UPPER TAILS AND INDEPENDENCE POLYNOMIALS IN RANDOM GRAPHS

BHASWAR B. BHATTACHARYA, SHIRSHENDU GANGULY, EYAL LUBETZKY, AND YUFEI ZHAO

Abstract. The upper tail problem in the Erdős-Rényi random graph $G \sim \mathcal{G}_{n,p}$ asks to estimate the probability that the number of copies of a graph H in G exceeds its expectation by a factor $1+\delta$. Chatterjee and Dembo (2014) showed that in the sparse regime of $p \to 0$ as $n \to \infty$ with $p \geq n^{-\alpha}$ for an explicit $\alpha = \alpha_H > 0$, this problem reduces to a natural variational problem on weighted graphs, which was thereafter asymptotically solved by two of the authors in the case where H is a clique.

Here we extend the latter work to any fixed graph H and determine a function $c_H(\delta)$ such that, for p as above and any fixed $\delta > 0$, the upper tail probability is $\exp[-(c_H(\delta) + o(1))n^2 p^\Delta \log(1/p)]$, where Δ is the maximum degree of H. As it turns out, the leading order constant in the large deviation rate function, $c_H(\delta)$, is governed by the independence polynomial of H, defined as $P_H(x) = \sum i_H(k)x^k$ where $i_H(k)$ is the number of independent sets of size k in H. For instance, if H is a regular graph on m vertices, then $c_H(\delta)$ is the minimum between $\frac{1}{2}\delta^{2/m}$ and the unique positive solution of $P_H(x) = 1+\delta$.

1. Introduction

1.1. The upper tail problem in the random graph. Let $\mathcal{G}_{n,p}$ be the Erdős-Rényi random graph on n vertices with edge probability p, and let X_H be the number of copies of a fixed graph H in it. The upper tail problem for X_H asks to estimate the large deviation rate function given by

$$R_H(n,p,\delta) := -\log \mathbb{P}\left(X_H \geq (1+\delta)\mathbb{E}[X_H]\right) \qquad \text{for fixed } \delta > 0\,,$$

a classical and extensively studied problem (cf. [6,11,12,15–18,24] and [1,14] and the references therein) which already for the seemingly basic case of triangles ($H = K_3$) is highly nontrivial and still not fully understood. It followed from works of Vu [24] and Kim and Vu [18] that[1]

$$n^2p^2 \lesssim R_{K_3}(n,p,\delta) \lesssim n^2p^2\log(1/p)$$

(the lower bound used the so-called "polynomial concentration" machinery, whereas the upper bound follows, e.g., from the fact that an arbitrary set of $s \sim \delta^{1/3}np$ vertices can form a clique in $\mathcal{G}_{n,p}$ with probability $p^{\binom{s}{2}} = p^{O(n^2p^2)}$, thus contributing $\binom{s}{3} \sim \frac{1}{6}\delta n^3p^3 \sim \delta\mathbb{E}[X_{K_3}]$ extra triangles); the correct order of the rate function was settled fairly recently by Chatterjee [6], and independently by DeMarco and Kahn [12], proving that $R_{K_3}(n,p,\delta) \asymp n^2p^2\log(1/p)$ for $p \geq \frac{\log n}{n}$. This was later extended in [11] to cliques ($H = K_k$ for $k \geq 3$), establishing that for $p \geq n^{-2/(k-1)+\varepsilon}$ with $\varepsilon > 0$ fixed[2],

$$R_{K_k}(n,p,\delta) \asymp n^2p^{k-1}\log(1/p)\,.$$

The methods of [6, 11, 12] did not allow recovering the exact asymptotics of this rate function, and in particular, one could ask, e.g., whether $R_{K_3}(n,p,\delta) \sim c(\delta)n^2p^2\log(1/p)$ with $c(\delta) = \frac{1}{2}\delta^{2/3}$, as the aforementioned clique upper bound for it may suggest (recalling its probability was $p^{\binom{s}{2}}$ for $s \sim \delta^{1/3}np$).

Much progress has since been made in that front, propelled by the seminal work of Chatterjee and Varadhan [9] that introduced a large deviation framework for $\mathcal{G}_{n,p}$ in the *dense regime* ($0 < p < 1$ fixed) via the theory of graph limits (cf. [8, 9, 22] for more on the many questions still open in that regime).

[1]We write $f \lesssim g$ to denote $f = O(g)$; $f \asymp g$ means $f = \Theta(g)$; $f \sim g$ means $f = (1+o(1))g$ and $f \ll g$ means $f = o(g)$.

[2]More precisely, DeMarco and Kahn [11] showed that $R_{K_k} \asymp \min\{n^2p^{k-1}\log(1/p), n^kp^{\binom{k}{2}}\}$ for $p \geq n^{-2/(k-1)}$.

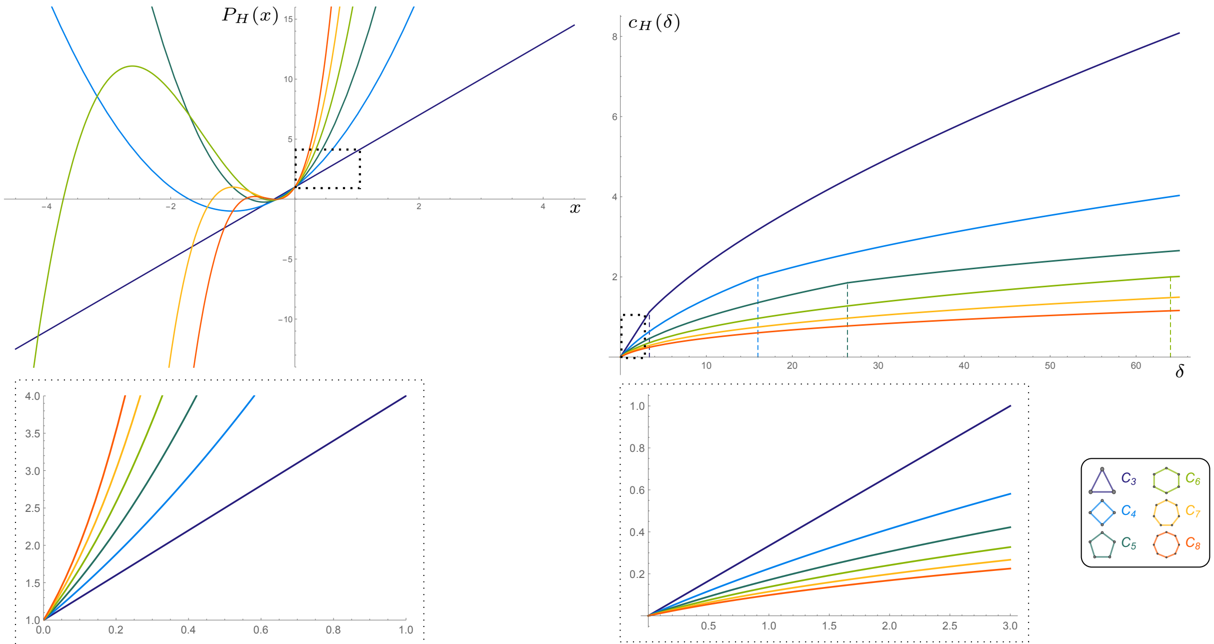


FIGURE 1. The leading order constant $c_H(\delta)$ for the upper tail rate function for k-cycles vs. their independence polynomials $P_H(x)$. Zoomed-in regions show $P_H(c_H(\delta)) = 1 + \delta$.

In the sparse regime ($p \to 0$), in the absence of graph limit tools, the understanding of large deviations for a fixed graph H (be it even a triangle) remained very limited until a recent breakthrough paper of Chatterjee and Dembo [7] that reduced it to a natural variational problem in a certain range of p (see Definition 1.3 and Theorem 1.4 below). In [23], this variational problem was solved asymptotically for triangles, thereby yielding the following conclusion: for fixed $\delta > 0$, if $n^{-1/42} \log n \le p = o(1)$, then

$$R_{K_3}(n, p, \delta) \sim c(\delta) n^2 p^2 \log(1/p) \quad \text{where} \quad c(\delta) = \min\left\{\tfrac{1}{2}\delta^{2/3}, \tfrac{1}{3}\delta\right\}, \tag{1.1}$$

and we see that the clique construction from above gives the correct leading order constant iff $\delta > 27/8$. More generally, for every $k \ge 3$ there is an explicit $\alpha_k > 0$ so that, for fixed $\delta > 0$, if $n^{-\alpha_k} \le p = o(1)$,

$$R_{K_k}(n, p, \delta) \sim c_k(\delta) n^2 p^{k-1} \log(1/p) \quad \text{where} \quad c_k(\delta) = \min\left\{\tfrac{1}{2}\delta^{2/k}, \tfrac{1}{k}\delta\right\}. \tag{1.2}$$

For a *general* fixed graph H with maximum degree $\Delta \ge 2$ (when $\Delta = 1$ the problem is nothing but the large deviation in the binomial variable corresponding to the edge-count) the order of the rate function was established up to a multiplicative $\log(1/p)$ factor by Janson, Oleszkiewicz, and Rucinński [15]; in the range $p \ge n^{-1/\Delta}$, their estimate (which involves a complicated quantity $M_H^*(n,p)$) simplifies into

$$n^2 p^\Delta \lesssim R_H(n, p, \delta) \lesssim n^2 p^\Delta \log(1/p)$$

(with constants depending on H and on δ). As a byproduct of the analysis of cliques in [23], it was shown [23, Corollary 4.5] that there is some explicit $\alpha_H > 0$ so that, for fixed $\delta > 0$, if $n^{-\alpha_H} \le p = o(1)$,

$$R_H(n, p, \delta) \asymp n^2 p^\Delta \log(1/p),$$

yet those bounds were not sharp already for the 4-cycle C_4. Here we extend that work and determine the precise asymptotics of $R_H(n, p, \delta)$ for any fixed graph H in the above mentioned range $n^{-\alpha_H} \le p = o(1)$ (as currently needed in the framework of [7]). Solving the variational problem for a general H requires significant new ideas atop [23], and turns out to involve the *independence polynomial* $P_H(x)$ (see Fig. 1).

Definition 1.1 (Independence polynomial)**.** *The* independence polynomial *of H is defined to be*

$$P_H(x) := \sum_k i_H(k) x^k \,, \tag{1.3}$$

where $i_H(k)$ is the number of k-element independent sets in H.

Definition 1.2 (Inducing on maximum degrees)**.** *For a graph H with maximum degree Δ, let H^* be the induced subgraph of H on all vertices whose degree in H is Δ. (Note that $H^* = H$ if H is regular.)*

Roots of independence polynomials were studied in various contexts (cf. [4,5,10] and their references); here, the unique positive x such that $P_{H^*}(x) = 1 + \delta$ will, perhaps surprisingly, give the leading order constant (possibly capped at some maximum value if H happens to be regular) of $R_H(n, p, \delta)$.

1.2. **Variational problem.** For graphs G and H, denote by $\hom(H, G)$ the number of homomorphisms from H to G (a graph homomorphism is a map $V(H) \to V(G)$ that carries every edge of H to an edge of G). The homomorphism density of H in G is defined as $t(H, G) := \hom(H, G)\, |V(G)|^{-|V(H)|}$, that is, the probability that a uniformly random map $V(H) \to V(G)$ is a homomorphism from H to G.

Henceforth, we will work with $t(H, G)$ for $G \sim \mathcal{G}_{n,p}$ instead of X_H for convenience (the two quantities are nearly proportional, as the only possible discrepancies—non-injective homomorphisms from H to G—are a negligible fraction of all homomorphisms when G is sufficiently large and not too sparse).

Chatterjee and Dembo [7] proved a non-linear large deviation principle, and in particular derived the exact asymptotics of the rate function for a general graph H in terms of a variational problem.

Definition 1.3 (Discrete variational problem)**.** *Let $\mathscr{G}_n$ denote the set of weighted undirected graphs on n vertices with edge weights in $[0, 1]$, that is, if $A(G)$ is the adjacency matrix of G then*

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

Let H be a fixed graph with maximum degree Δ. The variational problem for $\delta > 0$ and $0 < p < 1$ is

$$\phi(H, n, p, \delta) := \inf \left\{ I_p(G_n) : G_n \in \mathscr{G}_n \text{ with } t(H, G_n) \ge (1 + \delta) p^{|E(H)|} \right\} \tag{1.4}$$

where

$$t(H, G_n) := n^{-|V(H)|} \sum_{1 \le i_1, \cdots, i_k \le n} \prod_{(x,y) \in E(H)} a_{i_x i_y}$$

is the density of (labeled) copies of H in G_n, and $I_p(G_n)$ is the entropy relative to p, that is,

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

Theorem 1.4 (Chatterjee and Dembo [7])**.** *Let H be a fixed graph. There is some explicit $\alpha_H > 0$ such that for $n^{-\alpha_H} \le p < 1$ and any fixed $\delta > 0$,*

$$\mathbb{P}\left(t(H, \mathcal{G}_{n,p}) \ge (1 + \delta) p^{|E(H)|} \right) = \exp\bigl(-(1 + o(1)) \phi(H, n, p, \delta)\bigr) \,,$$

where $\phi(H, n, p, \delta)$ is as defined in (1.4) and the $o(1)$-term goes to zero as $n \to \infty$.

Thanks to this theorem, solving the variational problem $\phi(H, n, p, \delta)$ asymptotically would give the asymptotic rate function for H when $n^{-\alpha_H} \le p = o(1)$ (as was done for $H = K_k$ in [23], yielding (1.2)).

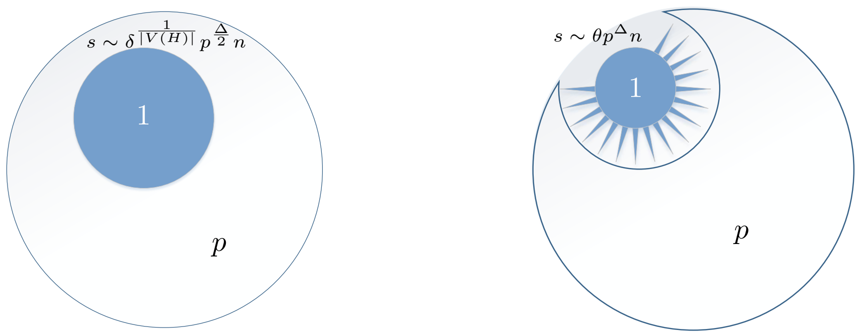


FIGURE 2. Solution candidates for discrete variational problem (clique and anti-clique).

1.3. **Main Result.** Let H be a graph with maximum degree $\Delta = \Delta(H)$; recall that H is *regular* (or Δ-*regular*) if all its vertices have degree Δ, and *irregular* otherwise. Starting with a weighted graph G_n with all edge-weights a_{ij} equal to p, we consider the following two ways of modifying the G_n so it would satisfy the constraint $t(H, G_n) \geq (1+\delta)p^{|E(H)|}$ of the variational problem (1.4) (see Figure 2).

(a) (Planting a clique) Set $a_{ij} = 1$ for all $1 \leq i, j \leq s$ for $s \sim \delta^{1/|V(H)|} p^{\Delta/2} n$. This construction is effective only when H is Δ-regular, in which case it gives $t(H, G_n) \sim (1+\delta)p^{|E(H)|}$.

(b) (Planting an anti-clique) Set $a_{ij} = 1$ whenever $i \leq s$ or $j \leq s$ for $s \sim \theta p^{\Delta} n$ for $\theta = \theta(H, \delta) > 0$ such that $P_{H^*}(\theta) = 1 + \delta$, in which case $t(H, G_n) \sim (1+\delta)p^{|E(H)|}$.

We postpone the short calculation that in each case $t(H, G_n) \sim (1+\delta)p^{|E(H)|}$ to §2. Our main result (Theorem 1.5 below) says that, for a connected graph H and $n^{-1/\Delta} \ll p \ll 1$, one of these constructions has $I_p(G_n)$ that is within a $(1+o(1))$-factor of the optimum achieved by the variational problem (1.4). For example, when $H = K_3$, the clique construction has $I_p(G_n) \sim \frac{1}{2}s^2 I_p(1) \sim \frac{1}{2}\delta^{2/3} n^2 p^2 \log(1/p)$, while $P_{K_3}(x) = 1 + 3x$ so $\theta = \delta/3$ and the anti-clique construction has $I_p(G_n) \sim s n I_p(1) \sim \frac{1}{3}\delta n^2 p^2 \log(1/p)$ (thus the clique is more advantageous iff $\delta > 27/8$), exactly the bounds that were featured in (1.1). The following result extends [23, Theorems 1.1 and 4.1] from cliques to the case of a general graph H.

Theorem 1.5. *Let H be a fixed connected graph with maximum degree $\Delta \geq 2$. For any fixed $\delta > 0$ and $n^{-1/\Delta} \ll p = o(1)$, the solution to the discrete variational problem* (1.4) *satisfies*

$$\lim_{n\to\infty} \frac{\phi(H, n, p, \delta)}{n^2 p^{\Delta} \log(1/p)} = \begin{cases} \min\left\{\theta\,,\, \frac{1}{2}\delta^{2/|V(H)|}\right\} & \text{if } H \text{ is regular} \\ \theta & \text{if } H \text{ is irregular} \end{cases},$$

where $\theta = \theta(H, \delta)$ is the unique positive solution to $P_{H^*}(\theta) = 1 + \delta$ (for $P_{H^*}(x)$ as per Defs. 1.1–1.2).

When combined with Theorem 1.4, this yields the following conclusion for the upper tail problem.

Corollary 1.6. *Let H be a fixed connected graph with maximum degree $\Delta \geq 2$. There exists $\alpha_H > 0$ such that for $n^{-\alpha_H} \leq p \ll 1$ the following holds. For any fixed $\delta > 0$,*

$$\lim_{n\to\infty} \frac{-\log \mathbb{P}\left(t(H, \mathcal{G}_{n,p}) \geq (1+\delta)p^{|E(H)|}\right)}{n^2 p^{\Delta} \log(1/p)} = \begin{cases} \min\left\{\theta\,,\, \frac{1}{2}\delta^{2/|V(H)|}\right\} & \text{if } H \text{ is regular} \\ \theta & \text{if } H \text{ is irregular} \end{cases},$$

where $\theta = \theta(H, \delta)$ is the unique positive solution to $P_{H^*}(\theta) = 1 + \delta$ (for $P_{H^*}(x)$ as per Defs. 1.1–1.2).

Observe that when H is regular, there exists a unique $\delta_0 = \delta_0(H) > 0$ such that

$$\theta(H,\delta) \le \tfrac{1}{2}\delta^{2/|V(H)|} \quad \text{iff} \quad \delta \le \delta_0(H)\,. \tag{1.5}$$

That is, the leading order constant of $\phi(H,n,p,\delta)$ (giving the asymptotic upper tail) is governed by the anti-clique for $\delta \le \delta_0$ and by the clique for $\delta \ge \delta_0$ (the above example of $H = K_3$ had $\delta_0 = 27/8$). Indeed, let $k = |V(H)| \ge 3$ and recall that $P_H(x)$ is an increasing (as it has nonnegative coefficients) bijection of $[0,\infty)$ onto $[1,\infty)$, whence (by definition) $\theta \le \frac{1}{2}\delta^{2/k}$ iff $1+\delta = P_H(\theta) \le P_H(\frac{1}{2}\delta^{2/k})$. Now, the function $f(\delta) = [P_H(\frac{1}{2}\delta^{2/k}) - 1]/\delta$ is decreasing in $[0,\infty)$ and satisfies $\lim_{\delta\to\infty} f(\delta) \le 2^{1-k/2} < 1$ and $\lim_{\delta\to 0} f(\delta) = \infty$, with a unique δ_0 in between at which $f(\delta_0) = 1$, and the observation follows.

We remark that our results extend (see Corollary 7.3) to any *disconnected* graph H. The interplay between different connected components can then cause the upper tail to be dominated, not by an exclusive appearance of either the clique or the anti-clique constructions (as was the case for any *connected* graph H, cf. Theorem 1.5), but rather by an interpolation of these; see §7.2 for more details.

The assumption $p \gg n^{-1/\Delta}$ in Theorem 1.5 is essentially tight in the sense that the upper tail rate function undergoes a phase transition at that location [15]: it is of order $n^{2+o(1)}p^{\Delta}$ for $p \ge n^{-1/\Delta}$, and below that threshold it becomes a function (denoted $M_H^*(n,p)$ in [15]) depending on all subgraphs of H. In terms of the discrete variational problem (1.4), again this threhold marks a phase transition, as the anti-clique construction ceases to be viable for $p \ll n^{-1/\Delta}$ (recall that $s \sim \theta p^{\Delta} n$ in that construction). Still, as in [23, Theorems 1.1 and 4.1], our methods show that if H is *regular* and $n^{-2/\Delta} \ll p \ll n^{-1/\Delta}$, the solution to the variational problem is $(1+o(1))\frac{1}{2}\delta^{2/|V(H)|}$ (i.e., governed by the clique construction).

Finally, several of the tools that were developed here to overcome the obstacles in extending the analysis of [23] to general graphs (arising already for the 4-cycle) may be of independent interest and find other applications, e.g., the crucial use of adaptively chosen degree-thresholds (see §4 for details).

1.4. **Examples.** We now demonstrate the solution of the variational problem (1.4), as provided by Theorem 1.5, for various families of graphs (adding to the previously known [23] case of cliques, cf. (1.2)).

Example 1.7 (k-cycle: $H = C_k$)**.** It is easy to verify that $P_{C_k}(x)$ satisfies the recursion[3]

$$P_{C_k}(x) = P_{C_{k-1}}(x) + xP_{C_{k-2}}(x)\,, \quad P_{C_2}(x) = 2x+1\,, \quad P_{C_3}(x) = 3x+1\,.$$

For instance, $P_{C_4}(x) = 2x^2+4x+1$ and $P_{C_5}(x) = 5x^2+5x+1$; by Theorem 1.5, if $n^{-1/2} \ll p \ll 1$,

$$\phi(C_4,n,p,\delta) \sim \min\left\{\theta(C_4,\delta)\,,\tfrac{1}{2}\delta^{1/2}\right\} n^2p^2\log(1/p) \qquad \text{for} \qquad \theta(C_4,\delta) = -1+\sqrt{1+\tfrac{1}{2}\delta}\,,$$

$$\phi(C_5,n,p,\delta) \sim \min\left\{\theta(C_5,\delta)\,,\tfrac{1}{2}\delta^{2/5}\right\} n^2p^2\log(1/p) \qquad \text{for} \qquad \theta(C_5,\delta) = -\tfrac{1}{2}+\tfrac{1}{2}\sqrt{1+\tfrac{4}{5}\delta}\,.$$

For general k, with square brackets denoting extraction of coefficients, $[x]P_{C_k}(x) = k$ (more generally, $[x]P_H(x) = |V(H)|$ for any H), while the closely related recursion for Chebyshev's polynomials yields

$$P_{C_k}(x) = 2^{1-k}\sum_{j=0}^{\lfloor k/2\rfloor}\binom{k}{2j}(1+4x)^j\,, \tag{1.6}$$

and so $[x^2]P_{C_k}(x) = \frac{1}{2}k(k-3)$; e.g., for any $k \ge 4$, the behavior of $\theta(C_k,\delta)$ for small δ (see Fig. 1) is

$$\theta(C_k,\delta) = \tfrac{1}{k-3}\left(-1+\sqrt{1+2\delta(k-3)/k}\right) + O(\delta^3) = \tfrac{1}{k}\delta + \tfrac{3-k}{2k^2}\delta^2 + O(\delta^3)\,.$$

[3] By the definition of the independence polynomial, for any graph H and vertex v in it, $P_H(x) = P_{H_1}(x) + xP_{H_2}(x)$, where H_1 is obtained from H by deleting v and H_2 is obtained from H by deleting v and all its neighbors.

Finally, observe that for even k we can write (1.6) as $P_{C_k}(x) = \left[\frac{1}{2}(\sqrt{1+4x}+1)\right]^k + \left[\frac{1}{2}(\sqrt{1+4x}-1)\right]^k$, and deduce that the value of $P_{C_k}(\frac{1}{2}\delta^{2/k})$ for $\delta = 2^k$ is simply $P_{C_k}(2) = 2^k + 1 = 1+\delta$. Thus, by the remark following Corollary 1.6, the transition addressed in (1.5) occurs at $\delta_0(C_k) = 2^k$ for even k; e.g.,

$$\lim_{n\to\infty} \frac{\phi(C_4,n,p,\delta)}{n^2p^2\log(1/p)} = \begin{cases} \frac{1}{2}\sqrt{\delta} & \text{if } \delta < 16\,, \\ -1 + \sqrt{1+\frac{1}{2}\delta} & \text{if } \delta \geq 16\,. \end{cases} \tag{1.7}$$

As mentioned above, $H = C_4$ is the simplest graph for which the arguments in [23] did not give sharp bounds on $\phi(H,n,p,\delta)$, and its treatment is instrumental for the analysis of general graphs (see §4.2).

Example 1.8 (binary tree)**.** Letting T_h denote the complete binary tree of height h ($|V(T_h)| = 2^h - 1$), observe that, by counting independent sets excluding/including the root, $P_{T_h}(x)$ satisfies the recursion

$$P_{T_h}(x) = P_{T_{h-1}}(x)^2 + xP_{T_{h-2}}(x)^4\,, \quad P_{T_0}(x) = 1\,, \quad P_{T_1}(x) = x+1\,.$$

The polynomial $P_{T_h^*}(x)$ restricts us to independent sets of T_h where all degrees are Δ, and therefore

$$P_{T_h^*}(x) = P_{T_{h-2}}(x)^2\,,$$

as the restriction excludes precisely the root and leaves. (More generally, for the b-ary tree ($b \geq 2$) one has $P_{T_h}(x) = P_{T_{h-1}}(x)^b + xP_{T_{h-2}}(x)^{b^2}$, and $P_{T_h^*}(x) = P_{T_{h-2}}(x)^b$.)

For instance, the binary tree on 15 vertices, T_4, has

$$P_{T_4^*}(x) = P_{T_2}(x)^2 = x^4 + 6x^3 + 11x^2 + 6x + 1\,,$$

and solving $P_{T_4^*}(\theta) = 1+\delta$ we find that, as per Theorem 1.5, if $n^{-1/3} \ll p \ll 1$ then

$$\lim_{n\to\infty} \frac{\phi(T_4,n,p,\delta)}{n^2p^3\log(1/p)} = -\tfrac{3}{2} + \tfrac{1}{2}\sqrt{5+4\sqrt{1+\delta}}\,. \tag{1.8}$$

For general h, we can for instance deduce from the recurrence above (and the facts $[x]P_H[x] = |V(H)|$ and $[x]P_{H^*}(x) = \#\{v : \deg(v) = \Delta\}$) that $[x]P_{T_h^*}(x) = 2^{h-1} - 2$ and $[x^2]P_{T_h^*}(x) = 2^{2h-3} - 7\cdot 2^{h-2} + 7$ for any $h \geq 3$, using which it is easy to write $\theta(T_h,\delta)$ explicitly up to an additive $O(\delta^3)$-term.

Example 1.9 (complete bipartite: $H = K_{k,\ell}$ for $k \geq \ell$)**.** In case $k > \ell$ we have $P_{K_{k,\ell}^*}(x) = (1+x)^\ell$ as we only count independent sets in the k-regular side (of size ℓ); thus, by Theorem 1.5, for $n^{-1/k} \ll p \ll 1$,

$$\lim_{n\to\infty} \frac{\phi(K_{k,\ell},n,p,\delta)}{n^2p^k\log(1/p)} = (1+\delta)^{1/\ell} - 1\,. \tag{1.9}$$

If $k = \ell$, the coefficients of x^j ($j \geq 1$) are doubled, so $P_{K_{k,\ell}}(x) = 2(1+x)^k - 1$ and for $n^{-1/k} \ll p \ll 1$,

$$\lim_{n\to\infty} \frac{\phi(K_{k,\ell},n,p,\delta)}{n^2p^k\log(1/p)} = \min\left\{\left(1+\tfrac{1}{2}\delta\right)^{1/k} - 1\,,\, \tfrac{1}{2}\delta^{1/k}\right\}\,. \tag{1.10}$$

1.5. **Organization.** In §2 we describe the graphon formulation of the variational problem. Section 3 contains preliminary estimates needed from [23]. Section 4 begins with a short proof for the known case of K_3, continues by illustrating some of the new techniques with the case of C_4, then outlines the approach for general graphs. Theorem 2.3 and Theorem 2.2 are proved in §5 and §6, respectively. Finally, §7 provides the details required to pass from the continuous to the discrete variational problems.

2. The graphon formulation of the variational problem

Following [23], we will analyze a continuous version of the discrete variational problem (1.4) that has the advantage of having no dependence on n. Recall that a *graphon* is a symmetric measurable function $W : [0,1]^2 \to [0,1]$ (where symmetric means $W(x,y) = W(y,x)$); in the continuous version of (1.4), W will replace the edge-weighted graph G_n, as the latter can be viewed as a discrete approximation of a graphon (see, e.g., [2, 3, 19, 21] for more on graph limits). As we will later see in §7 (see Lemma 7.1), the solution to this variational problem will immediately imply a lower bound on the problem (1.4). In what follows and throughout the paper, write $\mathbb{E}[f(W)] := \int_{[0,1]^2} f(W(x,y))\,\mathrm{d}x\mathrm{d}y$.

Definition 2.1 (Graphon variational problem)**.** *For $\delta > 0$ and $0 < p < 1$, let*

$$\phi(H,p,\delta) := \inf\left\{ \tfrac{1}{2}\mathbb{E}[I_p(W)] : \textit{graphon } W \textit{ with } t(H,W) \geq (1+\delta)p^{|E(H)|} \right\}, \tag{2.1}$$

where

$$t(H,W) := \int_{[0,1]^{|V(H)|}} \prod_{(i,j)\in E(H)} W(x_i,x_j)\,\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_{|V(H)|}\,.$$

For example, for $H = K_3$ we wish to minimize $\mathbb{E}[I_p(W)] := \int_{[0,1]^2} I_p(W(x,y))\mathrm{d}x\mathrm{d}y$ over all graphons W whose triangle density $t(K_3,W) = \int_{[0,1]^3} W(x,y)W(x,z)W(y,z)\,\mathrm{d}x\mathrm{d}y\mathrm{d}z$ is at least $(1+\delta)p^3$.

The solution of the graphon variational problem is given by the following two theorems.

Theorem 2.2. *Let H be a connected Δ-regular graph. Fix $\delta > 0$ and let $\theta = \theta(H,\delta)$ be the unique positive solution to $P_H(\theta) = 1+\delta$ (for $P_H(x)$ as defined in (1.3)). Then*

$$\lim_{p\to 0} \frac{\phi(H,p,\delta)}{p^{\Delta}\log(1/p)} = \min\left\{\theta\,,\,\tfrac{1}{2}\delta^{2/|V(H)|}\right\}\,.$$

Theorem 2.3. *Let H be a connected irregular graph with maximum degree Δ. Fix $\delta > 0$ and let $\theta = \theta(H,\delta)$ be the unique positive solution to $P_{H^*}(\theta) = 1+\delta$ (for $P_{H^*}(x)$ as per Defs. 1.1–1.2). Then*

$$\lim_{p\to 0} \frac{\phi(H,p,\delta)}{p^{\Delta}\log(1/p)} = \theta\,.$$

The proofs of the lower bounds are given in §5 and §6. The upper bounds for both theorems follow from the graphon analogs of the clique and anti-clique constructions in §1.3 (nearly identical computations yield the analogous upper bounds for the discrete variational problem (1.4)), as we establish in the following proposition.

Proposition 2.4. *For any fixed graph H with maximum degree Δ and any $\delta > 0$,*

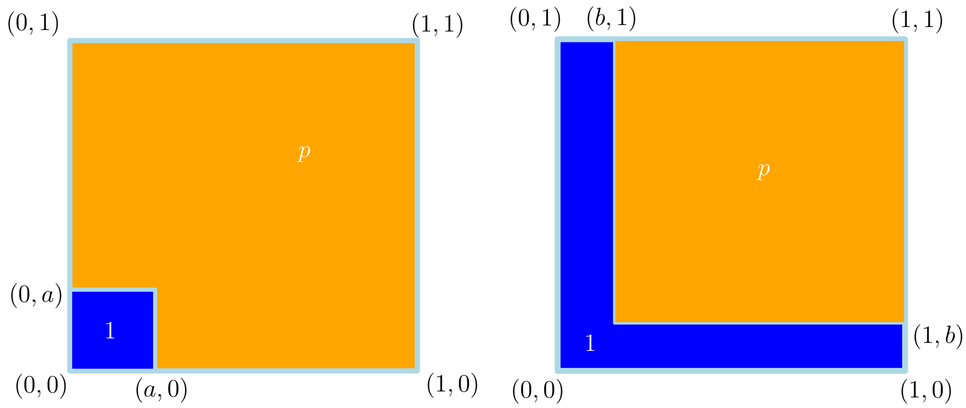
$$\phi(H,p,\delta) \leq (\theta + o(1))\,p^{\Delta}\log(1/p)\,, \tag{2.2}$$

for $\theta = \theta(H,\delta)$ the unique positive solution to $P_{H^*}(\theta) = 1+\delta$. Moreover, if H is regular and connected,

$$\phi(H,p,\delta) \leq \left(\tfrac{1}{2}\delta^{2/|V(H)|} + o(1)\right)p^{\Delta}\log(1/p)\,, \tag{2.3}$$

where in both cases the $o(1)$-terms vanish as $p \to 0$.

Throughout the paper, if H is a graph and $u \in V(H)$, let $N_H(v) = \{v : (u,v) \in E(H)\}$ be the neighborhood of v in H. For $S \subset V(H)$, write $N_H(S) = \cup_{v\in S} N_H(v)$ for the neighborhood of S, let H_S denote the induced subgraph of H on S, and let $\overline{S} = V \backslash S$.



(a) The clique graphon

(b) The anti-clique graphon

FIGURE 3. Solution candidates for the graphon variational problem ($a \asymp p^{\Delta/2}$ and $b \asymp p^{\Delta}$).

Proof of Proposition 2.4. Define the *clique graphon* W_1 and *anti-clique graphon* W_2 by

$$W_1 = p + (1-p)\mathbf{1}_{[0,a]^2} \qquad \text{for } a = \delta^{1/|V(H)|}\, p^{\Delta/2}\,, \tag{2.4}$$

$$W_2 = 1 - (1-p)\mathbf{1}_{[b,1]^2} \qquad \text{for } b = \theta\, p^{\Delta}\,. \tag{2.5}$$

(See Figure 3 for an illustration of these graphons.) Observe that

$$\mathbb{E}[I_p(W_1)] = a^2 \log(1/p)\,, \qquad \mathbb{E}[I_p(W_2)] = (2b - b^2)\log(1/p)\,;$$

thus, the desired statements (2.2)–(2.3) would follow once we show that

$$t(H, W_1) \sim (1+\delta)p^{|E(H)|} \qquad \text{for any regular connected graph } H\,, \tag{2.6}$$

$$t(H, W_2) \sim (1+\delta)p^{|E(H)|} \qquad \text{for any graph } H\,. \tag{2.7}$$

For (2.6), we sum over the subset $S \subset V(H)$ of vertices that get mapped to the interval $[0, a]$ and get

$$t(H, W_1) = \sum_{S \subset V(H)} a^{|S|}(1-a)^{|V(H)|-|S|} p^{|E(H)|-|E(H_S)|}\,.$$

Recalling $a \asymp p^{\Delta/2}$, the exponent of p in each summand is $E(H) + |S|\Delta/2 - |E(H_S)|$, hence the non-negligible terms (of order $p^{|E(H)|}$) are precisely those with $2|E(H_S)| = |S|\Delta$. This occurs for $S = \varnothing$, contributing $(1+o(1))p^{|E(H)|}$), and as H is regular and connected, the only other possibility is $S = V(H)$, contributing $a^{|V(H)|} \sim \delta p^{|E(H)|}$. This establishes (2.6).

For (2.7), by summing over the subset $S \subset V(H)$ of vertices that get mapped to the interval $[0, b]$, we obtain

$$t(H, W_2) = \sum_{S \subset V(H)} b^{|S|}(1-b)^{|V(H)|-|S|} p^{|E(H_{\overline{S}})|}\,.$$

Since $b \asymp p^{\Delta}$, ignoring summands that are $o(p^{|E(H)|}$ leaves only those with $|S|\Delta + |E(H_{\overline{S}})| \leq |E(H)|$. This holds (with an equality) iff S is an independent set where every vertex has degree Δ; thus,

$$t(H, W_2) \sim P_{H^*}(\theta)p^{|E(H)|} = (1+\delta)p^{|E(H)|}\,. \qquad \square$$

3. Preliminaries

In this section, we recall various relevant estimates from [23], used there to solve the continuous variational problem (2.1) for the case of cliques. A key inequality used both in [23] and in its prequel dealing with dense graphs [22] is the following generalization of Hölder's inequality ([13, Theorem 2.1]).

Lemma 3.1. *Let $\mu_1, \mu_2, \dots \mu_n$ be probability measures on $\Omega_1, \dots \Omega_n$ resp., and let $\mu = \prod_{i=1}^n \mu_i$. Let $A_1 \dots A_m$ be nonempty subsets of $[n] = \{1, \dots n\}$ and for $A \subset [n]$ put $\mu_A = \prod_{j \in A} \mu_j$ and $\Omega_A = \prod_{j \in A} \Omega_j$. Let $f_i \in L^{p_i}(\Omega_{A_i}, \mu_{A_i})$ for each $i \in [m]$, and further suppose that $\sum_{A_i \ni j} (1/p_i) \le 1$ for all $j \in [n]$. Then*

$$\int \prod_{i=1}^m |f_i| \,\mathrm{d}\mu \le \prod_{i=1}^m \left(\int |f_i|^{p_i} \,\mathrm{d}\mu_{A_i} \right)^{1/p_i} . \tag{3.1}$$

In particular, when $p_i = \Delta$ for every $i \in [m]$ we have $\int \prod_{i=1}^m |f_i| \,\mathrm{d}\mu \le \prod_{i=1}^m \left(\int |f_i|^\Delta \,\mathrm{d}\mu_{A_i} \right)^{1/\Delta}$.

Let H be any graph with maximum degree Δ, and let W be a graphon with $t(H, W) \ge (1+\delta) p^{|E(H)|}$. Since I_p is convex and decreasing from 0 to p and increasing from p to 1, we may assume $W \ge p$, i.e.,

$$U := W - p \qquad \text{satisfies} \qquad 0 \le U \le 1 - p\,, \qquad t(H, p + U) \ge (1+\delta) p^{|E(H)|} . \tag{3.2}$$

For $b \in (0, 1]$, define the set $B_b(U)$ of points x with high *normalized degree* $d_U(x)$ by

$$B_b(U) := \{x : d_U(x) \ge b\}\,, \qquad \text{where} \qquad d_U(x) := \int_0^1 U(x, y) \mathrm{d}y\,. \tag{3.3}$$

Hereafter, the dependence on U will be dropped from $B_b(U)$ and $d_U(x)$, whenever the graphon U is clear from the context. Letting λ denote Lebesgue measure, by the convexity of $I_p(\cdot)$, for any $b \gg p$,

$$\mathbb{E}\left[I_p(p+U)\right] = \int_{[0,1]^2} I_p(p + U(x, y)) \mathrm{d}x \mathrm{d}y \ge \int_0^1 I_p(p + d_U(x)) \mathrm{d}x \ge \lambda(B_b) I_p(p + b)\,. \tag{3.4}$$

Moreover, by Proposition 2.4, we may assume that, for some implicit constant depending on δ and H,

$$\mathbb{E}\left[I_p(p+U)\right] \lesssim p^\Delta I_p(1)\,. \tag{3.5}$$

The following estimates for $I_p(x)$ were given in [23]; the $\sim$ notation below is w.r.t. to limits as $p \to 0$.

Lemma 3.2 ([23, Lemma 3.3])**.** *If $0 \le x \ll p$, then $I_p(p + x) \sim \frac{1}{2} x^2/p$, whereas when $p \ll x \le 1 - p$ we have $I_p(p + x) \sim x \log(x/p)$.*

Lemma 3.3 ([23, Lemma 3.4])**.** *There exists $p_0 > 0$ such that for every $0 < p \le p_0$,*

$$I_p(p + x) \ge (x/b)^2 I_p(p + b) \qquad \text{for any } 0 \le x \le b \le 1 - p - \log(1 - p)\,.$$

Corollary 3.4 ([23, Corollary 3.5])**.** *There is some $p_0 > 0$ such that for every $0 < p \le p_0$,*

$$I_p(p + x) \ge x^2 I_p(1 - 1/\log(1/p)) \sim x^2 I_p(1) \qquad \text{for any } 0 \le x \le 1 - p\,.$$

As a consequence of these lemmas, observe that

$$x^{3/2} \lesssim I_p(p + x)/I_p(1) + o(p^2) \qquad \text{for any } 0 \le x \le 1 - p\,; \tag{3.6}$$

indeed, this is trivial for $x \ll p^{4/3}$ due to the $o(p^2)$ term; if $p^{2/3} \le x \le 1 - p$ then $I_p(p + x) \gtrsim x I_p(1)$ by Lemma 3.2; and in between, when $p^{4/3} \lesssim x \le p^{2/3}$, we have $I_p(p + x) \gtrsim x^{3/2} p^{-2/3} I_p(p + p^{2/3})$ by Lemma 3.3 (using $b = p^{2/3}$), while $I_p(p + p^{2/3}) \gtrsim p^{2/3} I_p(1)$ again by Lemma 3.2.

Recalling (3.4), it follows from Lemma 3.2 (combined with (3.5)) that for any $p \ll b \le 1-p$,

$$\lambda(B_b) \le \frac{\mathbb{E}\left[I_p(p+U)\right]}{I_p(p+b)} \lesssim \frac{p^{\Delta} I_p(1)}{b\log(b/p)} \lesssim \frac{p^{\Delta}}{b}\,. \tag{3.7}$$

Furthermore, letting $\overline{B}_b = [0,1]\backslash B_b$ (here and in what follows), by the convexity of $I_p(x)$ and Lemma 3.3,

$$\mathbb{E}\left[I_p(p+U)\right] \ge \int_{\overline{B}_b} I_p(p+d(x))\mathrm{d}x \ge I_p(p+b)\int_{\overline{B}_b} \left(d(x)/b\right)^2 \mathrm{d}x\,.$$

Combining these, we get

$$\int_{\overline{B}_b} d^2(x)\mathrm{d}x \le \frac{b^2\mathbb{E}\left[I_p(p+U)\right]}{I_p(p+b)} \lesssim p^{\Delta} b\,. \tag{3.8}$$

For a bound on $\mathbb{E}[U]$, recall first from Lemma 3.3 that $I_p\big(p + ap^{(\Delta+1)/2}\sqrt{\log(1/p)}\big) \sim \frac12 ap^{\Delta} I_p(1)$ for any $p \le p_0$ and fixed $a \ge 0$ and $\Delta \ge 2$. However, $I_p(p+\mathbb{E}[U]) \le \mathbb{E}\left[I_p(p+U)\right] \lesssim p^{\Delta} I_p(1)$ by convexity and (3.5). Therefore, by the monotonicity of $I_p(p+x)$ for $x \ge 0$ we may conclude that

$$\mathbb{E}[U] \lesssim p^{\frac{\Delta+1}{2}}\sqrt{\log(1/p)}\,. \tag{3.9}$$

Finally, using (3.5) and Corollary 3.4 we obtain that, for an implicit constant depending on δ and H,

$$\mathbb{E}[U^2] \lesssim p^{\Delta}\,. \tag{3.10}$$

4. Solving the variational problem: from K_3 to C_4 to general graphs

4.1. **Short solution of the variational problem for triangles.** The case of K_3 (and larger cliques) was resolved in [23] via a divide-and-conquer approach: roughly put, by setting a certain *degree threshold* $b = b(p)$ one finds that, in any graphon whose entropy is of the correct *order*, the Lebesgue measure of the set B_b of high degree points (defined in (3.3)) asymptotically determines the surplus of $K_{1,2}$ copies (as in the anti-clique graphon), whereas the points in $\overline{B}_b$ are left only with the possibility of contributing extra triangles through "cliques." We include a (slightly shorter) version of this proof, towards showing why a more sophisticated cut-off $b(p)$ (tailored to each U) is needed for a general H.

Theorem 4.1 ([23, Theorem 2.2])**.** *Fix $\delta > 0$. As $p \to 0$, $\phi(K_3, p, \delta) \sim \min\left\{\frac12\delta^{2/3}, \delta/3\right\} p^2\log(1/p)$.*

Proof. Let $W = p + U$ with $W \ge p$ and $t(K_3, W) \ge (1+\delta)p^3$. Expanding $t(K_3, W)$ in terms of U,

$$t(K_3, W) - p^3 = t(K_3, U) + 3pt(K_{1,2}, U) + 3p^2\mathbb{E}[U] \ge \delta p^3\,. \tag{4.1}$$

Now, $\mathbb{E}[U] \ll p$ by (3.9), and so (4.1) reduces to

$$t(K_3, U) + 3pt(K_{1,2}, U) \ge (\delta - o(1))\, p^3\,.$$

Letting $B_b = \{x : d(x) > b\}$ as in (3.3) and recalling (3.7), for any $p \ll b \ll 1-p$ we have

$$\int_{B_b\times B_b} U^2(x,y)\mathrm{d}x\mathrm{d}y \le \lambda(B_b)^2 \lesssim p^4/b^2 \ll p^2\,. \tag{4.2}$$

Let the degree threshold be some function $b = b(p)$ such that $\sqrt{p\log(1/p)} \ll b \ll 1$, and note that

$$\int_{[0,1]^3} U(x,y)U(y,z)U(y,z)\mathbf{1}\{x \in B_b \text{ or } y \in B_b \text{ or } z \in B_b\}\mathrm{d}x\mathrm{d}y\mathrm{d}z \le 3\lambda(B_b)\mathbb{E}[U] \ll p^3\,, \tag{4.3}$$

where the last step uses (3.7) and (3.9). Letting δ_1 and δ_2 be defined by

$$\delta_1 := p^{-2}\int_{B_b\times\overline{B}_b} U^2(x,y)\mathrm{d}x\mathrm{d}y, \qquad \delta_2 := p^{-2}\int_{\overline{B}_b\times\overline{B}_b} U^2(x,y)\mathrm{d}x\mathrm{d}y\,, \tag{4.4}$$

we deduce from (4.3) and Hölder's inequality (3.1) that

$$t(K_3,U) = \int_{\overline{B}_b\times\overline{B}_b\times\overline{B}_b} U(x,y)U(y,z)U(x,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z + o(p^3) \le \big(\delta_2^{3/2}+o(1)\big)p^3\,. \tag{4.5}$$

Similarly, it follows from (3.8) and (4.2) that (not only for b as above but for any $p \ll b \ll 1$),

$$t(K_{1,2},U) = \int_{B_b\times\overline{B}_b\times\overline{B}_b} U(x,y)U(x,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z + o(p^2) \le \big(\delta_1+o(1)\big)p^2\,. \tag{4.6}$$

Combining (4.5)–(4.6), the expansion (4.1) reduces to $3\delta_1+\delta_2^{3/2}\ge \delta - o(1)$, and using Corollary 3.4,

$$\tfrac12\mathbb{E}\left[I_p(p+U)\right] \ge (1+o(1))\left(\delta_1+\tfrac12\delta_2\right)p^2\log(1/p)\,. \tag{4.7}$$

Therefore,

$$\liminf_{p\to 0}\left\{\frac{\frac12\mathbb{E}\left[I_p(p+U)\right]}{p^2 I_p(1)} : t(K_3,p+U)\ge(1+\delta)p^3\right\} \ge \min_{\substack{\delta_1,\delta_2\ge 0\\ 3\delta_1+\delta_2^{3/2}\ge\delta}}\left\{\delta_1+\tfrac12\delta_2\right\} = \min\left\{\tfrac12\delta^{3/2},\tfrac13\delta\right\}\,,$$

since the minimum is attained either at $\delta_1 = 0$ or $\delta_2 = 0$. This together with the upper bound in Proposition 2.4 (recall that $P_{K_3}(x) = 3x+1$) completes the proof for the case of triangles. ☐

4.2. **Solution of the variational problem for the 4-cycle.** The argument in §4.1 can be applied to other graphs H and rule out certain subgraphs $F\subset H$ from having a non-negligible contribution to $t(H,W)$ in the expansion analogous to (4.1). However, as we next see, already for C_4 new ideas are required to tackle all subgraphs $F\subset C_4$ and deduce the correct lower bound on $\phi(H,p,\delta)$.

Let $W=p+U$ be such that $W\ge p$ and $t(C_4,W)\ge(1+\delta)p^4$. Expanding $t(C_4,W)$ à la (4.1),

$$t(C_4,W)-p^4 = t(C_4,U)+4p^2t(K_{1,2},U)+4pt(P_4,U)\ge(\delta-o(1))\,p^4\,, \tag{4.8}$$

where P_4 is the path on 4 vertices and we used that $\mathbb{E}[U]\ll p$ from (3.9). Let B_b be as in (3.3).

Observe that, by Hölder's inequality (3.1), embeddings $P_4\mapsto(w,x,y,z)\in[0,1]^4$ with $x\in\overline{B}_b$ satisfy

$$\begin{aligned}\int_{[0,1]\times\overline{B}_b\times[0,1]^2} U(w,x)U(x,y)U(y,z)\mathrm{d}w\mathrm{d}x\mathrm{d}y\mathrm{d}z &\le \int_{\overline{B}_b\times[0,1]^2} d(x)U(x,y)U(y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z\\ &\le \left(\int_{\overline{B}_b} d^2(x)\mathrm{d}x\right)^{1/2}\left(\int_0^1 U^2(x,y)\mathrm{d}x\mathrm{d}y\right) \lesssim p^3\sqrt{b}\ll p^3\end{aligned} \tag{4.9}$$

by (3.10) and (3.8), using $b\ll 1$ for the last inequality. Hence, embeddings of P_4 with a non-negligible contribution to $t(P_4,U)$ must place both of the interior (degree-2) vertices in B_b. The contribution from such embeddings is therefore at most $\lambda(B_b)^2\lesssim p^4/b^2\ll p^3$, provided $b\gg\sqrt{p}$; hence, $t(P_4,U)\ll p^3$.

We have already encountered the term $t(K_{1,2},U)$ in §4.1, and it should correspond to copies of C_4 arising from anti-cliques, as per (4.6). We therefore move our attention to the different embeddings of the 4-cycle $t(C_4,U)$, to which end we use $\tilde U(w,x,y,z) = U(w,x)U(x,y)U(y,z)U(z,w)$ for brevity. Clearly, any embedding placing two consecutive vertices of C_4 in B_b is negligible since, by (3.7) and (3.9),

$$\int_{B_b\times B_b\times[0,1]^2}\tilde U(w,x,y,z)\mathrm{d}w\mathrm{d}x\mathrm{d}y\mathrm{d}z \le \lambda(B_b)^2\mathbb{E}[U]\ll p^4\,. \tag{4.10}$$

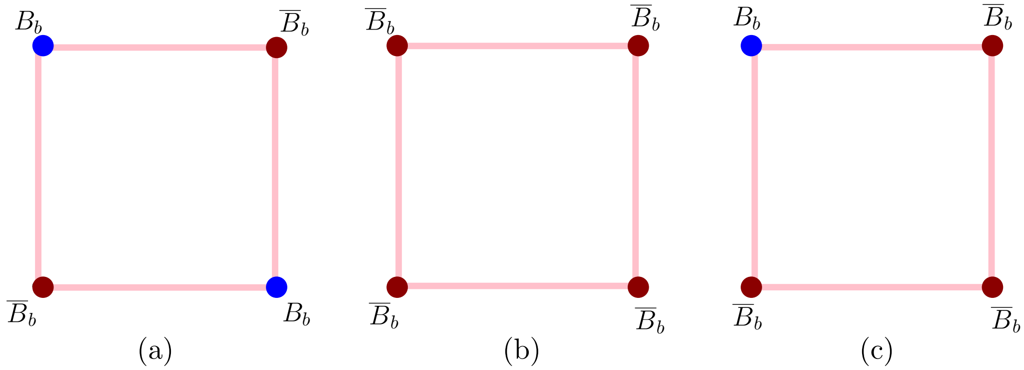


FIGURE 4. Different embeddings of the 4-cycle: (a) and (b) are non-negligible.

Recalling δ_1, δ_2 from (4.4), the three other possible embeddings of C_4 (see Fig. 4 for an illustration) are handled via Hölder's inequality as follows.

(a) two nonadjacent vertices in B_b (2 configurations):

$$\int_{B_b\times\overline{B}_b\times B_b\times\overline{B}_b} \tilde{U}(w,x,y,z)\mathrm{d}w\mathrm{d}x\mathrm{d}y\mathrm{d}z \le \left(\int_{B_b\times\overline{B}_b} U^2(x,y)\mathrm{d}x\mathrm{d}y\right)^2 \le \left(\delta_1^2+o(1)\right)p^4\,. \tag{4.11}$$

(b) no vertices in B_b (1 configuration):

$$\int_{\overline{B}_b^4} \tilde{U}(w,x,y,z)\mathrm{d}w\mathrm{d}x\mathrm{d}y\mathrm{d}z \le \left(\int_{\overline{B}_b\times\overline{B}_b} U^2(x,y)\mathrm{d}x\mathrm{d}y\right)^2 \le \left(\delta_2^2+o(1)\right)p^4\,. \tag{4.12}$$

(c) a single vertex in B_b (4 configurations):

$$\int_{B_b\times\overline{B}_b^3} \tilde{U}(w,x,y,z)\mathrm{d}w\mathrm{d}x\mathrm{d}y\mathrm{d}z \le \left(\delta_1\delta_2+o(1)\right)p^4\,. \tag{4.13}$$

Combining (4.10), (4.11)–(4.13) and the estimate (4.6) for $t(K_{1,2},U)$, the expansion (4.8) gives

$$2\delta_1^2+\delta_2^2+4\delta_1\delta_2+4\delta_1 \ge \delta - o(1), \tag{4.14}$$

valid for any $\sqrt{p} \ll b \ll 1$. As in the case of K_3, we wish to minimize $(\delta_1+\frac{1}{2}\delta_2)$ subject to this constraint. Unfortunately, the minimum of $(\delta_1+\frac{1}{2}\delta_2)$ subject to (4.14) is not attained at $\delta_1=0$ or $\delta_2=0$, thus the lower bound obtained in this way does not match the upper bound from Proposition 2.4.

Recall the clique and anti-clique graphons (2.4)–(2.5). The main contribution of the anti-clique to $t(C_4,U)$ is through embeddings of type (a), whereas for the clique it is through embeddings of type (b). For the correct lower bound, we must show that the contribution from embeddings of type (c) is negligible; however, this can no longer be achieved using any arbitrary $\sqrt{p}\ll b\ll 1$. To conclude the proof, we select the degree threshold b *adaptively* based on the graphon U.

Lemma 4.2 (Adaptive degree threshold for C_4)**.** *Fix $\varepsilon>0$. There exists b with $\sqrt{p}\ll b\ll 1$ such that*

$$\int_{B_b\times\overline{B}_b^3} \tilde{U}(w,x,y,z)\mathrm{d}w\mathrm{d}x\mathrm{d}y\mathrm{d}z \le \varepsilon p^4\,. \tag{4.15}$$

Proof. Fix $C>0$ such that $\mathbb{E}[U^2]^2 \le Cp^4$, as guaranteed by (3.10), and set $M=\lceil 2C/\varepsilon\rceil$. Further let $\sqrt{p}\ll b_1\ll b_2\ll\cdots\ll b_M\ll 1$. As usual, set $B_{b_i}=\{x: d(x)>b_i\}$, observing that $B_{b_M}\subset\ldots\subset B_{b_1}$.

For every $2 \le i \le M$, using (3.7) and (3.8) we find that

$$\int_{B_{b_i}\times \overline{B}_{b_i}\times \overline{B}_{b_{i-1}}\times \overline{B}_{b_i}} \tilde{U}(w,x,y,z)\mathrm{d}w\mathrm{d}x\mathrm{d}y\mathrm{d}z \le \lambda(B_{b_i})\int_{\overline{B}_{b_{i-1}}} d^2(x)\mathrm{d}x \lesssim \frac{b_{i-1}}{b_i}p^4 \ll p^4\,. \tag{4.16}$$

Thus, any one of the degree thresholds b_i will satisfy the required statement (4.15) if, say,

$$\int_{B_{b_i}\times \overline{B}_{b_i}\times (B_{b_{i-1}}\setminus B_{b_i})\times \overline{B}_{b_i}} \tilde{U}(w,x,y,z)\mathrm{d}w\mathrm{d}x\mathrm{d}y\mathrm{d}z \le \tfrac{1}{2}\varepsilon p^4\,. \tag{4.17}$$

The proof is concluded by noticing that there is necessarily some $2 \le i \le M$ satisfying (4.17), as otherwise—since the sets $\{B_{b_i}\times \overline{B}_{b_i}\times (B_{b_{i-1}}\setminus B_{b_i})\times \overline{B}_{b_i} : 2\le i\le M\}$ are mutually disjoint—we would get $t(C_4,U) > (M\varepsilon/2)p^4 \ge Cp^4$, a contradiction to our choice of C since $t(C_4,U)\le \mathbb{E}[U^2]^2$. ☐

Remark 4.3. Note the advantage of using multiple thresholds with $b_{i-1} \ll b_i$ in the proof of Lemma 4.2: a single threshold function $b_i \equiv b$ (as in §4.1) would have given a bound of $O(p^4)$ on the left-hand of (4.16) vs. the sought $o(p^4)$. This idea will be crucial in our arguments for general graphs.

Using b from Lemma 4.2 for an arbitrarily small ε, the constraint in (4.14) becomes $\delta_2^2+2\delta_1^2+4\delta_1 \ge \delta$, and, as in (4.7), we conclude by Corollary 3.4 that

$$\liminf_{p\to 0}\left\{\frac{\frac12\mathbb{E}[I_p(p+U)]}{p^2 I_p(1)} : t(C_4,p+U)\ge (1+\delta)p^4\right\} \ge \min_{\delta_1,\delta_2\ge 0}\left\{\delta_1+\frac{\delta_2}{2} : \delta_2^2+2\delta_1^2+4\delta_1\ge\delta\right\}.$$

This minimum is clearly attained at either $\delta_1 = 0$ or $\delta_2 = 0$ by the next observation, which concludes the bound for $\phi(C_4,p,\delta)$ matching the one from Proposition 2.4 (recall that $P_{C_4}(x) = 1+4x+2x^2$).

Observation 4.4. *Let f, g be convex nondecreasing functions on $[0,\infty)$ and let $a > 0$. The minimum of $x + y$ over the region $\{x, y \ge 0 : f(x) + g(y) \ge a\}$ is attained at either $x = 0$ or $y = 0$.*

Proof. By convexity, if $\gamma = \frac{y}{x+y}$ then $f(x) \le \gamma f(0)+(1-\gamma)f(x+y)$ and $g(y) \le (1-\gamma)g(0)+\gamma g(x+y)$, so $f(x)+g(y) \le \gamma[f(0)+g(x+y)]+(1-\gamma)[f(x+y)+g(0)]$, i.e., at most $\max\{f(0)+g(x+y),\ f(x+y)+g(0)\}$. Thus, the sought minimum is either $\min\{z : f(0) + g(z) \ge a\}$ or $\min\{z : f(z) + g(0) \ge a\}$. ☐

4.3. **General graphs.** The solution of the variational problem for a general graph H is substantially more delicate and, perhaps surprisingly, involves matchings and the polynomial $P_{H^*}(x)$.

By (3.2) assume $W \ge p$ and let $W = U + p$. Expanding $t(H,W)$ in terms of U (as in (4.1) and (4.8)),

$$t(H,W) - p^{|E(H)|} = \sum_{\substack{F\subset H\\ E(F)\neq\varnothing}} N(F,H)t(F,U)p^{|E(H)|-|E(F)|}\,, \tag{4.18}$$

where $N(F,H)$ is the number of copies of F in the labeled graph H up to automorphism, i.e., $N(F,H) := \hom(F,H)/\operatorname{Aut}(F)$ with $\operatorname{Aut}(F)$ the size of the automorphism group of the graph F (the subgraph $F\subset H$ consists of the edges $(x,y)\in E(H)$ involving $U(x,y)$ in the expansion of $t(H,p+U)$).

4.3.1. *Main contributing subgraphs.* As a first step towards identifying the subgraphs F that have a non-negligible contribution to the sum in (4.18), we have the following simple observation.

Observation 4.5. *Let $\Delta \ge 2$, and H be as in Theorem 2.3 and U a graphon satisfying (3.10). Then for any subgraph F of H with maximum degree $\Delta - 1$ we have $t(F,U) \ll p^{|E(F)|}$.*

Proof. If $\Delta = 2$, then F is a disjoint union of isolated edges, and by (3.9)

$$t(F,U) = \mathbb{E}[U]^{|E(F)|} \lesssim p^{3|E(F)|/2}(\log(1/p))^{|E(F)|/2} \ll p^{|E(F)|}\,.$$

On the other hand, if $\Delta \geq 3$, then using Hölder's inequality (3.1) we get

$$t(F,U) \leq \mathbb{E}[U^{\Delta-1}]^{\frac{|E(F)|}{\Delta-1}} \leq \mathbb{E}[U^2]^{\frac{|E(F)|}{\Delta-1}} \leq p^{\frac{\Delta|E(F)|}{\Delta-1}} \ll p^{|E(F)|}\,. \qquad \square$$

Having seen that $F \subset H$ with $\Delta(F) < \Delta(H)$ give a negligible contribution to (4.18), we next show that subgraphs F with a *vertex cover number* $\tau(F) > |E(F)|/\Delta(F)$ are also ruled out in this context.

Recall that a vertex cover of F is a subset $S \subset V(F)$ such that $V(F)\backslash S$ is an independent set, and $\tau(F)$ is the size of the minimum vertex cover of F (thus by definition $\tau(F) \geq |E(F)|/\Delta(F)$ for any F). Further recall that the *matching number* $\nu(F)$ is the size of a maximum matching in the graph F; the *fractional matching number*, denoted $\nu^*(F)$, is the maximum value of $w(E(F)) := \sum_{e\in E(F)} w(e)$ over all weight functions $w : E(F) \to [0,1]$ such that $\sum_{v\ni e} w(e) \leq 1$ for every vertex $v \in F$ (i.e., ν^* is the real-valued relaxation of the integer program ν, as it allows $0 \leq w(e) \leq 1$ rather than $w(e) \in \{0,1\}$).

Lemma 4.6. *If F is a connected irregular graph with maximum degree $\Delta \geq 2$ and $\tau(F) > |E(F)|/\Delta$,*

$$\nu^*(F) > |E(F)|/\Delta\,. \tag{4.19}$$

Proof. If F is bipartite, by König's Minimax Theorem $\nu(F) = \tau(F)$, and as $\nu^*(F)$ is always sandwiched between these two, we have $\nu^*(F) = \tau(F) > |E(F)|/\Delta$ by our hypothesis, as required.

Otherwise, let C be an odd cycle in F. Let u_0 be some vertex of degree $\deg(u_0) < \Delta$ (using that F is irregular) and let $P = (u_0, u_1, \ldots, u_r)$ ($r \geq 0$) be a shortest path in F from u_0 to C (using that F is connected); finally, write the vertices of C as $v_0, \ldots, v_{2k}$ ($k \geq 1$) with $v_0 = u_r$. We now perturb $w \equiv 1/\Delta$ into w' by adding $(-1)^j\varepsilon$ to (u_j, u_{j+1}) along the path, and adding $(-1)^{j+r}\varepsilon/2$ to (v_j, v_{j+1}) along the cycle ($v_{2k+1} \equiv v_0$). Since C is odd, $\sum_{v\ni e} w'(e) = \sum_{v\ni e} w(e)$ for all $v \neq u_0$; thus, w' is legal as long as $\varepsilon \leq \frac{1}{\Delta-1} - \frac{1}{\Delta}$. Finally, if r is even then $w'(P) = w(P)$ and $w'(C) = w(C) + \varepsilon/2$, and if r is odd then $w'(P) = w(P) + \varepsilon$ and $w'(C) = w(C) - \varepsilon/2$. Either way, $\nu^*(F) \geq w'(E(F)) = E(F)/\Delta + \varepsilon/2$. $\square$

Corollary 4.7. *Let F be a fixed connected irregular graph with maximum degree $\Delta \geq 2$ such that $\tau(F) > |E(F)|/\Delta$. There exists some $\kappa > 0$ such that, as $p \to 0$,*

$$t(F,U) \lesssim p^{(1+\kappa)|E(F)|} \ll p^{|E(F)|}\,. \tag{4.20}$$

Proof. By (4.19), the function $w \equiv 1/\Delta$ is not a local maximum of the linear program defining $\nu^*(F)$.

If $\Delta = 2$, one can perturb $w \equiv 1/2$ into $w' : E(F) \to [0, 2/3]$, such that $\sum_{v\ni e} w'(e) \leq 1$ and $\sum_{e\in E(F)} w'(e) > (1+\kappa)|E(F)|/2$ for some $\kappa > 0$. Using Hölder's inequality (3.1) with these weights,

$$t(F,U) \leq \prod_{e\in E(F)} \Big(\int_{[0,1]^2} U(x,y)^{1/w'(e)} \mathrm{d}x\mathrm{d}y\Big)^{w'(e)} \leq \Big(\int_{[0,1]^2} U(x,y)^{3/2} \mathrm{d}x\mathrm{d}y\Big)^{(1+\kappa)\frac{|E(F)|}{2}} \lesssim p^{(1+\kappa)|E(F)|}\,,$$

where the last inequality used (3.6) followed by (3.5) to replace the integral by $O(p^\Delta) = O(p^2)$.

If $\Delta \geq 3$, one can perturb $w \equiv 1/\Delta$ into $w' : E(F) \to [0, 1/2]$, such that $\sum_{v\ni e} w'(e) \leq 1$, and $\sum_{e\in E(F)} w'(e) > (1+\kappa)|E(F)|/\Delta$ for some $\kappa > 0$. Using Hölder's inequality (3.1) with these weights,

$$t(F,U) \leq \prod_{e\in E(F)} \Big(\int_{[0,1]^2} U(x,y)^{1/w'(e)} \mathrm{d}x\mathrm{d}y\Big)^{w'(e)} \leq \Big(\int_{[0,1]^2} U(x,y)^2 \mathrm{d}x\mathrm{d}y\Big)^{(1+\kappa)\frac{|E(F)|}{\Delta}} \lesssim p^{(1+\kappa)|E(F)|}\,,$$

using (3.10) for the last transition. $\square$

Observation 4.5 and Corollary 4.7 reduce (4.18) into the following key identity.

$$t(H,W) - p^{|E(H)|} = \sum_{F\in\mathscr{F}_H} N(F,H)t(F,U)p^{|E(H)|-|E(F)|} + o(p^{|E(H)|})\,, \tag{4.21}$$

where

$$\mathscr{F}_H := \{F \subset H \;:\; \Delta(F) = \Delta(H) \text{ and } \tau(F) = |E(F)|/\Delta(F)\}\ . \tag{4.22}$$

4.3.2. *The role of the independence polynomial.* In view of the above defined $\mathscr{F}_H$, observe the following.

Observation 4.8. *A graph F has $\tau(F) = |E(F)|/\Delta(F)$ iff it is bipartite and $\tau(F)$ is attained by an independent set A (one of the two color classes) such that $\deg_F(v) = \Delta$ for all $v \in A$.*

Consequently, in a connected graph H with maximum degree Δ, every $F \neq H$ in $\mathscr{F}_H$ contains a unique independent set A such that $\deg_H(v) = \Delta$ for all $v \in A$ (its size is thus $|E(F)|/\Delta$, and its uniqueness follows from the fact that H is connected); similarly, H itself admits 2 such sets if $H \in \mathscr{F}_H$ (i.e., if it is bipartite and regular), and none otherwise. Recalling Definition 1.1, it follows that

$$P_{H^*}(x) = 1 + \sum_{F\in\mathscr{F}_H} N(F,H)x^{|E(F)|/\Delta} + \mathbf{1}\{H\in\mathscr{F}_H\}x^{|E(H)|/\Delta}\,. \tag{4.23}$$

In §5–§6 we will relate each term $t(F,U)p^{|E(H)|-|E(F)|}$ in the right-hand of (4.21) to $\theta^{|E(F)|/\Delta}p^{|E(H)|}$ where θ is the analog of δ_1 from (4.4), thereby leading to $P_{H^*}(\theta)$ (just as $P_{C_4}(\delta_1)$ appeared in §4.2).

4.3.3. *Outline of the proof of Theorems 2.2 and 2.3.* A 2-*matching* M of a graph F is a union of 2 edge disjoint matchings of F (thus it is a disjoint union of cycles and paths).[4] Informally, the proof for a general graph H is obtained as follows. For every vertex v in the minimum vertex cover of F, we choose a 2-matching containing v.

- When F is irregular, the 2-matching is chosen such that the component containing v is a path (crucially, the existence of such a matching is guaranteed by Lemma 5.3). One can then show that the only embeddings of F with a non-negligible contribution are those where alternating degree-2 vertices of the path are all in B_b (see Definition 3.3) for a particular choice of b ; see §5 for details.
- When F is regular, the argument is broken into two cases depending on whether the graph is bipartite or not. If it is not bipartite, one essentially shows that the only embedding of F with a non-negligible contribution is when all the vertices are in $\overline{B}_b$ (i.e., they have low degree). In case F is bipartite with bipartition (F_1, F_2), the proof proceeds by showing that for an embedding to be non-negligible, the vertices in F_1 (and similarly F_2) must all be in either B_b or in $\overline{B}_b$ (e.g., Fig. 4). This is shown by a generalization of the argument for the 4-cycle applied on 2-matchings whose connected components are even cycles. To make these arguments work one needs to choose a degree threshold b adaptively (as in §4.2, and unlike the irregular case), via a nested subsequence of degree thresholds for the different 2-matchings involving the vertices of F; see §6 for details.

As in the case of the 4-cycle, these arguments imply a lower bound on $\phi(H,p,\delta)$ through an application of Corollary 3.4, which matches the upper bound from Proposition 2.4 and concludes the proof.

5. Irregular connected graphs: proof of Theorem 2.3

This section is devoted to the lower bound in Theorem 2.3 (the upper bound is by Proposition 2.4). The following combinatorial lemmas on 2-matchings will play a central role in our analysis further on.

[4]This differs slightly from the notion of 2-matchings in the literature (cf. [20]), the definition here being a special case.

5.1. **2-matchings.** The edges of a bipartite graph G of maximum degree Δ can be partitioned into Δ pairwise disjoint matchings by König's Edge Coloring Theorem (cf. [20, §1.4] for its proof via embedding G in a regular graph); each one saturates every vertex of degree Δ, thus Berge's criterion for maximum matchings via augmenting paths implies the following (see [20, §1] for more on these classical results).

Fact 5.1. *In any bipartite graph with maximum degree Δ, there exists a maximum matching which saturates all the vertices of degree Δ.*

In what follows, let $K_1 \cup K_2 = (V(K_1) \cup V(K_2), E(K_1) \cup E(K_2))$ denote the graph union of K_1, K_2. In addition, recalling the definition of $\mathscr{F}_H$ in (4.22), upon referring to a bipartition of $F \in \mathscr{F}_H$ by (F_A, F_B) we will let F_A be an independent set with $\deg_F(v) = \Delta$ for all $v \in F_A$.

Observation 5.2. *If $F \in \mathscr{F}_H$ is a bipartite graph with bipartition (F_A, F_B), then F has a 2-matching M of size $2|E(F)|/\Delta$. Moreover if $v \in F_B$ has $\deg_F(v) < \Delta$ then M can be chosen with $\deg_M(v) = 1$.*

Proof. By König's Edge Coloring Theorem, $E(F)$ can be partitioned into matchings $\{M_1, \dots, M_\Delta\}$; as F_A is an independent set with $\deg_F(v) = \Delta$ for all $v \in F_A$, we must have $M_i = |E(F)|/\Delta$ for all i, so $M_i \cup M_j$ (for any $i \neq j$) gives a 2-matching of size $2|E(F)|/\Delta$. Finally, if $v \in F_B$ has $\deg_F(v) < \Delta$ then there must exist i, j such that $v \in M_i$ and $v \notin M_j$, as needed. □

The next key lemma removes the assumption $\deg_F(v) < \Delta$ from Observation 5.2 for an irregular F.

Lemma 5.3. *Let $F \in \mathscr{F}_H$ be an irregular graph with bipartition (F_A, F_B) and $\Delta \geq 3$. Then for every $v \in F_B$, the graph F has a 2-matching M of size $2|E(F)|/\Delta$ such that $\deg_M(v) = 1$.*

Proof. We may assume F is connected (by restricting attention to the connected component of v—which is irregular as $F \in \mathscr{F}_H$ and H is connected so it cannot have a nontrivial Δ-regular component—and applying Observation 5.2 to other components) as well as that $\deg_F(v) = \Delta$ (again by Observation 5.2).

Consider the graph F' obtained from F by deleting v (and the edges incident to it). Every $S \subset F_A$ has at least $|S|$ neighbors in F (as the degrees in F_A are all Δ), so $|N_{F'}(S)| \geq |S|$ unless $|N_F(S)| = |S|$, which in turn is impossible if F is connected and irregular. By Hall's Theorem, $\nu(F') = |F_A| = |E(F)|/\Delta$, and by Fact 5.1, there is a maximum matching M_1 in F' saturating all every degree-Δ vertex in F'.

Let F'' be the graph obtained from F by deleting $E(M_1)$ and the edge (v, w), where w is an arbitrarily chosen neighbor of v in F (recall that $\deg_{M_1}(v) = 0$). Clearly, the maximum degree in F'' is $\Delta - 1$ (which is the degree in F'' of v and every vertex in F_A, except for w that has degree $\Delta - 2$), and so

$$\tau(F'') \geq \frac{|E(F'')}{\Delta - 1} = \frac{(1 - \frac{1}{\Delta})|E(F)| - 1}{\Delta - 1} = |F_A| - \frac{1}{\Delta - 1}\,.$$

As $1/(\Delta - 1) \leq 1/2$ (by the hypothesis $\Delta \geq 3$), $\tau(F'') \geq |F_A|$, but F_A is a vertex cover (already in F), so König's Minimax Theorem together with Fact 5.1 imply that there is a maximum matching M_2 in F'' of size $|F_A| = |E(F)|/\Delta$ that saturates all the degree $\Delta - 1$ vertices, in particular including v.

The desired 2-matching M is now obtained as the union of $M_1 \cup M_2$. □

Corollary 5.4. *Let $F \in \mathscr{F}_H$ be an irregular graph with bipartition (F_A, F_B). For any $u \in F_A$ there is a 2-matching M of size $2|E(F)|/\Delta$ in F so that the connected component of u in M is a path.*

Proof. If $\Delta = 2$ then there are no cycles in F, so the result follows from Observation 5.2; thus, assume $\Delta \geq 3$, let $v \in N_F(u)$, and let M be the 2-matching from Lemma 5.3 where $\deg_M(v) = 1$, i.e., v is the endpoint of a path P. If u lies on a path in M we are done; otherwise, it is on some cycle C and (u, v) is not part of M, so we may use it to merge P and C, unravelling C in the obvious way (see Fig. 5). □

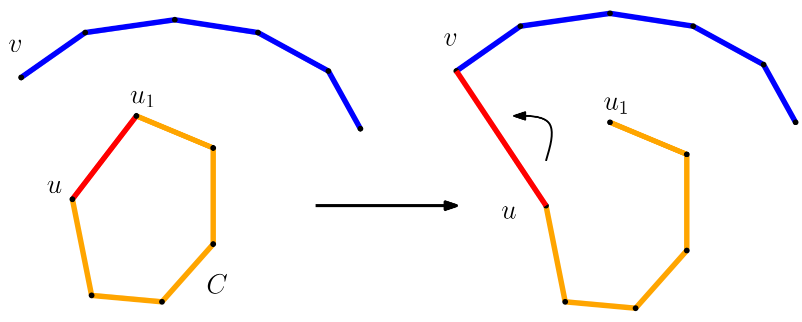


FIGURE 5. Merging a cycle and a path to get a longer path.

The 2-matching obtained in the above lemma will be used throughout the paper to determine contributing embeddings of $t(F, U)$. To this end we have the following definition:

Definition 5.5. *Let $M_F(u)$, for an irregular graph $F = (F_A, F_B) \in \mathscr{F}_H$ and $u \in F_A$, be a 2-matching of size $2|E(F)|/\Delta$ such that the connected component of u in $M_F(u)$ is a path (as per Corollary 5.4).*

5.2. **Proof of Theorem 2.3.** Using the 2-matching $M_F(u)$ from Definition 5.5 we can determine the embeddings of F which contribute to $t(F, U)$, for $F \in \mathscr{F}_H$. Hereafter, let $\boldsymbol{x} = (x_1, x_2, \cdots, x_{|V(H)|})$ denote the vector of variables indexed by the vertices of F. For brevity, define the notation

$$W(\boldsymbol{x}|F) := \prod_{(i,j)\in E(F)} W(x_i, x_j) \qquad \text{and} \qquad \mathrm{d}\boldsymbol{x}(F) := \prod_{i\in V(F)} \mathrm{d}x_i \,.$$

Recall the definition of B_b from (3.3). The following proposition shows that embeddings of $F \in \mathscr{F}_H$ which contribute to $t(F, U)$ must have all the vertices of F_A in B_b, for some appropriately chosen b.

Proposition 5.6. *Suppose $F \in \mathscr{F}_H$ with bipartition (F_A, F_B). For any $\varepsilon > 0$, $\gamma_0 \le 1/3$, and $b = p^{\gamma_0}$,*

$$\int_{[0,1]^{|V(F)|}} U(\boldsymbol{x}|F)\mathbf{1}\{\exists v \in F_A : x_v \in \overline{B}_b\}\mathrm{d}\boldsymbol{x}(F) \lesssim \varepsilon p^{|E(F)|} \tag{5.1}$$

holds for any graphon U such that $\mathbb{E}[U^2] \lesssim p^\Delta$ (as per (3.10)).

The proof of the above proposition is postponed to §5.3. We first show how to complete the proof of Theorem 2.3 using this proposition.

Proof of Theorem 2.3 assuming Proposition 5.6. The next lemma shows that the contribution to $t(F, U)$ from embeddings where $x_i, x_j \in B_b$ for $(i, j) \in E(F)$ is negligible.

Lemma 5.7. *Let $F \in \mathscr{F}_H$ and U be such that $\mathbb{E}[U^2] \lesssim p^\Delta$ (as per (3.10)). For any $p^{1/3} \ll b \ll 1$,*

$$\int_{[0,1]^{|V(F)|}} U(\boldsymbol{x}|F)\mathbf{1}\{\exists\, (i,j) \in E(F) \text{ such that } x_i, x_j \in B_b\}\mathrm{d}\boldsymbol{x}(F) \ll p^{|E(F)|} \,. \tag{5.2}$$

Proof. It suffices to show (5.2) for a fixed edge $(i, j) \in E(F)$. As argued in the proof of Observation 5.2, F is a disjoint union of Δ matchings of size $|E(F)|/\Delta$, thus there exists a 2-matching M of F of size $2|E(F)|/\Delta$ such that $(i, j) \in E(M)$. Let $M_{\setminus(i,j)}$ be the graph obtained by removing all the edges incident to the vertices i, j. Clearly, the number of edges removed is at least 2 and at most 3, since we cannot have $\deg_M(i) = \deg_M(j) = 1$ (recall that $\deg_H(i) = \Delta$ or $\deg_H(j) = \Delta$). Hence,

$$\int_{[0,1]^{|V(M)|}} U(\boldsymbol{x}|M)\mathbf{1}\{x_i, x_j \in B_b\}\mathrm{d}\boldsymbol{x}(M) \leq \lambda(B_b)^2 \int_{[0,1]^{|V(M)|-2}} U(\boldsymbol{x}|M_{\setminus(i,j)})\mathrm{d}\boldsymbol{x}(M_{\setminus(i,j)})$$
$$\lesssim \left(\frac{p^\Delta}{b}\right)^2 p^{\frac{\Delta(|E(M)|-3)}{2}} \lesssim b^{-2}p^{|E(F)|+\frac{\Delta}{2}} \ll p^{|E(F)|}\,.$$

The inequality between the lines is by (3.7),(3.10) and Lemma 3.1, while the last equality uses the assumption $b \gg p^{1/3} \gg p^{\Delta/4}$, since $\Delta \geq 2$. Finally, since $U(\boldsymbol{x}|F) \leq U(\boldsymbol{x}|M)$, the result follows. ☐

Let $\gamma_0 = 1/4$ and $b_0 = p^{\gamma_0} = p^{1/4}$. By Proposition 5.6, for all $F \in \mathscr{F}_H$,

$$\int_{[0,1]^{|V(F)|}} U(\boldsymbol{x}|F)\mathbf{1}\{\exists x_v \in \overline{B}_{b_0}\}\mathrm{d}\boldsymbol{x}(F) \leq \frac{\varepsilon}{\gamma(H)}p^{|E(F)|}\,,$$

where $\gamma(H) = \sum_{F\in\mathscr{F}_H} N(F,H)$ (absorbing the implicit constant from (5.1) in ε). This implies that

$$\sum_{F\in\mathscr{F}_H} N(F,H)p^{|E(H)|-|E(F)|}\int_{[0,1]^{|V(F)|}} U(\boldsymbol{x}|F)\mathbf{1}\{\exists x_v \in \overline{B}_{b_0}\}\mathrm{d}\boldsymbol{x}(F) \leq \varepsilon p^{|E(H)|}\,. \tag{5.3}$$

Now, let

$$\theta_{b_0}p^\Delta := \int_{B_{b_0}\times\overline{B}_{b_0}} U^2(x,y)\mathrm{d}x\mathrm{d}y\,, \qquad \theta'_{b_0}p^\Delta := \int_{\overline{B}_{b_0}\times\overline{B}_{b_0}} U^2(x,y)\mathrm{d}x\mathrm{d}y\,. \tag{5.4}$$

Set $\theta = \theta_{b_0}$ and $\theta' = \theta'_{b_0}$. Let δ_F be such that $t(F,U) = \delta_F p^{|E(F)|}$, for $F \in \mathscr{F}_H$. Combining (5.2) and (5.3) it follows that for $F \in \mathscr{F}_H$, with bipartition (F_A, F_B), the contribution to $t(F,U)$ comes from embeddings where every vertex in F_A is in B_{b_0} and every vertex in F_B is in $\overline{B}_{b_0}$. That is,

$$p^{|E(F)|}\left(\delta_F - \frac{\varepsilon}{\gamma(H)}\right) - o(p^{|E(H)|}) \leq \int_{[0,1]^{|V(F)|}} U(\boldsymbol{x}|F)\mathbf{1}\{x_v \in B_{b_0}, \forall v \in F_A, x_v \notin B_{b_0}, \forall v \in F_B\}\mathrm{d}\boldsymbol{x}(F)$$
$$\leq p^{|E(F)|}\theta^{\frac{|E(F)|}{\Delta}}\,, \tag{5.5}$$

where the last step uses Hölder's inequality (3.1). Therefore, from (4.21) and the definition of $\gamma(H)$,

$$(\delta - \varepsilon)p^{|E(H)|} - o(p^{|E(H)|}) = p^{|E(H)|}\sum_{F\in\mathscr{F}_H} N(F,H)\left(\delta_F - \frac{\varepsilon}{\gamma(H)}\right) - o(p^{|E(H)|})$$
$$\leq p^{|E(H)|}\sum_{F\in\mathscr{F}_H} N(F,H)\theta^{\frac{|E(H)|}{\Delta}} = p^{|E(H)|}\left(P_{H^*}(\theta) - 1\right)\,, \tag{5.6}$$

where the inequality between the lines used (5.5), and the last equality is by (4.23). It thus follows that $P_{H^*}(\theta) \geq 1 + \delta - \varepsilon - o(1) \geq 1 + \delta - 2\varepsilon$, for p small enough. Moreover, by Corollary 3.4,

$$\tfrac{1}{2}\mathbb{E}\left[I_p(p+U)\right] \geq \big(\theta + \theta'/2 - o(1)\big)\, p^\Delta \log(1/p)\,.$$

By the above discussion, since ε was arbitrary,

$$\liminf_{p\to0}\left\{\frac{\frac{1}{2}\mathbb{E}\left[I_p(p+U)\right]}{p^\Delta I_p(1)} : t(H, p+U) \geq (1+\delta)p^{|E(H)|}\right\} \geq \min_{\substack{\theta\in S(H),\\ \theta'\geq0}} \theta + \theta'/2\,,$$

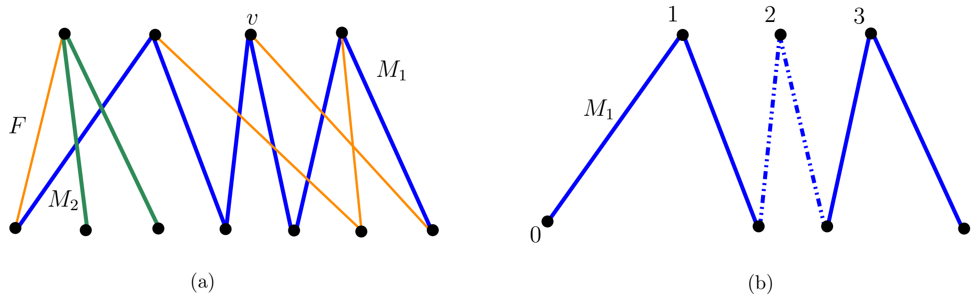


FIGURE 6. (a) Thick edges (colored blue and green) form a 2-matching of size $2|E(F)|/\Delta = 8$ for the graph F. The path M_1 is the connected component of v. (b) The alternating degree-2 vertices of the path M_1 are labeled $1, 2, \ldots,$ starting with the vertex 1 which is the neighbor of one of the endpoints (labeled 0).

where $S(H) := \{\theta \geq 0 : P_{H^*}(\theta) \geq 1 + \delta\}$. Since the constraint in the definition of $S(H)$ does not involve θ', the minimum in the last expression is attained when $\theta' = 0$. Thus, since ε is arbitrary,

$$\liminf_{p \to 0} \left\{ \frac{\frac{1}{2}\mathbb{E}\left[I_p(p+U)\right]}{p^{\Delta} I_p(1)} : t(H, p+U) \geq (1+\delta)p^{|E(H)|} \right\} \geq \theta$$

for $\theta = \theta(H, \delta)$ the unique positive solution to $P_{H^*}(\theta) = 1+\delta$, completing the proof of Theorem 2.3. □

5.3. **Proof of Proposition 5.6.** Fix $v \in F_A$. Consider the 2-matching $M_F(v)$ from Definition 5.5. Let $M_1, M_2, \ldots, M_q$ be the connected components of $M_F(v)$, such that M_1 (the connected component of $M_F(v)$ containing v) is a path (see Figure 6(a)). By Hölder's inequality (3.1) and by (3.5),

$$t(M_j, U) \leq (\mathbb{E}[U^2])^{\frac{|E(M_j)|}{2}} \lesssim p^{\frac{\Delta}{2}|E(M_j)|}\,, \tag{5.7}$$

for all $1 \leq j \leq q$.

To show (5.1) it suffices to prove that

$$\int_{[0,1]^{|V(F)|}} U(\boldsymbol{x}|F)\mathbf{1}\{x_v \in \overline{B}_b\}\mathrm{d}\boldsymbol{x}(F) \lesssim \varepsilon p^{|E(F)|}\,, \tag{5.8}$$

for every vertex $v \in F_A$, under the same assumptions of Proposition 5.6 (note that the choice of b does not depend on the vertex v). Since $U(\boldsymbol{x}|F) \leq U(\boldsymbol{x}|M_F(v))$, the inequality (5.8) holds whenever

$$\int_{[0,1]^{|V(M_F(v))|}} U(\boldsymbol{x}|M_F(v))\mathbf{1}\{x_v \in \overline{B}_b\}\mathrm{d}\boldsymbol{x}(M_F(v)) \lesssim \varepsilon p^{|E(F)|}\,. \tag{5.9}$$

Now, by (5.7), $t(M_j, U) \lesssim p^{\frac{\Delta}{2}|E(M_j)|}$ for every $2 \leq j \leq q$; thus, (5.9) would follow from showing that

$$\int_{[0,1]^{|V(M_1)|}} U(\boldsymbol{x}|M_1)\mathbf{1}\{x_v \in \overline{B}_b\}\mathrm{d}\boldsymbol{x}(M_1) \lesssim \varepsilon p^{\frac{\Delta}{2}|E(M_1)|} \tag{5.10}$$

holds for $b = b(p)$ such that $p^{\gamma_0} \leq b \ll 1$ for $\gamma_0 \leq 1/3$.

It therefore remains to prove (5.10). To this end, let $\beta = |E(M_1)|$ and $\beta_A = |V(M_1) \cap F_A|$. Since M_1 is a path, $|V(M_1)| = \beta + 1$. Label the vertices of $F_A \cap V(M_1)$ with $\{1, 2, \ldots \beta_A\}$, such that vertex 1 is the neighbor of one of the endpoints of M_1 (see Figure 6(b)).

Definition 5.8. *For $k \in [\beta_A]$ and sets $A_1, A_2, \dots A_k \subset [0,1]$, let*

$$t_k(A_1, A_2, \dots, A_k) = \int_{[0,1]^{\beta+1}} U(\boldsymbol{x}|M_1)\mathbf{1}\{x_1 \in A_1, x_2 \in A_2, \cdots, x_k \in A_k\}\mathrm{d}\boldsymbol{x}(M_1)\,,$$

be the contribution to $t(M_1, U)$ from the embeddings of M_1 where the vertex labeled i is in A_i for $i \in [k]$.

The main step in the proof of Proposition 5.6 is the following lemma. The ε dependence in the term $o(p^{\frac{1}{2}\Delta\beta})$ in (5.11) is suppressed and will be clear from the proof.

Lemma 5.9. *Let U be any graphon satisfying $\mathbb{E}[U^2] \lesssim p^\Delta$ (as per (3.10)) and $t(M_1, U) = \delta_\beta p^{\frac{1}{2}\Delta\beta}$. For any $\varepsilon > 0$, $\gamma_0 \le 1/3$, and $b_0 = p^{\gamma_0}$,*

$$t_{\beta_A}(B_{b_0}, B_{b_0}, \dots, B_{b_0}) \ge (\delta_\beta - \varepsilon)\, p^{\frac{\Delta}{2}\beta} - o(p^{\frac{\Delta}{2}\beta})\,. \tag{5.11}$$

This will immediately imply Proposition 5.6, since choosing $b = b_0 = p^{\gamma_0}$ in Lemma 5.9 yields

$$\begin{aligned}\int_{[0,1]^{|V(M_1)|}} U(\boldsymbol{x}|M_1)\mathbf{1}\{x_v \in \overline{B}_b\}\mathrm{d}\boldsymbol{x}(M_1) &\le \int_{[0,1]^{\beta+1}} U(\boldsymbol{x}|M_1)\mathbf{1}\{\exists s \in [\beta_A] : x_s \in \overline{B}_b\}\mathrm{d}\boldsymbol{x}(M_1)\\ &= \delta_\beta p^{\frac{\Delta}{2}\beta} - t_{\beta_A}(B_b, B_b, \dots, B_b) \le \varepsilon p^{\frac{\Delta}{2}\beta}\,.\end{aligned}$$

Proof of Lemma 5.9. Recalling (5.7), fix C throughout this proof such that $t(M_1, U) \le Cp^{\frac{\Delta}{2}|E(M_j)|}$. We will show that there exist $b_1 \gg b_2 \gg \dots \gg b_{\beta_A} \ge p^{\gamma_0}$ such that, for all $r \in [\beta_A]$,

$$t_r(B_{b_r}, B_{b_r}, \dots, B_{b_r}) \ge (\delta_\beta - \varepsilon r/\beta_A)\, p^{\frac{\Delta}{2}\beta} - o(p^{\frac{\Delta}{2}\beta})\,. \tag{5.12}$$

By choosing $r = \beta_A$ above, $t_{\beta_A}(B_{b_{\beta_A}}, B_{b_{\beta_A}}, \dots, B_{b_{\beta_A}}) \ge (\delta_\beta - \varepsilon)\, p^{\frac{\Delta}{2}\beta} - o(p^{\frac{\Delta}{2}\beta})$, and by monotonicity $t_{\beta_A}(B_{b_0}, B_{b_0}, \dots, B_{b_0}) \ge t_{\beta_A}(B_{b_{\beta_A}}, B_{b_{\beta_A}}, \dots, B_{b_{\beta_A}})$ (since $b_{\beta_A} \ge b_0 = p^{\gamma_0}$), which implies (5.11).

It remains to show (5.12); we proceed by induction on r. Let M_1^- be the path with $\beta - 1$ edges obtained by deleting the neighbor of vertex 1 in M (along with its edge). For $r = 1$ and any $b_1 \ll 1$,

$$\begin{aligned}t_1(\overline{B}_{b_1}) &= \int_{[0,1]^\beta} d(x_1)\mathbf{1}\{x_1 \in \overline{B}_{b_1}\}U(\boldsymbol{x}|M_1^-)\mathrm{d}\boldsymbol{x}(M_1^-)\\ &\le \left(\int_{\overline{B}_{b_1}} d^2(x_1)\mathrm{d}x_1\right)^{1/2}\left(\int_{[0,1]^2} U^2(x,y)\mathrm{d}x\mathrm{d}y\right)^{(\beta-1)/2} \lesssim p^{\frac{\Delta}{2}\beta}{b_1}^{1/2} \ll p^{\frac{\Delta}{2}\beta}\,,\end{aligned} \tag{5.13}$$

where in the first equality we integrated out the variable x_0 (corresponding to the neighbor of vertex 1 in M_1 which has degree 1) from $U(x_0, x_1)$, and the second used Hölder's inequality (3.1). Therefore, $t_1(B_{b_1}) = (\delta_\beta - o(1))p^{\frac{\Delta}{2}\beta}$, whenever $b_1 \ll 1$.

Now, assuming that (5.11) holds for all $r \in [k]$, we will extend it to $r = k+1$. Let $b', b'' \in [0,1]$, and consider the graph obtained by removing the two edges of M_1 incident to the vertex k (Figure 6(b)), which is a disjoint union of two paths. Let L_1 be the path which contains the vertex $k+1$, and L_2 the other path. Then integrating over the neighbor of $k+1$ in L_1, and using Hölder's inequality (3.1)

$$\int_{[0,1]^{|V(L_1)|}} U(\boldsymbol{x}|L_1)\mathbf{1}\{x_{k+1} \in \overline{B'_b}\}\mathrm{d}\boldsymbol{x}(L_1) = \left(\int_{\overline{B}_{b'}} d^2(x_{k+1})\mathrm{d}x_{k+1}\right)^{1/2}\left(\mathbb{E}[U^2]\right)^{\frac{|E(L)|-1}{2}}\,. \tag{5.14}$$

Furthermore, by Hölder's inequality (3.1), $t(L_2, U) \le \mathbb{E}[U^2]^{\frac{|E(L_2)|}{2}}$, which together with (5.14) above gives (since $|E(L_1)| + |E(L_2)| = \beta - 2$)

$$t_{k+1}([0,1],\dots,[0,1],B_{b''},\overline{B}_{b'}) \le \lambda(B_{b''})\left(\int_{\overline{B}_{b'}} d^2(x_{k+1})\mathrm{d}x_{k+1}\right)^{1/2}\left(\mathbb{E}[U^2]\right)^{\frac{\beta-3}{2}} \lesssim p^{\frac{\Delta}{2}\beta}\frac{\sqrt{b'}}{b''}\,, \tag{5.15}$$

where the last inequality uses (3.10) and (3.8). The above quantity is $o(p^{\frac{\Delta}{2}\beta})$ whenever $b' \ll b''^2$.

Define $M := \beta_a\lfloor\frac{C}{\varepsilon}\rfloor$. Given b_k, consider the following quantities for $a \in \{0,1,\dots,M\}$

$$D_a := t_{k+1}([0,1],\dots,[0,1],B_{b_k^{3^a}},\overline{B}_{b_k^{3^a}}) - t_{k+1}([0,1],\dots,[0,1],B_{b_k^{3^a}},\overline{B}_{b_k^{3^{a+1}}})\,. \tag{5.16}$$

Since D_a integrates the variable x_{k+1} over the set $B_{b_k^{3^{a+1}}}\backslash B_{b_k^{3^a}}$, which are disjoint for $0 \le a \le M$, we get $Cp^{\frac{\Delta}{2}\beta} \ge t(M_1, U) \ge \sum_{a=0}^{M} D_a$. Therefore, there exists $a_0 \in \{0,1,\dots,M\}$ such that $D_{a_0} \le \frac{\varepsilon}{\beta_a}p^{\frac{\Delta}{2}\beta}$. Let $b_{k+1} = b_k^{3^{a_0}}$. Thus,

$$\begin{aligned}t_{k+1}(B_{b_{k+1}},\dots,B_{b_{k+1}},\overline{B}_{b_{k+1}}) &\le t_{k+1}([0,1],\dots,[0,1],B_{b_{k+1}},\overline{B}_{b_{k+1}})\\ &\le \frac{\varepsilon}{\beta_a}p^{\frac{\Delta}{2}\beta} + O(p^{\frac{\Delta}{2}\beta}\sqrt{b_{k+1}}) = \frac{\varepsilon}{\beta_a}p^{\frac{\Delta}{2}\beta} + o(p^{\frac{\Delta}{2}\beta})\,,\end{aligned} \tag{5.17}$$

where the last inequality used (5.15) after adding and subtracting $t_{k+1}([0,1],\dots,[0,1],B_{b_{k+1}},\overline{B}_{b_{k+1}^3})$. Moreover, $t_k(B_{b_{k+1}},B_{b_{k+1}},\dots,B_{b_{k+1}}) \ge t_k(B_{b_k},B_{b_k},\dots,B_{b_k})$, since $b_{k+1} < b_k$. By this and (5.17),

$$\begin{aligned}t_{k+1}(B_{b_{k+1}},B_{b_{k+1}},\dots,B_{b_{k+1}}) &= t_k(B_{b_{k+1}},B_{b_{k+1}},\dots,B_{b_{k+1}}) - t_{k+1}(B_{b_{k+1}},\dots,B_{b_{k+1}},\overline{B}_{b_{k+1}})\\ &\ge t_k(B_{b_k},B_{b_k},\dots,B_{b_k}) - \frac{\varepsilon}{\beta_A}p^{\frac{\Delta}{2}\beta} - o(p^{\frac{\Delta}{2}\beta})\,,\end{aligned}$$

which, by the induction hypothesis, is at least

$$\big(\delta_\beta - k\varepsilon/\beta_A\big)p^{\frac{\Delta}{2}\beta} - \frac{\varepsilon}{\beta_A}p^{\frac{\Delta}{2}\beta} - o(p^{\frac{\Delta}{2}\beta}) \ge \big(\delta_\beta - (k+1)\varepsilon/\beta_A\big)p^{\frac{\Delta}{2}\beta} - o(p^{\frac{\Delta}{2}\beta})\,.$$

Notice that in (5.13) and subsequently all the $o(p^{\frac{\Delta}{2}\beta})$ terms are $O(p^{\frac{\Delta}{2}\beta}\sqrt{b_1})$. Also, observe that in the above proof b_{β_A} is obtained by raising b_1 to a large power depending on ε (see M, defined right before (5.16)). Thus, to ensure that $b_{\beta_A} \ge p^{\gamma_0}$ one can choose b_1 to be p raised to a sufficiently small power as a function of ε, which proves (5.12). ☐

6. Regular graphs: proof of Theorem 2.2

Here we prove the lower bound in Theorem 2.2 (the upper bound is by Proposition 2.4). Recall the key identity (4.21); the terms for $\mathscr{F}_H\backslash\{H\}$ will be treated as in §5, and our main focus will be $t(H,U)$.

6.1. **Embeddings of $t(H,U)$: adaptive choice of the degree threshold.**

Recall that H is a connected Δ-regular graph. To determine the embeddings of H which contribute to $t(H,U)$ we consider two cases: (a) $\tau(H) > |E(H)|/\Delta$, and (b) $\tau(H) = |E(H)|/\Delta$.

We begin with the case where H has minimum vertex cover of size greater than $|E(H)|/\Delta$. By Observation 4.8, H is not bipartite. In this case we show that the only non-negligible contribution to $t(H,U)$ comes from the embedding where all the vertices are in $\overline{B}_b$, for some appropriately chosen b.

Lemma 6.1. *Let H be a Δ-regular connected graph such that $\tau(H) > |E(H)|/\Delta$. Then there exists $\kappa := \kappa(H)$, such that for all $b \gg p^{\min\{\kappa,1/3\}}$ and any graphon U satisfying $\mathbb{E}[U^2] \lesssim p^\Delta$ (as per (3.10)),*

$$\int_{[0,1]^{|V(H)|}} U(\boldsymbol{x}|H)\mathbf{1}\{\exists j \in V(H), x_j \in B_b\}\mathrm{d}\boldsymbol{x}(H) \ll p^{|E(H)|}\,.$$

Proof. Fix $v \in V(H)$. Clearly it suffices to show that

$$s_v := \int_{[0,1]^{|V(H)|}} U(\boldsymbol{x}|H)\mathbf{1}\{x_v \in B_b\}\mathrm{d}\boldsymbol{x}(H) \ll p^{|E(H)|}\,.$$

Let E_v be the set of edges in $E(H)$ incident to the vertex v, and $H_v = (V(H)\backslash\{v\}, E(H)\backslash E_v)$.

(i) If $H = K_{\Delta+1}$ then $H_v = K_\Delta$. Let C_Δ be the cycle of length Δ. Then

$$s_v \le \lambda(B_b)t(K_\Delta, U) \lesssim \frac{p^\Delta}{b}\, t(C_\Delta, U) \lesssim \frac{p^\Delta}{b}\, p^{\frac{\Delta^2}{2}} \ll p^{\frac{\Delta(\Delta+1)}{2}}\,,$$

whenever $b \gg p^{1/3} \gg p^{\Delta/2}$. (If $\Delta = 2$, the above argument fails as the above bound gives $t(C_2, U) \lesssim p$, which is not enough. For this special case we recall the argument in [23, (3.6)],

$$s_v \le \lambda(B_b)\mathbb{E}[U] \lesssim b^{-1}p^{7/2}\log(1/p) \ll p^3\,,$$

where the second equality follows from (3.9) and the last equality by assumption $b \gg p^{1/3}$.)

(ii) Otherwise, since H is not a Δ-clique, H_v is an irregular graph of maximum degree Δ. Moreover $\tau(H_v) > |E(H)|/\Delta - 1$, since otherwise adding v to the vertex cover would mean that H has a vertex cover of size $|E(H)|/\Delta$, which contradicts hypothesis. Therefore, removing the edges incident to v and using the same argument as in Corollary 4.7, we get, for some $\kappa := \kappa(H)$

$$s_v \le \lambda(B_b)\int_{[0,1]^{|V(H)|-1}} \prod_{(x_i,x_j)\in E(H)\backslash E_v} U(x_i, x_j)\mathrm{d}\boldsymbol{x}(H) \lesssim \frac{p^\Delta}{b}\, p^{|E(H)|-\Delta+\kappa}\,,$$

since $\lambda(B_b) = p^\Delta/b$. Now, since $b \gg p^\kappa$, the result follows. ☐

Now assume that H is Δ-regular and has a vertex cover of size $|E(H)|/\Delta$ which implies that H is a bipartite graph with an equipartition (H_A, H_B).

Proposition 6.2. *Let H be a Δ-regular graph with $\tau = |E(H)|/\Delta$. Given $\varepsilon > 0$ and U satisfying $\mathbb{E}[U^2] \lesssim p^\Delta$ (as per (3.10)) there exists some $b_H = p^{\gamma(\varepsilon,H,U)} \gg p^{1/3}$ such that*

$$\int_{[0,1]^{V(H)}} U(\boldsymbol{x}|H)\mathbf{1}\{D_b\}\mathrm{d}\boldsymbol{x}(H) \le \varepsilon p^{|E(H)|}\,,$$

where $D_b = \{\exists u, w \in H_A : x_u \in B_b,\, x_w \in \overline{B}_b\}$.

The proof of the above proposition is postponed to §6.2. We first show how to complete the proof of Theorem 2.2 using this proposition.

Proof of Theorem 2.2 assuming Proposition 6.2. The next lemma shows that the contribution to $t(H, U)$ from embeddings where $x_i, x_j \in B_b$ for $(i, j) \in E(H)$ is negligible; we omit the proof as it is identical to that of Lemma 5.7.

Lemma 6.3. *Suppose H is Δ-regular with $\tau(H) = |E(H)|/\Delta$. For any $p^{1/3} \ll b \ll 1$*

$$\int_{[0,1]^{|V(H)|}} U(\boldsymbol{x}|H)\mathbf{1}\{\exists (i,j) \in E(H) \text{ such that } x_i, x_j \in B_b\} \mathrm{d}\boldsymbol{x}(H) \ll p^{|E(H)|} ,$$

where U is any graphon satisfying (3.10).

Let b_H be as in Proposition 6.2, and define $\theta = \theta_{b_H}$ and $\theta' = \theta'_{b_H}$ as in (5.4). Lemma 6.1 and Proposition 6.2 (together with Hölder's inequality (3.1)) will imply the following result.

Corollary 6.4. *Fix $\varepsilon > 0$. Let H be a Δ-regular graph and U be any graphon satisfying (3.10). Then*

$$t(H,U) - \varepsilon p^{|E(H)|} - o(p^{|E(H)|}) \le 2\theta^{\frac{|E(H)|}{\Delta}} \mathbf{1}\{H \text{ bipartite}\} p^{|E(H)|} + \theta'^{\frac{|E(H)|}{\Delta}} p^{|E(H)|} . \tag{6.1}$$

Proof. Consider the two cases depending on whether or not H is bipartite:

(i) H is non-bipartite: by Observation 4.8, $\tau(H) > |E(H)|/\Delta$ and by Lemma 6.1, the only non-negligible embedding places all vertices of H in $\overline{B}_{b_H}$. Hence, by Hölder's inequality (3.1),

$$t(H,U) \le \left(\int_{\overline{B}_{b_H} \times \overline{B}_{b_H}} U^2(x,y) \mathrm{d}x\mathrm{d}y \right)^{|E(H)|/\Delta} + o(p^{|E(H)|}) = \theta'^{\frac{|E(H)|}{\Delta}} p^{|E(H)|} + o(p^{|E(H)|}) ,$$

(ii) H is bipartite with equipartition (H_A, H_B): by Observation 4.8, $\tau(H) = |E(H)|/\Delta$. Now, by Proposition 6.2 up to an error of $\varepsilon p^{|E(H)|}$ all the contribution to $t(H,U)$ comes from embeddings where either all the vertices of H_A are in B_{b_H} or $\overline{B}_{b_H}$ and similarly for all vertices in H_B. To this end, consider the following integrals

$$\begin{aligned}
\Gamma_1 &= \int_{[0,1]^{|V(H)|}} U(\boldsymbol{x}|H)\mathbf{1}\{x_v \in B_{b_H}, \ \forall\, v \in H_A\} \mathrm{d}\boldsymbol{x}(H) , \\
\Gamma_2 &= \int_{[0,1]^{|V(H)|}} U(\boldsymbol{x}|H)\mathbf{1}\{x_v \in B_{b_H}, \ \forall\, v \in H_B\} \mathrm{d}\boldsymbol{x}(H) , \\
\Gamma_3 &= \int_{[0,1]^{|V(H)|}} U(\boldsymbol{x}|H)\mathbf{1}\{x_v \in \overline{B}_{b_H}, \ \forall\, v \in V(H)\} \mathrm{d}\boldsymbol{x}(H) .
\end{aligned}$$

By the above discussion and Hölder's inequality (3.1),

$$t(H,U) - \varepsilon p^{|E(H)|} \le \Gamma_1 + \Gamma_2 + \Gamma_3 \le 2\theta^{\frac{|E(H)|}{\Delta}} p^{|E(H)|} + \theta'^{\frac{|E(H)|}{\Delta}} p^{|E(H)|} .$$

We have thus established (6.1). ☐

Now, let δ_F be such that $t(F,U) = \delta_F p^{|E(F)|}$ for $F \in \mathscr{F}_H \cup \{H\}$. Using (4.21) we can write

$$\delta p^{|E(H)|} \le t(H,W) - p^{|E(H)|} = p^{|E(H)|} \sum_{F \in \mathscr{F}_H} \delta_F N(F,H) + o(p^{|E(H)|}) .$$

Since H is connected, any $F \in \mathscr{F}_H \backslash \{H\}$ must be irregular. Recalling $b_H = p^{\gamma(\varepsilon,H,U)} \gg p^{1/3}$ from Proposition 6.2, and notice that the proof of that proposition in fact allowed for $b_H \gg p^{\min\{\alpha,1/3\}}$, for any fixed $\alpha > 0$. In particular, taking $\alpha = \kappa := \kappa(H)$ is as in Lemma 6.1, we get $b_H \gg p^{\min\{\kappa,1/3\}}$. Now, by choosing $\gamma_0 = \gamma(\varepsilon, H, U)$ in Lemma 5.9, and using (5.2) and (5.3) it follows that for $F \in \mathscr{F}_H \backslash \{H\}$, with bipartition (F_A, F_B), the contribution to $t(F,U)$ comes from embeddings where every vertex in F_A is in B_{b_H} and every vertex in F_B is in $\overline{B}_{b_H}$. Arguments identical to those in (5.5) and (5.6) give

$$p^{|E(H)|} \sum_{F\in\mathscr{F}_H\setminus\{H\}} N(F,H)\left(\delta_F - \frac{\varepsilon}{\kappa(H)}\right) - o(p^{|E(H)|}) \le p^{|E(H)|} \sum_{F\in\mathscr{F}_H\setminus\{H\}} N(F,H)\theta^{\frac{|E(F)|}{\Delta}}\,, \tag{6.2}$$

where $\theta := \theta_{b_H}$ is as defined in (5.4). As for $F = H$, replacing ε by $\varepsilon/\kappa(H)$ in Corollary 6.4 gives

$$\begin{aligned}\left(\delta_H - \frac{\varepsilon}{\kappa(H)}\right)p^{|E(H)|} - o(p^{|E(H)|}) &= t(H,U) - \frac{\varepsilon}{\kappa(H)}p^{|E(H)|} - o(p^{|E(H)|})\\ &\le 2\theta^{\frac{|E(H)|}{\Delta}}\mathbf{1}\{H \text{ bipartite}\}p^{|E(H)|} + \theta'^{\frac{|E(H)|}{\Delta}}p^{|E(H)|}\,,\end{aligned} \tag{6.3}$$

where $\theta' = \theta'_{b_H}$ is defined in (5.4).

Combining (6.2) and (6.3), and revisiting (4.23), it follows that, for p small enough,

$$P_{H^*}(\theta) \ge 1 + \delta - \varepsilon - o(1) \ge \delta - 2\varepsilon\,.$$

Now, recall from Corollary 3.4 that $I_p(p+U) \ge (1+o(1))\big(\theta + \frac{\theta'}{2}\big)p^{\Delta}\log(1/p)$. Hence, for any $\varepsilon > 0$,

$$\liminf_{p\to 0}\left\{\frac{\frac{1}{2}\mathbb{E}\left[I_p(p+U)\right]}{p^{\Delta}I_p(1)} : t(H,p+U) \ge (1+\delta)p^{|E(H)|}\right\} \ge \min_{\theta,\theta'\in\Lambda(\mathcal{H})} \theta + \theta'/2\,,$$

where $\Lambda(\mathcal{H}) := \{\theta,\theta' \ge 0 : P_{H^*}(\theta) \ge 1+\delta-2\varepsilon\}$. By Observation 4.4, the minimum is attained either at $\theta = 0$ (as per the clique graphon from (2.4)) or $\theta' = 0$ (as per the anti-clique graphon from (2.5)). ☐

6.2. **Proof of Proposition 6.2.** Let $u, w \in H_A$ be such that $\mathrm{dist}_H(u,w)$, the distance between u and w, is 2. We will show

$$\int_{[0,1]^{|V(H)|}} U(\boldsymbol{x}|H)\mathbf{1}\{x_u \in B_b,\ x_w \in \overline{B}_b\}\mathrm{d}\boldsymbol{x}(H) \lesssim \varepsilon p^{|E(H)|}\,, \tag{6.4}$$

for all $u, w \in H_A$ with $\mathrm{dist}_H(u,w) = 2$. Clearly this suffices since the graph is connected.

Observe that H is a disjoint union of Δ matchings each of size $|H_A| = |E(H)|/\Delta$. By hypothesis there exists a path u, v, w in H. Let M_1 and M_2 be two disjoint matchings of H of size $|E(H)|/\Delta$, which has the edges (u,v) and (v,w), respectively. Consider the 2-matching $M_{u,w} = M_1 \cup M_2$. Since all the vertices have degree 2 in M, it is a disjoint union of cycles $C_{\beta_1}, C_{\beta_2}, \cdots, C_{\beta_k}$, where $\sum_{j=1}^k \beta_j = 2|E(H)|/\Delta$. W.l.o.g. assume that u, w lie on the cycle C_{β_1}. By Hölder's inequality,

$$t(C_{\beta_j}, U) \lesssim p^{\frac{\Delta}{2}|E(C_{\beta_j})|}\,. \tag{6.5}$$

As in (5.10), by considering the connected components of H, it is easy to show that Proposition 6.2 holds if for any $\varepsilon > 0$, there exists $b_H := p^{\gamma(\varepsilon,H,U)}$ (not depending on u, w) such that

$$\int_{[0,1]^{|V(C_{\beta_1})|}} U(\boldsymbol{x}|C_{\beta_1})\mathbf{1}\{x_u \in B_b,\ x_w \in \overline{B}_b\}\mathrm{d}\boldsymbol{x}(C_{\beta_1}) \lesssim \varepsilon p^{\frac{\Delta}{2}|E(C_{\beta_1})|}\,, \tag{6.6}$$

for any graphon U satisfying (3.10).

The rest of the section is devoted to the proof of (6.6). Label the vertices of C_{β_1} with $\{1, 2, \ldots, \beta_1\}$ in cyclic order. For $1 \le a \le \beta$ and sets $A_1, A_2 \subset [0,1]$, define

$$s_a(A_1, A_2) := \int_{[0,1]^{\beta}} U(\boldsymbol{x}|C_{\beta})\mathbf{1}\{x_a \in A_1, x_{a+2} \in A_2\}\mathrm{d}\boldsymbol{x}(C_{\beta_1})\,.$$

Lemma 6.5. *Let U be any graphon satisfying (3.10) and $t(C_{\beta_1}, U) = \delta_{\beta_1} p^{\frac{\Delta}{2}\beta_1}$. For any $\varepsilon > 0$, there exists $b_{u,w} := p^{\gamma(\varepsilon, C_{\beta_1}, U)}$ such that for all $1 \le a \le \beta_1$,*

$$s_a(B_b, \overline{B}_b) \le \frac{\varepsilon}{\beta_1} p^{\frac{\Delta}{2}\beta_1} + o(p^{\frac{\Delta}{2}\beta_1}), \tag{6.7}$$

where the ε dependence in $o(p^{\frac{\Delta}{2}\beta_1})$ is suppressed.

Proof. Note that by rotational symmetry of the cycle, $s_a(A_1, A_2)$ does not depend on a. Therefore, w.l.o.g. assume $a = 1$ and $x_1 \in B_b$.

Let $b', b'' \in [0, 1]$. Removing the edges incident to the vertex labeled 1 and by the same argument as in (4.9) (using Hölder's inequality (3.1) and bounds in (3.10), (3.8)), whenever $b'' \ll b'^2$ we get

$$s_1(B_{b'}, \overline{B}_{b''}) \lesssim p^{\frac{\Delta}{2}\beta_1} \frac{\sqrt{b''}}{b'} \ll p^{\frac{\Delta}{2}\beta_1}. \tag{6.8}$$

Fix $\varepsilon > 0$ and $0 \ll b_1 \ll 1$ and let $M := \beta_1 \lfloor \frac{C}{\varepsilon} \rfloor$, where C is constant in (6.5), i.e., $t(C_{\beta_1}, U) \le C p^{\frac{\Delta}{2}\beta_1}$. Consider the following quantities for $a \in \{0, 1, \ldots, M\}$

$$D_a := s_1(B_{b_1^{3^a}}, \overline{B}_{b_1^{3^a}}) - s_1(B_{b_1^{3^a}}, \overline{B}_{b_1^{3^{a+1}}}). \tag{6.9}$$

Since D_a integrates the variable x_3 over the set $B_{b_1^{3^{a+1}}} \backslash B_{b_1^{3^a}}$, which is disjoint for $0 \le a \le M$. Therefore, we get $\delta_{\beta_1} p^{\frac{\Delta}{2}\beta_1} = t(C_{\beta_1}, U) \ge \sum_{a=0}^{M} D_a$. Therefore, there exists $a_0 \in \{0, 1, \ldots, M\}$ such that $D_{a_0} \le \frac{\varepsilon}{\beta_1} p^{\frac{\Delta}{2}\beta_1}$. Let $b_{u,w} = b_1^{3^{a_0}}$. This choice of $b_{u,w}$ together with (6.8) implies that

$$\begin{aligned} s_1(B_{b_{u,w}}, \overline{B}_{b_{u,w}}) &= (s_1(B_{b_{u,w}}, \overline{B}_{b_{u,w}}) - s_1(B_{b_{u,w}}, \overline{B}_{b_{u,w}^3})) + s_1(B_{b_{u,w}}, \overline{B}_{b_{u,w}^3}) \\ &\le \frac{\varepsilon}{\beta_1} p^{\frac{\Delta}{2}\beta_1} + O(p^{\frac{\Delta}{2}\beta_1} b_{u,w}^{\frac{3a_0}{2}}) = \frac{\varepsilon}{\beta_1} p^{\frac{\Delta}{2}\beta_1} + o(p^{\frac{\Delta}{2}\beta_1}). \end{aligned}$$

□

Notice that in the above argument the chosen $b_{u,w}$ depends on the cycle C_{τ_1} (in particular on the vertices u and w) and the graphon U, in contrast with the case of irregular graphs. Recall the left-hand side in (5.10) and $s_a(B_b, \overline{B}_b)$ from (6.7), respectively. Notice that the left-hand side in (5.10) has an obvious monotonicity in b whereas the latter does not. Using this monotonicity in the irregular case, the choice of b could be made independent of the graph H and the graphon U (cf. the proof of Lemma 5.9).

In this case, for every pair $u, w \in H_A$, the 2-matching $M_{u,w}$ and the cycle C_{τ_1} could be distinct. Thus, to prove (6.6) and one has to provide additional arguments to show a single degree threshold b_H works for all such u, w. To this end, we make the following observation which strengthens (6.4).

Observation 6.6. *Given $\varepsilon > 0$, define $\ell = \lfloor \delta/\varepsilon \rfloor$ (where δ is as in Theorem 2.2). Then for $m \ge \ell$ and $u, w \in H_A$ as in (6.4) and for any sequence $b_1, b_2, \ldots b_m \gg p^{1/3}$ such that $b_{i+1} \le b_i^3$ for all $i \in [m]$, there exists a subsequence $b_{j_1} > b_{j_2} > \ldots > b_{j_{m-\ell+1}}$ all of which satisfy (6.4).*

Proof. Let $\mathscr{B} := \{b_1, b_2, \ldots, b_m\}$. By (6.8) and (6.9), there exists some b_{j_1} among $b_1, \ldots, b_m$ which satisfies (6.4), as $m > \ell$. Consider the largest such b_{j_1}. Now, look at the set $\mathscr{B} \backslash \{j_1\}$ and repeat the above argument. Clearly, this argument can be repeated as long as the length of the remaining sequence is at least ℓ. Thus, we obtain the sequence $b_{j_1} > b_{j_2} > \ldots > b_{j_{m-\ell+1}}$ which satisfies (6.4). □

Using the above result, the proof of Proposition 6.2 can be completed in the following way: Let $S_H = \{(u, w) : u, w \in H_A, \mathrm{dist}_H(u, w) = 2\}$, the set of pairs of vertices (u, w) in H_A at distance two from each other. All that is left is to find a b such that (6.4) holds for all $(u, w) \in S_H$. To this end, take $m = \ell \cdot (|S_H| + 1)$ (where ℓ appears in the statement of Observation 6.6) and $b_1, b_2, \dots b_m$ such that $p^{1/3} \ll b_{j+1} \le b_j^3$, for all $j \in [m]$. Fix $\{(u_i, w_i) : 1 \le i \le |S(H)|\}$ an arbitrary order on S_H.

The proof now follows by repeated applications of Observation 6.6. In the first step we apply the observation to the pair (u_1, v_1) and the sequence $b_1, b_2, \dots b_m$, to obtain a subsequence. In the next step we apply the observation to the pair (u_2, v_2) and the subsequence obtained in the previous step. We iterate the above over all the elements in S_H. Notice that m is chosen to be large enough to ensure that this iterative procedure can be run over all the elements of S_H and eventually we end with a subsequence of $b_1, b_2, \dots b_m$ of length at least ℓ . Clearly by construction, choosing b to be any of these ℓ elements, (6.4) holds for all $(u, w) \in S_H$. □

7. From the continuous to the discrete variational problem

The following simple lemma connects the discrete and the continuous variational problems:

Lemma 7.1 ([23, Lemma 2.1])**.** *For any fixed graph H and p, n, δ we have $\phi(H, p, \delta) \le n^{-2}\phi(H, n, p, \delta)$.*

Proof. Given $G_n \in \mathscr{G}_n$ form a graphon W^{G_n} as follows: divide $[0, 1]$ into n equal intervals $I_1, I_2, \dots, I_n$ and set $W^{G_n}(x, y) = a_{ij}$ if $x \in I_i, y \in I_j$ and $i \ne j$, and $W^{G_n}(x, y) = p$ otherwise. Clearly we have $t(H, G_n) \ge t(H, W^{G_n})$ and $I_p(W^{G_n}) = n^{-2}I_p(G_n)$ (diagonal entries contribute 0 to $I_p(W^{G_n})$). □

7.1. Proof of Theorem 1.5.

Fix a graph H with maximum degree $\Delta \ge 2$.

Suppose first that H is irregular. By Theorem 2.3 together with Lemma 7.1,

$$\lim_{n\to\infty} \frac{\phi(H, n, p, \delta)}{n^2 p^\Delta \log(1/p)} \ge \lim_{p\to 0} \frac{\phi(p, H, \delta)}{p^\Delta \log(1/p)} = \theta(H, \delta)\,,$$

and the anti-clique construction (from §1.3) gives a matching upper bound.

Next, suppose H is regular. By Theorem 2.2 combined with Lemma 7.1,

$$\lim_{n\to\infty} \frac{\phi(H, n, p, \delta)}{n^2 p^\Delta \log(1/p)} \ge \lim_{p\to 0} \frac{\phi(H, p, \delta)}{p^\Delta \log(1/p)} = \min\left\{\theta(H, \delta)\,,\, \tfrac{1}{2}\delta^{2/|V(H)|}\right\}\,,$$

and a matching upper bound is obtained by either the clique or the anti-clique constructions (from §1.3), whichever has the lower entropy. □

Remark 7.2. As mentioned below Theorem 1.5, when H is regular and $n^{-2/\Delta} \ll p \ll n^{-1/\Delta}$ (a range in which the anti-clique construction is no longer applicable), it follows from our arguments that

$$\lim_{n\to\infty} \frac{\phi(H, n, p, \delta)}{n^2 p^\Delta \log(1/p)} = \tfrac{1}{2}\delta^{2/|V(H)|}\,.$$

Indeed, the upper bound follows from the discrete clique G_A just as before, and it remains to verify the lower bound, to which end it suffices to consider—thanks to Theorem 2.2 and Lemma 7.1—the case where $\frac{1}{2}\delta^{2/|V(H)|} > \theta(H, \delta)$. Assume that (3.5) holds, and then choose b such that $np^\Delta \ll b \ll 1$. Now, recalling the definition of B_b from (3.3), by (3.7) we have

$$\lambda(B_b) \lesssim p^\Delta/b \ll n^{-1}\,,$$

which implies that the number of vertices with degree bigger than nb is 0. Thus, in all the calculations in §6 one can rule out the anti-clique graphon, and the above result follows.

7.2. **Disconnected graphs.** The arguments for connected graphs can be easily generalized to disconnected graphs. By considering the different connected components, the solution of the variational problem can be expressed as a two variable constrained optimization problem.

Corollary 7.3. *Let H be a graph with maximum degree Δ and connected components $H_1, H_2, \ldots, H_s$ for some $s \geq 1$. Define $P_{H_i^*}(x)$ as Defs. 1.1–1.2. Then for fixed $\delta > 0$,*

$$\lim_{p\to 0} \frac{\phi(H,p,\delta)}{p^\Delta \log(1/p)} = \inf \left\{ \theta + \tfrac{1}{2}\theta' \;:\; \prod_{i=1}^{s} \left(P_{H_i^*}(\theta) + \mathbf{1}\{H_i \ \Delta\text{-regular}\} \theta'^{|E(H_i)|/\Delta} \right) = 1 + \delta \right\},$$

For disconnected graphs the solution of the variational problem for large δ might not be attained by the clique or the anti-clique graphons, but by a mixture of these two.

Example 7.4. Let H be the disjoint union of a triangle (K_3) and a 2-star ($K_{1,2}$). By Corollary 7.3,

$$\begin{aligned} \lim_{p\to 0} \frac{\phi(H,p,\delta)}{p^2 \log(1/p)} &= \inf_{\theta,\theta' \geq 0} \left\{ \theta + \tfrac{1}{2}\theta' \;:\; (1 + 3\theta + \theta'^{3/2})(1+\theta) = 1 + \delta \right\} \\ &= \inf_{z_1,z_2 \geq 0} \left\{ z_1 + \tfrac{1}{2} z_2^2 \;:\; (1 + 3z_1 + z_2^3)(1 + z_1) = 1 + \delta \right\}. \end{aligned} \tag{7.1}$$

The constraint $(1 + 3z_1 + z_2^3)(1 + z_1) = 1 + \delta$ is a quadratic in z_1 with a unique positive solution

$$\lambda(z_2, \delta) := \frac{1}{6} \left(-4 - z_2^3 + \sqrt{12\delta + z_2^6 - 4z_2^3 + 16} \right). \tag{7.2}$$

From (7.1) it follows that

$$\lim_{p\to 0} \frac{\phi(H,p,\delta)}{p^2 \log(1/p)} = \inf \left\{ \lambda(z_2,\delta) + \tfrac{1}{2} z_2^2 \;:\; z_2 \in [0, \delta^{1/3}] \right\}.$$

The clique corresponds to $z_1 = 0$, so $z_2 = (1+\delta)^{\frac{1}{3}}$, and its objective value is $\frac{1}{2}(1+\delta)^{\frac{2}{3}} \sim \frac{1}{2}\delta^{\frac{2}{3}}$ for large δ. The anti-clique corresponds to $z_2 = 0$, so by (7.2) its objective value is $\sim \sqrt{\delta/3}$ for large δ. Thus, for large δ both solutions are inferior to $(\frac{1}{2}\sqrt{\delta}, 2^{-1/3}\delta^{1/6})$, which gives the objective value $\frac{1}{2}\sqrt{\delta} + O(\delta^{\frac{1}{3}})$.

Acknowledgements. This work was initiated when the first author was an intern at the Theory Group of Microsoft Research, Redmond. E.L. was supported in part by NSF grant DMS-1513403 and Y.Z. was supported by a Microsoft Research Ph.D. Fellowship.

References

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

[2] 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.

[3] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi. Convergent sequences of dense graphs II. Multiway cuts and statistical physics. *Ann. of Math. (2)*, 176(1):151–219, 2012.

[4] J. I. Brown, C. A. Hickman, and R. J. Nowakowski. The independence fractal of a graph. *J. Combin. Theory Ser. B*, 87(2):209–230, 2003.

[5] J. I. Brown and R. J. Nowakowski. Average independence polynomials. *J. Combin. Theory Ser. B*, 93(2):313–318, 2005.

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

[7] S. Chatterjee and A. Dembo. Nonlinear large deviations. Preprint, available at `arXiv:1401.3495`.
[8] S. Chatterjee and P. S. Dey. Applications of Stein's method for concentration inequalities. *Ann. Probab.*, 38(6):2443–2485, 2010.
[9] 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.
[10] M. Chudnovsky and P. Seymour. The roots of the independence polynomial of a clawfree graph. *J. Combin. Theory Ser. B*, 97(3):350–357, 2007.
[11] B. Demarco and J. Kahn. Tight upper tail bounds for cliques. *Random Structures Algorithms*, 41(4):469–487, 2012.
[12] B. DeMarco and J. Kahn. Upper tails for triangles. *Random Structures Algorithms*, 40(4):452–459, 2012.
[13] H. Finner. A generalization of Hölder's inequality and some probability inequalities. *Ann. Probab.*, 20(4):1893–1901, 1992.
[14] S. Janson, T. Łuczak, and A. Rucinski. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.
[15] S. Janson, K. Oleszkiewicz, and A. Ruciński. Upper tails for subgraph counts in random graphs. *Israel J. Math.*, 142:61–92, 2004.
[16] S. Janson and A. Ruciński. The infamous upper tail. *Random Structures Algorithms*, 20(3):317–342, 2002.
[17] S. Janson and A. Ruciński. The deletion method for upper tail estimates. *Combinatorica*, 24(4):615–640, 2004.
[18] 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.
[19] L. Lovász. *Large networks and graph limits*, volume 60 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2012.
[20] L. Lovász and M. D. Plummer. *Matching theory*, volume 121 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam; Akadémiai Kiadó (Publishing House of the Hungarian Academy of Sciences), Budapest, 1986. Annals of Discrete Mathematics, 29.
[21] L. Lovász and B. Szegedy. Limits of dense graph sequences. *J. Combin. Theory Ser. B*, 96(6):933–957, 2006.
[22] E. Lubetzky and Y. Zhao. On replica symmetry of large deviations in random graphs. *Random Structures Algorithms*, 47(1):109–146, 2015.
[23] E. Lubetzky and Y. Zhao. On the variational problem for upper tails in sparse random graphs. *Random Structures Algorithms*, to appear.
[24] 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.

B. B. Bhattacharya,
Department of Statistics, Stanford University, Stanford, CA 94305, USA.
E-mail address: `bhaswar@stanford.edu`

S. Ganguly,
Department of Mathematics, University of Washington, Seattle, WA 98195, USA
E-mail address: `sganguly@math.washington.edu`

E. Lubetzky
Courant Institute, New York University, 251 Mercer Street, New York, NY 10012, USA.
E-mail address: `eyal@courant.nyu.edu`

Y. Zhao
Mathematical Institute, University of Oxford, Oxford OX2 6GG, UK.
E-mail address: `yufei.zhao@maths.ox.ac.uk`