t+18} - 2^{2t+16}\right)\ell^4 \geqslant 2^{4t+17}\ell^4.$$

Finally, we disregard those good C_4s, counted in Z_g, that contain a $K_{1,2}$ of mixed edges that belongs to the family $\mathcal{C}$. For each element of $\mathcal{C}$, there are at most d_{t+1} such C_4s and therefore the number Z of good C_4s that contain no $K_{1,2}$s whose edges belong to $\mathcal{C}$ satisfies

$$Z \geqslant Z_g - d_{t+1} \cdot |\mathcal{C}| \geqslant \left(2^{4t+17} - 12\beta \cdot 2^{3t+23}\right)\ell^4 \geqslant \ell^4,$$

as required. This completes the proof of the theorem. $\qquad \square$

4. The number of split graphs and the non-structured regime

In this section, we prove assertions (a) and (b) of Theorem 1.2. We first establish two lower bounds on the cardinality of $\mathcal{F}_{n,m}^{\mathrm{ind}}(C_4)$: a stronger bound for all $m \ll n^{4/3}$ and a weaker bound for all $m \ll n^{4/3}(\log n)^{1/3}$. Second, we carefully estimate the number of split graphs with n vertices and m edges for all n and m with $n \ll m \ll n^2$. Third, we provide a simple upper bound on the number of graphs that are not ε-quasirandom. A straightforward comparison of these bounds yields the claimed results.

4.1. Lower bounds for $\mathcal{F}_{n,m}^{\mathrm{ind}}(C_4)$.

We first show that if $m \ll n^{4/3}$, then the family $\mathcal{F}_{n,m}^{\mathrm{ind}}(C_4)$ forms an $e^{-o(m)}$-proportion of all graphs with n vertices and m edges. In particular, as we shall later verify, if $m \gg n$, then for every fixed ε, graphs with no induced copy of C_4 outnumber the graphs that are not ε-quasirandom and thus a typical member of $\mathcal{F}_{n,m}^{\mathrm{ind}}(C_4)$ is ε-quasirandom.

Proposition 4.1. *For every $\gamma > 0$, there exists $\delta > 0$ such that for all sufficiently large n and all $m \leqslant \delta n^{4/3}$,*

$$\left|\mathcal{F}_{n,m}^{\mathrm{ind}}(C_4)\right| \geqslant e^{-\gamma m} \cdot \binom{\binom{n}{2}}{m}.$$

Proof. Fix a positive γ and choose $\delta > 0$ sufficiently small so that $17(1+\delta)^4\delta^3 < \delta/2$ and $(2e/\delta)^\delta < e^\gamma$. Suppose that $m \leqslant \delta n^{4/3}$, let $m' = \lfloor (1+\delta)m \rfloor$, and let G be the uniformly chosen random graph with vertex set $\{1, \ldots, n\}$ and precisely m' edges. Let X be the number of (not necessarily induced) copies of C_4 in G. As

$$\mathbb{E}[X] \leqslant n^4 \cdot (m')^4 \cdot \binom{n}{2}^{-4} \leqslant \frac{17(m')^4}{n^4} \leqslant 17(1+\delta)^4\delta^3 m \leqslant \frac{\delta m}{2},$$

Markov's inequality gives $\mathbb{P}(X \leqslant m' - m) = \mathbb{P}(X \leqslant \delta m) \geqslant 1/2$. In particular, at least half of all graphs with vertex set $\{1, \ldots, n\}$ and m' edges contain a subgraph with m edges and no copy of C_4. This implies that

$$\left|\mathcal{F}_{n,m}^{\mathrm{ind}}(C_4)\right| \geqslant \frac{1}{2} \cdot \binom{\binom{n}{2}}{m'}\binom{\binom{n}{2} - m}{m' - m}^{-1} = \frac{1}{2} \cdot \binom{\binom{n}{2}}{m}\binom{m'}{m}^{-1}.$$

Finally, by our assumption on δ,

$$\binom{m'}{m} \leqslant \binom{(1+\delta)m}{\delta m} \leqslant \left(\frac{e(1+\delta)}{\delta}\right)^{\delta m} \leqslant \frac{1}{2} \cdot e^{\gamma m}.$$

This completes the proof. $\qquad \square$

The derivation of our second lower bound on $|\mathcal{F}_{n,m}^{\text{ind}}(C_4)|$ follows a similar strategy, but the simple deletion argument is replaced with the following result of Kohayakawa, Kreuter, and Steger [32], stated here for the random graph $G_{n,m}$ rather than the binomial random graph $G(n,p)$. The heart of the proof of this theorem (which we shall not give here, but rather refer the reader to [32, Theorem 8] or to [24, Appendix A]) is a classical result of Ajtai, Komlós, Pintz, Spencer, and Szemerédi [1], or rather its corollary derived by Duke, Lefmann, and Rödl [19], that gives a lower bound on the independence number of a uniform hypergraph that contains few short cycles.

Theorem 4.2 ([32]). *There exists a constant c such that if $n^{4/3} \leqslant m \leqslant \binom{n}{2}$, then a.a.s.*

$$\text{ex}(G_{n,m}, C_4) \geqslant cn^{4/3}\big(\log\big(m/n^{4/3}\big)\big)^{1/3}.$$

Proposition 4.3. *For every $\gamma > 0$, there exists a $\delta > 0$ such that for all sufficiently large n and all $m \leqslant \delta n^{4/3}(\log n)^{1/3}$,*

$$\big|\mathcal{F}_{n,m}^{\text{ind}}(C_4)\big| \geqslant n^{-\gamma m} \cdot \binom{\binom{n}{2}}{m}.$$

Proof. Let c be the constant from the statement of Theorem 4.2. Given a positive γ, choose $\delta > 0$ sufficiently small so that $\delta \leqslant c(\gamma/2)^{1/3}$, let $m' = \lceil n^{4/3+\gamma/2} \rceil$, and observe that

$$cn^{4/3}\big(\log(m'/n^{4/3})\big)^{1/3} \geqslant c(\gamma/2)^{1/3}n^{4/3}(\log n)^{1/3} \geqslant \delta n^{4/3}(\log n)^{1/3}.$$

Suppose that $m \leqslant \delta n^{4/3}(\log n)^{1/3}$. It follows from Theorem 4.2 that at least half of all graphs with vertex set $\{1, \ldots, n\}$ and m' edges contain a subgraph with m edges and no copy of C_4, provided that n is sufficiently large. Therefore, similarly as in the proof of Proposition 4.1,

$$\big|\mathcal{F}_{n,m}^{\text{ind}}(C_4)\big| \geqslant \frac{1}{2} \cdot \binom{\binom{n}{2}}{m}\binom{m'}{m}^{-1} \geqslant \frac{1}{2} \cdot \Big(\frac{m}{em'}\Big)^m \binom{\binom{n}{2}}{m} \geqslant n^{-\gamma m} \cdot \binom{\binom{n}{2}}{m}.$$

This completes the proof. $\qquad\square$

4.2. **The number of split graphs.** As we shall need to compare the family of split graphs (and graphs that are close to a split graph) to various other families of graphs, we will need to derive some estimates on its cardinality. Let $\mathcal{S}_{n,m}$ denote the family of split graphs with vertex set $\{1, \ldots, n\}$ that have precisely m edges. Moreover, let $N_{n,m}(\ell)$ denote the number of those graphs that are complete on the set $\{1, \ldots, \ell\}$ and empty on its complement. Observe that

$$N_{n,m}(\ell) = \begin{cases} \binom{\ell(n-\ell)}{m-\binom{\ell}{2}}, & \text{if } \binom{\ell}{2} \leqslant m \leqslant \ell(n-\ell) + \binom{\ell}{2}, \\ 0, & \text{otherwise}, \end{cases}$$

and

$$\max_{\ell} N_{n,m}(\ell) \leqslant |\mathcal{S}_{n,m}| \leqslant \sum_{\ell} \binom{n}{\ell} N_{n,m}(\ell). \tag{32}$$

Since (32) is rather hard to work with due to its inexplicit form, we establish several asymptotic properties of the function $\ell \mapsto N_{n,m}(\ell)$, summarised in Proposition 4.4 below. We postpone the rather dull and technical proof of the proposition to Appendix A.

Proposition 4.4. *There is a positive constant λ such that the following holds for all sufficiently large n. If $n \ll m \leqslant \lambda n^2$, then the function $\ell \mapsto N_{n,m}(\ell)$ attains its maximum for some ℓ satisfying $\ell_{n,m}/2 < \ell < 2\ell_{n,m}$, where $\ell_{n,m}$ is defined by*

$$\ell_{n,m} = \left(\frac{m}{\log \left(\ell_{n,m} n/m \right)} \right)^{1/2}.$$

Moreover, $N_{n,m}(\ell_{n,m}) \geqslant 5^m$ and if $\ell \leqslant \ell_{n,m}/2$ or $\ell \geqslant 2\ell_{n,m}$, then

$$N_{n,m}(\ell) < \exp(-m/15) \cdot \max_{\ell} N_{n,m}(\ell).$$

4.3. The non-structured regime.

Proof of parts (a) and (b) of Theorem 1.2. Fix an arbitrary positive ε, suppose that $m \gg n$, and let G be the uniformly chosen random graph with vertex set $\{1, \ldots, n\}$ and exactly m edges. A standard averaging argument shows that if G is not ε-quasirandom, then it contains a subset A with exactly εn vertices and density differing from $m/\binom{n}{2}$ by more than $\varepsilon m/\binom{n}{2}$. Consequently, Hoeffding's inequality for the hypergeometric distribution [27] asserts the existence of a positive ρ that depends only on ε such that

$$\mathbb{P} \left(G \text{ is not } \varepsilon\text{-quasirandom} \right) \leqslant \binom{n}{\varepsilon n} \cdot \exp \left(-3\rho m \right).$$

It now follows from Proposition 4.1 invoked with $\gamma \leftarrow \rho$ that if δ is sufficiently small, then for all sufficiently large n and all m satisfying $n \ll m \leqslant \delta n^{4/3}$,

$$\mathbb{P} \left(G \text{ is not } \varepsilon\text{-quasirandom} \mid G \in \mathcal{F}_{n,m}^{\mathrm{ind}}(C_4) \right) \leqslant \frac{\mathbb{P} \left(G \text{ is not } \varepsilon\text{-quasirandom} \right)}{\mathbb{P} \left(G \in \mathcal{F}_{n,m}^{\mathrm{ind}}(C_4) \right)}$$

$$\leqslant \binom{n}{\varepsilon n} \cdot \exp(-3\rho m + \rho m) \leqslant \exp(-\rho m).$$

In other words, graphs that are not ε-quasirandom constitute only an exponentially small fraction of $\mathcal{F}_{n,m}^{\mathrm{ind}}(C_4)$.

Now, denote by $\mathcal{S}_{n,m}(\varepsilon)$ the family of graphs with vertex set $\{1, \ldots, n\}$ and m edges that are ε-close to a split graph. Each graph in $\mathcal{S}_{n,m}(\varepsilon)$ can be obtained from some graph in $\mathcal{S}_{n,m}$ by removing from it some εm edges and replacing them with arbitrarily chosen εm edges of K_n. Hence, if $m \gg n$ and n is sufficiently large, then

$$|\mathcal{S}_{n,m}(\varepsilon)| \leqslant |\mathcal{S}_{n,m}| \cdot \binom{m}{\varepsilon m} \cdot \binom{\binom{n}{2}}{\varepsilon m} \leqslant |\mathcal{S}_{n,m}| \cdot \left(\frac{em}{\varepsilon m} \cdot \frac{en^2}{2\varepsilon m} \right)^{\varepsilon m} \leqslant n^{\varepsilon m} \cdot |\mathcal{S}_{n,m}|. \qquad (33)$$

Moreover, it follows from (32) and Proposition 4.4 that

$$|\mathcal{S}_{n,m}| \leqslant 2^n \cdot \max_{\ell} N_{n,m}(\ell) = 2^n \cdot \max_{\ell \leqslant 2\ell_{n,m}} N_{n,m}(\ell)$$

$$\leqslant 2^n \cdot \binom{2\ell_{n,m} n}{m} \leqslant 2^n \cdot \left(\frac{2\ell_{n,m} n}{\binom{n}{2}} \right)^m \binom{\binom{n}{2}}{m}.$$

Suppose now that $n \ll m \leqslant \delta n^{4/3}(\log n)^{1/3}$. As $\ell_{n,m} \ll n^{2/3}$, it follows that

$$|\mathcal{S}_{n,m}(\varepsilon)| \leqslant n^{(\varepsilon - 1/3)m} \cdot \binom{\binom{n}{2}}{m}$$

34

for all sufficiently large n. Therefore, by Proposition 4.3 invoked with $\gamma = 1/24$ implies that if δ is sufficiently small, then

$$\mathbb{P}\left(G \in \mathcal{S}_{n,m}(1/4) \mid G \in \mathcal{F}_{n,m}^{\mathrm{ind}}(C_4)\right) \leqslant n^{(1/4-1/3)m} \cdot n^{\gamma m} = n^{-m/24}.$$

In other words, graphs that are 1/4-close to a split graph constitute only a super-exponentially small proportion of $\mathcal{F}_{n,m}^{\mathrm{ind}}(C_4)$, as required. $\qquad\square$

5. An approximate structural theorem

In this section, we shall use Theorems 1.4 and 3.1 to construct a collection of containers for the family $\mathcal{F}_{n,m}^{\mathrm{ind}}(C_4)$ whenever $n^{4/3}(\log n)^4 \leqslant m \ll n^2$. Our aim is to do this in such a way that all but a tiny proportion of the family will be covered by containers that describe predominantly graphs that are close to a split graph. To make this notion precise, let us say that a pregraph $\mathcal{P} = (M, E)$ on n vertices is an ε-*almost split pregraph* if there exists a partition $V(K_n) = U \cup W$ such that

$$e(E) \leqslant \binom{|U|}{2}, \qquad e_E(U) \geqslant (1-\varepsilon)\binom{|U|}{2}, \qquad \text{and} \qquad e_M(W) \leqslant 7\sqrt{\varepsilon}|U|n.$$

We will prove the following container theorem for sparse induced-C_4-free graphs. Recall from Section 1.2 that a graph G is contained in (described by) a pregraph $\mathcal{P} = (M, E)$ if $E \subseteq E(G) \subseteq E \cup M$.

Theorem 5.1. *For every $\varepsilon > 0$, there exists $\lambda > 0$ such that the following holds. For every $n \in \mathbb{N}$ and $n^{4/3}(\log n)^4 \leqslant m \leqslant \lambda n^2$, there exists a collection $\mathcal{C}$ of ε-almost split pregraphs on n vertices with $|\mathcal{C}| = e^{o(m)}$ such that all but at most $e^{-\lambda m} \cdot |\mathcal{F}_{n,m}^{\mathrm{ind}}(C_4)|$ of the graphs in $\mathcal{F}_{n,m}^{\mathrm{ind}}(C_4)$ are contained in some $\mathcal{P} \in \mathcal{C}$.*

To prove Theorem 5.1, we will apply Theorem 1.4 recursively, starting with the trivial container, which is defined by the 'complete' pregraph with $M = E(K_n)$ (and therefore E empty). We continue until we obtain a family of containers, each of which admits only few good copies of C_4; we will be able to control this process with the use of Theorem 3.1, which provides us with a precise structural description of such pregraphs. Finally, we will show that the containers that are not ε-almost split pregraphs contain at most $e^{-\lambda m} \cdot |\mathcal{F}_{n,m}^{\mathrm{ind}}(C_4)|$ members of $\mathcal{F}_{n,m}^{\mathrm{ind}}(C_4)$.

More formally, we shall build a rooted tree $\mathcal{T}$ whose vertices are pregraphs with n vertices. The root of $\mathcal{T}$ is the pregraph with $M = E(K_n)$ corresponding to the trivial container. The children (in $\mathcal{T}$) of a pregraph will correspond to refinements of it that we obtain by applying Theorem 1.4 to one of the hypergraphs $\mathcal{H}_i$ supplied by Theorem 3.1. This way, each graph in $\mathcal{F}_{n,m}^{\mathrm{ind}}(C_4)$ that is described by some pregraph $\mathcal{P}$ in $\mathcal{T}$ will be described by one of the children of $\mathcal{P}$ in $\mathcal{T}$. As a consequence, each graph in $\mathcal{F}_{n,m}^{\mathrm{ind}}(C_4)$ will be accounted for by one of the leaves of $\mathcal{T}$.

In order to decide whether a pregraph $\mathcal{P} = (M, E)$ should be a leaf of the tree or not (in which case we will apply Theorem 1.4 to it), we use the following definition.

Definition 5.2. A pregraph $\mathcal{P} = (M, E)$ on n vertices is a *leaf pregraph* (with respect to m, ε, and δ) if either $\mathcal{P}$ is an ε-almost split pregraph, or there exists $\ell \in \mathbb{N}$ such that

$$e(E) \geqslant \binom{\ell}{2} \qquad \text{and} \qquad e(M) \leqslant (1-\delta)\ell n, \tag{34}$$

or either of the following holds:

$$e(E) > m \quad \text{or} \quad e(M) < \left(\frac{n^2 m}{2^8 \log(n^2/m)} \right)^{1/2}. \tag{35}$$

Recall that, given a pregraph $\mathcal{P} = (M, E)$, the $(i, 4)$-uniform hypergraph $\mathcal{H}_i^{\mathcal{P}}$ comprises all pairs (A, B) such that B is a good copy of C_4 in $\mathcal{P}$ and A is the set of the remaining i mixed edges induced by the vertex set of this copy (which induces exactly $4 + i$ edges of M). Also, with foresight, let us set

$$r = \frac{m}{2^{13} \log n}. \tag{36}$$

We will use Theorem 3.1 to prove the following lemma.

Lemma 5.3. *For every $\varepsilon > 0$, there exist positive constants β, δ, and λ such that the following holds for every $n \in \mathbb{N}$ and $n(\log n)^2 \leqslant m \leqslant \lambda n^2$. Let $\mathcal{P} = (M, E)$ be a pregraph on n vertices that is not a leaf pregraph with respect to m, ε, and δ. Then there exist an integer ℓ with $\ell^2 \geqslant r$ and a hypergraph $\mathcal{H} \subseteq \mathcal{H}_i^{\mathcal{P}}$, for some $i \in \{0, 1, 2\}$, such that*

$$v(\mathcal{H}) \leqslant 5\ell n, \quad e(\mathcal{H}) \geqslant \beta \ell^4, \quad \Delta_{(0,1)}(\mathcal{H}) \leqslant \frac{\ell^3}{n}, \quad \text{and} \quad \Delta_{(0,2)}(\mathcal{H}) \leqslant \ell \tag{37}$$

and, if $i > 0$, then also $\Delta_{(1,0)}(\mathcal{H}) \leqslant \ell^2$.

Proof. Let $\beta_{3.1}$, $\delta_{3.1}$, $\lambda_{3.1}$, and $C_{3.1}$ be the constants given by Theorem 3.1 applied with $\varepsilon_{3.1} \leftarrow \varepsilon$, set $\beta = \beta_{3.1}$, $C = C_{3.1}$, $\delta = \delta_{3.1}/2$, and $\lambda = 2^{-8}(\lambda_{3.1}/C)^2$. We may assume that $C \geqslant 1$, $\delta \leqslant 1/4$, $\lambda \leqslant \delta^2$, and (by our bounds on m) that $n \geqslant 1/\lambda$.

Suppose first that there exists $\ell \geqslant C\sqrt{n}$ such that

$$e(E) \leqslant \binom{\ell}{2} \quad \text{and} \quad e(M) \geqslant 4\ell n \tag{38}$$

and choose $\ell \in \mathbb{N}$ maximal such that $e(M) \geqslant 4\ell n$. We claim that $\ell^2 \geqslant r$. Indeed, $\mathcal{P}$ is not a leaf pregraph and thus the maximality of ℓ and the second inequality in (35) give

$$2r \leqslant \frac{m}{2^{12} \log(n^2/m)} \leqslant \frac{e(M)^2}{16n^2} \leqslant (\ell + 1)^2.$$

In this case it follows immediately from Theorem 3.1 that there exists a hypergraph $\mathcal{H}$ with the claimed properties.

Next, suppose that there exists $\ell \geqslant C\sqrt{n}$ and a set U of size ℓ such that

$$e(E) \leqslant \binom{\ell}{2} \quad \text{and} \quad e_E(U) \geqslant (1 - \varepsilon)\binom{\ell}{2}. \tag{39}$$

Note that $e(M) < 4\ell n$, otherwise (38) holds and we are done as above. Since $\mathcal{P}$ is not a leaf pregraph, it follows that $\ell^2 \geqslant r$, as above, and $e_M(U^c) > 7\sqrt{\varepsilon}\ell n$, as $\mathcal{P}$ is not an ε-almost split pregraph. This means that $\mathcal{P}$ satisfies condition (M3) of Theorem 3.1 and so we obtain a hypergraph $\mathcal{H}$ with the claimed properties, as before.

Finally, let $\ell \in \mathbb{N}$ be minimal such that $e(M) \leqslant (1 - \delta)\ell n$ and observe that $e(E) \leqslant \binom{\ell}{2}$, since $\mathcal{P}$ is not a leaf pregraph, and that

$$\ell \geqslant \frac{e(M)}{(1 - \delta)n} \geqslant \left(\frac{m}{2^8 \log(n^2/m)} \right)^{1/2} \geqslant \max\left\{ C\sqrt{n}, \sqrt{r} \right\},$$

where the second inequality follows since $\mathcal{P}$ is not a leaf pregraph and the third by our bounds on m, since n is sufficiently large. It follows that E is not ε-close to K_ℓ, since if it were, then there would exist a set U of size ℓ such that $e_E(U) \geqslant (1 - \varepsilon)\binom{\ell}{2}$, in which case (39) would hold and we would be done as before. Note also that $e(M) \geqslant (1 - 2\delta)\ell n$, by our choice of ℓ and since $\delta\ell \geqslant \delta\sqrt{n} \geqslant \delta/\sqrt{\lambda} \geqslant 1$.

Now, observe that if (38) fails to hold, then either $e(M) \leqslant 4Cn^{3/2}$ or

$$ e(M) \leqslant 8n\sqrt{e(E)} \leqslant 8n\sqrt{m} \leqslant 8\sqrt{\lambda}n^2, $$

where in the second step we used the fact that $e(E) \leqslant m$ (which holds if $\mathcal{P}$ is not a leaf pregraph) and in the third we used our upper bound on m. In either case, it follows that $\ell \leqslant 2e(M)/n \leqslant \lambda_{3.1}n$, since $\lambda = 2^{-8}(\lambda_{3.1}/C)^2$ and $C \geqslant 1$. Hence $\mathcal{P}$ satisfies condition (M2) of Theorem 3.1 and we again obtain the desired hypergraph $\mathcal{H}$. This completes the proof of the lemma. $\qquad\square$

We next combine Theorem 1.4 and Lemma 5.3 to construct a rooted tree whose leaves correspond to a family of containers for the family $\mathcal{F}_{n,m}^{\mathrm{ind}}(C_4)$.

Lemma 5.4. *For every $\varepsilon > 0$, there exist positive constants δ and λ such that the following holds. For every $n \in \mathbb{N}$ and $n^{4/3}(\log n)^4 \leqslant m \leqslant \lambda n^2$, there exists a collection $\mathcal{C}$ of $e^{o(m)}$ pregraphs on n vertices such that*

(a) every $\mathcal{P} \in \mathcal{C}$ is a leaf pregraph with respect to m, ε, and δ and
(b) every graph $G \in \mathcal{F}_{n,m}^{\mathrm{ind}}(C_4)$ is contained in some $\mathcal{P} \in \mathcal{C}$.

Proof. We will construct a rooted tree $\mathcal{T}$ whose vertices are pregraphs on n vertices that has the following properties:

(i) the root of $\mathcal{T}$ is the complete pregraph with $M = E(K_n)$;
(ii) if $G \in \mathcal{F}_{n,m}^{\mathrm{ind}}(C_4)$ is contained in a pregraph $\mathcal{P} \in V(\mathcal{T})$ that is not a leaf of $\mathcal{T}$, then G is contained in some child of $\mathcal{P}$ in $\mathcal{T}$;
(iii) the height of $\mathcal{T}$ is $O(\log n)$;
(iv) the maximum degree of $\mathcal{T}$ is $\exp\big(o(m/\log n)\big)$;
(v) every leaf of $\mathcal{T}$ is a leaf pregraph with respect to m, ε, and δ.

It will then follow immediately that the leaves of $\mathcal{T}$ form a collection $\mathcal{C}$ as required.

To define the children of a vertex $\mathcal{P} \in V(\mathcal{T})$, we will apply Theorem 1.4 to the hypergraph given by Lemma 5.3. To begin, let $\beta = \beta_{5.3}$, $\delta = \delta_{5.3}$, and $\lambda = \lambda_{5.3}$ be the constants given by Lemma 5.3 applied with $\varepsilon_{5.3} \leftarrow \varepsilon$ and set $\xi(n) = (\log\log n)^{-1}$ (here we could use any function that tends to zero sufficiently slowly as $n \to \infty$). Note that, due to the form of the statement, we may assume throughout that n is sufficiently large.

Let $\mathcal{P} \in V(\mathcal{T})$ and suppose that $\mathcal{P}$ is not a leaf pregraph with respect to m, ε, and δ. By Lemma 5.3, there exist $\ell \in \mathbb{N}$ with $\ell^2 \geqslant r$ and an $(i, 4)$-uniform hypergraph $\mathcal{H} \subseteq \mathcal{H}_i^{\mathcal{P}}$, for some $i \in \{0, 1, 2\}$, satisfying the assertion of the lemma. We claim that we may apply Theorem 1.4 to the hypergraph $\mathcal{H}$ with

$$ K = \frac{5}{\beta} \qquad \text{and} \qquad b = \xi(n) \cdot \frac{m}{(\log n)^2}, $$

and r as defined in (36). To do so, we need to verify that (1) is satisfied for every pair $(\ell_0, \ell_1) \in \{0, \dots, i\} \times \{0, \dots, 4\}$ with $(\ell_0, \ell_1) \neq (0, 0)$.

Claim. For every $(\ell_0, \ell_1) \in \{0, \ldots, i\} \times \{0, \ldots, 4\}$ with $(\ell_0, \ell_1) \neq (0,0)$, we have

$$\Delta_{(\ell_0,\ell_1)}(\mathcal{H}) \leqslant K \cdot \frac{b^{\ell_0+\ell_1-1}}{m^{\ell_0} \cdot v(\mathcal{H})^{\ell_1}} \cdot e(\mathcal{H}) \cdot \left(\frac{m}{r}\right)^{\mathbb{1}[\ell_0>0]}. \tag{40}$$

Proof of claim. Observe that the right-hand side of (40) decreases when ℓ_0 or ℓ_1 increase, since $b \leqslant r \leqslant m$ and $v(\mathcal{H}) = e(M) \geqslant b$, the latter holding (with room to spare) since $\mathcal{P}$ is not a leaf pregraph and $m \leqslant \lambda n^2$. Assume first that $\ell_0 \geqslant 2$ or $\ell_1 \geqslant 3$ and note that in this case $\Delta_{(\ell_0,\ell_1)}(\mathcal{H}) \leqslant 1$. It thus suffices to show that the right-hand side of (40) is at least 1. Since $v(\mathcal{H}) \leqslant 5\ell n$ and $e(\mathcal{H}) \geqslant \beta \ell^4$, see (37), we have

$$K \cdot \frac{b^5}{m^2(5\ell n)^4} \cdot \beta \ell^4 \cdot \frac{m}{r} \geqslant \xi(n)^6 \cdot \frac{m^3}{n^4(\log n)^9} \geqslant 1,$$

since $m \geqslant n^{4/3}(\log n)^4$ and n is sufficiently large.

Next, recall that $\Delta_{(0,1)}(\mathcal{H}) \leqslant \ell^3/n$, by (37), and observe that if $(\ell_0, \ell_1) = (0,1)$, then the right-hand side of (40) is at least

$$K \cdot \frac{e(\mathcal{H})}{v(\mathcal{H})} \geqslant K \cdot \frac{\beta \ell^4}{5\ell n} \geqslant \frac{\ell^3}{n},$$

as required. Similarly, if $i \geqslant 1$ then $\Delta_{(1,0)}(\mathcal{H}) \leqslant \ell^2$, by (37), and if $(\ell_0, \ell_1) = (1,0)$, then the right-hand side of (40) is at least

$$K \cdot \frac{e(\mathcal{H})}{r} \geqslant K \cdot \frac{\beta \ell^4}{r} \geqslant \ell^2,$$

since $\ell^2 \geqslant r$. Finally, note that $\Delta_{(1,1)}(\mathcal{H}) = \Delta_{(1,2)}(\mathcal{H}) \leqslant \Delta_{(0,2)}(\mathcal{H}) \leqslant \ell$, by (37). In particular, if $(\ell_0, \ell_1) \in \{(1,1), (1,2), (0,2)\}$, then the right-hand side of (40) is at least

$$K \cdot \frac{b^2}{m \cdot (5\ell n)^2} \cdot \beta \ell^4 \cdot \frac{m}{r} \geqslant \xi(n)^2 \cdot \frac{m\ell^2}{n^2(\log n)^3} \geqslant \ell,$$

since $m\ell \geqslant m\sqrt{r} \geqslant m^{3/2}/\log n \geqslant n^2(\log n)^4$. $\qquad \square$

Observe that $\mathcal{F}^{\mathrm{ind}}_{n,m}(C_4) \cap \mathcal{P} \subseteq \mathcal{F}_{\leqslant m}(\mathcal{H})$, since each $G \in \mathcal{F}^{\mathrm{ind}}_{n,m}(C_4) \cap \mathcal{P}$ has m edges and each constraint in $\mathcal{H}$ corresponds to an induced copy of C_4. Therefore, by Theorem 1.4, there exists a collection $N(\mathcal{P})$ of at most

$$\binom{\binom{n}{2}}{2b}\binom{\binom{n}{2}}{4b} \leqslant \exp\left(12b \cdot \log n\right) = \exp\left(o\left(\frac{m}{\log n}\right)\right)$$

sub-pregraphs[12] $\mathcal{Q}$ of $\mathcal{P}$ with the following properties:

(a') if $\mathcal{Q} \in N(\mathcal{P})$, then either $M(\mathcal{Q}) \leqslant (1-c)M(\mathcal{P})$ or $E(\mathcal{Q}) \geqslant E(\mathcal{P}) + cr$, and

(b') each $G \in \mathcal{F}^{\mathrm{ind}}_{n,m}(C_4) \cap \mathcal{P}$ is contained in some $\mathcal{Q} \in N(\mathcal{P})$,

where $c = 2^{-42}K^{-1}$. We make $N(\mathcal{P})$ the set of children of $\mathcal{P}$ in $\mathcal{T}$, observing that the degree of $\mathcal{P}$ in $\mathcal{T}$ is $\exp\left(o(m/\log n)\right)$. By (a') and Definition 5.2, the height of the tree $\mathcal{T}$ obtained in this way is at most

$$\frac{1}{c} \log \binom{n}{2} + \frac{m}{cr} \leqslant 2^{42}K\left(2\log n + 2^{13}\log n\right) = O\left(\log n\right).$$

It follows that the total number of leaves of $\mathcal{T}$ is $e^{o(m)}$ and hence (by the definition of $\mathcal{T}$ and property (b')) the collection of leaves of $\mathcal{T}$ forms a family $\mathcal{C}$ as required. $\qquad \square$

---

[12]This means that $M(\mathcal{Q}) \subseteq M(\mathcal{P})$ and $E(\mathcal{P}) \subseteq E(\mathcal{Q}) \subseteq M(\mathcal{P}) \cup E(\mathcal{P})$.

To deduce Theorem 5.1, we will show that the containers $\mathcal{P} \in \mathcal{C}$ that are not ε-almost split pregraphs contain only an exponentially small proportion of the family $\mathcal{F}^{\mathrm{ind}}_{n,m}(C_4)$

Proof of Theorem 5.1. Let $\delta_{5.4}$, $\lambda_{5.4}$, and $\mathcal{C}_{5.4}$ be (respectively) the constants and the family of containers given by Lemma 5.4 applied with $\varepsilon_{5.4} \leftarrow \varepsilon$, let $\lambda_{4.4}$ be the constant given by Proposition 4.4, and set $\delta = \delta_{5.4}$, $\lambda = \min\{\lambda_{4.4}, \lambda_{5.4}, 2^{-8}\delta^2\}$, and $\mathcal{C}' = \mathcal{C}_{5.4}$. Note that we may assume (without loss of generality) that $\delta \leqslant 1$ and recall that $|\mathcal{C}'| = e^{o(m)}$. We claim that the collection

$$\mathcal{C} := \big\{\mathcal{P} \in \mathcal{C}' : \mathcal{P} \text{ is an } \varepsilon\text{-almost split pregraph}\big\}$$

has the property that all but at most $e^{-\lambda m} \cdot |\mathcal{F}^{\mathrm{ind}}_{n,m}(C_4)|$ of the graphs in $\mathcal{F}^{\mathrm{ind}}_{n,m}(C_4)$ are contained in some $\mathcal{P} \in \mathcal{C}$ and therefore $\mathcal{C}$ is the required family of 'almost' containers.

To prove this, we will give an upper bound on the number of graphs in $\mathcal{F}^{\mathrm{ind}}_{n,m}(C_4)$ that belong to a single container $\mathcal{P} \in \mathcal{C}' \setminus \mathcal{C}$. Recall that every $\mathcal{P} \in \mathcal{C}'$ is a leaf pregraph with respect to m, ε, and δ and therefore we may assume that $\mathcal{P} = (M, E)$ satisfies either (34) or (35).

Case 1. Either $e(E) > m$ or $e(M) < 2^{-4}\sqrt{n^2 m / \log(n^2/m)}$.

We may assume that $e(E) \leqslant m$, as otherwise $\mathcal{F}^{\mathrm{ind}}_{n,m}(C_4) \cap \mathcal{P}$ is empty. Therefore

$$|\mathcal{F}^{\mathrm{ind}}_{n,m}(C_4) \cap \mathcal{P}| \leqslant \binom{e(M)}{m - e(E)} \leqslant \binom{e(M) + e(E)}{m} \leqslant \binom{2^{-4}\sqrt{n^2 m / \log(n^2/m)} + m}{m}$$

and hence, since $m \leqslant 2^{-4}\sqrt{n^2 m / \log(n^2/m)}$ for every $m \leqslant \lambda n^2$, we obtain

$$|\mathcal{F}^{\mathrm{ind}}_{n,m}(C_4) \cap \mathcal{P}| \leqslant \binom{2^{-3}\sqrt{n^2 m / \log(n^2/m)}}{m}.$$

We claim that for some well-chosen $\ell \in \mathbb{N}$,

$$\binom{2^{-3}\sqrt{n^2 m / \log(n^2/m)}}{m} \leqslant 2^{-m} \cdot N_{n,m}(\ell) \leqslant 2^{-m} \cdot |\mathcal{F}^{\mathrm{ind}}_{n,m}(C_4)|, \tag{41}$$

where $N_{n,m}(\ell)$ (cf. Section 4.2) denotes the number of graphs with vertex set $\{1, \ldots, n\}$ and precisely m edges that are complete on the set $\{1, \ldots, \ell\}$ and empty on its complement. To prove (41), note first that

$$\binom{a}{c} \geqslant \left(\frac{a}{b}\right)^c \binom{b}{c} \qquad \text{and} \qquad \binom{a}{b} \leqslant \left(\frac{a}{b-c}\right)^c \binom{a}{b-c} \tag{42}$$

for every $a \geqslant b \geqslant c \geqslant 0$ and choose $\ell \in \mathbb{N}$ so that

$$\sqrt{\frac{3m}{2\log(n^2/m)}} \leqslant \ell \leqslant \sqrt{\frac{2m}{\log(n^2/m)}},$$

so, in particular, $\ell(n - \ell) \geqslant \sqrt{n^2 m / \log(n^2/m)}$. It follows that

$$\binom{2^{-3}\sqrt{n^2 m / \log(n^2/m)}}{m} \leqslant 2^{-3m} \left(\frac{\ell(n-\ell)}{m - \binom{\ell}{2}}\right)^{\binom{\ell}{2}} \binom{\ell(n-\ell)}{m - \binom{\ell}{2}} \tag{43}$$

and, since $\binom{\ell}{2} \leqslant m / \log(n^2/m)$ and $m \leqslant \lambda n^2$, the right-hand side of (43) is at most

$$2^{-3m} \left(\frac{n^2}{m}\right)^{\binom{\ell}{2}} \binom{\ell(n-\ell)}{m - \binom{\ell}{2}} \leqslant 2^{-3m} \cdot e^m \cdot N_{n,m}(\ell) \leqslant 2^{-m} \cdot N_{n,m}(\ell),$$

as claimed. It follows that there are at most $2^{-m} \cdot |\mathcal{F}^{\mathrm{ind}}_{n,m}(C_4)|$ graphs in $\mathcal{F}^{\mathrm{ind}}_{n,m}(C_4) \cap \mathcal{P}$.

Case 2. There exists $\ell \in \mathbb{N}$ such that $e(E) \geqslant \binom{\ell}{2}$ and $e(M) \leqslant (1-\delta)\ell n$.

We may again assume that $e(E) \leqslant m$, as otherwise $\mathcal{F}^{\mathrm{ind}}_{n,m}(C_4) \cap \mathcal{P}$ is empty. Since $\binom{\ell}{2} \leqslant m \leqslant \lambda n^2 \leqslant 2^{-8}\delta^2 n^2$, it follows (using (42)) that

$$|\mathcal{F}^{\mathrm{ind}}_{n,m}(C_4) \cap \mathcal{P}| \leqslant \binom{(1-\delta)\ell n}{m - e(E)} \leqslant \binom{(1-\delta/2)\ell(n-\ell)}{m - e(E)}$$

$$\leqslant \max\left\{ 4^m, \left(1 - \frac{\delta}{2}\right)^{m-\binom{\ell}{2}} \binom{\ell(n-\ell)}{m - \binom{\ell}{2}} \right\},$$

where the bound 4^m corresponds to the case $(1-\delta/2)\ell(n-\ell) \leqslant 2m$. However, since $\lambda \leqslant \lambda_{4.4}$, it follows from Proposition 4.4 that

$$4^m < e^{-m/5} \cdot 5^m \leqslant e^{-m/5} \cdot N_{n,m}(\ell_{n,m}) \leqslant e^{-m/5} \cdot |\mathcal{F}^{\mathrm{ind}}_{n,m}(C_4)|,$$

where $\ell_{n,m}$ is defined by $\ell_{n,m} = \left(m/\log(\ell_{n,m} n/m)\right)^{1/2}$, see Proposition 4.4. It will therefore suffice to bound the second term in the maximum above.

To do so, we will consider the cases $\ell \leqslant 2\ell_{n,m}$ and $\ell > 2\ell_{n,m}$ separately. If $\ell > 2\ell_{n,m}$, then it follows from Proposition 4.4 that

$$\binom{\ell(n-\ell)}{m - \binom{\ell}{2}} = N_{n,m}(\ell) \leqslant e^{-m/15} \cdot \max_{\ell} N_{n,m}(\ell) \leqslant e^{-m/15} \cdot |\mathcal{F}^{\mathrm{ind}}_{n,m}(C_4)|.$$

On the other hand, if $\ell \leqslant 2\ell_{n,m}$, then we will show that $\binom{\ell}{2} \leqslant m/2$. Indeed,

$$\binom{\ell}{2} \leqslant 2\ell_{n,m}^2 = \frac{2m}{\log(\ell_{n,m} n/m)} = \frac{4m}{\log(n^2/m) - \log\log(\ell_{n,m} n/m)} \leqslant \frac{m}{2},$$

since $\ell_{n,m} \leqslant n$, $m \leqslant \lambda n^2$, and $\lambda \leqslant 2^{-8}$. It follows that

$$\left(1 - \frac{\delta}{2}\right)^{m-\binom{\ell}{2}} \binom{\ell(n-\ell)}{m - \binom{\ell}{2}} \leqslant e^{-\delta m/4} \cdot N_{n,m}(\ell) \leqslant e^{-\delta m/4} \cdot |\mathcal{F}^{\mathrm{ind}}_{n,m}(C_4)|.$$

We have thus shown that there are at most $e^{-\delta m/4} \cdot |\mathcal{F}^{\mathrm{ind}}_{n,m}(C_4)|$ graphs in $\mathcal{F}^{\mathrm{ind}}_{n,m}(C_4) \cap \mathcal{P}$

Summing over the $e^{o(m)}$ pregraphs in $\mathcal{C}' \setminus \mathcal{C}$, it follows that at most $e^{-\lambda m} \cdot |\mathcal{F}^{\mathrm{ind}}_{n,m}(C_4)|$ of the graphs in $\mathcal{F}^{\mathrm{ind}}_{n,m}(C_4)$ are contained in some $\mathcal{P} \in \mathcal{C}'$ that is not an ε-almost split pregraph. Since, by property (b) of Lemma 5.4, the remaining graphs in $\mathcal{F}^{\mathrm{ind}}_{n,m}(C_4)$ are contained in some $\mathcal{P} \in \mathcal{C}$, the theorem follows. $\qquad\square$

We are finally ready to prove part (c) of Theorem 1.2. To deduce from Theorem 5.1 that almost all graphs in $\mathcal{F}^{\mathrm{ind}}_{n,m}(C_4)$ have the claimed structure, it only remains to bound the number of such graphs that are not ε-close to a split graph but are contained in a pregraph that is ε'-close to a split pregraph (for some well-chosen ε').

Proof of part (c) of Theorem 1.2. Assume (without loss of generality) that $\varepsilon > 0$ is sufficiently small, and set $\delta = \min\left\{\lambda_{5.1}, \varepsilon^4\right\}$, where $\lambda_{5.1}$ is the constant obtained by applying Theorem 5.1 with $\varepsilon_{5.4} \leftarrow \varepsilon^3$. Now, given $n^{4/3}(\log n)^4 \leqslant m \leqslant \delta n^2$, it follows from Theorem 5.1 that there exists a collection $\mathcal{C}$ of ε^3-almost split pregraphs on n vertices with $|\mathcal{C}| = e^{o(m)}$ and such that at most $e^{-\delta m} \cdot |\mathcal{F}^{\mathrm{ind}}_{n,m}(C_4)|$ graphs in $\mathcal{F}^{\mathrm{ind}}_{n,m}(C_4)$ are not contained in any $\mathcal{P} \in \mathcal{C}$.

We claim that, for each $\mathcal{P} \in \mathcal{C}$,

$$\left|\left\{G \in \mathcal{F}_{n,m}^{\mathrm{ind}}(C_4) \cap \mathcal{P} : G \text{ is not } \varepsilon\text{-close to a split graph}\right\}\right| \leqslant e^{-\varepsilon m} \cdot |\mathcal{F}_{n,m}^{\mathrm{ind}}(C_4)|. \quad (44)$$

Let $V(K_n) = U \cup W$ be a partition witnessing the fact that $\mathcal{P} = (M, E)$ is an ε^3-almost split pregraph and recall that

$$e(E) \leqslant \binom{|U|}{2}, \qquad e_E(U) \geqslant (1 - \varepsilon^3)\binom{|U|}{2}, \qquad \text{and} \qquad e_M(W) \leqslant 7\varepsilon^{3/2}|U|n. \quad (45)$$

Let ℓ be the largest integer such that $e_E(U) \geqslant \binom{\ell}{2}$ and note that $\ell \leqslant |U| \leqslant (1 + \varepsilon^3)(\ell + 2)$. As usual, we may assume that $e(E) \leqslant m$, since otherwise $\mathcal{F}_{n,m}^{\mathrm{ind}}(C_4) \cap \mathcal{P}$ is empty; note that therefore $\ell \leqslant \sqrt{3m} \leqslant \sqrt{3\delta} \cdot n$. We may also assume that $\ell \geqslant 2^{-5}\sqrt{m/\log(n^2/m)}$, since otherwise

$$e(M) \leqslant \varepsilon^3 \binom{|U|}{2} + |U|(n - |U|) + 7\varepsilon^{3/2}|U|n \leqslant 2\ell n \leqslant 2^{-4}\sqrt{\frac{n^2 m}{\log(n^2/m)}},$$

and in Case 1 of the proof of Theorem 5.1 we showed that if this is the case then $|\mathcal{F}_{n,m}^{\mathrm{ind}}(C_4) \cap \mathcal{P}| \leqslant 2^{-m} \cdot |\mathcal{F}_{n,m}^{\mathrm{ind}}(C_4)|$, as required. It follows that

$$m \leqslant 2^{10}\ell^2 \log \frac{n^2}{m} \leqslant 2^{10}\ell n \cdot \frac{\ell}{n} \log \frac{3n^2}{\ell^2} \leqslant 2^{10}\ell n \cdot \sqrt{3\delta} \log \frac{1}{\delta} \leqslant \varepsilon^{3/2}\ell n,$$

where the first inequality follows from $\ell \geqslant 2^{-5}\sqrt{m/\log(n^2/m)}$, the second since $\binom{\ell}{2} \leqslant e(E) \leqslant m$, the third since $\ell \leqslant \sqrt{3\delta} \cdot n$, and the fourth since $\delta \leqslant \varepsilon^4$.

Now, observe that, by (45), if $G \in \mathcal{F}_{n,m}^{\mathrm{ind}}(C_4) \cap \mathcal{P}$ is not ε-close to a split graph, then G has at least εm edges in the set W. It follows that the left-hand side of (44) is at most

$$\sum_{s \geqslant \varepsilon m} \binom{e_M(U) + e_M(U, W)}{m - s - e_E(U) - e_E(U, W)} \binom{e_M(W)}{s - e_E(W)}.$$

Noting that $e_M(U) \leqslant 2\varepsilon^3 \binom{\ell}{2}$, $e_M(U, W) \leqslant (1 + 2\varepsilon^3)\ell(n - \ell)$, and $e_M(W) \leqslant 8\varepsilon^{3/2}\ell n$, this is in turn at most

$$\sum_{s \geqslant \varepsilon m} \binom{\ell(n - \ell) + 2\varepsilon^3 \ell n}{m - s - e_E(U) - e_E(U, W)} \binom{8\varepsilon^{3/2}\ell n}{s - e_E(W)}. \quad (46)$$

To bound this sum, note first that the inequalities $m \leqslant \varepsilon^{3/2}\ell n$ and $\ell < \sqrt{3\delta}n < n/2$ imply that $\ell(n - \ell) \geqslant 2m$, and hence

$$\binom{\ell(n - \ell) + 2\varepsilon^3 \ell n}{m - s - e_E(U) - e_E(U, W)} \leqslant \binom{\ell(n - \ell) + 2\varepsilon^3 \ell n}{m - s - \binom{\ell}{2}}, \quad (47)$$

since $e_E(U) \geqslant \binom{\ell}{2}$. Now, using the inequalities $\binom{a+c}{b} \leqslant \left(\frac{a+c-b}{a-b}\right)^b \binom{a}{b}$ and $\binom{a}{b-c} \leqslant \left(\frac{b}{a-b}\right)^c \binom{a}{b}$ and the bounds $\ell(n - \ell) \geqslant 2m$ and $\ell < \sqrt{3\delta} \cdot n \leqslant n/3$, we can bound the right-hand side of (47) from above by

$$\left(1 + \frac{4\varepsilon^3 \ell n}{\ell(n - \ell)}\right)^m \left(\frac{2m}{\ell(n - \ell)}\right)^s \binom{\ell(n - \ell)}{m - \binom{\ell}{2}} \leqslant (1 + 6\varepsilon^3)^m \left(\frac{3m}{\ell n}\right)^s N_{n,m}(\ell).$$

Observe also that

$$\binom{8\varepsilon^{3/2}\ell n}{s - e_E(W)} \leqslant \binom{8\varepsilon^{3/2}\ell n}{s} \leqslant \left(\frac{8e\varepsilon^{3/2}\ell n}{s}\right)^s,$$

41

since $s = e(G[W]) \leqslant m \leqslant \varepsilon^{3/2}\ell n$. It follows that (46) is at most

$$\sum_{s \geqslant \varepsilon m} (1 + 6\varepsilon^3)^m \left(\frac{8e\varepsilon^{3/2}\ell n}{s}\right)^s \left(\frac{3m}{\ell n}\right)^s N_{n,m}(\ell),$$

which is easily bounded from above by

$$e^{6\varepsilon^3 m} N_{n,m}(\ell) \sum_{s \geqslant \varepsilon m} \left(\frac{24e\varepsilon^{3/2}m}{s}\right)^s \leqslant m\left(24e\varepsilon^{1/2}\right)^{\varepsilon m} e^{6\varepsilon^3 m} N_{n,m}(\ell) \leqslant e^{-\varepsilon m} N_{n,m}(\ell),$$

proving (44). Since $|\mathcal{C}| = e^{o(m)}$ and at most $e^{-\delta m} \cdot |\mathcal{F}_{n,m}^{\mathrm{ind}}(C_4)|$ graphs in $\mathcal{F}_{n,m}^{\mathrm{ind}}(C_4)$ are not contained in any $\mathcal{P} \in \mathcal{C}$, it follows that almost all graphs in $\mathcal{F}_{n,m}^{\mathrm{ind}}(C_4)$ are ε-close to a split graph, as required. This completes the proof of Theorem 1.2. $\qquad \square$

It only remains to prove Corollary 1.3. We will in fact use Theorem 5.1, together with Theorem 1.2 and Proposition 4.1, to prove the following slightly stronger (and more technical) statement.

Corollary 5.5. *For every $\varepsilon > 0$, there exists $\delta > 0$ such that the following holds a.a.s. for $G \sim G_{n,p}^{\mathrm{ind}}(C_4)$:*

(a) If $n^{-1} \ll p \leqslant \delta n^{-2/3}$, then G is ε-quasirandom, and

$$e(G) \in \left(1 \pm \varepsilon\right)p\binom{n}{2}. \tag{48}$$

(b) If $\delta n^{-2/3} \leqslant p \leqslant n^{-1/3}(\log n)^4$, then

$$\frac{\delta n^{4/3}}{4} \leqslant e(G) \leqslant n^{4/3}(\log n)^8.$$

(c) If $n^{-1/3}(\log n)^4 \leqslant p \ll 1$, then G is ε-close to a split graph and

$$e(G) = \Theta\left(\frac{p^2 n^2}{\log(1/p)}\right).$$

Proof. Let $G \sim G_{n,p}^{\mathrm{ind}}(C_4)$ be the (random) graph obtained by conditioning $G(n,p)$ to contain no induced 4-cycle and let E_m denote the event that G has exactly m edges. Observe that $\mathbb{P}(E_m) \propto \varphi(m)$, where

$$\varphi(m) = |\mathcal{F}_{n,m}^{\mathrm{ind}}(C_4)| \cdot \left(\frac{p}{1-p}\right)^m,$$

and that the distribution of G conditioned on E_m is uniform on $\mathcal{F}_{n,m}^{\mathrm{ind}}(C_4)$. Our main task will be to find the values of m for which $\varphi(m)$ is the largest. In order to show that $\mathbb{P}\left(\bigcup_{m \in R} E_m\right) \geqslant 1 - \alpha$ for some $R \subseteq \{0, \ldots, \binom{n}{2}\}$ and $\alpha \geqslant 0$, it is enough to prove that for some $m \in R$ and all $m' \notin R$, we have $\alpha \cdot \varphi(m) \geqslant n^2 \cdot \varphi(m')$. We first observe the following straightforward lower and upper bounds on $\varphi(m)$.

Claim 1. There exists an absolute constant c such that for every m,

$$\left(\frac{cnp}{\sqrt{m\log(n^2/m)}}\right)^m \leqslant \varphi(m) \leqslant \left(\frac{en^2 p}{2m(1-p)}\right)^m.$$

Proof. The upper bound follows by noting that

$$|\mathcal{F}^{\mathrm{ind}}_{n,m}(C_4)| \leqslant \binom{\binom{n}{2}}{m} \leqslant \left(\frac{e\binom{n}{2}}{m}\right)^m \leqslant \left(\frac{en^2}{2m}\right)^m.$$

The lower bound follows by observing that

$$|\mathcal{F}^{\mathrm{ind}}_{n,m}(C_4)| \geqslant |\mathcal{S}_{n,m}| \geqslant N_{n,m}\left(\sqrt{m/\log\left(n^2/m\right)}\right)$$

$$\geqslant \left(\frac{n\sqrt{m}}{2m\sqrt{\log\left(n^2/m\right)}}\right)^{(1-1/\log(n^2/m))\cdot m} \geqslant \left(\frac{cn}{\sqrt{m\log\left(n^2/m\right)}}\right)^m.$$

$\square$

We next choose the constant $\delta = \delta(\varepsilon) > 0$ as follows. First, note that the function $x \mapsto (e/x)^x$ is strictly increasing for $x \in (0,1]$ and strictly decreasing for $x \in [1,\infty)$, so there is a $\gamma > 0$ such that $(e/x)^x \leqslant e - 2\gamma$ whenever $x \notin [1 - \varepsilon/2, 1 + \varepsilon/2]$. Fix such a γ and let $\delta_{4.1}$ be the constant given by applying Proposition 4.1 with $\gamma_{4.1} \leftarrow \gamma/4$. Moreover, let $\delta_{1.2}$ be the constant given by Theorem 1.2 applied with $\varepsilon_{1.2} \leftarrow \varepsilon$ and set $\delta := \min\{\delta_{1.2}, \delta_{4.1}\}$. The following claim is an immediate consequence of Proposition 4.1.

Claim 2. If $1 \ll m \leqslant \delta n^{4/3}$, then

$$\varphi(m) \geqslant \left(\frac{(e-\gamma)n^2 p}{2m(1-p)}\right)^m$$

for all sufficiently large $n \in \mathbb{N}$.

Proof. By Proposition 4.1 and our choice of δ, we have

$$\left|\mathcal{F}^{\mathrm{ind}}_{n,m}(C_4)\right| \geqslant e^{-\gamma m/4} \cdot \binom{\binom{n}{2}}{m} \geqslant e^{-\gamma m/4} \cdot \left(\frac{(e-\gamma/4)n^2}{2m}\right)^m$$

for all sufficiently large n, where in the second inequality we used the fact that $1 \ll m \ll n^2$. Since $e^{-\gamma/4} \cdot (e - \gamma/4) \geqslant (1 - \gamma/4)(e - \gamma/4) > e - \gamma$, the claimed bound follows. $\square$

Finally, Theorem 5.1 gives the following upper bound on $\varphi(m)$.

Claim 3. There exists an absolute constant C such that if $n^{4/3}(\log n)^4 \leqslant m \ll n^2$, then

$$\varphi(m) \leqslant \left(\frac{Cnp}{\sqrt{m\log\left(n^2/m\right)}}\right)^m.$$

Proof. Suppose that $n^{4/3}(\log n)^4 \leqslant m \ll n^2$. By Theorem 5.1, almost all graphs in $\mathcal{F}^{\mathrm{ind}}_{n,m}(C_4)$ are contained in one of at most $e^{o(m)}$ pregraphs that are $(1/4)$-almost split. A given pregraph $\mathcal{P} = (M, E)$ contains exactly $\binom{e(M)}{m-e(E)}$ graphs with m edges and if $\mathcal{P}$ is $(1/4)$-almost split, then

$$\frac{3}{4}\binom{u}{2} \leqslant e(E) \leqslant \binom{u}{2} \qquad \text{and} \qquad e(M) \leqslant \frac{9}{2}un$$

for some integer u. Let ℓ be the smallest integer such that $e(E) \leqslant \binom{\ell}{2}$. It follows that $\binom{\ell-1}{2} \leqslant e(E) \leqslant \binom{\ell}{2}$ and $e(M) \leqslant 6\ell n$. In particular,

$$\binom{e(M)}{m-e(E)} \leqslant \max_{\ell} \binom{6\ell n}{m - \binom{\ell}{2}} \cdot (6\ell n)^\ell \leqslant 7^m \cdot \max_{\ell} \binom{\ell(n-\ell)}{m - \binom{\ell}{2}}.$$

By Proposition 4.4, the maximum above is attained at some ℓ satisfying $\ell_{n,m}/2 \leqslant \ell \leqslant 2\ell_{n,m}$, that is, $\ell = \Theta\left(\sqrt{m/\log(n^2/m)}\right)$. It follows that for some absolute constants C and C',

$$|\mathcal{F}_{n,m}^{\mathrm{ind}}(C_4)| \leqslant e^{C'm} \cdot N_{n,m}\left(\sqrt{m/\log(n^2/m)}\right) \leqslant \left(\frac{Cn}{2\sqrt{m\log(n^2/m)}}\right)^m,$$

which implies the claimed bound on $\varphi(m)$. $\qquad\square$

We will now use Claims 1, 2, and 3 to bound the ratios $\varphi(m)/\varphi(m')$ for various m and m'. Suppose first that $n^{-1} \ll p \leqslant \delta n^{-2/3}$, set $m = pn^2/2$, and let $m' \notin (1 \pm \varepsilon)p\binom{n}{2}$. By Claim 2,

$$\varphi(m) \geqslant \left(\frac{(e-\gamma)n^2 p}{2m(1-p)}\right)^m = \left(\frac{e-\gamma}{1-p}\right)^m$$

and by Claim 1,

$$\varphi(m') \leqslant \left(\frac{en^2 p}{2m'(1-p)}\right)^{m'} = \left(\frac{em}{m'(1-p)}\right)^{\frac{m'}{m}\cdot m}.$$

If $m' > 2em$, then $\varphi(m') \leqslant 1 \leqslant 2^{-m}\varphi(m)$, so we may assume that $m' \leqslant 2em$. Since $m' \notin (1 \pm \varepsilon/2)m$, then

$$\left(\frac{em}{m'}\right)^{\frac{m'}{m}} \leqslant e - 2\gamma$$

by our definition of γ, implying that

$$\frac{\varphi(m')}{\varphi(m)} \leqslant \left(\frac{e-2\gamma}{e-\gamma}\right)^m \cdot (1-p)^{m-m'} \leqslant \exp\left(-\gamma m/e + 2\varepsilon mp\right) \leqslant \exp\left(-\gamma m/3\right).$$

It follows that if $n^{-1} \ll p \leqslant \delta n^{-2/3}$, then $\mathbb{P}\left(e(G) \in (1 \pm \varepsilon)p\binom{n}{2}\right) \geqslant 1 - e^{-\gamma m/4}$. In particular, it follows from Theorem 1.2 that G is a.a.s. ε-quasirandom. This establishes part (a) of the corollary.

Suppose now that $\delta n^{-2/3} \leqslant p \leqslant n^{-1/3}(\log n)^4$ and let $m = \delta n^{4/3}/2$. By Claim 2,

$$\varphi(m) \geqslant \left(\frac{(e-\gamma)n^2 p}{2m(1-p)}\right)^m \geqslant 2^m,$$

where the second inequality holds as $m \leqslant pn^2/2$. If $m' \leqslant m/2$, then by Claim 1 and since $x \mapsto (a/x)^x$ is increasing for $x \leqslant a/e$,

$$\varphi(m') \leqslant \left(\frac{en^2 p}{2m'(1-p)}\right)^{m'} \leqslant \left(\frac{en^2 p}{m(1-p)}\right)^{m/2} \leqslant \left(\frac{4em(1-p)}{(e-\gamma)^2 n^2 p}\right)^{m/2} \cdot \varphi(m).$$

Since $m \leqslant pn^2/2$, it follows that

$$\varphi(m') \leqslant \left(\frac{2e(1-p)}{(e-\gamma)^2}\right)^{m/2} \cdot \varphi(m) \leqslant \left(\frac{3}{4}\right)^{m/2} \cdot \varphi(m).$$

On the other hand, if $m' \geqslant n^{4/3}(\log n)^8$, then by Claim 3,

$$\varphi(m') \leqslant \left(\frac{Cnp}{\sqrt{m'\log(n^2/m')}}\right)^{m'} \leqslant 1 \leqslant 2^{-m} \cdot \varphi(m).$$

It follows that if $\delta n^{-2/3} \leqslant p \leqslant n^{-1/3}(\log n)^4$, then $\mathbb{P}\left(\delta n^{4/3}/4 \leqslant e(G) \leqslant n^{4/3}(\log n)^8\right) \geqslant 1 - e^{-m/10}$, establishing part (b) of the corollary.

44

Finally, suppose that $n^{-1/3}(\log n)^4 \leqslant p \ll 1$, let $m = \beta p^2 n^2 / \log(1/p)$ for some small positive constant β, and observe that $m \geqslant n^{4/3}(\log n)^6$. By Claim 1,

$$\varphi(m) \geqslant \left(\frac{cnp}{\sqrt{m \log(n^2/m)}} \right)^m \geqslant \left(\frac{cnp\sqrt{\log(1/p)}}{\sqrt{\beta} \cdot pn \cdot \sqrt{3\log(1/p)}} \right)^m \geqslant e^m,$$

since β is small. If $m' \leqslant n^{4/3}(\log n)^4$, then by Claim 1,

$$\varphi(m') \leqslant \left(\frac{en^2 p}{2m'(1-p)} \right)^{m'} \leqslant n^{2m'} \leqslant \exp\left(2n^{4/3}(\log n)^5\right) \leqslant e^{m/2} \leqslant e^{-m/2} \cdot \varphi(m).$$

Let ξ be a small positive constant. If $m' \geqslant m/\xi$, then by Claim 3,

$$\varphi(m') \leqslant \left(\frac{Cnp}{\sqrt{m' \log(n^2/m')}} \right)^{m'} \leqslant 1 \leqslant e^{-m} \cdot \varphi(m).$$

Finally, suppose that $n^{4/3}(\log n)^4 \leqslant m' \leqslant \xi m$. By Claim 3,

$$\varphi(m') \leqslant \left(\frac{Cnp}{\sqrt{m' \log(n^2/m')}} \right)^{m'} \leqslant \left(\frac{Cnp}{\sqrt{m' \log(n^2/m)}} \right)^{m'}.$$

Since the function $x \mapsto (a/\sqrt{x})^x$ is increasing if $x \leqslant a^2/e$, then the right-hand side above is increasing in $m' \leqslant \xi m$ and thus

$$\varphi(m') \leqslant \left(\frac{Cnp}{\sqrt{\xi m \log(n^2/m)}} \right)^{\xi m} \leqslant \left(\frac{C}{\sqrt{\beta \xi}} \right)^{\xi m} \leqslant e^{m/2} \leqslant e^{-m/2} \cdot \varphi(m).$$

It follows that there is an absolute constant K such that if $n^{-1/3}(\log n)^4 \leqslant p \ll 1$, then

$$\mathbb{P}\left(\frac{p^2 n^2}{K \log(1/p)} \leqslant e(G) \leqslant \frac{Kp^2 n^2}{\log(1/p)} \right) \geqslant 1 - e^{-m/3}.$$

In particular, it follows from Theorem 1.2 that a.a.s. G is ε-close to a split graph. This completes the proof of the corollary. □

6. Concluding remarks and open problems

This paper makes a first step towards understanding the typical structure of a sparse member of a hereditary graph property. In this final section, we discuss a few of the many natural open problems suggested by our main results. We begin with the following conjecture on the typical structure of sparse induced-C_4-free graphs.

Conjecture 6.1. *Suppose that $n^{4/3}(\log n)^{1/3} \ll m \leqslant \binom{n}{2} - \Omega(n^2)$ and let G be a uniformly chosen random graph in $\mathcal{F}_{n,m}^{\mathrm{ind}}(C_4)$. Then a.a.s. G is a split graph.*

Note that Conjecture 6.1 would sharpen Theorem 1.2 in three different ways. First, the power of $\log n$ in the lower bound on m in part (c) of Theorem 1.2 would be reduced from 4 to 1/3, which is best possible, as shown by part (b) of Theorem 1.2. Second, the description of the typical members of $\mathcal{F}_{n,m}^{\mathrm{ind}}(C_4)$ would be more precise – the graphs are required to be split rather than only close to split. Finally, the upper bound on m is increased from $o(n^2)$ to $\binom{n}{2} - \Omega(n^2)$; in fact, we expect Conjecture 6.1 to remain true even when $\binom{n}{2} - m$ is much smaller than n^2, but then it is (arguably) more natural to consider the complements of graphs in $\mathcal{F}_{n,m}^{\mathrm{ind}}(C_4)$, which are sparse induced-$2K_2$-free graphs.

We made the assumption that $m = o(n^2)$ mainly for convenience (and to simplify the proof of Theorem 1.6) and it seems plausible that our techniques could be extended to all $m \leqslant \binom{n}{2} - \Omega(n^2)$, but we have not made any serious attempts to do so. A natural alternative approach of resolving Conjecture 6.1 for $m = \Theta(n^2)$ would be to generalise the method of Prömel and Steger [38], who characterised typical members of $\mathcal{F}_n^{\mathrm{ind}}(C_4)$.

A more substantial step towards resolving the conjecture would be to either determine the precise structure of a typical member of $\mathcal{F}_{n,m}^{\mathrm{ind}}(C_4)$ when $m \geqslant n^{4/3}(\log n)^{O(1)}$ or to determine the approximate structure in the range $n^{4/3}(\log n)^{1/3} \ll m \leqslant n^{4/3}(\log n)^4$. We remark that the methods of [11] may be helpful in achieving the former goal (determining the precise structure), though it appears that new ideas will be needed.

6.1. Sparse induced-H-free graphs.

Perhaps the most natural direction for further investigation would be to describe the typical structure of sparse induced-H-free graphs for an arbitrary graph H. The first step in this direction was made recently by Kalvari and Samotij [29], who used Theorem 1.4 to prove the following rough characterisation of a typical member of $\mathcal{F}_{n,m}^{\mathrm{ind}}(H)$ for all non-bipartite graphs H and all $m \ll n^2$.

Theorem 6.2. *Suppose that H is a non-bipartite graph and let G be uniformly chosen random graph in $\mathcal{F}_{n,m}^{\mathrm{ind}}(H)$. The following holds for every $\varepsilon > 0$:*

(a) If $n \ll m \ll n^{2-1/m_2(H)}$, then a.a.s. G is ε-quasirandom.
(b) If $n^{2-1/m_2(H)} \ll m \ll n^2$, then a.a.s. G is ε-close to $(\chi(H) - 1)$-partite.

We remark that the structural characterisation of typical sparse induced-H-free graphs provided by assertion (b) of Theorem 6.2 is not as precise as that given in Theorem 1.2 (c). We expect that for many H, the following holds for all $m \gg n^{2-1/m_2(H)}$: the vertex set of a typical member of $\mathcal{F}_{n,m}^{\mathrm{ind}}(H)$ can be partitioned into $\chi(H) - 1$ sets that are 'almost independent' (i.e., they induce $o(m)$ edges) and some number of 'almost-cliques' (that is, sets inducing graphs of density $1 - o(1)$) of size $\Theta(\sqrt{m/\log n})$. However, it is not true, in general, that the typical member of $\mathcal{F}_{n,m}^{\mathrm{ind}}(H)$ remains $o(1)$-close to $(\chi(H) - 1)$-partite when $m = \Omega(n^2)$. For example, Prömel and Steger [42] proved that almost all graphs in $\mathcal{F}_n^{\mathrm{ind}}(C_5)$ are so-called generalised split graphs. A graph G is a *generalised split graph* if the vertex set of either G or of its complement can be partitioned into sets V_1 and V_2 such that V_1 induces a union of pairwise disjoint cliques and V_2 induces a clique.

6.2. General hereditary properties of graphs.

A natural generalisation of the family of induced-H-free graphs that has been extensively studied in the literature (see, for example, [2, 3, 12]), are so-called *hereditary properties of graphs*, that is, properties of graphs that are closed under taking induced subgraphs. As we mentioned in the Introduction, the rough structure of a typical member of an arbitrary hereditary property of graphs was determined a few years ago by Alon, Balogh, Bollobás, and Morris [3]. It would be very interesting (and, most likely, extremely challenging) to obtain a corresponding statement for a typical sparse graph in a hereditary property.

In order to give the reader an idea of what it might be possible to prove in this very general setting, let us take this opportunity to state a theorem for *monotone properties of graphs* (that is, properties of graphs that are closed under taking subgraphs) which follows easily from the container theorems proved in [10, 44], but, as far as we are aware, has not previously been stated explicitly in the literature.

Given a monotone property of graphs $\mathcal{P}$, let $\mathcal{F}(\mathcal{P})$ denote the family of minimal forbidden subgraphs, i.e., the family of all graphs that are not in $\mathcal{P}$, but all of whose proper subgraphs are in $\mathcal{P}$. Theorem 6.4, below, gives an approximate structural description of a typical member of $\mathcal{P}$ with (essentially) any given order n and size m, as long as $\mathcal{F}(\mathcal{P})$ is finite. In order to state the theorem, we will need the following definition.

Definition 6.3. Given a non-trivial monotone property of graphs $\mathcal{P}$ such that $\mathcal{F}(\mathcal{P})$ is finite, we define the sequence

$$m(\mathcal{P}) = \big((a_1, r_1), \ldots, (a_s, r_s)\big)$$

as follows:

   (i) Set $a_0 = 0$ and $r_0 = \infty$.
   (ii) Let $i \geqslant 0$ and suppose that we have already defined (a_i, r_i). Let

$$a_{i+1} := \min \big\{ m_2(H) : H \in \mathcal{F}(\mathcal{P}),\ m_2(H) > a_i,\ \text{and}\ \chi(H) \leqslant r_i \big\},$$

provided that the above set is not empty; otherwise, set $s = i$ and stop.
   (iii) If a_{i+1} was defined in step (ii), then let

$$r_{i+1} := \min \big\{ \chi(H) - 1 : H \in \mathcal{F}(\mathcal{P})\ \text{and}\ m_2(H) = a_{i+1} \big\}.$$

If $r_{i+1} = 1$, then set $s = i+1$ and stop; otherwise, increase i by one and go to (ii).

Given integers n and m and a graph property $\mathcal{P}$, denote by $\mathcal{P}_{n,m}$ the family of all graphs with vertex set $\{1, \ldots, n\}$ and precisely m edges that belong to $\mathcal{P}$.

Theorem 6.4. *Let $\mathcal{P}$ be a non-trivial monotone property of graphs such that $\mathcal{F}(\mathcal{P})$ is finite and let G be a uniformly chosen random graph in $\mathcal{P}_{n,m}$. Suppose that $m(\mathcal{P}) = \big((a_1, r_1), \ldots, (a_s, r_s)\big)$. The following holds for every $\varepsilon > 0$:*

   (a) If $n \ll m \ll n^{2-1/a_1}$, then a.a.s. G is ε-quasirandom.
   (b) If $n^{2-1/a_i} \ll m \ll n^{2-1/a_{i+1}}$ for some $i \in \{1, \ldots, s-1\}$, then a.a.s. G is ε-close to r_i-partite.
   (c) If $m \gg n^{2-1/a_s}$ and $r_s \geqslant 2$, then a.a.s. G is ε-close to r_s-partite.

Since the proof of Theorem 6.4 is a (nowadays) standard application of the container method, using (a robust version of) the stability theorem of Erdős and Simonovits [20, 46] (cf. the proof of [10, Theorem 1.7]), we leave the details to the reader.

Finally, we remark that the assumption that $\mathcal{F}(\mathcal{P})$ is finite is essential. Indeed, suppose that $\mathcal{F}(\mathcal{P})$ contains all (minimal) non-bipartite graphs H with $m_2(H) \geqslant a$ for a given $a > 1$. If $m \geqslant an$, then $\mathcal{P}_{n,m}$ contains only bipartite graphs and thus if $\varepsilon > 0$ is sufficiently small, then there are no graphs in $\mathcal{P}$ that are ε-quasirandom.

References

1. M. Ajtai, J. Komlós, J. Pintz, J. Spencer, and E. Szemerédi, *Extremal uncrowded hypergraphs*, J. Combin. Theory Ser. A **32** (1982), 321–335.

2. V. E. Alekseev, *Range of values of entropy of hereditary classes of graphs*, Diskret. Mat. **4** (1992), 148–157.

3. N. Alon, J. Balogh, B. Bollobás, and R. Morris, *The structure of almost all graphs in a hereditary property*, J. Combin. Theory Ser. B **101** (2011), 85–110.

4. N. Alon and J. H. Spencer, *The probabilistic method*, fourth ed., Wiley Series in Discrete Mathematics and Optimization, John Wiley & Sons, Inc., Hoboken, NJ, 2016.

5. J. Balogh, B. Bollobás, and M. Simonovits, *The number of graphs without forbidden subgraphs*, J. Combin. Theory Ser. B **91** (2004), 1–24.

6. _____, *The typical structure of graphs without given excluded subgraphs*, Random Structures Algorithms **34** (2009), 305–318.

7. J. Balogh, B. Bollobás, and M. Simonovits, *The fine structure of octahedron-free graphs*, J. Combin. Theory Ser. B **101** (2011), 67–84.

8. J. Balogh and J. Butterfield, *Excluding induced subgraphs: critical graphs*, Random Structures Algorithms **38** (2011), 100–120.

9. J. Balogh, R. Morris, and W. Samotij, *The method of hypergraph containers*, arXiv:1801.04584 [math.CO].

10. _____, *Independent sets in hypergraphs*, J. Amer. Math. Soc. **28** (2015), 669–709.

11. 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), 6439–6485.

12. B. Bollobás and A. Thomason, *Hereditary and monotone properties of graphs*, The mathematics of Paul Erdős, II, Algorithms Combin., vol. 14, Springer, Berlin, 1997, pp. 70–78.

13. Y. Caro, *New results on the independence number*, Tech. report, Tel Aviv University, 1979.

14. D. Conlon and W. T. Gowers, *Combinatorial theorems in sparse random sets*, Ann. of Math. (2) **184** (2016), 367–454.

15. D. Dellamonica, Jr., Y. Kohayakawa, S. J. Lee, V. Rödl, and W. Samotij, *The number of B_3-sets of a given cardinality*, J. Combin. Theory Ser. A **142** (2016), 44–76.

16. _____, *The number of B_h-sets of a given cardinality*, J. Lond. Math. Soc. (2) **116** (2018), 629–669.

17. B. DeMarco and J. Kahn, *Turán's theorem for random graphs*, arXiv:1501.01340 [math.PR].

18. _____, *Mantel's theorem for random graphs*, Random Structures Algorithms **47** (2015), 59–72.

19. R. A. Duke, H. Lefmann, and V. Rödl, *On uncrowded hypergraphs*, Random Structures Algorithms **6** (1995), 209–212.

20. P. Erdős, *Some recent results on extremal problems in graph theory. Results*, Theory of Graphs (Internat. Sympos., Rome, 1966), Gordon and Breach, New York; Dunod, Paris, 1967, pp. 117–123 (English); pp. 124–130 (French).

21. P. Erdős, P. Frankl, and V. Rödl, *The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent*, Graphs Combin. **2** (1986), 113–121.

22. 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.

23. P. Erdős and A. H. Stone, *On the structure of linear graphs*, Bull. Amer. Math. Soc. **52** (1946), 1087–1091.

24. A. Ferber, G. McKinley, and W. Samotij, *Supersaturated sparse graphs and hypergraphs*, to appear in Int. Math. Res. Not. IMRN.

25. Z. Füredi and M. Simonovits, *The history of degenerate (bipartite) extremal graph problems*, Erdős centennial, Bolyai Soc. Math. Stud., vol. 25, János Bolyai Math. Soc., Budapest, 2013, pp. 169–264.

26. L. Gishboliner and A. Shapira, *Efficient removal without efficient regularity*, to appear in Combinatorica.

27. W. Hoeffding, *Probability inequalities for sums of bounded random variables*, J. Amer. Statist. Assoc. **58** (1963), 13–30.

28. 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.

29. T. Kalvari and W. Samotij, *The typical structure of sparse graphs with no induced copy of a given subgraph*, in preparation.

30. P. Keevash and W. Lochet, *The structure of typical eye-free graphs and a Turán-type result for two weighted colours*, to appear in Combin. Probab. Comput.

31. J. Kim, D. Kühn, D. Osthus, and T. Townsend, *Forbidding induced even cycles in a graph: typical structure and counting*, to appear in J. Combin. Theory Ser. B.

32. Y. Kohayakawa, B. Kreuter, and A. Steger, *An extremal problem for random graphs and the number of graphs with large even-girth*, Combinatorica **18** (1998), 101–120.

33. Y. Kohayakawa, T. Łuczak, and V. Rödl, *On K^4-free subgraphs of random graphs*, Combinatorica **17** (1997), 173–213.

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

35. T. Łuczak, *On triangle-free random graphs*, Random Structures Algorithms **16** (2000), 260–276.

36. R. Morris and D. Saxton, *The number of $C_{2\ell}$-free graphs*, Adv. Math. **298** (2016), 534–580.

37. D. Osthus, H. J. Prömel, and A. Taraz, *For which densities are random triangle-free graphs almost surely bipartite?*, Combinatorica **23** (2003), 105–150.

38. H. J. Prömel and A. Steger, *Excluding induced subgraphs: quadrilaterals*, Random Structures Algorithms **2** (1991), 55–71.

39. ———, *The asymptotic number of graphs not containing a fixed color-critical subgraph*, Combinatorica **12** (1992), 463–473.

40. H. J. Prömel and A. Steger, *Excluding induced subgraphs. III. A general asymptotic*, Random Structures Algorithms **3** (1992), 19–31.

41. ———, *Excluding induced subgraphs. II. Extremal graphs*, Discrete Appl. Math. **44** (1993), 283–294.

42. Hans Jürgen Prömel and Angelika Steger, *Almost all Berge graphs are perfect*, Combin. Probab. Comput. **1** (1992), 53–79.

43. ———, *On the asymptotic structure of sparse triangle free graphs*, J. Graph Theory **21** (1996), 137–151.

44. D. Saxton and A. Thomason, *Hypergraph containers*, Invent. Math. **201** (2015), 925–992.

45. M. Schacht, *Extremal results for random discrete structures*, Ann. of Math. (2) **184** (2016), 333–365.

46. M. Simonovits, *A method for solving extremal problems in graph theory, stability problems*, Theory of Graphs (Proc. Colloq., Tihany, 1966), Academic Press, New York, 1968, pp. 279–319.

47. P. Turán, *Eine Extremalaufgabe aus der Graphentheorie*, Mat. Fiz. Lapok **48** (1941), 436–452.

48. V. K. Wei, *A lower bound on the stability number of a simple graph*, Tech. Report 81-11217-9, Bell Laboratories, 1981.

Appendix A. The number of split graphs

Proof of Proposition 4.4. Suppose that $n \ll m \leqslant \lambda n^2$ for some positive constant λ whose value will be specified later. Fix an ℓ such that both $N_{n,m}(\ell)$ and $N_{n,m}(\ell+1)$ are nonzero and note that this implies that $0 \leqslant m - \binom{\ell+1}{2} \leqslant \ell(n-\ell)$. Therefore, $\ell < 2\sqrt{\lambda}n$ and

$$\frac{N_{n,m}(\ell+1)}{N_{n,m}(\ell)} = \frac{\binom{(\ell+1)(n-\ell-1)}{m-\binom{\ell+1}{2}}}{\binom{\ell(n-\ell)}{m-\binom{\ell+1}{2}}} \cdot \frac{\binom{\ell(n-\ell)}{m-\binom{\ell+1}{2}}}{\binom{\ell(n-\ell)}{m-\binom{\ell}{2}}}. \tag{49}$$

Denote the first and the second ratios in the right-hand side of (49) by $a(\ell)$ and $b(\ell)$, respectively. In other words, let

$$a(\ell) = \frac{\left((\ell+1)(n-\ell-1)\right)_{m-\binom{\ell+1}{2}}}{\left(\ell(n-\ell)\right)_{m-\binom{\ell+1}{2}}} \qquad \text{and} \qquad b(\ell) = \frac{\left(m - \binom{\ell}{2}\right)_\ell}{\left(\ell(n-\ell) - m + \binom{\ell+1}{2}\right)_\ell},$$

where $(\cdot)_k$ denotes the falling factorial, that is, $(a)_k = a!/(a-k)! = a(a-1)\dots(a-k+1)$. Routine calculation shows that

$$\left(1 + \frac{n-2\ell-1}{\ell(n-\ell)}\right)^{m-\binom{\ell+1}{2}} \leqslant a(\ell) \leqslant \left(1 + \frac{n-2\ell-1}{\ell(n-\ell)-m}\right)^{m-\binom{\ell+1}{2}}$$

and that

$$\left(\frac{m-\binom{\ell+1}{2}}{\ell(n-\ell)-m+\binom{\ell+1}{2}}\right)^{\ell} \leqslant b(\ell) \leqslant \left(\frac{m-\binom{\ell}{2}}{\ell(n-\ell)-m+\binom{\ell}{2}}\right)^{\ell}.$$

Let ε be a small positive constant. We claim that $\ell_{n,m} \leqslant \varepsilon\sqrt{m}$, provided that λ is sufficiently small. Indeed, otherwise we would have

$$\ell_{n,m}^2 = \frac{m}{\log(\ell_{n,m}n/m)} \leqslant \frac{m}{\log(\varepsilon n/\sqrt{m})} \leqslant \frac{m}{\log(\varepsilon/\sqrt{\lambda})} < \varepsilon^2 m,$$

contradicting our assumption. Moreover, we claim that $m \leqslant \varepsilon \ell_{n,m} n$, provided that λ is sufficiently small. Indeed, otherwise we would have

$$\ell_{n,m}^2 < \frac{m}{\varepsilon^2 n^2/m} \leqslant \frac{m}{\ell_{n,m}n/m} < \frac{m}{\log(\ell_{n,m}n/m)}.$$

We claim that the function $\ell \mapsto N_{n,m}(\ell)$ is increasing on the interval $[2m/n, 3\ell_{n,m}/4]$. Indeed, suppose that $2m/n \leqslant \ell \leqslant 3\ell_{n,m}/4$. Then $n-2\ell-1 \geqslant (1-\varepsilon)\cdot(n-\ell)$, whenever λ is sufficiently small, and hence

$$a(\ell) \geqslant \left(1 + \frac{1-\varepsilon}{\ell}\right)^{(1-\varepsilon)m} \geqslant \exp\left((1-3\varepsilon)\cdot\frac{m}{\ell}\right), \tag{50}$$

as $\ell > 2m/n \gg 1$. On the other hand, since $\ell \leqslant \ell_{n,m} \leqslant \varepsilon\sqrt{m}$, then

$$b(\ell) \geqslant \left(\frac{m-\ell^2}{\ell n}\right)^{\ell} \geqslant \left(\frac{(1-\varepsilon^2)m}{\ell n}\right)^{\ell}. \tag{51}$$

Finally, one easily checks that

$$\exp\left(\frac{\alpha m}{\ell}\right)\cdot\left(\frac{\beta m}{\ell n}\right)^{\ell} \geqslant 1 \qquad \Longleftrightarrow \qquad \ell \leqslant \sqrt{\frac{\alpha m}{\log(\ell n/(\beta m))}}, \tag{52}$$

and thus $a(\ell)\cdot b(\ell) \geqslant 1$, provided that ε is sufficiently small. Indeed, our assumption on ℓ implies that

$$\ell \leqslant 3\ell_{n,m}/4 = \frac{3}{4}\cdot\sqrt{\frac{m}{\log(\ell_{n,m}n/m)}} \leqslant \frac{3}{4}\cdot\sqrt{\frac{m}{\log(\ell n/m)}}.$$

Second, we show that the function $\ell \mapsto N_{n,m}(\ell)$ is decreasing for $\ell \geqslant 3\ell_{n,m}/2$. Indeed, if $\ell \geqslant 3\ell_{n,m}/2$, then $m < \varepsilon\ell n$ and hence, recalling that $\ell < 2\sqrt{\lambda}n \leqslant \varepsilon n$,

$$a(\ell) \leqslant \left(1 + \frac{n}{\ell(n-\ell)-m}\right)^{m} \leqslant \left(1 + \frac{n}{\ell n - 2\varepsilon\ell n}\right)^{m} \leqslant \exp\left((1+3\varepsilon)\cdot\frac{m}{\ell}\right) \tag{53}$$

and

$$b(\ell) \leqslant \left(\frac{m}{\ell n - 2m}\right)^{\ell} \leqslant \left(\frac{m}{(1-2\varepsilon)\ell n}\right)^{\ell}. \tag{54}$$

Therefore, $a(\ell)\cdot b(\ell) < 1$, provided that ε is sufficiently small, see (52).

We conclude that if $m \leqslant \lambda n^2$ for some sufficiently small positive λ, then the function $[2m/n, \infty) \ni \ell \mapsto N_{n,m}(\ell)$ attains its maximum value for some ℓ satisfying $3\ell_{n,m}/4 <$

$\ell < 3\ell_{n,m}/2$. In particular, in order to complete the proof of the first assertion of the proposition, it suffices to check that $N_{n,m}(\ell) < N_{n,m}(\ell_{n,m})$ whenever $\ell \leqslant 2m/n$. To this end, first note that if $\ell < 2m/n$, then

$$N_{n,m}(\ell) \leqslant 2^{\ell(n-\ell)} \leqslant 2^{2m},$$

whereas

$$N_{n,m}(\ell_{n,m}) \geqslant \binom{(1-\varepsilon)\ell_{n,m}n}{(1-\varepsilon)m} \geqslant \binom{(1-\varepsilon)m/\varepsilon}{(1-\varepsilon)m} \geqslant 5^m,$$

provided that ε is sufficiently small.

Finally, we establish the remaining 'large deviation' assertions of the proposition. To this end, we first note that if $m/n \leqslant \ell \leqslant 3\ell_{n,m}/4$, then

$$\exp\left(\frac{m}{\ell}\right) \cdot \left(\frac{m}{\ell n}\right)^{\ell} \geqslant \exp\left(\frac{4m}{3\ell_{n,m}}\right) \cdot \left(\frac{m}{\ell_{n,m}n}\right)^{\ell_{n,m}} = \exp\left(\frac{m}{3\ell_{n,m}}\right).$$

Consequently, if ε is sufficiently small, then by (50) and (51),

$$\frac{N_{n,m}(3\ell_{n,m}/4)}{N_{n,m}(\ell_{n,m}/2)} \geqslant \prod_{\ell=\ell_{n,m}/2}^{3\ell_{n,m}/4} a(\ell)b(\ell) \geqslant \exp\left(\frac{4}{5} \cdot \frac{m}{3\ell_{n,m}} \cdot \frac{\ell_{n,m}}{4}\right) = \exp\left(\frac{m}{15}\right). \qquad (55)$$

Similarly, if $\ell \geqslant 3\ell_{n,m}/2 \geqslant m/n$, then

$$\exp\left(\frac{m}{\ell}\right) \cdot \left(\frac{m}{\ell n}\right)^{\ell} \leqslant \exp\left(\frac{2m}{3\ell_{n,m}}\right) \cdot \left(\frac{m}{\ell_{n,m}n}\right)^{\ell_{n,m}} = \exp\left(-\frac{m}{3\ell_{n,m}}\right)$$

and consequently if ε is sufficiently small, then by (53) and (54),

$$\frac{N_{n,m}(2\ell_{n,m})}{N_{n,m}(3\ell_{n,m}/2)} \leqslant \prod_{\ell=3\ell_{n,m}/2}^{2\ell_{n,m}} a(\ell)b(\ell) \leqslant \exp\left(-\frac{3}{4} \cdot \frac{m}{3\ell_{n,m}} \cdot \frac{\ell_{n,m}}{2}\right) = \exp\left(-\frac{m}{8}\right). \qquad (56)$$

Since the function $\ell \mapsto N_{n,m}(\ell)$ is increasing when $2m/n \leqslant \ell \leqslant \ell_{n,m}/2$ and decreasing when $\ell \geqslant 2\ell_{n,m}$, and since

$$N_{n,m}(\ell) \leqslant 2^{2m} \leqslant e^{-m/15} \cdot 5^m \leqslant e^{-m/15} \cdot N_{n,m}(\ell_{n,m})$$

for all $\ell < 2m/n$, we conclude from (55) and (56) that if $\ell \leqslant \ell_{n,m}/2$ or $\ell \geqslant 2\ell_{n,m}$, then

$$N_{n,m}(\ell) \leqslant \exp(-m/15) \cdot \max_{\ell} N_{n,m}(\ell),$$

as claimed. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

IMPA, Estrada Dona Castorina 110, Jardim Botânico, Rio de Janeiro, RJ, Brasil
E-mail address: rob@impa.br

School of Mathematical Sciences, Tel Aviv University, Tel Aviv 6997801, Israel
E-mail address: samotij@post.tau.ac.il

DeepMind, London, UK
E-mail address: saxton@google.com