



Simplex links in determinantal hypertrees

Andrew Vander Werf¹

Received: 5 February 2023 / Revised: 10 October 2023 / Accepted: 28 December 2023 /
Published online: 7 February 2024
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2024

Abstract

We deduce a structurally inductive description of the determinantal probability measure associated with Kalai’s celebrated enumeration result for higher-dimensional spanning trees of the $(n - 1)$ -simplex. As a consequence, we derive the marginal distributions of the simplex links in such random trees. Along the way, we also characterize the higher-dimensional spanning trees of every other simplicial cone in terms of the higher-dimensional rooted forests of the underlying simplicial complex. We also apply these new results to random topology, the spectral analysis of random graphs, and the theory of high dimensional expanders. One particularly interesting corollary of these results is that the fundamental group of a union of $o(\log n)$ determinantal 2-trees has Kazhdan’s property (T) with high probability.

Keywords Random simplicial complexes · Spanning trees · High dimensional expanders · Determinantal measure

Mathematics Subject Classification 05C05 · 60C05 · 60B99 · 55U05 · 55U15

1 Introduction

Forty years ago, Kalai (1983) introduced, to spectacular effect, a generalization of the graph-theoretic notion of a tree to higher-dimensional simplicial complexes, called \mathbb{Q} -acyclic simplicial complexes for the triviality of their rational reduced homology groups in every dimension. However, recent authors (Linial and Peled 2019; Kahle and Newman 2022; Mészáros 2022) appear to have settled on simply calling these *hypertrees*. For $0 \leq k < n$, let $\mathcal{T}_{n,k}$ denote the set of k -dimensional hypertrees on the vertex set $[n] := \{1, 2, \dots, n\}$. Kalai noticed, among other things, that $\hat{H}_{k-1}(T)$ (assume integer coefficients throughout) is a finite group for all $T \in \mathcal{T}_{n,k}$ and moreover that

✉ Andrew Vander Werf
andrew_vander_werf@brown.edu

¹ Department of Applied Mathematics, Brown University, Providence, RI, USA

$$\sum_{T \in \mathcal{T}_{n,k}} |\tilde{H}_{k-1}(T)|^2 = n^{\binom{n-2}{k}},$$

which is seen to be a generalization of Caley’s formula by recalling that $|\tilde{H}_0(T)| = 1$ for all $T \in \mathcal{T}_{n,1}$, due to trees being connected. This suggests a natural probability measure (Lyons 2003, 2009) $\nu = \nu_{n,k}$, on $\mathcal{T}_{n,k}$ defined on atoms by $\nu_{n,k}(T) = n^{-\binom{n-2}{k}} |\tilde{H}_{k-1}(T)|^2$.

Seemingly unrelated to this measure, consider the k -dimensional *Linial-Meshulam complex*, denoted $\mathcal{Y}_k(n, p)$ and defined (Linial and Meshulam 2006; Meshulam and Wallach 2009) to be the random k -dimensional simplicial complex on $[n]$ with full $(k - 1)$ -skeleton wherein each k -face is included independently and with probability p . Let $\mu_{n,k}$ denote the probability density for $\mathcal{Y}_k(n, (n + 1)^{-1})$.

Let $\mathcal{T}_{n,k}$ denote a random complex distributed according to $\nu_{n,k}$, and let $\mathcal{Y}_{n,k}$ denote a random complex distributed according to $\mu_{n,k}$. Our main result is the following structure theorem for random hypertrees distributed according to $\nu_{n,k}$ (see Sect. 2 for definitions):

Theorem 1 *Assume $n > k + 1$. Then*

$$\text{Link}(n, \mathcal{T}_{n,k}) \stackrel{\mathcal{L}}{=} \mathcal{T}_{n-1,k-1} \cup \mathcal{Y}_{n-1,k-1}$$

where $\mathcal{T}_{n-1,k-1}, \mathcal{Y}_{n-1,k-1}$ are statistically independent. Also

$$\{f \in \mathcal{T}_{n,k} : n \notin f\} \stackrel{\mathcal{L}}{=} \mathcal{T}_{n-1,k} \setminus \mathcal{Y}_{n-1,k}$$

where $\mathcal{T}_{n-1,k}, \mathcal{Y}_{n-1,k}$ are statistically independent.

The first claim of this theorem states that a vertex link in $\mathcal{T}_{n,k}$ can be simulated by first sampling $\mathcal{T}_{n-1,k-1}$ and then adding each missing $(k - 1)$ -face independently with probability $1/n$. This effectively gives us the law of the set of faces in $\mathcal{T}_{n,k}$ which contain the vertex n . The second claim tells us that the remaining faces of $\mathcal{T}_{n,k}$ which do not contain the vertex n can be simulated by sampling $\mathcal{T}_{n-1,k}$ and then deleting each of its k -faces independently with probability $1/n$.

This simple idea of decomposing a hypertree into two collections of faces—those which contain a designated vertex and those which do not—turns out to be quite powerful. However, the idea is nothing new. For example, we can see this decomposition in use by Linial and Peled (2019) at Kalai’s suggestion to inductively construct collapsible hypertrees.

By iterating our formula for the law of a vertex link, we can also determine a similar expression for the law of the link of a simplex of arbitrary dimension.

Theorem 2 *The law of the link of any $(j - 1)$ -dimensional simplex in $\mathcal{T}_{n,k}$ is equivalent to that of $\mathcal{T}_{n-j,k-j} \cup \mathcal{Y}_{k-j}(n - j, j/n)$ where these two $(k - j)$ -dimensional complexes are independent.*

Proof We prove this by induction on the dimension of the simplex. By relabeling vertices, we may assume that the $(j - 1)$ -simplex we are interested in is $\{n - j + 1, \dots, n\}$. The base case $j = 1$ follows from Theorem 1. Inducting, we have

$$\begin{aligned} &\text{Link}(\{n - j + 1, \dots, n\}, \mathcal{T}_{n,k}) \\ &= \text{Link}(n - j + 1, \text{Link}(\{n - j + 2, \dots, n\}, \mathcal{T}_{n,k})) \\ &\stackrel{\mathcal{L}}{=} \text{Link}(n - j + 1, \mathcal{T}_{n-j+1,k-j+1} \cup \mathcal{Y}_{k-j+1}(n - j + 1, (j - 1)/n)). \end{aligned}$$

It is well known and easily verified from the definitions that the law of a vertex link of a Linial–Meshulam complex is a codimension-1 Linial–Meshulam complex on one fewer vertices with the same face probability. So, since the link operation is distributive over unions of k -complexes (see Sect. 2), this is equivalent to

$$\begin{aligned} &\text{Link}(n - j + 1, \mathcal{T}_{n-j+1,k-j+1}) \cup \text{Link}(n - j + 1, \mathcal{Y}_{k-j+1}(n - j + 1, (j - 1)/n)) \\ &\stackrel{\mathcal{L}}{=} \mathcal{T}_{n-j,k-j} \cup \mathcal{Y}_{k-j}(n - j, 1/(n - j + 1)) \cup \mathcal{Y}_{k-j}(n - j, (j - 1)/n) \\ &\stackrel{\mathcal{L}}{=} \mathcal{T}_{n-j,k-j} \cup \mathcal{Y}_{k-j}\left(n - j, \frac{1}{n - j + 1} + \frac{j - 1}{n} - \frac{1}{n - j + 1} \frac{j - 1}{n}\right) \\ &= \mathcal{T}_{n-j,k-j} \cup \mathcal{Y}_{k-j}\left(n - j, \frac{j}{n}\right), \end{aligned}$$

where the last equality in law follows from the independence of the two Linial–Meshulam complexes and the inclusion–exclusion principle. □

1.1 Applications to random topology

Theorem 2 has several applications to random topology. Indeed, Garland’s method (Garland 1973) and its refinements (see Žuk 1996, 2003; Oppenheim 2018, 2020), Žuk’s criterion among them, have proven to be highly effective tools for extracting global information about a pure k -dimensional simplicial complex using only information found in the $(k - 2)$ -dimensional links of the complex.

Theorem 3 (Garland’s method) *Let X be a pure k -dimensional simplicial complex. Suppose that, for all $(k - 2)$ -faces $\tau \in X$, we have $\lambda^{(0)}(\text{Link}(\tau, X)) \geq 1 - \varepsilon > 0$. Then $\lambda^{(k-1)}(X) \geq 1 - k\varepsilon$.*

Theorem 4 (Žuk’s criterion) *Let X be a pure 2-dimensional simplicial complex. Suppose that, for all 0-faces $\tau \in X$, we have that $\lambda^{(0)}(\text{Link}(\tau, X)) > 1/2$ and $\text{Link}(\tau, X)$ is connected. Then the fundamental group $\pi_1(X)$ has Kazhdan’s property (T) .*

The $\lambda^{(k-1)}(X)$ mentioned in the above statement of Garland’s method refers to the smallest nonzero eigenvalue of the top-dimensional up-down Laplacian of X under a particular weighted inner product (see Lubotzky 2018; Gundert and Wagner 2014 for greater detail). This eigenvalue will be referred to as the *spectral gap* of X . In the special case where $k = 1$ and X is a connected graph, $\lambda^{(0)}(X)$ corresponds to the

second smallest eigenvalue of the reduced Laplacian of X . Fortunately, the spectral gap of a random graph, in one form or another, is already quite well studied (Feige and Ofek 2005; Coja-Oghlan and Lanka 2009; Oliveira 2009; Tikhomirov and Youssef 2019; Hoffman et al. 2019). Combining our characterization of the links in $\mathcal{T}_{n,k}$ with the best known techniques for the kind of spectral gap estimation we would like to do, we find the following:

Proposition 5 *Let $\delta > 0$ be an arbitrary fixed constant, and let \mathcal{X} be the union of $\lceil \delta \log n \rceil$ jointly independent copies of $\mathcal{T}_{n,k}$ with k fixed. Then, for any fixed $s > 0$, there exists a constant $C > 0$ so that $|\lambda^{(k-1)}(\mathcal{X}) - 1| \leq \frac{C}{\sqrt{\log n}}$, with probability $1 - o(n^{-s})$.*

Proof This follows immediately from Garland's method in combination with Lemma 36 and a union bound on the probability that any one-dimensional link of \mathcal{X} fails to meet the criteria of Garland's method. \square

Applying Żuk's criterion in the case $k = 2$ gives the following corollary:

Corollary 6 *Let $\delta > 0$ be an arbitrary fixed constant, and let \mathcal{X} be the union of $\lceil \delta \log n \rceil$ jointly independent copies of $\mathcal{T}_{n,2}$. Then, for any fixed $s > 0$, we have with probability $1 - o(n^{-s})$ that $\pi_1(\mathcal{X})$ has Kazhdan's property (T).*

Particularly in computer science, there is a growing interest in generating families of graphs with large spectral gap, which are usually called expander graphs, and probabilistic constructions have been offered as a way to do this quickly and successfully with exceedingly high probability. For example, the authors of Hoffman et al. (2019) prove that the Erdős–Rényi random graph $\mathcal{G}(n, p)$ (in particular the random infinite family $\{\mathcal{G}(n, p(n))\}_{n \geq 1}$) achieves the same spectral gap as found in Lemma 36 and with equally high probability when $np \geq (1/2 + \delta) \log n$ for any fixed $\delta > 0$. Moreover, they show that the assumption $\delta > 0$ is necessary for this to hold. The best known expander graph families have been constructed explicitly and have asymptotically (with respect to vertex count) constant average degree. So, while the Erdős–Rényi random graph can succeed at being a reliable expander, it is only able to do so if it is allowed an expected average degree far exceeding $\frac{1}{2} \log n$.

Our characterization of the one-dimensional links of a determinantal (n, k) -tree as the union of a determinantal $(n - k + 1, 1)$ -tree with an independent $\mathcal{G}(n - k + 1, (k - 1)/n)$ makes it, or perhaps the union of a small number of independent copies of it, a potential naturally occurring candidate for a random expander graph with constant expected average degree. Lemma 36 shows that the number of superimposed independent copies of this graph required to match the result for $\mathcal{G}(n, p)$ is no more than $\delta \log n$ for every $\delta > 0$. In particular, the resulting graph has an expected average degree of around $(k + 1)\delta \log n$, improving upon $\mathcal{G}(n, p)$'s required expected average degree by a factor of arbitrary finite size.

1.2 Outline

Section 2 is primarily devoted to establishing notation and basic homological definitions as well as providing background for the deterministic study of simplicial spanning

trees. The only truly novel result of this section is Theorem 18. In Sect. 3 the results of Sect. 2 are applied to the spanning trees of the $(n - 1)$ -simplex to give Theorem 1. Section 4 establishes the tools required to prove Proposition 5 and Corollary 6.

2 Homological trees: simplicial and relative

A *chain complex* is a sequence of abelian groups, $\{C_j\}_{j \in \mathbb{Z}}$, called *chain groups*, linked by group homomorphisms $\partial_j : C_j \rightarrow C_{j-1}$ called *boundary maps* which satisfy $\partial_j \partial_{j+1} = 0$ for all $j \in \mathbb{Z}$, or equivalently $\text{Ker } \partial_j \supseteq \text{Im } \partial_{j+1}$. The *jth homology group* of a chain complex such as this is defined to be the quotient group $\text{Ker } \partial_j / \text{Im } \partial_{j+1}$. Since we will only be considering finitely-generated free \mathbb{Z} -modules for our chain groups, we can represent these boundary maps by integer matrices. By the structure theorem for finitely-generated abelian groups, the *jth homology group* can be expressed as the direct sum of a free abelian group, which is isomorphic to \mathbb{Z}^{β_j} and called its *free part*, and a finite abelian group called its *torsion subgroup* which is a direct sum of finite cyclic groups. The rank β_j of the free part is called the *jth Betti number*. We will require the following standard fact from homological algebra which gives two equivalent formulas for what is commonly called the *Euler characteristic* of a chain complex.

Lemma 7 *Suppose $(C_\#, \partial_\#)$ is a chain complex such that each chain group is freely and finitely generated and only finitely many of the chain groups are nontrivial. Let f_j denote the rank of C_j for each $j \in \mathbb{Z}$. Then*

$$\sum_{j \in \mathbb{Z}} (-1)^j f_j = \sum_{j \in \mathbb{Z}} (-1)^j \beta_j.$$

2.1 Simplicial

Fix an integer $n \geq 1$. The set $[n] := \{1, 2, \dots, n\}$ will be our vertex set. For $-1 \leq j \leq n - 1$, a *j-dimensional abstract simplex*, or *j-face*, is a subset of $[n]$ with cardinality $j + 1$. We denote the set of all *j-faces* on $[n]$ by $\binom{[n]}{j+1}$. All faces will be oriented according to the usual ordering on $[n]$. As such, we will be denoting elements of $\binom{[n]}{j+1}$ by $\{\tau_0, \tau_1, \dots, \tau_j\}$, where it is to be implicitly understood that $1 \leq \tau_0 < \tau_1 < \dots < \tau_j \leq n$. Let $\partial = \partial_k^{[n]}$ be the matrix that rows and columns are indexed respectively by $\binom{[n]}{k}$ and $\binom{[n]}{k+1}$, and for which, given $\sigma \in \binom{[n]}{k}$ and $\tau \in \binom{[n]}{k+1}$, we set

$$\partial(\sigma, \tau) := \begin{cases} (-1)^j, & \sigma = \tau \setminus \{\tau_j\} \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

to account for our choice of orientation for each face. In particular, $\partial_0^{[n]}$ is the $\binom{[n]}{0} \times \binom{[n]}{1}$ all-ones matrix.

An *abstract simplicial complex with vertices in $[n]$* is a nonempty subset $X \subseteq \bigcup_{j \geq -1} \binom{[n]}{j+1}$ which, for every pair of subsets $\sigma \subseteq \tau$, satisfies $\tau \in X \implies \sigma \in X$. Let \mathcal{A}_n denote the set of all abstract simplicial complexes on $[n]$. It is easily verified with this definition that \mathcal{A}_n is closed under intersection and union. We write $X_j := \binom{[n]}{j+1} \cap X$ and define the *dimension* of X to be $\dim X := \sup\{k \geq -1 : X_k \neq \emptyset\}$. If X has dimension k , X is said to be *pure* if, for every $-1 \leq j < k$ and every j -face $\sigma \in X$, there exists a $\tau \in X_k$ such that $\sigma \subset \tau$. An important example is $K_n^k := \bigcup_{j=0}^{k+1} \binom{[n]}{j}$, the complete k -dimensional complex on $[n]$.

Given a matrix M with entries indexed over the set $S \times T$, for $A \subseteq S$ and $B \subseteq T$, we write $M_{A,B}$ to denote the submatrix of M with rows indexed by A and columns indexed by B . We will also occasionally write $M_{\bullet,B}$ for the case $A = S$. As a point of clarification for this notation, transposes are handled by the convention of writing $M_{A,B}^t$ to mean $(M_{A,B})^t = (M^t)_{B,A}$.

Given $X \in \mathcal{A}_n$, we define its chain complex $(C_\#, \partial_\#)$ by $C_j(X) := \mathbb{Z}^{X_j}$, and $\partial_j^X := \partial_{X_{j-1}, X_j}$. Note then that we have $C_j(X) = 0$ for all $j \geq n$ and all $j < -1$ —because $|X_{-1}| = |\{\emptyset\}| = 1$. We denote the j th homology group of this chain complex by $\tilde{H}_j(X)$. It is equivalent to the j th *reduced* simplicial homology group, hence the notation. Note that, with this chain sequence, it makes sense to consider $\tilde{H}_{-1} = \text{Ker } \partial_{-1}^{[n]} / \text{Im } \partial_0^{[n]}$ as well, although we can see that this group is always trivial since $\partial_0^{[n]}$ is surjective.

Lemma 8 $\partial_k^{[n]} \partial_k^{[n]t} + \partial_{k-1}^{[n]t} \partial_{k-1}^{[n]} = n \text{ Id}$.

Proof This follows by explicit computation of matrix entries via (1) combined with the chain sequence identity $\partial_{k-1}^{[n]} \partial_k^{[n]} = 0$ which can also be seen to hold by explicit computation of matrix entries via (1). A more detailed explanation along these lines for the cases $k \geq 2$ can be found in the proof of Lemma 3 in Kalai (1983), and the case $k = 0$ is straightforward. For the case $k = 1$, $\partial_1^{[n]} \partial_1^{[n]t}$ is the combinatorial Laplacian of the complete graph on $[n]$. Hence $\partial_1^{[n]} \partial_1^{[n]t} = n \text{ Id} - J$ where J is the $n \times n$ all-ones matrix. Noticing that $\partial_0^{[n]t} \partial_0^{[n]} = J$ then completes the proof. \square

Fix a $\Delta \in \mathcal{A}_n$ and set $\mathcal{C}_{n,k}(\Delta) := \{X \in \mathcal{A}_n : K_n^{k-1} \cap \Delta \subseteq X \subseteq K_n^k \cap \Delta\}$. We will call elements of this set the (n, k) -complexes of Δ , or just k -complexes when the context is clear. Note this definition makes sense even if $\dim \Delta > k$, but in this case $\mathcal{C}_{n,k}(\Delta) = \mathcal{C}_{n,k}(K_n^k \cap \Delta)$. Building upon work done by Duval et al. (2008, 2011, 2015), Bernardi and Klivans (2015) define a higher-dimensional forest of Δ to be a subset $F \subseteq \Delta_k$ such that $\partial_{\Delta_{k-1}, F}$ is injective. Our definition of a forest will be equivalent to this one except that we will consider F to be an entire element of $\mathcal{C}_{n,k}(\Delta)$ by adding to it the $(k - 1)$ -skeleton of Δ , more similar to the situation in Duval et al. (2008). That is to say, an $F \in \mathcal{C}_{n,k}(\Delta)$ is called a k -forest of Δ if $\beta_k(F) = 0$, in which case we write $F \in \mathcal{F}_{n,k}(\Delta)$. It is easily verified that this is equivalent to Definition 3 in Bernardi and Klivans (2015) by noting that $\tilde{H}_k(F) = \text{Ker } \partial_k^F$ is a free group and thus is 0 if and only if its rank is 0. A k -forest $T \in \mathcal{F}_{n,k}(\Delta)$ with $|T_k| = \text{rank } \partial_k^\Delta$ (the maximal possible value for a forest) is said to be a k -tree of Δ . Let $\mathcal{T}_{n,k}(\Delta)$ denote the set of k -trees of Δ —(Bernardi and Klivans 2015) calls these the *spanning forests* of Δ . In the case $\Delta = K_n^{n-1}$, or just $\Delta \supseteq K_n^k$, these are Kalai’s k -dimensional

\mathbb{Q} -acyclic simplicial complexes mentioned in the introduction, and in this case we will suppress the (Δ) in all the above notation. The following result from Duval et al. (2015) generalizes Proposition 2 from Kalai (1983) to general k -trees of $\Delta \in \mathcal{A}_n$:

Lemma 9 *For $X \in \mathcal{C}_{n,k}(\Delta)$, if any two of the following conditions hold, then so does the other condition and moreover $X \in \mathcal{T}_{n,k}(\Delta)$:*

- $|X_k| = \text{rank } \partial_k^\Delta,$
- $\beta_{k-1}(X) = \beta_{k-1}(\Delta),$
- $\beta_k(X) = 0.$

2.2 Relative

Given a pair $(X, Y) \in \mathcal{A}_n^2$ with $X \supseteq Y$, we can define another chain complex by $\mathcal{C}_j(X, Y) := \mathbb{Z}^{X_j \setminus Y_j}$ with boundary maps $\partial_j^{X/Y} := \partial_{X_{j-1} \setminus Y_{j-1}, X_j \setminus Y_j}$. The homology groups of this chain complex are called *relative homology groups* and denoted $H_j(X, Y)$. We can recover ordinary reduced homology from this by taking Y to be $\{\emptyset\}$ or any complex with a single 0-face and no larger faces. Moreover, since homology is homotopy invariant, we also have that $H_j(X, Y) \cong \tilde{H}_j(X)$ as long as Y is contractible. An important example of a contractible complex, the simplicial cone of a complex $\Delta \in \mathcal{A}_{n-1}$ is defined by

$$\text{Cone}(n, \Delta) := \Delta \cup \{\sigma \cup \{n\} : \sigma \in \Delta\} \in \mathcal{A}_n.$$

We also define for any $X \in \mathcal{C}_{n,k}(\text{Cone}(n, \Delta))$ the complexes

$$\begin{aligned} \text{Proj}(n, X) &:= X \cap \Delta \in \mathcal{C}_{n-1,k}(\Delta) \quad \text{and} \\ \text{Link}(n, X) &:= \{\sigma \in \Delta : \sigma \cup \{n\} \in X\} \in \mathcal{C}_{n-1,k-1}(\Delta). \end{aligned}$$

More generally, for any simplex τ of dimension at most k ,

$$\text{Link}(\tau, X) := \{\sigma \in \Delta : \sigma \cup \tau \in X\} \in \mathcal{C}_{n-|\tau|,k-|\tau|}(\Delta).$$

The next lemma defines what we will call the *binomial correspondence* $\mathcal{C}_{n,k}(\text{Cone}(n, \Delta)) \xrightarrow{\sim} \mathcal{C}_{n-1,k}(\Delta) \times \mathcal{C}_{n-1,k-1}(\Delta)$ for its relationship to the binomial recurrence formula $\binom{n}{k+1} = \binom{n-1}{k} + \binom{n-1}{k+1}$ in the special case that $\text{Cone}(n, \Delta) \supseteq K_n^k$.

Lemma 10 *For $\Delta \in \mathcal{A}_{n-1}$, the map $\varphi(X) := (\text{Proj}(n, X), \text{Link}(n, X))$ is an injection from $\mathcal{C}_{n,k}(\text{Cone}(n, \Delta))$ into $\mathcal{C}_{n-1,k}(\Delta) \times \mathcal{C}_{n-1,k-1}(\Delta)$, and its (left) inverse is given by $\varphi^{-1}(F, R) = F \cup \text{Cone}(n, R)$ which extends to an injection of $\mathcal{C}_{n-1,k}(\Delta) \times \mathcal{C}_{n-1,k-1}(\Delta)$ into $\mathcal{C}_{n,k}(\text{Cone}(n, \Delta))$.*

Proof For brevity, set $\Delta_{(j)} := K_n^j \cap \Delta$. Since φ^{-1} is clearly also a right inverse of φ , the only thing to show is that the images of these restrictions lie where we claim they do. Starting with φ , we first recall that

$$\mathcal{C}_{n,k}(\text{Cone}(n, \Delta)) = \mathcal{C}_{n,k}(\text{Cone}(n, \Delta)_{(k)}) = \mathcal{C}_{n,k}(\Delta_{(k)} \cup \{\sigma \cup \{n\} : \sigma \in \Delta_{(k-1)}\}).$$

So for $X \in \mathcal{C}_{n,k}(\text{Cone}(n, \Delta))$ we have

$$\Delta_{(k-1)} \cup \{\sigma \cup \{n\} : \sigma \in \Delta_{(k-2)}\} \subseteq X \subseteq \Delta_{(k)} \cup \{\sigma \cup \{n\} : \sigma \in \Delta_{(k-1)}\}.$$

Thus $\Delta_{(k-1)} \subseteq \text{Proj}(n, X) \subseteq \Delta_{(k)}$ and $\Delta_{(k-2)} \subseteq \text{Link}(n, X) \subseteq \Delta_{(k-1)}$ as desired. As for φ^{-1} , we have

$$\text{Cone}(n, \Delta)_{(k-1)} = \Delta_{(k-1)} \cup \text{Cone}(n, \Delta_{(k-2)}) \subseteq F \cup \text{Cone}(n, R) \subseteq \text{Cone}(n, \Delta)_{(k)}$$

due to the assumption that $(F, R) \in \mathcal{C}_{n-1,k}(\Delta) \times \mathcal{C}_{n-1,k-1}(\Delta)$. □

Theorem 11 (*Excision Theorem*) For any $A, B \in \mathcal{A}_n$, we have $H_*(A, A \cap B) \cong H_*(A \cup B, B)$.

Theorem 12 (*Long exact sequence of a triple*) For $A \subseteq B \subseteq C \in \mathcal{A}_n$, arrows exist for which the following sequence is exact:

$$\begin{array}{ccccccc} \dots & \longrightarrow & H_k(B, A) & \longrightarrow & H_k(C, A) & \longrightarrow & H_k(C, B) \\ & & & & & \searrow & \\ & & H_{k-1}(B, A) & \longrightarrow & \dots & \longrightarrow & H_0(C, B) \longrightarrow 0 \end{array}$$

That is, the kernel of each arrow in this diagram is equal to the image of the preceding arrow.

Lemma 13 For any $\Delta \in \mathcal{A}_n$, $X \in \mathcal{C}_{n,k}(\Delta)$, and $Y \in \mathcal{C}_{n,j}(X)$ where $j \leq k$, any two of the following conditions imply the other:

- $|X_k| - |Y_k| = \text{rank } \partial_k^{\Delta/Y}$,
- $\beta_{k-1}(X, Y) = \beta_{k-1}(\Delta, Y)$,
- $\beta_k(X, Y) = 0$.

Proof The long exact sequence of the triple (Δ, X, Y) contains the sequence

$$\begin{array}{ccccccc} 0 & \longrightarrow & H_{k+1}(\Delta, Y) & \longrightarrow & H_{k+1}(\Delta, X) & & \\ & & & & & \searrow & \\ & & H_k(X, Y) & \longrightarrow & H_k(\Delta, Y) & \longrightarrow & H_k(\Delta, X) \\ & & & & & \searrow & \\ & & H_{k-1}(X, Y) & \longrightarrow & H_{k-1}(\Delta, Y) & \longrightarrow & 0. \end{array}$$

The exactness of this sequence implies that

$$\beta_k(X, Y) + (\beta_{k+1}(\Delta, Y) - \beta_{k+1}(\Delta, X) - \beta_k(\Delta, Y) + \beta_k(\Delta, X)) - (\beta_{k-1}(X, Y) - \beta_{k-1}(\Delta, Y)) = 0. \tag{2}$$

It suffices to show that the first bulleted condition is equivalent to the vanishing of the middle term in (2). Indeed, by the definition of relative homology, the rank-nullity theorem, dimensional considerations, and the fact that Δ, X have the same $(k - 1)$ -skeleton, we have

$$\begin{aligned} \beta_k(\Delta, X) - \beta_{k+1}(\Delta, X) &= \dim \text{Ker } \partial_k^{\Delta/X} - (\text{rank } \partial_{k+1}^{\Delta/X} + \dim \text{Ker } \partial_{k+1}^{\Delta/X}) + \text{rank } \partial_{k+2}^{\Delta/X} \\ &= \dim \text{Ker } \partial_k^{\Delta/X} - |(\Delta/X)_{k+1}| + \text{rank } \partial_{k+2}^{\Delta/X} \\ &= |(\Delta/X)_k| - |\Delta_{k+1}| + \text{rank } \partial_{k+2}^{\Delta/X}. \end{aligned}$$

By most of the same considerations, $\beta_k(\Delta, Y) - \beta_{k+1}(\Delta, Y) = \dim \text{Ker } \partial_k^{\Delta/Y} - |\Delta_{k+1}| + \text{rank } \partial_{k+2}^{\Delta/Y}$. So

$$\begin{aligned} \beta_{k+1}(\Delta, Y) - \beta_{k+1}(\Delta, X) - \beta_k(\Delta, Y) + \beta_k(\Delta, X) &= |\Delta_k| - |X_k| - \dim \text{Ker } \partial_k^{\Delta/Y} \\ &= |Y_k| + (|(\Delta/Y)_k| - \dim \text{Ker } \partial_k^{\Delta/Y}) - |X_k| \\ &= |Y_k| + \text{rank } \partial_k^{\Delta/Y} - |X_k| \end{aligned}$$

is equal to 0 if and only if the first bulleted condition holds. The result therefore follows from (2). □

Whenever two, and therefore all three, of the conditions of this lemma hold, we will call (X, Y) a *relative (n, k) -forest of Δ* . We will denote the set of such pairs by $\mathcal{F}_{n,k}^{\text{rel}}(\Delta)$. In the special case that Y is contractible (or just that $\dim Y < k - 2$) we get that $X \in \mathcal{T}_{n,k}(\Delta)$ if and only if $(X, Y) \in \mathcal{F}_{n,k}^{\text{rel}}(\Delta)$. We therefore have the following corollary which gives the original (Kalai 1983) three equivalent necessary and sufficient pairs of conditions for a k -complex of a homologically connected Δ (i.e. $\beta_{k-1}(\Delta) = 0$) to be a k -tree of Δ :

Corollary 14 *For any $\Delta \in \mathcal{A}_n$ with $\beta_{k-1}(\Delta) = 0$ and $X \in \mathcal{C}_{n,k}(\Delta)$, any two of the following conditions imply the other, and in particular are equivalent to the statement that $X \in \mathcal{T}_{n,k}(\Delta)$:*

- $|X_k| = \binom{n-1}{k}$,
- $\beta_k(X) = 0$,
- $\beta_{k-1}(X) = 0$.

An $R \in \mathcal{C}_{n,k-1}(\Delta)$ is called a $(k - 1)$ -root of Δ if $\partial_k^{\Delta/R}$ is surjective and has full rank (Bernardi and Klivans 2015, Definition 9). Whenever any two of the conditions in the following lemma hold for a pair $(X, Y) \in \mathcal{C}_{n,k}(\Delta) \times \mathcal{C}_{n,k-1}(\Delta)$, we will call (X, Y) a *rooted (n, k) -forest of Δ* and write $(X, Y) \in \mathcal{F}_{n,k}^{\text{root}}(\Delta)$. Equivalently, a rooted (n, k) -forest of Δ is a pair $(X, Y) \in \mathcal{F}_{n,k}(\Delta) \times \mathcal{C}_{n,k-1}(\Delta)$ such that Y is a $(k - 1)$ -root of X (Bernardi and Klivans 2015, Definition 12).

Lemma 15 For $X \in \mathcal{C}_{n,k}(\Delta)$ and $Y \in \mathcal{C}_{n,k-1}(\Delta)$, any two of the following conditions imply the other:

- $|X_k| = |X_{k-1} \setminus Y_{k-1}|$,
- $\beta_{k-1}(X, Y) = 0$,
- $\beta_k(X, Y) = 0$.

Proof By Lemma 7, we have $\sum_{j=-1}^k (-1)^j (|X_j \setminus Y_j|) = \sum_{j=-1}^k (-1)^j \beta_j(X, Y)$, which simplifies to

$$|X_k| - |X_{k-1} \setminus Y_{k-1}| = \beta_k(X, Y) - \beta_{k-1}(X, Y).$$

Indeed, since $Y_j = X_j = \Delta_j$ for all $j < k - 1$, we have $\mathcal{C}_j(X, Y) = 0$ for all $j < k - 1$. As $H_j(X, Y)$ is a quotient of a subgroup of $\mathcal{C}_j(X, Y)$, the former vanishes with the latter. □

Corollary 16 For any $\Delta \in \mathcal{A}_n$, we have

$$\mathcal{F}_{n,k}^{\text{root}}(\Delta) = \left\{ (X, Y) \in \mathcal{F}_{n,k}^{\text{rel}}(\Delta) \cap (\mathcal{C}_{n,k}(\Delta) \times \mathcal{C}_{n,k-1}(\Delta)) : \beta_{k-1}(\Delta, Y) = 0 \right\}.$$

Proof Clearly $\mathcal{F}_{n,k}^{\text{root}}(\Delta) \supseteq \{(X, Y) \in \mathcal{F}_{n,k}^{\text{rel}}(\Delta) \cap (\mathcal{C}_{n,k}(\Delta) \times \mathcal{C}_{n,k-1}(\Delta)) : \beta_{k-1}(\Delta, Y) = 0\}$. For the other inclusion, suppose that $(X, Y) \in \mathcal{F}_{n,k}^{\text{root}}(\Delta)$. Then the long exact sequence in the proof of Lemma 13 implies that $0 = \beta_{k-1}(X, Y) \geq \beta_{k-1}(\Delta, Y) \geq 0$. Thus $(X, Y) \in \mathcal{F}_{n,k}^{\text{rel}}(\Delta)$ since $\beta_{k-1}(X, Y) = 0 = \beta_{k-1}(\Delta, Y)$ and $\beta_k(X, Y) = 0$. □

Lemma 17 Suppose $\Delta \in \mathcal{A}_n$ and that $F \in \mathcal{C}_{n,k}(\Delta)$, $R \in \mathcal{C}_{n,k-1}(\Delta)$ satisfy $|F_k| = |\Delta_{k-1} \setminus R_{k-1}|$, and write $\bar{R} := \Delta_{k-1} \setminus R_{k-1}$. Then $\det \partial_{\bar{R}, F_k} \neq 0 \iff |\det \partial_{\bar{R}, F_k}| = |H_{k-1}(F, R)| \iff (F, R) \in \mathcal{F}_{n,k}^{\text{root}}(\Delta)$.

Proof The relative chain complex for the pair (F, R) is $0 \longrightarrow \mathcal{C}_k(F) \xrightarrow{\partial_{\bar{R}, F_k}} \mathcal{C}_{k-1}(\bar{R}) \longrightarrow 0 \longrightarrow \dots$ since $R_k = \emptyset$, $F_{k-1} \setminus R_{k-1} = \Delta_{k-1} \setminus R_{k-1} = \bar{R}$, and $F_{k-2} = R_{k-2}$. We therefore have $H_{k-1}(F, R) = \mathbb{Z}^{\bar{R}} / \partial_{\bar{R}, F_k} \mathbb{Z}^{F_k}$ and so, provided $\det \partial_{\bar{R}, F_k} \neq 0$, we have $|H_{k-1}(F, R)| = |\det \partial_{\bar{R}, F_k}|$ —this is seen most easily by putting $\partial_{\bar{R}, F_k}$ in Smith normal form. Having $|H_{k-1}(F, R)| = |\det \partial_{\bar{R}, F_k}|$ clearly implies $|H_{k-1}(F, R)| < \infty$ which, by the previous lemma, implies that $(F, R) \in \mathcal{F}_{n,k}^{\text{root}}(\Delta)$. Finally, $(F, R) \in \mathcal{F}_{n,k}^{\text{root}}(\Delta) \implies \beta_k(F, R) = 0 \implies \text{Ker } \partial_{\bar{R}, F_k} = H_k(F, R) = 0 \implies \det \partial_{\bar{R}, F_k} \neq 0$. □

Essentially all of the content of this last lemma is proven in greater detail in Lemmas 14 and 17 of Bernardi and Klivans (2015).

Theorem 18 Set $T = F \cup \text{Cone}(n, R)$ where $F \in \mathcal{C}_{n-1,k}(\Delta)$ and $R \in \mathcal{C}_{n-1,k-1}(\Delta)$ for any $\Delta \in \mathcal{A}_{n-1}$. Then $\tilde{H}_*(T) \cong H_*(F, R)$, and $T \in \mathcal{T}_{n,k}(\text{Cone}(n, \Delta))$ if and only if $(F, R) \in \mathcal{F}_{n-1,k}^{\text{root}}(\Delta)$. Moreover the binomial correspondence restricts to a bijection $\varphi : \mathcal{T}_{n,k}(\text{Cone}(n, \Delta)) \xrightarrow{\sim} \mathcal{F}_{n-1,k}^{\text{root}}(\Delta)$.

Proof Since F has no k -faces containing n and $R \subseteq F$, we have $F \cap \text{Cone}(n, R) = R$. Thus, by excision, the pairs $(T, \text{Cone}(n, R))$ and (F, R) have the same relative homology in every dimension. But, since $\text{Cone}(n, R)$ is contractible, $\tilde{H}_*(T) \cong H_*(F, R)$. Therefore $T \in \mathcal{T}_{n,k}(\text{Cone}(n, \Delta))$ if and only if $(F, R) \in \mathcal{F}_{n-1,k}^{\text{root}}(\Delta)$ by Lemmas 9 and 15. The bijection then follows from Lemma 10. \square

Corollary 19 *Set $T = F \cup \text{Cone}(n, R)$, where $F \in \mathcal{C}_{n-1,k}$ and $R \in \mathcal{C}_{n-1,k-1}$. Then $\tilde{H}_*(T) \cong H_*(F, R)$, and $T \in \mathcal{T}_{n,k}$ if and only if $(F, R) \in \mathcal{F}_{n-1,k}^{\text{root}}$. Moreover the binomial correspondence restricts to a bijection $\varphi : \mathcal{T}_{n,k} \xrightarrow{\sim} \mathcal{F}_{n-1,k}^{\text{root}}$.*

Proof This follows by recalling that $\mathcal{T}_{n,k}(\text{Cone}(n, \Delta)) = \mathcal{T}_{n,k}((\Delta \cap K_{n-1}^k) \cup \{\sigma \cup \{n\} : \sigma \in \Delta \cap K_{n-1}^{k-1}\})$. Thus if $\Delta \supseteq K_{n-1}^k$, then $(\Delta \cap K_{n-1}^k) \cup \{\sigma \cup \{n\} : \sigma \in \Delta \cap K_{n-1}^{k-1}\} = K_n^k$. \square

3 Determinantal measure

In this section we will narrow our focus to the case $\Delta \supseteq K_n^k$ and begin considering the probabilistic aspects of k -trees and rooted k -forest of such a Δ . Specifically, after defining the measure $\nu_{n,k}$ mentioned in the introduction, we will show (Corollary 21) that it has an alternative expression in terms of Proj and Link. Following a brief exploration of determinantal probability measures during which we derive the containment probabilities for ν (Lemma 22), we leverage our alternative expression for $\nu_{n,k}$ to determine the probability densities for Proj $\mathcal{T}_{n,k}$ and Link $\mathcal{T}_{n,k}$ in terms of of the containment probabilities we just found for ν (Lemma 23). We then use this to prove Theorem 1.

For ease of reading, we will for this section abuse notation by identifying each j -complex of K_n^{n-1} with its set of j -dimensional faces. For any $X \in \mathcal{C}_{n,j}$, we also will let \bar{X} denote the j -complex that its set of j -faces is the complement (with respect to $\binom{[n]}{j+1}$) of X_j . Define the submatrix $\hat{\partial}$ of ∂ by deleting all rows of ∂ that correspond to elements of $\binom{[n]}{k}$ which contain the vertex n (thus $\hat{\partial}_0^{[n]} = \partial_0^{[n]}$). We have the following corollary of Lemma 17:

Corollary 20 *Suppose $T \in \mathcal{C}_{n,k}$ satisfies $|T_k| = \binom{n-1}{k}$. Then $\det \hat{\partial}_{\bullet,T} \neq 0 \iff |\det \hat{\partial}_{\bullet,T}| = |\tilde{H}_{k-1}(T)| \iff T \in \mathcal{T}_{n,k}$.*

Proof Let $R = \text{Cone}(n, K_{n-1}^{k-2})$ so that $|T_k| + |R_{k-1}| = \binom{n}{k}$, and $\partial_{\bar{R},T} = \hat{\partial}_{\bullet,T}$. This now follows from Corollary 14 and Lemma 17. \square

This last corollary was originally proven for the cases $k \geq 2$ (Kalai 1983, Lemma 2) by Kalai, who combined this with the Cauchy–Binet formula and some deft linear algebra to show (Kalai 1983, Theorem 1) that

$$n \binom{n-2}{k} = \det \hat{\partial} \hat{\partial}^t = \sum_{T \in \mathcal{T}_{n,k}} \det \hat{\partial}_{\bullet,T}^2 = \sum_{T \in \mathcal{T}_{n,k}} |\tilde{H}_{k-1}(T)|^2.$$

Recalling Caley’s formula and understanding that $\widehat{\partial}_0^{[n]} = \partial_0^{[n]}$ and $\widetilde{H}_{-1}(T) = 0$, the cases $k = 0, 1$ are also seen to hold. This gives us a natural probability measure $\nu = \nu_{n,k}$ on $\mathcal{T}_{n,k}$. Namely,

$$\nu_{n,k}(T) := \frac{\det \widehat{\partial}_{\bullet,T}^2}{\det \widetilde{\partial \partial^t}} = \frac{|\widetilde{H}_{k-1}(T)|^2}{n^{\binom{n-2}{k}}}.$$

This measure was originally formulated in greater generality by Lyons (Lyons (2003), Section 12) and was further expanded upon in Lyons (2009). This version of the measure has also been considered again recently by authors such as Kahle and Newman (2022) and Mészáros (2022).

Corollary 21 *Suppose $T \in \mathcal{T}_{n,k}$, and let $F = \text{Proj}(n, T)$ and $E = \text{Link}(n, T)$. Then*

$$\nu_{n,k}(T) = \nu_{n,k}(F \cup \text{Cone}(n, E)) = \frac{\det \partial_{E,F}^2}{n^{\binom{n-2}{k}}} = \frac{|H_{k-1}(F, E)|^2}{n^{\binom{n-2}{k}}}.$$

Proof The first equality follows from Lemma 10. The rest follows from Corollary 19, the original definition $\nu_{n,k}(T) = \frac{|\widetilde{H}_{k-1}(T)|^2}{n^{\binom{n-2}{k}}}$, and Lemma 17. □

Probability measures defined in the above manner are said to be *determinantal* (Lyons 2003), as are the random variables associated to them. The following lemma is a special case of Theorem 5.1 from Lyons (2003):

Lemma 22 *Let R, S be finite sets and let M be an $R \times S$ matrix of rank $|R|$. Let μ be the determinantal measure on S which is defined by $\mu(T) = \frac{\det M_{\bullet,T}^2}{\det(MM^t)}$ for all $T \subseteq S$ of size $|R|$. Let $P := M^t(MM^t)^{-1}M$ (this is the matrix of the projection onto the rowspace of M). Then, for any $B \subseteq S$,*

$$\mu(T : T \supseteq B) = \det P_{B,B} \quad \text{and} \quad \mu(T : T \subseteq S \setminus B) = \det(\text{Id} - P)_{B,B}.$$

Determinantal random variables enjoy the negative associations property (Lyons 2003, Theorem 6.5), which can be stated for random variables on $\mathcal{C}_{n,k}$ as follows: A function $f : \mathcal{C}_{n,k} \rightarrow \mathbb{R}$ is called *increasing* if $f(X) \leq f(X \cup Y)$ for every $X, Y \in \mathcal{C}_{n,k}$. The k -faces of a random variable $\mathcal{X} \in \mathcal{C}_{n,k}$ are said to be *negatively associated* if, for every pair of increasing functions f_1, f_2 and every $Y \in \mathcal{C}_{n,k}$, we have

$$\mathbb{E}[f_1(\mathcal{X} \cap Y) f_2(\mathcal{X} \cap \bar{Y})] \leq \mathbb{E}[f_1(\mathcal{X} \cap Y)] \mathbb{E}[f_2(\mathcal{X} \cap \bar{Y})].$$

We will make use of negative associations when we go to prove our applications to random topology.

As we see from Lemma 22, for any $B \in \mathcal{C}_{n,k}$ and $A \in \mathcal{C}_{n,k-1}$, we have

$$\nu_{n,k}(T : T \supseteq B) = \det(P_{n,k})_{B,B} \quad \text{and} \quad \nu_{n,k-1}(T : T \subseteq \bar{A}) = \det(\text{Id} - P_{n,k-1})_{A,A} \tag{3}$$

where $P_{n,k} := \widehat{\partial}^t (\widehat{\partial} \widehat{\partial}^t)^{-1} \widehat{\partial}$. Our use of $\widehat{\partial}$ is actually a choice of basis for the row space of ∂ , and an arbitrary one at that. Let A be any $(k - 1)$ -root of K_n^{n-1} —this implies that $|A_{k-1}| = \binom{n-1}{k}$ (Kalai 1983, Lemma 1). Then, by applying a change of basis, we also have the equivalent definition(s)

$$v_{n,k}(T) := \frac{\det \partial_{A,T}^2}{\det(\partial \partial^t)_{A,A}},$$

all of which correspond to the same $P_{n,k}$. Mészáros (2022, Lemma 14) determined that

$$P_{n,k} := \frac{1}{n} \partial_k^{[n]t} \partial_k^{[n]}.$$

By Lemma 8, we also have $\text{Id} - P_{n,k} = \frac{1}{n} \partial_{k+1}^{[n]} \partial_{k+1}^{[n]t} =: Q_{n,k}$. By Lemma 17 and Cauchy–Binet, we therefore have the following lemma:

Lemma 23 *Suppose $B \in \mathcal{C}_{n,k}$ and $A \in \mathcal{C}_{n,k-1}$. Then*

$$v_{n,k-1}(T : T \subseteq \overline{A}) = \det(Q_{n,k-1})_{A,A} = n^{-|\overline{A}|} \sum_{B:(B,\overline{A}) \in \mathcal{T}_{n,k}^{\text{root}}} \det \partial_{A,B}^2,$$

and $v_{n,k}(T : T \supseteq B) = \det(P_{n,k})_{B,B} = n^{-|B|} \sum_{A:(B,A) \in \mathcal{T}_{n,k}^{\text{root}}} \det \partial_{A,B}^2.$

Lemma 24 *For all $F \in \mathcal{C}_{n-1,k}$ and $E \in \mathcal{C}_{n-1,k-1}$, we have*

$$v_{n,k}(T : \text{Proj}(n, T) = F) = v_{n-1,k}(T' : T' \supseteq F) (1 - 1/n)^{|F|} n^{|F| - \binom{n-2}{k}}$$

and

$$v_{n,k}(T : \text{Link}(n, T) = E) = v_{n-1,k-1}(T'' : T'' \subseteq E) (1 - 1/n)^{|\overline{E}|} n^{|\overline{E}| - \binom{n-2}{k}}.$$

Proof By Lemma 23, Corollary 21, and the Cauchy–Binet formula, we have

$$(n - 1)^{|F|} v_{n-1,k}(T : T \supseteq F) = \det(\partial^t \partial)_{F,F} = n^{\binom{n-2}{k}} v_{n,k}(T : \text{Proj}(n, T) = F)$$

and

$$(n - 1)^{|\overline{E}|} v_{n-1,k-1}(T : T \subseteq E) = \det(\partial \partial^t)_{\overline{E},\overline{E}} = n^{\binom{n-2}{k}} v_{n,k}(T : \text{Link}(n, T) = E)$$

as desired. □

Theorem 1 makes the claim that there exist certain relationships between the laws of certain random variables. An equivalent way of saying this is that there are couplings of these random variables for which these relationships hold with probability 1. In this sense, the following corollary is just a more detailed restatement of Theorem 1:

Corollary 25 *Using the notation from Theorem 1, there are couplings $\pi_{n,k}$ and $\lambda_{n,k}$ of $\mathcal{T}_{n,k}, \mathcal{T}_{n-1,k}, \mathcal{Y}_{n-1,k}$ and $\mathcal{T}_{n,k}, \mathcal{T}_{n-1,k-1}, \mathcal{Y}_{n-1,k-1}$ respectively such that, marginally, $\mathcal{T}_{n-1,k}, \mathcal{T}_{n-1,k-1}$ are independent of $\mathcal{Y}_{n-1,k}, \mathcal{Y}_{n-1,k-1}$ respectively, and*

$$\begin{aligned} & \pi_{n,k}((T, T', Y') : \text{Proj}(n, T) = T' \setminus Y') \\ & = \lambda_{n,k}((T, T'', Y'') : \text{Link}(n, T) = T'' \cup Y'') = 1. \end{aligned}$$

Namely,

$$\pi_{n,k}(T, T', Y') := \mu_{n-1,k}(Y')v_{n-1,k}(T') \frac{v_{n,k}(T)\mathbf{1}\{\text{Proj}(n, T) = T' \setminus Y'\}}{v_{n,k}(S : \text{Proj}(n, S) = T' \setminus Y')}$$

and

$$\lambda_{n,k}(T, T'', Y'') := \mu_{n-1,k-1}(Y'')v_{n-1,k-1}(T'') \frac{v_{n,k}(T)\mathbf{1}\{\text{Link}(n, T) = T'' \cup Y''\}}{v_{n,k}(S : \text{Link}(n, S) = T'' \cup Y'')}.$$

Proof It suffices to show that $\pi_{n,k}$ has the correct marginal densities, as the proof for $\lambda_{n,k}$ is practically identical. Summing $\pi_{n,k}$ over all T clearly produces the independent coupling of $\mathcal{T}_{n-1,k}, \mathcal{Y}_{n-1,k}$. For the remaining marginal, Lemma 24 gives us that

$$\begin{aligned} & \frac{\mathbf{1}\{\text{Proj}(n, T) = T' \setminus Y'\}}{v_{n,k}(S : \text{Proj}(n, S) = T' \setminus Y')} \\ & = \frac{\mathbf{1}\{T' \supseteq \text{Proj}(n, T)\}\mathbf{1}\{\text{Proj}(n, T) = T' \setminus Y'\}}{v_{n-1,k}(S : S \supseteq T' \setminus Y')(1 - 1/n)^{|T' \setminus Y'|}n^{|T' \setminus Y'| - \binom{n-k}{k}}} \\ & = \frac{\mathbf{1}\{T' \supseteq \text{Proj}(n, T)\}\mathbf{1}\{T' \setminus Y' = \text{Proj}(n, T)\}}{v_{n-1,k}(S : S \supseteq T' \setminus Y')(1 - 1/n)^{|\text{Proj}(n, T)|}n^{-|\text{Proj}(n, T)|}} \\ & = \frac{\mathbf{1}\{T' \supseteq \text{Proj}(n, T)\}\mathbf{1}\{T' \setminus Y' = \text{Proj}(n, T)\}}{v_{n-1,k}(S : S \supseteq T' \setminus Y')\mu_{n-1,k}(S : T' \setminus S = \text{Proj}(n, T))}, \end{aligned}$$

where in the last line we used that $\mu_{n-1,k}(S : T' \setminus S = \text{Proj}(n, T))$ is the probability that $|\text{Proj}(n, T)|$ faces from T' were not chosen to be included in $\mathcal{Y}_k(n - 1, 1/n)$, which occurs with probability $(1 - 1/n)^{|\text{Proj}(n, T)|}$, and $|T' \setminus \text{Proj}(n, T)|$ faces were chosen to be included in $\mathcal{Y}_k(n - 1, 1/n)$, which occurs with probability $n^{-|T' \setminus \text{Proj}(n, T)|}$. Thus

$$\pi_{n,k}(T, T', Y') = v_{n,k}(T) \frac{v_{n-1,k}(T')\mathbf{1}\{T' \supseteq \text{Proj}(n, T)\}}{v_{n-1,k}(S : S \supseteq \text{Proj}(n, T))} \frac{\mu_{n-1,k}(Y')\mathbf{1}\{T' \setminus Y' = \text{Proj}(n, T)\}}{\mu_{n-1,k}(S : T' \setminus S = \text{Proj}(n, T))}$$

which, summed over Y' and then T' , gives the desired marginal. □

4 Spectral estimates

All of this section is in service of proving our applications to random topology, Proposition 5 and Corollary 6. Lemma 36 at the end of this section is sufficient to do just that. Corollary 27 allows us to prove Lemma 36 by connecting it to a different problem which is solved by the Kahn–Szemerédi argument. In Sects. 4.1 and 4.2 below, we go into detail about this argument and generalize it to a slightly larger class of random graph that includes the kinds described in Sect. 4.3 which are used to prove Lemma 36.

Before we can do this any of this, we need to introduce notation for this slightly larger class of random graphs and find suitable analogs for the inequalities of Bernstein and Chernoff which are necessary for the Kahn–Szemerédi argument.

Let A be the adjacency matrix of a random graph $\mathcal{G} = \mathcal{G}(n, p, E, M)$ on $[n]$ satisfying the following:

1. $\mathbb{P}[e \in \mathcal{G}] = p \in (0, 1)$ for all $e \in \binom{[n]}{2}$.
2. There is a fixed constant $E \geq 1$ and an $M = M(n) > 0$ such that, for every $t \in [0, M]^{\binom{[n]}{2}}$, we have

$$\mathbb{E} \exp \left(\sum_{1 \leq i < j \leq n} t_{ij} A_{ij} \right) \leq E \prod_{1 \leq i < j \leq n} (1 - p + pe^{t_{ij}}).$$

The choice to make E a fixed constant is just for convenience. All of the results of this section can be easily adapted to work for $E = E(n)$ an arbitrary fixed polynomial in the variable n .

For this section, we use asymptotic notation, $o()$ and $O()$, to describe the behavior of a function of n as $n \rightarrow \infty$.

Proposition 26 *Let \mathcal{G} be defined as above with $p = \frac{\delta \log n}{n}$ (δ an arbitrary constant), and $M \geq \frac{3}{5}(n/p)^{\frac{1}{2}}$. Then, for any fixed $s > 0$, we have with probability at least $1 - o(n^{-s})$ that*

$$v^t Av = O(\sqrt{np}) \text{ for all unit vectors } v \perp \mathbf{1}.$$

Proof This will follow from Lemmas 30, 31, 32, and 33. □

Corollary 27 *Let $L := \text{Id} - D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ where D is the degree matrix of $\mathcal{G} = \mathcal{G}\left(n, \frac{\delta \log n}{n}, O(1), \frac{3n}{5\sqrt{\delta \log n}}\right)$. Suppose there are positive integers $s = O(1)$ and $m = m(n)$ such that $\mathbb{P}[\min_{i \in [n]} \deg_{\mathcal{G}}(i) \geq m] = 1 - o(n^{-s})$. Let $\lambda_1(L) \leq \lambda_2(L) \leq \dots \leq \lambda_n(L)$ denote the eigenvalues of L . Then, with probability $1 - o(n^{-s})$, we have $\lambda_1(L) = 0$ and*

$$|\lambda_2(L) - 1| = O\left(\frac{\sqrt{\log n}}{m}\right).$$

Proof Since \mathcal{G} is connected with sufficiently high probability, we will treat \mathcal{G} as though it were connected almost surely. As such, we know that L has minimal eigenvalue 0 with multiplicity one, and the corresponding eigenvector is $D^{\frac{1}{2}}\mathbf{1}$. Thus we are interested in bounding the quantity

$$\sup \left\{ \frac{y^t D^{-\frac{1}{2}} A D^{-\frac{1}{2}} y}{y^t y} : 0 \neq y \perp D^{\frac{1}{2}}\mathbf{1} \right\}$$

from above by some $\lambda := O(\sqrt{\log n/m})$. Equivalently, we would like to show that

$$x^t A x \leq \lambda x^t D x \text{ for all } x \perp D\mathbf{1}.$$

Without loss of generality, we can assume that x is a unit vector. Let

$$x = \cos \theta u + \sin \theta v \text{ where } u = \frac{1}{\sqrt{n}}\mathbf{1}, v \perp \mathbf{1}, \text{ and } |v| = 1.$$

Noting that $u^t A x = u^t D x = 0$ and $\cos \theta \frac{\text{tr} D}{\sqrt{n}} = -\sin \theta v^t D\mathbf{1}$ (both of these follow from the assumption that $x \perp D\mathbf{1}$), we have

$$x^t A x = \sin^2 \theta v^t A v - \cos^2 \theta \frac{\text{tr} D}{n} \text{ and } x^t D x = \sin^2 \theta v^t D v - \cos^2 \theta \frac{\text{tr} D}{n}.$$

So we have

$$\begin{aligned} x^t(\lambda D - A)x &= v^t(\lambda D - A)v \sin^2 \theta + \frac{(1 - \lambda) \text{tr} D}{n} \cos^2 \theta \\ &= v^t(\lambda D - A)v \frac{(\text{tr} D)^2}{(\text{tr} D)^2 + n(v^t D\mathbf{1})^2} + \frac{(1 - \lambda) \text{tr} D}{n} \frac{n(v^t D\mathbf{1})^2}{(\text{tr} D)^2 + n(v^t D\mathbf{1})^2} \end{aligned}$$

which we would like to show is positive. Solving for $v^t A v$, it suffices to show that

$$v^t A v \leq (1 - \lambda) \frac{(v^t D\mathbf{1})^2}{\text{tr} D} + \lambda v^t D v \text{ for all unit vectors } v \perp \mathbf{1}.$$

By our minimum degree assumption, it would even suffice to show that $v^t A v \leq \lambda m$. The result now follows from Proposition 26. □

In order to prove Lemma 31, we are going to need a version of Bernstein’s inequality that works on weighted sums of centered edge indicators of \mathcal{G} .

Theorem 28 (Bernstein’s inequality) *Let \mathcal{G} be as above. Suppose $|c_{ij}| \leq c$ for some fixed c and all $i < j$, and that $\varepsilon \geq 0$ is such that $M \geq \frac{\varepsilon}{p \sum_{i < j} c_{ij}^2 + 2c\varepsilon/3}$. Then, for any*

$H \subseteq \binom{[n]}{2}$, we have

$$\mathbb{P} \left[\sum_{(i,j) \in H} c_{ij}(A_{ij} - p) \geq \varepsilon \right] \leq E \exp \left(-\frac{\varepsilon^2}{2p \sum_{(i,j) \in H} c_{ij}^2 + 2c\varepsilon/3} \right).$$

Proof By a Chernoff bound, the numerical bounds $1 + x \leq e^x \leq 1 + x + \frac{3x^2}{6-2x}$ for $x \leq 3$, and our assumptions about \mathcal{G} , we have for $t \leq \frac{3}{c}$ that

$$\begin{aligned} \mathbb{P} \left[\sum_{(i,j) \in H} c_{ij}(A_{ij} - p) \geq \varepsilon \right] &\leq E \inf_{t \in (0, M]} e^{-t\varepsilon} \prod_{(i,j) \in H} \left(1 + \mathbb{E} \frac{3t^2 c_{ij}^2 (A_{ij} - p)^2}{6 - 2ct} \right) \\ &\leq E \inf_{t \in (0, M]} \exp \left(-t\varepsilon + \sum_{(i,j) \in H} \mathbb{E} \frac{3t^2 c_{ij}^2 (A_{ij} - p)^2}{6 - 2ct} \right) \\ &\leq E \inf_{t \in (0, M]} \exp \left(\frac{(p \sum_{(i,j) \in H} c_{ij}^2 + 2c\varepsilon/3)t^2 - 2\varepsilon t}{2 - 2ct/3} \right) \\ &= E \inf_{t \in (0, M]} \exp \left(\frac{(\sigma^2 + 2c\varepsilon/3)t^2 - 2\varepsilon t}{2 - 2ct/3} \right), \end{aligned}$$

where $\sigma^2 := p \sum_{(i,j) \in H} c_{ij}^2$. Taking $t = \frac{\varepsilon}{\sigma^2 + 2c\varepsilon/3}$, we have $2 - 2ct/3 = \frac{2\sigma^2 + 2c\varepsilon/3}{\sigma^2 + 2c\varepsilon/3}$, and thus

$$\begin{aligned} \frac{(\sigma^2 + 2c\varepsilon/3)t^2 - 2\varepsilon t}{2 - 2ct/3} &= ((\sigma^2 + 2c\varepsilon/3)t - 2\varepsilon) \frac{\varepsilon}{\sigma^2 + 2c\varepsilon/3} \frac{\sigma^2 + 2c\varepsilon/3}{2\sigma^2 + 2c\varepsilon/3} \\ &= -\frac{\varepsilon^2}{2\sigma^2 + 2c\varepsilon/3}. \end{aligned}$$

Recalling that we were required to assume that $t \leq \frac{3}{c}$, what we have shown holds for our choice of t as long as $\varepsilon \geq -3\sigma^2 c^{-1}$. □

We will also want to be able to apply a near-optimal Chernoff bound to uniformly weighted sums of edge indicators.

Lemma 29 For \mathcal{G} as above with $M \geq \log \frac{(1-p)\varepsilon}{1-p\varepsilon}$ where $\varepsilon \geq 3$, we have

$$\mathbb{P} \left[\sum_{(i,j) \in H} A_{ij} \geq \varepsilon p |H| \right] \leq E \exp \left(-\frac{\varepsilon \log \varepsilon}{3} p |H| \right).$$

Proof As with the previous proof,

$$\begin{aligned} \mathbb{P} \left[\sum_{(i,j) \in H} A_{ij} \geq \varepsilon p |H| \right] &\leq E \inf_{t \in (0, M]} e^{-t\varepsilon p |H|} (1 - p + pe^t)^{|H|} \\ &= E \left(\frac{(1 - p)\varepsilon}{1 - p\varepsilon} \right)^{-\varepsilon p |H|} \left(\frac{1 - p}{1 - p\varepsilon} \right)^{|H|} \\ &= E \varepsilon^{-\varepsilon p |H|} \left(1 + \frac{(\varepsilon - 1)p}{1 - p\varepsilon} \right)^{(1 - p\varepsilon)|H|} \\ &\leq E \exp(-(\varepsilon \log \varepsilon - \varepsilon + 1)p |H|). \end{aligned}$$

For $\varepsilon \geq 3$, this is bounded above by $E \exp\left(-\frac{\varepsilon \log \varepsilon}{3} p |H|\right)$. □

To prove Proposition 26, we will adapt the Kahn–Szemerédi argument. Let

$$\begin{aligned} S &:= \{v \in \mathbb{R}^n : |v| = 1 \text{ and } v \perp \mathbf{1}\} \text{ and} \\ T &:= \left\{ x \in \frac{1}{\sqrt{4n}} \mathbb{Z}^n : |x| \leq 1 \text{ and } x \perp \mathbf{1} \right\}. \end{aligned}$$

The following lemma is a special case of Claim 2.4 in Feige and Ofek (2005).

Lemma 30 *Suppose $|x^t A y| \leq c$ for all $x, y \in T$. Then $|v^t A v| \leq 4c$ for all $v \in S$.*

We now write

$$\sum_{(i,j) \in [n]^2} |x_i A_{ij} y_j| = \sum_{(i,j) \in \mathcal{L}} |x_i A_{ij} y_j| + \sum_{(i,j) \in \mathcal{H}} |x_i A_{ij} y_j|$$

where $\mathcal{L} := \{(i, j) \in [n]^2 : (x_i y_j)^2 \leq \frac{\rho}{n}\}$ and $\mathcal{H} := [n]^2 \setminus \mathcal{L}$.

4.1 Light couples

Lemma 31 *Suppose $M \geq \frac{3}{5}(n/p)^{\frac{1}{2}}$ in the definition of \mathcal{G} . For any constant $s > 0$, we have*

$$\mathbb{P} \left[\sum_{(i,j) \in \mathcal{L}} |x_i A_{ij} y_j| \geq 7\sqrt{np} \text{ for some } x, y \in T \right] \leq E(18e^{-3})^n = o(n^{-s}).$$

Proof We can bound the contribution from the light couples as follows: It is known (Feige and Ofek 2005, Claim 2.9) that $|T| \leq 18^n$. So, by applying a union bound, we have

$$\mathbb{P} \left[\sum_{(i,j) \in \mathcal{L}} |x_i A_{ij} y_j| \geq 7\sqrt{np} \text{ for some } x, y \in T \right]$$

$$\leq 18^n \sup_{x, y \in T} \mathbb{P} \left[\sum_{(i, j) \in \mathcal{L}} |x_i A_{ij} y_j| \geq 7\sqrt{np} \right].$$

Towards applying Bernstein’s inequality, define centered random variables

$$B_{ij} := (|x_i y_j| \mathbf{1}\{(i, j) \in \mathcal{L}\} + |x_j y_i| \mathbf{1}\{(j, i) \in \mathcal{L}\}) (A_{ij} - p)$$

so that

$$\sum_{\{(i, j): (i, j) \in \mathcal{L} \text{ or } (j, i) \in \mathcal{L}\}} B_{ij} = \sum_{(i, j) \in \mathcal{L}} |x_i A_{ij} y_j| - \mathbb{E} \sum_{(i, j) \in \mathcal{L}} |x_i A_{ij} y_j|.$$

Note that each $B_{ij}^2 \leq \frac{4p}{n}$ a.s., and $\sum_{i < j} \mathbb{E} B_{ij}^2 \leq p \sum_{i < j} 2((x_i y_j)^2 + (x_j y_i)^2) \leq 2p$ by taking advantage of the fact that $|x|, |y| \leq 1$. Thus, by Bernstein’s inequality,

$$\begin{aligned} \mathbb{P} \left[\sum_{\{(i, j): (i, j) \in \mathcal{L} \text{ or } (j, i) \in \mathcal{L}\}} B_{ij} \geq 6\varepsilon\sqrt{np} \right] &\leq E \exp \left(-\frac{(6\varepsilon)^2 np}{4p + 6\frac{4\sqrt{p}}{3\sqrt{n}}\sqrt{np}\varepsilon} \right) \\ &= E \exp \left(\frac{-9\varepsilon^2}{1 + 2\varepsilon} n \right). \end{aligned}$$

Taking $\varepsilon = 1$, this shows that $\sum_{(i, j) \in \mathcal{L}} |x_i A_{ij} y_j| < 6\sqrt{np} + \mathbb{E} \sum_{(i, j) \in \mathcal{L}} |x_i A_{ij} y_j|$ with probability at least $1 - Ee^{-3n}$. By Lemma 2.6 of Feige and Ofek (2005), we also have $\mathbb{E} \sum_{(i, j) \in \mathcal{L}} |x_i A_{ij} y_j| \leq \sqrt{np}$, thus giving us the desired bound. Our use of Bernstein’s inequality requires us to have $M \geq \frac{6\sqrt{np}}{2p + \frac{8\sqrt{p}}{\sqrt{n}}\sqrt{np}} = \frac{3}{5}(n/p)^{\frac{1}{2}}$. \square

4.2 Heavy couples

For $B, C \subseteq [n]$, let $e(B, C) := |\{(i, j) \in \mathcal{G} : i \in B, j \in C\}|$ and $\mu(B, C) := p|B||C|$. The following is a weakened form of Lemma 9.1 in Hoffman et al. (2019).

Lemma 32 *Suppose we have constants $c_0, c_1, c_2 > 1$ and a graph G on $[n]$ with $\max_{i \in [n]} \deg_G(i) \leq c_0 np$, and, for all $B, C \subseteq [n]$, one or more of the following hold:*

- $e(B, C) \leq c_1 \mu(B, C)$
- $e(B, C) \log \frac{e(B, C)}{\mu(B, C)} \leq c_2 (|B| \vee |C|) \log \frac{n}{|B| \vee |C|}$

Then $\sum_{(i, j) \in \mathcal{H}} |x_i A_{ij} y_j| = O(\sqrt{np})$.

Lemma 33 *Suppose that $p = \frac{\delta \log n}{n}$ for some fixed but arbitrary $\delta > 0$, and $M \geq \frac{3}{5}(n/p)^{\frac{1}{2}}$. For any $s > 0$, there are fixed constants $c_0, c_1, c_2 > 1$ so that the conditions of Lemma 32 hold for \mathcal{G} with probability at least $1 - o(n^{-s})$.*

In terms of probabilistic bounds, the proof of this will only rely on Lemma 29. We note then that the condition $M \geq \frac{3}{5}(n/p)^{\frac{1}{2}}$ is overkill since Lemma 29 only ever requires that we have M be greater than a fixed constant.

Proof Assume without loss of generality that $c_0 \geq 3$. First note that, due to the monotonicity of $x \log x$ for $x \geq 1$, the second bulleted condition is equivalent to the statement $e(B, C) \leq r_1 \mu(B, C)$ where $r_1 \geq 1$ solves $r_1 \log r_1 = \frac{c_2(|B| \vee |C|) \log \frac{n}{|B| \vee |C|}}{\mu(B, C)}$. So we can rewrite the two bulleted conditions as the single condition $e(B, C) \leq r \mu(B, C)$ where $r := r_1 \vee c_1$. By Lemma 29 and a union bound, we have

$$\mathbb{P}[\max_{i \in [n]} \deg_G(i) > c_0 np] \leq n \mathbb{P}[\deg_G(n) \geq c_0 np] \leq En \exp\left(-\frac{\delta c_0 \log c_0}{3} \log n\right)$$

and

$$\begin{aligned} \mathbb{P}[e(B, C) > r \mu(B, C)] &\leq \mathbb{P}[e(B, C) > r_1 \mu(B, C)] \\ &\leq E \exp\left(-\frac{c_2(|B| \vee |C|) \log \frac{n}{|B| \vee |C|}}{3}\right). \end{aligned}$$

Thus we can choose a constant c_0 large enough to make $\mathbb{P}[\max_{i \in [n]} \deg_G(i) > c_0 np] = o(n^{-s})$. Using this fact and the resulting (high probability) inequality

$$\frac{e(B, C)}{\mu(B, C)} \leq \frac{c_0(|B| \wedge |C|)np}{p|B||C|} = \frac{c_0 n}{|B| \vee |C|},$$

we have, in the case $|B| \vee |C| \geq n/e$, that

$$\begin{aligned} \mathbb{P}[\exists B, C \text{ s.t. } e(B, C) > ec_0 \mu(B, C) \text{ and } |B| \vee |C| \geq n/e] \\ \leq \mathbb{P}[\max_{i \in [n]} \deg_G(i) > c_0 np] = o(n^{-s}). \end{aligned}$$

So suppose that $|B| \vee |C| < n/e$. By a union bound over all possible pairs $\{i, j\} \subset [[n/e]]$ ($i \leq j$, without loss of generality) of sizes for the sets B, C , it suffices to show that

$$\binom{n}{i} \binom{n}{j} \exp\left(-\frac{c_2 j \log \frac{n}{j}}{3}\right) \leq n^{-s-3}.$$

Recalling that $\binom{n}{j} \leq \left(\frac{ne}{j}\right)^j$ for all $j \in [n]$, this can be done by showing that

$$(s + 3) \log n + j \left(1 + \log \frac{n}{j}\right) + i \left(1 + \log \frac{n}{i}\right) \leq \frac{c_2}{3} j \log \frac{n}{j}$$

for all $1 \leq i \leq j < n/e$. Indeed, since $j \log \frac{n}{j}$ is monotone increasing for $1 \leq j < n/e$ (Feige and Ofek 2005, Lemma 2.12), we have

$$\begin{aligned} &(s + 3) \log n + j \left(1 + \log \frac{n}{j} \right) + i \left(1 + \log \frac{n}{i} \right) \\ &\leq (s + 3) \log n + 4j \log \frac{n}{j} \leq (s + 7)j \log \frac{n}{j}. \end{aligned}$$

Taking $c_2 \geq 3s + 21$ therefore gives the desired bound. □

4.3 Link unions

Let T_1, T_2, \dots, T_m be jointly independent copies of $\mathcal{T}_{n,k-j}$ with k fixed. Then, if T'_1, T'_2, \dots, T'_m are jointly independent copies of $\mathcal{T}_{n+j,k}$, we can use Theorem 2 to couple these random trees so that

$$\begin{aligned} L_m^{k,j}(n) &:= \text{Link} \left(\{n + 1, \dots, n + j\}, \bigcup_{i \in [m]} T'_i \right) \\ &= \bigcup_{i \in [m]} \text{Link} (\{n + 1, \dots, n + j\}, T'_i) \\ &= \mathcal{Y}_{k-j}(n, p) \cup \bigcup_{i \in [m]} T_i, \end{aligned}$$

where $p := 1 - \left(1 - \frac{j}{n+j} \right)^m \sim \frac{jm}{n}$. Since each T_i is determinantal, the k -faces of each T_i are negatively associated (to be abbreviated NA)—see the discussion following Lemma 22 for the definition of NA. Moreover, the k -faces in $T := \bigcup_{i \in [m]} T_i$ are also NA, as

$$\mathbf{1}\{f \in T\} = 1 - \prod_{i \in [m]} (1 - \mathbf{1}\{f \in T_i\})$$

is increasing in T as a function $\mathcal{C}_{n,k} \rightarrow \mathbb{R}$, and any set of increasing functions evaluated over disjoint subsets of a set of NA objects is itself a set of NA objects (Joag-Dev and Proschan 1983). By the same reasoning and the fact that sets of jointly independent random variables are NA sets, we also have that the set of $(k - j)$ -faces of $L_m^{k,j}(n)$ is NA. We now narrow our focus to the case $j = k - 1$, in which

$$G := L_m^{k,k-1}(n)$$

is a graph with negatively associated edges each appearing with probability $q := 1 - (1 - p)(1 - \frac{2}{n})^m \sim \frac{(k+1)m}{n}$. This implies that G is of type $\mathcal{G}(n, q, 1, \infty)$. Toward applying Corollary 27 to this G , we will assume that $m = \lceil \delta \log n \rceil$ for some arbitrary constant $\delta > 0$. In order to get a sufficiently strong lower bound on the minimum

degree of G , we need to have a very strong bound on the moment generating function of $\text{deg}_G(n)$ for negative inputs.

Lemma 34 *Let $p_0 := 1 - \left(1 - \frac{k-1}{n+k-1}\right)^m \left(1 - \frac{1}{n}\right)^m \sim \frac{km}{n}$. Then for all $t < 0$ we have*

$$\mathbb{E} \left[\exp (t \text{deg}_G(n)) \right] \leq e^{mt} \left(1 + p_0(e^t - 1)\right)^{n-1-m} \exp \left(\frac{m^2 e^{1-t}}{2n} \right).$$

Proof Let $\ell := \text{Link}(n, \cup_{i \in [m]} T_i)$, and let $B_i := \mathbf{1}\{(i, n) \in \mathcal{G}(n, p)\}$, where $\mathcal{G}(n, p)$ is as above and, in particular, independent of ℓ . Then

$$\begin{aligned} \mathbb{E} \left[\exp (t \text{deg}_G(n)) \mid \ell \right] &= \prod_{i \in [n-1]} \mathbb{E} \left[\exp (t (\mathbf{1}\{i \in \ell\} + \mathbf{1}\{i \notin \ell\} B_i)) \mid \ell \right] \\ &= \prod_{i \in [n-1]} \left((1 - p) \exp (t \mathbf{1}\{i \in \ell\}) + p e^t \right). \end{aligned}$$

Note that $\ell \stackrel{\mathcal{L}}{=} \text{Binom}([n - 1], p') \cup R$ where $p' := 1 - \left(1 - \frac{1}{n}\right)^m$, and R is a random subset of $[n - 1]$ formed by independently and uniformly choosing (with replacement) m vertices from $[n - 1]$ —this follows from the $k = 1$ case of Theorem 1. Thus

$$\mathbb{E} \left[\exp (t \text{deg}_G(n)) \mid R \right] = \prod_{i \in [n-1]} \mathbb{E} \left[(1 - p) \exp (t \mathbf{1}\{i \in \ell\}) + p e^t \mid R \right].$$

Now note that

$$\begin{aligned} \mathbb{E} \left[\exp (t \mathbf{1}\{i \in \ell\}) \mid R \right] &= \mathbb{E} \left[\exp (t (\mathbf{1}\{i \in R\} + \mathbf{1}\{i \notin R\} \text{Bernoulli}(p'))) \mid R \right] \\ &= \exp (t \mathbf{1}\{i \in R\}) \left(1 - p' + p' \exp (t \mathbf{1}\{i \notin R\}) \right) \\ &= (1 - p') \exp (t \mathbf{1}\{i \in R\}) + p' e^t. \end{aligned}$$

Thus

$$\begin{aligned} \mathbb{E} \left[\exp (t \text{deg}_G(n)) \mid R \right] &= \prod_{i \in [n-1]} \left((1 - p) \left((1 - p') \exp (t \mathbf{1}\{i \in R\}) + p' e^t \right) + p e^t \right) \\ &= \prod_{i \in [n-1]} \left((1 - p_0) \exp (t \mathbf{1}\{i \in R\}) + p_0 e^t \right). \end{aligned}$$

We therefore have

$$\begin{aligned} \mathbb{E} \left[\exp (t \text{deg}_G(n)) \right] &= \sum_{j \in [m]} e^{jt} \left(1 - p_0 + p_0 e^t \right)^{n-1-j} \mathbb{P}[|R| = j] \\ &= e^{mt} \left(1 - p_0 + p_0 e^t \right)^{n-1-m} \sum_{j \in [m]} e^{(j-m)t} \left(1 - p_0 + p_0 e^t \right)^{m-j} \mathbb{P}[|R| = j]. \end{aligned}$$

Since $t < 0$, we have $1 - p_0 + p_0e^t \leq 1$, and thus

$$\mathbb{E} [\exp (t \deg_G(n))] \leq e^{mt} (1 - p_0 + p_0e^t)^{n-1-m} \sum_{j \in [m]} e^{-t(m-j)} \mathbb{P}[|R| = j].$$

Toward controlling the size of R , let $R = R_m$, and let R_j be defined as R after only the first j independent samplings of $[n - 1]$. Then, letting v_1, v_2, \dots, v_m be the random vertex selections, we have

$$|R_{j+1}| = |R_j| + \mathbf{1}\{v_{j+1} \notin R_j\}.$$

Thus, for $s > 0$,

$$\begin{aligned} \mathbb{E} [\exp (-s|R_{j+1}|) | R_j] &= \exp (-s|R_j|) \left(\frac{|R_j|}{n-1} + \frac{n-1-|R_j|}{n-1} e^{-s} \right) \\ &\leq \exp (-s|R_j|) \left(\frac{j}{n-1} + \frac{n-1-j}{n-1} e^{-s} \right) \\ &= \exp (-s|R_j| - s) \left(1 + \frac{j(e^s - 1)}{n-1} \right). \end{aligned}$$

Taking expectations and iterating then gives

$$\mathbb{E} \exp (-s|R_m|) \leq e^{-ms} \prod_{i=1}^{m-1} \left(1 + \frac{i(e^s - 1)}{n-1} \right).$$

Thus, by Chernoff’s inequality, we have

$$\mathbb{P}[|R_m| \leq m - j] \leq e^{-js} \prod_{i=1}^{m-1} \left(1 + \frac{i(e^s - 1)}{n-1} \right) \leq \exp \left(-js + \binom{m}{2} \frac{e^s - 1}{n-1} \right)$$

for any $s > 0$. Taking $s = \log \frac{2j(n-1)}{m(m-1)}$, gives

$$\mathbb{P}[|R_m| \leq m - j] \leq \exp \left(j - \frac{m(m-1)}{2(n-1)} \right) \left(\frac{m(m-1)}{2j(n-1)} \right)^j \leq \left(\frac{em^2}{2jn} \right)^j.$$

Thus $|R| \geq m - j + 1$ with probability at least $1 - \left(\frac{em^2}{2jn} \right)^j$. Recall the summation by parts formula:

$$\sum_{j \in [m]} f_j g_j = f_m \sum_{j \in [m]} g_j - \sum_{j \in [m]} (f_j - f_{j-1}) \sum_{i \in [j-1]} g_i.$$

Setting $f_j = e^{-t(m-j)}$ and $g_j = \mathbb{P}[|R| = j]$, we have

$$\begin{aligned} \sum_{j \in [m]} e^{-t(m-j)} \mathbb{P}[|R| = j] &= 1 + (e^{-t} - 1) \sum_{j \in [m]} \mathbb{P}[|R| < j] e^{-t(m-j)} \\ &\leq 1 + (e^{-t} - 1) \sum_{j \in [m]} \left(\frac{em^2}{2(m-j+1)n} \right)^{m-j+1} e^{-t(m-j)} \\ &= 1 + (1 - e^t) \sum_{j \in [m]} \left(\frac{e^{1-t} m^2}{2jn} \right)^j \\ &\leq 1 + (1 - e^t) \left(\exp \left(\frac{m^2 e^{1-t}}{2n} \right) - 1 \right), \end{aligned}$$

where the last inequality uses the bound $j^j \geq j!$. Thus, for $t < 0$ we have

$$\begin{aligned} \mathbb{E} \left[\exp(t \deg_G(n)) \right] &\leq e^{mt} (1 - p_0 + p_0 e^t)^{n-1-m} \sum_{j \in [m]} e^{-t(m-j)} \mathbb{P}[|R| = j] \\ &\leq e^{mt} (1 + p_0(e^t - 1))^{n-1-m} \left(1 + (1 - e^t) \left(\exp \left(\frac{m^2 e^{1-t}}{2n} \right) - 1 \right) \right) \\ &\leq e^{mt} (1 + p_0(e^t - 1))^{n-1-m} \exp \left(\frac{m^2 e^{1-t}}{2n} \right). \end{aligned}$$

□

Lemma 35 For $0 \leq j \leq m$, we have $\mathbb{P}[\min_{i \in [n]} \deg_G(i) \leq m - j] \leq (1 + o(1))n^{1-j} e^{-km}$.

Proof For any $t > 0$, we have by a union bound that

$$\begin{aligned} \mathbb{P}[\min_{i \in [n]} \deg_G(i) \leq m - j] &\leq n \mathbb{P}[\exp(-t \deg_G(n)) \geq e^{(j-m)t}] \\ &\leq ne^{-jt} (1 + p_0(e^{-t} - 1))^{n-1-m} \exp \left(\frac{m^2 e^{1+t}}{2n} \right) \\ &\leq \exp \left(\log n - jt + (n - 1 - m)p_0(e^{-t} - 1) + \frac{m^2 e^{1+t}}{2n} \right). \end{aligned}$$

Letting $t = r \log n$ for any $r \in (0, 1)$ gives

$$\begin{aligned} \mathbb{P}[\min_{i \in [n]} \deg_G(i) \leq m - j] &\leq \exp \left((1 - jr) \log n - (n - 1 - m)p_0(1 - n^{-r}) + \frac{em^2}{2n^{1-r}} \right) \\ &\leq (1 + o(1)) \exp((1 - jr) \log n - km). \end{aligned}$$

Taking $r \rightarrow 1$ from the left then gives $(1 + o(1))n^{1-j} e^{-km}$.

□

Lemma 36 *Let L be the reduced Laplacian of G as defined above with $m = \lceil \delta \log n \rceil$ and $\delta > 0$ an arbitrary constant (so that $\mathbb{P}[\min_{i \in [n]} \deg_G(i) \geq m - s + 2] = 1 - o(n^{-s})$ for any fixed s). Then for any fixed s , we have $\lambda_1(L) = 0$ and $\lambda_2(L) = 1 - O(1/\sqrt{\log n})$ with probability $1 - o(n^{-s})$.*

Proof This follows from the previous lemma and Corollary 27. \square

On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

- Bernardi, O., Klivans, C.: Directed rooted forests in higher dimension. *Electron. J. Combin.* (2015). <https://doi.org/10.37236/5819>
- Coja-Oghlan, A., Lanka, A.: The spectral gap of random graphs with given expected degrees. *Electron. J. Combin.* **16**, 138–138 (2009)
- Duval, A., Klivans, C., Martin, J.: Simplicial matrix-tree theorems. *Trans. Am. Math. Soc.* (2008). <https://doi.org/10.1090/S0002-9947-09-04898-3>
- Duval, A.M., Klivans, C.J., Martin, J.L.: Cellular spanning trees and Laplacians of cubical complexes. *Adv. Appl. Math.* **46**(1), 247–274 (2011). <https://doi.org/10.1016/j.aam.2010.05.005>. (Special issue in honor of Dennis Stanton)
- Duval, A.M., Klivans, C.J., Martin, J.L.: Cuts and flows of cell complexes. *J. Algebraic Comb.* **41**(4), 969–999 (2015). <https://doi.org/10.1007/s10801-014-0561-2>
- Feige, U., Ofek, E.: Spectral techniques applied to sparse random graphs. *Random Struct. Algorithms* **27**(2), 251–275 (2005). <https://doi.org/10.1002/rsa.20089>
- Garland, H.: p-Adic curvature and the cohomology of discrete subgroups of p-adic groups. *Ann. Math.* **97**, 375 (1973)
- Gundert, A., Wagner, U.: On eigenvalues of random complexes. *Israel J. Math.* (2014). <https://doi.org/10.1007/s11856-016-1419-1>
- Hoffman, C., Kahle, M., Paquette, E.: Spectral gaps of random graphs and applications. *Int. Math. Res. Not.* **2021**(11), 8353–8404 (2019). <https://doi.org/10.1093/imrn/rnz077>
- Joag-Dev, K., Proschan, F.: Negative association of random variables with applications. *Ann. Stat.* **11**(1), 286–295 (1983). <https://doi.org/10.1214/aos/1176346079>
- Kahle, M., Newman, A.: Topology and geometry of random 2-dimensional hypertrees. *Discrete Comput. Geom.* **4**, 1229–1244 (2022)
- Kalai, G.: Enumeration of q-acyclic simplicial complexes. *Israel J. Math.* **45**, 337–351 (1983)
- Linial, N., Meshulam, R.: Homological connectivity of random 2-complexes. *Combinatorica* **26**, 475–487 (2006)
- Linial, N., Peled, Y.: Enumeration and randomized constructions of hypertrees. *Random Struct. Algorithms* **55**(3), 677–695 (2019). <https://doi.org/10.1002/rsa.20841>
- Lubotzky, A.: High dimensional expanders. In: *Proceedings of the International Congress of Mathematicians: Rio de Janeiro 2018*, pp. 705–730. World Scientific (2018)
- Lyons, R.: Determinantal probability measures. *Publ. Math. l’IHÉS* **98**, 167–212 (2003)
- Lyons, R.: Random complexes and ℓ^2 -betti numbers. *J. Topol. Anal.* **01**(02), 153–175 (2009). <https://doi.org/10.1142/S1793525309000072>
- Meshulam, R., Wallach, N.: Homological connectivity of random k-dimensional complexes. *Random Struct. Algorithms* **34**, 408–417 (2009)
- Mészáros, A.: The local weak limit of k-dimensional hypertrees. <https://doi.org/10.1090/tran/8711>
- Oliveira, R.L.: Concentration of the adjacency matrix and of the laplacian in random graphs with independent edges. arXiv preprint [arXiv:0911.0600](https://arxiv.org/abs/0911.0600) (2009)
- Oppenheim, I.: Local spectral expansion approach to high dimensional expanders part I: descent of spectral gaps. *Discrete Comput. Geom.* **59**(2), 293–330 (2018)
- Oppenheim, I.: Local spectral expansion approach to high dimensional expanders part II: mixing and geometrical overlapping. *Discrete Comput. Geom.* **64**(3), 1023–1066 (2020)

- Tikhomirov, K., Youssef, P.: The spectral gap of dense random regular graphs. *Ann. Probab.* **47**(1), 362–419 (2019). <https://doi.org/10.1214/18-AOP1263>
- Žuk, A.: La propriété (t) de kazhdan pour les groupes agissant sur les polyedres. *C. R. Acad. Sci. Sér. 1, Math.* **323**(5), 453–458 (1996)
- Žuk, A.: Property (t) and Kazhdan constants for discrete groups. *Geom. Funct. Anal.* **13**, 643–670 (2003). <https://doi.org/10.1007/s00039-003-0425-8>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.