How does the chromatic number of a random graph vary?

Annika Heckel

LMU München

Oxford Discrete Mathematics and Probability Seminar

4 May 2021

Joint work with Oliver Riordan.
**Colouring of G:** Colour vertices so that neighbours get different colours

**Chromatic number** $\chi(G)$: Minimum number of colours we need

Pick an $n$-vertex graph uniformly at random. Pick another one.

Does it have the same chromatic number?

If not, how different are their chromatic numbers likely to be?
Chromatic number of random graphs

\[ G_{n, \frac{1}{2}} = \text{choose a graph on } n \text{ labelled vertices uniformly at random} \]

What can we say about \( \chi(G_{n,p}) \)?

- Value?
  - Upper and lower bounds?
- Concentration?
  - How much does \( \chi(G_{n,p}) \) vary?
What can we say about $\chi(G_{n,p})$?

**Value?**
Upper and lower bounds?

**Concentration?**
How much does $\chi(G_{n,p})$ vary?

**Bollobás 1987:**

$$\chi(G_{n,\frac{1}{2}}) \sim \frac{n}{2 \log_2 n} \text{ whp.}$$

**Improvements:** McDiarmid '90, Panagiotou & Steger '09, Fountoulakis, Kang & McDiarmid '10.

**H. 2016:**

$$\chi(G_{n,\frac{1}{2}}) = \frac{n}{2 \log_2 n - 2 \log_2 \log_2 n - 2} + o\left(\frac{n}{\log^2 n}\right) \text{ whp.}$$

*Explicit interval* of length $o\left(\frac{n}{\log^2 n}\right)$ which contains $\chi(G_{n,\frac{1}{2}})$ whp.
Concentration?

**Shamir, Spencer 1987:** For any function $p = p(n)$, $\chi(G_{n,p})$ is whp contained in a sequence of intervals of length about $\sqrt{n}$.

$p = 1 - \frac{1}{10n}$: not concentrated on fewer than $\Theta(\sqrt{n})$ values

$p \leq \frac{1}{2}$: slight improvement to $\frac{\sqrt{n}}{\log n}$ (Alon)

$p < n^{-\frac{1}{2} - \varepsilon}$: 2 values (‘two-point concentration’) (Alon, Krivelevich 97, Łuczak 91)

$\implies \chi(G_{n,p})$ behaves almost deterministically
The opposite question

**Question (Bollobás)**

Can we show that $\chi(G_n, \frac{1}{2})$ is **not** concentrated on 100 consecutive values?

Upper bound: $\frac{\sqrt{n}}{\log n}$ (Alon)

**Theorem (H. 2019; H., Riordan 2021)**

Let $\varepsilon > 0$, and let $[s_n, t_n]$ be a sequence of intervals such that $\chi(G_{n,1/2}) \in [s_n, t_n]$ whp. Then there are infinitely many values $n$ such that

$$t_n - s_n \geq n^{1/2 - \varepsilon}.$$
Proof ingredients

**Ingredient 1:** A (weak) concentration type result

\[ |\chi(G_{n,1/2}) - f(n)| \leq \Delta(n) \text{ whp} \]

where \( f(n) \) is some function with slope

\[ \frac{d}{dn} f(n) > \frac{1}{\alpha} + \delta. \]

(Will specify \( \alpha = \alpha(n) \) later.)

**Ingredient 2:** A coupling result

Couple \( G_{n,1/2} \) and \( G_{n',1/2} \) with \( n' = n + \alpha r \) (same \( \alpha \) as above) so that

\[ \mathbb{P}(\chi(G_{n',1/2}) \leq \chi(G_{n,1/2}) + r) > \frac{1}{4}. \]
If all intervals short: Contradiction!

So there is at least one long interval. (Length $\approx \alpha \delta r$)
What’s $\alpha(n)$?

**Independence number** $\alpha(G)$: Size of the largest independent vertex set (= set without edges).

For most $n$: $\alpha(G_n, \frac{1}{2}) = \alpha(n) \approx 2 \log_2 n$ whp.
What does this have to do with colourings?

Every colour class is an independent set, so if there are \( n \) vertices,

\[
\chi(G) \geq \frac{n}{\alpha(G)}
\]

In fact:

\[
\chi(G_n, \frac{1}{2}) = \frac{n}{\alpha - O(1)}
\]

**Intuition:** An optimal colouring of \( G_{n, \frac{1}{2}} \) contains all or almost all independent \( \alpha \)-sets as colour classes.

\[
\chi(G_{n, \frac{1}{2}}) \text{ should vary at least as much as } X_\alpha.
\]

\[
X_\alpha = \# \text{ independent } \alpha \text{-sets}
\]

\[
X_\alpha \sim \text{Poi}_\mu \quad \rightarrow \text{varies by } \pm \sqrt{\mu}
\]

where \( \mu = n^\rho, \quad 0 \leq \rho(n) \leq 1. \)
Ingredient 1: The (weak) concentration type result

Want:

\[ \chi(G_n, \frac{1}{2}) = f(n) \pm \Delta(n) \]

\[ \frac{df}{dn} \geq \frac{1}{\alpha} + \delta \]

H. 2016:

\[ \chi\left(G_n, \frac{1}{2}\right) = \frac{n}{2 \log_2 n - 2 \log_2 \log_2 n - 2} + o\left(\frac{n}{\log^2 n}\right) \text{ whp.} \]

then (unless \( \mu_\alpha \) is very close to \( n \))

\[ \frac{df}{dn} \geq \frac{1}{\alpha} + \Theta\left(\frac{1}{\log^2 n}\right) \delta(n) \]
Ingredient 2: The coupling result

Want: Coupling of $G_{n,1/2}$ and $G_{n',1/2}$ with $n' = n + \alpha r$ so that

$$\mathbb{P}\left(\chi(G_{n',1/2}) \leq \chi(G_{n,1/2}) + r\right) > \frac{1}{4}. $$

$G'$: $n'$ vertices

$G$: $n$ vertices

plant $r$ random $\alpha$-sets

normal $G_{n,\frac{1}{2}}$
$G': n'$ vertices

$G: n$ vertices

- Inner random graph: $G \sim G_{n, \frac{1}{2}}$
- Also clear:

\[
\chi(G') \leq \chi(G) + r
\]

- Need to prove: $G'$ similar to $G_{n', \frac{1}{2}}$ if $r$ is not too big
**Key Lemma**

**Planted model** $G_{n,\frac{1}{2}}^{pl}$: Plant an independent $\alpha$-set uniformly at random, and include all other edges independently with probability $\frac{1}{2}$.

$d_{TV}$: Total variation distance

**Key Lemma**

$$d_{TV} \left( G_{n,\frac{1}{2}}, G_{n,\frac{1}{2}}^{pl} \right) = O \left( \frac{1}{\sqrt{\mu}} \right),$$

where $\mu = \mathbb{E}[X_\alpha]$.

This means: $G_{n,\frac{1}{2}}$ and $G_{n,\frac{1}{2}}^{pl}$ can be **coupled** so that they agree with probability

$$1 - O \left( \frac{1}{\sqrt{\mu}} \right).$$
Key Lemma

$$d_{TV} \left( G_{n, \frac{1}{2}}, G_{n, \frac{1}{2}}^{pl} \right) = O \left( \frac{1}{\sqrt{\mu}} \right),$$

where $$\mu = \mathbb{E}[X_\alpha].$$

Proof:

$$d_{TV} \left( G_{n, \frac{1}{2}}, G_{n, \frac{1}{2}}^{pl} \right) = \frac{1}{2} \sum_G \left| \mathbb{P} \left( G_{n, \frac{1}{2}}^{pl} = G \right) - \mathbb{P} \left( G_{n, \frac{1}{2}} = G \right) \right|$$

$$= \frac{1}{2} \sum_G \left| \frac{X_\alpha(G)}{\binom{n}{\alpha}} \left( \frac{1}{2} \right)^{\binom{n}{\alpha}} - \left( \frac{1}{2} \right)^{\binom{n}{2}} \right|$$

$$= \frac{1}{2} \sum_G \left( \frac{1}{2} \right)^{\binom{n}{2}} \frac{X_\alpha(G) - \binom{n}{\alpha} \left( \frac{1}{2} \right)^{\binom{\alpha}{2}}}{\binom{n}{\alpha} \binom{\alpha}{2}}$$

$$= \frac{1}{2} \mathbb{E} \left[ \frac{|X_\alpha - \mu|}{\mu} \right] = O \left( \frac{1}{\sqrt{\mu}} \right)$$
If \( r = o\left(\sqrt{\mu}\right) \), then

\[
d_{TV}\left(G', G_{n', \frac{1}{2}}\right) = o(1)
\]

So can couple \( G_{n, \frac{1}{2}} \) and \( G_{n', \frac{1}{2}} \) such that, whp,

\[
\chi(G_{n', \frac{1}{2}}) \leq \chi(G_{n, \frac{1}{2}}) + r
\]
So what’s the truth?

Recall: $X_\alpha \sim \text{Poi}_\mu$ where $\mu = n^\rho$, $0 \leq \rho(n) \leq 1$.

Conjecture: $\chi(G_n, \frac{1}{2})$ is not concentrated on fewer than $n^{\rho/2}/\log n$ values.
Conjecture: \( \chi(G_n, \frac{1}{2}) \) is not concentrated on fewer than \( n^{\rho/2}/\log n \) values.

**Theorem (H., Riordan 21)**

Let \([s_n, t_n]\) be a sequence of intervals and suppose that \( \chi(G_n, \frac{1}{2}) \in [s_n, t_n] \) whp. Then for every \( n \) with \( \rho(n) < 0.99 \), there is some \( n^* \sim n \) such that

\[
t_{n^*} - s_{n^*} \geq \frac{(n^*)^{\rho(n^*)/2}}{2000 \log n^*}.
\]
Sources of non-concentration?

- Number of $\alpha$-sets: conjectured lower bound $n^{\rho/2 + o(1)}$
- Number of edges?

  Don’t seem to matter much. Can couple $G_{n,p}$ and $G_{n,m}$ so that the chromatic numbers only differ by about $\log n$.
- Number of $(\alpha - 1)$-sets

  $X_{\alpha-1}$ roughly Poisson with mean about $n^{1+\rho+o(1)}$.

  If $X_{\alpha-1}$ decreases by $n^{1+\rho/2 + o(1)}$, we need about $n^{1-\rho/2 + o(1)}$ more colours to make the expected number of colourings 1.

  Is that all?

**H., Panagiotou 21+**: The $(\alpha - 2)$-bounded chromatic number of $G_{n,m}$, $m = \frac{1}{2}N$, takes one of at most 2 consecutive values whp.
Zig-zag conjecture

(Bollobás, H., Morris, Panagiotou, Riordan, Smith)
Finer conjectures

- Gaussian limiting distribution: $\frac{n^{1/4}}{\log^{7/4} n}$
- Best case: $\frac{n^{1/4}}{\log^{7/4} n}$
- Worst case: $\frac{\sqrt{n \log \log n}}{\log^3 n}$, at $\mu_\alpha = \Theta \left( \frac{n}{\log^2 n} \right)$
Best lower concentration bound?

At the moment: $n^{1/2-o(1)}$ for infinitely many $n$.

**Bottleneck:** Error term $o\left(\frac{n}{\log^2 n}\right)$ in the explicit estimate for $\chi(G_{n,1/2})$.

**H., Panagiotou 21+:** Sharper explicit estimate for $\chi(G_{n,1/2})$.

**H., Riordan 21:** Assuming this estimate, the lower bound on the interval length can be improved to

$$c \frac{\sqrt{n \log \log n}}{\log^3 n}$$

for infinitely many $n$. 
Open questions

- Does the correct concentration interval length zigzag between $n^{1/4+o(1)}$ and $n^{1/2+o(1)}$?

- The proof only finds some $n^*$ near $n$ where the chromatic number is not too concentrated. Can we prove something for every $n$?

- Alon’s upper bound: $\sqrt{n}/\log n$. Our lower bound: $n^{1/2-o(1)}$.

  Conditional lower bound: $\frac{\sqrt{n}\log\log n}{\log^3 n}$. Show that this is optimal?

- Other ranges of $p$?

  $p < n^{-\frac{1}{2}-\varepsilon}$: two-point concentration. How “far down” does non-concentration go?

Thank you!