RANDOM GRAPHS BY JOEL SPENCER

Notation 1. We say "$f(n) \gg g(n)$ as $n \to \infty$" if $\frac{f(x)}{g(x)} \to \infty$ as $n \to \infty$.

Notation 2. We say "$f(n) \sim g(n)$ as $n \to \infty$" if $\frac{f(n)}{g(n)} \to 1$ as $n \to \infty$

Notation 3. We say "$f(n) = o(g(n))$ as $n \to \infty$" if $\frac{f(n)}{g(n)} \to 0$ as $n \to \infty$

Notation 4. $(n)_k := n(n-1)(n-2)\ldots(n-(k-1))$ notice there are k terms in the product.

Definition 5. The *Erdos Renyi* random graph $G(n,p)$ is a graph on the vertex set $V = [n] = \{1,\ldots n\}$ where each of the $\binom{n}{2}$ possible connections between vertices is chosen independently with probability p, i.e. $\mathbf{P}(\{i,j\} \in G) = p$.

Remark 6. A nice way to couple $G(n,p)$ to $G(n,q)$ is to choose a uniform random variable at each possible edge and include it if $U < p$ in $G(n,p)$ and if $U < q$ in $G(n,q)$. This idea gives a graph *process* $G(n,t)$ for $t \in [0,1]$ which ranges from the empty graph at $t = 0$ to the complete graph at $t = 1$.

Remark 7. We will very often let $p = p(n)$ depend on n, and be looking at problems that happen in the limit $n \to \infty$.

0.1. **Threshold Functions.** We will see that for many properties of graphs A, there is a vary narrow window for values of p where $\mathbf{P}(G(n,p) \in A)$ ranges from 0 to 1. We make this more precise in this section. We start with an example to illustrate the idea.

Example 8. Let X be the number of triangles in $G(n,p)$. Each triangle of K_n has a probablity of $p^3$of being in $G(n,p)$. Since there are $\binom{n}{3}$ such triangles, we have by the linearity of the expectation that:

$$\mathbf{E}(X) = \binom{n}{3} p^3$$

This suggests a parametrization $p = p(n) = \frac{c}{n}$, so that $\lim_{n\to\infty} \mathbf{E}(X) = \frac{c^3}{6}$. We will see later that the distribution of X is a *Poisson* distribution. The above calculation identifies the rate for us! Look at the event $\{X = 0\}$ now. Have:

$$\lim_{n\to\infty} \mathbf{P}\{X = 0\} = e^{-c^3/6}$$

Notice that this probability ranges from 0 to 1 as c ranges from ∞ to 0. For example, if $c = 10^6$, $G(n,p)$ is very likely to have triangles for large n, while for $c = 10^{-6}$ $G(n,p)$ is very unlikely to have triangles. In the dynamic view (our graph process) this is saying that there is a sharp window around $p \sim \frac{1}{n}$ where $\mathbf{P}\{X = 0\}$ goes from 0 to 1.

Notice that if we take for example $p(n) = n^{-0.9} = \frac{n^{0.1}}{n}$ (notice $p(n) \gg n^{-1}$ as $n \to \infty$), then $\lim_{n\to\infty} \mathbf{P}\{X = 0\} = 0$ by a similar analysis! On the other hand if we take for example $p(n) = \frac{1}{n \log n}$ (so $p(n) \ll n^{-1}$ as $n \to \infty$), then $\lim_{n\to\infty} \mathbf{P}\{X = 0\} = 1$. This is the idea of a *threshold function* which we formally define below.\

Definition 9. $r(n)$ is called a *threshold* function for a graph property A if the following two conditions are met:

$$\begin{aligned} p(n) \ll r(n) \text{ as } n \to \infty \quad &\Rightarrow \quad \lim_{n\to\infty} \mathbf{P}\{G(n,p) \in A\} = 0 \\ p(n) \gg r(n) \text{ as } n \to \infty \quad &\Rightarrow \quad \lim_{n\to\infty} \mathbf{P}\{G(n,p) \in A\} = 1 \end{aligned}$$

Remark 10. In our above example about the number of triangles, $r(n) = \frac{1}{n}$ was a threshold function. Notice that threshold functions are not unique. For example multiplying by a constant is ok.

In the above example of triangles, we could not actually compute $\mathbf{P}\{X = 0\}$ because triangles in the graph are not independent (We instead stated a result we will prove later). In the next section we will develop some techniques to get around this kind of "non-independence"

0.2. **Variance.** Here we will develop some lemmas for non-negative integer valued random variables X that will be useful to us. The notation devolved here will be used throughout.

Lemma 11. *Let X be a non-negative integer valued random variable. Then:*

$$\mathbf{P}(X = 0) \leq \frac{\mathbf{Var}(X)}{\mathbf{E}(X)^2}$$

Proof. Let $\mu = \mathbf{E}(X)$, $\sigma = \sqrt{\mathbf{Var}(X)}$. Apply Chebyshev's Inequality with $\lambda = \frac{\mu}{\sigma}$. Have:

$$\mathbf{P}(X = 0) \leq \mathbf{P}(|X - \mu| \geq \lambda\sigma) \leq \frac{1}{\lambda^2} = \frac{\sigma^2}{\mu^2}$$

$\square$

We will usually apply this result when $X = X_n$ is a really a sequence of random variables. For example, here are some useful corallaries:

Corollary 12. *If* $\mathbf{Var}(X) = o\left(\mathbf{E}(X)^2\right)$ *as* $n \to \infty$ *then* $\mathbf{P}\left(X > 0\right) \to 1$ *as* $n \to \infty$.

Proof. Have:

$$\mathbf{P}(X > 0) = 1 - \mathbf{P}\left(X = 0\right) \geq 1 - \frac{\mathbf{Var}(X)}{\mathbf{E}(X)^2} \to 1 - 0 = 1$$

□

Corollary 13. *If* $\mathbf{Var}(X) = o\left(\mathbf{E}(X)^2\right)$ *as* $n \to \infty$ *then for all* $\epsilon > 0$, $\mathbf{P}\left(\left|\frac{X}{\mathbf{E}(X)} - 1\right| < \epsilon\right) \to 1$ *as* $n \to \infty$ *i.e.* $\frac{X}{\mathbf{E}(X)} \xrightarrow{\mathbf{P}} 1$. *(We abbreviate this as "$X \sim \mathbf{E}(X)$ a.s.")*

Proof. Following the idea of the original result:

$$\begin{aligned}\mathbf{P}\left(\left|\frac{X}{\mathbf{E}(X)} - 1\right| \geq \epsilon\right) &= \mathbf{P}\left(|X - \mathbf{E}(X)| \geq \epsilon \mathbf{E}(X)\right) \\ &\leq \frac{\mathbf{Var}(X)}{\epsilon^2 \mathbf{E}(X)^2} \to 0\end{aligned}$$

□

Definition 14. Take a sequence of events $A_1, A_2, \ldots, A_m$ and let their indicator functions be labeled $X_i = 1_{A_i}$. Let $X = X_1 + X_2 + \ldots + X_m$. For indices i, j write $i \sim j$ if $i \neq j$ and the events A_i and $A_j$are not independent. In this setup we define:

$$\Delta = \sum_{i \sim j} \mathbf{P}(A_i \cap A_j)$$

Proposition 15. *In the set up for* Δ *above, we have*$\mathbf{Var}(X) \leq \mathbf{E}(X) + \Delta$

Proof. Notice that when $i \sim j$ we have the inequality:

$$\mathbf{Cov}\left(X_i, X_j\right) = \mathbf{E}\left(X_i X_j\right) - \mathbf{E}\left(X_i\right)\mathbf{E}\left(X_j\right) \leq \mathbf{E}\left(X_i X_j\right) = \mathbf{P}\left(A_i \cap A_j\right)$$

When $i \neq j$and not $i \sim j$ then A_i, A_j are independent so we have $\mathbf{Cov}(X_i, X_j) = 0$. When $i = j$ $\mathbf{Cov}(X_i, X_j) = \mathbf{E}\left(X_i^2\right) - \mathbf{E}\left(X_i\right)^2 \leq \mathbf{E}(X_i)$ since X_i is an indicator random variable. Combining these we have:

$$\begin{aligned}\mathbf{Var}(X) &= \mathbf{Cov}(X, X) \\ &= \sum_{i,j} \mathbf{Cov}(X_i, X_j) \\ &= \sum_{i \sim j} \mathbf{Cov}(X_i, X_j) + \sum_{i \neq j, i \nsim j} \mathbf{Cov}(X_i, X_j) + \sum_{i = j} \mathbf{Cov}(X_i, X_j) \\ &\leq \Delta + 0 + \mathbf{E}(X)\end{aligned}$$

□

Corollary 16. *If* $\mathbf{E}(X) \to \infty$ *and* $\Delta = o(\mathbf{E}(X)^2)$,*then* $\lim_{n\to\infty} \mathbf{P}(X > 0) = 1$ *and* $\frac{X}{\mathbf{E}(X)} \xrightarrow{\mathbf{P}} 1$ *as* $n \to \infty$.

Proof. By the hypothesis and the last inequality, have $\mathbf{Var}(X) \leq \mathbf{E}(X) + \Delta = o(\mathbf{E}(X)^2)$ since $\mathbf{E}(X) \to \infty$ means that $\mathbf{E}(X) = o(\mathbf{E}(X)^2)$. By the previous corollaries then, we have $\frac{X}{\mathbf{E}(X)} \xrightarrow{\mathbf{P}} 1$ and $\lim_{n\to\infty} \mathbf{P}(X > 0) = 1$ □

This shows why the definition of Δ is useful. In cases where there is more symmetry, it sometimes makes even more sense to think about the following related quantity:

$$\Delta^\star = \sum_{j \sim i} \mathbf{P}\left(A_j | A_i\right)$$

In general this depends on i, but if there is an automorphism of the probablity space that interchanges A_i with A_j for $i \neq j$, then this will not depend on i. (The natural way this happens for random graph models is through a relabeling of the vertices) Notice that if this is indeed the case, and the A_i's are "interchangeable" in this sense, then:

$$\begin{aligned}\Delta &= \sum_{i \sim j} \mathbf{P}(A_i \cap A_j) \\ &= \sum_i \mathbf{P}(A_i) \sum_{j \sim i} \mathbf{P}(A_j | A_i) \\ &= \sum_i \mathbf{P}(A_i) \Delta^\star \\ &= \mathbf{E}(X)\Delta^\star\end{aligned}$$

Corollary 17. *In this symmetric set up, if* $\mathbf{E}(X) \to \infty$ *and* $\Delta^\star = o(\mathbf{E}(X))$ *then* $\lim_{n\to\infty} \mathbf{P}(X > 0)$ *and* $\frac{X}{\mathbf{E}(X)} \xrightarrow{\mathbf{P}} 1$.

Proof. By the above computatation, $\Delta^\star = o(\mathbf{E}(X))$ gives $\Delta = o\left(\mathbf{E}(X)^2\right)$ and we are reduced to the previous corollary. $\square$

0.3. **Appearance of Small Subgraphs.** In this section we will use the tools developed above to find the threshold function for the appearance of a given graph H to be a subgraph of $G(n,p)$. We start with an instructive example, and then we will generalize it.

Definition 18. A set $S \subset G$ is called a *clique* if every pair of vertices, $x, y \in S$ in adjacent in G. Let $\omega(G)$ be the size of the largest clique in G. Notice that $\omega(G)$ is the size of the largest complete graph we can find as a subgraph G.

Theorem 19. *Let A be the event $A = \{\omega(G) \geq 4\} = \{G$ contains a copy of $K_4\}$. Then A has a threshold function of $n^{-2/3}$.*

Proof. For every set of 4 vertices in $G(n,p)$, let $A_S$be the event "S is a clique" and X_S its indicator random variable. Then:

$$\mathbf{E}(X_S) = \mathbf{P}(A_S) = p^6$$

as there are six different possible edges between the vertices of S. Define:

$$X = \sum_{|S|=4} X_S$$

so that X is the number of 4-cliques in G and $\omega(G) \geq 4$ if and only if $X > 0$. (This is exactly the set up we had in our discussion above for X.) By linearity of the expectation, we know that:

$$\mathbf{E}(X) = \sum_{|S|=4} \mathbf{E}(X_S) = \binom{n}{4} p^6 \sim \frac{n^4 p^6}{24}$$

This simple calculation already proves *half* of the theorem. Namely if $p(n) \ll n^{-2/3}$, then $\lim_{n\to\infty} \mathbf{E}(X) = 0$ and so $\lim_{n\to\infty} \mathbf{P}(X > 0) \to 0$ too.

To see the other half (namely when $p(n) \gg n^{-2/3}$ the probability approaches 1), we use the techniques we worked on above, in particular the last result about $\Delta^\star$. (We are in the set up of $\Delta^\star$here as all the 4-sets S are the same up to relabeling the vertices). Suppose $p(n) \gg n^{-2/3}$. For two four sets S, T we know that $S \sim T$ if and only if $S \neq T$ and S and T share at least one possible edge. This only happens if they share at least two vertices. Fix S. There are $\binom{n-4}{2} = O(n^2)$ with $|S \cap T| = 2$ (two vertices in common, one edge in common) and for each of these $\mathbf{P}(A_T | A_S) = p^5$ (the edge in common comes for free, 5 edges remain). There are $\binom{n-4}{1} = O(n)$ sets T with $|S \cap T| = 3$ (three vertices in common, three edges n common) and for each of these $\mathbf{P}(A_T | A_S) = p^3$ (three of the edges come for free). Have then that:

$$\Delta^\star = \sum_{S \sim T} \mathbf{P}\left(A_Y | A_S\right) = O(n^2 p^5) + O(np^3) = o(n^4 p^6) = o\left(\mathbf{E}(X)\right)$$

The last line works since $p(n) \gg n^{-2/3}$. Notice also that $\mathbf{E}(X) \to \infty$ because of $p \gg n^{-2/3}$. By the corollary above, we know that $\lim_{n\to\infty} \mathbf{P}(X > 0) = 1$ which proves the other half of the threshold function property.

We can generalize this a little bit to other subgraphs H where the same type of calculation works out. $\square$

Definition 20. Let H be a graph with v vertices and e edges. We call $\rho(H) = e/v$ the *density* of H. We call H *balanced* if every subgraph $H'$of H has $\rho(H') \leq \rho(H)$. We call H *strictly balanced* if every proper subgraph H'has $\rho(H) < \rho(H)$.

Example 21. K_4 the complete graph on 4 vertices is strictly balanced. This is part of why the above calculation works. In general K_n is completly balanced.

Theorem 22. *Let H be a balanced graph with v vertcies and e edges. Let $A(G)$ be the even that H is a subgraph of G.Then $p(n) = n^{-v/e}$is the threshold function for A.*

Proof. We follow the argument we made above that was for the case $H = K_4$. For each $v-$set S let $A_S$be the even that G_S contains H as a subgraph. Since H is somewhere between the complete graph on v edges we have that:

$$p^e \leq \mathbf{P}(A_S) \leq v! p^e$$

Any particular placement of H has probability $p^e$of occuring and there are at most $v!$ possible placements. The precise calcultion of $\mathbf{P}(A_S)$ is in general complicated due to the overlapping of potential copies of H, but the above naive inequality will be enough for our purposes. Let X_S be the indicator random variable for A_S and define $X = \sum_{|S|=v} X_S$ so that A holds if and only if $X > 0$. By linearity of the expectation we have:

$$\mathbf{E}(X) = \sum_{|S|=v} \mathbf{E}(X_S) = \binom{n}{v} \mathbf{P}(A_S) = \Theta(n^v p^e)$$

As before, this gives quicky calculation gives one side of the threshold function result: if $p(n) \ll n^{-v/e}$ then we see from the above that $\lim_{n\to\infty} \mathbf{E}(X) = 0$ and so $\lim_{n\to\infty} \mathbf{P}(X > 0) = 0$ too.

Now assume $p(n) \gg n^{-v/e}$ so that $\mathbf{E}(X) \to \infty$ from the above calculation. Let us estimate $\Delta^\star$as we did before. For a fixed $v-$set S, we have that $S \sim T$ if and only if $S \neq T$ and S, T have some edge sin common. We divide up the sum for $\Delta^\star$ by the number of vertices S and T have in common:

$$\Delta^\star = \sum_{T \sim S} \mathbf{P}(A_T | A_S) = \sum_{i=2}^{v-1} \sum_{|T \cap S| = i} \mathbf{P}(A_T | A_S)$$

For each value of i, there are $\binom{n-v}{v-i} = O(n^{v-i})$ sets T with $|T \cap S| = i$. For fixed T, consider $\mathbf{P}(A_T | A_S)$. There are $O(1)$ possible copies of H on T (some number between 1 and $v!$ as mentioned before). Each has at most $\frac{ie}{v}$ edges with both vertices in S and thus at least $e - \frac{ie}{v}$ other edges (this fact depends critically on the fact that H is balanced, to get the upper bound $\frac{ie}{v}$, look at a subgraph of that copy of H, namely $S \cap T \cap H$. Since the density of this graph is less than $\frac{e}{v}$ and this graph has i vertices, the number of edges in this subgraph is less than $\frac{ie}{v}$). We have then that:

$$\mathbf{P}(A_T | A_S) = O(p^{e - \frac{ie}{v}})$$

So we get that:

$$\Delta^\star = \sum_{i=2}^{v-1} O(n^{v-i} p^{e - \frac{ie}{v}}) = \sum_{i=2}^{v-1} O\left((n^v p^e)^{1 - \frac{i}{v}} \right)$$

Hence, since $p(n) \gg n^{-v/e}$ we know that:

$$\Delta^\star = \sum_{i=2}^{v-1} o(n^v p^e) = o\left(\mathbf{E}(X)\right)$$

So we now have the hypothesis of the earlier corollary which gives the result. □

Theorem 23. *If H is not balanced, then $p = n^{-v/e}$is* not *the threshold function for A.*

Proof. Let H_1 be a subgraph of H with $v_1$vertices, $e_1$edges and $e_1/v_1 > e/v$. Take α so that $v/e < \alpha < v_1/e_1$ and set $p(n) = n^{-\alpha}$ so that $p \gg n^{-v/e}$ and $p \gg n^{-v_1/e_1}$. By our earlier proof, since $\alpha < e_1/v_1$ then $\mathbf{E}(\# H_1$ a subgraph of $G(n,p)) \to 0$ (this was in the easy part of the proof, we didn't even need H_1 is balanced) Hence $\mathbf{P}(H_1$a subgraph of $G(n,p)) \to 0$. If there are no copies of H_1 there are definitely no copies of H, so $\mathbf{P}(H$ is a subgraph$) \to 0$ too. Since $p \gg n^{-v/e}$, and yet $\mathbf{P}(H$ a subgraph$) \to 0$ this shows that $n^{-v/e}$ cannot be a threshold function for A. □

Remark 24. Erdos and Renyi proved that the threshold function for containg a subgraph of H for arbitrary H is n^{-v_1/e_1} where v_1/e_1 is the density of the subgraph $H_1 \subset H$ of *maximal* density. The methods we have done so far actually get us really close to this result. The next result is a slight improvement of the hard half of the threshould function result we proved earlier.

Theorem 25. *Let H be a strictly balanced graph with v vertices and e edges. Suppose that there are a automorphisms of H. Let X be the number of copies of H in $G(n,p)$. If $p \gg n^{-v/e}$ then we have that $X \sim \frac{n^{-v/e}}{a}$ almost always (in the sense that the ratio $\xrightarrow{\mathbf{P}} 1$).*

Proof. The proof is very similar to the proof of $n^{-v/e}$as a threshold function. This time we use a slightly more precise counting of the number of ways H can be put into the set $S$with $|S| = v$. We also use the conclusion that $X \sim \mathbf{E}(X)$ rather than just $X > 0$ a.a. from the corollary at the end. We fill in the details now.

Label the vertices of H by $1, \ldots, v$. For each ordered set of vertices $x_1, x_2, \ldots, x_v$ let $A_{x_1, \ldots, x_v}$be the even that $x_1, \ldots, x_v$ provides a copy of H in that order. (To be precise this is the event that $\{x_i, x_j\} \in E(G)$ for every $\{i, j\} \in E(H)$.) Let $I_{x_1, \ldots x_v}$be the corresponding indicator random variable. Notice that $\mathbf{P}(A_{x_1, \ldots x_v}) = \mathbf{E}(I_{x_1, \ldots, x_v}) = p^e$ since there are e edges in H.

Now, we define an *equivalence relation* on ordered v-tuples by $(x_1, \ldots, x_v) \equiv (y_1, \ldots y_v)$ if there is an automorphism $\sigma$of $V(H)$ so that $y_{\sigma(i)} = x_i$. (By the hypothesis of the problem, there are a elements in each equivalence class). Define now:,

$$X = \sum I_{x_1, ., x_v}$$

Where the sum is taken over one entry from each equivalence class. There are $(n)_v$ possible ordered $v-$tuples, and so there are $(n)_v / a$ equivalence classes.We have then by linearity of $\mathbf{E}$ that:

$$\mathbf{E}(X) = \frac{(n)_v}{a} \mathbf{E}(I_{x_1, \ldots, x_v}) \sim \frac{n^v p^e}{a}$$

Our assumption that $p(n) \gg n^{-v/e}$ tells us that $\mathbf{E}(X) \to \infty$ so it remains only to show that $\Delta^\star = o(\mathbf{E}(X))$ and then we can apply the corollary and get the result. Fix $(x_1, \ldots, x_v)$ and consider:

$$\Delta^\star = \sum_{(y_1, \ldots, y_v) \sim (x_1, \ldots, x_v)} \mathbf{P}(A_{y_1, \ldots, y_v} | A_{x_1, \ldots, x_v})$$

The tuples where $(y_1, \ldots, y_v) \sim (x_1, \ldots, x_v)$ can be divided into v disjoint pieces, where $|\{y_1, \ldots, y_v\} \cap \{x, \ldots, x_v\}| = i$ for $i = 1, \ldots v$. The arguments of the last section are now used to consider each piece seperatly and see that each contributes $o(\mathbf{E}(X))$ to the sum. Hence $\Delta^\star = o(\mathbf{E}(X))$ too, as desired. We now conclude that $X \sim \mathbf{E}(X)$ a.a. □

Theorem 26. *Let H be any fixed graph. For every subgraph H' of H (including H itself) let $X_{H'}$ denote the number of copies of H' in $G(n,p)$. If $p(n)$ is so that $\mathbf{E}(X_{H'}) \to \infty$ for every H', then almost always:*

$$X_H \sim \mathbf{E}(X_H)$$

Proof. Say H has v vertices and e edges. Set up $A_{x_1,\ldots x_v}, I_{x_1,\ldots,x_v}$ as in the above theorem. As in the above theorem, the difficulty lies in showing that $\Delta^\star = o(\mathbf{E}(X))$. We split the sum defining $\Delta^\star$into a finite number of terms as follows. For every $H'$with w vertices and f edges, we consider those $v-$tuples $(y_1,\ldots,y_v)$ that overlap with the fixed $(x_1,\ldots,x_v)$ in a copy of H' (that is the overlap of the two tuples has exactly w edges and there are f edges that "come for free" because we are conditioning on $A_{x_1,.,x_v}$) For such $y_1,\ldots y_v$ we have that:

$$\mathbf{P}(A_{y_1,\ldots,y_v}|A_{x_1,\ldots,x_v}) = p^{e-f}$$

There are $\sim n^{v-w}$ such $v-$tuples (since H'has w vertices) so the total contribution from this piece is:

$$n^{v-w}p^{e-f} = \Theta\left(\frac{\mathbf{E}[X_H]}{\mathbf{E}[X_{H'}]}\right) = o(\mathbf{E}(X_H))$$

Since $\mathbf{E}(X_{H'}) \to \infty$ by hypothesis. Summing over all these finitely many terms, we arrive at the conclusion that $\Delta^\star = o(\mathbf{E}(X_H))$ too. $\square$

0.4. **Connectivity.** In this section we give a relativly simple example of the Poisson Paradigm: the rough notion that if ther eare many rare and independent events, then the number of events has approximatly a Poisson distribution. This applies to the connectivity of an Erdos Renyi graph, as we will see.

Definition 27. A vertex v is called *isolated* if it is adjacent to no $w \in V$. In $G(n,p)$ let X count the numebr of isolated vertices.

Theorem 28. *Let $p = p(n)$ satisfy $n(1-p)^{n-1} = \mu$. Then:*

$$\lim_{n\to\infty} \mathbf{P}(X=0) = e^{-\mu}$$

Proof. Let $X_i$be the indicator random variable that the vertex i is isolated, so that $X = \sum_{i=1}^n X_i$. Notice that $\mathbf{P}(X_i) = (1-p)^{n-1}$ for every i, so by linearity and symmetry of the problem, $\mathbf{E}(X) = n\mathbf{E}(X_i) = n(1-p)^{n-1} = \mu$. Now consider the r-th factorial moment, $\mathbf{E}((X)_r) = \prod_{i=0}^{r-1}(X-i)$. By symetry, we have that: $\mathbf{E}(X) = (n)_r\mathbf{E}(X_1X_2\cdots X_r)$. The product $X_1X_2\cdots X_r$ is 1 if all of the vertices $1,\ldots,r$ are isolated, and 0 otherwise. In order for all these to be isolated, we need each of the $n-1$ edges coming from each of the r vertices to be not-selected. However, this double counts the $\binom{r}{2}$ edges amongst the vertices, so the total number of edges that we need to be not-selected is $r(n-1) - \binom{r}{2}$. Hence $\mathbf{E}(X_1X_2\ldots X_r) = \mathbf{P}(X_1X_2\ldots X_r = 1) = (1-p)^{r(n-1)-\binom{r}{2}}$. Have then:

$$\mathbf{E}((X)_r) = (n)_r(1-p)^{r(n-1)-\binom{r}{2}} \sim n^r(1-p)^{r(n-1)} = \mu^r \text{ as } n\to\infty$$

(This is saying that the dependece among the X_i was asymptotically negligible.) These are the same moments as for a Poisson distribution! We may now cite a fact that $\lim_{n\to\infty}\mathbf{E}((X)_r) = \mu^r$ for every r implies convergence in distribution to a Poisson random variable, $X \Rightarrow Poiss(\mu)$. In our case, where we only want to show $\lim_{n\to\infty}\mathbf{P}(X=0) \to e^{-\mu}$ it is not much more work to see it directly. The main idea is from the proof of the above convergence fact, namely that:

$$\mathbf{P}(X=k) = \sum_{r=0}^{\infty}(-1)^r\frac{\mathbf{E}[(X)_r]}{(r-k)!k!}$$

(This identity is proven by writing $1_{\{X=k\}} = \binom{X}{k}(1-1)^{X-k} = \binom{X}{k}\sum_{i=0}^{X-k}(-1)^i\binom{X-k}{i} = \sum_{i=0}^{\infty}(-1)^i\binom{X}{k}\binom{X-k}{i}$. Exapanding the binomial coefficients and taking expectation proves the above.) Right now we are interested only in $\lim_{n\to\infty}\mathbf{P}(X=0)$:

$$\begin{aligned}\lim_{n\to\infty}\mathbf{P}(X=0) &= \lim_{n\to\infty}\sum_{r=0}^{\infty}(-1)^r\frac{\mathbf{E}[(X)_r]}{r!} \\ &= \sum_{r=0}^{\infty}(-1)^r\frac{\lim_{n\to\infty}\mathbf{E}[(X)_r]}{r!} \\ &= \sum_{r=0}^{\infty}(-1)^r\frac{\mu^r}{r!} \\ &= e^{-\mu}\end{aligned}$$

(Not sure why exactly we can change the limit with the sum here...I can't get the estimates to work out for a LDCT swap, I think you might have to do it more "hands-on" using the alternating series test) $\square$

Theorem 29. *[Erdos-Renyi "Double Exponential"] Suppose $p = p(n)$ is given by:*

$$p(n) = \frac{\log n}{n} + \frac{c}{n} + o\left(\frac{1}{n}\right)$$

Then:

$$\lim_{n\to\infty}\mathbf{P}\left(G(n,p) \text{ is connected}\right) = e^{-e^{-c}}$$

Proof. For such p, we have that $n(1-p)^{n-1} \sim \mu = e^{-c}$. By the previous theorem, the probability that $X = 0$ (i.e. G has no isolated component) approaches $e^{-\mu} = e^{-e^{-c}}$. We will now argue that the probability that G is connected is asymptotically equal to the probability that G has no isolated components. Indeed, if G has no isolated vertices but is not connnected, then there is a connected component of k vertices for $2 \leq k \leq \frac{n}{2}$. Let B be the this event, the we have:

$$\mathbf{P}(B) \leq \sum_{k=2}^{n/2} \binom{n}{k} k^{k-2} p^{k-1} (1-p)^{k(n-1)-\binom{k}{2}}$$

The sum is over the size of the connected component.The first factor is the number of choices for the componet of size $|S| = k$, $S \subset V(G)$. The second factor is the choice of a tree on S (there are k^{k-2}such trees, and this provides an upper bound, because every connected component has a spanning tree). The third factor is the probability that the tree's edges are in $E(G)$ (every tree with k vertices has $k-1$ edges). Finally, the last factor is the probability that the tree is isolated with no edges from the tree to outside of the tree (similar to the calculation of $(X)_r$ in the above theorem). Some analysis here shows that $k = 2$ is the dominating term and that $\mathbf{P}(B) = o(1)$. This verifies indeed that $\mathbf{P}(G(n,p)$ is connected$)$ is the same as $\mathbf{P}(X = 0)$ in the limit $n \to \infty$. □

0.5. **The Janson Inequalities.** In many cases we would like to bound the probability that none of a set of bad events B_i $i \in I$ occur. If the events are independent, then we would have:

$$\mathbf{P}\left(\bigcap_{i\in I} \bar{B}_i\right) = \prod_{i\in I} \mathbf{P}\left(\bar{B}_i\right)$$

When the B_i are not independent but are "mostly independent" in some sense, the Janson inequalities allow us to say that the two quantities above are "mostly equal". Of course, we have to make this precise. We first describe the set up in which the Janson inequality will apply, and then we state and prove the inequality as a theorem. The set up is similar to our framework for Δ and $\Delta^\star$ from the last few sections, but not exactly the same.

Let Ω be a finite universal set and let R be a random subset of Ω defined in such a way so that:

$$\mathbf{P}\left(r \in R\right) = p_r$$

and moreover, so that $r \in R$ is independent of $s \in R$ for $r \neq s$. (For our purposes, we will be using $\Omega = E(K_n)$, the set of possible edges in the graph $G(n,p)$, and we have $R = E(G(n,p))$ so that $p_r = p$ is constant for us.) Let A_i,$i \in I$ be subsets of Ω for a finite set I (e.g. a subset of possible edges). Let B_i be the *event* $A_i \subset R$ (e.g. $\mathbf{P}(B_i) = \prod_{r\in A_i} p_r$). Let $X_i$be the indicator function for $B_i$and let $X = \sum_{i\in I} X_i$, the number of $A_i \subset R$. Notice that $\{X = 0\} = \cap_{i\in I}\bar{B}_i$ is the event that $A_i \not\subseteq R$ for every i.

Let us write $i \sim j$ whenever $i \neq j$ and $A_i \cap A_j \neq \emptyset$. Notice that when $i \sim j$ the two events are $B_i$and B_j are *dependent* $A_i$and A_j have some overlap, while if $i \nsim j$ and $i \neq j$ then B_i and B_j are *independent* since A_i and A_j have no overlap. Even further, if we have a whole collection of indices $J \subset I$ and if $i \notin J$ and $i \sim j$ for all $j \in J$ then B_i is mutually independent of $\{B_j | j \in J\}$ (i.e. indpendent of any combination of unions/intersections of these events.) As before, this happens since the event B_i depends only on $\{r \in R\}$ for $r \in A_i$ while any event from the family $\{B_j | j \in J\}$ can only depend on $\{r \in R\}$for $r \in \cup_{j\in J} A_j$. (We will use this type of fact later on) Since these are disjoint, by the definition of the random set R these are independent.

Define now:

$$\Delta = \sum_{i\sim j} \mathbf{P}\left(B_i \cap B_j\right)$$

Here the sum is over unordered pairs (i,j) so we notice that $\Delta/2$ is the same sum but over ordered pairs (i,j) by symmetry. (This is very close to the Δ defined earlier). Define for convenience now:

$$M = \prod_{i\in I} \mathbf{P}\left(\bar{B}_i\right)$$

If all the B_i's were independent, then all the $A_i's$ would be mutually disjoint and Δwould be the empty sum, and we would have $M = \prod\left(\cap_{i\in I}\bar{B}_i\right)$ by independence. When the B_i's are not independent, $\Delta > 0$ in general and $M < \mathbf{P}(\cap_{i\in I}\bar{B}_i)$ in general. The Janson inequality gives an upper bound to control $\mathbf{P}(\cap_{i\in I}\bar{B}_i)$ in terms of M and Δ. In practical problems, we calculate/estimate Δ(just as we have been doing until now) and then the Janson inequality gives a bound on $\mathbf{P}(\cap_{i\in I}\bar{B}_i)$ which we hope is useful to us.

Theorem 30. *[The Janson Inequality] Let $B_i, i \in I$, Δ,M be as above and assume that there is an $\epsilon > 0$ so that $\mathbf{P}(B_i) \leq \epsilon$ for each $i \in I$. Then:*

$$M \leq \mathbf{P}\left(\cap_{i\in I}\bar{B}_i\right) \leq M \exp\left(\frac{1}{1-\epsilon}\frac{\Delta}{2}\right)$$

Corollary 31. *In the above set up, let $\mu = \mathbf{E}(X) = \sum_{i\in I} \mathbf{P}(B_i)$. Then notice that $\mathbf{P}(\bar{B}_i) = 1 - \mathbf{P}(B_i) \leq \exp(-\mathbf{P}(B_i))$, so then, multiplying over $i \in I$ we have that $M = \prod_{i\in I} \mathbf{P}\left(\bar{B}_i\right) \leq \exp(-\mu)$. Hence the Janson inequality gives:*

$$\exp(-\mu) \leq \mathbf{P}\left(\cap_{i\in I}\bar{B}_i\right) \leq \exp\left(-\mu + \frac{1}{1-\epsilon}\frac{\Delta}{2}\right)$$

Example 32. Here is an example of the Janson inequality from a problem we did before with our less refined corollary (the $\Delta = o(\mathbf{E}(X))$ one). Say $p(n) = cn^{-2/3}$ and we are interested in the probability that $G(n,p)$ contains no K_4. Let B_i $1 \le i \le \binom{n}{4}$ range over all the possible K_4's in the graph $G(n,p)$. Each is a 6 element subset of $\Omega = \{\text{possible edges}\}$. As we calculated before, we have $\mathbf{P}(B_i) = p^6$ so that we can take $\epsilon = o(1)$. As we calculated before, $p(n) = cn^{-2/3}$ is chosen just right so that $\Delta = o(1)$ too, and we have that $\mu = \mathbf{E}(X) = \frac{c^6}{24}$. Hence the Janson inequality gives a perfect sandwhich from which we conclude that $\mathbf{P}(\text{no copy of } K_4) = \mathbf{P}\left(\cap_{i\in I}\bar{B}_i\right) \to \exp(-\mu) = \exp\left(-\frac{c^6}{24}\right)$

The above shows how well the Janson inequality can work if $\Delta = o(1)$. For large Δ , the Janson inequality becomes less precise. Indeed, when $\Delta \ge 2\mu(1-\epsilon)$ the Janson inequality upper bound is greater than 1; a useless upper bound. In these circumstances, the following generalized result is more useful.

Theorem 33. *[Generalized Janson Inequality] In the same set up as above, and still assuming that there is an $\epsilon > 0$ so that $\mathbf{P}(B_i) \le \epsilon$ for each $i \in I$, assume also that $\Delta \ge \mu(1-\epsilon)$. Then:*

$$\mathbf{P}\left(\cap_{i\in I}\bar{B}_i\right) \le \exp\left(-\frac{\mu^2(1-\epsilon)}{2\Delta}\right)$$

Remark 34. This generalized inequality really is much better than what we have been using so far (which ultimately boiled down to a Chebyshev inequality). For example, in our earlier framework, Chebyshev gave us the bound:

$$\mathbf{P}\left(\cap_{i\in I}\bar{B}_i\right) = \mathbf{P}\left(X=0\right) \le \frac{\mathbf{Var}(X)}{\mathbf{E}(X)^2} \le \frac{\mu+\Delta}{\mu^2}$$

For $\mu \ll \Delta$, this bound is $\mathbf{P}(X=0) \le \frac{\mu^2}{\Delta}$. Compare this to the generalized Janson inequality, $\mathbf{P}(X=0) \le \exp\left(-\frac{\mu^2}{\Delta}C\right)$. This is a much improved bound!

0.6. **The Proofs.** The original proofs are based on estimates using Laplace transforms. Instead of doing that, we will do the proof following Boppana and Spencer [1989]. We will use the inequalities:

$$\mathbf{P}\left(B_i \middle| \cap_{j\in J}\bar{B}_j\right) \le \mathbf{P}\left(B_i\right)$$

whenever $J \subset I$ and $i \notin J$. (This is clear because of the non-negative correlation between the events B_i and B_j. To see this, just write out the probabilities of $\mathbf{P}(B_i \cap B_j), \mathbf{P}(B_i)\mathbf{P}(B_j)$ in terms of events $r \in R$ and probabilities p_r to get $\mathbf{P}(B_i \cap B_j) = \mathbf{P}(A_i \cup A_j \in R) = \frac{\mathbf{P}(A_i\in R)\mathbf{P}(A_j\in R)}{\mathbf{P}(A_i\cap A_j\in R)} \ge \mathbf{P}(A_i \in R)\mathbf{P}(A_j \in R) = \mathbf{P}(B_i)\mathbf{P}(B_j)$)

We will also use:

$$\mathbf{P}\left(B_i \middle| B_k \cap \bigcap_{j\in J}\bar{B}_j\right) \le \mathbf{P}\left(B_i \middle| B_k\right)$$

This is really the same as the above inequality: we can view conditioning on B_k as the same as setting $p_r = 1$ for each $r \in A_k$. We are now in a position to prove Janson's inequality.

Theorem 35. *[Recap of the Janson Inequality] In our setup, let $B_i, i \in I$, $\Delta = \sum_{i\sim j}\mathbf{P}\left(B_i \cap B_j\right)$; $M = \prod_{i\in I}\mathbf{P}\left(\bar{B}_i\right)$ and assume that there is an $\epsilon > 0$ so that $\mathbf{P}(B_i) \le \epsilon$ for each $i \in I$. Then:*

$$M \le \mathbf{P}\left(\cap_{i\in I}\bar{B}_i\right) \le M\exp\left(\frac{1}{1-\epsilon}\frac{\Delta}{2}\right)$$

Proof. To get the lower bound, just apply the fact that the B_i's are positivly correlated directly (the first inequality above). Say WOLOG that $I = \{1,\ldots,m\}$, so then:

$$\mathbf{P}\left(B_i \middle| \bigcap_{1\le j\le i}\bar{B}_j\right) \le \mathbf{P}\left(B_i\right)$$

So then:

$$\mathbf{P}\left(\bar{B}_i \middle| \bigcap_{1\le j\le i}\bar{B}_j\right) \ge \mathbf{P}\left(\bar{B}_i\right)$$

And so, using a "telescoping product" we write:

$$\mathbf{P}\left(\bigcap_{1\le j\le m}\bar{B}_i\right) = \prod_{i=1}^{m}\mathbf{P}\left(\bar{B}_i \middle| \bigcap_{1\le j\le i-1}\bar{B}_j\right) \ge \prod_{i=1}^{m}\mathbf{P}\left(\bar{B}_i\right) = M$$

Now we work on the upper bound. For a given i, we will be interested in the connection between B_i and B_j for those indices $j < i$ (we are setting up for another telescoping product at the end, just like above). Some of the indices $j < i$ have $i \sim j$ and some have $i \nsim j$. For convenience, relabel the vertices so that $i \sim j$ for $1 \le j \le d$ and $i \nsim j$ for $d+1 \le j \le i-1$.

Now, we use the inequality $\mathbf{P}\left(A\,|B\cap C\right)\leq\mathbf{P}\left(A\cap B\,|C\right)$ which is valid for any A,B,C. If we let $A=B_i$, let$B=\bar{B}_1\cap\bar{B}_2\cap\ldots\cap\bar{B}_d$ and let $C=\bar{B}_{d+11}\cap\bar{B}_{d+2}\cap\ldots\cap\bar{B}_{i-1}$, then we have:

$$\mathbf{P}\left(B_i\left|\bigcap_{1\leq j\leq i-1}\bar{B}_j\right.\right)=\mathbf{P}\left(A\,|B\cap C\right)\geq\mathbf{P}\left(A\cap B\,|C\right)=\mathbf{P}(A\,|C)\mathbf{P}\left(B\,|A\cap C\right)$$

From the definition of C and since $i\nsim j$ for $d+1\leq j\leq i-1$ we know that A and C are independent. Hence $\mathbf{P}(A\,|C)=\mathbf{P}(A)$. To handle $\mathbf{P}\left(B\,|A\cap C\right)$, we do the bound:

$$\mathbf{P}\left(B\,|A\cap C\right)\geq 1-\sum_{j=1}^{d}\mathbf{P}\left(B_j\,|B_i\cap C\right)\geq 1-\sum_{j=1}^{d}\mathbf{P}\left(B_j\,|B_i\right)$$

The last bound follows again because the B_i's all have non-negative correlation. Multiplying through by $\mathbf{P}(B_i)=\mathbf{P}(A)$, this is (actually a bit more manipulation is needed again here, I think we use the correlation again):

$$\mathbf{P}\left(B_i\left|\bigcap_{1\leq j\leq i-1}\bar{B}_j\right.\right)\geq\mathbf{P}\left(B_i\right)-\sum_{j=1}^{d}\mathbf{P}\left(B_j\cap B_i\right)$$

Reversing this, and using $\mathbf{P}\left(\bar{B}_i\right)\geq 1-\epsilon$ gives:

$$\begin{aligned}\mathbf{P}\left(\bar{B}_i\left|\bigcap_{1\leq j\leq i-1}\bar{B}_j\right.\right) &\leq \mathbf{P}\left(\bar{B}_i\right)+\sum_{j=1}^{d}\mathbf{P}\left(B_j\cap B_i\right)\\ &\leq \mathbf{P}\left(\bar{B}_i\right)\left(1+\frac{1}{1-\epsilon}\sum_{j=1}^{d}\mathbf{P}\left(B_j\cap B_i\right)\right)\end{aligned}$$

Now use the inequality $1+x\leq\exp(x)$ to get:

$$\mathbf{P}\left(\bar{B}_i\left|\bigcap_{1\leq j\leq m-1}\bar{B}_j\right.\right)\leq\mathbf{P}\left(\bar{B}_i\right)\exp\left(\frac{1}{1-\epsilon}\sum_{j=1}^{d}\mathbf{P}\left(B_j\cap B_i\right)\right)$$

Finally, use this inequality in our "telescoping product" formula:

$$\begin{aligned}\mathbf{P}\left(\bigcap_{1\leq j\leq m}\bar{B}_i\right) &= \prod_{i=1}^{m}\mathbf{P}\left(\bar{B}_i\left|\bigcap_{1\leq j\leq i-1}\bar{B}_j\right.\right)\\ &\leq \prod_{i=1}^{m}\mathbf{P}\left(\bar{B}_i\right)\exp\left(\frac{1}{1-\epsilon}\sum_{j=1}^{d}\mathbf{P}\left(B_j\cap B_i\right)\right)\\ &= \left(\prod_{i=1}^{m}\mathbf{P}\left(\bar{B}_i\right)\right)\exp\left(\frac{1}{1-\epsilon}\sum_{i=1}^{m}\sum_{j=1}^{d_i}\mathbf{P}\left(B_j\cap B_i\right)\right)\end{aligned}$$

This is exactly the inequality we wanted, since $M=\prod_{i=1}^{m}\mathbf{P}\left(\bar{B}_i\right)$ and the double sum in the exponent sums over each *ordered* pair $(i,j),j<i$ with $i\sim j$ once, so its exactly equal to $\frac{\Delta}{2}$. ☐

Theorem 36. *[Generalized Janson Inequality] In the same set up as above, and still assuming that there is an $\epsilon>0$ so that $\mathbf{P}(B_i)\leq\epsilon$ for each $i\in I$, assume also that $\Delta\geq\mu(1-\epsilon)$. Then:*

$$\mathbf{P}\left(\cap_{i\in I}\bar{B}_i\right)\leq\exp\left(-\frac{\mu^2(1-\epsilon)}{2\Delta}\right)$$

Proof. As discussed early, the Janson inequality is often written as:

$$\mathbf{P}\left(\bigcap_{i\in I}\bar{B}_i\right)\leq\exp\left(-\mu+\frac{1}{1-\epsilon}\frac{\Delta}{2}\right)$$

Where $\mu=\mathbf{E}(X)=\sum_{i\in I}\mathbf{P}(B_i)$. Taking logarithms from this gives:

$$-\log\left(\mathbf{P}\left(\bigcap_{i\in I}\bar{B}_i\right)\right)\geq\sum_{i\in I}\mathbf{P}(B_i)-\frac{1}{2(1-\epsilon)}\sum_{i\sim j}\mathbf{P}\left(B_i\cap B_j\right)$$

The same inequality holds for a subset $S\subset I$. This is:

$$-\log\left(\mathbf{P}\left(\bigcap_{i\in S}\bar{B}_i\right)\right)\geq\sum_{i\in S}\mathbf{P}(B_i)-\frac{1}{2(1-\epsilon)}\sum_{i,j\in S,i\sim j}\mathbf{P}\left(B_i\cap B_j\right)$$

Now we will do something very sneaky, and we will look at *random* subsets $S \subset I$. This will let us use probabilistic methods to prove a probability theorem....very sneaky! Suppse $S \subset I$ is a random subset chosen so that $\mathbf{P}(i \in S) = p$ and the events $\{i \in I\}$ are all mutually independent (We will choose p more precisely later on) Now that S is random, we see that since above inequality holds for every S, it holds when we take $\mathbf{E}$ over the random subset S (we will denote this as $\mathbf{E}_S$ to keep it separate from the other $\mathbf{P}$'s going on). Have:

$$\mathbf{E}_S\left[-\log\left(\mathbf{P}\left(\bigcap_{i\in S}\bar{B}_i\right)\right)\right] \geq \mathbf{E}_S\left[\sum_{i\in S}\mathbf{P}(B_i)\right] - \frac{1}{2(1-\epsilon)}\mathbf{E}_S\left[\sum_{i,j\in S, i\sim j}\mathbf{P}\left(B_i\cap B_j\right)\right]$$

Each term $\mathbf{P}(B_i)$ for $i \in I$ appears with probability p in the first sum on the right and each term $\mathbf{P}(B_i \cap B_j)$ appears with probability p^2 in the second term on the right. This observation gives us:

$$\begin{aligned}\mathbf{E}_S\left[-\log\left(\mathbf{P}\left(\bigcap_{i\in S}\bar{B}_i\right)\right)\right] &\geq \sum_{i\in I} p\mathbf{P}(B_i) - \frac{1}{2(1-\epsilon)}\sum_{i\sim j} p^2\mathbf{P}\left(B_i\cap B_j\right)\\ &= p\mu - \frac{1}{2(1-\epsilon)}p^2\frac{\Delta}{2}\end{aligned}$$

The choice of $p = \frac{\mu(1-\epsilon)}{\Delta}$ gives us the best possible inequality now. Have:

$$\mathbf{E}_S\left[-\log\left(\mathbf{P}\left(\bigcap_{i\in S}\bar{B}_i\right)\right)\right] \geq \frac{\mu^2(1-\epsilon)}{2\Delta}$$

Finally, we observe that if the expectation over all the possible $S \subset I$ is $\geq \frac{\mu^2(1-\epsilon)}{2\Delta}$, then there is at least one subset S_0 that obeys this inequality too. Have:

$$-\log\left(\mathbf{P}\left(\bigcap_{i\in S_0}\bar{B}_i\right)\right) \geq \frac{\mu^2(1-\epsilon)}{2\Delta}$$

This completes the proof, as we have now:

$$\mathbf{P}\left(\bigcap_{i\in I}\bar{B}_i\right) \leq \mathbf{P}\left(\bigcap_{i\in S_0}\bar{B}_i\right) \leq \exp\left(-\frac{\mu^2(1-\epsilon)}{2\Delta}\right)$$

□

0.7. **Appearance of Small Subgraphs Revisited.** We can use Janson's inequality to improve our results about the appearance of strictly balanced graphs H appearing as a subgraph of $G(n,p)$. Before, we showed $n^{-v/e}$ was a threshold function and we understood the behavior when $p \gg n^{-v/e}$. Using Janson's inequality, we can examine the *fine threshold* behavior, that is when $p(n) = cn^{-v/e}$.

Theorem 37. *Let H be a* strictly *balanced graph with v vertices, e edges and a automorphisms. Let $c > 0$ be arbitary. Let A be the property that G contains no copy of H. Then with $p = cn^{-v/e}$we have that:*

$$\lim_{n\to\infty}\mathbf{P}\left(G(n,p)\in A\right) = \exp\left(-\frac{c^e}{a}\right)$$

Proof. Let A_α, with the index $1 \leq \alpha \leq \binom{n}{v}\frac{v!}{a}$ range over the edge sets of possible copies of H, and let B_α be the event that $G(n,p) \supset A_\alpha$ (This is consistent with our notation from the Janson inequalities) In that notation, we have that:

$$\begin{aligned}\mu &= \sum_{1\leq\alpha\leq\binom{n}{v}\frac{v!}{a}}\mathbf{P}(B_i)\\ &= \binom{n}{v}\frac{v!}{a}p^e\\ &\to \frac{c^e}{a}\text{ as } n\to\infty\end{aligned}$$

So then we have$\lim_{n\to\infty} M \leq \exp\left(-\frac{c^e}{a}\right)$. Now, let us examine $\Delta = \sum_{\alpha\sim\beta}\mathbf{P}(B_\alpha \cap B_\beta)$. We use the same idea as before to split the sum into pieces, so that $|\{\text{vertices of overlap between } \alpha \text{ and } \beta\}| = j$ is fixed on each piece (notice that this is different than $|A_\alpha \cap A_\beta| = \#\text{edges of intersection}$). If $j = 0$ or $j = 1$ then $A_\alpha \cap A_\beta = \emptyset$ so that $\alpha \sim \beta$ cannot occur (they are independent). For $2 \leq j \leq v$ let f_j be the maximal overlap of $|A_\alpha \cap A_\beta|$ where $\alpha \sim \beta$and α, β intersect in j vertices. As $\alpha \neq \beta$, we know that $f_j \leq e - 1$. When $2 \leq j \leq v - 1$ the critical observation is that $A_\alpha \cap A_\beta$ is a subgraph of H and hence, as H is stricly balanced:

$$\frac{f_j}{j} < \frac{e}{v}$$

There are $O(n^{2v-j})$ choices of α, β intersecting in j points since α, β are determined up to reordering by $2v - j$ points. For each such α, β we have:

$$\mathbf{P}\left(B_\alpha \cap B_\beta\right) = p^{|A_\alpha\cap A_\beta|} = p^{2e-|A_\alpha\cap A_\beta|} \leq p^{2e-f_j}$$

Thus:

$$\Delta = \sum_{j=2}^{v} O(n^{2v-j})O(n^{-\frac{v}{e}(2e-f_j)})$$

But

$$2v - j - \frac{v}{e}(2e - f_j) = \frac{vf_j}{e} - j < 0$$

so each of the terms in the sum for Δ is $o(1)$. Hence $\Delta = o(1)$! In this case Janson's inequality gets asymptoically close to equality, and so:

$$\lim_{n\to\infty} \mathbf{P}\left(\cap \bar{B}_\alpha\right) = \lim_{n\to\infty} M = \exp(-\frac{c^e}{a})$$

□

Remark 38. This kind of calculation has been worked out for arbitrary graphs, but the calculation gets very messy.

0.8. **Some Very Low Probabilities.** Let A be the property that G does not contain $K_4$and consider $\mathbf{P}(G(n,p) \in A)$ as p varies. We know that $p(n) = n^{-2/3}$is a threshold function, so that for $p \gg n^{-2/3}$the probability is $o(1)$. Here we want to estimate that probability. If we were to do the most naive estimate, and treat every potential copy of K_4 as independent, we would have $\mathbf{P}(G(n,p) \in A) \geq \left(1-p^6\right)^{\binom{n}{4}} = \exp(-n^4p^6 + o(1))$, for p small this turns out to be the right order, but for larger p, say $p = \frac{1}{2}$, we have an even more naive estimate that is better, namely $\mathbf{P}(G(n,p) \in A) \geq \mathbf{P}(G(n,p) \text{ is empty}) = (1-p)^{\binom{n}{2}} = 2^{\binom{n}{2}} = \exp(-n^2 + o(1))$. We can use Janson's inequalities to find the regimes where each estimate is the better one.

Theorem 39. *Say $p(n) = n^{-\alpha}$. For $\frac{2}{3} > \alpha > 0$, $p(n) \gg n^{-2/3}$ so we know that $\mathbf{P}(G(n,p) \in A) = o(1)$. Moreover we have:*

$$\begin{aligned}
\mathbf{P}(G(n,p) \in A) &= \exp\left(-n^{4-6\alpha+o(1)}\right) \text{ for } \frac{2}{3} > \alpha \geq \frac{2}{5} \\
\mathbf{P}(G(n,p) \in A) &= \exp\left(-n^{2-\alpha+o(1)}\right) \text{ for } \frac{2}{5} \geq \alpha > 0
\end{aligned}$$

Proof. The lower bound comes from $\mathbf{P}(G(n,p) \in A) \geq \max\left((1-p^6)^{\binom{n}{4}}, (1-p)^{\binom{n}{2}}\right)$ as described above (this can also be seen as the lower bound for Janson's inequality, using potential K^4s once and using potential K^2s once.)

The upper bound comes from the upper bound in Janson's inequalities. For each set A_α of 4 vertices, let B_α be the event that that 4−set gives a K_4. From our earlier work, we have $\mu = \Theta\left(n^4p^6\right)$ and $-\ln M \sim \mu$ and $\Delta = \theta\left(\mu\Delta^\star\right)$with $\Delta^\star = \Theta\left(n^2p^5 + np^3\right)$. For $p = n^{-\alpha}$ and $\frac{2}{3} > \alpha > \frac{2}{5}$ we have $\Delta^\star = o(1)$ so that:

$$\mathbf{P}\left(\bigcap \bar{B}_\alpha\right) \leq \exp\left(-\mu(1+o(1)\right) = \exp\left(-n^{4-6\alpha+o(1)}\right)$$

When $\frac{2}{5} \geq \alpha > 0$, then $\Delta^\star = \Theta\left(n^2p^5\right)$ so we use the *extended* Janson inequality to get:

$$\mathbf{P}\left(\bigcap \bar{B}_\alpha\right) \leq \exp\left(-\Theta\left(\frac{\mu^2}{\Delta}\right)\right) = \exp\left(-n^{2-\alpha+o(1)}\right)$$

□

Remark 40. There is a general result for arbitary graphs H due to Luczak, Rucinski and Janson (1990).