COMBINATORICA
Bolyai Society – Springer-Verlag

COMBINATORICA **26** (4) (2006) 475–487

HOMOLOGICAL CONNECTIVITY OF RANDOM 2-COMPLEXES

NATHAN LINIAL*, ROY MESHULAM*

Received May 25, 2003

Let Δ_{n-1} denote the $(n-1)$-dimensional simplex. Let Y be a random 2-dimensional subcomplex of Δ_{n-1} obtained by starting with the full 1-dimensional skeleton of Δ_{n-1} and then adding each $2-$simplex independently with probability p. Let $H_1(Y;\mathbb{F}_2)$ denote the first homology group of Y with mod 2 coefficients. It is shown that for any function $\omega(n)$ that tends to infinity

$$\lim_{n\to\infty} \mathrm{Prob}[H_1(Y;\mathbb{F}_2)=0] = \begin{cases} 0 & p=\frac{2\log n-\omega(n)}{n} \\ 1 & p=\frac{2\log n+\omega(n)}{n} \end{cases}.$$

1. Introduction

Let $G(n,p)$ denote the probability space of graphs on the vertex set $[n]=\{1,\dots,n\}$ with independent edge probabilities p. Let log denote the natural logarithm. A classical result of Erdős and Rényi [2] asserts that the threshold probability for connectivity of $G\in G(n,p)$ coincides with the threshold for the non-existence of isolated vertices in G. In particular, for any function $\omega(n)$ that tends to infinity

$$\lim_{n\to\infty} \mathrm{Prob}[G\in G(n,p)\;:\;G \text{ connected}] = \begin{cases} 0 & p=\frac{\log n-\omega(n)}{n} \\ 1 & p=\frac{\log n+\omega(n)}{n} \end{cases}.$$

In this paper we study an analogous problem for random $2-$dimensional complexes. Unlike the graphical case, there are several distinct notions of

Mathematics Subject Classification (2000): 55U10, 05C80

* Supported by an Israel Science Foundation grant.

1-dimensional connectivity of a (connected) 2-dimensional complex X. The strongest such notion is *simple connectivity*, i.e. the triviality of the fundamental group $\pi_1(X)$. Next comes *homological* 1*-connectivity*, i.e. the vanishing of the first integral homology $H_1(X;\mathbb{Z})$. Here we are concerned with the still weaker notion of $\mathbb{F}_2$*-homological* 1*-connectivity*, namely the vanishing of the first homology with mod 2 coefficients $H_1(X;\mathbb{F}_2)$.

We recall some topological terminology (see e.g. [3]). For a simplicial complex X, let $f_k(X)$ denote the number of k-dimensional simplices in X and let $X^{(k)}=\{\sigma\in X:\dim\sigma\le k\}$ denote the k-dimensional skeleton of X. Let $C^k(X)$ denote the $\mathbb{F}_2$-vector space of $\mathbb{F}_2$ k-cochains on X – i.e. the space of $\mathbb{F}_2$-valued functions on the k-simplices of X. The coboundary operator $d_k : C^k(X)\rightarrow C^{k+1}(X)$ is given as follows. For $f\in C^k(X)$ and a $(k+1)$-dimensional $\sigma\in X$

$$d_k(f)(\sigma)=\sum_{\tau\subset\sigma\,,\ \dim\tau=k} f(\tau)\,.$$

Let $Z^k(X)=\ker(d_k)$ denote the space of k-cocycles of X and let $B^k(X)=\mathrm{Im}(d_{k-1})$ denote the space of k-coboundaries of X. The k-th cohomology group of X with $\mathbb{F}_2$ coefficients is

$$H^k(X;\mathbb{F}_2)=\frac{Z^k(X)}{B^k(X)}\,.$$

The k-th homology group $H_k(X;\mathbb{F}_2)$ is isomorphic to $H^k(X;\mathbb{F}_2)$. We abbreviate $H^k(X)=H^k(X;\mathbb{F}_2)$.

Let Δ_{n-1} denote the $(n-1)$-dimensional simplex on the vertex set $[n]$. Let $Y(n,p)$ denote the probability space of subcomplexes $\Delta_{n-1}^{(1)}\subset Y\subset\Delta_{n-1}^{(2)}$ with probability measure

$$\Pr(Y)=p^{f_2(Y)}(1-p)^{\binom{n}{3}-f_2(Y)}.$$

An edge $ij\in\binom{[n]}{2}$ is *isolated* in Y if it is not contained in any of the 2-simplices of Y. If ij is isolated then the indicator function of ij is a non-trivial 1-cocycle of Y, hence $H^1(Y)\neq 0$. Our main result is that the threshold probability for the vanishing of $H^1(Y)$ coincides with the threshold for the non-existence of isolated edges in Y.

Theorem 1.1. *Let $\omega(n)$ be any function which satisfies $\omega(n)\rightarrow\infty$ then*

$$(1)\qquad \lim_{n\to\infty}\mathrm{Prob}[Y\in Y(n,p):H^1(Y;\mathbb{F}_2)=0]=\begin{cases}0 & p=\frac{2\log n-\omega(n)}{n}\\ 1 & p=\frac{2\log n+\omega(n)}{n}\end{cases}.$$

The case $p=\frac{2\log n-\omega(n)}{n}$ is straightforward: Let $g(Y)$ denote the number of isolated edges of Y. Then

$$E[g]=\binom{n}{2}(1-p)^{n-2}=\Omega(\exp(\omega(n)))\,.$$

A standard second moment argument then shows that

$$\mathrm{Prob}[H^1(Y)=0]\le \mathrm{Prob}[g=0]=o(1)\,.$$

Before proceeding to the main part of the paper we briefly outline the analogy between the proof of Theorem 1.1 and that of the classical theorem on the threshold for graph connectivity. From a topological perspective, graph connectivity is homological 0-connectivity and here we are concerned with $\mathbb{F}_2$-homological 1-connectivity. The coupon-collector arguments for establishing the thresholds for the existence of an isolated vertex/edge are identical in both cases. The proof that $p=\frac{\log n+\omega(n)}{n}$ guarantees a.e. connectivity of $G\in G(n,p)$ may be formulated as follows. For $g:[n]\to\mathbb{F}_2$, let $B(g)$ be the number of edges $e=uv$ in the complete graph K_n, such that $d_0(g)(e)=g(u)+g(v)=1$. In other words, $B(g)=|g^{-1}(0)||g^{-1}(1)|$ is the number of edges in the cut $(g^{-1}(0),g^{-1}(1))$ of the complete graph K_n. A graph $G=([n],E)$ is thus disconnected iff there exists a non-constant function $g:V\to\mathbb{F}_2$ with $d_0(g)(e)=0$ for all $e\in E$. For a given g, the probability of this event in the space $G(n,p)$ is $(1-p)^{B(g)}$. Limiting ourselves, as we may, to g's with $|g^{-1}(1)|\le n/2$, we now apply a union bound to derive the desired conclusion.

Something similar happens here with $p=\frac{2\log n+\omega(n)}{n}$. For a mapping $f:\binom{[n]}{2}\to\mathbb{F}_2$, let $B(f)$ be the number of triples $\sigma=uvw$ for which $d_1(f)(\sigma)=f(uv)+f(vw)+f(uw)=1$. A complex $\Delta^{(1)}_{n-1}\subset Y\subset\Delta^{(2)}_{n-1}$ is not $\mathbb{F}_2$-homologically 1-connected iff there exists a mapping $f:\binom{[n]}{2}\to\mathbb{F}_2$ that is not of the form $d_0(g)$ for any g, such $d_1(f)(\sigma)=0$ for all 2-dimensional $\sigma\in Y$. Again, for a given f, this event has probability $(1-p)^{B(f)}$ in the space $Y(n,p)$. At this point, the perfect analogy breaks and we need more refined arguments. In the graphical case, the functions g and $1-g$ are equivalent, whence we had the liberty of assuming that $|\mathrm{Supp}(g)|\le n/2$, by switching, if necessary from g to $1-g$. Likewise, here f is equivalent to any function of the form $f+d_0(h)$ for any $h:[n]\to\mathbb{F}_2$. (Note that $d_0(h)$ is simply the characteristic function of a cut in K_n.) We are thus allowed to consider only functions f such that $|\mathrm{Supp}(f)|\le|\mathrm{Supp}(f+d_0(h))|$ for all h. It turns out quite easily that we may further restrict ourselves and assume that the graph $([n], \mathrm{Supp}(f))$ has a single nontrivial connected component. (This simplification has no graphical analogue.) Let $\mathcal{F}'_n$ be the class of all functions f satisfying these two assumptions. As in the graphical case we estimate $\mathrm{Prob}[H_1(Y) \neq 0] \leq \sum_{f \in \mathcal{F}'_n} (1-p)^{B(f)}$. The proof that this sum is $o(1)$ is significantly more involved than in the graphical case. There are two main steps in the proof. First we show (Proposition 2.1) that $B(f) \geq cn \cdot |\mathrm{Supp} f|$ for every $f \in \mathcal{F}'_n$, where $c > 0$ is an absolute constant. We then derive an upper bound (Proposition 2.3) on the number of $f \in \mathcal{F}'_n$ for which $B(f) = (1-\theta)n \cdot |\mathrm{Supp} f|$ for θ bounded away from zero.

We turn to a more concrete discussion. For an $f \in C^1(\Delta_{n-1})$ denote by $[f]$ the image of f in $H^1(\Delta_{n-1}^{(1)})$. Let

$$B(f) = \left| \left\{ \sigma \in \binom{[n]}{3} \ : \ d_1(f)(\sigma) = 1 \right\} \right| .$$

For any complex $Y \supset \Delta_{n-1}^{(1)}$ we identify $H^1(Y)$ with its image under the natural injection $H^1(Y) \to H^1(\Delta_{n-1}^{(1)})$. It follows that for $f \in C^1(\Delta_{n-1})$

$$\mathrm{Prob}[[f] \in H^1(Y)] = (1-p)^{B(f)}.$$

Let $\mathrm{Supp} f = \{e \in \binom{[n]}{2} : f(e) = 1\}$. The *Hamming Weight* of the cohomology class $[f] \in H^1(\Delta_{n-1}^{(1)})$ is defined by

$$\begin{aligned} w([f]) &= \min\{|\mathrm{Supp} f'| \ : \ [f'] = [f]\} \\ &= \min\{|\mathrm{Supp}(f + d_0(h))| \ : \ h \in C^0(\Delta_{n-1})\}. \end{aligned}$$

Let

$$\mathcal{F}_n = \{f \in C^1(\Delta_{n-1}) \ : \ w([f]) = |\mathrm{Supp} f|\}.$$

Associate with any $f \in C^1(\Delta_{n-1})$ the simple undirected graph $G_f = ([n], E_f)$ with edge set $E_f = \mathrm{Supp} f$. Let $\mathcal{F}'_n$ consist of all $0 \neq f \in \mathcal{F}_n$ such that G_f has exactly one connected component which is not an isolated point.

If $H^1(Y) \neq 0$ then any $0 \neq f \in Z^1(Y)$ with minimal support must belong to $\mathcal{F}'_n$. Therefore

$$\mathrm{Prob}[H^1(Y) \neq 0] \leq \sum_{f \in \mathcal{F}'_n} \mathrm{Prob}[[f] \in H^1(Y)] .$$

Theorem 1.1 will thus follow from

Theorem 1.2. *For* $p = \frac{2 \log n + \omega(n)}{n}$

$$\sum_{f \in \mathcal{F}'_n} (1-p)^{B(f)} = o(1) . \tag{2}$$

In Section 2 we formulate the problem in graph theoretical terms and state Propositions 2.1 and 2.3 which are the main ingredients in the proof of Theorem 1.2. The proofs of these results are given in Sections 3 and 4. In Section 5 we combine Propositions 2.1 and 2.3 to derive Theorem 1.2. We conclude in Section 6 with some remarks on possible extensions and open problems.

2. A Graph Theoretic Formulation

The mapping $f \to G_f$ defines a 1-1 correspondence between the 1-cochains $C^1(\varDelta_{n-1})$ and the (simple undirected) graphs on the vertex set $[n]$. For a graph $G=([n],E)=G_f$ we denote $B(G)=B(f)$. Clearly $B(G)$ is the number of triangles $T\in\binom{[n]}{3}$ which contain either one or three edges of G. Let

$$\mathcal{G}_n = \{G_f \;:\; f\in\mathcal{F}_n\}\,, \qquad \mathcal{G}'_n = \{G_f \;:\; f\in\mathcal{F}'_n\}\,.$$

Suppose $G=([n],E)=G_f\in\mathcal{G}_n$. Then f satisfies $w([f])=|\mathrm{Supp}(f)|$. Hence for all $g\in C^0(\varDelta_{n-1})$

$$|\mathrm{Supp}(f+d_0(g))| \geq |\mathrm{Supp}(f)|\,. \tag{3}$$

Let $S=\mathrm{Supp}(g)=\{v\in[n]:g(v)=1\}$ then $\mathrm{Supp}(d_0(g))\subset\binom{[n]}{2}$ consists of all $|S||\overline{S}|$ edges of the cut $(S,\overline{S})$.

It follows that $G\in\mathcal{G}_n$ if and only if

$$|E\cap(S,\overline{S})| \leq \frac{|S||\overline{S}|}{2} \tag{4}$$

for all cuts $(S,\overline{S})$. In particular the maximal degree in $G\in\mathcal{F}_n$ is less then $\frac{n}{2}$. $\mathcal{G}'_n$ consists of all $G\in\mathcal{G}_n$ that have at most one connected component which is not an isolated point. The proof of Theorem 1.2 depends on two results. In section 3 we prove the following lower bound on $B(G)$:

Proposition 2.1. *There exists a constant $c\geq\frac{1}{120}$ such that for any $G=([n],E)\in\mathcal{G}_n$*

$$B(G)\geq c|E|n\,. \tag{5}$$

Remark. Proposition 2.1 is equivalent to the following result which roughly says that if Y is close to $\varDelta_{n-1}^{(2)}$ then any 1-cohomology class of Y must contain a cocycle with a small support.

Corollary 2.2. *Let c be the constant of Proposition 2.1. Then for any simplicial complex $\Delta_{n-1}^{(1)} \subset Y \subset \Delta_{n-1}^{(2)}$ and any $[f] \in H^1(Y;\mathbb{F}_2)$*

$$w([f]) \leq c^{-1} \frac{\binom{n}{3} - f_2(Y)}{n}. \tag{6}$$

Example. Suppose n is divisible by 3. Let $[n] = \bigcup_{i=0}^{2} V_i$ be a partition of $[n]$ with $|V_i| = \frac{n}{3}$ and let

$$Y = \Delta_{n-1}^{(2)} - \left\{ \sigma \in \binom{[n]}{3} \ : \ |\sigma \cap V_i| = 1 \text{ for all } 0 \leq i \leq 2 \right\}.$$

The characteristic function $f = 1_E \in C^1(Y)$ of the set

$$E = \{\{u,v\} \ : \ u \in V_0, \ v \in V_1\}$$

is a 1-cocycle of Y. It can be checked that f is a cocycle of minimal support in its cohomology class $[f]$. Hence

$$w([f]) = |\mathrm{Supp} f| = \frac{n^2}{9} = 3 \frac{\binom{n}{3} - f_2(Y)}{n}.$$

It follows that c cannot be replaced by any constant bigger then $\frac{1}{3}$.

For $k \leq \binom{n}{2}$ and $0 \leq \theta \leq 1$ Let

$$\begin{aligned} \mathcal{G}_n(k) &= \{G = ([n], E) \in \mathcal{G}_n \ : \ |E| = k\} \\ \mathcal{G}'_n(k) &= \{G = ([n], E) \in \mathcal{G}'_n \ : \ |E| = k\} \\ \mathcal{G}_n(k,\theta) &= \{G \in \mathcal{G}_n(k) \ : \ B(G) = (1-\theta)kn\}. \end{aligned}$$

In Section 4 we prove an estimate on the cardinality of $\mathcal{G}_n(k,\theta)$:

Proposition 2.3. *For any $0 < \epsilon < \frac{1}{2}$ there exists a constant $C(\epsilon)$ such that for any $\theta \geq \epsilon$ and $\frac{n}{5} \leq k \leq n^{2-\epsilon}$*

$$|\mathcal{G}_n(k,\theta)| \leq \left(C(\epsilon) \cdot n^{2(1-(2-\epsilon)\theta)} \right)^k. \tag{7}$$

3. A Lower Bound on $B(G)$

Proof of Proposition 2.1. We show (5) with $c=\frac{1}{120}$. For $e\in E$ let $\alpha(e)$ denote the number of 2-simplices $\sigma\in\binom{[n]}{3}$ which contain e but whose two other edges are not in E. Let $\beta(e)$ denote the number of $\sigma\in\binom{[n]}{3}$ which contain e and whose two other edges are both in E. Then

$$B(G)=\sum_{e\in E}\alpha(e)+\frac{1}{3}\sum_{e\in E}\beta(e)\,. \tag{8}$$

Suppose for contradiction that

$$B(G)<c|E|n\,. \tag{9}$$

Let $\delta<1$ be a constant whose value will be assigned later, and let

$$E'=\left\{e\in E\;:\;\alpha(e)+\frac{1}{3}\beta(e)\le\frac{cn}{\delta}\right\}.$$

Then (8) and (9) imply that $|E'|\ge(1-\delta)|E|$.
For $v\in[n]$ let $\Gamma_G(v)=\{u\in[n]:uv\in E\}$ and let $\deg_G(v)=|\Gamma_G(v)|$. Recall that

$$\deg_G(v)<\frac{n}{2} \tag{10}$$

for all $v\in[n]$. Let $e=uv\in E'$ then

$$\frac{cn}{\delta}\ge\alpha(e)=n-|\Gamma_G(u)\cup\Gamma_G(v)|\ge n-(\deg_G(u)+\deg_G(v)) \tag{11}$$

and

$$\frac{3cn}{\delta}\ge\beta(e)=|\Gamma_G(u)\cap\Gamma_G(v)|\,. \tag{12}$$

(10) and (11) imply that $\deg_G(u),\deg_G(v)\ge(1-\frac{2c}{\delta})\frac{n}{2}$.

Let $G'=([n],E')$ then

$$(1-\delta)\sum_{u\in[n]}\deg_G(u)=2(1-\delta)|E|\le 2|E'|=\sum_{u\in[n]}\deg_{G'}(u)\,.$$

Hence there exists a $u\in[n]$ such that

$$0<(1-\delta)\deg_G(u)\le\deg_{G'}(u)\,.$$

Since u is incident with at least one edge in E' it follows that $\deg_G(u)\ge(1-\frac{2c}{\delta})\frac{n}{2}$. Thus

$$\deg_{G'}(u)\ge(1-\delta)\left(1-\frac{2c}{\delta}\right)\cdot\frac{n}{2}\,.$$

Let $S = \mathit{\Gamma}_{G'}(u)$ then (12) implies that for each $v \in S$

$$|\mathit{\Gamma}_G(v) \cap S| \le |\mathit{\Gamma}_G(v) \cap \mathit{\Gamma}_G(u)| \le \frac{3cn}{\delta}.$$

It follows that

$$|\mathit{\Gamma}_G(v) \cap \overline{S}| \ge \deg_G(v) - \frac{3cn}{\delta} \ge \left(1 - \frac{2c}{\delta}\right)\frac{n}{2} - \frac{3cn}{\delta} = \left(1 - \frac{8c}{\delta}\right) \cdot \frac{n}{2}.$$

Therefore

$$|E \cap (S, \overline{S})| = \sum_{v \in S} |\mathit{\Gamma}_G(v) \cap \overline{S}| \ge$$
$$(1-\delta)\left(1 - \frac{2c}{\delta}\right)\frac{n}{2} \cdot \left(1 - \frac{8c}{\delta}\right)\frac{n}{2} \ge (1-\delta)\left(1 - \frac{10c}{\delta}\right) \cdot \frac{n^2}{4}.$$

Taking $\delta = \sqrt{10c}$ we obtain

$$|E \cap (S, \overline{S})| \ge (1 - \sqrt{10c})^2 \cdot |(S, \overline{S})| > \frac{1}{2}|(S, \overline{S})|$$

in contradiction with (4). ∎

4. Near Transversals

The proof of Proposition 2.3 depends on the following

Claim 4.1. *For any $0 < \epsilon < \frac{1}{2}$ there exist constants $C_1(\epsilon)$, $n_0(\epsilon)$ such that for any $\epsilon \le \theta \le 1$ and for any graph $G = ([n], E)$ with $n \ge n_0(\epsilon)$ and degree sequence $d_1, \ldots, d_n \le \frac{n}{2}$ which satisfies*

$$\frac{n}{5} \le |E| = k \le n^{2-\epsilon} \qquad \text{and} \qquad \sum_{i=1}^{n} d_i^2 \ge \theta k n \tag{13}$$

there exists a set of vertices $S \subset [n]$ such that

$$|S| \le \frac{C_1(\epsilon) \cdot k}{n} \qquad \text{and} \qquad |\{e \in E : S \cap e \ne \emptyset\}| \ge (2 - \epsilon)\theta k. \tag{14}$$

Proof. Pick S at random with $\mathrm{Prob}[i \in S] = \frac{2d_i}{n}$ and let

$$X = |\{e \in E \,:\, S \cap e \ne \emptyset\}|.$$

Then

$$E[X] = \sum_{ij\in E}\left(1-\left(1-\frac{2d_i}{n}\right)\left(1-\frac{2d_j}{n}\right)\right) =$$

$$\frac{2}{n}\sum_{ij\in E}(d_i+d_j) - \frac{4}{n^2}\sum_{ij\in E}(d_i\cdot d_j) \geq \frac{2}{n}\sum_{i=1}^{n} d_i^2 - \frac{2}{n^2}\left(\sum_{i=1}^{n} d_i\right)^2 \geq$$

$$2\theta k - \frac{8k^2}{n^2} \geq (2\theta - 8n^{-\epsilon})k\,.$$

It follows that for $n \geq n_0(\epsilon) = (\frac{100}{\epsilon^2})^{\frac{1}{\epsilon}}$

$$\text{Prob}[X \geq (2-\epsilon)\theta k] \geq \text{Prob}\left[X \geq \left(1-\frac{\epsilon}{3}\right)E[X]\right] \geq$$

$$\frac{\epsilon}{3}\cdot\frac{E[X]}{k} \geq \frac{\epsilon\theta}{3} \geq \frac{\epsilon^2}{3}\,. \tag{15}$$

Since

$$E[|S|] = \sum_{i=1}^{n}\frac{2d_i}{n} = \frac{4k}{n}$$

it follows by a large deviation estimate (see e.g. Theorem A.1.12 in [1]) that for $\beta > e$

$$\text{Prob}\left[|S| > \beta\frac{4k}{n}\right] < \left(\frac{e}{\beta}\right)^{\frac{4\beta k}{n}} \leq \left(\frac{e}{\beta}\right)^{\frac{4\beta}{5}}. \tag{16}$$

Hence for a sufficiently large $C_1(\epsilon)$

$$\text{Prob}\left[|S| > \frac{C_1(\epsilon)\cdot k}{n}\right] < \frac{\epsilon^2}{3}\,. \tag{17}$$

(15) and (17) imply that there exists an $S \subset [n]$ which satisfies (14). ∎

Remark. Claim 4.1 is not far from optimal in the sense that there exist graphs which satisfy (13), but such that any set of $O(\frac{k}{n})$ vertices intersects at most $2\theta k(1+o(1))$ edges. Indeed let $\theta < \frac{1}{2}$ be fixed and let $\Omega(n) \leq k \leq o(n^2)$. Let A, B, C be three disjoint subsets of $[n]$ such that $|A| = \frac{4\theta k}{n}$, $|B| = \frac{n}{2}$, and $|C| = t$ where $\binom{t}{2} = (1-2\theta)k$. Let $G = ([n], E)$ consist of the complete bipartite graph with sides A and B, together with the complete graph on C. Then $|E| = k$, $d_i \leq \frac{n}{2}$ and $\sum_{i=1}^{n} d_i^2 \geq \theta kn$. For any $S \subset [n]$ such that $|S| = O(\frac{k}{n})$

$$|\{e \in E : S \cap e \neq \emptyset\}| \leq 2\theta k + |S|\cdot t = 2\theta k(1+o(1))\,.$$

Proof of Proposition 2.3. It suffices to consider $n \geq n_0(\epsilon)$. Let $G = ([n],E) \in \mathcal{G}_n(k,\theta)$. Then $d_i = \deg_G(i) < \frac{n}{2}$ and

$$(1-\theta)kn = B(G) \geq \sum_{e\in E} \alpha(e) \geq \sum_{ij\in E} (n - \deg_G(i) - \deg_G(j)) = kn - \sum_{i=1}^{n} d_i^2 .$$

It follows that $\sum_{i=1}^n d_i^2 \geq \theta kn$ hence by Claim 4.1 there exists a subset $S \subset [n]$ which satisfies (14). We now estimate $|\mathcal{G}_n(k,\theta)|$ as follows: The number of possible subsets S is at most 2^n. The number of choices for the edges of G which are incident with a fixed S is at most

$$\sum_{\ell=(2-\epsilon)\theta k}^{k} \binom{|S|n - \binom{|S|+1}{2}}{\ell} < k\binom{|S|n}{k} < k\left(\frac{e|S|n}{k}\right)^k < k(eC_1(\epsilon))^k .$$

The number of choices for the edges that are not covered by S is at most $\binom{\binom{n}{2}}{k(1-(2-\epsilon)\theta)}$. It follows that

$$|\mathcal{G}_n(k,\theta)| \leq 2^n \cdot k \cdot (eC_1(\epsilon))^k \cdot n^{2k(1-(2-\epsilon)\theta)} \leq \left(C(\epsilon)\cdot n^{2(1-(2-\epsilon)\theta)}\right)^k . \qquad \blacksquare$$

5. Proof of Theorem 1.2

Proof of Theorem 1.2. Let $\omega(n) \to \infty$ and let $p = \frac{2\log n + \omega(n)}{n}$. Writing (2) in the equivalent graph theoretic formulation, we have to show that

$$\sum_{k\geq 1} \sum_{G\in \mathcal{G}_n'(k)} (1-p)^{B(G)} = \sum_{G\in\mathcal{G}_n'} (1-p)^{B(G)} = o(1)\,. \tag{18}$$

Let $c \geq \frac{1}{120}$ denote the constant of Proposition 2.1. We deal separately with three intervals of k:

(i) $1 \leq k \leq \frac{n}{5}$. Recall that $G \in \mathcal{G}_n'(k)$ has exactly one connected component which is not an isolated point. It follows that

$$|\mathcal{G}_n'(k)| \leq \binom{n}{k+1}\binom{\binom{k+1}{2}}{k} \leq \left(\frac{en}{k+1}\right)^{k+1}\left(\frac{e(k+1)}{2}\right)^k < (10n)^{k+1}.$$

Next note that $\alpha(e) \geq n-k-1$ for all edges e of G, hence $B(G) \geq k(n-k-1)$. It follows that

$$\sum_{k=2}^{n/5} \sum_{G\in\mathcal{G}_n'(k)} (1-p)^{B(G)} \leq \sum_{k=2}^{n/5} |\mathcal{G}_n'(k)|(1-p)^{k(n-k-1)} \leq$$

$$\sum_{k=2}^{n/5} (10n)^{k+1}\left(n^{-2}e^{-\omega(n)}\right)^{\frac{4k}{5}} \leq \sum_{k=2}^{n/5} (10n)^{k+1} n^{-\frac{8k}{5}} = O(n^{-\frac{1}{5}}).$$

Hence

$$\sum_{k=1}^{n/5} \sum_{G\in\mathcal{G}'_n(k)} (1-p)^{B(G)} \le \binom{n}{2}(1-p)^{n-2} + O(n^{-\frac{1}{5}})$$
$$\le \exp\left(-\frac{\omega(n)}{2}\right) + O(n^{-\frac{1}{5}})\,. \tag{19}$$

(ii) $\frac{n}{5}\le k\le n^{2-c}$. We need the following:

Claim 5.1. *For any* $\frac{n}{5}\le k\le n^{2-c}$ *and* $0\le\theta\le 1$

$$\sum_{G\in\mathcal{G}_n(k,\theta)} (1-p)^{B(G)} = O(n^{-\frac{k}{4}})\,. \tag{20}$$

Proof. If $0\le\theta\le\frac{1}{3}$ then

$$\sum_{G\in\mathcal{G}_n(k,\theta)} (1-p)^{B(G)} \le \binom{\binom{n}{2}}{k}(1-p)^{(1-\theta)kn}$$
$$\le \left(\frac{en^2}{2k}\right)^k n^{-2(1-\theta)k} \le (10n^{2\theta-1})^k \le (10n^{-\frac{1}{3}})^k.$$

Suppose now that $\frac{1}{3}\le\theta\le 1$. Applying Proposition 2.3 with $\epsilon = c$ we obtain

$$\sum_{G\in\mathcal{G}_n(k,\theta)} (1-p)^{B(G)} \le |\mathcal{G}_n(k,\theta)|\cdot(1-p)^{(1-\theta)kn}$$
$$\le \left(C(\epsilon)\cdot n^{2(1-(2-\epsilon)\theta)}\right)^k\cdot n^{-2(1-\theta)k}$$
$$= \left(C(\epsilon)\cdot n^{-2\theta(1-\epsilon)}\right)^k$$
$$\le \left(C(c)\cdot n^{-\frac{2(1-c)}{3}}\right)^k.$$

∎

For each k the number of θ's for which $\mathcal{G}_n(k,\theta)$ is non-empty is at most n^3. It follows by Claim 5.1 that

$$\sum_{k=\frac{n}{5}}^{n^{2-c}} \sum_{G\in\mathcal{G}_n(k)} (1-p)^{B(G)} = \sum_{k=\frac{n}{5}}^{n^{2-c}} \sum_{\theta} \sum_{G\in\mathcal{G}_n(k,\theta)} (1-p)^{B(G)}$$
$$\le n^3 \sum_{k=\frac{n}{5}}^{n^{2-c}} O\left(n^{-\frac{k}{4}}\right) = n^{-\Omega(n)}\,. \tag{22}$$

(iii) $k \geq n^{2-c}$. By Proposition 2.1

$$\sum_{k \geq n^{2-c}} \sum_{G \in \mathcal{G}_n(k)} (1-p)^{B(G)} \leq \sum_{k \geq n^{2-c}} \binom{\binom{n}{2}}{k} (1-p)^{ckn}$$

$$\leq \sum_{k \geq n^{2-c}} \left(\frac{en^2}{2k}\right)^k n^{-2ck} \leq \sum_{k \geq n^{2-c}} (2n^{-c})^k = n^{-\Omega(n)} . \tag{23}$$

Finally (18) follows from (19), (22) and (23). ∎

6. Concluding Remarks

We have shown that in the model $Y(n,p)$ of random 2-complexes on n vertices, the threshold for the vanishing of $H_1(Y;\mathbb{F}_2)$ occurs at $p=\frac{2\log n}{n}$. A straightforward extension of the proof shows that the same result holds for homology with coefficients in any *fixed* finite abelian group. We still do not know the answer to the next set of naturally-arising questions: Where is the threshold for the vanishing of $H_1(Y,\mathbb{Z})$? For being simply connected? What happens in the higher-dimensional situation?

We believe the methods of this paper will be relevant to the questions of homological connectivity of random complexes in higher dimensions. The problem of simple connectivity will probably require a different approach and is particularly intriguing, since this property is, in general, undecidable. One point worth mentioning here is this. The three conditions: Vanishing of the first homology over $\mathbb{F}_2$, over $\mathbb{Z}$, and simple connectivity, are progressively stronger, in this order. If the thresholds for their occurrence differ, this will supply us with a large set of instances where some, but not all of these conditions hold. This may be of interest also outside of combinatorics. If the thresholds coincide, it will be of interest to develop more refined probabilistic models which do differentiate between these criteria for connectivity.

At any event, we believe that further study of topological properties of random complexes will prove both interesting and useful.

Acknowledgement. We would like to thank Gil Kalai for helpful discussions.

References

[1] N. Alon and J. Spencer: *The Probabilistic Method*, 2nd Edition, Wiley-Intescience, 2000.

[2] P. Erdős and A. Rényi: On the evolution of random graphs, *Publ. Math. Inst. Hungar. Acad. Sci.* **5** (1960), 17–61.
[3] J. Munkres: *Elements of Algebraic Topology*, Addison-Wesley, 1984.

Nathan Linial

Department of Computer Science
Hebrew University
Jerusalem 91904
Israel
nati@cs.huji.ac.il

Roy Meshulam

Department of Mathematics
Technion
Haifa 32000
Israel
meshulam@math.technion.ac.il