A note on the chromatic number of the square of a sparse random graph

Alan Frieze* and Aditya Raut
Department of Mathematical Sciences
Carnegie Mellon University
Pittsburgh PA15217

Abstract

We show that w.h.p the chromatic number $\chi$ of the square of $G_{n,p}$, $p = c/n$ is asymptotically equal to the maximum degree $\Delta(G_{n,p})$. This improves an earlier result of Garapaty et al [5] who proved that $\chi(G_{n,p}^2) \leq 6 \cdot \Delta(G_{n,p})$ w.h.p.

1 Introduction

Let $p = c/n$ where $c > 0$ is a constant. The chromatic number of $G_{n,p}$ is well-understood, at least for sufficiently large $c$. Luczak [6] proved that if $G = G_{n,p}$ then $\chi(G) \sim \frac{c}{2\log c}$. This was refined by Achlioptas and Naor [1] and further refined later by Coja-Oghlan and Vilenchik [2].

The square of a graph $G$ is obtained from $G$ by adding edges for all pairs of vertices at distance two or less from each other. Atkinson and Frieze [3] showed that w.h.p. the independence number of $G_2 = G_{n,p}^2$ is asymptotically equal to $\frac{4n\log c}{c^2}$, for large $c$. Garapaty, Lokshtanov, Maji and Pothen [5] studied the chromatic number of powers of $G_{n,p}$. Let $\Delta = \Delta(G_{n,p}) \sim \frac{\log n}{\log \log n}$ be the maximum degree in $G = G_{n,p}$ (for a proof of this known claim about the maximum degree, see for example [4], Theorem 3.4). Garapaty et al. proved, in the case of the square $G_2$ of $G_{n,p}, p = c/n$ that $\chi(G_2) \leq 6 \cdot \frac{\log n}{\log \log n}$ w.h.p. We strengthen this and prove

Theorem 1. Let $p = c/n$, $c > 0$ constant. Let $G_2$ denote the square of $G_{n,p}$. Then, w.h.p. $\chi(G_2) \sim \Delta(G_{n,p}) \sim \frac{\log n}{\log \log n}$.

We will show that w.h.p. we can properly color $G_2$ with $q = \Delta(1 + 3\theta^{1/3})$ colors, where $\theta = o(1)$ is given in [1]. Note that the neighbors of a vertex form a clique in $G_2$ and so the lower bound in the theorem is trivial.

Remark 1. The value of $c$ does not contribute to the main term in the claim of Theorem 1. Thus we would expect that we could replace $p = c/n$ by $p \leq \omega/n$ for some slowly growing function $\omega = \omega(n) \to \infty$. Indeed, a careful examination of the proof below verifies this so long as $c = o(\log \log n)$.

*Research supported in part by NSF grant DMS1952285
2 Proof of Theorem 1

2.1 Structural properties

We can use the following high probability bounds for $\Delta$ taken from [4], Theorem 3.4:

\[
\frac{\log n}{\log \log n} \left(1 - \frac{3 \log \log \log n}{\log n}\right) \leq \Delta \leq \frac{\log n}{\log \log n} \left(1 + \frac{3 \log \log \log n}{\log n}\right)
\]

This implies that w.h.p.

\[
n^{1-\theta} \leq \Delta^\Delta \leq n^{1+\theta} \text{ where } \theta = \frac{4 \log \log n}{\log n}.
\] (1)

Let $d(v)$ denote the degree of $v$ in $G_{n,p}$. For $0 < \alpha \leq 1$, let $V_\alpha = \{v : d(v) \geq \alpha \Delta\}$ and let $W_\alpha$ denote the closed neighborhood of $V_\alpha$ i.e $V_\alpha$ plus the neighbors of $V_\alpha$.

Fix

\[\varepsilon = \theta^{1/2}.\]

The next few lemmas are needed to analyse the coloring of vertices in $W_\varepsilon$.

Lemma 2. W.h.p., $v, w \in V_{2/3}$ implies that $\text{dist}(v, w) \geq 10$. (Here $\text{dist}(., .)$ is graph distance in $G_{n,p}$.)

Define

\[
L_m = \left\{(\ell_1, \ell_2, \ldots, \ell_m) \in \{\varepsilon \Delta, \varepsilon \Delta + 1, \ldots, \Delta\}^m : \sum_{i=1}^m \ell_i \geq (1 + \theta^{1/3})\Delta\right\}
\]

Lemma 3. Suppose that $m \leq 2/\varepsilon$. Then w.h.p. there does not exist a connected subset $S \subseteq [n]$ of $G_{n,p}$ with at most $3m$ vertices containing vertices $w_i \in i = 1, 2, \ldots, m$ such that $(d(w_i), i = 1, 2, \ldots, m) \in L_m$.

Corollary 4. A vertex $v \notin V_\varepsilon$ has at most $\Delta_1 = (1 + 2\theta^{1/3})\Delta$ $G_2$-neighbors in $W_\varepsilon$, w.h.p.

Proof. Suppose $v$ has more than $\Delta_1$ $G_2$-neighbors in $W_\varepsilon$. Let $T$ be the tree obtained by Breadth-First-Search to depth three from $v$ in $G_1 = G_{n,p}$. Let this tree have levels $L_0 = \{v\}, L_1, L_2, L_3$. Let the $G_1$ neighbors of $v$ be $\{u_1, u_2, \ldots, u_k\}$. Let $F_{i,t}, t = 2, 3$ denote the vertices in $L_t$ separated from $v$ in $T$ by $u_i$.

We now define a subtree $T_1$ of $T$ that will take the place of $S$ in Lemma 3. To obtain $T_1$ we do the following: suppose that $u_1, u_2, \ldots, u_p$ are the neighbors of $v$ in $V_\varepsilon$. Delete the neighbors of $u_i, i \in [1, p]$, except for $v$. Suppose that $X_i = F_{i,2} \cap V_\varepsilon \neq \emptyset$ for $i \in [p+1, q]$ and that $F_{i,2} \cap V_\varepsilon = \emptyset$ for $i \in [q+1, k]$. Choose a vertex $x_i \in X_i$ for each $i \in [p+1, q]$ and delete $X_i \setminus \{x_i\}$ from $T$. Suppose also that $Y_i = F_{i,3} \cap V_\varepsilon \neq \emptyset$ for $i \in [q+1, r]$. Choose a vertex $x_i \in X_i$ with a neighbor $y_i$ in $Y_i$ for each $i \in [q+1, r]$ and delete $X_i \setminus \{x_i\}, Y_i \setminus \{y_i\}$ from $T$. $T_1$ is the tree that survives these deletions. For $i \in [r+1, k]$, we delete $u_i$ and the vertices $F_{i,2} \cup F_{i,3}$ from $T$.

Vertex $v$ has at most

\[
D = \sum_{i=1}^r d(u_i) + (k - r) \leq \sum_{i=1}^p d(u_i) + (r - p)\varepsilon\Delta + (k - r) \leq \sum_{i=1}^p d(u_i) + (r - p + 1)\varepsilon\Delta.
\]

$G_2$ neighbors in $W_\varepsilon$. Our assumption is that $D > \Delta_1$. 

1
Next let $M = \sum_{w \in V(T_1) \cap V_\varepsilon} d(w)$. Then,

$$M \geq \sum_{i=1}^{p} d(u_i) + (r - p)\varepsilon \Delta \geq D - \varepsilon \Delta > (1 + 2\theta^{1/3} - \theta^{1/2})\Delta > (1 + \theta^{1/3})\Delta.$$ 

We get a contribution of at least $\varepsilon \Delta$ from a member of $F_{i,2}$ for $i \in [p + 1, q]$ and a contribution of at least $\varepsilon \Delta$ from the surviving member of $F_{i,3}$ for $i \in [q + 1, r]$.

The number of vertices $N$ in the tree $T_1$ satisfies

$$N \leq 1 + p + 2(q - p) + 3(r - q) \leq 3|V(T_1) \cap V_\varepsilon|.$$ 

Putting $m = |V(T_1) \cap V_\varepsilon|$, we see that this contradicts Lemma $3$ provided $m \leq 2/\varepsilon$. Assume then that $m > 2/\varepsilon$. It follows from Lemma $2$ that either $T$ consists of $u_1 \in V_{2/3}$ and the neighbors of $u_1$ and then the corollary holds trivially. Otherwise, $M > m\varepsilon \Delta > 2\Delta$ and one can delete a vertex of degree less than $2\Delta/3$ and reduce $m$ by one keep $M > (1 + \theta^{1/3})\Delta$, eventually leading to a contradiction. □

A similar argument gives

**Corollary 5.** A vertex has at most $\Delta_1 G_2$-neighbors in $V_\varepsilon$.

**Proof.** Suppose $v$ has more than $\Delta_1 G_2$-neighbors in $V_\varepsilon$. Let $T$ be the tree obtained by Breadth-First-Search to depth two from $v$ in $G_1 = G_{n,p}$. Remove all leaves from $T$ that are not in $V_\varepsilon$ and repeat. We are left with a set of $G_1$-neighbors $W_0$ of $v$ in $V_\varepsilon$ and set of $G_1$-neighbors $u_1, u_2, \ldots, u_k$ of $v$ that are not in $V_\varepsilon$. In addition we have sets $W_1, W_2, \ldots, W_k \subseteq V_\varepsilon$ such that $u_i$ is a $G_1$-neighbor of all vertices in $W_i, i = 1, 2, \ldots, k$. The $G_2$-degree of $v$ is given by $D = \sum_{w \in W_0} d(w) + k + \sum_{i=1}^{k} |W_i| > \Delta_1$. The number of $G_2$-neighbors in $V_\varepsilon$ is

$m_1 = \sum_{i=0}^{k} |W_i|$ and the tree $T$ contains $1 + k + m_1 \leq 2m_1 + 1$ vertices. Let $W = \bigcup_{i=0}^{k} W_i$ and add $v$ to $W$ if $v \in V_\varepsilon$. Then let $W = \{w_1, w_2, \ldots, w_m\}$ where $m = m_1 + 1_{v \in V_\varepsilon}$. If $m \leq 2/\varepsilon$ then we contradict Lemma $3$.

Otherwise, $M = \sum_{i=1}^{m} d(w_i) \geq m \varepsilon \Delta$ where $d_{\min} = \min \{d(w) : w \in W\}$. But $d_{\min} \geq \varepsilon \Delta$ and so $M > 2\Delta$. It follows from Lemma $2$ that $d_{\min} < 2\Delta/3$ and so we can reduce $m$ by one and keep $M > 4\Delta/3$. Continuing in this way, we eventually reduce $m$ to below $2/\varepsilon$ and keep $M > (1 + \theta^{1/3})\Delta$. But now we contradict Lemma $3$. □

The next part of our strategy is to bound the number of $G_2$-edges contained in any set $S$ that is disjoint from $W_\varepsilon$. We prove a high probability bound of $(5.5 + 2c)\varepsilon \Delta |S|$.

**Remark 2.** This will imply that the vertices of $[n] \setminus W_\varepsilon$ can be list-colored using at most $(5.5 + 2c)\varepsilon \Delta + 1$ colors.

This remark follows from Lemmas $3$, $7$, and $8$.

For large sets we can use the following:

**Lemma 6.** The total number of edges in $G_2$ is less than $c(c + 1)n$ w.h.p.

For $2 \leq s \leq n$ let $\nu_s$ be the maximum number of $G_2$-edges in a set of size $s$. Then we have that

$$\text{if } k_0 = 2/\varepsilon^2 \text{ then } \nu_s \leq 10k_0\varepsilon^3 s \text{ for } n/(10ck_0) \leq s \leq n. \quad (2)$$

If $S \cap W_\varepsilon = \emptyset$ then a vertex outside $S$ has at most $\varepsilon \Delta$ neighbors in $S$. For a fixed set $S$ let $a_{k,S}$ denote the number of (vertex, set) pairs $v, T$ where $v \notin S$ and $|T| = k$ and $T = N(v) \cap S$. 

3
Lemma 7. The following holds w.h.p. Let \( A_1(S) = \sum_{k \leq \varepsilon a_{k,S}} k^2 a_{k,S} \) bound the number of \( G_2 \)-edges due to the \( a_{k,S} \). Then \( A_1(S) \leq 5\varepsilon \Delta |S| \) for all \(|S| \leq n/(10ck_0)\).

We now have to deal with the number of edges \( uv \) and the number of paths of length two \( uvw \) where \( \{u, v, w\} \subseteq S \). Denote this by \( A_2(S) \).

Lemma 8. W.h.p., \( A_2(S) \leq (2c + 1/2)\varepsilon \Delta |S| \) for all \(|S| \leq n/(10ck_0)\).

Thus,
\[
A_1(S) + A_2(S) \leq (5.5 + 2c)\varepsilon \Delta |S|.
\]
(3)
This, together with (2), verifies what we claimed in Remark 2.

3 Coloring \( G_2 \)

Given the above we color \( G_2 \) as follows:

(a) We color \( V_\varepsilon \) with \( q = \Delta(1 + 3\theta^{1/3}) \) colors. We do this greedily i.e we arbitrarily order the vertices in \( V_\varepsilon \) and then in this order, we color a vertex with the lowest index available color. Corollary 5 implies that any vertex \( v \) has at most \( \Delta \left(1 + 2\theta^{1/3}\right) \) \( G_2 \)-neighbors in \( V_\varepsilon \) and so there will be an unused color.

(b) We color \( \hat{W}_\varepsilon = W_\varepsilon \setminus V_\varepsilon \) with \( q = \Delta(1 + 3\theta^{1/3}) \) colors. We do this greedily i.e we arbitrarily order the vertices in \( \hat{W}_\varepsilon \) and then in this order, we color a vertex with the lowest index available color. Corollary 4 implies that any vertex \( v \notin V_\varepsilon \) has at most \( \Delta \left(1 + 2\theta^{1/3}\right) \) \( G_2 \)-neighbors in \( \hat{W}_\varepsilon \) and so there will be an unused color.

(c) We then color \( [n] \setminus W_\varepsilon \) with at most \( \Delta \left(1 + 2\theta^{1/3} + (5.5 + 2c)\theta^{2/3}\right) \) colors. This follows from (3) and Corollary 4.

3.1 Proof of Lemma 2

Let \( \ell_0 = 2\Delta/3 - 10 \). We have
\[
\mathbb{P}(\exists v, w \in V_{2/3} : dist(v, w) < 10) \leq \sum_{k=1}^{9} \binom{n}{k} k! p^{k-1} \left( \sum_{\ell=\ell_0}^{n-1} \binom{n}{\ell} p^\ell (1-p)^{n-10-\ell} \right)^2
\]
\[
\leq \sum_{k=1}^{9} n e^{k-1} n^{-4/3+o(1)} = o(1).
\]

3.2 Proof of Lemma 3

Then,
\[
\mathbb{P}(\exists S) \leq \sum_{m=2}^{2/\varepsilon} \sum_{s=m}^{3m} \binom{n}{s} s^{-2} p^{s-1} \binom{s}{m} \sum_{D \geq (1+\theta^{1/3})\Delta} \prod_{i=1}^{m} \left( \sum_{k=\ell_i}^{n-s} \binom{n-s}{k} p^k (1-p)^{n-s-k} \right).
\]
3.3 Proof of Lemma 6

Let \( d(i) \) denote the degree of vertex \( i \) in \( G_{n,p} \). The expected number of edges in \( G_2 \) is

\[
\mathbb{E}\left(\sum_{i=1}^{n} \frac{d(i)(d(i) + 1)}{2}\right) = \frac{n}{2} \sum_{j=1}^{n-1} j(j+1) \binom{n-1}{j} p^j (1-p)^{n-1-j} = \frac{c^2(n-1)(n-2) + cn^2}{2n}.
\]

To show concentration round the mean, we use the following theorem from Warnke [7]:

**Theorem 9.** Let \( X = (X_1, X_2, \ldots, X_N) \) be a family of independent random variables with \( X_k \) taking values in a set \( \Lambda_k \). Let \( \Omega = \prod_{k \in [N]} \Lambda_k \) and suppose that \( \Gamma \subseteq \Omega \) and \( f : \Omega \to \mathbb{R} \) are given. Suppose also that whenever \( \mathbf{x}, \mathbf{x}' \in \Omega \) differ only in the \( k \)th coordinate

\[
|f(\mathbf{x}) - f(\mathbf{x}')| \leq \begin{cases} 
c_k & \text{if } \mathbf{x} \in \Gamma, 
d_k & \text{otherwise.}
\end{cases}
\]

If \( W = f(X) \), then for all reals \( \gamma_k > 0 \),

\[
\mathbb{P}(W \geq \mathbb{E}(W) + t) \leq \exp\left\{-2\sum_{k \in [N]} \left(\frac{t^2}{(c_k + \gamma_k (d_k - c_k)^2)}\right)\right\} + \mathbb{P}(X \notin \Gamma) \sum_{k \in [N]} \gamma_k^{-1}.
\]

We use Theorem 9 with \( N = n \), \( W = \sum_{i=1}^{n} \frac{d(i)(d(i) + 1)}{2} \), \( X_i = \{j < i : \{j, i\} \text{ is an edge of } G_{n,p}\} \), \( i = 1, 2, \ldots, n \) and \( \Gamma = \{\Delta(G_{n,p}) \leq \log n\} \). In which case we can take \( c_k = \log^2 n \), \( d_k = n^2 \) and \( \mathbb{P}(X \notin \Gamma) \leq (\log n)^{-\frac{1}{2}} \log n \). Then we can take \( \gamma_k = n^{-4} \) for \( k \in [n] \) and \( t = n^{3/2} \) to complete the proof of Lemma 6.

3.4 Proof of Lemma 7

Let \( |S| \leq n/(10ck_0) \). We will prove:
(a) For $k_0 < k_1 < k_2 \leq \varepsilon \Delta$, for all sets $S \subseteq [n] \setminus V_{\varepsilon}$,

$$\sum_{k=k_1}^{k_2} a_{k,S} \leq \frac{(1 + \varepsilon)|S|}{k_1}.$$

(b) $a_{k,S} \leq (10c)^{k_0}|S|$ for $k \leq k_0$.

We have, where $M_{u,k_1,k_2} = \{(x_2, \ldots, x_{\varepsilon \Delta}) : \sum_{k=k_1}^{k_2} x_k = u\}$.

$$\Pr\left( \exists S, |S| = s \leq n/(10ck_0) : \sum_{k=k_1}^{k_2} a_{k,S} \geq t \right) \leq \sum_{s=k_1^{1/2}}^{n/(10ck_0)} \left( \frac{n}{s} \right)^s \sum_{u \geq t} \sum_{x \in M_{u,k_1,k_2}} \left( \frac{n}{x_{k_1}, \ldots, x_{k_2}}, n - u \right) \prod_{k=k_1}^{k_2} \left( \left( \frac{s}{k} \right) \left( \frac{c}{n} \right)^k \right)^x_k$$

$$\leq \sum_{s=k_1^{1/2}}^{n/(10ck_0)} \left( \frac{ne}{s} \right)^s \sum_{u \geq t} \sum_{x \in M_{u,k_1,k_2}} \left( \frac{n}{x_{k_1}, \ldots, x_{k_2}}, n - u \right) \prod_{k=k_1}^{k_2} \left( \frac{\sec k_1}{kn} \right)^{x_k}$$

$$= \sum_{s=k_1^{1/2}}^{n/(10ck_0)} \left( \frac{ne}{s} \right)^s \sum_{u \geq t} \left( \frac{\sec k_1}{kn} \right)^{x_k} \prod_{k=k_1}^{k_2} \left( \frac{n}{u} \right)$$

Putting $t = (1 + \varepsilon)s/k_1$, we have, for large $k_1$ i.e. for $k_1 > 2/\varepsilon$,

$$\Pr\left( \exists S, |S| = s \leq n/(10ck_0) : \sum_{k=k_1}^{k_2} a_{k,S} \geq \frac{(1 + \varepsilon)s}{k_1} \right) \leq \sum_{s=k_1^{1/2}}^{n/(10ck_0)} \left( \frac{ne}{s} \right)^s \sum_{u \geq t} \left( \frac{\sec k_1}{kn} \right)^{x_k} \prod_{k=k_1}^{k_2} \left( \frac{n}{u} \right)$$

$$\leq 2 \sum_{s=k_1^{1/2}}^{n/(10ck_0)} \left( \frac{ne}{s} \right)^s \left( \frac{\sec k_1}{kn} \right)^{x_k} \prod_{k=k_1}^{k_2} \left( \frac{n}{(1 + \varepsilon)s/k_1} \right)$$

$$\leq 2 \sum_{s=k_1^{1/2}}^{n/(10ck_0)} \left( \frac{ne}{s} \right)^s \left( \frac{\sec k_1}{kn} \right)^{1+\varepsilon} \left( \frac{n}{(1 + \varepsilon)s/k_1} \right)^{(1+\varepsilon)s/k_1}$$

$$= 2 \sum_{s=k_1^{1/2}}^{n/(10ck_0)} \left( \frac{ne}{s} \right)^s \left( \frac{\sec k_1}{kn} \right)^{1+\varepsilon} \left( \frac{n}{(1 + \varepsilon)s/k_1} \right)^{(1+\varepsilon)s/k_1}$$

$$= 2 \sum_{s=k_1^{1/2}}^{n/(10ck_0)} \left( \frac{s}{n} \right)^{(1+\varepsilon)/k_1} \frac{2+\varepsilon+(1+\varepsilon)/k_1 c^{1+\varepsilon}}{k_1^{(1+\varepsilon)(k_1-1)/k_1}}$$

$$= o(n^{-2}).$$

We also have

$$\Pr(\exists S, |S| = s \leq n/(10ck_0) : a_{k,S} \sim t) \leq \left( \frac{n}{s} \right) \left( \frac{n}{t} \right) \left( \frac{\left( \frac{c}{n} \right)^k}{k} \right)^t$$
If $k \sigma$ are more technically challenging.

While we have shown that $\chi(G) = \Delta + 1$ w.h.p. This would be quite pleasing, but we are not confident enough to make this a conjecture. It is of course interesting to further consider $\chi(G)$ when $np \to \infty$. Note that when $np \gg n^{1/2}$, the diameter of $G_{n,p}$ is equal to 2 w.h.p. In which case $G_2 = K_n$. One can also consider higher powers of $G_{n,p}$ as was done in [3] and [5]. Such considerations are more technically challenging.

4 Conclusions

While we have shown that $\chi(G_2) \sim \Delta$ w.h.p., it is possible that $\chi(G_2) = \Delta + 1$ w.h.p. This would be quite pleasing, but we are not confident enough to make this a conjecture. It is of course interesting to further consider $\chi(G_2)$ when $np \to \infty$. Note that when $np \gg n^{1/2}$, the diameter of $G_{n,p}$ is equal to 2 w.h.p. In which case $G_2 = K_n$. One can also consider higher powers of $G_{n,p}$ as was done in [3] and [5]. Such considerations are more technically challenging.
References


