

MODERATE DEVIATIONS OF SUBGRAPH COUNTS IN THE ERDŐS-RÉNYI RANDOM GRAPHS $G(n, m)$ AND $G(n, p)$

CHRISTINA GOLDSCHMIDT, SIMON GRIFFITHS, AND ALEX SCOTT

ABSTRACT. The main contribution of this article is an asymptotic expression for the rate associated with moderate deviations of subgraph counts in the Erdős-Rényi random graph $G(n, m)$. Our approach is based on applying Freedman's inequalities for the probability of deviations of martingales to a martingale representation of subgraph count deviations. In addition, we prove that subgraph count deviations of different subgraphs are all linked, via the deviations of two specific graphs, the path of length two and the triangle. We also deduce new bounds for the related $G(n, p)$ model.

1. INTRODUCTION

Deviations of subgraph counts in random graphs, and in particular in the Erdős-Rényi random graph $G(n, p)$, have been the focus of intense study in recent years. Almost all of the results have concerned either small deviations (of the order of the standard deviation) or large deviations (of the order of the mean). Less is known about the intermediate range of moderately large deviations.

Corresponding to the first category, deviations of the order of the standard deviation, Ruciński established [31] that for the entire range of densities p such that $np^{e(H)}, (1-p)n^2 \rightarrow \infty$ the number of copies of a fixed graph H in $G(n, p)$ is asymptotically normally distributed. Articles with results that are quantitatively stronger have followed [3, 21, 29, 30]. On the other hand Janson [14] (building on the earlier articles, Janson [13] and Janson and Nowicki [16]) gives a general framework in which to think about random graph statistics. Among other results, he proves a functional central limit theorem for the evolution of subgraph count deviations, and that subgraph counts in $G(n, m)$ are also asymptotically normally distributed.

In the second category, deviations of the order of the mean, usually referred to as large deviations, have become a major focus in recent years. Interest in these problems grew after the seminal articles of Vu [32] and Janson and Ruciński [17] in the early 2000s provided many results, using a large range of techniques, which were still far from best possible. Important subsequent advances include the translation of such deviation problems into variational problems for graphons (Chatterjee and Varadhan [7]) and solutions to these variational problems for certain values of the parameters (Lubetzky and Zhao [24] and Zhao [34]). We recommend the survey of Chatterjee [6] and the references therein for a more detailed overview. Note that the approach of Chatterjee and Varadhan [7], which is applied in the context of the model $G(n, p)$, has been generalised to apply in $G(n, m)$ by Dembo and Lubetzky [8]. Very recently, a major breakthrough by Harel, Mousset and Samotij [19] has greatly extended the range of such large deviation results.

In this article, we focus on deviation events of some intermediate size, usually called moderate deviations. We shall focus on the random graph model $G(n, m)$, with a fixed number of edges, which we believe to be the more natural context in which to study moderate deviations of subgraph counts. For example, in the dense case, the standard deviation of the number of triangles in $G(n, m)$ is of order $n^{3/2}$, while it is of order n^2 in $G(n, p)$. This expresses the fact that by far the easiest way for $G(n, p)$ to have extra triangles is simply to have extra edges. By fixing the number of edges and working in $G(n, m)$ one studies the finer problem of other possible causes of triangle count deviations.

Our main contributions are as follows:

- (i) We give a general martingale-type expression for subgraph count deviations in $G(n, m)$ (see Theorem 2.1).
- (ii) We prove that subgraph count deviations are generally well predicted by the deviations of the counts of two specific graphs, P_2 and K_3 (see Theorem 1.6).
- (iii) Using the above results, we determine the asymptotic rate associated with moderately large subgraph count deviations. That is we determine the function $r = r(n)$ such that a deviation of this type has probability $\exp(-r(1 + o(1)))$ (see Theorem 1.1).
- (iv) We deduce results concerning moderately large subgraph count deviations in $G(n, p)$ which are significantly stronger than previously known bounds (see Theorems 1.9 and 1.11).

We state other auxiliary results along the way, such as an approximate bound on deviation probabilities across the whole range of deviations, Theorem 1.7 and an estimate for the tail of the binomial distribution, Theorem 1.13.

We require the following notation. We write $N_H(G)$ for the number of embeddings of a graph H in a graph G . That is, the number of injective functions $\phi : V(H) \rightarrow V(G)$ such that

$$\phi(u)\phi(v) \in E(G) \quad \text{for all} \quad uv \in E(H).$$

This is also referred to in the literature as the number of *isomorphic copies* of H in G . When we count without multiplicity we write $\binom{G}{H}$, so that, for example

$$N_{K_3}(K_4) = 24 \quad \text{and} \quad \binom{K_4}{K_3} = 4.$$

We shall be interested in $N_H(G)$, where H is a fixed graph and G is a large random graph. For example, we think of a fixed graph H with $v = v(H)$ vertices, and $e = e(H)$ edges, and a large graph G with n vertices and m edges, where n is taken very large, and m behaves as a function of n . (In view of this choice of notation, we will never use e to denote the base of the natural logarithm, but will rather write $\exp(1)$.)

Let $N := \binom{n}{2}$. For a graph H with v vertices, and e edges, the expected number of embeddings (isomorphic copies) of H in $G(n, m)$ is

$$L_H(m) := \frac{\binom{n}{v} \binom{m}{e}}{\binom{N}{e}}, \quad (1.1)$$

where $(n)_k := n(n-1)\dots(n-k+1)$ denotes the falling factorial. It will be useful at times to note that

$$L_H(m) - L_H(m-1) = \frac{1}{N-m+1} \sum_{f \in E(H)} (L_{H \setminus f}(m-1) - L_H(m-1)). \quad (1.2)$$

The intuition behind the identity is that both sides represent the increase in the expected number of embeddings of H caused by the addition of an edge: the sum on the right hand side corresponds to the expected number of almost complete embeddings, in the sense that a single edge is not present. Alternatively, direct calculation shows that both sides have value $e(n)_v(m-1)_{e-1}/(N)_e$.

A natural way to generate $G(n, m)$ is to add the edges one at a time. The Erdős-Rényi random graph process $(G_i : i = 0, \dots, N)$ is defined as follows. Let G_0 be the empty graph, and for each $i \geq 0$ let G_{i+1} be obtained by adding a uniformly chosen edge to G_i . Clearly G_m is distributed as $G(n, m)$. The process ends with G_N being the complete graph K_n . We observe that the process is Markovian. We refer the reader to the books [5, 15] for further background on random graphs.

Our focus will be on subgraph count deviations in G_m . We write $D_H(G_m)$ for the deviation of the H -count in G_m . That is,

$$D_H(G_m) := N_H(G_m) - L_H(m). \quad (1.3)$$

We shall see that paths of length two, which we denote \wedge , and triangles, which we denote \triangle , play a particularly important role. We write $\binom{H}{\wedge}$ for the number of paths of length two in a graph H and $\binom{H}{\triangle}$ for the number of triangles in H .

Let us define the function $\gamma_H(t)$ for $t \in (0, 1)$ by

$$\gamma_H(t) := \left(4 \binom{H}{\wedge}^2 t^{2e-2} (1-t)^2 + 12 \binom{H}{\triangle}^2 t^{2e-3} (1-t)^3 \right)^{-1}. \quad (1.4)$$

We now state our main result concerning the asymptotic rate of moderate deviations of subgraph counts. We use the notation $f \ll g$ for $f = o(g)$. We express the deviation as a multiple of $n^{v-3/2}$ as this is the order of the standard deviation (in the dense case).

The model we consider is defined as follows. Let $(G_{n,m} : m = 0, \dots, N)$, $n \geq 1$ be independent copies of the Erdős-Rényi random graph process, and let $(G_{n,t})_{n \geq 1}$ denote the sequence of random graphs $(G_{n,m_n})_{n \geq 1}$, where $m_n = \lfloor tN \rfloor$. We will be interested in $G_{n,t}$ both in the case that $t \in (0, 1)$ is a constant, and the case that $t = t(n)$ is a function of n . We exclude the case that $t(n)$ converges to 1 (see Remark 1.4).

Theorem 1.1. *Let $t = t(n) \in (0, 1)$ be a sequence bounded away from 1, and let H be graph with v vertices, e edges, and $\binom{H}{\wedge} \geq 1$. Then*

$$\mathbb{P} \left(D_H(G_{n,t}) > \alpha_n n^{v-3/2} \right) = \exp \left(-\gamma_H(t) \alpha_n^2 (1 + o(1)) \right),$$

for every sequence $(\alpha_n : n \geq 1)$ which satisfies either

- (i) $1 \ll \alpha_n \ll n^{1/2}$ and $t(n) = t \in (0, 1)$ is constant, or
- (ii) $\max\{t^{1/2}n^{-1/2} \log n, t^{e-3/2}\} \ll \alpha_n \ll \min\{t^{2e-5/2}n^{1/2}, t^{e+2}n^{1/2}\}$.

Furthermore the same holds for $\mathbb{P} \left(D_H(G_{n,t}) < -\alpha_n n^{v-3/2} \right)$.

Remark 1.2. We initially proved the results of this article in the dense case (i.e., with $t \in (0, 1)$ a constant), and have now partially extended them to sparser regimes. The problem of finding the asymptotic rate across the whole range of sparse densities remains open.

Remark 1.3. In the sparse case, $t = o(1)$, we may simplify $\gamma_H(t)$ to

- (i) $(4\binom{H}{\wedge})^2 t^{2e-2}(1-t)^2)^{-1}$ in the case $\binom{H}{\Delta} = 0$, or
- (ii) $(12\binom{H}{\Delta})^2 t^{2e-3}(1-t)^3)^{-1}$ in the case $\binom{H}{\Delta} \geq 1$.

We may also note that the same dichotomy applies to $\Lambda_H(G_{n,t})$, see (1.5), in the sense that the term involving $D_\Delta(G_{n,t})$ dominates, in the sparse case, if $\binom{H}{\Delta} \geq 1$.

Remark 1.4. Our proof of Theorem 1.1 breaks down as t approaches 1. However, with an alternative approach one may obtain the same bound provided:

$$(1-t)^{e-3/2} \ll \alpha_n \ll (1-t)^{e+2} n^{1/2}.$$

The alternative approach is to approximate $D_H(G_{n,t})$ by $\Lambda_H(G_{n,t})$ (using Theorem 1.6), and apply Corollary 2.8 to each of $D_\wedge(G_{n,t})$ and $D_\Delta(G_{n,t})$ to express these deviations in terms of deviations in the complement, and then apply Theorem 1.1 to the complement. (Since the complement is sparse the deviation event is more easily achieved by D_\wedge and the contribution of D_Δ is essentially trivial.)

Remark 1.5. In the dense case, $t \in (0, 1)$ constant, the range of deviations considered $(\omega(n^{v-3/2}), o(n^{v-1}))$, corresponds to the range strictly between the orders of magnitude of the standard deviation of $D_H(G)$ for $G \sim G(n, m)$ and $G \sim G(n, p)$ respectively. This range is best possible, in the sense that the asymptotics of $\log(\mathbb{P}(D_H(G_{n,t}) > \alpha_n n^{v-3/2}))$ are different if $\alpha_n = O(1)$ or $\alpha_n = \Omega(n^{1/2})$. For $\alpha_n = O(1)$ this follows from the central limit theorem of Janson [14]. On the other hand, if $\alpha_n = \Omega(n^{1/2})$ then the asymptotic log probability is larger¹. Theorem 1.7 below gives an exponent which is best possible up to multiplication by constant (in the dense case) across the whole range of deviations $(\omega(n^{v-3/2}), \Theta(n^v))$.

A key step in proving Theorem 1.1 is to establish a relation between the subgraph count deviations $D_H(G_{n,t})$ of different graphs H . Specifically, we prove that $D_H(G_{n,t})$, the deviation of the H -count in $G_{n,t}$ is generally well predicted by a certain linear combination of $D_\wedge(G_{n,t})$, the deviation of the P_2 count, and $D_\Delta(G_{n,t})$, the deviation of the triangle count. Let us define

$$\Lambda_H(G_{n,t}) := n^{v-3} t^{e-2} \left(\binom{H}{\wedge} - 3 \binom{H}{\Delta} \right) D_\wedge(G_{n,t}) + n^{v-3} t^{e-3} \binom{H}{\Delta} D_\Delta(G_{n,t}) \quad (1.5)$$

to be this linear combination, where $v = v(H)$ and $e = e(H)$. Note that $\Lambda_H(G_{n,t})$ is n^{v-3} times a linear combination $\kappa D_\wedge(G_{n,t}) + \rho D_\Delta(G_{n,t})$, in which the coefficients depend only on H and t .

¹As a particular example, if any vertex has degree $n-1$ then $D_\wedge(G_{n,t}) \geq (1-t)^2 n^2$, and this has probability at least $\Omega(t^n)$ which is larger than $\exp(-\gamma_\wedge(t)(1-t)^4 n(1+o(1)))$ for certain values of $t \in (0, 1)$.

Theorem 1.6. *Let H be a graph with v vertices and e edges. There exists a constant $C = C(H)$ such that for all n , and all $t = t(n) \in (0, 1)$, we have*

$$\mathbb{P}\left(|D_H(G_{n,t}) - \Lambda_H(G_{n,t})| > Cbt^{1/2}n^{v-2}\right) \leq \exp(-b) \quad (1.6)$$

for all $3 \log n \leq b \leq t^{1/2}n$. Furthermore

$$\mathbb{P}\left(|D_H(G_{n,t}) - \Lambda_H(G_{n,t})| > Cbn^{v-2}\right) \leq \exp(-b) \quad (1.7)$$

for all $b \geq 3 \log n$.

We also state a weaker version of Theorem 1.1 which applies across the entire range of possible deviations.

Theorem 1.7. *Let H be graph with v vertices and e edges. Then there is a constant $c = c(H)$ such that for all $t = t(n) \in (0, 1)$, and for all $\alpha, n \geq c^{-1}$, we have*

$$\mathbb{P}\left(|D_H(G_{n,t})| > \alpha n^{v-3/2}\right) \leq \exp(-c\alpha \min\{\alpha, n^{1/2}\}).$$

A discussion of our approach. Our main results, and Theorem 1.1 in particular, are proved using a pair of lemmas of Freedman [12], stated in Section 3, which provided an upper and a lower bound on deviation probabilities of martingales. In particular, in certain circumstances, they imply that the probability that a martingale $(S_i)_{i=0}^m$ has a certain deviation α from its mean, is given by

$$\exp\left(\frac{-\alpha^2(1+o(1))}{2\beta}\right),$$

where β is asymptotic to the discrete quadratic variation

$$\sum_{i=1}^m \mathbb{E}[(S_i - S_{i-1})^2 | \mathcal{F}_{i-1}]$$

of the process.

In order to apply these results in our setting we are presented with two main challenges. The first is to give a martingale expression for subgraph count deviations $D_H(G_m)$. We state both a precise martingale expression for $D_H(G_m)$, see Theorem 2.1, and an approximate (but simpler) martingale expression for $D_H(G_m)$, see Theorem 2.4. The precise martingale expression, Theorem 2.1, is relatively easy to prove. To verify the accuracy of the approximate martingale expression, Theorem 2.4 is substantially more difficult and this is done in Section 5, as part of the proof of Theorem 1.6.

The second challenge is to understand the behaviour of the discrete quadratic variation of these martingale expressions. The relevant result, Proposition 7.1, which follows from the more precise Proposition 7.2, allows us to deduce that this discrete quadratic variation is very predictable – it is very likely to be close to a particular deterministic function.

Our proof of Proposition 7.2 makes use of Theorem 1.6, which concerns the relationship between subgraph count deviations, and Theorem 1.7.

We remark that the Hoeffding-Azuma inequality, Lemma 3.1, is simpler to use than Freedman's inequality and for this reason we use it to prove various auxiliary results. However, we stress that the main theorem itself, Theorem 1.1, could not be proved using

the Hoeffding-Azuma inequality. In essence, the Hoeffding-Azuma inequality gives substantially weaker bounds than Freedman's inequality when the martingale increments are typically much smaller than their maximum possible value; more precisely, when the conditional second moment of the increments, $\mathbb{E}[X_i^2 | \mathcal{F}_{i-1}]$, are typically much smaller than their essential supremum, $\|X_i\|_\infty$.

Remark 1.8. We developed this discrete martingale approach to understanding subgraph count deviations precisely because this approach combines well with results, such as those of Freedman, about discrete martingales. We would like to acknowledge that a continuous time martingale framework for subgraph counts, and random graph statistics in general, was developed by Janson [14] in the 1990s. There are number of connections between our results and those of Janson. In particular, the significance of P_2 and triangle counts is also evident from Janson's results. We encourage the interested reader to read [14] for results on the central limit theorem in $G(n, m)$, results on functional limits of random graph statistics, and much more.

1.1. Moderate deviations of subgraph counts in $G(n, p)$. Until this point we have focussed exclusively on deviation events in the model $G(n, m)$. We now deduce results concerning the probabilities of moderate deviations of subgraph counts in the Erdős-Rényi random graph $G(n, p)$. We write q for $1 - p$ here and throughout.

We shall suppress n from the notation and write G_p for a graph chosen according to the distribution $G(n, p)$, i.e., with each edge included in G_p independently with probability p . For a graph H with v vertices and e edges, we write

$$L_H(p) := \binom{n}{v} p^e$$

for the expected number of isomorphic copies of H in G_p , and

$$D_H(G_p) := N_H(G_p) - L_H(p)$$

for the deviation of the H -count $N_H(G_p)$ from its mean. We consider deviations of size $\delta_n L_H(p)$ (that is, δ_n times the mean) where $n^{-1} \ll \delta_n \ll 1$. This corresponds to the range strictly between the standard deviation and the regime of large deviations (i.e., the order of the mean).

Our first result corresponds to the range $\delta_n \ll n^{-1/2}$. In this range we obtain a precise asymptotic expression for the deviation probability. We remark that this result has already been obtained using a completely different approach by Féray, Méliot and Nikeghbali, see Theorem 10.1 in [11]. Their result, which is proved in the framework of mod- ϕ convergence, also gives an asymptotically tight expression for deviation probabilities in this range.

Theorem 1.9. *Let $p \in (0, 1)$, and let H be a graph with v vertices and e edges. Let $(\delta_n : n \geq 1)$ be a sequence such that $n^{-1} \ll \delta_n \ll n^{-1/2}$. Then*

$$\begin{aligned} & \mathbb{P}(D_H(G_{n,p}) > \delta_n p^e \binom{n}{v}) \\ &= (1 + o(1)) \sqrt{\frac{e^2 q}{\pi p}} \exp\left(-\frac{\delta_n^2 p n^2}{4e^2 q} + \frac{((3e-2) - (3e-1)p)\delta_n^3 p n^2}{12e^3 q^2} - \log(n\delta_n)\right). \end{aligned}$$

Remark 1.10. Observe that the only dependence on the graph H in this range is via the numbers of edges e . This is related to the fact that it is vastly easier to achieve this deviation by having extra edges in G_p than achieving the deviation in G_m for $m \approx pN$. In other words, the above expression corresponds to the probability of the appropriate deviation of the binomial distribution.

The range of larger deviations, $n^{1/2} \ll \delta_n \ll 1$, is more difficult to study in that there is a non-trivial interplay between the deviation probabilities of the binomial distribution and subgraph count deviations in G_m . In particular, we require the full strength of Theorem 1.1 to obtain the following result. It is for this reason that $\gamma_H(p)$ appears in the rate.

We will also require the following notation. Recall that $N := \binom{n}{2}$. Set

$$x_* := \left[(1 + \delta_n)^{1/e} - 1 \right] \sqrt{\frac{pN}{q}},$$

and for $0 < x < \sqrt{N}/2$ define

$$E(x, N) = \sum_{i=1}^{\infty} \frac{(p^{i+1} + (-1)^i q^{i+1}) x^{i+2}}{(i+1)(i+2) p^{i/2} q^{i/2} N^{i/2}}.$$

We can now state our result for larger deviations. In fact the result may be stated across the whole range $n^{-1} \ll \delta_n \ll 1$.

Theorem 1.11. *Let $p \in (0, 1)$, and let H be a graph with v vertices and e edges. Let $(\delta_n : n \geq 1)$ be a sequence such that $n^{-1} \ll \delta_n \ll 1$. Then*

$$\begin{aligned} & \mathbb{P}(D_H(G_{n,p}) > \delta_n p^e (n)_v) \\ &= \exp\left(-\frac{x_*^2}{2} + E(x_*, N) + (1 + o(1)) \frac{\delta_n^2 n}{16 \gamma_H(p) e^4 p^{2e-2} q^2} + O(\log n)\right). \end{aligned}$$

We remark that the asymptotic rate, which gives the bound

$$\exp\left(-\frac{x_*^2}{2} (1 + o(1))\right) = \exp\left(\frac{-\delta_n^2 p n^2}{4e^2 q} + o(\delta_n^2 n^2)\right), \quad (1.8)$$

already appears in the articles of Döring and Eichelsbacher [9] and [10]. The difference between the results is the order of magnitude of the error term. In the range $\delta_n \gg n^{-1/2} \sqrt{\log n}$ we have an error term of the form $o(\delta_n^2 n)$ in the exponent²

On the other hand, Döring and Eichelsbacher in [9] obtained the asymptotic rate for the range of parameters

$$\sqrt{\frac{q}{p}} n^{-1} \ll \delta_n \ll p^{3e-2} q^2$$

by an estimation of the log-Laplace transform and the Gartner-Ellis theorem. In [10] Döring and Eichelsbacher show that results may also be obtained through a moderate deviation principles via cumulants, in an approach based on a celebrated lemma of large

²We believe that it ought to be possible to reduce the error term in the missing range $\Omega(n^{-1/2}) \leq \delta_n \leq O(n^{-1/2} \sqrt{\log n})$ to be of the form $o(\delta_n^2 n)$, rather than $O(\log n)$. For example, one might prove this by combining our approach with the central limit theorem for subgraph count deviations.

deviations theory due to Rudzkiš, Saulis and Statulevičius. The results of [10] include the asymptotic rate for the range of parameters

$$\sqrt{\frac{q}{p}} n^{-1} \ll \delta_n \ll p^{(3e-4)/5} q^{4/5} n^{-4/5}.$$

It may be of interest to investigate for which ranges of δ_n and p our more precise expansion remains valid. Janson and Warnke [18] focussed on the lower tail and found the same asymptotic expression, (1.8), for the logarithm of $\mathbb{P}(D_H(G_{n,p}) < -\delta_n p^e(n)_v)$ across the whole range of moderate deviations and densities $p \ll n^{-1/m_2(H)}$ where $m_2(H) = \max_{J \subseteq H} (e(J) - 1)/(v(J) - 2)$. Furthermore, their result also applies in the setting of k -uniform hypergraphs.

We also remark that a weaker result with the $1 + o(1)$ replaced by $O(1)$ may be proved using only Theorem 1.7 to bound deviation probabilities for $D_H(G_m)$. In this sense Theorem 1.1 has a relatively minor impact on the strength of the bound obtained for deviations $D_H(G_p)$ in $G(n, p)$. On the other hand, we believe that this reinforces our argument that $G(n, m)$ is the more natural setting in which to study these subgraph count deviations in the first place.

Finally, the reader may wonder why we gave the implicit definition $-x_*^2/2 + E(x_*, N)$ rather than just writing out the expansion. The problem is that the number (and complexity) of the terms in the expansion grows as δ_n increases. We illustrate this by giving the expansion in the range $n^{-1/2} \log n \ll \delta_n \ll n^{-2/5}$.

Corollary 1.12. *Let $p \in (0, 1)$, and let H be a graph with v vertices and e edges. Let $(\delta_n : n \geq 1)$ be a sequence such that $n^{-1/2} \log n \ll \delta_n \ll n^{-2/5}$. Then*

$$\begin{aligned} & \mathbb{P}(D_H(G_{n,p}) > \delta_n p^e(n)_v) \\ &= \exp \left(-\frac{\delta_n^2 p n^2}{4q e^2} - \frac{p[(3e-1)q-1]\delta_n^3 n^2}{12q^2 e^3} + \frac{p[(e-1)q(8e+11)q-6] + 1 - 3pq}{48q^3 e^4} \delta_n^4 n^2 \right. \\ & \quad \left. + (1 + o(1)) \frac{\delta_n^2 n}{16\gamma_H(p) e^4 p^{2e-2} q^2} \right). \end{aligned}$$

Naturally, both results will rely on an estimate for tail probabilities of the binomial distribution. While estimates are available (Littlewood [23] for example, see also McKay [27]), we shall give a proof of the following estimate for completeness. This result is essentially due to Bahadur [2]. In addition to $E(x, N)$ defined above, let us also define a truncated version of the sum:

$$E(x, N, J) = \sum_{i=1}^J \frac{(p^{i+1} + (-1)^i q^{i+1}) x^{i+2}}{(i+1)(i+2) p^{i/2} q^{i/2} N^{i/2}}.$$

Adapting the argument from Theorem 2 of Bahadur [2], we obtain the following asymptotics for

$$b_N(k) = \mathbb{P}(\text{Bin}(N, p) = k)$$

and

$$B_N(k) = \mathbb{P}(\text{Bin}(N, p) \geq k).$$

in terms of $x_N = \frac{k-pN}{\sqrt{Npq}}$. The theorem is valid for $p \in (0, 1)$ a constant or $p = p_N$ a function.

Theorem 1.13. *Suppose that (x_N) is a sequence such that $1 \ll x_N \ll \sqrt{Npq}$. Then*

$$b_N(\lfloor pN + x_N \sqrt{Npq} \rfloor) = (1 + o(1)) \frac{1}{\sqrt{2\pi Npq}} \exp\left(-\frac{x_N^2}{2} - E(x_N, N)\right)$$

and

$$B_N(pN + x_N \sqrt{Npq}) = (1 + o(1)) \frac{1}{x_N \sqrt{2\pi}} \exp\left(-\frac{x_N^2}{2} - E(x_N, N)\right).$$

Furthermore, if $1 \ll x_N \ll (pqN)^{1/2}(pqN)^{-1/(J+3)}$ then the infinite sum $E(x_N, N)$ may be replaced by the finite sum $E(x_N, N, J)$ in both expressions.

The proof of Theorem 1.13 is given in the appendix.

Let us now return to Theorems 1.9 and 1.11 and give an overview of their proofs. We immediately observe, by conditioning on the number of edges of G_p , that we may express $\mathbb{P}(D_H(G_p) > \delta_n p^e(n)_v)$ as a sum:

$$\mathbb{P}(D_H(G_p) > \delta_n p^e(n)_v) = \sum_{m=0}^N b_N(m) \mathbb{P}(N_H(G_m) > (1 + \delta_n) p^e(n)_v).$$

For $m > pN$ we have that the first term $(b_N(m))$ is decreasing while the second is increasing. The proofs are therefore concerned with identifying which terms make the largest contribution.

In the case of Theorem 1.11 the problem reduces exactly to a calculation of the maximum, as all other effects are swallowed up in the $O(\log n)$ error term in the exponent.

In the case of Theorem 1.9 we exploit the fact that there is an interval $[m_-, m_+]$ over which the first term $(b_N(m))$ decreases very little and the second term grows from $o(1)$ to $1 - o(1)$.

Layout of the article. In Section 2 we present Theorem 2.1, the general martingale expression for the subgraph count deviation $D_H(G_m)$. We also present an important approximate representation, Theorem 2.4, and a lemma relating subgraph counts to subgraph counts in the complementary graph. In Section 3 we state the martingale inequalities that we shall use throughout the article. These include the Hoeffding-Azuma inequality, a related inequality adapted to $G(n, m)$ and, crucially, Freedman's inequalities for the probability of deviations of martingales. In Section 4 we prove bounds concerning the behaviour of degrees and codegrees in $G(n, m)$. In Section 5 we prove the approximate representation result, Theorem 2.4, and deduce Theorem 1.6.

We then turn our focus towards deviation probabilities themselves. In Section 6, we prove Theorem 1.7, which gives a general though not especially precise bound on subgraph count probabilities. In order to prove the tighter result Theorem 1.1, we must first understand better the variances and covariances of the increments of the martingale representation. In Section 7 we prove bounds on general covariances of increments in the martingale representation of $D_H(G_m)$, and in Section 8 we prove Theorem 1.1.

Finally, in Section 9 we deduce our results for subgraph count deviations in $G(n, p)$.

Notation. Throughout N denotes $\binom{n}{2}$. Let us also recall that $N_H(G)$ denotes the number of embeddings of a graph H in a graph G , and that $\binom{G}{H}$ denotes the number of copies of H in G counted without multiplicity.

Use of m and t : In Sections 2, 3 and 4, we work with the Erdős-Rényi random graph process and use m simply to denote the number of edges of G_m . In later sections m is used specifically to refer to $\lfloor tN \rfloor$. The latter use corresponds to the use in the definition of $G_{n,t}$ as $G_{n,m}$ with $m = \lfloor tN \rfloor$.

Use of i and s : We think of G_m as the result of a realisation of the random graph process $(G_i : i = 0, \dots, m)$. In this context we use s throughout to refer to i/N , the proportion of pairs that occur as edges of G_i . This usage occurs below in the definitions of $\mathbb{X}_H(G_m)$, $\Lambda_H^{**}(G_{n,t})$, $V_{F,F'}(i, n)$ and $W_{F,F'}(G_{i-1})$, for example.

Use of v and e : We use v and e to denote the number of vertices and edges of the small graph we are currently working with. The majority of the time this is the graph H . However, in Section 5.1 and Section 7 it is the graph F . When necessary we write $v(F)$ and $e(F)$, for example, to avoid ambiguity.

Introductory notation:

First introduced:

$$L_H(m) := \frac{\binom{n}{v} \binom{m}{e}}{\binom{N}{e}} \quad (1.1)$$

$$D_H(G_m) := N_H(G_m) - L_H(m) \quad (1.3)$$

$$A_H(G_m) := N_H(G_m) - N_H(G_{m-1}) \quad (2.1)$$

$$\gamma_H(t) := \left(4 \binom{H}{\wedge}^2 t^{2e-2} (1-t)^2 + 12 \binom{H}{\Delta}^2 t^{2e-3} (1-t)^3 \right)^{-1} \quad (1.4)$$

Martingale increments related to subgraph count deviations:

First introduced:

$$X_H(G_m) := A_H(G_m) - \mathbb{E}[A_H(G_m) \mid G_{m-1}] \quad (2.2)$$

$$X_F^*(G_i) := n^{v-3} s^{e(F)-2} \left(\binom{F}{\wedge} - 3 \binom{F}{\Delta} \right) X_{\wedge}(G_i) + n^{v-3} s^{e(F)-3} \binom{F}{\Delta} X_{\Delta}(G_i) \quad (5.1)$$

$$Y_F(G_i) := X_F(G_i) - X_F^*(G_i) \quad (5.2)$$

Additionally, $\mathbb{X}_H(G_i; t)$, see (2.5), is defined by

$$\begin{aligned} & \mathbb{X}_H(G_i; t) \\ & := n^{v-3} t^{e-3} \left(t \binom{H}{\wedge} \frac{(1-t)^2}{(1-s)^2} X_{\wedge}(G_i) + \binom{H}{\Delta} \frac{(1-t)^3}{(1-s)^3} (X_{\Delta}(G_i) - 3s X_{\wedge}(G_i)) \right). \end{aligned}$$

There are three random variables $\Lambda_H(G_{n,t})$, $\Lambda_H^*(G_{n,t})$ and $\Lambda_H^{**}(G_{n,t})$ that all approximate $D_H(G_{n,t})$ in some sense. They are first defined respectively as equations (1.5), (2.6) and (5.23). In the following definition m denotes $\lfloor tN \rfloor$ and $\mathbb{X}_H(G_i, t)$ and $X_F(G_i)$ are as

defined above.

$$\begin{aligned}\Lambda_H(G_{n,t}) &:= n^{v-3}t^{e-2} \left(\binom{H}{\wedge} - 3\binom{H}{\Delta} \right) D_{\wedge}(G_{n,t}) + n^{v-3}t^{e-3} \binom{H}{\Delta} D_{\Delta}(G_{n,t}) \\ \Lambda_H^*(G_{n,t}) &:= \sum_{i=1}^m \mathbb{X}_H(G_i; t) \\ \Lambda_H^{**}(G_{n,t}) &:= \sum_{i=1}^m \sum_{F \subseteq E(H)} \frac{(1-t)^{e(F)}(t-s)^{e-e(F)}}{(1-s)^e} X_F(G_i)\end{aligned}$$

Degrees, codegrees and their deviations:

First introduced:

$d_u(G_m) :=$ degree of u in G_m

$$D_u(G_m) := d_u(G_m) - \frac{2m}{n} \quad (4.1)$$

$d_{u,w}(G_m) :=$ codegree of u in G_m

$$D_{u,w}(G_m) := d_{u,w}(G_m) - \frac{(n-2)(m)_2}{(N)_2} \quad (4.2)$$

$$\Delta(e_i) := D_u(G_{i-1})^2 + D_w(G_{i-1})^2 + D_{uw}(G_{i-1})^2 \quad (5.7)$$

Functions related to covariance:

First introduced:

$$V_{F,F'}(i, n) := n^{v+v'-5} s^{e+e'-4} (1-s) (s\theta_1(F, F') + (1-s)\theta_2(F, F')) \quad (7.1)$$

$$\theta_1(F, F') := 8 \binom{F}{\wedge} \binom{F'}{\wedge} \quad (7.2)$$

$$\theta_2(F, F') := 36 \binom{F}{\Delta} \binom{F'}{\Delta} \quad (7.2)$$

$$W_{F,F'}(G_{i-1}) := 8n^{v+v'-7} s^{e+e'-4} \binom{F}{\wedge} \binom{F'}{\wedge} D_{\wedge}(G_{i-1}) \quad (7.3)$$

Finally, we write \log for the natural logarithm.

2. MARTINGALE EXPRESSION FOR $D_H(G_m)$

In this section we state and prove Theorem 2.1, our martingale expression for $D_H(G_m)$. In doing so, we also prove Lemma 2.3, which concerns the expected number of copies of H created with the addition of the m th edge. We also state an approximate expression for $D_H(G_m)$, see Theorem 2.4, which will be proved in Section 5 as part of the proof of Theorem 1.6.

For the duration of the section, let us fix n and let $(G_m : m = 0, \dots, N)$ be a realisation of the Erdős-Rényi random graph process on n vertices. It is helpful to think of G_m as also including the information of the order in which its edges were added. Let us define

$$A_H(G_m) := N_H(G_m) - N_H(G_{m-1}), \quad (2.1)$$

the number of embeddings (isomorphic copies) of H created with the addition of the m th edge. Our martingale expression for $D_H(G_m)$ will be based on centered versions of these

random variables. Let

$$X_H(G_m) := A_H(G_m) - \mathbb{E}[A_H(G_m) | G_{m-1}]. \quad (2.2)$$

Note that $X_H(G_m)$ is obtained from $A_H(G_m)$ by shifting it so that $\mathbb{E}[X_H(G_m) | G_{m-1}] = 0$.

We may now state the martingale expression for $D_H(G_m)$.

Theorem 2.1. *Let H be a graph with v vertices and e edges. Then*

$$D_H(G_m) = \sum_{i=1}^m \sum_{F \subseteq E(H)} \frac{(N-m)_{e(F)}(m-i)_{e-e(F)}}{(N-i)_e} X_F(G_i), \quad (2.3)$$

where the inner sum is taken over all 2^e graphs F with $V(F) = V(H)$ and $E(F) \subseteq E(H)$.

Remark 2.2. Equation (2.2) shows that $X_H(G_m)$ is a martingale increment with respect to the natural filtration of G_0, \dots, G_N . Since

$$\sum_{F \subseteq E(H)} \frac{(N-m)_{e(F)}(m-i)_{e-e(F)}}{(N-i)_e} X_F(G_i)$$

is a linear combination of the random variables $X_F(G_i)$, it too is a martingale increment, and so (2.3) is indeed a martingale.

We begin with a lemma about $\mathbb{E}[A_H(G_m) | G_{m-1}]$, the expected number of embeddings (isomorphic copies) of H created with the m th edge, given the graph G_{m-1} .

Lemma 2.3. *In the Erdős-Rényi random graph process $(G_m : m = 0, \dots, N)$,*

$$\begin{aligned} \mathbb{E}[A_H(G_m) | G_{m-1}] &= \frac{1}{N-m+1} \sum_{f \in E(H)} (N_{H \setminus f}(G_{m-1}) - N_H(G_{m-1})) \\ &= (L_H(m) - L_H(m-1)) + \frac{1}{N-m+1} \sum_{f \in E(H)} (D_{H \setminus f}(G_{m-1}) - D_H(G_{m-1})), \end{aligned}$$

where $H \setminus f$ denotes the graph obtained from H by removing the edge f .

Proof. Let us first observe that the second equality follows directly from the definitions. Indeed, one may simply expand $N_H(G_m)$ as $L_H(m) + D_H(G_m)$, and use (1.2).

We now prove the first equality. We may view $A_H(G_m)$, the number of embeddings (isomorphic copies) of H created with the addition of the m th edge, e_m , as a sum

$$A_H(G_m) = \sum_{f \in E(H)} A_{H,f}(G_m),$$

where $A_{H,f}(G_m)$ denotes the number of embeddings of H created with the addition of e_m , in which e_m is the image of the edge f of H . It therefore suffices to prove that

$$\mathbb{E}[A_{H,f}(G_m) | G_{m-1}] = \frac{1}{N-m+1} (N_{H \setminus f}(G_{m-1}) - N_H(G_{m-1})) \quad (2.4)$$

for each $f \in E(H)$.

Fix $f \in E(H)$. In order for an injective function $\phi : V(H) \rightarrow V(G_m)$ to represent an embedding of H in G_m , but not in G_{m-1} , and have $e_m = \phi(f)$, it is necessary and

sufficient that ϕ embeds $H \setminus \{f\}$ into G_{m-1} , that $\phi(f)$ is not an edge of G_{m-1} , and, finally, that e_m is chosen to be $\phi(f)$.

The number of injective functions obeying the first two conditions is precisely $N_{H \setminus f}(G_{m-1}) - N_H(G_{m-1})$, and, for any such ϕ , the probability that e_m is chosen to be $\phi(f)$ is $1/(N - m + 1)$. The required equation, (2.4), follows immediately. \square

We now prove Theorem 2.1.

Proof. The proof is by induction on $e = e(H)$, and on $m \in \{0, \dots, N\}$. If $e = 1$ or $m = 0$ the result holds trivially. Now consider a graph H with v vertices and $e \geq 2$ edges, and $m \in \{1, \dots, N\}$. We may expand $D_H(G_m)$ as follows

$$\begin{aligned} D_H(G_m) &= N_H(G_m) - L_H(m) \\ &= N_H(G_{m-1}) + A_H(G_m) - L_H(m-1) - (L_H(m) - L_H(m-1)) \\ &= D_H(G_{m-1}) + A_H(G_m) - (L_H(m) - L_H(m-1)) \\ &= D_H(G_{m-1}) + X_H(G_m) + \mathbb{E}[A_H(G_m) | G_{m-1}] - (L_H(m) - L_H(m-1)) \\ &= D_H(G_{m-1}) + X_H(G_m) + \frac{1}{N - m + 1} \sum_{f \in E(H)} (D_{H \setminus f}(G_{m-1}) - D_H(G_{m-1})), \end{aligned}$$

where we have used the definition of $X_H(G_m)$ in the fourth line and Lemma 2.3 in the last line.

We have an expression for $D_H(G_m)$ in terms of $X_H(G_m)$ and a linear combination of deviations $D_F(G_{m-1})$ with $F \subseteq H$. By the induction hypothesis each of these may be expressed as a linear combination of the $X_F(G_i)$ with $F \subseteq E(H)$ and $1 \leq i \leq m$. One may check that the resulting expression for $D_H(G_m)$ is that claimed. \square

We now give a simpler expression which approximates $D_H(G_m)$ very well. Since this expression is itself closely related to the quantity $\Lambda_H(G_{n,t})$ which appears in Theorem 1.6, we use the notation $\Lambda_H^*(G_{n,t})$. For a graph H with v vertices and e edges, let us first define

$$\mathbb{X}_H(G_i; t) := n^{v-3} \left(t^{e-2} \binom{H}{\wedge} \frac{(1-t)^2}{(1-s)^2} X_{\wedge}(G_i) + t^{e-3} \binom{H}{\triangle} \frac{(1-t)^3}{(1-s)^3} (X_{\triangle}(G_i) - 3sX_{\wedge}(G_i)) \right) \quad (2.5)$$

where $s = i/N$, as it is throughout the article, and define

$$\Lambda_H^*(G_{n,t}) := \sum_{i=1}^m \mathbb{X}_H(G_i; t). \quad (2.6)$$

We are now ready to state Theorem 2.4. The statement will be given for $t \in (0, 1/2]$. This form is sufficient for the proof of Theorem 1.6.

Theorem 2.4. *Let H be a graph with v vertices and e edges. There exists a constant $C = C(H)$ such that for all $t = t(n) \in (0, 1/2]$ we have*

$$\mathbb{P} \left(|D_H(G_{n,t}) - \Lambda_H^*(G_{n,t})| > Cbt^{1/2}n^{v-2} \right) \leq \exp(-b) \quad (2.7)$$

for all $3 \log n \leq b \leq t^{1/2}n$. Furthermore,

$$\mathbb{P}\left(|D_H(G_{n,t}) - \Lambda_H^*(G_{n,t})| > Cbn^{v-2}\right) \leq \exp(-b) \quad (2.8)$$

for all $b \geq 3 \log n$.

Remark 2.5. The curious reader may wonder why we express the terms $\mathbb{X}_H(G_i; t)$ of $\Lambda_H^*(G_{n,t})$ as a linear combination of X_\wedge and $X_\Delta - 3sX_\wedge$ rather than directly as a linear combination of X_\wedge and X_Δ . We consider this choice natural because X_\wedge and $X_\Delta - 3sX_\wedge$ are asymptotically orthogonal (i.e., uncorrelated) in the sense that

$$\mathbb{E}\left[X_\wedge(G_i)(X_\Delta(G_i) - 3sX_\wedge(G_i)) \mid G_{i-1}\right]$$

is typically $o(n)$, while their individual variances are $\Theta(n)$. See Section 7 for more details.

In fact, the result holds for all $t \in (0, 1)$. This follows directly from Theorem 1.6 and (5.34).

Theorem 2.6. *Theorem 2.4 holds for $t = t(n) \in (0, 1)$.*

Theorem 2.4 is proved in Section 5 as part of the proof of Theorem 1.6. After proving Theorem 1.6 we easily deduce Theorem 2.6.

2.1. An aside: subgraph counts from subgraph counts in the complement. We record a simple lemma that allows one to relate subgraph counts in G to subgraph counts in the complement G^c .

Lemma 2.7. *Let H and G be graphs, and let G^c be the complement of G , then*

$$N_H(G) = \sum_{H' \subseteq E(H)} (-1)^{e(H')} N_{H'}(G^c), \quad (2.9)$$

where the sum is over all $2^{e(H)}$ subgraphs of H .

Proof. Writing $N_H^{\text{ind}}(G)$ for the number of induced embeddings (isomorphic copies) of H in G we have, by inclusion-exclusion, that

$$N_F^{\text{ind}}(G) = \sum_{F \subseteq H \subseteq K_v} (-1)^{e(H) - e(F)} N_H(G),$$

and, in the other direction,

$$N_H(G) = \sum_{H \subseteq F \subseteq K_v} N_F^{\text{ind}}(G) = \sum_{F^c \subseteq H^c} N_{F^c}^{\text{ind}}(G^c)$$

where K_v is the complete graph on the vertex set of H .

We now have

$$\begin{aligned}
 N_H(G) &= \sum_{F^c \subseteq H^c} N_{F^c}^{\text{ind}}(G^c) \\
 &= \sum_{F^c \subseteq H^c} \sum_{F^c \subseteq H' \subseteq K_v} (-1)^{e(H') - e(F^c)} N_{H'}(G^c) \\
 &= \sum_{H' \subseteq K_v} (-1)^{e(H')} N_{H'}(G^c) \sum_{H \cup (H')^c \subseteq F \subseteq K_v} (-1)^{e(F^c)} \\
 &= \sum_{H' \subseteq H} (-1)^{e(H')} N_{H'}(G^c),
 \end{aligned}$$

where the last line follows since the sum over F in the line above gives 1 if $H \cup (H')^c = K_v$ and 0 otherwise. \square

By linearity, the same identity holds for deviations.

Corollary 2.8. *Let H and G be graphs, and let G^c be the complement of G , then*

$$D_H(G) = \sum_{H' \subseteq E(H)} (-1)^{e(H')} D_{H'}(G^c),$$

where the sum is over all $2^{e(H)}$ subgraphs of H .

Proof. This follows easy from Lemma 2.7 by linearity. Indeed, by taking expectation (with $G \sim G(n, e(G))$) on both sides of (2.9) we obtain that

$$L_H(e(G)) = \sum_{H' \subseteq E(H)} (-1)^{e(H')} L_{H'}(e(G^c)).$$

Subtracting this from (2.9) gives the required identity. \square

3. MARTINGALE DEVIATION INEQUALITIES

In this section we state the Hoeffding-Azuma inequality [1, 20] which bounds the probability of martingale deviations. The particular form of the Hoeffding-Azuma inequality we shall use is stated as Corollary 3.2.

For certain key results, including our main theorem, we need to use an inequality of Freedman [12] instead. Freedman's inequality gives significantly stronger bounds in certain contexts; in particular when the martingale increments, X_i , have conditional second moments, $\mathbb{E}[X_i^2 | \mathcal{F}_{i-1}]$, much smaller than $\|X_i\|_\infty^2$.

We begin with the Hoeffding-Azuma inequality and its corollary. The corollary is an application of the inequality to functions $f(G_m)$, where $G_m \sim G(n, m)$.

Let $(S_n)_{n \geq 0}$ be a martingale with respect to a filtration $(\mathcal{F}_n)_{n \geq 0}$. Write $X_i = S_i - S_{i-1}$, $i \geq 1$ for its *increments* and note that $\mathbb{E}[X_i | \mathcal{F}_{i-1}] = 0$ for all $i \geq 1$.

Lemma 3.1 (Hoeffding-Azuma inequality). *Let $(S_m)_{m=0}^M$ be a martingale with increments $(X_i)_{i=1}^M$, and let $c_i = \|X_i\|_\infty$ for each $1 \leq i \leq M$. Then, for each $a > 0$,*

$$\mathbb{P}(S_M - S_0 > a) \leq \exp\left(\frac{-a^2}{2 \sum_{i=1}^M c_i^2}\right).$$

Furthermore, the same bound holds for $\mathbb{P}(S_M - S_0 < -a)$.

Let us write $\mathcal{G}_{n,m}$ for the family of graphs with n vertices and m edges. One may think of $\mathcal{G}_{n,m}$ as endowed with an edit distance, in which graphs which differ in two edges, G and $G \setminus \{e_i\} \cup \{e_j\}$ for example, have distance 1. It is then natural to say a function $f : \mathcal{G}_{n,m} \rightarrow \mathbb{R}$ is C -Lipschitz, if $|f(G) - f(G')| \leq C$ for all pairs of adjacent graphs G, G' .

Given a function $\psi : E(K_n) \rightarrow \mathbb{R}^+$, let us say that a function $f : \mathcal{G}_{n,m} \rightarrow \mathbb{R}$ is ψ -Lipschitz if for every adjacent pair of graphs $G, G' \in \mathcal{G}_{n,m}$ we have

$$|f(G) - f(G')| \leq \psi(e_i) + \psi(e_j),$$

where $G \triangle G' = \{e_i, e_j\}$.

Corollary 3.2. *Given $\psi : E(K_n) \rightarrow \mathbb{R}^+$ and a ψ -Lipschitz function $f : \mathcal{G}_{n,m} \rightarrow \mathbb{R}$, we have*

$$\mathbb{P}(f(G_m) - \mathbb{E}[f(G_m)] \geq a) \leq \exp\left(\frac{-a^2}{8\|\psi\|_2^2}\right)$$

for all $a \geq 0$, where $\|\psi\|_2^2 := \sum_{e \in E(K_n)} \psi(e)^2$.

Furthermore, the same bound holds for $\mathbb{P}(f(G_m) - \mathbb{E}[f(G_m)] \leq -a)$.

Remark 3.3. While we include a proof of this corollary for completeness, we do not claim that it is an original result. The statement is very close in spirit to that of McDiarmid's concentration inequality [26], although in a slightly different setting, as we do not have independence. See also Warnke [33], where generalisations of McDiarmid's inequality are proved, including one where the independence condition may be weakened.

Proof. Let e_1, \dots, e_N be an ordering of the edges of K_n in which ψ is decreasing. Consider the martingale

$$Z_i = \mathbb{E}[f(G_m) \mid G_m \cap \{e_1, \dots, e_i\}],$$

where the conditioning indicates that we reveal the first i edges in the ordering. Observe that $Z_0 = \mathbb{E}[f(G_m)]$ and $Z_N = f(G_m)$. The result will follow immediately from the Hoeffding-Azuma inequality provided we prove that

$$|Z_i - Z_{i-1}| \leq 2\psi(e_i) \quad \text{almost surely.} \quad (3.1)$$

Let $G_- := G_m \cap \{e_1, \dots, e_{i-1}\}$ and let us set $m_- = |G_-|$ and $m' = m - m_- - 1$. We may generate G_m as follows. Let J be a uniformly random subset of $\{i+1, \dots, N\}$ of cardinality m' , and let k be chosen uniformly in $\{i+1, \dots, N\} \setminus J$. We promise that G_m will be given by either

$$G_- \cup \{e_i\} \cup \{e_j : j \in J\} =: G_{-,i,J} \quad (3.2)$$

if $e_i \in G_m$ or by

$$G_- \cup \{e_k\} \cup \{e_j : j \in J\} =: G_{-,k,J} \quad (3.3)$$

if $e_i \notin G_m$. With this in mind, we have

$$\begin{aligned} & \mathbb{E} [f(G_m) \mid G_m \cap \{e_1, \dots, e_i\}] \\ &= 1_{e_i \in G_m} \mathbb{E}_J [f(G_{-,i,J}) \mid G_-] + 1_{e_i \notin G_m} \mathbb{E}_{J,k} [f(G_{-,k,J}) \mid G_-] \\ &= \mathbb{E}_{J,k} [f(G_{-,k,J}) \mid G_-] + 1_{e_i \in G_m} \mathbb{E}_{J,k} [f(G_{-,i,J}) - f(G_{-,k,J}) \mid G_-] \end{aligned}$$

and

$$\begin{aligned} & \mathbb{E} [f(G_m) \mid G_m \cap \{e_1, \dots, e_{i-1}\}] \\ &= \mathbb{P}(e_i \in G_m \mid G_-) \mathbb{E}_J [f(G_{-,i,J}) \mid G_-] + \mathbb{P}(e_i \notin G_m \mid G_-) \mathbb{E}_{J,k} [f(G_{-,k,J}) \mid G_-] \\ &= \mathbb{E}_{J,k} [f(G_{-,k,J}) \mid G_-] + \mathbb{P}(e_i \in G_m \mid G_-) \mathbb{E}_{J,k} [f(G_{-,i,J}) - f(G_{-,k,J}) \mid G_-]. \end{aligned}$$

It follows that

$$Z_i - Z_{i-1} = \mathbb{E}_{J,k} [f(G_{-,i,J}) - f(G_{-,k,J}) \mid G_-] (1_{e_i \in G_m} - \mathbb{P}(e_i \in G_m \mid G_-)).$$

Since $|1_{e_i \in G_m} - \mathbb{P}(e_i \in G_m \mid G_-)| \leq 1$, we obtain

$$|Z_i - Z_{i-1}| \leq |\mathbb{E}_{J,k} [f(G_{-,i,J}) - f(G_{-,k,J}) \mid G_-]| \leq \mathbb{E}_{J,k} [|f(G_{-,i,J}) - f(G_{-,k,J})| \mid G_-].$$

and since $|f(G_{-,i,J}) - f(G_{-,k,J})| \leq \phi(e_i) + \phi(e_k) \leq 2\phi(e_i)$ we obtain (3.1). \square

We now state Freedman's inequality, and the related converse inequality.

Probabilistic intuition would suggest that deviation probabilities ought to be more closely connected to the second moment of the increments X_i than to $\|X_i\|_\infty$. Freedman's inequality [12] essentially allows us to replace $\|X_i\|_\infty^2$ by $\mathbb{E}[X_i^2 \mid \mathcal{F}_{i-1}]$, the conditional second moment.

Lemma 3.4 (Freedman's inequality). *Let $(S_m)_{m=0}^M$ be a martingale with increments $(X_i)_{i=1}^M$ with respect to a filtration $(\mathcal{F}_m)_{m=0}^M$, let $R \in \mathbb{R}$ be such that $\max_i |X_i| \leq R$ almost surely, and let*

$$V(m) := \sum_{i=1}^m \mathbb{E} [|X_i|^2 \mid \mathcal{F}_{i-1}].$$

Then, for every $\alpha, \beta > 0$, we have

$$\mathbb{P}(S_m - S_0 \geq \alpha \quad \text{and} \quad V(m) \leq \beta \quad \text{for some } m) \leq \exp\left(\frac{-\alpha^2}{2(\beta + R\alpha)}\right).$$

In addition, Freedman [12] proved that this exponent is often close to best possible. Before stating this converse, let us restate the above inequality. Define the stopping time m_α to be the least m such that $S_m > S_0 + \alpha$, and define

$$T_\alpha := V(m_\alpha).$$

The above inequality states that

$$\mathbb{P}(T_\alpha \leq \beta) \leq \exp\left(\frac{-\alpha^2}{2(\beta + R\alpha)}\right).$$

Freedman's converse inequality [12] is as follows.

Lemma 3.5 (Converse Freedman inequality). *Let $(S_m)_{m=0}^M$ be a martingale with increments $(X_i)_{i=1}^M$ with respect to a filtration $(\mathcal{F}_m)_{m=1}^M$, let R be such that $\max_i |X_i| \leq R$ almost surely, and let T_α be as defined above. Then, for every $\alpha, \beta > 0$, we have*

$$\mathbb{P}(T_\alpha \leq \beta) \geq \frac{1}{2} \exp\left(\frac{-\alpha^2(1+4\delta)}{2\beta}\right),$$

where $\delta > 0$ is minimal such that $\beta/\alpha \geq 9R\delta^{-2}$ and $\alpha^2/\beta \geq 16\delta^{-2} \log(64\delta^{-2})$.

From the point of view of our present applications, the essential content of these inequalities is that

$$\mathbb{P}(T_\alpha \leq \beta) = \exp\left(\frac{-\alpha^2(1+o(1))}{2\beta}\right)$$

when $\alpha R \ll \beta \ll \alpha^2$. See Section 3.4 of [26] for other martingale inequalities in a similar spirit.

3.1. A bound for the hypergeometric distribution. The hypergeometric distribution represents the number of successes in a series of draws without replacement. Given N, K, m , a random variable S_m has hypergeometric distribution with parameters N, K, m if $\mathbb{P}(S_m = k) = \binom{K}{k} \binom{N-K}{m-k} / \binom{N}{m}$. If $\mu = \mathbb{E}[S_m] = Km/N$ we have the following bounds on the upper tail:

$$\mathbb{P}(S_m \geq \mu + a) \leq \exp\left(\frac{-a^2}{2\mu + 2a/3}\right) \leq \exp\left(\frac{-a^2}{2\mu + a}\right) \quad (3.4)$$

and the lower tail:

$$\mathbb{P}(S_m \leq \mu - a) \leq \exp\left(\frac{-a^2}{2\mu}\right), \quad (3.5)$$

which were proved in [20]. They also appear in Theorem 2.10 of [15].

4. DEGREES AND CODEGREES IN $G(n, m)$

There are many results on degree sequences of random graphs, for more information see the articles of Bollobás [4], McKay and Wormald [28] and Liebenau and Wormald [22] and the references therein.

We are not aware of a direct reference for the degree and codegree bounds that we require. In this section we prove bounds on the probability of certain events related to degrees and codegrees in the model $G(n, m)$. All the proofs are straightforward applications of Corollary 3.2, a form of the Hoeffding-Azuma inequality.

The three degree deviation results we prove concern the largest degree deviation, the sum of fourth powers of degree deviations and the sum of squares of degree deviations. After stating these results in Section 4.1, we state the analogous codegree results in Section 4.2. We make no effort to optimise the constants in any of these results.

Although it differs from the standard notation, we write $d_u(G)$ for the degree of a vertex u in a graph G . In the case of $G_m \sim G(n, m)$ the expected degree of u is $2m/n$, and so

$$D_u(G_m) := d_u(G_m) - \frac{2m}{n} \quad (4.1)$$

is the deviation of the degree of u from its mean. We shall also consider codegrees, writing $d_{u,w}(G)$ for the number of common neighbours of vertices u and w in a graph G , and

$$D_{u,w}(G_m) := d_{u,w}(G_m) - \frac{(n-2)(m)_2}{(N)_2} \quad (4.2)$$

for the deviation of $d_{u,w}(G_m)$ from its mean.

4.1. Degrees. We prove bounds related to the maximum degree deviation (Lemma 4.1), the sum of fourth powers of degree deviations (Lemma 4.2), and the sum of squares of degree deviations (Lemma 4.3).

We first state the result about the maximum degree deviation. Let

$$D_{\max}(G_m) := \max_u D_u(G_m)$$

and

$$D_{\min}(G_m) := \min_u D_u(G_m).$$

Lemma 4.1. *For all $b \geq \log n$, and all $m \leq N$, we have*

$$\mathbb{P}\left(D_{\max}(G_m) > 4b^{1/2}t^{1/2}n^{1/2} + 4b\right) \leq \exp(-b).$$

Furthermore, the same bound holds for the event $D_{\min}(G_m) < -4b^{1/2}t^{1/2}n^{1/2} - 4b$.

Proof. Fix a vertex $u \in V(G_m)$ and let $a = 4b^{1/2}t^{1/2}n^{1/2} + 4b$. The degree $d_u(G_m)$ has hypergeometric distribution with parameters $N, n-1, m$. By the bound (3.4) on the tail of the hypergeometric distribution we have

$$\mathbb{P}(D_u(G_m) > a) \leq \exp\left(\frac{-a^2}{2tn+a}\right) \leq \exp(-2b).$$

A union bound over the $n \leq \exp(b)$ vertices completes the proof of the main statement. The lower tail bound follows by a similar argument. \square

Our result on the sum of fourth powers of degree deviations is as follows.

Lemma 4.2. *There is a constant $C > 0$ such that, for all $b \geq n^{1/2}$, and all $m \leq N$, we have*

$$\mathbb{P}\left(\sum_{u \in V(G_m)} D_u(G_m)^4 > Cbn^2 \min\{b, n\}\right) \leq \exp(-b).$$

Proof. Fix $b \geq n^{1/2}$. We define a family of events related to degree deviations. For each $j \geq 1$, let

$$a_j := 2^{3-j/2}b^{1/2}n^{1/2} + 2^{3-j/8}n^{5/8}.$$

We may immediately note that

$$a_j^2 \geq 2^{6-j}bn + 2^{6-j/4}n^{5/4} \quad (4.3)$$

and that

$$a_j^4 \leq 2^{15-2j}b^2n^2 + 2^{15-j/2}n^{5/2}. \quad (4.4)$$

For each $j \geq 1$ and for each set $U \subseteq V(G_m)$ of cardinality 2^j we define $E_{j,U}^+$ to be the event that

$$D_u(G_m) > a_j \quad \text{for all } u \in U,$$

and $E_{j,U}^-$ to be the event

$$D_u(G_m) < -a_j \quad \text{for all } u \in U.$$

Claim: $\mathbb{P}(E_{j,U}^+) \leq \exp(-2b - 2^{1+3j/4}n^{1/4})$.

Proof of Claim: Define

$$f(G_m) = \sum_{u \in U} D_u(G_m).$$

We observe that $\mathbb{E}[f(G_m)] = 0$ and that f is ψ -Lipschitz for the function $\psi(e) = |e \cap U|$. We have $\sum_{e \in E(K_n)} \psi(e)^2 \leq 4|U|n = 2^{j+2}n$. We are now ready to apply Corollary 3.2. Noting that $f(G_m) > 2^j a_j$ on the event $E_{j,U}^+$, we have

$$\mathbb{P}(E_{j,U}^+) \leq \mathbb{P}(f(G_m) > 2^j a_j) \leq \exp\left(\frac{-2^{2j}a_j^2}{2^{j+5}n}\right).$$

The claim now follows immediately from (4.3).

Naturally, the same bound holds for $\mathbb{P}(E_{j,U}^-)$.

Now, for each $j \geq 1$, a union bound allows us to bound the probability that any of the events $E_{j,U}^+$ or $E_{j,U}^-$ occurs for any set U of 2^j vertices. Indeed this probability is at most

$$\begin{aligned} 2 \binom{n}{2^j} \exp(-2b - 2^{1+3j/4}n^{1/4}) &\leq \exp(-2b) \exp(1 + 2^j + 2^j \log(n2^{-j}) - 2^{1+3j/4}n^{1/4}) \\ &\leq \exp(-2b), \end{aligned}$$

where we use the fact that $\binom{n}{2^j} \leq (\exp(1)n2^{-j})^{2^j}$ and the final inequality is obtained using the bound $\log x \leq x^{1/4}$ applied with $x = n2^{-j}$. Taking a union bound over $1 \leq j \leq \log_2 n$ there is probability at most

$$\log_2 n \exp(-2b) \leq \exp(-b)$$

that any of the events $E_{j,U}^+$ or $E_{j,U}^-$ occurs. On the complementary event, there is a partition of the vertices into groups V_1, V_2, \dots such that $|V_j| \leq 2^{j+1}$ and $|D_u(G_m)| \leq a_j$ for all $u \in V_j$. Thus, with probability at least $1 - \exp(-b)$, we have

$$\begin{aligned} \sum_{u \in V(G_m)} D_u(G_m)^4 &\leq \sum_{j=1}^{\log_2 n} 2^{j+1} a_j^4 \\ &\leq \sum_{j=1}^{\log_2 n} \left(2^{16-j} b^2 n^2 + 2^{16+j/2} n^{5/2}\right) \\ &\leq C b^2 n^2, \end{aligned}$$

where we have used (4.4) to prove the second inequality, and for the third we have taken $C \geq 2^{19}$ and used that $b \geq n^{1/2}$.

This proves the lemma in the case that $b \leq n$. If $b > n$, then let J be the smallest integer such that $2^J \geq b/n$. It follows that $2^{-J} \leq n/b$. We now argue as above, except using the trivial bound $|D_u(G_m)| \leq n$ for the vertices $u \in V_j$ for $j < J$. Now, with probability at least $1 - \exp(-b)$, we have

$$\begin{aligned} \sum_{u \in V(G_m)} D_u(G_m)^4 &\leq 2^{J+1}n^4 + \sum_{j \geq J} 2^{j+1}a_j^4 \\ &\leq 4bn^3 + \sum_{j \geq J} \left(2^{16-j}b^2n^2 + 2^{16+j/2}n^{5/2} \right) \\ &\leq 4bn^3 + 2^{17-J}b^2n^2 + 2^{16}n^3 \\ &\leq 4bn^3 + 2^{17}bn^3 + 2^{16}n^3 \\ &\leq Cbn^3, \end{aligned}$$

where C has been taken to be at least 2^{18} . This proves the inequality in the case $b \geq n$, completing the proof. \square

Lemma 4.3. *There is a constant C such that for all $b \geq 30$, and all $m \leq N$, we have*

$$\mathbb{P} \left(\sum_{u \in V(G_m)} D_u(G_m)^2 > Cbn^2 \right) \leq \exp(-bn).$$

Proof. Fix $b \geq 30$. Let $\ell = \lfloor \log_2 n \rfloor$. We shall consider a function f_σ for each sequence $\sigma \in \{0, \pm 1, \pm 2, \pm 4, \dots, \pm 2^\ell\}^{V(G_m)}$ defined as follows

$$f_\sigma(G_m) = \sum_{u \in V(G_m)} \sigma_u D_u(G_m).$$

Let us set

$$\|\sigma\|^2 := \sum_{u \in V(G_m)} \sigma_u^2.$$

To see the connection between these functions and the result of the lemma, consider the choice of σ^* defined by setting $\sigma_u^* = 0$ if $|D_u(G_m)| \leq n^{1/2}$, and otherwise defined so that σ_u^* has the same sign as $D_u(G_m)$ and $|\sigma_u^*|$ is the largest power of two such that $|\sigma_u^*|n^{1/2}$ is at most $|D_u(G_m)|$. With this choice of σ^* we have

$$f_{\sigma^*}(G_m) \geq \|\sigma^*\|^2 n^{1/2}. \quad (4.5)$$

Furthermore, if $\sum_u D_u(G_m)^2 > Cbn^2$ with constant $C \geq 129$ then

$$\|\sigma^*\|^2 = \sum_u (\sigma_u^*)^2 \geq \sum_u \frac{D_u(G_m)^2 - n}{4n} > 32bn. \quad (4.6)$$

By (4.5), (4.6), and a union bound, proving the lemma reduces to the problem of proving the following inequality:

$$\sum_{\sigma: \|\sigma\|^2 > 32bn} \mathbb{P} \left(f_\sigma(G_m) > \|\sigma\|^2 n^{1/2} \right) \leq \exp(-bn). \quad (4.7)$$

We first bound this probability for each fixed σ .

Claim: $\mathbb{P} \left(f_\sigma(G_m) > \|\sigma\|^2 n^{1/2} \right) \leq \exp(-\|\sigma\|^2/16)$.

Proof of Claim: The function f_σ is ψ -Lipschitz for the function $\psi(uv) = |\sigma_u| + |\sigma_v|$, for which $\sum_{e \in E(K_n)} \psi(e)^2 \leq 2n \sum_{u \in V(G_m)} \sigma_u^2 = 2n \|\sigma\|^2$. Since $\mathbb{E}[f_\sigma(G_m)] = 0$, it follows from Corollary 3.2 that

$$\mathbb{P}\left(f_\sigma(G_m) > \|\sigma\|^2 n^{1/2}\right) \leq \exp\left(\frac{-\|\sigma\|^4 n}{16n \|\sigma\|^2}\right) \leq \exp(-\|\sigma\|^2/16),$$

as required, completing the proof of the claim.

For σ with $\|\sigma\|^2 > 32bn$ it follows that

$$\mathbb{P}\left(f_\sigma(G_m) > \|\sigma\|^2 n^{1/2}\right) \leq \exp\left(\frac{-\|\sigma\|^2}{16}\right) \leq \exp(-bn) \exp\left(\frac{-\|\sigma\|^2}{32}\right).$$

Substituting this bound into (4.7) we need only prove that

$$\sum_{\sigma: \|\sigma\|^2 > 32bn} \exp\left(\frac{-\|\sigma\|^2}{32}\right) \leq 1. \quad (4.8)$$

We prove this bound by splitting into “types”. Given a sequence $x = (x_{-\ell-1}, \dots, x_{\ell+1})$ we say σ has *type* x if precisely x_0 vertices u have $\sigma_u = 0$, precisely x_j have $\sigma_u = 2^{j-1}$ and precisely x_{-j} have $\sigma_u = -2^{j-1}$ for each $j \in \{1, \dots, \ell+1\}$. Setting

$$S_x := \{\sigma : \sigma \text{ has type } x \text{ and } \|\sigma\|^2 > 32bn\}$$

and observing that there are at most $n^{3\ell} \leq \exp(n)$ choices of σ which have type x , it suffices to prove that

$$\sum_{\sigma \in S_x} \exp\left(\frac{-\|\sigma\|^2}{32}\right) \leq \exp(-n) \quad (4.9)$$

for each type x . Note that all σ of type x have the same $\|\sigma\|^2$, which is given by

$$\varphi(x) := \sum_{j \neq 0} x_j 4^{|j|+1}.$$

It follows that S_x is empty if $\varphi(x) \leq 32bn$. Fix a type x with $\varphi(x) \geq 32bn$, we prove (4.9) for this type x . We must prove that

$$|S_x| \leq \exp\left(\frac{\varphi(x)}{32} - n\right).$$

We have

$$|S_x| = \binom{n}{x_{-\ell-1}, \dots, x_{\ell+1}} \leq \prod_{j \neq 0} \binom{n}{x_j}$$

and, by the well known inequality $\binom{n}{k} \leq (en/k)^k$, we obtain

$$|S_x| \leq \exp\left(\sum_{j \neq 0} x_j \log(en/x_j)\right) \leq \exp\left(e^{1/2} n^{1/2} \sum_{j \neq 0} x_j^{1/2}\right),$$

where we have used the inequality $\log y \leq y^{1/2}$ for $y > 0$.

For each $j \neq 0$, we have

$$|x_j| \leq \min\{n, 4^{1-j} \varphi(x)\}.$$

Using $|x_j| \leq n$ for $|j| \leq 4$ and $|x_j| \leq 4^{1-j}\varphi(x)$ for $|j| \geq 5$, we have

$$|S_x| \leq \exp \left(16n + 2n^{1/2} \sum_{|j| \geq 5} 2^{1-j} \varphi(x)^{1/2} \right).$$

Since $\varphi(x) \geq 32bn$, we obtain

$$|S_x| \leq \exp \left(\frac{\varphi(x)}{2b} + \frac{\varphi(x)}{16b^{1/2}} \right) \leq \exp \left(\frac{\varphi(x)}{32} - n \right),$$

as required, completing the proof. \square

4.2. Codegrees. We now state and prove the analogous results for codegrees.

Recall that $d_{u,w}(G_m)$ denotes the number of common neighbours of vertices u and w in G_m , and

$$D_{u,w}(G_m) := d_{u,w}(G_m) - \frac{(n-2)(m)_2}{(N)_2}$$

is the deviation of $d_{u,w}(G_m)$ from its mean. Let

$$D'_{\max}(G_m) := \max_{u,w} D_{u,w}(G_m)$$

and

$$D'_{\min}(G_m) := \min_{u,w} D_{u,w}(G_m).$$

Lemma 4.4. *For all $b \geq 2 \log n$, and all $m \leq N$, we have*

$$\mathbb{P} \left(D'_{\max}(G_m) > 4b^{1/2}t^{1/2}n^{1/2} + 8b \right) \leq \exp(-b).$$

Furthermore, the same bound holds for the event $D'_{\min}(G_m) < -4b^{1/2}t^{1/2}n^{1/2} - 8b$.

We omit the proof, which is essentially identical to the proof of Lemma 4.1.

The codegree version of Lemma 4.2, on fourth powers of degree deviations, is as follows.

Lemma 4.5. *There is a constant $C > 0$ such that, for all $b \geq n^{1/2}$, and all $m \leq N$, we have*

$$\mathbb{P} \left(\sum_{u,w \in V(G_m)} D_{u,w}(G_m)^4 > Cbn^3 \min\{b, n\} \right) \leq \exp(-b).$$

The proof is very similar to that of Lemma 4.2. One difference is that in place of the events $E_{j,U}^+$ and $E_{j,U}^-$, we consider events of this type inside matchings. This may appear ad hoc, but if we do not make such a restriction the argument runs into problems when we arrive at the union bound.

Proof. Fix $b \geq n^{1/2}$. Let M_1, \dots, M_n be a sequence of matchings which partition $E(K_n)$. We define a family of events related to codegree deviations. For each $j \geq 1$, let

$$a'_j := 2^{4-j/2}b^{1/2}n^{1/2} + 2^{4-j/8}n^{5/8}.$$

We may immediately note that

$$(a'_j)^2 \geq 2^{8-j}bn + 2^{8-j/4}n^{5/4} \tag{4.10}$$

and that

$$(a'_j)^4 \leq 2^{19-2j} b^2 n^2 + 2^{19-j/2} n^{5/2}. \quad (4.11)$$

For each $1 \leq j \leq \log_2 n$ and for each set $U \subseteq M_1$ of cardinality 2^j we define $F_{j,U}^+$ to be the event that

$$D_{u,w}(G_m) > a_j \quad \text{for all } uw \in U,$$

and $F_{j,U}^-$ to be the event

$$D_{u,w}(G_m) < -a_j \quad \text{for all } uw \in U.$$

Claim: $\mathbb{P}(F_{j,U}^+) \leq \exp(-4b - 2^{2+3j/4} n^{1/4})$.

Proof of Claim: Define

$$f(G_m) = \sum_{uw \in U} D_{u,w}(G_m).$$

We observe that $\mathbb{E}[f(G_m)] = 0$ and that f is ψ -Lipschitz for the function $\psi(e) = |e \cap \bigcup U|$, where $\bigcup U$ denotes the set of vertices that occur in an edge of U . We have $\sum \psi(e)^2 \leq 2^{j+3} n$. We are now ready to apply Corollary 3.2. Noting that $f(G_m) > 2^j a'_j$ on the event $F_{j,U}^+$ we have

$$\mathbb{P}(F_{j,U}^+) \leq \mathbb{P}(f(G_m) > 2^j a'_j) \leq \exp\left(\frac{-2^{2j} (a'_j)^2}{2^{j+6} n}\right).$$

The claim now follows immediately from (4.10).

Naturally, the same bound holds for $\mathbb{P}(F_{j,U}^-)$.

Now, for each $j \geq 1$, a union bound allows us to bound the probability that any of the events $F_{j,U}^+$ or $F_{j,U}^-$ occurs for any set U of 2^j pairs of M_1 . Indeed this probability is at most

$$\begin{aligned} 2 \binom{n/2}{2^j} \exp(-4b - 2^{2+3j/4} n^{1/4}) &\leq \exp(-4b) \exp(1 + 2^j + 2^j \log(n2^{-j}) - 2^{2+3j/4} n^{1/4}) \\ &\leq \exp(-4b), \end{aligned}$$

where the final inequality is obtained using the bound $\log x \leq x^{1/4}$ applied with $x = n2^{-j}$.

Taking a union bound over $1 \leq j \leq \log_2 n$ there is probability at most

$$\log(n) \exp(-4b) \leq \exp(-2b)$$

that any of the events $F_{j,U}^+$ or $F_{j,U}^-$ occurs. The above argument also holds inside the remaining matchings M_2, \dots, M_n . Since $n \exp(-2b) \leq \exp(-b)$, we have with probability at least $1 - \exp(-b)$ that in each matching and for each $j \geq 1$, at most 2^j edges uw have $D_{u,w}(G_m) > a_j$ and at most 2^j have $D_{u,w}(G_m) < -a_j$.

In this case, there is a partition of the edges of K_n into groups E_1, \dots such that $|E_j| \leq 2^{j+1} n$ and $|D_{u,w}(G_m)| \leq a_j$ for all $uw \in E_j$. Thus, with probability at least $1 - \exp(-b)$,

we have

$$\begin{aligned}
 \sum_{uw} D_{u,w}^4(G_m) &\leq \sum_{j=1}^{\log_2 n} 2^{j+1} n (a'_j)^4 \\
 &\leq \sum_{j=1}^{\log_2 n} (2^{20-j} b^2 n^3 + 2^{20+j/2} n^{7/2}) \\
 &\leq C b^2 n^3,
 \end{aligned}$$

where we have used (4.11) to prove the second inequality, and taken $C \geq 2^{22}$.

This proves the lemma in the case that $b \leq n$. If $b > n$, then let J be the smallest integer such that $2^J \geq b/n$. It follows that $2^{-J} \leq n/b$. We now argue as above, except using the trivial bound $|D_{u,w}(G_m)| \leq n$ for the pairs $uw \in V_j$ for $j < J$. Now, with probability at least $1 - \exp(-b)$, we have

$$\begin{aligned}
 \sum_{uw} D_{u,w}^4(G_m) &\leq 2^{J+1} n^5 + \sum_{j=J}^{\log_2 n} 2^{j+1} n a_j^4 \\
 &\leq 4bn^4 + \sum_{j=J}^{\log_2 n} (2^{20-j} b^2 n^3 + 2^{20+j/2} n^{7/2}) \\
 &\leq 4bn^4 + 2^{21-J} b^2 n^3 + 2^{20} n^4 \\
 &\leq 4bn^4 + 2^{21} b n^4 + 2^{20} n^4 \\
 &\leq C b n^4,
 \end{aligned}$$

where C has been taken to be at least 2^{22} . This proves the inequality in the case $b \geq n$, completing the proof. \square

Finally, the generalisation of Lemma 4.3 to codegrees is as follows.

Lemma 4.6. *There is a constant C such that for all $b \geq 30$, and all $m \leq N$, we have*

$$\mathbb{P} \left(\sum_{u,w} D_{u,w}(G_i)^2 > C b n^3 \right) \leq \exp(-bn).$$

Proof. We describe how the proof may be obtained from ideas present in the above proofs. As in the proof of Lemma 4.5, let M_1, \dots, M_n be a family of matchings that partition K_n .

Claim: There is a constant C such that

$$\mathbb{P} \left(\sum_{uw \in M_1} D_{u,w}(G_m)^2 > C b n^2 \right) \leq \exp(-2bn).$$

Proof of Claim: In the same way that the proof of Lemma 4.2 was adjusted to bound deviation probabilities for $\sum_{uw \in M_1} D_{uw}(G_m)^4$ in the proof of Lemma 4.5, so Lemma 4.3 may easily be adjusted to prove the claim.

Applying a union bound over the matchings M_1, \dots, M_n one obtains that with probability at least $1 - \exp(-bn)$ we have

$$\sum_{uw \in M_k} D_{u,w}(G_m)^2 \leq Cbn^2$$

for all $k = 1, \dots, n$. In this case

$$\sum_{u,w} D_{u,w}(G_m)^2 = \sum_{k=1}^n \sum_{uw \in M_k} D_{u,w}(G_m)^2 \leq Cbn^3,$$

as required. \square

5. APPROXIMATING THE DEVIATION $D_H(G_{n,t})$ IN TERMS OF $D_\wedge(G_{n,t})$ AND $D_\triangle(G_{n,t})$ – THEOREM 1.6

The main aim of this section is to prove Theorem 1.6, which states that $D_H(G_{n,t})$ is well approximated by a certain linear combination $\Lambda_H(G_{n,t})$ of $D_\wedge(G_{n,t})$ and $D_\triangle(G_{n,t})$. This result will be extremely useful since, for the range of deviations for which it applies, it essentially reduces the study of all subgraph count deviations $D_H(G_{n,t})$ to the cases of two specific graphs, the path of length two and the triangle.

In order to prove Theorem 1.6, we first prove Theorem 2.4, which shows that $D_H(G_{n,t})$ is very well approximated by

$$\begin{aligned} & \Lambda_H^*(G_{n,t}) \\ &= n^{v-3} \sum_{i=1}^m \left(t^{e-2} \binom{H}{\wedge} \frac{(1-t)^2}{(1-s)^2} X_\wedge(G_i) + t^{e-3} \binom{H}{\triangle} \frac{(1-t)^3}{(1-s)^3} (X_\triangle(G_i) - 3sX_\wedge(G_i)) \right), \end{aligned}$$

a sum of terms each of which is a linear combination of $X_\wedge(G_i)$ and $X_\triangle(G_i)$, where $m = \lfloor tN \rfloor$ and $s := i/N$. We deduce Theorem 1.6 (for $t \in (0, 1/2]$) from Theorem 2.4 by showing that $\Lambda_H(G_{n,t})$ is very close to $\Lambda_H^*(G_{n,t})$ deterministically. It is then straightforward to deduce the remaining cases ($t \in (1/2, 1)$) using Corollary 2.8.

Let us now discuss the task of proving Theorem 2.4. Naturally, our proof that $D_H(G_{n,t})$ is well approximated by $\Lambda_H^*(G_{n,t})$ begins with the precise martingale expression for $D_H(G_m)$ (given by Theorem 2.1)

$$D_H(G_m) = \sum_{i=1}^m \sum_{F \subseteq E(H)} \frac{(N-m)_{e(F)} (m-i)_{e-e(F)}}{(N-i)_e} X_F(G_i).$$

In order to show that the precise expression is well approximated by $\Lambda_H^*(G_{n,t})$, we show that each $X_F(G_i)$ can be well approximated by

$$X_F^*(G_i) := n^{v-3} s^{e(F)-2} \left(\binom{F}{\wedge} - 3 \binom{F}{\triangle} \right) X_\wedge(G_i) + n^{v-3} s^{e(F)-3} \binom{F}{\triangle} X_\triangle(G_i). \quad (5.1)$$

This statement is made rigorous in Proposition 5.2.

Definition 5.1. For each graph F we define

$$Y_F(G_i) := X_F(G_i) - X_F^*(G_i). \quad (5.2)$$

We prove that $Y_F(G_i)$ is small, in particular in the L^2 sense. Since it is the graph G_{i-1} that determines the distribution of $Y_F(G_i)|G_{i-1}$, the result will state that it is very unlikely that G_{i-1} is such that $\mathbb{E}[Y_F(G_i)^2|G_{i-1}]$ is large.

Proposition 5.2. *Let F be a graph with $v(F)$ vertices and $e(F)$ edges, and let $t \in (0, 1/2]$. There is a constant $C = C(F)$ such that for all $1 \leq i \leq tN$ and $b \geq 3 \log n$, we have*

$$\mathbb{P}\left(\mathbb{E}[Y_F(G_i)^2 | G_{i-1}] > Cbn^{2v(F)-6}\right) \leq \exp(-b). \quad (5.3)$$

Remark 5.3. The result may be proved for all $t \in (0, 1)$, however for our purposes working for $t \in (0, 1/2]$ is sufficient.

In Section 5.1 we prove Proposition 5.2. We will then be ready to prove Theorem 2.4 in Section 5.2 and Theorem 1.6 in Section 5.3.

5.1. Proof of Proposition 5.2. Fix a graph F with $v(F)$ vertices and $e(F)$ edges. In this subsection we write v for $v(F)$ and e for $e(F)$. The proof of Proposition 5.2 depends on Lemma 5.4 and Lemma 5.5. We shall now motivate and state these two lemmas.

Let e_1, \dots, e_N be the order in which edges are added in the realisation of the Erdős-Rényi random graph process, so that $G_m = \{e_1, \dots, e_m\}$. In particular, in this notation e_i is the edge we add to go from G_{i-1} to G_i . Define $A_F^*(G_i)$, a linear combination involving the degree and codegree deviation of e_i , by

$$\begin{aligned} A_F^*(G_i) &:= 2es^{e-1}n^{v-2} + s^{e-2}n^{v-3} \left(2\binom{F}{\wedge} - 6\binom{F}{\Delta} \right) (D_u(G_i) + D_w(G_i)) \\ &\quad + 6s^{e-3}n^{v-3} \binom{F}{\Delta} D_{u,w}(G_i). \end{aligned}$$

We will prove that $A_F(G_i)$ (which was introduced in (2.1)) is usually well approximated by $A_F^*(G_i)$ (see Lemma 5.5). On the other hand, we prove that $Y_F(G_i)$ may be expressed in terms of the difference $A_F(G_i) - A_F^*(G_i)$.

Lemma 5.4.

$$Y_F(G_i) = (A_F(G_i) - A_F^*(G_i)) - \mathbb{E}[A_F(G_i) - A_F^*(G_i) | G_{i-1}].$$

Proof. This expression for $Y_F(G_i)$ follows almost directly from its definition as $Y_F(G_i) := X_F(G_i) - X_F^*(G_i)$. Indeed, the definition (2.2) of $X_F(G_i)$ is

$$X_F(G_i) = A_F(G_i) - \mathbb{E}[A_F(G_i) | G_{i-1}]$$

and so we need only prove that

$$X_F^*(G_i) = A_F^*(G_i) - \mathbb{E}[A_F^*(G_i) | G_{i-1}]. \quad (5.4)$$

As $X_F^*(G_i)$ is defined (5.1) as a linear combination of $X_\wedge(G_i)$ and $X_\Delta(G_i)$ it is useful to note that

$$\begin{aligned} X_\wedge(G_i) &= A_\wedge(G_i) - \frac{8(i-1)}{n} + \mathbb{E}\left[\frac{8(i-1)}{n} - A_\wedge(G_i) | G_{i-1}\right] \\ &= 2(D_u(G_{i-1}) + D_w(G_{i-1})) - \mathbb{E}[2(D_u(G_{i-1}) + D_w(G_{i-1})) | G_{i-1}] \end{aligned}$$

and

$$X_\Delta(G_i) = 6D_{uw}(G_{i-1}) - \mathbb{E}[6D_{uw}(G_{i-1}) | G_{i-1}],$$

where we have used that

$$\begin{aligned} A_{\wedge}(G_i) &= 2(d_u(G_{i-1}) + d_w(G_{i-1})) \\ &= \frac{8(i-1)}{n} + 2(D_u(G_{i-1}) + D_w(G_{i-1})). \end{aligned} \quad (5.5)$$

and

$$A_{\Delta}(G_i) = \frac{6(n-2)(i-1)_2}{(N)_2} + 6D_{u,w}(G_{i-1}). \quad (5.6)$$

The required equation (5.4) now follow simply by substituting these values in the definition of $X_F^*(G_i)$. \square

We now state Lemma 5.5. We shall use the quantity $\Delta(e_i)$ defined to be the sum of squares of the degree and codegree deviations associated with edge e_i . That is,

$$\Delta(e_i) := D_u(G_{i-1})^2 + D_w(G_{i-1})^2 + D_{uw}(G_{i-1})^2, \quad (5.7)$$

where $e_i = \{u, w\}$.

Lemma 5.5. *Let F be a graph with v vertices and e edges. There is a constant $C = C(F)$ such that, for all $1 \leq i \leq N$ and $b \geq 1$, the event that*

$$|A_F(G_i) - A_F^*(G_i)| > Cb^{1/2}n^{v-3} + Cn^{v-4}\Delta(e_i) \quad (5.8)$$

has probability at most $\exp(-b)$.

Proof. The vertex set of G_m is $[n] = \{1, \dots, n\}$. By symmetry we may assume that the pair 12 is added as the i th edge, i.e., $e_i = 12$. Thus, the event (5.8) may be viewed as an event concerning the first $i-1$ edges e_1, \dots, e_{i-1} . We may reveal this information as follows: we first reveal the neighbourhoods $N_1(G_{i-1})$ and $N_2(G_{i-1})$ of vertices 1 and 2 in G_{i-1} , and then we reveal the remaining edges. We shall prove, for any choice on the first step, of $N_1(G_{i-1})$ and $N_2(G_{i-1})$, that the conditional probability that (5.8) occurs is at most $\exp(-b)$. The result of the lemma then follows by taking expectations.

Let us now fix $N_1 := N_1(G_{i-1})$ and $N_2 := N_2(G_{i-1})$. We set $d_1 = |N_1|$ and $d_2 = |N_2|$. Let us also abbreviate $D_1(G_{i-1})$, $D_2(G_{i-1})$ and $D_{1,2}(G_{i-1})$ to D_1 , D_2 and $D_{1,2}$ respectively, for the duration of the proof.

Our aim is to show that in selecting the remaining $i-1-d_1-d_2$ edges, in $V(G_m) \setminus \{1, 2\}$, there is probability at most $\exp(-b)$ that (5.8) occurs.

The proof will use the triangle inequality, in the sense that we bound $|A_F(G_i) - A_F^*(G_i)|$ by introducing a third quantity $A_F^{**}(G_i)$ such that

$$\mathbb{P}\left(|A_F(G_i) - A_F^{**}(G_i)| > Cb^{1/2}n^{v-3} \mid N_1, N_2, e_i = 12\right) \leq \exp(-b) \quad (5.9)$$

and

$$|A_F^*(G_i) - A_F^{**}(G_i)| \leq Cb^{1/2}n^{v-3} + Cn^{v-4}\Delta(e_i) \quad (5.10)$$

deterministically. We set

$$A_F^{**}(G_i) := \mathbb{E}[A_F(G_i) \mid N_1, N_2, e_i = 12].$$

It is clear (by considering the triangle inequality) that proving the lemma reduces to verifying (5.9) and (5.10).

Let us subdivide $A_F(G_i)$ depending on which edge f of F corresponds to the new edge e_i , and its orientation with respect to $e_i = 12$, which one may think of as oriented $\vec{12}$. That is, we write

$$A_{F,\vec{f}}(G_i)$$

for the number of embeddings $\phi(F)$ of F created with the addition of $e_i = 12$ in which $\phi(\vec{f}) = \vec{12}$. Clearly

$$A_F(G_i) = \sum_{\vec{f}} A_{F,\vec{f}}(G_i), \quad (5.11)$$

where the sum is over orientations \vec{f} of edges $f \in E(F)$. We shall also define $A_{F,\vec{f}}^*(G_i)$ and $A_{F,\vec{f}}^{**}(G_{i-1})$ for each $\vec{f} \in \vec{E}(F)$, as follows. We shall write $\Gamma_1(\vec{f})$ and $\Gamma_2(\vec{f})$ for the neighbourhood in F of the start and end vertex of \vec{f} respectively, and we set

$$\beta(\vec{f}) := |\Gamma_1 \cap \Gamma_2|, \quad \alpha_1(\vec{f}) := |\Gamma_1 \setminus \Gamma_2| \quad \text{and} \quad \alpha_2(\vec{f}) := |\Gamma_2 \setminus \Gamma_1|.$$

We may now define

$$A_{F,\vec{f}}^*(G_i) = s^{e-1}n^{v-2} + s^{e-2}n^{v-3}(\alpha_1(\vec{f})D_1 + \alpha_2(\vec{f})D_2) + s^{e-3}n^{v-3}\beta(\vec{f})D_{1,2}, \quad (5.12)$$

One may easily verify that $\sum_{\vec{f}} \alpha_1(\vec{f}) = \sum_{\vec{f}} \alpha_2(\vec{f}) = 2\binom{F}{\wedge} - 6\binom{F}{\Delta}$ and $\sum_{\vec{f}} \beta(\vec{f}) = 6\binom{F}{\Delta}$, from which it follows that

$$A_F^*(G_i) = \sum_{\vec{f}} A_{F,\vec{f}}^*(G_i). \quad (5.13)$$

We may also define

$$A_{F,\vec{f}}^{**}(G_i) := \mathbb{E} \left[A_{F,\vec{f}}^*(G_i) \mid N_1, N_2, e_i = 12 \right].$$

It follows directly from linearity of expectation that

$$A_F^{**}(G_i) = \sum_{\vec{f}} A_{F,\vec{f}}^{**}(G_i). \quad (5.14)$$

Taken together, equations (5.11), (5.13) and (5.14) reduce the problem of proving (5.9) and (5.10) to the problem of proving, for each $\vec{f} \in \vec{E}(F)$,

$$\mathbb{P} \left(|A_{F,\vec{f}}(G_i) - A_{F,\vec{f}}^{**}(G_i)| > Cb^{1/2}n^{v-3} \mid N_1, N_2, e_i = 12 \right) \leq \exp(-b) \quad (5.15)$$

and

$$|A_{F,\vec{f}}^*(G_i) - A_{F,\vec{f}}^{**}(G_i)| \leq Cn^{v-3} + Cn^{v-4}\Delta(e_i) \quad (5.16)$$

deterministically.

Fix $\vec{f} \in \vec{E}(F)$. Let us write α_1, α_2 and β for $\alpha_1(\vec{f}), \alpha_2(\vec{f})$ and $\beta(\vec{f})$ respectively, and let $\alpha = \alpha_1 + \alpha_2$. Let us first prove (5.16) for this \vec{f} . We shall use the notation $\pm E$ to denote an error of up E . For example, we may express $|N_j|$ as $sn + D_j \pm 1$, for $j \in \{1, 2\}$ and $|N_1 \cap N_2|$ as $s^2n + D_{1,2} \pm 2$.

We begin with a discussion of $A_{F,\vec{f}}^{**}(G_i)$, the expected number of embeddings $\phi(F)$ created with the addition of edge $e_i = 12$ in which $\phi(\vec{f}) = \vec{12}$, given N_1 and N_2 . Let us observe that, writing F' for the graph obtained by removing the vertices of \vec{f} , this is precisely the number of embeddings $\phi(F')$ of F' in

$$G'_i := G_i[V(G_i) \setminus \{1, 2\}]$$

in which

$$\phi(\Gamma_1) \subseteq N_1 \quad \text{and} \quad \phi(\Gamma_2) \subseteq N_2.$$

We may thus calculate that

$$A_{F,\bar{f}}^{**}(G_i) = (|N_1 \cap N_2|)_\beta (|N_1| - \beta)_{\alpha_1} (|N_2| - \beta \pm \alpha_1)_{\alpha_2} (n - 2 - \alpha - \beta)_{v-2-\alpha-\beta} \cdot \frac{(i - 1 - |N_1| - |N_2| + |N_1 \cap N_2|)_{e-1-\alpha-2\beta}}{(N')_{e-1-\alpha-2\beta}},$$

where we have written N' for $\binom{n-2}{2}$. We may now expand each of these term to obtain a main contribution and error terms. For example, we may express $(|N_1 \cap N_2|)_\beta$ as

$$(s^2n + D_{1,2} \pm 2)_\beta = (s^2n + D_{1,2} \pm 2)^\beta \pm \beta^2 n^{\beta-1}.$$

Continuing, and using that $|D_{1,2}| \leq n$, we may express $(|N_1 \cap N_2|)_\beta$ as

$$s^{2\beta} n^\beta + \beta s^{2\beta-2} n^{\beta-1} D_{1,2} \pm 2^\beta n^{\beta-2} D_{1,2}^2 \pm 2(3^\beta + \beta^2) n^{\beta-1}.$$

In particular, there is a constant $C_1 = C_1(F)$, so that

$$(|N_1 \cap N_2|)_\beta = s^{2\beta} n^\beta + \beta s^{2\beta-2} n^{\beta-1} D_{1,2} \pm C_1 (n^{\beta-1} + n^{\beta-2} \Delta(e_i)).$$

We may assume that $C_1 = C_1(F)$ is also chosen so that the equivalent statements hold for the remaining terms. In particular,

$$(|N_j| - \beta \pm \alpha_j)_{\alpha_j} = s^{\alpha_j} n^{\alpha_j} + \alpha_j s^{\alpha_j-1} n^{\alpha_j-1} D_j \pm C_1 (n^{\alpha_j-1} + n^{\alpha_j-2} \Delta(e_i))$$

for $j = 1, 2$,

$$(n - 2 - \alpha - \beta)_{v-2-\alpha-\beta} = n^{v-2-\alpha-\beta} \pm C_1 n^{v-3-\alpha-\beta}$$

and

$$\frac{(i - 1 - |N_1| - |N_2| + |N_1 \cap N_2|)_{e-1-\alpha-2\beta}}{(N')_{e-1-\alpha-2\beta}} = s^{e-1-\alpha-2\beta} \pm C_1 n^{-1}.$$

Replacing C_1 by a larger constant C_2 if necessary, it follows that

$$A_{F,\bar{f}}^{**}(G_i) = s^{e-1} n^{v-2} + s^{e-2} n^{v-3} (\alpha_1 D_1 + \alpha_2 D_2) + s^{e-3} n^{v-3} \beta D_{1,2} \pm C_2 (n^{v-3} + n^{v-4} \Delta(e_i)),$$

completing the proof of (5.16).

All that remains is to prove (5.15). With $e_i = 12$ and the neighbourhoods N_1 and N_2 fixed we have that $A_{F,\bar{f}}^{**}(G_i)$ is a function $f(G)$ of the graph $G = G_{i-1}[V \setminus \{1, 2\}] \sim G(n - 2, i - 1 - d_1 - d_2)$, and by definition, see (5.12), we have $\mathbb{E}[f(G)] = A_{F,\bar{f}}^{**}(G_i)$. Furthermore $f(G)$ is n^{v-4} -Lipschitz, in the sense described in Section 3. By Corollary 3.2, we have that

$$\mathbb{P}\left(|f(G) - \mathbb{E}[f(G)]| > C b^{1/2} n^{v-3}\right) \leq 2 \exp\left(\frac{-C^2 b n^{2v-6}}{8 n^{2v-6}}\right) \leq \exp(-2b)$$

provided we choose $C \geq 5$. This precisely proves (5.15), completing the proof. \square

We now prove Proposition 5.2.

Proof of Proposition 5.2. Let $b \geq 3 \log n$ be fixed. By Lemma 5.5 there is a constant C_1 such that the event

$$|A_F(G_i) - A_F^*(G_i)| > C_1 b^{1/2} n^{v-3} + C_1 n^{v-4} \Delta(e_i) \quad (5.17)$$

has probability at most $\exp(-3b) \leq n^{-2} \exp(-2b)$. We say that G_{i-1} is *b-good* if (5.17) does not occur for any choice of e_i . Since there are fewer than n^2 choices for e_i , it follows that

$$\mathbb{P}(G_{i-1} \text{ is } b\text{-good}) \geq 1 - \exp(-2b).$$

If G_{i-1} is *b-good* then we have that

$$|A_F(G_i) - A_F^*(G_i)| \leq C_1 b^{1/2} n^{v-3} + C_1 n^{v-4} \Delta(e_i)$$

for all possible choices of e_i , and

$$|\mathbb{E}[A_F(G_i) - A_F^*(G_i) \mid G_{i-1}]| \leq C_1 b^{1/2} n^{v-3} + C_1 n^{v-4} \mathbb{E}[\Delta(e_i) \mid G_{i-1}].$$

It follows immediately from Lemma 5.4 that

$$|Y_F(G_i)| \leq 2C_1 b^{1/2} n^{v-3} + C_1 n^{v-4} \Delta(e_i) + C_1 n^{v-4} \mathbb{E}[\Delta(e_i) \mid G_{i-1}] \quad (5.18)$$

whenever G_{i-1} is *b-good*. And, since $(\alpha + \beta + \gamma)^2 \leq 3(\alpha^2 + \beta^2 + \gamma^2)$,

$$Y_F(G_i)^2 \leq 12C_1^2 b n^{2v-6} + 3C_1^2 n^{2v-8} \Delta(e_i)^2 + 3C_1^2 n^{2v-8} \mathbb{E}[\Delta(e_i) \mid G_{i-1}]^2 \quad (5.19)$$

whenever G_{i-1} is *b-good*.

The first term is already in an appropriate form; we now consider the other two terms. Recalling that $\Delta(e_i) = D_u^2(G_{i-1}) + D_w^2(G_{i-1}) + D_{uw}^2(G_{i-1})$, where $e_i = uw$, it follows that

$$\mathbb{E}[\Delta(e_i)^2 \mid G_{i-1}] \leq 3\mathbb{E}[D_u^4(G_{i-1}) + D_w^4(G_{i-1}) + D_{uw}^4(G_{i-1}) \mid G_{i-1}].$$

By the Lemmas 4.2 and 4.5, there is a constant C_2 such that each of the events

$$\sum_{u \in V(G_{i-1})} D_u(G_{i-1})^4 > C_2 n^3 + C_2 b n^2 \min\{b, n\}, \quad (5.20)$$

and

$$\sum_{u, w \in V(G_{i-1})} D_{u,w}(G_{i-1})^4 > C_2 n^4 + C_2 b n^3 \min\{b, n\}, \quad (5.21)$$

has probability at most $\exp(-2b)$. We say the G_{i-1} is *b-great*, if it is *b-good*, and neither of the events (5.20), (5.21) occurs. We have

$$\mathbb{P}(G_{i-1} \text{ is } b\text{-great}) \geq 1 - \exp(-2b) - 2\exp(-2b) \geq 1 - \exp(-b).$$

Finally, if G_{i-1} is *b-great*, then since each vertex has probability at most

$$\frac{n-1}{N-i+1} \leq \frac{n-1}{(1-s)N} = \frac{2}{(1-s)n}$$

of being included in e_i , and each remaining pair has probability at most $3/(1-s)n^2$ of being e_i , we have

$$\begin{aligned} \mathbb{E}[\Delta(e_i)^2 \mid G_{i-1}] &\leq \frac{6}{(1-s)n} \sum_{u \in V(G_{i-1})} D_u(G_{i-1})^4 + \frac{9}{(1-s)n^2} \sum_{u, w \in V(G_{i-1})} D_{u,w}(G_{i-1})^4 \\ &\leq 20C_2 n^2 + 20C_2 b n \min\{b, n\} \\ &\leq C_3 b n^2, \end{aligned}$$

where $C_3 = 40C_2$. Taking conditional expectations in (5.19), and using the bound on $\mathbb{E}[\Delta(e_i)^2 | G_{i-1}]$, we have

$$\mathbb{E}[Y_F(G_i)^2 | G_{i-1}] \leq Cbn^{2v-6},$$

where $C = 12C_1^2(C_3 + 1)$, whenever G_{i-1} is b -great. This completes the proof of the proposition. \square

5.2. Proof of Theorem 2.4. In this section we show how we may deduce Theorem 2.4 from Proposition 5.2.

Let $t \in (0, 1/2]$ and let H be a graph with v vertices and e edges. The main statement of Theorem 2.4 is that there exists a constant $C = C(H)$ such that

$$\mathbb{P}\left(|D_H(G_{n,t}) - \Lambda_H^*(G_{n,t})| > Cbt^{1/2}n^{v-2}\right) \leq \exp(-b) \quad (5.22)$$

for all $3 \log n \leq b \leq t^{1/2}n$, where $\Lambda_H^*(G_{n,t}) = \sum_{i=1}^m \mathbb{X}_H(G_i; t)$ is the sum of the increments $\mathbb{X}_H(G_i; t) = n^{v-3} \left(t^{e-2} \binom{H}{\wedge} \frac{(1-t)^2}{(1-s)^2} X_\wedge(G_i) + t^{e-3} \binom{H}{\Delta} \frac{(1-t)^3}{(1-s)^3} (X_\Delta(G_i) - 3sX_\wedge(G_i)) \right)$.

The proof will use the triangle inequality, bounding the difference between $D_H(G_{n,t})$ and $\Lambda_H^*(G_{n,t})$ via

$$\Lambda_H^{**}(G_{n,t}) := \sum_{i=1}^m \sum_{F \subseteq E(H)} \frac{(1-t)^{e(F)}(t-s)^{e-e(F)}}{(1-s)^e} X_F(G_i), \quad (5.23)$$

where as usual m denotes $\lfloor tN \rfloor$.

Notice that $\Lambda_H^{**}(G_{n,t})$ is close to the martingale expression for $D_H(G_m)$, given by Theorem 2.1, with $m = \lfloor tN \rfloor$, except with coefficients

$$\frac{(1-t)^{e(F)}(t-s)^{e-e(F)}}{(1-s)^e} \quad \text{in place of} \quad \frac{(N-m)_{e(F)}(m-i)_{e-e(F)}}{(N-i)_e}.$$

The following lemma bounds the difference between these coefficients. For fixed constants $0 \leq c \leq e$, and any $1 \leq i \leq m \leq N$,

$$\nu_{c,e}(i, m) := \frac{(N-m)_c(m-i)_{e-c}}{(N-i)_e} - \frac{(1-t)^c(t-s)^{e-c}}{(1-s)^e}.$$

Lemma 5.6. *Let $t \in (0, 1/2]$ and $c, e \in \mathbb{N}$. There is a constant $C = C(c, e)$ such that for all $1 \leq i \leq m \leq tN$, we have*

$$|\nu_{c,e}(i, m)| \leq \frac{C}{n^2}.$$

Proof. We will show that the constant $C = 24e^2$ works for all sufficiently large n . One may then adjust C so that the result holds trivially for all smaller values of n .

Set $k = N - i$ and $\ell = N - m$. We have

$$\nu_{c,e}(i, m) = \frac{(\ell)_c(k-\ell)_{e-c}}{(k)_e} - \frac{\ell^c(k-\ell)^{e-c}}{k^e} = \frac{(\ell)_c(k-\ell)_{e-c}k^{e-1} - \ell^c(k-\ell)^{e-c}(k-1)_{c-1}}{(k-1)_{e-1}k^e}.$$

The numerator of this expression may be written as

$$[(\ell)_c - \ell^c](k-\ell)_{e-c}k^{e-1} + [(k-\ell)_{e-c} - (k-\ell)^{e-c}]\ell^c k^{e-1} + [k^{e-1} - (k-1)_{e-1}]\ell^c(k-\ell)^{e-c}.$$

Since $\ell \leq k$, and the highest order term cancels in each of the square brackets, the numerator has absolute value at most

$$3e^2 k^{2e-2}.$$

On the other hand the denominator is at least

$$k^e (k-1)_{e-1} \geq \frac{1}{2} k^{2e-1}$$

for all sufficiently large n . And so

$$|\nu_{c,e}(i, m)| \leq 6e^2 k^{-1} = \frac{6e^2}{N-m} \leq \frac{12e^2}{(1-t)n^2}.$$

Since $t \in (0, 1/2]$, this complete the proof. \square

We are now nearly ready to prove Theorem 2.4. Before doing so we require one more lemma. We may view $Y_F(G_i)$ as a function of G_{i-1} and e_i . Let us write $\|Y_F|G_{i-1}\|_\infty$ for the maximum possible value of $|Y_F(G_i)|$ over the possible choices $e \in E(K_n) \setminus E(G_{i-1})$ of the i th edge.

Lemma 5.7. *Let F be a graph with $v(F)$ vertices and $e(F)$ edges. There is a constant $C = C(F)$ such that, for all $1 \leq i \leq N$ and all $b \geq 3 \log n$, the event that*

$$\|Y_F|G_{i-1}\|_\infty > Cb^{1/2}n^{v(F)-3} + Cbsn^{v(F)-3} + Cb^{3/2}s^{1/2}n^{v(F)-7/2} + Cb^2n^{v-4}$$

has probability at most $\exp(-b)$.

Proof. With the usage of b -good introduced in the proof of Proposition 5.2, we have that G_{i-1} is b -good with probability at least $1 - \exp(-2b)$ and we recall (5.18), which states that for some constant C_1 we have

$$|Y_F(G_i)| \leq 2C_1b^{1/2}n^{v(F)-3} + C_1n^{v(F)-4}\Delta(e_i) + C_1n^{v(F)-4}\mathbb{E}[\Delta(e_i) | G_{i-1}]$$

whenever G_{i-1} is b -good. Let F_Δ be the event that some $e \in E(K_n) \setminus E(G_{i-1})$ has

$$\Delta(e) > 1000\left(bsn + b^{3/2}s^{1/2}n^{1/2} + b^2\right).$$

It follows easily from Lemma 4.1 and Lemma 4.4 that $\mathbb{P}(F_\Delta) \leq 6\exp(-2b)$.

We may now observe that there exists a constant C such that the event that

$$\|Y_F|G_{i-1}\|_\infty > Cb^{1/2}n^{v(F)-3} + Cbsn^{v(F)-3} + Cb^{3/2}s^{1/2}n^{v(F)-7/2} + Cb^2n^{v(F)-4}$$

is contained in $F_\Delta \cup \{G_{i-1} \text{ is not } b\text{-good}\}$. This probability is at most

$$6\exp(-2b) + \exp(-2b) \leq \exp(-b),$$

as required. \square

Proof of Theorem 2.4. Let $t \in (0, 1/2]$. We shall focus on the proof of the first statement. To deduce the ‘‘Furthermore’’ statement, simply follow the same proof, with the variable t removed, and use $R = 2^{e+1}n^{v-2}$.

Fix $3 \log n \leq b \leq t^{1/2}n$. By the triangle inequality it clearly suffices to prove

$$|D_H(G_{n,t}) - \Lambda_H^{**}(G_{n,t})| \leq C_1tn^{v-2} \tag{5.24}$$

deterministically, and

$$\mathbb{P}\left(|\Lambda_H^{**}(G_{n,t}) - \Lambda_H^*(G_{n,t})| > C_2 b t^{1/2} n^{v-2}\right) \leq \exp(-b) \quad (5.25)$$

for all $3 \log n \leq b \leq t^{1/2} n$, for some constants C_1 and C_2 .

We begin with (5.24). By Theorem 2.1, we have the precise martingale expression for $D_H(G_{n,t})$ given by

$$D_H(G_{n,t}) = \sum_{i=1}^m \sum_{F \subseteq E(H)} \frac{(N-m)_{e(F)} (m-i)_{e-e(F)}}{(N-i)_e} X_F(G_i),$$

where $m = \lfloor tN \rfloor$. It follows that $D_H(G_{n,t}) - \Lambda_H^{**}(G_{n,t})$ is

$$\sum_{i=1}^m \sum_{F \subseteq E(H)} \nu_{e(F),e}(i, m) X_F(G_i).$$

Since each $X_F(G_i)$ is at most n^{v-2} deterministically and $\nu_{e(F),e}(i, m)$ is at most C_3/n^2 , where C_3 is the constant given by Lemma 5.6, it follows that this difference is at most

$$m n^{v-2} \frac{C_1}{n^2} \leq C_1 t n^{v-2}$$

deterministically, where $C_1 = 2^e C_3$.

We now prove (5.25). Let $3 \log n \leq b \leq t^{1/2} n$ be fixed. The proof proceeds by replacing the $X_F(G_i)$ in $\Lambda_H^{**}(G_{n,t})$ by

$$X_F^*(G_i) + Y_F(G_i)$$

where

$$X_F^*(G_i) := n^{v-3} s^{e(F)-2} \left(\binom{F}{\wedge} - 3 \binom{F}{\Delta} \right) X_{\wedge}(G_i) + n^{v-3} s^{e(F)-3} \binom{F}{\Delta} X_{\Delta}(G_i).$$

We claim that the $X_F^*(G_i)$ contribute exactly $\Lambda_H^*(G_{n,t})$, so that:

Claim:

$$\Lambda_H^{**}(G_{n,t}) - \Lambda_H^*(G_{n,t}) = \sum_{i=1}^m \sum_{F \subseteq E(H)} \frac{(1-t)^{e(F)} (t-s)^{e-e(F)}}{(1-s)^e} Y_F(G_i). \quad (5.26)$$

Proof of Claim: We must prove that

$$\sum_{i=1}^m \sum_{F \subseteq E(H)} \frac{(1-t)^{e(F)} (t-s)^{e-e(F)}}{(1-s)^e} X_F^*(G_i) = \Lambda_H^*(G_{n,t}).$$

That is, we must prove that

$$\sum_{F \subseteq E(H)} \frac{(1-t)^{e(F)} (t-s)^{e-e(F)}}{(1-s)^e} X_F^*(G_i) = \mathbb{X}_H(G_i; t),$$

for all $i = 1, \dots, m$. It is clear that both sides are linear combinations of the increments $X_{\wedge}(G_i)$ and $X_{\Delta}(G_i)$, so it suffices to prove they receive the same coefficients on each side. We begin with $X_{\wedge}(G_i)$, which receives coefficient

$$n^{v-3} t^{e-2} \binom{H}{\wedge} \frac{(1-t)^2}{(1-s)^2} - 3 n^{v-3} s t^{e-3} \binom{H}{\Delta} \frac{(1-t)^3}{(1-s)^3} \quad (5.27)$$

on the right hand side, and coefficient (5.28)–3×(5.29) on the left, where (5.28) and (5.29) are given by

$$n^{v-3} \sum_{F \subseteq E(H)} \frac{(1-t)^{e(F)}(t-s)^{e-e(F)}}{(1-s)^e} s^{e(F)-2} \binom{F}{\wedge} \quad (5.28)$$

and

$$n^{v-3} \sum_{F \subseteq E(H)} \frac{(1-t)^{e(F)}(t-s)^{e-e(F)}}{(1-s)^e} s^{e(F)-2} \binom{F}{\triangle}. \quad (5.29)$$

Since there is a contribution to (5.28) for each copy of P_2 contained in the subgraph $F \subseteq E(H)$, one may sum first over copies of P_2 contained in H , with each having a contribution equal to the total contribution of subgraphs $F \subseteq E(H)$ which contain it. Thus (5.28) is $n^{v-3} \binom{H}{\wedge}$ times

$$\begin{aligned} & \sum_{P \subseteq F \subseteq E(H)} \frac{(1-t)^{e(F)}(t-s)^{e-e(F)}}{(1-s)^e} s^{e(F)-2} \\ &= \frac{(1-t)^2(t-s)^{e-2}}{(1-s)^e} \sum_{F' \subseteq E(H) \setminus P} \left(\frac{s(1-t)}{t-s} \right)^{e(F')}, \end{aligned}$$

where P is some copy of P_2 of H . Summing, using the binomial identity, reveals that (5.28) is precisely

$$n^{v-3} \binom{H}{\wedge} \frac{t^{e-2}(1-t)^2}{(1-s)^2}.$$

Similarly, (5.29) is $n^{v-3} \binom{H}{\triangle}$ times

$$\begin{aligned} & \sum_{T \subseteq F \subseteq E(H)} \frac{(1-t)^{e(F)}(t-s)^{e-e(F)}}{(1-s)^e} s^{e(F)-2} \\ &= \frac{s(1-t)^3(t-s)^{e-3}}{(1-s)^e} \sum_{F' \subseteq E(H) \setminus T} \left(\frac{s(1-t)}{t-s} \right)^{e(F')}, \end{aligned}$$

where T is some triangle of H . By the binomial identity, we find that (5.29) is precisely

$$n^{v-3} \binom{H}{\triangle} \frac{st^{e-3}(1-t)^3}{(1-s)^3}.$$

The coefficient on the left, (5.28)–3×(5.29), is equal to that on the right, (5.27).

Similar calculations confirm that $X_{\triangle}(G_i)$ receives coefficient

$$n^{v-3} t^{e-3} \binom{H}{\triangle} \frac{(1-t)^3}{(1-s)^3}$$

on both sides, completing the proof of the Claim.

Now to complete the proof of (5.25), it suffices to prove, for some constant C_2 , that

$$\sum_{i=1}^m \sum_{F \subseteq E(H)} \frac{(1-t)^{e(F)}(t-s)^{e-e(F)}}{(1-s)^e} Y_F(G_i) \quad (5.30)$$

is at most $C_2 b t^{1/2} n^{v-2}$, in absolute value, with probability at least $1 - \exp(-b)$.

Let C_4 be a constant $3e(H)$ times larger than the largest constant required by Proposition 5.2 for a subgraph $F \subseteq E(H)$ and C_5 the equivalent for Lemma 5.7. With these choices we have

$$\mathbb{P}(\mathbb{E}[Y_F(G_i)^2 | G_{i-1}] > C_4 b n^{2v-6}) \leq \exp(-3e(H)b) \leq n^{-2} 2^{-e(H)} \exp(-2b) \quad (5.31)$$

for all $F \subseteq E(H)$. Let $E_H(m)$ be the event that for some $F \subseteq E(H)$ and $1 \leq i \leq m$ either

$$\mathbb{E}[Y_F(G_i)^2 | G_{i-1}] > C_4 b n^{2v-6}$$

or

$$\|Y_F | G_{i-1}\|_\infty > C_5 b^{1/2} n^{v-3} + C_5 b t n^{v-3} + C_5 b^{3/2} t^{1/2} n^{v-7/2} + C_5 b^2 n^{v-4}$$

occurs. By (5.31), Lemma 5.7 and a union bound, we have

$$\mathbb{P}(E_H(m)) \leq \exp(-2b).$$

Let us define

$$Y_F^*(G_i) = Y_F(G_i) 1_{E_H(i)^c}.$$

We observe that the $Y_F^*(i)$ are also martingale increments, in the sense that

$$\mathbb{E}[Y_F^*(G_i) | G_{i-1}] = 0.$$

We observe further that they satisfy

$$\mathbb{E}[Y_F^*(G_i)^2 | G_{i-1}] \leq C_4 b n^{2v-6}$$

and

$$|Y_F^*(G_i)| \leq C_5 b^{1/2} n^{v-3} + C_5 b t n^{v-3} + C_5 b^{3/2} t^{1/2} n^{v-7/2} + C_5 b^2 n^{v-4},$$

almost surely, and

$$\sum_{i=1}^m \sum_{F \subseteq E(H)} \frac{(1-t)^{e(F)} (t-s)^{e-e(F)}}{(1-s)^e} Y_F^*(G_i) \quad (5.32)$$

is equal to (5.30) on $\Omega \setminus E_H^Y(m)$.

We bound the probability that (5.32) is large using Freedman's inequality, applied to the martingale (5.32), with increments

$$\sum_{F \subseteq E(H)} \frac{(1-t)^{e(F)} (t-s)^{e-e(F)}}{(1-s)^e} Y_F^*(G_i).$$

Furthermore, since the coefficients are all at most 1, we have

$$\mathbb{E} \left[\left(\sum_{F \subseteq E(H)} \frac{(1-t)^{e(F)} (t-s)^{e-e(F)}}{(1-s)^e} Y_F^*(G_i) \right)^2 \middle| G_{i-1} \right] \leq 4^{e(H)} C_4 b n^{2v-6}$$

almost surely. We now apply Freedman's inequality, Lemma 3.4, to (5.32), with $\alpha = C_2 b t^{1/2} n^{v-2}$,

$$\beta = 4^{e(H)} C_4 b m n^{2v-6} \leq 4^{e(H)} C_4 b t n^{2v-4}$$

and

$$\begin{aligned} R &= C_6 b^{1/2} n^{v-3} + C_6 b t n^{v-3} + C_6 b^{3/2} t^{1/2} n^{v-7/2} + C_6 b^2 n^{v-4} \\ &\leq 4 C_6 t^{1/2} n^{v-2}, \end{aligned}$$

where the inequality relies on the condition $b \leq t^{1/2}n$, and where we have chosen $C_6 = 2^{e(H)}C_5$. We obtain that the probability that (5.32) exceeds $C_2bt^{1/2}n^{v-2}$ in absolute value is at most

$$\exp\left(\frac{-C_2^2b^2tn^{2v-4}}{4^{e(H)}C_4bt n^{2v-4} + 8C_2C_6bt n^{2v-4}}\right) \leq \exp(-b),$$

provided $C_2 \geq 2^{e(H)+1}C_4C_6$, completing the proof of the theorem. \square

5.3. Proof of Theorem 1.6. We now show how we may deduce Theorem 1.6 from Theorem 2.4. The main statement of Theorem 1.6 is that

$$\mathbb{P}\left(|D_H(G_{n,t}) - \Lambda_H(G_{n,t})| > Cbt^{1/2}n^{v-2}\right) \leq \exp(-b)$$

for some constant $C = C(H)$, and for all $3 \log n \leq b \leq t^{1/2}n$, where

$$\Lambda_H(G_{n,t}) := n^{v-3}t^{e-2} \left(\binom{H}{\wedge} - 3 \binom{H}{\Delta} \right) D_{\wedge}(G_{n,t}) + n^{v-3}t^{e-3} \binom{H}{\Delta} D_{\Delta}(G_{n,t}). \quad (5.33)$$

We shall use the triangle inequality to control the difference between $D_H(G_{n,t})$ and $\Lambda_H(G_{n,t})$ via $\Lambda_H^*(G_{n,t})$.

We prove Theorem 1.6 first for $t \in (0, 1/2]$, and then show how we may deduce the result for $t \in (1/2, 1)$.

Proof of Theorem 1.6 for $t \in (0, 1/2]$. Let $t \in (0, 1/2]$. We shall focus on proof of the main statement. To deduce the ‘‘Furthermore’’ statement simply follow the same proof, with the variable t removed, and use the ‘‘Furthermore’’ part of Theorem 2.4.

Let $3 \log n \leq b \leq t^{1/2}n$ be fixed. By Theorem 2.4, there is a constant $C = C(H)$, such that

$$\mathbb{P}\left(|D_H(G_{n,t}) - \Lambda_H^*(G_{n,t})| > Cbt^{1/2}n^{v-2}\right) \leq \exp(-b)$$

for all $3 \log n \leq b \leq t^{1/2}n$. So, by the triangle inequality, it suffices to prove the bound

$$|\Lambda_H^*(G_{n,t}) - \Lambda_H(G_{n,t})| \leq C'tn^{v-2} \quad (5.34)$$

deterministically, for some constant $C' = C'(H)$. Using Theorem 2.1, the precise martingale expression for $D_H(G_{n,t})$, to expand $D_{\wedge}(G_{n,t})$ and $D_{\Delta}(G_{n,t})$ in terms of $X_{\wedge}(G_i)$ and $X_{\Delta}(G_i)$ we may express $\Lambda_H(G_{n,t})$ as a sum of the form³

$$\sum_{i=1}^m (\alpha(i, m)X_{\wedge}(G_i) + \beta(i, m)X_{\Delta}(G_i)).$$

On the other hand $\Lambda_H^*(G_{n,t})$ is already of the form

$$\sum_{i=1}^m (\alpha'(i, m)X_{\wedge}(G_i) + \beta'(i, m)X_{\Delta}(G_i)).$$

We shall prove that

$$\alpha(i, m) - \alpha'(i, m) = n^{v-3}t^{e-2} \left(\binom{H}{\wedge} - 3 \binom{H}{\Delta} \right) \nu_{2,2}(i, m) + 3n^{v-3}t^{e-3} \binom{H}{\Delta} \nu_{2,3}(i, m) \quad (5.35)$$

³One does not need to include terms $X_F(G_i)$ for graphs F with $e(F) \leq 1$, as $X_F(G_i) = 0$ in all such cases.

and

$$\beta(i, m) - \beta'(i, m) = n^{v-3}t^{e-3}\nu_{3,3}(i, m). \quad (5.36)$$

By Lemma 5.6, the ν values are $O(n^{-2})$, and so, based on (5.35) and (5.36) these differences are $O(n^{v-5})$. Since each of $X_{\wedge}(G_i)$ and $X_{\Delta}(G_i)$ has absolute value at most n deterministically, and the sums each have $m \leq tN$ terms, this completes the proof of (5.34), and therefore the whole proof.

All that remains is to verify (5.35) and (5.36). We observe that (5.36) follows immediately from the definitions. In order to prove (5.35), let us calculate

$$\alpha(i, m) - n^{v-3}t^{e-2} \left(\binom{H}{\wedge} - 3 \binom{H}{\Delta} \right) \nu_{2,2}(i, m) - 3n^{v-3}t^{e-3} \binom{H}{\Delta} \nu_{2,3}(i, m). \quad (5.37)$$

We have that

$$\alpha(i, m) = n^{v-3}t^{e-2} \left(\binom{H}{\wedge} - 3 \binom{H}{\Delta} \right) \frac{(N-m)_2}{(N-i)_2} + 3n^{v-3}t^{e-3} \binom{H}{\Delta} \frac{(N-m)_2(m-i)}{(N-i)_3}.$$

And so, using the definition of $\nu_{c,e}(i, m)$, we have that (5.37) equals

$$n^{v-3}t^{e-2} \left(\binom{H}{\wedge} - 3 \binom{H}{\Delta} \right) \frac{(1-t)^2}{(1-s)^2} + 3n^{v-3}t^{e-3} \binom{H}{\Delta} \frac{(1-t)^2(t-s)}{(1-s)^3}.$$

Cancelling, we obtain

$$n^{v-3}t^{e-2} \binom{H}{\wedge} \frac{(1-t)^2}{(1-s)^2} - 3n^{v-3}st^{e-3} \binom{H}{\Delta} \frac{(1-t)^3}{(1-s)^3}$$

which is precisely $\alpha'(i, m)$, completing the verification of (5.35) and therefore the proof. \square

We have now proved Theorem 1.6 for $t \in (0, 1/2]$. We deduce the cases $t \in (1/2, 1)$ by considering the complementary graph and using Corollary 2.8.

Proof of Theorem 1.6 for $t \in (1/2, 1)$. For this range of t it suffices to prove the ‘‘Furthermore’’ statement. Indeed, up to a change of the constant this implies the main statement. Fix $t \in (1/2, 1]$ and $b \geq 3 \log n$. Let $t' = 1 - t \in (0, 1/2]$. Let C'' be the maximum over subgraphs $F \subseteq E(H)$ of the constant obtained by the proof of Theorem 1.6 in the case $t \in (0, 1/2]$, let $C' = eC''$ and $C = 2^e C'$. By Theorem 1.6 for $t \in (0, 1/2]$ we have

$$\mathbb{P}(|D_F(G_{n,t'}) - \Lambda_F(G_{n,t'})| > C'bn^{v-2}) \leq \exp(-eb) \leq 2^{-e} \exp(-b).$$

Thus, by a union bound there is probability at least $1 - \exp(-b)$ that

$$|D_F(G_{n,t'}) - \Lambda_F(G_{n,t'})| \leq C'bn^{v-2} \quad (5.38)$$

for all $F \subseteq E(H)$. We complete the proof by showing that if (5.38) holds for $G_{n,t'} = G_{n,t}^c$, then

$$|D_H(G_{n,t}) - \Lambda_H(G_{n,t})| \leq Cbn^{v-2}. \quad (5.39)$$

(It is elementary that the complement of $G_{n,t}$ is distributed as $G_{n,t'}$.)

We now prove (5.39) which will complete the proof of the theorem. We shall use Corollary 2.8 which allows us to relate subgraph count deviations to those in the complement. By Corollary 2.8 we have that

$$D_H(G_{n,t}) = \sum_{F \subseteq E(H)} (-1)^{e(F)} D_F(G_{n,t}^c).$$

If (5.38) holds in $G_{n,t}^c$ (which has the same distribution as $G_{n,t'}$) then

$$D_H(G_{n,t}) = \sum_{F \subseteq E(H)} (-1)^{e(F)} \Lambda_F(G_{n,t'}) \pm Cbn^{v-2}.$$

We claim that the main sum

$$\sum_{F \subseteq E(H)} (-1)^{e(F)} \Lambda_F(G_{n,t'}) \tag{5.40}$$

is equal to $\Lambda_H(G_{n,t})$. Clearly proving this fact will complete the proof.

By the definition of $\Lambda_F(G_{n,t'})$, see (5.33), we can rewrite (5.40) as (5.41) + (5.42), defined by:

$$n^{v-3} \sum_{F \subseteq E(H)} (-1)^{e(F)} (t')^{e(F)-2} \left(\binom{F}{\wedge} - 3 \binom{F}{\triangle} \right) D_{\wedge}(G_{n,t'}) \tag{5.41}$$

and

$$n^{v-3} \sum_{F \subseteq E(H)} (-1)^{e(F)} (t')^{e(F)-3} \binom{F}{\triangle} D_{\triangle}(G_{n,t'}). \tag{5.42}$$

Summing over P_2 s and triangles of H and using the binomial identity, as in the proof of Theorem 2.4, we obtain that (5.41) is equal to

$$n^{v-3} t^{e-2} \binom{H}{\wedge} D_{\wedge}(G_{n,t'}) + 3n^{v-3} t' t^{e-3} \binom{H}{\triangle} D_{\wedge}(G_{n,t'}), \tag{5.43}$$

while (5.42) is equal to

$$-n^{v-3} t^{e-3} \binom{H}{\triangle} D_{\triangle}(G_{n,t'}). \tag{5.44}$$

Again using Corollary 2.8, and using the fact that we are taking $G_{n,t'}$ to be the complement of $G_{n,t}$, we have

$$D_{\wedge}(G_{n,t'}) = D_{\wedge}(G_{n,t}) \quad \text{and} \quad D_{\triangle}(G_{n,t'}) = -D_{\triangle}(G_{n,t}) + 3D_{\wedge}(G_{n,t}).$$

Substituting these values in (5.43) and (5.44), we obtain that (5.40) is

$$n^{v-3} t^{e-2} \left(\binom{H}{\wedge} - 3 \binom{H}{\triangle} \right) D_{\wedge}(G_{n,t}) + n^{v-3} t^{e-3} \binom{H}{\triangle} D_{\triangle}(G_{n,t}).$$

This proves that (5.40) is equal to $\Lambda_H(G_{n,t})$, and therefore completes the proof. \square

5.4. Deducing Theorem 2.6. We recall that Theorem 2.4 is stated for $t = t(n) \in (0, 1/2)$. Having proved Theorem 1.6 we may now deduce that Theorem 2.4 applies for $t \in (0, 1)$. This was stated as Theorem 2.6.

Proof of Theorem 2.6. In proving (5.34) above, we established that

$$|\Lambda_H^*(G_{n,t}) - \Lambda_H(G_{n,t})| \leq C' t n^{v-2}$$

deterministically. It is now immediate by observation that Theorem 2.6 follows from Theorem 1.6 and the triangle inequality. \square

6. A GENERAL BOUND ON DEVIATION PROBABILITIES – THEOREM 1.7

In this section we prove Theorem 1.7. We recall that this theorem gives a weaker bound on subgraph count deviations than Theorem 1.1. However, it applies across the whole range of possible deviations and gives an exponent which is best possible up to multiplication by constant.

Our proof will rely on using Theorem 2.6, which states that $D_H(G_{n,t})$ is well approximated by $\Lambda_H^*(G_{n,t})$, and the following proposition.

Proposition 6.1. *Let H be a graph with v vertices and e edges. There is a constant $c = c(H) > 0$ such that for all $t \in (0, 1/2]$, and all $\eta, n \geq c^{-1}$, we have*

$$\mathbb{P}\left(|\Lambda_H^*(G_{n,t})| > \eta n^{v-3/2}\right) \leq \exp(-c\eta \min\{\eta, n^{1/2}\}).$$

Let us first show that Theorem 1.7 follows easily from these results.

Proof of Theorem 1.7. We prove the result for $t \in (0, 1/2]$; up to changing the constant, the result then follows for $t \in (1/2, 1)$ by Corollary 2.8. Now suppose that $t \in (0, 1/2]$, has been fixed. Let us also fix the graph H with v vertices and e edges.

For all α, n , we have, by the triangle inequality,

$$\mathbb{P}\left(|D_H(G_{n,t})| > \alpha n^{v-3/2}\right) \leq \mathbb{P}\left(|\Lambda_H^*(G_{n,t})| > \frac{\alpha}{2} n^{v-3/2}\right) + \mathbb{P}(F(\alpha/2)), \quad (6.1)$$

where $F(b)$ is the event that

$$|D_H(G_{n,t}) - \Lambda_H^*(G_{n,t})| > bn^{v-3/2}.$$

We bound the first probability by applying Proposition 6.1. We obtain

$$\mathbb{P}\left(|\Lambda_H^*(G_{n,t})| > \frac{\alpha}{2} n^{v-3/2}\right) \leq \exp(-c'\alpha \min\{\alpha, n^{1/2}\}),$$

where c' is a quarter of the constant of that proposition.

We bound the second probability using Theorem 2.6. Let $C = C(H)$, be the constant given by Theorem 2.6. By Theorem 2.6, we have that

$$\mathbb{P}(F(\alpha/2)) \leq \exp\left(\frac{-\alpha n^{1/2}}{2C}\right).$$

Substituting these bounds into (6.1), we obtain

$$\begin{aligned} \mathbb{P}\left(|D_H(G_{n,t})| > \alpha n^{v-3/2}\right) &\leq \exp(-c'\alpha \min\{\alpha, n^{1/2}\}) + \exp\left(\frac{-\alpha n^{1/2}}{2C}\right) \\ &\leq \exp(-c\alpha \min\{\alpha, n^{1/2}\}), \end{aligned}$$

where c is taken to be at most $\min\{c'/2, 1/4C\}$. □

All that remains to complete the section is to prove Proposition 6.1. We require the following lemma.

Lemma 6.2. *There is a constant C such that for all $1 \leq i \leq N/2$, and all $\eta \geq 1$ there is probability at least $1 - \exp(-\eta n^{1/2})$ that*

$$\mathbb{E} [X_\wedge(G_i)^2 \mid G_{i-1}] \leq C n^{1/2} \max\{\eta, n^{1/2}\} \quad (6.2)$$

and

$$\mathbb{E} [X_\Delta(G_i)^2 \mid G_{i-1}] \leq C n^{1/2} \max\{\eta, n^{1/2}\}. \quad (6.3)$$

Proof. Let C' be twice the larger of the constants given by Lemma 4.3 and Lemma 4.6. Let E_1 be the event that

$$\sum_u D_u(G_{i-1})^2 > C' n^{3/2} \max\{\eta, n^{1/2}\}$$

and E_2 the event that

$$\sum_{u,w} D_{u,w}(G_{i-1})^2 > C' n^{5/2} \max\{\eta, n^{1/2}\}.$$

By considering the cases $\eta \leq n^{1/2}$ and $\eta > n^{1/2}$, it follows from Lemma 4.3 that

$$\mathbb{P}(E_1) \leq \exp(-2n) 1_{\eta \leq n^{1/2}} + \exp(-2\eta n^{1/2}) 1_{\eta > n^{1/2}} \leq \exp(-2\eta n^{1/2}).$$

Using Lemma 4.6 one may obtain the same bound on $\mathbb{P}(E_2)$, so that

$$\mathbb{P}(E_1 \cup E_2) \leq 2 \exp(-2\eta n^{1/2}) \leq \exp(-\eta n^{1/2}).$$

It therefore suffices to prove that the event that (6.2) fails is contained in E_1 , and the event that (6.3) fails is contained in E_2 .

Let us now find a bound on $\mathbb{E} [X_\wedge(G_i)^2 \mid G_{i-1}]$ which will show that (6.2) holds in E_1^c . We recall that $X_\wedge(G_i)$ is defined by $X_\wedge(G_i) = A_\wedge(G_i) - \mathbb{E} [A_\wedge(G_i) \mid G_{i-1}]$, and so

$$\begin{aligned} \mathbb{E} [X_\wedge(G_i)^2 \mid G_{i-1}] &= \text{Var}(A_\wedge(G_i) \mid G_{i-1}) \\ &\leq \mathbb{E} \left[\left(A_\wedge(G_i) - \frac{8(i-1)}{n} \right)^2 \mid G_{i-1} \right]. \end{aligned}$$

Now, we recall from (5.5) that $A_\wedge(G_i) = 8(i-1)/n + 2(D_u(G_{i-1}) + D_w(G_{i-1}))$ where uw is the i th edge. It follows that, on the event E_1^c ,

$$\begin{aligned} \mathbb{E} [X_\wedge(G_i)^2 \mid G_{i-1}] &\leq \frac{1}{N-i+1} \sum_{uw \notin E(G_{i-1})} 4(D_u(G_{i-1}) + D_w(G_{i-1}))^2 \\ &\leq \frac{8}{N-i+1} \sum_{uw \notin E(G_{i-1})} (D_u(G_{i-1})^2 + D_w(G_{i-1})^2) \\ &\leq \frac{16(n-1)}{N} \sum_u D_u(G_{i-1})^2 \\ &\leq 32C' n^{1/2} \max\{\eta, n^{1/2}\}, \end{aligned}$$

For $C \geq 32C'$ it follows that the event

$$\mathbb{E} [X_\wedge(G_i)^2 \mid G_{i-1}] > C n^{1/2} \max\{\eta, n^{1/2}\}$$

is contained in E_1 , as required.

We recall from (5.6) that $A_\Delta(G_i) = 6(n-2)(i-1)_2/(N)_2 + 6D_{u,w}(G_{i-1})$. A calculation as above, using that

$$\begin{aligned} \mathbb{E} [X_\Delta(G_i)^2 | G_{i-1}] &= \text{Var}(A_\Delta(G_i) | G_{i-1}) \\ &\leq \mathbb{E} \left[\left(A_\Delta(G_i) - \frac{6(n-2)(i-1)_2}{(N)_2} \right)^2 \middle| G_{i-1} \right] \\ &= \frac{1}{N-i+1} \sum_{uw \notin E(G_{i-1})} (6D_{u,w}(G_{i-1}))^2, \end{aligned}$$

shows that the event (6.3) fails is contained in E_2 , provided $C \geq 160C'$, completing the proof. \square

We now present a proof of Proposition 6.1

Proof of Proposition 6.1. The proof is obtained by an application of Freedman's inequality, Lemma 3.4, to

$$\Lambda_H^*(G_{n,t}) = \sum_{i=1}^m \mathbb{X}_H(G_i; t)$$

where $m = \lfloor tN \rfloor$ and

$$\mathbb{X}_H(G_i; t) = n^{v-3} t^{e-3} \left(t \binom{H}{\wedge} \frac{(1-t)^2}{(1-s)^2} X_\wedge(G_i) + \binom{H}{\Delta} \frac{(1-t)^3}{(1-s)^3} (X_\Delta(G_i) - 3sX_\wedge(G_i)) \right).$$

Since the co-efficients of $X_\wedge(G_i)$ and $X_\Delta(G_i)$ are at most Cn^{v-3} in absolute value, for some constant $C = C(H)$, we have

$$\mathbb{E} [\mathbb{X}_H(G_i; t)^2 | G_{i-1}] \leq 2C^2 n^{2v-6} \mathbb{E} [X_\wedge(G_i)^2 | G_{i-1}] + 2C^2 n^{2v-6} \mathbb{E} [X_\Delta(G_i)^2 | G_{i-1}].$$

Writing $E_{\text{var}}(i-1)$ for the event that

$$\mathbb{E} [\mathbb{X}_H(G_i; t)^2 | G_{i-1}] > 4C_1 C^2 n^{2v-11/2} \max\{\eta, n^{1/2}\}, \quad (6.4)$$

where C_1 is taken to be the constant of Lemma 6.2, it follows, from Lemma 6.2, that $\mathbb{P}(E_{\text{var}}(i-1)) \leq \exp(-\eta n^{1/2})$. Now let E_{var} be the event

$$\sum_{i=1}^m \mathbb{E} [\mathbb{X}_H(G_i; t)^2 | G_{i-1}] > 4C_2^3 m n^{2v-11/2} \max\{\eta, n^{1/2}\},$$

where $C_2 = \max\{C, C_1\}$. By a union bound, we have

$$\mathbb{P}(E_{\text{var}}) \leq n^2 \exp(-\eta n^{1/2}) \leq \exp(-\eta n^{1/2}/2).$$

One may also note that, since $|X_\wedge(G_i)|, |X_\Delta(G_i)| \leq n$, we have

$$|\mathbb{X}_H(G_i; t)| \leq 2Cn^{v-2} \quad \text{almost surely.}$$

We now apply Freedman's inequality, Lemma 3.4, with $\alpha = \eta t^{1/2} n^{v-3/2}$, with $\beta = 4C_2^3 t n^{2v-7/2} \max\{\eta, n^{1/2}\}$ and with $R = 2Cn^{v-2}$. We obtain

$$\begin{aligned} & \mathbb{P}\left(\Lambda_H^*(G_{n,t}) > \eta n^{v-3/2}\right) \\ & \leq \exp\left(\frac{-\eta^2 t n^{2v-3}}{8C_2^3 t n^{2v-7/2} \max\{\eta, n^{1/2}\} + 4C\eta n^{2v-7/2}}\right) + \mathbb{P}(E_{\text{var}}) \\ & \leq \exp\left(\frac{-\eta \min\{\eta, n^{1/2}\}}{12C_2^3}\right) + \exp(-\eta n^{1/2}/2) \\ & \leq 2 \exp(-4c\eta \min\{\eta, n^{1/2}\}), \end{aligned}$$

where c was chosen to be at most $1/24C_2^3$. Since an identical argument applies to bound the probability that $\Lambda_H^*(G_{n,t}) < -\eta n^{v-3/2}$, we have

$$\mathbb{P}\left(|\Lambda_H^*(G_{n,t})| > \eta n^{v-3/2}\right) \leq 4 \exp(-4c\eta \min\{\eta, n^{1/2}\}) \leq \exp(-c\eta \min\{\eta, n^{1/2}\}).$$

This completes the proof. \square

7. VARIANCE AND COVARIANCE OF THE INCREMENTS $X_F(G_i)$

The aim of this section is to prove that the conditional variance $\text{Var}(X_F(G_i)|G_{i-1})$ of $X_F(i)$ and the conditional covariance

$$\mathbb{E}[X_F(G_i)X_{F'}(G_i) | G_{i-1}],$$

of $X_F(G_i)$ and $X_{F'}(G_i)$, are very predictable, in the sense that they are generally close to certain deterministic functions. Given two graphs, F with v vertices and e edges, and F' with v' vertices and e' edges, let us define

$$V_{F,F'}(i, n) := n^{v+v'-5} s^{e+e'-4} (1-s)(s\theta_1(F, F') + (1-s)\theta_2(F, F')), \quad (7.1)$$

where

$$\theta_1(F, F') := 8 \binom{F}{\wedge} \binom{F'}{\wedge} \quad \text{and} \quad \theta_2(F, F') := 36 \binom{F}{\triangle} \binom{F'}{\triangle}. \quad (7.2)$$

Proposition 7.1. *Let F, F' be graphs with v, v' vertices (respectively) and e, e' edges (respectively) and let $t \in (0, 1)$. There is a constant $C = C(F, F', t)$ such that, for all $1 \leq i \leq tN$ and all $3 \log n \leq b \leq n/2C$, we have*

$$\mathbb{P}\left(|\mathbb{E}[X_F(G_i)X_{F'}(G_i) | G_{i-1}] - V_{F,F'}(i, n)| > Cb^{1/2}n^{v+v'-11/2}\right) \leq \exp(-b).$$

Since the path of length two and the triangle play a particularly important role (see Theorem 2.4, for example), it is perhaps of interest to note that in these cases we have

$$V_{\wedge, \wedge}(i, n) = 8ns(1-s),$$

$$V_{\wedge, \triangle}(i, n) = 24ns^2(1-s) \quad \text{and}$$

$$V_{\triangle, \triangle}(i, n) = 36ns^2(1-s^2).$$

We may also now explain why we chose to express the terms of $\Lambda_H^*(G_{n,t})$ as multiples of $X_\wedge(G_i)$ and $X_\Delta(G_i) - 3sX_\wedge(G_i)$. This representation was chosen because these increments are asymptotically orthogonal in the sense that

$$\mathbb{E} [X_\wedge(G_i)(X_\Delta(G_i) - 3sX_\wedge(G_i)) \mid G_{i-1}] = o(n)$$

with high probability. This follows from Proposition 7.1 and the fact that

$$V_{\wedge,\Delta}(i, n) - 3sV_{\wedge,\wedge}(i, n) = 0.$$

We in fact prove an even more precise result, Proposition 7.2, which includes a second order term related to the current deviation $D_\wedge(G_{i-1})$. We define

$$W_{F,F'}(G_{i-1}) := 8n^{v+v'-7}s^{e+e'-4} \binom{F}{\wedge} \binom{F'}{\wedge} D_\wedge(G_{i-1}). \quad (7.3)$$

Proposition 7.2. *Let F, F' be graphs with v, v' vertices (respectively) and e, e' edges (respectively) and let $t \in (0, 1)$. There is a constant $C = C(F, F', t)$ such that, for all $1 \leq i \leq tN$ and all $3 \log n \leq b \leq n/2C$, we have*

$$\mathbb{P} \left(\left| \mathbb{E} [X_F(G_i)X_{F'}(G_i) \mid G_{i-1}] - (V_{F,F'}(i, n) + W_{F,F'}(G_{i-1})) \right| > Cbn^{v+v'-6} \right) \leq \exp(-b).$$

The term $W_{F,F'}(G_{i-1})$ is generally much smaller than the main term $V_{F,F'}(i, n)$. This follows from the fact that $D_\wedge(G_{i-1})$ is generally much smaller than n^2 , which follows from Theorem 1.7.

Let us observe that Proposition 7.1 follows from Proposition 7.2 and Theorem 1.7.

Proof of Proposition 7.1. Let F, F' and t be fixed. Writing C_1 for the constant of Proposition 7.2, and c for the constant associated with $H = \wedge$ in Theorem 1.7, we define

$$C = 2 \left(C_1 + 8c^{-1} \binom{F}{\wedge} \binom{F'}{\wedge} \right).$$

By the triangle inequality,

$$\begin{aligned} & \left| \mathbb{E} [X_F(G_i)X_{F'}(G_i) \mid G_{i-1}] - V_{F,F'}(i, n) \right| \\ & \leq \left| \mathbb{E} [X_F(G_i)X_{F'}(G_i) \mid G_{i-1}] - (V_{F,F'}(i, n) + W_{F,F'}(G_{i-1})) \right| + |W_{F,F'}(i, n)|, \end{aligned}$$

and so the event

$$\left| \mathbb{E} [X_F(G_i)X_{F'}(G_i) \mid G_{i-1}] - V_{F,F'}(i, n) \right| > Cb^{1/2}n^{v+v'-11/2} \quad (7.4)$$

may only occur if

$$\begin{aligned} \left| \mathbb{E} [X_F(G_i)X_{F'}(G_i) \mid G_{i-1}] - (V_{F,F'}(i, n) + W_{F,F'}(G_{i-1})) \right| & > 2C_1b^{1/2}n^{v+v'-11/2} \\ & \geq 2C_1bn^{v+v'-6} \end{aligned}$$

or

$$|D_\wedge(G_{i-1})| > 2c^{-1}b^{1/2}n^{3/2}.$$

By Proposition 7.2 and Theorem 1.7 respectively these events each have probability at most $\exp(-2b)$. By a union bound, the event (7.4) has probability at most $2 \exp(-2b) \leq \exp(-b)$, as required. \square

We now prove Proposition 7.2.

Proof of Proposition 7.2. Let F, F' and t be fixed. Let $i \leq tN$. Since

$$X_F(G_i) := A_F(G_i) - \mathbb{E}[A_F(G_i) \mid G_{i-1}],$$

the conditional covariance may be expressed as

$$\begin{aligned} & \mathbb{E}[X_F(G_i)X_{F'}(G_i) \mid G_{i-1}] \\ &= \mathbb{E}[A_F(G_i)A_{F'}(G_i) \mid G_{i-1}] - \mathbb{E}[A_F(G_i) \mid G_{i-1}]\mathbb{E}[A_{F'}(G_i) \mid G_{i-1}]. \end{aligned} \quad (7.5)$$

The proof consists of two main stages. In the first we express each of the terms of (7.5) as a linear combination of terms of the form $L_H(i)$ and $D_H(G_i)$. This first stage results in the expressions (7.7) and (7.8), which we combine to obtain (7.9), an expression for $\mathbb{E}[X_F(G_i)X_{F'}(G_i) \mid G_{i-1}]$ in terms of $L_H(i)$ and $D_H(G_i)$. In the second stage we expand these terms and, after some calculation, arrive at the conclusion.

We begin the first stage by calculating an expression⁴ for $\mathbb{E}[A_F(G_i) \mid G_{i-1}]$. Let Φ_F be the set of injective functions $\phi : V(F) \rightarrow V(G_{i-1})$. We write $\phi(e)$ for the image of an edge, i.e., $\phi(uw) = \{\phi(u), \phi(w)\}$. For $f \in E(F)$, we say that ϕ is *f-ready* if $\phi(e)$ is an edge of G_{i-1} for every edge of $F \setminus f$ and $\phi(f)$ is a non-edge of G_{i-1} . For $f \in E(F)$, we define

$$\Phi_F^f := \{\phi \in \Phi_F : \phi \text{ is } f\text{-ready}\}.$$

Since a copy of F may only be created at the moment we add its final edge, and each of the $N - i + 1$ remaining pairs is added as the i th edge with probability $1/(N - i + 1)$, we have

$$\mathbb{E}[A_F(G_i) \mid G_{i-1}] = \sum_{f \in E(F)} \frac{|\Phi_F^f|}{N - i + 1}. \quad (7.6)$$

Observe that $|\Phi_F^f|$, the number of f -ready injective functions ϕ , is given by

$$|\Phi_F^f| = N_{F \setminus f}(G_{i-1}) - N_F(G_{i-1}).$$

Substituting this into (7.6), and expanding each $N_H(G_{i-1})$ as $L_H(i - 1) + D_H(G_{i-1})$, we obtain

$$\begin{aligned} & \mathbb{E}[A_F(G_i) \mid G_{i-1}] \\ &= \sum_{f \in E(F)} \frac{L_{F \setminus f}(i - 1) - L_F(i - 1)}{N - i + 1} + \frac{1}{N - i + 1} \sum_{f \in E(F)} (D_{F \setminus f}(G_{i-1}) - D_F(G_{i-1})). \end{aligned}$$

Since it is easily checked that

$$\sum_{f \in E(F)} \frac{L_{F \setminus f}(i - 1) - L_F(i - 1)}{N - i + 1} = L_F(i) - L_F(i - 1),$$

we obtain

$$\mathbb{E}[A_F(G_i) \mid G_{i-1}] = (L_F(i) - L_F(i - 1)) + \frac{1}{N - i + 1} \sum_{f \in E(F)} (D_{F \setminus f}(G_{i-1}) - D_F(G_{i-1})). \quad (7.7)$$

⁴This expression is in fact already given by Lemma 2.3, however reproving it is a useful step towards the more difficult challenge of expressing $\mathbb{E}[A_F(G_i)A_{F'}(G_i) \mid G_{i-1}]$ in the desired form.

We continue the first stage by calculating the expression (7.8) for $\mathbb{E}[A_F(G_i)A_{F'}(G_i)|G_{i-1}]$. To abbreviate the notation we set

$$\mathbb{E}_{F,F'}(G_{i-1}) := \mathbb{E}[A_F(G_i)A_{F'}(G_i) | G_{i-1}].$$

Let $\Phi_{F,F'} = \Phi_F \times \Phi_{F'}$, be the set of pairs (ϕ, ϕ') of injective functions $\phi : V(F) \rightarrow V(G_{i-1})$ and $\phi' : V(F') \rightarrow V(G_{i-1})$. We say that such a pair (ϕ, ϕ') is (f, f', \rightarrow) -ready, for edges $f \in E(F)$, $f' \in E(F')$ and a relative orientation \rightarrow of the edges f and f' , if

- (i) $\phi(e) \in E(G_{i-1})$ for all $e \in F \setminus f$,
- (ii) $\phi'(e) \in E(G_{i-1})$ for all $e \in F' \setminus f'$ in G_{i-1} , and
- (iii) $\phi(f)$ and $\phi'(f')$ map to the same non-edge of G_{i-1} , and have relative orientation \rightarrow .

For $f \in E(F)$, $f' \in E(F')$ and a relative orientation \rightarrow , we define

$$\Phi_{F,F'}^{f,f',\rightarrow} := \{(\phi, \phi') \in \Phi_{F,F'} : (\phi, \phi') \text{ is } (f, f', \rightarrow)\text{-ready}\}.$$

Since embeddings of F and F' may only be simultaneously created at the moment we add their final edge, and each of the $N - i + 1$ remaining pairs is added as the i th edge with probability $1/(N - i + 1)$, we have

$$\mathbb{E}_{F,F'}(G_{i-1}) = \sum_{f,f',\rightarrow} \frac{|\Phi_{F,F'}^{f,f',\rightarrow}|}{N - i + 1}.$$

Observe that

$$|\Phi_{F,F'}^{f,f',\rightarrow}|,$$

the number of (f, f', \rightarrow) -ready pairs (ϕ, ϕ') includes a count over those pairs (ϕ, ϕ') whose images overlap in exactly two vertices, those whose images overlap in three vertices, and those pairs that overlap in four or more vertices. The count of pairs with overlap exactly two vertices is

$$N_{\Gamma^o(F,F':f,f',\rightarrow)}(G_{i-1}) - N_{\Gamma(F,F':f,f',\rightarrow)}(G_{i-1}).$$

where $\Gamma(F, F' : f, f', \rightarrow)$ is the graph obtained by joining F and F' by identifying f and f' using the relative orientation \rightarrow , and $\Gamma^o(F, F' : f, f', \rightarrow)$ is obtained by then removing the identified edge. The count of pairs with overlap exactly three vertices is

$$\sum_{u \in V(F) \setminus f, u' \in V(F') \setminus f'} (N_{\Gamma^o(F,F':f,f',u,u',\rightarrow)}(G_{i-1}) - N_{\Gamma(F,F':f,f',u,u',\rightarrow)}(G_{i-1}))$$

where $\Gamma(F, F' : f, f', u, u', \rightarrow)$ is the graph obtained by joining F and F' by identifying f and f' using the relative orientation \rightarrow and also identifying u and u' , and $\Gamma^o(F, F' : f, f', u, u', \rightarrow)$ is obtained by then removing the identified edge. It follows that

$$\begin{aligned} & \mathbb{E}_{F,F'}(G_{i-1}) \\ &= \frac{1}{N - i + 1} \sum_{f,f',\rightarrow} (N_{\Gamma^o(F,F':f,f',\rightarrow)}(G_{i-1}) - N_{\Gamma(F,F':f,f',\rightarrow)}(G_{i-1})) \\ & \quad + \frac{1}{N - i + 1} \sum_{f,f',u,u',\rightarrow} (N_{\Gamma^o(F,F':f,f',u,u',\rightarrow)}(G_{i-1}) - N_{\Gamma(F,F':f,f',u,u',\rightarrow)}(G_{i-1})) \\ & \quad + O(n^{v+v'-6}), \end{aligned}$$

where the error term $O(n^{v+v'-6})$ comes from the fact that there are $O(n^{v+v'-4})$ pairs that overlap in four or more vertices, and $N - i + 1 \geq N - tN + 1 = \Theta(n^2)$, as $t \in (0, 1)$ is fixed. Expanding the terms $N_H(G_{i-1})$ we obtain the desired expression for $\mathbb{E}[A_F(G_i)A_{F'}(G_i)|G_{i-1}]$:

$$\begin{aligned}
 & \mathbb{E}_{F,F'}(G_{i-1}) \\
 &= \frac{1}{N-i+1} \sum_{f,f',\rightarrow} (L_{\Gamma^o(F,F':f,f',\rightarrow)}(i-1) - L_{\Gamma(F,F':f,f',\rightarrow)}(i-1)) \\
 &+ \frac{1}{N-i+1} \sum_{f,f',u,u',\rightarrow} (L_{\Gamma^o(F,F':f,f',u,u',\rightarrow)}(i-1) - L_{\Gamma(F,F':f,f',u,u',\rightarrow)}(i-1)) \\
 &+ \frac{1}{N-i+1} \sum_{f,f',\rightarrow} (D_{\Gamma^o(F,F':f,f',\rightarrow)}(G_{i-1}) - D_{\Gamma(F,F':f,f',\rightarrow)}(G_{i-1})) \\
 &+ \frac{1}{N-i+1} \sum_{f,f',u,u',\rightarrow} (D_{\Gamma^o(F,F':f,f',u,u',\rightarrow)}(G_{i-1}) - D_{\Gamma(F,F':f,f',u,u',\rightarrow)}(G_{i-1})) \\
 &+ O(n^{v+v'-6}).
 \end{aligned} \tag{7.8}$$

Combining (7.7) and (7.8) and substituting into (7.5), we obtain the following expression for $\mathbb{E}[X_F(G_i)X_{F'}(G_i)|G_{i-1}]$:

$$\begin{aligned}
 & \mathbb{E}[X_F(G_i)X_{F'}(G_i) | G_{i-1}] \\
 &= \frac{1}{N-i+1} \sum_{f,f',\rightarrow} (L_{\Gamma^o(F,F':f,f',\rightarrow)}(i-1) - L_{\Gamma(F,F':f,f',\rightarrow)}(i-1)) \\
 &+ \frac{1}{N-i+1} \sum_{f,f',u,u',\rightarrow} (L_{\Gamma^o(F,F':f,f',u,u',\rightarrow)}(i-1) - L_{\Gamma(F,F':f,f',u,u',\rightarrow)}(i-1)) \\
 &+ \frac{1}{N-i+1} \sum_{f,f',\rightarrow} (D_{\Gamma^o(F,F':f,f',\rightarrow)}(G_{i-1}) - D_{\Gamma(F,F':f,f',\rightarrow)}(G_{i-1})) \\
 &+ \frac{1}{N-i+1} \sum_{f,f',u,u',\rightarrow} (D_{\Gamma^o(F,F':f,f',u,u',\rightarrow)}(G_{i-1}) - D_{\Gamma(F,F':f,f',u,u',\rightarrow)}(G_{i-1})) \\
 &- \left((L_F(i) - L_F(i-1)) + \frac{1}{N-i+1} \sum_{f \in E(F)} (D_{F \setminus f}(G_{i-1}) - D_F(G_{i-1})) \right) \\
 &\times \left((L_{F'}(i) - L_{F'}(i-1)) + \frac{1}{N-i+1} \sum_{f' \in E(F')} (D_{F' \setminus f'}(G_{i-1}) - D_{F'}(G_{i-1})) \right) \\
 &+ O(n^{v+v'-6}).
 \end{aligned} \tag{7.9}$$

We now begin the second stage of the proof. Essentially we must understand the terms in (7.9), and calculate what remains after cancellations. Our hope is that all the terms involving deviations reduce to $W_{F,F'}(G_{i-1})$, up to a small error term. To prove this we use Theorem 1.6 and Theorem 1.7.

By Theorem 1.6 and Theorem 1.7 there exists, for each v^* , a constant $C_1 = C_1(v^*)$ such that for all $1 \leq b \leq n/2C_1$ and all graphs H on at most v^* vertices, each of the events

$$|D_H(G_{n,s}) - \Lambda_H(G_{n,s})| \leq C_1 b n^{v(H)-2} \quad (7.10)$$

and

$$|D_H(G_{n,s})| \leq C_1 b^{1/2} n^{v(H)-3/2} \quad (7.11)$$

fail with probability at most $\exp(-((v^*)^2 + 2)b)$. Let $E_{v^*}(b)$ be the event that both of (7.10), (7.11) hold for any graph H on at most v^* vertices. By a straightforward union bound over the (at most $2^{(v^*)^2}$) graphs H on at most v^* vertices. We have that

$$\mathbb{P}(E_{v^*}(b)) \geq 1 - \exp(-b).$$

To complete the proof it suffices to prove that there is a constant C such that, on the event $E_{v+v'}(b)$, we have

$$|\mathbb{E}[X_F(G_i)X_{F'}(G_i) | G_{i-1}] - (V_{F,F'}(i, n) + W_{F,F'}(G_{i-1}))| \leq C b n^{v+v'-6}.$$

Let us continue our calculation of $\mathbb{E}[X_F(G_i)X_{F'}(G_i) | G_{i-1}]$ by expanding and cancelling the terms of (7.9). We begin by calculating the total contribution (7.15) of the terms involving only the $L_H(i-1)$. Using that both $\Gamma^o(F, F' : f, f', \rightarrow)$ and $\Gamma(F, F' : f, f', \rightarrow)$ have $v+v'-2$ vertices, and that $e(\Gamma^o(F, F' : f, f', \rightarrow)) = e+e'-2$ and $e(\Gamma(F, F' : f, f', \rightarrow)) = e+e'-1$, it is easily verified that

$$L_{\Gamma^o(F, F' : f, f', \rightarrow)}(i-1) - L_{\Gamma(F, F' : f, f', \rightarrow)}(i-1) = \frac{\binom{n}{v+v'-2} \binom{i-1}{e+e'-2} (N-i+1)}{\binom{N}{e+e'-1}},$$

for every choice of f, f', \rightarrow . It follows that

$$\begin{aligned} & \frac{1}{N-i+1} \sum_{f, f', \rightarrow} (L_{\Gamma^o(F, F' : f, f', \rightarrow)}(i-1) - L_{\Gamma(F, F' : f, f', \rightarrow)}(i-1)) \quad (7.12) \\ &= \frac{2ee'(n)_{v+v'-2} \binom{i-1}{e+e'-2}}{\binom{N}{e+e'-1}} \\ &= \frac{2ee'(n)_{v+v'-2} s^{e+e'-2}}{N} + O(n^{v+v'-6}) \\ &= 4ee'n^{-2} \left(1 + \frac{1}{n}\right) \left(n^{v+v'-2} - \binom{v+v'-2}{2} n^{v+v'-3}\right) s^{e+e'-2} + O(n^{v+v'-6}) \\ &= 4ee'n^{v+v'-4} s^{e+e'-2} + 2ee'(2 - (v+v'-2)(v+v'-3)) n^{v+v'-5} s^{e+e'-2} + O(n^{v+v'-6}). \end{aligned}$$

The main negative term comes from the product

$$-(L_F(i) - L_F(i-1))(L_{F'}(i) - L_{F'}(i-1)).$$

Since

$$L_F(i) - L_F(i-1) = \frac{e(n)_v \binom{i-1}{e-1}}{\binom{N}{e}} = 2en^{-2} \left(1 + \frac{1}{n}\right) (n)_v s^{e-1} + O(n^{v-4}),$$

this main negative term is

$$-4ee'n^{-4} \left(1 + \frac{2}{n}\right) (n)_v (n)_{v'} s^{e+e'-2} + O(n^{v+v'-6})$$

which may be expressed as

$$-4ee'n^{v+v'-4} s^{e+e'-2} + 2ee'(v(v-1) + v'(v'-1) - 4)n^{v+v'-5} s^{e+e'-2} + O(n^{v+v'-6}). \quad (7.13)$$

The final contribution from terms purely involving the terms $L_H(i-1)$ is

$$\frac{1}{N-i+1} \sum_{f,f',u,u',\rightarrow} (L_{\Gamma^o(F,F':f,f',u,u',\rightarrow)}(i-1) - L_{\Gamma(F,F':f,f',u,u',\rightarrow)}(i-1)).$$

The value of the summand depends on the number of extra overlaps of edges that occur in the identification. For $j = 0, 1, 2$, let λ_j be the number of sequences $f, f', u, u', \rightarrow$ in which $\Gamma^o(F, F' : f, f', u, u', \rightarrow)$ has $e + e' - 2 - j$ edges, meaning that j edges other than f and f' are lost in the identification. The contribution of

$$\frac{1}{N-i+1} (L_{\Gamma^o(F,F':f,f',u,u',\rightarrow)}(i-1) - L_{\Gamma(F,F':f,f',u,u',\rightarrow)}(i-1))$$

is

$$2n^{v+v'-5} s^{e+e'-2-j} + O(n^{v+v'-6})$$

in the case of j extra edges being lost in the identification. It follows that the total contribution of these terms is

$$2n^{v+v'-5} (\lambda_0 s^{e+e'-2} + \lambda_1 s^{e+e'-3} + \lambda_2 s^{e+e'-4}) + O(n^{v+v'-6}). \quad (7.14)$$

Summing all contributions to (7.9) from terms involving only the $L_H(i-1)$, i.e., summing (7.12), (7.13) and (7.14), we obtain

$$2n^{v+v'-5} \left((\lambda_0 - 2ee'(v-2)(v'-2)) s^{e+e'-2} + \lambda_1 s^{e+e'-3} + \lambda_2 s^{e+e'-4} \right) + O(n^{v+v'-6}).$$

Using that $\lambda_0 + \lambda_1 + \lambda_2 = 2ee'(v-2)(v'-2)$, this total contribution is

$$\begin{aligned} & 2n^{v+v'-5} \left((-\lambda_1 - \lambda_2) s^{e+e'-2} + \lambda_1 s^{e+e'-3} + \lambda_2 s^{e+e'-4} \right) + O(n^{v+v'-6}) \\ &= 2n^{v+v'-5} (\lambda_1 s^{e+e'-3} (1-s) + \lambda_2 s^{e+e'-4} (1-s^2)) + O(n^{v+v'-6}). \end{aligned}$$

We may now relate λ_1 and λ_2 to the parameters $\theta_1(F, F')$ and $\theta_2(F, F')$ that occur in the definition of $V_{F, F'}(i, n)$.

Claim: We have $2\lambda_1 = \theta_1(F, F') - 2\theta_2(F, F')$ and $2\lambda_2 = \theta_2(F, F')$.

Proof of Claim: Let ρ_1 be the number of pairs of an edge f of F and a disjoint vertex u such there is precisely one edge between the endpoints of f and u . Let ρ_2 be the number of such pairs in which both possible edges are present, i.e., $f \cup \{u\}$ is a triangle in F , and let ρ'_1 and ρ'_2 be the equivalent quantities in F' . It is easily verified that

$$\rho_1 = 2 \binom{F}{\wedge} - 6 \binom{F}{\triangle} \quad \text{and} \quad \rho_2 = 3 \binom{F}{\triangle}.$$

Let us now prove that $2\lambda_2 = \theta_2(F, F')$. We recall that λ_2 counts the number of choices $f, f', u, u', \rightarrow$ such that the overlap contains two extra edges. This occurs if and only if

$f \cup \{u\}$ and $f' \cup \{u'\}$ are triangles in their respective graphs, and so $\lambda_2 = 2\rho_2\rho'_2$, where the factor of 2 has come from counting the two possible orientations. It follows that

$$2\lambda_2 = 4\rho_2\rho'_2 = 36\binom{F}{\Delta}\binom{F'}{\Delta} = \theta_2(F, F').$$

We now turn to λ_1 , which counts the number of choices $f, f', u, u', \rightarrow$ such that the overlap contains exactly one extra edge. This occurs for one of the two orientations if there is one edge between f and u and likewise between f' and u' , and with both orientations if one of the two is a triangle. Thus

$$\lambda_1 = \rho_1\rho'_1 + 2\rho_1\rho'_2 + 2\rho_2\rho'_1.$$

Substituting in the values of $\rho_1, \rho_2, \rho'_1, \rho'_2$ we obtain

$$2\lambda_1 = 2\left(2\binom{F}{\Delta} - 6\binom{F}{\Delta}\right)\left(2\binom{F'}{\Delta} - 6\binom{F'}{\Delta}\right) + 12\left(2\binom{F}{\Delta} - 6\binom{F}{\Delta}\right)\binom{F'}{\Delta} + 12\left(2\binom{F'}{\Delta} - 6\binom{F'}{\Delta}\right)\binom{F}{\Delta},$$

which is $\theta_1(F, F') - 2\theta_2(F, F')$, completing the proof of the claim.

Using the claim, the total contribution of the terms involving only the $L_H(i-1)$ is

$$\begin{aligned} n^{v+v'-5}s^{e+e'-4}\left((s-s^2)(\theta_1(F, F') - 2\theta_2(F, F')) + (1-s^2)\theta_2(F, F')\right) = \\ n^{v+v'-5}s^{e+e'-4}(1-s)\left(s\theta_1(F, F') + (1-s)\theta_2(F, F')\right) = V_{F, F'}(i, n). \end{aligned} \quad (7.15)$$

We now turn to terms involving deviations $D_H(G_{i-1})$. On the event $E_{v+v'}(b)$, that both of (7.10) and (7.11) hold for all graphs on at most $v+v'$ vertices, we have that the deviation $D_H(G_{i-1})$ is given by

$$n^{v(H)-3}s^{e(H)-2}\binom{H}{\Delta}D_{\Delta}(G_{i-1}) + n^{v(H)-3}s^{e(H)-3}\binom{H}{\Delta}\left(D_{\Delta}(G_{i-1}) - 3sD_{\Delta}(G_{i-1})\right) \pm C_1bn^{v(H)-2} \quad (7.16)$$

for all graphs H on at most $v+v'$ vertices, and so, in particular for any graph included in (7.9). Thus, we need only to determine the coefficients of $D_{\Delta}(G_{i-1})$ and $D_{\Delta}(G_{i-1})$ obtained after summing the terms of (7.9) that involve deviations. We note that the terms with overlap at least 3, in the sum over $f, f', u, u', \rightarrow$ for example, are at most

$$C_1b^{1/2}n^{v+v'-13/2} = O(n^{v+v'-6})$$

on $E_{v+v'}(b)$. We may also safely ignore the term

$$-\frac{1}{(N-i+1)^2} \sum_{f \in E(F)} (D_{F \setminus f}(G_{i-1}) - D_F(G_{i-1})) \sum_{f' \in E(F')} (D_{F' \setminus f'}(G_{i-1}) - D_{F'}(G_{i-1}))$$

which has absolute value at most

$$C_1^2bn^{v+v'-7} = O(n^{v+v'-6})$$

on $E_{v+v'}(b)$. The remaining terms are

$$\begin{aligned} & \frac{1}{N-i+1} \sum_{f,f',\rightarrow} (D_{\Gamma^\circ(F,F':f,f',\rightarrow)}(G_{i-1}) - D_{\Gamma(F,F':f,f',\rightarrow)}(G_{i-1})) \\ & - (L_F(i) - L_F(i-1)) \frac{1}{N-i+1} \sum_{f' \in E(F')} (D_{F' \setminus f'}(G_{i-1}) - D_{F'}(G_{i-1})) \\ & - (L_{F'}(i) - L_{F'}(i-1)) \frac{1}{N-i+1} \sum_{f \in E(F)} (D_{F \setminus f}(G_{i-1}) - D_F(G_{i-1})) \end{aligned}$$

By expanding each $D_H(G_{i-1})$ using (7.16) we will find an expression for the remaining terms of (7.9) as a combination

$$\beta_1 D_\wedge(G_{i-1}) + \beta_2 (D_\Delta(G_{i-1}) - 3s D_\wedge(G_{i-1})) + O(bn^{v+v'-6})$$

on $E_{v+v'}(b)$.

Let us first calculate β_1 . Using that

$$L_F(i) - L_F(i-1) = \frac{e(n)_v(i-1)_{e-1}}{(N)_e} = 2en^{v-2}s^{e-1} + O(n^{v-3})$$

and expanding each $D_H(G_{i-1})$ using (7.16), we find that

$$\begin{aligned} (N-i+1)\beta_1 &= n^{v+v'-5}s^{e+e'-4} \sum_{f,f',\rightarrow} \left(\binom{\Gamma^\circ(F,F':f,f',\rightarrow)}{\wedge} - s \binom{\Gamma(F,F':f,f',\rightarrow)}{\wedge} \right) \\ & - 2en^{v+v'-5}s^{e+e'-4} \sum_{f' \in E(F')} \left(\binom{F' \setminus f'}{\wedge} - s \binom{F'}{\wedge} \right) \\ & - 2e'n^{v+v'-5}s^{e+e'-4} \sum_{f \in E(F)} \left(\binom{F \setminus f}{\wedge} - s \binom{F}{\wedge} \right). \end{aligned} \quad (7.17)$$

We may count the contribution of the first sum as follows, each P_2 in F is counted $2(e-2)e'$ times by the first term and $2ee'$ times by the second, while P_2 s in F' are counted $2e(e'-2)$ and $2ee'$ times respectively. The other way to find a P_2 in these graphs is crossing between F and F' ; a little thought shows that there are

$$4 \binom{F}{\wedge} \binom{F'}{\wedge}$$

such contributions to each of the two terms. Thus the result of the first sum is

$$2ee' \left(\binom{F}{\wedge} + \binom{F'}{\wedge} \right) (1-s) + 4 \binom{F}{\wedge} \binom{F'}{\wedge} (1-s) - 2e' \binom{F}{\wedge} - 2e \binom{F'}{\wedge}.$$

The equivalent results for the second and third terms are

$$e' \binom{F'}{\wedge} (1-s) - 2 \binom{F'}{\wedge}$$

and

$$e \binom{F}{\wedge} (1-s) - 2 \binom{F}{\wedge}$$

respectively. Substituting these values in (7.17) we obtain

$$(N-i+1)\beta_1 = 4n^{v+v'-5} \binom{F}{\wedge} \binom{F'}{\wedge} s^{e+e'-4} (1-s),$$

and so

$$\beta_1 = 8n^{v+v'-7} s^{e+e'-4} \binom{F}{\wedge} \binom{F'}{\wedge} + O(n^{v+v'-8}).$$

This is consistent with our aim to prove that the contribution of terms involving deviations is given by $W_{F,F'}(G_{i-1})$ up to $O(bn^{v+v'-6})$. All that remains is to prove that $\beta_2 = 0$. That is, the terms which contribute a multiple of $D_\Delta(G_{i-1}) - 3sD_\Delta(G_{i-1})$ in the expansion cancel. We have

$$\begin{aligned} \beta_2(N-i+1) &= n^{v+v'-5} s^{e+e'-5} \sum_{f,f',\rightarrow} \left(\left(\Gamma^\circ(F,F':f,f',\rightarrow) \right)_\Delta - s \left(\Gamma(F,F':f,f',\rightarrow) \right)_\Delta \right) \\ &\quad - 2en^{v+v'-5} s^{e+e'-5} \sum_{f' \in E(F')} \left(\left(\binom{F' \setminus f'}{\Delta} \right) - s \binom{F'}{\Delta} \right) \\ &\quad - 2e'n^{v+v'-5} s^{e+e'-5} \sum_{f \in E(F)} \left(\left(\binom{F \setminus f}{\Delta} \right) - s \binom{F}{\Delta} \right). \end{aligned}$$

The calculation is as above; however, since no triangles can cross between F and F' , we obtain only terms that cancel. The result is that $\beta_2 = 0$. This confirms that the contribution of terms in (7.9) that involve deviations contribute

$$W_{F,F'}(G_{i-1}) + O(bn^{v+v'-6})$$

on $E_{v+v'}(b)$. Combining this with (7.15) we obtain

$$\left| \mathbb{E} [X_F(G_i) X_{F'}(G_i) \mid G_{i-1}] - (V_{F,F'}(i, n) + W_{F,F'}(G_{i-1})) \right| \leq Cbn^{v+v'-6}$$

for an appropriately chosen constant C on $E_{v+v'}(b)$, an event with probability at least $1 - \exp(-b)$, as required. \square

8. PROBABILITY OF SUBGRAPH COUNT DEVIATIONS – THEOREM 1.1

In this section we bring together the various threads and complete our proof of Theorem 1.1. Armed with Theorem 2.6 it will suffice to prove the analogous statement with $\Lambda_H^*(G_{n,t})$ in place of $D_H(G_{n,t})$.

Proposition 8.1. *Let $t = t(n) \in (0, 1)$ be a sequence bounded away from 1, let H be graph with v vertices, e edges, and $\binom{H}{\Delta} \geq 1$. Then*

$$\mathbb{P} \left(\Lambda_H^*(G_{n,t}) > \alpha_n n^{v-3/2} \right) = \exp \left(-\gamma_H(t) \alpha_n^2 (1 + o(1)) \right),$$

for every sequence $(\alpha_n : n \geq 1)$ with $t^{-3/2} \ll \alpha_n \ll t^{e+2} n^{1/2}$. Furthermore, the same holds for $\mathbb{P} \left(\Lambda_H^*(G_{n,t}) < -\alpha_n n^{v-3/2} \right)$.

Here the expression $\gamma_H(t)$ is as defined in the introduction, namely:

$$\gamma_H(t) := \left(4 \binom{H}{\Delta}^2 t^{2e-2} (1-t)^2 + 12 \binom{H}{\Delta}^2 t^{2e-3} (1-t)^3 \right)^{-1}.$$

Let us observe that indeed Theorem 1.1 follows from Proposition 8.1 and Theorem 2.6.

Proof of Theorem 1.1. Let us fix $t = t(n) \in (0, 1)$ and a graph H with v vertices, e edges, and $\binom{H}{\Delta} \geq 1$. Let the sequence $(\alpha_n : n \geq 1)$ with

$$\max\{t^{1/2} n^{-1/2} \log n, t^{e-3/2}\} \ll \alpha_n \ll \min\{t^{2e-5/2} n^{1/2}, t^{e+2} n^{1/2}\}$$

be given. Finally, let us also fix $\varepsilon > 0$. We may suppose $\varepsilon < 1/10$.

We first show the upper bound on $\mathbb{P}(D_H(G_{n,t}) > \alpha_n n^{v-3/2})$. We begin by observing that

$$\begin{aligned} \mathbb{P}(D_H(G_{n,t}) > \alpha_n n^{v-3/2}) &\leq \mathbb{P}(\Lambda_H^*(G_{n,t}) > (1-\varepsilon)\alpha_n n^{v-3/2}) \\ &\quad + \mathbb{P}(|D_H(G_{n,t}) - \Lambda_H^*(G_{n,t})| > \varepsilon\alpha_n n^{v-3/2}). \end{aligned}$$

Now, by Proposition 8.1, we have

$$\begin{aligned} \mathbb{P}(\Lambda_H^*(G_{n,t}) > (1-\varepsilon)\alpha_n n^{v-3/2}) &\leq \exp(-\gamma_H(t)\alpha_n^2(1-\varepsilon+o(1))^2) \\ &\leq \exp(-\gamma_H(t)\alpha_n^2(1-3\varepsilon)) \end{aligned}$$

for all sufficiently large n . On the other hand, we shall apply Theorem 2.6 (as stated in Theorem 2.4) with $b = \varepsilon\alpha_n C^{-1}t^{-1/2}n^{1/2}$ to bound

$$\mathbb{P}(|D_H(G_{n,t}) - \Lambda_H^*(G_{n,t})| > \varepsilon\alpha_n n^{v-3/2}).$$

It is easily checked that the conditions on α_n ensure that $3\log n \leq b \leq t^{1/2}n$, for all n sufficiently large. By Theorem 2.6 we have

$$\mathbb{P}(|D_H(G_{n,t}) - \Lambda_H^*(G_{n,t})| > \varepsilon\alpha_n n^{v-3/2}) \leq \exp(-c\varepsilon t^{-1/2}\alpha_n n^{1/2})$$

for some constant $c > 0$. Since $\alpha_n \ll t^{2e-5/2}n^{1/2}$, we have that

$$c\varepsilon t^{-1/2}\alpha_n n^{1/2} \gg \gamma_H(t)\alpha_n^2(1-3\varepsilon)$$

and so, for all sufficiently large n ,

$$\mathbb{P}(D_H(G_{n,t}) > \alpha_n n^{v-3/2}) \leq (1+\varepsilon)\exp(-\gamma_H(t)\alpha_n^2(1-3\varepsilon)).$$

Since ε is arbitrary, and $\gamma_H(t)\alpha_n^2 \gg 1$, we have

$$\mathbb{P}(D_H(G_{n,t}) > \alpha_n n^{v-3/2}) \leq \exp(-\gamma_H(t)\alpha_n^2(1-o(1))).$$

The proof of the lower bound follows immediately by the same argument, and the fact that

$$\begin{aligned} \mathbb{P}(D_H(G_{n,t}) > \alpha_n n^{v-3/2}) &\geq \mathbb{P}(\Lambda_H^*(G_{n,t}) > (1+\varepsilon)\alpha_n n^{v-3/2}) \\ &\quad - \mathbb{P}(|D_H(G_{n,t}) - \Lambda_H^*(G_{n,t})| > \varepsilon\alpha_n n^{v-3/2}). \end{aligned}$$

This completes the proof. \square

Our remaining task is to prove Proposition 8.1. Let us recall that $\Lambda_H^*(G_{n,t})$ is the martingale expression

$$\Lambda_H^*(G_{n,t}) = \sum_{i=1}^{\lfloor tN \rfloor} (\kappa_{H,n}^t(i)X_\Lambda(G_i) + \rho_{H,n}^t(i)(X_\Delta(G_i) - 3sX_\Lambda(G_i))),$$

where

$$\kappa_{H,n}^t(i) := n^{v-3}t^{e-2} \binom{H}{\Lambda} \frac{(1-t)^2}{(1-s)^2}$$

and

$$\rho_{H,n}^t(i) := n^{v-3}t^{e-3} \binom{H}{\Delta} \frac{(1-t)^3}{(1-s)^3}.$$

We prove a general statement on the probability of deviations of martingales of this general form.

Proposition 8.2. *Let $t = t(n) \in (0, 1)$ be a sequence bounded away from 1, let $\kappa = (\kappa_n^t)_{n \geq 1}$ and $\rho = (\rho_n^t)_{n \geq 1}$ be two sequences of functions such that $\kappa_n^t : \{1, \dots, [tN]\} \rightarrow [-C, C]$ for some constant $C \in \mathbb{R}$ and $\rho_n^t : \{1, \dots, [tN]\} \rightarrow [-Ct^{-1}, Ct^{-1}]$, and suppose there exists $\eta > 0$ such that*

$$\sum_{i=1}^{[tN]} |\kappa_n^t(i)| + t|\rho_n^t(i)| \geq \eta tN \quad (8.1)$$

for all sufficiently large n . Then

$$S_n^t := \sum_{i=1}^{[tN]} (\kappa_n^t(i)X_{\wedge}(G_i) + \rho_n^t(i)(X_{\Delta}(G_i) - 3sX_{\wedge}(G_i)))$$

satisfies

$$\mathbb{P}\left(S_n^t > \alpha_n n^{3/2}\right) = \exp\left(\frac{-\alpha_n^2(1+o(1))}{2\tau_{\kappa,\rho}}\right),$$

for every sequence $(\alpha_n : n \geq 1)$ with $t^{1/2} \ll \alpha_n \ll t^4 n^{1/2}$, where

$$\tau_{\kappa,\rho} := n^{-2} \sum_{i=1}^{[tN]} \left(8s(1-s)\kappa_n^t(i)^2 + 36s^2(1-s)^2\rho_n^t(i)^2\right).$$

Furthermore the same holds for $\mathbb{P}(S_n^t < -\alpha_n n^{v-3/2})$.

Let us observe that indeed Proposition 8.1 follows from Proposition 8.2.

Proof of Proposition 8.1. Let κ be the sequence of functions

$$\kappa_n^t(i) = n^{3-v}t^{2-e}\kappa_{H,n}^t(i) = \binom{H}{\wedge} \frac{(1-t)^2}{(1-s)^2}$$

and let ρ be

$$\rho_n^t(i) = n^{3-v}t^{2-e}\rho_{H,n}^t(i) = t^{-1} \binom{H}{\Delta} \frac{(1-t)^3}{(1-s)^3}.$$

It is easily verified that the average $\frac{1}{[tN]} \sum_{i=1}^{[tN]} \kappa_n^t(i)$ is bounded away from 0.

Since $n^{v-3}t^{e-2}S_n^t = \Lambda_H^*(G_{n,t})$, we have

$$\mathbb{P}\left(\Lambda_H^*(G_{n,t}) > \alpha_n n^{v-3/2}\right) = \mathbb{P}\left(S_n^t > \alpha_n t^{2-e} n^{3/2}\right).$$

In order to apply Proposition 8.2 we must verify that $t^{1/2} \ll \alpha_n t^{2-e} \ll t^4 n^{1/2}$. This follows immediately from the condition that $t^{e-3/2} \ll \alpha_n \ll t^{e+2} n^{1/2}$. And so, by an application of Proposition 8.2, we have

$$\mathbb{P}\left(\Lambda_H^*(G_{n,t}) > \alpha_n n^{v-3/2}\right) = \exp\left(\frac{-\alpha_n^2 t^{4-2e}(1+o(1))}{2\tau_{\kappa,\rho}}\right),$$

where

$$\tau_{\kappa,\rho} := n^{-2} \sum_{i=1}^{\lfloor tN \rfloor} \left(8s(1-s)\kappa_n^t(i)^2 + 36s^2(1-s)^2\rho_n^t(i)^2 \right).$$

All that remains is to prove that

$$\gamma_H(t) = \frac{1 + o(1)}{2t^{2e-4}\tau_{\kappa,\rho}}. \quad (8.2)$$

Substituting the values of $\kappa_n^t(i)$ and $\rho_n^t(i)$ into the definition of $\tau_{\kappa,\rho}$ we obtain

$$\begin{aligned} t^{2e-4}\tau_{\kappa,\rho} &= t^{2e-4}n^{-2} \sum_{i=1}^{\lfloor tN \rfloor} 8s(1-s) \left(\binom{H}{\wedge} \frac{(1-t)^2}{(1-s)^2} \right)^2 \\ &\quad + t^{2e-4}n^{-2} \sum_{i=1}^{\lfloor tN \rfloor} 36s^2(1-s)^2 \left(t^{-1} \binom{H}{\Delta} \frac{(1-t)^3}{(1-s)^3} \right)^2. \end{aligned}$$

The contribution of the first term is

$$\begin{aligned} &\frac{(4 + o(1))}{N} t^{2e-4} (1-t)^4 \binom{H}{\wedge}^2 \sum_{i=1}^{\lfloor tN \rfloor} \frac{s}{(1-s)^3} \\ &= (4 + o(1)) t^{2e-4} (1-t)^4 \binom{H}{\wedge}^2 \int_0^t \frac{s}{(1-s)^3} ds \\ &= (2 + o(1)) t^{2e-2} (1-t)^2 \binom{H}{\wedge}^2, \end{aligned}$$

where we have used that the integral has value $t^2/2(1-t)^2$ (as may be seen using the substitution $u = 1 - s$, for example). The contribution of the second term is

$$\begin{aligned} &\frac{(18 + o(1))}{N} t^{2e-6} (1-t)^6 \binom{H}{\Delta}^2 \sum_{i=1}^{\lfloor tN \rfloor} \frac{s^2}{(1-s)^4} \\ &= (18 + o(1)) t^{2e-6} (1-t)^6 \binom{H}{\Delta}^2 \int_0^t \frac{s^2}{(1-s)^4} ds \\ &= (6 + o(1)) t^{2e-3} (1-t)^3 \binom{H}{\Delta}^2, \end{aligned}$$

where we have used that the integral has value $t^3/3(1-t)^3$. Summing these two contributions, we have

$$t^{2e-4}\tau_{\kappa,\rho} = (2 + o(1)) t^{2e-2} (1-t)^2 \binom{H}{\wedge}^2 + (6 + o(1)) t^{2e-3} (1-t)^3 \binom{H}{\Delta}^2.$$

By observation, (8.2) holds, and so the proof is complete. \square

Our final task is to prove Proposition 8.2. This proof will use the inequalities of Freedman stated in Section 3.

Proof of Proposition 8.2. Let $t \in (0, 1)$ and the sequences $\kappa = (\kappa_n^t)_{n \geq 1}$ and $\rho = (\rho_n^t)_{n \geq 1}$ be fixed. We may assume that $\kappa_n^t : \{1, \dots, \lfloor tN \rfloor\} \rightarrow [-C, C]$ and $\rho_n^t : \{1, \dots, \lfloor tN \rfloor\} \rightarrow$

$[-Ct^{-1}, Ct^{-1}]$ are such that (8.1) holds. Let us also fix $\varepsilon > 0$. It will be useful at times to note that

$$\Omega(t^2) \leq \tau_{\kappa, \rho} \leq O(t) \quad (8.3)$$

which follows easily from the definition of $\tau_{\kappa, \rho}$ and the conditions on κ_n^t and ρ_n^t .

We must prove an upper bound and a lower bound on the probability of a deviation of the final value S_n^t of the martingale

$$S_n^t := \sum_{i=1}^{\lfloor tN \rfloor} \left(\kappa_n^t(i) X_{\wedge}(G_i) + \rho_n^t(i) (X_{\Delta}(G_i) - 3sX_{\wedge}(G_i)) \right).$$

Fix the sequence $(\alpha_n : n \geq 1)$ with $t^{1/2} \ll \alpha_n \ll t^4 n^{1/2}$.

We first prove the upper bound on the probability

$$\mathbb{P} \left(S_n^t > \alpha_n n^{3/2} \right),$$

by an application of Freedman's inequality, Lemma 3.4. We have that S_n^t is the final value of a martingale

$$S_n^t := \sum_{i=1}^{\lfloor tN \rfloor} X(i),$$

with increments

$$X(i) := \kappa_n^t(i) X_{\wedge}(G_i) + \rho_n^t(i) (X_{\Delta}(G_i) - 3sX_{\wedge}(G_i)).$$

In order to apply Freedman's inequality we need to assess the quantity

$$V(\lfloor tN \rfloor) := \sum_{i=1}^{\lfloor tN \rfloor} \mathbb{E} [X(i)^2 | G_{i-1}].$$

Let $E_n^t(\varepsilon)$ be the event that

$$|V(\lfloor tN \rfloor) - n^3 \tau_{\kappa, \rho}| \leq \varepsilon t^2 n^3.$$

We bound the probability of $E_n^t(\varepsilon)$ using the following claim. The quantities $V_{F, F'}(i, n)$ are as defined by (7.1).

Claim: Let $\delta = \varepsilon/15C^2$. If

$$\begin{aligned} |\mathbb{E} [X_{\wedge}(G_i)^2 | G_{i-1}] - V_{\wedge, \wedge}(i, n)| &\leq \delta t n \\ |\mathbb{E} [X_{\wedge}(G_i) X_{\Delta}(G_i) | G_{i-1}] - V_{\wedge, \Delta}(i, n)| &\leq \delta t^2 n \quad \text{and} \\ |\mathbb{E} [X_{\Delta}(G_i) X_{\Delta}(G_i) | G_{i-1}] - V_{\Delta, \Delta}(i, n)| &\leq \delta t^3 n, \end{aligned}$$

for all $1 \leq i \leq \lfloor tN \rfloor$, then $E_n^t(\varepsilon)$ occurs.

Proof of Claim: We may express $\mathbb{E} [X(i)^2 | G_{i-1}]$ as

$$\begin{aligned} \kappa_n^t(i)^2 \mathbb{E} [X_{\wedge}(G_i)^2 | G_{i-1}] + 2\kappa_n^t(i)\rho_n^t(i) \mathbb{E} [X_{\wedge}(G_i)(X_{\Delta}(G_i) - 3sX_{\wedge}(G_i)) | G_{i-1}] \\ + \rho_n^t(i)^2 \mathbb{E} [(X_{\Delta}(G_i) - 3sX_{\wedge}(G_i))^2 | G_{i-1}]. \end{aligned}$$

By the assumption of the claim, it follows that

$$\begin{aligned} \mathbb{E} [X(i)^2 | G_{i-1}] &= \kappa_n^t(i)^2 V_{\wedge, \wedge}(i, n) \\ &\quad + 2\kappa_n^t(i)\rho_n^t(i) (V_{\wedge, \Delta}(i, n) - 3sV_{\wedge, \wedge}(i, n)) \\ &\quad + \rho_n^t(i)^2 (V_{\Delta, \Delta}(i, n) - 6sV_{\wedge, \Delta}(i, n) + 9s^2V_{\wedge, \wedge}(i, n)) \\ &\quad \pm 30C^2\delta tn. \end{aligned}$$

Substituting in the values

$$\begin{aligned} V_{\wedge, \wedge}(i, n) &= 8ns(1-s), \\ V_{\wedge, \Delta}(i, n) &= 24ns^2(1-s) \quad \text{and} \\ V_{\Delta, \Delta}(i, n) &= 36ns^2(1-s^2), \end{aligned}$$

we obtain that

$$\mathbb{E} [X(i)^2 | G_{i-1}] = 8ns(1-s)\kappa_n^t(i)^2 + 36ns^2(1-s)^2\rho_n^t(i)^2 \pm 30C^2\delta tn.$$

Now summing over $i = 1, \dots, \lfloor tN \rfloor$ we obtain

$$|V(\lfloor tN \rfloor) - n^3\tau_{\kappa, \rho}| \leq 30C^2\delta tn \lfloor tN \rfloor \leq \varepsilon t^2 n^3,$$

as required. This completes the proof of the claim.

Let $C_1 \geq \max\{15, C\}$ be at least the maximum constant of Proposition 7.1 for cases with $F, F' \in \{\wedge, \Delta\}$ and with $t = \sup_n t(n)$. By the claim, the event $E_n^t(\varepsilon)^c$ may only occur if one of the events in the condition of the claim fails to occur. Using Proposition 7.1 to bound the probability of such events, we obtain that, for all sufficiently large n ,

$$\begin{aligned} \mathbb{P}(E_n^t(\varepsilon)^c) &\leq 3tN \exp\left(\frac{-\varepsilon^2 t^6 n}{15^2 C^4 C_1^2}\right) \\ &\leq \exp\left(\frac{-\varepsilon^2 t^6 n}{C_1^8}\right). \end{aligned} \tag{8.4}$$

We are now ready to apply Lemma 3.4 and obtain our upper bound. For an upper bound on $\|X(i)\|_\infty$ we simply use that

$$|X(i)| \leq 4C|X_\wedge(i)| + Ct^{-1}|X_\Delta(i)| \leq 5Ct^{-1}n.$$

We observe that

$$\mathbb{P}(S_n^t > \alpha_n n^{3/2}) \leq \mathbb{P}(\{S_n^t > \alpha_n n^{3/2}\} \cap E_n^t(\varepsilon)) + \mathbb{P}(E_n^t(\varepsilon)^c).$$

We bound the first probability by applying Freedman's inequality with $\alpha = \alpha_n n^{3/2}$, $\beta = n^3(\tau_{\kappa, \rho} + \varepsilon t^2)$ and $R = 5Ct^{-1}n$, and the second by (8.4). We obtain

$$\mathbb{P}(S_n^t > \alpha_n n^{3/2}) \leq \exp\left(\frac{-\alpha_n^2 n^3}{2n^3(\tau_{\kappa, \rho} + \varepsilon t^2) + 10Ct^{-1}\alpha_n n^{5/2}}\right) + \exp\left(\frac{-\varepsilon^2 t^6 n}{C_1^8}\right).$$

for all sufficiently large n .

By the upper bound condition of α_n and (8.3) we have that $\alpha_n \ll t^4 n^{1/2} = O(t^3 \tau_{\kappa, \rho}^{1/2} n^{1/2})$. It follows that the second exponential above is $o(1)$ times the first exponential, and the

second term of the denominator in the first exponential is $o(1)$ times the first term. And so

$$\mathbb{P}\left(S_n^t > \alpha_n n^{3/2}\right) \leq \exp\left(\frac{-\alpha_n^2(1 - O(\varepsilon))}{2\tau_{\kappa,\rho}}\right)$$

for all sufficiently large n . Since $\varepsilon > 0$ is arbitrary this completes the proof of the upper bound

$$\mathbb{P}\left(S_n^t > \alpha_n n^{3/2}\right) \leq \exp\left(\frac{-\alpha_n^2(1 + o(1))}{2\tau_{\kappa,\rho}}\right).$$

We now prove the lower bound. In principle the proof of the lower bound should be straightforward, essentially equivalent to the proof of the upper bound, except with Lemma 3.5 being used instead of Lemma 3.4. One subtlety is that such a direct application of Lemma 3.5 would give a lower bound on the probability of a deviation occurring *before* a certain time, rather than *at* time $\lfloor tN \rfloor$. In particular, it will allow us to obtain a lower bound on the probability of the event $F_n^t(\varepsilon)$ that

$$\exists \ell \leq \lfloor tN \rfloor \quad \text{such that} \quad \sum_{i=1}^{\ell} X(i) > (1 + \varepsilon)\alpha_n n^{3/2}.$$

By an application of Freedman's inequality to the part of the martingale that occurs after first crossing $(1 + \varepsilon)\alpha_n n^{3/2}$, one easily verifies that there is at least probability $1/2$ that the martingale remains above $\alpha_n n^{3/2}$, for all sufficiently large n . And so,

$$\mathbb{P}\left(S_n^t > \alpha_n n^{3/2}\right) \geq \frac{1}{2} \mathbb{P}\left(F_n^t(\varepsilon)\right)$$

for all sufficiently large n . Thus, to complete the proof we need only prove that

$$\mathbb{P}\left(F_n^t(\varepsilon)\right) \geq \exp\left(\frac{-\alpha_n^2(1 + O(\varepsilon))}{2\tau_{\kappa,\rho}}\right)$$

for all sufficiently large n .

We recall that the statement of Lemma 3.5 provides a lower bound on the probability

$$\mathbb{P}(T_\alpha \leq \beta)$$

where T_α is defined by

$$T_\alpha = \sum_{i=1}^{m_\alpha} \mathbb{E}[|X(i)|^2 | \mathcal{F}_{i-1}].$$

where m_α is the least m such that the martingale exceeds α . If we take

$$\alpha = (1 + \varepsilon)\alpha_n n^{3/2}$$

and $\beta = n^3(\tau_{\kappa,\rho} - \varepsilon t^2)$, then it is easily observed that event $T_\alpha \leq \beta$ is contained in $E_n^t(\varepsilon) \cup F_n^t(\varepsilon)$. So we have that

$$\mathbb{P}\left(F_n^t(\varepsilon)\right) \geq \mathbb{P}(T_\alpha \leq \beta) - \mathbb{P}\left(E_n^t(\varepsilon)^c\right).$$

Applying Lemma 3.5, we obtain

$$\mathbb{P}\left(F_n^t(\varepsilon)\right) \geq \frac{1}{2} \exp\left(\frac{-\alpha^2(1 + 4\delta)}{2\beta}\right) - \exp\left(\frac{-\alpha_n^2}{\varepsilon\tau_{\kappa,\rho}}\right)$$

where $\delta > 0$ is minimal such that $\beta/\alpha \geq 9R\delta^{-2}$ and $\alpha^2/\beta \geq 16\delta^{-2} \log(64\delta^{-2})$. Substituting the values of α and β we obtain

$$\mathbb{P}(F_n^t(\varepsilon)) \geq \frac{1}{3} \exp\left(\frac{-\alpha_n^2(1 + O(\varepsilon) + O(\delta))}{2\tau_{\kappa,\rho}}\right).$$

From the definition of α_n it is easily verified that $\delta = o(1)$, and so

$$\mathbb{P}(F_n^t(\varepsilon)) \geq \exp\left(\frac{-\alpha_n^2(1 + O(\varepsilon))}{2\tau_{\kappa,\rho}}\right)$$

for all sufficiently large n , as required. This completes the proof of the main statement of the proposition.

The furthermore part of the statement follows immediately by multiplying the functions by -1 and applying the main part. \square

9. MODERATE DEVIATIONS OF SUBGRAPH COUNTS IN $G(n, p)$

As discussed in the sketch proof in the introduction, the proofs of both Theorem 1.9 and Theorem 1.11 are based around the identity

$$\mathbb{P}(D_H(G_p) > \delta_n p^e(n)_v) = \sum_{m=0}^N b_N(m) \mathbb{P}(N_H(G_m) > (1 + \delta_n)p^e(n)_v), \quad (9.1)$$

and in particular in finding which terms make the largest contribution to the sum. With this in mind we define

$$m_* := pN(1 + \delta_n)^{1/e}$$

and note that this is approximately (up to a small additive constant) the value of m at which no deviation is required in G_m in order that $N_H(G_m) > (1 + \delta_n)p^e(n)_v$. Indeed, since

$$\frac{(m - e)^e}{N^e} < \frac{(m)_e}{(N)_e} \leq \frac{m^e}{N^e}$$

we have $L_H(m_*) \leq (1 + \delta_n)p^e(n)_v$ and $L_H(m_* + e) > (1 + \delta_n)p^e(n)_v$.

It will be useful in the proof of Theorem 1.9 to also consider a value m_- slightly less than m_* and a value m_+ slightly larger. We define

$$m_- := \lfloor m_* - \delta_n^{-1/2} n^{1/4} \rfloor$$

and

$$m_+ := \lfloor m_* + \delta_n^{-1/2} n^{1/4} \rfloor.$$

The intuition behind the definitions of m_- and m_+ is simply that their difference from m_* is between order $n^{1/2}$ (the amount one must change m to have a significant effect on $L_H(m)$) and order δ_n^{-1} (the amount one can change m before it has a significant effect on the tail bound for binomial deviations). In other words, the probability that G_p has at least m_+ edges is asymptotically equivalent to the probability it has at least m_- edges, and yet the event $N_H(G_m) > (1 + \delta_n)p^e(n)_v$ changes from being very unlikely to very likely as m grows from m_- to m_+ .

In addition we define

$$x(m) := \frac{m - pN}{\sqrt{Npq}},$$

and we set $x_* := x(m_*)$, $x_- := x(m_-)$ and $x_+ := x(m_+)$.

We split the proof of Theorem 1.9 into two parts (Section 9.1 and Section 9.2), corresponding to the lower bound and upper bound. Theorem 1.11 is proved in Section 9.3.

9.1. Lower bound of Theorem 1.9. Let the sequence $n^{-1} \ll \delta_n \ll n^{-1/2}$ be given.

Since the second term, $\mathbb{P}(N_H(G_m) > (1 + \delta_n)p^e(n)_v)$, in the expression (9.1) is increasing in m it follows that

$$\begin{aligned} \mathbb{P}(D_H(G_p) > \delta_n p^e(n)_v) &\geq \sum_{m=m_+}^N b_N(m) \mathbb{P}(N_H(G_{m_+}) > (1 + \delta_n)p^e(n)_v) \\ &= B_N(m_+) \mathbb{P}(N_H(G_{m_+}) > (1 + \delta_n)p^e(n)_v) . \end{aligned}$$

The proof of the lower bound therefore reduces to proving the following two lemmas.

Lemma 9.1.

$$B_N(m_+) = \sqrt{\frac{e^2 q}{\pi p}} \exp\left(-\frac{\delta_n^2 p n^2}{4e^2 q} + \frac{((3e-2) - (3e-1)p)\delta_n^3 p n^2}{12e^3 q^2} - \log(n\delta_n) + o(1)\right).$$

Lemma 9.2.

$$\mathbb{P}(N_H(G_{m_+}) > (1 + \delta_n)p^e(n)_v) = 1 + o(1).$$

Let us first see that Lemma 9.2 follows easily from Theorem 1.7.

Proof. The event that

$$N_H(G_{m_+}) \leq (1 + \delta_n)p^e(n)_v \tag{9.2}$$

will correspond to a large negative deviation $D_H(G_{m_+})$. Indeed, we observe that

$$\begin{aligned} L_H(m_+) &= \frac{\binom{n}{v}(m_+)_e}{\binom{N}{e}} \\ &= \frac{\binom{n}{v}}{\binom{N}{e}} (\lfloor m_* + \delta_n^{-1/2} n^{1/4} \rfloor)_e \\ &\geq \frac{\binom{n}{v}}{\binom{N}{e}} (m_* + \delta_n^{-1/2} n^{1/4} - e)^e \\ &\geq \frac{\binom{n}{v}}{\binom{N}{e}} (m_*^e + \delta_n^{-1/2} n^{1/4} m_*^{e-1}) \\ &\geq \binom{n}{v} p^e (1 + \delta_n) + \Omega(\delta_n^{-1/2} n^{v-7/4}). \end{aligned}$$

And so (9.2) is contained in the event

$$D_H(G_{m_+}) \leq -\Omega(\delta_n^{-1/2} n^{v-7/4}).$$

Since $\delta_n^{-1/2} n^{v-7/4} = \omega(n^{v-3/2})$, this event has probability $o(1)$ by Theorem 1.7, as required. \square

Remark 9.3. In fact one does not need Theorem 1.7 to obtain Lemma 9.2. An alternative would be to use Chebyshev's inequality and the fact that the variance of $D_H(G_m)$ is $O(n^{2v-3})$.

Before proving Lemma 9.1 let us examine more closely the values of x_* and x_+ . We recall that $m_* := pN(1 + \delta_n)^{1/e}$, and so has expansion

$$m_* = pN \left(1 + \frac{\delta_n}{e} - \frac{\delta_n^2(e-1)}{2e^2} + O(\delta_n^3) \right).$$

It follows that x_* may be expressed as

$$\begin{aligned} x_* &= \frac{\delta_n p^{1/2} N^{1/2}}{e q^{1/2}} - \frac{\delta_n^2(e-1) p^{1/2} N^{1/2}}{2e^2 q^{1/2}} + O(\delta_n^3 n) \\ &= \frac{\delta_n p^{1/2} n}{e \sqrt{2} q^{1/2}} - \frac{\delta_n^2(e-1) p^{1/2} n}{2\sqrt{2} e^2 q^{1/2}} + O(\delta_n^3 n) + O(n^{-1}). \end{aligned} \quad (9.3)$$

As $m_+ - m_*$ is $\delta_n^{-1/2} n^{1/4} \pm 1$ it is clear that $x_+ - x_*$ is $\delta_n^{-1/2} n^{1/4} / \sqrt{Npq} \pm n^{-1}$. In fact we will only need that $x_+ - x_* = O(\delta_n^{-1/2} n^{-3/4})$, so that

$$x_+ = \frac{\delta_n p^{1/2} n}{e \sqrt{2} q^{1/2}} - \frac{\delta_n^2(e-1) p^{1/2} n}{2\sqrt{2} e^2 q^{1/2}} + O(\delta_n^{-1/2} n^{-3/4}). \quad (9.4)$$

Proof of Lemma 9.1. By Theorem 1.13 we have

$$B_N(m_+) = (1 + o(1)) \frac{1}{x_+ \sqrt{2\pi}} \exp \left(-\frac{x_+^2}{2} - E(x_+, N, 1) \right), \quad (9.5)$$

where we have used that $x_+ = \Theta(\delta_n n) \ll N^{1/4}$ to truncate the infinite sum $E(x_+, N)$ to $E(x_+, N, 1)$.

From our expression (9.4) for x_+ we have that

$$\frac{x_+^2}{2} = \frac{\delta_n^2 p n^2}{4e^2 q} - \frac{\delta_n^3(e-1) p n^2}{4e^3 q} + o(1)$$

and

$$x_+^3 = \frac{\delta_n^3 p^{3/2} n^3}{2^{3/2} e^3 q^{3/2}} + o(n).$$

It is straightforward to calculate that

$$\frac{x_+^2}{2} + E(x_+, N, 1) = \frac{\delta_n^2 p n^2}{4e^2 q} + \frac{((3e-1)p - (3e-2)) \delta_n^3 p n^2}{12e^3 q^2}.$$

Substituting this into (9.5) and using that $x_+ = \delta_n p^{1/2} n / e \sqrt{2} q^{1/2} + o(1)$, we obtain the desired result. \square

9.2. Upper bound of Theorem 1.9. A key observation is that the expression for $B_N(m_+)$ given in Lemma 9.1 is also an expression for $B_N(m_-)$ (as the difference between the two is contained in the $o(1)$ term). This follows easily from the proof of Lemma 9.1 and the observation that x_- may also be expressed as

$$x_- = \frac{\delta_n p^{1/2} n}{e \sqrt{2} q^{1/2}} - \frac{\delta_n^2(e-1) p^{1/2} n}{2\sqrt{2} e^2 q^{1/2}} - O(\delta_n^{-1/2} n^{-3/4}).$$

Applying the trivial upper bound $\mathbb{P}(N_H(G_m) > (1 + \delta_n)p^e(n)_v) \leq 1$ for $m \geq m_-$ we obtain from the identity (9.1) that

$$\begin{aligned} & \mathbb{P}(D_H(G_p) > \delta_n p^e(n)_v) \\ & \leq B_N(m_-) + \sum_{m=0}^{m_- - 1} b_N(m) \mathbb{P}(N_H(G_m) > (1 + \delta_n)p^e(n)_v) \\ & = \sqrt{\frac{e^2 q}{\pi p}} \exp\left(-\frac{\delta_n^2 p n^2}{4e^2 q} + \frac{((3e - 2) - (3e - 1)p)\delta_n^3 p n^2}{12e^3 q^2} - \log(n\delta_n) + o(1)\right) \\ & \quad + \sum_{m=0}^{m_- - 1} b_N(m) \mathbb{P}(N_H(G_m) > (1 + \delta_n)p^e(n)_v). \end{aligned}$$

It therefore suffices to prove that

$$\sum_{m=0}^{m_- - 1} b_N(m) \mathbb{P}(N_H(G_m) > (1 + \delta_n)p^e(n)_v) = o(1) \exp(-r_0(p, n, \delta_n)) \quad (9.6)$$

where we have set

$$r_0(p, n, \delta_n) := \frac{\delta_n^2 p n^2}{4e^2 q} + \frac{((3e - 2) - (3e - 1)p)\delta_n^3 p n^2}{12e^3 q^2} - \log(n\delta_n).$$

We bound the sum by showing that the contribution of terms with $m \leq m_- - n$ is small and by dividing the terms $m_- - n \leq m < m_-$ into intervals. If we write $m = m_* - f$, we may calculate that

$$\begin{aligned} L_H(m) &= \frac{(n)_v(m)_e}{(N)_e} \\ &\leq \frac{(n)_v(m_* - f)^e}{N^e} \\ &\leq \frac{(n)_v p^e(m_*^e - f m_*^{e-1})}{N^e} \\ &= (1 + \delta_n)p^e(n)_v - \Omega(f n^{v-2}). \end{aligned}$$

It follows that the event $N_H(G_m) > (1 + \delta_n)p^e(n)_v$ corresponds to a deviation $D_H(G_m) > \Omega(f n^{v-2})$ which has probability at most

$$\exp(-\Omega(f^2/n)) \quad (9.7)$$

by Theorem 1.7.

This gives us immediately that the contribution to (9.6) of terms with $m \leq m_- - n$ is at most $e^{-\Omega(n)}$ which is certainly $o(1) \exp(-r(p, n, \delta_n))$, as required.

We split the remaining values of m into intervals $I_k := \{m_{k+1}, \dots, m_k - 1\}$ where $m_k := \lfloor m_* - k\delta_n^{-1/2} n^{1/4} \rfloor$. To complete the proof it clearly suffices to show that

$$\sum_{k=1}^{\lceil \delta_n^{1/2} n^{3/4} \rceil} B_N(m_{k+1}) \mathbb{P}(N_H(G_{m_k}) > (1 + \delta_n)p^e(n)_v) = o(1) \exp(-r(p, n, \delta_n)).$$

To bound $B_N(m_{k+1})$ we calculate that $x_k := x(m_k)$ satisfies

$$x_k = \frac{\delta_n p^{1/2} n}{e\sqrt{2}q^{1/2}} - \frac{\delta_n^2(e-1)p^{1/2}n}{2\sqrt{2}e^2q^{1/2}} - O(k\delta_n^{-1/2}n^{-3/4}),$$

and so working as in the proof of Lemma 9.1 we obtain

$$B_N(m_k) \leq O(1) \exp\left(-r_0(p, n, \delta_n) + O(k\delta_n^{1/2}n^{1/4})\right).$$

On the other hand, we have from (9.7) that

$$\mathbb{P}(N_H(G_{m_k}) > (1 + \delta_n)p^e(n)_v) = \exp\left(-\Omega(k^2\delta_n^{-1}n^{-1/2})\right).$$

It follows that

$$\begin{aligned} & \sum_{k=1}^{\lceil \delta_n^{1/2}n^{3/4} \rceil} B_N(m_{k+1}) \mathbb{P}(N_H(G_{m_k}) > (1 + \delta_n)p^e(n)_v) \\ &= O(1) \sum_{k=1}^{\lceil \delta_n^{1/2}n^{3/4} \rceil} \exp\left(-r_0(p, n, \delta_n) + O((k+1)\delta_n^{1/2}n^{1/4}) - \Omega(k^2\delta_n^{-1}n^{-1/2})\right) \\ &\leq O(1) \sum_{k=1}^{\infty} \exp\left(-r_0(p, n, \delta_n) + k\delta_n^{-1}n^{-1/2}(O(\delta_n^{3/2}n^{3/4}) - \Omega(k))\right) \\ &\leq O(1) \sum_{k=1}^{\infty} \exp\left(-r_0(p, n, \delta_n) - \Omega(k^2\delta_n^{-1}n^{-1/2})\right) \\ &= o(1) \exp(-r_0(p, n, \delta_n)). \end{aligned}$$

9.3. Proof of Theorem 1.11. Let the sequence $n^{-1} \ll \delta_n \ll 1$ be given.

We set

$$r(p, n, \delta_n) := \frac{x_*^2}{2} + E(x_*, N) - \frac{\delta_n^2 n}{16\gamma_H(p)e^4 p^{2e-2} q^2}.$$

Our aim is to prove that

$$\mathbb{P}(D_H(G_{n,p}) > \delta_n p^e(n)_v) = \exp\left(-r(p, n, \delta_n) + o(\delta_n^2 n) + O(\log n)\right).$$

Since we have included a $O(\log n)$ error term in the exponent, which is equivalent to a multiplicative factor of $n^{O(1)}$, the sum given in (9.1) is equivalent to its largest term, and so it suffices to prove that

$$\max_m b_N(m) \mathbb{P}(N_H(G_m) > (1 + \delta_n)p^e(n)_v) = \exp\left(-r(p, n, \delta_n) + o(\delta_n^2 n) + O(\log n)\right).$$

The maximum is achieved with m slightly smaller than m_* . We explore values of m of the form $m_* - f$.

It will be useful to isolate a subset of the terms of $r(p, n, \delta_n)$, we set

$$s(p, n, \delta_n) = \frac{x_*^2}{2} + E(x_*, N),$$

and note that

$$b_N(m_*) = \exp\left(-s(p, n, \delta_n) + O(\log n)\right).$$

Let us calculate expressions for $b_N(m)$ and $\mathbb{P}(N_H(G_m) > (1 + \delta_n)p^e(n)_v)$ for $m = m_* - f$. We note that

$$x(m) = x_* - f\sigma^{-1}$$

where $\sigma = \sqrt{Npq}$, and so, by Theorem 1.13, we have

$$\begin{aligned} b_N(m) &= \exp\left(-\frac{(x_* - f\sigma^{-1})^2}{2} - E(x_* - f\sigma^{-1}, N) + O(\log n)\right) \\ &= \exp(-s(p, n, \delta_n) + (1 + o(1))fx_*\sigma^{-1} + O(\log n)). \end{aligned} \quad (9.8)$$

In order to get an expression for $\mathbb{P}(N_H(G_m) > (1 + \delta_n)p^e(n)_v)$ we first calculate

$$\begin{aligned} L_H(m) &= \frac{(n)_v(m)_e}{(N)_e} \\ &= \frac{(n)_v(m_* - f)^e}{N^e} + O(n^{v-2}) \\ &= (1 + \delta_n)p^e(n)_v - ef(n)_v \frac{m_*^{e-1}}{N^e} + O(n^{v-2}) \\ &= (1 + \delta_n)p^e(n)_v - (1 + o(1))\frac{ef(n)_vp^{e-1}}{N} + O(n^{v-2}) \\ &= (1 + \delta_n)p^e(n)_v - (1 + o(1))2efp^{e-1}n^{v-2} + O(n^{v-2}). \end{aligned}$$

Therefore the event $N_H(G_m) > (1 + \delta_n)p^e(n)_v$ corresponds to a deviation

$$D_H(G_m) > (1 + o(1))2efp^{e-1}n^{v-2}.$$

If $f = \Omega(n)$ then

$$\mathbb{P}(N_H(G_m) > (1 + \delta_n)p^e(n)_v) = \exp(-\Omega(f))$$

by Theorem 1.7. For $f \ll n$ we may apply Theorem 1.1 to obtain

$$\mathbb{P}(N_H(G_m) > (1 + \delta_n)p^e(n)_v) = \exp\left(\frac{-(4 + o(1))\gamma_H(p)e^2f^2p^{2e-2}}{n}\right). \quad (9.9)$$

We may immediately observe that the maximum will not occur with $f = \Omega(n)$. Indeed, for such f , we get

$$\exp(-s(p, n, \delta_n) + (1 + o(1))fx_*\sigma^{-1} - \Omega(f))$$

and $fx_*/\sigma - \Omega(f) = o(f) - \Omega(f) < 0$ for all sufficiently large n . We may therefore assume $f \ll n$.

In this range ($f \ll n$), we may combine (9.8) and (9.9) to obtain

$$\begin{aligned} b_N(m) \mathbb{P}(N_H(G_m) > (1 + \delta_n)p^e(n)_v) \\ = \exp\left(-s(p, n, \delta_n) + (1 + o(1))fx_*\sigma^{-1} - \frac{(4 + o(1))\gamma_H(p)e^2f^2p^{2e-2}}{n} + O(\log n)\right). \end{aligned}$$

If $\delta_n \ll n^{-1/2}\sqrt{\log n}$ then $x_* \ll \sqrt{n \log n}$ and the maximum of the terms involving f is $o(\log n)$, and can be absorbed into the error term. If $\delta_n = \Omega(n^{-1/2}\sqrt{\log n})$ then the maximum is obtained with

$$f_* = (1 + o(1))\frac{x_*n}{8\gamma_H(p)\sigma e^2p^{2e-2}}.$$

Setting $m = m_* - f_*$, this maximum is given by

$$\begin{aligned} b_N(m) \mathbb{P}(N_H(G_m) > (1 + \delta_n)p^e(n)_v) \\ = \exp\left(-s(p, n, \delta_n) + (1 + o(1))\frac{x_*^2 n}{16\gamma_H(p)\sigma^2 e^2 p^{2e-2}} + O(\log n)\right). \end{aligned}$$

Since $x_* \sim \delta_n p^{1/2} n / \sqrt{2qe}$ and $\sigma^2 \sim pqn^2/2$ we have shown that at the maximum we have

$$\begin{aligned} b_N(m) \mathbb{P}(N_H(G_m) > (1 + \delta_n)p^e(n)_v) \\ = \exp\left(-s(p, n, \delta_n) + (1 + o(1))\frac{\delta_n^2 n}{16\gamma_H(p)e^4 p^{2e-2} q^2} + O(\log n)\right), \end{aligned}$$

as required. This completes the proof.

Acknowledgements. C.G. and S.G. were supported by EPSRC grant EP/J019496/1. C.G. was also supported by EPSRC Fellowship EP/N004833/1. S.G. was also supported by research support from PUC-Rio, CNPq bolsa de produtividade em pesquisa (Proc. 310656/2016-8) and FAPERJ Jovem cientista do nosso estado (Proc. 202.713/2018). A.S. was supported by a Leverhulme Trust Research Fellowship. The authors would like to thank José D. Alvarado for giving some helpful comments on a draft of the paper.

REFERENCES

- [1] K. Azuma, Weighted sums of certain dependent random variables, *Tohoku Math. J.* **19** (1967), no. 3, 357–367.
- [2] R. R. Bahadur, Some approximations to the binomial distribution function, *Ann. Math. Statist.* **31** (1960), no. 1, 43–54.
- [3] A.D. Barbour, M. Karoński and A. Ruciński, A central limit theorem for decomposable random variables with applications to random graphs, *J. Combin. Theory Ser. B* **47** (1989), no. 2, 125–145.
- [4] B. Bollobás, Degree sequences of random graphs, *Discrete Math.* **33** (1981), no. 1, 1–19.
- [5] B. Bollobás, *Random graphs*, Cambridge University Press, second edition, 2001.
- [6] S. Chatterjee, An introduction to large deviations for random graphs, *Amer. Math. Soc.* **53** (2016), no. 4, 617–642.
- [7] S. Chatterjee and S.R.S. Varadhan, The large deviation principle for the Erdős-Rényi random graph, *European J. Combin.* **32** (2011), no.7, 1000–1017.
- [8] A. Dembo and E. Lubetzky, A large deviation principle for the Erdős-Rényi uniform random graph, [arXiv:1804.11327 \[math.PR\]](https://arxiv.org/abs/1804.11327) (2018).
- [9] H. Döring and P. Eichelsbacher, Moderate deviations in a random graph and for the spectrum of Bernoulli random matrices, *Electron. J. Probab.* **14** (2009), 2636–2656.
- [10] H. Döring and P. Eichelsbacher, Moderate Deviations via Cumulants, *J. Theoret. Probab.* **26** (2013), no. 2, 360–385.
- [11] V. Féray, P.-L. Méliot and A. Nikeghbali, *Mod- ϕ Convergence: Normality Zones and Precise Deviations*, Springer Briefs in Probability and Mathematical Statistics (2016)
- [12] D.A. Freedman, On tail probabilities for martingales, *Ann. Probab.* **3** (1975), 100–118.
- [13] S. Janson, A functional limit theorem for random graphs with applications to subgraph count statistics, *Random Structures Algorithms* **1** (1990), 15–37.
- [14] S. Janson, Orthogonal decompositions and functional limit theorems for random graph statistics, *Mem. Amer. Math. Soc.* **111** (1994), no. 534.
- [15] S. Janson, T. Łuczak and A. Ruciński, *Random graphs*, Wiley-Interscience, 2000.
- [16] S. Janson and K. Nowicki, The asymptotic distributions of generalized U -statistics with applications to random graphs, *Probab. Theory Related Fields* **90** (1991), no. 3, 341–375.

- [17] S. Janson and A. Ruciński, The infamous upper tail, *Random Structures Algorithms* **20** (2002), no. 3, 317–342.
- [18] S. Janson and L. Warnke, The Lower Tail: Poisson Approximation Revisited, *Random Structures Algorithms* **48** (2016), no. 2, 219–246.
- [19] M. Harel, F. Mousset, W. Samotij, Upper tails via high moments and entropic stability, [arXiv:1904.08212 \[math.PR\]](https://arxiv.org/abs/1904.08212) (2019)
- [20] W. Hoeffding, Probability inequalities for sums of bounded random variables, *J. Amer. Statist. Assoc.* **58** (1963), 13–30.
- [21] K. Krokowski, A. Reichenbachs and C. Thäle, Discrete Malliavin-Stein method: Berry-Esseen bounds for random graphs and percolation, *Ann. Probab.* **45** (2017), no. 2, 1071–1109.
- [22] A. Liebenau and N.C. Wormald, Asymptotic enumeration of graphs by degree sequence, and the degree sequence of a random graph, [arXiv:1702.08373 \[math.CO\]](https://arxiv.org/abs/1702.08373) (2017).
- [23] J.E. Littlewood, On the probability in the tail of a binomial distribution, *Adv. Appl. Probab.* **1** (1969), no. 1, 43–72.
- [24] E. Lubetzky and Y. Zhao, On the variational problem for upper tails in sparse random graphs, *Random Structures Algorithms* **50** (2017), 420–436.
- [25] C. McDiarmid, Concentration. In Probabilistic methods for algorithmic discrete mathematics, *Algorithms Combin.* **16**, Springer (1998), 195–248.
- [26] C. McDiarmid, On the method of bounded differences. In *Surveys in Combinatorics*, **141** (1989), Cambridge Univ. Press, 148–188.
- [27] B.D. McKay, On Littlewood’s estimate for the binomial distribution, *Adv. Appl. Probab.* **21** (1989), no. 2, 475–478.
- [28] B.D. McKay and N.C. Wormald, Asymptotic enumeration by degree sequence of graphs of high degree, *Eur. J. Combin.* **11** (1990), no. 6, 565–580.
- [29] G. Reinert and A. Röllin, Random subgraph counts and U -statistics: multivariate normal approximation via exchangeable pairs and embedding, *J. Appl. Probab.* **47** (2010), no. 2, 378–393.
- [30] A. Röllin, Kolmogorov bounds for the Normal approximation of the number of triangles in the Erdős–Rényi random graph, [arXiv:1704.00410 \[math.PR\]](https://arxiv.org/abs/1704.00410) (2017).
- [31] A. Ruciński, When are small subgraphs of a random graph normally distributed?, *Probab. Theory Related Fields* **78** (1988), no. 1, 1–10.
- [32] V. Vu, A large deviation result on the number of small subgraphs of a random graph, *Combin. Probab. Comput.* **10** (2001), no. 1, 79–94.
- [33] L. Warnke, On the method of typical bounded differences, *Combin. Probab. Comput.* **25** (2016), no. 2, 269–299.
- [34] Y. Zhao, On the lower tail variational problem for random graphs, *Combin. Probab. Comput.* **26** (2017), no. 2, 301–320.

10. APPENDIX

We prove Theorem 1.13.

In the context in which it was presented and applied it was more natural to state the theorem for $\text{Bin}(N, p)$ where N denotes $\binom{n}{2}$. Let us revert to lower case n for the proof, so that

$$b_n(k) := \mathbb{P}(\text{Bin}(n, p) = k)$$

and

$$B_n(k) := \mathbb{P}(\text{Bin}(n, p) \geq k) .$$

As we stated before the statement of Theorem 1.13, p may either be a constant $p \in (0, 1)$ or a function $p = p_n$.

Set $\sigma_n = \sqrt{npq}$. We must prove that

$$b_n(\lfloor pn + x_n\sigma_n \rfloor) = (1 + o(1)) \frac{1}{\sqrt{2\pi\sigma_n^2}} \exp\left(-\frac{x_n^2}{2} - E(x_n, n)\right) \quad (10.1)$$

and

$$B_n(pn + x_n\sigma_n) = (1 + o(1)) \frac{1}{x_n\sqrt{2\pi}} \exp\left(-\frac{x_n^2}{2} - E(x_n, n)\right), \quad (10.2)$$

for all $1 \ll x_n \ll \sigma_n$, where

$$E(x, n) = \sum_{i=1}^{\infty} \frac{(p^{i+1} + (-1)^i q^{i+1}) x^{i+2}}{(i+1)(i+2)p^{i/2} q^{i/2} n^{i/2}}.$$

Let us also remark that if we keep track of the error terms in the proof then we obtain

$$B_n(pn + x_n\sigma_n) = \left(1 + O\left(\frac{x_n}{\sqrt{np}} + \frac{1}{x_n^2}\right)\right) \frac{1}{x_n\sqrt{2\pi}} \exp\left(-\frac{x_n^2}{2} - E(x_n, n)\right).$$

Proof of Theorem 1.13. Let us note immediately that

$$E(x_n, n, J) = E(x_n, n) + o(1)$$

in the case that $x_n \ll (pqn)^{1/2}(pqn)^{-1/(J+3)}$, and so the ‘‘Furthermore’’ statement follows immediately from the main statements.

Both of the main asymptotic identities will follow from the fact that

$$A_n(np + x_n\sigma_n)C_n(x_n) \rightarrow 1$$

where

$$A_n(k) := \frac{(k+1)q}{k+1 - (n+1)p} \binom{n}{k} p^k q^{n-k},$$

and

$$C_n(x) := x\sqrt{2\pi} \exp\left(\frac{x^2}{2} + E(x, n)\right)$$

In fact we prove that

$$A_n(np + x_n\sigma_n)C_n^+(x_n) \rightarrow 1$$

where

$$C_n^+(x) := \left(x + \sqrt{\frac{q}{np}}\right) \sqrt{2\pi} \exp\left(\frac{x^2}{2} + E(x, n)\right)$$

which is clearly equivalent as $1 \leq C_n^+(x)/C_n(x) \leq 1 + O(\sqrt{q/px}) = 1 + o(1)$.

Setting $k_n = np + x_n\sigma_n$ we observe that

$$\frac{k_n + 1}{k_n + 1 - (n+1)p} = \frac{np + x_n\sigma_n + 1}{x_n\sigma_n + q} = \frac{\sqrt{\frac{np}{q}} + x_n + \frac{1}{\sigma_n}}{x_n + \sqrt{\frac{q}{np}}},$$

and so

$$\left(x_n + \sqrt{\frac{q}{np}}\right) \frac{k_n + 1}{k_n + 1 - (n+1)p} = \sqrt{\frac{np}{q}} \left(1 + x_n \sqrt{\frac{q}{np}} + \frac{1}{np}\right).$$

It follows that

$$\begin{aligned} & A_n(k_n) C_n^+(x_n) \\ &= \binom{n}{k_n} p^{k_n} q^{n-k_n} \left(1 + x_n \sqrt{\frac{q}{np}} + \frac{1}{np}\right) \sqrt{2\pi\sigma_n^2} \exp\left(x_n^2/2 + E(x_n, n)\right). \end{aligned}$$

We have $n, k_n, n - k_n \rightarrow \infty$ and so we may apply Stirling's approximation to the three factorials involved in the binomial coefficient to obtain that the right-hand side is asymptotically equivalent to

$$\frac{\left(1 + x_n \sqrt{\frac{q}{np}} + \frac{1}{np}\right)}{\left(1 + \frac{(q-p)x_n}{\sqrt{pq}} \frac{x_n}{\sqrt{n}} - \frac{x_n^2}{n}\right)^{1/2}} \cdot \frac{\exp\left(x_n^2/2 + E(x_n, n)\right)}{\left(1 + x_n \sqrt{\frac{q}{np}}\right)^{np+x_n\sigma_n} \left(1 - x_n \sqrt{\frac{p}{nq}}\right)^{nq-x_n\sigma_n}}.$$

The first term in this product is equal to $1 + O(x_n/\sqrt{np})$. On the other hand the logarithm of the denominator of the second term in the product is

$$\begin{aligned} & (np + x_n\sigma_n) \log\left(1 + x_n \sqrt{\frac{q}{np}}\right) + (nq - x_n\sigma_n) \log\left(1 - x_n \sqrt{\frac{p}{nq}}\right) \\ &= (np + x_n\sigma_n) \sum_{i=1}^{\infty} (-1)^{i+1} \frac{1}{i} \left(\frac{q}{p}\right)^{i/2} \left(\frac{x_n}{\sqrt{n}}\right)^i - (nq - x_n\sigma_n) \sum_{i=1}^{\infty} \frac{1}{i} \left(\frac{p}{q}\right)^{i/2} \left(\frac{x_n}{\sqrt{n}}\right)^i \\ &= \sum_{j=0}^{\infty} \frac{(p^{j+1} + (-1)^j q^{j+1}) x_n^{j+2}}{(j+1)(j+2)p^{j/2}q^{j/2}n^{j/2}} \\ &= \frac{x_n^2}{2} + E(x_n, n) \end{aligned}$$

provided n is such that $x_n/\sqrt{n} < \min\{\sqrt{q/p}, \sqrt{p/q}\}$ (which is certainly the case for all n sufficiently large). Hence,

$$A_n(k_n) C_n^+(x_n) = (1 + O(x_n/\sqrt{np})) \rightarrow 1.$$

Now observe that

$$b_n(np + x_n\sigma_n) = \frac{k_n + 1 - (n+1)p}{(k_n + 1)q} A_n(k_n)$$

and that

$$\frac{k_n + 1 - (n+1)p}{(k_n + 1)q} \frac{\sigma_n}{x_n} \rightarrow 1$$

as $n \rightarrow \infty$, from which it follows that

$$b_n(np + x_n\sigma_n) = (1 + o(1)) \frac{1}{\sqrt{2\pi\sigma_n^2}} \exp\left(-\frac{x_n^2}{2} - E(x_n, n)\right).$$

Finally, by Theorem 1 of Bahadur [2], we have

$$1 \leq \frac{A_n(k_n)}{B_n(k_n)} < 1 + x_n^{-2}.$$

It follows that

$$B_n(np + x_n\sigma_n) = (1 + o(1)) \frac{1}{x_n \sqrt{2\pi}} \exp\left(-\frac{x_n^2}{2} - E(x_n, n)\right). \quad \square$$

DEPARTMENT OF STATISTICS AND LADY MARGARET HALL, UNIVERSITY OF OXFORD, 24-29 ST GILES',
OXFORD OX1 3LB, UK

E-mail address: `goldschm@stats.ox.ac.uk`

DEPARTAMENTO DE MATEMÁTICA, PUC-RIO, RUA MARQUÊS DE SÃO VICENTE 225, GÁVEA, RIO DE
JANEIRO 22451-900, BRAZIL

E-mail address: `simon@mat.puc-rio.br`

MATHEMATICAL INSTITUTE, UNIVERSITY OF OXFORD, OXFORD OX2 6GG, UK

E-mail address: `scott@maths.ox.ac.uk`