RANDOM HYPERGRAPH IRREGULARITY

PAUL BALISTER*, BÉLA BOLLOBÁS†, JENŐ LEHEL‡, AND MICHAŁ MORAYNE §

Abstract. A hypergraph is k-irregular if there is no set of k vertices all of which have the same degree. We asymptotically determine the probability that a random uniform hypergraph is k-irregular.

Key words. random hypergraph, uniform hypergraph, vertex degrees, degree irregularity

AMS subject classifications. 05C65, 05C80, 05C07

1. Introduction. For $r \geq 2$, an *r-uniform hypergraph* is a pair $\mathcal{H} = (V, E)$ consisting of a finite set V, called the set of vertices, and a set E of distinct r-element subsets of V, called the set of edges. The *degree* of a vertex $v \in V$ is $d_H(v) = |\{e \in E : v \in e\}|$. The hypergraph H is called (degree) *irregular* if its vertex degrees are all distinct, i.e., $d_H(u) = d_H(v)$ implies $u = v$.

Repeated degrees in graphs (the case $r = 2$) was observed first by Behzad et al. in [1], and the repetitions were discussed by Bollobás in [2] (see also [3]). The study of degree irregular hypergraphs started with Faudree et al. in [5]. Gyárfás et al. show in [7] that there exist degree irregular r-uniform hypergraphs of order n if and only if $r \geq 3$ and $n \geq r + 3$. Furthermore, it is proved that almost every random r-uniform hypergraph is degree irregular for $r \geq 6$, when the edges emerge independently with probability $p = 1/2$.

The case when the probability p depends on n was discussed recently for $r = 3$ and 4. Balister et al. proved in [4] that in a random 3-uniform or 4-uniform hypergraph of order n the probability that some two vertices have the same degree tends to one as $n \to \infty$. The purpose of the present paper is to give a complete answer to the question for every $r \geq 2$ and $p = p_n$. As the corollary of a more general result we obtain the asymptotic value of the probability that a random r-uniform hypergraph has no repeated degrees.

THEOREM 1.1. *Let $\mathcal{H}$ be a random r-uniform hypergraph of order n with independent probability $p = p_n$ for each r-set being an edge. Then*

$$\mathbb{P}(\mathcal{H} \text{ is irregular}) \to \begin{cases} 0 & \text{if } p(1-p)n^{r-5} \to 0, \\ f_r(c) & \text{if } p(1-p)n^{r-5} \to c \in (0,\infty), \\ 1 & \text{if } p(1-p)n^{r-5} \to \infty, \end{cases}$$

*Department of Mathematical Sciences, University of Memphis, Tennessee, USA (pbalistr@memphis.edu). Research supported in part by NSF grant DMS 1301614

† Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, Centre for Mathematical Sciences, Wilberforce Road Cambridge CB3 0WB, UK (B.Bollobas@dpmms.cam.ac.uk), and Department of Mathematical Sciences, University of Memphis, Tennessee, USA (bollobas@msci.memphis.edu). Research supported in part by MULTIPLEX grant 317532

‡ Department of Mathematical Sciences, University of Memphis, Tennessee, USA (jlehel@memphis.edu), and Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences Budapest, P.O. Box 127 Budapest, Hungary, H-1364

§ Institute of Mathematics and Computer Science, Wrocław University of Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland (michal.morayne@pwr.wroc.pl)

where

$$f_r(c) = \exp\left(- \frac{1}{4\sqrt{\pi}} \sqrt{\frac{(r-1)!}{c}} \right) \in (0,1).$$

The formula implies that in a random r-uniform hypergraph asymptotically almost surely there exist repeating degrees for $r \leq 4$. In contrast, asymptotically almost surely there is no degree repetition for $r \geq 6$ and constant p, $0 < p < 1$. In this regard the value $r = 5$ behaves like a "threshold" with respect to the rank r for the property that an r-uniform hypergraph is asymptotically almost surely irregular or not.

For fixed $k \geq 2$, we call a hypergraph $\mathcal{H}$ *k-irregular* if there is no set of k vertices all of which having the same degree in $\mathcal{H}$. For the case $k = 2$, a 2-irregular hypergraph is just irregular, all vertices having distinct degrees. Our main result is the following theorem.

Theorem 1.2. *Let $\mathcal{H}$ be a random r-uniform hypergraph of order n with independent probability $p = p_n$ for each r-set being an edge. Then*

$$\mathbb{P}(\mathcal{H} \text{ is } k\text{-irregular}) = \exp\left(-\binom{n}{k} \frac{1}{\sqrt{k}(\sqrt{2\pi}\sigma)^{k-1}} \right) + o(1)$$

as $n \to \infty$, where $\sigma^2 = \binom{n}{r-1} p(1-p)$. Moreover, if $\binom{n}{k} \frac{1}{\sqrt{k}(\sqrt{2\pi}\sigma)^{k-1}} \to \lambda$ for some constant $\lambda \in [0,\infty)$ as $n \to \infty$, then the number of k-sets of vertices of equal degree converges in distribution to a Poisson random variable of mean λ.

Corollary 1.3. *For any fixed k and r,*

$$\mathbb{P}(\mathcal{H} \text{ is } k\text{-irregular}) \to \begin{cases} 0 & \text{if } p(1-p)n^{r-3-2/(k-1)} \to 0, \\ f_{k,r}(c) & \text{if } p(1-p)n^{r-3-2/(k-1)} \to c \in (0,\infty), \\ 1 & \text{if } p(1-p)n^{r-3-2/(k-1)} \to \infty, \end{cases}$$

where

$$f_{k,r}(c) = \exp\left(- \frac{1}{k!\sqrt{k}} \left(\frac{(r-1)!}{2\pi c} \right)^{(k-1)/2} \right) \in (0,1).$$

Theorem 1.1 above is the special case $k = 2$ of Corollary 1.3. It is worth noting that if we allow $c = 0, \infty$, then with the usual extension of arithmetic, Corollary 1.3 takes the following simpler form: if $p(1-p)n^{r-3-2/(k-1)} \to c \in [0,\infty]$, then $\mathbb{P}(\mathcal{H}$ is k-irregular$) \to f_{k,r}(c)$.

The proof of Theorem 1.2 uses the method of moments. It is prepared for by a sequence of technical lemmas in Section 2. The choice of the conditions and the assumptions in the lemmas might look arbitrary until Lemma 2.6 is obtained and applied to conclude the proof of Theorem 1.2 in Section 3.

It is tempting to conjecture that when $\binom{n}{k} \frac{1}{\sqrt{k}(\sqrt{2\pi}\sigma)^{k-1}} \to \infty$, the number W of k-sets of vertices of equal degree is asymptotically normally distributed. This should indeed hold under a fairly wide range of the parameters, but one should note that it is not universally true, even when the mean and variance of W are large. For example, if $p\binom{n}{r} \sim 1$ then the number X of edges in $\mathcal{H}$ is given by an approximately Poisson distribution and these edges are usually disjoint. But, with disjoint edges, $W = \binom{n-rX}{k} + \binom{rX}{k} \approx \binom{n}{k} - rX\binom{n}{k-1}$ has a probability distribution which is far from normal, even though it has large mean and variance.

Recall a few standard technical notations that we will use in the discussions. For positive functions $f(n)$ and $g(n)$, we write $f(n) = O(g(n))$ if there is a constant c such that $f(n) \le c \cdot g(n)$; we write $f(n) = \Omega(g(n))$ if there is a constant $c > 0$ such that $f(n) \ge c \cdot g(n)$; we write $f(n) = \Theta(g(n))$ provided both $f(n) = O(g(n))$ and $f(n) = \Omega(g(n))$; we write $f(n) = o(g(n))$ if $f(n)/g(n) \to 0$ as $n \to \infty$; we write $f(n) = \omega(g(n))$ if $f(n)/g(n) \to \infty$ as $n \to \infty$.

2. Bounds and estimations. Let $\mathcal{H} = \mathcal{H}^{(r)}(n,p)$ be a random r-uniform hypergraph on n vertices, each edge present independently with probability $p = p_n$. Denote the vertex set and the edge set of $\mathcal{H}$ by $V(\mathcal{H})$ and $E(\mathcal{H})$, respectively, and write $q = q_n = 1 - p_n$.

We wish to estimate the probability that $\mathcal{H}$ is k-irregular. As this probability is the same as for the complement of $\mathcal{H}$, we shall usually assume that $p \le \frac{1}{2}$. We shall use the method of moments, so we wish to estimate the probability that there are t fixed k-sets in $V(\mathcal{H})$, such that the degrees are constant within each k-set. We shall usually be dealing with the case when the t fixed k-sets are pairwise disjoint.

Fix $t \ge 0$, $k' \ge k \ge 2$, and let $V_1, \dots, V_t, V_{t+1}$ be pairwise disjoint subsets of vertices with $|V_i| = k$, for $1 \le i \le t$, and $|V_{t+1}| = k'$. (For technical reasons we shall need to consider the case $k' > k$.) Let $\mathcal{E}_i$ be the event that all vertices in V_i have the same degree in $\mathcal{H}$ and let $\mathcal{E}_{\le t} = \bigcap_{i=1}^t \mathcal{E}_i$. Set $V_{\le t} = \bigcup_{i=1}^t V_i$.

For $S \subseteq V_{\le t}$, let T_S be the number of edges $e \in E(\mathcal{H})$ with $e \cap V_{\le t} = S$. We may assume that $|S| \le r$ (otherwise $T_S = 0$). Clearly $\mathcal{E}_{\le t}$ is determined by the set of integers T_S, $S \ne \emptyset$, indeed, the degree of any vertex $v \in V_{\le t}$ is just $\sum_{S \ni v} T_S$. Let $V_{t+1} = \{v_1, \dots, v_{k'}\}$ and denote the degree of v_i by d_i. Let $d_i^{(0)}$ be the number of edges of $\mathcal{H}$ that meet both v_i and $V_{\le t}$.

Lemma 2.1. *Let X be the random variable $d_2^{(0)} - d_1^{(0)}$. Then $\mathbb{E}(X^2) \le 2ptk\binom{n}{r-2}$. Furthermore, if k and t are fixed and if $\mathbb{P}(\mathcal{E}_{\le t}) = n^{-O(1)}$ and $pn^{r-1} = \omega(\log n)$, then $\mathbb{E}(X^2 \mid \mathcal{E}_{\le t}) \le (2 + o(1))ptk\binom{n}{r-2}$.*

Proof. For the first result, note that there are $M = \binom{n-2}{r-1} - \binom{n-tk-2}{r-1}$ potential edges that contribute to $d_2^{(0)}$ but not $d_1^{(0)}$, and a disjoint set of M possible edges that contribute to $d_1^{(0)}$ but not $d_2^{(0)}$. Thus $X \sim \mathrm{Bin}(M,p) - \mathrm{Bin}(M,p)$ is the difference of two independent Binomial random variables. From this it is immediate that $\mathbb{E}(X) = 0$ and $\mathbb{E}(X^2) = \mathrm{Var}(X) = 2Mpq \le 2Mp$. Now

$$M = \binom{n-2}{r-1} - \binom{n-tk-2}{r-1} = \sum_{j=n-tk-2}^{n-3} \binom{j}{r-2} \le tk\binom{n}{r-2},$$

so the first result follows.

For the second result we first show that

$$\mathbb{E}(X^2 \mid \{T_S\}_{S \ne \emptyset}) \le \frac{2}{n-tk} \sum_{S \ne \emptyset} (r - |S|) T_S \le (2 + o(1)) \frac{r-1}{n} \sum_{S \ne \emptyset} T_S. \tag{2.1}$$

Write $X = \sum_{S \ne \emptyset} X_S$ where X_S is the contribution to X coming from edges e such that $e \cap V_{\le t} = S$ and $|S| \le r-1$. Fix S and let $N = \binom{n-tk}{r-|S|}$ be the number of potential edges e such that $e \cap V_{\leq t} = S$. Let $N_1 = \binom{n-tk-2}{r-|S|-1}$ be the number of potential edges e such that $e \cap V_{\leq t} = S$ and $v_1 \in e$ but $v_2 \notin e$ (or vice versa). Then conditioning on T_S means [1)] that $X_S = X_S^+ - X_S^-$, where X_S^+ counts the number of edges chosen from one set of size N_1 and X_S^- counts the number of edges chosen from a disjoint set of size N_1, inside a set of N elements, when T_S distinct elements are chosen uniformly at random from this N element set. Clearly $\mathbb{E}(X_S^+) = \mathbb{E}(X_S^-) = N_1 \cdot \frac{T_S}{N}$. Also

$$\mathbb{E}(X_S^+(X_S^+ - 1)) = \mathbb{E}(X_S^-(X_S^- - 1)) = N_1(N_1 - 1) \cdot \frac{T_S(T_S - 1)}{N(N-1)}$$

$$\mathbb{E}(X_S^+ X_S^-) = N_1^2 \cdot \frac{T_S(T_S-1)}{N(N-1)}.$$

From this one deduces that

$$\begin{aligned}\mathbb{E}(X_S^2 \mid T_S) &= \mathbb{E}((X_S^+ - X_S^-)^2) = 2\mathbb{E}(X_S^+(X_S^+ - 1)) + 2\mathbb{E}(X_S^+) - 2\mathbb{E}(X_S^+ X_S^-)\\ &= 2\frac{T_S(T_S-1)N_1(N_1-1)}{N(N-1)} + 2\frac{T_S N_1}{N} - 2\frac{T_S(T_S-1)N_1^2}{N(N-1)}\\ &\leq 2\frac{T_S N_1}{N}.\end{aligned}$$

Now $N_1/N = \frac{(r-|S|)(n-tk-r+|S|)}{(n-tk)(n-tk-1)} \leq (r-|S|)/(n-tk)$. Thus (2.1) follows by summing over all $S \neq \emptyset$.

The event $\mathcal{E}_{\leq t}$ is determined entirely by $\{T_S\}_{S \neq \emptyset}$, so to estimate $\mathbb{E}(X^2 \mid \mathcal{E}_{\leq t})$ it is enough to combine (2.1) with a bound on $\mathbb{E}(T \mid \mathcal{E}_{\leq t})$, where $T = \sum_{S \neq \emptyset} T_S$. The unconditioned distribution of T is $\mathrm{Bin}(N', p)$, where $N' = \binom{n}{r} - \binom{n-tk}{r}$ is the number of potential edges meeting $V_{\leq t}$.

Fix $\varepsilon > 0$. Then by the Chernoff bound, there exists $\delta = \delta(\varepsilon) > 0$ such that $\mathbb{P}(T > (1+\varepsilon)pN') \leq e^{-\delta N' p}$. Using the assumption that $\mathbb{P}(\mathcal{E}_{\leq t}) = n^{-O(1)}$ and $N'p = \Omega(pn^{r-1}) = \omega(\log n)$ we have that

$$\mathbb{P}(T > (1+\varepsilon)N'p \mid \mathcal{E}_{\leq t}) \leq e^{-\delta N'p + O(\log n)} = e^{-\omega(\log n)}.$$

As the maximum possible value of T is $n^{O(1)} = e^{O(\log n)} N'p$ we have

$$\mathbb{E}(T \mid \mathcal{E}_{\leq t}) \leq (1+\varepsilon)N'p + \mathbb{P}(T > (1+\varepsilon)N'p \mid \mathcal{E}_{\leq t})(\max T) \leq (1+2\varepsilon)N'p,$$

for all sufficiently large n. Finally, $N' = \binom{n}{r} - \binom{n-tk}{r} \leq tk\binom{n}{r-1}$, so using (2.1) and the chain rule for conditional expectation we obtain

$$\mathbb{E}(X^2 \mid \mathcal{E}_{\leq t}) \leq (2+o(1))\tfrac{r-1}{n}\mathbb{E}(T \mid \mathcal{E}_{\leq t}) \leq (2+o(1))\tfrac{r-1}{n} N'p \leq (2+o(1))ptk\tbinom{n}{r-2}.$$

□

Let $d_i^{(1)}$ be the number of edges of $\mathcal{H}$ that meet v_i and at least one other $v_j \in V_{t+1}$, but not $V_{\leq t}$. Let Y be the random variable $d_2^{(1)} - d_1^{(1)}$. As Y does not involve edges meeting $V_{\leq t}$, Y is independent of both $X = d_2^{(0)} - d_1^{(0)}$ and $\mathcal{E}_{\leq t}$.

Lemma 2.2. *If $Y = d_2^{(1)} - d_1^{(1)}$ then $\mathbb{E}(Y^2) \leq 2pk'\binom{n}{r-2}$.*

[1)] note that the conditional distribution here is multivariate hypergeometric

Proof. $Y \sim \mathrm{Bin}(N,p) - \mathrm{Bin}(N,p)$ where $N = \binom{n-tk-2}{r-1} - \binom{n-tk-k'}{r-1}$ counts the number of potential edges meeting, say v_1 and some other v_j, but not v_2 or $V_{\le t}$. As $\mathbb{E}(Y) = 0$, $\mathbb{E}(Y^2) = \mathrm{Var}(Y) = 2Npq$. The result follows as $N \le k'\binom{n}{r-2}$. □

Let $d'_i = d_i^{(0)} + d_i^{(1)}$ be the number of edges meeting v_i and at least one other vertex of $V_{\le t} \cup V_{t+1}$.

Corollary 2.3. *Set $Z_{ij} = d'_i - d'_j$. If $\mathbb{P}(\mathcal{E}_{\le t}) = n^{-O(1)}$ and $pn^{r-1} = \omega(\log n)$ as $n \to \infty$, then both $\mathbb{E}(Z_{ij}^2)$ and $\mathbb{E}(Z_{ij}^2 \mid \mathcal{E}_{\le t})$ are bounded by $(2+o(1))p(kt+k')\binom{n}{r-2} = O(pn^{r-2})$ for $i \neq j$.*

Proof. Note that $Z_{ij} \sim Z_{21} = X + Y$. As Y is independent of X and $\mathcal{E}_{\le t}$, and $\mathbb{E}(Y) = 0$,

$$\mathbb{E}(Z_{21}^2 \mid \mathcal{E}_{\le t}) = \mathbb{E}(X^2 \mid \mathcal{E}_{\le t}) + 2\mathbb{E}(X \mid \mathcal{E}_{\le t})\mathbb{E}(Y) + \mathbb{E}(Y^2) = \mathbb{E}(X^2 \mid \mathcal{E}_{\le t}) + \mathbb{E}(Y^2)$$

and similarly for the unconditioned expectation. The result follows from Lemmas 2.1 and 2.2. □

Before we estimate the probability that $\mathcal{E}_{t+1}$ holds, we shall need a simple estimate comparing the probabilities in a binomial distribution with that of a normal distribution. The lemma below follows from a result in Feller [6] (Theorem 1. p.170) by elementary calculations. Here we present an alternative direct proof.

Lemma 2.4. *There exists an absolute constant $C > 0$ such that for all $N, t \in \mathbb{Z}$, $N \ge 0$, $p \in (0,1)$, and $z + Np \in (t - \frac{1}{2}, t + \frac{1}{2}]$, we have*

$$\mathbb{P}(\mathrm{Bin}(N,p) = t) = \frac{1}{\sqrt{2\pi}\sigma}\big(e^{-z^2/2\sigma^2} + \varepsilon_{N,p}(z)\big)$$

where $\sigma^2 = Np(1-p)$, $|\varepsilon_{N,p}(z)| \le C/\sigma$, and $\int_{-\infty}^{\infty} |\varepsilon_{N,p}(z)|\, dz \le C$.□

Proof. Write $q = 1 - p$. First we note that the result is trivial for bounded σ, so we shall assume σ is large. We may also assume that $|z| < \sigma^{4/3}$ as otherwise both $\mathbb{P}(\mathrm{Bin}(N,p) = t)$ and $e^{-z^2/2\sigma^2}$ are bounded by any negative power of σ. (For the binomial this follows from the Chernoff bound.) Similarly, we can assume the integral of $|\varepsilon_{N,p}(z)|$ is restricted to the range $|z| < \sigma^{4/3}$ in the last statement. As $|z| < \sigma^{4/3} \ll \sigma^2 \le \min\{Np, Nq\}$, we also have that $t, N-t \ge \sigma^2/2$. Write $t = z_0 + Np$. From Stirling's formula and some calculation we have

$$\begin{aligned}
\mathbb{P}(\,\mathrm{Bin}(N,p) = t) &= \binom{N}{t} p^t q^{N-t} \\
&= \frac{1}{\sqrt{2\pi}\sigma}\big(\tfrac{Np}{t}\big)^{t+\frac{1}{2}}\big(\tfrac{Nq}{N-t}\big)^{N-t+\frac{1}{2}}\big(1 + O(\tfrac{1}{t} + \tfrac{1}{N-t})\big) \\
&= \frac{1}{\sqrt{2\pi}\sigma} e^{-\left(Np+z_0+\frac{1}{2}\right)\log\left(1+\frac{z_0}{Np}\right) - \left(Nq-z_0+\frac{1}{2}\right)\log\left(1-\frac{z_0}{Nq}\right) + O\left(\frac{1}{\sigma^2}\right)} \\
&= \frac{1}{\sqrt{2\pi}\sigma} e^{-z_0^2/2\sigma^2 + O(1/\sigma^2) + O(|z_0|/\sigma^2) + O(|z_0|^3/\sigma^4)} \\
&= \frac{1}{\sqrt{2\pi}\sigma} e^{-z^2/2\sigma^2 + O(1/\sigma^2) + O(|z|/\sigma^2) + O(|z|^3/\sigma^4)}\,.
\end{aligned}$$

As the $O(1/\sigma^2) + O(|z|/\sigma^2) + O(|z|^3/\sigma^4)$ terms are bounded and the maximum of $z^n e^{-z^2/2\sigma^2}$ occurs at $z = O(\sigma)$, we deduce that

$$\varepsilon_{N,p}(z) = O\big(e^{-z^2/2\sigma^2}(1/\sigma^2 + |z|/\sigma^2 + |z|^3/\sigma^4)\big) = O(1/\sigma).$$

Also, the integral of $z^n e^{-z^2/2\sigma^2}$ is $O(\sigma^{n+1})$, so

$$\int_{-\sigma^{4/3}}^{\sigma^{4/3}} |\varepsilon_{N,p}(z)|\, dz = O(1).$$

□

LEMMA 2.5. *Fixing k', k, t, and conditioning on the event $\mathcal{E}_{\le t}$ and the values of d'_i, $i = 1, \dots, k'$, the probability that all vertices in V_{t+1} have the same degree is*

$$\mathbb{P}(\mathcal{E}_{t+1} \mid \mathcal{E}_{\le t},\, \{d'_i\}_1^{k'}) = \frac{1}{\sqrt{k'}\big(\sqrt{2\pi}\sigma\big)^{k'-1}} \bigg(1 + O\big(\tfrac{1}{\sigma}\big) - O\Big(\tfrac{1}{\sigma^2}\sum_{i<j} Z_{ij}^2\Big)\bigg)$$

where $\sigma^2 := \binom{n-tk-k'}{r-1}pq$.

Proof. Note that the event $\mathcal{E}_{\le t}$ and the values of d'_i are determined by the set of edges that either meet $V_{\le t}$, or meet V_{t+1} in at least two vertices. Let $N = \binom{n-tk-k'}{r-1}$. Then there are N possible edges meeting $V_{\le t} \cup V_{t+1}$ only at v_i. We condition on $\mathcal{E}_{\le t}$ and on the values of d'_i. Observe that $d_i - d'_i \sim \mathrm{Bin}(N,p)$, and $d_i - d'_i$ is independent of $d_j - d'_j$ for all $j \ne i$. The probability that all vertices in V_{t+1} have the same degree is then given by

$$\mathbb{P}(\mathcal{E}_{t+1} \mid \mathcal{E}_{\le t}, \{d'_i\}_{i=1}^{k'}) = \sum_d \prod_{i=1}^{k'} \mathbb{P}(d_i = d \mid d'_i) = \sum_d \prod_{i=1}^{k'} \mathbb{P}(\mathrm{Bin}(N,p) = d - d'_i).$$

By Lemma 2.4, for every fixed d and for every $z \in (d - d'_i - Np - \frac{1}{2}, d - d'_i - Np + \frac{1}{2}]$ we have $\mathbb{P}(\mathrm{Bin}(N,p) = d - d'_i) = \frac{1}{\sqrt{2\pi}\sigma}\big(e^{-z^2/2\sigma^2} + \varepsilon_{N,p}(z)\big)$ or equivalently, for every $x \in (d - Np - \frac{1}{2}, d - Np + \frac{1}{2}]$,

$$\mathbb{P}(\mathrm{Bin}(N,p) = d - d'_i) = \frac{1}{\sqrt{2\pi}\sigma}\big(e^{-(x-d'_i)^2/2\sigma^2} + \varepsilon_{N,p}(x - d'_i)\big).$$

Note that, defining d to be the nearest integer to $x + Np$, the right hand side is a step function of x, and hence we obtain

$$\begin{aligned}\mathbb{P}(\mathcal{E}_{t+1} \mid \mathcal{E}_{\le t}, \{d'_i\}_{i=1}^{k'}) &= \int_{-\infty}^{\infty} \prod_i \frac{1}{\sqrt{2\pi}\sigma}\big(e^{-(x-d'_i)^2/2\sigma^2} + \varepsilon_{N,p}(x - d'_i)\big)\, dx \\ &= \frac{1}{(\sqrt{2\pi}\sigma)^{k'}}\bigg\{\int_{-\infty}^{\infty} e^{-\sum(x-d'_i)^2/2\sigma^2}\, dx + O(1)\bigg\},\end{aligned}$$

where we have used the fact that $e^{-(x-d'_i)^2/2\sigma^2}$ and $\varepsilon_{N,p}(z - d'_i)$ are bounded, so the integral of any product of these terms is bounded by a multiple of $\int |\varepsilon_{N,p}(z)|\, dz = O(1)$ when there is at least one $\varepsilon_{N,p}(z - d'_i)$ term. Thus

$$\begin{aligned}\mathbb{P}(\mathcal{E}_{t+1} \mid \mathcal{E}_{\le t}, \{d'_i\}_{i=1}^{k'}) &= \frac{1}{(\sqrt{2\pi}\sigma)^{k'}}\bigg\{e^{-(\sum d_i'^2 - k'\overline{d}^2)/2\sigma^2}\int_{-\infty}^{\infty} e^{-k'(x-\overline{d})^2/2\sigma^2}\, dx + O(1)\bigg\} \\ &= \frac{1}{\sqrt{k'}(\sqrt{2\pi}\sigma)^{k'-1}}\Big(e^{-(\sum d_i'^2 - k'\overline{d}^2)/2\sigma^2} + O(1/\sigma)\Big)\end{aligned}$$

where $\overline{d} = \frac{1}{k'}\sum_{i=1}^{k'} d'_i$. Now

$$\sum d_i'^2 - k'\overline{d}^2 = \frac{1}{k'}\Big(\sum (k'-1)d_i'^2 - 2\sum_{i<j} d'_i d'_j\Big) = \frac{1}{k'}\sum_{i<j}(d'_i - d'_j)^2 = \frac{1}{k'}\sum_{i<j} Z_{ij}^2,$$

and k' is fixed, so

$$\mathbb{P}(\mathcal{E}_{t+1} \mid \mathcal{E}_{\le t}, \{d'_i\}_{i=1}^{k'}) = \frac{1}{\sqrt{k'}(\sqrt{2\pi}\sigma)^{k'-1}}\bigg(1 + O\big(\tfrac{1}{\sigma}\big) - O\Big(\tfrac{1}{\sigma^2}\sum_{i<j} Z_{ij}^2\Big)\bigg).$$

□

Lemma 2.6. *Assume that $pn^{r-1} = \omega(\log n)$ as $n \to \infty$, and let $\sigma^2 = \binom{n}{r-1}p(1-p)$. Then $\mathbb{P}(\mathcal{E}_{t+1} \mid \mathcal{E}_{\le t}) = \frac{1+o(1)}{\sqrt{k'}(\sqrt{2\pi}\sigma)^{k'-1}}$, in particular, $\mathbb{P}(\mathcal{E}_1) = \frac{1+o(1)}{\sqrt{k}(\sqrt{2\pi}\sigma)^{k-1}}$ and $\mathbb{P}(\mathcal{E}_{\le t}) = (1+o(1))\mathbb{P}(\mathcal{E}_1)^t$.*

Proof. We apply Lemma 2.5 with $t = 0$ and $k' = k$ to obtain

$$\mathbb{P}(\mathcal{E}_1 \mid \{d'_i\}_{i=1}^{k}) = \frac{1}{\sqrt{k}(\sqrt{2\pi}\sigma')^{k-1}}\bigg(1 + O\big(\tfrac{1}{\sigma'}\big) - O\Big(\tfrac{1}{\sigma'^2}\sum_{i<j} Z_{ij}^2\Big)\bigg),$$

where $\sigma'^2 = \binom{n-k}{r-1}pq = (1+o(1))\sigma^2$ and $\sigma^2 = \Theta(pn^{r-1}) \to \infty$, by assumption (and since $p \le 1/2$). Now $Z_{ij} \sim Z_{21} = Y$ as $X = 0$, so by Lemma 2.2, we obtain $\mathbb{E}(Z_{ij}^2) = O(pn^{r-2})$. As $n \to \infty$, the result follows for one k-set of vertices by using the total probability formula. Note that $\mathbb{P}(\mathcal{E}_1) = n^{-O(1)}$.

Now by induction on t, we may assume that $\mathbb{P}(\mathcal{E}_{\le t}) = n^{-O(1)}$. By Corollary 2.3, we have $\mathbb{E}(Z_{ij}^2 \mid \mathcal{E}_{\le t}) = O(pn^{r-2})$. Then, by Lemma 2.5, we obtain

$$\begin{aligned}\mathbb{P}(\mathcal{E}_{t+1} \mid \mathcal{E}_{\le t}) &= \mathbb{E}(\mathbb{P}(\mathcal{E}_{t+1} \mid \mathcal{E}_{\le t},\ \{d_i\}_1^k) \mid \mathcal{E}_{\le t}) \\ &= \frac{1}{\sqrt{k'}(\sqrt{2\pi}\sigma'')^{k'-1}}\Big(1 + O\big(\tfrac{1}{\sigma''}\big) - O\big(\tfrac{1}{\sigma''^2}pn^{r-2}\big)\Big),\end{aligned}$$

where $\sigma''^2 = \binom{n-tk-k'}{r-1}p(1-p) = (1+o(1))\sigma^2$ and $\sigma^2 = \Theta(pn^{r-1}) \to \infty$. Thus the first result follows using the total probability formula, and the last result follows by taking $k' = k$ and using induction on t. □

Let $\mathcal{E}'_i \subseteq \mathcal{E}_i$ be the event that vertices of V_i have the same degree, but that no other vertex of $\mathcal{H}$ has this degree. Define $\mathcal{E}'_{\le t} = \bigcap_{i=1}^t \mathcal{E}'_i$.

Lemma 2.7. *Under the conditions of Lemma 2.6, $\mathbb{P}(\mathcal{E}'_{\le t}) = (1 - O(n/\sigma))\mathbb{P}(\mathcal{E}_{\le t})$.*

Proof. For $\mathcal{E}_{\le t} \setminus \mathcal{E}'_{\le t}$ to hold, either two of the same degrees of two k-sets are equal, or one of the k-sets has degree equal to some other vertex of $\mathcal{H}$. In the first case we have $t-2$ k-sets and $k' = 2k$ vertices in the unified set (obtained by merging two V_i), which we label by V_t; in the second case we have $t-1$ k-sets and $k' = k+1$ vertices in V_t. Let $\mathcal{E}_t(q)$ be the event that in V_t the q vertices have the same degree.

Then, using Lemma 2.6 several times, and as t is fixed, we obtain

$$\begin{aligned}\mathbb{P}(\mathcal{E}_{\le t} \setminus \mathcal{E}'_{\le t}) &\le \binom{t}{2}\mathbb{P}(\mathcal{E}_{\le t-2} \cap \mathcal{E}_t(2k)) + tn\mathbb{P}(\mathcal{E}_{\le t-1} \cap \mathcal{E}_t(k+1)) \\ &= \binom{t}{2}\mathbb{P}(\mathcal{E}_t(2k) \mid \mathcal{E}_{\le t-2})\mathbb{P}(\mathcal{E}_{\le t-2}) + tn\mathbb{P}(\mathcal{E}_t(k+1) \mid \mathcal{E}_{\le t-1})\mathbb{P}(\mathcal{E}_{\le t-1}) \\ &= O(t^2/\sigma + tn/\sigma)\mathbb{P}(\mathcal{E}_{\le t}) = O(n/\sigma)\mathbb{P}(\mathcal{E}_{\le t}).\end{aligned}$$

The result then follows since $\mathcal{E}'_{\le t} \subseteq \mathcal{E}_{\le t}$. □

3. Proof of Theorem 1.2. We must show that

$$\mathbb{P}(\mathcal{H} \text{ is } k\text{-irregular}) = \exp\left(-\binom{n}{k}\frac{1}{\sqrt{k}(\sqrt{2\pi}\sigma)^{k-1}}\right) + o(1) \tag{3.1}$$

as $n \to \infty$. Recall that by replacing $\mathcal{H}$ with its complement, we may assume that $p = p_n \le \frac{1}{2}$. The degree of any one vertex is given by a $\mathrm{Bin}(\binom{n-1}{r-1}, p)$ random variable.

If $pn^{r-1} = O(\log n)$, then the expected degree is $O(\log n)$. Assume there are no k vertices of identical degree. Then there are at least $n/2$ vertices with degree at least $n/(2(k-1))$. Since the expected sum of degrees is $n(np^{r-1}) = n \cdot O(\log n)$, by Markov's inequality, we have

$$n \cdot O(\log n) \ge \frac{n}{2} \cdot \frac{n}{2(k-1)} \cdot \mathbb{P}(\mathcal{H} \text{ is } k\text{-irregular}).$$

From here we obtain $\mathbb{P}(\mathcal{H} \text{ is } k\text{-irregular}) = O((\log n)/n)$. Hence in this case the result holds as the right hand side of (3.1) equals $o(1)$. (Recall that $\sigma^2 = \binom{n}{r-1}p(1-p) = O(\log n)$.) Hence if (4) fails, $pn^{r-1}/\log n$ must be unbounded. Thus, taking a subsequence if necessary, we may assume $pn^{r-1} = \omega(\log n)$.

Let $\alpha = \frac{1}{\sqrt{k}(\sqrt{2\pi}\sigma)^{k-1}}$. First we deal with the case when $c_n := \binom{n}{k}\alpha$ is bounded as $n \to \infty$. Note that in this case $\sigma^{k-1} = \Omega(n^k)$, so that $n/\sigma \to 0$ as $n \to \infty$. Let W' be the number of k-sets of vertices that have the same degree as each other, but degrees different from any other vertex of $\mathcal{H}$. Thus W' counts the number of events $\mathcal{E}'_1$ that occur as we let V_1 range over all k-subsets of vertices of $\mathcal{H}$. Moreover $W'(W'-1)\dots(W'-t+1)$ counts the ordered t-tuples of distinct such sets. Thus

$$\mathbb{E}\big(W'(W'-1)\dots(W'-t+1)\big) = \sum_{V_1,\dots,V_t \text{ distinct}} \mathbb{P}(\mathcal{E}'_1 \cap \dots \cap \mathcal{E}'_t).$$

However, for $i < j$, if $V_i \cap V_j \ne \emptyset$, then $\mathcal{E}_i \cap \mathcal{E}_j$ implies that all vertices in $V_i \cup V_j$ have the same degree, so $\mathcal{E}'_i$ and $\mathcal{E}'_j$ fail. Thus $\mathcal{E}'_1 \cap \dots \cap \mathcal{E}'_t \ne \emptyset$ only when the V_i are disjoint. Now using that $n/\sigma \to 0$, by Lemmas 2.7 and 2.6, we obtain

$$\begin{aligned}\mathbb{E}\big(W'(W'-1)\dots(W'-t+1)\big) &= \sum_{V_1,\dots,V_t \text{ disjoint}} \mathbb{P}(\mathcal{E}'_{\le t}) \\ &= \binom{n}{k}\binom{n-k}{k}\dots\binom{n-tk+k}{k}(1-O(n/\sigma))\mathbb{P}(\mathcal{E}_{\le t}) \\ &= (1+o(1))\left(\binom{n}{k}\mathbb{P}(\mathcal{E}_1)\right)^t = (1+o(1))c_n^t.\end{aligned}$$

From Theorem 1.21 of [3], this implies that $d_{TV}(W', \mathrm{Po}(c_n)) \to 0$ as $n \to \infty$, where $d_{TV}(W', \mathrm{Po}(c_n))$ is the total variation distance of W' from a Poisson distribution of mean c_n. Now let W be the number of k-sets of vertices with common degree. Then $W \geq W'$, and using Lemmas 2.7 and 2.6 again, we have

$$\mathbb{E}(W - W') = \sum_{V_1}(\mathbb{P}(\mathcal{E}_1) - \mathbb{P}(\mathcal{E}_1')) = O(n/\sigma)\sum_{V_1}\mathbb{P}(\mathcal{E}_1) = O(n/\sigma)c_n.$$

Thus as $n \to \infty$, $\mathbb{E}(W - W') \to 0$ and so $\mathbb{P}(W = W') \to 1$. Consequently, if $c_n \to c \in [0, \infty)$ then W tends in distribution to a $\mathrm{Po}(c)$ random variable, and in general $\mathbb{P}(W = 0) = e^{-c_n} + o(1)$.

Now we deal with the case when $\binom{n}{k}\alpha \to \infty$ as $n \to \infty$. In this case we need to show that $\mathbb{P}(\mathcal{H}$ is k-irregular$) = o(1)$. We use the second moment method. Let W be the number of k-sets of equal degree vertices. Then, by Lemma 2.6, $\mathbb{E}(W) = (1 + o(1))\binom{n}{k}\alpha$ and with $t = 2$ we obtain

$$\mathbb{E}(W(W-1)) = \sum_{V_1 \neq V_2}\mathbb{P}(\mathcal{E}_1 \cap \mathcal{E}_2) = (1 + o(1))\binom{n}{k}^2\mathbb{P}(\mathcal{E}_1)^2 + \sum_{V_1 \cap V_2 \neq \emptyset}\mathbb{P}(\mathcal{E}_1 \cap \mathcal{E}_2)\ .$$

To see that the second term is negligible, let $\mathcal{E}_1(q)$ be the event that in a given q-set the vertices have the same degree. From Lemma 2.6 it follows that $\mathbb{P}(\mathcal{E}_1(k+i)) = O(\alpha/\sigma^i)$. Thus we obtain

$$\begin{aligned}\sum_{k<q=|V_1\cup V_2|<2k}\mathbb{P}(\mathcal{E}_1 \cap \mathcal{E}_2) &\leq \sum_{q=k+1}^{2k-1}\binom{n}{q}\binom{q}{k}^2\mathbb{P}(\mathcal{E}_1(q)) \\ &\leq \binom{2k-1}{k}^2\sum_{i=1}^{k-1}\binom{n}{k+i}\mathbb{P}(\mathcal{E}_1(k+i)) \\ &= O(1)\binom{n}{k}\alpha\sum_{i=1}^{k-1}(n/\sigma)^i.\end{aligned}$$

Since $\binom{n}{k}\alpha = \Omega(n(n/\sigma)^{k-1})$, we have $\sum_{i=1}^{k-1}(n/\sigma)^i = o(\binom{n}{k}\alpha)$. Thus we obtain $\mathbb{E}(W^2) = (1 + o(1))\mathbb{E}(W)^2$ and $\mathbb{E}(W) \to \infty$, so $\mathbb{P}(W > 0) = 1 - o(1)$ follows. This implies (3.1) and concludes the proof of Theorem 1.2. $\square$

REFERENCES

[1] M. Behzad and G. Chartrand: No graph is perfect, Amer Math. Monthly, 74 (1967) 962–963.

[2] B. Bollobás: Vertices of given degree in a random graph, Journal of Graph Theory, 6 (1982) 147–155.

[3] B. Bollobás: *Random Graphs*, Academic Press Inc. (London) Ltd. 1985.

[4] P. Balister, B. Bollobás, J. Lehel, and M. Morayne: Repeated degrees in random uniform hypergraphs, SIAM J. Discrete Math. 27 (2013) 145–154.

[5] R. J. Faudree, R. H. Schelp, M. S. Jacobson, and J. Lehel, Irregular networks, regular graphs and integer matrices with distinct row and column sums, Discrete Math., 76 (1989) 223–240.

[6] W. Feller: *An Introduction to Probability Theory and Its Applications*, Vol.1 John Wiley and Sons, Inc., Second Edition, 1966, New York, London, Sydney.

[7] A. Gyárfás, M. S. Jacobson, L. Kinch, J. Lehel, and R. H. Schelp, Irregularity strength of uniform hypergraphs, J. Combin. Math. Combin. Comput., 11 (1992) 161–172.