

# STABILITY OF LARGE CUTS IN RANDOM GRAPHS

ILAY HOSHEN, WOJCIECH SAMOTIJ, AND MAKSIM ZHUKOVSKII

**ABSTRACT.** We prove that the family of largest cuts in the binomial random graph exhibits the following stability property: If  $1/n \ll p = 1 - \Omega(1)$ , then, with high probability, there is a set of  $n - o(n)$  vertices that is partitioned in the same manner by all maximum cuts of  $G_{n,p}$ . Moreover, the analogous statement remains true when one replaces maximum cuts with nearly-maximum cuts.

We then demonstrate how one can use this statement as a tool for showing that certain properties of  $G_{n,p}$  that hold in a fixed balanced cut hold simultaneously in all maximum cuts. We provide two example applications of this tool. First, we prove that maximum cuts in  $G_{n,p}$  typically partition the neighbourhood of every vertex into nearly equal parts; this resolves a conjecture of DeMarco and Kahn for all but a narrow range of densities  $p$ . Second, for all edge-critical, nonbipartite, and strictly 2-balanced graphs  $H$ , we prove a lower bound on the threshold density  $p$  above which every largest  $H$ -free subgraph of  $G_{n,p}$  is  $(\chi(H) - 1)$ -partite. Our lower bound exactly matches the upper bound on this threshold recently obtained by the first two authors.

## 1. INTRODUCTION

Let  $r \geq 2$  be an integer. An  $r$ -cut in a graph  $G$  is a partition of its vertex set into  $r$  subsets. The *size* of a cut is the number of edges of  $G$  with endpoints in different parts of the cut. A *maximum  $r$ -cut* of  $G$  is an  $r$ -cut that has the largest size (among all  $r$ -cuts in  $G$ ). Since the problem of determining the size of a maximum  $r$ -cut in a graph is NP-hard for every  $r \geq 2$ , much effort has been devoted to studying maximum cuts in random graphs, which can be viewed as the average-case variant of the general problem. In spite of this, we still know relatively little even about maximum 2-cuts in the binomial random graph  $G_{n,p}$ , except in the regime  $p = O(1/n)$ , see [5, 8, 10, 11]. In particular, the state-of-the-art result [3] falls short of determining the precise asymptotics of the second-order term<sup>1</sup> in the typical size of a maximum  $r$ -cut in  $G_{n,p}$  in the regime  $p \gg 1/n$ , determining it only up to a multiplicative constant.

The main obstacle in studying properties of maximum cuts in random graphs is that, even if a given property holds with high probability<sup>2</sup> for a fixed cut, commonly this probability is not sufficiently close to one to warrant a union bound over all (exponentially many in the number of vertices) choices of a cut. In an attempt to remedy this, Brightwell, Panagiotou, and Steger [2] proved that, if  $1/n \ll p \leq 1/2$ , then whp any two

---

This research was supported by: the Israel Science Foundation grant 2110/22; the grant 2019679 from the United States–Israel Binational Science Foundation (BSF) and the United States National Science Foundation (NSF); and the ERC Consolidator Grant 101044123 (RandomHypGra).

<sup>1</sup>It is easy to see that the first-order term is  $(1 - 1/r)n^2p/2$ .

<sup>2</sup>With probability approaching 1 as  $n \rightarrow \infty$ ; in what follows, we will write ‘whp’ for brevity.

maximum 2-cuts in  $G_{n,p}$  differ in a small number of vertices. Although this implies that the number of maximum 2-cuts in  $G_{n,p}$  is typically subexponential in  $n$ , this is still not enough to treat a maximum cut as a fixed cut. Brightwell, Panagiotou, and Steger also suggested that whp  $G_{n,p}$  has the following stronger property: There are disjoint vertex sets  $S_1$  and  $S_2$ , each containing nearly  $n/2$  vertices, that are contained in opposite parts of every maximum 2-cut in  $G_{n,p}$ . Intuitively, such property could prove useful in approximating a maximum cut by a fixed cut. This is because additionally requiring that  $S_1$  and  $S_2$  are inclusion-maximal with the above property ensures that the pair  $\{S_1, S_2\}$  is unique.

In this paper, we not only confirm the prediction of Brightwell, Panagiotou, and Steger but also show how one may exploit this ‘clustering’ property of maximum cuts while studying properties of  $G_{n,p}$ . Let  $b_r(G)$  denote the maximum size of an  $r$ -cut of  $G$ . The *deficit* of an  $r$ -cut in  $G$  is the difference between  $b_r(G)$  and the size of the cut. Given integers  $d \geq 0$  and  $r \geq 2$  and a real  $\alpha > 0$ , we say that a graph  $G$  *admits a  $(d, r, \alpha)$ -core* if there exist  $r$  pairwise-disjoint sets of vertices  $S_1, \dots, S_r$  of size exceeding  $(1/r - \alpha) \cdot |V(G)|$  each that are contained in different parts of every  $r$ -cut of  $G$  with deficit at most  $d$ . (We will give a precise definition of a core in Section 2.) The following statement is an abbreviated, qualitative version of our main result. We postpone the full, unabbreviated statement to Section 2.

**Theorem 1.1.** *The following hold for every integer  $r \geq 2$ , real  $\alpha > 0$ , and all  $p = p(n)$  satisfying  $1/n \ll p = 1 - \Omega(1)$ :*

- (i) *If  $d \ll \sqrt{np}$ , then whp  $G_{n,p}$  admits a  $(d, r, \alpha)$ -core.*
- (ii) *There exists a  $C = C(\alpha) > 0$  such that, for every  $d \geq C\sqrt{np}$ , whp no set of  $\alpha n$  vertices is entirely contained in a single part of every  $r$ -cut with deficit at most  $d$ .*

Theorem 1.1 implies, in particular, that whp each maximum  $r$ -cut in  $G_{n,p}$  ‘respects’ a single partition of all but some  $o(n)$  vertices into  $r$  parts. The aforementioned unabbreviated version of the theorem replaces this  $o(n)$  with  $\omega\sqrt{n/p}$ , where  $\omega$  is an arbitrary function that tends to infinity with  $n$ . While we do believe that this is optimal, in the sense that whp the largest set of vertices that is partitioned the same way by all largest  $r$ -cuts has size  $n - \Omega(\sqrt{n/p})$ , currently we cannot even show that the probability that  $G_{n,p}$  has a unique largest 2-cut is bounded away from one. In order to support our belief, we will at least show (in Section 2.6) that the *expected* number of vertices of  $G_{n,p}$  that lie outside of its core is  $\Omega(\sqrt{n/p})$ .

An important aspect of Theorem 1.1 is that the property of admitting a core is ‘stable’ under small perturbations of the edge set of a graph. More precisely, if a graph  $G$  admits a  $(d, r, \alpha)$ -core, then any graph  $G'$  that is obtained from  $G$  by adding/removing some  $t \leq d$  edges to/from  $G$  still admits a  $(d - t, r, \alpha)$ -core (moreover, the core of  $G'$  contains the core of  $G$  in a sense that will be made precise by Corollary 2.7 below). We will be able to exploit this fact when analysing certain properties of  $G_{n,p}$  that concern the distribution of its edges relative to a maximum cut. Roughly speaking, the stability

property of cores will allow one to change a small proportion of adjacencies in  $G_{n,p}$  while keeping the set of maximum cuts essentially unchanged (and thus determined up to the locations of  $o(n)$  vertices).

In the second part of this work, we present two applications of Theorem 1.1. First, we resolve (almost fully) a conjecture of DeMarco and Kahn [6] stating that maximum cuts partition neighbourhoods of all vertices almost equally, see Section 1.1 below. Second, for every nonbipartite edge-critical and strictly 2-balanced graph  $H$ , we prove an optimal lower bound on the threshold probability for the property that every largest  $H$ -free subgraph of  $G_{n,p}$  is  $(\chi(H) - 1)$ -partite; a matching upper bound on this threshold probability was recently proved by the first two authors [14]. This resolves another problem proposed by DeMarco and Kahn [6] and refutes their guess regarding the location of the said threshold. We present the relevant background, as well as our results, in Section 1.2.

**1.1. Neighbourhoods in maximum cuts.** It is easy to show that, as soon as  $p \gg \log n/n$ , a fixed 2-cut of  $G_{n,p}$  into sets of size  $n/2 + o(n)$  (we call such a cut *balanced*) whp divides the neighbourhood of every vertex into parts of size  $(1/2 + o(1))np$ . Obviously, this is not true for all balanced cuts, but is it true for all maximum cuts? It was conjectured by DeMarco and Kahn [6, Conjecture 13.2] that the answer is yes, for all  $p \gg \log n/n$ . In Section 3, we confirm the conjecture for all  $p \gg (\log n)^2/n$ . In fact, with only little extra work, we generalise this result to  $r$ -cuts and common neighbourhoods of sets of vertices of any constant size.

**Theorem 1.2.** *Let  $k \geq 1$ ,  $r \geq 2$  be integers and let  $\varepsilon > 0$ . There exists  $C > 0$  such that, for all  $p$  satisfying  $p \geq C(\log n/n)^{1/k}$  and  $p \gg (\log n)^2/n$ , whp every maximum  $r$ -cut of  $G_{n,p}$  partitions the common neighbourhood of every set of  $k$  vertices into parts of size  $(1/r \pm \varepsilon)np^k$  each.*

It is not hard to see that the lower-bound assumption on  $p$  is optimal (up to the value of  $C$ ) for all  $k \geq 2$  and  $r \geq 2$ . While we do believe that the assumption  $p \gg (\log n)^2/n$  (which comes in force only when  $k = 1$ ) may be weakened to  $p \gg \log n/n$ , proving this would likely involve significantly new ideas. The reason why we need this additional assumption is that, in order to overcome the union bound over the choice of the set of vertices whose neighbourhood we are analysing, we need to resample  $\Theta(\log n)$  edges of  $G_{n,p}$ ; however, this seems to require good control over cuts with deficit  $\Theta(\log n)$ , which we do not have unless  $p \gg (\log n)^2/n$ , see the second part of Theorem 1.1.

**1.2. Sharp thresholds for Simonovits's theorem in  $G_{n,p}$ .** The well-known theorem of Turán [20] states that, for every  $r \geq 2$  and all  $n$ , the largest  $K_{r+1}$ -free subgraphs of the complete graph  $K_n$  are its largest  $r$ -partite subgraphs. Simonovits [19] proved that a similar result holds also for edge-critical graphs; we say that a graph  $H$  is *edge-critical* if  $\chi(H \setminus e) = \chi(H) - 1$  for some edge  $e \in H$ , where  $\chi(H)$  is the chromatic number of  $H$ .

**Theorem 1.3** ([19]). *If a graph  $H$  is edge-critical and  $n$  is a sufficiently large integer, then every largest  $H$ -free subgraph of  $K_n$  is  $(\chi(H) - 1)$ -partite.*

Let us call a graph  $H$ -Simonovits, if each of its largest  $H$ -free subgraphs is  $(\chi(H) - 1)$ -partite. Note that the assumption that  $H$  is edge-critical in the above theorem is crucial. Indeed, adding one edge to a  $(\chi(H) - 1)$ -partite graph cannot introduce a copy of  $H$  unless  $H$  is edge-critical. Consequently, if  $H$  is not edge-critical, then no graph with chromatic number at least  $\chi(H)$  can be  $H$ -Simonovits (in particular, no  $K_n$  with  $n \geq \chi(H)$  is  $H$ -Simonovits).

To the best of our knowledge, the first to study the question of when  $G_{n,p}$  is whp  $H$ -Simonovits were Babai, Simonovits, and Spencer [1]. They proved that, for every  $\ell \geq 1$ ,  $G_{n,p}$  is whp  $C_{2\ell+1}$ -Simonovits as long as  $p \geq 1/2 - \varepsilon_\ell$  for some (small) positive constant  $\varepsilon_\ell$  that depends only on  $\ell$ . Answering a challenge raised by the authors of [1], Brightwell, Panagiotou, and Steger [2] proved that  $G_{n,p}$  is whp  $K_{r+1}$ -Simonovits, for every  $r \geq 2$ , already when  $p \geq n^{-c_r}$  for some (small) constant  $c_r > 0$ .

It is not hard to see that as soon as the expected number of copies of some subgraph  $F \subseteq H$  in  $G_{n,p}$  becomes significantly smaller than the expected number of edges,  $G_{n,p}$  cannot be  $H$ -Simonovits.<sup>3</sup> This implies that  $G_{n,p}$  cannot be  $H$ -Simonovits whp, unless  $p = \Omega(n^{-1/m_2(H)})$ , where (we write  $e_F$  and  $v_F$  for the numbers of edges and vertices of a graph  $F$ , respectively)

$$m_2(H) := \max \left\{ \frac{e_F - 1}{v_F - 2} : F \subseteq H, e_F \geq 2 \right\}$$

is the 2-density of  $H$ . Kohayakawa, Łuczak, and Rödl [16] conjectured that, for every nonbipartite graph  $H$  (not necessarily edge-critical), when  $p \gg n^{-1/m_2(H)}$ , then whp every largest  $H$ -free subgraph of  $G_{n,p}$  is close to being  $(\chi(H) - 1)$ -partite (i.e., it can be made  $(\chi(H) - 1)$ -partite by removing some  $o(n^2p)$  edges). This was proved in the breakthrough work of Conlon and Gowers [4], under the technical assumption that  $H$  is also *strictly 2-balanced* (i.e., the maximum in the definition of the 2-density is achieved uniquely at  $F = H$ ), which was later removed by the second author [17], using an adaptation of the argument of Schacht [18], who proved the slightly weaker assertion that whp every  $H$ -free subgraph of  $G_{n,p}$  has at most  $(1 - 1/(\chi(H) - 1) + o(1)) \binom{n}{2} p$  edges.

In another major development, DeMarco and Kahn [6, 7] showed that adding an extra polylogarithmic factor in the lower bound on  $p$  suffices to show that, when  $H$  is a clique, then whp  $G_{n,p}$  is  $H$ -Simonovits. Moreover, their lower-bound assumption on  $p$  is best possible up to a constant factor. This result was recently generalised by the first two authors [14] to every nonbipartite, strictly 2-balanced, and edge-critical graph  $H$ .

In order to state the main result of [14], we need an additional definition. First, let us denote by  $K_r(m)$  the complete, balanced  $r$ -partite graph with parts of size  $m$  and let  $K_r(m)^+$  be the graph obtained from  $K_r(m)$  by adding a single edge contained in one of the parts. (Note that  $H$  is edge-critical if and only if  $H \subseteq K_{\chi(H)-1}(m)^+$  for all  $m \geq v_H$ .) Letting  $\text{Cop}(H, G)$  be the number of copies of  $H$  in  $G$ , set

$$\pi_H := \lim_{m \rightarrow \infty} \frac{\text{Cop}(H, K_{\chi(H)-1}(m)^+)}{m^{v_H-2}} > 0. \quad (1)$$

<sup>3</sup>Unless  $p = O(1/n)$ , since then  $\chi(G_{n,p}) = O(1)$  whp and a more careful analysis is required.

Finally, let  $\theta_H$  be the positive real satisfying

$$(\chi(H) - 1)^{2-v_H} \cdot \pi_H \cdot \theta_H^{e_H-1} = 2 - \frac{1}{m_2(H)}. \quad (2)$$

**Theorem 1.4** ([14]). *If  $H$  is a nonbipartite, edge-critical, strictly 2-balanced graph and*

$$p \geq (\theta_H + \varepsilon) \cdot n^{-1/m_2(H)} (\log n)^{1/(e_H-1)}$$

*for some positive constant  $\varepsilon$ , then whp  $G_{n,p}$  is  $H$ -Simonovits.*

In Section 4, we show that  $\theta_H n^{-1/m_2(H)} (\log n)^{1/(e_H-1)}$  is in fact a sharp threshold for the property of being  $H$ -Simonovits. This confirms the prediction made in [14] (and refutes the suggestion of DeMarco and Kahn [6] in the case of when  $H$  is a complete graph).

**Theorem 1.5.** *If  $H$  is a nonbipartite, edge-critical, strictly 2-balanced graph and*

$$1/n \ll p \leq (\theta_H - \varepsilon) \cdot n^{-1/m_2(H)} (\log n)^{1/(e_H-1)}$$

*for some positive constant  $\varepsilon$ , then whp  $G_{n,p}$  is not  $H$ -Simonovits.*

It remains an interesting open problem to extend Theorems 1.4 and 1.5 to edge-critical graphs  $H$  that are not strictly 2-balanced. We note that the closely-related problem of describing the typical structure of  $H$ -free graphs with given order and size, for an arbitrary edge-critical graph  $H$ , was studied in [9].

**1.3. Relation to previous works.** The notions of cores and rigidity (defined in Section 2), which play a central role in our arguments, were first introduced by DeMarco and Kahn [6]. In fact, one can infer from the proof of [6, Lemma 12.3] that the random graph  $G_{n,p}$  is typically close (up to adding/removing a small number of edges) to a graph that has a  $(0, r, o(1))$ -core. Finally, even though the result of Brightwell, Panagiotou, and Steger [2] does not imply the existence of a core, our proof of the fact that the random graph typically has a  $(0, r, o(1))$ -core (the case  $d = 0$  in Theorem 1.1) was inspired by their arguments.

**Organisation.** The remainder of this paper is organised as follows. Section 2 is devoted to the proof of Theorem 1.1. In fact, the main results of this section are Theorem 2.1 and Corollary 2.2, the latter of which is a stronger, non-asymptotic version of Theorem 1.1. In the following two sections, we prove Theorems 1.2 and 1.5, respectively.

## 2. RIGIDITY

Recall the definitions of an  $r$ -cut and the size and the deficit of a cut. In order to quantify the ‘clustering’ property of the family of cuts with a small deficit in a graph  $G$ , we will count pairs of vertices of  $G$  that are never separated by such a cut. To make this precise, given integers  $r \geq 2$  and  $d \geq 0$ , we say that two vertices are  $(d, r)$ -equivalent in  $G$  if they are in the same part of every  $r$ -cut in  $G$  with deficit at most  $d$ . Further, we say that  $G$  is  $(d, r, \varepsilon)$ -rigid, for some  $\varepsilon > 0$ , if there are at least  $\frac{1-\varepsilon}{r} \binom{n}{2}$  pairs of

$(d, r)$ -equivalent vertices. The following theorem, which is the main result of this work, provides a lower bound on the probability that the uniform random graph  $G(n, m)$  is  $(d, r, \varepsilon)$ -rigid for given  $d, r$ , and  $\varepsilon$  (that are allowed to depend on  $n$  and  $m$ ).

**Theorem 2.1.** *There exists an absolute constant  $C$  such that, for all  $\delta, \varepsilon \in (0, 1)$  and all nonnegative integers  $d, m, n, r$  satisfying  $r \geq 2$  and  $1 \leq m \leq (1 - \delta) \binom{n}{2}$ ,*

$$\mathbb{P}(G(n, m) \text{ is not } (d, r, \varepsilon)\text{-rigid}) \leq \frac{Cr}{\varepsilon} \cdot \left( \left( \frac{d+1}{\delta} + r \right) \cdot \sqrt{\frac{n}{m}} + \sqrt[4]{\frac{n}{m}} \right).$$

Let us now relate the property of being rigid to the property of admitting a core from the statement of Theorem 1.1. To this end, note first that  $(d, r)$ -equivalence is an equivalence relation on the vertex set of  $G$ ; we shall call its equivalence classes the  $(d, r)$ -components of  $G$ . In the case when the  $r$  largest  $(d, r)$ -components  $C_1, \dots, C_r$  of  $G$  have strictly more than  $n/(r+1)$  vertices each (and thus all remaining components have fewer than  $n/(r+1)$  vertices), we call the set  $\text{core}_d^r(G) := \{C_1, \dots, C_r\}$  the  $(d, r)$ -core of  $G$ . We will show that, under mild assumptions on the distribution of edges of  $G$  (that  $G(n, m)$  fails to have with probability much smaller than the upper bound in Theorem 2.1), the fact that  $G$  is  $(d, r, \varepsilon)$ -rigid implies that it has a  $(d, r)$ -core whose each component has at least  $n/r - r\varepsilon n$  vertices. In particular, letting  $\text{CORE}_d^r(\alpha)$  denote the set of all graphs on  $\llbracket n \rrbracket$  that have a  $(d, r)$ -core whose each component has at least  $n/r - \alpha n$  vertices, Theorem 2.1 will imply the following statement.

**Corollary 2.2.** *There exists an absolute constant  $C$  such that, for all  $\alpha, \delta \in (0, 1)$  and all nonnegative integers  $d, m, n, r$  satisfying  $r \geq 2$ ,  $\alpha < 1/(r^2 + r)$ , and  $1 \leq m \leq (1 - \delta) \binom{n}{2}$ ,*

$$\mathbb{P}(G(n, m) \notin \text{CORE}_d^r(\alpha)) \leq \frac{Cr^2}{\alpha} \cdot \left( \left( \frac{d+1}{\delta} + r \right) \cdot \sqrt{\frac{n}{m}} + \sqrt[4]{\frac{n}{m}} \right).$$

**Organisation.** The remainder of this section is organised as follows. First, in Section 2.1, we establish two crucial properties of the dynamics of the sets of  $(d, r)$ -equivalent pairs of vertices with respect to addition/deletion of edges. Next, in Section 2.2, we derive several (standard) estimates regarding the concentration of the number of edges in various induced subgraphs of  $G(n, m)$  and sizes of the parts in cuts with small deficit. In Section 2.3, we use the results established in the previous two subsections to prove Theorem 2.1 and in Section 2.4, we derive Corollary 2.2. Finally, in Section 2.5 we prove Theorem 1.1.

**Notation.** Given integers  $d \geq 0$  and  $r \geq 2$ , the set of  $(d, r)$ -equivalent pairs in a graph  $G$  will be henceforth denoted by  $\text{eq}_d^r(G)$ . For brevity, we will identify a graph with its set of edges; in particular, we will write  $|G|$  to denote the number of edges in a graph  $G$ . Following DeMarco and Kahn [6], we will use the following notational conventions: First, given a collection  $\Pi$  of pairwise-disjoint sets of vertices in a graph (e.g., a cut), we will denote the set of pairs of vertices contained in a single set of  $\Pi$  by  $\text{int}(\Pi)$  and the set of remaining pairs of vertices of  $\bigcup \Pi$  by  $\text{ext}(\Pi)$ ; this way

$$b_r(G) = \max\{|\text{ext}(\Pi) \cap G| : \Pi \text{ is an } r\text{-cut}\}.$$

Second, given an integer  $r \geq 2$ , we will write  $\text{crit}_r(G)$  to denote the set of edges of  $G$  that cross all maximum  $r$ -cuts (the  $r$ -critical edges of  $G$ ). Finally, given two families  $\mathcal{C}$  and  $\mathcal{C}'$  of pairwise-disjoint sets of vertices, we will write  $\mathcal{C} \preceq \mathcal{C}'$  if each element of  $\mathcal{C}$  is contained in some element of  $\mathcal{C}'$ . In other words,

$$\mathcal{C} \preceq \mathcal{C}' \iff \forall X \in \mathcal{C} \exists X' \in \mathcal{C}' X \subseteq X' \iff \text{int}(\mathcal{C}) \subseteq \text{int}(\mathcal{C}').$$

**2.1. Dynamics of equivalent pairs.** Suppose that  $r \geq 2$  and that  $G$  is an arbitrary graph on  $\llbracket n \rrbracket$ . The proof of Theorem 2.1 will exploit the following crucial property of the dynamics of the sequence  $(\text{eq}_d^r(G))_d$  with respect to addition of edges.

**Lemma 2.3.** *The following holds for all integers  $r \geq 2$  and  $d \geq 0$  and every  $e \in K_n$ : If  $e \notin G \cup \text{eq}_{d+1}^r(G)$ , then  $e \notin \text{eq}_d^r(G \cup e)$ .*

Before we prove the lemma, it will be useful to observe the following alternative definitions of the set of  $(0, r)$ -equivalent pairs.

**Fact 2.4.** *The following statements are equivalent for all  $e \in K_n \setminus G$ :*

- $e \notin \text{eq}_0^r(G)$ ;
- $e \in \text{ext}(\Pi)$  for some maximum  $r$ -cut  $\Pi$  of  $G$ ;
- $e \in \text{ext}(\Pi)$  for all maximum  $r$ -cuts  $\Pi$  of  $G \cup e$  (that is,  $e \in \text{crit}_r(G \cup e)$ )
- $b_r(G \cup e) = b_r(G) + 1$ .

*Proof of Lemma 2.3.* Suppose that  $e \notin G \cup \text{eq}_{d+1}^r(G)$ . Let  $\Pi$  be an  $r$ -cut in  $G$  such that  $e \in \text{ext}(\Pi)$  and whose deficit  $d_e$  in  $G$  is the smallest possible. Note that  $d_e \leq d + 1$  as we assumed that  $e$  crosses some cut with deficit at most  $d + 1$ . If  $d_e = 0$ , then  $\Pi$  is a maximum  $r$ -cut of  $G \cup e$ , so in particular  $e \notin \text{eq}_d^r(G \cup e)$ , due to Fact 2.4. Otherwise,  $b_r(G \cup e) = b_r(G)$  and, consequently, the deficit of  $\Pi$  in  $G \cup e$  equals

$$b_r(G \cup e) - |\text{ext}(\Pi) \cap (G \cup e)| = b_r(G) - |\text{ext}(\Pi) \cap G| - 1 = d_e - 1 \leq d.$$

In particular,  $\Pi$  is a cut witnessing that  $e \notin \text{eq}_d^r(G \cup e)$ .  $\square$

Both of our applications of Corollary 2.2 will crucially use the fact that cores are ‘stable’ under small edge perturbations. Our next lemma, and its corollaries, formalise this notion of stability.

**Lemma 2.5.** *For every integer  $d \geq 0$ , graph  $G \subseteq K_n$ , and edge  $e \in K_n$ , we have*

$$\text{eq}_{d+1}^r(G) \subseteq \text{eq}_d^r(G \triangle e).$$

*Proof.* We prove the equivalent statement  $K_n \setminus \text{eq}_d^r(G \triangle e) \subseteq K_n \setminus \text{eq}_{d+1}^r(G)$ . Observe first that, for every fixed  $r$ -cut  $\Pi$ , the function<sup>4</sup>

$$\mathcal{P}(K_n) \ni H \mapsto \text{def}_H(\Pi) = b_r(H) - |H \cap \text{ext}(\Pi)|$$

is the difference of two nondecreasing, 1-Lipshitz functions, and thus it is also 1-Lipshitz. In particular, for every graph  $G$  and edge  $e \in K_n$ , we have  $\text{def}_G(\Pi) \leq \text{def}_{G \triangle e}(\Pi) + 1$ .

<sup>4</sup>We write  $\mathcal{P}(K_n)$  to denote the family of all subgraphs of  $K_n$ , i.e., the powerset of  $K_n$ , which we identify here with its set of edges.

Now, suppose that  $f \in K_n \setminus \text{eq}_d(G \triangle e)$ , that is,  $f \in \text{ext}(\Pi)$  for some  $\Pi$  with  $\text{def}_{G \triangle e}(\Pi) \leq d$ . Since  $\text{def}_G(\Pi) \leq \text{def}_{G \triangle e}(\Pi) + 1 \leq d + 1$ , we have  $f \notin \text{eq}_{d+1}(G)$ .  $\square$

**Corollary 2.6.** *For every integer  $d \geq 0$  and graphs  $G, T \subseteq K_n$ , we have*

$$\text{eq}_{d+e_T}(G) \subseteq \text{eq}_d(G \triangle T).$$

**Corollary 2.7.** *The following holds for every integer  $d \geq 0$  and all graphs  $G, T \subseteq K_n$ . If  $G$  has a  $(d + e_T, r)$ -core, then  $G \triangle T$  has a  $(d, r)$ -core and  $\text{core}_{d+e_T}^r(G) \preceq \text{core}_d^r(G \triangle T)$ .*

**2.2. Distribution of edges in  $G(n, m)$ .** Our arguments require various estimates of the number of edges in certain subgraphs of  $G(n, m)$ . Luckily, all these estimates can be easily deduced from the following lower-tail estimate for the number of edges in  $G(n, m)$  induced by subsets of its vertices.

**Lemma 2.8.** *Suppose that  $G \sim G(n, m)$  for some  $m \in \{0, \dots, \binom{n}{2}\}$ . Then, letting  $p = m / \binom{n}{2}$ ,*

$$\mathbb{P} \left( \exists U \ |G[U]| < \binom{|U|}{2} p - 2|U|\sqrt{np} \right) \leq e^{-n}.$$

*Proof.* We may clearly assume that  $m \geq 1$ . Fix a nonempty  $U \subseteq [n]$ . Standard estimates on lower tail probabilities of the binomial distribution yield<sup>5</sup>

$$\mathbb{P} \left( |G[U]| < \binom{|U|}{2} p - 2|U|\sqrt{np} \right) \leq \exp \left( -\frac{4|U|^2 np}{p|U|^2} \right) = e^{-4n}.$$

The union bound over all  $U$  finishes the proof.  $\square$

In fact, we will only use the following immediate corollary of Lemma 2.8.

**Corollary 2.9.** *Suppose that  $G \sim G(n, m)$  for some  $m \in \{0, \dots, \binom{n}{2}\}$ . With probability at least  $1 - e^{-n}$ , for every partition  $\Pi$  of  $[n]$ ,*

$$|\text{int}(\Pi) \cap G| \geq |\text{int}(\Pi)| \cdot p - 2n\sqrt{np}.$$

One important consequence of Corollary 2.9 is that every cut of  $G(n, m)$  with small deficit must be balanced.

**Lemma 2.10.** *Suppose that  $G \sim G(n, m)$  for some  $m \in \{0, \dots, \binom{n}{2}\}$ . With probability at least  $1 - e^{-n}$ , each part of every  $r$ -cut of  $G$  with deficit at most  $d$  has at most  $n/r + 4 \max\{(d/m)^{1/2}, (n/m)^{1/4}\} \cdot n$  vertices.*

The proof of Lemma 2.10 uses the following straightforward estimate on the number of edges that cross a non-balanced cut.

**Fact 2.11.** *If  $\Pi = (A_1, \dots, A_r)$  is an  $r$ -cut in  $K_n$  with  $\max_i |A_i| = n/r + \varepsilon n$  for some  $\varepsilon \geq 0$ , then*

$$|\text{ext}(\Pi)| \leq \left( 1 - \frac{1}{r} - \frac{r\varepsilon^2}{r-1} \right) \cdot \binom{n}{2} + \frac{n}{2}.$$

<sup>5</sup>See [13, Section 6], which argues that the hypergeometric distribution is at least as concentrated as the binomial distribution with the same parameters.

*Proof.* We may clearly assume that  $|A_1| = \max_i |A_i| = n/r + \varepsilon n$ . By convexity of  $x \mapsto \binom{x}{2} := \frac{x(x-1)}{2}$ ,

$$|\text{ext}(\Pi)| = \binom{n}{2} - \sum_{i=1}^r \binom{|A_i|}{2} \leq \binom{n}{2} - \binom{n/r + \varepsilon n}{2} - (r-1) \cdot \binom{n/r - \varepsilon n/(r-1)}{2}.$$

Further, as for all nonnegative  $\lambda_1, \dots, \lambda_r$  with  $\lambda_1 + \dots + \lambda_r = 1$ , we have

$$\sum_{i=1}^r \left( \lambda_i^2 \binom{n}{2} - \binom{\lambda_i n}{2} \right) = \sum_{i=1}^r \frac{(1 - \lambda_i) \lambda_i n}{2} \leq \frac{n}{2},$$

we may conclude that

$$\begin{aligned} |\text{ext}(\Pi)| &\leq \left( 1 - \left( \frac{1}{r} + \varepsilon \right)^2 - (r-1) \left( \frac{1}{r} - \frac{\varepsilon}{r-1} \right)^2 \right) \cdot \binom{n}{2} + \frac{n}{2} \\ &= \left( 1 - \frac{1}{r} - \frac{r\varepsilon^2}{r-1} \right) \cdot \binom{n}{2} + \frac{n}{2}, \end{aligned}$$

as claimed.  $\square$

*Proof of Lemma 2.10.* Assume the assertion of Corollary 2.9, which holds with probability at least  $1 - e^{-n}$ . Let  $\Pi = (A_1, \dots, A_r)$  be an  $r$ -cut with deficit at most  $d$  and let  $\varepsilon$  be the number satisfying  $\max_i |A_i| = n/r + \varepsilon n$ . On the one hand, since every graph  $G$  with  $m$  edges satisfies  $b_r(G) \geq \frac{r-1}{r} \cdot m$  (to see this, consider a uniformly random  $r$ -cut), the size of  $\Pi$  in  $G$  is at least  $\frac{r-1}{r} \cdot m - d$ . On the other hand, by the assumed assertion of Corollary 2.9 and Fact 2.11, letting  $p := m/\binom{n}{2}$ ,

$$|\text{ext}(\Pi) \cap G| \leq |\text{ext}(\Pi)| \cdot p + 2n\sqrt{np} \leq \left( \frac{r-1}{r} - \frac{r\varepsilon^2}{r-1} \right) \cdot m + \frac{np}{2} + 2n\sqrt{np}.$$

This yields

$$\frac{r\varepsilon^2 m}{r-1} \leq d + \frac{np}{2} + 2n\sqrt{np} \leq d + \frac{5n\sqrt{np}}{2} \leq d + 4\sqrt{nm},$$

which means that

$$\max_i |A_i| - \frac{n}{r} = \varepsilon n \leq 2 \max \left\{ \sqrt{\frac{d}{m}}, \left( \frac{16n}{m} \right)^{1/4} \right\} \cdot n,$$

as desired.  $\square$

**2.3. Proof of Theorem 2.1.** It will be convenient to define, for a graph  $G$  on  $[n]$ ,

$$\text{eq}_{-1}^r(G) := \text{int}(\Pi_G),$$

where  $\Pi_G$  is some (canonically chosen) maximum  $r$ -cut in  $G$ . By convexity of  $x \mapsto \binom{x}{2}$ ,

$$|\text{eq}_{-1}^r(G)| \geq r \cdot \binom{n/r}{2} = \frac{n^2}{2r} - \frac{n}{2}. \quad (3)$$

The following lemma lies at the heart of the proof of Theorem 2.1. Throughout this section, we denote  $N := \binom{n}{2}$ .

**Lemma 2.12.** *Let  $G_1 \sim G(n, m)$  and  $G_2 \sim G(n, m+1)$  for some  $m \in \{0, \dots, N-1\}$ . For every  $r \geq 2$  and all  $d \geq -1$ ,*

$$\mathbb{E} |\text{eq}_{d+1}^r(G_1) \setminus G_1| \geq \frac{N-m}{m+1} \cdot \mathbb{E} |\text{eq}_d^r(G_2) \cap G_2|.$$

*Proof.* We may clearly couple  $G_1$  and  $G_2$  so that  $G_2 = G_1 \cup e$  and  $e$  is a uniformly random edge of each of  $K_n \setminus G_1$  and  $G_2$ . Define, for every  $d \geq -1$ ,

$$q_d := \mathbb{P}(e \notin \text{eq}_d^r(G_1))$$

and observe that

$$q_d \cdot (N-m) = \mathbb{E} |(K_n \setminus G_1) \setminus \text{eq}_d^r(G_1)| = N-m - \mathbb{E} |\text{eq}_d^r(G_1) \setminus G_1|. \quad (4)$$

On the other hand, Lemma 2.3 implies that, for every  $d \geq 0$ ,

$$q_{d+1} \cdot (m+1) \leq \mathbb{P}(e \notin \text{eq}_d^r(G_2)) \cdot (m+1) = m+1 - \mathbb{E} |\text{eq}_d^r(G_2) \cap G_2|. \quad (5)$$

This inequality remains valid also when  $d = -1$  since, by Fact 2.4,

$$q_0 \cdot (m+1) = \mathbb{E} |\text{crit}_r(G_2)| \leq \mathbb{E} [b_r(G_2)] = m+1 - \mathbb{E} |\text{eq}_{-1}^r(G_2) \cap G_2|.$$

The claimed inequality follows by combining (4), with  $d$  replaced by  $d+1$ , and (5).  $\square$

We now use Corollary 2.9 to derive the following cleaner statement.

**Corollary 2.13.** *Let  $G_1 \sim G(n, m)$  and  $G_2 \sim G(n, m+1)$  for some  $m \in \{0, \dots, N-1\}$ . For every  $r \geq 2$ , all  $d \geq -1$ , and all sufficiently large  $n$ ,*

$$\mathbb{E} |\text{eq}_{d+1}^r(G_1)| \geq \mathbb{E} |\text{eq}_d^r(G_2)| - \frac{2n^{9/2}}{\sqrt{m+1} \cdot (N-m)}.$$

*Proof.* Since, for all  $d, r$ , and  $G$ , the graph  $\text{eq}_d^r(G)$  is a union of cliques, Corollary 2.9 gives, letting  $p_1 := m/N$  and  $p_2 := (m+1)/N$ ,

$$\mathbb{E} |\text{eq}_d^r(G_2) \cap G_2| \geq \mathbb{E} |\text{eq}_d^r(G_2)| \cdot p_2 - 2n\sqrt{np_2} - n^2e^{-n} \geq \mathbb{E} |\text{eq}_d^r(G_2)| \cdot p_2 - 3n\sqrt{np_2},$$

where the last inequality holds for all sufficiently large  $n$ , and, similarly,

$$\begin{aligned} \mathbb{E} |\text{eq}_{d+1}^r(G_1) \setminus G_1| &= \mathbb{E} |\text{eq}_{d+1}^r(G_1)| - \mathbb{E} |\text{eq}_{d+1}^r(G_1) \cap G_1| \\ &\leq \mathbb{E} |\text{eq}_{d+1}^r(G_1)| \cdot (1-p_1) + 3n\sqrt{np_1}. \end{aligned}$$

Consequently, Lemma 2.12 yields

$$\begin{aligned} \mathbb{E} |\text{eq}_{d+1}^r(G_1)| &\geq \frac{1}{1-p_1} \cdot (\mathbb{E} |\text{eq}_{d+1}^r(G_1) \setminus G_1| - 3n\sqrt{np_1}) \\ &\geq \frac{1}{1-p_1} \cdot \frac{N-m}{m+1} \cdot \mathbb{E} |\text{eq}_d^r(G_2) \cap G_2| - \frac{3n\sqrt{np_1}}{1-p_1} \\ &\geq \frac{p_2}{1-p_1} \cdot \frac{N-m}{m+1} \cdot \left( \mathbb{E} |\text{eq}_d^r(G_2)| - \frac{3n\sqrt{np_2}}{p_2} \right) - \frac{3n\sqrt{np_1}}{1-p_1} \\ &= \mathbb{E} |\text{eq}_d^r(G_2)| - 3n^{3/2} \cdot \left( \frac{1}{\sqrt{p_2}} + \frac{\sqrt{p_1}}{1-p_1} \right). \end{aligned}$$

The claimed inequality follows as

$$\frac{1}{\sqrt{p_2}} + \frac{\sqrt{p_1}}{1-p_1} \leq \frac{1}{\sqrt{p_2}} + \frac{\sqrt{p_2}}{1-p_1} = \frac{1-p_1+p_2}{\sqrt{p_2} \cdot (1-p_1)} = \frac{\sqrt{N} \cdot (N+1)}{\sqrt{m+1} \cdot (N-m)}. \quad \square$$

We conclude this subsection with the proof of Theorem 2.1, which now amounts to little more than summing the inequalities from the assertion of Corollary 2.13 over a range of  $d$ .

*Proof of Theorem 2.1.* Let  $\delta, \varepsilon \in (0, 1)$  be arbitrary and suppose that nonnegative integers  $d, m, n, r$  satisfy the assumptions of the theorem. Note that we may assume that  $d \leq \sqrt{m/n}$  and that  $n$  is sufficiently large, as otherwise the asserted upper bound on the probability is larger than one. Let  $G \sim G(n, m)$  and let  $G' \sim G(n, m+d+1)$ . It follows from Corollary 2.13 that

$$\begin{aligned} \mathbb{E} |\text{eq}_d^r(G)| - \mathbb{E} |\text{eq}_{-1}^r(G')| &\geq - \sum_{i=0}^d \frac{2n^{9/2}}{\sqrt{m+i+1} \cdot (N-m-i)} \\ &\geq - \frac{2(d+1)n^{9/2}}{\sqrt{m} \cdot (\delta N - d)} \geq - \frac{5(d+1)n^{5/2}}{\delta \sqrt{m}}. \end{aligned}$$

In particular, due to (3),

$$\mathbb{E} |\text{eq}_d^r(G)| \geq \frac{N}{r} - \frac{n}{2} - \frac{5(d+1)n^{5/2}}{\delta \sqrt{m}}. \quad (6)$$

On the other hand, since  $(d/m)^{1/2} \leq (mn)^{-1/4} \leq (n/m)^{1/4}$ , Lemma 2.10 implies that, with probability at least  $1 - e^{-n}$ , each part of every  $r$ -cut with deficit at most  $d$  has at most  $n/r + 4(n/m)^{1/4} \cdot n$  vertices. Thus, letting  $a := 4(n/m)^{1/4} \cdot n$ , we have

$$\begin{aligned} |\text{eq}_d^r(G)| &\leq r \cdot \binom{n/r + a}{2} \leq r \cdot \binom{n/r}{2} + r \cdot \left(\frac{n}{r} + \frac{a}{2}\right) a \leq \frac{N}{r} + na + \frac{ra^2}{2} \\ &= \frac{N}{r} + 4(n/m)^{1/4} \cdot n^2 + 8r(n/m)^{1/2} \cdot n^2. \end{aligned}$$

with probability at least  $1 - e^{-n}$ . Finally, let

$$P := \mathbb{P}(G \text{ is not } (d, r, \varepsilon)\text{-rigid}) = \mathbb{P}(\text{eq}_d^r(G) \leq (1 - \varepsilon)N/r).$$

We may conclude that, for some absolute constant  $C$ ,

$$\mathbb{E} |\text{eq}_d^r(G)| \leq P \cdot \frac{(1 - \varepsilon)N}{r} + (1 - P) \cdot \left(\frac{1}{r} + C\sqrt[4]{\frac{n}{m}} + Cr\sqrt{\frac{n}{m}}\right) \cdot N + e^{-n} \cdot N. \quad (7)$$

Combining (6) and (7), we obtain

$$P \leq \frac{r}{\varepsilon} \cdot \left(\frac{5(d+1)n^{5/2}}{\delta \sqrt{m}N} + C\sqrt[4]{\frac{n}{m}} + Cr\sqrt{\frac{n}{m}}\right),$$

provided that  $C$  is sufficiently large, as claimed.  $\square$

**2.4. Derivation of Corollary 2.2.** The following key lemma, which is a variant of [6, Proposition 10.1] (see also [14, Proposition 7.2]), implies that every graph that is

sufficiently rigid must contain a core, provided that each of its cuts with small deficit is balanced.

**Lemma 2.14.** *Suppose that  $r \geq 2$  and  $\alpha \in (0, 1/(r^2 + r))$  and let  $d \geq 0$  be an arbitrary integer. If an  $n$ -vertex graph  $G$  is  $(d, r, \alpha/r)$ -rigid and each part of every  $r$ -cut of  $G$  with deficit at most  $d$  has size at most  $(1 + \alpha)n/r$ , then the  $r$  largest  $(d, r)$ -components of  $G$  have size at least  $n/r - \alpha n$  each.*

*Proof.* Suppose that each part of every  $r$ -cut of  $G$  with deficit at most  $d$  has size at most  $(1 + \alpha)n/r$ . If only at most  $r - 1$  largest  $(d, r)$ -components of  $G$  had size at least  $n/r - \alpha n$ , then it would follow from convexity of  $x \mapsto \binom{x}{2}$  that

$$\begin{aligned} |\text{eq}_d^r(G)| &\leq (r - 1) \binom{(1 + \alpha)n/r}{2} + \binom{n/r - \alpha n}{2} + \binom{\alpha n/r}{2} \\ &\leq \left( (r - 1) \left( \frac{(1 + \alpha)}{r} \right)^2 + \left( \frac{1}{r} - \alpha \right)^2 + \left( \frac{\alpha}{r} \right)^2 \right) \binom{n}{2} \\ &= \frac{1 - 2\alpha/r + (r + 1)\alpha^2}{r} \binom{n}{2} < \frac{1 - \alpha/r}{r} \binom{n}{2}, \end{aligned}$$

which means that  $G$  would not be  $(d, r, \alpha/r)$ -rigid.  $\square$

*Proof of Corollary 2.2.* Let  $\alpha, \delta \in (0, 1)$  be arbitrary and suppose that nonnegative integers  $d, m, n, r$  satisfy the assumptions of the corollary. Note that we may assume that  $d \leq \sqrt{m/n}$  and  $4r^2/\alpha < \sqrt[4]{m/n}$ , as otherwise the asserted upper bound on the probability is larger than one. Since these assumptions imply that

$$\frac{\alpha n}{r} > 4 \cdot \sqrt[4]{\frac{n}{m}} \cdot n = 4 \cdot \max \left\{ \sqrt{\frac{d}{m}}, \sqrt[4]{\frac{n}{m}} \right\} \cdot n,$$

we may invoke Lemma 2.10 to conclude that, with probability  $1 - e^{-n}$ , each part of every  $r$ -cut in  $G(n, m)$  with deficit at most  $d$  has no more than  $(1 + \alpha)n/r$  vertices. Consequently, by Lemma 2.14,

$$\mathbb{P}(G(n, m) \notin \text{CORE}_d^r(\alpha)) \leq \mathbb{P}(G(n, m) \text{ is not } (d, r, \alpha/r)\text{-rigid}) + e^{-n}.$$

The assertion of the corollary now follows immediately from Theorem 2.1.  $\square$

**2.5. Proof of Theorem 1.1.** Since our assumption that  $pn \gg 1$  implies that whp  $G_{n,p}$  has  $(1 + o(1))\binom{n}{2}p$  edges, the first assertion of the theorem is a straightforward consequence of Corollary 2.2. It thus remains to prove the second assertion. The following observation is at the heart of the matter. Given a graph  $G$ , a vertex  $v \in V(G)$ , and a set  $W \subseteq V(G)$ , we write  $\deg_G(v, W)$  to denote the number of neighbours of  $v$  in  $W$ .

**Lemma 2.15.** *Suppose that some largest  $r$ -cut  $\Pi = \{V_1, \dots, V_r\}$  in a graph  $G$  and a vertex  $v \in V_1$  satisfy*

$$\deg_G(v, V_1) \geq \frac{\deg_G v}{r} - \frac{r - 1}{r} \cdot d$$

*for some nonnegative integer  $d$ . Then  $v$  is not  $(d, r)$ -equivalent to any other vertex of  $G$ .*

*Proof.* For each  $i \in \{2, \dots, r\}$ , denote by  $\Pi_i$  the  $r$ -cut obtained from  $\Pi$  by moving the vertex  $v$  from  $V_1$  to  $V_i$  and observe that

$$\begin{aligned} |\text{ext}(\Pi_i) \cap G| &= |\text{ext}(\Pi) \cap G| + \deg_G(v, V_1) - \deg_G(v, V_i) \\ &= b_r(G) + \deg_G(v, V_1) - \deg_G(v, V_i). \end{aligned}$$

Consequently, writing  $\text{def}_G(\Pi_i) := b_r(G) - |\text{ext}(\Pi_i) \cap G|$  for the deficit of  $\Pi_i$  in  $G$ , we have

$$\sum_{i=2}^r \text{def}_G(\Pi_i) = \sum_{i=2}^r (\deg_G(v, V_i) - \deg_G(v, V_1)) = \deg_G v - r \deg_G(v, V_1) \leq (r-1)d.$$

In particular, there must be some  $i \in \{2, \dots, r\}$  for which  $\text{def}_G(\Pi_i) \leq d$ . The assertion now follows as no vertex  $u \neq v$  shares with  $v$  the same part of both  $\Pi$  and  $\Pi_i$ .  $\square$

Suppose now that  $G \sim G_{n,p}$ . Let  $\Pi$  be an arbitrary largest  $r$ -cut in  $G$  and define, for each nonnegative integer  $d$ ,

$$X_d := \left\{ v \in V(G) : v \in V \in \Pi \text{ and } \deg_G(v, V) < \frac{\deg_G v}{r} - \frac{r-1}{r} \cdot d \right\}.$$

Since Lemma 2.15 implies that  $\text{eq}_d^r(G) \subseteq (X_d)$ , it suffices to show that whp  $|X_d| < \alpha n$  when  $d \geq C\sqrt{pn}$  for large enough  $C$ .

The assumption that  $\Pi$  is a largest  $r$ -cut in  $G$  implies that  $\deg_G(v, V) \leq (\deg_G v)/r$  for each  $v \in V \in \Pi$  and therefore

$$\frac{r-1}{r} \cdot (2|G| + |X_d| \cdot d) \leq \sum_{v \in V \in \Pi} (\deg_G v - \deg_G(v, V)) = 2|\text{ext}(\Pi) \cap G|.$$

This yields

$$|X_d| \leq \frac{2r \cdot |\text{ext}(\Pi) \cap G| - 2(r-1) \cdot |G|}{(r-1) \cdot d} = \frac{2|G| - 2r \cdot |\text{int}(\Pi) \cap G|}{(r-1) \cdot d}.$$

In particular, it follows from Corollary 2.9 that, with probability at least  $1 - e^{-n}$ ,

$$|X_d| \leq \frac{2|G| - 2r \cdot |\text{int}(\Pi)| \cdot |G|/\binom{n}{2} + 4r \cdot n\sqrt{n|G|/\binom{n}{2}}}{(r-1) \cdot d}.$$

Finally, since  $r \cdot |\text{int}(\Pi)| \geq r^2 \cdot \binom{n/r}{2} \geq \binom{n}{2} - \frac{rn}{2}$ , by convexity of  $x \mapsto \binom{x}{2}$ , and since whp we have  $|G| \leq 2\binom{n}{2}p$  and

$$|X_d| \leq \frac{2rnp + 6rn\sqrt{np}}{(r-1) \cdot d} < \frac{7rn\sqrt{np}}{(r-1) \cdot d}.$$

In particular, if  $d \geq 14\sqrt{np}/\alpha$ , then  $|X_d| < \alpha n$ , as desired.

**2.6. Non-uniqueness of maximum cuts and vertices outside of the core.** In this subsection, we present supporting evidence for our suspicion that whp  $G(n, m)$  does not have a unique maximum cut for all  $m$  satisfying  $n \ll m \leq N - \Omega(N)$ .<sup>6</sup> By [3, Lemma 14], with probability very close to one, the maximum size of an  $r$ -cut in  $G(n, m)$  is  $\frac{r-1}{r} \cdot m + \Theta(\sqrt{mn})$ . It is reasonable to believe that the expectation of this random

<sup>6</sup>This is vacuously true when  $m \leq n \log n/2 - \omega(n)$ , since then whp  $G(n, m)$  has isolated vertices.

variable grows ‘smoothly’ with  $m$  in the sense that  $\mathbb{E}[b_r(G(n, m+1)) - b_r(G(n, m))] = \frac{r-1}{r} + \Theta(\sqrt{n/m})$ . If this estimate held conditionally on  $G(n, m)$ , under the natural coupling for which  $G(n, m) \subseteq G(n, m+1)$ , that is, if

$$\partial b_r := \mathbb{E}[b_r(G(n, m+1)) \mid G(n, m)] - b_r(G(n, m)) \geq \frac{r-1}{r} + \Theta(\sqrt{n/m}), \quad (8)$$

then  $G(n, m)$  would not have a unique maximum  $r$ -cut. Indeed, it follows from Fact 2.4 that, if  $\Pi$  is the unique maximum  $r$ -cut of  $G(n, m)$ , then

$$\partial b_r = \frac{|\text{ext}(\Pi) \setminus G(n, m)|}{\binom{n}{2} - m} \leq \frac{r-1}{r}.$$

Unfortunately, we do not even know how to prove that (8) holds with probability  $\Omega(1)$ . Instead, we show that the expected number of vertices that lie outside of the core is large in either  $G(n, m)$  or in  $G(n, m+1)$ , and thus also in  $G_{n,p}$ .

Given a graph  $G$  and an integer  $r \geq 2$ , denote by  $X_r(G)$  the set of vertices not belonging to the  $(0, r)$ -core of  $G$  (so that  $X_r(G) = V(G)$  if  $G$  does not have a  $(0, r)$ -core). Further, set  $x_r(G) := |X_r(G)|$ . By the definition of the  $(0, r)$ -core,  $x_r(G) = 0$  if and only if  $G$  has a unique largest  $r$ -cut whose each part has more than  $|V(G)|/(r+1)$  vertices.<sup>7</sup> The following proposition is the main result of this section.

**Proposition 2.16.** *Let  $r \geq 2$  be an integer and let  $\delta$  be a positive constant. If  $1/n \ll p \leq 1 - \delta$ , then*

$$\mathbb{E}[x_r(G_{n,p})] = \Omega\left(\sqrt{n/p}\right).$$

*Proof.* Let  $N := \binom{n}{2}$ . We will show that, for all  $n \ll m \leq (1 - \delta)N$ ,

$$\mathbb{E}[x_r(G(n, m)) + x_r(G(n, m+1))] = \Omega\left(\sqrt{n^3/m}\right), \quad (9)$$

which clearly implies the assertion of the proposition. As in the proof of Lemma 2.12, consider the coupling of  $G_1 \sim G(n, m)$  and  $G_2 \sim G(n, m+1)$  such that  $e \in K_n$  is an edge chosen uniformly at random among the non-edges of  $G_1$  and the edges of  $G_2$ . Recall from Fact 2.4 that

$$\mathbb{P}(e \in \text{crit}_r(G_2)) = \mathbb{P}(e \notin \text{eq}_0^r(G_1)). \quad (10)$$

Writing  $\partial_{G_2}(X_r(G_2))$  for the set of edges of  $G_2$  with at least one endpoint in  $X_r(G_2)$ , we have

$$|\text{crit}_r(G_2)| \geq |\text{ext}(\text{core}_0^r(G_2)) \cap G_2| \geq b_r(G_2) - |\partial_{G_2}(X_r(G_2))|$$

and thus

$$\mathbb{P}(e \in \text{crit}_r(G_2)) \geq \mathbb{E}\left[\frac{b_r(G_2) - |\partial_{G_2}(X_r(G_2))|}{m+1}\right].$$

On the other hand,

$$\mathbb{P}(e \in \text{eq}_0^r(G_1)) \geq \mathbb{P}(e \in \text{int}(\text{core}_0^r(G_1))) = \mathbb{E}\left[\frac{|\text{int}(\text{core}_0^r(G_1)) \setminus G_1|}{N - m}\right]$$

<sup>7</sup>Lemma 2.10 implies that, for all  $m \gg n$ , with probability at least  $1 - e^{-n}$ , each part of every maximum  $r$ -cut of  $G(n, m)$  has  $n/r \pm o(n)$  vertices.

while

$$\begin{aligned}
|\text{int}(\text{core}_0^r(G_1)) \setminus G_1| &= |\text{int}(\text{core}_0^r(G_1))| - |\text{int}(\text{core}_0^r(G_1)) \cap G_1| \\
&\geq |\text{int}(\text{core}_0^r(G_1))| - \frac{m}{r} \geq r \cdot \binom{(n - x_r(G_1))/r}{2} - \frac{m}{r} \\
&= \frac{(n - x_r(G_1))(n - x_r(G_1) - r)}{2r} - \frac{m}{r} \\
&\geq \frac{N - n \cdot (x_r(G_1) + r) - m}{r},
\end{aligned}$$

where the first inequality is true since the edges in  $\text{int}(\text{core}_0^r(G_1)) \cap G_1$  do not cross any maximum  $r$ -cut of  $G_1$  and  $b_r(G_1) \geq \frac{r-1}{r} \cdot m$  with probability one. Hence,

$$\begin{aligned}
\mathbb{P}(e \notin \text{eq}_0^r(G_1)) &\leq 1 - \frac{N - n \cdot (\mathbb{E}[x_r(G_1)] + r) - m}{r \cdot (N - m)} \\
&= \frac{r-1}{r} + \frac{n \cdot (\mathbb{E}[x_r(G_1)] + r)}{r \cdot (N - m)}.
\end{aligned}$$

Substituting the above bounds into (10) and taking expectation, we conclude that

$$\frac{n \cdot (\mathbb{E}[x_r(G_1)] + r)}{r \cdot (N - m)} + \frac{\mathbb{E}|\partial_{G_2}(X_r(G_2))|}{m+1} \geq \frac{\mathbb{E}[b_r(G_2)]}{m+1} - \frac{r-1}{r}.$$

Since  $\mathbb{E}[b_r(G_2)] = \frac{r-1}{r} \cdot m + \Theta(\sqrt{nm})$  when  $n \ll m \leq (1-\delta)N$ , by [3, Lemma 14],

$$\frac{\mathbb{E}[x_r(G_1)]}{n} + \frac{\mathbb{E}|\partial_{G_2}(X_r(G_2))|}{m} \geq c \cdot \sqrt{\frac{n}{m}}, \quad (11)$$

for some  $c > 0$  that depends only on  $r$  and  $\delta$ . The following claim bounds the second summand in the left-hand side of (11) from above by an almost-linear function of  $\mathbb{E}[x_r(G_2)]$ .

**Claim 2.17.** *There exists a constant  $C$  such that*

$$\mathbb{E}|\partial_{G_2}(X_r(G_2))| \leq \frac{4m}{n} \cdot \mathbb{E}[x_r(G_2)] + C \sqrt{\frac{m}{n}} \cdot \mathbb{E}[x_r(G_2)] \log \left( \frac{en}{\mathbb{E}[x_r(G_2)]} \right) + 1.$$

We first show how (11) and Claim 2.17 imply (9). Substituting the upper bound on  $\mathbb{E}|\partial_{G_2}(X_r(G_2))|$  into (11), we obtain

$$\mathbb{E}[x_r(G_1)] + \mathbb{E}[x_r(G_2)] + \sqrt{\frac{n}{m}} \cdot \mathbb{E}[x_r(G_2)] \log \left( \frac{en}{\mathbb{E}[x_r(G_2)]} \right) \geq \tilde{c} \cdot \sqrt{\frac{n^3}{m}} \quad (12)$$

for some  $\tilde{c}$  that depends only on  $r$  and  $\delta$ . Suppose that  $\mathbb{E}[x_r(G_2)] \leq \tilde{c}/3 \cdot \sqrt{n^3/m} \ll n$ . Since the function  $x \mapsto x \log(\frac{en}{x})$  is increasing on  $(0, n]$ , we have

$$\sqrt{\frac{n}{m}} \cdot \mathbb{E}[x_r(G_2)] \log \left( \frac{en}{\mathbb{E}[x_r(G_2)]} \right) \leq \frac{\tilde{c}}{3} \cdot \sqrt{\frac{n^3}{m}} \cdot \sqrt{\frac{n}{m}} \log \left( \frac{3e}{\tilde{c}} \cdot \sqrt{\frac{m}{n}} \right) \leq \frac{\tilde{c}}{3} \cdot \sqrt{\frac{n^3}{m}},$$

where the second inequality holds as  $m \gg n$  and  $\lim_{x \rightarrow 0} x \log(1/x) = 0$ . This means, by (12), that  $\mathbb{E}[x_r(G_1)] \geq \tilde{c}/3 \cdot \sqrt{n^3/m}$ , as desired. We now prove the claim.

It suffices to show that, with probability at least  $1 - 1/n^2$ , for every nonempty set  $A \subseteq V(G_2)$ ,

$$|\partial_{G_2}(A)| \leq \frac{m}{N} \cdot |A|n + C\sqrt{\frac{m}{n}} \cdot |A| \log\left(\frac{en}{|A|}\right), \quad (13)$$

where  $C$  is a large constant. Indeed, this estimate yields

$$\begin{aligned} \mathbb{E}|\partial_{G_2}(X_r(G_2))| &\leq \frac{4m}{n} \cdot \mathbb{E}[x_r(G_2)] + C\sqrt{\frac{m}{n}} \cdot \mathbb{E}\left[x_r(G_2) \log\left(\frac{en}{x_r(G_2)}\right)\right] + \frac{1}{n^2} \binom{n}{2} \\ &\leq \frac{4m}{n} \cdot \mathbb{E}[x_r(G_2)] + C\sqrt{\frac{m}{n}} \cdot \mathbb{E}[x_r(G_2)] \log\left(\frac{en}{\mathbb{E}[x_r(G_2)]}\right) + 1, \end{aligned}$$

where the second inequality follows from Jensen's inequality applied to the concave function  $x \mapsto x \log(en/x)$ .

Note that, for every nonempty set  $A \subseteq V(G_2)$ , the random variable  $|\partial_{G_2}(A)|$  has hypergeometric distribution with mean at most  $m/N \cdot |A|n$ . In particular, standard tail estimates yield (writing  $p = m/N$  and  $|A| = k$ ), for every  $t \geq 0$ ,

$$\begin{aligned} \mathbb{P}\left(|\partial_{G_2}(A)| \geq \frac{m}{N} \cdot |A|n + t\right) &\leq \mathbb{P}(\text{Bin}(kn, p) \geq p \cdot kn + t) \leq \exp\left(\frac{-t^2}{2knp + t}\right) \\ &\leq \exp\left(\frac{-t^2}{4knp}\right) + \exp\left(-\frac{t}{2}\right). \end{aligned}$$

Invoking the above estimate with  $t := C\sqrt{m/n} \cdot |A| \log(en/|A|)$ , we conclude that the probability that (13) fails for some nonempty set  $A$  is at most

$$\begin{aligned} \sum_{k=1}^n \binom{n}{k} \left( \exp\left(-\frac{C^2 k N (\log(en/k))^2}{4n^2}\right) + \exp\left(-Ck \log\left(\frac{en}{k}\right)\right) \right) \\ \leq 2 \sum_{k=1}^n \left(\frac{en}{k}\right)^{k-Ck} \leq n^{-2}, \end{aligned}$$

where the last inequality is true whenever  $C$  is sufficiently large.  $\square$

### 3. NEIGHBOURHOOD IN MAXIMUM CUTS

In this section, we prove Theorem 1.2. It will be convenient to first introduce some notation. Given a graph  $G$ , a vertex  $v \in V(G)$ , and a set of vertices  $S \subseteq V(G)$ , we denote by  $N_G(v)$  and  $N_G(v; S)$  the sets of neighbours of  $v$  in  $V(G)$  and in  $S$ , respectively. Further, given a set  $W \subseteq V(G)$ , we denote by  $N_G(W)$  and by  $N_G(W; S)$  the set of common neighbours of  $W$  in  $V(G)$  and in  $S$ , respectively; note that usually  $N_G(W)$  denotes the union of  $N_G(v)$  over  $v \in W$ , rather than their intersection. That is,

$$N_G(W) := \bigcap_{v \in W} N_G(v) \quad \text{and} \quad N_G(W; S) := \bigcap_{v \in W} N_G(v; S).$$

Finally, we will denote by  $\text{Hyp}(N, s, t)$  the size of the intersection of a random  $t$ -element subset of  $\llbracket N \rrbracket$  with the set  $\llbracket s \rrbracket$  (this is a hypergeometric random variable).

Let us fix integers  $k \geq 1$ ,  $r \geq 2$  and a small  $\varepsilon > 0$ . Further, assume that  $p$  satisfies

$$C \left( \frac{\log n}{n} \right)^{1/k} \leq p \leq 1 - \varepsilon \quad \text{and} \quad p \gg \frac{(\log n)^2}{n}$$

for a sufficiently large  $C > 0$ . Our aim is to prove that whp every maximum  $r$ -cut  $(V_1, \dots, V_r)$  of  $\mathbf{G} \sim G_{n,p}$  and every  $k$ -element  $W \subseteq V(G)$  satisfy

$$|N_{\mathbf{G}}(W; V_i)| = \left( \frac{1}{r} \pm \varepsilon \right) np^k \quad \text{for all } i \in [r].$$

Denote by  $\mathcal{B}$  the event that there exists a set  $W$  of  $k$  vertices and a maximum  $r$ -cut  $(V_1, \dots, V_r)$  of  $\mathbf{G}$  such that, for some  $i \in [r]$ ,

$$|N_{\mathbf{G}}(W; V_i)| \leq \frac{1 - \varepsilon}{r} \cdot |N_{\mathbf{G}}(W)|.$$

Moreover, for a  $k$ -set  $W \subseteq V(\mathbf{G})$ , denote by  $\mathcal{Y}_W$  the event that  $|N_{\mathbf{G}}(W)| = (1 \pm \varepsilon^2)np^k$  and set  $\mathcal{Y} := \bigcap_W \mathcal{Y}_W$ . Since on the event  $\mathcal{B}^c \cap \mathcal{Y}$ , we have

$$\begin{aligned} \left( \frac{1}{r} - \varepsilon \right) np^k &\leq \frac{1 - \varepsilon}{r} \cdot (1 - \varepsilon^2) np^k \leq |N_{\mathbf{G}}(W; V_i)| \leq |N_{\mathbf{G}}(W)| - \sum_{j \neq i} |N_{\mathbf{G}}(W; V_j)| \\ &\leq (1 + \varepsilon^2) np^k - (r - 1) \cdot \frac{1 - \varepsilon}{r} np^k \leq \left( \frac{1}{r} + \varepsilon \right) np^k, \end{aligned}$$

in order to complete the proof of Theorem 1.2, it is sufficient to prove that

$$\mathbb{P}(\mathcal{Y}^c) + \mathbb{P}(\mathcal{B}) = o(1). \quad (14)$$

By the Chernoff bound and our assumption on  $p$  and the choice of  $C$ , for every  $k$ -element  $W \subseteq V(G)$ ,

$$\mathbb{P} \left( |N_{\mathbf{G}}(W)| \neq (1 \pm \varepsilon^2)np^k \right) \leq \exp \left( -\varepsilon^4 np^k / 5 \right) \leq n^{-2k};$$

the estimate  $\mathbb{P}(\mathcal{Y}^c) = o(1)$  follows from a straightforward union bound over all  $W$ .

We now turn to bounding  $\mathbb{P}(\mathcal{B})$ . Let  $\alpha$  be a small positive constant. Recall that  $\mathbf{G} \in \text{CORE}_d^r(\alpha)$  means that  $\mathbf{G}$  has a  $(d, r)$ -core whose each component has at least  $n/r - \alpha n$  vertices and that  $\text{core}_d^r(\mathbf{G})$  denotes the set of the  $r$  components of the  $(d, r)$ -core of  $\mathbf{G}$ . (If  $\mathbf{G} \notin \text{CORE}_d^r(\alpha)$ , then we set  $\text{core}_d^r(\mathbf{G}) = \emptyset$ .) For a  $k$ -set  $W \subseteq V(\mathbf{G})$ , denote by  $\mathcal{A}_W$  the event that

$$|N_{\mathbf{G}}(W; S)| \leq \frac{1 - \varepsilon}{r} \cdot |N_{\mathbf{G}}(W)| \quad \text{for some } S \in \text{core}_0^r(\mathbf{G}) \quad (15)$$

and let  $\mathcal{A} := \bigcup_W \mathcal{A}_W$ . Set  $t := \lceil \sqrt{C} \log n \rceil$  and let  $d := kt$ . The key observation is that, if  $\mathbf{G}$  has a  $(0, r)$ -core and  $\mathcal{B}$  holds, then  $\mathcal{A}$  holds as well. Therefore,

$$\mathbb{P}(\mathcal{B}) \leq (\mathbf{G} \notin \text{CORE}_{2d}^r(\alpha)) + \mathbb{P}(\mathbf{G} \in \text{CORE}_{2d}^r(\alpha) \wedge \mathcal{A}).$$

The first summand above is  $o(1)$ , by Corollary 2.2, whereas

$$\mathbb{P}(\mathbf{G} \in \text{CORE}_{2d}^r(\alpha) \wedge \mathcal{A}) \leq \mathbb{P}(\mathcal{Y}^c) + \sum_W \mathbb{P}(\mathbf{G} \in \text{CORE}_{2d}^r(\alpha) \wedge \mathcal{A}_W \mid \mathcal{Y}_W). \quad (16)$$

Since we have already shown that  $\mathbb{P}(\mathcal{Y}^c) = o(1)$ , it suffices to prove that each term in the sum is much smaller than  $n^{-k}$ .

To this end, fix a  $k$ -element set  $W$  and, assuming that  $\mathcal{Y}_W$  holds, let  $\mathbf{R}$  be a uniformly chosen  $t$ -element subset of  $N_{\mathbf{G}}(W)$  (clearly,  $\mathcal{Y}_W$  implies that  $|N_{\mathbf{G}}(W)| \geq t$ ). Denote by  $\mathcal{A}_{W,t}$  the event that

$$|\mathbf{R} \cap S| \leq \frac{1 - \varepsilon/2}{r} \cdot t \quad \text{for some } S \in \text{core}_0^r(\mathbf{G}) \quad (17)$$

and observe that

$$\begin{aligned} \mathbb{P}(\mathbf{G} \in \text{CORE}_{2d}^r(\alpha) \wedge \mathcal{A}_{W,t} \mid \mathcal{Y}_W) &\geq \mathbb{P}(\mathbf{G} \in \text{CORE}_{2d}^r(\alpha) \wedge \mathcal{A}_W \mid \mathcal{Y}_W) \\ &\quad \cdot \mathbb{P}(\mathcal{A}_{W,t} \mid \mathbf{G} \in \text{CORE}_{2d}^r(\alpha) \wedge \mathcal{A}_W \wedge \mathcal{Y}_W). \end{aligned} \quad (18)$$

If  $\mathcal{Y}_W$  holds (so that  $\mathbf{R}$  is defined) and  $\mathbf{G}$  has a  $(0, r)$ -core, then, for every  $S \in \text{core}_0^r(\mathbf{G})$ , the cardinality  $|\mathbf{R} \cap S|$  is distributed like  $\text{Hyp}(|N_{\mathbf{G}}(W)|, |N_{\mathbf{G}}(W; S)|, t)$ . In particular, it follows from standard tail bounds for the hypergeometric distribution [13, Section 6] that whp (in the choice of  $\mathbf{R}$ ) (15) implies (17). Thus, the second factor in the right-hand side of (18) approaches 1, implying

$$\mathbb{P}(\mathbf{G} \in \text{CORE}_{2d}^r(\alpha) \wedge \mathcal{A}_W \mid \mathcal{Y}_W) \leq (1 + o(1)) \cdot \mathbb{P}(\mathbf{G} \in \text{CORE}_{2d}^r(\alpha) \wedge \mathcal{A}_{W,t} \mid \mathcal{Y}_W). \quad (19)$$

Now, if the distribution of  $\mathbf{R}$ , conditioned on  $\text{core}_0^r(\mathbf{G})$ , was the uniform distribution on  $t$ -element subsets of  $V \setminus W$ , then, for every  $S \in \text{core}_0^r(\mathbf{G})$ , the cardinality  $|\mathbf{R} \cap S|$  would again have hypergeometric distribution, implying that  $\mathcal{A}_{W,t}$  has exponentially small probability. We will show something marginally weaker: The distribution of  $\mathbf{R}$  is approximately uniform after we condition on the graph  $\mathbf{G}^*$  that is obtained from  $\mathbf{G}$  by resampling all edges between  $W$  and  $\mathbf{R}$ . Since  $\mathbf{G}^*$  satisfies  $\text{core}_d^r(\mathbf{G}^*) \preceq \text{core}_0^r(\mathbf{G})$ , by Corollary 2.7, this will allow us to prove the following estimate.

**Claim 3.1.** *There exists a  $c = c(r, \alpha, \varepsilon) > 0$  such that, for every  $k$ -element set  $W$ ,*

$$\mathbb{P}(\mathbf{G} \in \text{CORE}_{2d}^r(\alpha) \wedge \mathcal{A}_{W,t} \mid \mathcal{Y}_W) \leq e^{-ct}.$$

Given Claim 3.1, it is now easy to finish the proof of Theorem 1.2. Indeed, by (19),

$$\mathbb{P}(\mathbf{G} \in \text{CORE}_{2d}^r(\alpha) \wedge \mathcal{A}_W \mid \mathcal{Y}_W) \leq (1 + o(1)) \cdot e^{-ct} \ll n^{-k}$$

due to our choice of  $t$ . Substituting this bound into (16) completes the proof.

*Proof of Claim 3.1.* Fix some  $k$ -element set of vertices  $W$ , assume that  $\mathcal{Y}_W$  holds, and let  $\mathbf{G}^*$  be the graph obtained from  $\mathbf{G}$  by independently resampling each edge between  $W$  and  $\mathbf{R}$  with probability  $p$ . We shall first show that, conditioned on  $\mathbf{G}^*$ ,  $\mathbf{R}$  is a nearly-uniformly random  $k$ -element subset of  $\llbracket n \rrbracket \setminus W$ .

To see this, fix some  $t$ -element  $R \subseteq \llbracket n \rrbracket \setminus W$  and a graph  $G^*$  such that  $\mathbf{G}^* = G^*$  with nonzero probability. Writing  $G^*[W, R]$  for the induced bipartite subgraph of  $G^*$  with parts  $W$  and  $R$  and  $[W, R]$  for the complete bipartite graph with parts  $W$  and  $R$ , we

have

$$\begin{aligned}
\mathbb{P}(\mathbf{R} = R \wedge \mathbf{G}^* = G^*) &= \mathbb{P}(\mathbf{G} = G^* \cup [W, R]) \cdot \mathbb{P}(\mathbf{R} = R \mid \mathbf{G} = G^* \cup [W, R]) \\
&\quad \cdot \mathbb{P}(\mathbf{G}^*[W, R] = G^*[W, R] \mid \mathbf{G} = G^* \cup [W, R] \wedge \mathbf{R} = R) \\
&= p^{|G^* \cup [W, R]|} (1-p)^{\binom{n}{2} - |G^* \cup [W, R]|} \cdot \binom{|N_{G^* \cup [W, R]}(W)|}{t}^{-1} \\
&\quad \cdot p^{|G^*[W, R]|} (1-p)^{|W||R| - |G^*[W, R]|} \\
&= p^{|G^*| + kt} (1-p)^{\binom{n}{2} - |G^*|} \cdot \binom{|N_{G^*}(W)| - |N_{G^*}(W) \cap R| + t}{t}^{-1}.
\end{aligned}$$

Since  $|N_{\mathbf{G}^*}(W)| \leq |N_{\mathbf{G}}(W)| \leq (1 + \varepsilon^2)np^k$  and

$$|N_{\mathbf{G}^*}(W)| - t \geq |N_{\mathbf{G}}(W)| - 2t \geq (1 - \varepsilon^2)np^k - 2t \geq (1 - 2\varepsilon^2)np^k,$$

we have, for any two  $t$ -element subsets  $R, R' \subseteq \llbracket n \rrbracket \setminus W$ ,

$$\frac{\mathbb{P}(\mathbf{R} = R \mid \mathbf{G}^* = G^*)}{\mathbb{P}(\mathbf{R} = R' \mid \mathbf{G}^* = G^*)} \leq \frac{\binom{|N_{G^*}(W)| + t}{t}}{\binom{|N_{G^*}(W)|}{t}} \leq \left( \frac{|N_{G^*}(W)|}{|N_{G^*}(W)| - t} \right)^t \leq \left( \frac{1 + \varepsilon^2}{1 - 2\varepsilon^2} \right)^t.$$

In particular, for every  $\mathcal{X} \subseteq \binom{\llbracket n \rrbracket \setminus W}{t}$ ,

$$\mathbb{P}(\mathbf{R} \in \mathcal{X} \mid \mathbf{G}^*) \leq |\mathcal{X}| \cdot \binom{n-k}{t}^{-1} \cdot e^{4\varepsilon^2 t}. \quad (20)$$

Now, since  $|\mathbf{G} \triangle \mathbf{G}^*| \leq |\mathbf{R}||W| = kt = d$ , Corollary 2.7 yields that  $\mathbf{G} \in \text{CORE}_{2d}^r(\alpha)$  implies that  $\mathbf{G}^* \in \text{CORE}_d^r(\alpha)$  and that  $\text{core}_d^r(\mathbf{G}^*) \preceq \text{core}_0^r(\mathbf{G})$ . In particular, the event  $\mathbf{G} \in \text{CORE}_{2d}^r(\alpha) \wedge \mathcal{A}_{W,t}$  implies the event

$$|\mathbf{R} \cap S^*| \leq \frac{1 - \varepsilon/2}{r} \cdot t \quad \text{for some } S^* \in \text{core}_d^r(\mathbf{G}^*),$$

which we will denote by  $\mathcal{A}_{W,t}^*$ . To summarise, we have

$$\begin{aligned}
\mathbb{P}(\mathbf{G} \in \text{CORE}_{2d}^r(\alpha) \wedge \mathcal{A}_{W,t} \mid \mathcal{Y}_W) &\leq \mathbb{P}(\mathbf{G}^* \in \text{CORE}_d^r(\alpha) \wedge \mathcal{A}_{W,t}^*) \\
&\leq \mathbb{P}(\mathcal{A}_{W,t}^* \mid \mathbf{G}^* \in \text{CORE}_d^r(\alpha)).
\end{aligned}$$

Finally, we conclude from (20) that

$$\begin{aligned}
\mathbb{P}(\mathcal{A}_{W,t}^* \mid \mathbf{G}^* \in \text{CORE}_d^r(\alpha)) &\leq e^{4\varepsilon^2 t} \max_{\mathbf{G}^* \in \text{CORE}_d^r(\alpha)} \sum_{S^* \in \text{core}_d^r(\mathbf{G}^*)} \mathbb{P}\left(\text{Hyp}(n-k, |S^* \setminus W|, t) \leq \frac{1 - \varepsilon/2}{r} \cdot t\right) \\
&\leq e^{4\varepsilon^2 t} \cdot r \cdot \mathbb{P}\left(\text{Hyp}(n-k, (1/r - \alpha)n - k, t) \leq \frac{1 - \varepsilon/2}{r} \cdot t\right) \leq e^{-ct},
\end{aligned}$$

for some constant  $c = c(r, \alpha, \varepsilon) > 0$ , provided that  $\alpha \leq \varepsilon/(4r)$ .  $\square$

#### 4. THE THRESHOLD FOR SIMONOVITS'S THEOREM IN $G_{n,p}$

In this section, we prove Theorem 1.5. We start by setting some notational conventions. As in Section 2, given a collection  $\Pi$  of pairwise-disjoint sets of vertices of a graph,

we denote by  $\text{int}(\Pi)$  the set of pairs of vertices contained in a single set of  $\Pi$  and by  $\text{ext}(\Pi)$  the set of remaining pairs of vertices of  $\bigcup \Pi$ . If  $\Pi$  is a collection of sets of vertices of  $K_n$ , we also define

$$\text{ext}^*(\Pi) := K_n \setminus \text{int}(\Pi)$$

and note that  $\text{ext}^*(\Pi) \supseteq \text{ext}(\Pi)$  and that  $\text{ext}^*(\Pi) = \text{ext}(\Pi)$  whenever  $\bigcup \Pi = K_n$ . Let  $H$  be a nonbipartite, edge-critical, strictly 2-balanced graph and denote  $r := \chi(H) - 1$ . As in [14], we denote by  $\mathcal{H}$  the collection of all copies of  $H$  in  $K_n$ , viewed as a hypergraph with vertex set  $K_n$ . Further, for each  $e \in K_n$ , let

$$\partial_e \mathcal{H} := \{K \setminus e : e \in K \in \mathcal{H}\}$$

denote the copies of all proper, spanning subgraphs of  $H$  that form a copy of  $H$  together with the edge  $e$ . Finally, we write  $\partial_e \mathcal{H}[G']$  to denote the subhypergraph of  $\partial_e \mathcal{H}$  that is induced by the subgraph  $G' \subseteq K_n$ .

Let  $\mathbf{G} \sim G_{n,p}$ . Suppose first that  $1/n \ll p \ll n^{-1/m_2(H)}$ . It follows from the definition of 2-density and Markov's inequality that whp  $\mathbf{G}$  contains only  $o(n^2 p)$  copies of  $H$ ; consequently, whp  $\mathbf{G}$  contains an  $H$ -free subgraph with  $(1 - o(1))\binom{n}{2}p$  edges. On the other hand, Corollary 2.9 implies that whp the largest size of an  $r$ -cut in  $\mathbf{G}$  is at most

$$\max_{\Pi: r\text{-cut}} |\text{ext}(\Pi)| \cdot p + 3n\sqrt{np} = \left(1 - \frac{1}{r} + o(1)\right) \binom{n}{2} p.$$

Therefore,  $\mathbf{G}$  is whp not  $H$ -Simonovits. We may thus assume from now on that

$$\varepsilon n^{-1/m_2(H)} \leq p \leq (\theta_H - \varepsilon) \cdot n^{-1/m_2(H)} \cdot (\log n)^{1/(\epsilon_H - 1)}$$

for some positive constant  $\varepsilon$ .

It is clearly sufficient to prove that, with high probability, for some largest  $r$ -cut  $\Pi$  of  $\mathbf{G}$  and an edge  $e \in \text{int}(\Pi) \cap \mathbf{G}$ , the graph  $(\text{ext}(\Pi) \cap \mathbf{G}) \cup e$  is  $H$ -free. We will prove a stronger statement.

Note that every largest  $r$ -cut  $\Pi$  of  $\mathbf{G}$  satisfies

$$\text{ext}(\Pi) = K_n \setminus \text{int}(\Pi) \subseteq K_n \setminus \text{int}(\text{core}_0^r(\mathbf{G})) = \text{ext}^*(\text{core}_0^r(\mathbf{G})),$$

provided that  $\mathbf{G}$  has a  $(0, r)$ -core. The aforementioned stronger statement that implies Theorem 1.5 is that whp  $\mathbf{G}$  has a  $(0, r)$ -core with minimum part size  $n/r - o(n)$  and there is an  $e \in \text{int}(\text{core}_0^r(\mathbf{G})) \cap \mathbf{G}$  such that  $\partial_e \mathcal{H}[\text{ext}^*(\text{core}_0^r(\mathbf{G})) \cap \mathbf{G}]$  is empty; note that this implies that  $(\text{ext}(\Pi) \cap \mathbf{G}) \cup e$  is  $H$ -free for every largest  $r$ -cut  $\Pi$  of  $\mathbf{G}$ . More precisely, set  $\alpha := (1/\log n)^2$  and define, for each  $e \in K_n$ , the event

$$\mathcal{Y}_e := \{G \in \text{CORE}_0^r(\alpha) : e \in \text{int}(\text{core}_0^r(G)) \cap G \wedge \partial_e \mathcal{H}[\text{ext}^*(\text{core}_0^r(G)) \cap G] = \emptyset\}.$$

Our goal is to prove that whp  $\mathbf{G} \in \mathcal{Y}_e$  for some  $e \in K_n$ . Denoting by  $Z$  the number of  $e \in K_n$  satisfying  $\mathbf{G} \in \mathcal{Y}_e$ , it will be enough to show that

$$\mathbb{E}[Z^2] \leq (1 + o(1)) \cdot \mathbb{E}[Z]^2. \quad (21)$$

Indeed, if (21) holds, then

$$\mathbb{P}(\mathbf{G} \in \mathcal{Y}_e \text{ for some } e \in K_n) = \mathbb{P}(Z \neq 0) \geq \frac{\mathbb{E}[Z]^2}{\mathbb{E}[Z^2]} = 1 - o(1)$$

by the Paley–Zygmund inequality.

**Proof outline.** In order to establish (21), we separately prove a lower bound on  $\mathbb{E}[Z]$  and an upper bound on  $\mathbb{E}[Z^2]$ . We obtain a lower bound on  $\mathbb{E}[Z]$  using a delicate switching argument that (roughly speaking) goes as follows. We first choose  $d \gg \log n$  so that whp  $\mathbf{G} \in \text{CORE}_{2d}^r(\alpha)$ ; this is possible thanks to Corollary 2.2. Fix some  $e \in K_n$  and assume that  $e \in \text{int}(\text{core}_{2d}^r(\mathbf{G})) \cap \mathbf{G}$ . Our upper-bound assumption on  $p$  and the fact that  $H$  is strictly 2-balanced imply that whp  $\partial_e \mathcal{H}[\mathbf{G}]$  is a matching of size  $O(\log n)$ , which in turn allows us to analyse the following ‘resampling’ process: Remove from  $\mathbf{G}$  all the edges of  $\bigcup \partial_e \mathcal{H}[\mathbf{G}]$ , denote the resulting graph by  $\mathbf{G}^*$ , and consider the conditional distribution of  $\mathbf{G}$  given  $\mathbf{G}^*$ . The fact that  $\partial_e \mathcal{H}[\mathbf{G}]$  is a matching with  $o(d)$  edges allows us to infer that  $\mathbf{G}^* \in \text{CORE}_d^r(\alpha)$  and to essentially couple the conditional distribution of  $\partial_e \mathcal{H}[\mathbf{G}]$  given  $\mathbf{G}^*$  with a  $p^{e_H-1}$ -random subset of  $\partial_e \mathcal{H}$ , giving

$$\mathbb{P}(\partial_e \mathcal{H}[\text{ext}^*(\text{core}_d^r(\mathbf{G}^*)) \cap \mathbf{G}] = \emptyset \mid \mathbf{G}^*) \geq (1 - o(1)) \cdot (1 - p^{e_H-1})^{|\partial_e \mathcal{H}[\text{ext}^*(\text{core}_d^r(\mathbf{G}^*))]|}.$$

Since  $\text{core}_d^r(\mathbf{G}^*)$  is a collection of  $r$  pairwise-disjoint sets of size at least  $n/r - \alpha n$  each, we have

$$|\partial_e \mathcal{H}[\text{ext}^*(\text{core}_d^r(\mathbf{G}^*))]| \leq (1 + O(\alpha)) \cdot \text{Cop}(H, K_r(n/r)^+) = (1 + O(\alpha)) \cdot \pi_H \cdot (n/r)^{v_H-2}.$$

Further, since  $\mathbf{G}$  and  $\mathbf{G}^*$  differ in  $o(d)$  edges, we have  $\text{core}_0^r(\mathbf{G}) \succeq \text{core}_d^r(\mathbf{G}^*)$ , by Corollary 2.7, and thus  $\text{ext}^*(\text{core}_0^r(\mathbf{G})) \subseteq \text{ext}^*(\text{core}_d^r(\mathbf{G}^*))$ ; consequently, since  $n^{v_H-2} p^{e_H-1} = O(\log n)$  and  $\alpha \ll 1/\log n$ ,

$$\begin{aligned} \mathbb{P}(\partial_e \mathcal{H}[\text{ext}^*(\text{core}_0^r(\mathbf{G})) \cap \mathbf{G}] = \emptyset \mid \mathbf{G} \in \text{CORE}_{2d}^r(\alpha) \wedge e \in \text{int}(\text{core}_{2d}^r(\mathbf{G})) \cap \mathbf{G}) \\ \geq \exp(-\pi_H \cdot (n/r)^{v_H-2} \cdot p^{e_H-1} - o(1)). \end{aligned}$$

A lower bound on  $\mathbb{E}[Z]$  now follows by multiplying the above inequality by the probability of the event in the conditioning and summing the result over all  $e \in K_n$ .

In order to prove an upper bound on  $\mathbb{E}[Z^2]$ , we adapt an elegant argument of DeMarco–Kahn [6], which allows us to bound, for every pair  $e, f$  of edges of  $K_n$ , the conditional probability

$$\mathbb{P}((\partial_e \mathcal{H} \cup \partial_f \mathcal{H})[\text{ext}(\text{core}_0^r(\mathbf{G})) \cap \mathbf{G}] = \emptyset \mid \mathbf{G} \in \text{CORE}_0^r(\alpha) \wedge e, f \in \text{int}(\text{core}_0^r(\mathbf{G})) \cap \mathbf{G})$$

from above by the (unconditional) probability of the same event with  $\text{core}_0^r(\mathbf{G})$  replaced by a fixed collection of  $r$  pairwise-disjoint sets of at least  $n/r - \alpha n$  vertices each. The latter probability can be easily shown, using Janson’s inequality, to be at most  $\exp(-2\pi_H \cdot (n/r)^{v_H-2} \cdot p^{e_H-1} + o(1))$ . An upper bound on  $\mathbb{E}[Z^2]$  is then deduced in a straightforward manner by summing the above estimate over all pairs  $e, f$ .

**Organisation.** The remainder of this section is organised as follows. In Section 4.1, we recall the statement of Janson's inequality, prove several useful estimates concerning the hypergraph  $\partial_e \mathcal{H}$ , and derive estimates on the moments of  $|\text{int}(\text{core}_d^r(\mathbf{G})) \cap \mathbf{G}| \cdot \mathbb{1}_{\mathbf{G} \in \text{CORE}_d^r(\alpha)}$  from Corollary 2.2. In the remaining two sections, we prove the lower bound on  $\mathbb{E}[Z]$  and the upper bound on  $\mathbb{E}[Z^2]$ .

**4.1. Preliminaries.** Given a set  $V$  and a real  $p \in [0, 1]$ , we denote by  $V_p$  the random subset of  $V$  obtained by independently retaining each element of  $V$  with probability  $p$ . Further, given a hypergraph  $\mathcal{G}$  with vertex set  $V$ , we define the following two quantities:

$$\mu_p(\mathcal{G}) := \sum_{A \in \mathcal{G}} p^{|A|} \quad \text{and} \quad \Delta_p(\mathcal{G}) := \sum_{\substack{A, B \in \mathcal{G} \\ A \neq B, A \cap B \neq \emptyset}} p^{|A \cup B|},$$

where the second sum is over unordered pairs of edges; in other words,  $\mu_p(\mathcal{G})$  is just the expected number of edges of  $\mathcal{G}[V_p]$  and  $\Delta_p(\mathcal{G})$  is the expected number of pairs of distinct edges of  $\mathcal{G}[V_p]$  that intersect. The following well-known inequality of Janson plays a key role in our proof of the upper bound on  $\mathbb{E}[Z^2]$ .

**Theorem 4.1** (Janson's inequality [15]). *Let  $\mathcal{G}$  be a hypergraph on a finite vertex set  $V$ . For all  $p \in [0, 1]$ ,*

$$\mathbb{P}(\mathcal{G}[V_p] = \emptyset) \leq \exp(-\mu_p(\mathcal{G}) + \Delta_p(\mathcal{G})).$$

We will now prove lower and upper bounds on the values of  $\mu_p$  and  $\Delta_p$  on certain induced subhypergraphs of  $\partial_e \mathcal{H}$  and  $\partial_e \mathcal{H} \cup \partial_f \mathcal{H}$ . We start with an estimate on the sizes of subhypergraphs of  $\partial_e \mathcal{H}$  induced by graphs that are close to a complete, balanced,  $r$ -partite graph.

**Lemma 4.2.** *Let  $\alpha$  be a nonnegative real and suppose that  $\mathcal{C}$  is a family of  $r$  pairwise-disjoint subsets of  $\llbracket n \rrbracket$  such that  $|X| \geq n/r - \alpha n$  for each  $X \in \mathcal{C}$ . There is a constant  $C_H$  that depends only on  $H$  such that, for every  $e \in \text{int}(\mathcal{C})$ ,*

$$\begin{aligned} |\partial_e \mathcal{H}[\text{ext}(\mathcal{C})]| &\geq (\pi_H - C_H \alpha) \cdot (n/r)^{v_H-2} - C_H n^{v_H-3}, \\ |\partial_e \mathcal{H}[\text{ext}^*(\mathcal{C})]| &\leq (\pi_H + C_H \alpha) \cdot (n/r)^{v_H-2}. \end{aligned}$$

*Proof.* Since  $\text{ext}(\mathcal{C}) \supseteq K_r(n/r - \alpha n)$  by our assumption on  $\mathcal{C}$ , when  $e \in \text{int}(\mathcal{C})$ , we have

$$|\partial_e \mathcal{H}[\text{ext}(\mathcal{C})]| \geq \text{Cop}(H, K_r(n/r - \alpha n)^+) \geq \pi_H \cdot (n/r - \alpha n)^{v_H-2} - O(n^{v_H-3}),$$

which implies the first inequality. (The reason why we may write such explicit error term is that  $\text{Cop}(H, K_r(m)^+)$  is a polynomial of degree  $v_H - 2$  in  $m$ .) Further, since<sup>8</sup>

$$\text{ext}^*(\mathcal{C}) \subseteq K_r(n/r - \alpha n) \vee K_{r\alpha n},$$

every copy of  $H$  minus an edge in  $\text{ext}^*(\mathcal{C})$  that is not fully contained in  $K_r(n/r - \alpha n)$  must have at least one vertex in  $K_{r\alpha n}$ . Consequently, there is a constant  $C'_H$  that depends

<sup>8</sup>We write  $G_1 \vee G_2$  for the graph obtained from the disjoint union of  $G_1$  and  $G_2$  by adding all edges joining  $V(G_1)$  and  $V(G_2)$ .

only on  $H$  such that, for each  $e \in \text{int}(\mathcal{C})$ ,

$$|\partial_e \mathcal{H}[\text{ext}^*(\mathcal{C})]| \leq \text{Cop}(H, K_r(n/r)^+) + C'_H r \alpha n \cdot n^{v_H-3} \leq \pi_H \cdot (n/r)^{v_H-2} + C'_H r \alpha n^{v_H-2},$$

which implies the second inequality.  $\square$

Our second lemma supplies an upper bound on  $\Delta_p(\partial_e \mathcal{H} \cup \partial_f \mathcal{H})$ , and thus also on  $\Delta_p(\partial_e \mathcal{H})$ . Even though this upper bound is implicit in [14, Lemma 5.2], we include a (self-contained) proof here for completeness.

**Lemma 4.3.** *For every pair of distinct edges  $e, f \in K_n$  and all  $p \geq \varepsilon n^{-1/m_2(H)}$ ,*

$$\Delta_p(\partial_e \mathcal{H} \cup \partial_f \mathcal{H}) \leq C n^{-\lambda} \cdot (n^{v_H-2} p^{e_H-1})^2$$

for some positive  $\lambda = \lambda(H)$  and  $C = C(H, \varepsilon)$ .

The following non-probabilistic inequality, which is [14, Lemma 4.1], encapsulates the key inequality in the proof of Lemma 4.3.

**Lemma 4.4** ([14]). *Let  $H$  be a nonempty graph and suppose that  $p \geq \varepsilon n^{-1/m_2(H)}$  for some  $\varepsilon > 0$ . Then,  $n^{v_{H'}-2} p^{e_{H'}-1} \geq \varepsilon^{e_{H'}-1}$  for every nonempty subgraph  $H' \subseteq H$ . Moreover, if  $H$  is strictly 2-balanced, then there exists a  $\lambda > 0$  that depends only on  $H$  such that  $n^{v_{H'}-2} p^{e_{H'}-1} \geq \varepsilon^{e_{H'}-1} n^\lambda$  for every  $H' \subseteq H$  with  $1 < e_{H'} < e_H$ .*

*Proof.* Let  $H'$  be a nonempty subgraph of  $H$ . Since the assertion of the lemma holds vacuously if  $e_{H'} = 1$ , we may assume that  $e_{H'} > 1$ . By our assumption on  $p$ ,

$$n^{v_{H'}-2} p^{e_{H'}-1} \geq n^{v_{H'}-2} \left( \varepsilon n^{-\frac{1}{m_2(H)}} \right)^{e_{H'}-1} = \varepsilon^{e_{H'}-1} n^{v_{H'}-2 - \frac{e_{H'}-1}{m_2(H)}}.$$

Now, recall that the definition of  $m_2(H)$  implies that  $v_{H'} - 2 \geq \frac{e_{H'}-1}{m_2(H)}$  and that, when  $H$  is strictly 2-balanced, this inequality is strict unless  $H' = H$ . Thus, the exponent of  $n$  is nonnegative and it is positive if  $H$  is strictly 2-balanced and  $1 < e_{H'} < e_H$ .  $\square$

*Proof of Lemma 4.3.* Observe first that

$$\Delta_p(\partial_e \mathcal{H} \cup \partial_f \mathcal{H}) \leq \Delta_p(\partial_e \mathcal{H}) + \Delta_p(\partial_f \mathcal{H}) + \Delta_p(\partial_e \mathcal{H}, \partial_f \mathcal{H}),$$

where

$$\Delta_p(\partial_e \mathcal{H}, \partial_f \mathcal{H}) := \sum_{K \in \partial_e \mathcal{H}} \sum_{\substack{K' \in \partial_f \mathcal{H} \\ K \cap K' \neq \emptyset}} p^{e(K \cup K')}.$$

Now, for some positive  $\lambda = \lambda(H)$ ,  $C' = C'(H)$ , and  $C = C(H, \varepsilon)$

$$\begin{aligned} 2\Delta_p(\partial_e \mathcal{H}) &= \sum_{K \in \partial_e \mathcal{H}} \sum_{\substack{K' \in \partial_e \mathcal{H} \setminus \{K\} \\ K \cap K' \neq \emptyset}} p^{e(K \cup K')} = \sum_{\substack{J \subseteq H \\ 2 \leq e_J < e_H}} \sum_{K \in \partial_e \mathcal{H}} \sum_{\substack{K' \in \partial_e \mathcal{H} \setminus \{K\} \\ (K \cap K') \cup e \cong J}} p^{2e_H - e_J - 1} \\ &\leq C' \sum_{\substack{J \subseteq H \\ 2 \leq e_J < e_H}} n^{2v_H - v_J - 2} p^{2e_H - e_J - 1} = C' \sum_{\substack{J \subseteq H \\ 2 \leq e_J < e_H}} \frac{n^2 p}{n^{v_J} p^{e_J}} \cdot (n^{v_H-2} p^{e_H-1})^2 \\ &\leq C n^{-\lambda} \cdot (n^{v_H-2} p^{e_H-1})^2, \end{aligned}$$

where the last inequality follows from Lemma 4.4; this estimate clearly remains true when we replace  $e$  with  $f$ . Similarly, for some  $C' = C'(H)$  and  $C = C(H, \varepsilon)$ ,

$$\begin{aligned} \Delta_p(\partial_e \mathcal{H}, \partial_f \mathcal{H}) &= \sum_{\substack{J \subseteq H \\ 1 \leq e_J < e_H - 1}} \sum_{K \in \partial_e \mathcal{H}} \sum_{\substack{K' \in \partial_f \mathcal{H} \setminus \{K\} \\ K \cap K' \cong J}} p^{2e_H - e_J - 2} \\ &\leq C' \sum_{\substack{J \subseteq H \\ 1 \leq e_J < e_H}} \frac{n^{|e \cap f|}}{n^{v_J} p^{e_J}} \cdot (n^{v_H - 2} p^{e_H - 1})^2 \leq \frac{C}{np} \cdot (n^{v_H - 2} p^{e_H - 1})^2, \end{aligned}$$

where the last inequality again follows from Lemma 4.4. Since  $np \geq \varepsilon n^{1-1/m_2(H)}$  and  $m_2(H) > 1$ , the claimed inequality follows.  $\square$

**Lemma 4.5.** *Suppose that  $\alpha, p \in (0, 1/2)$  and a nonnegative integer  $d$  satisfy*

$$\alpha \ll 1 \quad \text{and} \quad pn \gg \max \{ \alpha^{-4}, (d/\alpha)^2 \}.$$

*Then, for all fixed  $k \geq 0$  and  $r \geq 2$ , the random graph  $\mathbf{G} \sim G_{n,p}$  satisfies*

$$\mathbb{E} \left[ |\text{int}(\text{core}_d^r(\mathbf{G})) \cap \mathbf{G}|^k \cdot \mathbb{1}_{\mathbf{G} \in \text{CORE}_d^r(\alpha)} \right] = (1 + o(1)) \cdot \left( \frac{n^2 p}{2r} \right)^k.$$

*Proof.* Set  $m := \binom{n}{2} p$  and note that our assumption that  $np \gg 1$  guarantees that whp  $|\mathbf{G}| \in [m/2, 3m/2]$ . In particular, the assumed asymptotic relations between  $\alpha$ ,  $p$ , and  $d$  allow us to conclude from Corollary 2.2 that whp  $\mathbf{G} \in \text{CORE}_d^r(\alpha)$ . Since for every  $G \in \text{CORE}_d^r(\alpha)$ , the graph  $\text{core}_d^r(G)$  is a disjoint union of  $r$  complete graphs of order at least  $n/r - \alpha n$  each, we have

$$|\text{int}(\text{core}_d^r(G))| = \frac{n^2}{2r} \pm O(\alpha n^2)$$

and further, by the Chernoff bound and the union bound over the at most  $(r+1)^n$  possible graphs  $\text{core}_d^r(G)$ ,

$$\mathbb{P} \left( \forall G \in \text{CORE}_d^r(\alpha) \quad \left| |\text{int}(\text{core}_d^r(G)) \cap \mathbf{G}| - \frac{n^2 p}{2r} \right| = O \left( \alpha + (np)^{-1/2} \right) \cdot n^2 p \right) \geq 1 - e^{-n}.$$

The assertion of the lemma follows, as  $\alpha + (np)^{-1/2} \ll 1$  by our assumptions.  $\square$

**4.2. Proof of the lower bound on  $\mathbb{E}[Z]$ .** The following lemma, which is the main technical step in the proof of the lower bound on  $\mathbb{E}[Z]$ , abstracts the essence of the ‘resampling’ procedure that we described in the proof outline presented above. Given a hypergraph  $\mathcal{G}$ , we denote by  $\mathcal{I}(\mathcal{G})$  the family of its independent sets.

**Lemma 4.6.** *Suppose that  $\mathcal{G}$  is a  $k$ -uniform hypergraph on  $V$ , let  $t$  be a nonnegative integer, and define*

$$\mathcal{M} := \{R \subseteq V : \mathcal{G}[R] \text{ is a matching with } \leq t \text{ edges}\}.$$

Suppose further that  $p \in (0, 1)$ , let  $\mathbf{R} \sim V_p$ , and let  $\mathbf{R}^* := \mathbf{R} \setminus \bigcup \mathcal{G}[\mathbf{R}]$ . Then, for all  $\mathcal{A} \subseteq \mathcal{I}(\mathcal{G})$  and  $A: \mathcal{A} \rightarrow \mathcal{P}(V)$ , letting

$$\frac{q}{1-q} := \left( \frac{p}{1-p} \right)^k \quad \text{and} \quad a := \max \{ |\mathcal{G}[A(R^*)]| : R^* \in \mathcal{A} \},$$

we have

$$\mathbb{P}(\mathbf{R} \in \mathcal{M} \wedge \mathbf{R}^* \in \mathcal{A} \wedge \mathcal{G}[A(\mathbf{R}^*) \cap \mathbf{R}] = \emptyset) \geq (1-q)^a \cdot \mathbb{P}(\mathbf{R} \in \mathcal{M} \wedge \mathbf{R}^* \in \mathcal{A}).$$

*Proof.* Define the map  $\mathcal{S}: \mathcal{A} \rightarrow \mathcal{P}(\mathcal{P}(V))$  by

$$\mathcal{S}(R^*) := \{ S \subseteq V \setminus R^* : \mathcal{G}[S] \text{ is a matching spanning } S \text{ and } |\mathcal{G}[R^* \cup S]| = |\mathcal{G}[S]| \leq t \}.$$

Since, for every  $R^* \in \mathcal{A}$ , the family  $\mathcal{S}(R^*)$  comprises precisely those sets  $S \subseteq V \setminus R^*$  that are unions of edges of  $\mathcal{G}$  and satisfy  $R^* \cup S \in \mathcal{M}$ , we have

$$P_{R^*}(S) := \frac{\mathbb{P}(\mathbf{R} = R^* \cup S \wedge \mathbf{R}^* = R^*)}{\mathbb{P}(\mathbf{R} \in \mathcal{M} \wedge \mathbf{R}^* = R^*)} = \frac{1}{Z_{R^*}} \cdot \left( \frac{p}{1-p} \right)^{|S|} = \frac{1}{Z_{R^*}} \cdot \left( \frac{p}{1-p} \right)^{k \cdot |\mathcal{G}[S]|}$$

for all  $S \in \mathcal{S}(R^*)$ , where

$$Z_{R^*} := \sum_{S \in \mathcal{S}(R^*)} \left( \frac{p}{1-p} \right)^{|S|} = \sum_{S \in \mathcal{S}(R^*)} \left( \frac{q}{1-q} \right)^{|\mathcal{G}[S]|} = \sum_{\substack{M \subseteq \mathcal{G} \\ \bigcup M \in \mathcal{S}(R^*)}} \left( \frac{q}{1-q} \right)^{|M|}.$$

In other words, letting  $\mathbf{M}$  denote the  $q$ -random subset of  $\mathcal{G}$ , we have

$$P_{R^*}(S) = \mathbb{P}(\mathbf{M} = \mathcal{G}[S] \mid \bigcup \mathbf{M} \in \mathcal{S}(R^*)).$$

Since both  $\{M \subseteq \mathcal{G} : \bigcup M \in \mathcal{S}(R^*)\}$  and  $\{M \subseteq \mathcal{G} : M[A(R^*)] = \emptyset\}$  are decreasing families of subsets of  $\mathcal{G}$ , Harris's inequality [12] gives

$$\begin{aligned} \frac{\mathbb{P}(\mathbf{R} \in \mathcal{M} \wedge \mathbf{R}^* = R^* \wedge \mathcal{G}[A(R^*) \cap \mathbf{R}] = \emptyset)}{\mathbb{P}(\mathbf{R} \in \mathcal{M} \wedge \mathbf{R}^* = R^*)} &= \sum_{\substack{S \in \mathcal{S}(R^*) \\ \mathcal{G}[A(R^*) \cap S] = \emptyset}} P_{R^*}(S) \\ &= \mathbb{P}(\mathbf{M}[A(R^*)] = \emptyset \mid \bigcup \mathbf{M} \in \mathcal{S}(R^*)) \geq \mathbb{P}(\mathbf{M}[A(R^*)] = \emptyset) \\ &= (1-q)^{|\mathcal{G}[A(R^*)]|} \geq (1-q)^a. \end{aligned}$$

Multiplying the above inequality through by  $\mathbb{P}(\mathbf{R} \in \mathcal{M} \wedge \mathbf{R}^* = R^*)$  and summing the result over all  $R^* \in \mathcal{A}$  gives the desired inequality.  $\square$

**Corollary 4.7.** Suppose that  $p \leq Cn^{-1/m_2(H)}(\log n)^{1/(e_H-1)}$  for some constant  $C$  and let  $\mathbf{G} \sim G_{n,p}$ . For every  $e \in K_n$  and all  $\alpha > 0$ , letting  $d := \lceil \log n \rceil^2$ , we have

$$\begin{aligned} \mathbb{P}(\mathcal{Y}_e) &\geq \exp(-(\pi_H + O(\alpha + p)) \cdot (n/r)^{v_H-2} p^{e_H-1}) \\ &\quad \cdot (\mathbb{P}(\mathbf{G} \in \text{CORE}_{2d}^r(\alpha) \wedge e \in \text{int}(\text{core}_{2d}^r(\mathbf{G})) \cap \mathbf{G}) - o(p)). \end{aligned}$$

*Proof.* Let  $\omega := \mathbb{N} \rightarrow \mathbb{R}$  be an arbitrary function satisfying  $1 \ll \omega(n) \ll \log n$  and let

$$t := \omega(n) \cdot \mathbb{E}|\partial_e \mathcal{H}[\mathbf{G}]|.$$

It follows from our upper-bound assumption on  $p$  that, for some constants  $C_1 = C_1(H)$  and  $C_2 = C_2(H, C)$ , we have

$$t \leq \omega(n) \cdot C_1 n^{v_H-2} p^{e_H-1} \leq C_2 \omega(n) \log n \leq d/e_H, \quad (22)$$

provided that  $n$  is sufficiently large. As in Lemma 4.6, let

$$\mathcal{M} := \{G \subseteq K_n : \partial_e \mathcal{H}[G] \text{ is a matching with } \leq t \text{ edges}\}.$$

Further, let  $\mathbf{G}^* := \mathbf{G} \setminus \bigcup \partial_e \mathcal{H}[\mathbf{G}]$  and define

$$\mathcal{A} := \{G^* \subseteq K_n : G^* \in \text{CORE}_d^r(\alpha) \wedge e \in \text{int}(\text{core}_d^r(G^*)) \cap G^* \wedge \partial_e \mathcal{H}[G^*] = \emptyset\}$$

and the function  $A: \mathcal{A} \rightarrow \mathcal{P}(K_n)$  by  $A(G^*) := \text{ext}^*(\text{core}_d^r(G^*))$  for every  $G^* \in \mathcal{A}$ . By Lemma 4.2, for some constant  $C_H$  that depends only on  $H$ ,

$$\max_{G^* \in \mathcal{A}} |\partial_e \mathcal{H}[A(G^*)]| \leq (\pi_H + C_H \alpha) \cdot (n/r)^{v_H-2} =: a.$$

Further, since  $1 - x \geq \exp(-x/(1-x))$  for all  $x \in (0, 1)$ , Lemma 4.6 applied to the  $(e_H - 1)$ -uniform hypergraph  $\partial_e \mathcal{H}$  with vertex set  $K_n$  yields

$$\begin{aligned} \mathbb{P}(\mathbf{G} \in \mathcal{M} \wedge \mathbf{G}^* \in \mathcal{A} \wedge \partial_e \mathcal{H}[A(\mathbf{G}^*) \cap \mathbf{G}] = \emptyset) \\ \geq \exp\left(-\left(\frac{p}{1-p}\right)^{e_H-1} \cdot a\right) \cdot \mathbb{P}(\mathbf{G} \in \mathcal{M} \wedge \mathbf{G}^* \in \mathcal{A}). \end{aligned} \quad (23)$$

We now show that (23) implies the assertion of the corollary.

First, since  $(1-x)^{1-e_H} = 1 + (e_H - 1)x + O(x^2)$  as  $x \rightarrow 0$ , we have

$$\exp\left(-\left(\frac{p}{1-p}\right)^{e_H-1} \cdot a\right) \geq \exp\left(-(\pi_H + C_H \alpha) \cdot (1 + e_H p) \cdot (n/r)^{v_H-2} p^{e_H-1}\right).$$

Second, observe that  $\mathbf{G} \in \mathcal{M}$  implies that

$$|\mathbf{G}^*| = |\mathbf{G}| - (e_H - 1) \cdot |\partial_e \mathcal{H}[\mathbf{G}]| \geq |\mathbf{G}| - e_H t \geq |\mathbf{G}| - d, \quad (24)$$

see (22). Consequently, Corollary 2.7 implies that the event  $\mathbf{G} \in \text{CORE}_{2d}^r(\alpha) \cap \mathcal{M}$  is contained in the event  $\mathbf{G}^* \in \text{CORE}_d^r(\alpha)$  and  $\text{int}(\text{core}_{2d}^r(\mathbf{G})) \subseteq \text{int}(\text{core}_d^r(\mathbf{G}^*))$ . We thus have

$$\begin{aligned} \mathbb{P}(\mathbf{G} \in \mathcal{M} \wedge \mathbf{G}^* \in \mathcal{A}) &\geq \mathbb{P}(\mathbf{G} \in \text{CORE}_{2d}^r(\alpha) \cap \mathcal{M} \wedge e \in \text{int}(\text{core}_{2d}^r(\mathbf{G})) \cap \mathbf{G}) \\ &\geq \mathbb{P}(\mathbf{G} \in \text{CORE}_{2d}^r(\alpha) \wedge e \in \text{int}(\text{core}_{2d}^r(\mathbf{G})) \cap \mathbf{G}) - \mathbb{P}(\mathbf{G} \notin \mathcal{M} \wedge e \in \mathbf{G}). \end{aligned}$$

Since the events  $e \in \mathbf{G}$  and  $\mathbf{G} \notin \mathcal{M}$  are independent, we further have

$$\mathbb{P}(\mathbf{G} \notin \mathcal{M} \wedge e \in \mathbf{G}) = p \cdot \mathbb{P}(\mathbf{G} \notin \mathcal{M}) \leq p \cdot (\mathbb{P}(|\partial_e \mathcal{H}[\mathbf{G}]| > t) + \mathbb{P}(\Delta(\partial_e \mathcal{H}[\mathbf{G}]) \geq 2)).$$

The first probability in the right-hand side is at most  $1/\omega(n)$ , by Markov's inequality and the definition of  $t$ , whereas the second probability can be bounded using Lemma 4.3

as follows:

$$\begin{aligned} \mathbb{P}(\Delta(\partial_e \mathcal{H}[\mathbf{G}]) \geq 2) &\leq \mathbb{E}|\{(K, K') \in (\partial_e \mathcal{H}[\mathbf{G}])^2 : K \neq K', K \cap K' \neq \emptyset\}| \\ &= 2\Delta_p(\partial_e \mathcal{H}) \leq Cn^{-\lambda} \cdot (n^{v_H-2}p^{e_H-1})^2 \end{aligned}$$

for some positive  $C = C(H, \varepsilon)$  and  $\lambda = \lambda(H)$ ; since  $n^{v_H-2}p^{e_H-1} = O(\log n)$  under our upper-bound assumption on  $p$ , we may conclude that  $\mathbb{P}(\Delta(\partial_e \mathcal{H}[\mathbf{G}]) \geq 2) = o(1)$ .

Finally, since  $\mathbf{G} \in \mathcal{M}$  implies that  $|\mathbf{G}| \leq |\mathbf{G}^*| + d$ , see (24), Corollary 2.7 implies that the event  $\mathbf{G} \in \mathcal{M} \wedge \mathbf{G}^* \in \text{CORE}_d^r(\alpha)$  is contained in the event that  $\mathbf{G} \in \text{CORE}_0^r(\alpha)$  and  $\text{int}(\text{core}_d^r(\mathbf{G}^*)) \subseteq \text{int}(\text{core}_0^r(\mathbf{G}))$  (equivalently, that  $\text{ext}^*(\text{core}_0^r(\mathbf{G})) \subseteq \text{ext}^*(\text{core}_d^r(\mathbf{G}^*)) = A(\mathbf{G}^*)$ ). Further, since  $\mathbf{G}^* \in \mathcal{A}$  implies that  $e \in \text{int}(\text{core}_d^r(\mathbf{G}^*)) \cap \mathbf{G}^*$ , we conclude that

$$\mathbb{P}(\mathcal{Y}_e) \geq \mathbb{P}(\mathbf{G} \in \mathcal{M} \wedge \mathbf{G}^* \in \mathcal{A} \wedge \partial_e \mathcal{H}[A(\mathbf{G}^*) \cap \mathbf{G}] = \emptyset).$$

The assertion of the lemma follows by combining the above inequality with (23) and the lower bounds on the two terms in the right-hand side of (23).  $\square$

We are finally ready to complete the derivation of the lower bound on  $\mathbb{E}[Z]$ . Since  $n^{v_H-2}p^{e_H-1} = O(\log n)$ , we have

$$(\pi_H + O(\alpha + p)) \cdot (n/r)^{v_H-2}p^{e_H-1} = \pi_H \cdot (n/r)^{v_H-2}p^{e_H-1} + o(1).$$

Consequently, we may deduce from Corollary 4.7 that

$$\begin{aligned} \mathbb{E}[Z] &= \sum_{e \in K_n} \mathbb{P}(\mathcal{Y}_e) \geq \exp(-\pi_H \cdot (n/r)^{v_H-2}p^{e_H-1} - o(1)) \\ &\quad \cdot \left( \mathbb{E} \left[ |\text{int}(\text{core}_{2d}^r(\mathbf{G})) \cap \mathbf{G}| \cdot \mathbb{1}_{\mathbf{G} \in \text{CORE}_{2d}^r(\alpha)} \right] - o(n^2 p) \right), \end{aligned}$$

where  $d := \lceil \log n \rceil^2$ . Finally, Lemma 4.5 allows us to conclude that

$$\mathbb{E}[Z] \geq (1 + o(1)) \cdot \exp(-\pi_H \cdot (n/r)^{v_H-2}p^{e_H-1}) \cdot \frac{n^2 p}{2r}. \quad (25)$$

**4.3. Proof of the upper bound on  $\mathbb{E}[Z^2]$ .** Given distinct edges  $e, f \in K_n$  and a family  $\mathcal{C}$  of  $r$  pairwise-disjoint subsets of  $\llbracket n \rrbracket$ , define

$$\begin{aligned} I_{e,f}(\mathcal{C}) &:= \{G \subseteq K_n : e, f \in \text{int}(\mathcal{C}) \cap G\}, \\ E_{e,f}(\mathcal{C}) &:= \{G \subseteq K_n : (\partial_e \mathcal{H} \cup \partial_f \mathcal{H})[\text{ext}(\mathcal{C}) \cap G] = \emptyset\}. \end{aligned}$$

and note that, for every graph  $G \subseteq K_n$ ,

$$G \in \mathcal{Y}_e \cap \mathcal{Y}_f \implies G \in \text{CORE}_0^r(\alpha) \cap I_{e,f}(\text{core}_0^r(G)) \cap E_{e,f}(\text{core}_0^r(G)),$$

as  $\text{ext}(\mathcal{C}) \subseteq \text{ext}^*(\mathcal{C})$  for every family  $\mathcal{C}$ . We may thus conclude that

$$\mathbb{E}[Z^2] \leq \mathbb{E}[Z] + \sum_{\substack{e, f \in K_n \\ e \neq f}} \mathbb{P}(\mathbf{G} \in \text{CORE}_0^r(\alpha) \cap I_{e,f}(\text{core}_0^r(\mathbf{G})) \cap E_{e,f}(\text{core}_0^r(\mathbf{G}))). \quad (26)$$

Our next lemma, which is a variant of [6, Lemma 10.2], will allow us to bound from above the probabilities in the right-hand side of (26).

**Lemma 4.8.** *Let  $\alpha$  be a nonnegative real and let  $\mathfrak{C}$  be the collection of all  $r$ -element families  $\mathcal{C}$  of pairwise-disjoint subsets of  $\llbracket n \rrbracket$  satisfying  $|X| \geq n/r - \alpha n$  for all  $X \in \mathcal{C}$ . Suppose that, for each  $\mathcal{C} \in \mathfrak{C}$ , we have an event  $I(\mathcal{C})$  that is determined by  $\text{int}(\mathcal{C})$  and an event  $E(\mathcal{C})$  that is determined by  $\text{ext}(\mathcal{C})$  and decreasing, and satisfies  $\mathbb{P}(\mathbf{G} \in E(\mathcal{C})) \leq \xi$ . Then,*

$$\mathbb{P}(\mathbf{G} \in \text{CORE}_0^r(\alpha) \cap I(\text{core}_0^r(\mathbf{G})) \cap E(\text{core}_0^r(\mathbf{G}))) \leq \xi \cdot \mathbb{P}(\mathbf{G} \in \text{CORE}_0^r(\alpha) \cap I(\text{core}_0^r(\mathbf{G}))).$$

*Proof.* For any  $\mathcal{C} \in \mathfrak{C}$ , denote by  $\text{CORE}(\mathcal{C})$  the family of all  $G \subseteq K_n$  such that  $\text{core}_0^r(G) = \mathcal{C}$ . We first observe that  $\text{CORE}(\mathcal{C})$  is increasing in  $\text{ext}(\mathcal{C})$ . Indeed, since adding to  $G$  an edge of  $\text{ext}(\text{core}_0^r(G))$  does not change the set of largest  $r$ -cuts, the resulting graph has the same core as  $G$ . Let  $\mathcal{F}_{\mathcal{C}}$  be the  $\sigma$ -algebra generated by  $(K_n \setminus \text{ext}(\mathcal{C})) \cap \mathbf{G}$ . Since  $I(\mathcal{C})$  is determined by  $\text{int}(\mathcal{C}) \subseteq K_n \setminus \text{ext}(\mathcal{C})$ , we have

$$\mathbb{P}(\mathbf{G} \in \text{CORE}(\mathcal{C}) \cap I(\mathcal{C}) \cap E(\mathcal{C}) \mid \mathcal{F}_{\mathcal{C}}) = \mathbb{P}(\mathbf{G} \in \text{CORE}(\mathcal{C}) \cap E(\mathcal{C}) \mid \mathcal{F}_{\mathcal{C}}) \cdot \mathbb{1}_{\mathbf{G} \in I(\mathcal{C})}. \quad (27)$$

Further, since  $\text{CORE}(\mathcal{C})$  is increasing in  $\text{ext}(\mathcal{C})$  and  $E(\mathcal{C})$  is decreasing in  $\text{ext}(\mathcal{C})$ , Harris's inequality [12] implies that

$$\begin{aligned} \mathbb{P}(\mathbf{G} \in \text{CORE}(\mathcal{C}) \cap E(\mathcal{C}) \mid \mathcal{F}_{\mathcal{C}}) &\leq \mathbb{P}(\mathbf{G} \in \text{CORE}(\mathcal{C}) \mid \mathcal{F}_{\mathcal{C}}) \cdot \mathbb{P}(\mathbf{G} \in E(\mathcal{C}) \mid \mathcal{F}_{\mathcal{C}}) \\ &= \mathbb{P}(\mathbf{G} \in \text{CORE}(\mathcal{C}) \mid \mathcal{F}_{\mathcal{C}}) \cdot \mathbb{P}(\mathbf{G} \in E(\mathcal{C})), \end{aligned} \quad (28)$$

where the equality holds as  $E(\mathcal{C})$  is determined by  $\text{ext}(\mathcal{C})$  and thus the event  $\mathbf{G} \in E(\mathcal{C})$  is independent of  $\mathcal{F}_{\mathcal{C}}$ . Substituting (28) into (27), and using our assumption, yields

$$\begin{aligned} \mathbb{P}(\mathbf{G} \in \text{CORE}(\mathcal{C}) \cap I(\mathcal{C}) \cap E(\mathcal{C}) \mid \mathcal{F}_{\mathcal{C}}) &\leq \xi \cdot \mathbb{P}(\mathbf{G} \in \text{CORE}(\mathcal{C}) \mid \mathcal{F}_{\mathcal{C}}) \cdot \mathbb{1}_{\mathbf{G} \in I(\mathcal{C})} \\ &= \xi \cdot \mathbb{P}(\mathbf{G} \in \text{CORE}(\mathcal{C}) \cap I(\mathcal{C}) \mid \mathcal{F}_{\mathcal{C}}). \end{aligned}$$

Consequently, since  $\{\text{CORE}(\mathcal{C}) : \mathcal{C} \in \mathfrak{C}\}$  is a partition of  $\text{CORE}_0^r(\alpha)$ ,

$$\begin{aligned} \mathbb{P}(\mathbf{G} \in \text{CORE}_0^r(\alpha) \cap I(\text{core}_0^r(\mathbf{G})) \cap E(\text{core}_0^r(\mathbf{G}))) \\ = \sum_{\mathcal{C} \in \mathfrak{C}} \mathbb{P}(\mathbf{G} \in \text{CORE}(\mathcal{C}) \cap I(\mathcal{C}) \cap E(\mathcal{C})) &\leq \xi \cdot \sum_{\mathcal{C} \in \mathfrak{C}} \mathbb{P}(\mathbf{G} \in \text{CORE}(\mathcal{C}) \cap I(\mathcal{C})) \\ &= \xi \cdot \mathbb{P}(\mathbf{G} \in \text{CORE}_0^r(\alpha) \cap I(\text{core}_0^r(\mathbf{G}))), \end{aligned}$$

as desired.  $\square$

Returning to (26), since clearly  $I_{e,f}(\mathcal{C})$  is determined by  $\text{int}(\mathcal{C})$  whereas  $E_{e,f}(\mathcal{C})$  is determined by  $\text{ext}(\mathcal{C})$  and decreasing, Lemma 4.8 implies that

$$\begin{aligned} \mathbb{E}[Z^2] &\leq \mathbb{E}[Z] + \xi \cdot \sum_{\substack{e,f \in K_n \\ e \neq f}} \mathbb{P}(\mathbf{G} \in \text{CORE}_0^r(\alpha) \cap I_{e,f}(\text{core}_0^r(\mathbf{G}))) \\ &\leq \mathbb{E}[Z] + \xi \cdot \mathbb{E} \left[ |\text{int}(\text{core}_0^r(\mathbf{G})) \cap \mathbf{G}|^2 \cdot \mathbb{1}_{\mathbf{G} \in \text{CORE}_0^r(\alpha)} \right], \end{aligned} \quad (29)$$

where (writing  $\mathfrak{C}$  for the collection of all  $r$ -element families  $\mathcal{C}$  of pairwise-disjoint subsets of  $\llbracket n \rrbracket$  satisfying  $|X| \geq n/r - \alpha n$  for all  $X \in \mathcal{C}$ )

$$\xi := \max_{\mathcal{C} \in \mathfrak{C}} \mathbb{P}(\mathbf{G} \in E_{e,f}(\mathcal{C})) = \max_{\mathcal{C} \in \mathfrak{C}} \mathbb{P}((\partial_e \mathcal{H} \cup \partial_f \mathcal{H})[\text{ext}(\mathcal{C}) \cap \mathbf{G}] = \emptyset).$$

It follows from Lemma 4.2 that, for every  $\mathcal{C} \in \mathfrak{C}$ ,

$$\begin{aligned} \mu_p((\partial_e \mathcal{H} \cup \partial_f \mathcal{H})[\text{ext}(\mathcal{C})]) &= \mu_p(\partial_e \mathcal{H}[\text{ext}(\mathcal{C})]) + \mu_p(\partial_f \mathcal{H}[\text{ext}(\mathcal{C})]) \\ &\geq 2(\pi_H - O(\alpha + n^{-1})) \cdot (n/r)^{v_H-2} p^{e_H-1} \\ &\geq 2\pi_H \cdot (n/r)^{v_H-2} p^{e_H-1} - o(1), \end{aligned}$$

where we used that  $n^{v_H-2} p^{e_H-1} = O(\log n)$  whereas  $\alpha \ll 1/\log n$ . On the other hand, Lemma 4.3 gives that, for some  $\lambda = \lambda(H) > 0$  and  $C = C(H, \varepsilon)$ ,

$$\begin{aligned} \max_{\mathcal{C} \in \mathfrak{C}} \Delta_p((\partial_e \mathcal{H} \cup \partial_f \mathcal{H})[\text{ext}(\mathcal{C})]) &\leq \Delta_p(\partial_e \mathcal{H} \cup \partial_f \mathcal{H}) \\ &\leq Cn^{-\lambda} \cdot (n^{v_H-2} p^{e_H-1})^2 = o(1). \end{aligned}$$

Applying Janson's inequality (Theorem 4.1), we conclude that

$$\xi \leq \exp(-2\pi_H \cdot (n/r)^{v_H-2} p^{e_H-1} + o(1)).$$

Substituting this estimate into (29) and using Lemma 4.5, we obtain

$$\mathbb{E}[Z^2] \leq \mathbb{E}[Z] + (1 + o(1)) \cdot \left( \exp(-\pi_H \cdot (n/r)^{v_H-2} p^{e_H-1}) \cdot \frac{n^2 p}{2r} \right)^2.$$

Finally, recalling (25), in order to get the desired conclusion that  $\mathbb{E}[Z^2] \leq (1+o(1)) \cdot \mathbb{E}[Z]^2$ , it is enough to argue that

$$\exp(-\pi_H \cdot (n/r)^{v_H-2} p^{e_H-1}) \cdot \frac{n^2 p}{2r} \gg 1.$$

To see that this is the case, note that our upper-bound assumption on  $p$  and the assumption that  $H$  is 2-balanced, i.e.,  $m_2(H) = (e_H - 1)/(v_H - 2)$ , gives

$$\pi_H \cdot (n/r)^{v_H-2} p^{e_H-1} \leq \pi_H \cdot r^{2-v_H} \cdot (\theta_H - \varepsilon)^{e_H-1} \cdot \log n \leq \left( 2 - \frac{1}{m_2(H)} - \frac{\varepsilon}{\theta_H} \right) \cdot \log n,$$

whereas our lower-bound assumption on  $p$  is that  $n^2 p \geq \varepsilon n^{2-1/m_2(H)}$ .

## REFERENCES

- [1] L. Babai, M. Simonovits, and J. Spencer. Extremal subgraphs of random graphs. *J. Graph Theory*, 14(5):599–622, 1990.
- [2] G. Brightwell, K. Panagiotou, and A. Steger. Extremal subgraphs of random graphs. *Random Structures Algorithms*, 41(2):147–178, 2012.
- [3] A. Coja-Oghlan, C. Moore, and V. Sanwalani. MAX  $k$ -CUT and approximating the chromatic number of random graphs. *Random Structures Algorithms*, 28(3):289–322, 2006.
- [4] D. Conlon and W. T. Gowers. Combinatorial theorems in sparse random sets. *Ann. of Math. (2)*, 184(2):367–454, 2016.
- [5] D. Coppersmith, D. Gamarnik, M. T. Hajiaghayi, and G. B. Sorkin. Random MAX SAT, random MAX CUT, and their phase transitions. *Random Structures Algorithms*, 24(4):502–545, 2004.
- [6] B. DeMarco and J. Kahn. Turán's theorem for random graphs. arXiv:1501.01340.
- [7] B. DeMarco and J. Kahn. Mantel's theorem for random graphs. *Random Structures Algorithms*, 47(1):59–72, 2015.
- [8] A. Dembo, A. Montanari, and S. Sen. Extremal cuts of sparse random graphs. *Ann. Probab.*, 45(2):1190–1217, 2017.

- [9] O. Engelberg, W. Samotij, and L. Warnke. On the typical structure of graphs not containing a fixed vertex-critical subgraph. arXiv:2110.10931.
- [10] D. Gamarnik and Q. Li. On the max-cut of sparse random graphs. *Random Structures Algorithms*, 52(2):219–262, 2018.
- [11] L. Gishboliner, M. Krivelevich, and G. Kronenberg. On MAXCUT in strictly supercritical random graphs, and coloring of random graphs and random tournaments. *Random Structures Algorithms*, 52(4):545–559, 2018.
- [12] T. E. Harris. A lower bound for the critical probability in a certain percolation process. *Proc. Cambridge Philos. Soc.*, 56:13–20, 1960.
- [13] W. Hoeffding. Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.*, 58:13–30, 1963.
- [14] I. Hoshen and W. Samotij. Simonovits’s theorem in random graphs. arXiv:2308.13455.
- [15] S. Janson. Poisson approximation for large deviations. *Random Structures Algorithms*, 1(2):221–229, 1990.
- [16] Y. Kohayakawa, T. Łuczak, and V. Rödl. On  $K^4$ -free subgraphs of random graphs. *Combinatorica*, 17(2):173–213, 1997.
- [17] W. Samotij. Stability results for random discrete structures. *Random Structures Algorithms*, 44(3):269–289, 2014.
- [18] M. Schacht. Extremal results for random discrete structures. *Ann. of Math. (2)*, 184(2):333–365, 2016.
- [19] 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-London, 1968.
- [20] P. Turán. Eine Extremalaufgabe aus der Graphentheorie. *Mat. Fiz. Lapok*, 48:436–452, 1941.

SCHOOL OF MATHEMATICAL SCIENCES, TEL AVIV UNIVERSITY, TEL AVIV 6997801, ISRAEL  
*Email address:* ilayhoshen@gmail.com

SCHOOL OF MATHEMATICAL SCIENCES, TEL AVIV UNIVERSITY, TEL AVIV 6997801, ISRAEL  
*Email address:* samotij@tauex.tau.ac.il

DEPARTMENT OF COMPUTER SCIENCE, UNIVERSITY OF SHEFFIELD  
*Email address:* M.Zhukovskii@sheffield.ac.uk