

# On the KLR conjecture in random graphs

D. Conlon\*      W. T. Gowers†      W. Samotij‡      M. Schacht§

## Abstract

The KLR conjecture of Kohayakawa, Łuczak, and Rödl is a statement that allows one to prove that asymptotically almost surely all subgraphs of the random graph  $G_{n,p}$ , for sufficiently large  $p := p(n)$ , satisfy an embedding lemma which complements the sparse regularity lemma of Kohayakawa and Rödl. We prove a variant of this conjecture which is sufficient for most known applications to random graphs. In particular, our result implies a number of recent probabilistic versions, due to Conlon, Gowers, and Schacht, of classical extremal combinatorial theorems. We also discuss several further applications.

## 1 Introduction

Szemerédi’s regularity lemma [83], which played a crucial role in Szemerédi’s proof of the Erdős-Turán conjecture [82] on long arithmetic progressions in dense subsets of the integers, is one of the most important tools in extremal graph theory (see [57, 58, 73]). Roughly speaking, it says that the vertex set of every graph  $G$  may be divided into a bounded number of parts in such a way that most of the induced bipartite graphs between different parts are pseudorandom.

More precisely, a bipartite graph between sets  $U$  and  $V$  is said to be  $\epsilon$ -regular if, for every  $U' \subseteq U$  and  $V' \subseteq V$  with  $|U'| \geq \epsilon|U|$  and  $|V'| \geq \epsilon|V|$ , the density  $d(U', V')$  of edges between  $U'$  and  $V'$  satisfies

$$|d(U', V') - d(U, V)| \leq \epsilon.$$

We will say that a partition of the vertex set of a graph into  $t$  pieces  $V_1, \dots, V_t$  is an equipartition if, for every  $1 \leq i, j \leq t$ , we have the condition that  $||V_i| - |V_j|| \leq 1$ . We say that the partition is  $\epsilon$ -regular if it is an equipartition and, for all but at most  $\epsilon t^2$  pairs  $(V_i, V_j)$ , the induced graph between  $V_i$  and  $V_j$  is  $\epsilon$ -regular. Szemerédi’s regularity lemma can be formally stated as follows.

**Theorem 1.1.** *For every  $\epsilon > 0$  and every positive integer  $t_0$ , there exists a positive integer  $T$  such that every graph  $G$  with at least  $t_0$  vertices admits an  $\epsilon$ -regular partition  $V_1, \dots, V_t$  of its vertex set into  $t_0 \leq t \leq T$  pieces.*

Often the strength of the regularity lemma lies in the fact that it may be combined with a counting or embedding lemma that tells us approximately how many copies of a particular subgraph

---

\*Mathematical Institute, Oxford OX1 3LB, United Kingdom. E-mail: david.conlon@maths.ox.ac.uk. Research supported by a Royal Society University Research Fellowship.

†Department of Pure Mathematics and Mathematical Statistics, Wilberforce Road, Cambridge CB3 0WB, UK. E-mail: w.t.gowers@dpmmms.cam.ac.uk. Research supported by a Royal Society 2010 Anniversary Research Professorship.

‡School of Mathematical Sciences, Tel Aviv University, Tel Aviv, Israel; and Trinity College, Cambridge CB2 1TQ, UK. E-mail: samotij@post.tau.ac.il. Research supported in part by a Trinity College JRF.

§Fachbereich Mathematik, Universität Hamburg, Bundesstraße 55, D-20146 Hamburg, Germany. E-mail: schacht@math.uni-hamburg.de. Research supported by the Heisenberg programme of the DFG.

a graph contains, in terms of the densities  $d(V_i, V_j)$  arising in an  $\epsilon$ -regular partition. The so-called regularity method usually works as follows. First, one applies the regularity lemma to a graph  $G$ . Next, one defines an auxiliary graph  $R$  whose vertices are the parts of the regular partition of  $G$  one obtains, and whose edges correspond to regular pairs with non-negligible density. (For some applications we may instead take a weighted graph, where the weight of the edge between a regular pair  $V_i$  and  $V_j$  is the density  $d(V_i, V_j)$ .) If one can then find a copy of a particular subgraph  $H$  in  $R$ , the counting lemma allows one to find many copies of  $H$  in  $G$ . If  $R$  does not contain a copy of  $H$ , this information can often be used to deduce some further structural properties of the graph  $R$  and, thereby, the original graph  $G$ . Applied in this manner, the regularity and counting lemmas allow one to prove a number of well-known theorems in extremal graph theory, including the Erdős-Stone theorem [21], its stability version due to Erdős and Simonovits [81], and the graph removal lemma [4, 24, 29, 76].

For sparse graphs – that is, graphs with  $n$  vertices and  $o(n^2)$  edges – the regularity lemma stated in Theorem 1.1 is vacuous, since every equipartition into a bounded number of parts is  $\epsilon$ -regular for  $n$  sufficiently large. It was observed independently by Kohayakawa [47] and Rödl that the regularity lemma can nevertheless be generalized to an appropriate class of graphs with density tending to zero. Their result applies to a natural class of sparse graphs that is wide enough for the lemma to have several interesting applications. In particular, it applies to relatively dense subgraphs of random graphs – that is, one takes a random graph  $G_{n,p}$  of density  $p$  and a subgraph  $G$  of  $G_{n,p}$  of density at least  $\delta p$  (or relative density at least  $\delta$  in  $G_{n,p}$ ), where  $p$  usually tends to 0, while  $\delta > 0$  is usually independent of the number of vertices.

To make this precise, we say that a bipartite graph between sets  $U$  and  $V$  is  $(\epsilon, p)$ -regular if, for every  $U' \subseteq U$  and  $V' \subseteq V$  with  $|U'| \geq \epsilon|U|$  and  $|V'| \geq \epsilon|V|$ , the density  $d(U', V')$  of edges between  $U'$  and  $V'$  satisfies

$$|d(U', V') - d(U, V)| \leq \epsilon p.$$

That is, we alter the definition of regularity so that it is relative to a particular density  $p$ . This density is usually comparable to the total density between  $U$  and  $V$ . A partition of the vertex set of a graph into  $t$  pieces  $V_1, \dots, V_t$  is then said to be  $(\epsilon, p)$ -regular if it is an equipartition and, for all but at most  $\epsilon t^2$  pairs  $(V_i, V_j)$ , the induced graph between  $V_i$  and  $V_j$  is  $(\epsilon, p)$ -regular.

The class of graphs to which the Kohayakawa-Rödl regularity lemma applies are the so-called upper-uniform graphs [53]. Suppose that  $0 < \eta \leq 1$ ,  $D > 1$ , and  $0 < p \leq 1$  are given. We will say that a graph  $G$  is  $(\eta, p, D)$ -upper-uniform if for all disjoint subsets  $U_1$  and  $U_2$  with  $|U_1|, |U_2| \geq \eta|V(G)|$ , the density of edges between  $U_1$  and  $U_2$  satisfies  $d(U_1, U_2) \leq Dp$ . This condition is satisfied for many natural classes of graphs, including all subgraphs of random and pseudorandom graphs of density  $p$ . The regularity lemma of Kohayakawa and Rödl is the following.

**Theorem 1.2.** *For every  $\epsilon, D > 0$  and every positive integer  $t_0$ , there exist  $\eta > 0$  and a positive integer  $T$  such that for every  $p \in [0, 1]$ , every graph  $G$  with at least  $t_0$  vertices that is  $(\eta, p, D)$ -upper-uniform admits an  $(\epsilon, p)$ -regular partition  $V_1, \dots, V_t$  of its vertex set into  $t_0 \leq t \leq T$  pieces.*

The proof of this theorem is essentially the same as the proof of the dense regularity lemma, with the upper uniformity used to ensure that the iteration terminates after a constant number of steps. We note that different versions of the result have appeared in the literature where the  $(\eta, p, D)$ -upper-uniformity assumption is altered [3, 63] or dropped completely [80].

As we have already mentioned above, the usefulness of the dense regularity method relies on the existence of a corresponding counting lemma. Roughly speaking, a counting lemma says that if we start with an arbitrary graph  $H$  and replace its vertices by large independent sets and its edges by  $\epsilon$ -regular bipartite graphs with non-negligible density, then this blown-up graph will contain roughly the expected number of copies of  $H$ . Here is a precise statement to that effect.

**Lemma 1.3.** *For every graph  $H$  with vertex set  $\{1, 2, \dots, k\}$  and every  $\delta > 0$ , there exists  $\epsilon > 0$  and an integer  $n_0$  such that the following statement holds. Let  $n \geq n_0$  and let  $G$  be a graph whose vertex set is a disjoint union  $V_1 \cup \dots \cup V_k$  of sets of size  $n$ . Assume that for each  $ij \in E(H)$ , the bipartite subgraph of  $G$  induced between  $V_i$  and  $V_j$  is  $\epsilon$ -regular and has density  $d_{ij}$ . Then the number of  $k$ -tuples  $(v_1, \dots, v_k) \in V_1 \times \dots \times V_k$  such that  $v_i v_j \in E(G)$  whenever  $ij \in E(H)$  is  $n^k (\prod_{ij \in E(H)} d_{ij} \pm \delta)$ .*

In particular, if the density  $d_{ij}$  is large for every  $ij \in E(H)$ , then  $G$  contains many copies of  $H$ .

Let us define a *canonical copy* of  $H$  in  $G$  to be a  $k$ -tuple as in the lemma above: that is, a  $k$ -tuple  $(v_1, \dots, v_k)$  such that  $v_i \in V_i$  for every  $i \in V(H)$  and  $v_i v_j \in E(G)$  for every  $ij \in E(H)$ . Let us also write  $G(H)$  for the number of canonical copies of  $H$  in  $G$ . (Of course, the definitions of “canonical copy” and  $G(H)$  depend not just on  $G$  but also on the partition of  $G$  into  $V_1, \dots, V_k$ , but we shall suppress this dependence in the notation.)

In order to use Theorem 1.2, one would ideally like a statement similar to Lemma 1.3 but adapted to a sparse context. For this we would have an additional parameter  $p$ , which can tend to zero with  $n$ . We would replace the densities  $d_{ij}$  by  $d_{ij}p$  and we would like to show that  $G(H)$  is approximately  $n^k p^{e(H)} (\prod_{ij \in E(H)} d_{ij} \pm \delta)$ . In order to obtain this stronger conclusion (stronger because the error estimate has been multiplied by  $p^{e(H)}$ ), we need a stronger assumption, and the natural assumption, given the statement of Theorem 1.2 (which is itself natural), is to replace  $\epsilon$ -regularity by  $(\epsilon, p)$ -regularity.

Of course, we cannot expect such a result if  $p$  is too small. Consider the random graph  $G$  obtained from  $H$  by replacing each vertex of  $H$  by an independent set of size  $n$  and each edge of  $H$  by a random bipartite graph with  $pn^2$  edges. With high probability,  $G(H)$  will be about  $p^{e(H)} n^{v(H)}$ . Hence, if  $p^{e(H)} n^{v(H)} \ll pn^2$ , then one can remove all copies of  $H$  from  $G$  by deleting a tiny proportion of all edges. (We may additionally delete a further small proportion of edges to ensure that all bipartite graphs corresponding to the edges of  $H$  have the same number of edges.) It is not hard to see that with high probability the bipartite graphs that make up the resulting graph  $G'$  will be  $(\epsilon, q)$ -regular for some  $q = (1 - o(1))p$ , but that  $G'$  will contain no canonical copies of  $H$ .

Therefore, a sparse analogue of Lemma 1.3 cannot hold if  $p \leq cn^{-\frac{v(H)-2}{e(H)-1}}$  for some small positive constant  $c$ . Note that one can replace  $H$  in the above argument by an arbitrary subgraph  $H' \subseteq H$ , since removing all copies of  $H'$  from a graph also results in a  $H'$ -free subgraph. This observation naturally leads to the notion of *2-density*  $m_2(H)$  of a graph  $H$ , defined by

$$m_2(H) = \max \left\{ \frac{e(H') - 1}{v(H') - 2} : H' \subseteq H \text{ with } v(H') \geq 3 \right\}.$$

(We take  $m_2(K_2) = \frac{1}{2}$ .) With this notation, what we have just seen is that to have any chance of an appropriate analogue of Lemma 1.3 holding, we need to assume that  $p \geq Cn^{-1/m_2(H)}$  for some absolute constant  $C > 0$ .

Unfortunately, there is a more fundamental difficulty with finding a sparse counting lemma to match a sparse regularity lemma. Instead of sparse random graphs with many vertices, one can consider blow-ups of sparse random graphs with far fewer vertices. That is, one can pick a counterexample of the kind just described but with the sets  $V_i$  of size  $r$  for some  $r$  that is much smaller than  $n$ , and then one can replace each vertex of this small graph by an independent set of  $n/r$  vertices to make a graph with  $n$  vertices in each  $V_i$ . Roughly speaking, the counterexample above survives the blowing-up process, and the result is that the hoped-for sparse counting lemma is false whenever  $p = o(1)$ . (For more details, see [34, 52].)

However, these “block” counterexamples have a special structure, so, for  $p \geq Cn^{-1/m_2(H)}$ , it looks plausible that graphs for which the sparse counting lemma fails should be very rare. This intuition was formalized by Kohayakawa, Luczak, and Rödl [50], who made a conjecture that is usually known as the KLR conjecture. Before we state it formally, let us introduce some notation.

As above, let  $H$  be a graph with vertex set  $\{1, 2, \dots, k\}$ . We denote by  $\mathcal{G}(H, n, m, p, \epsilon)$  the collection of all graphs  $G$  obtained in the following way. The vertex set of  $G$  is a disjoint union  $V_1 \cup \dots \cup V_k$  of sets of size  $n$ . For each edge  $ij \in E(H)$ , we add to  $G$  an  $(\epsilon, p)$ -regular bipartite graph with  $m$  edges between the pair  $(V_i, V_j)$ . These are the only edges of  $G$ . Let us also write  $\mathcal{G}^*(H, n, m, p, \epsilon)$  for the set of all  $G \in \mathcal{G}(H, n, m, p, \epsilon)$  that do not contain a canonical copy of  $H$ .

Since the sparse regularity lemma yields graphs with varying densities between the various pairs of vertex sets, it may seem surprising that we are restricting attention to graphs where all the densities are equal (to  $m/n^2$ ). However, as we shall see later, it is sufficient to consider just this case. In fact, the KLR conjecture is more specific still, since it takes all the densities to be equal to  $p$ . Again, it turns out that from this case one can deduce the other cases that are needed.

**Conjecture 1.4.** *Let  $H$  be a fixed graph and let  $\beta > 0$ . Then there exist  $C, \epsilon > 0$  and a positive integer  $n_0$  such that*

$$|\mathcal{G}^*(H, n, m, m/n^2, \epsilon)| \leq \beta^m \binom{n^2}{m}^{\epsilon(H)}$$

for every  $n \geq n_0$  and every  $m \geq Cn^{2-1/m_2(H)}$ .

Note that  $\binom{n^2}{m}^{\epsilon(H)}$  is the number of graphs with vertex set  $V_1 \cup \dots \cup V_k$  with  $m$  edges between each pair  $(V_i, V_j)$  when  $ij \in E(H)$  and no edges otherwise. Thus, we can interpret the conjecture as follows: the probability that a random such graph belongs to the bad set  $\mathcal{G}^*(H, n, m, m/n^2, \epsilon)$  is at most  $\beta^m$ .

The rough idea of the conjecture is that the probability that a graph is bad is so small that a simple union bound tells us that with high probability a random graph does not contain any bad graph – which implies that we may use the sparse embedding lemma we need. In other words,  $(\epsilon, p)$ -regularity on its own does not suffice, but if you know in addition that your graph is a subgraph of a sparse random graph, then with high probability it does suffice.

More precisely, let  $G$  be a random graph with  $N$  vertices and edge probability  $p$  and let  $n = \eta N$  and  $m = dpn^2 \geq n$ . Then the expected number of subgraphs of  $G$  of the form  $\mathcal{G}^*(H, n, m, p, \epsilon)$  is at most

$$p^{m\epsilon(H)} \beta^m \binom{n^2}{m}^{\epsilon(H)} \binom{N}{n}^{v(H)} \leq p^{m\epsilon(H)} \beta^m \left(\frac{e}{dp}\right)^{m\epsilon(H)} \left(\frac{e}{\eta}\right)^{nv(H)} \leq \beta^m \left(\frac{e}{d}\right)^{m\epsilon(H)} \left(\frac{e}{\eta}\right)^{mv(H)}.$$

Therefore, choosing  $\beta$  to be sufficiently small in terms of  $d, \eta$ , and  $H$ , the probability that  $G$  contains a graph in  $\mathcal{G}^*(H, n, m, p, \epsilon)$  is very small. By summing over the possible values of  $n$  and  $m$ , we may rule out such bad subgraphs for all  $n$  and  $m$  with  $n \geq \eta N$  and  $m \geq dpn^2$ .

This does not give us a counting lemma for  $(\epsilon, p)$ -regular subgraphs of  $G$ , but it does at least tell us that every  $(\epsilon, p)$ -regular subgraph of  $G$  with sufficiently dense pairs in the right places contains a canonical copy of  $H$ . In other words, it gives us an embedding lemma, which makes it suitable for several applications to embedding results. For example, as noted in [50], it is already sufficiently strong that a straightforward application of the sparse regularity lemma then allows one to derive the following theorem, referred to as Turán’s theorem for random graphs, which was eventually proved in a different way by Conlon and Gowers [17] (for strictly balanced graphs, i.e., those for which  $m_2(H) > m_2(H')$  for every proper subgraph  $H'$  of  $H$ ) and, independently, Schacht [79] (see

also [10, 78]). We remark that this theorem was the original motivation behind Conjecture 1.4 – see Section 6 of [50]. Following [17], let us say that a graph  $G$  is  $(H, \epsilon)$ -Turán if every subgraph of  $G$  with at least

$$\left(1 - \frac{1}{\chi(H) - 1} + \epsilon\right) e(G)$$

edges contains a copy of  $H$ . Here  $\chi(H)$  is the chromatic number of  $H$ .

**Theorem 1.5.** *For every  $\epsilon > 0$  and every graph  $H$ , there exist positive constants  $c$  and  $C$  such that*

$$\lim_{n \rightarrow \infty} \mathbb{P}(G_{n,p} \text{ is } (H, \epsilon)\text{-Turán}) = \begin{cases} 0, & \text{if } p < cn^{-1/m_2(H)}, \\ 1, & \text{if } p > Cn^{-1/m_2(H)}. \end{cases}$$

The KLR conjecture has attracted considerable attention over the past two decades and has been verified for a number of small graphs. It is straightforward to verify that it holds for all graphs  $H$  that do not contain a cycle. In this case, the class  $\mathcal{G}^*(H, n, m, p, \epsilon)$  will be empty. The cases  $H = K_3, K_4$ , and  $K_5$  were resolved in [49], [32], and [33], respectively. In the case when  $H$  is a cycle, the conjecture was proved in [11, 30] (see also [48] for a slightly weaker version). Very recently, it was proved for all balanced graphs, that is, those graphs  $H$  for which  $m_2(H) = \frac{e(H)-1}{v(H)-2}$ , by Balogh, Morris, and Samotij [10] and by Saxton and Thomason [78] in full generality.

Besides implying Theorem 1.5, Conjecture 1.4 is also sufficient for transferring many other classical extremal results about graphs to subgraphs of the random graph  $G_{n,p}$ , including Ramsey’s theorem [67] and the Erdős-Simonovits stability theorem [81]. However, there are situations where an embedding result is not enough: rather than just a single copy of  $H$ , one needs to know that there are many copies. That is, one needs something more like a full counting lemma. In this paper, we shall state and prove such a “counting version” of the KLR conjecture for subgraphs of random graphs. Later in the paper we shall give examples of classical theorems whose sparse random versions do not follow from the KLR conjecture but do follow from our counting result.

Our main theorem is the following.

**Theorem 1.6.** *For every graph  $H$  and every  $\delta, d > 0$ , there exist  $\epsilon, \xi > 0$  with the following property. For every  $\eta > 0$ , there is a  $C > 0$  such that if  $p \geq Cn^{-1/m_2(H)}$ , then a.a.s. the following holds in  $G_{N,p}$ :*

(i) *For every  $n \geq \eta N$ ,  $m \geq dpn^2$ , and every subgraph  $G$  of  $G_{N,p}$  in  $\mathcal{G}(H, n, m, p, \epsilon)$ ,*

$$G(H) \geq \xi \left(\frac{m}{n^2}\right)^{e(H)} n^{v(H)}. \quad (1)$$

(ii) *Moreover, if  $H$  is strictly balanced, that is, if  $m_2(H) > m_2(H')$  for every proper subgraph  $H'$  of  $H$ , then*

$$G(H) = (1 \pm \delta) \left(\frac{m}{n^2}\right)^{e(H)} n^{v(H)}. \quad (2)$$

Note that strictly speaking the statements above depend not just on the graph  $G$  but on the partition  $V_1 \cup \dots \cup V_k$  that causes  $G$  to belong to  $\mathcal{G}(H, n, m, p, \epsilon)$ . Roughly speaking, (i) tells us that if  $G$  contains “many” edges in the right places, then there are “many” copies of  $H$ , while (ii) tells us that the number of copies of  $H$  is roughly what one would expect for a random graph with pairs of the same densities. We note that a result similar to (ii) holds for all graphs if one is willing to allow some extra logarithmic factors. We will say more about this in the concluding remarks.

The proof of part (i) employs the ideas of Schacht [79], as modified by Samotij [77], and as a result part (i) holds with probability at least  $1 - \exp(-bpN^2)$  for some  $b > 0$  depending on  $H, \eta$ , and  $d$ . Part (ii) is proved using the results of Conlon and Gowers [17] and hence hold with probability at least  $1 - N^{-B}$  for any fixed  $B > 0$ , provided that  $C$  and  $N$  are sufficiently large. Since part (ii) gives an upper bound as well as a lower bound, standard results on upper tail estimates imply that the result cannot hold with the same exponential probability as part (i) (see, for example, [44]).

We note that weaker versions of Theorem 1.6, applicable for larger values of  $p$ , may be found in some earlier papers on the KLR conjecture and Turán's theorem [31, 54] and more recent work on sparse regularity in pseudorandom graphs [16]. We also believe that a variant of part (i) of Theorem 1.6 may be derivable from the work of Saxton and Thomason [78], though they have not stated it in these terms.

## 1.1 Known applications

It is not hard to show that Theorem 1.6, like Conjecture 1.4, implies the best possible sparse random analogues of many classical theorems in extremal graph theory. In particular, it implies Theorem 1.5 above. It also implies the following sparse random version of the Erdős-Simonovits stability theorem, which was first proved by Conlon and Gowers [17] for all strictly balanced graphs and later extended to general  $H$  by Samotij [77], who adapted Schacht's method for this purpose.

**Theorem 1.7.** *For every graph  $H$  and every  $\delta > 0$ , there exist  $\epsilon, C > 0$  such that if  $p \geq Cn^{-1/m_2(H)}$ , then a.a.s. every  $H$ -free subgraph  $G' \subseteq G_{n,p}$  with  $e(G') \geq \left(1 - \frac{1}{\chi(H)-1} - \epsilon\right) \binom{n}{2} p$  may be made  $(\chi(H) - 1)$ -partite by removing at most  $\delta pn^2$  edges.*

Another easy consequence of Theorem 1.6 is the 1-statement of the following sparse Ramsey theorem, originally proved by Rödl and Ruciński [72] in 1995. We remark here that the 1-statement of Theorem 1.8 below also extends to general hypergraphs [17, 28]. Following [17], we say that a graph  $G$  is  $(H, r)$ -Ramsey if every  $r$ -colouring of the edges of  $G$  yields a monochromatic copy of  $H$ .

**Theorem 1.8.** *For every graph  $H$  that is not a star forest or a path of length 3 and for every positive integer  $r \geq 2$ , there exist constants  $c, C > 0$  such that*

$$\lim_{n \rightarrow \infty} \mathbb{P}(G_{n,p} \text{ is } (H, r)\text{-Ramsey}) = \begin{cases} 0, & \text{if } p < cn^{-1/m_2(H)}, \\ 1, & \text{if } p > Cn^{-1/m_2(H)}. \end{cases}$$

We will omit the deductions of Theorem 1.7 and the 1-statements of Theorems 1.5 and 1.8 from Theorem 1.6. These are fairly standard applications of the regularity method (see, for example, [34, 47]).

## 1.2 A sparse removal lemma for graphs

The triangle removal lemma of Ruzsa and Szemerédi [76] states that for every  $\delta > 0$  there exists an  $\epsilon > 0$  such that if  $G$  is any graph on  $n$  vertices that contains at most  $\epsilon n^3$  triangles, then  $G$  may be made triangle-free by removing at most  $\delta n^2$  edges. Despite its innocent appearance, this result has several striking consequences. Most notably, it easily implies Roth's theorem [75] on 3-term arithmetic progressions in dense subsets of the integers. For general graphs  $H$ , a similar statement holds [4, 24, 29] (see also [14, 27]): if an  $n$ -vertex graph contains  $o(n^{v(H)})$  copies of  $H$ , then it may be made  $H$ -free by removing  $o(n^2)$  edges. This result is known as the graph removal lemma. A sparse random version of the graph removal lemma was conjectured by Łuczak in [64] and proved, for strictly balanced  $H$ , by Conlon and Gowers [17]. Here, we apply our main result, Theorem 1.6, to extend this result to all graphs  $H$ .

**Theorem 1.9.** *For every  $\delta > 0$  and every graph  $H$ , there exist positive constants  $\epsilon$  and  $C$  such that if  $p \geq Cn^{-1/m_2(H)}$ , then the following holds a.a.s. in  $G_{n,p}$ . Every subgraph of  $G_{n,p}$  which contains at most  $\epsilon p^{e(H)} n^{v(H)}$  copies of  $H$  may be made  $H$ -free by removing at most  $\delta pn^2$  edges.*

Note that if  $p \leq cn^{-1/m_2(H)}$  for a sufficiently small positive constant  $c$  (depending on  $H$  and  $\delta$ ), then  $H$  has a subgraph  $H'$ , the expected number of copies of which is at most  $\delta pn^2$ , so we can remove all copies of  $H$  by deleting an edge from each copy of  $H'$ . Thus, it is natural to conjecture, as Łuczak did, that Theorem 1.9 actually holds for all values of  $p$ . For balanced graphs, we may close the gap by taking  $\epsilon$  to be sufficiently small in terms of  $C$ ,  $\delta$ , and  $H$ . For  $p \leq Cn^{-1/m_2(H)}$  and  $\epsilon < \delta C^{-e(H)}$ , the number of copies of  $H$  is at most  $\epsilon p^{e(H)} n^{v(H)} \leq \epsilon C^{e(H)} pn^2 < \delta pn^2$ . Deleting an edge from each copy of  $H$  yields the result.

### 1.3 The clique density theorem

For any  $\rho > 0$ , let  $g_k(\rho, n)$  be the minimum number of copies of  $K_k$  which are contained in any graph on  $n$  vertices with density at least  $\rho$ . We then take

$$g_k(\rho) = \lim_{n \rightarrow \infty} \frac{g_k(\rho, n)}{\binom{n}{k}}.$$

Complete  $(k-1)$ -partite graphs demonstrate that  $g_k(\rho) = 0$  for  $\rho \leq 1 - \frac{1}{k-1}$ . On the other hand, a robust version of Turán's theorem known as supersaturation [25] tells us that  $g_k(\rho) > 0$  for all  $\rho > 1 - \frac{1}{k-1}$ .

There has been much work [12, 22, 23, 36, 62] on determining  $g_k(\rho)$  above the natural threshold  $1 - \frac{1}{k-1}$ , but it is only in recent years that the exact dependency of  $g_k(\rho)$  on  $\rho$  has been found for any  $k \geq 3$ . For  $k = 3$  and the interval  $\frac{1}{2} \leq \rho \leq \frac{2}{3}$ , this was accomplished by Fisher [26, 35], while for a general  $\rho$  the value of  $g_3(\rho)$  was determined by Razborov [68, 69] using flag algebras. Employing different methods, Nikiforov [66] reproved the  $k = 3$  case and also solved the case  $k = 4$ . Finally, Reiher [70] recently resolved the general case.

We will show that an analogous theorem holds within the random graph  $G_{n,p}$ . We note that this theorem, Theorem 1.10 below, also follows as a direct application of the results of [17]. To state the result, we define, for a subgraph  $G'$  of  $G_{n,p}$ , the relative density of  $G'$  in  $G_{n,p}$  to be  $e(G')/p \binom{n}{2}$ .

**Theorem 1.10.** *For any  $k \geq 3$  and any  $\epsilon > 0$ , there exists a constant  $C > 0$  such that if  $p \geq Cn^{-2/(k+1)}$ , then the following holds a.a.s. in  $G_{n,p}$ . Any subgraph  $G'$  of  $G_{n,p}$  will contain at least  $(g_k(\rho) - \epsilon)p \binom{n}{k}$  copies of  $K_k$ , where  $\rho$  is the relative density of  $G'$  in  $G_{n,p}$ .*

We remark that, as  $m_2(K_k) = (k+1)/2$ , the assumption on  $p$  in the above theorem is best possible up to the value of the constant  $C$ . Using part (ii) of Theorem 1.6, we may show that a similar theorem also holds for any strictly balanced graph  $H$ . That is, if  $g_H(\rho)$  is the natural analogue of  $g_k(\rho)$  defined for  $H$ , then for any  $\epsilon > 0$ , there exists a  $C > 0$  such that if  $p \geq Cn^{-1/m_2(H)}$ , the random graph  $G_{n,p}$  will a.a.s. be such that any subgraph  $G'$  of  $G_{n,p}$  contains at least  $(g_H(\rho) - \epsilon)p^{e(H)} \frac{n^{v(H)}}{\text{Aut}(H)}$  copies of  $H$ , where again  $\rho$  is the relative density of  $G'$  in  $G_{n,p}$ . However, the function  $g_H(\rho)$  is only well understood for cliques and certain classes of bipartite graph (see, for example, [15, 61]).

### 1.4 The Hajnal-Szemerédi theorem

Let  $H$  be a fixed graph on  $h$  vertices. An arbitrary collection of vertex-disjoint copies of  $H$  in some larger graph is called an  $H$ -packing. A *perfect  $H$ -packing* (or  $H$ -factor) is an  $H$ -packing that

covers all vertices of the host graph. It has long been known for certain graphs  $H$  that if the minimum degree of an  $n$ -vertex graph  $G$  is sufficiently large and  $n$  is divisible by  $h$ , then  $G$  contains an  $H$ -factor. For example, Dirac's theorem [20] implies that if  $H$  is a path of length  $h - 1$ ,  $n$  is divisible by  $h$ , and  $\delta(G) \geq n/2$ , then  $G$  contains an  $H$ -factor. Corrádi and Hajnal [19] proved that  $\delta(G) \geq 2n/3$  implies the existence of a  $K_3$ -factor in  $G$ . A milestone in this area of research, the famous theorem of Hajnal and Szemerédi [42], states that the condition  $\delta(G) \geq (1 - \frac{1}{k})n$  is sufficient to guarantee a perfect  $K_k$ -packing in  $G$  for an arbitrary  $k$  (see [46] for a short proof of this theorem).

**Theorem 1.11.** *For any  $k \geq 3$ , every  $n$ -vertex graph  $G$  with  $\delta(G) \geq (1 - \frac{1}{k})n$  contains a  $K_k$ -factor, provided that  $n$  is divisible by  $k$ .*

This theorem has also been generalized to arbitrary  $H$  [7, 56, 59]. In particular, a result of Komlós [55] shows that a parameter known as the critical chromatic number governs the existence of almost perfect packings (i.e., packings covering all but a  $o(1)$ -fraction of the vertices of the host graph) in graphs of large minimum degree.

We will prove the following approximate version of the Hajnal-Szemerédi Theorem in the random graph  $G_{n,p}$ .

**Theorem 1.12.** *For any  $k \geq 3$  and any  $\gamma > 0$ , there exists a constant  $C > 0$  such that if  $p \geq Cn^{-2/(k+1)}$ , then the following holds a.a.s. in  $G_{n,p}$ . Any subgraph  $G'$  of  $G_{n,p}$  with  $\delta(G') \geq (1 - \frac{1}{k} + \gamma)pn$  contains a  $K_k$ -packing that covers all but at most  $\gamma n$  vertices.*

We remark that the assumption on  $p$  in the above theorem is best possible up to the value of the constant  $C$ . Using the result of Komlós [55], we may show that a similar theorem also holds for any graph  $H$ . That is, for any  $H$  and any  $\gamma > 0$ , there exists a  $C > 0$  such that if  $p \geq Cn^{-1/m_2(H)}$ , then a.a.s. every subgraph  $G'$  of  $G_{n,p}$  satisfying  $\delta(G') \geq (1 - \frac{1}{\chi_{\text{cr}}(H)} + \gamma)pn$ , where  $\chi_{\text{cr}}(H)$  is the critical chromatic number of  $H$ , contains an  $H$ -packing covering all but at most  $\gamma n$  vertices.

Finally, we remark that the problem of finding perfect packings in subgraphs of random graphs seems to be much more difficult. On the positive side, a best possible sparse random analogue of Dirac's theorem was recently proved by Lee and Sudakov [60]. On the negative side, it was observed by Huang, Lee, and Sudakov [43] that for every  $\epsilon > 0$ , there are  $c, C > 0$  such that if every vertex of  $H$  is contained in a triangle and  $Cn^{-1/2} \leq p \leq c$ , then  $G_{n,p}$  a.a.s. contains a spanning subgraph  $G'$  with  $\delta(G') \geq (1 - \epsilon)pn$  such that at least  $\epsilon p^{-2}/3$  vertices of  $G'$  are not contained in a copy of  $H$ . Therefore, the presence of the set of uncovered vertices in the statement of Theorem 1.12 is indispensable. For further discussion and related results, we refer the reader to [9, 43].

## 1.5 The Andrásfai-Erdős-Sós theorem

A result of Zarankiewicz [85] states that if a graph on  $n$  vertices has minimum degree at least  $(1 - \frac{1}{k-1})n$  then it contains a copy of  $K_k$ . This result follows immediately from Turán's theorem but is interesting because it has a surprisingly robust stability version, due to Andrásfai, Erdős, and Sós [8]. This theorem states that any  $K_k$ -free graph on  $n$  vertices with minimum degree at least  $(1 - \frac{3}{3k-4})n$  must be  $(k-1)$ -partite.

This result was extended to general graphs by Alon and Sudakov [6] (see also [1]), who showed that for any graph  $H$  and any  $\gamma > 0$ , every  $H$ -free graph on  $n$  vertices with minimum degree at least  $(1 - \frac{3}{3\chi(H)-4} + \gamma)n$  may be made  $(\chi(H) - 1)$ -partite by deleting  $o(n^2)$  edges. We prove a random analogue of this result, as follows.



**Theorem 1.13.** *For every graph  $H$  and every  $\gamma > 0$ , there exists  $C > 0$  such that if  $p \geq Cn^{-1/m_2(H)}$ , then a.a.s. every  $H$ -free subgraph  $G' \subseteq G_{n,p}$  with  $\delta(G') \geq \left(1 - \frac{3}{3\chi(H)-4} + \gamma\right)pn$  may be made  $(\chi(H) - 1)$ -partite by removing from it at most  $\gamma pn^2$  edges.*

More generally, for any  $H$  and any  $r$ , we may define

$$\delta_\chi(H, r) := \inf\{d: \delta(G) \geq d|G| \text{ and } H \not\subseteq G \Rightarrow \chi(G) \leq r\}.$$

The Andrásfai-Erdős-Sós theorem determines the value of  $\delta_\chi(K_k, k - 1)$ , but there are also some results known for other values of  $r$ . For example, it is known [13, 41, 45] that  $\delta_\chi(K_3, 3) = \frac{10}{29}$  and  $\delta_\chi(K_3, 4) = \frac{1}{3}$ . Our methods easily allow us to transfer any such results about  $\delta_\chi(H, r)$  to the random setting. These are approximate results, proving that any  $H$ -free graph with a certain minimum degree is close to a graph with bounded chromatic number. As noted in [2], one cannot hope to achieve a more exact result saying that the graph itself has bounded chromatic number.

## 2 KLR conjecture via multiple exposure

In this section, we prove part (i) of Theorem 1.6. The proof follows the ideas of [79, 77]. The main idea in these proofs is to expose the random graph in multiple rounds. In this context, this can be traced back to the work of Rödl and Ruciński in [72].

We are going to prove a somewhat stronger statement, Theorem 2.1 below, which easily implies part (i) of Theorem 1.6. A bipartite graph between sets  $U$  and  $V$  is  $(\epsilon, d)$ -lower-regular if, for every  $U' \subseteq U$  and  $V' \subseteq V$  with  $|U'| \geq \epsilon|U|$  and  $|V'| \geq \epsilon|V|$ , the density  $d(U', V')$  of edges between  $U'$  and  $V'$  satisfies  $d(U', V') \geq d$ . Given a graph  $H$  on the vertex set  $\{1, \dots, k\}$ , we denote by  $\mathcal{G}_\ell(H, n, d, \epsilon)$  the collection of all graphs  $G$  on the vertex set  $V_1 \cup \dots \cup V_k$ , where  $V_1, \dots, V_k$  are pairwise disjoint sets of size  $n$  each, whose edge set consists of  $e(H)$  different  $(\epsilon, d)$ -lower-regular bipartite graphs, one graph between  $V_i$  and  $V_j$  for each  $ij \in E(H)$ . For an arbitrary graph  $G$  and  $p \in [0, 1]$ , we will denote by  $G_p$  the random subgraph of  $G$ , where each edge of  $G$  is included with probability  $p$ , independently of all other edges. Finally, given  $H' \subseteq H$  and two graphs  $G, G' \in \mathcal{G}_\ell(H, n, 0, 1)$ , we denote by  $\mathcal{C}(H, G; H', G')$  the set of all canonical copies of  $H$  in  $G$  such that the edges of  $H'$  in these copies are in  $G'$ .

**Theorem 2.1.** *Let  $H$  be an arbitrary graph. For every  $H' \subseteq H$  and every  $d > 0$ , there exist  $\epsilon, \xi, b, C > 0$  and an integer  $n_0$  such that if  $n \geq n_0$  and  $p \geq Cn^{-1/m_2(H)}$ , then the following holds. For every  $G \in \mathcal{G}_\ell(H, n, d, \epsilon)$ , with probability at least  $1 - \exp(-bpn^2)$ , the random graph  $G_p$  has the following property: Every subgraph  $G'$  of  $G_p$  in  $\mathcal{G}_\ell(H, n, dp, \epsilon)$  satisfies*

$$|\mathcal{C}(H, G; H', G')| \geq \xi p^{e(H')} n^{v(H)}.$$

We now deduce the first part of Theorem 1.6 from Theorem 2.1.

*Proof of part (i) of Theorem 1.6.* Let  $H$  be an arbitrary  $k$ -vertex graph and let  $d > 0$ . We may assume that  $H$  contains a vertex of degree at least two (and hence  $m_2(H) \geq 1$ ) since otherwise the assertion of the theorem is trivial. Let

$$\epsilon = \min\{d/2, \epsilon_{2.1}(H, d/2)\}, \quad b' = \min\{1/40, b_{2.1}(H, d/2)\}, \quad \text{and} \quad C' = C_{2.1}(H, d/2).$$

Moreover, let  $\xi' = \xi_{2.1}(H, d/2)$  and let  $\xi = 2^{-e(H)}\xi'$ . Next, fix an arbitrary positive constant  $\eta$  and let  $C = \max\{C', 40k\} \cdot \eta^{-2}$ . Assume that  $p \geq CN^{-1/m_2(H)}$  and that  $N$  is sufficiently large. Finally,

let  $n \geq \eta N$  and let  $m \geq dpn^2$ . We estimate the probability of the event (we denote it by  $\mathcal{B}$ ) that  $G_{N,p}$  contains a subgraph  $G \in \mathcal{G}(H, n, m, p, \epsilon)$  with  $G(H) < \xi \left(\frac{m}{n^2}\right)^{e(H)} n^{v(H)}$ .

Let  $W_1, \dots, W_k$  be pairwise disjoint subsets of  $V(G_{N,p})$ , each of size  $n$ , and let  $\mathcal{B}(W_1, \dots, W_k)$  denote the event that  $G_{N,p}$  contains a subgraph  $G$  as above with sets  $W_1, \dots, W_k$  playing the role of  $V_1, \dots, V_k$  from the definition of  $\mathcal{G}(H, n, m, p, \epsilon)$ . By Chernoff's inequality (see below),

$$\Pr(e_{G_{N,p}}(W_i, W_j) \geq 2pn^2 \text{ for some } ij \in E(H)) \leq e(H) \cdot \exp(-pn^2/16) \leq \exp(-pn^2/20).$$

On the other hand, if  $m \leq 2pn^2$ , then by Theorem 2.1 with  $H = H'$ ,  $d_{2.1} = d/2$  and  $G_{2.1} = H(W_1, \dots, W_k)$ , the complete blow-up of  $H$  obtained by replacing the vertices of  $H$  by sets  $W_1, \dots, W_k$  and the edges by complete bipartite graphs, the probability that  $G_{N,p} \cap H(W_1, \dots, W_k)$  contains a subgraph  $G \in \mathcal{G}(H, n, m, p, \epsilon) \subseteq \mathcal{G}_\ell(H, n, pd/2, \epsilon)$  satisfying

$$G(H) < \xi \left(\frac{m}{n^2}\right)^{e(H)} n^{v(H)} \leq 2^{e(H)} \xi p^{e(H)} n^{v(H)} = \xi' p^{e(H)} n^{v(H)}$$

is at most  $\exp(-bpn^2)$ . To see this, note that  $H(W_1, \dots, W_k)_p$  has the same distribution as  $G_{N,p} \cap H(W_1, \dots, W_k)$  and that our assumptions imply that  $p \geq CN^{-1/m_2(H)} \geq C'n^{-1/m_2(H)}$ . It follows that

$$\Pr(\mathcal{B}) \leq \sum_{W_1, \dots, W_k} \Pr(\mathcal{B}(W_1, \dots, W_k)) \leq (k+1)^N [\exp(-pn^2/20) + \exp(-bpn^2)] \leq \exp(-b'\eta^2 pN^2/2)$$

since  $pn^2 \geq \eta^2 pN^2 \geq C\eta^2 N^{2-1/m_2(H)} \geq 40kN$ .  $\square$

Note that we used Chernoff's bound in the following standard form (see, for example, [5, Appendix A]).

**Lemma 2.2** (Chernoff's inequality). *Let  $t$  be a positive integer,  $p \in [0, 1]$ , and  $X \sim \text{Bin}(t, p)$ . For every positive  $a$ ,*

$$\Pr(X < pt - a) < \exp\left(-\frac{a^2}{pt}\right) \quad \text{and} \quad \Pr(X > pt + a) < \exp\left(-\frac{a^2}{2pt} + \frac{a^3}{2(pt)^2}\right).$$

In particular, if we apply the upper tail estimate with  $a = pt/2$ , we see that

$$\Pr(X > 2pt) \leq \Pr(X > 3pt/2) \leq \exp(-pt/16),$$

as required in the proof above with  $t = n^2$ .

In the proof of Theorem 2.1, we will also need the following approximate concentration result for random subgraphs of  $H(n)$ , the complete blow-up of  $H$  obtained by replacing the vertices of  $H$  by disjoint sets  $V_1, \dots, V_k$  of size  $n$  each and the edges of  $H$  by complete bipartite graphs. Lemma 2.3 below is [79, Proposition 3.6] with  $H_n$  being the  $e(H)$ -uniform hypergraph on the vertex set  $E(H(n))$  whose edges are the canonical copies of  $H$  in  $H(n)$ . The proof of the fact that this hypergraph is  $(K, n^{-1/m_2(H)})$ -bounded, see [79, Definition 3.2], is implicit in the proof of the 1-statement of [79, Theorem 2.7]. The definition of  $\deg_{H'}$  is given in Section 2.2.2.

**Lemma 2.3** ([72, 79]). *Let  $H$  be an arbitrary graph with  $\Delta(H) \geq 2$ . There exists a  $K > 0$  such that for every proper  $H' \subseteq H$  and every  $\eta > 0$ , there exist  $b > 0$  and an integer  $n_0$  such that for every  $n \geq n_0$ , if  $p \geq n^{-1/m_2(H)}$ , then with probability at least  $1 - \exp(-bpn^2)$ , for every  $ij \in E(H) \setminus E(H')$  there exists a subgraph  $X \subseteq H(n)_p$  with  $|X| \leq \eta pn^2$  satisfying*

$$\sum_{e \in (V_i, V_j)} \deg_{H'}^2(e, H(n), H(n)_p \setminus X) \leq Kp^{2e(H')} n^{2v(H)-2}.$$

Let  $H$  be an arbitrary graph and let  $d$  be a positive constant. In the remainder of this section, we prove Theorem 2.1 by induction on the number of edges in  $H'$ .

## 2.1 Induction base ( $e(H') = 0$ )

Let  $H'$  be the empty subgraph of  $H$  and note that, regardless of  $G'$ , we have  $|\mathcal{C}(H, G; H', G')| = G(H)$ . The base of the induction follows immediately from the following one-sided version of the counting lemma, Lemma 1.3, if we let  $\epsilon = \epsilon_{2.4}(d)$ ,  $\xi = \xi_{2.4}(d)$ ,  $b = 1$ ,  $C = 1$ , and  $n_0 = 1$ .

**Lemma 2.4.** *For every graph  $H$  and every  $d > 0$ , there exist  $\epsilon, \xi > 0$  such that for every  $n$  and every  $G \in \mathcal{G}_\ell(H, n, d, \epsilon)$ ,*

$$G(H) \geq \xi n^{v(H)}.$$

In fact, by choosing  $\epsilon$  sufficiently small, one can show that  $\xi \geq d^{e(H)} - \delta$ .

## 2.2 Induction step

Let  $H''$  be an arbitrary subgraph of  $H'$  with  $e(H') - 1$  edges. Let  $\epsilon, \xi', b', C'$ , and  $n'_0$  be the constants whose existence is asserted by the inductive assumption with  $H'$  replaced by  $H''$  and  $d$  replaced by  $d/4$ , i.e., let

$$\epsilon = \epsilon_{2.1}(H'', d/4), \quad \xi' = \xi_{2.1}(H'', d/4), \quad b' = b_{2.1}(H'', d/4), \quad \text{and} \quad C' = C_{2.1}(H'', d/4).$$

We also let  $K = K_{2.3}(H)$ ,  $\eta = \epsilon^2 d/8$ , and  $\hat{b} = b_{2.3}(H, H'', \eta)$ . Furthermore, let

$$R = \left\lceil \frac{4K}{(\xi')^2} + 1 \right\rceil \tag{3}$$

and let

$$\xi = \frac{\xi' \epsilon^2 d}{16(RL^R)^{e(H)}}, \quad b = \min \left\{ \frac{b^*}{2RL^R}, \frac{e(H)d}{21RL^R} \right\}, \quad \text{and} \quad C = \frac{32RL^R C'}{\epsilon^2 d},$$

where

$$b^* = \min \left\{ \frac{\epsilon^2 d}{16^2}, \frac{b'}{4}, \frac{\hat{b}}{4} \right\} \quad \text{and} \quad L = \frac{3e(H)}{b^*}.$$

Throughout the proof, we will assume that  $n \geq n_0$ , where  $n_0$  is sufficiently large. In particular, we will assume that  $n_0$  is larger than the values of  $n$  which come from applying Theorem 2.1 and Lemma 2.3 in context. Finally, assume that  $p$  satisfies  $p \geq Cn^{-1/m_2(H)}$  and fix some  $G \in \mathcal{G}_\ell(H, n, d, \epsilon)$ .

### 2.2.1 Multiple exposure trick

Let  $\mathcal{S}$  denote the event that the random graph  $G_p$  possesses the postulated property:

$$\mathcal{S}: \quad \text{Every subgraph } G' \subseteq G_p \text{ such that } G' \in \mathcal{G}_\ell(H, n, dp, \epsilon) \text{ satisfies } |\mathcal{C}(H, G; H', G')| \geq \xi p^{e(H')} n^{v(H)}.$$

Following Samotij [77], we will consider a richer probability space that is in a natural correspondence with the space  $\mathcal{P}(G)$  of all subgraphs of  $G$  equipped with the obvious probability measure  $\text{Pr}$ , i.e., the distribution of the random graph  $G_p$ . To this end, let  $p_1, \dots, p_R \in [0, 1]$  be the unique numbers that satisfy

$$1 - p = \prod_{s=1}^R (1 - p_s) \quad \text{and} \quad p_{s+1} = L p_s \text{ for every } s \in [R - 1] \tag{4}$$

and observe that

$$\sum_{s=1}^R p_s \geq p \quad \text{and consequently} \quad p_s \geq p_1 \geq \frac{p}{RL^R} \quad \text{for every } s \in [R]. \quad (5)$$

The richer probability space will be the space  $\mathcal{P}(G)^R$  equipped with the product measure  $\Pr^*$  that is the distribution of the sequence  $(G_{p_1}, \dots, G_{p_R})$  of independent random variables, where for each  $s$ , the variable  $G_{p_s}$  is a  $p_s$ -random subgraph of  $G$ . Crucially, observe that due to our choice of  $p_1, \dots, p_s$ , see (4), the natural mapping

$$\varphi: \mathcal{P}(G)^R \rightarrow \mathcal{P}(G) \quad \text{defined by} \quad \varphi(G_1, \dots, G_R) = G_1 \cup \dots \cup G_R$$

is measure preserving, i.e., for every  $G_0 \subseteq G$ ,

$$\Pr(G_0) = \Pr^*(\varphi^{-1}(G_0)).$$

In other words, the variables  $G_p$  and  $G_{p_1} \cup \dots \cup G_{p_R}$  have the same distribution. Finally, let  $d^* = d/2$  and consider the following event in the space  $\mathcal{P}(G)^R$ :

$$\mathcal{S}^*: \quad \text{For every } G'_1 \subseteq G_{p_1}, \dots, G'_R \subseteq G_{p_R} \text{ such that } G'_s \in \mathcal{G}_\ell(H, n, d^*p_s, \epsilon) \text{ for every } s \in [R], \\ \text{we have } |\mathcal{C}(H, G; H', G'_1 \cup \dots \cup G'_R)| \geq \xi p^{e(H')} n^{v(H)}.$$

There are two reasons why we consider the probability space  $\mathcal{P}(G)^R$ . The first reason is that the probability of  $\mathcal{S}^*$  is much easier to estimate than the probability of  $\mathcal{S}$ . The second reason is that a lower bound on  $\Pr^*(\mathcal{S}^*)$  implies a (marginally weaker) lower bound on  $\Pr(\mathcal{S})$ , which we show below.

**Claim 2.5.**  $1 - \Pr(\mathcal{S}) \leq 2 \cdot (1 - \Pr^*(\mathcal{S}^*)).$

*Proof.* Note that in order to prove the claim, it suffices to show that

$$\Pr^*(\mathcal{S}^* \mid \varphi^{-1}(\hat{G})) = \Pr^*(\mathcal{S}^* \mid G_{p_1} \cup \dots \cup G_{p_R} = \hat{G}) \leq 1/2 \quad (6)$$

for every  $\hat{G}$  that does *not* satisfy  $\mathcal{S}$ . Indeed, assuming that (6) holds for every such  $\hat{G}$ , we have

$$\begin{aligned} 1 - \Pr(\mathcal{S}) &= \sum_{\hat{G} \notin \mathcal{S}} \Pr(G_p = \hat{G}) = \sum_{\hat{G} \notin \mathcal{S}} \Pr^*(G_{p_1} \cup \dots \cup G_{p_R} = \hat{G}) \\ &\leq \sum_{\hat{G} \notin \mathcal{S}} 2 \left( 1 - \Pr^*(\mathcal{S}^* \mid G_{p_1} \cup \dots \cup G_{p_R} = \hat{G}) \right) \cdot \Pr^*(G_{p_1} \cup \dots \cup G_{p_R} = \hat{G}) \\ &\leq 2 \sum_{\hat{G}} \left( 1 - \Pr^*(\mathcal{S}^* \mid G_{p_1} \cup \dots \cup G_{p_R} = \hat{G}) \right) \cdot \Pr^*(G_{p_1} \cup \dots \cup G_{p_R} = \hat{G}) \\ &= 2(1 - \Pr^*(\mathcal{S}^*)). \end{aligned}$$

Consider an arbitrary  $\hat{G} \subseteq G$  that does not satisfy  $\mathcal{S}$ . By the definition of  $\mathcal{S}$ , there exists a subgraph  $G' \subseteq \hat{G}$  such that  $G' \in \mathcal{G}_\ell(H, n, dp, \epsilon)$  but  $|\mathcal{C}(H, G; H', G')| < \xi p^{e(H')} n^{v(H)}$ . Consider the event  $\varphi^{-1}(\hat{G})$ , i.e., the event  $G_{p_1} \cup \dots \cup G_{p_R} = \hat{G}$ . Now, for each  $s \in [R]$ , let  $G'_s = G' \cap G_{p_s}$ . Since clearly  $\mathcal{C}(H, G; H', G') = \mathcal{C}(H, G; H', G'_1 \cup \dots \cup G'_R)$ , it suffices to show that with probability at least  $1/2$ , we have  $G'_s \in \mathcal{G}_\ell(H, n, d^*p_s, \epsilon)$  for every  $s \in [R]$ .

To this end, observe that conditioned on the event  $G_{p_1} \cup \dots \cup G_{p_R} = \hat{G}$ , for each  $s \in [R]$ , the variable  $G_{p_s}$  has the same distribution as  $\hat{G}_{p'_s}$ , where  $p'_s = p_s/p$  (although  $G_{p_1}, \dots, G_{p_R}$  are

no longer independent). Since  $G' \in \mathcal{G}_\ell(H, n, dp, \epsilon)$  and for each  $s \in [R]$ , the graph  $G'_s$  is simply a  $p'_s$ -random subgraph of  $G'$ , it follows from Chernoff's inequality and the union bound that for fixed  $s \in [R]$ , the probability that  $G'_s \notin \mathcal{G}_\ell(H, n, d^*p_s, \epsilon) = \mathcal{G}_\ell(H, n, dp \cdot p'_s/2, \epsilon)$  is at most

$$e(H) \cdot 2^{2n} \cdot \exp(-\epsilon^2 dp_s n^2 / 16).$$

Since  $p_s \geq p/(RL^R) \geq Cn^{-1/m_2(H)}/(RL^R) \geq 32n^{-1}/(\epsilon^2 d)$ , the claimed estimate follows from the union bound, provided that  $n$  is sufficiently large.  $\square$

### 2.2.2 Estimating the probability of $\mathcal{S}^*$

In the remainder of the proof, we will work in the space  $\mathcal{P}(G)^R$  and estimate the probability of the event  $\mathcal{S}^*$ , that is,  $\Pr^*(\mathcal{S}^*)$ . Let  $G_{p_1}, \dots, G_{p_R}$  be independent random subgraphs of  $G$ . Given  $G'_1 \subseteq G_{p_1}, \dots, G'_R \subseteq G_{p_R}$ , let

$$G(s) = (G_{p_1}, \dots, G_{p_s}) \quad \text{and} \quad G'(s) = (G'_1, \dots, G'_s).$$

Let  $ij$  be the edge of  $H'$  that is missing in  $H''$  and let  $V_i$  and  $V_j$  be the subsets of  $V(G)$  corresponding to the vertices  $i$  and  $j$ , respectively. We consider the set  $Z_s$  of 'rich' edges of  $G(V_i, V_j)$  that belong to many copies of  $H$  from  $\mathcal{C}(H, G; H'', G'_s)$ , which we define by

$$Z_s = \left\{ e \in G(V_i, V_j) : \deg_{H''}(e, G, G'_s) \geq \frac{\xi'}{2} p_s^{e(H'')} n^{v(H)-2} \right\},$$

where

$$\deg_{H''}(e, G, G') = |\{J \in \mathcal{C}(H, G; H'', G') : e \in J\}|,$$

and let

$$Z(s) = Z_1 \cup \dots \cup Z_s.$$

Now comes the key step in the proof. We show that with very high probability, for every  $s \in [R]$ , regardless of  $G(s-1)$  and  $G'(s-1)$ , either  $\mathcal{C}(H, G; H', G'_1 \cup \dots \cup G'_s)$  is large or the set  $Z(s)$  of 'rich' edges grows by more than  $n^2/R$ , that is,  $|Z(s) \setminus Z(s-1)| \geq n^2/R$ .

**Claim 2.6.** *For every  $s \in [R]$  and every choice of  $G'(s-1) \subseteq G(s-1) \in \mathcal{P}(G)^{s-1}$ , let  $\mathcal{S}_{G'(s-1)}^*$  denote the event that  $G_{p_s}$  has the following property: For every  $G'_s \subseteq G_{p_s}$  that satisfies  $G'_s \in \mathcal{G}_\ell(H, n, d^*p_s, \epsilon)$ , either*

$$|\mathcal{C}(H, G; H', G'_1 \cup \dots \cup G'_s)| \geq \xi p^{e(H')} n^{v(H)} \quad (7)$$

or

$$|Z(s) \setminus Z(s-1)| \geq \frac{(\xi')^2}{4K} n^2. \quad (8)$$

Then for every  $\hat{G} \in \mathcal{P}(G)^{s-1}$ ,

$$\Pr^*(\mathcal{S}_{G'(s-1)}^* \mid G(s-1) = \hat{G}) \geq 1 - \exp(-2b^* p_s n^2),$$

where  $\Pr^*(\mathcal{S}_{G'(0)}^* \mid G(0) = \hat{G}) = \Pr^*(\mathcal{S}_{G'(0)}^*)$ .

### 2.2.3 Deducing Theorem 2.1 from Claims 2.5 and 2.6

For every  $s \in [R]$ , let  $\mathcal{A}_s$  denote the event that  $e(G_{p_s}) \leq 2p_s e(G)$  and let  $\mathcal{A}(s) = \mathcal{A}_1 \cap \dots \cap \mathcal{A}_s$ . Observe that by (4),

$$\sum_{t=1}^{s-1} 2p_t e(G) \leq 3p_{s-1} e(G) = \frac{3p_s e(G)}{L} \leq \frac{3e(H)p_s n^2}{L}. \quad (9)$$

By Chernoff's inequality, (4), (5), and the fact that  $e(G) \geq e(H)dn^2$ ,

$$\Pr^*(\neg \mathcal{A}(s-1)) \leq \sum_{t=1}^{s-1} \exp\left(-\frac{p_t e(G)}{16}\right) \leq s \cdot \exp\left(-\frac{p_1 e(G)}{16}\right) \leq \exp\left(-\frac{e(H)dp_1 n^2}{20}\right). \quad (10)$$

Now, for every  $s \in [R]$ , let  $\mathcal{S}_s^*$  denote the event that  $\mathcal{A}(s-1)$  holds and  $\mathcal{S}_{G'(s-1)}^*$  holds for all  $G'(s-1) \subseteq G(s-1)$ . Here,  $G'(s-1) \subseteq G(s-1)$  means that inclusion holds coordinate-wise. Observe that if  $\mathcal{S}_s^*$  holds for all  $s \in [R]$ , then  $\mathcal{S}^*$  must hold since (8) in Claim 2.6 can occur at most  $R-1$  times, see (3). Let

$$\hat{\mathcal{A}} = \left\{ \hat{G} \in \mathcal{P}(G)^{s-1} : |\hat{G}_t| \leq 2p_t e(G) \text{ for all } t \in [s-1] \right\}$$

and note that

$$\begin{aligned} \Pr^*(\neg \mathcal{S}_s^*) &\leq \Pr^*(\neg \mathcal{A}(s-1)) + \Pr^*(\neg \mathcal{S}_s^* \wedge \mathcal{A}(s-1)) \\ &= \Pr^*(\neg \mathcal{A}(s-1)) + \sum_{\hat{G} \in \hat{\mathcal{A}}} \Pr^*(\neg \mathcal{S}_s^* \wedge G(s-1) = \hat{G}). \end{aligned} \quad (11)$$

Now, by Claim 2.6, for every  $\hat{G} \in \hat{\mathcal{A}}$ ,

$$\begin{aligned} \Pr^*(\neg \mathcal{S}_s^* \wedge G(s-1) = \hat{G}) &= \Pr^*(\neg \mathcal{S}_s^* \mid G(s-1) = \hat{G}) \cdot \Pr^*(G(s-1) = \hat{G}) \\ &\leq \sum_{G'(s-1) \subseteq \hat{G}} \Pr^*(\neg \mathcal{S}_{G'(s-1)}^* \mid G(s-1) = \hat{G}) \cdot \Pr^*(G(s-1) = \hat{G}) \\ &\leq 2^{\sum_{t=1}^{s-1} 2p_t e(G)} \cdot \exp(-2b^* p_s n^2) \cdot \Pr^*(G(s-1) = \hat{G}). \end{aligned} \quad (12)$$

Since clearly  $\sum_{\hat{G} \in \hat{\mathcal{A}}} \Pr^*(G(s-1) = \hat{G}) \leq 1$ , it follows from (9), (10), (11), and (12) that

$$\begin{aligned} \Pr^*(\neg \mathcal{S}^*) &\leq \sum_{s=1}^R \Pr^*(\neg \mathcal{S}_s^*) \leq R \exp\left(-\frac{e(H)dp_1 n^2}{20}\right) + \sum_{s=1}^R 2^{\frac{3e(H)p_s n^2}{L}} \exp(-2b^* p_s n^2) \\ &\leq R \exp\left(-\frac{e(H)dp_1 n^2}{20}\right) + R \exp(-b^* p_1 n^2) \leq \frac{1}{2} \exp(-bpn^2). \end{aligned} \quad (13)$$

Now, Theorem 2.1 easily follows from (13) and Claim 2.5.

### 2.2.4 Proof of Claim 2.6

Let  $s \in [R]$ , condition on the event  $G(s-1) = \hat{G}$  for some  $\hat{G} \in \mathcal{P}(G)^{s-1}$ , and assume that  $G'(s-1)$  is given. Note that this uniquely defines the graph  $Z(s-1)$  of 'rich' edges. Also, observe that it

follows from the definition of  $Z(s-1)$  and (5) that

$$\begin{aligned}
|\mathcal{C}(H, G; H', G'_1 \cup \dots \cup G'_s)| &\geq \sum_{e \in G'_s \cap Z(s-1)} \deg_{H''}(e, G, G'_1 \cup \dots \cup G'_{s-1}) & (14) \\
&\geq e(G'_s \cap Z(s-1)) \cdot \frac{\xi'}{2} p_1^{e(H'')} n^{v(H)-2} \\
&\geq \frac{e(G'_s \cap Z(s-1))}{p_s} \cdot \frac{\xi'}{2} \frac{p^{e(H')}}{(RLR)^{e(H')}} n^{v(H)-2} \\
&\geq \frac{e(G'_s \cap Z(s-1))}{(\epsilon^2 d/8) p_s n^2} \cdot \xi p^{e(H')} n^{v(H)},
\end{aligned}$$

hence it will be enough if we show that

$$e(G'_s \cap Z(s-1)) \geq \frac{\epsilon^2 d}{8} p_s n^2. \quad (15)$$

Let  $B = G(V_i, V_j)$ . We consider two cases, depending on whether the bipartite graph  $B \setminus Z(s-1)$  of ‘poor’ edges is  $(\epsilon, d^*/2)$ -lower-regular.

**Case 1.  $B \setminus Z(s-1)$  is not  $(\epsilon, d^*/2)$ -lower-regular**

In this case, there exist sets  $X_i \subseteq V_i$  and  $X_j \subseteq V_j$  with  $|X_i|, |X_j| \geq \epsilon n$  such that

$$e_{B \setminus Z(s-1)}(X_i, X_j) < (d^*/2) |X_i| |X_j|.$$

By Chernoff’s inequality, with probability at least  $1 - \exp(-2b^* p_s n^2)$ , the graph  $B_{p_s}$  satisfies

$$e_{B_{p_s} \setminus Z(s-1)}(X_i, X_j) \leq (3d^*/4) p_s |X_i| |X_j|.$$

Consequently, since for every  $G'_s \subseteq G$  satisfying  $G'_s \in \mathcal{G}_\ell(H, n, d^* p_s, \epsilon)$ , the graph  $G'_s \cap B$  is  $(\epsilon, d^* p_s)$ -lower-regular, then

$$\begin{aligned}
e(G'_s \cap Z(s-1)) &\geq e_{G'_s \cap Z(s-1)}(X_i, X_j) \geq e_{G'_s}(X_i, X_j) - e_{B_{p_s} \setminus Z(s-1)}(X_i, X_j) \\
&\geq (d^*/4) p_s |X_i| |X_j| \geq (d^*/4) \epsilon^2 p_s n^2 = (d/8) \epsilon^2 p_s n^2,
\end{aligned}$$

which, by (14) and (15), proves (7).

**Case 2.  $B \setminus Z(s-1)$  is  $(\epsilon, d^*/2)$ -lower-regular**

In this case, we will apply the inductive assumption to the graph  $G \setminus Z(s-1)$ , which clearly is  $(\epsilon, d^*/2)$ -lower regular as  $Z(s-1) \subseteq B \subseteq G$ . First, observe that if  $e(G'_s \cap Z(s-1)) \geq (d/8) \epsilon^2 p_s n^2$ , then this, by (14) and (15), proves (7), so from now on we may assume that the opposite inequality holds, i.e., that

$$e(G'_s \cap Z(s-1)) < \frac{d}{8} \epsilon^2 p_s n^2. \quad (16)$$

Let  $\tilde{G} = G \setminus Z(s-1)$  and observe again that  $\tilde{G} \in \mathcal{G}_\ell(H, n, d/4, \epsilon)$ . Note that by (5) and our assumption on  $p$ , we have  $p_s \geq \frac{p}{RLR} \geq \frac{Cn^{-1/m_2(H)}}{RLR} \geq C'n^{-1/m_2(H)}$  and hence by the inductive assumption applied to  $\tilde{G}$ , with probability at least  $1 - \exp(-b' p_s n^2)$ , every subgraph  $\tilde{G}' \subseteq \tilde{G}_{p_s}$  in  $\mathcal{G}_\ell(H, n, (d/4) p_s, \epsilon)$  satisfies

$$|\mathcal{C}(H, \tilde{G}; H'', \tilde{G}')| \geq \xi' p_s^{e(H'')} n^{v(H)}.$$

Moreover, it follows from Lemma 2.3 that with probability at least  $1 - \exp(-\hat{b} p_s |V|)$ , there exists a subgraph  $X \subseteq \tilde{G}_{p_s}$  satisfying

$$e(X) \leq \eta p_s n^2 \quad (17)$$

and

$$\sum_{e \in \tilde{G}(V_i, V_j)} \deg_{H''}^2(e, \tilde{G}, \tilde{G}_{p_s} \setminus X) \leq K p_s^{2e(H'')} n^{2v(H)-2}. \quad (18)$$

Consider the graph  $\tilde{G}' \subseteq \tilde{G}_{p_s}$  defined by  $\tilde{G}' = G'_s \setminus (X \cup Z(s-1))$ . We verify that  $\tilde{G}' \in \mathcal{G}_\ell(H, n, (d/4)p_s, \epsilon)$ . It follows from (16) and (17) that for every  $i'j' \in E(H)$  and every pair of sets  $W_{i'} \subseteq V_{i'}$  and  $W_{j'} \subseteq V_{j'}$  with  $|W_{i'}| \geq \epsilon n$  and  $|W_{j'}| \geq \epsilon n$ ,

$$\begin{aligned} e_{\tilde{G}'}(W_{i'}, W_{j'}) &\geq e_{G'_s}(W_{i'}, W_{j'}) - e(G'_s \cap Z(s-1)) - e(X) \\ &\geq \left( d^* - \frac{d}{8} - \frac{\eta}{\epsilon^2} \right) p_s |W_{i'}| |W_{j'}| \geq (d/4) p_s |W_{i'}| |W_{j'}|. \end{aligned}$$

From the inductive assumption (which holds for  $\tilde{G}_{p_s}$  with probability at least  $1 - \exp(-b' p_s n^2)$ ), we infer that

$$\sum_{e \in \tilde{G}(V_i, V_j)} \deg_{H''}(e, \tilde{G}, \tilde{G}') = |\mathcal{C}(H, \tilde{G}; H'', \tilde{G}')| \geq \xi' p_s^{e(H'')} n^{v(H)}. \quad (19)$$

Let

$$Z'_s = \left\{ e \in \tilde{G}(V_i, V_j) : \deg_{H''}(e, \tilde{G}, \tilde{G}') \geq \frac{\xi'}{2} p_s^{e(H'')} n^{v(H)-2} \right\}$$

and note that, by definition,  $Z'_s \subseteq Z_s$  and, by (19),

$$\begin{aligned} \sum_{e \in Z'_s} \deg_{H''}(e, \tilde{G}, \tilde{G}') &\geq \sum_{e \in \tilde{G}(V_i, V_j)} \deg_{H''}(e, \tilde{G}, \tilde{G}') - e(\tilde{G}(V_i, V_j) \setminus Z'_s) \cdot \frac{\xi'}{2} p_s^{e(H'')} n^{v(H)-2} \\ &\geq \xi' p_s^{e(H'')} n^{v(H)} - n^2 \cdot \frac{\xi'}{2} p_s^{e(H'')} n^{v(H)-2} = \frac{\xi'}{2} p_s^{e(H'')} n^{v(H)}. \end{aligned} \quad (20)$$

It follows from (18), (20), and the Cauchy-Schwarz inequality that

$$\begin{aligned} K p_s^{2e(H'')} n^{2v(H)-2} &\geq \sum_{e \in \tilde{G}(V_i, V_j)} \deg_{H''}^2(e, \tilde{G}, \tilde{G}_{p_s} \setminus X) \geq \sum_{e \in \tilde{G}(V_i, V_j)} \deg_{H''}^2(e, \tilde{G}, \tilde{G}') \\ &\geq \frac{1}{|Z'_s|} \left( \sum_{e \in Z'_s} \deg_{H''}(e, \tilde{G}, \tilde{G}') \right)^2 \geq \frac{1}{|Z'_s|} \left( \frac{\xi' p_s^{e(H'')} n^{v(H)}}{2} \right)^2 \end{aligned}$$

and consequently,

$$|Z'_s| \geq \frac{(\xi')^2}{4K} n^2.$$

Since  $Z'_s \subseteq \tilde{G} = G \setminus Z(s-1)$ , the graphs  $Z'_s$  and  $Z(s-1)$  are disjoint. Therefore, (8) holds with probability at least

$$1 - \exp(-b' p_s n^2) - \exp(-\hat{b} p_s n^2),$$

which is at least  $1 - \exp(-2b^* p_s n^2)$ . This concludes the proof of the claim.

### 3 KLR conjecture via transference

The method employed by Conlon and Gowers [17] to determine probabilistic thresholds for combinatorial theorems hinges on proving a transference principle [39, 40, 71]. Suppose, for example, that one wishes to prove Turán's theorem for triangles within the random graph  $G_{n,p}$  for  $p \geq Cn^{-1/2}$ .



Then they show that asymptotically almost surely  $G_{n,p}$  has the following property. Any subgraph  $G$  of  $G_{n,p}$  may be modelled by a subgraph  $K$  of the complete graph on  $n$  vertices in such a way that the proportion of edges and triangles in the dense graph is close to the proportion of edges and triangles in the sparse graph. That is, if the sparse graph  $G$  contains  $c_1pn^2$  edges and  $c_2p^3n^3$  triangles then the dense model graph  $K$  should contain approximately  $c_1n^2$  edges and  $c_2n^3$  triangles.

To see that this implies Turán's theorem in the sparse graph, suppose that  $G$  is a subgraph of  $G_{n,p}$  with  $(\frac{1}{2} + \epsilon)p\binom{n}{2}$  edges. Then, provided the approximation is sufficiently good, the model graph  $K$  will have at least  $(\frac{1}{2} + \frac{\epsilon}{2})\binom{n}{2}$  edges. A robust version of Turán's theorem (as discussed in Section 1.3) then implies that  $K$  contains at least  $cn^3$  triangles for some  $c > 0$ . Provided again that the approximation is sufficiently good, this implies that the original graph  $G$  contains at least  $\frac{c}{2}p^3n^3$  triangles.

More generally, we have the following theorem (Corollary 9.7 in [17]) from which many such threshold results for graphs may be derived in a similar fashion. For any graph  $G$ , we let the function  $G: V(G)^2 \rightarrow \{0, 1\}$  be the characteristic function of  $G$ , given by  $G(x, y) = 1$  if  $xy \in E(G)$  and 0 otherwise. For any fixed graph  $H$  on  $k$  vertices, we also let

$$\mu_H(G) = N^{-k} \sum_{x_1, \dots, x_k} \prod_{ij \in E(H)} G(x_i, x_j)$$

be the normalized count of homomorphisms from  $H$  to  $G$ .

**Theorem 3.1.** *For any strictly balanced graph  $H$  and any  $\epsilon > 0$ , there exist positive constants  $C$  and  $\lambda$  such that if  $CN^{-1/m_2(H)} \leq p \leq \lambda$  then the following holds a.a.s. in the random graph  $G_{N,p}$ . For every subgraph  $G$  of  $G_{N,p}$ , there exists a subgraph  $K$  of  $K_N$  such that*

$$|p^{-e(H)}\mu_H(G) - \mu_H(K)| \leq \epsilon$$

and, for all pairs of disjoint vertex subsets  $U_1, U_2$  of  $V(K_N)$ ,

$$|p^{-1} \sum_{x,y} G(x, y) - \sum_{x,y} K(x, y)| \leq \epsilon N^2,$$

where the sums are taken over all  $x \in U_1$  and  $y \in U_2$ .

As stated, the result in [17] only gives the bound  $p^{-e(H)}\mu_H(G) \geq \mu_H(K) - \epsilon$ . This is all that is necessary for the applications given in that (and this) paper. However, the bound  $p^{-e(H)}\mu_H(G) \leq \mu_H(K) + \epsilon$  also follows from a more careful analysis.

To be more explicit, we need to say a little about the method used in [17]. For any collection of functions  $h_1, \dots, h_{e(H)}$ , we consider the function

$$\mu_H(h_1, \dots, h_{e(H)}) = N^{-k} \sum_{x_1, \dots, x_k} \prod_{ij \in E(H)} h_{ij}(x_i, x_j).$$

Let  $\gamma$  be the associated measure of the random graph  $G_{N,p}$ . We define this as being  $\gamma(x, y) = p^{-1}$  if  $xy$  is an edge of  $G_{N,p}$  and 0 otherwise. Note that a.a.s. the value of  $\|\gamma\|_1$  is  $1 + o(1)$ . Suppose that  $k$  and  $g$  are functions on the edge set of  $G_{N,p}$  with  $0 \leq k \leq 1$  and  $0 \leq g \leq \gamma$ . For example,  $g$  could be (and usually is) the characteristic function of a subgraph  $G$  of  $G_{N,p}$ , with each edge weighted by a factor of  $p^{-1}$ .

One of the main ideas of [17] was to define a norm  $\|\cdot\|$  with the property that if  $\|g_i - k_i\| = o(1)$  for all  $1 \leq i \leq e(H)$  then  $\mu_H(g_1, \dots, g_{e(H)}) \geq \mu_H(k_1, \dots, k_{e(H)}) - o(1)$ . Note that if  $G_i$  is a graph and  $g_i$  is the weighted characteristic function  $g_i = p^{-1}G_i$ , then

$$p^{-e(H)}\mu_H(G_1, \dots, G_{e(H)}) = \mu_H(g_1, \dots, g_{e(H)}) \geq \mu_H(k_1, \dots, k_{e(H)}) - o(1),$$

that is, the functions  $k_i$  serve as dense models for the  $G_i$  for the purposes of one-sided counting. The main result of [17] is that for any function  $g_i$  with  $0 \leq g_i \leq \gamma$  such a dense model function  $k_i$  with  $0 \leq k_i \leq 1$  exists. In particular, the constant function 1 serves as an appropriate model for the function  $\gamma$  corresponding to the random graph.

Suppose that  $k$  serves as a model for  $g$ . Then, by the triangle inequality,

$$\|(\gamma - g) - (1 - k)\| \leq \|\gamma - 1\| + \|g - k\| = o(1),$$

that is  $1 - k$  serves as a model for  $\gamma - g$ . It follows that, for any  $1 \leq i \leq e(H)$ ,

$$\mu_H(g, \dots, g, \gamma - g, \gamma, \dots, \gamma) \geq \mu_H(k, \dots, k, 1 - k, 1, \dots, 1) - o(1), \quad (21)$$

where in the  $\mu_H$  on the left-hand side the first  $i - 1$  terms are  $g$ , the  $i$ th term is  $\gamma - g$  and the remaining terms are  $\gamma$ . The same holds for the  $\mu_H$  on the right-hand side with  $g$  replaced by  $k$  and  $\gamma$  by 1. Note that, since  $\mu_H$  is additive in each variable,

$$\begin{aligned} \mu_H(g, \dots, g, g) &= \mu_H(g, \dots, g, \gamma) - \mu_H(g, \dots, g, \gamma - g) \\ &= \mu_H(g, \dots, g, \gamma, \gamma) - \mu_H(g, \dots, g, \gamma - g, \gamma) - \mu_H(g, \dots, g, \gamma - g) \\ &\vdots \\ &= \mu_H(\gamma, \dots, \gamma, \gamma) - \mu_H(\gamma - g, \dots, \gamma, \gamma) - \dots - \mu_H(g, \dots, g, \gamma - g). \end{aligned}$$

Therefore, by (21),

$$\begin{aligned} \mu_H(g, \dots, g, g) &\leq \mu_H(1, \dots, 1, 1) - \mu_H(1 - k, \dots, 1, 1) - \dots - \mu_H(k, \dots, k, 1 - k) + o(1) \\ &= \mu_H(k, \dots, k, k) + o(1). \end{aligned}$$

That is,  $|\mu_H(g) - \mu_H(k)| = o(1)$ . Here we used that  $\mu_H(\gamma, \dots, \gamma, \gamma) = \mu_H(1, \dots, 1, 1) + o(1)$ , which follows from standard tail estimates (see, for example, [44]).

To recover the statement of Theorem 3.1, where we refer to graphs rather than functions, we let  $g = p^{-1}G$ . This yields a function  $k$  with  $0 \leq k \leq 1$  such that  $|p^{-1}\mu_H(G) - \mu_H(k)| = o(1)$ . If we now choose a graph  $K$  randomly by picking each edge  $xy$  independently with probability  $k(x, y)$ , we will a.a.s. produce a graph  $K$  with  $H$ -count close to  $k$  (see, for example, the proof of Corollary 9.7 in [17]). Choosing such a graph, we have  $|p^{-e(H)}\mu_H(G) - \mu_H(K)| = o(1)$ , as required.

Because we are dealing with canonical homomorphisms of a graph  $H$  with  $k$  vertices to a  $k$ -partite graph  $G$  with vertex sets  $V_1, \dots, V_k$ , it would be quite useful to have another version of Theorem 3.1 which captures this situation. To this end, let  $H^*(V_1, \dots, V_k)$  be the class of graphs on vertex set  $V_1 \cup \dots \cup V_k$ , where  $V_1, \dots, V_k$  are disjoint sets, such that the only edges lie between sets  $V_i$  and  $V_j$  with  $ij \in E(H)$ . Then, for any  $G \in H^*(V_1, \dots, V_k)$ , we let

$$\mu_H^*(G) = N^{-k} \sum_{x_1 \in V_1, \dots, x_k \in V_k} \prod_{ij \in E(H)} G(x_i, x_j)$$

be the normalized count of canonical homomorphisms from  $H$  to  $G$ . The following theorem may be proved by a minor modification of the proof of Theorem 3.1.

**Theorem 3.2.** *For any strictly balanced graph  $H$  on  $k$  vertices and any  $\epsilon > 0$ , there exist positive constants  $C$  and  $\lambda$  such that if  $CN^{-1/m_2(H)} \leq p \leq \lambda$  then the following holds a.a.s. in the random graph  $G_{N,p}$ . For all disjoint vertex subsets  $V_1, \dots, V_k$  and every subgraph  $G$  of  $G_{N,p}$  in  $H^*(V_1, \dots, V_k)$ , there exists a subgraph  $K$  of  $K_N$  in  $H^*(V_1, \dots, V_k)$  such that*

$$|p^{-e(H)}\mu_H^*(G) - \mu_H^*(K)| \leq \epsilon \quad (22)$$

and, for all pairs of disjoint vertex subsets  $U_1, U_2$  of  $V(K_N)$ ,

$$|p^{-1} \sum_{x,y} G(x,y) - \sum_{x,y} K(x,y)| \leq \epsilon N^2, \quad (23)$$

where the sums are taken over all  $x \in U_1$  and  $y \in U_2$ .

In proving part (ii) of Theorem 1.6, we will use the following slight variant of the dense counting lemma, Lemma 1.3. We let  $\mathcal{G}(H, n, m, p, \theta, \epsilon)$  be defined in exactly the same way as  $\mathcal{G}(H, n, m, p, \epsilon)$ , except we now allow the number of edges between each pair  $V_i$  and  $V_j$  to be  $m \pm \theta pn^2$ .

**Lemma 3.3.** *For every graph  $H$  and every  $\delta > 0$ , there exist  $\theta, \epsilon > 0$  and an integer  $n_0$  such that for every  $n \geq n_0$ , every  $m$ , and every  $G \in \mathcal{G}(H, n, m, 1, \theta, \epsilon)$ ,*

$$G(H) = \left(\frac{m}{n^2}\right)^{e(H)} n^{v(H)} \pm \delta n^{v(H)}.$$

Part (ii) of Theorem 1.6 is now a relatively easy corollary of Theorem 3.2.

*Proof of part (ii) of Theorem 1.6.* Let  $H$  be a fixed graph on  $k$  vertices and  $d, \delta$  fixed positive constants. Choose  $\theta$  and  $\epsilon_{3.3}$  so that the conclusion of Lemma 3.3 holds with  $\delta_{3.3} = \frac{d^{e(H)}\delta}{4}$ . We let  $\epsilon = \frac{\epsilon_{3.3}}{6}$  and, for a fixed positive constant  $\eta$ ,

$$\epsilon_{3.2} = \min\left(\frac{\eta^2 \epsilon_{3.3}^3}{3}, \eta^2 \theta, \frac{d^{e(H)}\delta}{4}\right)$$

and then choose  $C$  and  $\lambda$  so that the conclusion of Theorem 3.2 holds with  $\epsilon_{3.2}$ .

Suppose now that  $CN^{-1/m_2(H)} \leq p \leq \lambda$  and  $G_{N,p}$  satisfies the conclusion of Theorem 3.2. Let  $G$  be a subgraph of  $G_{N,p}$  from the set  $\mathcal{G}(H, n, m, p, 2\epsilon)$  with vertex sets  $V_1, \dots, V_k$ , each of size  $n \geq \eta N$ . Let  $K$  be the dense graph given by Theorem 3.2. If  $V'_i \subseteq V_i$  and  $V'_j \subseteq V_j$  the triangle inequality tells us that  $|d_K(V'_i, V'_j) - d_K(V_i, V_j)|$  is at most

$$|d_K(V'_i, V'_j) - p^{-1}d_G(V'_i, V'_j)| + p^{-1}|d_G(V'_i, V'_j) - d_G(V_i, V_j)| + |p^{-1}d_G(V_i, V_j) - d_K(V_i, V_j)|.$$

Suppose that  $|V'_i| \geq \epsilon_{3.3}|V_i|$  and  $|V'_j| \geq \epsilon_{3.3}|V_j|$ . Then, since  $G$  is  $(2\epsilon, p)$ -regular between  $V_i$  and  $V_j$  and  $\epsilon_{3.3} \geq 2\epsilon$ , the middle term is at most  $2\epsilon$ . By (23), since  $V'_i$  and  $V'_j$  are disjoint sets of size at least  $\epsilon_{3.3}\eta N$ , the first and third terms are each at most  $(\epsilon_{3.3}\eta)^{-2}\epsilon_{3.2}$ . We therefore see that

$$|d_K(V'_i, V'_j) - d_K(V_i, V_j)| \leq 2\epsilon + 2(\epsilon_{3.3}\eta)^{-2}\epsilon_{3.2} \leq \epsilon_{3.3}.$$

Therefore,  $K$  is  $\epsilon_{3.3}$ -regular. Note also that the number of edges between  $V_i$  and  $V_j$  in  $K$  is

$$p^{-1}m \pm \epsilon_{3.2}N^2 = p^{-1}m \pm \theta n^2,$$

since  $\epsilon_{3.2} \leq \eta^2\theta$ . Applying Lemma 3.3, we see that for  $n$  sufficiently large

$$K(H) = \left(\frac{p^{-1}m}{n^2}\right)^{e(H)} n^{v(H)} \pm \delta_{3.3}n^{v(H)} = \left(1 \pm \frac{\delta}{4}\right) \left(\frac{p^{-1}m}{n^2}\right)^{e(H)} n^{v(H)},$$

where we used that  $m \geq dpn^2$ . Therefore, by (22),

$$\begin{aligned} G(H) = p^{e(H)}K(H) \pm \epsilon_{3.2}p^{e(H)}n^{v(H)} &= \left(1 \pm \frac{\delta}{4}\right) \left(\frac{m}{n^2}\right)^{e(H)} \pm \epsilon_{3.2}p^{e(H)}n^{v(H)} \\ &= \left(1 \pm \frac{\delta}{2}\right) \left(\frac{m}{n^2}\right)^{e(H)} n^{v(H)}, \end{aligned} \quad (24)$$

where we used that  $\epsilon_{3,2} \leq \frac{d^{e(H)}\delta}{4}$ .

Suppose now that  $p > \lambda$  and fix a  $G \subseteq G_{N,p}$  in  $\mathcal{G}(H, n, m, p, \epsilon)$  with  $n \geq \eta N$  and  $m \geq dpn^2$ . Form a random subgraph  $G'$  of  $G$  by choosing  $m' = \lambda m$  edges in each pair  $G(V_i, V_j)$  uniformly at random. Clearly,

$$\mathbb{E}[G'(H)] = \lambda^{e(H)} G(H). \quad (25)$$

A standard application of Hoeffding's inequality (see [34, Lemma 4.3]) proves that with probability at least  $1 - \exp(-cm') \geq 1 - \delta/2$ , the graph  $G'$  is in  $\mathcal{G}(H, n, m', p', 2\epsilon)$ , where  $p' = \lambda p$ , and hence, by (24) and (25),

$$G(H) \geq \lambda^{-e(H)} \left(1 - \frac{\delta}{2}\right)^2 \left(\frac{m'}{n^2}\right)^{e(H)} n^{v(H)} \geq (1 - \delta) \left(\frac{m}{n^2}\right)^{e(H)} n^{v(H)}.$$

We may assume, since it happens a.a.s., that the total number of copies of  $H$  in  $G_{N,p}$  does not exceed  $2p^{e(H)} N^{v(H)}$  (see, for example, [44]). Therefore,

$$G(H) \leq \lambda^{-e(H)} \cdot [\Pr(G' \in \mathcal{G}(H, n, m', p', 2\epsilon)) \cdot \left(1 + \frac{\delta}{2}\right) \left(\frac{m'}{n^2}\right)^{e(H)} n^{v(H)} + \Pr(G' \notin \mathcal{G}(H, n, m', p', 2\epsilon)) \cdot 2p^{e(H)} N^{v(H)}].$$

Since  $n \geq \eta N$ ,  $m \geq dpn^2$  and  $\Pr(G' \notin \mathcal{G}(H, n, m', p', 2\epsilon)) \leq \exp(-cm') \leq \delta(\lambda d)^{e(H)} \eta^{v(H)}/4$  if  $n$  is sufficiently large, it follows that

$$G(H) \leq (1 + \delta) \left(\frac{m}{n^2}\right)^{e(H)} n^{v(H)}.$$

The result follows.  $\square$

We note that the second case, where  $p > \lambda$ , also follows as an immediate corollary of the counting lemma for pseudorandom graphs proved in [16]. Indeed, this result is already strong enough to imply a counting lemma down to densities of about  $N^{-c/\Delta(H)}$ , where  $\Delta(H)$  is the maximum degree of  $H$ . In addition, if one is only interested in one-sided counting, that is, in showing that  $G(H) \geq (1 - \delta)(m/n^2)^{e(H)} n^{v(H)}$ , then the results of [16] apply for  $p \geq N^{-c/d(H)}$ , where  $d(H)$  is the degeneracy of  $H$ .

## 4 Applications

### 4.1 Preliminaries

After applying the sparse regularity lemma, it is usually helpful to clean up the the regular partition, removing all edges which are not contained in a dense regular pair. The following standard lemma, which incorporates both the regularity lemma and this cleaning process, is sufficient for our purposes. Recall that a graph is  $(\eta, p, D)$ -upper-uniform if for all disjoint subsets  $U_1$  and  $U_2$  with  $|U_1|, |U_2| \geq \eta|V(G)|$ , the density of edges between  $U_1$  and  $U_2$  satisfies  $d(U_1, U_2) \leq Dp$ . Here, for the sake of clarity of presentation, we make the additional assumption in this definition that if  $U_1 = U_2 = U$ , then  $e_G(U) \leq Dp \binom{|U|}{2}$ .

**Proposition 4.1.** *For every  $\epsilon, D > 0$  and every positive integer  $t_0$ , there exist  $\eta > 0$  and a positive integer  $T$  such that, for every  $d > 0$ , every graph  $G$  with at least  $t_0$  vertices which is  $(\eta, p, D)$ -upper-uniform contains a subgraph  $G'$  with*

$$e(G') \geq e(G) - \left(\frac{D}{t_0} + 2D\epsilon + d\right) \frac{pn^2}{2}$$

that admits an equipartition  $V_1, \dots, V_t$  of its vertex set into  $t_0 \leq t \leq T$  pieces such that the following conditions hold.

1. There are no edges of  $G'$  within  $V_i$  for any  $1 \leq i \leq t$ .
2. Every non-empty graph  $G'(V_i, V_j)$  is  $(\epsilon, p)$ -regular and has at least  $dp|V_i||V_j|$  edges.

*Proof.* Fix  $\epsilon$ ,  $D$ , and  $t_0$  as in the statement of the proposition and let  $T = T_{1.2}(\epsilon, D, t_0)$  and  $\eta = \min\{\eta_{1.2}(\epsilon, D, t_0), \frac{1}{2T}\}$ . Let  $d > 0$ , fix a  $G$  as above, and apply the sparse regularity lemma, Theorem 1.2, to obtain an  $(\epsilon, p)$ -regular partition  $V_1, \dots, V_t$  of the vertices of  $G$  into  $t_0 \leq t \leq T$  pieces. Let us delete from  $G$  all edges that are contained in:

- one of the sets  $V_1, \dots, V_t$ ,
- one of the at most  $\epsilon t^2$  pairs  $(V_i, V_j)$  that are not  $(\epsilon, p)$ -regular, or
- one of the pairs  $(V_i, V_j)$  that have fewer than  $dp|V_i||V_j|$  edges.

Denote the resulting graph by  $G'$ . Since  $G$  is  $(\eta, p, D)$ -upper-uniform, we have  $e_G(V_i) \leq Dp\frac{1}{2}(\frac{n}{t})^2$  and  $e_G(V_i, V_j) \leq Dp(\frac{n}{t})^2$  for all  $i$  and  $j$ . It follows that

$$e(G) - e(G') \leq t \cdot Dp\frac{1}{2} \left(\frac{n}{t}\right)^2 + \epsilon t^2 \cdot Dp \left(\frac{n}{t}\right)^2 + \binom{t}{2} \cdot dp \left(\frac{n}{t}\right)^2 \leq \left(\frac{D}{t_0} + 2D\epsilon + d\right) \frac{pn^2}{2}. \quad \square$$

In the proofs of our applications, we will need a version of the main theorem which allows us to have different densities between different pairs of vertex sets. To this end, given a graph  $H$  on the vertex set  $\{1, \dots, k\}$  and a sequence  $\mathbf{m} = (m_{ij})_{ij \in E(H)}$  of integers, we denote by  $\mathcal{G}(H, n, \mathbf{m}, p, \epsilon)$  the collection of all graphs  $G$  obtained in the following way. The vertex set of  $G$  is a disjoint union  $V_1 \cup \dots \cup V_k$  of sets of size  $n$ . For each edge  $ij \in E(H)$ , we add to  $G$  an  $(\epsilon, p)$ -regular bipartite graph with  $m_{ij}$  edges between the pair  $(V_i, V_j)$ . These are the only edges of  $G$ . As before, for any  $G \in \mathcal{G}(H, n, \mathbf{m}, p, \epsilon)$ , let us denote by  $G(H)$  the number of canonical copies of  $H$  in  $G$ .

**Proposition 4.2.** *For every graph  $H$  and every  $\delta, d > 0$ , there exist  $\epsilon, \xi > 0$  with the following property. For every  $\eta > 0$ , there is a  $C > 0$  such that if  $p \geq CN^{-1/m_2(H)}$  then a.a.s. the following holds in  $G_{N,p}$ :*

- (i) *For every  $n \geq \eta N$ ,  $\mathbf{m}$  with  $m_{ij} \geq dpn^2$  for all  $ij \in E(H)$  and every subgraph  $G$  of  $G_{N,p}$  in  $\mathcal{G}(H, n, \mathbf{m}, p, \epsilon)$ ,*

$$G(H) \geq \xi \left( \prod_{ij \in E(H)} \frac{m_{ij}}{n^2} \right) n^{v(H)}. \quad (26)$$

- (ii) *Moreover, if  $H$  is strictly balanced, that is, if  $m_2(H) > m_2(H')$  for every proper subgraph  $H'$  of  $H$ , then*

$$G(H) = (1 \pm \delta) \left( \prod_{ij \in E(H)} \frac{m_{ij}}{n^2} \right) n^{v(H)}. \quad (27)$$

*Proof.* Fix  $H$ ,  $\delta$ , and  $d$  as in the statement of the proposition. We may assume that  $\Delta(H) \geq 2$  (and hence  $m_2(H) \geq 1$ ) as otherwise the assertion of the proposition is trivial. Let  $\epsilon = \epsilon_{1.6}(H, \delta/2, d)/2$  and  $\xi = \xi_{1.6}(H, \delta/2, d)/2$ . Moreover, fix some  $\eta > 0$ , let  $C$  be a sufficiently large positive constant, and suppose that  $p \geq CN^{-1/m_2(H)}$ . First, we will show that if  $G_{N,p}$  satisfies part (i) of Theorem 1.6,

which happens a.a.s., then it also satisfies part (i) of Proposition 4.2. To this end, let  $n \geq \eta N$ , let  $\mathbf{m}$  satisfy  $m_{ij} \geq dpn^2$  for all  $ij \in E(H)$ , and fix a  $G \subseteq G_{N,p}$  in  $\mathcal{G}(H, n, \mathbf{m}, p, \epsilon)$ . Form a random subgraph  $G'$  of  $G$  by choosing  $m = dpn^2$  edges in each pair  $G(V_i, V_j)$  uniformly at random. Clearly,

$$\mathbb{E}[G'(H)] = \left( \prod_{ij \in E(H)} \frac{m}{m_{ij}} \right) \cdot G(H). \quad (28)$$

A standard application of Hoeffding's inequality (see [34, Lemma 4.3]) proves that with probability at least  $1 - \exp(-cm)$ , where  $c > 0$  is an absolute constant, the graph  $G'$  is in  $\mathcal{G}(H, n, m, p, 2\epsilon)$ . Hence, if  $G_{N,p}$  satisfies part (i) of Theorem 1.6, then with probability at least  $1/2$ ,

$$G'(H) \geq 2\xi \left( \frac{m^2}{n} \right)^{e(H)} n^{v(H)}$$

and therefore (28) implies that

$$G(H) = \left( \prod_{ij \in E(H)} \frac{m_{ij}}{m} \right) \cdot \mathbb{E}[G'(H)] \geq \left( \prod_{ij \in E(H)} \frac{m_{ij}}{m} \right) \cdot \xi \left( \frac{m}{n^2} \right)^{e(H)} n^{v(H)},$$

as claimed.

Now, we show that if  $G_{N,p}$  satisfies part (ii) of Theorem 1.6 and the total number of copies of  $H$  in  $G_{N,p}$  does not exceed  $2p^{e(H)}N^{v(H)}$ , which happens a.a.s. (see, for example, [44]), then part (ii) of this proposition is also satisfied. To this end, fix  $\mathbf{m}$  and  $G$  as above, recall the definition of  $G'$ , and observe that since with probability at least  $1 - \delta/2$ , the graph  $G'$  is in  $\mathcal{G}(H, n, m, p, 2\epsilon)$ , then by (2) and (28),

$$G(H) \geq \left( \prod_{ij \in E(H)} \frac{m_{ij}}{m} \right) \cdot \left(1 - \frac{\delta}{2}\right)^2 \left( \frac{m}{n^2} \right)^{e(H)} n^{v(H)} \geq (1 - \delta) \left( \prod_{ij \in E(H)} \frac{m_{ij}}{n^2} \right) n^{v(H)}.$$

On the other hand, we also have

$$G(H) \leq \left( \prod_{ij \in E(H)} \frac{m_{ij}}{m} \right) \cdot [\Pr(G' \in \mathcal{G}(H, n, m, p, 2\epsilon)) \cdot \left(1 + \frac{\delta}{2}\right) \left( \frac{m}{n^2} \right)^{e(H)} n^{v(H)} + \Pr(G' \notin \mathcal{G}(H, n, m, p, 2\epsilon)) \cdot 2p^{e(H)}N^{v(H)}].$$

Since  $n \geq \eta N$ ,  $m \geq dpn^2$  and  $\Pr(G' \notin \mathcal{G}(H, n, m, p, 2\epsilon)) \leq \exp(-cm) \leq \delta d^{e(H)} \eta^{v(H)} / 4$  if  $n$  is sufficiently large, it follows that

$$G(H) \leq (1 + \delta) \left( \prod_{ij \in E(H)} \frac{m_{ij}}{n^2} \right) n^{v(H)}. \quad \square$$

## 4.2 The sparse removal lemma

Let  $\delta > 0$  and let  $H$  be an arbitrary (not necessarily balanced) graph. The proof of Theorem 1.9 is a classical application of the regularity method. We start by defining a range of constants. For the sake of brevity, we let  $k = v(H)$ . Furthermore, let  $d = \delta/2$ ,  $t_0 = 2/d$ ,  $D = 2$ ,  $\epsilon' =$

$\min\{\epsilon_{4.2}(H, d/2), \delta/8\}$ ,  $\xi = \xi_{4.2}(H, d/2)$ ,  $T = T_{4.1}(\epsilon'/k, D, t_0)$ , and  $\eta = \min\{\eta_{4.1}(\epsilon'/k, D, t_0), 1/(kT)\}$ . Finally, let  $\epsilon = \xi(d/2)^{e(H)}(kT)^{-v(H)}$  and  $C = C_{4.2}(H, d/2, \eta)$ . Suppose that  $p \geq Cn^{-1/m_2(H)}$ . We will show that the sparse  $H$ -removal lemma holds in  $G_{n,p}$  a.a.s., that is, that every subgraph of  $G_{n,p}$  with fewer than  $\epsilon p^{e(H)}n^{v(H)}$  copies of  $H$  can be made  $H$ -free by removing from it at most  $\delta pn^2$  edges. It clearly suffices to show that part (i) of Proposition 4.2 and  $(\eta, p, D)$ -upper-uniformity, which holds in  $G_{n,p}$  a.a.s., imply the above property. To this end, assume that  $G_{n,p}$  is  $(\eta, p, D)$ -upper-uniform and that part (i) of Proposition 4.2 holds in  $G_{n,p}$ . Let  $G \subseteq G_{n,p}$  be a subgraph with fewer than  $\epsilon p^{e(H)}n^{v(H)}$  copies of  $H$  and let  $G'$  be a subgraph of  $G$  satisfying the assertion of Proposition 4.1 with  $\epsilon_{4.1} = \epsilon'/k$ . By our choice of parameters,

$$e(G) - e(G') \leq \left( \frac{D}{t_0} + \frac{2D\epsilon'}{k} + d \right) \frac{pn^2}{2} \leq \left( 2d + \frac{\delta}{2} \right) \frac{pn^2}{2} < \delta pn^2.$$

We claim that  $G'$  is  $H$ -free. Since all edges of  $G'$  lie in  $(\frac{\epsilon'}{k}, p)$ -regular pairs with edge density at least  $dp$ , if  $G'$  contained a copy of  $H$ , there would be a graph  $H'$  with  $k' \leq k$  vertices which is a homomorphic image of  $H$ , pairwise disjoint sets  $V'_1, \dots, V'_{k'}$  and a sequence  $\mathbf{m}' = (m'_{ij})_{ij \in E(H')}$  with  $m'_{ij} \geq dp(\frac{n}{t})^2$  such that  $G'[V'_1 \cup \dots \cup V'_{v(H')}] \in \mathcal{G}(H', \frac{n}{t}, \mathbf{m}', p, \epsilon'/k)$ . Consequently, there would be pairwise disjoint sets  $V_1, \dots, V_k$  and a sequence  $\mathbf{m} = (m_{ij})_{ij \in E(H)}$  with  $m_{ij} \geq \frac{dp}{2}(\frac{n}{kt})^2$  such that  $G'[V_1 \cup \dots \cup V_k] \in \mathcal{G}(H, \frac{n}{kt}, \mathbf{m}, p, \epsilon')$ . One can obtain such  $V_1, \dots, V_k$  by arbitrarily dividing each  $V'_i$  into  $k$  parts and choosing  $k$  of these parts according to the homomorphism from  $H$  to  $H'$ . It would follow that the number of copies of  $H$  in  $G'$ , and therefore also in  $G$ , would exceed

$$\xi \left( \prod_{ij \in E(H)} \frac{m_{ij}}{(n/(kt))^2} \right) \left( \frac{n}{kt} \right)^{v(H)} \geq \xi \left( \frac{dp}{2} \right)^{e(H)} \left( \frac{n}{kt} \right)^{v(H)} \geq \epsilon p^{e(H)}n^{v(H)},$$

a contradiction.

### 4.3 The clique density theorem

To begin, we note that if  $W$  is a weighted graph on  $n$  vertices with  $0 \leq W \leq 1$  for which  $\sum_{x,y} W(x,y) \geq \rho \binom{n}{2}$ , then, for any  $\theta$  and  $n$  sufficiently large depending on  $\theta$ ,

$$\sum_{x_1, \dots, x_k} \prod_{1 \leq a < b \leq k} W(x_a, x_b) \geq (g_k(\rho) - \theta)n^k.$$

This follows from choosing a random graph  $G$ , picking each edge  $xy$  independently with probability  $W(x,y)$ . The resulting graph will, with high probability, have a similar count of edges and  $K_k$ s to the weighted graph  $W$  (see, for example, Corollary 9.7 in [17]). The result then follows by applying the clique density theorem to the graph  $G$  (and using the fact that  $g_k(\rho)$  is uniformly continuous).

Let  $k \geq 3$  and  $\epsilon > 0$ . We start by defining constants. We choose  $t_1$  such that, for  $t \geq t_1$  and all  $\rho$ , any weighted graph  $W$  on  $t$  vertices with  $\sum_{x,y} W(x,y) \geq \rho \binom{t}{2}$  also satisfies

$$\sum_{i_1, \dots, i_k} \prod_{1 \leq a < b \leq k} W(i_a, i_b) \geq \left( g_k(\rho) - \frac{\epsilon}{4} \right) t^k.$$

Since  $g_k(\rho)$  is uniformly continuous in  $\rho$ , we may choose  $\delta'$  such that  $|g_k(\rho \pm \delta') - g_k(\rho)| \leq \frac{\epsilon}{4}$ . Let  $\delta = \min\{\delta', \frac{\epsilon}{2}\}$ ,  $d = \delta/16$ ,  $t_0 = \max\{t_1, 2/d\}$ ,  $D = 2$ ,  $\epsilon' = \min\{\epsilon_{4.2}(H, \delta, d), \delta/32\}$ ,  $T = T_{4.1}(\epsilon', D, t_0)$ , and  $\eta = \min\{\eta_{4.1}(\epsilon', D, t_0), 1/T\}$ . Finally, let  $C = C_{4.2}(H, \delta, d, \eta)$ .

Suppose that  $p \geq Cn^{-2/(k+1)}$ . Since  $(\eta, p, D)$ -upper-uniformity holds a.a.s. in  $G_{n,p}$ , we will assume that it is satisfied. Let  $G' \subseteq G_{n,p}$  be a subgraph of  $G_{n,p}$  of relative density  $\rho$ , that is, with  $\rho p \binom{n}{2}$  edges. By Proposition 4.1 with  $\epsilon_{4,1} = \epsilon'$ , we get an equipartition  $V_1, \dots, V_t$  of the vertex set of  $G'$  into  $t_0 \leq t \leq T$  pieces and a subgraph  $G''$  of  $G'$  all of whose edges lie in  $(\epsilon', p)$ -regular pairs with edge density at least  $dp$  and which satisfies

$$e(G') - e(G'') \leq \left( \frac{D}{t_0} + 2D\epsilon' + d \right) \frac{pn^2}{2} \leq \left( 2d + \frac{\delta}{8} \right) \frac{pn^2}{2} < \frac{\delta}{4} pn^2.$$

Consider the reduced weighted graph  $R$  on vertex set  $[t]$  with the weight of edge  $ij$  given by

$$R(i, j) = \min \left( \frac{e_{G''}(V_i, V_j)}{p|V_i||V_j|}, 1 \right).$$

Note that a.a.s. the random graph  $G_{n,p}$  satisfies  $e_G(U, V) \leq p|U||V| + \frac{\delta}{4T^2}pn^2$  for all disjoint subsets  $U$  and  $V$ . This in turn implies that  $e_{G''}(V_i, V_j) \leq p|V_i||V_j| + \frac{\delta}{4T^2}pn^2$ . Hence,

$$R(i, j) \geq \frac{e_{G''}(V_i, V_j) - \frac{\delta}{4}p\left(\frac{n}{T}\right)^2}{p|V_i||V_j|}.$$

Therefore, since  $e(G') \geq \rho p \binom{n}{2}$ ,

$$\sum_{1 \leq i < j \leq t} R(i, j) \geq \frac{e(G'') - \frac{\delta}{4}pn^2}{p\left(\frac{n}{t}\right)^2} \geq \frac{e(G') - \frac{\delta}{2}pn^2}{p\left(\frac{n}{t}\right)^2} \geq (\rho - \delta) \binom{t}{2}.$$

Therefore, by the choice of  $t_1$  and  $\delta'$ , we have that

$$\sum_{i_1, \dots, i_k} \prod_{1 \leq a < b \leq k} R(i_a, i_b) \geq \left( g_k(\rho - \delta) - \frac{\epsilon}{4} \right) t^k \geq \left( g_k(\rho) - \frac{\epsilon}{2} \right) t^k.$$

Now, for any particular  $1 \leq i_1 < \dots < i_k \leq t$ , consider the sets  $V_{i_1}, \dots, V_{i_k}$ . Then  $G''[V_{i_1} \cup \dots \cup V_{i_k}]$  is an element of  $\mathcal{G}(K_k, \frac{n}{t}, \mathbf{m}, p, \epsilon')$  with  $m_{i_a i_b} \geq R(i_a, i_b)p\left(\frac{n}{t}\right)^2$ . By part (ii) of Proposition 4.2 and the choice of  $\epsilon'$ ,  $\eta$ , and  $C$ , it follows that the number of copies of  $K_k$  between the sets  $V_{i_1}, \dots, V_{i_k}$  is at least

$$(1 - \delta) \prod_{1 \leq a < b \leq k} R(i_a, i_b) \cdot p^{\binom{k}{2}} \left( \frac{n}{t} \right)^k.$$

Adding over all choices of  $i_1, \dots, i_k$  gives

$$\begin{aligned} \mu_{G''}(K_k) &\geq (1 - \delta) \sum_{i_1, \dots, i_k} \prod_{1 \leq a < b \leq k} R(i_a, i_b) \cdot p^{\binom{k}{2}} \left( \frac{n}{t} \right)^k \\ &\geq \left( 1 - \frac{\epsilon}{2} \right) \left( g_k(\rho) - \frac{\epsilon}{2} \right) t^k p^{\binom{k}{2}} \left( \frac{n}{t} \right)^k \geq (g_k(\rho) - \epsilon) p^{\binom{k}{2}} n^k, \end{aligned}$$

as required.

#### 4.4 The Hajnal-Szemerédi theorem

We start this section with a brief outline of the proof of Theorem 1.12. Fix some  $k \geq 3$  and  $\gamma > 0$ . Given a subgraph  $G'$  of  $G_{n,p}$  with  $\delta(G') \geq (1 - \frac{1}{k} + \gamma)pn$ , we apply the regularity lemma to  $G'$  to obtain an  $(\epsilon, p)$ -regular partition of its vertex set into  $t$  parts. We then construct an auxiliary graph



$R$  on the vertex set  $[t]$  whose edges correspond to  $(\epsilon, p)$ -regular pairs of non-negligible density in the regular partition of  $G'$ . Since most of the edges of  $G'$  lie in such dense and regular pairs, the assumption on the minimum degree of  $G'$  implies that  $R$  contains an almost spanning subgraph  $R'$  with minimum degree at least  $(1 - \frac{1}{k})t$ . By the Hajnal-Szemerédi theorem,  $R'$  contains a  $K_k$ -factor. Finally, a fairly straightforward application of Theorem 1.6 implies that each of the cliques in this  $K_k$ -factor corresponds to a  $K_k$ -packing in  $G'$  that covers most of the vertices in the  $k$  parts of the regular partition that form this clique. As is typically the case with arguments employing the regularity method, the details of the argument are somewhat intricate.

We start the actual proof by fixing several constants. Let

$$\beta = \min \left\{ \frac{\gamma}{2}, \frac{1}{2k} \right\}, \quad \gamma' = \frac{\beta^2}{14}, \quad \text{and} \quad t_0 = \left\lceil \frac{1}{\gamma'} \right\rceil.$$

Let  $\epsilon'$  be the constant obtained by invoking Proposition 4.2 with  $d_{4,2} = \gamma'$ . Let  $\epsilon = \min\{\beta\epsilon', \gamma'/2\}$  and  $D = 1 + \gamma'$ . Furthermore, let  $T = T_{4.1}(\epsilon, D, t_0)$ , let  $\eta = \min\{\eta_{4.1}(\epsilon, D, t_0), 1/T\}$ , and let  $\eta' = \beta/T$ . Assume that  $p \geq Cn^{-1/m_2(H)}$  for some large constant  $C$  so that a.a.s. the random graph  $G_{n,p}$  is  $(\eta, p, D)$ -upper-uniform and every subgraph  $G$  of  $G_{n,p}$  in  $\mathcal{G}(K_k, n', \mathbf{m}, p, \epsilon')$ , where  $n' \geq \eta'n$  and  $m_{ij} \geq \gamma'p(n')^2$  for all  $ij \in E(K_k)$ , contains a canonical copy of  $K_k$ . We shall show that, conditioned on the above two events, every  $G' \subseteq G_{n,p}$  with  $\delta(G') \geq (1 - \frac{1}{k} + \gamma)$  contains a  $K_k$ -packing covering all but at most  $\gamma n$  vertices.

Fix a  $G'$  as above and apply Proposition 4.1 with  $d_{4,1} = 2\gamma'$  to obtain an equipartition  $V_1, \dots, V_t$  of the vertex set of  $G'$  into  $t_0 \leq t \leq T$  pieces and a subgraph  $G''$  of  $G'$  all of whose edges lie in  $(\epsilon, p)$ -regular pairs with edge density at least  $2\gamma'p$  and which satisfies

$$e(G') - e(G'') \leq \left( \frac{D}{t_0} + 2D\epsilon + 2\gamma' \right) \frac{pn^2}{2} \leq 3\gamma'pn^2.$$

Let  $R$  be the graph on the vertex set  $[t]$  whose edges are those pairs  $ij$  such that the bipartite graph  $G''(V_i, V_j)$  is non-empty. Recall that each such bipartite graph is  $(\epsilon, p)$ -regular and has at least  $2\gamma'p|V_i||V_j|$  edges.

**Claim 4.3.** *The cluster graph  $R$  contains a subgraph  $R'$  with  $\delta(R') \geq (1 - \frac{1}{k})t$  and  $t' \geq (1 - \beta)t$  vertices for some  $t'$  divisible by  $k$ .*

*Proof.* We construct such a graph  $R'$  greedily by sequentially removing from  $R$  vertices of degree smaller than  $(1 - \frac{1}{k})t + k$  and at most  $k - 1$  further vertices in order to guarantee that  $k$  divides  $t'$ . If this process terminates before we delete from  $R$  more than  $\beta t - k$  vertices, then we will arrive at a graph  $R'$  with the desired properties. Otherwise,  $R$  contains at least  $\beta t - k$  vertices with degree at most  $(1 - \frac{1}{k} + \beta)t$ . Denote this set by  $X$  and observe that

$$\begin{aligned} \sum_{i \in X} \sum_{v \in V_i} \deg_{G'}(v) &\leq \sum_{i \in X} \sum_{v \in V_i} \deg_{G''}(v) + 2(e(G') - e(G'')) = \sum_{i \in X} \sum_{j \neq i} e_{G''}(V_i, V_j) + 2(e(G') - e(G'')) \\ &\leq |X| \left( 1 - \frac{1}{k} + \beta \right) t \cdot Dp \left( \frac{n}{t} \right)^2 + 6\gamma'pn^2 \leq \frac{|X|}{t} \left( 1 - \frac{1}{k} + \beta + \gamma' + \frac{12\gamma'}{\beta} \right) pn^2 \\ &< \frac{|X|}{t} \left( 1 - \frac{1}{k} + 2\beta \right) pn^2. \end{aligned}$$

On the other hand,

$$\sum_{i \in X} \sum_{v \in V_i} \deg_{G'}(v) \geq |X| \cdot \frac{n}{t} \cdot \delta(G') \geq \frac{|X|}{t} \left( 1 - \frac{1}{k} + \gamma \right) pn^2,$$

a contradiction, as  $\gamma \geq 2\beta$ . □

By the Hajnal-Szemerédi theorem, Theorem 1.11, the graph  $R'$  contains a  $K_k$ -factor. Hence, it suffices to show that each subgraph of  $G''$  induced by sets  $V_{i_1}, \dots, V_{i_k}$ , where  $i_1, \dots, i_k \in [t]$  form a copy  $K_k$  in  $R$ , contains a  $K_k$ -packing covering at least  $(1 - \beta)$ -fraction of its vertices. Indeed, since the  $K_k$ -factor in  $R'$  covers at least a  $(1 - \beta)$ -proportion of all the vertices of the cluster graph  $R$  and, as we show below, each of its cliques  $i_1, \dots, i_k$  corresponds to a  $K_k$ -packing covering at least a  $(1 - \beta)$ -proportion of the vertices in  $V_{i_1} \cup \dots \cup V_{i_k}$ , Theorem 1.12 will easily follow, as  $(1 - \beta)^2 > 1 - \gamma$ .

**Claim 4.4.** *Suppose that  $i_1, \dots, i_k \in [t]$  induce a copy of  $K_k$  in  $R$ . Then the graph  $G''[V_{i_1} \cup \dots \cup V_{i_k}]$  contains a  $K_k$ -packing covering all but at most  $\beta \frac{n}{t}$  vertices in each of  $V_{i_1}, \dots, V_{i_k}$ .*

*Proof.* For simplicity, assume that  $i_1 = 1, \dots, i_k = k$ . It will be enough to show that for every choice of  $W_1 \subseteq V_1, \dots, W_k \subseteq V_k$  with  $n' = |W_1| = \dots = |W_k| \geq \beta \frac{n}{t}$ , the graph  $G''[W_1 \cup \dots \cup W_k]$  contains a canonical copy of  $K_k$ . To this end, observe that this graph belongs to  $\mathcal{G}(K_k, n', \mathbf{m}, p, \epsilon')$  for some  $\mathbf{m} = (m_{ij})_{ij \in E(K_k)}$  with  $m_{ij} \geq \gamma' p (n')^2$  for all  $ij \in E(K_k)$ . Indeed, since for each  $ij \in E(K_k)$ , the graph  $G''(V_i, V_j)$  is  $(\epsilon, p)$ -regular and has at least  $2\gamma' p |V_i||V_j|$  edges, it follows that the graph  $G''(W_i, W_j)$  is  $(\epsilon', p)$ -regular and has at least  $\gamma' p |V_i||V_j|$  edges. The claim now follows.  $\square$

## 4.5 The Andrásfai-Erdős-Sós theorem

Fix a graph  $H$  and  $\gamma > 0$ . We start by fixing several constants. Choose  $t_1$  so that, for any  $t \geq t_1$ , the Andrásfai-Erdős-Sós theorem [8] (or rather its generalization due to Alon and Sudakov [6]) holds in the sense that any  $H$ -free graph on  $t$  vertices with minimum degree at least  $\left(1 - \frac{3}{3\chi(H)-4} + \frac{\gamma}{2}\right)t$  may be made  $(\chi(H) - 1)$ -partite by removing at most  $\frac{\gamma}{2}t^2$  edges. Let

$$\beta = \frac{\gamma}{4}, \quad \gamma' = \frac{\beta^2}{14}, \quad \text{and} \quad t_0 = \max\{2t_1, \lceil 1/\gamma' \rceil\}.$$

Let  $\epsilon'$  be the constant obtained by invoking Proposition 4.2 with  $d_{4.2} = \gamma'$ . Let  $\epsilon = \min\{\epsilon', \gamma'/2\}$  and  $D = 1 + \gamma'$ . Furthermore, let  $T = T_{4.1}(\epsilon, D, t_0)$  and  $\eta = \min\{\eta_{4.1}(\epsilon, D, t_0), 1/T\}$ . Assume that  $p \geq Cn^{-1/m_2(H)}$  for some large constant  $C$  so that a.a.s. the random graph  $G_{n,p}$  is  $(\eta, p, D)$ -upper uniform and every subgraph  $G$  of  $G_{n,p}$  in  $\mathcal{G}(H, n', \mathbf{m}, p, \epsilon)$ , where  $n' \geq \eta n$  and  $m_{ij} \geq \gamma' p (n')^2$  for all  $ij \in E(H)$ , contains a canonical copy of  $H$ . We shall show that, conditioned on the above two events, every  $H$ -free subgraph  $G' \subseteq G_{n,p}$  with  $\delta(G') \geq \left(1 - \frac{3}{3\chi(H)-4} + \gamma\right)pn$  may be made  $(\chi(H) - 1)$ -partite by removing at most  $\gamma pn^2$  edges.

Fix a  $G'$  as above and apply Proposition 4.1 with  $d_{4.1} = \gamma'$  to obtain an equipartition  $V_1, \dots, V_t$  of the vertex set of  $G'$  into  $t_0 \leq t \leq T$  pieces and a subgraph  $G''$  of  $G'$  all of whose edges lie in  $(\epsilon, p)$ -regular pairs with edge density at least  $\gamma' p$  and which satisfies

$$e(G') - e(G'') \leq \left(\frac{D}{t_0} + 2D\epsilon + \gamma'\right) \frac{pn^2}{2} \leq 3\gamma' pn^2.$$

Let  $R$  be the graph on vertex set  $[t]$  whose edges are those pairs  $ij$  such that the bipartite graph  $G''(V_i, V_j)$  is non-empty. Recall that each such bipartite graph is  $(\epsilon, p)$ -regular and has at least  $\gamma' p |V_i||V_j|$  edges. The following claim is proved in the same way as Claim 4.3.

**Claim 4.5.** *The cluster graph  $R$  contains a subgraph  $R'$  with  $\delta(R') \geq \left(1 - \frac{3}{3\chi(H)-4} + \frac{\gamma}{2}\right)t$  and  $t' \geq (1 - \beta)t$  vertices.*

Suppose now that  $R'$  contained a copy of  $H$ . Then there would be pairwise disjoint sets  $V_1, \dots, V_k$  and a sequence  $\mathbf{m} = (m_{ij})_{ij \in E(H)}$  with  $m_{ij} \geq \gamma' p \left(\frac{n}{t}\right)^2$  such that  $G'[V_1 \cup \dots \cup V_k] \in \mathcal{G}(H, \frac{n}{t}, \mathbf{m}, p, \epsilon)$ . Consequently, by Proposition 4.2 and the choice of  $\epsilon, \eta$ , and  $C$ , the number of copies of  $H$  in  $G''$  would be positive, contradicting the assumption that  $G'$  was  $H$ -free.

Therefore,  $R'$  is  $H$ -free and by the choice of  $t_0$ , the graph  $R'$  may be turned into a  $(\chi(H) - 1)$ -partite graph  $R_0$  by removing at most  $\frac{\gamma}{2} t'^2$  edges. The corresponding sparse graph  $G_0$  whose edges consist of all those edges contained within an edge of  $R_0$  is also  $(\chi(H) - 1)$ -partite. It is obtained from  $G'$  by first deleting at most  $3\gamma' p n^2$  edges to form  $G''$ , then at most  $\beta t^2 Dp \left(\frac{n}{t}\right)^2$  edges in forming  $R'$  from  $R$  and, finally, at most  $\frac{\gamma}{2} t'^2 Dp \left(\frac{n}{t}\right)^2$  edges in forming  $R_0$  from  $R'$ . Overall, this is at most

$$3\gamma' p n^2 + \beta t^2 Dp \left(\frac{n}{t}\right)^2 + \frac{\gamma}{2} t'^2 Dp \left(\frac{n}{t}\right)^2 \leq \frac{\gamma}{16} p n^2 + \frac{5\gamma}{16} p n^2 + \frac{9\gamma}{16} p n^2 < \gamma p n^2.$$

The result follows.

## 5 Concluding remarks

### 5.1 Hypergraphs

The methods employed in this paper should also extend to work for hypergraphs. That is, one should be able to show that a.a.s any regular partition of a subgraph of the random hypergraph has a corresponding counting lemma. The main obstacle here is to prove a sparse counterpart to the hypergraph regularity lemma [37, 38, 65, 74, 84]. This should be a comparatively straightforward hybrid of the hypergraph regularity lemma and the sparse regularity lemma. Once this theorem is in place, either method used in this paper should extend to show that it is effective in the random setting.

For linear hypergraphs, that is, hypergraphs for which every pair of edges intersect in at most one vertex, the results of this paper generalize more easily. In this case, the relevant regularity and counting lemmas [51] (see also [18]) follow in a similar fashion to the usual regularity and counting lemmas. This is because it is enough to have control over edge density on large vertex sets, whereas, for general hypergraphs, we need more elaborate conditions.

An alternative approach was already used in [17] to prove an extension of the hypergraph removal lemma to sparse random hypergraphs. Roughly speaking, rather than applying a sparse regularity lemma, one maps the sparse hypergraph  $G$  to its dense model  $K$  and then applies the usual hypergraph regularity lemma to this hypergraph  $K$ . This produces a regular partition which is also regular for the original hypergraph.

The difference is a matter of quantifiers. If we have a sparse hypergraph regularity lemma then the correct analogue of the KLR conjecture would be that a.a.s. *any* regular partition of a subgraph of the random hypergraph has a corresponding counting lemma. This alternative method allows one to say that a.a.s. for any subgraph of the random hypergraph there exists *some* regular partition for which there is a corresponding counting lemma. Despite this difference, we believe that this method is likely to be sufficient for most applications in the random setting.

### 5.2 Removing the need for strict balance

In Theorem 1.6 (ii) we assumed that the graph  $H$  was strictly balanced, which was sufficient for the applications in this paper. However, the methods of [17] can be used to obtain a result without

this condition if we allow an extra logarithmic factor. That is, if  $p \geq C(\log N)^c N^{-1/m_2(H)}$ , for  $c > 0$  an absolute constant, then a.a.s. we have the conclusion

$$G(H) = (1 \pm \delta) \left(\frac{m}{n^2}\right)^{e(H)} n^{v(H)}$$

that we had in Theorem 1.6 (ii). We are not formally claiming this as a result, since a certain amount of modification is needed to the proofs in [17], and though this modification appears to be straightforward, we have not written it out in detail.

However, let us briefly indicate why we are confident that this result is true. It turns out that many of the difficulties involved in proving transference disappear if one allows some extra logarithmic factors in the thresholds. This is because we no longer have to worry about certain large deviations from the expectation. To give an example, recall that the threshold for Turán's theorem for triangles occurs at around  $p = n^{-1/2}$ . This is the point at which the number of triangles is about the same as the number of edges. So we expect that most edges will be contained in at most a constant number of triangles. However, it will happen that there are some edges which are in many more triangles than expected and this deviation is enough to spoil some of the required estimates.

To deal with these unwanted deviations, one has to show that they do not occur too often and this adds many extra technicalities to the proof. In [17], the main tool for circumventing these problems was the introduction of so-called capped convolutions. If we allow some extra slack, we no longer have to work with these capped convolutions and this simplifies the proof considerably. In particular, it appears to be straightforward to prove the following result, valid for all graphs  $H$ .

**Claim 5.1.** *There is a positive constant  $c$  such that, for any graph  $H$  on  $k$  vertices and any  $\epsilon > 0$ , there exist positive constants  $C$  and  $\lambda$  such that if  $C(\log n)^c N^{-1/m_2(H)} \leq p \leq \lambda$  then the following holds a.a.s. in the random graph  $G_{N,p}$ . For all disjoint vertex subsets  $V_1, \dots, V_k$  and every subgraph  $G$  of  $G_{N,p}$  in  $H^*(V_1, \dots, V_k)$ , there exists a subgraph  $K$  of  $K_N$  in  $H^*(V_1, \dots, V_k)$  such that*

$$|p^{-e(H)} \mu_H^*(G) - \mu_H^*(K)| \leq \epsilon$$

and, for all pairs of disjoint vertex subsets  $U_1, U_2$  of  $V(K_N)$ ,

$$|p^{-1} \sum_{x,y} G(x,y) - \sum_{x,y} K(x,y)| \leq \epsilon N^2,$$

where the sums are taken over all  $x \in U_1$  and  $y \in U_2$ .

The version of Theorem 1.6 without the strict balance condition is a straightforward consequence of this claim, proved in exactly the same fashion as Part (ii) of Theorem 1.6. We omit the details.

**Acknowledgements.** The three junior authors are indebted to Tibor Szabó and his group at the Freie Universität, Berlin for hosting us during part of the period when we were working on this paper. We would also like to thank Jacob Fox for pointing out how to bridge the gap in the sparse removal lemma for balanced graphs.

## References

- [1] P. Allen. Dense  $H$ -free graphs are almost  $(\chi(H) - 1)$ -partite. *Electron. J. Combin.*, 17(1):Research Paper 21, 11pp. (electronic), 2010.

- [2] P. Allen, J. Böttcher, S. Griffiths, Y. Kohayakawa, and R. Morris. The chromatic thresholds of graphs. *Adv. Math*, 235:261–295, 2013.
- [3] N. Alon, A. Coja-Oghlan, H. Hàn, M. Kang, V. Rödl, and M. Schacht. Quasi-randomness and algorithmic regularity for graphs with general degree distributions. *SIAM J. Comput.*, 39(6):2336–2362, 2010.
- [4] N. Alon, R. A. Duke, H. Lefmann, V. Rödl, and R. Yuster. The algorithmic aspects of the regularity lemma. *J. Algorithms*, 16(1):80–109, 1994.
- [5] N. Alon and J. H. Spencer. *The probabilistic method*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons Inc., Hoboken, NJ, third edition, 2008. With an appendix on the life and work of Paul Erdős.
- [6] N. Alon and B. Sudakov.  $H$ -free graphs of large minimum degree. *Electron. J. Combin.*, 13(1):Research Paper 19, 9 pp. (electronic), 2006.
- [7] N. Alon and R. Yuster.  $H$ -factors in dense graphs. *J. Combin. Theory Ser. B*, 66(2):269–282, 1996.
- [8] B. Andrásfai, P. Erdős, and V. T. Sós. On the connection between chromatic number, maximal clique and minimal degree of a graph. *Discrete Math.*, 8:205–218, 1974.
- [9] J. Balogh, C. Lee, and W. Samotij. Corrádi and Hajnal’s theorem for sparse random graphs. *Combin. Probab. Comput.*, 21:23–55, 2012.
- [10] J. Balogh, R. Morris, and W. Samotij. Independent sets in hypergraphs. Submitted. arXiv:1204.6530v1 [math.CO].
- [11] M. Behrisch. Random graphs without a short cycle. Master’s thesis, Humboldt-Universität zu Berlin, 2002.
- [12] B. Bollobás. Relations between sets of complete subgraphs. In *Proceedings of the Fifth British Combinatorial Conference (Univ. Aberdeen, Aberdeen, 1975)*, pages 79–84. Congressus Numerantium, No. XV, Utilitas Math., Winnipeg, Man., 1976.
- [13] 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*.
- [14] D. Conlon and J. Fox. Graph removal lemmas. In *Surveys in combinatorics 2013*, London Math. Soc. Lecture Note Ser. Cambridge Univ. Press, Cambridge, 2013.
- [15] D. Conlon, J. Fox, and B. Sudakov. An approximate version of Sidorenko’s conjecture. *Geom. Funct. Anal.*, 20(6):1354–1366, 2010.
- [16] D. Conlon, J. Fox, and Y. Zhao. Extremal results in sparse pseudorandom graphs. Submitted. arXiv:1204.6645v1 [math.CO].
- [17] D. Conlon and W. T. Gowers. Combinatorial theorems in sparse random sets. Submitted. arXiv:1011.4310v1 [math.CO].
- [18] D. Conlon, H. Hàn, Y. Person, and M. Schacht. Weak quasi-randomness for uniform hypergraphs. *Random Structures Algorithms*, 40(1):1–38, 2012.

- [19] K. Corrádi and A. Hajnal. On the maximal number of independent circuits in a graph. *Acta Math. Acad. Sci. Hung.*, 14:423–439, 1963.
- [20] G. A. Dirac. Some theorems on abstract graphs. *Proc. London Math. Soc.*, 2:69–81, 1952.
- [21] P. Erdős and A. H. Stone. On the structure of linear graphs. *Bull. Amer. Math. Soc.*, 52:1087–1091, 1946.
- [22] P. Erdős. On a theorem of Rademacher-Turán. *Illinois J. Math.*, 6:122–127, 1962.
- [23] P. Erdős. On the number of complete subgraphs and circuits contained in graphs. *Časopis Pěst. Mat.*, 94:290–296, 1969.
- [24] P. Erdős, P. Frankl, and V. Rödl. The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent. *Graphs Combin.*, 2(2):113–121, 1986.
- [25] P. Erdős and M. Simonovits. Supersaturated graphs and hypergraphs. *Combinatorica*, 3(2):181–192, 1983.
- [26] D. C. Fisher. Lower bounds on the number of triangles in a graph. *J. Graph Theory*, 13(4):505–512, 1989.
- [27] J. Fox. A new proof of the graph removal lemma. *Ann. of Math. (2)*, 174(1):561–579, 2011.
- [28] E. Friedgut, V. Rödl, and M. Schacht. Ramsey properties of random discrete structures. *Random Structures Algorithms*, 37(4):407–436, 2010.
- [29] Z. Füredi. Extremal hypergraphs and combinatorial geometry. In *Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Zürich, 1994)*, pages 1343–1352, Basel, 1995. Birkhäuser.
- [30] S. Gerke, Y. Kohayakawa, V. Rödl, and A. Steger. Small subsets inherit sparse  $\epsilon$ -regularity. *J. Combin. Theory Ser. B*, 97(1):34–56, 2007.
- [31] S. Gerke, M. Marcinişzyn, and A. Steger. A probabilistic counting lemma for complete graphs. *Random Structures Algorithms*, 31(4):517–534, 2007.
- [32] S. Gerke, H. J. Prömel, T. Schickinger, A. Steger, and A. Taraz.  $K_4$ -free subgraphs of random graphs revisited. *Combinatorica*, 27(3):329–365, 2007.
- [33] S. Gerke, T. Schickinger, and A. Steger.  $K_5$ -free subgraphs of random graphs. *Random Structures Algorithms*, 24(2):194–232, 2004.
- [34] S. Gerke and A. Steger. The sparse regularity lemma and its applications. In *Surveys in combinatorics 2005*, volume 327 of *London Math. Soc. Lecture Note Ser.*, pages 227–258. Cambridge Univ. Press, Cambridge, 2005.
- [35] M. Goldwurm and M. Santini. Clique polynomials have a unique root of smallest modulus. *Inform. Process. Lett.*, 75(3):127–132, 2000.
- [36] A. W. Goodman. On sets of acquaintances and strangers at any party. *Amer. Math. Monthly*, 66:778–783, 1959.

- [37] W. T. Gowers. Quasirandomness, counting and regularity for 3-uniform hypergraphs. *Combin. Probab. Comput.*, 15(1-2):143–184, 2006.
- [38] W. T. Gowers. Hypergraph regularity and the multidimensional Szemerédi theorem. *Ann. of Math. (2)*, 166(3):897–946, 2007.
- [39] W. T. Gowers. Decompositions, approximate structure, transference, and the Hahn-Banach theorem. *Bull. Lond. Math. Soc.*, 42(4):573–606, 2010.
- [40] B. Green and T. Tao. The primes contain arbitrarily long arithmetic progressions. *Ann. of Math. (2)*, 167(2):481–547, 2008.
- [41] R. Häggkvist. Odd cycles of specified length in nonbipartite graphs. In *Graph theory (Cambridge, 1981)*, volume 62 of *North-Holland Math. Stud.*, pages 89–99. North-Holland, Amsterdam, 1982.
- [42] A. Hajnal and E. Szemerédi. Proof of a conjecture of P. Erdős. In *Combinatorial theory and its applications, II (Proc. Colloq., Balatonfüred, 1969)*, pages 601–623. North-Holland, Amsterdam, 1970.
- [43] H. Huang, C. Lee, and B. Sudakov. Bandwidth theorem for random graphs. *J. Combin. Theory Ser. B*, 102:14–37, 2012.
- [44] S. Janson, K. Oleszkiewicz, and A. Ruciński. Upper tails for subgraph counts in random graphs. *Israel J. Math.*, 142:61–92, 2004.
- [45] G. P. Jin. Triangle-free four-chromatic graphs. *Discrete Math.*, 145(1-3):151–170, 1995.
- [46] H. A. Kierstead and A. V. Kostochka. A short proof of the Hajnal-Szemerédi theorem on equitable colouring. *Combin. Probab. Comput.*, 17(2):265–270, 2008.
- [47] Y. Kohayakawa. Szemerédi’s regularity lemma for sparse graphs. In *Foundations of computational mathematics (Rio de Janeiro, 1997)*, pages 216–230. Springer, Berlin, 1997.
- [48] Y. Kohayakawa and B. Kreuter. Threshold functions for asymmetric Ramsey properties involving cycles. *Random Structures Algorithms*, 11(3):245–276, 1997.
- [49] Y. Kohayakawa, T. Łuczak, and V. Rödl. Arithmetic progressions of length three in subsets of a random set. *Acta Arith.*, 75(2):133–163, 1996.
- [50] Y. Kohayakawa, T. Łuczak, and V. Rödl. On  $K^4$ -free subgraphs of random graphs. *Combinatorica*, 17(2):173–213, 1997.
- [51] Y. Kohayakawa, B. Nagle, V. Rödl, and M. Schacht. Weak hypergraph regularity and linear hypergraphs. *J. Combin. Theory Ser. B*, 100(2):151–160, 2010.
- [52] Y. Kohayakawa and V. Rödl. Regular pairs in sparse random graphs. I. *Random Structures Algorithms*, 22(4):359–434, 2003.
- [53] Y. Kohayakawa and V. Rödl. Szemerédi’s regularity lemma and quasi-randomness. In *Recent advances in algorithms and combinatorics*, volume 11 of *CMS Books Math./Ouvrages Math. SMC*, pages 289–351. Springer, New York, 2003.

- [54] Y. Kohayakawa, V. Rödl, and M. Schacht. The Turán theorem for random graphs. *Combin. Probab. Comput.*, 13(1):61–91, 2004.
- [55] J. Komlós. Tiling Turán theorems. *Combinatorica*, 20(2):203–218, 2000.
- [56] J. Komlós, G. N. Sárközy, and E. Szemerédi. Proof of the Alon-Yuster conjecture. *Discrete Math.*, 235(1-3):255–269, 2001.
- [57] J. Komlós, A. Shokoufandeh, M. Simonovits, and E. Szemerédi. The regularity lemma and its applications in graph theory. In *Theoretical aspects of computer science (Tehran, 2000)*, volume 2292 of *Lecture Notes in Comput. Sci.*, pages 84–112. Springer, Berlin, 2002.
- [58] D. Kühn and D. Osthus. Embedding large subgraphs into dense graphs. In *Surveys in combinatorics 2009*, volume 365 of *London Math. Soc. Lecture Note Ser.*, pages 137–167. Cambridge Univ. Press, Cambridge, 2009.
- [59] D. Kühn and D. Osthus. The minimum degree threshold for perfect graph packings. *Combinatorica*, 29(1):65–107, 2009.
- [60] C. Lee and B. Sudakov. Dirac’s theorem for random graphs. *Random Structures Algorithms*, 41:293–305, 2012.
- [61] J. Li and B. Szegedy. On the logarithmic calculus and Sidorenko’s conjecture. Submitted. arXiv:1107.1153 [math.CO].
- [62] L. Lovász and M. Simonovits. On the number of complete subgraphs of a graph. II. In *Studies in pure mathematics*, pages 459–495. Birkhäuser, Basel, 1983.
- [63] T. Łuczak. On triangle-free random graphs. *Random Structures Algorithms*, 16(3):260–276, 2000.
- [64] T. Łuczak. Randomness and regularity. In *International Congress of Mathematicians. Vol. III*, pages 899–909. Eur. Math. Soc., Zürich, 2006.
- [65] B. Nagle, V. Rödl, and M. Schacht. The counting lemma for regular  $k$ -uniform hypergraphs. *Random Structures Algorithms*, 28(2):113–179, 2006.
- [66] V. Nikiforov. The number of cliques in graphs of given order and size. *Trans. Amer. Math. Soc.*, 363(3):1599–1618, 2011.
- [67] F. P. Ramsey. On a problem of formal logic,. *Proc. London Math. Soc. Ser. 2*, 30:264–286, 1930.
- [68] A. A. Razborov. Flag algebras. *J. Symbolic Logic*, 72(4):1239–1282, 2007.
- [69] A. A. Razborov. On the minimal density of triangles in graphs. *Combin. Probab. Comput.*, 17(4):603–618, 2008.
- [70] C. Reiher. The clique density theorem. Submitted. arXiv:1212.2454v1 [math.CO].
- [71] O. Reingold, L. Trevisan, M. Tulsiani, and S. P. Vadhan. Dense subsets of pseudorandom sets. In *FOCS*, pages 76–85. IEEE Computer Society, 2008.
- [72] V. Rödl and A. Ruciński. Threshold functions for Ramsey properties. *J. Amer. Math. Soc.*, 8(4):917–942, 1995.



- [73] V. Rödl and M. Schacht. Regularity lemmas for graphs. In *Fete of combinatorics and computer science*, volume 20 of *Bolyai Soc. Math. Stud.*, pages 287–325. János Bolyai Math. Soc., Budapest, 2010.
- [74] V. Rödl and J. Skokan. Regularity lemma for  $k$ -uniform hypergraphs. *Random Structures Algorithms*, 25(1):1–42, 2004.
- [75] K. F. Roth. On certain sets of integers. *J. London Math. Soc.*, 28:104–109, 1953.
- [76] I. Z. Ruzsa and E. Szemerédi. Triple systems with no six points carrying three triangles. In *Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976), Vol. II*, volume 18 of *Colloq. Math. Soc. János Bolyai*, pages 939–945. North-Holland, Amsterdam, 1978.
- [77] W. Samotij. Stability results for discrete random structures. *Random Structures Algorithms*. To appear. arXiv:1111.6885v3 [math.CO].
- [78] D. Saxton and A. Thomason. Hypergraph containers. Submitted. arXiv:1204.6595v2 [math.CO].
- [79] M. Schacht. Extremal results for discrete random structures. Submitted.
- [80] A. Scott. Szemerédi’s regularity lemma for matrices and sparse graphs. *Combin. Probab. Comput.*, 20(3):455–466, 2011.
- [81] M. Simonovits. A method for solving extremal problems in graph theory, stability problems. In *Theory of Graphs (Proc. Colloq., Tihany, 1966)*, pages 279–319. Academic Press, New York, 1968.
- [82] E. Szemerédi. On sets of integers containing no  $k$  elements in arithmetic progression. *Acta Arith.*, 27:199–245, 1975.
- [83] E. Szemerédi. Regular partitions of graphs. In *Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976)*, volume 260 of *Colloq. Internat. CNRS*, pages 399–401. CNRS, Paris, 1978.
- [84] T. Tao. A variant of the hypergraph removal lemma. *J. Combin. Theory Ser. A*, 113(7):1257–1280, 2006.
- [85] K. Zarankiewicz. Sur les relations symétriques dans l’ensemble fini. *Colloquium Math.*, 1:10–14, 1947.