Regular Subgraphs of Random Graphs

Béla Bollobás* Jeong Han Kim† Jacques Verstraëte‡

June 13, 2005

Abstract

In this paper, we prove that there exists a function $\rho_k = (4 + o(1))k$ such that $G(n, \rho/n)$ contains a k-regular graph with high probability whenever $\rho > \rho_k$. In the case of $k = 3$, it is also shown that $G(n, \rho/n)$ contains a 3-regular graph with high probability whenever $\rho > \lambda \approx 5.1494$. These are the first constant bounds on the average degree in $G(n, p)$ for the existence of a k-regular subgraph. We also discuss the appearance of 3-regular subgraphs in cores of random graphs.

1 Introduction

In this paper, we study the appearance of k-regular subgraphs of random graphs. For $k = 2$ and $\lambda > 1$, it is known that with high probability $G(n, \lambda/n)$ contains a cycle. This problem is well-researched and precise results concerning the distribution of cycles may be found in Janson [15], Bollobás [11], Flajolet, Knuth and Pittel [13]. It appears to be substantially more difficult to analyse the appearance of k-regular subgraphs in random graphs for $k \geq 3$. The regular subgraph problem in graphs first appeared in the context of a conjecture of Berge [7], which states that every 4-regular graph contains a 3-regular subgraph. Using elegant algebraic techniques, Alon, Friedland and Kalai [2] showed that every 4-regular graph, to which an edge is added, contains a 3-regular graph. The full conjecture of Berge was verified by Táskinov [23], who determined further those integers k, r for which every r-regular graph contains a k-regular graph [24]. The algebraic techniques of Alon, Friedland and Kalai, which are based on the Chevalley-Warning theorem on roots of polynomials (see Alon [1]), were used to show that for every prime p, every graph of maximum degree at most $2p - 1$ and average degree greater than $2p - 2$ contains a p-regular subgraph. This result, together with known results on gaps between primes, are crucial to the proof of Pyber [21] that every graph of average degree at least $32k^2 \log n$ contains a k-regular subgraph. In contrast, Pyber, Rödl and Szemerédi [22] gave an ingenious probabilistic

*Department of Mathematical Sciences, University of Memphis, Memphis TN 38152. bbollobas@memphis.edu, and Trinity College, Cambridge CB2 1TQ, UK. Research supported by NSF grant DSM 9971788 and DARPA grant F33615-01-C-1900. Research undertaken while visiting Microsoft Research.

†Microsoft Research One Microsoft Way, Redmond, WA 98052. jehkim@microsoft.com

‡Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1. jverstraete@math.uwaterloo.ca Research undertaken during a postdoc at Microsoft Research.

construction of a bipartite random graph, with specified degree sequence and average degree at least $\frac{1}{10}\log\log n$, containing no 3-regular graph. The gap between the bounds $\frac{1}{10}\log\log n$ and $32k^2\log n$ seems to be very difficult to resolve.

We now consider the regular subgraph problem in random graphs $G(n,\lambda/n)$, with independent edge probability λ/n. By the result of Pyber, it is clear that $G(n,\lambda/n)$ contains a k-regular graph with high probability, whenever $\lambda \geq 64k^2\log n$. On the other hand, the number of k-regular graphs on $2t$ vertices is asymptotically

$$\sqrt{2}\mathrm{e}^{-(k-1)(k+1)/4}\left(\frac{k^k}{\mathrm{e}^k k!^2}\right)^t t^{kt}.$$

A short computation shows that the expected number of k-regular subgraphs of a graph $G(n,\lambda/n)$ is $o(1)$ as $n\to\infty$ whenever $\lambda < \frac{2}{\mathrm{e}}(k+(2\pi k/\mathrm{e}^4)^{1/2k})$. Therefore $G(n,\lambda/n)$ almost surely does not contain a k-regular subgraph when λ is appreciably less than $\frac{2}{\mathrm{e}}k$. A priori, it is not clear whether $G(n,\lambda/n)$ should contain a k-regular graph for any constant λ, depending only on k. One of the aims of this paper is to verify this, by showing that we may take λ linear in k:

Theorem 1.1 *Let $k\geq 3$ be an integer. Then, with high probability, $G(n,\lambda/n)$ contains a k-regular subgraph whenever $\lambda \geq 4k+c_0k^{21/40}$, where c_0 is sufficiently large. If k is prime, then the same is true if $\lambda \geq 4k+(\sqrt{8}+o(1))\sqrt{k\log k}$ for an appropriate function $o(1)$ of k that goes to 0 as k goes to infinity.*

We believe that the (monotone) property of containing a k-regular subgraph should have a sharp threshold; this would be an interesting problem to solve. We also consider a related problem, involving the concept of the k-core of a graph.

The *k-core* of a graph G is the unique largest subgraph of G of minimum degree at least k. Bollobás [8] was the first to define and study the k-core in random graphs. It had been asked if there exists a critical constant λ_k for the existence of a k-core in a random graph on n vertices. In other words, with high probability, for all $\varepsilon>0$, $G(n,(\lambda_k+\varepsilon)/n)$ contains a k-core whereas $G(n,(\lambda_k-\varepsilon)/n)$ contains no k-core. One observes that every graph G on n vertices of size at least $(k-1)n-\binom{k}{2}+1$ has a k-core. This is done by successively deleting vertices of degree at most $k-1$ from G. Note that this procedure terminates after at most $n-k-1$ steps, otherwise we obtain a subgraph G' of G with k vertices and

$$e(G') > (k-1)n-\binom{k}{2}-(k-1)(n-k)=\binom{k}{2},$$

which is a contradiction. It is clear, therefore, that $\lambda_k \leq 2(k-1)$, if the constant λ_k exists. Assuming the existence of the constant λ_k, estimates were given by Łuczak, Chvátal [12], Molloy and Reed [18] [19]. The existence was finally proved in Pittel, Spencer and Wormald [20]. The exact value of λ_k was determined too. (cf. Theorem 3.2 in Section 3.)

Their proof is still valid for $\varepsilon = n^{-1/2+\delta}$, $\delta>0$. The second author used a new model for random graphs called the *Poisson Cloning Model*, to simplify the proof and improve the range of ε, that is, the same is true as long as ε is much larger than $n^{-1/2}$, which is the best possible result

one may expect (see [25]). We will briefly present the model and a related theorem that will be used to prove Theorem 1.2 below.

As every k-regular subgraph is contained in the k-core, we deduce that $G(n, \lambda/n)$ does not contain a k-regular subgraph whenever $\lambda < \lambda_k$. On the other hand, we will show how to deduce the following theorem from the results of Kim [17]:

Theorem 1.2 *For all $\lambda > \lambda_4$, the 4-core of $G(n, \lambda/n)$ contains a 3-regular subgraph with high probability, whereas, for some $\lambda > \lambda_3$, the 3-core of $G(n, \lambda/n)$ does not.*

As a consequence of this result if there is a sharp threshold ρ_k for the property of containing a k-regular graph, then ρ_3 lies between $\lambda_3 = 3.3509\cdots$ and $\lambda_4 = 5.1494\cdots$. Theorem 1.2 supports the following conjecture:

Conjecture 1.3 *For all $\lambda > \lambda_{k+1}$, the $(k+1)$-core of $G(n, \lambda/n)$ contains a k-regular subgraph with high probability, whereas, for some $\lambda > \lambda_k$, the k-core of $G(n, \lambda/n)$ does not.*

We prove Theorem 1.1 in Section 2, and the two statements in Theorem 1.2 are proved in Sections 3 and 4 respectively.

2 Regular Subgraphs of Sparse Random Graphs.

In this section, we prove Theorem 1.1. We will show that $G(n, \lambda/n)$ with $\lambda \geq 2k + c_0 k^{21/40}$ contains a k-regular subgraph, provided c_0 is sufficiently large. First, we will concentrate on the random bipartite model $G(n, n, \lambda/n)$ in which each edge has probability p and both parts of the graph have size n and show that $G(n, n, \lambda/n)$ contains a subgraph of maximum degree at most $2k-1$ and average degree greater than $2k-2$. We then apply the following elegant result of Alon, Friedland and Kalai [2], based on algebraic methods, to complete the proof of Theorem 1.1 when k is prime:

Theorem 2.1 (Alon,Friedland,Kalai) *Let k be a prime, and let G be a graph of maximum degree at most $2k-1$ and average degree larger than $2k-2$. Then G contains a k-regular subgraph.*

If k is not prime, then we use the following recent result of Baker, Harman and Pintz [4] to reduce to the case that k is prime: if k is sufficiently large, then some integer between k and $k + \frac{c_0}{2} k^{21/40}$ is a prime.

We now show that with high probability, $G(n, n, \lambda)$ contains a subgraph of maximum degree at most $2k-1$ and average degree larger than $2k-2$ whenever $\lambda \geq 2k + (1+o(1))\sqrt{2k \log k}$ (for an appropriate function $o(1)$ of k that goes to 0 as k goes to infinity).

Theorem 2.2 *Let $k \geq 3$, and let $\lambda \geq k + (2+o(1))\sqrt{k \log k}$ as $k \to \infty$. Then, with high probability, $G(n, n, \lambda)$ contains a subgraph of maximum degree at most k and average degree greater than $k-1$.*

Proof. Let A and B be the parts of the random bipartite graph $G(n,n,\lambda)$. We first create an auxiliary directed graph from the random graph $G = G(n,n,\lambda/n)$, as follows: add new vertices a and b adjacent to all vertices of A and B respectively, and orient all paths of length three from a to b towards b. Let the directed graph so obtained be $\vec{G}$. To apply the max-flow min-cut theorem, we assign capacities to the arcs of $\vec{G}$. All arcs incident with a and b are assigned capacity k, and remaining arcs of $\vec{G}$ are assigned unit capacity. For each flow f in $\vec{G}$, let E_f denote the collection of edges e of G corresponding to arcs $\vec{e}$ of $\vec{G}$ with unit flow:

$$E_f = \{e \in G : f(\vec{e}) = 1\}.$$

Let H_f be the subgraph of G spanned by the edges in E_f. Then, as the capacities on the arcs incident from a and to b are k, every vertex of H_f has degree at most k. Furthermore, $|E_f|$ is precisely the capacity $c(f)$ of f, since

$$|E_f| = \sum_{v\in A, w\in B} f(v\vec{w}) = \sum_{v\in A} f(a\vec{v}).$$

It remains to show that with high probability, a maximum flow f has capacity greater than $(k-1)n$. By the max-flow min-cut theorem (see e.g. Bollobás [11], page 70), the maximum capacity of a flow from a to b equals the minimum capacity of a cut separating a from b. Therefore we show that with high probability, every cut in $\vec{G}$ has capacity greater than $(k-1)n$.

We now prove this statement. For each cut M of $\vec{G}$, we let

$$X = \{v \in A : \vec{av} \in M\} \quad Y = \{v \in B : \vec{vb} \in M\}.$$

Every arc from $A\backslash X$ to $B\backslash Y$ is contained in M, since M is a cut of $\vec{G}$. Let Z denote the set of arcs from $A\backslash X$ to $B\backslash Y$ in $\vec{G}$, and let $c(M)$ denote the capacity of M. Then

$$c(M) \geq k|X| + k|Y| + |Z|.$$

If there exists a cut M such that $c(M) \leq (k-1)n$, then there exists a pair of sets X, Y such that $|Z| \leq (k-1)n - k|X| - k|Y|$ and hence for $|X| = \alpha n$, $|Y| = \beta n$,

$$(k-1)n - k|X| - k|Y| \geq 0 \quad \Rightarrow \quad \alpha + \beta \leq 1 - 1/k$$

and

$$|Z| \leq (k-1)n - (\alpha+\beta)kn. \tag{1}$$

Notice that for a fixed pair X, Y with $|X| = \alpha n$, $|Y| = \beta n$,

$$\mathbb{E}[|Z|] = (1-\alpha)(1-\beta)\lambda n = (1-\alpha)(1-\beta)\lambda n$$

and that (1) yields

$$\begin{aligned} |Z| - (1-\alpha)(1-\beta)\lambda n &\leq (k-1)n - (\alpha+\beta)kn - (1-\alpha)(1-\beta)(k+(\lambda-k))n \\ &= -\Big((\lambda-k)(1-\alpha)(1-\beta) + 1 + \alpha\beta k\Big)n \\ &\leq -\Big((\lambda-k)(1-\alpha)(1-\beta) + \alpha\beta k\Big)n. \end{aligned}$$

For $\lambda - k = (2ck\log k)^{1/2}$, $c > 2$, and

$$\eta = (2ck\log k)^{1/2}(1-\alpha)(1-\beta) + \alpha\beta k,$$

applying the Chernoff bound (see e.g. Alon and Spencer [3], page 268), we obtain

$$\mathbb{P}\Big[|Z| - \mathbb{E}[|Z|] \le -\eta n\Big] \le \exp\Big(-\frac{\eta^2 n^2}{2\mathbb{E}[|Z|]}\Big) = \exp\Big(-\frac{((2ck\log k)^{1/2}(1-\alpha)(1-\beta)+\alpha\beta k)^2 n}{2(1-\alpha)(1-\beta)\lambda}\Big).$$

Let S denote the number of such pairs X, Y; then the probability that there exists M with $c(M) \le (k-1)n$ is at most

$$\begin{aligned}\mathbb{E}[S] &= \sum_{\substack{X\subset A, Y\subset B:\\ k(|X|+|Y|)\le (k-1)n}} \mathbb{P}\Big[|Z| \le (k-1)n - k|X| - k|Y|\Big] \\ &\le \sum_{\substack{\alpha,\beta:\\ \alpha+\beta\le 1-1/k}} \binom{n}{\alpha n}\binom{n}{\beta n}\exp\Big(-\frac{((2ck\log k)^{1/2}(1-\alpha)(1-\beta)+\alpha\beta k)^2 n}{2(1-\alpha)(1-\beta)\lambda}\Big).\end{aligned}$$

Clearly,

$$\begin{aligned}&\frac{1}{n}\log\Big(\binom{n}{\alpha n}\binom{n}{\beta n}\exp\Big(-\frac{((2ck\log k)^{1/2}(1-\alpha)(1-\beta)+\alpha\beta k)^2 n}{2(1-\alpha)(1-\beta)\lambda}\Big)\Big) \\ &\le H(\alpha) + H(\beta) - \frac{ck(1-\alpha)(1-\beta)\log k}{\lambda} - \frac{(\alpha\beta)^2 k^2}{2(1-\alpha)(1-\beta)\lambda},\end{aligned}$$

where the entropy function $H(x)$ is defined by

$$H(x) = -x\log x - (1-x)\log(1-x).$$

Since there are at most n^2 pairs (α, β), it is enough to show that

$$H(\alpha) + H(\beta) - \frac{ck(1-\alpha)(1-\beta)\log k}{\lambda} - \frac{(\alpha\beta)^2 k^2}{2(1-\alpha)(1-\beta)\lambda} \le -\frac{(c-2-o(1))\log k}{k},$$

for all $0 \le \alpha, \beta \le 1$ with $\alpha + \beta \le 1 - 1/k$. (So, $c > 2$ is required for this approach.)

First, if $\alpha\beta \ge (4/k)^{1/2}$, then

$$H(\alpha) + H(\beta) - \frac{(\alpha\beta)^2 k^2}{2(1-\alpha)(1-\beta)\lambda} \le 2\log 2 - \frac{4k}{2\lambda} \le -0.6 + o(1).$$

Suppose $\alpha\beta < (4/k)^{1/2}$. Then one of α or β, say α, is $O(k^{-1/4})$. In particular, $1 - \alpha = 1 - o(1)$. We will use

$$\frac{ck(1-\alpha)(1-\beta)\log k}{\lambda} = (1-o(1))c(1-\beta)\log k.$$

If $1 - \beta \ge (\log k)^{-1}$, then

$$H(\alpha) + H(\beta) - \frac{ck(1-\alpha)(1-\beta)\log k}{\lambda} \le 2\log 2 - (1+o(1))c \le -0.6 + o(1),$$

for $c > 2$. If $1 - \beta \leq (\log k)^{-1} = o(1)$, then using

$$\alpha + \beta \leq 1 - 1/k \quad \Rightarrow \quad \alpha \leq 1 - \beta \ \text{ and } \ 1 - \beta \geq 1/k$$

we have

$$\begin{aligned} & H(\alpha) + H(\beta) - (1 + o(1))c(1 - \beta) \log k \\ & = -(1 + o(1))\alpha \log \alpha - (1 + o(1))(1 - \beta) \log(1 - \beta) - c(1 + o(1))(1 - \beta) \log k \end{aligned}$$

for $\alpha \leq 1 - \beta = o(1)$ (here we may not use $(1 + o(1))f - (1 + o(1))g = (1 + o(1))(f - g)$ since it is simply not true in general, for example for $f = g$), and hence $\alpha \leq 1 - \beta$ and $1 - \beta \geq 1/k$ imply

$$\begin{aligned} H(\alpha) + H(\beta) - (1 + o(1))c(1 - \beta) \log k & \leq -(1 - \beta)(2 \log(1 - \beta) + c \log k) + o((1 - \beta) \log k) \\ & \leq -(c - 2 - o(1)) \log k/k. \end{aligned}$$

□

Proof of Theorem 1.1. Suppose k is prime and let $n_1 = \lfloor n/2 \rfloor$. Clearly, $G(n, \lambda/n)$ contains a random graph $G(n_1, n_1, \lambda/n)$. Since $\lambda/n = (1 + o(1))(\lambda/2)/n_1$,

$$\lambda/2 \geq 2k - 1 + (2 + o(1))\sqrt{(2k - 1) \log(2k - 1)}\,,$$

or

$$\lambda \geq 4k + (\sqrt{8} + o(1))\sqrt{k \log k}\,,$$

and Theorem 2.2 implies that $G(n_1, n_1, \lambda/n)$ contains a subgraph of maximum degree at most $2k-1$ and average degree larger than $2k-2$. By Theorem 2.1, $G(n_1, n_1, \lambda/n)$, and hence $G(n, \lambda/n)$, contains a k-regular graph, as required.

Now suppose k is not prime. By a recent result of Baker, Harman and Pintz [4], there exists a prime ℓ between k and $k + ck^{21/40}$ for some constant c. The same argument as above yields $G(n_1, n_1, \lambda/n)$ contains an ℓ-regular subgraph provided

$$\lambda \geq 4\ell + (4 + o(1))\sqrt{\ell \log \ell}\,,$$

which follows from

$$\lambda \geq 4k + 4c(1 + o(1))k^{21/40} = 4k + O(k^{21/40}).$$

Since every regular bipartite graph contains a perfect matching, one may keep removing perfect matching(s) from the ℓ-regular subgraph until the remaining graph is k-regular. □

3 Cubic Subgraphs of The 4-Core.

In this section, we show that the 4-core in $G(n,p)$, $pn > \lambda_4$, contains a 3-regular subgraph with high probability (Theorem 1.2). We use a model $G_{PC}(n,p)$ called *Poisson Cloning Model*, developed by Kim [17].

In the Poisson cloning model, the degrees $d(v)$ of vertices v are first chosen (without any graph yet) to be independent Poisson $\lambda = p(n-1)$ random variables. Then, for each vertex v, $d(v)$ copies, or *clones*, of v are to be chosen. If the sum of all degrees is even, a perfect matching on the set of all clones is a set of $\frac{1}{2}\sum_{v\in V} d(v)$ pairwise disjoint edges each consisting of two clones. The Poisson cloning model is obtained by projecting, or contracting, all clones of v to the single vertex v from the uniform random perfect matching. When the degree sequence $(d(v))$ is given as a sequence of deterministic numbers, this cloning model is the same as the configuration model considered in Bollobás [9] (see also [6]). For more details on models of random regular graphs, see Wormald [26].

Generally, at each step, an unmatched clone is chosen in a certain way to be matched to a clone chosen uniformly at random among all other unmatched clones. If $\sum_{v\in V} d(v)$ is even, this yields a uniform random perfect matching regardless the ways to choose the first unmatched clones. As described before, the Poisson cloning model is obtained by projecting, or contracting, all clones of v to the single vertex v from the perfect matching. If $\sum_{v\in V} d(v)$ is odd, the last unmatched clone, say of w, is projected to a loop on w. Unlikely all other loops, which increase degrees by 2, this special loop increases the degree by only 1. The Poisson cloning model consists of the loop and the edges obtained by projecting all clones of v to the single vertex v from the perfect matching on all clones but the clone of w. The choice of the first unmatched clone is prescribed by a choice function. Thus, a sequence of choice functions determines the model. Notice, however, that all sequences of choice functions yield the same model if $\sum_{v\in V} d(v)$ is even. If $\sum_{v\in V} d(v)$ is odd, the model produces a non-simple graph and varies depending on particular sequence of choice functions. We write $G^{(\mathcal{S})}_{PC}(n,p)$ for the Poisson cloning model with the sequence $\mathcal{S}$ of choice functions.

Theorem 3.1 *If $pn = O(1)$, then there are positive constants c_1 and c_2 so that for any collection $\mathcal{G}$ of simple graphs and any sequence of choice functions $\mathcal{S}$*

$$c_1 \Pr[G^{(\mathcal{S})}_{PC}(n,p) \in \mathcal{G}] - e^{-\Omega(pn)} \le \Pr[G(n,p) \in \mathcal{G}] \le c_2\Big(\Pr[G^{(\mathcal{S})}_{PC}(n,p)(n,p) \in \mathcal{G}]\Big)^{1/2} + e^{-\Omega(pn)}.$$

To prove Theorem 1.2, we may select choice functions that choose clones of vertices with 3 or less unmatched clones whenever possible. In (only) this section, $G_{PC}(n,p)$ represents $G^{(\mathcal{S})}_{PC}(n,p)$ for a (fixed) sequence $\mathcal{S}$ of such choice functions. All we have to do now is to prove that the 4-core of $G_{PC}(n,p)$, $pn > \lambda_4 + \delta$ with a fixed $\delta > 0$, contains a 3-regular subgraph with high probability. Then applying the upper bound in Theorem 3.1 for the collection $\mathcal{G}$ of all simple graphs the 4-cores of which contain no 3-regular subgraph, we obtained the desired result. The theorem will also be used when we show, in the next section, that $G(n,\lambda/n)$ contains no 3-regular subgraph for some $\lambda > \lambda_3$.

Let $\lambda = p(n-1)$ and X_ρ denote a generic Poisson random variable with mean ρ, and let $P(\rho, k)$ denote the cumulative distribution function $\Pr[X_\rho \geq k]$. To prove the first part of Theorem 1.2, we require the following descriptive theorem concerning the k-core. Define $\rho(\lambda, k)$ to be the largest solution of $\rho = \lambda P(\rho, k-1)$. One may check that $\rho(\lambda, k) > 0$ if and only if

$$\lambda \geq \lambda_{_k} := \inf_{\rho>0} \frac{\rho}{P(\rho, k-1)},$$

and $\rho(\lambda, k) = 0$ otherwise. Let

$$\begin{aligned}
\mathbf{n}(\lambda, k) &= \text{the number of vertices in the } k\text{-core of } G_{PC}(n, \lambda/(n-1)),\\
\mathbf{n}_{_i}(\lambda, k) &= \text{the number of vertices of degree } i \text{ in the } k\text{-core}
\end{aligned}$$

and

$$\mathbf{a}(\lambda, k) = \mathbf{n}(\lambda, k)/n, \quad \mathbf{a}_i(\lambda, k) = \mathbf{n}_{_i}(\lambda, k)/n\,.$$

The following parameters are needed:

$$\rho(k) = \rho(\lambda_{_k}, k), \quad \alpha(k) = \Pr\Big(\text{Poisson}(\rho(k)) \geq k\Big),$$

and

$$\alpha_{_i}(k) = \Pr\Big(\text{Poisson}(\rho(k)) = i\Big) = \frac{(\rho(k))^i e^{-\rho(k)}}{i!} \quad \forall\ i \geq k.$$

Numerically, $\lambda_3 = 3.3509\cdots$, $\lambda_4 = 5.1494\cdots$ and $\rho(3) = 1.7932\cdots$, $\rho(4) = 3.3836\cdots$.

Theorem 3.2 *For fixed k and $\lambda = \lambda_{_k} + \delta$, $\delta > 0$, we have*

$$\rho(\lambda, k) = \rho(k) + O(\delta^{1/2})$$

and with high probability in $G_{PC}(n, \lambda/n)$,

$$\mathbf{a}(\lambda, k) = (1 + O(\delta^{1/2}))\alpha(k), \quad \mathbf{a}_i(\lambda, k) = (1 + O(\delta^{1/2}))^i \alpha_{_i}(k) \quad \forall\ i \geq k. \tag{2}$$

Moreover, if $\mathbf{a}(\lambda, k)$ and $\mathbf{a}_i(\lambda, k)$ are given, the k-core of $G_{PC}(n, \lambda/n)$ is isomorphic to the cloning model conditioned on a degree sequence with $\mathbf{a}_i(\lambda, k)n$ vertices of degree i.

Proof of the first part of Theorem 1.2 Let $\lambda = \lambda_{_4} + \delta$. We will show that the 4-core in $G = G_{PC}(n, \lambda/n)$ contains a subgraph of maximum degree at most five and average degree larger than four with high probability. Thereafter, we apply Theorems 2.1 and 3.1 to deduce that the 4-core contains a 3-regular subgraph.

Suppose a degree sequence $(d(v))$ of the 4-core is given so that there are $\mathbf{a}_{_i}(\lambda, 4)n$ vertices of degree i, and all $\mathbf{a}_{_i}(\lambda, 4)$ together with $\mathbf{a}(\lambda, 4) := \sum_{i\geq 4} \mathbf{a}_{_i}(\lambda, 4)$ satisfy (2). We write a and $a_{_i}$ for $\mathbf{a}(\lambda, 4)$ and $\mathbf{a}_{_i}(\lambda, 4)$, respectively. Recall that, in the cloning model, $d(v)$ clones are to be chosen for each v of degree $d(v)$. To obtain the desired subgraph, we truncate vertices of degree larger than 5 as follows: For a vertex of degree larger than 5, color (any) 5 clones blue. Color all clones

of vertices of degree 5 or less blue too. All other clones are to be colored red. We now generate a uniform random perfect matching on the set of all clones. An edge in the perfect matching is called blue only if it consists of two blue clones. Let G be the subgraph obtained by the projection of all blue edges. Clearly, the maximum degree of G is at most 5. For the average degree, notice that the number of clones belonging to blue edges is at least the number of blue clones minus the number of edges containing exactly one blue (or red) clone, which is at most the number of red clones. That is, the average degree of G is at least

$$\frac{\text{the number of blue clones - the number of red clones}}{an},$$

or

$$\frac{4a_4 + 5\sum_{i\geq 5} a_i - \sum_{i\geq 5}(i-5)a_i}{a} = \frac{4a_4 + 10\sum_{i\geq 5} a_i - \sum_{i\geq 5} ia_i}{a}.$$

Since (2) gives

$$a_4 \geq 0.1852, \quad \sum_{i\geq 5} a_i = a - a_4 \geq 0.4380 - 0.1853 = 0.2527 \text{ and } \sum_{i\geq 5} ia_i \leq 1.4823$$

for sufficiently small δ, the average degree is at least

$$\frac{4 \cdot 0.1852 + 10 \cdot 0.2527 - 1.4823}{0.4381} \geq 4.07.$$

□

Remark. As we may see in the last inequality in the proof, one may use the same approach to show that $G(n, \lambda/n)$ contains a 3-regular subgraph even for λ slightly less than λ_4 – all that is required is a subgraph of *average* degree more than four and maximum degree at most five.

4 Absence of 3-Regular Subgraphs in the 3-Core

In this section, we show that there is $\delta > 0$ such that the 3-core of $G_{PC}(n, \lambda/n)$ with $\lambda = \lambda_3 + \delta$ contains no 3-regular subgraph with high probability, where $G_{PC}(n,p)$ represents $G_{PC}^{(\mathcal{S})}(n,p)$ for a (fixed) sequence $\mathcal{S}$ of choice functions that choose clones of vertices with 2 or less unmatched clones whenever possible. This will establish the second part of Theorem 1.2.

Throughout this section, we assume that $\mathbf{a}(\lambda, 3)$ and $\mathbf{a}_i(\lambda, 3)$ are given and satisfy (2) and we write $a = \mathbf{a}(\lambda, 3)$, $a_i = \mathbf{a}_i(\lambda, 3)$, $n_i = a_i n$, and $2m = \sum_{i\geq 3} i\mathbf{a}_i(\lambda, 3)$. Notice that m is the number of edges in the 3-core.

Let $t_3, t_4, \cdots$ be fixed non-negative integers with $\sum_{j\geq 3} t_j = t$. The expected number (with multiplicity) of 3-regular subgraphs of the 3-core of $G_{PC}(n, \lambda/n)$, containing t_j vertices of degree j in the 3-core, is precisely

$$\prod_{j\geq 3} \binom{n_j}{t_j} \binom{j}{3}^{t_j} \cdot \frac{(3t-1)!!(2m-3t)!!}{(2m-1)!!}.$$

Summing over all vectors $\underline{s}(t) = (t_3, t_4, \ldots)$ with $\sum_{j\geq 3} t_j = t$ and all $t \leq an$, we find that the expected number of 3-regular subgraphs in the 3-core is

$$\sum_{t\leq an} \sum_{\underline{s}(t)} \prod_{j\geq 3} \binom{n_j}{t_j} \binom{j}{3}^{t_j} \cdot \frac{(3t-1)!!(2m-3t)!!}{(2m-1)!!}.$$

We will show that this expression is $o(1)$ as $n \to \infty$. First of all, it is easy to check that the last term in this expression is at most $\binom{2m}{3t}^{-1/2}$ and

$$\frac{(3t-1)!!(2m-3t)!!}{(2m-1)!!} \leq m^{1/4} \exp\Big(- mH(3t/2m)\Big).$$

To bound the expectation, we note that for all $\tau > 0$,

$$\begin{aligned}
\sum_{\underline{s}(t)} \prod_{j\geq 3} \binom{n_j}{t_j} \binom{j}{3}^{t_j} &= \tau^{-t} \cdot \left(\prod_{j\geq 3} \sum_{i=0}^{n_j} \binom{n_j}{t_j} \binom{j}{3}^{i} \tau^i \right)[\tau^t] \\
&\leq \tau^{-t} \cdot \left(\prod_{j\geq 3} (1+\tau\tbinom{j}{3})^{n_j} \right) \\
&= \exp\left(-t\log\tau + \sum n_j \log(1+\tau\tbinom{j}{3}) \right).
\end{aligned}$$

Therefore the expected number of 3-regular subgraphs of the 3-core is at most

$$m^{1/4} \sum_{t\leq an} \exp\left(-t\log\tau + \textstyle\sum_{j\geq 3} n_j \log(1+\tau\tbinom{j}{3}) - mH(\frac{3t}{2m}) \right).$$

Fixing t, the exponent in the summand is minimized when τ satisfies

$$t = n\tau \sum_{j\geq 3} \frac{a_j \tbinom{j}{3}}{1+\tau\tbinom{j}{3}}.$$

In stead of regarding τ as a function of t, one may regard t as a function of τ. In terms of τ, $\mu := m/n$, and

$$T = T(\tau) = \sum_{j\geq 3} \frac{a_j \tau \tbinom{j}{3}}{1+\tau\tbinom{j}{3}},$$

we need to analyze

$$\phi(\tau) := -T\log\tau + \sum_{j\geq 3} a_j \log(1+\tau\tbinom{j}{3}) - \mu H\Big(\frac{3T}{2\mu}\Big).$$

We will show that $\phi(\tau) < 0$ for all τ independently of n except possibly $o(1)$ terms. Since number of possible values of t is at most n, this is enough.

First, we notice that as $\tau \to 0$,

$$\begin{aligned}
\phi(\tau) &= -(1+o(1)) \sum_{j\geq 3} a_j \tbinom{j}{3} \tau\log\tau + (1+o(1)) \sum_{j\geq 3} a_j \tbinom{j}{3}\tau + (3/2+o(1)) \sum_{j\geq 3} a_j \tbinom{j}{3}\tau\log\tau \\
&= (1/2+o(1)) \sum_{j\geq 3} a_j \tbinom{j}{3}\tau\log\tau,
\end{aligned}$$

and as $\tau \to \infty$

$$\begin{aligned}\phi(\tau) &= -(1+O(\tau^{-1}))a\log\tau + a\log\tau + (1+o(1))\sum_{j\geq 3} a_j \log\binom{j}{3} - (1+o(1))\mu H\Big(\frac{3a}{2\mu}\Big)\\ &\leq -0.07 + O(\delta^{1/2})\end{aligned}$$

(recall $a = \sum_{j\geq 3} a_j \approx 0.267$ and $2\mu = \sum_{i\geq 3} i a_i \approx 0.959$). Thus, it remains to show that

$$\phi(\tau) < 0 \quad \text{for all } \tau \text{ with } \phi'(\tau) = 0.$$

Writing T' for $T'(\tau)$, we have

$$\phi'(\tau) = -T'\Big(\log\tau + \frac{3}{2}\log\Big(\frac{2\mu}{3T} - 1\Big)\Big),$$

and, from $T' > 0$, $\phi'(\tau) = 0$ implies that

$$\log\tau + \frac{3}{2}\log\Big(\frac{2\mu}{3T} - 1\Big) = 0, \quad \text{or} \quad T = \frac{2\mu}{3(1+\tau^{-2/3})}.$$

Notice that for $\tau \geq 12$

$$T(\tau) \leq a < 0.268 \leq \frac{2\mu}{3(1+\tau^{-2/3})}$$

and for $\tau \leq 0.02$

$$\frac{T}{\tau} \leq \sum_{i\geq 3} a_j \binom{j}{3} < 0.962 \leq \frac{2\mu}{3\tau(1+\tau^{-2/3})}, \quad \text{or} \quad T < \frac{2\mu}{3(1+\tau^{-2/3})} \quad ,$$

assuming δ is small enough. We may now assume $0.02 \leq \tau \leq 12$. There seem to be two solutions of $\phi'(\tau) = 0$, namely, near $\tau = 0.134$ and $\tau = 4.51$ (assuming δ is small enough). We could not, however, find a reasonably simple way to prove that $\phi'(\tau) = 0$ has only two solutions on $(0.02, 12)$. For a complete proof without using any computer software, except when we compute specific values of a function at finite number of certain points, we apply Taylor theorem on the intervals: For $\beta_i = 0.02 \cdot (1.2)^i$ and $\beta_{i-1} \leq \tau \leq \beta_i$, $i = 1, 2, ..., 36$, we have

$$\begin{aligned}\phi(\tau) &= \phi(\beta_i) + \phi'(\tau^*)(\tau - \beta_i)\\ &= \phi(\beta_i) + T'(\tau^*)\Big(\log\tau^* + \frac{3}{2}\log(\frac{2\mu}{3T(\tau^*)} - 1)\Big)(\beta_i - \tau)\end{aligned}$$

for some τ^* with $\beta_{i-1} \leq \tau^* \leq \beta_i$. The number 36 is chosen to be the minimum with $\beta_i \geq 12$. Since $T > 0$ increases and $T'(\tau) > 0$ decreases, as τ increases,

$$T'(\tau^*)\Big(\log\tau^* + \frac{3}{2}\log\Big(\frac{2\mu}{3T(\tau^*)} - 1\Big)\Big) \leq T'(\beta_{i-1}) \cdot \Big|\log\beta_i + \frac{3}{2}\log\Big(\frac{2\mu}{3T(\beta_{i-1})} - 1\Big)\Big|$$

and

$$\phi(\tau) \leq \gamma_i := \phi(\beta_i) + T'(\beta_{i-1}) \cdot \Big|\log\beta_i + \frac{3}{2}\log(\frac{2\mu}{3T\Big(\beta_{i-1}\Big)} - 1\Big)\Big|(\beta_i - \tau).$$

By estimating all 36 values of γ_i's, we conclude that

$$\phi(\tau) \leq \max_{i=1,...,36} \gamma_i \leq -2.5 \times 10^{-4}, \quad \forall \tau : \ 0.02 \leq \tau \leq 12.$$

□

5 Closing Remarks

With the current techniques, we are still quite far from determining whether there is a critical constant ρ_k for the property of containing a k-regular subgraph, even when $k = 3$. The existence of such a constant, together with a verification of Conjecture 1.3, would give $\rho_k \leq \lambda_{k+1}$. It would follow that $\lambda_k \leq \rho_k \leq \lambda_{k+1}$ and, in particular, $\rho_k = k + \sqrt{k \log k} + O(\log k)$, whereas we have shown in this paper that $\rho_k \leq (4 + o(1))k$. The reason for the factor four is twofold: first, a factor two is introduced in applying the result of Alon, Friedland and Kalai, and second a factor two is lost by taking a spanning bipartite subgraph of $G(n, \lambda/n)$. It might be possible to eliminate each of these factors of two, and therefore show that if ρ_k exists, then $\rho_k = (1 + o(1))k$. Even further, perhaps it is true that the $(k+1)$-core in a random graph $G(n, \lambda/n)$ for $\lambda > \lambda_k$ contains a k-factor – a spanning k-regular subgraph.

The questions encountered in this paper are particular cases of the following problem. Let X be a set of positive integers. Determine whether there exists a critical constant $\lambda(X)$ for which, with high probability, $G(n, \lambda/n)$ contains a subgraph all of whose vertex degrees are in X whenever $\lambda > \lambda(X)$ and does not contain such a subgraph whenever $\lambda < \lambda(X)$. The k-core problem is the case $X = \{k, k+1, k+2, \ldots\}$, and the k-regular subgraph problem is the case $X = \{k\}$. In particular, the results of this paper show in this case that $\lambda(X) \leq (4 + o(1))k$, where $k = \min\{x : x \in X\}$. We note also that $\lambda(X) \geq \lambda_k$. It would be interesting to establish stronger results in the cases $X = m\mathbb{Z} + \ell$ – the set of all integers congruent to ℓ modulo m. Finally, it would be interesting to determine whether $\lambda(\{k\}) = (1 + o(1))\lambda_k$ as $k \to \infty$.

References

[1] Alon, N. Combinatorial Nullstellensatz. Recent trends in combinatorics (Mátraháza, 1995). Combin. Probab. Comput. 8 (1999), no. 1-2, 7–29.

[2] Alon, N., Friedland, S., Kalai, G. Regular subgraphs of almost regular graphs. J. Combin. Theory Ser. B 37 (1984), no. 1, 79–91.

[3] Alon, N., Spencer, J. The probabilistic method. Second edition. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.

[4] Baker, R., Harman, G., Pintz, J. The difference between consecutive primes. II. Proc. London Math. Soc. (3) 83 (2001), no. 3, 532–562.

[5] Békéssy, A., Békéssy, P., Komlós J. Asymptotic enumeration of regular matrices, *Studia Sci. Math. Hungar.*, **7** (1972), 343–353.

[6] Bender, E., Canfield, R. The asymptotic number of labeled graphs with given degree sequences, *J. Combinatorial Theory Ser. A* **24** (1978), 296–307.

[7] Bondy, J., Murty, U. Graph Theory and Applications, MacMillan, London (1976).

[8] Bollobás, B. The evolution of sparse graphs. Graph theory and combinatorics (Cambridge, 1983), 35–57, Academic Press, London, 1984.

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

[10] Bollobás, B. Random graphs. Second edition. Cambridge Studies in Advanced Mathematics, 73. Cambridge University Press, Cambridge, 2001

[11] Bollobás, B., Cooper, C., Fenner, T., Frieze, A. Edge disjoint Hamilton cycles in sparse random graphs of minimum degree at least k. J. Graph Theory 34 (2000), no. 1, 42–59.

[12] Chvátal, V. Almost all graphs with $1.44n$ edges are 3-colorable. Random Structures Algorithms 2 (1991) 11–28.

[13] Flajolet, P., Knuth, D., Pittel, B. The first cycles in an evolving graph. Discrete Math. 75 (1989) 167–215.

[14] Huxley, M. On the difference between consecutive primes. Invent. Math. 15 (1972), 164–170.

[15] Janson, S. Poisson convergence and Poisson processes with applications to random graphs. Stochastic Process. Appl. 26, (1987), 1–30.

[16] Łuczak, T. Size and connectivity of the k-core of a random graph. Discrete Math. 91 (1991), no. 1, 61–68.

[17] Kim, J-H. Poisson cloning model for random graphs. In preparation.

[18] Molloy, M. Math. Thesis, University of Waterloo, (1992).

[19] Molloy, M., Reed, B. Critical subgraphs of a random graph. Electron. J. Combin. 6 (1999), no. 1, Research Paper 35, 13 pp.

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

[21] Pyber, L. Regular subgraphs of dense graphs. Combinatorica 5 (1985), no. 4, 347–349.

[22] Pyber, L., Rödl, V., Szemerédi, E. Dense graphs without 3-regular subgraphs. J. Combin. Theory Ser. B 63 (1995), no. 1, 41–54.

[23] Táskinov, V. Regular subgraphs of regular graphs, Soviet Math. Dokl. 26 (1982) 37–38.

[24] Táskinov, V. Regular parts of regular pseudographs. Mat. Zametki 43 (1988), no. 2, 263–275, 303; translation in Math. Notes 43 (1988), no. 1-2, 149–156.

[25] Wilson, D. On the critical exponents of random k-SAT. Random Structures and Algorithms, to appear.

[26] Wormald N. Models of random regular graphs, In *Surveys in Combinatorics, 1999*, J.D. Lamb and D.A. Preece, eds, pp. 239–298.