

# The threshold for loose Hamilton cycles in random hypergraphs

Alan Frieze\* and Xavier Pérez-Giménez†

March 6, 2025

## Abstract

We show that w.h.p. the random  $r$ -uniform hypergraph  $H_{n,m}$  contains a loose Hamilton cycle, provided  $r \geq 3$  and  $m \geq \frac{(1+\epsilon)n \log n}{r}$ , where  $\epsilon$  is an arbitrary positive constant. This is asymptotically best possible, as if  $m \leq \frac{(1-\epsilon)n \log n}{r}$  then w.h.p.  $H_{n,m}$  contains isolated vertices.

## 1 Introduction

The thresholds for the existence of Hamilton cycles in the random graphs  $G_{n,m}, G_{n,p}$  have been known for many years, see [1], [3] and [15]. There have been many generalisations of these results over the years and the problem is well understood. It is natural to try to extend these results to hypergraphs.

An  $r$ -uniform hypergraph is a pair  $H = (V, E)$  where  $E \subseteq \binom{V}{r}$ . Let  $H_{n,p}$  be the binomial random  $r$ -uniform hypergraph on  $n$  vertices with edge probability  $p$ . (The uniformity is  $r$  throughout and so we drop  $r$  from some of the notation.) Let  $H_{n,m}$  be the random  $r$ -uniform hypergraph on  $n$  vertices with  $m$  randomly chosen  $E_m$  edges from  $\binom{V}{r}$ . We are also interested in oriented hypergraphs in which each edge has one distinguished vertex called its *tail* and  $r - 1$  *heads*. Let  $\vec{H}_{n,p}$  be the oriented counterpart of  $H_{n,p}$  in which each one of the  $r \binom{n}{r}$  possible oriented edges is picked independently with probability  $p$ . Note that if we take  $\vec{H}_{n,p}$

---

\*Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh PA, USA, 15213. Research supported in part by NSF grant DMS1952285.

†Department of Mathematics, University of Nebraska-Lincoln, Lincoln NE, USA, 68588. Research supported in part by Simons Foundation Grant #587019 and by NSF grant DMS2201590.

and ignore the orientations, then the outcome is distributed as  $H_{n,p'}$  with  $p' = 1 - (1 - p)^r$ , so  $p' \sim rp$  when  $p = o(1)$ .

We say that an  $r$ -uniform sub-hypergraph  $C$  of  $H$  is a Hamilton cycle of type  $\ell$ , for some  $1 \leq \ell \leq r - 1$ , if there exists a cyclic ordering of the vertices  $V$  such that every edge consists of  $r$  consecutive vertices and for every pair of consecutive edges  $e_{i-1}, e_i$  in  $C$  (in the natural ordering of the edges) we have  $|e_{i-1} \setminus e_i| = \ell$ . When  $\ell = r - 1$  we say that  $C$  is a *loose* Hamilton cycle. Frieze [10], Dudek and Frieze [5] and Dudek, Frieze, Loh and Speiss [6] established the threshold for the existence of a loose Hamilton cycle in  $H_{n,m}$  up to a constant factor. (See also, Ferber [7] and Frankston, Kahn, Narayanan and Park [8].) We will prove the following theorem:

**Theorem 1.** *Let  $\varepsilon > 0$  be an arbitrary small positive constant. If  $r \geq 3$  and  $(r - 1) \mid n$  and  $m \geq \frac{(1+\varepsilon)n \log n}{r}$  then w.h.p.  $H_{n,m}$  contains a loose Hamilton cycle.*

If  $m \leq \frac{(1-\varepsilon)n \log n}{r}$  then w.h.p.  $H_{n,m}$  contains isolated vertices and so this is asymptotically best possible. (The constraint  $(r - 1) \mid n$  is necessary for the existence of a loose Hamilton cycle. We should also remark that Narayanan and Schacht [16] obtained the precise threshold for type  $\ell$  cycles, except for the case of loose Hamilton cycles.)

We need the following result of Altman, Greenhill, Isaev and Ramadurai [2]. They proved the following theorem: let  $H_{n,d}$  denote a random  $d$ -regular,  $r$ -uniform hypergraph on vertex set  $[n]$ .

**Theorem 2.** *Suppose that  $r \mid dn$  and  $r - 1 \mid n$ . Then,*

$$\lim_{n \rightarrow \infty} \mathbb{P}(H_{n,d} \text{ contains a loose Hamilton cycle}) = \begin{cases} 1 & d > \rho(r). \\ 0 & d \leq \rho(r). \end{cases}$$

Here  $\rho = \rho(r)$  is the unique real in  $(2, \infty)$  such that

$$(\rho - 1)(r - 1) \left( \frac{\rho r - d - r}{\rho r - \rho} \right)^{(r-1)(\rho r - \rho - r)/r} = 1.$$

We also need (the proof of) the following theorem of Kahn [13]:

**Theorem 3.** *Let  $\varepsilon > 0$  be an arbitrary positive constant. If  $r \mid n$  and  $m \geq \frac{(1+\varepsilon)n \log n}{r}$  then w.h.p.  $H_{n,m}$  contains a perfect matching.*

## 2 Proof of Theorem 1

Let  $\mathcal{K} = \binom{[n]}{r}$  denote the edges of the complete  $r$ -uniform hypergraph with vertex set  $[n]$ . Let  $N = |\mathcal{K}| = \binom{n}{r}$ . Let  $\hat{H}_{n,d}$  be the random (multi)-hypergraph in which every vertex  $v$

picks a random set of  $d$  edges that contain  $v$ . Each edge is picked independently of other choices and with replacement, so in principle one vertex could pick the same edge twice or two different vertices may pick the same edge. However, neither of these situations will occur w.h.p for  $d = d^* = \varepsilon^2 \log n$ , which is what we are interested in. We can think of  $\hat{H}_{n,d}$  as a hypergraph generalization of the  $d$ -out model of random graphs. For convenience, we may sometimes regard  $\hat{H}_{n,d}$  as oriented by making each vertex be the tail of all the edges that it picked.

We will be concerned with  $H_{n,p}$  for  $p = (1 + \varepsilon)p_*$  where  $p_* = \log n / \binom{n-1}{r-1}$  and  $\varepsilon > 0$  is a small constant.

**Theorem 4.** *Let  $d^* = \varepsilon^2 \log n$ ,  $p = (1 + \varepsilon) \log n / \binom{n-1}{r-1}$  and  $\rho \in \mathbb{N}$  constant. Then, there is a coupling in which  $H_{n,p}$  contains  $\rho$  independent copies  $\hat{\Gamma}_1, \hat{\Gamma}_2, \dots, \hat{\Gamma}_\rho$  of  $\hat{H}_{n,d^*}$  w.h.p.*

*Proof.* We will consider three types of edges: 1, 2 and 3. Edges of types 1 and 3 are oriented, i.e. they have a distinguished vertex called tail and  $r - 1$  heads. Edges of type 2 are not oriented.  $\vec{E}_1, E_2, \vec{E}_3$  will be auxiliary random hypergraphs on  $[n]$  with edges of types 1, 2, 3, respectively, and built independently of each other.  $\vec{E}_1$  is built by picking each one of the  $r \binom{n}{r}$  possible oriented edges with probability  $p_1 = 1 - (1 - \frac{\varepsilon}{2} p_*)^{1/r} \sim \frac{\varepsilon}{2r} p_*$  independently from other choices. So  $\vec{E}_1$  is distributed as  $\vec{H}_{n,p_1}$ .  $E_2$  is distributed like  $H_{n,p_2}$  (unoriented) with  $p_2 = (1 + \varepsilon/2)p_*$ . Finally,  $\vec{E}_3$  is distributed as  $\vec{H}_{n,p_3}$  with  $(1 - p_3)^r = 1 - p_2$ , so  $p_3 \sim p_2/r$ . Note that if  $E_i$  ( $i \in \{1, 3\}$ ) is obtained from  $\vec{E}_1, \vec{E}_3$  by forgetting the orientation of the vertices, then it is distributed as  $H_{n,\tilde{p}_i}$ , where  $\tilde{p}_i = 1 - (1 - p_i)^r \sim r p_i$  (more precisely,  $\tilde{p}_1 = \frac{\varepsilon}{2} p_*$  and  $\tilde{p}_3 = p_2 = (1 + \varepsilon/2)p_*$ ). Therefore, there is no hope of embedding  $E_1 \cup E_2 \cup E_3$  inside of  $H_{n,p}$  since each one of  $E_2$  and  $E_3$  have almost the same number of edges as  $H_{n,p}$ , so the union contains almost twice the needed edges. So our approach will be to consider all edges of type 1 together with some of type 2 and some of type 3 in a way that we can build the desired coupling.

We generate  $\vec{E}_1$  as the union of  $\rho$  independent copies  $\vec{E}_1^{(1)}, \dots, \vec{E}_1^{(\rho)}$  of  $\vec{H}_{n,p'_1}$  with  $(1 - p'_1)^\rho = 1 - p_1$ , so  $p'_1 \sim p_1/\rho = \frac{\varepsilon}{2r\rho} p_*$ . Given a vertex  $v$  in an oriented hypergraph, its out-edges are the edges that have  $v$  as a tail. Out-degrees and other related notions are defined the obvious way. We say  $v$  is *good* if its out-degree is at least  $d^*$  in every  $\vec{E}_1^{(1)}, \dots, \vec{E}_1^{(\rho)}$ . Otherwise we call vertex  $v$  *bad*.

**Lemma 5.** *Let  $H_1, H_2, \dots, H_\rho$  be independent copies of  $\vec{H}_{n,\alpha p_*}$  for some  $\alpha > 0$ . Let  $V_\alpha$  denote the set of vertices of degree less than  $d^*$  in at least one of the  $L_i$ . Then  $\mathbb{E}(|V_\alpha|) \leq n^{\theta_\alpha}$  where  $\theta_\alpha = \varepsilon^2(\log 1/\varepsilon^2 + 1 + \log \alpha) - \alpha$ .*

*Proof.* Let  $p_\alpha = \alpha p_*$ . Then,

$$\begin{aligned}
E(|V_\alpha|) &\leq n\rho \sum_{k=0}^{d^*} \binom{n-1}{k} p_\alpha^k (1-p_\alpha)^{(n-1)-k} \\
&\leq 2n\rho \left( \frac{\binom{n-1}{r-1} p_\alpha e}{\varepsilon^2 \log n} \right)^{\varepsilon^2 \log n} \exp \left\{ - \left( \binom{n-1}{r-1} - k \right) p_\alpha \right\} \\
&\leq 3n\rho \left( \frac{e\alpha}{\varepsilon^2} \right)^{\varepsilon^2 \log n} n^{-\alpha} \\
&= O(n^{1+\theta_\alpha}).
\end{aligned}$$

□

Putting  $\alpha = \varepsilon/(2r\rho)$  in Lemma 5 and using the Markov inequality, we see that w.h.p. all but at most a  $1/n^{\varepsilon'}$  fraction of the vertices are good, where  $\varepsilon' = \varepsilon'(\varepsilon, r, \rho) > 0$ . We call this event  $\mathcal{F}_1$ .

Now we will consider a hypergraph that contains all edges of type 1, those edges of type 2 with at most one bad vertex and those of type 3 with more than one bad vertex. More precisely, let  $V_G$  be the set of good vertices and  $V_B = [n] \setminus V_G$  the set of bad ones. Let  $E'_2 = \{e \in E_2 : |e \cap V_B| \leq 1\}$ . We can orient each edge in  $E'_2$  with exactly one bad vertex by making that bad vertex the tail, and edges with no bad vertices are oriented following any arbitrary deterministic rule. The resulting oriented hypergraph is denoted  $\vec{E}'_2$ .

Let  $\mathcal{F}_2$  be the event that every bad vertex has degree at least  $\rho d^*$  in  $E'_2$  (or equivalently out-degree at least  $\rho d^*$  in  $\vec{E}'_2$ ). For good vertices we have already found enough edges in  $E_1$  so we don't care about their degrees in  $E_2$  or  $E_3$ . Putting  $\alpha = 1 + \varepsilon/2$  in Lemma 5, we see that  $\theta_\alpha < -1$  and so  $V_\alpha = \emptyset$  w.h.p and  $\mathcal{F}_2$  holds.

Then define  $\vec{E}'_3 = \{e \in \vec{E}_3 : |e \cap V_B| > 1\}$  and let  $E'_3$  be the unoriented version of  $\vec{E}'_3$  resulting from ignoring orientations. Edges in  $E'_2$  appear with probability  $p_2$  among those edges containing at most one bad vertex. Edges in  $E'_3$  appear with probability  $1 - (1 - p_3)^r = p_2$  among the other edges that are not eligible for  $E'_2$ . In other words, even though each one of  $E'_2$  and  $E'_3$  depend on the set of good vertices (and thus on  $E_1$ ),  $E'_2 \cup E'_3$  is distributed as  $H_{n,p_2}$  and is independent of  $E_1$ . Then  $E_1 \cup E'_2 \cup E'_3$  is distributed as  $H_{n,\tilde{p}_1+p_2-\tilde{p}_1 p_2}$  which can be trivially included inside of  $H_{n,p}$  since  $\tilde{p}_1 + p_2 - \tilde{p}_1 p_2 \leq \tilde{p}_1 + p_2 = p$ . Hence, all it remains to show is that in the event that  $\mathcal{F}_1 \cap \mathcal{F}_2$  holds, we can find  $\rho$  independent copies of  $\hat{H}_{n,d^*}$  (regarded as oriented) contained in  $\vec{E}_1 \cup \vec{E}'_2 \cup \vec{E}'_3$ , which after forgetting orientations we just showed is contained in  $H_{n,p}$ . We call such copies of  $\hat{H}_{n,d^*}$   $K_1, \dots, K_\rho$ . If  $\mathcal{F}_1 \cap \mathcal{F}_2$  fails, just pick  $K_1, \dots, K_\rho$  independently from everything else so the coupling fails, but this occurs with probability  $o(1)$ .

Suppose we reveal the out-degrees of  $\vec{E}_1^{(1)}, \dots, \vec{E}_1^{(\rho)}$  and degrees of  $E'_2$  (but nothing else) and

suppose that  $\mathcal{F}_1 \cap \mathcal{F}_2$  holds. In particular, we know which vertices are good or bad. For every vertex  $v \in [n]$  and “layer”  $i = 1, \dots, \rho$ , we want to pick (with replacement)  $d^*$  out-edges of  $v$  uniformly at random and independently of all other choices. These will be the out-edges of  $v$  in  $K_i$ . If  $v$  is good, this is straightforward, since  $v$  has out-degree  $d_{v,i} \geq d^*$  in  $E_1^{(i)}$  and all sets of  $d_{v,i}$  out-neighbours are equally likely and independent of those of other vertices and layers. Incorporating replacement is also straightforward. (In other words, we are coupling  $d^*$  random independent samples with replacement from a universe with a uniformly chosen random subset of size  $d_{v,i} \geq d^*$ .)

If  $v$  is bad, picking its  $d^*$  out-edges in each  $K_i$  is trickier. We will be using edges in  $\vec{E}'_2 \cup \vec{E}'_3$  with tail  $v$ . In particular, we will only use edges in  $\vec{E}'_2$  with exactly one bad vertex, and ignore those with no bad vertices. Recall that edges of type 2 were originally unoriented, but we oriented them in  $\vec{E}'_2$  so that their only bad vertex (if they have one) is the tail. We will also use edges in  $\vec{E}'_3$  that have tail  $v$ . Most bad vertices will not pick vertices from  $\vec{E}'_3$ , but some will. Note that the edge densities of  $\vec{E}'_2$  and  $\vec{E}'_3$  are different since each out-edge of  $v$  with only one bad vertex appears in  $\vec{E}'_2$  with probability  $p_2$  (before conditioning on degrees of  $E_2$ ) while each out-edge of  $v$  with more bad vertices appears in  $\vec{E}'_3$  with probability  $p_3 \sim p_2/r$ . This is a delicate issue that complicates things and is possibly the crux of the argument. We cannot make the density of  $\vec{E}'_2$  smaller a priori since we want to make sure that bad vertices have enough out-edges and we cannot make the density of  $\vec{E}'_3$  larger since otherwise the coupling with  $H_{n,p}$  would not work.

Let us proceed to do this carefully. Let  $v$  be a bad vertex. Let  $X = X_v$  be the set of all possible out-edges of  $v$  in  $\vec{E}'_2$ , i.e. edges with tail at  $v$  and  $r - 1$  good vertices as heads. Note  $x = |X| = \binom{|V_G|}{r-1} \sim \binom{n-1}{r-1}$ . Let  $Y = Y_v$  be the set of all possible out-edges of  $v$  in  $\vec{E}'_3$ , i.e. edges with tail at  $v$  that contain at least one additional bad vertex. Clearly,  $y = |Y| = \binom{n-1}{r-1} - x \sim \frac{|V_B|}{n}x \leq n^{-\epsilon'}x$  by  $\mathcal{F}_1$ . Since we are also assuming that  $\mathcal{F}_2$  holds, the out-degree of  $v$  in  $\vec{E}'_2$  is  $s_v \geq \rho d^*$ , so the set  $S = S_v$  of out-edges of  $v$  in  $\vec{E}'_2$  is a random subset of  $X$  of size  $s_v \geq \rho d^*$ . Let  $S'$  be a random subset of  $S$  of size exactly  $\rho d^*$  (which is still uniformly chosen among all subsets of  $X$  of that size). Also, let  $T = T_v$  be the set of out-edges of  $v$  in  $\vec{E}'_3$ . Now  $t = |T|$  is distributed as  $\text{Bin}(y, p_3)$ . Our goal is to find a coupling  $U \subseteq S' \cup T$  where  $U = \{u_1, \dots, u_{\rho d^*}\}$  is obtained by uniformly and independently sampling  $\rho d^*$  random elements of  $X \cup Y$  (where recall  $X \cup Y$  is the set of all possible out-edges of  $v$ ) with replacement (i.e.  $|U| \leq \rho d^*$  and could be strictly smaller if the same edge is sampled twice).

To do this, for each  $i = 1, \dots, \rho d^*$ , let  $a_i$  be an integer in  $[x + y] = \left[ \binom{n-1}{r-1} \right]$  chosen uniformly at random (and independent of other choices). We can think of the edges in  $X \cup Y$  as having distinct labels in  $[x + y]$  so we pick  $u_i$  by first picking its label  $a_i$ , and later revealing what edge corresponds to that label. However, all we know so far is that edges in  $Y$  have labels in  $[y]$  and edges in  $X$  have labels in  $[x + y] \setminus [y]$ . The label of each individual edge is still undecided. The reason to do this is to see: 1) how many different edges are in  $U$  after possible edge repetitions, and 2) how many of these fall in  $X$  or  $Y$ . If  $a_i \leq y$ , that means

that  $u_i$  will land in  $Y$  (which has probability  $y/(x+y) \leq n^{-\varepsilon'}$ ). Otherwise, it lands in  $X$ . Let  $A = \{a_1, \dots, a_{\rho d^*}\}$ ,  $A_X = A \setminus [y]$  and  $A_Y = A \cap [y]$ . Then  $u = |U| = |A|$ ,  $u_X = |U \cap X| = |A_X|$  and  $u_Y = |U \cap Y| = |A_Y|$ . Let  $\ell$  be the number of indices  $i = 1, \dots, \rho d^*$  such that  $a_i \in [y]$  (or equivalently  $u_i \in Y$ ). Clearly,  $\ell$  has distribution  $\text{Bin}(\rho d^*, y/(x+y))$  and moreover  $u_Y \leq \ell$  by construction. If  $u_Y \leq t$ , then we can assign the  $u_Y$  labels in  $A_Y$  to a random set of edges in  $T$ . Similarly, we can always assign the  $u_X \leq \rho d^* = |S'|$  labels in  $A_X$  to a random set of edges in  $S'$ . Doing so, the  $u_1, \dots, u_{\rho d^*}$  are i.i.d. uniformly chosen from  $X \cup Y$  as desired. For all of this to work, it is enough to find a coupling in which  $\ell \leq t$ , or in other words  $\text{Bin}(\rho d^*, y/(x+y)) \leq \text{Bin}(y, p_3)$ . Ignoring  $(1 + o(1))$  factors in the parameters, all we need is a coupling

$$\text{Bin}\left(\rho \varepsilon^2 \log n, 1/n^{\varepsilon'}\right) \leq \text{Bin}\left(n^{-\varepsilon'} \binom{n-1}{r-1}, \frac{1 + \varepsilon/2 \log n}{r} \frac{1}{\binom{n-1}{r-1}}\right). \quad (1)$$

The coupling in (1) exists from the lemma below with  $N = \rho \varepsilon^2 \log n$ ,  $P = 1/n^{\varepsilon'}$ ,  $K = \frac{r}{1 + \varepsilon/2} \frac{\binom{n-1}{r-1}}{n^{\varepsilon' \log n}}$  and  $L = \frac{1 + \varepsilon/2}{\varepsilon^2 r \rho}$ .

Summarizing, for each bad vertex  $v$  we were able to sample  $u_1, \dots, u_{\rho d^*}$  uniformly at random and independently from the set of all  $\binom{n-1}{r-1}$  edges that contain  $v$  (or oriented edges with tail at  $v$ ) in a way that  $u_1, \dots, u_{\rho d^*}$  are also in  $E'_2 \cup E'_3$  (assuming  $\mathcal{F}_1$  and  $\mathcal{F}_2$ ) and thus in  $H_{n,p}$ . Then for each  $i = 1, \dots, \rho$ , we just pick the edges from  $v$  in  $K_i$  to be  $u_{1+(i-1)d^*}, \dots, u_{d^*+(i-1)d^*}$ . This gives the desired coupling and completes the proof of the theorem.  $\square$

**Lemma 6.** *Let  $0 \leq P \leq 1/2$  and  $N, K, L \in \mathbb{N}$  with  $L \geq 2$ . Then there is a coupling*

$$\text{Bin}(N, P) \leq \text{Bin}(LKN, P/K)$$

*Proof.* Let  $Y = \sum_{i=1}^N \sum_{j=1}^{LK} Y_{i,j}$  with  $Y_{i,j}$  i.i.d. Bernoulli( $P/K$ ), so  $Y$  is  $\text{Bin}(LKN, P/K)$ . We can write  $Y = \sum_{i=1}^N Y'_i$  where  $Y'_i = \sum_{j=1}^{LK} Y_{i,j}$ . Putting  $Z = \sum_{i=1}^N Z_i$  with  $Z_i = 1_{Y'_i \geq 1}$ , we have that  $Z_i \leq Y'_i$  and thus  $Z \leq Y$ . Note that  $Z$  is distributed as  $\text{Bin}(N, P_0)$  with

$$1 - P_0 = (1 - P/K)^{LK} \leq e^{-LP} \leq e^{-2P} \leq 1 - P$$

(where we used that  $L \geq 2$  and  $0 \leq P \leq 1/2$ ). Since  $P \leq P_0$ , we can easily find a  $\text{Bin}(N, P)$  random variable  $X$  and a coupling such that  $X \leq Z$ . As a result,  $X \leq Y$  as desired.  $\square$

### 3 Shamir problem with a lower bound on minimum degree

It will be convenient to replace  $\varepsilon$  by  $2\varepsilon$  and work with  $H_{n,p}, p = (1 + 2\varepsilon)p_*$ . This can be coupled to contain the union of  $\rho = \lfloor \rho(r) \rfloor + 1$  random hypergraphs  $H_0 \cup \bigcup_{i=1}^{\rho} H_i$  where  $H_0 = H_{n, (1+\varepsilon)p_*}$  and  $H_i = H_{n, \varepsilon m_*/\rho}$  where  $m_* = Np_*$ . Next let  $\widehat{\Gamma}_i, i = 1, 2, \dots, \rho$ , be

subhypergraphs of  $H_0$  as in Theorem 4 and let  $\Gamma_i = \widehat{\Gamma}_i \cup H_i, i = 1, 2, \dots, \rho$ . The aim of this section is to prove the following:

**Theorem 7.**  $\Gamma_1, \Gamma_2, \dots, \Gamma_\rho$  contain (independent uniform) perfect matchings (assuming that  $r$  divides  $n$ ), w.h.p.

Theorem 1 will follow from Theorem 7. Symmetry implies that the perfect matchings in the  $\Gamma_i$  are uniform and independence follows from the construction and Theorem 4. We apply Theorem 2 with the matchings in  $\Gamma_1, \Gamma_2, \dots, \Gamma_\rho$  to obtain Theorem 1, under the restriction on  $n$  that  $r(r-1) \mid n$ . This restriction will be removed in Section 4. (We are implicitly assuming that the union of the matchings is contiguous to  $H_{n,d}$ . This follows from [4].)

We continue to the proof of Theorem 7 and concentrate on  $\Gamma_1$ . Put quite simply, our proof is a minor modification of Kahn's proof of Theorem 3 giving the correct constant in Shamir's problem, see [13]. We let  $e_1, e_2, \dots, e_N$  be a random permutation of  $\mathcal{K}$ . Then for  $t \geq 1$ , we let  $R_t = \{e_{t+1}, e_{t+2}, \dots, e_N\}$ . We let  $\mathcal{H}_t$  denote the hypergraph with edge-set  $E_1 \cup R_t$  where  $E_1 = E(\widehat{\Gamma}_1)$ . Thus, we begin with  $\mathcal{H}_0 = \mathcal{K}$  and in a step of the process we obtain  $\mathcal{H}_{t+1}$  from  $\mathcal{H}_t$  by deleting the edge  $e_t$  from  $R_t$ . We leave  $E_1$  untouched by this process, so that  $m_t = |\mathcal{H}_t| \sim N - t + nd^*t/N$  w.h.p. and  $m_t \in [N - t, N - t + nd^*]$ . We continue for  $T = N - \varepsilon n \log n$  steps. Thus the edges of  $\Gamma_1 = \mathcal{H}_T$  are  $E_1 \cup R_T$ . We let  $E_{1,v}$  denote the  $d^*$  choices of edge by vertex  $v$  for  $E_1$ .

In principle, we could just write out Kahn's proof and make a few changes. This is a difficult proof and it would require about another 20 pages. This does not seem necessary and instead, we will give an outline of Kahn's proof and indicate the places where we need to make a change.

**Some notation:** The degree  $d_{\mathcal{H}_t}(v) = d_t(v)$  is the number of edges of  $\mathcal{H}_t$  containing  $v$  and the co-degree  $d_{\mathcal{H}_t}(v, w) = d_t(v, w)$  of  $v, w$  is the number of edges in  $\mathcal{H}_t$  containing  $v, w$ . We let  $\mathcal{H}_t - Z$  denote the subgraph of  $\mathcal{H}_t$  induced by  $[n] \setminus Z$  where  $Z \in \mathcal{K}$ . Generally speaking, from now on, we use Fraktur fonts for events, e.g.  $\mathfrak{A}, \mathfrak{B}$  etc. and calligraphic fonts for hypergraphs (and sets of edges) e.g.  $\mathcal{F}, \mathcal{H}$  etc. (By and large we use the notation of [13].)

We use the following Chernoff bounds throughout the paper:

$$\mathbb{P}(|\text{Bin}(n, p) - np| \geq \varepsilon np) \leq 2e^{-\varepsilon^2 np/3}. \quad (2)$$

$$\mathbb{P}(\text{Bin}(n, p) \geq \alpha np) \leq \left(\frac{e}{\alpha}\right)^{\alpha np}. \quad (3)$$

These also hold for the hypergeometric distribution  $\text{Hyp}(M, K, n)$ , where  $M$  is the population size,  $K$  is the number of successes in the population and  $n$  is the number of draws. Then we claim that (2), (3) hold with  $\text{Bin}(n, p)$  replaced by  $\text{Hyp}(M, K, n)$  under the assumption that  $K/M = p$ , see for example Hoeffding [11], Theorem 4.

Kahn [13] defines the following events: for a hypergraph  $\mathcal{H}$ , we let  $\Phi(\mathcal{H})$  denote the number

of perfect matchings of  $\mathcal{H}$ .

$$\mathfrak{A}_t = \left\{ \log \Phi(\mathcal{H}_t) > \log \Phi(\mathcal{H}_0) - \sum_{i=1}^t \gamma_i - o(n) \right\} \quad \text{where } \gamma_i = \frac{n}{rm_i}.$$

$\mathfrak{R}_t = \{d_t(v) \sim r|\mathcal{H}_t|/n (\sim D_t = rm_t/n)$  for a.e.  $v$ ,  $d_t(v, w) = o(D_t)$  for all  $v, w$ ,  $\Delta_t = O(D_t)$ ,  $\delta_t = \Omega(D_t)\}$  where  $\Delta, \delta$  denote maximum and minimum degrees.

$$\mathfrak{B}_t = \{\max r \mathbf{w}_t = O(1)\}$$

$$\text{where } \mathbf{w}_t(S) = \mathbf{w}_{\mathcal{H}_t}(S) = \Phi(\mathcal{H}_t - S) \text{ for } S \subseteq \mathcal{K}, \bar{\mathbf{w}}_t = N^{-1} \sum_{S \in \binom{[n]}{r}} \mathbf{w}_t(S)$$

$$\text{and } \max r \mathbf{w}_t = \frac{\max_{S \subseteq V_t} \mathbf{w}_t(S)}{\bar{\mathbf{w}}_t}.$$

Kahn then writes

$$\mathbb{P} \left( \bigcup_{t \leq T} \bar{\mathfrak{A}}_t \right) \leq \mathbb{P} \left( \bigcup_{t \leq T} \bar{\mathfrak{R}}_t \right) + \sum_{t < T} \mathbb{P}(\mathfrak{A}_t \bar{\mathfrak{R}}_t \bar{\mathfrak{B}}_t) + \sum_{t \leq T} \mathbb{P} \left( \left( \bigcap_{i < t} \mathfrak{B}_i \right) \cap \bar{\mathfrak{A}}_t \right), \quad (4)$$

where  $T = N - \beta n \log n$  as defined above. Of course, it suffices to show is that  $\mathfrak{A}_T$  occurs w.h.p.

### 3.1 Property $\mathfrak{R}_t$

Fix a vertex  $v$ . Suppose first that  $m_t \geq n^{3/2}$ . In this case  $rm_t/n \gg d^*$  and so  $D_t \sim rm_t/n$ . A vertex in  $\mathcal{H}_t$  has a degree in  $R_t$  that is hypergeometrically distributed with mean  $rm_t/n \geq rn^{1/2}$ . The Chernoff bound (2) implies that w.h.p. every  $v \in [n]$  has degree asymptotic to  $rm_t/n$ . Thus  $d_t(v) \sim D_t$  for all  $v \in [n]$  and also that  $\Delta_t = O(D_t)$ ,  $\delta_t = \Omega(D_t)$ . The co-degree of  $v, w$  is bounded by  $d^* = o(D_t)$  plus the co-degree of  $v, w$  in  $R_t$ . But the latter is distributed hypergeometrically with mean  $O(m_t n^{-2})$ . The Chernoff bounds imply that this is  $o(D_t)$  q.s.<sup>1</sup> and this verifies  $\mathfrak{R}_t$  for  $t \leq N - 2n^{3/2}$ , say.

Now assume that  $m_t < n^{3/2}$ . Then  $\mathbb{E}(|E_1 \cap R_t|) = O(n \log n \times m_t n^{-r}) = O(m_t n^{1-r} \log n) = o(1)$  and so  $E_1 \cap R_t = \emptyset$  w.h.p. for  $t > N + n^{3/2}$ . ( $E_1$  is unchanged and  $R_t$  decreases as the construction proceeds.) The argument for degrees now follows that in the previous paragraph given that the degree of a vertex  $v$  is its degree in  $H_1$  plus its degree in  $R_t$ . The Chernoff bounds imply that  $d_t(v) \in (1 \pm \log^{-1/3} n)r(m_t/n + d^*)$  for all but  $o(n)$  vertices. Thus, this time we only claim that  $d_t(v) \sim D_t$  for almost all  $v \in [n]$ . We now check co-degrees. Fix  $v, w$ . Then, let  $Z_1$  denote the number of edges in  $E_1$  coming from  $v, w$ 's choices and let  $Z_2$  be the number of edges containing  $v, w$  due to choices of  $x \neq v, w$ . Then  $Z_1$  is bounded in

<sup>1</sup>A sequence of events  $\mathfrak{R}_n, n \geq 0$  occurs quite surely (q.s.) if  $\mathbb{P}(\mathfrak{R}_n) \geq 1 - O(n^{-K})$  for any constant  $K$ .



distribution by the sum of two copies of  $\text{Bin}(d^*, \frac{r-1}{n-1})$  and  $Z_2$  is bounded in distribution by  $\text{Bin}((n-2)d^*, \frac{(r-1)(r-2)}{(n-1)(n-2)})$ . So

$$\mathbb{P}(Z_1 \geq 4) \leq \left(\frac{d^*r}{n}\right)^4 \quad \text{and} \quad \mathbb{P}(Z_2 \geq 4) \leq \left(\frac{r^2nd^*}{n^2}\right)^4.$$

Now let  $\widehat{d}_t(v, w)$  denote the co-degree of  $v, w$  in  $R_t$ . Then, where  $p_t = (N-t+1)/N$ ,

$$\mathbb{P}(\exists v, w \text{ s.t. } \widehat{d}_t(v, w) \geq k) \leq \binom{n}{2} \frac{\binom{n-2}{k} \binom{N-k}{N-t+1}}{\binom{N}{N-t+1}} \leq \binom{n}{2} \binom{n-2}{k} p_t^k \leq n^2 \left(\frac{r^2 m_t}{kn^2}\right)^k. \quad (5)$$

We put  $k = 3r$  and see that  $d_t(v, w) = o(D_t)$  w.h.p., for all  $v, w, t$ . Thus  $\mathbb{P}(\bigcup_{t \leq T} \bar{\mathfrak{A}}_t) = o(1)$ .

We finish this section with the observation that

$$\text{if } Z \in \mathcal{K} \text{ and } \mathcal{H}_t \in \mathfrak{A}_t \text{ then } \mathcal{H}_t - Z \in \mathfrak{A}_t. \quad (6)$$

To see this note that if  $v \notin Z$  then its degree in  $\mathcal{H}_t - Z$  is at most  $r \max_{v,w} d_t(v, w)$  less than it is in  $\mathcal{H}_t$ .

(To avoid problems with applying (6) repeatedly, we only apply this when  $|\mathcal{H}_t| \rightarrow \infty$ .)

### 3.2 Property $\mathfrak{A}_t$

In this section we discuss the proof that

$$\sum_{t \in I} \mathbb{P}\left(\left(\bigcap_{i < t} \mathfrak{B}_i\right) \cap \bar{\mathfrak{A}}_t\right) = o(1).$$

The argument can be lifted from Section 3 of [13]. We have

$$\log \Phi(\mathcal{H}_t) = \begin{cases} \log \Phi(\mathcal{H}_{t-1}) & e_t \in E_1. \\ \log \Phi(\mathcal{H}_{t-1}) + \log(1 - \xi_t) & e_t \notin E_1. \end{cases}$$

Recall that  $\mathcal{H}_t = \mathcal{H}_{t-1}$  if  $e_t \in E_1$ . Thus,

$$\log \Phi(\mathcal{H}_t) \geq \log \Phi(\mathcal{H}_0) + \sum_{i=1}^t \log(1 - \xi_i),$$

where  $\mathbb{E}(\xi_i) = \gamma_i$ . After this there is a careful use of martingale tail inequalities whose validity relies on the occurrence of the  $\mathfrak{B}_i$  to show that we always have  $\xi_i = O(\gamma_i)$ . This gives us a concentration of  $\sum_i \xi_i$  around its mean and also  $\sum_i \xi_i^2 = o(n)$ . The argument in [13] remains valid in our case.

### 3.3 Property $\mathfrak{B}_t$

The aim here is to prove that

$$\sum_{t \in I} \mathbb{P}(\mathfrak{A}_t \mathfrak{R}_t \bar{\mathfrak{B}}_t) = o(1). \quad (7)$$

For this, Kahn [13] introduced several new events: in the following we have dropped the subscript  $t$ .  $D = D_{\mathcal{H}} = r|\mathcal{H}|/n$  is the average degree in  $\mathcal{H} = \mathcal{H}_t$ ;  $Z$  ranges over  $\mathcal{K}$  and  $A$  ranges over the edges of  $\mathcal{H}$ .

$$\begin{aligned} \mathfrak{P}_x &= \{\mathbf{w}_{\mathcal{H}}((Z \setminus x) \cup y) \gtrsim \mathbf{w}_{\mathcal{H}}(Z)d_{\mathcal{H}}(x)/D \text{ for a.e. } y \in [n] \setminus Z\}. \\ \mathfrak{M} &= \{x \in Z \text{ and } \mathbf{w}_{\mathcal{H}}(Z) \geq \Phi(\mathcal{H})e^{-o(n)} \text{ implies } \mathfrak{P}_x\}. \\ \mathfrak{N} &= \{\mathbf{w}_{\mathcal{H}}(A) \sim \Phi(\mathcal{H})/D \text{ for a.e. } A \in \mathcal{H}\}. \\ \mathfrak{F} &= \{\mathbf{w}_{\mathcal{H}}(Z) \sim \Phi(\mathcal{H})/D \text{ for a.e. } Z \in \mathcal{K}\}. \end{aligned}$$

Kahn proves four lemmas: for events  $\mathfrak{N}_1, \mathfrak{N}_2$ , the notation  $\mathfrak{N}_1 \xrightarrow{*} \mathfrak{N}_2$  means that  $\mathbb{P}(\mathfrak{N}_1 \cap \neg \mathfrak{N}_2) = n^{-\omega(1)}$ .

**KLemma 6.1** If  $\mathcal{H}$  satisfies  $\mathfrak{A}, \mathfrak{R}$  then it satisfies  $\mathfrak{N}$ .

**KLemma 6.2**  $\{\mathcal{H} \text{ satisfies } \mathfrak{A}, \mathfrak{R}\} \xrightarrow{*} \{\mathcal{H} \text{ satisfies } \mathfrak{F}\}$ .

**KLemma 6.3** For  $x \in Z \in \binom{[n]}{2}$ ,  $\{\mathcal{H} \text{ satisfies } \mathfrak{R}\} \wedge \{\mathcal{H} - Z \text{ satisfies } \mathfrak{F}\} \xrightarrow{*} \{(\mathcal{H}, Z, x) \text{ satisfies } \mathfrak{P}_x\}$ .

**KLemma 6.4** If  $\mathcal{H}$  satisfies  $\mathfrak{R}, \mathfrak{F}, \mathfrak{M}$  then it satisfies  $\mathfrak{B}$ .

The proof of (7) follows easily from these lemmas. Given  $\mathfrak{A}, \mathfrak{R}$ , KLemma 6.1 implies that we have  $\mathfrak{N}$ . Then  $\mathcal{H} - Z$  satisfies  $\mathfrak{R}$  (see (6)) and if  $\mathbf{w}_{\mathcal{H}}(Z) \geq \Phi(\mathcal{H})e^{-o(n)}$  then  $\mathcal{H} - Z$  satisfies  $\mathfrak{A}$ . Thus, given KLemma 6.2 we have that q.s.,  $\mathcal{H} - Z$  satisfies  $\mathfrak{F}$  whenever  $\mathbf{w}_{\mathcal{H}}(Z) \geq \Phi(\mathcal{H})e^{-o(n)}$ . Then, KLemma 6.3 gives us that  $\mathfrak{M}$  holds q.s. KLemma 6.4 then implies that  $\mathfrak{B}$  is satisfied and (7) follows. Property  $\mathfrak{P}_x$  is key here. Starting with  $Z$  maximizing  $\mathbf{w}_t$ , it shows that  $\max \mathbf{w}_t = O(\bar{\mathbf{w}}_t)$ .

Now  $\mathcal{H}$  has a different distribution to that considered in [13]. But by looking at the proofs in [13] we shall see that this does not matter. KLemmas 6.1 and 6.4 are actually deterministic statements. It is only the proofs of KLemmas 6.2, 6.3 that are sensitive to the distribution of  $\mathcal{H}$ . Fortunately, our  $\mathcal{H}$  meets the necessary requirements. We briefly indicate the main points of the proofs of these lemmas to show why the change in distribution does not matter.

**KLemma 6.1** The proof in [13] begins with "Here  $\mathcal{H}$  is a general  $m$ -edge ( $n$ -vertex)  $r$ -graph satisfying  $\mathfrak{A}, \mathfrak{R}$ ." Given  $\mathcal{H}$  satisfying  $\mathfrak{A}, \mathfrak{R}$ , Kahn first shows that  $h(v, \mathcal{H}) > \log d(v) - o(1)$  for a.e. vertex  $v$ , where  $d(v)$  is the degree of  $v$  in  $\mathcal{H}$  and  $h(v, \mathcal{H})$  is the entropy of the distribution of the edge containing  $v$  in a uniform random perfect matching of  $\mathcal{H}$ . This shows that in a random p.m.  $v$  is contained in a nearly random incident edge and  $\mathfrak{N}$  quickly follows.

**Klemma 6.2** This uses the technical assumption that

$$|\mathcal{K} \setminus \mathcal{H}| \geq cN \quad (8)$$

for some absolute constant  $c > 0$ . Kahn points out that without this assumption  $\mathfrak{N}$  and  $\mathfrak{F}$  are equivalent.

Given  $\mathcal{H}$  satisfying  $\mathfrak{A}, \mathfrak{R}$ , let  $\theta$  be an arbitrary positive constant, let  $\Phi' = \Phi(\mathcal{H})/D$  and let

$$\begin{aligned} \mathfrak{Q} &= \{\mathbf{w}_{\mathcal{H}}(A) \sim \Phi' \text{ for a.e. } A \in \mathcal{H}, \\ &\text{but } \mathbf{w}_{\mathcal{H}}(U) \neq (1 \pm 2\theta)\Phi' \text{ for at least a } (2\theta)\text{-fraction of the } U\text{'s in } \mathcal{K} \setminus \mathcal{H}\}. \end{aligned}$$

Kahn reduces the lemma to showing that

$$\mathbb{P}(\mathfrak{A}\mathfrak{R}\mathfrak{Q}) < n^{-\omega(1)}. \quad (9)$$

Let  $\mathcal{T}$  be chosen uniformly from  $\binom{\mathcal{H}}{\tau}$  for a suitable  $\log n \ll \tau \ll \log^2 n$  and let  $\mathcal{F} = \mathcal{H} \setminus \mathcal{T}$ . Let  $\zeta = e^{-\tau/D}$  and define the event

$$\begin{aligned} \mathfrak{B} &= \{\mathbf{w}_{\mathcal{F}}(A) \sim \zeta\Phi' \text{ for a.e. } A \in \mathcal{T}, \\ &\text{but } \mathbf{w}_{\mathcal{F}}(U) \neq (1 \pm \theta)\zeta\Phi' \text{ for at least a } \theta\text{-fraction of the } U\text{'s in } \mathcal{K} \setminus \mathcal{H}\}. \end{aligned}$$

Note that (9) is implied by

$$\mathbb{P}(\mathfrak{B} \mid \mathfrak{A}\mathfrak{R}\mathfrak{Q}) = 1 - o(1). \quad (10)$$

$$\mathbb{P}(\mathfrak{B}) = n^{-\omega(1)}. \quad (11)$$

For equation (10) Kahn carefully chooses  $\tau$  and then observes "For the proof of (10) we choose  $\mathcal{H}$  and then  $\mathcal{T}$ . We assume we have chosen  $\mathcal{H}$  satisfying  $\mathfrak{A}, \mathfrak{R}, \mathfrak{Q}$ ; so  $\mathbb{P}$  now refers just to the choice of  $\mathcal{T}$ , and (10) will follow from  $\mathbb{P}(\mathcal{H}, \mathcal{T} \text{ satisfies } \mathfrak{B}) = 1 - o(1)$ ".

For equation (11) we choose  $\mathcal{F} = \mathcal{H} \setminus \mathcal{T}$  and then  $\mathcal{T}$ . More precisely, we need to choose a disjoint pair  $\mathcal{F}, \mathcal{T}$  such that  $\mathcal{F} \cup \mathcal{T}$  is distributed as  $\mathcal{H}$  and  $\mathcal{T}$  is distributed as a uniform random subset of  $\mathcal{F} \cup \mathcal{T}$ . So we choose  $\mathcal{H}'$  with the same distribution as  $\mathcal{H}$ , then choose a random subgraph  $\mathcal{T}'$  of  $\mathcal{H}'$  of size  $\tau$ . The hypergraph  $\mathcal{H}' \setminus \mathcal{T}'$  will be our  $\mathcal{F}$ . Then we choose  $\mathcal{T} = \mathcal{T}_1 \cup \mathcal{T}_2$  independently as follows: we choose  $\mathcal{T}_1$  of size  $\text{Bin}(\tau, \alpha_1)$ ,  $\alpha_1 = (N - t + 1)/(N - t + 1 + d^*n)$  from  $\mathcal{K} \setminus \mathcal{F}$  and  $\mathcal{T}_2$  of size  $\tau - |\mathcal{T}_1|$  from  $E_1 \setminus \mathcal{F}$ . Note that  $\alpha_1 \geq \varepsilon/(\varepsilon + \varepsilon^2) \geq 0.99$  and so  $|\mathcal{T}_1| \geq 9\tau/10$  q.s..

Given  $\mathcal{F}$ , let  $U_1, U_2, \dots$ , be an ordering of  $\mathcal{K} \setminus \mathcal{F}$  with  $\mathbf{w}_{\mathcal{F}}(U_1) \leq \mathbf{w}_{\mathcal{F}}(U_2) \leq \dots$  and let  $\mathcal{Y}, \mathcal{Z}$  be the first and last  $\theta|\mathcal{K} \setminus \mathcal{F}|/3$  of the  $U_i$ 's. Now at least one of  $\mathcal{Y}, \mathcal{Z}$  is contained in  $\mathcal{W} = \{U : \mathbf{w}_{\mathcal{F}}(U) \neq (1 \pm \theta)\zeta\Phi'\}$ . Otherwise, using  $\tau \ll |\mathcal{K} \setminus \mathcal{H}|$  (from (8)), we have  $|\mathcal{W} \setminus \mathcal{H}| < |\mathcal{Y}| + |\mathcal{Z}| < 2\theta(|\mathcal{K} \setminus \mathcal{H}| + \tau)/3 < \theta|\mathcal{K} \setminus \mathcal{H}|$ , contradicting the second part of  $\mathfrak{B}$ .

Now the first part of  $\mathfrak{B}$  implies that  $\min\{|\mathcal{T} \cap \mathcal{Y}|, |\mathcal{T} \cap \mathcal{Z}|\} \leq \theta\tau/10$ , say. But  $|\mathcal{T}_1 \cap \mathcal{Y}|$  is distributed as  $\text{Bin}(|\mathcal{T}_1|, \theta/3)$  and because  $|\mathcal{T}_1| \geq 9\tau/10$  q.s. we have that  $\min\{|\mathcal{T}_1 \cap \mathcal{Y}|, |\mathcal{T}_1 \cap \mathcal{Z}|\} \geq \theta\tau/4$  q.s. This contradiction implies that (11) holds.

**Klemma 6.3** Let  $\mathcal{H}' = \{A \in \mathcal{H} : A \cap Z = \{x\}\}$  and  $\mathcal{H}'' = \mathcal{H} \setminus \mathcal{H}'$ . Let  $Y = Z \setminus x$  and  $W = [n] \setminus Z$ . Property  $\mathfrak{R}$  for  $\mathcal{H}$  implies that

$$D_{\mathcal{G}} \sim D_{\mathcal{H}}, \quad \text{where } \mathcal{G} = \mathcal{H} - Z.$$

$$d'(x) = |\mathcal{H}'| \left( = \left| \left\{ S \in \binom{W}{r-1} : S \cup x \in \mathcal{H} \right\} \right| \right) \sim d_{\mathcal{H}}(x) = \Omega(D). \quad (12)$$

We choose  $\mathcal{H}''$  first, which determines  $\mathcal{G}$ . We can assume (12) and

$$\mathbf{w}_{\mathcal{G}}(U) \sim \Phi' = \frac{\Phi(\mathcal{G})}{D} \text{ for a.e. } U \in \binom{V \setminus Z}{r}. \quad (13)$$

After this we choose  $\mathcal{H}'$  and KLemma 6.3 will follow from

$$\mathbb{P}(\mathfrak{P}_x \mid (12), (13)) = 1 - n^{-\omega(1)}. \quad (14)$$

From (13) we have that for a.e.  $y \in W$ ,

$$\mathbf{w}_{\mathcal{F}}(S \cup y) \sim \Phi' \text{ for a.e. } S \in \binom{W \setminus y}{r-1}. \quad (15)$$

For a  $y$  as in (15) and a random choice of  $\mathcal{H}'$  (as described in Section 1), the Chernoff bounds imply that q.s., for all but  $o(d'(x))$  sets  $S$  such that  $S \cup x \in \mathcal{H}$ , we have  $\mathbf{w}_{\mathcal{F}}(Y \cup y) \sim \Phi'$ . As this is another place where the distribution of  $\mathcal{H}$  matters, we go into a little more detail. So, given  $y$ , let  $\mathcal{S}$  denote the collection of sets  $S \in \binom{W \setminus y}{r-1}$  such that (15) does not hold and let  $\eta = |\mathcal{S}| / \binom{n-r-1}{r-1} = o(1)$ . Then the number  $Z_1$  of edges in  $R_1$  of the form  $S \cup x \in \mathcal{H}'$ ,  $S \in \mathcal{S}$  is dominated in distribution by  $\text{Bin}(N-t+1, r\eta/n)$  and so  $Z_1 = o(D)$  q.s. The same argument shows that the number  $Z_2$  of edges in  $E_{1,v}$  of the form  $S \cup x \in \mathcal{H}'$ ,  $S \in \mathcal{S}$  satisfies  $Z_3 = o(D)$  q.s. Because  $d'(x) = \Omega(D)$ , we have  $Z_1 + Z_2 = o(d'(x))$  as required.

Now

$$\text{if } y \in W \text{ then } \mathbf{w}_{\mathcal{H}}(Y \cup y) = \sum_{S \in \binom{W \setminus y}{r-1}, S \cup x \in \mathcal{H}} \mathbf{w}_{\mathcal{F}}(S \cup y). \quad (16)$$

And so  $\mathbf{w}_{\mathcal{H}}(Y \cup y) \gtrsim \Phi' d'(x) \sim \mathbf{w}_{\mathcal{H}}(Z) d_{\mathcal{H}}(x) / D$ .

**Klemma 6.4** We define the property

$$\mathfrak{D} = \{\text{If } \mathbf{w}_{\mathcal{H}}(Z_0) > \Phi e^{-o(n)} \text{ then } \mathbf{w}_{\mathcal{H}}(Z) \gtrsim \mathbf{w}_{\mathcal{H}}(Z_0) D^{-r} \prod_{z \in Z_0} d_{\mathcal{H}}(x) \text{ for a.e. } Z \in \mathcal{K}\}.$$

After this we assert:

$$\mathfrak{RM} \text{ implies } \mathfrak{D}. \quad (17)$$

$$\mathfrak{RD}\mathfrak{F} \text{ implies } \mathfrak{B}. \quad (18)$$

If  $Z_0 = \{x_1, \dots, x_r\}$  then we obtain (17) by induction on  $i \in [r]$ , for a.e.  $y_1, \dots, y_r \in [n] \setminus Z_0$ , we have with  $Z_i = (Z_{i-1} \setminus x_i) \cup y_i$ ,

$$\forall i \mathbf{w}_{\mathcal{H}}(Z_i) \gtrsim \mathbf{w}_{\mathcal{H}}(Z_{i-1}) d_{\mathcal{H}}(x_i) / D \gtrsim \mathbf{w}_{\mathcal{H}}(Z_0) D^{-i} \prod_{j \leq i} d_{\mathcal{H}}(x_j). \quad (19)$$

(The only thing to observe here is that (19) for  $Z_{i-1}$  implies  $\mathbf{w}_{\mathcal{H}}(Z_i) \geq \Phi(\mathcal{H}) e^{-o(n)}$ , since  $\delta_{\mathcal{H}} = \Omega(D)$  implies that the r.h.s. of (19) is  $\Omega(\mathbf{w}_{\mathcal{H}}(Z_0))$ .) This gives  $\mathfrak{D}$ , since it implies that a.e.  $Z \in \mathcal{K}$  is  $Z_r$  for some  $y_1 \dots y_r$  supporting (19).

For (18) choose  $Z_0 \in \mathcal{K}$  with  $\mathbf{w}_{\mathcal{H}}(Z_0)$  maximum and note that  $\mathbf{w}_{\mathcal{H}}(Z_0) > \Phi(\mathcal{H}) e^{-o(n)}$ . Thus  $\mathfrak{D}$  (and  $\delta_{\mathcal{H}} = \Omega(D)$ ) give  $\mathbf{w}_{\mathcal{H}}(Z) = \Omega(\mathbf{w}_{\mathcal{H}}(Z_0))$  for a.e.  $Z \in \mathcal{K}$ , which with  $\mathfrak{F}$  implies  $\Phi(\mathcal{H})/D = \Omega(\mathbf{w}_{\mathcal{H}}(Z_0))$ . But this gives  $\mathfrak{B}$ , since  $\overline{\mathbf{w}}_{\mathcal{H}}(\mathcal{H}) = \Phi(\mathcal{H})/D$ .

## 4 Dealing with the divisibility problem

We use an idea of Ferber [7]. Suppose that  $n = (r-1)(r\hat{n} + k)$  where  $0 \leq k < r$ . We can use  $k = 0$  as the base case for an inductive proof. So, suppose that  $\mathfrak{X}_k$  is the statement that  $H_{n,m}$  contains a loose hamilton cycle w.h.p. when  $(r-1) \mid n$  and  $n/(r-1) \equiv k \pmod{r}$  and  $m \geq \frac{(1+\varepsilon)n \log n}{r}$  for any fixed  $\varepsilon > 0$ . Assume then that  $\mathfrak{X}_k$  is true. We will show using Ferber's idea that if  $n = (r-1)(r\hat{n} + k + 1)$  then there exists a positive constant  $\alpha$  such that if  $m \geq \frac{(1+\varepsilon)n \log n}{r}$  then  $H_{n,m}$  contains a loose Hamilton cycle with probability at least  $\alpha$ . (In fact, we will show that  $\alpha = 1/(2r)$  works.) After this, we can use a result of Friedgut [9] that says that there are no coarse thresholds with  $m \sim n \log n$ . In particular, we apply the second remark following Theorem 2.1 of that paper to obtain w.h.p.

Following Ferber, we let  $p = m/N$  and then let  $p_1 = \log n/N$  and define  $p_2$  by  $1 - p = (1 - p_1)(1 - p_2)$ . Fix an arbitrary edge  $e = \{x_1, x_2, \dots, x_k\} \in E(H(n, p_1))$  and note that such an edge exists w.h.p. Now let  $V^* = ([n] \setminus e) \cup \{e^*\}$ , where  $e^*$  is a new vertex. Note that  $|V^*| = n - r + 1 = (r-1)(r\hat{n} + k)$ . Now consider the hypergraph  $H^*$  with vertex set  $V^*$  derived from  $H_{n,p_2}$  by replacing an  $x_i$  in any edge by  $e^*$  and accepting it with probability  $\frac{1 - (1 - p_2)^{1/r}}{p_2} \sim 1/r$ . The edge density of  $H^*$  is such that we can apply the induction hypothesis to argue that  $H^*$  contains a loose Hamilton cycle  $C^*$  w.h.p. Now by symmetry, there is a probability of at least  $1/r$  that  $\{e^*\} = e_1^* \cap e_2^*$  for two consecutive edges of  $C^*$ . Suppose that  $e_i^*$ ,  $i = 1, 2$ , are derived from  $e_1, e_2$  of  $H_{n,p_2}$ . Given this there is a further probability of at least  $1 - 1/r \geq 1/2$  that we can replace  $e_1^*, e_2^*$  by  $e_1, e_2$  to obtain a loose Hamilton cycle in  $H_{n,p_2}$ .  $\square$

## 5 Final remarks

Kahn [14] also proved a hitting time version for perfect matchings. It is not obvious how to tighten the current approach to get a hitting time version for loose hamilton cycles. This is because we need a minimum degree of  $\rho > 2$  to be able to apply Theorem 2.

## References

- [1] M. Ajtai, J. Komlós and E. Szemerédi, The first occurrence of Hamilton cycles in random graphs, *Annals of Discrete Mathematics* 27 (1985) 173-178.
- [2] D. Altman, C. Greenhill, M. Isaev and R. Ramadurai, A threshold result for loose Hamiltonicity in random regular uniform hypergraphs, *Journal of Combinatorial Theory B* 142 (2020) 307-373.
- [3] B. Bollobás, The evolution of random graphs, *Transactions of the American Mathematical Society* 286 (1984) 257-274.
- [4] C. Cooper, A.M. Frieze and B. Reed, Perfect matchings in random  $r$ -regular,  $s$ -uniform hypergraphs, *Combinatorics, Probability and Computing* 5 (1996) 1-15.
- [5] A. Dudek and A.M. Frieze, *Loose Hamilton cycles in random uniform hypergraphs*, *Electronic Journal of Combinatorics* 18 (2011), #P48.
- [6] A. Dudek, A.M. Frieze, P. Loh and S. Speiss, Optimal divisibility conditions for loose Hamilton cycles in random hypergraphs, *Electronic Journal of Combinatorics* 19 (2012).
- [7] A. Ferber, Closing gaps in problems related to Hamilton cycles in random graphs and hypergraphs, *Electronic Journal of Combinatorics*, 21, (2015).
- [8] K. Frankston, J. Kahn, B. Narayanan and J. Park, Thresholds versus fractional expectation-thresholds, *Annals of Mathematics* 194 (2021) 475-495.
- [9] E. Friedgut, Hunting for Sharp Thresholds, *Random Structures Algorithms* 26 (2005) 37-51.
- [10] A.M. Frieze, *Loose Hamilton cycles in random 3-uniform hypergraphs*, *Electronic Journal of Combinatorics* 17 (2010), #N28.
- [11] W. Hoeffding, Probability inequalities for sums of bounded random variables, *Journal of the American Statistical Association* 58 (1963) 13-30.
- [12] A. Johansson, J. Kahn and V. Vu, Factors in Random Graphs, *Random Structures and Algorithms* 33 (2008) 1-28.
- [13] J. Kahn, Asymptotics for Shamir’s problem, *Advances in Mathematics* 422 (2023)

- [14] J. Kahn, Hitting times for Shamir's Problem, *Trans. Amer. Math. Soc.* 375 (2022) 627-668.
- [15] J. Komlós and E. Szemerédi, Limit distributions for the existence of Hamilton circuits in a random graph, *Discrete Mathematics* 43 (1983) 55-63.
- [16] B. Narayanan and M. Schacht, Sharp thresholds for nonlinear Hamiltonian cycles in hypergraphs, *Random Structures and Algorithms* 57 (2020) 244-255.
- [17] N. Wormald, Models of random regular graphs, *In J.D. Lamb and D.A. Preece, editors, Surveys in Combinatorics, 1999, volume 267 of London Mathematical Society Lecture Note Series, Cambridge University Press* (1999) 239-298.