The choice number of dense random graphs

Michael Krivelevich *

Abstract

We prove that if the edge probability p(n) satisfies $n^{-1/4+\epsilon} \le p(n) \le 3/4$, where $0 < \epsilon < 1/4$ is a constant, then the choice number and the chromatic number of the random graph G(n,p) are almost surely asymptotically equal.

1 Introduction

A coloring of a graph G is an assignment of a color to each of its vertices so that adjacent vertices get different colors. The chromatic number $\chi(G)$ of G is the minimal possible number of colors used in its coloring. If $\chi(G) \leq k$, we say that G is k-colorable.

A closely related but much more complicated quantity is the *choice number* of G, introduced by Vizing [12] and independently by Erdős, Rubin and Taylor [9]. Given a graph G = (V, E) with $V = \{v_1, \ldots, v_n\}$ and a family of color lists $S = \{S_1, \ldots, S_n\}$, where $S_i \subset Z$, we say that G is S-choosable, if there exists a choice function $f: V \to Z$ such that $f(v_i) \in S_i$ for $1 \le i \le n$, and also $f(v_i) \ne f(v_j)$ for every edge $(v_i, v_j) \in E(G)$. Next, for a positive integer k, G is called k-choosable if G is S-choosable for any family $S = \{S_1, \ldots, S_n\}$, satisfying $|S_i| = k$ for $1 \le i \le n$. Finally, the choice number ch(G) of G is defined as the minimal value of k, for which G is k-choosable. We address the reader to an excellent survey of Alon [2], where a wealth of different results and approaches to choosability problems is presented.

In this paper we consider an asymptotic behavior of the choice number of random graphs. As usually, G(n,p) denotes a finite probability space, whose points are graphs on n labeled vertices $\{1,\ldots,n\}$, where every pair of vertices forms an edge randomly and independently with probability p=p(n). We say that G(n,p) has property A almost surely (abbreviated by a.s.), if the probability that it satisfies A tends to 1 as n tends to infinity.

Coloring properties of random graphs have been attracted a great deal of attention during the last twenty five years. Culminating a series of efforts of various researchers, results of Bollobás [6] and Luczak [10] provided an asymptotic formula for the chromatic number of the random graph G(n, p)

^{*}School of Mathematics, Institute for Advanced Study, Princeton, NJ 08540. Email: mkrivel@math.ias.edu. Research supported by an IAS/DIMACS Postdoctoral Fellowship. Current address: DIMACS Center, Rutgers University, Piscataway, NJ; e-mail: mkrivele@dimacs.rutgers.edu.

for all p = p(n) satisfying $C/n \le p(n) \le 9/10$, where C > 0 is a sufficiently large constant. Although many interesting questions about the chromatic number of random graphs still remain unsolved, the main problem in this direction can be regarded as settled.

Much less is known about the asymptotic value of the choice number of G(n, p). Already in their original paper [9] Erdős, Rubin and Taylor addressed this question and conjectured, in particular, that almost surely ch(G(n, 1/2)) = o(n). This was proved by Alon in [1]. Kahn (unpublished) succeeded to prove that a.s. $ch(G(n, 1/2)) = (1 + o(1))n/2 \log_2 n$, that is, the choice number and the chromatic number of G(n, 1/2) have the same asymptotic value. Later, Alon (see [2]) found a simpler proof of the same result. It should be noted that his proof is immediately extendable to all constant edge probabilities p.

Quite recently, further results about choosability in random graphs have been obtained by Alon, Krivelevich and Sudakov [3] and independently by Vu [13]. Alon et al. showed that there exist constants $c_1, c_2 > 0$ such that for all p = p(n) satisfying $2 \le np \le n/2$, a.s. $c_1\chi(G(n, p)) \le ch(G(n, p)) \le c_2\chi(G(n, p))$, that is, for all values of p(n) in this region the choice number and the chromatic number have a.s. the same order of magnitude. Vu obtained the same result for $p(n) \ge \log^{1+\delta} n/n$, where $\delta > 0$ is a constant. Both papers give also upper bounds for the choice number of pseudo-random graphs.

The main aim of this note is to extend Alon-Kahn's result for smaller values of p. We prove the following.

Theorem 1.1 Let $0 < \epsilon < 1/4$ be a constant. If the edge probability p(n) satisfies $n^{-1/4+\epsilon} \le p(n) \le 3/4$, then a.s. $ch(G(n,p)) = (1+o(1))\chi(G(n,p))$.

The rest of the paper is organized as follows. In the next section we present properties of random graphs to be used later in the proof. Section 3 contains a proof of Theorem 1.1. The final section is devoted to concluding remarks and a discussion of related open problems.

Throughout the paper we assume, whenever this is needed, that the number of vertices n is sufficiently large. We also omit routinely floor and ceiling sings for the sake of simplicity of presentation. As usually, we denote by d = (n-1)p the expected vertex degree in G(n, p).

2 Preliminaries

This section is aimed to supply necessary technical tools for the proof of our main theorem. Properties of random graphs to be described in this section fall roughly into two categories: one is about the distribution of almost optimal independent sets in G(n, p), while the other addresses the distribution of edges. It is worth recalling here that we assume that the edge probability p(n) satisfies $n^{-1/4+\epsilon} \le p(n) \le 3/4$.

The first proposition of this section shows that a.s. every large enough subset of vertices in G(n, p) contains an independent set of almost optimal size. Since an essentially identical lemma

played a crucial role in Bollobás' proof for the chromatic number of G(n, 1/2) and is probably quite well known by now, together with several different proofs (see, e.g., a survey paper of Spencer [11]), we present its proof in a somewhat sketchy form.

Let $k_0 = k_0(n, p)$ be defined by

$$k_0 = \min \left\{ k : inom{rac{n}{\ln^4 n}}{k} (1-p)^{inom{k}{2}} \geq n^3
ight\} \; .$$

For the purposes of the subsequent analysis notice that $k_0 = (1 - o(1))2 \ln d/p$ when p(n) = o(1) and $k_0 = (1 - o(1))2 \ln n / \ln(1/(1-p))$ when p(n) is a constant. Therefore, it follows from the results of Bollobás [6] and Luczak [10] that a.s. $\chi(G(n,p)) = (1 + o(1))n/k_0$.

Proposition 2.1 A.s. in G(n,p) every subset $U \subset V$ of size $|U| \geq n/\ln^4 n$ spans an independent set of size k_0 .

Proof. Denote $m = n/\ln^4 n$ and consider a random graph G(m,p). Let X be the random variable counting the number of independent sets of size k_0 in G(m,p). Denoting the expectation of X by μ and recalling the definition of k_0 , we get:

$$\mu=inom{m}{k_0}(1-p)^{inom{k_0}{2}}\geq n^3$$
 .

Let

$$\Delta=2\sum_{\stackrel{|S|,|S'|=k_0}{2\leq |S\cap S'|\leq k_0-1}} Pr[S,S' ext{ form an independent set in } G(m,p)]$$
 .

Then

$$\Delta = \binom{m}{k_0} (1-p)^{\binom{k_0}{2}} \sum_{i=2}^{k_0-1} \binom{k_0}{i} \binom{m-k_0}{k_0-i} (1-p)^{\binom{k_0}{2}-\binom{i}{2}}$$

$$= \mu^2 \sum_{i=2}^{k_0-1} \frac{\binom{k_0}{i} \binom{m-k_0}{k_0-i} (1-p)^{-\binom{i}{2}}}{\binom{m}{k_0}} = \mu^2 \sum_{i=2}^{k_0-1} g(i) .$$

One can check that $g(2) = \Theta(k_0^4/m^2)$ is the dominating term in the above sum, while the summands decrease quickly as i goes away from the ends of the feasible interval. This implies that $\Delta = \Theta(\mu^2 k_0^4/m^2)$. Also, as $\mu \geq n^3$ we get $\Delta \geq \mu$. Then by the generalized Janson inequality (see, e.g., [4], Ch. 7)

$$Pr[X=0] \leq e^{-\frac{\mu^2(1+o(1))}{2\Delta}} = e^{-\Omega\left(\frac{m^2}{k_0^4}\right)} \leq e^{-n^{1+4\epsilon-o(1)}}$$

(the assumption $p(n) \geq n^{-1/4+\epsilon}$ is used to derive the last inequality). Hence the probability that there exists a subset of size m in G(n, p), that does not span an independent set of size k_0 , is at most $\binom{n}{m} Pr[X=0] \leq 2^n e^{-n^{1+\epsilon-o(1)}} = o(1)$.

The next two propositions assert that the local edge distribution in G(n, p) a.s. does not deviate much from its expected behavior. Somewhat strangely looking expressions in the formulations of these propositions should not puzzle the reader as they are formulated so as to be plugged in directly in the proof of Theorem 1.1.

Proposition 2.2 A.s. every $s \leq 4n^2 \ln^7 d/d^2$ vertices of G(n,p) span less than $(6n \ln^7 d/d)s$ edges.

Proof. Define $r = 6n \ln^7 d/d$. The probability that there exists a subset $V_0 \subset V$ violating the assertion of the proposition is at most

$$\sum_{i=r}^{4n^2\ln^7 d/d^2} \binom{n}{i} \binom{\binom{i}{2}}{ri} p^{ri} \leq \sum_{i=r}^{4n^2\ln^7 d/d^2} \left[\frac{en}{i} \left(\frac{ei}{2r} \right)^r p^r \right]^i \leq \sum_{i=r}^{4n^2\ln^7 d/d^2} \left[\frac{e^2np}{2r} \left(\frac{eip}{2r} \right)^{r-1} \right]^i$$

$$\leq \sum_{i=r}^{4n^2\ln^7 d/d^2} \left[O(1) \frac{d^2}{n\ln^7 d} \left(\frac{ei}{\frac{12n^2\ln^7 d}{d^2}} \right)^{\frac{6n\ln^7 d}{d} - 1} \right]^i \leq \sum_{i=r}^{4n^2\ln^7 d/d} (n \cdot 0.95^{\frac{6n\ln^7 d}{d} - 1})^i = o(1) . \quad \Box$$

Before proceeding further, let us indicate why the above proposition is relevant to choosability questions. For an integer d, a graph G is called d-degenerate if every subgraph of it contains a vertex of degree at most d. The following is a well known fact (c.f., e.g., [2]), easily proven by induction.

Claim 2.3 Every d-degenerate graph is (d+1)-choosable.

Returning back to Proposition 2.2, we see that a.s. every subgraph of G(n,p), spanned by a subset $V_0 \subset V$ of size $|V_0| \leq 4n^2 \ln^7 d/d^2$, has a vertex of degree less than $12n \ln^7 d/d$. Therefore we get

Corollary 2.4 A.s. every subgraph of G(n,p), spanned by a subset $V_0 \subset V$ of size $|V_0| \leq 4n^2 \ln^7 d/d^2$, is $12n \ln^7 d/d$ -choosable.

Proposition 2.5 A.s. for every subset $U_0 \subset V$ of size $|U_0| = n/\ln^4 n$ the subgraph $G[U_0]$ has less than $n \ln^3 d/d$ vertices of degree at least $d/\ln^3 d$.

Proof. Denote $t = n \ln^3 d/d$. Then the probability of existence of a subset U_0 violating the proposition can be bounded from above by

$$egin{pmatrix} n \ rac{n}{\ln^4 n} \end{pmatrix} egin{pmatrix} rac{n}{\ln^4 n} \ t \end{pmatrix} egin{pmatrix} rac{n}{\ln^3 n} - t \ t \ rac{dt}{\ln^3 d} - 2inom{t}{2} \end{pmatrix} p^{rac{dt}{\ln^3 d} - 2inom{t}{2}}$$

(first choose a subset U_0 , then choose t vertices of high degree in the subgraph $G[U_0]$, and then require that at least $dt/\ln^3 d - 2\binom{t}{2}$ edges will go from the vertices of high degree to the rest of U_0). The above expression is at most

$$2^n \ 2^n \ \left(rac{O(1) rac{nt}{\ln^4 n}}{rac{dt}{\ln^3 d}} \cdot p
ight)^{(1-o(1))dt/\ln^3 d} = 4^n \left(rac{O(1)}{\ln n}
ight)^{(1-o(1))n} = o(1) \ . \hspace{1cm} \square$$

3 The proof of the main result

Having done with all the necessary technical preparations, we are now ready to prove Theorem 1.1. First, note that it follows trivially from the definition of the choice number that $ch(G) \geq \chi(G)$ for any graph G. Therefore, we need only to prove the upper bound for ch(G(n, p)). The desired upper bound is a direct consequence of the following deterministic statement.

Proposition 3.1 Let G be a graph on n vertices, satisfying all properties in the assertions of Propositions 2.1, 2.2 and 2.5. Then $ch(G) \leq n/k_0 + d/\ln^2 d$.

The last thing to notice before going into the proof of the above proposition is that according to the definition of k_0 we have $k_0 = \Theta(n \ln d/d)$. Therefore we will prove $ch(G) = (1 + o(1))n/k_0$ and thus indeed a.s. $ch(G(n,p)) = (1 + o(1))\chi(G(n,p))$.

Proof. Given a family of color lists $S = \{S_1, \ldots, S_n\}$ with $|S_i| = n/k_0 + d/\ln^2 d$, $i = 1, \ldots, n$, we need to show that G is S-choosable.

Our coloring procedure consists of two phases. The first phase is in a sense identical to the first phase of Alon's argument. As long as there exists a color c which appears in the lists of at least $n/\ln^4 n$ of yet uncolored vertices, we do the following. Denote by V_0 the set of those uncolored vertices whose color list contains c, then $|V_0| \ge n/\ln^4 n$. According to Proposition 2.1, V_0 spans an independent set I of size $|I| = k_0$. We color all vertices of I by c, discard I and delete c from all lists. The total number of deleted colors during the first phase cannot exceed n/k_0 , as each time we remove a subset of size k_0 .

Let U denote the set of all vertices that are still uncolored after the first phase has been completed. if $U=\emptyset$ we are done, therefore we may assume that the set U is non-empty. The lists of all vertices of U are still quite large, namely, $|S(v)| \ge d/\ln^2 d$ for each $v \in U$. Also, each color c appears in at most $n/\ln^4 n$ lists of vertices from U.

For a vertex $v \in U$ and a color c, we say that the pair (v,c) is dangerous if $c \in S(v)$ and also c belongs to the lists of at least $d/\ln^3 d$ neighbors of v in U. The reason for this name is quite simple – using c to color v forces c to be deleted from lists of too many vertices. We would like to show that the number of dangerous pairs cannot be too large. To this end, for a color c denote by W(c) the set of all vertices $u \in U$ for which c is included in the corresponding list of colors S(u). We know that $|W(c)| \leq n/\ln^4 n$ for each color c. If a pair (v,c) is dangerous, this means exactly that the degree of v in the spanned subgraph G[W(c)] is at least $d/\ln^3 d$. Therefore, if c participates in at least t dangerous pairs, the subgraph G[W(c)] contains at least t vertices of degree at least $d/\ln^3 d$. Applying Proposition 2.5, we derive that the number of dangerous pairs for c does not exceed $n \ln^3 d/d$. Another observation is that if a color c participates in a dangerous pair, then $|W(c)| \geq d/\ln^3 d$. Since $\sum_{v \in U} |S(v)| \leq |U|(n/k_0 + d/\ln^2 d) \leq nd/\ln d$ (here we use the assumption $p \leq 3/4$), we have $|\{c: |W(c) \geq d/\ln^3 d|\} \leq (nd/\ln d)/(d/\ln^3 d) = n\ln^2 d$. Putting these two observations together, we see that the total number of dangerous pairs is at most $(n \ln^2 d) \cdot (n \ln^3 d/d) = n^2 \ln^5 d/d$.

Ideally we would like not to use color c for coloring v, whenever (v,c) forms a dangerous pair. For some vertices this restriction may be too severe, as many colors in their lists form a dangerous pair. We can argue, however, that the number of such vertices cannot be too large. Let T_0 be the set of all vertices in U, which participate in at least $d/2 \ln^2 d$ dangerous pairs. Using previously obtained bound on the number of dangerous pairs, we get $|T_0| \leq (n^2 \ln^5 d/d)/(d/2 \ln^2 d) = 2n^2 \ln^7 d/d^2$.

Next, we find a subset $T \subset U$ of size $|T| = O(n^2 \ln^7 d/d^2)$, including T_0 and such that every vertex $v \in U \setminus T$ has a small number of neighbors inside T. We start with $T = T_0$ and as long as there exists a vertex $v \in U \setminus T$ such that v has more than $12n \ln^7 d/d$ neighbors inside T, we add v to T. This process is bound to stop with $|T| \leq 4n^2 \ln^7 d/d^2$, because otherwise we would get a subset T of size $|T| = 4n^2 \ln^7 d/d^2$ containing more than $(6n \ln^7 d/d)|T|$ edges, thus contradicting the assertion of Proposition 2.2.

Since the size of T falls within the range of Corollary 2.4, we can choose colors for vertices of T from their lists. Having chosen colors for vertices of T, for every $v \in U \setminus T$ we delete from S(v) all those colors which are used to color the neighbors of v in T (at most $12n \ln^7 d/d$ colors are deleted from S(v)). We also delete from each S(v) all colors c, with whom v forms a dangerous pair (at most $d/2 \ln^2 d$ colors, according to the definition of T_0). Even after these two deletions, the lists of yet uncolored vertices are long enough $-|S(v)| \geq d/\ln^2 d - d/2 \ln^2 d - 12n^2 \ln^7 d/d \geq d/3 \ln^2 d$. Now all dangerous pairs have been destroyed, which means that for every vertex v and color c, satisfying $c \in S(v)$, c appears in lists of at most $d/\ln^3 d$ neighbors of v. We want to show that $U \setminus T$ is colorable from the remaining lists.

Let $U \setminus T = \{v_1, \ldots, v_s\}$. Define an auxiliary s-partite graph $H = (A_1 \cup \ldots \cup A_s, F)$. Each A_i has cardinality $d/3 \ln^2 d$ and its vertices are labeled by colors from $S(v_i)$ (if $|S(v_i)| > d/3 \ln^2 d$, choose $d/3 \ln^2 d$ colors from $S(v_i)$ arbitrarily). Two vertices $a_1 \in A_i$, $a_2 \in A_j$ $(i \neq j)$ are connected by an edge in H, if they are labeled by the same color and $(v_i, v_j) \in E(G)$. A subset $R \subset V(H)$ is called a transversal in H, if $|R \cap A_i| = 1$ for $1 \leq i \leq s$. A transversal R is independent if it forms an independent set in H. Note that an independent transversal in H corresponds directly to a proper choice of colors for all $v \in U \setminus T$. Therefore, our aim is to show that H has an independent transversal. We achieve this goal by applying the Lovász Local Lemma [8], quite similarly to the paper of Erdős, Gyárfás and Łuczak [7]. Pick a transversal R in H uniformly at random from all transversals of H. For an edge $f \in F(H)$ with endpoints in A_i and A_j , let B_f be the event that both endpoints of f belong to R. Clearly, $Pr[B_f] = 1/(|A_i||A_j|) = (3\ln^2 d/d)^2$. Notice that B_f is independent of all events $B_{f'}$, but those for which f' has at least one of its endpoints in $A_i \cup A_j$. The degree of every vertex $a \in V(H)$ is at most $d/\ln^3 d$. Therefore the dependency graph of the events B_f has maximal degree at most $2(d/3 \ln^2 d)(d/\ln^3 d) = 2d^2/\ln^5 d$. Hence, by the symmetric version of the Local Lemma, $Pr[\bigcap_{f \in F(H)} \overline{B_f}] > 0$, which means that H has an independent transversal R. For every $v_i \in U \setminus T$ choose a color of the vertex from $A_i \cap R$. This finishes a coloring of G and the proof of our main theorem.

4 Concluding remarks

- The first issue we would like to comment on is the range of edge probabilities p = p(n), for which our main result (Theorem 1.1) is valid. Clearly, the constant 3/4 in the formulation of Theorem 1.1 can be replaced by any constant a < 1 without changing much in the presented proof. The lower bound on the edge probability $p(n) \ge n^{-1/4+\epsilon}$ deserves more attention. As the reader may have already seen from the proof, the only obstacle in extending the main result for smaller values of p(n) is that the proof of Proposition 2.1 stops working when $p(n) \le n^{-1/4}$. The generalized Janson inequality is the best suited out of the three standard by now large deviation tools (the other two being martingales in the spirit of Bollobás' proof [6] and the inequality of Talagrand) for the purpose of deriving upper bounds for the probability that a random graph G(m,p) does not contain an independent set of size k_0 (we use the notation of Proposition 2.1). However, even it ceases to provide the required bound (namely, $Pr[X=0] < {n \choose m}^{-1}$), when $p(n) \le n^{-1/4}$. Any improvement in the question about distribution of almost optimal independent sets will result immediately in an extension of our main result to smaller values of p(n). Some cosmetic changes still will need to be done, but the spirit of the proof will remain untouched.
- Our main result shows that for dense random graphs the chromatic number and the choice number have a.s. the same asymptotic value. There is no apparent reason for this phenomenon not to hold for all values of p(n). It is thus plausible to conjecture the following.

Conjecture 4.1 For all values of the edge probability p(n) a.s. $ch(G(n,p)) = (1+o(1))\chi(G(n,p))$.

In fact, one can try to prove a stronger result by showing that the difference between ch(G(n,p)) and $\chi(G(n,p))$ has much smaller order than, say, $\chi(G(n,p))$ (and thus both of these two parameters). This however may require completely different approaches than the one of this paper.

• The main idea of the present paper can be used to derive a new proof of the result of Alon, Krivelevich and Sudakov [3] and of Vu [13], asserting that a.s. $ch(G(n,p)) = O(\chi(G(n,p)))$. Indeed, in order to apply the approach of this paper one needs to show that a.s. every subset $U \subset V$ of size $|U| = n/\ln^4 d$ contains an independent set I of order $|I| = \Theta(n \ln d/d)$ (see Proposition 2.1). This can be done, for example, by bounding the total number of triangles in G(n,p), estimating the number of edges in each such U and then applying known lower bounds for the independence number of graphs with given number of vertices, average degree and number of triangles (see, e.g. [5], Ch. 12, Lemma 15). Of course, due to Proposition 2.1 we may assume that, say, $p(n) \leq n^{-1/5}$. We omit further details of the proof.

Acknowledgment. The author would like to thank Noga Alon for helpful discussions.

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, submitted.
- [4] N. Alon and J. H. Spencer, The probabilistic method, Wiley, New York, 1992.
- [5] B. Bollobás, Random graphs, Academic Press, New York, 1985.
- [6] B. Bollobás, The chromatic number of random graphs, Combinatorica 8 (1988), 49-55.
- [7] P. Erdős, A. Gyárfás and T. Łuczak, *Independent transversals in sparse partite hypergraphs*, Combinatorics, Probability and Computing 3 (1994), 293–296.
- [8] P. Erdős and L. Lovász, Problems and results on 3-chromatic hypergraphs and some related questions, in *Infinite and finite sets*, (A. Hajnal et al., eds.), North Holland, Amsterdam, 1975, 609–628.
- [9] 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.
- [10] T. Łuczak, The chromatic number of random graphs, Combinatorica 11 (1991), 45-54.
- [11] J. Spencer, Probabilistic methods in combinatorics, in *Proceedings Int. Congress of Math. Zürich*, 1994, 1375-1383.
- [12] 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.
- [13] V. H. Vu, On some simple degree conditions which guarantee the upper bound on the chromatic (choice) number of random graphs, submitted.