ON THE VARIATIONAL PROBLEM FOR UPPER TAILS IN SPARSE RANDOM GRAPHS

EYAL LUBETZKY AND YUFEI ZHAO

Abstract. What is the probability that the number of triangles in $\mathcal{G}_{n,p}$, the Erdős-Rényi random graph with edge density p, is at least twice its mean? Writing it as $\exp[-r(n,p)]$, already the order of the rate function $r(n,p)$ was a longstanding open problem when $p = o(1)$, finally settled in 2012 by Chatterjee and by DeMarco and Kahn, who independently showed that $r(n,p) \asymp n^2p^2 \log(1/p)$ for $p \gtrsim \frac{\log n}{n}$; the exact asymptotics of $r(n,p)$ remained unknown.

The following variational problem can be related to this large deviation question at $p \gtrsim \frac{\log n}{n}$: for $\delta > 0$ fixed, what is the minimum asymptotic p-relative entropy of a weighted graph on n vertices with triangle density at least $(1+\delta)p^3$? A beautiful large deviation framework of Chatterjee and Varadhan (2011) reduces upper tails for triangles to a limiting version of this problem for *fixed* p. A very recent breakthrough of Chatterjee and Dembo extended its validity to $n^{-\alpha} \ll p \ll 1$ for an explicit $\alpha > 0$, and plausibly it holds in all of the above sparse regime.

In this note we show that the solution to the variational problem is $\min\{\frac{1}{2}\delta^{2/3}, \frac{1}{3}\delta\}$ when $n^{-1/2} \ll p \ll 1$ vs. $\frac{1}{2}\delta^{2/3}$ when $n^{-1} \ll p \ll n^{-1/2}$ (the transition between these regimes is expressed in the count of triangles minus an edge in the minimizer). From the results of Chatterjee and Dembo, this shows for instance that the probability that $\mathcal{G}_{n,p}$ for $n^{-\alpha} \leq p \ll 1$ has twice as many triangles as its expectation is $\exp[-r(n,p)]$ where $r(n,p) \sim \frac{1}{3}n^2p^2\log(1/p)$. Our results further extend to k-cliques for any fixed k, as well as give the order of the upper tail rate function for an arbitrary fixed subgraph when $p \geq n^{-\alpha}$.

1. Introduction

The following question regarding upper tails for triangle counts in $\mathcal{G}_{n,p}$, the Erdős-Rényi random graph with edge density p, has been extensively studied, being a representing example of large deviations for subgraph counts in random graphs (see, e.g., [4, 7, 8, 12–15, 21] as well as [2, 11] and the references therein):

Question. *What is the probability that the number of triangles in $\mathcal{G}_{n,p}$ is at least twice its mean, or more generally, larger by a factor of $1+\delta$ for $\delta > 0$ fixed?*

In the dense case (p fixed), the limiting asymptotics of the rate function — the normalized logarithm of this probability, here denoted by $r(n,p,\delta)$ — was reduced to an analytic variational problem on symmetric functions $f : [0,1]^2 \to [0,1]$ (for a large class of large deviation questions) by Chatterjee and Varadhan [6]. However, for $p = o(1)$, obtaining the order of $r(n,p,\delta)$ was already a longstanding open problem. That $n^2p^2 \lesssim r(n,p,\delta) \lesssim n^2p^2\log(1/p)$ followed from the works of Vu [21] and Kim and Vu [15] (see also [12]), and this question was finally settled in 2012 by Chatterjee [4] and by DeMarco and Kahn [8], where it was independently shown that $r(n,p,\delta) \asymp n^2p^2\log(1/p)$ for $p \gtrsim \frac{\log n}{n}$ (see [4, 8] for an account of the rich related literature). The exact asymptotics of this rate function was not known for any $\frac{\log n}{n} \lesssim p \ll 1$.

Note that for the dense regime of fixed p, while [6] provided a closed form for the rate function in terms of the above variational problem, its solution is only known in a subset of the range of parameters (p,δ) known as the *replica symmetric* phase (where the excess in the number of triangles is explained by encountering too many edges that are essentially uniformly distributed), and little is known on its complement (the *symmetry breaking* phase; see our previous work [19] where this phase diagram was determined).

Y. Zhao was supported by a Microsoft Research Ph.D. Fellowship.

The variational problem in [6] can be viewed, via Szemerédi's regularity lemma [20] and the theory of graph limits by Lovász et al. [3, 17, 18], as the limit of the following problem.

Definition (Discrete variational problem for upper tails of triangles)**.** Let $\mathscr{G}_n$ denote the set of weighted undirected graphs on n vertices with edge weights in $[0,1]$, i.e.,

$$\mathscr{G}_n = \Big\{ G = (g_{ij})_{1\le i<j\le n} \ : \ 0 \le g_{ij} \le 1 \ , \ g_{ij} = g_{ji} \ , \ g_{ii} = 0 \ \text{ for all } i,j \Big\} \,.$$

The variational problem for $\delta > 0$ and $0 < p < 1$ is given by

$$\phi(n,p,\delta) := \inf \Big\{ I_p(G) : G \in \mathscr{G}_n \text{ with } t(G) \ge (1+\delta)p^3 \Big\} \,, \tag{1.1}$$

where

$$t(G) := n^{-3} \sum_{1\le i,j,k\le n} g_{ij}g_{jk}g_{ik}$$

is the density of (labeled) triangles in G, and $I_p(G)$ is its entropy relative to p, i.e.,

$$I_p(G) := \sum_{1\le i<j\le n} I_p(g_{ij}) \quad \text{ with } \quad I_p(x) := x \log \frac{x}{p} + (1-x)\log\frac{1-x}{1-p} \,.$$

Indeed, it follows from the powerful large deviation framework of [6] that for p fixed (the dense regime) $\frac{1}{n^2}\log \mathbb{P}\left(t(\mathcal{G}_{n,p}) \ge (1+\delta)p^3\right)$ tends as $n\to\infty$ to the limit of $-\phi(n,p,\delta)/n^2$.

However, in the sparse regime of $p = o(1)$, which lacks the rich set of tools that are based on Szemerédi's regularity lemma for dense graphs, there were no counterparts to this result until a very recent breakthrough by Chatterjee and Dembo [5]. There it was shown that the discrete variational problem (1.1) does govern the rate function of subgraph counts as long as $p \ge n^{-\alpha}$ for a suitable constant α. In particular, for triangle counts (see [5, Theorem 1.2] and the remark following it, yielding a slightly wider range than the one stated next) one has that

$$\mathbb{P}\left(t(\mathcal{G}_{n,p}) \ge (1+\delta)p^3\right) = \exp\left[-(1-o(1))\phi(n,p,\delta)\right] \tag{1.2}$$

whenever $n^{-1/42}\log n \le p \ll 1$ (this should extend to smaller p, as commented in [5]; in fact, it is plausible that this result holds throughout the sparse regime of $\frac{\log n}{n} \ll p \ll 1$.)

In this note we establish the following for the discrete variational problem (1.1).

Theorem 1.1. *Fix $\delta > 0$. If $n^{-1/2} \ll p \ll 1$, then*

$$\lim_{n\to\infty} \frac{\phi(n,p,\delta)}{n^2p^2\log(1/p)} = \min\left\{ \frac{\delta^{2/3}}{2} \ , \ \frac{\delta}{3} \right\} . \tag{1.3}$$

On the other hand, if $n^{-1} \ll p \ll n^{-1/2}$, then

$$\lim_{n\to\infty} \frac{\phi(n,p,\delta)}{n^2p^2\log(1/p)} = \frac{\delta^{2/3}}{2} \,. \tag{1.4}$$

One can then deduce the following from the above result (1.2) of Chatterjee and Dembo.

Corollary 1.2. *For any $\delta > 0$, if $n^{-1/42}\log n \le p \ll 1$ then*

$$\mathbb{P}\left(t(\mathcal{G}_{n,p}) \ge (1+\delta)p^3\right) = \exp\left[-(1-o(1))\min\left\{\tfrac12\delta^{2/3}\,,\ \tfrac13\delta\right\} n^2p^2\log(1/p)\right] \,.$$

The lower bound is explained by forcing either a set of $k = \delta^{1/3}np$ vertices to be a clique (with probability $p^{\binom{k}{2}} = p^{(\delta^{2/3}/2+o(1))n^2p^2}$) or a set of $\ell = \frac{1}{3}\delta np^2$ vertices to be connected to all other vertices (with probability $p^{\ell(n-\ell)} = p^{(\delta/3+o(1))n^2p^2}$), the latter being preferable if and only if $\delta < 27/8$.

In fact, these constructions for the lower bound on $\mathbb{P}\left(t(\mathcal{G}_{n,p}) \geq (1+\delta)p^3\right)$ further explain the two separate regimes in Theorem 1.1. When $p \ll 1/\sqrt{n}$, the second (bipartite) construction — involving $\ell \asymp np^2$ vertices — ceases to be a viable option, as then we have $\ell = o(1)$. As remarked next, this translates into a qualitative difference between the solutions of the variational problem in each of these regimes, expressed in terms of

$$s(G) := n^{-3} \sum_{1\leq i\leq n} \Big(\sum_{1\leq j\leq n} g_{ij}\Big)^2 ,$$

equivalent to the asymptotic density of triangles minus an edge (i.e., $K_{1,2}$ homomorphisms, which in $\mathcal{G}_{n,p}$ have average density p^2, and so an excess of $\frac{1}{3}\delta p^2$ in their density, of which a p-fraction forms triangles via an extra edge, translates to δp^3 additional labeled triangles).

Remark 1.3. *The proof of Theorem 1.1 shows that for any fixed* $0 < \delta < \frac{27}{8}$, *if* $G_n \in \mathscr{G}_n$ *is a sequence of weighted graphs satisfying* $t(G_n) \geq (1+\delta)p^3$ *and* $I_p(G_n) \sim \phi(n,p,\delta)$ *then*

$$\lim_{n\to\infty} \frac{s(G_n)}{p^2} = \begin{cases} 1+\delta/3 & \text{if } n^{-1/2} \ll p \ll 1\,, \\ 1 & \text{if } n^{-1} \ll p \ll n^{-1/2}\,. \end{cases}$$

For fixed $\delta > \frac{27}{8}$, *the term* $1+\delta/3$ *in the first case* ($n^{-1/2} \ll p \ll 1$) *is replaced by* 1.

Regarding the behavior when $p \asymp n^{-1/2}$, there one expects a similar structure: i.e., whenever the bipartite construction is preferable, the optimal solution should feature a large bipartite subgraph while adhering to the integrality restrictions. It is plausible that methods similar to those used in this work can establish the solution in that regime as well.

Our arguments extend to yield analogous results for k-clique counts, where, for instance, the right-hand side of (1.3) (giving the asymptotics of the rate function provided $n^{-\alpha'} \ll p \ll 1$ for $\alpha'(k) > 0$) is replaced by $\min\{\frac{1}{2}\delta^{2/k}, \delta/k\}$; see Theorem 4.1 and Corollary 4.2. For a general graph on k vertices, the *order* of the rate function at $p \geq n^{-\alpha''}$ is given by Corollary 4.5.[1]

Finally, it is worthwhile mentioning that even without appealing to the new machinery of [5], if p tends to 0 sufficiently slowly with n — namely, $(\log n)^{-1/6} \ll p \ll 1$ — then Eq. (1.2) (stating that the variational problem (1.1) gives the asymptotic rate function for large deviations of triangles) follows essentially from the framework of Chatterjee and Varadhan [6] (and similarly for any fixed subgraph); instead of using the theory of graph limits or Szemerédi's regularity lemma, one can derive this statement by appealing in their framework to the *weak* regularity lemma of Frieze and Kannan [10] (we include this reduction for completeness; see §5).

Notation and organization. On occasion we will write $f_n \lesssim g_n$ instead of $f_n = O(g_n)$ for brevity, as well as $f_n \ll g_n$ instead of $f_n = o(g_n)$ (similarly for $f_n \gtrsim g_n$ and $f_n \gg g_n$); we let $f_n \sim g_n$ denote $f_n = (1+o(1))g_n$, and $f \asymp g$ denotes $f_n \lesssim g_n \lesssim f_n$.

[1] In our follow-up work [1] jointly with Bhattacharya and Ganguly, we extend this and find the asymptotic rate function for every graph H. The rate is given in terms of a certain independence polynomial, and exhibits a dichotomy with respect to δ if and only if H is a regular graph. See [1] for the statements of these newer results.

This paper is organized as follows. In §2 we give upper and lower bounds for the discrete variational problem (1.1): the construction of a clique/bipartite subgraph, and a (relaxed) continuous variational problem, whose solution we denote by $\phi(\delta, p)$ (notice this variant no longer depends on n; see Eq. (2.1) below). The analysis of the latter appears in §3, and §4 contains the extension of these results to k-cliques for any fixed k. Finally, §5 contains the reduction of the upper tail to the variational problem (1.1) when $p \to 0$ as a poly-log of n.

2. A continuous variational problem

In this section we compare the optimum $\phi(n, p, \delta)$ of the variational problem (1.1) with an analogue, $\phi(p, \delta)$, that eliminates the dependence on n. Before introducing this variant, we begin with the straightforward upper bound on $\phi(n, p, \delta)$, which involves constructing $G \in \mathscr{G}_n$ with $I_p(G)$ that attains the right-hand side of (1.3). There are two competing candidates.

- Let $g_{ij} = 1$ whenever $1 \le i < j \le a$ for some integer a to be specified later, and $g_{ij} = p$ for all other i, j. Then we have
$$t(G) \ge n^{-3}\left[a(a-1)(a-2) + (n(n-1)(n-2) - a(a-1)(a-2))p^3\right]$$
and
$$I_p(G) = \tbinom{a}{2} I_p(1) = \tbinom{a}{2}\log(1/p)\,.$$
So, we can choose $a = (\delta^{1/3} + o(1))pn$ so that $t(G) \ge (1+\delta)p^3$ and
$$I_p(G) = \left(\frac{\delta^{2/3}}{2} + o(1)\right) n^2 p^2 \log(1/p)\,.$$
- Let $g_{ij} = 1$ whenever $1 \le i \le a$ and $i < j$ and $g_{ij} = p$ otherwise. Then
$$t(G) \ge n^{-3}\left[3a(n-a)(n-a-1)p + (n-a)(n-a-1)(n-a-2)p^3\right]$$
and
$$I_p(G) = a\left(n - \frac{a+1}{2}\right) I_p(1) = a\left(n - \frac{a+1}{2}\right)\log(1/p)\,.$$
So, we can choose $a = (\delta/3 + o(1))p^2 n$ so that $t(G) \ge (1+\delta)p^3$ and
$$I_p(G) = \left(\frac{\delta}{3} + o(1)\right) n^2 p^2 \log(1/p)\,.$$

When $p \gg n^{-1/2}$, both constructions are valid, and taking the one with smaller $I_p(G)$ (the choice depends on the value of δ; when $\delta \ge 27/8$ we use the first construction and when $\delta < 27/8$ we use the second construction) yields the upper bound on $\phi(n, p, \delta)$ in (1.3).

When $n^{-1} \ll p \ll n^{-1/2}$, the second construction is no longer valid (since $a \ll 1$), but the first construction remains valid. Thus, we obtain the upper bound on $\phi(n, p, \delta)$ in (1.4).

Next, consider the following variant of the above variational problem. Whereas in $\phi(n, p, \phi)$ the variational problem occurs in the space of weighted graphs on n vertices, in the new variational problem $\phi(p, \phi)$, we consider the space of graphons, so that n does not appear (and the dependence of p on n plays no role). Here a *graphon* is a symmetric measurable function $W\colon [0,1]^2 \to [0,1]$. Let $\mathcal{W}$ denote the set of all graphons.

Given any graphon W and function $f\colon \mathbb{R} \to \mathbb{R}$, we use the shorthand notation

$$\mathbb{E}[f(W)] := \int_{[0,1]^2} f(W(x,y))\,dxdy\,.$$

For example, $\mathbb{E}W^2 = \int_{[0,1]^2} W^2\,dxdy$, and $\mathbb{E}[I_p(W)] = \int_{[0,1]^2} I_p(W(x,y))\,dxdy$.

Definition (Continuous variational problem)**.** For $\delta > 0$ and $0 < p < 1$, let

$$\phi(p,\delta) := \inf\left\{\frac{1}{2}\mathbb{E}[I_p(W)] : W \in \mathcal{W} \text{ such that } t(W) \geq (1+\delta)p^3\right\}, \tag{2.1}$$

where the triangle density $t(W)$ of W is defined by

$$t(W) := \int_{[0,1]^3} W(x,y)W(x,z)W(y,z)\,dxdydz\,.$$

The two variational problems (1.1) and (2.1) are related by the following inequality.

Lemma 2.1. *For any p, n, δ, we have*

$$\phi(p,\delta) \leq \frac{1}{n^2}\phi(n,p,\delta) + \frac{1}{2n}I_p(0). \tag{2.2}$$

Proof. For any $G \in \mathscr{G}_n$, we can construct a $W^G \in \mathcal{W}$ by dividing $[0,1]$ into n equal intervals $I_1, \ldots, I_n$, and setting $W(x,y) = g_{ij}$ whenever $x \in I_i$ and $y \in I_j$. Then $t(W^G) = t(G)$ and $\frac{1}{2}\mathbb{E}[I_p(W^G)] = n^{-2}I_p(G) + (2n)^{-1}I_p(0)$, where the extra term $(2n)^{-1}I_p(0)$ is due to the zero entries $g_{ii} = 0$ which were not included in $I_p(G)$. ■

The following theorem, providing a solution to the variational problem $\phi(p,\delta)$, is proved in the next section (see §3.1).

Theorem 2.2. *Fix $\delta > 0$. Then*

$$\lim_{p\to 0} \frac{\phi(p,\delta)}{p^2\log(1/p)} = \min\left\{\frac{\delta^{2/3}}{2}\,,\;\frac{\delta}{3}\right\}. \tag{2.3}$$

It can already be seen that the solution to the variational problem (1.1) when $n^{-1/2} \ll p \ll 1$ (i.e., Eq. (1.3)) will readily follow from the combination of Lemma 2.1 and Theorem 2.2. We defer the full details — together with the treatment of the regime $n^{-1} \ll p \ll n^{-1/2}$ (which will entail a short modification of the proof of Theorem 2.2) to the next section following the proof of Theorem 2.2 (see §3.2).

For now, let us give the constructions that give tight upper bounds on $\phi(p,\delta)$ for (2.3) — precisely the graphon analogs of the above given constructions for Theorem 1.1.

- Let $W(x,y) = 1$ whenever $x, y \in [0,a]$ for some $a \in [0,1]$ to be specified later, and $W(x,y) = p$ elsewhere. Then we have

  $$t(W) \geq a^3 + (1-a)^3p^3$$

  and

  $$\frac{1}{2}\mathbb{E}[I_p(W)] = \frac{1}{2}a^2 I_p(1) = \frac{1}{2}a^2\log(1/p)\,.$$

So, we can choose $a = (\delta^{1/3} + o(1))p$ so that $t(W) \ge (1+\delta)p^3$ and

$$\frac{1}{2}\mathbb{E}[I_p(W)] = \left(\frac{\delta^{2/3}}{2} + o(1)\right)p^2 \log(1/p)\,.$$

- Let $W(x,y) = 1$ whenever $\min\{x,y\} \le a$ and $W(x,y) = p$ otherwise. Then

$$t(W) \ge 3a(1-a)^2 p + (1-a)^3 p^3$$

and

$$\frac{1}{2}\mathbb{E}[I_p(W)] = a\left(1 - \frac{a}{2}\right) I_p(1) = a\left(1 - \frac{a}{2}\right)\log(1/p)\,.$$

So, we can choose $a = (\delta/3 + o(1))p^2$ so that $t(G) \ge (1+\delta)p^3$ and

$$\frac{1}{2}\mathbb{E}[I_p(W)] = \left(\frac{\delta}{3} + o(1)\right)p^2 \log(1/p)\,.$$

Depending on the value of δ (when $\delta \ge 27/8$ use the first construction; when $\delta < 27/8$ use the second), these two examples together prove the upper bound to $\phi(p,\delta)$ in (2.3).

3. Solving the variational problem

3.1. **Proof of Theorem 2.2.** Throughout this proof, we will occasionally require various technical properties of the function I_p when $p \to 0$; the proofs of these are deferred to §3.3.

Let $W \in \mathcal{W}$ satisfy $t(W) \ge (1+\delta)p^3$. We wish to show that

$$\frac{1}{2}\mathbb{E}[I_p(W)] \ge (1 - o(1)) \min\left\{\frac{\delta^{2/3}}{2}\,,\,\frac{\delta}{3}\right\}p^2 I_p(1)\,.$$

Since I_p is decreasing in $[0,p]$ and increasing in $[p,1]$, we may assume without loss of generality that $W \ge p$ and $t(W) = (1+\delta)p^3$. Write $W = U + p$, so that $0 \le U \le 1-p$. Letting

$$s(U) := \int_{[0,1]^3} U(x,y)U(x,z)\,dxdydz = \int_{[0,1]}\left(\int_{[0,1]} U(x,y)\,dy\right)^2 dx\,,$$

we have

$$t(W) - p^3 = t(U) + 3ps(U) + 3p^2\mathbb{E}U = \delta p^3\,. \tag{3.1}$$

Now write

$$t(U) = \delta_1 p^3, \qquad s(U) = \delta_2 p^2, \quad \text{and} \quad \mathbb{E}U = \delta_3 p\,.$$

Then $\delta_1 + 3\delta_2 + 3\delta_3 = \delta$. We may assume, for instance, that

$$\delta_3 \le \sqrt{p}\log(1/p) = o(1), \quad \text{so that } \mathbb{E}U = o(p)\,,$$

since otherwise by the convexity of I_p and the fact that $I_p(p+x) \sim x^2/(2p)$ for $x \ll p$ (see Lemma 3.3 below) we would already have

$$\mathbb{E}[I_p(W)] \ge I_p(\mathbb{E}W) = I_p(p + \mathbb{E}U) \ge I_p(p + p^{3/2}\log(1/p)) \gg p^2 I_p(1)\,.$$

The above decomposition reduces the problem to studying the following:

$$\phi'(p,\delta_1,\delta_2) := \inf\left\{\frac{1}{2}\mathbb{E}[I_p(p+U)] : U \in \mathcal{W} \text{ so that } 0 \le U \le 1-p,\ t(U) \ge \delta_1 p^3,\ \text{and } s(U) \ge \delta_2 p^2\right\}.$$

The (asymptotic) solution to this variational problem is given by the following key lemma.

Lemma 3.1. *Fix $D > 0$. Then*

$$\phi'(p,\delta_1,\delta_2) = \left(\frac{\delta_1^{2/3}}{2} + \delta_2 + o(1)\right) p^2 I_p(1)$$

uniformly for all $\delta_1, \delta_2 \in [0, D]$ as $p \to 0$.

Assuming Lemma 3.1, let us finish the proof of Theorem 2.2. We have

$$\begin{aligned}\frac{1}{2}\mathbb{E}[I_p(W)] &\geq \min\{\phi'(p,\delta_1,\delta_2) : \delta_1 + 3\delta_2 = \delta - o(1)\} \\ &= (1-o(1)) \min\left\{\frac{\delta_1^{2/3}}{2} + \delta_2 : \delta_1 + 3\delta_2 = \delta - o(1)\right\} p^2 I_p(1)\,.\end{aligned}$$

Note that if we fix the value of $\delta_1 + 3\delta_2$, then $\delta_1^{2/3}/2 + \delta_2$ is minimized when one of δ_1 and δ_2 is set to zero. It follows that

$$\frac{1}{2}\mathbb{E}[I_p(W)] \geq (1-o(1)) \min\left\{\frac{\delta^{2/3}}{2}\,,\ \frac{\delta}{3}\right\} p^2 I_p(1)\,.$$

We have thus established the desired lower bound for $\phi(p,\delta)$ in Theorem 2.2, while the upper bound was already given in §2 (immediately after the statement of the theorem). This completes the proof of the Theorem 2.2 modulo Lemma 3.1.

Towards the proof of Lemma 3.1, we need the following result, showing how to lower bound $\mathbb{E}[I_p(p+U)]$ given $t(U)$.

Lemma 3.2. *For any $U \in \mathcal{W}$ with $0 \leq U \leq 1-p$ we have*

$$\mathbb{E}[I_p(p+U)] \geq (1-o(1))\, I_p(1) t(U)^{2/3}\,.$$

where $o(1)$ is some quantity that goes to zero as $p \to 0$.

Proof. For $p = o(1)$ and any $0 \leq x \leq 1-p$ one has $I_p(p+x) \geq (1+o(1))x^2 I_p(1)$ (as established in Corollary 3.5 below); thus,

$$\begin{aligned}\mathbb{E}[I_p(p+U)] &= \int_{[0,1]^2} I_p(p+U(x,y))\,dxdy \\ &\geq (1-o(1))\, I_p(1) \int_{[0,1]^2} U(x,y)^2\,dxdy \geq (1-o(1))\, I_p(1) t(U)^{2/3}\,,\end{aligned}$$

where we will justify the last inequality using the fact that

$$t(U) \leq \left(\int_{[0,1]^2} U(x,y)^2\,dxdy\right)^{3/2} \qquad \text{for any } U \in \mathcal{W}\,. \tag{3.2}$$

Indeed, (3.2) follows from the Cauchy–Schwarz inequality:

$$\begin{aligned}t(U) &= \int_{[0,1]^3} U(x,y)U(x,z)U(y,z)\,dxdydz \\ &\leq \int_{[0,1]^2} \left(\int_{[0,1]} U(x,y)^2\,dx\right)^{1/2} \left(\int_{[0,1]} U(x,z)^2\,dx\right)^{1/2} U(y,z)\,dydz\end{aligned}$$

which, by two more applications of the Cauchy–Schwarz inequality, is at most

$$\int_{[0,1]} \left(\int_{[0,1]^2} U(x,y)^2 \, dxdy \right)^{1/2} \left(\int_{[0,1]} U(x,z)^2 \, dx \right)^{1/2} \left(\int_{[0,1]} U(y,z)^2 \, dy \right)^{1/2} dz$$

$$\leq \left(\int_{[0,1]^2} U(x,y)^2 \, dxdy \right)^{1/2} \left(\int_{[0,1]^2} U(x,z)^2 \, dxdz \right)^{1/2} \left(\int_{[0,1]^2} U(y,z)^2 \, dydz \right)^{1/2},$$

as required. ∎

Lemma 3.2 already shows that $\phi'(p, \delta_1, \delta_2) \geq (\delta_1^{2/3}/2 - o(1))p^2 I_p(1)$. However, this is not enough. To obtain the additional $\delta_2 p^2 I_p(1)$ term in the lower bound of ϕ', we isolate the high degree vertices and consider their contributions.

Proof of Lemma 3.1. First we prove an upper bound on $\phi'(p, \delta_1, \delta_2)$. Let A be the union of the rectangles

$$[0, \delta_1^{1/3} p]^2, \quad [0, \delta_2 p^2] \times [0,1], \text{ and } \quad [0,1] \times [0, \delta_2 p^2].$$

Set U to be $1-p$ on A and 0 elsewhere. Then we have $t(U) \geq \delta_2 p^3$, and $s(U) \geq \delta_2 p^2$, whereas $\frac{1}{2}\mathbb{E}[I_p(p+U)] = \frac{1}{2}\lambda(A) I_p(1) = (\frac{1}{2}\delta_1^{2/3} + \delta_2 + o(1))p^2 I_p(1)$, where here and in what follows λ denotes Lebesgue measure. This proves the upper bound on $\phi'(p, \delta_1, \delta_2)$.

Assume that $\mathbb{E}[I_p(p+U)] = O(p^2 \log(1/p))$ (with an implicit constant that may depend on D), or else we are done.

Let $f(x) = \int_{[0,1]} U(x,y)\, dy$. Let $b = p^{1/3}$ (any choice of b with $\sqrt{p \log(1/p)} \ll b \ll 1$ suffices), and $B = \{x \mid f(x) > b\} \subseteq [0,1]$. By the convexity of I_p we have

$$\mathbb{E}[I_p(p+U)] = \int_{[0,1]^2} I_p(p + U(x,y)) \, dxdy \geq \int_{[0,1]} I_p(p + f(x)) \, dx \geq \lambda(B) I_p(p+b).$$

Since $I_p(p+b) = (1+o(1)) b \log(b/p)$ (see Lemma 3.3 below),

$$\lambda(B) \leq \frac{\mathbb{E}[I_p(p+U)]}{I_p(p+b)} = \frac{O(p^2 \log(1/p))}{(1+o(1)) b \log(b/p)} = O\left(\frac{p^2}{b}\right). \tag{3.3}$$

Next, we have $I_p(p+x) \geq (x/b)^2 I_p(p+b)$ for $x \in [0,b]$ (see Lemma 3.4 below); hence,

$$\mathbb{E}[I_p(p+U)] \geq \int_{[0,1]\setminus B} I_p(p+f(x)) \, dx \geq \frac{I_p(p+b)}{b^2} \int_{[0,1]\setminus B} f(x)^2 \, dx.$$

Therefore,

$$\int_{[0,1]\setminus B} f(x)^2 \, dx \leq \frac{\mathbb{E}[I_p(p+U)] b^2}{I_p(p+b)} = O(p^2 b), \tag{3.4}$$

where the last step is by (3.3). Since $\int_{[0,1]} f(x)^2 \, dx = s(U) \geq \delta_2 p^2$, we have

$$\int_B f(x)^2 \, dx \geq (\delta_2 - O(b)) p^2 = (\delta_2 - o(1)) p^2.$$

First applying the convexity of I_p, then the fact (shown in Corollary 3.5 below) that $I_p(p+x)$ is at least $(1-o(1))x^2 I_p(1)$ for $p = o(1)$, and finally (3.4), we obtain

$$\begin{aligned}\int_{B\times[0,1]} I_p(p+U(x,y))\,dxdy &\geq \int_B I_p(p+f(x))\,dx \\ &\geq (1-o(1))\int_B f(x)^2 I_p(1)\,dx \geq (\delta_2 - o(1))p^2 I_p(1)\,.\end{aligned}$$

Since $U(x,y) = U(y,x)$, we have

$$\frac{1}{2}\int_{B\times[0,1]\cup[0,1]\times B} I_p(p+U(x,y))\,dxdy \geq (\delta_2 - o(1))p^2 I_p(1) - \frac{1}{2}\lambda(B)^2 I_p(1) \geq (\delta_2 - o(1))p^2 I_p(1)\,, \tag{3.5}$$

where the last step is due to $\lambda(B) = O(p^2/b) = o(p)$.

We have $\mathbb{E}[I_p(p+U)] \geq I_p(p+\mathbb{E}U)$ by convexity of I_p. As $I_p(p+x)$ is increasing for $x \in [0, 1-p]$, and Lemma 3.3 tells us that $I_p(p + Cp^{3/2}\sqrt{\log(1/p)}) \sim \frac{1}{2}C^2p^2\log(1/p)$ for each fixed $C > 0$ as $p \to 0$, we see that $\mathbb{E}[I_p(p+U)] = O(p^2\log(1/p))$ implies that $\mathbb{E}U = O(p^{3/2}\sqrt{\log(1/p)})$. Let $U' = U\mathbf{1}_{B^c\times B^c}$ where $B^c = [0,1]\setminus B$. We have

$$\begin{aligned}t(U) - t(U') &\leq 3\int_{B\times[0,1]\times[0,1]} U(x,y)U(x,z)U(y,z)\,dxdydz \\ &\leq 3\int_{B\times[0,1]\times[0,1]} U(y,z)\,dxdydz = 3\lambda(B)\mathbb{E}U = O\left(b^{-1}p^{7/2}\sqrt{\log(1/p)}\right) = o(p^3)\,.\end{aligned} \tag{3.6}$$

Thus,

$$t(U') \geq (\delta_1 - o(1))p^3.$$

By Lemma 3.2,

$$\begin{aligned}\frac{1}{2}\int_{B^c\times B^c} I_p(p+U(x,y))\,dxdy &= \frac{1}{2}\mathbb{E}[I_p(p+U')] \\ &\geq \left(\frac{1}{2} - o(1)\right)I_p(1)t(U')^{2/3} \geq \left(\frac{\delta_1^{2/3}}{2} - o(1)\right)p^2 I_p(1)\,.\end{aligned} \tag{3.7}$$

Combining (3.5) and (3.7), we deduce that

$$\frac{1}{2}\int_{[0,1]^2} I_p(p+U(x,y))\,dxdy \geq \left(\frac{\delta_1^{2/3}}{2} + \delta_2 - o(1)\right)p^2 I_p(1)\,.$$

This proves the lower bound on $\phi'(p,\delta_1,\delta_2)$. ■

3.2. **Discrete variational problem — proof of Theorem 1.1.** First consider the case $n^{-1/2} \ll p \ll 1$. The upper bound on the left-hand side of (1.3) was already proved in §2. For the lower bound, by applying Lemma 2.1 and then Theorem 2.2 we have

$$\lim_{n\to\infty}\frac{\phi(n,p,\delta)}{n^2p^2\log(1/p)} \geq \lim_{p\to 0}\frac{\phi(p,\delta)}{p^2\log(1/p)} - \lim_{n\to\infty}\frac{I_p(0)}{2np^2\log(1/p)} = \min\left\{\frac{\delta^{2/3}}{2}, \frac{\delta}{3}\right\} - 0\,.$$

The last zero is due to $I_p(0)/(np^2\log(1/p)) \sim 1/(np\log(1/p)) \to 0$. This proves (1.3).

It remains to treat the regime $n^{-1} \ll p \ll n^{-1/2}$. When $\delta \geq 27/8$, so that $\delta^{2/3}/2 \leq \delta/3$, the desired result again follows from Theorem 2.2 by the same argument as given above. However, when $\delta < 27/8$, second upper bound construction (stated immediately following Theorem 1.1) is invalid. In order to prove a matching lower bound for (1.4), we need to eliminate the second construction as a possibility. We sketch the modifications to the proof here. It suffices to show that $s(U) = o(p^2)$ (using the notation of the previous subsection). Indeed, once we know that $s(U) = o(p^2)$, the decomposition (3.1) implies $t(U) = (\delta - o(1))p^3$, from which we obtain $\frac{1}{2}\mathbb{E}[I_p(p+U)] \geq (\delta^{2/3}/2 - o(1))p^2 I_p(1)$ by Lemma 3.2.

From now on assume that $n^{-1} \ll p \ll n^{-1/2}$. Assume b is chosen so that

$$\max\{p^2 n, \sqrt{p\log(1/p)}\} \ll b \ll 1\,.$$

Then (3.3) gives $\lambda(B) = O(p^2/b) \ll 1/n$. Since we are in the discrete setting of Theorem 1.1, $\lambda(B) \ll 1/n$ implies that B must be an empty set. Therefore, from (3.4) we can infer that $s(U) = \int_{[0,1]} f(x)^2\,dx = O(p^2 b) = o(p^2)$, as claimed. This completes the proof. ∎

3.3. **Properties of the function I_p as $p \to 0$.** Here we collect the various facts about I_p that were referred to throughout the proof of Theorem 2.2.

Lemma 3.3. *Let $p \to 0$. If $0 \leq x \ll p$, then $I_p(p+x) \sim x^2/(2p)$. If $p \ll x \leq 1-p$, then $I_p(p+x) \sim x\log(x/p)$.*

Proof. We use Taylor expansion for $I_p(x)$ around $x = p$, noting that $I_p(p) = I_p'(p) = 0$, $I_p''(p) = 1/(p(1-p))$ and $I_p'''(x) = 1/(1-x)^2 - 1/x^2$. We have $I_p(p+x) = x^2 I_p''(p)/2 + x^3 I_p'''(\xi)/6$ for some $\xi \in (p, p+x)$; thus, $I_p(p+x) = x^2/(2p(1-p)) + O(x^3/p^2) \sim x^2/(2p)$ when $0 \leq x \ll p$.

If $p \ll x < 1-p$ (the required statement trivially holds for $x = 1-p$), then

$$I_p(p+x) = (p+x)\log\frac{p+x}{p} + (1-p-x)\log\frac{1-p-x}{1-p} = (1+o(1))x\log\frac{x}{p} + O(x)\,, \quad (3.8)$$

where the bound $O(x)$ comes from $|\log y| \leq y^{-1} - 1$ which is valid for all $y \in (0,1]$. This shows that $I_p(p+x) \sim x\log(x/p)$ when $p \ll x \leq 1-p$. ∎

Lemma 3.4. *There exists $p_0 > 0$ so that for all $0 < p \leq p_0$ and $0 \leq x \leq b \leq 1-p-1/\log(1/p)$,*

$$I_p(p+x) \geq (x/b)^2 I_p(p+b)\,. \quad (3.9)$$

Proof. Let $x_p = 1-p-1/\log(1/p)$. We will show that the function $f(x) = I_p(p+\sqrt{x})$ is concave for $x \in [0, x_p^2]$. The inequality (3.9) then follows because for each $b \leq x_p$, the chord joining $(0,0)$ and $(b^2, I_p(p+b))$ lies below f, so that $f(x) \geq (x/b^2)I_p(p+b)$ for all $0 \leq x \leq b^2$. Replacing x by x^2 yields (3.9).

We have

$$f''(x) = \frac{1}{4(1-p-\sqrt{x})(p+\sqrt{x})x} + \frac{1}{4x^{3/2}}\log\left(\frac{(1-p-\sqrt{x})p}{(p+\sqrt{x})(1-p)}\right)\,.$$

Let

$$g(x) = 4x^3 f''(x^2) = \frac{x}{(1-p-x)(p+x)} + \log\left(\frac{(1-p-x)p}{(p+x)(1-p)}\right)\,.$$

It now suffices to show that $g(x) \le 0$ for $x \in [0, x_p]$, which implies that f is concave in $[0, x_p^2]$. We have $g(0) = 0$ and

$$\begin{aligned} g(x_p) &= \log(1/p)\frac{1-p-1/\log(1/p)}{1-1/\log(1/p)} + \log\left(\frac{p}{\log(1/p)(1-1/\log(1/p))(1-p)}\right) \\ &\le \log(1/p) - \log(1/p) - \log\log(1/p) + O\left(1/\log(1/p)\right) = -\log\log(1/p) + o(1)\,. \end{aligned}$$

So, we can choose p_0 so that $g(x_p) \le 0$ for all $p \le p_0$. Furthermore, we have

$$g'(x) = \frac{(-1+2p+2x)x}{(1-p-x)^2(p+x)^2}\,.$$

It follows that g is decreasing when $x < 1/2 - p$ and increasing when $x > 1/2 - p$. Since $g(0), g(x_p) \le 0$, we conclude that $g(x) \le 0$ for all $x \in [0, x_p]$. ■

Corollary 3.5. *There is some $p_0 > 0$ so that for all $0 < p \le p_0$ and all $0 \le x \le 1 - p$ one has*

$$I_p(p+x) \ge x^2 I_p(1 - 1/\log(1/p)) = (1+o(1))x^2 I_p(1) \tag{3.10}$$

where the $o(1)$-term goes to zero as $p \to 0$.

Proof. Let $b = 1 - p - 1/\log(1/p)$. When $0 \le x \le b$, the first inequality in (3.10) follows from Lemma 3.4 since $b < 1$, and when $b < x \le 1-p$, it follows from $I_p(p+x) \ge I_p(p+b) \ge x^2 I_p(p+b)$ since $I_p(p+x)$ is increasing for $x \in [0, 1-p]$. The last step in (3.10) follows from Lemma 3.3. ■

4. Extension to cliques

In this section we extend Theorem 1.1 and Corollary 1.2) to upper tails for clique counts.

Definition (Discrete variational problem for upper tails of H-counts)**.** Let H be a graph on k vertices. Recall that $\mathscr{G}_n$ denotes the set of weighted undirected graphs on n vertices with edge weights in $[0, 1]$. The corresponding variational problem for $\delta > 0$ and $0 < p < 1$ is given by

$$\phi_H(n, p, \delta) := \inf\left\{ I_p(G) : G \in \mathscr{G}_n \text{ with } t(H, G) \ge (1+\delta)p^{|E(H)|} \right\}, \tag{4.1}$$

where

$$t(H, G) := n^{-k} \sum_{1 \le x_1, \ldots, x_k \le n} \prod_{ij \in E(H)} g_{x_i x_j}$$

is the probability that a random map $V(H) \to V(G)$ is a graph homomorphism.

Theorem 4.1. *Let K_k be the k-clique for a fixed $k \ge 3$, and let $\delta > 0$. Then*

$$\lim_{n\to\infty} \frac{\phi_{K_k}(n, p, \delta)}{n^2 p^{k-1}\log(1/p)} = \begin{cases} \min\left\{\frac12 \delta^{2/k}\,,\ \delta/k\right\} & \text{if } n^{-1/(k-1)} \ll p \ll 1\,, \\ \frac12 \delta^{2/k} & \text{if } n^{-2/(k-1)} \ll p \ll n^{-1/(k-1)}\,. \end{cases}$$

Given Theorem 4.1, the analogue of Corollary 1.2 again follows from the new framework of Chatterjee and Dembo, which establishes (see [5, Theorem 1.2]) that for any fixed $k \ge 3$, the rate function of upper tails for K_k counts in $\mathcal{G}(n, p)$ is $(1 + o(1))\phi_{K_k}(n, p, \delta)$ provided that $p \ge n^{-\alpha}$ for some $\alpha = \alpha(k) > 0$ (in particular, any fixed $0 < \alpha < (4k^3 - 8k^2 + k + 3)^{-1}$ suffices).

Corollary 4.2. *For any fixed $k \ge 3$ there exists some $\alpha = \alpha(k) > 0$ so the following holds. For any fixed $\delta > 0$, if $n^{-\alpha} \le p \ll 1$ then*

$$\mathbb{P}\left(t(K_k, \mathcal{G}_{n,p}) \ge (1+\delta)p^{\binom{k}{2}}\right) = \exp\left[-(1-o(1))\min\left\{\tfrac12\delta^{2/k}\,,\ \delta/k\right\} n^2 p^{k-1}\log(1/p)\right].$$

4.1. **Proof of Theorem 4.1.** Let $K_{1,\ell-1}$ be the star on ℓ vertices, and let $e(H)$ and $\Delta(H)$ denote the number of edges and maximum degree in H, resp. The proof will follow from the same arguments used to prove Theorem 1.1, once we establish the next lemma.

Lemma 4.3. *Fix $k \geq 4$ and let H be a non-edgeless k-vertex graph other than K_k and $K_{1,k-1}$. If $U \in \mathcal{W}$ is a graphon with $0 \leq U \leq 1-p$ and $I_p(p+U) \lesssim p^{k-1}\log(1/p)$, then $t(H,U) \ll p^{e(H)}$.*

Towards the proof of this lemma, we need the following simple claim.

Claim 4.4. *Let $H=(V,E)$ be a nonempty graph on $k \geq 4$ vertices other than K_k and $K_{1,k-1}$. Then H has a spanning subgraph $H'=(V,E')$ with $\Delta(H') \leq 2$ and $e(H') > 2e(H)/(k-1)$.*

Proof. First, we may assume that $\Delta(H) > 2$, since if $\Delta(H) \leq 2$ then $H' = H$ suffices (as $e(H) > 2e(H)/(k-1)$ for $k \geq 4$). Second, if H is acyclic then $2e(H)/(k-1) \leq 2$, so one can form H' via 2 edges incident to a vertex (recall $\Delta > 2$), along with another edge if needed (either disjoint or extending that path, recalling $H \neq K_{1,k-1}$). Thus, if we suppose H is a counterexample to the claim with a minimum number of edges, then H must contains a cycle.

Let $C = (v_0, \ldots, v_{\ell-1})$ be a longest cycle of H (so that $v_i v_{i+1} \in E$, indices taken modulo ℓ). Then $\ell < k$, otherwise we could take $E(H') = E(C)$, since $k > 2e(H)/(k-1)$ for $H \neq K_k$.

Denote by ∂C the set of edges in H with at least one endpoint in C_ℓ. We claim that $|\partial C| < \ell(k-1)/2$. Indeed, for any i, the vertices v_i and v_{i+1} cannot have any common neighbors outside C (as otherwise a longer cycle can be formed). Hence, every $u \notin C$ can be connected to at most $\lfloor \ell/2 \rfloor$ vertices in C, and unless all $\binom{\ell}{2}$ potential edges between the vertices of C are present, $|\partial C| < (k-\ell)\lfloor \ell/2 \rfloor + \binom{\ell}{2} \leq \ell(k-1)/2$. On the other hand, if these $\binom{\ell}{2}$ edges all belong to H, then every $u \notin C$ can be connected to at most one vertex in C (otherwise a longer cycle exists), whence $|\partial C| \leq k - \ell + \binom{l}{2} < \ell(k-1)/2$ (the last inequality used $2 < \ell < k$).

It follows that $e(H) > |\partial C|$, or else $2e(H)/(k-1) < \ell$ and again we can take $E(H') = E(C)$. Finally, let $H_1 = (V, E(H) \setminus \partial C)$. As established above, $e(H_1) > e(H) - \ell(k-1)/2$, so it would suffice to find a subgraph H_1' of it with $\Delta(H_1') \leq 2$ and $e(H_1') \geq 2e(H_1)/(k-1)$, to which we can add the cycle C as a separate connected component. Indeed such a subgraph H_1' exists, since $0 < e(H_1) < e(H)$ and H was assumed to be a counterexample minimizing $e(H)$. ∎

Proof of Lemma 4.3. By Corollary 3.5 (as used in the first step in the proof of Lemma 3.2), $\mathbb{E}[U^2] \leq (1+o(1))\mathbb{E}[I_p(p+U)]/I_p(1) \lesssim p^{k-1}$. Next, as a consequence of the generalized Hölder's inequality [9] (see [19, Corollary 3.2]),

$$t(F,U) \leq \mathbb{E}\big[U^d\big]^{e(F)/d} \qquad \text{for any graph } F \text{ with } \Delta(F) \leq d\,. \tag{4.2}$$

So, by combining these inequalities, $t(H',U) \lesssim p^{(k-1)e(H')/2}$ holds for any H' with $\Delta(H') \leq 2$. Taking H' as provided by Claim 4.4, we find that $t(H,U) \leq t(H',U) \ll p^{e(H)}$, as desired. ∎

The upper bound of Theorem 4.1 on ϕ_{K_k} is obtained via the same constructions of §2, with modified part sizes: a copy of K_r for $r = \delta^{1/k} n p^{(k-1)/2}$ or a copy of $K_{r,n-r}$ for $r = (\delta/k) n p^{k-1}$. For the lower bound, one decomposes $t(K_k, W)$ as in (3.1), in which, by Lemma 4.3, all terms other than $t(K_k,U)$ and $t(K_{1,k-1},U)$ are negligible. The remaining terms, resp. analogous to $t(U)$ and $s(U)$ in §3, are treated as in §3 (e.g., $\lambda(B) \lesssim p^{k-1}/b$ replaces $\lambda(B) \lesssim p^2/b$ in Lemma 3.1) with one exception: instead of (3.6), write $t(K_k,U) - t(K_k,U') \leq k\lambda(B)t(K_{k-1},U)$; we wish this quantity to be $o\big(p^{\binom{k}{2}}\big)$, and indeed, since $t(K_{k-1},U) \leq t(H',U) \lesssim p^{(k-1)e(H)/2}$ for any $H' \subset K_{k-1}$ with $\Delta(H') \le 2$ (as in the proof of Lemma 4.3), letting $H' = C_{k-1}$ (recall that $k \ge 4$) yields $t(K_{k-1}, U) \lesssim p^{(k-1)^2/2}$, and using $\lambda(B) \lesssim p^{k-1}/b$ with $b \gg p^{(k-1)/2}$ completes the proof. ■

4.2. **General subgraph counts.** It is worthwhile noting that the analysis of cliques from the previous section readily implies that, for any fixed graph F with maximum degree Δ,

$$\phi_F(n,p,\delta) \asymp n^2 p^\Delta \log(1/p) \qquad \text{whenever } p \gg n^{-1/\Delta} . \tag{4.3}$$

Consequently (again via [5]), there is some $\alpha = \alpha(F) > 0$ such that the rate function $R(n,p,\delta)$ for observing a number of F-copies that is $(1+\delta)$ times its mean in $\mathcal{G}_{n,p}$ for $p \ge n^{-\alpha}$ is of order $n^2p^\Delta \log(1/p)$ (the best previous bounds here, cf. [12], were $n^2 p^\Delta \lesssim R(n,p,\delta) \lesssim n^2p^\Delta \log(1/p)$).

Corollary 4.5. *Let F be a fixed graph with maximum degree Δ. There exist $\alpha = \alpha(F) > 0$ such that, for any fixed $\delta > 0$ and any $p \ge n^{-\alpha}$,*

$$-\log \mathbb{P}\big(t(F, \mathcal{G}_{n,p}) \ge (1+\delta)p^{e(F)}\big) \asymp n^2 p^\Delta \log(1/p) .$$

Indeed, assume $\Delta \ge 2$ (the case $\Delta = 1$ is trivial). For the upper bound on ϕ_F in (4.3), take a copy of $K_{r,n-r}$ for $r = \delta n p^\Delta$ (as in §2). For the lower bound, let W be such that $t(F,W) \ge (1+\delta)p^{e(F)}$ and write $U = W - p$ (so $0 \le U \le 1-p$). As in (3.1), we decompose $t(F,W) - p^{e(F)}$ into $\sum_{H \subseteq F} \theta_{F,H}\, p^{e(F)-e(H)}\, t(H,U)$ for some positive constants $\{\theta_{F,H}\}$, and by the assumption on $t(F,W)$ there must exist some $H \subseteq F$ with $t(H,U) \gtrsim p^{e(H)}$. However, by (4.2), $t(H,U) \le \mathbb{E}[U^\Delta]^{e(H)/\Delta}$, which is at most $\mathbb{E}[U^2]^{e(H)/\Delta}$ as $\Delta \ge 2$. Combining these, $\mathbb{E}[U^2] \gtrsim p^\Delta$, and yet (by Corollary 3.5, as before) $\mathbb{E}[U^2] \lesssim \mathbb{E}[I_p(p+U)]/I_p(1)$, as claimed.

5. Weak regularity

In this section, we give a short proof establishing (1.2) and Corollary 1.2 for slowly decreasing p, namely $(\log n)^{-1/6} \ll p \ll 1$, without requiring the new results of Chatterjee and Dembo. The lower bound on the tail probability is explained in the paragraph immediately following Corollary 1.2. The upper bound is established through the following proposition.

Proposition 5.1. *Let $0 < \eta < \delta$ and $0 < p < 1$. Then*

$$\mathbb{P}(t(\mathcal{G}_{n,p}) \ge (1+\delta)p^3) \le R \exp\left(-\phi(n,p,\delta-\eta)\right) , \tag{5.1}$$

with $R = M^n \varepsilon^{-M^2}$ where $\varepsilon = \eta p^3/6 < 1$ and $M = 4^{1/\varepsilon^2}$.

Assume $\delta > 0$ is fixed and $(\log n)^{-1/6} \ll p \ll 1$. Take a slowly decreasing $\eta = \eta_n$ so that $p^{-3}(\log n)^{-1/2} \ll \eta \ll 1$. Then $\varepsilon = \eta p^3/6 \gg (\log n)^{-1/2}$, and so, $M = 4^{o(\log n)} = n^{o(1)}$. Thus,

$$\begin{aligned}\log R = n \log M + M^2 \log(1/\varepsilon) &\ll n \log n + n^{o(1)} \log\log n \\ &\ll n^2p^2 \log(1/p) \asymp \phi(n,p,\delta) \sim \phi(n,p,\delta - o(1)) .\end{aligned}$$

It then follows by Proposition 5.1 that

$$\mathbb{P}(t(\mathcal{G}_{n,p}) \ge (1+\delta)p^3) \le \exp\left(-(1-o(1))\phi(n,p,\delta)\right) ,$$

which implies the upper bound in (1.2). More generally, one needs $p \gg (\log n)^{-1/(2e(H))}$ in order to use this method for upper tails of H-counts, where $e(H)$ is the number of edges in H.

We proceed to prove Proposition 5.1. Define the relative edge-density between two nonempty subsets of vertices $A, B \subseteq V(G)$ as $d_G(A,B) := |\{(a,b) \in A \times B : ab \in E(G)\}|/(|A|\,|B|)$.

Lemma 5.2. *Let $A_1, \dots, A_m$ be a partition of $V = \{1, \dots, n\}$ into nonempty sets. Let $\delta > 0$, let $0 < p < 1$ and take $0 \le d_{ij} \le 1$ and $d_{ij} = d_{ji}$ for each $1 \le i, j \le m$. Suppose that*

$$\frac{1}{n^3} \sum_{i,j,k} |A_i|\,|A_j|\,|A_k|\, d_{ij} d_{ik} d_{jk} \ge (1+\delta) p^3 \,.$$

Then for a random graph $G \sim \mathcal{G}_{n,p}$ on the vertex set V we have

$$\mathbb{P}(d_G(A_i, A_j) \ge d_{ij} \text{ for all } 1 \le i \le j \le m) \le \exp\left(-\phi(n,p,\delta)\right) \,.$$

Proof. Define $I_p^{>}(x) := I_p(\max\{x, p\})$. We know that a binomial random variable $X \sim \mathrm{Bin}(N, p)$ satisfies $\mathbb{P}(X \ge \delta N) \le \exp(-N I_p^{>}(\delta))$. We have

$$\begin{aligned} &\mathbb{P}\left(d_G(A_i, A_j) \ge d_{ij}\right) \le \exp\left(-\,|A_i|\,|A_j|\, I_p^{>}(d_{ij})\right) \quad \text{ if } i \ne j\,, \\ &\mathbb{P}\left(d_G(A_i, A_i) \ge d_{ii}\right) \le \exp\left(-\tbinom{|A_i|}{2} I_p^{>}(d_{ii})\right) \,. \end{aligned}$$

Let

$$I_p^{>}(A, d) := \sum_{1 \le i < j \le m} |A_i|\,|A_j|\, I_p^{>}(d_{ij}) + \sum_{i=1}^{m} \tbinom{|A_i|}{2} I_p^{>}(d_{ii}) \,.$$

Since $A_1, \dots, A_m$ are disjoint, we have

$$\mathbb{P}\left(d_G(A_i, A_j) \ge d_{ij} \text{ for all } 1 \le i \le j \le m\right) \le \exp\left(-I_p^{>}(A, d)\right) \le \exp\left(-\phi(n,p,\delta)\right) \,,$$

where the last step follows from the following observation: if $G' \in \mathcal{G}_n$ is the weighted graph on vertex set V obtained by setting $g_{xy} = \max\{d_{ij}, p\}$ whenever $x \in A_i$ and $y \in A_j$, then $t(G') \ge (1+\delta)p^3 n^3$, so that $I_p^{>}(A, d) = I_p(G') \ge \phi(n, p, \delta)$ by our definition (1.1) of ϕ. ∎

The following lemma is a consequence of the Frieze–Kannan weak regularity lemma and an associated counting lemma (see [16, §9.1, §10.5]).

Lemma 5.3. *Let $\varepsilon > 0$ and let G be a graph with n vertices. Then there exists a partition $\mathcal{P}$ of the vertices of G into at most $4^{1/\varepsilon^2}$ parts $A_1, \dots, A_m$ so that if $d_{ij} = d_G(A_i, A_j)$, then*

$$\left| t(G) - n^{-3} \sum_{i,j,k=1}^{m} |A_i|\,|A_j|\,|A_k|\, d_{ij} d_{ik} d_{jk} \right| \le 3\varepsilon \,.$$

Proof of Proposition 5.1. Let G be any graph on n vertices satisfying $t(G) \ge (1+\delta)p^3$. By Lemma 5.3, there exists a partition of its vertices into $m \le M$ parts $A_1, A_2, \dots, A_m$, so that

$$n^{-3} \sum_{i,j,k=1}^{m} |A_i|\,|A_j|\,|A_k|\, d_{ij} d_{ik} d_{jk} \ge (1+\delta)p^3 - 3\varepsilon \,,$$

where $d_{ij} = d_G(A_i, A_j)$. Let d'_{ij} be d_{ij} rounded down to the nearest multiple of ε. Then

$$n^{-3} \sum_{i,j,k=1}^{m} |A_i|\,|A_j|\,|A_k|\, d'_{ij} d'_{ik} d'_{jk} \ge (1+\delta)p^3 - 6\varepsilon = (1 + \delta - \eta)p^3 \,.$$

For any fixed choice of $\{A_i\}_i$, $\{d'_{ij}\}_{i,j}$, by Lemma 5.2 we have

$$\mathbb{P}(d_G(A_i, A_j) \ge d'_{ij} \text{ for all } 1 \le i \le j \le m) \le \exp\left(-\phi(n, p, \delta - \eta)\right) \,.$$

A union bound over the A_i's ($\le M^n$ choices) and d'_{ij}'s ($\le \varepsilon^{-M^2}$ choices) now yields (5.1). ∎

References

[1] B. B. Bhattacharya, S. Ganguly, E. Lubetzky, and Y. Zhao. Upper tails and independence polynomials in random graphs. Preprint, available at `arXiv:1507.04074`.

[2] B. Bollobás. *Random graphs*, volume 73 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2001.

[3] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi. Convergent sequences of dense graphs. I. Subgraph frequencies, metric properties and testing. *Adv. Math.*, 219(6):1801–1851, 2008.

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

[5] S. Chatterjee and A. Dembo. Nonlinear large deviations. Preprint, available at `arXiv:1401.3495`.

[6] S. Chatterjee and S. R. S. Varadhan. The large deviation principle for the Erdős-Rényi random graph. *European J. Combin.*, 32(7):1000–1017, 2011.

[7] B. Demarco and J. Kahn. Tight upper tail bounds for cliques. *Random Structures Algorithms*, 41(4):469–487, 2012.

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

[9] H. Finner. A generalization of Hölder's inequality and some probability inequalities. *Ann. Probab.*, 20(4):1893–1901, 1992.

[10] A. Frieze and R. Kannan. Quick approximation to matrices and applications. *Combinatorica*, 19(2):175–220, 1999.

[11] S. Janson, T. Łuczak, and A. Rucinski. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.

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

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

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

[15] J. H. Kim and V. H. Vu. Divide and conquer martingales and the number of triangles in a random graph. *Random Structures Algorithms*, 24(2):166–174, 2004.

[16] L. Lovász. *Large networks and graph limits*, volume 60 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2012.

[17] L. Lovász and B. Szegedy. Limits of dense graph sequences. *J. Combin. Theory Ser. B*, 96(6):933–957, 2006.

[18] L. Lovász and B. Szegedy. Szemerédi's lemma for the analyst. *Geom. Funct. Anal.*, 17(1):252–270, 2007.

[19] E. Lubetzky and Y. Zhao. On replica symmetry of large deviations in random graphs. *Random Structures Algorithms*, 47:109—146, 2015.

[20] E. Szemerédi. Regular partitions of graphs. In *Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976)*, volume 260 of *Colloq. Internat. CNRS*, pages 399–401. CNRS, Paris, 1978.

[21] V. H. Vu. A large deviation result on the number of small subgraphs of a random graph. *Combin. Probab. Comput.*, 10(1):79–94, 2001.

E. Lubetzky

Courant Institute of Mathematical Sciences, New York University, New York, NY 10012, USA

E-mail address: `eyal@courant.nyu.edu`

Y. Zhao

Mathematical Institute, University of Oxford, Oxford OX2 6GG, United Kingdom

E-mail address: `yufei.zhao@maths.ox.ac.uk`