SHARP VANISHING THRESHOLDS FOR COHOMOLOGY OF RANDOM FLAG COMPLEXES

MATTHEW KAHLE

ABSTRACT. We exhibit a sharp threshold for vanishing of rational cohomology in random flag complexes, providing a generalization of the Erdős–Rényi theorem. As a corollary, almost all d-dimensional flag complexes have nontrivial (rational, reduced) homology only in middle degree $\lfloor d/2 \rfloor$.

1. Introduction

1.1. **Overview.** The edge-independent random graph $G(n,p)$ is a fundamental example in probability and combinatorics. Here n is the number of vertices, and p is the probability of each edge appearing. The notation $G \in G(n,p)$ means that G is a graph chosen according to the distribution $G(n,p)$.

Erdős and Rényi showed in 1959 that $p = \log n/n$ is the threshold for the property of connectedness [9].

Theorem 1.1 (Erdős – Rényi)**.** *Let $\epsilon > 0$ be fixed, and $G \in G(n,p)$.*

(1) *If*
$$p \geq \frac{(1+\epsilon)\log n}{n},$$
then
$$\mathbb{P}[G \text{ is connected}] \to 1,$$
(2) *and if*
$$p \leq \frac{(1-\epsilon)\log n}{n},$$
then
$$\mathbb{P}[G \text{ is connected}] \to 0,$$
as $n \to \infty$.

(The Erdős–Rényi Theorem is actually slightly sharper than this — see for example Chapter 7 of [6].)

Our main result is a generalization of Theorem 1.1 to higher-dimensional random simplicial complexes.

A *flag simplicial complex* or simply *flag complex* is a simplicial complex which is maximal with respect to its underlying graph. This is also sometimes called a *clique complex* since the faces of the simplicial complex correspond to complete subgraphs of the graph. For a graph H, let $X(H)$ denote the associated flag complex. Throughout the article we blur the distinction between an abstract simplicial complex Δ and its geometric realization $|\Delta|$.

Our main object of study is the flag complex of an edge-independent random graph, which we denote by $X \in X(n,p)$. Taking the geometric realization of X

puts a measure on a wide range of topologies — indeed, every simplicial complex is homeomorphic to a flag complex, e.g. by barycentric subdivision. The following is a rough statement of our main result, which provides a generalization of Theorem 1.1, the analogous $k = 0$ case.

Theorem 1.2. *Let $k \geq 1$ and $\epsilon > 0$ be fixed, and $X \in X(n, p)$.*

(1) *If*
$$p \geq \left(\frac{(k/2 + 1 + \epsilon) \log n}{n} \right)^{1/(k+1)},$$
then
$$\mathbb{P}[H^k(X, \mathbb{Q}) = 0] \to 1,$$
(2) *and if*
$$n^{-1/k+\epsilon} \leq p \leq \left(\frac{(k/2 + 1 - \epsilon) \log n}{n} \right)^{1/(k+1)},$$
then
$$\mathbb{P}[H^k(X, \mathbb{Q}) = 0] \to 0,$$
as $n \to \infty$.

By universal coefficients for homology and cohomology, $H^k(X, \mathbb{Q})$ is isomorphic to $H_k(X, \mathbb{Q})$, so these results may be interpreted for rational homology instead.

One complication is that for $k \geq 1$ the vanishing of $H^k(X, \mathbb{Q})$ is not a monotone property. Non-monotonicity was already observed in [17], where a number of facts were proved about the expected topology of $X \in X(n, p)$. In particular, a range for $p = p(n)$ was given in which $H^k(X, \mathbb{Q})$ is nontrivial with high probability. We use "with high probability" or "w.h.p." throughout the article to mean that the probability approaches 1 as $n \to \infty$.

Together with earlier results [17], one corollary is the following. For fixed d, if p is in the right regime then the flag complex is d-dimensional with high probability.

Roughly speaking, if $d \geq 1$ is fixed, and
$$n^{-2/d} \ll p \ll n^{-2/(d+1)},$$
then with high probability

(1) $X \in X(n, p)$ is d-dimensional, and
(2) $\widetilde{H}_i(X, \mathbb{Q}) = 0$ unless $i = \lfloor d/2 \rfloor$.

(Here we are using "$\ll$" loosely to mean "much less than," omitting factors which are only logarithmic in n — a precise statement is given in the next section.)

So according to this measure, almost all d-dimensional flag complexes have all their (rational, reduced) homology in middle degree.

This corollary may be viewed as given a measure-theoretic explanation of the fact that so many simplicial complexes and posets arising in combinatorics have homology concentrated in a small number of degrees. Indeed, many complexes are known to be homotopy equivalent to a wedge of spheres of equal dimension, and at the moment we can not rule out the possibility that almost all d-dimensional flag complexes are homotopy equivalent to a wedge of $\lfloor d/2 \rfloor$-spheres, at least for $d \geq 6$. We discuss this question in more detail in Section 7.

2. Statement of results

A word on notation: Throughout, we use Bachmann–Landau and related notations. This includes the standard big-O and little-o, as well as big-Ω, little-ω notations. The function $f = \Omega(g)$ if and only if $g = O(f)$, and $f = \omega(g)$ if and only $g = o(f)$. Asymptotics in this article are always as the number of vertices $n \to \infty$. In particular $\omega(1)$ is any function that tends to ∞ as $n \to \infty$.

The following is our main result. (Note that is a stronger version of Theorem 1.2.)

Theorem 2.1. *Let $X \in X(n,p)$. For every $k \geq 1$ there exists a constant $C_k > 0$ depending only on k, such that the following holds.*

(1) *If*

$$p \geq \left(\frac{(k/2+1)\log n + C_k \sqrt{\log n} \log\log n}{n} \right)^{1/(k+1)},$$

then

$$\mathbb{P}[H^k(X,\mathbb{Q}) = 0] \to 1,$$

(2) *and if*

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

then

$$\mathbb{P}[H^k(X,\mathbb{Q}) = 0] \to 0,$$

as $n \to \infty$.

So for all $k \geq 0$ there is an interval of p for which $H^k(X,\mathbb{Q})$ is nontrivial w.h.p. — for $k = 0$ this interval is only bounded above, and for $k \geq 1$ it is bounded above and below. The exponent in the lower bound of Part (2) of Theorem 2.1 is best possible by Theorem 3.6 in [17].

As a corollary, as long as $p = O(n^{-\epsilon})$ for an arbitrary fixed $\epsilon > 0$, $X \in X(n,p)$ w.h.p. has at most two nontrivial homology groups and in many cases only has one.

The proof of Theorem 2.1 is based on earlier work in cohomology of buildings by Garland [12], and by Ballman and Świątkowski [4]. See also work of Żuk [23] and Hoffman, Kahle, and Paquette [15] on random groups, where a similar method was earlier applied in probabilistic settings.

Together with earlier results on random flag complexes, and applying universal coefficients for homology and cohomology, one corollary is that many d-dimensional random flag complexes have all their (rational, reduced) homology in middle degree.

Corollary 2.2. *Let $d \geq 1$ and $\epsilon > 0$ be fixed. If*

$$\left(\frac{(d/4+1)\log n + (d/4+\epsilon)\sqrt{\log n}\log\log n}{n} \right)^{2/d} \leq p \leq o\left(n^{-2/(d+1)-\epsilon}\right),$$

then w.h.p. $X \in X(n,p)$ is d-dimensional, and

$$\widetilde{H}_i(X,\mathbb{Q}) = 0 \text{ unless } i = \lfloor d/2 \rfloor.$$

In Section 3 we prove lemmas for maximal k-cliques in random graphs which will be used in later sections. In Section 4 we prove Part (1) of Theorem 2.1, and in Section 5 we prove Part (2). In Section 6 we prove Corollary 2.2, and in Section 7 we close with comments and conjectures.

3. Preliminary calculations for maximal $(k+1)$-cliques

Let N_{k+1} denote the number of *maximal* $(k+1)$-cliques, i.e. $(k+1)$-cliques which are not contained in any $(k+2)$-cliques. It is useful to think of N_{k+1} as a sum of $\binom{n}{k+1}$ indicator random variables, as follows. For $i \in \binom{[n]}{k+1}$ let A_i be the event that the vertex set corresponding to i spans a maximal $(k+1)$-clique, and let Y_i be the indicator random variable for the event A_i. Then

$$N_{k+1} = \sum_{i \in \binom{[n]}{k+1}} Y_i.$$

Since the probability that i spans a $(k+1)$-clique is $p^{\binom{k+1}{2}}$, and the probability of the independent event that the vertices in i have no common neighbor is $(1 - p^{k+1})^{n-k-1}$, we have

$$E[Y_i] = p^{\binom{k+1}{2}}(1 - p^{k+1})^{n-k-1}.$$

By linearity of expectation we have

$$E[N_{k+1}] = \binom{n}{k+1} p^{\binom{k+1}{2}}(1 - p^{k+1})^{n-k-1}.$$

So roughly speaking, if $p \approx n^{-\alpha}$ with $2/k < \alpha < 1/(k+1)$ then $E[N_{k+1}] \to \infty$.

For a more refined estimate at the upper end of this interval, set

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

where $c \in \mathbb{R}$ is constant, and in this case we have

$$\begin{aligned}
E[N_{k+1}] &= \sum_{i \in \binom{[n]}{k+1}} E[Y_i] \\
&= \binom{n}{k+1} p^{\binom{k+1}{2}}(1 - p^{k+1})^{n-k-1} \\
&\approx \frac{n^{k+1}}{(k+1)!} p^{\binom{k+1}{2}} e^{-p^{k+1} n} \\
&= \frac{n^{k+1}}{(k+1)!} \left(\frac{(k/2+1+o(1))\log n}{n} \right)^{k/2} n^{-(k/2+1)} (\log n)^{-k/2} e^{-c},
\end{aligned}$$

and then

$$E[N_{k+1}] \to \frac{(k/2+1)^{k/2}}{(k+1)!} e^{-c}, \tag{1}$$

as $n \to \infty$.

3.1. **Zero expectation.** Letting $c \to \infty$ in Equation (1) gives that $E[N_{k+1}] \to 0$. By Markov's inequality, we conclude the following.

Lemma 3.1. *Let $G \in G(n,p)$, and N_{k+1} count the number of maximal $(k+1)$-cliques in G. If*

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

then $N_{k+1} = 0$ w.h.p.

3.2. **Infinite expectation.** Now set

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

In this case we have that $E[N_{k+1}] \to \infty$. By Chebyshev's inequality, if we also have $\mathrm{Var}[N_{k+1}] = o\left(E[N_{k+1}]^2\right)$, then

$$\mathbb{P}[N_{k+1} > 0] \to 1.$$

(See for example, Chapter 4 of [2].)

So once we bound the variance we have the following.

Lemma 3.2. *Let $0 < \epsilon < \frac{1}{k(k+1)}$ be fixed, and $G \in G(n,p)$. If*

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

then $N_{k+1} > 0$ w.h.p

As above, write N_{k+1} as a sum of indicator random variables.

$$N_{k+1} = \sum_{i \in \binom{[n]}{k+1}} Y_i.$$

Then

$$\mathrm{Var}[N_{k+1}] \leq E[N_{k+1}] + \sum_{i,j \in \binom{[n]}{k+1}} \mathrm{Cov}[Y_i, Y_j]$$

where the covariance is

$$\begin{aligned} \mathrm{Cov}[Y_i, Y_j] &= E[Y_i Y_j] - E[Y_i]E[Y_j] \\ &= \mathbb{P}[A_i \text{ and } A_j] - \mathbb{P}[A_i]\mathbb{P}[A_j], \end{aligned}$$

since Y_i are indicator random variables.

Let $I = I_{i,j} = |i \cap j|$ be the number of vertices in the intersection of subsets i and j. It is convenient to divide into cases depending on the cardinality of $0 \leq I < k+1$.

(1) **case:** $I = 0$. Given two disjoint subsets, $i, j \in \binom{[n]}{k+1}$,

$$\mathbb{P}[A_i \text{ and } A_j] = p^{2\binom{k+1}{2}} (1 - 2p^{k+1} + p^{2k+2})^{n-2k-2} \left(1 - O\left(p^k\right)\right),$$

and

$$\begin{aligned}\mathbb{P}[A_i]\mathbb{P}[A_j] &= \left(p^{\binom{k+1}{2}}(1-p^{k+1})^{n-k-1}\right)^2\\ &= p^{2\binom{k+1}{2}}\left(1-2p^{k+1}+p^{2k+2}\right)^{n-k-1},\\ &= p^{2\binom{k+1}{2}}\left(1-2p^{k+1}+p^{2k+2}\right)^{n-2k-2}\left(1-2p^{k+1}+p^{2k+2}\right)^{k+1},\\ &= p^{2\binom{k+1}{2}}\left(1-2p^{k+1}+p^{2k+2}\right)^{n-2k-2}\left(1-O\left(p^{(k+1)^2}\right)\right),\end{aligned}$$

so

$$\mathbb{P}[A_i \text{ and } A_j] - \mathbb{P}[A_i]\mathbb{P}[A_j] = p^{2\binom{k+1}{2}}(1-2p^{k+1}+p^{2k+2})^{n-2k-2}O\left(p^k\right).$$

The number of vertex-disjoint pairs i,j is $O\left(n^{2k+2}\right)$ so the total contribution S_0 to the variance of all the terms when $I=0$ is

$$S_0 = O\left(n^{2k+2}p^{2\binom{k+1}{2}}(1-2p^{k+1}+p^{2k+2})^{n-k-1}p^k\right)$$

Compare this to

$$E[N_{k+1}]^2 = \binom{n}{k+1}^2 p^{2\binom{k+1}{2}}(1-p^{k+1})^{2(n-k-1)}.$$

Clearly

$$S_0/E[N_{k+1}]^2 = O\left(p^k\right),$$

and since $p \to 0$ by assumption, we have that

$$S_0 = o\left(E[N_{k+1}]^2\right),$$

as desired.

(2) **case:** $I=1$. This case is similar. If $I=1$ then

$$\mathbb{P}[A_i \text{ and } A_j] = p^{2\binom{k+1}{2}}(1-2p^{k+1}+p^{2k+1})^{n-2k-1}(1-O(p^k)),$$

and

$$\begin{aligned}\mathbb{P}[A_i]\mathbb{P}[A_j] &= \left(p^{\binom{k+1}{2}}(1-p^{k+1})^{n-k-1}\right)^2\\ &= p^{2\binom{k+1}{2}}\left(1-2p^{k+1}+p^{2k+2}\right)^{n-k-1},\\ &= p^{2\binom{k+1}{2}}\left(1-2p^{k+1}+p^{2k+2}\right)^{n-2k-1}\left(1-2p^{k+1}+p^{2k+2}\right)^{k}\\ &= p^{2\binom{k+1}{2}}\left(1-2p^{k+1}+p^{2k+2}\right)^{n-2k-1}\left(1-O\left(p^{k(k+1)}\right)\right)\end{aligned}$$

So

$$\mathbb{P}[A_i \text{ and } A_j] - \mathbb{P}[A_i]\mathbb{P}[A_j] = p^{2\binom{k+1}{2}}(1-2p^{k+1}+p^{2k+2})^{n-2k-1}O\left(p^k\right).$$

There are $O\left(n^{2k+1}\right)$ such pairs of events, so

$$S_1 = O\left(n^{2k+1} p^{2\binom{k+1}{2}} (1-2p^{k+1}+p^{2k+2})^{n-2k-1} p^k\right).$$

Compare this to

$$E[N_{k+1}]^2 = \binom{n}{k+1}^2 p^{2\binom{k+1}{2}} (1-p^{k+1})^{2(n-k-1)}.$$

Now

$$S_1/E[N_{k+1}]^2 = O\left(n^{-1}p^k\right) = o(1),$$

since $n \to \infty$ and $p \to 0$. So we have that

$$S_1 = o\left(E[N_{k+1}]^2\right),$$

as desired.

(3) **case: $2 \le I \le k$.**
In this case,

$$\mathbb{P}[A_i \text{ and } A_j] = p^{2\binom{k+1}{2}-\binom{I}{2}} (1-2p^{k+1}+p^{2k+2-I})^{n-2k-2+I}(1-O(p^k)),$$

and

$$\begin{aligned}\mathbb{P}[A_i]\mathbb{P}[A_j] &= \left(p^{\binom{k+1}{2}}(1-p^{k+1})^{n-k-1}\right)^2 \\ &= p^{2\binom{k+1}{2}}(1-2p^{k+1}+p^{2k+2})^{n-k-1}.\end{aligned}$$

Comparing, we have

$$\begin{aligned}\frac{\mathbb{P}[A_i]\mathbb{P}[A_j]}{\mathbb{P}[A_i \text{ and } A_j]} &\le p^{\binom{I}{2}}\left(1+\frac{p^{2k+2}-p^{2k+2-I}}{1-2p^{k+1}+p^{2k+2-I}}\right)^n (1+o(1)) \\ &\le p^{\binom{I}{2}},\end{aligned}$$

and since $p \to 0$ and $I \ge 2$ by assumption,

$$\frac{\mathbb{P}[A_i]\mathbb{P}[A_j]}{\mathbb{P}[A_i \text{ and } A_j]} \to 0.$$

So

$$\mathbb{P}[A_i \text{ and } A_j] - \mathbb{P}[A_i]\mathbb{P}[A_j] = (1-o(1))\,\mathbb{P}[A_i \text{ and } A_j],$$

and now we bound the covariance

$$\mathrm{Cov}[Y_i, Y_j]$$

by bounding the probability $\mathbb{P}[A_i \text{ and } A_j]$.

For every $2 \le I < k+1$, there are $O\left(n^{2k+2-I}\right)$ pairs of events i, j with vertex intersection of cardinality I.

So the total contribution to variance from such pairs is at most

$$S_I = O\left(n^{2k+2-I} p^{2\binom{k+1}{2}-\binom{I}{2}} (1-2p^{k+1}+p^{2k+2-I})^{n-2k-2+I}\right).$$

Compare this to

$$E[N_{k+1}]^2 = \binom{n}{k+1}^2 p^{2\binom{k+1}{2}} (1-p^{k+1})^{2(n-k-1)}.$$

We have

$$S_I / E[N_{k+1}]^2 = O\left(n^{-I} p^{-\binom{I}{2}}\right).$$

Clearly

$$\begin{aligned} n^I p^{\binom{I}{2}} &= \left(np^{(I-1)/2}\right)^I \\ &\to \infty, \end{aligned}$$

as $n \to \infty$, since $I \leq k$ and $p = \omega(n^{-1/(k+1)})$. Hence

$$S_I = o\left(E[N_{k+1}]^2\right),$$

for $2 \leq I \leq k$.

3.3. **Finite expectation.** Using the "method of moments" the following can be shown. (See for example Section 6.1 of [16].)

Lemma 3.3. *If*

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

where $c \in \mathbb{R}$ is constant, then the number of maximal $(k+1)$-cliques N_{k+1} approaches a Poisson distribution

$$N_{k+1} \to Pois(\mu)$$

with mean

$$\mu = \frac{(k/2+1)^{k/2}}{(k+1)!} e^{-c}.$$

Since we do not use this Lemma anywhere, we state it without proof. However we record the combinatorial observation, for the sake of completeness, and also to give justification for a topological conjecture in Section 7.

4. Vanishing cohomology

In this section we aim to prove Part (1) of Theorem 2.1, so we assume that

$$p \geq \left(\frac{(k/2+1)\log n + C_k\sqrt{\log n}\log\log n}{n}\right)^{1/(k+1)},$$

where C_k is a constant depending only on k, to be chosen later.

For a finite graph H, let $C^0(H)$ denote the vector space of 0-forms on H, i.e. the vector space of functions $f : V(H) \to \mathbb{R}$. If all the vertex degrees are positive then the averaging operator A on $C^0(H)$ is defined by

$$Af(x) = \frac{1}{\deg x} \sum_{y \sim x} f(y),$$

where the sum is over all vertices y which are adjacent to vertex x. The identity operator on $C^0(H)$ is denoted by I. Then *the normalized graph Laplacian* $\mathcal{L} = \mathcal{L}(H)$ is a linear operator on $C^0(H)$ defined by $\mathcal{L} = I - A$.

The eigenvalues of $\mathcal{L}$ satisfy $0 = \lambda_1 \le \lambda_2 \le \cdots \le N \le 2$, where $N = |V(G)|$ is the number of vertices of H. Moreover, the multiplicity of the zero eigenvalue is equal to the number of connected components of H. In the case that H is connected then the smallest positive eigenvalue $\lambda_2[H]$ is sometimes called the *spectral gap* of H.

A simplicial complex Δ is said to be *pure D-dimensional* if every face of Δ is contained in a D-dimensional face. A special case of Theorem 2.1 in [4] is the following.

Theorem 4.1 (Ballman–Świątkowski). *Let Δ be a pure D-dimensional finite simplicial complex such that for every $(D-2)$-dimensional face σ, the link $lk_\Delta(\sigma)$ is connected and has spectral gap is at least $\lambda_2[lk_\Delta(\sigma)] > 1-1/D$. Then $H^{D-1}(\Delta, \mathbb{Q}) = 0$.*

For a simplicial complex Δ, the cohomology group $H^{D-1}(\Delta, \mathbb{Q})$ only depends on the D-skeleton of Δ. For us, $D = k + 1$. So to use Theorem 4.1 to show that $H^k(X, \mathbb{Q}) = 0$ we will show that given the hypothesis that edge probability p is large enough, with high probability

(1) the $(k + 1)$-skeleton of $X \in X(n, p)$ is pure dimensional, and
(2) for every $(k - 1)$-dimensional face $\sigma \in X$, the link $\mathrm{lk}_\Delta(\sigma)$ is connected and has spectral gap $\lambda_2[\mathrm{lk}_\Delta(\sigma)] > 1 - 1/k$.

4.1. **Pure-dimensional.** Let p be as above. We wish to check that w.h.p. the $(k + 1)$-skeleton of $X \in X(n, p)$ is w.h.p. pure $(k + 1)$-dimensional; in other words, that every face is contained in a $(k + 1)$-face.

Every k-face is contained in a $(k + 1)$-face, as follows. A k-face not contained in a $(k + 1)$-face would correspond to a maximal $(k + 1)$-clique. But by Lemma 3.1, for p in this regime the probability that there are any such cliques is tending to zero as $n \to \infty$.

The argument that for $0 \le i < k$ w.h.p. every i-dimensional face is contained in an $(i + 1)$-dimensional face is identical.

4.2. **Connectedness and spectral gap.** Finally we have to check that w.h.p. the link of every $(k - 1)$-dimensional face in the $(k + 1)$-skeleton is connected and has sufficiently large spectral gap. We require the following recent result for spectral gaps of Erdős–Rényi random graphs from [15].

Theorem 4.2. *Let $G \in G(n, p)$ be an Erdős-Rényi random graph. Let $\mathcal{L}$ denote the normalized Laplacian of G, and let $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$ be the eigenvalues of $\mathcal{L}$. For every fixed $\alpha \ge 0$, there is a constant $\widetilde{C}_\alpha$ depending only on α, so that if*

$$p \ge \frac{(\alpha + 1)\log n + \widetilde{C}_\alpha \sqrt{\log n} \log\log n}{n}$$

then G is connected and

$$\lambda_2(G) > 1 - o(1),$$

with probability $1 - o(n^{-\alpha})$.

To apply Theorem 4.1, we need to show that the link of every $(k-1)$-dimensional face has spectral gap larger than $1 - 1/k$ w.h.p. By standard concentration results, the number of $(k-1)$-dimensional faces is tightly concentrated around $\binom{n}{k}p^{\binom{k}{2}}$. The link of every $(k-1)$-face has approximately $(n-k)p^k$ vertices. Since k is fixed and $n \to \infty$, we will set $N = np^k$ and will treat every link of a $(k-1)$-dimensional face as a $G(N,p)$.

With foresight into the following calculation, we set

$$\alpha = k(k+3)/2.$$

We want to check first that

$$p \geq \frac{(\alpha+1)\log N + \widetilde{C}_\alpha \sqrt{\log N}\log\log N}{N}. \tag{2}$$

Since $\alpha + 1 = (k+1)(k/2+1)$ and $N \approx np^k$, this is equivalent to checking that

$$np^{k+1} \geq (k+1)(k/2+1)[\log n + k\log p] + \widetilde{C}_\alpha \sqrt{\log n}\,(\log\log n + O(1)) \tag{3}$$

We ignore the $O(1)$ term for now.

We consider n fixed and set

$$f(p) = np^{k+1} - (k+1)(k/2+1)[\log n + k\log p] + \widetilde{C}_\alpha \sqrt{\log n}\,(\log\log n + O(1))\,.$$

Then

$$f'(p) = (k+1)np^k - (k+1)(k/2+1)kp^{-1}.$$

Solving for $f'(p) = 0$ reveals only one critical point of the function f, at

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

Since

$$\lim_{p\to 0} f(p) = \infty,$$
$$\lim_{p\to\infty} f(p) = \infty,$$

and f is smooth on its domain $p \in (0,\infty)$, we conclude that this critical point must be a global minimum. In particular $f(p)$ is increasing on the interval

$$p \in \left[\left(\frac{k(k/2+1)}{n}\right)^{1/(k+1)}, 1\right].$$

So for sufficiently large n, to check that

$$p \geq \frac{(\alpha+1)\log N + \widetilde{C}_\alpha \sqrt{\log N}\log\log N}{N}$$

for

$$p \geq \left(\frac{(k/2+1)\log n + C_k\sqrt{\log n}\log\log n}{n}\right)^{1/(k+1)},$$

it suffices to check it for

$$p = \left(\frac{(k/2+1)\log n + C_k\sqrt{\log n}\log\log n}{n}\right)^{1/(k+1)}. \tag{4}$$

Then

$$\log p = \frac{1}{k+1}(\log\log n - \log n) + O(1). \tag{5}$$

Substitute the expressions for p and $\log p$ from (4) and (5) into (3) and subtract $(k/2+1)\log n$ from both sides to obtain

$$C_k\sqrt{\log n}\log\log n \geq \left(k(k/2+1)+\widetilde{C}_\alpha\right)\sqrt{\log n}\left(\log\log n + O(1)\right),$$

so as long as

$$C_k > k(k/2+1)+\widetilde{C}_\alpha$$

we have satisfied (2). Since $\alpha = k(k+3)/2$ and C_α only depends on α, C_k only depends on k.

By Theorem 4.2 we have that $G \in G(N,p)$ has spectral gap $\lambda_2[G] > 1-1/k$ with probability $1-o\left(N^{-\alpha}\right)$. The link of every $(k-1)$-dimensional face in the $(k+1)$-skeleton of $X \in X(n,p)$ is precisely such a random graph. (Here N is a random variable rather than a number, but we are treating it as a number for simplicity since it is tightly concentrated around its expectation.)

There are w.h.p. approximately $\binom{n}{k}p^{\binom{k}{2}}$ such $(k-1)$-dimensional faces. So applying a union bound, the probability P_f that the link of at least one $(k-1)$-dimensional face fails to have spectral gap $\lambda_2 > 1-1/k$ is bounded above by

$$\begin{aligned}
P_f &\leq \binom{n}{k}p^{\binom{k}{2}}N^{-\alpha}\\
&= \binom{n}{k}p^{\binom{k}{2}}(np^k)^{-k(k+3)/2}\\
&\leq \left(n^{k-k(k+3)/2}p^{\binom{k}{2}-k^2(k+3)/2}\right)\\
&= n^{-k(k+1)/2}p^{-k(k+1)^2/2}\\
&= \left(np^{k+1}\right)^{-k(k+1)/2},
\end{aligned}$$

and since $np^{k+1}\to\infty$ by assumption, we have $P_f \to 0$ as $n\to\infty$, as desired.

5. Non-vanishing cohomology

We prove Part (2) of Theorem 2.1. In particular we show that if $C_2 < k/2$ and $\epsilon > 0$ are fixed and

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

then w.h.p. $H^k(X,\mathbb{Q}) \neq 0$). The strategy is to show that in this regime there exist isolated k-faces which generate nontrivial cohomology classes — this is the higher-dimensional analogue of "isolated vertices" being the main obstruction to connectivity of the random graph $G(n,p)$; see for example Chapter 7 of [6].

First we show that if p is in the given regime, then w.h.p. there exist k-dimensional faces $\sigma \in X$ which are not contained in any $(k+1)$-dimensional faces — such faces generate cocycles in H^k (i.e. by considering the characteristic function of σ in $C^k(X)$). Then we show that if p is sufficiently large, then no k-dimensional face can be a coboundary. Putting these facts together, we find an interval of p for which there is at least one k-dimensional face which represents a nontrivial class in $H^k(X,\mathbb{Q})$.

For context, we note that two other approaches to showing that $H^k \neq 0$ for nearly the same regime of $p = p(n)$ are given in [17]. Both of these earlier approaches (finding embedded spheres which represent nontrivial classes, and a dimension argument) give the best possible exponents for the endpoints of the interval, but the approach here gives a more refined (and basically tight) estimate for the upper end of the interval of nontrivial homology. Since the upper end is our emphasis, we assume for convenience that $p = \omega\left(n^{-1/k+\epsilon}\right)$ — Theorem 3.8 in [17] extends this lower end of the nontrivial interval all the way to $p = \omega\left(n^{-1/k}\right)$.

5.1. **Cocycles.** Lemma 3.2 gives that for p in this regime, w.h.p. there are maximal $(k+1)$-cliques in $G \in G(n,p)$. But these represent isolated k-faces σ in $X \in X(n,p)$, and for such a σ the characteristic function of σ is a cocycle. The main point is to show that these are nontrivial — i.e. that they are not coboundaries.

5.2. **Non-coboundaries.** We have showed above that for p in the proper regime, there w.h.p. exist k-dimensional faces which are not contained in any $(k+1)$-dimensional face. Any such face generates a class in the vector space $C^k(X)$ of k-cocycles. Now we will show that in the same regime of p, w.h.p. no k-dimensional face represents a k-coboundary. Hence $H^k(X, \mathbb{Q}) \neq 0$.

Suppose that a k-dimensional face $\sigma \in X$ represents a k-coboundary, i.e. $\sigma = d\phi$ for some $(k-1)$-cochain ϕ. Then ϕ represents a nontrivial class in $H^{k-1}(X-\sigma)$. (This notation means X with the open face σ deleted). We claim that this extremely unlikely.

Lemma 5.1. *Fix $k \geq 1$ and $0 < \epsilon \leq 1/k$, and let $X \in X(n,p)$. If $p \geq n^{-1/k+\epsilon}$ then w.h.p. $H^{k-1}(X, \mathbb{Q}) = 0$, and the same holds for $X - \sigma$ for every k-dimensional face σ.*

Proof. The claim that $H^{k-1}(X, \mathbb{Q}) = 0$ is implied by Part (1) of Theorem 2.1 (with the index shifted by 1), proved in Section 4, so our focus is on the second part of the claim, that $H^{k-1}(X-\sigma, \mathbb{Q}) = 0$ for every k-dimensional face σ.

We apply Theorem 4.1 again. Since the proof here is so similar to what is in Section 4 we omit some details, and focus on what is new in this argument.

We may restrict our attention to the k-skeleton of X. Let σ be an arbitrary k-dimensional face of X.

Consider the link $\mathrm{lk}_{X-\sigma}(\tau)$ of an arbitrary $(k-2)$-dimensional face τ of $X - \sigma$. Since we are restricting to the k-skeleton, this is a graph. This graph is either equal to $\mathrm{lk}_X(\tau)$ exactly or to $\mathrm{lk}\tau(X)$ with a single edge deleted. Recall from Section 4 that $\mathrm{lk}_X(\Delta)$ is an Erdős-Rényi random graph $G(N,p)$, where $N = (n-k+1)p^{k-1}$.

We have control on the spectral gap of $\mathrm{lk}_X(\tau)$ by Theorem 4.2. From this we can control the spectral gap of $\mathrm{lk}_{X-\sigma}(\tau)$ by applying the Wielandt–Hoffman theorem.

Theorem 5.2 (Wielandt–Hoffman)**.** *Let A and B be normal matrices. Let their eigenvalues a_i and b_i be ordered such that $\sum_i |a_i - b_i|^2$ is minimized. Then we have*

$$\sum_i |a_i - b_i|^2 \leq \|A - B\|,$$

where $\|\cdot\|$ denotes the Frobenius matrix norm.

Here we have normalized Laplacians $A = \mathrm{lk}_X(\tau)$ and $B = \mathrm{lk}_{X-\sigma}(\tau)$ — since these matrices are symmetric, they are normal, and Theorem 5.2 applies. All eigenvalues of A and B are real, and putting them in increasing order minimizes the sum $\sum_i |a_i - b_i|^2$.

We have

$$\|A - B\| = \sqrt{\sum_i \sum_j |a_{ij} - b_{ij}|^2}.$$

In a normalized graph Laplacian,

$$a_{ij} = \frac{1}{\sqrt{\deg(v_i)\deg(v_j)}},$$

if v_i is adjacent to v_j, and $a_{ij} = 0$ otherwise.

The link of a $(k-2)$-face is a random graph conditioned on the vertices in the link, so standard results give that the degree of every vertex is exponentially concentrated around its mean $\approx np^k \geq n^{k\epsilon}$ (see Chapter 3 in[6]) and there are only polynomially many such vertices summed over all links. So w.h.p. every vertex in every link has degree $(1 + o(1))np^k \geq n^{k\epsilon}$. Then Theorem 5.2 gives that the Frobenius matrix norm of the normalized Laplacian can not shift by more than $O\left(n^{-k\epsilon}\right) = o(1)$ when an edge is deleted. Hence no single eigenvalue can shift by more than this.

Since we already have $\lambda_2[\mathrm{lk}_X(\tau)] > 1 - o(1)$ for every τ by Section 4.2, this gives that $\lambda_2[\mathrm{lk}_{X-\sigma}(\tau)] > 1 - o(1)$ for every τ and σ as well. Applying Theorem 4.1 again, we have that $H^{k-1}(X - \sigma, \mathbb{Q}) = 0$ for every k-dimensional face σ.

□

6. d-dimensional flag complexes for fixed d

Now we prove Corollary 2.2. We wish to show that if $d \geq 1$ and

$$\left(\frac{(1 + d/4)\log n + \omega(\sqrt{\log n}\log\log n)}{n}\right)^{2/d} \leq p \leq o\left(n^{-2/(d+1)-\epsilon}\right),$$

then w.h.p. $X \in X(n,p)$ is d-dimensional, and

$$\widetilde{H}_i(X, \mathbb{Q}) = 0 \text{ unless } i = \lfloor d/2 \rfloor.$$

If

$$p \leq o\left(n^{-2/(d+1)-\epsilon}\right),$$

then w.h.p. $\widetilde{H}_i(X, \mathbb{Q}) = 0$ for $i > \lfloor d/2 \rfloor$ by Theorem 3.6 in [17]. (This may even be true if

$$p \leq o\left(n^{-2/(d+1)}\right);$$

see for example a similar situation in [20].)

If

$$p \geq \left(\frac{(1 + d/4)\log n + \omega(\sqrt{\log n}\log\log n)}{n}\right)^{2/d}$$

then w.h.p. $\widetilde{H}_i(X, \mathbb{Q}) = 0$ for $i < \lfloor d/2 \rfloor$ by the proof of part (1) of Theorem 2.1 in Section 4.

That

$$\widetilde{H}_{\lfloor d/2 \rfloor}(X, \mathbb{Q}) \neq 0$$

for p in this regime follows from Theorem 3.8 in [17] — for some results on the limiting distribution of $\beta_{\lfloor d/2 \rfloor}$, see [18].

7. Comments

Besides the Erdős–Rényi Theorem, our main result here can be compared to earlier results of Linial and Meshulam [21] and of Meshulam and Wallach [22]. These earlier also exhibit sharp thresholds for cohomology to pass from non-vanishing to vanishing. The techniques in all these papers involve some kind of "expansion," whether combinatorial (i.e. $\mathbb{Z}/2$-coefficients) or spectral (i.e. $\mathbb{Q}$-coefficients). DeMarco, Hamm, and Kahn have parallel results to those here for cohomology of random flag complexes with $\mathbb{Z}/2$-coefficients, in the case $k = 1$ [8].

We use the word "sharp" in the title in the sense of Friedgut and Kalai [11], meaning that the phase transition happens in a narrow window. More precisely, we say for a monotone graph property $\mathcal{P}$ that f is a *sharp threshold* for $\mathcal{P}$ if there exists a function $g = o(f)$ such that $G \in G(n,p)$ has property $\mathcal{P}$ with probability $\to 1$ if $p \geq f + g$ and has $\mathcal{P}$ with probability $\to 0$ if $p \leq f - g$.

As commented before, the homological properties that we study here are not monotone. Nevertheless, a small modification of the above definition makes sense of our claim that non-vanishing of $H^k(X, \mathbb{Q})$ has a sharp *upper* threshold.

It is conceivable that Theorem 2.1 could be sharpened to the following.

Conjecture 7.1. *Let $k \geq 1$ be fixed. For $X \in X(n,p)$,*

(1) *if*

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

then

$$\mathbb{P}[H^k(X,\mathbb{Q}) = 0] \to 1,$$

(2) *and if*

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

then

$$\mathbb{P}[H^k(X,\mathbb{Q}) = 0] \to 0,$$

as $n \to \infty$.

Indeed, the following seems plausible.

Conjecture 7.2. *If*

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

where $c \in \mathbb{R}$ is constant, then the dimension of kth cohomology β^k approaches a Poisson distribution

$$\beta^k \to Pois(\mu)$$

with mean

$$\mu = \frac{(k/2+1)^{k/2}}{(k+1)!} e^{-c}.$$

In particular,

$$\mathbb{P}[H^k(X,\mathbb{Q}) = 0] \to \exp\left[-\frac{(k/2+1)^{k/2}}{(k+1)!} e^{-c} \right],$$

as $n \to \infty$.

Conjecture 7.2 should be compared with Lemma 3.3. The conjecture is that in this regime, characteristic functions on isolated k-faces generate cohomology with high probability. For some closely related work on limit theorems, see [18].

Many complexes in topological combinatorics are known to be homotopy equivalent to wedges of spheres [10, 5], and many others are known to have homology concentrated in a relatively small number of degrees [7]. The results here may be viewed as a measure-theoretic explanation of this seemingly ubiquitous phenomenon.

One attractive feature of the random flag complex model is that it puts a measure on a wide range of topologies — every simplicial complex is homeomorphic to a flag complex, i.e. by barycentric subdivision. If one could show that integral homology was torsion free w.h.p., then one would have the following.

Conjecture 7.3. *Let $d \geq 6$ and*

$$\left(\frac{(1 + d/4) \log n + \omega(\log \log n)}{n} \right)^{2/d} \leq p \leq o\left(n^{-2/(d+1)} \right).$$

Then w.h.p. $X \in X(n,p)$ is homotopy equivalent to a wedge of $\lfloor d/2 \rfloor$-dimensional spheres.

(The fact that torsion-free homology would imply this homotopy equivalence follows from "uniqueness of Moore spaces" – e.g. see example 4.34 in [14].)

Conjecture 7.3 should be compared with Corollary 2.2. The reason for the $d \geq 6$ is that this is sufficient to make $\pi_1(X)$ vanish with high probability, for example by Theorem 3.4 of [17], and there is reason to believe that this condition is also necessary [3].

My guess is that Conjecture 7.3 is close to the truth, but it is worth noting that certain types random complexes are known to have very large torsion groups on average [19].

This work can also be viewed in the context of higher-dimensional expanders; see for example the recent work of Gromov [13].

Acknowledgements

I thank Noga Alon, Eric Babson, Chris Hoffman, Roy Meshulam, Elliot Paquette, and Uli Wagner for helpful conversations. I learned of applications of Garland's method in topological combinatorics from [1], where a global analogue of Theorem 4.1 is developed.

I also thank the Institute for Advanced Study in Princeton, where some of this work was completed, and NSA grant # H98230-10-1-0227, for partial support.

References

[1] R. Aharoni, E. Berger, and R. Meshulam. Eigenvalues and homology of flag complexes and vector representations of graphs. *Geometric and functional analysis*, 15(3):555–566, 2005.

[2] Noga Alon and Joel H. Spencer. *The probabilistic method*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons Inc., Hoboken, NJ, third edition, 2008. With an appendix on the life and work of Paul Erdős.

[3] Eric Babson, Christopher Hoffman, and Matthew Kahle. The fundamental group of random 2-complexes. *J. Amer. Math. Soc.*, 24(1):1–28, 2011.

[4] W. Ballmann and J. Świątkowski. On L^2-cohomology and property (T) for automorphism groups of polyhedral cell complexes. *Geom. Funct. Anal.*, 7(4):615–645, 1997.

[5] A. Björner. Topological methods. In *Handbook of combinatorics, Vol. 1, 2*, pages 1819–1872. Elsevier, Amsterdam, 1995.
[6] Béla Bollobás. *Random graphs*, volume 73 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2001.
[7] S. Bouc. Homologie de certains ensembles de 2-sous-groupes des groupes symétriques. *Journal of Algebra*, 150(1):158–186, 1992.
[8] B. DeMarco and J. Kahn. Mantel's theorem for random graphs. *Arxiv preprint arXiv:1206.1016*, 2012.
[9] P. Erdős and A. Rényi. On random graphs. I. *Publ. Math. Debrecen*, 6:290–297, 1959.
[10] Robin Forman. A user's guide to discrete Morse theory. *Sém. Lothar. Combin.*, 48:Art. B48c, 35, 2002.
[11] Ehud Friedgut and Gil Kalai. Every monotone graph property has a sharp threshold. *Proc. Amer. Math. Soc.*, 124(10):2993–3002, 1996.
[12] Howard Garland. p-adic curvature and the cohomology of discrete subgroups of p-adic groups. *Ann. of Math. (2)*, 97:375–423, 1973.
[13] M. Gromov. Singularities, expanders and topology of maps. part 2: From combinatorics to topology via algebraic isoperimetry. *Geometric And Functional Analysis*, 20(2):416–526, 2010.
[14] Allen Hatcher. *Algebraic topology*. Cambridge University Press, Cambridge, 2002.
[15] Christopher Hoffman, Matthew Kahle, and Elliott Paquette. A sharp threshold for Kazhdan's Property (t). arXiv:1201.0425, submitted, 2012.
[16] Svante Janson, Tomasz Łuczak, and Andrzej Rucinski. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.
[17] Matthew Kahle. Topology of random clique complexes. *Discrete Math.*, 309(6):1658–1671, 2009.
[18] Matthew Kahle and Elizabeth Meckes. Limit theorems for Betti numbers of random simplicial complexes. arXiv:1009.4130, to appear in Homology, Homotopy and Applications, 2012.
[19] G. Kalai. Enumeration of Q-acyclic simplicial complexes. *Israel Journal of Mathematics*, 45(4):337–351, 1983.
[20] D.N. Kozlov. The threshold function for vanishing of the top homology group of random d-complexes. 138(12):4517–4527, 2010.
[21] Nathan Linial and Roy Meshulam. Homological connectivity of random 2-complexes. *Combinatorica*, 26(4):475–487, 2006.
[22] R. Meshulam and N. Wallach. Homological connectivity of random k-dimensional complexes. *Random Structures Algorithms*, 34(3):408–417, 2009.
[23] A. Żuk. Property (T) and Kazhdan constants for discrete groups. *Geom. Funct. Anal.*, 13(3):643–670, 2003.

E-mail address: Matthew Kahle <mkahle@math.ias.edu>

School of Mathematics, Institute for Advanced Study, Princeton NJ, 08540