Cores of Random r-Partite Hypergraphs

Fabiano C. Botelho[a,c], Nicholas Wormald[b], Nivio Ziviani[c]

[a] Department of Computer Engineering, Federal Center for Technological Education of Minas Gerais, Belo Horizonte, Brazil
[b] Department of Combinatorics and Optimization, University of Waterloo, Waterloo ON, Canada
[c] Department of Computer Science, Federal University of Minas Gerais, Belo Horizonte, Brazil

Abstract

We show that the threshold $c_{r,k}$ for appearance of a k-core in a random r-partite r-uniform hypergraph $G_{r,n,m}$ is the same as for a random r-uniform hypergraph with cn/r edges without the r-partite restriction, where $r, k \geq 2$. In both cases, the average degree is c. This is an important problem in the analysis of the algorithm presented in [2]. The algorithm constructs a family of minimal perfect hash functions based on random r-partite r-uniform hypergraphs with an empty k-core subgraph, for $k \geq 2$. The above claim was not proved but was provided with strong experimental evidence. For an input key set S with m keys, the algorithm was the first one capable of constructing a simple and efficient family of minimal perfect hash functions that can be stored in $O(m)$ bits, where the hidden constant is within a factor of two from the information theoretical lower bound. The case $r, k = 2$ was analyzed in [3] but the general case $r \geq 3, k \geq 2$ was still open.

Key words: Minimal Perfect Hashing, cores of hypergraphs

1. Introduction

The study of random graphs started with Erdős and Rényi [6, 7, 9, 8]. A modern treatment is given in [1, 13]. Many results describing statistical properties of random graphs were obtained [10, 11, 12, 16, 17, 19, 20, 21]. For instance, distribution of component sizes, existence and size of a giant component, vertex degree distributions, arising of cycles, existence and size of specific subgraphs, among others.

We now introduce the following definitions:

Definition 1 *Let $G_{r,n,m} = (V, E)$ be a random r-partite r-uniform hypergraph where V is a disjoint union of the r parts $U_1, \ldots, U_r$, $|U_i| = n$ for $i = 1, \ldots, r$, $|E| = m = cn$, $r \geq 2$ and $c > 0$. The edges are inserted into $G_{r,n,m}$ one at a time, each being picked at random from the all n^r possible edges, allowing repetitions.*

Definition 2 *The k-core of a hypergraph is the largest subgraph of minimum degree at least k.*

Definition 3 *A minimal perfect hash function is a bijection from a static key set S of size m to $\{0, 1, \ldots, m-1\} = [m]$.*

The objective of this paper is to prove that the threshold $c_{r,k}$ for appearance of a k-core in $G_{r,n,m}$ is the same as for a random r-uniform hypergraph with cn/r edges without the r-partite restriction, where $r, k \geq 2$. In both cases, the average degree of the hypergraph is c. This problem came up in the analysis of the algorithm presented in [2], where the above claim was not proved but was provided with strong experimental evidence. The algorithm constructs a family of minimal perfect hash functions based on random r-partite r-uniform hypergraphs with an empty k-core subgraph, for $k \geq 2$. The idea of basing minimal perfect hashing on random hypergraphs was not new, see e.g. [14], but Botelho, Pagh and Ziviani proceeded differently in [2] to construct near-optimal space functions that are stored in $O(m)$ bits rather than $O(m \log m)$ bits. The case $r, k = 2$ was analyzed in [3] but the general case $r \geq 3, k \geq 2$ was still open.

In Section 1.1 we present some basic concepts and definitions. In Section 1.2 we outline our results and contributions.

1.1. Preliminaries

In this section we introduce some definitions found in [4] in order to use the same approach to prove the results summarized in Section 1.2. The first determination of the threshold of existence of a k-core in a random graph was given by Pittel, Spencer and the second author [18].

As shown in [4] there is a connection between independent Poisson random variables and multinomials. Let $\mathbf{Multi}(n, s)$ be the probability space of nonnegative integer vectors $(X_1, \ldots, X_n)$ whose entries sum to s, such that for any vector $(d_1, \ldots, d_n)$ with the same sum restriction, we have:

$$\mathbf{P}(X_i = d_i \text{ for } 1 \leq i \leq n) = \frac{s!}{n^s \prod_{i=1}^{n} d_i!}.$$

We will be interested in vertices with degrees at least k, and accordingly define

$$H_{n,s,k} := \{(h_1, \ldots, h_n) : \sum h_i = s \text{ and } h_i \geq k \text{ for all } i\}, \quad (1)$$

and let $\mathbf{Multi}(n, s)|_{\geq k}$ be the probability space obtained by restricting $\mathbf{Multi}(n, s)$ to elements of $H_{n,s,k}$. For $k \geq 0$, denote

Email addresses: fabiano@decom.cefetmg.br (Fabiano C. Botelho), nwormald@uwaterloo.ca (Nicholas Wormald), nivio@dcc.ufmg.br (Nivio Ziviani)

by $Z(k,\lambda)$ a random variable which has a k-truncated Poisson distribution with parameter λ, that is:

$$\mathbf{P}(Z(k,\lambda)=j)=\begin{cases}\dfrac{\lambda^j}{j!\,f_k(\lambda)} & j\geq k\\ 0 & j<k\end{cases} \tag{2}$$

where f_k is defined as:

$$f_k(\lambda)=e^{\lambda}-\sum_{i=0}^{k-1}\frac{\lambda^i}{i!}=\sum_{i\geq k}\frac{\lambda^i}{i!}.$$

Let λ_b denote the positive root of the equation

$$\mathbf{E}\,Z(k,\lambda)=\frac{\lambda f_{k-1}(\lambda)}{f_k(\lambda)}=b. \tag{3}$$

It is easily seen that λ_b exists provided that $b>k$. Let $h_k(\mu)=\frac{\mu}{e^{-\mu}f_{k-1}(\mu)}$ and define:

$$c_k=\inf\{h_k(\mu):\mu>0\}. \tag{4}$$

Take $k\geq 3$. Then c_k is a positive real because $h_k(\mu)$ tends to ∞ if μ tends to 0 or ∞. It is easily checked that for $c>c_k$ the equation $h_k(\mu)=c$ has two positive roots (and just one for $c=c_k$). Define $\mu_{k,c}$ to be the larger one. Also define $c_2=1$.

For $k\geq 2$ and $r\geq 3$, we are interested in k-cores of the random r-partite r-uniform hypergraph as presented in Definition 1. Let $\mathcal{G}_{r,n,m}$ denote the uniform probability space of r-partite r-uniform rn-vertex m-edge hypergraphs, as in that definition. Note that the k-cores of $\mathcal{G}_{r,n,m}$ form a probability space, which we denote by $\mathcal{K}_{r,n,m,k}$. Then, we can generalize the definition of c_k to

$$c_{r,k}=\inf\{h_{r,k}(\mu):\mu>0\},$$

where $h_{r,k}(\mu)=\frac{\mu}{(e^{-\mu}f_{k-1}(\mu))^{r-1}}$. As for $h_k(\mu)$, $h_{r,k}(\mu)$ tends to ∞ if μ tends to 0 or ∞, so $c_{r,k}$ is a positive real (and this applies even when $k=2$). Define $\mu_{r,k,c}$ to be the larger solution of $h_{r,k}(\mu)=c$ for $c>c_{r,k}$. We note that $c_{r,k}$ is the threshold of appearance of a k-core in a random r-partite r-uniform hypergraph that lies in $\mathcal{G}_{r,n,m}$.

1.2. Results

The threshold $c_{r,k}$ for appearance of a k-core for a random r-uniform hypergraph with cn/r edges was first analyzed in [5]. In this paper we instead analyze it for r-partite r-uniform hypergraphs. So these results will finish the analysis of the algorithm presented in [2].

We first consider the case when $r=2$ and $k=2$, which has been analyzed in [3] and is added here for completeness. The result is summarized in Theorem 1.

Theorem 1 ([3], Theorem 3.5) *Let $G_{2,n,m}=(V,E)$ be a random bipartite 2-uniform hypergraph with $2n$ vertices and $m=cn$ edges. Then, if $c=m/n$ holds for $c\in(0,1)$ and $n\to\infty$, the probability that $G_{2,n,m}$ has an empty 2-core component, for $n\to\infty$, is*

$$\Pr_a=\sqrt{1-c^2}. \tag{5}$$

Although it was not mentioned in [2], the version of the algorithm presented there for $G_{2,n,m}$ can be sped up by allowing a single cycle with length multiple of four per connected component. This happens because the probability of generating $G_{2,n,m}$ over the condition that it does not have any cycle with a length that is not a multiple of four is 58% higher than the probability of generating it when cycles are not permitted (see Sections 2 and 3) and the runtime of the algorithm is inversely proportional to this probability. To analyze this version of the algorithm it is required the following theorem.

Theorem 2 *Let $G_{2,n,m}=(V,E)$ be a random bipartite 2-uniform hypergraph with $2n$ vertices and $m=cn$ edges. Then, if $c=m/n$ holds for $c\in(0,1)$ and $n\to\infty$, the probability that $G_{2,n,m}$ has no cycle of length congruent to 2 (mod 4), for $n\to\infty$, is:*

$$\Pr_b=\frac{\sqrt{1-c^2}}{(1-c^4)^{\frac{1}{4}}}. \tag{6}$$

We remark that for $c>1$ there are a.a.s.(asymptotically almost surely) many cycles of all short even lengths.

We now consider the case when r and k are not both 2. Note that in $\mathcal{G}_{r,n,m}$, multiple edges, and edges containing repeated vertices, are permitted. However, even if they were forbidden, the same results would hold: it is well known by standard methods that the probability that there are no repeated vertices within any edge, and no multiple edges, is bounded away from 0 (see, e.g., [4, 14]). Hence, once we prove the theorem for this definition of G, it follows also for the other variations where multiple edges or repetitions of vertices within edges are forbidden.

Theorem 3 *Let $c>0$ and integers $r\geq 2$, $k\geq 2$ be fixed, where r and k are not both 2. Suppose that $m\sim cn$, and $G\in\mathcal{G}_{r,n,m}$. For $c<c_{r,k}$, G has empty k-core a.a.s. For $c>c_{r,k}$, the k-core of G a.a.s. has $e^{-\mu_{r,k,c}}f_k(\mu_{r,k,c})rn(1+o(1))$ vertices and $\mu_{r,k,c}e^{-\mu_{r,k,c}}f_{k-1}(\mu_{r,k,c})n(1+o(1))$ hyperedges. Moreover, let $j\geq k$ be fixed, and assume $c>c_{r,k}$. Then the number of vertices of degree j in $\mathcal{K}(r,n,m,k)$ is a.a.s. $rne^{-\mu}\mu^j/j!+o(n)$, where $\mu=\mu_{r,k,c}$.*

2. Proof of Theorem 1

Let $G_{2,n,m}=(V,E)$ be a bipartite random graph, where $|V|=2n$ and $|E|=m=cn$, where $c=m/n$ is the average degree of $G_{2,n,m}$. To build $G_{2,n,m}$ each edge is independently taken at random with probability p from all n^2 possible edges. As there are $2n$ vertices, and each is connected to an average of c edges, then we can conclude that $p=c/n=2c/|V|$. Let $\mathbb{C}_{2l}$ be the set of cycles of length $2l$ in the complete bipartite graph K_{2n}, for $l\geq 1$ and each n. A cycle in $\mathbb{C}_{2l}$ can be represented as a sequence of $2l$ distinct vertices in K_{2n} by choosing a starting point. Therefore, the cardinality of $\mathbb{C}_{2l}$ is given by

$$|\mathbb{C}_{2l}|=\frac{1}{2l}((n)_l)^2, \tag{7}$$

where $(n)_l=n(n-1)\ldots(n-l+1)$. As each edge in $G_{2,n,m}$ is selected independently of the others and with probability $p=\frac{c}{n}$, then, each cycle in $\mathbb{C}_{2l}$ occurs with probability

$$\Pr_{2l}(c) = p^{2l}. \tag{8}$$

Let $C_{2l}(G_{2,n,m})$ be a random variable that measures the number of cycles of length $2l$ in $G_{2,n,m}$. Let $C_e(G_{2,n,m})$ be a random variable that measures the number of cycles of any even length in $G_{2,n,m}$. The probability distribution of $C_{2l}(G_{2,n,m})$ and $C_e(G_{2,n,m})$ converges to a Poisson distribution with parameters λ_{2l} and λ_e, respectively. For a more detailed proof of a similar statement, see [11, Page 16]. To end the proof we are going to show how to get λ_{2l} and λ_e, which represents the average number of cycles of length $2l$ in $G_{2,n,m}$ and the average number of cycles of even length in $G_{2,n,m}$, respectively. It is easy to see that for $n \to \infty$

$$\lambda_{2l} = \Pr_{2l}(c) \times |\mathbb{C}_{2l}| = \left(\frac{c}{n}\right)^{2l} \frac{1}{2l}((n)_l)^2 = \frac{1}{2l} c^{2l} \tag{9}$$

and

$$\lambda_e = \sum_{l=1}^{\infty} \lambda_{2l} = \frac{1}{2}c^2 + \frac{1}{4}c^4 + \sum_{l=3}^{\infty} \frac{1}{2l} c^{2l} = -\frac{1}{2}\ln(1-c^2), \tag{10}$$

As in [11] we use $\sum_{l=3}^{\infty} \frac{1}{2l} x^l = -\frac{1}{2}\ln(1-x) - \frac{1}{2}x - \frac{1}{4}x^2$, where $x = c^2$. Therefore, the probability that $G_{2,n,m}$ is a forest and, consequently, has an empty 2-core is:

$$\Pr_a(C_e(G_{2,n,m}) = 0) = e^{-\lambda_e} = \sqrt{1-c^2}. \tag{11}$$

Note that c is restricted to be in the range $(0, 1)$ and, therefore, $c_{2,2} = 1$. ∎

This matches the experimental results presented in [2]. For instance, when $c = 2/2.09$ we have $\Pr_a = 0.29$. This is very close to 0.294 that is the value obtained in [2] by generating 1,000 random bipartite 2-graphs with $n = 10^7$ edges.

3. Proof of Theorem 2

Let $G_{2,n,m}$, $C_e(G_{2,n,m})$ and $c > 0$ be defined as in the proof of Theorem 1 presented in Section 2. From there we now that the random variable $C_e(G_{2,n,m})$ that measures the number of cycles of any even length in $G_{2,n,m}$ converges to a Poisson distribution with parameter:

$$\lambda_e = \sum_{l=1}^{\infty} \frac{1}{2l} c^{2l} = -\frac{1}{2}\ln\left(1-c^2\right). \tag{12}$$

Corresponding results hold for cycles with lengths in a given subset of $\{2, 4, 6, \dots\}$, as can be derived from the results of [11]. Those cycles with a length that is not a multiple of four cannot be used to build MPHFs in the algorithm presented in [2]. Accordingly, we let define such cycles to be *bad*, and let $C_b(G)$ be the random variable that measures the number of bad cycles in $G_{2,n,m}$. This converges to a Poisson distribution with parameter

$$\lambda_b = \sum_{l=1,3,5,7,\dots} \frac{1}{2l} c^{2l}. \tag{13}$$

From Eq. (12) we know that:

$$\begin{aligned}
\lambda_e &= \sum_{l=1,3,5,7,\dots} \frac{1}{2l} c^{2l} + \sum_{l=2,4,6,8,\dots} \frac{1}{2l} c^{2l} = -\frac{1}{2}\ln\left(1-c^2\right) \\
\lambda_b &= -\frac{1}{2}\ln\left(1-c^2\right) - \sum_{l=2,4,6,8,\dots} \frac{1}{2l} c^{2l} \\
&= -\frac{1}{2}\ln\left(1-c^2\right) - \frac{1}{2}\sum_{l=1}^{\infty} \frac{1}{2l}\left(c^2\right)^{2l} \\
&= -\frac{1}{2}\ln\left(1-c^2\right) + \frac{1}{4}\ln\left(1-c^4\right)
\end{aligned}$$

Therefore, the probability that $G_{2,n,m}$ has no bad cycle is given by:

$$\Pr_b(C_b(G) = 0) = e^{-\lambda_b} = \frac{\sqrt{1-c^2}}{(1-c^4)^{\frac{1}{4}}}.$$

Note that c is restricted to be in the range $(0, 1)$. ∎

For $c = 2.09$ we have $\Pr_b = 0.458$. Experimentally, we obtained $\Pr_b = 0.463$ by generating 1,000 random bipartite 2-graphs with $n = 10^7$ edges, which is very close to the theoretical value.

4. Proof of Theorem 3

Analogous results were proved for ordinary (not multipartite) hypergraphs in [4, 5, 15]. We will assume the reader is familiar with [4], and point out the simple modifications of its proof so as to cover the present setting. A *heavy* vertex is one of degree at least k, and a vertex is *light* otherwise. In the present setting, we have to pay attention to the number and total degree of the heavy vertices in each of the parts. So, for $j = 1, \dots, r$, let t_j denote the number of heavy vertices in U_j, let s_j denote the sum of their degrees, and let ℓ_j denote the sum of the degrees of the light vertices in U_j.

Given the vectors $\mathbf{t} = (t_1, \dots, t_r)$, $\mathbf{s} = (s_1, \dots, s_r)$ and $\mathbf{e} = (\ell_1, \dots, \ell_r)$, we consider a different model, called the hybrid model $\mathcal{P}$. This has vertex sets $V_1, \dots, V_r$ with $|V_j| = t_j$, and also separate vertex sets $L_1, \dots, L_r$ with $|L_j| = \ell_j$. Now select the m edges randomly by choosing for each edge one vertex in each of $L_j \cup V_j$, $j = 1, \dots, r$, conditioning on the total degree of vertices in V_j being s_j, for each j, and each vertex in L_j having degree exactly 1, and also conditioning on all vertices in V_j receiving degree at least k (for each j). This model represents a hypergraph $G_{\mathcal{P}}$ in which each vertex in L_j corresponds to the end of an edge at a light vertex. The model does not record which of the light vertices in $G_{\mathcal{P}}$ the edges are actually incident with.

The k-core of $G_{\mathcal{P}}$ has the same distribution as for $G_{r,n,m}$, conditioning on $\mathbf{t}$, $\mathbf{s}$ and $\mathbf{e}$. (For a detailed proof of the analogous statement in the general setting, see [4]; the proof is exactly the same in the present setting. The same reasoning gives the following claims.) Moreover, under the same conditioning, the distribution of vertex degrees in the vertices in V_j is precisely multinomial conditioned on all degrees being at least k. We call this distribution truncated multinomial.

The process of iteratively deleting light vertices from $G_{r,n,m}$ corresponds to the process applied to $G_{\mathcal{P}}$ whereby the vertices in the sets L_j and their incident edges are iteratively deleted, and any vertex in V_j whose degree reduces to $k' < k$ immediately transforms into k' new vertices in L_j. The degree distribution in V_j, conditioned on the values of s_j and t_j at each step of the process, is always truncated multinomial, which can be approximated by independent truncated Poisson.

Suppose that at each step of the process applied to $G_{\mathcal{P}}$, one light vertex is randomly selected from each of the sets $L_1, \ldots, L_r$, and deleted along with its incident edge. Let $S_{j,i}$ and $T_{j,i}$ denote the random values of s_j and t_j respectively after i steps of this process. Then deleting the edge containing a light vertex in L_j will have an effect on each of the other $r-1$ parts of the partition, in each part either deleting one light vertex or reducing the degree of a heavy vertex by 1. If in such a step, the degree of a vertex drops to $k-1$, it immediately fragments into $k-1$ light vertices. Thus, after i steps, each part will have total degree $m - ri$. The expected value of $T_{j,i+1} - T_{j,i}$, conditioned on the values of $\mathbf{s}$ and $\mathbf{t}$ at step i, is (assuming $m - i \to \infty$) asymptotically

$$-\frac{(r-1)S_{j,i}}{m-ri}\left(1 - \frac{\lambda_{j,i}T_{j,i}}{S_{j,i}}\right)$$

and for $S_{j,i+1} - S_{j,i}$ the expected change is asymptotically

$$-\frac{(r-1)S_{j,i}}{m-ri}\left(k - (k-1)\frac{\lambda_{j,i}T_{j,i}}{S_{j,i}}\right).$$

Here $\lambda_{j,i}$ is defined as $\lambda_{S_{j,i}/T_{j,i}}$ in the terminology of [4], and the asymptotic truncated Poisson distribution of the heavy vertex degrees is used.

The argument of [4] can now be followed from just below its Equation (12). In place of its Equations (15) and (16) we now have

$$z' = -\frac{(r-1)y}{c-rx}\left(1 - \frac{\mu z}{y}\right), \tag{14}$$

$$y' = -\frac{(r-1)y}{c-rx}\left(k - (k-1)\frac{\mu z}{y}\right) = (k-1)z' - 1, \tag{15}$$

and [21, Theorem 1] gives that a.a.s.

$$S_{j,i} = ny(i/n) + o(n) \quad \text{and} \quad T_{j,i} = nz(i/n) + o(n).$$

These are exactly the same equations as occur for hypergraphs in the proof of [4, Theorem 3]. Hence, the equations have the same solution, and the solution reaches a point where $c - rx - y = 0$ before $y = 0$ if and only if $c > c_{r,k}$. It follows that for $c > c_{r,k}$ the k-core exists a.a.s. Its size in each of the r parts is given asymptotically by yn, and the value of y when $c - rx - y = 0$ is the same as for the standard hypergraph case. Moreover, the asymptotic distribution of degrees is also the same, which gives the other statements in the theorem on the number of edges and number of vertices of given degrees.

For the lower bound, we use the same argument, which shows that for $c < c_{r,k}$, the k-core, if it exists, a.a.s. has size at most ϵn for any fixed $\epsilon > 0$. One can also observe that if ϵ is sufficiently small, then a random r-partite hypergraph with $cn + o(n)$ edges a.a.s. has no k-core of size less than ϵn. This comes from a simple first moment calculation, which shows that the expected number of sets of at most ϵn vertices that contain at least $k\epsilon n/r$ hyperedges tends to 0 as $n \to \infty$. ∎

Acknowledgements

We thank the partial support given by the Brazilian National Institute of Science and Technology for the Web (grant MCT/CNPq 573871/2008-6), Project InfoWeb (grant MCT/-CNPq/CT-INFO 550874/2007-0) and CNPq Grant 30.5237/02-0 (Nivio Ziviani).

References

[1] Bollobás, B., 2001. Random graphs, 2nd Edition. Vol. 73 of Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge.

[2] Botelho, F., Pagh, R., Ziviani, N., 2007. Simple and space-efficient minimal perfect hash functions. In: Proceedings of the 10th Workshop on Algorithms and Data Structures (WADs'07). Springer LNCS vol. 4619, pp. 139–150.

[3] Botelho, F., Pagh, R., Ziviani, N., 2007. Practical Perfect Hashing in Nearly Optimal Space. Information Systems (Submitted in 2010).

[4] Cain, J., Wormald, N. C., 2006. Encores on cores. Electronic Journal of Combinatorics 13 (1).

[5] Cooper, C., 2004. The cores of random hypergraphs with a given degree sequence. Random Structures & Algorithms 25 (4), 353–375.

[6] Erdős, P., Rényi, A., 1959. On random graphs I. Pub. Math. Debrecen 6, 290–297.

[7] Erdős, P., Rényi, A., 1960. On the evolution of random graphs. Magyar Tud. Akad. Mat. Kutató Int. Közl. 5, 17–61.

[8] Erdős, P., Rényi, A., 1961. On the evolution of random graphs. Bull. Inst. Internat. Statist. 38, 343–347.

[9] Erdős, P., Rényi, A., 1961. On the strength of connectedness of a random graph. Acta Mathematica Scientia Hungary 12, 261–267.

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

[11] Janson, S., 1987. Poisson convergence and poisson processes with applications to random graphs. Stochastic Processes and their Applications 26, 1–30.

[12] Janson, S., Łuczak, T., Ruciński, A., 1990. An exponential bound for the probability of nonexistence of a specified subgraph in a random graph. In: Random graphs '87 (Poznań, 1987). Wiley, Chichester, pp. 73–87.

[13] Janson, S., Łuczak, T., Ruciński, A., 2000. Random graphs. Wiley-Inter.

[14] Majewski, B., Wormald, N., Havas, G., Czech, Z., 1996. A family of perfect hashing methods. The Computer Journal 39 (6), 547–554.

[15] Molloy, M., 2005. Cores in random hypergraphs and boolean formulas. Random Structures & Algorithms 27 (1), 124–135.

[16] Molloy, M., Reed, B., 1995. A critical point for random graphs with a given degree sequence. Random Structures and Algorithms 6, 161–179.

[17] Molloy, M., Reed, B., 1998. The size of the giant component of a random graph with a given degree sequence. Combinatorics, Probability and Computing 7, 295–305.

[18] Pittel, B., Spencer, J., Wormald, N., 1996. Sudden emergence of a giant k-core in a random graph. Journal of Combinatorial Theory Series B 67 (1), 111–151.

[19] Pittel, B., Wormald, N. C., 2005. Counting connected graphs inside-out. Journal of Combinatorial Theory Series B 93 (2), 127–172.

[20] Radhakrishnan, J., 1992. Improved bounds for covering complete uniform hypergraphs. Information Processing Letters 41, 203–207.

[21] Wormald, N. C., 1995. Differential equations for random processes and random graphs. Annals of Applied Probability (5), 1217–1235.