# The (Non-)Concentration of the Chromatic Number

Annika Heckel

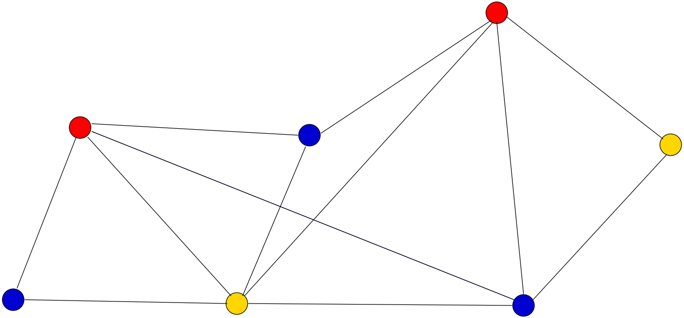
LMU München

**Concentration of Measure Phenomena**
Simons Institute

23 October 2020

# What is a colouring?

**Colouring of $G$:** Colour vertices so that neighbours get different colours



**Chromatic number $\chi(G)$:** Minimum number of colours where this is possible

$G_{n,p}$: $n$ vertices, include each edge independently with probability $p$

**What is the chromatic number of $G_{n,p}$?**

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

**Value?**

Upper and lower bounds?

**Concentration?**

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

$$p = \frac{1}{2}$$

**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:**

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

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

# How about 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 $\sqrt{n}$.

**Standard tool:** Azuma-Hoeffding inequality + vertex exposure martingale.

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

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

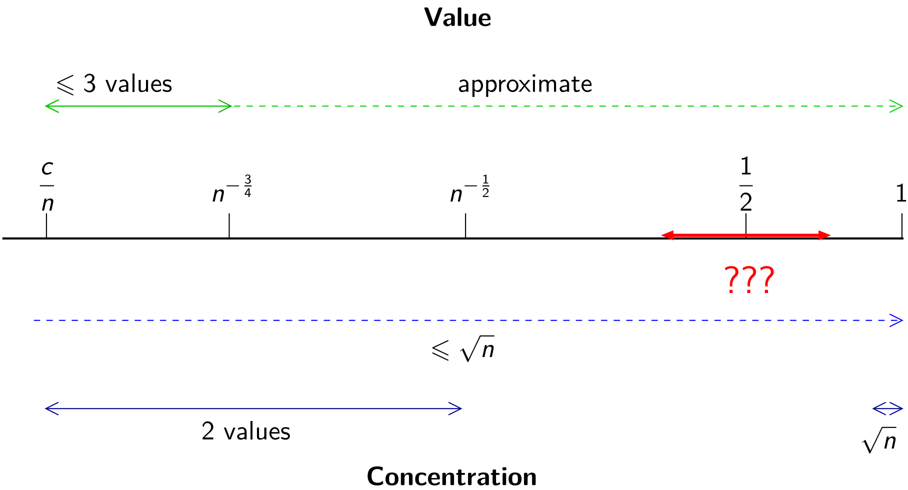
**Sparse random graphs:**

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

$p = \frac{C}{n}$: 2 *explicit* values. (Achlioptas, Naor 04)

$p < n^{-3/4-\varepsilon}$: 3 *explicit* values. (Coja-Oghlan, Panagiotou, Steger 08)

# $\chi(G_{n,p})$ for different $p = p(n)$




= minimum length of a series of intervals containing $\chi(G_{n,p})$

# The opposite question

Bollobás, Erdős, late 80s: Any non-concentration results?

**Erdős 1992, appendix to *The Probabilistic Method*:**
Can we show that $\chi(G_{n,\frac{1}{2}})$ is not concentrated on a constant number of values?

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

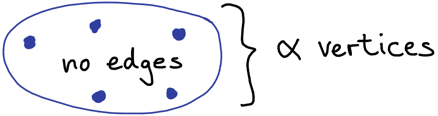
**Bollobás 2004:**
Any non-trivial examples of non-concentration?
*"even the weakest results claiming lack of concentration would be of interest"*

Theorem (H. 2019; H., Riordan 2020+):

$\chi(G_{n,\frac{1}{2}})$ is not contained whp in any sequence of intervals of length $n^{\frac{1}{2}-\varepsilon}$ for any fixed $\varepsilon > 0$.

# Independent sets

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

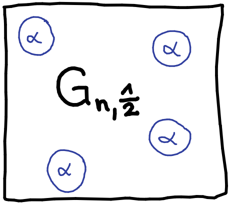



$$\alpha(G_{n,\frac{1}{2}}) = \lfloor \alpha_0 + o(1) \rfloor \text{ whp,}$$

where $\alpha_0 = 2\log_2 n - 2\log_2 \log_2 n + 2\log_2 (e/2) + 1$

$$X_\alpha = \#\text{ independent } \alpha\text{-sets}$$

$$X_\alpha \underset{\text{roughly}}{\sim} \mathrm{Poi}_\mu$$

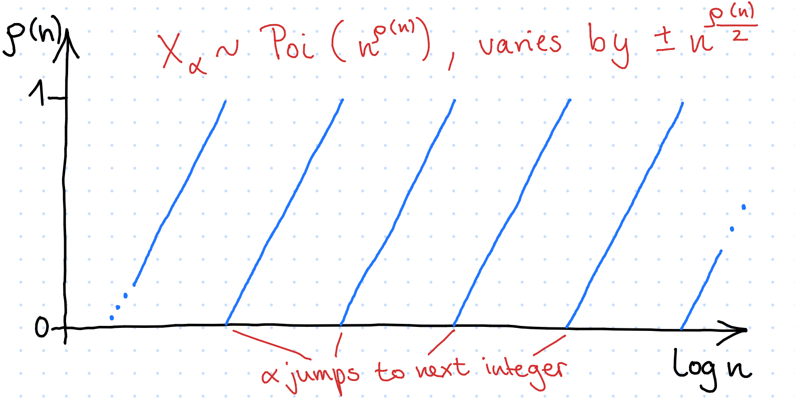


# $X_\alpha$ is not very sharply concentrated

$X_\alpha = \#$ independent $\alpha$-sets

$$X_\alpha \underset{\text{roughly}}{\sim} \mathrm{Poi}_\mu$$

$$\mu = n^\rho, \quad 0 \leqslant \rho(n) \leqslant 1.$$



# What does this have to do with colourings?

Every colour class is an independent set, so if there are $n$ vertices,

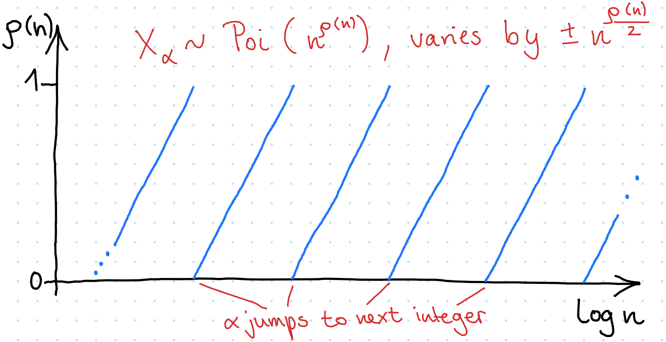
$$\chi(G) \geqslant \frac{n}{\alpha(G)}$$

We know:

$$\chi(G_{n,\frac{1}{2}}) \approx \frac{n}{\alpha_0 - 3.89}$$

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

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




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

## Theorem(H., Riordan 20+)

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

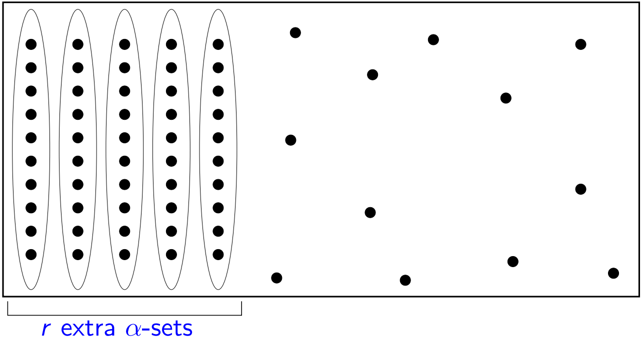
$$t_{n^*} - s_{n^*} \geqslant \frac{(n^*)^{\rho(n^*)/2}}{1000 \log n^*}.$$

# First proof attempt

Compare chromatic numbers of

- $G_{n,\frac{1}{2}}$ with $X_\alpha = A \approx \mu$
- $G_{n,\frac{1}{2}}$ with $X_\alpha = A + r$ where $r = \sqrt{\mu} = n^{\rho/2}$.

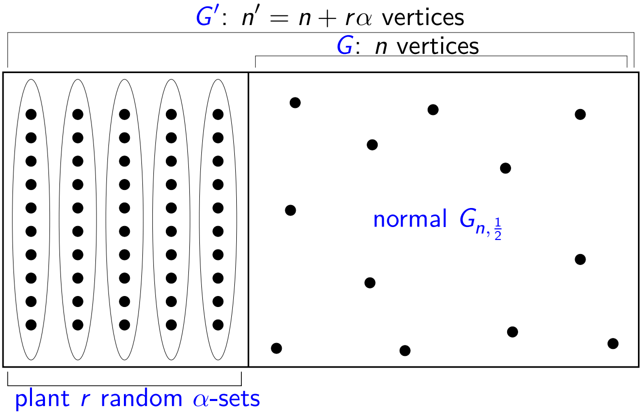
**Hope:** If $X_\alpha$ goes up, $\chi(G_{n,\frac{1}{2}})$ goes down.




**Problem:** Optimal colouring might not use these $\alpha$-sets.

# Second proof attempt

**Trick:** Compare $G_{n,\frac{1}{2}}$ for different $n$.




- Inner random graph: $G \sim G_{n,\frac{1}{2}}$
- Want to show: $G'$ similar to $G_{n',\frac{1}{2}}$

# Key Lemma

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

$d_{\mathrm{TV}}$: Total variation distance

## Key Lemma

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

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

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

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

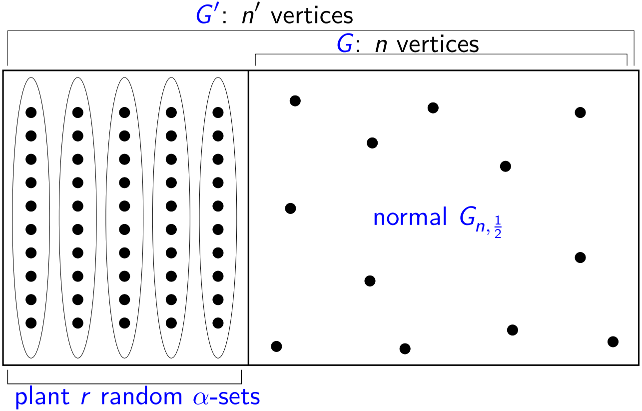
## Key Lemma

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

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

**Proof:**

$$\begin{aligned}
d_{\mathrm{TV}}\left(G_{n,\frac{1}{2}}, G^{\mathrm{pl}}_{n,\frac{1}{2}}\right) &= \frac{1}{2}\sum_G \left|\mathbb{P}\left(G^{\mathrm{pl}}_{n,\frac{1}{2}} = G\right) - \mathbb{P}\left(G_{n,\frac{1}{2}} = G\right)\right| \\
&= \frac{1}{2}\sum_G \left|\frac{X_\alpha(G)}{\binom{n}{\alpha}}\left(\frac{1}{2}\right)^{\binom{n}{2}-\binom{\alpha}{2}} - \left(\frac{1}{2}\right)^{\binom{n}{2}}\right| \\
&= \frac{1}{2}\sum_G \left(\frac{1}{2}\right)^{\binom{n}{2}} \frac{\left|X_\alpha(G) - \binom{n}{\alpha}\left(\frac{1}{2}\right)^{\binom{\alpha}{2}}\right|}{\binom{n}{\alpha}\left(\frac{1}{2}\right)^{\binom{\alpha}{2}}} \\
&= \mathbb{E}\left[\frac{|X_\alpha - \mu|}{\mu}\right] = O\left(\frac{1}{\sqrt{\mu}}\right)
\end{aligned}$$




Let $r = o\left(\sqrt{\mu}\right)$ and $n' = n + r\alpha$.

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

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

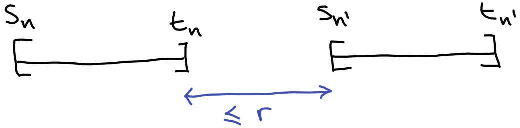
$$\chi(G_{n',\frac{1}{2}}) \leqslant \chi(G_{n,\frac{1}{2}}) + r$$

# Proof ingredients

**Ingredient 1:** A coupling of $G_{n,\frac{1}{2}}$ and $G_{n',\frac{1}{2}}$, $n' = n + \alpha r$, such that whp

$$\chi(G_{n',\frac{1}{2}}) \leqslant \chi(G_{n,\frac{1}{2}}) + r$$

Suppose that $\chi(G_{n,\frac{1}{2}}) \in [s_n, t_n]$ whp.



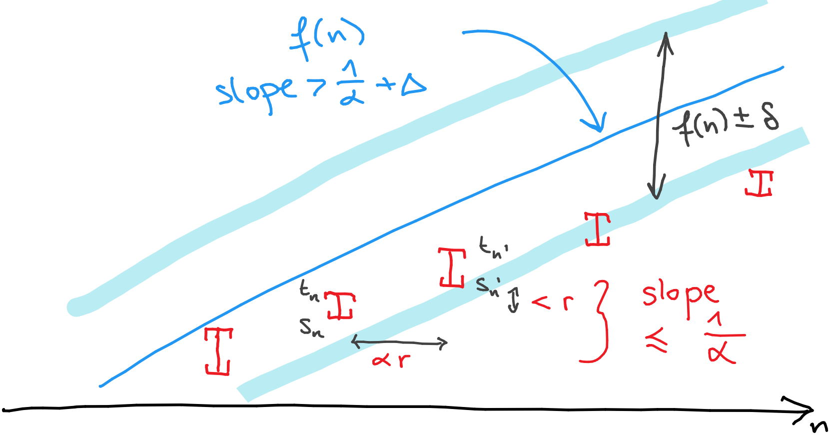

**Ingredient 2:** A (weak) concentration result

$$\chi(G_{n,\frac{1}{2}}) = f(n) \pm \delta(n)$$

with

$$\frac{\mathrm{d}f}{\mathrm{d}n} \geqslant \frac{1}{\alpha} + \Delta$$

Use known bounds: $\chi(G_{n,\frac{1}{2}}) = \dfrac{n}{2\log_2 n - 2\log_2\log_2 n - 2} + o\left(\dfrac{n}{\log^2 n}\right)$




If intervals **short**: Contradiction!

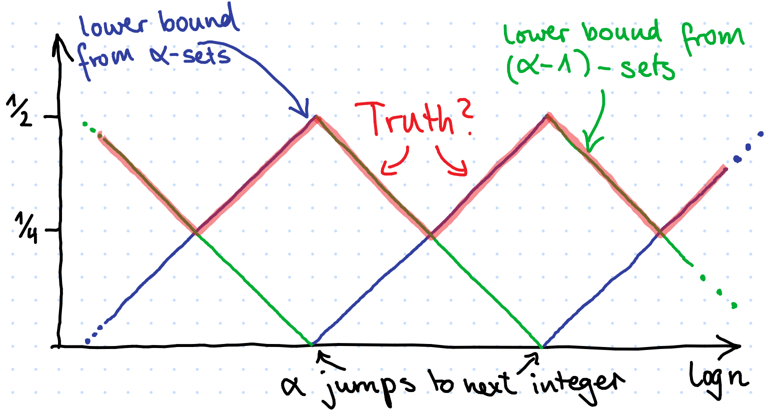
So there is at least one long interval.
(length $\approx r/\log n$)

# So what's the truth?

No colour classes of size $\alpha$ and $\alpha - 1$: two point concentration

**H., Panagiotou 20+:** The $(\alpha - 2)$-bounded chromatic number of $G_{n,\frac{1}{2}}$ takes one of at most 2 consecutive values whp.

**Zig-zag conjecture:** (Bollobás, H., Morris, Panagiotou, Riordan, Smith)



# Open questions

- Does the correct concentration interval length zigzag between $n^{1/4}$ and $n^{1/2}$?
- The proof only finds some $n^*$ where the chromatic number is not too concentrated. Can we prove something for every n?
- Alon's upper bound: $\frac{\sqrt{n}}{\log n}$. Our lower bound: $n^{\frac{1}{2}-o(1)}$. Close the gap?

  Preview: Might push lower bound to $\frac{\sqrt{n}}{\log^5 n}$.
- Other ranges of $p$?

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

Thank you!