The choice number of random bipartite graphs

Noga Alon *   Michael Krivelevich †

Abstract

A random bipartite graph $G(n,n,p)$ is obtained by taking two disjoint subsets of vertices A and B of cardinality n each, and by connecting each pair of vertices $a \in A$ and $b \in B$ by an edge randomly and independently with probability $p = p(n)$. We show that the choice number of $G(n,n,p)$ is, almost surely, $(1+o(1))\log_2(np)$ for all values of the edge probability $p = p(n)$, where the $o(1)$ term tends to 0 as np tends to infinity.

1 Introduction

A graph $G = (V,E)$ is called *k-choosable*, for an integer $k > 0$, if for every family of color lists $\mathcal{S} = \{S(v) \subset Z : v \in V(G)|\}$, satisfying $|S(v)| = k$ for every $v \in V$, there exists a choice function $f : V \rightarrow Z$ such that $f(v) \in S(v)$ for all $v \in V$, and also $f(u) \neq f(v)$ for every edge $e = (u,v) \in E(G)$. The *choice number* $ch(G)$ of G is the minimal integer k for which G is k-choosable.

The concept of choosability was introduced by Vizing in 1976 [6] and independently by Erdős, Rubin and Taylor in 1979 [5]. Although the choice number is a straightforward generalization of the more familiar notion of the chromatic number, after twenty years of research it appears to be a much more complicated quantity, and much less is known about it. The reader may consult the survey paper of the first author [2] for a discussion of various problems and results on choosability.

Our interest in choosability questions for bipartite graphs is stimulated in particular by the fact that bipartite graphs provide a standard example of a family of graphs for which the choice number can be much higher than the chromatic number. (It follows immediately from the definition of the choice number that $ch(G) \geq \chi(G)$ for every graph G). Indeed, Erdős , Rubin and Taylor noticed in their original paper that the choice number of the complete bipartite graph $K_{d,d}$ is $(1+o(1))\log_2 d$.

*Department of Mathematics, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, Israel and Institute for Advanced Study, Princeton, NJ 08540. Email: noga@math.tau.ac.il. Research supported in part by a USA Israeli BSF grant, by a grant from the Israel Science Foundation, by a Sloan Foundation grant No. 96-6-2 and by a State of New Jersey grant.

†School of Mathematics, Institute for Advanced Study, Princeton, NJ 08540. Email: mkrivel@math.ias.edu. Research supported by an IAS/DIMACS Postdoctoral Fellowship.

Mathematics Subject Classification (1991): 05C15, 05C35. Running title: Choosability of random bipartite graphs.

The (relatively simple) proof of this statement contains several useful ideas, some of which are applied in the present note.

The main goal of this note is to determine the asymptotic behavior of the choice number of random bipartite graphs. Formally, the *random bipartite graph* $G(m, n, p)$ is the probability space whose points are bipartite graphs on a fixed set of $m + n$ labeled vertices, partitioned into two color classes A and B of cardinalities $|A| = m$, $|B| = n$, respectively. Each pair of vertices $a \in A$ and $b \in B$ form an edge randomly and independently with probability $p = p(m, n)$. The color classes A and B form independent sets. By the term "the random bipartite graph" we mean a random point chosen in this probability space. In this note we confine ourselves to the case where the color classes A and B are of equal cardinality $|A| = |B| = n$, the corresponding model is denoted by $G(n, n, p)$. We use the usual notational conventions of the theory of random graphs. As is usually the case in this subject, we are interested in the behaviour of various parameters as n tends to infinity. We say that a graph property A holds *almost surely*, or a.s., for short, if the probability that $G(n, n, p)$ satisfies A tends to 1 as n tends to infinity. Note that $d = np$ is the expected degree of each vertex of $G(n, n, p)$.

The problem of determining the asymptotic value of the choice number of random bipartite graphs has been addressed already by Erdős, Rubin and Taylor [5]. They considered the model $G(n, n, 1/2)$ and proved that a.s. $\log n/\log 6 < ch(G(n, n, 1/2)) < 3\log n/\log 6$. In this paper we determine the typical asymptotic value of $ch(G(n, n, p))$ for all values of the edge probability $p = p(n)$ down to $p \geq C/n$ for some constant $C > 0$. We prove the following theorem.

Theorem 1.1 *There exists an absolute constant d_0 such that if the edge probability $p = p(n)$ satisfies $np > d_0$, then a.s.*

$$\log_2(np) - 4\log_2\log_2(np) \leq ch(G(n, n, p)) \leq \log_2(np) + \frac{5\log_2(np)\log_2\log_2\log_2(np)}{\log_2\log_2(np)} \ .$$

Thus, the choice number of $G(n, n, p)$ is almost surely $(1 + o(1))\log_2 d$ for all values of $p(n)$. We do not make here any serious attempt to optimize the error terms in Theorem 1.1, our main task is to find an asymptotic formula for the main term.

The rest of the paper is organized as follows. In the next section we present some properties of the edge distribution in random bipartite graphs, needed for the subsequent proofs. The proof of the lower bound is given in Section 3. Section 4 is devoted to the proof of the upper bound. Section 5, the final section of the paper, contains several concluding remarks and a discussion of relevant open problems.

Throughout the paper, all logarithms are in base 2. As mentioned above we denote by $d = np$ the expected vertex degree in $G(n, n, p)$ and assume, whenever needed, that d (and hence also n) is sufficiently large. We omit routinely floor and ceiling signs whenever these are not crucial, to simplify the presentation.

2 Preliminaries

In this section we prove two technical propositions about the edge distribution in bipartite random graphs. As shown in the next two sections, these propositions are essentially the only properties of random bipartite graphs that are needed to prove our result.

Proposition 2.1 *The random bipartite graph $G(n,n,p)$ has a.s. the following property: for every two subsets $X \subset A$, $Y \subset B$ of cardinalities $|X|, |Y| \geq 2n\log d/d$, there exists an edge $e \in E(G)$, connecting X and Y.*

Proof. The probability that there exists a pair X, Y, violating the assertion of the proposition, is at most

$$\binom{n}{\frac{2n\log d}{d}}^2 (1-p)^{\left(\frac{2n\log d}{d}\right)^2} \leq \left[\left(\frac{ed}{2\log d}\right)^2 e^{-\frac{2np\log d}{d}}\right]^{\frac{2n\log d}{d}} \leq \left[d^2 e^{-2\log_2 d}\right]^{\frac{2n\log d}{d}} = o(1). \qquad \Box$$

Proposition 2.2 *For every fixed constant $C > 0$ the random bipartite graph $G(n,n,p)$ has a.s. the following property: for every two subsets $X \subset A$, $Y \subset B$, satisfying $|X|, |Y| \leq Cn/d$, the spanned subgraph $G[X \cup Y]$ of G on $X \cup Y$ has at most $(|X|+|Y|)3\log d/\log\log d$ edges.*

Proof. The result is trivial for $d = \Omega(n)$, hence we may and will assume that $d = o(n)$. Denote $\epsilon_1 = \epsilon_1(d) = 3\log d/\log\log d$. The probability of existence of subsets X, Y, violating the proposition, can be bounded from above by

$$\sum_{i,j=1}^{C\frac{n}{d}} \binom{n}{i}\binom{n}{j}\binom{ij}{(i+j)\epsilon_1} p^{(i+j)\epsilon_1} \leq 2\sum_{1\leq i\leq j\leq Cn/d} \binom{n}{i}\binom{n}{j}\binom{ij}{(i+j)\epsilon_1} p^{(i+j)\epsilon_1}$$

$$\leq 2\sum_{1\leq i\leq j\leq Cn/d} \binom{n}{j}^2 \left(\frac{eij}{(i+j)\epsilon_1}\right)^{(i+j)\epsilon_1} p^{(i+j)\epsilon_1} \leq 2\sum_{1\leq i\leq j\leq Cn/d} \left(\frac{en}{j}\right)^{2j}\left(\frac{ejp}{\epsilon_1}\right)^{j\epsilon_1}$$

$$= 2\sum_{j=1}^{C\frac{n}{d}} j\left[\left(\frac{en}{j}\right)^2\left(\frac{ejp}{\epsilon_1}\right)^{\epsilon_1}\right]^j = 2\sum_{j=1}^{C\frac{n}{d}} j\left[\left(\frac{en}{j}\right)^2 (ejp)^2 \frac{(ejp)^{\epsilon_1-2}}{\epsilon_1^{\epsilon_1}}\right]^j$$

$$\leq 2\sum_{j=1}^{C\frac{n}{d}} j\left[\frac{e^4 d^2 (ejp)^{\epsilon_1-2}}{d^{3-o(1)}}\right]^j \leq 2\sum_{j=1}^{C\frac{n}{d}} j\left(\frac{(jp)^{\epsilon_1-2}}{d^{0.9}}\right)^j .$$

Denote the j-th summand of the last sum by s_j. Then, if $j \leq (n/d)^{1/2}$, we get $s_j \leq (d/n)^{(\epsilon_1-3)/2} = o(d/n)$, while if $j \geq (n/d)^{1/2}$, then $s_j \leq (Cn/d)(C^{\epsilon_1-2}/d^{0.9})^{(n/d)^{1/2}} = o(d/n)$, thus showing that the last sum tends to 0 as d tends to infinity. $\Box$

3 Proof of the lower bound

To prove the lower bound of Theorem 1.1, we argue *deterministically* that every bipartite graph $G = (A \cup B, E)$ with $|A| = |B| = n$, satisfying the assertion of Proposition 2.1 for some value of the parameter d, satisfies also $ch(G) \geq \log d - 4 \log\log d$. Again, we assume the parameter d (and hence n) to be large enough.

Let $t(k)$ denote the minimal number of edges in a k-uniform non-2-colorable hypergraph H. A tight connection between the problem of determining $t(k)$ and choosability questions for bipartite graphs has been exposed already in the original paper of Erdős, Rubin and Taylor [5]. As shown by Erdős [4], $t(k) \leq (1 + o(1))(e \ln 2/4)2^k k^2$. Given n and d, define an integer $k = k(n, d)$ by

$$k = \max\left\{ i : \frac{2in\log d}{d} \leq \left\lfloor \frac{n}{t(i)} \right\rfloor \right\} .$$

It is easy to see that $k \geq \log d - (1 + o(1))3 \log\log d > \log d - 4 \log\log d$. Therefore, to prove the desired lower bound on $ch(G)$ it suffices to prove that $ch(G) > k$.

Let us denote $t = t(k)$ to simplify the notation. Given a bipartite graph $G = (A \cup B, E)$, we partition the color class A into t color classes $A = A_1 \cup \ldots \cup A_t$ so that $|A_i| \geq \lfloor n/t \rfloor$, $1 \leq i \leq t$. Similarly, we partition $B = B_1 \cup \ldots \cup B_t$ with $|B_i| \geq \lfloor n/t \rfloor$. Let $H = (V, F)$ be a k-uniform non-2-colorable hypergraph with t edges. Denote $F = \{S_1, \ldots, S_t\}$. We view the vertices of $V(H)$ as colors, while the subsets S_i will be assigned to vertices of G as color lists. For each $1 \leq i \leq t$, every vertex $v \in A_i \cup B_i$ gets S_i as its list of colors. We claim that the assertion of Proposition 2.1 along with the non-2-colorability of H guarantee that G cannot be properly colored by assigning each vertex a color from its list. Indeed, let $f : A \cup B \to V(H)$ satisfy $f(v) \in S(v)$ for every $v \in V(G)$. Let C_1 be the subset of $V(H)$, formed by all the colors chosen by f at least $2n \log d/d$ times on A. For every $1 \leq i \leq t$, the list S_i of cardinality $|S_i| = k$ is assigned to all vertices $a \in A_i$. As $|A_i| \geq \lfloor n/t \rfloor \geq 2kn \log d/d$, we get that at least one color from S_i is chosen at least $2n \log d/d$ times on A_i. Thus, C_1 intersects every edge S_i of H. Now, as H is non-2-colorable, there exists an edge $S_{i_0} \in F$, satisfying $S_{i_0} \subseteq C_1$ (otherwise the partition $V = C_1 \cup (V \setminus C_1)$ forms a 2-coloring of H). The color list S_{i_0} is assigned to all vertices $b \in B_{i_0}$, hence there exists a color $c^* \in S_{i_0}$, chosen by f at least $|B_{i_0}/|S_{i_0}| \geq \lfloor n/t \rfloor/k \geq 2n \log d/d$ times on B_{i_0}. Let $X = \{a \in A : f(a) = c^*\}$, $Y = \{b \in B : f(b) = c^*\}$. As $c^* \in S_{i_0} \subseteq C_1$, we have $|X| \geq 2n \log d/d$. Also, $|Y| \geq 2n \log d/d$. But then according to the claim of Proposition 2.1, there exists an edge of G, connecting X and Y, thus showing that f does not form a proper coloring of G. This concludes the proof of the lower bound of Theorem 1.1. □

4 Proof of the upper bound

The proof here is also deterministic. We assume that a bipartite graph $G = (A \cup B, E)$ has the property given by Proposition 2.2 and show that every graph having this property has its choice number bounded from above by the upper bound of Theorem 1.1. Let us introduce the following notation. Define

$$\begin{aligned}\epsilon_1 = \epsilon_1(d) &= \frac{3\log d}{\log\log d}\ ,\\ \epsilon_2 = \epsilon_2(d) &= \frac{4\log d}{\log\log d}\\ \epsilon_3 = \epsilon_3(d) &= \frac{5\log d \log\log\log d}{\log\log d}\ ,\\ l = l(d) &= \lfloor \log d + \epsilon_3 \rfloor\ .\end{aligned}$$

Given a family $\mathcal{S} = \{S(v) : v \in A \cup B\}$, satisfying $|S(v)| = l$ for every $v \in A \cup B$, our aim is to prove the existence of a choice function f. Denote by $S = \bigcup_{v \in A\cup B} S(v)$ the union of all colors in all lists. Partition S randomly into two parts S_A and S_B by putting each color $c \in S$ into S_A or S_B independently and with probability $1/2$. Ideally, we would like to use colors from S_A to color vertices from A and those from S_B to color vertices from B. This strategy would eliminate any possible color conflict. The problem is that we cannot guarantee that every vertex from $A \cup B$ gets at least one eligible color under such a partition. However, we will be able to show that with positive probability the number of vertices that get only few colors is quite small.

We call a vertex $a \in A$ *poor* if $|S(a) \cap S_A| < \epsilon_2$. Similarly, a vertex $b \in B$ is poor if $|S(b) \cap S_B| < \epsilon_2$. Let T_0 denote the set of all poor vertices. The probability that $a \in A$ is poor is at most

$$\begin{aligned}\sum_{i=0}^{\epsilon_2}\binom{l}{i}2^{-l} &\le 2^{-l}\epsilon_2\binom{l}{\epsilon_2} \le \frac{1}{d}2^{-\epsilon_3}\epsilon_2\left(\frac{el}{\epsilon_2}\right)^{\epsilon_2}\\ &\le \frac{1}{d}2^{-\epsilon_3}\epsilon_2(\log\log d)^{\epsilon_2} = \frac{1}{d}2^{-\epsilon_3+\log\epsilon_2-\epsilon_2\log\log\log d} < \frac{1}{2d}\ .\end{aligned}$$

The same argument shows that for every $b \in B$, the probability that B is poor is less than $1/(2d)$. We conclude that with positive probability $|T_0| \le n/d$. Let us fix a partition $S = S_A \cup S_B$, for which indeed $|T_0| \le n/d$.

Now we find a (small) subset $T \subset A \cup B$, including T_0, such that every vertex of G outside T has less than ϵ_2 neighbors inside T. To find such a subset, we start with $T = T_0$,and as long as there exists a vertex $v \in (A \cup B) \setminus T$ having at least ϵ_2 neighbors in T, we add v to T. This process stops with $|T| \le 5n/d$ because otherwise we would get a subset T of size $|T| = 5n/d$ containing at least $(4n/d)\epsilon_2 = 16n\log d/(d\log\log d)$ edges, thus contradicting the assertion of Proposition 2.2.

We claim that we can find a proper coloring f by starting from choosing colors for vertices of T, and then by using colors from S_A to color vertices from $A \setminus T$ and colors from S_B to color those from $B \setminus T$. According to Proposition 2.2, the spanned subgraph $G[T]$ has a vertex of degree at most $2\epsilon_1$ in each of its subgraphs. This shows that $G[T]$ is $2\epsilon_1$-*degenerate* and thus $(2\epsilon_1 + 1)$-choosable (see, e.g., [2] for a discussion of the connection between degeneracy and choosability and for a very simple proof of the above statement). Thus we can use the original lists of colors $\{S(v) : v \in T\}$ to find colors for all vertices from T. Next, for every $a \in A \setminus T$ we delete from $S(a) \cap S_A$ the colors chosen for neighbors of a in T, and similarly, for every $b \in B \setminus T$ we delete from $S(b) \cap S_B$ the colors chosen for neighbors of b in T. As all poor vertices fall inside T and every vertex outside T has less than ϵ_2 neighbors inside T, even after this deletion each vertex in $(A \cup B) \setminus T$ still has at least one eligible color. We complete the choice of colors by choosing an arbitrary remaining color in $S(a) \cap S_A$ for each $a \in A$ and an arbitrary remaining color in $S(b) \cap S_B$ for each $b \in B$. This completes the proof of the upper bound of Theorem 1.1. □

5 Concluding remarks

- The proof in Section 3 shows that any bipartite graph with sufficiently strong expansion properties has a large choice number. More precisely, if G is a bipartite graph with n vertices in each of its two color classes A and B, and there is at least one edge between any subset of cardinality n/x of A and any subset of cardinality n/x of B, then the choice number of G is at least $(1 + o(1)) \log_2 x$, where the $o(1)$-term tends to 0 as x tends to infinity. By the known relation between the eigenvalues of the adjacency matrix of a graph and its expansion properties this implies, for example, that the choice number of any d-regular bipartite graph G in which the absolute value of every eigenvalue besides the largest and the smallest is at most λ, satisfies $ch(G) \geq (1 + o(1)) \log_2(d/\lambda)$. This is because in each such graph there is an edge between any two subsets of the two color classes provided each subset is of cardinality at least $n\lambda/d$.

- The arguments in Sections 3 and 4 can be extended to deal with the choice numbers of random r-partite graphs for $r \geq 2$. Let $G_r(n,p)$ denote the random r-partite graph obtained by taking r pairwise disjoint sets $A_1, A_2, \ldots, A_r$, each of cardinality n, and by connecting each pair of vertices in distinct sets A_i by an edge, randomly and independently, with probabiland every $p = p(n)$, the choice number of $G_r(n,p)$ is almost surely $(1 + o(1)) \log(np)/\log(r/(r-1))$, where the $o(1)$ term tends to 0 as np tends to infinity. The proof is similar to the one given here for the case $r = 2$. We omit the details. We note that it is known that the choice number of the complete r-partite graph with n vertices in each color class is $\Theta(r \log n)$ for all n and r, as proved in [1], and the choice number of the usual random graph $G(n, p(n))$ is almost surely $\Theta(np(n)/\log(np(n)))$ whenever $2 < np(n) < 9n/10$, as proved in [3].

- In [2] it is proved that the choice number of any graph with average degree at least d is at least $\Omega(\log d/\log\log d)$. It seems plausible that the $\log\log d$ term can be omitted, but at the moment this remains open. If true, this would, of course, be tight, up to a constant factor, (since, for example, $ch(K_{d,d}) = (1 + o(1)) \log_2 d$.)

  It is not difficult to prove that the choice number of any bipartite graph with maximum degree

d is at most $O(d/\log d)$, but we believe that the following much stronger result holds.

Conjecture 5.1 *The choice number of any bipartite graph with maximum degree d is at most $O(\log d)$.*

As far as we know it is even possible that the choice number of each such graph is at most $(1+o(1))\log_2 d$, where the $o(1)$ term tends to 0 as d tends to infinity. As a test case, it may be interesting to determine or estimate the choice number of the d-cube.

References

[1] N. Alon, Choice numbers of graphs: a probabilistic approach, Combinatorics, Probability and Computing 1 (1992), 107-114.

[2] N. Alon, Restricted colorings of graphs, in *Surveys in Combinatorics 1993*, London Math. Soc. Lecture Notes Series 187 (K. Walker, ed.), Cambridge Univ. Press, 1993, 1–33.

[3] N. Alon, M. Krivelevich and B. Sudakov, List coloring of random and pseudo-random graphs, to appear.

[4] P. Erdős, *On a combinatorial problem II*, Acta Math. Acad. Sci. Hungar. 15 (1964) 445-447.

[5] P. Erdős, A. L. Rubin and H. Taylor, Choosability in graphs, in *Proc. West Coast Conf. on Combinatorics, Graph Theory and Computing, Congressus Numerantium* XXVI, 1979, 125-157.

[6] V. G. Vizing, *Coloring the vertices of a graph in prescribed colors* (in Russian), Diskret. Analiz. No. 29, Metody Diskret. Anal. v. Teorii Kodov i Shem 101 (1976), 3-10.