AUSTRALASIAN JOURNAL OF COMBINATORICS
Volume **43** (2009), Pages 237–251

Cleaning random graphs with brushes

Paweł Prałat*

Department of Mathematics and Statistics
Dalhousie University
Halifax NS B3H 3J5
Canada
pralat@mathstat.dal.ca

Abstract

A model for *cleaning* a graph with brushes was recently introduced. In this paper, we consider the minimum number of brushes needed to clean a random graph $G(n, p = d/n)$ in this model, the so-called brush number. We show that the brush number of a random graph on n vertices is asymptotically almost surely (a.a.s.) $\frac{dn}{4}(1 + o(1))$ if the average degree is tending to infinity with n. For a constant $d > 1$, various upper and lower bounds are studied. For $d \leq 1$, we show that the number of brushes needed is a.a.s. $\frac{n}{4}(1 - \exp(-2d))(1 + o(1))$ and compute the probability that it attains its natural lower bound.

1 Introduction

The cleaning model, introduced in [4, 5], is a combination of the chip-firing game and edge-searching on a simple finite graph. Initially, every edge and vertex of a graph is *dirty* and a fixed number of brushes start on a set of vertices. At each step, a vertex v and all its incident edges which are dirty may be *cleaned* if there are at least as many brushes on v as there are incident dirty edges. When a vertex is cleaned, every incident dirty edge is traversed (that is, cleaned) by one and only one brush, and brushes cannot traverse a clean edge. See Figure 1 for an example of this cleaning process. The initial configuration has only 2 brushes, both at a. The solid edges are dirty and the dotted edges are clean. The circle indicates which vertex is cleaned next.

The assumption in [5], and taken here, is that *a graph is cleaned when every vertex has been cleaned.* If every vertex has been cleaned, it follows that every edge has been cleaned. It may be that a vertex v has no incident dirty edges at the time it is cleaned, in which case no brushes move from v. Although this viewpoint might seem unnatural, it simplified much of the analysis in [5].

* Research partially supported by grants from NSERC and MITACS

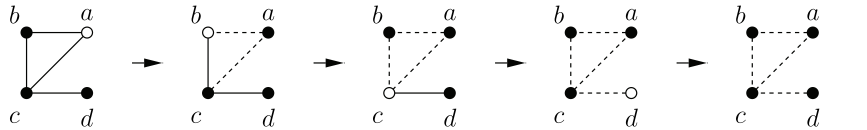


Figure 1: An example of the cleaning process for graph G.

In [1, 6], the minimum number of brushes needed to clean d-regular graphs in this model, focusing on the asymptotic number for random d-regular graphs, was considered. A degree-greedy algorithm was used to clean a random d-regular graph on n vertices (with dn even) and the differential equations method was used to find the (asymptotic) number of brushes needed to clean a random d-regular graph using this algorithm (for fixed d).

In this paper, we are interested in the asymptotic number of brushes needed to clean a binomial random graph. In Section 2, we give a definition of the model we study. In Section 3, we investigate dense graphs for which the average degree is growing with the size of a graph; an asymptotic almost sure value of the brush number is given. In Section 4, we deal with sparse graphs. If the average degree is greater than one, then various upper and lower bound are provided; otherwise an asymptotic almost sure value can be computed.

2 Definitions

The following cleaning algorithm and terminology was introduced in [5].

Formally, at each step t, $\omega_t(v)$ denotes the number of brushes at vertex v ($\omega_t : V \to \mathbb{N} \cup \{0\}$) and D_t denotes the set of dirty vertices. An edge $uv \in E$ is dirty if and only if both u and v are dirty: $\{u, v\} \subseteq D_t$. Finally, let $D_t(v)$ denote the number of dirty edges incident to v at step t:

$$D_t(v) = \begin{cases} |N(v) \cap D_t| & \text{if } v \in D_t \\ 0 & \text{otherwise.} \end{cases}$$

Definition 2.1 *The* ***cleaning process*** $\mathfrak{P}(G, \omega_0) = \{(\omega_t, D_t)\}_{t=0}^T$ *of an undirected graph* $G = (V, E)$ *with an* ***initial configuration of brushes*** ω_0 *is as follows:*

(0) *Initially, all vertices are dirty:* $D_0 = V$*; set* $t := 0$

(1) *Let* α_{t+1} *be any vertex in* D_t *such that* $\omega_t(\alpha_{t+1}) \geq D_t(\alpha_{t+1})$*. If no such vertex exists, then stop the process, set* $T = t$*, return the* ***cleaning sequence*** $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_T)$*, the* ***final set of dirty vertices*** D_T*, and the* ***final configuration of brushes*** ω_T

(2) *Clean α_{t+1} and all dirty incident edges by moving a brush from α_{t+1} to each dirty neighbour. More precisely, $D_{t+1} = D_t \setminus \{\alpha_{t+1}\}$, $\omega_{t+1}(\alpha_{t+1}) = \omega_t(\alpha_{t+1}) - D_t(\alpha_{t+1})$, and for every $v \in N(\alpha_{t+1}) \cap D_t$, $\omega_{t+1}(v) = \omega_t(v) + 1$ (the other values of ω_{t+1} remain the same as in ω_t)*

(3) $t := t + 1$ *and go back to* **(1)**

Note that for a graph G and initial configuration ω_0, the cleaning process can return different cleaning sequences and final configurations of brushes; consider, for example, an isolated edge uv and $\omega_0(u) = \omega_0(v) = 1$. It has been shown (see Theorem 2.1 in [5]), however, that the final set of dirty vertices is determined by G and ω_0. Thus, the following definition is natural.

Definition 2.2 *A graph $G = (V, E)$* ***can be cleaned*** *by the initial configuration of brushes ω_0 if the cleaning process $\mathfrak{P}(G, \omega_0)$ returns an empty final set of dirty vertices ($D_T = \emptyset$).*

Let the brush number, $b(G)$, be the minimum number of brushes needed to clean G, that is,

$$b(G) = \min_{\omega_0 : V \to \mathbb{N} \cup \{0\}} \Big\{ \sum_{v \in V} \omega_0(v) : G \text{ can be cleaned by } \omega_0 \Big\}.$$

Similarly, $b_\alpha(G)$ is defined as the minimum number of brushes needed to clean G using the cleaning sequence α.

It is clear that for every cleaning sequence α, we have $b_\alpha(G) \geq b(G)$ and $b(G) = \min_\alpha b_\alpha(G)$. (The last relation can be used as an alternative definition of $b(G)$.) In general, it is difficult to find $b(G)$, but $b_\alpha(G)$ can be easily computed. For this, it seems better not to choose the function ω_0 in advance, but to run the cleaning process in the order α, and compute the initial number of brushes needed to clean a vertex. We can adjust ω_0 along the way

$$\omega_0(\alpha_{t+1}) = \max\{2D_t(\alpha_{t+1}) - \deg(\alpha_{t+1}), 0\}, \tag{1}$$

for $t = 0, 1, \ldots, |V| - 1$, since that is the number of brushes we have to add over and above what we get for free. Indeed, at step t, vertex α_{t+1} has $D_t(\alpha_{t+1})$ dirty incident edges and $\deg(\alpha_{t+1}) - D_t(\alpha_{t+1})$ clean incident edges. So α_{t+1} must have received exactly $\deg(\alpha_{t+1}) - D_t(\alpha_{t+1})$ brushes from neighbouring vertices. If

$$(\deg(\alpha_{t+1}) - D_t(\alpha_{t+1})) - D_t(\alpha_{t+1}) \geq 0,$$

then α_{t+1} requires no additional brushes and we may set $\omega_0(\alpha_{t+1}) = 0$. Otherwise, α_{t+1} requires an additional

$$D_t(\alpha_{t+1}) - (\deg(\alpha_{t+1}) - D_t(\alpha_{t+1})) = 2D_t(\alpha_{t+1}) - \deg(\alpha_{t+1})$$

brushes in order to be cleaned at step t so we set $\omega_0(\alpha_{t+1}) = 2D_t(\alpha_{t+1}) - \deg(\alpha_{t+1})$.

Our main results refer to the probability space $\mathcal{G}(n,p) = (\Omega, \mathcal{F}, \mathbb{P})$ of random graphs, where Ω is the set of all graphs with vertex set $[n] = \{1, 2, \ldots, n\}$, $\mathcal{F}$ is the family of all subsets of Ω, and for every $G \in \Omega$

$$\mathbb{P}(G) = p^{|E(G)|}(1-p)^{\binom{n}{2}-|E(G)|}\,.$$

It can be viewed as a result of $\binom{n}{2}$ independent coin flipping, one for each pair of vertices, where the probability of success (that is, drawing an edge) is equal to p ($p = p(n)$ can tend to zero with n). We say that an event holds *asymptotically almost surely* (a.a.s.) if it holds with probability tending to 1 as $n \to \infty$.

We use $Po(\lambda)$ to denote the Poisson distribution. We write $X_n \xrightarrow{D} Z$ if X_n converges in distribution to Z, that is, $\mathbb{P}(X_n \le x) \to \mathbb{P}(Z \le x)$ for every real x that is a continuity point of $\mathbb{P}(Z \le x)$. The base of all logarithms is e.

3 Dense case

In this section we consider a brush number for a random graph $\mathcal{G}(n,p)$ for which the average degree is growing with the size of a graph. We state main result of this section below.

Theorem 3.1 *For $G \in \mathcal{G}(n,p)$, $p = p(n) = \omega(n)/n$ ($\omega(n)$ is any function tending to infinity with n) a.a.s.*

$$b(G) = \frac{pn^2}{4}(1 + o(1))\,.$$

We consider an upper and lower bound independently in the next subsections.

3.1 Upper bound

The following result provides an upper bound for the brush number of a general graph. Theorem 3.2 has been proved in [1, Theorem 3.7] but, since the proof is short, we present it here for completeness.

Theorem 3.2 ([1])

$$b(G) \le \frac{|E|}{2} + \frac{|V|}{4} - \frac{1}{4} \sum_{v \in V(G), \deg(v) \text{ is even}} \frac{1}{\deg(v)+1}$$

for any graph $G = (V, E)$.

Proof: Let π be a random permutation of the vertices of G taken with uniform distribution. We clean G according to this permutation to get the value of $b_\pi(G)$ (note that $b_\pi(G)$ is a random variable now). For a vertex $v \in V$, it follows from (1) that we have to assign to v exactly $X(v) = \max\{0, 2N^+(v) - \deg(v)\}$ brushes in the initial configuration, where $N^+(v)$ is the number of neighbors of v that follow it in the permutation (that is, the number of dirty neighbours of v at the time when v

is cleaned). The random variable $N^+(v)$ attains each of the values $0, 1, \dots, \deg(v)$ with probability $1/(\deg(v)+1)$; indeed, this follows from the fact that the random permutation π induces a uniform, random permutation on the set of $\deg(v)+1$ vertices consisting of v and its neighbors. Therefore the expected value of $X(v)$, for even $\deg(v)$, is

$$\frac{\deg(v) + (\deg(v)-2) + \cdots + 2}{\deg(v)+1} = \frac{\deg(v)+1}{4} - \frac{1}{4(\deg(v)+1)},$$

and for odd $\deg(v)$ it is

$$\frac{\deg(v) + (\deg(v)-2) + \cdots + 1}{\deg(v)+1} = \frac{\deg(v)+1}{4}.$$

Thus, by linearity of expectation,

$$\begin{aligned}\mathbb{E}b_\pi(G) &= \mathbb{E}\left(\sum_{v \in V} X(v)\right) = \sum_{v \in V} \mathbb{E}X(v) \\ &= \frac{|E|}{2} + \frac{|V|}{4} - \frac{1}{4} \sum_{v \in V(G), \deg(v) \text{ is even}} \frac{1}{\deg(v)+1},\end{aligned}$$

which means that there is a permutation π_0 such that $b(G) \le b_{\pi_0}(G) \le \mathbb{E}b_\pi(G)$ and the assertion holds. $\square$

For $G \in \mathcal{G}(n,p)$, $p = p(n) = \omega(n)/n$, let $X = |E(G)|$ (X is a random variable). The expected number of edges is

$$\mathbb{E}X = \binom{n}{2} p = \frac{n\omega(n)}{2}(1 + O(n^{-1})).$$

In order to finish the proof of the upper bound, we use the fact that a sum of independent random variables with large enough expected value is not too far from its mean (see, for example, Theorem 2.8 in [3]). From this it follows that, if $\varepsilon \le 3/2$, then

$$\mathbb{P}\left(|X - \mathbb{E}X| \ge \varepsilon \mathbb{E}X\right) \le 2\exp\left(-\frac{\varepsilon^2}{3}\mathbb{E}X\right). \tag{2}$$

Putting $\varepsilon = \log n/\sqrt{n\omega(n)}$ we get that a.a.s. $X \le \big(1+\varepsilon\big)\mathbb{E}X$ and thus a.a.s.

$$b(G) \le \frac{pn^2}{4}\big(1 + O(\varepsilon)\big) = \frac{pn^2}{4}(1 + o(1))$$

by Theorem 3.2.

3.2 Lower bound

A lower bound for random graphs can be obtained as follows. By [5, Theorem 3.2],

$$b(G) \geq \min_{S \subseteq V, |S| = \lfloor n/2 \rfloor} |E(S, V \setminus S)|\,, \tag{3}$$

where $E(S, V \setminus S)$ is the set of all edges between S and its complement. The proof is a simple corollary of the fact that the minimum obtained in (3) is a lower bound on the number of edges going from the first $\lfloor n/2 \rfloor$ vertices cleaned to elsewhere in the graph. (The minimum number of edges in a cut that splits the vertex set of a graph into two equal parts is called its bisection width.)

Now, fix any subset of vertices S of size $\lfloor n/2 \rfloor$ and let $Y = |E(S, V \setminus S)|$ (Y is a random variable). It is clear that

$$\mathbb{E}Y = \frac{pn^2}{4}\big(1 + O(n^{-1})\big) = \frac{\omega(n)n}{4}\big(1 + O(n^{-1})\big).$$

Using (2) with $\varepsilon = \omega(n)^{-1/3}$ we get that $Y \geq (1-\varepsilon)pn^2/4$ with probability $\exp(-\Omega(\omega(n)^{1/3}n)) = o(2^{-n})$. Thus, we can use the Stirling's formula ($n! = \sqrt{2\pi n}(n/e)^n(1 + o(1))$) to show that the expected number of sets S with less than $(1-\varepsilon)pn^2/4$ edges going from S to its complement is

$$\binom{n}{\lfloor n/2 \rfloor} \cdot o(2^{-n}) = \Theta(2^n/\sqrt{n}) \cdot o(2^{-n}) = o(1).$$

By Markov's inequality, we get an asymptotically almost sure lower bound on the brush number.

4 Sparse case

Perhaps the most studied phenomenon in the field of random graphs is the behaviour of the random graph when $p = d/n$ for d near 1. In the *subcritical phase* (that is, when $p = (1-s)/n$ and $s = s(n) \gg n^{-1/3}$), a.a.s. $G(n,p)$ consists of small trees and unicyclic components, and thus its structure is rather easy to study. We know that the giant component is formed from smaller ones during the so called *critical phase*, where $s = O(n^{-1/3})$. In the *supercritical phase* (that is, when $p = (1+s)/n$ and $s = s(n) \gg n^{-1/3}$), a.a.s. $G(n,p)$ consists of one complex component of size $\Omega(n^{2/3})$ and some number of small trees and unicyclic components of size $o(n^{2/3})$ each. Moreover, if $s = o(1)$, then the size of the largest component and the number of edges in this component are still $o(n)$. For $d > 1$ a.a.s. the size of the largest component is $\Theta(n)$ while the size of the second largest component is $O(\log n)$. For more information about random graphs see, for example, [3, 2].

4.1 The subcritical phase

We start our discussion from the subcritical phase but before we move to random structures we need some results on the brush number for deterministic graphs.

The model presented in this paper is one where the edges are continually recontaminated, say by algae, so that cleaning is regarded as an on-going process. Ideally, the final configuration of the brushes, after all the edges have been cleaned, should be a viable starting configuration to clean the graph again. We know that this is possible, even with the least number of brushes.

Theorem 4.1 ([5]) ***The Reversibility Theorem***
Given the initial configuration ω_0, suppose G can be cleaned yielding final configuration ω_n, $n = |V(G)|$. Then, given initial configuration $\tau_0 = \omega_n$, G can be cleaned yielding the final configuration $\tau_n = \omega_0$.

When a graph G is cleaned using the cleaning process described in Definition 2.1, each edge of G is traversed exactly once and by exactly one brush. Note that no brush may return to a vertex it has already visited, motivating the following definition.

Definition 4.2 *The **brush path** of a brush b is the path formed by the set of edges cleaned by b.*

By definition, G can be decomposed into $b(G)$ brush paths. (Since no brush can stay at its initial vertex in the minimal brush configuration, these paths each have at least one edge.) Thus, the minimum number of paths into which a graph G can be decomposed yields a lower bound for $b(G)$. This is only a lower bound because some path decompositions would not be valid in the cleaning process. For example, K_4 can be decomposed into two edge-disjoint paths, but $b(K_4) = 4$.

Following Definitions 2.1 and 4.2, every vertex of odd degree in a graph G will be the endpoint of (at least) one brush path. This leads to a natural lower bound for $b(G)$ since a graph with d_o odd vertices cannot be decomposed into less than $d_o/2$ paths (see [5] for more details). Note also that this lower bound is sharp since $b(T) = \frac{d_o(T)}{2}$ for any tree T:

Theorem 4.3 ([5]) *For any tree T with $d_o(T)$ vertices of odd degree, $b(T) = \frac{d_o(T)}{2}$.*

There are some unicyclic graphs that need a little bit more than $d_o/2$ brushes but one can show that two additional brushes suffice to clean any unicyclic graph.

Theorem 4.4 *For any unicyclic graph G with $d_o(G)$ vertices of odd degree and $d_t(G)$ vertices on the cycle of degree at least three,*

$$b(G) = \begin{cases} d_o(G)/2, & \text{if } d_t(G) \geq 2; \\ d_o(G)/2 + 1, & \text{if } d_t(G) = 1; \\ d_o(G)/2 + 2 = 2, & \text{if } d_t(G) = 0. \end{cases}$$

PROOF: Take any unicyclic graph G with cycle $C = (v_0, v_1, \ldots, v_{l-1} = v_0)$ of length l. First, let us outline a general strategy for cleaning a vertex on the cycle before moving into three possible cases. Consider vertex v_i $(0 \leq i \leq l-1)$ of degree $k+2$ $(k \geq 0)$ and suppose that $v_{(i-1) \bmod l}$ is already cleaned (the same argument will hold if $v_{(i+1) \bmod l}$ is assumed to be cleaned) – see Figure 2. Let $T_1, T_2, \ldots, T_k$

denote subtrees attached to vertex v_i. We clean independently $k_0 = \lfloor (k+1)/2 \rfloor$ trees $T_1, T_2, \ldots, T_{k_0}$ such that each process finishes with one brush in v_i (note that this is possible by the Reversibility Theorem). After this operation we have enough brushes to clean vertex v_i leaving one brush in v_i if $\deg(v_i)$ is odd; no brush gets stuck in v_i otherwise (again, by the Reversibility Theorem, we can clean each of the remaining trees by cleaning vertex v_i first).

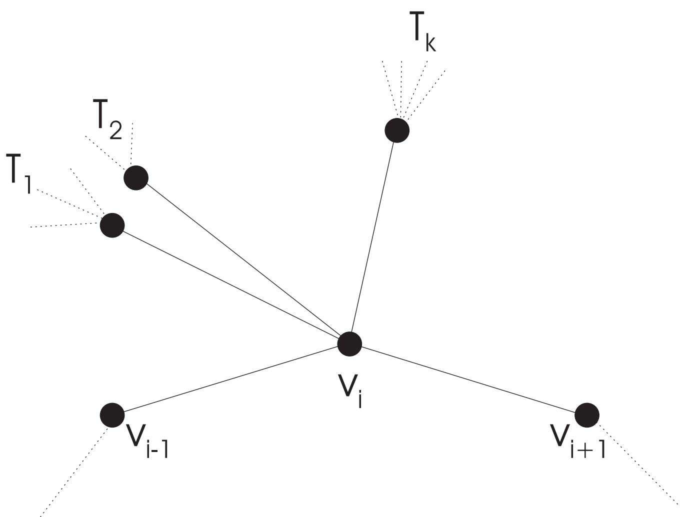


Figure 2: Cleaning unicyclic graphs.

Suppose first that $d_t(G) \geq 2$; vertices v_i and v_j $(i < j)$ have degrees $k_i + 2$ and $k_j + 2$, respectively, for some nonzero values of k_i, k_j. We start the cleaning process by cleaning $\lfloor (k_i + 2)/2 \rfloor$ trees adjacent to v_i leaving $\lfloor (k_i + 2)/2 \rfloor$ brushes in v_i. If $\deg(v_i)$ is even, then v_i is ready to be cleaned and no brush gets stuck in v_i. If $\deg(v_i)$ is odd, then we need to introduce one brush in the initial configuration (that is, $\omega_0(v_i) = 1$) to be able to clean v_i. Next we clean vertices $v_{i+1}, v_{i+2}, \ldots v_{j-1}$ and then vertices $v_{i-1}, \ldots v_1, v_0 = v_{l-1}, c_{l-2}, \ldots, v_{j+1}$ along the cycle. For that we use the strategy mentioned earlier. Finally, we clean $\lceil (k_j - 2)/2 \rceil$ trees adjacent to v_j, v_j itself, and the rest of the graph. Since every vertex of odd degree starts or finishes exactly one brush path, and no vertex of even degree is the endpoint of a brush path, the trivial lower bound of $d_o(G)/2$ for the brush number is obtained.

Suppose now that $d_t(G) = 1$; vertex v_i is the only vertex of degree $k_i + 2$ for some nonzero value of k_i. We can clean graph G as before but in this case, two brushes get stuck in vertex of degree 2 in the cycle C. Thus, $b(G) \leq d_o(G)/2 + 1$. On the other hand, at least one vertex in the cycle must start or finish at least two brush paths so $b(G) \geq d_o(G)/2 + 1$ and the assertion follows.

Finally, if $d_t(G) = 0$, G is a cycle and exactly two brushes are needed to clean G.

This completes the proof of the theorem. □

Now, we are ready to come back to random graphs.

Theorem 4.5 *For $G \in \mathcal{G}(n,p)$, $p = p(n) = d/n$ and $1 - d \gg n^{-1/3}$ (thus, $d \leq 1 - o(1)$)*

$$b(G) = \frac{d_o(G)}{2} + 2X + Y$$

where $X \xrightarrow{D} Po(\lambda_X)$, $Y \xrightarrow{D} Po(\lambda_Y)$ for

$$\lambda_X = -\frac{1}{2}\log(1 - d\exp(-d)) - \frac{(d\exp(-d))^2}{4} - \frac{d\exp(-d)}{2},$$

and

$$\lambda_Y = \frac{1 - \exp(-d)}{2(1 - d\exp(-d))} d^3 \exp(-2d).$$

In particular, a.a.s.

$$b(G) = \frac{d_o(G)}{2} + O(\log\log n) = n\frac{1 - e^{-2d}}{4}(1 + o(1)),$$

and $b(G) = d_o(G)/2$ with probability tending to $\exp(-\lambda_X - \lambda_Y)$ as $n \to \infty$.

Proof: Recall that in the subcritical phase all components are either trees or unicyclic graphs. Thus, by Theorem 4.4 we get that

$$b(G) = \frac{d_o(G)}{2} + 2X(G) + Y(G)$$

where $X(G)$ denotes the number of components that are cycles in G and $Y(G)$ denotes the number of components that are unicyclic graphs with exactly one tree attached to the cycle. Note that $d_o(G)$, $X(G)$, and $Y(G)$ are random variables and we can investigate an asymptotic distribution of $X(G)$ and $Y(G)$.

We know that there is no component of size more than $\log^2 n$ a.a.s. It is also not difficult to see that

$$\begin{aligned}
\mathbb{E}X &= \sum_{k=3}^{\lfloor \log^2 n \rfloor} \binom{n}{k} \frac{(k-1)!}{2} p^k (1-p)^{k(n-3)} \\
&\sim \sum_{k=3}^{\lfloor \log^2 n \rfloor} \frac{(d\exp(-d))^k}{2k} \\
&= \frac{1}{2} \sum_{k=1}^{\lfloor \log^2 n \rfloor} \frac{(d\exp(-d))^k}{k} - \frac{(d\exp(-d))^2}{4} - \frac{d\exp(-d)}{2} \\
&\sim -\frac{1}{2}\log(1 - d\exp(-d)) - \frac{(d\exp(-d))^2}{4} - \frac{d\exp(-d)}{2} \\
&=: \lambda_X
\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}Y &= \sum_{k=3}^{\lfloor \log^2 n \rfloor} \binom{n}{k} \frac{(k-1)!}{2} p^k (1-p)^{(k-1)(n-3)} k(1-\exp(-d)) \\ &\sim \frac{1-\exp(-d)}{2\exp(-d)} \sum_{k=3}^{\lfloor \log^2 n \rfloor} (d\exp(-d))^k \\ &\sim \frac{1-\exp(-d)}{2\exp(-d)} \frac{(d\exp(-d))^3}{1-d\exp(-d)} \\ &=: \lambda_Y .\end{aligned}$$

One can also check that, for a given $r \geq 2$, the rth factorial moments of $X(G)$ and $Y(G)$ tend to λ_X^r and λ_Y^r, respectively. Thus, both variables $X(G), Y(G)$ tend to a Poisson distribution with parameters λ_X, λ_Y, respectively. Moreover, these two random variables are asymptotically independent.

It is also straightforward to see that a.a.s. $X(G)+Y(G)$ is at most, say, $\log\log n$. Hence, in order to finish the proof, we have to estimate the number of odd vertices. For any $k = k(n) < \log n$, let p_k denote the probability that a given vertex has degree k. (It is known that there are no vertices of degree more than $\log n$ a.a.s.) It is clear that

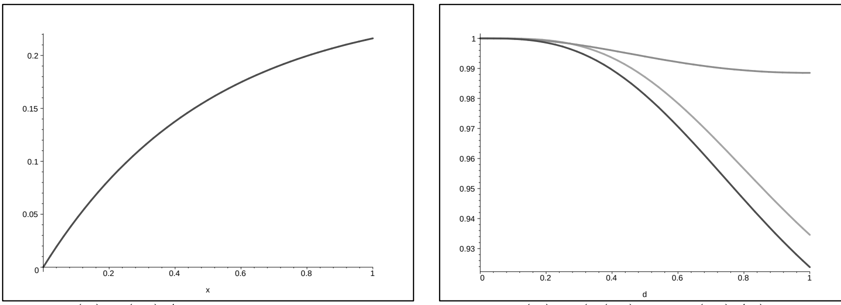
$$p_k = \binom{n-1}{k} p^k (1-p)^{n-1-k} = \frac{d^k}{k!} e^{-d} (1+o(1)) . \tag{4}$$

Thus, the expected number of odd vertices is equal to

$$\begin{aligned}\mathbb{E}d_o(G) &= n \sum_{1 \leq k < \log n, k \text{ is odd}} p_k \\ &= (1+o(1)) n e^{-d} \sum_{k, k \text{ is odd}} \frac{d^k}{k!} \\ &= (1+o(1)) n e^{-d} \frac{e^d - e^{-d}}{2} \\ &= (1+o(1)) n \frac{1-e^{-2d}}{2} .\end{aligned}$$

It is also not difficult to show that $d_o(G)$ is well concentrated around its expectation which finishes the proof of the theorem. □

As we already mentioned, if $p < (1+o(1))/n$, then the size of the giant component as well as the number of edges in this component are $o(n)$. So we can extend the previous theorem to the wider range of p (since adding one edge to the graph can increase the brush number by at most one) containing the critical phase and the 'early' supercritical phase, namely, it can be shown that $b(G) = (1+o(1))n(1-e^{-2d})/4$ for $p < (1+o(1))/n$. In Figure 3(a), the values of $b(G)/n$ have been presented for d-values between 0 and 1.



(a) $b(G)/n$ for $0 \leq d \leq 1$ (b) $\mathbb{P}(b(G) = d_o(G)/2)$

Figure 3: Behaviour of the brush number in the subcritical phase.

Finally, in Figure 3(b), we present three graphs: the first one (from the top) is the probability that there is no cycle in $G(n, d/p)$; the second one shows the probability that there is no unicyclic graph with exactly one tree; the third one expresses the probability that the brush number attains its trivial lower bound, that is, $b(G) = d_o(G)/2$.

4.2 The supercritical phase

In this subsection, we consider a random graph $G(n, p = d/n)$ for $d > 1$. For this range of the parameter p, we can determine the approximate size of the giant component as well as the structure of the graph formed by deleting it. It is known that the size of the giant component is a.a.s. $c_d n + o(n)$ where c_d is the unique solution to

$$c + e^{-dc} = 1\,, \tag{5}$$

and the graph formed by deleting the giant component is essentially equivalent to $G(n', p')$, where $n' = (1 - c_d + o(1))n$ and $p' = d/n = (d(1 - c_d + o(1))/n'$. (Note that $d(1 - c_d) < 1$ so, indeed, a.a.s. all other components are of order $O(\log n)$.) In Figure 4(a), the numerical values of c_d have been presented for d-values between 1 and 5.

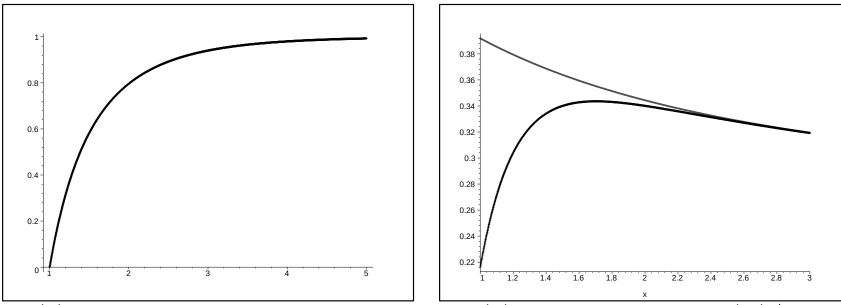
The knowledge about the size of the giant component helps us to get a stronger upper bound (numerical one) but we present also an exact formula which works well for relatively large values of d.

Theorem 4.6 *For $G \in \mathcal{G}(n, p)$, $p = p(n) = d/n$ and $d > 1$ a.a.s.*

$$b(G) \leq \frac{n}{4}\left(d + 1 - \frac{1 - e^{-2d}}{2d}\right)(1 + o(1))$$

and

$$b(G) \quad \leq \quad \frac{n}{4}\left(1 + dc_d(2 - c_d) - e^{-d(2-c_d)} + \frac{e^{-2d} - e^{-2d(1-c_d)}}{2d}\right)(1 + o(1))\,.$$



(a) Size of the giant component (b) Upper bounds on $b(G)/dn$

Figure 4: Behaviour of the brush number in the supercritical phase.

We give a graph of both upper bounds of $b(G)/dn$ in Figure 4(b).

PROOF: Using (2), we can show that the total number of edges is well concentrated around $dn/2$. Thus, from Theorem 3.2 (see also (4)) it follows that

$$\begin{aligned} b(G) &\leq \frac{|E|}{2} + \frac{|V|}{4} - \frac{1}{4} \sum_{v \in V(G), \deg(v) \text{ is even}} \frac{1}{\deg(v)+1} \\ &= \frac{n}{4} \left(d + 1 - \sum_{k, k \text{ is even}} \frac{1}{k+1} \cdot \frac{d^k}{k!} e^{-d} \right) (1+o(1)) \\ &= \frac{n}{4} \left(d + 1 - \frac{e^{-d}}{d} \sum_{k, k \text{ is odd}} \frac{d^k}{k!} \right) (1+o(1)) \\ &= \frac{n}{4} \left(d + 1 - \frac{1-e^{-2d}}{2d} \right) (1+o(1)). \end{aligned}$$

In order to get a better (but numerical) upper bound we can use Theorem 3.2 one more time to get a bound on the number of brushes needed to clean the giant component and give an asymptotically almost sure value for small components.

Similar calculations to ones we had before (see (4)) can be used to show that the probability a given vertex outside the giant component has degree k is equal to

$$p'_k = \frac{(d(1-c_d))^k}{k!} e^{-d(1-c_d)} (1+o(1)). \qquad (6)$$

Thus, the total number of brushes needed to clean all small components is equal to

$$\begin{aligned}
&(1+o(1))\frac{1}{2}(1-c_d)n \sum_{1\le k<\log n, k \text{ is odd}} p'_k \\
&\quad = (1+o(1))\frac{1}{2}e^{-dc_d} n e^{-d(1-c_d)} \sum_{k, k \text{ is odd}} \frac{(d(1-c_d))^k}{k!} \\
&\quad = (1+o(1))\frac{1}{2} n e^{-d} \frac{e^{d(1-c_d)} - e^{-d(1-c_d)}}{2} \\
&\quad = (1+o(1))\frac{n}{4}\left(1 - c_d - e^{-d(2-c_d)}\right). \qquad (7)
\end{aligned}$$

Since the number of edges in the giant component is equal to

$$(1+o(1))\left(\binom{n}{2}p - \binom{n'}{2}p'\right) = (1+o(1))\frac{dn}{2}c_d(2-c_d),$$

we get an upper bound on the number of brushes needed to clean the giant component G' of G, namely,

$$\begin{aligned}
b(G') &\le \frac{|E(G')|}{2} + \frac{|V(G')|}{4} - \frac{1}{4} \sum_{v \in V(G'), \deg(v) \text{ is even}} \frac{1}{\deg(v)+1} \\
&= \frac{n}{4}\Bigg(dc_d(2-c_d) + c_d - \sum_{k, k \text{ is even}} \frac{1}{k+1} \\
&\qquad \cdot \left(\frac{d^k}{k!} e^{-d} - (1-c_d)\frac{(d(1-c_d))^k}{k!} e^{-d(1-c_d)} \right) \Bigg)(1+o(1)) \\
&= \frac{n}{4}\left(dc_d(2-c_d) + c_d + \frac{e^{-2d} - e^{-2d(1-c_d)}}{2d} \right)(1+o(1)). \qquad (8)
\end{aligned}$$

In order to get a new upper bound and finish the proof, it is enough to add together (7) and (8). □

Now, let us move to a lower bound.

Theorem 4.7 *For $G \in \mathcal{G}(n,p)$, $p = p(n) = d/n$ and $d > 1$ a.a.s. the following lower bounds hold:*

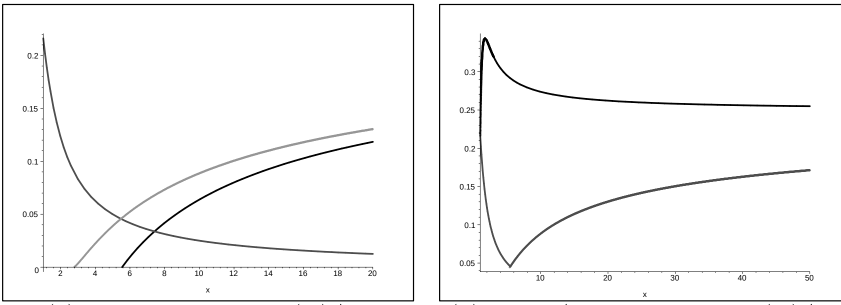
$$b(G) \ge \frac{d_o(G)}{2} = n\frac{1-e^{-2d}}{4}(1+o(1)), \qquad (9)$$

$$b(G) \ge \frac{n}{4}\left(d - \sqrt{8d\log 2}\right), \qquad (10)$$

$$b(G) \ge \frac{dn}{4}(d - t_d), \qquad (11)$$

where t_d is the unique solution to

$$(d-t)\log(1-t/d) + t = 4\log 2.$$



(a) Lower bounds on $b(G)/dn$ (b) Lower/upper bound on $b(G)/dn$

Figure 5: Behaviour of the brush number in the supercritical phase.

Before we prove the theorem, we provide a graph of all three normalized lower bounds (see Figure 5(a)) as well as a comparison of normalized lower/upper bound on the brush number (see Figure 5(b)).

PROOF: The first lower bound (9) is obvious and the formula has been already derived – see proof of the Theorem 4.5.

In order to show that (10) and (11) hold, we use the same argument as in the Subsection 3.2. The expected number of edges $\mathbb{E}Y$ coming from any set of $n/2$ vertices to its complement is $dn/4$. Moreover,

$$\mathbb{P}(Y \leq \mathbb{E}Y - t) \leq \exp\left(-\frac{t^2}{2\mathbb{E}Y}\right),$$

for $t \geq 0$ (see Theorem 2.1 in [3]). Putting $t_0 = n\sqrt{\frac{d \log 2}{2}}$ we get that $\mathbb{P}(Y \leq \mathbb{E}Y - t_0) \leq 2^{-n}$ and the expected number of sets of cardinality $n/2$ with less than $\mathbb{E}Y - t_0$ edges to its complement is tending to zero. Now, (10) holds by the Markov's inequality.

Finally, in order to get a better (numerical) lower bound of (11) one can use the whole power of [3, Theorem 2.1], namely

$$\mathbb{P}(Y \leq \mathbb{E}Y - t) \leq \exp\left(-\mathbb{E}Y \varphi(-t/\mathbb{E}Y)\right),$$

where $\varphi(x) = (1+x)\log(1+x) - x$, $x \geq -1$. □

Acknowledgment

The computations presented in the paper were performed by using Maple$^{\text{TM}}$ [7]. The worksheet can be found at the address: `http://www.mathstat.dal.ca/~pralat/`.

References

[1] N. Alon, P. Prałat and N. Wormald, Cleaning d-regular graphs with brushes, *SIAM J. Discrete Math.* (accepted).

[2] N. Alon and J.H. Spencer, *The Probabilistic Method*, Wiley, 1992 (Second Edition, 2000).

[3] S. Janson, T. Łuczak and A. Ruciński, *Random Graphs*, Wiley, New York, 2000.

[4] S. McKeil, Chip Firing Cleaning Processes, *MSc Thesis*, Dalhousie University, 2007.

[5] M.E. Messinger, R.J. Nowakowski and P. Prałat, Cleaning a Network with Brushes, *Theoretical Computer Science* **399** (2008), 191–205.

[6] M.E. Messinger, R.J. Nowakowski, P. Prałat and N. Wormald, Cleaning random d-regular graphs with brushes using a degree-greedy algorithm, *Proc. 4th Workshop Combinatorial and Algorithmic Aspects of Networking (CAAN2007)*, Lec. Notes Comp. Sci., Springer, 2007, 13–26.

[7] M. B. Monagan, K. O. Geddes, K. M. Heal, G. Labahn, S. M. Vorkoetter, J. McCarron and P. DeMarco, Maple 10 Programming Guide, Maplesoft, Waterloo ON, Canada, 2005.

(Received 27 Jan 2008; revised 26 July 2008)