Available online at www.sciencedirect.com
ScienceDirect
Discrete Mathematics 309 (2009) 1658–1671
DISCRETE MATHEMATICS
www.elsevier.com/locate/disc

Topology of random clique complexes ☆

Matthew Kahle [1]

Department of Mathematics, University of Washington, Seattle, United States

Received 27 June 2006; received in revised form 27 February 2008; accepted 27 February 2008
Available online 16 April 2008

Abstract

In a seminal paper, Erdős and Rényi identified a sharp threshold for connectivity of the random graph $G(n, p)$. In particular, they showed that if $p \gg \log n/n$ then $G(n, p)$ is almost always connected, and if $p \ll \log n/n$ then $G(n, p)$ is almost always disconnected, as $n \to \infty$.

The *clique complex* $X(H)$ of a graph H is the simplicial complex with all complete subgraphs of H as its faces. In contrast to the zeroth homology group of $X(H)$, which measures the number of connected components of H, the higher dimensional homology groups of $X(H)$ do not correspond to monotone graph properties. There are nevertheless higher dimensional analogues of the Erdős–Rényi Theorem.

We study here the higher homology groups of $X(G(n, p))$. For $k > 0$ we show the following. If $p = n^{\alpha}$, with $\alpha < -1/k$ or $\alpha > -1/(2k + 1)$, then the kth homology group of $X(G(n, p))$ is almost always vanishing, and if $-1/k < \alpha < -1/(k + 1)$, then it is almost always nonvanishing.

We also give estimates for the expected rank of homology, and exhibit explicit nontrivial classes in the nonvanishing regime. These estimates suggest that almost all d-dimensional clique complexes have only one nonvanishing dimension of homology, and we cannot rule out the possibility that they are homotopy equivalent to wedges of a spheres.
© 2008 Elsevier B.V. All rights reserved.

Keywords: Random graph; Phase transition; Clique complex; Flag complex; Discrete Morse theory

1. Introduction

A pioneering result in the theory of random graphs is the Erdős–Rényi theorem on the threshold for connectivity [4]. This is a primary inspiration for the line of research pursued here, and some of our results may be viewed as generalizations of the Erdős–Rényi theorem to higher dimensions, so we begin by defining random graphs and stating their result.

The random graph $G(n, p)$ is defined to be the probability space of all graphs on vertex set $[n] = \{1, 2, \dots, n\}$ with each edge inserted independently with probability p. Frequently, p is a function of n, and one asks whether a typical graph in $G(n, p)$ is likely to have a particular property as $n \to \infty$. We say that $G(n, p)$ *almost always (a.a.)* has property $\mathcal{P}$ if $\Pr[G(n, p) \in \mathcal{P}] \to 1$ as $n \to \infty$.

☆ Supported in part by NSA #H98230-05-1-0053 and NSF-VIGRE.
E-mail address: mkahle@gmail.com.
[1] Current address: Department of Mathematics, Stanford University, Stanford, United States.

0012-365X/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.disc.2008.02.037

Theorem 1.1 (*Erdős and Rényi*). *If* $p = (\log n + \omega(n))/n$ *and* $\omega(n) \to \infty$ *as* $n \to \infty$ *then* $G(n, p)$ *is almost always connected. If* $\omega(n) \to -\infty$ *then* $G(n, p)$ *is almost always disconnected.*

The number of connected components in a graph is a *monotone graph property*. In other words, adding edges to a graph can only decrease the number of components. (As with functions $f : \mathbb{R} \to \mathbb{R}$, we could talk about graph properties either being monotone increasing or decreasing.) Much random graph theory is concerned with monotone graph properties: chromatic number, clique number, subgraph containment, diameter, and so on [3].

The *clique complex* $X(G)$ of a graph G is the simplicial complex with all complete subgraphs of G as its faces. The 1-skeleton of $X(G)$ is G itself, so Erdős and Rényi's result may be interpreted as a statement about homology $\widetilde{H}_0(X(G(n, p)))$ or homotopy $\pi_0(X(G(n, p)))$. (For a brief introduction to the topological notions discussed in this article, please see Section 2.)

Note: To streamline notation, we will abbreviate $X(G(n, p))$ **by** $X(n, p)$ **for the rest of the article.**

Our main objects of study are $\widetilde{H}_k(X(n, p))$ and $\pi_k(X(n, p))$ for each fixed $k > 0$. We find that vanishing of higher homology is not monotone, as homology vanishes for large and small functions p, but is nonvanishing for some regime in between. Still, it is possible to make statements which generalize Theorem 1.1.

Another way to state our results is to fix the dimension d of the clique complex (by appropriately choosing p), rather than looking at a fixed homology group. In this case, we find that the homology of $X(n, p)$ is highly concentrated in its middle dimensions. Asymptotically, a d-dimensional random clique complex a.a. has trivial homology above dimension $\lfloor d/2 \rfloor$ and below dimension $\lfloor d/4 \rfloor$. On the other hand, homology is almost always nontrivial in dimension $\lfloor d/2 \rfloor$.

In fact we cannot rule out the possibility that the only nontrivial homology is in dimension $\lfloor d/2 \rfloor$. We give some evidence for this by estimating the expectation of the Betti numbers.

In Section 8 we briefly survey other papers concerning topology of random simplicial complexes.

2. Topological notions

The reader who is familiar with reduced homology and homotopy groups of topological spaces may feel free to skip this section. For anyone not so familiar, this will only serve as the very briefest of introductions, and will probably not be sufficient to understand the more technical parts of the proofs, but at the suggestion of an anonymous referee, we are including this section in order to make the article accessible to a wider audience. For anyone who wants to know more, a very nice introduction to algebraic topology is Allan Hatcher's book [6].

The *reduced homology groups* $\widetilde{H}_i(X, k)$, $i = 0, 1, 2, \ldots$, where $k = \mathbb{Z}$ or some field, are topological invariants associated with a topological space X. Very roughly, $\widetilde{H}_i(X, k)$ measures the number of i-dimensional holes in X. Suppose X is a finite simplicial complex of dimension d. The most important topological facts for the purposes of this article are the following.

- $\widetilde{H}_i(X, k)$ is a finitely generated abelian group. In the case the k is a field, it is a vector space over k.
- The ith *Betti number* is $\beta_i = \dim \widetilde{H}_i(X, \mathbb{Q})$. A classical fact is that if f_i is the number of i-dimensional faces of X, then the following Euler formula holds.

  $$f_0 - f_1 + \cdots + (-1)^d f_d = \beta_0 - \beta_1 + \cdots + (-1)^d \beta_d.$$

  Also, it follows directly from the definition of simplicial homology and dimensional considerations that for every i,

  $$-f_{i-1} + f_i - f_{i+1} \le \beta_i \le f_i.$$
- $\widetilde{H}_0(X, k) = 0$ if and only if X is connected.
- $\widetilde{H}_i(X, k) = 0$ for $i > d$.
- (Universal coefficients for homology) If $\widetilde{H}_i(X, \mathbb{Z}) = 0$ for $0 \le i \le m$ then $\widetilde{H}_i(X, k) = 0$ for any coefficients k.

We also briefly discuss the *homotopy groups* $\pi_i(X)$. Again, see [6] for a nice introduction, but the following facts will be more than sufficient to read this article.

- $\pi_i(X)$ is the set of homotopy classes of maps from the sphere $S^i \mapsto X$. In particular, we say $\pi_0(X) = \{0\}$ if and only if X is path connected.

- (Hurewicz Theorem) If $\pi_i(X) = \{0\}$ for $i \leq n$, (in which case we say that X is *n-connected*) then $\widetilde{H}_i(X, \mathbb{Z}) = 0$, $i = 0, 1, \ldots, n$.
- $\widetilde{H}_1(X, \mathbb{Z})$ is the abelianization of the fundamental group $\pi_1(X)$.

Reduced homology groups and homotopy groups are topological invariants, meaning that if two spaces are homeomorphic then their associated homology and homotopy groups are isomorphic. A stronger statement, also true, is that they are homotopy invariants, meaning that the same holds even if the spaces are only homotopy equivalent.

3. Statement of results

We discuss which groups $\widetilde{H}_i(X(n, p), \mathbb{Z})$ are nontrivial, then estimate Betti numbers. For comparison with the results, note that

$$\dim(X(n, p)) \approx -2 \log n / \log p.$$

For example, since $\dim(X(H)) \geq k$ if and only if H contains $(k + 1)$-cliques, standard random graph techniques for subgraph containment [3] give that if $p = n^\alpha$ with $\alpha < -2/k$ then a.a. $\dim(X(n, p)) < k$, and if $\alpha > -2/k$ then a.a. $\dim(X(n, p)) \geq k$.

We first show that if p is large enough then homology vanishes. A topological space X is said to be *k-connected* if every map from a sphere $\mathbb{S}^i \to X$ extends to a map from the ball $\mathbb{B}^{i+1} \to X$ for $i = 0, 1, \ldots, k$. Equivalently, X is k-connected if $\pi_i(X) = 0$ for $i \leq k$, and in particular 0-connected is equivalent to path connected. This implies, by the Hurewicz Theorem [6], that $\widetilde{H}_i(X, \mathbb{Z}) = 0$ for $i \leq k$.

The following is implicit in [10], although Meshulam's result was more general and stated for homology instead of homotopy groups. We prove the homotopy statement here for the sake of completeness, although the argument is similar to Meshulam's.

Theorem 3.1 (*Meshulam*). *If every $2k+2$ vertices of a graph H have a common neighbor then $X(H)$ is k-connected.*

In the case that H is a random graph, this can be improved. For example, in the case $k = 0$, Erdős and Rényi's theorem gives that the threshold for connectivity is the same as the threshold for every vertex having at least one neighbor. The threshold for every set of l vertices having a neighbor is given by the following.

Theorem 3.2. *If $p = \left(\frac{l \log n + \omega(n)}{n}\right)^{1/l}$ and $\omega(n) \to \infty$ then a.a., every l vertices of $G(n, p)$ have a common neighbor.*

Together with Meshulam's result, we immediately have the following:

Corollary 3.3. *If $p = \left(\frac{(2k+2) \log n + \omega(n)}{n}\right)^{1/(2k+2)}$ and $\omega(n) \to \infty$ then a.a. $X(n, p)$ is k-connected.*

Corollary 3.3 can be improved and we do so with Theorem 3.4. Note that when $k = 0$, this specializes to one direction of the Erdős–Renyi theorem.

Theorem 3.4. *If $p = \left(\frac{(2k+1) \log n + \omega(n)}{n}\right)^{1/(2k+1)}$ and $\omega(n) \to \infty$ then a.a. $X(n, p)$ is k-connected.*

As a consequence, we have a statement about vanishing of homology. In a different regime, we can make statements about nonvanishing homology by exhibiting nontrivial classes explicitly.

Theorem 3.5. *If $p^{k+1} n \to 0$ and $p^k n \to \infty$ as $n \to \infty$ then $X(n, p)$ a.a. retracts onto a sphere $\mathbb{S}^k$. Hence $\widetilde{H}_k(X(n, p), \mathbb{Z})$ a.a. has a $\mathbb{Z}$ summand.*

Theorem 3.4 gives a statement that if p is large enough, then homology vanishes. The same must be true when p is small enough, simply by dimensional considerations. But this kind of coarse argument will only give that $\alpha < -2/k$ then $\widetilde{H}_k(X(n, p), \mathbb{Z}) = 0$. By Theorem 3.5, the following is best possible.

Theorem 3.6. *If $p = n^\alpha$ with $\alpha < -1/k$ then $\widetilde{H}_k(X(n, p), \mathbb{Z}) = 0$ almost always.*

By Theorems 3.4–3.6, we have the following:

Corollary 3.7 (*Vanishing and Nonvanishing of Homology*). *If* $p = n^{\alpha}$ *then*

(1) *if* $\alpha < -1/k$ *or* $\alpha > -1/(2k+1)$ *then a.a.* $\widetilde{H}_k(X(G(n,p),\mathbb{Z})) = 0$,
(2) *and if* $-1/k < \alpha < -1/(k+1)$ *then a.a.* $\widetilde{H}_k(X(n,p),\mathbb{Z}) \neq 0$.

So rather than monotonicity, we have a kind of unimodality (in terms of p) for each fixed homology group as $n \to \infty$.

Corollary 3.7 does not address the case when $-1/(k+1) < \alpha < -1/(2k+1)$. We believe that Theorem 3.4 can probably be improved to say that if $p = n^{\alpha}$ with $\alpha > -1/(k+1)$ then a.a. $\widetilde{H}_k(X(n,p),\mathbb{Z}) = 0$.

To give evidence for this conjecture, we estimate the expected rank of homology, and show that it passes through phase transitions at $\alpha = -1/k$ and $-1/(k+1)$. Let f_k denote the number of k-dimensional faces of $X(n,p)$ and β_k its kth Betti number. That is, let

$$\beta_k = \dim \widetilde{H}_k(X(n,p),\mathbb{Q}),$$

although our result holds for coefficients in any field. By the definition of simplicial homology and dimensional considerations, $\beta_k \leq f_k$.

We show that, given the hypothesis of Theorem 3.5, f_k is actually a good approximation for β_k, but for p outside of this range, β_k is much smaller. We write $X \sim Y$ almost always if for every $\epsilon > 0$, as $n \to \infty$,

$$\mathbf{P}((1-\epsilon) \leq Y/X \leq (1+\epsilon)) \to 1.$$

Theorem 3.8. *If* $p^{k+1}n \to 0$ *and* $p^k n \to \infty$ *then* $E(\beta_k)/E(f_k) \to 1$. *Moreover* $\beta_k \sim E[\beta_k]$ *and* $f_k \sim E[f_k]$ *a.a., so* $\beta_k \sim f_k$ *a.a.*

Finally, we apply discrete Morse theory to show that $E[\beta_k]/E[f_k]$ passes through phase transitions at $p = n^{-1/(k+1)}$ and $p = n^{-1/k}$.

Theorem 3.9. *If* $p^{k+1}n \to \infty$ *or* $p^k n \to 0$ *then* $E(\beta_k)/E(f_k) \to 0$.

(Note that even the second case of Theorem 3.9 is not necessarily implied by Theorem 3.6, since the statement that a random variable is a.a. zero implies nothing about its expectation. Also, $p^k n \to 0$ is a slightly weaker hypothesis than $p = n^{\alpha}$ with $\alpha < -1/k$.)

As a corollary to Theorems 3.8 and 3.9 we have the following:

Corollary 3.10 (*Betti Numbers*). *If* $p = n^{\alpha}$ *then for any* $\epsilon > 0$,

(1) *if* $\alpha < -1/k$ *or* $\alpha > -1/(k+1)$ *then a.a.* $0 \leq \beta_k/f_k < \epsilon$,
(2) *if* $-1/k < \alpha < -1/(k+1)$ *then a.a.* $1 - \epsilon < \beta_k/f_k \leq 1$.

If Theorem 3.4 can be improved to say that if $\alpha > -1/(k+1)$ then $X(n,p)$ is a.a. k-connected, then the upshot is that a.a. d-dimensional clique complexes have only one nonvanishing dimension of homology. This might be a bit surprising, since it does not depend on d. In a sense. We discuss this more in Section 9.

In the next several sections we prove the results. Theorems 3.1, 3.2 and 3.4 are proved in Section 4, Theorem 3.6 in Section 5, Theorem 3.5 in Section 6, Theorems 3.8 and 3.9 in Section 7.

4. Connectivity

We use the following Nerve Theorem of Björner [2] throughout this section. The *nerve* of a family of nonempty sets $(\Delta_i)_{i\in I}$ is the simplicial complex $\mathcal{N}((\Delta_i)_{i\in I})$, defined on the vertex set I by the rule that $\sigma \in \mathcal{N}(\Delta_i)$ if and only if $\cap_{i\in\sigma} \Delta_i \neq \emptyset$. Note that the nerve depends on the whole family, but we denote it by $\mathcal{N}(\Delta_i)$ rather than $\mathcal{N}((\Delta_i)_{i\in I})$ for brevity.

Theorem 4.1 (*Björner*). *Let* Δ *be a simplicial complex, and* $(\Delta_i)_{i\in I}$ *a family of subcomplexes such that* $\Delta = \cup_{i\in I} \Delta_i$. *Suppose that every nonempty finite intersection* $\Delta_{i_1} \cap \Delta_{i_2} \cap \cdots \cap \Delta i_t$ *is* $(k-t+1)$*-connected,* $t \geq 1$. *Then* Δ *is* k*-connected if and only if* $\mathcal{N}(\Delta_i)$ *is* k*-connected.*

Proof of Theorem 3.1. We show that if every $2k+2$ vertices of a graph H have a neighbor then $X(H)$ is a.a. k-connected. Proceed by induction on k. The claim holds when $k=0$, i.e. if every pair of vertices of a graph have a common neighbor, then the graph is certainly connected. So suppose the claim holds for $k=0,\dots,i-1$ where $i\geq 1$. Further suppose that H is a graph such that every set of $2i+2$ vertices has some neighbor. We wish to show that $X(H)$ is i-connected.

Define the *star* of a vertex v in a simplicial complex Δ to be the subcomplex $\mathrm{st}_\Delta(v)$ of all faces in Δ containing v. Clearly we have $\Delta=\cup_{v\in\Delta}\mathrm{st}_\Delta(v)$. So to apply Theorem 4.1 we must check that each vertex star is itself i-connected, and that every t-wise intersection is $(i-t+1)$-connected for $t=2,\dots,i+1$.

Each star is a cone, hence contractible and in particular i-connected. Each t-wise intersection of vertex stars is a clique complex in which every $2i+2-t$ vertices share a neighbor, hence by induction is $i-\lfloor t/2\rfloor$-connected. Since $t\geq 2$, $i-\lfloor t/2\rfloor\geq i-t+1$, so the claim follows provided that $\mathcal{N}(\mathrm{st}_{X(H)}(v))$ is also i-connected. It is clear that since every $2i+2$ neighbors have a neighbor, the intersection of every $2i+2$ vertex stars is nonempty. So the $(2i+1)$-dimensional skeleton of $\mathcal{N}(\mathrm{st}_\Delta(v))$ is complete, and then $\mathcal{N}(\mathrm{st}_\Delta(v))$ is $2i$-connected. □

Proof of Theorem 3.2. We claim that if $p=\left(\frac{l\log n+\omega(n)}{n}\right)^{1/l}$ and $\omega(n)\to\infty$ then a.a. every l vertices of $G(n,p)$ have a neighbor. The expected number of l-tuples of vertices in $G(n,p)$ with no neighbor is

$$\begin{aligned}\binom{n}{l}(1-p^l)^{n-l}&\leq\binom{n}{l}\mathrm{e}^{-p^l(n-l)}\\&=\binom{n}{l}\mathrm{e}^{-\frac{l\log n+\omega(n)}{n}(n-l)}\\&=\binom{n}{l}n^{-l}\mathrm{e}^{-\omega(n)(n-l)/n}\\&\leq\mathrm{e}^{-\omega(n)(1-l/n)}\\&=o(1),\end{aligned}$$

since $\omega(n)\to\infty$. This proves Theorem 3.2. □

For a graph H and any subset of vertices $U\subseteq V(H)$, define

$$S(U):=\bigcap_{v\in U}\mathrm{st}_{X(H)}(v).$$

Lemma 4.2. *Let $k\geq 1$ and suppose H be any graph such that every $2k+1$ vertices share a neighbor, and for every set of $2k$ vertices $U\subseteq H$, $S(U)$ is connected. Then $X(H)$ is k-connected.*

Proof of Lemma 4.2. As in the proof of Theorem 3.2, cover $X(H)$ by its vertex stars $\mathrm{st}(v)$ and apply Theorem 4.1. The nerve $\mathcal{N}(\mathrm{st}(v))$ is k-connected since every $2k+1$ vertices sharing a neighbor implies that its $2k$-skeleton is complete, so it is in fact $(2k-1)$-connected. Then to check that $X(H)$ is k-connected, it suffices to check that every t-wise intersection of vertex stars is $(k-t+1)$-connected, $2\leq t\leq k+1$. We show something slightly stronger, that if $0\leq j<k$ and $i\leq 2k-2j$, then every i-wise intersection of vertex stars is j-connected.

The case $j=0$ is clear: if $|U|=2k$ then $S(U)$ is connected by assumption, and if $|U|<2k$ then $S(U)$ is still connected, since every pair of vertices in $S(U)$ shares a neighbor. Let $j=1$. The claim is that if $i\leq 2k-2$ and $|U|=i$ then $S(U)$ is 1-connected. Cover $S(U)$ by vertex stars $\mathrm{st}_{S(U)}(v)$, $v\in S(U)$ and again apply Theorem 4.1. We only need to check that every intersection $\mathrm{st}_{S(U)}(v)\cap\mathrm{st}_{S(U)}(v)$ is connected, but this is clear since

$$\mathrm{st}_{S(U)}(u)\cap\mathrm{st}_{S(U)}(v)=S(U\cup\{u,v\})$$

is the intersection of $i+2\leq 2k$ vertex stars, connected by assumption.

Similarly, let $j=2$, $i\leq 2k-4$, and $|U|=i$. Then to show that $S(U)$ is 2-connected, cover by vertex stars $\mathrm{st}_{S(U)}(v)$. Each 3-wise intersection of vertex stars

$$\mathrm{st}_{S(U)}(u)\cap\mathrm{st}_{S(U)}(v)\cap\mathrm{st}_{S(U)}(w)=S(U\cup\{u,v,w\})$$

is the intersection of at most $i+3\leq 2k-1$ vertex stars, connected by assumption. Each 2-wise intersection of vertex stars in $S(U)$ is the intersection of at most $i+2\leq 2k-2$ vertex stars in $X(H)$, 1-connected by the above. Again applying Theorem 4.1, we have that $S(U)$ is 2-connected as desired.

Proceeding in this way, the lemma follows by induction on j. $\square$

Proof of Theorem 3.4. The remainder of this section is a proof that if

$$p = \left(\frac{(2k+1)\log n + \omega(n)}{n}\right)^{1/(2k+1)}$$

and $\omega(n) \to \infty$ then a.a. $X(n, p)$ is k-connected. Our argument is inspired by a proof of Theorem 1.1 in [3]. Since $k = 0$ is Theorem 1.1 we assume that $k \geq 1$, Observe that for any graph H and vertex subset $U \subseteq V(H)$, $S(U)$ is the clique complex of a subgraph of H. Moreover, for any vertex $v \in S(U)$, $\mathrm{st}_{S(U)}(v) = S(U \cup \{v\})$. We use these facts repeatedly.

By Lemma 4.2 and Theorem 3.2, we need only check that a.a. the intersection of every $2k$ vertex stars in $X(n, p)$ is connected. It is convenient to instead check that the intersection of every $2k$ vertex links is connected. For a vertex v in a simplicial complex Δ define the *link* of v in Δ by

$$\mathrm{lk}_\Delta(v) := \{\sigma | v \notin \sigma \text{ and } \{v\} \cup \sigma \in \Delta\},$$

and for any vertex set U denote

$$L(U) := \bigcap_{v \in U} \mathrm{lk}(v).$$

Suppose H is as in the hypothesis of Lemma 4.2 and $|U| = 2k$. If $L(U)$ is connected, then $S(U)$ is connected also, as follows. If $S(U) - L(U) = \emptyset$ we are done, so suppose $x \in S(U) - L(U)$. Clearly $x \in U$. $L(U)$ is connected by assumption, and in particular nonempty, so let $v \in L(U)$. For $u \in U - \{x\}$, $v \sim u$ and $x \sim u$. So $\{u, v\}$, $\{u, x\}$, and $\{v, x\}$ are all edges in H, and $\{u, v, x\}$ is a face in $X(H)$, and $\{v, x\} \in \mathrm{st}_{X(H)}(u)$. So $\{v, x\} \in S(U)$ and x is connected to $L(U)$. This holds for every $x \in S(U) - L(U)$, so $S(U)$ is connected.

Now we check that if

$$p = \left(\frac{(2k+1)\log n + \omega(n)}{n}\right)^{1/(2k+1)},$$

then a.a., for every subset $U \subseteq [n]$ with $|U| = 2k$, $L(U)$ is connected. It suffices to consider the 1-dimensional skeleton $L(U)^{(1)}$, which is a random graph with independent edges. However the number of vertices in the graph is not constant but a distribution, and there are $\binom{n}{2k}$ such graphs, where edges in one are not necessarily independent of edges in another. However, the edges within each graph are still independent, and we may still apply linearity of expectation to show that the probability that at least one of these graphs is not connected goes to 0.

Let $U \subseteq [n]$ be any vertex set of cardinality $2k$. The number of vertices X in $L(U)$ is not constant, but it is tightly concentrated. X is the sum of $n - 2k$ independent indicator random variables, each with probability p^{2k}. So we have an the following estimate for the mean of X.

$$\mu = E[X] \sim p^{2k} n$$

since k is constant. It is convenient to assume that $p = o(1)$. A similar argument works for dense random graphs.

Standard large deviation bounds [1] give that

$$\mathbf{P}(|X - \mu| > \epsilon\mu) < \mathrm{e}^{-c_\epsilon \mu}$$

for some constant $c_\epsilon > 0$ depending only on ϵ. We set $\epsilon = 1/100$ and write $c = c_\epsilon$. Then

$$\begin{aligned} \mathrm{e}^{-c\mu} &\leq \mathrm{e}^{-cp^{2k}n} \\ &= \mathrm{e}^{-cp^{-1}p^{2k+1}n} \\ &\leq \mathrm{e}^{-cp^{-1}(2k+1)\log n} \\ &\leq n^{-c(2k+1)p^{-1}} \\ &\leq n^{-c(2k+1)\omega(n)}, \end{aligned}$$

where $\omega(n) \to \infty$. So, applying a union bound, the total probability that for any set U, $|X - p^{2k}n| > (1/100)p^{2k}n$ is no more than

$$\binom{n}{2k} n^{-c(2k+1)\omega(n)} = o(1).$$

We have shown that a.a., $0.99p^{2k}n < X < 1.01p^{2k}n$ holds for every U, so we assume this for the remainder of the proof. Note that $X \to \infty$ by our assumption on p.

Let $\mathbf{P}_i$ denote the probability that there are components of order i in $L(U)$ for at least one $2k$-subset U. $\mathbf{P}_1 = o(1)$ by Theorem 3.2. Next we bound $\mathbf{P}_2$. There are $\binom{n}{2k}$ choices for U, then conditioned on that choice of U, let X denote the number of vertices in $L(U)$, as above. Given that $u, v \in L(U)$, the probability that $\{u, v\}$ spans a component of order 2 in $L(U)$ is $p(1-p)^{2(X-2)}$. There are $\binom{X}{2}$ choices for $\{u, v\}$ so by our assumptions on X,

$$\begin{aligned}
\mathbf{P}_2 &\le \binom{n}{2k}\binom{X}{2} p(1-p)^{2(X-2)} \\
&\le n^{2k}\binom{\lceil 1.01p^{2k}n\rceil}{2} p\mathrm{e}^{-2p(X-2)} \\
&\le n^{2k}p^{4k}n^2 p\mathrm{e}^{-2p(X-2)} \\
&\le n^{2k+2}p^{4k+1}\mathrm{e}^{-2pX(1-o(1))} \\
&\le n^{2k+2}p^{4k+1}\mathrm{e}^{-1.98p^{2k+1}n(1-o(1))} \\
&\le n^{2k+2}p^{4k+1}\mathrm{e}^{-1.98(2k+1)\log n(1-o(1))} \\
&\le n^{2k+2}p^{4k+1}n^{-1.98(2k+1)(1-o(1))} \\
&= o(n^{-1}),
\end{aligned}$$

since $k \ge 1$.

For any U and subset of i vertices $S \subseteq L(U)$, for S to span a connected component of order i, it must at least contain a spanning tree. It is well known that the number of spanning trees on i vertices is i^{i-2} [13]. The probability that all $i-1$ edges in any particular tree appear is p^{i-1}, by independence. We first bound P_i from above, assuming $3 \le i \le 100$. Since $X \to \infty$,

$$\begin{aligned}
\mathbf{P}_i &\le \binom{n}{2k}\binom{X}{i} i^{i-2}p^{i-1}(1-p)^{i(X-i)} \\
&\le n^{2k}\frac{X^i}{i!} i^{i-2}p^{i-1}\mathrm{e}^{-ipX(1-o(1))} \\
&\le c_i n^{2k}X^i p^{i-1}\mathrm{e}^{-i(0.99p^{2k+1}n(1-o(1)))} \\
&\le c_i n^{2k}(1.01p^{2k}n)^i p^{i-1}\mathrm{e}^{-0.99i(2k+1)\log n(1-o(1))} \\
&= c_i \exp[(2k+i)\log n + i\log 1.01 + (2ki+i-1)\log p - 0.99(1-o(1))i(2k+1)\log n] \\
&\le c_i \exp[(2k+i-0.99i(2k+1)+o(1))\log n] \\
&\le c_i \exp[(2k+0.01i-1.98ik+o(1))\log n] \\
&\le c_i \exp[(2k+0.01i-ik-0.98ik+o(1))\log n] \\
&\le c_i \exp[(-k-0.97i+o(1))\log n],
\end{aligned}$$

where $c_i = i^{i-2}/i!$ is a constant that only depends on i. (The last line holds because $i \ge 3$ and $k \ge 1$.) So for large enough n,

$$\begin{aligned}
\mathbf{P}_i &\le c_i \exp[(-k-0.97i+o(1))\log n] \\
&\le c_i n^{-k/2-.97i}
\end{aligned}$$

and

$$\sum_{i=3}^{100}\mathbf{P}_i \le \sum_{i=3}^{100} c_i n^{-k/2-.97i} = o(n^{-3}).$$

Now suppose $100 < i \le \lfloor 0.6p^{2k}n\rfloor$. Here we need to be a bit more careful in our treatment of the $i^{i-2}/i!$ factor. Stirling's formula gives that $i^{i-2}/i! \le \mathrm{e}^i$ though, and this will be good enough. We have

$$\begin{aligned}
\mathbf{P}_i &\le \binom{n}{2k}\binom{X}{i} i^{i-2} p^{i-1}(1-p)^{i(X-i)} \\
&\le n^{2k}\frac{X^i}{i!} i^{i-2} p^{i-1} \mathrm{e}^{-p(0.4iX)} \\
&\le n^{2k} X^i \mathrm{e}^i p^{i-1} \mathrm{e}^{-0.4ipX} \\
&\le n^{2k}(1.01p^{2k}n)^i \mathrm{e}^i p^{i-1} \mathrm{e}^{-0.4i(0.99p^{2k+1}n)} \\
&= \exp[(2k+i)\log n + i(1+\log 1.01) + (2ki+i-1)\log p - 0.396i(2k+1)\log n] \\
&\le \exp[(2k+(0.604+o(1))i - 0.792ik)\log n].
\end{aligned}$$

Then by assumption that $k \ge 1$, $i > 100$, and for large enough n,

$$\begin{aligned}
\mathbf{P}_i &\le \exp[(2k+0.605i-0.792ik)\log n] \\
&= \exp[(2k+0.605i-0.092ik-0.7ik)\log n] \\
&\le \exp[(2k+0.605i-9.2k-0.7i)\log n] \\
&= \exp[(-7.2k-.095i)\log n] \\
&= n^{-7.2k-.095i},
\end{aligned}$$

and

$$\sum_{i=101}^{\lfloor 0.6p^{2k}n\rfloor} \mathbf{P}_i \le \sum_{i=101}^{\infty} n^{-7.2k-.095i} = o(n^{-15}).$$

Putting it all together, a.a. each $L(U)$ is of order X, $0.99p^{2k}n < X < 1.01p^{2k}n$, and there are no components of order i, $1 \le i \le \lfloor 0.6p^{2k}n\rfloor$ in any of the $L(U)$. We conclude that each $L(U)$ is connected, as desired. □

5. Vanishing homology

We show if $p = n^\alpha$ with $\alpha < -1/k$ then $\widetilde{H}_k(X(G(n,p),\mathbb{Z})) = 0$ almost always. In this section, we assume the reader is familiar with simplicial homology [6]. For a k-chain C, the *support*, supp(C), is the union of k-faces in C with nonzero coefficients. Similarly, the *vertex support*, vsupp(C), is the underlying vertex set of the support.

A pure k-dimensional subcomplex Δ is said to be *strongly connected* if every pair of k-faces $\sigma, \tau \in \Delta^d$ can be connected by a sequence of facets $\sigma = \sigma_o, \sigma_1, \sigma_2, \ldots, \sigma_j = \tau$ such that $\dim(\sigma_i \cap \sigma_{i+1}) = d-1$ for $0 \le i \le n-1$. Every k-cycle is a $\mathbb{Z}$-linear combination of k-cycles with strongly connected support. We show first that all strongly connected subcomplexes are supported on a bounded number of vertices, and then that all small cycles are boundaries.

Lemma 5.1. *Let $\alpha < -1/k$ and $0 < 1/N < -1/k-\alpha$. Then, there are a.a. no strongly connected pure k-dimensional subcomplexes of $X(n,p)$ with vertex support of more than $N+k+1$ vertices.*

Proof of Lemma 5.1. The vertices in the support of a strongly connected subcomplex can be ordered $v_1, v_2, \ldots, v_n$ such that $\{v_1, \ldots, v_{k+1}\}$ spans a k-face and v_i is connected to at least k vertices v_j with $j < i$. One way to see this is to order the k-faces $f_1, f_2, f_3, \ldots$, so that each has $(k-1)$-dimensional intersection with the union of the previous faces. This is possible because we have assumed that the subcomplex is strongly connected. Then let this ordering induce an ordering on vertices, since at most one new vertex gets added at a time in the sequence $f_1, f_1 \cup f_2, f_1 \cup f_2 \cup f_3 \ldots$

Suppose Δ has $N+k+1$ vertices. There are at least $\binom{k+1}{2} + Nk$ edges in Δ by the above. If the underlying graph of Δ is not a subgraph of $G(n,p)$ then Δ is not a subcomplex. Choose ϵ and N such that $1/N < \epsilon < -\alpha - 1/k$. We apply a union bound on the total probability that there are any subcomplexes isomorphic to Δ in $X(n,p)$. We have $p = n^\alpha < n^{-(1/k+\epsilon)}$ and $k < \epsilon N k$ by assumption, so

$$\begin{aligned}
\mathbf{P}(\exists \text{ subcomplex}) &\le (N+k+1)!\binom{n}{N+k+1} p^{\binom{k+1}{2}+Nk} \\
&\le (N+k+1)!\binom{n}{N+k+1} n^{-(1/k+\epsilon)\left(\binom{k+1}{2}+Nk\right)}
\end{aligned}$$

$$\le n^{N+k+1} n^{-(1/k)\left(\binom{k+1}{2}+Nk\right)} n^{-\epsilon\left(\binom{k+1}{2}+Nk\right)}$$
$$\le n^{1-(k+1)/2-\epsilon\binom{k+1}{2}}$$
$$= O(n^{-\epsilon}).$$

This last line holds since $k \geq 1$. There are only finitely many isomorphism types of strongly connected k-dimensional complexes Δ on $N+k+1$ vertices, and a.a. none of them are subcomplexes of $X(n, p)$ by repeating this argument for each of them. There are also no such subcomplexes on more than $N+k+1$ vertices, since each of these contains a strongly connected subcomplex on exactly $N+k+1$ vertices (e.g., by the ordering described above). □

Then homology is generated by cycles supported on small vertex sets. Let γ be a nontrivial k-cycle in a simplicial complex Δ, with minimal vertex support, and write it is a linear combination of faces

$$\gamma = \sum_{f \in \mathrm{supp}(\gamma)} \lambda_f f,$$

with $\lambda_f \in \mathbb{Z}$. For the remainder of this section we restrict our attention to the full induced subcomplex of Δ on $\mathrm{vsupp}(\gamma)$. Clearly γ is still a nontrivial cycle in this subcomplex. For $v \in \mathrm{vsupp}(\gamma)$ define the k-chain

$$\gamma \cap \mathrm{st}(v) := \sum_{f \in \mathrm{st}(v)} \lambda_f f,$$

and the $(k-1)$-chain

$$\gamma \cap \mathrm{lk}(v) := \sum_{f \in \mathrm{st}(v)} \lambda_f (f - \{v\}).$$

Order the vertices with v last and let this induce an orientation on every face. We observe that

$$\gamma \cap \mathrm{lk}(v) = \partial(\gamma \cap \mathrm{st}(v)),$$

and since $\partial \circ \partial = 0$ this gives that $\gamma \cap \mathrm{lk}(v)$ is a $(k-1)$-cycle.

Lemma 5.2. *With notation as above, $\gamma \cap lk(v)$ is a nontrivial $(k-1)$-cycle in $lk(v)$.*

Proof of Lemma 5.2. We need only check that $\gamma \cap \mathrm{lk}(v)$ is not a boundary. Suppose by way of contradiction that $\partial(\beta) = \gamma \cap \mathrm{lk}(v)$ for some k-chain β with $\mathrm{supp}(\beta) \subseteq \mathrm{lk}(v)$. In particular $v \notin \mathrm{vsupp}(\beta)$. Write

$$\beta = \sum_{f \in \mathrm{supp}(B)} \mu_f f$$

with $\mu_f \in \mathbb{Z}$ and define the $(k+1)$-chain

$$\beta * \{v\} := \sum_{f \in \mathrm{supp}(\beta)} \mu_f (f \cup \{v\}).$$

Then

$$\partial(\beta * \{v\}) = \gamma \cap \mathrm{st}(v) + (-1)^{k+2}\beta.$$

So

$$\gamma' := (\gamma - \gamma \cap \mathrm{st}(v)) + (-1)^{k+3}\beta$$

is a k-cycle homologous to γ, but with $\mathrm{vsupp}(\gamma') \subseteq \mathrm{vsupp}(\gamma) - \{v\}$, contradicting that γ has minimal vertex support. □

Lemma 5.3. *Let H be a graph and $X(H)$ its clique complex. Suppose γ is a nontrivial k-cycle in $X(H)$. Then* $|\mathrm{vsupp}(\gamma)| \geq 2k+2$.

Proof of Lemma 5.3. Proceed by induction on k. The claim is clear when $k = 0$. Suppose then that $|\mathrm{vsupp}(\gamma)| \leq 2k+1$, and $v \in \mathrm{vsupp}(\gamma)$. By Lemma 5.2, $\gamma \cap \mathrm{lk}(v)$ is a nontrivial cycle. By the induction hypothesis, $|\mathrm{vsupp}(\gamma \cap \mathrm{lk}(v))| \geq 2k$, so we must have equalities $|\mathrm{vsupp}(\gamma)| = 2k+1$ and $|\mathrm{vsupp}(\gamma \cap \mathrm{lk}(v))| = 2k$. Repeating this argument gives that every vertex in $\text{vsupp}(\gamma)$ has degree $2k$, so $\text{vsupp}(\gamma)$ spans a clique in H. But then $\text{vsupp}(\gamma)$ spans a $2k$-dimensional face in $X(H)$, a contradiction to γ nontrivial. $\square$

Proof of Theorem 3.6. Any nontrivial k-cycle with minimal vertex support must have minimum vertex degree at least $2k$ in its supporting subgraph, since each vertex link is a nontrivial $(k-1)$-cycle by Lemma 5.2, hence contains at least $2(k-1)+2=2k$ vertices by Lemma 5.3. (We discuss nontrivial k-cycles S^k with $|\text{vsupp}(S^k)| = 2k+2$ in Section 6.)

Let H be any fixed graph with minimal vertex degree $2k$. Let $m = |V(H)|$, and then $|E(H)| \geq m(2k)/2 = mk$. Then if $\alpha < -1/k$ and $p = n^\alpha$, H is a.a. not a subgraph of $G(n,p)$. We check this with a union bound. The probability that H is a subgraph is at most

$$\begin{aligned} m!\binom{n}{m}p^{mk} &\leq n^m n^{\alpha mk} \\ &= o(1), \end{aligned}$$

since $\alpha k < -1$. There are only finitely many isomorphism types of graphs of minimal degree $2k$ on $m = N+k$ vertices. Each has at least km edges. Applying this argument to each of them, we conclude $X(n,p)$ a.a. has no vertex minimal nontrivial k-cycles, so a.a. $\widetilde{H}_k(X(G(n,p),\mathbb{Z})) = 0$. $\square$

6. Spherical retracts

We prove Theorem 3.5, that if $p^{k+1}n \to 0$ and $p^k n \to \infty$ as $n \to \infty$ then $X(n,p)$ a.a. retracts onto a sphere $\mathbb{S}^k$.

Let S^d denote the d-dimensional octahedral sphere (i.e. the d-fold repeated join of two isolated points), and $(S^d)^{(1)}$ its 1-skeleton. An alternate description of $(S^d)^{(1)}$ as a graph is

$$V((S^d)^{(1)}) = \{u_1, u_2, \ldots, u_{d+1}\} \cup \{v_1, v_2, \ldots, v_{d+1}\}$$

and

$$E((S^d)^{(1)}) = -\{\{u_i, v_j\} \mid i = j\},$$

where the – denotes complement in the set of all possible edges. Hence $(S^k)^{(1)}$ has $2(k+1)$ vertices and $\binom{2(k+1)}{2} - (k+1)$ edges.

$(S^k)^{(1)}$ is a *strictly balanced* graph, meaning that the ratio of edges to vertices is strictly smaller for every proper subgraph. A standard result in random graph theory [3] gives that $n^{-2(k+1)/\left(\binom{2(k+1)}{2}-(k+1)\right)} = n^{-1/k}$ is a sharp threshold function for $G(n,p)$ containing a $(S^k)^{(1)}$ subgraph. In particular, if $p^k n \to \infty$, $G(n,p)$ a.a. contains such a subgraph.

With notation as above, let $S = \{u_1, u_2, \ldots, u_{k+1}\} \cup \{v_1, v_2, \ldots, v_{k+1}\}$ be the vertices of such a subgraph. The conditional probability that vertices $\{u_1, u_2, \ldots, u_{k+1}\}$ have a common neighbor is no more than

$$(k+1)p + (n-2k-2)p^{k+1} = o(1),$$

since $p^k n \to 0$ (so $p \to 0$) and $p^{k+1}n \to 0$. So a.a. $G(n,p)$ contains a $(S^k)^{(1)}$ subgraph S such that $\{u_1, u_2, \ldots, u_{k+1}\}$ has no common neighbor. Note that in this case u_i is never adjacent to v_i for any choice of i. Then define a retraction of $X(n,p)$ onto $X(S)$ by defining a map $r : G(n,p) \to S$ on vertices and extending simplicially. (In particular, the $(S^k)^1$ subgraph is induced.)

For $x \in S$, set $r(x) = x$ and for $x \notin S$, let i be chosen so that x is not adjacent to u_i and set $r(x) = u_i$. Such a choice exists for every $x \notin S$ almost always, by the above. There's no obstruction to extending r simplicially to a retraction $\tilde{r} : X(n,p) \to X(S)$, and $X(S)$ is homeomorphic to $\mathbb{S}^k$.

7. Betti numbers

First assume that $p^{k+1}n \to 0$ and $p^k n \to \infty$. We wish to prove Theorem 3.8 and in particular to show that a.a. $\beta_k \sim f_k$. For every simplicial complex δ, we have the Morse inequality [6]:

$$-f_{k-1} + f_k - f_{k+1} \leq \beta_k \leq f_k.$$

The point is that when p is in this interval, f_k is much larger than $f_{k-1} + f_{k+1}$.

By linearity of expectation, we have

$$-E[f_{k-1}]+E[f_k]-E[f_{k+1}]\le E[\beta_k]\le E[f_k], \tag{1}$$

and then expanding each term gives

$$-\binom{n}{k}p^{\binom{k}{2}}+\binom{n}{k+1}p^{\binom{k+1}{2}}-\binom{n}{k+2}p^{\binom{k+2}{2}}\le E[\beta_k]\le\binom{n}{k+1}p^{\binom{k+1}{2}}.$$

Since $p^{k+1}n\to 0$ and $p^k n\to\infty$, we also have that $\binom{n}{k}p^{\binom{k}{2}}\to 0$ and $\binom{n}{k+2}p^{\binom{k+2}{2}}\to 0$. Let $Y_k=-f_{k-1}+f_k-f_{k+1}$. Then, we have shown so far that

$$E[Y_k]\sim E[\beta_k]\sim E[f_k]. \tag{2}$$

We strengthen this by applying the Second Moment Method. A standard application of Chebyshev's inequality [1] gives that if $E[X]\to\infty$ and $\mathrm{Var}[X]=o(E[X]^2)$ then a.a. $X\sim E[X]$. To prove Theorem 3.5 it suffices to check that $\mathrm{Var}[f_k]=o(E[f_k]^2)$ and $\mathrm{Var}[Y_k]=o(E[Y_k]^2)$.

Let $\mu=E[f_k]$ and we have

$$\mu^2=\binom{n}{k+1}^2p^{2\binom{k+1}{2}}$$

and

$$\mathrm{Var}[f_k]=E[f_k^2]-\mu^2.$$

Label the $(k+1)$-subsets of $[n]$, $1,2,\ldots,\binom{n}{k+1}$. Let A_i be the event that subset i spans a k-face in $X(n,p)$, and $A_i\wedge A_j$ the event that both A_i and A_j occur. Then

$$\begin{aligned}E[f_k^2]&=\sum_{i=1}^{\binom{n}{k+1}}\sum_{j=1}^{\binom{n}{k+1}}\Pr[A_i\wedge A_j]\\&=\binom{n}{k+1}\sum_{j=1}^{\binom{n}{k+1}}\Pr[A_1\wedge A_j],\end{aligned}$$

by symmetry. By grouping together A_j by the size of their intersections with A_1 we have

$$\begin{aligned}E[f_k^2]&=\binom{n}{k+1}\sum_{m=0}^{k+1}\binom{k+1}{m}\binom{n-k-1}{k+1-m}p^{2\binom{k+1}{2}-\binom{m}{2}}\\&=\binom{n}{k+1}p^{2\binom{k+1}{2}}\sum_{m=0}^{k+1}\binom{k+1}{m}\binom{n-k-1}{k+1-m}p^{-\binom{m}{2}}\\&\le\mu^2+\binom{n}{k+1}p^{2\binom{k+1}{2}}\sum_{m=1}^{k+1}\binom{k+1}{m}\binom{n-k-1}{k+1-m}p^{-\binom{m}{2}},\end{aligned}$$

since $\binom{n-k-1}{k+1}\le\binom{n}{k+1}$. Then we have that

$$\begin{aligned}\frac{E[f_k^2]-\mu^2}{\mu^2}&\le\frac{\sum_{m=1}^{k+1}\binom{k+1}{m}\binom{n-k-1}{k+1-m}p^{-\binom{m}{2}}}{\binom{n}{k+1}}\\&=\sum_{m=1}^{k+1}O(n^{-m}p^{-\binom{m}{2}})\end{aligned}$$

$$= \sum_{m=1}^{k+1} O((n^{-1}p^{-(m-1)/2})^m)$$
$$= o(1),$$

since $n^{-1}p^{-k-1} = o(1)$ by assumption. We conclude that a.a. $f_k \sim E[f_k]$. This did not depend on any assumption about k, so we also have that a.a. $-f_{k-1} \sim E(-f_{k-1})$ and $-f_{k+1} \sim E(-f_{k+1})$, and adding these three gives that a.a. $Y_k \sim E(Y_k)$.

By Eqs. (1) and (2), a.a. $\beta_k \sim E(\beta_k)$. The conclusion is that a.a. $\beta_k \sim f_k$, so this completes the proof of Theorem 3.8.

Now we use discrete Morse Theory to prove Theorem 3.9, that if $p^{k+1}n \to \infty$ or $p^k n \to 0$ as $n \to \infty$ then $E[\beta_k]/E[f_k] = o(1)$. For this, a few definitions are in order. We will write $\sigma < \tau$ if σ is a face of τ of codimension 1.

Definition 7.1. A discrete vector field V of a simplicial complex Δ is a collection of pairs of faces of $\Delta\{\alpha < \beta\}$ such that each face is in at most one pair.

Given a discrete vector field V, a *closed V-path* is a sequence of faces

$$\alpha_0 < \beta_0 > \alpha_1 < \beta_1 > \ldots < \beta_n > \alpha_{n+1},$$

such that $\{\alpha_i < \beta_i\} \in V$ for $i = 0, \ldots, n$ and $\alpha_{n+1} = \alpha_o$. (Note that $\{\beta_i > \alpha_{i+1}\} \notin V$ since each face is in at most one pair.) We say that V is a *discrete gradient vector field* if there are no closed V-paths.

Call any simplex not in any pair in V *critical*. The main theorem of discrete Morse Theory is the following [5].

Theorem 7.2 (*Forman*). *Suppose Δ is a simplicial complex with a discrete gradient vector field V. Then Δ is homotopy equivalent to a CW complex with one cell of dimension k for each critical k-dimensional simplex.*

First assume that $p^{k+1}n \to \infty$. Since we are assuming the vertex set of $G(n, p)$ is labeled by $[n]$, we can let this induce a total ordering of the vertices. This induces a lexicographic ordering on the faces of $X(n, p)$. For two faces σ and τ of a simplicial complex, we write $\sigma <_{lex} \tau$ if σ comes before τ in the lexicographic ordering. For any set of faces S let lexmin(S) denote the lexicograhically first element of S.

Define a discrete gradient vector field on $X(n, p)$ as follows.

$$V := \{\{\alpha < \beta\} | \dim(\alpha) = k \text{ and } \beta = \text{lexmin}(\{b | \alpha < b \text{ and } \alpha <_{lex} b\})\}.$$

It is clear that no face is in more than one pair, and there are no closed V-paths. Let $\sigma := \{v_1, v_2, \ldots, v_{k+1}\} \subset [n]$, with the vertices listed in increasing order, and set $m := v_{k+1}$. Then σ is a critical k-dimensional face of $X(n, p)$ if and only if $\sigma \in X(n, p)$ and $\sigma \cup \{x\} \notin X(n, p)$ for every $x >_{lex} m$. These events are independent by independence of edges in $G(n, p)$. So

$$\mathbf{P}(\sigma \text{ is a critical } k - \text{face}) = p^{\binom{k+1}{2}}(1 - p^{k+1})^{n-m}.$$

There are $\binom{i-1}{k}$ possible choices for σ with $v_{k+1} = i$. Let the number of critical k-faces be denoted by $\tilde{f}_k$. We have

$$E(\tilde{f}_k) = \sum_{i=m+1}^{n} \binom{i-1}{k} p^{\binom{k+1}{2}}(1 - p^{k+1})^{n-i}$$
$$\le \binom{n}{k} p^{\binom{k+1}{2}} \sum_{i=m+1}^{n} (1 - p^{k+1})^{n-i}$$
$$\le \binom{n}{k} p^{\binom{k+1}{2}} \sum_{i=-\infty}^{n} (1 - p^{k+1})^{n-i}$$
$$= \binom{n}{k} p^{\binom{k+1}{2}} \frac{1}{p^{k+1}},$$

so

$$
\begin{aligned}
\frac{E(\tilde{f}_k)}{E(f_k)} &\leq \frac{\binom{n}{k} p^{\binom{k+1}{2}} \frac{1}{p^{k+1}}}{\binom{n}{k+1} p^{\binom{k+1}{2}}} \\
&= O\left(\frac{1}{np^{k+1}}\right) \\
&= o(1),
\end{aligned}
$$

since $np^{k+1} \to \infty$. By Theorem 7.2, $X(n, p)$ is homotopy equivalent to a CW complex with at most $\tilde{f}_k$ faces, and by cellular homology, $\beta_k \leq \tilde{f}_k$ [6]. So $E(\beta_k)/E(f_k) \to 0$ and this proves the first part of Theorem 3.9.

Now assume $np^k \to 0$. For each k-face $\tau = \{v_1, v_2, \ldots, v_{k+1}\}$ choose $i(\tau) \in \{1, 2, \ldots, k+1\}$ randomly, uniformly and independently. We'd like to set

$$V = \{\{\tau - v_{i(\tau)}, \tau\} | \tau \in X(n, p) \text{ and } \dim(\tau) = k\},$$

but this might not be a discrete gradient vector field. There are two things that might go wrong. Some $(k-1)$-faces might be in more than one pair, and there might be closed V-paths. If we remove one pair from V for each such bad event though, we are left with a proper discrete gradient vector field, with at most one critical cell for each bad event. So we compute the expected number of bad events.

Each bad event contains at least one pair of k-faces of $X(n, p)$ meeting in a $(k-1)$-face, either resulting in a $(k-1)$-face being in more than one pair, or a closed V-path. Let d denote the number of such pairs, which is also the number of pairs of K_{k+1} subgraphs in $G(n, p)$ which intersect in exactly k vertices. In such a situation there are $k+2$ vertices and at least $\binom{k+2}{2} - 1$ edges total, and given a set of $k+2$ vertices there are $\binom{k+2}{2}$ possible choices for a pair of K_{k+1} intersecting in k vertices, so

$$
\begin{aligned}
E(d) &= \binom{k+2}{2}\binom{n}{k+2} p^{\binom{k+2}{2}-1} \\
&= \binom{k+2}{2}\binom{n}{k+2} p^{\frac{k(k+3)}{2}}.
\end{aligned}
$$

Then

$$
\begin{aligned}
\frac{E(d)}{E(f_k)} &= \frac{\binom{k+2}{2}\binom{n}{k+2} p^{\frac{k(k+3)}{2}}}{\binom{n}{k+1} p^{\frac{k(k+1)}{2}}} \\
&= O(np^k) \\
&= o(1),
\end{aligned}
$$

since $np^k \to 0$ by assumption. Again, by Theorem 7.2 and cellular homology, $\beta_k \leq d$, so this completes the proof of Theorem 3.9.

8. Random simplicial complexes

$X(n, p)$ seems to us a natural probability space of simplicial complexes to study topologically, in part because every simplicial complex is homeomorphic to a clique complex, e.g. by barycentric subdivision [2]. But of course there are many other possible definitions of random simplicial complexes.

Linial and Meshulam give a definition for random 2-complexes $Y(n, p)$ which "locally" look like $G(n, p)$, and exhibited a sharp $\mathbb{Z}_2$-homological analogue of Theorem 1.1 [9]. This was subsequently generalized to d-dimensional complexes and arbitrary fixed finite coefficients by Meshulam and Wallach [11]. In [7], it is shown that the threshold for vanishing of $\pi_1(Y(n, p))$ is much larger than the Linial-Meshulam-Wallach threshold for $H_1(Y(n, p), \mathbb{Z}_2)$.

Pippenger and Schleich study a different sort of random 2-complexes, made by gluing edges of triangles together randomly [12]. Their 2-complexes are pseudomanifolds, and the main motivation is giving quantitative results about fluctuations in the topology of spacetime in theories of quantum gravity.

In another article [8], we study the *neighborhood complex* of a random graph $\mathcal{N}[G(n, p)]$. The results are comparable to what we find here: each fixed homology group is roughly unimodal in p, and the nontrivial homology of a random d-complex is concentrated in a small number of dimensions. Applications are discussed to topological bounds on chromatic number.

9. Future directions

Although Theorem 3.4 is technically a generalization of one direction of Theorem 1.1, it is not clear if it is best possible and we are of the opinion that it probably is not. We conjecture that Theorem 3.5 is tight instead, and that if $p = n^{\alpha}$ with $\alpha > -1/(k + 1)$ then $X(n, p)$ is k-connected, almost always. (The work in [7] suggests that this may not be true when $k = 1$, but we believe that it probably holds for $k > 1$.)

In a sense, this would be close to determining the homotopy type of $X(n, p)$ when $-1/k < \alpha < -1/(k + 1)$. In particular, if one could establish this conjecture, and also show that $\widetilde{H}_k(X(n, p), \mathbb{Z})$ is torsion free, then standard results in combinatorial homotopy theory [2] (Theorem 9.18) would imply that if $p = n^{\alpha}$ with $-1/k < \alpha < -1/(k + 1)$ then $X(n, p)$ is a.a. homotopy equivalent to a wedge of k-dimensional spheres. However, note that even showing that $\widetilde{H}_k(X(n, p), \mathbb{Z})$ is free of m-torsion for every fixed m would not be good enough, since it is still possible that there is m-torsion, with m tending to infinity along with n.

Many simplicial complexes arising in combinatorics are homotopy equivalent to wedges of spheres, and Robin Forman, among others, has asked if there is any good reason why [5]. Such complexes frequently arise as order complexes of posets, hence naturally arise as clique complexes, and we believe that the results in this article are a step toward answering this question.

Acknowledgements

The author thanks his thesis advisors Eric Babson and Chris Hoffman for many illuminating conversations, without which this work would not have been possible. Also, thanks to Tristram Bogart, Anton Dochtermann, Dmitry Kozlov, Nati Linial, and Roy Meshulam, for insight and many helpful comments. The author gratefully acknowledges support by NSA grant #H98230-05-1-0053, the University of Washington's NSF-VIGRE grant, and MSRI's program in Computational Applications of Algebraic Topology.

References

[1] N. Alon, J. Spencer, The Probabilistic Method, second edition, John Wiley, 2000.
[2] A. Björner, Topological methods, in: Handbook of Combinatorics, vol. 2, 1996, pp. 1819–1872.
[3] B. Bollobás, Random Graphs, second edition, Cambridge University Press, 2001.
[4] P. Erdős, A. Rényi, On random graphs I, Publ. Math. Debrecen 6 (1959) 290–297.
[5] R. Forman, A user's guide to discrete Morse theory, in: Proc. of the 2001 Internat. Conf. on Formal Power Series and Algebraic Combinatorics, in: A special volume of Advances in Applied Mathematics, 2001.
[6] A. Hatcher, Algebraic Topology, Cambridge University Press, 2002.
[7] E. Babson, C. Hoffman, M. Kahle, Simple connectivity of random 2-complexes (submitted for publication), arXiv:math.CO/0711.2704.
[8] M. Kahle, The neighborhood complex of a random graph, J. Combin. Theory Ser. A 114 (2007) 380–387.
[9] N. Linial, R. Meshulam, Homological connectivity of random 2-complexes, Combinatorica 26 (2006) 475–487.
[10] R. Meshulam, The clique complex and hypergraph matching, Combinatorica 21 (2001) 89–94.
[11] R. Meshulam, N. Wallach, Homological connectivity of random k-dimensional complexes (submitted for publication), arXiv:math.CO/0609773.
[12] N. Pippenger, K. Schleich, Topological characteristics of random triangulated surfaces, Random Structures Algorithms 28 (2006) 247–288.
[13] Richard Stanley, Enumerative Combinatorics, Vol. 2, Cambridge University Press, 1999.