Almost perfect matchings in random uniform hypergraphs

Michael Krivelevich *
Department of Mathematics,
Raymond and Beverly Sackler Faculty of Exact Sciences,
Tel Aviv University, Tel Aviv, 69978, Israel

April 20, 2001

Abstract

We consider the following model $\mathcal{H}_r(n,p)$ of random r-uniform hypergraphs. The vertex set consists of two disjoint subsets V of size $|V| = n$ and U of size $|U| = (r-1)n$. Each r-subset of $V \times \binom{U}{r-1}$ is chosen to be an edge of $H \in \mathcal{H}_r(n,p)$ with probability $p = p(n)$, all choices being independent. It is shown that for every $0 < \epsilon < 1$ if $p = \frac{C \ln n}{n^{r-1}}$ with $C = C(\epsilon)$ sufficiently large, then almost surely *every* subset $V_1 \subset V$ of size $|V_1| = \lfloor (1-\epsilon)n \rfloor$ is matchable, that is, there exists a matching M in H such that every vertex of V_1 is contained in some edge of M.

An *r-uniform hypergraph* H is an ordered pair $H = (V, E)$, where $V = V(H)$ is a finite set (the set of *vertices*) and $E = E(H)$ is a collection of distinct subsets of V of size r, called *edges*. In this paper the parameter $r \geq 2$ is assumed to be a fixed number. A subset $M \subseteq E(H)$ is called

*This research forms part of a Ph. D. thesis written by the author under the supervision of Professor Noga Alon.

a *matching* if every pair of edges from M has an empty intersection. A matching M is called *perfect* if $|M| = |V|/r$ (clearly, a perfect matching can exist only if r divides $|V|$). A subset $V_1 \subseteq V$ is *matchable* in $H = (V, E)$ if there exists a matching M in H so that every vertex from V_1 is contained in some edge from M.

A *random r-uniform hypergraph* $\mathcal{H}_r(n,p)$ is an r-uniform hypergraph with vertex set V of size $|V| = n$, in which each r-subset of V is chosen to be an edge of $H \in \mathcal{H}_r(n,p)$ with probability p (where p may depend on n), all choices being independent.

One of the central problems in probabilistic combinatorics is that of determining the minimal probability $p = p(n)$, for which a random hypergraph $H \in \mathcal{H}_r(n,p)$ has **whp** [1] a perfect matching (assuming of course that r always divides n). This problem was posed by Schmidt and Shamir in [6], they managed to prove that if $p(n) = n^{-r+3/2}w(n)$, where $w(n)$ is any function tending to infinity arbitrarily slowly, then **whp** $H \in \mathcal{H}_r(n,p)$ has a perfect matching. This result has recently been improved by Frieze and Janson [3], who showed that it suffices to take $p(n) = n^{-r+4/3}w(n)$. Both papers used the second moment method and the Chebyshev inequality. A fractional version of this problem is considered in [5], where it is shown that if $p(n) = (\ln n + w(n)) \Big/ \binom{n-1}{r-1}$, then a random hypergraph $H \in \mathcal{H}_r(n,p)$ has **whp** a perfect *fractional* matching of size n/r (that is, an assignment $f : E(H) \to R^+$ of non-negative

[1] An event $\mathcal{E}_n$ happens **whp** (with high probability) if the probability of $\mathcal{E}_n$ tends to 1 as n tends to infinity.

weights to the edges of H such that $\sum_{v\in e} f(e) \le 1$ for every $v \in V$ and also $\sum_{e\in E} f(e) = n/r$. As for the existence of an almost perfect matching (that is, a matching covering all but $o(n)$ vertices), de la Vega proved in [7], using Markov chains and the Chebyshev inequality, that if $p(n) = w(n)/n^{r-1}$ for any function $w(n) \to \infty$, then a random hypergraph $H \in \mathcal{H}_r(n,p)$ **whp** contains a matching M of size $(1-o(1))n/r$.

In this paper we treat a different model of random r-uniform hypergraphs, which however has many features similar to those of $\mathcal{H}_r(n,p)$. In this new model which we denote by $\mathcal{H}'_r(n,p)$, the vertex set consists of two disjoint subsets V of size $|V| = n$ and U of size $|U| = (r-1)n$. Each r-subset of $V \times \binom{U}{r-1}$ is chosen to be an edge of a random hypergraph $H \in \mathcal{H}'_r(n,p)$ with probability $p = p(n)$, all choices being independent. This model can be considered as a model of *random bipartite r-uniform hypergraphs* (adopting the terminology of [1]). We try to estimate the minimal probability $p = p(n)$, for which different subsets of V (including V itself) are matchable **whp** in $H \in \mathcal{H}'_r(n,p)$. Clearly, if H contains a perfect matching, then every subset $V_1 \subseteq V$ is matchable, so the most important problem is to determine the threshold probability $p = p(n)$ for the existence of a perfect matching in $\mathcal{H}'_r(n,p)$. If $H \in \mathcal{H}'_r(n,p)$ has a perfect matching, then every vertex $v \in V \cup U$ is contained in at least one edge, therefore the condition of non-existence of isolated vertices is a necessary condition for the existence of a perfect matching. It can be proven rather easily, using standard methods of random graphs theory (see, e.g., [2]), that if $p \geq (\ln n + w(n)) \Big/ \binom{(r-1)n}{r-1}$, then **whp** every vertex $v \in V \cup U$ is incident with at least one edge. This motivates the following conjecture.

Conjecture 1 *If* $p \geq \frac{\ln n + w(n)}{\binom{(r-1)n}{r-1}}$*, where* $w(n)$ *is any function tending to infinity arbitrarily slowly with* n*, then* **whp** *a random hypergraph* $H \in \mathcal{H}'_r(n,p)$ *contains a perfect matching.*

At this stage, we can not prove this conjecture for any $r \geq 3$. (For the case $r = 2$, that is, when $\mathcal{H}'_r(n,p)$ consists of bipartite graphs with equal sides, the conjecture is true, and the proof can be found, e.g., in [2]). The main difficulty in proving it seems to originate from the lack of appropriate combinatorial results (such as the Hall-König theorem), providing sufficient conditions for the existence of a perfect matching in an r-uniform hypergraph. Instead, we can prove that, for every constant $0 < \epsilon < 1$, if the probability p from Conjecture 1 is multiplied by a constant factor, then **whp** *every* subset $V_1 \subset V$ of size $|V_1| = \lfloor (1-\epsilon)n \rfloor$ is matchable in $H \in \mathcal{H}'_r(n,p)$. This is formally stated in the following theorem.

Theorem 1 *For every constant* $0 < \epsilon < 1$*, if* $C = C(\epsilon) = \left(\frac{15}{\epsilon}\right)^{r-1}$*, then* **whp** *in a random hypergraph* $H \in \mathcal{H}'_r(n,p)$*, where* $p = p(n) = \frac{C \ln n}{n^{r-1}}$*,* **every** *subset* $V_1 \subset V$ *of size* $|V_1| = \lfloor (1-\epsilon)n \rfloor$ *is matchable.*

(The constant 15 in the expression for $C(\epsilon)$ can certainly be improved. We do not make any attempt to optimize it here).

Note that Conjecture 1 implies immediately Theorem 1.

We remark that in contrast with the above mentioned result of de la Vega [7], we need to show not only the existence of a single large matching, but the existence of a matching for *every* large subset $V_1 \subset V$.

Our proof is based on a recent result of Haxell [4], providing a sufficient condition for matchability for a certain class of uniform hypergraphs. Let us reformulate her result in a form, convenient for our purposes.

Theorem 2 *Suppose V_0 and U_0 are disjoint sets of vertices, and suppose every edge e of a hypergraph $H = (V_0 \cup U_0, E)$ satisfies $|e \cap V_0| = 1$, $|e \cap U_0| = r-1$. Then if for every non-empty subset $W_0 \subseteq V_0$ there exist $m_0 = (2r-3)|W_0|$ edges $e_1, \ldots, e_{m_0} \in E(H)$ such that $e_i \cap W_0 \neq \emptyset$ for all $1 \leq i \leq m_0$ and also $e_i \cap e_j \subset W_0$ for every $1 \leq i \neq j \leq m_0$, then V_0 is matchable in H.*

One may see easily that Theorem 2 reduces to Hall's theorem for the case $r = 2$.

We can not apply Theorem 2 directly to a random hypergraph $H \in \mathcal{H}_r'(n,p)$, because if we take $W_0 \subset V$, $|W_0| = \lfloor (1-\epsilon)n \rfloor$, then there is no place in U for $(2r-3)|W_0|$ edges satisfying the conditions of Theorem 2. Instead, we act as follows. Fix a subset $U_0 \subset U$ of size $|U_0| = \lfloor \epsilon |U| \rfloor = \lfloor \epsilon (r-1)n \rfloor$. Then, for a subset $V_1 \subset V$ of size $|V_1| = \lfloor (1-\epsilon)n \rfloor$, we first match almost all (say, $|V_1| - k$, for $k = \left\lfloor \frac{\epsilon n}{10r} \right\rfloor$) vertices of V_1 inside $U \setminus U_0$. Denote the subset of yet unmatched vertices of V_1 by V_0, since $|V_0| < |U_0|/((2r-3)(r-1))$, we can now apply (as shown in Lemma 1 below) Theorem 2 to show that

V_0 is matchable inside U_0, thus obtaining the desired matching for V_1. The realization of this idea is based on the following

Lemma 1 *Denote* $k = \left\lfloor \frac{\epsilon n}{10r} \right\rfloor$. *Let* $U_0 \subset U$ *be a fixed set of* $|U_0| = \lfloor \epsilon(r-1)n \rfloor$ *vertices. Then* **whp** *a random hypergraph* $H \in \mathcal{H}_r'(n,p)$ *has the following properties:*

1. *for every subset* $W_1 \subseteq V$ *of size* $|W_1| > k$ *and for every subset* $U_1 \subseteq U$ *of size* $|U_1| = (r-1)|W_1|$ *there exists an edge of* H *inside* $W_1 \cup U_1$*;*

2. *for every subset* $W_0 \subset V$ *of size* $1 \leq |W_0| \leq k$ *there exist* $m_0 = (2r-3)|W_0|$ *edges* $e_1, \ldots, e_{m_0}$ *such that* $e_i \cap W_0 \neq \emptyset$ *for all* $1 \leq i \leq m_0$*, and* $e_i \cap U \subset U_0$ *for all* $1 \leq i \leq m_0$*, and also* $e_i \cap e_j \subset W_0$ *for all* $1 \leq i \neq j \leq m_0$.

Proof. 1) Clearly it suffices to prove the required statement for the case $|W_1| = k$, $|U_1| = (r-1)k$. The probability of the existence of a pair W_1, U_1 of sizes $|W_1| = k$, $|U_1| = (r-1)k$, contradicting Part 1 of the lemma, can be bounded from above by

$$\binom{n}{k}\binom{(r-1)n}{(r-1)k}(1-p)^{k\binom{(r-1)k}{r-1}} \leq 2^{n+(r-1)n} e^{-k\binom{(r-1)k}{r-1}p}$$
$$\leq 2^{rn} e^{-k \cdot k^{r-1} \frac{C \ln n}{n^{r-1}}} \leq 2^{rn} e^{-\Theta(n \ln n)} = o(1) \ ;$$

2) Denote by p_i, $1 \leq i \leq k$, the probability of the existence of a subset $W_0 \subset V$ of size $|W_0| = i$, contradicting Part 2 of the lemma. We will show that $p_i = o(n^{-1})$. Let $W_0 \subset V$, $|W_0| = i$, and assume W_0 contradicts

the lemma statement. Denote by E_0 a maximum set of edges lying inside $W_0 \cup U_0$ and having their pairwise intersections in W_0, then according to our assumption about W_0 we have $|E_0| < (2r-3)i$. Let $U' \subset U$ be the set of all vertices of U, lying on any edge from E_0, then $|U'| < (2r-3)(r-1)i$. Since E_0 has maximal cardinality, every other edge of H inside $W_0 \cup U_0$ intersects U'. This implies that the subset $W_0 \cup (U_0 \setminus U')$ spans no edges from H. Therefore we may estimate

$$
\begin{aligned}
p_i &\le \binom{n}{i}\binom{\lfloor \epsilon(r-1)n \rfloor}{(2r-3)(r-1)i}(1-p)^{i\left(\frac{\lfloor \epsilon(r-1)n\rfloor-(2r-3)(r-1)i}{r-1}\right)} \\
&\le \binom{n}{i}\binom{\lfloor \epsilon(r-1)n \rfloor}{(2r-3)(r-1)i} e^{-i\left(\frac{\lfloor \epsilon(r-1)n\rfloor-(2r-3)(r-1)i}{r-1}\right)p} \\
&\le \left(\frac{en}{i}\right)^i \left(\frac{\epsilon e n}{(2r-3)i}\right)^{(2r-3)(r-1)i} \exp\left\{-i\left(\epsilon n - \frac{(2r-3)\epsilon n}{10r}\right)^{r-1} \frac{C\ln n}{n^{r-1}}\right\} \\
&\le \left[\frac{en}{i}\left(\frac{\epsilon e n}{(2r-3)i}\right)^{(2r-3)(r-1)} n^{-\frac{C}{n^{r-1}}\left(\frac{4\epsilon n}{5}\right)^{r-1}}\right]^i .
\end{aligned}
$$

(We use inequalities $\left(\frac{a}{b}\right)^b \le \binom{a}{b} \le \left(\frac{ea}{b}\right)^b$).

Substituting $i = 1$ in the first and the second fractions in the brackets and then assigning $C = C(\epsilon) = \left(\frac{15}{\epsilon}\right)^{r-1}$, we get

$$
\begin{aligned}
p_i &\le \left[O\left(n^{1+(2r-3)(r-1)-\frac{C}{n^{r-1}}(\epsilon n)^{r-1}\left(\frac{4}{5}\right)^{r-1}}\right)\right]^i \\
&= \left[O\left(n^{2r^2-5r+4-C\left(\frac{4\epsilon}{5}\right)^{r-1}}\right)\right]^i \\
&= \left[O\left(n^{3r^2-1-12^{r-1}}\right)\right]^i = o(n^{-1}) ,
\end{aligned}
$$

as claimed. □

Proof of Theorem 1. Assume $H \in \mathcal{H}'_r(n,p)$ has the properties stated in Lemma 1. Let $V_1 \subset V$ be a subset of size $|V_1| = \lfloor (1-\epsilon)n \rfloor$. First, using the edges of H inside $V_1 \cup (U \setminus U_0)$, we build a matching M_1 greedily, adding edges one by one. According to Part 1 of the lemma, this process can not stop until $|V_1| - k$ vertices of V_1 will be matched. Denote by $V_0 \subset V_1$ the vertices of V_1, not covered by M_1, then $|V_0| \leq k$. By Part 2 of the lemma, the subhypergraph of H spanned by $V_0 \cup U_0$ satisfies the requirements of Haxell's Theorem 2, therefore there exists a matching M_0 of size $|M_0| = |V_0|$ inside $V_0 \cup U_0$. The union $M = M_0 \cup M_1$ forms the desired matching for V_1.
$\Box$

Acknowledgement. The author would like to thank the anonymous referee for many helpful comments.

References

[1] R. Aharoni, O. Kessler, *On a possible extension of Hall's theorem to bipartite hypergraphs*, Discrete Math. 84 (1990), 309–313.

[2] B. Bollobás, **Random graphs**, Academic Press, New York, 1985.

[3] A. Frieze and S. Janson, *Perfect matchings in random s-uniform hypergraphs*, Random Struct. Alg. 7 (1995), 41–57.

[4] P. E. Haxell, *A condition for matchability in hypergraphs*, Graphs and Comb. 11 (1995), 245–248.

[5] M. Krivelevich, *Perfect fractional matchings in random hypergraphs*, Random Struct. Alg., in press.

[6] J. Schmidt and E. Shamir, *A threshold for perfect matchings in random d-pure hypergraphs*, Discrete Math. 45 (1983), 287–295.

[7] W. F. de la Vega, *Sur la cardinalite maximum des couplages d'hypergraphes aleatoires uniformes*, Discrete Math. 40 (1982), 315–318.