

A Counterexample to the DeMarco-Kahn Upper Tail Conjecture

Matas Šileikis* and Lutz Warnke†

9 October 2018; revised February 8, 2019

Abstract

Given a fixed graph H , what is the (exponentially small) probability that the number X_H of copies of H in the binomial random graph $G_{n,p}$ is at least twice its mean? Studied intensively since the mid 1990s, this so-called infamous upper tail problem remains a challenging testbed for concentration inequalities. In 2011 DeMarco and Kahn formulated an intriguing conjecture about the exponential rate of decay of $\mathbb{P}(X_H \geq (1 + \varepsilon)\mathbb{E}X_H)$ for fixed $\varepsilon > 0$. We show that this upper tail conjecture is false, by exhibiting an infinite family of graphs violating the conjectured bound.

1 Introduction

Understanding the distribution of subgraph counts is one of the central topics in random graph theory. Ever since the seminal paper of Erdős and Rényi [11] from 1960 it has served as a rich source of intriguing probabilistic challenges and conjectures — repeatedly stimulating the development of new insights and tools in combinatorial probability theory (in particular concentration inequalities).

In this note we focus on the tails of the number $X_H = X_H(n, p)$ of copies of a fixed graph H in the binomial random graph $G_{n,p}$, which have been intensively studied for decades. Indeed, in the 1980s the need for exponentially small tail probabilities emerged in applications, and the behaviour of the *lower tail* $\mathbb{P}(X_H \leq (1 - \varepsilon)\mathbb{E}X_H)$ was eventually resolved by the celebrated Janson’s inequality [16, 15, 24, 22]. In the early 1990s the need for also understanding the exponential decay of the *upper tail* $\mathbb{P}(X_H \geq (1 + \varepsilon)\mathbb{E}X_H)$ became evident, and since then the following ‘infamous’ upper tail problem has proven to be much more challenging than its lower tail counterpart (see [20] and [15, 25] as well as [33, Section 4.8] and [21, Problem 6.1]).

Problem 1 (Upper tail problem for subgraph counts). *Given a fixed graph H with $e_H \geq 1$ edges, determine for fixed $\varepsilon > 0$ and arbitrary $p = p(n) \in (0, 1)$ the order of magnitude of*

$$-\log \mathbb{P}(X_H \geq (1 + \varepsilon)\mathbb{E}X_H). \quad (1)$$

In 2002 Janson, Oleszkiewicz and Ruciński [18] finally determined the exponential rate of decay (1) up to a factor of $O(\log(1/p))$. This breakthrough solved Problem 1 for constant edge-probabilities $p \in (0, 1)$, but closing the logarithmic gap for $p = o(1)$ remained an elusive technical challenge.

Shortly after the upper tail problem was settled for triangles $H = K_3$ [5, 8], in 2011 DeMarco and Kahn solved the more general case of fixed-size cliques $H = K_r$ with $r \geq 3$ [9], and also formulated a plausible *conjecture* on the general solution of Problem 1; see Conjecture 1 below. This ‘upper tail conjecture’ has been verified for large $p = p(n)$ of form $p \geq n^{-\delta_H}$ via large deviation machinery [6, 23, 2, 10], and for small $p = p(n)$ of form $p \leq n^{-v/e}(\log n)^{C_H}$ for so-called strictly balanced graphs H [31, 28, 34] (where $e_F/v_F < e_H/v_H$ for any non-empty $F \subsetneq H$); see also [1, 27, 28, 30] for further supporting results. In fact, this conjecture was also described as ‘likely to be true’ in the recent random graphs book by Frieze and Karoński [12, Section 5.4].

*Department of Theoretical Computer Science, Institute of Computer Science of the Czech Academy of Sciences, 182 07 Prague, Czech Republic. E-mail: matas.sileikis@gmail.com. With institutional support RVO:67985807. Research supported by the Czech Science Foundation, grant number GJ16-07822Y.

†School of Mathematics, Georgia Institute of Technology, Atlanta GA 30332, USA. E-mail: warnke@math.gatech.edu. Research partially supported by NSF Grant DMS-1703516 and a Sloan Research Fellowship.

In this note we show that the 7-year-old DeMarco–Kahn upper tail conjecture for subgraph counts is *false*, by exhibiting an infinite family of graphs which violate the conjectured behavior of the upper tail (1); see Theorem 1 below. On a conceptual level, our results shed new light on the upper tail behaviour for small edge-probabilities $p = p(n)$, indicating that close to the threshold of appearance the reason for having ‘too many’ copies of H can be more complicated than previously anticipated (see Sections 1.2 and 4). In retrospect this might perhaps not seem so surprising, taking into account that at the appearance threshold the limiting distribution of X_H can be quite complicated, as discovered in the 1980s [3, 14, 4, 26].

1.1 Main result

Turning to the details, we now formally state¹ the upper tail conjecture from [9], which proposes a compelling solution to Problem 1. Let $\mu_H := \mathbb{E}X_H$, $\sigma_H^2 := \text{Var} X_H$, and $m_H := \max_{F \subseteq H: v_F \geq 1} e_F/v_F$, as usual. As pointed out in [18, 9, 27, 12], to avoid degenerate behaviour of the upper tail it is natural and convenient to assume (i) that $p = p(n)$ is above the appearance threshold n^{-1/m_H} of H , and (ii) that $(1 + \varepsilon)\mathbb{E}X_H$ is at most the number of copies of H in the complete graph K_n , which is equivalent to $(1 + \varepsilon)p^{e_H} \leq 1$.

Conjecture 1 (DeMarco and Kahn, 2011). *Let H be a graph with $e_H \geq 1$ edges. For fixed $\varepsilon > 0$ and any $p = p(n)$ with $n^{-1/m_H} < p \leq (1 + \varepsilon)^{-1/e_H}$ we have*

$$-\log \mathbb{P}(X_H \geq (1 + \varepsilon)\mathbb{E}X_H) = \Theta\left(\min\{\Phi_H, M_H \log(1/p)\}\right), \quad (2)$$

where the implicit constants in (2) may depend on ε and H , with

$$\Phi_H = \Phi_H(n, p) := \min_{G \subseteq H: e_G \geq 1} \mu_G, \quad (3)$$

$$M_H = M_H(n, p) := \begin{cases} \min_{G \subseteq H: e_G \geq 1} \mu_G^{1/\alpha_G^*} & \text{if } p < n^{-1/\Delta_H}, \\ n^2 p^{\Delta_H} & \text{if } p \geq n^{-1/\Delta_H}, \end{cases} \quad (4)$$

where Δ_G is the maximum degree of G , and α_G^* is the fractional independence number² of G .

One conceptual contribution of the above conjecture was to enhance the exponent (2) by the Φ_H term, whose inclusion only matters for $p \leq n^{-1/m_H} (\log n)^{O(1)}$ unless $\Delta_H = 1$ holds (cf. [18, Remark 8.3] and Section 1.2).

Our main result shows that Conjecture 1 is false, by proving that there are infinitely many graphs H which violate the conjectured exponential rate of decay (2) close to the appearance threshold n^{-1/m_H} .

Theorem 1 (Counterexamples to Conjecture 1). *There is an infinite family \mathcal{H} of graphs such that the following holds for any $H \in \mathcal{H}$. There exists a constant $c_H > 0$ such that for fixed $\varepsilon > 0$ and any $p = p(n) \in [0, 1]$ with $n^{-1/m_H} \ll p \ll n^{-1/m_H} (\log n)^{c_H}$ we have*

$$-\log \mathbb{P}(X_H \geq (1 + \varepsilon)\mathbb{E}X_H) = o\left(\min\{\Phi_H, M_H \log(1/p)\}\right). \quad (5)$$

Remark 2. *Given $H \in \mathcal{H}$, for fixed $\varepsilon > 0$ and any $p = p(n) \in [0, 1]$ with $n^{-1/m_H} \ll p \ll 1$ we also have*

$$-\log \mathbb{P}(X_H \geq (1 + \varepsilon)\mathbb{E}X_H) = o(\Phi_H). \quad (6)$$

The proof of Theorem 1 and Remark 2 is given in Section 2. As we shall see, it uses the family of graphs \mathcal{H} illustrated in Figure 1, which are all connected and balanced (i.e., satisfy $e_H/v_H = m_H$). Remark 2 demonstrates that their upper tail probabilities are significantly larger than the lower tail probabilities for virtually all edge-probabilities p of interest, since [15, 22] gives under analogous assumptions that

$$-\log \mathbb{P}(X_H \leq (1 - \varepsilon)\mathbb{E}X_H) = \Theta(\Phi_H).$$

We find this complete separation of the decay of the two tails conceptually interesting (it was previously only known a bit above the appearance threshold, i.e., for $n^{-1/m_H} \log n \ll p \ll 1$; see [18, Remark 8.3]).

¹In the spirit of earlier questions and examples in the area (see, e.g., [32, Section 4] and [21, Section 6]), DeMarco and Kahn formally stated [9, Conjecture 10.1] for $\varepsilon = 1$ only, tacitly assuming the necessary condition $p \leq (1 + \varepsilon)^{-1/e_H}$. The natural variant (2) for arbitrary fixed $\varepsilon > 0$ is of course also attributed to them; cf. [12, Section 5.4] and [8, Section 4]. In (4) we also use a simplified (but up to constant factors equivalent) definition of M_H ; cf. [9, (46)] and [18, Theorem 1.5 and Remark 1.6].

²The fractional independence number is defined as $\alpha_G^* := \max \sum_{v \in V(G)} f(v)$, where maximum is taken over all functions $f : V(G) \rightarrow [0, 1]$ satisfying $f(u) + f(v) \leq 1$ for every edge $\{u, v\} \in E(G)$; see, e.g., [18, Appendix A].

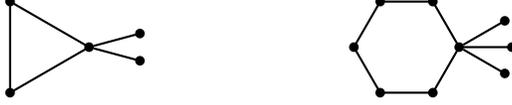


Figure 1: Examples of the graph C_{ℓ}^{+r} with $(\ell, r) = (3, 2)$ and $(\ell, r) = (6, 3)$, obtained by attaching r pendant edges to some vertex of an ℓ -vertex cycle. Theorem 1 shows that any graph in $\mathcal{H} := \{C_{\ell}^{+r} : \ell \geq 3, r \geq 2\}$ is a counterexample to the DeMarco–Kahn upper tail conjecture (see Section 2 for the full details).

1.2 Discussion

Conjecture 1 can be interpreted as an educated guess to a variant of the following question: what is the most likely way to get at least $(1 + \varepsilon)\mathbb{E}X_H$ copies of H in $G_{n,p}$? Indeed, as we shall see, it is based on two different mechanisms that each enforce $X_H \geq (1 + \varepsilon)\mathbb{E}X_H$ in $G_{n,p}$, giving two lower bounds with exponents $M_H \log(1/p)$ and Φ_H . Hence (2) intuitively predicts that the dominating (more likely) mechanism determines the exponential decay of the upper tail, ignoring constant factors in the exponent.

The first *clustered mechanism* is based on the idea that suitable ‘local’ clustering of the edges can enforce many copies of H (e.g., a clique K_z contains $\binom{z}{3} > 2\binom{n}{3}p^3$ triangles for suitable $z \asymp np$). In particular, if $F \subseteq K_n$ contains at least $(1 + \varepsilon)\mathbb{E}X_H$ copies of H , then by simply enforcing $F \subseteq G_{n,p}$ we obtain

$$\mathbb{P}(X_H \geq (1 + \varepsilon)\mathbb{E}X_H) \geq \mathbb{P}(F \subseteq G_{n,p}) = p^{e_F}.$$

Janson, Oleszkiewicz and Ruciński [18] noted that one does not need to directly enforce copies of H : it is enough if $F \subseteq K_n$ contains unusually many copies of some subgraph $J \subseteq H$ (say at least $2(1 + \varepsilon)\mathbb{E}X_J$ many), since after planting $F \subseteq G_{n,p}$ the rare upper tail event $\{X_H \geq (1 + \varepsilon)\mathbb{E}X_H\}$ becomes ‘typical’. By minimizing the number e_F of edges over all such special graphs $F \subseteq K_n$, this eventually gives a lower bound of form

$$\mathbb{P}(X_H \geq (1 + \varepsilon)\mathbb{E}X_H) \geq \max_{F \subseteq K_n} \mathbb{P}(X_H \geq (1 + \varepsilon)\mathbb{E}X_H \mid F \subseteq G_{n,p}) p^{e_F} \geq p^{\Theta(M_H)}, \quad (7)$$

see [18, Theorems 1.5 and 3.1] for the full details. This explains the exponent $M_H \log(1/p)$ in (2).

The second *disjoint mechanism* is based on many mutually exclusive ‘global’ configurations of the edges, which each contain many disjoint copies of H . Let $\mathcal{D}_{H,\varepsilon}$ denote the event that $G_{n,p}$ contains exactly $k := \lceil (1 + \varepsilon)\mathbb{E}X_H \rceil$ disjoint copies of H (either vertex-disjoint or edge-disjoint), and write $N = N(n, H)$ for the number of H -copies in K_n . Summing over distinct k -sets of disjoint H -copies, for strictly-balanced graphs this eventually gives a binomial-like lower bound in some range, which turns out to be roughly of form

$$\mathbb{P}(X_H \geq (1 + \varepsilon)\mathbb{E}X_H) \geq \mathbb{P}(\mathcal{D}_{H,\varepsilon}) \approx \binom{N}{k} \cdot p^{ke_H} \cdot (1 - p^{e_H})^{N-k} = \exp\{-\Theta(\mu_H)\},$$

see [9, 27, 28] for the full details. As in the clustered mechanism, it turns out that we can again optimize the resulting bound over all relevant subgraphs $J \subseteq H$, eventually leading to a lower bound of form

$$\mathbb{P}(X_H \geq (1 + \varepsilon)\mathbb{E}X_H) \geq \exp\{-\Theta(\Phi_H)\}, \quad (8)$$

see [27, Theorem 4.4] for the full details (note that, due to subgraphs consisting of a single edge, the optimization leading to (8) includes the mechanism which is based on enforcing $G_{n,p}$ to have ‘too many edges’). This explains the exponent Φ_H in (2). Alternatively, by combining the intuition that X_H should have *subgaussian tails* (in some range) with the standard variance estimate

$$\sigma_H^2 \asymp \mu_H^2 / \Phi_H \quad (9)$$

from [17, Lemma 3.5], for fixed $\varepsilon > 0$ we again (heuristically) arrive at an exponent of order $(\varepsilon\mu_H)^2 / \sigma_H^2 \asymp \Phi_H$.

A key message of Theorem 1 and our counterexamples from Sections 2–3 is that for some graphs H there is a yet another mechanism (which we tentatively call *locally-disjoint mechanism*), whose lower bound can beat both aforementioned mechanisms close to the appearance threshold n^{-1/m_H} . Remark 2 also shows that for certain graphs the disjoint mechanism (with exponent Φ_H) never wins, complementing the known fact that the clustered mechanism (with exponent $M_H \log(1/p)$) never wins for matchings [18, 9].

1.3 Organization

The remainder of this note is organized as follows. In Section 2 we prove Theorem 1 and Remark 2, i.e., present a simple set of counterexamples and describe how they contradict the upper tail conjecture. In Section 3 we elaborate the basic idea: we describe a larger set of counterexamples, and also give a new lower bound for the upper tail. The final Section 4 contains some concluding remarks and conjectures.

2 Simple counterexamples: Proof of Theorem 1 and Remark 2

In this section we prove Theorem 1 and Remark 2 by considering the graphs $H = C_\ell^{+r}$ illustrated in Figure 1, which are constructed from an ℓ -vertex cycle C_ℓ by connecting r additional vertices to the same vertex of the cycle (so $v_H = e_H = \ell + r$). These graphs have a history of exemplifying non-trivial behaviour of subgraph counts: (i) in 1987 Janson used C_ℓ^{+2} to demonstrate that at the threshold X_H can converge to complicated distributions [14, Section 10], and (ii) in 2000 Janson and Ruciński used C_3^{+3} to demonstrate that near the threshold X_H need not always have subgaussian tails [19, Example 6.14]. As we shall see, the following auxiliary result demonstrates yet another non-trivial behaviour of the graphs $H = C_\ell^{+r}$, since the lower bound (10) will contradict Conjecture 1 (and establish Theorem 1). Note that $m_H = 1$ and $\mu_H \asymp (np)^{\ell+r}$.

Lemma 3. *Given integers $\ell \geq 3$ and $r \geq 1$, let $H := C_\ell^{+r}$ be the graph defined above. For fixed $\varepsilon > 0$ and any $p = p(n) \in [0, 1]$ with $1 \ll np \ll n^{1/(1+\ell/r)}$ we have*

$$\mathbb{P}(X_H \geq (1 + \varepsilon)\mathbb{E}X_H) \geq \exp\left\{-O\left(\mu_H^{1/r} \log(np)\right)\right\}, \quad (10)$$

where the implicit constant in (10) may depend on ε and H .

One basic strategy for proving lower bounds is to enforce $F \subseteq G_{n,p}$ for some graph F which itself contains at least $(1 + \varepsilon)\mu_H$ copies of H . For example, $F := C_\ell^{+z}$ contains $\binom{z}{r} \geq (1 + \varepsilon)\mu_H$ copies of $H = C_\ell^{+r}$ for suitable $z \asymp (\mu_H)^{1/r}$. Following [32, 18], by enforcing F on the first v_F vertices of $G_{n,p}$ we would obtain

$$\mathbb{P}(X_H \geq (1 + \varepsilon)\mu_H) \geq \mathbb{P}(F \subseteq G_{n,p}) \geq p^{e_F} \geq \exp\left\{-O\left(\mu_H^{1/r} \log(1/p)\right)\right\}.$$

Here we shall improve the $\log(1/p) \asymp \log n$ in the exponent to $\log(np)$ by enforcing F *somewhere* in $G_{n,p}$. To this end, much in the spirit of a sequential embedding idea from [3], we will below use a two-round exposure of the edges of $G_{n,p}$ to first find a ‘random’ copy of C_ℓ , which we then extend to a copy of $F = C_\ell^{+z}$. We believe that, for $r \geq 2$ and $\varepsilon = \Theta(1)$, the resulting rate of decay (10) is best possible when $np \rightarrow \infty$ slowly.

Proof of Lemma 3. Set $p_2 := p/2$, and pick $p_1 \in [p/2, p]$ such that $(1 - p_1)(1 - p_2) = 1 - p$. We expose the edges in two rounds: for $i \in [2]$ we insert each of the $\binom{n}{2}$ possible edges into \mathcal{E}_i independently with probability p_i ; their union $\mathcal{E}_1 \cup \mathcal{E}_2$ then gives $G_{n,p}$. To establish the lower bound (10), the strategy is to (i) first use the \mathcal{E}_1 -edges to find one copy G' of $G := C_\ell$, and (ii) then use the \mathcal{E}_2 -edges to extend G' to at least $\binom{z}{r} \geq (1 + \varepsilon)\mu_H$ copies of H , by enforcing that (in \mathcal{E}_2) one vertex of G' has z neighbours outside of $V(G')$, where

$$z := \left\lceil r((1 + \varepsilon)\mu_H)^{1/r} \right\rceil \asymp (np)^{1+\ell/r} = o(n).$$

Turning to the details, for step (i) let X_G^* be the number of copies of $G = C_\ell$ in \mathcal{E}_1 . Since $m_G = 1$ and $p_1 \geq p/2 \gg n^{-1}$, it is well-known (see, e.g., [17, Theorem 3.4]) that

$$\mathbb{P}(X_G^* \geq 1) = 1 - o(1). \quad (11)$$

For step (ii), we henceforth condition on the edge-set \mathcal{E}_1 , and assume that $X_G^* \geq 1$; we also fix a copy G' of C_ℓ in \mathcal{E}_1 , and one vertex $v \in V(G')$. Defining Z as the number of vertices in $[n] \setminus V(G')$ that are neighbours of v in \mathcal{E}_2 , note that $Z = z$ implies $X_H \geq \binom{Z}{r} \geq (Z/r)^r \geq (1 + \varepsilon)\mu_H$. Hence

$$\mathbb{P}(X_H \geq (1 + \varepsilon)\mu_H \mid \mathcal{E}_1) \geq \mathbb{P}(Z = z \mid \mathcal{E}_1) = \binom{n - \ell}{z} (p_2)^z (1 - p_2)^{n - \ell - z} \geq \left(\frac{np}{4z}\right)^z e^{-np} \geq (np)^{-O(z)}. \quad (12)$$

It follows that $\mathbb{P}(X_H \geq (1 + \varepsilon)\mu_H \mid X_G^* \geq 1) \geq (np)^{-O(z)}$, which together with (11) implies inequality (10). \square

It is easy to check that the exponent of (10) is of order $[(1 + \varepsilon)\mu_H]^{1/r} \log[(1 + \varepsilon)^{1/\ell} np]$. In Section 3 we will give a variant of the above argument which not only gives a better dependence on ε (when $\varepsilon \rightarrow 0$), but also applies to a significantly larger family of graphs H . We are now ready to prove Theorem 1 and Remark 2.

Proof of Theorem 1. Define $\mathcal{H} := \{C_\ell^{+r} : \ell \geq 3, r \geq 2\}$. We henceforth fix $H = C_\ell^{+r} \in \mathcal{H}$. Since every subgraph of H with fewer than ℓ vertices is acyclic, for $1 \ll np \ll n^{1/(\ell-1)}$ we have

$$\Phi_H = \min_{G \subseteq H: e_G \geq 1} \mu_G \asymp \min \left\{ \min_{2 \leq k \leq \ell-1} \{n^k p^{k-1}\}, \min_{\ell \leq k \leq \ell+r} \{n^k p^k\} \right\} \asymp (np)^\ell \gg 1. \quad (13)$$

Turning to the parameter M_H defined in (4), note that $n^2 p^{\Delta_H} \geq n$ for $p \geq n^{-1/\Delta_H}$. Since $\alpha_G^* \leq v_G \leq \ell + r$ holds by definition (cf. Footnote 2 on page 2 or [18, Appendix A]), using (13) it follows that

$$M_H \log(1/p) = \Omega(\min\{\Phi_H^{1/(\ell+r)}, n\}) \cdot \log(1/p) \gg \log n. \quad (14)$$

Using $\ell \geq 3$ and $r \geq 2$, it now is routine to check that there is a constant $c_H > 0$ such that

$$\mu_H^{1/r} \log(np) \asymp (np)^{1+\ell/r} \log(np) \ll \min\{(np)^\ell, \log n\} \quad (15)$$

for $1 \ll np \ll (\log n)^{c_H}$, which in view of (13)–(14), Lemma 3 and $m_H = 1$ implies inequality (5). \square

Proof of Remark 2. Note that the above proof shows $\mu_H^{1/r} \log(np) \ll \Phi_H$ for $1 \ll np \ll n^{1/(\ell-1)}$, so Lemma 3 implies (6) for $n^{-1/m_H} \ll p \ll n^{-1/m_H+1/\ell}$, say (recall that $m_H = 1$ and $r \geq 2$). In the remaining range of p then [18, Remark 8.3] already states that inequality (6) holds (even for $n^{-1/m_H} \log n \ll p \ll 1$). \square

3 Extensions and generalizations

In this section we generalize the lower bound construction from Section 2. First, in Section 3.1 we show that many graphs H are not only counterexamples to Conjecture 1, but also fail to have subgaussian upper tails in some range. Next, in Section 3.2 we state a new lower bound for the upper tail, which complements the two clustered/disjoint mechanism based lower bounds from Section 1.2 used in Conjecture 1. We believe that our new lower bounds will not only serve as a testbed for future refinements of the upper tail conjecture (cf. Section 4), but also stimulate the development of new upper bounds (here the importance of having non-trivial lower bounds was already highlighted by Vu [33, Section 4.8] more than 15 years ago).

To state our results, we now introduce some terminology on the structure of the graph H . We say that a subgraph $G \subseteq H$ is *primal (for H)* if $e_G/v_G = m_H$. Clearly all primal subgraphs are induced, and thus we can treat them as a family \mathcal{L}_H of subsets of $V(H)$; see Claim 7 below for further properties. We say that G_2 *covers* a primal G_1 if $G_1 \subsetneq G_2$ and there is no further primal F with $G_1 \subsetneq F \subsetneq G_2$.

3.1 Further counterexamples and a general lower bound construction

The first inequality (16) of the following result generalizes Theorem 1, by showing that many graphs H violate Conjecture 1. The second inequality (17) conceptually generalizes Remark 2, by showing that the upper tail of these graphs is also not of a subgaussian type, no matter how close $p = p(n)$ is to the appearance threshold n^{-1/m_H} (even if we allow $\varepsilon \rightarrow 0$ reasonably slowly; for $H = C_3^{+3}$ this was already shown in [19, Example 6.14]). The assumption $\lambda\sigma_H \leq t = O(\mu_H)$ below means that we are considering large deviations, i.e., deviations that are of higher order than the standard deviation $\sigma_H = \sqrt{\text{Var } X_H}$.

Theorem 4. *Suppose that there is $G \in \mathcal{L}_H$ and distinct $J_1, \dots, J_r \in \mathcal{L}_H$ covering G , such that $K := J_1 \cup \dots \cup J_r$ satisfies $v_K/r < \min_{F \in \mathcal{L}_H} v_F$. Then there are constants $c_H, \beta_H > 0$ such that the following holds. For fixed $\varepsilon > 0$ and any $p = p(n) \in [0, 1]$ with $1 \ll np^{m_H} \ll (\log n)^{c_H}$ we have*

$$-\log \mathbb{P}(X_H \geq (1 + \varepsilon)\mathbb{E}X_H) = o\left(\min\{\Phi_H, M_H \log(1/p)\}\right). \quad (16)$$

Furthermore, there is $\lambda = \lambda(n, p, H) \gg 1$ with $\lambda\sigma_H \ll \mu_H$ such that, whenever $1 \ll np^{m_H} \ll n^{\beta_H}$ and $\lambda\sigma_H \leq t = O(\mu_H)$ holds, we have

$$-\log \mathbb{P}(X_H \geq \mathbb{E}X_H + t) = o(t^2/\sigma_H^2). \quad (17)$$

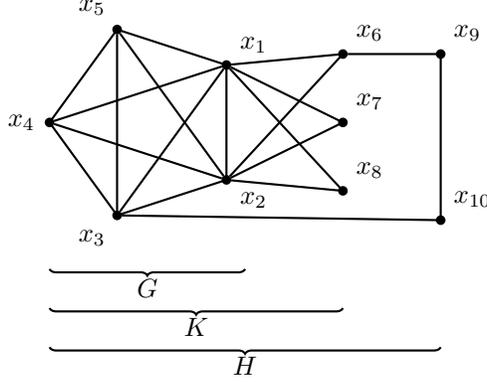


Figure 2: Example of a graph with $r = 3$, $G = H[\{x_1, \dots, x_5\}]$, $J_i = H[V(G) \cup \{x_{5+i}\}]$, and $K = J_1 \cup J_2 \cup J_3$, where $m(H) = e_G/v_G = e_{J_i}/v_{J_i} = e_K/v_K = 2$ and $e_H/v_H = 19/10 < 2$. Theorem 4 shows that this graph is another counterexample to the DeMarco–Kahn upper tail conjecture (and also fails to have subgaussian upper tails in some range).

Before giving the proof, we first use Theorem 4 to argue that counterexamples to Conjecture 1 are abundant, by describing an abstract way of generating them. Suppose that we have a balanced graph J and a primal subgraph G with the property that J covers G (using as illustration Figure 2, consider $G \cong K_5$ and construct J by connecting two vertices of G to a common outside neighbour). Then we construct K by ‘gluing’ r distinct copies J_1, \dots, J_r of J in a consistent way³ onto $G = \bigcap_{i \in [r]} J_i$ (see Figure 2 for an example with $r = 3$). Let P be a primal of J with the minimum number of vertices (in Figure 2 we have $P = G$). The resulting graph K is easily seen⁴ to (i) be balanced with density m_J , and (ii) have no primal with fewer vertices than P . If $v_J - v_G < v_P$ holds, then $v_K/r = v_J - v_G + v_G/r < v_P$ for sufficiently large r , in which case Theorem 4 implies that $H := K$ is a counterexample to Conjecture 1 (in fact, this is true for any graph $H \supseteq K$ for which P remains a vertex-minimal primal, as in Figure 2).

We shall prove Theorem 4 as a corollary of the following more general result, which qualitatively extends Lemma 3 to any graph H that is not strictly balanced (and also allows for $\varepsilon \rightarrow 0$). Here we are again considering large deviations, since by (9) the assumption $\varepsilon^2 \Phi_H \gg 1$ is equivalent to $\varepsilon \mathbb{E}X_H \gg \sqrt{\text{Var } X_H}$.

Lemma 5. *For any graph H with $e_H \geq 1$ there is a constant $\beta_H > 0$ such that the following holds for all $\varepsilon = \varepsilon(n) > 0$ and $p = p(n) \in [0, 1]$ with $\varepsilon^2 \Phi_H \gg 1$, $\varepsilon = O(1)$, and $1 \ll np^{m_H} \leq n^{\beta_H}$. If $G \in \mathcal{L}_H$ and distinct $J_1, \dots, J_r \in \mathcal{L}_H$ cover G , then we have, writing $K := J_1 \cup \dots \cup J_r$,*

$$\mathbb{P}(X_H \geq (1 + \varepsilon)\mathbb{E}X_H) \geq \exp\left\{-O\left((\varepsilon\mu_K)^{1/r} \log(np^{m_H})\right)\right\}, \quad (18)$$

where the implicit constant in (18) may depend on H .

Remark 6. *The proof shows that for $\varepsilon = \Theta(1)$ the condition $\varepsilon^2 \Phi_H \gg 1$ is redundant (as in Lemma 3, where $m_H = 1$), and that $\Phi_H \asymp (np^{m_H})^{\min_{F \in \mathcal{L}_H} v_F}$ holds for $1 \ll np^{m_H} \leq n^{\beta_H}$ (cf. inequalities (36)–(38)).*

Refining the proof strategy of Lemma 3, inspired by [32, 18, 22, 35] the idea is to first enforce $y = \Theta(\varepsilon\mu_K)$ copies of K via some special $F \subseteq G_{n,p}$, which we again find via two exposure rounds. Then we simultaneously (a) extend these y copies of K to $2\varepsilon\mu_H$ copies of H , and (b) also find additional $(1 - \varepsilon)\mu_H$ ‘random’ copies of H . The routine proof of the following auxiliary claim is deferred to Appendix B.

³To make the gluing precise, writing $V(G) = \{u_1, \dots, u_{v_G}\}$ and $V(J) \setminus V(G) = \{w_1, \dots, w_{v_J - v_G}\}$, the vertex-set of K consists of $V(G)$ and r new vertices $\{w_{j,1}, \dots, w_{j,r}\}$ for each $w_j \in V(J) \setminus V(G)$. The edge-set of K consists of $E(G)$ and $\{\{u_i, w_{j,k}\} : k \in [r]\}$ for every $\{u_i, w_j\} \in E(J) \setminus E(G)$ as well $\{\{w_{i,k}, w_{j,k}\} : k \in [r]\}$ for every $\{w_i, w_j\} \in E(J) \setminus E(G)$.

⁴For a formal proof of claims (i)–(ii) note that, for any $Q \subseteq K$ with $v_Q \geq 1$, using $m_{J_i} = m_J = m_G = e_G/v_G$ we have

$$e_Q = e_{Q \cap G} + \sum_{i \in [r]} (e_{G \cup (Q \cap J_i)} - e_G) \leq m_G [v_{Q \cap G} + \sum_{i \in [r]} (v_{G \cup (Q \cap J_i)} - v_G)] = m_J v_Q,$$

which holds with equality for $Q = K$ and thus establishes (i). For any primal $Q \subseteq K$ the above inequality must also hold with equality, and in view of $(Q \cap J_i) \cap G = Q \cap G$ it follows that $e_{Q \cap J_i} = e_{G \cup (Q \cap J_i)} + e_{Q \cap G} - e_G = m_J v_{Q \cap J_i}$ for all $i \in [r]$. Hence any $Q \cap J_i \neq \emptyset$ (at least one such subgraph must exist) is a primal of $J_i \cong J$, so $v_Q \geq v_{Q \cap J_i} \geq v_P$ establishes (ii).

Claim 7. *The following holds:*

- (i) For distinct $G_1, G_2 \in \mathcal{L}_H$ we have $G_1 \cup G_2 \in \mathcal{L}_H$.
- (ii) If $G \in \mathcal{L}_H$ and $J \in \mathcal{L}_H$ covers G , then the graph $J \setminus G := J[V(J) \setminus V(G)]$ is connected.
- (iii) If $G \in \mathcal{L}_H$ and distinct $J_1, \dots, J_r \in \mathcal{L}_H$ cover G , then the $J_i \setminus G$ are pairwise vertex-disjoint.

Proof-Sketch of Lemma 5. Deferring the choices of the constants $C_H \geq 1 \geq c_H > 0$, let

$$z := \left\lceil (C_H \varepsilon \mu_K)^{1/r} \right\rceil, \quad (19)$$

$$\delta := c_H \min\{\varepsilon, 1\}. \quad (20)$$

Similarly as in the proof of Lemma 3, we expose the edges of $G_{n,p}$ in three rounds: for $i \in [3]$ we insert each of the $\binom{n}{2}$ possible edges into \mathcal{E}_i independently with probability p_i , where

$$p_1 := p_2 := \delta p \quad \text{and} \quad p_3 := 1 - \frac{1-p}{(1-p_1)(1-p_2)} = (1 - O(\delta))p. \quad (21)$$

To establish the lower bound (18), the strategy is to (i) first use the \mathcal{E}_1 -edges to find one copy G' of G . Next, we (ii) partition the remaining vertex-set $[n] \setminus V(G')$ into r sets V_1, \dots, V_r of approximately equal sizes, and use the \mathcal{E}_2 -edges to simultaneously extend G' to z copies of each J_i which (a) embed $V(J_i \setminus G)$ into V_i , and (b) are pairwise vertex-disjoint outside of $V(G')$. This clearly enforces $y := z^r$ copies of $K = J_1 \cup \dots \cup J_r$ extending G' (by Claim 7(iii) all subgraphs $J_i \setminus G$ are pairwise vertex-disjoint). Finally, we (iii) use the \mathcal{E}_3 -edges to show that we can simultaneously (a) extend $y = \Theta(C_H \varepsilon \mu_K)$ of the aforementioned special copies of K via the \mathcal{E}_3 -edges to at least $2\varepsilon \mu_H$ copies of H , and (b) also find at least $(1 - \varepsilon)\mu_H$ additional copies of H in \mathcal{E}_3 itself, so that we overall obtain $X_H \geq 2\varepsilon \mu_H + (1 - \varepsilon)\mu_H \geq (1 + \varepsilon)\mu_H$ copies of H .

While some care is needed, the technical details of the outlined steps are mostly elementary, and thus deferred to Appendix B. Here we just mention that, analogously to Lemma 3, the probability of the ‘disjoint construction’ from step (ii) again gives the main contribution to our lower bound. In particular, by a more involved variant of the ‘enforcing z neighbours’ argument from (12), the aforementioned probability of step (ii) that G' has z ‘non-overlapping extensions’ to each J_i will turn out to be (noting that $\prod_{i \in [r]} n^{v_{J_i} - v_G} p^{e_{J_i} - e_G} \asymp \prod_{i \in [r]} (\mu_{J_i} / \mu_G) \asymp \mu_K / \mu_G$ by Claim 7(iii), and that $z^r \asymp \varepsilon \mu_K$) roughly of form

$$\prod_{i \in [r]} \binom{|V_i|}{v_{J_i} - v_G} p_2^{(e_{J_i} - e_G)z} \geq \left(\prod_{i \in [r]} \frac{\Theta(n^{v_{J_i} - v_G} (\delta p)^{e_{J_i} - e_G})}{z} \right)^z \geq \left(\frac{\Theta(\prod_{i \in [r]} \delta^{e_{J_i} - e_G})}{\varepsilon \mu_G} \right)^z. \quad (22)$$

Using $\delta \asymp \varepsilon$, $\varepsilon^2 \gg 1/\Phi_H \geq 1/\mu_G$ and $\mu_G \asymp (np^{m_H})^{v_G} \gg 1$ (by primality of G), this in turn is at least

$$\left(\frac{\Theta(\varepsilon^{\sum_i (e_{J_i} - e_G) - 1})}{\mu_G} \right)^z \geq \left(\frac{1}{\Theta(1)} \right)^z \geq (np^{m_H})^{-O(z)}, \quad (23)$$

making the right-hand side of inequality (18) plausible (see Appendix B for the full details). \square

Proof of Theorem 4. Let $\omega := np^{m_H}$ and $v_0 := \min_{F \in \mathcal{L}_H} v_F$. Note that $\Phi_H \asymp \omega^{v_0} \gg 1$ by Remark 6. Since the graph $K = J_1 \cup \dots \cup J_r$ is primal by Claim 7(i), it follows easily that $\mu_K \asymp \omega^{v_K}$ (see, e.g., (36)).

We are now ready to prove (16). Since $\Phi_H = O((\log n)^{v_0 c_H})$ holds by assumption, we have $\Phi_H \ll \log n \ll M_H \log(1/p)$ for $c_H > 0$ small enough. Since $v_K/r < v_0$ holds by assumption, we also have $\mu_K^{1/r} \log(np^{m_H}) \asymp \omega^{v_K/r} \log \omega \ll \omega^{v_0} \asymp \Phi_H$, so that inequality (16) follows from Lemma 5 (as $\varepsilon^2 \Phi_H \asymp \Phi_H \gg 1$).

We next turn to (17). Pick positive $c \in (v_0/2 - (rv_0 - v_K)/(2r - 1), v_0/2)$, and define $\lambda := \omega^c$. Using the variance estimate (9) we infer $\lambda \sigma_H / \mu_H \asymp \lambda / \Phi_H^{1/2} \asymp \omega^{c - v_0/2} \ll 1$ and thus $\lambda \sigma_H \ll \mu_H$. Defining $\varepsilon := t/\mu_H = O(1)$, using (9) we also infer $\varepsilon^2 \Phi_H \asymp t^2 / \sigma_H^2 \geq \lambda^2 \gg 1$, so Lemma 5 applies. Combining $t^2 / \sigma_H^2 \asymp \varepsilon^2 \Phi_H$ and $\varepsilon \geq \lambda \sigma_H / \mu_H \asymp \omega^{c - v_0/2}$ with $\Phi_H \asymp \omega^{v_0}$ and $\mu_K \asymp \omega^{v_K}$, it follows by choice of c that, say,

$$(t^2 / \sigma_H^2)^r \asymp \varepsilon \cdot \varepsilon^{2r-1} (\Phi_H)^r \geq \varepsilon \cdot \Omega\left(\omega^{(c - v_0/2)(2r-1) + rv_0}\right) \gg \varepsilon \cdot \omega^{v_K} (\log \omega)^r \asymp \varepsilon \mu_K (\log \omega)^r.$$

This readily implies $(\varepsilon \mu_K)^{1/r} \log \omega \ll t^2 / \sigma_H^2$, which in view of (18) establishes inequality (17). \square

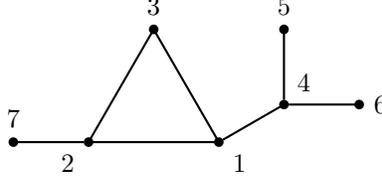


Figure 3: The snail graph H , which is balanced and satisfies $m_H = 1$. Example 9 demonstrates that no graph in the Bollobás–Wierman grading decomposition $H[123] \subset H[12347] \subset H$ minimizes $\zeta_H(G)$ in (24).

3.2 Optimizing the lower bound for the upper tail

In this subsection we optimize the lower bound (18) for the upper tail over all possible choices of G and $K = J_1 \cup \dots \cup J_r$, restricting to the important case where $\varepsilon > 0$ is fixed (as in Problem 1); see Lemma 8 below. To state our result, given $G \in \mathcal{L}_H$, let $J_1, \dots, J_{s(G)}$ be *all* primals of H which cover G , ordered by the increasing number of vertices (how the ties are broken is irrelevant for our purposes). Then, for graphs H which are not strictly balanced (which implies that there is $G \in \mathcal{L}_H$ with $s(G) \geq 1$), we define

$$\zeta_H(G) := \min_{r \in [s(G)]} \left\{ \frac{v_G + \sum_{i=1}^r (v_{J_i} - v_G)}{r} \right\} \quad \text{and} \quad \zeta_H := \min_{G \in \mathcal{L}_H} \{ \zeta_H(G) : s(G) \geq 1 \}. \quad (24)$$

Lemma 8. *For every graph H that is not strictly balanced there is a constant $\beta_H > 0$ such that the following holds. For fixed $\varepsilon > 0$ and any $p = p(n) \in [0, 1]$ with $1 \ll np^{m_H} \leq n^{\beta_H}$ we have*

$$\mathbb{P}(X_H \geq (1 + \varepsilon)\mathbb{E}X_H) \geq \exp \left\{ -O \left((np^{m_H})^{\zeta_H} \log(np^{m_H}) \right) \right\}, \quad (25)$$

where the implicit constant in (25) may depend on ε and H .

Proof. Fix arbitrary $G \in \mathcal{L}_H$ with $s(G) \geq 1$. Combining Lemma 5 with $\mu_K^{1/r} \asymp (np^{m_H})^{v_K/r} \gg 1$ (cf. the proof of Theorem 4), it suffices to show that the minimum of $v_{K_S}/|S|$ over all $S \subseteq [s(G)]$ with $S \neq \emptyset$ equals $\zeta_H(G)$, where $K_S := \cup_{i \in S} J_i$. By Claim 7(iii) the graphs J_i share no vertices except for those in $V(G)$, so

$$v_{K_S} = v_G + \sum_{i \in S} (v_{J_i} - v_G). \quad (26)$$

Recalling $v_{J_1} \leq \dots \leq v_{J_{s(G)}}$, a moment's thought reveals that the minimum is always attained by one of the sets $S \in \{[1], [2], \dots, [s(G)]\}$, which establishes $\min_S v_{K_S}/|S| = \zeta_H(G)$ and thus completes the proof. \square

It seems difficult to give a simple combinatorial description of the $G \in \mathcal{L}_H$ which minimize $\zeta_H(G)$ in (24). For balanced graphs H it is natural to first focus on the so-called ‘grading decomposition’ $\{G_0, \dots, G_s\} \subseteq \mathcal{L}_H$ of Bollobás and Wierman [4], which determines the limit distribution of X_H at the appearance threshold (i.e., when $p \sim cn^{-1/m_H}$ for some $c \in (0, \infty)$). Turning to the inductive definition of their decomposition, let G_0 be the union of minimal primal subgraphs of H . Then, given $G_i \neq H$, let G_{i+1} be the union of all primal subgraphs covering G_i . For balanced graphs H the resulting grading $G_0 \subset \dots \subset G_s$ always terminates with $G_s = H$ (and Claim 7(i) implies $G_j \in \mathcal{L}_H$). In [4] the distribution of X_H at the threshold is then determined inductively: first counting G_0 -subgraphs, then G_1 -subgraphs that contain the G_0 subgraphs, etc, continuing until all H -subgraphs are counted. Moving a tiny bit above the appearance threshold (as in Lemma 8), it thus sounds plausible that the exponential decay of $\mathbb{P}(X_H \geq (1 + \varepsilon)\mathbb{E}X_H)$ could potentially be determined by one of the ‘transitions’ from G_i to G_{i+1} , which in turn suggests that the minimum in ζ_H might perhaps be attained by some G_j . The following example shows that this speculation is false.

Example 9. Consider the graph H in Figure 3 with $v_H = 7$. Its primals (as vertex sets) are 123, 1234, 1237, 12347, 12345, 12346, 123456, 123457, 123467, and 1234567. Straightforward case checking reveals that ζ_H is attained by 1234, which is covered by the three primals 12345, 12346, and 12347, so that $\zeta_H(H[1234]) = \min\{5/1, 6/2, 7/3\} = 7/3$. However, the Bollobás–Wierman grading decomposition is $G_0 := H[123] \subset G_1 := H[12347] \subset G_2 := H$, and both $\zeta_H(G_0) = 5/2$ and $\zeta_H(G_1) = 7/2$ are suboptimal.

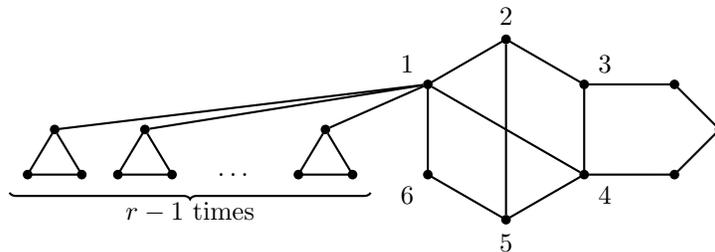


Figure 4: The graph H_r , which is balanced and satisfies $m_{H_r} = 4/3$. Theorem 10 illustrates that the upper tail behaviour of H_r is extremely complicated for $r \geq 7$ (see also Appendix A).

4 Concluding remarks

In this note we showed that the DeMarco–Kahn upper tail conjecture is false. Nevertheless we believe that its prediction is true when H is strictly balanced or $p = p(n)$ is sufficiently above the appearance threshold.

Conjecture 2. *Conjecture 1 is true for any strictly balanced graph H . Furthermore, for any fixed $\gamma > 0$, Conjecture 1 is true under the additional assumption $p \geq n^{-1/m_H + \gamma}$.*

We leave it as an intriguing open problem to formulate an upper tail conjecture for graphs which are not strictly balanced (this would already be interesting for balanced graphs). Combining the new ‘locally-disjoint mechanism’ based lower bound (25) from Lemma 8 with the previously known clustered/disjoint mechanism based lower bounds (7)–(8) from Section 1.2, it is tempting to speculate that we might perhaps have

$$-\log \mathbb{P}(X_H \geq (1 + \varepsilon)\mu_H) = \Theta\left(\min\{\Phi_H, M_H \log(1/p), (np^{m_H})^{\zeta_H} \log(np^{m_H})\}\right), \quad (27)$$

which we believe to be correct for many graphs (e.g. for the graphs C_ℓ^{+r} from Section 2). However, the following result shows that the natural guess (27) is false for the balanced graphs H_r illustrated in Figure 4, indicating that for subgraph counts a general upper tail conjecture is most likely quite complicated.

Theorem 10. *Let $\mathcal{H} := \{H_r : r \geq 7\}$. For any $H \in \mathcal{H}$ there are constants $1 > d_H > c_H > 0$ such that the following holds. For fixed $\varepsilon > 0$ and any $p = p(n) \in [0, 1]$ with $(\log n)^{c_H} \ll np^{m_H} \ll (\log n)^{d_H}$ we have*

$$-\log \mathbb{P}(X_H \geq (1 + \varepsilon)\mathbb{E}X_H) = o\left(\min\{\Phi_H, M_H \log(1/p), (np^{m_H})^{\zeta_H} \log(np^{m_H})\}\right). \quad (28)$$

The proof of inequality (28) is based on the observation that different kinds of extensions (for H_r from Figure 4 the dangling triangle and the rooted path) can have different ranges of $p = p(n)$ where the disjoint mechanism beats the clustered one, which means that in some transitional range of $p = p(n)$ a mixture of both mechanisms can potentially give better bounds (which turns out to be the case for H_r). More precisely, adapting the framework of Lemma 5 for $H = H_r$ with $G := H[123456]$ and $K := H$, after planting one copy of G here the idea is to (a) enforce z vertex-disjoint triangles which are each connected to vertex 1 of G , and (b) enforce at least z^* clustered copies of 5-vertex paths with endvertices 3, 4 of G (by planting a complete bipartite graph which connects a fixed vertex-set U of size $2\sqrt{z^*}$ with the vertex-set $\{w, 3, 4\}$, where the extra vertex $w \notin V(G) \cup U$ is also fixed). Analyzing these two mechanisms, it turns out that this way we obtain at least $\binom{z}{r-1} \cdot z^*$ copies of H_r with probability at least $(np^{m_H})^{-O(z)} \cdot p^{\Theta(\sqrt{z^*})}$, which for suitable $z \ll \mu_H^{1/r} \ll z^*$ and $p = p(n)$ eventually gives inequality (28); see Appendix A for the details.

Of course, one could augment (27) by the above-discussed new mix of the disjoint/clustered mechanisms (by adapting Lemmas 5 and 8), but we are not sure if the resulting bound would be optimal (in general).

Finally, it would also be interesting to explore if Stein’s method, large deviation theory (possibly after altering the variational problem from [7, 6, 10]), or some other probabilistic approach could yield an educated guess for the solution to the upper tail problem (Problem 1) close to the appearance threshold n^{-1/m_H} .

Acknowledgements. We thank the referees for helpful suggestions.

References

- [1] R. Adamczak and P. Wolff. Concentration inequalities for non-Lipschitz functions with bounded derivatives of higher order. *Probab. Theory Related Fields* **162** (2015), 531–586.
- [2] B. Bhattacharya, S. Ganguly, E. Lubetzky, and Y. Zhao. Upper tails and independence polynomials in random graphs. *Adv. Math.* **319** (2017), 313–347.
- [3] B. Bollobás. Threshold functions for small subgraphs. *Math. Proc. Cambridge Philos. Soc.* **90** (1981), 197–206.
- [4] B. Bollobás and J.C. Wierman. Subgraph counts and containment probabilities of balanced and unbalanced subgraphs in a large random graph. In *Graph theory and its applications: East and West (Jinan, 1986)*, Annals of the New York Academy Vol. **141**. pp. 763–70, New York Acad. Sci., New York (1989).
- [5] S. Chatterjee. The missing log in large deviations for triangle counts. *Random Struct. Alg.* **40** (2012), 437–451.
- [6] S. Chatterjee and A. Dembo. Nonlinear large deviations. *Adv. Math.* **299** (2016), 396–450.
- [7] S. Chatterjee and S.R.S. Varadhan. The large deviation principle for the Erdős-Rényi random graph. *European J. Combin.* **32** (2011), 1000–1017.
- [8] B. DeMarco and J. Kahn. Upper tails for triangles. *Random Struct. Alg.* **40** (2012), 452–459.
- [9] B. DeMarco and J. Kahn. Tight upper tail bounds for cliques. *Random Struct. Alg.* **41** (2012), 469–487.
- [10] R. Eldan. Gaussian-width gradient complexity, reverse log-Sobolev inequalities and nonlinear large deviations. *Geom. Funct. Anal.* **28** (2018) 1548–1596.
- [11] P. Erdős and A. Rényi. On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl* **5** (1960), 17–61.
- [12] A. Frieze and M. Karoński. *Introduction to random graphs*. Cambridge University Press, Cambridge (2016).
- [13] T.E. Harris. A lower bound for the critical probability in a certain percolation process. *Math. Proc. Cambridge Philos. Soc.* **56** (1960), 13–20.
- [14] S. Janson. Poisson convergence and Poisson processes with applications to random graphs. *Stochastic Process. Appl.* **26** (1987), 1–30.
- [15] S. Janson. Poisson approximation for large deviations. *Random Struct. Alg.* **1** (1990), 221–229.
- [16] S. Janson, T. Łuczak, and A. Ruciński. An exponential bound for the probability of nonexistence of a specified subgraph in a random graph. In *Random graphs '87 (Poznań, 1987)*, pp. 73–87, Wiley, Chichester (1990).
- [17] S. Janson, T. Łuczak, and A. Ruciński. *Random graphs*. *Wiley-Interscience Series in Discrete Mathematics and Optimization*. Wiley-Interscience, New York (2000).
- [18] S. Janson, K. Oleszkiewicz, and A. Ruciński. Upper tails for subgraph counts in random graphs. *Israel J. Math.* **142** (2004), 61–92.
- [19] S. Janson and A. Ruciński. The deletion method for upper tail estimates. Preprint (2000). http://www2.math.uu.se/~svante/papers/sj135_ppt.pdf
- [20] S. Janson and A. Ruciński. The infamous upper tail. *Random Struct. Alg.* **20** (2002), 317–342.
- [21] S. Janson and A. Ruciński. The deletion method for upper tail estimates. *Combinatorica* **24** (2004), 615–640.
- [22] S. Janson and L. Warnke. The lower tail: Poisson approximation revisited. *Random Struct. Alg.* **48** (2016), 219–246.
- [23] E. Lubetzky and Y. Zhao. On the variational problem for upper tails in sparse random graphs. *Random Struct. Alg.* **50** (2017), 420–436.
- [24] O. Riordan and L. Warnke. The Janson inequalities for general up-sets. *Random Struct. Alg.* **46** (2015), 391–395.
- [25] V. Rödl and A. Ruciński. Random graphs with monochromatic triangles in every edge coloring. *Random Struct. Alg.* **5** (1994), 253–270.
- [26] A. Ruciński. Small subgraphs of random graphs—a survey. In *Random graphs '87 (Poznań, 1987)*, pp. 283–303, Wiley, Chichester (1990).
- [27] M. Šileikis. *Inequalities for Sums of Random Variables: a combinatorial perspective*. PhD thesis, AMU Poznań (2012). Available from <https://sites.google.com/site/matassileikis/>
- [28] M. Šileikis. On the upper tail of counts of strictly balanced subgraphs. *Electron. J. Combin.* **19** (1) Paper 4 (2012), 14 pages.
- [29] M. Šileikis and L. Warnke. Counting extensions revisited. Manuscript (2019).
- [30] M. Šileikis and L. Warnke. Upper tail bounds for Stars. Preprint (2019). [arXiv:1901.10637](https://arxiv.org/abs/1901.10637).
- [31] V.H. Vu. On the concentration of multivariate polynomials with small expectation. *Random Struct. Alg.* **16** (2000), 344–363.
- [32] V.H. Vu. A large deviation result on the number of small subgraphs of a random graph. *Combin. Probab. Comput.* **10** (2001), 79–94.
- [33] V.H. Vu. Concentration of non-Lipschitz functions and applications. *Random Struct. Alg.* **20** (2002), 262–316.
- [34] L. Warnke. On the missing log in upper tail estimates. *J. Combin. Theory Ser. B*, to appear. [arXiv:1612.08561](https://arxiv.org/abs/1612.08561).
- [35] L. Warnke. Upper tails for arithmetic progressions in random subsets. *Israel J. Math.* **221** (2017), 317–365.

A Appendix: Proof of Theorem 10

Proof of Theorem 10. Fix $H = H_r \in \mathcal{H}$, with $v_H = 3r + 6$. Let $\omega := np^{m_H}$ and $\gamma := 1/r^3$. Define

$$c_H := \frac{2}{v_H/r - (r-1)\gamma} \quad \text{and} \quad d_H := \frac{1}{v_H/r - 2 + \gamma/2},$$

noting that $c_H < d_H$ as $\gamma < (4 - v_H/r)/r = (1 - 6/r)/r$. Since G has the smallest number of vertices among primals, we obtain $\Phi_H \asymp \omega^6$ by Remark 6. Using $1 \ll \omega \leq n^{o(1)}$ and $m_H \leq \Delta_H/2$, it is not difficult to verify that $M_H = \min_{G \subseteq H: e_G \geq 1} \mu_G^{v_G/\alpha_G^*}$ and thus $M_H \asymp \omega^{\min_{F \in \mathcal{L}_H} v_F/\alpha_F^*}$ holds (e.g., by combining (36)–(37) with $1/\alpha_F^* \in [1/v_F, 1]$). Since every $F \in \mathcal{L}_H$ is a union of G and some (possibly empty) subset of the J_i , using [18, Proposition A.4] it turns out that $\alpha_F^* = v_F/2$, so $M_H \asymp \omega^2$. It is routine to check that $\zeta_H = v_H/r = 3 + 6/r$. It follows that $\Phi_H \gg \omega^{\zeta_H} \log \omega$ and $M_H \log(1/p)/\omega^{\zeta_H} \asymp (\log n)/\omega^{v_H/r-2} \gg \log \omega$, so the minimum in (28) satisfies

$$\min\{\Phi_H, M_H \log(1/p), (np^{m_H})^{\zeta_H} \log(np^{m_H})\} = \omega^{v_H/r} \log \omega. \quad (29)$$

We are now ready to establish (28) by adapting the proof of Lemma 3, exposing the edges of $G_{n,p}$ via $\mathcal{E}_1 \cup \mathcal{E}_2$ in two independent rounds with edge-probabilities $p_2 := p/2$ and $p_1 \in [p/2, p]$. For the desired lower bound, the strategy is to (i) first use the \mathcal{E}_1 -edges to find one copy G' of $G := H[123456]$, where the vertices v_j of G' correspond to vertices j of G (see Figure 4). Next we (ii) partition the vertex-set $[n] \setminus V(G') = V_1 \cup V_2$ into two sets with $|V_i| \approx n/2$, and then use the \mathcal{E}_2 -edges to simultaneously (a) create z vertex-disjoint triangles in V_1 , which are each connected to vertex v_1 of G' , and (b) create z^* ‘clustered’ copies of a 5-vertex-path whose internal vertices are in V_2 and whose endpoints are v_3, v_4 of G' . This together enforces at least $\binom{z}{r-1} \cdot z^* > (1 + \varepsilon)\mu_H$ copies of $H = H_r$ extending G' (see Figure 4), where

$$z := \left\lceil r((1 + \varepsilon)\mu_H)^{1/r} \omega^{-\gamma} \right\rceil \asymp \omega^{v_H/r - \gamma} \quad \text{and} \quad z^* := \left\lceil ((1 + \varepsilon)\mu_H)^{1/r} \omega^{(r-1)\gamma} \right\rceil \asymp \omega^{v_H/r + (r-1)\gamma}.$$

Turning to the details, in step (i) we find with probability $1 - o(1)$ at least one copy of $G := H[123456]$ in \mathcal{E}_1 , since $m_G = 4/3 = m_H$ and $p_1 \geq p/2 \gg n^{-1/m_H}$ is above the appearance threshold. For step (ii), we henceforth condition on the edge-set \mathcal{E}_1 and fix one copy G' of G in \mathcal{E}_1 . Mimicking the calculations leading to (43)–(45) in Appendix B, it turns out that the probability of step (ii).(a) is at least

$$\frac{1}{z!} \prod_{0 \leq j < z} \left[\binom{|V_1| - 3j}{3} p_2^4 \right] \cdot \omega^{-o(z)} \geq \left(\frac{\Theta(n^3 p^4 \omega^{-o(1)})}{z} \right)^z \geq \left(\frac{\Theta(\omega^{3-o(1)})}{\omega^{v_H/r - \gamma}} \right)^z \geq \omega^{-o(\omega^{v_H/r})}, \quad (30)$$

where we used $v_H/r = 3 + 6/r$. Turning to step (ii).(b), after fixing a vertex-set $U \subseteq V_2$ of size $|U| = \lceil 2\sqrt{z^*} \rceil$ and a vertex $w \in V_2 \setminus U$, we define F as the complete bipartite graph between U and $\{v_3, v_4, w\}$. Note that the union of G' and F contains at least $\binom{|U|}{2} \geq z^*$ different 5-vertex-paths with endpoints v_3, v_4 and internal vertices from V_2 . Recalling $p_2 = n^{-1/m_H + o(1)}$, the probability of step (ii).(b) is thus at least

$$\mathbb{P}(F \subseteq \mathcal{E}_2) = p_2^{3|U|} \geq n^{-\Theta(\sqrt{z^*})} \geq \omega^{-o(\omega^{v_H/r})}, \quad (31)$$

where we used $\omega^{v_H/r}/\sqrt{z^*} \asymp \omega^{(v_H/r - (r-1)\gamma)/2} \geq \log n$. Noting that the step (ii) events lower bounded by (30)–(31) are independent (as they depend on disjoint edge-sets), it follows that $\mathbb{P}(X_H \geq (1 + \varepsilon)\mathbb{E}X_H) \geq \omega^{-o(\omega^{v_H/r})}$, which together with (29) implies inequality (28). \square

B Appendix: Proof of Lemma 5 and Claim 7

Proof of Claim 7. For property (i), using $e_{G_i}/v_{G_i} = m_H \geq e_{G_1 \cap G_2}/v_{G_1 \cap G_2}$ it routinely follows that

$$\frac{e_{G_1 \cup G_2}}{v_{G_1 \cup G_2}} = \frac{e_{G_1} + e_{G_2} - e_{G_1 \cap G_2}}{v_{G_1} + v_{G_2} - v_{G_1 \cap G_2}} \geq m_H, \quad (32)$$

which implies that $G_1 \cup G_2 \subseteq H$ is primal.

For property (iii), suppose that $J_i \setminus G$ and $J_j \setminus G$ with $i \neq j$ are not vertex-disjoint. Clearly $G \subsetneq J_i \cap J_j \subsetneq J_i$. Since J_k covers G , this implies $e_{J_i \cap J_j} / v_{J_i \cap J_j} < m_H$. Since $e_{J_k} / v_{J_k} = m_H$, analogously to (32) we infer $e_{J_i \cup J_j} / v_{J_i \cup J_j} > m_H$, reaching the desired contradiction (since $J_i \cup J_j \subseteq H$).

For property (ii), suppose that $J \setminus G$ is not connected. Then we can partition $V(J \setminus G) = V(J) \setminus V(G)$ into two non-empty vertex-sets V_j such that there are no edges between V_1 and V_2 in J . Since the graphs $F_j := J[V(G) \cup V_j]$ are not primal (as $G \subsetneq F_j \subsetneq J$), we have $e_{F_j} / v_{F_j} < m_H = e_G / v_G$. It follows that

$$\frac{e_J}{v_J} = \frac{e_{F_1} + e_{F_2} - e_G}{v_{F_1} + v_{F_2} - v_G} < m_H,$$

reaching the desired contradiction (since $J \subseteq H$ is primal). \square

Proof of Lemma 5. We keep the setup from the sketch in Section 3.1: in particular, we shall expose the edges of $G_{n,p}$ via $\mathcal{E}_1 \cup \mathcal{E}_2 \cup \mathcal{E}_3$ in three independent rounds with edge-probabilities $p_1 = p_2 = \delta p$ and $p_3 = (1 - O(\delta))p$, where $\delta = c_H \min\{\varepsilon, 1\}$ and $c_H \leq 1$. Adding an extra initial reduction step, we claim that it suffices to prove Lemma 5 for graphs $K = J_1 \cup \dots \cup J_r$ which satisfy, for all $i \in [r]$,

$$\mu_{J_i} / \mu_G \leq (\varepsilon \mu_K)^{1/r}. \quad (33)$$

To see that this implies Lemma 5 for arbitrary $K = J_1 \cup \dots \cup J_r$, we use induction on the number of J_1, \dots, J_r (formally allowing the implicit constant in inequality (18) to depend on $1 \leq r \leq v_H$). The base case $r = 1$ is immediate, since (33) always holds due to $(\varepsilon \mu_K)^{1/r} = \varepsilon \mu_G \cdot (\mu_{J_1} / \mu_G)$ and $\varepsilon \mu_G \geq \varepsilon \Phi_H \gg \varepsilon^{-1} = \Omega(1)$. For $r \geq 2$ it suffices to consider the case where (33) fails for some $i \in [r]$. Set $K' := \bigcup_{j \neq i} J_j$. Applying induction (with K replaced by K' , and thus r replaced by $r - 1$), the lower bound (18) holds with $(\varepsilon \mu_{K'})^{1/(r-1)} \log(np^{m_H})$ in the exponent. It thus remains to check that

$$(\varepsilon \mu_{K'})^{1/(r-1)} = O((\varepsilon \mu_K)^{1/r}). \quad (34)$$

Using Claim 7(iii) we obtain $\mu_K \asymp \mu_{K'} \cdot \mu_{J_i} / \mu_G$. Since we assumed that (33) fails (i.e., that $\mu_{J_i} / \mu_G > (\varepsilon \mu_K)^{1/r}$ holds) we infer $\varepsilon \mu_{K'} \asymp \varepsilon \mu_K \cdot \mu_G / \mu_{J_i} = O((\varepsilon \mu_K)^{1-1/r})$ and thus establish (34), completing the proof of the claimed reduction.

To facilitate our three-step proof strategy, we henceforth assume that (33) holds for all $i \in [r]$. Furthermore, we fix an ordering u_1, \dots, u_{v_H} of the vertices of H such that the first v_G vertices are vertices of G , the following $v_{J_1} - v_G$ vertices are vertices of $J_1 \setminus G$, followed by the vertices of $J_2 \setminus G$, and so on up to $J_r \setminus G$ (this is possible since the subgraphs $J_i \setminus G$ are pairwise vertex-disjoint, see Claim 7(iii)), while the final $v_H - v_K$ vertices are the remaining vertices of $H \setminus K$. We also introduce the shorthand notation

$$\omega := np^{m_H} \quad \text{with} \quad 1 \ll \omega \leq n^{\beta_H}. \quad (35)$$

We assume $\beta_H < 1/v_H$, so that every primal subgraph $F \subseteq H$ satisfies

$$\mu_F \asymp (np^{e_F/v_F})^{v_F} = (np^{m_H})^{v_F} = \omega^{v_F} \leq n^{v_F \beta_H} \ll n. \quad (36)$$

Furthermore, for any non-primal subgraph $F \subseteq H$ we have $B_{F,H} := m_H - e_F/v_F > 0$, so that, say,

$$\mu_F \asymp (np^{m_H} \cdot p^{-(m_H - e_F/v_F)})^{v_F} \geq (\omega \cdot n^{B_{F,H}(1-\beta_H)/m_H})^{v_F} \gg n^{2v_H^2 \beta_H} \geq \omega^{2v_H^2} \quad (37)$$

for $\beta_H > 0$ small enough (the ad hoc $2v_H^2$ -term is convenient later on). From (35)–(37) we easily deduce

$$\Phi_G \geq \Phi_H \gg 1. \quad (38)$$

Using $\varepsilon^2 \Phi_H \gg 1$ and (36) we obtain

$$\delta \asymp \min\{\varepsilon, 1\} \gg (\Phi_H)^{-1/2} \geq (\mu_G)^{-1/2} = \Omega(\omega^{-v_G/2}). \quad (39)$$

Finally, recalling the definition (19) of z , note that $\varepsilon^2 \Phi_H \gg 1$ and $\varepsilon = O(1)$ imply $z^r \asymp \varepsilon \mu_K \geq \varepsilon \Phi_H \gg \varepsilon^{-1} = \Omega(1)$ and $z^r = O(\mu_K)$. Since $K \subseteq H$ is primal (by Claim 7(i)), using (36) it follows that

$$1 \ll z = O(\omega^{v_K/r}) \ll n^{1/r}. \quad (40)$$

Turning to the technical details of step (i), let X_G^* be the number of copies of G in \mathcal{E}_1 . We claim that

$$\mathbb{P}(X_G^* \geq 1) \gg \omega^{-v_G e_G}. \quad (41)$$

For the proof we use a version of the Paley–Zygmund inequality (see, e.g., [17, (3.3)–(3.4)]) and the standard estimate $\text{Var } X_G^*/(\mathbb{E}X_G^*)^2 \asymp 1/\Phi_G(n, p_1)$ (see, e.g., [17, Lemma 3.5]), so that $p_1 = \delta p$ and $\delta \leq 1$ imply

$$\mathbb{P}(X_G^* \geq 1) \geq \frac{(\mathbb{E}X_G^*)^2}{(\mathbb{E}X_G^*)^2 + \text{Var } X_G^*} \asymp \min\{1, \Phi_G(n, p_1)\} \geq \min\{1, \delta^{e_G} \Phi_G\}.$$

Now inequality (41) follows, since $\delta \gg \omega^{-v_G/2}$ by (39) and $\Phi_G \gg 1$ by (38).

For step (ii), we henceforth condition on the edge-set \mathcal{E}_1 , and assume that $X_G^* \geq 1$. We also fix an *ordered* copy G' of G in \mathcal{E}_1 , i.e., a copy of G with $E(G') \subseteq \mathcal{E}_1$ and an ordering u'_1, \dots, u'_{v_G} of $V(G')$ that is consistent with the above-fixed ordering u_1, \dots, u_{v_G} of G (i.e., the injection $u_j \mapsto u'_j$ maps edges of $E(G)$ into edges of \mathcal{E}_1). We partition $[n] \setminus V(G')$ into r vertex-sets V_1, \dots, V_r of approximately equal sizes $n_i := |V_i| \approx n/r$. We say that an $(e_{J_i} - e_G)$ -element edge-set $\mathcal{S} \subseteq \binom{V_i \cup V(G')}{2} \setminus \binom{V(G')}{2}$ is an (G', J_i) -*edge-extension* if there is an injection from $V(J_i)$ to $W(\mathcal{S}) := V(G') \cup \bigcup_{f \in \mathcal{S}} f$ with $u_j \mapsto u'_j$ for $j \in [v_G]$ that maps every edge $E(J_i) \setminus E(G)$ to an edge in \mathcal{S} (this definition makes sense since $J_i \setminus G = J_i[V(J_i) \setminus V(G)]$ contains no isolated vertices, see Claim 7(ii)). Note that $|W(\mathcal{S}) \setminus V(G')| = v_{J_i} - v_G$, and that $\mathcal{S} \cup E(G')$ corresponds to (the edge-set of) a copy of J_i which contains G' . Let Z_{G', J_i} be the number of (G', J_i) -edge-extensions $\mathcal{S} \subseteq \mathcal{E}_2$. Noting that the random variables $Z_{G', J_i}, i \in [r]$ depend on disjoint sets of independent \mathcal{E}_2 -edges, we infer

$$\mathbb{P}(Z_{G', J_i} = z \text{ for all } i \in [r] \mid \mathcal{E}_1) = \prod_{i \in [r]} \mathbb{P}(Z_{G', J_i} = z \mid \mathcal{E}_1). \quad (42)$$

Fix $i \in [r]$. We claim that

$$\mathbb{P}(Z_{G', J_i} = z \mid \mathcal{E}_1) \geq \omega^{-O(z)}. \quad (43)$$

The following proof of (43) is fairly standard (similar to, e.g., [9, Proposition 9.1], [28, Theorem 1] or [35, Lemma 23]), and we shall omit the conditioning on \mathcal{E}_1 from our notation to avoid clutter. Let \mathfrak{S}_i denote the set of all (G', J_i) -edge-extensions \mathcal{S} . Since $\mathcal{S} \subseteq \binom{V_i \cup V(G')}{2} \setminus \binom{V(G')}{2}$ and $z \ll n$ by (40), the number of z -element collections $\mathcal{C} \subseteq \mathfrak{S}_i$ of edge-extensions with pairwise disjoint vertex-sets $W(\mathcal{S}) \setminus V(G')$ is at least

$$\frac{1}{z!} \prod_{0 \leq j < z} \binom{n_i - j(v_{J_i} - v_G)}{v_{J_i} - v_G} \geq \frac{1}{z!} \left[\left(\frac{n_i - z(v_{J_i} - v_G)}{v_{J_i} - v_G} \right)^{v_{J_i} - v_G} \right]^z \geq \left(\frac{\Theta(n^{v_{J_i} - v_G})}{z} \right)^z. \quad (44)$$

For any such collection \mathcal{C} , for brevity we introduce the events

$$\mathcal{I}_{\mathcal{C}} := \{\mathcal{E}_2 \text{ contains all } \mathcal{S} \in \mathcal{C}\} \quad \text{and} \quad \mathcal{D}_{\mathcal{C}} := \{\mathcal{E}_2 \text{ contains no } \mathcal{S} \in \mathfrak{S}_i \setminus \mathcal{C}\}.$$

We trivially have $\mathbb{P}(\mathcal{I}_{\mathcal{C}}) \geq p_2^{(e_{J_i} - e_G)z}$ (in fact, this holds with equality), and defer the proof of

$$\mathbb{P}(\mathcal{D}_{\mathcal{C}} \mid \mathcal{I}_{\mathcal{C}}) \geq \omega^{-o(z)}. \quad (45)$$

Since there are at least (44) many such collections \mathcal{C} , using disjointness of the events $\mathcal{I}_{\mathcal{C}} \cap \mathcal{D}_{\mathcal{C}}$ we obtain

$$\mathbb{P}(Z_{G', J_i} = z) \geq \sum_{\mathcal{C}} \mathbb{P}(\mathcal{I}_{\mathcal{C}}) \mathbb{P}(\mathcal{D}_{\mathcal{C}} \mid \mathcal{I}_{\mathcal{C}}) \geq \left(\frac{n^{v_{J_i} - v_G} p_2^{e_{J_i} - e_G} \omega^{-o(1)}}{z} \right)^z.$$

Note that (36) gives $\mu_{J_i}/\mu_G \asymp \omega^{v_{J_i} - v_G}$. Since $\delta \gg \omega^{-v_G/2}$ by (39) and $z = O(\omega^{v_{J_i}/r})$ by (40), we infer

$$\frac{n^{v_{J_i} - v_G} p_2^{e_{J_i} - e_G}}{z} \asymp \frac{\mu_{J_i}}{\mu_G} \cdot \frac{\delta^{e_{J_i} - e_G}}{z} \geq \omega^{-\Theta(1)},$$

and (recalling that we omitted the conditioning on \mathcal{E}_1 from our notation) inequality (43) follows. It remains to give the deferred proof of estimate (45). To this end observe that

$$\mathcal{D}_{\mathcal{C}} = \bigcap_{\mathcal{S} \in \mathfrak{S}_i \setminus \mathcal{C}} \{\mathcal{S} \not\subseteq \mathcal{E}_2\} \quad \text{and} \quad \mathcal{I}_{\mathcal{C}} = \{E_{\mathcal{C}} \subseteq \mathcal{E}_2\} \quad \text{with} \quad E_{\mathcal{C}} := \bigcup_{\mathcal{S} \in \mathcal{C}} \mathcal{S}.$$

Noting that the $\{\mathcal{S} \setminus E_C \not\subseteq \mathcal{E}_2\}$ are all decreasing events with respect to the independent \mathcal{E}_2 -edge indicators, using Harris' inequality [13] (a special case of the FKG-inequality) it follows that

$$\mathbb{P}(\mathcal{D}_C \mid \mathcal{I}_C) = \mathbb{P}\left(\bigcap_{\mathcal{S} \in \mathfrak{S}_i \setminus \mathcal{C}} \{\mathcal{S} \setminus E_C \not\subseteq \mathcal{E}_2\}\right) \geq \prod_{\mathcal{S} \in \mathfrak{S}_i \setminus \mathcal{C}} \mathbb{P}(\mathcal{S} \setminus E_C \not\subseteq \mathcal{E}_2) = \prod_{\mathcal{S} \in \mathfrak{S}_i \setminus \mathcal{C}} \left(1 - p_2^{|\mathcal{S} \setminus E_C|}\right). \quad (46)$$

Recall that each edge-extension $\mathcal{S} \in \mathfrak{S}_i$ is isomorphic to $E(J_i) \setminus E(G)$. Combining that $J_i \setminus G = J_i[V(J_i) \setminus V(G)]$ is connected (see Claim 7(ii)) with the fact that all vertex-sets $W(\mathcal{S}) \setminus V(G')$ with $\mathcal{S} \in \mathcal{C}$ are pairwise disjoint, it follows that E_C contains no further edge-extension $\mathcal{S} \in \mathfrak{S}_i \setminus \mathcal{C}$. Therefore in every factor in (46) we have $|\mathcal{S} \setminus E_C| \geq 1$ and thus $\mathcal{S} \setminus E_C$ is isomorphic to $E(J_i) \setminus E(F)$ for some $G \subseteq F \subsetneq J_i$. As $p_2 \leq p \ll 1$, $n_i \leq n$ and $|\mathcal{C}| = z$, it follows that

$$-\log \mathbb{P}(\mathcal{D}_C \mid \mathcal{I}_C) \leq 2 \sum_{\mathcal{S} \in \mathfrak{S}_i \setminus \mathcal{C}} p_2^{|\mathcal{S} \setminus E_C|} \leq 2 \sum_{G \subseteq F \subsetneq J_i} (v_{J_i} |\mathcal{C}|)^{v_F - v_G} n^{v_{J_i} - v_F} p^{e_{J_i} - e_F} = O\left(\sum_{G \subseteq F \subsetneq J_i} z^{v_F - v_G} \frac{\mu_{J_i}}{\mu_F}\right).$$

Our initial reduction step ensures $\mu_{J_i}/\mu_G \ll z \log \omega$, see (33) and (19). Furthermore, (40) gives $z = O(\omega^{v_K/r})$ and (36) gives $\mu_G \asymp \omega^{v_G}$. As no $G \subsetneq F \subsetneq J_i$ is primal (since J_i covers G), using (37) it follows that

$$-\log \mathbb{P}(\mathcal{D}_C \mid \mathcal{I}_C) = O\left(\frac{\mu_{J_i}}{\mu_G} \left[1 + \sum_{G \subsetneq F \subsetneq J_i} \omega^{v_F v_K/r} \frac{\omega^{v_G}}{\omega^{2v_H^2}}\right]\right) \ll z \log \omega,$$

which completes the proof of (45) and thus inequality (43).

For the final step (iii), we further (in addition to the conditioning on \mathcal{E}_1 from step (ii) above) condition on the edge-set \mathcal{E}_2 , assuming that $Z_{G', J_i} = z$ for all $i \in [r]$. Recalling that the subgraphs $J_i \setminus G$ are vertex-disjoint (see Claim 7(iii)), note that if we pick any r copies of J_1, \dots, J_r counted by $Z_{G', J_1}, \dots, Z_{G', J_r}$ (which are all vertex-disjoint outside of G'), then their union gives a copy of $K = J_1 \cup \dots \cup J_r$ (here it matters that the shared copy G' is ordered). For each such copy of K we henceforth fix *one* ordered copy K' with vertex-ordering $u'_1, \dots, u'_{v_G}, u'_{v_G+1}, \dots, u'_{v_K}$, say. Let \mathcal{K} denote the collection of all such ordered K' (each of which satisfies $E(K') \subseteq \mathcal{E}_1 \cup \mathcal{E}_2$), and define $V(\mathcal{K})$ as the union of all their vertex-sets. Note that

$$|\mathcal{K}| = z^r \asymp C_H \varepsilon \mu_K \asymp C_H \varepsilon n^{v_K} p^{e_K}. \quad (47)$$

Given $K' \in \mathcal{K}$, we say that a copy H' of H in $\mathcal{E}_1 \cup \mathcal{E}_2 \cup \mathcal{E}_3$ is an (K', H) -extension if H' contains the ordered copy K' with $V(K') = \{u'_1, \dots, u'_{v_K}\}$, satisfies $V(H') \setminus V(K') \subseteq [n] \setminus V(\mathcal{K})$, and there is an injection from $V(H)$ to $V(H')$ with $u_j \mapsto u'_j$ for $j \in [v_K]$ that maps every edge $E(H) \setminus E(K)$ to an edge in \mathcal{E}_3 . Let X'_H denote the number of copies of H which are (K', H) -extensions for some $K' \in \mathcal{K}$. Let X''_H denote the number of copies of H with vertices in $[n] \setminus V(G')$ and all edges in \mathcal{E}_3 . As the sets of H -copies counted by X'_H and X''_H are disjoint (the former contain G' , and the latter share no vertices with G'), we have $X_H \geq X'_H + X''_H$. Noting that X'_H and X''_H are both increasing functions of the independent \mathcal{E}_3 -edge indicators, using Harris' inequality it follows that

$$\mathbb{P}(X_H \geq (1 + \varepsilon)\mu_H \mid \mathcal{E}_1, \mathcal{E}_2) \geq \mathbb{P}(X'_H \geq 2\varepsilon\mu_H \mid \mathcal{E}_1, \mathcal{E}_2) \cdot \mathbb{P}(X''_H \geq (1 - \varepsilon)\mu_H \mid \mathcal{E}_1, \mathcal{E}_2). \quad (48)$$

To establish inequality (18) it thus suffices to prove

$$\mathbb{P}(X'_H \geq 2\varepsilon\mu_H \mid \mathcal{E}_1, \mathcal{E}_2) \gg \omega^{-v_K}, \quad (49)$$

$$\mathbb{P}(X''_H \geq (1 - \varepsilon)\mu_H \mid \mathcal{E}_1, \mathcal{E}_2) = 1 - o(1). \quad (50)$$

Indeed, since we conditioned on \mathcal{E}_1 satisfying $X_G^* \geq 1$ and \mathcal{E}_2 satisfying $Z_{G', J_i} = z$ for all $i \in [r]$, by combining (48)–(50) with estimates (41) and (42)–(43), then inequality (18) follows readily.

In the remaining proofs of (49)–(50) we shall again omit the conditioning (on $\mathcal{E}_1, \mathcal{E}_2$) from our notation. Turning to the crude estimate (49), we define $Y_{K', H}$ as the number of (K', H) -extensions, so that

$$X'_H = \sum_{K' \in \mathcal{K}} Y_{K', H}.$$

Note that (47) and (40) imply the rough bound $|V(\mathcal{K})| \leq v_K |\mathcal{K}| \asymp z^r \ll n$, so that $|[n] \setminus V(\mathcal{K})| \asymp n$, say. Combining (47) with $p_3 = (1 - O(\delta))p \asymp p$ (which due to $\delta = c_H \min\{\varepsilon, 1\}$ holds for $c_H > 0$ sufficiently small) and $\mu_H = \Theta(n^{v_H} p^{e_H})$, it follows for $C_H > 0$ sufficiently large that

$$\mathbb{E}X'_H = \sum_{K' \in \mathcal{K}} \mathbb{E}Y_{K',H} = |\mathcal{K}| \cdot \Theta(n^{v_H - v_K} p_3^{e_H - e_K}) = C_H \cdot \Theta(\varepsilon \mu_H) \geq 4\varepsilon \mu_H.$$

Similarly, for all $K'_1, K'_2 \in \mathcal{K}$ we also have the routine upper bound

$$\mathbb{E}(Y_{K'_1,H} Y_{K'_2,H}) \leq n^{v_H - v_K} p_3^{e_H - e_K} \sum_{K \subseteq F \subseteq H} n^{v_H - v_F} p_3^{e_H - e_F} = \prod_{i \in [2]} \mathbb{E}Y_{K'_i,H} \cdot O\left(\sum_{K \subseteq F \subseteq H} \frac{\mu_K}{\mu_F}\right).$$

Since K is primal (see Claim 7(i)), by combining $\mu_F \geq \Phi_H$ with estimates (36) and (38) it follows that

$$\mathbb{E}(X'_H)^2 = \sum_{K'_1, K'_2 \in \mathcal{K}} \mathbb{E}(Y_{K'_1,H} Y_{K'_2,H}) \leq (\mathbb{E}X'_H)^2 \cdot O(\mu_K / \Phi_H) \ll (\mathbb{E}X'_H)^2 \cdot \omega^{v_K}.$$

Using a version of the Paley–Zygmund inequality (see, e.g., [18, Lemma 3.2]) we infer

$$\mathbb{P}(X'_H \geq 2\varepsilon \mu_H) \geq \mathbb{P}(X'_H \geq \frac{1}{2} \mathbb{E}X'_H) \geq \frac{1}{4} \cdot \frac{(\mathbb{E}X'_H)^2}{\mathbb{E}(X'_H)^2} \gg \omega^{-v_K},$$

which (recalling that we omitted the conditioning on $\mathcal{E}_1, \mathcal{E}_2$ from our notation) implies inequality (49).

Turning to the final estimate (50), for any $F \subseteq H$ with $e_F \geq 1$ we define Y_F as the number of copies of F with vertex-set in $[n] \setminus V(G')$ and edge-set in \mathcal{E}_3 , so that $X''_H = Y_H$. Note that Y_F has the same distribution as the number of copies of F in the (unconditional) binomial random graph $G_{n-v(G), p_3}$. Furthermore, $\delta \gg n^{-1}$ follows from (39) and (36), with room to spare (since G is primal). Recalling the definitions of $p_3 = (1 - O(\delta))p$ and $\delta = c_H \min\{\varepsilon, 1\}$, for $c_H > 0$ sufficiently small it thus is routine to see that

$$\frac{\mathbb{E}Y_F}{\mu_F} = \frac{\binom{n-v(G)}{v_F}}{\binom{n}{v_F}} \left(\frac{p_3}{p}\right)^{e_F} = (1 - O(n^{-1})) \cdot (1 - O(\delta)) \geq 1 - \varepsilon/2.$$

Since also $\mathbb{E}Y_F \asymp \mu_F$, standard variance estimates for random graphs (see, e.g., (9) or [17, Lemma 3.5]) imply

$$\text{Var } Y_H \asymp \frac{(\mathbb{E}Y_H)^2}{\min_{F \subseteq H: e_F \geq 1} \mathbb{E}Y_F} \asymp \frac{(\mu_H)^2}{\Phi_H}.$$

Using $X''_H = Y_H$, Chebychev's inequality, and the assumption $\varepsilon^2 \Phi_H \gg 1$ it follows that

$$\mathbb{P}(X''_H \leq (1 - \varepsilon)\mu_H) \leq \mathbb{P}(Y_H \leq \mathbb{E}Y_H - \frac{1}{2}\varepsilon\mu_H) \leq \frac{\text{Var } Y_H}{(\frac{1}{2}\varepsilon\mu_H)^2} \asymp \frac{1}{\varepsilon^2 \Phi_H} = o(1),$$

which (as we omitted the conditioning on $\mathcal{E}_1, \mathcal{E}_2$) completes the proof of (49)–(50) and thus Lemma 5. \square