

No Dense Subgraphs Appear in the Triangle-free Graph Process

Stefanie Gerke

Royal Holloway College
University of London
Egham TW20 0EX, United Kingdom
`stefanie.gerke@rhul.ac.uk`

Tamás Makai

Royal Holloway College
University of London
Egham TW20 0EX, United Kingdom
`tamas.makai.2009@live.rhul.ac.uk`

Submitted: Nov 22, 2010; Accepted: Aug 3, 2011; Published: Aug 26, 2011

Mathematics Subject Classification: 05C80

Abstract

Consider the triangle-free graph process, which starts from the empty graph on n vertices and in every step an edge is added that is chosen uniformly at random from all non-edges that do not form a triangle with the existing edges. We will show that there exists a constant c such that asymptotically almost surely no copy of any fixed finite triangle-free graph on k vertices with at least ck edges appears in the triangle-free graph process.

1 Introduction

The Erdős-Rényi random graph process starts from the empty graph on n vertices $G_{n,0}$ and at the i th step, $G_{n,i}$ is obtained from $G_{n,i-1}$ by adding an edge chosen uniformly at random from the set of non-edges in $G_{n,i-1}$. We are interested in typical structural properties of $G_{n,m}$ when n is large. We say an event holds asymptotically almost surely (a.a.s.), if the probability that it occurs tends to one as the number of vertices n tends to infinity. The random graph process is well understood, partly as $G_{n,m}$ has strong connections to the random graph model $G_{n,p}$ when $p \approx m/\binom{n}{2}$. In $G_{n,p}$ each edge is present independently of the presence or absence of all other edges with probability p ; see [4], [7].

A variant of this process, namely the triangle-free graph process where at step i of the random graph process an edge is chosen uniformly at random from the set of non-edges in $G_{n,i-1}$ fulfilling the additional condition that when added to $G_{n,i-1}$ the graph remains triangle free. The process terminates when no more edges can be inserted. In this paper we show that there exists a constant c such that, given any fixed finite triangle-free

graph on k vertices with at least ck edges, a.a.s. no copy of it appears in the triangle-free graph process. For instance, large complete bipartite graphs a.a.s. do not appear in the triangle-free graph process.

The triangle-free process was first considered by Bollobás and Erdős [3]. The first results to appear in print were by Erdős, Suen and Winkler [5] who showed that the triangle-free graph process terminates a.a.s. in between $\Omega(n^{3/2})$ and $O(n^{3/2} \log n)$ steps. More recently, Bohman [1] strengthened this result by proving a conjecture of Spencer [9], namely that the final graph a.a.s. contains $\Theta(n^{3/2} \sqrt{\log n})$ edges. He also proved that the size of a maximum independent set a.a.s. is $O(\sqrt{n \log n})$. This implies Kim's result [8] on the lower bound of the Ramsey number $R(3, t) = \Omega(t^2 / \log t)$.

In his proof Bohman uses a variant of the differential equation method to track various variables closely through $m = \mu n^{3/2} \sqrt{\log n}$ steps for some small constant μ . The variable, crucial for determining bounds on the length of the process, is the number of non-edges that can be added. These non-edges are called *open pairs*. A non-edge that is not open is called *closed*. Additional variables are needed to track the process of closing a pair, some of which will be discussed in the following section.

Based on Bohman's results, subgraph counts have been established until m edges have been inserted. Bohman and Keevash [2] gave bounds on the number of copies of any fixed triangle-free graph F with $\max_{H \subset F} e_H / v_H < 2$ which hold a.a.s. for every step. Here, v_F denotes the number of vertices of F and e_F denotes the number of edges. They also proved that by step m a.a.s. a copy of every finite triangle-free graph with $\max_{H \subset F} e_H / v_H = 2$ is present, but no copy exists if $\max_{H \subset F} e_H / v_H > 2$. Wolfowitz [10] gave similar bounds, which hold a.a.s. for a given step, but his results include bounds for the case when $\max_{H \subset F} e_H / v_H = 2$.

However, dense triangle-free graphs could appear later in the process when Bohman's result does not apply anymore. Using Bohman's estimates, we will show that this is not the case. More precisely, we will show that for any placement of a fixed dense graph F into the vertex set of the random graph process, by the time $\mu n^{3/2} \sqrt{\log n}$ edges have been inserted a.a.s. one of its edges is closed. Therefore, no copy of F can be completed later in the process. Let us note that a related process, the random planar graph process (where at each step an edge is inserted if the graph remains planar) behaves differently. Gerke, Schlatter, Steger and Taraz [6] showed that a.a.s. the planar graph process contains a copy of any fixed planar graph after inserting just $(1 + \varepsilon)n$ edges.

2 Main Results

In the following we denote the edge set of $G_{n,i}$ by E_i . The vertex set of the process is denoted by V . The set of open pairs in $G_{n,i}$ is called O_i and we set $Q(i) = |O_i|$.

In order to track the number of open pairs, some additional variables are needed. For any pair of non-adjacent vertices u, v , Bohman [1] tracked the number of open, partial and complete vertices. Recall that a non-edge is open at step i if it can be inserted in the triangle-free process at step i without creating a triangle. A vertex w is *open* with

respect to the non-edge $\{u, v\}$ if both pairs $\{u, w\}$ and $\{v, w\}$ are open. A vertex w is *partial* with respect to $\{u, v\}$ if exactly one of $\{u, w\}$ and $\{v, w\}$ is open and the other is an edge. Finally, a vertex w is *complete* with respect to $\{u, v\}$ if $\{u, w\}$ and $\{v, w\}$ are edges. In particular, if there is a vertex which is complete with respect to $\{u, v\}$, then $\{u, v\}$ is a non-edge and it is closed, see Figure 1.

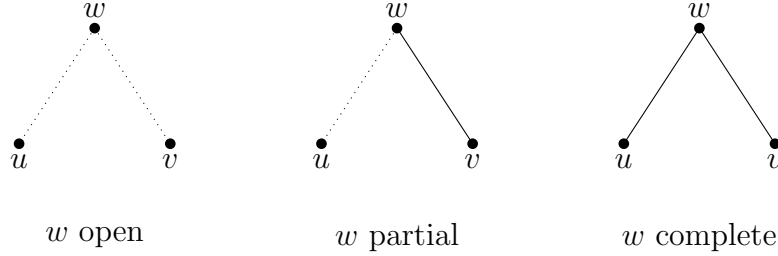


Figure 1. The vertex w is open/partial/complete with respect to $\{u, v\}$.
Dotted lines indicate open pairs, solid lines indicate edges.

In addition to the number of open pairs, our proof only requires the number of partial vertices for all non-adjacent pairs $\{u, v\}$. We denote the set of partial vertices of a non-edge $\{u, v\}$ at step i by $Y_{u,v}(i)$.

The following bounds were proven by Bohman [1] for the first $\mu n^{3/2} \sqrt{\log n}$ steps. (Bohman sets $\mu = 1/32$, however no effort was made to optimize the value.) For the remainder of the paper we set $\mu = 1/32$ and $m = \mu n^{3/2} \sqrt{\log n}$.

Definition 1. Let $\mathcal{H}(i)$ be the event that the following bounds hold for all $j \leq i$ and for all pairs $\{u, v\} \notin E_j$:

$$\begin{aligned} |Q(j) - n^2 q(t(j))| &\leq n^2 g_q(t(j)) \\ \left| |Y_{u,v}(j)| - \sqrt{n} y(t(j)) \right| &\leq \sqrt{n} g_y(t(j)) \end{aligned}$$

where

$$\begin{aligned} t(i) &= i/n^{3/2} \\ q(t) &= \exp(-4t^2)/2 \\ y(t) &= 4t \exp(-4t^2) \\ g_q(t) &= \begin{cases} \exp(41t^2 + 40t)n^{-1/6} & : t \leq 1 \\ \frac{\exp(41t^2 + 40t)}{t} n^{-1/6} & : t > 1 \end{cases} \\ g_y(t) &= \exp(41t^2 + 40t)n^{-1/6}. \end{aligned}$$

Theorem 1. [1] The event $\mathcal{H}(m)$ holds a.a.s..

Let $W \subset V$ and denote the number of edges spanned by W in $G_{n,i}$ by $e_W(i)$.

Lemma 2. Fix k . Let $\mathcal{S}_k(i)$ be the event that there exists $W \subset V$ with $|W| = k$ and $e_W(i) \geq 3k$. Then a.a.s. $\mathcal{S}_k(m)$ does not hold.

Note that $\overline{\mathcal{S}_k(m)}$ is decreasing and thus if it holds then it implies $\overline{\mathcal{S}_k(i)}$ for every $i < m$.

A sharper result showing that a.a.s. $\frac{e_W(m)}{|W|} \leq 2$ for any finite set of vertices W can be found in [2] and [10], however for completeness a short proof of Lemma 2 has been included.

Proof. Fix k vertices in V and denote this set by W . Let $\mathcal{A}_i$ be the event that an edge is added between two vertices in W at step i . For an index set T , let $\mathcal{A}_T^* = \bigcap_{j \in T} \mathcal{A}_j$.

Fix $i < m$ and let $T \subset [i]$ be such that $\mathcal{A}_T^* \cap \mathcal{H}(i) \neq \emptyset$ (where $[i]$ denotes $\{1, \dots, i\}$). Then since $q(t(i)) > q(t(m))$ for $i < m$ we have

$$\begin{aligned} P(\mathcal{A}_{i+1} | \mathcal{A}_T^* \cap \mathcal{H}(i)) &\leq \frac{k^2}{Q(i)} \leq \frac{k^2}{n^2(q(t(i)) - g_q(t(i)))} \leq \frac{2k^2}{n^2 q(t(i))} \\ &\leq \frac{2k^2}{n^2 q(t(m))} \leq \frac{4k^2}{n^2 \exp(-4\mu^2 \log n)} = \frac{4k^2}{n^{2-4\mu^2}}. \end{aligned}$$

For $i \leq j \leq m$, we have $\mathcal{H}(j) \subseteq \mathcal{H}(i)$ and thus $P(\mathcal{A}_{i+1} \cap \mathcal{A}_T^* \cap \mathcal{H}(i)) \geq P(\mathcal{A}_{i+1} \cap \mathcal{A}_T^* \cap \mathcal{H}(j))$. In addition, a.a.s. $P(\mathcal{H}(j)) = (1 + o(1))P(\mathcal{H}(i))$ as $\mathcal{H}(m)$ holds a.a.s. and $\mathcal{H}$ is monotone. Thus

$$(1 + o(1))P(\mathcal{A}_{i+1} \cap \mathcal{A}_T^* | \mathcal{H}(i)) \geq P(\mathcal{A}_{i+1} \cap \mathcal{A}_T^* | \mathcal{H}(j)).$$

Hence, letting t_1 denote the largest element of T , we have

$$\begin{aligned} P(e_W(m) \geq 3k | \mathcal{H}(m)) &\leq \sum_{T \subset [m], |T|=3k} P(\mathcal{A}_T^* | \mathcal{H}(m)) \\ &\leq \sum_{T \subset [m], |T|=3k} (1 + o(1))P(\mathcal{A}_T^* | \mathcal{H}(t_1 - 1)) \\ &= \sum_{T \subset [m], |T|=3k} (1 + o(1))P(\mathcal{A}_{t_1} | \mathcal{A}_{T \setminus \{t_1\}}^* \cap \mathcal{H}(t_1 - 1))P(\mathcal{A}_{T \setminus \{t_1\}}^* | \mathcal{H}(t_1 - 1)) \\ &\leq \left((1 + o(1)) \frac{4k^2}{n^{2-4\mu^2}} \right) \sum_{T \subset [m], |T|=3k} P(\mathcal{A}_{T \setminus \{t_1\}}^* | \mathcal{H}(t_1 - 1)). \end{aligned}$$

Repeating this argument yields

$$\begin{aligned} P(e_W(m) \geq 3k | \mathcal{H}(m)) &\leq \binom{m}{3k} \left((1 + o(1)) \frac{4k^2}{n^{2-4\mu^2}} \right)^{3k} \leq \left(\frac{(1 + o(1))em4k^2}{3kn^{2-4\mu^2}} \right)^{3k} \\ &\leq \left(\frac{(1 + o(1))4ke\mu n^{3/2} \sqrt{\log n}}{3n^{2-4\mu^2}} \right)^{3k} = o\left(\frac{1}{n^{3k/2-20k\mu^2}} \right). \end{aligned}$$

Since there are $\binom{n}{k}$ ways to select k vertices, it follows from the union bound that

$$P(\mathcal{S}_k(m) | \mathcal{H}(m)) \leq \binom{n}{k} o\left(\frac{1}{n^{3k/2-20k\mu^2}} \right) = o\left(\frac{n^k}{n^{3k/2-20k\mu^2}} \right) = o(1)$$

as μ^2 is sufficiently small. □

Given a fixed graph F and a graph G , we say that there exists a copy of F in G if an injective function $f : V(F) \rightarrow V(G)$ exists such that $\{f(u), f(v)\} \in E(G)$ for all $\{u, v\} \in E(F)$. We have just shown that no copy of a dense graph appears in the triangle-free process until the first m edges are inserted. We will now show that when m edges have been inserted at least one edge of any placement of a dense graph F is closed.

Theorem 3. *Let $\mathcal{T}$ be the event that in the triangle-free graph process there exists a copy of a graph F with k vertices and e edges satisfying $e \geq 10k/\mu^2$. Then a.a.s. $\mathcal{T}$ does not hold.*

Proof. Fix a set of vertices $W \subset V$ with $|W| = k$, and a set of pairs of vertices $E_F \subset W \times W$ such that if the pairs in E_F were inserted as edges they would form a copy of F on W . Let $\mathcal{C}_F(i)$ be the event that at least one pair in E_F is closed in $G_{n,i}$ and $\mathcal{O}_F(i)$ be the event that no pair is closed in $G_{n,i}$. Thus if $\mathcal{O}_F(i)$ holds then every pair in E_F is either an edge of $G_{n,i}$ or is in O_i . For the following, we assume that we are in the event $\mathcal{O}_F(i)$.

Note that a pair $\{u, v\}$ is closed at step $i + 1$ if and only if there is a partial vertex $w \in Y_{u,v}(i)$ and the missing edge is chosen. Thus the probability of closing a pair $s \in O_i$ is $|Y_s(i)|/Q(i)$. The problem is that an edge can close several pairs of vertices in E_F . The subset of $W \times W$ closed by $\{w_j, v\} \in O_i$ at step $i + 1$, with $v \notin W$, has size $|N_i(v) \cap W|$ where $N_i(v)$ denotes the neighbourhood of v in $G_{n,i}$ (see Figure 2).

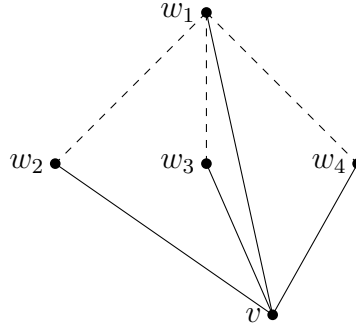


Figure 2. The edge $\{v, w_1\}$ closes $\{w_1, w_2\}, \{w_1, w_3\}$ and $\{w_1, w_4\}$

Let D_i be the set of vertices not in W that have more than 6 neighbours in W in $G_{n,i}$. Excluding the pairs with both vertices in W and the pairs with a vertex in D_i the remaining pairs close at most 6 pairs in $W \times W$ and in particular at most 6 pairs in E_F . Since we assume $\mathcal{O}_F(i)$ holds, $\sum_{f \in E_F \setminus E_i} |Y_f(i) \setminus (D_i \cup W)|$ counts any pair that closes a pair in E_F at step $i + 1$ at most 6 times.

Assume we are in $\mathcal{S}_{2k}(i)$, then the set D_i can have size at most k for otherwise $W \cup D_i$ would contain a set of $2k$ vertices that span more than $6k$ edges. Let $\mathcal{B}(i) = \overline{\mathcal{S}_k(i)} \cap \mathcal{S}_{2k}(i) \cap \mathcal{H}(i)$. Then

$$P(\mathcal{C}_F(i+1) | [\mathcal{O}_F(i) \cap \mathcal{B}(i)]) \geq \frac{\sum_{f \in E_F \setminus E_i} |Y_f(i) \setminus (D_i \cup W)|}{6Q(i)} \geq \frac{\sum_{f \in E_F \setminus E_i} (|Y_f(i)| - 2k)}{6Q(i)}.$$

Since the event $\overline{\mathcal{S}_k(i)}$ holds, there are less than $3k$ edges in E_F . Therefore, the sum in the previous inequality is over at least

$$(|E_F| - 3k) \geq (10/\mu^2 - 3)k \geq 9k/\mu^2$$

open pairs. Since $\mathcal{H}(i)$ holds, we have

$$P(\mathcal{C}_F(i+1)|[\mathcal{O}_F(i) \cap \mathcal{B}(i)]) \geq \frac{9k}{\mu^2} \frac{\sqrt{n}(y(t(i)) - g_y(t(i))) - 2k}{6n^2(q(t(i)) + g_q(t(i)))}.$$

If n is large then $q(t(i)) + g_q(t(i)) \leq 2q(t(i))$. If $m \geq i \geq n^{4/3}$ and n is large, then

$$y(t(i)) - g_y(t(i)) \geq \frac{y(t(i))}{2} \geq 2t(n^{4/3}) \exp(-4t^2(m)) = 2n^{-\frac{1}{6}-4\mu^2},$$

and since k is a constant

$$\frac{\sqrt{n}y(t(i))}{2} - 2k \geq \frac{7}{15}\sqrt{n}y(t(i)).$$

Thus if $m \geq i \geq n^{4/3}$ and n is large,

$$\begin{aligned} P(\mathcal{C}_F(i+1)|[\mathcal{O}_F(i) \cap \mathcal{B}(i)]) &\geq \frac{9k}{\mu^2} \frac{7\sqrt{n}y(t(i))/15}{6n^2(q(t(i)) + g_q(t(i)))} \\ &\geq \frac{7k}{\mu^2} \frac{\sqrt{n}y(t(i))}{20n^2q(t(i))} = \frac{7k}{\mu^2} \frac{4t(i) \exp(-4t^2(i))}{10n^{3/2} \exp(-4t^2(i))} \\ &= \frac{14ki}{5\mu^2n^3}, \end{aligned}$$

and hence

$$P(\mathcal{O}_F(i+1)|[\mathcal{O}_F(i) \cap \mathcal{B}(i)]) \leq 1 - \frac{14ki}{5\mu^2n^3} \leq \exp\left(-\frac{14ki}{5\mu^2n^3}\right).$$

Using $\mathcal{O}_F(i) \subset \mathcal{O}_F(i+1)$ and $\mathcal{B}(i) \subset \mathcal{B}(i+1)$, we have for sufficiently large n

$$\begin{aligned} P(\mathcal{O}_F(m) \cap \mathcal{B}(m)) &= \prod_{i=0}^{m-1} P(\mathcal{O}_F(i+1) \cap \mathcal{B}(i+1)|[\mathcal{O}_F(i) \cap \mathcal{B}(i)]) \\ &\leq \prod_{i=0}^{m-1} P(\mathcal{O}_F(i+1)|[\mathcal{O}_F(i) \cap \mathcal{B}(i)]) \\ &\leq \prod_{i=\lceil n^{4/3} \rceil}^{m-1} \exp\left(-\frac{14ki}{5\mu^2n^3}\right) = \exp\left(-\sum_{i=\lceil n^{4/3} \rceil}^{m-1} \frac{14ki}{5\mu^2n^3}\right) \\ &= \exp\left(-\frac{14k}{5\mu^2n^3} \left(\frac{m(m-1)}{2} - \frac{\lceil n^{4/3} \rceil(\lceil n^{4/3} \rceil - 1)}{2}\right)\right) \end{aligned}$$

$$\leq \exp\left(-\frac{4k}{3} \frac{m^2}{\mu^2 n^3}\right) = \exp\left(-\frac{4k}{3} \log n\right) = n^{-4k/3}.$$

As there are $\binom{n}{k} \frac{k!}{\text{aut}(F)}$ possible placements of F , applying the union bound gives

$$P(\mathcal{T} \cap \mathcal{B}(m)) \leq \binom{n}{k} k! n^{-4k/3} \leq n^{k-4k/3} = o(1).$$

□

Acknowledgement: We would like to thank the referee for helpful comments.

References

- [1] Tom Bohman. The triangle-free process. *Adv. Math.*, 221(5):1653–1677, 2009.
- [2] Tom Bohman and Peter Keevash. The early evolution of the H -free process. *Inventiones Mathematicae*, 181:291–336, 2010. 10.1007/s00222-010-0247-x.
- [3] Béla Bollobás. Personal communication.
- [4] Béla Bollobás. *Random graphs*, volume 73 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2001.
- [5] Paul Erdős, Stephen Suen, and Peter Winkler. On the size of a random maximal graph. In *Proceedings of the Sixth International Seminar on Random Graphs and Probabilistic Methods in Combinatorics and Computer Science, “Random Graphs ’93” (Poznań, 1993)*, volume 6, pages 309–318, 1995.
- [6] Stefanie Gerke, Dirk Schlatter, Angelika Steger, and Anusch Taraz. The random planar graph process. *Random Structures Algorithms*, 32(2):236–261, 2008.
- [7] Svante Janson, Tomasz Łuczak, and Andrzej Ruciński. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.
- [8] Jeong Han Kim. The Ramsey number $R(3, t)$ has order of magnitude $t^2/\log t$. *Random Structures Algorithms*, 7(3):173–207, 1995.
- [9] Joel Spencer. Maximal triangle-free graphs and Ramsey $R(3, t)$. 1995.
- [10] Guy Wolfowitz. Triangle-free subgraphs at the triangle-free process. *Random Structures and Algorithms*, to appear. arXiv:0903.1756.