THE ROYAL SWEDISH ACADEMY OF SCIENCES
INSTITUT MITTAG-LEFFLER

Auravägen 17, SE-182 60 Djursholm, Sweden
Tel. +46 8 622 05 60 Fax. +46 8 622 05 89
info@mittag-leffler.se www.mittag-leffler.se

Embedding the Erdos-Renyi Hypergraph into the Random Regular Hypergraph and Hamiltonicity

A. Dudek, A. Frieze, A. Rucinski and M. Sileikis

REPORT No. 41, 2013/2014, spring
ISSN 1103-467X
ISRN IML-R- -41-13/14- -SE+spring

Embedding the Erdős-Rényi Hypergraph into the Random Regular Hypergraph and Hamiltonicity

ANDRZEJ DUDEK[1*¶] ALAN FRIEZE[2†] ANDRZEJ RUCIŃSKI[3‡¶]
MATAS ŠILEIKIS[4§¶]

[1]Department of Mathematics, Western Michigan University, Kalamazoo, MI
andrzej.dudek@wmich.edu

[2]Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA
alan@random.math.cmu.edu

[3]Department of Discrete Mathematics, Adam Mickiewicz University, Poznań, Poland
rucinski@amu.edu.pl

[4]Mathematical Institute, the University of Oxford, United Kingdom
matas.sileikis@gmail.com

May 2, 2015

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 $n \log n \ll m \ll n^k$ then one can couple $\mathbb{G}^{(k)}(n,m)$ and $\mathbb{R}^{(k)}(n,d)$ with $d \sim km/n$ so that the latter contains the former with probability tending to one as $n \to \infty$. This complements some previous results 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 $n^{k-2} \ll d \ll n^{k-1}$, then a.a.s. $\mathbb{R}^{(k)}(n,d)$ contains a tight Hamilton cycle.

*Supported in part by Simons Foundation grant #244712.

†Supported in part by NSF grant CCF 1013110.

‡Supported in part by the Polish NSC grant N201 604940 and NSF grant DMS 1102086.

§Supported in part by the Polish NSC grant N201 604940 and the Knut and Alice Wallenberg Foundation.

¶Part of research performed during a visit to the Institut Mittag-Leffler (Djursholm, Sweden).

1 Introduction

1.1 Background

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

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

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 a k-graph 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 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 (d, m), or zero (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 $n \to \infty$.

In 2004, Kim and Vu proved that if $d = o(n^{1/3})$ then there exists a coupling of the random graphs $\mathbb{G}^{(2)}(n, p)$ and $\mathbb{R}^{(2)}(n, d)$ with $p = \frac{d}{n}\left(1 - (\log n/d)^{1/3}\right)$ 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 $\mathbb{G}(n, p)$ 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 the reverse embedding which is not of our interest here).

In [8] the authors 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 a sufficiently small constant $c = c(k) > 0$. Although (2) is stated for the uniform k-graph $\mathbb{G}^{(k)}(n, m)$, it 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 Remark 2).

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 [8]. Since the new result is also valid for $k = 2$ (for the proof of this case see [7, Section 10.3]), via asymptotic equivalence it extends (1).

Theorem 1. *For each $k \geq 2$ there is a positive constant C such that for $\gamma = \gamma(n)$ such that*

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

and $m = (1-\gamma)nd/k$ is an integer 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 1. In the assumption of Theorem 1 the term $1/n$ is not necessary when $k \leq 7$. Indeed, the inequality of arithmetic and geometric means implies that $(d/n^{k-1} + (\log n)/d)^{1/3} \geq (2/n^{(k-1)/2})^{1/3} \geq \sqrt[3]{2}/n$.

A graph property is any family of graphs closed under isomorphism. A graph property $\mathcal{P} = \mathcal{P}(n)$ consisting of k-graphs on n vertices is called *monotone increasing* if it preserved under edge addition.

Corollary 2. *Let $\mathcal{P}$ be a monotone increasing property of k-graphs such that $\mathbb{G}^{(k)}(n,m)$ a.a.s. satisfies $\mathcal{P}$ for some $m = m(n)$, $n \log n \ll m \ll n^k$. Then there is a sequence $d = d(n) \sim \frac{km}{n}$ such that $\mathbb{R}^{(k)}(n,d)$ a.a.s. satisfies $\mathcal{P}$.*

1.3 Hamilton Cycles in Hypergraphs

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

Define an *ℓ-overlapping cycle* (or, for short, *ℓ-cycle*) 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. A 1-cycle is called a *loose* cycle and $(k-1)$-cycle – a *tight* cycle. 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 in the case $k = 2$. Asymptotically almost sure hamiltonicity of $\mathbb{G}^{(2)}(n,d)$ was proved by Robinson and Wormald [14] in 1994 for fixed $d \geq 3$, by Krivelevich, Sudakov, Vu and Wormald [12] in 2001 for $d \geq n^{1/2}\log n$, and by Cooper, Frieze and Reed [3] in 2002 for $C \leq d \leq n/C$ and some large constant C.

Much less is known for random regular hypergraphs. First results on the loose hamiltonicity of $\mathbb{G}^{(k)}(n,p)$ were obtained in a sequence of papers by Frieze [9] (for $k=3$), Dudek and Frieze [4] (for $k \geq 4$) under a divisibility condition $2(k-1)|n$, which was ultimately relaxed to the necessary $(k-1)|n$ by Dudek, Frieze, Loh and Speiss [6]. 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 [10]) allows us to reformulate the aforementioned results for the k-graph $\mathbb{G}^{(k)}(n,m)$.

Theorem 3 ([9, 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.*

Fro Theorem 3 and the older embedding result (2), in [8] the authors derived a condition under which $\mathbb{R}^{(k)}(n,d)$ contains a loose Hamilton cycle a.a.s.

Theorem 4 ([8])**.** *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 5 ([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 5 immediately imply the following result that was already anticipated by the authors in [8].

Theorem 6.

(i) There is a constant $C > 0$ such that if $C\log n \leq d \leq n^{k-1}/C$, then a.a.s. $\mathbb{G}^{(3)}(n,d)$ 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 a ℓ-Hamilton cycle.

We believe that the restrictions on d in parts (ii) and (iii) of Theorem 6 are basically optimal. We discuss them in more detail in Section 4.

2 Proof of Theorem 1

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

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

Of course, $\mathbb{G}^{(k)}(M)$ has the same distribution as $\mathbb{G}^{(k)}(n, M)$.

Our approach is to represent $\mathbb{R}^{(k)}(n, d)$ as the outcome of a k-graph process which behaves similarly to $\mathbb{G}^{(k)}(t), t = 0, \dots, M$ and thus permits a good coupling. For this, generate $\mathbb{R}^{(k)}(n, d)$ and choose a random ordering of its M edges to get a random ordered graph $(\eta_1, \dots, \eta_M)$. Revealing them one by one, we obtain a *regular k-graph* process

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

From now on we will often drop the superscript $^{(k)}$ and write 'graph' instead of 'k-graph', whenever k is clear from the context.

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

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

This is not true if we replace $\mathbb{G}$ by $\mathbb{R}$, except for the very first step $t = 0$. However, it turns out that for the most of the time 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 a several definitions.

Given an ordered graph G , let $\mathcal{G}_G(n, d)$ be the family of *extensions* of G, that is, ordered d-regular graphs the first edges of which are equal to G. We say that a graph G with $t \leq M$ edges is *admissible*, if $\mathcal{G}_G(n, d)$ is not empty. For such graphs we have $\mathbb{P}\left(\mathbb{R}(t) = G\right) > 0$ and define

$$p_{t+1}(e|G) := \mathbb{P}\left(\eta_{t+1} = e \,|\, \mathbb{R}(t) = G\right), \quad t = 0, \dots, M-1. \tag{4}$$

For $\epsilon \in (0, 1)$ and $t = 0, \dots, M-1$, consider an event

$$\mathcal{A}_t = \left\{p_{t+1}(e|\mathbb{R}(t)) \geq \frac{1-\epsilon}{\binom{n}{k} - t} \quad \text{for every} \quad e \in K_n \setminus \mathbb{R}(t)\right\}. \tag{5}$$

Lemma 7. *For every $k \geq 2$ there is a positive constant C' such that for $\epsilon = \epsilon(n)$ such that $(1-\epsilon)M$ is an integer and*

$$C'\left((d/n^{k-1} + (\log n)/d)^{1/3} + 1/n\right) \leq \epsilon < 1, \tag{6}$$

then the event $\mathcal{A} := \mathcal{A}_1 \cap \dots \cap \mathcal{A}_{(1-\epsilon)M}$ occurs a.a.s.

From Lemma 7, which is proved in Section 3, we deduce Theorem 1 using a coupling similar to the one which was used in [8].

Proof of Theorem 1. Let $C = 3C'$, where C' is the constant from Lemma 7 and let $\epsilon = \gamma/3$.

Let us outline the plan of the proof. We will define a graph process $\mathbb{R}'(t) := (\eta'_1, \dots, \eta'_t), t = 0, \dots, M$ such that for every admissible graph R with $t \le M-1$ edges we have

$$\mathbb{P}\left(\eta'_{t+1} = e \,|\, \mathbb{R}'(t) = R\right) = p_{t+1}(e|R). \tag{7}$$

Then we will show that $\mathbb{G}(n, m)$ can be sampled from the subgraph $\mathbb{R}'((1-\varepsilon)M)$ a.a.s. In view of (7), the distribution of $\mathbb{R}'(M)$ is the same as the one of $\mathbb{R}(M)$ and thus we can identify $\mathbb{R}(n, d)$ with the graph underlying $\mathbb{R}'(M)$.

We set $\mathbb{R}'(0)$ to be an empty vector and define $\mathbb{R}'(t)$ inductively (for $t = 1, 2, \dots$) as follows. Suppose that 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 $\mathcal{A}_t$ occurs, that is,

$$p_{t+1}(e|R_t) \ge \frac{1-\epsilon}{\binom{n}{k} - t} \quad \text{for every} \quad e \in K_n \setminus R_t \tag{8}$$

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

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

where the inequality holds because of the assumption (8). 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

$$\eta'_{t+1} = \begin{cases} \varepsilon_{t+1}, & \text{if } \xi_{t+1} = 1, \varepsilon_{t+1} \in K_n \setminus R_t, \\ f_{R_t,G_t}(\varepsilon_{t+1}), & \text{if } \xi_{t+1} = 1, \varepsilon_{t+1} \in R_t, \\ \zeta_{t+1}, & \text{if } \xi_{t+1} = 0. \end{cases} \tag{9}$$

If the event $\mathcal{A}_t$ fails, we nevertheless generate ξ_t, whereas η'_{t+1} is then sampled directly (without defining ζ_{t+1}) according to probabilities (4). Such definition of η_t makes sure that

$$\mathcal{A}_t \cap \{\xi_{t+1} = 1\} \quad \Longrightarrow \quad \varepsilon_{t+1} \in \mathbb{R}'(t+1). \tag{10}$$

Further, define a set of edges

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

The crucial thing is that by (10) we have

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

Since the vectors (ξ_i) and (ε_i) are independent, conditioning on $|S| \ge m$, the first m edges in S comprise a graph which is distributed as $\mathbb{G}(n,m)$. Therefore

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

Since by Lemma 7 we have $\mathcal{A}$ a.a.s., to completes the proof it suffices to show that $\mathbb{P}\left(|S| < m\right) \to 0$.

Note that $|S|$ is a binomial random variable, namely,

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

with

$$\mathbb{E}|S| \ge (1-2\epsilon)M \quad \text{and} \quad \mathrm{Var}|S| \le (1-\epsilon)^2 \epsilon M \le \epsilon M. \tag{11}$$

Recall that $\epsilon = \gamma/3$ and thus $m = (1-\gamma)M = (1-3\epsilon)M$. By (11), Chebyshev's inequality, and the inequality $\epsilon \ge C' \log n/d$ which follows from (6)), we get

$$\mathbb{P}\left(|S| < m\right) \le \mathbb{P}\left(|S| - \mathbb{E}|S| < -\epsilon M\right) \le \frac{\epsilon M}{(\epsilon M)^2} = \frac{k}{\epsilon n d} \le \frac{k}{C' n \log n} \to 0. \tag{12}$$

□

3 Proof of Lemma 7

In all proofs of this section we will assume the conditions of Lemma 7, that is, $(1-\varepsilon)M$ is an integer and (6) holds with a sufficiently large $C' = C'(k) \ge 1$. In particular, we have

$$\epsilon \ge C'(\log n/d)^\alpha, \tag{13}$$

$$\epsilon \ge C'(d/n^{k-1})^\alpha \tag{14}$$

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

$$\epsilon \ge C'/n. \tag{15}$$

Given a graph G with maximum degree at most d, let us define the *residual vertex degrees* as

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

To prove Lemma 7 we will start with a fact which allows one to control the residual degrees of the evolving graph $\mathbb{R}(t) = (\eta_1, \dots, \eta_t)$. For a vertex $v \in [n]$ and $t = 0, \dots, M$, define

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

Claim 8. *Let $\tau = 1 - t/M$. For every $k \geq 2$ there is a constant $a = a(k) > 0$ so that a.a.s.*

$$\forall t \leq (1-\epsilon)M, \quad \forall v \in [n], \quad |R_t(v) - \tau d| \leq \sqrt{a\tau d \log n}. \tag{16}$$

In particular, a.a.s.

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

Proof. The idea is that the concentration of the degrees depends completely on the random ordering of the edges and not on the structure of the graph $\mathbb{R}(M)$. Observe that if we fix any d-regular graph H and condition $\mathbb{R}(M)$ to be a random permutation of the edges of H, then $R_t(v)$ is a hypergeometric random variable. Using Remark 2.11 in [10] together with inequalities (2.14) and (2.16) therein, we get

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

$$\mathrm{Var}R_t(v) = \frac{td(M-t)(M-d)}{M^2(M-1)} \leq \frac{d(M-t)}{M} = \tau d,$$

and

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

Let $a = 4(k+2)$ and $x = \sqrt{a\tau d \log n}$. Condition (13) 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,

$$\mathbb{P}\left(|R_t(v) - \tau d| \geq \sqrt{a\tau d \log n}\right) \leq 2\exp\left\{-\frac{a}{4}\log n\right\} = 2n^{-k-2}.$$

Since we have fewer than $nM \leq n^{k+1}$ choices of t and v, (16) follows by taking a union bound.

Finally, (17) follows from (16) and (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 $\mathcal{G}_G(n,d)$ is the family of extensions of G to a d-regular ordered graph and $\mathbb{G}_G = \mathbb{G}_G(n,d)$ is a graph chosen uniformly at random from $\mathcal{G}_G(n,d)$.

Further, for a graph $H \in \mathcal{G}_G(n,d)$ define the *common degree* (relative to subgraph $G \subseteq H$) 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 in general symmetric in u and v. Also, define the degree of a pair of vertices u,v as

$$\deg_H(uv) = \left|\left\{W \in \tbinom{[n]}{k-2} : W \cup uv \in H\right\}\right|.$$

The next fact is used in the proof of Claim 10 only.

Claim 9. *Let G be a graph with $t+1 \le (1-\epsilon)M$ edges such that $\mathcal{G}_G(n,d)$ is nonempty. Let $\tau = 1 - t/M$. Suppose that*

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

Then, for each $e \in K_n \setminus G$ we have

$$\mathbb{P}\left(e \in \mathbb{G}_G\right) \le \frac{(2k)^k k! \tau d}{n^{k-1}}. \tag{21}$$

Moreover, if $\ell \ge \ell_0 := 2^{k+2}(k-1)!k^{k-1}\tau d^2/n^{k-1}$, then for every $u,v \in [n], u \ne v$, we have

$$\mathbb{P}\left(\operatorname{cod}_{\mathbb{G}_G|G}(u,v) > \ell\right) \le 2^{-(\ell-\ell_0)}. \tag{22}$$

Also, if $p \ge p_0 := 16k!\tau d/n$, then for every $u,v \in [n], u \ne v$, we have

$$\mathbb{P}\left(\deg_{\mathbb{G}_G \setminus G}(uv) > p\right) \le 2^{-(p-p_0)}. \tag{23}$$

Proof. To prove (21) define the families

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

Let us define an auxiliary bipartite graph B between $\mathcal{G}_{e\in}$ and $\mathcal{G}_{e\notin}$ in which $H \in \mathcal{G}_{e\in}$ is connected to $H' \in \mathcal{G}_{e\notin}$ whenever H' can be obtained from H by the following operation (known as switching in the literature dating back to McKay [13]). Let $e = e_1 = v_{1,1} \dots v_{1,k}$ and pick $k-1$ more edges $e_i = v_{i,1} \dots v_{i,k} \in H \setminus G$, $i = 2, \dots, k$ (with vertices 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 Figure 1). Let $f(H)$ be the number of ways to apply a *forward switching*, i.e., number of graphs $H' \in \mathcal{G}_{e\notin}$ which can be obtained from H, and $b(H')$ be the number of ways

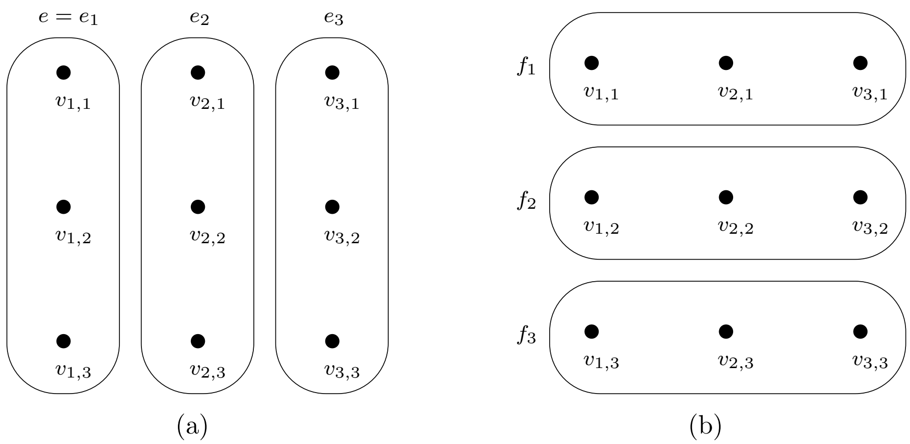

Figure 1: Switching (for $k = 3$): before (a) and after (b).

to apply a *backward switching*, i.e., the number of graphs $H \in \mathcal{G}_{e\in}$ from which H' can be obtained. Thus,

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

Note that $H \setminus G$ and $H' \setminus G$ both have $\tau M - 1$ edges and, by (20), 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 $H \setminus G$, there are at least

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

ways to choose a $(k-1)$-tuple $e_2, \ldots, e_k$ such that $e_1, \ldots, e_k$ are pairwise disjoint. The number of such $(k-1)$-tuples that may lead to a double edge after the switching (by repeating some edge of H which intersects e_1), is at most $kd \cdot (2\tau d)^{k-1}$. Thus,

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

Since we can clearly assume that $2k^4/n \le 1/4$ and (14) with $\alpha = 1$ and $C' \ge 4(2k)^{2k}$ implies

$$\frac{(2k)^{2k} d}{n^{k-1}} \le \frac{\epsilon (2k)^{2k}}{C'} \le 1/4,$$

we conclude that

$$f(H) \ge \frac{(\tau M)^{k-1}}{2(k-1)!}. \tag{25}$$

Moreover, we have $b(H') \le (2\tau d)^k$. Therefore (24) implies that

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

which concludes the proof of (21).

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

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

We compare sizes of $\mathcal{G}(\ell)$ and $\mathcal{G}(\ell-1)$ in a similar way as above. For this we define the following switching which maps a graph $H \in \mathcal{G}(\ell)$ to a graph $H' \in \mathcal{G}(\ell-1)$. Select two 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 $k-1$ more edges $e_2, \ldots, e_k \in H \setminus G$ so that $e_1, \ldots, e_k$ are mutually 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 $f_j = v_{1,j} \ldots v_{k,j}$, $j = 1, \ldots, k$ (see Figure 2).

By a similar counting as in the previous case, the number $f(H)$ of graphs $H' \in \mathcal{G}(\ell-1)$ which can be obtained from $H \in \mathcal{G}(\ell)$ by this switching is bounded by

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

due to (14) with $\alpha = 1$ and $C' \ge 4(k+1)!(2k)^k$.

Conversely, H can be obtained 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

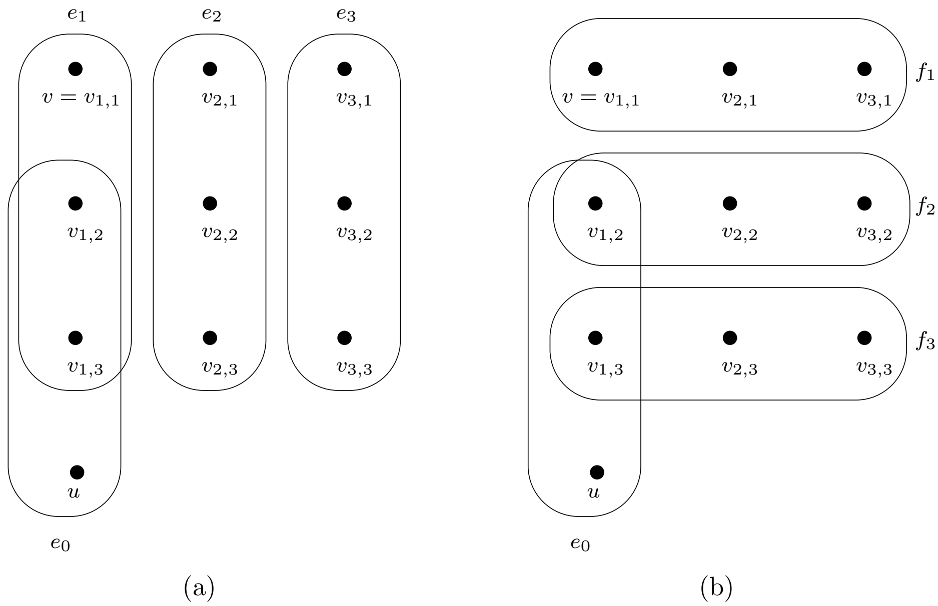

Figure 2: Switching (for $k = 3$): before (a) and after (b).

one vertex from $(e_0 \setminus u) \cup v$. Therefore the number of ways to apply a backward switching is $b(H') \le d \cdot (2\tau d)^k$. Clearly,

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

by the assumption $\ell \ge \ell_0 = 2^{k+2}(k-1)!k^{k-1}\tau d^2/n^{k-1}$. Then

$$\mathbb{P}\left(\mathrm{cod}_{\mathbb{G}_G|G}(u,v) > \ell\right) \le \sum_{i>\ell} \frac{|\mathcal{G}(i)|}{|\mathcal{G}_G(n,d)|} \le \sum_{i>\ell} \frac{|\mathcal{G}(i)|}{|\mathcal{G}(\ell_0)|}$$
$$= \sum_{i>\ell} \prod_{j=\ell_0+1}^{i} \frac{|\mathcal{G}(j)|}{|\mathcal{G}(j-1)|} \le \sum_{i>\ell} 2^{-(i-\ell_0)} = 2^{-(\ell-\ell_0)}, \quad (26)$$

which completes the proof of (22).

It remains to show (23). Fix $u, v \in [n]$ and define the families

$$\mathcal{G}(p) = \left\{H \in \mathcal{G}_G(n,d) : \deg_{H\setminus G}(uv) = p\right\}, \qquad p = 0, 1, \ldots.$$

In order to compare sizes of $\mathcal{G}(p)$ and $\mathcal{G}(p-1)$ we define the following switching which maps a graph $H \in \mathcal{G}(p)$ to a graph $H' \in \mathcal{G}(p-1)$. Select $e_1 \in H \setminus G$ contributing to $\deg_{H\setminus G}(uv)$ and pick $k-1$ edges $e_2, \ldots, e_k \in H \setminus G$ so that $e_1, \ldots, e_k$ are mutually disjoint. Writing $e_i = v_{i,1} \dots v_{i,k}, i = 1, \dots, k$ with $u = v_{1,1}$ and $v = v_{1,2}$, replace $e_1, \dots, e_k$ by $f_j = v_{1,j} \dots v_{k,j}$, $j = 1, \dots, k$ (similarly as in Figure 1).

Let $f(H)$ be the number of graphs $H' \in \mathcal{G}(p-1)$ which can be obtained from $H \in \mathcal{G}(p)$ and let $b(H')$ be the number of graphs $H \in \mathcal{G}(p)$ from which H' can be obtained. As in (25) we get that

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

For the converse we choose two disjoint edges in $H' \setminus G$ containing u and v, respectively, and then $k-2$ mutually disjoint edges in $H' \setminus G$ not containing u and v. Thus, $b(H') \leq (2\tau d)^2 \cdot (\tau M)^{k-2}$ and

$$\frac{|\mathcal{G}(p)|}{|\mathcal{G}(p-1)|} \leq \frac{\max_{H' \in \mathcal{G}(p-1)} b(H')}{\min_{H \in \mathcal{G}(p)} f(H)} \leq \frac{(2\tau d)^2(\tau M)^{k-2} \cdot 2(k-1)!}{p(\tau M)^{k-1}} \leq \frac{8k!\tau d}{pn} \leq \frac{1}{2},$$

since by assumption $p \geq p_0 = 16k!\tau d/n$. Now (23) follows from similar computations to (26). This finishes the proof of Claim 9. □

For the last claim, which will be directly used in the proof of Lemma 7, we need to provide a few more definitions regarding random d-regular multigraphs.

Let G be an ordered graph with t edges. Let $\mathbb{M}_G(n,d)$ be a random *multigraph extension* of an ordered graph G to an ordered d-regular multigraph. Namely, $\mathbb{M}_G(n,d)$ is a sequence of M edges, 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 of such permutations is

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

Although $\mathbb{M}_G(n,d)$ is not uniformly distributed over all multigraph extensions of G (multigraphs with loops are given by fewer permutations than simple graphs), but it is uniform over $\mathcal{G}_G(n,d)$. That is, $\mathbb{M}_G(n,d)$, conditioned on simplicity, has the same distribution as $\mathbb{G}_G(n,d)$. Further, for every two edges $e \in K_n \setminus G$, let us write

$$\mathbb{M}_e = \mathbb{M}_{G \cup e}(n,d) \quad \text{and} \quad \mathcal{G}_e = \mathcal{G}_{G \cup e}(n,d). \tag{27}$$

The next claim shows that the probabilities of simplicity $\mathbb{P}(\mathbb{M}_e \in \mathcal{G}_e)$ are asymptotically the same for all $e \in K_n \setminus G$.

Claim 10. *Let G be a graph with $t \leq (1-\epsilon)M - 1$ edges, such that $\mathcal{G}_G(n,d)$ is nonempty. Let $\tau = 1 - t/M$ and suppose that*

$$\tau d/2 + 1 \leq r_G(v) \leq 2\tau d \qquad \forall v \in [n]. \tag{28}$$

Then, for every $e', e'' \in K_n \setminus G$ we have

$$\frac{\mathbb{P}(\mathbb{M}_{e''} \in \mathcal{G}_{e''})}{\mathbb{P}(\mathbb{M}_{e'} \in \mathcal{G}_{e'})} \geq 1 - \epsilon/2.$$

Proof. Set

$$\mathbb{M}' = \mathbb{M}_{e'} \quad \mathbb{M}'' = \mathbb{M}_{e''}, \quad \mathcal{G}' = \mathcal{G}_{e'}, \quad \text{and} \quad \mathcal{G}'' = \mathcal{G}_{e''},$$

for convenience. We start by constructing a coupling of $\mathbb{M}'$ and $\mathbb{M}''$ in which they differ in at most $k+1$ positions (counting in the replacement of e' by e'' in the fixed part).

Let $e' = v'_1 \dots v'_k$ and $e'' = v''_1 \dots v''_k$. Let $r = k - |e' \cap e''|$ and suppose without loss of generality that $v'_1 \dots v'_r$ are disjoint from $v''_1 \dots v''_r$. Let Π' be the permutation underlying the multigraph $\mathbb{M}'$. Let Π^* be obtained from Π' by replacing, for each $i = 1, \dots, r$, a random uniform copy of v''_i by v'_i. Define $\mathbb{M}^*$ by chopping Π^* into consecutive k-tuples and appending them to $G \cup e''$.

It is easy to see that Π^* is uniform over permutations of the multiset $\{1, \dots, 1, \dots, n, \dots, n\}$ with multiplicities $r_{G \cup e''}(v), v \in [n]$, and therefore $\mathbb{M}^*$ has the same distribution as $\mathbb{M}''$ and therefore we further identify $\mathbb{M}^*$ and $\mathbb{M}''$.

Observe that if we condition $\mathbb{M}'$ on being a simple graph H, then $\mathbb{M}''$ can be equivalently obtained by choosing, uniformly at random, for each $i = 1, \dots, r$, an edge incident to v''_i in $H \setminus (G \cup e')$, say, e_i and replacing it by $(e_i \setminus v''_i) \cup v'_i$. The crucial idea is that such a switching of edges is unlikely to create loops or multiple edges.

It is, however, possible, that for certain H this is not true. For example, if $e'' \in H \setminus (G \cup e')$, then the random choice of two edges described above is unlikely to destroy this e'', but e' in the non-random part will be replaced by e'' thus creating a double edge. Moreover, if almost every $(k-1)$-tuple of vertices extending v''_i to an edge in $H \setminus (G \cup e')$ also extends v'_i to an edge in H, then most likely the replacement of v''_i by v'_i will create a double edge. To avoid such instances, we want to assume that

$$e'' \in K_n \setminus H \tag{29}$$

$$\max_{i=1,\dots,r} \operatorname{cod}_{H|G \cup e'}(v'_i, v''_i) \leq \ell_0 + k \log_2 n, \tag{30}$$

$$\max_{i=1,\dots,r} \deg_{H \setminus (G \cup e')}(v'_i v''_i) \leq p_0 + k \log_2 n, \tag{31}$$

where $\ell_0 = 2^{k+2}(k-1)!k^{k-1}\tau d^2/n^{k-1}$ and $p_0 = 16k!\tau d/n$ is as in Claim 9. Define the following subfamily of simple extensions of $G \cup e'$:

$$\mathcal{G}'_{\text{nice}} = \{H \in \mathcal{G}' : H \text{ satisfies (29)-(31)}\}.$$

Note that $\mathbb{M}'$, conditioned on $\mathbb{M}' \in \mathcal{G}'$, is distributed as $\mathbb{G}_{G \cup e'}(n, d)$. By Claim 9 applied to $G' = G \cup e'$ we have

$$\mathbb{P}(\mathbb{M}' \notin \mathcal{G}'_{\text{nice}} \,|\, \mathbb{M}' \in \mathcal{G}') \leq \frac{(2k)^k k! \tau d}{n^{k-1}} + 2 \cdot r 2^{-k \log_2 n} \leq \frac{((2k)^k k! + k)d}{n^{k-1}} \leq \frac{\epsilon}{4}, \tag{32}$$

by (14) with $\alpha = 1$ and $C \geq 4((2k)^k k! + k)$. By standard probability, we have

$$\frac{\mathbb{P}(\mathbb{M}'' \in \mathcal{G}'')}{\mathbb{P}(\mathbb{M}' \in \mathcal{G}')} \geq \mathbb{P}(\mathbb{M}'' \in \mathcal{G}'' \mid \mathbb{M}' \in \mathcal{G}') \\ \geq \mathbb{P}(\mathbb{M}'' \in \mathcal{G}'' \mid \mathbb{M}' \in \mathcal{G}'_{\text{nice}})\, \mathbb{P}(\mathbb{M}' \in \mathcal{G}'_{\text{nice}} \mid \mathbb{M}' \in \mathcal{G}'). \quad (33)$$

It suffices to show that

$$\mathbb{P}(\mathbb{M}'' \in \mathcal{G}'' \mid \mathbb{M}' \in \mathcal{G}'_{\text{nice}}) \geq 1 - \epsilon/4, \quad (34)$$

since in view of (32) and (34), inequality (33) completes the proof of the statement.

Now we prove (34). Fix $H \in \mathcal{G}'_{\text{nice}}$ and condition on $\mathbb{M}' = H$. The event that $\mathbb{M}''$ is not simple is contained in the union of the following four events:

$$\mathcal{E}_1 = \{ \text{ two of the randomly chosen edges } e_1, \ldots, e_r \text{ coincide } \},$$

$$\mathcal{E}_2 = \{ \text{ some edge } (e_i \setminus v_i'') \cup v_i' \text{ is a loop } \},$$

$$\mathcal{E}_3 = \{ \ (e_i \setminus v_i'') \cup v_i' \in H \text{ for some } i = 1, \ldots, r \ \},$$

$$\mathcal{E}_4 = \{ \ (e_i \setminus v_i'') \cup v_i' = (e_j \setminus v_j'') \cup v_j' \text{ for some } i < j \ \}.$$

The purpose of the event $\mathcal{E}_1$ is to avoid complicated analysis of the cases when a double edge is created by replacement of several vertices in the same edge. Events $\mathcal{E}_3$ and $\mathcal{E}_4$ concern the creation of multiple edges.

In what follows we will several time use the fact that

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

which is immediate from (28) and $\tau \geq \epsilon$. To bound the probability of $\mathcal{E}_1$, observe that, given $1 \leq i, j \leq r$, the number of choices of a coinciding pair $e_i'' = e_j''$ is crudely bounded by $\deg_{H \setminus G}(v_i'')$. Therefore,

$$\mathbb{P}(\mathcal{E}_1 | \mathbb{M}' = H) \leq \sum_{1 \leq i < j \leq r} \frac{1}{\deg_{H \setminus (G \cup e')}(v_j'')} \leq \frac{2\binom{k}{2}}{\epsilon d} \leq \frac{\epsilon}{16}, \quad (35)$$

where the last inequality follows from (13) with $\alpha = 1/2$ and $C' \geq 4\sqrt{2\binom{k}{2}}$.

To bound the probability of $\mathcal{E}_2$, we note that a loop can only be created in $\mathbb{M}''$ when for some $i = 1, \ldots, r$ the randomly chosen edge e_i'' contains both v_i'' and v_i'. There are at most $\deg_{H \setminus (G \cup e')}(v_i'' v_i')$ such edges. Therefore, from (31), (28), and $\tau \geq \epsilon$ we get

$$\mathbb{P}(\mathcal{E}_2 \mid \mathbb{M}' = H) \leq \sum_{i=1}^{r} \frac{\deg_{H \setminus (G \cup e')}(v_i' v_i'')}{\deg_{H \setminus (G \cup e')}(v_i'')} \leq \frac{2k(p_0 + k \log_2 n)}{\tau d} \\ \leq \frac{2kp_0}{\tau d} + \frac{2k^2 \log_2 n}{\varepsilon d} = \frac{32kk!}{n} + \frac{2k^2 \log_2 n}{\varepsilon d} \leq \frac{\epsilon}{16}, \quad (36)$$

where the last inequality is implied by (15) with $C' \geq 32\cdot 32kk!$ and (13) with $\alpha = 1/2$ and $C' \geq 8k/\sqrt{\log 2}$.

Similarly we bound the probability of $\mathcal{E}_3$. We choose $e_i \in H \setminus G$ containing v'' such that $(e_i \setminus v_i'') \cup v_i' \in H$. There are at most $\mathrm{cod}_{H|G\cup e'}(v_i', v_i'')$ such edges. Thus, by (30), (28), and $\tau \geq \epsilon$, we obtain

$$\begin{aligned}\mathbb{P}\left(\mathcal{E}_3 \,|\, \mathbb{M}' = H\right) &\leq \sum_{i=1}^{r} \frac{\mathrm{cod}_{H|G\cup e'}(v_i', v_i'')}{\deg_{H\setminus(G\cup e')}(v_i'')} \leq \frac{2k(\ell_0 + k\log_2 n)}{\tau d} \\ &\leq \frac{2k\ell_0}{\tau d} + \frac{2k^2 \log_2 n}{\epsilon d} \leq \frac{2^{k+3}k^k(k-1)!d}{n^{k-1}} + \frac{2k\log_2 n}{\varepsilon d} \leq \frac{\epsilon}{16}, \quad (37)\end{aligned}$$

where the last inequality follows from (14) applied with $\alpha = 1$ and $C' \geq 2^{k+8}k^k(k-1)!$ and (13) applied with $\alpha = 1/2$ and $C' \geq 8\sqrt{k/\log 2}$.

Finally, note that, given $1 \leq i < j \leq r$, the number of pairs $e_i, e_j \in H \setminus (G \cup e')$ for which event $\mathcal{E}_4$ holds is at most $\deg_{H\setminus(G\cup e')}(v_i'')$. Hence we get exactly the same bound as in (35):

$$\mathbb{P}\left(\mathcal{E}_4 \,|\, \mathbb{M}' = H\right) \leq \sum_{1\leq i<j\leq r} \frac{1}{\deg_{H\setminus(G\cup e')}(v_j'')} \leq \frac{\epsilon}{16}. \quad (38)$$

Combining (35)-(38), we obtain (34), as required. □

Proof of Lemma 7. In view of Claim 8 it suffices to show that

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

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

$$\tau d/2 + 1 \leq d(\tau - \delta) \leq r_G(v) \leq d(\tau + \delta) \leq 2\tau d, \qquad v \in [n],$$

where

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

In particular, G satisfies (28), therefore we can apply Claim 10.

For every $e', e'' \in K_n \setminus G$ we have (recall the definition (27))

$$\frac{\mathbb{P}\left(\eta_{t+1} = e'' \,|\, \mathbb{R}(t) = G\right)}{\mathbb{P}\left(\eta_{t+1} = e' \,|\, \mathbb{R}(t) = G\right)} = \frac{|\mathcal{G}_{G\cup e''}(n,d)|}{|\mathcal{G}_{G\cup e'}(n,d)|} = \frac{|\mathcal{G}''|}{|\mathcal{G}'|}. \quad (39)$$

Since each simple extension is given by $(k!)^{M-t}$ permutations, we have

$$\mathbb{P}\left(\mathbb{M}' \in \mathcal{G}'\right) = \frac{|\mathcal{G}'|(k!)^{M-t}}{N_G} = \frac{|\mathcal{G}'|(k!)^{M-t}\prod_{v\in[n]}(d - \deg_{G\cup e'}(v))!}{(k(M-t))!},$$

and similarly for the family $\mathcal{G}''$. This yields, after a few cancellations, that

$$\frac{|\mathcal{G}''|}{|\mathcal{G}'|} = \frac{\prod_{v\in e''\setminus e'} r_G(v)}{\prod_{v\in e'\setminus e''} r_G(v)} \cdot \frac{\mathbb{P}\left(\mathbb{M}'' \in \mathcal{G}''\right)}{\mathbb{P}\left(\mathbb{M}' \in \mathcal{G}'\right)}. \tag{40}$$

The ratio of products in (40) is at least

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

where the last inequality holds by (13) with $\alpha = 1/3$ and $C' \geq \sqrt[3]{16ak^2}$. On the other hand, the ratio of probabilities in (40) is at least $1-\epsilon/2$ by Claim 10, so we have obtained that (39) is at least $1-\epsilon$.

Finally, noting that

$$\max_{e'\in K_n\setminus G} \mathbb{P}\left(\eta_{t+1} = e' \,|\, \mathbb{R}(t) = G\right)$$

is at least as large as the average over all $e' \in K_n \setminus G$, which is $\frac{1}{\binom{n}{k}-t}$, we get that for every $e \in K_n \setminus G$

$$\mathbb{P}\left(\eta_{t+1} = e \,|\, \mathbb{R}(t) = G\right) \geq (1-\epsilon) \max_{e'\in K_n\setminus G} \mathbb{P}\left(\eta_{t+1} = e' \,|\, \mathbb{R}(t) = G\right) \geq \frac{1-\epsilon}{\binom{n}{k}-t},$$

which concludes the proof. □

4 Concluding Remarks

Remark 2. Theorem 1 remains valid if we replace 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 $B_n \sim \mathrm{Bin}(\binom{n}{k}, p)$ a random variable independent of the process $(\mathbb{G}^{(k)}(t))_t$. If $B_n \leq m \leq |S|$, sample $\mathbb{G}^{(k)}(n,p)$ by taking the first B_n edges of S (which are uniformly distributed over all B_n-edge k-graphs). Otherwise sample $\mathbb{G}^{(k)}(n,p)$ among B_n-edge graphs independently. In view of assumption (3), Chernoff's inequality (see [10, (2.5)]) and (12) imply

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

Remark 3. 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 $\Delta = \Delta(n)$ such that $d = o(\Delta)$ and the maximum degree $\mathbb{G}^{(k)}(n,m)$ is at least Δ a.a.s.

Remark 4. In view of Remark 3, our approach cannot be extended to $d = O(\log n)$ in part (i) of Theorem 6. Nevertheless, we believe (as it was already stated in [8]) that for loose Hamilton cycles suffices only to assume that $d = \Omega(1)$.

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

Remark 5. We also believe that the lower bounds on d in parts (ii) and (iii) of Theorem 6 are optimal.

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.*

References

[1] P. Allen, J. Böttcher, Y. Kohayakawa, and Y. Person. Tight Hamilton cycles in random hypergraphs. *Random Structures Algorithms* 46(3) 446-465, 2015.

[2] B. Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European J. Combin.*, 1(4):311–316, 1980.

[3] C. Cooper, A. Frieze, and B. Reed. Random regular graphs of non-constant degree: connectivity and Hamiltonicity. *Combin. Probab. Comput.*, 11(3):249–261, 2002.

[4] A. Dudek and A. Frieze. Loose Hamilton cycles in random uniform hypergraphs. *Electron. J. Combin.*, 18(1):Paper 48, pp. 14, 2011.

[5] A. Dudek and A. Frieze. Tight Hamilton cycles in random uniform hypergraphs. *Random Structures Algorithms*, 42(3):374–385, 2013.

[6] A. Dudek, A. Frieze, P.-S. Loh, and S. Speiss. Optimal divisibility conditions for loose Hamilton cycles in random hypergraphs. *Electron. J. Combin.*, 19(4):Paper 44, pp. 17, 2012.

[7] A. Frieze, and M. Karoński. *Introduction to Random Graphs.* Cambridge University Press, 2015 (to appear).

[8] A. Dudek, A. Frieze, A. Ruciński, and M. Šileikis. Loose hamilton cycles in regular hypergraphs. *Combin. Probab. Comput.* 24:179–194, 2015.

[9] A. Frieze. Loose Hamilton cycles in random 3-uniform hypergraphs. *Electron. J. Combin.*, 17(1):Note 28, pp. 4, 2010.

[10] S. Janson, T. Łuczak, and A. Ruciński. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.

[11] J. H. Kim and V. H. Vu. Sandwiching random graphs: universality between random graph models. *Adv. Math.*, 188(2):444–469, 2004.

[12] M. Krivelevich, B. Sudakov, V. H. Vu, and N. C. Wormald. Random regular graphs of high degree. *Random Structures Algorithms*, 18(4):346–363, 2001.

[13] B. D. McKay. Asymptotics for symmetric 0-1 matrices with prescribed row sums. *Ars Combin.*, 19(A):15–25, 1985.

[14] R. W. Robinson and N. C. Wormald. Almost all regular graphs are Hamiltonian. *Random Structures Algorithms*, 5(2):363–374, 1994.

[15] A. Steger and N. C. Wormald. Generating random regular graphs quickly. *Random graphs and combinatorial structures (Oberwolfach, 1997). Combin. Probab. Comput.* 8(4):377-396, 1999.