The concentration of the chromatic number of random graphs

Noga Alon * Michael Krivelevich †

Abstract

We prove that for every constant $\delta > 0$ the chromatic number of the random graph $G(n,p)$ with $p = n^{-1/2-\delta}$ is asymptotically almost surely concentrated in two consecutive values. This implies that for any $\beta < 1/2$ and any integer valued function $r(n) \leq O(n^{\beta})$ there exists a function $p(n)$ such that the chromatic number of $G(n,p(n))$ is precisely $r(n)$ asymptotically almost surely.

1 Introduction

Let $G(n,p)$ denote the random graph on n labeled vertices in which every edge is chosen randomly and independently with probability $p = p(n)$. We say that $G(n,p)$ has a property A *asymptotically almost surely* (a.a.s.) if the probability it satisfies A tends to 1 as n tends to infinity.

One of the most interesting early discoveries in the study of random graphs is that of the fact that many natural graph invariants are highly concentrated. One of the first striking results of this type was proved by Matula [9] and strengthened by various researchers; for fixed values of p almost all graphs $G(n,p)$ have the same *clique number*. The proof of this result is not difficult, and is based on the second moment method.

In this paper we study the concentration of the chromatic number of the random graph $G(n,p)$. This parameter is far more complicated than the clique number, and its asymptotic behavior is much less understood, despite the results of Bollobás [3] and Łuczak [7] that provide an asymptotic formula for its expectation. Shamir and Spencer [10] proved that there is always a choice of an interval $I = I(n,p)$ of length roughly $\sqrt{n}$, such that the chromatic number of $G(n,p)$ lies, asymptotically almost surely, in I. More surprisingly, they proved that for every constant $\alpha > 1/2$, if $p = n^{-\alpha}$ then the chromatic number of $G(n,p)$ is asymptotically almost surely concentrated in some fixed number of values. That is, there exists a function $t = t(n,p)$ and a constant $s = s(\alpha)$ which is at most the

*Department of Mathematics, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, 69978, Israel. Email: noga@math.tau.ac.il. Research supported in part by a USA Israeli BSF grant and by a grant from the Israel Science Foundation.

†Department of Mathematics, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, 69978, Israel. Email: krivelev@math.tau.ac.il. Research supported in part by a Charles Clore Fellowship.
Mathematics Subject Classification (1991): 05C80, 05C15

smallest integer strictly larger than $(2\alpha+1)/(2\alpha-1)$, such that a.a.s. $t \leq \chi(G(n,p)) \leq t+s$. A further step in this direction was made by Łuczak [8] who showed that if $\alpha > 5/6$, then $\chi(G(n,p))$ is a.a.s. *two point* concentrated. It is not difficult to see that the two point width of the concentration interval is best possible for a general p. Additional results on this problem were given in [6].

Here we extend the two point concentration result of Łuczak by proving the following result which shows that the bound of [10] for $s(\alpha)$ mentioned above can, in fact, be improved to 1 for all $\alpha > 1/2$.

Theorem 1.1 *For every positive constants ϵ, δ there exists an integer $n_0 = n_0(\epsilon, \delta)$ such that for every $n > n_0$ and probability $p = n^{-1/2-\delta}$ there is an integer $t = t(n,p,\epsilon)$ such that*

$$Pr[t \leq \chi(G(n,p)) \leq t+1] \geq 1-\epsilon \ .$$

In other words, for every constant $\alpha > 1/2$ the chromatic number of $G(n,p)$ with $p = n^{-\alpha}$ takes a.a.s. one of two consecutive values.

The above result and its proof imply the following.

Proposition 1.2 *For every fixed $\beta < 1/2, \epsilon > 0$ and every integer valued function $r(n)$ satisfying $1 \leq r(n) \leq n^{\beta}$, there exists an n_0 and a function $p(n)$ such that the chromatic number of $G(n,p(n))$ is precisely $r(n)$ with probability at least $1-\epsilon$ for all $n > n_0$.*

Therefore, for such values of $p(n)$, almost all graphs $G(n,p(n))$ have the same chromatic number !

Our proof uses a martingale approach much in the spirit of the papers of Shamir and Spencer [10] and of Łuczak [8], combined with additional probabilistic and combinatorial arguments. The presentation of the basic ideas follows closely that of [2], Chapter 7.

The rest of this paper is organized as follows. In the next section we prove several technical lemmas required for the proof of the main result. This proof is presented in Section 3. The final Section 4 contains a discussion of some related questions and open problems.

2 Preliminaries

The proof of Theorem 1.1 requires several preparations. We assume, whenever this is needed, that the number of vertices n is sufficiently large. Relying on the result of Łuczak, we may assume that, say, $\delta \leq 3/8$. Denote $d = np = n^{1/2-\delta}$. In the course of the proof floor and ceiling signs are occasionally omitted for the sake of convenience.

In the proof we apply some simple properties of the concept of k-choosability (see, e.g., [1], or [5], pp. 18–21). A graph $G = (V,E)$ is called *k-choosable* if for every family of lists $\{S(v) \subseteq \mathbf{Z}, |S(v)| = k; v \in V\}$ there exists a proper vertex-coloring $f : V \to \mathbf{Z}$ of G such that $f(v) \in S(v)$ for every $v \in V$. Clearly, the k-choosability of a graph G implies its k-colorability, but the converse is not true in general. A graph is *d-degenerate* if every subgraph of it contains a vertex of degree at most d. The following is a simple, well known fact (c.f., e.g., [1]):

Proposition 2.1 *Every d-degenerate graph is $(d+1)$-choosable.* $\square$

In the proof of our main result we need the following simple though somewhat technical lemma.

Lemma 2.2

- *(i) For every $\delta > 0$ there exists a constant $r = r(\delta) > 0$ such that for every $C > 0$, a.a.s. every $i \le C\sqrt{n}$ vertices of the random graph $G(n,p)$ with $p = n^{-1/2-\delta}$ span less than ri edges. Therefore, any subgraph of this graph induced by a subset $V_0 \subset V$ of size $|V_0| \le C\sqrt{n}$, is $2r$-choosable.*

- *(ii) Let $\delta > 1/5$. Then for every constant $C > 0$, a.a.s. every $i \le Cn^{1-\delta}$ vertices of the random graph $G(n,p)$ with $p = n^{-1/2-\delta}$ span less than $in^{1/10}$ edges.*

- *(iii) The random graph $G(n,p)$ with $p = n^{-1/2-\delta}$, $0 < \delta \le 3/8$, has a.a.s. the following properties:*

  1. *Every vertex $v \in V(G)$ has degree at most $3d = 3np$;*
  2. *$\chi(G) \ge d/2\ln n$;*
  3. *If $\delta \ge 1/6$, then the number of paths of length three (edges) between any two (not necessarily distinct) vertices of G is at most $\ln n$; if $0 < \delta < 1/6$, then the number of paths of length three between any two vertices of G is at most $d^3 \ln n/n$.*

Proof.
(i) Fix $r = \left\lceil \frac{1}{\delta} \right\rceil$. Then the probability of existence of a subset $V_0 \subset V$ violating the assertion of the lemma is at most

$$\begin{aligned}
& \sum_{i=r}^{C\sqrt{n}} \binom{n}{i} \binom{\binom{i}{2}}{ri} p^{ri} \\
\le & \sum_{i=r}^{C\sqrt{n}} \left[O(1) \frac{n}{i} i^r n^{\left(-\frac{1}{2}-\delta\right)r} \right]^i \\
\le & \sum_{i=r}^{C\sqrt{n}} \left[O(1)^r n^{1+\frac{1}{2}(r-1)-\left(\frac{1}{2}+\delta\right)r} \right]^i \\
= & \sum_{i=r}^{C\sqrt{n}} \left[O(1)^r n^{\frac{1}{2}-\delta r} \right]^i = o(1) \ .
\end{aligned}$$

The additional claim about the choosability now follows from Proposition 2.1.
(ii) The probability of existence of a subset $V_0 \subset V$ violating the claim of the lemma is at most

$$\sum_{i=2}^{Cn^{1-\delta}} \binom{n}{i} \binom{\binom{i}{2}}{in^{\frac{1}{10}}} p^{in^{\frac{1}{10}}}$$

$$\leq \sum_{i=2}^{Cn^{1-\delta}} \left[O(1)\frac{n}{i}\left(\frac{O(1)i}{n^{\frac{1}{10}}}\right)^{n^{\frac{1}{10}}} p^{n^{\frac{1}{10}}} \right]^i$$

$$= \sum_{i=2}^{Cn^{1-\delta}} \left[O(1)\frac{n}{i}\left(\frac{O(1)ip}{n^{\frac{1}{10}}}\right)^{n^{\frac{1}{10}}} \right]^i$$

$$\leq \sum_{i=2}^{Cn^{1-\delta}} \left[O(1)n\left(O(1)n^{\frac{2}{5}-2\delta}\right)^{n^{\frac{1}{10}}} \right]^i = o(1)\ .$$

(iii) 1. Indeed,

$$Pr[\exists v : d(v) > 3d] \leq n\binom{n-1}{3d}p^{3d}$$

$$\leq n\left(\frac{en}{3d}\right)^{3d} p^{3d} = n\left(\frac{en}{3d}\frac{d}{n}\right)^{3d} = o(1).$$

(iii) 2. The probability that G has an independent set of size $s = 2n\ln n/d$ is at most

$$\binom{n}{s}(1-p)^{\binom{s}{2}} \leq \left[O(1)\frac{d}{\ln n}e^{-p\frac{n\ln n}{d}}\right]^s$$

$$= \left[O(1)\frac{d}{\ln n}\frac{1}{n}\right]^{\frac{2n\ln n}{d}} = o(1)\ .$$

Therefore the chromatic number of G is a.a.s. at least $n/(2n\ln n/d) = d/2\ln n$;
(iii) 3. Let us first show that for every $\delta > 0$ a.a.s. any two vertices of the random graph $G(n,p)$ with $p = n^{-1/2-\delta}$ are connected by less than $c_0 = \lceil 2/\delta \rceil$ paths of length two. The probability that this is not so can be bounded from above by

$$\binom{n}{2}\binom{n-2}{c_0}p^{2c_0} \leq n^2[O(1)np^2]^{c_0} = n^2[O(1)n^{-2\delta}]^{c_0} = o(1)\ .$$

Therefore a.a.s. every edge $e = (w_1, w_2)$ of $G(n,p)$ participates in at most c_0 paths of length three between any two vertices $u, v \in V(G)$. Indeed, if $\{w_1, w_2\} \cap \{u, v\} = \emptyset$, then there are only two potential paths of length three from u to v, containing (w_1, w_2), i.e., the paths uw_1w_2v and uw_2w_1v. If, say, $w_1 = u$, $w_2 \neq v$, then every path of length three from u to v starting with e corresponds to a path of length two between w_2 and v, and the number of such paths is a.a.s. bounded from above by c_0. We conclude that a.a.s. every path of length three between any pair of vertices u, v has an edge in common with at most $3c_0$ other such paths.

Now, let $X_{u,v}$ be the number of paths of length three between u and v, then a.a.s. the number of edge disjoint paths of length three between u and v is at least $X_{u,v}/(3c_0 + 1)$. (This can be seen by defining an auxiliary graph A whose vertices correspond to the paths of length three between u and v and whose edges connect paths sharing an edge in $G(n,p)$. This graph has maximum degree at most

$3c_0$ and therefore is $(3c_0+1)$-colorable, thus it has an independent set of size at least $X_{u,v}/(3c_0+1)$). Hence we get that $Pr[X_{u,v} \geq l]$ is asymptotically at most the probability that the number of edge disjoint paths of length three between u and v is at least $l_0 = l/(3c_0+1)$. The latter probability is at most

$$\binom{n}{l_0}^2 l_0! p^{3l_0} \tag{1}$$

(first choose l_0 neighbors of u and l_0 neighbors of v, then fix a bijection between the vertices of the chosen sets, and then require all $3l_0$ edges of the chosen paths to be present in $G(n,p)$).

The probability in (1) is at most

$$\left[O(1)\frac{n^2}{l^2} l p^3\right]^{l_0} = \left[O(1)\frac{n^2p^3}{l}\right]^{\frac{l}{3c_0+1}} .$$

Taking $l = \ln n$ for the case $n^2p^3 \leq 1$ ($\delta \geq 1/6$) or $l = n^2p^3 \ln n = d^3 \ln n/n$ for the case $n^2p^3 \geq 1$ ($\delta \leq 1/6$), we get $Pr[X_{u,v} \geq l] = o(n^{-2})$. $\Box$

Our final preliminary lemma utilizes the idea used in the paper [8] of Łuczak (who attributes it to Frieze).

Lemma 2.3 *For every $\epsilon_0 > 0$ define $t = t(n,p,\epsilon_0)$ to be the least integer for which*

$$Pr[\chi(G) \leq t] \geq \epsilon_0. \tag{2}$$

Let X be the random variable whose value is the minimum number of vertices that have to be deleted from $V(G)$ to get a t-colorable graph. Let, further, λ be defined by $e^{-\lambda^2/2} = \epsilon_0$, then

$$Pr[X \geq 2\lambda\sqrt{n}] < \epsilon_0 .$$

Proof.

By the definition of t

$$Pr[\chi(G) < t] < \epsilon_0 . \tag{3}$$

It is easy to see that the random variable X satisfies the vertex Lipschitz condition, that is, if two graphs G and G' differ from each other only in edges containing some fixed vertex v, then $|X(G) - X(G')| \leq 1$. Therefore by considering the vertex exposure martingale on $G(n,p)$ as in, e.g., [2], Chapter 7, and by letting $\mu = EX$, we conclude that for every $\lambda > 0$

$$Pr[X \leq \mu - \lambda\sqrt{n}] < e^{-\frac{\lambda^2}{2}} , \quad Pr[X \geq \mu + \lambda\sqrt{n}] < e^{-\frac{\lambda^2}{2}} .$$

In particular, since for our choice of λ, $e^{-\lambda^2/2} = \epsilon_0$, it follows that these tail events both have probability less than ϵ_0. On the other hand, $Pr[X = 0] \geq \epsilon_0$, hence we derive from the first inequality that $\mu \leq \lambda\sqrt{n}$. Therefore, by the second inequality,

$$Pr[X \geq 2\lambda\sqrt{n}] < \epsilon_0 . \ \Box$$

Denoting $c = 2\lambda$, we summarize the above arguments as follows: with probability at least $1 - \epsilon$ the random graph $G = G(n,p) = (V,E)$ has all the properties stated in Lemma 2.2, satisfies $\chi(G) \geq t$ and also contains a subset $U_0 \subset V$ of size $|U_0| \leq c\sqrt{n}$ such that $G[V \setminus U_0]$ is t-colorable. Note also that by Lemma 2.2, part (iii) 2, $t \geq d/2 \ln n$.

3 The proof of the main result

Having finished all the necessary preparations, we are now ready to complete the proof of Theorem 1.1. In view of the last paragraph of the previous section it suffices to prove the following *deterministic* statement.

Proposition 3.1 *Let $G = (V,E)$ be a graph on n vertices satisfying all properties in the assertions of Lemma 2.2. Suppose, further, that $\chi(G) \geq t \geq d/2 \ln n$ and there is a subset $U_0 \subset V$ of size $|U_0| \leq c\sqrt{n}$ such that $G[V \setminus U_0]$ is t-colorable. Then G is $(t+1)$-colorable.*

We prove this (fully deterministic) proposition using probabilistic techniques.

As the first stage of the proof we find a subset $U \subset V$ of size $|U| = O(\sqrt{n})$ including U_0, such that every vertex $v \in V \setminus U$ has at most $10r$ neighbors in U, with r from Lemma 2.2. A similar idea plays a crucial role in the proof of Luczak. (Note that the number 10 can be easily reduced, and we make no attempt to optimize the multiplicative constants here and in what follows.) To find U as above, start with $U = U_0$, and as long as there exists a vertex $v \in V \setminus U$ having at least $10r$ neighbors in U, join it to U and update U by defining $U := U \cup \{v\}$. This process stops with $|U| \leq 2c\sqrt{n}$ because otherwise we would get a subset $U \subset V$ of size $|U| = \lceil 2c\sqrt{n} \rceil$, containing at least $10rc\sqrt{n}$ edges, thus contradicting the assertion of Lemma 2.2, part (i).

Let $U = \{u_1, \ldots, u_k\}$ with $k = O(\sqrt{n})$. Note that by Lemma 2.2, part (i), the subgraph $G[U]$ is $2r$-choosable. For every i, $1 \leq i \leq k$, put $N(u_i) = \{v \in V \setminus U : (v, u_i) \in E(G)\}$ and let $N(U)$ denote the union $N(U) = \bigcup_{i=1}^{k} N(u_i)$. Define an auxiliary graph H, whose vertex set W is a disjoint union of k sets $W_1, \ldots, W_k$, where $|W_i| = |N(u_i)|$. For each vertex $v \in N(U)$ and each neighborhood $N(u_i)$ in which it participates there is a vertex in W_i corresponding to v. For every edge $(v,w) \in E(G)$ with $v, w \in N(U)$ and for each copy of v and each copy of w in W, there is an edge in H between these two copies. Note that since every $v \in N(U)$ has at most $10r$ neighbors in U, there are at most $10r$ copies of v in H and thus each edge in G yields at most $(10r)^2$ edges in H. Note also that by our choice of H each stable set in H corresponds to a stable set in $N(U)$.

Let $f : V \setminus U \to \{1, \ldots, t\}$ be a fixed proper t-coloring of the subgraph $G[V \setminus U]$. Then f induces a t-coloring $f' : W \to \{1, \ldots, t\}$ of the vertices of H in a natural way.

The crucial idea is as follows. For every $1 \leq i \leq k$ we aim to recolor $2r$ color classes in $N(u_i)$ (or equivalently, in W_i) by a fresh color $t+1$, thus making $2r$ colors available for u_i for a coloring of $G[U]$. We need to show that such a recoloring is possible, that is, $2r$ color classes for each vertex u_i

can be chosen in such a way that their union is a stable set in G. Once this task is accomplished, we would be able to color $G[U]$ using the lists of available colors for each $u \in U$ and exploiting the fact that $G[U]$ is $2r$-choosable.

Let us first consider the case $\delta \leq 1/5$. In this case we apply an argument similar to the one in the proof of Proposition 5.3 in [2], Chapter 5. For each $1 \leq i \leq k$ choose randomly and independently $2r$ numbers from $\{1, \ldots, t\}$ without repetitions, and denote the chosen set by I_i. We claim that with positive probability the subset $W_0 = \bigcup_{i=1}^{k} \{w \in W_i : f'(w) \in I_i\}$ is stable in H. This will imply that the subset of vertices of $N(U)$ with at least one copy in W_0 is also stable, thus making $2r$ colors available for each vertex $u \in U$. To prove this claim, we use the Lovász Local Lemma (c.f., e.g., [2], Chapter 5). Consider an edge $e = (w_1, w_2) \in E(H)$ with $w_1 \in W_{i_1}$ and $w_2 \in W_{i_2}$ (where possibly $i_1 = i_2$). Denote by A_e the event ($f(w_1) \in I_{i_1}$ and $f(w_2) \in I_{i_2}$), that is, "the colors of both w_1 in W_{i_1} and w_2 in W_{i_2} are chosen". The probability of A_e is at most $(2r/t)^2 = O(\ln^2 n/d^2)$. Also, A_e is mutually independent of all other events $A_{e'}$ but those for which $e' \cap (W_{i_1} \cup W_{i_2}) \neq \emptyset$. By the assertion of Lemma 2.2, part (iii) 1, for every $1 \leq i \leq k$ $|N(u_i)| = O(d) = o(\sqrt{n})$ and therefore by the assertion of Lemma 2.2, part (i), the number of edges spanned by $N(u_i)$ and thus by W_i is $O(d)$. Also, by Lemma 2.2, part (iii) 3, the number of edges between $N(u_i)$ and $N(u_j)$ is at most $\ln n$ for $\delta \geq 1/6$ and is at most $d^3 \ln n/n$ for $0 < \delta < 1/6$, implying that the number of edges between any two color classes in H is at most $(10r)^2 \ln n = O(\ln n)$ for $\delta \geq 1/6$ and is at most $(10r)^2 d^3 \ln n/n = O(d^3 \ln n/n)$ for $0 < \delta < 1/6$. Therefore, for every part W_i the number of edges of H incident with W_i is at most $O(d) + k \cdot O(\ln n) = O(n^{1/2} \ln n)$ for $\delta \geq 1/6$ and is at most $O(d) + k \cdot O(d^3 \ln n/n) = O(d^3 \ln n/n^{1/2})$ for $0 < \delta < 1/6$. Returning to the "bad" event A_e we see that it is mutually independent of all but $O(n^{1/2} \ln n)$ events $A_{e'}$ for the case $\delta \geq 1/6$ and of all but $O(d^3 \ln n/n^{1/2})$ events $A_{e'}$ for $0 < \delta < 1/6$. Hence in both cases (since $d \geq n^{2/5}$)

$$Pr[A_e] \cdot |\{e' : e' \cap (W_{i_1} \cup W_{i_2}) \neq \emptyset\}| = o(1) \ . \tag{4}$$

Therefore, applying the symmetric version of the Lovász Local Lemma (see, e.g., [2], Chapter 5, Corollary 1.2), we get the desired result.

Now we treat the case $\delta > 1/5$. For this case the Local Lemma cannot be applied directly since the estimate (4) is not necessarily valid. Therefore we use a different approach.

For every $s \leq k$ subsets $W_{i_1}, \ldots, W_{i_s}$, their union, according to the assertion of Lemma 2.2, part (iii) 1, has $m \leq s \cdot 3d = O(n^{1-\delta})$ vertices and thus, by the assertion of Lemma 2.2(ii), spans at most $(10r)^2 mn^{1/10}$ edges in H. Therefore there exists a subset W_{i_l}, connected by at most $2(10r)^2 mn^{1/10}/s = O(dn^{1/10})$ edges to the rest of the subsets. This implies that the vertices $u_1, \ldots, u_k$ can be reordered in such a way that for every $1 < i \leq k$ there are $O(dn^{1/10})$ edges from W_i to $\bigcup_{i'<i} W_{i'}$. We assume in the sequel that $u_1, \ldots, u_k$ are indeed ordered to satisfy this restriction.

Now, we choose sequentially for every i from 1 to k a set J_i of $2 \ln n + 2r$ colors from $\{1, \ldots, t\}$ at random without repetitions. Each set J_i is chosen from the set of colors *available for* i, where a color j is *available* for i, if the corresponding color class in W_i has no connections with color classes having been chosen for previous indices. More formally, j is available for i if there does not exist an edge $(w_1, w_2) \in E(H)$ with $w_1 \in W_i$, $f'(w_1) = j$, $w_2 \in W_{i'}$ for some $i' < i$ and $f'(w_2) \in J_{i'}$.

Denote by x_i, $1 \leq i \leq k$, the probability that for some $i' \leq i$ while choosing the set $J_{i'}$ there are less than $t/2$ colors available for i'. Clearly, if $x_k < 1$, then there exists a family $\{J_i : 1 \leq i \leq k, |J_i| = 2\ln n + 2r\}$ for which there are no edges between the corresponding color classes of distinct subsets $W_{i'}, W_i$. Once such a family is indeed found, for every $1 \leq i \leq k$ we delete from J_i those colors for which the corresponding color class in W_i is incident with some edge inside W_i. By Lemma 2.2, part (iii) 3, every u_i participates in at most $\ln n$ triangles, hence the number of edges spanned by W_i is at most $\ln n$. Therefore, after this deletion we get a family $\{I_i : 1 \leq i \leq k, |I_i| \geq 2r\}$ for which the union $\bigcup_{i=1}^{k}\{w \in W_i : f'(w) \in I_i\}$ is stable in H, and can complete the proof as before.

In order to estimate x_i, $1 \leq i \leq k$, note first that $x_1 = 0$. Also, according to our reordering of the sets W_i, for each $1 \leq i \leq k$ there are at most $O(dn^{1/10})$ edges from W_i to the previous parts $W_{i'}$, $i' < i$. By Lemma 2.2, part (iii) 3, there are $O(\ln n)$ edges between $W_{i'}$ and W_i, therefore each color chosen to be included in $J_{i'}$ causes $O(\ln n)$ colors to become unavailable for i. The probability of each color to be chosen into $J_{i'}$ is at most $2\ln n + 2r$ divided by the number of available colors for i' at the moment of choosing $J_{i'}$. Hence,

$$
\begin{aligned}
x_i &\leq x_{i-1} + (1 - x_{i-1})\binom{O(dn^{\frac{1}{10}})}{\frac{\frac{t}{2}}{O(\ln n)}}\left(\frac{2\ln n + 2r}{\frac{t}{2}}\right)^{\frac{\frac{t}{2}}{O(\ln n)}} \\
&\leq x_{i-1} + \left[O(1)\frac{dn^{\frac{1}{10}}\ln^2 n}{t^2}\right]^{\Omega(\frac{t}{\ln n})} \\
&\leq x_{i-1} + \left[O(1)\frac{n^{\frac{1}{10}}\ln^4 n}{d}\right]^{\Omega(\frac{d}{\ln^2 n})} \\
&< x_{i-1} + e^{-n^{\frac{1}{10}}}
\end{aligned}
$$

(the next to last inequality uses the assertion of Lemma 2.2, part (iii) 2, while the last inequality uses the assumption that $\delta \leq 3/8$ and thus $d = np \geq n^{1/8}$).

Therefore $x_k \leq ke^{-n^{1/10}} << 1$, establishing the desired result. This completes the proof of the proposition and hence also the proof of Theorem 1.1. $\square$

Proof of Proposition 1.2. Let $p(n)$ be a real so that the probability that the chromatic number of $G(n, p(n))$ is strictly smaller than $r(n)$ is precisely $\epsilon/2$ (such a $p(n)$ clearly exists by continuity, as for every fixed n the above probability is simply a polynomial in p). By Theorem 1.1 (and the fact that from its proof it follows that $n_0(\epsilon, \delta)$ can be uniformly bounded for all $\delta > \delta_0$), there exists an $n_0 = n_0(\beta, \epsilon)$ such that if $n > n_0$ the chromatic number of $G(n, p(n))$ is one of two consecutive values with probability that exceeds $1 - \epsilon/2$. It follows that these two consecutive values must be $r(n) - 1$ and $r(n)$, and the desired result follows, since the probability that the chromatic number is $r(n) - 1$ is at most $\epsilon/2$. $\square$

4 Concluding remarks and open problems

- By continuity, for any n, any $\epsilon \leq 1/2$ and any two consecutive integers $t, t+1 \leq n$ there are values of the probability p such that the chromatic number of $G(n,p)$ is at most t with probability at least ϵ and at least $t+1$ with probability at least ϵ. Therefore, the two-point concentration result is optimal. It seems that for most values of the probability p in the range covered by Theorem 1.1 the chromatic number is in fact concentrated in one point; this is certainly the case for some values of p, as shown in Proposition 1.2. It would be interesting to decide if indeed, in an appropriately defined sense, a one point concentration is more typical in this range than a two-point concentration, and this question remains open.

- The problem of determining or estimating the correct behavior of the concentration of the chromatic number of the random graphs $G(n,p)$ for values of p that exceed $n^{-1/2-\delta}$ remains open, and seems to be very interesting. The case $p = 1/2$ is of particular interest. It seems plausible that there is some fixed $\mu > 0$ so that for infinitely many values of n there is no interval $I(n)$ of length smaller than n^{μ} so that the chromatic number of $G(n, 1/2)$ is in $I(n)$ a.a.s. It is also possible, however, that this is indeed the case, and yet for any such $\mu > 0$ there are infinitely many values of n for which such an interval does exist. Although there are some heuristic arguments that suggest that both these statements may well hold simultaneously, we are unable to prove any of them.

- The arguments in our proof of Theorem 1.1 imply the following result about the non-uniqueness of optimal vertex colorings of random graphs. Suppose $\epsilon > 0$, $\alpha > 1/2$, $p = n^{-\alpha}$ and $t = t(n,p,\epsilon)$ is the least integer for which the probability that the chromatic number of $G = G(n,p)$ is at most t, exceeds ϵ. Suppose, further, that the above mentioned probability is not greater than, say, $1-\epsilon$. Then, if n is sufficiently large (as a function of ϵ and α), the probability that G has (much) more than one proper $(t+1)$-vertex coloring exceeds $1-\epsilon/10$. Note that this implies that with probability greater than $\epsilon/2$ the chromatic number of G is $t+1$ and it has many proper $(t+1)$-vertex colorings.

- For values of $p(n)$ which are very close to 1 (e.g., $p(n) = 1 - 1/(10n)$), there is no interval of length smaller than $\Omega(\sqrt{n})$ for which the chromatic number of $G(n, p(n))$ lies in the interval a.a.s. This follows from some simple facts about the distribution of the size of the largest matching in the complement of $G(n, p(n))$. Thus the concentration result of [10] mentioned in the introduction cannot be improved in general, but it will be interesting to decide how close to the truth it is for values of p up to $1/2$.

- The problem of understanding the correct concentration of the chromatic number of the random graph is equivalent to that of understanding the concentration of the minimum number of cliques that cover all its vertices, since such a covering by cliques corresponds to a coloring of the complement. A related quantity is the *clique cover number* of a graph G, denoted $cc(G)$, which is the minimum number of cliques required to cover all its *edges*. Frieze and Reed [4] proved that this quantity, for $G(n,p)$ for fixed values of p, is, a.a.s., $\Theta(n^2/\ln^2 n)$. For this quantity, we have an argument that shows that for *some* values of p between, say, 0.001 and 0.999, it is *not* concentrated in any interval of length $n/\ln^2 n$. (It is not difficult to see that this is tight, up to a logarithmic factor.) This argument does not seem to provide any information for the concentration of the chromatic number, but since it is simple and applies to *any* graph invariant whose expectation changes considerably for $G(n,p)$ as p changes from ϵ to $1-\epsilon$ we close this paper with a sketch of this proof. The idea is that if $p = p(n)$ is bounded away from 0 and 1, say, $\epsilon < p < 1-\epsilon$, then there is a positive $\delta = \delta(\epsilon)$, such that any family of graphs on n vertices whose total probability is at least $1-\delta$ in $G(n, p(n))$ has probability at least 2δ in $G(n, p(n) + 1/n)$. This implies that if for every edge probability $p = p(n)$ between ϵ and $1-\epsilon$ there is an interval I_p, for which the probability that $cc(G(n,p))$ lies in I_p is larger than $1-\delta$, then for values of p between ϵ and $1-\epsilon$, I_p must intersect $I_{p+1/n}$. Since the mid-point of the interval I_ϵ differs from that of $I_{1-\epsilon}$ by $\Omega(n^2/\log^2 n)$ (as follows from the proof in [4]), the desired result follows.

References

[1] N. Alon, *Restricted colorings of graphs*, in "Surveys in Combinatorics", Proc. 14^{th} British Combinatorial Conference, London Mathematical Society Lecture Notes Series 187, edited by K. Walker, Cambridge University Press, 1993, 1-33.

[2] N. Alon and J. H. Spencer, **The probabilistic method**, Wiley, New York, 1992.

[3] B. Bollobás, *The chromatic number of random graphs*, Combinatorica 8 (1988), 49-55.

[4] A. Frieze and B. Reed, *Covering the edges of a random graph by cliques*, Combinatorica 15 (1995), 489–497.

[5] T. R. Jensen and B. Toft, **Graph coloring problems**, Wiley, New York, 1995.

[6] M. Krivelevich, *On the minimal number of edges in color-critical graphs*, Combinatorica, to appear.

[7] T. Łuczak, *The chromatic number of random graphs*, Combinatorica 11 (1991), 45–54.

[8] T. Łuczak, *A note on the sharp concentration of the chromatic number of random graphs*, Combinatorica 11 (1991), 295–297.

[9] D. W. Matula, On the complete subgraph of a random graph, *Combinatory Mathematics and its Applications*, Chapel Hill, North Carolina (1970), 356–369.

[10] E. Shamir and J. Spencer, *Sharp concentration of the chromatic number of random graphs* $G_{n,p}$, Combinatorica 7 (1987), 124–129.