Journal of Combinatorial Theory, Series B 122 (2017) 719–740

Contents lists available at ScienceDirect

Journal of Combinatorial Theory, Series B

www.elsevier.com/locate/jctb

Embedding the Erdős–Rényi hypergraph into the random regular hypergraph and Hamiltonicity

Andrzej Dudek [a,1,5], Alan Frieze [b,2], Andrzej Ruciński [c,3,5], Matas Šileikis [d,4,5]

[a] Department of Mathematics, Western Michigan University, Kalamazoo, MI, United States
[b] Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA, United States
[c] Department of Discrete Mathematics, Adam Mickiewicz University, Poznań, Poland
[d] Department of Applied Mathematics, Charles University, Prague, Czech Republic

ARTICLE INFO

Article history:
Received 27 August 2015
Available online 16 September 2016

Keywords:
Random regular graph
Random hypergraph
Hamilton cycle
Monotone graph property

ABSTRACT

We establish an inclusion relation between two uniform models of random k-graphs (for constant $k \geq 2$) on n labeled vertices: $\mathbb{G}^{(k)}(n,m)$, the random k-graph with m edges, and $\mathbb{R}^{(k)}(n,d)$, the random d-regular k-graph. We show that if $n \log n \ll m \ll n^k$ we can choose $d = d(n) \sim km/n$ and couple $\mathbb{G}^{(k)}(n,m)$ and $\mathbb{R}^{(k)}(n,d)$ so that the latter contains the former with probability tending to one as $n \to \infty$. This extends an earlier result of Kim and Vu about "sandwiching random graphs". In view of known threshold theorems on the existence of different types of Hamilton cycles in $\mathbb{G}^{(k)}(n,m)$, our result allows us to find conditions under which $\mathbb{R}^{(k)}(n,d)$ is Hamiltonian. In particular, for $k \geq 3$ we conclude that if

E-mail address: matas.sileikis@gmail.com (M. Šileikis).

[1] Project sponsored by the National Security Agency under Grant Number H98230-15-1-0172. The United States Government is authorized to reproduce and distribute reprints notwithstanding any copyright notation hereon.
[2] Supported in part by NSF grant CCF 1013110.
[3] Supported in part by the Polish NSC grant 2014/15/B/ST1/01688 and NSF grant DMS 1102086.
[4] Part of research performed at Uppsala University (Sweden) and the University of Oxford (United Kingdom).
[5] Part of research performed during a visit to the Institut Mittag-Leffler (Djursholm, Sweden).

http://dx.doi.org/10.1016/j.jctb.2016.09.003

$n^{k-2} \ll d \ll n^{k-1}$, then a.a.s. $\mathbb{R}^{(k)}(n,d)$ contains a tight Hamilton cycle.

1. Introduction

1.1. Background

A *k-uniform hypergraph* (or *k-graph* for short) on a vertex set $V = [n] = \{1, \ldots, n\}$ is an ordered pair $G = (V, E)$ where E is a family of k-element subsets of V. The *degree* of a vertex v in G is defined as

$$\deg_G(v) := |\{e \in E : v \in e\}|.$$

A k-graph is *d-regular* if the degree of every vertex is d. Let $\mathcal{R}^{(k)}(n,d)$ be the family of all d-regular k-graphs on V. Throughout, we tacitly assume that k divides nd. By $\mathbb{R}^{(k)}(n,d)$ we denote the *d-regular random k-graph* which is chosen uniformly at random from $\mathcal{R}^{(k)}(n,d)$.

Let us recall two more standard models of random k-graphs on n vertices. For $p \in [0,1]$, the *binomial random k-graph* $\mathbb{G}^{(k)}(n,p)$ is obtained by including each of the $\binom{n}{k}$ possible edges with probability p, independently of others. Further, for an integer $m \in [0, \binom{n}{k}]$, the *uniform random k-graph* $\mathbb{G}^{(k)}(n,m)$ is chosen uniformly at random among all $\binom{\binom{n}{k}}{m}$ k-graphs on V with precisely m edges.

We study the behavior of these random k-graphs as $n \to \infty$. Parameters d, m, p are treated as functions of n and typically tend to infinity in case of d, m, or zero, in case of p. Given a sequence of events $(\mathcal{A}_n)$, we say that $\mathcal{A}_n$ holds *asymptotically almost surely* (*a.a.s.*) if $\mathbb{P}(\mathcal{A}_n) \to 1$, as $n \to \infty$. Also, we write $a_n \ll b_n$ and $b_n \gg a_n$ for $a_n = o(b_n)$.

In 2004, Kim and Vu [11] proved that if $\log n \ll d \ll n^{1/3}/\log^2 n$ then there exists a coupling (that is, a joint distribution) of the random graphs $\mathbb{G}^{(2)}(n,p)$ and $\mathbb{R}^{(2)}(n,d)$ with $p = \frac{d}{n}\left(1 - O\left((\log n/d)^{1/3}\right)\right)$ such that

$$\mathbb{G}^{(2)}(n,p) \subset \mathbb{R}^{(2)}(n,d) \quad \text{a.a.s.} \tag{1}$$

They pointed out several consequences of this result, emphasizing the ease with which one can carry over known properties of $\mathbb{G}^{(2)}(n,p)$ to the harder to study regular model $\mathbb{R}^{(2)}(n,d)$. Kim and Vu conjectured that such a coupling is possible for all $d \gg \log n$ (they also conjectured a reverse embedding which is not of our interest here). In [7] we considered a (slightly weaker) extension of Kim and Vu's result to k-graphs, $k \geq 3$, and proved that

$$\mathbb{G}^{(k)}(n,m) \subset \mathbb{R}^{(k)}(n,d) \quad \text{a.a.s.} \tag{2}$$

whenever $C \log n \leq d \ll n^{1/2}$ and $m \sim cnd$ for some absolute large constant C and a sufficiently small constant $c = c(k) > 0$. Although (2) is stated for the uniform k-graph $\mathbb{G}^{(k)}(n, m)$, it is easy to see that one can replace $\mathbb{G}^{(k)}(n, m)$ by $\mathbb{G}^{(k)}(n, p)$ with $p = m/\binom{n}{k}$ (see Section 5).

1.2. The main result

In this paper we extend (2) to larger degrees, assuming only $d \leq cn^{k-1}$ for some constant $c = c(k)$. Moreover, our result implies that, provided $\log n \ll d \ll n^{k-1}$, we can take $m \sim nd/k$, that is, the embedded k-graph contains almost all edges of the regular k-graph rather than just a positive fraction, as in [7]. The new result is also valid for $k = 2$ (for the proof of this case alone, see also [10, Section 10.3]), and thus extends (1).

Theorem 1. *For each $k \geq 2$ there is a positive constant C such that if for some real $\gamma = \gamma(n)$ and positive integer $d = d(n)$,*

$$C\left(\left(d/n^{k-1} + (\log n)/d\right)^{1/3} + 1/n\right) \leq \gamma < 1, \tag{3}$$

and $m = (1-\gamma)nd/k$ is an integer, then there is a joint distribution of $\mathbb{G}^{(k)}(n, m)$ and $\mathbb{R}^{(k)}(n, d)$ with

$$\lim_{n\to\infty} \mathbb{P}\left(\mathbb{G}^{(k)}(n, m) \subset \mathbb{R}^{(k)}(n, d)\right) = 1.$$

Remark. In the assumption (3) of Theorem 1 the term $1/n$ can be omitted when $k \leq 7$. Indeed, the inequality of arithmetic and geometric means implies that

$$(d/n^{k-1} + (\log n)/d)^{1/3} \geq (2/n^{(k-1)/2})^{1/3} \geq 1/n.$$

For a given $k \geq 2$, *a k-graph property* is a family of k-graphs closed under isomorphisms. A k-graph property $\mathcal{P}$ is called *monotone increasing* if it is preserved by adding edges (but not necessarily by adding vertices, as the example of, say, perfect matching shows).

Corollary 2. *Let $\mathcal{P}$ be a monotone increasing property of k-graphs and $\log n \ll d \ll n^{k-1}$. If for some $m \leq (1-\gamma)nd/k$, where γ satisfies (3), $\mathbb{G}^{(k)}(n, m) \in \mathcal{P}$ a.a.s., then $\mathbb{R}^{(k)}(n, d) \in \mathcal{P}$ a.a.s.*

1.3. Comparison with the proof by Kim and Vu

Kim and Vu [11] proved (1) by analyzing a certain algorithm that generates a random graph $\mathbb{R}_A$, coupling it with $\mathbb{R}(n, d)$ so that $\mathbb{R}_A = \mathbb{R}(n, d)$ a.a.s., and then embedding $\mathbb{G}(n, p)$ into $\mathbb{R}_A$ a.s.s.

The algorithm can be described concisely as sequentially generated configuration model which rejects a chosen edge (with replacement), if it violates the simplicity of the graph. Note that the algorithm may run out of admissible edges before it produces a d-regular graph. Refining analysis of Steger and Wormald [17], Kim and Vu [11] proved a coupling of $\mathbb{R}_A$ and $\mathbb{R}(n,d)$ for $d \ll n^{1/3}/\log^2 n$. It is worth mentioning that in another paper Kim and Vu [12] proved, for $d = n^{1/3-\varepsilon}$ with arbitrary $\varepsilon > 0$, a slightly stronger statement that $\mathbb{R}_A$ is asymptotically uniform, that is,

$$\mathbb{P}\left(\mathbb{R}_A = G\right) = (1 + o(1))|\mathcal{R}(n,d)|^{-1} \tag{4}$$

uniformly over *all* $G \in \mathcal{R}(n,d)$. The last section in [12] reflects some beliefs that this result cannot be extended to d larger than $n^{1/3}$. Although for the coupling of $\mathbb{R}_A$ and $\mathbb{R}(n,d)$ it is enough to prove weaker uniformity, when (4) is allowed to fail for $o(|\mathcal{R}(n,d)|)$ graphs, an attempt to extend the approach of Kim and Vu did not seem to be very promising.

Another looming obstacle was the dependence of the proof of asymptotic uniformity in [11] on an asymptotic formula for $|\mathcal{R}(n,d)|$ due to McKay and Wormald [15], which is valid for $d \ll n^{1/2}$. The problem of asymptotically enumerating $\mathcal{R}(n,d)$ had been open in the range $n^{1/2} \le d \ll n/\log n$ since 1991 (see [15]).

In the present paper we avoid both explicit generation of random regular graphs and enumeration of regular graphs. Instead we embed $\mathbb{G}(n,m)$ directly into $\mathbb{R}(n,d)$. For this we show that if $\mathbb{R}(n,d)$ is revealed edge by edge (by first sampling the graph and then exposing its edges in a random order), then the conditional distribution of the next edge is nearly uniform over the complement of the current graph (unless we are close to the end).

Still, getting a fair estimate for the conditional distribution of the next edge is as hard as enumerating graphs with a given degree sequence. We deal with this issue by instead estimating ratios of (conditional) probabilities. This allows us to replace asymptotic enumeration by *relative* enumeration, by which we mean comparison of the number of ways to extend two graphs G_1, G_2 (differing just by two edges) to a d-regular graph.

In April 2016, well after the present paper was submitted, Wormald [18] announced a proof (as a joint result with Anita Liebenau) of asymptotic enumeration in the missing range of d. This makes it more likely that an approach relying on enumeration could lead to another proof of our result. However, we have not attempted this.

For the outline of our proof, see Subsection 1.5.

1.4. Hamilton cycles in hypergraphs

To show a more specific application of Theorem 1 we consider Hamilton cycles in random regular hypergraphs.

For integers $1 \le \ell < k$, define an *ℓ-overlapping cycle* (or *ℓ-cycle*, for short) as a k-graph in which, for some cyclic ordering of its vertices, every edge consists of k consecutive vertices, and every two consecutive edges (in the natural ordering of the edges induced by the ordering of the vertices) share exactly ℓ vertices. (For $\ell > k/2$ it implies, of course, that some nonconsecutive edges intersect as well.) A 1-cycle is called *loose* and a $(k-1)$-cycle is called *tight*. A spanning ℓ-cycle in a k-graph H is called an *ℓ-Hamilton* cycle. Observe that a necessary condition for the existence of an ℓ-Hamilton cycle is that n is divisible by $k - \ell$. We will assume this divisibility condition whenever relevant.

Let us recall the results on Hamiltonicity of random regular graphs, that is, the case $k = 2$. Asymptotically almost sure Hamiltonicity of $\mathbb{R}^{(2)}(n, d)$ was proved by Robinson and Wormald [16] for fixed $d \geq 3$, by Krivelevich, Sudakov, Vu and Wormald [13] for $d \geq n^{1/2} \log n$, and by Cooper, Frieze and Reed [3] for $C \leq d \leq n/C$ and some large constant C.

Much less is known for random hypergraphs. For the binomial models, the thresholds were found only recently. First, results on loose Hamiltonicity of $\mathbb{G}^{(k)}(n, p)$ were obtained by Frieze [8] (for $k = 3$), Dudek and Frieze [4] (for $k \geq 4$ and $2(k-1)|n$), and by Dudek, Frieze, Loh and Speiss [6] (for $k \geq 3$ and $(k-1)|n$). As usual, the asymptotic equivalence of the models $\mathbb{G}^{(k)}(n, p)$ and $\mathbb{G}^{(k)}(n, m)$ (see, e.g., Corollary 1.16 in [9]) allows us to reformulate the aforementioned results for the random k-graph $\mathbb{G}^{(k)}(n, m)$.

Theorem 3 *([8,4,6]). There is a constant $C > 0$ such that if $m \geq Cn \log n$, then a.a.s. $\mathbb{G}^{(3)}(n, m)$ contains a loose Hamilton cycle. Furthermore, for every $k \geq 4$ if $m \gg n \log n$, then a.a.s. $\mathbb{G}^{(k)}(n, m)$ contains a loose Hamilton cycle.*

From Theorem 3 and the older embedding result (2), in [7] we concluded that there is a constant $C > 0$ such that if $C \log n \leq d \ll n^{1/2}$, then a.a.s. $\mathbb{G}^{(3)}(n, d)$ contains a loose Hamilton cycle. Furthermore, for every $k \geq 4$ if $\log n \ll d \ll n^{1/2}$, then a.a.s. $\mathbb{R}^{(k)}(n, d)$ contains a loose Hamilton cycle.

Thresholds for ℓ-Hamiltonicity of $\mathbb{G}^{(k)}(n, m)$ in the remaining cases, that is, for $\ell \geq 2$, were recently determined by Dudek and Frieze [5] (see also Allen, Böttcher, Kohayakawa, and Person [1]).

Theorem 4 *([5]).*

(i) *If $k > \ell = 2$ and $m \gg n^2$, then a.a.s. $\mathbb{G}^{(k)}(n, m)$ is 2-Hamiltonian.*
(ii) *For all integers $k > \ell \geq 3$, there exists a constant C such that if $m \geq Cn^\ell$ then a.a.s. $\mathbb{G}^{(k)}(n, m)$ is ℓ-Hamiltonian.*

In view of Corollary 2, Theorems 3 and 4 immediately imply the following result that was already anticipated by the authors in [7].

Theorem 5.

(i) *There is a constant $C > 0$ such that if $C \log n \leq d \leq n^{k-1}/C$, then a.a.s. $\mathbb{R}^{(3)}(n, d)$ contains a loose Hamilton cycle. Furthermore, for every $k \geq 4$ there is a constant*

$C > 0$ such that if $\log n \ll d \leq n^{k-1}/C$, then a.a.s. $\mathbb{R}^{(k)}(n,d)$ contains a loose Hamilton cycle.

(ii) *For all integers $k > \ell = 2$ there is a constant C such that if $n \ll d \leq n^{k-1}/C$ then a.a.s. $\mathbb{R}^{(k)}(n,d)$ contains a 2-Hamilton cycle.*

(iii) *For all integers $k > \ell \geq 3$ there is a constant C such that if $Cn^{\ell-1} \leq d \leq n^{k-1}/C$ then a.a.s. $\mathbb{R}^{(k)}(n,d)$ contains an ℓ-Hamilton cycle.*

We conjecture that in the cases (ii) and (iii) (but not (i)) the assumed lower bound for d is actually a threshold for Hamiltonicity in $\mathbb{R}^{(k)}(n,d)$, see Section 5.

1.5. Structure of the paper

In the following section we define a k-graph process $(\mathbb{R}(t))_t$ which reveals edges of the random d-regular k-graph one at a time. Then we state a crucial Lemma 6, which says, loosely speaking, that unless we are very close to the end of the process, the conditional distribution of the $(t+1)$-th edge is approximately uniform over the complement of $\mathbb{R}(t)$. Based on Lemma 6, we show that a.a.s. $\mathbb{G}^{(k)}(n,m)$ can be embedded in $\mathbb{R}^{(k)}(n,d)$, by refining a coupling similar to the one we used in [7] and thus proving Theorem 1.

In Section 3 we prove auxiliary results needed in the proof of Lemma 6. They mainly reflect the phenomenon that a typical trajectory of the d-regular process $(\mathbb{R}(t))_t$ has concentrated local parameters. In particular, concentration of vertex degrees is deduced from a Chernoff-type inequality (the only "external" result used in the paper), while (one-sided) concentration of common degrees of sets of vertices is obtained by an application of the switching technique (a similar application appeared in [13]).

In Section 4 we prove Lemma 6. First we rephrase it as an enumerative problem (counting the number of d-regular extensions of a given k-graph). We prove Lemma 6 by estimating the ratio of the numbers of extensions of two k-graphs which differ just in two edges. For this we define two random multi-k-graphs (via the configuration model) and couple them using yet another switching.

2. Proof of Theorem 1

We often drop the superscript in notations like $\mathbb{G}^{(k)}$ and $\mathbb{R}^{(k)}$ whenever k is clear from the context.

Let K_n denote the complete k-graph on vertex set $[n]$. Recall the standard k-graph process $\mathbb{G}(t)$, $t = 0, \ldots, \binom{n}{k}$ which starts with the empty k-graph $\mathbb{G}(0) = ([n], \emptyset)$ and at each time step $t \geq 1$ adds an edge ε_t drawn from $K_n \setminus \mathbb{G}(t-1)$ uniformly at random. We treat $\mathbb{G}(t)$ as an *ordered k-graph* (that is, with an ordering of edges) and write

$$\mathbb{G}(t) = (\varepsilon_1, \ldots, \varepsilon_t), \qquad t = 0, \ldots, \binom{n}{k}.$$

Of course, the random uniform k-graph $\mathbb{G}(n,m)$ can be obtained from $\mathbb{G}(m)$ by ignoring the ordering of the edges.

Our approach is to represent $\mathbb{R}(n,d)$ as an outcome of another k-graph process which, to some extent, behaves similarly to $(\mathbb{G}(t))_t$. For this, generate a random d-regular k-graph $\mathbb{R}(n,d)$ and choose an ordering $(\eta_1,\dots,\eta_M)$ of its

$$M := \frac{nd}{k}$$

edges uniformly at random. Revealing the edges of $\mathbb{R}(n,d)$ in that order one by one, we obtain a *regular k-graph* process

$$\mathbb{R}(t) = (\eta_1,\dots,\eta_t), \qquad t = 0,\dots,M.$$

For every ordered k-graph G with t edges and every edge $e \in K_n \setminus G$ we clearly have

$$\mathbb{P}\left(\varepsilon_{t+1} = e \,|\, \mathbb{G}(t) = G\right) = \frac{1}{\binom{n}{k} - t}.$$

This is not true for $\mathbb{R}(t)$, except for the very first step $t = 0$. However, it turns out that for the most of the time, the conditional distribution of the next edge in the process $\mathbb{R}(t)$ is approximately uniform, which is made precise by the lemma below. To formulate it we need some more definitions.

Given an ordered k-graph G, let $\mathcal{R}_G(n,d)$ be the family of *extensions* of G, that is, ordered d-regular k-graphs the first edges of which are equal to G. More precisely, setting $G = (e_1,\dots,e_t)$,

$$\mathcal{R}_G(n,d) = \{H = (f_1,\dots,f_M) : f_i = e_i, i = 1,\dots,t, \text{ and } H \in \mathcal{R}^{(k)}(n,d)\}.$$

We say that a k-graph G with $t \le M$ edges is *admissible*, if $\mathcal{R}_G(n,d) \neq \emptyset$ or, equivalently, $\mathbb{P}\left(\mathbb{R}(t) = G\right) > 0$. We define, for admissible G,

$$p_{t+1}(e|G) := \mathbb{P}\left(\eta_{t+1} = e \,|\, \mathbb{R}(t) = G\right), \quad t = 0,\dots,M-1. \tag{5}$$

Given $\epsilon \in (0,1)$, we define events

$$\mathcal{A}_t = \left\{p_{t+1}(e|\mathbb{R}(t)) \ge \frac{1-\epsilon}{\binom{n}{k} - t} \text{ for every } e \in K_n \setminus \mathbb{R}(t)\right\}, \quad t = 0,\dots,M-1. \tag{6}$$

Now we are ready to state the main ingredient of the proof of Theorem 1.

Lemma 6. *Suppose that $\epsilon = \epsilon(n) \in (0,1)$ is such that $(1-\epsilon)M$ is an integer, and consider the event*

$$\mathcal{A} := \mathcal{A}_0 \cap \dots \cap \mathcal{A}_{(1-\epsilon)M-1}.$$

For every $k \ge 2$ there is a positive constant C' such that whenever ϵ and $d = d(n)$ satisfy

$$C'\left((d/n^{k-1}+(\log n)/d)^{1/3}+1/n\right)\leq\epsilon<1 \tag{7}$$

then the event $\mathcal{A}$ occurs a.a.s.

From Lemma 6, which is proved in Section 4, we deduce Theorem 1 using a coupling similar to the one which was used in [7].

Proof of Theorem 1. Clearly, we can pick $\epsilon \leq \gamma/3$ such that $(1-\epsilon)M$ is integer and (3) implies (7) with C' being some constant multiple of C.

Let us first outline the proof. We will define a k-graph process $\mathbb{R}'(t) := (\eta'_1, \ldots, \eta'_t)$, $t = 0, \ldots, M$ such that for every admissible k-graph G with $t \leq M-1$ edges,

$$\mathbb{P}\left(\eta'_{t+1}=e\,|\,\mathbb{R}'(t)=G\right)=p_{t+1}(e|G). \tag{8}$$

In view of (8), the distribution of $\mathbb{R}'(M)$ is the same as the one of $\mathbb{R}(M)$ and thus we can define $\mathbb{R}(n,d)$ as the k-graph $\mathbb{R}'(M)$ with order of edges ignored. Then we will show that a.a.s. $\mathbb{G}(n,m)$ can be sampled from the subhypergraph $\mathbb{R}'((1-\epsilon)M)$ of $\mathbb{R}'(M)$.

Now come the details. Set $\mathbb{R}'(0)$ to be an empty vector and define $\mathbb{R}'(t)$ inductively (for $t = 1, 2, \ldots$) as follows. Suppose that k-graphs $R_t = \mathbb{R}'(t)$ and $G_t = \mathbb{G}(t)$ have been exposed. Draw ε_{t+1} uniformly at random from $K_n \setminus G_t$ and, independently, generate a Bernoulli random variable ξ_{t+1} with the probability of success $1-\epsilon$. If event $\mathcal{A}_t$ has occurred, that is,

$$p_{t+1}(e|R_t)\geq\frac{1-\epsilon}{\binom{n}{k}-t}\qquad\text{for every}\qquad e\in K_n\setminus R_t, \tag{9}$$

then draw a random edge $\zeta_{t+1} \in K_n \setminus R_t$ according to the distribution

$$\mathbb{P}\left(\zeta_{t+1}=e|\mathbb{R}'(t)=R_t\right):=\frac{p_{t+1}(e|R_t)-(1-\epsilon)/(\binom{n}{k}-t)}{\epsilon}\geq 0,$$

where the inequality holds by (9). Observe also that

$$\sum_{e\in K_n\setminus R_t}\mathbb{P}\left(\zeta_{t+1}=e|\mathbb{R}'(t)=R_t\right)=1,$$

so ζ_{t+1} has a well-defined distribution. Finally, fix an arbitrary bijection $f_{R_t,G_t} : R_t\setminus G_t \to G_t \setminus R_t$ between the sets of edges and define

$$\eta'_{t+1}=\begin{cases}\varepsilon_{t+1}, & \text{if } \xi_{t+1}=1, \varepsilon_{t+1}\in K_n\setminus R_t,\\ f_{R_t,G_t}(\varepsilon_{t+1}), & \text{if } \xi_{t+1}=1, \varepsilon_{t+1}\in R_t,\\ \zeta_{t+1}, & \text{if } \xi_{t+1}=0.\end{cases} \tag{10}$$

If the event $\mathcal{A}_t$ fails, then η'_{t+1} is sampled directly (without defining ζ_{t+1}) according to probabilities (5). Such a definition of η'_{t+1} ensures that

$$\mathcal{A}_t \cap \{\xi_{t+1} = 1\} \quad \Longrightarrow \quad \varepsilon_{t+1} \in \mathbb{R}'(t+1). \tag{11}$$

Further, define a random subsequence of edges of $\mathbb{G}((1-\epsilon)M)$,

$$S := \{\varepsilon_i : \xi_i = 1\,, i \le (1-\epsilon)M\}\,.$$

Conditioning on the vector (ξ_i) determines $|S|$. If $|S| \ge m$, we define $\mathbb{G}(n,m)$ to have the edge set consisting of the first m edges of S (note that since the vectors (ξ_i) and (ε_i) are independent, these m edges are uniformly distributed), and if $|S| < m$, then we define $\mathbb{G}(n,m)$ as a graph with edges $\{\varepsilon_1, \dots, \varepsilon_m\}$.

Let event $\mathcal{A}$ be as in Lemma 6. The crucial thing is that by (11) we have

$$\mathcal{A} \quad \Longrightarrow \quad S \subset \mathbb{R}'(M).$$

Therefore

$$\mathbb{P}\left(\mathbb{G}(n,m) \subset \mathbb{R}(n,d)\right) \ge \mathbb{P}\left(\{|S| \ge m\} \cap \mathcal{A}\right).$$

Since by Lemma 6 event $\mathcal{A}$ holds a.a.s., to complete the proof it suffices to show that $\mathbb{P}\left(|S| < m\right) \to 0$.

To this end, note that $|S|$ is a binomial random variable, namely,

$$|S| = \sum_{i=1}^{(1-\epsilon)M} \xi_i \sim \mathrm{Bin}((1-\epsilon)M, 1-\epsilon),$$

with

$$\mathbb{E}|S| \ge (1-2\epsilon)M \quad \text{and} \quad \mathrm{Var}|S| = (1-\epsilon)^2 \epsilon M \le \epsilon M. \tag{12}$$

Recall that $\epsilon \le \gamma/3$ and thus $m = (1-\gamma)M \le (1-3\epsilon)M$. By (12), Chebyshev's inequality, and the inequality $\epsilon \ge C' \log n/d$, which follows from (7), we get

$$\mathbb{P}\left(|S| < m\right) \le \mathbb{P}\left(|S| - \mathbb{E}|S| < -\epsilon M\right) \le \frac{\epsilon M}{(\epsilon M)^2} = \frac{k}{\epsilon n d} \le \frac{k}{C' n \log n} \to 0. \qquad \square \tag{13}$$

3. Preparations for the proof of Lemma 6

Throughout this section we adopt the assumptions of Lemma 6, that is, $(1-\epsilon)M$ is an integer and (7) holds with a sufficiently large $C' = C'(k) \ge 1$. In particular,

$$\epsilon \geq C'(\log n/d)^{\alpha}, \tag{14}$$

$$\epsilon \geq C'(d/n^{k-1})^{\alpha} \tag{15}$$

for every $\alpha \geq 1/3$, and

$$\epsilon \geq C'/n. \tag{16}$$

Given a k-graph G with maximum degree at most d, let us define the *residual degree* of a vertex $v \in V(G)$ as

$$r_G(v) = d - \deg_G(v).$$

We begin our preparations toward the proof of Lemma 6 with a fact which allows one to control the residual degrees of the evolving k-graph $\mathbb{R}(t) = (\eta_1, \ldots, \eta_t)$. For a vertex $v \in [n]$ and $t = 0, \ldots, M$, define random variables

$$X_t(v) = r_{\mathbb{R}(t)}(v) = |\{i \in (t, M] : v \in \eta_i\}|.$$

Given an integer $t \in [0, M]$, we will use a shorthand

$$\tau = 1 - \frac{t}{M}.$$

We will usually assume $t \leq (1-\epsilon)M$, which implies $\tau \geq \epsilon$.

Claim 7. *For every $k \geq 2$ there is a constant $a = a(k) > 0$ such that a.a.s.*

$$\forall t \leq (1-\epsilon)M, \quad \forall v \in [n], \quad |X_t(v) - \tau d| \leq \sqrt{a\tau d \log n} \leq \tau d/2 - 1. \tag{17}$$

Proof. A crucial observation is that the concentration of the degrees depends solely on the random ordering of the edges and not on the structure of the k-graph $\mathbb{R}(M)$. If we fix a d-regular k-graph H and condition $\mathbb{R}(M)$ to be a random permutation of the edges of H, then $X_t(v)$ is a hypergeometric random variable with expectation

$$\mathbb{E}X_t(v) = \frac{(M-t)d}{M} = \tau d.$$

Using Theorem 2.10 in [9] together with inequalities (2.5) and (2.6) therein, we get

$$\mathbb{P}\left(|X_t(v) - \tau d| \geq x\right) \leq 2\exp\left\{-\frac{x^2}{2\tau d\,(1 + x/(3\tau d))}\right\}.$$

Let $a = 3(k+2)$ and $x = \sqrt{a\tau d \log n}$. Condition (14) with $\alpha = 1$ and $C' \geq 9a$ implies that

$$\tau d \geq \epsilon d \geq C' \log n. \tag{18}$$

Therefore

$$x/(\tau d) = \sqrt{a \log n/(\tau d)} \leq \sqrt{a \log n/(\epsilon d)} \leq \sqrt{a/C'} \leq 1/3. \tag{19}$$

Hence,

$$\mathbb{P}\left(|X_t(v) - \tau d| \geq \sqrt{a\tau d \log n}\right) \leq 2\exp\left\{-\frac{a}{3}\log n\right\} = 2n^{-k-2}.$$

Since we have fewer than $nM \leq n^{k+1}$ choices of t and v, the first inequality in (17) follows by taking the union bound.

The second inequality in (17) follows from (19), since

$$\sqrt{a\tau d \log n} = x \leq \tau d/3 \leq \tau d/2 - 1,$$

where the last inequality holds (for large enough n) by (18). □

Recall that $\mathcal{R}_G(n,d)$ is the family of extensions of G to a d-regular ordered k-graph. For a k-graph $H \in \mathcal{R}_G(n,d)$ define the *common degree* (relative to subhypergraph $G \subseteq H$) of an ordered pair (u,v) of vertices as

$$\operatorname{cod}_{H|G}(u,v) = \left|\left\{W \in \tbinom{[n]}{k-1} : W \cup u \in H, W \cup v \in H \setminus G\right\}\right|.$$

Note that $\operatorname{cod}_{H|G}(u,v)$ is not symmetric in u and v. Also, define *the degree of a pair of vertices* u, v as

$$\deg_H(u,v) = |\{e \in H : \{u,v\} \subset e\}|.$$

Claim 8. *Let G be an admissible k-graph with $t+1 \leq (1-\epsilon)M$ edges such that*

$$r_G(v) \leq 2\tau d \qquad \forall v \in [n]. \tag{20}$$

Suppose that $\mathbb{R}_G$ is a k-graph chosen uniformly at random from $\mathcal{R}_G(n,d)$. There are constants C_0, C_1, and C_2, depending on k only such that the following holds.

For each $e \in K_n \setminus G$,

$$\mathbb{P}(e \in \mathbb{R}_G) \leq \frac{C_0 \tau d}{n^{k-1}}. \tag{21}$$

Moreover, if $\ell \geq \ell_1 := C_1\tau d/n$, then for every $u, v \in [n]$, $u \neq v$,

$$\mathbb{P}\left(\deg_{\mathbb{R}_G \setminus G}(u,v) > \ell\right) \leq 2^{-(\ell-\ell_1)}. \tag{22}$$

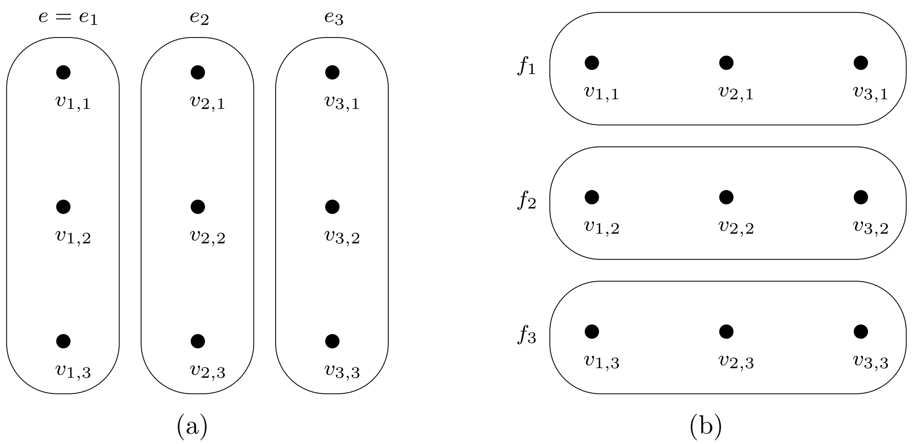

Fig. 1. Switching (for $k = 3$): before (a) and after (b).

Also, if $\ell \geq \ell_2 := C_2\tau d^2/n^{k-1}$, *then for every* $u, v \in [n]$, $u \neq v$,

$$\mathbb{P}\left(\mathrm{cod}_{\mathbb{R}_G|G}(u,v) > \ell\right) \leq 2^{-(\ell-\ell_2)}. \tag{23}$$

Proof. To prove (21), fix $e \in K_n \setminus G$ and define families of ordered k-graphs

$$\mathcal{R}_{e\in} = \{H \in \mathcal{R}_G(n,d) : e \in H\} \quad \text{and} \quad \mathcal{R}_{e\notin} = \{H \in \mathcal{R}_G(n,d) : e \notin H\}.$$

In order to compare the sizes of $\mathcal{R}_{e\in}$ and $\mathcal{R}_{e\notin}$, define an auxiliary bipartite graph B between $\mathcal{R}_{e\in}$ and $\mathcal{R}_{e\notin}$ in which $H \in \mathcal{R}_{e\in}$ is connected to $H' \in \mathcal{R}_{e\notin}$ whenever H' can be obtained from H by the following operation (known as *switching* in the literature dating back to McKay [14]). Let $e = e_1 = \{v_{1,1} \dots v_{1,k}\}$ and pick $k-1$ more edges

$$e_i = \{v_{i,1} \dots v_{i,k}\} \in H \setminus G, \qquad i = 2, \dots, k$$

(with vertices labeled in the increasing order within each edge) so that all k edges are disjoint. Replace, for each $j = 1, \dots, k$, the edge e_j by

$$f_j := \{v_{1,j} \dots v_{k,j}\}$$

to obtain H' (see Fig. 1).

Let $f(H) = \deg_B(H)$ be the number of k-graphs $H' \in \mathcal{R}_{e\notin}$ which can be obtained from H, and $b(H') = \deg_B(H')$ be the number of k-graphs $H \in \mathcal{R}_{e\in}$ from which H' can be obtained. Thus,

$$|\mathcal{R}_{e\in}| \cdot \min_{H \in \mathcal{R}_{e\in}} f(H) \leq |E(B)| \leq |\mathcal{R}_{e\notin}| \cdot \max_{H' \in \mathcal{R}_{e\notin}} b(H'). \tag{24}$$

Note that $H \setminus G$ and $H' \setminus G$ each have $\tau M - 1$ edges and, by (20), maximum degrees at most $2\tau d$. To estimate $f(H)$, note that because each edge intersects at most $k \cdot 2\tau d$ other edges of $H \setminus G$, the number of ways to choose an unordered $(k-1)$-tuple $\{e_2, \dots, e_k\}$ is at least

$$\frac{1}{(k-1)!}\prod_{i=1}^{k-1}(\tau M - 1 - ik \cdot 2\tau d) \geq (\tau M - k^2 \cdot 2\tau d)^{k-1}/(k-1)!. \tag{25}$$

We have to subtract the $(k-1)$-tuples which are not allowed since they would create a double edge after the switching (by repeating some edge of H which intersects e_1). Their number is at most $kd \cdot (2\tau d)^{k-1}$. Thus,

$$\begin{aligned} f(H) &\geq \frac{(\tau M - 2k^2\tau d)^{k-1}}{(k-1)!} - k(2\tau)^{k-1}d^k \\ &= \frac{(\tau M)^{k-1}}{(k-1)!}\left(\left(1 - \frac{2k^2 d}{M}\right)^{k-1} - \frac{k!(2\tau)^{k-1}d^k}{(\tau M)^{k-1}}\right) \\ &= \frac{(\tau M)^{k-1}}{(k-1)!}\left(\left(1 - \frac{2k^3}{n}\right)^{k-1} - \frac{k!(2k)^{k-1}d}{n^{k-1}}\right) \\ &\geq \frac{(\tau M)^{k-1}}{(k-1)!}\left(1 - \frac{2k^4}{n} - \frac{(2k)^{2k}d}{n^{k-1}}\right). \end{aligned}$$

By (15) with $\alpha = 1$, (16), and sufficiently large C', we have

$$\frac{2k^4}{n} + \frac{(2k)^{2k}d}{n^{k-1}} \leq \frac{\epsilon(2k^4 + (2k)^{2k})}{C'} \leq 1/2.$$

Hence,

$$f(H) \geq \frac{(\tau M)^{k-1}}{2(k-1)!}. \tag{26}$$

Since G is admissible, either $\mathcal{R}_{e\notin}$ or $\mathcal{R}_{e\in}$ is non-empty. If $\mathcal{R}_{e\in}$ is non-empty, then by (24) and the fact that the right-hand side of (26) is positive we get that $\mathcal{R}_{e\notin}$ is also non-empty.

In order to bound $b(H')$ from above note that there are at most $(2\tau d)^k$ ways to choose a sequence $f_1, \dots, f_k \in H' \setminus G$ such that $v_{1,i} \in f_i$ and we can reconstruct the $k-1$-tuple $e_2, \dots, e_k$ in at most $((k-1)!)^{k-1}$ ways (by fixing an ordering of vertices of f_1 and permuting vertices in other f_i's). Therefore $b(H') \leq ((k-1)!)^{k-1} \cdot (2\tau d)^k$. This, with (24) and (26) implies that

$$\mathbb{P}(e \in \mathbb{R}_G) = \frac{|\mathcal{R}_{e\in}|}{|\mathcal{R}_G(n,d)|} \leq \frac{|\mathcal{R}_{e\in}|}{|\mathcal{R}_{e\notin}|} \leq \frac{\max_{H' \in \mathcal{R}_{e\notin}} b(H')}{\min_{H \in \mathcal{R}_{e\in}} f(H)} \leq \frac{2((k-1)!)^k(2\tau d)^k}{(\tau M)^{k-1}} = \frac{C_0 \tau d}{n^{k-1}},$$

for some constant $C_0 = C_0(k)$. This concludes the proof of (21).

To prove (22), fix distinct $u, v \in [n]$, $u < v$, and define the families

$$\mathcal{R}_1(\ell) = \left\{ H \in \mathcal{R}_G(n,d) : \deg_{H\setminus G}(u,v) = \ell \right\}, \qquad \ell = 0, 1, \ldots.$$

Since G is admissible, $\mathcal{R}_G(n,d)$ is non-empty and thus $\mathcal{R}_1(\ell)$ is nonempty for some $\ell \geq 0$. Let L_1 be the largest such ℓ. From the argument below we will see that actually $\mathcal{R}_1(\ell)$ is non-empty for every $\ell = 0, \ldots, L_1$.

In order to compare sizes of $\mathcal{R}_1(\ell)$ and $\mathcal{R}_1(\ell-1)$, $\ell \in [1, L]$, we define the following switching which maps a k-graph $H \in \mathcal{R}_1(\ell)$ to a k-graph $H' \in \mathcal{R}_1(\ell-1)$. Select $e_1 \in H \setminus G$ contributing to $\deg_{H\setminus G}(u,v)$ and pick $k-1$ edges $e_2, \ldots, e_k \in H \setminus G$ so that $e_1, \ldots, e_k$ are disjoint. Writing $e_i = \{v_{i,1} \ldots v_{i,k}\}$, $i = 1, \ldots, k$ and, for definiteness, labeling vertices inside each e_i in the increasing order, replace $e_1, \ldots, e_k$ by $f_1, \ldots, f_k$, where $f_j = \{v_{1,j} \ldots v_{k,j}\}$, for $j = 1, \ldots, k$ (as in Fig. 1).

Noting that e_1 can be chosen in ℓ ways, we get a lower bound on $f(H)$ very similar to that in (26):

$$f(H) \geq \ell \left((\tau M - 2k^2 \tau d)^{k-1}/(k-1)! - k(2\tau)^{k-1} d^k \right) \geq \frac{\ell (\tau M)^{k-1}}{2(k-1)!}. \tag{27}$$

Since this implies $f(H) > 0$, we get that whenever $\mathcal{R}_1(\ell)$, $\ell \geq 1$, then also $\mathcal{R}_1(\ell-1)$ is non-empty. Thus, $\mathcal{R}_1(\ell)$ is non-empty for every $\ell = 0, \ldots, L_1$, as mentioned above.

For the upper bound for $b(H')$ we choose two disjoint edges in $H' \setminus G$ containing u and v, respectively, and then $k-2$ more edges in $H' \setminus G$ not containing u and v so that all edges are disjoint. Crudely bounding number of permutations of vertices inside each of $f_1, \ldots, f_k$ by $(k!)^k$, we get $b(H') \leq (k!)^k (2\tau d)^2 (\tau M)^{k-2}$. We obtain, for $\ell = \ell_1, \ldots, L_1$,

$$\frac{|\mathcal{R}_1(\ell)|}{|\mathcal{R}_1(\ell-1)|} \leq \frac{\max_{H' \in \mathcal{R}_1(\ell-1)} b(H')}{\min_{H \in \mathcal{R}_1(\ell)} f(H)} \leq \frac{2(k!)^{k+1}(2\tau d)^2 (\tau M)^{k-2}}{\ell (\tau M)^{k-1}} \leq \frac{8(k!)^{k+1}\tau d}{\ell n} \leq \frac{1}{2},$$

by assumption $\ell \geq \ell_1 = C_1 \tau d/n$ and appropriate choice of constant C_1. Further,

$$\begin{aligned} \mathbb{P}\left(\deg_{\mathbb{R}_G \setminus G}(u,v) > \ell \right) &= \sum_{i=\ell+1}^{L_1} \frac{|\mathcal{R}_1(i)|}{|\mathcal{R}_G(n,d)|} \leq \sum_{i=\ell+1}^{L_1} \frac{|\mathcal{R}_1(i)|}{|\mathcal{R}_1(\ell_1)|} \\ &= \sum_{i=\ell+1}^{L_1} \prod_{j=\ell_1+1}^{i} \frac{|\mathcal{R}_1(j)|}{|\mathcal{R}_1(j-1)|} \leq \sum_{i>\ell} 2^{-(i-\ell_1)} = 2^{-(\ell-\ell_1)}, \end{aligned} \tag{28}$$

which completes the proof of (22).

It remains to show (23). Fix an ordered pair (u, v) of distinct vertices and define the families

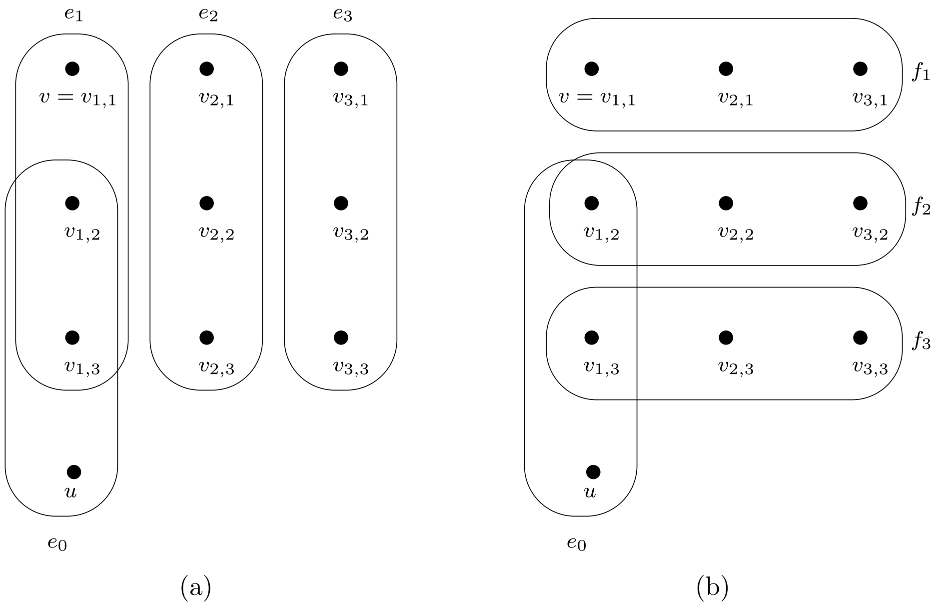

Fig. 2. Switching (for $k = 3$): before (a) and after (b).

$$\mathcal{R}_2(\ell) = \left\{H \in \mathcal{R}_G(n, d) : \mathrm{cod}_{H|G}(u, v) = \ell\right\}, \qquad \ell = 0, 1, \ldots .$$

We compare sizes of $\mathcal{R}_2(\ell)$ and $\mathcal{R}_2(\ell - 1)$ using the following switching. Select two distinct edges $e_0 \in H$ and $e_1 \in H \setminus G$ contributing to $\mathrm{cod}_{H|G}(u, v)$, that is, such that $e_0 \setminus \{u\} = e_1 \setminus \{v\}$; pick $k-1$ other edges $e_2, \ldots, e_k \in H \setminus G$ so that $e_1, \ldots, e_k$ are disjoint. Writing $e_i = \{v_{i,1} \ldots v_{i,k}\}$, $i = 1, \ldots, k$ with $v = v_{1,1}$, replace $e_1, \ldots, e_k$ by $f_1, \ldots, f_k$, where $f_j = \{v_{1,j} \ldots v_{k,j}\}$ for $j = 1, \ldots, k$ (see Fig. 2). We estimate $f(H)$ by first fixing a pair e_0, e_1 in one of ℓ ways. The number of choices of $e_2, \ldots, e_k$ is bounded as in (25). However, we subtract not just at most $kd \cdot (2\tau d)^{k-1}$ $(k-1)$-tuples which may create double edges, but also $(k-1)$-tuples for which $(f_1 \setminus \{v\}) \cup \{u\} \in H$ which prevents $\mathrm{cod}(u, v)$ from being decreased. There are at most $d \cdot (2\tau d)^{k-1}$ of such $(k-1)$-tuples. Hence the bound is very similar to (27) and, omitting very similar calculations, we get

$$f(H) \geq \ell \left(\frac{(\tau M - k^2 \cdot 2\tau d)^{k-1}}{(k-1)!} - (k+1)d \cdot (2\tau d)^{k-1} \right) \geq \frac{\ell (\tau M)^{k-1}}{2(k-1)!}.$$

Writing L_2 for the largest ℓ such that $\mathcal{R}_2(\ell)$ is non-empty, we get that $\mathcal{R}_2(\ell)$ is nonempty for $\ell = 0, \ldots, L_2$, by a similar argument as with the previous switching.

Conversely, H can be reconstructed from H' by choosing an edge $e_0 \in H'$ containing u but not containing v and then k disjoint edges $f_j \in H' \setminus G$, each containing exactly one vertex from $(e_0 \setminus \{u\}) \cup \{v\}$ and permuting the vertices inside $f_2 \setminus \{v_{1,2}\}, \ldots, f_k \setminus \{v_{1,k}\}$ in at most $((k-1)!)^{k-1}$ ways. Therefore $b(H') \leq ((k-1)!)^{k-1} d(2\tau d)^k$. Clearly, for $\ell = \ell_2, \ldots, L_2$,

$$\frac{|\mathcal{R}_2(\ell)|}{|\mathcal{R}_2(\ell-1)|} \le \frac{\max_{H' \in \mathcal{R}_2(\ell-1)} b(H')}{\min_{H \in \mathcal{R}_2(\ell)} f(H)} \le \frac{d(2\tau d)^k \cdot 2((k-1)!)^k}{\ell(\tau M)^{k-1}}$$
$$\le \frac{2^{k+1}((k-1)!)^k k^{k-1} \tau d^2}{n^{k-1}\ell} \le \frac{1}{2}$$

by the assumption $\ell \ge \ell_2 = C_2 \tau d^2 / n^{k-1}$ and appropriate choice of constant C_2. Now (23) follows from similar computations to (22).

This finishes the proof of Claim 8. $\square$

4. Proof of Lemma 6

In this section we prove the crucial Lemma 6. In view of Claim 7 it suffices to show that

$$\mathbb{P}\left(\eta_{t+1} = e \,|\, \mathbb{R}(t) = G\right) \ge \frac{1-\epsilon}{\binom{n}{k} - t}, \qquad \forall\, e \in K_n \setminus G, \tag{29}$$

for every $t \le (1-\epsilon)M - 1$ and every admissible G such that

$$d(\tau - \delta) \le r_G(v) \le d(\tau + \delta), \qquad v \in [n], \tag{30}$$

where

$$\tau = 1 - t/M \qquad \text{and} \qquad \delta = \sqrt{a\tau(\log n)/d}.$$

In some cases the following simpler bounds (implied by the second inequality in (17)) on $r_G(v)$ will suffice:

$$\tau d/2 + 1 \le r_G(v) \le 2\tau d, \qquad v \in [n]. \tag{31}$$

Since the average of $\mathbb{P}\left(\eta_{t+1} = e \,|\, \mathbb{R}(t) = G\right)$ over $e \in K_n \setminus G$ is exactly $1/\left(\binom{n}{k} - t\right)$, there is $f \in K_n \setminus G$ such that

$$\mathbb{P}\left(\eta_{t+1} = f \,|\, \mathbb{R}(t) = G\right) \ge \frac{1}{\binom{n}{k} - t}. \tag{32}$$

Fix any such f and let $e \in K_n \setminus G$ be arbitrary. We write $G \cup f$ for an ordered graph obtained by appending edge f at the end of G. Setting $\mathcal{R}_f := \mathcal{R}_{G \cup f}(n,d)$ and $\mathcal{R}_e := \mathcal{R}_{G \cup e}(n,d)$, we have

$$\frac{\mathbb{P}\left(\eta_{t+1} = e \,|\, \mathbb{R}(t) = G\right)}{\mathbb{P}\left(\eta_{t+1} = f \,|\, \mathbb{R}(t) = G\right)} = \frac{|\mathcal{R}_{G \cup e}(n,d)|}{|\mathcal{R}_{G \cup f}(n,d)|} = \frac{|\mathcal{R}_e|}{|\mathcal{R}_f|}. \tag{33}$$

To bound this ratio, we need to appeal to the configuration model for hypergraphs. Let $\mathbb{M}_G(n,d)$ be a random *multi-k-graph extension* of G to an ordered d-regular multi-k-graph. Namely, $\mathbb{M}_G(n,d)$ is a sequence of M edges (each of which is a k-element multiset of vertices), the first t of which comprise G, while the remaining ones are generated by taking a random uniform permutation Π of the multiset

$$\{1,\dots,1,\dots,n,\dots,n\}$$

with multiplicities $r_G(v)$, $v\in[n]$, and splitting it into consecutive k-tuples.

The number N_G of such permutations is a multinomial coefficient:

$$N_G := \binom{k(M-t)}{r_G(1),\dots,r_G(n)} = \frac{(k(M-t))!}{\prod_{v\in[n]} r_G(v)!}.$$

A *loop* is an edge with at least one repeated vertex. We say that an extension is *simple*, if all its edges are distinct and not loops.

Since each simple extension of G is given by the same number of permutations (namely $(k!)^{M-t}$), $\mathbb{M}_G(n,d)$ is uniform over $\mathcal{R}_G(n,d)$. That is, $\mathbb{M}_G(n,d)$, conditioned on simplicity, has the same distribution as $\mathbb{R}_G(n,d)$.

Set

$$\mathbb{M}_e = \mathbb{M}_{G\cup e}(n,d) \quad \text{and} \quad \mathbb{M}_f = \mathbb{M}_{G\cup f}(n,d),$$

for convenience. Noting that $G\cup f$ has $t+1$ edges, we have

$$\mathbb{P}\left(\mathbb{M}_f\in\mathcal{R}_f\right) = \frac{|\mathcal{R}_f|(k!)^{M-t-1}}{N_{G\cup f}} = \frac{|\mathcal{R}_f|(k!)^{M-t-1}\prod_{v\in[n]} r_{G\cup f}(v)!}{(k(M-t-1))!},$$

and similarly for $\mathbb{M}_e$ and $\mathcal{R}_e$. This yields, after a few cancellations, that

$$\frac{|\mathcal{R}_e|}{|\mathcal{R}_f|} = \frac{\prod_{v\in e\setminus f} r_G(v)}{\prod_{v\in f\setminus e} r_G(v)} \cdot \frac{\mathbb{P}\left(\mathbb{M}_e\in\mathcal{R}_e\right)}{\mathbb{P}\left(\mathbb{M}_f\in\mathcal{R}_f\right)}. \tag{34}$$

The ratio of the products in (34) is, by (30), at least

$$\left(\frac{\tau-\delta}{\tau+\delta}\right)^k \ge \left(1-\frac{2\delta}{\tau}\right)^k \ge 1-2k\sqrt{\frac{a\log n}{\tau d}} \ge 1-2k\sqrt{\frac{a\log n}{\epsilon d}} \ge 1-\epsilon/2,$$

where the last inequality holds by (14) with $\alpha=1/3$ and $C'\ge\sqrt[3]{16ak^2}$. On the other hand, the ratio of probabilities in (34) will be shown in Claim 9 below to be at least $1-\epsilon/2$. Consequently, the entire ratio in (34), and thus in (33), will be at least $1-\epsilon$, which, in view of (32), will imply (29) and yield the lemma.

Hence, to complete the proof of Lemma 6 it remains to show that the probabilities of simplicity $\mathbb{P}\left(\mathbb{M}_e\in\mathcal{R}_e\right)$ are asymptotically the same for all $e\in K_n\setminus G$. Recall that for every edge $e\in K_n\setminus G$ we write

$$\mathbb{M}_e = \mathbb{M}_{G\cup e}(n,d) \quad \text{and} \quad \mathcal{R}_e = \mathcal{R}_{G\cup e}(n,d). \tag{35}$$

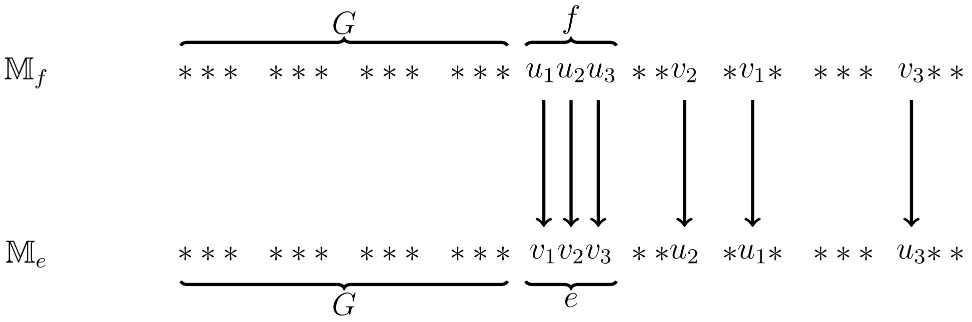

Fig. 3. Obtaining $\mathbb{M}_e$ from $\mathbb{M}_f$ for $k = s = 3$ by altering the underlying permutation.

Claim 9. *If G, e, and f are as above, then, for every $e \in K_n \setminus G$,*

$$\frac{\mathbb{P}\left(\mathbb{M}_e \in \mathcal{R}_e\right)}{\mathbb{P}\left(\mathbb{M}_f \in \mathcal{R}_f\right)} \geq 1 - \epsilon/2.$$

Proof. We start by constructing a coupling of $\mathbb{M}_e$ and $\mathbb{M}_f$ in which they differ in at most $k+1$ edges (counting in the replacement of f by e at the $(t+1)$-th position).

Let $f = \{u_1, \ldots, u_k\}$ and $e = \{v_1, \ldots, v_k\}$. Further, let $s = k - |f \cap e|$ and suppose without loss of generality that $\{u_1, \ldots, u_s\} \cap \{v_1, \ldots, v_s\} = \emptyset$. Let Π_f be a random permutation underlying the multi-k-graph $\mathbb{M}_f$. Note that Π_f differs from any permutation Π_e underlying $\mathbb{M}_e$ by having the multiplicities of $v_1, \ldots, v_s$ greater by one, and the multiplicities of $u_1, \ldots, u_s$ smaller by one than the corresponding multiplicities in Π_e.

Let Π^* be obtained from Π_f by replacing, for each $i = 1, \ldots, s$, a copy of v_i selected uniformly at random by u_i. Define $\mathbb{M}^*$ by chopping Π^* into consecutive k-tuples and appending them to $G \cup e$ (see Fig. 3).

It is easy to see that Π^* is uniform over all permutations of the multiset

$$\{1, \ldots, 1, \ldots, n, \ldots, n\}$$

with multiplicities $r_{G \cup e}(v)$, $v \in [n]$. This means that $\mathbb{M}^*$ has the same distribution as $\mathbb{M}_e$ and thus we will further identify $\mathbb{M}^*$ and $\mathbb{M}_e$.

Observe that if we condition $\mathbb{M}_f$ on being a simple k-graph H, then $\mathbb{M}_e$ can be equivalently obtained by the following switching: (i) replace edge f by e; (ii) for each $i = 1, \ldots, s$, choose, uniformly at random, an edge $e_i \in H \setminus (G \cup f)$ incident to v_i and replace it by $(e_i \setminus \{v_i\}) \cup \{u_i\}$ (see Fig. 4). Of course, some of e_i's may coincide. For example, if $e_{i_1} = \cdots = e_{i_l}$, then the effect of the switching is that e_{i_1} is replaced by $(e_{i_1} \setminus \{v_{i_1}, \ldots, v_{i_l}\}) \cup \{u_{i_1}, \ldots, u_{i_l}\}$.

The crucial idea is that such a switching is unlikely to create loops or multiple edges. However, for certain H this might not true. For example, if $e \in H \setminus (G \cup f)$, then the random choice of e_i's in step (ii) is unlikely to destroy e, but in step (i) edge f has been deterministically replaced by an additional copy of e, thus creating a double edge.

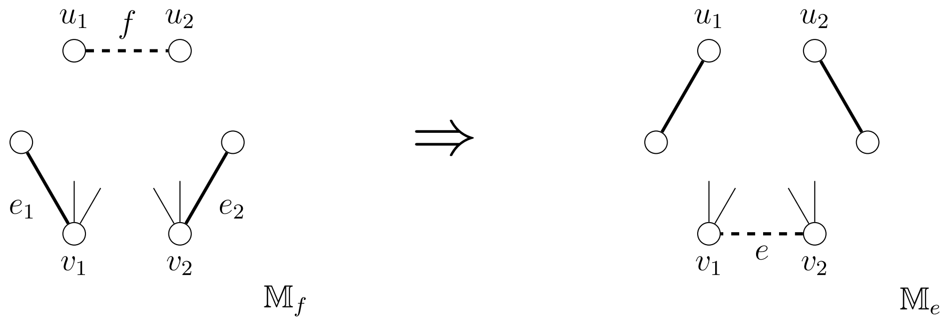

Fig. 4. Obtaining $\mathbb{M}_e$ from $\mathbb{M}_f$ for $k = s = 2$: only relevant edges are displayed; the ones belonging to $\mathbb{M}_f \setminus (G \cup f)$ are shown as solid lines.

Moreover, if almost every $(k-1)$-tuple of vertices extending v_i to an edge in $H \setminus (G \cup f)$ also extends u_i to an edge in H, then most likely the replacement of v_i by u_i will create a double edge, too. To avoid such and other bad instances, we say that $H \in \mathcal{R}_f$ is *nice* if the following three properties hold:

$$e \notin H \tag{36}$$

$$\max_{i=1,\dots,s} \deg_{H \setminus (G \cup f)}(u_i, v_i) \le \ell_1 + k \log_2 n, \tag{37}$$

$$\max_{i=1,\dots,s} \operatorname{cod}_{H | G \cup f}(u_i, v_i) \le \ell_2 + k \log_2 n, \tag{38}$$

where $\ell_1 = C_1 \tau d/n$ and $\ell_2 = C_2 \tau d^2/n^{k-1}$ are as in Claim 8. Note that $\mathbb{M}_f$, conditioned on $\mathbb{M}_f \in \mathcal{R}_f$, is distributed uniformly over $\mathcal{R}_{G \cup f}(n, d)$. Since we chose f such that by (32) is satisfied, we have that k-graph $G \cup f$ is admissible. Therefore by Claim 8 we have

$$\begin{aligned} \mathbb{P}\left(\mathbb{M}_f \text{ is not nice} \,|\, \mathbb{M}_f \in \mathcal{R}_f\right) &\le \frac{C_0 \tau d}{n^{k-1}} + 2 \cdot s 2^{-k \log_2 n} \\ &\le \frac{C_0 d + 2k}{n^{k-1}} \le \frac{\epsilon}{4}, \end{aligned} \tag{39}$$

where the last inequality follows by (15) with $\alpha = 1$ and sufficiently large constant C'. We have

$$\begin{aligned} \frac{\mathbb{P}\left(\mathbb{M}_e \in \mathcal{R}_e\right)}{\mathbb{P}\left(\mathbb{M}_f \in \mathcal{R}_f\right)} &\ge \mathbb{P}\left(\mathbb{M}_e \in \mathcal{R}_e \,|\, \mathbb{M}_f \in \mathcal{R}_f\right) \\ &\ge \mathbb{P}\left(\mathbb{M}_e \in \mathcal{R}_e \,|\, \mathbb{M}_f \text{ is nice}\right) \mathbb{P}\left(\mathbb{M}_f \text{ is nice} \,|\, \mathbb{M}_f \in \mathcal{R}_f\right). \end{aligned} \tag{40}$$

It suffices to show that

$$\mathbb{P}\left(\mathbb{M}_e \in \mathcal{R}_e \,|\, \mathbb{M}_f \text{ is nice}\right) \ge 1 - \epsilon/4, \tag{41}$$

since in view of (39) and (41), inequality (40) completes the proof of Claim 9.

Now we prove (41). Fix a nice k-graph $H \in \mathcal{R}_f$ and condition on the event $\mathbb{M}_f = H$. The event that $\mathbb{M}_e$ is not simple is contained in the union of the following four events:

$$\mathcal{E}_1 = \{ \text{ two of the randomly chosen edges } e_1, \ldots, e_s \text{ coincide } \},$$
$$\mathcal{E}_2 = \{ (e_i \setminus v_i) \cup u_i \text{ is a loop for some } i = 1, \ldots, s \},$$
$$\mathcal{E}_3 = \{ (e_i \setminus v_i) \cup u_i \in H \text{ for some } i = 1, \ldots, s \},$$
$$\mathcal{E}_4 = \{ (e_i \setminus v_i) \cup u_i = (e_j \setminus v_j) \cup u_j \text{ for some distinct } i \text{ and } j \}.$$

Event $\mathcal{E}_1$ covers all cases when a double edge is created by replacing several vertices in the same edge. Creation of multiple edges in other ways is addressed by events $\mathcal{E}_3$ and $\mathcal{E}_4$.

In what follows we will several times use the fact that

$$\deg_{H\setminus(G\cup f)}(v) \geq \tau d/2 \geq \epsilon d/2, \qquad \forall v \in [n], \tag{42}$$

which is immediate from (31) and $\tau \geq \epsilon$. To bound the probability of $\mathcal{E}_1$, observe that, given $1 \leq i < j \leq s$, the number of choices of a coinciding pair $e_i = e_j$ is $\deg_{H\setminus(G\cup f)}(v_i, v_j) \leq \deg_{H\setminus(G\cup f)}(v_i)$ and the probability that both v_i and v_j actually select a fixed common edge is $(\deg_{H\setminus(G\cup f)}(v_i)\deg_{H\setminus(G\cup f)}(v_j))^{-1}$. Therefore using (42) we obtain

$$\mathbb{P}(\mathcal{E}_1|\mathbb{M}_f = H) \leq \sum_{1\leq i<j\leq s} \frac{\deg_{H\setminus(G\cup f)}(v_i, v_j)}{\deg_{H\setminus(G\cup f)}(v_i)\deg_{H\setminus(G\cup f)}(v_j)}$$
$$\leq \sum_{1\leq i<j\leq s} \frac{1}{\deg_{H\setminus(G\cup f)}(v_j)} \leq \frac{2\binom{k}{2}}{\epsilon d} \leq \frac{\epsilon}{16}, \tag{43}$$

where the last inequality follows from (14) with $\alpha = 1/2$ and sufficiently large C'.

To bound the probability of $\mathcal{E}_2$, note that a loop in $\mathbb{M}_e$ can only be created when for some $i = 1, \ldots, s$, the randomly chosen edge e_i contains both v_i and u_i. There are at most $\deg_{H\setminus(G\cup f)}(u_i, v_i)$ such edges. Therefore, by (37) and (42) we get

$$\mathbb{P}(\mathcal{E}_2 \,|\, \mathbb{M}_f = H) \leq \sum_{i=1}^{s} \frac{\deg_{H\setminus(G\cup f)}(u_i, v_i)}{\deg_{H\setminus(G\cup f)}(v_i)} \leq \frac{2k(\ell_1 + k\log_2 n)}{\tau d}$$
$$\leq \frac{2k\ell_1}{\tau d} + \frac{2k^2\log_2 n}{\epsilon d} = \frac{2kC_1}{n} + \frac{2k^2\log_2 n}{\epsilon d} \leq \frac{\epsilon}{16}, \tag{44}$$

where the last inequality is implied by (14) with $\alpha = 1/2$, (16) and sufficiently large C'.

Similarly we bound the probability of $\mathcal{E}_3$, the event that for some i we will choose $e_i \in H \setminus (G \cup f)$ with $(e_i \setminus v_i) \cup u_i \in H$. There are $\mathrm{cod}_{H|G\cup f}(u_i, v_i)$ such edges. Thus, by (38) and (42) we obtain

$$\mathbb{P}(\mathcal{E}_3 \mid \mathbb{M}_f = H) \le \sum_{i=1}^{s} \frac{\mathrm{cod}_{H|G\cup f}(u_i, v_i)}{\deg_{H\setminus(G\cup f)}(v_i)} \le \frac{2k(\ell_2 + k\log_2 n)}{\tau d}$$
$$\le \frac{2k\ell_2}{\tau d} + \frac{2k^2\log_2 n}{\tau d} \le \frac{2kC_2 d}{n^{k-1}} + \frac{2k^2\log_2 n}{\epsilon d} \le \frac{\epsilon}{16}, \quad (45)$$

where the last inequality follows from (14) with $\alpha = 1/2$, (15) with $\alpha = 1$ and sufficiently large C'.

Finally, note that, given $1 \le i < j \le s$, if a pair $e_i, e_j \in H \setminus (G \cup f)$ satisfies the condition in $\mathcal{E}_4$, then $u_j \in e_i \setminus v_i$ and $e_j = (e_i \setminus \{v_i, u_j\}) \cup \{v_j, u_i\}$. This means that e_j is uniquely determined by e_i, u_i, u_j, v_i, and v_j. Therefore the number of such pairs e_i, e_j is at most $\deg_{H\setminus(G\cup f)}(v_i, u_j) \le \deg_{H\setminus(G\cup f)}(v_i)$ and we get exactly the same bound as in (43):

$$\mathbb{P}(\mathcal{E}_4 \mid \mathbb{M}_f = H) \le \sum_{1\le i<j\le s} \frac{\deg_{H\setminus(G\cup f)}(v_i, u_j)}{\deg_{H\setminus(G\cup f)}(v_i)\deg_{H\setminus(G\cup f)}(v_j)}$$
$$\le \sum_{1\le i<j\le s} \frac{1}{\deg_{H\setminus(G\cup f)}(v_j)} \le \frac{\epsilon}{16}. \quad (46)$$

Combining (43)–(46) and averaging over nice H, we obtain (41), as required. □

5. Concluding remarks

Theorem 1 remains valid if we replace the random hypergraph $\mathbb{G}^{(k)}(n,m)$ by $\mathbb{G}^{(k)}(n,p)$ with $p = (1-2\gamma)d/\binom{n-1}{k-1}$, say. To see this one can modify the proof of Theorem 1 as follows. Let $B_n \sim \mathrm{Bin}(\binom{n}{k}, p)$ be a random variable independent of the process $(\mathbb{G}(t))_t$. If $B_n \le m \le |S|$, sample $\mathbb{G}^{(k)}(n,p)$ by taking the first B_n edges of S (which are uniformly distributed over all k-graphs with B_n edges). Otherwise sample $\mathbb{G}^{(k)}(n,p)$ among k-graphs with B_n edges independently. In view of the assumption (3), Chernoff's inequality (see [9, (2.5)]) and (13) imply

$$\mathbb{P}\left(\mathbb{G}^{(k)}(n,p) \not\subset \mathbb{R}^{(k)}(n,d)\right) \le \mathbb{P}(B_n > m) + \mathbb{P}(|S| < m) \to 0, \quad \text{as} \quad n \to \infty.$$

The lower bound on d in Theorem 1 is necessary because the second moment method applied to $\mathbb{G}^{(k)}(n,p)$ (cf. Theorem 3.1(ii) in [2]) and asymptotic equivalence of $\mathbb{G}^{(k)}(n,p)$ and $\mathbb{G}^{(k)}(n,m)$ yields that for $d = o(\log n)$ and $m \sim cM$ there is a sequence $\Delta = \Delta(n) \gg d$ such that the maximum degree $\mathbb{G}^{(k)}(n,m)$ is at least Δ a.a.s.

In view of the above, our approach cannot be extended to $d = O(\log n)$ in part (i) of Theorem 5. Nevertheless, we believe (as it was already stated in [7]) that for loose Hamilton cycles it suffices to assume that $d = \Omega(1)$.

Conjecture 1. *For every $k \ge 3$ there is a constant d_k such that if $d \ge d_k$, then a.a.s. $\mathbb{R}^{(k)}(n,d)$ contains a loose Hamilton cycle.*

We also believe that the lower bounds on d in parts (ii) and (iii) of Theorem 5 are of optimal order.

Conjecture 2. *For all integers $k > \ell \geq 2$ if $d \ll n^{\ell-1}$, then a.a.s. $\mathbb{R}^{(k)}(n,d)$ is* not *ℓ-Hamiltonian.*

Acknowledgments

The authors would like to thank the anonymous referees for careful reading of the manuscript and numerous helpful suggestions.

References

[1] Peter Allen, Julia Böttcher, Yoshiharu Kohayakawa, Yury Person, Tight Hamilton cycles in random hypergraphs, Random Structures Algorithms 46 (3) (2015) 446–465.
[2] Béla Bollobás, Random Graphs, second edition, Cambridge Studies in Advanced Mathematics, vol. 73, Cambridge University Press, Cambridge, 2001.
[3] Colin Cooper, Alan Frieze, Bruce Reed, Random regular graphs of non-constant degree: connectivity and Hamiltonicity, Combin. Probab. Comput. 11 (3) (2002) 249–261.
[4] Andrzej Dudek, Alan Frieze, Loose Hamilton cycles in random uniform hypergraphs, Electron. J. Combin. 18 (1) (2011), Paper 48, 14 pp.
[5] Andrzej Dudek, Alan Frieze, Tight Hamilton cycles in random uniform hypergraphs, Random Structures Algorithms 42 (3) (2013) 374–385.
[6] Andrzej Dudek, Alan Frieze, Po-Shen Loh, Shelley Speiss, Optimal divisibility conditions for loose Hamilton cycles in random hypergraphs, Electron. J. Combin. 19 (4) (2012), Paper 44, 17 pp.
[7] Andrzej Dudek, Alan Frieze, Andrzej Ruciński, Matas Šileikis, Loose Hamilton cycles in regular hypergraphs, Combin. Probab. Comput. 24 (1) (2015) 179–194.
[8] Alan Frieze, Loose Hamilton cycles in random 3-uniform hypergraphs, Electron. J. Combin. 17 (1) (2010), Note 28, 4 pp.
[9] Svante Janson, Tomasz Łuczak, Andrzej Ruciński, Random Graphs, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2000.
[10] Michał Karoński, Alan Frieze, Introduction to Random Graphs, Cambridge University Press, 2015. Available at http://www.math.cmu.edu/~af1p/Book.html.
[11] Jeong Han Kim, Van H. Vu, Sandwiching random graphs: universality between random graph models, Adv. Math. 188 (2) (2004) 444–469.
[12] Jeong Han Kim, Van H. Vu, Generating random regular graphs, Combinatorica 26 (6) (2006) 683–708.
[13] Michael Krivelevich, Benny Sudakov, Van H. Vu, Nicholas C. Wormald, Random regular graphs of high degree, Random Structures Algorithms 18 (4) (2001) 346–363.
[14] Brendan D. McKay, Asymptotics for symmetric 0–1 matrices with prescribed row sums, Ars Combin. 19 (A) (1985) 15–25.
[15] Brendan D. McKay, Nicholas C. Wormald, Asymptotic enumeration by degree sequence of graphs with degrees $o(n^{1/2})$, Combinatorica 11 (4) (1991) 369–382.
[16] Robert W. Robinson, Nicholas C. Wormald, Almost all regular graphs are Hamiltonian, Random Structures Algorithms 5 (2) (1994) 363–374.
[17] A. Steger, N.C. Wormald, Generating random regular graphs quickly, in: Random Graphs and Combinatorial Structures, Oberwolfach, 1997, Combin. Probab. Comput. 8 (4) (1999) 377–396.
[18] Nick Wormald, The degree sequence of a random graph, Technical Report 22/2016, Oberwolfach Research Institute for Mathematics, 2016.