# Intuitive Exposition of Erdős-Rado Sunflower

by throwing pokeballs $W$ to catch just half a pokemon $S \in \mathcal{F}$

Daniel Rui 6/30/25

*Definition 0.1: Erdős-Rado sunflower function*

Fix integers $r, k \geqslant 2$. $\mathrm{Sun}(r,k) :=$ **smallest** positive number s.t.
EVERY ("sufficiently large") set-family $|\mathcal{F}| \geqslant \mathrm{Sun}(r,k)$ of [size $\leqslant k$]-sets $\subseteq U$ (any universe $U$; no relation to $r, k$, i.e. $\mathrm{Sun}(r,k)$ doesn't care what $U$ is! **w.l.o.g.** $U$ is as large as we want.) is guaranteed to contain a $r$-sunflower (i.e. $S_1, \ldots, S_r \in \mathcal{F}$ s.t. $\forall i, j \in [1..r]$, $S_i \cap S_j = \bigcap_{l=1}^{r} S_l$).

Another way of thinking about it: let $\mathcal{F}^{\heartsuit}$ denote the/a **largest** family of ($\leqslant k$)-sets $\subseteq U$ (any possible universe) that contains NO $r$-sunflowers. Then $\mathrm{Sun}(r,k) := |\mathcal{F}^{\heartsuit}| + 1$.

Erdős-Rado 1962 proved that $\mathrm{Sun}(r,k) \leqslant (r-1)^k \cdot k!$ (so in particular, finite), by looking at the "link"/"petals" of singleton sets in $\mathcal{F}$, i.e. $\mathcal{F}_{\{a\}} := \{S \setminus \{a\} : \{a\} \subseteq S \in \mathcal{F}\}$, i.e. "cut away one element at a time", until getting to a situation where they could guarantee $r$ disjoint sets (the simplest kind of $r$-sunflower). Sunflower conjecture: $\mathrm{Sun}(r,k) \leqslant (C(r))^k$, or even $\lesssim r^k$.
As for the lower bound, an easy explicit example gives $\mathrm{Sun}(r,k) > (r-1)^k$.

The Erdős-Rado proof proceeds by induction using "petals", a notion very well-suited to induction on problems about sunflowers, because a sunflower in the "petals set" $\mathcal{F}_I$ lifts to a sunflower in $\mathcal{F}$ (just add $I$ to the core). We will continue to use "petals" for the purpose of induction in the below proof, but "petals" of more than just singleton sets.

Their idea of inductively cutting away until you get a situation where you can guarantee $r$ disjoint sets is also fruitful; it remains at the heart of the below proof.

Blue highlight for skeleton/landmarks to help structure/navigate the proof.

*Theorem 0.2: AWZ'20 Erdös-Rado sunflower bound*

$\mathrm{Sun}(r,k) < s^k$ for a parameter $s = s(r,k) \geqslant 2r$ I'll reveal by the end.
I.e. we show that every family $|\mathcal{F}| \geqslant s^k$ of ($\leqslant k$)-sets is guaranteed to contain a $r$-sunflower.

*Proof:* by induction on $k$.

For $k = 2$, follows from the Erdős-Rado 1962 result.

Now if there is some set $I \subseteq U$ s.t. the set of "petals" of $I$ in $\mathcal{F}$, i.e. $\mathcal{F}_I := \{S \setminus I : I \subseteq S \in \mathcal{F}\}$, has size $\geqslant s^{k-|I|}$, then we'd be done by the induction hypothesis. So in the rest of the proof, we may assume that

$$\forall I \subseteq U, \quad |\mathcal{F}_I| \leqslant \frac{1}{s^{|I|}} \cdot |\mathcal{F}| \iff \Pr_{S \in \mathcal{F}}(I \subseteq S) \leqslant \left(\frac{1}{s}\right)^{|I|}$$

i.e. "asking to contain any element cuts down the size of $\mathcal{F}$ by constant factor of at least $\frac{1}{s}$", e.g. $\forall i \in U, \quad \mathcal{F}_{\{i\}} \leqslant \frac{1}{s} \cdot |\mathcal{F}|$ and $\forall i, j \in U, \quad \mathcal{F}_{\{i,j\}} \leqslant (\frac{1}{s})^2 \cdot |\mathcal{F}|$.

This is a notion known as spreadness, which has since become a staple of other combinatorial results (e.g. see 3AP and corners work in additive combinatorics), where finding the right definition of spreadness is the/a key behind the proof; but in this problem at least, it comes out very naturally from just the induction.

As I said/foreshadowed above, we have boiled the problem down to a special case of the original theorem:

*Theorem 0.3: AWZ'20 spreadness implies many disjoints*

Just assuming $\mathcal{F}$ is $s$-spread (which actually implies $|\mathcal{F}| \geqslant s^k$, where $k = \max_{S\in\mathcal{F}}|S|$), we will produce $r$ disjoint sets $S_1, \dots, S_r \in \mathcal{F}$ (i.e. a degenerate $r$-sunflower).

We do this by the **probablistic method**: choosing $W^1, \dots, W^{2r}$ ("witnesses for disjointness") a uniform random partition of $U$ to pieces of size $\frac{1}{2r}|U|$ (**w.l.o.g.**, by making $U$ bigger, may assume $|U|$ is divisible by anything we want), we hope that in expectation, $\geqslant r$ of them contain an $S \in \mathcal{F}$.

---

So, we have boiled down to the goal: for a unif. random $W$ of size $\frac{1}{2r}|U|$, with probability $\geqslant \frac{1}{2}$, we can get/"catch" an $S \in \mathcal{F}$ inside. "Toss pokeball randomly, catch a pokemon with high probability."

This turns out to be true, but how on earth do we prove it?

MAIN PLAN: **cut the problem in half ("dyadic decomposition")**: "toss pokeball $W_1$ (of some size $< |W| = \frac{1}{2r}|U|$) randomly, catch **half a pokemon** with high probability" (a little barbaric ... lol).

In other words, after "Round 1", we hope to find many "Round 1" remains/leftovers $\mathcal{L}_1 := \mathcal{F} - W_1 = \{S \setminus W_1 : S \in \mathcal{F}\}$ that are "halfsmall", i.e. size $\leqslant \frac{k}{2}$.

And then our hope is that by repeating this iteratively $\bar{\log}_2 k := \lceil \log_2 k \rceil$ many times, we will have caught one pokemon spread dyadically over $\bar{\log}_2 k$ many pokeballs $W_1, \dots, W_{\bar{\log}_2 k}$ (aha! so we realize we want each of these "pokeballs" $W$ to have size $\frac{1}{2r}|U|/\bar{\log}_2 k$).

In order for this plan to work, we'd in particular want to show something like [there are *(lots of)* halfsmall leftovers $L$ in $\mathcal{L}_1$] $\iff$ [there are very *few* LARGE leftovers $L$]. This is still hard to prove (perhaps even false). But, we *can* show [there are very *few* LARGE minimal leftovers $L \not\supseteq L$].

I.e. [w/ high prob. $\mathbf{Pr}_{W_1}$, many leftovers $L \in \mathcal{L}_1 := \mathcal{F} - W_1$ will contain $\supseteq$ a halfsmall leftover $L$].

---

For a given $W_1$ (randomly chosen subset $\subseteq U_1 := U$ of size $m := \frac{1}{2r}|U|/\bar{\log}_2 k$), we define the set of leftovers (see "big picture" Fig. 1 below)

$$\mathcal{L}_1 := \mathcal{F} - W_1 := \{S \setminus W_1 : S \in \mathcal{F}\},$$

and the set $\mathcal{E}_1$ of "exceptionally LARGE" yet "minimal" leftovers (see "big picture" Fig. 1 below) where "minimal" is w.r.t. set containment (i.e. for $E \in \mathcal{E}_1$, there is no other $L \in \mathcal{L}_1$ s.t. $L \subseteq E$ — notated as $E \not\supseteq L$ for short), and "exceptionally large" just means size $\geqslant \frac{k}{2}$:

$$(\mathcal{L}_1 \supseteq)\mathcal{E}_1 := \{E \in \mathcal{L}_1 : |E| \geqslant \frac{k}{2}, E \not\supseteq L \ (\forall L \in \mathcal{L}_1)\}$$

[See also Round 1 of the BIG illustration below for a visual focused on these definitions.]

Remember, our goal now is to show LARGE minimal leftovers $\mathcal{E}_1 \ni E \not\supseteq L$ are *rare*. More precisely, we will upper bound $\mathbf{E}_{W_1 \subseteq U_1} |\mathcal{E}_1|$.

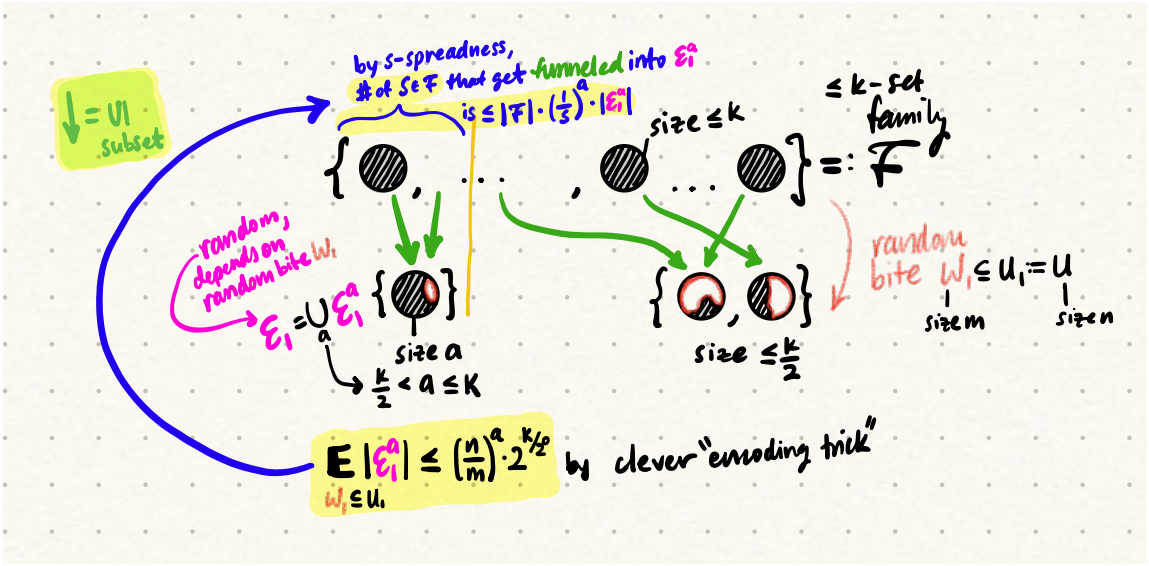



Figure 1: GAMEPLAN/BIG PICTURE: Each $S \in \mathcal{F}$ has green arrow to (arbitrarily chosen, if there are many candidates) smallest contained leftover $S \supseteq L \in \mathscr{L}_1$. Some $S$ have green arrow to $E \in \mathscr{C}_1^a$, LARGE minimal leftover. How many such $S$? Use $s$-spreadness!

To upper bound $\mathbf{E}_{W_1 \subseteq U_1} |\mathscr{C}_1|$ (Fig. 2) (recall the set $\mathscr{C}_1 = \mathscr{C}_1(W_1)$ of LARGE minimal-leftovers $L \not\supseteq L$ depends on the unif. random $m$-set $W_1 \subseteq U_1$), we use the idea of **lossless encoding**. Notation $|\mathscr{C}_1|$ to visually indicate it is a small-cardinality set of LARGE minimal-leftovers.

I.e. we show that pairs $(W_1, E)$ for $E \in \mathscr{C}_1 = \mathscr{C}_1(W_1)$ can be losslessly encoded as tuples of data of a certain "format", of which there are not that many possible.

We further stratify by size $|E| = a$ for $\frac{k}{2} \leqslant a \leqslant k$ (*surprisingly critical* in the finale!). Denote by $\mathscr{C}_1^a \subseteq \mathscr{C}_1$ the collection of such $E$ of size $a$. We also explicitly set $U_1 := U = \{1, \ldots, n_1\} =: [n_1]$ where $n_1 := n := |U|$, so as to fix an order of the elements ($\rightsquigarrow$ lexical order for sets of such elts. too).

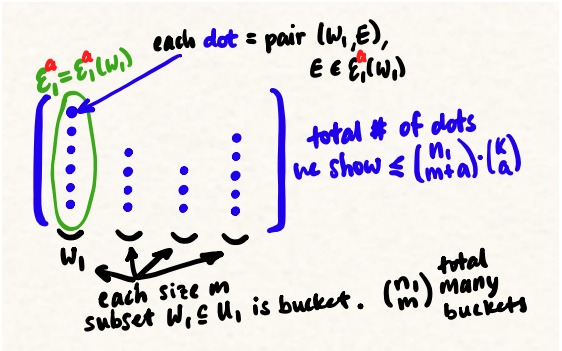



Figure 2: $\mathbf{E}_{W_1 \subseteq U_1} |\mathscr{C}_1^a|$ as avg. height of tower

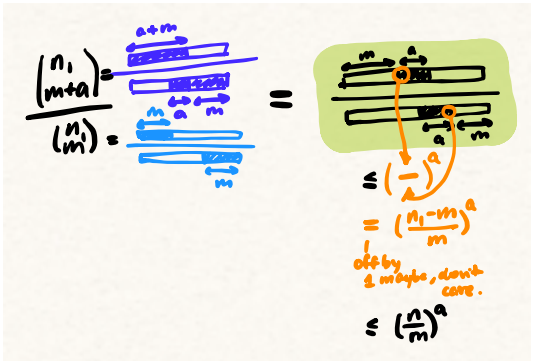



Figure 3: Binomial coefficient upper bound

Without further ado, let us define and then explain the extremely clever encoding of $(W_1, E)$ for $|E| = a$ (note it is clearly invertible/lossless) [see also big illustrated diagram on next page]:

$$\Phi := (W_1, E) \mapsto (\underbrace{W_1 \sqcup E}_{\in \binom{U_1}{m+a}}, \underbrace{S'}_{\substack{\text{completely} \\ \text{determined} \\ \text{by } W_1 \sqcup E}}, \underbrace{E}_{\in \binom{S'}{a}}), \quad \text{i.e. the data of } \left(\begin{smallmatrix} m+a \\ \text{numbers} \\ \text{in } [n] \end{smallmatrix}, \varnothing, \begin{smallmatrix} a \\ \text{numbers} \\ \text{in } [k] \end{smallmatrix}\right),$$

where $S' :=$ the lexicographically first set in $\mathscr{F}$ contained in $W_1 \sqcup E$ (there is such an $S'$ because recall $E$ is in particular a leftover in $\mathscr{L}_1 := \mathscr{F} - W_1$, meaning that there is some "pokemon" $S_E \in \mathscr{F}$ that was partially eaten by "bite" $W_1$ leaving behind remains/leftover $E = S_E \setminus W_1 \implies S_E \subseteq W_1 \sqcup E$).

Furthermore, $E$ is indeed a subset of $S'$ (and hence can be encoded as $a$ numbers in $S' \hookrightarrow [k]$, as indicated (see underbrace under $E$) in the definition of $\Phi$ above), because

$$S' \subseteq W_1 \sqcup E \iff S' \setminus W_1 \subseteq E, \text{ which by MINIMALITY(!!!) is } \iff S' \setminus W_1 = E.$$

So you can alternatively define $S'$ as the (lexicographically) first set in $\mathscr{F}$ whose leftover (after "bite" $W_1$) is $E$. But the **crucial** point of the original definition of $S'$ (I even wrote it in its underbrace in $\Phi$!) is that you can determine $S'$ **solely** from the set $W_1 \sqcup E$; i.e. you do NOT need to know $W_1$ and $E$ separately, just their union!

The number of such tuples of data [data of the above "format"] (again, $n_1 := n$, $k, a$ are fixed) is:

$$\#\left\{\left(\begin{smallmatrix} m+a \\ \text{numbers} \\ \text{in } [n_1] \end{smallmatrix}, \varnothing, \begin{smallmatrix} a \\ \text{numbers} \\ \text{in } [k] \end{smallmatrix}\right)\right\} = \binom{n_1}{m+a} \cdot \binom{k}{a}.$$

And so we may upper bound (see Fig. 2 above again for illustration)

$$\mathop{\mathbf{E}}_{W_1 \subseteq U_1} |\mathscr{C}_1^a| \leqslant \frac{\binom{n_1}{m+a} \cdot \binom{k}{a}}{\binom{n_1}{m}} \leqslant \left(\frac{n}{m}\right)^a \cdot \binom{k}{a} \leqslant \left(\frac{n}{m}\right)^a \cdot 2^{k/2^0} = (2r\bar{\log}_2 k)^a \cdot 2^{k/2^0}, \qquad (❦)$$

by the manipulation of binom. coeffs. illustrated in Fig. 3 above (also $\frac{k}{2^1} \leqslant a \leqslant \frac{k}{2^0} \implies \binom{k}{a} \leqslant 2^{k/2^0}$).

---

Strictly speaking, **this section is UNNECESSSARY**; but I've written it to foreshadow the final descent later. Also, note that I have not $\sum_a$ to bound $\mathbf{E}_{W_1} |\mathscr{C}_1|$; in fact, one **MUST NOT** sum in $a$ now!

Interlude: let's now check back in on our plan of catching half a pokemon.

Recall: we threw one "pokeball" $W_1$, and then we hoped that it took a big bite out of (lots of) "pokemon" $S \in \mathscr{F}$, i.e. for (lots of) "pokemon" $S \in \mathscr{F}$ are s.t. the leftover $S \setminus W_1$ (after the "bite" $W_1$) is halfsmall.

The upper bound above now allows us to prove a weaker statement (we need the extra assumption of minimality for the above encoding to work!): that an average $S \in \mathscr{F}$ and average "bite" $W_1$ [leaves a leftover $L := S \setminus W_1 \in \mathscr{L}_1$ that contains $\supseteq$ a halfsmall leftover $\iota$] $\impliedby$ [$S$ contains no LARGE minimal leftover $E \in \mathscr{C}_1$] (" $\impliedby$ " true b/c leftover $L := S \setminus W_1$ not containing LARGE minimal leftover, means it must contain halfsmall leftover).

We do this in turn by defining $\mathscr{C}_1(W_1)^{\subseteq S} := \{E \in \mathscr{C}_1(W_1) : E \subseteq S\}$ (which we hope is likely empty), and upper bounding (using spreadness $\mathop{\mathbf{Pr}}_{S \in \mathscr{F}}(E \subseteq S) \leqslant (\frac{1}{s})^{|E|}$):

$$\mathop{\mathbf{E}}_{S \in \mathscr{F}} \mathop{\mathbf{E}}_{W_1 \subseteq U_1} |\mathscr{C}_1^{\subseteq S}| := \mathop{\mathbf{E}}_{W_1} \sum_{\frac{k}{2} \leqslant a \leqslant k} \sum_{E \in \mathscr{C}_1^a} \mathop{\mathbf{E}}_{S \in \mathscr{F}} [\mathbb{1}_{E \subseteq S}] \leqslant \mathop{\mathbf{E}}_{W_1} \sum_{\frac{k}{2} \leqslant a \leqslant k} \sum_{E \in \mathscr{C}_1^a} \left(\frac{1}{s}\right)^a \leqslant \sum_{\frac{k}{2} \leqslant a \leqslant k} \left(\frac{1}{s}\right)^a \cdot \mathop{\mathbf{E}}_{W_1} |\mathscr{C}_1^a|;$$

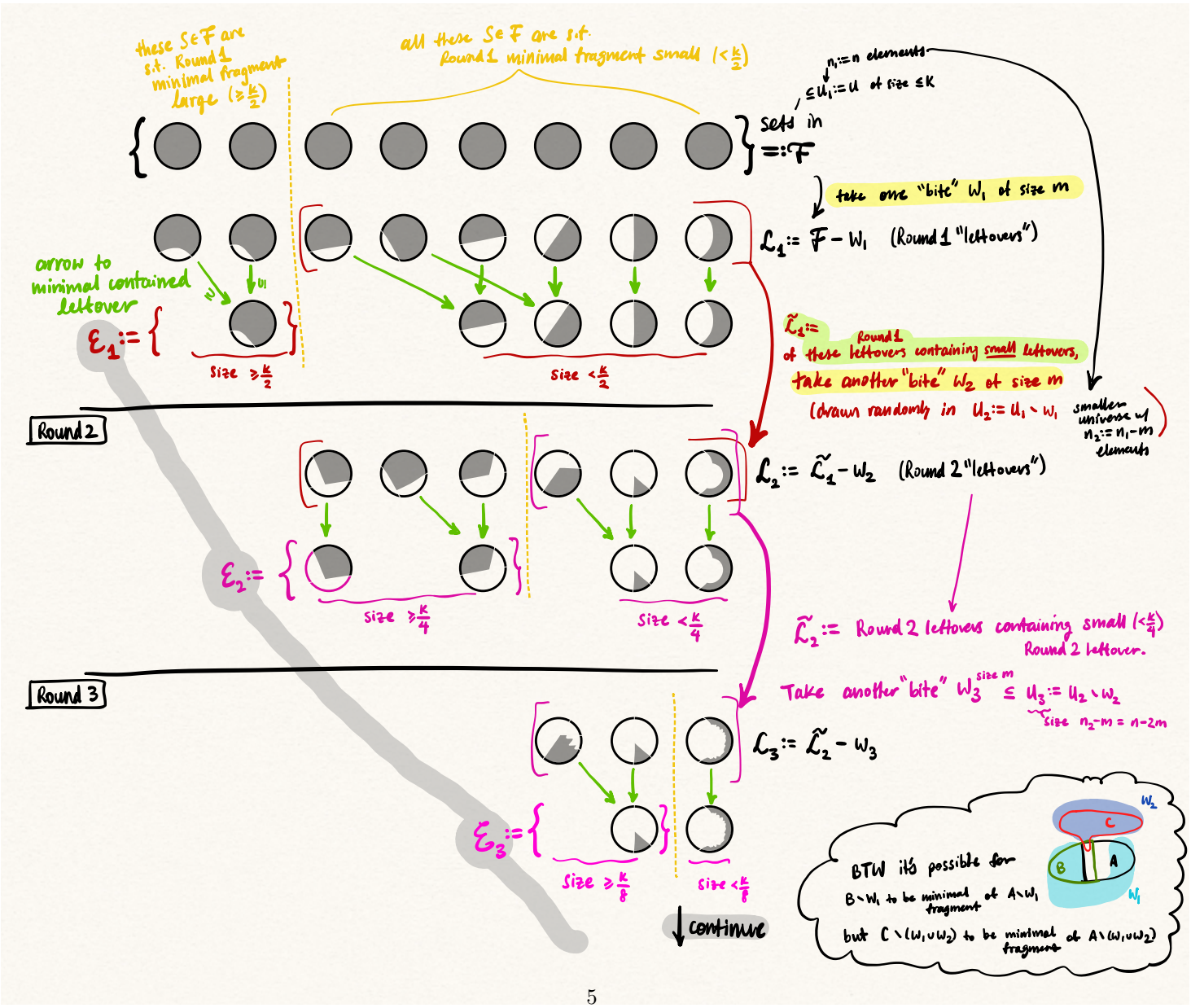
which upon applying the above upper bound (❦), taking $s := 128r\bar{\log}_2 k$, and using that the sum of a geometric series (ratio $< \frac{1}{2}$) is within a factor of 2 of the largest term, produces:

$$\mathop{\mathbf{E}}_{S\in\mathcal{F}} \mathop{\mathbf{E}}_{W_1\subseteq U_1} |\mathscr{C}_1^{\subseteq S}| \leqslant \ldots(\text{above})\ldots \leqslant 2^{k/2^0}\cdot \sum_{\frac{k}{2}\leqslant a\leqslant k} \left(\frac{2r\bar{\log}_2 k}{s}\right)^a \leqslant 2\cdot 4^{k/2^1}\cdot\left(\frac{1}{64}\right)^{k/2^1} = 2\cdot\left(\frac{1}{16}\right)^{k/2^1}.$$

So indeed with high probability, **we do indeed catch at least half a pokemon** (more precisely, the inequality says that for random $W_1 \subseteq U_1$ and $S \in \mathcal{F}$, with high probability $\mathscr{C}_1(W_1)^{\subseteq S}$ is empty, implying in particular that the minimal-leftover $\iota$ contained in the leftover $L := S \setminus W_1$ must be halfsmall, i.e. size $< \frac{k}{2}$, meaning $W_1$ actually "bit" off at least half of some $S' \in \mathcal{F}$)

---

To conclude the proof, we just have to iterate this $\lceil \log_2 k \rceil$ many times. **Warning: below, sometimes I don't draw an arrow $\overset{\supseteq}{\to}$, even though based on the drawing, one leftover may look like it contains some other leftover**. The arrows are to be considered more accurate than the drawings.



The average # of $E \in \mathscr{C}_i$ contained in avg. $S \in \mathscr{F}$ can be upper bounded exactly as above (Exercise; just go through the Interlude and replace sub/superscript 1 or 0 with $i$ or $i-1$), to get:

$$\mathop{\mathbf{E}}_{S\in\mathscr{F}} \mathop{\mathbf{E}}_{W_i\subseteq U_i} |\mathscr{C}_i^{\subseteq S}| \leqslant \sum_{\frac{k}{2^i}\leqslant a\leqslant \frac{k}{2^{i-1}}} \left(\frac{1}{s}\right)^a \cdot \mathop{\mathbf{E}}_{W_i} |\mathscr{C}_i^a| \leqslant \sum_{\frac{k}{2^i}\leqslant a} \left(\frac{1}{s}\right)^a \cdot \left(\frac{n_i}{m}\right)^a 2^{k/2^{i-1}} \leqslant 2\Big(\frac{2r\bar{\log}_2 k}{s}\Big)^{k/2^i} 4^{k/2^i}$$

$$= 2\cdot\left(\frac{1}{16}\right)^{k/2^i}$$

At the end of $\bar{\log}_2 k$ many iterations, define $W \coloneqq W_1 \sqcup \ldots \sqcup W_{\bar{\log}_2 k}$ and $\mathscr{C} \coloneqq \mathscr{C}_1 \sqcup \ldots \sqcup \mathscr{C}_{\bar{\log}_2 k}$.

Define $\mathbf{E}_W$ to be expectation over $W$ sampled in this way (i.e. uniform random $W_1 \subseteq U_1$, then uniform random $W_2 \subseteq U_2 \coloneqq U_1 \smallsetminus W_1$ etc.; then take union of all $W_i$) — turns out this is identical to just uniformly sampling $W \subseteq U$ of size $\bar{\log}_2 k \cdot m = \frac{1}{2r}|U|$.

Also think of $\mathscr{C} \coloneqq \mathscr{C}(W)$, a random (stratified) collection depending on random (stratified) set $W$.

We can now make our final descent :

$$\mathop{\mathbf{E}}_{W}\mathop{\mathbf{E}}_{S\in\mathscr{F}} |\mathscr{C}^{\subseteq S}| \coloneqq \mathop{\mathbf{E}}_{W}\mathop{\mathbf{E}}_{S\in\mathscr{F}} \sum_{i=1}^{\bar{\log}_2 k} |\mathscr{C}_i(W_i)^{\subseteq S}| = \sum_{i=1}^{\bar{\log}_2 k} 2\cdot\left(\frac{1}{16}\right)^{k/2^i} < \frac{1}{2};$$

where the bound "$< \frac{1}{2}$" comes from using the "geometric series factor of 2" bound (note the series is even better than geometric) — the greatest term is $(\frac{1}{16})^{k/2^{\bar{\log}_2 k}} \leqslant (\frac{1}{16})^{1/2} = \frac{1}{4}$. Of course ("1st moment method"!!!),

$$\mathop{\mathbf{E}}_{W}\mathop{\mathbf{E}}_{S\in\mathscr{F}} |\mathscr{C}^{\subseteq S}| \geqslant \mathop{\mathbf{Pr}}_{W}\left(\mathop{\mathbf{E}}_{S\in\mathscr{F}} |\mathscr{C}(W)^{\subseteq S}| \geqslant 1\right)\cdot 1.$$

Putting the two pieces together tells us

$$\mathop{\mathbf{Pr}}_{W}\left(\mathop{\mathbf{E}}_{S\in\mathscr{F}} |\mathscr{C}(W)^{\subseteq S}| < 1\right) \geqslant \frac{1}{2}.$$

And so indeed with probability $\geqslant \frac{1}{2}$, picking a uniform random $W \subseteq U$ of size $\bar{\log}_2 k \cdot m \coloneqq \frac{1}{2r}|U| \iff$ uniform random $W_1, \ldots, W_{\bar{\log}_2 k}$ (picked iteratively, i.e. $W_1 \subseteq U_1$, then $W_2 \subseteq U_2 \coloneqq U_1 \smallsetminus W_1$, etc.), and forming the collections of large minimal leftovers $\mathscr{C}_1, \ldots, \mathscr{C}_{\bar{\log}_2 k}$, we conclude there MUST be some $S \in \mathscr{F}$ containing NONE of these large minimal leftovers...

...meaning that $S$ must have "passed all the way through" the rightmost side of my illustration/diagram above...

...i.e. each $W_i$ "bit" off at least half of what remained of $S$ at the beginning of Round $i$...

...meaning $W$ ate all of $S$, i.e. $W \supseteq S$ ("$W$ caught a pokemon $S$") for some $S \in \mathscr{F}$! ■

To be super clear: for a fixed $W$ ($\subseteq U$ of size $\frac{1}{2r}|U|$), by construction of $\mathscr{C}$ (as unravelled in the preceding ¶),

$$\left(\mathop{\mathbf{E}}_{S\in\mathscr{F}} |\mathscr{C}(W)^{\subseteq S}| < 1\right) \implies \exists S \in \mathscr{F} \text{ s.t. } \underbrace{|\mathscr{C}(W)^{\subseteq S}| = 0}_{\implies S\subseteq W}.$$

Let's bring it back to pg. 1: we proved that an $s$-spread family $\mathscr{F}$ of ($\leqslant k$)-sets for $s \coloneqq 128\cdot r\cdot \bar{\log}_2 k$ (any $r, k \geqslant 2$) MUST contain $r$ disjoint sets (which we can find by considering a random partition of $U$ into $2r$ equal-sized pieces; in expectation $\geqslant r$ of them contain an $S \in \mathscr{F}$). For $r = 2$, cf. Erdős-Ko-Rado.

I should say that I first learned the above proof in a course in Wi25 by R.M. (the Big Illustration is in particular heavily inspired by R.M.'s illustration from lecture). I also leaned heavily on Anup Rao's existing exposition on the topic, "From Soil to Oil".

Thanks to Siddarth Mulherkar for pointing out an issue with an earlier version of this article, where I summed in $a$ too early, leading to a premature bound of

Summing (❦) over $\frac{k}{2} \leqslant a \leqslant k$ and recalling that a sum of a geometric series is proportional (within a factor of 2 here b/c $\frac{n}{m} \geqslant 2$) to its largest term (also recalling $m \coloneqq \frac{1}{2r}|U|/\bar{\log}_2 k$), we get the upper bound

$$\mathop{\mathbf{E}}_{W_1 \subseteq U_1} |\mathscr{C}_1| \leqslant 2 \cdot \left(\frac{n}{m}\right)^{k/2^0} \cdot 2^{k/2^0} \leqslant 2 \cdot \left(2r\bar{\log}_2 k\right)^{k/2^0} \cdot 2^{k/2^0} \leqslant 2 \cdot \left(2 \cdot 2 \cdot r\bar{\log}_2 k\right)^{k/2^0}.$$

It may be instructive to see these exact ideas written out in a cleaner and more professional (but in my opinion, less intuitive) way, in the form of a proof for the Kahn-Kalai conjecture. The following is taken from my notes on a class lecture given by A.B. on 10/25/25. (For background, see **Appendix**)

Set-up, $X$ with size $N$.

$Y \sim X_p$ means $Y$ is a $p$-random subset of $X$ (every element $x \in X$ is included independently with probability $p$).

For $\mathcal{F} \subseteq \mathcal{P}ow(X)$, the notation $\langle \mathcal{F} \rangle$ denotes upward closure.

Def: $\mathcal{F}$ is cover of an increasing family $\mathcal{H}$ if $\langle \mathcal{F} \rangle \supseteq \mathcal{H}$. All $\mathcal{H}$ in this lecture are increasing and nontrivial.

$\mathcal{H}^{\downarrow} := \{\subseteq\text{-minimal member of } \mathcal{H}\}$.

$\mathcal{H}$ is $L$-bounded if minimal elts are size $\leqslant L$, i.e. $|S| \leqslant L$ for all $s \in \mathcal{H}^{\downarrow}$.

DEF: $p \in [0,1]$, $\mathcal{F} \subseteq \mathcal{P}ow(X)$, the $p$-capture of $\mathcal{F}$ is $c_p(\mathcal{F}) := \sum_{S \in \mathcal{F}} p^{|S|} = \mathbb{E}_{Z \sim X_p}[\#$ of sets $S \in \mathcal{F}$ included in $Z]$ — "the expected number of $S \in \mathcal{F}$ captured by a $p$-random subset of $X$".

*Theorem 0.4: Kahn-Kalai conjecture (Park-Pham '22)*

$\mathcal{H} \subseteq \mathcal{P}ow(X)$ increasing nontrivial family, $L$-bounded ($L \geqslant 2$), then for all $p \in [0,1]$, at least one of the following happens: (2 winning conditions)
(winCase1) $\mathcal{H}$ has a cover $\mathcal{F}$ with $c_p(\mathcal{F}) \leqslant \frac{1}{2}$. (In terms of thresholds, the expectation threshold $p \leqslant t^*_{\mathcal{H}}$ by 1stMM)
(winCase2) For $q := \min\{p \cdot 100 \log L, 1\}$, then $\mathbb{P}_{Y \sim X_q}[Y \in \mathcal{H}] \geqslant \frac{1}{2}]$ (this means that $q \geqslant t_{\mathcal{H}}$).
We assume $100p \log L < 1$ ("suff. small"), since otherwise the theorem is trivial.

The reason the theorem is stated like this, is because you are oblivious to which one will happen (you don't care). Very elegant proof.

Main lemma — the shape of the lemma will make it clear how the proof will go.

*Lemma 0.5: Main lemma*

Fix $\mathcal{H} \subseteq \mathcal{P}ow(X)$ increasing, $L$-bounded for $L \geqslant 1$. Fix $p \in [0,1]$ ("suff small").
Pick and fix $Y \sim X_{50p}$.
Then, there are families $\mathcal{F}, \mathcal{H}' \subseteq \mathcal{P}ow(X)$ ("$\mathcal{F}$ is our attempt reach winCase1) s.t.
(1) $\mathcal{F} \cup \mathcal{H}'$ is a cover of $\mathcal{H}$.
(2a) $\mathcal{H}'$ is increasing
(2b) $\mathcal{H}'$ is $\frac{L}{2}$-bounded (inclusion-minimal members are small sets)
(2c) (there has to be some relation between the things we're constructing and the random $Y$ we have; that's this bullet point) $\forall A \in \mathcal{H}'$, $A \cup Y \in \mathcal{H}$ (i.e. $Y$ is almost in $\mathcal{H}$, just need to add a tiny $A$).
(3) $\mathbb{E}[c_p(\mathcal{F})] \leqslant 10^{-L}$ ($\mathcal{F}$ is cheap)

*Proof:* $\mathcal{H} := \mathcal{H}_Y := \{A \subseteq X : A \cup Y \in \mathcal{H}\}$, i.e. sets that are almost in $\mathcal{H}$, but just need a bit of help from $Y$. Like $\mathcal{H} - \mathcal{P}ow(Y) = \{H \setminus \mathrm{y} : H \in \mathcal{H}, \mathrm{y} \in \mathcal{P}ow(Y)\}$. In other words, $A \in \mathcal{H}_Y$ means that $A \in \mathcal{H}$ "modulo" discrepancies from/[contained in] $Y$.

$\mathcal{H}$ clearly contains $\supseteq \mathcal{H}$, and is also an increasing family.

$\mathcal{H}^{\downarrow} = \{\subseteq\text{-minimal elts. of } \mathcal{H}\}$ — sets $A \subseteq X$ that "got in" to $\mathcal{H}$ with the help of $Y$, but remove a single element and $Y$ is not enough to help them into $\mathcal{H}$ (note: this forces $A \not\subset Y$). Is a cover of $\mathcal{H}$, hence also of $\mathcal{H}$.

Imagine $Y$ is the "grading curve", the free padding on select assignment grades that a professor gives everyone, and these minimal sets are the ones that are 1pt above passing threshold lol.

Remember we want $\mathscr{H}'$ to be $\frac{L}{2}$-bounded (2b). So,

- define ("$M$" for minimal) $\mathscr{H}' := \langle \{M \in \mathscr{H}^{\downarrow}_{\sim Y} : |M| \leqslant \frac{L}{2}\} \rangle$. Obviously $\subseteq \tilde{\mathscr{H}}$ (2c) and increasing (2a), and by def. $\frac{L}{2}$-bounded (2b).
- $\mathscr{F} := \{M \in \mathscr{H}^{\downarrow}_{\sim Y} : |M| > \frac{L}{2}\}$ — LARGE minimal/"just barely" "helped$_Y$-into-$\mathscr{H}$" sets. Imagine these as sad "borderline pass even with $Y$-padding" student, that did a LOT of work that just wasn't valued in the grading scheme $\mathscr{H}$.
- Note $\mathscr{F} \cup \mathscr{H}' \supseteq \mathscr{C} := \mathscr{H}^{\downarrow}_{\sim Y}$ (1).

(1), (2abc) hold by definition. So (3) ("$\mathscr{F}$ is evasive, hard to $p$-capture") is the entire crux of the Kahn-Kalai conjecture.

The tension is that sets in $\mathscr{F}$ are inclusion-minimal, but large. That is why we should "expect" to catch/capture very few of them.

Note that $\mathscr{F}$ is a function of random $Y$. For fixed $W$, we write $\mathscr{F}(W) :=$ the value of $\mathscr{F}$ given that $Y = W$.

**Claim** (deterministic, no randomness!): for fixed $w$ and $m > \frac{L}{2}$, the [# of pairs $(W, M) \in \binom{X}{w} \times \binom{X}{m}$ s.t. $M \in \mathscr{F}(W)$] is $\leqslant \binom{N}{w+m}\binom{L}{m}$.

Note, the trivial upper bound is $\leqslant \binom{N}{w+m}\binom{w+m}{m}$. The magic of the Claim is that somehow we can replace $w+m$ by $L$.

*Proof of Claim:* fix $(W, M)$ with $M \in \mathscr{F}(W) \subseteq \mathscr{H}^{\downarrow}_{\sim W} \subseteq \mathscr{H}_{\sim W} \supseteq \mathscr{H}$.

We do not use that $M$ is large in the first half; it only matters in the Big Calculation. I will use notation $M$ for now, and then bring back $M$ for the Big Calculation.

Recall that $M \in \mathscr{H}^{\downarrow}_{\sim W}$ means that "with the help of $W$" it "gets in": $M \cup W \in \mathscr{H}$, but is "inclusion-minimal", meaning that removing a single element would "disqualify" it (no longer getting into $\mathscr{H}$, even with the help of $W$).

Note that $M$ (in fact any elt. of $\mathscr{H}^{\downarrow}_{\sim W}$) must be disjoint from $W$, because o.w. I could remove a common element from $M$ and keep the union in $\mathscr{H}$, ※.

Fix (in advance, independent of $W, M$, globally) an arbitrary function $f : \mathscr{H} \to \mathscr{H}^{\downarrow}$ s.t. for all $A \in \mathscr{H}$, $f(A) \subseteq A$.

Well, we have $M \sqcup W \in \mathscr{H}$, so $S := f(M \sqcup W) \in \mathscr{H}^{\downarrow} \subseteq \mathscr{H}$.

Draw picture $X$ horizontal oval, $W$ = left hemisphere of oval, $M$ is little barnacle grabbing onto (right edge of) $W$. Draw $S \subseteq W \cup M$ as thin tube connecting $W$ and barnacle $M$.

$(S \setminus W) \sqcup W \supseteq S$ (obviously; or look at picture!) where $S \in \mathscr{H}$, implying that $S \setminus W \in \tilde{\mathscr{H}}(W)$ (by def; indeed $S \setminus W$ union $W$ "gets in" to $\mathscr{H}$)

Ok, $S \setminus W \subseteq M$ (look at picture!!!) $\implies S \setminus W = M$ by MINIMALITY(!!!) of $M$. So $M \subseteq S$. (Revise picture so that $S$ contains $M$.)

COUNT: # of such pairs $(W, M)$.

- Pick union $M \sqcup W$, $\quad \leqslant \binom{N}{w+m}$ choices.
- Compute $S := f(M \sqcup W)$.

- Choose $M$ from $\binom{S}{m}$, $\leqslant \binom{L}{m}$ choices.
- Set $W \coloneqq (M \sqcup W) \setminus M$.

Total: $\leqslant \binom{N}{w+m}\binom{L}{m}$. $\boxtimes$*(Claim)*

Finally, we show (3) by calculating the expected $p$-capture of $\mathscr{F}$:

$$\begin{aligned}
\mathbb{E}_{Y\sim X_p}[c_p(\mathscr{F}(Y))] &\coloneqq \mathbb{E}_{Y\sim X_p}\left[\sum_{M\in\mathscr{F}(Y)} p^{|M|}\right] \\
&= \sum_{w=0}^{N}\sum_{W\subseteq X,|W|=w} \underbrace{(50p)^w(1-50p)^{N-w}}_{=\mathbb{P}[Y=W]} \cdot \sum_{M\in\mathscr{F}(W)} p^{|M|} \\
&\leqslant \sum_{w=0,(|W|=w)}^{N} \underbrace{(50p)^w(1-50p)^{N-w}}_{=\mathbb{P}[Y=W]} \cdot \sum_{m>\frac{L}{2},(|M|=m)} \underbrace{\binom{N}{w+m}\binom{L}{m}}_{\geqslant\#\text{ pairs }(W,M)\text{ size }(w,m)} \cdot \underbrace{p^m}_{=p^{|M|}} \\
&= \sum_{m>\frac{L}{2}} \binom{L}{m} p^m \cdot \sum_{w=0}^{N} \binom{N}{w+m}(50p)^w(1-50p)^{N-w} \\
&= \sum_{m>\frac{L}{2}} \binom{L}{m} \frac{(1-50p)^m}{50^m} \underbrace{\sum_{w=0}^{N} \binom{N}{w+m}(50p)^{w+m}(1-50p)^{N-w-m}}_{\leqslant 1\text{ (by binomial formula!)}} \\
&\leqslant \sum_{m>\frac{L}{2}} 2^L \cdot 50^{-m} \leqslant 2^L \cdot 50^{-L/2} \cdot \frac{50}{49} < 10^{-L}
\end{aligned}$$

for all $L \geqslant 1$. ■

*Proof of Theorem:* let $K \coloneqq \bar{\log}_2 L$.
Independently pick and fix $Y_1, \ldots, Y_K \sim X_{50p}$.
Let $L_0 \coloneqq L$, and inductively $L_{i+1} \coloneqq \frac{L_i}{2}$. ($K$ chosen so that $L_{K-1} \geqslant 1, L_K \leqslant 1$)
**Claim:** we can inductively construct $\mathscr{H}_0, \mathscr{F}_1, \mathscr{H}_1, \ldots, \mathscr{F}_k, \mathscr{H}_k \subseteq \mathscr{P}\mathscr{o}\mathscr{w}(X)$ s.t.
(1) $\mathscr{F}_1 \cup \ldots \cup \mathscr{F}_i \cup \mathscr{H}_i$ is cover of $\mathscr{H}$.
(2a) $\mathscr{H}_i$ is increasing
(2b) $\mathscr{H}_i$ is $L_i$-bounded
(2c) for all $A \in \mathscr{H}_i$, $A \cup Y_1 \cup \ldots \cup Y_i \in \mathscr{H}$
(3) $\mathbb{E}[c_p(\mathscr{F}_{i+1})] \leqslant 10^{-L_i}$.
*Proof of Claim:* well, start with $\mathscr{H}_0 \coloneqq \mathscr{H}$. Each step use lemma to "split" $\mathscr{H}_i$ into $\mathscr{F}_{i+1} \cup \mathscr{H}_{i+1}$, where $L$-boundedness halves each time. $\boxtimes$

*Penultimate remarks:* defining $\mathscr{F} \coloneqq \mathscr{F}_1 \cup \ldots \cup \mathscr{F}_K$, we can bound $\mathbb{E}_{Y_1,\ldots,Y_K}[c_p(\mathscr{F})] \leqslant \sum_{i=1}^{K-1} \mathbb{E}[c_p(\mathscr{F}_i)] \leqslant \sum_{i=1}^{K-1} 10^{-L_i} < \frac{1}{4}$.
Then by Markov, $\mathbb{P}_{Y_1,\ldots,Y_K}[c_p(\mathscr{F}) > \frac{1}{2}] < 0.5 \iff \mathbb{P}[c_p(\mathscr{F}) \leqslant \frac{1}{2}] \geqslant 0.5$. $\boxtimes$

What does it mean that $\mathscr{H}_K$ is $L_K$-bounded (where $L_K < 1$)? Two cases (exactly the 2 cases of the theorem!)

(I) $\mathscr{H}_K = \mathscr{H}_K(Y_1, \ldots, Y_K) = \varnothing$.

Using (1), this means that $\mathscr{F} := \mathscr{F}_1 \cup \ldots \cup \mathscr{F}_K$ is a cover of $\mathscr{H}$.

So if case(I) happens for "$> 50\%$ of $Y_1, \ldots, Y_K \sim X_{50p}$" (i.e. $\mathbb{P}_{Y_1,\ldots,Y_K \sim X_{50p}}[\text{case(I) happens}] > 0.5$), then (by the penultimate remark, and the fact that a probability-$[> 0.5]$ event and a probability-$[\geqslant 0.5]$ event must overlap) there is a positive probability that $Y_1, \ldots, Y_K \sim X_{50p}$ generates a cover $\mathscr{F}$ of $\mathscr{H}$ that is very "evasive", i.e. has small $p$-capture $c_p(\mathscr{F}) \leqslant \frac{1}{2}$.

(II) $\mathscr{H}_K = \mathscr{H}_K(Y_1, \ldots, Y_K) = \mathscr{P}\mathscr{o}\mathscr{w}(X)$ ($\mathscr{H}_K^{\downarrow} = \{\varnothing\}$).

Now using (2c), may take $A = \varnothing$ to get $Y := Y_1 \cup \ldots \cup Y_K \in \mathscr{H}$.

But then the r.v. $Y$ has distribution $Y \sim X_t$ for $t \leqslant K \cdot 50p \leqslant 100p \log L$.

So if case(II) happens with probability $\mathbb{P}_{Y_1,\ldots,Y_K \sim X_{50p}} \geqslant \frac{1}{2}$, then $\mathbb{P}_{Y \sim X_{100p \log L}}[Y \in \mathscr{H}] \geqslant \frac{1}{2}$, and we win!

One of the cases must happen more than 50% of the time (pigeonhole!), so we win either way! ■

## Appendix

General theory of thresholds.

- Setup: $X$ a set, $N := |X|$, $Y \sim X_p$ ($p$-random subset). Example: fix $n$-element set $V$. $X := \binom{V}{2} = E(K_n)$ so $N := \binom{n}{2}$. Then $Y \sim X_p$ identified with $G \sim \mathbb{G}(n,p)$.
- $\mathscr{H} \subseteq \mathscr{P}\mathscr{o}\mathscr{w}(X)$ an increasing family of sets. ($A \supseteq B \in \mathscr{H} \implies A \in \mathscr{H}$) Can of course talk about decreasing families, mutatis mutandis.
- define $f(p) := \mathbb{P}[Y \in \mathscr{H}]$ (function from $[0,1] \to [0,1]$). "How does the probability that a $p$-random set belong in $\mathscr{H}$, as $p$ changes?"
- We assume $\mathscr{H}$ is non-trivial, so $\mathscr{H} \neq \varnothing, \mathscr{P}\mathscr{o}\mathscr{w}(X)$. Equivalently (by increasingness), we want $X \in \mathscr{H}, \varnothing \notin \mathscr{H}$. Stupid technicality, but important to state out loud.

Simple observation: (1) $f(0) = 0$. (0-random subset have no elements). $f(1) = 1$, because $X \in \mathscr{H}$.

(2) $f$ is strictly increasing.

(3) $f$ is continuous!

Proof (2): 2 arguments. "Coupling" — $\theta \in [0,1]$ trick used in random permutation?. First, weakly increasing. $0 \leqslant p \leqslant q \leqslant 1$, want $f(p) \leqslant f(q)$. A priori, different distributions ("separate") of $p$-random subset vs. $q$-random subset. We couple these distributions to make the monotonicity obvious: assign to each $x \in X$ an indep. unif. real $\theta_x \in [0,1]$, and $Y_1 := \{x \in X : \theta_x \leqslant p\}, Y_2 := \{x \in X : \theta_x \leqslant q\}$ yields $Y_1 \sim X_p, Y_2 \sim X_q$, and ALWAYS $Y_1 \subseteq Y_2$. Trivial now containment of events $[Y_1 \in \mathscr{H}] \subseteq [Y_2 \in \mathscr{H}]$ so $f(p) := \mathbb{P}[Y_1 \in \mathscr{H}] \leqslant \mathbb{P}[Y_2 \in \mathscr{H}] := f(q)$.

2nd argument, "maybe less elegant". "Sprinkling" sample $Y_1 \sim X_p$. Sample (independently) $Z \sim X_t$ for $t =$???. Let $Y_2 := Y_1 \cup Z$. "$Z$ is the sprinkles you sprinkle on top of $Y_1$". Of course by def. $Y_1 \subseteq Y_2$. What is distr? Every element of $X$ is added to $Y_2$ indep. with probability $1 - (1-p)(1-t)$ (or $p + (1-p)t$ — prob. $p$ in $Y_1$, and $(1-p)t$ not in $Y_1$ but is in $Z$). Take $t = \frac{q-p}{1-p} \in [0,1]$ to get $X_2 \sim X_q$.

Strictly incr is Exc. Lol.

(3) $f$ is continuous. Indeed, it is a poly of degree $\leqslant N$ (this implies strictly incr. lol). Indeed, $f(p) = \sum_{A \in \mathscr{H}} \mathbb{P}[Y = A]$ (for $Y \sim X_p$), and $\mathbb{P}[Y = A] = p^{|A|}(1-p)^{N-|A|}$. "This is manifestly a polynomial".

"We have formula, so why are we being so vague? Well, the formula is not super helpful, especially these $(1-p)$ factors".

*Definition 0.6: General threshold*

The threshold for $\mathscr{H}$ is the unique $t_{\mathscr{H}} \in [0,1]$ s.t. $f(t_{\mathscr{H}}) = \frac{1}{2}$.
Exists by intermediate value theorem, unique by strict incr.
Remark: $\frac{1}{2}$ not essential, the next theorem tells us any fixed value $> 0, < 1$ yields the same conclusion.

*Theorem 0.7: Threshold actually displays threhold phenomenon (Bollobás-Thomason '87)*

For any $\epsilon > 0$, there is $C = C(\epsilon) > 0$ (indep. of $\mathscr{H}$!) s.t.

$$f(p) = \begin{cases} \leqslant \epsilon & \text{if } p \leqslant \frac{t_{\mathscr{H}}}{C} \\ \geqslant 1-\epsilon & \text{if } p \geqslant C t_{\mathscr{H}} \end{cases}$$

(In general, $t_{\mathscr{H}} \to \infty$ as $N = |X| \to \infty$.)

*Proof:* "sprinkling" (also known as "nibbling"). Let $C = C(\epsilon)$ be largish natural number (we will calculate later).

For brevity, denote $t := t_{\mathscr{H}}$. Take $p \geqslant Ct$ (the other case is an Exercise)

Independently generate $t$-random sets $Y_1, \ldots, Y_c \sim X_t$, let $Y := Y_1 \cup \ldots \cup Y_C$. "Sprinkle $C$ small sets to build big set"

Then $Y \sim X_q$ for some value $q \in [0,1]$ where we can compute $q$ explicitly $= 1-(1-t)^C$. But we can be crude, overcount (union bound) $\leqslant Ct \leqslant p$. For small threshold, this double counting is actually not bad, since everything is so sparse things are almost disjoint.

By increasing,

$$1 - f(p) \leqslant 1 - f(q) = \mathbb{P}[Y \notin \mathscr{H}] \leqslant \mathbb{P}[Y_i \notin \mathscr{H}, \forall i = 1..C] = \prod_{i=1}^{C} \underbrace{\mathbb{P}[Y_i \notin \mathscr{H}]}_{1-f(t)=\frac{1}{2}} = \left(\frac{1}{2}\right)^C \leqslant \epsilon$$

for $C \geqslant \log_2(\frac{1}{\epsilon})$. ■

How to compute $t_{\mathscr{H}}$? Computing hard, maybe ask "what drives $t_{\mathscr{H}}$ behavior".

1stMM: for $\mathscr{F} \subseteq \mathscr{Pow}(X)$, write $\langle \mathscr{F} \rangle$ increasing family generated by by $\mathscr{F}$, i.e. $\{A \subseteq X : A \supseteq S \text{ for some } S \in \mathscr{F}\}$.

Example: $\langle \mathscr{H} \rangle = \mathscr{H}$. But $\mathscr{H}$ not most efficient way of generating $\mathscr{H}$ using bracket; instead, $\mathscr{H}^{\downarrow} := \{\text{inclusion-minimal sets in } \mathscr{H}\}$

Examples: for $\mathbb{G}(n,p)$,

- $\{\text{graphs with triangle}\}^{\downarrow} = \{\text{triangles}\}$.
- $\{\text{graphs with Hamiltonian cycle}\}^{\downarrow} = \{\text{Hamiltonian cycles}\}$.
- $\{\text{connected graphs}\}^{\downarrow} = \{\text{spanning trees}\}$.

A cover of $\mathscr{H}$ is $\mathscr{F} \subseteq \mathscr{Pow}(X)$ s.t. $\mathscr{H} \subseteq \langle \mathscr{F} \rangle$.

*Example 0.8: Why cover?*

Example: for $\mathscr{H} \coloneqq \{\text{graphs with a copy of } [K_4 \text{ with edge sticking out}]\}$ (from a couple lectures ago), then $\mathscr{H}^{\downarrow} = \{\text{copies of } [K_4 \text{ with edge sticking out}]\}$ but this didn't control $t_{\mathscr{H}}$. Instead, the family of sets we need to look at to compute correct threshold was $\mathscr{F} = \{\text{copies of } K_4\}$, which is a cover of $\mathscr{H}$.

If $\mathscr{F}$ is a cover of $\mathscr{H}$, then $f(p) \coloneqq \mathbb{P}[Y \sim X_p \text{ is in } \mathscr{H}] \leqslant \mathbb{P}[Y \supseteq S \text{ for some } S \in \mathscr{F}]$, (1stMM) $\leqslant \mathbb{E}[\#$ of sets $S \in \mathscr{F}$ s.t. $S \subseteq Y]$. (Again, return to above Exmpl, that is how we computed correct threshold.)

$$= \sum_{S \in \mathscr{F}} p^{|S|}$$

Notice that looking at covers $\rightsquigarrow$ no $(1-p)$ terms!!! In ptclr all coeffs positive, bounds are easier.

Hence, if $\mathscr{H}$ has some cover $\mathscr{F}$ with $\sum_{S \in \mathscr{F}} p^{|S|} \leqslant \frac{1}{2}$, then $t_{\mathscr{H}} \geqslant p$.

*Definition 0.9: Expectation threshold*

The expecation threshold for $\mathscr{H}$ is "maximum lower bound you can attain from above method", $t^*_{\mathscr{H}} = \max p$ s.t. $\exists$ cover $\mathscr{F}$ of $\mathscr{H}$ with $\sum_{S \in \mathscr{F}} p^{|S|} \leqslant \frac{1}{2}$.
And "by definition" $t^*_{\mathscr{H}} \leqslant t_{\mathscr{H}}$.