On a Question of Erdős and Gimbel on the Cochromatic Number

Annika Heckel[a]

Submitted: Aug 30, 2024; Accepted: Nov 11, 2024; Published: Dec 27, 2024

Abstract

In this note, we show that the difference between the chromatic and the cochromatic number of the random graph $G_{n,1/2}$ is not whp bounded by $n^{1/2-o(1)}$, addressing a question of Erdős and Gimbel.

Mathematics Subject Classifications: 05C15, 05C80

1 Introduction

The cochromatic number $\zeta(G)$ of a graph G is the minimum number of colours needed for a vertex colouring where every colour class is either an independent set or a clique. If $\chi(G)$ denotes the usual chromatic number, then clearly $\zeta(G) \leqslant \chi(G)$. Using classical methods, it is not hard to show that for the random graph $G_{n,1/2}$, with high probability[1] we have $\zeta(G_{n,1/2}) \sim \chi(G_{n,1/2}) \sim \frac{n}{2\log_2 n}$.

Erdős and Gimbel [4] (see also [5]) asked the following question: For $G \sim G_{n,1/2}$, does the difference $\chi(G) - \zeta(G)$ tend to infinity as $n \to \infty$? In other words, is there a function $f(n) \to \infty$ such that, with high probability,

$$\chi(G) - \zeta(G) > f(n)?$$

At a conference on random graphs in Poznań[2], Erdős offered \$100 for the solution if the answer was 'yes', and \$1000 if the answer was 'no' (although later said to Gimbel that perhaps \$1000 was too much) [5]. The question is listed as Problem #625 on Thomas Bloom's Erdős Problems website [1].

In this note, we show that it is *not* the case that the difference $\chi(G) - \zeta(G)$ is whp bounded, and that in fact, it is not whp bounded by $n^{1/2-o(1)}$. It turns out that any

[a] Matematiska institutionen, Uppsala universitet, Box 480, 751 06 Uppsala, Sweden (`annika.heckel@math.uu.se`).

[1] As usual, we say that a sequence $(E_n)_{n\geqslant 0}$ of events holds *with high probability (whp)* if $\mathbb{P}(E_n) \to 1$ as $n \to \infty$.

[2] most likely in 1991, or possibly in 1989, based on Erdős's and Gimbel's participation records

function $g(n)$ so that whp $\chi(G) - \zeta(G) \leqslant g(n)$ cannot be smaller than the concentration interval length of the chromatic number $\chi(G_{n,1/2})$, for which corresponding lower bounds were recently obtained [6, 8, 7]. Formally we prove the following statement.

Theorem 1. *Let $G \sim G_{n,1/2}$. There is a constant $c > 0$ so that for any sequence of integers $g(n)$ such that*

$$\mathbf{P}\big(\chi(G) - \zeta(G) \leqslant g(n)\big) > 0.999, \tag{1}$$

there is a sequence of integers n^* such that

$$g(n^*) > c\frac{\sqrt{n^*}\log\log n^*}{\log^3 n^*}.$$

The proof relies on a comparison of the chromatic and co-chromatic numbers of $G \sim G_{n,1/2}$ and of the complement graph $\bar{G} \sim G_{n,1/2}$, which is obtained from G by exchanging all the edges and non-edges. Clearly $\zeta(G) = \zeta(\bar{G})$, so for a function $g(n)$ as above, $\chi(G)$ and $\chi(\bar{G})$ are likely to be at most $g(n)$ apart from each other. Of course $\chi(G)$ and $\chi(\bar{G})$ have the same distribution $\chi(G_{n,1/2})$. Informally speaking, to see why $g(n)$ has to be at least the concentration interval length of this distribution, suppose that we have some lower bound on the concentration interval length; for example, a statement saying that any interval containing $\chi(G_{n,1/2})$ with probability at least 0.9 has length at least $\ell(n)$. Then if $X_1, X_2 \sim \chi(G_{n,1/2})$ were independent samples of this distribution, they would be reasonably likely to be at least about $\ell(n)$ apart from each other. Of course $\chi(G)$ and $\chi(\bar{G})$ are not independent, but the first is an increasing and the other a decreasing function of the edges of $G \sim G_{n,1/2}$, and so with the help of Harris's Lemma we can draw the same conclusion. So we know that $\chi(G)$ and $\chi(\bar{G})$ are both likely to be at most $g(n)$ apart, but also reasonably likely to be at least $\ell(n)$ apart and it follows that $g(n) \geqslant \ell(n)$.

Independently from this note, Raphael Steiner recently also discovered the connection between $\chi(G) - \zeta(G)$ and the concentration interval length of $\chi(G_{n,1/2})$. For his work on this and two other questions of Erdős, Gimbel and Straight, see [10].

2 Proof of Theorem 1

Turning to the details, let us first state the non-concentration result for $\chi(G_{n,1/2})$ that we will use, which follows by combining Theorem 8 from [8] and Theorem 1.2 from [7].

Theorem 2 ([8, 7])**.** *There is a constant $c > 0$ so that for any sequence of intervals $[s_n, t_n]$ such that $\mathbf{P}\big(\chi(G_{n,1/2}) \in [s_n, t_n]\big) > 0.9$, there is a sequence of integers n^* such that*

$$t_{n^*} - s_{n^*} > c\frac{\sqrt{n^*}\log\log n^*}{\log^3 n^*}.$$

Theorem 1 follows directly from Theorem 2 and the following proposition.

Proposition 3. *Let $g(n)$ be a sequence of integers which satisfy* (1)*, then there is a sequence of intervals $[s_n, t_n]$ with $t_n - s_n = g(n)$ so that*

$$\mathbf{P}\big(\chi(G_{n,1/2}) \in [s_n, t_n]\big) > 0.9.$$

Proof of Proposition 3. Let $G \sim G_{n,1/2}$ and let $\bar{G}$ be the complement graph of G (which contains exactly the edges which are missing in G). Then $\bar{G} \sim G_{n,1/2}$ and $\zeta(\bar{G}) = \zeta(G)$, so with probability at least 0.999,

$$\chi(\bar{G}) \leqslant \zeta(\bar{G}) + g(n) = \zeta(G) + g(n) \leqslant \chi(G) + g(n). \tag{2}$$

Now let s_n be the smallest integer k such that $\mathbf{P}(\chi(G) \leqslant k) \geqslant 0.05$, and define the following events:

$$\mathcal{D} = \{\chi(G) \leqslant s_n\},$$
$$\mathcal{U} = \{\chi(\bar{G}) \leqslant s_n + g(n)\}.$$

Then $\mathcal{D}$ is a down-set and $\mathcal{U}$ is an up-set in the edges of $G \sim G_{n,1/2}$. By the definition of s_n,

$$\mathbf{P}(\mathcal{D}) \geqslant 0.05.$$

Furthermore, if $\mathcal{D}$ holds, then either (2) does not hold (which has probability at most 0.001), or (2) holds and implies $\mathcal{U}$, so

$$\mathbf{P}(\mathcal{U} \cap \mathcal{D}) \geqslant \mathbf{P}(\mathcal{D}) - \mathbf{P}\big(\chi(\bar{G}) > \chi(G) + g(n)\big) \geqslant \mathbf{P}(\mathcal{D}) - 0.001.$$

By Harris's Lemma (see for example §2, Lemma 3 in [3]),

$$\mathbf{P}(\mathcal{U}) \geqslant \mathbf{P}(\mathcal{U} \cap \mathcal{D})/\mathbf{P}(\mathcal{D}) \geqslant 1 - \frac{0.001}{\mathbf{P}(\mathcal{D})} \geqslant 1 - \frac{0.001}{0.05} = 0.98.$$

But since G and $\bar{G}$ have the same distribution, and by the definition of s_n, this implies

$$\begin{aligned}\mathbf{P}\big(s_n \leqslant \chi(G) \leqslant s_n + g(n)\big) &= 1 - \mathbf{P}\big(\chi(G) \leqslant s_n - 1\big) - \mathbf{P}\big(\chi(G) > s_n + g(n)\big) \\ &\geqslant 1 - 0.05 - 0.02 > 0.9.\end{aligned}$$

The claim follows. □

3 Discussion

So how about Erdős and Gimbel's original question: does $\chi(G) - \zeta(G)$ tend to infinity whp for $G \sim G_{n,1/2}$? Theorem 1 suggests that the answer is 'yes', but of course does not imply this.

If we had a result like Theorem 1, but with the conclusion that $g(n) \geqslant h(n)$ for every n and some $h(n) \gg \sqrt{n}/\log n$, this would imply that the answer to the original question is 'yes': by an argument of Alon [2, 9], both $\chi(G_{n,1/2})$ and $\zeta(G_{n,1/2})$ are contained in a sequence of intervals of length about $\sqrt{n}/\log n$, respectively, and consequently so is their difference.[3] Let $[s_n, t_n]$ be such a sequence of intervals, then taking $g(n) = t_n$ would give

[3]Formally the statement is: For $G \sim G_{n,1/2}$ and any function $\omega(n) \to \infty$, there is a sequence of intervals of length at most $\omega(n)\sqrt{n}/\log n$ that contains $\chi(G) - \zeta(G)$ whp.

that $t_n \geqslant h(n) \gg \sqrt{n}/\log n$, which implies that $s_n \gg \sqrt{n}/\log n$ as well[4], and so whp $\chi(G) - \zeta(G) \geqslant s_n \gg \sqrt{n}/\log n$.

It is reasonable to expect that the chromatic number $\chi(G_{n,1/2})$ is close to its *first moment threshold*, that is, the smallest k such that the expected number of k-colourings is at least 1.[5] Using the same heuristic for the cochromatic number $\zeta(G_{n,1/2})$, the first moment threshold there should be of order $n/\log^3 n$ smaller than that of the chromatic number: for any $k \sim n/(2\log_2 n)$, the expected number of k-cocolourings is multiplied by a factor $2^k = \exp(\Theta(n/\log n))$ when compared to the expected number of k-colourings (since we may choose for each colour class whether it is a clique or an independent set); and decreasing the number of colours by 1 should multiply the expectation by a factor $\exp(-\Theta(\log^2 n))$ (like it does for the chromatic number[6]). We therefore make the following conjecture.

Conjecture 4. For $G \sim G_{n,1/2}$, whp,

$$\chi(G) - \zeta(G) = \Theta(n/\log^3 n).$$

Acknowledgements

I would like to thank John Gimbel for bringing this problem to my attention and for helpful discussions, and Svante Janson for his help in tracking down past conference proceedings.

This research was funded by the Swedish Research Council, Starting Grant 2022-02829.

References

[1] Erdős Problem #625. http://www.erdosproblems.com/625.

[2] N. Alon and J. Spencer. *The Probabilistic Method*. Wiley, 4th edition, 2015.

[3] B. Bollobás and O. Riordan. *Percolation*. Cambridge University Press, 2006.

[4] P. Erdős and J. Gimbel. Some problems and results in cochromatic theory. In *Annals of Discrete Mathematics*, volume 55, pages 261–264. Elsevier, 1993.

[5] J. Gimbel. Some of my favorite coloring problems for graphs and digraphs. *Graph Theory: Favorite Conjectures and Open Problems-1*, pages 95–108, 2016.

[6] A. Heckel. Non-concentration of the chromatic number of a random graph. *Journal of the American Mathematical Society*, 34(1):245–260, 2021.

[7] A. Heckel and K. Panagiotou. Colouring random graphs: Tame colourings. arXiv:2306.07253, 2023.

[8] A. Heckel and O. Riordan. How does the chromatic number of a random graph vary? *Journal of the London Mathematical Society*, 108(5):1769–1815, 2023.

[4] for a suitable choice of the arbitrary function $\omega(n) \to \infty$

[5] This is an oversimplification and there are several complications with this heuristic, for a detailed discussion see §1.3 and the appendix in [8].

[6] Again an oversimplification, and we refer to the discussion in [8].

[9] A. Scott. On the concentration of the chromatic number of random graphs. arXiv:0806.0178, 2008.

[10] R. Steiner. On the difference between the chromatic and cochromatic number. arXiv:2408.02400v2, 2024.