Thresholds and expectation thresholds in random combinatorics

Jinye He

Wolfson College

University of Oxford

A thesis submitted for the degree of

Master of Science in Mathematics and Foundations of Computer Science

September 2023

Abstract

Park and Pham successfully proved the renowned Kahn-Kalai conjecture, offering an elegant upper-bound solution for thresholds which is a core problem in random discrete structures. In this survey focusing on the Kahn-Kalai conjecture, we will offer many examples to make the conjecture more motivated. Additionally, we will give various applications to vividly illustrate the profound significance of this conjecture.

Contents

Chapter 1

Introduction

The field of random graphs was initiated by Erdős and Rényi in their pioneering paper [11] in 1960. In this seminal work, they introduced two models of random graphs $G = G(V, E)$ where $V = [n] := \{1, ..., n\}$ and E a subset of $\binom{[n]}{2}$.

1. In the $G(n, p)$ model, each of the possible $\binom{n}{2}$ edges in the random graph on n vertices turns up with probability p independently;

2. In the $G(n, m)$ model, the m edges in the random graph on n vertices are chosen uniformly from $\binom{\binom{[n]}{2}}{m}$.

We will refer to the first model the density model and the second model the size model. For most natural questions, for example, the threshold phenomenon we will talk about, these two models behave very similarly. We can correspondingly choose the size parameters, i.e., we take $m \sim p\binom{n}{2}$ and use concentration inequality to convert from the density model to the size model.

Also in this significant paper [11], the central problem, what is the 'typical' structure in a given stage (i.e. p is equal to a given function $p(n)$ of n) of a random graph, was proposed by Erdős and Rényi. A 'typical' structure means that such a structure turns up in the random graph $G(n, p(n))$ with probability tending to 1 as n tends to ∞. In other words, if $\mathcal{F}$ is the property containing a 'typical' structure, 'almost all' $G(n, p(n))$ has this property. In their study, they found the 'threshold phenomenon' for lots of structural properties such as connectedness, containing a given subgraph and so on. This phenomenon describes the sudden change of the appearance and disappearance of certain properties.

Though threshold functions were given for many particular properties since 1960, we didn't know why a property should have a threshold function until Bollobás and Thomason [6] showed that every non-trivial (monotone) increasing property has a threshold in 1987. A trivial property is one that always holds or never holds. For an increasing property, a graph has this property whenever one of its subgraphs has. Many familiar graph properties are increasing properties, for example, a graph is connected, a graph contains a triangle and so on. There is another interesting question about the sharpness of the threshold. We refer interested readers to [15, 14, 13].

Calculating the threshold for a particular property can be very difficult such as the threshold for perfect matchings in hypergraph (known as 'Shamir's problem' and proved by Johansson, Kahn, and Vu [17] in 2008) and bounded degree spanning tree in random graphs (proved by Montgomery [22] in 2019). Both of these two proofs are very difficult (actually, the paper for bounded-degree spanning trees is 71 pages long). In 2007, Kahn and Kalai [19] proposed their conjecture saying that the threshold is not far from its natural lower bound—expectation threshold which can imply these difficult results above. The inspiration of the expectation threshold is from when the number of copies of some subgraph is expected to be larger than zero. This conjecture is so difficult and remarkable that Kahn and Kalai write in [19]

'It would probably be more sensible to conjecture that it is not true.'

In 2010, Talagrand [29] proposed the fractional version of the Kahn-Kalai conjecture which is weaker than the original Kahn-Kalai conjecture. In 2021, Frankston, Kahn, Narayanan and Park [12] proved this weaker conjecture. Their proof is inspired by the work on the sunflower conjecture due to Alweiss, Lovett, Wu and Zhang [1]. In 2022, this conjecture was proved by Park and Pham [24] elegantly within 7 pages which is an excellent breakthrough in random graphs.

This survey effectively introduces the motivation behind the conjecture by presenting various examples, rendering it more natural to understand. Furthermore, the survey underscores the significance of the conjecture by showing a series of applications in historically difficult problems.

Overview In chapter 2, we will introduce the threshold phenomenon after some basic results in Probability Theory and Erdős-Rényi model. In chapter 3, we will give many inspiring examples preparing for the introduction of the Kahn-Kalai conjecture in chapter 4. In chapter 4, we will introduce the Kahn-Kalai conjecture and the weaker but useful result, the fractional Kahn-Kalai conjecture with many applications to show how powerful these results are. After exhibiting the proof of the Kahn-Kalai conjecture in chapter 5, we end this survey with some further work and two open conjectures.

Chapter 2

Preliminaries

In this chapter, we will formally state notions mentioned in chapter 1 and introduce some basic and useful techniques and results. First, we introduce results in Probability Theory such as Markov's inequality and Chebyshev's inequality. Then we will show the relationship between the two models. Finally, we use an example to introduce the threshold phenomenon and show increasing properties always have a threshold.

2.1 Results in Probability Theory

The first important result is Markov's inequality which gives an upper bound of the probability of the event that a non-negative random variable is larger than or equal to a positive constant. This inequality has many sources and is named after Andrey Markov.

Theorem 2.1 (Markov's inequality)**.** *Let X be a non-negative random variable and a be a positive real. The probability*

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}[X]}{a}.$$

Proof. Let $I_{X\geq a}$ be the indicator of the event that the random variable X is larger than a, that is

$$I_{X\geq a} = \begin{cases} 0 & \text{if} \quad X < a \\ 1 & \text{if} \quad X \geq a. \end{cases}$$

Then we have $X \geq aI_{X\geq a}$. Take the expectations of the two sides and we have

$$\mathbb{E}[X] \geq \mathbb{E}[aI_{X\geq a}] = a\mathbb{P}(X \geq a)$$

which implies the inequality we want. □

The second result can be implied from Markov's inequality and it describes the deviation around the mean of a random variable. It is named after Markov's teacher, Pafnuty Chebyshev.

Corollary 2.2 (Chebyshev's inequality)**.** *Let X be a random variable and a be a positive real. The probability*

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

Proof. Let $Y = (X - \mathbb{E}[X])^2$ and Y is a non-negative random variable. By Markov's inequality, the probability

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq a) = \mathbb{P}(Y \geq a^2) \leq \frac{\mathbb{E}[Y]}{a^2} = \frac{\mathrm{Var}[X]}{a^2}.$$

□

These two results seem very simple but show that the moments of a random variable are closely related to its distribution. For some distributions such as Poisson distribution and Normal distribution, the distribution is uniquely determined by its finite-order moments. We introduce the following theorem without proof and refer interested readers to section 30 in [3].

Theorem 2.3 (Method of Moments)**.** *Suppose that the distribution of X is determined by its moments. If there is a sequence of random variables $\{X_n\}$ such that*

$$\lim_{n \to \infty} \mathbb{E}[X_n^r] = \mathbb{E}[X^r]$$

for $\forall r \in \mathbb{Z}^+$, then X_n converges in distribution to X.

Remark 2.4. *Since finite-order moments can be written as linear combinations of binomial moments, the theorem above also shows that the distribution is uniquely determined by binomial moments.*

In random graphs, one of the most important distributions is the binomial distribution. The binomial distribution with parameters n and p, write $B(n, p)$, describes the number of successes in n independent experiments and the probability of success for an individual experiment is p. An obvious observation is that the number of edges in the random graph $G(n, p)$ has a distribution $B(\binom{n}{2}, p)$. By the linearity of expectation, it is easy to get that the mean of $B(n, p)$ is np but the median is not so easy. Fortunately, it was proved by Kaas and Buhrman [18] that medians are very close to the mean for binomial distributions.

Theorem 2.5. *Let X be a random variable with distribution $B(n,p)$. Let $\mu_{\frac{1}{2}}$ be a median of X which means*

$$\mathbb{P}(X \le \mu_{\frac{1}{2}}) \ge \frac{1}{2} \quad \text{and} \quad \mathbb{P}(X \ge \mu_{\frac{1}{2}}) \ge \frac{1}{2}.$$

Then the mean μ and the median $\mu_{\frac{1}{2}}$ satisfy

$$\lfloor \mu \rfloor \le \mu_{\frac{1}{2}} \le \lceil \mu \rceil.$$

In particular, the median equals to the mean when np is an integer.

2.2 Models of random graphs

We mention two models of random graphs in the chapter 1 and we restate them here. A random graph on n vertices (for the sake of convenience, let the vertices set $V = [n]$) is a labelled graph.

- In the density model $G(n,p)$, each of the possible $\binom{n}{2}$ edges in the random graph on n vertices turns up with probability p independently;
- In the size model $G(n,m)$, the m edges in the random graph on n vertices are chosen uniformly from $\binom{\binom{[n]}{2}}{m}$.

In this section, we will show that these two models behave similarly when we take $m \sim p\binom{n}{2}$.

Let $|G(n,p)|$ be the number of edges in $G(n,p)$. The following simple result shows that the density model $G(n,p)$, conditioned on the event that $\{|G(n,p)| = m\}$, is equivalent to the size model $G(n,m)$.

Theorem 2.6. *Let $N = \binom{n}{2}$. The probability that $G(n,p)$ is a given graph G with m edges conditioned on the number of edges is $1/\binom{N}{m}$.*

Proof. Let G be any certain graph on n vertices with m edges. Then we have

$$\begin{aligned}
\mathbb{P}\left(G(n,p) = G \;\middle|\; |G(n,p)| = m\right) &= \frac{\mathbb{P}\left(G(n,p) = G\right)}{\mathbb{P}\left(|G(n,p)| = m\right)} \\
&= \frac{p^m(1-p)^{N-m}}{\binom{N}{m}p^m(1-p)^{N-m}} \\
&= \binom{N}{m}^{-1}
\end{aligned}$$

□

As we mentioned above, the core problem in random graphs is what is the typical structure in a given stage of the random graph. In other words, for a given structure, when does it turn up with high probability? We use 'graph has a property' to refer to the appearance of a particular structure in the graph. If we view the random graph $G(n,p)$ as a random variable, then the sample space is the set of all possible graphs on $[n]$. For a given structure, some possible graphs have this structure but others don't. It follows that when we say that the random graph $G(n,p)$ contains a given structure we mean that the random graph is one of those graphs containing this structure. For labelled graphs on $[n]$, a graph is determined uniquely by its edges set. Formally, a graph property $\mathcal{F}$ is a subset of $2^{\binom{[n]}{2}}$ where $\binom{[n]}{2}$ indicates the set of all possible edges on vertices set $[n]$.

If the property $\mathcal{F}$ such that

$$A \subseteq B \text{ and } A \in \mathcal{F} \Rightarrow B \in \mathcal{F},$$

then we call $\mathcal{F}$ an increasing property. Increasing properties are quite common, for example, a graph containing a triangle, a graph having no isolated vertex, and a graph being connected. So it is natural to focus on increasing properties, and we have the following results.

First, we will introduce the coupling technique and show an intuitively obvious fact as an application of the coupling technique. Let $p_1 \leq p$ and p_2 be defined by the following equation

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

Imagine that G is a random graph in which u, v are adjacent if and only if they are adjacent in $G(n, p_1)$ or $G(n, p_2)$ where $G(n, p_1)$ and $G(n, p_2)$ independent random graphs on the same vertices set $[n]$. It follows that each possible edge turns up in G with probability p independently. Then we have

$$G(n,p) = G(n,p_1) \cup G(n,p_2)$$

where $G(n, p_1)$ and $G(n, p_2)$ are independent. Similarly, in the size model, we have

$$G(n,m) = G(n,m_1) \cup H$$

where $m_1 \leq m$ and H is a random graph on $[n]$ with $m_2 = m - m_1$ edges chosen uniformly from $\binom{[n]}{2} \backslash E\left(G(n, m_1)\right)$. Intuitively, the probability $\mathbb{P}(G(n,p) \text{ has } \mathcal{F})$ increases as p gets larger for increasing property $\mathcal{F}$. By coupling technique, we show this fact formally.

Proposition 2.7. *Let $\mathcal{F}$ be an increasing property. If $p_1 \le p_2$, then*

$$\mathbb{P}\left((G(n,p_1) \text{ has } \mathcal{F}\right) \le \mathbb{P}\left((G(n,p_2) \text{ has } \mathcal{F}\right).$$

Proof. Let $p = \frac{p_2 - p_1}{1-p_1}$. It is easy to check that

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

Let G_1, G be independent copies of $G(n,p_1)$ and $G(n,p)$ respectively. By the coupling technique, we have that $G_2 := G_1 \cup G$ has a distribution as $G(n,p_2)$ and $E(G_1) \subseteq E(G_2)$. Since $\mathcal{F}$ is an increasing property, we have

$$\mathbb{P}\left(G_1 \text{ has } \mathcal{F}\right) \le \mathbb{P}\left(G_2 \text{ has } \mathcal{F}\right).$$

Equivalently,

$$\mathbb{P}\left((G(n,p_1) \text{ has } \mathcal{F}\right) \le \mathbb{P}\left((G(n,p_2) \text{ has } \mathcal{F}\right).$$

□

Similarly, we have the following result in the size model.

Proposition 2.8. *Let $\mathcal{F}$ be an increasing property. If $m_1 \le m_2$, then*

$$\mathbb{P}\left((G(n,m_1) \text{ has } \mathcal{F}\right) \le \mathbb{P}\left((G(n,m_2) \text{ has } \mathcal{F}\right).$$

Recall the theorem 2.6 and let $N = \binom{n}{2}$. Using the law of total probability, we have

$$\mathbb{P}\left(G(n,p) \text{ has } \mathcal{F}\right) = \sum_{k=0}^{N} \mathbb{P}\left(G(n,k) \text{ has } \mathcal{F}\right) \mathbb{P}\left(|G(n,p)| = k\right).$$

Chebyshev's inequality shows that a random variable is far from its mean with a bounded probability. Suppose that $Np \to \infty$ as $n \to \infty$. Recall that $\mathbb{E}[|G(n,p)|] = Np$ and $\text{Var}\left[|G(n,p)|\right] = Np(1-p)$, we have the following observation

$$\begin{aligned}\mathbb{P}\left(\left||G(n,p)| - Np\right| \ge \frac{1}{2}Np\right) &\le \frac{4\text{Var}\left[|G(n,p)|\right]}{(Np)^2} \\ &= \frac{4Np(1-p)}{(Np)^2} \to 0\end{aligned}$$

as $n \to \infty$ which means the probability concentrates around the mean. For increasing properties, we can get a better result using the theorem 2.5.

Theorem 2.9. *Let $\mathcal{F}$ be an increasing property, $p \in [0,1]$ and $m = \lfloor p\binom{n}{2} \rfloor$, we have*

$$\mathbb{P}\left(G(n,p) \text{ has } \mathcal{F}\right) \ge \frac{1}{2}\mathbb{P}\left(G(n,m) \text{ has } \mathcal{F}\right).$$

Proof. Let $N = \binom{n}{2}$. The probability

$$\begin{aligned}\mathbb{P}\left(G(n,p) \text{ has } \mathcal{F}\right) &= \sum_{k=0}^{N} \mathbb{P}\left(G(n,k) \text{ has } \mathcal{F}\right) \mathbb{P}\left(|G(n,p)| = k\right) \\ &\geq \sum_{k=m}^{N} \mathbb{P}\left(G(n,k) \text{ has } \mathcal{F}\right) \mathbb{P}\left(|G(n,p)| = k\right) \\ &\geq \mathbb{P}\left(G(n,m) \text{ has } \mathcal{F}\right) \mathbb{P}\left(|G(n,p)| \geq m\right)\end{aligned}$$

where the second inequality holds by proposition 2.8. Since $|G(n,p)| = B(N,p)$, by the theorem 2.5 we have the median $\mu_{\frac{1}{2}}$ of $|G(n,p)|$ satisfy

$$\lfloor Np \rfloor \leq \mu_{\frac{1}{2}} \leq \lceil Np \rceil.$$

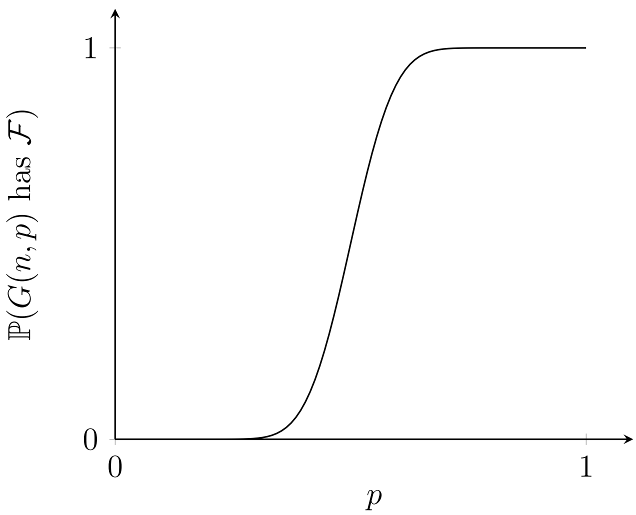
It follows that

$$\mathbb{P}(|G(n,p)| \geq m) = \mathbb{P}(|G(n,p)| \geq \lfloor Np \rfloor) \geq \mathbb{P}(|G(n,p)| \geq \mu_{\frac{1}{2}}) \geq \frac{1}{2}.$$

Therefore, we have $\mathbb{P}\left(G(n,p) \text{ has } \mathcal{F}\right) \geq \frac{1}{2}\mathbb{P}\left(G(n,m) \text{ has } \mathcal{F}\right)$. □

2.3 Threshold phenomenon

In the research of Erdős and Rényi [11, 10], an interesting observation is that the appearance of some increasing properties is abrupt such as connectedness and containing given subgraph. That is to say, random graphs $G(n,p)$ have a property with high probability as long as p is slightly larger than some p^*. Typically, we have the following picture.



We use a simple example to show this phenomenon. This example is common, we reference the lecture note by Riordan [28].

Theorem 2.10. *Let $p^* = n^{-\frac{2}{3}}$ and $\mathcal{F}$ be the property that $G(n,p)$ contains a K_4 as subgraph.*

1. *If $p/p^* \to \infty$ as $n \to \infty$, the the probability $\mathbb{P}(G(n,p)$ has $\mathcal{F}) \to 1$;*

2. *If $p/p^* \to 0$ as $n \to \infty$, the the probability $\mathbb{P}(G(n,p)$ has $\mathcal{F}) \to 0$.*

Proof. Let X be the number of K_4s in $G(n,p)$. Let S be a given subset of $[n]$ with $|S| = 4$ and A_S be the event that S induces a K_4 in $G(n,p)$. Let I_{A_S} be the indicator of the event A_S, that is,

$$I_{A_S} = \begin{cases} 0 & \text{if the event } A_S \text{ does not occur} \\ 1 & \text{if the event } A_S \text{ occurs.} \end{cases}$$

It follows that $X = \sum_S I_{A_S}$. The event A_S occurs if and only if 6 specific edges are present. It follows that

$$\mathbb{P}(A_S) = p^6.$$

By the linearity of expectation, we have

$$\mu := \mathbb{E}[X] = \sum_S \mathbb{P}(A_S) = \binom{n}{4} p^6 \sim \frac{n^4 p^6}{24}. \tag{2.1}$$

In the case of $p/p^* \to 0$, the expectation

$$\mathbb{E}[X] = \Theta\left(n^4 p^6\right) = \Theta\left(n^4 \left(n^{-\frac{2}{3}}\right)^6 (p/p^*)^6\right) \to 0$$

as $n \to \infty$. Thus, by Markov's inequality, we have that

$$\mathbb{P}(G(n,p) \text{ has } \mathcal{F}) = \mathbb{P}(X \geq 1) \leq \frac{\mathbb{E}[X]}{1} \to 0$$

as $n \to \infty$.

In the case of $p/p^* \to \infty$, notice that $\mathbb{E}[X] \to \infty$. But it is not enough to show that $\mathbb{P}(X \geq 1) \to 1$. We will use Chebyshev's inequality to show what we want. First, we consider the second moment of X which will imply $\mathrm{Var}[X]$ with $\mathbb{E}[X]$. Expand the second moment of X and we have

$$X^2 = \left(\sum_S I_{A_S}\right)^2 = \sum_S \sum_T I_{A_S} I_{A_T}.$$

Since $I_{A_S} I_{A_T} = 1$ if and only is A_S and A_T occur, we have

$$\mathbb{E}\left[X^2\right] = \sum_S \sum_T \mathbb{P}(A_S \cap A_T).$$

Notice that $\mathbb{P}(A_S \cap A_T)$ only depends on the size of $S \cap T$. We calculate the contribution in $\mathbb{E}[X^2]$ of different cases of $|S \cap T|$ as follows.

$\lvert S \cap T \rvert$	contribution
0	$\binom{n}{4}p^6\binom{n-4}{4}p^6 \sim \frac{n^4p^6}{24}\frac{n^4p^6}{24} = \mu^2 + o(\mu^2)$
1	$\binom{n}{4}p^6\binom{4}{1}\binom{n-4}{3}p^6 = \Theta(n^7p^{12}) = o(\mu^2)$
2	$\binom{n}{4}p^6\binom{4}{2}\binom{n-4}{2}p^5 = \Theta(n^6p^{11}) = o(\mu^2)$
3	$\binom{n}{4}p^6\binom{4}{3}\binom{n-4}{1}p^3 = \Theta(n^5p^9) = o(\mu^2)$
4	$\binom{n}{4}p^6 = \mu.$

Since

$$\mathbb{P}(G(n,p) \text{ doesn't have } \mathcal{F}) = \mathbb{P}(X = 0) \le \mathbb{P}\left(\left|X - \mu\right| \ge \mu\right),$$

we estimate the probability $\mathbb{P}\left(\left|X-\mu\right| \ge \mu\right)$ by Chebyshev's inequality,

$$\begin{aligned}\mathbb{P}\left(\left|X - \mu\right| \ge \mu\right) &\le \frac{\operatorname{Var}[X]}{\mu^2}\\ &= \frac{\mathbb{E}[X^2] - \mathbb{E}[X]^2}{\mu^2}\\ &= \frac{\mu^2 + o(\mu^2) + \mu - \mu^2}{\mu^2} \to 0\end{aligned}$$

as $n \to \infty$. It follows that $\mathbb{P}(G(n,p) \text{ has } \mathcal{F}) \to 1$ as $n \to \infty$ when $p/p^* \to \infty$ and the proof is complete. □

Now we formally state the definition of the threshold.

Definition 2.11. Let $\mathcal{F}$ be a property of random graphs. A function $p^*(n)$ is called a threshold function for $\mathcal{F}$ if

1. $\mathbb{P}(G(n,p(n)) \text{ has } \mathcal{F}) \to 0$ when $\frac{p(n)}{p^*(n)} \to 0$ and
2. $\mathbb{P}(G(n,p(n)) \text{ has } \mathcal{F}) \to 1$, when $\frac{p(n)}{p^*(n)} \to \infty$.

A threshold function is not unique since we can easily get a new one by multiplying a constant. By the simple example above, we use Markov's inequality and Chebyshev's inequality which are the most basic and useful methods to get the threshold. We call them the first moment method and the second moment method.

Theorem 2.12 (First moment method)**.** *Let X be a non-negative integer random variable. The probability*

$$\mathbb{P}(X \geq 1) \leq \mathbb{E}[X].$$

Theorem 2.13 (Second moment method)**.** *Let X be a non-negative integer random variable. The probability*

$$\mathbb{P}(X = 0) \leq \frac{\mathrm{Var}[X]}{\mathbb{E}[X]^2}.$$

Calculating the expectation of a random variable tends to be easy since its linearity. A function $p(n)$ such that $\mathbb{E}[X] \to 0$ gives a natural lower bound of the threshold. A natural idea is whether the lower bound is exactly the threshold or how far is the threshold from the lower bound. In later chapters, we will show that although the threshold can't be gotten easily just from the expectation, it is not far from this lower bound provided we are careful about what we are counting.

In addition to what the threshold is, another important problem is when the threshold exists. In 1987, Bollobás and Thomason [6] showed why non-trivial increasing properties always have a threshold. We end this chapter with this result. The proof here references [16] by Frieze and Karoński.

Theorem 2.14. *If $\mathcal{F}$ is a non-trivial increasing property, there is a threshold for random graphs to have property $\mathcal{F}$.*

Proof. Let $G_1, \cdots, G_k$ be independent copies of $G(n,p)$. Then the graph $G = \bigcup_{i=1}^{k} G_k$ is distributed as $G(n, 1-(1-p)^k)$. Since $1-(1-p)^k \leq kp$, $G(n,kp)$ can be obtained from $G(n, 1-(1-p)^k)$ by coupling with $G(n,p')$ where p' is such that

$$(1-p)^k(1-p') = 1 - kp.$$

Notice that G doesn't have property $\mathcal{F}$ implies none of $\{G_i\}_{i=1}^k$ has property $\mathcal{F}$. It follows that

$$\begin{aligned}\mathbb{P}\left(G(n,kp) \text{ doesn't have } \mathcal{F}\right) &\leq \prod_{i=1}^{k} \mathbb{P}\left(G_i \text{ doesn't have } \mathcal{F}\right) \\ &= \left(\mathbb{P}\left(G(n,p) \text{ doesn't have } \mathcal{F}\right)\right)^k\end{aligned}$$

Let $N = \binom{n}{2}$. For each element S in $\mathcal{F}$, the probability of the event that the random graph $G(n,p)$ is the graph exactly containing these edges is $p^{|S|}(1-p)^{N-|S|}$. Since different elements in $\mathcal{F}$ are mutually exclusive, we have

$$\mu_p(\mathcal{F}) := \mathbb{P}\left(G(n,p) \text{ has } \mathcal{F}\right) = \sum_{S \in \mathcal{F}} p^{|S|}(1-p)^{N-|S|}$$

which is a polynomial in p. As we show in proposition 2.7, $\mu_p(\mathcal{F})$ increases as p increases from 0 to 1. For $\epsilon \in [0,1]$, let $p(\epsilon)$ be defined by $\mu_{p(\epsilon)}(\mathcal{F}) = \epsilon$. Let $p^* = p(\frac{1}{2})$ and $\omega(n) \to \infty$ arbitrarily slowly. Without loss of generality, we suppose $\omega(n) > 0$. We have

$$\mathbb{P}\left(G(n, \omega p^*) \text{ doesn't have } \mathcal{F}\right) \leq \left(\mathbb{P}\left(G(n, p^*) \text{ doesn't have } \mathcal{F}\right)\right)^{\omega} = \left(\frac{1}{2}\right)^{\omega} \to 0$$

as $n \to \infty$ and

$$\frac{1}{2} = \mathbb{P}\left(G(n, p^*) \text{ doesn't have } \mathcal{F}\right) \leq \left(\mathbb{P}\left(G(n, p^*/\omega) \text{ doesn't have } \mathcal{F}\right)\right)^{\omega}$$

which implies that

$$\mathbb{P}\left(G(n, p^*/\omega) \text{ doesn't have } \mathcal{F}\right) \geq 2^{-\frac{1}{\omega}} \to 1$$

as $n \to \infty$. It follows that $\mathbb{P}\left(G(n, \omega p^*) \text{ has } \mathcal{F}\right) \to 1$ and $\mathbb{P}\left(G(n, p^*/\omega) \text{ has } \mathcal{F}\right) \to 0$ as $n \to \infty$. Since that $\omega(n) \to \infty$ arbitrarily slowly, p^* is a threshold for property $\mathcal{F}$. $\square$

Remark 2.15. *The value $\frac{1}{2}$ of $p^* = p(\frac{1}{2})$ in the proof is not essential. Any constant strictly between 0 and 1 would work.*

Chapter 3

Examples of threshold

In this chapter, we will give some examples of thresholds. Through these examples, we want to show that the first moment method provides a natural lower bound of the threshold easily but in many cases, it is not exactly a threshold. Fortunately, the threshold seems not far from this lower bound.

The first section will show the threshold concerning a small subgraph with finite vertices and edges. These cases are relatively easy to analyze.

In the second section, we move our attention to large subgraphs such as spanning trees, that is, the connectedness of the random graph. This is one of the important earliest results in random graphs given by Erdős and Rényi [10]. In fact, they prove that $\frac{\log n}{n}$ is the sharp threshold for connectedness which is stronger than the threshold.

Finally, based on the threshold concerning connectedness, we show another two results concerning the Hamilton cycles and perfect matchings. In terms of perfect matchings, due to Erdős and Rényi's work [9], the threshold was given in 1964. The threshold for Hamilton cycles is much more difficult and left open until 1976 solved by Pósa [25]. Since the aim of this chapter is to show some motivations for Kahn and Kalai to propose their conjecture, we first show the threshold for Hamilton cycles and deduce the threshold for perfect matchings. Interestingly, it is found that both of them have similar relations with the lower bound given by the first moment method.

3.1 Small subgraph

Recall the threshold for random graphs containing a K_4 is $n^{-\frac{2}{3}}$ which is equal to the lower bound provided by the first moment method. However, this coincidence will vanish after a tiny change. The discussion here references [28].

Let H be a kite, K_4 with an extra edge hanging out. Since K_4 is a subgraph of H, we have that $G(n, p)$ contains a K_4 whenever it contains an H. As we do in the

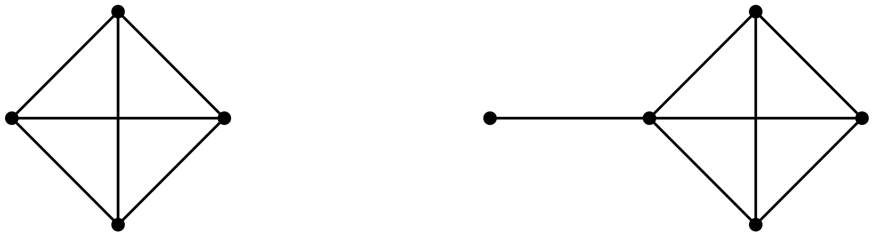


Figure 3.1: K_4 and H

example of K_4, for a copy of H, the probability of the event that this copy turns up in $G(n,p)$ if and only if 7 specific edges are present. In terms of number of copies, since the number of automorphisms of H is 3!, there are $\frac{n(n-1)\cdots(n-4)}{3!}$ different copies of H. It follows that

$$\mathbb{E}\left[\text{number of copies of } H \text{ in } G(n,p)\right] = \frac{n(n-1)\cdots(n-4)}{3!}p^7.$$

Using the first moment method, we suppose $n^{-\frac{5}{7}}$ is a threshold. Notice that $n^{-\frac{5}{7}}/n^{-\frac{2}{3}} = n^{-\frac{1}{21}} \to 0$ as $n \to \infty$ and we can find some $p(n)$ between $n^{-\frac{5}{7}}$ and $n^{-\frac{2}{3}}$ such that

$$p(n)/n^{-\frac{2}{3}} \to 0 \quad \text{and} \quad p(n)/n^{-\frac{5}{7}} \to \infty$$

as $n \to \infty$. By the definition of thresholds, we have

$$\mathbb{P}\left(G(n,p(n)) \text{ contains } K_4\right) \to 0 \quad \text{and} \quad \mathbb{P}\left(G(n,p(n)) \text{ contains } H\right) \to 1$$

as $n \to \infty$ which is contradictory to the fact that $G(n,p)$ containing H implies $G(n,p)$ containing K_4. Though this example shows the threshold is not always given by the first moment method, for small subgraphs with finite vertices and edges, the threshold can be solved by the first and second moment method. Before introducing the result, we define the edge density of a graph and simplify the second moment method.

Definition 3.1. The edge density $d(G)$ of a graph G is $|E(G)|/|V(G)|$.

Let $\{A_{n,i}\}$ be a sequence of events and X_n be the number of events which occur. Taking independence into consideration, we have the following lemma.

Lemma 3.2. *Let $\Delta_n = \sum_i \sum_{i \sim j} \mathbb{P}(A_{n,i} \cap A_{n,j})$ where $i \sim j$ means events $A_{n,i}$ and $A_{n,j}$ are dependent and $i \neq j$. Let $\mathbb{E}\left[X_n\right] = \mu_n$. If $\Delta_n/\mu_n^2 \to 0$ and $\mu_n \to \infty$, then we have $\mathbb{P}\left(X_n = 0\right) \to 0$ as $n \to \infty$.*

Proof. By the definition of variance, we have $\mathrm{Var}[X_n] = \mathbb{E}\left[(\sum_i I_{n,i})^2\right] - (\mathbb{E}\left[\sum_i I_{n,i}\right])^2$ where $I_{n,i}$ is the indicator of the event $A_{n,i}$. It follows that

$$\begin{aligned}\mathrm{Var}[X_n] &= \sum_i \sum_j (\mathbb{P}(A_{n,i} \cap A_{n,j}) - \mathbb{P}(A_{n,i})\mathbb{P}(A_{n,j})) \\ &= \sum_i \left(\mathbb{P}(A_{n,i}) - \mathbb{P}(A_{n,i})^2\right) + \sum_i \sum_{i \neq j} (\mathbb{P}(A_{n,i} \cap A_{n,j}) - \mathbb{P}(A_{n,i})\mathbb{P}(A_{n,j})) \\ &\leq \mathbb{E}[X_n] + \sum_i \sum_{i \sim j} \mathbb{P}(A_{n,i} \cap A_{n,j}) = \mu_n + \Delta_n.\end{aligned}$$

If $\frac{\Delta_n}{\mu_n^2} \to 0$ and $\mu_n \to \infty$, then we have $\frac{\mathrm{Var}[X_n]}{\mu_n^2} \to 0$. By the theorem 2.13, we have $\mathbb{P}(X_n = 0) \to 0$ as $n \to \infty$. □

Theorem 3.3. *Let H be a graph and let H' be one of the subgraphs of H with maximum edge density. One of thresholds of $G(n,p)$ containing H is $p^* = n^{-\frac{1}{d(H')}}$.*

Proof. Let $e = |E(H)|, v = |V(H)|$ and $e' = |E(H')|, v' = |V(H')|$. So we have $d(H') = \frac{e'}{v'}$ and $\frac{e'}{v'} \geq \frac{e}{v}$. Let X be the number of copies of H in $G(n,p)$. Let $\{H_i\}$ be all possible copies of H and $\{A_i\}$ be the corresponding event that H_i turn up in $G(n,p)$. The event A_i occurs if and only if specific e edges are present. It follows that

$$\mathbb{P}(A_i) = p^e.$$

By the linearity of expectation, we have

$$\mu := \mathbb{E}[X] = \frac{n(n-1)\cdots(n-v+1)}{aut(H)} p^e = \Theta(n^v p^e)$$

where $aut(H)$ is the number of automorphism of H. Notice that events A_i and A_j are dependent if and only if $E(H_i) \cap E(H_j) \neq \emptyset$. We have that

$$\Delta = \sum_i \sum_{i \sim j} \mathbb{P}(A_i \cap A_j) = \sum_i \sum_{i \sim j} \mathbb{P}(H_i \cup H_j \subseteq G(n,p)).$$

Let r be the number of vertices in $H_i \cap H_j$ and s be the number of edges in $H_i \cap H_j$. Since $H_i \cap H_j$ is also a subgraph of H, we have $\frac{s}{r} \leq \frac{e'}{v'}$. There are $2v - r$ vertices and $2e - s$ edges in $H_i \cup H_j$. For fixed r, s, the contribution of these terms are

$$\Theta(n^{2v-r} p^{2e-s}) = \Theta(\mu^2 n^{-r} p^{-s}) = \Theta\left(\mu^2 \frac{1}{n^r p^{*s}} \left(\frac{p}{p^*}\right)^{-s}\right).$$

Since $\frac{s}{r} \leq \frac{e'}{v'}$ and $\frac{e}{v} \leq \frac{e'}{v'}$, we have

$$n^r p^{*s} = n^{r - s\frac{v'}{e'}} \geq 1 \quad \text{and} \quad n^v p^{*e} = n^{v - e\frac{v'}{e'}} \geq 1.$$

In the case of $p/p^* \to \infty$ as $n \to \infty$, we have

$$\Theta\left(\mu^2 \frac{1}{n^r p^{*s}} \left(\frac{p}{p^*}\right)^{-s}\right) / \mu^2 \leq \Theta\left(\left(\frac{p}{p^*}\right)^{-s}\right) \to 0 \tag{3.1}$$

and

$$\mu = \Theta(n^v p^e) = \Theta\left(n^v p^{*e} \left(\frac{p}{p^*}\right)^e\right) \geq \Theta\left(\left(\frac{p}{p^*}\right)^e\right) \to \infty$$

as $n \to \infty$. Since H is a graph with finite vertices and edges, the number of different pairs of r, s is finite. It follows that $\Delta = o(\mu^2)$ by (3.1). Using the lemma 3.2, we have

$$\mathbb{P}\left(X = 0\right) \to 0$$

as $n \to \infty$ which implies that $\mathbb{P}(G(n,p)$ contains $H) \to 1$.

In terms of the case $p/p^* \to 0$, let X' be the number of H's in $G(n,p)$. We have

$$\mathbb{E}[X'] = \frac{n(n-1)\cdots(n-v'+1)}{aut(H')} p^{e'} = \Theta(n^{v'} p^{e'}) = \Theta(n^{v'} n^{-e' \frac{v'}{e'}} \left(\frac{p}{p^*}\right)^{e'}) \to 0$$

as $n \to \infty$ which implies that $\mathbb{P}(X' \geq 1) \to 0$ by Markov's inequality. Since H' is a subgraph of H, we have

$$\mathbb{P}(G(n,p) \text{ contains } H) \leq \mathbb{P}(G(n,p) \text{ contains } H') = \mathbb{P}(X' \geq 1) \to 0.$$

Therefore, we complete our proof and show that $p^* = n^{-\frac{1}{d(H')}}$ is a threshold of $G(n,p)$ containing H. □

3.2 Connectedness

Connectedness may be the most basic property of a graph. Since a graph is connected if and only if it contains a spanning tree, a natural idea to get the threshold for connectedness is to count the number of spanning trees as we do in the last section. For a labelled graph on n vertices, there are n^{n-2} possible spanning trees by the famous Cayley's formula and each of them contains $n-1$ edges. Let X be the number of spanning trees in the random graph $G(n,p)$. It is easy to check that

$$\mathbb{E}[X] = n^{n-2} p^{n-1}.$$

Take $p^* = n^{-1}$ and we have $\mathbb{E}[X] \to 0$ as $n \to \infty$ which is a lower bound of the threshold. However, it is too difficult to consider the second moment which leads us to another idea.

A component of a graph is a connected subgraph and there is no larger connected subgraph containing it. A k-component is a component on k vertices. It is obvious that a graph is connected if and only if there is no k-component with $2 \leq k < n$ and isolated vertices. First, let's focus on the isolated vertices.

Theorem 3.4. *Let $\mathcal{F}$ be the property that there is no isolated vertex in the graph. Let*

$$p = \frac{\log n + \omega(n)}{n}.$$

The probability

$$\mathbb{P}\left(G(n,p) \ has \ \mathcal{F}\right) = \begin{cases} 1 & if \quad \omega(n) \to \infty \\ 0 & if \quad \omega(n) \to -\infty \end{cases}$$

where $|\omega(n)| \to \infty$ arbitrarily slowly as $n \to \infty$.

Proof. Since $|\omega(n)| \to \infty$ arbitrarily slowly, we assume that $|\omega(n)| = o(\log n)$ without loss of generality. Let X be the number of isolated vertices in $G(n,p)$. Let A_i be the event that the vertex i is isolated and we have $\mathbb{P}(A_i) = (1-p)^{n-1}$. By the linearity of expectation, we have

$$\mathbb{E}[X] = \sum_i \mathbb{P}(A_i) = n(1-p)^{n-1}.$$

Since $p \to 0$ as $n \to \infty$, we use Taylor expansion to have

$$1 - p = e^{-p+O(p^2)}.$$

Using this estimate, we have

$$\mathbb{E}[X] = n(1-p)^{n-1} = e^{\log n - pn + p + O(p^2 n)}.$$

Substituting $p = \frac{\log n + \omega(n)}{n}$, we have

$$\begin{aligned} \mathbb{E}[X] &= e^{\log n - pn + p + O(p^2 n)} \\ &= e^{-\omega(n) + o(1)} \\ &= (1 + o(1)) e^{-\omega(n)}. \end{aligned}$$

In the case of $\omega(n) \to \infty$, we have

$$\mathbb{P}(G(n,p) \text{ doesn't have } \mathcal{F}) = \mathbb{P}(X \geq 1) \leq \frac{\mathbb{E}[X]}{1} \to 0$$

as $n \to \infty$. Equivalently, we have $\mathbb{P}(G(n,p) \text{ has } \mathcal{F}) \to 1$ as $n \to \infty$ if $\omega(n) \to \infty$.

In the case of $\omega(n) \to -\infty$, we have

$$\mu := \mathbb{E}[X] = (1 + o(1))e^{-\omega(n)} \to \infty.$$

The second moment of X can be write as $\mathbb{E}[X^2] = \sum_i \sum_j \mathbb{P}(A_i \cap A_j) = \sum_i \mathbb{P}(A_i) + \sum_i \sum_{i \neq j} \mathbb{P}(A_i \cap A_j)$ and the event $A_i \cap A_j$ for $i \neq j$ occurs if and only if all $2n - 4$ edges between $\{i, j\}$ and the other vertices are absent as well as ij. It follows that

$$\mathbb{P}(A_i \cap A_j) = (1 - p)^{2n-3}$$

and

$$\begin{aligned}
\mathbb{E}[X^2] &= \sum_i \mathbb{P}(A_i) + \sum_{i \neq j} \mathbb{P}(A_i \cap A_j) \\
&= \mu + \mu^2 \frac{1}{1-p} + O(\mu) \\
&= (1 + o(1))\mu^2 + o(\mu^2),
\end{aligned}$$

we have $\mathrm{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = o(\mu^2)$. By the theorem 2.13, we have

$$\mathbb{P}(X = 0) \to 0$$

as $n \to \infty$ which implies that $\mathbb{P}(G(n, p) \text{ has } \mathcal{F}) \to 0$ as $n \to \infty$. □

The threshold can be implied by this result and actually, it shows $\frac{\log n}{n}$ is a sharp threshold. From the definition of a sharp threshold, it is easy to find that a sharp threshold is also a threshold.

Definition 3.5. Let $\mathcal{F}$ be an increasing property of random graphs. A function $p^*(n)$ is called a sharp threshold function for $\mathcal{F}$ if for $\forall \epsilon > 0$

1. $\lim_{n\to\infty} \mathbb{P}(G(n, p(n)) \text{ has } \mathcal{P}) = 0$ when $\lim_{n\to\infty} \frac{p(n)}{p^*(n)} \leq 1 - \epsilon$ and
2. $\lim_{n\to\infty} \mathbb{P}(G(n, p(n)) \text{ has } \mathcal{P}) = 1$, when $\lim_{n\to\infty} \frac{p(n)}{p^*(n)} \geq 1 + \epsilon$.

The theorem 3.4 shows the probability of the random graph containing no isolated vertices when $|\omega(n)| \to \infty$. Actually, for $|\omega(n)| < \infty$, we have a finer description, that is, the distribution of the number of isolated vertices in $G(n, p)$ converges to Poisson distribution. We maintain X denote the number of isolated vertices and $p = \frac{\log n + \omega(n)}{n}$. There are two methods to show this result, referencing [16] by Frieze and Karoński.

The first one uses the method of moments. As we state in theorem 2.3 and remark 2.4, if finite order binomial moments of X tend to the corresponding binomial moments of Poisson distribution, then X converges in distribution to Poisson distribution. Since the k^{th} binomial moment of Poisson distribution with parameter e^{-c} is $\frac{e^{-ck}}{k!}$, we will show that

$$\mathbb{E}\left[\binom{X}{k}\right] \to \frac{e^{-ck}}{k!} \quad \text{as} \quad n \to \infty.$$

Let A_i denote the event that the vertex i is isolated. As the combinatorial meaning of binomial coefficients, $\binom{X}{k}$ describes the number of different choices of unordered k isolated vertices. Then the k^{th} binomial moment of X satisfy

$$\begin{aligned}\mathbb{E}\left[\binom{X}{k}\right] &= \sum_{1 \le i_1 < \cdots < i_k \le n} \mathbb{P}\left(\bigcap_{s=1}^{k} A_{i_s}\right) \\ &= \binom{n}{k}(1-p)^{k(n-k)+\binom{k}{2}} \\ &= \frac{1}{k!}\frac{n!}{(n-k)!}(1-p)^{k(n-k)+\binom{k}{2}}\end{aligned}$$

The second equation comes from the equivalence between the event that given k vertices are isolated and the event that there is no edge between these k vertices and the other $n-k$ vertices as well as among these k vertices. By the estimate that $1-p = e^{-p+O(p^2)}$ and Stirling formula, we have the follow estimate

$$\begin{aligned}\frac{n!}{(n-k)!}(1-p)^{k(n-k)+\binom{k}{2}} &\sim \sqrt{\frac{n}{n-k}}\left(\frac{n}{e}\right)^n\left(\frac{e}{n-k}\right)^{n-k} e^{k(-p+O(p^2))(n-k+\frac{k-1}{2})} \\ &\sim \left(\frac{n}{e}\right)^k\left(1+\frac{k}{n-k}\right)^{\frac{n-k}{k}k}\left(n^{-1}e^{-\omega(n)}\right)^k \\ &\sim e^{-ck}\end{aligned}$$

and $\mathbb{E}\left[\binom{X}{k}\right] \to \frac{e^{-ck}}{k!}$ as $n \to \infty$. It follows that X converges to Poisson distribution with parameter e^{-c} by the method of moments.

The second method shows the distribution of X directly. Before our proof, we first give some lemmas.

A Boolean function is a function of n variables, each of which is in $\{0,1\}$ and only has two values in $\{0,1\}$. Let $A_1, \cdots, A_n$ be events and I_{A_i}s be the indicators of A_is.

Lemma 3.6. *Let $A_1, \cdots, A_n$ be events. Let $f_1, \cdots, f_k$ be Boolean polynomials in n variables and $\alpha_1, \cdots \alpha_k$ be real constants. If*

$$\mathbb{E}\left[\sum_{i=1}^{k} \alpha_i f_i(I_{A_1}, \cdots, I_{A_n})\right] \ge 0 \tag{3.2}$$

whenever $\mathbb{P}(A_i) = 0$ or 1 for $\forall i$, then (3.2) holds for every choice of events.

Proof. First, the event $\{f_i(I_{A_1}, \cdots, I_{A_n}) = 1\}$ can be written as

$$f_i^{-1}(\{1\}) = \bigsqcup_{S \subseteq \mathcal{I}} \left(\left(\bigcap_{i \in S} A_i \right) \cap \left(\bigcap_{i \notin S} A_i^C \right) \right)$$

for some $\mathcal{I} \subseteq 2^{[n]}$. It follows that

$$\mathbb{P}(f_i(I_{A_1}, \cdots, I_{A_n}) = 1) = \sum_{S \subseteq \mathcal{I}} \mathbb{P} \left(\left(\bigcap_{i \in S} A_i \right) \cap \left(\bigcap_{i \notin S} A_i^C \right) \right)$$

and we can rewrite the left side of (3.2) as follows

$$\begin{aligned} \mathbb{E}\left[\sum_{i=1}^{k} \alpha_i f_i(I_{A_1}, \cdots, I_{A_n}) \right] &= \sum_{i=1}^{k} \alpha_i \sum_{S \subseteq \mathcal{I}} \mathbb{P} \left(\left(\bigcap_{i \in S} A_i \right) \cap \left(\bigcap_{i \notin S} A_i^C \right) \right) \\ &= \sum_{S \subseteq [n]} \beta_S \mathbb{P} \left(\left(\bigcap_{i \in S} A_i \right) \cap \left(\bigcap_{i \notin S} A_i^C \right) \right). \end{aligned} \tag{3.3}$$

Since we suppose (3.3) is non-negative whenever $\mathbb{P}(A_i) = 0$ or 1, we take $A_i = \Omega$ (Ω is the sample space) for $i \in T$ and $A_i = \emptyset$ for $i \notin T$ for $T \subseteq [n]$. We have

$$\sum_{S \subseteq [n]} \beta_S \mathbb{P} \left(\left(\bigcap_{i \in S} A_i \right) \cap \left(\bigcap_{i \notin S} A_i^C \right) \right) = \beta_T \geq 0.$$

Therefore, we have $\beta_T \geq 0$ for $\forall T \subseteq [n]$ which implies that

$$\mathbb{E}\left[\sum_{i=1}^{k} \alpha_i f_i(I_{A_1}, \cdots, I_{A_n}) \right] \geq 0$$

in general. $\square$

This lemma was proposed by Rényi [27] in 1958 and the principle of inclusion-exclusion can be implied from this lemma. Using this lemma, we prove the following result.

Lemma 3.7. *Let $A_1, \cdots, A_n$ be events in the sample space Ω and X be the number of events among $\{A_i\}$ that occur. Let $B_k = \mathbb{E}\left[\binom{X}{k}\right]$. Then we have*

$$\mathbb{P}(X = j) \begin{cases} \leq \sum_{k=j}^{s} (-1)^{k-j} \binom{k}{j} B_k & \text{if} \quad s - j \text{ is even} \\ \geq \sum_{k=j}^{s} (-1)^{k-j} \binom{k}{j} B_k & \text{if} \quad s - j \text{ is odd} \\ = \sum_{k=j}^{s} (-1)^{k-j} \binom{k}{j} B_k & \text{if} \quad s = n. \end{cases} \tag{3.4}$$

Proof. To use the lemma 3.6, we just need to check the cases when $\mathbb{P}(A_i) = 0$ or 1 for $\forall i \in [n]$. Suppose that exactly ℓ given events occur with probability 1 and the others never occur for $1 \le \ell \le n$. We have $B_k = \binom{\ell}{k}$. It follows that

$$\begin{aligned}
\sum_{k=j}^{s}(-1)^{k-j}\binom{k}{j}B_k &= \sum_{k=j}^{s}(-1)^{k-j}\binom{k}{j}\binom{\ell}{k} \\
&= \sum_{k=j}^{s}(-1)^{k-j}\frac{k!\ell!}{j!(k-j)!k!(\ell-k)!} \\
&= \sum_{k=j}^{s}(-1)^{k-j}\frac{\ell!}{j!(\ell-j)!}\frac{(\ell-j)!}{(k-j)!(\ell-k)!} \\
&= \binom{\ell}{j}\sum_{k=j}^{s}(-1)^{k-j}\binom{\ell-j}{k-j}. \qquad (3.5)
\end{aligned}$$

If $\ell < j$, we have $\binom{\ell-j}{k-j} = 0$ which equals to $\mathbb{P}(X = j) = 0$. In terms of $\ell = j$, we have (3.5) only have the first term $\binom{j}{j}(-1)^{j-j}\binom{j-j}{j-j} = 1$ non-zero and $\mathbb{P}(X = j) = 1$. Finally, if $j < \ell \le n$, we have $\mathbb{P}(X = j) = 0$ and

$$\begin{aligned}
\sum_{k=j}^{s}(-1)^{k-j}\binom{\ell-j}{k-j} &= \sum_{t=0}^{s-j}(-1)^{t}\binom{\ell-j}{t} \\
&= \binom{\ell-j}{0} + \sum_{t=1}^{s-j}(-1)^{t}\left(\binom{\ell-j-1}{t} + \binom{\ell-j-1}{t-1}\right) \\
&= (-1)^{s-j}\binom{\ell-j-1}{s-j}.
\end{aligned}$$

The second equality uses the simple fact that $\binom{\ell-j}{t} = \binom{\ell-j-1}{t} + \binom{\ell-j-1}{t-1}$. Therefore, we have shown that (3.4) holds for $\mathbb{P}(A_i) = 0$ or 1. By lemma 3.6, we have it holds in general. □

This lemma provides us with an upper bound and a lower bound of the probability $\mathbb{P}(X = j)$. Then squeeze theorem shows what we want.

Theorem 3.8. *Let X be the number of isolated vertices in $G(n,p)$ with $p = \frac{\log n + \omega(n)}{n}$ and $\omega(n) \to c < \infty$ as $n \to \infty$. The probability*

$$\mathbb{P}(X = j) \to e^{-e^{-c}}\frac{e^{-cj}}{j!}$$

as $n \to \infty$ which implies that X converges in distribution to the Poisson distribution.

Proof. As we mentioned above, we have $B_k = \mathbb{E}\left[\binom{X}{k}\right] \to \frac{e^{-ck}}{k!}$ as $n \to \infty$. By lemma 3.7, for $\ell \geq 0$, just consider

$$\sum_{k=j}^{j+2\ell+1} (-1)^{k-j}\binom{k}{j}\frac{e^{-ck}}{k!} \leq \lim_{n\to\infty} \mathbb{P}(X=j) \leq \sum_{k=j}^{j+2\ell}(-1)^{k-j}\binom{k}{j}\frac{e^{-ck}}{k!}.$$

It is easy to check that

$$\begin{aligned}\sum_{k=j}^{j+m}(-1)^{k-j}\binom{k}{j}\frac{e^{-ck}}{k!} &= \sum_{t=0}^{m}(-1)^t\frac{e^{-c(j+t)}}{j!t!}\\ &= \frac{e^{-cj}}{j!}\sum_{t=0}^{m}(-1)^t\frac{e^{-ct}}{t!}\\ &\to \frac{e^{-cj}}{j!}e^{-e^{-c}}\end{aligned}$$

as $m \to \infty$ which implies that $\mathbb{P}(X=j) \to e^{-e^{-c}}\frac{e^{-cj}}{j!}$. □

Now we return to our starting problem when the random graph $G(n,p)$ becomes connected. Amazingly, it is found that random graphs become connected as long as isolated vertices vanish.

Theorem 3.9. *Let* $p = \frac{\log n + \omega(n)}{n}$. *The probability*

$$\mathbb{P}\left(G(n,p) \text{ is connected}\right) \to \begin{cases} 1 & \text{if} \quad \omega(n)\to\infty \\ e^{-e^{-c}} & \text{if} \quad \omega(n)\to c<\infty \\ 0 & \text{if} \quad \omega(n)\to-\infty\end{cases}$$

as $n \to \infty$.

Proof. Suppose $\mathbb{P}\left(G(n,p) \text{ is connected}\right) \to e^{-e^{-c}}$ when $\omega(n) \to c < \infty$. Since connectedness is an increasing property, we have $\mathbb{P}\left(G(n,p) \text{ is connected}\right) \to 1$ when $\omega(n) \to \infty$ and $\mathbb{P}\left(G(n,p) \text{ is connected}\right) \to 0$ when $\omega(n) \to -\infty$.

Therefore, we only need to show the case of $\omega(n) \to c < \infty$. Without loss of generality, suppose $|\omega(n)| \leq \log n$ for $n \geq 1$. It is known that a graph is connected if and only if it has no k-components for $2 \leq k < n$ as well as isolated vertices. Let X_k be the number of k-component. Since it is impossible that there is no k-component for $2 \leq k \leq \lfloor \frac{n}{2} \rfloor$ nor isolated vertices while ℓ-component exists for $\ell > \lfloor \frac{n}{2} \rfloor$, we only need to consider k-component for $k \leq \lfloor \frac{n}{2} \rfloor$ and isolated vertices. For the sake of convenience, we view isolated vertices as 1-components.

First, we consider k-component for $2 \leq k \leq \lfloor \frac{n}{2} \rfloor$. Let X_k be the number of k-component. Let S be a set of vertices with k elements and A_S be the event that S

induces a k-component. Given k vertices, the event that they induce a k-component means there is a spanning tree on these vertices meanwhile there is no edge between these k vertices and the others. By Cayley's formula, we have the union bound of the probability

$$\mathbb{P}(A_S) \le (1-p)^{k(n-k)} k^{k-2} p^{k-1}.$$

Then we have

$$\mathbb{E}[X_k] \le \binom{n}{k} (1-p)^{k(n-k)} k^{k-2} p^{k-1}.$$

By the following estimates[1]

$$\binom{n}{k} \le \left(\frac{en}{k}\right)^k \qquad 1 - x \le e^{-x},$$

we have

$$\begin{aligned}
\mathbb{E}[X_k] &\le \left(\frac{en}{k}\right)^k e^{-pk(n-k)} k^{k-2} p^{k-1} \\
&= \frac{1}{k^2} \frac{n}{\log n + \omega(n)} \left(\frac{en}{k} k \frac{\log n + \omega(n)}{n} e^{-\frac{\log n + \omega(n)}{n}(n-k)}\right)^k \\
&= \frac{1}{k^2} O\left(\frac{n}{\log n}\right) \left(O(\log n)(1 + o(1)) \frac{1}{n} e^{k \frac{\log n}{n}}\right)^k.
\end{aligned}$$

When n is large enough, and $2 \le k \le 5$, we have

$$\begin{aligned}
\frac{1}{k^2} O\left(\frac{n}{\log n}\right) \left(O(\log n)(1 + o(1)) \frac{1}{n} e^{k \frac{\log n}{n}}\right)^k &= O\left(\frac{n}{\log n}\right) \left(O(\log n) \frac{1}{n}\right)^k \\
&= O(n \log n) \frac{1}{n^2} \\
&\ll n^{-\frac{1}{2}}.
\end{aligned}$$

In terms of $6 \le k \le \lfloor \frac{n}{2} \rfloor$, we have

$$\begin{aligned}
\frac{1}{k^2} O\left(\frac{n}{\log n}\right) \left(O(\log n)(1 + o(1)) \frac{1}{n} e^{k \frac{\log n}{n}}\right)^k &= O\left(\frac{n}{\log n}\right) \left(O(\log n) \frac{1}{n} \sqrt{n}\right)^k \\
&= O\left(n (\log n)^5\right) n^{-3} \\
&\ll n^{-\frac{3}{2}}.
\end{aligned}$$

Combining the above, we have

$$\sum_{k=2}^{\lfloor \frac{n}{2} \rfloor} \mathbb{E}[X_k] \ll 4 \cdot n^{-\frac{1}{2}} + \frac{n}{2} n^{-\frac{3}{2}} = \Theta(n^{-\frac{1}{2}}) \to 0$$

[1] the first one is the standard upper bound for binomial coefficients

as $n \to \infty$. Then, by Markov's inequality, we have

$$\mathbb{P}\left(\sum_{k=2}^{\lfloor \frac{n}{2} \rfloor} X_k \geq 1\right) \leq \frac{\sum_{k=2}^{\lfloor \frac{n}{2} \rfloor} \mathbb{E}\left[X_k\right]}{1} \to 0$$

as $n \to \infty$.

Recall that

$$\mathbb{P}\left(G(n,p) \text{ is connected}\right) = \mathbb{P}\left(\bigcap_{k=1}^{\lfloor \frac{n}{2} \rfloor} \{G(n,p) \text{ has no } k\text{-component}\}\right).$$

Equivalently, the probability

$$\begin{aligned}
\mathbb{P}\left(G(n,p) \text{ is not connected}\right) &= \mathbb{P}\left(\bigcup_{k=1}^{\lfloor \frac{n}{2} \rfloor} \{X_k \geq 1\}\right) \\
&\leq \mathbb{P}\left(X_1 \geq 1\right) + \mathbb{P}\left(\sum_{k=2}^{\lfloor \frac{n}{2} \rfloor} X_k \geq 1\right) \\
&= \mathbb{P}\left(X_1 \geq 1\right) + o(1).
\end{aligned}$$

Since $\mathbb{P}\left(G(n,p) \text{ is not connected}\right) \geq \mathbb{P}\left(X_1 \geq 1\right)$, we have

$$\mathbb{P}\left(G(n,p) \text{ is not connected}\right) \to \mathbb{P}\left(X_1 \geq 1\right)$$

as $n \to \infty$. Equivalently, we have

$$\mathbb{P}\left(G(n,p) \text{ is connected}\right) \to \mathbb{P}\left(X_1 = 0\right).$$

By theorem 3.8, we have

$$\mathbb{P}\left(X_1 = 0\right) \to e^{-e^{-c}}$$

as $n \to \infty$ when $p = \frac{\log n + \omega(n)}{n}$ with $\omega(n) \to c < \infty$ and our proof completes. □

Theorem 3.10. *Let $\mathcal{F}$ be the property that the random graph is connected. Then $p^* = \frac{\log n}{n}$ is a threshold for $\mathcal{F}$.*

3.3 Hamilton cycles and perfect matchings

A Hamiltonian cycle is a cycle that passes each vertex exactly once. A graph is Hamiltonian if it contains a Hamilton cycle. The threshold concerning Hamilton cycles was first given by Pósa [25]. For a cycle on n labelled vertices, we can get $2n$

different permutations by different choices of the first elements and the direction. It follows that there are $\frac{n!}{2n} = \frac{(n-1)!}{2}$ possible Hamilton cycles in a graph on $[n]$. Let X be the number of Hamilton cycles in the random graph $G(n, p)$. By the first moment method, we get

$$\mathbb{E}[X] = \frac{(n-1)!}{2} p^n$$

and it follows that n^{-1} is a lower bound of the threshold. It is obvious that n^{-1} is not a threshold since a Hamilton cycle in a graph implies that the graph is connected and n^{-1} is much less than the threshold for connectedness.

Pósa provided an observation of the longest path in graphs. Let P be one of the longest paths in graph G on $[n]$ with end vertices u, v. We also use P to denote the vertices set or edges set of the path if there is no ambiguity. Since P is the longest path in G, the end vertices have no neighbourhood in $[n]\backslash P$. If there is some $x \in P$

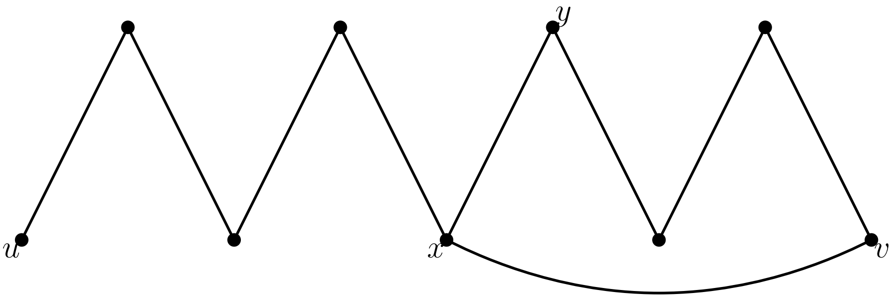


Figure 3.2: Path P

such that xv is an edge in G, we have a transformation to get a new path P_{uy} like fig. 3.2 to fig. 3.3. Notice that u is the common end vertices of P and P_{uy} and P also can be gotten from P_{uy} by a similar transformation. Allowing transformation several

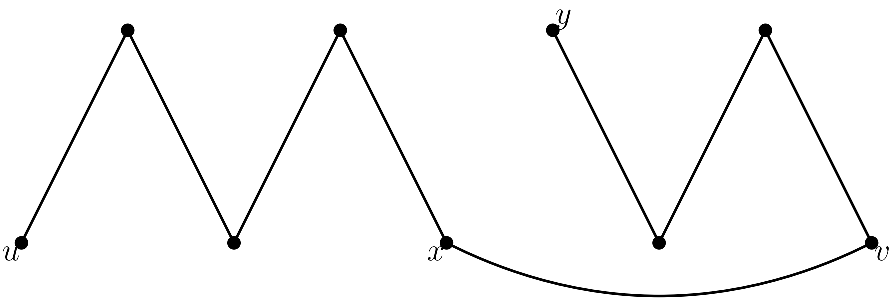


Figure 3.3: Path P_{uy}

times successively, let A be the set of vertices a such that there is a path P_{ua} from u to a can be obtained from P by transformation with common end vertex u. Let B

be the set of vertices such that they are different from u or any vertices in A and are not adjacent to any vertex in A on the original path P. Thus all vertices in $[n]\backslash P$ are in B. We have the following result.

Lemma 3.11. *There is no edge between A and B.*

Proof. For $\forall b \in [n]\backslash P$, suppose to the contrast that there is an $a \in A$ such that a, b are connected by an edge in G. It follows that there is a path P_{ua} from u to a that can be prolonged by adding the new edge ab which is contradictory to the assumption that P is one of the longest paths.

For $\forall b \in P \cap B$, we will show that if b is adjacent to some $a \in A$, then there is a $\tilde{a} \in A$ such that $\tilde{a}b \in P$. Suppose to the contrast that α, β are neighbors of b on P with $\alpha, \beta \notin A$ and b is adjacent to $a \in A$. Consider the path P_{ua} from u to a (since $a \in A$, we can always obtain P_{ua} from P by a series of transformation). Let $\{s, t\}$ be the neighbours of b on P_{ua}. We have $\{s, t\} = \{\alpha, \beta\}$, otherwise $b\alpha$, or $b\beta$ is deleted by one transformation which implies that one of α, β, b belongs to A which is a contradiction. Since $\{s, t\} = \{\alpha, \beta\}$ and b is adjacent to a, it follows that one of $\{\alpha, \beta\}$ belongs to A by a transformation from P_{ua} which is also a contradiction.

Therefore, there is no edge between A and B. □

Since for each $a \in A$, a has no more than two neighbours on the path P, there are at most $2|A| + |A| + 1$ elements not in B. It follows that

$$|B| \geq n - 3|A| - 1.$$

Through the longest path, we can construct such a structure in a graph that there are two disjoint sets of vertices and there are no edges between them. In the random graph $G\left(n, \frac{c \log n}{n}\right)$, such structures turn up with low probability if the size of one set is small and c is large enough.

Lemma 3.12. *Let $p = \frac{c \log n}{n}$. In the random graph $G(n, p)$, the probability that there are two disjoint nonempty sets of vertices S, T with $|S| = k \leq \lfloor \frac{n}{4} \rfloor$ and $|T| = n-3k-1$ such that there are no edges between S and T tends to 0 as $n \to \infty$ when c is large enough.*

Proof. For $1 \leq k \leq \lfloor \frac{n}{4} \rfloor$, let $\mathcal{E}_{S,T;k}$ be the event that for given two disjoint sets of vertices S, T with $|S| = k$ and $|T| = n - 3k - 1$, there are no edges between S and T. The event occurs if and only if the specific $k(n - 3k - 1)$ edges are absent. It follows that

$$\mathbb{P}(\mathcal{E}_{S,T;k}) = (1 - p)^{k(n-3k-1)}.$$

Let X_k be the number of different pairs S and T such that the event $\mathcal{E}_{S,T;k}$ occur and the expectation of X_k

$$\mathbb{E}\left[X_k\right] = \binom{n}{k}\binom{n-k}{n-3k-1}(1-p)^{k(n-3k-1)}.$$

By the estimate $1 - x \leq e^{-x}$, we have

$$\begin{aligned}\binom{n}{k}\binom{n-k}{n-3k-1}(1-p)^{k(n-3k-1)} &\leq \binom{n}{k}\binom{n}{2k+1}e^{-pk(n-3k-1)}\\ &\leq n^{3k+1}e^{-\frac{c\log n}{n}(n-3k-1)k}\\ &\leq \left(n^4 n^{-\frac{c}{5}}\right)^k.\end{aligned}$$

Set $c = 30$, and we have $\mathbb{E}\left[X_k\right] \leq O(n^{-2})$. Let $\mathcal{E}$ be the event that for some $1 \leq k \leq \lfloor \frac{n}{4} \rfloor$, there are two disjoint sets of vertices S, T with $|S| = k$ and $|T| = n - 3k - 1$ such that there are no edges between S and T. It follows that

$$\mathbb{P}(\mathcal{E}) = \mathbb{P}\left(\bigcup_{k=1}^{\lfloor \frac{n}{4} \rfloor}\{X_k \geq 1\}\right) \leq \sum_{k=1}^{\lfloor \frac{n}{4} \rfloor}\mathbb{P}(X_k \geq 1).$$

By Markov's inequality, we have that

$$\sum_{k=1}^{\lfloor \frac{n}{4} \rfloor}\mathbb{P}(X_k \geq 1) \leq \frac{\sum_{k=1}^{\lfloor \frac{n}{4} \rfloor}\mathbb{E}\left[X_k\right]}{1} \leq \frac{n}{4}n^{-2} \to 0$$

as $n \to \infty$ and our proof complete. □

By this lemma, we will show there is a Hamilton path in $G\left(n, \frac{c\log n}{n}\right)$ with high probability when c is large enough.

Theorem 3.13. *Let $p = \frac{c\log n}{n}$ with c large enough. In the random graph $G(n,p)$, the probability that there is a Hamilton path tends to 1 as $n \to \infty$.*

Proof. We define events as follows

- $\mathcal{E}$: for some $1 \leq k \leq \lfloor \frac{n}{4} \rfloor$, there exist two disjoint sets of vertices S, T with $|S| = k$ and $|T| = n - 3k - 1$ such that there are no edges between S and T;
- $\mathcal{L}_x$: for given vertex x, any longest path in $G(n,p)$ passes through x.

We use $\overline{\mathcal{A}}$ to denote the complement of the event $\mathcal{A}$. Suppose $\overline{\mathcal{L}_x}$ occurs, that is, there exists a longest path on G that doesn't pass x. Let's consider one of the longest paths of G in $G - x$ and call it P. We define the vertices set A, B as we state above. If

$|A| \leq \lfloor \frac{n}{4} \rfloor$, we have $|B| \geq n - 3k - 1$ which implies the event $\mathcal{E}$ occurs. In the case of $|A| > \lfloor \frac{n}{4} \rfloor$, we notice that x isn't adjacent to any $a \in A$ otherwise P can be prolonged. The possibility

$$\begin{aligned}\mathbb{P}\left(x \text{ isn't adjacent to any } a \in A\right) &\leq \left(1 - \frac{c \log n}{n}\right)^{\frac{n}{4}} \\ &\leq e^{-\frac{n}{4}\frac{c \log n}{n}} \\ &= n^{-\frac{c}{4}}.\end{aligned}$$

It follows that $\mathbb{P}\left(\overline{\mathcal{L}_x} \cap \overline{\mathcal{E}}\right) \leq n^{-\frac{c}{4}}$ which implies that $\mathbb{P}\left(\bigcup_x (\overline{\mathcal{L}_x} \cap \overline{\mathcal{E}})\right) \leq n \cdot n^{-\frac{c}{4}}$. By the lemma 3.12, the probability of event $\mathcal{E}$ is $o(1)$ when c is large enough. It is easy to check that

$$\begin{aligned}\mathbb{P}\left(\bigcup_x \overline{\mathcal{L}_x}\right) &= \mathbb{P}\left(\bigcup_x \overline{\mathcal{L}_x} \cap \overline{\mathcal{E}}\right) + \mathbb{P}\left(\bigcup_x \overline{\mathcal{L}_x} \cap \mathcal{E}\right) \\ &\leq \mathbb{P}\left(\bigcup_x (\overline{\mathcal{L}_x} \cap \overline{\mathcal{E}})\right) + \mathbb{P}\left(\mathcal{E}\right) \\ &\leq n^{-\frac{c}{4}+1} + o(1) \to 0\end{aligned}$$

as $n \to \infty$ for large c which implies that

$$\mathbb{P}\left(\bigcap_x \mathcal{L}_x\right) \to 1$$

as $n \to \infty$. Equivalently, the longest path in $G(n, p)$ passes every vertex with high probability which means there is a Hamilton path in $G(n, p)$ with high probability. □

Finally, we use the coupling technique and show that there is a Hamilton cycle with high probability.

Theorem 3.14. *Let $p = \frac{C \log n}{n}$ with C large enough. In the random graph $G(n, p)$, the probability that there is a Hamilton cycle tends to 1 as $n \to \infty$.*

Proof. Let $p_1 = \frac{c \log n}{n}$ with c large enough and $p_2 = \frac{\log n}{n}$. Let G_1, G_2 are independent copies of $G(n, p_1), G(n, p_2)$ respectively. Let G be the graph obtained from G_1 by coupling G_2. We define events as follows

- $\mathcal{E}$: for some $1 \leq k \leq \lfloor \frac{n}{4} \rfloor$, there exist two disjoint sets of vertices S, T with $|S| = k$ and $|T| = n - 3k - 1$ such that there are no edges between S and T in G_1;

- $\mathcal{M}$: the random graph G_1 contains a Hamilton path;
- $\mathcal{H}$: the random graph G contains a Hamilton cycle.

Suppose $\mathcal{M}$ occurs. Let P be a Hamilton path in G_1. As stated above, we have a fixed end vertex u and two vertices sets A, B such that there are no edges between A and B. If $|A| \le \lfloor \frac{n}{4} \rfloor$, then the event $\mathcal{E}$ occurs. In terms of $|A| > \lfloor \frac{n}{4} \rfloor$, if there is no Hamilton cycle in G, then there are no edges between u and A in G_2. The probability of this event is less than

$$\left(1 - \frac{\log n}{n}\right)^{\frac{n}{4}} \le e^{-\frac{n}{4}\frac{\log n}{n}} \le n^{-\frac{1}{4}}.$$

Using the law of total probability, we have

$$\begin{aligned}
\mathbb{P}\left(\overline{\mathcal{H}}\right) &= \mathbb{P}\left(\overline{\mathcal{H}} \cap \mathcal{M}\right) + \mathbb{P}\left(\overline{\mathcal{H}} \cap \overline{\mathcal{M}}\right) \\
&\le \mathbb{P}\left(\overline{\mathcal{H}} \cap \mathcal{M}\right) + \mathbb{P}\left(\overline{\mathcal{M}}\right) \\
&= \mathbb{P}\left(\overline{\mathcal{H}} \cap \mathcal{M} \cap \mathcal{E}\right) + \mathbb{P}\left(\overline{\mathcal{H}} \cap \mathcal{M} \cap \overline{\mathcal{E}}\right) + \mathbb{P}\left(\overline{\mathcal{M}}\right) \\
&\le \mathbb{P}\left(\overline{\mathcal{H}} \cap \mathcal{M} \cap \overline{\mathcal{E}}\right) + \mathbb{P}\left(\mathcal{E}\right) + \mathbb{P}\left(\overline{\mathcal{M}}\right).
\end{aligned}$$

By the theorem 3.13 and lemma 3.12, we have

$$\mathbb{P}\left(\overline{\mathcal{H}} \cap \mathcal{M} \cap \overline{\mathcal{E}}\right) + \mathbb{P}\left(\mathcal{E}\right) + \mathbb{P}\left(\overline{\mathcal{M}}\right) \le n^{-\frac{1}{4}} + o(1) + o(1) \to 0$$

as $n \to \infty$ when c is large enough. It follows that $\mathbb{P}\left(\mathcal{H}\right) \to 1$ as $n \to \infty$. Notice that

$$\begin{aligned}
\left(1 - \frac{c\log n}{n}\right)\left(1 - \frac{\log n}{n}\right) &= 1 - (c+1)\frac{\log n}{n} + c\left(\frac{\log n}{n}\right)^2 \\
&\ge 1 - (c+1)\frac{\log n}{n}.
\end{aligned}$$

The random graph G has a distribution as $G(n, p')$ with $p' \le (c+1)\frac{\log n}{n}$. Take $C = c+1$ and $p = \frac{C\log n}{n}$. We have

$$\mathbb{P}\left(G\left(n, p\right) \text{ has a Hamilton cycle}\right) \ge \mathbb{P}\left(G(n, p') \text{ has a Hamilton cycle}\right) \to 1$$

as $n \to \infty$. □

Recall that the threshold for the property that there is no isolated vertex in $G(n, p)$ is $\frac{\log n}{n}$. Since a random graph containing a Hamilton cycle can imply that it has no isolated vertex, we have a lower bound of the threshold for the Hamilton cycle $\frac{\log n}{n}$. As we show above, $G(n, \frac{C\log n}{n})$ contains a Hamilton cycle with high probability. Combining them, it follows that $\frac{\log n}{n}$ is a threshold for the Hamilton cycle.

Theorem 3.15. *Let $\mathcal{F}$ be the property that the random graph contains a Hamilton cycle. One of thresholds for $\mathcal{F}$ is $\frac{\log n}{n}$.*

After Pósa proved this result, Komlós and Szemerédi showed that the sharp threshold for the existence of Hamilton cycles is $\frac{\log n+\log\log n}{n}$ in [21]. Bollobás [4] improved this to a hitting time result. If we add edges one by one then with probability tending to 1 the graph becomes Hamiltonian at exactly the point when the minimum degree becomes at least 2.

Notice that the existence of a Hamilton cycle implies the existence of a perfect matching for graphs on even vertices. Meanwhile, a perfect matching implies that there is no isolated vertex. We can get the threshold for perfect matchings directly.

Theorem 3.16. *Let $\mathcal{F}$ be the property that the random graph contains a perfect matching. One of thresholds for $\mathcal{F}$ is $\frac{\log n}{n}$.*

This result was proven first by Erdős and Rényi in [9]. Actually, they proved that $\frac{\log n}{n}$ is a sharp threshold. Bollobás and Frieze improved their work and we refer interested readers to [5].

Interestingly, if we try to use the first moment method to consider the threshold for the existence of perfect matchings, we will find that n^{-1} is a natural lower bound since there are $\frac{n!}{2^{n/2}\frac{n}{2}!}$ copies of perfect matchings and each of them contains exactly $\frac{n}{2}$ edges. Recall the examples of connectedness and Hamilton cycles. All of them show that there is just a logarithmic gap between the lower bound and the threshold. These examples motivated Kahn and Kalai to propose their conjecture.

Chapter 4

Kahn-Kalai Conjecture

In this chapter, we will introduce the Kahn-Kalai conjecture which gives an easy but good estimate of the thresholds for increasing properties. As we mentioned in the last chapter, the first moment method provides a lower bound of the threshold. However, for the upper bound, there is no good estimate until 2007. Kahn and Kalai [19] proposed their conjecture saying the upper bound is not far from the trivial lower bound. Park and Pham [24] proved this conjecture in 2022. Before that, Frankston, Kahn, Narayanan and Park [12] proved a weaker theorem in 2021 called the fractional Kahn-Kalai conjecture which has many significant applications. This conjecture was modified by Talagrand [29]. The proofs of these two conjectures are similar by iteration. Park and Pham introduced a technical notion, *minimum fragment*, which enables them to prove the Kahn-Kalai conjecture elegantly. This technique can also be applied in [12] to simplify the proof.

We will introduce the Kahn-Kalai conjecture including the fractional version conjecture with applications and leave detailed proof to the next chapter. The applications reference [16] by Frieze and Karoński.

4.1 Kahn-Kalai Conjecture

Recall that the lower bound of the threshold for the existence of spanning trees, Hamilton cycles, and perfect matchings provided by the first moment method are all at least n^{-1}. Meanwhile, they have the same threshold $\frac{\log n}{n}$. All of them show that there is just a logarithmic gap between the lower bound and the threshold. These examples are part of the motivations for Kahn and Kalai to propose their conjecture. Now we state this conjecture formally.

As we mentioned above, a graph property is a subset of the power set $2^{\binom{[n]}{2}}$. Generally, let X be the ground set and a non-trivial increasing property $\mathcal{F}$ is a subset of 2^X such that

$$A \subseteq B \text{ and } A \in \mathcal{F} \Rightarrow B \in \mathcal{F}$$

and $\mathcal{F} \neq 2^X, \emptyset$. We use X_p to denote the random subset of X and each element of X is contained in X_p with probability p independently for $p \in [0,1]$. Correspondingly, X_m is a uniform m-element subset of X for $m \leq |X|$. If the random subset has the property $\mathcal{F}$, we write $X_p \in \mathcal{F}$ or $X_m \in \mathcal{F}$. Let μ_p be the product measure on 2^X given by

$$\mu_p(S) = p^{|S|}(1-p)^{|X \setminus S|}.$$

Different elements in $\mathcal{F}$ are mutually exclusive, so we have

$$\mathbb{P}(X_p \in \mathcal{F}) = \sum_{S \in \mathcal{F}} \mathbb{P}(X_p = S) = \sum_{S \in \mathcal{F}} p^{|S|}(1-p)^{|X \setminus S|}.$$

Let $\mu_p(\mathcal{F}) = \sum_{S \in \mathcal{F}} p^{|S|}(1-p)^{|X \setminus S|}$. Similar to the proposition 2.7, we have $\mu_p(\mathcal{F})$ increases as p increases.

Recall that the threshold is not unique in the original definition. The proof of theorem 2.14 offered us a insight that $\forall \epsilon \in (0,1)$, $p(\epsilon)$ defined by

$$\mathbb{P}(G(n, p(\epsilon)) \text{ has } \mathcal{F}) = \epsilon$$

is a threshold. By this observation, we redefine the threshold uniquely as follows.

Definition 4.1. Let $\mathcal{F}$ be a non-trivial increasing property. The function $p_c(\mathcal{F})$ is the threshold for $\mathcal{F}$ if

$$\frac{1}{2} = \mu_{p_c(\mathcal{F})}(\mathcal{F}).$$

Recall that the first moment method provides a lower bound in chapter 3. Especially motivated by the examples 2.10 and 3.3, we write this lower bound formally.

Definition 4.2 (expectation threshold in graph)**.** Let H be a subgraph of the complete graph K_n. The expectation threshold for the event that $G(n,p)$ contains H as a subgraph is defined by the following equation

$$p_E(H) = \min\{p : \mathbb{E}\left[\text{number of } H' \text{ in } G(n,p)\right] \geq \frac{1}{2}^{1} \quad \forall H' \subseteq H\}.$$

This definition has a clear meaning and we can rewrite it as

$$p_E(H) = \max\{p : \mathbb{E}\left[\text{number of } H' \text{ in } G(n,p)\right] \leq \frac{1}{2} \quad \exists H' \subseteq H\}.$$

[1] we use $\frac{1}{2}$ here to keep the definition compatible with abstract expectation thrshold

Since $\mathbb{E}$ [number of H' in $G(n,p)$] $= \sum_{\text{copies of } H'} p^{|E(H')|}$, Kahn and Kalai generalise this concept into more abstract settings in [19]. Let $\mathcal{G}$ be a subset of 2^X and if

$$\mathcal{F} \subseteq \langle \mathcal{G} \rangle := \bigcup_{S \in \mathcal{G}} \{T : T \supseteq S\},$$

we say that $\mathcal{G}$ is a cover of $\mathcal{F}$. It is easy to find that the set of all copies of H' is a cover of the property that $G(n,p)$ contains H as a subgraph. Following [29], we say $\mathcal{F}$ is *p-small* if there is a cover $\mathcal{G}$ of $\mathcal{F}$ such that

$$\sum_{S \in \mathcal{G}} p^{|S|} \leq \frac{1}{2}.$$

Definition 4.3 (abstract expectation threshold)**.** $q(\mathcal{F})$ is the expectation threshold of $\mathcal{F}$ if it is the maximum p such that $\mathcal{F}$ is *p-small.*

It is easy to check that $q(\mathcal{F})$ is a lower bound of $p_c(\mathcal{F})$. We refer interested readers to the proof of the proposition 4.8.

Kahn and Kalai conjectured that there is a universal constant K such that

$$p_c(\mathcal{F}) \leq K q(\mathcal{F}) \log |X|.$$

In Park and Pham's proof, they slightly improve this result. Let $\ell_1(\mathcal{F})$ be the maximum size of the minimal element of $\mathcal{F}$ and $\ell(\mathcal{F}) = \max\{2, \ell_1(\mathcal{F})\}$.

Theorem 4.4 (Kahn-Kalai Conjecture)**.** *There is a universal constant K such that for every set X and non-trivial increasing property $\mathcal{F} \subseteq 2^X$,*

$$p_c(\mathcal{F}) \leq K q(\mathcal{F}) \log \ell(\mathcal{F}).$$

In Park and Pham's proof, they first reduced the Kahn-Kalai conjecture to the following theorem 4.5. Let $\mathcal{H}$ be a subset of 2^X. We say $\mathcal{H}$ is ℓ-bounded if the size of any element in $\mathcal{H}$ is no more than ℓ.

Theorem 4.5. *Let $\ell \geq 2$. There is a universal constant L such that for any nonempty ℓ-bounded subset $\mathcal{H}$ of 2^X that is* **not** *p-small, a uniformly random $((Lp\log\ell)|X|)$-element subset of X belongs to $\langle\mathcal{H}\rangle$ with probability tending to 1 as $\ell \to \infty$.*

Starting from this theorem, let $\mathcal{H}$ be the set of the minimal elements in property $\mathcal{F}$. Then we have $\langle\mathcal{H}\rangle = \mathcal{F}$ and by Chebyshev's inequality, we can get the theorem 4.4. We leave the proof of reduction to the next chapter. This theorem is not easy to apply, since it is not easy to show a subset **not** *p-small.* We introduce the following fractional Kahn-Kalai conjecture which is convenient to apply.

4.2 Fractional Kanh-Kalai Conjecture

The remarkable breakthrough before Park and Pham's proof is the fractional version of the Kahn-Kalai conjecture. Frankston, Kahn, Narayanan and Park gave their proof with many applications in [12]. This conjecture was proposed by Talagrand [29] and he introduced the fractional expectation threshold by relaxing notion *p-small* to *weakly p-small* as follows.

For an increasing property $\mathcal{F}$ in X, we say $\mathcal{F}$ is *weakly p-small* if there is a $g : 2^X \to [0,1]$ such that

- $\sum_{S \subseteq T} g(S) \geq 1 \ \ \forall T \in \mathcal{F}$ and,
- $\sum_{S \subseteq X} g(S)p^{|S|} \leq \frac{1}{2}$.

Definition 4.6. $q_f(\mathcal{F})$ is the fractional expectation-threshold of $\mathcal{F}$ if it is the maximum p such that $\mathcal{F}$ is *weakly p-small.*

Theorem 4.7 (Fractional Kahn-Kalai Conjecture)**.** *There is a universal constant K such that for every set X and non-trivial increasing property $\mathcal{F} \subseteq 2^X$,*

$$p_c(\mathcal{F}) \leq K q_f(\mathcal{F}) \log \ell(\mathcal{F}).$$

The following proposition shows that the fractional Kahn-Kalai conjecture can be implied from the Kahn-Kalai conjecture.

Proposition 4.8. *For every finite set X and non-trivial increasing property $\mathcal{F}$ in X, we have*

$$q(\mathcal{F}) \leq q_f(\mathcal{F}) \leq p_c(\mathcal{F}).$$

Proof. To show the first inequality right, we just need to show that if $\mathcal{F}$ is *p-small*, then $\mathcal{F}$ is *weakly p-small.* Supposing that $\mathcal{F}$ is *p-small*, there is a $\mathcal{G} \subseteq 2^X$ such that $\mathcal{F} \subseteq \langle \mathcal{G} \rangle$ and

$$\sum_{S \in \mathcal{G}} p^{|S|} \leq \frac{1}{2}.$$

Let $g : 2^X \to [0,1]$ be the function as follows

$$g(S) = \begin{cases} 1 & \text{if } S \in \mathcal{G} \\ 0 & \text{if } S \notin \mathcal{G}. \end{cases}$$

Since $\mathcal{G}$ is a cover of $\mathcal{F}$, for each $T \in \mathcal{F}$, there is a $S \in \mathcal{G}$ such that $S \subseteq T$. Thus, we have

$$\sum_{S \subseteq T} g(S) \geq 1 \ \ \forall T \in \mathcal{F}.$$

Meanwhile, we have

$$\sum_{S\subseteq X} g(S)p^{|S|} = \sum_{S\in\mathcal{G}} p^{|S|} \leq \frac{1}{2}$$

which shows that $\mathcal{F}$ is *weakly p-small*.

In terms of the second inequality, suppose that $\mathcal{F}$ is *weakly p-small* and we have

$$\begin{aligned}
\mu_p(\mathcal{F}) = \sum_{T\in\mathcal{F}} \mu_p(T) &\leq \sum_{T\in\mathcal{F}} \mu_p(T) \sum_{S\subseteq T} g(S) \\
&\leq \sum_{T\subseteq X} \mu_p(T) \sum_{S\subseteq T} g(S) \\
&= \sum_{S\subseteq X} g(S) \sum_{S\subseteq T} \mu_p(T) = \sum_{S\subseteq X} g(S)p^{|S|} \leq \frac{1}{2}
\end{aligned}$$

which implies that $q_f(\mathcal{F}) \leq p_c(\mathcal{F})$. □

Talagrand also conjectured that every *weakly p-small* $\mathcal{F}$ is (p/K)*-small* for some universal constant K which would serve as a bridge to the Kahn-Kalai conjecture from the fractional version of it. Unfortunately, this conjecture hasn't been solved yet though the Kahn-Kalai Ccnjecture has been proven.

Weakly p-small offers us an important property about *spread*. We say a probability measure ν on 2^X is *q-spread* if

$$\nu(\langle S\rangle) \leq q^{|S|} \quad \forall \ S \subseteq X.$$

Meanwhile, we say a subset $\mathcal{H}$ of 2^X is *q-spread* if

$$|\mathcal{H} \cap \langle S\rangle\,| \leq q^{|S|}|\mathcal{H}| \quad \forall \ S \subseteq X.$$

It is easy to check that $\mathcal{H}$ is *q-spread* if and only if the uniform measure on $\mathcal{H}$ is *q-spread*. As Talagrand showed, the following proposition links the notions of *weakly p-small* and *spread*.

Proposition 4.9. *If an increasing property $\mathcal{F}$ in X is* **not** *weakly p-small, there is a $(2p)$-spread probability measure on 2^X supported on $\mathcal{F}$.*

From a $(2p)$*-spread* probability measure supported on $\mathcal{F}$, we can get a $(2p)$*-spread* measure on $\mathcal{G}$ which is the set of minimal elements of $\mathcal{F}$ (just transfer the weight of $S \in \mathcal{F}$ to the minimal element $T \subseteq S$). We can relax the $(2p)$*-spread* measure to $(2p+\epsilon)$*-spread* measure to make it take values in $\mathbb{Q}$ where $\epsilon > 0$ can be arbitrarily small. Finally, we can get a uniform $(2p+\epsilon)$*-spread* measure on $\mathcal{H}$ whose elements are copies of elements of $\mathcal{G}$ (allowing multi-element) which implies that $\mathcal{H}$ is $(2p+\epsilon)$-*spread.*

Similar to theorem 4.5 for the Kahn-Kalai conjecture, theorem 4.7 is reduced to the following theorem which is more convenient to apply.

Theorem 4.10. *There is a universal constant L such that for any nonempty ℓ-bounded and p-spread subset $\mathcal{H}$ of 2^X, a uniformly random $((Lp\log\ell)|X|)$-element subset of X belongs to $\langle\mathcal{H}\rangle$ with probability $1-o_{\ell\to\infty}$.*

Now we apply this theorem to show some results. We will begin with Hamilton cycles and perfect matchings and then generalise them to hypergraphs. Finally, we will show a result about bounded-degree spanning trees which is very difficult historically.

Perfect matchings in graphs: Recall the example of perfect matchings. For a graph on n vertices, suppose $2 \mid n$. There are $\frac{n!}{\frac{n}{2}!2^{\frac{n}{2}}}$ possible perfect matchings. Let $\mathcal{H}$ be the subset of $2^{\binom{[n]}{2}}$ and each of its elements exactly corresponds to a perfect matching. The size of each elements in $\mathcal{H}$ is $\frac{n}{2}$ and we say $\mathcal{H}$ is $\frac{n}{2}$-uniform which implies that $\mathcal{H}$ is $\frac{n}{2}$-bounded. We will show that $\mathcal{H}$ is $\frac{C}{n}$-*spread* for some large C. Recall Stirling's approximation

$$\sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{\frac{1}{12n+1}} \le n! \le \sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{\frac{1}{12n}}.$$

For $\forall S \subseteq X$ with $|S| = s$, we have

$$\begin{aligned}\frac{|\mathcal{H}\cap\langle S\rangle|}{|\mathcal{H}|} &\le \frac{\frac{n}{2}!2^{\frac{n}{2}}}{n!}\frac{(n-2s)!}{\left(\frac{n}{2}-s\right)!2^{n/2-s}}\\ &\le e\frac{\sqrt{2\pi n/2}}{\sqrt{2\pi n}}\frac{(n/2e)^{n/2}2^{n/2}}{(n/e)^n}\frac{\sqrt{2\pi(n-2s)}}{\sqrt{2\pi(n/2-s)}}\frac{((n-2s)/e)^{n-2s}}{((n-2s)/2e)^{n/2-s}2^{n/2-s}}\\ &\le e\left(\frac{e}{n-2s}\right)^s\\ &\le \left(\frac{C}{n}\right)^s.\end{aligned}$$

Notice that $\left(\frac{C}{n}\log\frac{n}{2}\right)\binom{n}{2} \le L'n\log n$ for some large L'. Using the theorem 4.10, there exists a constant L such that a $(Ln\log n)$-element subset uniformly chosen from $\binom{[n]}{2}$ corresponds to a graph containing a perfect matching with high probability. As we mentioned above, this result was first proved by Erdős and Rényi in [9].

Shamir's problem: This problem asks when the random hypergraph has a perfect matching. The model of random hypergraphs is similar to random graphs. A hypergraph is k-uniform if each edge of the hypergraph consists of k vertices. For a k-uniform random hypergraph on n vertices (for the sake of convenience, we say the vertices set $[n]$), each possible edge in $\binom{[n]}{k}$ turns up with probability p independently. We use $H(n,p;k)$ to denote such a random hypergraph. It is necessary to suppose $k \mid n$ to have perfect matchings. Let $r = \frac{n}{k}$ and $\mathcal{H}$ be the subset of 2^X where $X = \binom{[n]}{k}$ and each of its elements corresponds to a perfect matching in the k-uniform hypergraph. Since each perfect matching corresponds to a partition that divides the n vertices into r disjoint k-element parts, there are $\frac{n!}{r!(k!)^r}$ elements in $\mathcal{H}$ and $\mathcal{H}$ is r-uniform. To apply theorem 4.10, we show that there is a constant C large enough such that $\mathcal{H}$ is $Cn^{-(k-1)}$-*spread*. For $\forall S \subseteq X$ with $|S| = s$, using Stirling's formula we have

$$\begin{aligned}
\frac{|\mathcal{H} \cap \langle S \rangle|}{|\mathcal{H}|} &= \frac{r!(k!)^r}{n!} \frac{(n-sk)!}{(r-s)!(k!)^{r-s}} \\
&\le e \frac{\sqrt{2\pi n/k}}{\sqrt{2\pi n}} \frac{(n/ek)^{n/k}(k!)^{n/k}}{(n/e)^n} \frac{\sqrt{2\pi(n-sk)}}{\sqrt{2\pi(n-sk)/k}} \frac{((n-sk)/e)^{n-sk}}{((n-sk)/ke)^{(n-sk)/k}(k!)^{n/k-s}} \\
&\le e \left(\frac{e(k-1)!}{(n-sk)^{k-1}} \right)^s \\
&\le \left(\frac{C}{n^{k-1}} \right)^s
\end{aligned}$$

Using theorem 4.10, there exists a constant L such that a uniformly random $(Ln \log n)$-element subset of X belongs to $\langle \mathcal{H} \rangle$ with high probability. This result was first proved by Johansson, Kahn, and Vu [17].

Hamilton cycles in graphs: Recall the example of Hamilton cycles in the last chapter. There are $\frac{(n-1)!}{2}$ possible Hamilton cycles in a random graph. Let $X = \binom{[n]}{2}$ and $\mathcal{H}$ be the subset of 2^X, each of whose elements corresponds to a Hamilton cycle on n vertices. Notice that $\mathcal{H}$ is n-uniform. We will show that $\mathcal{H}$ is $\frac{C}{n}$-*spread* for some large C. For $\forall S \subseteq X$ with $|S| = s$, using Stirling's approximation we have

$$\begin{aligned}\frac{|\mathcal{H} \cap \langle S \rangle|}{|\mathcal{H}|} &\leq \frac{2^s (n-s)!/2(n-s)}{n!/2n} \\ &\leq 2^s e \frac{\sqrt{2\pi(n-s-1)}}{\sqrt{2\pi(n-1)}} \frac{((n-s-1)/e)^{n-s-1}}{((n-1)/e)^{n-1}} \\ &\leq 2^s e^{s+1} \left(\frac{1}{n-1}\right)^s \\ &\leq \left(\frac{C}{n}\right)^s\end{aligned}$$

where C is large enough. Then by the theorem 4.10, we have that a random uniform $Ln \log n$-element subset of X belongs to $\langle \mathcal{H} \rangle$ with high probability. As we mentioned above, Pósa proved this result in [25].

Loose Hamilton cycles: In k-uniform random hypergraph on n vertices. A loose Hamilton cycle is a collection of $r = \frac{n}{k-1}$ (suppose $(k-1) \mid n$) edges such that for some cyclic order of $[n]$, each edge contains consecutive k vertices and each pair of consecutive edges has exactly one common vertex. In the fig. 4.1, we give an example

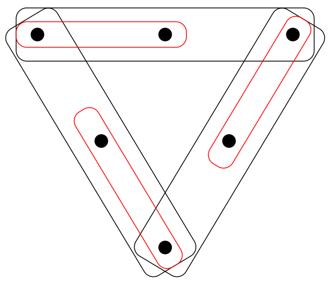


Figure 4.1: Loose Hamilton cycle

of 3-uniform loose Hamilton cycle on 6 vertices. Notice that we can divide the vertices into 3 disjoint parts according to the loose Hamilton cycle. In other words, a loose k-uniform Hamilton cycle on n vertices can induce a partition of the cyclic ordered $[n]$ and each part contains $k-1$ vertices. There are $\frac{n!}{2n} \frac{k-1}{((k-2)!)^r}$ possible loose Hamilton cycles. The term $\frac{n!}{2n}$ comes from cyclic permutation. The term $k-1$ comes from how we divide the cyclic ordered $[n]$ into r parts. Within each hyperedge, the order is trivial, so we divide $((k-2)!)^r$. Let $X = \binom{[n]}{k}$ and $\mathcal{H}$ be the subset of 2^X, each of whose elements corresponds to a loose Hamilton cycle on n vertices. Notice that $\mathcal{H}$ is r-bounded and we will show that $\mathcal{H}$ is $\frac{C}{n^{k-1}}$-*spread* for some large C. For $\forall S \subseteq X$

with $|S| = s$, using Stirling's approximation we have

$$\begin{aligned}
\frac{|\mathcal{H} \cap \langle S \rangle|}{|\mathcal{H}|} &\leq \frac{2n((k-2)!)^r}{(k-1)n!} \frac{(n-ks+s)!}{2(n-ks+s)} \frac{(k(k-1))^s}{((k-2)!)^{r-s}} \\
&= \frac{((k-2)!)^s(k(k-1))^s}{k-1} \frac{(n-(k-1)s-1)!}{(n-1)!} \\
&\leq \frac{(k!)^s e}{k-1} \frac{\sqrt{2\pi(n-(k-1)s-1)}}{\sqrt{2\pi(n-1)}} \frac{((n-(k-1)s-1)/e)^{n-(k-1)s-1}}{((n-1)/e)^{n-1}} \\
&\leq \frac{(k!)^s e^{(k-1)s+1}}{k-1} \left(\frac{1}{n-1}\right)^{(k-1)s} \\
&\leq \left(\frac{C}{n^{k-1}}\right)^s
\end{aligned}$$

where C is large enough. Then by the theorem 4.10, we have that a random uniform $Ln \log n$-element subset of X belongs to $\langle \mathcal{H} \rangle$ with high probability. This result was first given by Dudek and Frieze [7].

Bounded degree spanning trees: Let $\mathcal{H}$ be the family of edges set of all spanning trees of the complete graph K_n with maximum degree $\Delta = O(1)$. In other words, $\mathcal{H}$ is a $(n-1)$-uniform subset of 2^X where $X = \binom{[n]}{2}$ and each element of $\mathcal{H}$ corresponds to a Δ-bounded spanning tree. To use theorem 4.10, we will show that $\mathcal{H}$ is $\frac{\Delta}{n-1}$-*spread*. If $S \subseteq X$ is not a subset of any element in $\mathcal{H}$, then $\mathcal{H} \cap \langle S \rangle = \emptyset$. Suppose that $S \subseteq X$ is a subset of some element T in $\mathcal{H}$. Let π be a random permutation of $[n]$ and

$$\pi(T) = \{\pi(u)\pi(v) : uv \in T\}.$$

It is found that $\pi(T)$ induces a Δ-bounded spanning tree isomorphic to T and $\pi(T) \in \mathcal{H}$ trivially. Since isomorphism induces a partition of $\mathcal{H}$, we have

$$\frac{|\mathcal{H} \cap \langle S \rangle|}{|\mathcal{H}|} \leq \max_T \mathbb{P}(S \subseteq \pi(T)) \leq \left(\frac{\Delta}{n-1}\right)^{|S|}$$

and the second inequality holds since for each edge in S we have at most Δ choices for one end after the other end has been determined. Then by theorem 4.10, a uniformly random $(Ln \log n)$-element subset of X belongs to $\langle \mathcal{H} \rangle$ with high probability. This result was first given by Montgomery [22].

Chapter 5

Proofs

In this chapter, we will first prove that the theorem 4.4 can be implied from the theorem 4.5. Then we use iteration to prove the theorem 4.5. In the proof, we try to find a small partial cover of $\mathcal{H}$, i.e. cover part of $\mathcal{H}$ through the uniformly chosen set W_i in the i^{th} step. The iteration will result in two cases

1. We find a cover of $\mathcal{H}$ whose expectation cost is small, or
2. there is a uniformly chosen $((Lp\log\ell)|X|)$-element subset of X belonging to $\langle\mathcal{H}\rangle$.

If $\mathcal{H}$ is not *p-small*, the case 1 can never occur. Then the proof of the theorem 4.5 is completed. We will first state how to find the cover in each step and show the cover is expected small. After that, we will show how the iteration leads to the exactly two final cases and complete the proof we want. The proofs reference [24].

5.1 Proof of reduction

Proof. (theorem 4.4 can be implied from the theorem 4.5) Let $\mathcal{H}$ be the set of minimal elements of the property $\mathcal{F}$, i.e. $\langle\mathcal{H}\rangle = \mathcal{F}$. Recall theorem 4.5, let $\ell = C\ell(\mathcal{F})$ and $m = (Lp\log\ell)|X|$, where $p > q(\mathcal{F})$ and C is large enough to make

$$\mathbb{P}(X_m \in \langle\mathcal{H}\rangle) \geq \frac{3}{4}.$$

Let $p' = 2Lp\log\ell$, we have

$$\mathbb{P}(X_{p'} \in \mathcal{F}) = \mathbb{P}(X_{p'} \in \langle\mathcal{H}\rangle) \geq \mathbb{P}(X_m \in \langle\mathcal{H}\rangle)\mathbb{P}(|X_{p'}| \geq m).$$

To show $\mathbb{P}(X_{p'} \in \mathcal{F}) \geq \frac{1}{2}$, we just need to show that $\mathbb{P}(|X_{p'}| \geq m) \geq \frac{2}{3}$. Notice that $\mathbb{E}[|X_{p'}|] = p'|X| = 2m$, we have

$$\begin{aligned}\mathbb{P}(|X_{p'}| \geq m) &\geq \mathbb{P}(||X_{p'}| - 2m)| \leq m) \\ &\geq 1 - \frac{\mathrm{Var}(|X_{p'}|)}{m^2} \\ &= 1 - \frac{2m(1-p')}{m^2}\end{aligned}$$

by Chebyshev's inequality. Consider that the subset of 2^X consisting of all single-element sets is a trivial cover of $\langle \mathcal{H} \rangle$. We have $p|X| > \frac{1}{2}$ since $\mathcal{H}$ is not *p-small*. It follows that $2m = p'|X| \geq L \log \ell$ and $\mathbb{P}(|X_{p'}| \geq m) \geq \frac{2}{3}$ when L large enough. Therefore, we have

$$\mathbb{P}(X_{p'} \in \mathcal{F}) \geq \frac{3}{4}\mathbb{P}(|X_{p'}| \geq m) \geq \frac{1}{2}$$

which implies that

$$p' = 2Lp \log \ell \geq p_c(\mathcal{F}).$$

Recall that p can be any value greater than $q(\mathcal{F})$, we have completed our proof. □

5.2 Proof of the theorem 4.5

Proof. Our proof is divided into three parts and we prove the result by **iteration**. In the first part, we will show how to construct a small partial cover of $\mathcal{H}$ in each iteration step. In the second part, we will analyze the iteration and state the final cases when the iteration ends. Finally, we show that one of the final cases occurs with high probability which is exactly the result we want to prove. Let $|X| = n$ in our proof.

Part 1. In the i^{th} step, let $w_i = L_i pn$ where L_i is a large constant and W_i be the random subset uniformly chosen from $\binom{X_{i-1}}{w_i}$ where $X_i = X_{i-1} \backslash W_i$ and $X_0 = X$. Let $\mathcal{H}_0 = \mathcal{H}$ and $\mathcal{H}_i$ is a $0.9^i\ell$-bounded subset of 2^{X_i} for $i \geq 1$.

Given $S_{i-1} \in \mathcal{H}_{i-1}$, we define the *minimum fragment* $T(S_{i-1}, W_i)$ as the set with the smallest size that can be written as the form of $S'_{i-1} \backslash W_i$ where $S'_{i-1} \subseteq S_{i-1} \cup W_i$ and $S'_{i-1} \in \mathcal{H}_{i-1}$. Let $t(S_{i-1}, W_i) = |T(S_{i-1}, W_i)|$. Let $\mathcal{G}(W_i)$ be the set of elements in $\mathcal{H}_{i-1}$ such that $T(S_{i-1}, W_i)$ is relatively large as follow,

$$\mathcal{G}(W_i) := \{S_{i-1} \in \mathcal{H}_{i-1} : t(S_{i-1}, W_i) \geq 0.9^i \ell\}.$$

Since $T(S_{i-1}, W_i) \subseteq S_{i-1}$, we have a cover $\mathcal{U}(W_i)$ of $\mathcal{G}(W_i)$ as follows,

$$\mathcal{U}(W_i) := \{T(S_{i-1}, W_i) : S_{i-1} \in \mathcal{G}(W_i)\}.$$

In terms of $\mathcal{H}_{i-1}\backslash\mathcal{G}(W_i)$, it is not always covered by $\mathcal{U}(W_i)$. Let

$$\mathcal{H}_i := \{T(S_{i-1}, W_i) : S_{i-1} \in \mathcal{H}_{i-1}\backslash\mathcal{G}(W_i)\},$$

and we have that $\mathcal{H}_i$ is a $0.9^i\ell$-bounded subset of $2^{X_{i-1}\backslash W_i} = 2^{X_i}$. Notice that $T(S_{i-1}, W_i) \subseteq S_{i-1}$, it follows that a cover of $\mathcal{H}_i$ is also a cover of $\mathcal{H}_{i-1}\backslash\mathcal{G}(W_i)$.

Part 2. Notice that in each iteration step, the size bound of elements in $\mathcal{H}_i$ strictly decreases. Thus there is a k such that $0.9 \leq 0.9^k\ell < 1$, which implies that $\mathcal{H}_k$ is either an empty set or only contains the empty set. So, our iteration can end within k steps. Now let's analyze the two different cases when the iteration ends.

1. If $\mathcal{H}_k = \emptyset$, then $\bigcup_{i=1}^k \mathcal{U}(W_i)$ is a cover of $\mathcal{H}$. We show this fact by induction. Since $\mathcal{H}_k = \emptyset$, we have

   $$\mathcal{H}_{k-1} = \mathcal{H}_k \cup \mathcal{G}(W_k) = \mathcal{G}(W_k).$$

   It follows that $\mathcal{U}(W_k)$ is a cover of $\mathcal{H}_{k-1}$.

   Suppose that $\bigcup_{i=j}^k \mathcal{U}(W_i)$ is a cover of $\mathcal{H}_{j-1}$ which implies that it is also a cover of $\mathcal{H}_{j-2}\backslash\mathcal{G}(W_{j-1})$. Since $\mathcal{U}(W_{j-1})$ is a cover of $\mathcal{G}(W_{j-1})$, we have $\bigcup_{i=j}^k \mathcal{U}(W_i) \cup \mathcal{U}(W_{j-1})$ is a cover of $\mathcal{H}_{j-2}$, that is, $\bigcup_{i=j-1}^k \mathcal{U}(W_i)$ is a cover of $\mathcal{H}_{j-2}$.

   Finally, we have $\bigcup_{i=1}^k \mathcal{U}(W_i)$ is a cover of $\mathcal{H}_0 = \mathcal{H}$.

2. If $\mathcal{H}_k = \{\emptyset\}$, then $\bigsqcup_{i=1}^k W_i \in \langle\mathcal{H}\rangle$. We also prove this fact by induction. Since $\mathcal{H}_k = \{\emptyset\}$, we have that

   $$\emptyset = T(S_{k-1}, W_k) = S'_{k-1}\backslash W_k \quad \text{for some} \quad S'_{k-1} \in \mathcal{H}_{k-1}.$$

   It follows that $S'_{k-1} \subseteq W_k$. In other words, we have $W_k \in \langle S'_{k-1}\rangle \subseteq \langle\mathcal{H}_{k-1}\rangle$.

   Suppose that $\bigsqcup_{i=j}^k W_i \in \langle\mathcal{H}_{j-1}\rangle$ which implies that there is an $S_{j-1} \in \mathcal{H}_{j-1}$ such that $S_{j-1} \subseteq \bigsqcup_{i=j}^k W_i$. Since

   $$S_{j-1} = T(S_{j-2}, W_{j-1}) = S'_{j-2}\backslash W_{j-1} \quad \text{for some} \quad S'_{j-2} \in \mathcal{H}_{j-2},$$

   we have $S'_{j-2} \subseteq S_{j-1} \cup W_{j-1}$. It follows that $S'_{j-2} \subseteq \bigsqcup_{i=j}^k W_i \sqcup W_{j-1}$. In other words, $\bigsqcup_{i=j-1}^k W_i \in \langle S'_{j-2}\rangle \subseteq \langle\mathcal{H}_{j-2}\rangle$.

   Finally, we have $\bigsqcup_{i=1}^k W_i \in \langle\mathcal{H}_0\rangle = \langle\mathcal{H}\rangle$.

Part 3. We first show that each partial cover $\mathcal{U}(W_i)$ is expected *small* enough. Let $\ell_i = 0.9^i\ell$.

Lemma 5.1. $\mathbb{E}\left[\sum_{U\in\mathcal{U}(W_i)} p^{|U|}\right] \le L_i^{-0.8\ell_{i-1}}$.

Proof. Recall the definition of $\mathcal{U}(W_i)$, we have that $\ell_i \le |U| \le \ell_{i-1}$ and $U = T(S_{i-1}, W_i)$ for some $S_{i-1} \in \mathcal{H}_{i-1}$. Expand the expectation, and we have

$$\begin{aligned}\mathbb{E}\left[\sum_{U\in\mathcal{U}(W_i)} p^{|U|}\right] &= \sum_{W_i\in\binom{X_{i-1}}{w_i}} \frac{1}{\binom{|X_{i-1}|}{w_i}} \sum_{U\in\mathcal{U}(W_i)} p^{|U|} \\ &= \frac{1}{\binom{|X_{i-1}|}{w_i}} \sum_{\ell_i\le h\le\ell_{i-1}} \left(p^h \times N_h\right)\end{aligned}$$

where $N_h = \#\{(W_i, T(S_{i-1}, W_i)) : W_i \in \binom{X_{i-1}}{w_i}, T(S_{i-1}, W_i) = h, S_{i-1} \in \mathcal{H}_{i-1}\}$. Then we estimate N_h as follows.

1. Specify $Z = W_i \sqcup T(S_{i-1}, W_i)$. The number of possible Z is at most
$$\begin{aligned}\binom{|X_{i-1}|}{w_i+h} &= \binom{|X_{i-1}|}{w_i}\frac{(|X_{i-1}|-w_i)\cdots(|X_{i-1}|-w_i-h+1)}{(w_i+1)\cdots(w_i+h)} \\ &\le \binom{|X_{i-1}|}{w_i}\left(\frac{n}{w_i}\right)^h \\ &= \binom{|X_{i-1}|}{w_i}\left(\frac{1}{L_i p}\right)^h.\end{aligned}$$

2. Specify $T(S_{i-1}, W_i)$. For any $S''_{i-1} \subseteq Z$ and $S''_{i-1} \in \mathcal{H}_{i-1}$, we claim that $T(S_{i-1}, W_i) \subseteq S''_{i-1}$. This key observation is from *minimum fragment*. Since Z is the disjoint union of W_i and $T(S_{i-1}, W_i)$, we have $S''_{i-1}\backslash W_i \subseteq T(S_{i-1}, W_i)$. Suppose that $S''_{i-1}\backslash W_i \subsetneq T(S_{i-1}, W_i)$, it is easy to check that $S''_{i-1}\backslash W_i$ with smaller size than $T(S_{i-1}, W_i)$ which is contradictory to the definition of $T(S_{i-1}, W_i)$. Thus, we can pick a $\mathcal{H}_{i-1} \ni S''_{i-1} \subseteq Z$ arbitrarily and $T(S_{i-1}, W_i)$ is always contained in it which implies that we have at most $2^{\ell_{i-1}}$ possible $T(S_{i-1}, W_i)$.

3. After specify Z and $T(S_{i-1}, W_i)$, we have $W_i = Z\backslash T(S_{i-1}, W_i)$.

Therefore, we have
$$N_h \le \binom{|X_{i-1}|}{w_i}\left(\frac{1}{L_i p}\right)^h 2^{\ell_{i-1}}$$
and
$$\begin{aligned}\mathbb{E}\left[\sum_{U\in\mathcal{U}(W_i)} p^{|U|}\right] &\le \sum_{\ell_i\le h\le\ell_{i-1}}\left(\left(\frac{1}{L_i}\right)^h 2^{\ell_{i-1}}\right) \\ &\le L_i^{-0.8\ell_{i-1}}.\end{aligned}$$

The last inequality holds since L_i is large. □

In **Part 2.**, we have shown that there are exactly two final cases which imply that

$$\mathbb{P}\left(\bigcup_{i=1}^{k}\mathcal{U}(W_i) \text{ is a cover of } \mathcal{H}\right)+\mathbb{P}\left(\bigsqcup_{i=1}^{k}W_i\in\langle\mathcal{H}\rangle\right)\geq 1$$

and equivalently

$$\mathbb{P}\left(\bigsqcup_{i=1}^{k}W_i\in\langle\mathcal{H}\rangle\right)\geq 1-\mathbb{P}\left(\bigcup_{i=1}^{k}\mathcal{U}(W_i) \text{ is a cover of } \mathcal{H}\right). \tag{5.1}$$

By the assumption that '$\mathcal{H}$ is **not** *p-small*', we have

$$\mathbb{P}\left(\bigcup_{i=1}^{k}\mathcal{U}(W_i) \text{ is a cover of } \mathcal{H}\right)\leq\mathbb{P}\left(\sum_{U\in\bigcup_{i=1}^{k}\mathcal{U}(W_i)}p^{|U|}>\frac{1}{2}\right). \tag{5.2}$$

Now we estimate the expectation of $\sum_{U\in\bigcup_{i=1}^{k}\mathcal{U}(W_i)}p^{|U|}$ by lemma 5.1. Let

$$L_i=\begin{cases}C & \text{if } i<k-\sqrt{\log_{0.9}(0.9/\ell)}\\ C\sqrt{\log\ell} & \text{if } k-\sqrt{\log_{0.9}(0.9/\ell)}\leq i\leq k\end{cases}$$

where C is a large constant and let $\log_{0.9}(0.9/\ell)=c\log\ell$ where $c>0$. We have

$$\begin{aligned}\mathbb{E}\left[\sum_{U\in\bigcup_{i=1}^{k}\mathcal{U}(W_i)}p^{|U|}\right]&\leq\sum_{i=1}^{k}\mathbb{E}\left[\sum_{U\in\mathcal{U}(W_i)}p^{|U|}\right]\\&\leq\sum_{i=1}^{k}L_i^{-0.8\ell_{i-1}}\\&=\sum_{i<k-\sqrt{c\log\ell}}C^{-0.8\cdot0.9^{i-1}\ell}+\sum_{i\geq k-\sqrt{c\log\ell}}^{k}\left(C\sqrt{\log\ell}\right)^{-0.8\cdot0.9^{i-1}\ell}\end{aligned}$$

Recall that $0.9\leq0.9^k\ell<1$, we have

$$\begin{aligned}&c\log\ell-1=\log_{0.9}\frac{1}{\ell}<k\leq\log_{0.9}\frac{0.9}{\ell}=c\log\ell\\&\exp\left(c'\sqrt{\log\ell}\right)\leq0.9^{i-1}\ell\quad\text{ if } i<k-\sqrt{c\log\ell}\\&1\leq0.9^{k-1}\ell\quad\text{ and }\quad 1<0.8\cdot0.9^{k-4}\ell\end{aligned}$$

where $c'>0$ is a constant. It follows that

$$\begin{aligned}\mathbb{E}\left[\sum_{U\in\bigcup_{i=1}^{k}\mathcal{U}(W_i)}p^{|U|}\right]&\leq(c\log\ell)\,C^{-0.8\cdot\exp\left(c'\sqrt{\log\ell}\right)}+\sqrt{c\log\ell}\left(C\sqrt{\log\ell}\right)^{-0.8\cdot0.9^{k-4}\ell}+3\left(C\sqrt{\log\ell}\right)^{-0.8}\\&=O((\log\ell)^{-\hat{c}})\\&=o_{\ell\to\infty}(1)\end{aligned}$$

where $\hat{c} > 0$ is a constant and $\sqrt{c \log \ell} \geq 3$ when ℓ is large. Use Markov's inequality in (5.2), we have

$$\begin{aligned}
\mathbb{P}\left(\bigcup_{i=1}^{k} \mathcal{U}(W_i) \text{ is a cover of } \mathcal{H}\right) &\leq \mathbb{P}\left(\sum_{U \in \bigcup_{i=1}^{k} \mathcal{U}(W_i)} p^{|U|} > \frac{1}{2}\right) \\
&\leq 2\mathbb{E}\left[\sum_{U \in \bigcup_{i=1}^{k} \mathcal{U}(W_i)} p^{|U|}\right] \\
&= o_{\ell \to \infty}(1)
\end{aligned}$$

which implies that

$$\mathbb{P}\left(\bigsqcup_{i=1}^{k} W_i \in \langle \mathcal{H} \rangle\right) \geq 1 - \mathbb{P}\left(\bigcup_{i=1}^{k} \mathcal{U}(W_i) \text{ is a cover of } \mathcal{H}\right) \geq 1 - o_{\ell \to \infty}(1).$$

by (5.1). Notice that $\bigsqcup_{i=1}^{k} W_i$ is a $\left(\sum_{i=1}^{k} L_i pn\right)$-element uniform subset of X where $\sum_{i=1}^{k} L_i pn = (Lp \log \ell) |X|$ for some L, we complete our proof. $\square$

Chapter 6

Discussion

Park and Pham's proof of the Kahn-Kalai conjecture is a remarkable breakthrough in random combinatorics, attracting widespread attention. Recall that we choose the $p(\frac{1}{2})$ to be the threshold and $p(\epsilon)$ for $\epsilon \in (0,1)$ is defined as follows

$$\mathbb{P}(G(n, p(\epsilon)) \text{ has } \mathcal{F}) = \epsilon.$$

For fixed ϵ, Park and Pham's work also can provide an upper bound. But their bound dependent on ϵ is not good enough. Bell [2] proposed this problem and showed a better ϵ-dependent upper bound.

Theorem 6.1. *Let $\mathcal{H}$ be an ℓ-bounded subset of 2^X that is **not** q-small and let $\epsilon \in (0,1)$. Let $p = 48q \log \frac{\ell}{\epsilon}$. Then $\mathbb{P}(X_p \in \langle \mathcal{H} \rangle) > 1 - \epsilon$.*

The arguments in [2] also give a slightly better bound for the threshold as follows

$$p_c(\mathcal{F}) \leq 1 - (1 - Kq(\mathcal{F}))^{\log_2 \ell(\mathcal{F})}.$$

Przybyłowski and Riordan [26] slightly improved this bound by a technique *clone*

$$p_c(\mathcal{F}) \leq 1 - e^{-Kq(\mathcal{F}) \log_2 \ell(\mathcal{F})}.$$

Other endeavours tend to give bounds on the multiplicative constant K. The best known work due to Vu and Tran [30] is $K \approx 3.998$. They proposed a simplification of Park and Pham's proof by technically strengthening the theorem 4.5 and applying induction. This reduces the proof to only one page and also applies to ϵ-dependent problem.

Kahn-Kalai conjecture (now Park-Pham theorem) is extremely powerful, implying very difficult results historically such as Shamir's problem and bounded degree spanning tree as we mentioned before. We end this dissertation with two open conjectures as Park mentioned in [23].

The first conjecture is the graph version of the Kahn-Kalai conjecture. Recall the definition of abstract expectation threshold and expectation threshold in graphs (definition 4.3 and definition 4.2). We claim that

$$p_E(H) \leq q(\mathcal{F}_H)$$

where $\mathcal{F}_H$ is the property that the random graph $G(n,p)$ contains H as a subgraph. This claim is easy to check since $\langle H' \rangle$ is a cover of $\mathcal{F}_H$ for $\forall H' \subseteq H$. Though the abstract Kahn-Kalai conjecture has been proved, the graph version is open. This conjecture was also proposed by Kahn and Kalai in [19].

Conjecture 6.2. *Let $\mathcal{F}_H$ be the property that the random graph $G(n,p)$ contains H as a subgraph. There is a universal constant K such that*

$$p_c(\mathcal{F}_H) \leq K p_E(H) \log n.$$

The second conjecture tries to solve the problem that it is not easy to compute $q(\mathcal{F})$. Recall that the fractional expectation threshold $q_f(\mathcal{F})$ has a close relationship with *spread* (refer to proposition 4.9 and [29]). Through finding an α-spread a bound probability measure on $\mathcal{F}$, we can get an upper bound of $q_f(\mathcal{F})$. Many applications use this idea to apply the fractional Kahn-Kalai conjecture. Talagrand proposes the following conjecture which implies the equivalence between the fractional Kahn-Kalai conjecture and the Kahn-Kalai conjecture.

Conjecture 6.3. *Let $\mathcal{F}$ be an increasing property in finite set X. There exists a constant K such that*

$$q_f(\mathcal{F}) \leq K q(\mathcal{F}).$$

References

[1] Ryan Alweiss, Shachar Lovett, Kewen Wu, and Jiapeng Zhang. Improved bounds for the sunflower lemma. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, pages 624–630, 2020.

[2] Tolson Bell. The park-pham theorem with optimal convergence rate. *arXiv preprint arXiv:2210.03691*, 2022.

[3] Patrick Billingsley. *Probability and measure*. John Wiley & Sons, 2017.

[4] Béla Bollobás. The evolution of sparse graphs. *Graph theory and combinatorics (Cambridge, 1983)*, pages 35–57, 1984.

[5] Béla Bollobás and Alan M Frieze. *On matchings and Hamiltonian cycles in random graphs*. Management Sciences Research Group, Graduate School of Industrial ..., 1983.

[6] Béla Bollobás and Arthur G Thomason. Threshold functions. *Combinatorica*, 7(1):35–38, 1987.

[7] Andrzej Dudek and Alan Frieze. Loose hamilton cycles in random uniform hypergraphs. *arXiv preprint arXiv:1006.1909*, 2010.

[8] Pál Erdős and Alfréd Rényi. On random matrices. *Magyar Tud. Akad. Mat. Kutató Int. Közl*, 8:455–461, 1964.

[9] Pál Erdős and Alfréd Rényi. On the existence of a factor of degree one of a connected random graph. *Acta Math. Acad. Sci. Hungar*, 17:359–368, 1966.

[10] Paul Erdős and Alfréd Rényi. On random graphs i. *Publ. Math. Debrecen*, 6:290–297, 1959.

[11] Paul Erdős and Alfréd Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5(1):17–60, 1960.

[12] Keith Frankston, Jeff Kahn, Bhargav Narayanan, and Jinyoung Park. Thresholds versus fractional expectation-thresholds. *Annals of Mathematics*, 194(2):475–495, 2021.

[13] Ehud Friedgut. Hunting for sharp thresholds. *Random Structures & Algorithms*, 26(1-2):37–51, 2005.

[14] Ehud Friedgut, Jean Bourgain, et al. Sharp thresholds of graph properties, and the k-sat problem. *Journal of the American mathematical Society*, 12(4):1017–1054, 1999.

[15] Ehud Friedgut and Gil Kalai. Every monotone graph property has a sharp threshold. *Proceedings of the American mathematical Society*, 124(10):2993–3002, 1996.

[16] Alan Frieze and Michał Karoński. *Introduction to random graphs*. 2023. https://www.math.cmu.edu/~af1p/BOOK.pdf.

[17] Anders Johansson, Jeff Kahn, and Van Vu. Factors in random graphs. *Random Structures & Algorithms*, 33(1):1–28, 2008.

[18] Rob Kaas and Jan M Buhrman. Mean, median and mode in binomial distributions. *Statistica Neerlandica*, 34(1):13–18, 1980.

[19] Jeff Kahn and Gil Kalai. Thresholds and expectation thresholds. *Combinatorics, Probability and Computing*, 16(3):495–502, 2007.

[20] Jeff Kahn, Bhargav Narayanan, and Jinyoung Park. The threshold for the square of a hamilton cycle. *Proceedings of the American Mathematical Society*, 149(8):3201–3208, 2021.

[21] János Komlós and Endre Szemerédi. Limit distribution for the existence of hamiltonian cycles in a random graph. *Discrete mathematics*, 43(1):55–63, 1983.

[22] Richard Montgomery. Spanning trees in random graphs. *Advances in Mathematics*, 356:106793, 2019.

[23] Jinyoung Park. Threshold phenomena for random discrete structures. *arXiv preprint arXiv:2306.13823*, 2023.

[24] Jinyoung Park and Huy Tuan Pham. A proof of the kahn-kalai conjecture. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 636–639. IEEE, 2022.

[25] Lajos Pósa. Hamiltonian circuits in random graphs. *Discrete Mathematics*, 14(4):359–364, 1976.

[26] Tomasz Przybyłowski and Oliver Riordan. Thresholds and expectation thresholds for larger p. *arXiv preprint arXiv:2302.03327*, 2023.

[27] A Rényi. Quelques remarques ser les probabilites des evenements dependants. *Journal de Mathematique*, 37(393-398):216, 1958.

[28] Oliver Riordan. *Probabilistic Combinatorics*. 2019. https://courses.maths.ox.ac.uk/pluginfile.php/27059/mod_resource/content/1/LectureNotes.pdf.

[29] Michel Talagrand. Are many small sets explicitly small? In *Proceedings of the forty-second ACM symposium on Theory of computing*, pages 13–36, 2010.

[30] Van Vu and Phuc Tran. A short proof of kahn-kalai conjecture. *arXiv preprint arXiv:2303.02144*, 2023.