ELSEVIER

Discrete Mathematics 207 (1999) 1–8

DISCRETE MATHEMATICS

On the number of nearly perfect matchings in almost regular uniform hypergraphs

A.S. Asratian[a,*], N.N. Kuzjurin[b]

[a] *Department of Mathematics, Lulea University, S-971 87 Lulea, Sweden*

[b] *Institute for System Programming, Russian Academy of Sciences, B. Kommunisticheskaya 25, 109004 Moscow, Russia*

Received 5 February 1999; accepted 15 February 1999

Abstract

Strengthening the result of Rődl and Frankl (Europ. J. Combin 6 (1985) 317–326), Pippenger proved the theorem stating the existence of a nearly perfect matching in almost regular uniform hypergraph satisfying some conditions (see J. Combin. Theory A 51 (1989) 24–42). Grable announced in J. Combin. Designs 4 (4) (1996) 255–273 that such hypergraphs have exponentially many nearly perfect matchings. This generalizes the result and the proof in Combinatorica 11 (3) (1991) 207–218 which is based on the Rődl nibble algorithm (European J. Combin. 5 (1985) 69–78). In this paper, we present a simple proof of Grable's extension of Pippenger's theorem. Our proof is based on a comparison of upper and lower bounds of the probability for a random subgraph to have a nearly perfect matching. We use the Lovasz Local Lemma to obtain the desired lower bound of this probability.

Keywords: Hypergraph; Nearly perfect matching; Enumeration; Local lemma

1. Introduction

An important problem in integer linear programming is to find tight bounds on the ratio of an integral optimum and an optimum of its linear relaxation for different classes of integer linear programs. For packing and covering integer programs such bounds were obtained in [1,9,4]. It is interesting to find subclasses of packing and covering integer programs with better approximation ratio of integral and rational optima.

One interesting example of such subclasses gives the theorem of Pippenger originally formulated in terms of hypergraphs and presented in [10] as a part of joint work. This theorem can be formulated as a statement which guarantees that for integer packing

* Corresponding author.

E-mail address: asratian@sm.luth.se (A.S. Asratian)

0012-365X/99/$ - see front matter
PII: S0012-365X(99)00127-2

(covering) linear programs with special balancing conditions on (0,1)-matrix of constraints the integer optimum is close to the optimum of its linear relaxation (see [7]). Moreover, it was proved in [12,13] that under these conditions the so-called random greedy algorithm always finds a near-optimal integer solution.

Now we give some definitions and notation. A *hypergraph* H is a pair (V,E), where V is a finite set of *vertices* and E is a finite family of subsets of V, called *edges*. A hypergraph is *r-uniform* if every edge contains precisely r vertices. The number of edges of a hypergraph H containing a vertex v is called the *degree* of v and denoted by $d_H(v)$ or simply $d(v)$. A hypergraph H is called d-regular if $d_H(v)=d$ for each vertex v of H. For two distinct vertices u and v of a hypergraph H, the number of edges containing both u and v is denoted by $d_H(u,v)$ or simply $d(u,v)$. A *matching* in a hypergraph is a collection of pairwise disjoint edges.

For two sequences f_n and g_n we write $g_n = \mathrm{o}(f_n)$ if $g_n/f_n \to 0$ as $n \to \infty$ and $f_n \sim g_n$ if $|f_n - g_n| = \mathrm{o}(g_n)$.

Let H be an r-uniform hypergraph on n vertices and M be a matching of H. M is called a *perfect matching* if it contains precisely n/r edges and a *nearly perfect matching* if it contains at least $(n-\mathrm{o}(n))/r$ edges.

We consider the following version of the theorem of Pippenger:

Theorem 1 (Pippenger and Spencer [10]). *Let r be fixed and H_n be an r-uniform hypergraph on n vertices satisfying the following conditions: for some sequence $d=d_n$ $(d_n \to \infty)$*

$$\text{(1)} \quad d(v) \sim d \quad \text{as } n \to \infty \tag{1}$$

for each vertex v of H_n,

$$\text{(2)} \quad d(u,v) = \mathrm{o}(d) \tag{2}$$

for every two distinct vertices u,v of H_n.

Then H_n has a nearly perfect matching.

Let $N(H_n)$ denote the number of matchings of H_n. Our contribution is a simple proof of the following theorem.

Theorem 2. *Let r be fixed and H_n be an r-uniform hypergraph on n vertices satisfying the conditions of Theorem 1.*

Then

$$N(H_n) \geqslant \exp\{((n-\mathrm{o}(n))/r)\ln d\}.$$

Moreover, for any constant δ, $0<\delta<1/2$, and sufficiently large n the number of matchings of H_n each containing at least $(1-\delta)n/r$ edges is at least

$$\exp\{(1-2\delta)(n/r)\ln d\}.$$

This means that H_n has at least $\exp\{((n-\mathrm{o}(n))/r)\ln d\}$ nearly perfect matchings.

Corollary 1. $\ln N(H_n) \sim (n/r)\ln d$ *as* $n \to \infty$.

This result was announced by Grable in [6]. His proof is a direct generalization of the proof in [4] and is based on the Rődl nibble algorithm [11] (see also [2,3,5,13]). This is a very powerful but not easy technique which has proven to be successful in solving a few well-known problems in combinatorics.

In this paper we present simple and quite different argument sufficient to prove Theorem 2. We use the Lovasz Local Lemma (see [2]) almost as it was done in [8]. The difference with [8] (and the most other known applications of the Local Lemma) is that we use not only the assertion that the probability of a 'good event' is positive but also the lower bound of this probability. Thus, we may consider our proof as another application of the Local Lemma in its full generality.

2. Proof of Theorem 2

Our main tools will be the following two lemmas.

Lemma 1 (Lovasz Local Lemma, Alon and Spencer [2]). *Let* $A_1,\dots,A_m$ *be events in an arbitrary probability space. A directed graph* $D=(V,E)$ *on the set of vertices* $V=\{1,2,\dots,m\}$ *is called a dependency digraph for the events* $A_1,\dots,A_m$ *if for each* i, $1\leqslant i\leqslant m$, *the event* A_i *is mutually independent of all the events* $\{A_j\colon (i,j)\notin E\}$. *Suppose that* $D=(V,E)$ *is a dependency digraph for the above events and suppose there are real numbers* $x_1,\dots,x_m$ *such that* $0\leqslant x_i<1$ *and*

$$P(A_i)\leqslant x_i \prod_{(i,j)\in E}(1-x_j)$$

for all $1\leqslant i\leqslant m$. *Then*

$$P\left(\bigwedge_{i=1}^{m}\bar{A}_i\right)\geqslant\prod_{i=1}^{m}(1-x_i).$$

Lemma 2 (Alon and Spencer [2], Srinivasan [14]). *Let* $z_1,\dots,z_t$ *be independent random variables such that* z_i *takes two values* 0 *and* 1, *and*

$$P\{z_i=1\}=p, \qquad P\{z_i=0\}=1-p.$$

Then for $Z=\sum_{i=1}^{t} z_i$ *and* $EZ=pt$ *the following inequalities hold*:

$$P\{|Z-EZ|>\gamma EZ\}\leqslant 2\exp\{-(\gamma^2/3)EZ\} \tag{3}$$

if $0\leqslant\gamma\leqslant 1$,

$$P\{Z-EZ>\gamma EZ\}\leqslant\exp\{-(((1+\gamma)\ln(1+\gamma))/4)EZ\} \tag{4}$$

if $\gamma\geqslant 1$.

Proof of Theorem 2. Let δ be a constant, $0<\delta<\frac{1}{2}$, and $N=N_\delta(H_n)$ be the number of matchings containing at least $T=\lceil(1-\delta)n/r\rceil$ edges in the hypergraph H_n satisfying the conditions of Theorem 2. Assume that $E=\{E_1,\dots,E_q\}$ is the set of edges of H_n. We define the random subgraph $H_n(p)$ of H_n obtained by choosing independently each edge E_i with probability $p=d^{-1+\delta}$.

Let X be the random variable equal to the number of matchings each containing at least $T=\lceil(1-\delta)(n/r)\rceil$ edges in $H_n(p)$. The idea of the proof is simple. First we show that

$$P(X\geqslant 1)\leqslant EX\leqslant\exp\{\ln N-(1-\delta)^2(n/r)\ln d\}.$$

Then, using Lemma 1, we find a lower bound of the probability $P(X\geqslant 1)$. Combining these two inequalities we obtain a desired lower bound

$$\ln N\geqslant(1-2\delta)(n/r)\ln d$$

which holds for any $\delta>0$ and sufficiently large n. Note, that we will omit, for simplicity, the expression 'for sufficiently large n' in some other inequalities.

(1) It is clear, that $EX\leqslant Np^{\mathrm{T}}$ because each matching in $H_n(p)$ is contained in H_n. We have

$$\begin{aligned}EX &\leqslant Np^{\mathrm{T}}=\exp\{\ln N-T\ln p^{-1}\}\\ &\leqslant \exp\{\ln N-(1-\delta)^2(n/r)\ln d\}.\end{aligned}$$

Using Chebyshev's inequality we conclude that

$$P\{X\geqslant 1\}\leqslant EX\leqslant\exp\{\ln N-(1-\delta)^2(n/r)\ln d\}. \tag{5}$$

(2) Now, we will find a lower bound of the probability that the random subgraph $H_n(p)$ satisfies all the conditions of Theorem 1.

For each vertex v of H_n, let $D(v)$ be the random variable equal to the degree of v in $H_n(p)$. Furthermore, let $y_1,\dots,y_{d(v)}$ be independent random variables such that $y_j=1$ iff the jth edge containing v is in $H_n(p)$, and $y_j=0$, otherwise. It is clear that

$$D(v)=\sum_{j=1}^{d(v)}y_j.$$

Therefore,

$$ED(v)=\sum_{j=1}^{d(v)}Ey_j=(d+\mathrm{o}(d))p=d^\delta+\mathrm{o}(d^\delta).$$

Furthermore, by (3) with $Z=D(v)$ and $\gamma=\gamma_n=(\ln d)^{-1}$, we have

$$\begin{aligned}P\{|D(v)-ED(v)|>\gamma_n ED(v)\} &\leqslant 2\exp\left\{-(d+\mathrm{o}(d))\frac{\gamma_n^2}{3}p\right\}\\ &= 2\exp\left\{-(d^\delta+\mathrm{o}(d^\delta))\frac{\gamma_n^2}{3}\right\}\leqslant d^{-3}.\end{aligned} \tag{6}$$

Now for each pair of distinct vertices u and v with $d_{H_n}(u,v)\geqslant 1$ we define the random variable $D(u,v)$ which is equal to the number of edges in $H_n(p)$ containing both u

and v. Furthermore, we define independent random variables $x_1,\ldots,x_{d(u,v)}$ such that $x_j=1$ iff the jth edge containing both u and v is in $H_n(p)$, and $x_j=0$ otherwise. It is clear that

$$D(u,v)=\sum_{j=1}^{d(u,v)} x_j.$$

Therefore,

$$ED(u,v)=\sum_{j=1}^{d(u,v)} Ex_j \leqslant p \max_{a,b:a\neq b} d(a,b)=\mathrm{o}(d^{\delta}).$$

Let $d^{\delta}=g_n ED(u,v)$. By the previous inequality we have $g_n\to\infty$ as $n\to\infty$.

Hence, by (4) with $Z=D(u,v)$ and $\gamma=\gamma_n'=\max\{(\ln d)^{-1},(\ln g_n)^{-1}\}$, we have

$$\begin{aligned} &P\{D(u,v)-ED(u,v)>\gamma_n' d^{\delta}\}\\ &\quad =P\{D(u,v)-ED(u,v)>\gamma_n' g_n ED(u,v)\}\\ &\quad \leqslant \exp\{-((1+\gamma_n' g_n)(\ln(1+\gamma_n' g_n))/4)ED(u,v)\}\\ &\quad \leqslant \exp\{-\gamma_n' d^{\delta}(\ln(1+\gamma_n' g_n))/4\}\leqslant d^{-3}. \end{aligned} \tag{7}$$

For each vertex v of H_n let $A(v)$ denote the event: $|D(v)-ED(v)|>\gamma_n ED(v)$. Furthermore, for each pair of vertices u,v in H_n with $d_{H_n}(u,v)\geqslant 1$ let $B(u,v)$ denote the event: $D(u,v)-ED(u,v)>\gamma_n' d^{\delta}$.

We will use Lemma 1 to prove that with positive probability none of the events $A(v)$ and $B(u,v)$ occurs. Our analysis is almost the same as in [8] (see Lemma 4 of [8]). Consider the dependency digraph for the events $A(v)$ and $B(u,v)$ for all vertices u and v of H_n such that $d_{H_n}(u,v)\geqslant 1$. Clearly, the event $A(v)$ does not depend on all events $A(x)$ such that $x\notin\bigcup_{v\in E_i} E_i$. Analogously, the event $A(v)$ does not depend on all events $B(u,w)$ such that $\{u,w\}$ is not contained in an edge containing also v. It follows that $A(v)$ is independent of all but

$$(r-1)d(v)+d(v)\binom{r}{2}<(rd)^2$$

other events $A(x)$ and $B(u,w)$. Similarly, $B(v,w)$ is independent of all but

$$2+(r-2)d(v,w)+d(v,w)\binom{r}{2}<(rd)^2$$

other events $A(x)$ and $B(x,y)$. Hence, the degree of each vertex in the dependency digraph is at most $(rd)^2$.

Setting $x_i=(rd)^{-2}$ we can conclude that for each i the inequality

$$x_i\prod_{(i,j)\in E}(1-x_j)\geqslant (rd)^{-2}\left(1-\frac{1}{(rd)^2}\right)^{(rd)^2}\geqslant \frac{1}{6}(rd)^{-2}$$

holds. By (6) and (7) we have that $P\{A(v)\}\leqslant d^{-3}$ for all v and $P\{B(u,v)\}\leqslant d^{-3}$ for all u and v. Therefore, for each pair of vertices u and v of H_n, $P\{A(v)\}$ and $P\{B(u,v)\}$ are not less than $x_i\prod_{(i,j)\in E}(1-x_j)$. It means that all the conditions of Lemma 1 hold.

Clearly, the number m of vertices of the dependency digraph, that is, the number of different events $A(v)$ and $B(u,v)$ is at most $n+nr(d+\mathrm{o}(d))$. It is well-known that $(1-1/(k+1))^k \geqslant 1/e$ for each natural $k \geqslant 2$. Using this inequality with $k=(dr)^2-1$ we have, by Lemma 1,

$$P\left(\bigwedge_v \bar{A}(v) \wedge \bigwedge_{u \neq v, d(u,v) \geqslant 1} \bar{B}(u,v)\right) \geqslant \left(1-\frac{1}{(rd)^2}\right)^{n(r(d+\mathrm{o}(d))+1)}$$
$$\geqslant \exp\left\{-\frac{n(rd+1+\mathrm{o}(rd))}{(dr)^2-1}\right\}$$
$$\geqslant \exp\left\{-\frac{(dr+1)(n+\mathrm{o}(n))}{(dr)^2-1}\right\}$$
$$\geqslant \exp\left\{-\frac{n+\mathrm{o}(n)}{dr-1}\right\}.$$

This means that all the conditions of Theorem 1 hold for $H_n(p)$ with positive probability which is not less than $\exp\{-(n+\mathrm{o}(n))/(dr-1)\}$. This and Theorem 1 imply that the probability $P(X \geqslant 1)$ that $H_n(p)$ has a matching with at least $(1-\delta)n/r$ edges also satisfies the inequality

$$P(X \geqslant 1) \geqslant \exp\left\{-\frac{n+\mathrm{o}(n)}{dr-1}\right\}.$$

Taking into account (5) we have

$$\ln N - (1-\delta)^2 \frac{n}{r} \ln d \geqslant -\frac{n+\mathrm{o}(n)}{dr-1},$$

which implies (in our notations) the inequality

$$\ln N_\delta(H_n) \geqslant (1-2\delta)\frac{n}{r}\ln d.$$

The last inequality holds for any fixed $\delta > 0$ and sufficiently large n and implies the assertion of Theorem 2. The proof of Theorem 2 is complete. □

Corollary 2. $\ln N(H_n) \sim (n/r)\ln d$ *as* $n \to \infty$.

Proof. Let q denote the number of edges in H_n. We use the following trivial upper bound:

$$N(H_n) \leqslant \sum_{k=1}^{\lfloor n/r \rfloor} \binom{q}{k}.$$

Since H_n is r-uniform and almost d-regular (in view of (1)),

$$q = (d+\mathrm{o}(d))n/r.$$

Since $q \geqslant 2n/r$ for sufficiently large n, $\binom{q}{k} < \binom{q}{k+1}$ for $k < n/r$. Using also the inequality $\binom{q}{k} \leqslant (eq/k)^k$, we obtain

$$N(H_n) \leqslant \frac{n}{r}\binom{q}{\frac{n}{r}} \leqslant \frac{n}{r}\binom{(d+\mathrm{o}(d))n/r}{\frac{n}{r}} \leqslant n((d+\mathrm{o}(d))e)^{n/r}$$
$$= \exp\left\{\frac{n+\mathrm{o}(n)}{r}\ln d\right\}.$$

Combining this with the bound obtained in the proof of Theorem 2 we have

$$(1-2\delta)\frac{n}{r}\ln d \leqslant \ln N_\delta(H_n) \leqslant \ln N(H_n) \leqslant \frac{n+\mathrm{o}(n)}{r}\ln d. \tag{8}$$

The last inequality holds for any fixed $\delta > 0$ and sufficiently large n. Therefore, it implies the desired asymptotics for $\ln N(H_n)$. □

3. Concluding remarks

Sometimes, the theorem of Pippenger is considered in the more general form:

Theorem. *Let r and K be positive constants and H_n be an r-uniform hypergraph on n vertices satisfying the following conditions: for some sequence $d = d_n$ ($d_n \to \infty$ as $n \to \infty$)*

(1) $d(v) \leqslant Kd$

for each vertex v of H_n,

(2) $d(v) \sim d$

for all but at most $\mathrm{o}(n)$ vertices,

(3) $d(u,v) = \mathrm{o}(d)$

for all two distinct vertices u, v of H_n.

Then H_n has a nearly perfect matching.

The assertion of Theorem 2 is true with the conditions of this theorem as well. The proof does not differ significantly from the proof of Theorem 2.

Acknowledgements

We would like to thank Svenska Institutet for the financial support which made our cooperation possible. We also thank Alexander Kostochka for many valuable remarks.

References

[1] R. Aharoni, P. Erdos, N. Linial, Optima of dual integer linear programs, Combinatorica 8 (1) (1988) 476–483.
[2] N. Alon, J.H. Spencer, The Probabilistic Method, Wiley, New York, 1992.
[3] N. Alon, H.-J. Kim, J.H. Spencer, Nearly perfect matchings in regular simple hypergraps, Preprint, 1996.
[4] C.J. Colbourn, D.G. Hoffman, K.T. Phelps, V. Rődl, P.M. Winkler, The number of t-wise balanced designs, Combinatorica 11 (3) (1991) 207–218.
[5] P. Frankl, V. Rődl, Near perfect coverings in graphs and hypergraphs, European J. Combin. 6 (1985) 317–326.
[6] D.A. Grable, K.T. Phelps, Random methods in design theory: a survey, J. Combin. Designs 4 (4) (1996) 255–273.
[7] J. Kahn, A linear programming perspective on the Frankl–Rődl–Pippenger theorem, Random Struct. Algorithms 8 (1996) 149–157.
[8] A.V. Kostochka, V. Rődl, Partial Steiner systems and matchings in hypergraphs, Random Struct. Algorithms, to appear.
[9] L. Lovasz, On the ratio of optimal integral and fractional covers, Discrete Math 13 (1975) 383–390.
[10] N. Pippenger, J. Spencer, Asymptotic behavior of the chromatic index for hypergraphs, J. Combin. Theory A 51 (1989) 24–42.
[11] V. Rődl, On a packing and covering problem, European J. Combin. 5 (1985) 69–78.
[12] V. Rődl, L. Thoma, Asymptotic packing and the random greedy algorithm, Random Struct. Algorithms 8 (1996) 161–177.
[13] J. Spencer, Asymptotic packing via a branching process, Random Struct. Algorithms 7 (1995) 167–172.
[14] A. Srinivasan, Improved Approximation Guarantees for Packing and Covering Integer Programs, Proceedings of the ACM Symposium on the Theory of Computing, 1995, pp. 268–276.