Combinatorics, Probability and Computing (2020), **29**, pp. 213–233
doi:10.1017/S0963548319000397

CAMBRIDGE
UNIVERSITY PRESS

ARTICLE

Sharp concentration of the equitable chromatic number of dense random graphs

Annika Heckel†

Mathematical Institute, University of Oxford, Andrew Wiles Building, Woodstock Road, Oxford OX2 6GG, UK
Email: heckel@maths.ox.ac.uk

(Received 20 December 2017; revised 29 July 2019; first published online 26 November 2019)

Abstract

An *equitable colouring* of a graph G is a vertex colouring where no two adjacent vertices are coloured the same and, additionally, the colour class sizes differ by at most 1. The *equitable chromatic number* $\chi_=(G)$ is the minimum number of colours required for this. We study the equitable chromatic number of the dense random graph $\mathcal{G}(n, m)$ where $m = \lfloor p\binom{n}{2}\rfloor$ and $0 < p < 0.86$ is constant. It is a well-known question of Bollobás [3] whether for $p = 1/2$ there is a function $f(n) \to \infty$ such that, for any sequence of intervals of length $f(n)$, the normal chromatic number of $\mathcal{G}(n, m)$ lies outside the intervals with probability at least 1/2 if n is large enough. Bollobás proposes that this is likely to hold for $f(n) = \log n$. We show that for the *equitable* chromatic number, the answer to the analogous question is negative. In fact, there is a subsequence $(n_j)_{j\in\mathbb{N}}$ of the integers where $\chi_=(\mathcal{G}(n_j, m_j))$ is concentrated on exactly one explicitly known value. This constitutes surprisingly narrow concentration since in this range the equitable chromatic number, like the normal chromatic number, is rather large in absolute value, namely asymptotically equal to $n/(2\log_b n)$ where $b = 1/(1-p)$.

2010 MSC Codes: Primary 05C80; Secondary 05C15

1. Introduction

An assignment of colours to the vertices of a graph G is called a *(proper) colouring* if no two adjacent vertices are coloured the same. The *chromatic number* $\chi(G)$ is the minimum number of colours required for this.

An *equitable colouring* is a colouring where, additionally, the colour classes (*i.e.* the sets of vertices of each colour) are as equal in size as possible. Since the number of colours does not necessarily divide the number of vertices, this means that the colour class sizes may differ by at most 1. The least number of colours where this is possible is called the *equitable chromatic number* $\chi_=(G)$.

Finally, the *equitable chromatic threshold* $\chi^*(G)$ is defined as the smallest k such that, for all $l \geqslant k$, G allows an equitable colouring with exactly l colours. Note that, for any graph G,

$$1 \leqslant \chi(G) \leqslant \chi_=(G) \leqslant \chi^*_=(G) \leqslant n.$$

The famous Hajnal–Szemerédi theorem [6] states that if G has maximum degree $\Delta(G)$, then

$$\chi^*_=(G) \leqslant \Delta(G) + 1.$$

†This research was partially funded by ERC grant 676632.

In this paper we will study the equitable chromatic number of dense random graphs. For $p \in [0,1]$, we let $G \sim \mathcal{G}(n,p)$ denote the *binomial random graph* with n labelled vertices where each of the $N = \binom{n}{2}$ possible edges is present independently with probability p. Given integers n and $m \in \{0, \ldots, N\}$, the *uniform random graph* $G \sim \mathcal{G}(n,m)$ is chosen uniformly at random from all graphs with n labelled vertices and exactly m edges.

Determining the chromatic number is a classic challenge in random graph theory and was raised in one of the earliest papers by Erdős and Rényi [4]. In a landmark contribution, Bollobás first determined the asymptotic value of the chromatic number of the dense random graph $G \sim \mathcal{G}(n,p)$ where p is constant [2]. Using martingale concentration inequalities, he proved that w.h.p.[1]

$$\chi(G) = (1 + o(1)) \frac{n}{2 \log_b n},$$

where $b = 1/(1-p)$. This result has been sharpened several times [5, 9, 10, 11], most recently in [7], where the first and second moment methods were used to show that w.h.p.

$$\chi(G) = \frac{n}{2 \log_b n - 2 \log_b \log_b n - 2 \log_b 2 - x_0(n) + o(1)} \tag{1.1}$$

for a certain function $x_0(n) \in [0,1]$. If $p \leqslant 1 - 1/e^2 \approx 0.86$, $x_0(n) = 0$ for all n.

In general, $\chi(G)$ and $\chi_=(G)$ can be far apart from each other: for example, if G is the star $K_{1,n-1}$, then $\chi(G) = 2$ and $\chi_=(G) = 1 + \lceil (n-1)/2 \rceil$. However, for dense random graphs $G \sim \mathcal{G}(n,p)$ where $n^{-1/5+\varepsilon} \leqslant p \leqslant 0.99$, Krivelevich and Patkós [8] proved that w.h.p.

$$\chi_=(G) \sim \chi(G).$$

They also showed (amongst other things) that if $p \leqslant 0.99$ and $\log \log n \ll \log(np)$, then

$$\chi_=^*(G) \leqslant (2 + o(1)) \chi(G).$$

Rombach and Scott have announced further results.

While the results above concern the likely *values* of $\chi(\mathcal{G}(n,p))$, $\chi_=(\mathcal{G}(n,p))$ and $\chi_=^*(\mathcal{G}(n,p))$, the *concentration of the chromatic number* is similarly well studied. Shamir and Spencer [13] showed that for any function $p = p(n)$, $\chi(\mathcal{G}(n,p))$ is w.h.p. concentrated on an interval of length about $\sqrt{n}$. If p tends to 1 sufficiently quickly, for example if $p = 1 - 1/(10n)$, this result is asymptotically optimal (see [1]). In contrast, for the equitable chromatic number no general concentration results are known. This is because the martingale concentration arguments used in [13] do not apply in the equitable setting.

For sufficiently sparse random graphs, there are much sharper concentration results: Alon and Krivelevich [1] proved that if $p = n^{-1/2-\varepsilon}$ with $\varepsilon > 0$, the chromatic number of $\mathcal{G}(n,p)$ is concentrated on only two consecutive values, which is generally best possible. For the wide range of values of p which are larger than $n^{-1/2-\varepsilon}$ but not close to 1 – in particular for $p = 1/2$ – the question of the concentration of $\chi(\mathcal{G}(n,p))$ is still wide open. Alon and Krivelevich speculate in [1] that the answer for $\mathcal{G}(n, 1/2)$ might be different for n in different subsequences of the integers.

One very simple reason why extremely sharp concentration of $\chi(\mathcal{G}(n,p))$ would be surprising is that the chromatic number is quite large in absolute value, namely of order $\Theta(n/\log n)$. In [1], Alon and Krivelevich give an argument showing non-concentration of any graph parameter which changes considerably as p increases from ε to $1-\varepsilon$. This argument only works for graph parameters of order at least n.

For p constant, the number of edges in $\mathcal{G}(n,p)$ is of order n^2 with standard deviation of order n, so part of the variance of $\chi(\mathcal{G}(n,p))$ may simply be due to variations in the number of edges.

[1] We say that an event $E = E(n)$ holds *with high probability* (w.h.p.) if $\lim_{n\to\infty} \mathbb{P}(E) = 1$.

https://doi.org/10.1017/S0963548319000397 Published online by Cambridge University Press

Therefore, to study the concentration of the chromatic number, it makes sense to focus our attention on the random graph $\mathcal{G}(n, m)$ where the number of edges is fixed at $m = \lfloor pN \rfloor$.

In [3], Bollobás asks whether there is a function $f(n) \to \infty$ such that, for any sequence of intervals I_n where I_n is of length $f(n)$, $\chi(\mathcal{G}(n, \lfloor \frac{1}{2}N \rfloor))$ does not lie in any interval of length $f(n)$ with probability at least 1/2, and conjectures that $f(n) = \log n$, and maybe even $f(n) = n^{\varepsilon}$ could do.

In this paper we will show that the corresponding statement about the *equitable chromatic number* does not hold. In fact, we will show that for constant $p \in (0, 0.86)$, there is a subsequence $(n_j)_{j \geqslant 1}$ of the integers such that the equitable chromatic number of $\mathcal{G}(n_j, m_j)$ with $m_j = \lfloor p\binom{n_j}{2} \rfloor$ is concentrated on exactly one explicitly known value.

Theorem 1.1. *Let $0 < p < 1 - 1/e^2 \approx 0.86$ be constant. There exists a strictly increasing sequence of integers $(n_j)_{j \geqslant 1}$ such that:*

(a) *for all $j \geqslant 1$, $j | n_j$,*
(b) *letting $b = 1/(1-p)$ and $\gamma(n) = 2\log_b n - 2\log_b \log_b n - 2\log_b 2$,*

$$\gamma(n_j) = j + o(1) \text{ as } j \to \infty,$$

(c) *letting $G \sim \mathcal{G}(n_j, m_j)$ with $m_j = \lfloor p\binom{n_j}{2} \rfloor$, with high probability as $j \to \infty$,*

$$\chi_{=}(G) = \frac{n_j}{j}.$$

In other words, we can pick a subsequence $(n_j)_{j \geqslant 1}$ of the integers so that w.h.p. as $j \to \infty$, the equitable chromatic number of $G \sim \mathcal{G}(n_j, m_j)$ with $m_j = \lfloor p\binom{n_j}{2} \rfloor$ is exactly n_j/j. This concentration is perhaps surprisingly sharp because, like the normal chromatic number, the equitable chromatic number of these dense random graphs is of order $\Theta(n/\log n)$. For a discussion of the condition that $p < 1 - 1/e^2$, see the end of the next section.

We prove Theorem 1.1 in Sections 2–5. The proof is based on a very accurate calculation of the second moment of the number of equitable k-colourings, and relies on choosing the sequence $(n_j)_{j \geqslant 1}$ in such a way that just as the expected number of equitable colourings starts tending to infinity, all colour classes are of exactly the same size. We will use several lemmas from [7], where the second moment method was recently used to obtain the currently best bounds for the normal chromatic number of $\mathcal{G}(n, p)$ where p is constant.

2. Outline and notation

From now on, fix $p < 1 - 1/e^2$, let $q = 1 - p$ and $b = 1/q$. Let $N = \binom{n}{2}$, $m = m(n) = \lfloor pN \rfloor$ and $G \sim \mathcal{G}(n, m(n))$.

For two functions $f = f(n)$, $g = g(n)$, we say that f is asymptotically at most g, denoted by $f \lesssim g$, if $f(n) \leqslant (1 + o(1))g(n)$ as $n \to \infty$. We write $f = O(g)$ if there are constants C and n_0 such that $|f(n)| \leqslant Cg(n)$ for all $n \geqslant n_0$. We write $f = \Omega(g)$ if $g(n) \geqslant 0$ and there are constants $c > 0$ and n_0 such that $f(n) \geqslant cg(n)$ for all $n \geqslant n_0$. Furthermore, we use the notation $f = \Theta(g)$ if $f = O(g)$ and $g = O(f)$.

For $k \geqslant 1$, we call an ordered partition of n vertices into k parts an *ordered k-equipartition* if all k parts have size $\lceil n/k \rceil$ or $\lfloor n/k \rfloor$ and decrease in size (so the parts of size $\lceil n/k \rceil$ come first, followed by the parts of size $\lfloor n/k \rfloor$). Let $X_{n,k}$ denote the number of ordered k-equipartitions of G which induce valid colourings.

We start with a straightforward analysis of the first moment of $X_{n,k}$ in Section 3. Next, in Section 4, we show that there is a strictly increasing sequence $(n_j)_{j \geqslant 1}$ which fulfils parts (a) and (b) of Theorem 1.1, so that, letting $k_j = n_j/j$, $\mathbb{E}[X_{n_j,k_j}]$ tends to infinity slowly as $j \to \infty$. Using a first

https://doi.org/10.1017/S0963548319000397 Published online by Cambridge University Press

moment argument, we then show that w.h.p. G has no equitable k-colouring if $k < k_j$, so w.h.p.

$$\chi_=(G) \geqslant k_j = \frac{n_j}{j}.$$

We prove the matching upper bound through the second moment method. By the Paley–Zygmund inequality, for any non-negative integer random variable Z,

$$\mathbb{P}(Z > 0) \geqslant \frac{\mathbb{E}[Z]^2}{\mathbb{E}[Z^2]}.$$

Hence, to show that an equitable k_j-colouring exists w.h.p., it suffices to show that as $j \to \infty$

$$\mathbb{E}[X^2_{n_j,k_j}]/\mathbb{E}[X_{n_j,k_j}]^2 \leqslant 1 + o(1). \tag{2.1}$$

This will be proved in Section 5, using a number of lemmas from [7] where the corresponding ratio was bounded for colourings of $\mathcal{G}(n,p)$ which could have two different colour class sizes. In that context, it was sufficient to bound the corresponding expression by $\exp(n/(\log^7 n))$.

We need much more accurate calculations in this paper to obtain the bound $1 + o(1)$, which only holds in $\mathcal{G}(n,m)$ (see the discussion below). We will also require all colour class sizes to be exactly equal, which simplifies calculations considerably. This enables us to use lemmas from [7], which in that paper either gave worse bounds or only held if the expected number of colourings was very large, to prove the bound $1 + o(1)$ in our case.

Remark 2.1. We conclude this section with some remarks on why the method fails for $\mathcal{G}(n,p)$, and also for $\mathcal{G}(n, \lfloor p\binom{n}{2}\rfloor)$ in the case $p > 1 - 1/e^2$.

Since $\mathbb{E}[X^2] \geqslant \mathbb{E}[X]^2$ holds for any random variable X, (2.1) implies $\mathbb{E}[X^2_{n_j,k_j}] \sim \mathbb{E}[X_{n_j,k_j}]^2$. Let π be an arbitrary fixed ordered k-equipartition; then it is not hard to show this is equivalent to

$$\mathbb{E}[X_{n_j,k_j} \mid \pi \text{ induces a valid colouring}] \sim \mathbb{E}[X_{n_j,k_j}]. \tag{2.2}$$

In other words, if we condition the random graph on the presence of one particular equitable colouring, this should not change the expected number of equitable colourings significantly; otherwise, the second moment method breaks down.

If $\mathcal{G}(n,p)$ is conditioned on the event that π induces a valid colouring, this decreases the expected number of edges by $\Theta(n \log n)$. This is more than the standard deviation of the number of edges in $\mathcal{G}(n,p)$, and other equitable colourings are now more likely simply because there are significantly fewer edges than in the unconditional random graph. This boosts the expectation of X_{n_j,k_j} (by a factor of size $\exp(\Theta(\log^2 n))$), so (2.2) does not hold.

A similar situation occurs if the equitable colouring contains very large colour classes. If $p > 1 - 1/e^2$, it follows from the work in [7] that an optimal equitable colouring of $\mathcal{G}_{n,\lfloor pN\rfloor}$ contains colour classes which are close to the likely value of the independence number of $\mathcal{G}_{n,\lfloor pN\rfloor}$. Since the expected number of such large independent sets is relatively small, conditioning on π increases their number significantly. This makes other colourings more likely, and increases the expectation of X_{n_j,k_j} so much that (2.2) does not hold. There is no particular reason to believe that the corresponding statement to Theorem 1.1 is true if $p \geqslant 1 - 1/e^2$.

3. The first moment

We start by analysing the first moment of $X_{n,k}$ and give a number of technical lemmas which will be needed to choose the sequence $(n_j)_{j\geqslant 1}$. Recall that

$$\gamma = \gamma(n) = 2\log_b n - 2\log_b \log_b n - 2\log_b 2. \tag{3.1}$$

We will only consider k-colourings where

$$k = \frac{n}{\gamma + O(1)}. \tag{3.2}$$

Let

$$\delta = \delta_{n,k} = \frac{n}{k} - \left\lfloor \frac{n}{k} \right\rfloor.$$

If k does not divide n, an ordered k-equipartition of n vertices consists of $k_L = k_L(n) = \delta_{n,k}k$ larger parts of size $\lceil n/k \rceil$, followed by $k_S = k_S(n) = (1 - \delta_{n,k})k$ smaller parts of size $\lfloor n/k \rfloor$. If k divides n, then all $k = k_S(n)$ parts are of size exactly $n/k = \lfloor n/k \rfloor$. By Stirling's formula $n! \sim \sqrt{2\pi n} n^n/e^n$, the total number of ordered k-equipartitions with $k = O(n/\log n)$ as in (3.2) is

$$P_{n,k} = \frac{n!}{\lceil n/k \rceil!^{k_L} \lfloor n/k \rfloor!^{k_S}} = k^n \exp\left(o(n)\right). \tag{3.3}$$

A k-equipartition induces a valid equitable colouring if and only if exactly

$$f = f_{n,k} = k_L \binom{\lceil n/k \rceil}{2} + k_S \binom{\lfloor n/k \rfloor}{2} = \frac{n(n/(k) - 1)}{2} + \frac{\delta_{n,k}(1 - \delta_{n,k})}{2} k \sim n \log_b n \tag{3.4}$$

forbidden edges within the parts of the partition are not present in G. Let

$$\varepsilon = Np - m \in [0, 1).$$

Since $f = f_{n,k} \sim n \log_b n$, using Stirling's formula, the probability that a given k-equipartition induces a valid colouring is

$$\begin{aligned}
\binom{N-f}{m} \Big/ \binom{N}{m} &= \frac{(N-f)!(qN+\varepsilon)!}{N!(qN-f+\varepsilon)!} \\
&\sim \frac{(N-f)^{N-f}(qN+\varepsilon)^{qN+\varepsilon}}{N^N(qN-f+\varepsilon)^{qN-f+\varepsilon}} \\
&= q^f \frac{(1-f/N)^{N-f}(1+\varepsilon/(qN))^{qN+\varepsilon}}{(1-f/(qN)+\varepsilon/(qN))^{qN-f+\varepsilon}}.
\end{aligned}$$

Using $\log(1+x) = x - x^2/2 + O(x^3)$ for $x \to 0$ and that $\varepsilon \in [0, 1)$ and $f = \Theta(n \log n)$,

$$\log\left(\binom{N-f}{m} \Big/ \binom{N}{m}\right) = f \log q - \frac{f^2 p}{2qN} + O\left(\frac{f^3}{N^2}\right).$$

Since $f^3/N^2 = o(1)$, the *expected number of ordered k-equipartitions which induce a valid colouring* is

$$\mu_{n,k} := \mathbb{E}[X_{n,k}] = P_{n,k}\left(\binom{N-f_{n,k}}{m} \Big/ \binom{N}{m}\right) \sim P_{n,k} q^{f_{n,k}} \exp\left(-\frac{f_{n,k}^2 p}{2qN}\right). \tag{3.5}$$

The expected number of *unordered* equitable partitions which induce valid colourings is

$$\bar{\mu}_{n,k} = \frac{\mu_{n,k}}{k_L! k_S!} \sim \frac{P_{n,k} q^{f_{n,k}}}{k_L! k_S!} \exp\left(-\frac{f_{n,k}^2 p}{2qN}\right). \tag{3.6}$$

If k is close to n/γ, we can approximate $\bar{\mu}_{n,k}$ with the following lemma, which is proved in the Appendix.

Lemma 3.1. *Given n and k and $x = O(1)$ such that*

$$k = \frac{n}{\gamma + x},$$

https://doi.org/10.1017/S0963548319000397 Published online by Cambridge University Press

then

$$\bar{\mu}_{n,k} = b^{-(x/2)n+o(n)}.$$

Corollary 3.2. *Given integers n and k such that* $|n/k - \gamma(n)| \leqslant 10$,

- *if* $\bar{\mu}_{n,k} \geqslant 1$, *then* $k \geqslant n/(\gamma(n) + o_n(1))$,
- *if* $\bar{\mu}_{n,k} \leqslant n$, *then* $k \leqslant n/(\gamma(n) + o_n(1))$.

In Lemma 4.1, we will pick a sequence $(n_j)_{j\geqslant 1}$ such that $\bar{\mu}_{n_j,k_j}$ starts tending to infinity just as all parts of a k_j-equipartition are of size exactly j. For this we will need some technical lemmas which examine how much $\mu_{n,k}$ and $\bar{\mu}_{n,k}$ change if we increase n or k by 1. Both are proved in the Appendix.

Lemma 3.3. *Given n and k such that* $k = n/(\gamma + O(1))$,

$$\frac{\bar{\mu}_{n,k+1}}{\bar{\mu}_{n,k}} \geqslant \exp\left(\Omega(\log n \log\log n)\right).$$

Lemma 3.4. *Given n and k such that* $k = n/(\gamma + O(1))$,

$$\frac{\mu_{n+1,k}}{\mu_{n,k}} = \Theta\left(\frac{\log n}{n}\right).$$

Note that for any graph G, $\chi_{=}(G) \geqslant \chi(G)$. So the lemma below follows directly from (1.1) and standard arguments about the equivalence of $\mathcal{G}(n, m)$ and $\mathcal{G}(n, p)$; it is also easy to prove directly through the first moment method (see [11]).

Lemma 3.5. *Let* $p \in (0, 1)$ *be constant and let* $G \sim \mathcal{G}(n, m)$ *with* $m = \lfloor pN \rfloor$. *Then w.h.p.*

$$\chi_{=}(G) \geqslant \frac{n}{\gamma(n) + o(1)}.$$

□

4. Choice of the subsequence

We are now ready to choose the sequence $(n_j)_{j\geqslant 1}$ from Theorem 1.1.

Proposition 4.1. *There is a strictly increasing sequence* $(n_j)_{j\geqslant 1}$ *such that the following holds for all* $j \geqslant 1$:

(a) $k_j := n_j/j \in \mathbb{Z}$,
(b) $\gamma(n_j) = j + o(1)$ *as* $j \to \infty$, *where* $\gamma(n_j) = 2\log_b n_j - 2\log_b \log_b n_j - 2\log_b 2$,
(c) $\bar{\mu}_{n_j,k_j} \to \infty$ *as* $j \to \infty$,
(d) *let* $G \sim \mathcal{G}(n_j, m_j)$ *with* $m_j = \left\lfloor p\binom{n_j}{2}\right\rfloor$; *then w.h.p. as* $j \to \infty$, *for all* $k \leqslant k_j - 1$, G *has no equitable* k-*colouring.*

Proof. It suffices to show that there is some constant $j_0 > 0$ and a strictly increasing sequence $(n_j)_{j\geqslant j_0}$ which fulfils (a)–(d) for all $j \geqslant j_0$: given such a sequence, by (b) n_j grows exponentially in j, so without loss of generality we can assume that $n_{j_0} \geqslant j_0$ and let $n_j = j$ for $1 \leqslant j < j_0$.

For $j \geqslant 1$, let n_j be the smallest multiple of j such that, letting $k_j = n_j/j$:

(1) $\gamma(n_j) \in [j - 10, j + 10]$,
(2) $\bar{\mu}_{n_j,k_j} \geqslant \log j$.

https://doi.org/10.1017/S0963548319000397 Published online by Cambridge University Press

Claim 1. *If j is large enough, n_j is well-defined.*

Proof. Roughly speaking, n_j is well-defined because $\gamma(n)$ is logarithmic in n and grows slowly as n is increased (so we are able to find a multiple $n = tj$ of j so that $\gamma(tj)$ lies in the interval $[j+1, j+2]$, say), and because by Lemma 3.1, letting $k = n/j$, $\bar{\mu}_{n,k}$ is exponentially large if $\gamma(n) - j$ is positive and bounded away from zero.

In more detail, fix j and consider the sequence $(\gamma(tj))_{t\geqslant j}$. If j is large enough and $t \geqslant j$, then

$$0 < \gamma((t+1)j) - \gamma(tj) \leqslant 3(\log_b((t+1)j) - \log_b(tj)) = 3\log_b\left(1 + \frac{1}{t}\right) < \frac{1}{2}. \tag{4.1}$$

Furthermore, if j is large enough, then $\gamma(j^2) \leqslant j$ and $\gamma(tj) \to \infty$ as $t \to \infty$. Therefore, there is an integer $t_0 \geqslant j$ such that

$$\gamma(t_0 j) \in [j+1, j+2].$$

By Lemma 3.1, letting $n' = t_0 j$,

$$\bar{\mu}_{n',t_0} \geqslant b^{(\frac{1}{2}+o(1))n'} \geqslant \log n' \geqslant \log j.$$

Therefore, n' is a multiple of j such that the two conditions (1) and (2) from the definition of n_j are fulfilled, so n_j is well-defined. □

The definition of n_j immediately implies (a) and (c). We now show that (b) holds.

Claim 2. *As $j \to \infty$, $\gamma(n_j) = j + o(1)$.*

Proof. By the definition of n_j, if j is large enough we have

$$\bar{\mu}_{n_j,k_j} \geqslant \log j \geqslant 1.$$

Since $k_j = n/j$, together with Corollary 3.2 this implies that $j \leqslant \gamma(n_j) + o(1)$. In particular, this implies that n_j grows (at least) exponentially in j.

For the other direction, first note that by (4.1), for j large enough,

$$\gamma(n_j) \geqslant \gamma(n_j - j) \geqslant \gamma(n_j) - \frac{1}{2} \geqslant j - o(1) - \frac{1}{2} \geqslant j - 1.$$

In particular,

$$\gamma(n_j - j) \in [j-10, j+10].$$

So $n_j - j$ is another multiple of j for which part (1) of the definition of n_j holds. As n_j is defined as the smallest multiple of j such that (1) and (2) hold, we must have

$$\bar{\mu}_{n_j-j,k_j-1} < \log j. \tag{4.2}$$

Since n_j grows (at least) exponentially in j, for j large enough we have $\log j \leqslant n_j - j$, and therefore by (4.2),

$$\bar{\mu}_{n_j-j,k_j-1} < n_j - j.$$

By Corollary 3.2,

$$\frac{n_j - j}{j} = k_j - 1 \leqslant \frac{n_j - j}{\gamma(n_j - j) + o(1)},$$

so $j \geqslant \gamma(n_j - j) + o(1)$. Note that

$$\gamma(n_j) - \gamma(n_j - j) = O\left(\log\left(\frac{n_j}{n_j - j}\right)\right) = O(\log(1 + o(1))) = o(1),$$

https://doi.org/10.1017/S0963548319000397 Published online by Cambridge University Press

and thus

$$j \geqslant \gamma(n_j) + o(1)$$

as required, so (b) holds. □

By (b), $\gamma(n_j) = j + o(1)$, and hence n_j grows exponentially in j. In particular, this implies that n_j is strictly increasing if j is large enough.

It only remains to prove (d). By Lemma 3.5, w.h.p. $G \sim \mathcal{G}(n_j, m_j)$ has no equitable colouring with fewer than $n_j/(\gamma(n_j) + o(1))$ colours. To prove that (d) holds, it is therefore sufficient to show that w.h.p. G has no equitable colouring with more than, say, $n_j/(\gamma(n_j) + 1)$ and at most $k_j - 1$ colours. We first show that the expected number $\bar{\mu}_{n_j,k_j-1}$ of (unordered) equitable colourings of $\mathcal{G}(n_j, m_j)$ with $k_j - 1$ colours tends to 0.

Claim 3. *If j is large enough, then*

$$\bar{\mu}_{n_j,k_j-1} \leqslant \frac{1}{j}.$$

Proof. As $k_j = n_j/j$, an equitable partition of n_j vertices into $k_j - 1$ parts consists of exactly j larger parts of size $j + 1$ and $k_j - 1 - j$ smaller parts of size j, so

$$\bar{\mu}_{n_j,k_j-1} = \frac{\mu_{n_j,k_j-1}}{j!(k_j - 1 - j)!}. \tag{4.3}$$

By Lemma 3.4 and as $k_j - 1$ divides $n_j - j$, we have

$$\mu_{n_j,k_j-1} = \mu_{n_j-j,k_j-1}\left(\Theta\left(\frac{\log n_j}{n_j}\right)\right)^j = \bar{\mu}_{n_j-j,k_j-1}\,(k_j-1)!\left(\Theta\left(\frac{\log n_j}{n_j}\right)\right)^j.$$

Note that by (4.2), $\bar{\mu}_{n_j-j,k_j-1} < \log j$, so with (4.3),

$$\bar{\mu}_{n_j,k_j-1} < \log j\,\frac{(k_j-1)!}{j!(k_j-1-j)!}\left(\Theta\left(\frac{\log n_j}{n_j}\right)\right)^j = \log j\binom{k_j-1}{j}\left(\Theta\left(\frac{\log n_j}{n_j}\right)\right)^j.$$

As

$$\binom{k_j-1}{j} < \left(\frac{e(k_j-1)}{j}\right)^j$$

and $k_j = \Theta(n_j/\log n_j)$, if j is large enough this gives

$$\bar{\mu}_{n_j,k_j-1} \leqslant \log j\left(\Theta\left(\frac{k_j \log n_j}{jn_j}\right)\right)^j = \log j(\Theta(1/j))^j < 1/j.$$ □

By Lemma 3.3 and Claim 3, the expected number of unordered equitable colourings with between $n_j/(\gamma(n_j) + 1)$ and $k_j - 1$ colours is

$$\sum_{n/(\gamma(n_j)+1)\leqslant k\leqslant k_j-1} \bar{\mu}_{n_j,k} = O(\bar{\mu}_{n_j,k_j-1}) = O(1/j) = o(1),$$

so w.h.p. G has no such colouring. This completes the proof of Proposition 4.1. □

5. The second moment

For the proof of Theorem 1.1, it remains to show that for the sequence $(n_j)_{j\geqslant 1}$ from Proposition 4.1,

$$\mathbb{E}[X^2_{n_j,k_j}]/\mathbb{E}[X_{n_j,k_j}]^2 \leqslant 1 + o(1) \text{ as } j \to \infty. \tag{5.1}$$

https://doi.org/10.1017/S0963548319000397 Published online by Cambridge University Press

We will be able to re-use large parts of the calculations from [7], but will need to be more accurate here than in [7], where it was sufficient to bound the corresponding ratio by $\exp(n/\log^8 n)$ rather than $1 + o(1)$.

Let $j \geqslant 1$, $N_j = \binom{n_j}{2}$, $m_j = \lfloor pN_j \rfloor$ and $f_j = f_{n_j,k_j}$ as in (3.4) and $P_j = P_{n_j,k_j}$ as in (3.3), and $G \sim \mathcal{G}(n_j, m_j)$. To simplify notation, we will omit the indices of n_j, m_j, k_j and so on when the context is clear.

Note that since $j = n/k$ is an integer, we can simplify the expressions for P and f:

$$P = \frac{n!}{j!^k} \quad \text{and} \quad f = \frac{n(j-1)}{2}. \tag{5.2}$$

As

$$X_{n,k} = \sum_{\pi \text{ ordered } k\text{-equipartition}} \mathbb{1}_{\pi \text{ induces a valid colouring}},$$

by linearity of the expectation,

$$\mathbb{E}[X_{n,k}^2] = \sum_{\pi_1,\pi_2 \text{ ordered } k\text{-equipartitions}} \mathbb{P}(\text{both } \pi_1 \text{ and } \pi_2 \text{ induce proper colourings}).$$

The terms in the last sum may vary considerably depending on how similar π_1 and π_2 are. To quantify this, we define the *overlap sequence* $\mathbf{r} = \mathbf{r}(\pi_1, \pi_2) = (r_i)_{i=2}^{j}$ of π_1 and π_2. We let r_i denote the number of pairs of parts (with the first part in π_1 and the second part in π_2) which intersect in exactly i vertices.

Conversely, given an *overlap sequence* $\mathbf{r}$, let $P_{\mathbf{r}}$ denote the number of ordered pairs π_1, π_2 which overlap according to $\mathbf{r}$. We call an intersection of size at least 2 between two parts an *overlap block*. Let

$$v = v(\mathbf{r}) = \sum_{i=2}^{j} i r_i \tag{5.3}$$

be the number of *vertices involved in the overlap*, and denote the proportion of such vertices in the graph by

$$\rho = v/n.$$

If π_1 and π_2 overlap according to $\mathbf{r}$, then they share exactly

$$d = d(\mathbf{r}) = \sum_{i=2}^{j} r_i \binom{i}{2} \tag{5.4}$$

forbidden edges. Therefore, π_1 and π_2 with overlap sequence $\mathbf{r}$ both induce valid colourings at the same time if and only if exactly $2f - d(\mathbf{r})$ given forbidden edges are not present in G, so by (3.5),

$$\mathbb{E}[X_{n,k}^2] = \sum_{\mathbf{r}} P_{\mathbf{r}} \left(\binom{N - 2f + d(\mathbf{r})}{m} \Big/ \binom{N}{m} \right) = \mu_{n,k}^2 \sum_{\mathbf{r}} \frac{P_{\mathbf{r}}}{P^2} \cdot \frac{\binom{N-2f+d}{m}\binom{N}{m}}{\binom{N-f}{m}^2}.$$

Let

$$Q_{\mathbf{r}} = \frac{P_{\mathbf{r}}}{P^2},$$
$$S_{\mathbf{r}} = \left(\binom{N - 2f + d(\mathbf{r})}{m} \binom{N}{m} \right) \Big/ \binom{N-f}{m}^2. \tag{5.5}$$

https://doi.org/10.1017/S0963548319000397 Published online by Cambridge University Press

Then, to prove (5.1), we need to show that as $j \to \infty$,

$$\sum_{\mathbf{r}} Q_{\mathbf{r}} S_{\mathbf{r}} \leqslant 1 + o(1). \tag{5.6}$$

We will determine the asymptotic value of $S_{\mathbf{r}}$ in Section 5.1. To bound the sum (5.6), we will split up the calculations into three different ranges in Sections 5.2–5.4, corresponding to the different cases of the calculations in [7]. In the first case we need to obtain a much more accurate bound than in [7], and this is where most of the work in calculating the second moment lies. The other two cases follow more or less directly from lemmas in [7].

More specifically, in Section 5.2 we will consider all those overlap sequences $\mathbf{r}$ where $\rho = v/n \leqslant c$ for some constant $c > 0$. Most pairs of k-equipartitions belong in this range, and this is also where the main contribution to (5.6) comes from. While in [7] it was sufficient to bound the contribution from this range to the equivalent of (5.6) by $\exp(n/\log^8 n)$, in this paper we will show how to obtain the sharpest possible bound $1 + o(1)$ (which only holds in $\mathcal{G}(n, m)$).

In Section 5.3 we consider those $\mathbf{r}$ where there are at least $v = cn$ vertices involved in the overlap, but there are either many vertices not in the overlap or many small overlap blocks. We will use a simplification of arguments in [7] to show that the contribution from these overlap sequences $\mathbf{r}$ to (5.6) is $o(1)$.

Finally, Section 5.4 concerns overlap sequences $\mathbf{r}$ corresponding to pairs of partitions which are very similar to each other. A lemma from [7] shows that if all colour classes are of exactly the same size, then the contribution from this range of $\mathbf{r}$ is $O(1/\bar{\mu}_{n,k})$, which tends to 0 since $\bar{\mu}_{n,k} \to \infty$. This will conclude the proof of Theorem 1.1.

5.1 Asymptotics of $S_{\mathbf{r}}$

Consider an overlap sequence $\mathbf{r}$. Again letting $\varepsilon = Np - m \in [0, 1]$, and $d = d(\mathbf{r})$,

$$S_{\mathbf{r}} = \frac{(N-2f+d)!\, N!\, (N-m-f)!^2}{(N-f)!^2\, (N-m)!\, (N-m-2f+d)!} = \frac{(N-2f+d)!\, N!\, (qN+\varepsilon-f)!^2}{(N-f)!^2\, (qN+\varepsilon)!\, (qN+\varepsilon-2f+d)!}.$$

Since $d \leqslant f = O(n \log n) = o(N)$, applying Stirling's formula $n! \sim \sqrt{2\pi n}\, n^n/e^n$ gives

$$\begin{aligned} S_{\mathbf{r}} &\sim \frac{(N-2f+d)^{N-2f+d} N^N (qN+\varepsilon-f)^{2qN+2\varepsilon-2f}}{(N-f)^{2N-2f}(qN+\varepsilon)^{qN+\varepsilon}(qN+\varepsilon-2f+d)^{qN+\varepsilon-2f+d}} \\ &= q^{-d} \cdot \frac{(1-(2f-d)/N)^{N-2f+d}(1-(f-\varepsilon)/(qN))^{2qN+2\varepsilon-2f}}{(1-f/N)^{2N-2f}(1+\varepsilon/(qN))^{qN+\varepsilon}(1-(2f-d-\varepsilon)/(qN))^{qN+\varepsilon-2f+d}}. \end{aligned}$$

Using $\log(1+x) = x - x^2/2 + O(x^3)$ for $x \to 0$, and as $d \leqslant f = O(n \log n)$ and $\varepsilon < 1$, we get

$$\log S_{\mathbf{r}} = -d \log q + \frac{-p(d^2 + 2f^2 - 4df)}{2qN} + o(1). \tag{5.7}$$

Details of this calculation can be found in the Appendix. Therefore,

$$S_{\mathbf{r}} \sim q^{-d} \exp\left(-\frac{p(d^2+2f^2-4df)}{2qN}\right) = b^d \exp\left(-\frac{p(d^2+2f^2-4df)}{2qN}\right), \tag{5.8}$$

where $d = d(\mathbf{r})$ is given in (5.4).

https://doi.org/10.1017/S0963548319000397 Published online by Cambridge University Press

5.2 The typical overlap case

In this section we bound the contribution to (5.6) from overlap sequences $\mathbf{r}$ corresponding to pairs of equipartitions with relatively few vertices involved in the overlap. Most pairs of equipartitions fall in this range. To see this, note that if we pick an equipartition with $k = \Theta(n/\log n)$ parts uniformly at random, the probability that $t = O(1)$ given vertices are in the same part is of order $O((\log n/n)^{t-1})$. Therefore, if we sample two equipartitions independently and uniformly at random, the expected number of joint forbidden edges is of order $O(\log^2 n)$, and the expected number of triples of vertices which are in the same part in both equipartitions is $O(\log^4 n/n) = o(1)$. So most pairs of equipartitions share $O(\log^2 n)$ (disjoint) forbidden edges and have no larger overlap blocks in common; in our notation, this means $r_2 = O(\log^2 n)$ and $r_3 = r_4 = \cdots = r_j = 0$. We will cover a much larger range of sequences $\mathbf{r}$ here, namely all those $\mathbf{r}$ where only a certain constant fraction cn of all vertices is contained in overlap clusters, that is, $v = \sum_i ir_i \leqslant cn$. As this is where the main contribution of $1 + o(1)$ to (5.6) comes from, our calculations will need to be quite accurate, at least in the case $r_2 = O(\log^2 n)$ and $r_i = 0$ for $i \geqslant 3$.

For any $\mathbf{r}$, Lemma 9 from [7] gives

$$Q_{\mathbf{r}} b^d \lesssim \prod_{i=2}^{j} \left(\frac{1}{r_i!} \left(\frac{e^{\rho i} b^{\binom{i}{2}} k^2 j!^2}{n^i i! (j-i)!^2} \right)^{r_i} \right) \exp\left(-\frac{1}{2} \left(\frac{n-v}{k} - 1 \right)^2 \right).$$

(Note that the quantity x_0 from [7] is 0 as $p < 1 - 1/e^2$.) Together with (5.8), this gives

$$Q_{\mathbf{r}} S_{\mathbf{r}} \lesssim \prod_{i=2}^{j} \left(\frac{1}{r_i!} \left(\frac{e^{\rho i} b^{\binom{i}{2}} k^2 j!^2}{n^i i! (j-i)!^2} \right)^{r_i} \right) \exp\left(-\frac{1}{2} \left(\frac{n-v}{k} - 1 \right)^2 - \frac{p(d^2 + 2f^2 - 4df)}{2qN} \right).$$

Letting

$$T_i := \frac{e^{\rho i} b^{\binom{i}{2}} k^2 j!^2}{n^i i! (j-i)!^2},$$

we have

$$Q_{\mathbf{r}} S_{\mathbf{r}} \lesssim \exp\left(-\frac{1}{2} \left(\frac{n-v}{k} - 1 \right)^2 - \frac{p(d^2 + 2f^2 - 4df)}{2qN} \right) \prod_{i=2}^{j} \frac{T_i^{r_i}}{r_i!}. \tag{5.9}$$

If $\rho = \rho(\mathbf{r}) = v/n$ is small enough, the terms $T_3, \ldots, T_j$ are all small. This is the content of the following lemma, which is proved in the Appendix. Noting that $p < 1 - 1/e^2$ and therefore $\log b < 2$, let

$$c = \frac{1}{2} \left(1 - \frac{\log b}{2} \right) \in (0, 1). \tag{5.10}$$

Lemma 5.1. *If j (and therefore $n = n_j$) is large enough and $\rho \leqslant c$, then for all $3 \leqslant i \leqslant j$,*

$$T_i \leqslant n^{-\tilde{c}},$$

where

$$\tilde{c} = \min\left(\frac{1}{2} \left(\frac{1}{\log b} - \frac{1}{2} \right), \frac{1}{2} \right) \in (0, 1).$$

In this section we will bound the contribution to (5.6) from all sequences $\mathbf{r}$ so that Lemma 5.1 applies. So let

$$\mathcal{R}_1 = \{ \mathbf{r} \mid \rho = \rho(\mathbf{r}) \leqslant c \}.$$

https://doi.org/10.1017/S0963548319000397 Published online by Cambridge University Press

For $\mathbf{r} \in \mathcal{R}_1$, the terms $T_3, \ldots, T_j$ are small, leaving the main contribution to (5.6) to come from the terms T_2. Using Lemma 5.1, we bound the terms $T_3, \ldots, T_j$ by a single term depending only on R_3, where

$$R_3 = \sum_{i=3}^{j} r_i.$$

By Lemma 5.1, (5.9) and the definition (5.2) of f, for $\mathbf{r} \in \mathcal{R}_1$,

$$Q_{\mathbf{r}} S_{\mathbf{r}} \lesssim \exp\left(-\frac{1}{2}\left(\frac{n-v}{k}-1\right)^2 - \frac{p(d^2+2f^2-4df)}{2qN}\right)\frac{T_2^{r_2}}{r_2!} n^{-\tilde{c}R_3}. \quad (5.11)$$

Before summing (5.11) over $\mathbf{r} \in \mathcal{R}_1$, we give a simple lemma which will enable us to take the sum over R_3 rather than $r_3, \ldots, r_j$.

Lemma 5.2. *Given R_3, there are at most $(2e \log_b n)^{R_3}$ ways to select $r_3, \ldots, r_j$ so that $\sum_{i=3}^{j} r_i = R_3$.*

Proof. Since $j \leqslant 2 \log_b n$, there are at most

$$\binom{R_3 + j - 3}{R_3} \leqslant \left(\frac{e(R_3 + j - 3)}{R_3}\right)^{R_3} \leqslant (e(1 + j - 3))^{R_3} \leqslant (2e \log_b n)^{R_3}$$

ways to write R_3 as an ordered sum of $j-2$ non-negative summands. □

We will see below that if v and d are not too large, the term T_2 in (5.11) is roughly $(b/2)(j-1)^2$. Also, the exponential term in (5.11) is then roughly $\exp(-(b/2)(j-1)^2)$; this will exactly cancel out $\sum_{r_2}(T_2^{r_2}/r_2!)$ when summing (5.11) over $\mathbf{r}$, giving an overall sum of $1+o(1)$. Before giving the details of this calculation, we need to handle the case where either v or d are very large (too large for the approximations of T_2 and the exponential term to hold), and show that the contribution from this case is $o(1)$.

So let $\mathcal{R}_1^{\mathrm{ex}}$ be the set of all $\mathbf{r} \in \mathcal{R}_1$ where $v = v(\mathbf{r}) \geqslant n/(\log^3 n)$ or $d = d(\mathbf{r}) \geqslant n/(\log^3 n)$.

Lemma 5.3.

$$\sum_{\mathbf{r} \in \mathcal{R}_1^{\mathrm{ex}}} Q_{\mathbf{r}} S_{\mathbf{r}} = o(1).$$

Proof. Again we assume throughout that j is large enough for all estimates to be valid. Let $\mathbf{r} \in \mathcal{R}_1^{\mathrm{ex}}$. As $d \leqslant f = O(n \log n)$ and $n/k = j = O(\log n)$,

$$\exp\left(-\frac{1}{2}\left(\frac{n-v}{k}-1\right)^2 - \frac{p(d^2+2f^2-4df)}{2qN}\right) = \exp(O(\log^2 n)), \quad (5.12)$$

and

$$T_2 \leqslant \frac{e^2 b^{\binom{2}{2}} k^2 j^4}{n^2 2!} = \Theta(\log^2 n). \quad (5.13)$$

Since

$$v = \sum_{i=2}^{j} i r_i \leqslant 2r_2 + 2R_3 \log_b n \quad \text{and} \quad d = \sum_{i=2}^{j} \binom{i}{2} r_i \leqslant r_2 + 2R_3 \log_b^2 n,$$

if $\mathbf{r} \in \mathcal{R}_1^{\mathrm{ex}}$, then either $r_2 \geqslant n/\log^6 n$ or $R_3 \geqslant n/\log^6 n$.

https://doi.org/10.1017/S0963548319000397 Published online by Cambridge University Press

Case 1. $r_2 \geqslant n/\log^6 n$. Then from (5.11), (5.12) and (5.13), and as $r_2! \geqslant r_2^{r_2}/e^{r_2}$,

$$\begin{aligned} Q_{\mathbf{r}} S_{\mathbf{r}} &\lesssim \exp(O(\log^2 n)) \frac{(\Theta(\log^2 n))^{r_2}}{r_2!} n^{-\tilde{c}R_3} \\ &\leqslant \exp(O(\log^2 n)) \left(\frac{\Theta(\log^2 n)}{r_2}\right)^{r_2} n^{-\tilde{c}R_3} \\ &\lesssim \exp(O(\log^2 n)) \left(\frac{\log^9 n}{n}\right)^{r_2} n^{-\tilde{c}R_3}. \end{aligned}$$

With Lemma 5.2, we can bound the sum of (5.11) over those $\mathbf{r}$ with $r_2 \geqslant n/\log^6 n$ crudely by a sum over r_2 and R_3:

$$\begin{aligned} \sum_{\substack{\mathbf{r} \in \mathcal{R}_1^{\mathrm{ex}} \\ r_2 \geqslant n/\log^6 n}} Q_{\mathbf{r}} S_{\mathbf{r}} &\leqslant \exp(O(\log^2 n)) \sum_{r_2 \geqslant n/\log^6 n,\ R_3} \left(\left(\frac{\log^9 n}{n}\right)^{r_2} \left(\frac{2e \log_b n}{n^{\tilde{c}}}\right)^{R_3} \right) \\ &= o(1). \end{aligned}$$

Case 2. $R_3 \geqslant n/\log^6 n$. By Lemma 5.2, (5.12) and (5.13),

$$\begin{aligned} \sum_{\substack{\mathbf{r} \in \mathcal{R}_1^{\mathrm{ex}} \\ R_3 \geqslant n/\log^6 n}} Q_{\mathbf{r}} S_{\mathbf{r}} &\leqslant \exp(O(\log^2 n)) \sum_{r_2} \frac{O(\log^2 n)^{r_2}}{r_2!} \sum_{R_3 \geqslant n/\log^6 n} \left(\frac{2e \log_b n}{n^{\tilde{c}}}\right)^{R_3} \\ &= \exp(O(\log^2 n)) \sum_{R_3 \geqslant n/\log^6 n} \left(\frac{2e \log_b n}{n^{\tilde{c}}}\right)^{R_3} \\ &= o(1). \end{aligned}$$

□

We will now sum (5.11) for all $\mathbf{r} \in \mathcal{R}_1 \setminus \mathcal{R}_1^{\mathrm{ex}}$. Note that $p/q = b-1$ and that by (5.2), $f^2/N = \frac{1}{2}(j-1)^2 + o(1)$, so if $v < n/(\log^3 n)$ and $d < n/(\log^3 n)$,

$$\begin{aligned} \exp\left(-\frac{1}{2}\left(\frac{n-v}{k} - 1\right)^2 - \frac{p(d^2 + 2f^2 - 4df)}{2qN}\right) &\sim \exp\left(-\frac{1}{2}\left(\frac{n}{k} - 1\right)^2 - \frac{pf^2}{qN}\right) \\ &\sim \exp\left(-\frac{b}{2}(j-1)^2\right). \end{aligned} \tag{5.14}$$

Note that for $\mathbf{r} \in \mathcal{R}_1 \setminus \mathcal{R}_1^{\mathrm{ex}}$, $\rho = v/n < 1/(\log^3 n)$, so

$$T_2 = \frac{e^{2\rho} b^{\binom{2}{2}} k^2 j!^2}{n^2 2!(j-2)!^2} = e^{2\rho} \frac{b}{2}(j-1)^2 \leqslant e^{2/(\log^3 n)} \frac{b}{2}(j-1)^2 =: T.$$

Therefore, from (5.11) and together with (5.14) and Lemma 5.2,

$$\begin{aligned} \sum_{\mathbf{r} \in \mathcal{R}_1 \setminus \mathcal{R}_1^{\mathrm{ex}}} Q_{\mathbf{r}} S_{\mathbf{r}} &\lesssim \exp\left(-\frac{b}{2}(j-1)^2\right) \sum_{r_2, R_3 \geqslant 0} \frac{T^{r_2}}{r_2!} (2en^{-\tilde{c}} \log_b n)^{R_3} \\ &\sim \exp\left(-\frac{b}{2}(j-1)^2\right) \sum_{r_2 \geqslant 0} \frac{T^{r_2}}{r_2!} \\ &= \exp\left(-\frac{b}{2}(j-1)^2 + T\right) \end{aligned}$$

https://doi.org/10.1017/S0963548319000397 Published online by Cambridge University Press

$$= \exp\left(-\frac{b}{2}(j-1)^2(1-e^{2/(\log^3 n)})\right)$$
$$= \exp\left(-\frac{b}{2}(j-1)^2 O(\log^{-3} n)\right)$$
$$= 1 + o(1),$$

since $j = O(\log n)$. Together with Lemma 5.3, it follows that the contribution from $\mathcal{R}_1$ to (5.6) is $1 + o(1)$.

5.3 The intermediate overlap case

Now let $0 < c' < 1$ be an arbitrary constant. As in [7], we let $\mathcal{R}_2^{c'}$ denote the set of all overlap sequences which are not in $\mathcal{R}_1$ and where there are either at least $c'n$ vertices not involved in the overlap at all, or at least $c'n$ vertices in 'small' overlap blocks of size at most 0.6γ (the maximum size of any overlap block is $j = \gamma + o(1)$), that is,

$$\mathcal{R}_2^{c'} = \left\{ \mathbf{r} \,\middle|\, \rho > c \wedge \left(\sum_{2 \leqslant i \leqslant 0.6\gamma} i r_i \geqslant c'n \vee \rho \leqslant 1 - c' \right) \right\}.$$

We will simplify arguments from [7] to show that the contribution to (5.6) from overlap sequences $\mathbf{r} \in \mathcal{R}_2^{c'}$ is $o(1)$. Fix an arbitrary ordered k-equipartition π_1, and let

$$\mathcal{P}_2^{c'} = \{\text{ordered } k\text{-equipartitions } \pi_2 \text{ such that } \mathbf{r}(\pi_1, \pi_2) \in \mathcal{R}_2^{c'}\}.$$

Following [7], for an overlap sequence $\mathbf{r}$, we let $P'_{\mathbf{r}}$ denote the number of ordered k-equipartitions with overlap $\mathbf{r}$ with π_1. By symmetry, $P'_{\mathbf{r}}$ does not depend on the choice of π_1, and in particular $P_{\mathbf{r}} = PP'_{\mathbf{r}}$. By the definition (5.5) of $Q_{\mathbf{r}}$,

$$Q_{\mathbf{r}} = \frac{P_{\mathbf{r}}}{P^2} = \frac{P'_{\mathbf{r}}}{P}.$$

Using (5.8) and that by (3.3), $P = k^n \exp(o(n))$,

$$\begin{aligned} \sum_{\mathbf{r} \in \mathcal{R}_2^{c'}} Q_{\mathbf{r}} S_{\mathbf{r}} &= \sum_{\mathbf{r} \in \mathcal{R}_2^{c'}} \frac{P'_{\mathbf{r}}}{P} b^d \exp(o(n)) \\ &= \sum_{\pi_2 \in \mathcal{P}_2^{c'}} P^{-1} b^{d(\pi_1, \pi_2)} \exp(o(n)) \\ &= \sum_{\pi_2 \in \mathcal{P}_2^{c'}} k^{-n} b^{d(\pi_1, \pi_2)} \exp(o(n)). \end{aligned} \tag{5.15}$$

In [7], three sets of partitions $\mathcal{P}^{\mathrm{I}}$, $\mathcal{P}^{\mathrm{II}}$ and $\mathcal{P}^{\mathrm{III}}$ were defined. We will not repeat the exact definitions of $\mathcal{P}^{\mathrm{I}}$ and $\mathcal{P}^{\mathrm{II}}$ here, as they are quite technical; they can be found in the Appendix. Note that

$$\mathcal{P}^{\mathrm{III}} := \{\text{ordered } k\text{-equipartitions } \pi_2 \text{ such that } c < \rho \leqslant 1 - c'\} \setminus \mathcal{P}^{\mathrm{I}} \setminus \mathcal{P}^{\mathrm{II}}. \tag{5.16}$$

In Lemma 13 of [7] – which is also given in the Appendix – it was proved that

$$\mathcal{P}_2^{c'} \subset \mathcal{P}^{\mathrm{I}} \cup \mathcal{P}^{\mathrm{II}} \cup \mathcal{P}^{\mathrm{III}}, \tag{5.17}$$

and Lemma 15 showed that

$$\sum_{\pi_2 \in \mathcal{P}^{\mathrm{I}} \cup \mathcal{P}^{\mathrm{II}}} k^{-n} b^{d(\pi_1, \pi_2)} \exp(o(n)) = o(1). \tag{5.18}$$

https://doi.org/10.1017/S0963548319000397 Published online by Cambridge University Press

Therefore, to bound the contribution of $\mathcal{R}_2^{c'}$ to (5.6) by $o(1)$, we only need to consider the partitions $\pi_2 \in \mathcal{P}^{\mathrm{III}}$. In [7], two types of vertices involved in the overlap were distinguished: those v_1 vertices which are in parts of size a in π_1, and those v_2 vertices which are in parts of size at most $a-1$ in π_1, where

$$a = \lfloor \gamma \rfloor + 1.$$

With this notation, of course $v = v_1 + v_2$. Similarly, there are d_1 shared forbidden edges in parts of size a in π_1, and d_2 shared forbidden edges in parts of size at most $a-1$ in π_1, so that $d_1 + d_2 = d$. Fixing integers v_1, v_2, d_1 and d_2 and letting

$$\mathcal{P}'(v_1, v_2, d_1, d_2) = \{\pi_2 \in \mathcal{P}^{\mathrm{III}} \mid v_i(\pi_1, \pi_2) = v_i, d_i(\pi_1, \pi_2) = d_i, i = 1, 2\}, \tag{5.19}$$

Lemma 17 of [7] states that

$$\sum_{\pi_2 \in \mathcal{P}'(v_1, v_2, d_1, d_2)} k^{-n} b^{d(\pi_1, \pi_2)} \leqslant b^{n(1-\rho)\log_b(1-\rho) + v_1/2 - (\Delta v)/2} \exp(o(n)).$$

In the context of this paper, this last expression can be simplified as all parts are of size exactly j. As $j = \gamma + o(1)$, $a = \lfloor \gamma \rfloor + 1$ is either $j+1$ or j. In the first case, we have $v_1 = 0$, $v_2 = v$ and $\Delta = \gamma - \lfloor \gamma \rfloor = o(1)$. In the second case, we have $v_1 = v$, $v_2 = 0$ and $\Delta = 1 - o(1)$. In both cases,

$$\sum_{\pi_2 \in \mathcal{P}'(v_1, v_2, d_1, d_2)} k^{-n} b^{d(\pi_1, \pi_2)} \leqslant b^{n(1-\rho)\log_b(1-\rho)} \exp(o(n)).$$

By the definitions (5.16) and (5.19), we have $c \leqslant \rho \leqslant 1 - c'$, so

$$\sum_{\pi_2 \in \mathcal{P}'(v_1, v_2, d_1, d_2)} k^{-n} b^{d(\pi_1, \pi_2)} \leqslant b^{-c'' n},$$

where

$$c'' = \frac{1}{2} \min(|(1-c)\log(1-c)|, |c' \log c'|) > 0.$$

Bounding crudely, there are at most n^2 choices for the values of v_1, v_2 and (as $d \leqslant f \leqslant n^2$) at most n^4 choices for d_1 and d_2, so

$$\sum_{\pi_2 \in \mathcal{P}^{\mathrm{III}}} k^{-n} b^{d(\pi_1, \pi_2)} \leqslant n^6 b^{-c'' n} = o(1).$$

Together with (5.15), (5.17) and (5.18), this gives

$$\sum_{\mathbf{r} \in \mathcal{R}_2^{c'}} Q_{\mathbf{r}} S_{\mathbf{r}} = o(1).$$

5.4 The large overlap case

Given $c' \in (0, 1)$, let

$$\mathcal{R}_3^{c'} = \left\{ \mathbf{r} \,\middle|\, \rho > 1 - c', \sum_{2 \leqslant i \leqslant 0.6\gamma} i r_i \leqslant c' n \right\}$$

be the remaining set of overlap sequences $\mathbf{r}$, where c' is still an arbitrary constant. To finish the proof of Theorem 1.1, it remains to show that we can choose $c' \in (0, 1)$ so that the contribution from the overlap sequences $\mathbf{r} \in \mathcal{R}_3^{c'}$ to (5.6) is $o(1)$. This is a direct consequence of the following lemma from [7].

https://doi.org/10.1017/S0963548319000397 Published online by Cambridge University Press

Lemma 5.4 (Lemma 22 of [7]). *There is a constant $c' > 0$ such that*

$$\sum_{\mathbf{r}\in\mathcal{R}_3^{c'}} Q_{\mathbf{r}} b^d = O\left(\frac{k_S!k_L!}{Pq^f}\binom{k}{k_S}\right) = O\left(\frac{k!}{Pq^f}\right).$$

As in our case all k colour classes are of the same size, by (3.6) we have

$$\bar{\mu} = \frac{\mu}{k!} = \frac{Pq^f}{k!}\exp\left(-\frac{f^2p}{2qN}\right),$$

and thus

$$\sum_{\mathbf{r}\in\mathcal{R}_3^{c'}} Q_{\mathbf{r}} b^d = O\left(\frac{1}{\bar{\mu}}\right)\exp\left(-\frac{f^2p}{2qN}\right).$$

By (5.8) and since $d \leqslant f$,

$$\begin{aligned}\sum_{\mathbf{r}\in\mathcal{R}_3^{c'}} Q_{\mathbf{r}} S_{\mathbf{r}} &= O\left(\frac{1}{\bar{\mu}}\right)\exp\left(-\frac{p(3f^2-4df+d^2)}{2qN}\right)\\ &= O\left(\frac{1}{\bar{\mu}}\right)\exp\left(-\frac{p((2f-d)^2-f^2)}{2qN}\right)\\ &= O\left(\frac{1}{\bar{\mu}}\right).\end{aligned}$$

This expression is $o(1)$ as soon as $\bar{\mu}\to\infty$, which is indeed the case by our choice of $(n_j)_{j\geqslant 1}$: see Proposition 4.1(c). This completes the proof of Theorem 1.1. □

6. Open problems

While Theorem 1.1 establishes one point concentration of $\chi_=(\mathcal{G}(n,m))$ on a subsequence of the integers, we have no general result which holds for all n. For the normal chromatic number, Shamir and Spencer [13] showed concentration on at most about $\sqrt{n}$ integers for every function $p = p(n)$. It would be interesting to obtain a more general concentration result for the equitable chromatic number which holds for all n, or indeed for different functions $p(n)$ or $m(n)$.

We also have not addressed the equitable chromatic threshold of dense random graphs in Theorem 1.1. It may be possible to extend the second moment arguments to equitable k'-colourings where $k' \geqslant k$. However, this would only show that an equitable k'-colouring exists w.h.p. for one particular value of k', not that equitable k'-colourings exist w.h.p. for all such values k' *simultaneously*. To prove such a statement via the second moment method, we would need good bounds on the rate of convergence of $\mathbb{E}[X_{n,k'}^2]/\mathbb{E}[X_{n,k'}]^2 \to 1$. Nevertheless, it seems like a reasonable conjecture that in the context of Theorem 1.1, w.h.p.

$$\chi^*_=(G_{n_j,m_j}) = \chi_=(G_{n_j,m_j}) = n/j.$$

Acknowledgements

I would like to thank my DPhil supervisor Oliver Riordan and my thesis examiners Colin McDiarmid and Andrew Thomason for their comments and suggestions on earlier versions of this work. I would also like to thank the anonymous referees for their detailed remarks which greatly improved the presentation of this paper.

https://doi.org/10.1017/S0963548319000397 Published online by Cambridge University Press

References

[1] Alon, N. and Krivelevich, M. (1997) The concentration of the chromatic number of random graphs. *Combinatorica* **17** 303–313.
[2] Bollobás, B. (1988) The chromatic number of random graphs. *Combinatorica* **8** 49–55.
[3] Bollobás, B. (2004) How sharp is the concentration of the chromatic number? *Combin. Probab. Comput.* **13** 115–117.
[4] Erdős, P. and Rényi, A. (1960) On the evolution of random graphs. *Publ. Math. Inst. Hungar. Acad. Sci.* **5** 17–61.
[5] Fountoulakis, N., Kang, R. and McDiarmid, C. (2010) The t-stability number of a random graph. *Electron. J. Combin.* **17** R59.
[6] Hajnal, A. and Szemerédi, E. (1970) Proof of a conjecture of P. Erdős. *Combin. Theory Appl.* **2** 601–623.
[7] Heckel, A. (2018) The chromatic number of dense random graphs. *Random Struct. Alg.* **53** 140–182.
[8] Krivelevich, M. and Patkós, B. (2009) Equitable coloring of random graphs. *Random Struct. Alg.* **35** 83–99.
[9] McDiarmid, C. (1989) On the method of bounded differences. In *Surveys in Combinatorics, 1989* (J. Siemons, ed.), Vol. 141 of London Mathematical Society Lecture Note Series, Cambridge University Press, pp. 148–188.
[10] McDiarmid, C. (1990) On the chromatic number of random graphs. *Random Struct. Alg.* **1** 435–442.
[11] Panagiotou, K. and Steger, A. (2009) A note on the chromatic number of a dense random graph. *Discrete Math.* **309** 3420–3423.
[12] Rombach, P. and Scott, A. Equitable colourings of random graphs. Preprint.
[13] Shamir, E. and Spencer, J. (1987) Sharp concentration of the chromatic number on random graphs $G_{n,p}$. *Combinatorica* **7** 121–129.

Appendix

Proof of Lemma 3.1. Since $1 \leqslant \binom{k}{k_S} \leqslant 2^k$ and $k \sim n/(2\log_b n)$,

$$k_S! k_L! = k! \exp(o(n)) = b^{n/2} \exp(o(n)). \tag{A.1}$$

Furthermore, by (3.1) and (3.4),

$$q^{f_{n,k}} = b^{-(n/2)(n/k-1)} \exp(o(n)) = b^{(n/2)(1-\gamma-x)} \exp(o(n)) = b^{n((1-x)/2)} \left(\frac{2\log_b n}{n}\right)^n \exp(o(n)). \tag{A.2}$$

Finally, note that

$$\exp\left(-\frac{f_{n,k}^2 p}{2qN}\right) = \exp(O(\log^2 n)) = \exp(o(n)).$$

Together with (3.3), (3.6), (A.1) and (A.2), this gives

$$\bar{\mu}_{n,k} = \frac{k^n b^{n((1-x)/2)}}{b^{n/2}} \left(\frac{2\log_b n}{n}\right)^n \exp(o(n)) = b^{-(x/2)n}(1+o(1))^n \exp(o(n)) = b^{-(x/2)n+o(n)}.$$

□

Proof of Lemma 3.3. We first examine how much $\mu_{n,k}$ increases if we increase k in the following lemma.

Lemma A.1. *Given n and k such that $k = n/(\gamma + O(1))$,*

$$\frac{\mu_{n,k+1}}{\mu_{n,k}} \gtrsim b^{n^2/(2k(k+1))-n/(2k)}.$$

Proof. Recall that by (3.5),

$$\mu_{n,k} \sim P_{n,k} q^{f_{n,k}} \exp\left(-\frac{f_{n,k}^2 p}{2qN}\right), \tag{A.3}$$

where

$$P_{n,k} = \frac{n!}{\lceil n/k \rceil!^{k_L} \lfloor n/k \rfloor!^{k_S}}.$$

The product $\lceil n/k\rceil!^{k_{\rm L}}\lfloor n/k\rfloor!^{k_{\rm S}}$ contains exactly n factors, and increasing k by 1 can only decrease those n factors, so

$$P_{n,k+1}\geqslant P_{n,k}. \tag{A.4}$$

Next, we will compare $f_{n,k}$ and $f_{n,k+1}$ using (3.4). For this, let

$$x=\frac{n}{k}-\frac{n}{k+1}=\frac{n}{k(k+1)}.$$

Recall that $\delta_{n,k}:=n/k-\lfloor n/k\rfloor$. If $\lfloor n/k\rfloor=\lfloor n/(k+1)\rfloor$, then $\delta_{n,k+1}=\delta_{n,k}-x$. Otherwise, $\lfloor n/k\rfloor=\lfloor n/(k+1)\rfloor+1$ and $\delta_{n,k}+(1-\delta_{n,k+1})=x$. In both cases,

$$|\delta_{n,k+1}(1-\delta_{n,k+1})-\delta_{n,k}(1-\delta_{n,k})|\leqslant x.$$

Therefore, by (3.4),

$$f_{n,k}-f_{n,k+1}\geqslant\frac{n^2}{2k(k+1)}-\frac{x}{2}(k+1)=\frac{n^2}{2k(k+1)}-\frac{n}{2k}. \tag{A.5}$$

Furthermore,

$$|f_{n,k}-f_{n,k+1}|=\frac{n^2}{2k(k+1)}+O\biggl(\frac{n}{k}\biggr)=O(\log^2 n),$$

so as $f_{n,k}\sim n\log_b n$,

$$\exp\biggl(-\frac{f_{n,k+1}^2p}{2qN}+\frac{f_{n,k}^2p}{2qN}\biggr)=\exp\left(o(1)\right)\sim 1.$$

Plugging this, (A.4) and (A.5) into (A.3), and using $b=1/q$, we have

$$\mu_{n,k+1}\sim P_{n,k+1}q^{f_{n,k+1}}\exp\biggl(-\frac{f_{n,k+1}^2p}{2qN}\biggr)\gtrsim\mu_{n,k}\,b^{n^2/(2k(k+1))-n/(2k)}. \qquad\square$$

It should be noted that $\bar\mu_{n,k}$ behaves in a slightly irregular way: while $\mu_{n,k}$ increases steadily if we increase k, as seen in Lemma A.1, the increases in $\bar\mu_{n,k}$ are not as large as one might expect if n/k is close to an integer. This is because the product $k_{\rm L}!k_{\rm S}!$ is larger when n/k is close to an integer (*i.e.* if either $k_{\rm L}$ or $k_{\rm S}$ is close to k) than when n/k is sufficiently far away from any integers. However, we will see that even in the 'worst-case scenario', $\bar\mu_{n,k}$ still increases by a sufficient amount.

For this, let $k_{\rm L}'=\delta_{n,k+1}(k+1)$ and $k_{\rm S}'=(1-\delta_{n,k+1})(k+1)$, and then

$$k_{\rm L}'+k_{\rm S}'=k+1=k_{\rm L}+k_{\rm S}+1.$$

If $k_{\rm L}>\lfloor n/k\rfloor$, then given a k-equipartition of n vertices, we can form a $(k+1)$-equipartition by removing one vertex from $\lfloor n/k\rfloor$ parts of size $\lceil n/k\rceil$ and forming a new part of size $\lfloor n/k\rfloor$ from the removed vertices. In this case, $k_{\rm L}'=k_{\rm L}-\lfloor n/k\rfloor$ and $k_{\rm S}'=k_{\rm S}+\lfloor n/k\rfloor+1$, and therefore

$$\frac{k_{\rm L}'!k_{\rm S}'!}{k_{\rm L}!k_{\rm S}!}=\frac{\prod_{t=1}^{\lfloor n/k\rfloor+1}(k_{\rm S}+t)}{\prod_{t=0}^{\lfloor n/k\rfloor-1}(k_{\rm L}-t)}\leqslant\frac{(k+1)^{\lfloor n/k\rfloor+1}}{\lfloor n/k\rfloor!}. \tag{A.6}$$

Otherwise, if $k_{\rm L}\leqslant\lfloor n/k\rfloor$, then starting with a k-equipartition, we can form a $(k+1)$-equipartition by removing one vertex from each of the $k_{\rm L}$ parts of size $\lceil n/k\rceil$ and from $\lfloor n/k\rfloor-k_{\rm L}$ parts of size $\lfloor n/k\rfloor$, and forming a new part of size $\lfloor n/k\rfloor$ from the removed vertices. In this case, $k_{\rm S}'=\lfloor n/k\rfloor-k_{\rm L}$ and $k_{\rm L}'=k+1-\lfloor n/k\rfloor+k_{\rm L}$. Note that if $k_{\rm L}\leqslant k_{\rm S}'$, we also have $k_{\rm S}=k-k_L\geqslant k-k_{\rm S}'=k_L'-1$, and therefore

$$k_{\rm L}\leqslant k_{\rm S}'\leqslant k_L'-1\leqslant k_{\rm S}.$$

https://doi.org/10.1017/S0963548319000397 Published online by Cambridge University Press

As for any integers $1 \leqslant x_1 \leqslant x_2 \leqslant x_3 \leqslant x_4$ with $x_1 + x_4 = x_2 + x_3$, we have $x_1!x_4! \geqslant x_2!x_3!$, this implies

$$(k+1)k_L!k_S! \geqslant (k+1)(k'_L - 1)!k'_S! \geqslant k'_L!k'_S!. \tag{A.7}$$

Otherwise, if $k'_S = \lfloor n/k \rfloor - k_L < k_L \leqslant \lfloor n/k \rfloor$, then

$$\frac{k'_L!k'_S!}{k_L!k_S!} = \frac{k'_L!/k_S!}{k_L!/k'_S!} \leqslant \frac{(k+1)^{k'_L - k_S}}{(k_L - k'_S)!} = \frac{(k+1)^{1+k_L-k'_S}}{(k_L - k'_S)!} \leqslant \frac{(k+1)^{\lfloor n/k \rfloor + 1}}{\lfloor n/k \rfloor !}.$$

Comparing this to (A.6) and (A.7), we can see that in every case,

$$\begin{aligned} \frac{k'_L!k'_S!}{k_L!k_S!} &\leqslant \frac{(k+1)^{\lfloor n/k \rfloor + 1}}{\lfloor n/k \rfloor !} \\ &\leqslant \frac{e^{\lfloor n/k \rfloor}(k+1)^{\lfloor n/k \rfloor + 1}}{\lfloor n/k \rfloor^{\lfloor n/k \rfloor}} \\ &\leqslant \left(\frac{e(k+1)}{n/k - 1} \right)^{\lfloor n/k \rfloor} (k+1) \\ &\leqslant \left(\frac{e(k+1)k}{n-k} \right)^{n/k} (k+1). \end{aligned}$$

Together with Lemma A.1, and since $n/k = \gamma + O(1) = \Theta(\log n)$, this gives

$$\begin{aligned} \frac{\bar{\mu}_{n,k+1}}{\bar{\mu}_{n,k}} &\gtrsim b^{n^2/(2k(k+1)) - n/(2k)} \left(\frac{n-k}{(k+1)k} \right)^{n/k} n^{O(1)} \\ &= b^{(\gamma n)/(2(k+1))} \left(\frac{n}{k^2} \right)^{\gamma} n^{O(1)} \\ &= \left(\frac{n}{2 \log_b n} \right)^{n/(k+1)} \left(\frac{n}{k^2} \right)^{\gamma} n^{O(1)} \\ &= \left(\frac{n^2}{k^2 \log_b n} \right)^{\gamma} n^{O(1)} \\ &= (\Theta(\log n))^{\gamma} n^{O(1)} \\ &= \exp\left(\Theta(\log n \log\log n)\right). \end{aligned}$$

□

Proof of Lemma 3.4. Given a k-equipartition of n vertices, adding a vertex to a part of size $\lfloor n/k \rfloor$ yields a k-equipartition of $n+1$ vertices, so

$$f_{n+1,k} = f_{n,k} + \left\lfloor \frac{n}{k} \right\rfloor. \tag{A.8}$$

Therefore, since $\lfloor n/k \rfloor = \gamma(n) + O(1)$ and by (3.1),

$$q^{f_{n+1,k} - f_{n,k}} = \Theta(1) \left(\frac{2 \log_b n}{n} \right)^2. \tag{A.9}$$

Furthermore, as $f_{n,k} = O(n \log n)$ and from (A.8),

$$-\frac{f^2_{n+1,k}p}{2qN} + \frac{f^2_{n,k}p}{2qN} = O\left(\frac{\log^2 n}{n} \right) = o(1). \tag{A.10}$$

Finally, note that by (3.3),

$$\frac{P_{n+1,k}}{P_{n,k}} \sim \frac{n}{2 \log_b n},$$

https://doi.org/10.1017/S0963548319000397 Published online by Cambridge University Press

since the factorial in the numerator is multiplied by $n+1 \sim n$ and the product in the denominator is multiplied by exactly one factor $\lfloor n/k \rfloor + 1 \sim \gamma \sim 2\log_b n$ if n is increased to $n+1$. Together with (3.5), (A.9) and (A.10), this completes the proof. □

Calculations for equation (5.7). Note that $\log(1+x) = x - x^2/2 + O(x^3)$ for $x \to 0$. Therefore, for any y and z so that $y/z \to 0$, we have

$$\log\left(\left(1-\frac{y}{z}\right)^{z-y}\right) = -y + \frac{y^2}{2z} + O\left(\frac{y^3}{z^2}\right).$$

Using this, and recalling that $d \leqslant f = O(n \log n)$ and $0 \leqslant \varepsilon < 1$,

$$\begin{aligned}\log S_{\mathbf{r}} &= -d\log q - 2f + d + \frac{(2f-d)^2}{2N} - 2f + 2\varepsilon + \frac{(f-\varepsilon)^2}{qN} + 2f - \frac{f^2}{N} - \varepsilon - \frac{\varepsilon^2}{2qN} \\ &\quad + 2f - d - \varepsilon - \frac{(2f-d-\varepsilon)^2}{2qN} + O\left(\frac{f^3}{N^2}\right) \\ &= -d\log q + \frac{q(2f-d)^2 + 2(f-\varepsilon)^2 - 2qf^2 - (2f-d-\varepsilon)^2}{2qN} + O\left(\frac{f^3}{N^2}\right) \\ &= -d\log q + \frac{-p(d^2+2f^2-4df)}{2qN} + o(1).\end{aligned}$$

Proof of Lemma 5.1. First note that $p < 1 - 1/e^2$ and therefore $\log b < 2$, so indeed $\tilde{c} \in (0,1)$. We assume throughout that j is large enough for all bounds to hold. We will first check that the claim holds for the first and last terms, T_3 and T_j. Next, we will see that $T_{i+1}/T_i \leqslant 1$ if i is small ($i < 0.8\log_b n$) and $T_{i+1}/T_i \geqslant 1$ is i is large ($i > 1.2\log_b n$), which means we can bound T_i by $\max(T_3, T_j)$ for all such i. Finally we check that the claim holds for the intermediate terms T_i where $0.8\log_b n < i < 1.2\log_b n$.

So consider the case $i = 3$. Note that

$$T_3 \leqslant \frac{e^3 b^3 k^2 j^6}{n^3 3!} = n^{-1+o(1)} \leqslant n^{-\tilde{c}}.$$

Next, consider $i = j$. Since $j = \gamma(n_j) + o(1) = 2\log_b n - 2\log_b \log_b n - 2\log_b 2 + o(1)$,

$$b^{\binom{j}{2}} = \left(\frac{n}{2\log_b n}\right)^{j-1} n^{o(1)},$$

so with Stirling's formula and as $k = n/j = n^{1+o(1)}$ and $j \sim 2\log_b n$, we have

$$T_j = \frac{e^{\rho j} b^{\binom{j}{2}} k^2 j!}{n^j} = n^{o(1)} \frac{e^{\rho j} n j^j}{(2\log_b n)^{j-1} e^j} = n^{1+o(1)} e^{-j+\rho j}.$$

As $e^j = n^{2/(\log b)+o(1)}$, and $\rho \leqslant c = \frac{1}{2}(1-(\log b)/2)$ since $\mathbf{r} \in \mathcal{R}_1$,

$$T_j \leqslant n^{1-(1-\rho)(2/\log b)+o(1)} = n^{\frac{1}{2}-(1/\log b)+o(1)} \leqslant n^{-\tilde{c}}.$$

Note that

$$\frac{T_{i+1}}{T_i} = \frac{e^{\rho} b^i (j-i)^2}{n(i+1)}.$$

In particular, for all $i \leqslant 0.8\log_b n$,

$$\frac{T_{i+1}}{T_i} \leqslant n^{-0.2+o(1)} \leqslant 1,$$

so for all $3 \leqslant i \leqslant 0.8 \log_b n$, we have $T_i \leqslant T_3 \leqslant n^{-\tilde{c}}$. If $i \geqslant 1.2 \log_b n$, then

$$\frac{T_{i+1}}{T_i} \geqslant n^{0.2+o(1)} \geqslant 1,$$

so for all $1.2 \log_b n \leqslant i \leqslant j$, we have $T_i \leqslant T_j \leqslant n^{-\tilde{c}}$. For the remaining case $0.8 \log_b < i < 1.2 \log_b n$, note that as $j \leqslant 2 \log_b n$,

$$T_i \leqslant \frac{e^i b^{i^2/2} n^2 j^{2i}}{n^i} \leqslant n^{O(1)} \frac{n^{0.6i} j^{2i}}{n^i} \leqslant n^{O(1)-0.3i} = n^{-\Theta(\log n)} \leqslant n^{-\tilde{c}}. \qquad \square$$

On the sets $\mathcal{P}^{\mathrm{I}}$, $\mathcal{P}^{\mathrm{II}}$ and $\mathcal{P}^{\mathrm{III}}$. We give the definitions of the sets $\mathcal{P}^{\mathrm{I}}$, $\mathcal{P}^{\mathrm{II}}$ and $\mathcal{P}^{\mathrm{III}}$, as well as the statement of Lemma 13 of [7].

Let $a = \lfloor \gamma \rfloor + 1$. In the context of this paper, we have $\gamma = j + o(1)$ and hence $a \in \{j, j+1\}$. Given some $\pi_2 \in \mathcal{P}_2^{c'}$, let

$V_1 = V_1(\pi_2) =$ set of vertices in the overlap of π_1 and π_2 which are in parts of size a in π_1,
$V_2 = V_2(\pi_2) =$ set of vertices in the overlap of π_1 and π_2 which are in parts of size at most $a-1$ in π_1,
$D_1 = D_1(\pi_2) =$ set of forbidden edges between vertices in V_1,
$D_2 = D_2(\pi_2) =$ set of forbidden edges between vertices in V_2.

For $i \in \{1, 2\}$, let $v_i = |V_i|$ and $d_i = |D_i|$. Finally, let

$$\beta_1 = \frac{2d_1}{v_1(a-1)} \leqslant 1, \quad \beta_2 = \frac{2d_2}{v_2(a-2)} \leqslant 1.$$

The parameters β_1 and β_2 quantify how close the overlap of π_1 and π_2 is to consisting only of large overlap blocks of size a or $a-1$ – for details; see §5.3.1 of [7]. We are now ready to state Lemma 13.

Lemma A.2 (Lemma 13 of [7]). *If $\pi_2 \in \mathcal{P}_2^{c'}$ and n is large enough, then at least one of the following three conditions applies:*

(I) $v_1 \geqslant \dfrac{n}{(\log\log n)^2}$ *and* $\beta_1 \leqslant 1 - \dfrac{(\log\log n)^4}{\log n}$,
(II) $v_2 \geqslant \dfrac{n}{(\log\log n)^2}$ *and* $\beta_2 \leqslant 1 - \dfrac{(\log\log n)^4}{\log n}$,
(III) *neither (I) nor (II) holds, and $c < \rho \leqslant 1 - c'$.*

Now let

$$\mathcal{P}^{\mathrm{I}} = \left\{ \text{ordered } k\text{-equipartitions } \pi_2 \text{ such that } v_1 \geqslant \frac{n}{(\log\log n)^2} \text{ and } \beta_1 \leqslant 1 - \frac{(\log\log n)^4}{\log n} \right\},$$
$$\mathcal{P}^{\mathrm{II}} = \left\{ \text{ordered } k\text{-equipartitions } \pi_2 \text{ such that } v_2 \geqslant \frac{n}{(\log\log n)^2} \text{ and } \beta_2 \leqslant 1 - \frac{(\log\log n)^4}{\log n} \right\},$$
$$\mathcal{P}^{\mathrm{III}} = \{\text{ordered } k\text{-equipartitions } \pi_2 \text{ such that } c < \rho \leqslant 1 - c'\} \setminus \mathcal{P}^{\mathrm{I}} \setminus \mathcal{P}^{\mathrm{II}}.$$

By the lemma above,

$$\mathcal{P}_2^{c'} \subset \mathcal{P}^{\mathrm{I}} \cup \mathcal{P}^{\mathrm{II}} \cup \mathcal{P}^{\mathrm{III}}.$$

Cite this article: Heckel A (2020). Sharp concentration of the equitable chromatic number of dense random graphs. *Combinatorics, Probability and Computing* **29**, 213–233. https://doi.org/10.1017/S0963548319000397