Random regular graphs of non-constant degree: Concentration of the chromatic number

Sonny Ben-Shimon * Michael Krivelevich †

December 16, 2008

Abstract

In this work we show that with high probability the chromatic number of a graph sampled from the random regular graph model $\mathcal{G}_{n,d}$ for $d = o(n^{1/5})$ is concentrated in two consecutive values, thus extending a previous result of Achlioptas and Moore. This concentration phenomena is very similar to that of the binomial random graph model $\mathcal{G}(n,p)$ with $p = \frac{d}{n}$. Our proof is largely based on ideas of Alon and Krivelevich who proved this two-point concentration result for $\mathcal{G}(n,p)$ for $p = n^{-\delta}$ where $\delta > 1/2$. The main tool used to derive such a result is a careful analysis of the distribution of edges in $\mathcal{G}_{n,d}$, relying both on the switching technique and on bounding the probability of exponentially small events in the configuration model.

1 Introduction

The most widely used random graph model is the binomial random graph, $\mathcal{G}(n,p)$. In this model, which was introduced in a slightly modified form by Erdős and Rènyi, we start with n vertices, labeled, say, by $\{1,\dots,n\} = [n]$, and select a graph on these n vertices by going over all $\binom{n}{2}$ pairs of vertices, deciding uniformly at random with probability p for a pair to be an edge. $\mathcal{G}(n,p)$ is thus a probability space of all labeled graphs on the vertex set $[n]$ where the probability of such a graph, $G = ([n], E)$, to be selected is $p^{|E|}(1-p)^{\binom{n}{2}-|E|}$. This product probability space gives us a wide variety of probabilistic tools to analyze the behavior of various random graph properties of this probability space. (See monographs [9] and [17] for a thorough introduction to the subject of random graphs).

In this paper, we consider a different random graph model. Our probability space, which is denoted by $\mathcal{G}_{n,d}$ (where dn is even), is the uniform space of all d-regular graphs on n vertices labeled by the set $[n]$. In this model, one cannot apply the techniques used to study $\mathcal{G}(n,p)$ as these two models do not share the same probabilistic properties. Whereas the appearances of edges in $\mathcal{G}(n,p)$ are independent, the appearances of edges in $\mathcal{G}_{n,d}$ are not. Nevertheless, many results obtained thus far for the random regular graph model $\mathcal{G}_{n,d}$

*School of Computer Science, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, 69978, Israel. Email: sonny@tau.ac.il.

†School of Mathematical Sciences, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, 69978, Israel. Email: krivelev@tau.ac.il. Research supported in part by a USA-Israel BSF grant, by a grant from the Israel Science Foundation, and by Pazy Memorial Award.

are in some sense equivalent to the results obtained in $\mathcal{G}(n,p)$ with suitable expected degrees, namely, $d = np$ (see, e.g. [32] and [23] for a collection of results). This relation between the two random graph models was partially formalized in [21]. The interested reader is referred to [32] for a thorough survey of the random regular graph model $\mathcal{G}_{n,d}$.

The main research interest in random graph models is the asymptotic behavior of properties as we let the number of vertices of our graph grow to infinity. We say that an event $\mathcal{A}$ in our probability space occurs *with high probability* (or w.h.p. for brevity) if $\Pr[\mathcal{A}] \to 1$ as n goes to infinity. Therefore, from now on and throughout the rest of this work, we will always assume n to be large enough. We use the usual asymptotic notation, that is, for two functions of n, $f(n)$ and $g(n)$, we write $f = O(g)$ if there exists a constant $C > 0$ such that $f(n) \le C \cdot g(n)$ for large enough values of n; $f = o(g)$ if $f/g \to 0$ as n goes to infinity; $f = \Omega(g)$ if $g = O(f)$; $f = \omega(g)$ if $g = o(f)$; $f = \Theta(g)$ if both $f = O(g)$ and $f = \Omega(g)$.

As far as notation goes, we will always assume, unless specified otherwise, that the set of vertices of our graph is $[n]$. We use the usual notation of $N_G(U)$ for the set of neighbors of a vertex set U in a graph G, that is, $N_G(U) = \{v \in V(G) \setminus U \ : \ \exists u \in U \ \{v,u\} \in E(G)\}$. For a single vertex v, we abuse slightly this notation by writing $N_G(v)$ for $N_G(\{v\})$. We denote the degree of a vertex v in a graph G by $d_G(v)$, namely, $d_G(v) = |N_G(v)|$. The set of edges spanned by a set of vertices U, or between two disjoint sets, U and W, is denoted by $E(U)$ and $E(U,W)$, respectively, and the cardinalities of these sets are denoted by $e(U)$ and $e(U,W)$, respectively. We use the notation $\mathcal{G}(n,p)$ or $\mathcal{G}_{n,d}$ to denote both the corresponding probability space, and a random graph generated in this probability space, where the actual meaning is clear from the context.

A *vertex coloring* of a graph G is an assignment of a color to each of its vertices. The coloring is *proper* if no two adjacent vertices are assigned the same color. The *chromatic number* of the graph G, denoted by $\chi(G)$ is the minimum number of colors used in a proper coloring of it. The chromatic number of a graph is one of the most widely researched graph parameters. A major result of Bollobás [8] that was later extended by Łuczak [24] showed that w.h.p. $\chi(\mathcal{G}(n,p)) = (1+o(1))\frac{n \ln \frac{1}{1-p}}{2\ln(np)}$, where $p \le c$ and $np \to \infty$. Frieze and Łuczak in [16] proved a similar result for $\mathcal{G}_{n,d}$.

Theorem 1.1 (Frieze and Łuczak [16])**.** *For any* $0 < \delta < \frac{1}{3}$ *w.h.p.*

$$\left| \chi(\mathcal{G}_{n,d}) - \frac{d}{2\ln d} - \frac{8d \ln\ln d}{\ln^2 d} \right| \le \frac{8d\ln\ln d}{\ln^2 d}, \tag{1}$$

where $d_0 < d < n^\delta$ *for some fixed positive constant* d_0.

Krivelevich et al. [23] and Cooper et al. [13] extended the range of d for which (1) holds.

We say that a random variable X in some discrete probability space Ω is *highly concentrated* if for every $\varepsilon > 0$ it takes one of a finite set of values (not depending on the cardinality of Ω which we think of as tending to infinity) with probability at least $1-\varepsilon$. It has been shown that in the binomial random graph model $\mathcal{G}(n,p)$ the chromatic number is highly concentrated when the graphs are sparse enough. A series of papers, starting with the seminal work of Shamir and Spencer [29], and succeeded by Łuczak [25], and Alon and Krivelevich [3] prove that $\chi(\mathcal{G}(n,p))$ is concentrated w.h.p. in two consecutive values for $p = n^{-1/2-\varepsilon}$ where ε is any positive real. Achlioptas and Naor [2] went even further to compute the two values on which $\chi(\mathcal{G}(n,p))$ is concentrated for $p = \frac{c}{n}$ for a constant c. Finally, Coja-Oghlan, Panagiotou and Steger [12], building upon the foundations of [2], computed three consecutive values on which $\chi(\mathcal{G}(n,p))$ is concentrated for $p = n^{-3/4-\varepsilon}$ where ε is a positive constant.

We prove a similar concentration result in the random regular graph model $\mathcal{G}_{n,d}$. In the course of the proof in [3], the authors prove that in $\mathcal{G}(n,p)$ subsets of vertices that are not "too large" cannot be "too dense" with high probability (where "too large" and "too dense" are quantified as functions of n and p). We follow this recipe combined with some structural results on the "typical" random regular graphs, and specifically, the number of edges spanned by a single vertex set of an appropriate size in $\mathcal{G}_{n,d}$, to prove the following concentration result.

Theorem 1.2. *For every positive constant ε there exists an integer $n_0 = n_0(\varepsilon)$ such that for every $n > n_0$ and $d = o(n^{1/5})$ there exists an integer $t = t(n,d,\varepsilon)$ such that*

$$\Pr\left[\chi(\mathcal{G}_{n,d}) \in \{t, t+1\}\right] \geq 1 - \varepsilon.$$

In other words, Theorem 1.2 states that for large enough values of n, the chromatic number of $\mathcal{G}_{n,d}$ for every $d = o(n^{1/5})$ w.h.p. takes one of two consecutive values. This result extends a previous result by Achlioptas and Moore [1], who prove the same concentration result for a smaller range of values of $d = d(n)$.

The rest of this paper is organized as follows. In Section 2 we give a brief introduction to random regular graphs and state some known results which will be of use in the succeeding sections. In Section 3 we perform a somewhat technical analysis of some structural properties of $\mathcal{G}_{n,d}$, and in particular analyze the distribution of edges for various orders of subsets of vertices and ranges of d. We then utilize these results to give a proof of Theorem 1.2 in Section 4.

2 Preliminaries

We start by analyzing and exploring the setting of random regular graphs and the techniques that we have to tackle problems in this probability space.

2.1 The Configuration Model

One of the major obstacles posed by the random regular graph model is the lack of a random generation process of all d-regular graphs for some given value of d. The following generation process, called the *Configuration Model* was introduced by Bollobás in [7], and implicitly by Bender and Canfield in [6]. Consider a set of dn elements (assuming dn is even), $\{e_1, ..., e_{dn}\}$. Let this set be partitioned into n *cells*, $c_i = \{e_j \mid d(i-1)+1 \leq j \leq di\}$, where $1 \leq i \leq n$. A perfect matching of the elements into $\frac{dn}{2}$ pairs is called a *pairing*. We denote by $\mathcal{P}_{n,d}$ the uniform probability space of these $(dn)!! = \frac{(dn)!}{\left(\frac{dn}{2}\right)!2^{dn/2}}$ possible pairings. Let $P \in \mathcal{P}_{n,d}$, and e_a be some element. We denote by e_a^P the element that is paired with e_a in the pairing P, that is if $\{e_a, e_b\} \in P$ then $e_a^P = e_b$ and $e_b^P = e_a$. We define a multigraph $G(P)$, where $V(G(P)) = [n]$, and for every pair $\{e_a, e_b\} \in P$ where $e_a \in c_i$ and $e_b \in c_j$ we add and edge connecting i to j in $G(P)$. For a general pairing P, $G(P)$ can obviously have loops and multiple edges. We can define a random d-regular multigraph model that assigns to each such multigraph G the accumulated probability of all pairings P from $\mathcal{P}_{n,d}$ such that $G(P) = G$. Although this probability space is not of simple d-regular graphs, but of d-regular multigraphs, it can be easily shown (see e.g. [32]) that all d-regular simple graphs on n vertices are equiprobable in this space. Now, one can generate a regular graph in $\mathcal{G}_{n,d}$ by sequentially generating random pairings $P \in \mathcal{P}_{n,d}$ and taking the first $G(P)$ that is a simple graph.

We define the event SIMPLE as the event that the pairing generated in $\mathcal{P}_{n,d}$ corresponds to a simple graph. By the uniformity of the two models, it follows that for any event $\mathcal{A}$ in $\mathcal{G}_{n,d}$, and any event $\mathcal{B}$ in $\mathcal{P}_{n,d}$ where $P \in \mathcal{B} \cap \textsc{Simple} \Leftrightarrow G(P) \in \mathcal{A}$ we have

$$\Pr[\mathcal{A}] = \Pr[\mathcal{B} \mid \textsc{Simple}] \leq \frac{\Pr[\mathcal{B}]}{\Pr[\textsc{Simple}]}. \tag{2}$$

McKay and Wormald [28] managed to compute the probability of SIMPLE for $d = o(\sqrt{n})$.

Theorem 2.1 (McKay and Wormald [28])**.** *For $d = o(\sqrt{n})$,*

$$\Pr[\textsc{Simple}] = \exp\left(\frac{1-d^2}{4} - \frac{d^3}{12n} + O\left(\frac{d^2}{n}\right)\right). \tag{3}$$

This estimate on the probability of SIMPLE combined with (2) will enable us to bound the probability of events in our regular graph model $\mathcal{G}_{n,d}$ based on bounds on the probabilities of the corresponding events in the “easier” Configuration Model $\mathcal{P}_{n,d}$.

We now introduce a few basic facts on $\mathcal{P}_{n,d}$ which will be useful for later computations on this model. Let e_a and e_b be two distinct elements of our set of dn elements. We define the indicator variable $I_{\{e_a,e_b\}}$ over $\mathcal{P}_{n,d}$ for the event that the pair $\{e_a, e_b\}$ is part of the random pairing. By the symmetry on the model, e_a is equally likely to be paired with any other element, thus $\Pr\left[I_{\{e_a,e_b\}} = 1\right] = \frac{1}{dn-1}$.

For any subset of indices $I \subseteq [n]$ we define the set $T_I = \bigcup_{i \in I} c_i$. Fix two such disjoint subsets, I and J, where $|I| = t$ and $|J| = s$. Given a random pairing P in $\mathcal{P}_{n,d}$, let $X_{T_I}(P)$ be the random variable counting the number of pairs in P that use only elements from T_I, and let $X_{T_I,T_J}(P)$ be the random variable counting the number of pairs in P that use one element from T_I and the other from T_J. By linearity of expectation we have that

$$\mathbf{E}[X_{T_I}] = \sum_{\{e_a,e_b\} \in \binom{T_I}{2}} \mathbf{E}\left[I_{\{e_a,e_b\}}\right] = \binom{dt}{2}\frac{1}{dn-1}; \tag{4}$$

$$\mathbf{E}[X_{T_I,T_J}] = \sum_{(e_a,e_b) \in T_I \times T_J} \mathbf{E}\left[I_{\{e_a,e_b\}}\right] = \frac{d^2 st}{dn-1}. \tag{5}$$

The expected value of X_{T_I,T_J} conditioned on the event $X_{T_I} = i$ can be also be computed using the linearity of expectation. There are $(dt - 2i)ds$ potential pairs, and the probability of each of these pairs to be in a random pairing is $\frac{1}{dn-dt}$, thus

$$\mathbf{E}[X_{T_I,T_J} \mid X_{T_I} = i] = \frac{ds(dt-2i)}{dn-dt}. \tag{6}$$

Let P and P' be two distinct pairings in $\mathcal{P}_{n,d}$. We write

$$P \sim P' \;\Leftrightarrow\; \exists e_a, e_b \quad P' = P \setminus \{\{e_a, e_a^P\}, \{e_b, e_b^P\}\} \cup \{\{e_a, e_b\}, \{e_a^P, e_b^P\}\}, \tag{7}$$

that is, $P \sim P'$ if P and P' differ only by a single simple *switch*. The following is a well known concentration result in the Configuration Model which makes use of martingales, and the Azuma-Hoeffding inequality (see e.g [4], [26]).

Theorem 2.2 ([32] Theorem 2.19)**.** *If X is a random variable defined on $\mathcal{P}_{n,d}$ such that $|X(P)-X(P')| \leq c$ whenever $P \sim P'$, then for all $\lambda > 0$*

$$\Pr\left[X \geq \mathbf{E}\left[X\right]+\lambda\right] \leq \exp\left(\frac{-\lambda^2}{dnc^2}\right);$$
$$\Pr\left[X \leq \mathbf{E}\left[X\right]-\lambda\right] \leq \exp\left(\frac{-\lambda^2}{dnc^2}\right).$$

A direct corollary of Theorem 2.2 and (2) gives us the following concentration result for $\mathcal{G}_{n,d}$.

Corollary 2.3 ([32])**.** *Let Y be a random variable defined on $\mathcal{G}_{n,d}$ such that $Y(G(P)) = X(P)$ for all $P \in$ Simple where X is a random variable defined on $\mathcal{P}_{n,d}$ that satisfies the conditions of Theorem 2.2. Then for all $\lambda > 0$*

$$\Pr\left[Y \geq \mathbf{E}\left[Y\right]+\lambda\right] \quad \leq \quad \frac{\exp\left(\frac{-\lambda^2}{dnc^2}\right)}{\Pr\left[\textsc{Simple}\right]};$$
$$\Pr\left[Y \leq \mathbf{E}\left[Y\right]-\lambda\right] \quad \leq \quad \frac{\exp\left(\frac{-\lambda^2}{dnc^2}\right)}{\Pr\left[\textsc{Simple}\right]}.$$

2.2 Working directly with $\mathcal{G}_{n,d}$ - the Switching Technique

A more recent approach, introduced by McKay in [27], that has come to be known as the *Switching Technique*, enables us to work directly on the random regular graph model, $\mathcal{G}_{n,d}$, without passing through the Configuration model, $\mathcal{P}_{n,d}$, which becomes futile for large values of $d(n)$. This technique enables us to overcome the basic difficulty of counting elements in $\mathcal{G}_{n,d}$ by giving an alternative "relative" counting technique. The basic operation is the following:

Definition 2.4. *Let G be a d-regular graph, and $S = (v_0, \ldots, v_{2r-1}) \subseteq V(G)$ be some ordered set of $2r$ vertices of G such that for any $1 \leq i \leq r$ $\{v_{2i}, v_{2i+1}\} \in E(G)$ and $\{v_{2i+1}, v_{2i+2}\} \notin E(G)$ (where the addition in the indices is done modulo $2r$). A* r-switch *of G by S is the removal of all r edges $\{v_{2i}, v_{2i+1}\}$ and the addition of the r non-edges $\{v_{2i+1}, v_{2i+2}\}$ to G as edges.*

The result of applying an r-switch operation on a d-regular graph is a d-regular graph, and this r-switch operation is obviously reversible. Now, let Q be some integer-valued graph parameter, and denote by $\mathcal{Q}_k$ the set of all graphs for which $Q(G) = k$. We can now bound the ratio $\frac{|\mathcal{Q}_k|}{|\mathcal{Q}_{k+1}|}$ by bounding the ratio of the number of r-switch operations that take us from a graph G' where $Q(G') = k+1$ to a graph G where $Q(G) = k$ and the number of r-switch operations that take us from a graph G where $Q(G) = k$ to a graph G' where $Q(G') = k+1$ for any integer r. For a more detailed explanation, the interested reader is referred to the proofs of Lemma 3.6 and Corollary 3.12, or to [23], [13], [19], and [20] where the Switching Technique is used extensively to prove various results on the random regular graph model. Throughout this paper, we will only make use of the 2-switch operation, but there are some cases (e.g. [23], [13]) where the more involved 3-switch was used to overcome technical difficulties.

3 Some structural properties of $\mathcal{G}_{n,d}$

We proceed to prove a series of standard probabilistic claims about the structure of a typical graph sampled from $\mathcal{G}_{n,d}$ which will be needed in the course of the proof of Theorem 1.2.

3.1 Edges spanned by a subset of vertices

In this section we will analyze the number of edges spanned by a set of vertices. Our main motivation for this is to be able to prove some technical lemmas on the distribution of edges spanned by subsets of these cardinalities, in the spirit very similar to lemmas proved in [3] for the model $\mathcal{G}(n,p)$, that will be used in the course of the proof of Theorem 1.2 in the succeeding section.

The *adjacency matrix* of a d-regular graph G on n vertices labeled by $[n]$, is the $n \times n$ binary matrix, $A(G)$, where $A(G)_{ij} = 1$ iff $(i,j) \in E(G)$. As $A(G)$ is real and symmetric it has an orthogonal basis of real eigenvectors and all its eigenvalues are real. We denote the eigenvalues of $A(G)$ in descending order by $\lambda_1 \geq \lambda_2 \ldots \geq \lambda_n$, where $\lambda_1 = d$ and its corresponding eigenvector is j_n (the $n \times 1$ all ones vector). Finally, let $\lambda = \lambda(G) = \max\{|\lambda_2(G)|, |\lambda_n(G)|\}$, and call such a graph G an (n,d,λ)-graph. For an extensive survey of fascinating properties of (n,d,λ)-graphs the reader is referred to [22]. The celebrated expander mixing lemma (see e.g. [4] or [11]) states roughly that the smaller λ is, the more random-like is the graph. Here we present a simple variant of this lemma, which bounds the number of edges spanned by any subset of vertices in an (n,d,λ)-graph.

Lemma 3.1 (The Expander Mixing Lemma - Corollary 9.2.6 in [4])**.** *For every (n,d,λ)-graph $G=(V,E)$, every subset of vertices $U \subseteq V$ satisfies*

$$\left| e(U) - \binom{|U|}{2}\frac{d}{n} \right| \leq \lambda |U|.$$

Broder et al. [10], extending a previous result of Friedman, Kahn and Szemerédi [15], who used the so-called trace method, estimate the "typical" second eigenvalue of the $\mathcal{G}_{n,d}$ for $d = o(\sqrt{n})$.

Theorem 3.2 (Broder et al. [10])**.** *For $d = o(\sqrt{n})$, w.h.p. $\lambda(\mathcal{G}_{n,d}) = O(\sqrt{d})$.*

With Lemma 3.1 and Theorem 3.2 at hand, the following theorem is an immediate consequence.

Theorem 3.3. *For every $d = o(\sqrt{n})$ if $G=(V,E)$ is sampled from $\mathcal{G}_{n,d}$ then w.h.p. every subset of vertices $U \subseteq V$ satisfies*

$$\left| e(U) - \binom{|U|}{2}\frac{d}{n} \right| = O(|U|\sqrt{d}). \tag{8}$$

The authors in [5] give an alternative proof to Theorem 3.3 based purely on combinatorial techniques and not relying on the spectral properties of the random graph $\mathcal{G}_{n,d}$.

Although Theorem 3.3 bounds the number of edges spanned by a single set of vertices, the bound obtained does not meet our needs to prove Theorem 1.2, and we will need tighter bounds. Therefore, we will start by analyzing the distribution of the number of pairs spanned by a single set of indices in $\mathcal{P}_{n,d}$.

Let I be a subset of indices from $[n]$ of cardinality $|I| = t$. Denote by $\mathcal{C}_{T_I,k}$ be the set of all pairings in $\mathcal{P}_{n,d}$ where there are exactly k pairs that use only elements from T_I. The possible values for k are

$$\max\left\{0, dt - \frac{dn}{2}\right\} \leq k \leq \frac{dt}{2}. \tag{9}$$

To compute the cardinality of the set $\mathcal{C}_{T_I,k}$ we count all possible ways to choose $2k$ elements from T_I and pair them up, pair the rest of the elements from T_I to elements outside T_I, and pair the rest of the elements. It follows that

$$|\mathcal{C}_{T_I k}| = \binom{dt}{2k}\frac{(2k)!}{k!2^k}\binom{dn-dt}{dt-2k}(dt-2k)!\frac{(dn-2dt+2k)!}{\left(\frac{dn}{2}-dt+k\right)!2^{\frac{dn}{2}-dt+k}}.$$

Now, set

$$f(k) = \frac{|\mathcal{C}_{T_I,k+1}|}{|\mathcal{C}_{T_I,k}|} = \frac{1}{4}\cdot\frac{1}{k+1}\cdot\frac{(dt-2k)(dt-2k+1)}{\frac{dn}{2}-dt+k+1}, \tag{10}$$

and for convenience extend $f(k)$ to be a real valued function. Note that if $k_1 > k_2$, both values in the range of k then

$$\frac{f(k_1)}{f(k_2)} = \frac{k_2+1}{k_1+1}\cdot\frac{dt-2k_1}{dt-2k_2}\cdot\frac{dt-2k_1+1}{dt-2k_2+1}\cdot\frac{\frac{dn}{2}-dt+k_2+1}{\frac{dn}{2}-dt+k_1+1} \leq \frac{k_2+1}{k_1+1}. \tag{11}$$

Lemma 3.4. *For any subset $I \subseteq [n]$ where $|I| = t$, $|\mathcal{C}_{T_I,k}|$ is monotonically increasing from $k = \max\left\{0, dt - \frac{dn}{2}\right\}$ to $\lfloor \mathbf{E}[X_{T_I}]\rfloor$, and monotonically decreasing from $k = \lceil \mathbf{E}[X_{T_I}]\rceil$ to $k = \frac{dt}{2}$.*

Proof. Set $k_0 = \mathbf{E}[X_{T_I}]$, and let k' satisfy $f(k') = 1$. Trivially, $f(k)$ is monotonically decreasing when k ranges from $\max\left\{0, dt - \frac{dn}{2}\right\}$ to $\frac{dt}{2}$, therefore it is enough to show $k_0 - 1 \leq k' \leq k_0$. By solving for k' we get

$$k' = \frac{(dt)^2 - 2dn + 3t - 4}{2dn+6} = \frac{dt(dt+3)}{2(dn+3)} - \frac{dn+2}{dn+3}.$$

Applying (4) we have

$$k_0 - k' = \frac{dt(dt-1)}{2(dn-1)} - \frac{dt(dt+3) - 2(dn+2)}{2(dn+3)} = \frac{2dt(dt-dn) + 2(dn+2)(dn-1)}{(dn-1)(dn+3)}.$$

The above function gets its minimal value for $t = \frac{n}{2}$ thus $k_0 - k' \geq \frac{\frac{3}{2}(dn)^2+2dn-4}{(dn-1)(dn+3)} \geq 0$. On the other hand, again, by applying (4),

$$k' - k_0 + 1 = \frac{dt(dt+3)}{2(dn+3)} - \frac{dn+2}{dn+3} - \frac{dt(dt-1)}{2(dn-1)} + 1 \geq \frac{2dt(\frac{dn}{2}-dt)}{(dn-1)(dn+3)} \geq 0,$$

completing the proof of our claim. □

Lemma 3.5. *Let $P \in \mathcal{P}_{n,d}$ and $I \subseteq [n]$ be a fixed set of indices. Then,*

(a) *for any value of Δ,* $\Pr\left[|X_{T_I} - \mathbf{E}[X_{T_I}]| \geq \Delta\right] \leq \frac{d|I|}{2} e^{-\frac{\Delta^2}{4\mathbf{E}[X_{T_I}]+2\Delta+4}}$;

(b) *if $\Delta > \mathbf{E}[X_{T_I}]$ then* $\Pr\left[|X_{T_I} - \mathbf{E}[X_{T_I}]| \geq \Delta\right] \leq \frac{d|I|}{2}\cdot e^{-\frac{\Delta}{2}\ln\frac{2\mathbf{E}[X_{T_I}]+\Delta+2}{2\mathbf{E}[X_{T_I}]+2}}$.

Proof. Set $|I| = t$, $k_0 = \mathbf{E}[X_{T_I}] = \binom{dt}{2}\frac{1}{dn-1}$, $k_1 = k_0 + \Delta$, and $k_2 = k_0 - \Delta$. Following Lemma 3.4 we know that $f(k_0 - 1) > 1 > f(k_0)$. We start by proving claim (a).

$$\Pr[X_{T_I} = k_1] \leq \frac{|\mathcal{C}_{T_I,k_1}|}{|\mathcal{C}_{T_I,k_0}|} = f(k_0)^{\Delta}\cdot\prod_{i=k_0}^{k_1-1}\frac{f(i)}{f(k_0)} \leq \left(\frac{f\left(k_0+\frac{\Delta}{2}\right)}{f(k_0)}\right)^{\frac{\Delta}{2}} \leq \left(\frac{k_0+1}{k_0+\frac{\Delta}{2}+1}\right)^{\frac{\Delta}{2}}, \tag{12}$$

where the last inequality follows from (11). Similarly,

$$\Pr\left[X_{T_I}=k_2\right] \le \frac{|\mathcal{C}_{T_I,k_2}|}{|\mathcal{C}_{T_I,k_0}|} = \left(\frac{1}{f(k_0)-1}\right)^{\Delta} \cdot \prod_{i=k_2}^{k_0-1} \frac{f(k_0-1)}{f(i)} \le \left(\frac{f\left(k_0-1\right)}{f(k_0-\frac{\Delta}{2}-1)}\right)^{\frac{\Delta}{2}} \le \left(\frac{k_0-\frac{\Delta}{2}}{k_0}\right)^{\frac{\Delta}{2}}, \quad (13)$$

where, again, the last inequality follows from (11). To prove claim (a) we note that $\Pr\left[X_{T_I}=k_1\right] \le e^{-\frac{\Delta^2}{4k_0+2\Delta+4}}$. By Lemma 3.4 we know that $|\mathcal{C}_{T_I,k}|$ is monotonically increasing for $k=\max\left\{0, dt-\frac{dn}{2}\right\}$ to k_0, and monotonically decreasing for $k=k_0$ to $\frac{dt}{2}$, therefore, by (12) and (13), we have

$$\Pr\left[|X_{T_I}-k_0| \ge \Delta\right] = \Pr\left[X_{T_I} \ge k_1\right] + \Pr\left[X_{T_I} \le k_2\right] \le$$
$$\left(\frac{dt}{2}-k_1\right) \cdot e^{-\frac{\Delta^2}{4k_0+2\Delta+4}} + k_2 \cdot e^{-\frac{\Delta^2}{4k_0}} \le \frac{dt}{2} e^{-\frac{\Delta^2}{4k_0+2\Delta+4}}.$$

The proof of claim (b) is very similar. We note that $\Pr\left[X_{T_I}=k_1\right] \le e^{-\frac{\Delta}{2}\ln\frac{2k_0+\Delta+2}{2k_0+2}}$ and that $\Pr\left[X_{T_I}=k_2\right] \le e^{-\frac{\Delta}{2}\ln\frac{2k_0}{2k_0-\Delta}}$. By Lemma 3.4 we know that $|\mathcal{C}_{T_I,k}|$ is monotonically increasing for $k=\max\left\{0, dt-\frac{dn}{2}\right\}$ to k_0, and monotonically decreasing for $k=k_0$ to $\frac{dt}{2}$, therefore, by (12) and (13), we have

$$\Pr\left[|X_{T_I}-k_0| \ge \Delta\right] = \Pr\left[X_{T_I} \ge k_1\right] \le \frac{dt}{2} \cdot e^{-\frac{\Delta}{2}\ln\frac{2k_0+\Delta+2}{2k_0+2}}.$$

□

We now apply the Switching Technique to analyze the distribution of edges in a single subset of vertices. Fix a set $U \subseteq V(G)$ of size u. Let $\mathcal{C}_i$ denote the set of all d-regular graphs where exactly i edges have both ends in U.

Lemma 3.6. *For every $d, u = o(n)$,*

$$\frac{|\mathcal{C}_i|}{|\mathcal{C}_{i-1}|} \le \frac{1}{i}\binom{u}{2}\frac{d}{n}\left(1+\frac{2u+4d}{n}\right).$$

Proof. Let F_i be a bipartite graph whose vertex set is composed of $\mathcal{C}_i \cup \mathcal{C}_{i-1}$ and $\{G_1, G_2\} \in E(F_i)$ if and only if $G_1 \in \mathcal{C}_i$, $G_2 \in \mathcal{C}_{i-1}$, and G_2 can be derived from G_1 by a 2-switch (and vice versa). Let $G \in \mathcal{C}_{i-1}$. To go to a graph in $\mathcal{C}_i$ with a single 2-switch operation, we need to choose the vertices $a, b \in U$, where $\{a,b\} \notin E(G)$, and $x, y \in V(G) \setminus U$ such that $\{a,x\},\{b,y\} \in E(G)$ and $\{x,y\} \notin E(G)$, and perform the switch as follows: $\{a,x\},\{b,y\} \to \{a,b\},\{x,y\}$. This yields the upper bound $d_{F_i}(G) \le \binom{u}{2}d^2$.

In order to give a lower bound on $d_{F_i}(G')$, where $G' \in \mathcal{C}_i$, we note that any choice of $a, b \in U$ such that $\{a,b\} \in E(G)$ and an edge $\{x,y\} \in E(G')$ such that $x, y \notin U$ and $x, y \notin N_{G'}(\{a,b\})$, gives us two ways to perform the switch: $\{a,b\},\{x,y\} \to \{a,x\},\{b,y\}$ or $\{a,b\},\{x,y\} \to \{a,y\},\{b,x\}$ resulting in a graph in C_{i-1}. We note that $|N_{G'}(U)| \le ud$ and $|N_{G'}(\{a,b\})| \le 2d$, thus $d_{F_i}(G') \ge 2i\left(\frac{dn}{2}-ud-d^2\right)$.

Combining the upper and lower bounds gives us

$$|\mathcal{C}_i| \cdot 2i\left(\frac{dn}{2}-ud-2d^2\right) \le \sum_{G' \in \mathcal{C}_i} d_{F_i}(G') = e(F_i) = \sum_{G \in \mathcal{C}_{i-1}} d_{F_i}(G) \le |\mathcal{C}_{i-1}| \cdot \binom{u}{2}d^2.$$

Using the fact that $\frac{dn}{2}-ud-2d^2 > ud+2d^2 > 0$ we conclude that

$$\frac{|\mathcal{C}_i|}{|\mathcal{C}_{i-1}|} \le \frac{1}{i} \cdot \binom{u}{2}\frac{d}{n} \cdot \left(1+\frac{ud+2d^2}{\frac{dn}{2}-ud-2d^2}\right) \le \frac{1}{i} \cdot \binom{u}{2}\frac{d}{n} \cdot \left(1+\frac{2u+4d}{n}\right).$$

□

The previous lemma give us the necessary ingredient to prove the following concentration result for $e(U)$.

Corollary 3.7. *For every $d, u = o(n)$, $\Delta \geq \binom{u}{2}\frac{d}{n}$ and fixed set of u vertices U in $\mathcal{G}_{n,d}$,*
(a)$\Pr\left[e(U) \geq \mathbf{E}\left[e(U)\right] + \Delta\right] \leq \frac{du}{2} \cdot \exp\left(-\frac{\Delta^2}{4\mathbf{E}[e(U)]+2\Delta} + o(\Delta)\right)$;
(b)$\Pr\left[e(U) \geq \mathbf{E}\left[e(U)\right] + \Delta\right] \leq \frac{du}{2} \cdot \exp\left(-\frac{\Delta}{2}\ln\left(1 + \frac{\Delta}{2\mathbf{E}[e(U)]}\right) + o(\Delta)\right)$.

Proof. Set $k_0 = \mathbf{E}\left[e(U)\right] = \binom{u}{2}\frac{d}{n}$ and $k = k_0 + \Delta$. By Lemma 3.6 it follows that $\frac{|\mathcal{C}_i|}{|\mathcal{C}_{i-1}|} \leq \frac{1}{i}\binom{u}{2}\frac{d}{n}(1 + o(1))$, and that $|\mathcal{C}_i|$ is monotonically decreasing for $i \geq 2\binom{u}{2}\frac{d}{n}$ and hence,

$$\Pr\left[e(U) \geq k\right] \leq \sum_{j=k}^{\frac{du}{2}} \frac{|\mathcal{C}_j|}{|\mathcal{C}_{k_0}|} \leq \frac{du}{2} \cdot \prod_{i=k_0+1}^{k} \frac{|\mathcal{C}_i|}{|\mathcal{C}_{i-1}|} =$$
$$\frac{du}{2} \cdot (1 + o(1))^{\Delta} \prod_{i=1}^{\Delta} \frac{k_0}{k_0 + i} \leq \frac{du}{2} \cdot e^{o(\Delta)} \cdot \left(\frac{k_0}{k_0 + \frac{\Delta}{2}}\right)^{\frac{\Delta}{2}}. \tag{14}$$

On the one hand, the right term of (14) equals $\frac{du}{2} \cdot \exp\left(-\frac{\Delta}{2}\ln\frac{2k_0+\Delta}{2k_0} + o(\Delta)\right)$, proving claim (b). On the other hand, $\frac{du}{2} \cdot e^{o(\Delta)} \cdot \left(\frac{k_0}{k_0+\frac{\Delta}{2}}\right)^{\frac{\Delta}{2}} \leq \frac{du}{2} \cdot \exp\left(-\frac{\Delta^2}{4k_0+2\Delta} + o(\Delta)\right)$, proving claim (a). □

The concentration of $e(U)$ provided by Corollary 3.7 enables us to obtain some upper bounds on the number of edges spanned by subsets of vertices of different cardinalities.

Corollary 3.8. *For every constant $C > 0$ and $d = o(n^{1/5})$, w.h.p. every subset of $u \leq C\sqrt{nd^3}$ vertices of $\mathcal{G}_{n,d}$ spans less than $5u$ edges.*

Proof. Let C be some positive constant. For every vertex subset U of cardinality $u \leq C\sqrt{nd^3} = o(n^{4/5})$, let $\Delta = \Delta(U) = 4u$, and $k_0 = \mathbf{E}\left[e(U)\right] = \binom{u}{2}\frac{d}{n} \ll \Delta$. Note that we can assume that $u > 5$, since for $u \leq 5$ the claim is trivial. Summing over all possible values of u, and applying Corollary 3.7 (b) and the union bound, we have that the probability that there exists a set U of at most $C\sqrt{nd^3}$ vertices, that spans more than $4u + k_0 < 5u$ edges is bounded by

$$\sum_{u=6}^{C\sqrt{nd^3}} \binom{n}{u}\frac{du}{2} \cdot \exp\left(-2u\ln\left(1 + \frac{4n}{(u-1)d}\right) + o(u)\right) \leq$$
$$\sum_{u=6}^{C\sqrt{nd^3}} \exp\left(-2u\ln\frac{4n}{ud} + u\ln\frac{n}{u} + u + \ln(du) + o(u)\right) \leq$$
$$n \cdot \exp\left(-5\ln n\right) = o(1).$$

□

Corollary 3.9. *For every $d = o(n)$ w.h.p. every subset of $u \geq \frac{n\ln n}{d}$ vertices of $\mathcal{G}_{n,d}$ spans less than $\frac{5u^2d}{n}$ edges.*

Proof. For every vertex subset U of cardinality $u \geq \frac{n\ln n}{d}$, let $\Delta = \Delta(U) = \frac{4u^2d}{n}$, and $k_0 = \mathbf{E}\left[e(U)\right] = \binom{u}{2}\frac{d}{n} < \frac{u^2d}{2n}$. Summing over all values of u, and applying Corollary 3.7 (a) and the union bound, we have that the probability that there exists a set U of at least $\frac{n \ln n}{d}$ vertices, that spans more than $\frac{4u^2 d}{n} + k_0 < \frac{5u^2 d}{n}$ vertices is bounded by

$$\sum_{u=\frac{n \ln n}{d}}^{n} \binom{n}{u} \frac{du}{2} \cdot \exp\left(-\frac{8u^2 d}{5n} + o\left(\frac{u^2 d}{n}\right)\right) \leq$$
$$\sum_{u=\frac{n \ln n}{d}}^{n} \exp\left(-\frac{3u^2 d}{2n} + \ln(du) + u + u \ln \frac{n}{u}\right) \leq$$
$$\sum_{u=\frac{n \ln n}{d}}^{n} \exp\left(-\frac{3}{2} u \ln n + \ln(du) + u + u \ln \frac{n}{u}\right) \leq$$
$$n \cdot \exp\left(-\frac{n \ln^2 n}{3d}\right) = o(1).$$

□

3.2 Other structural properties of $\mathcal{G}_{n,d}$

In the course of the proof of Theorem 1.2 we will need some other asymptotic properties of $\mathcal{G}_{n,d}$ which we proceed to prove. We note that Lemma 3.10 and Corollary 3.12 can also be obtained using a general result of Kim, Sudakov and Vu [20, Theorem 1.3] using density considerations, but we prefer to give direct proofs using the Switching Technique, similarly to what we have previously seen.

Lemma 3.10. *For every $d = o(n^{1/5})$ w.h.p. every vertex participates in at most 4 triangles in $\mathcal{G}_{n,d}$.*

Proof. Fix some vertex $v \in [n]$, and let $\mathcal{D}_i$ denote the set of all d-regular graphs on the set of vertices $[n]$ where $e(N_G(v)) = i$. We denote by H_i the bipartite graph whose vertex set is composed of $\mathcal{D}_i \cup \mathcal{D}_{i-1}$ and $\{G_1, G_2\} \in E(H_i)$ if and only if $G_1 \in \mathcal{D}_i$, $G_2 \in \mathcal{D}_{i-1}$, G_2 can be derived from G_1 by a 2-switch (and vice versa), and $N_{G_1}(v) = N_{G_2}(v)$ (that is, we require that performing a 2-switch does not affect the set of v's neighbors). This additional restriction enables us to apply Lemma 3.6[1] to get that $\frac{|\mathcal{D}_i|}{|\mathcal{D}_{i-1}|} \leq \frac{1}{i}\binom{d}{2}\frac{d}{n}\left(1 + \frac{6d}{n}\right) < \frac{d^3}{in} = o(n^{-2/5})$. It follows that for the given values of d, $|\mathcal{D}_i|$ is monotonically decreasing as i goes from 0 to $\binom{d}{2}$. The probability that there are more than four edges spanned by $N_G(v)$ can be bounded by

$$\Pr\left[e(N_G(v)) \geq 5\right] \leq \sum_{k=5}^{\binom{d}{2}} \frac{|\mathcal{D}_k|}{|\mathcal{D}_0|} \leq \frac{|\mathcal{D}_5|}{|\mathcal{D}_0|} \cdot \sum_{j=0}^{\binom{d}{2}-5} \left(n^{-2/5}\right)^j =$$
$$(1 + o(1)) \prod_{i=1}^{5} \frac{|\mathcal{D}_i|}{|\mathcal{D}_{i-1}|} \leq (1 + o(1)) \left(n^{-2/5}\right)^5 = o(n^{-1}),$$

and applying the union bound over all vertices from $[n]$ completes the proof. □

Fix $U, W \subseteq V(G)$ to be two not necessarily disjoint subsets of vertices of cardinality u and w respectively. Let $\mathcal{E}_i$ denote the subset of all d-regular graphs where there are exactly i edges with one endpoint in U and the other in W. The following proof follows closely the proof of Lemma 3.6.

[1]This statement requires some caution. Recall that in Lemma 3.6, if taking U as the set of neighbors of some vertex v, then counting the number of 2-switches that take a graph from $\mathcal{C}_i$ to $\mathcal{C}_{i-1}$ the neighbor set of v may change. Nevertheless, the bound in the proof avoids counting edges leaving U, so it applies to this case as well.

Lemma 3.11. *For $w \leq u \leq \frac{n}{10}$ and $d = o(n)$,*

$$\frac{|\mathcal{E}_i|}{|\mathcal{E}_{i-1}|} < \frac{2uwd}{in}.$$

Moreover, $|\mathcal{E}_i|$ is monotonically decreasing for $i > \frac{2uwd}{n}$.

Proof. Let H_i be a bipartite graph whose vertex set is composed of $\mathcal{E}_i \cup \mathcal{E}_{i-1}$ and $\{G_1, G_2\} \in E(H_i)$ if and only if $G_1 \in \mathcal{E}_i$, $G_2 \in \mathcal{E}_{i-1}$, and G_2 can be derived from G_1 by a 2-switch (and vice versa). Let $G \in \mathcal{E}_{i-1}$. To give an upper bound on $d_{H_i}(G)$ we need to find $a \in U$ and $b \in W$ such that $\{a,b\} \notin E(G)$, and find $x \in V(G) \setminus W$ and $y \in V(G) \setminus U$ such that $\{a,x\},\{b,y\} \in E(G)$ and $\{x,y\} \notin E(G)$. Now we can perform the switch $\{a,x\},\{b,y\} \to \{a,b\},\{x,y\}$ resulting in a graph in $\mathcal{E}_i$. This gives us an upper bound $d_{H_i}(G) \leq uwd^2$.

On the other hand, let $G' \in \mathcal{E}_i$. To derive a lower bound on $d_{H_i}(G')$ we can first choose $a \in U$ and $b \in W$ such that $\{a,b\} \in E(G')$ in i ways. Next we find an edge $\{x,y\} \in E(G')$ such that $x, y \notin U \cup W$ and $x, y \notin N_{G'}(\{a,b\})$. The number of edges with an endpoint in $U \cup W$ is at most $(u+w)d$ and $|N_{G'}(\{a,b\})| \leq 2d$. After choosing such x and y we have two ways to perform the switch: $\{a,b\},\{x,y\} \to \{a,x\},\{b,y\}$ or $\{a,b\},\{x,y\} \to \{a,y\},\{b,x\}$ which gives us a graph in D_{i-1}. It follows that $d_{H_i}(G') \geq 2i\left(\frac{dn}{2} - (u+w)d - 2d^2\right)$.

We now compute the ratio, much like we have previously seen.

$$|\mathcal{E}_i| \cdot 2i\left(\frac{dn}{2} - (u+w)d - 2d^2\right) \leq \sum_{G' \in \mathcal{E}_i} d_{H_i}(G') = \sum_{G \in \mathcal{E}_{i-1}} d_{H_i}(G) \leq |\mathcal{E}_{i-1}| \cdot uwd^2.$$

Since $\frac{dn}{2} - (u+w)d - 2d^2 > (u+w)d + 2d^2 > 0$, we have

$$\frac{|\mathcal{E}_i|}{|\mathcal{E}_{i-1}|} \leq \frac{uwd}{in} \cdot \left(\frac{dn}{dn - 2(u+w) - 4d^2}\right) < \frac{2uwd}{in},$$

as claimed. □

Corollary 3.12. *For every $d = o(n^{1/5})$ w.h.p. the number of paths of length three between any two distinct vertices in $\mathcal{G}_{n,d}$ is at most 4.*

Proof. Fix $u, w \in [n]$ be two distinct vertices. Let $\mathcal{F}_i$ denote the set of all d-regular graphs, where there are exactly i paths of length three between u and w. We denote by H_i the bipartite graph whose vertex set is composed of $\mathcal{F}_i \cup \mathcal{F}_{i-1}$ and $\{G_1, G_2\} \in E(H_i)$ if and only if $G_1 \in \mathcal{F}_i$, $G_2 \in \mathcal{F}_{i-1}$, G_2 can be derived from G_1 by a 2-switch (and vice versa), $N_{G_1}(u) = N_{G_2}(u)$ and $N_{G_1}(w) = N_{G_2}(w)$ (that is, performing the 2-switch does not affect the neighbor sets of u and w). We may therefore set $U = N_{G_1}(u)$ and $W = N_{G_1}(w)$ and use Lemma 3.11[2] to derive that $\frac{|\mathcal{F}_i|}{|\mathcal{F}_{i-1}|} < \frac{2d^3}{in} = o(n^{-2/5})$. It follows, that for the given values of d, $\mathcal{F}_i$ is monotonically decreasing. Now, we bound the probability that there are more than 5 paths of length three between u and w as follows:

$$\begin{aligned}
&\Pr[\text{There are more than 4 paths of length three between } u \text{ and } w] \leq \\
&\sum_{k=5}^{n^2} \frac{|\mathcal{F}_k|}{|\mathcal{F}_0|} \leq \frac{|\mathcal{F}_5|}{|\mathcal{F}_0|} \cdot \sum_{j=0}^{n^2-5} \left(n^{-2/5}\right)^j \leq \\
&(1+o(1)) \cdot \prod_{i=1}^{5} \frac{|\mathcal{F}_i|}{|\mathcal{F}_{i-1}|} \leq (1+o(1)) \cdot o\left(n^{-2/5}\right)^5 = o(n^{-2}).
\end{aligned}$$

[2]Same as previous footnote. The way we count the numbers of performing the 2-switch in Lemma 3.11 will not affect the sets of neighbors of u and w.

Using the union bound over all pairs of vertices from $[n]$ completes our proof. $\square$

4 Proof of Theorem 1.2

In this section we prove Theorem 1.2. The proof will follow closely the proof of Alon and Krivelevich in [3] for the random graph model $\mathcal{G}(n,p)$.

Let us introduce the notion of graph choosability. A graph $G = (\{v_1, \dots, v_n\}, E)$ is $\mathcal{S}$-*choosable*, for a family of color lists $\mathcal{S} = \{S_1, ..., S_n\}$, if there exists a proper coloring f of G that satisfies for every $1 \le i \le n$, $f(v_i) \in S_i$. G is k-*choosable*, for some positive integer k, if G is $\mathcal{S}$-choosable for any family $\mathcal{S}$ such that $|S_i| = k$ for every $i \in \{1, ..., n\}$. The *choice number* of G, which is denoted by $ch(G)$, is the minimum integer k such that G is k-choosable.

A graph is d-*degenerate* if every subgraph of it contains a vertex of degree at most d. The following gives a trivial upper bound on the choice number of a graph.

Proposition 4.1. *Every d-degenerate graph is $(d+1)$-choosable.*

For every $1 \ge \varepsilon > 0$ define $\tau = \tau(n, d, \varepsilon)$ to be the least integer for which

$$\Pr\left[\chi(\mathcal{G}_{n,d}) \le \tau\right] \ge \varepsilon \tag{15}$$

and let $Y(G)$ be the random variable defined over $\mathcal{G}_{n,d}$ that denotes the minimal size of a set of vertices S for which $G \setminus S$ can be τ-colored.

Lemma 4.2. *For every integer n, $\varepsilon > 0$ and $d = o(\sqrt{n})$, there exists a constant, $C = C(\varepsilon)$ such that*

$$\Pr\left[Y(\mathcal{G}_{n,d}) \ge C\sqrt{nd^3}\right] \le \varepsilon.$$

Proof. Let G be a random graph in $\mathcal{G}_{n,d}$ and fix some $\varepsilon > 0$. By the minimality of τ it follows that $\Pr\left[\chi(G) < \tau\right] < \varepsilon$. Define a proper coloring of the multigraph generated by a pairing P, $G(P)$, as a proper coloring of the this multigraph discarding its loops, and let Y' be the random variable defined over $\mathcal{P}_{n,d}$ such that for every $P \in \mathcal{P}_{n,d}$ $Y'(P)$ is the minimal size of a set of vertices S' for which $G(P) \setminus S'$ can be τ-colored. Obviously, for $P \in \textsc{Simple}$, $Y(G(P)) = Y'(P)$. Let $P_0 \in \textsc{Simple}$, and denote by $P_0(m)$ the subset of pairs from P_0 which covers all of the first m elements. We define the following random variables over $\mathcal{P}_{n,d}$:

$$\forall 1 \le m \le dn \qquad Y'_m(P_0) = \mathbf{E}_{P \in \mathcal{P}_{n,d}}\left[Y'(P) \mid P_0(m) \subseteq P\right], \tag{16}$$

i.e., $Y'_m(P_0)$ is the expectation of the size of S' conditioned on all the pairings in $\mathcal{P}_{n,d}$ that have the same first m pairs as P_0. $Y'_0(P_0), ..., Y'_{dn-1}(P_0)$ is indeed a martingale, and the random variable Y' satisfies $|Y'(P) - Y'(P')| \le 2$ for all $P \sim P'$ (recalling (7)). Setting $\lambda = 2\sqrt{d \ln\left(\varepsilon \cdot \Pr\left[\textsc{Simple}\right]\right)^{-1}}$, and applying Corollary 2.3 implies the following concentration result on Y:

$$\Pr\left[Y(G) \ge \mathbf{E}\left[Y(G)\right] + \lambda\sqrt{n}\right] \le \frac{e^{-\lambda^2/4d}}{\Pr\left[\textsc{Simple}\right]} = \varepsilon;$$

$$\Pr\left[Y(G) \le \mathbf{E}\left[Y(G)\right] - \lambda\sqrt{n}\right] \le \frac{e^{-\lambda^2/4d}}{\Pr\left[\textsc{Simple}\right]} = \varepsilon. \tag{17}$$

Notice that $\Pr[Y(G) = 0] = \Pr[\chi(G) \le \tau] > \varepsilon$, therefore, by (17), $\mathbf{E}[Y(G)] < \lambda\sqrt{n}$, and thus,

$$\Pr\left[Y(G) \ge 2\lambda\sqrt{n}\right] \le \Pr\left[Y(G) \ge \mathbf{E}[Y(G)] + \lambda\sqrt{n}\right] \le \varepsilon. \tag{18}$$

Theorem 2.1 implies an upper bound on λ,

$$\lambda = 2\sqrt{-d(\ln\varepsilon + \ln\Pr[\textsc{Simple}])} \le 2\sqrt{-d\left(\ln\varepsilon + \left(\frac{1-d^2}{4} - \frac{d^3}{12n} - O\left(\frac{d^2}{n}\right)\right)\right)} = O\left(\frac{d^{3/2}}{2}\right).$$

Returning to (18), $\Pr\left[Y(G) \ge C\sqrt{nd^3}\right] \le \varepsilon$ where C is some constant depending on ε, as claimed. □

Coming to prove Theorem 1.2, we can use the previous result of Achlioptas and Moore [1][3], whose proof of the same two-point concentration result can be shown to go through for $d = n^{1/9-\delta}$ for any $\delta > 0$, and we can thus assume that $d > n^{1/10}$. Let $\Gamma(n, d)$ denote the set of d-regular graphs on n vertices that satisfy the following properties:

1. For every constant $C > 0$, every subset of $u \le C\sqrt{nd^3}$ vertices spans less than $5u$ edges.
2. For every constant $C > 0$, every subset of $u \le Cn^{9/10}$ vertices spans less than $O(u\sqrt{d})$ edges.
3. Every subset of $u \ge \frac{n\ln n}{d}$ vertices spans less than $O\left(\frac{u^2 d}{n}\right)$ edges.
4. For every vertex v, the number of edges spanned by $N(v)$ is at most 4.
5. The number of paths of length three between any two vertices is at most 4.

We have already proved that all properties of $\Gamma(n, d)$ occur w.h.p. in $\mathcal{G}_{n,d}$ for $d = o(n^{1/5})$ (Corollary 3.8, Theorem 3.3, Corollary 3.9, Lemma 3.10 and Corollary 3.12 respectively), and by Theorem 1.1 we know that for $n^{1/10} < d \ll n^{1/5}$ w.h.p. $\chi(\mathcal{G}_{n,d}) \ge \frac{d}{2\ln d} > 10$. The proof of Theorem 1.2 now follows from the following deterministic proposition, Proposition 4.3, by taking $t = \tau(n, d, \frac{\varepsilon}{3})$, since by (15) and Lemma 4.2 we have:

$$\begin{aligned}&\Pr[\chi(\mathcal{G}_{n,d}) < t \text{ or } \chi(\mathcal{G}_{n,d}) > t+1] \le \\ &\Pr[\mathcal{G}_{n,d} \notin \Gamma(n,d)] + \frac{\varepsilon}{3} + \Pr\left[Y(\mathcal{G}_{n,d}) \ge c\sqrt{nd^3}\right] \le \varepsilon.\end{aligned}$$

Proposition 4.3. *There exists a positive integer n_0 such that for every $n \ge n_0$, $n^{1/10} < d \ll n^{1/5}$, if $G \in \Gamma(n, d)$ such that $\chi(G) \ge t \ge \frac{d}{2\ln d}$ and that there is a subset $U_0 \subseteq V$ of size $|U_0| = O(\sqrt{nd^3})$ such that $G[V \setminus U_0]$ is t-colorable, then G is $(t+1)$-colorable.*

Proof. First, we find a subset $U \subseteq V$ of size $O(\sqrt{nd^3})$ including U_0 such that every vertex $v \in V \setminus U$ has at most 50 neighbors in U. To find such a set U we proceed as follows. We start with $U = U_0$, and as long as there exists a vertex $v \in V \setminus U$ with at least 50 neighbors in U, we add v to U and iterate the process again. After r iterations of this process we have $|U| \le c\sqrt{nd^3} + r$ and $e(U) \ge 50r$. It follows that the number of iterations is at most $\frac{c\sqrt{nd^3}}{9}$, since otherwise, we would get a set of $\frac{10c\sqrt{nd^3}}{9}$ vertices spanning at least $\frac{50c\sqrt{nd^3}}{9}$

[3] In [1] the authors formally state this claim for the case where d is a constant (Theorem 1 in their paper), but in the first paragraph of Section 2 of their paper they state it for $d = O\left(n^{1/7-\delta}\right)$ for all $\delta > 0$. It appears that in their computations there may have been an oversight of the fact that the probability of Simple needs to be taken into consideration when moving from $\mathcal{P}_{n,d}$ to $\mathcal{G}_{n,d}$. Nevertheless, when correcting this apparent oversight it can be shown that the proof holds for $d = O\left(n^{1/9-\delta}\right)$ for all $\delta > 0$ which is the formulation we use here.

edges, a contradiction of Property 1 of the set $\Gamma(n,d)$. Let $U=\{u_1,\ldots,u_k\}$ be the set at the end of the process, with $k=O(\sqrt{nd^3})$. Since by Property 1 of $\Gamma(n,d)$ every subset of $i=O(\sqrt{nd^3})$ vertices spans less than $5i$ edges, then for every $U'\subseteq U$ there is a vertex $v\in U'$ with $d_{G[U']}(v)\leq\frac{2e(G[U'])}{|U'|}<10$. $G[U]$ is thus 9-degenerate and by Proposition 4.1, 10-choosable.

Let $f:V\setminus U\to\{1\ldots,t\}$ be a fixed proper t-coloring of the subgraph $G[V\setminus U]$. Given this coloring of $G[V\setminus U]$, we show that there exists a choice of 10 color classes in each neighborhood $N_G(u_i)$ for every $i\in\{1,\ldots,k\}$, such that the union of these $10k$ color classes is an independent set. We can recolor the 10 color classes in each $N_G(u_i)$ by a new color $t+1$, yielding a proper $(t+1)$-coloring, $g:V\setminus U\to\{1\ldots,t+1\}$, of the vertices of $G[V\setminus U]$, and making 10 colors from $\{1,\ldots,t\}$ available for u_i. Since $G[U]$ is 10-choosable, there exists a proper coloring of $G[U]$ which colors each u_i from the set of 10 colors available for it, that extends g to all of G, thus proving G is $(t+1)$-colorable as claimed.

We define an auxiliary graph $H=(W,F)$ whose vertex set is a disjoint union of k sets, $W_1,\ldots,W_k$, where for each vertex $x\in N_G(U)$ and each neighborhood $N_G(u_i)$ in which it participates there is a vertex in $w_{x,i}\in W_i$ corresponding to x, and thus $|W_i|=|N_G(u_i)|=d$. We define $\{w_{x,i},w_{y,j}\}\in F\Leftrightarrow\{x,y\}\in E(G[N_G(U)])$, that is for every edge $\{x,y\}$ spanned in the edge set of $N_G(U)$, we define the corresponding edges between all copies of x and y in H. Since every $x\in N_G(U)$ has at most 50 neighbors in U, there are at most 50 copies of x in H, and thus every edge in $E(G[N_G(U)])$ yields at most 2500 edges in H. Furthermore, every independent set in H corresponds to an independent set in $G[N_G(U)]$, therefore, f induces a proper t-coloring, $f':W\to\{1,\ldots,t\}$, of the vertices of H.

For every $s\leq k$ subsets $W_{i_1},\ldots,W_{i_s}$, the union has $m=sd$ vertices. If $d\leq n^{4/25}$, then $m=O(n^{9/10})$ and thus, by Property 2 of $\Gamma(n,d)$, this union of sets spans at most $O(m\sqrt{d})$ edges in H. Now, for $n^{4/25}<d\ll n^{1/5}$, if $s\leq n^{7/10}$ then $sd<n^{9/10}$ and, again by Property 2 of $\Gamma(n,d)$, this union spans at most $O(m\sqrt{d})$ edges, and if $s>n^{7/10}$ then $sd>\frac{n\ln n}{d}$, and hence spans at most $O\left(\frac{m^2d}{n}\right)$ edges in H by Property 3 of $\Gamma(n,d)$. It follows, that for every $s\leq k$ subsets $W_{i_1},\ldots,W_{i_s}$ there exists a set W_{i_l} connected by at most $O(m\sqrt{d}/s)=O(d^{3/2})$ edges to the rest of the subsets if $d\leq n^{4/25}$ or $n^{4/25}<d\ll n^{1/5}$ and $s\leq n^{7/10}$, and that there exists a set W_{i_l} connected by at most $O\left(\frac{m^2d}{sn}\right)=O\left(\frac{d^{9/2}}{n^{1/2}}\right)$ to the rest of the subsets if $n^{4/25}<d\ll n^{1/5}$ and $s>n^{7/10}$. This implies that if $d\leq n^{4/25}$ the vertices $u_1,\ldots,u_k$ can be reordered in such a way that for every $1<i\leq k$ there are $O(d^{3/2})$ edges from W_i to $\cup_{i'<i}W_{i'}$, and if $n^{4/25}<d\ll n^{1/5}$ the vertices $u_1,\ldots,u_k$ can be reordered in such a way that or every $1<i\leq n^{7/10}$ there are $O(d^{3/2})$ edges from W_i to $\cup_{i'<i}W_{i'}$, and for every $n^{7/10}<i\leq k$ there are $O\left(\frac{d^{9/2}}{n^{1/2}}\right)$ edges from W_i to $\cup_{i'<i}W_{i'}$. Assume that the vertices of U are ordered in such a way.

Now, according to the given order, we choose for each u_i, for i ranging from 1 to k, a set J_i of 14 colors. We say that a color $c\in\{1,\ldots,t\}$ is *available* for u_i if there does not exist an edge $\{w_{x,i},w_{y,i'}\}$ for some $i'<i$ such that $f'(w_{x,i})=c$ and $f'(w_{y,i'})\in J_{i'}$, i.e., a color c is available for u_i, if the corresponding color class in W_i, has no connection with color classes having been chosen for previous indices. The color lists $\{J_i\}_{i=1}^k$ are sequentially chosen uniformly at random from the set of *available* colors for u_i.

Denote by p_i, for $1\leq i\leq k$, the probability that for some $i'\leq i$ while choosing the set $J_{i'}$, there are less than $\frac{t}{2}$ colors available for i'. Let us estimate p_i. Obviously, $p_1=0$. First, assume $d\leq n^{4/25}$ and $1\leq i\leq k$ or $n^{4/25}<d\ll n^{1/5}$ and $1<i\leq n^{7/10}$. In this case, there are at most $O(d^{3/2})$ edges from W_i to the previous sets $W_{i'}$ for $i'<i$. Assuming $n^{4/25}<d\ll n^{1/5}$ and $n^{7/10}<i\leq k$, it follows that there are at most $O\left(\frac{d^{9/2}}{n^{1/2}}\right)$ edges from W_i to the previous sets $W_{i'}$ for $i'<i$. By Properties 4 and 5 of $\Gamma(n,d)$, there are $\Theta(1)$ edges between $W_{i'}$ and W_i, therefore each color chosen to be included in $J_{i'}$ causes $\Theta(1)$ colors to

become unavailable for u_i. The probability of each color to be chosen into $J_{i'}$ is at most 14 divided by the number of available colors for $u_{i'}$ at the moment of choosing $J_{i'}$. Hence, if $n^{1/10} < d \leq n^{4/25}$ and $1 \leq i \leq k$ or $n^{4/25} < d \ll n^{1/5}$ and $1 < i \leq n^{7/10}$, then

$$p_i \leq p_{i-1} + (1-p_{i-1})\binom{O(d^{3/2})}{\frac{t/2}{\Theta(1)}}\left(\frac{14}{t/2}\right)^{\frac{t/2}{\Theta(1)}} \leq p_{i-1} + \left(O(1)\frac{d^{3/2}}{t^2}\right)^{\Theta(t)} \leq$$
$$p_{i-1} + \exp\left(-C_1 t \ln \frac{t^2}{d^{3/2}}\right) \leq p_{i-1} + \exp\left(-\frac{C_1 d}{2\ln n}\ln\frac{d^{1/2}}{4\ln^2 n}\right) \leq$$
$$p_{i-1} + e^{-C_2 n^{1/10}},$$

where C_1 and C_2 are positive constants. If $n^{4/25} < d \ll n^{1/5}$ and $n^{7/10} < i \leq k$, then

$$p_i \leq p_{i-1} + (1-p_{i-1})\binom{O\left(\frac{d^{9/2}}{n^{1/2}}\right)}{\frac{t/2}{\Theta(1)}}\left(\frac{14}{t/2}\right)^{\frac{t/2}{\Theta(1)}} \leq p_{i-1} + \left(O(1)\frac{d^{9/2}}{n^{1/2}t^2}\right)^{\Theta(t)} \leq$$
$$p_{i-1} + \exp\left(-C_3 t \ln \frac{n^{1/2}t^2}{d^{9/2}}\right) \leq p_{i-1} + \exp\left(-\frac{C_3 d}{2\ln n}\ln\frac{n^{1/2}}{4d^{5/2}\ln^2 n}\right) \leq$$
$$p_{i-1} + e^{-C_4 n^{4/25}} \leq p_{i-1} + e^{-C_4 n^{1/10}},$$

where C_3 and C_4 are positive constants.

Since $p_k < O(ke^{-Cn^{1/10}}) = o(1)$ for some constant C, there exists a family of color lists $\{J_i : 1 \leq i \leq k,\ |J_i| = 14\}$ for which there are no edges between the corresponding color classes of distinct subsets $W_{i'}, W_i$. Once such a family is indeed found, for every $1 \leq i \leq k$, we go over all edges inside W_i, and for every edge choose one color class that is incident with it, and delete its corresponding color from J_i. By Property 4 of $\Gamma(n,d)$, each W_i spans at most four edges, therefore, we deleted at most four colors from each J_i and thus, we get a family $\{I_i : 1 \leq i \leq k\ ,\ |I_i| \geq 10\}$, for which the union $\cup_{i=1}^k \{w \in W_i\ :\ f'(w) \in I_i\}$ is an independent set in H, completing the proof. □

5 Concluding remarks and open problems

In this paper we proved that for $d = o(n^{1/5})$ the chromatic number of a random d-regular graph on n vertices is w.h.p. concentrated on two consecutive integers. We propose here further questions which may be of interest to pursue, but will most likely need new ideas to resolve.

- Alon and Krivelevich in [3] noted, using a continuity argument, that for $\mathcal{G}(n,p)$ the two-value concentration is best possible for a general $p \leq n^{1/2-\varepsilon}$ where $\varepsilon > 0$. On the other hand, they showed there exists a series of values of p in this range for which $\chi(\mathcal{G}(n,p))$ is in fact concentrated on a single value. This may as well be the case for $\mathcal{G}_{n,d}$, but as d must be an integer, the arguments of Alon and Krivelevich cannot be applied trivially.

- Theorem 1.2 does not give any evidence as to what are the actual values on which $\chi(\mathcal{G}_{n,d})$ is concentrated. Achlioptas and Moore in [1], following ideas of Achlioptas and Naor [2], showed that for a constant d, w.h.p. $\chi(\mathcal{G}_{n,d})$ is concentrated on three consecutive integers $\{k, k+1, k+2\}$ where $k = k(d) = \min\{t \in \mathbb{N} : 2t\ln t > d\}$. In addition if $d > (2k-1)\log k$ then $\chi(\mathcal{G}_{n,d})$ is concentrated on two consecutive integers $\{k+1, k+2\}$. Recently, Kemkes, Pérez-Giménez and Wormald [18] showed that w.h.p. $\chi(\mathcal{G}_{n,d}) < k+2$ (extending [30, 31]; see also related result on $\chi(\mathcal{G}_{n,5})$ in [14]) thus determining exactly or up to two consecutive integers $\chi(\mathcal{G}_{n,d})$. Albeit the above, locating the concentration interval for non-constant values of d remains open.

- Lastly, the range of d for which Theorem 1.2 holds, seems to be far from optimal. The main obstacle to increase the range of d so as to match the corresponding result of Alon and Krivelevich (i.e. taking d to be as high as $n^{1/2-\varepsilon}$ for any $\varepsilon > 0$) is the fact the in $\mathcal{P}_{n,d}$ we could not find a "vertex-exposure" martingale which satisfied a Lipschitz condition. By using an analogue of the "edge-exposure" martingale, our concentration result was much more restrictive on d. This seems more of a technicality of our proof approach, and we believe that Theorem 1.2 holds for a larger range of d.

Acknowledgement

We would like to thank Alan Frieze for bringing Theorem 3.2 to our attention, Nick Wormald and Xavier Pérez-Giménez for providing us several of the references above and the anonymous referees for helpful comments and corrections.

References

[1] D. Achlioptas and C. Moore. The chromatic number of random regular graphs. In *RANDOM: International Workshop on Randomization and Approximation Techniques in Computer Science*, pages 219–228. LNCS, 2004.

[2] D. Achlioptas and A. Naor. The two possible values of the chromatic number of a random graph. *Annals of Mathematics. Second Series*, 162(3):1335–1351, 2005.

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

[4] N. Alon and J. H. Spencer. *The Probabilistic Method, 2nd edition*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons, 2000.

[5] S. Ben-Shimon. On the distribution of edges in random regular graphs. Master's thesis, Tel Aviv University, 2005. http://cs.tau.ac.il/research/sonny.benshimon/papers/MScThesis.pdf.

[6] E. A. Bender and E. R. Canfield. The asymptotic number of non-negative integer matrices with given row and column sums. *Journal of Combinatorial Theory, Series A*, 24(3):296–307, 1978.

[7] B. Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European Journal of Combinatorics*, 1:311–316, 1980.

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

[9] B. Bollobás. *Random Graphs*. Cambridge University Press, 2001.

[10] A. Broder, A. Frieze, S. Suen, and E. Upfal. Optimal construction of edge-disjoint paths in random graphs. *SIAM Journal on Computing*, 28(2):541–573, 1999.

[11] F. Chung. Discrete isoperimetric inequalities. In A. Grigor'yan and S. T. Yau, editors, *Surveys in Differential Geometry: Eigenvalues of Laplacians and other geometric operators*, volume IX. International Press, 2004.

[12] A. Coja-Oghlan, K. Panagiotou, and A. Steger. On the chromatic number of random graphs. *Journal of Combinatorial Theory, Series B*, 98(5):980–993, 2008.

[13] C. Cooper, A. Frieze, B. Reed, and O. Riordan. Random regular graphs of non-constant degree: Independence and chromatic number. *Combinatorics, Probability and Computing*, 11(4):323–342, 2002.

[14] J. Díaz, A. C. Kaporis, G. D. Kemkes, L. M. Kirousis, X. Pérez, and N. C. Wormald.

[15] J. Friedman, J. Kahn, and E. Szemerédi. On the second eigenvalue in random regular graphs. In *Proceedings of the Twenty First Annual ACM Symposium on Theory of Computing*, pages 587–598, 1989.

[16] A. M. Frieze and T. Łuczak. On the independence and chromatic numbers of random regular graphs. *Journal of Combinatorial Theory, Series B*, 54(1):123–132, 1992.

[17] S. Janson, T. Łuczak, and A. Ruciński. *Random Graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons, 2000.

[18] G. D. Kemkes, X. Pérez-Gimenéz, and N. C. Wormald. On the chromatic number of d-regulat graphs. http://arxiv.org/abs/0812.2973.

[19] J. H. Kim, B. Sudakov, and V. H. Vu. On the asymmetry of random regular graphs and random graphs. *Random Structures & Algorithms*, 21(3-4):216–224, 2002.

[20] J. H. Kim, B. Sudakov, and V. H. Vu. Small subgraphs of random regular graphs. *Discrete Mathematics*, 307(15):1961–1967, 2007.

[21] J. H. Kim and V. H. Vu. Sandwiching random graphs: Universality between random graph models. *Advances in Mathematics*, 188(2):444–469, 2004.

[22] M. Krivelevich and B. Sudakov. Pseudo-random graphs. In E. Győri, G. O. H. Katona, and L. Lovász, editors, *More sets, graphs and numbers: A Salute to Vera Sòs and András Hajnal*, volume 15 of *Bolyai Society Mathematical Studies*, pages 199–262. Springer, 2006.

[23] M. Krivelevich, B. Sudakov, V. H. Vu, and N. C. Wormald. Random regular graphs of high degree. *Random Structures & Algorithms*, 18(4):346–363, 2001.

[24] T. Łuczak. The chromatic number of random graphs. *Combinatorica*, 11(1):45–54, 1991.

[25] T. Łuczak. A note on the sharp concentration of the chromatic number of random graphs. *Combinatorica*, 11(3):295–297, 1991.

[26] C. McDiarmid. Concentration. In M. Habib, C. McDiarmid, J. Ramirez-Alfonsin, and B. Reed, editors, *Probabilistic Methods for Algorithmic Discrete Mathematics*, pages 195–248. Springer, 1998.

[27] B. D. McKay. Subgraphs of random graphs with specified degrees. *Congressus Numerantium*, 33:213–223, 1981.

[28] B. D. McKay and N. C. Wormald. Asymptotic enumeration by degree sequence of graphs with degrees $o(n^{1/2})$. *Combinatorica*, 11(4):369–382, 1991.

[29] E. Shamir and J. Spencer. Sharp concentration of the chromatic number on random graphs $\mathcal{G}(n,p)$. *Combinatorica*, 7(1):131–138, 1987.

[30] L. Shi and N. Wormald. Colouring random 4-regular graphs. *Combinatorics, Probability and Computing*, 16(2):309–344, 2007.

[31] L. Shi and N. C. Wormald. Colouring random regular graphs. *Combinatorics, Probability and Computing*, 16(3).

[32] N. C. Wormald. Models of random regular graphs. In J. Lamb and D. Preece, editors, *Surveys in Combinatorics*, volume 276 of *London Mathematical Society Lecture Note Series*, pages 239–298. Cambridge University Press, 1999.