Stability for maximal independent sets

Jeff Kahn[‡] Jinyoung Park[*]

Department of Mathematics
Rutgers University
New Jersey, U.S.A.

{jkahn, jp1324}@math.rutgers.edu

Submitted: Feb 25, 2019; Accepted: Feb 9, 2020; Published: Mar 20, 2020

Abstract

Answering questions of Y. Rabinovich, we prove "stability" versions of upper bounds on maximal independent set counts in graphs under various restrictions. Roughly these say that being close to the maximum implies existence of a large induced matching or triangle matching (depending on assumptions).

A mild strengthening of one of these results is a key ingredient in a proof (to appear elsewhere) of a conjecture of L. Ilinca and the first author giving asymptotics for the number of maximal independent sets in the graph of the Hamming cube.

Mathematics Subject Classifications: 05C69

1 Introduction

Denote the number of maximal independent sets in a graph G by $\text{mis}(G)$. We recall two well-known bounds for these numbers:

Theorem 1 (Moon-Moser [10]). *For any n-vertex graph G,*

$$\text{mis}(G) \leqslant 3^{n/3},$$

with equality iff G is the disjoint union of $n/3$ triangles.

Theorem 2 (Hujter-Tuza [5]). *For any n-vertex, triangle-free graph G,*

$$\text{mis}(G) \leqslant 2^{n/2},$$

with equality iff G is a perfect matching.

---

[*]Supported by the National Science Foundation under Grant Award DMS1501962

[†]Supported by a Simons Fellowship

As usual M is an *induced matching* of G if it is an induced subgraph of G that is a matching. Similarly, T is an *induced triangle matching* of G if it is an induced subgraph of G that is a vertex disjoint union of triangles.

Write $\mathrm{itm}(G)$ for the number of triangles in a largest induced triangle matching in G, and $\mathrm{im}(G)$ for the number of edges in a largest induced matching.

In what follows we will usually prefer to work with $\log \mathrm{mis}$ ($\log = \log_2$), thought of as the number of bits needed to specify a maximal independent set. Note that $\mathrm{itm}(G) \log 3$ and $\mathrm{im}(G)$ are obvious lower bounds on $\log \mathrm{mis}(G)$. We will be interested in questions suggested to us a few years ago by Yuri Rabinovich [11] concerning "stability" aspects of upper bounds on mis, meaning, roughly: does large mis imply existence of a large induced triangle matching or large induced matching (as appropriate)? Formally, his conjectures were unquantified versions of the following three statements, whose proofs are the content of the present work. (The questions were motivated by [12], which includes a proof of Theorem 4 for bipartite graphs.)

Theorem 3. *For any $\varepsilon > 0$, there is a $\delta = \delta(\varepsilon) = \Omega(\varepsilon)$ such that for an n-vertex graph G, if $\mathrm{itm}(G) < (1-\varepsilon)\frac{n}{3}$ then $\log \mathrm{mis}(G) < (\frac{1}{3}\log 3 - \delta)n$.*

Theorem 4. *For any $\varepsilon > 0$, there is a $\delta = \delta(\varepsilon) = \Omega(\varepsilon)$ such that for a triangle-free n-vertex graph G, if $\mathrm{im}(G) < (1-\varepsilon)\frac{n}{2}$ then $\log \mathrm{mis}(G) < (\frac{1}{2} - \delta)n$.*

One reason to be interested in Theorem 4—or in what its proof actually gives; see Theorem 14 below—is its key role in a proof of the following statement, which was conjectured in [6] (see also [2]) and whose proof is completed in [7] and [8].

Theorem 5. *With Q_n denoting the n-dimensional Hamming cube,*

$$\mathrm{mis}(Q_n) \sim 2n \exp_2[2^{n-2}].$$

While Theorem 14 is one of the easier ingredients in the proof of Theorem 5, it is in some sense the basis for the whole; in particular, it was understanding the connection between induced matchings and stability that first suggested that the conjecture of [6], which had seemed out of reach, might in fact be manageable.

Theorem 4 applies to bipartite graphs, of course. If G is bipartite with bipartition $X \cup Y$, then $\log \mathrm{mis}(G)$ is trivially at most $\min\{|X|, |Y|\}$ (since a maximal independent set is determined by its intersection with either of X, Y); so the statement is uninteresting unless G is close to balanced. But Rabinovich asked whether something analogous also holds for unbalanced (bipartite) G; more precisely, whether something along the following lines is true.

Theorem 6. *For any $\varepsilon > 0$, there is a $\delta = \delta(\varepsilon) = 2^{-O(1/\varepsilon)}$ such that for a bipartite graph G on $X \cup Y$ with $|X| = n$ and $|Y| = 2n$, if $\mathrm{im}(G) < (1-\varepsilon)n$ then $\log \mathrm{mis}(G) < (1-\delta)n$.*

The proof of this is easily adapted to $|Y| = Bn$ (with δ then $\delta(\varepsilon, B)$), but to keep things simple we just state the result for $B = 2$.

Rabinovich suspected that, as in Theorems 3 and 4, $\delta(\varepsilon)$ should be linear in ε, but this is not true. In fact, Theorem 6 is tight (up to the implied constant); a construction to show this will be given in Section 4.2.

The rest of the paper is organized as follows. Section 2 recalls some background, in particular Füredi's upper bounds on mis for paths and cycles, and Shearer's entropy lemma. Section 3 gives the proofs of Theorems 3 and 4. The proof of Theorem 6 and the example to show its tightness are given in Section 4.

The proofs of Theorems 3 and 4 are similar, while that of Theorem 6 is related but somewhat trickier. The general approach has its roots in an idea for counting (ordinary) independent sets due to A.A. Sapozhenko [13], [14].

Strictly speaking we prove the theorems only for sufficiently large n, since we occasionally hide minor terms in $o(1)$'s. Of course combined with the characterizations of equality in Theorems 1 and 2 this does give the stated versions, though the δ's we produce may not be valid for small n. Since we are really interested in large n anyway, this approach seems preferable to carrying explicit error terms.

Notation. We use "$\sim$" for adjacency, $N(x)$ for the neighborhood of x, $N(S) = \cup_{x\in S} N(x)$, and $d_S(x) = |N(x) \cap S|$. As usual, $G[S]$ is the subgraph of G induced by $S \subseteq V(G)$.

2 Preliminaries

For the proof of Theorem 4 we need the following upper bounds on mis for paths and cycles, given by Z. Füredi [4].

Proposition 7. *Let γ (≈ 1.325) be the unique real solution of the equation $1 + \gamma = \gamma^3$.*

1. *For P_n, the path with n vertices,*

$$\mathrm{mis}(P_n) \leqslant 2\gamma^{n-2}.$$

2. *For C_n, the cycle with n vertices,*

$$\mathrm{mis}(C_n) \leqslant 3\gamma^{n-3}.$$

We very briefly recall a few entropy basics (see also e.g. [9]).

For discrete random variables X, Y, the (binary) entropy of X is

$$H(X) = \sum_x p(x) \log \frac{1}{p(x)},$$

and *conditional entorpy of X given Y* is

$$H(X|Y) = \sum_y p(y) \sum_x p(x|y) \log \frac{1}{p(x|y)}$$

(where $p(x) = \mathbb{P}(X = x)$ and $p(x|y) = \mathbb{P}(X = x|Y = y)$).

Lemma 8.

(a) $H(X) \leqslant \log |Range(X)|$, *with equality iff* X *is uniform from its range;*

(b) $H(X, Y) = H(X) + H(Y|X)$.

In addition to these very basic properties we need the following version of Shearer's Lemma [1].

Lemma 9. *If* $\psi = (\psi_1, \ldots, \psi_m)$ *is a random vector and* $\alpha : 2^{[m]} \to \mathbb{R}_{\geqslant 0}$ *satisfies*

$$\sum_{A \ni i} \alpha_A = 1 \quad \forall i \in [m], \tag{1}$$

then

$$H(\psi) \leqslant \sum_{A \subseteq [m]} \alpha_A H(\psi_A) \tag{2}$$

(where $\psi_A = (\psi_i : i \in A)$*).*

Finally, we will need the following standard fact (see e.g. Lemma 16.19 in [3]; this is also implied by Lemma 9 with α_A equal to 1 if $|A| = 1$ and zero otherwise).

Proposition 10. *For* $k \leqslant \frac{1}{2}n$,

$$\sum_{i=0}^{k} \binom{n}{i} \leqslant 2^{H(\frac{k}{n})n}.$$

3 Proofs of Theorems 3 and 4

In this section, I is always a maximal independent set in G. The basis for what we do is the following algorithm, which, given G and I, encodes some portion of Ias a string $\xi(I)$, with the numbers of possibilities for both $\xi(I)$ and the full specification of I given $\xi(I)$ not too large.

3.1 Algorithm

Given G, fix an order "$\prec$" on $V(G)$.

For a given maximal independent set I, let $X_0 = V(G)$ and repeat for $i = 1, 2, \ldots$:

1. Let x_i be the first vertex of X_{i-1} in $\prec$ among those with largest degree in X_{i-1}.

2. If $x_i \in I$ then let $X_i = X_{i-1} \setminus (\{x_i\} \cup N(x_i))$; otherwise, let $X_i = X_{i-1} \setminus \{x_i\}$.

3. Terminate the process if $d_{X_i}(x) \leqslant 2$ for all $x \in X_i$.

Let $X^* = X^*(I) = X_t$ be the final X_i, $t(I) = t$, and $G^* = G^*(I) = G[X^*]$. Define the sequence $\xi = \xi(I) = (\xi_1, \xi_2, \cdots, \xi_t)$ by $\xi_i := \mathbf{1}_{\{x_i \in I\}}$. Notice that ξ encodes a complete description of the run of the algorithm (so we may also write $G^* = G^*(\xi)$), including, in particular, the identities of the x_i's. Finally, let $s = s(I) = |\text{supp}(\xi)|$.

3.2 Proof of Theorem 3

The argument for Theorem 3 goes roughly as follows. Noting that

$$\xi(I) \text{ determines both } X^* \text{ and } I \setminus X^*, \tag{3}$$

and

$$I \cap X^* \text{ is a maximal independent set of } G^*(I)\ (= G^*(\xi)), \tag{4}$$

we find that

$$\mathrm{mis}(G) \leqslant \sum_{\xi} \mathrm{mis}(G^*(\xi)) \tag{5}$$

(If we restrict the sum to *possible* ξ's—those corresponding to actual I's—then we have equality in (5).)

It turns out that running the algorithm for very long is "expensive" in the sense that the loss in $|X^*|$, and so in possibilities for $I \cap X^*$, outweighs what is contributed to (5) by possibilities for ξ; this limits the number of I's with $t(I)$ large. Similarly, the difference between the bounds in Theorems 1 and 2 says there are "few" I's for which the triangle-free part of G^* is large. (Note G^*, having maximum degree at most two, is a disjoint union of triangles and a triangle-free part, below called R.) But the part of $\mathrm{mis}(G)$ corresponding to I's for which both t and R are small must come mainly from counting choices for the restriction of I to the triangles of G^*, and these are limited by our assumption on $\mathrm{itm}(G)$.

To begin with, the following lemma bounds the number of I's with large $t(I)$.

Lemma 11. *Let $\alpha = -\log(4 \cdot 3^{-4/3})$ (≈ 0.113). For any $x \in [0, 1]$,*

$$\log |\{I : t(I) \geqslant xn\}| \leqslant (\frac{1}{3}\log 3 - \alpha x + o(1))n. \tag{6}$$

Proof. For given t and s, consider I's for which $t(I) = t$ and $s(I) = s$. Note that for each such I, $|V(G^*)| \leqslant n - (t + 3s)$, so by Theorem 1 we have $\mathrm{mis}(G^*) \leqslant 3^{(n-(t+3s))/3}$. Also, there are at most $\binom{t}{s}$ possibilities for $\xi(I)$, so by (3) and (4) we have

$$|\{I : t(I) = t, s(I) = s\}| \leqslant \binom{t}{s} 3^{(n-(t+3s))/3}, \tag{7}$$

so

$$|\{I : t(I) = t\}| \leqslant \sum_{s=0}^{t} \binom{t}{s} 3^{(n-(t+3s))/3} = 3^{n/3}\alpha_1^t,$$

where $\alpha_1 = 4 \cdot 3^{-4/3}$. Thus,

$$|\{I : t(I) \geqslant xn\}| \leqslant 3^{n/3}\alpha_1^{xn}/(1 - \alpha_1),$$

yielding (6). □

Let $T = T(I)$ be the union of the triangles in X^* (so the unique maximal induced triangle matching in G^*), $R = R(I) = G^*[X^* \setminus V(T)]$, and $r = r(I) = |V(R)|$. Note that there are no edges between $V(T)$ and $V(R)$, since G^* has maximum degree at most 2, so

$$\mathrm{mis}(G^*) = \mathrm{mis}(T)\mathrm{mis}(R). \tag{8}$$

Note also that R is triangle-free, so

$$\log \mathrm{mis}(R) \leqslant r/2 \tag{9}$$

by Theorem 2. Now, the following lemma bounds the number of I's with large r.

Lemma 12. *Let $\beta = -\log(2^{1/2}3^{-1/3})$ (≈ 0.028). For any $y \in [0,1]$,*

$$\log |\{I : r(I) \geqslant yn\}| \leqslant (\frac{1}{3}\log 3 - \beta y + o(1))n. \tag{10}$$

Proof. By (8) and (9), we have

$$|\{I : r(I) = r, t(I) = t, s(I) = s\}| \leqslant \binom{t}{s} 3^{(n-(t+3s+r))/3} 2^{r/2},$$

so

$$\begin{aligned} |\{I : r(I) = r\}| &\leqslant \sum_{t=0}^{n}\sum_{s=0}^{t}\binom{t}{s}3^{(n-(t+3s+r))/3}2^{r/2} \\ &\leqslant 3^{n/3}\beta_1^r/(1-\alpha_1), \end{aligned}$$

where $\alpha_1 = 4 \cdot 3^{-4/3}$ (as in Lemma 11) and $\beta_1 = 2^{1/2}3^{-1/3}$. Thus,

$$|\{I : r(I) \geqslant yn\}| \leqslant 3^{n/3}\beta_1^{yn}/((1-\alpha_1)(1-\beta_1)),$$

which gives (10). □

Lemma 13. *If* $\mathrm{itm}(G) < (1-\varepsilon)n/3$ *then for any* $x, y \in [0,1]$,

$$\log |\{I : t(I) < xn, r(I) < yn\}| \leqslant ((1-\varepsilon)\frac{1}{3}\log 3 + x + y/2 + o(1))n. \tag{11}$$

Proof. For any I, with $G^* = G^*(I)$ and $r = r(I)$, we have (using (8), (9) and $|V(T(I))| = 3\mathrm{itm}(G^*) < (1-\varepsilon)n$)

$$\mathrm{mis}(G^*) \leqslant 3^{(1-\varepsilon)n/3}2^{r/2}.$$

Therefore,

$$|\{I : t(I) = t, r(I) = r\}| \leqslant 2^t 3^{(1-\varepsilon)n/3}2^{r/2},$$

so

$$\begin{aligned} |\{I : t(I) < xn, r(I) < yn\}| &\leqslant \sum_{t<xn}\sum_{r<yn} 3^{(1-\varepsilon)n/3}2^{r/2+t} \\ &\leqslant 3^{(1-\varepsilon)n/3} \cdot 2^{xn+1} \cdot (\sqrt{2}-1)^{-1}2^{(yn+1)/2}, \end{aligned}$$

giving (11). □

Proof of Theorem 3. This is now just a matter of combining the above bounds. With $\delta_1 = \varepsilon\alpha/8$ and $\delta_2 = \varepsilon\beta/4$, Lemmas 11 - 13 give (respectively)

$$\log|\{I : t(I) \geqslant \delta_1 n/\alpha\}| \leqslant (\frac{1}{3}\log 3 - \delta_1 + o(1))n,$$

$$\log|\{I : r(I) \geqslant \delta_2 n/\beta\}| \leqslant (\frac{1}{3}\log 3 - \delta_2 + o(1))n$$

and (using $\frac{1}{3}\log 3 > 1/2$)

$$\log|\{I : t(I) < \delta_1 n/\alpha,\ r(I) < \delta_2 n/\beta\}| \leqslant (\frac{1}{3}\log 3 - \varepsilon/4 + o(1))n.$$

Thus, with $\delta = \min\{\delta_1, \delta_2, \varepsilon/4\}$ $(= \Omega(\varepsilon))$, we have

$$\log \mathrm{mis}(G) \leqslant (\frac{1}{3}\log 3 - \delta + o(1))n. \qquad \square$$

3.3 Proof of Theorem 4

We first give the slightly stronger version of Theorem 4 mentioned in Section 1. For $I \subseteq V(G)$, write $m(I) = m_G(I)$ for the maximum size of an induced matching M satisfying

- each edge of M meets I and
- there are no edges joining $V(M)$ (the set of vertices covered by M) and $I \setminus V(M)$.

Given G we now write $\mathcal{I} = \mathcal{I}(G)$ for the collection of maximal independent sets of G and set

$$\mathcal{I}_\varepsilon = \mathcal{I}(G, \varepsilon) = \{I \in \mathcal{I}(G) : m(I) < (1-\varepsilon)n/2\}$$

and $\mathrm{mis}(G, \varepsilon) = |\mathcal{I}_\varepsilon|$.

Theorem 14. *For any $\varepsilon > 0$ there is a $\delta = \Omega(\varepsilon)$ such that for any n-vertex, triangle free G,*

$$\log \mathrm{mis}(G, \varepsilon) < (1-\delta)n/2.$$

(We have omitted the corresponding strengthening of Theorem 3.)

As mentioned earlier, the argument for Theorem 14 is similar to the one in Section 3.2, so we will try to be brief. We again start from the algorithm in Section 3.1, and continue to use the notation (X^*, G^* etc.) defined in the paragraph following the algorithm's description. (For most of this we just need $I \in \mathcal{I}$; the role of $\mathcal{I}_\varepsilon$ will appear in Lemma 18.)

Lemma 15. *Let $\alpha = -\log(\frac{1}{\sqrt{2}} + \frac{1}{4})$ (≈ 0.063). For any $x \in [0, 1]$,*

$$\log|\{I \in \mathcal{I} : t(I) \geqslant xn\}| \leqslant (\frac{1}{2} - \alpha x + o(1))n. \qquad (12)$$

Proof. Arguing as for (7) in Section 3.2, we obtain

$$|\{I \in \mathcal{I} : t(I) = t, s(I) = s\}| \leqslant \binom{t}{s} 2^{(n-(t+3s))/2}, \tag{13}$$

where we used $\mathrm{mis}(G^*) \leqslant 2^{(n-(t+3s))/2}$, as given by Theorem 2 (since G is triangle-free). Thus

$$|\{I \in \mathcal{I} : t(I) = t\}| \leqslant \sum_{s=0}^{t} \binom{t}{s} 2^{(n-(t+3s))/2} = 2^{n/2}\alpha_1^t,$$

where $\alpha_1 = \frac{1}{\sqrt{2}} + \frac{1}{4}$, and

$$|\{I \in \mathcal{I} : t(I) \geqslant xn\}| \leqslant 2^{n/2}\alpha_1^{xn}/(1-\alpha_1),$$

yielding (12). □

Say an edge vw of G^* is *isolated* if $G^*[\{v,w\}]$ is a component of G^*. Let $M = M(I)$ be the set of isolated edges in G^*, $R = R(I) = G^*[X^* \setminus V(M)]$, and $r = r(I) = |V(R)|$. Notice that M satisfies the two $\bullet$'s from the definition of $m(I)$ (the first by maximality of I, the second by the definition of M and the fact that there are no edges joining X^* and $I \setminus X^*$); so if $I \in \mathcal{I}_\varepsilon$ then $|M| < (1-\varepsilon)n/2$. Also, since there are no edges between $V(M)$ and $V(R)$,

$$\mathrm{mis}(G^*) = \mathrm{mis}(M)\mathrm{mis}(R). \tag{14}$$

Note that R is triangle-free, so is a vertex-disjoint union of isolated vertices, cycles with at least 4 vertices, and paths with at least 3 vertices. Combining this with Proposition 7, we obtain an upper bound for $\mathrm{mis}(R)$. (Recall that $\gamma \approx 1.325$ was defined in Proposition 7.)

Lemma 16. *With R and r as above,* $\mathrm{mis}(R) \leqslant (3\gamma)^{r/4}$.

Proof. Let l_p (resp. l_c) be the number of vertices in the union of all paths (resp. cycles) in R. Clearly $l_p + l_c \leqslant r$, while the number of paths (resp. cycles) in R is at most $l_p/3$ (resp. $l_c/4$). Thus

$$\begin{aligned} \mathrm{mis}(R) &\leqslant (2/\gamma^2)^{l_p/3}(3/\gamma^3)^{l_c/4}\gamma^r \\ &< (3/\gamma^3)^{r/4}\gamma^r = (3\gamma)^{r/4}, \end{aligned}$$

where the first inequality is given by Proposition 7 and the second follows from the fact that $(2\gamma^{-2})^{1/3} < (3\gamma^{-3})^{1/4}$. □

Lemma 17. *Let $\beta = -\log(2^{-1/2}(3\gamma)^{1/4})$ (≈ 0.0023). For any $y \in [0,1]$,*

$$\log|\{I \in \mathcal{I} : r(I) \geqslant yn\}| \leqslant (\frac{1}{2} - \beta y + o(1))n. \tag{15}$$

Proof. By (14) and Lemma 16,

$$|\{I \in \mathcal{I} : r(I) = r, t(I) = t, s(I) = s\}| \leqslant \binom{t}{s} 2^{(n-(t+3s+r))/2}(3\gamma)^{r/4},$$

and summing this over t and s gives

$$|\{I \in \mathcal{I} : r(I) = r\}| \leqslant 2^{n/2}\beta_1^r/(1-\alpha_1),$$

where $\alpha_1 = \frac{1}{\sqrt{2}} + \frac{1}{4}$ (as in Lemma 15) and $\beta_1 = 2^{-1/2}(3\gamma)^{1/4}$. Thus,

$$|\{I \in \mathcal{I} : r(I) \geqslant yn\}| \leqslant 2^{n/2}\beta_1^{yn}/((1-\alpha_1)(1-\beta_1)),$$

which gives (15). □

Lemma 18. *For any $x, y \in [0,1]$,*

$$\log|\{I \in \mathcal{I}_\varepsilon : t(I) < xn, r(I) < yn\}| \leqslant ((1-\varepsilon)/2 + x + (\log(3\gamma)/4)y + o(1))n. \quad (16)$$

Proof. As in the proof of Lemma 13 (now using $|M| < (1-\varepsilon)n/2$),

$$\mathrm{mis}(G^*) \leqslant 2^{(1-\varepsilon)n/2}(3\gamma)^{r/4}$$

for any $I \in \mathcal{I}_\varepsilon$ with $r(I) = r$. Therefore,

$$|\{I \in \mathcal{I}_\varepsilon : t(I) = t, r(I) = r\}| \leqslant 2^t 2^{(1-\varepsilon)n/2}(3\gamma)^{r/4},$$

and summing over the relevant t's and r's gives

$$|\{I \in \mathcal{I}_\varepsilon : t(I) < xn, r(I) < yn\}| \leqslant 2^{(1-\varepsilon)n/2} \cdot 2^{xn+1} \cdot ((3\gamma)^{1/4} - 1)^{-1}(3\gamma)^{(yn+1)/4};$$

so we have (16). □

Proof of Theorem 14. With $\delta_1 = \varepsilon\alpha/8$ and $\delta_2 = \varepsilon\beta/(2\log(3\gamma))$, Lemmas 15, 17 and 18 give (respectively)

$$\log|\{I \in \mathcal{I} : t(I) \geqslant \delta_1 n/\alpha\}| \leqslant (\frac{1}{2} - \delta_1 + o(1))n,$$

$$\log|\{I \in \mathcal{I} : r(I) \geqslant \delta_2 n/\beta\}| \leqslant (\frac{1}{2} - \delta_2 + o(1))n,$$

and

$$\log|\{I \in \mathcal{I}_\varepsilon : t(I) < \delta_1 n/\alpha, r(I) < \delta_2 n/\beta\}| \leqslant (\frac{1}{2} - \varepsilon/4 + o(1))n.$$

Thus, with $\delta = \min\{\delta_1, \delta_2, \varepsilon/4\}$, we obtain

$$\log \mathrm{mis}(G, \varepsilon) \leqslant (\frac{1}{2} - \delta + o(1))n. \qquad □$$

4 Proof of Theorem 6

For a bipartite graph G on $X \cup Y$, say $X' \subseteq X$ is *irredundant* if $\forall x \in X'$, $N(x) \nsubseteq N(X' \setminus \{x\})$. (So for this discussion "irredundant" sets are always subsets of X.) Denote the number of irredundant sets in G by $\mathrm{irr}(G)$.

Proposition 19. *For any G as above,* $\mathrm{mis}(G) \leqslant \mathrm{irr}(G)$.

Proof. This follows from the observation that for each maximal independent set I there is an irredundant set $J \subseteq I \cap X$ with $N(J) = N(I \cap X)$ $(= Y \setminus I)$; namely, this is true whenever $J \subseteq I \cap X$ is minimal with $N(J) = N(I \cap X)$. □

Thus the following statement implies Theorem 6.

Theorem 20. *For any $\varepsilon > 0$, there is a $\delta = \delta(\varepsilon) = 2^{-O(1/\varepsilon)}$ such that for a bipartite graph G on $X \cup Y$ with $|X| = n$ and $|Y| = 2n$, if $\mathrm{im}(G) < (1-\varepsilon)n$ then $\log \mathrm{irr}(G) < (1-\delta)n$.*

For the rest of this section, G is as in Theorem 20.

4.1 Proof

The algorithm we use for Theorem 20 is slightly different from the one in section 3.1. In what follows, I is always an irredundant set (thus $I \subseteq X$).

Algorithm Let $X_0 = X$, $Y_0 = Y$ and $M = M_\varepsilon = 12/\varepsilon$. Fix an order "$\prec$" on X. For a given I, repeat for $i = 1, 2, \ldots$:

1. Let x_i be the first vertex of X_{i-1} in $\prec$ among those with largest degree in X_{i-1}.
2. If $x_i \in I$ then set $Y_i = Y_{i-1} \setminus N(x_i)$; otherwise, set $Y_i = Y_{i-1}$. In either case, set $X_i = X_{i-1} \setminus \{x_i\}$.
3. Terminate the process if $d_{Y_i}(x) < M$ for all $x \in X_i$.

Let $X^* = X^*(I) = X_t$ and $Y^* = Y^*(I) = Y_t$ be the final X_i and Y_i, respectively. Set $t = t(I)$ and $G^* = G^*(I) = G[X^* \cup Y^*]$. As in section 3.1, define $\xi = \xi(I) = (\xi_1, \xi_2, \cdots, \xi_t)$ by $\xi_i := \mathbf{1}_{\{x_i \in I\}}$, and let $|\xi|$ be the length of ξ (so $|\xi(I)| = t(I)$). Finally, let $s = s(I) = |\mathrm{supp}(\xi)|$ and define $\psi = \psi(I) = I \cap X^*$.

Notice that I is determined by (ξ, ψ), namely (as earlier) $I \setminus X^*$ is determined by ξ (and $I \cap X^* = \psi$).

Consider a random (uniform) irredundant set $\mathbf{I}$. Our various parameters $(\xi, \psi, \ldots)$ are then random variables, which will be denoted by $\boldsymbol{\xi}$ and so on. Since each of $\mathbf{I}$ and $(\boldsymbol{\xi}, \boldsymbol{\psi})$ determines the other and $\boldsymbol{\xi}$ determines $\mathbf{t}$, we have (using parts (a) and (b) of Lemma 8)

$$\begin{aligned} H(\mathbf{I}) &= H(\boldsymbol{\xi}) + H(\boldsymbol{\psi}|\boldsymbol{\xi}) \\ &= H(\mathbf{t}) + H(\boldsymbol{\xi}|\mathbf{t}) + H(\boldsymbol{\psi}|\boldsymbol{\xi}) \\ &\leqslant \log n + H(\boldsymbol{\xi}|\mathbf{t}) + H(\boldsymbol{\psi}|\boldsymbol{\xi}). \end{aligned} \tag{17}$$

Notice that, by Lemma 8 (a),

$$H(\boldsymbol{\xi}|\mathbf{t}=t) \leqslant t$$

for any t and

$$H(\boldsymbol{\psi}|\boldsymbol{\xi}=\xi) \leqslant n-|\xi| \ (=n-t) \tag{18}$$

for any ξ. Thus the sum of the last two terms in (17) is at most

$$\sum_t \mathbb{P}(\mathbf{t}=t)\left[H(\boldsymbol{\xi}|\mathbf{t}=t)+\sum_{|\xi|=t}\mathbb{P}(\boldsymbol{\xi}=\xi|\mathbf{t}=t)H(\boldsymbol{\psi}|\boldsymbol{\xi}=\xi)\right] \leqslant n,$$

and we would like to somewhat improve these bounds. (Since we aim for $H(\mathbf{I}) < n-\Omega(n)$, the $\log n$ in (17) is irrelevant.) The next lemma, giving such a gain in (18) when t is small, is our main point.

Lemma 21. *For any ξ with $|\xi| = t < \varepsilon n/2$,*

$$H(\boldsymbol{\psi}|\boldsymbol{\xi}=\xi) \leqslant n-t-\vartheta n,$$

where $\vartheta = \vartheta(\varepsilon) = 2^{-O(1/\varepsilon)}$.

Proof. Given ξ as in the Lemma, set

$$\tilde{X} = \tilde{X}(\xi) = \{x \in X^* : N_{Y^*}(x) \subseteq N_{Y^*}(X^* \setminus \{x\})\}.$$

We have

$$(1-\varepsilon)n > \mathrm{im}(G) \geqslant \mathrm{im}(G^*) \geqslant n-t-|\tilde{X}|,$$

where the last inequality holds since for each $x \in X^* \setminus \tilde{X}$ there is some $y_x \in Y^*$ with $N_{X^*}(y_x) = \{x\}$, and $\{(x, y_x) : x \in X^* \setminus \tilde{X}\}$ is an induced matching of G^* of size $|X^* \setminus \tilde{X}| = n-t-|\tilde{X}|$. Thus

$$|\tilde{X}| > \varepsilon n - t > \varepsilon n/2. \tag{19}$$

For each $x \in \tilde{X}$ fix some $Z_x \subseteq X^* \setminus \{x\}$ such that

$$N_{Y^*}(x) \subseteq N_{Y^*}(Z_x), \tag{20}$$

$$|Z_x| < M \tag{21}$$

and

$$\forall z \in X^* \quad |\{x \in \tilde{X} : z \in Z_x\}| < 2M. \tag{22}$$

To see that we can do this: For each $y \in N_{Y^*}(\tilde{X})$ let Π_y be a partition of $N_{X^*}(y)$ into blocks of size 2 or 3. (Note $y \in N_{Y^*}(\tilde{X})$ implies $d_{X^*}(y) \geqslant 2$.) Then to form Z_x, for each $y \in N_{Y^*}(x)$ choose one $x' \neq x$ from the block of Π_y containing x and take $x' \in Z_x$. Note that each $x \in X^*$ has degree less than M in G^* (see step 3 of the algorithm), so we have (21) and (22) (and (20) is clear).

Let $W_x = Z_x \cup \{x\}$ $(x \in \tilde{X})$, and $\psi_{_A} = \psi \cap A$ for any $A \subseteq X$. Note that for each $x \in X^*$,

$$\begin{aligned} H(\boldsymbol{\psi}_{W_x}|\boldsymbol{\xi} = \xi) &\leqslant \log[2^{|W_x|} - 1] \\ &= |W_x| + \log(1 - 2^{-|W_x|}) \\ &< |W_x| - 2^{-M}\log e. \end{aligned} \tag{23}$$

(The first inequality follows from irredundancy: we cannot have $\psi_{W_x} = W_x$.)

Now aiming to use Lemma 9, form $\alpha : 2^{X^*} \to \mathbb{R}_{\geqslant 0}$ by assigning weight $1/(2M)$ to each W_x (thus assigning each set weight some multiple of $1/(2M)$, with the total weight of the sets containing any given x' at most 1 by (22)) and supplementing with weights on the singletons to get to (1). Then by Lemma 9,

$$\begin{aligned} H(\boldsymbol{\psi}|\boldsymbol{\xi} = \xi) &\leqslant \sum_{A \subseteq X^*} \alpha_A H(\boldsymbol{\psi}_A|\boldsymbol{\xi} = \xi) \\ &= \sum_{x \in \tilde{X}} \alpha_{W_x} H(\boldsymbol{\psi}_{W_x}|\boldsymbol{\xi} = \xi) + \sum_{x \in X^*} \alpha_{\{x\}} H(\boldsymbol{\psi}_{\{x\}}|\boldsymbol{\xi} = \xi). \end{aligned} \tag{24}$$

Now (23) and the fact that α assigns total weight $|\tilde{X}|/(2M)$ to the W_x's give

$$\sum_{x \in \tilde{X}} \alpha_{W_x} H(\boldsymbol{\psi}_{W_x}|\boldsymbol{\xi} = \xi) < \sum_{x \in \tilde{X}} \alpha_{W_x} |W_x| - |\tilde{X}|(2M2^M)^{-1}\log e,$$

while the second sum in (24) is at most $\sum_{x \in X^*} \alpha_{\{x\}}$ (since $H(\boldsymbol{\psi}_{\{x\}}|\boldsymbol{\xi} = \xi) \leqslant 1$). Thus the entire bound in (24) is at most

$$\begin{aligned} \sum_{x \in \tilde{X}} \alpha_{W_x}|W_x| + \sum_{x \in X^*} \alpha_{\{x\}} - |\tilde{X}|(2M2^M)^{-1}\log e &= |X^*| - |\tilde{X}|(2M2^M)^{-1}\log e \\ &< n - t - \vartheta n, \end{aligned}$$

where $\vartheta = (\varepsilon/2)(2M2^M)^{-1}\log e = 2^{-O(1/\varepsilon)}$ (see (19)) and we use

$$\sum_{x \in \tilde{X}} \alpha_{W_x}|W_x| + \sum_{x \in X^*} \alpha_{\{x\}} = \sum_{x \in X^*} \sum_{A \ni x} \alpha_A = |X^*|. \qquad \square$$

Corollary 22. *Let $\zeta = \mathbb{P}(\mathbf{t} < \varepsilon n/2)$. Then with ϑ as in Lemma 21,*

$$H(\boldsymbol{\psi}|\boldsymbol{\xi}) \leqslant n - \mathbb{E}\mathbf{t} - \zeta\vartheta n.$$

Proof. Using Lemma 21 and (18) we have

$$\begin{aligned} H(\boldsymbol{\psi}|\boldsymbol{\xi}) &= \sum_t \sum_{|\xi|=t} \mathbb{P}(\boldsymbol{\xi} = \xi) H(\boldsymbol{\psi}|\boldsymbol{\xi} = \xi) \\ &\leqslant \sum_{t < \varepsilon n/2} \mathbb{P}(\mathbf{t} = t)(n - t - \vartheta n) + \sum_{t \geqslant \varepsilon n/2} \mathbb{P}(\mathbf{t} = t)(n - t) \\ &= n - \mathbb{E}\mathbf{t} - \zeta\vartheta n. \end{aligned} \qquad \square$$

The gain for larger t is easier. Noting that

$$\mathbf{s} \leqslant s_0 := 2n/M = \varepsilon n/6,$$

setting $H(1/3) = 1 - \gamma$ and using Proposition 10, we have, for any $t \geqslant \varepsilon n/2$,

$$H(\boldsymbol{\xi}|\mathbf{t} = t) \leqslant \log \sum_{s \leqslant s_0} \binom{t}{s} \leqslant H(1/3)t = (1-\gamma)t,$$

whence (recall $\zeta = \mathbb{P}(\mathbf{t} < \varepsilon n/2)$)

$$\begin{aligned} H(\boldsymbol{\xi}|\mathbf{t}) &\leqslant \sum_{t<\varepsilon n/2} \mathbb{P}(\mathbf{t} = t)t + \sum_{t \geqslant \varepsilon n/2} \mathbb{P}(\mathbf{t} = t)(1-\gamma)t \\ &\leqslant \mathbb{E}\mathbf{t} - (1-\zeta)\gamma\varepsilon n/2. \end{aligned}$$

Finally, combining this with (17) and Corollary 22 yields

$$\begin{aligned} H(\mathbf{I}) &\leqslant \log n + n - [\zeta\vartheta + (1-\zeta)\gamma\varepsilon/2]n \\ &\leqslant \log n + n - \vartheta n \end{aligned}$$

(since the ϑ produced in Lemma 21 is much smaller than $\gamma\varepsilon/2$), proving Theorem 20.

4.2 Tightness

Define a bipartite graph B_m on $X \cup Y = [m] \cup [2m]$ (disjoint copies, of course) as follows.

1. If $x \in X$ and $x \leqslant m-1$, then $x \sim y$ iff $y = x$ or $y = m - 1 + x$.
2. If $x = m \in X$, then $x \sim y$ iff $m \leqslant y \leqslant 2m - 2$.

It is easy to see that $\mathrm{im}(B_m) = m - 1$, and $\mathrm{mis}(B_m) = 2^m - 1$.

Now, for $\varepsilon > 0$ and n with $1/\varepsilon$ and εn integers, let G be the union of εn disjoint copies of $B_{1/\varepsilon}$. Then G is bipartite on $[n] \cup [2n]$, $\mathrm{im}(G) = (1-\varepsilon)n$, and $\mathrm{mis}(G) = (2^{1/\varepsilon} - 1)^{\varepsilon n}$. So,

$$\begin{aligned} \log \mathrm{mis}(G) &= \varepsilon n \log(2^{1/\varepsilon} - 1) \\ &= \varepsilon n(\frac{1}{\varepsilon} + \log(1 - 2^{-1/\varepsilon})) \\ &= n(1 - 2^{-1/\varepsilon}\varepsilon \log e + O(2^{-2/\varepsilon})) \end{aligned}$$

(where the implied constant does not depend on ε).

Acknowledgements

We would like to thank Yuri Rabinovich for suggesting these questions.

References

[1] F.R.K. Chung, P. Frankl, R. Graham and J.B. Shearer, Some intersection theorems for ordered sets and graphs, pp. 23-37 in *J. Comb. Theory Ser. A* **48**, 1986.

[2] D. Duffus, P. Frankl and V. Rödl, Maximal independent sets in the covering graph of the cube, pp. 1203-1208 in *Discrete Applied Mathematics* **161**, 2013.

[3] J. Flum and M. Grohe, Parameterized Complexity Theory, Springer, New York, 2006.

[4] Z. Füredi, The number of maximal independent sets in connected graphs, pp. 463-470 in *J. Graph Theory* **11**, 1987.

[5] M. Hujter and Z. Tuza, The number of maximal independent sets in triangle-free graphs, pp. 284-288 in *SIAM J. Discrete Math* **6**, 1993.

[6] L. Ilinca and J. Kahn, Counting maximal antichains and independent sets, pp. 427-435 in *Order* **30**, 2013.

[7] J. Kahn and J. Park, An isoperimetric inequality for the Hamming cube and some consequences, preprint. arXiv:1909.04274v2 [math.CO]

[8] J. Kahn and J. Park, The number of maximal independent sets in the Hamming cube, preprint. arXiv:1909.04283v2 [math.CO]

[9] R.J. McEliece, The Theory of Information and Coding, *Addison-Wesley*, London, 1977.

[10] J.W. Moon and L. Moser, On cliques in graphs, pp. 23-28 in *Israel J. Math.* **3**, 1965.

[11] Y. Rabinovich, personal communication, 2013.

[12] Y. Rabinovich, J.A. Telle and M. Vatshelle, Upper Bounds on Boolean-Width with Applications to Exact Algorithms, pp. 308-320 in *IPEC 2013: Parameterized and Exact Computation*, Springer, 2013.

[13] A.A. Sapozhenko, On the Number of Independent Sets in Extenders, pp. 56-62 in *Diskret. Matem.* **13** (1), 2001 (in Russian).

[14] A.A. Sapozhenko, The Number of Independent Sets in Graphs, pp. 116-118 in *Moscow Univ. Math. Bull.* **62**, 2007.