Thresholds

Talk by Jinyoung Park

September 15–16, 2022

§1 Background and Results

§1.1 Random Graphs

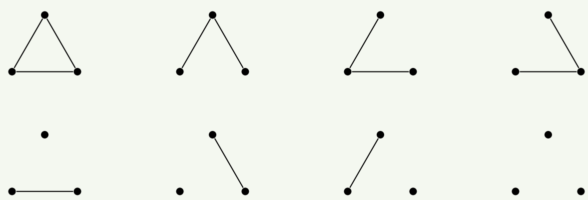
We will begin by discussing *random graphs* — to have a concrete picture in mind, and because historically, random graph theory was the starting point for our story. Later we will see that the setting for Prof. Park's work is more general.

Definition 1.1

The **Erdős–Rényi random graph** has vertex set $[n]$, and each of the $\binom{n}{2}$ potential edges is included with probability p independently.

Example 1.2

When $n = 3$, there are 8 possible graphs:



If $p = 1/2$, all these graphs are equally likely. But if $p = 0.001$ then the sparse graphs are much more likely, while if p is very close to 1 then the dense graphs are much more likely.

Usually, the p we are interested in will be a function of n, and $p \to 0$ as $n \to \infty$ — for example, $p \approx 1/n$ or $p \approx (\log n)/n$.

Note that the random graph is not a fixed graph; rather, it is a probability distribution. So it makes sense to ask for probabilities such as $\mathbb{P}(G_{n,p} \text{ is planar})$ or $\mathbb{P}(G_{n,p} \text{ is connected})$. We're generally not interested in the precise answer, but in *typicality* — as p varies, what's the dominant behavior of $G_{n,p}$ regarding our property?

Definition 1.3

We say $G_{n,p}$ does A **with high probability** (abbreviated whp) if

$$\mathbb{P}(G_{n,p} \text{ does } A) \to 1 \text{ as } n \to \infty.$$

§1.1.1 The Evolution of $G_{n,p}$

One striking thing about $G_{n,p}$ is that the appearance and disappearance of certain properties are "abrupt" — this leads to thresholds.

We can imagine starting with $p = 0$, where we have an empty graph, and increasing p up to $p = 1$, where we have K_n.

Example 1.4

The typical maximal size of connected components of $G_{n,p}$ is

$$\begin{cases} \lesssim \log n & \text{if } np < 1 - \varepsilon \\ \asymp n & \text{if } np > 1 + \varepsilon. \end{cases}$$

So we say that $p = 1/n$ is the **threshold** for $G_{n,p}$ having a giant component, because the behavior of whether $G_{n,p}$ has a giant component changes abruptly at $1/n$.

This happens for many other interesting properties as well, and it's a central interest in probabilistic combinatorics to find thresholds for various properties. Many results have been found for *specific* properties — for example, for $G_{n,p}$ to be connected, or to have long paths, or to have long cycles. We'll see later that the Kahn–Kalai conjecture gives a *unified* result — it actually implies most of these results.

§1.2 Definition of Thresholds

The setting for thresholds is much more general than random graphs. We first fix a few definitions:

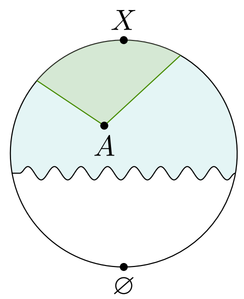
- X is a finite set, and 2^X is the set of subsets of X.
- μ_p is the p-biased product probability measure on 2^X — for each $A \subseteq X$, we have
$$\mu_p(A) = p^{|A|}(1-p)^{|X \setminus A|}.$$
- $X_p \sim \mu_p$ (in words, X_p is a random variable with probability distribution μ_p) — so X_p is a p-random subset of X, which we can think of as choosing elements of X at random with probability p.

Example 1.5

If $X = \binom{[n]}{2}$ (this notation denotes sets of two elements of $[n]$), then $X_p = G_{n,p}$.

Definition 1.6

We say $\mathcal{F} \subseteq 2^X$ is an **increasing property** if for all $A \in \mathcal{F}$ and all $B \supseteq A$, we have $B \in \mathcal{F}$.



In other words, a property is increasing if we can't destroy the property by adding an element (if a subset $A \subseteq X$ satisfies the property $\mathcal{F}$, and we add elements to A, the resulting subset should still satisfy the property).

Example 1.7

For $G_{n,p}$, examples of increasing properties include $\mathcal{F} = \{\text{connected}\}$ and $\mathcal{F} = \{\text{contains a triangle}\}$ — we can't destroy these properties by adding an edge to a graph.

The following is a well-known fact:

Fact 1.8 — For any increasing property $\mathcal{F}$ (other than $\varnothing$ and 2^X), $\mu_p(\mathcal{F})$ is continuous and strictly increasing in p.

Continuity is obvious, because $\mu_p(\mathcal{F})$ is a polynomial in p — it's a sum of terms of the form $p^{|A|}(1-p)^{|X \setminus A|}$ for $A \in \mathcal{F}$, which are all polynomials.

Now we can imagine increasing p from 0 to 1. Then $\mu_p(\mathcal{F})$ increases from 0 to 1 as well. So there must exist a unique value of p for which $\mu_p(\mathcal{F})$ is exactly $1/2$.

Definition 1.9

The **threshold** for $\mathcal{F}$, denoted $p_c(\mathcal{F})$, is the value of p for which $\mu_p(\mathcal{F}) = 1/2$.

Intuitively, for p below the threshold, it's unlikely that X_p satisfies our increasing property $\mathcal{F}$, while for p above the threshold, it's likely that X_p satisfies $\mathcal{F}$.

There's two main directions of study regarding thresholds:

- The *location* of thresholds — historically most work was on thresholds for *specific* properties, but the Kahn–Kalai conjecture suggests a *general* bound. As a preview of what we'll see later, suppose we're given an increasing property $\mathcal{F}$, and we want to find $p_c(\mathcal{F})$. We'll see that there's an *expectation* threshold $q(\mathcal{F})$, which gives a lower bound on $p_c(\mathcal{F})$ and is often easy to compute. Then the Kahn–Kalai conjecture states that it gives an *upper* bound as well, up to a small error — $p_c(\mathcal{F})$ is at most $q(\mathcal{F})$ times a small error.
- The *sharpness* of thresholds — how steep the curve is at the threshold. The main tool in this area is usually Fourier analysis. In fact, people have tried to prove the Kahn–Kalai conjecture using Fourier analysis, but such methods have not been successful — eventually, the tool that Prof. Park used (which will be explained tomorrow) is not at all connected to Fourier analysis.

§1.3 The Kahn–Kalai Conjecture

Now we'll get to the statement of the Kahn–Kalai conjecture. It's a really strong conjecture — in fact, the authors stated that it would be more sensible to conjecture that it is *not* true!

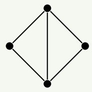
§1.3.1 Some Motivating Examples

Question 1.10. What drives $p_c(\mathcal{F})$?

We'll look at a few examples from random graphs.

Example 1.11

Let $X = \binom{[n]}{2}$, so $X_p = G_{n,p}$, and let $\mathcal{F}_H$ be the property that a graph contains a copy of

$H =$  .

We're interested in whether $G_{n,p}$ has a copy of H, so as usual, we can start by finding the expected value of the *number* of copies of H. We have

$$\mathbb{E}[\# \text{ copies of } H \text{ in } G_{n,p}] \asymp n^4 p^5$$

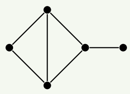
by linearity of expectation (we can choose the four vertices, and then the probability that they form a copy of H is p^5, since we need 5 edges to appear). This means the threshold for the *expected value* $\mathbb{E}$ is on the order of $n^{-4/5}$ — for $p \ll n^{-4/5}$ the expected value goes to 0, while for $p \gg n^{-4/5}$ it grows large.

This gives a trivial lower bound on the threshold as well — we must have $p_c(\mathcal{F}) \gtrsim n^{-4/5}$, since if the expected number of copies of H is very small, usually we must have 0 copies. More precisely, if $\mathbb{E}[X] \to 0$ as $n \to \infty$, then $X = 0$ with high probability.

But it turns out that this lower bound is exactly the right answer! It can be shown that $p_c(\mathcal{F}) \asymp n^{-4/5}$ (by using the second moment method). So at this point, we might dream that $\mathbb{E}$ predicts $p_c(\mathcal{F})$.

Example 1.12

Again consider random graphs, and let $\mathcal{F}_K$ be the property that a graph contains a copy of

$K =$  .

We can again start with the expectation calculation — we have 5 vertices and 6 edges, so

$$\mathbb{E}[\# \text{ copies of } K \text{ in } G_{n,p}] \asymp n^5 p^6,$$

giving the lower bound $p_c(\mathcal{F}_K) \gtrsim n^{-5/6}$. But in this case, that's *not* the correct answer — instead, we have $p_c(\mathcal{F}_K) \asymp n^{-4/5}$, which is much larger! But there's an obvious reason for this lower bound as well — any graph containing K also has to contain H. For $p \ll n^{-4/5}$ the graph typically has no copies of H, and if it doesn't have any copies of H, then it certainly can't have any copies of K.

For *fixed* graphs, the question has been fully answered, and the answer is the same as in this example:

Theorem 1.13

For a fixed graph K, the threshold $p_c(\mathcal{F}_K)$ is equal, up to constant factors, to the expectation threshold for the densest subgraph of K.

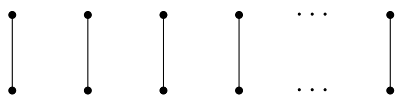
There's two important takeaways from this example — the expectation threshold gives us the right answer, but we need to look at the *core* part and not just the entire graph.

Now let's consider an example where the graph we're looking for is *not* fixed.

Example 1.14

What is the threshold for $G_{n,p}$ to contain a perfect matching (assuming $2 \mid n$)?

Here we're attempting to find a copy of the following graph, which grows with n:



In this case, we have

$$\mathbb{E}[\#\text{ perfect matchings}] \asymp \left(\frac{np}{e}\right)^{n/2},$$

which means the threshold for $\mathbb{E}$ is $1/n$. But we actually have $p_c(\mathcal{F}) \asymp (\log n)/n$, not $1/n$.

But it turns out $(\log n)/n$ is another trivial lower bound — it turns out that if $p \ll (\log n)/n$ then $G_{n,p}$ has an isolated vertex with high probability, and if there's an isolated vertex, there can't be a perfect matching.

One way to think about this behavior is in terms of the *coupon collector problem:*

> **Problem 1.15.** There are n different types of coupons, and each cereal box contains a coupon (the coupons are distributed uniformly at random). How many boxes of cereal do we typically need to buy to collect all n coupons?

The answer is $\asymp n \log n$. In our situation, the coupons are the n vertices, and placing down an edge collects two coupons. So if the number of edges we place is $\ll \log n$, then there's typically an uncollected coupon, meaning an isolated vertex. This gives us a second lower bound of $p_c(\mathcal{F}) \gg (\log n)/n$, which turns out to give the correct answer.

> **Example 1.16** (Shamir's Problem)
> Now take $X = \binom{[n]}{r}$ — the set of r-element subsets of n vertices. Then X_p (where we choose each r-tuple with probability p independently) is the *random r-uniform hypergraph*, denoted $\mathcal{H}^r_{n,p}$. For $r \geq 3$ with $r \mid n$, what is the threshold for $\mathcal{H}^r_{n,p}$ to contain a perfect matching?

The $r = 2$ case was solved by Erdős–Rényi in 1966, but for $r \geq 3$ the problem is much harder (since we no longer have an analog of Hall's theorem).

For $r = 3$, we can try a similar approach. The lower bound we get using expectation is $\asymp 1/n^2$, while the lower bound we get using the coupon collector argument (the threshold for not having isolated vertices) is $\asymp (\log n)/n^2$. People expected the second bound to be the correct value, and this was proved by Johansson–Kahn–Vu in 2008.

§1.3.2 The Kahn–Kalai Conjecture

In the first two examples, we saw that the threshold for $\mathbb{E}$ drives the threshold for $p_c(\mathcal{F})$, while in the last two examples, we saw that coupon collector-ish behavior pushes $p_c(\mathcal{F})$ up from the expectation threshold by a factor of $\log n$. This leads to the Kahn–Kalai conjecture, that this behavior is true in general:

> **Conjecture 1.17** (Kahn–Kalai Conjecture, 2006)
> For any increasing property $\mathcal{F}$, the threshold $p_c(\mathcal{F})$ is at most $\log |X|$ times the expectation threshold.

This conjecture is very strong! The proof by Johansson–Kahn–Vu for Shamir's problem was very hard, but if the conjecture is true, then the answer is a one-line corollary — it's easy to compute that the expectation threshold is $1/n^{r-1}$, so the Kahn–Kalai conjecture would give an upper bound of $(\log n)/n^{r-1}$, which is exactly the lower bound from coupon collector. It would also answer a different problem, the "tree conjecture" on the threshold for bounded-degree spanning trees.

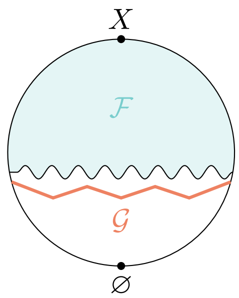
Both of these are difficult problems — they were unanswered at the time the Kahn–Kalai conjecture was stated, making the conjecture difficult to believe — it seemed more like wishful thinking. But it turns out that it's actually true!

§1.3.3 The Expectation Threshold

First, for an abstract property $\mathcal{F}$, it's unclear whose expectation we want to compute. So we need a careful definition of the expectation threshold.

Note that (as a generalization of the expectation calculation) we have $p_c(\mathcal{F}) \geq q$ if there exists $\mathcal{G} \subseteq 2^X$ such that the following two properties hold:

- $\mathcal{G}$ covers $\mathcal{F}$ — for all $A \in \mathcal{F}$, there exists $B \in \mathcal{G}$ with $A \supseteq B$.



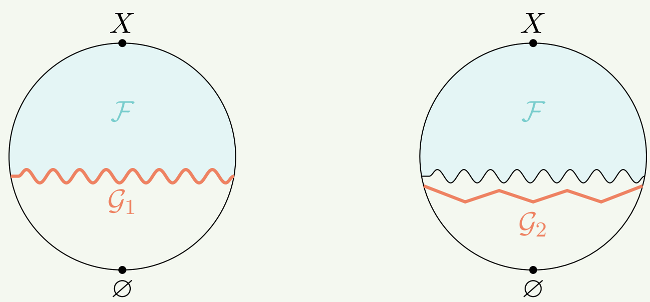
- We have $\sum_{S\in\mathcal{G}} q^{|S|} \leq 1/2$.

Intuitively, the quantity $\sum_{S\in\mathcal{G}} q^{|S|}$ is the expected number of elements of $\mathcal{G}$ that cover X_q — for any $S \in \mathcal{G}$ to cover X_q, we must choose all its $|S|$ elements in X_q, and each has probability q of being chosen. So then if $\sum_{S\in\mathcal{G}} q^{|S|} \leq 1/2$, the probability that *any* $S \in \mathcal{G}$ covers X_q is also at most $1/2$, and since every $A \in \mathcal{F}$ is covered by some $B \in \mathcal{G}$, this means the probability $X_q \in \mathcal{F}$ is at most $1/2$ as well. So we must have $p_c(\mathcal{F}) \geq q$.

Example 1.18

In Example 1.12, where $\mathcal{F}$ is the property of containing a copy of K, a trivial cover for $\mathcal{F}$ is the set $\mathcal{G}_1$ of labelled copies of K. In this case, $\sum_{S\in\mathcal{G}} q^{|S|} \asymp q^5 n^6$ (this is the expected number of copies of K in X_q). We have $q^5 n^6 \lesssim 1/2$ for $q \asymp n^{-5/6}$, so this gets the bound $p_c(\mathcal{F}) \gtrsim n^{-5/6}$.

But we can do better — instead, choose $\mathcal{G}_2$ to be the set of all labelled copies of H. Then $\mathcal{G}_2$ still covers $\mathcal{F}$ — any element of $\mathcal{F}$ contains a copy of H as well — and now $\sum_{S\in\mathcal{G}} q^{|S|} = q^4 n^5$, giving the better lower bound $p_c(\mathcal{F}) \gtrsim n^{-4/5}$.



So any cover gives a lower bound, and we want to take the most useful one:

Definition 1.19

The **expectation threshold** of $\mathcal{F}$, denoted $p_E(\mathcal{F})$, is the greatest q such that there exists a cover $\mathcal{G}$ of $\mathcal{F}$ with $\sum_{S\in\mathcal{G}} q^{|S|} \leq 1/2$.

Now we can properly state the Kahn–Kalai conjecture:

Conjecture 1.20 (Kahn–Kalai Conjecture)

There exists a universal constant $K > 0$ such that for every finite set X and increasing property $\mathcal{F}$,

$$p_E(\mathcal{F}) \leq p_c(\mathcal{F}) \leq Kp_E(\mathcal{F}) \log |X|.$$

In random graph theory, this is very meaningful because the expectation threshold is easy to compute. On the other hand, if $\mathcal{F}$ is very abstract, then computing $p_E(\mathcal{F})$ may be hard.

One interpretation of the conjecture is that a cover is the most naive way to approximate our increasing propert, and the Kahn–Kalai conjecture states that even this naive approximation gives us the correct answer up to a factor of $\log |X|$.

§1.4 Results

The first result was the *fractional* version of the Kahn–Kalai conjecture. There is also a *fractional expectation threshold* — we won't define this, but it replaces the cover $\mathcal{G}$ with a *fractional cover*. Let this threshold be p_E^*; then it's clear that $p_E(\mathcal{F}) \leq p_E^*(\mathcal{F}) \leq p_c(\mathcal{F})$.

Theorem 1.21 (Conjectured by Talagrand 2010, proved by Frankston–Kahn–Narayanan–Park 2019)

There exists $K > 0$ such that for every finite X and increasing $\mathcal{F} \subseteq 2^X$,

$$p_c(\mathcal{F}) \leq Kp_E^*(\mathcal{F}) \log \ell(\mathcal{F}),$$

where $\ell(\mathcal{F})$ is the size of a largest minimal element of $\mathcal{F}$.

This is weaker than the original Kahn–Kalai conjecture, but it's still strong enough to get all the known applications, since in those applications we know $p_E(\mathcal{F}) \asymp p_E^*(\mathcal{F})$. In fact, it was conjectured by Talagrand that $p_E(\mathcal{F}) \asymp p_E^*(\mathcal{F})$ is true in general; this would imply the equivalence of the Kahn–Kalai conjecture and the fractional Kahn–Kalai conjecture.

It was expected that a proof of the full Kahn–Kalai conjecture would use this, but in fact it doesn't. So by the Kahn–Kalai conjecture we now know $p_E^*(\mathcal{F}) \leq Kp_E(\mathcal{F}) \log \ell(\mathcal{F})$ is always true (for a constant K) — in particular if $\ell(\mathcal{F})$ is constant, then there's no gap between p_E^* and p_E. But the general case of that conjecture remains open.

Theorem 1.22 (Conjectured by Kahn–Kalai 2006, proved by Park–Pham 2022)

There exists $K > 0$ such that for every finite X and increasing $\mathcal{F} \subseteq 2^X$,

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

The proof uses a simple and direct argument. It's only 6 pages long (and the last page is the references, and the first two pages the introduction)! The full idea of the proof will be explained tomorrow.

§1.4.1 Further Questions

We've now seen that

$$p_E(\mathcal{F}) \leq p_c(\mathcal{F}) \lesssim p_E(\mathcal{F}) \log \ell(\mathcal{F}).$$

> **Question 1.23.** What characterizes the gap between $p_c(\mathcal{F})$ and $p_E(\mathcal{F})$?

In many cases, the $\log \ell(\mathcal{F})$ gap is tight. But there are some cases where it isn't, and those serve as good test cases. One such example is the property of containing the *square* of a Hamiltonian cycle. Here the expectation threshold is $\asymp n^{-1/2}$, and it was conjectured in 2012 by Kühn–Osthus that this is the correct value (i.e. there is no extra log factor). This was proven by Kahn–Narayanan–Park in 2020, using similar methods to the proof of the fractional Kahn–Kalai conjecture.

Another example is the property of containing a triangle factor — disjoint triangles that cover all the vertices. Here the threshold has been proven to be $n^{-2/3}(\log n)^{1/3}$ — we have a strange exponent of $1/3$ (which comes from the threshold for all vertices to be in a triangle). The simpler argument used to prove the Kahn–Kalai conjecture can't get this fractional exponent of the log factor.

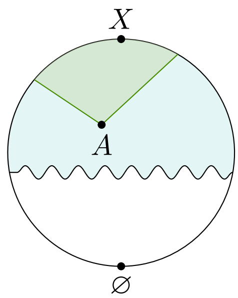
Another question is what would happen if we changed the random graph model to not have isolated vertices (since these are in some sense what push up the threshold from the expectation threshold). This line of question is still open.

§2 Proof of the Kahn–Kalai Conjecture

§2.1 Definitions

We'll begin by restating the definitions from yesterday's talk.

- X is a finite set, and 2^X the set of subsets of X.
- μ_p is the p-biased product probability measure on 2^X — for each subset $A \subseteq X$, we have $\mu_p(A) = p^{|A|}(1-p)^{|X \setminus A|}$.
- $X_p \sim \mu_p$ (meaning that X_p is the random variable with distribution given by μ_p) — so X_p is a p-random subset of X, meaning that we choose every element of X with probability p independently.
- $\mathcal{F} \subseteq 2^X$ is an *increasing property* — whenever $A \in \mathcal{F}$, its *up-set* $\langle A \rangle := \{B \subseteq X \mid B \supseteq A\}$ must also be contained in $\mathcal{F}$.



- We define $\mu_p(\mathcal{F}) := \sum_{A \in \mathcal{F}} \mu_p(A)$.
- If $\mathcal{F}$ is not $\varnothing$ or 2^X, then as p increases from 0 to 1, so does $\mu_p(\mathcal{F})$. So there exists a unique $p_c(\mathcal{F})$ where $\mu_p(\mathcal{F}) = 1/2$; this is called the **threshold** for $\mathcal{F}$.

Theorem 2.1 (Kahn–Kalai Conjecture)
There exists a constant K such that for all X and for all $\mathcal{F} \supset 2^X$,

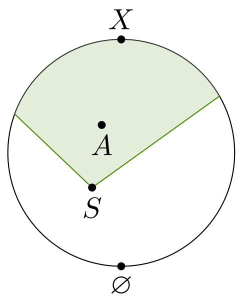
$$p_c(\mathcal{F}) \leq K \cdot q(\mathcal{F}) \cdot \log \ell(\mathcal{F}),$$

where $q(\mathcal{F})$ is the *expectation threshold* and $\ell(\mathcal{F})$ the size of the largest minimal element of $\mathcal{F}$.

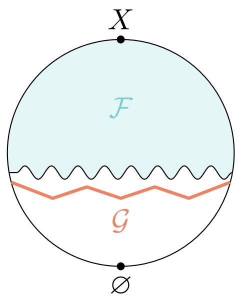
We'll now review the definition of the expectation threshold. In what follows, assume that $A, S \subseteq X$ and $\mathcal{F}, \mathcal{G} \subseteq 2^X$.

Definition 2.2
We say that S **covers** A if $S \subseteq A$.

In other words, S covers A if A is contained in the up-set $\langle S \rangle$.



Definition 2.3
We say $\mathcal{G}$ covers $\mathcal{F}$ if for all $A \in \mathcal{F}$, there exists $S \in \mathcal{G}$ such that S covers A.



In other words, $\mathcal{G}$ covers $\mathcal{F}$ if $\langle \mathcal{G} \rangle \supseteq \mathcal{F}$ (where $\langle \mathcal{G} \rangle$ is the union of the up-sets $\langle A \rangle$ over all $A \in \mathcal{G}$).

Yesterday, we saw the observation that $p_c(\mathcal{F}) \geq q$ if there exists $\mathcal{G}$ that covers $\mathcal{F}$ such that

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

If some $\mathcal{G}$ satisfies ($*$), then we say $\mathcal{G}$ is *q-cheap*.

Definition 2.4
The **expectation threshold** $q(\mathcal{F})$ is the maximal q for which there exists a q-cheap cover $\mathcal{G}$.

§2.2 Overview of the Proof

The theorem we'll actually prove is the following standard reformulation of the Kahn–Kalai conjecture:

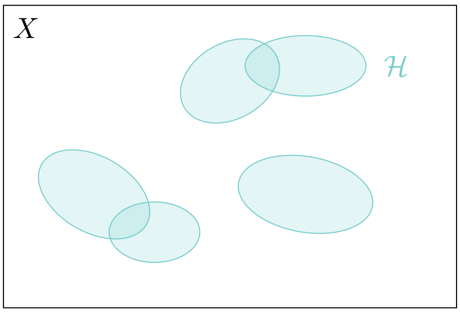
Theorem 2.5 (Reformulation of the Kahn–Kalai Conjecture)
There exists $L > 0$ such that for all ℓ-bounded $\mathcal{H}$, if $p > q(\langle \mathcal{H} \rangle)$, then if we let $m = Lp \log \ell \cdot |X|$,

$$\mathbb{P}(X_m \text{ contains a member of } \mathcal{H}) = 1 - o(1) \text{ as } \ell \to \infty.$$

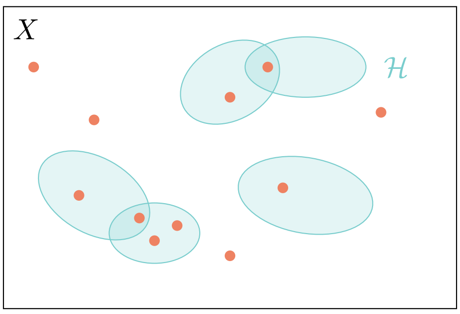
To say $\mathcal{H}$ is ℓ-bounded means that for all $A \in \mathcal{H}$ we have $|A| \le \ell$. Here X_m means a random subset of X with m elements.

For some intuition on why this implies the Kahn–Kalai conjecture, think of $\mathcal{H}$ as the collection of minimal elements of $\mathcal{F}$, so that $\langle \mathcal{H} \rangle = \mathcal{F}$. Then we're considering random subsets X_m with $Lp \log \ell \cdot |X|$ elements; this is similar to choosing random subsets where we choose each element with probability around $Lp \log \ell$. Now if X_m is very likely to contain a member of $\mathcal{H}$, it's also very likely to contain a member of $\mathcal{F}$ (since the members of $\mathcal{H}$ are all members of $\mathcal{F}$). So this provides an upper bound $p_c(\mathcal{F}) \lesssim Lp \log \ell$ as well. (In the statement ℓ is fixed, but L is an absolute constant, so we can take ℓ to be the size of the largest minimal element of $\mathcal{F}$.)

To visualize this, we can think of X as a universe, and $\mathcal{H}$ as a collection of subsets of X.



We then sprinkle in m random elements, and the statement says that typically these m elements contain some member of $\mathcal{H}$.



First, here is an overview of the proof.

(1) We'll use W to denote X_m. We choose W little by little — at each step we choose a W_i with $|W_i| = Lp|X|$ at random (such that the W_i are disjoint), and we set $W = W_1 \sqcup W_2 \sqcup \cdots$. We can choose as many as $\log \ell$ of these intermediate sets W_i.

(2) As we choose W little by little, $\mathcal{H}$ will evolve as well — we'll have $\mathcal{H} = \mathcal{H}_0 \to \mathcal{H}_1 \to \mathcal{H}_2 \to \cdots$.

(3) In the end, we want to have $W \supseteq S$ for some $S \in \mathcal{H}$ with high probability.

(4) We'll apply a randomized algorithm (where the randomness comes from the choice of W_i) — we iteratively produce a *partial cover* $\mathcal{U}(W)$ of $\mathcal{H}$, by building a partial cover $\mathcal{U}_i(W_i)$ at each step of the algorithm and taking $\mathcal{U}(W) = \bigcup \mathcal{U}_i(W_i)$. (A *cover* $\mathcal{G}$ of $\mathcal{F}$ would mean that every subset in $\mathcal{F}$ contains some subset in $\mathcal{G}$; a *partial cover* means that we only cover some part of $\mathcal{F}$, not necessarily all of it.)

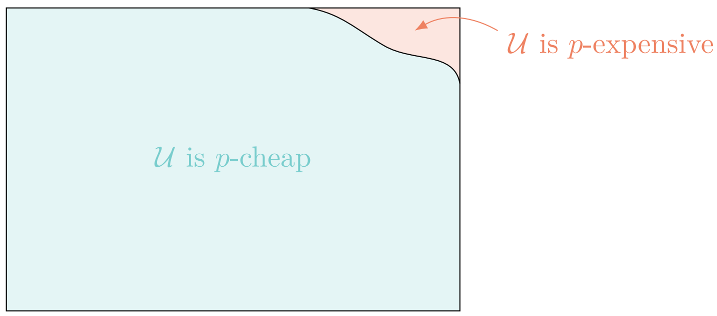
(5) The main point of the proof is that our partial cover $\mathcal{U}(W)$ will be p-cheap with high probability.

(6) When the algorithm terminates (we'll see the termination condition later), either:

   (1) $\mathcal{U}$ entirely covers $\mathcal{H}$, or

   (2) W contains an element $S \in \mathcal{H}$.

If we believe this, then we're done with the proof — consider the sample space for the choice of W. Then by (5), most of the time $\mathcal{U}$ is p-cheap; it's very unlikely that $\mathcal{U}$ is expensive.



But now we can apply our assumption that $p > q(\langle \mathcal{H} \rangle)$ — we know that $q(\langle \mathcal{H} \rangle)$ is the largest q that admits a q-cheap cover of $\mathcal{H}$. So if $p > q(\langle \mathcal{H} \rangle)$, then there does not exist a p-cheap cover of $\langle H \rangle$. This means if $\mathcal{U}$ covers $\mathcal{H}$, then it must be expensive!

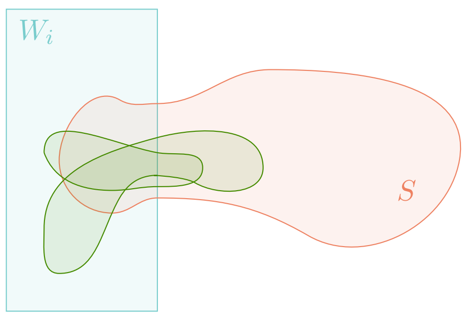
But it is unlikely that $\mathcal{U}$ is expensive. So when the algorithm terminates, (1) must be unlikely, which means (2) occurs with high probability. But (2) is exactly what we're looking for.

§2.3 Constructing $\mathcal{U}_i(W_i)$

First we'll describe how to construct $\mathcal{U}_i(W_i)$ in a given step; we'll then iterate this construction at most $\log \ell$ times.

In this step, our host hypergraph is $\mathcal{H}_{i-1}$; suppose $\mathcal{H}_{i-1}$ is s-bounded for some s (initially $\mathcal{H}$ is ℓ-bounded, but s will change as $\mathcal{H}$ changes).

Suppose we've chosen W_i (this is done at random). Then we examine all $S \in \mathcal{H}_{i-1}$ and decide whether we want to cover them or not. First, for each S, we look at all elements of $\mathcal{H}_{i-1}$ that sit inside $W_i \cup S$.



Now of these subsets in $\mathcal{H}_{i-1}$ inside $W_i \cup S$, we let S' be the subset with minimal $|S' \setminus W_i|$ (if there's multiple, we choose arbitrarily).

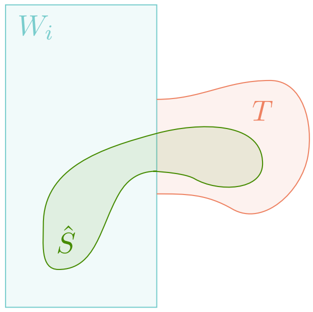
Definition 2.6

Given S and W_i, we define $T = S' \setminus W_i$ as the **minimal (S, W_i)-fragment**.

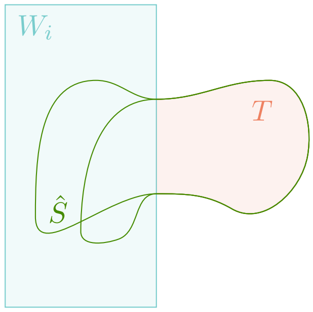
This minimal fragment will be the key gadget of the proof. There are a few key observations:

(1) There must exist some $\hat{S} \in \mathcal{H}_{i-1}$ with $\hat{S} \subseteq W_i \cup T$ (by definition — since T is the piece of some S' that lies outside W_i).

(2) For every $\hat{S} \subseteq W \cup T$, we must have $T \subseteq \hat{S}$ — otherwise, this would violate the minimality of T, as $\hat{S} \setminus W$ would be a smaller fragment:



So we have the following picture:



This will be the core of what makes our partial cover cheap.

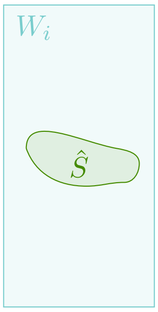
Now we say that S is *good* if T is large — if $|T| \geq 0.9s$. In this case, we put T in $\mathcal{U}_i$. Otherwise, if T is small then it's not affordable, so we don't place it in $\mathcal{U}_i$, and instead we place T in our next hypergraph $\mathcal{H}_i$.

So in this step, for each $S \in \mathcal{H}_{i-1}$ we either add its minimal fragment to $\mathcal{U}_i(W_i)$ — which covers S — or we replace it with a subset whose bound is smaller by a factor of at least 0.9. In particular, this is why our algorithm will perform at most $\log \ell$ steps — at the start $\mathcal{H}$ is ℓ-bounded, after one step it's 0.9ℓ-bounded, after two steps it's $0.9^2\ell$-bounded, and so on.

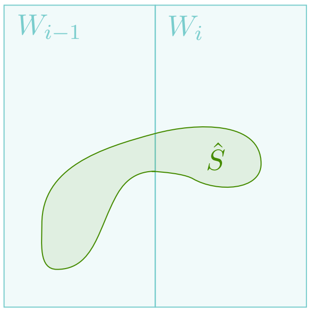
§2.4 Termination Conditions

Now that we've described how to construct our partial cover, we'll describe when the algorithm terminates.

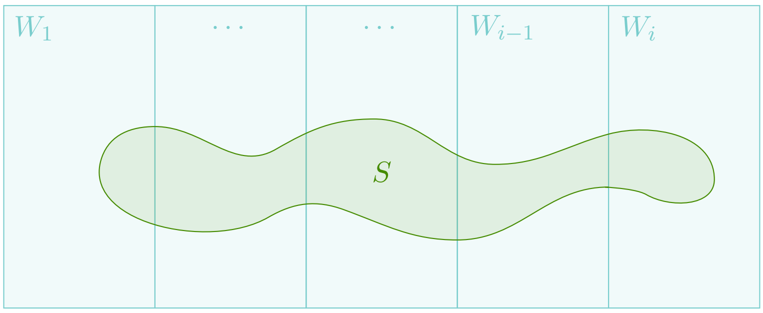
We terminate the algorithm as soon as some set has minimal fragment $T = \varnothing$. If this occurs, then we must have some $\hat{S} \in \mathcal{H}_{i-1}$ sitting entirely in W_i.



Then we claim we've reached our second goal of (2) — that $W = W_1 \sqcup W_2 \sqcup \cdots$ contains some element of $\mathcal{H}$. To see this, note that since $\hat{S}$ was in $\mathcal{H}_{i-1}$, it must have been the minimal fragment of some set in $\mathcal{H}_{i-2}$ (since we only ever insert minimal fragments into $\mathcal{H}_{i-1}$). This means we had a set in $\mathcal{H}_{i-2}$, and then W_{i-1} "ate" a piece of it and left us with $\hat{S}$.



But *this* set is in $\mathcal{H}_{i-2}$, so it's the minimal fragment of some set in $\mathcal{H}_{i-3}$. We can keep extending backwards to get a set $S \in \mathcal{H}$ from the beginning of the process:



So then we started off with $S \in \mathcal{H}$, and some part of it got eaten by W_1, then W_2, and so on; and its final part got eaten by W_i. This means $W = W_1 \sqcup W_2 \sqcup \cdots$ covers S.

On the other hand, suppose this never happens. Then we keep running the process until there's nothing left in $\mathcal{H}$. But in every step, we look at all S in $\mathcal{H}_{i-1}$, and either we add their minimal fragment T to $\mathcal{U}_i(W_i)$ — which covers S — or we add T to $\mathcal{H}_i$ (and if we later cover T, that set covers S as well). So if $\mathcal{H}$ ends up empty, then $\mathcal{U}(W) = \mathcal{U}_1(W_1) \cup \mathcal{U}_2(W_2) \cup \cdots$ covers all sets S originally in $\mathcal{H}$. So we've reached our first goal of (1).

§2.5 p-Cheapness of $\mathcal{U}(W)$

Now there's one remaining piece of the proof — that our partial cover $\mathcal{U}(W)$ is cheap with high probability. The key point is the following:

Lemma 2.7

Let $|X| = n$, $w = Lpn$, and suppose $\mathcal{H}$ is s-bounded. Then

$$\sum_{W_i \in \binom{X}{w}} \sum_{U \in \mathcal{U}_i(W_i)} p^{|U|} < \binom{n}{w} L^{-0.8s}.$$

This implies that the *average* cost of $\mathcal{U}_i(W_i)$ (over all possibilities for W_i) is less than $L^{-0.8s}$. This can be used to show that the average cost of $\mathcal{U}(W)$ is small as well, and then Markov's inequality shows that $\mathcal{U}(W)$ is usually p-cheap.

Proof. We use double-counting. Recall that if $U \in \mathcal{U}_i(W_i)$, then we must have $|U| \geq 0.9s$ (since this is our condition for adding T to our partial cover). For simplicity assume $|U| = 0.9s$ (the calculations are messier in the general case). Then the left-hand side becomes

$$p^{0.9s} \sum_{W_i \in \binom{X}{w}} \sum_{U \in \mathcal{U}_i(W_i)} 1.$$

But the double summation simply counts pairs (W_i, U) where U is the minimal fragment of somebody (satisfying the size condition) — so we just want to count pairs $(W_i, T(S, W_i))$ where $W_i \in \binom{X}{w}$ and $S \in \mathcal{H}$, and $|T(S, W_i)| = 0.9s$. But this means

$$p^{0.9s} \sum_{W_i \in \binom{X}{w}} \sum_{U \in \mathcal{U}_i(W_i)} 1 \leq p^{0.9s} \cdot \binom{n}{w + 0.9s} \cdot 2^s.$$

To see this, there are $\binom{n}{w+0.9s}$ ways to choose $W \cup T$. Once we've chosen $W \cup T$, we can use our observations about minimal fragments — for T to be somebody's minimal fragment, we must have some $\hat{S} \in \mathcal{H}$ with $\hat{S} \subseteq W \cup T$, and then T must sit inside $\hat{S}$. There's at most 2^s subsets of $\hat{S}$, so at most 2^s possible T.

Now this sum is at most

$$p^{0.9s} \cdot \binom{n}{w} \cdot (Lp)^{-0.9s} \cdot 2^s = \binom{n}{w} \cdot L^{-0.9s} \cdot 2^s < \binom{n}{w} L^{-0.8s}.$$

So we are done. $\square$