Peter Allen, Julia Böttcher, Yoshiharu Kohayakawa, and Barnaby Roberts

Triangle-free subgraphs of random graphs

Article (Accepted version) (Refereed)

TRIANGLE-FREE SUBGRAPHS OF RANDOM GRAPHS

PETER ALLEN, JULIA BÖTTCHER, YOSHIHARU KOHAYAKAWA, AND BARNABY ROBERTS

Abstract. Recently there has been much interest in studying random graph analogues of well known classical results in extremal graph theory. Here we follow this trend and investigate the structure of triangle-free subgraphs of $G(n,p)$ with high minimum degree. We prove that asymptotically almost surely each triangle-free spanning subgraph of $G(n,p)$ with minimum degree at least $\big(\frac{2}{5}+o(1)\big)pn$ is $\mathcal{O}(p^{-1}n)$-close to bipartite, and each spanning triangle-free subgraph of $G(n,p)$ with minimum degree at least $(\frac{1}{3}+\varepsilon)pn$ is $\mathcal{O}(p^{-1}n)$-close to r-partite for some $r=r(\varepsilon)$. These are random graph analogues of a result by Andrásfai, Erdős, and Sós [Discrete Math. 8 (1974), 205–218], and a result by Thomassen [Combinatorica 22 (2002), 591–596]. We also show that our results are best possible up to a constant factor.

1. Introduction

In a 1948 edition of the recreational math journal Eureka, Blanche Descartes proved that triangle-free graphs can have arbitrarily large chromatic number, and thus be complex in structure. This motivates the question of which additional restrictions on the class of triangle-free graphs allow for a bound on the chromatic number. By Mantel's theorem [17], the densest triangle-free graphs are balanced complete bipartite graphs. So we may first ask whether triangle-free graphs H with minimum degree somewhat below $\frac{1}{2}v(H)$ are still necessarily bipartite. This is true, as Andrásfai, Erdős and Sós showed in 1974.

Theorem 1 (Andrásfai, Erdős, Sós [4])**.** *All triangle-free graphs H with $\delta(H) > \frac{2}{5}v(H)$ are bipartite.*

Triangle-free graphs of smaller minimum degree do not need to be bipartite, as blow-ups of a 5-cycle illustrate. But one may still ask whether their chromatic number is bounded (questions of this type were first addressed by Erdős and Simonovits in [11]). In 2002 Thomassen [19] proved that this is the case for triangle-free graphs of minimum degree at least $(\frac{1}{3}+\varepsilon)n$.

Theorem 2 (Thomassen [19])**.** *For any $\varepsilon > 0$ there exists r_ε such that if H is triangle-free and $\delta(H) > (\frac{1}{3}+\varepsilon)v(H)$ then H is r_ε-colourable.*

A construction of Hajnal (see [11]) shows that the minimum degree bound in this theorem cannot be replaced by $(\frac{1}{3}-\varepsilon)n$. A much stronger result was established by Brandt and Thomassé [7], who showed that triangle-free graphs H with $\delta(H) > \frac{1}{3}n$ are 4-colourable.

1991 *Mathematics Subject Classification.* 05C35,05C80.

The third author was partially supported by FAPESP (2013/03447-6, 2013/07699-0), CNPq (310974/2013-5 and 459335/2014-6), NSF (DMS 1102086) and NUMEC/USP (Project MaCLinC/USP).

In this paper we are interested in random graph analogues of Theorem 1 and Theorem 2. Establishing such analogues for prominent results in extremal graph theory has been a particularly fruitful area of study in the last few years. A good overview can be found in Conlon's survey paper [8].

In order to study these kinds of questions systematically, Kohayakawa [13] and Rödl (unpublished) developed a sparse analogue of Szemerédi's Regularity Lemma, and, together with Łuczak [14] formulated the KŁR conjecture which asserts the existence of a corresponding 'counting lemma'. Recently Conlon, Samotij, Schacht and Gowers [9] proved this conjecture (see also [5, 18]). It is easy (as observed in [9]) to use these results to prove 'approximate' random versions of Theorems 1 and 2, as well as to re-prove Mantel's theorem for random graphs. Thus if $p \gg n^{-1/2}$ then asymptotically almost surely (a.a.s.) the random graph $G(n,p)$ has the property that all subgraphs with minimum degree a little larger than $\frac{2}{5}pn$ can be made bipartite by deleting $o(pn^2)$ edges. Similarly, the sparse random version of Mantel's theorem obtained states that any subgraph with a little more than half the edges of $G(n,p)$ contains a triangle.

One might expect that all subgraphs of $G(n,p)$ with minimum degree a little larger than $\frac{2}{5}pn$ are bipartite. Indeed, an alternative sparse random version of Mantel's theorem, proved by DeMarco and Kahn [10], states that a largest triangle-free subgraph of $G(n,p)$ coincides exactly with a largest bipartite subgraph for $p \gg (\log n/n)^{1/2}$. However, subgraphs of $G(n,p)$ with minimum degree larger than $\frac{2}{5}pn$ which are not bipartite do exist (see Theorem 5 below). In this paper we determine for all p how far from bipartite such graphs can be.

Theorem 3. *For any $\gamma > 0$, there exists C such that for any $p(n)$ the random graph $\Gamma = G(n,p)$ a.a.s. has the property that all triangle-free spanning subgraphs $H \subseteq \Gamma$ with $\delta(H) \geqslant (\frac{2}{5} + \gamma)pn$ can be made bipartite by removing at most $\min\big(Cp^{-1}n, (\frac{1}{4} + \gamma)pn^2\big)$ edges.*

In addition we derive an analogous random graph version of Theorem 2.

Theorem 4. *For any $\gamma > 0$, there exist C and r such that for any $p(n)$ the random graph $\Gamma = G(n,p)$ a.a.s. has the property that all triangle-free spanning subgraphs $H \subseteq \Gamma$ with $\delta(H) \geqslant (\frac{1}{3} + \gamma)pn$ can be made r-partite by removing at most $\min\big(Cp^{-1}n, (\frac{1}{2r} + \gamma)pn^2\big)$ edges.*

Up to the values of C, these theorems are best possible.

Theorem 5. *For any $\gamma > 0$ and $r \in \mathbb{N}$, there exist constants $c, c' > 0$ such that if $n^{-1/2}/c' \leqslant p(n) \leqslant c'$ then $\Gamma = G(n,p)$ a.a.s has a triangle-free spanning subgraph H with $\delta(H) \geqslant (\frac{1}{2} - \gamma)pn$ which cannot be made r-partite by removing fewer than $cp^{-1}n$ edges.*

Note that for $p \ll n^{-1/2}$ the minimum in each of Theorems 3 and 4 is achieved by the second term and that these statements are easy: For such values of p only a tiny fraction of the edges of $G(n,p)$ are in triangles and the question reduces to asking for the largest bipartite (respectively, r-partite) subgraph of $G(n,p)$. For p close to 1, by the original Theorems 1 and 2, the conclusion of Theorem 5 becomes false, so that we need the condition $p \leqslant c'$.

It would be interesting to obtain analogous results for K_r-free subgraphs of $G(n,p)$ for $r > 3$. It would also be interesting to know whether Theorem 4 could be improved to generalise the result of Brandt and Thomassé. We conjecture that this is the case.

Organisation. In Section 2 we will introduce some of the main tools that will be used throughout the paper. Section 3 of this paper will give a method of constructing a triangle-free subgraph from a given, randomly generated graph. We will then prove a series of results about this construction which will result in proving Theorem 5. In Section 4 we will state and prove some properties that a.a.s. $\Gamma = G(n,p)$ possesses. We will then use these properties in Section 5 to prove Theorem 3, and in Section 6 to prove Theorem 4.

2. Tools

Notation. We write $[n]$ for the set $\{1, ..., n\}$, and the notation $x = (1 \pm \varepsilon)$ is used to mean $x \in [1-\varepsilon, 1+\varepsilon]$.

In a graph G we say a vertex is a *common neighbour* of a pair of vertices if it is adjacent to both of them. For disjoint sets of vertices X and Y in G we will use $E_G(X,Y)$ to denote the set of edges between X and Y in G and $E_G(X)$ to denote the set of edges of G with both ends in X. We denote the sizes of these sets by $e_G(X,Y)$ and $e_G(X)$ respectively. We will use $N_G(v,X)$ to denote the set of vertices in X which are adjacent to a vertex v of G and $\deg_G(v,X)$ for the number of vertices in $N_G(v,X)$. For two vertices u, v we will write $N_G(u,v,X)$ for the common neighbourhood $N_G(u,X) \cap N_G(v,X)$ of u and v in X, and $\deg_G(u,v,X)$ for its size. For $X = V(G)$ we will simply use $N_G(v)$, $\deg_G(v)$ and $N_G(u,v)$. Often, when it is clear which graph is being referred to, we also omit the subscripts.

Throughout the paper we shall omit floor and ceiling symbols when this does not affect our argument.

Probability. We write $\mathrm{Bin}(n,p)$ for the binomial distribution with n trials and success probability p. Our proofs we will make frequent use of the following Chernoff bound, which is an immediate corollary of [12, Theorem 2.1].

Lemma 6 (Chernoff bound)**.** *Let X be a random variable with distribution* $\mathrm{Bin}(n,p)$ *and* $0 < \delta < \frac{3}{2}$*. Then*

$$\mathbb{P}(X < (1-\delta)\mathbb{E}X) < \exp\left(\tfrac{-\delta^2}{3}\mathbb{E}X\right) \quad \textit{and} \quad \mathbb{P}(X > (1+\delta)\mathbb{E}X) < \exp\left(\tfrac{-\delta^2}{3}\mathbb{E}X\right). \tag{1}$$

Sparse regularity. We define the *density* $d(U,V)$ of a pair of disjoint vertex sets (U,V) to be the value $e(U,V)/|U||V|$. A pair (U,V) is called (ε,d,p)*-lower-regular* if for any sets $U' \subseteq U$, $V' \subseteq V$ satisfying $|U'| \geqslant \varepsilon|U|$, $|V'| \geqslant \varepsilon|V|$ we have $d(U',V') \geqslant (d-\varepsilon)p$. We say a pair (U,V) is (ε,d,p)*-regular* if $d(U,V) \geqslant dp$ and for any sets $U' \subseteq U$, $V' \subseteq V$ satisfying $|U'| \geqslant \varepsilon|U|$, $|V'| \geqslant \varepsilon|V|$ we have $d(U',V') = (d(U,V) \pm \varepsilon)$. We say (U,V) is (ε,p)*-regular* if it is (ε,d,p)*-regular* for some d.

An (ε,p)*-regular-partition* of a graph H is a vertex partition $V_0 \cup V_1 \cup \cdots \cup V_t$ of $V(G)$ with $|V_0| \leqslant \varepsilon|V|$ and $|V_1| = |V_2| = \cdots = |V_t|$ such that all but at most $\varepsilon\binom{t}{2}$ pairs (V_i, V_j)

with $i, j \geqslant 1$ are (ε, p)-regular. The corresponding (ε, d, p)-*reduced graph* R is the graph with vertex set $[t]$ where ij is an edge precisely if (V_i, V_j) is an (ε, d, p)-lower-regular pair in H. The following version of the Sparse Regularity Lemma can be deduced from [1, Lemma 12] [1].

Lemma 7 (Sparse Regularity Lemma, Minimum Degree Form Version)**.** *For all $\beta \in [0, 1]$, $\varepsilon > 0$ and every integer t_0 there exists $t_1 \geqslant 1$ such that for all $d \in [0, 1]$ the following holds for any $p > 0$. For any graph G on n vertices with minimum degree βpn, such that for any $X, Y \subseteq V(G)$ with $|X|, |Y| \geqslant \frac{\varepsilon n}{t_1}$ we have $e(X, Y) \leqslant (1 + \frac{1}{1000}\varepsilon^2)p|X||Y|$, there is a regular-partition of $V(G)$ with (ε, d, p)-reduced graph R satisfying $\delta(R) \geqslant (\beta - d - \varepsilon)|V(R)|$ and $t_0 \leqslant |V(R)| \leqslant t_1$. Furthermore, for each $i \in V(R)$ the number of $j \in V(R)$ such that (V_i, V_j) is not (ε, p)-regular is at most $\varepsilon v(R)$, and for each $i \in V(R)$ and $v \in V_i$, at most $(d + \varepsilon)pn$ neighbours of v lie in $\bigcup_{j:ij \notin R} V_j$.*

Note that the regularity lemma above is not specifically for $G(n, p)$ but for graphs in which the density edges between pairs of large sets is never much greater than p. For $p = \omega(\frac{\log n}{n})$ the random graph $G(n, p)$ a.a.s. satisfies this, see for example Lemma 14 part (c).

When applying the Sparse Regularity Lemma we will wish to say that if H is triangle-free then the reduced graph is also triangle-free. In order to do this we use the following regularity inheritance lemma, which is [3, Lemma 1.27] and is based on techniques from [15].

Lemma 8 (Regularity Inheritance)**.** *For any $0 < \varepsilon', d$ there exist ε_0 and C' such that for any $0 < \varepsilon < \varepsilon_0$ and any $0 < p = p(n) < 1$ the random graph $\Gamma = G(n, p)$ a.a.s. has the following property. For any $X, Y \subseteq V(\Gamma)$ with $|X|, |Y| \geqslant C' \max\{p^{-2}, p^{-1}\log n\}$ and any subgraph H of $\Gamma[X, Y]$ which is (ε, d, p)-lower-regular, there are at most $C' \max\{p^{-2}, p^{-1}\log n\}$ vertices v of $V(\Gamma)$ such that $(X \cap N_\Gamma(v), Y \cap N_\Gamma(v))$ is not (ε', d, p)-lower-regular in H.*

We shall also want the following consequence of this lemma, stating that for every regular partition of every $H \subseteq G(n, p)$ the neighbourhoods of most vertices induce lower-regular subgraphs on the regular pairs of the partition.

Lemma 9. *For any $0 < \varepsilon', d < 1$ there exist ε_0 and C' such that for any $t_1 \in \mathbb{N}$ and any $p > 2C't_1 n^{-1/2}$ the random graph $\Gamma = G(n, p)$ a.a.s. satisfies the following. For any $0 < \varepsilon < \varepsilon_0$, any spanning subgraph H of Γ and any (ε, d, p)-regular-partition $V_0 \cup V_1 \cup \cdots \cup V_t$ of H with $t \leqslant t_1$ and reduced graph R, all but at most $\binom{t_1}{2} C' \max\{p^{-2}, p^{-1}\log n\}$ vertices*

[1]The statement is identical to that in [1] except for the final 'Furthermore' conclusion. That we can assume no part is in many irregular pairs follows from the proof there. The final condition can be obtained by applying the statement in [1] with $\varepsilon/100$ replacing ε and removing vertices from $V_1, \ldots, V_{v(R)}$ to V_0, keeping the sizes of the V_i equal, until no vertices failing the condition remain. Initially, by regularity and by the upper bound on densities in G, we remove at most $\frac{\varepsilon}{20}n$ vertices. Thereafter, we remove vertices only because they have at least $\varepsilon pn/2$ neighbours in the current set V_0. If at some point in the process V_0 has $\varepsilon n/(10)$ vertices, then it contains at least $\varepsilon^2 pn^2/(40)$ edges, so contains a bipartite subgraph with at least $\varepsilon^2 pn^2/(80)$ edges, in contradiction to the density assumption on G. We conclude the process stops before this point, as desired.

v of H have the property that for each $ij \in E(R)$ the pair $(N_\Gamma(v) \cap V_i, N_\Gamma(v) \cap V_j)$ is (ε', d, p)-lower-regular in H.

Proof. By applying Lemma 8 with ε' and d we are given ε_0 and C'. Suppose $p \geqslant 2C'tn^{-1/2}$ and that Γ satisfies the probable event of Lemma 8. Now let $H \subseteq \Gamma$ and a partition $V_0 \cup V_1 \cup \cdots \cup V_t$ of H with reduced graph R be given. Let $ij \in E(R)$. For large enough n we have $C' \max\{p^{-2}, p^{-1}\log n\} \leqslant C' \max\{\frac{n}{4C'^2t_1^2}, \frac{\sqrt{n}\log n}{2C't_1}\} \leqslant \frac{n}{2t_1} \leqslant |V_i|, |V_j|$. So we conclude from Lemma 8 that for all but at most $C'\max\{p^{-2}, p^{-1}\log n\}$ vertices $v \in V(H)$ the pair $(N_\Gamma(v) \cap V_i, N_\Gamma(v) \cap V_j)$ is (ε', d, p)-lower-regular in H. The lemma follows by summing over all $ij \in E(R)$. □

The following lemma combines Lemma 7 with Lemma 8 to give a regular partition of a triangle-free subgraph H for which the reduced graph is triangle-free.

Lemma 10. *For any $0 < \varepsilon, d, \beta < 1$ and any t_0 there exist c and t_1 such that for $p \geqslant cn^{-1/2}$ in $\Gamma = G(n,p)$ a.a.s. any triangle-free subgraph H with $\delta(H) > \beta pn$ has an (ε, d, p)-regular-partition $V_0 \cup V_1 \cup \cdots \cup V_t$ with $t_0 \leqslant t \leqslant t_1$ such that the corresponding reduced graph R is triangle-free and has minimum degree at least $(\beta - d - \varepsilon)v(R)$.*

Proof. Suppose we are given $\varepsilon, d, \beta, t_0$ as in the lemma statement. Set $\varepsilon' = \frac{d}{3}$ and apply Lemma 8 (Regularity Inheritance) to ε' and d to obtain ε_0 and C'. Now apply Lemma 7 (Sparse Regularity, Minimum Degree Form) with d, β, t_0 as given and with ε also required to be smaller than ε_0. This gives t_1. Take $c = 6t_1C'$.

Lemma 7 has given us an (ε, d, p)-regular-partition of H with reduced graph R that satisfies all the conditions we require except that of R being triangle free. Suppose for a contradiction there is a triangle in R. This corresponds to an (ε, d, p)-lower-regular triple (X, Y, Z). First observe that $|X| = |Y| \geqslant \frac{n}{2t_1}$ and for $p(n) \geqslant cn^{-1/2}$ we have $\frac{n}{4t_1} > C'\max\{p^{-2}, p^{-1}\log n\}$. By lower-regularity of $(X < Z)$ and (Y, Z), at least $\frac{1}{2}|Z|$ vertices z of Z have $\deg_H(z, X) \geqslant \frac{d}{2}p|X|$ and also $\deg_H(z, Y) \geqslant \frac{d}{2}p|Y|$. Furthermore, for all but at most $C'\max\{p^{-2}, p^{-1}\log n\} \leqslant \frac{|Z|}{3}$ vertices z of Z, the pair $\big(N_\Gamma(z, X), N_\Gamma(z, Y)\big)$ is (ε', d, p)-lower-regular. Choosing a vertex $z \in Z$ which satisfies both conditions, by regularity of $\big(N_\Gamma(z, X), N_\Gamma(z, Y)\big)$ the edge density of $\big(N_H(z, X), N_H(z, Y)\big)$ is at least $(d - \varepsilon)p > 0$. This gives a triangle, the desired contradiction. □

Finally, we need the following special case of the Slicing Lemma.

Lemma 11 (Slicing Lemma)**.** *Let (V_i, V_j) be (ε, d, p)-lower-regular. For any $X \subseteq V_i$, $Y \subseteq V_j$ such that $|X| \geqslant d|V_i|, |Y| \geqslant d|V_j|$ the pair (X, Y) is $(\frac{\varepsilon}{d}, d, p)$-lower-regular.*

Proof. Let $X' \subseteq X$, $Y' \subseteq Y$ satisfy $|X'| \geqslant \frac{\varepsilon}{d}|X| \geqslant \varepsilon|V_i|$ and $|Y'| \geqslant \frac{\varepsilon}{d}|Y| \geqslant \varepsilon|V_j|$. So $d(X', Y') \geqslant (d - \varepsilon)p \geqslant \big(d - \frac{\varepsilon}{d}\big)p$. □

3. Proof of Theorem 5

Recall that Theorem 5 asserts that for any $\gamma > 0$ and $r \in \mathbb{N}$, there are $c, c' > 0$ such that for any $n^{-1/2}/c' \leqslant p \leqslant c'$ the random graph $G(n,p)$ a.a.s. contains a subgraph which is triangle-free, whose minimum degree is at least $\big(\frac{1}{2}-\gamma\big)pn$, and which cannot be made r-partite by removing any $cp^{-1}n$ edges.

The idea of the proof of this theorem is as follows. Let $\Gamma = G(n,p)$ and partition $[n]$ into sets $B=[n/2]$ and $A=[n]\setminus B$. We remove all edges in A. We further 'sparsify' $\Gamma[B]$, keeping edges with a suitable probability p'. The goal of this 'sparsification' is to obtain a subgraph of $\Gamma[B]$ which is still complex enough for the rest of the argument, but is such that for each vertex a in A the number of edges in $N(a,B)$ is negligible compared to the degree of a (see Lemma 12(*b*)). Observe that this subgraph is distributed as the following inhomogeneous random graph model. We define $G(n,p,p')$ to be the random graph on $[n]$ obtained by letting pairs of vertices within $[n/2]$ be edges independently with probability pp', letting pairs in $[n]\setminus[n/2]$ all be non-edges, and letting all other pairs be edges independently with probability p.

We next use the fact, first proved in [6], that there exists a triangle-free graph F which is not r-partite. Let $[\ell]$ be the vertex set of F. We place a 'random blow-up' of F into B as follows: We partition B into ℓ equal sets $B_1,\dots,B_\ell$ and keep only those edges in B running between B_i and B_j with $ij\in F$. Finally, we remove in B all edges with an endpoint whose degree in B deviates too much from expectation, and then all edges between A and B which are in a triangle with a vertex from A. This last step is the only step in which we delete edges between A and B.

It is easy to check that the resulting graph is triangle-free by construction. Using some properties of $G(n,p,p')$ and the blow-up of F we can also show that it cannot be made r-partite by deleting $cp^{-1}n$ edges. Moreover, using the fact that for each vertex a in A the number of edges in $N(a,B)$ is small and hence in the last step not many edges were deleted at any vertex, we can also conclude that the minimum degree of the resulting graph is at least $\big(\frac{1}{2}-\gamma\big)pn$.

The typical properties of $G(n,p,p')$ we need are the following.

Lemma 12. *For any $\varepsilon>0$ and $K\geqslant 10$, there exists $0<c<\varepsilon$ such that the following holds. If $Kn^{-1/2}\leqslant p(n)\leqslant \varepsilon^2c/(10^4K^2)$ and $p'=cK^2p^{-2}n^{-1}$, then a.a.s. the random graph $G(n,p,p')$ has the following properties. Let $B=[n/2]$ and $A=[n]\setminus B$.*

(*a*) $\deg(b,A),\deg(a,B)=\big(\frac{1}{2}\pm\varepsilon\big)pn$ *for every $a\in A$ and $b\in B$.*

(*b*) *For each $a\in A$, at most $p'p^3n^2$ edges have both ends in $N(a,B)$.*

(*c*) *For each $b\in B$ with $\deg(b,B)\geqslant\frac{1}{10}p'pn$, the number of vertices $a\in A$ such that there exists $b'\in B$ with abb' a triangle is at most $pn\big(1-(1-p)^{\deg(b,B)}\big)$.*

(*d*) *At most $cp^{-1}n$ edges in B are incident to some $b\in B$ with $\deg(b,B)\geqslant pp'n$ or $\deg(b,B)\leqslant\frac{1}{10}p'pn$.*

(*e*) *$e(U,V)>2cp^{-1}n$ for every pair of disjoint sets $U,V\subseteq B$ with $|U|,|V|\geqslant 2n/K$.*

We delay the proof of this lemma to after the proof of Theorem 5.

Proof of Theorem 5. Given $\gamma>0$ and $r\in\mathbb{N}$, let F be a triangle-free graph which is not r-partite. Let $\ell=v(F)$. We set $K=8r\ell$ and

$$\varepsilon=\tfrac{1}{400}\gamma r^{-2}\ell^{-2}\,. \qquad (2)$$

Now we let $c > 0$ with $c < \varepsilon$ be returned by Lemma 12 for input ε and K. We choose $c' = \min\big(\frac{1}{K}, \frac{c}{10^4}\big)$.

Given $n^{-1/2}/c' \leqslant p(n) \leqslant c'$, let $p' = cK^2p^{-2}n^{-1}$. Observe that $p' \leqslant 1$ by choice of p. Let $B = [n/2]$, and $A = [n] \setminus B$. We generate $\Gamma = G(n,p)$, and let G_1 be the subgraph of Γ obtained by sparsifying B, keeping edges independently with probability p' and removing all edges of A. Since G_1 is distributed as $G(n,p,p')$, by Lemma 12 it a.a.s. satisfies the properties (a)–(e). We now condition on G_1 satisfying these properties.

Partition B into ℓ equal sets $B_1, \dots, B_\ell$. Let G_2 be the subgraph of G_1 obtained by keeping only edges of the form ab with $a \in A$ and $b \in B$, or of the form bb' with $b \in B_i$ and $b' \in B_j$ for some $ij \in F$. We claim that $G_2[B]$ is far from r-partite.

Claim 13. *$G_2[B]$ cannot be made r-partite by deleting any $2cp^{-1}n$ edges.*

Proof. Given a (not necessarily proper) r-colouring $\chi : B \to [r]$, we define a majority r-colouring $\chi' : [\ell] \to [r]$ by setting $\chi'(i)$ equal to the smallest j such that $\big|\chi^{-1}(j) \cap B_i\big| \geqslant |B_i|/r$. Since F is not r-partite, the colouring χ' is not proper, and hence there exists $ij \in F$ such that $\chi'(i) = \chi'(j)$. The subsets B'_i and B'_j of B_i and B_j respectively which are given colour $\chi'(i)$ by χ are by construction disjoint and each of size at least $n/(4r\ell) = 2n/K$. Thus by Lemma 12(e) we have $e(B'_i, B'_j) > 2cp^{-1}n$, and the claim follows. □

Now we let G_3 be obtained from G_2 by deleting all edges of $G_2[B]$ which use a vertex $b \in B$ with $\deg(b,B) \geqslant pp'n$ or $\deg(b,B) \leqslant pp'n/10$. By Lemma 12$(d)$ the number of edges deleted is at most $cp^{-1}n$.

Finally, we let H be obtained from G_3 by deleting all edges ab of G_3 with $a \in A$ and $b \in B$ such that there exists $b' \in B$ with abb' a triangle of G_3. Observe that since A is independent in H, any triangle of H has at most one vertex in A. By construction of H, there are no triangles with exactly one vertex in A, so any triangle of H has all three vertices in B. But then the three vertices of a triangle in H would lie in sets B_i, B_j and B_k with ijk a triangle in F, and we chose F to be a triangle-free graph. We conclude that H is triangle-free. Furthermore, if H can be made r-partite by deleting $cp^{-1}n$ edges, then certainly $H[B]$ can be made r-partite by deleting $cp^{-1}n$ edges. But since we deleted at most $cp^{-1}n$ edges from $G_2[B]$ in order to obtain $G_3[B]$, and no further edges to obtain $H[B]$, this implies $G_2[B]$ can be made r-partite by deleting at most $2cp^{-1}n$ edges, in contradiction to Claim 13.

It remains only to show that $\delta(H) \geqslant \big(\frac{1}{2} - \gamma\big)pn$. First consider any vertex $b \in B$. By Lemma 12(a) we have $\deg_{G_1}(b,A) \geqslant \big(\frac{1}{2} - \varepsilon\big)pn$. By construction, no edge from b to A was deleted in creating G_2 from G_1, or G_3 from G_2. By construction of G_3, either $\deg_{G_3}(b,B) = 0$, in which case no edge from b to A was deleted in creating H, or we have $\frac{1}{10}pp'n \leqslant \deg_{G_1}(b,B) \leqslant pp'n$. By Lemma 12$(c)$ we conclude that the total number of edges deleted from b to A in forming H from G_3 is at most

$$pn\big(1 - (1-p)^{pp'n}\big) \leqslant p^3p'n^2 \leqslant 64r^2\ell^2cpn \overset{(2)}{\leqslant} \tfrac{1}{2}\gamma pn\,,$$

because $c < \varepsilon$. Thus we have

$$d_H(b) \geqslant \big(\tfrac{1}{2} - \varepsilon\big)pn - \tfrac{1}{2}\gamma pn \overset{(2)}{\geqslant} \big(\tfrac{1}{2} - \gamma\big)pn$$

as desired.

Now consider any $a \in A$. Again by Lemma 12(a) we have $\deg_{G_1}(a, B) \geqslant \big(\tfrac{1}{2} - \varepsilon\big)pn$. Again no edges from a to B are deleted in forming G_2 or G_3. In forming H from G_3, we delete edges from a to each of b and b' in B whenever abb' forms a triangle in G_3. Since $G_3[B]$ is a subgraph of $G_1[B]$, this means that we delete at most $2e\big(N_{G_1}(a;B)\big)$ edges from a to B, which by Lemma 12(b) is at most $2p'p^3n^2$. Thus we have

$$d_H(a) \geqslant \big(\tfrac{1}{2} - \varepsilon\big)pn - 2p'p^3n^2 \overset{(2)}{\geqslant} \big(\tfrac{1}{2} - \tfrac{1}{2}\gamma\big)pn - \tfrac{1}{2}\gamma pn = \big(\tfrac{1}{2} - \gamma\big)pn\,,$$

which completes the proof. □

We now give the proof of Lemma 12.

Proof of Lemma 12. Choose $c = \min\{\frac{1}{2}\varepsilon, K^{-2}\}$. These properties follow from easy applications of the Chernoff bound, Lemma 6. We omit the proof of (a) as it is standard.

(b): By property (a) we may assume that there are at most $(\frac{1}{2} + \varepsilon)pn$ vertices in $N(a, B)$ for each $a \in A$. Now consider an arbitrary set S of $(\frac{1}{2} + \varepsilon)pn$ vertices in B. The expected number of edges in S is $\binom{|S|}{2}p'p \leqslant \frac{1}{2}|S|^2p'p$. By Lemma 6 the probability that S has more than $|S|^2p'p \leqslant p'p^3n^2$ edges is less than $\exp(\frac{-1}{6}|S|^2p'p) \leqslant \exp(-\frac{1}{100}p'p^3n^2) = \exp(-\frac{1}{100}K^2cpn)$. Hence the claimed property follows by taking a union bound over all $a \in A$.

(c): Assume that we first only reveal the edges of $G(n, p, p')$ in B and consider a vertex $b \in B$ for which $\deg(b, B) \geqslant \frac{1}{10}p'pn$. Now reveal also the edges between A and B. Then a fixed $a \in A$ forms a triangle with b in which the third vertex is also in B with probability $p \cdot (1 - (1-p)^{\deg(b,B)})$. Therefore the expected number of such $a \in A$ is

$$\frac{1}{2}np(1 - (1-p)^{\deg(b,B)}) \geqslant \frac{1}{2}np \cdot (1 - (1-p)^{p'pn/10}) \geqslant \frac{1}{40}p'p^3n^2\,,$$

where the inequality follows from $1 - (1-p)^{p'pn/10} \geqslant \frac{1}{10}p'p^2n - \frac{1}{100}p'^2p^4n^2 \geqslant \frac{1}{20}p'p^2n$, which uses $p' = K^2cp^{-2}n^{-1}$. Hence by Lemma 6 the probability that there are more than $pn(1 - (1-p)^{\deg(b,B)})$ such $a \in A$ is less than $\exp(-10^{-3}p'p^3n^2) = \exp(-10^{-3}K^2cpn)$. Taking a union bound over vertices in B the claimed property follows.

(d): Two applications of Lemma 6 and simple union bounds show that a.a.s. for any $S \subseteq B$ with $|S| = n/(2K^2)$ we have

$$e(S) \leqslant (1+\varepsilon)p'p\binom{|S|}{2} \quad \text{and} \tag{3}$$

$$e(S, B \setminus S) = (1 \pm \varepsilon)p'p|S||B \setminus S|\,, \tag{4}$$

since $p \leqslant \varepsilon^2 c/(10^4 K^2)$. This implies that for any $S \subseteq B$ with $|S| \leqslant n/(2K^2)$ the number of edges in B adjacent to S is at most

$$(1+\varepsilon)p'p\binom{n/(2K^2)}{2} + (1+\varepsilon)p'p\frac{n}{2K^2}\Big(\frac{n}{2} - \frac{n}{2K^2}\Big) \leqslant (1+\varepsilon)p'p\frac{n}{2K^2}\cdot\frac{n}{2} \leqslant \frac{1}{2}cp^{-1}n\,.$$

Hence, with $C = \{b \in B\colon \deg(b,B) \leqslant \frac{1}{10}p'pn\}$ and $D = \{b \in B\colon \deg(b,B) \geqslant p'pn\}$, the claimed property follows if $|C| \leqslant n/(2K^2)$ and $|D| \leqslant n/(2K^2)$.

So assume that there is $C' \subseteq C$ with $|C'| = n/(2K^2)$. But then $e(C', B\setminus C') \leqslant |C'|\frac{1}{10}p'pn \leqslant \frac{1}{20K^2}p'pn^2$, contradicting (4). Similarly, assuming there is $D' \subseteq D$ with $|D'| = n/(2K^2)$ and using (3) we get

$$e(D', B\setminus D') \geqslant |D'|p'pn - 2e(D') \geqslant \frac{n^2p'p}{2K^2} - (1+\varepsilon)p'p\Big(\frac{n}{2K^2}\Big)^2 \geqslant \frac{1}{3K^2}p'pn^2\,,$$

contradicting (4).

(e): For any disjoint $U, V \subseteq B$ each with at least $\frac{2n}{K}$ vertices the expected number of edges between U and V is $|U||V|p'p \geqslant \frac{4n^2}{K^2}p'p = 4cp^{-1}n$, so the result follows from another application of Lemma 6 and a union bound (using $p \leqslant \varepsilon^2 c/(10^4K^2)$). □

4. Auxiliary properties of $G(n,p)$

In this section we list some typical properties of $G(n,p)$, which we shall use in the proofs of Theorems 3 and 4.

Lemma 14. *For any $0 < \varepsilon < \frac{3}{2}$ and $M \in \mathbb{N}$ and any $p = \omega\big(\frac{\ln n}{n}\big)$, the graph $\Gamma = G(n,p)$ a.a.s. satisfies the following.*

(a) $\deg_\Gamma(v) = (1\pm\varepsilon)pn$ *for every* $v \in V(\Gamma)$.

(b) $e_\Gamma(A) \leqslant \max\{|A|^2p, 9n\}$ *for every* $A \subseteq V(\Gamma)$.

(c) $e_\Gamma(A,B) = (1\pm\varepsilon)p|A||B|$ *for every disjoint* $A, B \subseteq V(\Gamma)$ *with* $|A|, |B| \geqslant \frac{n}{M}$. *If on the other hand* $|A| < M^{-1}n$, *then* $e_\Gamma(A,B) \leqslant (1+\varepsilon)pM^{-1}n^2$.

(d) *For any* $A \subseteq V(\Gamma)$ *with* $|A| \geqslant \frac{n}{M}$ *all but at most* $10M\varepsilon^{-2}p^{-1}$ *vertices in* $V(\Gamma)$ *have* $(1\pm\varepsilon)p|A|$ *neighbours in* A.

Proof. These properties follow from standard applications of the Chernoff bound, Lemma 6. Here we only show (b); the other properties follow similarly.

Suppose that A is an arbitrarily chosen vertex subset. The expected number of edges in A is $\binom{|A|}{2}p \leqslant |A|^2p$. By Lemma 6 the probability that there are more than $|A|^2p$ edges in A is less than $\exp(\frac{-1}{3}\binom{|A|}{2}p) \leqslant \exp(\frac{-1}{7}|A|^2p)$. For $|A| \geqslant 3p^{-1/2}n^{1/2}$ this probability is less than $\exp(\frac{-9}{7}n)$ and so taking a union bound over all subsets the probability that Property (b) fails for a set of size at least $3p^{-1/2}n^{1/2}$ is less than $2^n\exp(\frac{-9}{7}n)$, which tends to zero. A set A with $|A| < 3p^{-1/2}n^{1/2}$ is less likely to have more than $9n$ edges than a set B with $|B| = 3p^{-1/2}n^{1/2} \leqslant n$. Therefore, since $|B|^2p = 9n$ and by the previous argument, the probability that a set A of size less than $3p^{-1/2}n^{1/2}$ has more than $9n$ edges tends to zero. □

The next lemma shows that for any partition $V\big(G(n,p)\big) = A \cup B$ with neither A nor B very small, most edges of $G(n,p)$ have 'typical' neighbourhoods in each set.

Lemma 15. *For any $0 < \varepsilon < \frac{1}{2}$, $M \in \mathbb{N}$ and $p = \omega\big(\frac{\ln n}{n}\big)$ in $\Gamma = G(n,p)$ a.a.s. for any two subsets A, B of $V(\Gamma)$ with $\frac{n}{M} \leqslant |A|, |B|$ all but at most $10^3 M\varepsilon^{-2}p^{-1}n$ edges uv in Γ satisfy all of the following:*

- $\deg_\Gamma(u, A), \deg_\Gamma(v, A) = (1 \pm \varepsilon)p|A|$.
- $\deg_\Gamma(u, B), \deg_\Gamma(v, B) = (1 \pm \varepsilon)p|B|$.
- $\deg_\Gamma(u, v, B) \geqslant (1 - \varepsilon)p^2|B|$.

Proof. By Lemma 14(d) we may assume that all but a set S of at most $20M\varepsilon^{-2}p^{-2}$ vertices in Γ have $(1 \pm \varepsilon)p|B|$ neighbours in B and $(1 \pm \varepsilon)p|A|$ neighbours in A. By Lemma 14(c) we further may assume that we have

$$e(S, A) \leqslant (1 + \varepsilon)p \cdot 20M\varepsilon^{-2}p^{-2}n = 20(1 + \varepsilon)M\varepsilon^{-2}p^{-1}n\,. \tag{5}$$

We now consider an arbitrary vertex v in $V \setminus S$ and two arbitrary sets $P, Q \subseteq N(v)$ satisfying $|P| \geqslant (1 - \frac{1}{2}\varepsilon)p|B|$ and $|Q| \geqslant 100M\varepsilon^{-2}p^{-1}$. The probability that all vertices in Q have fewer than $(1 - \varepsilon)p^2|B| \leqslant (1 - \frac{1}{2}\varepsilon)p|P|$ neighbours in P is less than

$$\exp\Big(-\frac{\varepsilon^2}{12}p|P||Q|\Big) \leqslant \exp\Big(-\frac{\varepsilon^2}{12}p \cdot \frac{1}{2}p\frac{n}{M} \cdot 100M\varepsilon^{-2}p^{-1}\Big) \leqslant \exp(-3pn)\,.$$

Since $P, Q \subseteq N(v)$ we have $|P|, |Q| \leqslant (1+\varepsilon)pn$. So, taking a union bound, the probability that there exist v, P, Q as above is less than $n2^{(1+\varepsilon)pn}2^{(1+\varepsilon)pn}\exp(-3pn)$ which tends to zero as n tends to infinity for $p = \omega(\log n/n)$. Hence a.a.s. each vertex v in $V \setminus S$ has at most $100M\varepsilon^{-2}p^{-1}$ neighbours u such that $\deg(u, v, B) < (1-\varepsilon)p^2|B|$. Summing over v we obtain at most $100M\varepsilon^{-2}p^{-1}n$ such edges, which along with the edges incident to S by (5) gives at most $10^3M\varepsilon^{-2}p^{-1}n$ edges. □

The following lemma is crucial in the proofs of Theorems 3 and 4. Before stating it we need some definitions. For any $s \in \mathbb{N}$, the *s-star* is the star $K_{1,s}$. The vertex of degree s in the s-star is called its *centre*, all other vertices are its *leaves*. For $A \subseteq V(\Gamma)$ and $0 < q, \varepsilon < 1$ we say that an s-star with centre x is *(q, ε)-bad* for A if there is $S \subseteq N_\Gamma(x, A)$ with $|S| \leqslant qp|A|$ such that each leaf y of the s-star satisfies $\deg_\Gamma(y, S) \geqslant (1 + \varepsilon)qp^2|A|$; in other words y has substantially more neighbours in S than expected. We also say that S *witnesses* this badness.

When we use this definition, we will choose a star with centre x and set $S = N_\Gamma(x, A) \setminus N_H(x, A)$, where H is a triangle-free subgraph of Γ with large minimum degree, and we will choose our star such that that $N_\Gamma(y, S)$ is quite large for each leaf y. Now if the star is good it follows that S itself must be quite large, so that the degree of x in H cannot be too large, leading to a contradiction to the minimum degree of H. The following lemma however implies that bad stars cover only $\mathcal{O}(p^{-1}n)$ edges, which is where the sharp bounds in Theorems 3 and 4 come from.

Lemma 16. *For every $0 < \varepsilon < 1$ and every p the random graph $G(n,p)$ a.a.s. satisfies the following. For every $A \subseteq V(\Gamma)$ with $\frac{n}{3} \leqslant |A|$, every q with $\varepsilon < q < 1$, and every*

$s \geqslant 100q^{-1}\varepsilon^{-2}p^{-1}$ *there are fewer than* $\frac{1}{2}p^{-1}$ *vertex disjoint* s*-stars in* $V(\Gamma) \setminus A$ *which are* (q,ε)*-bad for* A.

Proof. First let A be fixed. Consider an s-star with centre x and a set $S \subseteq N_\Gamma(x, A)$ with $|S| \leqslant qp|A|$. By the Chernoff bound, Lemma 6, the probability that S witnesses that this star is (q,ε)-bad for A is less than $\exp\big(\frac{-\varepsilon^2}{3} \cdot qp^2|A|s\big)$. Observe that $|S| \leqslant qp|A| \leqslant pn$ and that we may assume $s \leqslant \deg_\Gamma(x) \leqslant 2pn$ by Lemma 14(a). So by taking a union bound over choices of S for a single s-star, and then considering collections of $\frac{1}{2}p^{-1}$ vertex disjoint s-stars, and taking another union bound over all such collections, we obtain that the probability that there are at least $\frac{1}{2}p^{-1}$ disjoint (q,ε)-bad stars for A in $V(\Gamma) \setminus A$ is less than

$$\Big(n \cdot 2^{2pn}\Big)^{\frac{1}{2}p^{-1}} \cdot \Big(2^{pn} \exp\big(\tfrac{-\varepsilon^2}{3} qp^2|A|s\big)\Big)^{\frac{1}{2}p^{-1}} \leqslant \big(2^{4pn} \exp\big(\tfrac{-\varepsilon^2}{9} qp^2 ns\big)\big)^{\frac{1}{2}p^{-1}} .$$

By taking a union bound over choices of A we find that the probability that there is A such that $\frac{1}{2}p^{-1}$ stars $K_{1,s}$ outside A are (q,ε)-bad for A is less than

$$2^n \big(2^{4pn} \exp\big(\tfrac{-\varepsilon^2}{9} qp^2 ns\big)\big)^{\frac{1}{2}p^{-1}} \leqslant \exp\big(n + 2n - \tfrac{\varepsilon^2}{18} qpns\big) ,$$

which tends to zero for $s \geqslant 100\varepsilon^{-2}q^{-1}p^{-1}$. (Observe that we do not have to take a union bound over s, because for $s' > s$ any s-star which is a subgraph of a (q,ε)-bad s'-star is also (q,ε)-bad.) □

5. Proof of Theorem 3

Recall that Theorem 3 states the following.

Theorem 3. *For any* $\gamma > 0$*, there exists* C *such that for any* $p(n)$ *the random graph* $\Gamma = G(n,p)$ *a.a.s. has the property that all triangle-free spanning subgraphs* $H \subseteq \Gamma$ *with* $\delta(H) \geqslant (\frac{2}{5} + \gamma)pn$ *can be made bipartite by removing at most* $\min\big(Cp^{-1}n, (\frac{1}{4} + \gamma)pn^2\big)$ *edges.*

The main strategy of the proof is as follows. We first apply Lemma 10 (which is a consequence of the Sparse Regularity Lemma) to H to obtain a dense triangle-free reduced graph R of H with minimum degree above $\frac{2}{5}v(R)$, which by the Andrásfai–Erdős–Sós Theorem, Theorem 1, is bipartite. We conclude that H can be made bipartite by removing $o(pn^2)$ edges. Hence in a maximum cut $X \cup Y$ of H we have $e_H(X), e_H(Y) = o(pn^2)$. Our goal will then be to improve this bound on $e_H(X)$ and $e_H(Y)$ by distinguishing between 'typical' and 'atypical' edges in these sets and applying the results established in the previous section to count these, using that $X \cup Y$ is a maximum cut and that H is triangle-free.

Proof of Theorem 3. Let

$$\varepsilon = \frac{\gamma^4}{10^7}, \quad d = \frac{\gamma^2}{10^3}, \quad \eta = d + 3\varepsilon, \quad \beta = \frac{2}{5} + \gamma, \quad t_0 = \frac{1}{\varepsilon} \tag{6}$$

and let c and t_1 be the values attained by applying Lemma 10 with inputs ε, d, β and t_0. Let $M = t_1^2$, and let

$$C = \max\left(10^{10}\varepsilon^{-2}, c^2\right). \tag{7}$$

We first consider the easy case that p is small. If $p \leqslant n^{-7/4}$, then the expected number of paths with two edges in $G(n,p)$ is at most $p^2n^3 \leqslant n^{-1/2}$. In particular a.a.s there are no such paths, so a.a.s. $G(n,p)$ is bipartite and the statement of Theorem 3 holds trivially. We may therefore assume $p \geqslant n^{-7/4}$, so by Lemma 6 a.a.s. $G(n,p)$ has at most $\left(\frac{1}{2}+\gamma\right)pn^2$ edges. Now if G is any graph with at most $\left(\frac{1}{2}+2\gamma\right)pn^2$ edges, then we can make G bipartite by removing all the edges of G not in a maximum cut. Since a maximum cut of G contains at least half its edges, we remove at most $\left(\frac{1}{4}+\gamma\right)pn^2$ edges. Again, if $\min\left(Cp^{-1}n, (\frac{1}{4}+\gamma)pn^2\right) = (\frac{1}{4}+\gamma)pn^2$, which occurs when $p \leqslant cn^{-1/2}$, the statement of Theorem 3 follows.

It remains to consider the hard case that $p \geqslant cn^{-1/2}$. We now assume $\Gamma = G(n,p)$ satisfies the properties stated in Lemma 14 with input ε and M, Lemma 15 with input ε and M, Lemma 16 with input ε and Lemma 10 for the parameters given above.

Consider any triangle-free $H \subseteq \Gamma$ with $\delta(H) \geqslant (\frac{2}{5}+\gamma)pn$ and let $X \cup Y$ be a maximum cut of the vertex set of H. Assume without loss of generality that $e_H(X) \geqslant e_H(Y)$. Our goal is to show $e_H(X) \leqslant \frac{1}{2}Cp^{-1}n$. We start with the following observation.

Claim 17. *$e_H(X) \leqslant \eta pn^2$.*

Proof of Claim 17. By the property asserted by Lemma 10 we obtain an $\left(\varepsilon, d, p\right)$-regular partition $V(\Gamma) = V_0 \cup V_1 \cup \cdots \cup V_t$ of H with $t_0 \leqslant t \leqslant t_1$ whose corresponding reduced graph R is triangle-free and has minimum degree at least $(\frac{2}{5}+\gamma-d-\varepsilon)v(R) > \frac{2}{5}v(R)$. Therefore, by the Andrásfai–Erdős–Sós Theorem, Theorem 1, R is bipartite.

By Lemma 14(a) at most $\varepsilon n(1+\varepsilon)pn$ edges have at least one end in V_0. Moreover, since at most an ε-fraction of all pairs are irregular, by Lemma 14(c) at most $\varepsilon(1+\varepsilon)pn^2$ edges are contained in irregular pairs. Finally, at most dpn^2 edges are in pairs with density less than d. We conclude that at most $(d+2(1+\varepsilon)\varepsilon)pn^2 \leqslant \eta pn^2$ edges of H do not lie in pairs corresponding to edges of R, which proves the claim. □

We next bound the sizes of X and Y.

Claim 18. *$\left(\frac{2}{5}+\frac{1}{2}\gamma\right)n \leqslant |X|, |Y| \leqslant \left(\frac{3}{5}-\frac{1}{2}\gamma\right)n$.*

Proof of Claim 18. Suppose for a contradiction that X satisfies $|X| > (\frac{3}{5}-\frac{1}{2}\gamma)n$ and hence $|Y| < (\frac{2}{5}+\frac{1}{2}\gamma)$. Then by Lemma 14$(c)$ we see that $e_H(X,Y) \leqslant e_\Gamma(X,Y) \leqslant (1+\varepsilon)(\frac{3}{5}-\frac{1}{2}\gamma)(\frac{2}{5}+\frac{1}{2}\gamma)pn^2$.

On the other hand, by our minimum degree condition $2e_H(X)+e_H(X,Y) \geqslant (\frac{2}{5}+\gamma)pn|X|$, and similarly $2e_H(Y)+e_H(X,Y) \geqslant (\frac{2}{5}+\gamma)pn|Y|$. Since $e_H(X), e_H(Y) \leqslant \eta pn^2$ this gives $e_H(X,Y) \geqslant (\frac{2}{5}+\gamma)pn \cdot \max\{|X|,|Y|\} - 2\eta pn^2$. Since $\max\{|X|,|Y|\} \geqslant (\frac{3}{5}-\frac{1}{2}\gamma)n$ we obtain $e_H(X,Y) \geqslant \left((\frac{3}{5}-\frac{1}{2}\gamma)(\frac{2}{5}+\gamma)-2\eta\right)pn^2$, a contradiction.

So $|X| \leqslant (\frac{3}{5}-\frac{1}{2}\gamma)n$, and analogously $|Y| \leqslant (\frac{3}{5}-\frac{1}{2}\gamma)n$, proving the claim. □

We next define

$$\tilde{X} = \left\{x \in X : \deg_H(x, X) \geqslant \gamma \cdot \deg_H(x)\right\},$$

a set of vertices with high degree in X, which require special treatment later on. The next claim shows that $\tilde{X}$ is small and contains at most half of the edges in X.

Claim 19. *$|\tilde{X}| \leqslant \frac{1}{100}\gamma n$, and if $e_H(X) > \frac{1}{2}Cp^{-1}n$ then $e_H(\tilde{X}) \leqslant \frac{1}{2}e_H(X)$.*

Proof of Claim 19. By Claim 17 and the definition of $\tilde{X}$ we have

$$\eta p n^2 \geqslant e_H(X) \geqslant \frac{1}{2}|\tilde{X}|\gamma\delta(H) \geqslant \frac{\gamma}{2}\Big(\frac{2}{5} + \gamma\Big)pn|\tilde{X}|\,, \tag{8}$$

hence $|\tilde{X}| \leqslant \frac{2\eta n}{\gamma(2/5+\gamma)} \leqslant 5\gamma^{-1}\eta n \leqslant \gamma n/100$ by (6).

For the second part of the claim assume that $e_H(X) > \frac{1}{2}Cp^{-1}n$. By Lemma 14(*b*) we have $e_H(\tilde{X}) \leqslant e_\Gamma(\tilde{X}) \leqslant \max\{|\tilde{X}|^2p, 9n\}$. If this maximum is attained by $9n$, then we are done because $9n \leqslant \frac{1}{4}Cp^{-1}n < \frac{1}{2}e_H(X)$. Otherwise $e_H(\tilde{X}) \leqslant |\tilde{X}|^2p$, and since $|\tilde{X}| \leqslant \frac{1}{100}\gamma n$, we have

$$|\tilde{X}|^2p \leqslant \frac{1}{100}\gamma pn|\tilde{X}| \leqslant \frac{\gamma}{4}\Big(\frac{2}{5} + \gamma\Big)pn|\tilde{X}| \overset{(8)}{\leqslant} \frac{1}{2}e_H(X)\,,$$

and we are also done. □

We continue by removing 'atypical' edges from H. Let H' be the graph obtained from H by removing edges from $E_H(X)$ which do not satisfy the conditions of Lemma 15 with respect to the partition $X \cup Y$. We also remove the edges in $E_H(\tilde{X})$. By Lemma 15 and Claim 19 we have $e_H(X) \leqslant \frac{1}{2}Cp^{-1}n$ or

$$e_H(X) - e_{H'}(X) \leqslant 10^3\varepsilon^{-2}p^{-1}n + \frac{1}{2}e_H(X) \overset{(7)}{\leqslant} \frac{1}{10}Cp^{-1}n + \frac{1}{2}e_H(X)\,. \tag{9}$$

Our goal in the remainder is to bound the number of H'-edges in X.

Let xz be any H'-edge in X. We have

$$\deg_\Gamma(x, z, Y) \geqslant (1-\varepsilon)p^2|Y| \tag{10}$$

by construction of H', so this common neighbourhood constitutes many Γ-triangles xzy, for each of which either xy or zy is not present in H'. We now would like to direct the edges in X according which of these two cases is more common – however, it turns out that we need to favour vertices not in $\tilde{X}$ in this process; so we direct with a bias.

More precisely, for any H'-edge in X, if one of its vertices is in $\tilde{X}$ call it x, otherwise let x be any vertex of the edge. Let x' be the other vertex of the edge. We direct xx' towards x if

$$|N_\Gamma(x, x', Y) \setminus N_{H'}(x, Y)| \geqslant \frac{2}{3}\deg_\Gamma(x, x', Y)\,,$$

that is if many edges from x to $N_\Gamma(x, x', Y)$ were deleted. We direct xx' towards x' otherwise, in which case we have

$$|N_\Gamma(x, x', Y) \setminus N_{H'}(x', Y)| > \frac{1}{3}\deg_\Gamma(x, x', Y)\,,$$

An *s-in-star* in this directed graph is an s-star such that all edges are directed towards the centre. Recall that an s-star with centre x is (q,ε)-bad for Y if there is a witness $S \subseteq N_\Gamma(x,Y)$ with $|S| \leqslant qp|Y|$ such that each leaf z of the s-star satisfies $\deg_\Gamma(z,S) \geqslant (1+\varepsilon)qp^2|Y|$. The next claim shows that in-stars in $H'[X]$ are bad. We define

$$s = 10^3\varepsilon^{-2}p^{-1}\,, \quad \tilde{q} = (1-2\varepsilon)\frac{2}{3}\,, \quad q = (1-2\varepsilon)\frac{1}{3}\,.$$

Claim 20. *Each s-in-star in $H'[X]$ with centre $x \in \tilde{X}$ is $(\tilde{q},\varepsilon)$-bad for Y, and each s-in-star in $H'[X]$ with centre $x \notin \tilde{X}$ is (q,ε)-bad for Y.*

Proof of Claim 20. First assume F is an s-in-star with centre $x \in \tilde{X}$ which is not $(\tilde{q},\varepsilon)$-bad. We first show that this implies

$$|N_\Gamma(x,Y) \setminus N_{H'}(x,Y)| > \tilde{q}p|Y|\,. \tag{11}$$

Indeed, assume otherwise. Then, since F is not $(\tilde{q},\varepsilon)$-bad for Y we have for $S = N_\Gamma(x,Y) \setminus N_{H'}(x,Y)$ that there is a leaf z of F such that

$$|N_\Gamma(x,z,Y) \setminus N_{H'}(x,Y)| = \deg_\Gamma(z,S) < (1+\varepsilon)\tilde{q}p^2|Y| \leqslant \frac{2}{3}(1-\varepsilon)p^2|Y|\,.$$

This however contradicts the fact that F is an in-star and thus

$$|N_\Gamma(x,z,Y) \setminus N_{H'}(x,Y)| \geqslant \frac{2}{3}\deg_\Gamma(x,z,Y) \overset{(10)}{\geqslant} \frac{2}{3}(1-\varepsilon)p^2|Y|\,.$$

Accordingly (11) holds.

Since $\deg_H(x,Y) = \deg_{H'}(x,Y)$ we conclude that

$$\deg_H(x,Y) \leqslant \deg_\Gamma(x,Y) - \tilde{q}p|Y| \leqslant (1+\varepsilon)p|Y| - (1-2\varepsilon)\frac{2}{3}p|Y| \leqslant \Big(\frac{1}{3}+3\varepsilon\Big)p|Y|\,.$$

Because $X \cup Y$ is a maximum cut this implies by Claim 18 that

$$\deg_H(x) \leqslant 2\Big(\frac{1}{3}+3\varepsilon\Big)p\Big(\frac{3}{5}-\frac{1}{2}\gamma\Big)n < \Big(\frac{2}{5}+\gamma\Big)pn\,,$$

contradicting the minimum degree of H.

For the second part of the claim assume that F is an s-in-star with centre $x \notin \tilde{X}$ which is not (q,ε)-bad. By similar logic to the proof of (11), this implies that

$$|N_\Gamma(x,Y) \setminus N_{H'}(x,Y)| > qp|Y|$$

by using that for any leaf z of F we have $|N_\Gamma(x,z,Y) \setminus N_{H'}(x,Y)| > \frac{1}{3}\deg_\Gamma(x,z,Y)$. Also analogously, this implies that $\deg_H(x,Y) \leqslant (\frac{2}{3}+3\varepsilon)p|Y|$. Recall that $x \notin \tilde{X}$ means that $\deg_H(x,X) < \gamma\deg_H(x)$ and hence $\deg_H(x) \leqslant \frac{1}{1-\gamma}\deg_H(x,Y) \leqslant (1+2\gamma)\deg_H(x,Y)$. Thus, by Claim 18,

$$\deg_H(x) \leqslant (1+2\gamma)\Big(\frac{2}{3}+3\varepsilon\Big)p\Big(\frac{3}{5}-\frac{1}{2}\gamma\Big)n \leqslant \Big(\frac{2}{3}+\frac{5}{3}\gamma\Big)p\Big(\frac{3}{5}-\frac{1}{2}\gamma\Big)n < \Big(\frac{2}{5}+\gamma\Big)pn\,,$$

again contradicting the minimum degree of H. □

By Lemma 16, however, the number of s-stars in Γ which are either $(\tilde{q},\varepsilon)$-bad or (q,ε)-bad is less than p^{-1}. So Claim 20 implies that the number of s-in-stars in $H'[X]$ is less than p^{-1}. The following claim shows that this implies that $e_{H'}(X)$ is small.

Claim 21. *$e_{H'}(X) \leqslant \frac{1}{10}Cp^{-1}n$.*

Proof of Claim 21. Assume for a contradiction that $e_{H'}(X) > \frac{1}{10}Cp^{-1}n \geqslant 10^4\varepsilon^{-2}p^{-1}n$. Using a greedy argument, we will show that we then can find more than p^{-1} stars in $H'[X]$ which are s-in-stars (with $s = 10^3\varepsilon^{-2}p^{-1}$). Indeed, the average in-degree is at least $10^4\varepsilon^{-2}p^{-1}$, so we can find at least one $(10^3\varepsilon^{-2}p^{-1})$-in-star. If we remove from $H'[X]$ this star and all edges adjacent to it this accounts for at most $(1+s)(1+\varepsilon)pn \leqslant 2spn$ edges. So we can repeat this process p^{-1} times, after which at most $2sn = 2\cdot 10^3\varepsilon^{-2}p^{-1}n$ edges have been deleted from $H'[X]$, hence $H[X]$ still contains more than $10^3\varepsilon^{-2}p^{-1}n$ edges in X, still giving an average in-degree of at least $10^3\varepsilon^{-2}p^{-1}$, and hence we can find another $(10^3\varepsilon^{-2}p^{-1})$-in-star, which is the desired contradiction. □

Now (9) and Claim 21 imply $e_H(Y) \leqslant e_H(X) \leqslant \frac{1}{2}Cp^{-1}n$, hence H can be made bipartite by removing at most $Cp^{-1}n$ edges as claimed. □

6. Proof of Theorem 4

The proof of Theorem 4 adds the techniques developed for the proof of Theorem 3 to ideas used in [2, 16]. Our strategy is as follows. Given a subgraph H of $\Gamma = G(n,p)$ with $\delta(H) \geqslant \big(\frac{1}{3}+\gamma\big)pn$, we will apply the sparse regularity lemma to obtain a regular partition $V(H) = V_0 \cup \dots \cup V_t$ with (ε,d,p)-reduced graph R. We let W be the set of all vertices whose degree to some set V_i is far from the expected $p|V_i|$, and then for each $I \subseteq [t]$ we let N_I be the subset of vertices in $V(H)\setminus W$ with many neighbours in exactly the clusters $\{V_i : i \in I\}$, which gives a partition of $V(H)$ into 2^t+1 sets. We will show that there are $O(p^{-1}n)$ edges in W and in each N_I, hence we can remove all such edges to obtain a graph with bounded chromatic number. We do this by showing that W is too small to contain many edges, and that the same is true for any N_I such that $R[I]$ contains an edge. If on the other hand $R[I]$ is independent, we use an argument similar to that in the proof of Theorem 3.

Proof of Theorem 4. Given $\gamma > 0$, let

$$d = \frac{\gamma}{20}\,, \quad \varepsilon' = \frac{d^3}{30}\,, \quad \beta = \frac{1}{3}+\gamma\,, \quad t_0 = \frac{1}{\varepsilon'}\,. \tag{12}$$

Let ε_0, C_{L9} be the outputs if Lemma 9 is applied with ε' and d. We take $\varepsilon = \min\{\varepsilon_0,\varepsilon'\}$ and let t_1 be the output if Lemma 7 is applied with β, ε and t_0. We require as well that $t_1 \geqslant 10$. We choose $c = 2C_{\mathrm{L9}}t_1$ (which is needed for the application of Lemma 9). Finally we choose

$$M = 2t_1\,, \quad r = 2^{t_1}+1\,, \quad C' = 10^4\cdot 2^{10t_1}\varepsilon^{-3}\,, \quad C = \max(rC'^2, c^2)\,. \tag{13}$$

As in the proof of Theorem 3, if $p \leqslant n^{-7/4}$ a.a.s. $G(n,p)$ is bipartite and the statement is trivially true, while for any graph G a maximum r-partition of G contains at least $\frac{r-1}{r}e(G)$

edges, so that when $p \geqslant n^{-7/4}$ a.a.s. we can make any subgraph of $G(n,p)$ r-partite by deleting at most $\big(\frac{1}{2r}+\gamma\big)pn^2$ edges. Again, this leaves the hard case when $p \geqslant cn^{-1/2}$.

Now sample $\Gamma = G(n,p)$. Since $p > cn^{-1/2} = \omega(\frac{\ln n}{n})$ we can assume that Γ satisfies the properties of Lemmas 7, 14, 15, and 16 with the parameters chosen above.

Let H be a triangle-free spanning subgraph of Γ with $\delta(H) \geqslant \big(\frac{1}{3}+\gamma\big)n$. By Lemma 7 there is an (ε,d,p)-regular partition $V_0 \cup V_1 \cup \cdots \cup V_t$ of H with $t \leqslant t_1$ such that the reduced graph R has $\delta(R) \geqslant \big(\frac{1}{3}+\gamma-d-3\varepsilon\big)v(R) \geqslant \big(\frac{1}{3}+\frac{\gamma}{2}\big)v(R)$, and such that for each i and each $v \in V_i$, the vertex v has at most $(d+\varepsilon)pn$ neighbours in $\bigcup_{j:ij\notin R} V_j$.

Let W consist of all vertices which either have more than $(1+\varepsilon)p|V_i|$ neighbours in V_i for some i, or more than $2\varepsilon pn$ neighbours in V_0. By Lemma 14(d) we have $|W| \leqslant 10M(t+1)\varepsilon^{-2}p^{-1}$, and by Lemma 14$(b)$ the number of edges in W is therefore at most $\max\big(100M^2(t+1)^2\varepsilon^{-4}p^{-1}, 9n\big) \leqslant 10p^{-1}n$, where the inequality holds for all sufficiently large n. Now for each $I \subseteq [t]$, let N_I be the set of vertices of H with many neighbours exactly in the clusters V_i with $i \in I$, that is,

$$N_I = \{v \in V(H)\colon |N(v) \cap V_i| > 10dp|V_i| \text{ if and only if } i \in I\}\,.$$

Claim 22. *$\{N_I\colon |I| > \frac{t}{3}\}$ partitions $V(H) \setminus W$.*

Proof. The sets $\{N_I : I \subseteq [t]\}$ are disjoint and partition $V(H)\setminus W$ by definition. If $|I| \leqslant \frac{t}{3}$ then any vertex $v \in N_I$ has at most $\sum_{i\in I}(1+\varepsilon)p|V_i| + \sum_{i\notin I} 10dp|V_i| + 2\varepsilon pn < \big(\frac{1}{3}+\gamma\big)pn$ neighbours since $v \notin W$ and by definition of N_I, which is a contradiction, so $N_I = \emptyset$ if $|I| \leqslant \frac{t}{3}$. □

Our goal is thus to show that $e_H(N_I) \leqslant C'^2p^{-1}n$ for any I with $|I| > \frac{t}{3}$, since this implies that H can be made r-partite with $r = 2^{t_1+1}$ by removing at most $rC'^2p^{-1}n \leqslant Cp^{-1}n$ edges. This is established by the following two claims.

Claim 23. *If $R[I]$ contains an edge, then $e_H(N_I) \leqslant C'^2p^{-1}n$.*

Proof of Claim 23. Suppose that $ij \in R[I]$. If $v \in N_I$ is such that $\big(N_\Gamma(v,V_i), N_\Gamma(v,V_j)\big)$ is (ε',d,p)-lower-regular in H. Since $v \notin W$, the pair $\big(N_H(v,V_i), N_H(v,V_j)\big)$ is $\big(\varepsilon'\frac{1+\varepsilon}{10d},d,p\big)$-lower-regular in H. Since $d > \varepsilon'\frac{1+\varepsilon}{10d}$, there is an edge of H in this latter pair and hence H contains a triangle, a contradiction.

We conclude that there are no such vertices in N_I, so by Lemma 9 we have $|N_I| \leqslant C' \max\big(p^{-2}, p^{-1}\log n\big)$. By Lemma 14$(b)$ the number of edges in N_I is therefore at most $\max\big(C'^2p^{-3}, C'^2p^{-1}\log^2 n, 9n\big) \leqslant C'^2p^{-1}n$ by choice of p and C'. □

Claim 24. *If $R[I]$ is independent, then $e_H(N_I) \leqslant C'p^{-1}n$.*

Proof of Claim 24. Since $\delta(R) \geqslant \big(\frac{1}{3}+\frac{\gamma}{2}\big)t$, if $R[I]$ is independent then $|I| < \frac{2t}{3}$. Let $S_I := \bigcup_{i\in I} V_i$. We first show that S_I and N_I are disjoint. Indeed, if $v \in N_i$ were in some V_i with $i \in I$, then by definition of N_I the vertex v has at least $\sum_{j\in I} 10dp|V_j| \geqslant 5dpn/3$ neighbours in $\bigcup_{j\in I} V_j$, where the inequality follows since $|I| > t/3$. Since ij is not an edge of R for any $j \in I$, this is in contradiction to the guarantee that v has at most $(d+\varepsilon)pn$ neighbours in $\bigcup_{j:ij\notin R} V_j$.

We now delete some 'atypical' edges from $H[N_I]$. Remove from $H[N_I]$ each edge uv with $\deg_\Gamma(u,v,S_I) < (1-\varepsilon)|S_I|p^2$. to obtain the graph H'. By Lemma 15 this accounts for at most $10^3 \cdot 4\varepsilon^{-2}p^{-1}n \leqslant \frac{\varepsilon}{10}C'p^{-1}n$ edges.

Let Z be the set of vertices $v \in N_I$ such that $\deg_H(v) - \deg_{H'}(v) \geqslant \varepsilon pn$. By double counting we have $|Z| \leqslant \frac{\varepsilon C'p^{-1}n}{5\varepsilon pn} = \frac{1}{5}C'p^{-2}$.

We now proceed similarly as in the proof of Theorem 3. We orient the edges uv in $H'[N_I]$ towards u if $|N_\Gamma(u,v,S_I) \setminus N_{H'}(u,S_I)| \geqslant \frac{1}{2}\deg_\Gamma(u,v,S_I)$ and towards v otherwise. Again, for $s = 10^3q^{-1}\varepsilon^{-2}p^{-1}$ and $q = (1-2\varepsilon)\frac{1}{2}$ any s-in-star with centre x not in Z is (q,ε)-bad with respect to S_I. Indeed, otherwise, analogously to the proof of (11), we have $|N_\Gamma(x,S_I) \setminus N_{H'}(x,S_I)| > qp|S_I|$, which implies

$$\deg_{H'}(x,S_I) < (1+\varepsilon)p|S_I| - qp|S_I| = \frac{1}{2}p|S_I| \leqslant \frac{1}{2}p\frac{2}{3}n = \frac{1}{3}pn$$

Since $x \notin Z$, we have $\deg_H(x) \leqslant \deg_{H'}(x) + \varepsilon pn < \left(\frac{1}{3}+\gamma\right)pn$, a contradiction.

We now pick greedily vertex disjoint s-in-stars whose centres are not in Z until no more remain. By Lemma 16, since S_I and N_I are disjoint, this process terminates having found less than $\frac{1}{2}p^{-1}$ such stars. Let Y be the set of vertices contained in all these stars; then $|Y| \leqslant \frac{1}{2}p^{-1}s \leqslant 10^3q^{-1}\varepsilon^{-2}p^{-2}$. Now $e_{H'}\big(N_I \setminus (Y \cup Z)\big) \leqslant s|N_I|$ since $N_I \setminus (Y \cup Z)$ contains no s-in-star, so we conclude

$$e_H(N_I) \leqslant (1+\varepsilon)pn|Y \cup Z| + s|N_I| + \tfrac{1}{10}C'p^{-1}n \leqslant C'p^{-1}n\,,$$

as desired. □

Finally, these claims show that deleting all edges internal to any of the sets W and N_I for $I \subseteq [t]$ yields a $2^t + 1 = r$-partite graph, and that the number of edges deleted is at most $Cp^{-1}n$, as desired. □

References

1. P. Allen, J. Böttcher, J. Ehrenmüller, and A. Taraz, *The Bandwidth Theorem in sparse graphs*, submitted (arXiv:1612.00661), 2016.
2. P. Allen, J. Böttcher, S. Griffiths, Y. Kohayakawa, and R. Morris, *The chromatic thresholds of graphs*, Adv. Math. **235** (2013), 261–295.
3. P. Allen, J. Böttcher, H. Hàn, Y. Kohayakawa, and Y. Person, *Blow-up lemmas for sparse graphs*, submitted (arXiv:1612.00622), 2016.
4. B. Andrásfai, P. Erdős, and V. T. Sós, *On the connection between chromatic number, maximal clique and minimal degree of a graph*, Discrete Math. **8** (1974), 205–218.
5. J. Balogh, R. Morris, and W. Samotij, *Independent sets in hypergraphs*, J. Amer. Math. Soc., to appear.
6. Blanche Descartes, *A three-colour problem*, Eureka **9** (April 1947), 21, Solution March 1948.
7. S. Brandt and S. Thomassé, *Dense triangle-free graphs are four colorable: A solution to the Erdős-Simonovits problem*, To appear in J. Combin. Theory Ser. B.
8. D. Conlon, *Combinatorial theorems relative to a random set*, Proceedings of the 2014 ICM, to appear.
9. D. Conlon, T. Gowers, W. Samotij, and M. Schacht, *On the KŁR conjecture in random graphs*, Israel J. Math, to appear, 2013.
10. B. DeMarco and J. Kahn, *Mantel's theorem for random graphs*, arXiv:1206.1016, 2012.

11. P. Erdős and M. Simonovits, *On a valence problem in extremal graph theory*, Discrete Math. **5** (1973), 323–334.
12. S. Janson, T. Łuczak, and A. Ruciński, *Random graphs*, Wiley-Interscience, New York, 2000.
13. Y. Kohayakawa, *Szemerédi's regularity lemma for sparse graphs*, Foundations of computational mathematics (Rio de Janeiro, 1997), Springer, Berlin, 1997, pp. 216–230.
14. Y. Kohayakawa, T. Łuczak, and V. Rödl, *On K^4-free subgraphs of random graphs*, Combinatorica **17** (1997), no. 2, 173–213.
15. Y. Kohayakawa, V. Rödl, M. Schacht, and E. Szemerédi, *Sparse partition universal graphs for graphs of bounded degree*, Adv. Math. **226** (2011), no. 6, 5041–5065.
16. J. Lyle, *On the chromatic number of H-free graphs of large minimum degree*, Graphs and Combinatorics (2010), 1–14.
17. W. Mantel, *Problem 28*, Wiskundige Opgaven **10** (1907), 60–61, (solution by H. Gouwentak, W. Mantel, J. Texeira de Mattes, F. Schuh and W.A. Wythoff).
18. D. Saxton and A. Thomason, *Hypergraph containers*, Invent. Math. **201** (2015), no. 3, 925–992.
19. C. Thomassen, *On the chromatic number of triangle-free graphs of large minimum degree*, Combinatorica **22** (2002), no. 4, 591–596.

Department of Mathematics, The London School of Economics and Political Science, Houghton Street, London WC2A 2AE, U.K.
E-mail address: p.d.allen|j.boettcher|b.j.roberts@lse.ac.uk

Instituto de Matemática e Estatística, Universidade de São Paulo, Rua do Matão 1010, 05508–090 São Paulo, Brazil.
E-mail address: yoshi@ime.usp.br