SERIE B — INFORMATIK

A Lower Bound on the Independence Number of General Hypergraphs in Terms of the Degree Vectors

Torsten Thiele*

B 95–02
March 1995

Abstract

This paper proves a lower bound on the independence number of general hypergraphs in terms of the degree vectors. The degree vector of a vertex v is given by $d(v) = (d_1(v), d_2(v), \ldots)$ where $d_m(v)$ is the number of edges of size m containing v. We define a function f with the property that any hypergraph $H = (V, E)$ satisfies $\alpha(H) \geq \sum_{v \in V} f(d(v))$. This lower bound is sharp when H is a matching. Furthermore this bound generalizes known bounds of Wei/Caro and Caro/Tuza for ordinary graphs and uniform hypergraphs.

*Freie Universität Berlin, Fachbereich Mathematik und Informatik, Arnimallee 3, 14195 Berlin, Germany

1 Introduction

Wei and Caro independently discovered the following nice lower bound for the independence number of a graph in terms of the degrees (see also [G]).

Theorem 1 **[W, C]** *Let $G = (V, E)$ be a graph with independence number $\alpha(G)$. Then*

$$\alpha(G) \geq \sum_{v \in V} \frac{1}{d(v) + 1} ,$$

where $d(v)$ is the degree of the vertex v.

This bound is tight if (and only if) G is the union of disjoint cliques. This result raises the question if a similar lower bound can be found for the independence number of hypergraphs. Before stating the results we have to make some definitions.

A *hypergraph* is a pair $H = (V, E)$ where V is a finite set and E is a collection of non-empty subsets of V, i.e. $E \subseteq 2^V \setminus \{\emptyset\}$. The *rank* r of a hypergraph $H = (V, E)$ is the maximal size of an edge in E. The hypergraph H is *k-uniform* if all edges in E have size k. A set $I \subseteq V$ is called *independent* if $2^I \cap E = \emptyset$, i.e. the set I contains no edge of E. The maximal size of an independent set of H is defined as the *independence number* $\alpha(H)$.

Caro and Tuza proved the following result, which is an extension of Theorem 1.

Theorem 2 **[CT]** *Let $H = (V, E)$ be a k-uniform hypergraph with $k \geq 2$. Then*

$$\alpha(H) \geq \sum_{v \in V} f(d(v)) ,$$

where $d(v)$ is the degree of v, i.e. the number of edges containing v and the function f is given by

$$f(d) := \prod_{i=1}^{d} \left(1 - \frac{1}{i\,(k-1) + 1}\right) .$$

In fact the result of Caro and Tuza is slightly more general.

Remark. The function f in Theorem 2 can be simplified to $f(d) = \binom{d+1/(k-1)}{d}^{-1}$. Thus we may rewrite the result as

$$\alpha(H) \geq \sum_{v \in V} \binom{d(v) + \frac{1}{k-1}}{d(v)}^{-1} .$$

For $k = 2$ (ordinary graphs) this is the Wei/Caro bound.

In order to generalize this result to arbitrary (non-uniform) hypergraphs we have to generalize the concept of the degree of a vertex. The first idea maybe simply to define the degree of a vertex v similarly as the number of edges containing v. But we will run into troubles with this approach, since we don't have any information about the sizes of the edges containing v. More useful is the following approach.

Let $H = (V, E)$ be a hypergraph of rank r. For every vertex $v \in V$ define the *degree vector* $d(v) = (d_1(v), d_2(v), \ldots, d_r(v)) \in \mathbf{N}_0^r$ where $d_m(v)$ is the number of edges of size m containing v for $1 \leq m \leq r$.

Definition. Let $r \geq 1$ be an integer. Define the function $f_r : \mathbf{N}_0^r \to \mathbf{R}$ by

$$f_r(d) = \sum_{i \in \mathbf{N}_0^r} \left[\prod \binom{d_m}{i_m}\right] \frac{(-1)^{\sum i_m}}{\sum (m-1) \cdot i_m + 1} .$$

The product and the inner sums are taken over all $1 \leq m \leq r$. Note that the outer sum is finite since all summands are zero unless $i \in [0, d] := \{j \in \mathbf{N}_0^r : 0 \leq j_m \leq d_m \text{ for all } 1 \leq m \leq r\}$. Now we are in the position to state our main theorem.

Theorem 3 *Let $H = (V, E)$ be a hypergraph of rank r. Then*

$$\alpha(H) \geq \sum_{v \in V} f_r(d(v)) \ .$$

Suppose $H = (V, E)$ is k-uniform, $k \geq 2$. Let $v \in V$ be arbitrary, e_k denotes the k-th unit vector. Since H is k-uniform $f(d(v))$ reduces to

$$f(d(v)) = f(d_k(v) \cdot e_k) = \sum_{i=0}^{d_k(v)} \binom{d_k(v)}{i} \frac{(-1)^i}{(k-1) \cdot i + 1} = \binom{d_k(v) + \frac{1}{k-1}}{d_k(v)}^{-1}$$

(see Concrete Mathematics [GKP] p. 188). Thus the theorem is a generalization of the results of Wei/Caro and Caro/Tuza. Let us also consider the case $k = 1$, i.e. H is 1-uniform. Then $f(d(v))$ reduces to

$$f(d(v)) = f(d_1(v)) = \sum_{i=0}^{d_1(v)} \binom{d_1(v)}{i} (-1)^i = \begin{cases} 1 \text{ if } d_1(v) = 0 \\ 0 \text{ if } d_1(v) \geq 1 \end{cases} .$$

This is what we expect: The unique maximum independent set is given by the set of vertices of degree 0.

Observation. Let $H = (V, E)$ be a matching of rank r, i.e. H is a hypergraph with the property $e \neq e' \in E \Rightarrow e \cap e' = \emptyset$. Then

$$\alpha(H) = \sum_{v \in V} f_r(d(v)).$$

Proof. Since H is a matching the independence number of H is given by

$$\alpha(H) = \#\text{vertices of degree vector zero} + \sum_{e \in E} (|e| - 1) \ .$$

On the other hand $f_r(0) = 1$ and for every edge $e \in E$ we have $\sum_{v \in e} f_r(d(v)) = |e| (1 - 1/|e|) = |e| - 1$. Thus $\alpha(H) = \sum_{v \in V} f_r(d(v))$ □

Lemma 4 *Let $r \in \mathbf{N}$, $C_1, C_2, \ldots, C_r \geq 0$ and $C_0 > 0$ be given. The function $g : \mathbf{N}_0^r \to \mathbf{R}$ given by*

$$g(d) = \sum_i \left[\prod \binom{d_m}{i_m} \right] \frac{(-1)^{\sum i_m}}{\sum C_m i_m + C_0}$$

is the solution of the recurrence

$$g(d) = \frac{\sum_k C_k d_k g(d - e_k)}{\sum_k C_k d_k + C_0}$$

with $g(0) = C_0^{-1}$. In particular $g(d)$ is non-negative for all $d \in \mathbf{N}_0^r$.

By this lemma we infer that our function f satisfies the recurrence

$$f(d) = \frac{\sum (k-1) \cdot d_k f(d - e_k)}{\sum (k-1) \cdot d_k + 1}$$

with $f(0) = 1$. In particular $0 \leq f(d) \leq 1$ for all $d \in \mathbf{N}_0^r$. For later purposes we need the following equivalent **partial difference equation** for f

$$f(d) = \sum_m (m-1) \cdot d_m \left[f(d - e_m) - f(d) \right] \tag{1}$$

for $d \neq 0$.

2 Proof of the Main Theorem

For convenience let us define the function $F(H) := \sum_{v \in V} f(d(v))$ for every hypergraph $H = (V, E)$, where $f = f_r$ and $r = \mathrm{rank}(H)$. Suppose x is a vertex of H. Let $H \setminus x$ denote the resulting hypergraph after removing x together with all incident edges from H. The key to the proof of our main theorem is

Lemma 5 *Let $H = (V, E)$ be a hypergraph with $E \neq \emptyset$. Then there exists a vertex $x \in V$ with $F(H \setminus x) \geq F(H)$.*

The main work will be the proof of this lemma.

Proof of Theorem 3. Lemma 5 enables us to use the following algorithm to find an independent set I in H.

```
WHILE E(H) ≠ ∅ DO
    Choose x ∈ V(H) with F(H \ x) ≥ F(H);
    H := H \ x;
END;
Output independent set I = V(H).
```

Since $f(0) = 1$ we know that $F(I) = |I|$. On the other hand the value of F never decreases by the choice of the deleted vertices. Thus $F(H) \leq F(I) = |I| \leq \alpha(H)$.

□

We remark that the proof implies a polynomial algorithm that computes an independent set of size at least $F(H)$ in an arbitrary hypergraph H of constant rank. In particular, for uniform hypergraph, this is the so-called max-algorithm (see also [CT, G]): Successively remove vertices of maximum degree with all incident edges until no edges are left. It is easy to see that a vertex x with maximum degree in a uniform hypergraph has always the property $F(H \setminus x) \geq F(H)$.

3 Proofs of Lemmas

For the proof of Lemma 5 we need

Lemma 6 *Let $r \in \mathbf{N}$, $d \in \mathbf{N}_0^r$ and $\Delta \in [0, d]$ be given. Then*

$$f(d - \Delta) - f(d) \geq \sum_{m=1}^{r} \Delta_m \cdot [f(d - e_m) - f(d)] \ .$$

Proof of Lemma 5. Let $H = (V, E)$ be a hypergraph of rank r with $E \neq \emptyset$. Define V^* to be the set of all non-isolated vertices, i.e. vertices x with $d(x) \neq 0$. By assumption, $V^* \neq \emptyset$. Furthermore for two distinct vertices $x, w \in V$ the *co-degree vector* is given by $d(x, w) = (d_1(x, w), d_2(x, w), \ldots, d_r(x, w)) \in \mathbf{N}_0^r$, where $d_m(x, w)$ is the number of edges of size m containing both x and w. Set $d(w, w) := 0$. Now let $x \in V^*$ be arbitrary, then

$$F(H \setminus x) - F(H) = \sum_{w \in V^*} [f(d(w) - d(x, w)) - f(d(w))] - f(d(x)) \ .$$

Consider one summand. Lemma 6 implies

$$[f(d(w) - d(x, w)) - f(d(w))] \geq \sum_m d_m(x, w) \cdot [f(d(w) - e_m) - f(d(w))] \ .$$

Thus
$$F(H \setminus x) - F(H) \geq \sum_{w \in V^*} \sum_m d_m(x,w) \cdot [f(d(w) - e_m) - f(d(w))] - f(d(x)) \; .$$

We sum these differences up over all $x \in V^*$:

$$\begin{aligned}
\sum_{x \in V^*} [F(H \setminus x) \;\; &- \;\; F(H)] \\
&\geq \sum_{x \in V^*} \sum_{w \in V^*} \sum_m d_m(x,w) \cdot [f(d(w) - e_m) - f(d(w))] - \sum_{x \in V^*} f(d(x)) \\
&= \sum_m \sum_{x \in V^*} \sum_{w \in V^*} d_m(x,w) \cdot [f(d(w) - e_m) - f(d(w))] - \sum_{x \in V^*} f(d(x)) \\
&= \sum_m \sum_{w \in V^*} \left(\sum_{x \in V^*} d_m(x,w) \right) [f(d(w) - e_m) - f(d(w))] - \sum_{x \in V^*} f(d(x)) \\
&= \sum_m \sum_{w \in V^*} (m-1) \cdot d_m(w) \, [f(d(w) - e_m) - f(d(w))] - \sum_{x \in V^*} f(d(x)) \\
&= \sum_{x \in V^*} \left(\sum_m (m-1) \cdot d_m(x) \, [f(d(x) - e_m) - f(d(x))] - f(d(x)) \right) \\
&= 0 \; .
\end{aligned}$$

There we made use of the following observation

$$\sum_{x \in V^*} d_m(x,w) = (m-1) \cdot d_m(w)$$

and the fact that $f(d)$ satisfies the partial difference equation (1) for $d \neq 0$. By definition, $d(x) \neq 0$ for all $x \in V^*$.

We infer that for a random $x \in V^*$ the expectation of $F(H \setminus x) - F(H)$ is non-negative. Thus there exists an $x \in V^* \subset V$ with $F(H \setminus x) \geq F(H)$. □

Lemma 7 *For $r \in \mathbf{N}$, $1 \leq k, l \leq r$ and $d \in \mathbf{N}_0^r$ with $d_k \geq 1$ we have*

$$f(d - e_k) - f(d) \geq f((d + e_l) - e_k) - f(d + e_l) \; .$$

Proof. We will show that

$$[f(d - e_k) - f(d)] - [f((d + e_l) - e_k) - f(d + e_l)] \geq 0 \; .$$

Consider the case $k \neq l$ first.

$$\begin{aligned}
f(d - e_k) - f(d) \;\; &= \;\; \sum_i \binom{d_k - 1}{i_k} \prod_{m \neq k} \binom{d_m}{i_m} \frac{(-1)^{\sum i_m}}{\sum (m-1)\, i_m + 1} \\
&\quad - \sum_i \binom{d_k}{i_k} \prod_{m \neq k} \binom{d_m}{i_m} \frac{(-1)^{\sum i_m}}{\sum (m-1)\, i_m + 1} \\
&= \;\; - \sum_i \binom{d_k - 1}{i_k - 1} \prod_{m \neq k} \binom{d_m}{i_m} \frac{(-1)^{\sum i_m}}{\sum (m-1)\, i_m + 1} \; .
\end{aligned}$$

Similarly

$$f((d + e_l) - e_k) - f(d + e_l) = - \sum_i \binom{d_k - 1}{i_k - 1} \binom{d_l + 1}{i_l} \prod_{m \neq k,l} \binom{d_m}{i_m} \frac{(-1)^{\sum i_m}}{\sum (m-1)\, i_m + 1} \; .$$

Putting this together yields

$$\begin{aligned}[f(d-e_k)-f(d)] \quad - \quad & [f((d+e_l)-e_k)-f(d+e_l)] \\ = \quad & \sum_i \binom{d_k-1}{i_k-1}\binom{d_l}{i_l-1}\prod_{m\neq k,l}\binom{d_m}{i_m}\frac{(-1)^{\sum i_m}}{\sum(m-1)\,i_m+1} \\ = \quad & \sum_i \binom{d_k-1}{i_k}\prod_{m\neq k}\binom{d_m}{i_m}\frac{(-1)^{\sum i_m}}{\sum(m-1)\,i_m+\underbrace{[(k-1)+(l-1)+1]}_{=:C_0>0}} \\ = \quad & g(d-e_k)\,,\end{aligned}$$

where g is given by the recurrence

$$g(d)=\frac{\sum(m-1)\cdot d_m g(d-e_m)}{\sum(m-1)\cdot d_m+C_0}$$

with $g(0)=C_0^{-1}>0$ according to Lemma 4. In particular $g(d-e_k)$ is non-negative which proofs the claim for $k\neq l$.

Now let $k=l$. We have to prove that $[f(d-e_k)-f(d)]-[f(d)-f(d+e_k)]\geq 0$. Consider again

$$f(d-e_k)-f(d)=-\sum_i\binom{d_k-1}{i_k-1}\prod_{m\neq k}\binom{d_m}{i_m}\frac{(-1)^{\sum i_m}}{\sum(m-1)\,i_m+1}$$

and similarly

$$f(d)-f(d+e_k)=-\sum_i\binom{d_k}{i_k-1}\prod_{m\neq k}\binom{d_m}{i_m}\frac{(-1)^{\sum i_m}}{\sum(m-1)\,i_m+1}\,.$$

We infer that

$$\begin{aligned}[f(d-e_k)-f(d)] \quad - \quad & [f(d)-f(d+e_k)] \\ = \quad & \sum_i \binom{d_k-1}{i_k-2}\prod_{m\neq k}\binom{d_m}{i_m}\frac{(-1)^{\sum i_m}}{\sum(m-1)\,i_m+1} \\ = \quad & \sum_i \binom{d_k-1}{i_k}\prod_{m\neq k}\binom{d_m}{i_m}\frac{(-1)^{\sum i_m}}{\sum(m-1)\,i_m+\underbrace{[2\,(k-1)+1]}_{=:C_0>0}} \\ = \quad & g(d-e_k)\,,\end{aligned}$$

where g is again given by the recurrence

$$g(d)=\frac{\sum(m-1)\cdot d_m g(d-e_m)}{\sum(m-1)\cdot d_m+C_0}$$

with $g(0)=C_0^{-1}>0$ according to Lemma 4. In particular $g(d-e_k)$ is non-negative and the claim follows also for $k=l$. □

Remark. Lemma 7 tells us that for any d and k the difference $f(d-e_k)-f(d)$ decreases whenever we increase any component of d. This is essential for the proof of Lemma 6

Proof of Lemma 6. Let $r\in\mathbf{N}$, $d\in\mathbf{N}_0^r$ and $\Delta\in[0,d]$ be given. Consider the points $(d-\Delta)$ and d on the $\mathbf{N}_0^r$ grid. A *monotonical path* between these points is a sequence of grid points starting with $(d-\Delta)$ and terminating with d where two neighboring points are of the form $(d'-e_m)$, d' for some

$1 \leq m \leq r$. Each monotonical path between $(d-\Delta)$ and d has length $\sigma := \sum \Delta_m$ and the number of such paths is given by the multinomial coefficient $\binom{\sum \Delta_m}{\Delta_1,\ldots,\Delta_r}$. Now let $P = p_0, p_1, \ldots, p_\sigma$ be such a monotonical path, $p_0 = d - \Delta$ and $p_\sigma = d$. According to this path we rewrite $f(d-\Delta) - f(d)$ as the telescoping sum

$$f(d-\Delta) - f(d) = \sum_{j=1}^{\sigma} [f(p_{j-1}) - f(p_j)] \ .$$

Note that all differences have the form $f(d'-e_m) - f(d')$ for some $1 \leq m \leq r$ and $d' \in [d-\Delta+e_m, d]$.

For each $1 \leq m \leq r$ there are exactly Δ_m differences of the form $f(d'-e_m) - f(d')$ in the telescoping sum since P is monotonic. By Lemma 7 we see that each such difference satisfies

$$f(d'-e_m) - f(d') \geq f(d-e_m) - f(d) \ .$$

Thus we can estimate

$$f(d-\Delta) - f(d) \geq \sum_m \Delta_m \left[f(d-e_m) - f(d) \right] \ .$$

□

It remains to proof Lemma 4.

Proof of Lemma 4. Let $r \in \mathbf{N}$, $C_1, C_2, \ldots, C_r \geq 0$ and $C_0 > 0$ be given. We have to show that the function $g : \mathbf{N}_0^r \to \mathbf{R}$ given by

$$g(d) = \sum_i \left[\prod \binom{d_m}{i_m} \right] \frac{(-1)^{\sum i_m}}{\sum C_m i_m + C_0}$$

satisfies the recurrence

$$g(d) = \frac{\sum C_k d_k g(d-e_k)}{\sum C_k d_k + C_0}$$

with $g(0) = C_0^{-1}$. It is easy to check that $g(0) = C_0^{-1}$ holds.

Let us rewrite the recurrence as a partial difference equation

$$C_0 g(d) = \sum_k C_k d_k [g(d-e_k) - g(d)]$$

for $d \neq 0$. Suppose $d_k > 0$ then we have

$$\begin{aligned} g(d-e_k) - g(d) &= -\sum_i \binom{d_k - 1}{i_k - 1} \prod_{m \neq k} \binom{d_m}{i_m} \frac{(-1)^{\sum i_m}}{\sum C_m i_m + C_0} \\ &= -\frac{1}{d_k} \sum_i \prod \binom{d_m}{i_m} (-1)^{\sum i_m} \frac{i_k}{\sum C_m i_m + C_0} \ . \end{aligned}$$

Hence,

$$C_k d_k [g(d-e_k) - g(d)] = -\sum_i \prod \binom{d_m}{i_m} (-1)^{\sum i_m} \frac{C_k i_k}{\sum C_m i_m + C_0}$$

and therefore

$$\begin{aligned} \sum_k C_k d_k [g(d-e_k) - g(d)] &= -\sum_i \prod \binom{d_m}{i_m} (-1)^{\sum i_m} \frac{\sum_k C_k i_k}{\sum C_m i_m + C_0} \\ &= -\sum_i \prod \binom{d_m}{i_m} (-1)^{\sum i_m} \left(1 - \frac{C_0}{\sum C_m i_m + C_0} \right) \end{aligned}$$

$$\begin{aligned}
&= \underbrace{-\sum_i \prod \binom{d_m}{i_m} (-1)^{\sum i_m}}_{=0 \text{ for } d\neq 0} \\
&\quad + C_0 \underbrace{\sum_i \prod \binom{d_m}{i_m} \frac{(-1)^{\sum i_m}}{\sum C_m i_m + C_0}}_{=g(d)} \\
&= C_0 g(d)
\end{aligned}$$

as desired. □

References

[C] Y. Caro, New Results on the Independence Number, *Tech. Report, Tel-Aviv University* (1979).

[CT] Y. Caro, Z. Tuza, Improved Lower Bounds on k-Independence, *J. of Graph Theory,* (1991), Vol. 15, p. 99-107.

[G] J. Griggs, Lower Bounds on the Independence Number in Terms of the Degrees, *J. of Comb. Theory, Ser. B 34* (1983), p. 22-39.

[GKP] R. L. Graham, D. E. Knuth, O. Patashnik, Concrete Mathematics, *Addison-Wesley,* (1992), Eight printing.

[W] V. K. Wei, A Lower Bound on the Stability Number of a Simple Graph, *Bell Lab. Tech. Memo.* No. 81-11217-9 (1981).