PARK AND PHAM'S PROOF OF THE KAHN-KALAI CONJECTURE

HAFSAH AAMER

MATH 490, FALL 2022

ABSTRACT. This paper aims to expound the argument used in Pham and Park's proof of the Kahn-Kalai's "expectation-threshold" conjecture. Before doing so, we provide necessary background and results needed to understand the statement of the theorem, as well as the proof thereof.

1. PRELIMINARIES

We have used [5], [6], [4], and [3] to describe preliminary definitions and results.

We start with a brief introduction to random graphs. The foundations of the theory of random graphs were laid down by Erdős and Rényi in their seminal paper [1], thus initiating the study of "typical" properties of graphs equipped with a certain probability distribution. We describe the the most encountered probability spaces (models) of random graphs. First, we give the following preliminary definitions:

Definition 1.1. A *probability space* is a triple $(\Omega, \Sigma, \mathbb{P})$, where Ω is a set, Σ is a σ-algebra of subsets of Ω, and $\mathbb{P}$ is a non-negative measure on Σ with $\mathbb{P}(\Omega) = 1$. In other words, Σ is the set of all measurable sets of Ω under the probablity measure $\mathbb{P}$. (A probability measure $\mathbb{P}$ is a measure such that $\mathbb{P}(\Omega) = 1$.)

A simple case of a probability space is where Ω is a set, Σ is 2^Ω - the power set of Ω. The probability measure $\mathbb{P}$ is the function $\mathbb{P} : \Sigma \to [0, 1]$ that assigns to each event in Σ a probability, which is a number between 0 and 1, namely

$$\mathbb{P}(A) = \sum_{a \in A} P(\{a\}), A \subset \Omega.$$

Definition 1.2. A *random variable* X on $(\Omega, \Sigma, \mathbb{P})$ is a measurable function $X : \Omega \longrightarrow E$ from sample space Ω to a measurable space E. The probability that X takes on some value $S \in E$ is given as:

$$\mathbb{P}(X \in S) = \mathbb{P}(\{\omega \in \Omega \mid X(\omega) \in S\})$$

Definition 1.3. Given a random variable X with countably many possible outcomes x_i. Then the *expectation of X* is given as

$$\mathbb{E}(X) = \sum_{i=1}^{\infty} x_i \mathbb{P}(X = x_i)$$

Theorem 1.4 (Markov's Inequality)**.** *Let $X \geq 0$ be a random variable, and $a > 0$, then*

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

1.1. **Random Graphs.** Consider the vertex set $V = [n] = \{1, 2, \ldots, n\}$. Then the maximum number of edges on $[n]$ is $\binom{n}{2}$, i.e. the complete graph K_n on $[n]$ has $\binom{n}{2}$ edges. Let $\mathscr{G}_{n,m}$ be the family of all labelled graphs on $[n]$ with exactly m edges, $0 \leq m \leq \binom{n}{2}$, so that $|\mathscr{G}_{n,m}| = \binom{\binom{n}{2}}{m}$. We assign to each graph G in $\mathscr{G}_{n,m}$ the probability

$$\mathbb{P}(G) = \mathbb{P}(G \in \mathscr{G}_{n,m}) = \left(\binom{\binom{n}{2}}{m} \right)^{-1}$$

We call this graph G the *uniform random graph* - since the possibility of picking each graph in $\mathscr{G}_{n,m}$ is equally likely - and denote it by $G_{n,m}$.

The second graph model $\mathscr{G}_{n,p}$ is the family of all graphs G on $[n]$ where each of the possible $\binom{n}{2}$ edges is included independently with the fixed probability p, $0 \leq p \leq 1$. Hence, if G is a graph in $\mathscr{G}_{n,p}$ with $[n]$ vertices and it has m edges, $0 \leq m \leq \binom{n}{2}$, we assign to G the probability

$$\mathbb{P}(G) = \mathbb{P}(G \in \mathscr{G}_{n,p}) = p^m (1-p)^{\binom{n}{2}-m}$$

We call this graph the *binomial random graph* and denote it by $G_{n,p}$.

With the definitions above, the random graph models $\mathscr{G}_{n,m}$ and $\mathscr{G}_{n,p}$ form a probability space under the assosciated probability functions $\mathbb{P}$. Upon observing the descriptions of these random graph models, there seems to be a close relationship between the two models. Indeed we have:

Lemma 1.5. *Given that a random graph $G_{n,p}$ has m edges, then it is equally likely to be one of the $\binom{\binom{n}{2}}{m}$ choices of graphs that have m edges.*

Proof. Let H be a labelled graph with m edges. We will denote the size of the edge set of H by $|H| = m$. Now, since

$$\{G_{n,p} = H\} \subseteq \{|G_{n,p}| = m\}$$

we have

$$\mathbb{P}\left(G_{n,p} = H \,||\, G_{n,p}| = m\right) = \frac{\mathbb{P}\left(G_{n,p} = H, |G_{n,p}| = m\right)}{\mathbb{P}\left(|E_{n,p}| = m\right)}$$

$$= \frac{\mathbb{P}\left(G_{n,p} = H\right)}{\mathbb{P}\left(|G_{n,p}| = m\right)}$$

$$= \frac{p^m (1-p)^{\binom{n}{2}-m}}{\binom{\binom{n}{2}}{m} p^m (1-p)^{\binom{n}{2}-m}}$$

$$= \binom{\binom{n}{2}}{m}^{-1}$$

$\square$

This just says that $G_{n,p}$ shares the same probability distribution as $G_{n,m}$ when conditioned on the event that $G_{n,p}$ has m edges. Thus, for large n, we can expect $G_{n,p}$ and $G_{n,m}$ to behave similarly, given that the fixed m in $G_{n,m}$ is equal to or close to the expected number of edges of $G_{n,p}$, i.e. $m = \binom{n}{2}p \approx \frac{n^2 p}{2}$.

We imagine our random graphs evoluting on the initially isolated vertex set $V = [n]$. As we increase the values m and p, respectively, our graphs successively amass more edges. A very natural question and a central point of interest in the theory of random graphs is, at what stage in this evolution a particular interesting property emerges in our random graph?

Write $2^{\binom{n}{2}}$ for the set of all labeled graphs on the vertex set $[n]$. A *graph property* $\mathcal{F}$ is defined to be a subset of $2^{\binom{n}{2}}$ having the property $\mathcal{F}$. The statement that 'the graph G has a property $\mathcal{F}$' is then equivalent to '$G \in \mathcal{F}$'. For example, all planar graphs, all connected graphs, graphs containing a triangle, graphs containing a Hamiltonian cycle are specific 'graph properties' (on n vertices) and for instance, the graph property of being connected is the set $\left\{ G \in 2^{\binom{n}{2}} : G \text{ is connected} \right\}$.

We call $\mathcal{F}$ a *monotone increasing property* if whenever $G \in \mathcal{F}$ and $G \subseteq H$ then $G \in \mathcal{F}$ as well. This means adding an edge e to a graph G does not destroy the property, since $G \subseteq G + e \in \mathcal{F}$. Thus the property of containment of a specific subgraph, such as a triangle, or a Hamiltonian cycle, is monotone increasing. We say a monotone increasing property $\mathcal{F}$ is *non-trivial* if the empty graph $\emptyset \notin \mathcal{F}$ and the complete graph $K_n \in \mathcal{F}$. Hereon, we will consider only non-trivial properties.

To see the behavior of a 'typical' element acquiring a graph property, we will work in a sequence of probability spaces (graph models). Let $\mathscr{G}_n$ be a random graph model of order n, (could be $\mathscr{G}_n = \mathscr{G}_{n,m}$ or $\mathscr{G}_n = \mathscr{G}_{n,p}$). We say that a *typical element in $\mathscr{G}_n$ has a property $\mathcal{F}$* when for any random graph $G_n \in \mathscr{G}_n$, $\mathbb{P}(G_n \in \mathcal{F}) \to 1$ as $n \to \infty$. When this happens, we also say that *almost every* (a.e) graph in $\mathscr{G}_n$ has property $\mathcal{F}$.

We have the following fact about the probabilities of containment of a property $\mathcal{F}$:

Theorem 1.6. *Let $\mathcal{F}$ be a monotone increasing property, and $0 \leq m_1 < m_2 \leq \binom{n}{2}$ and $0 \leq p_1 < p_2 \leq 1$. Then*

$$\mathbb{P}\left(G_{n,p_1} \in \mathcal{F}\right) \leq \mathbb{P}\left(G_{n,p_2} \in \mathcal{F}\right) \ \ and \ \mathbb{P}\left(G_{n,m_1} \in \mathcal{F}\right) \leq \mathbb{P}\left(G_{n,m_2} \in \mathcal{F}\right)$$

Proof.   (i) We select the m_1 edges successively. If the random graph obtained from selecting only m_1 edges has the property $\mathcal{F}$, then the random graph obtained from selecting m_2 will also have the property $\mathcal{F}$.

   (ii) Let p be such that $1 - p_2 = (1 - p_1)(1 - p)$, and pick independently $G_1 \in \mathscr{G}_{n,p_1}$ and $G \in \mathscr{G}_{n,p}$, and put $G_2 = G_1 \cup G$. Then an edge is not included in the graph G_2 if it is not included in both G_1 and G. This means the *failure* probability of an edge in G_2 is exactly $1 - p_2 = (1 - p_1)(1 - p)$. This implies exactly that $G_2 \in \mathscr{G}_{n,p_2}$. Now $\mathcal{F}$ monotone increasing implies that if G_1 has the property $\mathcal{F}$, then G_2 also has it and so we are done.

$\square$

1.2. Thresholds. Erdős and Rényi observed a very striking result that the appearance and disappearance of most monotone graph properties are rather "abrupt": for a certain m (or p) value, almost no $G_{n,m}$ (or $G_{n,p}$) has the property $\mathcal{F}$, but being slightly above m (or p) renders graphs $G_{n,m}$ (or $G_{n,p}$) that almost always have $\mathcal{F}$. We formalize this by defining *threshold functions*, or simply *thresholds*.

Definition 1.7. Given a monotone increasing property $\mathcal{F}$ in $G_{n,m}$ a function $m^* = m^*(n)$ is called a *threshold function* for $\mathcal{F}$ if

$$\lim_{n\to\infty} \mathbb{P}\left(G_{n,m} \in \mathcal{F}\right) = \begin{cases} 0 & \text{if } m/m^* \to 0, \\ 1 & \text{if } m/m^* \to \infty, \end{cases}$$

as $n \to \infty$.

We have a similar definition for the other model.

Definition 1.8. Given a monotone increasing property $\mathcal{F}$ in $G_{n,p}$ a function $p^* = p^*(n)$ is called a *threshold function* for $\mathcal{F}$ if

$$\lim_{n \to \infty} \mathbb{P}\left(G_{n,p} \in \mathcal{F}\right) = \begin{cases} 0 & \text{if } p/p^* \to 0, \\ 1 & \text{if } p/p^* \to \infty, \end{cases}$$

as $n \to \infty$.

A huge amount of research is centered around finding these thresholds for various graph properties. An important result in the theory, shown by Bollabás and Thomason [4], says that every non-trivial monotone graph property has a threshold. Their results also show that $m^*(n) = \max\left\{l : \mathbb{P}(G_{n,l} \in \mathcal{F}) \leq 1/2\right\}$ and $p^*(n) = \max\left\{q : \mathbb{P}(G_{n,q} \in \mathcal{F}) \leq 1/2\right\}$ are respective thresholds for a property $\mathcal{F}$ on $G_{n,m}$ and $G_{n,p}$. Clearly, threshold functions are not unique, but they are unique up to factors: if m_1^* is a threshold of $\mathcal{F}$, then m_2^* is another threshold of $\mathcal{F}$ iff $m_2^* = O\left(m_1^*\right)$ and $m_1^* = O\left(m_2^*\right)$. Following this, we shall talk about *the* threshold of a property.

The *Kahn-Kalai Conjecture* in substance provides bounds on thresholds of increasing graph properties. Although motivation of the subject arises from random graphs and their evolution, the setting of the conjecture is more general. Random graphs help to serve a more concrete picture in mind, but the definitions can be easily carried to a more general finite set X.

2. The Kahn-Kalai Conjecture

We have used [2], [3], and [7] to build up to the statement of the main result of the paper. Most of the notation used closely follows that of the aforementioned papers.

In the setting of this result, it is irrelevant that we have graph properties and graph thresholds. We give analogous definitions to $G_{n,m}$ and $G_{n,p}$.

Let X be a finite set and 2^X the power set of X. Let $|X| = n$ where n is large. For any integer m, we call any subset of X of size m an *m-subset* of X, and we denote by X_m a *uniformly random m*-subset of X.

For $p \in [0, 1]$, let μ_p be the product measure on 2^X given by $\mu_p(A) = p^{|A|}(1-p)^{n-|A|}$ for $A \subseteq X$. We denote by X_p the *random variable* whose distribution is μ_p. In other words, X_p is the random set obtained from X when we include each element of X independently with the probability p.

As before $\mathcal{F} \in 2^X$ is an increasing property of subsets of X, meaning that if $B \supseteq A \in \mathcal{F} \Rightarrow B \in \mathcal{F}$. Then we have

$$\mu_p(\mathcal{F}) := \sum_{A \in \mathcal{F}} \mu_p(A) = \sum_{A \in \mathcal{F}} p^{|A|}(1-p)^{n-|A|}$$
$$= \sum_{A \in \mathcal{F}} \mathbb{P}\left(X_p = A\right)$$
$$= \mathbb{P}\left(X_p \in \mathcal{F}\right)$$

From Theorem 1.6 it is easy to see that that $\mu_p(\mathcal{F}) = \mathbb{P}\left(X_p \in \mathcal{F}\right)$ is strictly increasing in p, and since it is also a polynomial in p, we have that there is unique p for which $\mu_p(\mathcal{F}) = 1/2$. Call this p the *threshold* $p_c(\mathcal{F})$ whose existence is ensured by the results in [4].

Recall since we are working in sequences of probability spaces, $p_c = p_c(n)$ is a threshold function means that if $\mathcal{F}_n$ are a sequence of monotone properties of n-sized sets X, $(n = 1, 2, \ldots)$, then

$$\mu_{p(n)}\left(\mathcal{F}_n\right) \to \begin{cases} 0 & \text{if } p(n)/p_c(n) \to 0 \\ 1 & \text{if } p(n)/p_c(n) \to \infty \end{cases}$$

We note that equivalent definitions can be easily given on m in the X_m model, but we avoid stating them for now.

Let $\langle A \rangle := \{B : B \supseteq A\}$. We say $\mathcal{G} \subseteq 2^X$ is a *cover* of $\mathcal{F}$ (or *covers* $\mathcal{F}$) if

$$\mathcal{F} \subseteq \langle \mathcal{G} \rangle := \bigcup_{S \in \mathcal{G}} \langle S \rangle \tag{2.0.1}$$

We have

$$\mu_p(\mathcal{F}) \leq \mu_p(\langle \mathcal{G} \rangle) = \sum_{S \in \langle \mathcal{G} \rangle} \mathbb{P}(S = X_p) \leq \sum_{S \in \mathcal{G}} \mathbb{P}(S \subseteq X_p) = \sum_{S \in \mathcal{G}} p^{|S|}. \tag{2.0.2}$$

where $\sum_{S \in \mathcal{G}} p^{|S|} = \mathbb{E}[|\{S \in \mathcal{G} : S \subseteq X_p\}|]$.

For any cover $\mathcal{G}$, we define the *p-cost* of $\mathcal{G}$ as $\text{cost}_p(\mathcal{G}) := \sum_{S \in \mathcal{G}} p^{|S|}$. We say the cover $\mathcal{G}$ is *p-cheap* (or just *cheap*) if $\text{cost}_p(\mathcal{G}) \leq 1/2$. When a property $\mathcal{F}$ has a *p*-cheap cover, we say $\mathcal{F}$ is *p-small*.

Define the *expectation-threshold* of a property $\mathcal{F}$, $q(\mathcal{F})$, to be the maximum p such that $\mathcal{F}$ has a p-cheap cover. We have that $q = q(\mathcal{F})$ provides a lower bound, since

(2.0.2) implies $\mu_q(\mathcal{F}) \leq 1/2$.

The question we are dealing with in our main result is about locating the threshold for our random variable X_p to contain an increasing property $\mathcal{F}$. Take, for example, the case where the increasing property $\mathcal{F}$ is containment of a fixed subgraph H in the usual random graph $G_{n,p}$. Earlier basic results [3] show that the threshold for such an $\mathcal{F}$ is obtained by observing the threshold for *expectation of the number of copies of densest part of the subgraph H*. Roughly speaking, the definition of expectation-threshold that we provided above makes this notion of containing copies of densest parts of subgraphs a little better and encompasses more general cases.

Given an increasing property $\mathcal{F}$, consider the set of minimal elements of $\mathcal{F}$ such that if A and B are minimal elements, then $A \not\subset B$ and $B \not\subset A$. Write $\ell(\mathcal{F})$ for the size of the largest minimal element of $\mathcal{F}$. Thus, we have acquired all the technology to state the main theorem of this paper.

Theorem 2.1 (The Kahn-Kalai Conjecture). *There is a universal constant $K > 0$ such that for every finite X and increasing property $\mathcal{F} \subseteq 2^X$,*

$$p_c(\mathcal{F}) \leq Kq(\mathcal{F}) \log \ell(\mathcal{F})$$

Notation. In the following text all logarithms are base 2, unless specified otherwise.

To attack this problem better, Park and Pham [2] reformulate this problem into a one which consists the X_m model instead of the X_p model, enabling the brilliant use of covers that we have defined above.

Denote by $\mathcal{H}$ a *hypergraph* on X which is a collection of subsets of X. $\mathcal{H}$ is called ℓ-bounded if each of its elements, or *edges*, have size at most ℓ. We can apply our earlier definitions to $\mathcal{H}$, so that we have $\langle \mathcal{H} \rangle = \bigcup_{S \in \mathcal{H}} \langle S \rangle$, and we allow $\mathcal{H}$ to be p-cheap in the same sense.

Theorem 2.2. *There is a universal constant $L > 0$ such that if $\mathcal{H}$ is any ℓ-bounded hypergraph on X and $p > q(\mathcal{F})$, then with $m = (Lp \log \ell)|X|$*

$$\mathbb{P}\left(X_m \in \langle \mathcal{H} \rangle \right) = 1 - o_{\ell \to \infty}(1) \tag{2.2.1}$$

In what follows, with high probability (w.h.p.) refers to the probability tending to 1 as $\ell \to \infty$.

Theorem 2.2 is a stronger result in that if we adjust the ℓ in (2.2.1) we get Theorem 2.1 as an implication. In the original statement (Theorem 2.1) it is enough to show

that the containment probaility is at least $1/2$, but in the reformulation (Theorem 2.2) the containment probability tends to 1 as ℓ tends to ∞. We will now show this implication.

Theorem 2.2 $\implies$ Theorem 2.1. Let $\mathcal{F}$ be as in Theorem 2.1. We assume Theorem 2.2, and let $q > q(\mathcal{F})$. We want to show that with $p = (Kq\log\ell)|X|$, the probability of X_p having property $\mathcal{F}$ is at least $1/2$, i.e. $\mathbb{P}(X_p \in \mathcal{F}) > 1/2$. Note that if we take $\mathcal{H}$ in Theorem 2.2 to be the set of minimal elements of $\mathcal{F}$, then $\langle\mathcal{H}\rangle = \mathcal{F}$. Since $\ell(\mathcal{F})$ is the size of the largest element of $\mathcal{F}$, we have that $\mathcal{H}$ is $\ell(\mathcal{F})$-bounded. Being above the expectation threshold i.e. $q > q(\mathcal{F})$, $\mathcal{F}$ is not p-small and so $\mathcal{H}$ is not p-small (see argument below (3.0.1)), upon which the implication in Theorem 2.2 tells us that $\mathbb{P}(X_m \in \langle\mathcal{H}\rangle) = 1 - o_{\ell\to\infty}(1)$. Let $\ell = C\ell(\mathcal{F})$ where C is a universal constant such that ℓ adjusts the containment probability to be at least $3/4$, i.e. $\mathbb{P}(X_m \in \langle\mathcal{H}\rangle) \geq 3/4$.

Let $m = (Lq\log\ell)|X|$ and $p' = 2m/|X|$. Here we make the distinction that $X_{p'} \sim \mu_{p'}$ and X_m is the uniform random m-subset of X.

$$\mathbb{P}(X_{p'} \in \langle\mathcal{H}\rangle) = \sum_{k=0}^{n} \mathbb{P}(X_{p'} \in \langle\mathcal{H}\rangle \mid |X_{p'}| = k)\,\mathbb{P}(|X_{p'}| = k)$$
$$= \sum_{k=0}^{n} \mathbb{P}(X_k \in \langle\mathcal{H}\rangle)\,\mathbb{P}(|X_{p'}| = k)$$
$$\geq \mathbb{P}(X_m \in \langle\mathcal{H}\rangle)\,\mathbb{P}(|X_{p'}| \geq m)$$

where the second equality follows from Lemma (1.5), and X_k is the uniform random k-subset of X. We have

$$\mathbb{P}(X_{p'} \in \langle\mathcal{H}\rangle) \geq \mathbb{P}(|X_{p'}| \geq m)\,\mathbb{P}(X_m \in \langle\mathcal{H}\rangle) \geq (3/4)\mathbb{P}(|X_{p'}| \geq m) > 1/2$$

Note that $\mathcal{H}$ is not q-cheap means $\mathrm{cost}_q(\mathcal{H}) > 1/2$ which implies $|X|q > 1/2$ and hence $m > (L\log\ell)/2$, i.e. m is somewhat large, and so $\mathbb{P}(|X_{p'}| \geq m) \approx 1$. This gives us the last inequality. Now taking K such that $p \geq p'$ ensures the containment probability of at least $1/2$, so we are done. $\square$

We prove the reformulated statement of the Kahn-Kalai Conjecture, that is Theorem 2.2.

3. Proof of Theorem 2.2

Our main aim is to present the proof given by Pham and Park [2] in an expository manner. Hence, we follow their argument. Firstly, we provide an informal overview of the strategy employed.

Observe that if we take $\mathcal{H}$ to be the set of minimal elements of $\mathcal{F}$, then any cover of $\mathcal{H}$ will also cover $\mathcal{F}$. This follows, because given a cover $\mathcal{G}$ of $\mathcal{H}$, then we have $\forall S \in \mathcal{H}, \ \exists G \in \mathcal{G}$ such that $G \subseteq S$, and

$$\mathcal{F} \subseteq \langle \mathcal{H} \rangle = \bigcup_{S \in \mathcal{H}} \langle S \rangle \subseteq \bigcup_{G \in \mathcal{G}} \langle G \rangle = \langle \mathcal{G} \rangle \tag{3.0.1}$$

Thus, if $\mathcal{F}$ does not admit a cheap cover, then $\mathcal{H}$ does not admit one either, where cheap (or p-cheap) means $\mathrm{cost}_p(\mathrm{cover}) < 1/2$.

Note that the implication of Theorem 2.2 says that $\mathbb{P}\left(X_m \in \langle \mathcal{H} \rangle\right) = 1 - o_{\ell \to \infty}(1)$, where $\mathbb{P}\left(X_m \in \langle \mathcal{H} \rangle\right) = \mathbb{P}\left(X_m \supseteq S \in \langle \mathcal{H} \rangle\right)$. Therefore, we would be successful in proving this theorem, if we can show X_m captures some $S \in \mathcal{H}$ w.h.p. Even if this is our main goal, we cannot hope to sprinkle a random set X_m and hope to capture in it some S in $\mathcal{H}$, since we have very minimal assumptions on $\mathcal{H}$. So we embark on an alternative route where we try to cover $\mathcal{H}$ and see if that could somehow lead us to capturing some element $S \in \mathcal{H}$.

To begin with, we run an randomized algorithm where we iteratively try to form a cover of the hypergraph $\mathcal{H} = \mathcal{H}_0$. At each step in the iteration, we renew the hypergraph from the previous one and construct a cover of the new one, so we have a sequence of hypergraphs:

$$\mathcal{H}_0 \longrightarrow \mathcal{H}_1 \longrightarrow \mathcal{H}_2 \longrightarrow \cdots$$

At i-th step, we are working on $\mathcal{H}_{i-1}$ produced at $(i-1)$-th step. Here we select a sub-graph $\mathcal{G}_i$ (*good sets*) of $\mathcal{H}_{i-1}$ which admit a cheap cover $\mathcal{U}_i$ and let $\mathcal{H}_i \subseteq \mathcal{H}_{i-1} \setminus \mathcal{G}_i$ be the next updated hypergraph. So, in the $(i+1)$-th step, we now want to find the *good sets* of $\mathcal{H}_i$ and so on. A crucial point here is that since $\mathcal{H}_i \subseteq \mathcal{H}_{i-1} \setminus \mathcal{G}_i$, the cover we wish to construct of $\mathcal{H}_i$ should also cover the "leftovers" $\mathcal{H}_{i-1} \setminus \mathcal{G}_i$ from the previous hypergraph,

$$\mathcal{H}_{i-1} \setminus \mathcal{G}_i \subseteq \langle \mathcal{H}_i \rangle$$

Hence, at each step we reduce the task of finding a cover of $\mathcal{H}_{i-1} \setminus \mathcal{G}_i$ to just finding a cover of "smaller" $\mathcal{H}_i$.

Due to lack of assumptions on the hypergraph $\mathcal{H}$, we allow two outcomes at the termination of the algorithm. The first one is exactly the implication of the Theorem 2.2, that we contain in a random set X_m some $S \in \mathcal{H}$ w.h.p. We show that if the first outcome fails, then we are forced into the second outcome which says that $\bigcup_i \mathcal{U}_i$ cheaply covers $\mathcal{H}$. But the fact $\mathcal{H}$ is not p-small in the assumption of our theorem contradicts the second outcome, pulling us back into the first outcome w.h.p. This is what we want, hence Theorem 2.2 follows.

We first show how to construct the cover $\mathcal{U}$ in Section 3.1. For now, we will consider a single iteraion where we throw a random set W of size $Lp|X|$ in $\mathcal{H}$, for some $p \in [0,1]$. Then, we look at the outcome of all the iterations combined and their termination in Section 3.2.

3.1. Constructing the cover.

Let $|X| = n$, $L \geq 1024$, $\mathcal{H}$ an ℓ-bounded hypergraph. Denote by $\binom{X}{w}$ the collection of w-subsets of X, where $w := Lpn$. Let W be uniformly chosen from $\binom{X}{w}$.

We consider the proceeding of a single iteration. Suppose in this iteration we start with the hypergraph $\mathcal{H}$. We sprinkle the set W inside $\mathcal{H}$; for all $S \in \mathcal{H}$ and W, we look at $S \cup W$. Consider all the members S' of $S \cup W$ and for each one call the set $S' \setminus W$ an (S,W)-*fragment*. Given S and W, for the (S,W)-fragment whose size is the minimum out of all the subsets of $S \cup W$, define $T = T(S,W)$ to be that (S,W)-fragment. Thus, $|T(S,W)| = |S' \setminus W|$, where S' is the subset of $S \cup W$ that forms the minimum (S,W)-fragment.

For a given W, let $\mathcal{G} = \mathcal{G}(W)$, (the *good set*) be the collection of S in $\mathcal{H}$ such that the size of the their minimum fragment with W is "large"; we can state this more formally,

$$\mathcal{G}(W) := \{S \in \mathcal{H} : |T(S,W)| \geq .9\ell\}$$

and $\mathcal{G}(W)$ is covered by $\mathcal{U}(W)$, which is defined as:

$$\mathcal{U}(W) := \{T(S,W) : S \in \mathcal{G}(W)\}$$

The motivation behind such definitions is that if we use large $T(S,W)$'s, i.e the minimum fragments, to form a cover, then this cover would be cheap, as we shall see shortly. Observe that $\mathcal{U}(W)$ does in fact cover $\mathcal{G}(W)$ from the way both objects are defined.

Since the leftover edges $\mathcal{H} \setminus \mathcal{G}(W)$ are not necessarily covered by $\mathcal{U}(W)$, since $\mathcal{U}(W)$ is mainly designed to cover only $\mathcal{G}(W)$, we use the leftovers to construct the next hypergraph such that

$$\mathcal{H}' = \mathcal{H}'(W) = \{T(S,W) : S \in \mathcal{H} \setminus \mathcal{G}(W)\} \tag{3.0.2}$$

This new hypergraph $\mathcal{H}'$ contains edges of size at most 0.9ℓ. As we highlighted in the proof overview, we have the crucial property that $\mathcal{H} \setminus \mathcal{G}(W) \subseteq \langle \mathcal{H}' \rangle$, and so a cover of $\mathcal{H}'$ also covers $\mathcal{H} \setminus \mathcal{G}$.

Note that we uptil now, in the construction of this cover, we have not made any assumptions on p. Thus, for W such that $|W| = Lpn$, different sizes of W give different sizes of intersections with $S \in \mathcal{H}$ and hence, different good sets $\mathcal{G}(W)$ and different covers $\mathcal{U}(W)$. The covers $\mathcal{U}(W)$ can be both cheap or expensive, but the following

result shows that for a "typical" W, the random set that we sprinkle, the cover we obtain is cheap.

Lemma 3.1. *For W uniformly chosen from $\binom{X}{w}$, then with a probability of at least $1 - L^{-0.1\ell}$,*

$$\sum_{U \in \mathcal{U}(W)} p^{|U|} < L^{-.5\ell}.$$

With probability of at least $1 - L^{-0.1\ell}$ means that if $A := \left(\sum_{U \in \mathcal{U}(W)} p^{|U|} < L^{-.5\ell}\right)$, then

$$\mathbb{P}(A < L^{-.5\ell}) \geq 1 - L^{-0.1\ell}.$$

Note that this statement follows from

$$\mathbb{E}(A) < L^{-.6\ell} \quad \text{(expectation is over the choice of } W) \tag{3.1.1}$$

since then we have by Markov's Inequality (Theorem 1.4):

$$\mathbb{P}\left(A \geq L^{-0.5\ell}\right) \leq \frac{\mathbb{E}(A)}{L^{-0.5\ell}} < L^{-0.1\ell}.$$

Thus, it is enough to prove 3.1.1. Observe that 3.1.1 is equivalent to

$$\sum_{W \in \binom{X}{w}} \sum_{U \in \mathcal{U}(W)} p^{|U|} < \binom{n}{w} L^{-.6\ell} \tag{3.1.2}$$

Therefore, we shall prove 3.1.2.

Proof of 3.1.2. Given W and $m \geq 0.9\ell$, define

$$\mathcal{G}_m(W) := \{S \in \mathcal{H} : |T(S, W)| = m\}$$

and

$$\mathcal{U}_m(W) := \{T(S, W) : S \in \mathcal{G}_m(W)\}$$

so that sets $U \in \mathcal{U}_m(W)$ have size at most m. Thus, $p^{|U|} = p^m$ and the expression $\sum_{W \in \binom{X}{w}} \sum_{U \in \mathcal{U}(W)} p^{|U|}$ is equal to p^m multiplied by

$$\left|\left\{(W, T(S, W)) : W \in \binom{X}{w}, S \in \mathcal{H}, \text{ and } |T(S, W)| = m\right\}\right| \tag{3.1.3}$$

Hence what is left is to bound the number of choices of W and $T = T(S, W)$ in 3.1.3, such that T is "large" enough to be specified cheaply. We put the bound in the following steps.

Step 1. Firstly we count the number of ways we can pick W and T, independent of the cost of specifying them. Hence, we pick $Z := W \cup T$. We have W and T are always disjoint by construction, so that $|Z| = w + m$. Thus, the number of

ways we can pick Z is bounded above by:

$$\binom{n}{w+m} = \binom{n}{w} \cdot \prod_{j=1}^{m} \frac{n-w-j}{w+j} \leq \binom{n}{w}(Lp)^{-m}$$

Step 2. Now pick any edge of $\mathcal{H}$, $\hat{S} \subseteq Z$. Note that this always possible, since by the definition of minimum fragment $T(S,W) = S' \setminus W$ for some S' that makes the size of $T(S,W)$ the minimum and so, $(S' \setminus W) \cup W = Z \implies Z \supseteq S'$. Hence, the choice of $\hat{S}$ comes for free. But, more importantly $T \subseteq \hat{S}$ by the minimality of T. Now $|\hat{S}| \leq \ell$ implies that we can specify T as a subset of S in 2^ℓ ways.

Since the specifying the pairs (W,T) is determined by fixing choices of Z and T, we have the following bound on our sum

$$\sum_{W \in \binom{X}{w}} \sum_{U \in \mathcal{U}_m(W)} p^{|U|} \leq p^m \binom{n}{w}(Lp)^{-m}2^\ell = \binom{n}{w}L^{-m}2^\ell$$

Lastly for $L \geq 1024$, if we sum over all the $m \geq 0.9\ell$, we have

$$\sum_{m \geq .9\ell} \binom{n}{w}L^{-m}2^\ell \leq \binom{n}{w}L^{-.6\ell} \tag{3.1.4}$$

where the right hand side above is equal to that of the inequality in 3.1.2

$\square$

We restate that we have just shown that a cover produced from a typical choice of W_i (in a single iteration) is cheap, and hence if we sum up the costs of each cover $\mathcal{U}_i$ over all the iterations, the union of the covers $\bigcup_i \mathcal{U}_i$ is still cheap. Recall that the random set W_i that we used in the iteration had size $Lp|X|$. Eventually, we want to have thrown a random m-set of X ($m = (Lp \log \ell)|X|$) that would capture some $S \in \mathcal{H}$. So if W is a disjoint union of all the W_i's, i.e. $W = \bigsqcup_i W_i$, we wish that the number of total iterations is $C \log \ell$ (where C is a universal constant), so that $|W(= X_m)| = (Lp \log \ell)|X|$. ($L$ is again a universal constant, it is irrelevant that we use the same constant from earlier.) Now, we show this formally in the following section.

3.2. **Iteration.** We have as before $|X| = n, L \geq 1024$, and $\ell \to \infty$. Recall that our starting hypergraph $\mathcal{H}_0$ is ℓ-bounded, the next hypergraph $\mathcal{H}_1$ is 0.9ℓ-bounded, and $\mathcal{H}_i$ is $0.9^i\ell$-bounded. Note that a hypergraph $\mathcal{H}_k$ contains an *empty* edge if $\lfloor 0.9^k\ell \rfloor < 1$. Hence, let $i = 1, 2, \cdots, \lfloor \log_{0.9} 1/\ell + 1 \rfloor =: \gamma$, and $\ell_i = 0.9^i\ell$, so that

$$0 < \ell_\gamma < 1. \tag{3.1.5}$$

In order to form disjoint sets W_i, let $X = X_0$ and choose W_i uniformly from $\binom{X_{i-1}}{w_i}$, where $X_i = X_{i-1} \setminus W_i$ and $w_i = L_i pn$ with,

$$L_i = \begin{cases} L & \text{if} \quad i < \gamma - \sqrt{\log_{.9}(1/\ell)} \\ L\sqrt{\log \ell} & \text{if} \quad \gamma - \sqrt{\log_{.9}(1/\ell)} \leq i \leq \gamma \end{cases}$$

These definitions render $W = \bigsqcup_i^\gamma W_i$ such that it is a uniformly random subset of X of size $(KLp\log \ell)n$, where $K \leq K'$ for some absolute constant $K' > 0$. This follows upon after carrying out some calculations which show that there is an absolute constant $k > 0$ such that

$$\ell_i > \exp(k\sqrt{\log \ell}) \quad \forall i < \gamma - \sqrt{\log_{.9}(1/\ell)} \tag{3.1.6}$$

Now, we apply the procedure described in the previous section (3.1) to each iteraton. Doing so produces a sequence $\{\mathcal{H}_k\}$, for which $\mathcal{H}_0 = \mathcal{H}$ is the starting hypergraph and

$$\mathcal{H}_i = \mathcal{H}'_{i-1} \tag{3.1.7}$$

where we use the definition of the next hypergraph $\mathcal{H}'$ as in 3.0.2.

For each i, $\mathcal{H}_i$ is ℓ_i-bounded, and for each W_i, we construct the corresponding *good set* $\mathcal{G}_i = \mathcal{G}_i(W_i)$ and the *cover* $\mathcal{U}_i = \mathcal{U}_i(W_i)$ of $\mathcal{G}_i$. We highlight that $\mathcal{G}_i$ is a sub-graph of $\mathcal{H}_{i-1}$. For $i = 1, 2, \cdots, \gamma$, we call W_i *successful* if it admits a cheap cover $\mathcal{U}_i$ of $\mathcal{G}_i$, that is

$$\sum_{U \in \mathcal{U}(W_i)} p^{|U|} < L_i^{-.5\ell_i}; \text{ and}$$
$$\mathcal{H}_i \text{ does not contain } \emptyset. \tag{3.1.8}$$

Remark 3.2. Note that the second condition for W_i to be successful is essential in our theory. Considering, if $\emptyset$ is contained in our next hypergraph $\mathcal{H}_i$, our algorithm would try to find its cover, but $\emptyset$ has only itself as its cover, for which $\text{cost}_p(\emptyset) = p^0 = 1$ is too expensive for us to afford. Moreover, if we leave it as a *bad set* (the set which does not admit a cheap cover), it will be thrown in next hypergraph and so on, leaving the algorithm interminable.

We terminate our process once we have

$$\mathcal{G}_i = \mathcal{H}_{i-1} \setminus \{\emptyset\}$$

This means that we have to terminate, as soon as we have covered everything in $\mathcal{H}_{i-1}$, except for some empty sets. Strictly speaking, $\mathcal{H}_i \subseteq \{\emptyset\}$ for some $i(=: i_{\max} \leq \gamma)$. If our host hypergraph $\mathcal{H}_{i-1}$ contains some empty edges, that is because they were reduced to size 0 by choosing minimum fragments and only allowing 90% of the previous edge sizes. Note that if, for instance, there are some other *bad* sets (not covered yet) that are not empty, in our current hypergraph, we put them in the next hypergraph and

continue the iterations. However, this is exactly the situation we want to avoid if we want W_i to be successful.

We say the algorithm *terminates successfully*, if W_i's are successful, for all $i \leq i_{\max}$.

Proposition 3.3. *If our process terminates successfully, then the union of the (partial) covers,* $\bigcup_{i \leq i_{max}} \mathcal{U}(W_i)$*, covers* $\mathcal{H}$

Proof. Suppose $\mathcal{U}$ does not fully cover $\mathcal{H}$. Then there is some $S \in \mathcal{H}$ such that $\mathcal{U}$ does not cover S. Consider the evolution of S in our iteration: $S = S_0, S_1, \cdots, S_i \in \mathcal{H}_i$, where $S_i := T(S_{i-1}, W_i)$. As mentioned in the argument above, S not being covered by $\mathcal{U}$ implies that S was a bad set, which eventually was reduced to size 0 (since in each iteration either a set is reduced in size, or it is covered). Thus, there exists some i_0 for which $S_{i_0} = \emptyset$. Since, S_{i_0} lives in $\mathcal{H}_{i_0}$, this means W_{i_0} is not successful, which is a contradiction to successful termination of the process. $\qquad\square$

Suppose we arrive at a successful termination, then the cost of our cover $\mathcal{U}$ is

$$\mathrm{cost}_p(\mathcal{U}) = \sum_{U \in \mathcal{U}} p^{|U|} \overset{(3.1.8)}{<} \sum_{i \leq i_{\max}} L_i^{-.5\ell_i} \overset{(3.1.6)}{\leq} 2L^{-.5\exp(k\sqrt{\log \ell})} + O\left(\left(L\sqrt{\log \ell}\right)^{-k'}\right) \overset{(\star)}{\leq} 1/2,$$

where $k' > 0$ is some constant, and the inequality in ($\star$) holds for ℓ sufficiently large. Hence, we have shown the cover $\mathcal{U}$ is p-cheap. Now in the premises of Theorem 2.2, we have that $\mathcal{H}$ is not p-small, which means $\mathcal{H}$ does not admit any p-cheap cover. Thus, the event that the p-cheap cover $\mathcal{U}$ that we constructed in our randomized algorithm covers $\mathcal{H}$ happens with low probability. Therefore, we will now show that the other case, that $W \in \langle \mathcal{H} \rangle$, happens with high probability. First we have the following proposition,

Proposition 3.4. *If* $\mathcal{H}_i$ *contains* $\emptyset$ *for some i, then $W \in \mathcal{H}$*

Proof. Recall that $W \in \mathcal{H}$ is equivalent to saying $W \supseteq S \in \mathcal{H}$, that W contains some S in $\mathcal{H}$. Suppose there exists some j for which $\mathcal{H}_j$ contains $\emptyset$. From the way we constructed $\{\mathcal{H}_i\}$, we have $\emptyset = S_j, S_{j-1}, \cdots, S_1, S_o := S$ with $S_i \in \mathcal{H}_i$ and $S_i = S_{i-1} \setminus W_i$. (Recall $T(S_{i-1}, W_i) = S_{i-1} \setminus W_i$ and taking the union $W_i \cup T(S_{i-1}, W_i)$ recovers exactly the edge $S_{i-1} \in \mathcal{H}_{i-1}$.) Thus, we have $S_i = S \setminus \left(\bigcup_{k \leq i} W_k \right)$ for $i \in \{1, 2, \cdots, j\}$, which gives $S \subseteq \bigcup_{k \leq j} W_k = W$. $\qquad\square$

We will show Theorem 2.2 by contradiction. If we suppose that the implication in the statement of Theorem 2.2 is not true, then we will show that the event that all W_i's are successful happens with a positive probability. Since all successful W_i's mean we can find a p-cheap cover of $\mathcal{H}$, this gives the contradiction.

Notation.

$$\bigwedge_{i=1}^{n} A_i = A_1 \wedge A_2 \wedge \cdots \wedge A_n \text{ means "all events } A_i \text{ happen"}$$

$$\bigvee_{i=1}^{n} A_i = A_1 \wedge A_2 \wedge \cdots \wedge A_n \text{ means "some event } A_i \text{ happens".}$$

Let η_i denote the event that for W_i, $\sum_{U \in \mathcal{U}(W_i)} p^{|U|} \geq L_i^{-.5\ell_i}$ (Note that this is the event where the first condition for W_i to be successful fails in (3.1.8)). Due to Proposition 3.4, we then have

$$\mathbb{P}\left(\bigvee_{i \leq i_{\max}} \{W_i \text{ not successful }\} \right) \leq \sum_{i \leq i_{\max}} \mathbb{P}(\eta_i) + \mathbb{P}(W \in \langle \mathcal{H} \rangle)$$

$$\leq \sum_{i \leq \gamma} L_i^{-.1\ell_i} + \mathbb{P}(W \in \langle \mathcal{H} \rangle)$$

Further, we have

$$\sum_{i \leq \gamma} L_i^{-.1\ell_i} = \sum_{i < \gamma - \sqrt{\log_{.9}(1/\ell)}} L_i^{-.1\ell_i} + \sum_{i \geq \gamma - \sqrt{\log_{.9}(1/\ell)}}^{\gamma} L_i^{-.1\ell_i}$$

$$\overset{(3.1.5),(3.1.6)}{\leq} 2L^{-.1\exp(k\sqrt{\log \ell})} + O\left((L\sqrt{\log \ell})^{-k''} \right) = o_\ell(1)$$

where k'' is some positive constant. Now, suppose the implication 2.2.1 of our main theorem fails. This implies that there exists a fixed $\varepsilon > 0$ such that

$$\mathbb{P}(W \in \langle \mathcal{H} \rangle) \leq 1 - \varepsilon,$$

This together with the inequality 3.2 implies

$$\mathbb{P}\left(\bigvee_{i \leq i_{\max}} \{W_i \text{ not successful }\} \right) < 1,$$

which in turn implies

$$\mathbb{P}\left(\bigwedge_{i \leq i_{\max}} \{W_i \text{ successful }\} \right) > 0.$$

This concludes the argument, and Theorem 2.2 is proved.

4. APPLICATIONS

When Kahn and Kalai first conjectured this theorem in [3], they remarked: "It would probably be more sensible to conjecture that it is not true"; the reason for such skepticism being the strength of consequences of the conjecture. Since these consequences are out of the scope of this paper, we refer the reader to the original

paper, in which the conjecture first came up [3], and [7] which discusses the significance of the conjecture in light of its consequences for special *hard* problems. Lastly, the proof discussed here also simplifies that of the main lemma in [7].

References

[1] P. ERDŐS AND A. RÈNYI, On random graphs, *I, Publ. Math.*, **6**(1959), 290-297.

[2] J. PARK AND H. T. PHAM, A proof of the kahn-kalai conjecture, *arXiv preprint arXiv:2203.17207.*

[3] J. KAHN AND G. KALAI, Thresholds and expectation thresholds, *Combin. Probab. Comput.*, **16**(2007), no. 3, 495–502.

[4] B. BOLLOBÁS AND A. THOMASON, Threshold functions, *Combinatorica*, **7**(1987), 35-38.

[5] A. FRIEZE AND M. KAROŃSKI, *Introduction to random graphs*, Cambridge University Press, 2016.

[6] B. BOLLOBÁS, *Random Graphs. 2nd ed.*, Cambridge University Press, 2001.

[7] K. FRANKSTON, J. KAHN, B. NARAYANAN, AND J. PARK, Thresholds versus fractional expectation-thresholds, *Ann. of Math. (2)*, **194**(2021), no. 2, 475–495.

[8] R. ALWEISS, L. SHACHAR, K. WU, AND J. ZHANG, Improved bounds for the sunflower lemma, *Ann. of Math. (2)*, **194**(2021), no. 3, 795–815