Approximate counting of regular hypergraphs

Andrzej Dudek [a,*,1], Alan Frieze [b,2], Andrzej Ruciński [c,3], Matas Šileikis [d,4]

[a] *Department of Mathematics, Western Michigan University, Kalamazoo, MI, USA*
[b] *Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA, USA*
[c] *Department of Discrete Mathematics, Adam Mickiewicz University, Poznań, Poland*
[d] *Department of Mathematics, Uppsala University, Sweden*

ARTICLE INFO

ABSTRACT

In this paper we asymptotically count d-regular k-uniform hypergraphs on n vertices, provided k is fixed and $d = d(n) = o(n^{1/2})$. In doing so, we extend to hypergraphs a switching technique of McKay and Wormald.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

We consider *k-uniform hypergraphs* (or *k-graphs*, for short) on the vertex set $V = [n] := \{1, \dots, n\}$. A k-graph $H = (V, E)$ is d-regular, if the degree of every vertex $v \in V$, $\deg_H(v) := \deg(v) := |\{e \in E: v \in e\}|$ equals d.

Let $\mathcal{H}^{(k)}(n, d)$ be the class of all d-regular k-graphs on $[n]$. Note that each $H \in \mathcal{H}^{(k)}(n, d)$ has $m := nd/k$ edges (throughout, we implicitly assume that $k|nd$). We treat d as a function of n, possibly constant.

A result of McKay [8] contains an asymptotic formula for the number of n-vertex d-regular graphs, when $d \leqslant \varepsilon n$ for any constant $\varepsilon < 2/9$. In this paper we present an asymptotic enumeration of all d-regular k-graphs on a given set of n vertices, where $k \geqslant 3$ and $d = d(n)$ is ei-

ther a constant or does not grow with n too quickly. Let $\kappa = \kappa(k) = 1$ for $k \geqslant 4$ and $\kappa(3) = 1/2$.

Theorem 1. *For every $k \geqslant 3$, $1 \leqslant d = o(n^\kappa)$, and*

$$
\begin{aligned}
&\left|\mathcal{H}^{(k)}(n, d)\right| \\
&= \frac{(nd)!}{(nd/k)!(k!)^{nd/k}(d!)^n} \\
&\quad \times \exp\left\{-\frac{1}{2}(k-1)(d-1) + O\left((d/n)^{1/2} + d^2/n\right)\right\}.
\end{aligned}
$$

The error term in the exponent tends to zero (thus giving the asymptotics of $|\mathcal{H}^{(k)}(n, d)|$) if and only if $d = o(n^{1/2})$. Cf. an analogous formula for $k = 2$ by McKay [8], which gives the asymptotics if and only if $d = o(n^{1/3})$. Recently, Blinovsky and Greenhill [2] obtained more general results counting sparse uniform hypergraphs with given degrees.

Theorem 1 extends a result from [5] where Cooper, Frieze, Molloy and Reed proved that formula for d fixed using the by now standard *configuration model* (see [1,3, 11] for the graph case). Already for graphs, in [8], and later in [9] and [10], this technique was combined with the idea of *switchings*, a sequence of operations on a graph

* Corresponding author.
E-mail address: andrzej.dudek@wmich.edu (A. Dudek).

which eliminate loops and multiple edges, while keeping the degrees unchanged and leading to an *almost* uniform distribution of the simple graphs obtained as the ultimate outcome (but see Remark 3 in Section 3).

To prove Theorem 1 we apply these ideas together with a modification from [4], where instead of configurations, permutations were used to generate graphs with a given degree sequence. To describe this modification, consider a generalization of a k-graph in which edges are multisets of vertices rather than just sets. By a k-*multigraph* we mean a pair $H = (V, E)$ where V is a set and E is a multiset of k-element multisubsets of V. Thus we allow both multiple edges and loops, a *loop* being an edge which contains more than one copy of a vertex. We call an edge *proper* if it is not a loop. We say that a k-multigraph is *simple* if it is a k-graph, that is, if it contains neither multiple edges nor loops. Henceforth, for brevity of notation, we denote an edge of a k-multigraph by $v_1 \ldots v_k$ rather than $\{v_1, \ldots, v_k\}$.

Given a sequence $\mathbf{x} \in [n]^{ks}$, $s \in \mathbb{N}$, let $H(\mathbf{x})$ stand for the k-multigraph with edge multiset $E = \{x_{ki+1}, \ldots, x_{ki+k} : i = 0, \ldots, s-1\}$ and let $\lambda(\mathbf{x})$ be the number of loops in $H(\mathbf{x})$.

Let $\mathcal{P} = \mathcal{P}(n, d) \subset [n]^{nd}$ be the family of all permutations of the sequence

$$(\underbrace{1, \ldots, 1}_{d}, \underbrace{2, \ldots, 2}_{d}, \ldots, \underbrace{n, \ldots, n}_{d}).$$

Note that $|\mathcal{P}| = (nd)!(d!)^{-n}$. Let $\mathbf{Y} = (Y_1, \ldots, Y_{nd})$ be chosen uniformly at random from $\mathcal{P}$.

In the next section we sketch a proof of Theorem 1 together with some auxiliary results.

2. Proof of Theorem 1

2.1. Setup

Let $\mathcal{E}$ be the family of those permutations $\mathbf{y} \in \mathcal{P}$ for which the k-multigraph $H(\mathbf{y})$ has no multiple edges and contains at most

$$L := \sqrt{nd}$$

loops, but no loops with less than $k-1$ distinct vertices. Let

$$\mathcal{E}_l = \{\mathbf{y} \in \mathcal{E} : \lambda(\mathbf{y}) = l\}, \quad l = 0, \ldots, L.$$

Note that

$$\mathcal{E}_0 = \{\mathbf{y} \in \mathcal{P} : H(\mathbf{y}) \in \mathcal{H}^{(k)}(n, d)\}$$

is precisely the family of those permutations from $\mathcal{P}$ which represent simple k-graphs. In turn, for each $H \in \mathcal{H}^{(k)}(n, d)$ there are $(nd/k)!(k!)^{nd/k}$ permutations $\mathbf{y} \in \mathcal{E}_0$ with $H(\mathbf{y}) = H$. Therefore, in order to prove Theorem 1, it suffices to show that

$$|\mathcal{P}|/|\mathcal{E}_0|$$

$$= \exp\left\{\frac{1}{2}(k-1)(d-1) + O\left(\sqrt{d/n} + d^2/n\right)\right\}. \quad (1)$$

Our plan is as follows. First, in Proposition 2, we prove that

$$|\mathcal{P}| \sim \left(1 + O\left(\sqrt{d/n} + d^2/n^{k-2}\right)\right)|\mathcal{E}|. \quad (2)$$

Note that for $d = o(n^\kappa)$, the error term in (2) tends to zero and is at most the error term in (1). Thus, it is enough to show (1) with $|\mathcal{E}|$ in place of $|\mathcal{P}|$, which we do by writing

$$\frac{|\mathcal{E}|}{|\mathcal{E}_0|} = \sum_{l=0}^{L} \prod_{i=1}^{l} \frac{|\mathcal{E}_i|}{|\mathcal{E}_{i-1}|}, \quad (3)$$

and estimating the ratio $|\mathcal{E}_l|/|\mathcal{E}_{l-1}|$ uniformly for every $1 \leqslant l \leqslant L$.

In what follows it will be convenient to work directly with permutation $\mathbf{Y}$ rather than with the k-multigraph $H(\mathbf{Y})$ generated by it. Recycling the notation, we still call consecutive k-tuples $(Y_{ki+1}, \ldots, Y_{ki+k})$ of $\mathbf{Y}$ *edges*, *proper edges*, or *loops*, whatever appropriate. E.g., we say that $\mathbf{Y}$ contains *multiple edges*, if $H(\mathbf{Y})$ contains multiple edges, that is, some two edges of $\mathbf{Y}$ are identical as multisets. We use the standard notation $(x)_a = x(x-1)\cdots(x-a+1)$.

The following proposition implies (2), because $\mathbb{P}(\mathbf{Y} \in \mathcal{E}) = |\mathcal{E}|/|\mathcal{P}|$.

Proposition 2. *If* $k \geqslant 3$, *then* $\mathbb{P}(\mathbf{Y} \in \mathcal{E}) = 1 - O(\sqrt{d/n} + d^2/n^{k-2})$.

A simple proof of Proposition 2 (details can be found in the appendix of [6]) is based on the first moment method. In particular, the expected numbers of pairs of multiple edges, loops with less than $k-1$ distinct vertices, and all loops are, respectively, $O(d^2/n^{k-2})$, $O(d/n)$, and $\mathbb{E}\lambda(\mathbf{Y}) \sim \frac{k-1}{2}(d-1)$. The last formula implies that $\mathbb{P}(\lambda(\mathbf{Y}) > L) \leqslant \frac{\mathbb{E}\lambda(\mathbf{Y})}{L} = O(\sqrt{d/n})$.

2.2. Switchings

Now we define an operation, called *switching*, which generalizes to k-graphs a graph switching introduced in [9] (see also [10]). Permutations $\mathbf{y} \in \mathcal{E}_l$, $\mathbf{z} \in \mathcal{E}_{l-1}$ are said to be *switchable*, if $\mathbf{z}$ can be obtained from $\mathbf{y}$ by the following operation. From the edges of $\mathbf{y}$, choose a loop f and two proper edges e_1, e_2 that are disjoint from f and share at most $k-2$ vertices (see Fig. 1(a)). Letting $s = |e_1 \cap e_2|$, write

$$f = vvx_1 \ldots x_{k-2}, \qquad e_1 = w_1 \ldots w_s y_1 \ldots y_{k-s},$$

$$e_2 = w_1 \ldots w_s z_1 \ldots z_{k-s}.$$

Select vertices $y_* \in \{y_1, \ldots, y_{k-s}\}$ and $z_* \in \{z_1, \ldots, z_{k-s}\}$, and replace f, e_1, and e_2 by three proper edges

$$e_1' = e_1 \cup \{v\} - \{y_*\}, \qquad e_2' = e_2 \cup \{v\} - \{z_*\},$$

$$e_3' = f \cup \{y_*, z_*\} - \{v, v\}$$

as in Fig. 1(b). Since we are dealing with permutations, for definiteness let us assume that the procedure is performed by swapping with y_* the copy of v which appears in $\mathbf{y}$ further to the left and with z_* the one further to the right.

We can reconstruct permutations in $\mathcal{E}_{l+1}$ which are switchable with $\mathbf{y}$ as follows. Pick a vertex $v \in [n]$, two edges e_1', e_2' containing v, and one more edge e_3' (consult with Fig. 1 again). Choose a pair $\{y_*, z_*\}$ of vertices
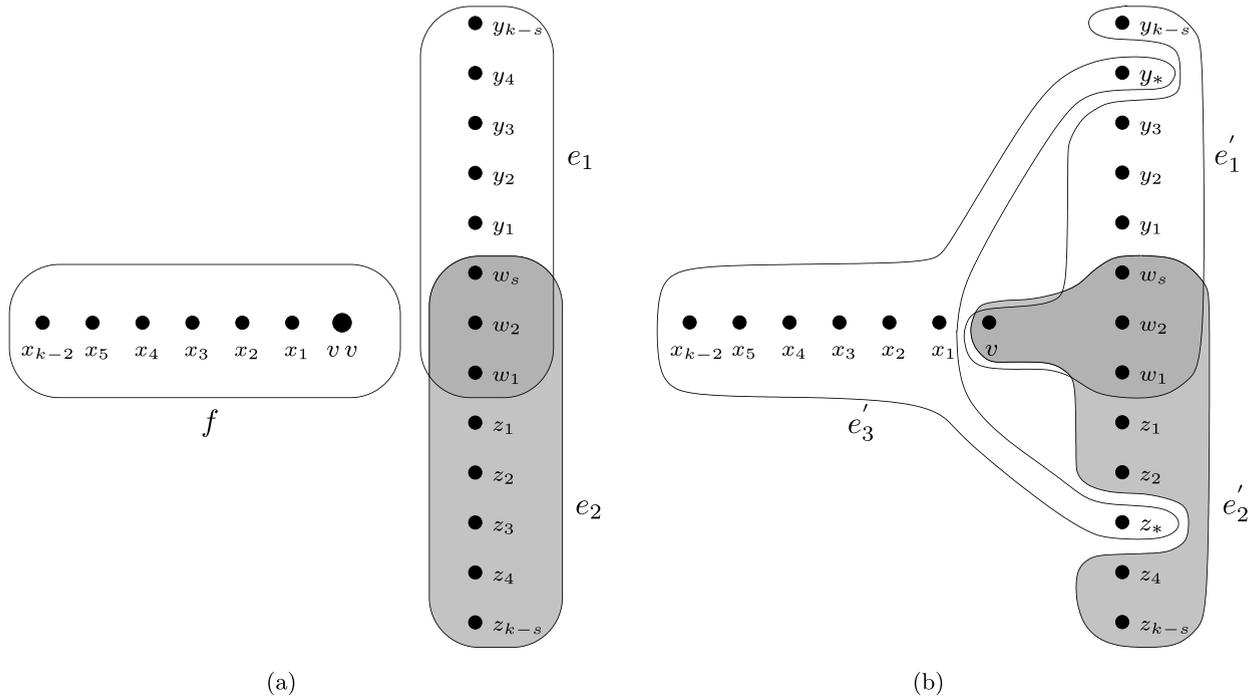
Fig. 1. Switching (a) before and (b) after.

from e'_3; replace e'_i, $i = 1, 2, 3$, by a loop and two edges defined as

$$f = e'_3 \cup \{v, v\} \setminus \{y_*, z_*\}, \qquad e_1 = e'_1 \cup \{y_*\} \setminus \{v\},$$

$$e_2 = e'_2 \cup \{z_*\} \setminus \{v\}.$$

Given $\mathbf{y} \in \mathcal{E}_l$, let $F(\mathbf{y})$ and $B(\mathbf{y})$ stand, respectively, for the number of ways to perform the forward and backward switching, or, in other words, the number of permutations $\mathbf{x} \in \mathcal{E}_{l-1}$ and $\mathbf{z} \in \mathcal{E}_{l+1}$ which are switchable with $\mathbf{y}$. Recall that $L = \sqrt{nd}$ and set $F_l = d^2 n^2 l$, $l = 1, \ldots, L$, and $B = \frac{k-1}{2} n^2 d^2 (d-1)$.

Proposition 3. *There is a sequence $\delta = \delta(n) = O((L + d^2)/dn)$ such that for all $\mathbf{y} \in \mathcal{E}_l$, $0 < l \leqslant L$*

$$(1 - \delta) F_l \leqslant F(\mathbf{y}) \leqslant F_l \quad \text{and} \quad (1 - \delta) B \leqslant B(\mathbf{y}) \leqslant B.$$

Proof. Clearly $F(\mathbf{y}) \leqslant lm^2 k^2 = n^2 d^2 l$. We say that two edges e', e'' of a k-graph are *distant* from each other if their distance in the intersection graph of $H(\mathbf{y})$ is at least three. Note that given f, e_1, and e_2, some choice of y_* and z_* might not yield a permutation $\mathbf{z} \in \mathcal{E}_{l-1}$, because one or more of e''_i's might already be present in $\mathbf{y}$. However, all k^2 choices of (y_*, z_*) are allowed, if $e_1 \cap e_2 = \emptyset$ and both e_1 and e_2 are distant from f. Therefore,

$$F(\mathbf{y}) \geqslant k^2 (m - l - 2k^2 d^2)^2 l$$
$$= k^2 m^2 l \big(1 - O\big((L + d^2)/m\big)\big).$$

Clearly $B(\mathbf{y}) \leqslant n(d)_2 m\binom{k}{2} = B$. To bound $B(\mathbf{y})$ from below, we estimate the number of choices of (v, e'_1, e'_2, e'_3), for

which at least one pair $\{y_*, z_*\}$ does not yield a permutation in $\mathcal{E}_{l+1}$. This can only happen when one of e'_1, e'_2, e'_3 is a loop, which occurs for at most $2kldm + ln(d)_2$ choices, or when e'_3 is not distant from both e'_1 and e'_2, which occurs for at most $n(d)_2 \cdot 2k^2 d^2$ choices. We have $B = \Theta(n^2 d^3)$, therefore

$$B(\mathbf{y}) \geqslant B - \binom{k}{2}\big(2kldm + ln(d)_2 + 2k^2 nd^4\big)$$

$$= B\left(1 - O\left(\frac{L + d^2}{nd}\right)\right). \qquad \square$$

Proof of Theorem 1. Counting the switchable pairs $\mathbf{y} \in \mathcal{E}_l$, $\mathbf{z} \in \mathcal{E}_{l-1}$ in two ways, from Proposition 3 we conclude that

$$\frac{(1 - \delta)B}{F_l} \leqslant \frac{|\mathcal{E}_l|}{|\mathcal{E}_{l-1}|} \leqslant \frac{B}{(1 - \delta)F_l}. \qquad (4)$$

Since $B/F_l = (k-1)(d-1)/2l$, from (3) and (4) we get

$$\sum_{l=0}^{L} \frac{x^l}{l!} \leqslant \frac{|\mathcal{E}|}{|\mathcal{E}_0|} \leqslant \sum_{l=0}^{L} \frac{y^l}{l!}$$

where $x = \frac{1}{2}(1 - \delta)(k-1)(d-1)$ and $y = \frac{1}{2}(k-1)(d-1)/(1 - \delta)$. Therefore by Taylor's theorem $|\mathcal{E}|/|\mathcal{E}_0|$ is at most e^y and at least

$$e^x\big(1 - x^L/L!\big) \geqslant e^x\big(1 - (ex/L)^L\big) = \exp\big\{x - o(\sqrt{d/n})\big\},$$

the inequality following from a standard fact $L! \geqslant (L/e)^L$. Since $x, y = (k-1)(d-1)/2 + O(\sqrt{d/n} + d^2/n)$, we get

$$\frac{|\mathcal{E}|}{|\mathcal{E}_0|} = \exp\left\{\frac{1}{2}(k-1)(d-1) + O\big(\sqrt{d/n} + d^2/n\big)\right\}$$

which together with (2) implies (1), hereby completing the proof. □

3. Concluding remarks

Remark 1. We believe that for $k = 3$ the constraint $d = o(n^{1/2})$ in Theorem 1 can be relaxed to $d = o(n)$ by allowing $O(d^2/n)$ multiple edges in $\mathbf{y} \in \mathcal{E}$ and applying an appropriate switching technique to eliminate them along with the loops.

Remark 2. In a forthcoming paper [7] we apply the switching technique presented here to embed asymptotically almost surely (*a.a.s.*) an ordinary Erdős–Rényi random k-graph $\mathbb{H}^{(k)}(n, m')$, $k \geqslant 3$, into a random d-regular k-graph $\mathbb{H}^{(k)}(n, d)$ for $d = \Omega(\log n)$, $d = o(\sqrt{n})$ and $m' = cnd/k$, for some constant $c > 0$. Consequently, *a.a.s.* $\mathbb{H}^{(k)}(n, d)$ inherits from $\mathbb{H}^{(k)}(n, m')$ all increasing properties held by the latter model.

Remark 3. An algorithm of McKay and Wormald [9] can be easily adapted to k-graphs, yielding an expected polynomial time uniform generation of d-regular k-graphs in $\mathcal{H}^{(k)}(n, d)$. The algorithm keeps selecting a random permutation $\mathbf{y} \in \mathcal{P}$ until $\mathbf{y} \in \mathcal{E}$. Then, iteratively, a random switching is applied $\lambda(\mathbf{y})$ times to eliminate all loops and finally yield a random element of $\mathcal{E}_0$. This leads to an *almost* uniform distribution over $\mathcal{H}^{(k)}(n, d)$. To make it *exactly* uniform, McKay and Wormald applied an ingenious trick of restarting the whole algorithm after every iteration of switching, say from $\mathbf{y} \in \mathcal{E}_l$ to $\mathbf{z} \in \mathcal{E}_{l-1}$, with probability $1 - (F(\mathbf{y})(1 - \delta_1)B)/(B(\mathbf{z})F_l) \leqslant 2\delta_1$. However, the assumption on d has to be strengthened, so that the reciprocal of the probability of not restarting the algorithm before its successful termination, or $(1 - \phi_k(n))^{-1}(1 - 2\delta_1(n))^{-L} = e^{O(\delta_1(n)L)}$, is at most a polynomial function of n. With our choice of L this imposes the bound $d = O(n^{1/3}(\log n)^{2/3})$. We may push it up to $d = O(\sqrt{n \log n})$ by redefining $L = kd + \omega(n)$ for any (sufficiently slow) sequence $\omega(n) \to \infty$. This change requires that in the last part of the proof of Proposition 2, instead of the first moment, Chebyshev's inequality is used (see the appendix of [6]).

References

[1] E. Bender, R. Canfield, The asymptotic number of labeled graphs with given degree sequences, J. Combinatorial Theory Ser. A 24 (3) (1978) 296–307.
[2] V. Blinovsky, C. Greenhill, Asymptotic enumeration of sparse uniform hypergraphs with given degrees, http://arxiv.org/abs/1306.2012.
[3] B. Bollobás, A probabilistic proof of an asymptotic formula for the number of labelled regular graphs, European J. Combin. 1 (4) (1980) 311–316.
[4] C. Cooper, A.M. Frieze, M. Krivelevich, Hamilton cycles in random graphs with a fixed degree sequence, SIAM J. Discrete Math. 24 (2) (2010) 558–569.
[5] C. Cooper, A.M. Frieze, M. Molloy, B. Reed, Perfect matchings in random r-regular, s-uniform hypergraphs, Combin. Probab. Comput. 5 (1) (1996) 1–14.
[6] A. Dudek, A. Frieze, A. Ruciński, M. Šileikis, Approximate counting of regular hypergraphs, http://arxiv.org/abs/1303.0400.
[7] A. Dudek, A. Frieze, A. Ruciński, M. Šileikis, Loose Hamilton cycles in random regular hypergraphs, Combin. Probab. Comput., in press.
[8] B.D. McKay, Asymptotics for symmetric 0–1 matrices with prescribed row sums, Ars Combin. 19 (1985) 15–25.
[9] B.D. McKay, N.C. Wormald, Uniform generation of random regular graphs of moderate degree, J. Algorithms 11 (1) (1990) 52–67.
[10] B.D. McKay, N.C. Wormald, Asymptotic enumeration by degree sequence of graphs with degrees $o(n^{1/2})$, Combinatorica 11 (4) (1991) 369–382.
[11] N.C. Wormald, The asymptotic distribution of short cycles in random regular graphs, J. Combin. Theory Ser. B 31 (2) (1981) 168–182.