

THE THRESHOLD FOR STACKED TRIANGULATIONS

EYAL LUBETZKY AND YUVAL PELED

ABSTRACT. A *stacked triangulation* of a d -simplex $\mathfrak{o} = \{1, \dots, d+1\}$ ($d \geq 2$) is a triangulation obtained by repeatedly subdividing a d -simplex into $d+1$ new ones via a new vertex (the case $d = 2$ is known as an Apollonian network). We study the occurrence of such a triangulation in the Linial–Meshulam model, i.e., for which p does the random simplicial complex $Y \sim \mathcal{Y}_d(n, p)$ contain the faces of a stacked triangulation of the d -simplex \mathfrak{o} , with its internal vertices labeled in $[n]$. In the language of bootstrap percolation in hypergraphs, it pertains to the threshold for K_{d+2}^{d+1} , the $(d+1)$ -uniform clique on $d+2$ vertices.

Our main result identifies this threshold for every $d \geq 2$, showing it is asymptotically $(\alpha_d n)^{-1/d}$, where α_d is the growth rate of the Fuss–Catalan numbers of order d . The proof hinges on a second moment argument in the supercritical regime, and on Kalai’s algebraic shifting in the subcritical regime.

1. INTRODUCTION

We study the following question, which can be phrased, and different flavors of it were studied, in the context of (i) stacked triangulations of a simplex in a random simplicial complex, or (ii) the homotopy-equivalent closure of a simplicial complex under elementary expansions, or (iii) bootstrap percolation on random hypergraphs. For simplicity, we first phrase it for dimension $d = 2$, noting the objects of interest, *stacked triangulations*, are in that case also known as *Apollonian networks*.

Question 1.1. A *stacked triangulation* of a triangle (a.k.a. *Apollonian network*) is obtained by repeatedly subdividing a triangle into three via a new vertex. For a (random) set of triangles $Y \subseteq \binom{[n]}{3}$, is there such a triangulation of $\{1, 2, 3\}$ and a labeling of its vertices in $[n]$ so that all these labeled faces belong to Y ?

If Y contains each triangle in $\binom{[n]}{3}$ independently with probability p , this addresses such triangulations in the Linial–Meshulam [26] random complex $\mathcal{Y}_2(n, p)$. See §1.1.1 for related work on triangulations and simple connectivity of $\mathcal{Y}_2(n, p)$.

The equivalent formulation obtained from reversing this process is as follows.

Question 1.1’. Given a (random) set of triangles $Y \subseteq \binom{[n]}{3}$, repeatedly add to Y every triangle $f = \{v_1, v_2, v_3\}$ for which there exists some $u \in [n]$ such that all three triangles $\{v_1, v_2, u\}$, $\{v_1, u, v_3\}$, $\{u, v_2, v_3\}$ belong to Y . When this process terminates, does the final set Y_∞ contain $\{1, 2, 3\}$? When does one have $Y_\infty = \binom{[n]}{3}$?

In the language of weak saturation and bootstrap percolation of hypergraphs introduced by Bollobás [13] in the late 1960’s, this is K_4^3 -bootstrap percolation: one starts with a random 3-uniform hypergraph, and repeatedly adds a hyperedge if it completes a new copy of K_4^3 , the complete hypergraph on 4 vertices (see §1.1.2).

Another way of viewing Question 1.1’ is in terms of Whitehead’s *elementary moves* between simplicial complexes [36]. A *3-expansion* in a simplicial complex is the addition of a 3-simplex and the single 2-face missing from it—a fundamental operation that preserves its homotopy type. In this case, one adds the tetrahedron $\{v_1, v_2, v_3, u\}$ along with the missing face f . The above question thus asks what complex we arrive at by after performing all possible 3-expansions on $\mathcal{Y}_2(n, p)$.

We now state the setting of the problem formally for any fixed dimension $d \geq 2$. A bistellar 0-move on a d -dimensional simplex is a subdivision of the simplex into $d + 1$ simplicies using a new vertex in its interior. Combinatorially, the d -simplex¹ $f = \{v_1, \dots, v_{d+1}\}$ is replaced by $f \setminus \{v_j\} \cup \{u\}$, $j = 1, \dots, d + 1$, where u is the new vertex. A *stacked triangulation* is a triangulation of the d -dimensional simplex that is obtained from the simplex by a sequence of bistellar 0-moves. Let Y be a simplicial complex containing the boundary of the d -simplex $\{v_1, \dots, v_{d+1}\}$. We say that Y admits a *stacked contraction* of f if there is a simplicial map from a stacked triangulation T to Y that maps the boundary to f , or in other words, if there is a labelling of the vertices of T in which the boundary vertices are labeled v_1, \dots, v_{d+1} and the label-set of every face of T belongs to Y .

In this paper we study a stacked triangulations and contractions in the binomial Linial–Meshulam random simplicial complex. More precisely, we let $Y \sim \mathcal{Y}_d(n, p)$ be an n -vertex d -dimensional complex with a full skeleton in which every d -simplex appear independently with probability $p = p(n)$, and ask whether there exists a stacked contraction of the d -simplex $\{1, \dots, d + 1\}$ in Y with high probability (a probability tending to 1 as $n \rightarrow \infty$). Moreover, how many d -simplices in $\binom{[n]}{d+1}$ have a stacked contraction in Y ?

For the alternate point-of-view on this problem, let $Y \subseteq \binom{[n]}{d+1}$ be an initial set of d -dimensional faces. As long as there exist $f = \{v_1, \dots, v_{d+1}\} \notin Y$ and u such that

$$f \setminus \{v_i\} \cup \{u\} \in Y \quad \text{for all } i,$$

the face f is added to Y . This deterministic process terminates if no such f exists.

In the language of Whitehead’s elementary moves, this is phrased as follows.

An elementary $(d + 1)$ -collapse is the removal of a $(d + 1)$ -dimensional face \hat{f} and one of its d -dimensional facets f from a complex in which f is not contained in any other $(d + 1)$ -dimensional face. The inverse of this move, in which a $(d + 1)$ -simplex \hat{f} and its d -facet f are added to a complex that contains all the boundary of \hat{f} except f , is called an elementary $(d + 1)$ -expansion. Both these moves preserve the homotopy type of the complex, and simple-homotopy theory studies the conditions under which homotopy equivalent complexes can be obtained one from the other by a sequence of elementary collapses and expansions [15].

Elementary collapses play a big role in the study of random simplicial complexes (see §1.1.1). In our context, we start with the random simplicial complex $\mathcal{Y}_d(n, p)$, and carry out an elementary $(d + 1)$ -expansion as long as such a move is possible. The above questions address, e.g., whether $\{1, \dots, d + 1\}$ is in the final complex.

The H -bootstrap percolation process is a deterministic process which starts with an initially infected set $Y \subseteq \binom{[n]}{d+1}$, and in every step a new hyperedge that is the only non-infected hyperedge in a copy of some fixed hypergraph H gets infected. In our setting, the initially infected hypergraph is a binomial random hypergraph in which every hyperedge appears independently with probability p and the fixed hypergraph H is K_{d+2}^{d+1} — the $(d + 1)$ -uniform clique on $d + 2$ vertices. We wish to estimate the probability that a given hyperedge is infected, as well as the probability for *percolation*, i.e., the infection of all hyperedges in $\binom{[n]}{d+1}$. Similar problems in graph bootstrap percolation were investigated thoroughly (see §1.1.2 for details). Here we settle these problems for the cliques K_{d+2}^{d+1} , perhaps one of the most basic problems in random hypergraph bootstrap percolation.

¹As a member of a simplicial complex, a set of size $d + 1$ is called a d -dimensional face or simplex.

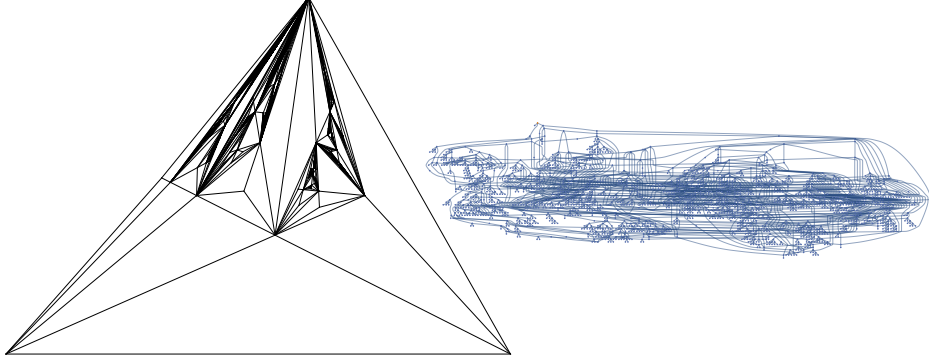


FIGURE 1. A stacked triangulation of $\{1, 2, 3\}$ in $\mathcal{Y}_2(n, p)$ at criticality for $n = 200$, and the corresponding directed acyclic graph (see Def. 3.4); it has 2097 vertices, 1275 of which are leaves, using 184 labels from $[n]$.

Our main result determines, for every $d \geq 2$, the asymptotic threshold p_c for the property that $\mathcal{Y}_d(n, p)$ admits stacked contractions of all $f \in \binom{[n]}{d+1}$, or equivalently, the asymptotic threshold for K_{d+2}^{d+1} -percolation (see Fig. 2 for a triangulation at p_c). In what follows, let

$$\alpha_d = (d+1)^{d+1}/d^d \quad (1.1)$$

denote the growth rate of the Fuss–Catalan numbers of order d (see §2.2).

Theorem 1. *Let $Y \sim \mathcal{Y}_d(n, p)$ for fixed $d \geq 2$, and set α_d as in (1.1).*

- (i) [Subcritical] *If $p = (1 - \frac{1}{\log n})(\alpha_d n)^{-1/d}$ then the probability that Y admits a stacked contraction of the face $\mathbf{o} = \{1, \dots, d+1\}$ is $n^{-1/d+o(1)}$.*
- (ii) [Supercritical] *There exists $\delta_d > 0$ such that, if $p = (1 + n^{-\delta_d})(\alpha_d n)^{-1/d}$ then, with high probability, Y admits a stacked contraction of every face $f \in \binom{[n]}{d+1}$.*

For a heuristic argument suggesting that $(\alpha_d n)^{-1/d}$ is the correct threshold, let us completely ignore the delicate dependencies between the stacked contractions of various faces; suppose instead that each face has such a contraction independently with probability $\theta = \theta(n)$, setting aside the trivial equilibrium $\theta = 1$. Then a given face \mathbf{o} has such a contraction either by being in Y , which has probability p , or via a new vertex, which has probability $1 - (1 - \theta^{d+1})^{n-(d+1)} \approx n\theta^{d+1}$ (we should have $\theta \ll n^{-1/(d+1)}$ so the final expression could give back θ). This yields the equation

$$\theta = p + n\theta^{d+1},$$

and we see that if $p = \gamma n^{-1/d}$ and $\theta = \hat{\gamma} n^{-1/d}$ then this becomes $\hat{\gamma}^{d+1} - \hat{\gamma} + \gamma = 0$. The polynomial $Q_\gamma(x) = x^{d+1} - x + \gamma$ is convex in \mathbb{R}_+ with $Q_\gamma(x_0) = -\alpha_d^{1/d}$ at its local minimum $x_0 = (d+1)^{-1/d}$. If $\gamma > \alpha_d^{1/d}$ then $Q_\gamma(x)$ has no roots in \mathbb{R}_+ , leaving us only with the trivial solution $\theta = 1$. If $\gamma < \alpha_d^{1/d}$ then $Q_\gamma(x)$ has two roots in \mathbb{R}_+ , and the solution for θ should correspond to the smaller root $0 < \hat{\gamma} < x_0$.

The next more detailed result on the subcritical regime confirms this heuristic.

Theorem 2. *Consider $Y \sim \mathcal{Y}_d(n, \gamma n^{-1/d})$ for fixed $d \geq 2$ and $0 \leq \gamma < \alpha_d^{-1/d}$. Let \hat{X} be the fraction of sets $\mathbf{o} \in \binom{[n]}{d+1}$ so that Y admits a stacked contraction of \mathbf{o} . Then $n^{1/d} \hat{X} \rightarrow \hat{\gamma}$ in probability as $n \rightarrow \infty$, where $\hat{\gamma}$ is the smallest positive root of the polynomial $Q_\gamma(x) = x^{d+1} - x + \gamma$.*

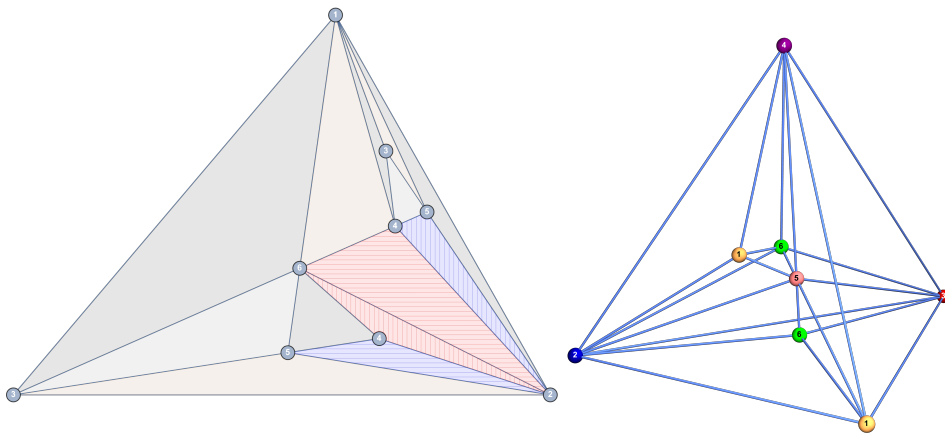


FIGURE 2. Stacked contraction in dimension $d = 2$ (left; the labels 3, 4, 5 and the shaded faces $\{2, 4, 6\}, \{2, 4, 5\}$ are repeated) and $d = 3$ (right; the labels 1, 6 and faces $\{1, 2, 4, 5\}, \{1, 2, 5, 6\}, \{2, 3, 5, 6\}$ are repeated).

Theorems 1 and 2 show that if $p = \gamma n^{-1/d}$ for $\gamma < \alpha_d^{1/d}$ then the K_{d+2}^{d+1} -bootstrap percolation process starting from a binomial random hypergraph with edge density p does not have a real impact — the density of the infected hyperedges remains of order $n^{-1/d}$ as it was initially. On the other hand, if $p = \gamma n^{-1/d}$ for $\gamma > \alpha_d^{1/d}$ then the process percolates and with high probability all the hyperedges get infected. In fact, numerical simulations suggest that if we increase the initially infected set one random hyperedge at a time, this cascade occurs at one critical step.

The supercritical regime in Theorem 1, where $\mathcal{Y}_d(n, p)$ admits a stacked contraction of every $\mathfrak{o} \in \binom{[n]}{d+1}$, is handled by establishing the stronger property that \mathfrak{o} has a stacked triangulation in which all labels are distinct. The standard way of approaching this would be via the second moment method; however, already for a single face \mathfrak{o} , the vanilla method fails: the number of stacked triangulations of \mathfrak{o} is not concentrated at $p = (1 + o(1))p_c$, nor in fact at $p = Cp_c$ for any fixed $C > 1$, due to correlations between triangulations. To bypass this, and further show that with high probability there is a stacked contraction of every face, we use a modified second moment, restricted to a set of *balanced* stacked triangulations where each of the d -simplices subdividing \mathfrak{o} have the same number of internal vertices.

Conversely, in the subcritical regime we are required to exclude the presence of stacked contractions with *repeated labels*, whose combinatorial complexity presents a substantial challenge (see Fig. 2). This is an inherent difficulty in bootstrap percolation and related problems; e.g., the 2-dimensional non-stacked variant of this problem — determining whether a given simplicial complex is simply-connected — is undecidable. In a random setting, determining the sharp threshold for simple connectivity of $\mathcal{Y}_2(n, p)$, is an important open problem in combinatorial topology.

In the stacked case, an adaptation of the existing graph bootstrap percolation machinery due to Balogh, Bollobás and Morris [8] shows $p_c > n^{-1/d}(\log n)^{-c_d}$, where $c_d > 0$ increases with the dimension d . While we could decrease c_d in the superfluous poly-log factor by a careful analysis of triangulations with repeated labels, neither this refinement, nor any of the other combinatorial first moment arguments that we attempted, could recover the correct order of the critical probability.

Miraculously, Kalai’s algebraic shifting procedure [23] enabled us to establish new non-trivial bounds on the number and complexity of stacked contractions (roughly put, one aims to bound the tree excess in the directed acyclic graph corresponding to a stacked contraction via the number of labels it features — e.g., in Fig. 1 the former is 370 and the latter is 184; see Theorem 3.5), leading to the sharp threshold.

1.1. Related works. Next we discuss a number of lines of research that pertain to the problem studied here, and mention related literature on these.

1.1.1. Random simplicial complexes. The systematic study of the topology of random complexes was initiated by Linial and Meshulam [26] who introduced $\mathcal{Y}_2(n, p)$ and studied its homological connectivity. The d -dimensional generalization of this model $\mathcal{Y}_d(n, p)$ was introduced in [32].

Babson, Hoffman and Kahle [6] studied the fundamental group of $\mathcal{Y}_2(n, p)$ and showed it is non-trivial and hyperbolic if $p = n^{-1/2+o(1)}$. In our terminology, they proved that with high probability $\mathcal{Y}_2(n, p)$ does not admit a (not necessarily stacked) contraction of a fixed $\{v_1, v_2, v_3\}$. In addition, the property of $\mathcal{Y}_2(n, p)$ having a triangulation with distinct labels of every $\{v_1, v_2, v_3\}$ was shown in [30] to have the asymptotic threshold $p = (c_0 n)^{-1/2}$ for an explicit constant c_0 . However, it remains a challenging open problem to exclude the presence of a contraction with repeated labels of $\{v_1, v_2, v_3\}$ in $\mathcal{Y}_2(n, p)$ where $p = \Theta(n^{-1/2})$, thereby improving the lower bound for simple-connectivity of $\mathcal{Y}_2(n, p)$. In our analysis of stacked triangulation, the main challenge is again controlling the complicated effect of repeated labels.

It is worthwhile mentioning that in higher dimensions $d \geq 3$, determining whether all of the d -simplices admit a (not necessarily stacked) contraction in $Y \sim \mathcal{Y}_d(n, p)$ is purely algebraic: the homotopy groups $\pi_i(Y)$ are trivial for all $0 \leq i < d - 1$, and therefore, by the Hurewicz Theorem, $\pi_{d-1}(Y)$ coincides with the well-studied homology group of Y [35]. In particular, such contractions do exist with high probability if $p \gg d \frac{\log n}{n}$, starkly below the critical threshold found in Theorem 1.

One of the most useful techniques in studying topological invariants of $\mathcal{Y}_d(n, p)$ is to simplify it via elementary collapses. The threshold probability for d -collapsibility of $\mathcal{Y}_d(n, p)$ — the property that all d -faces of the complex can be removed by a sequence of elementary collapses — was studied in [5]. Subsequently, it was discovered in [4, 27] that, throughout the entire regime $p = \Theta(1/n)$, d -collapses play a decisive role in computing the Betti numbers of $\mathcal{Y}_d(n, p)$ (see the survey [28]). Elementary collapses were also studied on other models of random complexes such as the multi-parameter model [16] and the random clique complex [31]. The study of elementary expansions in random complexes is not as developed, but Kahle, Paquette and Roldán [19] used a cubical analog of elementary expansions to determine the sharp threshold probability for homological and simple connectivity of $\mathcal{Q}_2(n, p)$ — a random 2-dimensional cubical subcomplex of the n -dimensional cube.

1.1.2. Hypergraph bootstrap percolation. Bollobás introduced hypergraph bootstrap percolation in [13] under the notion of weakly saturated graphs; for some of the many related notable results in extremal combinatorics, see, e.g., [2, 17, 21]. The name H -bootstrap percolation was given in [9], taking after a related model known as r -neighbor bootstrap percolation, introduced by the physicists Chapula, Leath and Reich [14]. The latter, in which a vertex of a graph becomes infected if it has r or more infected neighbors, was studied extensively (see [1, 7] for example). Balogh, Bollobás and Morris [8] were the first to study H -bootstrap percolation started at

the Erdős–Rényi random graph $\mathcal{G}(n, p)$; they focused on $H = K_r$, a clique on r vertices, for $r \geq 4$, and found the corresponding critical probability $p_c(K_r)$ up to a poly-logarithmic factor. Thereafter, Bartha and Kolesnik [10] determined $p_c(K_r)$ up to a constant factor for every $r \geq 5$. Angel and Kolesnik [3, 24] investigated the special case $r = 4$, and obtained the sharp asymptotics $p_c(K_4) \sim (3n \log n)^{-1/2}$. Other questions regarding graph bootstrap percolation and weak saturation were studied, e.g., in [11, 18, 25]. A random H -bootstrap percolation process in which the infection of a hyperedge missing from a copy of H occurs randomly, rather than deterministically, was studied in [33]. From this result, one can deduce an upper bound of the form $p_c(K_{d+2}^{d+1}) < C_d n^{-1/d}$, off by a constant factor; no lower bound of the correct order was previously known.

1.2. Open problems. Theorems 1 and 2 establish the phase transition for the event that $Y \sim \mathcal{Y}_d(n, p)$ admits a stacked contraction of $\mathbf{o} = \{1, \dots, d+1\}$ —and more generally for the behavior of the density \bar{X} of simplices having a stacked contraction—identifying its asymptotic location to be $p \sim (\alpha_d n)^{-1/d}$. Here are some concrete problems aiming to investigate this phase transition more closely.

Question 1.2. Find $p(n)$ such that, with high probability, $Y \sim \mathcal{Y}_d(n, p)$ admits a stacked contraction but not a stacked triangulation (in which the labels are distinct) of \mathbf{o} . Numerical experiments (cf. Figure 1) suggest such a probability p exists.

Question 1.3. Find $w(n)$ such that at $p = (\alpha_d n)^{-1/d} + w(n)$ the probability that $Y \sim \mathcal{Y}_d(n, p)$ admits a stacked contraction of \mathbf{o} converges to $1/2$. If the stacked contraction does exist at this p , how many labels does it use?

Question 1.4. Let π be a uniform random ordering of the elements of $\binom{[n]}{d+1}$. Denote by $\hat{X}(t)$ the fraction of simplices in $\binom{[n]}{d+1}$ that have a stacked contraction whose faces are among the first t elements of π . Is there, with high probability, a critical step $\tau \approx (\alpha_d n)^{-1/d} \binom{n}{d+1}$ such that $\hat{X}(\tau) = O(n^{-1/d})$ and $\hat{X}(\tau + 1) = 1$?

The rest of the paper is organized as follows. Section 2 gives some necessary background material about simplicial complexes and algebraic shifting. In Sections 3 and 4 we prove the subcritical and supercritical items of Theorem 1 respectively. Finally, in Section 5 we reexamine the subcritical regime and prove Theorem 2.

2. PRELIMINARIES

2.1. Simplicial complexes. A simplicial complex Y is a family of sets that is closed under taking subsets. A set $f \in Y$ is called a face, and its dimension is $|f| - 1$. The dimension of Y is defined as the largest dimension of its faces. The set of d -dimensional faces in a simplicial complex Y is denoted Y_d and the \mathfrak{f} -vector of Y is defined by $\mathfrak{f}_d(Y) = |Y_d|$. We fix an order of the 0-dimensional faces, aka vertices, of Y , and define the d -dimensional boundary operator $\partial_d = \partial_d(Y) : \mathbb{R}^{Y_d} \rightarrow \mathbb{R}^{Y_{d-1}}$ by mapping $f = \{v_0 < \dots < v_d\}$ to

$$\partial_d(f) = \sum_{j=0}^d (-1)^j (f \setminus \{v_j\}).$$

Here faces are identified with their characteristic vectors. A basic fact in algebraic topology is that $\partial_d \partial_{d+1} = 0$, and the d -th real Betti number of Y is defined by

$$\beta_d = \beta_d(Y) = \dim(\ker(\partial_d) / \text{im}(\partial_{d+1})).$$

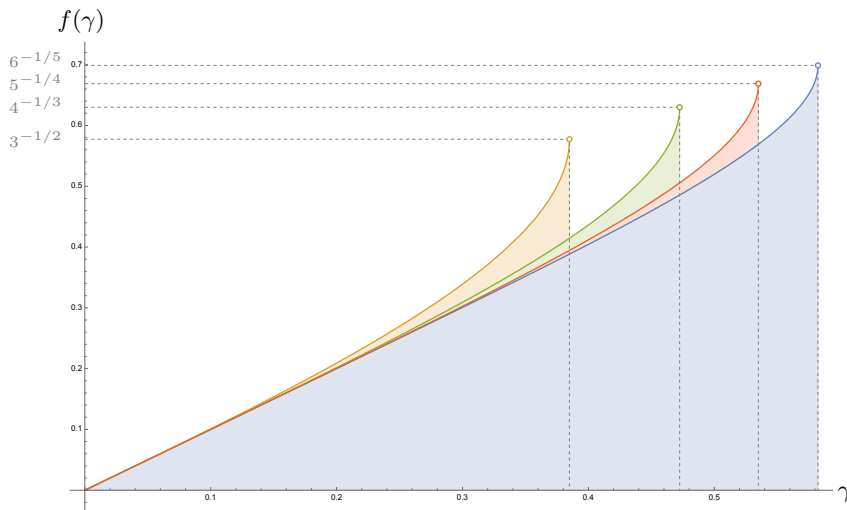


FIGURE 3. The generating function $f(\gamma)$ of the Fuss–Catalan numbers of order $d = 2, 3, 4, 5$; marked points are at $(\alpha_d^{-1/d}, (d+1)^{-1/d})$.

In particular, if Y is a d -dimensional complex then $\beta_d = \dim \ker \partial_d$.

2.2. Stacked triangulations and the Fuss–Catalan numbers. Fix $d \geq 2$ and $f = \{1, \dots, d+1\}$. Let $C_d(s)$ be the number of combinatorially distinct stacked triangulations of the d -simplex f with s *unlabeled* internal vertices. We have that

$$C_d(0) = 1, \quad C_d(s) = \sum_{s_1 + \dots + s_{d+1} = s-1} \prod_{j=1}^{d+1} C_d(s_j), \quad s \geq 1.$$

This recursive relation, that also describes the number of $(d+1)$ -ary trees with s internal vertices and $ds+1$ leaves, defines the Fuss–Catalan numbers:

$$C_d(s) = \frac{1}{ds+1} \binom{(d+1)s}{s} = \left(\frac{d+1}{2\pi d^3} + o_s(1) \right)^{1/2} s^{-3/2} (\alpha_d)^s$$

for $\alpha_d = (d+1)^{d+1}/d^d$ as defined in (1.1).

Using the recursive definition of the Fuss–Catalan number $C_d(s)$, it can be shown that the generating function $f(\gamma) := \sum_{s=0}^{\infty} C_d(s) \gamma^{ds+1}$ satisfies

$$Q_\gamma(f(\gamma)) = f(\gamma)^{d+1} - f(\gamma) + \gamma = 0,$$

for every $0 \leq \gamma < \alpha_d^{-1/d}$. In fact, since $f(\gamma)$ is non-negative, continuous and monotone increasing in γ , it is the smaller among the two positive roots of this equation. See Figure 3 for an illustration of $f(\gamma)$ for $d = 2, 3, 4, 5$.

2.3. Algebraic shifting. Algebraic shifting, that was introduced by Kalai [20,22], is a procedure that, given a simplicial complex Y , produces another simplicial complex $\Delta(Y)$ with special and useful properties. There are two main variants of algebraic shifting: exterior and symmetric. Both variants can suit our purposes in this paper, and we choose to work with exterior shifting. For an extensive overview, we refer the reader to the interesting exposition in [23].

Let $K \subseteq \binom{[n]}{d+1}$ and X be an $n \times n$ real matrix with generic entries. Let M be a $|K| \times \binom{n}{d+1}$ matrix, whose rows are indexed by K and columns are indexed

by $\binom{[n]}{d+1}$ ordered lexicographically. Each entry of M is equal to the corresponding $(d+1) \times (d+1)$ minor of X . To construct $\Delta(K)$ we greedily choose a basis for the column space of M by taking the columns that are not spanned by their predecessors in the lexicographic order, and add the indices of these columns to $\Delta(K)$. It is clear that $|\Delta(K)| = |K|$. In addition, $\Delta(K)$ is shifted, i.e., if $f \in \Delta(K)$, $u \notin f$, $v \in f$ and $u < v$ then $f \setminus \{v\} \cup \{u\} \in \Delta(K)$. We write $f' \leq_p f$ if f' can be obtained from f by a sequence of shifts — in which an element is replaced by a smaller one. In other words, $\{u_0 < \dots < u_d\} \leq_p \{v_0 < \dots < v_d\}$ if and only if $u_i \leq v_i \forall i$. Note that the lexicographic order is a linear extension of \leq_p .

To compute the algebraic shifting of a simplicial complex Y the above procedure is performed independently in each and every dimension, i.e., $\Delta(Y) = \bigcup_d \Delta(Y_d)$. Björner and Kalai [12] showed that $\Delta(Y)$ is a simplicial complex with the same f -vector and the same Betti numbers as Y . Moreover, the Betti numbers of $\Delta(Y)$ are given combinatorially by

$$\beta_d(Y) = \beta_d(\Delta(Y)) = b_d(\Delta(Y)) := |\{f \in \Delta(Y)_d : f \cup \{1\} \notin \Delta(Y)\}|.$$

Note that if Y is a d -dimensional complex then so is $\Delta(Y)$ and

$$b_d(\Delta(Y)) = |\{f \in \Delta(Y)_d : 1 \notin f\}|.$$

3. THE SUBCRITICAL REGIME

Our goal in this section is to establish the following result on the stacked contractions of $\{1, 2, \dots, d+1\}$ in $\mathcal{Y}_d(n, p)$, including those that allow repeated labels.

Theorem 3.1. *Fix $d \geq 2$ and let $p = (1 - \delta)(\alpha_d n)^{-1/d}$ for some $\delta = \delta(n) \in [0, \frac{1}{2}]$. Let \mathcal{E}_s° be the event that $\mathcal{Y}_d(n, p)$ admits a stacked contraction of $\mathbf{o} = \{1, 2, \dots, d+1\}$ that features $s + d + 1$ distinct labels. Then for every $0 \leq s \leq (\log n)^{5/2}$,*

$$\mathbb{P}(\mathcal{E}_s^\circ) \leq e^{-\delta ds} (p + n^{-2/d+o(1)}).$$

The lower bound in our main theorem will follow immediately from the above theorem in light of the following simple fact.

Observation 3.2. *Suppose $K \subseteq \binom{[n]}{d+1}$ admits a stacked contraction of a face f using $s + d + 1$ distinct labels. Then for every $0 \leq k < s$, there is some f' which has a stacked contraction in K using $s' + d + 1$ distinct labels, where $k/(d+1) \leq s' \leq k$.*

Proof. Take $s \geq 1$ (or there is nothing to show), and consider a stacked contraction for f , where $z \in [n] \setminus f$ subdivides f into $f^{(1)}, \dots, f^{(d+1)}$, and each $f^{(i)}$ has a stacked contraction in K via $s^{(i)} + d + 1$ distinct labels. Then $s \leq \sum s^{(i)} + 1$ (the last term accounting for the label z), hence there exists an $f^{(i)}$ with $s^{(i)} \geq (s-1)/(d+1)$ and at least one fewer internal vertex in its underlying triangulation than f . Set $f_0 = f$ and $s_0 = s$, and let f_i be the face obtained in this manner from f_{i-1} ; namely, it has m_i internal vertices and $s_i + d + 1$ distinct labels in its stacked triangulation in K , and

$$\frac{s_{i-1} - 1}{d+1} \leq s_i \leq s_{i-1} \quad , \quad m_i < m_{i-1}.$$

Let $t \geq 1$ be the minimal index such that $s_t \leq k$ (since the m_i 's are strictly decreasing, and $m_i = 0$ would imply $s_i = 0$, we have $t \leq m_0$). Then we also have that $s_t \geq (s_{t-1} - 1)/(d+1) \geq k/(d+1)$ by definition, thus $f' = f_t$ is as required. \blacksquare

Corollary 3.3. Fix $d \geq 2$ and let $p = (1 - \frac{1}{\log n})(\alpha_d n)^{-1/d}$. Then the probability that $\mathcal{Y}_d(n, p)$ admits a stacked contraction of $\{1, 2, \dots, d+1\}$ is at most $n^{-1/d+o(1)}$.

Proof. A union bound on the estimate from Theorem 3.1 shows that

$$\mathbb{P}\left(\bigcup\{\mathcal{E}_s^\circ : s \leq (\log n)^{5/2}\}\right) \leq n^{-1/d+o(1)}.$$

Furthermore, if \mathcal{E}_s^f is the analog of \mathcal{E}_s° corresponding to a stacked contraction of a face $f = \{x_1, \dots, x_{d+1}\}$ (rather than $\mathbf{o} = \{1, \dots, d+1\}$) then

$$\mathbb{P}\left(\bigcup\{\mathcal{E}_s^f : f \in \binom{[n]}{d+1}, \frac{1}{d+1}(\log n)^{5/2} \leq s \leq (\log n)^{5/2}\}\right) \leq n^{-(\frac{d}{d+1}-o(1))\sqrt{\log n}}, \quad (3.1)$$

as the probability estimate from Theorem 3.1 is at most $n^{-\frac{d}{d+1}\sqrt{\log n}}$ for any f, s as above. Combining these estimates with Observation 3.2 concludes the proof. ■

Theorem 3.1 will be derived by a first moment argument on the number of stacked contractions of $\{1, 2, \dots, d+1\}$. It is natural to view a stacked contraction of a d -simplex f whose vertices have distinct labels as a $(d+1)$ -ary tree of d -simplices: the root of the tree is f , and the $(d+1)$ children of every node in the tree are the d -simplices that subdivided it during the recursive construction of the stacked triangulation. In consequence, the leaves of this tree are the d -faces of the stacked triangulation. More generally, an arbitrary stacked contraction can be viewed as a directed acyclic graph (DAGs) whose vertices are a subset of $\binom{[n]}{d+1}$, as follows.

Definition 3.4 (an (m, l, s) -pedigree). Let G be a DAG on a vertex set $V \subseteq \binom{[n]}{d+1}$. We say that G is an (m, l, s) -pedigree for a vertex $\mathbf{o} \in V$ if the following hold:

- (1) [root] \mathbf{o} is the unique vertex in G with in-degree 0.
- (2) [out-degrees] G has $m+l$ vertices, of which m have out-degree $d+1$ (internal) and the remaining l have out-degree 0 (leaves).
- (3) [labels] the cardinality of $\bigcup_{\mathbf{u} \in V} \mathbf{u}$ (all labels from $[n]$ appearing in G) is $s+d+1$.
- (4) [stacking] If a vertex $\mathbf{u} = \{x_1 < x_2 < \dots < x_{d+1}\}$ has outgoing edges to $\mathbf{v}_1, \dots, \mathbf{v}_{d+1}$, then there is some $z \notin \mathbf{u}$ so that $\mathbf{v}_i = \mathbf{u} \setminus \{x_i\} \cup \{z\}$ for all i . The label z is called is the outgoing label of \mathbf{u} .

One can easily verify (by the recursive definition of a stacked triangulation) that if a simplicial complex Y admits a stacked contraction of a d -simplex f , then there exists a DAG as above whose leaves are all members of Y , and whose root \mathbf{o} is equal to f .

A crucial ingredient in the proof of Theorem 3.1 will be the following result, which bounds the tree excess of an (m, l, s) -pedigree in terms of the surplus in s , its total number of labels.

Theorem 3.5. Let $d \geq 2$. For every (m, l, s) -pedigree G , if \mathbf{a} is given by

$$\mathbf{a} := l - (ds + 1) \quad (3.2)$$

(the excess in leaves) and \mathbf{b} is the tree excess

$$\mathbf{b} := |E(G)| - (|V(G)| - 1) = dm - (l - 1), \quad (3.3)$$

then $\mathbf{a} \geq 0$ and

$$\mathbf{b} \leq d\mathbf{a}^{(d+1)/d} + (d^2 - 1)\mathbf{a}. \quad (3.4)$$

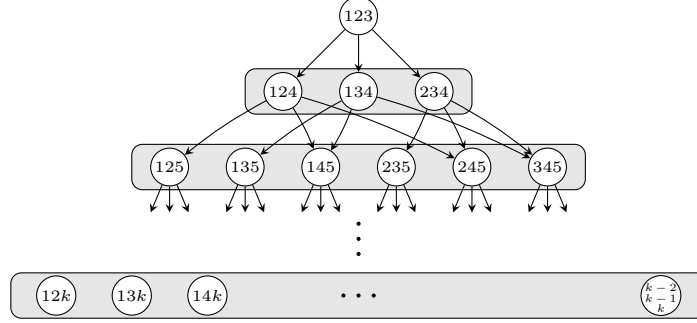


FIGURE 4. A family of (m, l, s) -pedigrees $\{G_k\}_{k \geq 3}$ for $d = 2$ where \mathbf{b} is of order $\mathbf{a}^{3/2}$ (namely, $\mathbf{b} \sim m = \sum_{j=2}^{k-2} \binom{j}{2}$, $\mathbf{a} \sim l = \binom{k-1}{2}$ and $s = k - 3$).

Remark 3.6. *The bound $\mathbf{b} \leq c_d \mathbf{a}^{(d+1)/d}$ in (3.4) is tight up to the value of the constant prefactor c_d , as one may consider, for any fixed $d \geq 2$ and any $k > d + 1$, the (m, l, s) -pedigree G_k where the j -th level of the breadth-first-search tree from \mathbf{o} consists of $\{j\} \cup f$ for all $f \in \binom{[j-1]}{d}$. Here $m = \sum_{j \leq k-2} \binom{j}{d}$, $l = \binom{k-1}{d+1}$, $s = k - (d + 1)$, and \mathbf{b} is of order $\mathbf{a}^{(d+1)/d}$. (Figure 4 illustrates this for $d = 2$.)*

3.1. Proof of Theorem 3.5 via algebraic shifting. Whereas the fact that $\mathbf{b} \geq 0$, that is, $dm \geq l - 1$, is trivial from the representation of \mathbf{b} as the tree excess of (the undirected underlying graph of) G , the fact that $\mathbf{a} \geq 0$ is less obvious. We note that it can be obtained by a direct adaptation of the methods in [8]. An additional proof of this fact appears in Proposition A.1, which has the benefit of also expressing \mathbf{a} in terms of the rank of an explicit matrix. Unfortunately, both these arguments do not provide any quantitative information. However, the matrix in Proposition A.1 is a rigidity matrix, which suggests the relevance of algebraic shifting to this problem. Indeed, we will eventually obtain the fact that $\mathbf{a} \geq 0$ as a byproduct of the argument establishing (3.4) via algebraic shifting.

Claim 3.7. *Let $K \subseteq \binom{[s+d+1]}{d+1}$ be the vertex set of an (m, l, s) -pedigree with $s \geq 1$, and let \bar{K} denote the d -dimensional complex consisting of K and all of its subsets. Then the shifted complex $\Delta(\bar{K})$ contains $\mathbf{v}_0 := [d + 1] \cup \{s + d + 1\} \setminus \{2\}$.*

Proof. We will establish the claim for every (m, l, s) -pedigree G with $s \geq 1$ by induction on the number of vertices $m + l$, with the base case of $m = 1$ (the root) and $l = d + 1$ (the case where $s = 1$) containing \mathbf{v}_0 as one of the leaves. Since \bar{K} is in this case the boundary of a $(d + 1)$ -simplex, we have $\Delta(\bar{K}) = \bar{K} = \binom{[d+2]}{d+1} \ni \mathbf{v}_0$.

For a general (m, l, s) -pedigree G with $m + l > d + 2$ (and $s \geq 2$), consider some vertex $\mathbf{u} \in V(G)$ such that all of its $d + 1$ children in G are leaves (such a vertex exists in every finite DAG). Let z be the outgoing label of \mathbf{u} , let G' be the DAG obtained by deleting \mathbf{u} and its $d + 1$ children from G , and—denoting its vertex set by $K' \subseteq \binom{[s+d+1]}{d+1}$ —let \bar{K}' be the complex consisting of K' and all of its subsets.

If the label z still appears in G' (i.e., G' is an $(m - 1, l - (d + 1), s)$ -pedigree), then $\mathbf{v}_0 \in \Delta(\bar{K}')$ by the induction hypothesis. In this case, the induction step is a direct consequence of the following monotonicity property of the shifted complex:

Fact 3.8. *Let $\bar{K}' \subseteq \bar{K}$ be such that $\mathfrak{f}_0(\bar{K}') = \mathfrak{f}_0(\bar{K})$. Then $\Delta(\bar{K}') \subseteq \Delta(\bar{K})$.*

To verify this property, recall the algorithm for generating $\Delta(\bar{K})$ by testing the appropriate columns of a $\mathfrak{f}_d(\bar{K}) \times \binom{\mathfrak{f}_0(\bar{K})}{d+1}$ matrix M (recall $\mathfrak{f}_d(\bar{K}) = |K|$) for linear

dependence on their predecessors. Having $\mathfrak{f}_0(\bar{K}') = \mathfrak{f}_0(\bar{K})$ implies that the matrix M' used to generate $\Delta(\bar{K}')$ in this manner is the submatrix of M corresponding to a $|K'|$ -subset of its rows; thus, if $f \notin \Delta(\bar{K})$, then its corresponding column in M is linearly dependent on its predecessors, so this also holds for M' , and $f \notin \Delta(\bar{K}')$.

It remains to treat the case where the label z does not appear in G' . By permuting the labels in G we may assume without loss of generality that $z = s + d + 1$. Since G' has $s - 1 \geq 1$ distinct labels, our induction then asserts that $\mathbf{v}'_0 \in \Delta(\bar{K}')$ for $\mathbf{v}'_0 = [d + 1] \cup \{s + d\} \setminus \{2\}$.

Write $\mathbf{u} = \{x_1 < \dots < x_{d+1}\}$, and define

$$L = \{\mathbf{u}\} \cup \{\mathbf{u} \cup \{z\} \setminus \{x_i\}\},$$

so that $K = K' \cup L$. As usual, write \bar{L} for the complex obtained by the closure of L w.r.t. subsets. We will appeal to a useful characterization due to Nevo [34] for the faces of the algebraic shift of a union of two complexes:

Theorem 3.9 ([34, Thm 4.6]). *Let Γ_1, Γ_2 be two d -dimensional simplicial complexes with $\mathfrak{f}_d(\Gamma_1 \cap \Gamma_2) = 1$ (i.e., their intersection contains a single d -dimensional face). Then for every $f = \{y_1 < \dots < y_{d+1}\} \subseteq (\Gamma_1 \cup \Gamma_2)_0$,*

$$f \in \Delta(\Gamma_1 \cup \Gamma_2) \iff y_{d+1} - y_d \leq \mathcal{D}_{\Gamma_1}(f) + \mathcal{D}_{\Gamma_2}(f) - \mathcal{D}_{\Gamma_1 \cap \Gamma_2}(f),$$

in which $\mathcal{D}_{\Gamma}(f) = \#\{w : \{y_1 < \dots < y_d < w\} \in \Delta(\Gamma)\}$.

In our application, let Γ_1 and Γ_2 correspond to \bar{K}' and \bar{L} , respectively, noting that $K' \cap L = \{\mathbf{u}\}$ by the assumption that z does not appear in G' . Further take $f = \mathbf{v}_0$, whence $y_{d+1} - y_d = s$, and $\mathcal{D}_{\Gamma}(\mathbf{v}_0)$ counts the number of faces of Γ of the form $[d + 1] \cup \{w\} \setminus \{2\}$ for some $w > d + 1$. Then

- (i) $\mathcal{D}_{\bar{K}'}(\mathbf{v}_0) = s - 1$ (corresponding to $w = d + 1 + i$ for $i \in [s - 1]$, as $\mathbf{v}'_0 \in \Delta(\bar{K}')$);
- (ii) $\mathcal{D}_{\bar{L}}(\mathbf{v}_0) = 1$ (since $\Delta(\bar{L})$ is the complete simplicial complex $\binom{[d+2]}{[d+1]}$);
- (iii) $\mathcal{D}_{\bar{K}' \cap \bar{L}}(\mathbf{v}_0) = 0$ (since the shifting of a single d -face \mathbf{u} is $[d + 1]$).

We see that $\mathcal{D}_{\bar{K}'}(\mathbf{v}_0) + \mathcal{D}_{\bar{L}}(\mathbf{v}_0) - \mathcal{D}_{\bar{K}' \cap \bar{L}}(\mathbf{v}_0) = s = y_{d+1} - y_d$, and thus $\mathbf{v}_0 \in \Delta(\bar{K})$ by the above theorem, as required. \blacksquare

Proof of Theorem 3.5. Assume $s \geq 1$ (otherwise $l = 1$ and $\mathbf{a} = \mathbf{b} = 0$), and that without loss of generality the outgoing labels in G are taken from $[s + d + 1]$ (recalling that \mathbf{a} and \mathbf{b} are invariant under replacing every occurrence of an outgoing label z by a previously unused z'). Let \bar{K} denote the complex containing the vertices of G and all of their subsets, and consider the Betti number $\beta_d := \beta_d(\bar{K}) = \beta_d(\Delta(\bar{K}))$.

On one hand, we claim that $\beta_d \geq m$. This follows from the basic fact that if \mathbf{u} is an internal vertex in G with children $\mathbf{v}_1, \dots, \mathbf{v}_{d+1}$, then $\partial_d \mathbf{u}$ is a linear combination of $\{\partial_d \mathbf{v}_i : i \in [d + 1]\}$. To see this, consider an internal vertex $\mathbf{u} = \{x_1 < \dots < x_{d+1}\}$, let z be the outgoing label of \mathbf{u} , and suppose that k is the index such that $x_{k-1} < z < x_k$ (setting $k = 0$ if $z < x_1$ and $k = d + 2$ if $z > x_{d+1}$). If $\mathbf{w} := \{x_1 < \dots < x_{k-1} < z < x_k < \dots < x_{d+1}\}$ then the well-known property $\partial_d \partial_{d+1} = 0$ produces the sought linear combination, since

$$0 = \partial_d \partial_{d+1} \mathbf{w} = \sum_{i=1}^{k-1} (-1)^i \partial_d \mathbf{v}_i + (-1)^k \partial_d \mathbf{u} + \sum_{i=k}^{d+1} (-1)^{i+1} \partial_d \mathbf{v}_i.$$

With this fact in hand, one can construct the matrix corresponding to ∂_d as follows: begin with the columns $\partial_d \mathbf{v}_1, \dots, \partial_d \mathbf{v}_l$ where $\mathbf{v}_1, \dots, \mathbf{v}_l$ are the leaves of G ; then repeatedly add the column $\partial_d \mathbf{u}$ for an internal vertex \mathbf{u} all of whose children were

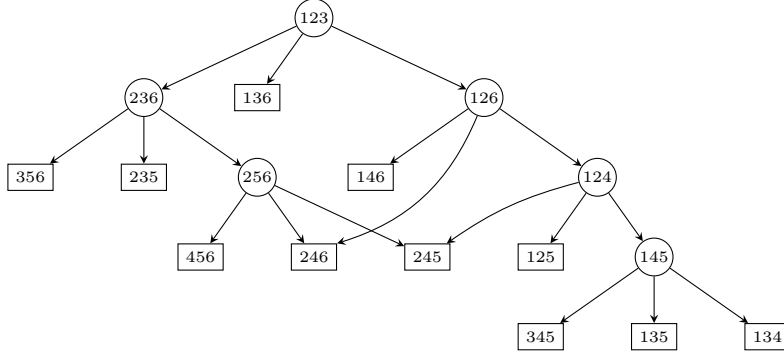


FIGURE 5. A $(6, 11, 3)$ -pedigree G for $d = 2$ whose simplicial complex has $\beta_2 > m$: as the dimension of the leaves is at most $\binom{s+d}{2}$, in this case it is at most $10 < l$, not linearly independent (here $\beta_2 = 7$ vs. $m = 6$).

all added. As each such iteration adds a column which is a linear combination of its predecessors, we deduce that

$$\beta_d = \dim \ker(\partial_d) \geq m. \quad (3.5)$$

(Note that this inequality may be strict: see Figure 5.)

On the other hand, an upper bound on β_d can be derived from its characterization

$$\beta_d = b_d(\Delta(\bar{K})) = |L| \quad \text{where} \quad L := \{f \in \Delta(\bar{K}) : |f| = d+1, 1 \notin f\}.$$

Recalling that $\mathbf{v}_0 \in \Delta(\bar{K})$ by Claim 3.7, define

$$\begin{aligned} F &:= \{f \in \binom{[s+d+1]}{d+1} : f \leq_P \mathbf{v}_0, 1 \in f\} \subseteq \Delta(\bar{K}) \quad \text{and} \\ \hat{F} &:= \{f \in \Delta(\bar{K}) : |f| = d+1, 1 \in f\} \setminus F, \end{aligned}$$

where $F \subseteq \Delta(\bar{K})$ as per the definition of a shifted complex containing \mathbf{v}_0 , and set

$$\hat{\mathbf{a}} := |\hat{F}| = f_d(\Delta(\bar{K})) - \beta_d - |F|.$$

We readily see that

$$F = \{[d+1]\} \cup \{[d+1] \cup \{x+d+1\} \setminus \{y\} : x \in [s], y = 2, \dots, d+1\},$$

so that

$$|F| = ds + 1$$

and substituting this (as well as $|\Delta(\bar{K})| = m + l$) in the expression for $\hat{\mathbf{a}}$ we get

$$\hat{\mathbf{a}} = l - (ds + 1) - (\beta_d - m) = \mathbf{a} - (\beta_d - m) \leq \mathbf{a}, \quad (3.6)$$

with the last inequality using (3.5). In particular, $\mathbf{a} \geq |\hat{F}| \geq 0$.

We next aim to bound $|L|$ in terms of $\hat{\mathbf{a}}$. Write L as the disjoint union of

$$\begin{aligned} L_1 &= \{\{x_1 < \dots < x_{d+1}\} \in L : x_d \leq d+1\}, \\ L_2 &= \{\{x_1 < \dots < x_{d+1}\} \in L : x_d > d+1, x_1 \leq d+1\}, \\ L_3 &= \{\{x_1 < \dots < x_{d+1}\} \in L : x_1 \geq d+2\}. \end{aligned}$$

Recalling $1 \notin f$ for $f \in L_1$, clearly L_1 is $\{(2, \dots, d+1, x+d+1) : x \in [s]\}$, and so

$$|L_1| = s.$$

To treat L_2 , consider the map T which replaces the smallest element of $f \in L_2$ by 1. Since every $f \in L_2$ has at least 2 elements that exceed $d + 1$, we see that $Tf \not\prec_P \mathbf{v}_0$, whence $TL_2 \subseteq \widehat{F}$. Moreover, every element of \widehat{F} has at most d preimages of T in L_2 thanks the requirement $2 \leq x_1 \leq d + 1$, hence

$$|L_2| \leq d\widehat{\mathbf{a}}.$$

Lastly, to handle L_3 , set

$$\widehat{F}' = \{f \setminus \{1\} : f \in \widehat{F}\}.$$

For every $f = \{x_1, \dots, x_{d+1}\} \in L_3$ and every i , we have that

$$f \setminus \{x_i\} \in \widehat{F}',$$

since it has d elements (and in particular, at least 2 elements) that exceed $d + 1$ by the definition of L_3 , whence $f \cup \{1\} \setminus \{x_i\} \not\prec_P \mathbf{v}_0$. Applying the Kruskal–Katona theorem in its simpler form due to Lovász (see [29, p. 95, Ex. 13.31(b)]) we find that, if $|L_3| = \binom{\theta}{d+1}$ for $\theta \in \mathbb{R}_+$ then $\binom{\theta}{d} \leq |\widehat{F}'| \leq \widehat{\mathbf{a}}$ (and $|L_3| = \binom{\theta}{d} \frac{\theta-d}{d+1}$), whence

$$|L_3| \leq \widehat{\mathbf{a}} \frac{(d!\widehat{\mathbf{a}})^{1/d}}{d+1} \leq (\widehat{\mathbf{a}})^{(d+1)/d}.$$

Combining these estimates shows that $\beta_d = |L|$ satisfies

$$\beta_d \leq s + d\widehat{\mathbf{a}} + (\widehat{\mathbf{a}})^{(d+1)/d}. \quad (3.7)$$

Using this estimate we may now conclude the proof. To bound $\mathbf{b} = dm - (l - 1)$ from above we can increase m to β_d (with (3.5) in mind) and get

$$\mathbf{b} \leq d\beta_d - (l - 1) = d(\beta_d - s) - (l - (ds + 1)) \leq d\mathbf{a}^{(d+1)/d} + (d^2 - 1)\mathbf{a},$$

where the last inequality used (3.7) and the fact that $\widehat{\mathbf{a}} \leq \mathbf{a}$. \blacksquare

3.2. Proof of Theorem 3.1. Let $\mathcal{E}_{m,l,s}$ be the event that $Y \sim \mathcal{Y}_d(n, p)$ admits a stacked contraction of $\mathbf{o} = \{1, 2, \dots, d+1\}$ corresponding to some (m, l, s) -pedigree, so that $\mathcal{E}_s^\mathbf{o} = \bigcup_{m,l,s} \mathcal{E}_{m,l,s}$.

Define \mathbf{a}, \mathbf{b} as (3.2)–(3.3), and note that if $\mathbf{a} = 0$ then also $\mathbf{b} = 0$ by Theorem 3.5 (while this follows from (3.4), it can be readily proved via an elementary argument examining the pedigree under consideration; we will need the delicate bound (3.4) later in order to treat $\mathbf{b} > 0$). In particular, in this case $m = s$ and $l = ds + 1$, and the (m, l, s) -pedigrees are precisely the set of $(d + 1)$ -ary trees with s internal vertices. As the number of such trees is $\mathbf{C}_d(s) \leq \alpha_d^s$, we thus have the upper bound

$$\mathbb{P}(\mathcal{E}_{s, ds+1, s}) \leq \binom{n}{s} s! \mathbf{C}_d(s) p^{ds+1} \leq (1 - \delta)^{ds} p \leq e^{-\delta ds} p.$$

Since $\mathbf{a} \geq 0$ holds for any (m, l, s) -pedigree as per (the easy part of) Theorem 3.5, it remains to treat $\mathbf{a} \geq 1$, which we henceforth assume, noting this implies $s \geq 2$.

Let us first narrow down the relevant values of the parameter $2 \leq s \leq (\log n)^{5/2}$ and $\mathbf{a} = l - (ds + 1)$ via a crude union bound that ignores the DAG structure (including its number of internal vertices) and merely looks at the probability that Y contains l faces using at most $s + d + 1$ different labels:

$$\begin{aligned} \mathbb{P}\left(\bigcup_m \mathcal{E}_{m,l,s}\right) &\leq \binom{n}{s} \binom{s+d+1}{d+1} p^l \leq \left(n^{\frac{1}{d}} p (s + d + 1)^{d+1}\right)^{ds} ((s + d + 1)^{d+1} p)^{\mathbf{a}+1} \\ &\leq s^{c_d(s+\mathbf{a}+1)} n^{-(\mathbf{a}+1)/d} \leq s^{c_d s} n^{-(\frac{1}{d}-o(1))(\mathbf{a}+1)}, \end{aligned} \quad (3.8)$$

for some $c_d > 0$, where it was possible to increase $c(s+c)^{cs}$ into $s^{c's}$ having $s \geq 2$, and we thereafter absorbed the term $s^{O(\mathbf{a})} = \exp[O(\mathbf{a} \log \log n)]$ into the $o(1)$ -term. If $s \leq \sqrt{\log n}$ then the right-hand of (3.8) is in this case at most

$$\exp \left[O(\sqrt{\log n} \log \log n) - (\mathbf{a} + 1) \left(\frac{1}{d} - o(1) \right) \log n \right] \leq n^{-2/d+o(1)},$$

with the last step using the fact that $\mathbf{a} \geq 1$. We may thus assume henceforth that

$$\sqrt{\log n} \leq s \leq (\log n)^{5/2}. \quad (3.9)$$

Reexamining (3.8) for this adjusted range of s , we see that if $\mathbf{a} \geq s^{5/8}$ then

$$s \log s \leq \mathbf{a} s^{3/8} \log s \leq \mathbf{a} (\log n)^{\frac{15}{16}+o(1)} = o(\mathbf{a} \log n),$$

so the right-hand of (3.8) is in this case at most $n^{-(1/d-o(1))\mathbf{a}} \leq n^{-\sqrt{\log n}}$ for large enough n ; thus, we can assume

$$\mathbf{a} \leq s^{5/8}. \quad (3.10)$$

Recalling the definitions of \mathbf{a}, \mathbf{b} from (3.2)–(3.3), we have $d(m-s) = \mathbf{a} + \mathbf{b}$, and so (3.4) yields that, in the above range of \mathbf{a} ,

$$0 \leq m-s \leq d(\mathbf{a}^{(d+1)/d} + \mathbf{a}) = O(\mathbf{a}^{(d+1)/d}) = O(\mathbf{a}^{3/2}) = O(s^{\frac{15}{16}}) = o(s)$$

and we deduce that

$$m = (1 + o(1))s.$$

We now wish to bound $\mathbb{P}(\mathcal{E}_{l,s})$ by enumerating over (m, l, s) -pedigrees. To this end, we argue that the number of unlabeled DAGs which can be obtained by erasing the vertex labels of an (m, l, s) -pedigree is at most

$$((d+1)m)^{2\mathbf{b}} \mathbf{C}_d(m);$$

indeed, given a DAG that arises from an (m, l, s) -pedigree, one can repeatedly (for \mathbf{b} steps) take a tree excess edge and replace it by an out-edge to a new leaf, arriving at a $(d+1)$ -ary tree T with m internal vertices (and $dm+1$ leaves). To recover the original DAG, one can repeatedly (for \mathbf{b} steps) choose a leaf (of those we had added, identifying its parent as the source of an excess edge) and a vertex (the target of said edge), having $(dm+1)((d+1)m+1) \leq ((d+1)m)^2$ options per step (this inequality holds for all m large enough); the above estimate now follows from the fact that the number of $(d+1)$ -ary trees with m internal vertices is $\mathbf{C}_d(m)$.

For each unlabeled DAG as above, one can assign vertex labels from $\binom{[n]}{d+1}$ using a total of s distinct outgoing labels, by selecting said s labels from $[n]$, then choosing the s internal vertices which will have those outgoing labels, and thereafter labeling the remaining $m-s$ vertices; carrying the labeling process in this way leverages the fact that $m-s = (\mathbf{a} + \mathbf{b})/d$, which we have good control over.

Combining these two arguments, we conclude that

$$\begin{aligned} \mathbb{P}(\mathcal{E}_{m,l,s}) &\leq \binom{n}{s} \binom{m}{s} s! s^{m-s} ((d+1)m)^{2\mathbf{b}} \mathbf{C}_d(m) p^l \\ &\leq (\alpha_d n p^d)^s \binom{m}{m-s} ((d+1)m)^{2\mathbf{b}} (\alpha_d s)^{(\mathbf{a}+\mathbf{b})/d} p^{\mathbf{a}+1}, \end{aligned}$$

where the last inequality used that $\mathbf{C}_d(m) \leq \alpha_d^n$ and that $m = s + (\mathbf{a} + \mathbf{b})/d$. Substituting the value of p , we have $\alpha_d n p^d = (1-\delta)^d$, and using $\binom{m}{m-s} \leq m^{(\mathbf{a}+\mathbf{b})/d}$

we further find that

$$\mathbb{P}(\mathcal{E}_{m,l,s}) \leq \frac{(1-\delta)^{ds}}{(\alpha_d n)^{1/d}} \left(\frac{ms}{n}\right)^{a/d} \left(\alpha_d (d+1)^{2d} m^{2d+1} s\right)^{b/d}.$$

In what is our central application of Theorem 3.5, we can next bound \mathbf{b} in terms of \mathbf{a} and find that

$$\mathbf{b} \leq d\mathbf{a}^{(d+1)/d} + d^2\mathbf{a} \leq 2d^2\mathbf{a}s^{5/16} \leq \mathbf{a}s^{1/3},$$

and plugging this in the last bound on $\mathbb{P}(\mathcal{E}_{m,l,s})$, as well as $\alpha_d = \frac{(d+1)^{d+1}}{d^d} \leq e(d+1)$ for all $d \geq 2$, we get that

$$\begin{aligned} \mathbb{P}(\mathcal{E}_{m,l,s}) &\leq \frac{(1-\delta)^{ds}}{(\alpha_d n)^{1/d}} \left[\frac{ms}{n} \left(e(d+1)^{2d+1} m^{2d+1} s \right)^{s^{1/3}} \right]^{a/d} \\ &\leq \frac{(1-\delta)^{ds}}{(\alpha_d n)^{1/d}} \left[n^{-1} \exp\left((2(d+1) + o(1)) s^{1/3} \log s \right) \right]^{a/d} \\ &\leq e^{-\delta ds} n^{(-\frac{1}{d} + o(1))(a+1)}, \end{aligned}$$

with the last equality following from the fact that

$$s^{1/3} \log s \leq (\log n)^{5/6 + o(1)} = o(\log n).$$

We conclude with a union bound over $\bigcup_{m,l} \mathcal{E}_{m,l,s}$ (going over l via $\mathbf{a} = l - (ds + 1)$). The probability of this union of events for a given \mathbf{a} and all $m \leq (1 + o(1))s = n^{o(1)}$ is at most $e^{-\delta ds} n^{(-1/d + o(1))(a+1)}$; summing this over all $\mathbf{a} \geq 1$ concludes the proof. ■

4. THE SUPERCRITICAL REGIME

We establish the supercritical regime in Theorem 1 by proving that a stronger property occurs with high probability. Namely, that $\mathcal{Y}_d(n, p)$ admits a so-called balanced stacked triangulation for every $f \in \binom{[n]}{d+1}$. To be consistent with the previous section, we state this result in the language of pedigrees.

Definition 4.1. A *proper* s -pedigree is an (m, l, s) -pedigree T in which all the outgoing labels of its internal vertices are distinct, and differ from those on the root of T . In such a case $m = s$, T is a tree and we denote the set consisting of its $l = ds + 1$ leaves by $\mathcal{L}(T)$.

Given a vertex \mathbf{v} of T , let the subtree $T_{\mathbf{v}}$ of T rooted at T be the subgraph induced by \mathbf{v} and all its descendants. If the numbers of leaves in each of the $d + 1$ subtrees of T rooted at the children of T 's root are equal, T is called *balanced*. In such a case $s = (d + 1)s' + 1$ for some s' , and each such subtree has $ds' + 1$ leaves.

Theorem 4.2. Fix $d \geq 2$ and let $\delta = 1/(18(d+1)^2)$, $\varepsilon = n^{-\delta}$, $p = (1 + \varepsilon)(\alpha_d n)^{-1/d}$, $s' = n^{2\delta}$ and $s = (d + 1)s' + 1$. Let \mathcal{A}_s denote the event that there is a balanced proper s -pedigree for $\{1, \dots, d + 1\}$ whose leaves are all members of $Y \sim \mathcal{Y}_d(n, p)$. Then, for every sufficiently large n ,

$$\mathbb{P}(\mathcal{A}_s) \geq 1 - \exp(-n^{12\delta}).$$

The upper bound in our main theorem will follow immediately from the above theorem by taking a union bound over the $\binom{n}{d+1}$ possible roots.

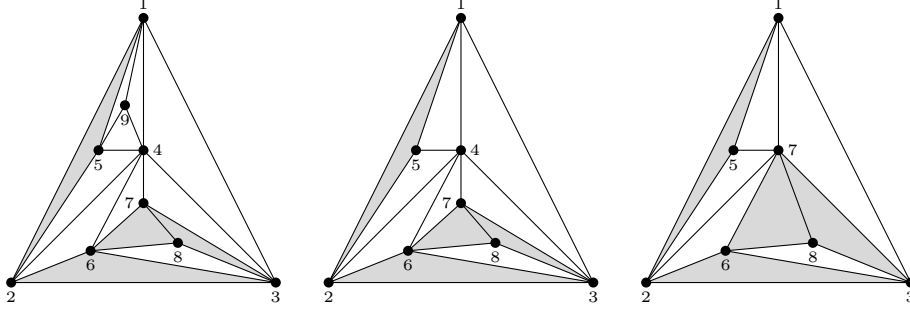


FIGURE 6. An illustration of Lemma 4.3: given a subset P of the faces of a proper stacked triangulation of $\mathbf{o} = \{1, 2, 3\}$ (shaded faces on left picture), there exists such a triangulation of \mathbf{o} containing these faces and no other labels (on right). In this example, to generate the latter, one first deletes the vertex 9 (middle picture), then contracts the vertices 4, 7.

4.1. Combinatorics of proper pedigrees. We begin the proof with some combinatorial properties of proper pedigrees. The following lemma shows that, for any proper pedigree of \mathbf{o} and any prescribed subset P of its faces, there exists a proper pedigree of \mathbf{o} that contains this subset P and no additional labels (see Figure 6). This fact will play a key role in characterizing (and enumerating) the pairs of proper pedigrees with a given intersection, as part of the second moment analysis.

Lemma 4.3. *Let T be a proper s -pedigree for $\mathbf{o} = \{x_1, \dots, x_{d+1}\}$, $P \subseteq \mathcal{L}(T)$, $R = \bigcup_{\mathbf{v} \in P} \mathbf{v} \setminus \mathbf{o}$ and $r = |R|$. Then,*

- (1) *There exists a proper r -pedigree $H = H(P, T)$ for \mathbf{o} such that $P \subseteq \mathcal{L}(H)$. Furthermore, $P = \mathcal{L}(H)$ if and only if $P = \mathcal{L}(T)$.*
- (2) *If $P \subsetneq \mathcal{L}(T)$ then $|P| \leq dr$, and equality holds if and only if $P = \mathcal{L}(T) \setminus \mathcal{L}(T_{\mathbf{v}})$ for some vertex \mathbf{v} of T .*

Proof. We start with the proof of item (1). Let z be the outgoing label in T of the root \mathbf{o} . For every $1 \leq j \leq d+1$, the subtree T_j of T rooted at $\mathbf{v}_j = \mathbf{o} \setminus \{x_j\} \cup \{z\}$ is a proper s_j -pedigree for \mathbf{v}_j , for some $s_j < s$. Let $P_j := P \cap \mathcal{L}(T_j)$, $R_j = \bigcup_{\mathbf{v} \in P_j} \mathbf{v} \setminus \mathbf{v}_j$ and $r_j = |R_j|$. By induction on s , we find $H_j = H(P_j, T_j)$ — an $|R_j|$ -pedigree for \mathbf{v}_j such that $P_j \subseteq \mathcal{L}(H_j)$, and $P_j = \mathcal{L}(H_j)$ if and only if $P_j = \mathcal{L}(T_j)$. The label sets R_1, \dots, R_{d+1} are disjoint and contained in R . If $z \notin R$ then $R = \bigcup_j R_j$ and otherwise $R = \{z\} \cup \bigcup_j R_j$. In particular, $r = \mathbf{1}_{\{z \in R\}} + \sum_j r_j$.

Let H' be the proper $(1 + \sum_j r_j)$ -pedigree for \mathbf{o} obtained from T by replacing each T_j with H_j . If $z \in R$ it is straightforward to see that setting H as H' concludes the proof. Otherwise, the label $z \notin R$ needs to be removed from H' . We further assume that $P \neq \emptyset$ or else we set H as the isolated vertex \mathbf{o} . Without loss of generality, suppose that $P_{d+1} \neq \emptyset$.

Let $\mathbf{u} = \{y_1, \dots, y_d, z\}$ be an internal vertex in H_{d+1} containing z of maximal distance from \mathbf{o} . Such an internal vertex exists because $z \notin R$ whence the root \mathbf{v}_{d+1} of H_{d+1} is an internal vertex containing z . Consequently, if we denote the outgoing label of \mathbf{u} by y , then all the vertices $\mathbf{u} \setminus \{y_i\} \cup \{y\}$, $i = 1, \dots, d$ contain z and of greater distance to \mathbf{o} than \mathbf{u} and, therefore, are leaves. Let $T_{\mathbf{u}'}$ be the subtree of H_{d+1} rooted at $\mathbf{u}' = \{y_1, \dots, y_d, y\}$.

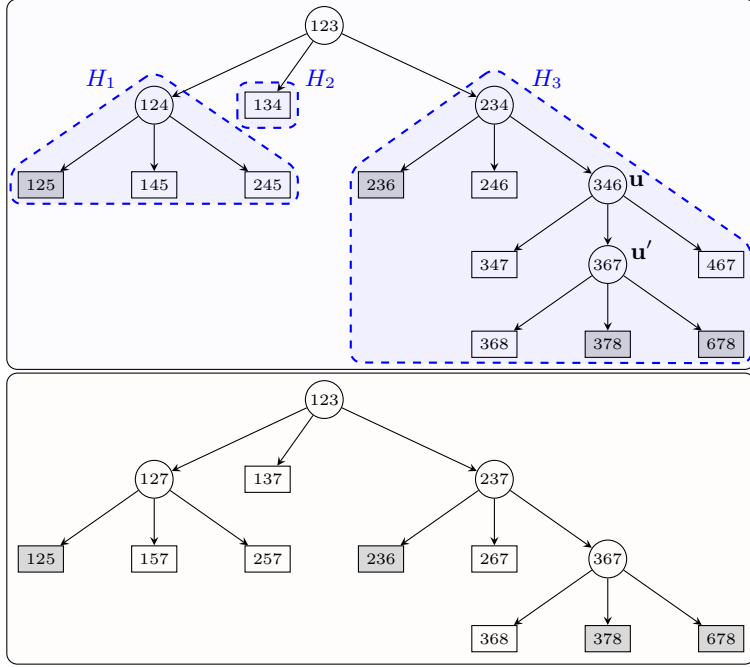


FIGURE 7. An example of the construction of H (bottom) from H' (top) in the non-trivial case in which $P \neq \emptyset$ (shaded) and $z = 4 \notin R$. We use the rightmost tree to replace z , so, in the notations of the proof, $\mathbf{u} = 346$, $y = 7$ and $T_{\mathbf{u}'}$ is the tree descending from $\mathbf{u}' = 367$. To obtain H we delete the descendants of 346 , replace all the appearances of 4 by 7 , and attach $T_{\mathbf{u}'}$ to the vertex 367 that replaced 346 . Alternatively, the leftmost subtree could have been used with $\mathbf{u} = 124$ and $y = 5$.

We construct $H = H(P, G)$ from H' by the following procedure. First remove all the descendants of \mathbf{u} from H' . Afterwards, replace every vertex \mathbf{v} in H' containing z with $\mathbf{v} \setminus \{z\} \cup \{y\}$. In particular, the vertex \mathbf{u} is replaced by \mathbf{u}' . The final step is to append the subtree $T_{\mathbf{u}'}$ to \mathbf{u}' (See Figure 7).

The fact that $P \subseteq \mathcal{L}(H)$ follows from observing that the only leaves in $\mathcal{L}(H') \setminus \mathcal{L}(H)$ are the d children of \mathbf{u} which were deleted from H , namely $\mathbf{u} \setminus \{y_i\} \cup \{y\}$, $i = 1, \dots, d$. All these leaves contain $z \notin R$ hence they do not belong to P . In addition, $P \neq \mathcal{L}(T)$ since $z \notin R$, so we also need to show that $P \neq \mathcal{L}(H)$. To prove this, we note that there exists $\mathbf{v} \in \mathcal{L}(H_1)$ that contains z (we could have chosen a leaf from any of the trees H_1, \dots, H_d — but not from H_{d+1} — and H_1 is selected arbitrarily here). Indeed, since the root \mathbf{v}_1 of H_1 contains z , one of its leaves must also contain it. We claim that $\mathbf{v}' = \mathbf{v} \setminus \{z\} \cup \{y\} \in \mathcal{L}(H) \notin P$. First, $\mathbf{v}' \in \mathcal{L}(H)$ since our construction replaces every z with a y . Second, \mathbf{v}' contains a label that appears in T only in vertices of T_1 . On the other hand, the label y appears in T only in vertices of T_{d+1} . Therefore \mathbf{v}' is not a leaf of T , and does not belong to P .

We turn to the proof of item (2). By the first item, we have that $P \subsetneq \mathcal{L}(H)$ whence $|P| < |\mathcal{L}(H)| = dr + 1$. Furthermore, if $P = \mathcal{L}(T) \setminus \mathcal{L}(T_{\mathbf{v}})$ then, by construction, H is obtained from T by removing all the descendants of \mathbf{v} . Therefore, \mathbf{v} is the only leaf of H that is not in P whence $|P| = dr$. On the other hand, if

$|P| = dr$ then P is obtained from $\mathcal{L}(H)$ by removing one leaf. In consequence, for $j = 1, \dots, d$, w.l.o.g., $P_j = \mathcal{L}(H_j)$ whence, by the first item, $P_j = \mathcal{L}(T_j)$. In particular, $z \in R$ and we have that

$$|P_{d+1}| = dr - \sum_{j=1}^d (dr_j + 1) = d \left(\mathbf{1}_{\{z \in R\}} + \sum_{j=1}^{d+1} r_j \right) - \sum_{j=1}^d (dr_j + 1) = dr_{d+1}.$$

The claim follows by induction. \blacksquare

Fix the root $\mathbf{o} = \{1, \dots, d+1\}$, and consider the families

$$\mathcal{L} = \{\mathcal{L}(T) : T \text{ is a balanced proper } s\text{-pedigree for } \mathbf{o}\}$$

and

$$\mathcal{P} = \{P : \exists L \in \mathcal{L}, \emptyset \neq P \subsetneq L\}.$$

Let $z \in [n]$ and $K \subseteq \binom{[n]}{d+1}$. The link of z in K is

$$\text{lk}_z(K) = \{\mathbf{v} \setminus \{z\} : z \in \mathbf{v} \in K\}.$$

Definition 4.4. Let $P \in \mathcal{P}$. We say that an internal label $z \in \cup_{\mathbf{v} \in P} \mathbf{v} \setminus \mathbf{o}$ is an encircled label of P if $\text{lk}_z(P) = \text{lk}_z(L)$ for every $L \in \mathcal{L}$ containing P . Otherwise, z is called a boundary label of P .

Claim 4.5. For every $L \in \mathcal{L}$ and $P \subseteq L$, every internal label $z \in \cup_{\mathbf{v} \in L} \mathbf{v} \setminus \mathbf{o}$ of L satisfies exactly one of the following: z is an encircled label of P , z is an encircled label of $L \setminus P$, or z is a boundary label of both P and $L \setminus P$.

Proof. By viewing a set $L \in \mathcal{L}$ as the d -dimensional faces of a stacked triangulation of the simplex \mathbf{o} , and an internal label z as an internal vertex in this triangulation, we observe that the link $\text{lk}_z(L)$ consists of the $(d-1)$ -faces of a triangulation of the $(d-1)$ -dimensional sphere \mathbb{S}^{d-1} (and in fact this triangulated sphere is the combinatorial boundary of a stacked polytope). This observation gives rise to the following intrinsic characterization of an encircled label. Let $P \in \mathcal{P}$ and z be an internal label that appear in P . We claim that z is encircled in P if and only if $\text{lk}_z(P)$ consists of the $(d-1)$ -faces of a triangulation of \mathbb{S}^{d-1} . This condition is clearly necessary since the property that the link of z is a triangulation of \mathbb{S}^{d-1} holds true for every $L \in \mathcal{L}$. In addition, if for some $L \in \mathcal{L}$ that contains P there holds $\text{lk}_z(P) \subsetneq \text{lk}_z(L)$ then $\text{lk}_z(P)$ is not a triangulation of \mathbb{S}^{d-1} — since no strict subset of \mathbb{S}^{d-1} is homeomorphic to \mathbb{S}^{d-1} . Hence, this condition is also sufficient.

We now turn to the proof of the claim. For every internal label $z \in \cup_{\mathbf{v} \in L} \mathbf{v} \setminus \mathbf{o}$ of L , its link $\text{lk}_z(P)$ in P is either equal to $\text{lk}_z(L)$, empty, or neither of these. The first case holds true if and only if $\text{lk}_z(P)$ is a triangulation of \mathbb{S}^{d-1} whence z is an encircled label of P . Similarly, the second case is satisfied if and only if $\text{lk}_z(L \setminus P) = \text{lk}_z(L)$ is a triangulation of \mathbb{S}^{d-1} whence z is an encircled label of $L \setminus P$. In the third case neither $\text{lk}_z(P)$ nor $\text{lk}_z(L \setminus P)$ is a triangulation of \mathbb{S}^{d-1} and z is a boundary label of both sets. \blacksquare

Lemma 4.6. The following holds for every integers $0 \leq t \leq r \leq s$.

- (1) The number of sets $P \in \mathcal{P}$ with $r-t$ encircled labels and t boundary labels is at most $2(\alpha_d n)^r (2^d r)^t$.
- (2) For every set $P \in \mathcal{P}$ with $r-t$ encircled labels and t boundary labels, the number of $L \in \mathcal{L}$ containing P is at most $2(\alpha_d n)^{s-r} (\alpha_d 2^d (s-r+t))^t$.

Proof. For the first item, suppose that the set $B \in \binom{\{d+2, \dots, n\}}{t}$ of the t boundary labels of P is fixed. Recall that, by item (1) of Lemma 4.3, every $P \in \mathcal{P}$ with r internal labels is contained in the set of leaves a proper r -pedigree for \mathbf{o} . To construct such a pedigree H we (i) select a $(d+1)$ -ary trees with r internal nodes, (ii) split the nodes to the t nodes that will get a boundary label from B and the $r-t$ node that will get an encircled label from $[n] \setminus (B \cup \{1, \dots, d+1\})$, (iii) select the encircled labels and (iv) assign the labels to the nodes accordingly. We construct P from the pedigree H as following. First, we must include in P all the leaves containing an encircled label. The remaining leaves are spanned by the t boundary labels and by Lemma 4.3 (in which we now let H and t play the roles of T and r , respectively) there are at most $dt + \mathbf{1}_{\{r=t\}}$ of them. We conclude the construction of P by selecting a subset of these leaves.

In summary, the number of sets $P \in \mathcal{P}$ with a specific set B of t boundary labels and $r-t$ encircled labels is at most

$$C_d(r) \binom{r}{t} \binom{n-d-1-t}{r-t} t!(r-t)! 2^{dt+1} \leq 2(\alpha_d n)^{r-t} (\alpha_d 2^d r)^t, \quad (4.1)$$

and the first item is proved by additionally selecting the set B .

For the second item, fix a set $P \in \mathcal{P}$ with the above parameters, and let B denote the set of its t boundary labels. Every $L \in \mathcal{L}$ containing P is determined by the set $P' := L \setminus P$. By Claim 4.5, B is the set of boundary labels of P' , and in addition P' has $s-r$ encircled labels that do not appear in P . Thus, by (4.1), the number of such sets P' (and in turn, the number of $L \in \mathcal{L}$ with $L \supseteq P$) is at most

$$2(\alpha_d n)^{s-r} (\alpha_d 2^d (s-r+t))^t,$$

as claimed. ■

4.2. Proof of Theorem 4.2. Let $Y \sim \mathcal{Y}_d(n, p)$. Throughout the proof, we denote by c_d a constant depending on d that may change from line to line. We define

$$\begin{aligned} \mu &= \sum_{L \in \mathcal{L}} \mathbb{P}(\{L \subseteq Y\}), \\ \rho &= \sum_{\substack{L_1 \neq L_2 \in \mathcal{L}, \\ L_1 \cap L_2 \neq \emptyset}} \mathbb{P}(\{L_1, L_2 \subseteq Y\}). \end{aligned}$$

The cardinality of \mathcal{L} is $C_d(s')^{d+1} (n-d-1)_s$ and every $L \in \mathcal{L}$ is of size $ds+1$ whence

$$\begin{aligned} \mu &= C_d(s')^{d+1} (n-d-1)_s p^{ds+1} \\ &= (c_d + o(1)) (\alpha_d n)^{-1/d} (s')^{-3(d+1)/2} (1+\varepsilon)^{ds+1} = e^{(d(d+1)+o(1))n^\delta}. \end{aligned} \quad (4.2)$$

We can express ρ as

$$\begin{aligned} \rho &= \sum_{P \in \mathcal{P}} \sum_{\substack{L_1 \neq L_2 \in \mathcal{L}, \\ L_1 \cap L_2 = P}} \mathbb{P}(\{L_1, L_2 \subseteq Y\}) = \sum_{P \in \mathcal{P}} \sum_{\substack{L_1 \neq L_2 \in \mathcal{L}, \\ L_1 \cap L_2 = P}} p^{2(ds+1)-|P|} \\ &\leq \sum_{P \in \mathcal{P}} |\{L \in \mathcal{L} : P \subseteq L\}|^2 p^{2(ds+1)-|P|}. \end{aligned}$$

The inequality is obtained by summing over all L_1, L_2 whose intersection contains P rather than distinct L_1, L_2 whose intersection is equal to P . By Lemma 4.6 we find

$$\rho \leq 8 \sum_{m,r,t} (\alpha_d n)^{m/d-r-2/d} (1+\varepsilon)^{2(ds+1)-m} (\alpha_d 2^d s)^{3t}, \quad (4.3)$$

where the summation is over all admissible tuples m, r and t for which there exists $P \in \mathcal{P}$ with $|P| = m$, $r - t$ encircled labels and t boundary labels. Combining (4.3) and the LHS of (4.2) yields

$$\begin{aligned} \rho/\mu^2 &\leq c_d \sum_{m,r,t} (\alpha_d n)^{f/d-r} (1+\varepsilon)^{-m} (\alpha_d 2^d s)^{3(t+d+1)} \\ &\leq c_d \max_{m,r,t} (\alpha_d n)^{m/d-r} (1+\varepsilon)^{-m} (\alpha_d 2^d s)^{3(t+d+2)}. \end{aligned} \quad (4.4)$$

The last inequality follows from the fact that $t \leq r \leq s$ and $m \leq ds$.

We now turn to bound the expression in (4.4), where the power $m/d - r$ of n plays a key role. Lemma 4.3 asserts that $m/d - r \leq 0$, and we claim that if equality holds then $m \geq d(ds' + 1)$. Indeed, consider a set $P \in \mathcal{P}$ for which $m = dr$, and let T be a balanced proper s -pedigree satisfying $P \subseteq \mathcal{L}(T)$. By item (2) of Lemma 4.3, the equality $m = dr$ implies that $P = \mathcal{L}(T) \setminus \mathcal{L}(T_{\mathbf{v}})$ for some vertex \mathbf{v} of T . Since T is balanced — and this is the crucial place we use this assumption — $T_{\mathbf{v}}$ contains at most $ds' + 1$ leaves, and

$$m = |P| \geq ds + 1 - (ds' + 1) = d(ds' + 1).$$

In addition, recall that P is obtained from a proper r -pedigree H by removing $dr + 1 - m$ leaves. Every leaf removal increases the number of the boundary labels of P by at most $d + 1$. Therefore, we find that $t \leq (d + 1)(dr + 1 - m)$, and equivalently, $m/d - r \leq 1/d - t/(d(d + 1))$.

Denote the expression in (4.4) by

$$\Xi := (\alpha_d n)^{m/d-r} (1+\varepsilon)^{-m} (\alpha_d 2^d s)^{3(t+d+2)},$$

and consider the following three cases.

Case (i). $m = dr$. By the discussion above, this can occur only if $t \leq d + 1$ and $m \geq d(ds' + 1)$. Therefore,

$$\Xi \leq e^{-(d^2+o(1))n^\delta}.$$

Case (ii). $t \geq 2(d + 1)$. We recall that $m/d - r \leq 1/d - t/(d(d + 1))$ and find

$$\Xi \leq (\alpha_d n)^{1/d-t/(d(d+1))} (\alpha_d 2^d s)^{3(t+d+2)}. \quad (4.5)$$

The derivative of the logarithm of the RHS of (4.5) with respect to t is

$$-\frac{\log n}{d(d+1)} + 3 \log s + O_d(1) = \log n \left(-\frac{1}{d(d+1)} + 6\delta \right) + O_d(1),$$

which is negative for all sufficiently large integers n . Therefore, we can bound Ξ by assigning $t = 2(d + 1)$ in the RHS of (4.5), yielding

$$\Xi \leq n^{-\frac{2d+3}{3d(d+1)^2}+o(1)}.$$

Case (iii). $m \leq dr - 1$ and $t \leq 2d + 1$. By a direct assignment we have that

$$\Xi \leq n^{-\frac{1}{d(d+1)}+o(1)}.$$

In summary, and using (4.4), we find that

$$\rho/\mu^2 \leq n^{-\frac{2d+3}{3d(d+1)^2} + o(1)}. \quad (4.6)$$

Substituting (4.2) and (4.6) in Janson's inequality then establishes that

$$\mathbb{P}\left(\bigcap_{L \in \mathcal{L}} \{L \not\subseteq Y\}\right) \leq \exp\left(-\frac{\mu^2}{\mu + \rho}\right) = \exp\left(-n^{\frac{2d+3}{3d(d+1)^2} + o(1)}\right) < \exp\left(-n^{12\delta}\right),$$

with the last inequality valid for large enough n . This completes the proof. \blacksquare

5. A CLOSER LOOK AT THE SUBCRITICAL REGIME

In this section we prove Theorem 2. Let $0 \leq \gamma < \alpha_d^{-1/d}$, $p = \gamma n^{-1/d}$ and $Y \sim \mathcal{Y}_d(n, p)$. Recall from §2.2 that the smallest positive root $\hat{\gamma}$ of $Q_\gamma(x) = x^{d+1} - x + \gamma$ is given by

$$\hat{\gamma} := \sum_{s=0}^{\infty} C_d(s) \gamma^{ds+1}.$$

For an integer $s \geq 0$, we denote by \mathfrak{T}_s the set of $(d+1)$ -ary trees with s internal vertices, and let $\mathfrak{T}_{\leq s} = \cup_{s' \leq s} \mathfrak{T}_{s'}$. In addition, for every $s \geq 0$ and a tree $T \in \mathfrak{T}_s$, let X_T be the number of proper s -pedigrees whose leaves are all members of Y and whose underlying tree is T . We prove Theorem 2 using the following lemmas.

Lemma 5.1. *Fix an integer $s \geq 0$ and $T \in \mathfrak{T}_s$. Then,*

$$n^{1/d} \frac{X_T}{\binom{n}{d+1}} \rightarrow \gamma^{ds+1}$$

in probability as $n \rightarrow \infty$.

Proof. We enumerate the proper s -pedigrees whose underlying tree is T by its root $\mathbf{o} \in \binom{[n]}{d+1}$ and the internal labels z_1, \dots, z_s given in some predetermined order (e.g., by the order of their appearance in a Breadth-First-Search on T).

The expectation of X_T is given by

$$\begin{aligned} \mathbb{E}[X_T] &= \binom{n}{d+1} (n-d-1)_s (\gamma n^{-1/d})^{ds+1} \\ &= (1+o(1)) \binom{n}{d+1} n^{-1/d} \gamma^{ds+1}, \end{aligned} \quad (5.1)$$

where the transition to the second line follows from $(n-d-1)_s n^{-s} \rightarrow 1$.

We proceed to compute the variance of X_T . Let T_1, T_2 be proper s -pedigrees whose underlying tree is T , let q denote the total number of common labels (internal and root) in T_1 and T_2 , and write $P := \mathcal{L}(T_1) \cap \mathcal{L}(T_2)$ and $t = |P|$.

Clearly, $q = 0$ implies $t = 0$. In addition, we claim that if $0 < q \leq s + d + 1$ then

$$t \leq dq - 1.$$

Indeed, we first observe that if $t = ds + 1$ then T_1 and T_2 have the same label set, thereby $q = s + d + 1$ and the inequality is satisfied. Otherwise, $P \subsetneq \mathcal{L}(T_1)$. Let R denote the internal labels of T_1 that are contained in one of the sets in P . Note that the labels in R must also appear in T_2 — either as internal labels or in the root. Furthermore, by Lemma 4.3, $|P| \leq d|R|$. If this inequality is strict, then the claim follows from $|R| \leq q$. Otherwise, Lemma 4.3 asserts that $P = \mathcal{L}(T_1) \setminus \mathcal{L}((T_1)_{\mathbf{v}})$ for some $\mathbf{v} \in T_1$. In such a case, besides the t/d common labels of R , all the $d+1$

labels in the root of T_1 also appear in one of the sets of P — and therefore also in T_2 . Consequently, an even stronger inequality of $q \geq t/d + d + 1$ is obtained.

Therefore, by summing over the integers $0 \leq q \leq d + s + 1$, we find

$$\mathbb{E}[X_T^2] \leq \mathbb{E}[X_T]^2 + \sum_{q=1}^{s+d+1} C n^{2s+2d+2-q} (\gamma n^{-1/d})^{2ds+2-(dq-1)},$$

where the constant $C = C_{s,d}$ accounts for assigning the selected $2s + 2d + 2 - q$ labels to the nodes of T that creates T_1, T_2 . Therefore $\text{Var}(X_T) = O(n^{2d+2-3/d})$, and the lemma is derived via Chebyshev's inequality. \blacksquare

Lemma 5.2. *Let $s_0 \geq 0$ be an integer, and let Z_{s_0} denote the number of pairs of proper pedigrees with at most s_0 internal vertices that have the same root and whose leaves are contained in Y . Then,*

$$\mathbb{E}[Z_{s_0}] \leq C n^{d+1-2/d},$$

where C is a constant depending on d, s_0 .

Proof. Let T_1, T_2 be distinct proper pedigrees sharing the same root \mathbf{o} . Denote by R_i the set of internal labels of T_i , $i = 1, 2$, and let $s = |R_1 \cup R_2|$. We claim that

$$|\mathcal{L}(T_1) \cup \mathcal{L}(T_2)| \geq ds + 2. \quad (5.2)$$

Indeed, since $|\mathcal{L}(T_i)| = d|R_i| + 1$, (5.2) reformulates to

$$|\mathcal{L}(T_1) \cap \mathcal{L}(T_2)| \leq d|R_1 \cap R_2|,$$

which is derived by Lemma 4.3. We conclude the lemma by the following crude bound.

$$\begin{aligned} \mathbb{E}[Z_{s_0}] &\leq \binom{n}{d+1} |\mathfrak{T}_{\leq s_0}|^2 \sum_{s=1}^{2s_0} \binom{n-d-1}{s} s^{2s_0} (\gamma n^{-1/d})^{ds+2} \\ &\leq C n^{d+1-2/d}. \end{aligned}$$

Here we chose a pair T_1, T_2 by selecting the common root, the two underlying $(d+1)$ -ary trees and the label set $R_1 \cup R_2$, and afterwards assigning the outgoing labels to the (at most $2s_0$) internal vertices. Finally, we use (5.2) to get a lower bound for the number of leaves that are required to be in Y . \blacksquare

Proof of Theorem 2. Let $\varepsilon > 0$, and consider a sufficiently large integer s_0 satisfying

$$\sum_{s=0}^{s_0} C_d(s) \gamma^{ds+1} > \hat{\gamma} - \varepsilon.$$

We start with the lower bound. Note that

$$\binom{n}{d+1} \hat{X} \geq \sum_{T \in \mathfrak{T}_{\leq s_0}} X_T - Z_{s_0}. \quad (5.3)$$

Indeed, the summation on the RHS counts the number of special pedigrees (i.e., proper pedigrees with at most s_0 internal vertices) whose leaves are all members of Y , from which we subtract the number of pairs of such pedigrees with the same root. Therefore, every root of such a pedigree is accounted for at most once.

Since $|\mathfrak{T}_{\leq s_0}|$ does not depend on n , by Lemma 5.1 we have that

$$\begin{aligned} n^{1/d} \frac{1}{\binom{n}{d+1}} \sum_{T \in \mathfrak{T}_{\leq s_0}} X_T &= \sum_{s=0}^{s_0} \sum_{T \in \mathfrak{T}_s} n^{1/d} \frac{X_T}{\binom{n}{d+1}} \rightarrow \sum_{s=0}^{s_0} \sum_{T \in \mathfrak{T}_s} \gamma^{ds+1} \\ &= \sum_{s=0}^{s_0} C_d(s) \gamma^{ds+1} > \hat{\gamma} - \varepsilon, \end{aligned} \quad (5.4)$$

where the convergence is in probability as $n \rightarrow \infty$. In addition, by Lemma 5.2, $\mathbb{E}[n^{1/d} Z_{s_0} / \binom{n}{d+1}] \rightarrow 0$ as $n \rightarrow \infty$, and we conclude the lower bound using (5.3).

To establish a matching upper bound, we bound the number of sets $\mathbf{o} \in \binom{[n]}{d+1}$ that have a stacked contraction in Y by

$$\binom{n}{d+1} \hat{X} \leq \sum_{T \in \mathfrak{T}_{n-d-1}} X_T + W, \quad (5.5)$$

where W is the number of *improper* (m, l, s) -pedigrees (i.e., $\mathbf{a} := l - ds - 1 \geq 1$) whose leaves are all members in Y .

We split the first summation by considering big and small trees in separate. Namely, using (5.1) we find that

$$\begin{aligned} \mathbb{E} \left[\sum_{s=s_0+1}^{n-d-1} \sum_{T \in \mathfrak{T}_s} X_T \right] &< \binom{n}{d+1} n^{-1/d} \sum_{s>s_0} C_d(s) \gamma^{ds+1} \\ &< \binom{n}{d+1} n^{-1/d} \varepsilon, \end{aligned}$$

whence

$$\mathbb{P} \left(n^{1/d} \frac{1}{\binom{n}{d+1}} \sum_{s=s_0+1}^{n-d-1} \sum_{T \in \mathfrak{T}_s} X_T > \sqrt{\varepsilon} \right) \leq \sqrt{\varepsilon}. \quad (5.6)$$

In addition, since $\sum_{s \leq s_0} C_d(s) \gamma^{ds+1} < \hat{\gamma}$, we find by (5.4) that

$$\mathbb{P} \left(n^{1/d} \frac{1}{\binom{n}{d+1}} \sum_{s \leq s_0} \sum_{T \in \mathfrak{T}_s} X_T > \hat{\gamma} \right) \rightarrow 0, \quad (5.7)$$

as $n \rightarrow \infty$.

To bound the contribution of W , we recall several facts from Section 3. Let \mathcal{B} denote the event that there is no (m, l, s) -pedigree with $s > (\log n)^{5/2}$ distinct internal labels whose leaves are all members of Y . By Observation 3.2 and (3.1), \mathcal{B} occurs with probability $1 - o(1)$.

In addition, recall that (as stated in the last paragraph of the proof of Theorem 3.1) for every fixed $\mathbf{o} \in \binom{[n]}{d+1}$, $s \leq (\log n)^{5/2}$ and $\mathbf{a} \geq 1$, the probability that there exists an (l, m, s) -pedigree, such that $l = ds + 1 + \mathbf{a}$, whose root is \mathbf{o} and whose leaves are all members in Y is at most $e^{-\delta s} n^{-\frac{\mathbf{a}+1}{d} + o(1)}$, where $\delta = 1 - \gamma \alpha_d^{1/d}$. Therefore, by summing over \mathbf{o}, s and $\mathbf{a} \geq 1$ we find that

$$\mathbb{E}[W \mathbf{1}_{\mathcal{B}}] \leq \binom{n}{d+1} n^{-2/d+o(1)},$$

whence

$$\mathbb{P} \left(n^{1/d} \frac{1}{\binom{n}{d+1}} W > \sqrt{\varepsilon} \right) \leq n^{-1/d+o(1)} / \sqrt{\varepsilon} + (1 - \mathbb{P}(\mathcal{B})) \rightarrow 0, \quad (5.8)$$

as $n \rightarrow \infty$. In summary, and using (5.5),(5.6),(5.7) and (5.8), we find that as $n \rightarrow \infty$,

$$\mathbb{P}(n^{1/d}\hat{X} > \hat{\gamma} + 2\sqrt{\varepsilon}) \leq \sqrt{\varepsilon} + o(1),$$

as claimed. ■

APPENDIX A. THE LEAF-TO-LABEL DIFFERENTIAL IN THE DAGS

Proposition A.1. *Let $d \geq 2$, and suppose T is a stacked contraction with l faces and a total of $s + d + 1$ distinct labels. Then*

$$l \geq ds + 1. \tag{A.1}$$

Moreover, there is an explicit $A \in \mathbb{R}^{d(s+d+1) \times l}$ such that $\dim(\ker A) = l - (ds + 1)$.

Proof. Suppose without loss of generality that the label set is $\{1, \dots, s + d + 1\}$, and let $v_1, \dots, v_{s+d+1} \in \mathbb{R}^d$ be in general position. Let $x_1 < x_2 < \dots < x_{d+1}$; we represent a face $f = \{x_1, \dots, x_{d+1}\}$ via a vector $w_f \in \mathbb{R}^{d(s+d+1)}$ as follows. Define the matrix

$$A_f := \left(\begin{array}{c|ccc|c} v_{x_1} & \cdots & v_{x_{d+1}} & \\ \hline 1 & & & 1 \end{array} \right), \tag{A.2}$$

whose determinant is — up to the sign — the volume of the simplex whose vertices are $v_{x_1}, \dots, v_{x_{d+1}}$. The cofactor $C_{i,j}(A_f)$ of the matrix A_f , for every $1 \leq j \leq d + 1$ and $1 \leq i \leq d$, is the partial derivative of $\det(A_f)$ with respect to the i -th coordinate of v_{x_j} . Representing the vector $w_f \in \mathbb{R}^{d(s+d+1)}$ as an $(s + d + 1) \times d$ matrix, we define w_f to have

$$(w_f)_{x_j, i} := C_{i,j}(A_f) \quad 1 \leq j \leq d + 1, 1 \leq i \leq d, \tag{A.3}$$

and 0 elsewhere. To a set $K \subset \binom{[s+d+1]}{d+1}$ we associate the matrix A_K whose columns are w_f , $f \in K$.

Let $z \in \mathbb{R}^{d(s+d+1)}$ be a vector that is also represented as an $(s + d + 1) \times d$ matrix whose columns are z_1, \dots, z_{s+d+1} , and consider a continuous motion of the vertices

$$v_j^{(t)} = v_j + t \cdot z_j, \quad j = 1, \dots, s + d + 1.$$

Let $A_f^{(t)}$ be defined as (A.2) with respect to $v_{x_j}^{(t)}$, $j = 1, \dots, d + 1$. Then, the derivative of the determinant of $A_f^{(t)}$ at time $t = 0$ equals the inner product of z and w_f . In words, z is orthogonal to w_f if and only if the above motion infinitesimally preserves the volume of the simplex whose vertices start at $v_{x_1}, \dots, v_{x_{d+1}}$. Clearly, any translation of \mathbb{R}^d , where $z_j = u$ for all $j = 1, \dots, s + d + 1$, preserves the volume of every simplex. Additionally, we claim that for every matrix $M \in \mathbb{R}^{d \times d}$ having $\text{tr}(M) = 0$, setting $z_j = Mv_j$, $j = 1, \dots, s + d + 1$ infinitesimally preserves the volume of every simplex f . Indeed, in such a case $v_j^{(t)} = (I + tM)v_j$ whence the volume of the simplex f at time t satisfies

$$\det(A_f^{(t)}) = \det(I + tM) \det(A_f).$$

We conclude the claim by observing that the derivative of $\det(I + tM)$ equals $\text{tr}(M) = 0$. In summary, we showed that the left kernel of A_K is of dimension at least $d + d^2 - 1$, so its rank is at most $d(s + d + 1) - (d + d^2 - 1) = ds + 1$. In the remainder of the appendix, we show that this bound is attained by the faces of every stacked contraction.

Claim A.2. Let $x_1 < \dots < x_{d+1}$ and z be such that $x_{k-1} < z < x_k$ for some $1 \leq k \leq d+2$. If $f = (x_1, \dots, x_{d+1})$ and f_i ($i = 1, \dots, d+1$) is the face obtained from it by replacing x_i with z , then for every $v_{x_1}, \dots, v_{x_{d+1}}, v_z \in \mathbb{R}^d$,

$$w_f = \sum_{r=1}^{k-1} (-1)^{k+r+1} w_{f_r} + \sum_{r=k}^{d+1} (-1)^{k+r} w_{f_r} \quad (\text{A.4})$$

Proof. Let $1 \leq i, j \leq d+1$, and consider the matrix

$$E_i = \left(\begin{array}{c|c|c|c|c|c|c|c} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ v_z^{(\neg j)} & v_{x_1}^{(\neg j)} & \dots & v_{x_{i-1}}^{(\neg j)} & v_{x_{i+1}}^{(\neg j)} & \dots & v_{x_{d+1}}^{(\neg j)} & \\ \hline 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{array} \right),$$

where $v^{(\neg j)}$ represents the vector obtained from v by omitting its j -th coordinate. Observe that the submatrix of E_i obtained by omitting the first row and the column containing v_z is precisely the submatrix obtained by omitting from A_f the j -th row and the column containing v_{x_i} . Furthermore, omitting the first row and the column of v_{x_r} from E_i (for $1 \leq r \leq d+1$, $r \neq i$) gives the submatrix of A_{f_r} obtained by omitting its j -th row and the column of v_{x_i} , followed by moving the column of v_z to be the first one (retaining the inner ordering between the remaining columns). Observe that the order of the columns in A_{f_r} is such that:

- (a) the vector v_{x_i} appears in column $i - \mathbf{1}_{\{r < i\}} + \mathbf{1}_{\{k \leq i\}}$;
- (b) if we omit the column of v_{x_i} , then v_z appears in column $k - \mathbf{1}_{\{r < k\}} - \mathbf{1}_{\{i < k\}}$.

It then follows from the observations above and (A.3) that, if $i \geq k$, then

$$\sum_{r=1}^{k-1} (-1)^r C_{1,r+1}(E_i) = \sum_{r < k} (-1)^{r+i+j} (w_{f_r})_{x_i, j}$$

(for each r the corresponding cofactor of A_{f_r} contributes $(-1)^{i+j}$), and

$$\begin{aligned} \sum_{r=k}^d (-1)^{r+1} C_{1,r+1}(E_i) &= \sum_{r=k}^{i-1} (-1)^{r+i+j} (w_{f_r})_{x_i, j} + \sum_{r=i}^d (-1)^{r+(i+1)+j} (w_{f_{r+1}})_{x_i, j} \\ &= \sum_{\substack{r \geq k \\ r \neq i}} (-1)^{r+i+j} (w_{f_r})_{x_i, j}. \end{aligned}$$

Combining these, along with the fact that $(w_{f_i})_{x_i, j} = 0$ for all j (since f_i does not contain the label x_i , having replaced it by z), we find that for $i \geq k$,

$$\sum_{r=1}^d (-1)^r C_{1,r+1}(E_i) = \sum_{r < k} (-1)^{r+i+j} (w_{f_r})_{x_i, j} - \sum_{r \geq k} (-1)^{r+i+j} (w_{f_r})_{x_i, j}. \quad (\text{A.5})$$

On the other hand, if $i < k$ then

$$\sum_{r=k-1}^d (-1)^r C_{1,r+1}(E_i) = \sum_{r=k}^{d+1} (-1)^{(r+1)+i+j} (w_{f_r})_{x_i, j},$$

whereas

$$\begin{aligned} \sum_{r=1}^{k-2} (-1)^{r+1} C_{1,r+1}(E_i) &= \sum_{r=1}^{i-1} (-1)^{r+(i-1)+j} (w_{f_r})_{x_i,j} + \sum_{r=i}^{k-2} (-1)^{r+i+j} (w_{f_{r+1}})_{x_i,j} \\ &= \sum_{r=1}^{k-1} (-1)^{r+i+j+1} (w_{f_r})_{x_i,j}, \end{aligned}$$

and we see that (A.5) also holds for $i < k$ (again using that $(w_{f_i})_{x_i,j} = 0$ for all j).

Recalling that $C_{1,1}(E_i) = (-1)^{i+j} (w_f)_{x_i,j}$, we may now, for every $i = 1, \dots, d+1$, substitute (A.5) in the expansion $\sum_r C_{1,r}(E_i) = \det(E_i) = 0$, and obtain that

$$(w_f)_{x_i,j} = \sum_{r < k} (-1)^{k+r+1} (w_{f_r})_{x_i,j} + \sum_{r \geq k} (-1)^{k+r} (w_{f_r})_{x_i,j}.$$

It remains to verify (A.4) at the coordinate z, j . To this end, note first that $(w_f)_{z,j} = 0$ for all j (as the face f does not contain the label z), and let

$$E = \left(\begin{array}{c|ccc|c} 1 & & & & 1 \\ v_{x_1}^{(\neg j)} & & \dots & & v_{x_{d+1}}^{(\neg j)} \\ & & & & \\ \hline & & & & 1 \end{array} \right).$$

For every $r = 1, \dots, d+1$, the $(1, r)$ -minor of E is exactly the minor corresponding to omitting the j -th row and the column containing v_z from A_{f_r} . Therefore,

- for $r = 1, \dots, k-1$ (the vector v_z appears in column $k-1$ of A_{f_r}) we have

$$C_{1,r}(E) = (-1)^{r+k+j} (w_{f_r})_{z,j};$$

- for $r = k, \dots, d+1$ (the vector v_z appears in column $k-1$ of A_{f_r}) we have

$$C_{1,r}(E) = (-1)^{r+1+k+j} (w_{f_r})_{z,j}.$$

Altogether, recalling definition (A.3) and that $\sum_r C_{1,r}(E) = \det(E) = 0$ shows that

$$\sum_{r < k} (-1)^{k+r} (w_{f_r})_{z,j} + \sum_{r \geq k} (-1)^{k+r+1} (w_{f_r})_{z,j} = 0,$$

at which point the fact that $(w_f)_{z,j} = 0$ for all j establishes (A.4) at coordinate (z, j) and thereby concludes the proof. \blacksquare

Claim A.3. *Let $x_1 < \dots < x_{d+1}$ and z be such that $z \neq x_i$ for all i . Let f_i ($i = 1, \dots, d+1$) be the face with labels $\{x_1, \dots, x_{d+1}, z\} \setminus \{x_i\}$. If \mathcal{F} is any collections of faces that do not contain the label z , and \tilde{w} is any non-zero vector in the span of $\{w_f : f \in \mathcal{F}\}$, then $\tilde{w}, w_{f_1}, \dots, w_{f_d}$ are linearly independent.*

Proof. Denote by $1 \leq k \leq d+2$ the integer such that $x_{k-1} < z < x_s$, and suppose that for some $a_0, \dots, a_d \in \mathbb{R}$ we have

$$a_0 \tilde{w} + \sum_{i=1}^d a_i w_{f_i} = 0.$$

Since \tilde{w} is a combination of $\{w_f : f \in \mathcal{F}\}$ where \mathcal{F} does not feature the label z , every $f \in \mathcal{F}$ has $(w_f)_{z,j} = 0$ for all j , from which we infer that

$$\sum_{i=1}^d a_i (w_{f_i})_{z,j} = 0 \quad \text{for all } j = 1, \dots, d. \quad (\text{A.6})$$

Define

$$E = \begin{pmatrix} v_{x_1}^\top & 1 \\ \vdots & \vdots \\ v_{x_{d+1}}^\top & 1 \end{pmatrix} \quad \text{and} \quad \hat{v} = \begin{pmatrix} \hat{a}_1 \\ \vdots \\ \hat{a}_d \\ 0 \end{pmatrix}, \quad (\text{A.7})$$

where

$$\hat{a}_i := (-1)^{i+k+1_{\{i < k\}}} a_i \quad (i = 1, \dots, d), \quad (\text{A.8})$$

and let E_j be the matrix obtained from E by replacing its j -th column by \hat{v} .

With (A.2) in mind, observe that for every $1 \leq i, j \leq d$, omitting the i -th row and j -th column of E_j gives the transpose of the submatrix obtained by omitting the column of v_z and the j -th row in A_{f_i} . Thus, by (A.3) and the fact that v_z appears in column $k - 1_{\{i < k\}}$ in A_{f_i} , the (i, j) -minor of E_j is $(-1)^{j+k-1_{\{i < k\}}}(w_{f_i})_{z,j}$. We conclude that for every $j = 1, \dots, d$,

$$\det(E_j) = \sum_{i=1}^d \hat{a}_i (-1)^{i+k-1_{\{i < k\}}}(w_{f_i})_{z,j} = \sum_{i=1}^d a_i (w_{f_i})_{z,j} = 0,$$

with the last equality due to (A.6). At the same time, we have that $u = (u_j)_{j=1}^{d+1}$ given by

$$u_j := \frac{\det(E_j)}{\det(E)}$$

is the unique solution of $Eu = \hat{v}$ by Cramer's rule, relying here on the hypothesis that v_1, \dots, v_k are in general position in \mathbb{R}^d , whence $\det(E) \neq 0$. We have shown above that $u_j = 0$ for all $1 \leq j \leq d$, and so $u = u_{d+1}e_{d+1}$, where e_{d+1} is the $(d+1)$ -th standard basis vector in \mathbb{R}^d . But $Ee_{d+1} = \mathbf{1}$, the all-1 vector, whence

$$\hat{v} = Eu = u_{d+1}\mathbf{1}.$$

From this and the fact that $\hat{v}_{d+1} = 0$ we can then deduce that $\hat{v} = 0$, and in turn, that $a_i = 0$ for all $i = 1, \dots, d$. This reduces the original hypothesis to having $a_0 \tilde{w} = 0$, whence $a_0 = 0$ as well, and the proof is concluded. \blacksquare

We claim that if the oriented faces g_1, \dots, g_l of the triangulation T form a stacked triangulation of f_0 then the $d(s+d+1) \times l$ matrix A whose columns are $\{w_{g_i}\}_{i=1, \dots, l}$ has rank at least $ds+1$, which will imply (A.1).

This will be proved by induction on the number of internal vertices m in the stacked triangulation T . The base case of $m = 0$ internal vertices (where T is the single face f_0) corresponds to $s = 0$ and $l = 1 = ds+1$.

Every stacked triangulation with $m \geq 1$ faces is obtained by subdividing some face f of a stacked triangulation \tilde{T} with $m-1$ internal vertices. If $f = (x_1, \dots, x_{d+1})$, this corresponds to replacing the face f by the $d+1$ faces f_i ($i = 1, \dots, d+1$) as in where f_i has x_i swapped by z .

Claim A.2 immediately implies that $\text{rank}(A) \geq \text{rank}(\tilde{A})$, where \tilde{A} is the matrix corresponding to \tilde{T} , the triangulation before subdividing f via the label z . In

particular, if the label z has already appeared in \tilde{T} , then $\text{rank}(A) \geq \text{rank}(\tilde{A}) \geq ds+1$ holds true by our induction.

On the other hand, if z is a new label, then $\text{rank}(A) \geq \text{rank}(\tilde{A}) + d$ by virtue of Claim A.3. \blacksquare

Acknowledgment. The authors thank Eran Nevo for useful discussions. E.L. was supported in part by NSF grants DMS-1812095 and DMS-2054833.

REFERENCES

- [1] M. Aizenman and J. L. Lebowitz. Metastability effects in bootstrap percolation. *J. Phys. A*, 21(19):3801–3813, 1988.
- [2] N. Alon. An extremal problem for sets with applications to graph theory. *J. Combin. Theory Ser. A*, 40(1):82–89, 1985.
- [3] O. Angel and B. Kolesnik. Sharp thresholds for contagious sets in random graphs. *Ann. Appl. Probab.*, 28(2):1052–1098, 2018.
- [4] L. Aronshtam and N. Linial. When does the top homology of a random simplicial complex vanish? *Random Structures Algorithms*, 46(1):26–35, 2015.
- [5] L. Aronshtam, N. Linial, T. Łuczak, and R. Meshulam. Collapsibility and vanishing of top homology in random simplicial complexes. *Discrete Comput. Geom.*, 49(2):317–334, 2013.
- [6] E. Babson, C. Hoffman, and M. Kahle. The fundamental group of random 2-complexes. *J. Amer. Math. Soc.*, 24(1):1–28, 2011.
- [7] J. Balogh, B. Bollobás, H. Duminil-Copin, and R. Morris. The sharp threshold for bootstrap percolation in all dimensions. *Trans. Amer. Math. Soc.*, 364(5):2667–2701, 2012.
- [8] J. Balogh, B. Bollobás, and R. Morris. Graph bootstrap percolation. *Random Structures Algorithms*, 41(4):413–440, 2012.
- [9] J. Balogh, B. Bollobás, R. Morris, and O. Riordan. Linear algebra and bootstrap percolation. *J. Combin. Theory Ser. A*, 119(6):1328–1335, 2012.
- [10] Z. Bartha and B. Kolesnik. Weakly saturated random graphs. 2020. Preprint, [arXiv:2007.14716](https://arxiv.org/abs/2007.14716).
- [11] E. Bayraktar and S. Chakraborty. $K_{r,s}$ graph bootstrap percolation. 2019. Preprint, [arXiv:1904.12764](https://arxiv.org/abs/1904.12764).
- [12] A. Björner and G. Kalai. An extended Euler-Poincaré theorem. *Acta Math.*, 161(3-4):279–303, 1988.
- [13] B. Bollobás. Weakly k -saturated graphs. In *Beiträge zur Graphentheorie (Kolloquium, Manebach, 1967)*, pages 25–31. Teubner, Leipzig, 1968.
- [14] J. Chalupa, P. L. Leath, and G. R. Reich. Bootstrap percolation on a Bethe lattice. *J. Phys. C: Solid State Phys.*, 12(1):L31, 1979.
- [15] M. M. Cohen. *A course in simple-homotopy theory*. Graduate Texts in Mathematics, Vol. 10. Springer-Verlag, New York-Berlin, 1973.
- [16] A. Costa and M. Farber. Homological domination in large random simplicial complexes. 2015. Preprint, [arXiv:1503.03253](https://arxiv.org/abs/1503.03253).
- [17] P. Frankl. An extremal problem for two families of sets. *European J. Combin.*, 3(2):125–127, 1982.
- [18] K. Gunderson, S. Koch, and M. Przykucki. The time of graph bootstrap percolation. *Random Structures Algorithms*, 51(1):143–168, 2017.
- [19] M. Kahle, E. Paquette, and E. Roldán. Topology of random 2-dimensional cubical complexes. *Forum Math. Sigma*, 9:Paper No. e76, 24, 2021.
- [20] G. Kalai. Characterization of f -vectors of families of convex sets in \mathbf{R}^d . I. Necessity of Eckhoff’s conditions. *Israel J. Math.*, 48(2-3):175–195, 1984.
- [21] G. Kalai. Weakly saturated graphs are rigid. In *Convexity and graph theory (Jerusalem, 1981)*, volume 87 of *North-Holland Math. Stud.*, pages 189–190. North-Holland, Amsterdam, 1984.
- [22] G. Kalai. Characterization of f -vectors of families of convex sets in \mathbf{R}^d . II. Sufficiency of Eckhoff’s conditions. *J. Combin. Theory Ser. A*, 41(2):167–188, 1986.
- [23] G. Kalai. Algebraic shifting. In *Computational commutative algebra and combinatorics (Osaka, 1999)*, volume 33 of *Adv. Stud. Pure Math.*, pages 121–163. Math. Soc. Japan, Tokyo, 2002.

- [24] B. Kolesnik. Sharp threshold for K_4 -percolation. 2017. Preprint, [arXiv:1705.08882](#).
- [25] D. Korándi and B. Sudakov. Saturation in random graphs. *Random Structures Algorithms*, 51(1):169–181, 2017.
- [26] N. Linial and R. Meshulam. Homological connectivity of random 2-complexes. *Combinatorica*, 26(4):475–487, 2006.
- [27] N. Linial and Y. Peled. On the phase transition in random simplicial complexes. *Ann. of Math. (2)*, 184(3):745–773, 2016.
- [28] N. Linial and Y. Peled. Random simplicial complexes: around the phase transition. pages 543–570, 2017.
- [29] L. Lovász. *Combinatorial problems and exercises*. AMS Chelsea Publishing, Providence, RI, second edition, 2007.
- [30] Z. Luria and Y. Peled. On simple connectivity of random 2-complexes. *Discrete Comput. Geom.*, 2021.
- [31] G. Malen. Collapsibility of random clique complexes. 2019. Preprint, [arXiv:1903.05055](#).
- [32] R. Meshulam and N. Wallach. Homological connectivity of random k -dimensional complexes. *Random Structures Algorithms*, 34(3):408–417, 2009.
- [33] N. Morrison and J. A. Noel. A sharp threshold for bootstrap percolation in a random hypergraph. *Electron. J. Probab.*, 26:Paper No. 97, 85, 2021.
- [34] E. Nevo. Algebraic shifting and basic constructions on simplicial complexes. *J. Algebraic Combin.*, 22(4):411–433, 2005.
- [35] A. Newman and E. Paquette. The integer homology threshold in $Y_d(n, p)$. 2018. Preprint, [arXiv:1808.10647](#).
- [36] J. H. C. Whitehead. Simplicial Spaces, Nuclei and m -Groups. *Proc. London Math. Soc. (2)*, 45(4):243–327, 1939.

E. LUBETZKY

COURANT INSTITUTE, NEW YORK UNIVERSITY, 251 MERCER STREET, NEW YORK, NY 10012, USA.

Email address: eyal@courant.nyu.edu

Y. PELED

EINSTEIN INSTITUTE OF MATHEMATICS, HEBREW UNIVERSITY, JERUSALEM 91904, ISRAEL.

Email address: yuval.peled@mail.huji.ac.il