On the Chromatic Number in the Stochastic Block Model

Mikhail Isaev*
School of Mathematics
Monash University
3800 Clayton, Australia
mikhail.isaev@monash.edu

Mihyun Kang†
Institute of Discrete Mathematics
Graz University of Technology
8010 Graz, Austria
kang@math.tugraz.at

Submitted: Sep 14, 2021; Accepted: Apr 12, 2023; Published: Jun 30, 2023

Abstract

We prove a generalisation of Bollobás' classical result on the asymptotics of the chromatic number of the binomial random graph to the stochastic block model. In addition, by allowing the number of blocks to grow, we determine the chromatic number in the Chung-Lu model. Our approach is based on the estimates for the weighted independence number, where weights are specifically designed to encapsulate inhomogeneities of the random graph.

Mathematics Subject Classifications: 05C80, 05C15

1 Introduction

The chromatic number $\chi(G)$ of a graph G, denoted by $\chi(G)$, is the smallest number of colours needed for the assignment of colours to the vertices of G so that no two adjacent vertices have the same colour. Understanding properties of the distribution of $\chi(\boldsymbol{G})$ for random $\boldsymbol{G}$ is one of the most prominent problems in the random graph theory since the seminal paper [12] by Erdős and Rényi.

The binomial random graph $\boldsymbol{G}(n,p)$ is the most studied in the literature. Recall that $\boldsymbol{G}(n,p)$ is a graph on vertex set $[n] := \{1,2,\ldots,n\}$ and each pair of distinct vertices is connected by an edge independently of each other with probability p. A long line of research led to many breakthrough results on the asymptotic behaviour and concentration of $\chi(\boldsymbol{G}(n,p))$; see [2, 3, 7, 8, 10, 16, 17, 22, 25, 26, 28, 30, 33] — this list is far from being

*Supported by Australian Research Council Discovery Project DP190100977 and by Australian Research Council Discovery Early Career Researcher Award DE200101045
†Supported by Austrian Science Fund (FWF): I3747

exhaustive. In particular, it is well known that if $p = p(n) \in [0,1]$ is such that $np \to \infty$ as $n \to \infty$ and $p \leqslant 1 - \varepsilon$ for some fixed $\varepsilon > 0$, then, whp (meaning *with probability tending to one*) as $n \to \infty$,

$$\chi(\boldsymbol{G}(n,p)) = (1+o(1))\frac{n\log\left(\frac{1}{1-p}\right)}{2\log(pn)}. \tag{1}$$

Formally, "$\boldsymbol{X}(n) = (1+o(1))\boldsymbol{Y}(n)$ holds whp as $n \to \infty$" means that, for any fixed $\epsilon > 0$, the probability of the event that $(1-\epsilon)\boldsymbol{Y}(n) \leqslant \boldsymbol{X}(n) \leqslant (1+\epsilon)\boldsymbol{Y}(n)$ tends to 1 as $n \to \infty$. Throughout the paper, we use log to denote the natural logarithm.

Our paper focuses on generalising formula (1) to a random graph $\boldsymbol{G}$ from *the stochastic block model*, in which all vertices are distributed between several different blocks and the probabilities of adjacencies of vertices depend only on the block they belong to; see Section 2 for formal definitions. The chromatic number in this random graph model was recently studied by Martinsson et al. [27]. Under the condition that the number of blocks is fixed and all probabilities are constants from $(0,1)$, namely, they are all independent of the number of vertices n, Martinsson et al. proved that, whp as $n \to \infty$,

$$\chi(\boldsymbol{G}) = (1+o(1))\frac{n}{c^* \log n},$$

where constant c^* is the solution of a certain convex optimisation problem, which depends only on the matrix of probabilities and the proportions for the distribution of n vertices between the blocks.

In this paper, we extend the above result by Martinsson et al. [27] in two directions:

(1) the edge probabilities can be functions of n (in particular, vanishing or tending 1),

(2) the number of blocks can grow as a function of n.

We defer the exact statement of our main result (Theorem 5) to Section 2 in order to obviate introducing the technical notations in the introduction. In the rest of this section, we discuss several consequences of Theorem 5, which are interesting of their own.

1.1 A very dense binomial random graph

The classical binomial random graph $\boldsymbol{G}(n,p)$ can be considered as a random graph from the stochastic block model with a single block. Even in this case, our main result (Theorem 5) implies new information on the chromatic number of a very dense binomial random graph when $p = p(n) \to 1$ as $n \to \infty$ which was not treated in the literature. Namely, as a straightforward application of Theorem 5, we obtain the following result.

Theorem 1. *If $p = p(n) \in [0,1]$ such that $p \to 1$ and $1 - p = n^{o(1)}$, then (1) holds whp.*

We believe that $n^{o(1)}$ in Theorem 1 can not be improved. For example, if $p_1 = 1 - \frac{1}{n\log n}$ then whp $\boldsymbol{G}(n,p_1)$ has a clique of size $(1+o(1))n$ since its complement contains $o(n)$ edges. Thus, whp as $n \to \infty$

$$\chi(\boldsymbol{G}(n,p_1)) = (1+o(1))n.$$

On the other hand, if $p_2 = 1 - \frac{\log^2 n}{n}$ then whp the complement of $\boldsymbol{G}(n, p_2)$ contains a perfect matching as shown by Erdős and Rényi [13]. Thus, whp as $n \to \infty$

$$\chi(\boldsymbol{G}(n, p_2)) \leqslant (1 + o(1))\frac{n}{2}.$$

Note that formula (1) is not valid for $p = p_1$, but it might still be true for $p = p_2$, because

$$\frac{\log \frac{1}{1-p_1}}{\log(p_1 n)} = 1 + o(1) \qquad \text{and} \qquad \frac{\log \frac{1}{1-p_2}}{\log(p_2 n)} = 1 + o(1).$$

More generally, for the case when $p = 1 - n^{O(1)}$, we conjecture the following.

Conjecture 2. Let $r \geqslant 2$ be a fixed integer and $p = p(n) \in [0, 1]$ be such that

$$n^{-\frac{2}{r+1}} \gg 1 - p \gg n^{-\frac{2}{r}}.$$

Then, $\chi(\boldsymbol{G}(n, p)) = (1 + o(1))\frac{n}{r}$ whp as $n \to \infty$.

As observed above, $\boldsymbol{G}(n, p)$ can be coloured in $\frac{n}{2}$ colours if its complement has a perfect matching. In fact, to achieve a colouring with at most $(1 + o(1))\frac{n}{2}$ colours, it is sufficient that the complement of $\boldsymbol{G}(n, p)$ contains an almost perfect matching covering $n - o(n)$ vertices. Similarly, for any fixed integer $r \geqslant 2$, in order to show that $\chi(\boldsymbol{G}(n, p)) \leqslant (1 + o(1))\frac{n}{r}$, it is sufficient to find an almost perfect K_r-matching in the complement of $\boldsymbol{G}(n, p)$. (Throughout the paper, K_r denotes the complete graph with vertex set $[r]$ or the clique of size r.) For an arbitrary graph G, the thresholds for the existence of perfect G-matchings and almost perfect G-matchings was studied by Ruciński [31] and by Johansson, Kahn, and Vu [15]. In particular, [31, Theorem 4] establishes the existence of an almost perfect K_r-matching if $n(1 - p)^{r/2} \gg 1$ which implies the upper bound of Conjecture 2. However, the lower bound for $\chi(\boldsymbol{G}(n, p))$ does not follow from the known results on G-matchings since an optimal colouring migh have colour classes of different sizes.

Conjecture 2 was recently confirmed by Surya and Warnke; see [32, Theorem 13].

1.2 Percolations on blow-up graphs

Given a graph $G = (V(G), E(G))$ and $p \in (0, 1)$, the percolated random graph G_p, which is also known as a random subgraph of G, is generated from G by keeping each edge in $E(G)$ independently with probability p. In particular, if $G = K_n$, then G_p is equivalent to the binomial random graph $\boldsymbol{G}(n, p)$. In this case, formula (1) can read as follows: whp

$$\chi(G_p) = (1 + o(1))\frac{\log(\frac{1}{1-p})}{2\log(pn)}\,\chi(G), \qquad \text{as } n = |V(G)| \to \infty. \tag{2}$$

In this paper we show that (2) holds when G is a *blow-up graph* $G_H(\boldsymbol{n})$ constructed as follows. Given a graph H on vertex set $[k]$ and a vector $\boldsymbol{n} = (n_1, \ldots, n_k)^T \in \mathbb{N}^k$, we denote by $G_H(\boldsymbol{n})$ the graph obtained from H by replacing each vertex $i \in [k]$ with K_{n_i}. An edge between any two vertices from different cliques appears in $G_H(\boldsymbol{n})$ if the corresponding edge is present in H. One can consider the blow-up graph $G_H(\boldsymbol{n})$ as a special case of a "random" graph from the stochastic block model by setting all probabilities 1 or 0 according to the adjacency matrix of the graph H.

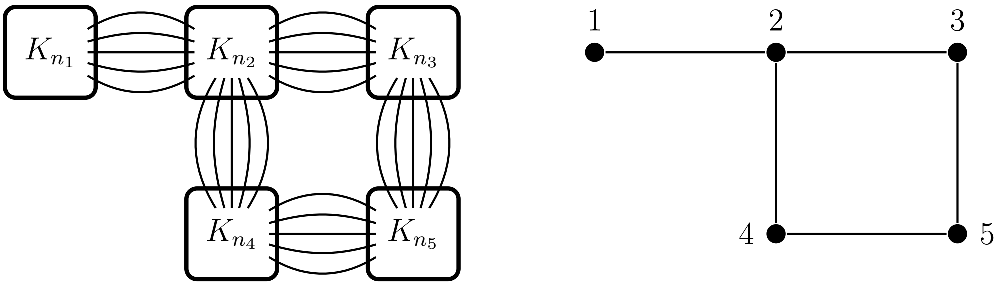

Figure 1: A blow-up graph $G_H(\boldsymbol{n})$ (left) for a graph H on 5 vertices (right).

Everywhere in this paper the norm notation $\|\cdot\|$ stands for the 1-norm:

$$\|\boldsymbol{n}\| = n_1 + \cdots + n_k.$$

Theorem 3. *Let $\epsilon \in (0, \frac{1}{4})$ be fixed and H be a graph with vertex set $[k]$. Assume $\boldsymbol{n} = \boldsymbol{n}(n) \in \mathbb{N}^k$ and $p = p(n) \in (0,1)$ are such that as $n \to \infty$,*

$$\|\boldsymbol{n}\| \to \infty, \qquad p \geqslant \|\boldsymbol{n}\|^{-\frac{1}{4}+\epsilon}, \qquad 1-p = \|\boldsymbol{n}\|^{-o(1)}.$$

Then, (2) *with $G = G_H(\boldsymbol{n})$ holds whp.*

We prove Theorem 3 in Section 3.3. Note that Theorem 3 with $k = 1$ and $n_1 = n$ (and thus $G_H(\boldsymbol{n}) = K_n$) recovers Theorem 1.

Determining the chromatic number of a random subgraph G_p for a general graph G is a much harder problem; see, for example, [4–6,29,34]. In particular, Bukh asks [6] whether for any graph G, there exists a positive constant c such that $\mathbb{E}\chi(G_{1/2}) \geqslant \frac{c}{\log(\chi(G))}\,\chi(G)$. Using standard concentration results, Bukh's question for blow-up graphs is equivalent to that whp

$$\chi(G_{1/2}) \geqslant \frac{c}{\log(|V(G)|)}\,\chi(G).$$

Theorem 3 establishes this bound for blow-up graphs. It would be interesting to find other classes of graphs that satisfy (2) (or at least its lower bound).

1.3 Chung-Lu model

As mentioned, our main result (Theorem 5) allows the number of blocks to grow. Thus, one can study $\chi(\boldsymbol{G})$ for general inhomogenous random graphs $\boldsymbol{G}$ using approximations by the stochastic block model. To demonstrate the idea, we consider the following two random graph models. Given $\boldsymbol{u} = (u_1, \ldots, u_n)^T \in [0,1]^n$ and $p \in [0,1]$, a random graph

$\boldsymbol{G}_p^\times \sim \mathcal{G}^\times(\boldsymbol{u}, p)$ has vertex set $[n]$ and edges ij are generated independently of each other with probabilities

$$p_{ij}^\times = p\, u_i u_j \qquad i, j \in [n].$$

Similarly, given $\boldsymbol{u} \in [0,1]^n$ and $p \in [0, \frac{1}{2}]$, a random graph $\boldsymbol{G}_p^+ \sim \mathcal{G}^+(\boldsymbol{u}, p)$ has vertex set $[n]$ and edges ij are generated independently of each other with probabilities

$$p_{ij}^+ = p\,(u_i + u_j) \qquad i, j \in [n].$$

The model $\mathcal{G}^\times(\boldsymbol{u}, p)$ is known as the Chung-Lu random graph model and it is of central importance in the network analysis; for more extensive background, see, for example, [9] and references therein. For decreasing $p = p(n)$, the model $\mathcal{G}^+(\boldsymbol{u}, p)$ is asymptotically equivalent to the complement of the Chung-Lu model.

Theorem 4. *Let $\epsilon > 0$ be fixed and $p = p(n)$ be such $1 \gg p \geqslant n^{-\frac{1}{4}+\epsilon}$ as $n \to \infty$. Then, whp uniformly over $\boldsymbol{u} \in [0,1]^n$ satisfying $\sum_{i\in[n]} u_i = \Omega(n)$, the following hold:*

(a) $\chi(\boldsymbol{G}_p^\times) = (1 + o(1)) \dfrac{p}{2\log(pn)} \max_{U \subseteq [n]} \dfrac{1}{|U|} \left(\sum_{i \in U} u_i \right)^2$, *where* $\boldsymbol{G}_p^\times \sim \mathcal{G}^\times(\boldsymbol{u}, p)$;

(b) $\chi(\boldsymbol{G}_p^+) = (1 + o(1)) \dfrac{p}{\log(pn)} \sum_{i \in [n]} u_i$, *where* $\boldsymbol{G}_p^+ \sim \mathcal{G}^+(\boldsymbol{u}, p)$.

We prove Theorem 4 in Section 3.4. Theorem 4 applies to the case when a constant fraction of expected degrees of the random graphs $\boldsymbol{G}_p^\times$ and $\boldsymbol{G}_p^+$ are within a multiplicative constant of the maximum expected degree. We believe that the formulas of Theorem 4 can be extended to allow a larger variation of components of $\boldsymbol{u}$ covering, for example, power-law degree sequences.

2 Stochastic block model

Before stating our main result on the chromatic number of a random graph from the stochastic block model, we first define the stochastic block model formally. For a positive integer k, a vector $\boldsymbol{n} = (n_1, \ldots, n_k)^T \in \mathbb{N}^k$, and a $k \times k$ symmetric matrix $P = (p_{ij})_{i,j\in[k]}$ with $p_{ij} \in [0,1]$, a random graph $\boldsymbol{G}$ from the stochastic block model $\mathcal{G}(\boldsymbol{n}, P)$, denoted by $\boldsymbol{G} \sim \mathcal{G}(\boldsymbol{n}, P)$, is constructed as follows:

- the vertex set $V(\boldsymbol{G})$ is partitioned into k disjoint blocks $B_1, \ldots, B_k$ of sizes $|B_i| = n_i$ for $i \in [k]$ (and we write $V(\boldsymbol{G}) = B_1 \cup \cdots \cup B_k$);
- each pair $\{u, v\}$ of distinct vertices $u, v \in V(\boldsymbol{G})$ is included in the edge set $E(\boldsymbol{G})$, independently of one another, with the probability

  $$p(u, v) := p_{ij},$$

  where $i = i(u) \in [k]$ and $j = j(v) \in [k]$ are such $u \in B_i$ and $v \in B_j$.

Throughout the paper, for all asymptotic notation, we implicitly consider sequences of vectors $\boldsymbol{n} = \boldsymbol{n}(n) \in \mathbb{N}^k$ and $k \times k$ symmetric matrices $P = P(n)$, where

$$k = k(n), \qquad \boldsymbol{n}(n) = (n_1(n), \dots n_k(n))^T, \qquad P = \Big(p_{ij}(n)\Big)_{i,j\in[k]}.$$

Our bounds (including whp results) hold uniformly over all sequences $\boldsymbol{n}(n)$ and $P(n)$, where $n \to \infty$, satisfying stated assumptions where the implicit functions like in $o(\cdot)$ depend on n only. Apart from the standard Landau notation $o(\cdot)$ and $O(\cdot)$, we also use the notation $a_n = \omega(b_n)$ or $a_n = \Omega(b_n)$ if $a_n > 0$ always and $b_n = o(a_n)$ or $b_n = O(a_n)$, respectively. We write $a_n = \Theta(b_n)$ if $a_n = O(b_n)$ and $b_n = O(a_n)$. If both a_n and b_n are positive sequences, we also write $a_n \ll b_n$ if $a_n = o(b_n)$, and $a_n \gg b_n$ if $a_n = \omega(b_n)$. For example, $k = \|\boldsymbol{n}\|^{o(1)}$ means that $\frac{\log k(n)}{\log \|\boldsymbol{n}(n)\|} \to 0$.

In the following, we always assume that $p_{ij} = p_{ji}$ and $0 \leqslant p_{ij} < 1$ for all $i, j \in [k]$. Define the $k \times k$ symmetric matrix $Q = Q(P)$ by

$$Q := (q_{ij})_{i,j\in[k]}, \qquad \text{where} \;\; q_{ij} := \log\left(\tfrac{1}{1-p_{ij}}\right). \tag{3}$$

Let $\mathbb{R}_+ := [0, +\infty)$ and, for $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^k$, we denote

$$\boldsymbol{y} \preceq \boldsymbol{x} \text{ whenever} \boldsymbol{x} - \boldsymbol{y} \in \mathbb{R}_+^k.$$

Let $w(\cdot, Q) : \mathbb{R}_+^k \to \mathbb{R}_+$ be defined by

$$w(\boldsymbol{x}, Q) := \max_{\boldsymbol{0}\preceq\boldsymbol{y}\preceq\boldsymbol{x}} \frac{\boldsymbol{y}^T Q\, \boldsymbol{y}}{\|\boldsymbol{y}\|}, \qquad \boldsymbol{x} \in \mathbb{R}_+^k, \tag{4}$$

where $\boldsymbol{0} = (0, \dots, 0)^T \in \mathbb{R}^k$ and $\|\boldsymbol{y}\| := |y_1| + \cdots + |y_k|$. In (4), we take $\frac{\boldsymbol{y}^T Q\, \boldsymbol{y}}{\|\boldsymbol{y}\|}$ to be zero for $\boldsymbol{y} = \boldsymbol{0}$, so it is a continuous function of $\boldsymbol{y}$, which achieves its the maximal value on the compact set $\{\boldsymbol{y} \in \mathbb{R}_+^k : \boldsymbol{y} \preceq \boldsymbol{x}\}$. In fact, it is always achieved at a corner, where $y_i \in \{0, x_i\}$ for all $i \in [k]$; seeTheorem 10(b).

The quantity $w(\boldsymbol{x}, Q)$ is closely related to the minimum number of colours required to properly colour an inhomogeneous graph with "balanced" colour classes. To illustrate it, let us consider a random graph $\boldsymbol{G} \sim \mathcal{G}(\boldsymbol{n}, P)$, where $\boldsymbol{n} = (nx_1, nx_2, \dots, nx_k)^T = n\boldsymbol{x}$ and number of blocks k and all probabilities $p_{ij} \in (0, 1)$ are fixed. In order to determine the size of the largest "balanced" independent set, we will present here some rough first moment calculations, while the full details are given in Section 4 and Section 6.

The expected number of collections of k disjoint sets S_i, each of which takes sx_i vertices from each block B_i, such that $\cup_{i\in[k]} S_i$ is an independent set in $\boldsymbol{G}$ (see Figure 2) is given by

$$\prod_{i\in[k]} \frac{\binom{nx_i}{sx_i}}{(1-p_{ii})^{sx_i}} \cdot \prod_{i,j\in[k]} (1-p_{ij})^{s^2 x_i x_j} = \exp\left(-\frac{s^2}{2}\boldsymbol{x}^T Q \boldsymbol{x} + O(s\|\boldsymbol{x}\|)\right) \left(\frac{en}{s}\right)^{s\|\boldsymbol{x}\|}, \tag{5}$$

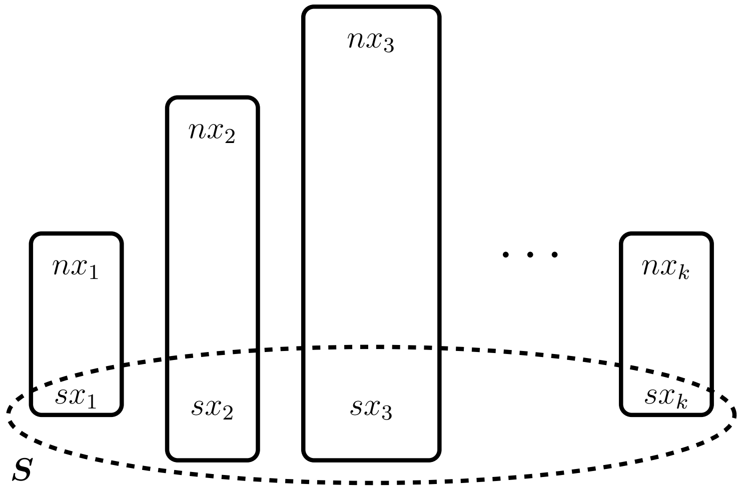

Figure 2: A "balanced" set $S = \bigcup_{i\in[k]} S_i$ in $\boldsymbol{G} \sim \mathcal{G}(\boldsymbol{n}, P)$, where $|S_i| = sx_i$ and $n_i = nx_i$.

where the RHS is derived via Stirling's formula for any slowly growing $s = s(n) \ll \sqrt{n}$. The first moment threshold corresponds to

$$\exp\left(-\frac{1}{2}s\boldsymbol{x}^T Q\boldsymbol{x}\right)\left(\frac{en}{s}\right)^{\|\boldsymbol{x}\|} = 1,$$

which gives

$$s \approx 2\log n \cdot \frac{\|\boldsymbol{x}\|}{\boldsymbol{x}^T Q\boldsymbol{x}}.$$

However, this might be significantly above the existence threshold due to the fact that our random graph model is inhomogeneous. In particular, the appearance of "balanced" independent sets in $\boldsymbol{G}$ implies the existence of a "balanced" independent set (with the same size proportion s/n) in its subgraph $\boldsymbol{G}' \sim \mathcal{G}(\boldsymbol{n}', P)$ where $\boldsymbol{n}' = n\boldsymbol{y}$ for any $\boldsymbol{0} \preceq \boldsymbol{y} \preceq \boldsymbol{x}$. Repeating the arguments of (5) for such $\boldsymbol{G}'$, we conclude that whp s can not exceed

$$2\log n \cdot \min_{\boldsymbol{0}\preceq\boldsymbol{y}\preceq\boldsymbol{x}} \frac{\|\boldsymbol{y}\|}{\boldsymbol{y}^T\, Q\, \boldsymbol{y}} = 2\log n \cdot \frac{1}{w(\boldsymbol{x}, Q)}.$$

In Section 4, we show that it is indeed the existence threshold (for a more general setting that allows vanishing probabilities); see Theorem 19.

Define $w_*(\cdot, Q) : \mathbb{R}_+^k \to \mathbb{R}_+$ by

$$w_*(\boldsymbol{x}, Q) := \inf_{\mathcal{S}\in\mathcal{F}(\boldsymbol{x})} \sum_{\boldsymbol{y}\in\mathcal{S}} w(\boldsymbol{y}, Q), \qquad \boldsymbol{x} \in \mathbb{R}_+^k, \tag{6}$$

where $\mathcal{F}(\boldsymbol{x})$ consists of finite systems $\mathcal{S}$ of vectors from $\mathbb{R}_+^k$ such that $\sum_{\boldsymbol{y}\in\mathcal{S}} \boldsymbol{y} = \boldsymbol{x}$. In fact, the infimum of $\sum_{\boldsymbol{y}\in\mathcal{S}} w(\boldsymbol{y})$ in (6) is always achieved by a system $\mathcal{S} \in \mathcal{F}(\boldsymbol{x})$ consisting of at most k vectors; see Theorem 10(f).

Similarly to $w(\boldsymbol{x}, Q)$, the quantity $w_*(\boldsymbol{x}, Q)$ has a combinatorial meaning as follows. Consider again a random graph $\boldsymbol{G} \sim \mathcal{G}(\boldsymbol{n}, P)$, where $\boldsymbol{n} = (nx_1, nx_2, \ldots, nx_n)^T = n\boldsymbol{x}$

and all probabilities $p_{ij} \in (0,1)$ are fixed. Then whp the *minimum* number of colours required to properly colour $\boldsymbol{G}$ utilising at most k different types of independent sets is asymptotically equal to

$$\frac{n}{2\log n} w_*(\boldsymbol{x}, Q).$$

The next theorem shows that such colourings are asymptotically optimal, that is, no more than k different types are required to determine $\chi(\boldsymbol{G})$ (for a more general setting that allows vanishing probabilities). Let

$$q^* := \max_{i \in [k]} q_{ii} \qquad \text{and} \qquad \hat{q}(\boldsymbol{x}) := \frac{\sum_{i \in [k]} x_i q_{ii}}{\|\boldsymbol{x}\|}, \quad \boldsymbol{x} \neq \mathbf{0}. \tag{7}$$

For convenience, we also set $\hat{q}(\mathbf{0}) := q^*$.

Theorem 5. *Let $\sigma \in [0, \sigma_0]$ for some fixed $0 < \sigma_0 < \frac{1}{4}$ and let $P = (p_{ij})_{i,j \in [k]}$ be such that $p_{ij} = p_{ji}$ and $0 \leqslant p_{ij} < 1$ for all $i, j \in [k]$. Let $Q := (q_{ij})_{i,j \in [k]}$ where $q_{ij} := \log\left(\frac{1}{1-p_{ij}}\right)$. Let q^*, $\hat{q}(\cdot)$, and $w_*(\cdot, Q)$ be defined by (7) and (6). Assume that the following asymptotics hold:*

$$\|\boldsymbol{n}\| \to \infty, \qquad k = \|\boldsymbol{n}\|^{o(1)}, \qquad q^* = \|\boldsymbol{n}\|^{-\sigma + o(1)}, \qquad \hat{q}(\boldsymbol{n}) = \|\boldsymbol{n}\|^{-\sigma + o(1)}. \tag{8}$$

Assume also that

$$\left(1 + \frac{1}{q^*}\right) \max_{i,j \in [k]} q_{ij} \ll \log \|\boldsymbol{n}\| \tag{9}$$

and

$$w_*(\boldsymbol{n}, Q) \gg k \hat{q}(\boldsymbol{n}) q^* \frac{\|\boldsymbol{n}\|}{\log \|\boldsymbol{n}\|}. \tag{10}$$

Then, whp

$$\chi(\boldsymbol{G}) = (1 + o(1)) \frac{w_*(\boldsymbol{n}, Q)}{2(1-\sigma) \log \|\boldsymbol{n}\|}, \qquad \textit{where } \boldsymbol{G} \sim \mathcal{G}(\boldsymbol{n}, P).$$

Remark 6. In Theorem 5, the parameter $\sigma \in [0, \sigma_0]$ governs the density of $\boldsymbol{G}$. It is convenient for our examples to have it not fixed, but treat σ as a bounded parameter appearing in the formula for the chromatic number. We believe that the condition $\sigma_0 < \frac{1}{4}$ is an artefact of our proof techniques. Similarly to the dense case in [14, Section 7.4] and also to [27], we rely on Janson's inequality to find sufficiently large independent sets inside any subset of remaining vertices. Generalisations of the techniques used by Łuczak [25] should extend Theorem 5 to any $\sigma_0 < 1$.

Remark 7. Informally, the assumptions of (8) say that the number of blocks in $\mathcal{G}(\boldsymbol{n}, P)$ is not too big (sublinear in $\|\boldsymbol{n}\|$) and the maximum edge probability within a block deviates not too much (also by a sublinear in $\|\boldsymbol{n}\|$ factor) from the average probability within blocks. Next, the behaviour of edge probabilities between blocks is limited by assumption (9): they can vary much more significantly than the diagonal probabilities, but we prohibit them to converge to 1 too quickly. Note that we allow some edge probabilities to be small and even 0, but the upper bounds on the maximal probabilities are essential as demonstrated in Section 1.1. Finally, (10) is a technical assumption that is usually not very hard to verify. In particular, it follows from a stronger but more explicit condition $(kq^*)^2 \ll \hat{q}(\boldsymbol{n}) \log \|\boldsymbol{n}\|$; see the lower bound of Theorem 10(d).

2.1 Proof of Theorem 5

In this section, we provide the proof of Theorem 5 based on two explicit probability estimates for $\chi(\boldsymbol{G})$ to satisfy the upper and the lower bound stated below.

Theorem 8. *Let $\boldsymbol{G} \sim \mathcal{G}(\boldsymbol{n}, P)$, where $P = (p_{ij})_{i,j\in[k]}$ is such that $p_{ij} = p_{ji}$ and $0 \leqslant p_{ij} < 1$ for all $i, j \in [k]$. Assume $\|\boldsymbol{n}\| \to \infty$ and (9) holds. Then, for any $\varepsilon > 0$,*

$$\Pr\left(\chi(\boldsymbol{G}) < (1-\varepsilon)\frac{w_*(\boldsymbol{n})}{2\log(q^*\|\boldsymbol{n}\|)}\right) \leqslant \exp\left(-\Omega\left(\frac{\log(q^*\|\boldsymbol{n}\|)}{\max_{i,j\in[k]} q_{ij}}\right)\right). \tag{11}$$

We prove Theorem 8 in Section 4.1. This lower tail bound follows from the existence of large weighted independent sets, similarly to arguments of Bollobás [7] and Łuczak [25]. Comparing to the assumptions of Theorem 5, we note that Theorem 8 also applies for sparser graphs with $q^* < \|\boldsymbol{n}\|^{-\frac{1}{4}}$,because it does not require assumption (8) to hold.

Theorem 9. *Let $\boldsymbol{G} \sim \mathcal{G}(\boldsymbol{n}, P)$, where $P = (p_{ij})_{i,j\in[k]}$ is such that $p_{ij} = p_{ji}$ and $0 \leqslant p_{ij} < 1$ for all $i, j \in [k]$. Let $\sigma \in [0, \sigma_0]$ for some fixed $0 < \sigma_0 < \frac{1}{4}$. Assume that (8) and (10) hold. Then, for any $\varepsilon > 0$,*

$$\Pr\left(\chi(\boldsymbol{G}) > (1+\varepsilon)\,\frac{w_*(\boldsymbol{n})}{2\log(q^*\|\boldsymbol{n}\|)}\right) \leqslant \exp\left(-\|\boldsymbol{n}\|^{2-4\sigma+o(1)}\right). \tag{12}$$

We prove Theorem 9 in Section 6.3, using/extending some results and standard arguments on the chromatic numer of the classical binomial random graph. Comparing to the assumptions of Theorem 5, we note that Theorem 9 allows more variation in the off-diagonal probabilities p_{ij},because it does not require assumption (9) to hold.

Now, we are ready to prove Theorem 5.

Proof of Theorem 5. Using the assumption $q^* = \|\boldsymbol{n}\|^{-\sigma+o(1)}$ from (8), we observe that

$$\log(q^*\|\boldsymbol{n}\|) = (1+o(1))(1-\sigma)\log\|\boldsymbol{n}\|.$$

Note that all assumptions of Theorems 8 and 9 hold as they appear as the assumptions of Theorem 5. Also, the quantities on the right hand sides of (11) and (12) satisfy

$$\frac{\log(q^*\|\boldsymbol{n}\|)}{\max_{i,j\in[k]} q_{ij}} \to \infty \qquad \text{and} \qquad \|\boldsymbol{n}\|^{2-4\sigma} \to \infty.$$

Thus, applying Theorems 8 and 9, we get that, for any fixed $\varepsilon > 0$, whp

$$(1-\varepsilon)\frac{w_*(\boldsymbol{n})}{2(1-\sigma)\log\|\boldsymbol{n}\|} \leqslant \chi(\boldsymbol{G}) \leqslant (1+\varepsilon)\frac{w_*(\boldsymbol{n})}{2(1-\sigma)\log\|\boldsymbol{n}\|}.$$

This completes the proof. □

2.2 Properties of $\boldsymbol{w}$ and $\boldsymbol{w}_*$

In this section, we collect some facts about the quantities $w(\cdot) = w(\cdot, Q)$ and $w_*(\cdot) = w_*(\cdot, Q)$, defined by (4) and (6) for a general matrix Q. These properties are helpful for applications of Theorem 5 and will also be repeatedly used in the proofs.

Theorem 10. *Let $Q = (q_{ij})_{i,j\in[k]}$ be a symmetric $k \times k$ matrix with non-negative entries. Let q^* and $\hat{q}(\cdot)$ be defined according (7). Then, the following hold for any $\boldsymbol{x} = (x_1, \dots, x_k)^T \in \mathbb{R}_+^k$.*

(a) [Scaling and monotonicity]. *If $\boldsymbol{x}' \in \mathbb{R}_+^k$ and $\boldsymbol{x}' \preceq s\boldsymbol{x}$ for some $s > 0$, then $w(\boldsymbol{x}') \leqslant sw(\boldsymbol{x})$ and $w_*(\boldsymbol{x}') \leqslant sw_*(\boldsymbol{x})$. In particular, $w(s\boldsymbol{x}) = sw(\boldsymbol{x})$ and $w_*(s\boldsymbol{x}) = sw_*(\boldsymbol{x})$.*

(b) [Corner maximiser]. *There is $\boldsymbol{z} = (z_1, \dots, z_k)^T$ with $z_i \in \{0, x_i\}$ for all $i \in [k]$ such that*
$$\frac{\boldsymbol{z}^T Q\, \boldsymbol{z}}{\|\boldsymbol{z}\|} = w(\boldsymbol{x}) := \max_{\boldsymbol{0}\preceq\boldsymbol{y}\preceq\boldsymbol{x}} \frac{\boldsymbol{y}^T Q\, \boldsymbol{y}}{\|\boldsymbol{y}\|}.$$

(c) [Pseudodefinite property]. *If $\boldsymbol{y}^T Q\boldsymbol{y} \geqslant 0$ for all $\boldsymbol{y} \in \mathbb{R}^k$ with $\sum_{i\in[k]} y_i = 0$, then*
$$w(\boldsymbol{x}) = w_*(\boldsymbol{x}) := \inf_{\mathcal{S}\in\mathcal{F}(\boldsymbol{x})} \sum_{\boldsymbol{y}\in\mathcal{S}} w(\boldsymbol{y}).$$

(d) [Upper and lower bounds]. *We have*
$$q^*\|\boldsymbol{x}\| \geqslant \hat{q}(\boldsymbol{x})\|\boldsymbol{x}\| \geqslant w_*(\boldsymbol{x}) \geqslant \frac{(\hat{q}(\boldsymbol{x}))^2}{\sum_{i\in[k]} q_{ii}}\|\boldsymbol{x}\| \geqslant \frac{(\hat{q}(\boldsymbol{x}))^2}{kq^*}\|\boldsymbol{x}\|,$$
where the lower bounds for $w_*(\boldsymbol{x})$ hold under the additional condition that $q^* > 0$.

(e) [Triangle inequality]. *For any $\boldsymbol{x}' \in \mathbb{R}_+^k$, we have $w_*(\boldsymbol{x}) + w_*(\boldsymbol{x}') \geqslant w_*(\boldsymbol{x} + \boldsymbol{x}')$.*

(f) [Minimal system of k vectors]. *There exists a system of vectors $(\boldsymbol{x}^{(t)})_{t\in[k]}$, each from $\mathbb{R}_+^k$, such that $\sum_{t\in[k]} \boldsymbol{x}^{(t)} = \boldsymbol{x}$ and $\sum_{t\in[k]} w(\boldsymbol{x}^{(t)}) = w_*(\boldsymbol{x})$.*

(g) [Near-optimal integer system]. *If $\boldsymbol{x} \in \mathbb{N}^k$ then there exists a system of vectors $(\boldsymbol{x}^{(t)})_{t\in[k]}$, each from $\mathbb{N}^k$, such that $\sum_{t\in[k]} \boldsymbol{x}^{(t)} = \boldsymbol{x}$ and $\sum_{t\in[k]} w(\boldsymbol{x}^{(t)}) \leqslant w_*(\boldsymbol{x}) + k^2q^*$.*

The proof of Theorem 10 is technical and not very insightful, but for completeness it is provided at the end of the paper in Section 7.

Remark 11. An interesting question not covered in this paper is how to compute or at least approximate $w_*(\cdot)$ efficiently. We give some examples in Section 3, but the question remains open in general. We believe that the optimization problems of finding $w_*(\cdot)$ and $w(\cdot)$ can be efficiently solved by fast converging iterative methods such as gradient descent and analogues of the simplex method.

2.3 Structure of the rest of the paper

Section 3 covers applications of our main result, Theorem 5. We consider first the case of two blocks in detail and then prove Theorems 3 and 4. In addition, we study the unions of two independent random graphs from the stochastic block model. In Section 4, we introduce the weighted independence number and prove the lower tail probability estimate of Theorem 8. Sections 5 and 6 are devoted to the upper tail probability estimate of Theorem 9. In Secton 5 we derive some preliminary estimates based on idea of separately colouring the blocks of $\mathcal{G}(\boldsymbol{n}, P)$. In Section 6, we derive an asymptotically optimal bound on the chromatic number using the estimates for the existence of large weighted independent sets given in Section 4.2. Finally, we prove Theorem 10 in Section 7.

3 Applications of the main theorem

In this section, we discuss some applications of Theorem 5. Specifically, we consider the case of two blocks (Section 3.1), the union of two independent random graphs from the stochastic block model (Section 13), percolations on a blow-up graph (Section 3.3), and Chung-Lu model and its complement (Section 3.4).

3.1 Two blocks

Consider the random graph $\boldsymbol{G} \sim \mathcal{G}(\boldsymbol{n}, P)$ with two blocks, where

$$k = 2, \qquad \boldsymbol{n} = (n_1, n_2)^T \in \mathbb{N}^2, \qquad P = \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} \quad \text{with } p_{12} = p_{21}.$$

Let B_1 and B_2 denote the two blocks of $\boldsymbol{G}$, i.e., a partition of the vertex set $V(\boldsymbol{G})$, and let $\boldsymbol{G}_1 := \boldsymbol{G}[B_1] \sim \mathcal{G}(n_1, p_{11})$ and $\boldsymbol{G}_2 := \boldsymbol{G}[B_2] \sim \mathcal{G}(n_2, p_{22})$ denote the induced subgraphs of $\boldsymbol{G} \sim \mathcal{G}(\boldsymbol{n}, P)$ on B_1 and B_2, respectively. Since B_1 and B_2 are disjoint, we have

$$\max\{\chi(\boldsymbol{G}_1), \chi(\boldsymbol{G}_2)\} \leqslant \chi(\boldsymbol{G}) \leqslant \chi(\boldsymbol{G}_1) + \chi(\boldsymbol{G}_2). \tag{13}$$

For fixed $p_{11}, p_{22} \in (0, 1)$, Martinsson et al. observed in [27, Section 4.1] that there are two threshold values $\underline{p}$ and $\overline{p}$ such that whp $\chi(\boldsymbol{G})$ is asymptotically equal to the lower bound of (13) if $p_{12} \leqslant \underline{p}$, but it is equal to the upper bound of (13) if $p_{12} \geqslant \overline{p}$. Using Theorem 5, we extend this result to non-fixed $p_{11} = p_{11}(n)$ or $p_{22} = p_{22}(n)$ that are allowed to vanish asymptotically. We also obtain the asymptotic formula for $\chi(\boldsymbol{G})$ when $\underline{p} \leqslant p_{12} \leqslant \overline{p}$.

To state our results, define

$$\begin{aligned} \overline{p} &= \overline{p}(p_{11}, p_{22}) := 1 - (1 - p_{11})^{\frac{1}{2}} (1 - p_{22})^{\frac{1}{2}}, \\ \underline{p} &= \underline{p}(\boldsymbol{n}, p_{11}, p_{22}) := 1 - \min\left\{ (1 - p_{11})^{\frac{1}{2}} \cdot (1 - p_{22})^{-\frac{n_2}{2n_1}}, (1 - p_{22})^{\frac{1}{2}} \cdot (1 - p_{11})^{-\frac{n_1}{2n_2}} \right\}. \end{aligned}$$

Obviously, $1 \geqslant \overline{p} \geqslant \underline{p}$ since $p_{11}, p_{22} \in (0,1)$. Observe also $\underline{p} \geqslant 0$ since

$$\min\left\{\frac{(1-p_{11})^{\frac{1}{2}}}{(1-p_{22})^{\frac{n_2}{2n_1}}}, \frac{(1-p_{22})^{\frac{1}{2}}}{(1-p_{11})^{\frac{n_1}{2n_2}}}\right\} \leqslant \left(\frac{(1-p_{11})^{\frac{1}{2}}}{(1-p_{22})^{\frac{n_2}{2n_1}}}\right)^{\frac{n_1}{n_1+n_2}} \left(\frac{(1-p_{22})^{\frac{1}{2}}}{(1-p_{11})^{\frac{n_1}{2n_2}}}\right)^{\frac{n_2}{n_1+n_2}} = 1.$$

Recall from (3) that $Q = Q(P) = (q_{ij})_{i,j\in\{1,2\}}$ is defined by $q_{ij} := \log(\frac{1}{1-p_{ij}})$.

Theorem 12. *Let $\sigma \in [0, \sigma_0]$ for some fixed $0 < \sigma_0 < \frac{1}{4}$. Assume that*

$$\|\boldsymbol{n}\| \to \infty, \qquad q_{11} = \|\boldsymbol{n}\|^{-\sigma+o(1)}, \qquad q_{22} = \|\boldsymbol{n}\|^{-\sigma+o(1)}, \qquad \frac{q_{11}^2}{q_{22}} + \frac{q_{22}^2}{q_{11}} \ll \log \|\boldsymbol{n}\|.$$

Then the following hold whp.

(i) If $\underline{p} \leqslant p_{12} \leqslant \overline{p}$ then

$$\chi(\boldsymbol{G}) = (1+o(1))\frac{\boldsymbol{n}^T Q \boldsymbol{n}}{2(1-\sigma)\|\boldsymbol{n}\| \log \|\boldsymbol{n}\|}.$$

(ii) If $\overline{p} \leqslant p_{12} \leqslant 1$ then

$$\chi(\boldsymbol{G}) = (1+o(1))\left(\chi(\boldsymbol{G}_1) + \chi(\boldsymbol{G}_2)\right) = (1+o(1))\frac{n_1 q_{11} + n_2 q_{22}}{2(1-\sigma)\log\|\boldsymbol{n}\|}.$$

(iii) If $0 \leqslant p_{12} \leqslant \underline{p}$ then

$$\chi(\boldsymbol{G}) = (1+o(1))\max\{\chi(\boldsymbol{G}_1), \chi(\boldsymbol{G}_2)\} = (1+o(1))\frac{\max\{n_1 q_{11}, n_2 q_{22}\}}{2(1-\sigma)\log\|\boldsymbol{n}\|}.$$

Proof. Let q^*, $\hat{q}(\cdot)$, $w(\cdot, Q)$ and $w_*(\cdot, Q)$ be defined by (7), (4), and (6). We will first check that the assumptions of Theorem 5 are satisfied in part (i). To this end, note that (8) are given in Theorem 12 and observe that

$$p_{12} \leqslant \overline{p} \quad \Longleftrightarrow \quad q_{12} \leqslant \tfrac{1}{2}q_{11} + \tfrac{1}{2}q_{22}.$$

In particular, we get that $q_{12} \leqslant q^*$, thus

$$\left(1 + \frac{1}{q^*}\right)\max_{i,j\in\{1,2\}} q_{ij} \leqslant q^* + 1 \leqslant \frac{(q^*)^2}{\hat{q}(\boldsymbol{n})} + 1 \leqslant \frac{q_{11}^2}{q_{22}} + \frac{q_{22}^2}{q_{11}} + 1 \ll \log\|\boldsymbol{n}\|.$$

Using alsothe bounds of Theorem 10(d), we find that

$$w_*(\boldsymbol{n}, Q) \geqslant \frac{(\hat{q}(\boldsymbol{n}))^2}{2q^*}\|\boldsymbol{n}\| \gg \frac{\hat{q}(\boldsymbol{n}) q^* \|\boldsymbol{n}\|}{\log\|\boldsymbol{n}\|}.$$

This establishes (9) and (10).

To prove part (i) by applying Theorem 5, it remains to show that if $\underline{p} \leqslant p_{12} \leqslant \overline{p}$ then

$$w_*(\boldsymbol{n}, Q) = \frac{\boldsymbol{n}^T Q \boldsymbol{n}}{\|\boldsymbol{n}\|}.$$

The inequality $q_{12} \leqslant \frac{1}{2}q_{11} + \frac{1}{2}q_{22}$ is also equivalent to $\boldsymbol{y}^T Q \boldsymbol{y} \geqslant 0$ for any $\boldsymbol{y} \in \mathbb{R}^2$ with $y_{11}+y_{22} = 0$ (clearly, one only needs to consider $\boldsymbol{y} = (1, -1)^T$). Usingthe corner maximiser and the pseudodefinite properties in Theorem 10(b,c), we get that

$$w_*(\boldsymbol{n}, Q) = w(\boldsymbol{n}, Q) = \max\left\{n_1 q_{11}, n_2 q_{22}, \frac{\boldsymbol{n}^T Q \boldsymbol{n}}{\|\boldsymbol{n}\|}\right\}.$$

Observe that

$$n_1 q_{11} \leqslant \frac{\boldsymbol{n}^T Q \boldsymbol{n}}{\|\boldsymbol{n}\|} = \frac{n_1^2 q_{11} + 2 n_1 n_2 q_{12} + n_2^2 q_{22}}{n_1 + n_2}$$

holds whenever $q_{12} \geqslant \frac{1}{2}q_{11} - \frac{n_2}{2n_1} q_{22}$. Similarly, $n_2 q_{22} \leqslant \frac{\boldsymbol{n}^T Q \boldsymbol{n}}{\|\boldsymbol{n}\|}$ holds whenever $q_{12} \geqslant \frac{1}{2}q_{22} - \frac{n_1}{n_2} q_{11}$. Next, we recall the second assumption of part (i) that $p_{12} \geqslant \underline{p}$, which is equivalent to

$$q_{12} \geqslant \max\left\{\frac{1}{2}q_{11} - \frac{n_2}{2n_1} q_{22}, \frac{1}{2}q_{22} - \frac{n_1}{2n_2} q_{11}\right\}.$$

Thus, we conclude that

$$w_*(\boldsymbol{n}, Q) = \max\left\{n_1 q_{11}, n_2 q_{22}, \frac{\boldsymbol{n}^T Q \boldsymbol{n}}{\|\boldsymbol{n}\|}\right\} = \frac{\boldsymbol{n}^T Q \boldsymbol{n}}{\|\boldsymbol{n}\|},$$

which completes the proof of (i).

For part (ii), we consider the random graph $\overline{\boldsymbol{G}} \sim \mathcal{G}(n, \overline{P})$ such that $\overline{\boldsymbol{G}} \subset \boldsymbol{G}$, where the diagonal entries of $\overline{P}$ are the same as of P while the off-digonal entries of $\overline{P}$ equal $\overline{p}$. Using (13), we get that

$$\chi(\overline{\boldsymbol{G}}) \leqslant \chi(\boldsymbol{G}) \leqslant \chi(\boldsymbol{G}_1) + \chi(\boldsymbol{G}_2).$$

Thus, it is sufficient to show that whp

$$\chi(\overline{\boldsymbol{G}}) = (1 + o(1)) \frac{n_1 q_{11} + n_2 q_{22}}{2(1-\sigma) \log \|\boldsymbol{n}\|}, \tag{14}$$

$$\chi(\boldsymbol{G}_1) + \chi(\boldsymbol{G}_2) = (1 + o(1)) \frac{n_1 q_{11} + n_2 q_{22}}{2(1-\sigma) \log \|\boldsymbol{n}\|}. \tag{15}$$

Applying part (i) to $\overline{\boldsymbol{G}}$, we get that

$$\chi(\overline{\boldsymbol{G}}) = (1 + o(1)) \frac{\boldsymbol{n}^T \overline{Q} \boldsymbol{n}}{2(1-\sigma)\|\boldsymbol{n}\| \log \|\boldsymbol{n}\|},$$

where $\overline{Q}$ is the matrix corresponding to $\overline{P}$. Note that

$$\frac{\boldsymbol{n}^T \overline{Q} \boldsymbol{n}}{\|\boldsymbol{n}\|} = \frac{n_1^2 q_{11} + n_1 n_2 (q_{11} + q_{22}) + n_2^2 q_{22}}{n_1 + n_2} = n_1 q_{11} + n_2 q_{22}.$$

Thus, (14) holds.

Next, observe that (15) is implied by Theorem 1 if $n_1 = \|\boldsymbol{n}\|^{1+o(1)}$ and $n_2 = \|\boldsymbol{n}\|^{1+o(1)}$. Otherwise, if one of the parts is very small, say $\boldsymbol{G}_1$, then we have

$$\|\boldsymbol{n}\| = (1+o(1))n_2 \quad \text{and} \quad n_1 q_{11} + n_2 q_{22} = (1+o(1))n_2 q_{22}.$$

Applying Theorem 1 to $\boldsymbol{G}_2$, we get whp

$$\chi(\boldsymbol{G}_2) = (1+o(1))\frac{n_2 q_{22}}{2\log(p_{22}n_2)} = (1+o(1))\frac{n_1 q_{11} + n_2 q_{22}}{2(1-\sigma)\log\|\boldsymbol{n}\|}.$$

Let $n_1' = \frac{n_2 q_{22}}{q_{11}\log\|\boldsymbol{n}\|}$. By the assumptions, $n_1' = \|\boldsymbol{n}\|^{1+o(1)} \gg n_1$. Using the embedding $\boldsymbol{G}_1 \subset \boldsymbol{G}(n_1', p_{11})$, we estimate

$$\chi(\boldsymbol{G}_1) \leqslant \chi(\boldsymbol{G}(n_1', p_{11})) = (1+o(1))\frac{n_1' q_{11}}{2\log(p_{11}n_1')} = o(1)\frac{n_1 q_{11} + n_2 q_{22}}{\log\|\boldsymbol{n}\|}.$$

The above two bounds for $\chi(\boldsymbol{G}_1)$ and $\chi(\boldsymbol{G}_2)$ prove (15), completing the proof of part (ii).

Part (iii) is proved in a similar way to part (ii). □

3.2 Union of two independent random graphs

Consider two independent binomial random graphs $\boldsymbol{G}_1 = \boldsymbol{G}(n, p_1)$ and $\boldsymbol{G}_2 = \boldsymbol{G}(n, p_2)$ on the same vertex set $[n]$, where p_1, p_2 are some constants from $(0, 1)$. It is easy to show that their union $\boldsymbol{G}_1 \cup \boldsymbol{G}_2$ is also a binomial random graph $\boldsymbol{G}(n, p)$, where $1-p = (1-p_1)(1-p_2)$. This is equivalent to

$$\log\left(\tfrac{1}{1-p}\right) = \log\left(\tfrac{1}{1-p_1}\right) + \log\left(\tfrac{1}{1-p_2}\right). \tag{16}$$

Then, by formula (1), we get that whp

$$\chi(\boldsymbol{G}_1 \cup \boldsymbol{G}_2) = (1+o(1))\left(\chi(\boldsymbol{G}_1) + \chi(\boldsymbol{G}_2)\right). \tag{17}$$

That is, the chromatic number of the union of two independent random graphs is whp asymptotically equal to the sum of the chromatic numbers of the two binomial random graphs. In this section, we prove a generalisation of this observation to the stochastic block modelbased on Theorem 5. Apart from the assumptions of Theorem 5, we also insist that both random graph models satisfy the pseudodefinite property of Theorem 10(c).

Theorem 13. *Let $\boldsymbol{G}_1 \sim \mathcal{G}(\boldsymbol{n}, P_1)$ and $\boldsymbol{G}_2 \sim \mathcal{G}(\boldsymbol{n}, P_2)$ be independent random graphs from the stochastic block models, where P_1 and P_2 satisfy the assumptions of Theorem 5 (with the same σ). Assume also that $\boldsymbol{y}^T Q_1 \boldsymbol{y} \geqslant 0$ and $\boldsymbol{y}^T Q_2 \boldsymbol{y} \geqslant 0$ for all $\boldsymbol{y} \in \mathbb{R}^k$ with $y_1 + \ldots + y_k = 0$, where $Q_1 = Q(P_1)$ and $Q_2 = Q(P_2)$ are defined by (3). Then, whp*

$$\chi(\boldsymbol{G}_1 \cup \boldsymbol{G}_2) \leqslant (1+o(1))\left(\chi(\boldsymbol{G}_1) + \chi(\boldsymbol{G}_2)\right). \tag{18}$$

In addition, if $w(\boldsymbol{n}, Q_1) = \frac{\boldsymbol{n}^T Q_1 \boldsymbol{n}}{\|\boldsymbol{n}\|}$ and $w(\boldsymbol{n}, Q_2) = \frac{\boldsymbol{n}^T Q_2 \boldsymbol{n}}{\|\boldsymbol{n}\|}$ then (17) holds whp.

Proof. Applying Theorem 5, we find that whp

$$\chi(\boldsymbol{G}_1) = (1+o(1))\frac{w_*(\boldsymbol{n}, Q_1)}{2(1-\sigma)\log\|\boldsymbol{n}\|}, \qquad \chi(\boldsymbol{G}_2) = (1+o(1))\frac{w_*(\boldsymbol{n}, Q_2)}{2(1-\sigma)\log\|\boldsymbol{n}\|}. \tag{19}$$

Observe that the union $\boldsymbol{G}_1 \cup \boldsymbol{G}_2$ also belongs to the stochastic block model $\mathcal{G}(\boldsymbol{n}, P)$with the entries of P defined similarly to (16). Observe that

$$Q = Q_1 + Q_2,$$

where $Q = Q(P)$ is defined by (3). In particular, we get that $w_*(\boldsymbol{n}, Q) \geqslant w_*(\boldsymbol{n}, Q_1)$ and $w_*(\boldsymbol{n}, Q_2)$. It is straightforward to check that P and Q satisfy the assumptions of Theorem 5 (with the same σ). Thus, we get whp

$$\chi(\boldsymbol{G}_1 \cup \boldsymbol{G}_2) = (1+o(1))\frac{w_*(\boldsymbol{n}, Q_1+Q_2)}{2(1-\sigma)\log\|\boldsymbol{n}\|}. \tag{20}$$

Next, by the pseudodefinite property in Theorem 10(c), we get that

$$w_*(\boldsymbol{n}, Q_1) = w(\boldsymbol{n}, Q_1), \qquad w_*(\boldsymbol{n}, Q_2) = w(\boldsymbol{n}, Q_2).$$

Note also $\boldsymbol{y}^T Q \boldsymbol{y} = \boldsymbol{y}^T Q_1 \boldsymbol{y} + \boldsymbol{y}^T Q_2 \boldsymbol{y} \geqslant 0$ for all $\boldsymbol{y} \in \mathbb{R}^n$ with $y_1 + \ldots + y_n = 0$. Using Theorem 10(c) again, we find that

$$\begin{aligned} w_*(\boldsymbol{n}, Q_1+Q_2) &= w(\boldsymbol{n}, Q_1+Q_2) \\ &= \max_{\boldsymbol{0} \preceq \boldsymbol{y} \preceq \boldsymbol{n}} \frac{\boldsymbol{y}^T(Q_1+Q_2)\boldsymbol{y}}{\|\boldsymbol{y}\|} \leqslant \max_{\boldsymbol{0} \preceq \boldsymbol{y} \preceq \boldsymbol{n}} \frac{\boldsymbol{y}^T Q_1 \boldsymbol{y}}{\|\boldsymbol{y}\|} + \max_{\boldsymbol{0} \preceq \boldsymbol{y} \preceq \boldsymbol{n}} \frac{\boldsymbol{y}^T Q_2 \boldsymbol{y}}{\|\boldsymbol{y}\|} \\ &= w(\boldsymbol{n}, Q_1) + w(\boldsymbol{n}, Q_2) = w_*(\boldsymbol{n}, Q_1) + w_*(\boldsymbol{n}, Q_2). \end{aligned}$$

Combining the above, we prove (18).

Now assume that $w(\boldsymbol{n}, Q_1) = \frac{\boldsymbol{n}^T Q_1 \boldsymbol{n}}{\|\boldsymbol{n}\|}$ and $w(\boldsymbol{n}, Q_2) = \frac{\boldsymbol{n}^T Q_2 \boldsymbol{n}}{\|\boldsymbol{n}\|}$. To establish (17), we will show that

$$w_*(\boldsymbol{n}, Q_1+Q_2) = w_*(\boldsymbol{n}, Q_1) + w_*(\boldsymbol{n}, Q_2).$$

Then, the result would follow by (19) and (20). We already proved that $w_*(\boldsymbol{n}, Q_1) = w(\boldsymbol{n}, Q_1)$, $w_*(\boldsymbol{n}, Q_2) = w(\boldsymbol{n}, Q_2)$, $w_*(\boldsymbol{n}, Q_1+Q_2) = w(\boldsymbol{n}, Q_1+Q_2)$, and $w(\boldsymbol{n}, Q_1+Q_2) \leqslant w(\boldsymbol{n}, Q_1) + w(\boldsymbol{n}, Q_2)$. Thus, it remains to prove that

$$w(\boldsymbol{n}, Q_1+Q_2) \geqslant w(\boldsymbol{n}, Q_1) + w(\boldsymbol{n}, Q_2). \tag{21}$$

Note that

$$\frac{\boldsymbol{n}^T Q_1 \boldsymbol{n}}{\|\boldsymbol{n}\|} + \frac{\boldsymbol{n}^T Q_2 \boldsymbol{n}}{\|\boldsymbol{n}\|} = \frac{\boldsymbol{n}^T(Q_1+Q_2)\boldsymbol{n}}{\|\boldsymbol{n}\|} \leqslant \max_{\boldsymbol{0} \preceq \boldsymbol{y} \preceq \boldsymbol{n}} \frac{\boldsymbol{y}^T(Q_1+Q_2)\boldsymbol{y}}{\|\boldsymbol{y}\|} = w(\boldsymbol{n}, Q_1+Q_2).$$

Recalling the assumptions that $w(\boldsymbol{n}, Q_1) = \frac{\boldsymbol{n}^T Q_1 \boldsymbol{n}}{\|\boldsymbol{n}\|}$ and $w(\boldsymbol{n}, Q_2) = \frac{\boldsymbol{n}^T Q_2 \boldsymbol{n}}{\|\boldsymbol{n}\|}$, we derive (21), thus completing the proof. □

3.3 Percolations on blow-up graphs: proof of Theorem 3

In order to prove Theorem 3 by applying Theorem 5, we need the following auxiliary result on the chromatic number of any deterministic graph which can be found as a solution of a discrete optimisation problem similar to (6).

For a graph G, let $\mathrm{mad}(G)$ denote *the maximum average degree* over all subgraphs of G.

Lemma 14. *For any graph G, we have*

$$\chi(G) = \min_{\mathcal{U}} \sum\nolimits_{S\in\mathcal{U}} (1 + \mathrm{mad}(G[S])),$$

where the minimum is over all partitions $\mathcal{U}$ of the vertex set $V(G)$ and $G[S]$ denotes the induced subgraph of G.

Proof. It is a standard fact from the graph theory that

$$\chi(G) \leqslant 1 + \max_{U\subseteq V(G)} \delta_G(U), \tag{22}$$

where $\delta_G(U)$ is the minimum degree in the induced graph $G[U]$. The proof of (22) is by a straightforward induction on $|V(G)|$; see, for example, [14, Lemma 7.12].

Clearly, we have that

$$\max_{U\subseteq V(G)} \delta_G(U) \leqslant \mathrm{mad}(G),$$

which together with (22) implies that, for any $S \in V(G)$,

$$\chi(G[S]) \leqslant 1 + \mathrm{mad}(G[S]).$$

Thus, for any partition $\mathcal{U}$ of $V(G)$, we get

$$\chi(G) \leqslant \sum\nolimits_{S\in\mathcal{U}} \chi(G[S]) \leqslant \sum\nolimits_{S\in\mathcal{U}} (1 + \mathrm{mad}(G[S]))$$

by colouring all parts of $\mathcal{U}$ in different colours.

On the other hand, for the partition $\mathcal{U}$ of $V(G)$ corresponding to the colour classes of an optimal colouring of G, we observe that $\mathrm{mad}(G[S]) = 0$ for all $S \in \mathcal{U}$. Thus, we get that

$$\sum\nolimits_{S\in\mathcal{U}} (1 + \mathrm{mad}(G[S])) = \chi(G).$$

This completes the proof. □

We proceed to the proof of Theorem 3. The three key proof ingredients are the following:

- the asymptotic formula for the chromatic number of the stochcastic block model in terms of $w_*(\cdot)$ given in Theorem 5;

- Lemma 14 that expresses the chromatic number of an arbitrary graph in terms of the maximum average degree that is similar to the underlying optimisation problem for $w_*(\cdot)$;
- the existence of a small near-optimal integer system given by Theorem 10(g) that approximates $w_*(\cdot)$.

Proof of Theorem 3. Let $\epsilon \in (0, \frac{1}{4})$ and H be a graph on vertex set $[k]$. Assume that $p = p(n)$ satisfies the conditions in Theorem 3. Let A denote the adjacency matrix of H and I be the $k \times k$ identity matrix.

First, we note that the percolated random graph G_p, where $G = G_H(\boldsymbol{n})$, is distributed according to $\mathcal{G}(\boldsymbol{n}, P)$ with $P = p(I + A)$. Let $Q = Q(P)$ be defined according to (3). We apply Theorem 5 with $\sigma := -\frac{\log p}{\log \|\boldsymbol{n}\|} \leqslant \sigma_0 := \frac{1}{4} - \epsilon$. All assumptions of Theorem 5 are straightforward to check. Since $(1-\sigma)\log\|\boldsymbol{n}\| = \log(p\|\boldsymbol{n}\|)$ by definition of σ, Theorem 5 implies that

$$\chi(G_p) = (1 + o(1))\frac{w_*(\boldsymbol{n}, Q)}{2(1-\sigma)\log\|\boldsymbol{n}\|} = (1 + o(1))\frac{w_*(\boldsymbol{n}, Q)}{2\log(p\|\boldsymbol{n}\|)}.$$

Note that all elements of matrix Q are $\log(\frac{1}{1-p})$ or 0. More precisely, $Q = \log(\frac{1}{1-p})\tilde{Q}$, where $\tilde{Q} := I + A$. By the scaling property in Theorem 10(a), we have

$$w_*(\boldsymbol{n}, Q) = \log(\tfrac{1}{1-p})\ w_*(\boldsymbol{n}, \tilde{Q}).$$

Thus, to prove (2), it remains to show that

$$\chi(G) = (1 + o(1))w_*(\boldsymbol{n}, \tilde{Q}) \tag{23}$$

To show (23), we employ Lemma 14. To this end, for any $S \subseteq V(G)$, we define

$$\boldsymbol{b}(S) = (b_1(S), \ldots, b_k(S))^T \in \mathbb{N}^k \quad \text{with} \quad b_i(S) := |S \cap B_i| \quad \text{for } i \in [k], \tag{24}$$

and observe that, for any $U \subseteq S$

$$\begin{aligned} \boldsymbol{b}(U)^T\tilde{Q}\boldsymbol{b}(U) &= \sum_{i\in[k]} b_i(U)^2 + 2\sum_{ij\in H} b_i(U)b_j(U) \\ &= \|\boldsymbol{b}(U)\| + 2\left(\sum_{i\in[k]} \tfrac{b_i(U)(b_i(U)-1)}{2} + \sum_{ij\in H} b_i(U)b_j(U)\right) \\ &= |U| + 2|E(G[U])|. \end{aligned} \tag{25}$$

Using (25) and the corner maximiser property in Theorem 10(b), we find that

$$w(\boldsymbol{b}(S), \tilde{Q}) = \max_{U\subseteq S} \frac{\boldsymbol{b}(U)^T\tilde{Q}\ \boldsymbol{b}(U)}{\|\boldsymbol{b}(U)\|} = \max_{U\subseteq S}\frac{|U| + 2|E(G[U])|}{|U|} = 1 + \mathrm{mad}(G[S]).$$

By the definition of $w_*(\boldsymbol{n},\tilde{Q})$, we obtain that

$$w_*(\boldsymbol{n},\tilde{Q}) \leqslant \min_{\mathcal{U}} \sum\nolimits_{S\in\mathcal{U}} (1+\mathrm{mad}(G[S])), \tag{26}$$

where the minimum is over all partitions $\mathcal{U}$ of the vertex set $V(G)$. On the other hand, due to the near-optimal integer system of Theorem 10(g), there exists a partition $\mathcal{U}^*$ of $V(G)$ such that

$$\min_{\mathcal{U}} \sum\nolimits_{S\in\mathcal{U}} (1+\mathrm{mad}(G[S])) \leqslant \sum\nolimits_{S\in\mathcal{U}^*} (1+\mathrm{mad}(G[S])) \leqslant w_*(\boldsymbol{n},\tilde{Q}) + k^2\tilde{q}^*.$$

This together with (26) gives

$$w_*(\boldsymbol{n},\tilde{Q}) \leqslant \min_{\mathcal{U}} \sum\nolimits_{S\in\mathcal{U}} (1+\mathrm{mad}(G[S])) \leqslant w_*(\boldsymbol{n},\tilde{Q}) + k^2\tilde{q}^*.$$

Note that the bounds of Theorem 10(d) imply $w_*(\boldsymbol{n},\tilde{Q}) \to \infty$. Because $\tilde{q}^* = 1$ and $k = |V(H)|$ is a fixed constant, using Lemma 14, we get that

$$\chi(G) = \min_{\mathcal{U}} \sum\nolimits_{S\in\mathcal{U}} (1+\mathrm{mad}(G[S])) = w_*(\boldsymbol{n},\tilde{Q}) + O(1),$$

which implies (23) and completes the proof. □

3.4 Chung-Lu model: proof of Theorem 4

Let $k = k(n) \in \mathbb{N}$ be such that $1 \ll k \ll \log n$. Let $S_1 = [0, \frac{1}{k}]$, $S_2 = (\frac{1}{k}, \frac{2}{k}], \ldots, S_k = (\frac{k-1}{k}, 1]$. Define $\boldsymbol{n}(\boldsymbol{u}) = (n_1, \ldots, n_k)^T$ by

$$n_i = n_i(\boldsymbol{u}) := |\{t \in [n] \,:\, u_t \in S_i\}|\,.$$

Define two $k \times k$ matrices $P^L = (p^L_{ij})_{i,j\in[k]}$ and $P^U = (p^U_{ij})_{i,j\in[k]}$ by

$$p^L_{ij} := p \cdot \tfrac{(i-1)(j-1)}{k^2}, \qquad p^U_{ij} := p \cdot \tfrac{ij}{k^2}.$$

Then, for any two vertices $a, b \in V(\boldsymbol{G}^\times_p) = [n]$, we have

$$p^L_{ij} \leqslant p^\times_{ab} \leqslant p^U_{ij},$$

where $i = i(a)$ and $j = j(b)$ are such that $u_a \in S_i$ and $u_b \in S_j$. Therefore, there are two random graphs $\boldsymbol{G}^L \sim \mathcal{G}(\boldsymbol{n}(\boldsymbol{u}), P^L)$ and $\boldsymbol{G}^U \sim \mathcal{G}(\boldsymbol{n}(\boldsymbol{u}), P^U)$ such that $\boldsymbol{G}^L \subseteq \boldsymbol{G}^\times_p \subseteq \boldsymbol{G}^U$. Furthermore, we find that

$$\chi\left(\boldsymbol{G}^U\right) \leqslant \chi\left(\boldsymbol{G}^\times_p\right) \leqslant \chi\left(\boldsymbol{G}^L\right).$$

Let $Q^L = Q(P^L)$ and $Q^U = Q(P^U)$ be defined according to (3).

Next, we show that $w_*\left(\boldsymbol{n}(\boldsymbol{u}), Q^L\right) = \Omega(n)$ and $w_*\left(\boldsymbol{n}(\boldsymbol{u}), Q^U\right) = \Omega(n)$. Then the assumptions of Theorem 5 hold for both random graphs $\boldsymbol{G}^L$ and $\boldsymbol{G}^U$. Indeed, setting

$\sigma := -\frac{\log p}{\log n}$ and using the assumptions of Theorem 4, that is, $1 \gg p \geqslant n^{-1/4+\epsilon}$ and $\sum_{i\in[n]} u_i = \Omega(n)$, we get

$$q^*, \hat{q}((\boldsymbol{n}(\boldsymbol{u})) = \Theta(p) = \Theta(n^{-\sigma}) = n^{-\sigma+o(1)} \ll 1$$

for $\boldsymbol{G}^L$ and $\boldsymbol{G}^U$. Recalling also $k \ll \log n$, we obtain (8) and (9). Finally (10) holds if $w_*\left(\boldsymbol{n}(\boldsymbol{u}), Q^L\right)$ and $w_*\left(\boldsymbol{n}(\boldsymbol{u}), Q^U\right)$ are $\Omega(n)$ since $k\hat{q}(\boldsymbol{n}(\boldsymbol{u}))q^* \ll \log n$. Thus, to complete the proof of Theorem 4(a), it remains to establish the following lemma.

Lemma 15. *Suppose the assumptions of Theorem 4 hold. Then*

$$\max_{U\subseteq[n]} \frac{1}{|U|}\left(\sum\nolimits_{t\in U} u_t\right)^2 = \frac{1+o(1)}{p} w_*\left(\boldsymbol{n}(\boldsymbol{u}), Q^L\right) = \frac{1+o(1)}{p} w_*\left(\boldsymbol{n}(\boldsymbol{u}), Q^U\right) = \Omega(n).$$

Proof. Since $\sum_{t\in[n]} u_t = \Omega(n)$, we find that

$$M := \max_{U\subseteq[n]} \frac{1}{|U|}\left(\sum\nolimits_{t\in U} u_t\right)^2 \geqslant \frac{1}{n}\left(\sum\nolimits_{t\in[n]} u_t\right)^2 = \Omega(n). \tag{27}$$

Since $p = o(1)$, we have $\log \frac{1}{1-pxy} = (1+o(1))pxy$ uniformly over all $x, y \in [0,1]$. Then, by the definition of $w_*(\cdot)$, we derive that

$$w_*\left(\boldsymbol{n}(\boldsymbol{u}), Q^L\right) = (1+o(1)) w_*\left(\boldsymbol{n}(\boldsymbol{u}), P^L\right),$$
$$w_*\left(\boldsymbol{n}(\boldsymbol{u}), Q^U\right) = (1+o(1)) w_*\left(\boldsymbol{n}(\boldsymbol{u}), P^U\right).$$

For any $\boldsymbol{x} \in \mathbb{R}^k$, we have

$$\boldsymbol{x}^T P^L \boldsymbol{x} = \frac{p}{k^2}\left(\sum\nolimits_{i\in[k]} x_i(i-1)\right)^2 \geqslant 0. \tag{28}$$

Using the pseudodefinite property in Theorem 10(c), we find that $w_*\left(\boldsymbol{n}(\boldsymbol{u}), P^L\right) = w\left(\boldsymbol{n}(\boldsymbol{u}), P^L\right)$. Similarly, we get that $w_*\left(\boldsymbol{n}(\boldsymbol{u}), P^U\right) = w\left(\boldsymbol{n}(\boldsymbol{u}), P^U\right)$. Thus, it remains to show that

$$M = \frac{1+o(1)}{p} w\left(\boldsymbol{n}(\boldsymbol{u}), P^L\right) = \frac{1+o(1)}{p} w\left(\boldsymbol{n}(\boldsymbol{u}), P^U\right). \tag{29}$$

Let $U^* \subseteq [n]$ be the set that maximises $\frac{1}{|U|}\left(\sum_{t\in U} u_t\right)^2$, that is,

$$M = \frac{1}{|U^*|}\left(\sum\nolimits_{t\in U_*} u_t\right)^2.$$

Using the trivial bound $\sum_{t\in U^*} u_t \leqslant |U^*|$ and (27), we get that

$$\sum\nolimits_{t\in U^*} u_t \geqslant \frac{1}{|U^*|}\left(\sum\nolimits_{t\in U^*} u_t\right)^2 = \Omega(n). \tag{30}$$

Let

$$\boldsymbol{x}(U^*) := (x_1, \ldots, x_k)^T \in \mathbb{N}^k, \qquad x_i := |\{t \in [n] \,:\, u_t \in S_i \cap U^*\}.$$

Combining (30) and the trivial bound $\|\boldsymbol{x}(U^*)\| = |U^*| \leqslant n$, we get that

$$\textstyle\sum_{t\in U^*} u_t = (1+O(k^{-1}))\displaystyle\sum_{i\in[k]} \frac{i-1}{k}x_i = (1+O(k^{-1}))\sum_{i\in[k]} \frac{i}{k}x_i.$$

Due to (28) and a similar formula for P^U, we get that $(\sum_{t\in U^*} u_t)^2$ is equivalent to $\boldsymbol{x}^T(U^*)P^L\boldsymbol{x}(U^*)$ and $\boldsymbol{x}^T(U^*)P^U\boldsymbol{x}(U^*)$ up to the factor p. Recalling $\|\boldsymbol{x}(U^*)\| = |U^*|$, we get that

$$pM = (1+o(1))\frac{\boldsymbol{x}^T(U^*)P^L\boldsymbol{x}(U^*)}{\|\boldsymbol{x}(U^*)\|} = (1+o(1))\frac{\boldsymbol{x}^T(U^*)P^U\boldsymbol{x}(U^*)}{\|\boldsymbol{x}(U^*)\|}.$$

This implies

$$w\left(\boldsymbol{n}(\boldsymbol{u}), P^U\right) \geqslant w\left(\boldsymbol{n}(\boldsymbol{u}), P^L\right) \geqslant (1+o(1))pM.$$

For the other direction, using the corner maximiser property in Theorem 10(b), we get that there is $W \subseteq [k]$ such that

$$w\left(\boldsymbol{n}(\boldsymbol{u}), P^U\right) = \frac{p}{\sum_{i\in W} n_i}\left(\sum_{i\in W}\frac{i}{k}n_i\right)^2.$$

Let $U(W) := \{t \in [n] \,:\, u_t \in \cup_{i\in W} S_i\}$. Then, $\sum_{i\in W} n_i = |U(W)|$. We also have

$$\sum_{i\in W}\frac{i-1}{k}n_i \leqslant \sum_{t\in U(W)} u_t \leqslant \sum_{i\in W}\frac{i}{k}n_i.$$

We already established that $w\left(\boldsymbol{n}(\boldsymbol{u}), P^U\right) \geqslant (1+o(1))pM = \Omega(pn)$. Thus,

$$\sum_{i\in W}\frac{i}{k}n_i \geqslant \frac{1}{\sum_{i\in W} n_i}\left(\sum_{i\in W}\frac{i}{k}n_i\right)^2 = \Omega(n).$$

Therefore,

$$\frac{1}{\sum_{i\in W} n_i}\left(\sum_{i\in W}\frac{i}{k}n_i\right)^2 = (1+O(k^{-1}))\frac{1}{|U|}\left(\sum_{t\in U} u_t\right)^2.$$

This implies

$$w\left(\boldsymbol{n}(\boldsymbol{u}), P^L\right) \leqslant w\left(\boldsymbol{n}(\boldsymbol{u}), P^U\right) \leqslant (1+o(1))pM.$$

This completes the proof of required bound (29) and of the lemma. □

We proceed to the proof of Theorem 4(b). Define two $k \times k$ matrices $\widehat{P}^L = (\widehat{p}^L_{ij})_{i,j\in[k]}$ and $\widehat{P}^U = (\widehat{p}^U_{ij})_{i,j\in[k]}$ by

$$\widehat{p}^L_{ij} := p \cdot \tfrac{(i-1)+(j-1)}{k}, \qquad \widehat{p}^U_{ij} := p \cdot \tfrac{i+j}{k}.$$

Then, for any two vertices $a, b \in V(\boldsymbol{G}_p^+) = [n]$, we have

$$\widehat{p}_{ij}^L \leqslant p_{ab}^+ \leqslant \widehat{p}_{ij}^U,$$

where $i = i(a)$ and $j = j(b)$ are such that $u_a \in S_i$ and $u_b \in S_j$. Therefore, there are two random graphs $\widehat{\boldsymbol{G}}^L \sim \mathcal{G}(\boldsymbol{n}(\boldsymbol{u}), \widehat{P}^L)$ and $\widehat{\boldsymbol{G}}^U \sim \mathcal{G}(\boldsymbol{n}(\boldsymbol{u}), \widehat{P}^U)$ such that $\widehat{\boldsymbol{G}}^L \subseteq \boldsymbol{G}_p^+ \subseteq \widehat{\boldsymbol{G}}^U$. Furthermore, we find that

$$\chi\left(\widehat{\boldsymbol{G}}^U\right) \leqslant \chi\left(\boldsymbol{G}_p^+\right) \leqslant \chi\left(\widehat{\boldsymbol{G}}^L\right).$$

Then Theorem 4(b) follows immediately by combining Theorem 5 and the following lemma.

Lemma 16. *Let the assumptions of Theorem 4 hold. Let $\widehat{Q}^L = Q(\widehat{P}^L)$ and $Q^U = Q(\widehat{P}^U)$ be defined according to* (3)*. Then*

$$\sum_{t\in[n]} u_t = \frac{1+o(1)}{p} w_*\left(\boldsymbol{n}(\boldsymbol{u}), \widehat{Q}^L\right) = \frac{1+o(1)}{p} w_*\left(\boldsymbol{n}(\boldsymbol{u}), \widehat{Q}^U\right) = \Omega(n).$$

Proof. Since $p = o(1)$, we have $\log \frac{1}{1-p(x+y)} = (1+o(1))p(x+y)$ uniformly over $x, y \in [0,1]$. Then, by the definition of $w_*(\cdot)$, we observe that

$$w_*\left(\boldsymbol{n}(\boldsymbol{u}), \widehat{Q}^L\right) = (1+o(1)) w_*\left(\boldsymbol{n}(\boldsymbol{u}), \widehat{P}^L\right),$$
$$w_*\left(\boldsymbol{n}(\boldsymbol{u}), \widehat{Q}^U\right) = (1+o(1)) w_*\left(\boldsymbol{n}(\boldsymbol{u}), \widehat{P}^U\right).$$

For any $\boldsymbol{x} \in \mathbb{R}_+^k$, we have

$$w\left(\boldsymbol{x}, \widehat{P}^L\right) = \max_{\boldsymbol{0}\prec\boldsymbol{y}\prec\boldsymbol{x}} \frac{\boldsymbol{y}^T \widehat{P}^L \boldsymbol{y}}{\|\boldsymbol{y}\|} = p \cdot \max_{\boldsymbol{0}\prec\boldsymbol{y}\prec\boldsymbol{x}} \sum_{i\in[k]} \frac{i-1}{k} y_i = p \cdot \sum_{i\in[k]} \frac{i-1}{k} x_i.$$

By the definition of $w_*(\cdot)$, we find that

$$w_*\left(\boldsymbol{n}(\boldsymbol{u}), \widehat{P}^L\right) = p \cdot \sum_{i\in[k]} \frac{i-1}{k} n_i.$$

Observing that

$$\sum_{i\in[k]} \frac{i}{k} n_i \geqslant \sum_{t\in[n]} u_t \geqslant \sum_{i\in[k]} \frac{i-1}{k} n_i$$

and recalling $\sum_{t\in[n]} u_t = \Omega(n)$, we derive

$$w_*\left(\boldsymbol{n}(\boldsymbol{u}), \widehat{P}^L\right) = (1 + O(k^{-1}))p \cdot \sum\nolimits_{t\in[n]} u_t.$$

Similarly, we prove $w_*\left(\boldsymbol{n}(\boldsymbol{u}), \widehat{P}^U\right) = (1 + O(k^{-1}))p \cdot \sum_{t\in[n]} u_t$. This completes the proof. ∎

4 Weighted independence number

A set $U \subseteq V(G)$ is an *independent set* of a graph G if the induced graph $G[U]$ has no edges. Let $\mathcal{I}(G)$ denote the set of all the independent sets of G. The *independence number* $\alpha(G)$ equals the size of a largest independent set of G. It is well known that (see, for example [19, 25]) if $np \to \infty$ and $p < 1 - \varepsilon$ for a constant $\varepsilon \in (0, 1)$, then whp

$$\chi(\boldsymbol{G}(n,p)) = (1 + o(1))\frac{n}{\alpha(\boldsymbol{G}(n,p))}. \tag{31}$$

That is, for an asymptotically optimal colouring of $\boldsymbol{G}(n,p)$, almost all vertices are covered with colour classes of approximately equal size $\alpha(\boldsymbol{G}(n,p))$.

One may think that, to approach the chromatic number of inhomogeneous random graphs, one can also start with its independence number. In fact, Doležal et al. [11] studied the clique number in inhomogeneous random graphs. Note that the clique number of a graph equals the independence number of its complement. However, we find little use of the results of [11] in determining the chromatic number of a random graph $\boldsymbol{G} \sim \mathcal{G}(\boldsymbol{n}, P)$ from the stochastic block model. Unlike the homogeneous binomial random graph $\boldsymbol{G}(n,p)$, some parts of the random graph $\boldsymbol{G} \sim \mathcal{G}(\boldsymbol{n}, P)$ will typically contain substantially larger independent sets than other parts of the graph so one can not achieve an optimal colouring using colour classses of approximately same size.

To take the inhomogeneity of $\boldsymbol{G} \sim \mathcal{G}(\boldsymbol{n}, P)$ into account, we assign special weights to subsets of vertices (depending on the edge probabilities) and introduce a new parameter, called *weighted independence number*, which is the maximal weight of an independent set. Formally, for a set $U \subseteq V(\boldsymbol{G})$, define

$$h(U) = h(U, \boldsymbol{n}, P) := \frac{-\log(\Pr(U \in \mathcal{I}(\boldsymbol{G})))}{|U|}.$$

Then, for a graph G on vertex set $V(G) = V(\boldsymbol{G})$, let

$$\alpha_h(G) = \alpha_h(G, \boldsymbol{n}, P) := \max_{U \in \mathcal{I}(G),\, U \neq \emptyset} h(U). \tag{32}$$

It might be not obvious but nevertheless true that the weights $h(U)$ are designed in such a way that all maximal independent sets U in the random graph $\boldsymbol{G}$ have similar weights whp. This is a natural generalisation of the idea of the balanced colouring of $\boldsymbol{G}(n,p)$ except we use the weight instead of the size of a colour class.

In this section, we show, in particular that, under the assumptions of Theorem 5 and provided that not all blocks B_i are very small, the quantity $\alpha_h(\boldsymbol{G})$ is concentrated around $(1 - \sigma)\log\|\boldsymbol{n}\|$ whp; see Theorem 19. Moreover, we establish fast decreasing tail bounds for the probability of $\alpha_h(\boldsymbol{G})$ being too large or too small; see Lemmas 17 and 18, respectively. Lemma 17 almost immediately leads to the proof of Theorem 8. Even though, Lemma 18 does not immediately give Theorem 9, it will be the crucial instrument for our construction of an optimal colouring of $\boldsymbol{G}$ in further sections.

Let $Q = Q(P)$ be defined by (3), where P is the matrix of edge probabilities for $\boldsymbol{G} \sim \mathcal{G}(\boldsymbol{n}, P)$. For simplicity, everywhere in this section, let

$$w(\cdot) \equiv w(\cdot, Q) \quad \text{and} \quad w_*(\cdot) \equiv w_*(\cdot, Q);$$

see (4), (6) for definitions. Let q^* and $\hat{q}(\cdot)$ be defined according to (7). In addition, we consider the vector-valued function $\boldsymbol{b} : 2^{V(\boldsymbol{G})} \to \mathbb{N}^k$ that maps $U \subseteq V(\boldsymbol{G})$ into $\boldsymbol{b}(U)$ defined by

$$\boldsymbol{b}(U) = (b_1(U), \dots, b_k(U))^T \quad \text{with} \quad b_i(U) := |U \cap B_i| \quad \text{for } i \in [k].$$

Here, B_i are the blocks of vertices in the stochastic block model $\mathcal{G}(\boldsymbol{n}, P)$. Note that, for any $U \subseteq V(\boldsymbol{G})$, we have that $\|\boldsymbol{b}(U)\| = |U|$ and

$$\boldsymbol{b}(U)^T Q\, \boldsymbol{b}(U) = -2 \log\left(\Pr(U \in \mathcal{I}(\boldsymbol{G}))\right) + \sum_{i \in [k]} q_{ii} b_i(U) \tag{33}$$

$$\leqslant -2 \log\left(\Pr(U \in \mathcal{I}(\boldsymbol{G}))\right) + q^* |U|. \tag{34}$$

4.1 Lower tail bound: proof of Theorem 8

First, we estimate the probability of $\alpha_h(\boldsymbol{G})$ to be large for a general random graph $\boldsymbol{G}$ with independent adjacencies.

Lemma 17. *Let $\boldsymbol{G}$ be a random graph on n vertices where edges appear independently of each other. Assume $s_t e^t \geqslant 6n$ for some $t > 0$, where*

$$s_t := \min\left\{|U| \,:\, \emptyset \neq U \subseteq V(\boldsymbol{G}),\ \Pr(U \in \mathcal{I}(\boldsymbol{G})) \leqslant e^{-t|U|}\right\}.$$

Then

$$\Pr(\alpha_h(\boldsymbol{G}) \geqslant t) \leqslant 2^{1-s_t}.$$

Proof. Let X_s denote the number of independent sets U of size s in $\boldsymbol{G}$ such that

$$\Pr(U \in \mathcal{I}(\boldsymbol{G})) \leqslant e^{-t|U|}.$$

By definition of s_t, we have that $X_s = 0$ for any $s < s_t$. If $s \geqslant s_t$ then we bound

$$\Pr(X_s > 0) \leqslant \mathbb{E}\, X_s \leqslant \sum_U \Pr(U \in \mathcal{I}(\boldsymbol{G}))$$

$$\leqslant \binom{n}{s} e^{-ts} \leqslant \left(\frac{en}{se^t}\right)^s \leqslant 2^{-s},$$

where the sum is over all U that contribute to X_s. Thus, we can bound

$$\Pr(\alpha_h(\boldsymbol{G}) \geqslant t) \leqslant \sum_{s=s_t}^{n} \Pr(X_s > 0) \leqslant \sum_{s=s_t}^{n} 2^{-s} \leqslant 2^{1-s_t},$$

which concludes the proof. ∎

Next, applying Lemma 17 to $\boldsymbol{G} \sim \mathcal{G}(\boldsymbol{n}, P)$, we derive the required probability bound for the event that the chromatic number $\chi(\boldsymbol{G})$ is small.

Proof of Theorem 8. Take $t := \log(q^*\|\boldsymbol{n}\|)$. To apply Lemma 17, we need to bound the quantity s_t in Lemma 17. If U is such that $\Pr(U \in \mathcal{I}(\boldsymbol{G})) \leqslant e^{-t|U|}$, then using (33), we get that

$$t|U| \leqslant -\log\left(\Pr(U \in \mathcal{I}(\boldsymbol{G}))\right) \leqslant \tfrac{1}{2}|U|^2 \max_{i,j\in[k]} q_{ij}.$$

Then, by the assumptions, we get that

$$s_t \geqslant \frac{2t}{\max_{i,j\in[k]} q_{ij}} = \frac{2\log(q^*\|\boldsymbol{n}\|)}{\max_{i,j\in[k]} q_{ij}} \to \infty \tag{35}$$

and

$$s_t e^t \geqslant \frac{2te^t}{\max_{i,j\in[k]} q_{ij}} = \frac{2q^*\|\boldsymbol{n}\|\log(q^*\|\boldsymbol{n}\|)}{\max_{i,j\in[k]} q_{ij}} \gg \|\boldsymbol{n}\|.$$

Applying Lemma 17, we find that

$$\Pr(\alpha_h(\boldsymbol{G}) \geqslant \log(q^*\|\boldsymbol{n}\|)) \leqslant 2^{1-s_t}. \tag{36}$$

Next, using the corner maximiser property in Theorem 10(b) and (34), we find that, for any $U \in \mathcal{I}(\boldsymbol{G})$,

$$w(\boldsymbol{b}(U)) = \max_{\emptyset\neq W\subseteq U} \frac{\boldsymbol{b}(W)^T Q \boldsymbol{b}(W)}{|W|} \leqslant \max_{\emptyset\neq W\subseteq U} \left(-\frac{2\Pr(W\in\mathcal{I}(\boldsymbol{G}))}{|W|} + q^*\right) \leqslant 2\alpha_h(\boldsymbol{G}) + q^*. \tag{37}$$

In the above, we also used that if $W \subseteq U \in \mathcal{I}(\boldsymbol{G})$ then $W \in \mathcal{I}(\boldsymbol{G})$. Recall that $\log(q^*\|\boldsymbol{n}\|) = \Theta(\log\|\boldsymbol{n}\|)$ by (8) and $q^* \leqslant \max_{i,j\in[k]} q_{ij} \ll \log\|\boldsymbol{n}\|$ by (9). Thus, if $\alpha_h(\boldsymbol{G}) \leqslant \log(q^*\|\boldsymbol{n}\|)$ then, by (37), we have that

$$\max_{U\in\mathcal{I}(\boldsymbol{G})} w(\boldsymbol{b}(U)) \leqslant 2\log(q^*\|\boldsymbol{n}\|)) + q^* = (2+o(1))\log(q^*\|\boldsymbol{n}\|)).$$

Let $\{U_i\}_{i=1,\ldots,\chi(\boldsymbol{G})}$ be the partition of $V(\boldsymbol{G})$ into colour classes of any optimal colouring of $\boldsymbol{G}$. Cosidering the system consisting of vectors $\boldsymbol{b}(U_i)$ for $i = 1, \ldots, \chi(\boldsymbol{G})$, and recalling definition (6), we find that

$$w_*(\boldsymbol{n}) \leqslant \sum_{i=1}^{\chi(\boldsymbol{G})} w(\boldsymbol{b}(U_i)).$$

We conclude that, with probability at least $1 - 2^{1-s_t}$,

$$\chi(\boldsymbol{G}) \geqslant \frac{\sum_{i=1}^{\chi(\boldsymbol{G})} w(\boldsymbol{b}(U_i))}{\max_i w(\boldsymbol{b}(U_i))} \geqslant \frac{w_*(\boldsymbol{n})}{(2+o(1))\log(q^*\|\boldsymbol{n}\|)} \geqslant (1-\varepsilon)\frac{w_*(\boldsymbol{n})}{2\log(q^*\|\boldsymbol{n}\|)}.$$

Using (35), we get that

$$2^{1-s_t} = \exp\left(-\Omega\left(\frac{\log(q^*\|\boldsymbol{n}\|)}{\max_{i,j\in[k]} q_{ij}}\right)\right).$$

This completes the proof. ∎

4.2 Existence of heavy independent sets

We consider a special class of sets distributed between the blocks $B_1, \ldots, B_k$ proportionally to its sizes (up to rounding). For a vector $\boldsymbol{x} = (x_1, \ldots, x_k)^T \in \mathbb{R}^k$, denote

$$\lfloor \boldsymbol{x} \rfloor := (\lfloor x_1 \rfloor, \ldots, \lfloor x_k \rfloor)^T \qquad \text{and} \qquad x_* := \min_{i \in [k]} x_i.$$

For a positive real ν, let $\mathcal{I}_\nu(\boldsymbol{G})$ denote the family of independent sets $U \subseteq \mathcal{I}(\boldsymbol{G})$ such that $\boldsymbol{b}(U) = \lfloor \nu \boldsymbol{n} \rfloor$.

Lemma 18. *Let $\boldsymbol{G} \sim \mathcal{G}(\boldsymbol{n}, P)$, where $P = (p_{ij})_{i,j \in [k]}$ is such that $p_{ij} = p_{ji}$ and $0 \leqslant p_{ij} < 1$ for all $i, j \in [k]$. Let $Q = Q(P)$ be as in (3) and $\sigma \in [0, \sigma_0)$ for some fixed $0 < \sigma_0 < \frac{1}{2}$. Assume that $\|\boldsymbol{n}\| \to \infty$, $w(\boldsymbol{n}) \geqslant \|\boldsymbol{n}\|^{1-\sigma}$,*

$$n_* = \|\boldsymbol{n}\|^{1+o(1)}, \qquad n_* \gg \frac{w(\boldsymbol{n})}{\log \|\boldsymbol{n}\|}, \tag{38}$$

where $n_* := \min_{i \in [k]} n_i$. Then, there exists $\nu \in \mathbb{R}_+$ such that $\nu = (2 + o(1)) \frac{\log(w(\boldsymbol{n}))}{w(\boldsymbol{n})}$ and

$$\Pr\left(\mathcal{I}_\nu(\boldsymbol{G}) = \emptyset\right) \leqslant \exp\left(-\|\boldsymbol{n}\|^{2-4\sigma+o(1)}\right).$$

Proof. Since $n_* \gg \frac{w(\boldsymbol{n})}{\log \|\boldsymbol{n}\|}$, we can find some $r(\boldsymbol{n})$ such that

$$\frac{w(\boldsymbol{n})}{n_*} \ll r(\boldsymbol{n}) \ll \log \|\boldsymbol{n}\|. \tag{39}$$

For example, one can take $r(\boldsymbol{n}) := \left(\frac{w(\boldsymbol{n})}{n_*} \log \|\boldsymbol{n}\|\right)^{\frac{1}{2}}$. Define

$$\nu := \frac{2}{w(\boldsymbol{n})} \Bigg(\log(w(\boldsymbol{n})) - 2 \log \log(w(\boldsymbol{n})) - \log\left(\frac{\|\boldsymbol{n}\|}{n_*}\right) - r(\boldsymbol{n}) \Bigg). \tag{40}$$

Note that the assumptions imply that

$$\log(w(\boldsymbol{n})) \geqslant \frac{1}{2} \log \|\boldsymbol{n}\|, \qquad \nu = (2 + o(1)) \frac{\log(w(\boldsymbol{n}))}{w(\boldsymbol{n})}. \tag{41}$$

Let $\boldsymbol{\ell} = (\ell_1, \ldots, \ell_k) = \lfloor \nu \boldsymbol{n} \rfloor$. That is, we have $\boldsymbol{\ell} = \boldsymbol{b}(U)$ for all $U \in \mathcal{I}_\nu(\boldsymbol{G})$. Using the assumptions, we get, for all $i \in [k]$,

$$\ell_i = (2 + o(1)) \frac{n_i \log(w(\boldsymbol{n}))}{w(\boldsymbol{n})} \geqslant (2 + o(1)) \frac{n_* \log(w(\boldsymbol{n}))}{w(\boldsymbol{n})} \gg 1. \tag{42}$$

Observe that the number of ways to pick the set $U \in V(\boldsymbol{G})$ such that $\boldsymbol{b}(U) = \boldsymbol{\ell}$ equals $\prod_{i=1}^k \binom{n_i}{\ell_i}$. Then, using (33) that relates $\Pr(U \in \mathcal{I}_\nu(\boldsymbol{G}))$ and $e^{-\frac{\boldsymbol{b}(U)^T Q \, \boldsymbol{b}(U)}{2}} = e^{-\frac{\boldsymbol{\ell}^T Q \, \boldsymbol{\ell}}{2}}$, we get that

$$\mathbb{E}\,|\mathcal{I}_\nu(\boldsymbol{G})| = \Pr(U \in \mathcal{I}_\nu(\boldsymbol{G})) \prod_{i=1}^{k} \binom{n_i}{\ell_i} = e^{-\frac{\boldsymbol{\ell}^T Q \boldsymbol{\ell}}{2}} \prod_{i=1}^{k} \binom{n_i}{\ell_i} (1-p_{ii})^{-\ell_i/2}$$

$$\geqslant e^{-\frac{\boldsymbol{\ell}^T Q \boldsymbol{\ell}}{2}} \prod_{i=1}^{k} \left(\frac{n_i}{\ell_i}\right)^{\ell_i} \geqslant e^{-\frac{\boldsymbol{\ell}^T Q \boldsymbol{\ell}}{2}} \nu^{-\|\boldsymbol{\ell}\|}.$$

Using the scaling property in Theorem 10(a) and the definition (40) of ν, we get

$$\frac{\boldsymbol{\ell}^T Q\, \boldsymbol{\ell}}{2\|\boldsymbol{\ell}\|} = \nu \frac{(\nu^{-1}\boldsymbol{\ell})^T Q\, (\nu^{-1}\boldsymbol{\ell})}{2\|\nu^{-1}\boldsymbol{\ell}\|} \leqslant \frac{\nu\, w(\boldsymbol{n})}{2} = \log\left(\frac{w(\boldsymbol{n}) n_*}{\log^2(w(\boldsymbol{n}))\|\boldsymbol{n}\|}\right) - r(\boldsymbol{n}). \tag{43}$$

From (42), we also get that

$$\|\boldsymbol{\ell}\| = (2+o(1))\frac{\|\boldsymbol{n}\| \log(w(\boldsymbol{n}))}{w(\boldsymbol{n})}.$$

Using (39), (41), (43), the obvious inequality $n_* \leqslant \|\boldsymbol{n}\|$, and $w(\boldsymbol{n}) \geqslant \|\boldsymbol{n}\|^{1-\sigma}$, we get that

$$\begin{aligned}\mathbb{E}\,|\mathcal{I}_\nu(\boldsymbol{G})| &\geqslant e^{-\frac{\boldsymbol{\ell}^T Q \boldsymbol{\ell}}{2}} \nu^{-\|\boldsymbol{\ell}\|} \geqslant \left(\left(\tfrac{1}{2}+o(1)\right) \log(w(\boldsymbol{n})) \frac{\|\boldsymbol{n}\|}{n_*} e^{r(\boldsymbol{n})}\right)^{\|\boldsymbol{\ell}\|} \\ &\gg e^{\|\boldsymbol{\ell}\| r(\boldsymbol{n})} = \exp\left(\omega\left(\frac{\|\boldsymbol{\ell}\| w(\boldsymbol{n})}{n_*}\right)\right) = \|\boldsymbol{n}\|^{\omega(1)}.\end{aligned} \tag{44}$$

Next, let

$$\Delta := \sum_{|U\cap W|\geqslant 2} \Pr\left(U \in \mathcal{I}_\nu(\boldsymbol{G}) \text{ and } W \in \mathcal{I}_\nu(\boldsymbol{G})\right),$$

where the sum is over all possible ordered pairs (U, W) of subsets of $V(\boldsymbol{G})$ such that $|U \cap W| \geqslant 2$. Note that if $|U \cap W| \leqslant 1$ then the events $\{U \in \mathcal{I}_\nu(\boldsymbol{G})\}$ and $\{W \in \mathcal{I}_\nu(\boldsymbol{G})\}$ are independent. By Janson's inequality, see [20, Theorem 1], we have

$$\Pr\left(\mathcal{I}_\nu(\boldsymbol{G}) = \emptyset\right) \leqslant \exp\left(-\frac{(\mathbb{E}\,|\mathcal{I}_\nu(\boldsymbol{G})|)^2}{2\,\mathbb{E}\,|\mathcal{I}_\nu(\boldsymbol{G})| + 2\Delta}\right). \tag{45}$$

We have already established a lower bound for $\mathbb{E}\,|\mathcal{I}_\nu(\boldsymbol{G})|$ in (44). Thus, it remains to bound $\frac{\Delta}{(\mathbb{E}\,|\mathcal{I}_\nu(\boldsymbol{G})|)^2}$ from the above. Using (33), we find that

$$\frac{\Delta}{(\mathbb{E}\,|\mathcal{I}_\nu(\boldsymbol{G})|)^2} = \sum_{\boldsymbol{m}} e^{\frac{\boldsymbol{m}^T Q \boldsymbol{m}}{2}} \prod_{i=1}^{k} \frac{\binom{\ell_i}{m_i}\binom{n_i-\ell_i}{\ell_i-m_i}}{\binom{n_i}{\ell_i}} (1-p_{ii})^{m_i/2},$$

where the sums are over $\boldsymbol{m} = (m_1, \ldots, m_k)^T \in \mathbb{N}^k$ with $\|\boldsymbol{m}\| \geqslant 2$ and $\boldsymbol{m} \preceq \boldsymbol{\ell}$. Observe that

$$\begin{aligned}\frac{\binom{\ell_i}{m_i}\binom{n_i-\ell_i}{\ell_i-m_i}}{\binom{n_i}{\ell_i}} &= \frac{((\ell_i)_{m_i})^2 (n_i-\ell_i)_{\ell_i-m_i}}{m_i!(n_i)_{\ell_i}} \leqslant \frac{((\ell_i)_{m_i})^2}{m_i!(n_i)_{m_i}} \leqslant \frac{1}{m_i!}\left(\frac{\ell_i^2}{n_i}\right)^{m_i} \\ &= \frac{\left((1+o(1))\nu^2 n_i\right)^{m_i}}{m_i!} \leqslant \frac{1}{m_i!}\left(5n_i\left(\frac{\log(w(\boldsymbol{n}))}{w(\boldsymbol{n})}\right)^2\right)^{m_i}.\end{aligned}$$

Denote

$$\theta_m := \max_{\|\boldsymbol{m}\|=m} \frac{\boldsymbol{m}^T Q \boldsymbol{m}}{2\|\boldsymbol{m}\|},$$

where the maximum is over $\boldsymbol{m} \in \mathbb{N}^k$ with $\|\boldsymbol{m}\| = m$ and $\boldsymbol{m} \preceq \boldsymbol{\ell}$. Then, we obtain

$$\begin{aligned} \frac{\Delta}{(\mathbb{E}\,|\mathcal{I}_\nu(\boldsymbol{G})|)^2} &\leqslant \sum_{m=2}^{\|\boldsymbol{\ell}\|} \left(5\left(\frac{\log(w(\boldsymbol{n}))}{w(\boldsymbol{n})}\right)^2 e^{\theta_m}\right)^{m_i} \prod_{i=1}^{k} \frac{n_i^{m_i}}{m_i!} \\ &= \sum_{m=2}^{\|\boldsymbol{\ell}\|} \frac{1}{m!}\left(5\|\boldsymbol{n}\|\left(\frac{\log(w(\boldsymbol{n}))}{w(\boldsymbol{n})}\right)^2 e^{\theta_m}\right)^{m}. \end{aligned} \tag{46}$$

There are two ways we can estimate the quantity θ_m. First, repeating the arguments of (43) with $\boldsymbol{\ell}$ replaced by any $\boldsymbol{m} \preceq \boldsymbol{\ell}$, we find that

$$\theta_m \leqslant \frac{\nu\, w(\boldsymbol{n})}{2} = \log\left(\frac{w(\boldsymbol{n})n_*}{\log^2(w(\boldsymbol{n}))\|\boldsymbol{n}\|}\right) - r(\boldsymbol{n}). \tag{47}$$

Second, observing $\frac{n_*}{\|\boldsymbol{m}\|}\boldsymbol{m} \preceq \boldsymbol{n}$ and using the monotonicity property in Theorem 10(a), we get

$$\frac{\boldsymbol{m}^T\, Q\, \boldsymbol{m}}{2\|\boldsymbol{m}\|} \leqslant \frac{\|\boldsymbol{m}\|\, w\left(\frac{n_*}{\|\boldsymbol{m}\|}\boldsymbol{m}\right)}{2\, n_*} \leqslant \frac{\|\boldsymbol{m}\| w(\boldsymbol{n})}{2\, n_*}.$$

Thus, we get

$$\theta_m \leqslant \frac{m\, w(\boldsymbol{n})}{2n_*}, \tag{48}$$

which is better than (47) for small m.

Using (39), we can find $m_0 \in \mathbb{N}$ such that

$$1 \ll \frac{\log\|\boldsymbol{n}\|}{r(\boldsymbol{n})} + \frac{n_*}{w(\boldsymbol{n})} \ll m_0 \ll \frac{n_* \log\|\boldsymbol{n}\|}{w(\boldsymbol{n})}. \tag{49}$$

Using the inequality $m! \geqslant m^m e^{-m}$ and the bound $e^{\theta_m} \leqslant \frac{w(\boldsymbol{n})n_*}{\log^2(w(\boldsymbol{n}))\|\boldsymbol{n}\|} e^{-r(\boldsymbol{n})}$ implied by (47), we find that

$$\begin{aligned} \sum_{m=m_0}^{\|\boldsymbol{\ell}\|} \frac{1}{m!}\left(5\|\boldsymbol{n}\|\left(\frac{\log(w(\boldsymbol{n}))}{w(\boldsymbol{n})}\right)^2 e^{\theta_m}\right)^m &\leqslant \sum_{m=m_0}^{\|\boldsymbol{\ell}\|}\left(\frac{5e^{1-r(\boldsymbol{n})}n_*}{m\, w(\boldsymbol{n})}\right)^m \\ &\ll \sum_{m=m_0}^{\|\boldsymbol{\ell}\|} e^{-mr(\boldsymbol{n})} \leqslant \|\boldsymbol{\ell}\| e^{-\omega(\log\|\boldsymbol{n}\|)} = \|\boldsymbol{n}\|^{-\omega(1)}, \end{aligned} \tag{50}$$

where the last two inequalities used the lower bounds of (49): first $m \geqslant m_0 \gg \frac{n_*}{w(\boldsymbol{n})}$ and then $m \geqslant m_0 \gg \frac{\log\|\boldsymbol{n}\|}{r(\boldsymbol{n})}$. we have $\theta_m \ll \log\|\boldsymbol{n}\|$ by (48). Recalling our assumptions that $w(\boldsymbol{n}) \geqslant \|\boldsymbol{n}\|^{1-\sigma}$ and $\|\boldsymbol{n}\|^{1+o(1)} = n_* \gg \frac{w(\boldsymbol{n})}{\log\|\boldsymbol{n}\|}$, we find that the following sum is dominated by the first term:

$$\sum_{m=2}^{m_0-1} \frac{1}{m!}\left(\frac{5(\log(w(\boldsymbol{n})))^2\|\boldsymbol{n}\|e^{\theta_m}}{(w(\boldsymbol{n}))^2}\right)^m = \left(\tfrac{1}{2}+o(1)\right)\left(\frac{\|\boldsymbol{n}\|e^{o(\log\|\boldsymbol{n}\|)}}{(w(\boldsymbol{n}))^2}\right)^2 \leqslant \|\boldsymbol{n}\|^{4\sigma-2+o(1)}. \tag{51}$$

Putting (50) and (51) in (46), we obtain that

$$\frac{\Delta}{(\mathbb{E}\,|\mathcal{I}_\nu(\boldsymbol{G})|)^2} \leqslant \|\boldsymbol{n}\|^{2-4\sigma+o(1)} + \|\boldsymbol{n}\|^{-\omega(1)} = \|\boldsymbol{n}\|^{2-4\sigma+o(1)}.$$

Recalling from (44) that $\mathbb{E}\,|\mathcal{I}_\nu(\boldsymbol{G})| = \|\boldsymbol{n}\|^{\omega(1)}$, we conclude that

$$\frac{\Delta + \mathbb{E}\,|\mathcal{I}_\nu(\boldsymbol{G})|}{(\mathbb{E}\,|\mathcal{I}_\nu(\boldsymbol{G})|)^2} \leqslant \|\boldsymbol{n}\|^{2-4\sigma+o(1)}.$$

Applying (45), we complete the proof. □

4.3 Concentration of the weighted independence number

The estimates of Sections 4.1 and 4.2 lead to the following result.

Theorem 19. *Let $\boldsymbol{G} \sim \mathcal{G}(\boldsymbol{n}, P)$, where $P = (p_{ij})_{i,j\in[k]}$ is such that $p_{ij} = p_{ji}$ and $0 \leqslant p_{ij} < 1$ for all $i, j \in [k]$. Let $Q = Q(P)$ be as in (3). Let $\sigma \in [0, \sigma_0]$ for some fixed $0 < \sigma_0 < \frac{1}{2}$. Assume that (8), (9) hold and*

$$n_* = \|\boldsymbol{n}\|^{1+o(1)}, \qquad n_* \gg \frac{w(\boldsymbol{n})}{\log\|\boldsymbol{n}\|},$$

where $n_* := \min_{i\in[k]} n_i$. Then, whp

$$\alpha_h(\boldsymbol{G}) = (1-\sigma+o(1))\log\|\boldsymbol{n}\|.$$

Proof. All the assumptions of Theorem 8 also present in this theorem, so we can use the formulas and arguments given in its proof. Using (36) and the assumption $q^* = \|\boldsymbol{n}\|^{-\sigma+o(1)}$ by (8), we find that whp

$$\alpha_h(\boldsymbol{G}) \leqslant \log(q^*\|\boldsymbol{n}\|) = (1-\sigma+o(1))\log\|\boldsymbol{n}\|.$$

Next, using Theorem 10(d) and the assumptions $\hat{q}(\boldsymbol{n}), q^* = \|\boldsymbol{n}\|^{-\sigma+o(1)}$ by (8), we have that

$$w(\boldsymbol{n}) \geqslant w_*(\boldsymbol{n}) \geqslant \frac{(\hat{q}(\boldsymbol{n}))^2}{kq^*}\|\boldsymbol{n}\| = \|\boldsymbol{n}\|^{1-\sigma+o(1)}.$$

Thus, all assumptions of Lemma 18 hold. Applying Lemma 18, we find that whp $\mathcal{I}_\nu(\boldsymbol{G}) \neq \emptyset$ for some $\nu = (2+o(1))\frac{\log(w(\boldsymbol{n}))}{w(\boldsymbol{n})}$. If $U \in \mathcal{I}_\nu(\boldsymbol{G})$ then for all $i \in [k]$

$$b_i(U) = \lfloor \nu n_i \rfloor \geqslant \left(\nu - \frac{1}{n_*}\right) n_i.$$

Combining the above, the monotonicity property in Theorem 10(a), and (8), we get that

$$w(\boldsymbol{b}(U)) \geqslant (\nu - \frac{1}{n_*})w(\boldsymbol{n}) = (2 + o(1)) \log(w(\boldsymbol{n})) \geqslant (2 - 2\sigma + o(1)) \log \|\boldsymbol{n}\|.$$

Thus, using (37) and the arguments below (37) showing that $q^* \ll \log \|\boldsymbol{n}\|$, we find that, whp

$$\alpha_h(\boldsymbol{G}) \geqslant \tfrac{1}{2} \left(w(\boldsymbol{b}(U)) - q^* \right) = (1 - \sigma + o(1)) \log \|\boldsymbol{n}\|.$$

This completes the proof. □

We note that the proof of our main result, Theorem 5, does not rely on Theorem 19, but the study of the distribution of the parameter $\alpha_h(\boldsymbol{G})$ is of independent interest. Observe that definition (32) extends to any random graph model. We believe that Theorem 19 carries over as well. In particular, we conjecture the following.

Conjecture 20. Let $\boldsymbol{G} = \boldsymbol{G}(n)$ be a random graph on vertex set $[n]$ where edges ij appear independently of each other with probabilities $p_{ij} = p_{ij}(n) \in (0, 1)$. Assume that there exist $q = q(n)$ and constants $c_1, c_2 > 0$ such that

$$qn \to \infty, \qquad q \ll \log n,$$

as $n \to \infty$, and, for all edges ij,

$$e^{-c_1 q} \leqslant 1 - p_{ij} \leqslant e^{-c_2 q}.$$

Then, whp

$$\alpha_h(\boldsymbol{G}) = (1 + o(1)) \log(qn).$$

In fact, proceeding from Lemma 17 similarly to (36), one can derive that $\alpha_h(\boldsymbol{G}) \leqslant \log(qn)$ whp under the assumptions of Conjecture 20. However, proving the counterpart would require significant modifications of the arguments given in Section 4.2.

5 Crude upper bound

In this section, we establish a crude upper bound on $\chi(\boldsymbol{G})$, where $\boldsymbol{G} \sim \mathcal{G}(\boldsymbol{n}, P)$ based on a simple idea of colouring each block separately. To do so, we only need the results for the classical binomial random graph $\boldsymbol{G}(n, p)$.

Lemma 21. *Let $\sigma \in [0, \sigma_0]$ for some fixed $0 < \sigma_0 < \frac{1}{4}$ and $p = p(n) \in (0, 1)$ is such that*

$$\log n \gg q := \log \tfrac{1}{1-p} \geqslant n^{-\sigma}.$$

Then, for any $\varepsilon > 0$ and any $s = n^{1+o(1)}$, with probability at least $1 - \exp\left(-n^{2-4\sigma+o(1)}\right)$, there is a colouring of $\boldsymbol{G}(n, p)$ with at least $n - s$ vertices using at most $(1 + \varepsilon) \frac{qn}{2 \log(qn)}$ colours.

Proof. This argument is well known for a constant $p \in (0,1)$; see for example, [14, Section 7.4]. For the sake of completeness, we repeat it here and check that it extends to $p = p(n)$ satisfying the assumptions of Lemma 21.

We will apply Lemma 18 to subgraphs $\boldsymbol{G}(n', p)$ of $\boldsymbol{G}(n,p)$ with $n' \geqslant s$, by setting $k = 1$, $\boldsymbol{n} = (n')$, and $P = (p)$. Then, by definition, we have $n_* = \|\boldsymbol{n}\| = n'$ and $w(\boldsymbol{n}) = qn'$ so all assumptions of Lemma 18 hold. Using Lemma 18, we show that the probability that there is a subgraph in $\boldsymbol{G}(n,p)$ with at least s vertices without an independent set of size $(2-\varepsilon)\frac{\log(qn)}{q}$ is at most

$$2^n \exp(-s^{2-4\sigma+o(1)}) = \exp\left(-n^{2-4\sigma+o(1)}\right).$$

Thus, we can keep colouring such independent sets and deleting them from the graph until we are left with fewer than s vertices. The number of colours used in this process is bounded above by

$$\frac{n}{(2-\varepsilon)\frac{\log(qn)}{q}} \leqslant \left(\frac{1}{2} + \varepsilon\right) \frac{qn}{\log(qn)}.$$

Note that, in the above, we can assume that $\varepsilon < 1$ since the statement of the lemma becomes stronger. Then, the inequality $\frac{1}{2-\varepsilon} \leqslant \frac{1+\varepsilon}{2}$ holds. □

To colour the remaining vertices, we use the following lemma.

Lemma 22. *Let $\boldsymbol{G} = \boldsymbol{G}(n,p)$, where $p = p(n) \in (0,1)$ is such that $pn \to \infty$ as $n \to \infty$. Then, for any positive integer $s \geqslant p^{-1}$, we have*

$$\Pr\Big(\exists\ W \subseteq V(\boldsymbol{G}) \,:\, |W| = s \text{ and } \chi(\boldsymbol{G}[W]) \geqslant ps\log n + 1\Big) \leqslant \exp\left(-\omega(p^2 s^2 \log^2 n)\right).$$

Proof. First, for any postive integer $u \leqslant s$, we estimate the probability of the event that the minimal degree of $\boldsymbol{G}' = \boldsymbol{G}(u,p)$ is at least $ps\log n$. This event implies that $\boldsymbol{G}'$ has at least $\frac{1}{2}psu\log n$ edges. Let $N_u := \binom{u}{2}$. Since the distribution of the number of edges in $\boldsymbol{G}'$ is $\mathrm{Bin}(N_u, p)$, we find that

$$\Pr\left(\boldsymbol{G}' \text{ has at least } \tfrac{1}{2}psu\log n \text{ edges}\right) = \sum_{i \geqslant \frac{1}{2}psu\log n} \binom{N_u}{i} p^i (1-p)^{N_u - i}$$
$$\leqslant \sum_{i \geqslant \frac{1}{2}psu\log n} \left(\frac{epN_u}{i}\right)^i \leqslant 2\left(\frac{e}{\log n}\right)^{\frac{1}{2}psu\log n} = n^{-\omega(psu)}.$$

To derive the last inequality in the above, we observe that

$$\frac{pN_u}{i} \leqslant \frac{pu^2}{psu\log n} \leqslant \frac{1}{\log n}.$$

Note also that if $u \leqslant ps\log n$ then $\boldsymbol{G}'$ has less than $\frac{1}{2}psu\log n$ edges with probability 1.

Using the union bound over all choices for $W \subseteq V(\boldsymbol{G})$ with $|W| = s$, for u such that $ps \log n < u \leqslant s$, and for $U \subseteq W$ with $|U| = u$, we get that

$$\begin{aligned}\Pr\left(\exists W \subseteq V(\boldsymbol{G}) \,:\, |W| = s \text{ and } \max_{U \subseteq W} \delta_{\boldsymbol{G}}(U) \geqslant ps \log n\right) \\ \leqslant 2\binom{n}{s} \sum_{u > ps \log n} \binom{s}{u} n^{-\omega(psu)} \\ \leqslant 2\left(\frac{en}{s}\right)^s \sum_{u > ps \log n} \left(\frac{es}{u} n^{-\omega(ps)}\right)^u = \exp\left(-\omega(p^2 s^2 \log^2 n)\right).\end{aligned}$$

Combining this and the upper bound (22) on the chromatic number in terms of the minimal degree of subgraphs, we complete the proof. □

Combining Lemma 21 and Lemma 22, we get the following result for $\boldsymbol{G} \sim \mathcal{G}(\boldsymbol{n}, P)$. This result will be important in the proof of Theorem 9 to show that the number of colours required for the remaining vertices (given by a set $U \subseteq V(\boldsymbol{G})$) after a certain "optimal" colouring process is negligible.

Theorem 23. *Let $\boldsymbol{G} \sim \mathcal{G}(\boldsymbol{n}, P)$, where $P = (p_{ij})_{i,j \in [k]}$ is such that $p_{ij} = p_{ji}$ and $0 \leqslant p_{ij} < 1$ for all $i, j \in [k]$. Let q^* and $\hat{q}$ be defined in (7). Let $\sigma \in [0, \sigma_0]$ for some fixed $0 < \sigma_0 < \frac{1}{4}$. Assume that n is such that*

$$n \geqslant \|\boldsymbol{n}\|^{1+o(1)}, \qquad k = n^{o(1)}, \qquad \log n \gg q^* \geqslant n^{-\sigma}.$$

Let $\boldsymbol{u} = (u_1, \ldots, u_k)^T \in \mathbb{R}_+^k$ be such that $\hat{q}(\boldsymbol{u}) = q^* n^{o(1)}$. Then, for any $\varepsilon > 0$,

$$\Pr\left(\max_{\boldsymbol{b}(U) \leqslant \boldsymbol{u}} \chi(\boldsymbol{G}[U]) > (1+\varepsilon)\, \frac{\hat{q}(\boldsymbol{u}) \|\boldsymbol{u}\|}{2 \log (q^* n)}\right) \leqslant \exp\left(-n^{2 - 4\sigma + o(1)}\right),$$

where the maximum is over all subsets $U \subseteq V(\boldsymbol{G})$ such that $|U \cap B_i| \leqslant u_i$ for all $i \in [k]$.

Proof. For each $i \in [k]$, we let $\boldsymbol{G}_i = \boldsymbol{G}(n_i, p_{ii})$ denote the induced subgraph of $\boldsymbol{G}[B_i]$. Let

$$n_i' := n_i + s, \qquad \text{where } s := \left\lfloor \frac{\hat{q}(\boldsymbol{u}) \|\boldsymbol{u}\|}{k q^* \log^3 \|\boldsymbol{n}\|} \right\rfloor.$$

Define $p_{ii}' \in (0, 1)$ to be such that

$$q_{ii}' := \log \frac{1}{1 - p_{ii}'} = q_{ii} + q_0, \qquad \text{where } q_0 := \frac{\hat{q}(\boldsymbol{u})}{k \log \|\boldsymbol{n}\|}.$$

Since $\log n \gg q^* \geqslant n^{-\sigma}$ and $\hat{q}(\boldsymbol{u}) = q^* n^{o(1)}$, we get that

$$\log n \gg q_{ii} + o(q^*) \geqslant q_{ii}' \geqslant q_0 \geqslant n^{-\sigma + o(1)}. \tag{52}$$

By adding s dummy vertices to each block B_i and introducing some rejection probability, for each $i \in [k]$ there is a coupling $(\boldsymbol{G}_i', \boldsymbol{G}_i)$ such that $\boldsymbol{G}_i' = \boldsymbol{G}(n_i', p_{ii}')$ and $\boldsymbol{G}_i$ is a subgraph of $\boldsymbol{G}_i'$ with probability 1.

Consider any $U_i \subseteq V(\boldsymbol{G}_i)$ such that $|U_i| \leqslant u_i$ and let U_i' consist of the union of U_i and s dummy vertices of $\boldsymbol{G}_i'$. Note that, by assumptions,

$$n^{1+o(1)} \leqslant s \leqslant u_i' := |U_i'| \leqslant u_i + s.$$

Using (52), we find that $sq_0 = q^*\|\boldsymbol{u}\|n^{o(1)}$ and

$$\begin{aligned}(u_i+s)(q_{ii}+q_0) &= u_i q_{ii} + o\left(\frac{\hat{q}(\boldsymbol{u})\|\boldsymbol{u}\|n^{o(1)}}{k}\right),\\ (u_i+s)(q_{ii}+q_0) &\leqslant (1+o(1))q^*\|\boldsymbol{u}\|.\end{aligned} \tag{53}$$

Since $q^* \geqslant n^{-\sigma}$, we get that

$$\log(q_{ii}'u_i') = (1+o(1))\log(q^*n).$$

Using (52), we find that the assumptions of Lemma 21 hold for $\boldsymbol{G}_i[U_i'] = \boldsymbol{G}(u_i', p_{ii}')$. Applying Lemma 21 with $\varepsilon' = \frac{\varepsilon}{2}$, we find that there is a colouring of $u_i' - s = u_i$ vertices of $\boldsymbol{G}_i'$ using at most

$$(1+\varepsilon')\frac{q_{ii}'u_i'}{2\log(q_{ii}'u_i')} \leqslant \left(1+\frac{\varepsilon}{2}+o(1)\right)\frac{q_{ii}'u_i'}{2\log(q^*n)}$$

colours with probability at least

$$1-\exp\left(-(u_i')^{2-4\sigma+o(1)}\right) \geqslant 1-\exp\left(-n^{2-4\sigma+o(1)}\right).$$

Recalling that $2-4\sigma > 1$ and applying the union bound, we get that the probability that there are some $i \in [k]$ and $U_i \subseteq V(\boldsymbol{G}_i)$ with $|U_i| \leqslant u_i$ for which such colouring does not exist is bounded above by

$$\sum_{i=1}^{k} 2^{n_i+s}\exp\left(-n^{2-4\sigma+o(1)}\right) = \exp\left(-n^{2-4\sigma+o(1)}\right).$$

Next, we show that only a small number of colours is needed to colour the remaining s vertices from each $V(\boldsymbol{G}_i')$. Applying Lemma 22, we get that any subset $W \subseteq V(\boldsymbol{G}_i')$ with $s \geqslant n^{1+o(1)} \geqslant (p_{ii}')^{-1}$ vertices can be coloured using at most

$$p_{ii}'s\log n_i' + 1 \leqslant q^*s\log n_i' + 1 \ll \frac{\hat{q}(\boldsymbol{u})\|\boldsymbol{u}\|}{k\log(q^*n)}$$

colours with probability at least

$$1-\exp\left(-\omega((p_{ii}')^2s^2\log^2 n_i')\right) \geqslant 1-\exp\left(-n^{2-2\sigma+o(1)}\right).$$

The last inequality is clear for $p_{ii}' \geqslant \frac{1}{2}$. For $p_{ii}' < \frac{1}{2}$, one can use (52) together with the inequality $p_{ii}' \geqslant \frac{q_{ii}'}{2\log 2}$, which follows from the fact that $t^{-1}\log\frac{1}{1-t}$ is monotonically increasing for $t \in (0,1)$. Applying the union bound, we can complete the colouring of all sets U_i' for $i \in [k]$ using at most

$$\sum_{i\in[k]}(p_{ii}'s\log n_i' + 1) \ll \frac{\hat{q}(\boldsymbol{u})\|\boldsymbol{u}\|}{\log(q^* n)}$$

colours with probability at least

$$1 - k\exp\left(-n^{2-2\sigma+o(1)}\right) \geqslant 1 - \exp\left(-n^{2-4\sigma+o(1)}\right).$$

Now, consider any $U \subseteq V(\boldsymbol{G})$ such that $|U \cap B_i| \leqslant u_i$ for all $i \in [k]$. Combining the above bounds and using the inequality in the second line of (53), we get that U can be coloured with at most

$$\begin{aligned}
&\sum_{i\in[k]}\left(1+\frac{\varepsilon}{2}+o(1)\right)\frac{q_{ii}'u_i'}{2\log(q^*n)} + \sum_{i\in[k]}(p_{ii}'s\log n_i' + 1)\\
&\quad = \left(1+\frac{\varepsilon}{2}+o(1)\right)\frac{\sum_{i\in k}\left(u_i q_{ii} + o\left(\frac{\hat{q}(\boldsymbol{u})\|\boldsymbol{u}\|}{k}\right)\right)}{2\log(q^*n)} + o\left(\frac{\hat{q}(\boldsymbol{u})\|\boldsymbol{u}\|}{\log(q^*n)}\right)\\
&\quad \leqslant (1+\varepsilon)\frac{\hat{q}(\boldsymbol{u})\|\boldsymbol{u}\|}{2\log(q^*n)}
\end{aligned}$$

colours with probability at least $1 - \exp\left(-n^{2-4\sigma+o(1)}\right)$. □

6 Optimal colouring: proof of Theorem 9

In this section we prove Theorem 9. First, applying Lemma 18 multiple times, we find there are approximately $\frac{w(\boldsymbol{n})}{2\log(q^*\|\boldsymbol{n}\|)}$ independent sets covering almost all vertices of $\boldsymbol{G} \sim \mathcal{G}(\boldsymbol{n}, P)$. Then, we use Theorem 23 to estimate the number of colours for the remaining vertices, proving that

$$\chi(\boldsymbol{G}) \leqslant (1+o(1))\frac{w(\boldsymbol{n})}{2\log(q^*\|\boldsymbol{n}\|)} + O\left(\frac{k\hat{q}(\boldsymbol{n})q^*\|\boldsymbol{n}\|}{\log^2(q^*\|\boldsymbol{n}\|)}\right)$$

with probability sufficiently close to 1. Finally, we obtain Theorem 9 by applying this upper bound to each random graph corresponding to an optimal system of k vectors $(\boldsymbol{x}^{(t)})_{t\in[k]}$ from $\mathbb{R}_+^k$ such that

$$\boldsymbol{n} = \sum\nolimits_{t\in[k]}\boldsymbol{x}^{(t)} \qquad \text{and} \qquad w_*(\boldsymbol{n}) = \sum\nolimits_{t\in[k]} w(\boldsymbol{x}^{(t)}).$$

Everywhere in this section, we use notations $w(\cdot)$ and $w_*(\cdot)$ in place of $w(\cdot, Q)$ and $w_*(\cdot, Q)$, where $Q = Q(P)$ is the matrix defined by (3), and $q^*, \hat{q}(\cdot)$ are the same as in (7).

6.1 Covering by independent sets

Recall that, for $\boldsymbol{x} = (x_1, \dots, x_k)^T \in \mathbb{R}_+^k$, we defined

$$\lfloor \boldsymbol{x} \rfloor := (\lfloor x_1 \rfloor, \dots, \lfloor x_k \rfloor)^T \qquad \text{and} \qquad x_* := \min_{i \in [k]} x_i.$$

Provided $x_* > 0$, we have, for any $s > 0$,

$$s\|\boldsymbol{x}\| \geqslant \|\lfloor s\boldsymbol{x} \rfloor\| \ \geqslant \left(s - \frac{1}{x_*}\right) \|\boldsymbol{x}\|. \tag{54}$$

Similarly, using the monotonicity and scaling properties in Theorem 10(a), we get that

$$sw(\boldsymbol{x}) \geqslant w(\lfloor s\boldsymbol{x} \rfloor) \geqslant \left(s - \frac{1}{x_*}\right) w(\boldsymbol{x}). \tag{55}$$

The next lemma shows that we can cover almost all vertices of a random graph from the stochastic block model with the large "balanced" independent sets provided by Lemma 18.

Lemma 24. *Let $\boldsymbol{G} \sim \mathcal{G}(\boldsymbol{n}, P)$, where $P = (p_{ij})_{i,j \in [k]}$ is such that $p_{ij} = p_{ji}$ and $0 \leqslant p_{ij} < 1$ for all $i, j \in [k]$. Let $\sigma \in [0, \sigma_0]$ for some fixed $0 < \sigma_0 < \frac{1}{4}$. Assume that $w(\boldsymbol{n})$ and n_* satisfy the following as $\|\boldsymbol{n}\| \to \infty$:*

$$w(\boldsymbol{n}) \geqslant \|\boldsymbol{n}\|^{1-\sigma}, \qquad n_* = \|\boldsymbol{n}\|^{1+o(1)}, \qquad n_* \gg \frac{w(\boldsymbol{n})}{\log \|\boldsymbol{n}\|}.$$

Then, for any fixed constant $\varepsilon \in (0, 1)$, with probability at least

$$1 - \exp\left(-\|\boldsymbol{n}\|^{2-4\sigma+o(1)}\right),$$

there is a colouring of $\boldsymbol{G}$ with at most $(1+\varepsilon)\frac{w(\boldsymbol{n})}{2\log(w(\boldsymbol{n}))}$ colours covering at least

$$n_i \left(1 - \frac{1}{\log^2 \|\boldsymbol{n}\|} - 5\varepsilon^{-1} \frac{w(\boldsymbol{n})}{n_* \log(w(\boldsymbol{n}))}\right)$$

vertices from each block B_i for all $i \in [k]$.

Throughout this section we let $\varepsilon \in (0, 1)$ be fixed and set

$$\nu := (2 - \varepsilon) \frac{\log(w(\boldsymbol{n}))}{w(\boldsymbol{n})} \qquad \text{and} \qquad \theta := \frac{1}{2\log^2 \|\boldsymbol{n}\|} + 3\varepsilon^{-1} \frac{w(\boldsymbol{n})}{n_* \log(w(\boldsymbol{n}))}.$$

To prove Lemma 24 we first claim some auxiliary results, whose proofs we defer to the end of this section.

Claim 25. *With probability at least*

$$1 - \exp\left(-\|\boldsymbol{n}\|^{2-4\sigma+o(1)}\right)$$

there exists a sequence $(U_1, \dots, U_\ell)$ of disjoint independent sets in $\boldsymbol{G}$ satisfying the following.

(i) *For all $j \in [\ell]$, we have*

$$\boldsymbol{b}(U_j) = \left\lfloor \frac{\nu\|\boldsymbol{n}\|}{\|\boldsymbol{n}^{(j)}\|}\boldsymbol{n}^{(j)} \right\rfloor,$$

where

$$\boldsymbol{n}^{(j)} = (n_1^{(j)}, \dots, n_k^{(j)})^T := \boldsymbol{n} - \sum_{i=1}^{j-1} \boldsymbol{b}(U_i) \tag{56}$$

satisfies $\|\boldsymbol{n}^{(j)}\| \geqslant \theta\|\boldsymbol{n}\|$.

(ii) *The set $\bigcup_{j=1}^{\ell} U_j$ covers all but at most $\theta e^{\frac{1}{2}} n_i$ vertices from each block B_i. That is, for all $i \in [k]$, we have $\sum_{j=1}^{\ell} b_i(U_j) \geqslant \left(1 - \theta e^{\frac{1}{2}}\right) n_i$.*

Our next claim gives the upper bound on the length of the sequence $(U_1, \dots, U_\ell)$ of disjoint independent sets in $\boldsymbol{G}$ from Claim 25.

Claim 26. *Suppose there exists a sequence $(U_1, \dots, U_\ell)$ of disjoint independent subsets in $\boldsymbol{G}$ such that condition (i) of Claim 25 holds. Then*

$$\ell \;\leqslant\; (1+\varepsilon)\,\frac{w(\boldsymbol{n})}{2\log(w(\boldsymbol{n}))}.$$

We are ready to establish Lemma 24 based on the claims given above.

Proof of Lemma 24. We take the independent sets $U_1, \dots, U_\ell$ provided by Claim 25 as our colour classes. Note that $\frac{1}{2}e^{\frac{1}{2}} \leqslant 1$ and $3\, e^{\frac{1}{2}} \leqslant 5$ so the condition (ii) of Claim 25 ensures that this colouring covers all but at most $n_i \left(\frac{1}{\log^2 \|\boldsymbol{n}\|} + 5\varepsilon^{-1}\frac{w(\boldsymbol{n})}{n_* \log(w(\boldsymbol{n}))}\right)$ vertices from each block B_i. Claim 26 establishes the upper bound for the number of colours as desired. $\square$

In the rest of this section we will prove first Claim 26 and then Claim 25. To this end we need the following lower bounds on $n_*^{(j)}$ defined by

$$n_*^{(j)} := \min_{i\in[k]} n_i^{(j)}.$$

Claim 27. *Suppose there exists a sequence $(U_1, \dots, U_\ell)$ of disjoint independent subsets in $\boldsymbol{G}$ such that condition (i) of Claim 25 holds. Then, for all $j \in [\ell]$ such that $j \leqslant \frac{w(\boldsymbol{n})}{\log(w(\boldsymbol{n}))}$, we have*

$$n_*^{(j)} \geqslant \left(1 - \tfrac{\varepsilon}{3}\right) n_* \frac{\|\boldsymbol{n}^{(j)}\|}{\|\boldsymbol{n}\|}.$$

Proof of Claim 27. It is sufficient to prove that, for all $j \in [\ell]$ such that $j \leqslant \frac{w(\boldsymbol{n})}{\log(w(\boldsymbol{n}))}$,

$$n_*^{(j)} \geqslant n_* \frac{\|\boldsymbol{n}^{(j)}\|}{\|\boldsymbol{n}\|} - j + 1. \tag{57}$$

Indeed, by the condition (i) of Claim 25 we have $\|\boldsymbol{n}^{(j)}\| \geqslant \theta\|\boldsymbol{n}\|$. From the definition of θ, we get that, for every $j \in [\ell]$,

$$\|\boldsymbol{n}^{(j)}\| \geqslant \theta\|\boldsymbol{n}\| = \frac{\|\boldsymbol{n}\|}{2\log^2\|\boldsymbol{n}\|} + 3\varepsilon^{-1}\frac{\|\boldsymbol{n}\|w(\boldsymbol{n})}{n_*\log(w(\boldsymbol{n}))}. \tag{58}$$

Using the trivial bound $\|\boldsymbol{n}\| \geqslant n_*$, we immediately get from (58) that

$$n_*\frac{\|\boldsymbol{n}^{(j)}\|}{\|\boldsymbol{n}\|} \geqslant 3\varepsilon^{-1}\frac{w(\boldsymbol{n})}{\log(w(\boldsymbol{n}))}. \tag{59}$$

Thus, if $j \leqslant \frac{w(\boldsymbol{n})}{\log(w(\boldsymbol{n}))}$ then Claim 27 follows from (57) and (59).

We will prove (57) by induction on j. Clearly, it is true for $j = 1$ since $\boldsymbol{n}^{(1)} = \boldsymbol{n}$. Suppose, we established the claim for $j = i$ such that $i < \ell$. By definition and using (54) with $\boldsymbol{x} = \boldsymbol{n}^{(i)}$ and $s = \frac{\nu\|\boldsymbol{n}\|}{\|\boldsymbol{n}^{(i)}\|}$, we get that

$$\frac{\|\boldsymbol{n}^{(i+1)}\|}{\|\boldsymbol{n}^{(i)}\|} \leqslant 1 - \frac{\nu\|\boldsymbol{n}\|}{\|\boldsymbol{n}^{(i)}\|} + \frac{1}{n_*^{(i)}}. \tag{60}$$

Combining the induction hypothesis, the bound of (60), and $\frac{\|\boldsymbol{n}^{(i+1)}\|}{\|\boldsymbol{n}^{(i)}\|} \leqslant 1$, we find that

$$\begin{aligned} n_*^{(i+1)} &\geqslant \left(1 - \frac{\nu\|\boldsymbol{n}\|}{\|\boldsymbol{n}^{(i)}\|}\right) n_*^{(i)} \geqslant \left(\frac{\|\boldsymbol{n}^{(i+1)}\|}{\|\boldsymbol{n}^{(i)}\|} - \frac{1}{n_*^{(i)}}\right) n_*^{(i)} \\ &\geqslant n_*\frac{\|\boldsymbol{n}^{(i+1)}\|}{\|\boldsymbol{n}\|} - (i-1)\frac{\|\boldsymbol{n}^{(i+1)}\|}{\|\boldsymbol{n}^{(i)}\|} - 1 \geqslant n_*\frac{\|\boldsymbol{n}^{(i+1)}\|}{\|\boldsymbol{n}\|} - i. \end{aligned}$$

Note also that the induction hypothesis and (59) imply that $n_*^{(i)}$ is positive, since it is at least $\left(1 - \frac{\varepsilon}{3}\right) n_*\frac{\|\boldsymbol{n}^{(i)}\|}{\|\boldsymbol{n}\|}$. Thus, the claim is true for $j = i+1$ and, by induction, for all $j \in [\ell]$ such that $j \leqslant \frac{w(\boldsymbol{n})}{\log(w(\boldsymbol{n}))}$. □

Proof of Claim 26. Assume otherwise that $\ell > (1+\varepsilon)\,\frac{w(\boldsymbol{n})}{2\log(w(\boldsymbol{n}))}$. By definition (24) we have $|U_j| = \|\boldsymbol{b}(U_j)\|$ and by Claim 25(i) and (54) with $\boldsymbol{x} = \boldsymbol{n}^{(j)}$ and $s = \frac{\nu\|\boldsymbol{n}\|}{\|\boldsymbol{n}^{(j)}\|}$ we have

$$\|\boldsymbol{b}(U_j)\| \geqslant \left(\frac{\nu\|\boldsymbol{n}\|}{\|\boldsymbol{n}^{(j)}\|} - \frac{1}{n_*^{(j)}}\right)\|\boldsymbol{n}^{(j)}\|.$$

Note further that, by definition of ν and the assumptions $n_* \gg \frac{w(\boldsymbol{n})}{\log\|\boldsymbol{n}\|}$, $w(\boldsymbol{n}) \geqslant \|\boldsymbol{n}\|^{1-\sigma}$, we have $\nu \gg 1/n_*$. Then, using Claim 27, we get that $\frac{\nu\|\boldsymbol{n}\|}{\|\boldsymbol{n}^{(j)}\|} \gg \frac{1}{n_*^{(j)}}$ for all $j \in [\ell]$ such that $j \leqslant \frac{w(\boldsymbol{n})}{\log(w(\boldsymbol{n}))}$. Therefore,

$$|U_j| = \|\boldsymbol{b}(U_j)\| \geqslant \left(\frac{\nu\|\boldsymbol{n}\|}{\|\boldsymbol{n}^{(j)}\|} - \frac{1}{n_*^{(j)}}\right)\|\boldsymbol{n}^{(j)}\| = (1-o(1))\nu\|\boldsymbol{n}\|. \tag{61}$$

Using (61) and our assumption that $\ell > (1+\varepsilon)\frac{w(\boldsymbol{n})}{2\log(w(\boldsymbol{n}))}$, we get

$$\begin{aligned}\|\boldsymbol{n}\| \geqslant \sum_{j\in[\ell]} |U_j| &\geqslant (1-o(1))\nu\|\boldsymbol{n}\|\,(1+\varepsilon)\,\frac{w(\boldsymbol{n})}{2\log(w(\boldsymbol{n}))}\\ &= (1-o(1))(2-\varepsilon)\,(1+\varepsilon)\,\frac{\|\boldsymbol{n}\|}{2} > \|\boldsymbol{n}\|.\end{aligned}$$

The last inequality is true for any fixed $\varepsilon \in (0,1)$ when $o(1)$ gets sufficiently small. This contradiction proves Claim 26. □

Proof of Claim 25. In order to show the existence of such a sequence $(U_1,\dots,U_\ell)$ of disjoint independent sets in $\boldsymbol{G}$, we repeatedly apply Lemma 18. Suppose we already constructed sets $U_1,\dots,U_{j-1}$. We will show that if $\|\boldsymbol{n}^{(j)}\| \geqslant \theta\|\boldsymbol{n}\|$ then, with sufficiently high probability, we can find another independent set U_j in the induced subgraph of $\boldsymbol{G}$ on remaining vertices, which satisfies condition (i) of Claim 25. By Claim 26, we get that

$$j \leqslant (1+\varepsilon)\,\frac{w(\boldsymbol{n})}{2\log(w(\boldsymbol{n}))} \leqslant \frac{w(\boldsymbol{n})}{\log(w(\boldsymbol{n}))}. \tag{62}$$

For all $i<j$, using (54) with $\boldsymbol{x} = \boldsymbol{n}^{(i)}$ and $s = \frac{\nu\|\boldsymbol{n}\|}{\|\boldsymbol{n}^{(i)}\|}$, we find that

$$\frac{\|\boldsymbol{n}^{(i+1)}\|}{\|\boldsymbol{n}^{(i)}\|} \geqslant 1 - \frac{\nu\|\boldsymbol{n}\|}{\|\boldsymbol{n}^{(i)}\|}. \tag{63}$$

From (55) (with the same $\boldsymbol{x}$ and s), we obtain

$$w(\boldsymbol{n}^{(i+1)}) \leqslant \left(1 - \frac{\nu\|\boldsymbol{n}\|}{\|\boldsymbol{n}^{(i)}\|} + \frac{1}{n_*^{(i)}}\right) w(\boldsymbol{n}^{(i)}). \tag{64}$$

Due to (62), we can apply Claim 27 to obtain

$$n_*^{(i)} \geqslant \left(1-\frac{\varepsilon}{3}\right) n_* \frac{\|\boldsymbol{n}^{(i)}\|}{\|\boldsymbol{n}\|} \geqslant \frac{3-\varepsilon}{\varepsilon}\,\frac{w(\boldsymbol{n})}{\log(w(\boldsymbol{n}))}\,\frac{\|\boldsymbol{n}^{(i)}\|}{\|\boldsymbol{n}^{(i+1)}\|},$$

where the last inequality follows from (59) by taking $j = i+1$ which gives

$$\frac{n_*}{\|\boldsymbol{n}\|} \geqslant \frac{3}{\varepsilon}\,\frac{w(\boldsymbol{n})}{\log(w(\boldsymbol{n}))}\,\frac{1}{\|\boldsymbol{n}^{(i+1)}\|}.$$

Then, combining this with (63), we have

$$1 - \frac{\nu\|\boldsymbol{n}\|}{\|\boldsymbol{n}^{(i)}\|} + \frac{1}{n_*^{(i)}} \leqslant \frac{\|\boldsymbol{n}^{(i+1)}\|}{\|\boldsymbol{n}^{(i)}\|}\left(1 + \frac{\varepsilon\log(w(\boldsymbol{n}))}{(3-\varepsilon)w(\boldsymbol{n})}\right),$$

which in (64) implies

$$\frac{w(\boldsymbol{n}^{(i+1)})}{w(\boldsymbol{n}^{(i)})} \leqslant \frac{\|\boldsymbol{n}^{(i+1)}\|}{\|\boldsymbol{n}^{(i)}\|}\left(1 + \frac{\varepsilon\log(w(\boldsymbol{n}))}{(3-\varepsilon)w(\boldsymbol{n})}\right). \tag{65}$$

Multiplying (65) together for $i = 1, \dots, j-1$, we obtain

$$\frac{w(\boldsymbol{n}^{(j)})}{w(\boldsymbol{n})} = \prod_{i=1}^{j-1} \frac{w(\boldsymbol{n}^{(i+1)})}{w(\boldsymbol{n}^{(i)})} \leqslant \frac{\|\boldsymbol{n}^{(j)}\|}{\|\boldsymbol{n}\|} \left(1 + \frac{\varepsilon \log(w(\boldsymbol{n}))}{(3-\varepsilon) w(\boldsymbol{n})}\right)^{j-1}$$
$$\leqslant \frac{\|\boldsymbol{n}^{(j)}\|}{\|\boldsymbol{n}\|} \exp\left(\frac{\varepsilon(1+\varepsilon)}{2(3-\varepsilon)}\right) \leqslant \frac{\|\boldsymbol{n}^{(j)}\|}{\|\boldsymbol{n}\|} e^{\frac{\varepsilon}{2}}, \tag{66}$$

where the penultimate inequality is due to $1 + x \leqslant e^x$ and the first inequality in (62). Similarly to (65) and (66), we get

$$\frac{w(\boldsymbol{n}^{(i+1)})}{w(\boldsymbol{n}^{(i)})} \geqslant \left(\frac{\|\boldsymbol{n}^{(i+1)}\|}{\|\boldsymbol{n}^{(i)}\|} - \frac{1}{n_*^{(i)}}\right) \geqslant \frac{\|\boldsymbol{n}^{(i+1)}\|}{\|\boldsymbol{n}^{(i)}\|} \left(1 - \frac{\varepsilon \log(w(\boldsymbol{n}))}{(3-\varepsilon) w(\boldsymbol{n})}\right),$$

which leads to the bound

$$\frac{w(\boldsymbol{n}^{(j)})}{w(\boldsymbol{n})} \geqslant \frac{\|\boldsymbol{n}^{(j)}\|}{\|\boldsymbol{n}\|} e^{-\frac{\varepsilon}{2}}. \tag{67}$$

Thus, we obtain from (66) and (67) that

$$\frac{\|\boldsymbol{n}^{(j)}\|}{\|\boldsymbol{n}\|} e^{-\frac{\varepsilon}{2}} \leqslant \frac{w(\boldsymbol{n}^{(j)})}{w(\boldsymbol{n})} \leqslant \frac{\|\boldsymbol{n}^{(j)}\|}{\|\boldsymbol{n}\|} e^{\frac{\varepsilon}{2}}. \tag{68}$$

From (58), we have $\frac{\|\boldsymbol{n}^{(j)}\|}{\|\boldsymbol{n}\|} \geqslant \theta \geqslant \frac{1}{2\log^2 \|\boldsymbol{n}\|}$ and obviously $\frac{\|\boldsymbol{n}^{(j)}\|}{\|\boldsymbol{n}\|} \leqslant 1$. Using the assumption $n_* = \|\boldsymbol{n}\|^{1+o(1)}$, we get that

$$\log(\|\boldsymbol{n}\|) = (1 + o(1)) \log \|\boldsymbol{n}^{(j)}\|. \tag{69}$$

By Claim 27 and the assumption $n_* \gg \frac{w(\boldsymbol{n})}{\log \|\boldsymbol{n}\|}$ we find that

$$n_*^{(j)} \geqslant \left(1 - \frac{\varepsilon}{3}\right) n_* \frac{\|\boldsymbol{n}^{(j)}\|}{\|\boldsymbol{n}\|} \gg \frac{\|\boldsymbol{n}^{(j)}\|}{\|\boldsymbol{n}\|} \cdot \frac{w(\boldsymbol{n})}{\log \|\boldsymbol{n}\|)} \geqslant e^{-\frac{\varepsilon}{2}} \frac{w(\boldsymbol{n}^{(j)})}{\log \|\boldsymbol{n}\|},$$

where the last inequality follows from (66). Combining this with (69), we get

$$n_*^{(j)} \gg \frac{w(\boldsymbol{n}^{(j)})}{\log \|\boldsymbol{n}^{(j)}\|}.$$

Combining, Claim 27, the assumption $n_* = \|\boldsymbol{n}\|^{1+o(1)}$, and (69), we find that

$$1 \geqslant \frac{n_*^{(j)}}{\|\boldsymbol{n}^{(j)}\|} \geqslant \left(1 - \frac{\varepsilon}{3}\right) \frac{n_*}{\|\boldsymbol{n}\|} = \|\boldsymbol{n}\|^{o(1)} = \|\boldsymbol{n}^{(j)}\|^{o(1)}.$$

Furthermore, the assumption $w(\boldsymbol{n}) \geqslant \|\boldsymbol{n}\|^{1-\sigma}$ together with (68) and (69) implies

$$w(\boldsymbol{n}^{(j)}) \geqslant w(\boldsymbol{n}) \frac{\|\boldsymbol{n}^{(j)}\|}{\|\boldsymbol{n}\|} e^{-\frac{\varepsilon}{2}} \geqslant e^{-\frac{\varepsilon}{2}} \|\boldsymbol{n}\|^{1-\sigma-o(1)} = \|\boldsymbol{n}^{(j)}\|^{1-\sigma-o(1)}.$$

Thus, all assumptions of Lemma 18 hold for the random graph $\boldsymbol{G}^{(j)} \sim \mathcal{G}(\boldsymbol{n}^{(j)}, P)$.

Applying Lemma 18 to $\boldsymbol{G}^{(j)} \sim \mathcal{G}(\boldsymbol{n}^{(j)}, P)$ we show the existence of an independent set $U' \subset V(\boldsymbol{G}^{(j)})$ in $\boldsymbol{G}^{(j)}$ with probability at least $1 - \exp\left(-\|\boldsymbol{n}\|^{2-4\sigma+o(1)}\right)$ such that $\boldsymbol{b}(U') = \lfloor \nu' \boldsymbol{n} \rfloor$, where

$$\nu' = (2 + o(1)) \frac{\log\left(w(\boldsymbol{n}^{(j)})\right)}{w(\boldsymbol{n}^{(j)})}.$$

Moreover, since $2 - 4\sigma > 1$, the probability that there exists $W \subseteq V(\boldsymbol{G})$ such that $\boldsymbol{b}(W) = \boldsymbol{n}^{(j)}$ and $\boldsymbol{G}[W]$ does not contain such an independent set U' is at most

$$\binom{\|\boldsymbol{n}\|}{\|\boldsymbol{n}^{(j)}\|} \exp\left(-\|\boldsymbol{n}\|^{2-4\sigma+o(1)}\right) = \exp\left(-\|\boldsymbol{n}\|^{2-4\sigma+o(1)}\right).$$

In particular, we get that the graph obtained from $\boldsymbol{G}$ by removing $U_1, \ldots, U_{j-1}$ contains such U' with probability at least $1 - \exp\left(-\|\boldsymbol{n}\|^{2-4\sigma+o(1)}\right)$.

Next, we show that it is possible to find $U_j \subseteq U'$ such that $\boldsymbol{b}(U_j) = \left\lfloor \frac{\nu\|\boldsymbol{n}\|}{\|\boldsymbol{n}^{(j)}\|} \boldsymbol{n}^{(j)} \right\rfloor$. To do this, it is sufficient to show that $\nu' \geqslant \frac{\nu\|\boldsymbol{n}\|}{\|\boldsymbol{n}^{(j)}\|}$. Using (68), (69), and the assumption $w(\boldsymbol{n}) \geqslant \|\boldsymbol{n}\|^{1-\sigma}$, we find that

$$\log\left(w(\boldsymbol{n}^{(j)})\right) = (1 + o(1)) \log\left(w(\boldsymbol{n})\right).$$

Observe that $g(\varepsilon) := (2-\varepsilon)e^{\varepsilon/2}$ is decreasing on $\mathbb{R}_+$, so $g(\varepsilon) < g(0) = 2$. Therefore, using (68) and the first inequality in (55), we get that

$$\nu' \geqslant (2-\varepsilon)e^{\varepsilon/2} \frac{\log\left(w(\boldsymbol{n})\right)}{w(\boldsymbol{n}^{(j)})} \geqslant (2-\varepsilon) \frac{\|\boldsymbol{n}\| \log(w(\boldsymbol{n}))}{\|\boldsymbol{n}^{(j)}\| w(\boldsymbol{n})} = \frac{\nu\|\boldsymbol{n}\|}{\|\boldsymbol{n}^{(j)}\|}.$$

The probability that there exists the required sequence $(U_1, \ldots, U_\ell)$ can be estimated as follows. Using Claim 26 and applying the union bound for the event that there is no suitable choice for U_{j+1} after removing $U_1, \ldots, U_j$ from $\boldsymbol{G}$, we get that

$$\begin{aligned}\sum_{j\in[\ell]} \exp\left(-\|\boldsymbol{n}\|^{2-4\sigma+o(1)}\right) &\leqslant \left(\frac{1}{2} + \varepsilon\right) \frac{w(\boldsymbol{n})}{\log(w(\boldsymbol{n}))} \exp\left(-\|\boldsymbol{n}\|^{2-4\sigma+o(1)}\right) \\ &= \exp\left(-\|\boldsymbol{n}\|^{2-4\sigma+o(1)}\right).\end{aligned}$$

To derive the last inequality, we use the assumptions to estimate $\frac{w(\boldsymbol{n})}{\log(w(\boldsymbol{n}))} \ll n_* \leqslant \|\boldsymbol{n}\|$ and recall that $2 - 4\sigma > 1$.

The construction of the sequence $(U_1, \ldots, U_\ell)$ is terminated when $\|\boldsymbol{n}^{(\ell+1)}\| < \theta\|\boldsymbol{n}\|$. Note that, for any $i \in [k]$,

$$n_i^{(j+1)} \leqslant n_i^{(j)} - \frac{\nu\|\boldsymbol{n}\|}{\|\boldsymbol{n}^{(j)}\|} n_i^{(j)} + 1 \leqslant \left(1 - \frac{\nu\|\boldsymbol{n}\|}{\|\boldsymbol{n}^{(j)}\|} + \frac{1}{n_*^{(j)}}\right) n_i^{(j)}.$$

Repeating the arguments of (65) and (66), we find that

$$n_i^{(\ell+1)} \leqslant e^{\frac{\varepsilon}{2}} \frac{\|\boldsymbol{n}^{(\ell+1)}\|}{\|\boldsymbol{n}\|} n_i \leqslant \theta e^{\frac{1}{2}} n_i.$$

Thus, condition (ii) of Claim 25 is satisfied. This completes the proof of Claim 25. □

6.2 Final ingredient for colouring completion

In this section, we combine Lemma 24 and Theorem 23 to estimate the chromatic number of $\boldsymbol{G} \sim \mathcal{G}(\boldsymbol{n}, P)$ under the additional condition that $w_*(\boldsymbol{n})$ is asymptotically equal to $w(\boldsymbol{n})$. In the general case of Theorem 9, this additional condition will be satisfied by each part of the random graph $\boldsymbol{G}$ corresponding to a near-optimal integer system given by Theorem 10(g); see Section 6.3.

Lemma 28. *Let $\boldsymbol{G} \sim \mathcal{G}(\boldsymbol{n}, P)$, where $P = (p_{ij})_{i,j\in[k]}$ is such that $p_{ij} = p_{ji}$ and $0 \leqslant p_{ij} < 1$ for all $i, j \in [k]$. Let $\sigma \in [0, \sigma_0]$ for some fixed $0 < \sigma_0 < \frac{1}{4}$. Assume that, as $\|\boldsymbol{n}\| \to \infty$:*

$$k = \|\boldsymbol{n}\|^{o(1)}, \qquad \log \|\boldsymbol{n}\| \gg kq^* \geqslant \|\boldsymbol{n}\|^{-\sigma}.$$

Assume also that

$$w_*(\boldsymbol{n}) = (1 + o(1))w(\boldsymbol{n}) \geqslant (q^*\|\boldsymbol{n}\|)^{1+o(1)}.$$

Then, for any fixed $\varepsilon \in (0, 1)$,

$$\Pr\left(\chi(\boldsymbol{G}) > (1+\varepsilon)\,\frac{w(\boldsymbol{n})}{2\log(q^*\|\boldsymbol{n}\|)} + 20\varepsilon^{-2}\frac{k\hat{q}(\boldsymbol{n})q^*\|\boldsymbol{n}\|}{\log^2(q^*\|\boldsymbol{n}\|)}\right) \leqslant \exp\left(-\|\boldsymbol{n}\|^{2-4\sigma+o(1)}\right).$$

Proof. Let

$$n_0 := \frac{\varepsilon\, w(\boldsymbol{n})}{2kq^*}.$$

Consider the vector $\tilde{\boldsymbol{n}} = (\tilde{n}_1, \dots, \tilde{n}_k)^T \in \mathbb{N}^k$ defined by

$$\tilde{n}_i := \begin{cases} n_i, & \text{if } n_i \geqslant n_0, \\ 0, & \text{otherwise.} \end{cases}$$

Let U_{big} be the union of blocks B_i for which $n_i \geqslant n_0$. We will apply Lemma 24 for the induced subgraph $\widetilde{\boldsymbol{G}} := \boldsymbol{G}[U_{\text{big}}] \sim \mathcal{G}(\tilde{\boldsymbol{n}}, P)$ (ignoring zero components of $\tilde{\boldsymbol{n}}$). Then, we will use Theorem 23 to colour the rest of the vertices of $\boldsymbol{G}$.

First, we check that $\widetilde{\boldsymbol{G}}$ satisfies the assumptions of Lemma 24. From the triangle inequality in Theorem 10(e) and the assumptions, we find that

$$w(\tilde{\boldsymbol{n}}) \geqslant w_*(\tilde{\boldsymbol{n}}) \geqslant w_*(\boldsymbol{n}) - w_*(\boldsymbol{n} - \tilde{\boldsymbol{n}}) \geqslant (1 + o(1))w(\boldsymbol{n}) - w_*(\boldsymbol{n} - \tilde{\boldsymbol{n}}).$$

Using the upper bound of Theorem 10(d) and by the definitions of $\tilde{\boldsymbol{n}}, n_0$, we get

$$w_*(\boldsymbol{n} - \tilde{\boldsymbol{n}}) \leqslant \sum_{i\in[k]} n_0 q_{ii} = \sum_{i\in[k]} \frac{\varepsilon\, w(\boldsymbol{n})}{2kq^*} q_{ii} \leqslant \frac{\varepsilon}{2} w(\boldsymbol{n}).$$

Therefore, by the assumptions

$$w(\tilde{\boldsymbol{n}}) \geqslant \left(1 - \frac{\varepsilon}{2} + o(1)\right) w(\boldsymbol{n}) \geqslant (q^*\|\boldsymbol{n}\|)^{1+o(1)} \geqslant \|\boldsymbol{n}\|^{1-\sigma+o(1)} \geqslant \|\tilde{\boldsymbol{n}}\|^{1-\sigma+o(1)}. \tag{70}$$

Using our assumption that $w(\boldsymbol{n}) = (1+o(1))w_*(\boldsymbol{n})$ and Theorem 10(d) again, we get

$$\|\boldsymbol{n} - \tilde{\boldsymbol{n}}\| \leqslant kn_0 = \frac{\varepsilon\, w(\boldsymbol{n})}{2q^*} = (1+o(1))\frac{\varepsilon\, w_*(\boldsymbol{n})}{2q^*} \leqslant (1+o(1))\frac{\varepsilon}{2}\|\boldsymbol{n}\|. \tag{71}$$

In particular, we get $\|\tilde{\boldsymbol{n}}\| = \|\boldsymbol{n}\|^{1+o(1)}$. By the definition of $\tilde{\boldsymbol{n}}$, all non-zero components of $\tilde{\boldsymbol{n}}$ are at least n_0. Using the bounds of Theorem 10(d) and the assumptions, we have that

$$q^*\|\boldsymbol{n}\| \geqslant w_*(\boldsymbol{n}) \geqslant (1+o(1))w(\boldsymbol{n}) \geqslant (q^*\|\boldsymbol{n}\|)^{1+o(1)}.$$

Thus, $\frac{w(\boldsymbol{n})}{q^*\|\boldsymbol{n}\|} = \|\boldsymbol{n}\|^{o(1)}$. Recalling also $k = \|\boldsymbol{n}\|^{o(1)}$, we find that

$$n_0 = \frac{\varepsilon\, w(\boldsymbol{n})}{2kq^*} = \frac{\varepsilon}{2k} \cdot \frac{w(\boldsymbol{n})}{q^*\|\boldsymbol{n}\|} \cdot \|\boldsymbol{n}\| = \|\boldsymbol{n}\|^{1+o(1)} = \|\tilde{\boldsymbol{n}}\|^{1+o(1)} \tag{72}$$

and, since $kq^* \ll \log\|\boldsymbol{n}\|$ and $\tilde{\boldsymbol{n}} \preceq \boldsymbol{n}$,

$$n_0 \gg \frac{w(\boldsymbol{n})}{\log\|\boldsymbol{n}\|} \geqslant \frac{w(\tilde{\boldsymbol{n}})}{\log\|\boldsymbol{n}\|} = \frac{w(\tilde{\boldsymbol{n}})}{(1+o(1))\log\|\tilde{\boldsymbol{n}}\|}.$$

Thus, all assumptions of Lemma 24 for $\widetilde{\boldsymbol{G}} \sim \mathcal{G}(\tilde{\boldsymbol{n}}, P)$ hold.

Applying Lemma 24 with $\tilde{\varepsilon} := \frac{\varepsilon}{3}$, we show that, with probability at least

$$1 - \exp\left(-\|\tilde{\boldsymbol{n}}\|^{2-4\sigma+o(1)}\right) = 1 - \exp\left(-\|\boldsymbol{n}\|^{2-4\sigma+o(1)}\right),$$

there is a colouring of $\widetilde{\boldsymbol{G}}$ with at most

$$(1+\tilde{\varepsilon})\,\frac{w(\tilde{\boldsymbol{n}})}{2\log(w(\tilde{\boldsymbol{n}}))} \leqslant (1+\tilde{\varepsilon}+o(1))\,\frac{w(\boldsymbol{n})}{2\log(q^*\|\boldsymbol{n}\|)}$$

colours covering all vertices from each block B_i that $n_i \geqslant n_0$ except at most

$$n_i\left(\frac{1}{\log^2\|\tilde{\boldsymbol{n}}\|} + 5(\tilde{\varepsilon})^{-1}\frac{w(\tilde{\boldsymbol{n}})}{n_0\log(w(\tilde{\boldsymbol{n}}))}\right)$$

vertices.

Using (70) and recalling $\sigma < \frac{1}{4}$, we find that $w(\tilde{\boldsymbol{n}}) = (w(\boldsymbol{n}))^{1+o(1)} = (q^*\|\boldsymbol{n}\|)^{1+o(1)}$. Using also the assumption $kq^* \ll \log\|\boldsymbol{n}\|$, we conclude that the set of remaining uncoloured vertices (with sufficiently high probability) has at most

$$\begin{aligned} u_i &:= n_0 + n_i\left(\frac{1}{\log^2\|\tilde{\boldsymbol{n}}\|} + 5(\tilde{\varepsilon})^{-1}\frac{w(\tilde{\boldsymbol{n}})}{n_0\log(w(\tilde{\boldsymbol{n}}))}\right) \\ &= n_0 + (1+o(1))\frac{30kq^*n_i}{\varepsilon^2\log(q^*\|\boldsymbol{n}\|)} \end{aligned}$$

vertices in each block B_i. Since $n_0 = \|\boldsymbol{n}\|^{1+o(1)}$ by (72) and $k = \|\boldsymbol{n}\|^{o(1)}$ by our assumptions, we find that $u_i = \|\boldsymbol{n}\|^{1+o(1)}$ and $\|\boldsymbol{u}\| = \|\boldsymbol{n}\|^{1+o(1)}$, where $\boldsymbol{u} = (u_1, \ldots, u_k)^T$. Then, we get that

$$q^* \geqslant \hat{q}(\boldsymbol{u}) := \frac{\sum_{i\in[k]} u_i q_{ii}}{\|\boldsymbol{u}\|} \geqslant \frac{q^* n_0}{\|\boldsymbol{u}\|} = q^*\|\boldsymbol{n}\|^{o(1)}.$$

By the definitions of u_i and n_0, we observe that

$$\begin{aligned}\hat{q}(\boldsymbol{u})\|\boldsymbol{u}\| &\leqslant kq^* n_0 + (1+o(1))\frac{30kq^*\|\boldsymbol{n}\|\hat{q}(\boldsymbol{n})}{\varepsilon^2\log(q^*\|\boldsymbol{n}\|)} \\ &= \frac{\varepsilon}{2}w(\boldsymbol{n}) + (1+o(1))\frac{30kq^*\|\boldsymbol{n}\|\hat{q}(\boldsymbol{n})}{\varepsilon^2\log(q^*\|\boldsymbol{n}\|)}.\end{aligned}$$

Using Theorem 23 with $n := \|\boldsymbol{n}\|$ with any $\varepsilon' < \frac{1}{3}$, we can colour the remaining vertices using at most

$$(1+\varepsilon')\,\frac{\hat{q}(\boldsymbol{u})\|\boldsymbol{u}\|}{2\log(q^*\|\boldsymbol{n}\|)} \leqslant \frac{(1+\varepsilon')\varepsilon}{4}\cdot\frac{w(\boldsymbol{n})}{\log(q^*\|\boldsymbol{n}\|)} + 20\varepsilon^{-2}\frac{kq^*\|\boldsymbol{n}\|\hat{q}(\boldsymbol{n})}{\log(q^*\|\boldsymbol{n}\|)}$$

colours with a probability at least $1-\exp\left(-\|\boldsymbol{n}\|^{2-4\sigma+o(1)}\right)$. Thus, the total number of colours is at most

$$\left(1+\tilde{\varepsilon}+\frac{(1+\varepsilon')\varepsilon}{2}+o(1)\right)\frac{w(\boldsymbol{n})}{2\log(q^*\|\boldsymbol{n}\|)} + 20\varepsilon^{-2}\frac{kq^*\|\boldsymbol{n}\|\hat{q}(\boldsymbol{n})}{\log(q^*\|\boldsymbol{n}\|)}.$$

The claimed bound on $\chi(\boldsymbol{G})$ follows since $\tilde{\varepsilon} = \varepsilon/3$ and $\frac{(1+\varepsilon')\varepsilon}{2} < 2\varepsilon/3$. □

6.3 Upper tail bound: proof of Theorem 9

By the near-optimal integer system property given in Theorem 10(g), we can find k vectors $(\boldsymbol{n}^{(t)})_{t\in[k]}$ from $\mathbb{N}^k$ such that

$$\boldsymbol{n} = \sum_{t\in[k]}\boldsymbol{n}^{(t)} \qquad \text{and} \qquad \sum_{t\in[k]} w(\boldsymbol{n}^{(t)}) \leqslant w_*(\boldsymbol{n}) + k^2 q^*. \tag{73}$$

We treat our graph $\boldsymbol{G}$ as the union of the vertex disjoint random graphs $\boldsymbol{G}^{(t)} \sim \mathcal{G}(\boldsymbol{n}^{(t)}, P)$, for $t \in [k]$. Since we can colour them with different colours, we have that, with probability 1,

$$\chi(\boldsymbol{G}) \leqslant \sum_{t\in[k]}\chi(\boldsymbol{G}^{(t)}). \tag{74}$$

Let

$$T_{\text{small}} = \left\{t\in[k] : w(\boldsymbol{n}^{(t)}) < \frac{w_*(\boldsymbol{n})}{k^2\log\|\boldsymbol{n}\|}\right\}.$$

The proof of Theorem 9 consists of two parts. First, applying Theorem 23, we show that, with sufficiently high probability, $\sum_{t\in T_{\text{small}}}\chi(\boldsymbol{G}^{(t)}) \ll \frac{w_*(\boldsymbol{n})}{\log(q^*\|\boldsymbol{n}\|)}$. Second, we use Lemma 28 to estimate $\chi(\boldsymbol{G}^{(t)})$ for $t \notin T_{\text{small}}$.

Before proceeding, we derive some preliminary bounds implied by our assumptions. Since $k = \|\boldsymbol{n}\|^{o(1)}$ and $\hat{q}(\boldsymbol{n}), q^* = \|\boldsymbol{n}\|^{-\sigma+o(1)}$, we find that

$$\frac{(\hat{q}(\boldsymbol{n}))^2}{kq^*} = \|\boldsymbol{n}\|^{-\sigma+o(1)} = \|\boldsymbol{n}\|^{o(1)}kq^*.$$

Then, using the bounds of Theorem 10(d), we get

$$q^* \geqslant \hat{q}(\boldsymbol{n}) \geqslant \frac{w_*(\boldsymbol{n})}{\|\boldsymbol{n}\|} \geqslant \frac{(\hat{q}(\boldsymbol{n}))^2}{\sum_{i\in[k]} q_{ii}} \geqslant \frac{(\hat{q}(\boldsymbol{n}))^2}{kq^*} = \|\boldsymbol{n}\|^{o(1)} kq^*. \tag{75}$$

Since $kq^*\|\boldsymbol{n}\| = \|\boldsymbol{n}\|^{1-\sigma+o(1)}$ we derive from (75) that

$$w_*(\boldsymbol{n}) = (q^*\|\boldsymbol{n}\|)^{1+o(1)} = \|\boldsymbol{n}\|^{1-\sigma+o(1)} \gg k^2 \log(q^*\|\boldsymbol{n}\|). \tag{76}$$

Using (75) and assumption (10), we get that

$$\hat{q}(\boldsymbol{n}) \geqslant \frac{w_*(\boldsymbol{n})}{\|\boldsymbol{n}\|} \gg \frac{kq^*\hat{q}(\boldsymbol{n})}{\log\|\boldsymbol{n}\|},$$

which implies

$$q^* \leqslant kq^* \ll \log\|\boldsymbol{n}\|,$$

which is needed to apply Theorem 23 and Lemma 28. Also, by the definition of $w_*(\cdot)$, we know that

$$w_*(\boldsymbol{n}) \leqslant \sum_{t\in[k]} w_*(\boldsymbol{n}^{(t)}).$$

which, together with (73), implies that

$$\sum_{t\in[k]} (w(\boldsymbol{n}^{(t)}) - w_*(\boldsymbol{n}^t)) \leqslant k^2 q^*$$

Since every term of the sum above is non-negative, using the assumption $k = \|\boldsymbol{n}\|^{o(1)}$ and the estimate $q^* \ll \log\|\boldsymbol{n}\|$, we derive that, for any $t \in [k]$,

$$w(\boldsymbol{n}^{(t)}) \leqslant w_*(\boldsymbol{n}^{(t)}) + k^2 q^* = w_*(\boldsymbol{n}^{(t)}) + \|\boldsymbol{n}\|^{o(1)}.$$

Then, using the first equality of (76) and our assumption $k = \|\boldsymbol{n}\|^{o(1)}$, for any $t \in [k] \setminus T_{\text{small}}$, we get

$$w(\boldsymbol{n}^{(t)}) \geqslant \frac{w_*(\boldsymbol{n})}{k^2 \log\|\boldsymbol{n}\|} = (q^*\|\boldsymbol{n}\|)^{1+o(1)} \geqslant (q^*\|\boldsymbol{n}^{(t)}\|)^{1+o(1)}. \tag{77}$$

This implies that

$$w_*(\boldsymbol{n}^{(t)}) = (1+o(1)) w(\boldsymbol{n}^{(t)}) \geqslant (q^*\|\boldsymbol{n}^{(t)}\|)^{1+o(1)}$$

as required by Lemma 28.

Now, consider any $t \in T_{\text{small}}$. Define $\boldsymbol{u} = (u_1, \ldots, u_k)^T \in \mathbb{R}_+^k$

$$u_i := n_i^{(t)} + \frac{w_*(\boldsymbol{n})}{k^2 q^* \log\|\boldsymbol{n}\|}.$$

Using the bound $w_*(\boldsymbol{n}) \leqslant q^*\|\boldsymbol{n}\|$ of Theorem 10(d) and the assumption $k = \|\boldsymbol{n}\|^{o(1)}$, we get

$$\begin{aligned}\|\boldsymbol{n}^{(t)}\| \leqslant \|\boldsymbol{u}\| &\leqslant \|\boldsymbol{n}^{(t)}\| + \sum_{i\in[k]} \frac{w_*(\boldsymbol{n})}{k^2 q^* \log\|\boldsymbol{n}\|} \\ &\leqslant \|\boldsymbol{n}^{(t)}\| + \frac{\|\boldsymbol{n}\|}{k\log\|\boldsymbol{n}\|} \leqslant \|\boldsymbol{n}\| + \|\boldsymbol{n}\|^{1+o(1)} = \|\boldsymbol{n}\|^{1+o(1)}.\end{aligned} \tag{78}$$

Using the lower bounds of Theorem 10(d) and the inequality $\hat{q}(\boldsymbol{n}^{(t)}) \leqslant q^*$, we find that

$$w(\boldsymbol{n}^{(t)}) \geqslant w_*(\boldsymbol{n}^{(t)}) \geqslant \frac{(\hat{q}(\boldsymbol{n}^{(t)}))^2\|\boldsymbol{n}^{(t)}\|}{kq^*} \geqslant \frac{\hat{q}(\boldsymbol{n}^{(t)})\|\boldsymbol{n}^{(t)}\|}{k}.$$

This implies $\hat{q}(\boldsymbol{n}^{(t)})\|\boldsymbol{n}^{(t)}\| \leqslant kw(\boldsymbol{n}^{(t)}) < \frac{w_*(\boldsymbol{n})}{k\log\|\boldsymbol{n}\|}$ because $w(\boldsymbol{n}^{(t)}) < \frac{w_*(\boldsymbol{n})}{k^2\log\|\boldsymbol{n}\|}$ since $t \in T_{\text{small}}$. Using the definition of u_i we get that

$$\begin{aligned}\hat{q}(\boldsymbol{u})\|\boldsymbol{u}\| &= \sum_{i\in[k]} q_{ii}u_i = \sum_{i\in[k]} q_{ii}\left(n_i^{(t)} + \frac{w_*(\boldsymbol{n})}{k^2q^*\log\|\boldsymbol{n}\|}\right) \\ &= \hat{q}(\boldsymbol{n}^{(t)})\|\boldsymbol{n}^{(t)}\| + \frac{w_*(\boldsymbol{n})}{k^2q^*\log\|\boldsymbol{n}\|}\sum_{i\in[k]} q_{ii} \ll \frac{w_*(\boldsymbol{n})}{k}.\end{aligned} \tag{79}$$

Using (76), the inequality $\hat{q}(\boldsymbol{u}) \leqslant q^*$, and $q^* = \|\boldsymbol{n}\|^{-\sigma+o(1)}$ by (8), observe also that

$$\hat{q}(\boldsymbol{u})\|\boldsymbol{u}\| \geqslant \frac{w_*(\boldsymbol{n})}{k^2q^*\log\|\boldsymbol{n}\|}\sum_{i\in[k]} q_{ii} \geqslant \frac{w_*(\boldsymbol{n})}{k^2\log\|\boldsymbol{n}\|} = q^*\|\boldsymbol{n}\|^{1+o(1)}.$$

Recalling from (78) that $\|\boldsymbol{u}\| \leqslant \|\boldsymbol{n}\|^{1+o(1)}$ and using $\hat{q}(\boldsymbol{u}) \leqslant q^*$, we derive that

$$q^* = \hat{q}(\boldsymbol{u})\|\boldsymbol{n}\|^{o(1)}.$$

Applying Theorem 23 with $n := \|\boldsymbol{n}\|$ and using $\hat{q}(\boldsymbol{u})\|\boldsymbol{u}\| \ll \frac{w_*(\boldsymbol{n})}{k}$ from (79) we get that

$$\chi(\boldsymbol{G}^{(t)}) = O\left(\frac{\hat{q}(\boldsymbol{u})\|\boldsymbol{u}\|}{2\log(q^*\|\boldsymbol{n}\|)}\right) \ll \frac{w_*(\boldsymbol{n})}{k\log(q^*\|\boldsymbol{n}\|)}$$

with probability at least $1 - \exp\left(-\|\boldsymbol{n}\|^{2-4\sigma+o(1)}\right)$. Applying the union bound, it follows that, with sufficiently high probability,

$$\sum_{t\in T_{\text{small}}} \chi(\boldsymbol{G}^{(t)}) \ll \frac{w_*(\boldsymbol{n})}{\log(q^*\|\boldsymbol{n}\|)}. \tag{80}$$

Next, we consider any $t \in [k] \setminus T_{\text{small}}$. Since $w_*(\boldsymbol{n}^{(t)}) \leqslant q^*\|\boldsymbol{n}^{(t)}\|$ by Theorem 10(d) and $w_*(\boldsymbol{n}^{(t)}) = (1+o(1))w(\boldsymbol{n}^{(t)})$, we have $\|\boldsymbol{n}^{(t)}\| \geqslant (1+o(1))\frac{w(\boldsymbol{n}^{(t)})}{q^*}$. Using (77) and the bound $k = \|\boldsymbol{n}\|^{o(1)}$, we find also that

$$\|\boldsymbol{n}\| \geqslant \|\boldsymbol{n}^{(t)}\| \geqslant (1+o(1))\frac{w(\boldsymbol{n}^{(t)})}{q^*} \geqslant (1+o(1))\frac{w_*(\boldsymbol{n})}{k^2q^*\log\|\boldsymbol{n}\|} = \|\boldsymbol{n}\|^{1+o(1)}.$$

That is, we get $\|\boldsymbol{n}^{(t)}\| = \|\boldsymbol{n}\|^{1+o(1)}$. Applying Lemma 28 with any $0 < \varepsilon' < \varepsilon$, we derive that,

$$\chi(\boldsymbol{G}^{(t)}) \leqslant (1+\varepsilon')\,\frac{w(\boldsymbol{n}^{(t)})}{2\log(q^*\|\boldsymbol{n}^{(t)}\|)} + O\left(\frac{kq^*\|\boldsymbol{n}^{(t)}\|\hat{q}(\boldsymbol{n}^{(t)})}{\log^2(q^*\|\boldsymbol{n}^{(t)}\|)}\right),$$

with probability at least $1 - \exp\left(-\|\boldsymbol{n}\|^{2-4\sigma+o(1)}\right)$. Using the union bound for all such event over $t \notin T_{\text{small}}$, we get that, with sufficiently high probability,

$$\begin{aligned}\sum_{t\in[k]\setminus T_{\text{small}}} \chi(\boldsymbol{G}^{(t)}) &\leqslant (1+\varepsilon'+o(1))\,\frac{w_*(\boldsymbol{n}) + k^2q^*}{2\log(q^*\|\boldsymbol{n}\|)} + O\left(\frac{kq^*\|\boldsymbol{n}\|\hat{q}(\boldsymbol{n})}{\log^2(q^*\|\boldsymbol{n}\|)}\right)\\ &\leqslant (1+\varepsilon'+o(1))\,\frac{w_*(\boldsymbol{n})}{2\log(q^*\|\boldsymbol{n}\|)}.\end{aligned} \tag{81}$$

For the first inequality in (81), we estimated the $O(\cdot)$ term using

$$\sum_{t\in[k]\setminus T_{\text{small}}} \|\boldsymbol{n}^{(t)}\|\hat{q}(\boldsymbol{n}^{(t)}) = \sum_{t\in[k]\setminus T_{\text{small}}}\sum_{i\in[k]} q_{ii}n_i^{(t)} \leqslant \sum_{i\in[k]} q_{ii}n_i = \|\boldsymbol{n}\|\hat{q}(\boldsymbol{n}).$$

Finally, substituting the bounds of (80) and (81) into (74) and bounding

$$1+\varepsilon'+o(1) \leqslant 1+\varepsilon,$$

we complete the proof of Theorem 9.

7 Proof of Theorem 10

The part (a) follows directly by the definition.

The part (b) is trivial if $w(\boldsymbol{x}) = 0$ as we can take $\boldsymbol{y} = \boldsymbol{0}$. Thus, we can assume that $w(\boldsymbol{x}) > 0$. Observe that the function $\boldsymbol{y} \mapsto \frac{\boldsymbol{y}^TQ\boldsymbol{y}}{\|\boldsymbol{y}\|}$ (defined to be 0 at origin) is continuous on the compact set $K_{\boldsymbol{x}}$ defined by

$$K_{\boldsymbol{x}} := \{\boldsymbol{y} \in \mathbb{R}_+^k \,:\, \boldsymbol{y} \preceq \boldsymbol{x}\}.$$

Therefore, there exists a maximiser $\boldsymbol{y}^* = (y_1^*, \ldots, y_k^*) \in K_{\boldsymbol{x}}$. Furthermore, for each $i \in [k]$,

$$f_i(y_i) := \frac{\boldsymbol{y}^T\,Q\,\boldsymbol{y}}{\|\boldsymbol{y}\|} = q_{ii}(y_i - a) + \frac{b}{y_1 + \cdots + y_n},$$

where $\boldsymbol{y}$ differ from $\boldsymbol{y}^*$ in the i-th component only and $a = a(\boldsymbol{y}^*)$, $b = b(\boldsymbol{y}^*)$. If $b \leqslant 0$ then $q_{ii} \neq 0$ because $f_i(y_i^*) = w(\boldsymbol{x}) > 0$. This implies that $q_{ii} > 0$ so the function $y_i \mapsto f_i(y_i)$ is strictly increasing. If $b > 0$ then the function $y_i \mapsto f_i(y_i)$ is strictly convex. In any case, the maximum lies on the boundary $\{0, x_i\}$. Repeating the same argument for the other components, we get (b).

Before proceeding, we introduce an additional notation. For a positive integer ℓ and $\boldsymbol{x} \in \mathbb{R}_+^k$, let

$$w_\ell(\boldsymbol{x}) := \min_{(\boldsymbol{x}^{(t)})_{t\in[\ell]}} \quad \sum_{t=1}^{\ell} w_Q(\boldsymbol{x}^{(t)}) \tag{82}$$
$$\text{subject to } \sum_{t=1}^{\ell} \boldsymbol{x}^{(t)} = \boldsymbol{x} \text{ and } \boldsymbol{x}^{(t)} \in \mathbb{R}_+^k \text{ for all } t \in [\ell].$$

A minimum system of ℓ vectors $(\boldsymbol{x}^{(t)})_{t\in[\ell]}$ in (82) exists since $\sum_{t=1}^{\ell} w_Q(\boldsymbol{x}^{(t)})$ is a continuous function on the compact set of all systems that $\sum_{t=1}^{\ell} \boldsymbol{x}^{(t)} = \boldsymbol{x}$ and $\boldsymbol{x}^{(t)} \in \mathbb{R}_+^k$. Using definitions (4) and (6), we find that

$$w(\boldsymbol{x}) = w_1(\boldsymbol{x}) \geqslant w_\ell(\boldsymbol{x}) \geqslant w_*(\boldsymbol{x}) \quad \text{and} \quad w_*(\boldsymbol{x}) = \lim_{\ell\to\infty} w_\ell(\boldsymbol{x}).$$

We will also use the following identity.

Lemma 29. *For any $\boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}_+^k$, we have*

$$\frac{\boldsymbol{y}^T Q\, \boldsymbol{y}}{\|\boldsymbol{y}\|} + \frac{\boldsymbol{z}^T Q\, \boldsymbol{z}}{\|\boldsymbol{z}\|} - \frac{(\boldsymbol{y}+\boldsymbol{z})^T Q\, (\boldsymbol{y}+\boldsymbol{z})}{\|\boldsymbol{y}\| + \|\boldsymbol{z}\|}$$
$$= \frac{\|\boldsymbol{y}\|\|\boldsymbol{z}\|}{\|\boldsymbol{y}\| + \|\boldsymbol{z}\|} \left(\frac{\boldsymbol{y}}{\|\boldsymbol{y}\|} - \frac{\boldsymbol{z}}{\|\boldsymbol{z}\|}\right)^T Q \left(\frac{\boldsymbol{y}}{\|\boldsymbol{y}\|} - \frac{\boldsymbol{z}}{\|\boldsymbol{z}\|}\right),$$

where $\frac{\boldsymbol{x}^T Q \boldsymbol{x}}{\|\boldsymbol{x}\|}$ is 0 if $\|\boldsymbol{x}\| = 0$ and the RHS is taken to be 0 if $\|\boldsymbol{y}\| = 0$ or $\|\boldsymbol{z}\| = 0$.

Proof. This follows by expanding $\left(\frac{\boldsymbol{y}}{\|\boldsymbol{y}\|} - \frac{\boldsymbol{z}}{\|\boldsymbol{z}\|}\right)^T Q \left(\frac{\boldsymbol{y}}{\|\boldsymbol{y}\|} - \frac{\boldsymbol{z}}{\|\boldsymbol{z}\|}\right)$ and $(\boldsymbol{y}+\boldsymbol{z})^T Q\, (\boldsymbol{y}+\boldsymbol{z})$, using linearity, and direct substitution. ☐

We proceed to part (c). Take any $\boldsymbol{y} \in K_{\boldsymbol{x}}$ such that $\frac{\boldsymbol{y}^T Q \boldsymbol{y}}{\|\boldsymbol{y}\|} = w(\boldsymbol{x})$, which exists by part (b). Using Lemma 29, we find that

$$w_\ell(\boldsymbol{x}) \geqslant w_\ell(\boldsymbol{y}) = \sum_{t=1}^{\ell} w(\boldsymbol{y}^{(t)}) \geqslant \sum_{t=1}^{\ell} \frac{(\boldsymbol{y}^t)^T Q \boldsymbol{y}^t}{\|\boldsymbol{y}^t\|} \geqslant \frac{\boldsymbol{y}^T Q \boldsymbol{y}}{\|\boldsymbol{y}\|} = w(\boldsymbol{x}).$$

Taking the limit $\ell \to \infty$, we prove (c).

The upper bound

$$w_*(\boldsymbol{x}) \leqslant \hat{q}(\boldsymbol{x})\|\boldsymbol{x}\| = \sum_{i\in[k]} x_i q_{ii}$$

follows by definition (6) taking the system of k vectors $(\boldsymbol{x}^{(t)})_{t\in[k]}$, where, for each $t \in [k]$, the t-th component of $\boldsymbol{x}^{(t)}$ equals x_t while other components are 0. Also we have $\hat{q}(\boldsymbol{x})\|\boldsymbol{x}\| \leqslant q^*\|\boldsymbol{x}\|$. Next we prove the lower bound for $w_*(\boldsymbol{x})$ of part (d). Let $(\boldsymbol{x}^{(t)})_{t\in\ell}$ be such that $w_\ell(\boldsymbol{x}) = \sum_{t\in[\ell]} w(\boldsymbol{x}^{(t)})$ and $\sum_{t\in[\ell]} \boldsymbol{x}^{(\ell)} = \boldsymbol{x}$. By the Cauchy-Schwarz inequality, we find that

$$w(\boldsymbol{x}^{(t)}) \geqslant \frac{(\boldsymbol{x}^{(t)})^T Q \boldsymbol{x}^{(t)}}{\|\boldsymbol{x}^{(t)}\|} \geqslant \frac{\sum_{i\in[k]} q_{ii}(x_i^{(t)})^2}{\|\boldsymbol{x}^{(t)}\|} \geqslant \frac{\left(\sum_{i\in[k]} q_{ii} x_i^{(t)}\right)^2}{\|\boldsymbol{x}^{(t)}\| \sum_{i\in[k]} q_{ii}}.$$

Using the Cauchy-Schwarz inequality again, we obtain

$$\sum_{t\in[\ell]} \|\boldsymbol{x}^{(t)}\| \cdot \sum_{t\in[\ell]} \frac{\left(\sum_{i\in[k]} q_{ii} x_i^{(t)}\right)^2}{\|\boldsymbol{x}^{(t)}\|} \geqslant \left(\sum_{t\in[\ell]} \sum_{i\in[k]} q_{ii} x_i^{(t)}\right)^2 = \left(\sum_{i\in[k]} q_{ii} x_i\right)^2 = (\hat{q}(\boldsymbol{x})^2)\|\boldsymbol{x}\|^2.$$

Therefore,

$$w_\ell(\boldsymbol{x}) = \sum_{t\in[\ell]} \boldsymbol{x}^{(\ell)} \geqslant \frac{(\hat{q}(\boldsymbol{x})^2)}{\sum_{i\in[k]} q_{ii}} \|\boldsymbol{x}\|.$$

Taking the limit $\ell \to \infty$ and observing $\sum_{i\in[k]} q_{ii} \leqslant kq^*$, we complete the proof of (d).

For (e), consider any two vector systems $\mathcal{S} \in \mathcal{F}(\boldsymbol{x})$ and $\mathcal{S}' \in \mathcal{F}(\boldsymbol{x}')$; see definition (6). Then, the union system $\mathcal{S} \cup \mathcal{S}'$ belongs to $\mathcal{F}(\boldsymbol{x} + \boldsymbol{x}')$. Thus,

$$\sum_{\boldsymbol{y}\in\mathcal{S}} w(\boldsymbol{y}) + \sum_{\boldsymbol{y}\in\mathcal{S}'} w(\boldsymbol{y}) = \sum_{\boldsymbol{y}\in\mathcal{S}\cup\mathcal{S}'} w(\boldsymbol{y}) \geqslant w_*(\boldsymbol{x} + \boldsymbol{x}').$$

Taking the infimum over $\mathcal{S}, \mathcal{S}'$, we get (e).

For (f), assume $\ell > k$. Then, we can find real constants $c^{(t)}$, $t = 1, \ldots, \ell$, such that

$$\sum_{t=1}^{\ell} c^{(t)} \boldsymbol{x}^{(t)} = 0.$$

Next, we show that $\boldsymbol{x}_\varepsilon^{(t)} = (1 + \varepsilon c^{(t)})\boldsymbol{x}^{(t)}$ gives another optimal solution of (82). Observe that $\sum_{t=1}^{\ell} \boldsymbol{x}_\varepsilon^{(t)} = \boldsymbol{x}$. If $|\varepsilon|$ is sufficiently small that $\boldsymbol{x}_\varepsilon^{(t)} \in \mathbb{R}_+^k$ then

$$f(\varepsilon) := \sum_{t=1}^{\ell} w(\boldsymbol{x}_\varepsilon^{(t)}) = \sum_{t=1}^{\ell} (1 + \varepsilon c^{(t)}) w(\boldsymbol{x}^{(t)}),$$

that is, $f(\varepsilon)$ is a linear function of ε. Since $\varepsilon = 0$ gives the minimum value of $f(\varepsilon)$, it should be a constant function. Then, we can make at least one of $\boldsymbol{x}_\varepsilon^{(t)}$ to be trivial while others remain in $\mathbb{R}_+^k$ without changing the value of the target function $\sum_{t=1}^{\ell} w(\boldsymbol{x}_\varepsilon^{(t)})$. This implies $w_\ell(\boldsymbol{x}) = w_{\ell-1}(\boldsymbol{x})$. Repeating these arguments several times we find that

$$w_\ell(\boldsymbol{x}) = w_{\ell-1}(\boldsymbol{x}) = \cdots = w_k(\boldsymbol{x}).$$

Taking the limit $\ell \to \infty$, we get (f).

Finally, we proceed to part (g). Using part (f), we can find a system $(\boldsymbol{y}^{(t)})_{t\in[k]}$ such that

$$\sum_{t=1}^{k} w(\boldsymbol{y}^{(t)}) = w_*(\boldsymbol{x}) \qquad \text{and} \qquad \sum_{t=1}^{k} \boldsymbol{y}^{(t)} = \boldsymbol{x}.$$

In particular, by definition of $w_*(\cdot)$, we find that

$$w_*(\boldsymbol{y}^{(t)}) = w(\boldsymbol{y}^{(t)}). \tag{83}$$

Define $\boldsymbol{x}^{(t)} := \lfloor \boldsymbol{y}^{(t)} \rfloor$. Combining (83) and parts (d), (e), we also find that for all $t \in [k]$

$$w(\boldsymbol{y}^{(t)}) = w_*(\boldsymbol{y}^{t}) \leqslant w_*(\boldsymbol{x}^{(t)}) + w_*(\boldsymbol{y}^{(t)} - \boldsymbol{x}^{(t)}) \leqslant w_*(\boldsymbol{x}^{(t)}) + q^*\|\boldsymbol{y}^{(t)} - \boldsymbol{x}^{(t)}\| \leqslant w_*(\boldsymbol{x}^{(t)}) + kq^*.$$

Thus, we get that

$$\sum_{t=1}^{k} w(\boldsymbol{x}^{(t)}) \geqslant \sum_{t=1}^{k} (w(\boldsymbol{y}^{(t)}) - kq^*) = w_*(\boldsymbol{x}) - k^2 q^*.$$

Now, we can increase some components of $\boldsymbol{x}^{(t)}$ to ensure that $\sum_{t\in[k]} \boldsymbol{x}^{(t)} = \boldsymbol{x}$. By part (a), this would only increase the values of $w(\boldsymbol{x}^{(t)})$. This completes the proof of part (g) and Theorem 10.

Acknowledgements

We thank the referee for valuable comments and suggestions that helped us improve the paper.

References

[1] E. Abbe, Community detection and stochastic block models: recent developments, *Journal of Machine Learning Research*, **18** (2017), 6446–6531.

[2] D. Achlioptas, A. Naor, The two possible values of the chromatic number of a random graph, *Annals of Mathematics*, **162** (2005), 1333–1349.

[3] N. Alon and M. Krivelevich, The concentration of the chromatic number of random graphs, *Combinatorica*, **17** (1997), 303–313.

[4] N. Alon, M. Krivelevich, B. Sudakov, Subgraphs with a large cochromatic number, *Journal of Graph Theory*, **26** (1997), 295–297.

[5] R. Berkowitz, P. Devlin, C. Lee, H. Reichard, D. Townley, Expected chromatic number of random subgraphs, arXiv:1811.02018.

[6] B. Bukh, Interesting problems, N.2, http://www.borisbukh.org/problems.html.

[7] B. Bollobás, The chromatic number of random graphs, *Combinatorica*, **8**(1) (1988), 49–55.

[8] B. Bollobás, How sharp is the concentration of the chromatic number?, *Combinatorics, Probability and Computing*, **13** (2004), 115–117.

[9] F. Chung, L. Lu. Connected components in random graphs with given expected degrees. *Annals of Combinatorics*, **6** (2002), 125–145.

[10] A. Coja-Oghlan, K. Panagiotou, A. Steger, On the chromatic number of random graphs, *Journal of Combinatorial Theory, Series B*, **98** (2008), 980–993.

[11] M. Doležal, J. Hladký, A. Máthé, Cliques in dense inhomogeneous random graphs, *Random Structures and Algorithms,* **51**(2) (2017), 275–314.

[12] P. Erdős, A. Rényi, On the evolution of random graphs, *A Magyar Tudományos Akadémia Matematikai Kutató Intézetének Közleményei*, **5** (1960), 17–61.

[13] P. Erdős, A. Rényi, On the existence of a factor of degree one of a connected random graph. *Acta Mathematica Academiae Scientiarum Hungaricae*, **17** (1966), 359–368.

[14] A. Frieze, M. Karoński, *Introduction to Random Graphs*, Cambridge University Press, 2016.

[15] A. Johansson, J. Kahn, V. Vu, Factors in random graphs, *Random Structures and Algorithms*, **33** (2008), 1–28.

[16] A. Heckel, The chromatic number of dense random graphs, *Random Structures and Algorithms*, **53** (2018), 140–182.

[17] A. Heckel, Sharp concentration of the equitable chromatic number of dense random graphs, *Combinatorics, Probability and Computing*, **29** (2020), 213–233.

[18] J. Hladký, I. Rocha, Independent sets, cliques, and colorings in graphons, *European Journal of Combinatorics*, **88** (2020), 103108, 18 pp.

[19] A. Frieze, On the independence number of random graphs, *Discrete Mathematics,* **81** (1990), 171–175.

[20] S. Janson, Poisson approximation for large deviations, *Random Structures and Algorithms*, **1** (1990), 221–229.

[21] S. Janson, T. Łuczak, A. Ruciński, *Random Graphs*, John Wiley and Sons, New York, 2000.

[22] M. Krivelevich, B. Sudakov, The chromatic numbers of random hypergraphs, *Random Structures and Algorithms*, **12** (1998), 381–403.

[23] L. Lovász, *Large Networks and Graph Limits.* American Mathematical Society, 2012.

[24] L. Lovász, B. Szegedy, Szemerédi's lemma for the analyst, *Geometric and Functional Analysis*, **17** (2007), 252–270.

[25] T. Łuczak, The chromatic number of random graphs, *Combinatorica*, **11** (1991), 45–54.

[26] T. Łuczak, A note on the sharp concentration of the chromatic number of random graphs, *Combinatorica*, **11** (1991), 295–297.

[27] A. Martinsson, K. Panagiotou, P. Su, M. Trujić, The chromatic number of dense random block graphs, arXiv:2007.07700.

[28] C. McDiarmid, On the chromatic number of random graphs, *Random Structures and Algorithms*, **1** (1990), 435–442.

[29] B. Mohar, H. Wu, Fractional chromatic number of a random subgraph, *Journal of Graph Theory*, **95** (2020), 467–472.

[30] K. Panagiotou, A. Steger, A note on the chromatic number of a dense random graph, *Discrete Mathematics*, **309** (2009), 3420–3423.

[31] A. Ruciński, Matching and covering the vertices of a random graph by copies of a given graph, *Discrete Mathematics*, **105** (10002), 185–197.

[32] E. Surya, L. Warnke, On the concentration of the chromatic number of random graph, arXiv:2201.00906.

[33] E. Shamir, J. Spencer, Sharp concentration of the chromatic number on random graphs $G_{n,p}$, *Combinatorica*, **7** (1987), 121–129.

[34] I. Shinkar, On coloring random subgraphs of a fixed graph, arXiv:1612.04319.