DIMACS Technical Report 2000-25
August 2000

On the Number of Independent Sets in Bipartite Graphs with Large Minimum Degree

by

A.A. Sapozhenko
Moscow State University

DIMACS is a partnership of Rutgers University, Princeton University, AT&T Labs-Research, Bell Labs, NEC Research Institute and Telcordia Technologies (formerly Bellcore).

DIMACS is an NSF Science and Technology Center, funded under contract STC–91–19999; and also receives support from the New Jersey Commission on Science and Technology.

ABSTRACT

The asymptotics of the independent set number is found for bipartite graphs which minimum degree is equal in order to the number of vertices.

All graphs under consideration are finite, undirected, and simple. Let $I(G)$ denote the number of independent sets of a graph G. For a graph $G=(V;\,E)$ and a set $A\subseteq V$, we call the set $\partial(A)=\{u\in V\setminus A:\exists v\in A\ \ \{u,v\}\in E\}$ the *boundary* of A. In case $A=\{v\}$, we write $\partial(v)$ instead of $\partial(\{v\})$. A bipartite graph $G=(X,Z;E)$ with the vertex parts X, Z and the edge set E will be called an (ϵ,δ)-*expander* if $|A|\le|\partial(A)|(1-\delta)$ for any $A\subseteq X$ such that $|A|\le\lceil\epsilon\,|X|\rceil$ and for any $A\subseteq Z$ such that $|A|\le\lceil\epsilon\,|Z|\rceil$.

Some historical comments. The asymptotics for the number of independent sets in the n-cube were found in [KS]. The asymptotics for the number of independent sets in expanders were obtained in [S1]. The last result is founded on the upper bounds (see [S3]) for the number of connected sets with given size of the boundary in bipartite graphs. The similar result for arbitrary (non-bipartite) graphs was obtained in [S4]. In [S2] these asymptotics were used for solving the Dedekind problem on the number of monotone Boolean functions. In [A], N.Alon established the logarithmic asymptotics of $I(G)$ for k-regular graphs and for "almost regular" bipartite graphs. Besides, in [A] these results were applied to find the logarithmic asymptotics of the number of sum-free sets in finite groups of even order.

This paper presents bounds for the number of sets with given boundary size in bipartite graphs of large minimum vertex degree. Using these bounds, we obtain the asymptotics for $I(G)$ under weaker restrictions on vertex degrees of G in comparison with [S1]. The assumption that the vertex degree is sufficiently large allows us to simplify the proof and generalize it for the irregular graphs.

Denote by $\widetilde{I}(G)$ the number of independent sets C of graph G such that $C\cap X\neq\emptyset$ and $C\cap Z\neq\emptyset$. Throughout of the paper, we assume that the number vertices of considered graphs is sufficiently large. The main result of this paper is

Theorem 1 *Let graph $G=(X,Z;E)$ on n vertex with minimum degree k and maximum degree $k+\theta$ be an $(1/2,\delta)$-expander, $k\ge cn$, where $c>0$ is an absolute constant. Further, let z be the largest of the solutions of the equation*[1] $x=\log(2ex/c\delta)$. *Then*

$$2^{|X|}+2^{|Z|}-1\le I(G)\le\left(2^{|X|}+2^{|Z|}\right)\left(1+2^{-k\delta/z+O\left(\sqrt{k}\log k+\theta\right)}\right),\quad(1)$$

[1] Everywhere $\log a$ denotes $\log_2 a$

and
$$\widetilde{I}(G) \le 2^{-k\delta/z + O\left(\sqrt{k}\log k + \theta\right)}. \tag{2}$$

Corollary 1 *Let the conditions of Theorem* 1 *hold and*
$$\theta/k + k^{-1/2}\log k = o(\delta) \tag{3}$$
as $n \to \infty$. *Then*
$$I(G) \sim 2^{|X|} + 2^{|Z|}. \tag{4}$$

We need some auxiliary assertions to prove Theorem 1. Let $G = (X, Z; E)$ be a bipartite graph, $A \subseteq X$ or $A \subseteq Z$. Define a *closure* of A as the set $[A] = \{v \in X \cup Z \,:\, \partial(v) \subseteq \partial(A)\}$. A subset A is called *closed* if $A = [A]$.

Lemma 1 *Let a bipartite graph* $G = (X, Z; E)$ *satisfy the conditions:*
(1) $|\partial(A)| \ge |A|$ *for all* $A \subseteq X$ *such that* $|A| \le \lceil |X|/2 \rceil$,
(2) $|\partial(B)| \ge |B|$ *for all* $B \subseteq Z$ *such that* $|B| \le \lceil |X|/2 \rceil$.
Then for an arbitrary independent set C *at least one of the following inequalities holds:*
$$|[C \cap X]| \le |X|/2, \qquad |[C \cap Z]| \le |Z|/2. \tag{5}$$

Proof. Assuming the opposite, we have $|[C \cap X]| > |X|/2$ and $|[C \cap Z]| > |Z|/2$. Note that for every independent set $C \subseteq X \cup Z$ the set $[C \cap X] \cup [C \cap Z]$ is also independent. Without loss of generality we may assume $|Z| \ge |X|$. For an arbitrary $B \subseteq [C \cap Z]$ such that $|B| = \lceil |Z|/2 \rceil$, we have
$$|\partial([C \cap Z])| + |[C \cap X]| \ge |\partial(B)| + |[C \cap X]| > |Z|/2 + |X|/2 \ge |X|.$$
It follows that $\partial([C \cap Z]) \cap [C \cap X] \ne \emptyset$. This contradicts the independence of C. $\square$

For a bipartite graph $G = (X, Z; E)$, $M \in \{X, Z\}$, and integers a, g, we set
$$\mathcal{A}(M) = \{A \subseteq M : \, |[A]| \le \lceil |M|/2 \rceil\},$$
$$\mathcal{A}(M, g) = \{A \subseteq \mathcal{A}(M) : \, |\partial(A)| = g\},$$
$$\mathcal{A}(M, g, a) = \{A \in \mathcal{A}(M, g) : \, |[A]| = a\}.$$

Lemma 1 implies the following

Corollary 2 *Let $G = (X, Z; E)$ be the bipartite graph satisfying the conditions of Lemma* 1*. Then*

$$I(G) \leq \sum_{A \in \mathcal{A}(X)} 2^{|Z| - |\partial(A)|} + \sum_{A \in \mathcal{A}(Z)} 2^{|X| - |\partial(A)|} \tag{6}$$

Lemma 2 *Let $G = (X, Z; E)$ be a bipartite graph with minimum degree k and maximum degree $k + \theta$, $0 < \varphi \leq k$, $A \subseteq X$. Then there exists a set $T \subseteq A$ such that*

$$|T| \leq |\partial(A)|/\varphi, \tag{7}$$

$$A \subseteq D, \tag{8}$$

where $D = \{v \in X : |\partial(v) \setminus \partial(T)| < \varphi\}$, and

$$|D| \leq |\partial(T)| \frac{k + \theta}{k - \varphi}. \tag{9}$$

Proof. We construct the set T using the following step-by-step procedure.

Step 1. Let u_1 be an arbitrary vertex from A. Set $T_1 = \{u_1\}$.
Suppose m steps are already made and a set $T_m = \{u_1, \ldots, u_m\}$ is constructed.

Step $m+1$. If there exists $u_{m+1} \in A$ such that $|\partial(u_{m+1}) \setminus \partial(T_m)| \geq \varphi$, we assume $T_{m+1} = T_m \cup \{u_{m+1}\}$. Otherwise, the procedure is complete and the result is $T = T_m$. The inequality (7) and the inclusion (8), evidently, hold for the set T constructed as above. The inequality (9) follows from the facts that $|\partial(v) \cap \partial(T)| \geq k - \varphi$ for every $v \in D$ and $|\partial(u)| \leq k + \theta$ for every $u \in \partial(T)$. □

Theorem 2 *Let $G = (X, Z; E)$ be abipartite graph on n vertices with minimum degree k and maximum degree $k + \theta$, where $k \geq cn$, $c > 0$ does not depend on n. Moreover, let a and g be integers, $a \leq g$, $M \in \{X, Z\}$, and y be the largest of the solutions of the equation $x = \log(2enx/(g - a))$. Then*

$$|\mathcal{A}(M, g, a)| \leq 2^{g\left(1 - \frac{g-a}{gy} + O\left(\frac{\log k}{\sqrt{k}} + \frac{\theta}{k}\right)\right)}. \tag{10}$$

Proof. Set $0 < \varphi \leq k$. Suppose $A \in \mathcal{A}(M, g)$. Then by Lemma 2 there exists the set $T \subseteq M$ satisfying conditions (7) – (9). The number of sets $T \subseteq M$ which satisfy (7) does not exceed[2]

$$\sum_{i \le g/\sqrt{k}} \binom{n}{i} \le 2^{\frac{g}{\sqrt{k}} \log k}. \tag{11}$$

Let $T \subseteq M$. Assume $\mathcal{A}(M, g, a, T) = \{B \subseteq A : \ A \in \mathcal{A}(M, g, a) : T \subseteq A\}$. Note that

$$|\mathcal{A}(M, g, a)| \le \sum_{T \subseteq M, \, |T| \le g/\sqrt{k}} |\mathcal{A}(M, g, a, T)|. \tag{12}$$

Now we bound $|\mathcal{A}(M, g, a, T)|$ from above. Two cases are possible.

1) $|\partial(T)| \le g - (g - a)/y$. Then using the notations of Lemma 2 and the inequality (9), we have

$$|D| \le |\partial(T)| \frac{k + \theta}{k - \sqrt{k}} \le g \left(1 - \frac{g - a}{gy}\right) \frac{k + \theta}{k - \sqrt{k}}.$$

Taking into account (8), we have

$$|\mathcal{A}(M, g, a, T)| \le 2^{|D|} \le 2^{g\left(1 - \frac{g-a}{gy}\right)\frac{k+\theta}{k-\sqrt{k}}} = 2^{g\left(1 - \frac{g-a}{gy} + O\left(\frac{1}{\sqrt{k}} + \frac{\theta}{k}\right)\right)}. \tag{13}$$

2) $|\partial(T)| > g - (g - a)/y$. In this case, choosing a certain subset $H \subseteq Z$ for a given T as $\partial(A) \setminus \partial(T)$, we obtain the set $\partial(A) = \partial(T) \cup H$, that uniquely determines $[A]$. The number of possibilities to pick H does not exceed $\sum_{i \le (g-a)/y} \binom{n}{i}$, and the number of of possibilities to pick B, such that $[B] = A$ and $T \subseteq B$, does not exceed 2^a. Hence using the equality $y = \log(2eny/(g - a)$, we have

$$|\mathcal{A}(M, g, a, T)| \le \sum_{i \le (g-a)/y} \binom{n}{i} 2^a \le 2^{a + \frac{g-a}{y} \log\left(eny/(g-a)\right)} = 2^{g - \frac{g-a}{y}}. \tag{14}$$

Now (10) follows from (11) – (14). $\square$

Remark 1 *Let y be the largest of the solutions of the equation $x = \log(bx)$, $e \ln 2 \le b \le d$, and z be the largest of the solutions of the equation $x = \log(dx)$. Then $y \le z$.*

[2]Hereinafter we use the inequality $\sum_{i \le k} \binom{n}{i} \le (en/k)^k$ (for example see [S4]).

Proof of Theorem 1. The lower bound in (1) is trivial. The upper one will follow from (2). So it is sufficient to prove (2). Since G is a $(1/2, \delta)$-expander, similarly to (6), we have

$$\widetilde{I}(G) \le \sum_{A \in \mathcal{A}(X),\, A \ne \emptyset} 2^{|Z| - |\partial(A)|} + \sum_{A \in \mathcal{A}(Z),\, A \ne \emptyset} 2^{|X| - |\partial(A)|}. \tag{15}$$

Now we get the upper bound for the sum $S_1 = \sum_{A \in \mathcal{A}(X),\, A \ne \emptyset} 2^{-|\partial(A)|}$. Note that by the condition $|\partial(A)| \ge k \ge cn$ for any nonempty set A. Using (10), denotations of Theorem 1, Remark 1, and taking into account that a graph $G = (X, Z; E)$ is $(1/2, \delta)$-expander, we have

$$S_1 = \sum_{g \ge k} \sum_{a \le g(1-\delta)} \sum_{A \in \mathcal{A}(X,g,a)} 2^{-|\partial(A)|} = \sum_{g \ge k} \sum_{a \le g(1-\delta)} |\mathcal{A}(X, g, a)|\, 2^{-g} \le$$

$$\le \sum_{g \ge k} \sum_{g-a \ge g\delta} 2^{-\frac{g-a}{y} + O\left(\frac{\log k}{\sqrt{k}} + \frac{\theta}{k}\right)g} \le \sum_{g \ge k} 2^{O\left(\frac{\log k}{\sqrt{k}} + \frac{\theta}{k}\right)g} \sum_{i \ge g\delta} 2^{-i/z} =$$

$$= \frac{1}{1 - 2^{-1/z}} \sum_{g \ge k} 2^{g\left(-\frac{\delta}{z} + O\left(\frac{\log k}{\sqrt{k}} + \frac{\theta}{k}\right)\right)} \le 2^{-k\delta/z + O\left(\sqrt{k}\log k + \theta\right)}. \tag{16}$$

Similarly for sum $S_2 = \sum_{A \in \mathcal{A}(Z),\, A \ne \emptyset} 2^{-|\partial(A)|}$ we obtain

$$S_2 \le 2^{-k\delta/z + O\left(\sqrt{k}\log k + \theta\right)}. \tag{17}$$

Now (15), (16) and (17) imply (2), and, hence, the upper bound in (1). $\square$

Remark 2 *In fact, the upper bound in (4) remains still valid under much weaker restrictions on k and δ than in Corollary 1. Namely, the inequality $k \ge cn$ can be replaced by the inequality $k \ge f(n)$, where $f(n) = o(n)$, and a constant δ be can be replaced by a function $\delta(n) = o(1)$. It is only important that the following condition holds $\sqrt{k}\log k + \theta = o(k\delta/z)$, where z is the largest of the solutions of the equation $x = \log(2exn/f(n)\delta(n))$.*

This research is supported by the RFFI grant No 98-01-00509.

Bibliography

[A] N.Alon. Independent sets in regular graphs and sum-free subsets of finite groups.// Israel J. of Math. Vol. 73, No 2, 1991, 247-256.

[KS] A.D.Korshunov, A.A. Sapozhenko. On the number of binary codes with distance two.// Problemy Kibernetiki, M.Nauka, Vol. 40, 1993, 111-140 (in Russian).

[S1] A.A. Sapozhenko. On the number of antichains in ranked posets (in Russian).// Discrete Mathematics - Moscow, Nauka, v.1, No 1, 1989 p.74-93. Engl. Translation in Discrete Mathematics and Applications - Utrecht, The Netherlands, Tokio, Japan, - v.1, No 1, 35-59.

[S2] A.A. Sapozhenko, On the number of antichains in multilevelled ranked posets (in Russian)// Discrete Mathematics - Moscow, Nauka, v.1, No.2 - 1989 - 110-128. Engl. Translation in Discrete Mathematics and Applications - Utrecht, The Netherlands, Tokio, Japan, - v.1,No.2 149-171.

[S3] A.A.Sapozhenko, On the number of connected sets of given boundary size,// Metody diskretnogo analiza v reshenii combinatornykh zadach, — 1987, No 3, 42-70 (in Russian).

[S4] A.A.Sapozhenko, On the number of connected sets of given neighborhood size,// Diskretnyj analiz i issledovanie operacij, — 1997, Serie 1, Tom 4, No 3, 18-34 (in Russian).