10

Colouring random graphs

ROSS J. KANG and COLIN MCDIARMID

1. Introduction
2. Dense random graphs
3. Sparse random graphs
4. Random regular graphs
5. Random geometric graphs
6. Random planar graphs and related classes
7. Other colourings
References

Typically how many colours are required to colour a graph? In other words, given a graph chosen randomly, what can we expect its chromatic number to be? We survey the classic interpretation of this question, with the binomial or Erdős–Rényi random graph and the usual chromatic number. We also treat a few variations, not only of the random graph model, but also of the chromatic parameter.

1. Introduction

How many colours are typically necessary to colour a graph?

We survey a number of perspectives on this natural question, which is central to random graph theory and to probabilistic and extremal combinatorics. It has stimulated a vibrant area of research, with a rich history extending back through more than half a century.

Erdős and Rényi [36] asked a form of this question in a celebrated early paper on random graphs in 1960. Let $G_{n,m}$ be a graph chosen uniformly at random from the set of graphs with vertex-set $[n] = \{1, 2, \ldots, n\}$ and m edges. In this probabilistic model, we cannot rule out the possibility that $G_{n,m}$ is, for example, the disjoint union of one large clique and some isolated vertices, or perhaps one Turán graph (a balanced complete multipartite graph) and some isolated vertices. The resulting

https://doi.org/10.1017/CBO9781139519793.013 Published online by Cambridge University Press

range is large: in the former situation the chromatic number could be about $\sqrt{2m}$, while in the latter it could be 2 if $m \leq n^2/4$. These outcomes are unlikely, however, and we are interested in the most probable ones. To state the question properly, we say that an event A_n (which here always describes a property of a random graph on the vertex-set $[n]$) holds *asymptotically almost surely* (*a.a.s.*) if the probability that A_n holds satisfies $\mathbb{P}(A_n) \to 1$ as $n \to \infty$. Erdős and Rényi asked the following question.

Suppose that $m \sim cn$ for some $c \geq \frac{1}{2}$. Is there a positive integer function $f = f_c(n)$ for which $\chi(G_{n,m}) = f$ a.a.s., and if so how large is it?

They already noted that $f = 3$ if $c = \frac{1}{2}$, which tells us that $\mathbb{P}(\chi(G_{n,n/2}) = 3) \to 1$ as $n \to \infty$; however, as we will see, handling larger fixed values of c is a very challenging task.

If, however, we instead consider $G_{n,m}$ with the symmetric choice $m = \frac{1}{2}\binom{n}{2}$, then this is nearly the same as the basic model $G_{n,1/2}$, chosen uniformly at random from graphs with vertex-set $[n]$. Equivalently, form $G_{n,1/2}$ by including as an edge each of the $\binom{n}{2}$ possible pairs from $\binom{[n]}{2}$ independently at random with probability $\frac{1}{2}$. Even the following seemingly innocent question was open for decades in spite of serious and sustained efforts, thus taking the pattern of other hard problems in combinatorics and graph theory.

What is the a.a.s. first-order approximate behaviour of the chromatic number of a uniformly chosen graph with vertex-set $[n]$?

In other words, this question asks for an elementary function $f(n)$ (if there is one) that satisfies $\chi(G_{n,1/2}) \sim f(n)$ a.a.s. The typical value of the chromatic number is an attractive concept in its own right, but there is also a strong interplay with other research areas. This provides some contrasting viewpoints we now briefly discuss.

Researchers in 'deterministic' (chromatic) graph theory frequently look to random graphs for examples and counter-examples. In 1959 Erdős [34] elegantly proved the existence of graphs with both girth and chromatic number arbitrarily large – a clean constructive proof came somewhat later. Also, Erdős and Fajtlowicz [35] showed that 'almost all' graphs are counter-examples to Hajós's conjecture, which, as an over-strengthened form of Hadwiger's conjecture, claimed that every k-chromatic graph contains a subdivision of K_k. Both of these well-known applications of the probabilistic method used lower bounds for the chromatic number of a random graph.

In theoretical computer science, determining the chromatic number of a given graph is an archetypal NP-hard optimization problem. One might wonder whether the situation is simpler in a probabilistic setting. The chromatic number of random graphs (as well as the performance of algorithms that compute or approximate it) may be viewed as a crude indication of 'average-case' computational behaviour.

Surprisingly, the study of the chromatic number of random graphs is equivalent to an important model of spin glasses studied in statistical physics. It has been referred

https://doi.org/10.1017/CBO9781139519793.013 Published online by Cambridge University Press

to variously as the 'diluted mean-field antiferromagnetic Potts model' and the 'zero-temperature Curie–Weiss–Potts antiferromagnet'. Spin glass models are insightful representations for the physical phenomenon of a *phase transition*, such as the change between water and ice.

This topic thus provides common ground for combinatorialists, probabilists, theoretical computer scientists and statistical physicists. Later we will see examples where the mixing of fields in this area has been stimulating and fruitful.

A main focus of ours is to address the two questions in italics above, which correspond to standard density regimes in the theory of classical Erdős–Rényi random graphs. The latter question, treated in Section 2, is representative of the 'dense' regime of random graphs, since the expected number of edges in $G_{n,1/2}$ is $\Theta(n^2)$. The former question (discussed in Section 3) belongs to the 'sparse' regime, since there are only a linear number of edges. Random graphs have very interesting properties and go through a number of well-defined phases according to the graph's density. For more on the theory and evolution of Erdős–Rényi random graphs, see the standard texts by Bollobás [15] and Janson, Łuczak and Ruciński [49].

The overarching question posed at the very beginning can be ascribed a broader meaning. It is well separated into its constituent parts, one of which concerns 'random graphs', the other of which concerns 'colouring'. The second half of this chapter is devoted to selectively developing these parts independently.

Random graph theory continues to develop rapidly and steadily, partly through the influence of networks. Many probability spaces for graphs can be interpreted as models for real networks such as the internet, protein–protein interaction, mathematical collaboration and telecommunication. We treat a selection of random graph models (apart from Erdős–Rényi) for which the chromatic number is important. Random regular graphs are discussed in Section 4; random geometric graphs, which are considered a standard model for frequency allocation in ad hoc communication networks, are treated in Section 5; and random graph models related to the colouring of planar graphs are considered in Section 6.

It is more than evident from the chapters in this book that chromatic graph theory is a broad and mature field. Chromatic *random* graph theory is also well developed. In the last section, Section 7, we present results on a few specific chromatic parameters, concentrating for brevity on the dense Erdős–Rényi random graph. First we discuss strengthenings of the chromatic number, then we cover edge and total colourings, and last we consider generalizations of the chromatic number.

We must mention at least two other major research areas that combine probability with graph colouring. Although they are closely related to this topic, they are outside the scope of this chapter. The generation of a uniformly random colouring of a graph, often with the Markov chain Monte Carlo method, has been extensively studied (see [44]). The application of the probabilistic method in graph colouring, using the Lovász local lemma, for instance, has been very successful and is covered in depth in the monograph of Molloy and Reed [75].

https://doi.org/10.1017/CBO9781139519793.013 Published online by Cambridge University Press

In this chapter, the symbols $\mathbb{P}$ and $\mathbb{E}$ denote probability and expectation, respectively. The level of probabilistic expertise needed for this material is not burdensome, but for more background see Alon and Spencer [9] or one of the random graph texts [15], [49]. Mostly we just use the first moment (Markov's inequality) and second moment (Chebyshev's or the Paley–Zygmund inequality) methods. Given a non-negative random variable X, Markov's inequality asserts that $\mathbb{P}(X \geq a) \leq \mathbb{E}(X)/a$ for any $a > 0$, while the Paley–Zygmund inequality implies that $\mathbb{P}(X > 0) \geq (\mathbb{E}(X))^2/\mathbb{E}(X^2)$. We refer to other inequalities for the establishment of concentrated probability measure – Chernoff, Azuma–Hoeffding, Janson's and Talagrand's inequalities – all of which are covered in the above-mentioned texts. We also use the approximation $(n/k)^k \leq \binom{n}{k} \leq (en/k)^k$, which is a consequence of the fact that $e^x \geq x^k/k!$ for $x \geq 0$, and the basic inequality $(1-a)^n \leq \exp(-an)$ for $a \in (0,1)$. When we write an unadorned logarithm symbol log, it is assumed to have base 2; ln has the natural base e. We omit symbols for floors and ceilings, unless they are important. We have adopted standard asymptotic (Landau) notation: O, Ω, Θ, $\sim$, o and ω. Unless otherwise specified, the asymptotics are as $n \to \infty$. Recall that, for functions $f(n)$ and $g(n) > 0$, we write $f = O(g)$ if $f(n) \leq cg(n)$ for some constant c, $f = \Theta(g)$ if both $f = O(g)$ and $g = O(f)$, $f \sim g$ if $f(n)/g(n) \to 1$, $f = o(g)$ if $f(n)/g(n) \to 0$, and $f = \omega(g)$ if $g = o(f)$.

2. Dense random graphs

By the dense regime of Erdős–Rényi random graphs, we mean $G_{n,p}$ for fixed $p \in (0,1)$. For clarity, we have chosen to restrict our attention to $p = \frac{1}{2}$ – that is, 'almost all graphs', though all results in this section extend naturally across the entire dense regime, to other fixed $p \in (0,1)$.

Upon encountering the problem for the first time, one might observe that the graph $G_{n,1/2}$ has 'global' rather than 'local' structure. In particular, any given vertex has a reasonable chance to be adjacent (or not) to any other vertex in the graph. This suggests that its chromatic number might not necessarily be determined by a local quantity, such as a large clique or a small maximum degree. As it happens, both the clique number of $G_{n,1/2}$ (which is about $2\log n$ a.a.s., as we shall see) and the maximum degree of $G_{n,1/2}$ (which is concentrated around $\frac{1}{2}n$ a.a.s.) are quite far away from $\chi(G_{n,1/2})$.

Instead, it turns out that the value of $\chi(G_{n,1/2})$ strongly depends upon large stable sets in $G_{n,1/2}$. (Throughout this chapter, we exclusively use *stable set*, rather than *independent set*, to refer to a vertex subset that induces a subgraph having no edges.) Recall that in any proper colouring of $G_{n,1/2}$, each colour class has order at most $\alpha(G_{n,1/2})$ (where $\alpha(G)$ denotes the number of vertices in a largest stable set of G), and thus $\chi(G_{n,1/2}) \geq n/\alpha(G_{n,1/2})$ always. Fortunately, as we see in the next subsection, $\alpha(G_{n,1/2})$ can be precisely determined without much difficulty. On the other hand, a thorough understanding of the 'global' behaviour of large stable sets in $G_{n,1/2}$ is

https://doi.org/10.1017/CBO9781139519793.013 Published online by Cambridge University Press

necessary for a good upper bound on $\chi(G_{n,1/2})$. This vague narrative should become clearer later.

The stability number

Let $\mathcal{S}_k$ be the collection of stable sets of order k in $G_{n,1/2}$. To determine $\alpha(G_{n,1/2})$, we first study the expectation of $|\mathcal{S}_k|$. There are $\binom{n}{k}$ sets of order k, and the probability that such a set is stable (that is, each of its $\binom{k}{2}$ edges is absent) is $2^{-\binom{k}{2}}$. We thus have

$$\mathbb{E}(|\mathcal{S}_k|) = \binom{n}{k} 2^{-\binom{k}{2}}.$$

For relevant values of k, this expression for the first moment of $|S_k|$ is decreasing in k. We are interested in the values of k over which it changes from a large quantity (much greater than 1) to a small quantity (somewhat less than 1). In the latter case, there is positive probability that there are no stable sets of order k (that is, $\mathbb{P}(|\mathcal{S}_k| = 0) > 0$), whereas in the former, if the expectation is large enough, we would hope that the distribution of such sets is sufficiently well behaved that we could then pin down such a set. As we now show, the change in $\mathbb{E}(|\mathcal{S}_k|)$ is large and the range of k over which it happens is small.

With foresight, we consider, for some fixed $\varepsilon \in (0, 1)$, the choice

$$k = k_{+\varepsilon} = \lfloor k_0 + \varepsilon \rfloor, \quad \text{where } k_0 = 2\log n - 2\log\log n + 2\log e - 1.$$

Now, by an approximation of the binomial coefficient,

$$\mathbb{E}(|\mathcal{S}_{k_{+\varepsilon}+1}|) \le \left(\frac{en}{k_{+\varepsilon}+1}\right)^{k_{+\varepsilon}+1} 2^{-\binom{k_{+\varepsilon}+1}{2}},$$

and so, since $k_{+\varepsilon} + 1 \ge k_0 + \varepsilon$,

$$\frac{2\log \mathbb{E}(|\mathcal{S}_{k_{+\varepsilon}+1}|)}{k_{+\varepsilon}+1} \le 2\log e + 2\log n - 2\log k_0 - k_0 - \varepsilon + 1 = -\varepsilon + o(1),$$

since $\log k_0 = \log\log n + 1 + o(1)$. As we are interested in an asymptotic statement, we may allow an arbitrarily large choice of n. Then this last expression becomes at most $-\frac{1}{2}\varepsilon$, and $k_{+\varepsilon} \ge \log n$. It follows that, for sufficiently large n, $\mathbb{E}(|\mathcal{S}_{k_{+\varepsilon}+1}|) \le \exp\left(-\frac{1}{4}\varepsilon \log n\right)$, a vanishing quantity. We immediately deduce from an application of Markov's inequality that

$$\mathbb{P}(\alpha(G_{n,1/2}) > k_{+\varepsilon}) = \mathbb{P}(|\mathcal{S}_{k_{+\varepsilon}+1}| \ge 1) \le \mathbb{E}(|\mathcal{S}_{k_{+\varepsilon}+1}|) \to 0 \text{ as } n \to \infty.$$

In other words, $\alpha(G_{n,1/2}) \le k_{+\varepsilon}$ a.a.s.

Next let us consider, for some fixed $\varepsilon > 0$, the choice $k = k_{-\varepsilon} = \lfloor k_0 - \varepsilon \rfloor$, where k_0 is as above. Using another approximation of the binomial coefficient (in the other direction), we obtain by almost identical manipulations that, for n large enough,

https://doi.org/10.1017/CBO9781139519793.013 Published online by Cambridge University Press

$\mathbb{E}(|\mathcal{S}_{k_{-\varepsilon}}|) \geq \exp\left(\frac{1}{4}\varepsilon \log n\right)$, an unbounded quantity. So we see that as a result of a small decrease in k (from $k_{+\varepsilon}+1$ to $k_{-\varepsilon}$) there is a large increase in $\mathbb{E}(|\mathcal{S}_k|)$. Around 1970 Matula [62], [63] showed that moreover $|\mathcal{S}_{k_{-\varepsilon}}|$ is concentrated around its mean. In particular, it is possible to show that the second moment $\mathbb{E}(|\mathcal{S}_{k_{-\varepsilon}}|^2)$ of $|\mathcal{S}_{k_{-\varepsilon}}|$ is very small compared to $(\mathbb{E}(|\mathcal{S}_{k_{-\varepsilon}}|))^2$, by using calculations similar to (but more involved than) those just given above for the first moment. Consequently, an application of Chebyshev's or the Paley–Zygmund inequality yields $\alpha(G_{n,1/2}) \geq k_{-\varepsilon}$ a.a.s. Thus the value of the stability number of $G_{n,1/2}$ is very precisely determined a.a.s.

Theorem 2.1 *For fixed $\varepsilon \in (0,1)$, $\alpha(G_{n,1/2}) \in [k_{-\varepsilon}, k_{+\varepsilon}]$ a.a.s.*

A slightly sharper form of this was proved independently by Bollobás and Erdős [17] in 1976.

A greedy algorithm

In the last subsection, we saw with a first moment argument that $\alpha(G_{n,1/2}) \leq k_{+\varepsilon}$ and so the lower bound $\chi(G_{n,1/2}) \geq n/k_{+\varepsilon} \sim n/(2\log n)$ holds a.a.s. We now show how another first moment argument, due to Grimmett and McDiarmid [48] in 1975, also gives a reasonable upper bound on $\chi(G_{n,1/2})$, one that is only twice as large as the lower bound. This argument uses a simple colouring algorithm. It relies on the following observation for maximal stable sets – that is, stable sets that cannot be enlarged by adding a vertex.

Theorem 2.2 *With probability at least $1-\exp(-\Omega(\log^3 n))$, every maximal stable set of $G_{n,1/2}$ has order greater than $\log n - 3\log\log n$.*

Proof Given a stable set in $G_{n,1/2}$ of order k, the probability that it is maximal is $(1-2^{-k})^{n-k}$, this being the probability that each vertex outside the stable set is adjacent to at least one of the vertices within the set. As there are at most $\binom{n}{k}$ stable sets of order k, and so fewer than n^k, the expected number of maximal stable sets of order at most $\log n - 3\log\log n$ is at most

$$\begin{aligned}\sum_{k=1}^{\log n - 3\log\log n} \binom{n}{k}(1-2^{-k})^{n-k} &\leq \sum_{k=1}^{\log n - 3\log\log n} n^k \exp(-(n-k)2^{-k}) \\ &\leq (\log n) n^{\log n} \exp(-(1-k/n)\log^3 n) \\ &= \exp(-\Omega(\log^3 n)),\end{aligned}$$

where we have used the fact that $\exp(-(n-k)2^{-k})$ is maximized over $k \in [1, \log n - 3\log\log n]$ when $k = \log n - 3\log\log n$. So, by Markov's inequality, the probability of there being a maximal stable set of order at most $\log n - 3\log\log n$ is also at most $\exp(-\Omega(\log^3 n))$. ■

https://doi.org/10.1017/CBO9781139519793.013 Published online by Cambridge University Press

We see now that the following greedy procedure quickly and reliably produces a stable set S whose order is almost half of $\alpha(G_{n,1/2})$: initialize S to be the empty set; at each step, choose a previously unchosen vertex of $[n] \setminus S$ and add it to S if it is adjacent to none of the vertices in S. The greedy colouring algorithm borrows this as a subprocedure to generate new colour classes iteratively. The validity of the analysis relies on the observation that the generation of S relies only on the examination (or 'exposure') of candidate edges that have an endpoint in S, and thus the graph obtained by the removal of S from some $G_{n',1/2}$ may be viewed as an independent copy of the random graph $G_{n'-|S|,1/2}$, upon which we may iterate. We continue until fewer than $n/\log^2 n$ vertices remain and we give each of these vertices its own colour. Thus the total number of colours used is less than

$$\frac{n}{\log n - 5\log\log n} + \frac{n}{\log^2 n} \sim \frac{n}{\log n},$$

assuming that each iteration successfully generates a colour class of order $\log n - 5\log n$. The probability that any iteration of the algorithm fails to extract such a colour class from $G_{n',1/2}$, with $n' \geq n/\log^2 n$, is at most $\exp(-\Omega(\log^3 n))$, by Theorem 2.2. Trivially, there are at most n stages, and thus by the union bound (Boole's inequality) the algorithm succeeds with probability at least $1 - n\exp(-\Omega(\log^3 n)) \to 1$ as $n \to \infty$. It is easily checked that this algorithm runs in polynomial time, and so we have the following result.

Theorem 2.3 *There is a polynomial-time algorithm that a.a.s. produces a proper colouring of $G_{n,1/2}$ with at most $(1 + o(1))n/\log n$ colours.*

In fact, McDiarmid [66] showed that many of the most common polynomial-time colouring heuristics a.a.s. produce colourings of $G_{n,1/2}$ with at most $(1+o(1))n/\log n$ colours.

Concentration and the chromatic number

We have seen that $(1 + o(1))n/(2\log n) \leq \chi(G_{n,1/2}) \leq (1 + o(1))n/\log n$ a.a.s. Grimmett and McDiarmid also showed that the expected number of proper $((1 + \varepsilon)n/(2\log n))$-colourings of $G_{n,1/2}$ tends to infinity, and then conjectured that the lower bound in the above range should be the correct first-order behaviour of $\chi(G_{n,1/2})$ a.a.s. This remained one of the most important open problems in random graph theory for well over a decade.

A turning point was the introduction of simple yet powerful tools from probability theory for proving concentration of measure. We owe this important insight to Shamir and Spencer [82] in 1987. With a direct application of the Azuma–Hoeffding martingale inequality (to the 'vertex exposure' martingale), they showed that the value of $\chi(G_{n,1/2})$ is a.a.s. concentrated within an interval roughly of length $O(\sqrt{n})$. Frustratingly, however, this method gave no clue as to the location of this interval!

https://doi.org/10.1017/CBO9781139519793.013 Published online by Cambridge University Press

In spite of this, the discovery of this connection proved decisive and the problem was resolved soon after. Several methods were developed, independently or in quick succession. In the next section, we discuss the 'expose-and-merge' algorithm of Matula [64]. The elegant method of Bollobás [14] applied martingale concentration to prove that a non-algorithmic form of the greedy process described in the last subsection can properly colour $G_{n,1/2}$ with large colour classes. This approach depends on the following strong upper bound on the probability of $G_{n,1/2}$ to have all stable sets slightly smaller than the values given by Theorem 2.1.

Theorem 2.4 $\mathbb{P}(\alpha(G_{n,1/2}) < 2\log n - 3\log\log n) < \exp(-\Omega(n^2/\log^5 n))$.

Before discussing this bound further, let us first see how it is used to determine the a.a.s. first-order behaviour of $\chi(G_{n,1/2})$.

Theorem 2.5 $\chi(G_{n,1/2}) \sim n/(2\log n)$ *a.a.s.*

Proof As the lower bound follows from the upper bound on $\alpha(G_{n,1/2})$ of Theorem 2.1, it suffices to prove the upper bound. To this end, let $\mathcal{A}_n$ be the set of graphs G on $[n]$ with $\alpha(G[S]) \geq 2\log n - 7\log\log n$ for all $S \subseteq [n]$ with $|S| \geq n/\log^2 n$. Then Theorem 2.4 yields, as $n \to \infty$, that

$$\begin{aligned}\mathbb{P}(G_{n,1/2} \notin \mathcal{A}_n) &\leq 2^n\, \mathbb{P}(\alpha(G_{n/\log^2 n,1/2}) < 2\log n - 7\log\log n)\\ &\leq 2^n \exp(-\Omega(n^2/\log^9 n)) \to 0.\end{aligned}$$

Thus $G_{n,1/2} \in \mathcal{A}_n$ a.a.s.

For any graph in $\mathcal{A}_n$, an iterative process similar to the greedy algorithm yields a good colouring. We start with $T = [n]$. As long as $|T| \geq n/\log^2 n$, the condition for membership in $\mathcal{A}_n$ guarantees that we can extract a new colour class S from T with at least $2\log n - 7\log\log n$ vertices and then set $T = T \setminus S$. After stopping, we give each vertex left in T its own colour, so fewer than $n/(2\log n - 7\log\log n) + n/\log^2 n \sim n/(2\log n)$ colours are used. ■

By our earlier use of the term 'non-algorithmic' (which was not strictly correct), we meant that, while Theorem 2.4 ensures that there is a large stable set, it does not provide any efficient method for producing one. Certainly, a brute-force test of every subset with the appropriate number of vertices is far too slow. In 1976 Karp [52] asked whether there is a polynomial-time algorithm that produces from $G_{n,1/2}$ a stable set of size $(1+\varepsilon)\log n$ a.a.s. This is open, even if (for example, by Theorem 2.2) there are fast algorithms to produce stable sets of size $(1-\varepsilon)\log n$.

As mentioned above, tools for proving strong concentration of measure were crucial. By now, there are a few ways to prove Theorem 2.4. Bollobás's proof made a clever choice of two random variables that track large stable sets. One of these never exceeds the other, but they serve different purposes: the former has expectation easily bounded from below, while the latter is a martingale amenable to the Azuma–Hoeffding inequality. More direct proofs use more sophisticated concentration

https://doi.org/10.1017/CBO9781139519793.013 Published online by Cambridge University Press

bounds, such as Talagrand's inequality. We have chosen to outline the use of Janson's inequality – a powerful probabilistic tool partly inspired by Bollobás's method – to prove Theorem 2.4. The reader may choose to skip the following outline, and take away the rough message that, for sets that are large but not too large, Janson's inequality provides the required probability bound if we can control the pairwise dependence (measured by a parameter Δ below) among the events of being stable.

Outline of proof of Theorem 2.4 As before, let $\mathcal{S}_k$ be the collection of stable sets of order k in $G_{n,1/2}$. By Janson's inequality (Theorem 2.18(ii) in [49]),

$$\mathbb{P}(\alpha(G_{n,1/2}) < k) = \mathbb{P}(|\mathcal{S}_k| = 0) \le \exp\left(-\frac{\mu^2}{\mu + \Delta}\right),$$

where $\mu = \mathbb{E}(|\mathcal{S}_k|)$. On summing over ordered pairs $A, B \subseteq [n]$ with $2 \le |A \cap B| \le k-1$,

$$\Delta = \sum \mathbb{P}(A, B \in \mathcal{S}_k).$$

Recall that $\mu = \binom{n}{k} 2^{-\binom{k}{2}}$, and observe that $2^{-2\binom{k}{2}+\binom{\ell}{2}}$ is the probability that two given subsets of $[n]$ of order k that overlap in exactly ℓ vertices are both in $\mathcal{S}_k$. We may then write

$$\Delta = \sum_{\ell=2}^{k-1} \binom{n}{k}\binom{k}{\ell}\binom{n-k}{k-\ell} 2^{-2\binom{k}{2}+\binom{\ell}{2}} = \mu \sum_{\ell=2}^{k-1} \binom{k}{\ell}\binom{n-k}{k-\ell} 2^{-\binom{k}{2}+\binom{\ell}{2}}.$$

It is a routine exercise to check, under the assumption that $k \le 2\log n - 3\log\log n$, that the first term is the largest in the sum, and therefore

$$\Delta \le \mu(k-2)\binom{k}{2}\binom{n-k}{k-2} 2^{-\binom{k}{2}+1} \le \mu^2 O\left(\frac{k^5}{n^2}\right) = O\left(\frac{\mu^2 \log^5 n}{n^2}\right).$$

The choice of k also implies that $\mu = \exp(\Omega((\log\log n)\log n))$, so $\mu^2/(\mu+\Delta) = \Omega(n^2/\log^5 n)$, as required. ■

Colouring rate and anti-concentration

Theorem 2.5 answers the latter question in the Introduction, but the story does not end here. We mentioned that Shamir and Spencer proved that $\chi(G_{n,1/2})$ is contained in an interval of length roughly $O(\sqrt{n})$ a.a.s., but it remains unclear precisely where the interval could be situated. (Note that this interval can be narrowed by a $\log n$ factor, see [9, Ex. 7.9].) Theorem 2.5 has been tightened recently, as we see in the following theorem, but the resulting explicit interval is still comparatively large, having length $\Theta(n/\log^2 n)$.

Theorem 2.6 *Let* $k_1 = 2\log n - 2\log\log n - 2$ $(= k_0 - 2\log e - 1)$. *Then*

$$\frac{n}{k_1 + o(1)} \le \chi(G_{n,1/2}) \le \frac{n}{k_1 - 1 - o(1)} \text{ a.a.s.}$$

https://doi.org/10.1017/CBO9781139519793.013 Published online by Cambridge University Press

The lower bound, due to Panagiotou and Steger [78] in 2009, and the upper bound, due to Fountoulakis, Kang and McDiarmid [38] a year later, improved upon the best previously known range (in [67]) from some two decades earlier. We prefer to restate the result in a modified form. We define the *colouring rate* $\overline{\alpha}(G)$ of a graph G to be $|V(G)|/\chi(G)$, which is the maximum average size of a colour class in a proper colouring of G. The above theorem says that the colouring rate of $G_{n,1/2}$ a.a.s. lies in the following interval of length $1 + o(1)$:

$$k_1 - 1 - o(1) \leq \overline{\alpha}(G_{n,1/2}) \leq k_1 + o(1).$$

Notice in particular that $\overline{\alpha}(G_{n,1/2})$ is strictly smaller than $\alpha(G_{n,1/2})$ a.a.s. A natural problem is to determine the a.a.s. behaviour of $\overline{\alpha}(G_{n,1/2})$ up to a $o(1)$ additive error.

What is the length of a smallest interval I for which $\mathbb{P}(\chi(G_{n,1/2}) \in I) \geq \varepsilon$? Lower bounds on this quantity are called 'anti-concentration' results. Various behaviours for anti-concentration of the chromatic number have been proposed (see [7], [16]). This problem is likely to be far from easy: while the length of the interval is at most about $O(\sqrt{n}/\log n)$, there are currently no results that show that $\chi(G_{n,1/2})$ is not contained a.a.s. in an interval of *constant* length. The possibility has been suggested [7] ('conjecture' would be too strong a word) that there may even be non-monotonicity in n – that is, it might be the case that very sharp concentration holds for infinitely many values of n, while for infinitely many other values it does not.

3. Sparse random graphs

In this section we consider the chromatic number for $G_{n,p}$ when $p = p(n)$ satisfies $p \to 0$ as $n \to \infty$. First we see how the chromatic number becomes more sharply concentrated for sparser graphs – that is, for $p(n)$ tending to 0 more quickly. In the second subsection we consider the specific case $p = d/n$, for some fixed choice of $d > 1$. By a standard equivalence, this amounts to the first displayed question in the Introduction. This question has recently benefited from the influence of analysis and statistical physics, leading to dramatic progress.

Expose-and-merge and concentration

A subtle colouring method of Matula [64], dubbed the 'expose-and-merge' algorithm, was used to prove Theorem 2.5 without the need for martingale- or Talagrand-type concentration, but rather with the more elementary inequalities of Chebyshev and Chernoff, as shown by Matula and Kučera [65]. For sparse $G_{n,p}$, Bollobás's method of extracting the colour classes one by one could not be extended further, though it worked down to edge probability about $p = n^{-1/3}$ (that is, expected average degree about $n^{2/3}$). However, Łuczak [59] realized that Matula's algorithm and the concentration bounds key to Bollobás's approach could be combined effectively to

https://doi.org/10.1017/CBO9781139519793.013 Published online by Cambridge University Press

prove the following stronger result. This significantly extended the range of $p = p(n)$ for the known a.a.s. first-order term of $\chi(G_{n,p})$.

Theorem 3.1 *There is a constant d_0 such that, if $p = p(n)$ satisfies $np > d_0$ for n sufficiently large and $p = o(1)$, then a.a.s.*

$$\chi(G_{n,p}) = \frac{np}{2\ln(np)}\left(1 + \Theta\left(\frac{\ln\ln(np)}{\ln(np)}\right)\right).$$

Outline of proof For simplicity, we assume throughout that $p < \ln^{-7} n$. The lower bound follows by noticing that the first moment argument used for the upper bound in Theorem 2.1 extends routinely to the sparse regime:

$$\alpha(G_{n,p}) \leq (2\ln(np) - 2\ln\ln(np) + 2 - 2\ln 2 + \varepsilon)/p \text{ a.a.s.}$$

For the upper bound, we employ the expose-and-merge method. We are somewhat rough with our description of this; for a more precise and detailed account, see [49, Sec. 7.5]. The engine behind this technique is a result of the following type, which a.a.s. guarantees the existence of many disjoint stable sets of nearly maximum order.

Theorem 3.2 *Fix $\varepsilon > 0$. There exists C_ε such that if $C_\varepsilon/n \leq p \leq \ln^{-7} n$ then, with probability at least $1 - 1/n$, $G_{n,p}$ contains a disjoint collection of y stable sets, each of order at least k, where*

$$k = k(n) = (2\ln(np) - 2\ln\ln(np) + 2 - 2\ln 2 - \varepsilon)/p \text{ and } y = y(n) = \frac{n}{k\ln^5(np)}.$$

Matula and Kučera proved a dense version of this result by a lengthy though uncomplicated second moment computation. For the form stated here, specific to sparse random graphs, the second moment is not as well behaved. Luckily, it does not grow too quickly, so the standard argument succeeds when it is strengthened by the appropriate application of a martingale concentration inequality. This neat trick was first employed by Frieze [41] to determine the stability number of sparse $G_{n,p}$.

With Theorem 3.2 in hand, we may attempt to extract from $G_{n,p}$ several large stable sets at a time. Set $\hat{n} = n/\ln^C(np)$ for a fixed positive constant C to be determined and, with a view to applying Theorem 3.2 repeatedly, set

$$\hat{k} = k(\hat{n}) \sim (2\ln(np) - 2(C+1)\ln\ln(np))/p \text{ and } \hat{y} = y(\hat{n}) \sim \frac{n}{\hat{k}\ln^{C+5}(np)},$$

where the asymptotics implied in $\sim$ are as $np \to \infty$.

To begin, choose, or rather *expose*, some subset A_1 of $[n]$ uniformly at random among all subsets of order $\hat{n}$. The induced subgraph $H_1 = G_{n,p}[A_1]$ may be viewed as $G_{\hat{n},p}$, so by Theorem 3.2, with probability at least $1 - 1/\hat{n}$, it contains a collection of

https://doi.org/10.1017/CBO9781139519793.013 Published online by Cambridge University Press

$\hat{y}$ disjoint stable sets each of order $\hat{k}$. Let us choose one such collection, $S_1^1, S_1^2, \ldots, S_1^{\hat{y}}$, uniformly at random (giving each subset of A_1 of order $\hat{k}$ a fair chance to be part of this collection). We label all vertices in $\bigcup_j S_1^j$ as *used* and all pairs of vertices in A_1 as *exposed*. Due to our extraction of a particular collection of stable sets from H_1, we can no longer view the exposed pairs as random edges.

Next we choose a new subset A_2 of order $\hat{n}$ uniformly at random from only the unused vertices. Unfortunately, we may not view the induced subgraph $G_{n,p}[A_2]$ as $G_{\hat{n},p}$ because its structure depends possibly on some exposed pairs. The bald but critical idea now is simply to forget about this problem for a while. More precisely, for each exposed pair from A_2, we resample by performing a new independent random experiment to include that pair as an edge with probability p. Let H_2 denote the graph on A_2 formed in this way. The key property of H_2 is that it may now be viewed as an independent copy of $G_{\hat{n},p}$ and, using Theorem 3.2 as above, we may choose a collection, $S_2^1, S_2^2, \ldots, S_2^{\hat{y}}$, of disjoint stable sets each of order $\hat{k}$. We now label all vertices in $\bigcup_j S_2^j$ as *used* and all pairs of vertices in A_2 as *exposed*.

Over $\hat{x} = \ln^{C+5}(np) - \ln^{C+2}(np)$ iterations, the probability that some iteration fails to generate a collection of $\hat{y}$ disjoint stable sets, each of order $\hat{k}$, is at most $\hat{x}/\hat{n} = o(1)$. With all but negligible probability, we have at the end a collection of $\hat{x}\hat{y} \sim np/(2\ln(np))$ disjoint sets S_i^j such that $S_i^1, S_i^2, \ldots, S_i^{\hat{y}}$ are all stable in H_i. It now remains to show how to *merge* these sets into a proper colouring of nearly all of $G_{n,p}$. In particular, we sketch how the sets S_i^j can be pruned slightly to form a proper colouring of all but at most $O(n/\ln^3(np))$ vertices. (It can then be checked that the graph induced by the remaining vertices has small enough degeneracy that it can be polished off greedily, using only a few extra colours — recall that *degeneracy* is defined as the maximum over all induced subgraphs of the minimum degree.) The crux here is that, due to our temporary amnesia, although S_i^j is a stable set in H_i, it might not be stable in $G_{n,p}$. We are in trouble if S_i^j contains some pair of adjacent vertices that was exposed in a previous iteration but resampled as a non-edge in the ith iteration. With an appropriate choice of C and a first moment estimation, it is indeed possible (although we omit the details) to show that the number Y of such troublesome pairs is small, satisfying $\mathbb{P}(Y > n/\ln^3(np)) = O(1/\ln^2(np))$. Thus a.a.s. we may safely remove Y vertices from $\bigcup_{i,j} S_i^j$, one from each troublesome pair, to obtain the required proper colouring of nearly all of $G_{n,p}$. ■

As in the dense regime, in the sparse case one can prove strong concentration results for $\chi(G_{n,p})$ without knowing an asymptotic value. In addition to the aforementioned bound for the dense case, Shamir and Spencer showed with martingales that, for $p = n^{-\alpha}$ with $\alpha > \frac{1}{2}$, $\chi(G_{n,p})$ is a.a.s. confined to an interval whose width is a constant dependent upon α. A few years later, their proof was sharpened by Łuczak [60] to demonstrate *two*-point concentration when $\alpha > \frac{5}{6}$. (By continuity considerations, there must exist edge probability functions $p(n)$ that evenly balance the probability mass between the events that the chromatic number evaluates to one of two consecutive integers; so two-point concentration is best possible in general.)

https://doi.org/10.1017/CBO9781139519793.013 Published online by Cambridge University Press

Another few years later, Alon and Krivelevich [7] extended the two-point result to hold for all (fixed) $\alpha > \frac{1}{2}$, using recolouring via the Lovász local lemma.

This brings us close to the situation that we encountered in the dense case. Martingale techniques show that there is a small range of values that the chromatic number can take, but they do not give any hint of where it is located. The explicit interval implied by Theorem 3.1 is of width $\Theta(np \ln\ln(np)/\ln^2(np))$. This is unbounded (and certainly larger than 2) if np is unbounded. On the other hand, if np is bounded, then we have the intuition (or at least some hope) that it would be possible to identify (nearly) exact values. We pursue this next.

Sharp thresholds for k-colourability

Random graphs exhibit behaviours that might seem counter-intuitive at first sight. Many properties or parameters of $G_{n,p}$, such as the stability number or the number of proper k-colourings, experience a sudden change across a small interval of p. This is sometimes referred to as a *threshold* (or *phase transition*, as mentioned earlier). It is well known that thresholds arise in wider contexts in mathematics, physics, economics, sociology and elsewhere. In $G_{n,p}$, their study is of great importance. The demonstration by Bollobás and Thomason [18] that thresholds exist in a general sense for graph properties (families of graphs that are closed under isomorphism) that are monotone (closed under edge-removal) is one of the fundamental results in random graphs.

In 1999 Friedgut [40] investigated how suddenly a given phase transition takes place. Imprecisely, given a property of $G_{n,p}$ that undergoes a transition, consider the ratio of the following two quantities:

- the length of the window of p over which the transition takes place
- the critical probability p_c for which G_{n,p_c} possesses the property with probability exactly $\frac{1}{2}$ (which is guaranteed to exist).

The property's threshold is *sharp* if as $n \to \infty$ this ratio tends to 0, whereas if the ratio is bounded away from 0 it is a called a *coarse* threshold. (The result of Bollobás and Thomason showed the ratio to be bounded above for any monotone property.) The foundational example of a sharp threshold is the critical probability for connectivity of $G_{n,p}$: it is located at around $p = (\ln n)/n$, but the length of the interval over which $G_{n,p}$ changes from being disconnected to connected is of order $1/n$. The seminal work of Friedgut used Fourier analytic techniques to give a deep criterion for monotone graph properties in $G_{n,p}$ to have sharp thresholds. In rough terms, a monotone graph property with a coarse threshold can be approximated by the property that it contains one of a fixed collection of graphs as a subgraph. With this characterization, Achlioptas and Friedgut [1] showed that, for $k \geq 3$, the property of k-colourability (that is, of having chromatic number at most k) has a sharp threshold sequence.

https://doi.org/10.1017/CBO9781139519793.013 Published online by Cambridge University Press

Theorem 3.3 *Let $k \geq 3$ be an integer. There exists $d_k = d_k(n)$ such that, for any $\varepsilon > 0$,*

$$\lim_{n\to\infty} \mathbb{P}\left(\chi\left(G_{n,(d_k-\varepsilon)/n}\right) \leq k\right) = 1 \text{ and } \lim_{n\to\infty} \mathbb{P}\left(\chi\left(G_{n,(d_k+\varepsilon)/n}\right) \leq k\right) = 0.$$

(It is known that the threshold for 2-colourability is coarse, since the probabilities of $G_{n,d/n}$ having an odd cycle and not having an odd cycle are both bounded away from 0 for $d \in (0, 1)$.) From Theorem 3.3, we deduce that $\chi(G_{n,d/n})$ would be one-point concentrated for all but a vanishing proportion of $d > 1$ (in Lebesgue measure) if it could be shown that $d_k(n)$ converges as $n \to \infty$, or even that $d_k(n)$ is constrained to an interval of length independent of k for all large enough n. Unfortunately, Achlioptas and Friedgut's result does not directly yield anything on the actual value of the chromatic number.

The question of convergence of $d_k(n)$ as $n \to \infty$ is open, but in a breakthrough in 2005 Achlioptas and Naor [3] found an explicit interval within which the limit must lie if it exists.

Theorem 3.4 *Let $k \geq 3$ be an integer. For any $\varepsilon > 0$,*

$$\lim_{n\to\infty} \mathbb{P}\left(\chi\left(G_{n,(2(k-1)\ln(k-1)-\varepsilon)/n}\right) \leq k\right) = 1$$
$$\text{and} \quad \lim_{n\to\infty} \mathbb{P}\left(\chi\left(G_{n,((2k-1)\ln k)/n}\right) \leq k\right) = 0.$$

Thus the sharp threshold sequence $d_k = d_k(n)$ for k-colourability (as guaranteed by Theorem 3.3) satisfies, for all n,

$$2(k-1)\ln(k-1) \leq d_k \leq (2k-1)\ln k. \tag{1}$$

An immediate corollary of Theorem 3.4 is that a.a.s. the chromatic number of $G_{n,d/n}$ must take either one explicit value or one of two values.

Corollary 3.5 *Given $d > 0$, let k_d be the least integer k for which $d < 2(k-1)\ln(k-1)$. Then $\chi(G_{n,d/n})$ is $k_d - 1$ or k_d a.a.s. If $d \geq (2k_d - 3)\ln(k_d - 1)$, then $\chi(G_{n,d/n}) = k_d$ a.a.s.*

Building upon this work and combining it carefully with some other ideas used for existential sharp concentration of the chromatic number, Coja-Oghlan, Panagiotou and Steger [26] have obtained an explicit two- or three-point concentration result for denser $G_{n,p}$, specifically for $p = n^{-\alpha}$ with $\alpha > \frac{3}{4}$. Recently, Dyer, Frieze and Greenhill [33] have obtained an analogue of Theorem 3.4 for the weak chromatic number of binomial random k-uniform hypergraphs.

We briefly outline some of the ideas of the proof of Theorem 3.4. First of all, rather than $G_{n,d/n}$, it is easier, and also does not hurt, to work with the random *multigraph* $\widetilde{G}_{n,m}$ with $m = \frac{1}{2}dn$, where we throw in m edges independently and uniformly at random, allowing repetitions. Counting the number of proper k-colourings of $\widetilde{G}_{n,m}$ is equivalent to summing, over all possible partitions of $[n]$

https://doi.org/10.1017/CBO9781139519793.013 Published online by Cambridge University Press

into k parts, the indicator variable that the given partition is a proper colouring. The right-hand side of (1), as first observed by Devroye (see [23]), is a simple first moment estimate. It follows from the fact that a given partition into k parts is a proper colouring with probability at most $(1-1/k)^m$, and thus the expected number of proper k-colourings of $\widetilde{G}_{n,m}$ is at most $k^n(1-1/k)^m$.

For the much more difficult part, the lower bound on d_k, Achlioptas and Naor made a consummate application of the second moment method, which they reduced to a high-dimensional optimization problem. A powerful simplification is to restrict attention to partitions of $[n]$ that are *balanced* – that is, each part has n/k vertices (where we have also assumed that k divides n). These account for most proper k-colourings of $\widetilde{G}_{n,m}$. The number of balanced proper k-colourings of $\widetilde{G}_{n,m}$ is a sum of $n!/((n/k)!)^k$ indicators, and it follows that its first moment is

$$\frac{n!}{((n/k)!)^k}\left(1-\frac{1}{k}\right)^m.$$

To bound the second moment of this random variable, we bound the probability, for a given pair of balanced partitions, that both are proper k-colourings of $\widetilde{G}_{n,m}$. Using the property that the partitions are balanced, we may see that this probability is

$$\left(1-\frac{2}{k}+\sum_{i,j}\frac{{\ell_{ij}}^2}{n^2}\right)^m,$$

where ℓ_{ij} is the number of vertices of $[n]$ with colour i in one of the balanced partitions and colour j in the other. So the correlation between two variables that indicate whether a given balanced partition is proper is governed by k^2 parameters. Consider the set $\mathcal{L}$ of all $k \times k$ matrices $\mathbf{L} = (\ell_{ij})$ of non-negative integers for which the sum of each row and each column is n/k. The second moment of the number of balanced proper k-colourings can be expressed as

$$\sum_{\mathbf{L}\in\mathcal{L}}\frac{n!}{\prod_{i,j}\ell_{ij}!}\left(1-\frac{2}{k}+\sum_{i,j}\frac{{\ell_{ij}}^2}{n^2}\right)^{dn/2}.$$

Optimization over $\mathcal{L}$ can be translated into optimization over the Birkhoff polytope $\mathcal{B}_k$, the set of $k \times k$ doubly stochastic matrices. Thus, to bound the second moment appropriately, it suffices to show a particular entropy–energy inequality over $\mathcal{B}_k$. Achlioptas and Naor did so by first relaxing to singly stochastic matrices and then using sophisticated geometric and analytic ideas. Their argument then shows that the probability of being k-colourable is bounded above 0, whereupon an application of Theorem 3.3 implies the theorem.

Recent work of Coja-Oghlan and Vilenchik [27] and of Coja-Oghlan [24] has markedly improved both ends of the interval (1) (which we remark has length $(1+o(1))\ln k$ as $k \to \infty$). Guided by important developments in statistical physics, which used applications of the so-called 'cavity method' and 'belief/survey

https://doi.org/10.1017/CBO9781139519793.013 Published online by Cambridge University Press

propagation' algorithms [21], [76], they have obtained a range of length that is independent of k:

$$(2k-1)\ln k - 2\ln 2 - o(1) \leq d_k \leq (2k-1)\ln k - 1 + o(1), \tag{2}$$

where the $o(1)$ asymptotics are as $k \to \infty$. The above range for the k-colourability sharp threshold sequence (particularly the improved lower bound) is quite close to completely settling the original question of Erdős and Rényi from 1960. It shows that $\chi(G_{n,d/n})$ is concentrated on an explicit value – that is, the least integer k for which $d < (2k-1)\ln k - 2\ln 2$ holds – for all but a vanishing proportion of $d > 1$.

4. Random regular graphs

Given an integer $r \geq 1$, let $G_{n,r}$ be a graph chosen uniformly at random from the set of r-regular graphs with vertex-set $[n]$ (where rn is even, for feasibility). This model of random graphs has received much attention since the late 1970s, when Bollobás [13] probabilistically reinterpreted an enumerative formula of Bender and Canfield [11] using the now-standard 'configuration' or 'pairing' model. In some aspects, $G_{n,r}$ is qualitatively quite distinct from $G_{n,p}$; for example, $G_{n,r}$ has stronger connectivity properties at lower edge densities: notably, $G_{n,3}$ is Hamiltonian a.a.s. However, for larger r, $G_{n,r}$ seems to behave more like $G_{n,p}$; this was partially formalized by Kim and Vu [54]. (The rough intuition behind this is that, above a certain density, all the degrees of $G_{n,p}$ are highly concentrated around np.) Techniques for proving properties of $G_{n,r}$ are usually more difficult though, and are close in spirit to those used in analytic and enumerative combinatorics. The refinement of these techniques is useful in the study of degree-constrained random graphs, with relevance to real-world network modelling. Wormald [88] has a broad account of the theory of random regular graphs. It is important to note that, unlike $G_{n,p}$, the model $G_{n,r}$ is not generally one that is monotone in density. In other words, it is not true in general that, if a monotone property holds in $G_{n,r}$ a.a.s., then it also holds in $G_{n,r+1}$ a.a.s. (as can be seen by considering the property of containment of a perfect matching); however, for *fixed* $r \geq 2$, this statement is true as a consequence of 'contiguity' (see, for example, [49] or [88]).

Two of the first results on the chromatic number of $G_{n,r}$ were given in 1992. Molloy and Reed bounded the expected number of proper k-colourings in the pairing model of $G_{n,r}$, with r fixed, to conclude that, if $k(1-1/k)^{r/2} < 1$, then $\chi(G_{n,r}) > k$ a.a.s. This argument was given in Molloy's University of Waterloo Master's thesis, but it has been reiterated in the more easily obtained paper [84]. Frieze and Łuczak [43] further analyzed the pairing model to show that, if $r = r(n) < n^{1/3-\varepsilon}$ for some $\varepsilon > 0$ and $r \geq r_0$ for some constant r_0, then a.a.s.

$$\chi(G_{n,r}) = \frac{r}{2\ln r} + \Theta\left(\frac{r\ln\ln r}{\ln^2 r}\right).$$

https://doi.org/10.1017/CBO9781139519793.013 Published online by Cambridge University Press

About a decade later, with an 'edge-switching' strategy, Cooper, Frieze, Reed and Riordan [28] extended this asymptotic statement to hold if $r < n^{1-\varepsilon}$ for some $\varepsilon > 0$ (and $r \geq r_0$ for some constant r_0). In an independent and complementary project, Krivelevich, Sudakov, Vu and Wormald [57] used an expression for the probability that $G_{n,p}$ is r-regular to show in the range $n^{6/7+\varepsilon} < r < 0.9n$ that $\chi(G_{n,r}) \sim n/(2\log_b r)$, where $b = n/(n-r)$. An easy calculation checks that these estimates of $\chi(G_{n,r})$ agree asymptotically as $r \to \infty$ with those of $\chi(G_{n,p})$ with $p = r/n$.

As it did for $G_{n,p}$, significant progress on the determination of $\chi(G_{n,r})$ in the sparse case (when r is constant) came later. The values $r \leq 3$ are easy: trivially, $\chi(G_{n,1}) = 2$ a.a.s. For $r \geq 2$, $G_{n,r}$ contains an odd cycle a.a.s. and so $\chi(G_{n,2}) = 3$ a.a.s.; also a.a.s. $\chi(G_{n,3}) = 3$, since $G_{n,3}$ contains no K_4-subgraph and the statement follows from Brooks's theorem. The next values are more involved. For $4 \leq r \leq 10$, Shi and Wormald [83], [84] provided a fine analysis in 2007, using the *differential equations method* [87], of a tailored colouring algorithm, to obtain explicit one- or two-point determinations. (The idea behind this general analytic method is to show that a random variable of interest can be tracked accurately by some function expressed as the solution to a set of differential equations.) In particular, Shi and Wormald proved that $\chi(G_{n,4}) = 3$, $\chi(G_{n,5})$ is 3 or 4, and $\chi(G_{n,6}) = 4$ a.a.s.

A further advance was obtained a few years later. Improving on a previously-announced bound of Achlioptas and Moore [2], Kemkes, Pérez-Giménez and Wormald [53] used *subgraph conditioning* (see [88]) to prove the following theorem.

Theorem 4.1 *Given $r > 0$, let k_r be the least integer k for which $r < 2(k-1)\ln(k-1)$. Then $\chi(G_{n,r})$ is $k_r - 1$ or k_r a.a.s. If $r \geq (2k_r - 3)\ln(k_r - 1)$, then $\chi(G_{n,r}) = k_r$ a.a.s.*

The lower bound is just the aforementioned bound of Molloy and Reed. As we have often seen, the most difficult part in obtaining the upper bound is controlling the second moment. Even after an adaptation of the sophisticated analysis of Achlioptas and Naor, the second moment is too large for the application of either Chebyshev's inequality or the Paley–Zygmund inequality directly, and no sharp threshold-type result is available. In particular, the second moment is some constant times the square of the first moment. The broad intuition for the subgraph conditioning method is that much of the magnitude of the second moment may depend on the presence of some specific small but not too common subgraphs, such as short cycles. When the method applies, conditioning on counts of enough of these small subgraphs does not change the first moment by too much, but it reduces the second moment to any desired fraction of the original. Kemkes, Pérez-Giménez and Wormald showed that this method is applicable to the number of balanced k-colourings of $G_{n,r}$, in order to obtain the upper bound on $\chi(G_{n,r})$ in this theorem.

A comparison of Theorem 4.1 and Corollary 3.5 suggests that $G_{n,r}$ and $G_{n,r/n}$ have a close affinity with respect to k-colourability, which contrasts with the broader qualitative differences we pointed out earlier. In a recent manuscript, Coja-Oghlan,

https://doi.org/10.1017/CBO9781139519793.013 Published online by Cambridge University Press

Efthymiou and Hetterich [25] have proved a more direct connection between these parameters, and used that link to tighten the bounds on $\chi(G_{n,r})$ to match the bounds on $\chi(G_{n,p})$ in (2). In particular, there is some ε_k with $\varepsilon_k \to 0$ as $k \to \infty$ such that if $r \leq (2k-1)\ln k - 2\ln 2 - \varepsilon_k$, then $\chi(G_{n,r}) \leq k$ a.a.s., whereas if $r \geq (2k-1)\ln k - 1 + \varepsilon_k$, then $\chi(G_{n,r}) > k$ a.a.s.

For the upper bound, they combined the second moment calculations in [27] with the subgraph conditioning methodology in [53]. The lower bound demands more technical care, and for this they study the geometry of the set of proper k-colourings and use large deviations analysis.

Note that the result that $\chi(G_{n,4}) = 3$ a.a.s. is not captured by Theorem 4.1 or the more recent work [25]. The conjecture that $\chi(G_{n,5}) = 3$ a.a.s. remains open, but there is some positive evidence for it in [30]. Ben-Shimon and Krivelevich [10] have used the edge-switching method to prove existential two-point concentration of $\chi(G_{n,r})$ for $r = o(n^{1/5})$, though since the continuity argument fails for $G_{n,r}$ it is unclear even whether two-point concentration is best possible. For fixed r, it seems very likely that $\chi(G_{n,r})$ is always uniquely determined a.a.s. This is true with an explicit determination for 'most' values of r, by the new results in [25].

Random Cayley graphs

So far in this section, we have focused on the uniform model of a regular graph. We now mention briefly a very different algebraically defined family of random regular graph models that has recently garnered heightened interest, including in regard to the chromatic number. Let B be a finite group of order n and let $k \in [\lfloor \frac{1}{2}n \rfloor]$. Choose a subset $A \subseteq B$ by first choosing k elements of B independently and uniformly at random (with repetitions), and then letting A be these elements together with their inverses, without the identity. The *random Cayley graph* with respect to B and k is the graph with vertex-set B, where b_1 and b_2 are adjacent when $b_1 \cdot {b_2}^{-1} \in A$; this graph is regular of degree $|A| \leq 2k$. The study of such random graphs is important in many settings, such as information theory, computational complexity and additive combinatorics.

The behaviour of the chromatic number of random Cayley graphs is currently an active research area; we refer the interested reader to Alon [6] for a systematic study, where, for instance, it is shown that the chromatic number of the random Cayley graph with respect to any B and k is $O(k/\ln k)$. (Alon furthermore gives general lower bounds that nearly match the upper bound in the dense case $k = \Theta(n)$ but worsen for smaller k.) We also point to a recent manuscript of Green and Morris [47] on *random Cayley sum graphs* with respect to $\mathbb{Z}/n\mathbb{Z}$. In this model, $A \subseteq \mathbb{Z}/n\mathbb{Z}$ is chosen by including each element independently at random with probability $\frac{1}{2}$; then $x, y \in \mathbb{Z}/n\mathbb{Z}$ are adjacent in the graph when $x+y \in A$. With methods from additive combinatorics, such as an 'arithmetic regularity lemma', they prove that such graphs a.a.s. have no clique or stable set of order more than $(2+o(1))\log n$. A matching upper bound on

https://doi.org/10.1017/CBO9781139519793.013 Published online by Cambridge University Press

the chromatic number, although with a technical restriction that $\gcd(n, 6) = 1$, was recently proved by Green [46].

5. Random geometric graphs

Suppose that we wish to assign bands of radio frequencies to a collection of active transmitters scattered in some region. Minimizing the amount of spectrum allocated can be modelled (albeit with many simplifying assumptions) as follows. Given a collection $\mathcal{X} = \mathcal{X}_n$ of independent random points $X_1, X_2, \ldots, X_n \in \mathbb{R}^d$ with common probability distribution ν and a positive distance parameter $r = r(n) > 0$, we first construct a *random geometric graph* $G_n(\mathcal{X}, r)$ with vertex-set $[n]$, where i and j are adjacent when $\|X_i - X_j\| < r$. Then, taking $\|\cdot\|$ as the ordinary Euclidean norm and ν as a uniform distribution over $[0, 1]^2$, we may imagine the points X_i as transmitters in the plane and r as their common transmission (and interference) range. We are interested in the asymptotic behaviour of $G_n(\mathcal{X}, r)$ as $n \to \infty$ and assume that $r = r(n) = o(n)$. (In fact, the norm $\|\cdot\|$ may be any norm and ν may be any probability distribution with a bounded density function over $\mathbb{R}^d$.) We may thus consider a proper colouring as a good spectrum allocation, so that $\chi(G_n(\mathcal{X}, r))$ indicates the minimum amount needed.

The random geometric graph is a much-studied model because of its simple and natural definition, its history going back more than half a century, and its well-established links to areas such as statistics, probability theory and electrical engineering. Background and further properties of $G_n(\mathcal{X}, r)$ are covered in detail in the monograph of Penrose [79]. Intuitively, the parameter $r(n)$ plays the same role that the edge probability parameter $p(n)$ does in $G_{n,p}$. It controls density: the average degree of $G_n(\mathcal{X}, r)$ is $\Theta(r^d n)$ in general, and is about $\pi r^2 n$ for the Euclidean norm and a uniform distribution over $[0, 1]^2$. As we might expect, $G_n(\mathcal{X}, r)$ goes through a number of phases as r increases. We remark here that most of the results in this section are given with stronger convergence than a.a.s.: a property holds *almost surely* (*a.s.*) if it holds with probability 1.

We will see that there is a threshold in the a.s. behaviour of $\chi(G_n(\mathcal{X}, r))$. Loosely speaking, the chromatic number behaves very 'locally' in the sparse regime when the average degree is $o(\ln n)$, whereas it behaves essentially as if the points are uniformly spread in the dense regime when the average degree is $\omega(\ln n)$; the behaviour at the threshold $\Theta(\ln n)$ is most interesting. Moreover, the chromatic number of $G_n(\mathcal{X}, r)$ is quite close to both the clique number and the maximum degree, in contrast with $G_{n,p}$. This bolsters our natural intuition from the fact that the structure of $G_n(\mathcal{X}, r)$ is locally determined. (We comment here that, for a 'deterministic' geometric graph G in the plane with Euclidean norm, it is known that $\chi(G) \leq 3\omega(G) - 2$ and $\Delta(G) \leq 6\omega(G) - 7$.)

Before stating the main result, we define the parameters required of the common probability measure ν and the norm $\|\cdot\|$. The *maximum density* σ of ν is the essential

https://doi.org/10.1017/CBO9781139519793.013 Published online by Cambridge University Press

supremum of the probability density function f of ν – that is to say, $\sigma = \sup\{t : \mathrm{vol}(\{x : f(x) > t\}) > 0\}$, where vol denotes the (Lebesgue measure) volume. For $K > 0$, let $N(K)$ be the maximum cardinality of a collection of pairwise disjoint translates of the unit ball $B = \{x \in \mathbb{R}^d : \|x\| < 1\}$ with centres in $(0, K)^d$. The (*translational*) *packing density* δ of the unit ball B with respect to $\|\cdot\|$ is defined as

$$\delta = \lim_{K\to\infty} \frac{N(K)\,\mathrm{vol}(B)}{K^d}.$$

This limit always exists and satisfies $0 < \delta \le 1$. For the Euclidean norm in $\mathbb{R}^2$, it is known that $\delta = \pi/(2\sqrt{3}) \approx 0.907$. Informally, δ is the greatest proportion of $\mathbb{R}^d$ that can be filled with disjoint translates of B. We are now ready to see the full picture of phases for $\chi(G_n(\mathcal{X}, r))$.

Theorem 5.1 *For the random geometric graph $G_n(\mathcal{X}, r)$, the following hold.*

Very sparse regime: *Suppose that $r^d n \le n^{-\alpha}$ for some fixed $\alpha > 0$, and let $k = k(n) = \left\lfloor |\ln n/\ln(r^d n)| + \frac{1}{2} \right\rfloor$. Then a.s. $\chi(G_n(\mathcal{X}, r)) = \omega(G_n(\mathcal{X}, r))$ is k or $k+1$ for all but finitely many n.*
Sparse regime: *Suppose that $r^d n = \omega(n^{-\varepsilon})$ for all $\varepsilon > 0$ but $r^d n = o(\ln n)$. Then a.s. both $\chi(G_n(\mathcal{X}, r))$ and $\omega(G_n(\mathcal{X}, r))$ are asymptotic to $\ln n/\ln(\ln n/(r^d n))$.*
Intermediate regime: *Suppose that $\sigma r^d n/\ln n \to t \in (0, \infty)$. Then a.s.*

$$\chi(G_n(\mathcal{X}, r)) \sim f_\chi(t) \cdot \sigma r^d n \ \text{ and } \ \omega(G_n(\mathcal{X}, r)) \sim f_\omega(t) \cdot \sigma r^d n$$

where the functions f_χ and f_ω are given explicitly in [72]. *They depend only on d and $\|\cdot\|$, and are continuous; f_χ is non-increasing, and satisfies $f_\chi(t) \to \mathrm{vol}(B)/(2^d\delta)$ as $t \to \infty$, and $f_\chi(t) \to \infty$ as $t \downarrow 0$; and f_ω is strictly decreasing, and satisfies $f_\omega(t) \to \mathrm{vol}(B)/2^d$ as $t \to \infty$, and $f_\omega(t) \to \infty$ as $t \downarrow 0$.*
Dense regime: *Suppose that $r^d n = \omega(\ln n)$ but $r \to 0$. Then a.s.*

$$\chi(G_n(\mathcal{X}, r)) \sim \frac{\mathrm{vol}(B)}{2^d\delta} \cdot \sigma r^d n \ \text{ and } \ \omega(G_n(\mathcal{X}, r)) \sim \frac{\mathrm{vol}(B)}{2^d} \cdot \sigma r^d n.$$

The results on $\omega(G_n(\mathcal{X}, r))$ essentially already appeared in Penrose's book [79] of 2003, except that an assumption has been dropped and some results amplified. Both Penrose [79] and McDiarmid [68] studied the chromatic number in the sparse and dense regimes, but prior to the 2011 work of McDiarmid and Müller [72] little was known about the chromatic number in the most difficult intermediate regime, $r^d n = \Theta(\ln n)$. The proof of Theorem 5.1 took an optimization viewpoint. In particular, McDiarmid and Müller thoroughly investigated the fractional chromatic number, which is the solution to the linear programming relaxation of the natural integer linear programme for the chromatic number. This could be reduced to analyzing weighted integrals that correspond to feasible solutions of the dual linear programme for the fractional chromatic number. The result then follows on showing that the chromatic and fractional chromatic numbers are close.

https://doi.org/10.1017/CBO9781139519793.013 Published online by Cambridge University Press

The most striking consequence of Theorem 5.1 is that $\chi(G_n(\mathcal{X}, r))$ and $\omega(G_n(\mathcal{X}, r))$ have the same asymptotic value when $r^d n = o(\ln n)$, but the two quantities differ by a multiplicative factor (if the packing density δ is less than 1) when $r^d n = \omega(\ln n)$. McDiarmid and Müller also found a sharp threshold between these behaviours (again, if $\delta < 1$).

Theorem 5.2 *If $\delta < 1$, there is a critical parameter t_0, depending only on d and $\|\cdot\|$, for which the following hold for the random geometric graph $G_n(\mathcal{X}, r)$:*

- *if* $\limsup_{n\to\infty} \sigma r^d n / \ln n \le t_0$, *then* $\chi(G_n(\mathcal{X}, r))/\omega(G_n(\mathcal{X}, r)) \to 1$ *a.s.*
- *if* $\liminf_{n\to\infty} \sigma r^d n / \ln n > t_0$, *then* $\liminf_{n\to\infty} \chi(G_n(\mathcal{X}, r))/\omega(G_n(\mathcal{X}, r)) > 1$ *a.s.*

Müller [77] established a general principle of existential concentration for parameters of $G_n(\mathcal{X}, r)$. Specific consequences of this principle are that the clique number $\omega(G_n(\mathcal{X}, r))$, the chromatic number $\chi(G_n(\mathcal{X}, r))$ and the degeneracy $\delta^*(G_n(\mathcal{X}, r))$ of the random geometric graph are each a.a.s. two-point concentrated if $r^d n = o(\ln n)$. The two consecutive values upon which the chromatic number is concentrated remain unknown except in the very sparse regime as specified in Theorem 5.1. Furthermore, in the light of Theorem 5.2, Müller suggested that it would be interesting to determine the location of the threshold (if one exists) for the property that the difference $\chi(G_n(\mathcal{X}, r)) - \omega(G_n(\mathcal{X}, r))$ is bounded.

6. Random planar graphs and related classes

Given the storied history of chromatic graph theory, it would seem unjust to omit mention of planar and other embeddable graphs in a colouring survey. A considerable amount of work has been done on random planar graphs and related models, much of it in recent years. Let $\mathscr{P}_n$ be the set of planar graphs with vertex-set $[n]$ and let P_n be a graph chosen uniformly at random from $\mathscr{P}_n$. We already see the difficulty of this model in its definition: to get a good grip on this random object, we essentially need to enumerate planar graphs as combinatorial structures, regardless of any particular topological embedding. An advance in our understanding of P_n came about a decade ago, when McDiarmid, Steger and Welsh [74] used a superadditivity argument to show the existence of the (*labelled*) *planar graph growth constant* γ, satisfying

$$\lim_{n\to\infty} \left(\frac{|\mathscr{P}_n|}{n!} \right)^{1/n} = \gamma.$$

More recently, Giménez and Noy [45] built on work of Bender, Gao and Wormald [12] in a masterful application of analytic combinatorics. They applied classical graph-theoretic decompositions according to connectivity, showed that these translate into basic functional relationships among the corresponding generating functions, and then applied singularity analysis to these functions to determine γ exactly. The mere existence of γ, though, was already enough to derive a range of

https://doi.org/10.1017/CBO9781139519793.013 Published online by Cambridge University Press

properties of P_n, including results on the numbers of vertices of a given degree, of faces of a given size in any embedding, and of automorphisms. These results were given as consequences of what is now known as the 'appearances theorem' [74, Theorem 4.1], which says roughly that a.a.s. any connected planar graph appears in P_n as a pendant copy a linear number of times. Another consequence of the appearances theorem is that P_n fails to contain a K_4 subgraph with probability $e^{-\Omega(n)}$. This implies the following, the first part of which is immediate.

Theorem 6.1 *A.a.s.* $\chi(P_n) = 4$. *Furthermore, there is an expected quadratic-time algorithm to provide an optimal colouring of* P_n.

The second part of this theorem is an observation of Taraz and Krivelevich (see [74]). In linear time, we can distinguish between two cases. The first is that there is a K_4 subgraph, in which case we apply the quadratic-time 4-colouring algorithm for planar graphs. In the second case, which occurs with probability $e^{-\Omega(n)}$, we colour the graph optimally using the planar separator theorem, and this takes $O(c^{\sqrt{n}})$ time. This amounts to a quadratic expected running time. One might wonder whether there is a faster colouring algorithm or a shorter proof of 4-colourability specific to random planar graphs; however, the appearances theorem suggests this to be unlikely. With the above in mind, we might ask what more can be said about colouring random planar graphs. We take several tacks.

One natural direction, towards a probabilistic analogue of the Ringel–Youngs theorem, is to consider graphs embeddable in a fixed surface S. Let S_n be a graph chosen uniformly at random from the set of graphs embeddable in S with vertex-set $[n]$. These graph classes were first considered in [69], where it was shown, for example, that they all possess a growth constant, and indeed all have the *same* growth constant γ (as we met for planar graphs). It then follows by an appearances statement (linearly many pendant copies of each connected planar graph) that $\chi(S_n) \geq 4$ a.a.s. due to K_4 subgraphs. With additional ingredients from topological graph theory and generating function analysis, Chapuy, Fusy, Giménez, Mohar and Noy [22] obtained more precise results on S_n. One consequence of their work was that a.a.s. S_n has arbitrarily large *face-width*, indicating local planarity. A result of Thomassen [85] then implies that a.a.s. S_n is properly 5-colourable. Thus the following holds.

Theorem 6.2 *A.a.s.* $\chi(S_n)$ *is* 4 *or* 5.

It was conjectured in [22] that 4 is a.a.s. the correct value. Recall from Chapter 6 that the *choosability* $\mathrm{ch}(G)$ of a graph G is the least k for which, for any assignment of lists of at least k colours at each vertex, there is always a proper colouring of G so that each vertex is given a colour from its list. (We discuss choosability further in Section 7.) The choosability of S_n has been determined exactly: a list-colouring extension [29] of the aforementioned result of Thomassen, together with the appearances statement for some non-4-choosable (connected) planar graph, imply that $\mathrm{ch}(S_n) = 5$ a.a.s.

https://doi.org/10.1017/CBO9781139519793.013 Published online by Cambridge University Press

From surfaces, one could continue further by studying graphs chosen uniformly from a minor-closed class, as was done in [70]. We comment on one particular aspect of this study which is relevant to chromatic graph theory. In the Introduction, we mentioned that $G_{n,1/2}$ generates a plethora of counter-examples to an over-strengthened form of Hadwiger's conjecture. Let us rather consider Hadwiger's conjecture itself from the viewpoint of random graphs. Recall that this conjecture asserts that, for any positive integer k, each k-chromatic graph has K_k as a minor. It has been proved for $k \leq 6$. Fix an integer $k \geq 7$, and let $\mathrm{Ex}(K_k)$ denote the class of graphs without K_k as a minor. Since the class $\mathrm{Ex}(K_k)$ is 'addable' (meaning that it satisfies some simple decomposition closure properties), it then has a growth constant and so admits an appearances theorem (linearly many pendant copies of each connected member of $\mathrm{Ex}(K_k)$). So, if Hadwiger's conjecture fails for k, then it should be 'easy' to find counter-examples: if G_n is chosen uniformly at random from the set of graphs in $\mathrm{Ex}(K_k)$ on the vertex-set $[n]$, then $\chi(G_n) \geq k$ with probability $1 - e^{-\Omega(n)}$.

Next, we consider random planar graphs of a given density. Let $P_{n,m}$ be a graph chosen uniformly at random from the set of all planar graphs on $[n]$ with m edges; any such graph satisfies $m \leq 3n - 6$. This model is directly inspired by $G_{n,m}$ or $G_{n,p}$, although the condition of planarity clearly imposes extra difficulties. As a by-product of their analytic work, Giménez and Noy exactly determined the growth constant for the asymptotic number of planar graphs on $[n]$ with cn edges, for $c \in (1, 3)$. Recently, M. Kang and Łuczak [50], investigating what happens around $c = 1$, proved (by a combination of counting and analytic methods) a phenomenon akin to, though qualitatively distinct from, the emergence of the 'giant component' in $G_{n,p}$. Around $c = 1$, something also occurs with respect to the chromatic number of $P_{n,m}$.

Theorem 6.3
If $\limsup_{n\to\infty} m/n < 1$, *then* $\chi(P_{n,m}) \leq 3$ *a.a.s.*
If $\liminf_{n\to\infty} m/n > 1$, *then* $\chi(P_{n,m}) = 4$ *a.a.s. and there is an expected quadratic-time algorithm for an optimal colouring of* $P_{n,m}$.

Kang and Łuczak showed the first part as a corollary to structural properties that they proved for $P_{n,m}$ in the sparse regime. The second part follows from work of Dowden; in particular, the argument is similar to that above for P_n, but with an appearances theorem for random planar graphs of a given edge density (up to and including 3) (see [31, Theorem 61] and [32]). This demonstrates a phase transition in the behaviour of $\chi(P_{n,m})$. It would be of interest to understand more precisely how $\chi(P_{n,m})$ behaves when $m/n \to 1$.

Our last tack is to consider the stability number of P_n. Obviously, $\alpha(P_n) \geq \frac{1}{4}n$, but is it a.a.s. much higher? We do not determine $\alpha(P_n)$, but we shall quickly show that the answer is *yes*. Recall that the *core* (or 2-*core*) $\mathrm{core}(G)$ of a graph G is the unique maximal subgraph with minimum degree at least 2. Call $\mathrm{fr}(G) = G \setminus \mathrm{core}(G)$ the *fringe* of G. Take G to be a 4-colourable graph with n vertices. Since $\mathrm{core}(G)$ is also

https://doi.org/10.1017/CBO9781139519793.013 Published online by Cambridge University Press

4-colourable, it has a bipartite induced subgraph B with $|B| \geq \frac{1}{2}(n - |\operatorname{fr}(G)|)$. The vertices of $\operatorname{fr}(G)$ can be added one by one to B so that it remains bipartite, yielding a bipartite graph B' with $|B'| \geq \frac{1}{2}(n-|\operatorname{fr}(G)|)+|\operatorname{fr}(G)| = \frac{1}{2}(n+|\operatorname{fr}(G)|)$. Thus $\alpha(G) \geq \frac{1}{4}(n + |\operatorname{fr}(G)|)$. Applying this to P_n, using the fact shown in [71] that $|\operatorname{fr}(P_n)|/n \to \delta \approx 0.038157$ a.a.s., we conclude that $\alpha(P_n) \geq (0.25 + 0.019)n = 0.269n$ a.a.s. The question remains, *how large is* $\alpha(P_n)$ *a.a.s.*?

7. Other colourings

In this section we hint at the variety of research, even within random graph theory, on other types of graph colouring. We have had to limit ourselves to brief accounts for a small selection of topics, showing the tip of the iceberg. Moreover, we mainly consider just the symmetric case $G_{n,1/2}$, the dense Erdős–Rényi random graph, as in Section 2. Some of the results mentioned below extend to other (sparser) choices of p, to varying extents.

Strengthened colourings

Equitable chromatic number

A proper colouring is *equitable* if the orders of the colour classes differ by at most 1. The celebrated Hajnal–Szemerédi theorem states that every graph G has an equitable proper k-colouring, for each $k \geq \Delta(G) + 1$. The least k for which G has an equitable proper k-colouring is the *equitable chromatic number* $\chi_=(G)$ of G. The property of equitable k-colourability is not monotone in k, but in the light of the above statement the following definition is natural. The least k for which G has an equitable proper k'-colouring for every $k' \geq k$ is the *equitable chromatic threshold* $\chi^*_=(G)$ of G. Note that $\chi(G) \leq \chi_=(G) \leq \chi^*_=(G)$ always.

Krivelevich and Patkós [56] re-purposed the greedy processes, both of Grimmett and McDiarmid and of Bollobás, to prove the following.

Theorem 7.1
A.a.s. $\chi_=(G_{n,1/2}) \leq (1 + o(1))\chi(G_{n,1/2})$
and $\chi^*_=(G_{n,1/2}) \leq (2 + o(1))\chi(G_{n,1/2})$.

They obtained partial results for other values of p and have furthermore conjectured that $\chi^*_=(G_{n,p}) \leq (1 + o(1))\chi(G_{n,p})$ a.a.s. if $np \to \infty$. Further results have been obtained by Rombach and Scott [80].

Choosability

We defined the choosability $\operatorname{ch}(G)$ of a graph G in Section 6 – see also Chapter 6. In a first paper on choosability in 1979, Erdős, Rubin and Taylor conjectured that

https://doi.org/10.1017/CBO9781139519793.013 Published online by Cambridge University Press

$\mathrm{ch}(G_{n,1/2}) = o(n)$ a.a.s. This conjecture was proved by Alon [4] in 1992. Shortly thereafter, Kahn proved the following theorem (see [5]).

Theorem 7.2 *A.a.s.* $\mathrm{ch}(G_{n,1/2}) = (1 + o(1))\chi(G_{n,1/2})$.

Kahn's proof is elegant. The idea is to extract repeatedly (using Theorem 2.4) a large colour class for any colour that appears on more than $n/\log^2 n$ vertices, and at the end to match the remaining vertices with the remaining colours using Hall's theorem. Later, Krivelevich [55] extended Theorem 7.2 to hold for p as low as $n^{-\alpha}$ with $\alpha < \frac{1}{4}$, which was soon after improved to $\alpha < \frac{1}{3}$ in [58]. Alon, Krivelevich and Sudakov [8] conjectured that Theorem 7.2 extends to the case $np \to \infty$. They obtained partial progress on this conjecture, getting the correct value up to a constant multiple; this was proved independently by Vu [86].

Edge-colourings and total colourings

Chromatic index

Recall from Chapter 5 that a *proper edge-colouring* of a graph G is an assignment of colours to the edges so that incident edges receive different colours; the *chromatic index* $\chi'(G)$ is the least number of colours needed for such a colouring. Vizing's theorem states that $\chi'(G)$ is $\Delta(G)$ or $\Delta(G) + 1$; those graphs with the smaller value are of *class 1* while those with the larger are of *class 2*. By noticing that $G_{n,1/2}$ a.a.s. has a unique vertex of maximum degree, Erdős and Wilson [37] in 1977 proved the following.

Theorem 7.3 *A.a.s.* $G_{n,1/2}$ *is of class 1.*

A decade later, Frieze, Jackson, McDiarmid and Reed [42] improved on this by showing that the probability of $G_{n,1/2}$ being of class 2 satisfies, for any fixed $\varepsilon > 0$ and large enough n,

$$n^{-(1/2+\varepsilon)n} \leq \mathbb{P}(G_{n,1/2} \text{ is of class 2}) \leq n^{-(1/8-\varepsilon)n}.$$

They proved the upper bound by analyzing a certain polynomial-time edge-colouring algorithm and bounding its failure probability. The lower bound relied on bounding the probability that $G_{n,1/2}$ is regular.

Total chromatic number

A *total colouring* of a graph G is an assignment of a colour to each vertex and each edge so that no pair of adjacent vertices or adjacent edges receive the same colour and every edge is coloured differently from its endpoints. The *total chromatic number* $\chi''(G)$ is the least number of colours needed for such a colouring. It is obvious that $\chi''(G) \geq \Delta(G) + 1$ for all G. The total colouring conjecture, due independently to Behzad and Vizing, asserts that $\chi''(G) \leq \Delta(G) + 2$ for all G. Molloy and Reed have

https://doi.org/10.1017/CBO9781139519793.013 Published online by Cambridge University Press

shown by probabilistic means that there is a (not incredibly large) constant C for which $\chi''(G) \leq \Delta(G) + C$ for all G (see [75, Chs. 17–18]). Before that, McDiarmid and Reed [73] in 1993 studied $\chi''(G_{n,1/2})$. Partly by adapting the above-mentioned algorithm for edge-colouring and partly by invoking the probabilistic method, they proved the following.

Theorem 7.4

$$\mathbb{P}(\chi''(G_{n,1/2}) > \Delta(G_{n,1/2}) + 1) \leq n^{-(1/8-\varepsilon)n}.$$

$$\mathbb{P}(\chi''(G_{n,1/2}) > \Delta(G_{n,1/2}) + 2) = e^{-\Omega(n^2)}.$$

The second bound in particular implies that $G_{n,1/2}$ satisfies the total colouring conjecture up to a minuscule failure probability.

Generalized colourings

Hereditary properties

A graph property is *hereditary* if it is closed under vertex deletion. (Note, for example, that every monotone graph property is hereditary.) Given some fixed hereditary property $\mathcal{P}$, the (*generalized*) $\mathcal{P}$*-chromatic number* $\chi_{\mathcal{P}}(G)$ of a graph G is the least number of parts in a vertex partition of G for which each part induces a subgraph satisfying $\mathcal{P}$. By taking $\mathcal{P}$ to be the set of all edgeless graphs, we recover the ordinary chromatic number. If $\mathcal{P}$ is the set of all graphs or is finite, then $\chi_{\mathcal{P}}(G)$ is uninteresting; so we assume otherwise, calling such properties *non-trivial*. In 1992 Scheinerman [81] discovered that a non-trivial general statement is possible. In particular, for non-trivial hereditary $\mathcal{P}$, $\chi_{\mathcal{P}}(G_{n,1/2}) = O(n/\log n)$ a.a.s. A few years later, Bollobás and Thomason [19] proved the following tighter result.

Theorem 7.5 *For any non-trivial hereditary property $\mathcal{P}$, there is a constant $c = c_{1/2}(\mathcal{P})$ for which $\chi_{\mathcal{P}}(G_{n,1/2}) = (c + o(1))n/(2\log n)$ a.a.s.*

To prove this, they used a well-known asymptotic characterization of hereditary properties according to a particular extremal parameter. Given non-negative integers a and b, we say that an (a, b)*-colouring* of a graph G is a partition of the vertex-set into a cliques and b stable sets. The *colouring number* $\tau(\mathcal{P})$ of a hereditary property $\mathcal{P}$ is the largest k for which there are a and b with $a + b = k$ for which every (a, b)-colourable graph has property $\mathcal{P}$. It was known that the parameter $\tau(\mathcal{P})$ governs the asymptotic behaviour of the cardinality $|\mathcal{P}_n|$ of graphs on $[n]$ satisfying $\mathcal{P}$. Bollobás and Thomason showed moreover that $\tau(\mathcal{P})$ determines the constant c in Theorem 7.5. In follow-up work [20], they extended this theorem to other fixed values of $p \in (0, 1)$, with $c_p(\mathcal{P})$ in the place of $c_{1/2}(\mathcal{P})$. This problem gave exception to the rule that a result for $G_{n,1/2}$ is easily extendable to the entire dense regime. An application of Szemerédi's regularity lemma was needed to approximate the given hereditary

https://doi.org/10.1017/CBO9781139519793.013 Published online by Cambridge University Press

property structurally. Marchant and Thomason [61] recently improved this structural approximation to shed more light on how to compute specific values of $c_p(\mathcal{P})$.

Improper colourings

Given a non-negative integer t, a *t-improper colouring* of G is a vertex partition in which each part induces a subgraph of maximum degree at most t; the *t-improper chromatic number* $\chi^t(G)$ is the least number of parts needed for such a partition. This type of chromatic parameter can be used to measure the effects of a controlled amount of error in our graph colourings, which can be helpful in certain applications, such as in channel assignment or in distributed computing. Clearly, $t = 0$ corresponds to the ordinary chromatic number. It is known (and fairly easy to see) that

$$\frac{\chi(G)}{t+1} \le \chi^t(G) \le \min\left\{\left\lceil \frac{\Delta(G)+1}{t+1} \right\rceil, \chi(G)\right\}.$$

The authors in [51] considered $\chi^t(G_{n,1/2})$ when t is allowed to increase as a function of n. In particular, the above theorem for hereditary properties does not apply in this situation. If $t = o(\log n)$, then $\chi^t(G_{n,1/2})$ has the same a.a.s. first-order behaviour as $\chi(G_{n,1/2})$, while $\chi^t(G_{n,1/2}) \sim np/t$ a.a.s. if $t = \omega(\log n)$. This implies that the above upper bound on $\chi^t(G)$ for general G is usually the right asymptotic answer for $G_{n,1/2}$. The following is a fuller and more detailed result.

Theorem 7.6 *There exists a function $\kappa = \kappa_{1/2}(\tau)$ that is continuous and strictly increasing for $\tau \in [0,\infty)$, with $\kappa_{1/2}(0) = 2/\ln 2$ and $\kappa_{1/2}(\tau) \sim 2\tau$ as $\tau \to \infty$, for which, if $t(n) = o(n)$, then a.a.s.*

$$\chi^t(G_{n,1/2}) \sim \frac{n}{\kappa_{1/2}(t/\ln n)\ln n}.$$

This result was proved by a careful study of the nearly maximum *t-stable sets* – vertex subsets that induce subgraphs of maximum degree at most t – using tools from large deviations analysis. Such tools were critical in computing both the first and second moments, especially at the threshold $t = \Theta(\log n)$. Other more precise results on both the t-stability and t-improper chromatic numbers of $G_{n,p}$ can be found in [38] and [39].

References

1. D. Achlioptas and E. Friedgut, A sharp threshold for k-colorability, *Random Structures Algorithms* **14** (1999), 63–70.
2. D. Achlioptas and C. Moore, The chromatic number of random regular graphs, *Proc. 7th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems (APPROX) and 8th International Workshop on Randomization and Computation (RANDOM)*, Lecture Notes in Computer Science, Springer (2004), 219–228.
3. D. Achlioptas and A. Naor, The two possible values of the chromatic number of a random graph, *Ann. of Math. (2)* **162** (2005), 1335–1351.

https://doi.org/10.1017/CBO9781139519793.013 Published online by Cambridge University Press

4. N. Alon, Choice numbers of graphs: a probabilistic approach, *Combin. Probab. Comput.* **1** (1992), 107–114.
5. N. Alon, Restricted colorings of graphs, *Surveys in Combinatorics, 1993* (Keele) (ed. K. Walker), *London Math. Soc. Lecture Notes* **187**, Cambridge Univ. Press (1993), 1–33.
6. N. Alon, The chromatic number of random Cayley graphs, *Europ. J. Combin.* **34** (2013), 1232–1243.
7. N. Alon and M. Krivelevich, The concentration of the chromatic number of random graphs, *Combinatorica* **17** (1997), 303–313.
8. N. Alon, M. Krivelevich and B. Sudakov, List coloring of random and pseudo-random graphs, *Combinatorica* **19** (1999), 453–472.
9. N. Alon and J. Spencer, *The Probabilistic Method* (3rd edn.), with an appendix on the life and work of Paul Erdős, Wiley–Interscience Series in Discrete Mathematics and Optimization, 2008.
10. S. Ben-Shimon and M. Krivelevich, Random regular graphs of non-constant degree: concentration of the chromatic number, *Discrete Math.* **309** (2009), 4149–4161.
11. E. A. Bender and E. R. Canfield, The asymptotic number of labeled graphs with given degree sequences, *J. Combin. Theory (A)* **24** (1978), 296–307.
12. E. A. Bender, Z. Gao and N. C. Wormald, The number of labeled 2-connected planar graphs, *Electron. J. Combin.* **9**, Research Paper 43, 13 pp. (electronic), 2002.
13. B. Bollobás, A probabilistic proof of an asymptotic formula for the number of labelled regular graphs, *Europ. J. Combin.* **1** (1980), 311–316.
14. B. Bollobás, The chromatic number of random graphs, *Combinatorica* **8** (1988), 49–55.
15. B. Bollobás, *Random Graphs* (2nd edn.), Cambridge Studies in Advanced Math. **73**, 2001.
16. B. Bollobás, How sharp is the concentration of the chromatic number?, *Combin. Probab. Comput.* **13** (1) (2004), 115–117.
17. B. Bollobás and P. Erdős, Cliques in random graphs, *Math. Proc. Cambridge Philos. Soc.* **80** (1976), 419–427.
18. B. Bollobás and A. Thomason, Threshold functions, *Combinatorica* **7** (1987), 35–38.
19. B. Bollobás and A. Thomason, Generalized chromatic numbers of random graphs. *Random Structures Algorithms* **6** (1995), 353–356.
20. B. Bollobás and A. Thomason, The structure of hereditary properties and colourings of random graphs, *Combinatorica* **20** (2000), 173–202.
21. A. Braunstein, R. Mulet, A. Pagnani, M. Weigt and R. Zecchina, Polynomial iterative algorithms for coloring and analyzing random graphs, *Phys. Rev. E (3)* **68** (2003), 036702, 15.
22. G. Chapuy, É. Fusy, O. Giménez, B. Mohar and M. Noy., Asymptotic enumeration and limit laws for graphs of fixed genus, *J. Combin. Theory (A)* **118** (2011), 748–777.
23. V. Chvátal, Almost all graphs with 1.44*n* edges are 3-colorable, *Random Structures Algorithms* **2** (1991), 11–28.
24. A. Coja-Oghlan, Upper-bounding the *k*-colorability threshold by counting covers, *Electron. J. Combin.* **20** (2013), paper 32, 28.
25. A. Coja-Oghlan, C. Efthymiou and S. Hetterich, On the chromatic number of random regular graphs, ArXiv *e-prints*, Aug. 2013.
26. A. Coja-Oghlan, K. Panagiotou and A. Steger, On the chromatic number of random graphs, *J. Combin. Theory (B)* **98** (2008), 980–993.
27. A. Coja-Oghlan and D. Vilenchik, Chasing the k-colorability threshold, ArXiv *e-prints*, Apr. 2013.
28. C. Cooper, A. Frieze, B. Reed and O. Riordan, Random regular graphs of non-constant degree: independence and chromatic number, *Combin. Probab. Comput.* **11** (2002), 323–341.

https://doi.org/10.1017/CBO9781139519793.013 Published online by Cambridge University Press

29. M. DeVos, K. Kawarabayashi and B. Mohar, Locally planar graphs are 5-choosable, *J. Combin. Theory (B)* **98** (2008), 1215–1232.
30. J. Díaz, A. C. Kaporis, G. D. Kemkes, L. M. Kirousis, X. Pérez-Giménez and N. C. Wormald, On the chromatic number of a random 5-regular graph, *J. Graph Theory* **61** (2009), 157–191.
31. C. Dowden, *Uniform Random Planar Graphs with Degree Constraints*, Ph.D. thesis, University of Oxford, 2008, *http://ora.ox.ac.uk/objects/uuid%3Af8a9afe3-30ad-4672-9a6c-4fb9ac9af041*.
32. C. Dowden, The evolution of uniform random planar graphs, *Electron. J. Combin.* **17** (2010), R7, 20 pp.
33. M. Dyer, A. Frieze and C. Greenhill, On the chromatic number of a random hypergraph, ArXiv *e-prints*, Aug. 2012.
34. P. Erdős, Graph theory and probability, *Canad. J. Math.* **11** (1959), 34–38.
35. P. Erdős and S. Fajtlowicz, On the conjecture of Hajós, *Combinatorica* **1** (1981), 141–143.
36. P. Erdős and A. Rényi, On the evolution of random graphs, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **5** (1960), 17–61.
37. P. Erdős and R. J. Wilson, On the chromatic index of almost all graphs, *J. Combin. Theory (B)* **23** (1977), 255–257.
38. N. Fountoulakis, R. J. Kang and C. McDiarmid, The t-stability number of a random graph, *Electron. J. Combin.* **17** (2010), R59, 29 pp.
39. N. Fountoulakis, R. J. Kang and C. McDiarmid, Largest sparse subgraphs of random graphs, *Europ. J. Combin.* **35** (2014), 232–244.
40. E. Friedgut, Sharp thresholds of graph properties, and the k-SAT problem, with an appendix by Jean Bourgain, *J. Amer. Math. Soc.* **12** (1999), 1017–1054.
41. A. Frieze, On the independence number of random graphs, *Discrete Math.* **81** (1990), 171–175.
42. A. Frieze, B. Jackson, C. McDiarmid and B. Reed, Edge-colouring random graphs, *J. Combin. Theory (B)* **45** (1988), 135–149.
43. A. Frieze and T. Łuczak, On the independence and chromatic numbers of random regular graphs, *J. Combin. Theory (B)* **54** (1992), 123–132.
44. A. Frieze and E. Vigoda, A survey on the use of Markov chains to randomly sample colourings, *Combinatorics, Complexity, and Chance* (eds. G. Grimmett and C. McDiarmid), Oxford Lecture Ser. Math. Appl. **34**, Oxford Univ. Press (2007), 53–71.
45. O. Giménez and M. Noy, Asymptotic enumeration and limit laws of planar graphs, *J. Amer. Math. Soc.* **22** (2009), 309–329.
46. B. Green, On the chromatic number of random Cayley graphs, ArXiv *e-prints*, Aug. 2013.
47. B. Green and R. Morris, Counting sets with small sumset and applications, ArXiv *e-prints*, May 2013.
48. G. Grimmett and C. McDiarmid, On colouring random graphs, *Math. Proc. Cambridge Philos. Soc.* **77** (1975), 313–324.
49. S. Janson, T. Łuczak and A. Ruciński, *Random Graphs*, Wiley–Interscience Series in Discrete Math. and Optimization, 2000.
50. M. Kang and T. Łuczak, Two critical periods in the evolution of random planar graphs, *Trans. Amer. Math. Soc.* **364** (2012), 4239–4265.
51. R. J. Kang and C. McDiarmid, The t-improper chromatic number of random graphs, *Combin. Probab. Comput.* **19** (2010), 87–98.
52. R. M. Karp, The probabilistic analysis of some combinatorial search algorithms, *Algorithms and Complexity* (Carnegie-Mellon Univ., 1976) (ed. J. F. Traub), Academic Press (1976), 1–19.
53. G. Kemkes, X. Pérez-Giménez and N. C. Wormald, On the chromatic number of random d-regular graphs, *Adv. Math.* **223** (2010), 300–328.

https://doi.org/10.1017/CBO9781139519793.013 Published online by Cambridge University Press

54. J. H. Kim and V. H. Vu, Sandwiching random graphs: universality between random graph models, *Adv. Math.* **188** (2004), 444–469.
55. M. Krivelevich, The choice number of dense random graphs, *Combin. Probab. Comput.* **9** (2000), 19–26.
56. M. Krivelevich and B. Patkós, Equitable coloring of random graphs, *Random Structures Algorithms* **35** (2009), 83–99.
57. M. Krivelevich, B. Sudakov, V. H. Vu and N. C. Wormald, Random regular graphs of high degree, *Random Structures Algorithms* **18** (2001), 346–363.
58. M. Krivelevich, B. Sudakov, V. H. Vu and N. C. Wormald, On the probability of independent sets in random graphs, *Random Structures Algorithms* **22** (2003), 1–14.
59. T. Łuczak, The chromatic number of random graphs, *Combinatorica* **11** (1991), 45–54.
60. T. Łuczak, A note on the sharp concentration of the chromatic number of random graphs, *Combinatorica* **11** (1991), 295–297.
61. E. Marchant and A. Thomason, The structure of hereditary properties and 2-coloured multigraphs, *Combinatorica* **31** (2011), 85–93.
62. D. W. Matula, On the complete subgraphs of a random graph, *Proc. Second Chapel Hill Conf. on Combinatorial Mathematics and its Applications* (Univ. of North Carolina, 1970), Univ. of North Carolina, Chapel Hill (1970), 356–369.
63. D. W. Matula, The employee party problem, *Notices Amer. Math. Soc.* **19** (1972), A–382.
64. D. W. Matula, Expose-and-merge exploration and the chromatic number of a random graph, *Combinatorica* **7** (1987), 275–284.
65. D. W. Matula and L. Kučera, An expose-and-merge algorithm and the chromatic number of a random graph, *Random Graphs '87* (Poznań, 1987) (eds. M. Karoński, J. Jaworski and A. Ruciński), Wiley (1990), 175–187.
66. C. McDiarmid, Colouring random graphs badly, *Graph Theory and Combinatorics* (Open Univ., Milton Keynes, 1978) (ed. R. J. Wilson), Research Notes in Math. **34**, Pitman (1979), 76–86.
67. C. McDiarmid, On the chromatic number of random graphs, *Random Structures Algorithms* **1** (1990), 435–442.
68. C. McDiarmid, Random channel assignment in the plane, *Random Structures Algorithms* **22** (2003), 187–212.
69. C. McDiarmid, Random graphs on surfaces, *J. Combin. Theory (B)* **98** (2008), 778–797.
70. C. McDiarmid, Random graphs from a minor-closed class, *Combin. Probab. Comput.* **18** (2009), 583–599.
71. C. McDiarmid, Random graphs from a weighted minor-closed class, *Electron. J. Combin.* **20** (2013), R52, 39 pp.
72. C. McDiarmid and T. Müller, On the chromatic number of random geometric graphs, *Combinatorica* **31** (2011), 423–488.
73. C. McDiarmid and B. Reed, On total colourings of graphs, *J. Combin. Theory (B)* **57** (1993), 122–130.
74. C. McDiarmid, A. Steger and D. Welsh, Random planar graphs, *J. Combin. Theory (B)* **93** (2005), 187–205.
75. M. Molloy and B. Reed, *Graph Colouring and the Probabilistic Method*, *Algorithms and Combinatorics* **23**, Springer-Verlag, 2002.
76. R. Mulet, A. Pagnani, M. Weigt and R. Zecchina, Coloring random graphs, *Phys. Rev. Lett.* **89** (2002), 268701, 4.
77. T. Müller, Two-point concentration in random geometric graphs, *Combinatorica* **28** (2008), 529–545.

https://doi.org/10.1017/CBO9781139519793.013 Published online by Cambridge University Press

78. K. Panagiotou and A. Steger, A note on the chromatic number of a dense random graph, *Discrete Math.* **309** (2009), 3420–3423.
79. M. Penrose, *Random Geometric Graphs*, Oxford Studies in Probability **5**, Oxford University Press, 2003.
80. M. P. Rombach and A. Scott, Equitable colourings of random graphs, in preparation, 2014.
81. E. R. Scheinerman, Generalized chromatic numbers of random graphs, *SIAM J. Discrete Math.* **5** (1992), 74–80.
82. E. Shamir and J. Spencer, Sharp concentration of the chromatic number on random graphs $G_{n,p}$, *Combinatorica* **7** (1987), 121–129.
83. L. Shi and N. C. Wormald, Colouring random 4-regular graphs, *Combin. Probab. Comput.* **16** (2007), 309–344.
84. L. Shi and N. C. Wormald, Colouring random regular graphs, *Combin. Probab. Comput.* **16** (2007), 459–494.
85. C. Thomassen, Five-coloring maps on surfaces, *J. Combin. Theory (B)* **59** (1993), 89–105.
86. V. H. Vu, On some simple degree conditions that guarantee the upper bound on the chromatic (choice) number of random graphs, *J. Graph Theory* **31** (1999), 201–226.
87. N. C. Wormald, Differential equations for random processes and random graphs, *Ann. Appl. Probab.* **5** (1995), 1217–1235.
88. N. C. Wormald, Models of random regular graphs, *Surveys in Combinatorics, 1999* (Canterbury) (eds. J. D. Lamb and D. A. Preece), *London Math. Soc. Lecture Notes* **267**, Cambridge Univ. Press (1999), 239–298.

https://doi.org/10.1017/CBO9781139519793.013 Published online by Cambridge University Press