Triangle factors in random graphs

Michael Krivelevich
Department of Mathematics,
Raymond and Beverly Sackler Faculty of Exact Sciences,
Tel-Aviv University, Tel-Aviv, Israel

September 9, 1996

Abstract

For a graph $G = (V, E)$ on n vertices, where 3 divides n, a triangle factor is a subgraph of G, consisting of $n/3$ vertex disjoint triangles (complete graphs on three vertices). We discuss the problem of determining the minimal probability $p = p(n)$, for which a random graph $G \in \mathcal{G}(n,p)$ contains almost surely a triangle factor. This problem (in a more general setting) has been studied by Alon and Yuster and by Ruciński, their approach implies $p = O((\log n/n)^{1/2})$. Our main result is that $p = O(n^{-3/5})$ already suffices. The proof is based on a multiple use of the Janson inequality. Our approach can be extended to improve known results about the threshold for the existence of an H-factor in $\mathcal{G}(n,p)$ for various graphs H.

1 Introduction

Let H be a graph on h vertices. If h divides n, we say that a graph G on n vertices contains an *H-factor*, if G contains n/h vertex disjoint copies of H. Thus, for example, a K^2-factor is a perfect matching.

As usual, we define $\mathcal{G}(n,p)$ as the probability space, consisting of all labelled graphs with vertex set $V = \{1, \ldots, n\}$, where a graph $G = (V, E) \in$

$\mathcal{G}(n,p)$ has probability $P[G] = p^{|E|}(1-p)^{\binom{n}{2}-|E|}$, the probability p may depend on n. In other words, each edge $(i,j) \in \binom{V}{2}$ is chosen to be an edge of $G \in \mathcal{G}(n,p)$ with probability p, all choices being independent. For a graph property $\mathcal{A}$ we say that a random graph $G \in \mathcal{G}(n,p)$ **whp** (*with high probability*) has $\mathcal{A}$, if the probability of $\mathcal{A}$ tends to 1 as n tends to infinity.

In this paper we consider the following problem. Let H be a fixed graph on h vertices, and assume h divides n. What is then the minimal probability $p = p(n)$ asserting that a random graph $G \in \mathcal{G}(n,p)$ **whp** has an H-factor? For the case $H = K^2$ the solution has been given by Erdős and Rényi in [3], they showed that $p = O(\log n/n)$ suffices to have **whp** a perfect matching in $G \in \mathcal{G}(n,p)$. For the case of a general graph H the problem remains unsolved. Some partial results have been obtained by Alon and Yuster [2] and by Ruciński [5]. However, the problem is still open for many important classes of graphs, in particular, for the case $H = K^r$ for every $r \geq 3$.

In this paper we are mostly concerned with the case $H = K^3$, this graph is actually the smallest one for which the problem is yet unsolved. This interesting problem is mentioned by Erdős in his Appendix to the monograph of Alon and Spencer [1]. We would like to note, however, that the approach developed in this paper can be applied as well to other graphs H.

For the case $H = K^3$, the method of Alon and Yuster and of Ruciński shows that it suffices to take $p = C(\ln n/n)^{1/2}$ for some constant $C > 0$. Here we improve this result by proving the following

Theorem 1 *A random graph $G \in \mathcal{G}(n, 1200n^{-3/5})$* **whp** *contains a triangle factor, assuming 3 divides n.*

The constant 1200 in the statement of the theorem can certainly be improved. We do not make any special attempt to optimize it.

Throughout the paper, we assume n to be sufficiently large where necessary. We will occasionally omit floor and ceiling signs for the sake of convenience.

The rest of the paper is organized as follows. In Section 2 we describe the framework of the Janson inequality used as the main technical tool in our proof. In Section 3 we show that a random graph $G \in \mathcal{G}(n,p)$ **whp** contains an almost triangle factor, that is, loosely speaking, a family of vertex disjoint triangles covering all but n^ϵ vertices for some constant $0 < \epsilon < 1$. Then we define a certain family of graphs, which we call H_0-trees (Section 4) and show that our random graph **whp** contains large H_0-trees (Section 5). Finally (Section 6), these H_0-trees are used to turn an almost triangle factor into the desired triangle factor. In Section 7, we discuss the applications of the presented approach for estimating the threshold for the existence of an H-factor for other graphs H, in particular, for the case $H = K^r$, $r > 3$.

2 The Janson inequality

In the course of the proof we will make a multiple use of the powerful inequality of Janson, first described in [4] (see also [1], Ch. 8). The following particular scheme of the inequality will suffice for our purposes.

Let $\mathcal{S}$ be a family of labelled subgraphs of a complete graph on n labelled vertices. Each edge of this complete graph is chosen to be an edge of a random graph $G \in \mathcal{G}(n,p)$ with probability p, all choices being independent. For each member $S \in \mathcal{S}$ we define the corresponding indicator random variable X_S which takes the value 1 if $S \subseteq G$, and the value 0 otherwise. Now define

$$X = \sum_{S \in \mathcal{S}} X_S \ ,$$

that is, X counts the number of subgraphs from $\mathcal{S}$ that turn out to be subgraphs of G. Also, let

$$\Delta = 2 \sum_{\substack{S, S' \in \mathcal{S} \\ E(S) \cap E(S') \neq \emptyset}} P[(X_S = 1) \wedge (X_{S'} = 1)] \ . \tag{1}$$

Then, assuming $P[X_S = 1] \le \delta$ for every $S \in \mathcal{S}$, Janson's inequality states that

$$P[X = 0] \le \exp\left\{-EX + \frac{\Delta}{2(1-\delta)}\right\} . \tag{2}$$

3 Covering almost all vertices by triangles

As the starting point of our proof, we show that if $G \in \mathcal{G}(n,p)$ with $p = \Theta(n^{-3/5})$, then **whp** almost all vertices can be covered by a family of vertex disjoint triangles.

Proposition 1 **whp** *every set of at least $n^{0.95}$ vertices of a random graph $G \in \mathcal{G}(n,p)$, where $p = Cn^{-3/5}$ for any absolute constant $C > 0$, contains a copy of a triangle.*

Proof. For a fixed subset $V_0 \subset V(G)$ of size $|V_0| = n^{0.95}$, denote by $\mathcal{S}$ the family of all triangles of a complete graph on V_0. For each triangle $S \in \mathcal{S}$ we denote by X_S the corresponding indicator random variable. Let $X = \sum_{S \in \mathcal{S}} X_S$, then

$$EX = \binom{|V_0|}{3} p^3 = \Theta(n^{1.05}) .$$

In order to apply Janson's inequality, define as in (1)

$$\begin{aligned} \Delta &= 2 \sum_{\substack{S,S' \in \mathcal{S} \\ E(S) \cap E(S') \neq \emptyset}} P[(X_S = 1) \wedge (X_{S'} = 1)] \\ &= \binom{|V_0|}{3} p^3 (|V_0| - 3) p^2 = O(n^{4/5}) = o(EX) . \end{aligned}$$

Since $\delta = P[X_S = 1] = p^3 = o(1)$, we have by (2)

$$P[X = 0] \le \exp\left\{-EX + \frac{\Delta}{2(1-\delta)}\right\} = \exp(-\Theta(n^{1.05})) .$$

Hence the probability of the existence of a set V_0 of size $|V_0| = \lceil n^{0.95} \rceil$, spanning no triangles, can be bounded from above by

$$\binom{n}{n^{0.95}} P[X = 0] \leq 2^n e^{-\Theta(n^{1.05})} = o(1) ,$$

that is, **whp** every set of $n^{0.95}$ (and therefore of at least $n^{0.95}$) vertices spans a copy of a triangle. □

Corollary 1 *For any constant $C > 0$, a random graph $G \in \mathcal{G}(n,p)$ with $p = Cn^{-3/5}$* **whp** *contains a family of vertex disjoint triangles, covering all but at most $n^{0.95}$ vertices.*

Proof. The desired family can be obtained by picking triangles one by one greedily. The above proposition shows that this process **whp** will not stop before less than $n^{0.95}$ vertices will remain uncovered. □

4 The graph H_0 and H_0-trees

The following simple graph H_0 plays a crucial role in our proof. H_0 has four vertices v_0, v_1, v_2 and v_3 and five edges $(v_0, v_1), (v_0, v_2), (v_1, v_2), (v_1, v_3)$ and (v_2, v_3) (that is, H_0 is a complete graph on four vertices with one edge deleted). An important property of H_0 is that if we remove one of the vertices v_0, v_3, then the remaining graph forms a triangle. For this reason, we will call the vertices v_0, v_3 *removable*, while the vertices v_1, v_2 will be called the *kernel* of H_0.

Now, based on H_0, we define the following family of graphs, each member of it being called an H_0-*tree*. With each H_0-tree $T = (V, E)$ we associate a vertex subset $R \subseteq V(T)$, which is called the *set of removable vertices.* Here is the recursive definition of H_0-trees.

Definition 1 **1)** *A graph H_0 is H_0-tree with the set of removable vertices $R = \{v_0, v_3\}$;*

2) *If* $T = (V, E)$ *is an* H_0*-tree with the set of removable vertices* $R \subseteq V(T)$ *and* H *is a copy of* H_0 *with removable vertices* u_0, u_3 *and kernel* $\{u_1, u_2\}$ *so that* $V(H) \cap V(T) = V(H) \cap R = \{u_0\}$*, then the graph* $T' = (V', E')$*, defined by* $V' = V(T) \cup V(H)$*,* $E' = E(T) \cup E(H)$*, is an* H_0*-tree with the set of removable vertices* $R' = R(T) \cup \{u_3\}$*;*

3) *Every* H_0*-tree can be obtained from* H_0 *by applying 2).*

In words, an H_0-tree T' can be obtained by taking a union of an H_0-tree T and a copy of H_0, sharing exactly one vertex that is removable for both of them.

The following proposition states some properties of H_0-trees.

Proposition 2 *Every* H_0*-tree* $T = (V, E)$ *with the set of removable vertices* $R \subseteq V(T)$ *has the following properties:*

1) $|V(T)| \equiv 1 \pmod 3$*;*

2) $|R| \geq |V(T)|/3$*;*

3) *For every* $v \in R$*, the graph* $T \setminus \{v\}$ *contains a triangle factor.*

(That is why the set R is called the set of removable vertices.)

Proof. The proof is by induction on $|V(T)|$.

1) and **2)** If $T = H_0$, then $|V(T)| \equiv 1 \pmod 3$ and also $|R(T)| = 2 > |V(T)|/3$. Each application of part 2) of Definition 1 adds three new vertices to T, one of them being added to the set of removable vertices;

3) Let T be obtained by joining an H_0-tree $T_0 = (V_0, E_0)$ with the set of removable vertices $R_0 \subseteq V_0$ and a copy H of H_0 on vertices u_0, u_1, u_2, u_3, where u_0 and u_3 are removable and $V_0 \cap V(H) = \{u_0\}$. If a vertex $r \in R_0 \setminus \{u_0\}$ is deleted from T, then by induction the graph $T_0 \setminus \{r\}$ contains a triangle factor, which can be completed to a triangle factor of $T \setminus \{r\}$ by adding the triangle u_1, u_2, u_3. If u_0 is deleted, then a desired triangle factor is obtained by adding the triangle u_1, u_2, u_3 to a triangle factor in $T_0 \setminus \{u_0\}$, existing by the induction hypothesis. Finally, if u_3 is deleted from T, then the graph $T_0 \setminus \{u_0\}$ contains by induction a triangle factor, to which we add the triangle u_0, u_1, u_2. □

The above proposition shows that, having a family $\mathcal{F}$ of vertex disjoint triangles and an H_0-tree T such that $V(\mathcal{F}) \cap V(T) = \emptyset$, we can use any removable vertex $v \in R(T)$ to build a new triangle with vertices yet uncovered by $\mathcal{F}$. The deletion of v from T results in the new graph T', having a triangle factor. This triangle factor can be then added to $\mathcal{F}$.

5 Finding large H_0-trees

In this section we show that when $p_0 = 6n^{-3/5}$, then **whp** a random graph $G \in \mathcal{G}(n, p_0)$ contains large vertex disjoint H_0-trees. Since in the sequel we will make use of the existence of disjoint H_0-trees of various sizes, the result is presented in the following parametric form.

Lemma 1 *If $p_0 = 6n^{-3/5}$, then for every integer k, satisfying $4 \le k \le n/6$ and $k \equiv 1 \pmod 3$, a random graph $G \in \mathcal{G}(n, p_0)$* **whp** *contains $\left\lfloor \frac{n}{6k} \right\rfloor$ vertex disjoint copies of H_0-trees, each having k vertices.*

The proof of the lemma is based on the following

Proposition 3 *Let $p_0 = 6n^{-3/5}$. Then* **whp** *for every triple of disjoint subsets U', U'', W of the vertex set of a random graph $G \in \mathcal{G}(n, p_0)$, satisfying $|U'| \ge n/18$, $|U''| \ge n/6$, $|W| \ge n/3$, there exists in G a copy of the graph H_0, having its kernel vertices in W, one of its removable vertices in U' and the other one in U''.*

The constants 1/18, 1/6, 1/3 in the above proposition are chosen somewhat arbitrarily, they are specific for its use in the proof of Lemma 1.

Proof. Clearly, it suffices to prove the proposition for the case $|U'| = n/18$, $|U''| = n/6$, $|W| = n/3$.

For a fixed triple U', U'', W satisfying the above restriction, let us denote by $\mathcal{S}$ the family of all copies of the graph H_0 in the complete graph on $U' \cup U'' \cup W$, satisfying the proposition requirements. For each such copy S we denote by X_S the corresponding indicator random variable, taking the value 1 iff $S \subset G$. Let also

$$X = \sum_{S \in \mathcal{S}} X_S \ .$$

Then

$$EX = \sum_{S \in \mathcal{S}} EX_S = \frac{n}{18}\frac{n}{6}\binom{\frac{n}{3}}{2} p^5 > 3n \ .$$

Let

$$\begin{aligned} \Delta &= 2 \sum_{\substack{S,S' \in \mathcal{S} \\ E(S) \cap E(S') \neq \emptyset}} P[(X_S = 1) \wedge (X_{S'} = 1)] \\ &= 2 \sum_{S \in \mathcal{S}} P[X_S = 1] \sum_{\substack{S' \in \mathcal{S} \\ E(S') \cap E(S) \neq \emptyset}} P[X_{S'} = 1 | X_S = 1] \\ &= EX \sum_{\substack{S' \in \mathcal{S} \\ E(S') \cap E(S) \neq \emptyset}} P[X_{S'} = 1 | X_S = 1] \end{aligned}$$

for some fixed $S \in \mathcal{S}$. A routine consideration of all possible mutual configurations of S and S' yields

$$\Delta = EX(O(n^2p^4 + np^2 + np^3)) = o(EX) \ .$$

Since $\delta = P[X_S = 1] = o(1)$, we have by Janson's inequality (2)

$$P[X = 0] \leq e^{-EX(1+o(1))} < e^{-3n} \ .$$

Hence the probability of the existence of a triple U', U'', W satisfying the above requirements and not containing a desired copy of H_0, can be bounded from above by

$$\binom{n}{|U'|}\binom{n}{|U''|}\binom{n}{|W|}P[X=0]$$
$$\leq (2^n)^3 e^{-3n} = o(1)\ . \quad \Box$$

Proof of Lemma 1. Assuming that the assertion of the above proposition holds, we prove the lemma deterministically. Let k be an integer, satisfying the conditions of the lemma. Suppose we have already found t such H_0-trees, each with a vertex set of size k and where $0 \leq t < n/6k$. Denote the union of their vertex sets by V_0. Clearly, $|V_0| = kt \leq n/6$. We put V_0 aside and show that the remaining vertices $V_1 = V \setminus V_0$ still contain an H_0-tree on k vertices.

Let us consider the following iterative process. At each stage of the process we have a family $\mathcal{T}_i = \{T_1, \ldots, T_m\}$ of vertex disjoint subgraphs of G, a vertex subset $U_j \subseteq V(T_j)$, $1 \leq j \leq m$, being associated with each member of $\mathcal{T}_i$. Initially, we choose arbitrarily $n/6$ vertices $v_1, \ldots, v_{n/6}$ from V_1 and set $T_j = \{v_j\}$, $U_j = \{v_j\}$, $\mathcal{T}_1 = \{T_1, \ldots, T_{n/6}\}$. Assume now that at stage $i \geq 1$ we have a family $\mathcal{T}_i = \{T_j\}_{j=1}^{m(i)}$, satisfying the following conditions:

(1) every member $T \in \mathcal{T}_i$ is either a single vertex or an H_0-tree with the set of removable vertices $R(T) = U(T)$;

(2) $\frac{n}{6} \leq \sum_{T \in \mathcal{T}_i} |V(T)| \leq \frac{n}{3}$.

Note that these conditions are clearly satisfied for our initial choice of $\mathcal{T}_1$. We first check whether there exists a subgraph $T_j \in \mathcal{T}_i$ with $|T_j| \geq k$. If such a subgraph indeed exists, the process terminates with T_j as the output. Otherwise, set $U' = \bigcup_{T \in \mathcal{T}_i} U(T)$. By Proposition 2 and conditions (1), (2), $|U'| \geq n/18$. Then we choose arbitrarily $n/6$ vertices from $V_1 \setminus \bigcup_{T \in \mathcal{T}_i} V(T)$ and denote this set by U'', let also $W = V_1 \setminus (\bigcup_{T \in \mathcal{T}_i} V(T) \cup U'')$. It follows from the above discussion that $|W| \geq n/3$. According to Proposition 3, there exists a copy H of the graph H_0 with its kernel vertices u_1, u_2 in W, one removable vertex u_0 in some set $U(T_j)$ and the other u_3 in U''. Now we replace the subgraph T_j in the family $\mathcal{T}_i$ by a new subgraph T_j' with $V(T_j') = V(T_j) \cup V(H)$ and $E(T_j') = E(T_j) \cup E(H)$, we set also $U(T_j') = U(T_j) \cup \{u_3\}$. Clearly, T_j' is an H_0-tree with the set of removable vertices $R(T_j') = U(T_j')$, also $|V(T_j')| = |V(T_j)| + 3$ and $|U(T_j')| = |U(T_j)| + 1$. Hence a new family $\mathcal{T}_i' = \mathcal{T}_i - T_j + T_j'$ satisfies condition (1). If $\sum_{T \in \mathcal{T}_i'} |V(T)| \leq n/3$, then we set $\mathcal{T}_{i+1} = \mathcal{T}_i'$, otherwise we get $\mathcal{T}_{i+1}$ by deleting from $\mathcal{T}_i'$ a subgraph T having the smallest number of vertices, if several such subgraphs exist, we choose one of them arbitrarily. Since all subgraphs of $\mathcal{T}_i$ are of size at most $n/6$, a new family $\mathcal{T}_{i+1}$ satisfies $n/6 \leq \sum_{T \in \mathcal{T}_{i+1}} |V(T)| \leq n/3$, and we may proceed to the next step.

One can easily see from the description of the process that the average size of a subgraph in $\mathcal{T}_i$ increases at each step. Therefore after a finite number of steps (which may depend on n) the process terminates. Since at each step the size of some member of $\mathcal{T}_i$ increases by 3, we can not 'skip' over the desired size k of an H_0-tree. $\Box$

6 Increasing the size of an almost triangle factor and covering all vertices

The proof of Theorem 1 is based on the following key lemma.

Lemma 2 *Define sequences $\{p_l\}_{l=0}^{\infty}$ and $\{\epsilon_l\}_{l=0}^{\infty}$ by $p_0 = 6n^{-3/5}$, $p_l = p_0 + \frac{6}{5}p_{l-1} - \frac{6}{5}p_0 p_{l-1}$ for $l \geq 1$; and $\epsilon_l = 0.95 - 0.05l$ for $l \geq 0$. Then for every integer l satisfying $0 \leq l \leq 18$, a random graph $G \in \mathcal{G}(n, p_l)$* **whp** *contains a family of vertex disjoint triangles covering all but at most n^{ϵ_l} vertices.*

Proof. By induction on l. For $l = 0$, the assertion of the lemma follows from Corollary 1.

Assume now $l \geq 1$. Note that $1-p_l = (1-p_0)(1-\frac{6}{5}p_{l-1})$. This enables us to represent a random graph $G \in \mathcal{G}(n,p_l)$ as a union of two random graphs: $G' \in \mathcal{G}(n,p_0)$ and $G'' \in \mathcal{G}(n,\frac{6}{5}p_{l-1})$. Note also that $p_l = \Theta(n^{-3/5})$ for every $1 \leq l \leq 18$.

Let us first expose the edges of G'. Let $m_1 = \lfloor n^{\epsilon_{l-1}} \rfloor$. Let m_2 be the smallest integer satisfying $m_2 \geq 2m_1 + n^{\epsilon_l}$ and $m_1 + m_2 \equiv 0 \pmod 3$. In the calculations below we substitute $m_1 = n^{\epsilon_{l-1}}$, $m_2 = 2n^{\epsilon_{l-1}} + n^{\epsilon_l}$, this does not affect the correctness of our proofs. Also, let k be the largest integer satisfying $km_2 \leq n/6$ and $k \equiv 1 \pmod 3$. Clearly, $k = (1+o(1))(n^{1-\epsilon_{l-1}}/12)$. According to Lemma 1 G' **whp** contains $\lfloor n/6k \rfloor \geq m_2$ vertex disjoint H_0-trees, each having k vertices. We choose m_2 of these subgraphs and denote this family by $\mathcal{T}_0 = \{T_1, \dots, T_{m_2}\}$. Let

$$
\begin{aligned}
V_0 &= \bigcup_{j=1}^{m_2} V(T_j) \ , \\
V_1 &= V \setminus V_0 \ ,
\end{aligned}
$$

then $|V_1| \geq 5n/6$.

Now we expose the edges of G''. Let us first look at the edges of G'' inside V_1. By the induction hypothesis the subgraph $G''[V_1]$, considered as a random graph on $|V_1| \geq \frac{5}{6}n$ vertices with edge probability $\frac{6}{5}p_{l-1}$, **whp** contains a family of vertex disjoint triangles covering all but at most $|V_1|^{\epsilon_{l-1}} < m_1$ vertices. Let us fix a subfamily $\mathcal{F}_0$ of triangles, covering all but *exactly* m_1 vertices of V_1. Here we use the assumption $m_1 + m_2 \equiv 0 \pmod 3$. Since every $T_j \in \mathcal{T}_0$ is an H_0-tree, we have $|V(T_j)| \equiv 1 \pmod 3$, therefore $|V_0| \equiv m_2 \pmod 3$. Hence we can drop from the family of triangles, covering all but at most $(5n/6)^{\epsilon_{l-1}} < m_1$ vertices of V_1 some triangles in order to cover all but exactly m_1 vertices. We denote the set of vertices of V_1, uncovered by $\mathcal{F}_0$, by $W = \{v_1, \dots, v_{m_1}\}$.

Let us define the following process. At the beginning of step i ($1 \leq i \leq m_1$) we have a subfamily $\mathcal{T}_{i-1} \subseteq \mathcal{T}_0$, $|\mathcal{T}_{i-1}| = m_2 - 2(i-1)$, and a family of vertex disjoint triangles $\mathcal{F}_{i-1} \supseteq \mathcal{F}_0$. We aim to find a triangle v_i, u_1, u_2, where $u_1 \in R(T_{j_1})$ and $u_2 \in R(T_{j_2})$, $T_{j_1}, T_{j_2} \in \mathcal{T}_{i-1}$, $j_1 \neq j_2$. If such a triangle exists, then we add it to $\mathcal{F}_{i-1}$, obtaining $\mathcal{F}'_{i-1}$. Since u_l and u_2 are removable vertices of T_{j_1} and T_{j_2}, respectively, both subgraphs $T_{j_1} \setminus \{u_1\}$ and $T_{j_2} \setminus \{u_2\}$ contain triangle factors. We add these triangle factors to $\mathcal{F}'_{i-1}$, thus obtaining $\mathcal{F}_i$. The subgraphs T_{j_1} and T_{j_2} are then removed from $\mathcal{T}_{i-1}$, resulting in a new family $\mathcal{T}_i$.

Now we claim that **whp** this process can be performed successfully for m_1 steps, that is, for all points of W. To prove this, we look at the edges of G'' inside V_0 and between V_0 and V_1. At step i of the above process the family $\mathcal{T}_{i-1}$ contains at least $m_2 - 2m_1 \geq n^{\epsilon_l}$ subgraphs. Choose $m_2 - 2m_1$ subgraphs from $\mathcal{T}_{i-1}$ arbitrarily and denote them by $T_1, \dots, T_{m_2-2m_1}$. Note that $\sum_{j=1}^{m_2-2m_1} |V(T_j)| = k(m_2 - 2m_1) = \Theta(n^{0.95})$. By Proposition 2, part 2, $|R(T_j)| \geq |V(T_j)|/3$, so $|R(T_j)| \geq k/3$ and thus $|R(T_j)| = \Theta(n^{1-\epsilon_{l-1}})$ for every $1 \leq j \leq m_2 - 2m_1$.

The random variable Y, counting the number of edges of G'' with endpoints in distinct removable sets of $T_1, \dots, T_{m_2-2m_1}$, is binomially distributed with parameters $\sum_{1 \leq j_1 < j_2 \leq m_2 - 2m_1} |R(T_{j_1})||R(T_{j_2})| = \Theta(n^{1.9})$ and $\frac{6}{5}p_{l-1}$. Using standard large deviation inequalities (see, e.g., [1], Appendix A, Theorem A.13) it can be shown that **whp** for *every* choice of $m_2 - 2m_1$ subgraphs from $\mathcal{T}_0$ the corresponding random variable Y satisfies $Y \geq EY(1 + o(1))$ and therefore $Y = \Omega(n^{1.9}p_{l-1})$. Also, for a vertex $u \in \bigcup_{j=1}^{m_2-2m_1} R(T_j)$ the expectation of the number of edges of G'', connecting u with vertices from V_0, is $\frac{6}{5}|(V_0| - 1)p_{l-1} \leq np_{l-1}/4$, and again large deviation inequalities imply that **whp** this number of edges does not exceed np for every such u.

Let us examine now the subgraph G_i of G'', whose vertex set is $\bigcup_{j=1}^{m_2-2m_1} R(T_j)$ and whose edge set consists of all edges having endpoints in distinct $R(T_j)$. According to the above discussion, this subgraph **whp** has $\Omega(n^{1.9}p_{l-1})$ edges, and the maximal degree in it does not exceed np_{l-1}.

Now we expose the edges of G'', connecting the current vertex v_i with the vertices of G_i. If v_i is connected with both endpoints of $e = (u_1, u_2) \in E(G_i)$, then v_i, u_1, u_2 form the desired triangle. Therefore, we need to prove that such an edge e exists with probability $1 - o(m_1^{-1})$. We show it again through the Janson inequality. For every $e = (u_1, u_2) \in E(G_i)$ let $S_e = \{(v_i, u_1), (v_i, u_2)\}$. Let $\mathcal{S} = \{S_e : e \in E(G_i)\}$. For every $S_e \in \mathcal{S}$ the corresponding indicator random variable X_{S_e} takes the value 1 iff $(v_i, u_1), (v_i, u_2) \in E(G'')$, where $e = (u_1, u_2)$. Clearly, $EX_{S_e} = (\frac{6}{5}p_{l-1})^2 = o(1)$. Denoting $X = \sum_{S_e \in \mathcal{S}} X_{S_e}$, we see that

$$EX = |E(G_i)|(\frac{6}{5}p_{l-1})^2 = \Omega(n^{1.9}p_{l-1}^3) .$$

Recalling definition (1) of Δ, we have

$$\begin{aligned}
\Delta &= 2 \sum_{\substack{e,e' \in E(G_i), \\ e \cap e' \neq \emptyset}} P[(X_{S_e} = 1) \wedge (X_{S_{e'}} = 1)] \\
&= 2 \sum_{v \in V_0} \sum_{\substack{e,e' \in E(G_i), \\ v = e \cap e'}} P[(X_{S_e} = 1) \wedge (X_{S_{e'}} = 1)] \\
&= 2 \sum_{\bigcup_{j=1}^{m_2 - 2m_1} V(T_j)} \binom{d_{G_i}(v)}{2} \left(\frac{6}{5}p_{l-1}\right)^3 = O(n^{0.95}\, n^2 p_{l-1}^2\, p_{l-1}^3) = O(n^{2.95} p_{l-1}^5) .
\end{aligned}$$

But $p_{l-1} = \Theta(n^{-3/5})$, and therefore $EX = \Omega(n^{0.1})$ and $\Delta = o(1)$. Hence Janson's inequality (2) implies that

$$P[X = 0] \leq e^{-EX(1+o(1))} = e^{-\Omega(n^{0.1})} ,$$

showing that the i-th step of the above described process is successful with probability $1 - e^{-\Omega(n^{0.1})}$. Since the number of steps is $m_1 < n$, the process **whp** can be performed successfully for all vertices of W.

After the process has finished, we have a family $\mathcal{T}_{m_1} = \{T_1, \ldots, T_{m_2 - 2m_1}\}$ of H_0-trees. Deleting a removable vertex $u_j \in R(T_j)$ from every $T_j \in \mathcal{T}_{m_1}$, we then find a triangle factor in each T_j. Their union is added to $\mathcal{F}_{m_1}$. Now

$\mathcal{F}_{m_1}$ covers all vertices of V but $\{u_1, \ldots, u_{m_2-2m_1}\}$, thus forming a desired family of triangles. $\Box$

Proof of Theorem 1. The proof is very similar to that of Lemma 2, therefore we indicate only some differences, leaving the details to the reader. We refer to the notation of Lemma 2 and its proof. Consider a random graph $G \in \mathcal{G}(n, p_{19})$ and represent it as a union of $G' \in \mathcal{G}(n, p_0)$ and $G'' \in \mathcal{G}(n, p_{18})$. Note that by Lemma 2, G'' **whp** contains a family of vertex disjoint triangles, covering all but at most $n^{0.05}$ vertices. Define m_1 to be the largest integer not exceeding $n^{0.05}$ and satisfying $m_1 \equiv 0 \pmod 3$, and define crucially $m_2 = 2m_1$. Then if the m_1–step process of Lemma 2 will be successful, the family $\mathcal{T}_{m_1}$ will be empty. Therefore, after step m_1 we will have a triangle factor. Note that at each step i, $1 \leq i \leq m_1$, the family $\mathcal{T}_i$ contains at least $m_2 - 2m_1 + 2 = 2$ H_0-trees with $k = \Theta(n^{0.95})$ vertices. Arguments similar to those of the proof of Lemma 2 show that step i is successful with probability $1 - o(m_1^{-1})$. We omit detailed calculations.

Note that we have shown that $G \in \mathcal{G}(n, p_{19})$ **whp** has a triangle factor. In order to estimate p_{19}, we write $p_l \leq p_0 + \frac{6}{5} p_{l-1}$ for every $l \geq 1$. This implies $p_l \leq 5\left(\left(\frac{6}{5}\right)^{l+1} - 1\right) p_0$. Thus $p_{19} < 187 p_0 < 1200 n^{-3/5}$. $\Box$

7 Concluding remarks

1) One may wonder why the asymptotic order of probability $p = \Theta(n^{-3/5})$ in Theorem 1 cannot be reduced. Indeed, at almost all steps of the above proof a smaller order of probability would suffice. It is easy to see that when $p = o(n^{-3/5})$ the proof of Proposition 3 ceases to work. Moreover, if $p = o(n^{-3/5})$, for every fixed vertex $v \in V$ the expected number of copies of H_0, containing v, is o(1), and therefore we expect the copies of H_0 in $G \in \mathcal{G}(n, p)$ to be mostly vertex disjoint. It can be shown that in this case G **whp** does not contain an H_0-tree of size n^ϵ for any fixed $\epsilon > 0$.

2) A natural question is what is the right value of the threshold for the

existence of a triangle factor in a random graph $G \in \mathcal{G}(n,p)$. Our Theorem 1 shows that this threshold is at most $O(n^{-3/5})$. On the other hand, it can be proven, using for example Janson's inequality and the second moment method, that when $\binom{n-1}{2}p^3 = \ln n + w(n)$ for any function $w(n)$ tending to infinity with n, that **whp** every vertex of a random graph $G \in \mathcal{G}(n,p)$ participates in at least one triangle, while if $\binom{n-1}{2}p^3 = \ln n - w(n)$, then **whp** there exists at least one vertex that does not participate in any triangle. This motivates the following

Conjecture 1 $p(n) = (\log n)^{1/3} n^{-2/3}$ *is the threshold function for the existence of a triangle factor in a random graph* $G \in \mathcal{G}(n,p)$.

3) The approach of this paper can be used as well to get new results in the problem of determining the threshold for the appearance of an H-factor for graphs H other than a triangle. Let us illustrate this by taking $H = K^r$, a complete graph on r vertices, where $r \geq 4$ is fixed. Here is a brief outline of the proof.

a) Prove that if $p = Cn^{-\frac{2r}{(r-1)(r+2)}}$ for any constant $C > 0$, then **whp** every $n^{1-\frac{1}{10r}}$ vertices of a random graph $G \in \mathcal{G}(n,p)$ span a copy of K^r (cf. Proposition 1). Therefore **whp** a random graph $G \in \mathcal{G}(n,p)$ with $p = Cn^{-\frac{2r}{(r-1)(r+2)}}$ contains a family of vertex disjoint copies of K^r covering all but at most $n^{1-\frac{1}{10r}}$ vertices (cf. Corollary 1).

b) Define the graph $H_0 = K^{r+1} \setminus \{e\}$. The two removable vertices of H_0 are those having degree $r-1$, the other vertices form a kernel. Then define the notion of an H_0-tree in a manner similar to that of Definition 1. Every H_0-tree $T = (V,E)$ with the set of removable vertices $R \subset V(T)$ has the following properties (cf. Proposition 2):

1. $|V(T)| \equiv 1 \pmod r$;

2. $|R| \geq |V(T)|/r$;

3. For every $v \in R$, the graph $T \setminus \{v\}$ contains a K^r-factor.

Note that if $p = \Theta(n^{-\frac{2r}{(r-1)(r+2)}})$, then the expectation of the number of copies of H_0 in $G \in \mathcal{G}(n,p)$ is linear in n. This will enable us to build H_0-trees containing n^ϵ vertices for $0 < \epsilon < 1$.

c) Prove that if $p_0 = 6n^{-\frac{2r}{(r-1)(r+2)}}$, then **whp** for every triple of disjoint subsets U', U'', W of $V(G)$, $G \in \mathcal{G}(n,p)$, satisfying $|U'| = n/18$, $|U''| = n/6$ and $|W| = n/3$, there exists in G a copy of H_0 having its kernel vertices in W, one of its removable vertices in U' and the other in U'' (cf. Proposition 3). Derive from this that for every integer k satisfying $r+1 \le k \le n/6$ and $k \equiv 1 \pmod r$, a random graph $G \in \mathcal{G}(n,p)$ **whp** contains $\left\lfloor \frac{n}{6k} \right\rfloor$ vertex disjoint copies of H_0-trees, each having k vertices (cf. Lemma 1).

d) Define two sequences $\{p_l\}_{l=0}^\infty$ and $\{\epsilon_l\}_{l=0}^\infty$ by $p_0 = 6n^{-\frac{2r}{(r-1)(r+2)}}$, $p_l = p_0 + \frac{6}{5}p_{l-1} - \frac{6}{5}p_0p_{l-1}$ for $l \ge 1$; and $\epsilon_l = 1 - \frac{1}{10r}(l+1)$ for $l \ge 0$. Then prove that for every l satisfying $0 \le l \le 10r-2$, a random graph $G \in \mathcal{G}(n,p_l)$ **whp** contains a family of vertex disjoint copies of K^r covering all but at most n^{ϵ_l} vertices. The proof of this statement is very similar to that of Lemma 2. We mention only some differences. We define the numbers m_1 and m_2 so that they satisfy $m_1 + m_2 \equiv 0 \pmod r$ and $m_2 - (r-1)m_1 \approx n^{\epsilon_l}$. Also, k should satisfy $k \equiv 1 \pmod r$. Then we find in $G' \in \mathcal{G}(n,p_0)$ m_2 vertex disjoint H_0-trees, each having k vertices. Denote the union of their vertex sets by V_0, let also $V_1 = V \setminus V_0$. Now we look at the edges of $G'' \in \mathcal{G}(n, \frac{6}{5}p_{l-1})$. First, we find in V_1 a family $\mathcal{F}$ of of vertex disjoint copies of K^r, covering all but m_1 vertices. Denote these unmatched vertices by W. Then, for each vertex $v_i \in W$ we find a copy of K^r, containing v_i and having remaining $r-1$ vertices in removable sets of distinct H_0-trees.

e) Prove that $G \in \mathcal{G}(n, p_{10r-1})$ **whp** contains a K^r-factor. The proof can be shaped after that of Theorem 1. We define $m_1 \approx n^{\frac{1}{10r}}$, $m_1 \equiv 0 \pmod r$, $m_2 = (r-1)m_1$. Then k satisfies $k = \Theta(n^{1-\frac{1}{10r}})$. At each step of the m_1-step process the family $\mathcal{T}_i$ has at least $r-1$ H_0-trees, each having k vertices.

Estimating p_l from above gives $p_l \leq p_0 + \frac{6}{5} p_{l-1}$, hence $p_l \leq 5\left(\left(\frac{6}{5}\right)^{l+1} - 1\right) p_0$. Thus $p_{10r-1} \leq 5\left(\left(\frac{6}{5}\right)^{10r} - 1\right) p_0 < 30\left(\frac{6}{5}\right)^{10r} n^{-\frac{2r}{(r-1)(r+2)}}$. Summing the above, we claim the following

Theorem 2 *Let $r \geq 4$ be a fixed number. Then if $p(n) = 30\left(\frac{6}{5}\right)^{10r} n^{-\frac{2r}{(r-1)(r+2)}}$, then a random graph $G \in \mathcal{G}(n,p)$* **whp** *contains a K^r-factor, assuming r divides n.*

Acknowledgements. This research forms part of a Ph. D. thesis written by the author under the supervision of Professor Noga Alon. The author's work was supported in part by a Charles Clore Fellowship.

The author would like to thank the anonymous referee for many helpful comments.

References

[1] N. Alon and J. H. Spencer, **The probabilistic method**, Wiley, New York, 1992.

[2] N. Alon and R. Yuster, *Threshold functions for H-factors*, Comb., Prob. and Computing 2 (1993), 137–144.

[3] P. Erdős and A. Rényi, *On the existence of a factor of degree one in a connected random graph*, Acta Math. Acad. Sci. Hungar. 17 (1966), 359–368.

[4] S. Janson, T. Łuczak and A. Ruciński, An exponential bound for the probability of nonexistence of a specified subgraph in a random graph, in *Random Graphs' 87* (M. Karoński et al., eds.), Wiley, Chichester, 1990, 73–87.

[5] A. Ruciński, *Matching and covering the vertices of a random graph by copies of a given graph*, Discrete Math. 106 (1992), 185–197.