Threshold functions for asymmetric Ramsey properties involving cycles

Y. Kohayakawa*
Instituto de Matemática e Estatística, Universidade de São Paulo
Rua do Matão 1010, 05508–900 São Paulo SP, Brazil
E-mail address: <yoshi@ime.usp.br>

B. Kreuter
Institut für Informatik, Humboldt-Universität zu Berlin
Unter den Linden 6, 10099 Berlin, Germany
E-mail address: <kreuter@informatik.hu-berlin.de>

Abstract

We consider the binomial random graph $G_{n,p}$ and determine a sharp threshold function for the edge-Ramsey property

$$G_{n,p} \to (C^{\ell_1}, \ldots, C^{\ell_r})$$

for all $\ell_1, \ldots, \ell_r$, where C^ℓ denotes the cycle of length ℓ. As deterministic consequences of our results, we prove the existence of sparse graphs having the above Ramsey property as well as the existence of infinitely many critical graphs with respect to the property above.

AMS subject classification code (1991): 05C55, 05C80, 05C38
Key words and phrases: Ramsey theory, random graphs, threshold functions, Szemerédi's lemma, cycles

1 Introduction and Results

Ramsey properties of random graphs have been investigated only recently. The aim of these investigations has been to find sharp thresholds for various Ramsey properties that a random graph may satisfy. Let us concentrate here on the binomial model of random graphs, the so-called $G_{n,p}$- or G_p-model. Thus, our random graph $G_p = G_{n,p}$ has n vertices

*Partially supported by CNPq (Proc. 300334/93–1 and ProTeM-CC-II Project ProComb). Part of this work was done while this author was visiting Humboldt-Universität zu Berlin through PROBRAL project number 026/95, a CAPES–DAAD exchange programme.

and its edges are present independently with probability p. See Bollobás [1] for details concerning random graphs and Graham, Rothschild, and Spencer [2] for Ramsey theory.

The thresholds for vertex-Ramsey properties, *i.e.*, the case in which the vertices of the random graph are coloured, were established by Łuczak, Ruciński, and Voigt [12]. The case in which edges are coloured turned out to be rather more involved, but Rödl and Ruciński [13, 14, 15] succeeded in settling this case quite recently as well. Now, the results in [12, 13, 14, 15] cover the case of *symmetric* Ramsey properties, that is, the case in which we seek a monochromatic copy of one given graph. Here we shall consider the *asymmetric* case, namely, when we have one target graph for each colour. The threshold functions for asymmetric Ramsey properties with respect to vertex colourings are determined in [10]. In this note we consider edge-Ramsey properties of G_p when the target graphs are cycles.

We need some definitions before we may proceed. For an integer n, we write $[n]$ for $\{1, \ldots, n\}$. All logarithms in this note are with respect to base e. For a graph G, we denote by $V(G)$ and $E(G)$ the sets of vertices and edges of G. The sizes of these sets will be abbreviated by $|G| = |V(G)|$ and $e(G) = |E(G)|$. The *maximum 2-density* of a graph G is

$$m_2(G) \quad = \quad \max\left\{\frac{e(J)-1}{|J|-2} \colon J \subseteq G \text{ with } |J| \geq 3\right\}.$$

We now recall the arrow notation for Ramsey properties. For two graphs F and H and an integer $r \geq 2$, we write $F \to (H)_r$ to mean that in any colouring of the edges of F with r colours there is a subgraph of F isomorphic to H whose edges are all coloured by the same colour. Given r graphs $H_1, \ldots, H_r$, we may generalize this definition by letting $F \to (H_1, \ldots, H_r)$ stand for the fact that, for any colouring of the edges of F with colours $1, \ldots, r$, there is a colour i for which there is a copy of H_i in F whose edges are all coloured with colour i.

Using this notation, the results of Rödl and Ruciński read as follows, where P^4 denotes the path on four vertices.

Theorem 1 *Let H be a graph containing a cycle or a P^4. Let $\beta = 1/m_2(H)$ and let $r \geq 2$ be an integer. Then there exist constants b and $B > 0$ such that*

(*a*) *for $p = p(n) \leq bn^{-\beta}$ we have*

$$\mathbf{P}\left[G_p \to (H)_r\right] \quad = \quad o(1),$$

(*b*) *for $p = p(n) \geq Bn^{-\beta}$ we have*

$$\mathbf{P}\left[G_p \to (H)_r\right] \quad = \quad 1 - o(1).$$

Now consider the problem of determining the threshold for the property $G_p \to (H_1, H_2)$. For graphs H_1 and H_2 with $m_2(H_1) \le m_2(H_2)$, put

$$\beta = \beta(H_1, H_2) = \min\left\{\frac{|J| - 2 + 1/m_2(H_1)}{e(J)} \colon J \subseteq H_2 \text{ with } e(J) \ge 1\right\}. \tag{1}$$

Observe that β is not symmetric in H_1 and H_2 but depends to a larger extent on the denser of the two graphs. The following theorem contains the so-called 1-statement corresponding to the case in which H_1 is the cycle C^ℓ and H_2 is an arbitrary 2-balanced graph H with $m_2(H) > m_2(C^\ell) = (\ell - 1)/(\ell - 2)$.

Theorem 2 *Let $\ell \ge 3$ be a fixed integer, and let H be a 2-balanced graph with $m_2(H) = (e(H) - 1)/(|H| - 2) > m_2(C^\ell) = (\ell - 1)/(\ell - 2)$. Then there is a constant B such that, setting $p = p(n) = Bn^{-\beta}$ where*

$$\beta = \beta(C^\ell, H) = \frac{|H| - 2 + 1/m_2(C^\ell)}{e(H)} = \frac{|H| - 1 - 1/(\ell - 1)}{e(H)}, \tag{2}$$

we have

$$\mathbf{P}\left[G_p \to (C^\ell, H)\right] = 1 - o(1). \tag{3}$$

In the result above, the condition that H should be 2-balanced and that $m_2(H) > m_2(C^\ell)$ may be replaced by the condition that H should be *strictly* 2-balanced (*i.e.*, the maximum in the definition of $m_2(H)$ should be attained by H only) and that $m_2(H) \ge m_2(C^\ell)$. The main lemma in the proof of Theorem 2, Lemma 17, concerns the number of C^ℓ-free graphs on n vertices fulfilling some additional properties. This lemma is a generalization of Lemma 11 from [8], which deals with the case $\ell = 3$. Furthermore, Lemma 17 may be used to unify the proofs of the results in [3] and [4]. However, the main result from [7], which is similar in spirit, does not follow from Lemma 17.

The next theorem says that the exponent β in Theorem 2 is involved in the corresponding 0-statement, thus establishing a sharp threshold for the property $G_p \to (C^\ell, C^k)$.

Theorem 3 *Let $3 \le k \le \ell$ be integers, and let*

$$\beta = \beta(C^\ell, C^k) = \frac{(k-1)(\ell-1) - 1}{k(\ell - 1)} = 1 - \frac{\ell}{k(\ell - 1)}.$$

Then there exists a constant $b > 0$ such that, for $p = p(n) = bn^{-\beta}$, we have

$$\mathbf{P}\left[G_p \to (C^\ell, C^k)\right] = o(1).$$

We shall also sketch the proof of the following generalized version of Theorem 2.

Theorem 4 *Let $3 \leq \ell_2 \leq \cdots \leq \ell_r$ $(r \geq 2)$ be integers, and suppose H is a 2-balanced graph with $m_2(H) > m_2(C^{\ell_2})$. Let $\beta = \beta(C^{\ell_2}, H)$. Then there are constants B and $\kappa > 0$ such that, for $p = p(n) = Bn^{-\beta}$, almost every G_p satisfies the property that any subgraph $G' \subseteq G_p$ of G_p with $e(G') \geq e(G_p) - \kappa n^{1+1/(\ell_2-1)}/(\log n)^{|H|}$ is such that*

$$G' \to (H, C^{\ell_2}, \ldots, C^{\ell_r}).$$

Similarly to the case of Theorem 2, the result above also holds if H is strictly 2-balanced and $m_2(H) = m_2(C^{\ell_2})$. We further remark that the log-factor in the number of edges that one is allowed to delete from G_p can be got rid of.

The following result is an immediate corollary to Theorems 3 and 4.

Corollary 5 *Let $3 \leq \ell_1 \leq \cdots \leq \ell_r$ $(r \geq 2)$ be integers, and let $\beta = \beta(C^{\ell_2}, C^{\ell_1})$. Then there exist constants b and $B > 0$ such that*

(*a*) *for $p = p(n) \leq bn^{-\beta}$ we have*

$$\mathbf{P}\left[G_p \to (C^{\ell_1}, \ldots, C^{\ell_r})\right] = o(1),$$

(*b*) *for $p = p(n) \geq Bn^{-\beta}$ we have*

$$\mathbf{P}\left[G_p \to (C^{\ell_1}, \ldots, C^{\ell_r})\right] = 1 - o(1).$$

It would be interesting to determine the threshold functions for the properties $G_p \to (H_1, H_2)$ for arbitrary graphs H_1 and H_2, or more generally for the analogous problem when more than two graphs are involved. We formulate the following conjecture.

Conjecture 6 *Let H_1, H_2 be graphs with $m_2(H_1) \leq m_2(H_2)$ and let $\beta = \beta(H_1, H_2)$ be as defined in (1). Then $p = p(n) = n^{-\beta}$ is a sharp threshold for the property $G_p \to (H_1, H_2)$.*

Unfortunately, the techniques used in this note are by far not sufficient to yield a proof of this more general case.

We now single out two deterministic consequences of our methods. Let integers t, $\ell_2 \geq 3$ and a graph H be given. Let $\mathcal{V}_{\ell_2,H,t}$ be the class of graphs F with at most t vertices satisfying $|F| - \beta(C^{\ell_2}, H)e(F) < 1 + 1/(\ell_2 - 1)$.

Corollary 7 *For any fixed integers $3 \leq \ell_2 \leq \cdots \leq \ell_r$ $(r \geq 2)$ and $t \geq 1$ and any 2-balanced graph H with $m_2(H) > m_2(C^{\ell_2})$, there is a graph G that satisfies $G \to (H, C^{\ell_2}, \ldots, C^{\ell_r})$ and contains no member of $\mathcal{V}_{\ell_2,H,t}$.*

An immediate consequence of the corollary above is that, for any $3 \leq \ell_1 \leq \cdots \leq \ell_r$, there are graphs G of girth ℓ_1 satisfying $G \to (C^{\ell_1}, \dots, C^{\ell_r})$.

Our second deterministic result states that the r-tuple $(C^{\ell_1}, \dots, C^{\ell_r})$ is Ramsey-infinite for any $\ell_1, \dots, \ell_r \geq 3$. A graph G is called *Ramsey-critical*, or simply *critical*, for an r-tuple of graphs $(H_i)_{1\leq i\leq r}$ $(r \geq 2)$ if $G \to (H_i)_{1\leq i\leq r}$ but, for any proper subgraph $G' \subset G$ of G, the relation $G' \to (H_i)_{1\leq i\leq r}$ fails. The r-tuple $(H_i)_{1\leq i\leq r}$ is called *Ramsey-finite* if the class of all graphs that are critical for $(H_i)_{1\leq i\leq r}$ is finite and *Ramsey-infinite* otherwise. If $(H_i)_{1\leq i\leq r}$ is Ramsey-finite then there clearly is a polynomial time algorithm that, given a graph G, decides whether or not $G \to (H_i)_{1\leq i\leq r}$ holds.

Corollary 2 of [15] implies that the pair (H, H) is Ramsey-infinite for any graph H containing a cycle, and Luczak [11] showed that the pair (F, H) is Ramsey-infinite if F is a forest and H contains a cycle. Our result is as follows.

Theorem 8 *For any* $3 \leq \ell_1 \leq \cdots \leq \ell_r$ $(r \geq 2)$*, the* r*-tuple* $(C^{\ell_i})_{1\leq i\leq r}$ *is Ramsey-infinite.*

The organisation of this note is as follows. In Section 2 we give the proof of Theorem 2 and sketch the proof of Theorem 4. Theorem 3 is proved in Section 3. The deterministic consequences are proved in Section 4.

2 The proofs of Theorems 2 and 4

We shall need a fair amount of preparation to prove Theorems 2 and 4. In Sections 2.1 and 2.2, we set out some well known lemmas and some preliminary results that we shall need. In Section 2.3, we prove our crucial lemma, Lemma 17.

2.1 Basic tools

2.1.1 Janson's inequality

Let E be a finite set. Let $(J_e)_{e\in E}$ be a family of independent random indicator variables indexed by E such that $\mathbf{E}(J_e) = p$ for all $e \in E$. Suppose $(E_\alpha)_{\alpha\in\mathcal{A}}$ is a family of subsets of E. Let $I_\alpha = \prod_{e\in E_\alpha} J_e$, and $S = \sum_{\alpha\in\mathcal{A}} I_\alpha$. Thus, the random variable S counts how many of the sets E_α $(\alpha \in \mathcal{A})$ have all its elements $e \in E_\alpha$ such that $J_e = 1$. Let

$$\Delta = \sum_{\alpha,\beta} \mathbf{E}(I_\alpha I_\beta), \tag{4}$$

where the sum extends over all ordered pairs (α, β) with $\alpha \neq \beta$ but $E_\alpha \cap E_\beta \neq \emptyset$. Janson's inequality [5] is then the following.

Lemma 9 *For all* $0 \leq \varepsilon \leq 1$*, writing* $\mu = \mathbf{E}(S)$*, we have*

$$\mathbf{P}[S \leq (1-\varepsilon)\mu] \leq \exp\left\{-\frac{1}{2(1+\Delta/\mu)}\varepsilon^2\mu\right\}. \tag{5}$$

2.1.2 Szemerédi's lemma

The following notions will be needed in what follows. Let a graph J be fixed. In the sequel we let $V = V(J)$. Suppose $0 < \eta \le 1$, $D > 1$, and $0 < p \le 1$. We say that J is (η, D, p)*-upper-uniform* if for all U, $W \subseteq V$ with $U \cap W = \emptyset$ and $|U|$, $|W| \ge \eta n$, we have $e_J(U, W) \le Dp|U||W|$. In the sequel, for any two disjoint non-empty sets U, $W \subseteq V$, let the p*-density* of J between U and W be

$$d_{J,p}(U, W) = e_J(U, W)/p|U||W|.$$

Now suppose $0 < \varepsilon \le 1$ is a real number, and U, $W \subseteq V$ are two non-empty disjoint sets of vertices of J. We say that the pair (U, W) is (ε, J, p)*-regular* if for all $U' \subseteq U$, $W' \subseteq W$ with $|U'| \ge \varepsilon|U|$ and $|W'| \ge \varepsilon|W|$ we have

$$|d_{J,p}(U', W') - d_{J,p}(U, W)| \le \varepsilon.$$

We say that a partition $P = (V_i)_{i=0}^k$ of $V = V(J)$ is (ε, k)*-equitable* if $|V_0| \le \varepsilon n$, and $|V_1| = \ldots = |V_k|$. Also, we say that V_0 is the *exceptional* class of P. When the value of ε is not relevant, we refer to an (ε, k)-equitable partition as a k*-equitable* partition. Similarly, P is an *equitable* partition of V if it is a k-equitable partition for some k. Finally, we say that an (ε, k)-equitable partition $P = (V_i)_{i=0}^k$ of V is (ε, J, p)*-regular* if at most $\varepsilon\binom{k}{2}$ pairs (V_i, V_j) with $1 \le i < j \le k$ are not (ε, J, p)-regular. We may now state an extension of Szemerédi's lemma [16] to subgraphs of (η, D, p)-upper-uniform graphs, observed independently by Rödl and Kohayakawa [6].

Lemma 10 *For any given $\varepsilon > 0$, $k_0 \ge 1$, and $D > 1$, there are constants $\eta = \eta(\varepsilon, k_0, D) > 0$ and $K_0 = K_0(\varepsilon, k_0, D) \ge k_0$ that depend only on ε, k_0, and D such that any (η, D, p)-upper-uniform graph J admits an (ε, J, p)-regular (ε, k)-equitable partition of its vertex set with $k_0 \le k \le K_0$.*

2.1.3 Nearly regular subgraphs

Let us introduce a piece of notation. Let a positive integer t be given. If $U_1, \ldots, U_t \subseteq V(J)$ are pairwise disjoint sets of vertices of a given graph J, we write $J[U_1, \ldots, U_t]$ for the t-partite subgraph of J naturally defined by the U_i $(1 \le i \le t)$. Thus, $J[U_1, \ldots, U_t]$ has vertex set $\bigcup_{i=1}^t U_i$ and two of its vertices are adjacent in $J[U_1, \ldots, U_t]$ if and only if they are adjacent in J and, moreover, they belong to distinct U_i.

Now suppose we have real numbers $0 < p \le 1$, $0 < \varepsilon \le 1$, and $0 < \gamma_0 \le 1$. Let also $m_i = |U_i|$ $(1 \le i \le t)$, and write γ_{ij} for the p-density $d_{J,p}(U_i, U_j)$ for all distinct i and j. Suppose the pair (U_i, U_{i+1}) is (ε, J, p)-regular and $\gamma_{i,i+1} \ge \gamma_0$ for any $1 \le i < t$.

We may now state our next lemma. In what follows, we write $O_1(x)$ for any term y satisfying $|y| \le x$. We write $\Gamma_J(x)$ for the J-neighbourhood of a vertex $x \in V(J)$.

Lemma 11 *Let J and the sets U_i $(1 \le i \le t)$ be as above. Suppose $0 < \varepsilon \le 1/5$ and put $A = 4 + 1/\gamma_0$. Then there are sets $\bar{U}_i \subseteq U_i$ with $|\bar{U}_i| \ge (1 - 4\varepsilon)m_i$ for all $1 \le i \le t$ such that, for all i, $j \in [t]$ with $|i - j| = 1$ and any $x \in \bar{U}_i$, we have*

$$d_{ij}(x) = |\Gamma_J(x) \cap \bar{U}_j| = (1 + O_1(A\varepsilon))\gamma_{ij} p m_j.$$

Lemma 11 above is very similar to Lemma 2 from [8] and Lemma 5 from [9], and hence its proof is omitted.

2.1.4 Ramsey's theorem

Here we give the following easy consequence of Ramsey's theorem.

Lemma 12 *Let graphs $H_1, \dots, H_r$ $(r \ge 1)$ be given. Then there exist positive constants $c = c(H_1, \dots, H_r) > 0$ and $k_0 = k_0(H_1, \dots, H_r)$ for which the following holds. If $k \ge k_0$ and K^k is given an arbitrary r-edge-colouring, then we necessarily have, for some $1 \le i \le r$, at least $ck^{|H_i|}$ monochromatic copies of H_i of colour i.*

Proof. Let $k_0 = q$ and $c = (1/r)q^{-\max_i |H_i|}$, where

$$q = R(H_1, \dots, H_r) = \min\{n \colon K^n \to (H_1, \dots, H_r)\},$$

the Ramsey number for $(H_1, \dots, H_r)$. Now let K^k with $k \ge k_0$ be given an arbitrary r-edge-colouring. Then, each q-subset of the vertex set of K^k gives, for some i, a monochromatic copy of H_i of colour i. For some i, at least $(1/r)\binom{k}{q}$ of these q-subsets gives a monochromatic H_i of colour i. Each such copy is counted at most $\binom{k-|H_i|}{q-|H_i|}$ times. The number of such copies is, therefore, at least

$$\frac{1}{r}\binom{k}{q} \Big/ \binom{k-|H_i|}{q-|H_i|} = \frac{1}{r}\frac{(k)_{|H_i|}}{(q)_{|H_i|}} \ge \frac{1}{r}\left(\frac{k}{q}\right)^{|H_i|} \ge ck^{|H_i|},$$

as required. □

2.2 Preliminary lemmas

Throughout this section we fix an integer $\ell \ge 3$ and a graph H as in the statement of Theorem 2. Moreover, here we shall always suppose that $p = p(n) = Bn^{-\beta}$, where β is as defined in (2) and $B > 0$ is some constant. We also sometimes write h for the order of our graph H. For the sake of definiteness, we shall assume that $V(G_p) = [n]$.

If H_1 is a copy of H in G_p we say that H_1 is *isolated* if it does not intersect any other copy of H present in G_p on an edge. In other words, if we have $E(H_1) \cap E(H_2) = \emptyset$ for any other copy H_2 of H in G_p.

2.2.1 Isolated copies of H in G_p

Let us write $T = T(G_p)$ for the number of isolated copies of H in G_p. Let $h = |H|$. Using the fact that H is 2-balanced and that $m_2(C^\ell) < m_2(H)$, it is easy to see that almost all copies of H in G_p are isolated but we shall not need this here. Let us write p_0 for the expected number of (not necessarily isolated) copies of H in G_p that contain some fixed edge $e \in \binom{[n]}{2}$. Clearly, the expected number μ_H of copies of H in G_p is $|\mathrm{Aut}(H)|^{-1}(n)_h p^{e(H)}$, where $\mathrm{Aut}(H)$ denotes the automorphism group of H. A moment's thought then shows that

$$p_0 = \frac{(n)_h e(H)}{|\mathrm{Aut}(H)|\binom{n}{2}} p^{e(H)} = (2 + o(1)) \frac{e(H) B^{e(H)}}{|\mathrm{Aut}(H)|} n^{-1+1/(\ell-1)}. \tag{6}$$

We now consider the following construction we shall use later. Let $\mathcal{H} = \{H_1, \ldots, H_T\}$ be the set of isolated copies of H in G_p. We shall always write $\widetilde{H} = \widetilde{H}(G_p)$ for the spanning subgraph of G_p whose edge set is the union of the edge sets of the H_i $(1 \le i \le T)$. Clearly, we have $e(\widetilde{H}) = Te(H)$.

If we denote by $\widetilde{p}$ the probability that a given edge $e \in \binom{[n]}{2}$ is contained in $\widetilde{H}$ then $\widetilde{p} = (1 + o(1))p_0$ as almost all copies of H are isolated. In the sequel we shall only use p_0, which may be considered to be the edge probability of $\widetilde{H}$.

Next we show that the edges of $\widetilde{H}$ satisfy some negative correlation properties. For $E \subseteq \binom{[n]}{2}$, write $E \sqsubset E(\widetilde{H})$ if $E \subseteq E(\widetilde{H})$ and all edges from E belong to *different* isolated copies of H in G_p.

Lemma 13 *For any set $E \subseteq \binom{[n]}{2}$,*

$$\mathbf{P}[E \sqsubset E(\widetilde{H})] \le {p_0}^{|E|}.$$

Proof. Suppose $E = \{e_1, \ldots, e_a\}$ where $a = |E|$. Let I be the set of all a-tuples $(H_1, \ldots, H_a)$ of H-subgraphs of K^n with $e_i \in E(H_i)$ for all $1 \le i \le a$. Also, let J be the set of such a-tuples $(H_1, \ldots, H_a) \in I$ with the H_i $(1 \le i \le a)$ pairwise edge-disjoint. Finally, let $\mathcal{H}_i$ be set of H-subgraphs of K^n that contain the edge e_i $(1 \le i \le a)$. Thus, in particular, $I = \mathcal{H}_1 \times \cdots \times \mathcal{H}_a$. Then

$$\begin{aligned} \mathbf{P}[E \sqsubset E(\widetilde{H})] &\le \sum_{(H_1,\ldots,H_a)\in J} \mathbf{P}[H_1 \subseteq G_p, \ldots, H_a \subseteq G_p] \\ &= \sum_{(H_1,\ldots,H_a)\in J} \mathbf{P}[H_1 \subseteq G_p] \cdots \mathbf{P}[H_a \subseteq G_p] \\ &\le \sum_{(H_1,\ldots,H_a)\in I} \mathbf{P}[H_1 \subseteq G_p] \cdots \mathbf{P}[H_a \subseteq G_p] \end{aligned}$$

$$
\begin{aligned}
&= \sum_{H_1 \in \mathcal{H}_1} \mathbf{P}[H_1 \subseteq G_p] \cdots \sum_{H_a \in \mathcal{H}_a} \mathbf{P}[H_a \subseteq G_p] \\
&= {p_0}^a,
\end{aligned}
$$

as required. □

2.2.2 Upper-uniformity results for $\widetilde{H}$

In this section we show that the edges of $\widetilde{H}$ are fairly well distributed in the sense that $\widetilde{H}$ is, almost surely, locally sparse. Our first lemma implies that almost every $\widetilde{H}$ is $(\eta, 3e(H), p_0)$-upper-uniform for any fixed $0 < \eta \leq 1$.

Lemma 14 *Let $\omega = \omega(n) \to \infty$ as $n \to \infty$. Then, for almost every G_p, the graph $\widetilde{H} = \widetilde{H}(G_p)$ is such that the following holds. Suppose U, $W \subseteq V(\widetilde{H})$ with $U \cap W = \emptyset$ are such that $|U||W| \geq \omega n^{2-1/(\ell-1)}$. Then we have*

$$
e_{\widetilde{H}}(U, W) \leq 3e(H)p_0|U||W|. \tag{7}
$$

Proof. If a fixed pair (U, W) with $u = |U|$ and $w = |W|$ does not satisfy (7) then there is a set $E \subseteq U \times W$ with $E \subseteq E(\widetilde{H})$ and $|E| = \lceil 3e(H)p_0uw \rceil$. This implies the existence of a set $E' \subseteq U \times W$ of size $|E'| = \lceil 3p_0uw \rceil$ such that all edges from E' belong to different isolated copies of H in G_p. The probability of this event is by Lemma 13 at most

$$
\binom{uw}{\lceil 3p_0uw \rceil} {p_0}^{\lceil 3p_0uw \rceil} \leq \left(\frac{\mathrm{e}}{3}\right)^{3p_0uw}.
$$

As there are at most 2^{2n} choices for the sets U and W, the lemma follows from the fact that $3p_0uw/n \to \infty$ as $n \to \infty$ (cf. (6)). □

Our next lemma concerns the number of $\widetilde{H}$-edges between sets U and W smaller than the ones considered in Lemma 14, and hence may be thought of as a more local condition on the sparseness of $\widetilde{H}$. To state our lemma concisely, we introduce a definition. Let J be a graph on n vertices and suppose $C \geq 1$ and $\ell \geq 1$ are given. We say that J is *locally (p_0, C, ℓ)-upper-uniform* if for all pairs (U, W) of disjoint sets of vertices U, $W \subseteq V(J)$ satisfying

$$
|U| \leq |W| \leq d_0|U| \leq {d_0}^{\ell-2}, \tag{8}
$$

where $d_0 = p_0 n$, we have

$$
e_J(U, W) \leq C|W|. \tag{9}
$$

Lemma 15 *Almost every G_p is such that $\widetilde{H} = \widetilde{H}(G_p)$ is locally $(p_0, 2\ell e(H), \ell)$-upper uniform.*

Proof. The probability that a fixed pair (U, W) of disjoint sets of vertices with $u = |U|$ and $w = |W|$ does not satisfy (9) can be bounded similarly as in the proof of Lemma 14 by

$$\binom{uw}{\lceil 2\ell w\rceil} {p_0}^{\lceil 2\ell w\rceil} \leq \left(\frac{\mathrm{e}u}{2\ell} p_0\right)^{2\ell w}.$$

Observe that, for pairs (U, W) as in (8), we have $u \leq (p_0 n)^{\ell-3} \leq An^{(\ell-3)/(\ell-1)}$ for some constant $A = A(H, B) > 0$. Moreover, we have $u \leq w = o(n)$ as $n \to \infty$, and hence $\binom{n}{u}\binom{n}{w} \leq \binom{n}{w}^2$. Therefore, the expected number of pairs (U, W) satisfying (8) but not (9) is at most

$$\sum_{w=1}^{\lfloor n/2\rfloor} w \binom{n}{w}^2 \left(\frac{\mathrm{e}u}{2\ell} p_0\right)^{2\ell w} \leq \sum_{w=1}^{\lfloor n/2\rfloor} w \left(n^{1/\ell} \frac{\mathrm{e}^{1+1/\ell}}{2\ell} A n^{(\ell-3)/(\ell-1)} p_0\right)^{2\ell w} = o(1)$$

(cf. (6)) and the lemma follows from Markov's inequality. $\square$

2.2.3 The distribution of isolated copies of H in G_p

In this section we prove a simple lemma concerning the existence of a large number of isolated copies of H in G_p when p is as given in Theorem 2. In fact, for technical reasons, we shall need to look at the existence of such copies of H with each of their vertices taken from specified subsets of vertices of G_p.

Label the vertices of H by $v_1, \dots, v_h$ where $h = |H|$. Suppose $W_1, \dots, W_h$ are pairwise disjoint subsets of vertices of G_p. Put $\mathbf{W} = (W_i)_{i=1}^h$. Here we are interested in the number $Z = Z_{\mathbf{W}}(G_p)$ of injections $\iota\colon V(H) \to V(G_p)$ taking v_i into W_i for all $1 \leq i \leq h$, and such that $\iota(v_i)\iota(v_j) \in E(G_p)$ whenever $v_iv_j \in E(H)$. We shall refer to such maps ι as $\mathbf{W}$*-embeddings* of H in G_p. Moreover, we say that the subgraph of G_p with vertex set $\iota(V(H))$ and edge set $\{\iota(v_i)\iota(v_j)\colon v_iv_j \in E(H)\}$ is the *image* $\operatorname{im}\iota$ of ι.

Finally, let us define $Y = Y_{\mathbf{W}}(G_p)$ to be the number of $\mathbf{W}$-embeddings ι of H in G_p with $\operatorname{im}\iota$ an isolated H-subgraph of G_p.

Lemma 16 *Suppose $B > 0$ is a constant and β is as defined in* (2). *Then, for $p = Bn^{-\beta}$, almost every G_p has the following property. For any vector $\mathbf{W} = (W_i)_{i=1}^h$ of pairwise disjoint sets $W_i \subseteq [n]$ $(1 \leq i \leq h)$, each of cardinality at least $n/\log n$, we have $Y = Y_{\mathbf{W}}(G_p) \geq (1/2)\mathbf{E}(Z_{\mathbf{W}})$.*

Proof. Define $\mathcal{U}$ to be the set of all graphs that can be written as a union of two copies of H intersecting in at least one edge, not including H itself. Let then $X = X(G_p)$ count the number of graphs from $\mathcal{U}$ present in G_p. We first aim at bounding X. From $m_2(C^\ell) <$

$m_2(H)$ it is easy to see that $\beta > 1/m_2(H)$. Let $\delta = \beta - 1/m_2(H) > 0$. As H is 2-balanced we have that for any proper subgraph J of H containing at least one edge

$$\frac{|H|-|J|}{e(H)-e(J)} \leq \frac{1}{m_2(H)} = \beta - \delta.$$

Therefore, writing $\sum_{K^2 \subseteq J \subset H}$ for the sum over all proper subgraphs $J \subset H$ containing at least one edge, we have

$$\begin{aligned}
\mathbf{E}(X) &= \sum_{U \in \mathcal{U}} \frac{(n)_{|U|}}{|\mathrm{Aut}(U)|} p^{e(U)} \\
&= O\Big(\sum_{K^2 \subseteq J \subset H} n^{2|H|-|J|} p^{2e(H)-e(J)} \Big) \\
&= O\Big(\sum_{K^2 \subseteq J \subset H} n^{2|H|-|J|-\beta(2e(H)-e(J))} \Big) \\
&= O\left(n^{|H|-\beta e(H)-\delta(e(H)-e(J))} \right) \\
&= O(n^{1+1/(\ell-1)-\delta(e(H)-e(J))}).
\end{aligned}$$

Next we use Janson's inequality (5) to show that $Z_{\mathbf{W}}$ is with high probability not too small. So fix $\mathbf{W} = (W_i)_{i=1}^h$ and let $(H_\alpha : \alpha \in \mathcal{A})$ be the family of all graphs which are isomorphic to H and whose vertex set can be written as $\{v_1, \ldots, v_h\}$ with $v_i \in W_i$ for all $1 \leq i \leq h$. Thus $|\mathcal{A}| = |W_1| \ldots |W_h| \geq (n/\log n)^h$, and the random variable S defined in Section 2.1.1 coincides with $Z_{\mathbf{W}}$. We have that

$$\mathbf{E}(Z_{\mathbf{W}}) \geq \left(\frac{n}{\log n}\right)^{|H|} B^{e(H)} n^{-\beta e(H)} = B^{e(H)} n^{1+1/(\ell-1)-|H|\log\log n/\log n}.$$

By Markov's inequality we have that $X = o(\mathbf{E}(Z_{\mathbf{W}}))$ almost surely and therefore also $\Delta \leq 2X = o(\mathbf{E}(Z_{\mathbf{W}}))$ where Δ is as defined in (4). Then Janson's inequality implies that

$$\mathbf{P}\left[Z_{\mathbf{W}} \leq \frac{3}{4}\mathbf{E}(Z_{\mathbf{W}})\right] \leq \exp\left(-\frac{1}{33}\mathbf{E}(Z_{\mathbf{W}})\right) \leq \exp(-n^{1+1/\ell}),$$

if n is large enough. On the other hand, there are at most 2^{hn} choices for $\mathbf{W}$. Thus for almost every G_p we have $Z_{\mathbf{W}} \geq (3/4)\mathbf{E}(Z_{\mathbf{W}})$ for every $\mathbf{W}$. Furthermore we almost surely have that $X \leq \mathbf{E}(Z_{\mathbf{W}})/8$ and therefore also $Y_{\mathbf{W}}(G_p) \geq Z_{\mathbf{W}} - 2X \geq \mathbf{E}(Z_{\mathbf{W}})/2$ for every $\mathbf{W}$, almost surely, as required. $\square$

2.3 The key counting lemma

Throughout this section, $\ell \geq 3$ is a fixed integer and $\mathbf{V} = (V_1, \ldots, V_\ell)$ is a fixed vector of pairwise disjoint sets $V_1, \ldots, V_\ell$, all of the same cardinality, say m.

In this section we shall be interested in counting certain graphs F whose vertex set is $V(F) = \bigcup_{i=1}^{\ell} V_i$ and all whose edges join vertices in consecutive classes V_i. More precisely, if e is an edge of F, then its two endpoints will belong to V_{i-1} and V_i, respectively, for some $1 \leq i \leq \ell$. Here and below, the indices of the sets V_i will be taken modulo ℓ. For conciseness, given $\mathbf{V} = (V_i)_{i=1}^{\ell}$, let us say that a graph F as described above is a $\mathbf{V}$-*graph*. To define the precise subclass of $\mathbf{V}$-graphs F that we shall be interested in, however, we need to introduce some more parameters and definitions.

Let $0 < \varepsilon \leq 1$, $\bar{B} > 0$, $0 < \gamma_0 \leq 1$, $C \geq 1$, and $D \geq 1$ be real numbers and let $\bar{T}$ be a positive integer. We shall write

$$\mathcal{F}_m(\varepsilon, \bar{B}, \gamma_0, C, D; \mathbf{V}, \bar{T})$$

for the set of $\mathbf{V}$-graphs F satisfying conditions (i)–(iv) given below.

(i) The graph F has $\bar{T}$ edges.

(ii) The graph F is such that all the pairs (V_{i-1}, V_i) $(1 \leq i \leq \ell)$ are $(\varepsilon, F, \bar{p})$-regular, where $\bar{p} = \bar{p}(m) = \bar{B} m^{-1+1/(\ell-1)}$.

(iii) The $\bar{p}$-density $\gamma_{i-1,i} = d_{F,\bar{p}}(V_{i-1}, V_i)$ of F between the sets V_{i-1} and V_i is such that, for all $1 \leq i \leq \ell$, we have

$$\gamma_0 \leq \gamma_{i-1,i} \leq D. \tag{10}$$

(iv) For all $U \subseteq V_{i-1}$ and $W \subseteq V_i$ $(2 \leq i < \ell)$ with

$$|U| \leq |W| \leq \bar{d}|U| \leq \bar{d}^{\ell-2}, \tag{11}$$

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

$$e_F(U, W) \leq C|W|. \tag{12}$$

Let us now define the class

$$\mathcal{F}'_m(\varepsilon, \bar{B}, \gamma_0, C, D; \mathbf{V}, \bar{T}) \subseteq \mathcal{F}_m(\varepsilon, \bar{B}, \gamma_0, C, D; \mathbf{V}, \bar{T})$$

as the class of graphs $F \in \mathcal{F}_m(\varepsilon, \bar{B}, \gamma_0, C, D; \mathbf{V}, \bar{T})$ that do not contain a copy of C^ℓ as a subgraph. Our aim in this section is to estimate the cardinality $f'_m(\varepsilon, \bar{B}, \gamma_0, C, D; \bar{T})$ of $\mathcal{F}'_m(\varepsilon, \bar{B}, \gamma_0, C, D; \mathbf{V}, \bar{T})$ from above.

Lemma 17 *Let an integer $\ell \geq 3$ be fixed, and let constants $0 < \alpha \leq 1$, $0 < \gamma_0 \leq 1$, $C \geq 1$, and $D \geq 1$ be given. Then there are constants $0 < \varepsilon \leq 1$, $\bar{B}_0 > 0$, and m_0 that* *depend only on ℓ, α, γ_0, C, and D such that, for all integers $m \geq m_0$ and $\bar{T} \geq 1$ and real numbers $\bar{B} \geq \bar{B}_0$, we have*

$$f'_m(\varepsilon, \bar{B}, \gamma_0, C, D; \bar{T}) \leq \alpha^{\bar{T}} \binom{(\ell+2)m^2}{\bar{T}}. \tag{13}$$

The proof of Lemma 17 is based on a technical lemma, Lemma 18, that we now describe. Let a constant $0 < \delta \leq 1$ be given, and suppose we have real numbers $0 < \varepsilon \leq 1$, $0 < \gamma_0 \leq 1$, $\bar{B}$, C, $D \geq 1$, and an integer $\bar{T} \geq 1$. Let

$$F \in \mathcal{F}_m(\varepsilon, \bar{B}, \gamma_0, C, D; \mathbf{V}, \bar{T})$$

be fixed. Lemma 18 will concern the $(\ell-1)$-partite subgraph $\dot{F} = F[V_1, \ldots, V_{\ell-1}]$ of F induced by the vertex classes $V_1, \ldots, V_{\ell-1}$. Suppose $y_1, \ldots, y_{j-1} \in V_1$ are $j-1$ given vertices of V_1, where $1 \leq j \leq \delta m/2$. Let

$$U_1 = V_1 \setminus \{y_1, \ldots, y_{j-1}\} \tag{14}$$

and

$$U_{\ell-1} = V_{\ell-1} \setminus \bigcup_{q=1}^{j-1} \dot{\Gamma}_F^{(\ell-2)}(y_q), \tag{15}$$

where $\dot{\Gamma}_F^{(\ell-2)}$ denotes the $(\ell-2)$-fold iteration of $\dot{\Gamma}_F = \Gamma_{\dot{F}}$. Suppose $|U_{\ell-1}| \geq \delta m$. Our technical lemma tells us that then, provided ε is small enough, most vertices in U_1 are 'good', in the sense that their $(\ell-2)$nd $\dot{F}$-neighbourhood within $U_{\ell-1}$ is large.

Lemma 18 *Let $0 < \delta \leq 1$, $0 < \gamma_0 \leq 1$, $\bar{B} > 0$, $C \geq 1$, and $D \geq 1$ be given. Then if $0 < \varepsilon \leq \min\{\delta/4, \gamma_0/(8\gamma_0+2)\}$, and m and $\bar{T}$ are integers, any graph $F \in \mathcal{F}_m(\varepsilon, \bar{B}, \gamma_0, C, D; \mathbf{V}, \bar{T})$ is such that at least $(1-\delta)|U_1|$ vertices $y \in U_1$ satisfy*

$$\left|\dot{\Gamma}_F^{(\ell-2)}(y) \cap U_{\ell-1}\right| \geq 2^{-\ell+2} \frac{\gamma_0^{\ell-2}}{C^{\ell-3}} \delta \bar{d}^{\ell-2}, \tag{16}$$

where $\bar{d} = \bar{p}m$.

Proof of Lemma 17. Let ℓ, α, γ_0, C, and D as in the statement of our lemma be given. Let us define the constants

$$\begin{aligned}
\delta &= \min\{1/2, (\alpha^{\ell D} 2^{-5D})^{8/\gamma_0}\}, \\
\varepsilon &= \min\{\delta/4, \gamma_0/(8\gamma_0+2)\}, \\
\bar{B}_0 &= \left(\frac{4C^{\ell-3}}{2^{-\ell+2}\gamma_0^{\ell-1}\delta}\right)^{1/(\ell-1)}, \\
m_0 &= \left(\frac{3}{2} D \bar{B}_0\right)^{(\ell-1)/(\ell-2)}.
\end{aligned}$$

Our aim is to show that $\varepsilon, \bar{B}_0$ and m_0 given above will do in Lemma 17. Thus, let integers $m \geq m_0$ and $\bar{T}$ and a real $\bar{B} \geq \bar{B}_0$ be given. We have to show that (13) holds. We estimate $f'_m(\varepsilon, \bar{B}, \gamma_0, C, D; \bar{T})$ from above by estimating the number of ways we may 'generate' an element from $\mathcal{F}'_m(\varepsilon, \bar{B}, \gamma_0, C, D; \mathbf{V}, \bar{T})$.

In order to generate a graph F from $\mathcal{F}'_m(\varepsilon, \bar{B}, \gamma_0, C, D; \mathbf{V}, \bar{T})$, we first choose $\bar{T}_1 = e_F(V_1, V_\ell)$ and $\bar{T}_{\ell-1} = e_F(V_{\ell-1}, V_\ell)$. Later on we shall sum over all possible values of $\bar{T}_1$ and $\bar{T}_{\ell-1}$. Observe that for $i \in \{1, \ell\}$ we have

$$\gamma_0 \bar{B} m^{1+1/(\ell-1)} \leq \bar{T}_i = d_{F,\bar{p}}(V_i, V_\ell) \bar{p} m^2 \leq D \bar{B} m^{1+1/(\ell-1)}. \tag{17}$$

Put $\bar{T}_0 = \bar{T} - \bar{T}_1 - \bar{T}_{\ell-1}$. Next we fix $\dot{F} = F[V_1, \ldots, V_{\ell-1}]$, for which we have not more than $\binom{(\ell-2)m^2}{\bar{T}_0}$ possibilities. By Lemma 11 for $t = 3$, we know that for any choice of the edges of F incident to V_ℓ there have to be sets $\bar{V}_i \subseteq V_i$ $(i \in \{1, \ell-1, \ell\})$ such that for any $x \in \bar{V}_\ell$ and any $i \in \{1, \ell-1\}$ we have

$$d_i(x) = |\Gamma_F(x) \cap \bar{V}_i| = (1 + O_1(A\varepsilon)) \frac{\bar{T}_i}{\bar{p}|V_\ell||V_i|} \bar{p}|V_i| = (1 + O_1(A\varepsilon)) \frac{\bar{T}_i}{m}, \tag{18}$$

where $A = 4 + 1/\gamma_0$. Call the F-edges in $\bar{V}_\ell \times \bar{V}_i$ $(i \in \{1, \ell-1\})$ *regular* edges and the remaining F-edges between V_ℓ and V_i *irregular* edges. For $i \in \{1, \ell-1\}$, denote by $\bar{T}_i^{(\mathrm{irr})}$ the number of irregular edges incident to V_i. Summing up we have that, for given $\bar{T}_0$ and $\bar{T}_i^{(\mathrm{irr})}$, there are at most

$$\binom{(\ell-2)m^2}{\bar{T}_0} \binom{m^2}{\bar{T}_1^{(\mathrm{irr})}} \binom{m^2}{\bar{T}_{\ell-1}^{(\mathrm{irr})}} \tag{19}$$

ways of choosing the set of edges of F, except for the regular edges incident to V_ℓ.

Now we have to bound the number of choices for the regular edges. Let us first choose the degree sequences $(d_i(x))_{x \in \bar{V}_\ell}$ of the vertices in $\bar{V}_\ell$, where $d_i(x) = |\Gamma_F(x) \cap \bar{V}_i|$ for $i \in \{1, \ell-1\}$. For a vertex $x \in \bar{V}_\ell$ we may assume by (17), (18) and the definition of m_0 that

$$d_{\min} := \frac{1}{2} \gamma_0 \bar{B} m^{1/(\ell-1)} \leq d_i(x) \leq m. \tag{20}$$

Let us now fix a vertex x in $\bar{V}_\ell$, and let us say that a set $Y \subseteq V_1$ of size $d_{\min} \leq d' = |Y| \leq m$ is *good* if $|\dot{\Gamma}_F^{(\ell-2)}(Y) \cap V_{\ell-1}| \geq (1-\delta)m$. The idea is that if we choose a good set Y to be the neighbourhood of x within V_1, then the number of vertices in $V_{\ell-1}$ that are allowed to be adjacent to x in F is very small, namely, at most δm, since we are supposed to generate a C^ℓ-free graph F. Now, if a d'-set $Y \subseteq V_1$ as above is not good then, for any ordering $y_1, y_2, \ldots, y_{d'}$ of the vertices of Y, there are at least $d'/2$ indices i with

$$\left| \dot{\Gamma}_F^{(\ell-2)}(y_i) \setminus \bigcup_{q=1}^{i-1} \dot{\Gamma}_F^{(\ell-2)}(y_q) \right| < \frac{(1-\delta)m}{d'/2} \leq 2^{-\ell+2} \frac{\gamma_0^{\ell-2}}{C^{\ell-3}} \delta \bar{d}^{\ell-2},$$

where the last inequality follows from the definition of $\bar{B}_0$ and (20). For these indices i, by Lemma 18, there are only δm choices for y_i once $y_1, \ldots, y_{i-1}$ are chosen. Therefore the number of bad sets Y of size $d_{\min} \leq d' \leq m$ is at most

$$\binom{d'}{\lceil d'/2 \rceil} \delta^{d'/2} \frac{m^{d'}}{d'!} \leq 2^{d'} \delta^{d_{\min}/2} \mathrm{e}^{d'} \binom{m}{d'}.$$

If the neighbourhood $\Gamma_F(x) \cap V_1$ of the vertex $x \in \bar{V}_\ell$ in V_1 is good then, since $C^\ell \not\subseteq F$, there are only $\binom{\delta m}{d_{\ell-1}(x)} \leq \delta^{d_{\ell-1}(x)} \binom{m}{d_{\ell-1}(x)}$ choices for the neighbourhood $\Gamma_F(x) \cap V_{\ell-1}$ of x within $V_{\ell-1}$. On the other hand, if $\Gamma_F(x) \cap V_1$ happens to be bad, then trivially there are at most $\binom{m}{d_{\ell-1}(x)}$ choices for this neighbourhood. Therefore there are at most

$$\begin{aligned} \binom{m}{d_1(x)} \delta^{d_{\ell-1}(x)} \binom{m}{d_{\ell-1}(x)} &+ \delta^{d_{\min}/2} 2^{3d_1(x)} \binom{m}{d_1(x)} \binom{m}{d_{\ell-1}(x)} \\ &\leq \delta^{d_{\min}/2} 2^{3d_1(x)+1} \binom{m}{d_1(x)} \binom{m}{d_{\ell-1}(x)} \end{aligned}$$

ways of choosing the regular edges incident to our vertex $x \in \bar{V}_\ell$.

Let us write $\sum'_{\bar{T}}$ for the sum over all partitions $\bar{T}_0 + \bar{T}_1 + \bar{T}_{\ell-1} = \bar{T}$ of $\bar{T}$, and similarly $\sum'_{\bar{T}_1}$ and $\sum'_{\bar{T}_{\ell-1}}$ for the sums over all partitions $\bar{T}_1^{(\mathrm{irr})} + d_1 + \ldots + d_m = \bar{T}_1$ of $\bar{T}_1$ and $\bar{T}_{\ell-1}^{(\mathrm{irr})} + d'_1 + \ldots + d'_m = \bar{T}_{\ell-1}$ of $\bar{T}_{\ell-1}$. Let us also write $\sum_{\bar{V}_\ell}$ for the sum over all possible choices of $\bar{V}_\ell$. Then, recalling the bound (19), we see that $f'_m(\varepsilon, \bar{B}, \gamma_0, C, D; \bar{T})$ is at most

$$\begin{aligned} \sum\nolimits'_{\bar{T}} \sum\nolimits'_{\bar{T}_1} \sum\nolimits'_{\bar{T}_{\ell-1}} \sum\nolimits_{\bar{V}_\ell} & \binom{(\ell-2)m^2}{\bar{T}_0} \binom{m^2}{\bar{T}_1^{(\mathrm{irr})}} \binom{m^2}{\bar{T}_{\ell-1}^{(\mathrm{irr})}} \\ & \times \delta^{|\bar{V}_\ell| d_{\min}/2} \prod_{j=1}^{m} \left(2^{3d_j+1} \binom{m}{d_j} \binom{m}{d'_j} \right) \\ \leq \delta^{m d_{\min}/4} 2^{3\bar{T}_1 + 2m} \binom{(\ell+2)m^2}{\bar{T}} & \leq \alpha^{\bar{T}} \binom{(\ell+2)m^2}{\bar{T}}, \end{aligned}$$

where we apply Vandermonde's identity and the fact that, trivially, there are at most 2^m choices for $\bar{V}_\ell$ to deduce the first inequality and we use (17), (20), that $\bar{T} \leq \ell D \bar{B} m^{1+1/(\ell-1)}$, and the definition of δ to deduce the second inequality. □

Proof of Lemma 18. Let ε, $\bar{B}$, γ_0, C, D, and $\bar{T}$ be as in the statement of our lemma. Fix a graph $F \in \mathcal{F}_m(\varepsilon, \bar{B}, \gamma_0, C, D; \mathbf{V}, \bar{T})$ and let us consider $\dot{F} = F[V_1, \ldots, V_{\ell-1}]$. Recall (14) and (15) and put $U_i = V_i$ for all $2 \leq i \leq \ell - 2$. The first step is to apply Lemma 11 with $t = \ell - 1$ to get rid of vertices of small degree. We obtain sets $\bar{U}_i \subseteq U_i$ $(1 \le i < \ell)$ with

$$|\bar{U}_i| \ge (1 - 4\varepsilon)|U_i| \ge (1 - \delta)|U_i| \tag{21}$$

for all $1 \le i < \ell$, and such that, for all $2 \le i < \ell$ and all $x \in \bar{U}_{i-1}$, we have

$$\begin{aligned} d_{i-1,i}(x) &= |\Gamma_F(x) \cap \bar{U}_i| \\ &\ge \left(1 - \left(\frac{4\gamma_0 + 1}{\gamma_0}\right)\varepsilon\right) d_{F,\bar{p}}(U_{i-1}, U_i)\bar{p}|U_i| \\ &\ge (\gamma_0/2)\bar{p}|U_i|. \end{aligned} \tag{22}$$

In the case $\ell = 3$ we are already finished if we let $i = 2$ in (22). If $\ell > 3$ we use property (*iv*) of F to prove the following assertion by induction on i.

Claim 1 *For any $y \in \bar{U}_1$ and any $2 \le i \le \ell - 2$, we have*

$$\left|\dot{\Gamma}_F^{(i-1)}(y) \cap \bar{U}_i\right| \ge 2^{-i+1}\frac{\gamma_0^{i-1}}{C^{i-2}}\bar{d}^{i-1}.$$

The claim for $i = 2$ is just (22). For $2 < i \le \ell - 2$, let $U \subseteq \dot{\Gamma}_F^{(i-2)}(y) \cap \bar{U}_{i-1}$ be a set of size $\lceil 2^{-i+2}\gamma_0^{i-2}C^{-i+3}\bar{d}^{i-2}\rceil$. Assume that $W = \Gamma_F(U) \cap \bar{U}_i$ has less than $2^{-i+1}\gamma_0^{i-1}C^{-i+2}\bar{d}^{i-1}$ elements. Because $|U| \le \bar{d}^{\ell-3}$, we may apply (12) to yield

$$e_F(U, W) \le C|W| < \frac{\gamma_0 \bar{d}}{2}|U|.$$

Because $\bar{d} = \bar{p}|U_i|$ for $i \le \ell - 2$ this is a contradiction to (22), and hence the claim is proven.

Finally, using the fact that $\bar{p}|U_{\ell-1}| \ge \bar{p}\delta m = \delta\bar{d}$, the same expansion argument implies that for any $y \in \bar{U}_1$ we have

$$\left|\dot{\Gamma}_F^{(\ell-2)}(y) \cap \bar{U}_{\ell-1}\right| \ge 2^{-\ell+2}\frac{\gamma_0^{\ell-2}}{C^{\ell-3}}\delta\bar{d}^{\ell-2}. \tag{23}$$

We have thus shown that any $y \in \bar{U}_1$ is such that (16) holds. Since by (21) we have $|\bar{U}_1| \ge (1 - \delta)|U_1|$, the proof of the lemma is finished. $\square$

2.4 Proof of Theorem 2

Proof of Theorem 2. Let ℓ and H be as in the statement of the theorem. We shall start by defining some constants. Let first k_0 and c be the constants from Lemma 12 applied to the triple (C^ℓ, H, K^2). Set $C = 2\ell e(H)$, $D = 3e(H)$,

$$\gamma_0 = \min\{1, c2^{-h-1}/e(H)\}, \tag{24}$$

$$\alpha = \min\{1, \gamma_0/(\ell + 2)\}, \tag{25}$$

where $h = |H|$ as usual. Apply Lemma 17 to the set of constants ℓ, α, γ_0, C and D to obtain the constants ε, $\bar{B}_0$, and m_0 whose existence is guaranteed by that lemma. Clearly, we may and shall assume that $\varepsilon \le \min\{1/2, 2c\}$. Finally, let $K_0 = K_0(\varepsilon, k_0, D)$ and $\eta = \eta(\varepsilon, k_0, D)$ be the constants from Lemma 10 and set

$$B = B_{\ell,H} = B = |\mathrm{Aut}(H)|K_0\bar{B}_0. \tag{26}$$

In the remaining of the proof, we shall show that this choice of $B = B_{\ell,H}$ will do. Recall that we consider $p = p(n) = Bn^{-\beta}$, where β is as given by (2). In the sequel we shall also use p_0 and (6).

As before, let $\widetilde{H} = \widetilde{H}(G)$ be the spanning subgraph of G with edge set $E(\widetilde{H}) = \bigcup_{1\le i\le T} E(H_i)$, where $H_1, \ldots, H_T$ are the isolated copies of H in G. Let us now define $\mathcal{G} = \mathcal{G}^n$ to be the family of graphs G of order n satisfying properties (a)–(c) below.

- (a) $\widetilde{H}(G)$ is (η, D, p_0)-upper-uniform.
- (b) If $\mathbf{W} = (W_i)_{i=1}^h$ is a vector of pairwise disjoint subsets of vertices of G with $|W_i| \ge n/\log n$ for all $1 \le i \le h$, then $Y_{\mathbf{W}} = Y_{\mathbf{W}}(G) \ge (1/2)|W_1|\ldots|W_h|p^{e(H)}$, where $Y_{\mathbf{W}}$ is as defined in Section 2.2.3.
- (c) For any vector $\mathbf{V} = (V_i)_{i=1}^\ell$ of pairwise disjoint m-subsets of vertices of G, where $n/2K_0 \le m \le n/k_0$, if a $\mathbf{V}$-graph F (cf. Section 2.3) with $E(F) \sqsubset E(\widetilde{H}(G))$ is such that all pairs (V_{i-1}, V_i) with $1 \le i \le \ell$ are (ε, F, p_0)-regular and satisfy $d_{F,p_0}(V_{i-1}, V_i) \ge \gamma_0$, then F contains a copy of C^ℓ.

The proof of Theorem 2 will be finished if we establish the following two claims.

Claim 2 *Almost every G_p belongs to $\mathcal{G}$.*

Claim 3 *Every member G of $\mathcal{G}$ satisfies $G \to (C^\ell, H)$ provided n is large enough.*

Proof of Claim 2. Lemma 14 implies that almost every $G = G_p$ satisfies condition (a) above, whereas Lemma 16 implies that almost every $G = G_p$ satisfies (b). To prove that (c) is almost surely true for $G = G_p$, we first want to argue that any F as in the statement of (c) is an element of $\mathcal{F}_m(\varepsilon, \bar{B}, \gamma_0, C, D; \mathbf{V}, \bar{T})$, where $\bar{T} = e(F)$ and $\bar{B} = \bar{B}_m > 0$ is such that $p_0 = \bar{B}m^{-1+1/(\ell-1)}$. Note that by the choices of our constants we have $\bar{B} = p_0 m^{1-1/(\ell-1)} \ge e(H)B^{e(H)}/(|\mathrm{Aut}(H)|K_0) \ge \bar{B}_0$ (cf. (26)). Conditions (i) and (ii) in the definition of $\mathcal{F}$ are clearly satisfied. Condition (iii) follows from the upper uniformity of $\widetilde{H}$ (cf. (a)) and from the assumption on the density of the pairs (V_{i-1}, V_i). Finally, condition (iv) in the definition of $\mathcal{F}$ follows from Lemma 15. Now if any such F does not contain a copy of C^ℓ then this implies that $\widetilde{H}$ contains an element of $\mathcal{F}'_m(\varepsilon, \bar{B}, \gamma_0, C, D; \mathbf{V}, \bar{T})$ for some choice of m, $\bar{T}$, and $\mathbf{V}$. The probability of this event is by Lemmas 13 and 17 at most

$$
\begin{aligned}
|\mathcal{F}'_m(\varepsilon,\bar{B},\gamma_0,C,D;\mathbf{V},\bar{T})|{p_0}^{\bar{T}} &\leq \alpha^{\bar{T}}\binom{(\ell+2)m^2}{\bar{T}}{p_0}^{\bar{T}} \\
&\leq \left(\alpha\frac{\mathrm{e}(\ell+2)m^2}{\gamma_0\ell p_0 m^2}p_0\right)^{\bar{T}} \quad \leq \quad \left(\frac{\mathrm{e}}{\ell}\right)^{\bar{T}}. \qquad (27)
\end{aligned}
$$

Summing (27) over all possible choices of m, $\bar{T}$, and $\mathbf{V}$ gives us an additional factor of at most $n^3 2^{\ell n}$. Using the fact that $\bar{T}$ is superlinear in n, we deduce by Markov's inequality that (c) holds for almost every G_p. This finishes the proof of Claim 2.

Proof of Claim 3. Let $G \in \mathcal{G}$ be fixed, and suppose G is given a blue-red edge-colouring that contains no red H. Our aim is to show that then there is a blue copy of C^ℓ in our graph G. Recall that G contains T isolated copies $H_1,\dots,H_T$ of H. Clearly, we may pick one blue edge $e_i \in E(H_i)$ for each $1 \leq i \leq T$. We shall prove that, however we choose these T blue edges e_i, we shall necessarily select the edges of a $C^\ell \subseteq G$.

We start by letting J be the spanning subgraph of G with edge set $E(J) = \{e_i\colon 1 \leq i \leq T\}$. From property (a) above, we know that $\widetilde{H}$ is (η, D, p_0)-upper-uniform, and hence so is J. We may thus apply Lemma 10 to J and obtain an (ε,k)-equitable (ε, J, p_0)-regular partition $\Pi = (V_i)_{i=0}^k$ of $V(J)$ with $k_0 \leq k \leq K_0$. Write m for the cardinality of the V_i $(1 \leq i \leq k)$.

We now consider a complete graph K^k on the vertex set $[k]$, and consider the 3-edge-colouring defined as follows: colour the edge ij with colour 3 if the pair (V_i,V_j) is not (ε, J, p_0)-regular; otherwise, colour the edge ij colour 1 if $\gamma_{ij} = d_{J,p_0}(V_i,V_j) \geq \gamma_0$, and colour 2 if $\gamma_{ij} < \gamma_0$. We then apply Lemma 12 for the triple (C^ℓ, H, K^2) to the 3-edge-coloured complete graph K^k. For later convenience, for all $i \neq j$ $(i,\ j \in [k])$, let us colour the edges of G that join the classes V_i and V_j with the colour used to colour the edge $ij \in E(K^k)$. Note that we now have two colours associated to each edge of G.

Suppose first that there is a monochromatic copy of C^ℓ of colour 1 in our K^k. Let $\mathbf{V} = (V_1,\dots,V_\ell)$ be the ℓ vertex classes corresponding to this copy of C^ℓ and let F be the $\mathbf{V}$-graph on $\bigcup_{i=1}^\ell V_i$ containing the J-edges between consecutive classes. By property (c) we have that F contains a C^ℓ, and we are home.

We now observe that our K^k must contain fewer than $\varepsilon\binom{k}{2} \leq 2c\binom{k}{2} < ck^2$ edges coloured with colour 3, since the partition Π is (ε, J, p_0)-regular. Therefore, we know by Lemma 12 that we may find at least ck^h monochromatic copies of H of colour 2 in our K^k. We shall see in the sequel that this leads us to a contradiction, and the proof of Claim 3 will be finished.

Observe first that $m \geq n/2k \geq n/\log n$ if n is large enough. In what follows, we may and shall assume that $m \geq n/\log n$. Therefore we may deduce from property (b) above that each H-subgraph of K^k of colour 2 determines at least $m^h p^{e(H)}/2$ isolated copies of H in G. Clearly, distinct copies of H in K^k determine distinct isolated copies of H in G. We may thus conclude that G contains at least

$$\frac{1}{2}ck^h m^h p^{e(H)} \geq \frac{1}{2}ck^h \left(\frac{n}{2k}\right)^h p^{e(H)} \geq c2^{-h-1}B^{e(H)}n^{1+1/(\ell-1)}$$

isolated copies of H, all edges of which are coloured with colour 2. Hence the number T' of edges of J that are coloured 2 satisfies

$$T' \geq c2^{-h-1}B^{e(H)}n^{1+1/(\ell-1)}. \tag{28}$$

However, we also have an upper bound for T', coming from the fact that this quantity counts only edges of G that are coloured 2. Let us write $\sum_{ij}^2$ for the sum over all edges ij of K^k that are coloured 2. Then, by the definition of the colouring of K^k, by (6), and using that $m \leq n/k$, we may conclude that

$$T' \leq \sum\nolimits_{ij}^2 d_{J,p_0}(V_i,V_j)p_0m^2 < \binom{k}{2}\gamma_0 p_0 m^2 \leq \gamma_0 e(H)B^{e(H)}n^{1+1/(\ell-1)},$$

which, together with the definition of γ_0 given in (24), contradicts (28).

This finishes the proof of Claim 3, and hence the proof of Theorem 2 is complete. □

2.5 Sketch of the proof of Theorem 4

In this section we describe the alterations in the proof of Theorem 2 that yield a proof of Theorem 4. We start by defining a new class $\mathcal{G}'$ of graphs to replace $\mathcal{G}$ from the proof of Theorem 2. In what follows, (a), (b), and (c) refer to conditions (a), (b), and (c) in the definition of the class $\mathcal{G}$.

Let us write (b') for the statement we obtain from (b) by replacing the constant $1/2$ in (b) by $1/3$. Moreover, to stress the length of the cycle that is involved in statement (c), let us write (c_ℓ) for that condition.

The n-vertex graphs that we put in the new class $\mathcal{G}'$ are the ones that satisfy (a), (b') and $(c_{\ell_2}) \wedge \cdots \wedge (c_{\ell_r})$. We now consider the following two claims.

Claim 4 *There is a constant $\kappa > 0$ such that almost every G_p has the property that, however we delete $\leq \kappa n^{1+1/(\ell_2-1)}/(\log n)^{|H|}$ edges from G_p, we obtain a subgraph G' of G_p that belongs to $\mathcal{G}'$.*

Claim 5 *Every member G' of $\mathcal{G}'$ satisfies $G' \to (H, C^{\ell_2}, \ldots, C^{\ell_r})$ provided n is large enough.*

Proof of Claim 4. (Sketch) Condition (a) holds for any subgraph of G_p almost surely. Condition (b') may be made to hold almost surely by choosing $\kappa > 0$ sufficiently small. We now consider condition $(c_{\ell_2}) \wedge \cdots \wedge (c_{\ell_r})$. Let $q \geq \ell_2$ be be a fixed integer. We argue that (c_q) holds for every subgraph G' of G_p as in the statement of our claim almost surely. Let $p' = p'(n) = B_q n^{-\beta_q}$, where $B_q = B_{q,H}$ is as given in (26) and $\beta_q = \beta(C^q, H)$. We already know that (c_q) holds for almost every $G_{p'}$. Notice that, since we are now talking about a random graph of different edge-probability, the value of p_0 in (c_q) is different from the value of p_0 in (c_ℓ).

Suppose for a contradiction that (c_q) does not hold almost surely for every subgraph G' of G_p as specified. We may generate the random graph $G_{p'}$ by first generating G_p, and then randomly deleting edges from G_p with probability $1 - p'/p$. It is then easy to check that our hypothesis on (c_q) failing for G_p implies that (c_q) does not hold almost surely for $G_{p'}$, which is a contradiction. The interested reader may find the missing details in the proof of a closely related statement proved in [9] (see Lemma 19 in [9]). $\square$

Proof of Claim 5. (Sketch) Following the proof of Claim 3, we define graphs J_i $(2 \leq i \leq r)$ in a similar way: each isolated copy of H in G_p contains an edge of colour i for some $2 \leq i \leq r$. Use these edges to define the graphs J_i. Now apply Lemma 10 simultaneously to all the J_i to obtain a partition $(V_j)_{j=0}^k$ of $V(G_p)$ that is ε-regular with respect to all the J_i $(2 \leq i \leq r)$.

We now encode the regularity/non-regularity and density of the pairs (V_a, V_b) with respect to J_i in an edge-coloured complete graph K^k. If a pair (V_a, V_b) is not ε-regular with respect to some J_i, we colour the edge ab of K^k colour $r+1$ right away. Otherwise, if (V_a, V_b) is a 'dense' pair with respect to J_i, for some $2 \leq i \leq r$, we colour $ab \in E(K^k)$ colour i. The remaining edges we colour with colour 1. We then apply Lemma 12 to the $(r+1)$-tuple of graphs $(H, C^{\ell_2}, \ldots, C^{\ell_r}, K^2)$.

As in the proof of Theorem 2, if we find in our edge-coloured K^k a monochromatic cycle C^{ℓ_i} of colour i, for some $2 \leq i \leq r$, then we are done by invoking statement (c_{ℓ_i}). The possibility that we shall find many edges of K^k coloured with colour $r+1$ is ruled by the ε-regularity of our partition $(V_j)_{j=0}^k$ (here we need that ε is small enough). Finally, the possibility that we shall find many monochromatic copies of H of colour 1 may be ruled out as in the proof of Claim 3 (here we need that γ_0 should be small enough). We omit the details. $\square$

Theorem 4 is immediate from Claims 4 and 5.

3 Proof of the 0-statement

Proof of Theorem 3. Since the case $\ell = k$ is contained in Theorem 1, we assume in the sequel that $\ell > k$. Similarly to [10] and [13], given a graph G, we define a hypergraph $\mathcal{G} = \mathcal{G}(G)$ whose hyperedges correspond to the ℓ- and k-cycles of G. Let

$$\begin{aligned}\mathcal{E}_1 = \mathcal{E}_1(G) &= \{E(C)\colon C \text{ is an } \ell\text{-cycle in } G\},\\ \mathcal{E}_2 = \mathcal{E}_2(G) &= \{E(C)\colon C \text{ is a } k\text{-cycle in } G\}.\end{aligned}$$

We put $\mathcal{G} = \mathcal{G}(G) = \mathcal{E}_1 \cup \mathcal{E}_2$. We refer to the hyperedges in $\mathcal{E}_i$ as hyperedges of *type* i ($i \in \{1,2\}$).

If $\mathcal{H}$ is a subhypergraph of $\mathcal{G} = \mathcal{G}(G)$, that is, if $\mathcal{H} \subseteq \mathcal{G}$, then we define the *underlying graph* $G(\mathcal{H})$ of $\mathcal{H}$ to be the subgraph of G spanned by $\bigcup \mathcal{H}$. In other words, $G(\mathcal{H}) = G[\bigcup \mathcal{H}]$. Moreover, if $E \in \mathcal{G}$, then we write $V(E)$ for $V(G(E))$, the set of vertices of G incident to the edges in E. We now notice that $G \to (C^\ell, C^k)$ simply means that

(*) *for every colouring of the hypervertices of $\mathcal{G}$ with two colours, blue and red, say, there is either a blue hyperedge of type 1 or a red hyperedge of type 2.*

A subhypergraph $\mathcal{G}_0 \subseteq \mathcal{G}$ of $\mathcal{G}$ is said to be *critical* with respect to property (*) if $\mathcal{G}_0$ fulfils (*) but if we remove any hyperedge of any type from $\mathcal{G}_0$ then the resulting hypergraph fails to have (*). The following very simple but useful claim is similar to Claim 1 in the proof of Theorem 2(a) in [10].

Claim 6 *If $\mathcal{G}_0$ is critical with respect to (*) then for any hyperedge E of type i ($i \in \{1,2\}$) and any hypervertex $e \in E$ there exists a hyperedge F of type $3-i$ such that $E \cap F = \{e\}$.*

We now need to define some constants. Set

$$\alpha = k - 2 - \beta(k-1) = \frac{k-\ell}{k(\ell-1)}. \tag{29}$$

Observe that

$$-1 < \alpha < 0. \tag{30}$$

Furthermore, let

$$\delta = \min\left\{-\alpha, \frac{1}{\ell-1}, 1-\beta\right\}, \tag{31}$$

and notice that then $0 < \delta < 1+\alpha$. Indeed, the first inequality follows from the definition of β and (30), and the second follows from $1-\beta < 1+\alpha$.

Now assume that G is a graph and that $\mathcal{G}_0 \subseteq \mathcal{G}(G)$ is critical with respect to (*). As in [10], we will use Claim 6 in order to look for a subhypergraph $\mathcal{H}$ of $\mathcal{G}_0$ whose underlying graph $G(\mathcal{H})$ is dense.

We shall give a procedure, Procedure HYPERTREE, that will build a subhypergraph of $\mathcal{G}_0$ that may contain some hypercycles, but will generally have a 'hypertree structure'. Any hypergraph that may be produced by this procedure will be said to be an *alternating hypertree*. The key observation concerning hypertrees is that *the number of those alternating hypertrees whose underlying graph is not dense is small.*

Let us now describe Procedure HYPERTREE.

Procedure HYPERTREE
Input: *a hypergraph* $\mathcal{G}_0 = \mathcal{E}_1 \cup \mathcal{E}_2$ *critical with respect to* (*)
Output: *a triple* $(m, \mathcal{H}_m, I_m)$ *where* $\mathcal{H}_m = \{E_0, \ldots, E_m\} \subseteq \mathcal{G}_0$
and $I_m \subseteq [m] \times \{1, 2\}$
$\langle$*main body*$\rangle$

Thus, the input to Procedure HYPERTREE is a subhypergraph $\mathcal{G}_0$ of $\mathcal{G}(G)$, for some graph G, with $\mathcal{G}_0$ critical with respect to (*). Its output is a triple $(m, \mathcal{H}_m, I_m)$, where m is an integer, $\mathcal{H}_m = \{E_0, \ldots, E_m\}$ is a subhypergraph of $\mathcal{G}_0$, and $I_m \subseteq [m] \times \{1, 2\}$ is a certain set indicating the *degeneracies* present in $\mathcal{H}_m$. Loosely speaking, the set I_m says how far $\mathcal{H}_m$ is from being a 'genuine' hypertree.

Below we shall show that (†) *either the set I_m is large and hence the underlying graph of $\mathcal{H}_m$ is dense, or else I_m is small and the number of alternating hypertrees with such a small set of degeneracies is small.* As an illustration, we mention that if $\mathcal{H}_m$ has no degeneracies at all, then $\mathcal{H}_m$ is uniquely determined up to isomorphism, as is the underlying graph $G(\mathcal{H}_m)$ of $\mathcal{H}_m$, regardless of the input hypergraph $\mathcal{G}_0$. Assertion (†) above will imply that G_p, with $p = p(n) = bn^{-\beta}$ and sufficiently small b, almost surely does not contain a copy of the underlying graph of any alternating hypertree. This assertion will establish Theorem 3.

The main body of Procedure HYPERTREE is given in Figure 1.

$\langle$*main body*$\rangle \equiv$ **begin** $i := 0$; *pick* $E_0 \in \mathcal{E}_2$ *arbitrarily*; $\mathcal{H}_0 := \{E_0\}$;
(1) *rank and direct the edges in* E_0;
$I_0 := \emptyset$; $A_0 := \emptyset$;
repeat $\langle$*extend current hypertree*$\rangle$
until $|G(\mathcal{H}_i)| - \beta e(G(\mathcal{H}_i)) \leq -1/500$ *or* $i \geq \log n$
end

Figure 1: Main body of Procedure HYPERTREE

Thus, Procedure HYPERTREE starts by setting up a hypergraph $\mathcal{H}_0 = \{E_0\}$ containing a single hyperedge that corresponds to a k-cycle of G, and then goes into an iterative process. Before we go on to detail this iterative procedure, we need to explain what is meant in line (1) in $\langle$*main body*$\rangle$. The hypervertices of the hypertree will be ranked with ranks starting from 1. Every time we add a new hypervertex to our hypertree, it gets a rank that is by one larger than that of its predecessor. As the hyperedges correspond to cycles in G, we can first direct the edges of a new hyperedge in a cyclic order and then rank the edges following the same order. For simplicity, we shall not make this directing and ranking procedure explict.

We now turn to the iterative procedure. This process will create sequences $\mathcal{H}_0 \subseteq \mathcal{H}_1 \subseteq \cdots$ and $I_0 \subseteq I_1 \subseteq \cdots$ that will eventually produce the desired output $(m, \mathcal{H}_m, I_m)$. For all $0 \le i \le m$, we shall have $\mathcal{H}_i = \{E_0, \dots, E_i\}$ and $I_i \subseteq [i] \times \{1, 2\}$. In the ith iteration we add the hyperedge E_i to $\mathcal{H}_{i-1}$ and we *may* add one of the pairs $(i, 1)$ or $(i, 2)$ to I_{i-1}. The latter addition to the set of degeneracies happens if E_i intersects $\mathcal{H}_{i-1}$ 'substantially'. The procedure will also need to construct a sequence $A_0, A_1, \dots$ of sets of edges, with $A_i \subseteq \bigcup \mathcal{H}_i$ for all $0 \le i \le m$. The edges in A_i are said to be the *assets* of $\mathcal{H}_i$.

The general iterative step is as follows.

$\langle$*extend current hypertree*$\rangle \equiv$
  if *there is* $F \in \mathcal{E}_2$ *such that* $F \not\subseteq \bigcup \mathcal{H}_i$ *and* $|V(F) \cap V(G(\mathcal{H}_i))| \ge 2$
    then $\langle$*add* F *to* $\mathcal{H}_i\rangle$
    else $\langle$*look for and add a type 1 hyperedge to* $\mathcal{H}_i\rangle$

When we try to extend our current hypertree $\mathcal{H}_i$, if it so happens that there is a type 2 hyperedge F 'for free' (that is, as specified in the **if** condition of $\langle$*extend current hypertree*$\rangle$), then we take it and add it to $\mathcal{H}_i$. The idea here is that this addition will increase the density of the underlying graph of the current hypertree.

$\langle$*add* F *to* $\mathcal{H}_i\rangle \equiv i := i + 1$; $E_i := F$; $\mathcal{H}_i := \mathcal{H}_{i-1} \cup \{E_i\}$;
  rank and direct the edges in $E_i \setminus \bigcup \mathcal{H}_{i-1}$;
  $I_i := I_{i-1} \cup \{(i, 1)\}$; $A_i := \emptyset$

If such a 'free' type 2 hyperedge F does not exist, then we have to do some more work, which we describe in two stages:

$\langle$*look for and add a type 1 hyperedge to* $\mathcal{H}_i\rangle \equiv \langle$*stage I*$\rangle$; $\langle$*stage II*$\rangle$

In $\langle$*stage I*$\rangle$ we look for and add a type 1 hyperedge to our current hypertree $\mathcal{H}_i$. This will, unfortunately, decrease the density of the current underlying graph. Therefore, in $\langle$*stage II*$\rangle$, we try to add as many type 2 hyperedges to our hypertree as possible; these hyperedges all 'hang' on hypervertices that were added to the hypertree by the addition of the type 1 hyperedge. These hypervertices are the 'assets' that we acquired with the addition of the type 1 hyperedge. One may think of the process of adding type 2 hyperedges in $\langle$*stage II*$\rangle$ as a process of turning these assets into 'cash' (*i.e.*, density). See Figures 2 and 3.

$\langle$*stage I*$\rangle \equiv$ *let* f *be the edge of smallest rank in* $\bigcup \mathcal{H}_i$
  for which there is a hyperedge $E \in \mathcal{E}_1$ *with* $E \not\subseteq \bigcup \mathcal{H}_i$;
  {NB. The existence of the edge f above is proved in Claim 10.}
 $i := i + 1$;
 pick $E_i \in \mathcal{E}_1$ *with* $E_i \not\subseteq \bigcup \mathcal{H}_i$ *and* $f \in E_i$;
 $\mathcal{H}_i := \mathcal{H}_{i-1} \cup \{E_i\}$;
 rank the edges in $E_i \setminus \bigcup \mathcal{H}_{i-1}$ *following the cyclic order of* $E_i = C^\ell$
  given by the orientation of f, *starting from the successor of* f;
  suppose these edges are, in increasing order of rank, $g_1, \ldots, g_a$;
 if $|V(E_i) \cap V(G(\mathcal{H}_{i-1}))| = 2$ **then** $I_i := I_{i-1}$ **else** $I_i := I_{i-1} \cup \{(i, 2)\}$;
(1) $A_i := \{g_1, \ldots, g_a\}$; {NB. $|A_{i-1}| = 0$ and $a = |A_i| = \ell - 1$ if $(i, 2) \notin I_i$}

Figure 2: Adding a type 1 hyperedge

$\langle$*stage II*$\rangle \equiv F := E_i$; {NB. This is the ℓ-cycle we have just added}
 for $j := 1$ **to** a **do**
  if $g_j \in A_i$ **then begin**
   $i := i + 1$; *pick* $E_i \in \mathcal{E}_2$ *with* $E_i \cap F = \{g_j\}$;
   $\mathcal{H}_i := \mathcal{H}_{i-1} \cup \{E_i\}$;
   rank and direct the edges in $E_i \setminus \bigcup \mathcal{H}_{i-1}$ *following*
    the cyclic order of $E_i = C^k$ *given by the orientation of* g_j,
    starting from the successor of g_j;
   if $|V(E_i) \cap V(G(\mathcal{H}_{i-1}))| = 2$
    then begin $I_i := I_{i-1}$; $A_i := A_{i-1} \setminus \{g_j\}$ **end**
    else begin $I_i := I_{i-1} \cup \{(i, 2)\}$; $\langle$*compute* $A_i\rangle$ **end**
   end

Figure 3: Adding as many type 2 hyperedges as possible

In $\langle$*stage II*$\rangle$, the computation of A_i requires a little care. Roughly speaking, if a k-cycle E_i we add in $\langle$*stage II*$\rangle$ contains a vertex w added to the underlying graph of the current hypertree by the addition of the ℓ-cycle in $\langle$*stage I*$\rangle$, then we have to consider an asset $g \in A_{i-1}$ incident to w no longer an asset. In this case we say that g was *spoiled* by E_i. See Figure 4.

The integer $m = |\mathcal{H}_m| - 1 \leq \log n$ returned by Procedure HYPERTREE is the *length*

$\langle \textit{compute } A_i \rangle \equiv A'_i := A_{i-1} \setminus \{g_j\};$
for $p := j+1$ **to** a **do**
if *the head* w_p *of* g_p *belongs to* $V(E_i)$ **then** $A'_i := A'_i \setminus \{g_p\};$
$\{$NB. In this case we say that E_i *spoiled* $g_p\}$
$A_i := A'_i$

Figure 4: Computing A_i

of $\mathcal{H}_m$. Our alternating hypertree $\mathcal{H}_m$ is said to be *degenerate* at level i if $(i,1)$ or $(i,2) \in I_m$.

Let $f(i) = |G(\mathcal{H}_i)| - \beta e(G(\mathcal{H}_i)) = |G(\mathcal{H}_i)| - \beta|\bigcup \mathcal{H}_i|$. Then, for $p = p(n) = n^{-\beta}$, the expected number of copies of $G(\mathcal{H}_i)$ in G_p is $O(n^{f(i)})$. In the series of claims given below, we investigate the effect of the degeneracies on the behaviour of the function f.

For brevity, let $a_i = |A_i|$ $(0 \leq i \leq m)$. In the sequel, it will be important that $\mathcal{H}_i$ may have a non-zero number of assets only from the point in which we have just added a type 1 hyperedge up to the point in which we have cashed all available assets in the **for** loop in $\langle \textit{stage II} \rangle$. In other words, $a_i > 0$ can only occur between line (1) in $\langle \textit{stage I} \rangle$ and the end of $\langle \textit{stage II} \rangle$. In particular, $a_0 = a_m = 0$.

Claim 7 *For all* $1 \leq i \leq m$*, if* $(i,1)$*,* $(i,2) \notin I_i$ *then* $f(i) + \alpha a_i = f(i-1) + \alpha a_{i-1}$*.*

Proof. Consider first the case in which E_i is of type 1. Then $a_{i-1} = 0$ and moreover $a_i = \ell - 1$, since $(i,2) \notin I_i$ (cf. the remark in line (1) of $\langle \textit{stage I} \rangle$). Therefore E_i contributes exactly $\ell - 2$ new vertices and exaclty $\ell - 1$ new edges to $\mathcal{H}_{i-1}$. Hence,

$$\begin{aligned} f(i) + \alpha a_i &= f(i-1) + \ell - 2 - \beta(\ell - 1) + (k - 2 - \beta(k-1))(\ell - 1) \\ &= f(i-1) + (k-1)(\ell-1) - 1 - \beta k(\ell - 1) \\ &= f(i-1) + \alpha a_{i-1}. \end{aligned}$$

If E_i is of type 2 then $a_i = a_{i-1} - 1$ and

$$\begin{aligned} f(i) + \alpha a_i &= f(i-1) + (k-2) - \beta(k-1) + \alpha(a_{i-1} - 1) \\ &= f(i-1) + \alpha a_{i-1}, \end{aligned}$$

as required.

Claim 8 *For all* $1 \leq i \leq m$*, if* $(i,1) \in I_i$ *then* $f(i) + \alpha a_i \leq f(i-1) + \alpha a_{i-1} - \delta$*.*

Proof. Let $H \subseteq C^k$ be the graph with vertex set $V(E_i) \cap V(G(\mathcal{H}_{i-1}))$ and edge set $E_i \cap \bigcup \mathcal{H}_{i-1}$, that is, H is the *intersection* of $(V(E_i), E_i)$ and $G(\mathcal{H}_{i-1})$. Clearly, H is a disjoint union of paths. Moreover, $G(\mathcal{H}_i)$ has $k - |H|$ new vertices and $k - e(H)$ new edges. Therefore, by $|H| \geq e(H) + 1$, $|H| \geq 2$, and $\beta \leq 1$, and by the definitions of α and δ (cf. (29) and (31)), we have

$$\begin{aligned} f(i) &= f(i-1) + k - |H| - \beta(k - e(H)) \\ &\leq f(i-1) + k - 2 - \beta(k-1) \\ &\leq f(i-1) - \delta, \end{aligned}$$

which implies the claim, since $a_{i-1} = a_i = 0$.

Claim 9 *For all* $1 \leq i \leq m$*, if* $(i, 2) \in I_i$ *then* $f(i) + \alpha a_i \leq f(i-1) + \alpha a_{i-1} - \delta$.

Proof. Consider first the case in which E_i is of type 1. Let $H \subseteq C^\ell$ be the intersection of the two graphs $(V(E_i), E_i)$ and $G(\mathcal{H}_{i-1})$. Then $a_i = \ell - e(H)$ and

$$\begin{aligned} f(i) + \alpha a_i &= f(i-1) + \ell - |H| - \beta(\ell - e(H)) + \alpha(\ell - e(H)) \\ &= f(i-1) + (\ell - e(H))(k - 1 - k\beta) + e(H) - |H| \\ &= f(i-1) + \frac{\ell - e(H)}{\ell - 1} + e(H) - |H|. \end{aligned} \tag{32}$$

If $e(H) \geq 2$ then we conclude from $|H| \geq e(H) + 1$ that $(\ell - e(H))/(\ell - 1) + e(H) - |H| \leq -1/(\ell - 1) \leq -\delta$ (cf. (31)). If $e(H) = 1$ then $|H| \geq 3$ and thus $(\ell - e(H))/(\ell - 1) + e(H) - |H| \leq -\delta$. Because of $a_{i-1} = 0$ we have that (32) is in both cases at most $f(i-1) + \alpha a_{i-1} - \delta$.

Let now E_i be of type 2 and let $H \subseteq C^k$ be the intersection of the two graphs $(V(E_i), E_i)$ and $G(\mathcal{H}_{i-1})$. We first deduce from the fact that g_j is an asset before we add E_i, that is, $g_j \in A_{i-1}$, that $H \neq C^k$. Indeed, let $E_{i'}$ be the hyperedge of type 1 that E_i is attached to; *i.e.*, $E_i \cap E_{i'} = \{g_j\}$. Because of the **if** condition in $\langle$*extend current hypertree*$\rangle$, the hyperedge E_i is disjoint from $\bigcup \mathcal{H}_{i'-1}$. Let w be the head of g_j. Then if $E_i \subseteq \bigcup \mathcal{H}_{i-1}$, the vertex w must be incident to some edge in $\bigcup \mathcal{H}_{i-1} \setminus \bigcup \mathcal{H}_{i'}$. This means that g_j was spoiled before, *i.e.*, $g_j \notin A_{i-1}$, which is a contradiction. Hence $H \neq C^k$.

Let b be such that $a_i = a_{i-1} - b$. Observe that $b \geq 1$. If E_i spoils an edge g_h then the head w_h of g_h is isolated in H. Indeed, E_i contains no edge from $\bigcup \mathcal{H}_{i'-1}$, no edge from $E_{i'}$ incident to w_h, and, since g_h has not been spoiled before, w_h is incident to no edge in $\bigcup \mathcal{H}_{i-1} \setminus \bigcup \mathcal{H}_{i'}$. Therefore, there does not exist any edge in $E_i \cap \bigcup \mathcal{H}_{i-1}$ incident to w_h. Hence w_h is indeed isolated in H. Setting $J := H \setminus \{w_f \colon g_f \text{ is spoiled by } E_i\}$ we get that $|H| = |J| + b - 1$. Moreover,

$$
\begin{aligned}
&f(i) + \alpha a_i \\
&\quad = \; f(i-1) + k - |J| - (b-1) - \beta(k - e(J)) + \alpha(a_{i-1} - b) \\
&\quad = \; f(i-1) + \alpha a_{i-1} - (b-1)(\alpha+1) + k - |J| - \beta(k - e(J)) - \alpha \\
&\quad \leq \; f(i-1) + \alpha a_{i-1} - (b-1)(\alpha+1) + (e(J) - |J| + 1) + (e(J) - 1)(\beta - 1).
\end{aligned}
$$

The three last terms are non-positive. If $b \geq 2$ then the first term is smaller than $-\delta$ because of $\delta < 1 + \alpha$. If $b = 1$ then J is a proper subgraph of $E_i = C^k$ with $J \neq K^2$ and therefore either $e(J) > 1$ or $|J| > e(J) + 1 = 2$ implying that one of the two last terms is smaller or equal than $-\delta$, and the claim is proven.

Claim 10 *The edge f as specified in $\langle$stage I$\rangle$ always exists.*

Proof. Assume that Procedure HYPERTREE gets stuck at the beginning of $\langle$*stage I*$\rangle$ at iteration i because the edge f as specified does not exist. Our aim is to reach a contradiction.

Let $G_i = G(\mathcal{H}_i)$ be the underlying graph of $\mathcal{H}_i$, so that G_i is spanned by its edge set $E(G_i) = \bigcup \mathcal{H}_i$. We start with the following assertion.

(**) *For both $q \in \{1, 2\}$, if $e \in E \subseteq E(G_i)$ and $E \in \mathcal{E}_q$, then there exists a hyperedge $F \in \mathcal{E}_{3-q}$ such that $e \in F \subseteq E(G_i)$ and moreover $E \cap F = \{e\}$.*

Let us prove assertion (**). Suppose first that $q = 1$. Then Claim 6 gives a hyperedge $F \in \mathcal{E}_2$ with $E \cap F = \{e\}$. The **if** condition in $\langle$*extend current hypertree*$\rangle$ implies that $F \subseteq \bigcup \mathcal{H}_i = E(G_i)$, as required. If $q = 2$, then Claim 6 gives $F \in \mathcal{E}_1$ with $E \cap F = \{e\}$. Since we did not find an edge f as specified on the first line of $\langle$*stage I*$\rangle$, we must have $F \subseteq \bigcup \mathcal{H}_i = E(G_i)$. Assertion (**) is proved.

An immediate corollary to (**) is the following assertion.

(‡) *Every edge e of G_i is contained in an ℓ-cycle C_1 and in a k-cycle C_2 with $E(C_q) \in \mathcal{E}_q$ ($q \in \{1, 2\}$) and $E(C_1) \cap E(C_2) = \{e\}$.*

In particular, we have the following observation.

(i) *The minimal degree of G_i is at least 3.*

Let us now consider the simpler case in which $k \geq 4$. Because of $\ell > k$, it is easily seen that $\beta(C^\ell, C^k) \geq \beta(C^5, C^4) = 11/16$. On the other hand, by (i), we have $e(G_i) \geq (3/2)|G_i|$ and therefore in this case we have

$$
f(i) \; = \; |G_i| - \beta e(G_i) \; \leq \; |G_i| \left(1 - \frac{3}{2} \times \frac{11}{16}\right) \; \leq \; -\frac{1}{500}.
$$

We have thus reached a contradiction, since this implies that Procedure HYPERTREE should have stopped with $m = i$.

We now turn to the case $k = 3$. We shall have to argue more carefully here. We introduce some notation in order to avoid repetition. Suppose x_1, x_2, and x_3 are three vertices of G_i. We shall write $T(x_1, x_2, x_3)$ when these vertices induce a triangle in G_i and moreover $\{x_1x_2, x_2x_3, x_1x_3\} \in \mathcal{E}_2$. Similarly, if $x_1, \ldots, x_\ell$ define, in this order, an ℓ-cycle C^ℓ of G_i and furthermore $E(C^\ell) \in \mathcal{E}_1$, then we shall write $C(x_1, \ldots, x_\ell)$.

Let $A(x_1x_2, T(x_1, x_2, x_3))$ stand for the application of (**) to the edge x_1x_2 and the type 2 hyperedge corresponding to $\{x_1, x_2, x_3\}$. Thus $A(x_1x_2, T(x_1, x_2, x_3))$ gives us the existence of vertices $y_3, \ldots, y_\ell \in V(G_i) \setminus \{x_1, x_2\}$ with $y_3 \neq x_3 \neq y_\ell$ and such that $C(x_1, x_2, y_3, \ldots, y_\ell)$. Similarly, let $A(x_1x_2, C(x_1, \ldots, x_\ell))$ stand for the corresponding statement guaranteeing the existence of a vertex $y_3 \in V(G_i) \setminus \{x_1, x_2, x_3, x_\ell\}$ with $T(x_1, x_2, y_3)$. Now we prove in a couple of steps that G_i has large density.

(*ii*) *A vertex of degree* 3 *in* G_i *is only adjacent to vertices of degree at least* 4.

To prove the assertion above, assume that there are two vertices of degree 3, say, x and y, that are adjacent in G_i. By (‡), we have a vertex z such that $T(x, y, z)$. We now argue as follows.

(*a*) $A(xy, T(x, y, z))$ gives 2 distinct vertices v and $w \notin \{x, y, z\}$ with $C(v, x, y, w, \ldots)$.
(*b*) $A(yz, T(x, y, z))$ gives $C(w, y, z, \ldots)$, since $d(y) = 3$.
(*c*) $A(wy, C(w, y, z, \ldots))$ gives $T(w, y, x)$, since $d(x) = d(y) = 3$.

However, $T(w, y, x)$ implies that $v = w$, which is a contradiction. Assertion (*ii*) is proved.

(*iii*) *Every vertex of* G_i *has a neighbour of degree at least* 4.

Indeed, let x be a vertex of G_i, and let y be a neighbour of x. Then, by (‡), there exists a vertex z with $T(x, y, z)$. By (*ii*), either y or z has degree at least 4, and (*iii*) is proven.

Consider first the case $\ell \geq 5$. Denote by a and b the number of vertices of degree exactly 3 and at least 4 in G_i, respectively. Assertions (*ii*) and (*iii*) imply that $\sum_{d(v)\geq 4} d(v) \geq 3a + b$. Therefore

$$\frac{2e(G_i)}{|G_i|} \geq \max\left\{\frac{3a+4b}{a+b}, \frac{6a+b}{a+b}\right\} \geq \min_{x\geq 0} \max\{4 - x, 1 + 5x\} = \frac{7}{2}.$$

Using $\beta(C^\ell, C^3) \geq \beta(C^5, C^3) = 7/12$ we have that

$$f(i) \;=\; |G_i| - \beta e(G_i) \;\leq\; |G_i|\left(1 - \frac{7}{12} \times \frac{7}{4}\right) \;\leq\; -\frac{1}{500}.$$

In the sequel we assume that $\ell = 4$.

(*iv*) *Every vertex of degree* 3 *in* G_i *has a neighbour of degree at least* 5.

Assume to the contrary that there is a vertex $v \in V(G_i)$ of degree 3 with neighbours x, y, and z all having degree 4. Recalling (‡), we may assume without loss of generality that $T(v,x,y)$ and $T(v,y,z)$. Now we argue as follows.

(*a*) Since $d(v)=3$, we have that $A(vy, T(v,y,z))$ gives a vertex $w \in V(G_i) \setminus \{v,x,y,z\}$ such that $C(w,x,v,y)$.

(*b*) $A(vy, T(v,x,y))$ similarly gives a vertex $w' \in V(G_i) \setminus \{v,x,y,z\}$ such that $C(w',z,v,y)$. However, $d(y)=4$, and hence $w=w'$.

(*c*) Since $d(v)=3$, we have that $A(vz, C(w,y,v,z))$ gives that $T(v,z,x)$.

(*d*) $A(yz, T(v,y,z))$ gives either $C(y,z,x,w)$ or else $C(y,z,w,x)$.

We reach a contradiction analysing the two possibilities in (*d*) separately. If $C(y,z,x,w)$, consider $A(xw, C(y,z,x,w))$ to get $T(x,w,v)$. But then $vw \in E(G_i)$, contradicting $d(v)$=3. On the other hand, if $C(y,z,w,x)$, consider $A(xw, C(x,y,z,w))$ to deduce $T(x,w,v)$ again. Therefore assertion (*iv*) is proved.

We now estimate the density of G_i from below. Let a and b denote the number of vertices of degree 3 and 4 in G_i, respectively, and let c be the number of vertices of degree at least 5. By (*ii*) and (*iii*) we have that $\sum_{d(v)\geq 4} d(v) \geq 3a+b+c$. Assertions (*iii*) and (*iv*) imply that $\sum_{d(v)\geq 5} d(v) \geq a+c$. Therefore we have

$$\frac{2e(G_i)}{|G_i|} \quad \geq \quad \max\left\{\frac{3a+4b+5c}{a+b+c}, \frac{6a+b+c}{a+b+c}, \frac{4a+4b+c}{a+b+c}\right\}.$$

Setting $x=a/(a+b+c)$ and $y=c/(a+b+c)$, this implies

$$\frac{2e(G_i)}{|G_i|} \quad \geq \quad \min_{x,y\geq 0} \max\{4-x+y, 1+5x, 4-3y\} \quad = \quad \frac{83}{23},$$

as one may easily check. Furthermore $\beta(C^4, C^3)=5/9$. Thus

$$f(i) \;=\; |G_i| - \beta e(G_i) \;\leq\; |G_i|\left(1-\frac{83}{46}\times\frac{5}{9}\right) \;\leq\; -\frac{1}{500},$$

and Claim 10 is proven.

Claim 11 *Let* $(m, \mathcal{H}_m, I_m)$ *be the output of Procedure* HYPERTREE. *Then the following assertions hold:*

(*a*) $f(m) \leq f(0) = 1+1/(\ell-1) \leq 3/2$, $m \leq \log n + \ell$, *and*

$$|I_m| \quad \leq \quad M \quad := \quad \ell + \left\lceil \frac{f(0)+1/500}{\delta} \right\rceil - 1. \tag{33}$$

(b) *Either* (i) $m \geq \log n$ *or else* (ii) $f(m) \leq -1/500$.

Proof. (a) The inequality for $f(m)$ follows easily from Claims 7, 8, and 9, and the fact that $a_0 = a_m = 0$. Suppose now that $m \geq \log n$. Then, letting i be the largest integer strictly smaller than $\log n$, we see that $m \leq i + \ell < \log n + \ell$, as required. To see (33), suppose for a contradiction that $|I_m| \geq M+1$. This time let i be the maximal integer not greater than m for which $a_i = 0$. Then $m \leq i + \ell$, and hence $|I_i| = |I_m \cap ([i] \times \{1,2\})| \geq M+1-\ell \geq \lceil (f(0)+1/500)/\delta \rceil$. Observing that $a_0 = a_m = 0$, we have, applying Claims 7, 8, and 9, that

$$f(i) \quad \leq \quad f(0) - \delta \left\lceil \frac{f(0)+1/500}{\delta} \right\rceil \quad \leq \quad -\frac{1}{500}.$$

However, this shows that Procedure Hypertree should have stopped after the ith iteration (see Figure 1), which is a contradiction.

Assertion (b) follows trivially from the **until** condition in $\langle$*main body*$\rangle$. This finishes the proof of Claim 11.

The heart of the proof of Theorem 3 is Claim 13 below. Loosely speaking, the idea is as follows. If $G_p \to (C^\ell, C^k)$, then $\mathcal{G}(G_p)$ contains a subhypergraph $\mathcal{G}_0$ critical with respect to (*). We may then feed this hypergraph $\mathcal{G}_0$ into Procedure Hypertree, and obtain a hypertree $\mathcal{H}_m \subseteq \mathcal{G}_0$ satisfying the conditions given in Claim 11. Using these conditions, we show in Claim 13 that a.e. G_p does not contain the underlying graph $G(\mathcal{H}_m)$ of $\mathcal{H}_m$ if $p = p(n) = bn^{-\beta}$ and b is small enough. However, this is a contradiction, since clearly $G(\mathcal{H}_m) \subseteq G_p$. We shall have thus proved Theorem 3.

To detail the argument sketched above, we need to introduce some definitions. Fix a graph G on n vertices such that $G \to (C^\ell, C^k)$. Consider all possible choices for $\mathcal{G}_0 \subseteq \mathcal{G}(G)$, the (*)-critical subhypergraph of $\mathcal{G}(G)$. Recall that Procedure Hypertree is a non-deterministic procedure, and, for each $\mathcal{G}_0$ as above, consider all sequences $\mathcal{H}_0 \subseteq \mathcal{H}_1 \subseteq \cdots$ that may be generated by Procedure Hypertree. Now, for each $i \geq 0$, let $\mathcal{F}(i)$ be the set of all hypergraphs $\mathcal{H}_i$ that may be obtained in the process described above, letting the graph G vary over all graphs G' with vertex set $[n]$ and satisfying $G' \to (C^\ell, C^k)$.

For each $i \geq 0$, let $\mathbf{G}(i)$ be the family of all graphs, *up to isomorphism*, that are underlying graphs of the hypergraphs $\mathcal{H}_i$ in $\mathcal{F}(i)$. Now recall that each $\mathcal{H}_i \in \mathcal{F}(i)$ has an associated set of degeneracies $I_i = I_i(\mathcal{H}_i) \subseteq [i] \times \{1,2\}$. For all $0 \leq j \leq i$, let $\mathbf{G}(i,j)$ be the collection of all graphs $G \in \mathbf{G}(i)$ that are underlying graphs of hypergraphs $\mathcal{H}_i \in \mathcal{F}(i)$ with j degeneracies, that is, with $|I_i(\mathcal{H}_i)| = j$. Finally, let $\mathbf{G}_{\leq}(m,d) = \bigcup \mathbf{G}(i,j)$, where the union is taken over all $0 \leq i \leq m$ and $0 \leq j \leq d$.

The crucial lemma in the proof of Claim 13 is as follows.

Claim 12 (i) *For all* $0 \leq j \leq i$, *we have* $|\mathbf{G}(i,j)| \leq i^j (i\ell)^{j\ell}$.

(ii) *For all* $m \geq 1$ *and* $0 \leq d \leq m$, *we have* $|\mathbf{G}_{\leq}(m,d)| \leq (m+1)m^{d+1}(m\ell)^{\ell(d+1)}$.

Proof. (i) For a hypergraph $\mathcal{H}_i$ in $\mathcal{F}(i)$, let $J_i(\mathcal{H}_i) \subseteq [i]$ be the set of integers j for which $\mathcal{H}_i$ is degenerate at level j. In other words, $j \in J_i(\mathcal{H}_i)$ if and only if one of $(j,1)$, $(j,2)$ belongs to $I_i(\mathcal{H}_i)$. For a set $J \subseteq [i]$, let $\mathbf{G}(i,J)$ be the set of all graphs $G \in \mathbf{G}(i,|J|)$ that are underlying graphs of $\mathcal{H}_i \in \mathcal{F}(i)$ with $J_i(\mathcal{H}_i) = J$. For simplicity, put $g(i,j) = |\mathbf{G}(i,j)|$ and $g(i,J) = |\mathbf{G}(i,J)|$.

Let us first fix $J \subseteq [i]$ and estimate $g(i,J)$. We claim that

$$g(i,J) \quad \leq \quad \begin{cases} g(i-1,J) & \text{if } i \notin J \\ g(i-1,J\setminus\{i\})(i\ell)_\ell & \text{if } i \in J, \end{cases} \tag{34}$$

where, as usual, $(a)_b = a(a-1)\cdots(a-b+1)$. Let us first suppose that $i \notin J$. Let $G \in \mathbf{G}(i,J)$, and suppose that $G = G(\mathcal{H}_i)$ with $J_i(\mathcal{H}_i) = J$. Since $i \notin J$, the ranks of the elements of the ith hyperedge E_i of $\mathcal{H}_i$ are uniquely determined once $\mathcal{H}_{i-1} = \mathcal{H}_i \setminus \{E_i\}$ is given. At this point it is crucial that in Procedure Hypertree, when adding a non-degenerate hyperedge to our current hypertree, we choose the hypervertex *with lowest rank* to 'hang' our new hyperedge on. A moment's thought now shows that we have thus implicitly defined an injection $\mathbf{G}(i,J) \hookrightarrow \mathbf{G}(i-1,J)$, since here we are dealing with graphs up to isomorphism.

Suppose now that $i \in J$. Let G, $\mathcal{H}_i$, and E_i be as before. Then the number of ways we can choose the cycle C of G that corresponds to the hyperedge E_i can be bounded very generously by $(i\ell)_\ell$. Indeed, the underlying graph $G(\mathcal{H}_{i-1})$ of $\mathcal{H}_{i-1} = \mathcal{H}_{i-1} \setminus \{E_i\}$ has at most $i\ell$ vertices, and to define C it suffices to select its vertices from $V(G(\mathcal{H}_{i-1}))$. This completes the proof of (34).

By induction, we may deduce from (34) that $g(i,J) \leq \{(i\ell)_\ell\}^{|J|}$, since $g(i,\emptyset) \leq 1$ for any $i \geq 0$. Therefore

$$g(i,j) = |\mathbf{G}(i,j)| = \Big|\bigcup_{J\subseteq[i]} \mathbf{G}(i,J)\Big| \leq \binom{i}{j}\{(i\ell)_\ell\}^j \leq i^j(i\ell)^{j\ell},$$

completing the proof of (i). Assertion (ii) follows easily from (i), since

$$|\mathbf{G}_{\leq}(m,d)| \leq \sum_{i=0}^{m}\sum_{j=0}^{d} g(i,j) \leq \sum_{i=0}^{m}\sum_{j=0}^{d} i^j(i\ell)^{j\ell} \leq \sum_{i=0}^{m}\{i(i\ell)^\ell\}^{d+1} \leq (m+1)m^{d+1}(m\ell)^{\ell(d+1)},$$

as required.

In the sequel, we apply Claim 12(ii) for $m = O(\log n)$ and $d = O(1)$. Our claim tells us that, in this case, $\mathbf{G}_{\leq}(m,d)$ has polylogarithmic size only.

Before stating our last claim, we need to introduce a piece of notation. Below, we shall write $\widetilde{\mathbf{G}}_{\leq}(m,d)$ for the family of graphs $G \in \mathbf{G}_{\leq}(m,d)$ that are isomorphic to underlying graphs $G(\mathcal{H}_m)$ of *outputs* $\mathcal{H}_m$ of Procedure Hypertree.

Claim 13 *Put* $p = p(n) = \mathrm{e}^{-2} n^{-\beta}$, $m' = \lfloor \log n + \ell \rfloor$, *and let* M *be as in* (33). *Then the expected number of subgraphs of* G_p *that are isomorphic to members of* $\widetilde{\mathbf{G}}_{\leq}(m', M)$ *is* $o(1)$.

Proof. An element G of $\widetilde{\mathbf{G}}_{\leq}(m', M)$ corresponds to an output $\mathcal{H}_m$ of Procedure HYPERTREE, and hence the assertions in Claim 11 hold for $\mathcal{H}_m$. In particular, we may write $\widetilde{\mathbf{G}}_{\leq}(m', M) = \widetilde{\mathbf{G}}^{(i)}_{\leq}(m', M) \cup \widetilde{\mathbf{G}}^{(ii)}_{\leq}(m', M)$, where the elements G in $\widetilde{\mathbf{G}}^{(i)}_{\leq}(m', M)$ correspond to $\mathcal{H}_m$ satisfying Claim 11$(b)(i)$, and similarly to $\widetilde{\mathbf{G}}^{(ii)}_{\leq}(m', M)$.

We now claim that the expected number of subgraphs of G_p that are isomorphic to members of $\widetilde{\mathbf{G}}^{(i)}_{\leq}(m', M)$ is $o(1)$. Fix $G \in \widetilde{\mathbf{G}}^{(i)}_{\leq}(m', M)$. Writing $\#\{G \subseteq G_p\}$ for the number of subgraphs of G_p isomorphic to G, we have

$$\mathbf{E}(\#\{G \subseteq G_p\}) \leq \left(\frac{1}{\mathrm{e}^2}\right)^{e(G)} n^{|G| - \beta e(G)} \leq \left(\frac{1}{\mathrm{e}^2}\right)^{\log n} n^{f(0)} \leq n^{-1/2}.$$

Our claim then follows as $|\widetilde{\mathbf{G}}^{(i)}_{\leq}(m', M)|$ is, by Claim 12, only polylogarithmic in n.

Similarly, we claim that the expected number of subgraphs of G_p that are isomorphic to members of $\widetilde{\mathbf{G}}^{(ii)}_{\leq}(m', M)$ is $o(1)$. Indeed, for any $G \in \widetilde{\mathbf{G}}^{(ii)}_{\leq}(m', M)$, we have

$$\mathbf{E}(\#\{G \subseteq G_p\}) \leq n^{|G| - \beta e(G)} \leq n^{-1/500},$$

but again $|\widetilde{\mathbf{G}}^{(ii)}_{\leq}(m', M)|$ is only polylogarithmic in n. This finishes the proof of Claim 13.

To finish the proof of Theorem 3, it suffices to recall that, by Claim 11, the event $G_p \to (C^\ell, C^k)$ implies that G_p contains a subgraph isomorphic to an element of $\widetilde{\mathbf{G}}_{\leq}(m', M)$, where m' and M are as in Claim 13. □

4 Deterministic consequences

Proof of Corollary 7. Let $\beta = \beta(C^{\ell_2}, H)$. Let B and κ be the constants from Theorem 4 for $\ell_2, \ldots, \ell_r$ and H. Let

$$\varepsilon = \min\{1 + 1/(\ell_2 - 1) - |F| + \beta(C^{\ell_2}, H)e(F)\},$$

where the minimum is taken over all graphs $F \in \mathcal{V}_{\ell_2,H,t}$. Then $\varepsilon > 0$ and the expected number of copies of graphs from $\mathcal{V}_{\ell_2,H,t}$ in G_p, where $p = Bn^{-\beta}$, is $\mathcal{O}(n^{1+1/(\ell_2-1)-\varepsilon})$. Therefore, by Markov's inequality, the number of copies of graphs from $\mathcal{V}_{\ell_2,H,t}$ is almost surely at most $\kappa n^{1+1/(\ell_2-1)}/(\log n)^{|H|}$. If G_p is a random graph having this property and fulfilling the statement of Theorem 4 then we define a graph G by removing one edge from every copy of a graph of $\mathcal{V}_{\ell_2,H,t}$ in G_p. The graph G has the required properties. □

Proof of Theorem 8. We show that for any t there exists a graph on more than t vertices which is critical for $(C^{\ell_i})_{1\le i\le r}$. Let B and $\beta = \beta(C^{\ell_2}, C^{\ell_1})$ be the constants from Corollary 5. Let $\mathcal{T} = \{F: |F| \le t \text{ and } |F| - \beta e(F) \le -1/500\}$. For $p = Bn^{-\beta}$, almost all graphs G_p do not contain any graphs from $\mathcal{T}$. Furthermore $G_p \to (C^{\ell_1}, \dots, C^{\ell_r})$ almost surely by Corollary 5. Therefore there exists a graph H on $n \ge t$ vertices with $H \to (C^{\ell_1}, \dots, C^{\ell_r})$ and not containing any graph from $\mathcal{T}$ as a subgraph. Assume that there exists a subgraph K of H on t vertices such that $K \to (C^{\ell_1}, \dots, C^{\ell_r})$. Hence in particular $K \to (C^{\ell_1}, C^{\ell_2})$. Let $\mathcal{K}$ be the hypergraph with hyperedges being the ℓ_1-and ℓ_2-cycles of K. Let $\mathcal{K}_0$ be a critical subgraph of $\mathcal{K}$ with respect to property (*) from Section 3. We now run **Procedure** Hypertree on the input $\mathcal{K}_0$ but where we exit the **repeat**-loop in $\langle$*main body*$\rangle$ only if $|G(\mathcal{H}_i)| - \beta e(G(\mathcal{H}_i)) \le -1/500$ thus possibly creating hypertrees that are of length more than $\log t$. It follows easily from the proof of the 0-statement (cf., *e.g.*, Claim 11) that the output $\mathcal{H}_m$ is such that its underlying graph $G(\mathcal{H}_m)$ is in $\mathcal{T}$, contradicting the fact that H is $\mathcal{T}$-free. It follows that some subgraph of H containing more than t vertices is critical for $(C^{\ell_i})_{1\le i\le r}$. □

References

[1] B. Bollobás. *Random Graphs.* Academic Press, 1985.

[2] R.L. Graham, B. Rothschild, and J.H. Spencer. *Ramsey Theory.* Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons, New York, second edition, 1990.

[3] P.E. Haxell, Y. Kohayakawa, and T. Łuczak. Turán's extremal problem in random graphs: forbidding odd cycles. *Combinatorica.* To appear.

[4] P.E. Haxell, Y. Kohayakawa, and T. Łuczak. Turán's extremal problem in random graphs: forbidding even cycles. *J. Combinatorial Theory*, Series B, 64:273–288, 1995.

[5] S. Janson. Poisson approximation for large deviations. *Random Structures and Algorithms*, 1(2):221–229, 1990.

[6] Y. Kohayakawa. The regularity lemma of Szemerédi for sparse graphs. Manuscript, 1993.

[7] Y. Kohayakawa, B. Kreuter, and A. Steger. An extremal problem for random graphs and the number of graphs with large even-girth. Submitted, 1995.

[8] Y. Kohayakawa, T. Łuczak, and V. Rödl. Arithmetic progressions of length three in subsets of a random set. *Acta Arithmetica.* To appear.

[9] Y. Kohayakawa, T. Łuczak, and V. Rödl. On K^4-free subgraphs of random graphs. Submitted, 1995.

[10] B. Kreuter. Threshold functions for asymmetric Ramsey properties with respect to vertex colorings. Submitted, 1995.

[11] T. Łuczak. On Ramsey minimal graphs. *The Electronic Journal of Combinatorics*, 1:R#4, 1994.

[12] T. Łuczak, A. Ruciński, and B. Voigt. Ramsey properties of random graphs. *J. Combinatorial Theory*, Series B, 56:55–68, 1992.

[13] V. Rödl and A. Ruciński. Lower bounds on probability thresholds for Ramsey properties. In *Combinatorics—Paul Erdős is Eighty* (Volume 1), Bolyai Soc. Math. Studies, pages 317–346, Budapest, 1993.

[14] V. Rödl and A. Ruciński. Random graphs with monochromatic triangles in every edge coloring. *Random Structures and Algorithms*, 5(2):253–270, 1994.

[15] V. Rödl and A. Ruciński. Threshold functions for Ramsey properties. *J. Amer. Math. Soc.*, 8(4):917–942, 1995.

[16] E. Szemerédi. Regular partitions of graphs. In J.-C. Bermond, J.-C. Fournier, M. Las Vergnas, and D. Sotteau, editors, *Problèmes Combinatoires et Théorie des Graphes, Proc. Colloque Inter. CNRS*, pages 399–401, Paris, 1978. CNRS.