INSIDE THE CRITICAL WINDOW FOR COHOMOLOGY OF RANDOM k-COMPLEXES

MATTHEW KAHLE AND BORIS PITTEL

Abstract. We prove sharper versions of theorems of Linial–Meshulam and Meshulam–Wallach which describe the behavior for $(\mathbb{Z}/2)$-cohomology of a random k-dimensional simplicial complex within a narrow transition window. In particular, we show that within this window the Betti number β^{k-1} is in the limit Poisson distributed. For $k = 2$ we also prove that in an accompanying growth process, with high probability, first cohomology vanishes exactly at the moment when the last isolated $(k-1)$-simplex gets covered by a k-simplex.

1. Introduction

In 1959 Erdős and Rényi pioneered a systematic study of a graph $G(n, m)$ chosen uniformly at random among all graphs on vertex set $[n] = \{1, 2, \dots, n\}$ with exactly m edges. They found the threshold value $\bar{m}$ for connectedness of $G(n, m)$ [6].

Here and throughout the paper *"with high probability (w.h.p)"* means that the probability of an event approaches 1 as the number of vertices $n \to \infty$.

Theorem 1.1 (Erdős–Rényi)**.** *If*

$$m = \frac{n}{2}(\log n + c),$$

where $c \in \mathbb{R}$ is constant, then w.h.p. $G(n, m)$ consists of a giant component and isolated vertices, and the number of isolated vertices converges in distribution to Poisson with mean e^{-c}. In particular

$$\mathbb{P}[G(n, m) \text{ is connected}] \to e^{-e^{-c}},$$

as $n \to \infty$. of $G(n, m)$.

Consequently, $\bar{m} = (n/2) \log n$ is a sharp threshold for connectedness, in the following sense.

Theorem 1.2. *Let $\omega \to \infty$ arbitrarily slowly. If*

$$m = \frac{n}{2}(\log n + \omega),$$

then w.h.p. $G(n, m)$ is connected, and if

$$m = \frac{n}{2}(\log n - \omega),$$

then w.h.p. $G(n, m)$ is disconnected.

The first author gratefully acknowledges DARPA grant # N66001-12-1-4226.
The second author gratefully acknowledges NSF grant # DMS-1101237.

In 1969 Stepanov [13] considered the Bernoulli counterpart $G(n,p)$, a random graph on $[n]$ such that a pair (i,j) forms an edge with probability p independently of all other pairs. He determined the threshold value $\bar{p}$ for connectedness of $G(n,p)$, and $\bar{m} \sim \bar{p}\binom{n}{2}$. Informally, this had to be expected because $G(n,m)$ is distributed as $G(n,p)$ *conditioned* on the number of edges being equal m, and for $\bar{p}$ the number of edges in $G(n,p)$ is sharply concentrated around its expected value, i.e. $\bar{p}\binom{n}{2}$.

Theorem 1.3 (Stepanov)**.** *If*

$$p = \frac{\log n + c}{n},$$

where $c \in \mathbb{R}$ is constant, then w.h.p. $G(n,p)$ consists of a giant component and isolated vertices, and the number of isolated vertices is asymptotic to Poisson with mean e^{-c}. In particular

$$\mathbb{P}[G(n,p) \text{ is connected}] \to e^{-e^{-c}},$$

as $n \to \infty$.

Consequently, $\bar{p} = \log n / n$ is a sharp threshold for connectedness.

Theorem 1.4. *Let $\omega \to \infty$ arbitrarily slowly. If*

$$p = \frac{\log n + \omega}{n},$$

then w.h.p. $G(n,p)$ is connected, and if

$$p = \frac{\log n - \omega}{n},$$

then w.h.p. $G(n,p)$ is disconnected.

Nowadays, Theorem 1.1 and Theorem 1.3 may be viewed as essentially equivalent, thanks to general "transfer" theorems, e.g. Janson, Łuczak and Ruciński [8], Propositions 1.12, 1.13.

An important advantage of the Erdős–Rényi random graph $G(n,m)$ is that it can be gainfully viewed as a snapshot of a natural random graph process $\{G(n,M)\}$, $0 \le M \le \binom{n}{2}$, at "time" $M = m$. Here $G(n,M)$ is obtained from $G(n,M-1)$ by selecting the location of M-th edge uniformly at random among all $\binom{n}{2} - (M-1)$ still available options. As a special case of a result of Bollobás and Thomasson [5], we have

Theorem 1.5. *For almost all realizations of the $\{G(n,M)\}$ process,*

$$\min\{M : \text{min degree of } G(n,M) > 0\} = \min\{M : G(n,M) \text{ is connected}\}.$$

In a seminal paper [10] Linial and Meshulam defined random 2-dimensional simplicial complexes and found a two-dimensional cohomological analogue of Theorem 1.4. Subsequently Meshulam and Wallach [12] managed to extend the result of [10] to all dimensions $k \ge 2$. These papers have inspired several other articles exploring the topology of random simplicial complexes, e.g. see Aronshtam et al [1], Babson et al. [2], Bollobás and Riordan [4], and Kozlov [9].

Our main goal in this article is to establish some k-dimensional analogues of Theorems 1.3 and 1.5, based on, or inspired by, the Linial–Meshulam and Meshulam–Wallach theorems.

1.1. **Topological preliminaries.** In this subsection we define simplicial complexes and simplicial cohomology with $(\mathbb{Z}/2)$-coefficients. For a more complete introduction we refer the reader to the first two chapters of Hatcher's book [7].

An abstract *simplicial complex* is a finite set V, called the *vertices* of S and a collection S of subsets of V such that

- $\{v\} \in S$ for every $v \in V$,
- if $A \in S$ and $B \subset A$ is nonempty then $B \in S$.

Elements $\{x, y\} \in S$ of cardinality 2 are sometimes called *edges*, and elements $\{x, y, z\}$ of cardinality 3 *triangles*. In general, elements of S are called *faces*.

The *dimension* of a face $f \in S$ is $|f| - 1$, where $|f|$ denotes the cardinality of f. (So vertices are 0-dimensional, edges are 1-dimensional, etc.) The dimension of S is the maximum cardinality of its faces. Note that a simplicial complex of dimension 1 corresponds to a simple graph (i.e. a graph with no loops or multiple edges).

One may also consider the *geometric realization* of S, sometimes denoted $|S|$, as a topological space. We will abuse notation and identify S with $|S|$. In practice, it is clear whether one is talking about a combinatorial feature of S or a topological feature.

Let $F_k = F_k(S)$ denote the set of k-dimensional faces of S. Then the *k-cochains* C^k is the vector space of functions $f : F_k \to \mathbb{Z}/2$. There is a *coboundary map* $d_k : C^k \to C^{k+1}$, defined by

$$d_k f(\sigma) = \sum_{\tau} f(\tau),$$

where the sum is over all faces $\tau \subset \sigma$ such that $\dim \tau = \dim \sigma - 1 = k$. If we introduce an $|F_k| \times |F_{k+1}|$ incidence matrix I_k such that $I_k(\tau, \sigma) =$ if and only if $\tau \subset \sigma$, and view $f := \{f(\tau)\}$, $g := \{d_k f(\sigma)\}$ as vectors, then $g^T = f^T I_k$.

The k-cocycles is defined to be the subspace $Z^k = \ker d_k$, i.e. the left null-space of I_k, and the k-coboundaries is the subspace $B^k = \operatorname{im} d_{k-1}$, i.e. the row space of I_{k-1}. For each $f \in B^k$ there exists $A \subset F_{k-1}$ such that f is supported by the faces $\beta \in F_k(S)$ with a property: β has an odd number of $(k-1)$-faces $\alpha \in F_{k-1}$. In particular, for $k = 2$, f is supported by the cut/set separating A and $A^c = [n] \setminus A$.

It is easy to verify that $B^k \subseteq Z^k$, i.e. $d_k \circ d_{k-1} = 0$ for every k, or equivalently $I_{k-1} I_k = 0$; indeed, given $\alpha \in F_{k-1}$, $\beta \in F_{k+1}$, α is either a face of exactly two k-faces of β, or not a face of any k-face of β.

Then the kth cohomology is defined to be the quotient vector space $H^k = Z^k / B^k$. (We might write $H^k(S, \mathbb{Z}/2)$ to emphasize that this is the cohomology for the simplicial complex S, and that we mean cohomology with $\mathbb{Z}/2$ coefficients.) We are especially interested in the case $H^k = 0$, which means that every k-cocycle is a k-coboundary.

If we view $\emptyset$ as (-1)-dimensional face, this is sometimes called *reduced* cohomology, and denoted by $\widetilde{H}^k$ rather than H^k. Note that $\widetilde{H}^k = H^k$ except in the case $k = 0$. In the reduced cohomology case I_0 is a single row with all entries 1, while I_1 is a vertex-edge incidence matrix of a simple graph G on $[n]$. Consequently $\widetilde{H}^0 = 0$ if and only if G is connected. In general, the dimension of $\widetilde{H}^0$ is $c(G) - 1$, where $c(G)$ is the number of connected components of G.

A topological aside: One may just as easily talk about *homology* H_k rather than cohomology H^k, but in the cases we are interested in the results would be exactly the same. (It is pointed out in [10] that this equivalence follows from universal coefficients.) However the main argument seems to be easier to make in terms of cohomology than in terms of homology.

1.2. **The Linial–Meshulam and Meshulam–Wallach theorems.** The random k-dimensional simplicial complex $Y \sim Y(n,p)$ has vertices $[n] = \{1, 2, \ldots, n\}$, and complete $(k-1)$-skeleton, meaning that Y contains all subsets of $[n]$ of cardinality k. Then each k-face (subset of cardinality $k+1$) is included in Y with probability p, independently of all other such faces. The following is a cohomological analogue of Theorem 1.1, $k = 2$ and $k \geq 2$ versions being proved by Linial and Meshulam [10] and by Meshulam and Wallach [12] respectively.

Theorem 1.6. *Let $\omega \to \infty$ arbitrarily slowly and $Y \sim Y(n,p)$.*

(1) *If*
$$p \geq \frac{k \log n + \omega}{n},$$
then
$$\mathbb{P}[H^{k-1}(Y, \mathbb{Z}/2) = 0] \to 1;$$
(2) *if*
$$p \leq \frac{k \log n - \omega}{n},$$
then
$$\mathbb{P}[H^{k-1}(Y, \mathbb{Z}/2) = 0] \to 0,$$
as $n \to \infty$.

(In [12] the same statement is shown to hold, even when $\mathbb{Z}/2$ is replaced by any finite abelian group of coefficients.)

Part (i) is the heart of Theorem 1.6, as part (ii) is relatively straightforward. Indeed, for $p \leq (k \log n - \omega)/n$, w.h.p. at least one $\alpha \in F_{k-1}$ is not a face of any k-face in $Y(n,p)$, i.e. α is *isolated*. The characteristic function f of such an α then is a cocycle, by default, but it is not a coboundary; indeed, for the k-faces $\sigma \notin F_k(Y(n,p))$, such that $\alpha \subset \sigma$,
$$\sum_{\tau \subset \sigma} f(\tau) = 1 \neq 0.$$
On the other hand, one sees that, for $p \geq n^{-1}(k \log n + \omega)$, w.h.p. there are no isolated $(k-1)$-faces. This suggests that isolated $(k-1)$-faces might hopefully be the *most* likely obstruction to cohomological connectedness of $Y(n,p)$ if $p \geq n^{-1}(k \log n + w)$. That was exactly the motivation behind the statement and the proof of the key part (i) in [10], [12]. In fact, we shall see that, in a closer analogy with $G(n,p)$, isolated $(k-1)$-faces are the *only* likely obstruction to such connectedness even "earlier", when $p = n^{-1}(k \log n + O(1))$.

1.3. **Notions of connectivity.** Linial and Meshulam introduced the terminology "(co)homological connectedness" to emphasize that Theorem 1.6 should be viewed as a 2-dimensional analogue of the Erdős–Rényi theorem. Spaces where every (co)cycle is a (co)boundary are also sometimes called "acyclic", or to have "vanishing (co)homology."

We call a k-dimensional simplicial complex S *hypergraph connected* if for every two $(k-1)$-faces α, $\alpha' \in F_{k-1}(S)$, there exists a sequence of $(k-1)$-faces
$$\alpha = \alpha_1, \alpha_2, \ldots, \alpha_k = \alpha'$$

that joins α and α', in a sense that for i, $\alpha_i \cup \alpha_{i+1} \in F_k(S)$.

Theorem 1.7. *Let S be a k-dimensional complex with complete $(k-1)$-skeleton (i.e. $|F_{k-1}(S)| = \binom{n}{k}$). If $H^{k-1}(S, \mathbb{Z}/2) = 0$, then S is hypergraph connected.*

Proof of Theorem 1.7. We use induction on k. The statement obviously holds for $k = 1$. Suppose it is true for some $k \geq 1$. Let S be a $(k+1)$-dimensional complex such that $H^k(S, \mathbb{Z}/2) = 0$. Define the link of a vertex $v \in [n]$ by

$$\mathrm{lk}_S(v) = \{\sigma - v : \sigma \in S,\, v \in \sigma\}.$$

Note that $\mathrm{lk}_S(v)$ is itself a simplicial complex, and

$$\dim(\mathrm{lk}_S(v)) \leq \dim(S) - 1.$$

Then $H^{k-1}(\mathrm{lk}_S(v)) = 0$ for all $v \in [n]$. If not, there is v^* such that $H^{k-1}(\mathrm{lk}_S(v^*)) \neq 0$. Thus there exists a cocycle $g : F_{k-1}(\mathrm{lk}_S(v^*)) \to \{0, 1\}$ which is not a coboundary, the latter meaning that for some k-face σ, that does not contain v^*,

$$\sum_{\tau \in \sigma} g(\tau) = 1.$$

Define $f : F_k(S) \to \{0, 1\}$ by the conditions

(a) $f((v^*, \tau)) = g(\tau)$ for $\tau \in F_{k-1}(\mathrm{lk}_S(v^*))$;

(b) $f(\alpha) = 0$ for all k-faces $\alpha \neq (v^*, \tau)$, with $\tau \in F_{k-1}(\mathrm{lk}_S(v^*))$.

Then f is a cocycle of S, but f is not a coboundary, because

$$\sum_{\alpha \in (v^*, \sigma)} f(\alpha) = \sum_{\tau \in \sigma} g(\tau) = 1.$$

Contradiction! So indeed $H^{k-1}(\mathrm{lk}_S(v)) = 0$ for all $v \in [n]$. By induction hypothesis, each $\mathrm{lk}_S(v)$ is hypergraph connected, so that in $\mathrm{lk}_S(v)$ every two $(k-1)$-faces are joined by a path of $(k-1)$-faces.

It remains to show that S itself is hypergraph connected. Let α, $\alpha' \in F_k(S)$. Define $t = |\alpha \cap \alpha'|$. If $t > 0$, then $\alpha = (v_1, v_2, \ldots, v_k)$, $\alpha' = (v_1, v_2', \ldots, v_k)$. Since $\mathrm{lk}_S(v_1)$ is hypergraph connected, $(v_2, \ldots, v_k)$ and $(v_2', \ldots, v_k')$ are joined by a path of $(k-1)$-faces in $\mathrm{lk}_S(v_1)$. Augmenting these intermediate faces with v_1 we obtain a path joining α and α' in S. Suppose $t = 0$, so that $(v_1, \ldots, v_k)$ and $(v_1', \ldots, v_k')$ do not overlap. By the preceding argument, $(v_1, v_2, \ldots, v_k)$ and $(v_1, v_2', \ldots, v_k')$ are joined by a path in S, and so are $(v_1, v_2', \ldots, v_k')$ and $(v_1', v_2', \ldots, v_k')$. Concatenating these two paths we get a path from $(v_1, \ldots, v_k)$ to $(v_1', \ldots, v_k')$ in S. □

In light of Theorem 1.7, Theorem 1.6 effectively shows that $\bar{p} = n^{-1}k \log n$ is both the threshold for $H_{k-1}(Y, \mathbb{Z}/2) = 0$ and the threshold for the hypergraph connectedness of the underlying hypergraph.

1.4. **Main results.** Our first result is a hypergraph analogue of Theorem 1.3. Let $k \geq 1$. Let $HG(n, p)$ denote the random hypergraph induced by the random complex $Y(n, p)$. The hypervertex set and the hyperedge set of $HG(n, p)$ are

$$F_{k-1} = \binom{[n]}{[k]}$$

and

$$F_k(Y(n, p)) \subseteq F_k = \binom{[n]}{[k+1]},$$

respectively.

Theorem 1.8. *If*
$$p = \frac{k \log n + c}{n},$$
where $c \in \mathbb{R}$ is constant, then w.h.p. $HG(n,p)$ consists of a giant component and isolated vertices, and the number of those is asymptotically Poisson, with mean $e^{-c}/k!$, and hence the probability of hypergraph connectedness approaches $\exp(e^{-c}/k!)$ as $n \to \infty$. Consequently, $p = n^{-1} k \log n$ is a sharp threshold probability for connectedness property of $HG(n,p)$.

A key estimate in an unexpectedly simple proof is obtained by using links and induction on k.

Analogously to $\{G(n,M)\}$, let us introduce the random complex process $\{Y(n,M)\}$, where $Y(n,M)$ is a uniformly random k-dimensional complex with M k-faces. Each $Y(n,M)$ is obtained from $Y(n,M-1)$ by choosing the location of M-th k-face uniformly at random among all $\binom{n}{k+1} - (M-1)$ possibilities. Here is a k-dimensional extension of Bollobás-Thomasson's Theorem 1.5.

Theorem 1.9. *For almost all realizations of the k-dimensional process $\{Y(n,M)\}$, the random k-face, that eliminates the chronologically last isolated $(k-1)$-face, also makes the resulting complex hypergraph connected.*

Furthermore we sharpen the Linial-Meshulam and Meshulam-Wallach theorems.

Theorem 1.10. *If*
$$p = \frac{k \log n + c}{n},$$
where $c \in \mathbb{R}$ is constant, then $\beta^{k-1} := \dim(H^{k-1}(Y, \mathbb{Z}/2))$ is asymptotically Poisson, with mean $e^{-c}/k!$. In particular, $H^{k-1}(Y, \mathbb{Z}/2)$ vanishes with limiting probability $\exp(-e^{-c}/k!)$.

Since $e^{-e^{-c}/k!} \to 0$ if $c \to -\infty$, and $e^{-e^{-c}/k!} \to 1$ if $c \to \infty$, Theorem 1.10 implies Theorem 1.6. We should also note that Theorem 1.10 in combination with Theorem 1.7, imply Theorem 1.3 as a direct byproduct.

For $k = 2$ we prove a cohomological extension of Theorem 1.5.

Theorem 1.11. *For almost all realizations of the 2-dimensional process $\{Y(n,M)\}$, the 2-face (triangle), that eliminates the chronologically last isolated 1-face (edge), also makes $H^1(Y, \mathbb{Z}/2) = 0$.*

Thus, while the three random moments,
$$M_1 =: \min\{m : Y(n,m) \text{ has no isolated edge}\},$$
$$M_2 =: \min\{m : Y(n,m) \text{ is hypergraph connected}\},$$
$$M_3 =: \min\{m : H^1(Y(n,m)) \text{ vanishes}\},$$
(always obeying $M_1 \le M_2 \le M_3$), may generally be distinct, the event $\{M_1 = M_2 = M_3\}$ has probability approaching 1 as $n \to \infty$.

A key part of our proof is based on counting non-trivial cocycles by the degree sequences of their supports, an approach considerably simpler than deep counting arguments in [10], [12]. We conjecture that the extension of Theorem 1.11 holds for all $k \geq 2$, and for cohomology with coefficients in any finite abelian group. Since the proofs of Theorems 1.8 and 1.9 are relatively simple, we wonder whether such an extension could be proved by using links and induction on k. It may well be possible also to get it done by a proper modification of the method in [12], but we haven't explored this route.

2. Proofs of Theorem 1.8 and Theorem 1.9.

(1) Let A_n be the event that all non-isolated $(k-1)$-faces of $Y \sim Y_k(n,p)$ belong to the same component, or equivalently that every two non-isolated $(k-1)$-faces are joined by a path in Y, in the sense of "hypergraph connected" described above.

Given a vertex $v \in [n]$, define the vertex link

$$\mathrm{lk}_Y(v) = \{\sigma - v : \sigma \in Y(n,p),\, v \in \sigma]\}.$$

So each $\mathrm{lk}_Y(v)$ is a $(k-1)$-dimensional complex, distributed as $Y_{k-1}(n-1,p)$. Of course, $\mathrm{lk}_Y(v)$, $v \in [n]$, are interdependent. Let $A_n(v)$ be the the event that every two non-isolated $(k-2)$-faces of $\mathrm{lk}_Y(v)$ are joined by a path in $\mathrm{lk}_Y(v)$. Let B_n be the event that for every α, $\alpha' \in F_{k-1}(Y(n;p))$, with $\alpha \cap \alpha' = \emptyset$, there exist $v \in \alpha$ and $v' \in \alpha'$ such that $(\alpha' \setminus \{v'\}) \cup \{v\}$ is a non-isolated $(k-1)$-face of $Y(n,p)$. Then

$$A_n \supseteq \left(\bigcap_{v \in [n]} A_n(v)\right) \bigcap B_n. \tag{2.1}$$

Indeed, suppose that the RHS event in (2.1) holds. Let $\alpha, \alpha' \in F_{k-1}(Y(n,p))$ be non-isolated. If there is $v \in \alpha \cap \alpha'$ then $\alpha \setminus \{v\}$ and $\alpha' \setminus \{v\}$ are non-isolated $(k-2)$-faces in $\mathrm{lk}_Y(v)$, whence they are joined by a path in $\mathrm{lk}_Y(v)$. By the definition of $\mathrm{lk}_Y(v)$, augmenting the edges of this path with v, we get a path joining α and α' in $Y(n,p)$. Suppose that $\alpha \cap \alpha' = \emptyset$. Then, by the definition of B_n, there exist $v \in \alpha$ and $v' \in \alpha'$ such that $\alpha'' := (\alpha' \setminus \{v'\}) \cup \{v\}$ is non-isolated. By the first part, α and α'' are joined by a path in $Y(n,p)$, and likewise so are α'' and α'.

Let $g_k(n;p) = \mathrm{P}(A_n^c)$, and $h_k(n;p) = \mathrm{P}(B_n^c)$. Then (2.1) implies a recurrence inequality

$$g_t(n;p) \leq n g_{t-1}(n-1;p) + h_t(n;p), \quad t \geq 2. \tag{2.2}$$

Let us bound $h_t(n;p)$. We observe that the number of of ordered pairs of disjoint α, $\alpha' \in F_{t-1}(Y(n,p)$ is less than $\binom{n}{t}^2$, and the number of pairs (v,v'), $v \in \alpha$, $v' \in \alpha$, is t^2. The probability that for every such pair (v,v') there does not exist $u \in [n] \setminus (\alpha \cup \alpha')$ such $(\alpha' \setminus \{v'\}) \cup \{v\} \cup \{u\}$ is in $F_t(Y(n,p))$ is $q^{(n-2t)t^2}$. Therefore, for $t \leq k$,

$$h_t(n;p) \leq \binom{n}{t}^2 q^{(n-2t)t^2} \leq a n^{2t} q^{nt^2}, \quad a := \frac{e^{2k^2}}{k!}.$$

So (2.2) simplifies to

$$g_t(n;p) \leq n g_{t-1}(n-1;p) + a n^{2t} q^{nt^2}. \tag{2.3}$$

Since $q^n = O(n^{-k})$, an easy induction shows that

$$g_t(n;p) \leq n^{t-1} g_1(n-t+1;p) + 2a n^{t+2} q^{4n}, \quad 2 \leq t \leq k. \tag{2.4}$$

Consider (2.4) for $t = k$. If

$$p = \frac{k\log n + c(n)}{n}, \quad |c(n)| = o(\log n), \tag{2.5}$$

then

$$n^{k+2}q^{4k} = O(n^{k+2}n^{-4k}e^{4|c(n)|}) = O(n^{-2k+2}e^{4|c(n)|}) \to 0,$$

as $k \ge 2$. Furthermore, $g_1(n-k+1;p)$ is the probability that $G(n-k+1;p)$ has a component of size from 2 to $(n-k+1)/2$, which—for p in question—is bounded by twice the expected number of components of size 2. And this expected value is of order

$$n^2p(1-p)^{2n} \le n^2pe^{-2np} = O\left(\frac{e^{2|c(n)|}\log n}{n^{2k-1}}\right).$$

So the first term in the RHS of (2.4) is of order $O(n^{-k}e^{2|c(n)|}\log n)$. In summary,

$$\mathbb{P}(A_n^c) = O(n^{-k}e^{2|c(n)|}\log n).$$

Thus, under condition (2.5),

$$\mathbb{P}(A_n) = 1 - O(n^{-k}e^{2|c(n)|}\log n) \to 1. \tag{2.6}$$

It remains to show that, for $p = (k\log n + c)/n$, X_n the total number of isolated $(k-1)$-faces of $Y(n,p)$ is asymptotically Poisson, with mean $e^{-c}/k!$. This is done by a standard argument based on factorial moments. So $Y(n,p)$ is connected with the limiting probability $e^{-e^{-c}/k!}$. The proof of Theorem 1.3 is complete. □

Note. If $c(n) = -\log\log n$, say, then the second order moment method shows that

$$\frac{X_n}{(\log n)^{1/k!}} \to 1, \quad \text{in probability}. \tag{2.7}$$

(2) Let us prove Theorem 1.9. We embed the random complex process $\{Y(n,M)\}$ into a continuous-time Markov process $\{Y_t(n)\}$, $t \ge 0$. To do so, we introduce i.i.d. random variables T_σ, $\sigma \in F_k$, with $\mathbb{P}(T_\sigma \le t) = 1 - e^{-t}$. T_σ can be interpreted as a waiting time till "birth" of the k-face σ. We define

$$Y_t(n) = \{\sigma \in F_k : T_\sigma \le t\}.$$

So $Y_t(n)$ is a complex whose k-faces have been born up to time t; $Y_0(n)$ is the complete $(k-1)$-dimensional complex, and $Y_\infty(n)$ is the complete k-dimensional complex. Clearly, $Y_t(n)$ is a Bernoulli complex $Y(n,p)$ with $p = p(t)$. Also, $\{Y_t(n)\}_{t\ge 0}$ is a Markov process, thanks to memoryless property of the exponential distribution. Introduce a sequence $\{t(M)\}$ of stopping times such

$$t(M) = \min\{t \ge 0 : |\{\sigma : T_\sigma \le t\}| = M\};$$

in words, $t(M)$ is the first time t the number of k-faces reaches M. Then (1) each $Y_{t(M)}(n)$ is distributed uniformly on the set of all complexes with M k-faces, and (2) conditioned on $Y_{t(M-1)}(n)$, the location of M-th k-face in $Y_{t(M)}(n)$ is distributed uniformly on the set of all $\binom{n}{k+1} - (M-1)$ available locations. Thus $\{Y_{t(M)}(n)\}$ is distributed as $\{Y(n,M)\}$.

Introduce

$$p_1 = \frac{k\log n - \log\log n}{n}, \quad p_2 = \frac{k\log n + \log\log n}{n},$$

and t_i defined by $p_i = 1 - e^{-t_i}$. Since $Y_t(n)$ is distributed as $Y(n, p(t))$, it follows from (2.7) that w.h.p. $Y_{t_1}(n)$ consists of $X_n \sim (\log n)^{1/k!}$ isolated $(k-1)$-faces and a single component on the remaining $\left[\binom{n}{k} - X_n\right]$ $(k-1)$-faces. As for $Y_{t_2}(n)$, w.h.p. it consists of a single component. Let τ be the first time t when the number of isolated $(k-1)$-faces drops down by two or more. Then

$$\begin{aligned}\mathbb{P}(\tau \le t_2 \mid Y_{t_1}(n)) &\le X_n^2 \int_{t_1}^{t_2} e^{-t}\,dt \\ &\le X_n^2 (t_2 - t_1) = O\big(n^{-1}(\log n)^{2/k!} \log\log n\big) \to 0.\end{aligned}$$

Here X_n^2 is a crude upper bound for the number of pairs of $(k-1)$-faces, isolated at time t_1, that happen to be the faces of the same k-simplex. And e^{-t} is the probability density of the birth time for such a k-simplex.

So, w.h.p. throughout $[t_1, t_2]$ the complex $Y_t(n)$ continues to be a giant component plus a set of isolated $(k-1)$-faces, gradually swallowed, one such face at a time, by the current giant component. Thus, w.h.p. $Y_t(n)$ becomes connected when the last isolated $(k-1)$-face gets joined by a newly born k-simplex to the current giant component. Consequently, the same property holds for the subprocess $\{Y_{t(M)}(n)\} \overset{\mathcal{D}}{\equiv} \{Y(n, M)\}$. □

3. Proof of Theorem 1.10 for $k = 2$

The reason we present an argument for $k = 2$ separately is that our proof of Theorem 1.11 is essentially this argument's follow-up.

Thus we consider $Y(n, p) := Y_2(n, p)$, the Bernoulli 2-dimensional complex with the complete 1-dimensional skeleton. Our main task is to bound the expected number of non-trivial cocycles for p close to $(2 \log n)/n$.

A 1-cocycle f induces a graph $G = G(f)$ on the vertex set $[n]$ with the edge set $E(G) = \{\mathbf{u} \in F_1 \,:\, f(\mathbf{u}) = 1\}$, i.e. the support of f. Let $\mathbf{d} = \mathbf{d}_G = \{d_G(\mathbf{v})\} = \{d(\mathbf{v})\}$ be the degree sequence of $G = G(f)$. A key idea of [10], [12] was to focus on non-trivial cocycles f with the smallest $|E(G(f))|$. These extremal cocycles have three crucial properties.

First of all, it turned out that, for every such cocycle f,

$$\max_{\mathbf{v}\in[n]} d\big(\mathbf{v}(G(f))\big) \le D := \left\lfloor \frac{n-1}{2} \right\rfloor, \tag{3.1}$$

a crucial improvement of the trivial bound $n-1$. (For $k \ge 2$, the bound is $\lfloor (n-k+1)/2 \rfloor$.)

Second, the graph $G(f)$ has a single non-trivial component.

To formulate the third, rather subtle, property, introduce $X(f)$, the number of such triangles that contain an odd number, 1 or 3, edges from $E(G(f))$. Then

$$X(f) \ge \frac{n|E(G(f))|}{3}. \tag{3.2}$$

(For the k-dimensional complex, the lower bound is $n|E(f)|/(k+1)$, [12].)

Why does $X(f)$ matter so much? Because the probability P_n that $Y(n,p)$ has a non-trivial 1-cocycle is *at most* the expected number of 1-*cochains* f, having those three properties, such that $Y(n,p)$ does not contain any one of $X(f)$ triangles. The probability of this event is

$$(1-p)^{X(G(f))} \le e^{-pX(G(f))}.$$

Therefore

$$P_n \le \sum_{m\ge 1} s(m), \quad s(m) := \sum_{G: e(G)=m} e^{-pX(G)}; \tag{3.3}$$

here $X(G)$ is the total number of triangles that contain an odd number of edges of G, and the sum is over all graphs G with $e(G) = m$ edges, of maxdegree $\le \lfloor (n-1)/2 \rfloor$, with a single non-trivial component, and $X(G) \ge ne(G)/3$.

Lemma 3.1. *Let*

$$p = \frac{2\log n + x_n}{n}, \quad |x_n| = o(\log n). \tag{3.4}$$

Then (1)

$$\sum_{G\,:\,e(G)>1} \exp\bigl[-pX(G)\bigr] \le_b n^{-1} e^{2|x_n|} \to 0, \tag{3.5}$$

and (2)

$$\sum_{G\,:\,e(G)=1} \exp\bigl[-pX(G)\bigr] \sim \frac{e^{-x_n}}{2}. \tag{3.6}$$

Proof of Lemma 3.1. The cases $e(G) = O(n)$ and $e(G) > n^{1+\varepsilon}$ are relatively simple, and it is the intermediate values of $e(G)$ where our argument truly differs from those in [10], [12]. To be sure, our treatment of $e(G) = O(n)$ is different enough to cover $p = (2\log n + x_n)$, $|x_n| = o(\log n)$, compared with $x_n \to \infty$ in [10]-[12].

Let $\nu = \nu(G)$ denote the number of vertices, and $m = e(G)$ the number of edges in a non-trivial component $C = C(G)$ of a generic graph G in question. Then $m \ge \nu - 1$.

(1) Let $\nu \le an$, where $a \in (0, 1/4)$. Obviously

$$X(G) \ge (n-\nu)m. \tag{3.7}$$

The total number of graphs G on $[n]$ with m edges and $|V(C(G))| = \nu$ is at most

$$\binom{n}{\nu}\binom{\binom{\nu}{2}}{m} \le \binom{n}{\nu}\left(\frac{e\binom{\nu}{2}}{m}\right)^m.$$

So

$$\sum_{G:\,|V(C)|=\nu,\,|E(C)|=m} \exp\bigl[-pX(G)\bigr] \le S(\nu, m),$$

where

$$S(\nu,m) := \binom{n}{\nu}\left(\frac{e\binom{\nu}{2}}{m}\right)^m \exp\bigl[-p(n-\nu)m\bigr].$$

Now, using $m+1 \ge \nu$,

$$\frac{S(\nu,m+1)}{S(\nu,m)} \le \frac{\nu^2}{m+1}\exp\bigl[-p(n-\nu)\bigr]$$
$$\le \nu\exp\bigl[-p(n-\nu)\bigr] \le \frac{e^{|x_n|}}{n^{1-2a}} \to 0,$$

as $a<1/2$. Hence, for $\nu \le an$,

$$\sum_{m\ge \nu-1} S(\nu,m) \le 2S(\nu,\nu-1) \le_b S(\nu);$$
$$S(\nu) := (4n)^\nu \exp\bigl[-p(n-\nu)(\nu-1)\bigr].$$

Now, for $\nu \le an$,

$$\frac{S(\nu+1)}{S(\nu)} = 4n\exp\bigl[-p(n-2\nu)\bigr] \le 4\frac{e^{|x_n|}}{n^{1-4a}} \to 0,$$

as $a<1/4$. Therefore

$$\sum_{G\,:\,3\le|V(C)|\le an} \exp\bigl[-pX(G)\bigr] = O\bigl(n^{-1}e^{2|x_n|}\bigr). \tag{3.8}$$

And, of course,

$$\sum_{G\,:\,|V(C)|=2,\,|E(C)|=1} \exp\bigl[-pX(G)\bigr] = \binom{n}{2}\exp\bigl[-p(n-2)\bigr] \sim \frac{1}{2}e^{-x_n}. \tag{3.9}$$

(2) Let $m \ge m_n := 3n^{4/3}e^{|x_n|}$. Using $X(G)\ge nm/3$, we have

$$s(m) \le \sum_\nu \binom{n}{\nu}\binom{\binom{\nu}{2}}{m}\cdot\exp\bigl[-pnm/3\bigr]$$
$$\le 2^n\left(\frac{2n^2}{m}\right)^m\cdot\exp\bigl[-pnm/3\bigr] = 2^n\left(\frac{2n^2}{me^{pn/3}}\right)^m$$
$$\le 2^n\left(\frac{2n^{4/3}e^{|x_n|}}{m}\right)^m.$$

Consequently

$$\sum_{m\ge m_n} s(m) \le \exp\bigl(n\log 2 - n^{4/3}\log 2\bigr) \to 0, \tag{3.10}$$

superexponentially fast.

(3) It remains to consider $\nu \ge an$ and $m \le m_n$. It is crucial that $m/\nu^2 \to 0$ in this range. By symmetry,

$$\sum_{G:e(G)=m,|V(C(G))|=\nu} \exp\bigl[-pX(G)\bigr] = \binom{n}{\nu}\sum_{G:e(G)=m,V(C(G))=[\nu]} \exp\bigl[-pX(G)\bigr]. \tag{3.11}$$

We need to show that the RHS tends to 0 as $n\to\infty$.

Let $\mathbf{d} = (d_1,\dots,d_\nu)$ be the generic vertex degrees of $C(G)$; so

$$\|\mathbf{d}\| := \sum_u d_u = 2\mu, \quad 1 \le d_u \le n/2,\ \ \forall\, u\in[\nu]. \tag{3.12}$$

Notice upfront that the total number of such graphs, connected or not, is bounded above by

$$(2\mu-1)!!\prod_u \frac{1}{d_u!}, \tag{3.13}$$

see Bender and Canfield [3].

Our next step, logically, is to find a lower bound for $X(G)$ in terms of $\mathbf{d}$. To this end, we first write

$$X(G) = (n-\nu)\mu + Y_1(C) + Y_3(C), \tag{3.14}$$

where $Y_1(C)$ ($Y_3(C)$ resp.) is the total number of triples (u,v,w) from the ν vertices such that (u,v) is an edge, and (u,w) and (v,w) are not edges ((u,w), (v,w) are edges, resp.). For a given edge (u,v) the number of $w\neq u,v$ such that at least one of (u,w), (v,w) is an edge is

$$\begin{aligned}&(d_u-1)+(d_v-1)\\ &-|\{w\neq u,v\,:\,(u,w),\,(v,w)\text{ edges, both}\}|\\ &=(\nu-2)-|\{w\neq u,v\,:\,(u,w),\,(v,w)\text{ not edges}\}|,\end{aligned}$$

whence

$$\begin{aligned}|\{w\neq u,v\,:\,(u,w),\,(v,w)\text{ edges, both}\}| &= (d_u+d_v)-\nu\\ &\quad+|\{w\neq u,v\,:\,(u,w),\,(v,w)\text{ not edges}\}|.\end{aligned}$$

So, summing over all edges (u,v), and noticing that every triangle in C will be counted thrice,

$$\begin{aligned}3Y_3(C) &= \sum_{(u,v)\text{ edge}}(d_u+d_v)-\nu m+Y_1(C)\\ &=\frac12\sum_{\{u,v\}:\,(u,v)\text{ edge}}(d_u+d_v)-\nu m+Y_1(C)\\ &=\sum_u d_u\sum_{v\neq u}\mathbf{1}_{\{(u,v)\text{ edge}\}}-\nu m+Y_1(C)\\ &=\sum_u d_u^2-\nu m+Y_1(C).\end{aligned}$$

Consequently

$$Y_1(C)\ge \nu m-\sum_u d_u^2,\quad Y_3(C)\ge\frac13\left(\sum_u d_u^2-\nu m\right). \tag{3.15}$$

The second inequality is known, see Lovász [11], Solution of Exercise 10.33. Since

$$\sum_u d_u^2\ge\nu\left(\frac{2m}{\nu}\right)^2=\frac{4m^2}{\nu},$$

it follows from (3.15) that

$$Y_3(C)\ge\frac13\left(\frac{4m^2}{\nu}-\nu m\right), \tag{3.16}$$

a classic inequality, due to Tuŕan, useful when $m > \nu^2/4$. However, in our case $m = o(\nu^2)$, so we pin our hopes on the lower bound for $Y_1(C)$ in (3.15). It gives

$$\exp[-pX(G)] \le \exp\left(-pnm + p\sum_u d_u^2\right). \tag{3.17}$$

So (3.11), (3.13) and (3.17) yield

$$\begin{aligned}\sum_{G:e(G)=m,\,|V(C(G))|=\nu} &\exp\bigl[-pX(G)\bigr]\\ &\le e^{-pnm}\binom{n}{\nu}(2m-1)!!\sum_{\mathbf{d}\text{ meets (3.12)}}\prod_u \frac{\exp(p\,d_u^2)}{d_u!}.\end{aligned} \tag{3.18}$$

Using $d! \ge (d/e)^d$,

$$\prod_u \frac{\exp(p\,d_u^2)}{d_u!} \le e^{H(\mathbf{d})},$$

where

$$H(\mathbf{d}) = \sum_u \phi(d_u), \quad \phi(d) = d\log\frac{e}{d} + pd^2.$$

Let us show that $\sum_v \phi(d_v)$ is "negligible", uniformly for $\{d_v\}$ in question. To this end, notice that

$$\phi^{(2)}(d) = -\frac{1}{d} + 2p$$

is negative (positive resp.) for $d < \bar{d}$ ($d > \bar{d}$ resp.), where

$$\bar{d} = \frac{1}{2p} = \frac{n}{2(2\log n + x_n)} = \Theta\left(\frac{n}{\log n}\right). \tag{3.19}$$

That is, $\phi(d)$ is strictly concave if $d \ge \bar{d}$. Since $d \in [\bar{d}, n/2]$ is a convex combination of $\bar{d}$ and D,

$$d = \frac{D-d}{D-\bar{d}}\,\bar{d} + \frac{d-\bar{d}}{D-\bar{d}}\,D,$$

we have then

$$\phi(d) \le \frac{D-d}{D-\bar{d}}\,\phi(\bar{d}) + \frac{d-\bar{d}}{D-\bar{d}}\,\phi(D).$$

So, introducing

$$\nu_1 = \nu_1(\mathbf{d}) := |\{v\,:\,d_v \ge \bar{d}\}|, \qquad \mu_1 = \mu_1(\mathbf{d}) = \sum_{v\,:\,d_v\ge\bar{d}} d_v, \tag{3.20}$$

we have

$$\begin{aligned}\sum_{\mathbf{v}\,:\,d_{\mathbf{v}}\ge\bar{d}} \phi(d) &\le \frac{\phi(\bar{d})}{D-\bar{d}}\,(\nu_1 D - \mu_1) + \frac{\phi(D)}{D-\bar{d}}\,(\mu_1 - \nu_1\bar{d})\\ &= -\nu_1\,\frac{\phi(D)\bar{d} - \phi(\bar{d})D}{D-\bar{d}} + \mu_1\,\frac{\phi(D)-\phi(\bar{d})}{D-\bar{d}}.\end{aligned} \tag{3.21}$$

Direct computation shows that

$$\alpha_n := \frac{\phi(D)\bar{d} - \phi(\bar{d})D}{D - \bar{d}} = \frac{n}{4}\bigl[1 + O\bigl((\log\log n + |x_n|)/\log n\bigr)\bigr],$$
$$\beta^n := \frac{\phi(D) - \phi(\bar{d})}{D - \bar{d}} = O(1 + |x_n|);$$

in particular, $\alpha_n > 0$, and crucially $\beta^n = o(\log n)$. Consequently, since $\mu_1 \le 2m$, (3.21) yields

$$\sum_{\mathbf{v}\,:\,d_{\mathbf{v}} \ge \bar{d}} \phi(d_{\mathbf{v}}) \le O\bigl(m(1 + |x_n|)\bigr). \tag{3.22}$$

Next, if $d_v \in [0, \bar{d}\,)$, then, by (3.19)

$$\phi(d_v) \le 2d_v + p\bar{d}\,d_v = \frac{5}{2}\,d_v.$$

so

$$\sum_{v\,:\,d_v \le \bar{d}} \phi(d_{\mathbf{v}}) \le \frac{5}{2}\sum_{v\,:\,d_v \le \bar{d}} d_v \le 5m. \tag{3.23}$$

Combining (3.22) and (3.23) we obtain

$$\sum_{v} \phi(d_v) \le O\bigl(m(1 + |x_n|)\bigr). \tag{3.24}$$

Now the number of summands in the sum on the RHS of (3.18) is, at most, the total number of positive solutions of $\|\mathbf{d}\| = 2m$, which is

$$\binom{2m-1}{\nu-1} \le_b \left(\frac{2em}{\nu}\right)^{\nu} = \exp\bigl[\nu\log(m/\nu) + O(\nu)\bigr]. \tag{3.25}$$

Also

$$\binom{n}{\nu} \le 2^n, \quad (2m-1)!! = O((2m/e)^m). \tag{3.26}$$

Combining (3.18), (3.24)-(3.26), we obtain

$$\begin{aligned}\sum_{G\colon |V(C)|=\nu,\, |E(C)|=m} &\exp\bigl[-pX(G)\bigr]\\ &\le\ \exp\bigl[-pnm + m\log m + \nu\log(m/\nu) + O(m(1 + |x_n|))\bigr].\end{aligned} \tag{3.27}$$

Here, since $m \le m_n := 3n^{4/3}e^{|x_n|}$,

$$\begin{aligned}-\,pnm + m\log m + \nu\log(m/\nu) &= -m\left(pn - \log m - \frac{\nu}{m}\log\frac{m}{\nu}\right)\\ &\le -m\left(2\log n + 2|x_n| - \frac{4}{3}\log n - \log 3 + O(n^{-1/3}\log n)\right) \le -\beta m\log n,\end{aligned}$$

for a $\beta \in (0, 2/3)$ and all large enough n. So, for $\beta^* \in (0, \beta)$,

$$\sum_{G\colon |V(C)|=\nu,\, |E(C)|=m} \exp\bigl[-pX(G)\bigr] \le \exp\bigl[-\beta^* m\log n\bigr], \tag{3.28}$$

uniformly for $\nu \ge an$, $\nu - 1 \le m \le m_n$, and n large enough. It follows from (3.28) that

$$\sum_{\substack{\nu \ge an \\ \nu-1 \le m \le m_n}} \sum_{G\colon |V(C)|=\nu,\, |E(C)|=m} \exp\bigl[-pX(G)\bigr] \le_b \exp(-\gamma n \log n), \tag{3.29}$$

for $\gamma \in (0, a\beta^*)$.

Combining (3.8), (3.9), (3.10) and (3.29) we complete the proof of Lemma 3.1. □

Thus, for $p = (2\log n + x_n)/n$, and $|x_n| = o(\log n)$, whp there are no extremal non-trivial cocycles with support size 2 or more. Let $\mathcal{C}_n$ be the total number of cocycles with support size 1, i.e. isolated edges. By part (2) of this Lemma, if $x_n = c$, then

$$\lim_{n\to\infty} \mathbb{E}[\mathcal{C}_n] = \lambda := \frac{e^{-c}}{2}.$$

And, as we had mentioned in the proof of Theorem 1.8, it can be shown that, in general,

$$\lim_{n\to\infty} \mathbb{E}\bigl[(\mathcal{C}_n)_t\bigr] = \lambda^t, \quad t \ge 1.$$

So $\mathcal{C}_n$ is in the limit Poisson(λ), which completes the proof of Theorem 1.10. □

4. Proof of Theorem 1.11

As in the proof of Theorem 1.9, we embed the 2-dimensional process $\{Y(n, M)\}$ into the continuous-time Markov process $\{Y_t(n)\}$. We introduce

$$p_1 = \frac{2\log n - \log\log n}{n}, \quad p_2 = \frac{2\log n + \log\log n}{n},$$

and t_i defined by $p_i = 1 - e^{-t_i}$. For $k = 2$, our argument showed that w.h.p. throughout $[t_1, t_2]$ the complex $Y_t(n)$ continues to be a giant component plus a set of isolated edges, whose number can decrease by 1 only. That is, w.h.p. there is a random moment $\tau \in (t_1, t_2)$ when the last isolated edge disappears, being swallowed by the giant component. We need to show that w.h.p. $H^1(Y_\tau(n))$ vanishes as well. Suppose not. Then there exists a 1-cochain f, of support size 2 or more, meeting the three conditions necessary for a non-trivial cocycle, such that none of $X(G(f))$ triangles is present in $Y_{\tau-}(n)$, i.e. $Y_\tau(n)$ minus the triangle born at time τ. Let (u, v) denote a generic value of the last isolated edge that disappeared at time τ. Then

$$X_{(u,v)}(G(f)) \le X(G(f)) \le X_{(u,v)}(G(f)) + n - 2, \tag{4.1}$$

where $X_{(u,v)}(G(f))$ is the number of triangles, except those containing (u, v), that contain an odd number of edges of $G(f)$. Introduce $Y_t(n; (u, v))$, a subcomplex of $Y_t(n)$, with all the triangles in $Y_t(n)$, if any, that contain (u, v) being deleted. Then

$$\begin{aligned} &\mathbb{P}(H^1(Y_\tau(n)) \text{ does not vanish}) \\ &\quad \le o(1) + \sum_{f,\,(u,v)} \int_{t_1}^{t_2} (1 - p(t))^{X_{(u,v)}(G(f))}\, (n-2)e^{-t}(1 - p(t))^{n-3}\, dt. \end{aligned} \tag{4.2}$$

Explanation. $o(1)$ stands for the probability that $\tau \in (t_1, t_2)$ does not exist. t is the generic value of a time when the edge (u,v) stops being isolated. $(1-p(t))^{X_{(u,v)}(G(f))}$ is the probability that none of $X_{(u,v)}(G(f))$ triangles are present in $Y_t(n;(u,v))$.

Let $m = m(f) \geq 2$ be support size for f. By Meshulam-Wallach inequality and (4.1),

$$X_{(u,v)}(G(f)) \geq X(G(f)) - (n-2) \geq \frac{nm}{3} - (n-2).$$

So, like part **(2)** of the proof of Lemma 3.1, for $m \geq 2m_n = 6n^{4/3}e^{|c|}$,

$$\sum_{f:e(G(f))\geq 2m_n} (1-p(t))^{X_{(u,v)}(G(f))} \leq \exp(-n^{4/3}), \tag{4.3}$$

uniformly for $t \in [t_1, t_2]$. Hence the contribution of all such f's to the RHS of (4.2) is superexponentially small.

Suppose that $m \leq 2m_n$. Let $\nu = |V(C(f))|$, $C(f) := C(G(f))$. Then

$$X_{(u,v)}(G(f)) \geq (\nu-1)m - \sum_{v\in V(C(f))} d_v^2.$$

Indeed, using the proof of (3.15) we see the following. (1) If (w,x) from the support of f is not in a triangle containing (u,v), then the number of triangles that contain (w,x) as the only edge supporting f is $\nu - (d_w + d_x)$, at least. (2) If on the other hand (w,x) and (u,v) are edges from the same triangle, then this triangle is unique, and so the number of the triangles containing (w,x) as the only edge supporting f is $(\nu-1) - (d_w + d_x)$, at least. Hence

$$\begin{aligned} X_{(u,v)}(G(f)) &\geq (\nu-1)m - \sum_{(w,x)\in E(C(f))} (d_w + d_x) \\ &= (\nu-1)m - \sum_{v\in V(C(f))} d_v^2. \end{aligned}$$

So, with only trivial changes, we obtain a counterpart of (3.29): for $a > 0$,

$$\sum_{\substack{\nu\geq an \\ \nu-1\leq m\leq 2m_n}} \sum_{|V(C(f))|=\nu,\,|E(C(f))|=m} (1-p(t))^{X_{(u,v)}(G(f))} \leq_b \exp(-\gamma n \log n), \tag{4.4}$$

$\gamma = \gamma(a) > 0$.

Finally, suppose that $\nu \leq an$, $a < 1$. A counterpart of the bound (3.7) is

$$X_{(u,v)}(G(f)) \geq (n-\nu-1)m,$$

since, given an edge supporting f, there can be at most one triangle that contains this edge and (u,v). So, with only minor changes in the part **(1)** of Lemma 3.1, we obtain: for $a < 1/2$,

$$\begin{aligned} \sum_{\substack{3\leq\nu\leq an \\ \nu-1\leq m}} \sum_{|V(C(f))|=\nu,\,|E(C(f))|=m} & (1-p(t))^{X_{(u,v)}(G(f))} \\ &= O(n^{-1}e^{2\log\log n}) = O(n^{-1}\log^2 n), \end{aligned} \tag{4.5}$$

uniformly for $t \in [t_1, t_2]$.

Combining (4.3), (4.4) and (4.5), we have

$$\sum_f (1-p(t))^{X_{(u,v)}(G(f))} = O(n^{-1}\log^2 n).$$

Consequently, the bound (4.2) becomes

$$\begin{aligned}
&\mathbb{P}(H^1(Y_\tau(n)) \text{ does not vanish})\\
&\qquad \le_b o(1) + n^3 e^{-np_1}(t_2-t_1)n^{-1}\log^2 n\\
&\qquad\qquad \le_b o(1) + n^{-1}\log^4 n \to 0, \quad n\to\infty.
\end{aligned}$$

This completes the proof of Theorem 1.11. □

5. Proof of Theorem 1.10 for $k \geq 2$

Our proof is a minor refinement of the proofs in [10]-[12]. Like the case $k=2$ in Section 3, we need to show that, for

$$p = \frac{k\log n + x_n}{n}, \quad |x_n| = o(\log n),$$

$$\sum_{H:|E(H)|>1} \exp\bigl[-pX(H)\bigr] \to 0, \tag{5.1}$$

$$\sum_{H:|E(H)|=1} \exp\bigl[-pX(H)\bigr] \sim \frac{e^{-x_n}}{k!}. \tag{5.2}$$

Here H is a hypergraph on hypervertex-set $\binom{[n]}{[k-1]}$, with hyperedge-set $E(H) \subseteq \binom{[n]}{[k]}$, and $X(H)$ is the total number of k-simplexes that contain an odd number of hyperedges of H. Further, an admissible H meets the conditions:

(1) maximum hypervertex degree is at most $\lfloor (n-k+1)\rfloor/2$;

(2) H has a single non-trivial component;

(3) $X(H) \geq n|E(H)|/(k+1)$.

For $|E(H)|$ "small", the argument is analogous to the part **(1)** of the proof of Lemma 3.1. For completeness, here it is.

Let $\nu = \nu(H)$ denote the range of H, i. e.

$$\nu(H) := |\{i \in [n] : i \in e \text{ for some } e \in E(H)\}|;$$

then $X(H) \geq (n-\nu)m$. Since H has a single non-trivial component,

$$m := |E(H)| \geq \nu - (k-1).$$

Let $k \leq \nu \leq an$, $0 < a < 1/(2k)$. Then $X(H) \geq (n-\nu)m$. The total number of H with $|E(H)| = m$ and $\nu(H) = \nu$, is at most

$$\binom{n}{\nu}\binom{\binom{\nu}{k}}{m} \leq \binom{n}{\nu}\left(\frac{e\binom{\nu}{k}}{m}\right)^m.$$

So

$$\sum_{H:\nu(H)=\nu,\,|E(H)|=m} \exp\bigl[-pX(H)\bigr] \leq S(\nu,m),$$

where

$$S(\nu,m) := \binom{n}{\nu}\left(\frac{e\binom{\nu}{k}}{m}\right)^m \exp\bigl[-p(n-\nu)m\bigr].$$

Now, using $m \ge \nu-(k-1)$,

$$\begin{aligned}\frac{S(\nu,m+1)}{S(\nu,m)} &\le \frac{\nu^2}{m+1}\exp\bigl[-p(n-\nu)\bigr]\\ &\le k\nu\exp\bigl[-p(n-\nu)\bigr] \le_b \frac{e^{|x_n|}}{n^{(k-1)-ka}} \to 0,\end{aligned}$$

as $k-1-ka>0$. Hence

$$\begin{aligned}\sum_{m\ge\nu-1} S(\nu,m) \le& 2S(\nu,\nu-k+1) \le_b S(\nu);\\ S(\nu) :=& (\beta n)^\nu \nu^{(k-2)\nu}\exp\bigl[-p(n-\nu)(\nu-k+1)\bigr],\end{aligned}$$

for some $\beta>0$. It is easy to check that $\rho(\nu) := S(\nu+1)/S(\nu)$ increases with ν, and

$$\rho(an) \le_b \frac{e^{|x_n|}}{n^{1-2ak}} \to 0,$$

as $a<1/2k$. So

$$\sum_{\nu=k+1}^{an} S(\nu) \le_b S(k+1) \le_b \frac{e^{2|x_n|}}{n^{k-1}},$$

which proves that

$$\sum_{H:|E(H)|>1,\,\nu(H)\le an}\exp\bigl[-pX(H)\bigr]\to 0.$$

In addition,

$$\sum_{H:|E(H)|=1}\exp\bigl[-pX(H)\bigr] = \binom{n}{k}\exp\bigl[-p(n-k)\bigr] \sim \frac{e^{-x_n}}{k!},$$

which proves (5.2).

The relation (5.1) will follow if we can show that

$$\sum_{H:|E(H)|>1,\,\nu(H)\ge an}\exp\bigl[-pX(H)\bigr]\to 0,$$

as well. For those H,

$$|E(H)| \ge \nu(H)+k-1 \ge an+k-1 \ge an/2,$$

and all we need is to cite a remarkable bound established in [12]: for a given $\alpha>0$,

$$\sum_{H:|E(H)|\ge\alpha n}\exp\bigl[-pX(H)\bigr] \le \exp\bigl[-\Omega(n\log n)\bigr]. \tag{5.3}$$

(To be sure, (5.3) was stated and proved for $x_n\to\infty$ arbitrarily slow. However, only obvious changes in the proof are required to cover $|x_n| = o(\log n)$.) This highly non-trivial estimate was obtained by showing, via probabilistic method, that for every hypergraph H, either there are "many" k-simplexes that contain exactly one of hyperedges of H, or there is a "small" subset $\hat{E}\subset E(H)$ such that almost all other hyperedges are incident to hyperedges in $\hat{E}$.

With (5.1)-(5.2), the proof of Theorem 1.10 for $k > 2$ is completed exactly like for $k = 2$ in Section 3. $\square$

Acknowledgment. The authors are very grateful to Nati Linial and Roy Meshulam for several inspiring, helpful conversations, both in person and through email.

References

[1] L. Aronshtam, N. Linial, T. Luczak, and R. Meshulam. Collapsibility and vanishing of top homology in random simplicial complexes. To appear in Discrete and Computational Geometry, 2011.

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

[3] Edward A. Bender and E. Rodney Canfield. The asymptotic number of labeled graphs with given degree sequences. *J. Combinatorial Theory Ser. A*, 24(3):296–307, 1978.

[4] Béla Bollobás and Oliver Riordan. Clique percolation. *Random Structures Algorithms*, 35(3):294–322, 2009.

[5] Béla Bollobás and Andrew Thomason. Random graphs of small order. In *Random graphs '83 (Poznań, 1983)*, volume 118 of *North-Holland Math. Stud.*, pages 47–97. North-Holland, Amsterdam, 1985.

[6] P. Erdős and A. Rényi. On random graphs. I. *Publ. Math. Debrecen*, 6:290–297, 1959.

[7] Allen Hatcher. *Algebraic topology*. Cambridge University Press, Cambridge, 2002.

[8] Svante Janson, Tomasz Łuczak, and Andrzej Rucinski. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.

[9] Dmitry N. Kozlov. The threshold function for vanishing of the top homology group of random d-complexes. *Proc. Amer. Math. Soc.*, 138(12):4517–4527, 2010.

[10] Nathan Linial and Roy Meshulam. Homological connectivity of random 2-complexes. *Combinatorica*, 26(4):475–487, 2006.

[11] László Lovász. *Combinatorial problems and exercises*. North-Holland Publishing Co., Amsterdam, second edition, 1993.

[12] R. Meshulam and N. Wallach. Homological connectivity of random k-dimensional complexes. *Random Structures Algorithms*, 34(3):408–417, 2009.

[13] V. E. Stepanov. Combinatorial algebra and random graphs. *Teor. Verojatnost. i Primenen*, 14:393–420, 1969.