CRITICALITY: FROM RANDOM GRAPHS TO SIMPLICIAL COMPLEXES

ZOYA BRAHIMZADEH

Abstract. Random graphs are graphs that are generated according to some specified probability distribution. They have a particularly fascinating feature: their phase transitions. Interestingly, despite their canonical randomness, they have many critical points that reliably and asymptotically predict their properties as they evolve. Of recent interest have been attempts to create higher-dimensional analogues of such random graphs while preserving important properties and phase transitions. Most popular of these attempts is the Linial-Meshulam model for random simplicial complexes. This paper reviews some of the most important properties that emerge from random graphs of the Erdős–Rényi model and draws parallels to analogous properties of algebraic topology within the Linial-Meshulam model. Particularly, we explore topological analogues of the exciting phase transitions in the Erdos Renyi model around two exciting thresholds that contain the vanishing of acyclicity, vanishing of collapsibility, emergence of a giant component, and emergence of full connectivity. In higher dimensions, the nature of parallel properties—homology groups, cycles, and collapses— and their corresponding thresholds have nuanced differences.

1. Introduction

Random graphs play a dual role for modeling both chaos and order. As a model of randomness, they are frequently used as null models; to discern which network structures are inherent to how networks are formed and which structures are truly non-random patterns. An easy example is social networks—are the observed patterns due to intrinsic behavior of graphs and noise in the dataset, or is there some meaningful idea at play? On the other hand, these graphs can also be used to model order—when phenomena are too complex, a random graph can sometimes be a great tool to generate networks of similar complexity and structure. One can, again, imagine a highly dense and convoluted social network where a randomly generated graph could yield a similar structure. Using random graphs as such models allows us to easily compute certain properties.

These random graphs can evolve: given a number of nodes n, one can iteratively adding edges, modeling the growth of the network. As they evolve, certain *phase transitions* occur. In physical systems, phase transitions are sudden changes in some physical properties of a system. The same definition can be applied to objects of mathematics—a phase transition in a random graph is the sudden change in its properties. Examples of potential properties can be: acyclicity, connectivity, and collapsibility. However, the story of phase transitions doesn't stop with random graphs—we can expand our horizons and translate to higher dimensions with the notion of random simplicial complexes. Scaling from random graphs to higher dimensions yields much greater power when using random objects for practical purposes—they can model more complex phenomena; for example, patterns within structural organization of the brain in response to stimulus [RNS^{+}17]. The area of magic

Date: July 20, 2024.

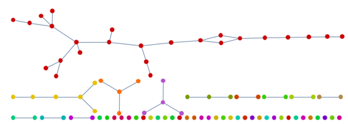


Figure 1. An example of a graph within $G(n,p)$ model with $n = 90$ and $p = \frac{1}{89}$.

and mystique comes when we attempt to try to make this translation to higher dimensions as smooth as possible. In higher dimensions, things are different. Sometimes, there is no obvious analog for a graph property—such as connected components—and sometimes the translation is very natural—such as with the notion of cycles. It's fascinating to see where we can make an easy translation and where properties of random simplicial complexes—and their transitions likewise—come with different characteristics than their graph theory counterparts. The motivation of this paper is with this spirit: to compare phase transitions of properties within random graphs and random simplicial complexes, finding where their properties and the thresholds for property emergence diverge. This is an incredibly important piece of information to hold as we try to use random objects to both model and compare real world phenomena. Comparing random simplicial complexes—a very abstract concept—to random graphs—more concretely defined and studied— allows us to give some grounding to the meaning of all these abstract topological properties. Thus, phase transitions are an incredibly important piece of the puzzle to characterizing these otherworldly, chaotic, random, unpredictable objects; they predict properties of randomness. These peculiar phase transitions happen at a specific threshold and accurately predict properties of complexes. We usually express these transitions as occurring with a suffix "w.h.p"—with high probability—or a.a.s—asymptotically almost surely. When looking at transitions, we are typically looking at the behavior of a complex or graph asymptotically—with respect to the number of nodes n as $n \to \infty$. Occasionally, a.s—almost surely—is also used. The paper will continue as follows: first, some preliminary definitions are provided; then, we go on an in-depth exploration of phase transitions within random graphs; finally, we finish by exploring phase transitions within random simplicial complexes and drawing connections between the two objects.

2. Preliminaries

2.1. **Erdős–Rényi Model.** Let $G(n,p)$ represent the class of Erdős-Rényi random undirected graphs where n is the number of nodes and p is the independent probability that two vertexes will have an edge between them. An example of a random graph is pictured in Figure 1.

2.2. **N-dimensional equivalents.** As the curious beings we are, we always want to expand further—to push the limits of our objects. A similar thing happens with this study of random graphs: is there an d-dimensional equivalent for a graph? The answer comes from algebraic topology: simplicial complexes.

2.2.1. *Simplicial Complexes.* A n-simplex, σ_n, is an n-dimensional simple shape generated by $n+1$ points, where one of the points is the origin. A 0-simplex is a point, a 1-simplex is a line, a 2-simplex is a triangle, and a 3-simplex is a tetrahedron; one can continue to list higher dimensional simple shapes. A **simplicial complex**, K, is then simply a set of n-simplicies

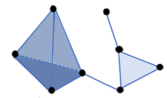


Figure 2. Simplicial complex with 8 0-faces, 11 1-faces, and 4 2-faces.

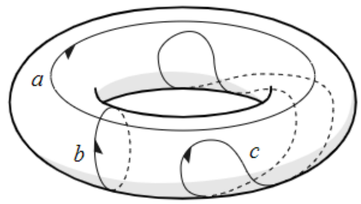


Figure 3. Image of a torus, a 2-dimensional topological object, and some of its cycles labeled a, b. and c.

and their skeletons: $K = \{\sigma_1, \sigma_2, \sigma_3, \ldots, \sigma_n\}$. An example of a simplicial complex is pictured in Figure 2. A p-skeleton of a simplicial complex K is $K^{(p)}$, the set of all p-simplicies $\sigma_p \in K$ in dimension $n \leq p$.

2.2.2. *Homology.* A topological *hole* is a discontinuity in space, it prevents a topological object from being deformed to a point. The most common example is a torus which has 1 0-dimensional hole, 2 1-dimensional holes, and 1 2-dimensional hole. These holes all prevent the Torus from being compressed in different dimensions. Figure 3 shows an example of holes in a torus: the two 1-dimensional holes are A and B, the one 2-dimensional hole is labeled C.

Let $H_k(X, G)$ denote the kth homology group of space X with coefficents in G. The *kth* homology group stores information regarding nontrivial holes within dimension k. The k-th Betti number, $\beta_k = \dim H_k(X, G)$ "counts" the number of holes of such dimension. These holes are identified by finding topological cycles of dimension k that are not part of a lower dimensional simplex ("bounded") in the simplicial complex.

2.2.3. *Linial-Meshulam Model.* With this background in mind, let us now introduce the Linial-Meshulam model, one of the most well-studied models of random simplicial complexes. The beauty of this model comes from the incredibly intuitive way that it expands random graphs into higher dimensional analogues. $Y_d(n, p)$ is d-dimensional Linial-Meshulam random simplicial complex n vertices (0-faces), a full (k-1) skeleton, and a probability p that each d-dimensional simplex is drawn in. Evidently, $Y_1(n, p)$ is analogous to the Erdős–Rényi random graph, where $1-1 = 0$ dimensional simplicies (vertices) are connected by 1-simplices (edges). An example of a random 2-complex is pictured in Figure 4

2.3. **Asymptotic and approximate notation.** Now let's side-step to establish some preliminaries important when analyzing the evolution of random objects. Because all of our results consider the asymptotic nature of models: as $n \to \infty$, its important to use the right notation to accurately capture the energy of our results.

(1) We use big-O notation to describe that a function f is growing proportionally to function g. Formally, $f(x) = O(g(x)) : |f(x)| \leq K|g(x)|$ for some constant $k > 0$ and all $x \in \mathbb{R}$.

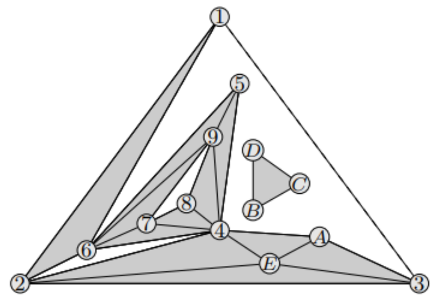


Figure 4. A planar embedding of a random 2-complex. The vertices are labeled using hexadecimal notation for convenience. There is a full $(d-1) = 1$-skeleton of lines to encapsulate all the $d = 2$-faces. The shaded triangles are the 2-faces with probability p of being drawn and are described as $\{AE3, DCB, \ldots\}$.

(2) We use small-o notation to indicate that one function f grows at a significantly slower rate than another function g. This upper bound on the growth of f is not tight. Formally, $f(x) = o(g(x))$ as $x \to a : \lim_{x\to a} f(x)/g(x) = 0$.
(3) We use small-omega notation to indicate that one function f grows at a strictly faster rate than another function g. Formally, $f(x) = \omega(g(x))$ as $x \to a$: $\lim_{x\to a} f(x)/g(x) = \infty$.
(4) We use $\ll$ to denote infinite asymptotic behavior with respect to n. Formally, $A \ll B : A/B \to 0$ as $n \to \infty$.
(5) We use $\gg$ to denote infinite asymptotic behavior with respect to n. Formally, $A \gg B : A/B \to \infty$ as $n \to \infty$.
(6) We use $\approx$ to denote limiting behavior with respect to some parameter. Formally, $A \approx B : A/B \to 1$ as some parameter converges to 0 or ∞ or some other limit.
(7) We use big-Omega notation to indicate that function f asymptotically grows at least as fast as function g, it is a lower bound on the growth of f. Formally, for constants $c > 0$ and $x_0 \geq 0, f(x) = \Omega(g(x)) : f(x) \geq g(x)$ for all $x \geq x_0$.

2.4. **Phase transitions.** The main subject of this paper—and the most fascinating—is the notion of criticality within these random models. Both random graphs and random simplicial complexes have *phase transitions* as they evolve. A phase transition is a rapid change in the observable properties of an object. In our context, this would be computable properties of a graph or complex—such as its homology or the number of edges. These phase transitions occur at a specific threshold in their evolution, described as a certain probability p with respect to the number of nodes n and the dimension d of the random simplicial complex (when discussing random graphs, $d = 1$ always). More precisely, let $\mathcal{P}$ be a graph property. We say that f is a threshold function for $\mathcal{P}$ if for any random simplicial complex $Y \in Y_d(n,p)$ whenever $p = o(f), Y$ does not have property $\mathcal{P}$ w.h.p, and whenever $p = \omega(f), Y$ does have property $\mathcal{P}$ w.h.p.

However, some thresholds may be one-sided sharp. This means t a property may or may not exist before a threshold with positive probability, however w.h.p the property $\mathcal{P}$ exists after the threshold.

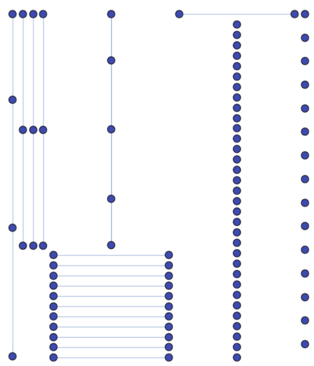


Figure 5. Acyclic graph with $n = 90$ and $p = \frac{5}{801} < \frac{1}{90}$.

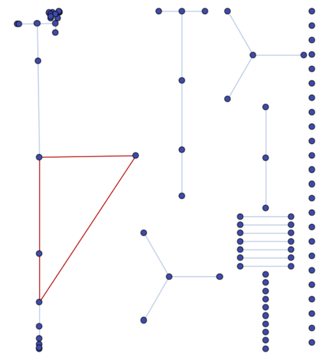


Figure 6. Cyclic graph with $n = 90$, $p = \frac{10}{801} > \frac{1}{90}$ and the cycle of length 4 highlighted.

Throughout this paper, we will discover many phase transitions. We will begin in the case of $d = 1$, phase transitions within random graphs. Then, we will scale to higher dimensions and explore phase transitions within random simplicial complexes of higher dimensions, drawing analogues between $d = 1$ and beyond. The discovery of these transitions requires multiple steps—generally, a loose bound or existence of a transition is first found, then upper and lower bounds are found more precisely to make the threshold tighter. We will explore examples of many of these steps.

2.5. **Probabilistic Tools.** Now let us also introduce some useful lemmas that will be utilized in the proofs to follow.

Lemma 2.1. *First Moment Method Let X be a non-negative integer valued random variable. Then*

$$\mathbb{P}(X > 0) \leq \mathbb{E}[X].$$

Lemma 2.2. *Chebyshev Inequality Let X be a random variable with a finite mean and variance. For $t > 0$,*

$$\mathbb{P}(|X_k - \mathbb{E}X_k| \geq t) \leq \frac{Var[X]}{t^2}.$$

3. Phase transitions in Erdős–Rényi

There are a multitude of fascinating transitions within this model. In this section, we explore a couple of the best-studied properties of graphs and their critical emergence.

Definition 3.1. A graph is said to be acyclic if no cycles exist within it. A cycle is a path with at least one step from a start vertex back to itself; no edge can be repeatedly visited.

Important to note, these cycles do not need to encapsulate the whole graph—cycles on connected components make a graph cyclic. As an example, Figure 5 and 6 depict a random graph immediately before and after the vanishing of this property. Let us begin our exploration of phase transitions by looking at a lovely proof for step towards the lower bound of the threshold for vanishing of acyclicity. This proof characterizes the *deep sub-critical phase*—when there are many more nodes n than edges m. Again, this is when $p = o(n^{-1})$ and $n/m \to 0$ as $n \to \infty$. Let $j = j(n)$ be some function growing slowly with n, for example $j = \log \log n$.

Theorem 3.2. *[FK23, Theorem 4.1] Let $m = \binom{n}{2} \times p$, the number of edges in $G_{n,p}$. If $m \ll n$, then $G_{n,p}$ is a forest w.h.p.*

Proof. Recall that a forest is a graph with no cycles. By calculating the expected value $\mathbb{E}$ of the number of cycles X and bound X via the First Moment Method, we can determine $\mathbb{P}(\mathbf{G}_{n,p}$ is not a forest) with a given p.

Let X be the number of cycles in $\mathbf{G}_{n,p}$, $p \leq \frac{3}{jn}$. The expectation of X within this interval can be expressed as

$$\mathbb{E}X = \sum_{k=3}^{n} \binom{n}{k} \frac{(k-1)!}{2} p^k.$$

We know that $\binom{n}{k} = \frac{n^k}{k!(n-k)!}$. To give an upper bound to $\mathbb{E}$, we can remove a term from this expression in our equation for $\mathbb{E}X$.

$$\mathbb{E}X = \sum_{k=3}^{n} \frac{n^k}{k!(n-k)!} \frac{(k-1)!}{2} p^k \leq \sum_{k=3}^{n} \frac{n^k}{k!} \frac{(k-1)!}{2} p^k.$$

Now let us continue to reduce this expression, substituting p^k and reducing the factorial terms:

$$\mathbb{E}X \leq \sum_{k=3}^{n} \frac{n^k}{2k} \frac{3^k}{j^k n^k}.$$

Finally, asymptotically, we can write this as

$$\mathbb{E}X = O(j^{-3}) \to 0.$$

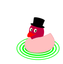
Here, the big-O notation denotes that asymptotically, the expected number of cycles within $\mathbf{G}_{n,p}$ will tightly approach zero. Intuitively, this is a sensible result considering that the presence of a forest implies a lack of cycles.
Furthermore, having formalized the expected value, we can now use the First Moment Method to complete the proof:

$$\mathbb{P}(\mathbf{G}_{n,p} \text{ is not a forest}) = \mathbb{P}(X \geq 1) \leq \mathbb{E}X = o(1).$$

Asymptotically, this implies that

$$\mathbb{P}(\mathbf{G}_{n,p} \text{ is a forest}) \to 1 \text{ as } n \to \infty.$$

Thus, within the interval $p \leq \frac{3}{jn}$, $G(n,p)$ is asymptotically w.h.p a forest, indicating a threshold at such point.



However, the threshold function for acyclicity can be more sharp than that. The classical theorem, proved by Pittel in the 1980s provides the sharpest threshold.

Theorem 3.3. *[Pit88]*
Let $G \sim G(n, c/n)$ with constant $c \in (0, \infty)$.

(1) *If $c < 1$ then w.h.p the probability that G contains cycles is*

$$\mathbb{P}(G \textit{ contains cycles}) = 1 - \sqrt{1-c}\exp(c/2 + c^2/4).$$

(2) *If $c > 1$, then w.h.p $G \sim G(n, c/n)$ has cycles.*

From this perspective, Theorem 3.2 only says that w.h.p $G(n,p)$ is acyclic for very small values of c. However, interestingly, this classical theorem highlights that as we approach the threshold for cyclicity, there is an increasingly non-zero probability that there may be cycles in the graph. Interesting to note, this highlights that the actual threshold is one-sided: there is non-zero (but not incredibly great) probability that there are cycles before the threshold $p = 1/n$; However, with incredible confidence, a.a.s, there are cycles after the threshold.

Now, there is a lot more than just the vanishing of acyclicity at $p = 1/n$. This moment is vital in the evolution of a random graph. Synonymous to an adolescent's rapid growth around its toddler years, a lot of notable phenomena emerges at $p = 1/n$. Let us continue to look at k-collapsibility, a property that also emerges at this threshold.

Definition 3.4. A graph G is said to be *k-collapsible* where the minimum degree $\delta(G) = k$ and any proper induced sub graph has a smaller minimum degree. An elementary collapse is a step where a vertex of degree 1 and the single edge that contains it are eliminated.

Collapsibility is a useful property to allow networks to be reduced and allow for more time-efficient computations regarding other graph properties. For graphs, this has the same threshold as acyclicity. Forests—-groups of trees—-and acyclic graphs are synonymous definitions. Trees are components of a graph where each two vertices are connected by exactly one path. Thus, they are able to be fully destroyed via a set of elementary collapses. Since acyclicity is a one-sided transition, for $G(n,p)$ k-collapsibility is as well.

There is even more that occurs at $p = 1/n$. Let us next look at a mysterious concept: the emergence of a giant component.

Definition 3.5. A **"giant" component** is a connected component of significant size in a graph.

In $G(n,p)$ this property has a gathered a lot of thorough study, yielding quite precise characterizations of the size of this component and its threshold. This giant component is unique and distinctive, as we will explore momentarily. However, it may behoove the reader to think of why this is the case intuitively: as p increases and a large component grows in size, the chance that a new independently added edge is formed between a vertex outside the component with any vertex in the component (subsequently making the giant component larger) increases greatly. This explains the criticality surrounding the rise of a giant component. As an example, Figure 7 and 8 depict a random graph immediately before and after the vanishing of this property.

Interestingly, the threshold is tight and sharp for the emergence of such a giant component. Let us make the characterization of the giant components and other components in the Erdős–Rényi model as follows,

Theorem 3.6. *[Łuc90, Bol84, Erd61]*
Let error term $\varepsilon = \varepsilon(n) > 0$ satisfy $\varepsilon \to 0$ and $\varepsilon^3 n \to \infty$, as $n \to \infty$

1. *If $p \leq \frac{1-\varepsilon}{n}$, then w.h.p all components of G(n,p) have size at most $O(\varepsilon^{-2} \log \varepsilon^3 n)$*
2. *If $p \geq \frac{1+\varepsilon}{n}$ then w.h.p the size of the largest component of $G(n,p)$ is $(1 + o(1))2\varepsilon n$ while all other components have size at most $O(\varepsilon^{-2} \log \varepsilon^3 n)$.*

Here, we prove an integral part of this theorem:

Theorem 3.7. *[FK23, Theorem 4.10]*
W.h.p $\mathbf{G}(n,p)$, at $c > 1$, $p > \frac{c}{n}$, consists of a unique giant component with $(1 - \frac{x}{c} + o(1))n$

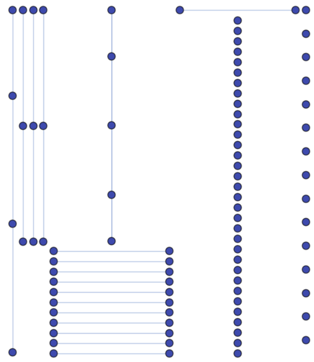


Figure 7. Graph without a distinct giant component, with $n = 90$ and $p = \frac{5}{801} < \frac{1}{90}$

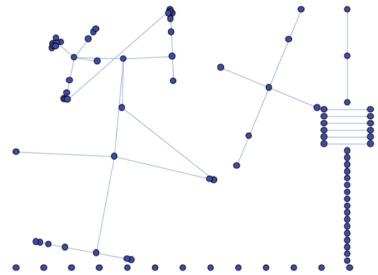


Figure 8. Graph with a distinct giant component (leftmost giant component), with $n = 90$ and $p > \frac{1}{90}$

vertices. Here $0 < x < 1$ is the unique solution of the equation $xe^{-x} = ce^{-c}$. The remaining components are of order at most $O(\log n)$.

First, some extra tools to add to our toolbelt for this proof.

Lemma 3.8. *Chebyshev's inequality Let $k = O(\log n)$, X be a random variable with finite mean and variance with $\mathbb{E}X_k > 0$. The probability that the difference between the actual and expected value X is greater than the expectation of X multiplied by a constant $\epsilon > 0$ can be bounded as*

$$\mathbb{P}(|X_k - \mathbb{E}X_k| \geq \epsilon \mathbb{E}X_k) \leq \frac{1}{\epsilon^2 \mathbb{E}X_k} + \frac{2ck^2}{\epsilon^2 n} = o(1).$$

Where this expression is derived from Chebyshev inequality (Lemma 2.2) with $t = \varepsilon \mathbb{E}X_k$ and $VarX \leq \mathbb{E}X_k + 2ck^2(\mathbb{E}X_k)^2/n$.

Note that the c and k used above are the same as used throughout the rest of the proof.

Lemma 3.9. *The Markov Inequality Let X be a non-negative random variable. Then, for all t ¿0,*

$$\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}X}{t}$$

Lemma 3.10. *If $c > 0, c \neq 1$ is a constant, k is the order the component, and $x = x(c)$ is as defined above, then*

$$\frac{1}{x} \sum_{k=1}^{\infty} \frac{k^{k-1}}{k!} (ce^{-c})^k = 1.$$

Proof. Proof of Theorem 3.7 Let Z_k be the number of components of order k in $\mathbf{G}_{n,p}$, and let $A > 0$ be a constant. We can bound Z_k by the number of trees with k vertices that span the component, getting us

$$\mathbb{E}Z_k \leq \binom{n}{k} k^{k-2} p^{k-1} (1-p)^{k(n-k)}$$

We can then continue to simplify this as follows:

$$\mathbb{E}Z_k \leq A(\frac{ne}{k})^k k^{k-2} (\frac{c}{n})^{k-1} e^{-ck+ck^2/2}$$

$$\mathbb{E}Z_k \leq \frac{An}{k^2}(ce^{1-c+ck/n})^k.$$

Now let's define some constants! Let $\beta_1 = \beta_1(c)$ be small enough to satisfy

$$ce^{1-c+c\beta_1} < 1,$$

and let $\beta_0 = \beta_0(c)$ be large enough so that

$$(ce^{1-c+o(1)})^{\beta_0 \log n} < \frac{1}{n^2}.$$

Finally, let $k_0 = \frac{1}{2\alpha}\log n, \alpha = c - 1 - \log c$

With such definitions of β_0, β_1, there is no component of order $k \in [\beta_0 \log n, \beta_1 n]$; ie. w.h.p there are no components that have order k that fall within the range $k \in [\beta_0 \log n, \beta_1 n]$. Thus we transition to define "small" and "giant" components: small components have order $1 \leq k < \beta_0 \log n$, giant components have order $k > \beta_1 n$.

Now, we will estimate the total number of vertices on small components by splitting small components into three categories: trees of order $1 \leq k \leq k_0$, trees of order $k_0 + 1 \leq k \leq \beta_0 \log n]$, and connected components of order $1 \leq k \leq \beta_0 \log n$. Where $k_0 = \frac{1}{2\alpha}\log n$.

Lets begin with the smallest!

Lemma 3.11. *The expected number of vertices within small tree components of order* $1 \leq k \leq k_0$ *is* $\frac{nx}{c}$

Proof. The expected value of the number of vertices of such trees of order k can be approximated as

$$\mathbb{E}(\sum_{k=1}^{k_0} kX_k) \approx \frac{n}{c}\sum_{k=1} k_0 \frac{k^{k-1}}{k!}(ce^{-c})^k.$$

This sum can be expanded from k_0 to infinity for ease of computation by using $\frac{k^{k-1}}{k!} < e^k$ and $ce^{-c} < e^{-1}$ for $c \neq 1$:

$$\mathbb{E}(\sum_{k=1}^{k_0} kX_k) \approx \frac{n}{c}\sum_{k=1}^{\infty} \frac{k^{k-1}}{k!}(ce^{-c})^k.$$

Now, we can use this bound for our task to compare the expectation to the actual value of X_K. Using Lemma 3.8 and setting $\epsilon = 1/\log n$, we can approximate the probability X_k deviates from its mean more than $1 \pm \epsilon$ on the interval $1 \leq k \leq k_0$ is at most

$$\sum_{k=1}^{k_0}[\frac{(\log n)^2}{n^{1/2-o(1)}} + O(\frac{(\log n)^4}{n}] = o(1).$$

X_k does not deviate from its mean greatly!

If $x = x(c), 0 < x < 1$ is the unique solution within (0,1) of the equation $x_1 e^{-x_1} = c_1 e^{-c_1}$, then we can approximate the number of vertices in trees on the interval $[1, k_0]$ as

$$\mathbb{E}(\sum_{k=1}^{k_0} kX_k) \approx \frac{n}{c}\sum_{k=1}^{\infty} \frac{k^{k-1}}{k!}(xe^{-x})^k.$$

Now by applying Lemma 3.10 and some arithmetic, we can simplify this to get w.h.p

$$\frac{n}{c}\sum_{k=1}^{\infty} \frac{k^{k-1}}{k!}(xe^{-x})^k = \frac{nx}{c}.$$

Thus, we find that the number of components within this interval to be approximately $\frac{nx}{c}$. Now, lets continue to consider the next interval of tree components: $k_0 < k \leq \beta_0 \log n$.

Lemma 3.12. *The number of vertices kX_k of small tree components with order: $k_0 < k \leq \beta_0 \log n$ is approximately $o(n)$.*

Proof. Again, we will begin by bounding the expected value of the number of vertices.

$$\begin{aligned}\mathbb{E}(\sum_{k=k_0+1}^{\beta_0 \log n} kX_k \leq & \frac{n}{c} \sum_{k=k_0+1}^{\beta_0 \log n} (ce^{1-c+ck/n})^k \\ =& O(n(ce^{1-c})_0^k \\ =& O(n^{1/2+o(1)}).\end{aligned}$$

Now, by Markov's inequality (Theorem 3.9), w.h.p.

$$\sum_{k=k_0+1}^{\beta_0 \log n} kX_k = o(n).$$

This is fascinating—-within this interval, the number of vertices in components grows very slowly, at the very best proportionally less than the number of nodes.

Lemma 3.13. *The number of vertices kY_k of small connected components with order $1 < k \leq \beta_0 \log n$ is approximately $o(n)$.*

Proof. Now consider the number of non-tree components Y_k with vertices k, $1 \leq k \leq \beta_0 \log n$, and the number of vertices within these components kY_k.

$$\begin{aligned}\mathbb{E}(\sum_{k=1}^{\beta_0 \log n} kX_k) &\leq \sum_{k=1}^{\beta_0 \log n} \binom{n}{k} k^{k-1} \binom{k}{2} (\frac{c}{n})^k (1-\frac{c}{n})^{k(n-k)} \\ &\leq \sum_{k=1}^{\beta_0 \log n} k(ce^{1-c+ck/n})^k \\ &= O(1)\end{aligned}$$

Similar to the last interval, we again use the Markov's inequality (Theorem 3.9) to get w.h.p

$$\sum_{k=1}^{\beta_0 \log n} kY_k = o(n).$$

Fascinatingly, we have found a similar result in both small connected components and small tree components of order in the interval $[k_0, \beta_0 \log n]$ regarding their rate of growth—less than $o(n)$ with respect to n, indicating that it is growing very slowly.

Thus, in summary, so far we have proved the number of vertices within the interval $1 \le k \le \beta_0 \log n$ to be approximately $\frac{nx}{c}$ since the number of vertices in trees of order $1 \le k \le k_0$ is the only significantly contributing interval given that the other two grow loosely slower than proportional to the growth of number of nodes. The upper bound for our definition of "small components" was $\beta_0 \log n$, we can generalize this to say that small components have a maximum order of $O(\log n)$.

Another question: what if there's multiple "giant components"? Well, we can say with high probability that you won't encounter multiple—the "giant component" is a separate, distinct entity. Let

$$c_1 = c - \frac{\log n}{n} \text{ and } p_1 = \frac{c_1}{n}.$$

And let $\mathbf{G}_{n,p_1}$ be the Erdős–Rényi graph with probability p_1.

Additionally, let us define p_2 in terms of the general probability p and other probability constant p_1,

$$1 - p = (1 - p_1)(1 - p_2).$$

Where $p_2 \ge \frac{\log n}{n^2}$.

We can adapt our earlier definition of $x(c)$ to this step as well. If $x_1 e^{-x_1} = c_1 e^{-c_1}$ then $x_1 \approx x$ and thus our previous analysis can be applied; w.h.p $\mathbf{G}_{n,p_1}$ has no components with size in the range $[\beta_0 \log n, \beta_1 n]$. Suppose there are components $C_1, C_2, \ldots, C_l$ where any $|C_i| > \beta_1 n$ and $l \le 1/\beta_1$. Now for the fun! We proceed to add edges of $\mathbf{G}_{n,p_2}$ to $\mathbf{G}_{n,p_1}$ (note the order); we take their union:

$$\mathbf{G}_{n,p} = \mathbf{G}_{n,p_1} \cup \mathbf{G}_{n,p_2}.$$

To prove the uniqueness of the "giant component", we must see if this union of graphs continues to yield many "large" ($k > \beta_1 n$) components or if eventually most components merge into one; in other words, whether or not $\mathbf{G}_{n,p_1}$ is the roots of a growing giant component that $\mathbf{G}_{n,p_2}$ adds to. This can be expressed as the probability of the existence of values $i, j \in \mathbb{N}$ such that performing the union operation, adding edges of $\mathbf{G}_{n,p_2}$ to $\mathbf{G}_{n,p_1}$, will *not* connect components C_i and C_j. Then

$$\mathbb{P}(\exists\, i, j \in \mathbb{N} : \text{no } \mathbf{G}_{n,p} \text{ edge joins } C_i \text{ with } C_j) \le \binom{l}{2}(1 - p_2)^{(\beta_1 n)^2}$$

$$\le l^2 e^{-\beta_1^2 \log n} = o(1).$$

Which highlights that the probability that the union of $\mathbf{G}_{n,p_1}$ and $\mathbf{G}_{n,p_2}$ will not form a giant component is very asymptotically unlikely; the elements of such set will grow loosely less than constant 1. Thus, this highlights the unique nature of the formation of a giant component, as highlighted earlier. As the random graph evolves, this giant component C_i will be uniquely destined to grow in size.

Given that we've proven the giant component exists and is singular, we can use our results from above to formalize information regarding its size. The number of vertices in $G(n, p)$, n, is composed of the number of vertices of small components—components within the interval $1 \le k \le \beta_0 \log n$—and large components—of order $k > \beta_0 \log n$. From our previous results, $n = \frac{nx}{c} + L$, where L is the number of vertices within the giant component. $L = (1 - \frac{x}{c})n$ vertices.

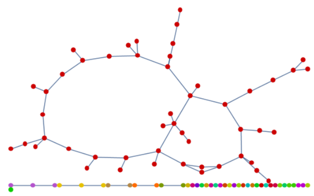


Figure 9. Graph with $n = 90$ and $\frac{1}{90} < p < \frac{\log 90}{90}$.

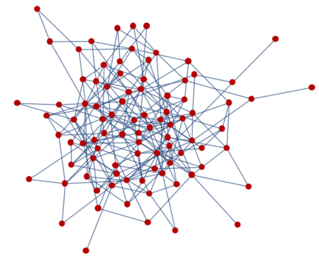


Figure 10. Graph with $n = 90$ immediately after $p < \frac{\log 90}{90}$.

Remark 3.14. This proof is quite fascinating—it answers many questions regarding the nature of the giant component. It highlights the uniqueness of the "giant" component. Furthermore, these results imply that this giant component is distinctive and that it will continue to merge smaller components into itself as the probability p in $G(n,p)$ grows past the threshold. This characterizes random graphs with quite detail: around $p = \frac{1}{n}$, the graph transitions from a scattered collection of trees and small connected components to a giant lump of components. Within the context of phase transitions, the unique, predictable nature of this component is fascinating given the random network it is embedded in.

We are now jumping away from the excitement at $p = \frac{1}{n}$ to a new fascinating point in the evolution of a random graph: $p = \frac{\log n}{n}$. This is where the transition to full connectivity emerges!

Definition 3.15. A graph is **connected** if a path exists between every pair of vertices u and v within the graph.

The notion of connectivity and the minimal number of edges needed to achieve connectivity is a core studied area of graph theory. For concreteness, Figure 9 depicts a random graph before being fully connected, and Figure 10 depicts a random graph immediately after $p = \frac{\log n}{n}$. Let us now analyze this threshold in further specificity,

Theorem 3.16. *[ER59]*
W.h.p $G(n,p)$ is connected if

$$p \geq \frac{\log n + \omega(1)}{n},$$

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

$$p \leq \frac{\log n - \omega(1)}{n}.$$

Note that $\omega(1)$ is any function that grows asymptotically greater than 1, $\omega(1) \to \infty$ as $n \to \infty$.

A measure of connected components in a graph is the 0th-homology group. A graph becoming fully connected is analogous with the vanishing of the 0-th homology group. Let X_1 be a 1-dimensional topological space, the graph X_1 will be connected when $H_0(X,G) = 0$ with any group of coefficients G. We will explore this analogy in further depth when analyzing thresholds within the Linial-Meshulam model. Though the standard proof treats this as a homology-vanishing problem, an alternate method can also be used as a close approximation and reveals additional qualities of the graph that emerge at $p = \frac{\log n}{n}$. Having isolated vertices implies disconnectivity, so it can be used as a starting point for rough lower bound for the connectivity threshold.

Theorem 3.17. *[Roc24]*
For $G(n,p)$, not having any isolated vertices has threshold function $p = \frac{\log n}{n}$. Let $0 \leq p_n \leq 1$,

(1) *If $p_n \ll \frac{\log n}{n}$, then a.a.s $G(n,p)$ has isolated vertices*
(2) *If $p_n \gg \frac{\log n}{n}$, then a.a.s $G(n,p)$ does not have isolated vertices.*

We will use the following definition and lemma to prove this theorem.

Definition 3.18. Let $a \sim b$ denotes two events, a and b, where $a \neq b$ and a and b are not independent of each other. For example, when analyzing events in graph theory, knowing that vertex i is isolated increases the probability that vertex j is isolated since there surely cannot have an edge between i and j.

Lemma 3.19. *Precursor to second moment method for sums of indicators Let X be a non-negative random variable (not identically zero). Let $A_1, \ldots, A_m$ be a collection of events where $A_i \sim A_j$ denotes that the two events are not independent of each other. Then, letting*

$$\mu_m := \sum_i \mathbb{P}[A_i],\ \gamma_m := \sum_{i \sim j} \mathbb{P}[A_i \cap A_j],$$

can yield the following expression for a version of the second moment method:

$$\mathbb{P}[X > 0] \geq \frac{\mathbb{E}_{n,p_n}[X_n^2]}{(\mathbb{E}_{n,p_n}[X_n])^2} = 1 - \frac{Var[X]}{\mathbb{E}X^2 + Var[X]} = \frac{\mu_n + \gamma_n}{\mu_n^2}.$$

Proof. Proof of Theorem 3.17 Let X_n be the number of isolated vertices in the Erdős–Rényi graph $\mathbf{G}_n \sim \mathbf{G}_{n,p_n}$. The expectation of X_n can be written as

$$\mathbb{E}_{n,p_n}[X_n] = n(1 - p_n)^{n-1}.$$

The rest of the proof will proceed as follows: we begin by proving the absence of isolated vertices when $p \gg \frac{\log n}{n}$, then transition to prove the high probability of a non-zero number of isolated vertices in the when $p \ll \frac{\log n}{n}$.

Using $1 - x \leq e^{-x}$ for $x \in \mathbb{R}$ and $x \geq 0$, we can bound $\mathbb{E}_{n,p_n}[X_n]$ and observe its asymptotic behavior :

$$\mathbb{E}_{n,p_n}[X_n] \leq e^{\log n - (n-1)p_n} \to 0.$$

Using the first moment method (Lemma 2.1),

$$\mathbb{P}_{n,p_n}[X_n > 0] \to 0.$$

Thus, the probability that there is a non-zero amount of isolated vertices following the identified threshold is asymptotically zero.

Now that we've proven one direction of the threshold, lets transition to the other and prepare to apply a version of the second moment method (Lemma 3.19). Let A_j be all the subspace of events in which vertex j is isolated and μ_n be the union of probabilities of each event occurring,

$$\mu_n = \sum_i \mathbb{P}_{n,p_n}[A_i] = n(1 - p_n)^{n-1}.$$

Using the same principle from above, we can substitute p_n for $x \in [0, 1/2]$,

$$\mu_n \geq e^{\log n - np_n - np_n^2}.$$

Notably, $\mu_n \to +\infty$ when $p_n \ll \frac{\log n}{n}$ since $\mu_n \geq e^{\log n - n\frac{\log n}{n} - n(\frac{\log n}{n})^2} = e^{n(\frac{\log n}{n})^2}$.

The probability that both sets of events A_i and A_j for $i \sim j$ occur is expressed as

$$\mathbb{P}_{n,p_n}[A_i \cap A_j] = (1 - p_n)^{2(n-1)+1}.$$

We can generalize this to all events A_i, A_j with a two-varied summation that runs through all permutations of events i, j:

$$\gamma_n = \sum_{i \neq j} \mathbb{P}n, p_n[A_i \cap A_j] = n(n-1)(1 - p_n)^{2(n-1)+1}.$$

Now, we shall use Lemma 3.19, a version of the second moment method, to complete the proof. Let us do

$$\mathbb{P}[X > 0] \geq \frac{\mathbb{E}_{n,p_n}[X_n^2]}{(\mathbb{E}_{n,p_n}[X_n])} = \frac{\mu_n + \gamma_n}{\mu_n^2}.$$

Substituting for μ_n and γ_n,

$$\leq \frac{n(1-p_n)^{n-1} + n^2(1-p)n)^{2n-3}}{n^2(1-p_n)^{2n-2}}$$

$$\leq \frac{1}{n(1-p_n)^{n-1}} + \frac{1}{1-p_n} \to 1 + o(1).$$

In summary, for the case of $p_n \ll \frac{\log n}{n}$, the probability that there is a non-zero amount of isolated vertices is asymptotically highly probable.

Theorem 3.20. *[Roc24]*
For $G(n, p)$, connectivity has threshold function $p = \frac{\log n}{n}$. Let $0 \leq p_n \leq 1$,

(1) *If $p_n \ll \frac{\log n}{n}$, then a.a.s $G(n, p)$ is disconnected*
(2) *If $p_n \gg \frac{\log n}{n}$, then a.a.s $G(n, p)$ is fully connected.*

Proof. Let us begin with proving Part 1 of Theorem 3.2. If $p_n \ll \frac{\log n}{n}$, Theorem 3.17implies that $G(n, p)$ is disconnected since there are isolated vertices. Now lets look at the other direction, Part 2. Assume $p_n \gg \frac{\log n}{n}$ and let $\mathcal{D}_n$ be the event in which $G(n, p_n)$ is disconnected. The goal of this proof is to show that $\mathbb{P}_{n,p_n}[\mathcal{D}_n]$ is a.a.s unlikely. We will accomplish this by bounding $\mathbb{P}_{n,p_n}[\mathcal{D}_n]$ by the number of subsets of vertices that are disconnected from all other vertices in the graph, that have degree 0. Formally, let $Y_k, k \in \{1, \ldots, n/2\}$ be the number of subsets of k vertices with degree 0. We can bound the probability that $Y_k \neq 0$, implying $G(n, p)$ is disconnected, using the first moment method:

$$\mathbb{P}_{n,p_n}[\sum_{k=1}^{n/2} Y_k > 0] \leq \sum_{k=1}^{n/2} \mathbb{E}_{n,p_n}[Y_k].$$

Now, we can use this information to bound $\mathbb{P}_{n,p_n}[\mathcal{D}_n]$,

$$\mathbb{P}_{n,p_n}[\mathcal{D}_n] \leq \mathbb{P}_{n,p_n}[\sum_{k=1}^{n/2} Y_k > 0] \leq \sum_{k=1}^{n/2} \mathbb{E}_{n,p_n}[Y_k].$$

Let's step away from the full expression and solely consider the expectation of Y_k; $\mathbb{E}_{n,p_n}[Y_k]$ is easy to estimate. Using earlier provided information that $k \leq n/2$ and $k! \geq (k/e)^k$,

$$\mathbb{E}_{n,p_n}[Y_k] = \binom{n}{k}(1-p_n)^{k(n-k)}.$$

Simplifying for $\binom{n}{k}$ and reducing terms, we get that

$$\leq \frac{n^k}{k!}(1-p_n)^{kn/2} \leq (en(1-p_n)^{n/2})^k.$$

Let us now return to summing over k.

$$\mathbb{P}_{n,p_n}[\mathcal{D}_n] \leq \sum_{k=1}^{+\infty}(en(1-p_n)^{n/2})^k. \tag{3.1}$$

This geometric series is dominated asymptotically by its first term. Thus, we get the first term, set $k = 1$, and summarize this expressions' growth (ignoring constants) using big-O notation,

$$= O(n(1-p_n)^{n/2}).$$

Considering that $en(1-p_n)^{n/2} = o(1)$ when $p_n \gg \frac{\log n}{n}$, we get that

$$= o(1).$$

Also note that we set the upper bound in (3.1) to $+\infty$ instead of $n/2$ since it is dominated by the first term in either case and we're trying to analyze asymptotic behavior.

Remark 3.21. Interestingly, connectivity and not having isolated vertices have the same threshold of transition. The two have a slight distinction—previously isolated vertices could consolidate be in multiple smaller components, not just one giant connected component. Yet, the vanishing of isolated vertices and emergence of connectivity still share the same threshold. This raises praise to the power of the giant component within the evolution of random graphs—it is incredibly distinct if it continues to grow to encompass the connectivity of the whole network instead of permitting the other components to grow.

4. Phase transitions in Linial-Meshulam

Despite the similarity in definition, these random simplicial complexes have many different—and still deliberated—properties; we must refine our definitions of such properties discussed above to adapt to higher dimensional behavior.

Definition 4.1. The classification **acyclic space** is a space X without homological cycles, where $H_i(X) = 0$ for each and every i, $0 \leq i \leq d$. This means that there are no nontrivial holes. An element of $H_i(X)$ is a d-cycle.

These acyclic spaces, though simple, can be used to build more fascinating spaces/objects. At $p = \frac{c_d^*}{n}, c_d^* = d + 1 - o(1)$ there is w.h.p a transition to cyclicity. Interesting to note, this transition is *one-sided*, before the threshold the simplicial complex may or may not contain cycles. Intuitively, this makes sense: in the sub critical phase, there is a full d-1 skeleton with sparse d-faces being added according to probability p. Some d faces will cover a (d-1) face. After a certain threshold, there will be more d faces than needed to cover all the (d-1) faces and thus some will have no boundary. Comparing back to Erdős–Rényi, interestingly, the one-sided nature of this transition remains even as we scale to higher dimensions.

Theorem 4.2. *[LP16, AL15]*
Let t_d^* be the unique root in (0,1) of

$$(d+1)(1-t_d^*) + (1+dt_d^*)\log t_d^* = 0,$$

and let c_d be defined as

$$c_d^* := \frac{-\log t_d^*}{(1-t_d^*)^d}.$$

Let $Y = Y_d(n,p)$.

(1) *If $p < c_d^*/n$, then w.h.p $H_d(Y,G)$ is generated by simplex boundaries. So*

$$\mathbb{P}[H_d(Y,\mathbb{R}) = 0] \to \exp(-c^{d+2}/(d+2)!).$$

(2) *If $p > c_d^*/n$, then w.h.p $H_d(Y,G) \neq 0$.*

Note that this highlights the one-sided nature of the threshold for cyclicty by providing the exact probability that $H_d(Y,\mathbb{R}) = 0$ instead of claiming it as a.a.s acyclic before the threshold. Mirroring our analysis for the Erdős–Rényi model, let us examine a proof for the lower bound of this threshold in further depth.

Theorem 4.3. *[LP16]*
For every $c < c_d^*$, asymptotically almost surely, $H_d(Y_d(n, \frac{c}{n}); \mathbb{R})$ is either trivial or generated by at most a bounded number of copies of the boundary of a (d+1) simplex.

This theorem establishes a tight threshold for acyclicity within random simplicial complexes. Before we begin, it is important to include a couple of tibbits of extra information.

Lemma 4.4. *Let $Y_0 \in Y_d(n, \frac{c'}{n})$ for some $c < c' < c_d^*$.* $\mathbb{E}[\dim(H_d(Y_0;\mathbb{R})] = o(n^d)$

The proof for this lemma involves a much denser mathematical background than the scope of this paper. Nonetheless, this is a powerful lemma to characterize the number of holes in a random simplicial complex before the threshold. Note the asymptotic notation—this is still a one sided transition.

Lemma 4.5. *For every $c > 0$ a.a.s every minimal core in $Y_d(n, \frac{c}{n})$ is either the boundary of a (d+1)-simplex, or it has a cardinality of at least δn^d, where $\delta > 0$ depends only on c.*

This lemma is quite powerful when we consider that every d-cycle is a core. We can substitute these phrases, bringing us to

Lemma 4.6. *For every $c > 0$, a.a.s every d-cycle of $Y_d(n, \frac{c}{n})$ is either the boundary of a (d+1) simplex or it is a big cycle; the cycle is either trivial or it has at least δn^d d-faces. Here $\delta > 0$ depends only on c.*

This lemma will be the main basis of the proof to come. Using this information, we can prove that big cycles do not exist within a tight range before the threshold in order to prove that all cycles within this interval are bounded by a (d+1)-simplexes.

The following is a high probability proof regarding the threshold of acyclicity.

Proof. Proof of Theorem 4.2 Using the Lemma 4.4, we can apply Markov inequality (Theorem 3.9): since $\mathbb{E}[\dim H_d(Y_0;\mathbb{R})] = o(n^d)$, a.a.s $\dim H_d(Y_0;\mathbb{R}) = o(n^d)$. Let k be the number of samples from a uniformly random sample of $|F_d(Y_0)|$ values from the range $[0,1]$ that are $< 1 - c/c'$. Furthermore, allow *d-complexes* from Y_0 to be defined as the following: $Y_0 \supset Y_1 \supset Y_2 \supset \cdots \supset Y_k$ where Y_{i+1} results by removing a random d-face σ_i from Y_i for

$i = 0, 1, \ldots, k-1$. Looking at the bigger picture, we are iteratively removing the d-faces of a certain random simplicial complex, where the reduced simplicial complex after i deletions is Y_i. Clearly, $Y_i \in Y_d(n, \frac{c}{b})$.

If Y_i *does* contain a big d-cycle, then with a constant probability bounded away from zero, the random d-face σ_{i+1} is in it. This would be the next d-face to be removed. In such a case,

$$\dim H_d(Y_{i+1}; \mathbb{R}) = \dim H_d(Y_i, \mathbb{R}) - 1.$$

Interpolating this equation across a range of values of i, if Y_k has a big cycle, then $\dim H_d(Y_i; \mathbb{R})_{i=1}^{k}$ is a random sequence of $\Omega(n^d)$ non negative integers. The sequence would begin with a value of $o(n^d)$ and have a constant probability of dropping by 1 after each step. However, this is a clear contradiction: beginning with a number that is $o(n^d)$, then subtract 1 $\Omega(n^d)$ times will eventually return a negative number; this is a nonsensical result for Betti numbers. Thus, there are no big cycles within this interval and only cycles that are part of a boundary of a (d+1)-simplex.

Remark 4.7. This proof is particularly fascinating—it precomputes the results of an algorithm that would be applied to a big cycle in order to prove that the results are not valid. A scenario with big cycles is imagined, then proved to be solely a fantasy, not a mathematical reality. Though both the proof for the threshold in Erdős–Rényi and Linial-Meshulam utilize the expected value of the number of cycles to support their arguments, they use it in vastly different—and fascinating—ways. More broadly, the transitions for both $G(n, p)$ and $Y_d(n, p)$ are very similar in their one-sided nature. This is a fascinating parallel across dimensions. One of the biggest struggles, as we will see shortly, with phase transitions in random simplicial complexes is finding suitable higher dimensional equivalents of properties within graphs given that these two objects—graphs and topological spaces—are usually regarded as different. The fact that the one-sided nature of the transition is preserved through this higher dimensional equivalent is a quite satisfactory result.

The next graph property where topological equivalents must be analyzed for is giant component emergence. A question posed by Linial Meshulam was the implications of the rise of a giant component for random simplicial complex. Though there was diverging ideas earlier on, a "giant" shadow is now the dominant higher dimensional equivalent within the field.

Definition 4.8. [LNPR14]
A **shadow** $SH(X)$ is a set of d-simplicies that aren't in the set of X, but are in the span of X. It is the set of all d-simplicies that, if added, would create cycles in the random simplicial complex. Formally,

$$SH_{\mathbb{R}}(X) = \{\sigma \notin Y : H_d(X, \mathbb{R}) \text{ is a proper subspace of } H_d(X \cup \{\sigma\}; \mathbb{R})\}.$$

For the relentless reader curious for a *why*: since an d-simplex in a shadow must form a cycle, its implied that it must be added to a "connected component", a component of a complex in which there exists a collection of d-1 faces with the potential to form cycles; it is a measure of connectivity and cyclicity. Two vertices must be in the same component to have the potential to form cycles. In the case of $G(n, p)$, the shadow with $d = 1$ reaches positive density at the same time the giant component emerges, making this "giant" shadow (so to speak) a scalable and natural equivalent of the giant component in higher dimensions.

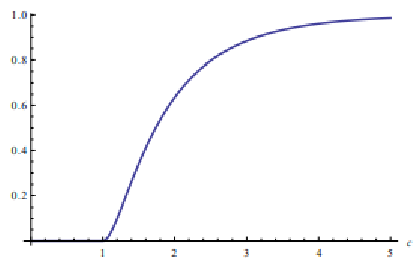


Figure 11. Density of the shadow of $G(n, \frac{c}{n})$ as c increases.

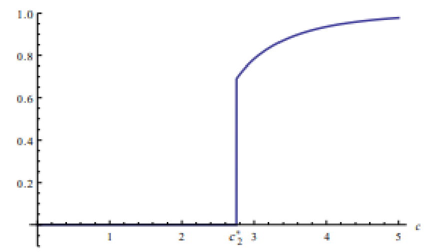


Figure 12. Density of the shadow of $Y_2(n, \frac{c}{n})$ as c increases.

The tightest bound for the threshold of this "giant shadow" is—interestingly—at the same threshold for acyclicity in $Y_d(n,p)$, as explored in Theorem 4.2. Referring back to analogies in $G(n,p)$, this makes sense: a giant component emerges at the same time that cycles do. There is a correlation between the idea of cycles and connectivity, no matter the dimension. A fascinating theorem characterizing the phase transition regarding the density and size of this shadow is as follows,

Theorem 4.9. *[LP16] Let $Y \in Y_d(n, \frac{c}{n})$ for some $d \geq 2 \in \mathbb{Z}$ and $c > 0 \in \mathbb{R}$. Furthermore, let c_d^* be defined same as above.*

(1) *If $c < c_d^*$, then a.a.s.,*

$$|SH_{\mathbb{R}}(Y)| = \Theta(n).$$

(2) *If $c > c_d^*$, let t_c be the smallest root in (0,1) of $t = e^{-c(1-t)^d}$ then a.a.s.,*

$$|SH_{\mathbb{R}}(Y)| = \binom{n}{d+1}((1-t_c)^{d+1} + o(1)).$$

These values may seem very arbitrary—it is fascinating to see an example of this theorem in action. An example of the evolution of the shadow is highlighted in Figure 12. Evidently, before the threshold, the shadow does not have positive density since in this interval, the size of the shadow grows with dependence only on the number nodes n . After the threshold, the density of the shadow suddenly spikes.

Remark 4.10. A lot is still unknown about the implications of this "shadow", however, it is the dominant concept for a higher-dimensional equivalent of the birth of a giant, connected component in the Erdős–Rényi model. This definition of the shadow may also imply that the giant shadow could be dominated by a bunch of tiny components, instead of being a metric for the giant component. However, Dotterrer-Guth-Kahle are currently working to prove that after a certain threshold the sub-complexes within the shadow are not small.

We began our drawing of parallels between d=1 and $d > 1$ with our discussion regarding cyclicity. Given that the concept of a shadow was specifically conceived to be an analogy of the giant component in higher dimensions, it has a lot of well crafted properties, such as quantifying the size and presence of "connected components" within a random simplicial complex. However, there are some differences: referring to Figure 11, we can see that though the transition to positive density of the shadow in $G(n,p)$ has a continuous derivative, while in $Y_2(n,p)$, Figure 12, it has a discontinuous derivative around the point $p = \frac{c_d^*}{n}$.

Next, we will look at the vanishing of d-collapsibility within $Y_d(n,p)$. Interestingly, d-collapsibility is no longer analogous to acyclicity.

Definition 4.11. A topological space is ***d*-collapsible** if it is possible to eliminate all d-simplices from the simplicial complex series of elementary collapses. Let σ be a $(d-1)$-dimensional face that is contained in only one d-dimensional face, τ; we can also describe σ as an *exposed* or "free" d-face. An elementary collapse is then a single deletion that removes both σ and τ. On the other hand, a d-core is a d-complex with no exposed $(d-1)$-faces, the terminating result of a series of elementary collapses on a simplicial complex not d-collapsible. These d-cores are also d-cycles.

Hand in hand, it has a different threshold for vanishing compared to acyclicity.

Theorem 4.12. *[AL16, ALLM13]*
Let $\eta > 0$ be some fixed value for the following equation of x:

$$0 = e(1-x)^d(-\eta).$$

And n_d be any value of η for which the equation above has some root $x < 1$. There is a tight threshold for the vanishing of d-collapsibility at $p = \frac{\eta_d}{n}$. Namely,

(1) *Fix $c < \eta_d$. A d-dimensional complex $Y \in Y_d(n, \frac{c}{n})$ is a.s. d-collapsible or it contains a copy of $\partial\Delta_{d+1}$.*
(2) *For every $c > n_d$ a d-dimensional complex $Y \in Y_d(n, \frac{c}{n})$ is a.s. not d-collapsible.*

The proof of this theorem will be collapsed (pun not intended) for the sake of technical detail. However, the techniques used are quite fascinating. In particular, the Part 1 of Theorem 4.12, similar to the proof of Theorem3.2, involves something more algorithmic in nature. The proof splits its techniques into two *epochs* or iterations—the two are the same, the second epoch can just be thought of as a slowed down version of the first to yield more rich analysis. Let $r > 0 \in \mathbb{Z}$. In the first epoch, *r phases* of collapses are carried out simultaneously. In the second epoch, each (d-1) face removal is carried out one-by-one in a random order. Let X_i be the number of free $(d-1)$ faces at the i-th collapse step in the second epoch). A simplicial complex is not d-collapsible if $X_i = 0$ but the number of $(d-1)$-faces not eliminated is not zero. The authors analyze activity during the second epoch thoroughly, showing that the expected drop in free faces in one step, $\mathbb{E}(X_i - X_{i+1})$, is a.s. sufficiently large such that at some moment $X_i = 0$ but some (d-1) faces still exist. In other words, the number of free faces available declines faster than the number of (d-1) simplicies present in the simplicial complex. This is because τ will contain other (d-1) simplicies that are removed as a result of the elementary collapse, complicating the relationship between the two.

Remark 4.13. Just as in the Erdős–Rényi model, any d-collapsible simplicial complex is also acyclic: if Y is d-collapsible then Y is homotopy equivalent to a (d-1) dimensional complex and it is acyclic: $H_d(Y) = 0$, no d-dimensional holes are preventing it from being deformed to a (d-1) dimensional shape. However, the other direction does not hold up for higher dimensions: an acyclic simplicial complex is not necessarily d-collapsible. The phase transition to acyclicity and the vanishing of collapsability do not share the same threshold for $d > 1$. However, interestingly, they both maintain their one-sided threshold.

We will now transition to explore higher dimensional parallels of full connectivity in $G(n,p)$. Let us expand on the usage of homology groups to quantify the number of connected components in a graph. The $\dim H_{d-1}(X,G)$ counts the number of connected components in a d-dimensional random simplicial complex. Intuitively, this also makes sense: holes are a measure of discontinuity. If a space is continuous it can be continuously deformed because all of its points are connected. Since graph connectivity is synonymous to the vanishing of the 0th-homology, it is very natural to suggest the vanishing of the $(d-1)$th-homology as an easy, abstract generalization of graph connectivity for higher dimensions. And the threshold for higher dimensions has a similarly intuitive threshold to its d =1 counterpart.

Theorem 4.14. *[MW09] Fix $d \geq 1$, and let $Y = Y_d(n,p)$. Let G be any finite abelian group. W.h.p $H_{d-1}(Y,G) = 0$ if*

$$p \geq \frac{d\log n + \omega(1)}{n},$$

and w.h.p $H_{d-1}(Y,G) \neq 0$ if

$$p \leq \frac{d\log n + \omega(1)}{n}.$$

Remark 4.15. A couple of words regarding the connectivity threshold:

(1) Important to note is the distinction in the scope of G: this theorem has not yet been proven beyond coefficients in a finite abelian group.
(2) Connectivity is fascinating because it has a very natural higher-dimensional counterpart, while as explored with giant component emergence, there has been much greater of a struggle to find a suitable d-dimensional counterpart. Homology groups prove to be a very natural bridge between properties in $d = 1$ and $d > 1$.
(3) Broadly, these properties in higher dimensions need to continue to provide valuable information regarding the structure of the simplicial complex just as their d=1 counterparts provide rich information regarding the graph—homology groups are a great example of such a meaningful property for topological objects.
(4) In higher dimensions, there is no longer great mystique around one specific point p, instead thresholds for phase transitions are more spread out across the range of p.

5. Acknowledgments

The author would like to thank Simon Rubinstein-Salzedo and Christina Yu for all the helpful (and endearingly endless) discussions.

References

[AL15] Lior Aronshtam and Nathan Linial. When does the top homology of a random simplicial complex vanish? *Random Structures & Algorithms*, 46(1):26–35, 2015.
[AL16] Lior Aronshtam and Nathan Linial. The threshold for d-collapsibility in random complexes. *Random Structures & Algorithms*, 48(2):260–269, 2016.
[ALŁM13] Lior Aronshtam, Nathan Linial, Tomasz Łuczak, and Roy Meshulam. Collapsibility and vanishing of top homology in random simplicial complexes. *Discrete & Computational Geometry*, 49(2):317–334, 2013.
[Bol84] Béla Bollobás. The evolution of random graphs. *Transactions of the American Mathematical Society*, 286(1):257–274, 1984.
[ER59] OSP ERD and A Renyi. On random graphs. *Publ. Math*, 6:290–297, 1959.
[Erd61] Paul Erdos. On the evolution of random graphs. *Bulletin of the Institute of International Statistics*, 38:343–347, 1961.
[FK23] Alan Frieze and Michał Karoński. *Random graphs and networks: a first course*. Cambridge University Press, 2023.

[LNPR14] Nati Linial, Ilan Newman, Yuval Peled, and Yuri Rabinovich. Extremal problems on shadows and hypercuts in simplicial complexes. *arXiv preprint arXiv:1408.0602*, 2014.

[LP16] Nathan Linial and Yuval Peled. On the phase transition in random simplicial complexes. *Annals of mathematics*, pages 745–773, 2016.

[Łuc90] Tomasz Łuczak. Component behavior near the critical point of the random graph process. *Random Structures & Algorithms*, 1(3):287–310, 1990.

[MW09] Roy Meshulam and Nathan Wallach. Homological connectivity of random k-dimensional complexes. *Random Structures & Algorithms*, 34(3):408–417, 2009.

[Pit88] Boris Pittel. A random graph with a subcritical number of edges. *Transactions of the American Mathematical Society*, 309(1):51–75, 1988.

[RNS+17] Michael W Reimann, Max Nolte, Martina Scolamiero, Katharine Turner, Rodrigo Perin, Giuseppe Chindemi, Paweł Dłotko, Ran Levi, Kathryn Hess, and Henry Markram. Cliques of neurons bound into cavities provide a missing link between structure and function. *Frontiers in computational neuroscience*, 11:266051, 2017.

[Roc24] Sebastien Roch. Moments and tails. In *Modern discrete probability: An essential toolkit*. Cambridge University Press, 2024.