A note on randomly colored matchings in random graphs

Alan Frieze

Abstract
We are given a bipartite graph that contains at least one perfect matching and where each edge is colored from a set \( Q = \{c_1, c_2, \ldots, c_q\} \). Let \( Q_i = \{e \in E(G) : c(e) = c_i\} \), where \( c(e) \) denotes the color of \( e \). The perfect matching color profile \( mcp(G) \) is defined to be the set of vectors \((m_1, m_2, \ldots, m_q) \in [n]^q\) such that there exists a perfect matching \( M \) such that \(|M \cap Q_i| = m_i\). We give bounds on the matching color profile for a randomly colored random bipartite graph.

1 Introduction

We consider the following problem: we are given a random bipartite graph \( G \) in which each edge is given a random color from a set \( Q = \{c_1, c_2, \ldots, c_q\} \). An edge \( e \) is colored \( c(e) = c_i \) with probability \( \alpha_i \) where \( \alpha_i > 0 \) is a constant. Let \( Q_i = \{e \in E(G) : c(e) = c_i\} \), where \( c(e) \) denotes the color of \( e \). The perfect matching color profile \( mcp(G) \) is defined to be the set of vectors \((m_1, m_2, \ldots, m_q) \in [n]^q\) such that there exists a perfect matching \( M \) such that \(|M \cap Q_i| = m_i\). We give bounds on the matching color profile for a randomly colored random bipartite graph.

Randomly colored random graphs have been studied recently in the context of (i) rainbow matchings and Hamilton cycles, see for example \([2], [3], [7], [11]\); (ii) rainbow connection see for example \([5], [9], [10], [13], [12]\); (iii) pattern colored Hamilton cycles, see for example \([1], [6]\). This paper can be considered to be a contribution in the same genre. One can imagine a possible interest in the color profile via the following scenario: suppose that \( A \) is a set of tools and \( B \) is a set of jobs where edge \( \{a, b\} \) indicates that \( b \) can be completed using \( a \). If colors represent

Alan Frieze
Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh PA15213, USA.
e-mail: alan@random.math.cmu.edu. Research supported in part by NSF grant DMS1661063
people, then one might be interested in equitably distributing jobs. I.e. determining whether \((n/q, n/q, \ldots, n/q) \in mcp(G)\). In any case, we find the problem interesting.

We will consider \(G\) to be the random bipartite graph \(G_{n,n,p}\) where \(p = \frac{\log n + \omega}{n}, \omega = \omega(n) \to \infty\) where \(\omega = o(\log n)\). Erdős and Rényi [4] proved that \(G\) has a perfect matching w.h.p. We will prove the following theorem: let \(\alpha_1, \alpha_2, \ldots, \alpha_q\) be positive constants such that \(\alpha_1 + \alpha_2 + \cdots + \alpha_q = 1\). Let

\[
\alpha_{\text{min}} = \min\{\alpha_i : i \in [q]\} \quad \text{and} \quad n_1 = 2n/(\alpha_{\text{min}} \log n)^{1/2}.
\]

**Theorem 1.** Let \(G\) be the random bipartite graph \(G_{n,n,p}\) where \(p = \frac{\log n + \omega}{n}, \omega = \omega(n) \to \infty\) where \(\omega = o(\log n)\). Suppose that the edges of \(G\) are independently colored with colors from \(C = \{c_1, c_2, \ldots, c_q\}\) where \(\Pr(e(c) = c_i) = \alpha_i\) for \(e \in E(G)\), \(i \in [q]\). Let \(m_1, m_2, \ldots, m_q\) satisfy: (i) \(m_1 + \cdots + m_q = n\) and (ii) \(n_1 \leq m_i \leq n - n_1, i \in [q]\). Then w.h.p., there exists a perfect matching \(M\) in which exactly \(m_i\) edges are colored with \(c_i\), \(i = 1, 2, \ldots, q\).

It is clear that w.h.p. \((n, 0, \ldots, 0) \notin mcp(G)\). This is because the bipartite graph induced by edges of color \(c_1\) is distributed as \(G_{n,n,\alpha_1,p}\) and this contains isolated vertices w.h.p. On the other hand, if \(p \geq \frac{2(\log n + \omega)}{\alpha_{\text{min}} n}\) then w.h.p. \(mcp(G) = [n]^q\). To see this, suppose that \(m_1 \leq m_2 \leq \cdots m_q \leq n\). Suppose we have found a matching that uses \(m_i\) edges of color \(c_i\) for \(i \geq 0\). Let \(n' = n - m_1 - \cdots - m_i\). Then the random bipartite graph induced by vertices not in \(M\) and has edges of color \(c_i\) has density at least \(\frac{\beta \log n'}{\alpha_{\text{min}} n} \geq \frac{\log n' + \omega}{n^{1/2}}\) and so has a perfect matching w.h.p.

**Open Question:** What is the threshold for \(mcp(G) = [0, n]^q\)?

### 2 Structural Lemma

Suppose that the bipartition of \(V(G)\) is denoted \(A, B\). Let \(\text{SMALL}_i = \{x \in A : x\) is incident to at most \((\alpha_i \log n)/20\) edges of color \(c_i\}\) and \(\text{SMALL} = \bigcup_{i=1}^q \text{SMALL}_i\) and \(\text{LARGE} = A \setminus \text{SMALL}\). Let \(\Psi(\alpha) = 1 - \alpha + \alpha(\log 20 + 1 - \log \alpha)/20\).

**Lemma 1.** Let \(p = \frac{\log n + \omega}{n}, \omega = \omega(n) \to \infty\) where \(\omega = o(\log n)\). Then w.h.p.

(a) \(|\text{SMALL}_i| \leq n^{1-\Psi(\alpha_i)/2}\) for \(i \in [q]\).

(b) \(S \subseteq A, T \subseteq B\) and \(|S| \leq n_0 = \gamma n/\log n\) where \(\gamma = \alpha_{\text{min}}^2/100\), \(|T| \leq \beta |S| \log n\) where \(\beta \leq \alpha_{\text{min}}/4\) implies that \(e_G(S : T) \leq 2\beta |S| \log n\).

(c) There do not exist sets \(S \subseteq A, T \subseteq B\) and \(i \in [q]\) such that \(|S|, |T| \geq n_1 = 2n/(\alpha_{\text{min}} \log n)^{1/2}\) and \(e_i(S, T) = 0\).

(d) There does not exist a vertex \(a \in A\) together with \(L = 2/(1 - \Psi(\alpha_{\text{min}})/2)\) paths \((a,b_1,a_i)\) such that \(a_i \in \text{SMALL}\) for \(i = 1, 2, \ldots, L\). Here we can assume that \(a_1, \ldots, a_L, b_1, \ldots, b_L\) are distinct.

**Proof** (a) Let \(\ell_i = (\alpha_i \log n)/20\). We have
The probability that the condition is violated can be bounded by

\[ \mathbb{E}(\text{SMALL}_i) \leq n \sum_{j=0}^{\ell_i} \binom{n}{j} (\alpha p)^j (1 - \alpha p)^{n-1-j} \leq n \sum_{j=1}^{\ell_i} \left( \frac{ne\alpha p}{j} \right)^j n^{-\alpha + o(1)} \leq n^{1-\alpha + o(1)} \left( \frac{20e^{1+o(1)}}{\alpha} \right)^{\alpha \log n} \leq n^{1-2\psi(\alpha)/3}. \]

The result follows from the Markov inequality and the union bound over \(i \in [q] \).

(b) The probability that the condition is violated can be bounded as follows:

\[
\sum_{s=2\beta \log n}^{m_0} \sum_{t=1}^{n \choose s} \left( \frac{s}{2 \beta \log n} \right) \beta^{s \log n} \left( \frac{s}{\beta \log n} \right)^{1/2} \left( \frac{e^{1+o(1)}}{2 \beta n} \right)^2 \beta^{s \log n} \leq o(1).
\]

(c) Suppose that \(|S|, |T| \geq 2n/\left(\alpha_{\max} \log n\right)^{1/2}\) then

\[ \mathbf{Pr}(e(S, T) = 0) \leq (1 - \alpha p)^{|S|+|T|} \leq e^{-2n}. \]

The result now follows from the union bound over the at most \(4^n\) choices for \(S, T\).

(d) The expected number of vertices violating the condition can be bounded by

\[
n \binom{n}{L}^2 (\log n)^L \sum_{i=1}^{q} \binom{\alpha \log n}{20} \binom{n}{j} (\alpha p)^j (1 - \alpha p)^{n-1-j} L = O(n(\log n)^2 e^{-L(1 - \psi(\alpha_{\max}/2))}) = o(1).
\]

3 Proof of Theorem 1

**Proof** Assume from now on the high probability conditions of Lemma 1 are in force. Let \(X\) be a set of edges and let \(\mu_i = |X \cap Q_i|\). For a set of edges \(X\) we let
We prove the theorem, by showing that if \( M \) is a perfect matching of \( G \) and \( \Delta(M) > 0 \) then there is an alternating cycle \( C \) such that the matching \( M' = M \oplus C \) satisfies \( \Delta(M') \leq \Delta(M) - 2 \).

Suppose then that \( \Delta(M) > 0 \). For \( a \in A \) let \( \{a, M(a)\} \) be the edge of \( M \) that contains \( a \). We prove the existence of \( e_1, f_1, \ldots, e_r, f_r \) such that if \( E_i = (M \cup \{e_1, e_2, \ldots, e_i\}) \setminus \{f_1, f_2, \ldots, f_i\} \) then (i) \( \Delta(E_i) < \Delta(M) \), (ii) \( e_i \cap f_i = \{x_j\} \subseteq A, e_i \cap f_{i+1} = \{y_j\} \subseteq B \), and because we add edges that are in \( f \), after this we let \( a \cap M = 1 \) and \( a \cap M = 2 \). For each \( a \in S_j \) there is an alternating path \( P_a = (x_0, a_0, y_0, x_1, y_1, \ldots, x_s = a \in S_j) \) such that \( x_k = M(x_k) \in T_k \) and \( x_k = \phi(x_{k-1}) \in S_k \) for \( k \leq j \). Then let \( X_a = M \oplus E(P_a) \).

Having chosen \( S_j \) for some \( j \geq 0 \), we define

\[
T_j = \left\{ b \in B \setminus \bigcup_{k<j} T_k : \exists a \in S_j \text{ s.t. } \{a, b\} \in Q_-(X_a) \text{ and } \{b, M^{-1}(b)\} \notin Q_1 \right\}.
\]

After this we let \( S_{j+1} = M^{-1}(T_j) \cap LARGES \). We note that \( |X_a| = |M| \) for \( a \in S_j \), \( i \geq 0 \) and because \( X_a \cap Q_1 = M \cap Q_1 \) we see that \( \Delta(X_a) > 0 \) and so \( Q_-(X_a) \neq \emptyset \). We note also, that because we add edges that are in \( Q_i, i \in Q_- \) that \( \Delta(X_a) \leq \Delta(M) \).

We claim that for \( i \geq 0 \),

\[
|S_i| \leq \frac{ym}{\log n} \text{ implies that } |S_{i+1}| \geq \frac{\alpha_{\min} \log n}{10} |S_i|.
\]

We verify (1) below. Assuming its truth, there exists a smallest \( k \) such that \( |S_{k-1}| \geq y\alpha_{\min}n/10 \). It then follows from Lemma 1(c) that

\[
|S_{k-1} \cup S_k| \geq n - n_1.
\]

Starting with \( S_0 = \{b_0\} \), we can similarly construct a sequence of sets \( T_1, S_1 \), \( \ldots, \hat{T}_k, \hat{S}_k \) and corresponding paths \( P_b, b \in \hat{T}_j \) and sets \( X_b \). Having defined \( \hat{T}_j \) we let

\[
\hat{S}_j = \left\{ a \in A \setminus \bigcup_{k<j} \hat{S}_k : \exists b \in \hat{T}_j \text{ s.t. } \{a, b\} \in Q_-(X_b) \text{ and } \{a, M(a)\} \notin Q_1 \right\}.
\]
and then let \( \tilde{T}_{i+1} = M(\hat{S}_j) \). The equivalent of (1) will be
\[
|\tilde{T}_i| \leq \frac{\gamma n}{\log n} \text{ implies that } |\tilde{T}_{i+1}| \geq \frac{\alpha_{\min} \log n}{10} |\tilde{T}_i|.
\] (3)

Assuming its truth, there exists \( \ell \) such that \( |\tilde{T}_{\ell-1}| \geq \gamma \alpha_{\min} n / 10 \). It then follows from Lemma 1(c) that
\[
|\tilde{T}_{\ell-1} \cup \tilde{T}_\ell| \geq n - n_1.
\] (4)

It follows from (2), (4) that there exists \( a \in S_{k-1} \cup S_k \) such that \( b = M(a) \in \tilde{T}_{\ell-1} \cup \tilde{T}_\ell \). The paths \( \hat{P}_a, \hat{P}_b \) define an alternating path from \( a_0 \) to \( b_0 \) (perhaps only using parts of the paths) which together with the edge \{\( a_0, b_0 \)\} defines an alternating cycle \( C \) such that \( \Delta(M \oplus C) < \Delta(M) \). We have \( < \) here because we have removed an edge of \( M \cap Q_1 \).

**Verification of (1), (3):** It follows from Lemma 1(d) that
\[
|S_1| = |T_1| \geq \frac{\alpha_{\min} \log n}{20} - O(1).
\]

Now suppose that \( |S_1| \leq |S_i| \leq \gamma n / \log n \). Let
\[
Z = \{ b \in B : \exists a \in S_i \text{ s.t. } \{a, b\} \in Q \setminus (X_a) \}.
\]

Then \( e_G(S_i : Z) \geq |S_i|(|\alpha_{\min} \log n|)/20 \) and applying Lemma 1(b) we see that \( |Z| \geq |S_i|(|\alpha_{\min} \log n|)/40 \). Applying Lemma 1(d) we see that \( |S_{i+1}| \geq |S_i|(|\alpha_{\min} \log n|)/40 - L|S_i| \) and (1) follows. The argument for (3) is similar.

\[\square\]

4 Concluding Remarks

We have established that w.h.p. \( mcp(G) \) is almost all of \([0, n]^q\) and posed the question of finding the exact threshold for \( mcp(G) = [0, n]^q \). It is straightforward to extend our results to randomly colored \( G_{n,p} \). Basically, one can take \( B = A \). It would be of some interest to analyse other spanning subgraphs from this point of view e.g. Hamilton cycles.

References