

DIFFERENTIAL METHODS FOR FINDING INDEPENDENT SETS IN HYPERGRAPHS*

YUSHENG LI[†] AND WENAN ZANG[‡]

Abstract. It is shown by using differential methods that if \mathcal{H} is a double linear, r -uniform hypergraph with degree sequence $\{d_v\}$ such that any subhypergraph induced by a neighborhood has maximum degree less than m , then its independence number is at least $\sum_v f_{r,m}(d_v)$, where $f_{r,m}(x)$ is a convex function satisfying $f_{r,m}(x) \sim (\log x)/x$ if $r = 2$ and $c/x^{1/(r-1)}$ if $r \geq 3$, as $x \rightarrow \infty$, and $c = c(r, m) > 0$ is a constant. The proof yields a polynomial-time algorithm for finding such an independent set in \mathcal{H} .

Key words. hypergraph, independent set, differential method, convex function, algorithm

AMS subject classifications. 05C65, 05C69, 05C85

DOI. 10.1137/S0895480104442571

1. Introduction. Hypergraphs are systems of sets which are conceived as natural extensions of graphs: elements correspond to vertices and sets correspond to edges which are allowed to connect more than two vertices. Hypergraph theory is a part of the general study of combinatorial properties of families of sets; for in-depth accounts of the subject, see Berge [5] and Duchet [8]. The present paper concerns itself with the independent set problem on hypergraphs.

A *hypergraph* $\mathcal{H} = (V, \mathcal{E})$ consists of a vertex set V and an edge set \mathcal{E} such that each edge is a nonempty subset of V . Throughout this paper we assume that each edge contains at least two vertices. For each vertex v , the *degree* of v , denoted by d_v , is the number of edges containing v , and the *neighborhood* of v , denoted by $N(v)$, is the set of all neighbors of v , where a vertex u is a *neighbor* of (or is *adjacent* to) v if $u \neq v$ and there is an edge that contains both u and v . Let U be a subset of V . Set $\mathcal{E}_U = \{E \in \mathcal{E} : E \subseteq U\}$. The hypergraph (U, \mathcal{E}_U) is called the *subhypergraph of \mathcal{H} induced by U* . We say that U is an *independent set* of \mathcal{H} if it contains no edge. The *independence number* of \mathcal{H} , denoted by $\alpha(\mathcal{H})$, is the maximum number of vertices in an independent set of \mathcal{H} . The *independent set problem* is to find an independent set with the largest size. As is well known, this *NP-hard* problem arises in a rich variety of applications, so it has attracted tremendous research efforts.

Let $G = (V, E)$ be a graph on N vertices with average degree d . A classical theorem of Turán asserts that $\alpha(G) \geq \frac{N}{d+1}$, which was strengthened independently by Caro [6] and Wei [17] as $\alpha(G) \geq \sum_{v \in V} \frac{1}{d_v+1}$ (this bound is better than the former since function $1/(1+x)$ is strictly convex); a nice probabilistic proof of this theorem can be found in Alon and Spencer [4]. In case G is triangle-free, Turán's lower bound can be improved substantially. As shown by Ajtai et al. [1, 2] and Ajtai, Komlós, and Szemerédi [3], $\alpha(G) \geq \frac{cN \log d}{d}$, where (and throughout this paper) $\log x$ stands for the

*Received by the editors March 27, 2004; accepted for publication (in revised form) November 11, 2005; published electronically February 21, 2006.

<http://www.siam.org/journals/sidma/20-1/44257.html>

[†]Department of Mathematics, Tongji University, Shanghai 200092, China (li_yusheng@mail.tongji.edu.cn). The work of this author was supported in part by the National Natural Science Foundation of China.

[‡]Department of Mathematics, University of Hong Kong, Hong Kong, China (wzang@maths.hku.hk). The work of this author was supported in part by the Research Grants Council of Hong Kong.

natural logarithmic function, and constant c can be set equal to $1/2.4$ (cf. Griggs [9]). Shearer [15] confirmed a conjecture of Ajtai, Komlós, and Szemerédi [3] and managed to improve c to $1 - o(1)$ by establishing that $\alpha(G) \geq Ng(d)$, where

$$(1) \quad g(x) = \frac{x \log x - x + 1}{(x - 1)^2};$$

he [16] further improved the bound as $\alpha(G) \geq \sum_v \bar{g}(d_v)$, where the function $\bar{g}(x)$ is asymptotically equal to $g(x)$ as $x \rightarrow \infty$. In his proofs, Shearer first introduced the appealing differential methods, which are proved to be very powerful in applications. Shearer's results can be extended [11, 12, 13] as follows: if in a graph G with N vertices and average degree d , any subgraph induced by a neighborhood has no vertex of degree at least m , then $\alpha(G) \geq \sum_v g_m(d_v) \geq Ng_m(d)$, where

$$(2) \quad g_m(x) = \int_0^1 \frac{(1-t)^{1/m}}{m + (x-m)t} dt.$$

(Notice that $g_1(x)$ is exactly Shearer's function $g(x)$ as specified in (1).) This result has interesting applications in Ramsey theory [12, 14]; for instance, it yields $R(m, n) \leq (1 + o(1))n^{m-1}/(\log n)^{m-2}$, where Ramsey number $R(m, n)$ is the smallest integer N such that for any graph G of order N , either $\alpha(G) \geq m$ or $\alpha(\bar{G}) \geq n$ holds. It is worthwhile pointing out that since the order of magnitude of $R(3, n)$ is $n^2/\log n$ (see Kim [10]), the above-mentioned lower bound due to Ajtai, Komlós, and Szemerédi [3] cannot be improved more than a constant factor; we believe Shearer's bound is asymptotically sharp on extremal graphs for $R(3, n)$.

The independent set problem on hypergraphs is much more difficult and intractable than that on graphs. So it is natural to restrict our attentions to some special classes of hypergraphs. A hypergraph \mathcal{H} is called *r-uniform* if each edge of \mathcal{H} contains exactly r vertices (so a 2-uniform hypergraph is a graph), and called *triangle-free* if \mathcal{H} contains no three distinct vertices v_1, v_2, v_3 and three distinct edges E_1, E_2, E_3 such that $\{v_1, v_2, v_3\} - \{v_i\}$ is a subset of E_i for $i = 1, 2, 3$. We say that a hypergraph \mathcal{H} is *linear* if any two edges of \mathcal{H} have at most one vertex in common. A linear hypergraph \mathcal{H} is said to be *double linear* if for any two nonadjacent vertices u and v , each edge containing u contains at most one neighbor of v . Caro and Tuza [7] proposed a problem on extending the lower bound of Ajtai, Komlós, and Szemerédi [3] to triangle-free hypergraphs; as a solution to this problem, Zhou and Li [18] proved that every r -uniform linear triangle-free hypergraph \mathcal{H} satisfies $\alpha(\mathcal{H}) \geq Nf_{r,1}(d)$, where function $f_{r,1}(x)$ is much bigger than $(\log x)/x$ when $r \geq 3$. Observe that if a linear hypergraph \mathcal{H} is triangle-free, then its subhypergraph induced by any neighborhood has maximum degree zero. However, the converse need not hold in general. In this paper we consider hypergraphs whose subhypergraphs induced by neighborhoods may have edges.

Let us define some functions before presenting our main result. As usual, let $B(p, q) = \int_0^1 (1-t)^{p-1} t^{q-1} dt$ denote the beta function with $p, q > 0$. For integers $r \geq 2$ and $m \geq 1$, set constants

$$a = \frac{1}{(r-1)^2}, \quad b = \frac{r-2}{r-1},$$

and

$$B = B(a/m, 1-b) = \int_0^1 (1-t)^{a/m-1} t^{-b} dt.$$

Clearly, $0 < a \leq 1$, $0 \leq b < 1$, and $B > 0$. For the above r , m and $x \geq 0$, define

$$f_{r,m}(x) = \frac{m}{B} \int_0^1 \frac{(1-t)^{a/m}}{t^b[m+(x-m)t]} dt.$$

Since

$$\frac{(1-t)^{a/m}}{t^b[m+(x-m)t]} \leq \frac{(1-t)^{a/m}}{t^b[m(1-t)]} = \frac{1}{m}(1-t)^{a/m-1}t^{-b},$$

we see that $f_{r,m}(x)$ is bounded above by 1 and thus is well defined.

THEOREM. *Let $\mathcal{H} = (V, \mathcal{E})$ be an r -uniform, double linear hypergraph with degree sequence $\{d_v\}$. If the maximum degree of any subhypergraph induced by a neighborhood is less than m , then*

$$\alpha(\mathcal{H}) \geq \sum_{v \in V} f_{r,m}(d_v).$$

Note that if $r = 2$, then $a = 1$, $b = 0$, and $B = m$, so $f_{2,m}$ is the function $g_m(x)$ defined in (2). And $f_{r,1}(x)$ is precisely the function involved in the above Zhou–Li bound. Since any graph and any linear triangle-free hypergraph are double linear, our theorem generalizes all the results cited above, including Turán’s theorem and the Caro–Wei theorem as long as graphs in consideration satisfy the conditions.

For any fixed integers $r \geq 3$ and $M \geq 1$, it was shown in [18] that $f_{r,1}(x) \geq (\log^M x)/x$ provided x is large enough. We shall verify that $f_{r,m}(x)$ is a convex function for $x \geq 0$ and that $f_{r,m}(x) \sim (\log x)/x$ if $r = 2$ and $c/x^{1/(r-1)}$ if $r \geq 3$, as $x \rightarrow \infty$, where $c = c(r, m) > 0$ is a constant and \sim means an asymptotic equality. By convexity of $f_{r,m}(x)$, we have $f_{r,m}(d) \leq \frac{1}{|V|} \sum_{v \in V} f_{r,m}(d_v)$, where $d = \frac{1}{|V|} \sum_{v \in V} d_v$. Thus the following is an immediate consequence of the above theorem.

COROLLARY. *For fixed integers $r \geq 3$ and $m \geq 1$, let $c = c(r, m) > 0$ be the constant as described above. Then for any $\epsilon > 0$, there exists a constant $D = D(r, m, \epsilon)$ such that if a hypergraph $\mathcal{H} = (V, \mathcal{E})$ is double linear, r -uniform, and the subhypergraph induced by any neighborhood has maximum degree less than m , then*

$$\alpha(\mathcal{H}) \geq (1 - \epsilon) \frac{cN}{d^{1/(r-1)}},$$

where $N = |V|$ and d is the average degree of \mathcal{H} with $d \geq D$.

2. Properties of the function $f_{r,m}$. The purpose of this section is to exhibit some properties satisfied by the function $f_{r,m}$ defined in the preceding section.

LEMMA 1. *For fixed integers $r \geq 2$ and $m \geq 1$ and for $x \geq 0$, the function $f(x) = f_{r,m}(x)$ satisfies the differential equation*

$$(3) \quad (r-1)^2 x(x-m)f'(x) + [(r-1)x+1]f(x) = 1.$$

Moreover, $f(x)$ is strictly and completely monotonic, that is, $(-1)^k f^{(k)}(x) > 0$ for all $x \geq 0$. In particular, $f(x)$ is positive, strictly decreasing, and strictly convex.

Proof. By differentiating x under the integral and then integrating by parts,

we have

$$\begin{aligned}
 & x(x-m)f'(x) \\
 &= \frac{-mx(x-m)}{B} \int_0^1 \frac{(1-t)^{a/m} t^{1-b}}{[m+(x-m)t]^2} dt \\
 &= \frac{mx}{B} \int_0^1 (1-t)^{a/m} t^{1-b} \frac{d}{dt} \left(\frac{1}{m+(x-m)t} \right) dt \\
 &= \frac{-mx}{B} \int_0^1 \frac{1}{m+(x-m)t} \left[(1-b)(1-t)^{a/m} t^{-b} - \frac{a}{m} (1-t)^{a/m-1} t^{1-b} \right] dt \\
 &= -(1-b)xf(x) + \frac{ax}{B} \int_0^1 \frac{(1-t)^{a/m-1} t^{1-b}}{m+(x-m)t} dt \\
 &= -(1-b)xf(x) + \frac{am}{B} \int_0^1 \left(\frac{1}{m(1-t)} - \frac{1}{m+(x-m)t} \right) (1-t)^{a/m} t^{-b} dt \\
 &= -(1-b)xf(x) + \frac{a}{B} \int_0^1 (1-t)^{a/m-1} t^{-b} dt - af(x) \\
 &= -(1-b)xf(x) + a - af(x) \\
 &= - \left(\frac{x}{r-1} + \frac{1}{(r-1)^2} \right) f(x) + \frac{1}{(r-1)^2},
 \end{aligned}$$

so the desired differential equation follows. The strict and complete monotonicity of $f(x)$ can be seen by repeatedly differentiating x under the integral. \square

Let us now proceed to the asymptotic behavior of the function $f_{2,m}(x)$.

LEMMA 2. *For any fixed integer $m \geq 1$ and for $x > 1$, we have*

$$\frac{\log(x/m) - 1}{x} \leq f_{2,m}(x) \leq \frac{x \log x - x + 1}{(x-1)^2}.$$

Therefore $f_{2,m}(x) \sim (\log x)/x$ as $x \rightarrow \infty$.

Proof. We first claim that for fixed $x \geq 1$, function

$$f_{2,m}(x) = \int_0^1 \frac{(1-t)^{1/m} dt}{m+(x-m)t} = \int_0^1 \frac{t^{1/m} dt}{mt+x(1-t)}$$

decreases as $m \geq 1$ increases. To justify the claim, setting $t = u^m$ gives

$$f_{2,m}(x) = \int_0^1 \frac{mu^m du}{mu^m + x(1-u^m)}.$$

So it suffices to show that if $\delta > 0$ and $0 < u < 1$, then

$$\frac{mu^m}{mu^m + x(1-u^m)} > \frac{(m+\delta)u^{m+\delta}}{(m+\delta)u^{m+\delta} + x(1-u^{m+\delta})}.$$

Equivalently,

$$\delta u^{m+\delta} + m - (m+\delta)u^\delta > 0.$$

For this purpose, set $h(u) = \delta u^{m+\delta} + m - (m+\delta)u^\delta$. Then $h(1) = 0$ and $h'(u) = \delta(m+\delta)u^{\delta-1}(u^m - 1) < 0$ for $0 < u < 1$, and thus the claim follows.

Since for $x > 1$, we have

$$f_{2,1}(x) = \int_0^1 \frac{(1-t)dt}{1+(x-1)t} = \frac{x \log x - x + 1}{(x-1)^2},$$

by the above claim $f_{2,m}(x) \leq f_{2,1}(x)$, and so the upper bound is established.

To derive the lower bound, note that

$$\begin{aligned} f_{2,m}(x) &= \int_0^1 \frac{(1-t)^{1/m} dt}{m+(x-m)t} > \int_0^1 \frac{(1-t)dt}{m+(x-m)t} \\ &= \frac{x \log(x/m) - x + m}{(x-m)^2} \geq \frac{\log(x/m) - 1}{x}, \end{aligned}$$

where the last inequality amounts to $(2x-m)\log(x/m) \geq x-m$, or equivalently $(2t-1)\log t \geq t-1$. Set $\phi(t) = (2t-1)\log t - t + 1$. Then $\phi(1) = 0$ and $\phi'(t) = 2\log t + (1-1/t)$, which is less than 0 if $0 < t < 1$, equal to 0 if $t = 1$, and greater than 0 if $t > 1$. Hence $\phi(t) \geq 0$ for $t > 0$, implying the lower bound. \square

Our next lemma concerns the case when $r \geq 3$; it shows that the asymptotic behavior of $f_{r,m}$ is dramatically different from that of $f_{2,m}$.

LEMMA 3. For fixed integers $r \geq 3$ and $m \geq 1$, function $f_{r,m}(x) \sim \frac{c}{x^{1/(r-1)}}$ as $x \rightarrow \infty$, where $c = c(r, m) > 0$ is defined to be

$$\frac{m}{B(m+1)^{a/m}} \int_0^1 \frac{(1-t)^{a/m}}{t^b(m+t)} dt + a \int_{m+1}^{\infty} \frac{dt}{t^{1+a/m}(t-m)^{b-a/m}}.$$

Proof. Our proof relies heavily on the theorem that a linear first-order differential equation

$$\frac{dy}{dx} = p(x)y + q(x)$$

has a unique solution

$$y = e^{\phi(x)} \left(y_0 + \int_{x_0}^x q(t)e^{-\phi(t)} dt \right)$$

satisfying $y_0 = y(x_0)$, where $\phi(x) = \int_{x_0}^x p(t)dt$. Now let us transform the differential equation (3) in Lemma 1 into the above standard form. Then we get

$$p(x) = -\frac{a((r-1)x+1)}{x(x-m)} \quad \text{and} \quad q(x) = \frac{a}{x(x-m)}.$$

Set $x_0 = m+1$ and

$$y_0 = f_{r,m}(m+1) = \frac{m}{B} \int_0^1 \frac{(1-t)^{a/m}}{t^b(m+t)} dt.$$

It follows from the uniqueness of the solution that

$$f_{r,m}(x) = e^{\phi(x)} \left(y_0 + \int_{m+1}^x q(t)e^{-\phi(t)} dt \right) \quad \text{for } x \geq m+1.$$

Since

$$\begin{aligned}\phi(x) &= - \int_{m+1}^x \frac{a((r-1)t+1)}{t(t-m)} dt \\ &= -a \log \left(\left(\frac{m+1}{x} \right)^{1/m} (x-m)^{r-1+1/m} \right),\end{aligned}$$

we obtain

$$(4) \quad e^{\phi(x)} = \frac{x^{a/m}}{(m+1)^{a/m} (x-m)^{1/(r-1)+a/m}}$$

$$(5) \quad \sim \frac{1}{(m+1)^{a/m} x^{1/(r-1)}},$$

and hence

$$e^{-\phi(x)} \sim (m+1)^{a/m} x^{1/(r-1)}.$$

Thus there exists a constant $M > 0$ such that for all $t \geq m+1$,

$$0 \leq q(t)e^{-\phi(t)} \leq \frac{Mt^{1/(r-1)}}{t^2} = \frac{M}{t^{1+b}}.$$

Recall that $b > 0$ as $r \geq 3$, so $\int_{m+1}^{\infty} q(t)e^{-\phi(t)} dt < \infty$ and

$$\int_{m+1}^x q(t)e^{-\phi(t)} dt = \int_{m+1}^{\infty} q(t)e^{-\phi(t)} dt - o(1)$$

as $x \rightarrow \infty$. It follows from (5) that

$$\begin{aligned}f_{r,m}(x) &= e^{\phi(x)} \left(y_0 + \int_{m+1}^{\infty} q(t)e^{-\phi(t)} dt - o(1) \right) \\ &\sim \frac{c}{x^{1/(r-1)}},\end{aligned}$$

where $c = \frac{1}{(m+1)^{a/m}} (y_0 + \int_{m+1}^{\infty} q(t)e^{-\phi(t)} dt)$. Using (4) and plugging y_0 , we see that c is as defined in the lemma. \square

3. Proof of the theorem. Let us introduce some notions before presenting the proof. For each $v \in V$, let \mathcal{H}_v be the subhypergraph of \mathcal{H} induced by $V - (N(v) \cup \{v\})$, and let $\{d'_u\}$ denote the degree sequence of \mathcal{H}_v . For simplicity, write $f_{r,m}(x)$ as $f(x)$. Set $S(\mathcal{H}) = \sum_{u \in V(\mathcal{H})} f(d_u)$ and $S(\mathcal{H}_v) = \sum_{u \in V(\mathcal{H}_v)} f(d'_u)$. The default value of $S(\mathcal{H}_v)$ is zero if $V - (N(v) \cup \{v\}) = \emptyset$.

The key step of our proof is to establish the following statement.

LEMMA 4. *There exists a vertex v in \mathcal{H} such that $1 + S(\mathcal{H}_v) \geq S(\mathcal{H})$.*

To show that \mathcal{H} contains an independent set I with size at least $\sum_{v \in V} f(d_v)$, we may apply the following algorithm: Initially set $I = \emptyset$. Let v be the vertex exhibited in Lemma 4. Set $I = I \cup \{v\}$ and $\mathcal{H} = \mathcal{H}_v$. Repeat the process until \mathcal{H} contains no vertex.

So Lemma 4 serves as a criterion for selecting vertices in I . Let us now prove that such an independent set I is indeed as desired.

Proof of the Theorem (assuming Lemma 4). We apply induction on $|V|$, the number of vertices in \mathcal{H} . Since $f(0) = 1$ by (3), the assertion holds trivially for $|V| = 1$. So we proceed to the induction step.

Note that $\alpha(\mathcal{H}) \geq 1 + \alpha(\mathcal{H}_u)$ for any vertex u of \mathcal{H} . Let v be a vertex as described in Lemma 4. Then, by induction hypothesis, we have $\alpha(\mathcal{H}) \geq 1 + \alpha(\mathcal{H}_v) \geq 1 + S(\mathcal{H}_v) \geq S(\mathcal{H})$, completing the proof.

It therefore remains to prove the above lemma.

Proof of Lemma 4. For each $v \in V$, set

$$N_2(v) = \{x \in V - (N(v) \cup \{v\}) : N(x) \cap N(v) \neq \emptyset\}$$

and

$$\begin{aligned} Y(v) &= 1 + S(\mathcal{H}_v) - S(\mathcal{H}) \\ &= 1 + \sum_{x \in V(\mathcal{H}_v)} [f(d'_x) - f(d_x)] - f(d_v) - \sum_{u \in N(v)} f(d_u). \end{aligned}$$

Besides, for each $x \in N_2(v)$, set $n_{v,x} = |N(v) \cap N(x)|$. Let us consider the terms in $Y(v)$. Since any vertex $x \in V(\mathcal{H}_v) - N_2(v)$ satisfies $d'_x = d_x$ and any vertex $x \in N_2(v)$ satisfies $d'_x = d_x - n_{v,x}$ (for \mathcal{H} is double linear),

$$Y(v) = 1 - f(d_v) - \sum_{u \in N(v)} f(d_u) + \sum_{x \in N_2(v)} [f(d_x - n_{v,x}) - f(d_x)].$$

Clearly, (6) is equivalent to saying that $Y(v) \geq 0$ for some vertex v of \mathcal{H} . So to prove the lemma it suffices to show that

$$(6) \quad \sum_{v \in V(\mathcal{H})} Y(v) \geq 0.$$

Since \mathcal{H} is linear and r -uniform,

$$\sum_{v \in V(\mathcal{H})} \sum_{u \in N(v)} f(d_u) = (r-1) \sum_{v \in V(\mathcal{H})} d_v f(d_v).$$

So

$$\begin{aligned} & \sum_{v \in V(\mathcal{H})} Y(v) \\ &= \sum_{v \in V(\mathcal{H})} \{1 - [(r-1)d_v + 1]f(d_v)\} + \sum_{v \in V(\mathcal{H})} \sum_{x \in N_2(v)} [f(d_x - n_{v,x}) - f(d_x)]. \end{aligned}$$

Observe that $x \in N_2(v)$ if and only if $v \in N_2(x)$ and that $n_{v,x} = n_{x,v}$; exchanging the variables in the sum gives

$$\sum_{v \in V(\mathcal{H})} \sum_{x \in N_2(v)} [f(d_x - n_{v,x}) - f(d_x)] = \sum_{v \in V(\mathcal{H})} \sum_{x \in N_2(v)} [f(d_v - n_{v,x}) - f(d_v)].$$

Let

$$Z(v) = \sum_{x \in N_2(v)} [f(d_v - n_{v,x}) - f(d_v)].$$

Then

$$(7) \quad \sum_{v \in V(\mathcal{H})} Y(v) = \sum_{v \in V(\mathcal{H})} \{1 - [(r-1)d_v + 1]f(d_v)\} + \sum_{v \in V(\mathcal{H})} Z(v).$$

Now comes the technical part of our proof, the analysis of the term $\sum_{v \in V(\mathcal{H})} Z(v)$.

Since $f(x)$ is convex, we have

$$(8) \quad f(x-1) - f(x) \geq f(y-1) - f(y) \text{ whenever } 1 \leq x \leq y.$$

(To see this, write $x = \alpha(x-1) + (1-\alpha)y$ and $y-1 = \beta(x-1) + (1-\beta)y$, where $0 \leq \alpha, \beta \leq 1$. By convexity, $f(x) \leq \alpha f(x-1) + (1-\alpha)f(y)$ and $f(y-1) \leq \beta f(x-1) + (1-\beta)f(y)$. Summing up these two inequalities yields $f(x) + f(y-1) \leq f(x-1) + f(y)$ as $\alpha + \beta = 1$.) From (8) we deduce that

$$f(d_v - n_{v,x}) - f(d_v) = \sum_{i=1}^{n_{v,x}} [f(d_v - i) - f(d_v - (i-1))] \geq [f(d_v - 1) - f(d_v)]n_{v,x},$$

and so

$$Z(v) \geq [f(d_v - 1) - f(d_v)] \sum_{x \in N_2(v)} n_{v,x}.$$

Note that \mathcal{H} is double linear, r -uniform, and each vertex $u \in N(v)$ is incident to at most $m-1$ edges in $N(v)$. Moreover, there is precisely one edge in \mathcal{H} containing both u and v . So

$$\sum_{x \in N_2(v)} n_{v,x} \geq (r-1) \sum_{u \in N(v)} (d_u - m).$$

Write $A_v = f(d_v - 1) - f(d_v)$. Then

$$\begin{aligned} \sum_{v \in V(\mathcal{H})} Z(v) &\geq (r-1) \sum_{v \in V(\mathcal{H})} \sum_{u \in N(v)} (d_u - m)A_v \\ &= (r-1) \sum_{E \in \mathcal{E}} \sum_{u, v \in E} \{(d_u - m)A_v + (d_v - m)A_u\} \\ &= (r-1) \sum_{E \in \mathcal{E}} \sum_{u, v \in E} \{(d_v - m)A_v + (d_u - m)A_u + (d_u - d_v)(A_v - A_u)\}. \end{aligned}$$

By (8), we get $(d_u - d_v)(A_v - A_u) \geq 0$. Thus

$$\begin{aligned} \sum_{v \in V(\mathcal{H})} Z(v) &\geq (r-1) \sum_{E \in \mathcal{E}} \sum_{u, v \in E} \{(d_v - m)A_v + (d_u - m)A_u\} \\ &= (r-1) \sum_{v \in V(\mathcal{H})} \sum_{u \in N(v)} (d_v - m)A_v \\ &= (r-1)^2 \sum_{v \in V(\mathcal{H})} d_v (d_v - m)A_v. \end{aligned}$$

From the convexity of $f(x)$, it follows that $f(y) \geq f(x) + f'(x)(y-x)$ for any $x, y \geq 0$. So $A_v = f(d_v - 1) - f(d_v) \geq -f'(d_v)$ and hence

$$(9) \quad \sum_{v \in V(\mathcal{H})} Z(v) \geq -(r-1)^2 \sum_{v \in V(\mathcal{H})} d_v (d_v - m)f'(d_v).$$

Finally, combining (7) with (9) and using differential equation (3) in Lemma 1, we obtain

$$\begin{aligned} & \sum_{v \in V(\mathcal{H})} Y(v) \\ \geq & \sum_{v \in V(\mathcal{H})} \{1 - [(r-1)d_v + 1]f(d_v) - (r-1)^2 d_v(d_v - m)f'(d_v)\} \\ = & 0. \end{aligned}$$

This completes the proof of (6) and hence the lemma.

It is easy to see that our proof yields a polynomial-time algorithm for finding an independent set in \mathcal{H} with at least $\sum_{v \in V} f_{r,m}(d_v)$ vertices.

REFERENCES

- [1] M. AJTAI, P. ERDŐS, J. KOMLÓS, AND E. SZEMERÉDI, *On Turán's theorem for sparse graphs*, *Combinatorica*, 1 (1981), pp. 313–317.
- [2] M. AJTAI, P. ERDŐS, J. KOMLÓS, AND E. SZEMERÉDI, *A dense infinite Sidon sequence*, *European J. Combin.*, 2 (1981), pp. 1–11.
- [3] M. AJTAI, J. KOMLÓS, AND E. SZEMERÉDI, *A note on Ramsey numbers*, *J. Combin. Theory Ser. A*, 29 (1980), pp. 354–360.
- [4] N. ALON AND J. SPENCER, *The Probabilistic Method*, Wiley-Interscience, New York, 1992.
- [5] C. BERGE, *Hypergraphs*, North-Holland, Amsterdam, 1989.
- [6] Y. CARO, *New Results on the Independence Number*, Technical report, Tel-Aviv University, Tel-Aviv, Israel, 1979.
- [7] Y. CARO AND Z. TUZA, *Improved lower bounds on k -independence*, *J. Graph Theory*, 15 (1991), pp. 99–107.
- [8] P. DUCHET, *Hypergraphs Handbook of Combinatorics*, R. L. Graham, M. Grötschel, and L. Lovász, eds., Elsevier, Amsterdam, 1995, pp. 381–432.
- [9] J. R. GRIGGS, *An upper bound on the Ramsey number $R(3, k)$* , *J. Combin. Theory Ser. A*, 35 (1983), pp. 145–153.
- [10] J. KIM, *The Ramsey number $R(3, t)$ has order of magnitude $t^2/\log t$* , *Random Structures Algorithms*, 7 (1995), pp. 174–207.
- [11] Y. LI AND C. ROUSSEAU, *On book-complete graph Ramsey numbers*, *J. Combin. Theory Ser. B*, 68 (1996), pp. 36–44.
- [12] Y. LI, C. ROUSSEAU, AND W. ZANG, *Asymptotic upper bounds for Ramsey functions*, *Graphs Combin.*, 17 (2001), pp. 123–128.
- [13] Y. LI, C. ROUSSEAU, AND W. ZANG, *The lower bound on independence number*, *Sci. China Ser. A*, 45 (2002), pp. 64–69.
- [14] Y. LI AND W. ZANG, *Ramsey numbers involving large dense graphs and bipartite Turán numbers*, *J. Combin. Theory Ser. B*, 87 (2003), pp. 280–288.
- [15] J. SHEARER, *A note on the independence number of triangle-free graphs*, *Discrete Math.*, 46 (1983), pp. 83–87.
- [16] J. SHEARER, *A note on the independence number of triangle-free graphs, II*, *J. Combin. Theory Ser. B*, 53 (1991), pp. 300–307.
- [17] A. K. WEI, *A Lower Bound on the Stability Number of a Simple Graph*, Bell Laboratories Technical Memorandum, No. 81-11217-9, Murray Hill, NJ, 1981.
- [18] G. ZHOU AND Y. LI, *Independence numbers of hypergraphs with sparse neighborhoods*, *European J. Combin.*, 25 (2004), pp. 355–362.