Sandwiching a densest subgraph by consecutive cores

Pu Gao*
University of Toronto
pu.gao@utoronto.ca

Abstract

In this paper, we show that in the random graph $\mathcal{G}(n, c/n)$, with high probability, there exists an integer $\widetilde{k}$ such that a subgraph of $\mathcal{G}(n, c/n)$, whose vertex set differs from a densest subgraph of $\mathcal{G}(n, c/n)$ by $O(\log^2 n)$ vertices, is sandwiched by the $\widetilde{k}$ and the $(\widetilde{k}+1)$-core, for almost all sufficiently large c. We determine the value of $\widetilde{k}$. We also prove that (a), the threshold of the k-core being balanced coincides with the threshold that the average degree of the k-core is at most $2(k-1)$, for all sufficiently large k; (b) with high probability, there is a subgraph of $\mathcal{G}(n, c/n)$ whose density is significantly denser than any of its non-empty cores, for almost all sufficiently large $c > 0$.

1 Introduction

For any graph G and any positive integer k, the *k-core* of G is the maximum subgraph of G with minimum degree at least k. The density of G, denoted by $\rho(G)$, is defined by the number of edges of G divided by the number of vertices in G. We say G is balanced if there is no subgraph of G with density greater than G.

Let $\mathcal{G}(n, p)$ denote the binomial random graph on vertex set $[n]$, which includes each edge in K_n independently with probability p. The relevant range of p considered in this paper is $\Theta(1/n)$; hence with high probability the density of $\mathcal{G}(n, p)$ is bounded. Let $k \geq 1$ be a fixed integer. In the work of determining the threshold of k-orientability of the random graph $\mathcal{G}(n, c/n)$, done by Cain, Sanders and Wormald [3] and Fernholz and Ramachandran [5], it was proved that the following three graph properties have the same sharp threshold: (i) the $(k+1)$-core has density at most k; (ii) no subgraph of the $(k+1)$-core has density more than k; (iii) no subgraph of $\mathcal{G}(n, c/n)$ has density more than k. (Note that the latter two thresholds coincide trivially and the non-trivial part is that these two thresholds coincide with the one in (i).) Given a graph G, let $\rho^*(G)$ denote the density of a densest subgraph of G. It is then natural to ask the question: "what is the threshold for $\rho^*(\mathcal{G}(n, c/n)) \leq \alpha$, where α is a given real number?" The work in [3, 5] answers this question if α is an integer, but there is no answer otherwise. We can even further ask: "Where is the densest subgraph

*Research supported by the Humboldt Foundation and the NSERC PDF fellowship. Major part of this work was done when this author worked in Max Planck Institüt für Informatik, Saarbrücken, Germany.

of $\mathcal{G}(n, c/n)$ and what is the distribution of the random variable $\rho^*(\mathcal{G}(n, c/n))$?" Intuition tells that $\rho^*(\mathcal{G}(n, c/n))$ is likely to be "close to" (meaning allowing an $o(1)$ error) the density of the densest non-empty core of $\mathcal{G}(n, c/n)$. If that is the case, then we can easily deduce the concentration of $\rho^*(\mathcal{G}(n, c/n))$ and determine its asymptotic value by applying results in [12], in which the density of each non-empty core of $\mathcal{G}(n, c/n)$ is determined. In fact, it is true for certain values of c (see Proposition 2.2). However, we will prove that this is not true in general. In fact, we will show that for almost all sufficiently large c, $\rho^*(\mathcal{G}(n, c/n))$ is significantly greater than the density of the densest core of $\mathcal{G}(n, c/n)$. The goal of this paper is to answer the question of where the densest subgraph of G is, rather than gaining new knowledge of $\rho^*(G)$, for $G \sim \mathcal{G}(n, c/n)$ (here and throughout the paper, $G \sim \mathcal{G}(n, p)$ means that G is random graph distributed as $\mathcal{G}(n, p)$). We will show that for almost all sufficiently large constants $c > 0$ (except that c lies in a countable set $\mathbf{C}$, defined later), asymptotically almost surely (a.a.s.) there are constant $\widetilde{k} > 0$ and a densest subgraph H of $\mathcal{G}(n, c/n)$ such that H is contained in the $\widetilde{k}$-core and H contains all but at most $O(\log^3 n)$ vertices of the $(\widetilde{k} + 1)$-core. Moreover, H is significantly denser than any non-empty core of $\mathcal{G}(n, c/n)$.

The key part of the proofs in [3] was to show that if a random $(k+1)$-core has density at most $k-\epsilon$, for some absolute constant $\epsilon > 0$, then a.a.s. it does not contain any subgraph with density more than k. In other words, the $(k+1)$-core is "loosely" balanced, and the density of a densest subgraph of $G \sim \mathcal{G}(n, c/n)$ is close to the density of the $(k+1)$-core by letting $\epsilon \to 0$. In our current work, we strengthen this result by showing that if k is sufficiently large, then a.a.s. a random $(k+1)$-core with density $k - \epsilon$ is "strictly" balanced, i.e. it contains no denser subgraphs. We also prove that a.a.s. a random $(k + 1)$-core with density $k + \epsilon$ is not balanced, when k is sufficiently large. (In fact, we will prove a stronger result. See Theorem 2.4.) Therefore, the threshold of the $(k+1)$-core being balanced coincides with the threshold that the $(k + 1)$-core has density at most k.

Using the threshold for a random k-core being balanced, we chase the densest subgraph of $\mathcal{G}(n, c/n)$ by first searching for its densest core. We will determine the value of $\widetilde{k}$ such that the densest core is either the $\widetilde{k}$ or the $(\widetilde{k} + 1)$-core. Moreover, we will prove that, roughly speaking, the $\widetilde{k}$ and the $(\widetilde{k} + 1)$-core sandwich the densest subgraph of $\mathcal{G}(n, c/n)$. Given a graph G, consider the graph sequence $(G_0, G_1, \ldots, G_\ell)$, where $G_0 = G$, G_k is the k-core of G for every $k \geq 1$, and ℓ is the largest integer such that G_ℓ is non-empty. Then this graph process enumerates all non-empty cores of G. For every $k \geq 0$, going from G_k to G_{k+1}, vertices with relatively low degrees are removed. We would expect that the density of the graphs G_k will first increase as k increases. For instance, if a graph has average degree greater than $2k$, then removing a vertex with degree at most k would increase the density of the graph. This monotonicity will terminate before step t when the t-core becomes balanced. If $G_0 \sim \mathcal{G}(n, c/n)$, t is a concentrated random variable whose value can be determined by applying Theorem 2.4 (stated in Section 2). We use this information to determine which core is the densest and then further determine where a densest subgraph of G_0 should be located.

For future work, it is interesting to answer the question: "What is the structure of the densest subgraph of $G \sim \mathcal{G}(n, c/n)$, where is it located exactly and how dense it is?" Our result in this paper shows that we only need to restrict our search to the subgraphs sandwiched by the $\widetilde{k}$ and the $(\widetilde{k} + 1)$-core. One possible approach is to consider the stripping process

$(G'_0, G'_1, \ldots, G'_\tau)$, where G'_0 and G'_τ are the $\widetilde{k}$ and the $(\widetilde{k}+1)$-cores of G and G_t is obtained from G_{t-1} by removing a vertex with minimum degree. What is the density of the densest graph among the G_t and how close is it from a densest subgraph of G? Indeed, our proof indicates that there is a graph G_t in the stripping process whose density is significantly denser than any non-empty core of G. But we believe that none of these G'_t is a densest subgraph of G, even allowing an $o(1)$ error. Thus, new techniques will be required to further study the graph structures between two consecutive cores.

We state lower and upper bounds of $\rho^*(\mathcal{G}(n, c/n))$ in Section 2 as a simple corollary of our main results (for large c). Note that these bounds are not new. The previous work in [3, 5] yields the same bounds, which also work for small c. These bounds are not tight. Very little is known about this random variable, including the existence of the limit of its expectation. In very special cases (see Proposition 2.2), this variable is concentrated and its value is determined. This makes us believe that in general cases, not only the limit of its expectation shall exist but the variable itself is likely to be concentrated.

2 Main results

Recall that for a given graph G, $\rho(G)$ denotes the density of G and $\rho^*(G)$ denotes the density of a densest subgraph of G. Given an integer k, let G_k denote the k-core of G.

It is well known that given $k \geq 3$, there is a constant $c_k > 0$: for any $c > c_k$, there exist two constants $\alpha_{c,k}$ and $\beta_{c,k}$ such that a.a.s. the k-core of $\mathcal{G}(n, c/n)$ contains $\alpha_{c,k} n + o(n)$ vertices and $\beta_{c,k} n + o(n)$ edges; for any $c < c_k$, a.a.s. the k-core of $\mathcal{G}(n, c/n)$ is empty. This was first proved by Pittel, Spencer and Wormald [12], and was later re-proved in several papers, including Cain and Wormald [4], Kim [7] and Molloy [11]. The constants c_k, $\alpha_{c,k}$ and $\beta_{c,k}$ are defined as follows.

For any integer $k \geq 0$ and real $\lambda \geq 0$, define

$$f_k(\lambda) = e^{-\lambda} \sum_{i \geq k} \frac{\lambda^i}{i!}.$$

For any $k \geq 3$, define

$$h_k(\mu) = \frac{\mu}{f_{k-1}(\mu)},$$

and let

$$c_k = \inf\{h_k(\mu), \mu > 0\}.$$

For any $c > c_k$, define $\mu_{c,k}$ to be the larger solution of $h_k = c$, and define

$$\alpha_{c,k} = f_k(\mu_{c,k}), \quad \beta_{c,k} = \frac{1}{2} \mu_{c,k} f_{k-1}(\mu_{c,k}). \tag{2.1}$$

Let $\mathcal{G}(n, m)$ denote the random graph with n vertices and m edges, with uniform distribution. The following theorem is from [12].

Theorem 2.1 *Let $k \geq 3$ and $c > 0$ be fixed and suppose that $m = cn/2$. If $c > c_k$ then a.a.s. $\mathcal{G}(n,m)$ has a non-empty k-core with $\alpha_{c,k}n + o(n)$ vertices and $\beta_{c,k}n + o(n)$ edges. If $c < c_k$ then a.a.s. the k-core of $\mathcal{G}(n,m)$ is empty.*

Remark: It follows easily that the same conclusion holds in $\mathcal{G}(n,p)$ with $p = c/n$ by conditioning on the number of edges in $\mathcal{G}(n,p)$. A direct analogous statement for $\mathcal{G}(n,p)$ can also be found in [7].

It is easy to see that c_k is an increasing sequence and $\beta_{c,k}/\alpha_{c,k}$ is an increasing function of $c > c_k$, given $k \geq 3$. It is also easy to check that $\beta_{c,k}/\alpha_{c,k} \to \infty$ as $c \to \infty$ and $\lim_{c \to c_k} \beta_{c,k}/\alpha_{c,k} < k - 1$ (by using the precise estimate of c_k and the calculation leading to it in [13, Lemma 1]). Thus, it is valid to define $\widetilde{c}_k$ to be the unique c satisfying $\beta_{c,k}/\alpha_{c,k} = k - 1$ for every $k \geq 3$. Define $\mathbf{C} = \{\widetilde{c}_k : k \geq 3\}$. It was shown in [3, 5] that for every $k \geq 3$ and every $\epsilon > 0$, if the k-core G_k of $\mathcal{G}(n, c/n)$ has density less than $k-1$ (corresponding to $c < \widetilde{c}_k$), then a.a.s. G_k contains no subgraph with density greater than k. The following proposition follows by letting $\epsilon \to 0$ and by Theorem 2.1 together with a simple coupling argument (to deal with the case that the density of the k-core is between $k - 1 - \epsilon$ and $k - 1 + \epsilon$). For the complete proof of the proposition, refer to [8, Corollary 31].

Proposition 2.2 *Let $k \geq 3$ be fixed. If $c = \widetilde{c}_k + o(1)$ and $G \sim \mathcal{G}(n, c/n)$, then a.a.s. $\rho^*(G) = \rho(G_k) + o(1) = k - 1 + o(1)$.*

Proposition 2.2 leads us to guess that the same conclusion holds for all $c > c_3$ rather than only for $c \in \mathbf{C}$. Surprisingly, as we state in the next theorem, this is not true. Let $\mathbf{R}_0 = \mathbf{R}_+ \setminus \{\widetilde{c}_k : \; k \geq 3\}$, where $\mathbf{R}_+$ is the set of all positive real numbers. Clearly, $\mathbf{R}_0$ is $\mathbf{R}_+$ except for a countable set of real numbers. Let $\{G(n)\}$ and $\{H(n)\}$ be two sequences of graphs. We say $G(n)$ is *significantly denser* than $H(n)$, if there exists an absolute constant $\eta > 0$ such that $\rho(G(n)) \geq \rho(H(n)) + \eta$, for all sufficiently large n.

Theorem 2.3 *For all sufficiently large $c \in \mathbf{R}_0$, a densest subgraph of $G \sim \mathcal{G}(n, c/n)$ is significantly denser than any non-empty core of G.*

By Theorem 2.3, a densest subgraph H of $G \sim \mathcal{G}(n, c/n)$, for almost all sufficiently large c, cannot be any non-empty core of G. Moreover, H is likely to differ from every non-empty core of G by at least, say, $\Theta(n)$ vertices. However, gaining information of which core is the densest will help us to chase the densest subgraphs. The result in [3, 5] only compares $\rho^*(G_k)$ with k (instead of $\rho(G_k)$), if $\rho(G_k) < k$. Thus, it does not imply that G_k contains no denser cores and also gives no information of which core is the densest. We strengthen their result by proving that G_k is indeed a.a.s. balanced and thus contains no denser $(k+1)$-core. Using that, we further specify two cores and show that the densest core of G must be one of them.

Theorem 2.4 *The threshold at which the k-core of $G \sim \mathcal{G}(n, c/n)$ is balanced coincides with the threshold at which the density of the k-core is at most $k - 1$, i.e. at $\widetilde{c}_k$, provided k is sufficiently large. Moreover, if $c > \widetilde{c}_k$, then a.a.s. there exists $H \subseteq G_k$ which is significantly denser than G_k.*

Corollary 2.5 *For all sufficiently large $c \in \mathbf{R}_0$, there exists an integer $\widetilde{k} = \widetilde{k}(c)$ and a real $\widetilde{d} = \widetilde{d}(c) < \widetilde{k}$, such that a.a.s. the $\widetilde{k}$-core is not balanced and the $(\widetilde{k}+1)$-core is balanced in $G \sim \mathcal{G}(n, c/n)$ and $\rho(G_{\widetilde{k}+1}) = \widetilde{d} + o(1)$. Moreover, if the k-core is the densest core of G, then a.a.s. $k \in \{\widetilde{k}, \widetilde{k}+1\}$.*

Proof. Let $\widetilde{k}(c)$ denote the maximum integer k such that $c > c_k$ and $\beta_{c,k}/\alpha_{c,k} \geq k-1$. It is easy to check that c_k is an increasing sequence and $c_k \to \infty$ as $k \to \infty$. Thus, there are only finitely many k's such that $c_k < c$. Moreover, $\beta_{c,k}/\alpha_{c,k} \geq k-1$ holds for $k = 3$, provided that c is sufficiently large. This verifies that $\widetilde{k}(c)$ is well defined for all sufficiently large c. By Theorem 2.1 and the definition of $\widetilde{k}$, and since $c \in \mathbf{R}_0$, there is $\widetilde{d} = \widetilde{d}(c) < \widetilde{k}$ such that a.a.s. the density of the $(\widetilde{k}+1)$-core is $\widetilde{d} + o(1)$, whereas the density of the $\widetilde{k}$-core is greater than $\widetilde{k}-1$. A.a.s. the $(\widetilde{k}+1)$-core is balanced whereas the $\widetilde{k}$-core is not balanced by Theorem 2.4. Suppose that the k-core is the densest core of G. Then, a.a.s. $k \leq \widetilde{k}+1$, since a.a.s. the $(\widetilde{k}+1)$-core is balanced. Since we have shown that a.a.s. $\rho(G_{\widetilde{k}}) > \widetilde{k}-1$, it follows then that a.a.s. $\rho(G_k) \geq \rho(G_{\widetilde{k}}) > \widetilde{k}-1$ since G_k is the densest core. We prove next that this implies $k \geq \widetilde{k}$. Assume $k \leq \widetilde{k}-1$. Then, repeatedly removing all vertices of degree at most $k \leq \widetilde{k}-1$ from G_k will only increase the density of the resulting graph as $\rho(G_k) > \widetilde{k}-1$. This means that $\rho(G_{k+1}) > \rho(G_k)$, contradicting with G_k being the densest core. Therefore, we have shown that a.a.s. $k \geq \widetilde{k}$. It follows then that a.a.s. $k \in \{\widetilde{k}, \widetilde{k}+1\}$. ∎

Corollary 2.5 states that either the $\widetilde{k}$ or the $(\widetilde{k}+1)$-core is the densest core of $\mathcal{G}(n, c/n)$, whereas Theorem 2.3 states that neither of them is the densest subgraph of $\mathcal{G}(n, c/n)$. In the next theorem, we prove that there is a densest subgraph which is almost sandwiched by these two cores. For any two sets A and B, let $A \Delta B$ denote the symmetric difference of A and B, i.e. $A \Delta B = (A \setminus B) \cup (B \setminus A)$.

Theorem 2.6 *Consider $G \sim \mathcal{G}(n, c/n)$. For all sufficiently large $c \in \mathbf{R}_0$, a.a.s. $\rho^*(G) \leq \widetilde{k}$ and there exists $G_{\widetilde{k}+1} \subseteq H \subseteq G_{\widetilde{k}}$ with $\rho(H) = \rho^*(G) + O(\log^3 n/n)$, where $\widetilde{k} = \widetilde{k}(c)$ is defined as in Corollary 2.5. Moreover, a.a.s. there exists a densest subgraph G' of G, such that $|V(H) \Delta V(G')| \leq \log^2 n$.*

By Corollary 2.5, either $G_{\widetilde{k}}$ or $G_{\widetilde{k}+1}$ may be the densest core, even though it is not clear which one is denser between these two cores. The following corollary, which follows directly from Theorems 2.6 and 2.3, gives bounds of $\rho^*(\mathcal{G}(n, c/n))$. As we commented in the introduction, these bounds are not new.

Corollary 2.7 *For all sufficiently large $c \in \mathbf{R}_0$, a.a.s.*

$$\max\left\{\frac{\beta_{c,\widetilde{k}}}{\alpha_{c,\widetilde{k}}}, \frac{\beta_{c,\widetilde{k}+1}}{\alpha_{c,\widetilde{k}+1}}\right\} \leq \rho^*(\mathcal{G}(n, c/n)) \leq \widetilde{k}.$$

Note that the above bounds are not tight. We conjecture that $\rho^*(\mathcal{G}(n, c/n))$ is concentrated and has a limit as $n \to \infty$.

Conjecture 2.8 *There exists $\rho^* = \rho^*(c)$ such that a.a.s. $\rho^*(\mathcal{G}(n, c/n)) = \rho^* + o(1)$.*

In the proof of Theorems 2.4 and 2.6, we consider only large k and c to simplify the analysis. However, we conjecture that the same conclusions hold for small k and c as well.

Conjecture 2.9 *Theorem 2.4 holds for all $k \geq 2$ and Theorem 2.6 holds for all $c \in \mathbf{R}_0$ with $c > c_3$.*

3 Generating a random core

Given three positive integers n, m and k with $2m \geq kn$, let $\mathcal{P}(n, m, k)$ denote the probability space of equiprobable functions $f : [2m] \to [n]$ such that $|f^{-1}(i)| \geq k$ for every $i \in [n]$. Let $\mathcal{M}(n, m, k)$ denote the random multigraph generated by representing each element $i \in [n]$ in $\mathcal{P}(n, m, k)$ as a vertex and each $\{f(2i-1), f(2i)\}$ as an edge for $1 \leq i \leq m$. We say $\mathcal{M}(n, m, k)$ is generated from $\mathcal{P}(n, m, k)$. An alternative way to generate $\mathcal{M}(n, m, k)$ is to throw $2m$ balls uniformly at random into n bins, conditioned on that each bin receives at least k balls. Take a uniform pairing of all balls. Represent each bin as a vertex and each pair of balls as an edge. The latter way of generating $\mathcal{M}(n, m, k)$ is usually referred to as the allocation-pairing model [4], a variation of the configuration model, first introduced by Bollobás [1].

Let $\mathcal{H}(n, m, k)$ denote the probability space of $\mathcal{M}(n, m, k)$ conditioning on graphs being simple. The following result was proved by Cain and Wormald [4].

Lemma 3.1 *Let $G \sim \mathcal{G}(n, c/n)$. Then for any $k \geq 0$, conditional on the number of vertices and edges of G_k being n' and m', we have $G_k \sim \mathcal{H}(n', m', k)$.*

By Lemma 3.1, we can analyse G_k by analysing $\mathcal{H}(n, m, k)$. Note that $\mathcal{H}(n, m, k)$ is $\mathcal{M}(n, m, k)$ conditioned to simple graphs. It is relatively much easier to compute probabilities of events in $\mathcal{M}(n, m, k)$ than in $\mathcal{H}(n, m, k)$. The following results allow us to translate properties of $\mathcal{M}(n, m, k)$ to $\mathcal{H}(n, m, k)$.

Proposition 3.2 *Let $k \geq 0$ be a fixed integer. Assume $m = O(n)$ and $2m \geq kn$. Then the probability that a graph in $\mathcal{M}(n, m, k)$ is simple is $\Omega(1)$.*

This proposition follows from [14, Theorem 3], which is an analog to the well known fact that the probability that a graph in $\mathcal{M}(n, m, 0)$ is simple is $\Omega(1)$, proved by Chvátal [2]. Proposition 3.2 immediately yields the following useful corollary.

Corollary 3.3 *Let $k \geq 0$ be a fixed integer. Assume $m = O(n)$ and $2m \geq kn$. Then for any event A_n such that $\mathbf{P}_{\mathcal{M}(n,m,k)}(A_n) = 1 - o(1)$, we have $\mathbf{P}_{\mathcal{H}(n,m,k)}(A_n) = 1 - o(1)$.*

We will use the following result from [4] to analyse the degree sequence of $\mathcal{M}(n, m, k)$.

Lemma 3.4 *Let $j \geq k$ be fixed and let X_j denote the number of vertices with degree j in $\mathcal{M}(n,m,k)$, where $2m/n > k$. Then a.a.s.*

$$X_j = p_{k,2m/n}(j)n + o(n),$$

where

$$p_{k,b}(j) = e^{-\lambda_b}\frac{\lambda_b^j}{f_k(\lambda_b)j!},$$

where $b > k$ and λ_b is the positive root λ of $\lambda f_{k-1}(\lambda)/f_k(\lambda) = b$.

The following corollary follows immediately.

Corollary 3.5 *Let $j \geq k$ be fixed and let X_j denote the number of vertices with degree j in $\mathcal{M}(n,m,k)$, where $2m/n > k$. Then a.a.s. $X_j = \Theta(n)$.*

4 Proof of Theorem 2.4

For a vertex set S, let $d(S)$ denote the degree sum of vertices in S. Let $t(S)$ denote the number of edges contained in S, or the number of pairs with both end-points contained in S, if the allocation-pairing model is under consideration.

Recall the definition of $\alpha_{c,k}$ and $\beta_{c,k}$ in (2.1) and recall that $\widetilde{c}_k$ denotes the unique c satisfying $\beta_{c,k}/\alpha_{c,k} = k-1$. Theorem 2.4 follows immediately from the following two theorems.

Theorem 4.1 *Let k be a positive integer and assume $c < \widetilde{c}_k$. Then a.a.s. $G_k \subseteq G \sim \mathcal{G}(n, c/n)$ is balanced provided k is sufficiently large.*

Theorem 4.2 *Let k be a positive integer and assume $c > \widetilde{c}_k$. Then a.a.s. $G_k \subseteq G \sim \mathcal{G}(n, c/n)$ is not balanced provided k is sufficiently large. Moreover, there exists $H \subseteq G_k$ which is significantly denser than G_k.*

We first prove Theorem 4.1. Consider the allocation-pairing model that generates $\mathcal{M}(n, m(n), k)$. Let $p(q,t)$ denote the probability that there are at least t pairs among a given set of q points in $\mathcal{M}(n, m(n), k)$. A special case of [9, Claim 5.9] gives the following lemma. As the proof is short, we include this proof.

Lemma 4.3

$$p(q,t) \leq \begin{cases} \big(eq(q-t)/2t(2m-1)\big)^t & \text{if } (q-t)/(2m-1) \leq 1/2 \\ \big(eq(q-2t+1)/2t(2m-2t+1)\big)^t & \text{if } (q-t)/(2m-1) > 1/2. \end{cases}$$

Proof. Given a set of q points, there are at most

$$\binom{q}{2t}\frac{(2t)!}{2^t t!}$$

ways to locate t pairs within the given q points. Such t pairs form a partial pairing of the $2m$ points. The probability that a given partial pairing with t pairs occurs is

$$\prod_{i=0}^{t-1} \frac{1}{2m-1-2i}.$$

So

$$p(q,t) \le \binom{q}{2t} \frac{(2t)!}{2^t t!} \cdot \prod_{i=0}^{t-1} \frac{1}{2m-1-2i},$$

which is at most

$$\frac{\prod_{i=0}^{t-1}(q-i)}{2^t t!} \prod_{i=0}^{t-1} \frac{q-t-i}{2m-1-2i}$$
$$\le \begin{cases} \big(eq(q-t)/2t(2m-1)\big)^t & \text{if } (q-t)/(2m-1) \le 1/2 \\ \big(eq(q-2t+1)/2t(2m-2t+1)\big)^t & \text{if } (q-t)/(2m-1) > 1/2. \end{cases}$$ ∎

The following lemma was proved in [9].

Lemma 4.4 *([9, Corollary 5.4]) Let $G \sim \mathcal{M}(n, m(n), k)$ with $2m \ge kn$. For any $0 < \delta < 1$, there exists $N(\delta) > 0$ and $0 < \alpha(\delta) < 1$ such that provided $k > N$, for any $S \subseteq G$ with $|S| \ge \log^2 n$,*

$$\mathbf{P}(|d(S) - 2\rho s| \ge \delta 2\rho s) < \alpha^{\rho s},$$

where $\rho = m(n)/n$.

Next, we prove that a.a.s. all small subgraphs of $G(n, c/n)$ are sparse.

Lemma 4.5 *For any fixed $k \ge 0$ and $c > 0$, there exists $\epsilon_c = 8/e^3c^2$ such that a.a.s. there is no $S \subseteq G \sim \mathcal{G}(n, c/n)$ with $|S| \le \epsilon_c n$ and $t(S) > 2|S|$.*

Proof. The expected number of S with $|S| = s$ and $t(S) > 2s$ is at most

$$\binom{n}{s}\binom{s^2/2}{2s}(c/n)^{2s} \le \left(\frac{en}{s}\right)^s \left(\frac{esc}{4n}\right)^{2s} = \left(\frac{e^3c^2}{16} \cdot \frac{s}{n}\right)^s.$$

Let $\epsilon_c = 8/e^3c^2$. Then $e^3c^2\epsilon_c/16 = 1/2$ and thus the expected number of S with $|S| \le \epsilon_c n$ and $t(S) > 2|S|$ is at most

$$\sum_{s \le \ln n} \frac{e^3c^2}{16} \cdot \frac{\ln n}{n} + \sum_{\ln n < s \le \epsilon_c n} 2^{-s} = o(1).$$ ∎

Lemma 4.6 *For any $c > 0$, the maximum degree in $\mathcal{G}(n, c/n)$ is a.a.s. $O(\log n)$.*

Proof. The expected number of vertices in $\mathcal{G}(n, c/n)$ with degree at least $\ln n$ is at most

$$n\binom{n}{\ln n}\left(\frac{c}{n}\right)^{\ln n} \leq n\left(\frac{en}{\log n}\cdot\frac{c}{n}\right)^{\ln n} \leq \exp\left(\ln n - \ln n(\ln\ln n - \ln(ce))\right) = o(1). \blacksquare$$

Let $S \subseteq G$ be a vertex set. Recall that $d(S)$ denote the sum of degrees of vertices of S in G.

Lemma 4.7 *Let k, n and $m(n)$ be positive integers such that $2m \geq kn$ and $m/n = O(1)$. Let $\epsilon_0 > 0$ be a fixed constant. Then, a.a.s. there exists no $S \subseteq G \sim \mathcal{H}(n, m(n), k)$ with $\epsilon_0 n \leq |S| < (1.8/e)n$ and $t(S) > \rho(G)|S|$, provided k is sufficiently large.*

Proof. Let $\rho = \rho(G)$. Consider the allocation-pairing model that generates $\mathcal{M}(n, m, k)$. Given s, the expected number of $S \subseteq G \sim \mathcal{M}(n, m, k)$ with $|S| = s$ and $t(S) \geq \rho s$ is

$$\sum_q \binom{n}{s}\mathbf{P}(d(S) = q)p(q, \rho s) \leq \sum_q \mathbf{P}(d(S) = q)\left(\frac{en}{s}\right)^s p(q, \rho s). \tag{4.1}$$

Let $r_1 = 1.8/e$. Since $\rho \geq k$, we may assume that k is sufficiently large so that $(e/r_1)^{1/\rho} \leq 9.5/9$. Let $\delta > 0$ be $\sup\{\delta :\ (1+\delta)(1+2\delta) < 99/95\}$. By Lemma 4.4, there exists $N(\delta) > 0$ and $0 < \alpha < 1$ such that provided $k > N$ (therefore, ρ is sufficiently large),

$$\mathbf{P}(\exists S, |S| \geq \epsilon_0 n, d(S) \geq (1+\delta)2\rho s) \leq \sum_{s \geq \epsilon_0 n}\binom{n}{s}\alpha^{\rho s} \leq \sum_{s \geq \epsilon_0 n}\left(e\epsilon_0^{-1}\alpha^{\rho}\right)^s = o(1). \tag{4.2}$$

By (4.1) and (4.2), the expected number of S with $\epsilon_0 n \leq |S| \leq r_1 n$ and $t(S) > \rho|S|$, denoted by Y, is at most

$$\mathbf{E}(Y) \leq o(1) + \sum_{\epsilon_0 n \leq s \leq r_1 n}\ \sum_{q \leq (1+\delta)2\rho s}\mathbf{P}(d(S) = q)\left(\frac{en}{s}\right)^s p(q, \rho s). \tag{4.3}$$

Since for any $q \leq (1+\delta)2\rho s$, $p(q, \rho s) \leq p((1+\delta)2\rho s, \rho s)$ as $p(q, t)$ is a non-decreasing function of q, (4.3) is at most

$$o(1) + \sum_{\epsilon_0 n \leq s \leq r_1 n}\left(\frac{en}{s}\right)^s p((1+\delta)2\rho s, \rho s),$$

which is at most

$$o(1) + \sum_{\epsilon_0 n \leq s \leq r_1 n}\left(\frac{en}{s}\right)^s\left(\frac{e(1+\delta)2\rho s((1+\delta)2\rho s - \rho s)}{2\rho s \cdot (2\rho n - 1)}\right)^{\rho s},$$

by Lemma 4.3 (Note that $(q-t)/(2m-1) \leq 1/2$ by taking $q = (1+\delta)2\rho s$ and $t = \rho s$.). Let $r = s/n$. The above is

$$o(1) + \sum_{\epsilon_0 n \leq s \leq r_1 n}\left(\left(\frac{e}{r}\right)^{1/\rho}\frac{e(1+\delta)(1+2\delta)r}{2(1-1/2\rho n)}\right)^{\rho s}.$$

By our choice of r_1 and δ,

$$\left(\frac{e}{r}\right)^{1/\rho}\frac{e(1+\delta)(1+2\delta)r}{2}\leq\left(\frac{e}{r_1}\right)^{1/\rho}\frac{er_1}{2}(1+\delta)(1+2\delta)\leq\frac{9.5}{9}\cdot\frac{1.8}{2}\cdot\frac{99}{95}\leq 0.99,$$

and so $\mathbf{E}Y=o(1)$. By Corollary 3.3, the expected number of $S\subseteq G\sim\mathcal{H}(n,m(n),k)$ with $\epsilon_0 n\leq|S|<r_1 n$ and $t(S)>\rho|S|$ is $o(1)$. ∎

Lemma 4.8 *Assume $\delta_2>\delta_1>0$ are absolute constants which do not depend on k. Let k, n and $m(n)$ be positive integers with $2m\geq kn$ and $\delta_1<k-1-m/n<\delta_2$. Then a.a.s. there exists no $S\subseteq G\sim\mathcal{M}(n,m,k)$ with $\log^2 n\leq|S|\leq n/2$ and $t(S)\geq d(S)-(k-1)|S|$, provided k is sufficiently large.*

Proof. Let $\rho=m/n$ and let $\delta=\delta(n)=k-1-\rho$. Then, $\rho=k-1-\delta$ and $\delta_1<\delta<\delta_2$. The expected number of S with $|S|=s$ and $t(S)\geq d(S)-(k-1)|S|$, denoted by Y_s, is

$$\mathbf{E}Y_s=\binom{n}{s}\sum_q\mathbf{P}(d(S)=q)p(q,q-(k-1)s).$$

Let $r=s/n$. Since $r\leq 1/2$ and thus

$$\frac{q-(q-(k-1)s)}{2m-1}=\frac{(k-1)s}{2m-1}\leq\frac{(k-1)s}{kn-1}\leq 1/2,$$

we have that

$$\mathbf{E}Y_s\leq\sum_q\mathbf{P}(d(S)=q)\left(\frac{e}{r}\right)^s\left(\frac{eq\cdot(k-1)s}{2(q-(k-1)s)(2\rho n-1)}\right)^{q-(k-1)s}. \tag{4.4}$$

by Lemma 4.3. Let

$$\delta'=\sup\left\{x:\ \frac{(1-x)2\rho}{k-1}>1+5e/18\approx 1.755\right\}.$$

Since $\rho=k-1-\delta>k-1-\delta_2$, we have $0<\delta'<1$, provided k sufficiently large. Let

$$S_1=\sum_{q<(1-\delta')2\rho s}\mathbf{P}(d(S)=q)\left(\frac{e}{r}\right)^s\left(\frac{eq\cdot(k-1)s}{2(q-(k-1)s)(2\rho n-1)}\right)^{q-(k-1)s},$$

$$S_2=\sum_{q\geq(1-\delta')2\rho s}\mathbf{P}(d(S)=q)\left(\frac{e}{r}\right)^s\left(\frac{eq\cdot(k-1)s}{2(q-(k-1)s)(2\rho n-1)}\right)^{q-(k-1)s}.$$

Then

$$\mathbf{E}Y_s\leq S_1+S_2. \tag{4.5}$$

We first estimate S_2. Clearly,

$$\left(\frac{eq\cdot(k-1)s}{2(q-(k-1)s)(2\rho n-1)}\right)^{q-(k-1)s}$$

is a decreasing function of q. Hence,

$$\begin{aligned}
S_2 &\le \left(\frac{e}{r}\right)^s \left(\frac{e(1-\delta')2\rho s\cdot(k-1)s}{2((1-\delta')2\rho s-(k-1)s)(2\rho n-1)}\right)^{(1-\delta')2\rho s-(k-1)s} \\
&= \left(\frac{e}{r}\right)^s \left(\frac{e(1-\delta')r}{((1-\delta')2\rho/(k-1)-1)(2-1/\rho n)}\right)^{(1-\delta')2\rho s-(k-1)s} \\
&\le \left(\frac{e}{r}\right)^s \left(\frac{er}{(5e/18)\cdot 2}\right)^{(k-1)s\cdot 5e/18} = \left((e/r)^{1/C}(9r/5)\right)^s,
\end{aligned}$$

where $C = 5e(k-1)/18$, by the choice of δ'. Clearly, $(e/r)^{1/C}(9r/5)$ is an increasing function of r and $r \le r_1 := 1/2$. We may assume that k is sufficiently large so that $(e/r_1)^{1/C} < 9.5/9$. Hence,

$$S_2 \le ((e/r_1)^{1/C}(9r_1/5))^s < \left(\frac{9.5}{9}\cdot\frac{9}{10}\right)^s \le \left(\frac{9.5}{10}\right)^s. \tag{4.6}$$

Next, we estimate S_1. By Lemma 4.4, there exists $0<\alpha<1$ such that

$$\mathbf{P}(d(S) < (1-\delta')2\rho s) \le \alpha^{\rho s}.$$

Since all vertices in G has degree at least k, we have $q \ge ks$. Thus,

$$\begin{aligned}
S_1 &\le \alpha^{\rho s}\left(\frac{e}{r}\right)^s \left(\frac{eks\cdot(k-1)s}{2(ks-(k-1)s)(2\rho n-1)}\right)^{ks-(k-1)s} \\
&= \alpha^{\rho s}\left(\frac{e}{r}\right)^s \left(\frac{ek\cdot(k-1)r}{2\rho(2-1/\rho n)}\right)^s \le \left(\alpha^\rho\frac{e^2k^2}{4\rho}\right)^s.
\end{aligned}$$

Since $\rho = k-1-\delta > k-1-\delta_2$, we may assume that k is sufficiently large so that

$$\alpha^\rho\frac{e^2k^2}{4\rho} < 1/2.$$

Then,

$$S_1 \le 2^{-s}. \tag{4.7}$$

By (4.5), (4.6) and (4.7), we have

$$\mathbf{E}Y_s \le 0.95^s + 0.5^s.$$

So, the expected number of S with $\log^2 n \le |S| \le r_1 n$ and $t(S) \ge d(S)-(k-1)|S|$ is

$$\sum_{s=\log^2 n}^{r_1 n} \mathbf{E}Y_s = o(1). \blacksquare$$

The condition $k-1-m/n < \delta_2$ can be removed if we replace $t(S) \ge d(S)-(k-1)|S|$ by $t(S) \ge d(S)-\rho(G)|S|$ and restrict to slightly smaller S in the above lemma. This immediately gives the following, with almost the same proof as Lemma 4.8.

Lemma 4.9 *Let k, n and $m(n)$ be positive integers with $2m \geq kn$ and $\limsup_{n\to\infty} m(n)/n < k-1$. Let $r_1 = 1 - 1.78/e$. Then a.a.s. there exists no $S \subseteq G \sim \mathcal{M}(n, m(n), k)$ with $\log^2 n \leq |S| \leq r_1 n$ and $t(S) \geq d(S) - \rho(G)|S|$, provided k is sufficiently large.*

Proof. We follow almost the same proof as in Lemma 4.8. Then (4.8) becomes

$$\mathbf{E}Y_s \leq \sum_q \mathbf{P}(d(S) = q) \left(\frac{e}{r}\right)^s \left(\frac{eq \cdot \rho s}{2(q - \rho s)(2\rho n - 1)}\right)^{q-\rho s}.$$

Define δ' to be

$$\delta' := \sup\{x : \frac{1-x}{1-2x} < \frac{1}{e - 1.75} \approx 1.032\}.$$

Then $0 < \delta' < 1/2$. Define S_1 and S_2 and bound S_1 in the same way as in Lemma 4.8, using the fact that $k/2 \leq \rho < k-1$. Now it only remains to bound S_2. Following the same calculation in Lemma 4.8, but using $\rho \geq k/2$ rather than $\rho > k - 1 - \delta_2$, we have

$$\begin{aligned} S_2 &\leq \left(\frac{e}{r}\right)^s \left(\frac{e(1-\delta')(k-1)r}{((1-\delta')2\rho - \rho)(2 - 1/\rho n)}\right)^{(1-\delta')2\rho s - \rho s} \\ &\leq \left(\frac{e}{r}\right)^s \left(\frac{e(1-\delta')r}{1-2\delta'}\right)^{(1-2\delta')\rho s} = \left((e/r)^{1/C_1}(C_2 r)\right)^s, \end{aligned}$$

where $C_1 = (1-2\delta')\rho \geq (1/2 - \delta')k$ and $C_2 = e/(e - 1.75)$. Comparing with r_1, we have $C_2 r_1 < \gamma$ for some constant $\gamma < 1$. We may assume that k is sufficiently large so that $(e/r_1)^{1/C_1} < 1/\sqrt{\gamma}$. Since S_2 is an increasing function of r, we have

$$S_2 \leq \left((e/r_1)^{1/C_1}(C_2 r_1)\right)^s < \gamma^{s/2}.$$

Then we can complete the proof of Lemma 4.9 as in Lemma 4.8. ∎

We will use Lemma 4.9 in the proof of Theorem 4.1 and use Lemma 4.8 in the proof of Theorem 2.6.

Proof of Theorem 4.1. Since $c < \widetilde{c}_{k+1}$, by Theorem 2.1, there exist an absolute constant $\delta > 0$ such that a.a.s. the density of $G_{k+1} \subseteq G \sim G(n, c/n)$ is at most $k - \delta$. Let $\epsilon_c = 8/e^3c^2$. By Lemma 4.5, a.a.s. there is no subgraph of G whose size is at most $\epsilon_c n$ and whose density is at least 2. Let n' denote the number of vertices in G_{k+1}. Then $n' \leq n$. It follows that a.a.s. there is no subgraph of G_{k+1} whose size is at most $\epsilon_c n'$ and whose density is at least 2.

We first prove that if G_{k+1} is not balanced, then the following statements are true.

(a) There exists $S \subseteq G_{k+1}$ with $t(S) > \rho(G_{k+1})|S|$;

(b) There exists $S \subseteq G_{k+1}$ with $t(\overline{S}) > d(\overline{S}) - \rho(G_{k+1})|\overline{S}|$, where $\overline{S} = V(G_{k+1}) \setminus S$ and $V(G_{k+1})$ denotes the vertex set of G_{k+1}.

Statement (a) is obvious. We show that (a) and (b) are indeed equivalent. Let $e(S, \overline{S})$ denote the number of edges between S and $\overline{S}$ and let $\rho = \rho(G_{k+1})$. Then we have

$$\begin{aligned} \rho|V(G_{k+1})| &= t(S) + t(\overline{S}) + e(S, \overline{S}), \\ d(S) &= e(S, \overline{S}) + 2t(S) \\ 2\rho|V(G_{k+1})| &= d(S) + d(\overline{S}). \end{aligned}$$

Hence,

$$\rho|V(G_{k+1})| = t(S)+t(\overline{S})+\big(d(S)-2t(S)\big) = t(\overline{S})+d(S)-t(S) = t(\overline{S})+\big(2\rho|V(G_{k+1})|-d(\overline{S})\big)-t(S),$$

i.e.

$$t(S) = t(\overline{S}) + \rho|V(G_{k+1})| - d(\overline{S}).$$

Thus,

$$t(S) > \rho|S|$$

if and only if

$$t(\overline{S}) + \rho|V(G_{k+1})| - d(\overline{S}) > \rho|S|,$$

i.e.

$$t(\overline{S}) > d(\overline{S}) - \rho(|V(G_{k+1})| - |S|) = d(\overline{S}) - \rho|\overline{S}|.$$

Next, we show that a.a.s. any $S \subseteq G_{k+1}$ will violates either (a) or (b) and then it follows that a.a.s. G_{k+1} is balanced. Since a.a.s. any subgraph of G_{k+1} whose size is at most $\epsilon_c n$ has density at most 2, it follows immediately that a.a.s. there is no set S with size at most $\epsilon_c n'$ that satisfies (a), provided $k \geq 3$. By Lemma 3.1, by conditioning on the values of n', the number of vertices and m', the number of edges in G_{k+1}, we have $G_{k+1} \sim \mathcal{H}(n', m', k+1)$. Consider any n' and m' such that $k - \delta_0 < m'/n' \leq k - \delta$. Let $r_1 = 1.8/e$ and $r_2 = 1.78/e$. Then $r_1 > r_2$.

By Lemma 4.7, the expected number of S with $\epsilon_c n' \leq |S| < r_1 n'$ that satisfies (a) is $o(1)$. By Lemma 4.9 and Corollary 3.3, the expected number of $S \subseteq G \sim \mathcal{H}(n', m', k+1)$ with $r_2 n' < |S| \leq n' - \log^2 n'$ that satisfies (b) is $o(1)$. Thus, it only remains to show that the expected number of $S \subseteq G \sim \mathcal{H}(n', m', k+1)$ with $|S| > n' - \log^2 n'$ that satisfies (b) is $o(1)$, and then the proof of Theorem 4.1 is complete.

Consider $G_{k+1} \in \mathcal{M}(n', m', k+1)$, where $m'/n' < k - \delta$ and let $\rho = \rho(G_{k+1}) = m'/n'$. The expected number of S with $|S| = s$ that satisfies (b), denoted by Y_s, is at most

$$\mathbf{E}Y_s = \binom{n'}{\overline{s}} \sum_{\overline{q}} \mathbf{P}(d(\overline{S}) = \overline{q}) p(\overline{q}, \overline{q} - \rho\overline{s}),$$

where $\overline{s} = n' - s$. Let $\overline{r} = 1 - r$. The above is at most

$$\sum_{\overline{q}} \mathbf{P}(d(\overline{S}) = \overline{q}) \left(\frac{e}{\overline{r}}\right)^{\overline{s}} \left(\frac{e\overline{q} \cdot \overline{r}}{2(\overline{q} - \rho\overline{s})(2 - 1/\rho n')}\right)^{\overline{q} - \rho\overline{s}}. \qquad (4.8)$$

Hence the expected number of S with $|S| > n' - \log^2 n'$ which satisfies (b) is $\sum_{s > n' - \log^2 n'} \mathbf{E}Y_s$. Since in $\mathcal{M}(n', m', k+1)$, we always have $\overline{q} \geq (k+1)\overline{s}$ and we have $\rho \leq k - \delta$ by assumption, it follows that $\overline{q} - \rho\overline{s} \geq (1 + \delta)\overline{s}$. By (4.8) and the fact that

$$\left(\frac{e\overline{q} \cdot \overline{r}}{2(\overline{q} - \rho\overline{s})(2 - 1/\rho n')}\right)^{\overline{q} - \rho\overline{s}}$$

is a decreasing function of $\overline{q}$, we obtain

$$\begin{aligned}\sum_{\overline{s}<\log^2 n'} \mathbf{E}Y_s &\leq \sum_{\overline{s}<\log^2 n'} \left(\frac{en'}{\overline{s}}\right)^{\overline{s}} \left(\frac{e(k+1)}{2(1+\delta)(2-1/\rho n')}\cdot\frac{\overline{s}}{n'}\right)^{(1+\delta)\overline{s}} \\ &= \sum_{\overline{s}<\log^2 n'} \left(e^{2+\delta}\left(\frac{(k+1)}{2(1+\delta)(2-1/\rho n')}\right)^{1+\delta}\cdot\left(\frac{\log^2 n'}{n'}\right)^{\delta}\right)^{\overline{s}} = o(1),\end{aligned}$$

where the upper bound is obtained by taking $\overline{q} = (k+1)\overline{s}$. Our claim then follows by Corollary 3.3. ∎

Next, we prove Theorem 4.2. Let S_{k+1} denote the set of vertices with degree $k+1$ in $G \sim \mathcal{M}(n, m(n), k+1)$ and let $G_{[S_{k+1}]}$ denote the graph induced by S_{k+1}.

Lemma 4.10 *Let $k > 0$ be a positive integer. Assume $m(n) = \Theta(n)$. Then a.a.s. for any constant integer $\ell > 0$, there are $\Theta(n)$ vertex disjoint paths in $G_{[S_{k+1}]}$, each with length ℓ, provided k is sufficiently large.*

Proof. Consider the allocation-pairing model that generates $\mathcal{M}(n, m, k)$. Let $n_{k+1} = |S_{k+1}|$. By Corollary 3.5, $n_{k+1} = \Theta(n)$. For any fixed integer $\ell > 0$, let Z_ℓ denote the number of paths of length ℓ in $G_{[S_{k+1}]}$. Then

$$\begin{aligned}\mathbf{E}Z_\ell &= \binom{n_{k+1}}{\ell+1}\frac{(\ell+1)!}{2}(k+1)^2((k+1)k)^{\ell-2}\prod_{i=0}^{\ell-1}\frac{1}{2m-2i-1} \\ &\sim \frac{n_{k+1}^{\ell+1}}{2}\frac{(k+1)^\ell k^{\ell-2}}{(2m)^\ell} = \Theta(n_{k+1}) = \Theta(n).\end{aligned}$$

Similarly, the expected number of ordered pairs of vertex disjoint paths of length ℓ in $G_{[S_{k+1}]}$ is asymptotically $(\mathbf{E}Z_\ell)^2 = \Theta(n^2)$. Let Y denote the number of subgraphs contained in $G_{[S_{k+1}]}$ for which there are at most $2\ell+1$ vertices, and the maximum degree is at most 4, and the number of edges is at least the number of vertices minus one. Clearly, the number of pairs of paths of length ℓ that share at least one vertex is at most Y. Since

$$\mathbf{E}Y \leq \sum_{i\leq 2\ell+1}\binom{n_{k+1}}{i}\binom{\binom{i}{2}}{i-1}k^{4i}\prod_{j=0}^{i-1}\frac{1}{2m-2j-1} = O(n_{k+1}) = o((\mathbf{E}Z_\ell)^2),$$

where $\binom{\binom{i}{2}}{i-1}$ is an upper bound of the number of ways to locate the $i-1$ edges, and k^{4i} upper bounds the number of ways to pick the end points of a given set of $i-1$ edges, by the method of the second moment, we immediately have that a.a.s. for any fixed $\ell > 0$, the number of paths in $G_{[S_{k+1}]}$ with length ℓ is $\Theta(n)$. Since each vertex in $G_{[S_{k+1}]}$ has a bounded degree ($\leq k+1$), and each path under discussion has bounded length ($\leq \ell$), each of these paths is vertex disjoint with all but at most constant number of other paths. Therefore, there are $\Theta(n)$ vertex disjoint paths of length ℓ in $G_{[S_{k+1}]}$. ∎

Proof of Theorem 4.2. By Corollary 3.3, we only need to show that for any fixed $\delta_0 > 0$, if $G_{k+1} \in \mathcal{M}(n, m(n), k+1)$ with $m(n) \geq (k + \delta_0)n$, then a.a.s. there exists a subgraph of G_{k+1} that is significantly denser than G_{k+1}. Note that this is a stronger conclusion than G_{k+1} being not balanced. Let $\delta = \delta_n$ be that $m = (k + \delta)n$. Then $\delta \geq \delta_0$. By Lemma 4.10, a.a.s. there exists a set $\mathcal{P} = \{p_1, \ldots, p_s\}$ of vertex disjoint paths in $G_{[S_{k+1}]} \subseteq G_{k+1}$, each of length $\ell = 2/\delta \leq 2/\delta_0$, where $s = \Theta(n)$. Let G' be obtained from G_{k+1} by removing all vertices in $\mathcal{P}$. Then

$$\rho(G') \geq \frac{m - ((k+1)(\ell+1) - \ell)s}{n - (\ell+1)s}.$$

Since $\ell = 2/\delta$, we have

$$\frac{(k+1)(\ell+1) - \ell}{\ell+1} = k + \frac{1}{\ell+1} < k + \frac{\delta}{2} = \frac{m}{n} - \delta/2 \leq \frac{m}{n} - \delta_0/2.$$

Together with $s = \Omega(n)$, we have that there exists an absolute constant $\eta > 0$ such that $\rho(G') > \rho(G_{k+1}) + \eta$. ∎

5 Proof of Theorem 2.3

In the rest of the paper, let $\widetilde{k} = \widetilde{k}(c)$ and $\widetilde{d} = \widetilde{d}(c)$ be as stated in Corollary 2.5. The following proposition follows from Theorems 2.4 and 2.1.

Proposition 5.1

(a) $\widetilde{k}(c)$ *is a monotone increasing function of* c*;*

(b) For all $k \leq \widetilde{k}(c)$, $|G_{k+1}|/|G_k| \to 1$ *as* $c \to \infty$.

Note that Theorem 2.3 holds if there exists a constant $\eta > 0$ such that the following statements are true.

(a) a.a.s. $\rho^*(G) > \rho(G_k) + \eta$, for any integer $0 \leq k \leq \widetilde{k}$;

(b) a.a.s. $\rho^*(G) > \rho(G_{\widetilde{k}+1}) + \eta$;

(c) a.a.s. $\rho^*(G) > \rho(G_k) + \eta$, for all $k > \widetilde{k} + 1$,

where $G \sim \mathcal{G}(n, c/n)$.

By the choice of $\widetilde{k}$ and Theorem 4.2, there exists $\eta' > 0$ such that for all $0 \leq k \leq \widetilde{k}$, a.a.s. there exists a subgraph of G_k whose density is greater than $\rho(G_k) + \eta'$, which implies (a). If (b) is true, then (c) will be true because $G_{\widetilde{k}+1}$ is balanced. Therefore, we only need to prove (b).

Let $\delta > 0$ be a constant chosen such that a.a.s. $\rho(G_{\widetilde{k}+1}) < \widetilde{k} - \delta$ and $\rho(G_{\widetilde{k}}) > \widetilde{k} - 1 + \delta$. (Note that by the choice of $\widetilde{k}$, such δ exists.) If $\rho(G_{\widetilde{k}+1}) \leq \widetilde{k} - 1 + \delta$, then (b) follows by noting that $\rho^*(G) > \rho(G_{\widetilde{k}}) + \eta' > \rho(G_{\widetilde{k}+1}) + \eta'$. Hence, we may assume that $\rho(G_{\widetilde{k}+1}) > \widetilde{k} - 1 + \delta$.

Consider the following vertex deletion algorithm. Let G be a graph with minimum degree at least k. At step 0, let $G_0 = G$. At each step $t \geq 1$, choose randomly a vertex v with degree k from G_{t-1}, remove v and all its incident edges from G_{t-1}. If there is some vertex with degree less than k being produced, repeatedly remove all these vertices and their incident edges until either there are no vertices with degree less than k or $n^{1/3}$ vertices has already been removed in this single step. Let G_t be the resulting graph. The algorithm terminates when there is no vertex with degree k. Let τ denote the stopping time.

We will show that a.a.s. the algorithm never removes $n^{1/3}$ vertices repeatedly in each step, and therefore, a.a.s. the algorithm outputs the $(k+1)$-core of G_0.

The following proposition has qualitatively the same nature as Theorem 2.1 and indeed it can be proved with almost the same argument in Theorem 2.1 by analysing the edge-deletion algorithm, defined in [12], applied to $G_0 \sim \mathcal{H}(n, m, k)$ (instead of $\mathcal{H}(n, m, 0) = \mathcal{G}(n, m)$).

Proposition 5.2 *Given n, $m = O(n)$ and k with $m/n > k - 1$, assume $G_0 \sim \mathcal{H}(n, m, k)$. Let n' and m' denote the number of vertices and edges in G_τ. Assume the minimum degree of G_τ is at least $k + 1$. Then there exist absolute constants α and β, such that a.a.s. $n' = \alpha n + o(n)$, $m' = \beta m + o(m)$. Moreover, if we condition on the values of n' and m', then $G_\tau \sim \mathcal{H}(n', m', k + 1)$.*

We complete the proof of Theorem 2.3 using the following lemma.

Lemma 5.3 *Let n, m and $\delta > 0$ be such that $\alpha > 0$, $\beta > 0$ and $k - 1 - \delta < \beta m/\alpha n < k - \delta$, where α and β are constants determined by n, m and k as in Proposition 5.2. Assume $G_0 \sim \mathcal{H}(n, m, k)$. Then, a.a.s. G_τ is the $(k + 1)$-core of G_0 and there exists a constant $\eta > 0$ and $G_\tau \subseteq G \subseteq G_0$ such that $\rho(G) \geq \rho(G_\tau) + \eta$.*

By the previous argument, there exists $\delta > 0$ such that a.a.s. $\rho(G_{\widetilde{k}+1}) < \widetilde{k} - \delta$, $\rho(G_{\widetilde{k}}) > \widetilde{k} - 1 + \delta$ and $\rho(G_{\widetilde{k}+1}) > \widetilde{k} - 1 + \delta$. Conditional on the number of vertices and the number of edges in $G_{\widetilde{k}}$ being n_1 and m_1, we have $G_{\widetilde{k}} \sim \mathcal{H}(n_1, m_1, \widetilde{k})$. Let n_2 and m_2 denote the numbers of vertices and edges in $G_{\widetilde{k}+1}$. By Proposition 5.2, there exists α and β such that a.a.s. $n_2 = \alpha n_1 + o(n_1)$, $m_2 = \beta m_1 + o(n_1)$ and $\widetilde{k} - 1 + \delta < \beta m_1/\alpha n_1 < \widetilde{k} - \delta$. Conditional on the values of n_2 and m_2, $G_{\widetilde{k}+1} \sim \mathcal{H}(n_2, m_2, \widetilde{k} + 1)$. Apply Lemma 5.3 with $k = \widetilde{k}$, $n = n_1$ and $m = m_1$. Theorem 2.3 follows. ∎

It only remains to prove Lemma 5.3. We first show a technical lemma.

Lemma 5.4 *Let δ be a constant. Let $(c_t)_{t\geq 1}$ be positive reals such that $c_t^2 = o(t)$ and $(Y_t)_{t\geq 1}$ be independent random variables such that $|Y_t| \leq c_t$ and $\mathbf{E}Y_t \geq \delta$ for all $t \geq 1$. Let $X_0 = 0$ and $X_t = \sum_{i \leq t} Y_i$ for all $t \geq 1$. Then for any $\epsilon > 0$, a.a.s. $X_n \geq \delta n - \epsilon|\delta|n$.*

Proof. Clearly $(\delta t - X_t)_{t\geq 0}$ is a supermartingle. Let $n \to \infty$. Then by [15, Lemma 4.2],

$$\mathbf{P}(\delta n - X_n \geq \epsilon|\delta|n) \leq \exp\left(-\frac{\epsilon^2\delta^2 n^2}{\sum_{t=1}^{n} c_t^2}\right) = o(1),$$

as $\sum_{t=1}^{n} c_t^2 = o(n^2)$. Thus, a.a.s. $X_n \geq \delta n - \epsilon|\delta|n$. ∎

Proof of Lemma 5.3. Instead of letting $G_0 \sim \mathcal{H}(n,m,k)$, we let $G_0 \sim \mathcal{M}(n,m,k)$ and run the vertex deletion algorithm on G_0. For all $0 \le t \le \tau$, let L_t denote the degree sum of all vertices with degree less than $k+1$ in G_t. Let V_t and B_t denote the number of vertices and the degree sum of all vertices in G_t.

Claim 5.5 *A.a.s. the algorithm never removes $n^{1/3}$ vertices repeatedly in one step.*

By Claim 5.5, a.a.s. G_τ is the $(k+1)$-core of G_0 and so a.a.s. the numbers of edges of G_τ is $\beta m + o(n)$. Thus, a.a.s. for all $t \le \tau$, $B_t \ge 2\beta m + o(n)$.

Let $\epsilon > 0$ be a small constant. (We will choose $\epsilon \le \min\{1/k^5, \delta/8k^2\}$.) By Claim 5.5, a.a.s. $L_\tau = 0$. Since $|L_{t+1} - L_t| = O(n^{1/3})$ always, we have that a.a.s. there exists t_0 such that $L_{t_0}/2\beta m > \epsilon/2$ and $L_t/2\beta m < \epsilon$ for all $t \ge t_0$, as long as ϵ is sufficiently small.

Let $\Delta B_t = B_{t+1} - B_t$ and similarly we define ΔV_t. Clearly, $\Delta B_t \le -2k$ always. Thus, for any $t > t_0$, $B_t - B_{t_0} \le -2k(t - t_0)$. Next, we analyse a lower bound of V_t. At each step, the algorithm removes one vertex v with degree k and repeatedly all vertices with degree less than k unless more than $n^{1/3}$ vertices are removed. We label v with 0, and we label a vertex with $i \ge 1$ if its degree drops below $k+1$ after the removal of a vertex labelled $i-1$. Let X_i denote the number of vertices labelled i. Then the total number of vertices removed at this step is $\sum_{0 \le i \le n^{1/3}} X_i$, since in each step, at most $n^{1/3}$ light vertices (vertices with degree at most k) can be removed. Consider the pairing model. At step $t+1$, the degree sum of light vertices is $L_t + O(n^{1/3})$ and the degree sum of all vertices is $B_t + O(n^{1/3})$. The probability that a uniformly chosen point lies in a light vertex is $L_t/B_t + O(n^{-2/3})$. Thus, $X_0 = 1$ and for all $i \ge 1$, $\mathbf{E}X_i \le (kL_t/B_t + O(n^{-2/3}))^i$. Since for all $t \ge t_0$, $L_t/B_t \le L_t/B_\tau$ and a.a.s. $B_\tau = 2\beta m + o(n)$, we have $L_t/B_t < \epsilon$ for all $t \ge t_0$ and so

$$\mathbf{E}(\Delta V_t \mid G_t) \ge -1 - \sum_{i \ge 1} \left(k \frac{L_t}{B_t}\right)^i + O(n^{-2/3}) \ge -1 - \delta/4k,$$

by choosing $\epsilon \le \delta/8k^2$. Since $(\Delta V_t)^2 \le n^{2/3} = o(n)$, by Lemma 5.4, for any $t_0 < t \le \tau$ such that $t - t_0 = \Omega(n)$, a.a.s. $V_t - V_{t_0} \ge (-1 - \delta/2k)(t - t_0)$.

Claim 5.6 *A.a.s. $\tau - t_0 = \Omega(n)$.*

It follows then by Claim 5.6 that a.a.s.,

$$\frac{B_\tau}{V_\tau} \le \frac{-2k(\tau - t_0) + B_{t_0}}{(-1 - \delta/2k)(\tau - t_0) + V_{t_0}}. \tag{5.1}$$

If $\rho(G_{t_0}) \ge k - \delta/2$, then a.a.s. $\rho(G_{t_0}) > \rho(G_\tau) + \delta/2$ since a.a.s. $\rho(G_\tau) = \beta m/\alpha n + o(1) < k - \delta$. Now assume $\rho(G_{t_0}) < k - \delta/2$. Let $Z = (\tau - t_0)/V_{t_0}$. By Claim 5.6, a.a.s. there exists $\gamma > 0$, such that $Z > \gamma$. By (5.1), a.a.s.,

$$2\rho(G_\tau) = \frac{B_\tau}{V_\tau} \le \frac{2\rho(G_{t_0}) - 2kZ}{1 - (1 + \delta/3k)Z} \le 2\rho(G_{t_0}) - \bar{\epsilon},$$

where

$$\bar{\epsilon} = \frac{\epsilon\delta\gamma}{2k(1-(1+\delta/2k)\gamma)},$$

since $Z > \gamma$ a.a.s. and $\rho(G_{t_0}) < k - \delta/2$ by assumption. This completes the proof that a.a.s. there exists $G = G_{t_0}$ such that $G_\tau \subseteq G \subseteq G_0$ and $\rho(G) \geq \rho(G_\tau) + \eta$ by choosing $\eta = \min\{\delta/2, \bar{\epsilon}/2\}$. ∎

Proof of Claim 5.5. Recall that in each step, X_i denote the number of vertices removed that are labelled i. If there are $n^{1/3}$ vertices removed, then there exists a vertex labelled $\ell = \log_k^{n^{1/3}}$ that is removed. The expected number of such vertices is $(kL_t/B_t + O(n^{-2/3}))^\ell \leq n^{-4/3}$, because $L_t/B_t < \epsilon \leq 1/k^5$ by the choice of ϵ. Hence, the probability that at some step, there are more than $n^{1/3}$ vertices being removed is $o(n^{-1/3})$ by the union bound. ∎

Proof of Claim 5.6. Consider $\Delta L_t = L_{t+1} - L_t$. Clearly, $-\Delta L_t \leq k\sum_{i\geq 0} X_i$, where X_i denote the number of vertices removed that are labelled i. We have shown that $\mathbf{E}(X_i \mid G_t) = (kL_t/B_t + O(n^{-2/3}))^i$. Thus,

$$\mathbf{E}(\Delta L_t \mid G_t) \geq -k - k\sum_{i\geq 1}\left(\frac{kL_t}{B_t}\right)^i + O(n^{-2/3}) \geq -k-1,$$

by the choice of ϵ. Since $(\Delta L_t)^2 \leq k^2 n^{2/3} = o(n)$ and $L_{t_0} > \epsilon\beta m = \Omega(n)$, by Lemma 5.4, it follows that a.a.s. $\tau - t_0 = \Omega(n)$. ∎

6 Proof of Theorems 2.6

Given a graph G, let $d_{\min}(G)$ denote the minimum degree of G. For a graph G, we call a subgraph H of G *considerable* if $d_{\min}(H) \geq \lfloor\rho(H)\rfloor + 1$. The following proposition is obvious.

Proposition 6.1 *For any (multi)graph G, there exists a considerable subgraph H that is a densest subgraph of G.*

Recall that $\rho^*(G)$ denote the density of a densest subgraph of G.

Lemma 6.2 *Consider $G \sim \mathcal{G}(n, c/n)$. Then for all sufficiently large $c \in \mathbf{R}_0$, a.a.s. there exists $H \subseteq G_{\widetilde{k}}$ such that $\rho(H) = \rho^*(G)$ and $\lfloor\rho^*(G)\rfloor \leq \widetilde{k} - 1$.*

Proof. By Proposition 6.1, there exists a densest subgraph H of G, such that H is considerable. Then we have $d_{\min}(H) \geq \lfloor\rho^*(G)\rfloor + 1$. So H is contained in the $(\lfloor\rho^*(G)\rfloor + 1)$-core. By Theorem 2.3, a.a.s. H is not the $(\lfloor\rho^*(G)\rfloor + 1)$-core, i.e. the $(\lfloor\rho^*(G)\rfloor + 1)$-core is not balanced. By the definition of $\widetilde{k}$, a.a.s. $\lfloor\rho^*(G)\rfloor + 1 \leq \widetilde{k}$, i.e. $\lfloor\rho^*(G)\rfloor \leq \widetilde{k} - 1$. ∎

Lemma 6.3 *Let $\mathcal{G}_{\widetilde{k}}$ denote the $\widetilde{k}$-core of $\mathcal{G}(n, c/n)$. Then a.a.s. $\rho(G_{\widetilde{k}}) > \widetilde{k} - 1$.*

Proof. It follows from Theorem 2.4 and the choice of $c \in \mathbf{R}_0$, that $\rho(G_{\widetilde{k}}) > \widetilde{k} - 1$. ∎

Lemma 6.3 implies that a.a.s. $\rho^*(\mathcal{G}(n, c/n)) > \widetilde{k} - 1$. So a.a.s.

$$\lceil \rho^*(\mathcal{G}(n, c/n)) \rceil + 1 \geq \widetilde{k} + 1. \tag{6.1}$$

Assume H is a subgraph of G. Recall that $d(S)$ denote the degree sum of vertices in S (inside G). Let $d_H(S)$ denote the sum of degrees of vertices of S in the subgraph H (instead of inside G), if $S \subseteq H$.

Lemma 6.4 *Let $G \sim \mathcal{G}(n, c/n)$. A.a.s. there exists $H \supseteq G_{\widetilde{k}+1}$ such that $\rho(H) = \rho^*(G) + O(\log^3 n/n)$, provided c is sufficiently large. Moreover, there is a densest subgraph G' of G such that $|V(G')\Delta V(H)| \leq \log^2 n$.*

Proof. Assume G' is a densest subgraph of $G_{\widetilde{k}}$ and assume $U := G_{\widetilde{k}+1} \setminus G'$ is nonempty. Let $\overline{U}$ denote $G_{\widetilde{k}} \setminus U$. Let $\rho' = \rho(G')$. By Lemma 6.2, a.a.s. $\rho' = \rho^*(G)$. We first prove that if G' is a densest subgraph of $G_{\widetilde{k}}$, then the following conditions are implied.

(a) $t(U) \geq d_{G_{\widetilde{k}+1}}(U) - \rho'|U|$;

(b) a.a.s. $|G'| \geq (1.8/e)|G_{\widetilde{k}+1}|$.

We first prove (b). By Theorem 2.1, a.a.s. $|G_{\widetilde{k}+1}| = \Theta(n)$. By Lemma 4.5, a.a.s. $|G'| \geq \epsilon|G_{\widetilde{k}+1}|$, for some small $\epsilon > 0$, since a.a.s. there is no subgraph of $G \sim \mathcal{G}(n, c/n)$ with size at most $\epsilon_c n$ (ϵ_c is the small constant defined in the statement of Lemma 4.5) and density more than 2, whereas $\rho' > 2$ as long as c is sufficiently large. By Lemma 3.1, $G_{\widetilde{k}+1} \sim \mathcal{H}(n_{\widetilde{k}+1}, m_{\widetilde{k}+1}, \widetilde{k}+1)$ conditional on the number of vertices and edges in $G_{\widetilde{k}+1}$ being $n_{\widetilde{k}+1}$ and $m_{\widetilde{k}+1}$. By the choice of $\widetilde{k}$, we may assume that $m_{\widetilde{k}+1}/n_{\widetilde{k}+1} \leq \widetilde{k} - \delta$ for some absolute constant $\delta > 0$. Then part (b) follows immediately from Lemma 4.7 by noting that $|G_{\widetilde{k}+1}| \leq |G_{\widetilde{k}}|$. Next, we show (a). Let $G' \cup U$ denote the subgraph induced by all vertices in $G' \cup U$. Since G' is a densest subgraph, $\rho' = \rho(G') \geq \rho(G' \cup U)$. Since

$$\rho(G') = \frac{t(G' \cup U) - d_{G' \cup U}(U) + t(U)}{|V(G' \cup U)| - |U|} \leq \frac{t(G' \cup U) - d_{G_{\widetilde{k}+1}}(U) + t(U)}{|V(G' \cup U)| - |U|},$$

we have

$$\rho(G' \cup U) \leq \rho(G') \leq \frac{t(G' \cup U) - d_{G_{\widetilde{k}+1}}(U) + t(U)}{|V(G' \cup U)| - |U|},$$

and thus,

$$\frac{d_{G_{\widetilde{k}+1}}(U) - t(U)}{|U|} \leq \rho(G' \cup U) \leq \rho',$$

which implies (a). We will show that (a) fails for all small U except that $|U| \leq \log^2 n$ and cases of all large U lead to contradiction with (b). Then the proof of Lemma 6.4 is complete.

Let $n_{\widetilde{k}+1} = |V(G_{\widetilde{k}+1})|$ and $m_{\widetilde{k}+1} = |E(G_{\widetilde{k}+1})|$. Let ϵ be that $|V(G_{\widetilde{k}})| = (1+\epsilon)n_{\widetilde{k}+1}$. Note that ϵ is a random variable here, which a.a.s. goes to 0 as $c \to \infty$ by Proposition 5.1 (b) (or equivalently, as $\widetilde{k} \to \infty$ by Proposition 5.1 (a).). As we have discussed, by conditioning on the values of $n_{\widetilde{k}+1}$ and $m_{\widetilde{k}+1}$, $G_{\widetilde{k}+1} \sim \mathcal{H}(n_{\widetilde{k}+1}, m_{\widetilde{k}+1}, \widetilde{k}+1)$ and we may assume that $m_{\widetilde{k}+1}/n_{\widetilde{k}+1} \leq \widetilde{k} - \delta$ for some absolute constant $\delta > 0$, by the choice of $\widetilde{k}$,. Let $r_0 = 1/2$. By Lemma 6.2, a.a.s. $\lfloor \rho^* \rfloor \leq \widetilde{k} - 1$, which implies a.a.s. $\rho^* \leq \widetilde{k}$, whereas in $G_{\widetilde{k}+1}$, the minimum degree is at least $\widetilde{k}+1$. Thus,

$$\begin{aligned}
&\mathbf{P}(\exists U \subseteq G_{\widetilde{k}+1} : \log^2 n \leq |U| \leq r_0 n_{\widetilde{k}+1}, t(U) \geq d_{G_{\widetilde{k}+1}}(U) - \rho'|U|) \\
&\quad \leq \mathbf{P}(\rho' \neq \rho^*) + \mathbf{P}(\rho^* > \widetilde{k}) \\
&\qquad + \mathbf{P}(\rho^* \leq \widetilde{k},\ \exists U \subseteq G_{\widetilde{k}+1} : \log^2 n \leq |U| \leq r_0 n_{\widetilde{k}+1}, t(U) \geq d_{G_{\widetilde{k}+1}}(U) - \rho^*|U|) \\
&\leq o(1) + \mathbf{P}(\exists U \subseteq G_{\widetilde{k}+1} : \log^2 n \leq |U| \leq r_0 n_{\widetilde{k}+1}, t(U) \geq d_{G_{\widetilde{k}+1}}(U) - \widetilde{k}|U|) = o(1),
\end{aligned}$$

where the last equality holds by Lemma 4.8 (with $k = \widetilde{k}+1$).

Next we show that $|U| \geq r_0 n_{\widetilde{k}+1}$ implies $|G'| < (1.8/e) n_{\widetilde{k}+1}$, leading to a contradiction with (b). This follows immediately from the following observation

$$|G'| \leq |\overline{U}| = |G_{\widetilde{k}}| - |U| \leq (1+\epsilon) n_{\widetilde{k}+1} - n_{\widetilde{k}+1}/2 < (1.8/e) n_{\widetilde{k}+1},$$

provided $\widetilde{k}$ is sufficiently large. This is because $\epsilon \to 0$ as $c \to \infty$.

We have shown that if G' is a densest considerable subgraph, then G' contains all but at most $\log^2 n$ vertices of $G_{\widetilde{k}+1}$. Let $H = G' \cup G_{\widetilde{k}+1}$, i.e. a.a.s. G' misses at most $\log^2 n$ vertices of H. By Lemma 4.6, a.a.s. $\rho(H) - \rho(G') = O(\log^3 n/n)$. Thus, a.a.s. there is $G_{\widetilde{k}+1} \subseteq H \subseteq G_{\widetilde{k}}$ such that $\rho(H) = \rho^*(G) + O(\log^3 n/n)$, and $|V(H) \Delta V(G')| \leq \log^2 n$, where G' is a densest subgraph of $\mathcal{G}(n, c/n)$. ∎

Proof of Theorem 2.6. It follows from Lemmas 6.2 and 6.4. ∎

7 Acknowledgement

The author would like to thank Konstantinos Panagiotou and Reto Spöhel for helpful discussions.

References

[1] B. Bollobás, A probabilistic proof of an asymptotic formula for the number of labelled regular graphs, *European J. Combin.* 1 (1980), 311–316.

[2] V. Chvátal, Almost all graphs with 1.44n edges are 3-colorable, *Random Structures Algorithms* 2 (1991), no. 1, 11-28.

[3] J. Cain and P. Sanders and N. Wormald, The random graph threshold for k-orientiability and a fast algorithm for optimal multiple-choice allocation, *Proc. ACM-SIAM Symposium on Discrete Algorithms (SODA)*, January 2007, 469–476.

[4] J. Cain and N. Wormald, Encores on cores, *Electronic Journal of Combinatorics*, 13, RP 81, 2006.

[5] D. Fernholz and V. Ramachandran, The k-orientability thresholds for $G_{n,p}$, *Proc. ACM-SIAM Symposium on Discrete Algorithms (SODA)*, January 2007, 459–468.

[6] S. L. Hakimi, On the degrees of the vertices of a directed graph, *J. Franklin Inst.*, 279, (1965), 290–308.

[7] Jeong Han Kim, Poisson cloning model for random graphs, *International Congress of Mathematicians*, Vol. III, 873-897, *Eur. Math. Soc., Zürich*, 2006.

[8] P. Gao, X. Pérez-Giménez and C. M. Sato, Arboricity and spanning-tree packing in random graphs with an application to load balancing, *arXiv:1303.3881*, extended abstract accepted by *SODA 2014*.

[9] P. Gao and N. Wormald, Orientability thresholds for random hypergraphs, preprint.

[10] B. Pittel and J. Spencer and N. Wormald, Sudden emergence of a giant k-core in a random graph, *J. Combin. Theory Ser. B*,67, no.1, 1996, 111–151.

[11] M. Molloy, Cores in random hypergraphs and Boolean formulas, *Random Structures Algorithms* 27 (2005), no. 1, 124-135.

[12] B. Pittel and J. Spencer and N. Wormald, Sudden emergence of a giant k-core in a random graph, *J. Combin. Theory Ser. B*,67, no.1, 1996, 111–151.

[13] PawełPrałat, Jacques Verstraëte and Nicholas Wormald, On the threshold for k-regular subgraphs of random graphs, *Combinatorica* 31 (2011), no. 5, 565-581.

[14] B. Pittel and N. C. Wormald, Asymptotic enumeration of sparse graphs with a minimum degree constraint, *J. Combin. Theory Ser. A* 101 (2003), no. 2, 249-263.

[15] N.C. Wormald, The differential equation method for random graph processes and greedy algorithms, *Lectures on Approximation and Randomized Algorithms*, (editor: M. Karonski and H.J. Proemel), PWN, Warsaw, 1999, 73–155.