On disjoint perfect tree-matchings in random graphs

Justyna Kurkowiak

Department of Discrete Mathematics
Adam Mickiewicz University
Poznań, Poland

Abstract

For an arbitrary tree T, a T-matching in G is a set of vertex-disjoint subgraphs of G which are isomorphic to T. A T-matching which is a spanning subgraph of G is called a perfect T-matching. For any t-vertex tree T we find a threshold probability function $\tilde{p} = \tilde{p}(n)$ for the existence of r edge-disjoint perfect T-matchings in a random graph $G(n, p)$.

1. Introduction

A random graph $G(n,p)$ is a graph obtained from the complete graph K_n by independent deletion of each edge with probability $1-p$. Similarly, a random bipartite graph $G(n,n,p)$ is a graph obtained from the complete bipartite graph $K_{n,n}$ by independent deletion of each edge with probability $1-p$. We say that a random graph possesses a property Q *asymptotically almost surely (a.a.s.)* if the probability of possessing Q converges to 1 as $n \to \infty$.

A *matching* in a graph G is a set of vertex-disjoint edges. A *perfect matching* is a matching which covers every vertex of G. A necessary condition for the existence of a perfect matching is the lack of isolated vertices. It turns out that for most graphs this is also sufficient - the threshold for containing a perfect matching coincides with that for disappearance of isolated vertices. The random graph $G(n,n,p)$, as well as $G(n,p)$, *a.a.s.* has isolated vertices when $np - \log n \to -\infty$. These vertices *a.a.s.* disappear when $np - \log n \to \infty$.

Graphs which contain a perfect matching are characterized by Hall's theorem in the bipartite case and Tutte's theorem in the arbitrary case. Because the condition of Tutte's theorem is more complex, the random bipartite graph was treated first. In 1964 Erdős and Rényi [1] proved the following result.

Research supported by KBN grant 2 PO3A 023 09.

Theorem 1. *If $np - \log n \to \infty$ then*

$$\lim_{n\to\infty} P(G(n,n,p) \textit{ has a perfect matching}) = 1.$$

Two years later the same authors showed a similar result for the random graph $G(n,p)$ [3]. Their proof was based on Tutte's theorem.

Theorem 2. *Assume n is even. If $np - \log n \to \infty$ then*

$$\lim_{n\to\infty} P(G(n,p) \textit{ has a perfect matching}) = 1.$$

A natural extension of Theorem 1 is the following result of Erdős and Rényi from 1968 [2].

Theorem 3. *For each r, if $np - \log n - (r-1)\log\log n \to \infty$ then*

$$\lim_{n\to\infty} P(G(n,n,p) \textit{ has } r \textit{ disjoint perfect matchings}) = 1.$$

Also in this case a necessary condition turned out to be asymptotically sufficient — the threshold for the existence of r disjoint perfect matchings coincides with that for the disappearance of vertices of degree less than r. An analogous extension of Theorem 2 was obtained by Shamir and Upfal in 1981. In fact the theorem below follows from the proof of their main theorem [6].

Theorem 4. *Assume n is even. For each r, if $np - \log n - (r-1)\log\log n \to \infty$ then*

$$\lim_{n\to\infty} P(G(n,p) \textit{ has } r \textit{ disjoint perfect matchings}) = 1.$$

Let H be an arbitrary graph. An *H-matching* in a graph G is a set of vertex-disjoint subgraphs of G which are all isomorphic to H. An H-matching will be called a *perfect H-matching* if it is a spanning subgraph of G. (A perfect K_2-matching is just a perfect matching).

In [4] Łuczak and Ruciński proved a result for perfect T-matchings in $G(n,p)$, where T is a tree. The result below is a corollary of the main theorem from [4].

Theorem 5. *Let T be a tree on t vertices and assume n is divisible by t. If $np - \log n \to \infty$ then*

$$\lim_{n\to\infty} P(G(n,p) \textit{ has a perfect } T\textit{-matching}) = 1.$$

In the case $t = 2$, Theorem 5 reduces to Theorem 2. In the proof of Theorem 5 a perfect T-matching was built from $t-1$ perfect matchings in suitably defined bipartite subgraphs of $G(n,p)$. Thus, although in the general case perfect matchings are characterized by Tutte's theorem, Łuczak and Ruciński proposed in fact an alternative proof of Theorem 2 using Hall's theorem. In the present paper their approach is combined with an idea from [2] to derive the following result about edge-disjoint perfect T-matchings.

Theorem 6. *Let T be a tree on t vertices ($t \geq 2$) and assume n is divisible by t. For each r, if $np - \log n - (r-1)\log\log n \to \infty$ then*

$$\lim_{n\to\infty} P(G(n,p) \textit{ has } r \textit{ edge-disjoint perfect } T\textit{-matchings}) = 1.$$

2. Preliminaries

Assume that $(r-1)\log\log n \leq np - \log n \leq r \log\log n$. Firstly we will prove several properties of the random graph $G(n,p)$.

Denote by $\Delta_{G(n,p)}$ the maximum degree in $G(n,p)$.

LEMMA 1. *A.a.s. $\Delta_{G(n,p)} < 4\log n$.*

Proof. Denote by X the number of vertices with degree at least $4\log n$. Then, from Markov's inequality

$$\begin{aligned} P(X>0) &\leq EX \leq n\binom{n}{4\log n} p^{4\log n} \\ &\leq n\left(\frac{enp}{4\log n}\right)^{4\log n} = O\left(n\left(\frac{e}{4}\right)^{4\log n}(\log n)^{4r}\right) = o(1). \quad \square \end{aligned}$$

LEMMA 2. *For all fixed integers k and d, a.a.s. there is no k-vertex tree in $G(n,p)$ with more than one vertex of degree at most d in $G(n,p)$.*

Proof. Denote by Y the number of k-vertex trees in $G(n,p)$ with at least two vertices of degree at most d. From Markov's inequality we have:

$$\begin{aligned} P(Y>0) \leq EY &\leq \binom{n}{k} k^{k-2} p^{k-1} \binom{k}{2} \left[\sum_{i=0}^{d} \binom{n-k}{i} p^i (1-p)^{n-k-i}\right]^2 \\ &= O\left(n^k p^{k-1}\left[\sum_{i=0}^{d} n^i p^i e^{-np}\right]^2\right) = O(np^k[(np)^d n^{-1}(\log n)^{-r+1}]^2) \\ &= O((\log n)^{k-1+2d} n^{-1} (\log n)^{-2r+2}) = o(1). \quad \square \end{aligned}$$

Let $B(n,p)$ be a random variable having binomial distribution with parameters n and p. Suppose that $m < np(1-\epsilon)$, for some $\epsilon > 0$. Since

$$\frac{P(B(n,p)=m-1)}{P(B(n,p)=m)} = \frac{m(1-p)}{(n-m+1)p} < \frac{m}{np},$$

then

$$(*)\qquad \begin{aligned} P(B(n,p)\le m) &= \sum_{l=0}^{m} P(B(n,p)=l)\\ &< \sum_{l=0}^{m}\left(\frac{m}{np}\right)^{l} P(B(n,p)=m) = O(1)P(B(n,p)=m). \end{aligned}$$

We will use the above property in the next lemma.

Let T be a tree on t vertices ($t \ge 2$) and assume that n is divisible by t. Let us arbitrarily partition $[n] = V_1 \cup V_2 \cup \cdots \cup V_t, where|V_i| = \frac{n}{t}, for i = 1, 2, \ldots, t.$ We will call a vertex v *bad* if $d_i(v) < \frac{\log n}{4t^3}$ for some $i = 1, 2, \ldots, t$ where $d_i(v) = |N_{G(n,p)}(v) \cap V_i|$.

LEMMA 3. *For all fixed integers k and t ($k > t \ge 2$), a.a.s.*

(1) *there is no k-vertex tree in $G(n,p)$ with more than t bad vertices;*
(2) *there are no more than $n^{1-0.72/t}$ bad vertices.*

Proof.

(1) Denote by Z the number of k-vertex trees in $G(n,p)$ with at least $t+1$ *bad* vertices. Then, using (*),

$$\begin{aligned} P(Z>0) \le EZ &\le \binom{n}{k} k^{k-2} p^{k-1} \binom{k}{t+1}\left[tP\left(B\left(\frac{n}{t}-k;p\right) \le \frac{\log n}{4t^3}\right)\right]^{t+1}\\ &= O\left(n^k p^{k-1}\left[\binom{\frac{n}{t}}{\frac{\log n}{4t^3}} p^{\frac{\log n}{4t^3}} (1-p)^{\frac{n}{t}-\frac{\log n}{4t^3}}\right]^{t+1}\right)\\ &= O\left(n(np)^{k-1}\left[(4t^2 e)^{\frac{\log n}{4t^3}} (\log n)^{\frac{r}{4t^3}} e^{-\frac{np}{t}}\right]^{t+1}\right)\\ &= O\left(n(\log n)^{k-1}\left(n^{\frac{3}{4t^3}} n^{\frac{logt}{2t^3}} (\log n)^{\frac{r}{4t^3}} n^{-\frac{1}{t}}\right)^{t+1}\right)\\ &= O\left((\log n)^{k-1+\frac{(t+1)r}{4t^3}} n^{\frac{3(t+1)+2(t+1)\log t}{4t^3}} n^{-\frac{1}{t}}\right) = o(1). \end{aligned}$$

(2) Denote by W the number of *bad* vertices in $G(n,p)$. Again by (*)

$$\begin{aligned} EW &\le ntP\left(B\left(\frac{n}{t}-1;p\right) \le \frac{\log n}{4t^3}\right)\\ &= O(1)n(4t^2e)^{\frac{\log n}{4t^3}} (\log n)^{\frac{r}{4t^3}} n^{-\frac{1}{t}}\\ &= O(1)nn^{-\frac{1}{t}+\frac{2\log t+3}{4t^3}}\\ &= o(n^{1-0.72/t}). \end{aligned}$$

Hence $P(W > n^{1-0.72/t}) = o(1)$. $\square$

By the first moment method it is easy to prove the following lemma:

LEMMA 4.

(1) *A.a.s. for every pair* (S_1, S_2) *of disjoint subsets of sizes* $|S_1| \geq |S_2| > \frac{n(\log\log n)^2}{\log n}$ *there are at least* $\log\log\log n |S_1|$ *edges between them in* $G(n,p)$

(2) *A.a.s. every subset* S *such that* $|S| \leq 2\frac{n(\log\log n)^2}{\log n}$ *induces in* $G(n,p)$ *less than* $(\log\log n)^3|S|$ *edges.* □

The next lemma and its proof can be found in [4].

LEMMA 5. *For every* $c > 0$ *there is* n_0 *such that the following holds. Let* G *be a graph on* $n \geq n_0$ *vertices which satisfies the following two conditions:*

(1) *For every pair* (S_1, S_2) *of disjoint subsets of sizes* $|S_1| \geq |S_2| > \frac{n(\log\log n)^2}{\log n}$ *there is an edge between them in* G

(2) *Every subset* S *with* $|S| \leq 2\frac{n(\log\log n)^2}{\log n}$ *induces in* G *less than* $(\log\log n)^3|S|$ *edges.*

Then every bipartite subgraph B *of* G *induced by the bipartition* $W_1, W_2 \subset [n]$, $w = |W_1| = |W_2| \geq \frac{n}{2t}$ *and such that the minimum degree* δ_B *in* B *is at least* $c\log n$ *has a perfect matching.*

Proof. Denote by $u = \frac{n(\log\log n)^2}{\log n}$.

Suppose there is a bipartite subgraph B with bipartition (W_1, W_2): $|W_1| = |W_2| = w \geq n/2t$, $\delta_B > c\log n$ which satisfies the conditions (1) and (2) and without a perfect matching. By Hall's theorem there exist $S \subseteq W_1$ such that $|N_B(S)| < |S|$. Let us consider three cases:

A.$|S| \leq u$. Then $|S| \cup |N_B(S)| < 2|S| \leq 2u$. But since all the edges from S go to $N_B(S)$, there are at least $c\log n|S|$ edges within a set of order less than $2u$ — a contradiction with (2).

B. $|N_B(S)| \geq w - u$. Then $|W_1 - S| < |W_2 - N_B(S)| < u$ and $W_1 - S$ is a neighborhood of $W_2 - N_B(S)$. Similiarly as in the previous case, $|W_1 - S| \cup |W_2 - N_B(S)| < 2u$ which gives a contradiction again with (2).

C. $|S| > u$ and $|N_B(S)| < w - u$. Hence $|W_2 - N_B(S)| > u$, so there are two sets each of size at least u with no edge between them — a contradiction with (1). □

3. The main proof

We will call a vertex v *small* if $deg_{G(n,p)}(v) < t^2 + (r-1)(t-1)$ (otherwise v is called *large*).

Denote by $\mathcal{A}_T$ the event that $G(n,p)$ contains r edge-disjoint perfect T-matchings $M_1, M_2, \ldots, M_r$ and by $\mathcal{A}'_T$ the event that such matchings $M_1, M_2, \ldots, M_r$ exist and in addition, that every *small* vertex in $G(n,p)$ is a pendant vertex in each of $M_1, M_2, \ldots, M_r$. (The condition that every *small* vertex in $G(n,p)$ has to be pendant in each of $M_1, M_2, \ldots, M_r$ is needed because otherwise it could happen that there would not be enough edges to have r edge-disjoint perfect T-matchings

— for example if T is a star on t vertices). To prove that $P(\mathcal{A}_T) \to 1$ as $n \to \infty$ it is enough to show that $P(\mathcal{A}'_T) \to 1$ as $n \to \infty$.

It was observed in [2] that if a graph G does not have r disjoint perfect matchings, then there is a way of deleting some edges in such a way that every degree decreases by at most $r-1$, and the remaining graph has no perfect matchings at all. Simply remove all edges of a maximal family of disjoint perfect matchings in G. Now we apply this observation in the context of perfect T-matchings.

Given a graph G, let $\mathcal{H}_G$ be the family of all spanning subgraphs H of G such that for all v, $deg_H(v) \geq \max\{1, deg_G(v) - (r-1)(t-1)\}$. Let A''_T be the event that for every $H \in \mathcal{H}_{G(n,p)}$ there is a perfect T-matching M in H such that every *small* vertex (the notion of *a small* vertex remains with respect to $G(n,p)$) is a pendant vertex in M. The event A''_T implies A'_T. Indeed, assume to the contrary that A''_T holds but A'_T does not, i.e. there are at most $r-1$ edge-disjoint perfect T-matchings in $G(n,p)$ in which every *small* vertex is pendant. Let us remove them all from the edge set of $G(n,p)$. The remaining subgraph H belongs to $\mathcal{H}_{G(n,p)}$ and there is no perfect T-matching in H such that every *small* vertex is pendant — a contradiction. Thus to prove that $P(A'_T) \to 1$ it is enough to show that $P(A''_T) \to 1$.

Let us arbitrarily partition, as for Lemma 3, $[n] = V_1 \cup V_2 \cup \cdots \cup V_t$, $|V_i| = \frac{n}{t}$, $i = 1, 2, \ldots, t$. Let G be a graph from the space $G(n,p)$, $n \geq n_0$, which satisfies the conditions of Lemmas 1–4 for $d = t^2 + (r-1)(t-1)$, $k = 2t-1$ in Lemma 2, $k = t^2 + t$ in Lemma 3, and let $H \in \mathcal{H}_G$. Firstly we shall find a partial T-matching in H containing all *bad* vertices and such that every *small* vertex is pendant. Then, after removing the vertices of this partial T-matching from H, for the remnant partition $V'_1, V_2, \ldots, V'_t$ we might have $|V'_i| \neq |V'_j|$ for some $i, j \in \{1, 2, \ldots, t\}$. To make the partition balanced again, we will move as few vertices as possible trying not to lower the degrees of the remaining vertices too extensively. Then to complete the proof of Theorem 6 it will be enough to apply Lemma 5.

For fixed k and for every vertex v, let us denote by $N_G{}^{(k)}(v)$ the set of vertices of distance at most k from v. We order all *bad* vertices by increasing degrees $deg(v_1) \leq deg(v_2) \leq \cdots \leq deg(v_b)$. As a matter of fact it is enough to take the *small* vertices first. Assume we have already found vertex-disjoint copies of T in H, $T_{v_1}, T_{v_2}, \ldots, T_{v_{i-1}}$, such that for every $j \in \{1, 2, \ldots, i-1\}$, v_j belongs to T_{v_j} and, moreover, if v_j is *small* then it is a pendant vertex of T_{v_j}. Consider a vertex v_i.

Assume further that v_i is *small.* Then no vertex from $N_H{}^{(t-1)}(v_i)$ is in $V(T_{v_1}) \cup V(T_{v_2}) \cup \cdots \cup V(T_{v_{i-1}})$, since otherwise there would be at least two *small* vertices in a tree on $2t-1$ vertices (a contradiction with Lemma 2). Thus v_i has at least one neighbor, say v, and v, as well as every vertex in $N_H{}^{(t-2)}(v)$, is *large*, so they have enough neighbors to build a copy of T, which is vertex-disjoint from T_{v_1}, $T_{v_2}, \ldots, T_{v_{i-1}}$. In this way one can build a copy T_{v_i} of a tree T in which v_i is a pendant vertex.

Assume now that v_i is *bad* and *large*, which in particular implies that $deg_H(v_i) \geq t^2$. The vertex v_i, as well as every vertex in $N_H{}^{(t-1)}(v_i)$, has at most $(t-1)t$

neighbors in the set $V(T_{v_1}) \cup V(T_{v_2}) \cup \cdots \cup V(T_{v_{i-1}})$, since otherwise there would be at least $t+1$ *bad* vertices in a tree on t^2+t vertices (see Lemma 3(1)). Hence again one can build a copy of T which contains v_i and is vertex-disjoint from T_{v_1}, $T_{v_2}, \ldots, T_{v_{i-1}}$.

In this way one can match all *bad* vertices into vertex-disjoint copies $T_{v_1}, T_{v_2}, \ldots,$ T_{v_b} of T in H so that every *small* vertex is pendant. Some of these copies may coincide. Let us remove the set $V_0 = V(T_1) \cup V(T_2) \cup \cdots \cup V(T_b)$ from H. By Lemma 3(2), $|V_0| \le tn^{1-0.72/t}$. Let $V_i' = V_i - V_0, \quad i = 1, 2, \ldots, t$. This remnant partition is no longer balanced, i.e. possibly $|V_i'| \ne |V_j'|$ for some $i, j \in \{1, 2, \ldots, t\}$. Set $s_i = |V_i'| - \frac{n-|V_0|}{t}$ and observe that $|s_i| \le |V_0| \le tn^{1-0.72/t}$ (the worst case is when all copies $T_{v_1}, T_{v_2}, \ldots, T_{v_b}$ are in the set V_i).

We have to make this new partition balanced again by moving some vertices around. By doing so we do not want to lower the degrees $d_i(v)$ for $i = 1, 2, \ldots, t$, too extensively. Hence the best strategy is to move vertices which form a 2-independent set in H, i.e. no two of them have a common neighbor. This way the parameters $d_i(v)$ will be affected by at most 1.

In every graph F on n vertices and with maximum degree at most Δ there exists an independent set of order at least $\frac{n}{\Delta+1}$. It can be greedily found by repeated inclusion of available vertices whose neighbors are immediately discarded. In fact, given a balanced partition $(U_1, \ldots, U_t)$ of $V(F)$, this construction can be easily adapted to yield an independent set I such that for each $i = 1, \ldots, t$, $|I \cap U_i| \ge n/(t^2\Delta + t) - 1$ (simply, include the vertices to I in an order alternating the elements of $U_1, \ldots, U_t$; then after each round of t inclusions, at most $t\Delta + 1$ vertices will be deleted from any given set U_i).

This applied to the square of H yields, via Lemma 1, the presence of a 2-independent set in H, intersecting each set V_i' in at least $\frac{n}{16t^2\log^2 n + t} - |V_0| \ge \frac{n}{17t^2\log^2 n}$.

That is more than we need. If $s_i > 0$, we have to move from the set V_i' exactly s_i vertices to the other sets. Let I_i be a 2-independent set of order s_i in $G[V_i']$ if $s_i > 0$, and let $\{J_i \ : \ s_i < 0\}$ be an arbitrary partition of the set $\bigcup_{s_i > 0} I_i$ such that $|J_i| = -s_i$. Let

$$V_i'' = \begin{cases} V_i' - I_i & \text{if} \quad s_i > 0, \\ V_i' \cup J_i & \text{if} \quad s_i < 0, \\ V_i'' = V_i' & \text{otherwise}. \end{cases}$$

For every pair i, j consider the bipartite graph $B = B_{ij}$ induced in H by (V_i'', V_j''). Now all we need is a perfect matching in $t-1$ bipartite graphs generated by pairs V_i, V_j for which the vertex i is joined with vertex j in a tree T. These perfect matchings will together constitute a perfect T-matching in $H[V \setminus V_0]$, completing the proof. The existence of a perfect matching in B will be shown by Lemma 5. We have $|V_i'| \ge \frac{n}{t} - |V_0| \ge \frac{n}{2t}$, for sufficiently large n. Let us estimate the degrees of vertices in B. For every $v \in V_i''$, v was a *good* vertex in G (with respect to the original partition $V_1, V_2, \ldots, V_t$). Hence $d_j(v) \ge \frac{\log n}{4t^3}$. By the definition of H, the degrees of vertices in G and in H differ by at most $(r-1)(t-1)$. By removing the set V_0 we could lower a degree by at most t^2. The worst case is when a vertex v is joined to every vertex in each of some t trees from $T_{v_1}, T_{v_2}, \ldots, T_{v_b}$. It cannot be joined to more than t trees, since otherwise there would be at least $t+1$ *bad* vertices in a tree on $t(t+1)+1$ vertices. Finally, by passing from $V_1', V_2', \ldots, V_t'$ to $V_1'', V_2'', \ldots, V_t''$ we could have decreased the number of neighbors of a vertex by at most 1. Hence, for every vertex $v \in V_i'' \cup V_j''$, and for n large enough

$$deg_B(v) > \frac{\log n}{4t^3} - (r-1)(t-1) - t^2 - 1 > \frac{\log n}{8t^3}.$$

Condition (2) of Lemma 5 is the same as in Lemma 4. Because the degrees of every vertex in G and H differ by at most $(r-1)(t-1)$, the condition (1) of Lemma 5 holds as well, by Lemma 4(1). Finally, by applying Lemma 5 with $c = \frac{1}{8t^3}$, we complete the proof of Theorem 6.

Acknowledgements

The author thanks Andrzej Ruciński for suggesting the problem and for many helpful remarks.

References

1. P.Erdős,A.Rényi, *On random matrices*, Publ. Math. Inst. Hung. Acad. Sci. **8** (1964), 455–461.
2. P.Erdős,A.Rényi, *On random matrices II*, Studia Sci. Math. Hung. **3** (1968), 459–464.
3. P.Erdős,A.Rényi, *On the existance of a factor of degree one of a connected random graph*, Acta Math. Acad. Sci. Hung. **17** (1966), 359–368.
4. T.Łuczak,A.Ruciński, *Tree matching in random graphs*, Report 87489-OR, Bonn (1991).
5. T.Łuczak,A.Ruciński, *Tree matching in graph processes*, SIAM J.Disc. Math **4, No.1** (1991), 107–120.
6. E.Shamir, E.Upfal, *On factors in random graphs*, Israel J. Math. **39** (1981), 296–302.

(Received 6/11/97)