General deletion lemmas via the Harris inequality

Reto Spöhel[1]

Technik und Informatik
Berner Fachhochschule
3400 Burgdorf, Switzerland
reto.spoehel@bfh.ch

Angelika Steger

Institute of Theoretical Computer Science
ETH Zürich
8092 Zurich, Switzerland
steger@inf.ethz.ch

Lutz Warnke[2]

Department of Pure Mathematics and Mathematical Statistics
University of Cambridge
Cambridge CB3 0WB, UK
L.Warnke@dpmms.cam.ac.uk

ABSTRACT. We present a new approach to the deletion method that is based on the Harris inequality. We obtain deletion lemmas that are similar in spirit to those of Rödl and Ruciński, but hold for arbitrary decreasing properties. That is, we show that under appropriate conditions, with very high ('Janson-like') probability it suffices to delete a small fraction of the edges of a random graph $G_{n,p}$ to ensure that a given decreasing property holds. We also obtain stronger results for decreasing properties that only depend on edges involved in copies of some given graph H.

As an application of our methods, we present a new deletion lemma that concerns *local* subgraph counts, i.e., the number of copies of H each edge or vertex of $G_{n,p}$ is contained in.

1 Introduction

1.1 Tail bounds for subgraph counts

The subgraph containment problem is a well-studied problem in random graph theory. It was shown in [2] that the threshold for the existence of a given fixed graph H in the binomial random graph $G_{n,p}$ is determined by the densest subgraph of H. A related line of research is the study of *subgraph counts*, i.e., of the distribution of the random variable X_H that counts the number of copies of H in $G_{n,p}$. Throughout, we denote by v_H and e_H the number of vertices and edges of H, and by

$$\mu_H = \mu_H(n, p) := \mathbb{E}[X_H] = \binom{n}{v_H} \frac{(v_H)!}{\mathrm{Aut}(H)} p^{e_H} \tag{1}$$

the expected number of copies of H in $G_{n,p}$. Note that $\mu_H = \Theta(n^{v_H} p^{e_H})$. We are mostly interested in settings where μ_H (and Φ_H defined below) is a growing function of n.

Janson's inequality [10, 16] gives very good upper bounds on the probability that X_H is significantly smaller than its expectation. Letting

$$\Phi_H = \Phi_H(n, p) := \min_{J \subseteq H:\, e_J \geq 1} \mu_J(n, p),$$

---

it yields that for any fixed $\varepsilon > 0$ there exists a constant $c = c(H, \varepsilon) > 0$ such that

$$\Pr[X_H \leq (1 - \varepsilon)\mathbb{E}[X_H]] \ \leq \ \mathrm{e}^{-c \cdot \Phi_H(n,p)}. \tag{2}$$

For the corresponding upper tail, however, such bounds are not easily obtained. Much research has been devoted to proving results of the form

$$\Pr[X_H \geq (1 + \varepsilon)\mathbb{E}[X_H]] \ \leq \ \mathrm{e}^{-f(n,p,H,\varepsilon)}$$

in our and similar combinatorial settings [4, 6, 5, 12, 13, 14]. In [12], a general exponent $f(n, p, H, \varepsilon)$ was given that is best possible up to logarithmic factors. As it turns out, in general this upper tail probability is simply not as small as the lower tail probability given by (2). Roughly speaking, this is due to the fact that a reasonably small number of edges that cluster in an appropriate way can give rise to many copies of H, see, e.g., [13] for an explicit example.

1.2 The deletion method

In order to better control the upper tail of X_H, Rödl and Ruciński [17] showed that, with 'Janson-like' probability, the number of copies of H can be reduced to at most $(1 + \varepsilon)$ times its expectation by deleting only a few edges.

Lemma 1 ([11, Lemma 2.51], 'Deletion Lemma'). *Let H be a graph and $0 \leq p \leq 1$. Then for every $\varepsilon > 0$ there exists $c = c(H, \varepsilon)$ such that for every integer k, with probability at least $1 - \mathrm{e}^{-ck}$, there exists a set $E_0 \subseteq E(G_{n,p})$ of size k such that $G_{n,p} \setminus E_0$ contains at most $(1 + \varepsilon)\mathbb{E}[X_H]$ copies of H.*

Typically, this lemma is applied in settings where $\Phi_H(n, p) = \Theta(n^2 p)$, with $k := \alpha n^2 p$ set to delete some small fraction of all edges. Note that then the resulting bound is similar to the lower tail bound given by (2).

Usually, this deletion lemma is used in conjunction with a second lemma that states that other relevant properties of the random graph (that hold with probability very close to 1) are *not* destroyed by the removal of a few edges. Such a lemma is often called 'robustness lemma'. The robustness lemma used in [17], see also [11, Lemma 2.52], works for arbitrary increasing graph properties (see Section 2 for the formal definition), but it is quite technical and only useful if k is of order $n^2 p$. A more generally applicable version can be found in [1], see also [8, Theorem 2.54].

Lemma 2 ([1, Lemma 4.2], 'Robustness Lemma'). *Given an increasing graph property $\mathcal{I}$, let $\mathcal{I}_{(k)} \subseteq \mathcal{I}$ denote the property that for every set $E_0 \subseteq E(G_{n,p})$ of size k the graph $G_{n,p} \setminus E_0$ satisfies $\mathcal{I}$. Then for any integer k and any real numbers $0 < p, \gamma < 1$ we have*

$$\Pr[G_{n,p} \notin \mathcal{I}_{(k)}] \ \leq \ \gamma^{-k} \Pr[G_{n,(1-\gamma)p} \notin \mathcal{I}].$$

In typical applications, the robustness lemma implies that if $\mathcal{I}$ holds with probability exponentially close in k to 1, then the same holds for $\mathcal{I}_{(k)}$, the 'robust version' of $\mathcal{I}$.

Combining these two lemmas with Janson's inequality (Theorem 13 below), one obtains in particular that if $\Phi_H(n, p) = \omega(1)$, then for any $\alpha > 0$, with probability as in (2) there exists a set E_0 of at most $\alpha \cdot \Phi_H(n, p)$ edges such that $G_{n,p} \setminus E_0$ contains between $(1 - \varepsilon)\mathbb{E}[X_H]$ and $(1 + \varepsilon)\mathbb{E}[X_H]$ many copies of H, i.e., roughly as many as the expectation predicts. (The requirement that $\Phi_H(n, p) = \omega(1)$ is equivalent to the requirement that $p = p(n)$ is above the threshold for the appearance of H in $G_{n,p}$.)

1.3 Our results

In this work, we present a new approach to the deletion method that is based on the Harris inequality [9]. We obtain deletion lemmas similar in spirit to Lemma 1 that hold for *arbitrary* decreasing graph properties (see Section 2 for the formal definition). Moreover, our approach provides a direct guarantee that desirable increasing events still hold after edge deletion. This removes the need for a separate robustness lemma, and thus the technicality of having to consider random graphs with slightly smaller edge probability (as Lemma 2 and the robustness lemma of [17] do). Furthermore, we obtain stronger results for decreasing properties that 'only depend on H-edges' (see Definition 6 below). As an application of our methods, we present a new deletion lemma that concerns *local* subgraph counts, i.e., the number of copies of H each edge or vertex of $G_{n,p}$ is contained in.

For the sake of presentation, we state our results in increasing order of sophistication – the results given later in this introduction will subsume some of the results stated earlier. We will use the following notation.

Definition 3. For any graph property $\mathcal{E}$ and any $k > 0$, we denote by $\mathcal{E}_k$ the property that there exists a set of at most $\lfloor k \rfloor$ edges such that after deleting these edges the remaining subgraph satisfies $\mathcal{E}$.

a) A general deletion lemma Our general deletion lemma can be stated as follows.

Theorem 4. *Let $\mathcal{D}$ be a decreasing graph property, and let $n^{-2} \ll p = p(n) \leq 1$ be such that there exists a constant $\delta > 0$ such that*

$$\Pr[G_{n,p} \in \mathcal{D}] \geq \delta$$

for n large enough. Then for any $\alpha > 0$ there exists $c = c(\alpha) > 0$ such that

$$\Pr[G_{n,p} \in \mathcal{D}_{\alpha n^2 p}] \geq 1 - \mathrm{e}^{-cn^2 p}$$

for n large enough.

Applying Theorem 4 with the decreasing property $\mathcal{D} := \{X_H \leq (1 + \varepsilon)\mathbb{E}[X_H]\}$ yields essentially the statement of the original deletion lemma (Lemma 1) with $k := \alpha n^2 p$. (The fact that $\mathcal{D}$ holds with constant probability follows from Markov's inequality.) Results similar to Theorem 4 were also proven by Bollobás and Leader [3, Section 3].

b) A direct guarantee Our next theorem is a version of Theorem 4 that additionally provides a guarantee that desirable increasing properties 'survive' the deletion of edges, essentially at no extra cost. Note that if $\mathcal{E} = \mathcal{D} \cap \mathcal{I}$ for some decreasing property $\mathcal{D}$ and some increasing property $\mathcal{I}$ in Definition 3, then $\mathcal{E}_k \subseteq \mathcal{I}$ is the event that we can delete a set of at most k edges to ensure $\mathcal{D}$ *without destroying $\mathcal{I}$.*

Theorem 5. *Let $\mathcal{D}$ be a decreasing graph property, let $\mathcal{I}$ be an increasing graph property, and let $n^{-2} \ll p = p(n) \leq 1$ be such that there exists constants $\delta, c' > 0$ such that*

$$\Pr[G_{n,p} \in \mathcal{D}] \geq \delta$$

and

$$\Pr[G_{n,p} \in \mathcal{I}] \geq 1 - \mathrm{e}^{-c'n^2 p} \tag{3}$$

for n large enough. Then for any $\alpha > 0$ there exists $c = c(\alpha, c') > 0$ such that

$$\Pr[G_{n,p} \in (\mathcal{D} \cap \mathcal{I})_{\alpha n^2 p}] \geq 1 - e^{-cn^2 p}$$

for n large enough.

Theorem 5 provides an alternative to the combination of Lemmas 1 and 2 in typical applications.

c) Properties that only depend on H-edges   If $\mathcal{D}$ only depends on edges that take part in copies of some specific graph H, then the event $\mathcal{D}_k$ is only of interest for values of k that are significantly smaller than $\mu_H = \mathbb{E}[X_H]$. Our next result is a refinement of Theorem 4 that provides better bounds for such a scenario.

Definition 6. Let H be a graph. We say that a graph property $\mathcal{E}$ *only depends on H-edges* if a graph G on vertex set V is in $\mathcal{E}$ if and only if the graph on V obtained as the union of all copies of H in G is in $\mathcal{E}$.

Theorem 7. *Let H be a graph, and let $\mathcal{D}$ be a decreasing graph property that only depends on H-edges. Let $n^{-v_H/e_H} \ll p = p(n) \leq 1$ be such that*

$$\Phi_H(n, p) = \mu_H(n, p) \tag{4}$$

for n large enough, and such that there exists a constant $\delta > 0$ such that

$$\Pr[G_{n,p} \in \mathcal{D}] \geq \delta$$

for n large enough. Then for any $\alpha > 0$ there exists $c = c(H, \alpha) > 0$ such that

$$\Pr[G_{n,p} \in \mathcal{D}_{\alpha \mu_H}] \geq 1 - e^{-c\mu_H}$$

for n large enough.

Note that the assumption that $p \gg n^{-v_H/e_H}$ is equivalent to $\mu_H(n, p) = \omega(1)$.

A version similar to Theorem 5 (which also provides a guarantee for an arbitrary increasing property $\mathcal{I}$ that only depends on H-edges and holds with probability at least $1 - e^{-c'\mu_H}$ for some $c' > 0$) can also be proven. We will come back to this in Remark 17 below; we state the simpler version here because this is how we will use Theorem 7 in the following.

Also note that every graph property 'only depends on K_2-edges', and that the condition (4) is trivially satisfied for $H = K_2$. Thus Theorem 7 contains Theorem 4 as a special case.

d) A bootstrapping trick   So far, our theorems are only applicable if the desired property $\mathcal{D}$ holds with constant probability. As the next corollary shows, we may even allow ourselves to delete edges to reach such a constant probability lower bound:

Corollary 8. *Let H be a graph, and let $\mathcal{D}$ be a decreasing graph property that only depends on H-edges. Let $n^{-v_H/e_H} \ll p = p(n) \leq 1$ be such that $\Phi_H(n, p) = \mu_H(n, p)$ for n large enough, and such that there exist constants $\alpha, \delta > 0$ such that*

$$\Pr[G_{n,p} \in \mathcal{D}_{\alpha \mu_H}] \geq \delta$$

for n large enough. Then there exists $c = c(H, \alpha) > 0$ such that

$$\Pr[G_{n,p} \in \mathcal{D}_{2\alpha \mu_H}] \geq 1 - e^{-c\mu_H}$$

for n large enough.

Proof. Apply Theorem 7 with $\mathcal{D} := \mathcal{D}_{\alpha\mu_H}$, and observe that $\mathcal{D}_{2\alpha\mu_H} \supseteq (\mathcal{D}_{\alpha\mu_H})_{\alpha\mu_H}$. $\qquad\square$

Corollary 8 provides a bootstrapping tool: In order to prove that a deletion statement of the desired kind holds with probability exponentially close to 1, it suffices to prove that the statement holds with some positive probability, and then slightly adjust the number of edges we allow to delete.

e) Strictly 2-balanced graphs A connected graph H with $e_H \geq 1$ and $v_H \geq 3$ is said to be *strictly 2-balanced* if $(e_J - 1)/(v_J - 2) < (e_H - 1)/(v_H - 2)$ for all proper subgraphs $J \subsetneq H$ with $v_J \geq 3$. We also define K_2 to be strictly 2-balanced. Complete graphs, complete bipartite graphs, and cycles of arbitrary size are strictly 2-balanced.

It is well-known (see, e.g., [11, Remark 3.17]) that if H is strictly 2-balanced, then there exists $n_0 = n_0(H)$ such that

$$\Phi_H(n, p) = \min\left\{\mu_H(n, p), \binom{n}{2}p\right\} \tag{5}$$

for all $0 < p \leq 1$ and $n \geq n_0$. Thus in that case, Theorems 4 and 7 cover all $p = p(n)$ for which $\mu_H = \omega(1)$, and we obtain:

Corollary 9. *Let H be a strictly 2-balanced graph, and let $\mathcal{D}$ be a decreasing graph property that only depends on H-edges. Let $n^{-v_H/e_H} \ll p = p(n) \leq 1$ be such that there exists a constant $\delta > 0$ such that*

$$\Pr[G_{n,p} \in \mathcal{D}] \geq \delta$$

for n large enough. Then for any $\alpha > 0$ there exists $c = c(H, \alpha) > 0$ such that

$$\Pr[G_{n,p} \in \mathcal{D}_{\alpha\Phi_H}] \geq 1 - \mathrm{e}^{-c\Phi_H}$$

for n large enough.

Note that the exponent in Corollary 9 matches the one in (2).

f) Application: Local subgraph counts As an application of our methods, we prove a strengthening of Lemma 1 which concerns not only global but also local subgraph counts.

Let $V = [n]$ denote the vertex set of $G_{n,p}$. For a vertex $v \in V$, let the random variable $X_{H,v}$ denote the number of copies of H in $G_{n,p}$ that contain v. Similarly, for two vertices $v_1, v_2 \in V$ and $e = \{v_1, v_2\}$, let $X_{H,e}$ denote the number of copies of H in the graph $G_{n,p} \cup \{v_1, v_2\}$ (i.e., in the graph obtained by inserting the edge $\{v_1, v_2\}$ into $G_{n,p}$ if it is not already present) that contain the edge $\{v_1, v_2\}$. It is easy to see that the expectations of these two random variables satisfy

$$\begin{aligned}
\mu_v = \mu_v(H, n, p) &:= \mathbb{E}[X_{H,v}] = \Theta(n^{v_H-1}p^{e_H}), \\
\mu_e = \mu_e(H, n, p) &:= \mathbb{E}[X_{H,e}] = \Theta(n^{v_H-2}p^{e_H-1}),
\end{aligned} \tag{6}$$

with leading constants that depend on H only. With these notations, our result reads as follows.

Theorem 10. *Let H be a strictly 2-balanced graph. Then for any $\varepsilon > 0$ there exist constants $c = c(H, \varepsilon) > 0$ and $C = C(H, \varepsilon) > 0$ such that the following holds: for $n^{-v_H/e_H} \ll p = p(n) \leq 1$ arbitrary and n large enough, with probability at least $1 - \mathrm{e}^{-c\Phi_H}$ there exists a set $E_0 \subseteq E(G_{n,p})$ of size at most $\varepsilon\Phi_H$ edges such that in $G_{n,p} \setminus E_0$, every vertex is contained in at most $\max\{C, (1+\varepsilon)\mu_v\}$ copies of H, and every edge is contained in at most $\max\{C, (1+\varepsilon)\mu_e\}$ copies of H.*

The constant C is necessary here: The analogous statement with the stronger bound $(1 + \varepsilon)\mu_v$ instead of $\max\{C, (1 + \varepsilon)\mu_v\}$ would clearly be wrong in the regime where $\mu_v = o(1)$ but $\mu_H = \omega(1)$.

Combining Theorem 10 with Lemma 2 and Janson's inequality, we obtain similar to our discussion at the end of Section 1.2 that, for H and p as in the theorem, with probability $1 - \mathrm{e}^{-\Theta(\Phi_H(n,p))}$ there exists a set E_0 of at most $\alpha\Phi_H(n, p)$ edges such that in $G_{n,p} \setminus E_0$, the global number of copies of H is close to its expectation, and all local subgraph counts are bounded as in Theorem 10.

1.4 Organization of the paper

After introducing some definitions and tools in Section 2, we prove Theorems 5 and 7 in Section 3. Finally, Section 4 is devoted to the proof of Theorem 10.

2 Preliminaries

Throughout this paper, we assume, as usual, that the random graph $G_{n,p}$ is generated on the vertex set $[n] = \{1, \ldots, n\}$. For the purposes of this paper, a graph property is a family of *labelled* graphs on the vertex set $[n]$ (which is not necessarily closed under isomorphism), where n will be clear from the context. We say that a graph property $\mathcal{A}$ is *decreasing* if for any two graphs G and H on vertex set $[n]$ the following holds: if $G \in \mathcal{A}$ and $H \subseteq G$, we also have $H \in \mathcal{A}$. Similarly, we say that a graph property $\mathcal{A}$ is *increasing* if the complement of $\mathcal{A}$ is decreasing.

In our proofs, we will make use of the following well-known tools.

Theorem 11 (Harris' inequality [9]). *For any two decreasing (increasing) graph properties $\mathcal{A}$ and $\mathcal{B}$ and any $n \in \mathbb{N}$ and $0 \le p \le 1$, we have*

$$\Pr[G_{n,p} \in \mathcal{A} \cap \mathcal{B}] \ge \Pr[G_{n,p} \in \mathcal{A}] \Pr[G_{n,p} \in \mathcal{B}].$$

Lemma 12 (Chernoff bounds [11]). *Let X be a binomially distributed random variable with parameters N and p. Then for any $0 \le k \le Np$ we have*

$$\Pr[X \ge Np + k] \le \mathrm{e}^{-k^2/(3Np)},$$

$$\Pr[X \le Np - k] \le \mathrm{e}^{-k^2/(2Np)}.$$

Theorem 13 (Janson's inequality [10]). *Consider a family $\mathcal{H} = \{H_i \mid i \in I\}$ of subgraphs of the complete graph on the vertex set $[n]$. For each $H_i \in \mathcal{H}$, let X_i denote the indicator random variable for the event $H_i \subseteq G_{n,p}$, and, for each ordered pair $(H_i, H_j) \in \mathcal{H} \times \mathcal{H}$ with $i \ne j$, write $H_i \sim H_j$ if H_i and H_j are not edge-disjoint. Let*

$$
\begin{aligned}
X &:= \sum_{H_i \in \mathcal{H}} X_i, \\
\mu &:= \mathbb{E}[X] = \sum_{H_i \in \mathcal{H}} p^{e(H_i)}, \\
\Delta &:= \sum_{\substack{(H_i, H_j) \in \mathcal{H} \times \mathcal{H} \\ H_i \sim H_j}} \mathbb{E}[X_i X_j] = \sum_{H_i \in \mathcal{H}} p^{e(H_i)} \sum_{\substack{H_j \in \mathcal{H} \\ H_i \sim H_j}} p^{e(H_j) - e(H_i \cap H_j)}.
\end{aligned}
$$

Then for all $0 \le \delta \le 1$ we have

$$\Pr[X \le (1 - \delta)\mu] \le \mathrm{e}^{-\frac{\delta^2 \mu^2}{2(\mu + \Delta)}}. \tag{7}$$

3 Deletion lemmas via the Harris inequality

To the best of our knowledge, the Harris inequality has not been used to prove deletion lemmas before. On a related note however, it was shown in [15] that certain upper tail difficulties, which are traditionally dealt with via the deletion method (as outlined in Section 1.2), can in some cases be avoided by applying the Harris inequality. The arguments in this section hinge on Lemma 14, which may be of independent interest.

3.1 Basic inequality

The following consequence of the Harris inequality is at the heart of our approach.

Lemma 14. *Let $n \in \mathbb{N}$ and $0 \leq p \leq 1$ be arbitrary. Let $\mathcal{I}$ denote an increasing graph property, and let $\mathcal{D}$ denote a decreasing graph property satisfying $\Pr[G_{n,p} \in \mathcal{D}] > 0$. Moreover, let $\mathcal{S} = \mathcal{S}(\mathcal{I}, \mathcal{D})$ denote the property of containing a* subgraph *that satisfies $\mathcal{I}$ and $\mathcal{D}$ simultaneously. Then we have*

$$\Pr[G_{n,p} \notin \mathcal{I}] \ \leq \ \Pr[G_{n,p} \notin \mathcal{S}] \ \leq \ \frac{\Pr[G_{n,p} \notin \mathcal{I}]}{\Pr[G_{n,p} \in \mathcal{D}]}.$$

Proof. The lower bound on $\Pr[G_{n,p} \notin \mathcal{S}]$ follows trivially from the fact that $\mathcal{S} \subseteq \mathcal{I}$. The definition of $\mathcal{S}$ also implies that $\mathcal{I} \cap \mathcal{D} \subseteq \mathcal{S}$ and thus $\neg \mathcal{S} \cap \mathcal{D} \subseteq \neg \mathcal{I}$.

By construction $\mathcal{S}$ is an increasing graph property, and thus $\neg \mathcal{S}$ is a decreasing graph property. Hence the Harris inequality (Theorem 11) implies that

$$\Pr[G_{n,p} \notin \mathcal{S}] \ \leq \ \frac{\Pr[G_{n,p} \in (\neg \mathcal{S} \ \cap \ \mathcal{D})]}{\Pr[G_{n,p} \in \mathcal{D}]} \ \leq \ \frac{\Pr[G_{n,p} \notin \mathcal{I}]}{\Pr[G_{n,p} \in \mathcal{D}]}.$$

$\square$

Note that whenever $\Pr[G_{n,p} \in \mathcal{D}]$ is bounded from below by a constant, Lemma 14 implies that the probability that $\mathcal{S}$ fails is within a constant factor of the probability that $\mathcal{I}$ fails. In many asymptotic scenarios, this means that it suffices to prove a bound for $\mathcal{I}$ to derive a good bound for $\mathcal{S}$.

3.2 Proof of Theorem 5

We infer Theorem 5 from Lemma 14 as follows.

Proof of Theorem 5. W.l.o.g. we assume that $\alpha \leq 0.01$. We will prove the claim for $c := 0.99 \cdot \min\{\alpha^2/6, c'\}$. Set $k := \alpha n^2 p$, and let $\mathcal{I}^*$ denote the property of satisfying $\mathcal{I}$ as given *and* containing at least $\binom{n}{2}p - k/2$ edges. We apply Lemma 14 with $\mathcal{D}$ as given and $\mathcal{I}^*$ as just defined. Together with the Chernoff bounds stated in Lemma 12, we obtain that, for n large enough, with probability at least

$$1 - \frac{\Pr[G_{n,p} \notin \mathcal{I}^*]}{\Pr[G_{n,p} \in \mathcal{D}]} \geq 1 - \delta^{-1}\left(e^{-(\alpha^2/4) \cdot n^2 p} + e^{-c' n^2 p}\right),$$

$G_{n,p}$ contains a subgraph that satisfies $\mathcal{D} \cap \mathcal{I}$ and has at least $\binom{n}{2}p - k/2$ edges. Lemma 12 also yields that with probability at least $1 - e^{-(\alpha^2/6) \cdot n^2 p}$, $G_{n,p}$ contains at most $\binom{n}{2}p + k/2$ edges. Hence with the claimed probability it suffices to remove at most k edges from $G_{n,p}$ to obtain a subgraph satisfying $\mathcal{D} \cap \mathcal{I}$. $\square$

Remark 15. A closer inspection of the preceding proof reveals that in fact, the conclusion of Theorem 5 still holds under the weaker assumption that, for n large enough, $\Pr[G_{n,p} \in \mathcal{D}] \geq e^{-c''n^2p}$ for some constant $c'' < \min\{\alpha^2/4, c'\}$. The constant c then needs to be adjusted to $c = c(\alpha, c', c'') := 0.99 \cdot \min\{\alpha^2/6, \min\{\alpha^2/4, c'\} - c''\}$.

3.3   Proof of Theorem 7

The proof of Theorem 7 uses Lemma 14 in a way similar to the previous proof, but is quite a bit more technical.

Proof of Theorem 7. We begin by defining all constants and events needed in the proof. Let H and α be given, and assume w.l.o.g. that $\alpha \leq 0.01$. Set

$$\gamma = \gamma(H, \alpha) := \frac{\alpha}{6(e_H)^2}, \tag{8}$$

$$c_\Delta = c_\Delta(H) := 2^{v_H + e_H}(v_H)^{2v_H}, \tag{9}$$

$$c_1 = c_1(H, \alpha) := \frac{\alpha^2(1-\gamma)^{e_H}}{144(1+c_\Delta)(e_H)^2}, \tag{10}$$

$$\varepsilon = \varepsilon(H, \alpha) := \min\{\alpha/3, \frac{c_1}{-\log\gamma}\}, \tag{11}$$

and let $c_2 = c_2(H, \alpha)$ denote the constant guaranteed by Theorem 4 for $\alpha := \varepsilon$.

We shall prove the theorem for

$$c = c(H, \alpha) := 0.99 \cdot \min\{c_1, c_2\}. \tag{12}$$

Let $p = p(n)$ as in the theorem be given, and consider the increasing properties

$$\mathcal{L} := \{G \mid G \text{ contains at least } (1 - \alpha/(3e_H))\mu_H \text{ copies of } H\}$$

(where G ranges over all graphs on n vertices) and

$$\mathcal{I} := \mathcal{L}_{(\varepsilon\mu_H)}$$

as defined in Lemma 2. Let $\mathcal{S} := \mathcal{S}(\mathcal{I}, \mathcal{D})$ be defined as in Lemma 14.

Furthermore, let

$$\mathcal{T} := \{G \mid \text{ there exists } X \subseteq E(G), |X| \leq \varepsilon\mu_H, \text{ such that}$$
$$G \setminus X \text{ contains at most } (1 + \alpha/(3e_H))\mu_H \text{ copies of } H\}.$$

We will prove Theorem 7 by showing that

$$\mathcal{S} \cap \mathcal{T} \subseteq \mathcal{D}_{\alpha\mu_H}, \tag{13}$$

and that

$$\Pr[G_{n,p} \notin \mathcal{S} \cap \mathcal{T}] \leq e^{-c\mu_H} \tag{14}$$

for n large enough.

To see that (13) holds, let $G \in \mathcal{S} \cap \mathcal{T}$ be given. By definition of $\mathcal{T}$, we can remove a set X of at most $\varepsilon\mu_H$ edges from G to obtain a graph $G' := G \setminus X$ that contains at most $(1 + \alpha/(3e_H))\mu_H$ copies of

8

H. Furthermore, by definition of $\mathcal{S}$, G contains a subgraph J that satisfies $\mathcal{I}$ and $\mathcal{D}$. Note that the graph $J' := J \setminus X \subseteq G'$ satisfies $\mathcal{D}$ (trivially, as $\mathcal{D}$ is decreasing) and $\mathcal{L}$ (by definition of $\mathcal{I} = \mathcal{L}_{(\varepsilon\mu_H)}$). Thus the number of copies of H in G' and J' differs by at most $2\alpha/(3e_H) \cdot \mu_H$, and we need to remove at most $2\alpha/3 \cdot \mu_H$ edges from G' to obtain a graph G'' whose set of H-edges is exactly the set of H-edges of J'. Since $\mathcal{D}$ only depends on H-edges, it follows from the fact that J' satisfies $\mathcal{D}$ that also G'' satisfies $\mathcal{D}$. Moreover, G'' was obtained by removing at most $\varepsilon\mu_H + 2\alpha/3 \cdot \mu_H \leq \alpha\mu_H$ edges from G (see (11)). This shows that G is in $\mathcal{D}_{\alpha\mu_H}$ and concludes the proof of (13).

To show that (14) holds, we first bound the probability of $\mathcal{I} = \mathcal{L}_{(\varepsilon\mu_H)}$ using Janson's inequality (Theorem 13) and the robustness lemma stated in the introduction (Lemma 2). For convenience of notation, we set $p' = (1-\gamma)p$, and abbreviate $\mu_H(n, p')$ as defined in (1) by μ'_H.

Note that

$$\mu'_H = (1-\gamma)^{e_H}\mu_H \overset{(8)}{\geq} (1 - \alpha/(6e_H))\mu_H.$$

Therefore, if $G_{n,p'}$ is not in $\mathcal{L}$ then the random variable that counts the number of copies of H in $G_{n,p'}$ is at least $\alpha/(6e_H) \cdot \mu_H \geq \alpha/(6e_H) \cdot \mu'_H$ below its expectation μ'_H. We use Janson's inequality (Theorem 13) to bound the probability that this happens. It is not difficult to see that for every $J \subseteq H$ with $e_J \geq 1$ we have $(p')^{e_H - e_J} \leq p^{e_H - e_J}$ and $n^{v_J}p^{e_J}/(v_J)^{v_J} \leq \mu_J \leq n^{v_J}p^{e_J}$. By standard arguments it follows that

$$\Delta \leq \mu'_H \sum_{J \subseteq H:\, e_J \geq 1} (v_H)^{v_J} n^{v_H - v_J} (p')^{e_H - e_J}$$

$$\leq \mu'_H \cdot 2^{v_H + e_H}(v_H)^{2v_H} \max_{J \subseteq H:\, e_J \geq 1} (\mu_H/\mu_J) \overset{(9)}{\leq} c_\Delta \mu'_H,$$

where in the last step we also used that $\mu_H \leq \mu_J$ by assumption (4).

Thus Janson's inequality (7) yields

$$\Pr[G_{n,p'} \notin \mathcal{L}] \leq \exp\left(-\frac{(\alpha/(6e_H))^2(\mu'_H)^2}{2(\mu'_H + \Delta)}\right) \overset{(10)}{=} e^{-2c_1\mu'_H/(1-\gamma)^{e_H}} = e^{-2c_1\mu_H},$$

and applying Lemma 2 to $\mathcal{I} = \mathcal{L}_{(\varepsilon\mu_H)}$ we obtain

$$\Pr[G_{n,p} \notin \mathcal{I}] \leq \gamma^{-\varepsilon\mu_H}\Pr[G_{n,p'} \notin \mathcal{L}] \overset{(11)}{\leq} e^{-c_1\mu_H}.$$

Furthermore, Lemma 14 implies with our assumption on $\mathcal{D}$ that

$$\Pr[G_{n,p} \notin \mathcal{S}] \leq \frac{\Pr[G_{n,p} \notin \mathcal{I}]}{\Pr[G_{n,p} \in \mathcal{D}]} \leq \delta^{-1}e^{-c_1\mu_H}.$$

Lastly, by Theorem 4 and our definition of c_2 we have

$$\Pr[G_{n,p} \notin \mathcal{T}] \leq e^{-c_2 \cdot n^2 p} \overset{(4)}{\leq} e^{-c_2\mu_H}.$$

Together with the definition of c in (12), and observing that $\mu_H = \omega(1)$, the last two inequalities imply that $\Pr[G_{n,p} \notin \mathcal{S} \cap \mathcal{T}]$ is bounded as claimed in (14). This concludes the proof of Theorem 7. $\square$

Remark 16. As can be seen from the preceding proof, the conclusion of Theorem 7 still holds under the weaker assumption that, for n large enough, $\Pr[G_{n,p} \in \mathcal{D}] \geq \mathrm{e}^{-c''\mu_H}$ for a small enough constant $c'' = c''(H, \alpha)$. We omit the details.

Remark 17. A version of Theorem 7 that also provides a guarantee for an increasing property $\mathcal{I}$ that only depends on H-edges and holds with probability at least $1 - \mathrm{e}^{-c'\mu_H}$ can be proved very similarly. It suffices to replace the event $\mathcal{S} = \mathcal{S}(\mathcal{L}_{(\varepsilon\mu_H)}, \mathcal{D})$ defined above with $\mathcal{S}((\mathcal{I} \cap \mathcal{L})_{(\varepsilon\mu_H)}, \mathcal{D})$ and adjust a few details.

4   Proof of Theorem 10

To simplify notation, we introduce the abbreviations

$$
\begin{aligned}
L_v &= L_v(H, n, p, C, \varepsilon) := \max\{C, (1+\varepsilon)\mu_v\}, \\
L_e &= L_e(H, n, p, C, \varepsilon) := \max\{C, (1+\varepsilon)\mu_e\}.
\end{aligned}
\tag{15}
$$

In view of Corollary 9 and the bootstrapping trick stated in Corollary 8, it suffices to prove the following statement.

Lemma 18. *Let H be a strictly 2-balanced graph. Then for any $\varepsilon > 0$ there exists $C = C(H, \varepsilon) > 0$ such that the following holds: for $n^{-v_H/e_H} \ll p = p(n) \leq 1$ arbitrary and n large enough, with probability at least $1/2$ there exists a set $E_0 \subseteq E(G_{n,p})$ of size at most $\varepsilon\Phi_H$ edges such that in $G_{n,p} \setminus E_0$, every vertex is contained in at most L_v copies of H, and every edge is contained in at most L_e copies of H.*

We assume w.l.o.g. that $\varepsilon \leq 0.01$. Clearly, to ensure that a given vertex v is contained in at most L_v copies of H, it suffices to delete one edge from $(X_{H,v} - L_v)$ copies of H (that contain v), if $X_{H,v} > L_v$. The expected number of edges we delete by this is thus at most

$$
\sum_{k > L_v} \Pr[X_{H,v} = k] \cdot (k - L_v) = \sum_{k > L_v} \Pr[X_{H,v} \geq k] =: z_v.
$$

Similarly, to ensure that a fixed vertex pair $e = \{v_1, v_2\}$ is contained in at most L_e extensions to copies of H, it suffices to delete at most $(X_{H,e} - L_e)$ edges from $G_{n,p}$ if $X_{H,e} > L_e$. Note that we only care about $X_{H,e}$ if e is an edge of $G_{n,p}$, and that $X_{H,e}$ does *not* depend on the status of e. Since e is present with probability p, the expected number of edges we delete is thus at most

$$
p \sum_{k > L_e} \Pr[X_{H,e} \geq k] =: pz_e.
$$

It follows that we can reach our goal by deleting an expected number of

$$
z = z(H, n, p, C, \varepsilon) := nz_v + \binom{n}{2}pz_e
$$

edges in total. By Markov's inequality, to infer Lemma 18 it suffices to show that $z \leq \varepsilon\Phi_H/2$ for an appropriate choice of the constant C.

To do so, it suffices to show that z_v is bounded by $\varepsilon\Phi_H/(4n)$, and that z_e is bounded by $\varepsilon\Phi_H/(4\binom{n}{2}p)$. Note that the last term is of order $\min\{\mu_e, 1\}$ by (5). Recalling that H is connected, and observing that trees T with $v_T \geq 3$ are not strictly 2-balanced, it follows that $H \neq K_2$ satisfies

$$
e_H \geq v_H.
\tag{16}
$$

Hence the term $\varepsilon \Phi_H/(4n)$ is of order $\min\{\mu_v, np\} = \Omega(\min\{\mu_v, 1\})$, where the asymptotic inequality comes from the fact that (16) implies $p \gg n^{-v_H/e_H} \geq n^{-1}$.

Before we summarize in the next claim what we are left with to prove, we make one additional observation: Using that H is strictly 2-balanced, it follows from the results of Spencer [18] that if $p = p(n)$ is such that $\mu_e = \omega(\log n)$ then for any $\varepsilon > 0$, with high probability (i.e., with probability tending to 1 as $n \to \infty$) every edge of $G_{n,p}$ is contained in at most $(1 + \varepsilon)\mu_e$ copies of H (and thus there is no need to delete any edges to fix the number of copies containing each edge). Hence when bounding z_e we may restrict our attention to the case where $\mu_e = O(\log^2 n)$, say. By an analogous argument we may assume that $\mu_v = O(\log^2 n)$ when bounding z_v; to apply Spencer's result here we use the well-known fact that every strictly 2-balanced graph is also strictly 1-balanced, i.e., satisfies $e_J/(v_J - 1) < e_H/(v_H - 1)$ for all proper subgraphs $J \subsetneq H$ with $v_J \geq 2$ (this inequality is trivial if $H = K_2$ or $v_J = v_H$, and otherwise equivalent to $\frac{e_H - e_J}{v_H - v_J} > \frac{e_H}{v_H - 1}$, which in turn holds by the proof of Lemma 20 below).

With these observations, and after some simplification of the constants, it remains to prove the following claim.

Claim 19. *For any strictly 2-balanced graph H and any $\varepsilon > 0$ there exists $C = C(H, \varepsilon)$ such that*

(i) *for any function $p = p(n)$ for which $\mu_v = O(\log^2 n)$ we have $\sum_{k > L_v} \Pr[X_{H,v} \geq k] \leq \varepsilon \min\{\mu_v, 1\}$ for n large enough,*

(ii) *for any function $p = p(n)$ for which $\mu_e = O(\log^2 n)$ we have $\sum_{k > L_e} \Pr[X_{H,e} \geq k] \leq \varepsilon \min\{\mu_e, 1\}$ for n large enough.*

The assumption of strict 2-balancedness comes into play via the following lemma.

Lemma 20. *For every strictly 2-balanced graph H, there exists $\gamma = \gamma(H)$ such that the following hold.*

(i) *If $p = p(n)$ is such that $\mu_v = O(\log^2 n)$, then for every proper subgraph $J \subsetneq H$ with $e_J \geq 1$ we have $n^{v_H - v_J} p^{e_H - e_J} \leq n^{-\gamma}$ for n large enough.*

(ii) *If $p = p(n)$ is such that $\mu_e = O(\log^2 n)$, then for every proper subgraph $J \subsetneq H$ with $e_J \geq 2$ we have $n^{v_H - v_J} p^{e_H - e_J} \leq n^{-\gamma}$ for n large enough.*

Proof. As there is nothing to prove for $H = K_2$, we may assume that $e_H \geq v_H \geq 3$, see (16). The assumption that H is strictly 2-balanced implies that every subgraph $J \subseteq H$ with $3 \leq v_J < v_H$ satisfies

$$\frac{e_H - e_J}{v_H - v_J} = \frac{(e_H - 1) - (e_J - 1)}{(v_H - 2) - (v_J - 2)} > \frac{e_H - 1}{v_H - 2} > \frac{e_H}{v_H - 1},$$

where in the last inequality we used that $e_H \geq v_H$. Noting that the inequality $\frac{e_H - e_J}{v_H - v_J} > \frac{e_H}{v_H - 1}$ also holds for $J = K_2$, we may define

$$\gamma_v = \gamma_v(H) := \frac{1}{2} \min_{J \subsetneq H : e_J \geq 1} \left(\frac{(e_H - e_J)(v_H - 1)}{e_H} - (v_H - v_J) \right) > 0,$$

$$\gamma_e = \gamma_e(H) := \frac{1}{2} \min_{J \subsetneq H : e_J \geq 2} \left(\frac{(e_H - e_J)(v_H - 2)}{e_H - 1} - (v_H - v_J) \right) > 0.$$

Assume that we have $\mu_v = O(\log^2 n)$. Then this implies that $p = O(n^{-(v_H-1)/e_H}(\log n)^{2/e_H})$, and thus for any proper subgraph $J \subsetneq H$ with $e_J \geq 1$ we have

$$n^{v_H-v_J}p^{e_H-e_J} = O\left(n^{(v_H-v_J)-\frac{(e_H-e_J)(v_H-1)}{e_H}}(\log n)^{2(e_H-e_J)/e_H}\right) = O(n^{-2\gamma_v}\log^2 n) \leq n^{-\gamma_v}$$

for n large enough. Similarly, if $\mu_e = O(\log^2 n)$ then $p = O(n^{-(v_H-2)/(e_H-1)}(\log n)^{2/(e_H-1)})$, and thus for any proper subgraph $J \subsetneq H$ with $e_J \geq 2$ we have

$$n^{v_H-v_J}p^{e_H-e_J} = O\left(n^{(v_H-v_J)-\frac{(e_H-e_J)(v_H-2)}{e_H-1}}(\log n)^{2(e_H-e_J)/(e_H-1)}\right) = O(n^{-2\gamma_e}\log^2 n) \leq n^{-\gamma_e}$$

for n large enough. The claim follows for $\gamma := \min\{\gamma_v, \gamma_e\}$. $\qquad\square$

To prove Claim 19, we apply the so-called 'approximation by a disjoint subfamily' technique [13], which can be traced back to Spencer [18]. For a graph $\mathcal{G}$ we denote by $\alpha(\mathcal{G})$ the independence number of $\mathcal{G}$, by $\beta(\mathcal{G})$ the size of a largest induced matching in $\mathcal{G}$, and by $\Delta(\mathcal{G})$ the maximum degree of $\mathcal{G}$. It follows from basic graph theory (see, e.g., Section 2.3.4. in [13]) that $V(\mathcal{G}) \leq \alpha(\mathcal{G}) + 2\beta(\mathcal{G})\Delta(\mathcal{G})$.

In order to prove part (i) of Claim 19 we define an auxiliary graph $\mathcal{G} = \mathcal{G}(G_{n,p})$ as follows. The vertex set of $\mathcal{G}$ is the set of all copies of H in $G_{n,p}$ that contain the vertex v, and two vertices of $\mathcal{G}$ are connected by an edge if and only if the two corresponding copies of H share at least one edge. Then

$$X_{H,v} = V(\mathcal{G}) \leq \alpha(\mathcal{G}) + 2\beta(\mathcal{G})\Delta(\mathcal{G}). \tag{17}$$

Intuitively, we will show that the dominant term on the right hand side of (17) is $\alpha(\mathcal{G})$, and that the probability for a contribution of the term $2\beta(\mathcal{G})\Delta(\mathcal{G})$ decreases exponentially. More precisely, we will show that there exist positive constants $A = A(H)$, $B = B(H,\varepsilon)$ and $\beta = \beta(H,\varepsilon)$ such that the following estimates hold for all n large enough and $k > L'_v$, where $L'_v = L_v(H, n, p, C', \varepsilon)$ with $C' = C'(H, \varepsilon)$ chosen appropriately:

$$\Pr[\Delta(\mathcal{G}) \geq A] \leq \varepsilon/2 \cdot n^{-v_H} \cdot \min\{\mu_v, 1\}, \tag{18}$$

$$\Pr[\alpha(\mathcal{G}) \geq (1 - 2\beta A)k] \leq e^{-Bk} \cdot \min\{\mu_v, 1\}, \tag{19}$$

$$\Pr[\beta(\mathcal{G}) \geq \beta k] \leq e^{-Bk} \cdot \min\{\mu_v, 1\}. \tag{20}$$

Note that together these estimates imply by (17) that for all such k we have

$$\Pr[X_{H,v} \geq k] \leq \left(2e^{-Bk} + \varepsilon/2 \cdot n^{-v_H}\right)\min\{\mu_v, 1\}.$$

For any $C \geq C'$ we have $L_v \geq L'_v$, and thus this bound holds for all $k > L_v$. As moreover $X_{H,v}$ cannot exceed n^{v_H}, summing up yields that

$$\sum_{k > L_v} \Pr[X_{H,v} \geq k] \leq \left(\frac{2e^{-BC}}{1 - e^{-B}} + \varepsilon/2\right)\min\{\mu_v, 1\}$$

for n large enough. For $C = C(H, \varepsilon)$ chosen suitably large, this is bounded by $\varepsilon\min\{\mu_v, 1\}$ as desired.

It remains to establish (18), (19), and (20). To do so, we introduce some additional notation. Let $\mathcal{S}$ be an arbitrary family of subsets of $E(K_n)$. For $\alpha \in \mathcal{S}$, let X_α denote the indicator random variable for the event that all edges of α are present in $G_{n,p}$, and set $X_\mathcal{S} = \sum_{\alpha \in \mathcal{S}} X_\alpha$. By $\mathcal{H}_p(\mathcal{S})$ we denote

the *random intersection graph* given by $\mathcal{S}$, which is defined as follows: The vertex set is the set of all $\alpha \in \mathcal{S}$ with $X_\alpha = 1$, and two vertices $\alpha, \beta \in \mathcal{S}$ are connected by an edge if and only if $\alpha \cap \beta \neq \emptyset$.

For $\mathcal{S} = \mathcal{S}_1$ being the family of all edge sets corresponding to copies of H containing v, this is exactly the auxiliary graph $\mathcal{G}$ defined above. In particular, we have $X_{\mathcal{S}_1} = X_{H,v}$ and $\alpha(\mathcal{G}) = \alpha(\mathcal{H}_p(\mathcal{S}_1))$. Moreover, for $\mathcal{S} = \mathcal{S}_2$ being the family of all edge sets obtained as the union of two intersecting copies of H, both containing v, this is exactly the line graph of the auxiliary graph $\mathcal{G}$ defined above, and thus we have $\beta(\mathcal{G}) = \alpha(\mathcal{H}_p(\mathcal{S}_2))$.

With these notations in hand, we prove (18). One mild difficulty arising here is that $\alpha_j \subseteq \bigcup_{i \in \mathcal{F}} \alpha_i$ with $j \notin \mathcal{F}$ is possible, i.e., that the union of several α_i can contain other α_j subsets. To overcome this, we first restrict our attention to the cases where this effect does not occur. We call a sequence $(\alpha_i)_{1 \leq i \leq r}$ of distinct sets $\alpha_i \in \mathcal{S}$ *nice* if for each $1 < j \leq r$, the set α_j is neither contained in $\bigcup_{1 \leq i < j} \alpha_i$ nor completely edge-disjoint from $\bigcup_{1 \leq i < j} \alpha_i$. Note that this means that this intersection is isomorphic to some nonempty proper subgraph $J \subsetneq H$.

Let $T = T(H) := \lceil 2v_H/\gamma \rceil + 2$, where $\gamma = \gamma(H)$ is the constant guaranteed by Lemma 20. We will derive a bound on the probability that a nice sequence of length T exists in $G_{n,p}$. Observe that any graph W spanned by the union of all α_i of such a sequence satisfies

$$e(W) = e_H + \sum_{i=2}^{T} (e_H - e(J_i)) \qquad \text{and} \qquad v(W) = v_H + \sum_{i=2}^{T} (v_H - v(J_i)), \tag{21}$$

where each J_i is a nonempty proper subgraph of H. Since $\mu_v \leq n^{v_H}$ implies $\mu_v n^{-v_H} \leq \min\{\mu_v, 1\}$, it follows from Markov's inequality and Lemma 20 that the probability for the existence of such a graph W containing v is at most

$$c_1 n^{v(W)-1} p^{e(W)} \leq c_1 n^{v_H-1} p^{e_H} (n^{-\gamma})^{T-1} \leq c_2 \mu_v n^{-2v_H-\gamma} \leq c_2 \min\{\mu_v, 1\} n^{-v_H-\gamma},$$

where c_1 and c_2 are constants depending only on H. Taking a union bound over all possible graphs W (there are only finitely many), we conclude that for n large enough, the probability that a nice sequence of length T exists is bounded by the right hand side of (18).

It remains to show that this implies the desired bound on $\Delta(\mathcal{G})$. Consider any *maximal* nice sequence $\alpha_1, \ldots, \alpha_r$. Then clearly, every $\alpha \in \mathcal{S} \setminus \{\alpha_1, \ldots, \alpha_r\}$ that intersects any one of these α_i is contained in the graph W spanned by $\alpha_1, \ldots, \alpha_r$. It follows that $\Delta(\mathcal{G}) \leq v(W)^{v_H} \leq (rv_H)^{v_H}$, which establishes (18) for $A = A(H) := (Tv_H)^{v_H}$.

The bounds for $\alpha(\mathcal{G}) = \alpha(\mathcal{H}_p(\mathcal{S}_1))$ and $\beta(\mathcal{G}) = \alpha(\mathcal{H}_p(\mathcal{S}_2))$ claimed in (19) and (20) follow from the following tail inequality, which combines estimates of Erdős and Tetali [7] and Janson [10] in a convenient way.

Lemma 21. *Let $\mathcal{S}$ be a family of subsets of $E(K_n)$, and set $\lambda = \mathbb{E}[X_{\mathcal{S}}]$. For $\xi > 0$, there exists $D = D(\xi) > 0$ such that for any $x \geq \max\{(1+\xi)\lambda, 2e\}$ we have*

$$\Pr[\alpha(\mathcal{H}_p(\mathcal{S})) \geq x] \leq \min\{\lambda^x, 1\} e^{-Dx}.$$

Proof. Lemma 1 in [7] and the estimate $x! \geq (x/e)^x \geq 2^x$ imply that

$$\Pr[\alpha(\mathcal{H}_p(\mathcal{S})) \geq x] \leq \lambda^x/x! \leq \lambda^x 2^{-x}.$$

By Lemma 2.46 in [11], which slightly extends Lemma 2 in [10], using $x \geq (1+\xi)\lambda$ a short calculation yields

$$\Pr[\alpha(\mathcal{H}_p(\mathcal{S})) \geq x] = \Pr[\alpha(\mathcal{H}_p(\mathcal{S})) \geq \lambda + (x - \lambda)] \leq e^{-\frac{(x-\lambda)^2}{2(\lambda+(x-\lambda)/3)}} \leq e^{-\frac{\xi^2 x}{2(1+\xi)(1+\xi/3)}}.$$

13

Lemma 21 follows for $D := \min\{\log 2, \frac{\xi^2}{2(1+\xi)(1+\xi/3)}\}$. $\qquad\square$

Setting $\beta := \varepsilon/(5A)$, and $C' := \max\{3\mathrm{e}, 10\mathrm{e}A\varepsilon^{-1}\}$, it is easily verified that, for n large enough,

$$(1 - 2\beta A)L_v' \geq \max\{(1 + \varepsilon/2)\mu_v, 2\mathrm{e}\} = \max\{(1 + \varepsilon/2)\mathbb{E}[X_{\mathcal{S}_1}], 2\mathrm{e}\},$$
$$\beta L_v' \geq \max\{\beta\mu_v, 2\mathrm{e}\} \geq \max\{2\mathbb{E}[X_{\mathcal{S}_2}], 2\mathrm{e}\},$$

where the last inequality follows with Lemma 20 from the observation that

$$\mathbb{E}[X_{\mathcal{S}_2}] = O(\max_{J \subsetneq H : e_J \geq 1} n^{2v_H - v_J - 1} p^{2e_H - e_J}) = O(\mu_v n^{-\gamma}).$$

Hence Lemma 21 implies the bounds (19) and (20) for $B := \min\{(1 - 2\beta A)D(\varepsilon/2), \beta D(1)\}$, where $D(\cdot)$ is the function given by Lemma 21. (Note that $\min\{\lambda^x, 1\} \leq \min\{\lambda, 1\}$ for all $\lambda > 0$ and $x \geq 1$.)

This concludes the proof of part (i) of Claim 19. Essentially the same arguments also yield part (ii). For the definition of the auxiliary graph $\mathcal{G} = \mathcal{H}_p(\mathcal{S}_1)$, we now consider the family $\mathcal{S}_1$ of all edge sets α of size $e_H - 1$ that together with $\{v_1, v_2\}$ form a copy of H. The bounds (18), (19), and (20), with μ_v replaced by μ_e, follow from almost the same arguments as before, and imply part (ii) of Claim 19 in exactly the same way. (For (19), note that (21) holds for the graph W obtained as the union of all α_i of a nice sequence of length T and the edge $\{v_1, v_2\}$, with $e(J_i) \geq 2$ for all $2 \leq i \leq T$. The probability for the existence of all these α_i is then bounded by $c_1 n^{v(W)-2} p^{e(W)-1}$, and from there the argument continues as before, using part (ii) of Lemma 20.)

References

[1] M. Aizenmann, J.T. Chayes, L. Chayes, J. Fröhlich, and L. Russo. On a sharp transition from area law to perimeter law in a system of random surfaces. *Communications in Mathematical Physics*, 92:16–69, 1983.

[2] B. Bollobás. Threshold functions for small subgraphs. *Mathematical Proceedings of the Cambridge Philosophical Society*, 90(2):197–206, 1981.

[3] B. Bollobás and I. Leader. Isoperimetric inequalities and fractional set systems. *Journal of Combinatorial Theory. Series A*, 56(1):63–74, 1991.

[4] S. Chatterjee. The missing log in large deviations for triangle counts. *Random Structures & Algorithms*, 40(4):437–451, 2012.

[5] B. DeMarco and J. Kahn. Upper tails for triangles. *Random Structures & Algorithms*, 40(4):452–459, 2012.

[6] R. DeMarco and J. Kahn. Tight upper tail bounds for cliques. *Random Structures & Algorithms*, 41(4):469–487, 2012.

[7] P. Erdős and P. Tetali. Representations of integers as the sum of k terms. *Random Structures & Algorithms*, 1(3):245–261, 1990.

[8] G. Grimmett. *Percolation*, volume 321 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, Second edition, 1999.

[9] T. E. Harris. A lower bound for the critical probability in a certain percolation process. *Proceedings of the Cambridge Philosophical Society*, 56:13–20, 1960.

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

[11] S. Janson, T. Łuczak, and A. Ruciński. *Random graphs*. Wiley-Interscience, New York, 2000.

[12] S. Janson, K. Oleszkiewicz, and A. Ruciński. Upper tails for subgraph counts in random graphs. *Israel Journal of Mathematics*, 142:61–92, 2004.

[13] S. Janson and A. Ruciński. The infamous upper tail. *Random Structures & Algorithms*, 20(3):317–342, 2002.

[14] S. Janson and A. Ruciński. The deletion method for upper tail estimates. *Combinatorica*, 24(4):615–640, 2004.

[15] Y. Kohayakawa, M. Schacht, and R. Spöhel. Upper bounds on probability thresholds for asymmetric Ramsey properties. To appear in *Random Structures & Algorithms*.

[16] O. Riordan and L. Warnke. The Janson inequalities for general up-sets. To appear in *Random Structures & Algorithms*. `arXiv:1203.1024`.

[17] V. Rödl and A. Ruciński. Threshold functions for Ramsey properties. *Journal of the American Mathematical Society*, 8(4):917–942, 1995.

[18] J. Spencer. Counting extensions. *Journal of Combinatorial Theory. Series A*, 55(2):247–255, 1990.