ON THE PROBABILITY THAT A RANDOM SUBGRAPH CONTAINS A CIRCUIT

PETER NELSON

ABSTRACT. Let $\mu > 2$ and $\varepsilon > 0$. We show that, if G is a sufficiently large simple graph of average degree at least μ, and H is a random spanning subgraph of G formed by including each edge independently with probability $p \geq \frac{1}{\mu-1} + \varepsilon$, then H contains a cycle with probability at least $1 - \varepsilon$.

1. INTRODUCTION

We prove the following theorem:

Theorem 1.1. *For all $\varepsilon > 0$ and $\mu > 2$, there exists $N \in \mathbb{Z}$ so that, if G is a graph on at least N vertices with average degree at least μ, and H is a random spanning subgraph of G formed by including each edge of G independently with probability $p \geq \frac{1}{\mu-1} + \varepsilon$, then H contains a circuit with probability at least $1 - \varepsilon$.*

The same result was shown by Alon and Bachmat [1] in the special case where G is regular. They also show that, for any $\varepsilon > 0$ and integer $d \geq 3$, there is a d-regular graph G such that a random spanning subgraph of G in which edges are chosen with probability $\frac{1}{d-1} - \varepsilon$ contains a circuit with probability at most ε; in other words, $\frac{1}{\mu-1}$ cannot be replaced by any smaller value in the above theorem if $\mu \in \mathbb{Z}$. On the other hand, it can be shown that if $\mu \notin \mathbb{Z}$, then the value $\frac{1}{\mu-1}$ *can* be lowered; we discuss this in the next section.

Our proof is based on that of the main theorem of [12], which concerns the related question of the probability that H contains at least half of the edges of some circuit of G. The aim of this paper is to present the relatively simple proof of Theorem 1.1 without the inherent and numerous technicalities of [12].

Similar questions concerning cycles in random subgraphs of graphs with a given *minimum* degree were considered in [6] and [9].

Date: May 26, 2015.

2. Preliminaries

All graphs are simple and finite unless otherwise stated. We use some standard graph theory terminology such as *path*, *walk*, *girth* and *adjacency matrix* (for a directed graph); see [4] for a reference. For $p \in [0,1]$, a *p-random* subset of a set E refers to a set $X \subseteq E$ obtained by including each element of E independently at random with probability p. We also use some basic probability theory; for a reference, see [5].

Zero-one Laws. Let E be a finite set. For each $\mathcal{W} \subseteq 2^E$, let $f_{\mathcal{W}} \colon [0,1] \to [0,1]$ be defined by $f_{\mathcal{W}}(p) = \sum_{X \in \mathcal{W}} p^{|X|}(1-p)^{|E|-|X|}$; i.e. $f_{\mathcal{W}}(p)$ is the probability that a p-random subset of E is in $\mathcal{W}$. An important tool in our proof is a theorem of Margulis that gives a very general sufficient condition for $f_{\mathcal{W}}(p)$ to display a zero-one-law type behaviour. In [11], Margulis states his theorem in very high generality (and in Russian). Our statement follows a convenient 'discrete' formulation found in ([13], Section 2).

We say that $\mathcal{W} \subseteq 2^E$ is *increasing* if, whenever $X \in \mathcal{W}$ and $X \subseteq X'$, we have $X' \in \mathcal{W}$. Given $X \in 2^E$, let $s(X)$ denote the collection of subsets of E that differ from X by a single addition or removal (i.e. the Hamming sphere of radius 1 around X in 2^E). Let $\Delta(\mathcal{W})$ denote the minimum nonzero value of $|s(X) \setminus \mathcal{W}|$ over all $W \in \mathcal{W}$.

Note that, if $\varnothing \neq \mathcal{W} \neq 2^E$ and $\mathcal{W}$ is increasing, the function $f_{\mathcal{W}}(p)$ is monotonely increasing with $f_{\mathcal{W}}(0) = 0$ and $f_{\mathcal{W}}(1) = 1$. Margulis' theorem states that if $\Delta(\mathcal{W})$ is sufficiently large, then the value of $f_{\mathcal{W}}(p)$ is nearly always close to zero or one.

Theorem 2.1. *For all $\varepsilon > 0$ there exists $s \in \mathbb{Z}$ so that, if E is a finite set, and $\mathcal{W} \subseteq 2^E$ is increasing and satisfies $\Delta(W) \geq s$, then the interval $\{p \in [0,1] \colon \varepsilon \leq f_{\mathcal{W}}(p) \leq 1-\varepsilon\}$ has length less than ε.*

We will apply this result in the very special case where E is the edge set of a graph G, and $\mathcal{W}$ is the collection of edge-sets of subgraphs of G that contain a circuit. In this setting, it is easy to see that the parameter $\Delta(\mathcal{W})$ is exactly the girth of G.

Non-backtracking walks. A *non-backtracking walk* of length ℓ in a graph G is a walk $(v_0, v_1, \ldots, v_\ell)$ of G so that $v_{i+1} \neq v_{i-1}$ for all $i \in \{1, \ldots, \ell-1\}$. In all nontrivial cases, the number of such walks grows roughly exponentially in ℓ; in this section we state a result of Alon et al. that estimates the base of this exponent. Let $G = (V, E)$ be a connected graph of minimum degree at least 2. Let $\overline{E} = \{(u,v) \in V^2 : u \sim_G v\}$ be the $2|E|$-element set of arcs of G. Let $B = B(G) \in \{0,1\}^{\overline{E} \times \overline{E}}$ be

the matrix so that $B_{(u,v),(u',v')} = 1$ if and only if $u' = v$ and $u \neq v'$. It is easy to see that

(1) B is the adjacency matrix of a strongly connected digraph (essentially the 'line digraph' of G), and
(2) For each integer $\ell \geq 2$, the entry $(B^{\ell-1})_{e,f}$ is the number of non-backtracking walks of length ℓ in G with first arc $(v_0, v_1) = e$ and last arc $(v_\ell, v_{\ell+1}) = f$.

By (1) and the Perron-Frobenius theorem (see [7], section 8.8), there is a positive real eigenvalue λ_* of B and an associated positive real eigenvector w_*, so that $|\lambda_*| \geq |\lambda|$ for every eigenvalue λ of B. Furthermore, by Gelfand's formula [8] we have $\lambda_* = \lim_{n\to\infty} \|B^n\|^{1/n}$, where $\|B^n\|$ denotes the sum of the absolute values of the entries of B^n. By (2), the parameter $\lambda_* = \lambda_*(B(G))$ thus governs the growth of non-backtracking walks in G. It is clear that if G is d-regular we have $\lambda_*(B(G)) = d-1$; the following result of Alon et al. [2] shows that the average degree gives a similar lower bound for general graphs:

Lemma 2.2. *Let $\mu \geq 2$. If G is a connected graph of average degree at least μ and minimum degree at least 2, then $\lambda_*(B(G)) \geq \mu - 1$.*

In fact, the argument in [2] shows that $\lambda_*(B(G)) \geq \Lambda(G)$, where $\Lambda(G)$ is a certain symmetric function in the degree sequence of G that is bounded below by $\mu - 1$. When $\mu \notin \mathbb{Z}$, the inequality $\Lambda(G) \geq \mu - 1$ cannot hold with equality; in fact one can show (see [12], Lemma 3.2) that $\lambda_*(B(G)) \geq \mu - 1 + \frac{\eta(\mu)^3}{8\mu^3}$, where $\eta(\mu)$ denotes the distance from μ to the nearest integer. This can easily be shown to lead to an improved version of Theorem 1.1 where $\frac{1}{\mu-1}$ is replaced by a strictly smaller value for nonintegral μ.

3. Covering Trees

Given a connected graph $G = (V, E)$ of minimum degree at least 2, we denote the set of arcs of G, as before, by $\bar{E}$. Given an arc $e_0 = (u_0, u_1) \in \bar{E}$, the *covering tree of G at e_0*, for which we write $\Gamma_{e_0}(G)$, is the infinite rooted tree Γ whose root is the length-zero walk (u_0), the other vertices are the non-backtracking walks of G whose first arc is e_0, and the children of each walk $(u_0, u_1, \ldots, u_k) \in V(\Gamma)$ are exactly its extensions by a single arc: that is, the nonbacktracking walks of the form $(u_0, u_1, \ldots, u_k, u_{k+1})$. Note that the unique child of the root is the walk $(e_0) = (u_0, u_1)$.

Given such a Γ with root r, we borrow some terminology from [10]. For $x \in V(\Gamma)$, we write $|x|$ for the distance from x to r in Γ. For $x, y \in V(\Gamma)$, we write $x \preceq y$ if x is on the path from r to y, and $x \wedge y$

for the *join* of x and y in Γ: that is, the unique vertex z on both the path from r to x and the path from r to y for which $|z|$ is maximized.

The map $\theta\colon V(\Gamma) \to V(G)$ that assigns each walk to its final vertex is a graph homomorphism that is injective when restricted to the neighbourhood of any vertex. To analyse the probability that a p-random subset of $E(G)$ contains a circuit, we consider the probability that a p-random subset of $E(\Gamma)$ contains a long path containing the root. The following lemma is a stronger 'constructive' version of Theorem 6.2 of [10] (which applies to general infinite rooted trees) in the case where Γ is the covering tree of a finite graph.

Lemma 3.1. *Let G be a connected graph of minimum degree at least 2 and let $\lambda = \lambda_*(B(G))$. There is an arc e_0 of G so that, if $p \in [0,1]$ and X is a p-random subset of $E(\Gamma_{e_0}(G))$, then, for each integer $n \geq 1$, the probability that X contains a n-edge path of $\Gamma_{e_0}(G)$ containing the root is at least $p - \frac{1}{\lambda}$.*

Proof. Let $B = B(G)$, and let $w \in \mathbb{R}^{\overline{E}(G)}$ be the (positive, real) eigenvector of $B(G)$ corresponding to λ whose largest entry is 1. Let $e_0 \in \overline{E}(G)$ be such that $w(e_0) = 1$; we show that e_0 satisfies the lemma. We may assume that $p > \frac{1}{\lambda}$. Let $\Gamma = \Gamma_{e_0}(G)$, let r be the root of G, and let $\pi\colon V(\Gamma) \setminus \{r\} \to \overline{E}(G)$ be the map associating each walk with its last arc.

Let $\phi\colon V(\Gamma) \to \mathbb{R}_{>0}$ be defined by $\phi(r) = 1$ and $\phi(v) = \lambda^{1-|v|} w(\pi(v))$ for all $v \neq r$. Note that $\phi((e_0)) = w(e_0) = 1$ and that, since λ is an eigenvalue of B, the sum of $\phi(x)$ over the children x of a vertex v (or over the descendants x of v at any fixed level) is equal to $\phi(v)$. In other words, ϕ is a *unit flow* of Γ.

For $X \subseteq E(\Gamma)$, let R_X denote the vertex set of the component of $\Gamma[X]$ containing r. Define a random variable $Q = Q(X)$ by

$$Q = p^{-n} \sum_{|x|=n} \phi(x) 1_{R_X}(x).$$

Since $\sum_{|x|=n} \phi(x) = 1$ and each vertex at distance n from the root is in R_X with probability exactly p^n, we have $\mathbf{E}(Q) = 1$. We now bound the variance of Q.

Claim 3.1.1. $\mathbf{E}(Q^2) \leq (p - \frac{1}{\lambda})^{-1}$.

Proof. Note that $|x \wedge y| \geq 1$ whenever $|x|, |y| \neq 0$. We have

$$\mathbf{E}(Q^2) = p^{-2n} \sum_{|x|=|y|=n} \phi(x)\phi(y)\mathbf{P}(x, y \in R_X)$$

$$\begin{aligned}
&= p^{-2n} \sum_{|x|=|y|=n} \phi(x)\phi(y)p^{2n-|x\wedge y|} \\
&= \sum_{|x|=|y|=n} \phi(x)\phi(y)p^{-|x\wedge y|} \\
&= \sum_{1\le|z|\le n} p^{-|z|} \sum_{\substack{|x|=|y|=n \\ x\wedge y=z}} \phi(x)\phi(y) \\
&\le \sum_{1\le|z|\le n} p^{-|z|} \left(\sum_{\substack{|x|=n \\ x\succ z}} \phi(x) \right)^2 \\
&= \sum_{1\le|z|\le n} p^{-|z|}\phi(z)^2 \\
&= \sum_{i=1}^{n} p^{-i} \sum_{|z|=i} \phi(z)^2
\end{aligned}$$

Now by definition of ϕ and the fact that $w(e) \le 1$ for all e we have

$$\phi(z)^2 = \lambda^{2-2|z|} w(\pi(z))^2 \le \lambda^{2-2|z|} w(\pi(z)).$$

For each $e \in \overline{E}$, let $b_e \in \mathbb{R}^{\overline{E}}$ be the corresponding standard basis vector. For each $1 \le i \le n$, the number of $z \in V(\Gamma)$ with $|z| = i$ and $\pi(z) = e$ is exactly $b_{e_0}^T B^{i-1} b_e$, so it follows from $Bw = \lambda w$ and $b_{e_0}^T w = 1$ that

$$\sum_{|z|=i} \phi(z)^2 \le \lambda^{2-2i} b_{e_0}^T B^{i-1} \sum_{e\in\overline{E}} b_e w(e) = \lambda^{2-2i} b_{e_0}^T B^{i-1} w = \lambda^{1-i}.$$

Therefore $\mathbf{E}(Q^2) \le \sum_{i=1}^{n} p^{-i}\lambda^{1-i} < (p - \frac{1}{\lambda})^{-1}$, since $p\lambda > 1$. □

The Cauchy-Schwartz inequality now gives

$$1 = \mathbf{E}(Q)^2 = \mathbf{E}(Q \cdot 1_{Q>0})^2 \le \mathbf{E}(Q^2)\mathbf{P}(Q > 0) \le (p - \tfrac{1}{\lambda})^{-1}\mathbf{P}(Q > 0),$$

so $\mathbf{P}(Q > 0) \ge p - \frac{1}{\lambda}$. If $Q > 0$ then X contains an n-edge path of Γ containing the root; the lemma follows. □

4. Circuits

In this section, we prove our main theorem. First we show that, if G is 'non-degenerate' and $\lambda = \lambda_*(B(G))$, then a $(\frac{1}{\lambda} + \varepsilon)$-random subset of $E(G)$ contains a circuit with non-negligible probability.

Theorem 4.1. *Let G be a connected graph with minimum degree at least 2 and let $\lambda = \lambda_*(B(G))$. Let $\varepsilon > 0$ and $p \in [\frac{1}{\lambda} + \varepsilon, 1]$. If X is a p-random subset of $E(G)$, then X contains a circuit of G with probability at least $\frac{\varepsilon^2}{4}$.*

Proof. Let $p_1 = p - \frac{\varepsilon}{2} \geq \lambda^{-1} + \frac{\varepsilon}{2}$, and let $p_2 \geq \frac{\varepsilon}{2}$ be such that $1 - p = (1 - p_1)(1 - p_2)$. Let X_1 be a p_1-random subset of $E(G)$ and X_2 be a p_2-random subset of $E(G)$ independent of X_1; note that $X_1 \cup X_2$ is identically distributed to a p-random subset of $E(G)$.

Let $e_0 = (u, v)$ be an arc of G given by Lemma 3.1 for p_1 and λ and let $\Gamma = \Gamma_e(G)$. Let $\theta \colon V(\Gamma) \to V(G)$ denote the natural homomorphism from Γ to G. For each set $U \subseteq V(G)$, let $G(U)$ denote the subgraph of G induced by U.

Given X_1, let $R(X_1)$ denote the set of vertices x of G for which there is a non-backtracking walk of $G[X_1]$ from u to x that either contains no arc (that is, $x = u$), or has first arc (u, v). Observe that either $R(X_1) = \{u\}$, or $R(X_1)$ is the vertex set of a connected subgraph of G containing the edge uv. Similarly, for a p_1-random subset Y_1 of $E(\Gamma)$, let $R(Y_1)$ denote the vertex set of the component of $\Gamma[Y_1]$ containing the root. Observe that $G(R(X_1))$ and $G(\theta(R(Y_1)))$ are both connected subgraphs of G containing u (in general these graphs have edges not in X_1 or $\theta(Y_1)$). Let C_G denote the event that $G(R(X_1))$ contains a circuit, and C_Γ the event that $G(\theta(R(Y_1)))$ contains a circuit.

Claim 4.1.1. $\mathbf{P}(C_G) = \mathbf{P}(C_\Gamma)$.

Proof. Let $\mathcal{Z}'$ denote the collection of subsets of $V(G)$ that induce an acyclic connected subgraph of G containing u and v, and let $\mathcal{Z} = \mathcal{Z}' \cup \{\{u\}\}$. The event C_G fails to hold exactly when $R(X_1) \in \mathcal{Z}$, so

$$1 - \mathbf{P}(C_G) = \sum_{Z \in \mathcal{Z}} \mathbf{P}(R(X_1) = Z).$$

Similarly, we have

$$1 - \mathbf{P}(C_\Gamma) = \sum_{Z \in \mathcal{Z}} \mathbf{P}(\theta(R(Y_1)) = Z).$$

Clearly $\mathbf{P}(R(X_1) = \{u\}) = \mathbf{P}(\theta(R(Y_1)) = \{u\}) = 1 - p$. Let $Z \in \mathcal{Z}'$. By acyclicity of $G(Z)$, there is a unique subtree Γ_Z of Γ that contains the root of Γ and satisfies $\theta(V(\Gamma_Z)) = Z$, and moreover $G(Z)$ and Γ_Z are isomorphic finite trees. Now $|E(G(Z))| = |E(\Gamma_Z)|$, and the number of edges of G with exactly one end in $Z \setminus \{u\}$ is equal to the number of edges of Γ with exactly one end in $V(\Gamma_Z)$, so

$$\mathbf{P}(R(X_1) = Z) = \mathbf{P}(R(Y) = V(\Gamma_Z)) = \mathbf{P}(\theta(R(Y)) = Z).$$

The claim now follows from the above two summations. □

If Y_1 contains a $|V(G)|$-edge path of Γ that contains the root, then this path is mapped by θ to a non-backtracking walk of G visiting some vertex twice, so $G(\theta(R(Y_1)))$ contains a circuit of G. Thus, by the claim above and Lemma 3.1, we have $\mathbf{P}(C_G) = \mathbf{P}(C_\Gamma) \geq p_1 - \lambda \geq \frac{\varepsilon}{2}$.

Suppose that C_G holds; then $H = G(R(X_1))$ is a graph containing a cycle and having a spanning tree contained in X_1. Thus, there is some edge f of H such that $X_1 \cup \{f\}$ contains a circuit of H; such an f exists for every X_1 satisfying C_G. Now $f \in X_2$ with probability p_2, so the probability that $X_1 \cup X_2$ contains a circuit of G is at least $p_2\mathbf{P}(C_G) \geq \left(\frac{\varepsilon}{2}\right)^2 = \frac{\varepsilon^2}{4}$. This gives the result. □

We now restate and prove our main theorem.

Theorem 4.2. *For all $\varepsilon > 0$ and $\mu > 2$, there exists $N = N(\varepsilon, \mu) \in \mathbb{Z}$ so that, if G is a graph on at least N vertices with average degree at least μ, and X is a p-random subset of $E(G)$ for some $p \in [\frac{1}{\mu-1} + \varepsilon, 1]$, then X contains a circuit of G with probability at least $1 - \varepsilon$.*

Proof. Let $\varepsilon > 0$, $\mu > 2$, and $p \in [\frac{1}{\mu-1} + \varepsilon, 1]$. Let $s = s(\frac{\varepsilon^2}{16})$ be given by Theorem 2.1. Let $p_1 = p - \frac{\varepsilon^2}{16}$. Let t be an integer so that $(1-p_1^s)^t \leq \varepsilon$. Let $\mu_1 \in (2, \mu)$ be such that $\frac{1}{\mu_1 - 1} + \frac{\varepsilon}{2} \leq p_1$. Set $N = \left\lceil \frac{2st}{\mu_1 - \mu} \right\rceil$.

Let G be a graph on at least N vertices with average degree at least μ. For each $x \in [0, 1]$, let $f(x)$ denote the probability that an x-random subset of $E(G)$ contains a circuit of G. We will show that $f(p) \geq 1 - \varepsilon$.

Claim 4.2.1. *Either*

- *$f(p_1) \geq 1 - \varepsilon$, or*
- *G has a connected subgraph H with girth at least s, minimum degree at least 2, and average degree at least μ_1.*

Proof of claim: Let $\mathcal{C}$ be a maximal collection of pairwise edge-disjoint circuits of size less than s in G. If $|\mathcal{C}| \geq t$ then $f(p_1) \geq 1 - (1 - {p_1}^s)^t \geq 1 - \varepsilon$. If $|\mathcal{C}| < t$, then let G' be obtained by removing the edges of all circuits in $\mathcal{C}$ from G. By the maximality of $\mathcal{C}$, the graph G' has girth at least s. Now by choice of $N \leq |V(G)|$, we have

$$|E(G')| \geq |E(G)| - st \geq \tfrac{\mu}{2}|V(G)| - \tfrac{\mu - \mu_1}{2}N \geq \tfrac{\mu_1}{2}|V(G)|,$$

so G' has average degree at least μ_1. Now, let G'' be obtained from G' by deleting degree-1 vertices until no more such deletions are possible. It is easy to see (since $\mu_1 \geq 2$) that G'' has minimum degree at least 2 and average degree at least μ_1. Any connected component H of G'' with largest-possible average degree will satisfy the claim. □

Since $f(p) \geq f(p_1)$, we may assume that the above subgraph H exists. Let $\mathcal{W}$ be the collection of edge sets of subgraphs of H that contain a circuit; let $g(x) = \sum_{X \in \mathcal{W}} x^{|X|}(1-x)^{|E(H)|-|X|}$. This is the probability that an x-random subset of $E(H)$ contains a circuit, so we have $f(x) \geq g(x)$. By Lemma 2.2 we have $\lambda_*(B(H)) \geq \mu_1 - 1$, so $p_1 \geq (\lambda_*(B(H)))^{-1} + \frac{\varepsilon}{2}$ and therefore $g(p_1) \geq \frac{\varepsilon^2}{16}$ by Theorem 4.1.

Now $\Delta(\mathcal{W})$ is the girth of H, so $\Delta(\mathcal{W}) \geq s$ and therefore the interval $\{x \in [0,1] : \frac{\varepsilon^2}{16} \leq g(x) \leq 1 - \frac{\varepsilon^2}{16}\}$ has length at most $\frac{\varepsilon^2}{16}$ by Theorem 2.1. Therefore $g(p) = g(p_1 + \frac{\varepsilon^2}{16}) \geq 1 - \frac{\varepsilon^2}{16} \geq 1 - \varepsilon$. Since $f(p) \geq g(p)$, the result follows. □

References

[1] N. Alon and E. Bachmat, Regular graphs whose subgraphs tend to be acyclic, Random Struct. Algo. 29 (2006), 324–337.

[2] N. Alon, S. Hoory and N. Linial, The Moore Bound for Irregular Graphs, Graph Combinator. 18 (2002), 53–57.

[3] L. Decreusefond and G. Zémor, On the error-correcting capabilities of cycle codes of graphs, Combin. Probab. Comput. 6 (1997), 27–38.

[4] R. Diestel, Graph Theory, Springer, 2000.

[5] R. Durrett, Probability: Theory and Examples (4th edition), Cambridge University Press, 2010.

[6] A. Frieze and M. Krivelevich, On the non-planarity of a random subgraph, Combin. Probab. Comput. 22 (2013), 722–732.

[7] C. Godsil and G. Royle, Algebraic Graph Theory, Springer, 2001.

[8] I. Gelfand, Normierte ringe, Rech. Math. [Mat. Sbornik] N.S., 9 (1941), 3–24

[9] M. Krivelevich and W. Samotij, Long paths and cycles in random subgraphs of H-free graphs, Electron. J. Combin. 21 (2014), P1.30.

[10] R. Lyons, Random walks and percolation on trees, Ann. Probab. 18 (1990), 931–958.

[11] G. Margulis, Probabilistic characterisations of graphs with large connectivity, Problemy Peredachi Informatsii 10 (1974), 101–108.

[12] P. Nelson and Stefan H.M. van Zwam, On the maximum-likelihood decoding threshold for graphic codes, In preparation.

[13] G. Zémor, Threshold effects in codes, In *Algebraic Coding: Lecture Notes in Computer Science 781.* Springer-Verlag, 278–286.