Sharp Thresholds of Ramsey Triangle Properties

Anqi Li

Part III Essay 2024

Contents

1 Introduction

We begin with the following motivating question: for which $p = p(n)$ does $G(n, p)$ contain a triangle with probability $1 - o(1)$? By computing first moments, it is not difficult to deduce that if $np \to 0$ then $G(n, p)$ is triangle free with probability $1 - o(1)$. In this article, we will often be quite loose with asymptotic notation; for instance we might write the previous sentence as "when $p = o(n^{-1})$ then $G(n, p)$ is triangle-free with high probability".

Let us study this question from another angle. Let $\mathcal{A}$ be the set of all graphs which contain a triangle as a subgraph.[1] Then we are interested in $\mu_{\mathcal{A}}(p) := \mathbb{P}_{Z \in G(n,p)}[Z \in \mathcal{A}]$ as p changes. Clearly, when $\mu_{\mathcal{A}}(0) = 0$ and $\mu_{cA}(1) = 1$. As mentioned, first moment calculations shows that when $p = o(n^{-1})$, we have $\mu_{\mathcal{A}}(p) \ll 1$. What happens to $\mu_{\mathcal{A}}(p)$ as we increase p? In general, we show that $\mu_{\mathcal{A}}(p)$ is increasing in p when $\mathcal{A}$ is an monotone graph property, which we define next.

Definition 1.1. A set of graphs $\mathcal{A}$ is *monotone* if $G \in \mathcal{A}$ and $G \subset H$ implies that $H \in \mathcal{A}$.

Claim 1.2 (Monotonicity of $\mu_{\mathcal{A}}(p)$)**.** *Let $\mathcal{A}$ be a monotone graph property. Then $\mu_{\mathcal{A}}(p)$ (as defined earlier) is a strictly increasing function of p.*

Proof. Let $0 \leq p < q \leq 1$. Note that $B := G(n, q)$ has the same distribution as the union of two independent $A := G(n, p)$ and $A' := G(n, p')$ where p' is such that $1 - q = (1 - p)(1 - p')$. Therefore,

$$\mathbb{P}[A \in \mathcal{A}] < \mathbb{P}[A \cup A' \in \mathcal{A}] = \mathbb{P}[B \in \mathcal{A}]$$

where the inequality is strict because with positive probability, $A \notin \mathcal{A}$ while $A \cup A' \in \mathcal{A}$. $\square$

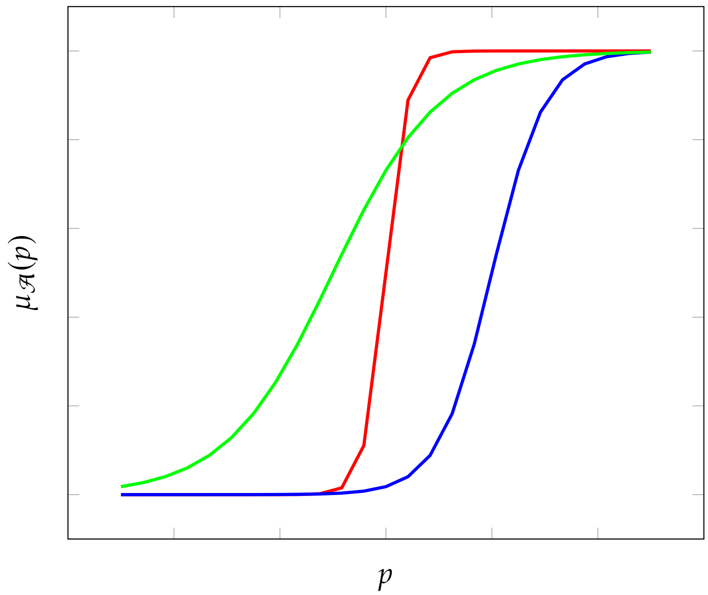
This intuitively seems to suggest that there should be a "phase transition" where for some $p_1 < p_2$ we have that $\mu_{\mathcal{A}}(p_1)$ is very close to 0 while $\mu_{\mathcal{A}}(p_2)$ is very close to 1. We make this precise in the following definition.

Definition 1.3. Let $\mathcal{A}$ be a graph property. We say that $p(n)$ is a *threshold* for $\mathcal{A}$ if

$$\mathbb{P}[G(n, q(n)) \in \mathcal{A}] \to \begin{cases} 0 & \text{if } \frac{q(n)}{p(n)} \to 0, \\ 1 & \text{if } \frac{q(n)}{p(n)} \to \infty. \end{cases}$$

It is a well-known theorem of Bollobás and Thomason [BT87] that every sequence of nontrivial monotone graph properties has a threshold. A way to see this is by plotting $\mu_{\mathcal{A}}(p)$ against p.

[1]Given a set of graphs $\mathcal{A}$ satisfying some property, we will oftentimes call $\mathcal{A}$ a *graph property*.



Instead of being concerned with exactly where the threshold behavior occurs, we instead endeavor to study the behavior of the probability plot. In particular, given any graph property $\mathcal{A}$ we want to be able to make a guess for which of the plots above (green, red, blue) is most likely to be its "probability-plot". Now, the key difference between the three plots above is how sharply the function transitions from 0 to 1. So our goal is to study what properties of $\mathcal{A}$ ensures that $\mu_{\mathcal{A}}$ turns from 0 to 1 in a small interval, which we will often call having a *sharp transition*. We make this notion precise in the following definition.

Definition 1.4. We say that a monotone graph property $\mathcal{A}$ has a *sharp threshold* at $p(n)$ if for every $\delta > 0$, we have

$$\mathbb{P}[G(n, q(n)) \in \mathcal{A}] \to \begin{cases} 0 & \text{if } \frac{q(n)}{p(n)} \leq 1 - \delta, \\ 1 & \text{if } \frac{q(n)}{p(n)} \geq 1 - \delta. \end{cases}$$

On the other hand, if there is some fixed $\varepsilon > 0$ and $0 < c < C$ such that $\mathbb{P}[G(n, q(n)) \in \mathcal{A}] \in (\varepsilon, 1 - \varepsilon)$ whenever $c \leq \frac{q(n)}{p(n)} \leq C$ then we say that $p(n)$ is a *coarse threshold*.

Another way to study this problem is to encode the graph as a Boolean string $\{0,1\}^{\binom{n}{2}}$: label the edges of the graph $1, \ldots, \binom{n}{2}$ and the ith bit of this Boolean string is 1 if the ith edge is present. Then $\mathcal{A}$ corresponds to a Boolean function $f\colon \{0,1\}^{\binom{n}{2}} \to \{0,1\}$. In the input string, the ith bit is 1 with probability p and the bits are all independent. In other words, we are endowing $\{0,1\}^{\binom{n}{2}}$ with the p-biased measure. It is therefore conceivable that p-biased Fourier analysis would be relatively useful. In the next subsection we give a quick overview of some concepts from Boolean Fourier analysis and refer the interested reader to [O'D14] for a more thorough overview.

1.1 Boolean Fourier analysis background

Henceforth we often freely interchance between the additive $\{0,1\}^n$ Boolean hypercube and the multiplicative $\{\pm 1\}^n$ version where to go from the former to the latter we consider the map $b \mapsto (-1)^b$.

We write $\sigma = \sqrt{p(1-p)}$. For each $i \in [n]$ we define $\chi_i \colon \{0,1\}^n \to \mathbb{R}$ by $\chi_i(x) = \frac{x_i - p}{\sigma}$. In particular, χ_i has mean 0 and variance 1).

It is well known that there is the following orthonormal Fourier basis $\{\chi_S\}_{S\subset[n]}$ of $L^2\left(\{0,1\}^n, \mu_p\right)$, where each $\chi_S := \prod_{i\in S}\chi_i$. Any $f : \{0,1\}^n \to \mathbb{R}$ has a unique expression $f = \sum_{S\subset[n]} \widehat{f}(S)\chi_S$ where $\{\hat{f}(S)\}_{S\subset[n]}$ are the *p-biased Fourier coefficients of* f. When $p = \frac{1}{2}$ this Fourier expansion has a rather nice interpretation: the Fourier expansion of $f \colon \{\pm1\}^n \to \mathbb{R}$ as

$$f(x) = \sum_{S\subset[n]} \widehat{f}(S)\chi_S(x) = \sum_{S\subset[n]} \widehat{f}(S)\prod_{i\in S} x_i$$

corresponds to the multilinear expansion of f.

The Plancherel's identity $\langle f, g\rangle = \sum_{S\subset[n]} \widehat{f}(S)\widehat{g}(S)$ is a consequence of the orthonormality of the Fourier characters. A special case of Plancherel's identity is Parseval's identity $\mathbb{E}[f^2] = \|f\|_2^2 = \langle f, f\rangle = \sum_{S\subset[n]} \widehat{f}(S)^2$. As shorthand we will also often write $\mu_p(f) = \mathbb{E}_{x\sim\mu_p^{\otimes n}}[f(x)]$. We also define the notion of restricting a function as follows: for a function $f \colon \{\pm1\}^n \to \mathbb{R}$, a set of coordinates $J \subset [n]$ and an assignment to them $z \in \{\pm1\}^J$. The restricted function $f_{J\to z} \colon \{\pm1\}^{\overline{J}} \to \mathbb{R}$ is defined by

$$f_{J\to z}(y) = f(x_J = z, x_{\overline{J}} = y).$$

We will often we studying the low-degree component of functions $f \colon \{0,1\}^n \to \mathbb{R}$, which corresponds to a suitable truncation of f which we define next. For $\mathcal{S} \subset \{0,1\}^n$ we define the $\mathcal{S}$-truncation $f^{\mathcal{S}} := \sum_{S\in\mathcal{S}} \widehat{f}(S)\chi_S$. When we talk about the low degree component of a function we are referring to truncating according to some degree threshold r, for which we write $f^{\le r} = \sum_{|S|\le r} \widehat{f}(S)\chi_S$.

For $i \in [n]$, the *i-derivative* $\partial_i f$ and *i-influence* $I_i(f)$ of f are defined as

$$\partial_i f = \sigma\left(f_{i\to 1} - f_{i\to 0}\right) = \sum_{S : i\in S} \widehat{f}(S)\,\chi_{S\setminus\{i\}}$$

and

$$I_{i,p}(f) = \|f_{i\to 1} - f_{i\to 0}\|_2^2 = \sigma^{-2}\,\mathbb{E}[f_i^2] = \tfrac{1}{p(1-p)} \sum_{S : i\in S} \widehat{f}(S)^2.$$

We define the (total) *influence* of f as

$$I_p(f) = \sum_i I_i(f) = (p(1-p))^{-1} \sum_S |S|\widehat{f}(S)^2. \tag{1}$$

where we sometimes suppress the subscript of p on I_p when it is clear what the parameter p is.

An important operator in the theory of Boolean function is the noise operator T_ρ which we next define. It will show up implicitly in the version of hypercontractive inequalities that we state (see Remark 2.11).

Definition 1.5 (Noise operator)**.** For $x \in \{0,1\}^n$, we define the ρ-correlated distribution $N_\rho(x)$ on $\{0,1\}^n$: a sample $y \sim N_\rho(x)$ is obtained by setting $y_i = x_i$ with probability ρ and otherwise with probability $1-\rho$ we resample y_i according to μ_p, executing this process independently for each coordinate i. We define the noise operator T_ρ on $L^2(\{0,1\}^n, \mu_p^{\otimes n})$ to be

$$T_\rho(f)(x) = \mathop{\mathbb{E}}_{Y \sim N_\rho(x)}[f(y)].$$

Each of the above identities can be checked easily; we omit their proofs and refer the reader to [O'D14].

1.2 Why p-biased Boolean Fourier analysis is relevant to thresholds: Russo-Margulis Lemma

As a taster for the power of Boolean Fourier analysis techniques, we show how understanding the total influence of a function gives us information about thresholds.

Theorem 1.6 (Russo-Margulis Lemma [Mar74, Rus82])**.** *Let $f\colon \{0,1\}^n \to \{0,1\}$ be a monotone function. Then*

$$\frac{d}{dp}\mu_p(f) = I_p[f].$$

Effectively, the Russo-Margulis Theorem states that if f has large total influence then it has a sharp threshold. I learnt of the following proof from the class I took with Minzer [Min21].

Proof. Take ε to be very small and we sample (x, y) in a coupled way so that marginally $x \sim \mu_p^{\otimes n}$, $y \sim \mu_{p+\varepsilon}^{\otimes n}$ and $x \le y$ always. (This can be done for instance by sampling $x \sim \mu_p^{\otimes n}$, and then for each i, if $x_i = 1$ we set $y_i = 1$, while if $x_i = 0$ then we take $y_i = 1$ with probability $\varepsilon/(1-p)$.)

Now, we have

$$\mu_{p+\varepsilon}(f) - \mu_p(f) = \mathop{\mathbb{E}}_{(x,y)}[f(x) - f(y)] = \mathop{\mathbb{E}}_{(x,y)}[(f(x) - f(y))\mathbf{1}_{x \neq y}].$$

Since the probability that x and y differ in more than a single coordinate is at most $n^2\varepsilon^2$, it follows that

$$\mu_{p+\varepsilon}(f) - \mu_p(f) - \sum_{i=1}^{n} \mathop{\mathbb{E}}_{(x,y)}[(f(x) - f(y))\mathbf{1}_{x,y \text{ differ only at } i}] \le n^2\varepsilon^2.$$

But we also have

$$\mathop{\mathbb{E}}_{(x,y)}[(f(x) - f(y))\mathbf{1}_{x,y \text{ differ only at } i}] = (\varepsilon - \mathbb{P}[x, y \text{ differ in } \ge 2 \text{ coordinates}])\, I_{i,p}[f].$$

Putting it all together, we have

$$\mu_{p+\varepsilon}(f) - \mu_p(f) - \varepsilon I_p[f] \le n^2\varepsilon^2 + n^3\varepsilon^2.$$

Now divide by ε and take $\varepsilon \to 0$ to get the desired conclusion. □

2 Friedgut-Bourgain sharp threshold theorem

The seminal work [Fri99] completely characterizes when a threshold of a graph property is sharp/coarse. Roughly speaking, the theorem states that all monotone graph properties with a coarse threshold can be approximated by a "local property". In this section, we aim to give an exposition of this result. In the upcoming sections, we will apply this result to prove the sharp threshold of several graph properties.

In the following subsections, we aim to give a full proof of the following theorem.

Theorem 2.1. *Let $\mathcal{A}$ be a graph property with a coarse threshold. Then there exists $p = p(n)$, $\varepsilon > 0$ and a fixed graph M with $\mathbb{P}[M \in G(n,p)] > \varepsilon$ such that with probability at least ε both of the following hold for $Z \sim G(n,p)$:*

$$\mathbb{P}[Z \cup G(n, \varepsilon p) \in \mathcal{A} | Z] \leq \frac{1}{2}$$

(here the randomness is over $G(n, \varepsilon p)$) and

$$\mathbb{P}_{\varphi}[Z \cup \varphi(M) \in \mathcal{A} | Z] \geq \varepsilon,$$

(here the randomness is over φ) where $\varphi(M)$ is a uniformly random copy of M in K_n.

Remark 2.2. A word on notation: $a_Z := \mathbb{P}[Z \cup G(n, \varepsilon p) \in \mathcal{A} | Z]$ and $b_Z := \mathbb{P}_{\varphi}[Z \cup \varphi(M) \in \mathcal{A} | Z] \geq \varepsilon$ is a random variable in Z. The statement is saying that $\mathbb{P}_{Z \sim G(n,p)}[a_Z \leq \frac{1}{2}$ and $b_Z \geq \varepsilon] \geq \varepsilon$.

An informal way to think about this theorem is that sprinkling constantly many edges "in a particular pattern" can boost the probability of being $\mathcal{A}$ much more than adding $\varepsilon p \binom{n}{2}$ random edges, and this pattern is given by the **M**agical graph M.

Remark 2.3. The version of the theorem we stated here is actually weaker than the version which Friedgut proves. In [Fri99], Friedgut shows that if $\mathcal{A}$ is a monotone graph property, let $0 < \alpha < 1$ and suppose p is such that $\alpha < \mathbb{P}_{Z \sim G(n,p)}[Z \in \mathcal{A}] < 1 - \alpha$, then for every $\varepsilon > 0$, we can find a fixed graph M such that $\mathbb{P}_{Z \sim G(n,p)}[Z \in \mathcal{A} | M \in Z] \geq 1 - \varepsilon$ where the notation means the property that $Z \in \mathcal{A}$ conditioned on the appearance of a specific copy of M. That is, Friedgut manages to find a graph M that gives a boost in probability of belonging in $\mathcal{A}$ till it becomes close to 1, while we only get a δ boost. Friedgut's proof relies rather heavily on the inherent graph symmetry. For all applications Theorem 2.1 suffices and so we contend ourselves to proving this weaker version.

As indicated in Section 1.1, we can study the threshold of a graph property using Boolean Fourier analysis. In order to prove Theorem 2.1, we will first prove a sharp threshold theorem that holds for any *monotone* Boolean function – this is (a strengthened version of) *Bourgain's sharp threshold theorem*.

Theorem 2.4 ((Strengthened) Bourgain's Sharp Threshold Theorem)**.** *For any monotone Boolean function $f\colon \{0,1\}^n \to \mathbb{R}$ such that $0 < \mu_p(f) =: \alpha < 1$ and $pI[f] \leq K\alpha(1-\alpha)$, there exists a set J of $O(K)$ coordinates such that $\mu_p(f_{J \to 1}) \geq \alpha + \exp(-O(K))$.*

Remark 2.5. Bourgain's original proof gives the bound $\mu_p(f_{J \to 1}) \geq \alpha + \exp(-O(K^2))$. This was strengthened to Theorem 2.4 in [KLLM24]; it is also shown in [KLLM24] that this bound is best possible.

We then show how to utilize the additional symmetries inherited by a Boolean function on graph properties (provided by automorphisms in the underlying graph) in order to boost the conclusions of Theorem 2.4 to obtain Theorem 2.1. We also note that the condition of $pI[f] \leq K\mu_p(f)(1 - \mu_p(f))$ is not as artificial as it looks: by the Russo-Margulis formula, if f was the indicator of an increasing graph property then it would automatically imply this condition.

2.1 Proof of Bourgain's sharp threshold theorem

In this subsection, we prove Theorem 2.4. There exists many expositions of Bourgain's sharp threshold theorem (see [Bal13, O'D14, Lac22], just to name a few). We will take a historically inaccurate and therefore slightly different approach as compared to these expositions. That said, our exposition is very much inspired by [O'D14].

Roadmap. The p-biased Fourier analysis is quite unintuitive, and so we will start off by studying the situation with a uniform distribution over the Boolean hypercube (i.e. $p = \frac{1}{2}$). By invoking what we have seen in Subsection 1.2, we show that in this context a sharp threshold result is implied by *Friedgut's junta theorem* [Fri98] (see Theorem 2.9). We then give a proof of Friedgut's junta theorem by using the *hypercontractivity inequality* (Bonami-Beckner inequality) [Bon70, Bec75, Gro75] (see Lemma 2.10). We next explore why these fail to generalize in the regime where $p \lesssim n^{-1}$. This will motivate the correct form of the hypercontractivity inequality (Theorem 2.16). Finally, we deduce Theorem 2.4 from the corresponding hypercontractivity inequality, in the same spirit that we derived Friedgut's junta theorem from the classical hypercontractivity inequality.

Remark 2.6. Historically the developments of the p-biased hypercontractivity inequality/global hypercontractivity inequality [KLLM24, Zha21]) came after [Fri99].

Remark 2.7. Another reason for presenting the proof in this slight roundabout fashion is that variants of Theorem 2.16 have found many applications in additive combinatorics [BKM23, EKL24], group theory/probability [KLS23, EKLM24], extremal combinatorics [KLLM21, Zak23], and various other fields of theoretical computer science [KM22].

2.1.1 Uniform distribution over $\mathbb{F}_2^n$.

As expounded upon in Section 1.1, the Fourier basis for the uniform distribution - which corresponds to the characters of the additive group - are more interpretable and the Fourier expansion corresponds to the multilinear expansion.

Let us recall the set-up: we have a monotone function $f: \mathbb{F}_2^n \to \{0, 1\}$. As we increase p from 0 to 1, we are interested in whether $\mathbb{P}_{x\sim\mu_p^n}[f(x) = 1]$ has a sharp threshold around the critical probability $p_c = \frac{1}{2}$. Henceforth we work with the multiplicative notion and consider $f: \{\pm 1\}^n \to \{\pm 1\}$. As we have seen in Sub-section 1.2, vaguely speaking a sharp transition occurs if and only if $I_{p_c}[f]$ is large (on the order of $\Omega(1)$) at the threshold probability p_c. This motivates us to understand the converse problem: what is the structure of Boolean functions with influence $O(1)$?

The most obvious examples that come to mind are the k-juntas for some $k = O(1)$:

Definition 2.8. A function $f: \{\pm 1\}^n \to \{\pm 1\}$ is a *k-junta* for $k \in \mathbb{N}$ if it depends on at most k coordinates. That is, $f(x) = g(x_{i_1}, \ldots, x_{i_k})$ for some $g: \{\pm 1\}^k \to \{\pm 1\}$ and $i_1, \ldots, i_k \in [n]$.

Are there any other such examples? It turns out that (in some sense) there are not, and this is the content of Friedgut's junta theorem.

Theorem 2.9. *Let $f: \{\pm 1\}^n \to \{\pm 1\}$ Then for every $\varepsilon > 0$, there exists a k-junta $g: \{\pm 1\}^k \to \{\pm 1\}$ such that $k \leq 2^{O\left(\frac{I(f)}{\varepsilon}\right)}$ and $\|f - g\|_2 \leq \varepsilon$.*

The most intuitive construction for g is to build it out of the most influential coordinates of f and then round appropriately to ± 1. This is exactly the approach that we take. It will become clear in the proof that we need to handle the Fourier coefficients $\widehat{f}(S)$ where $|S|$ is large differently from those when $|S|$ is small. When $|S|$ is small, the key idea is (roughly) the fact that "low-degree polynomials are smooth", captured in the following *hypercontractive inequality*.

Lemma 2.10 (Bonami-Beckner/Hypercontractive inequality). *For any $f: \{\pm 1\}^n \to \mathbb{R}$ and any $r \le n$, we have*

$$\left\|f^{\le r}\right\|_4 \le \sqrt{3}^r \left\|f^{\le r}\right\|_2 .$$

Remark 2.11. We will not prove this fact, but it turns out that the hypercontractive inequality is equivalent to *small-set expansion*, the latter is more geometric and is perhaps a more illuminative illustration of content of the theorem. Roughly speaking, we say that A has ρ-small set expansion if the probability of the following random walk starting in A landing outside A is high: the random walk where each bit is marked with probability $1-\rho$ and the removed bit is re-sampled uniformly. Precisely, suppose y is the next step of this random walk from x, then if $|A| = \alpha 2^n$, we say that A satisfies ρ-small set expansion if

$$\mathbb{P}[y \in A | x \in A] \le \alpha^{\frac{1-\rho}{1+\rho}}.$$

In the final stages, we will need the following simple consequence of Hölder's inequality.

Lemma 2.12. *Let $f: \{0,1\}^n \to \{-1,0,1\}$. Let $\mathcal{S} \subset \mathcal{P}([n])$ and let $g(x) = \sum_{S\in\mathcal{S}} \widehat{f}(S)\chi_S(x)$. Then*

$$\|g\|_2^2 \le \|g\|_4 \|f\|_2^{3/2} .$$

Here we let the image of f be in $\{-1,0,1\}$ because as we will see, we apply this lemma to $\partial_i f$ for a Boolean function f which has image not in $\{-1,1\}$ but rather $\{-1,0,1\}$.

Proof. By first applying Plancherel's and Hölder's inequality, we can write

$$\begin{aligned} \mathbb{E}[g^2] &= \sum_{S\in\mathcal{S}} \widehat{f}(S)^2 \\ &= \langle f, g\rangle \\ &\le \|f\|_{4/3} \|g\|_4 . \end{aligned}$$

But since the image of f lies in $\{-1,0,1\}$, it follows that $\|f\|_{4/3} = (\mathbb{E} f^2)^{3/4} = \|f\|_2^{3/2}$. Putting these together gives the desired inequality. □

Remark 2.13. We did not use any property of the norm defined over $\mu_{1/2}$. In particular, this means that Lemma 2.12 continues to hold in the p-biased setting as well.

By combining Lemma 2.12 and Lemma 2.10, we obtain the following version of the hypercontractivity inequality which we utilize in the sequel.

Corollary 2.14. *For any $f: \{\pm 1\}^n \to \{-1,0,1\}$ and any $r \le n$, we have*

$$\left\|f^{\le r}\right\|_2 \le \sqrt{3}^r \|f\|_2^{3/2} .$$

Proof. We write

$$\left\|f^{\le r}\right\|_2^2 \le \left\|f^{\le r}\right\|_4 \|f\|_2^{3/2} \le \sqrt{3}^r \left\|f^{\le r}\right\|_2 \|f\|_2^{3/2}.$$

Rearranging, we get

$$\left\|f^{\le r}\right\|_2 \le \sqrt{3}^r \|f\|_2^{3/2},$$

as desired. □

Proof of Theorem 2.9. Let $C > 0$ be an absolute constant to be determined, let $\delta = 2^{-C\frac{I(f)}{\varepsilon}}$ and define the set of influential coordinates

$$J = \{i : I_i[f] \ge \delta\}.$$

Let $G(x) = \sum_{S \subset J, |S| \le \frac{2I[f]}{\varepsilon}} \widehat{f}(S)\chi_S(x)$ and define $g(x) = \mathrm{sign}(G(x))$. It is by definition clear that $|J| \le 2^{(C+1)\frac{I[f]}{\varepsilon}}$ and $\|f - g\|_2 \le 2\|f - G\|_2$.

A vanilla way to bound $\|f - G\|_2$ is to write using Parseval's and the fact that $I_i[f] = \sum_{i \in S} \widehat{f}(S)^2 \le \delta$ for $i \notin J$:

$$\|f - G\|_2^2 = \sum_{S \not\subset J} \widehat{f}(S)^2 \le \sum_{i \notin J} \sum_{S \ni i} \widehat{f}(S)^2 \le n\delta$$

which is unfortunately too lossy for our purposes. Upon closer inspection, it is because we are doing way too much double counting in the first inequality, especially for Fourier coefficients $\widehat{f}(S)$ where $|S|$ is large. Let $M = \frac{2I[f]}{\varepsilon}$. Then, for the large Fourier coefficients we can write

$$\sum_{|S| \ge M} \widehat{f}(S)^2 \le \frac{1}{M} \sum_S |S| \widehat{f}(S)^2 = \frac{I[f]}{M}.$$

It remains to bound $\sum_{\substack{S \not\subset J \\ |S| \le M}} \widehat{f}(S)^2 \le \sum_{i \notin J} \sum_{\substack{S \ni i \\ |S| \le M}} \widehat{f}(S)^2$. To that end, we apply Lemma 2.14 to $\partial_i f^{\le M}$. This is because we have $\partial_i x_S = \begin{cases} x_{S \setminus \{i\}} & \text{if } i \in S, \\ 0 & \text{otherwise,} \end{cases}$ which combined with linearity furnishes the identity

$$I_i[f] = \sum_{S \ni i} \widehat{f}(S)^2 = \|\partial_i f\|_2^2$$

indicating that $\partial_i f$ are good functions to work with to pick our Fourier coefficients containing a particular element. That is, by Lemma 2.14 it follows that

$$\sum_{\substack{S \ni i \\ |S| \le M}} \widehat{f}(S)^2 = \left\|\partial_i f^{\le M}\right\|_2^2 \le \sqrt{3}^M \|\partial_i f\|_2^{3/2} = \sqrt{3}^M I_i[f]^{3/2},$$

which in turn implies that

$$\sum_{\substack{S \not\subset J \\ |S| \le M}} \widehat{f}(S)^2 \le \sqrt{3}^M \sum_{i \notin J} I_i[f]^{3/2} \le \sqrt{3}^M \sqrt{\delta} I[f].$$

Now, putting all these estimates together, we obtain

$$\|f - G\|_2^2 \le \frac{I[f]}{M} + \sqrt{3}^M \sqrt{\delta} I[f] \le \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

by taking $M = \frac{2I[f]}{\varepsilon}$. □

2.1.2 p-biased setting.

First, let us consider whether the above argument generalizes. Consider the following "dictatorship function" (this name arising because we can think of a Boolean function as effectively giving a scheme to aggregate votes in an election) $f: \{\pm 1\}^n \to \{\pm 1\}$ given by $f(x) = x_1$. Let us consider Corollary 2.14 for this choice of f with the biased μ_p-measure. The easiest way to see that we should not expect a statement like Corollary 2.14 to hold for the dictatorship is by using the equivalence with the geometric property of small-set expansion as stated in Remark 2.11. Indeed, the subset A of the Boolean hypercube corresponding to where f evaluates to 1 is small: it is at most p fraction of the entire cube. However, the ρ-noisy-process stays within A with probability ρ, instead of leaving the set A almost surely, with the latter what we would expect from a function satisfying a hypercontractivity inequality. Given that the hypercontractivity inequality is at the heart of our proof of Theorem 2.9, this suggests a real difficulty in generalizing the previous argument to the μ_p setting.

Upon closer examination, there is another more subtle reason why the analogy of Theorem 2.9 breaks: consider the function $f: \{\pm 1\}^n \to \{\pm 1\}$ that evaluates to 1 if and only if its input string x contains at least one coordinate i such that $x_i = 1$. It's not difficult to see that when $p = O(n^{-1})$, f is very "diffuse" in the sense that f is not 0.001-close to any junta on $o(n)$ coordinates.

More precisely, there is this sort of slightly counterintuitive dichotomy in the setting of μ_p with $p = O(n^{-1})$: we expect "global functions" (such as the `OR` example above) to have large total influence/satisfy a hypercontractive inequality of some sort; on the other hand, we would expect "local functions" (such as the dictator) to be not expanding/not satisfy a hypercontractive inequality/be "obstructions" of some form. And vaguely we can see the threshold theorem taking form – Russo-Margulis implies that graph properties with coarse thresholds should have corresponding indicator functions with small total influence, and our heuristic implies that this means the indicator function must be "local" and these "local" coordinates correspond to the Magic subgraph M.

Remark 2.15. We chose to present the proof following [Zha21] rather than [KLLM24] because upon closer inspection one can note that the proof directly generalizes to any product distribution; conversely, the proof in [KLLM24] is more specific to the p-biased set-up.

As in the case of $\mu_{1/2}$, the most important step is to establish a suitable "global hypercontractivity" inequality.

Theorem 2.16 (Global Hypercontractivity)**.** *[Zha21, Lemma 2.1.4] For any $f \in L^2(\{\pm 1\}^n, \mu_p^{\otimes n})$ and any $r \le n$, we have*

$$\left\|f^{\le r}\right\|_4^4 \le C^i \left\|f^{\le r}\right\|_2^2 \max_{\substack{J,y \\ |J|=r}} \left\|f_{J \to y_J}\right\|_2^2$$

for some absolute constant $C > 0$.

Here $f_{J\to y_J}$ encodes the heuristic that we should only expect such a hypercontractivity statement to hold in the case of f being a global function: if f was a "local" Boolean function (that "depends on the setting of few coordinates"), then we would expect the existence of some J and y_J such that $f_{J\to y_J}$ to be constant 1 and therefore the right hand side of Theorem 2.16 to be large.

To prove Theorem 2.16, the idea is to construct a "cousin" $\widetilde{f}$ of f through a trick known as *symmetrization* so that $\widetilde{f}$ can be thought of as living on the Boolean hypercube. This symmetrization trick works on any product space.

In the following, let $f = \sum_{S\subset[n]} f^{=S}$ be the Efron-Stein decomposition over this product space (see [O'D14, Chapter 8] for an extended exposition). Over the Boolean hypercube, this decomposition corresponds to the usual Fourier decomposition; that is, over the Boolean hypercube we have $f^{=S} = \widehat{f}(S)\chi_S$. In the general product space context, the noise operator T_ρ can be defined to be $T_\rho(f)(x) = \sum_{T\subset[n]}(1-\rho)^{n-|T|}\rho^{|T|}\,\mathbb{E}[f(z)|z_T = x_T]$ and it is an exercise to check that over the cube we recover Definition 1.5.

Definition 2.17. Let $f \in L^2(\Omega^n, \pi^n)$ be any function over a product space. The *symmetrization* of f, denoted by $\widetilde{f} \in L^2(\{-1,1\}^n \times \Omega^n, \mu_{1/2}^{\otimes n} \otimes \pi^{\otimes n})$ is given by

$$\widetilde{f}(r,x) = \sum_{S\subset[n]} r_S f^{=S}(x).$$

In this way, we can then apply the usual hypercontractivity inequality (Lemma 2.10) to $\widetilde{f}$, and then we show that the qth moments of f and $\widetilde{f}$ are roughly the same and so we can suitably "pullback" Lemma 2.10 to obtain a hypercontractivity inequality for f.

But first, let us work out an explicit example of symmetrization to gain some intuition for how it works.

Example 2.18. Let us study what happens to a Boolean function $f\colon \{\pm 1\}^n \to \mathbb{R}$ under the process of symmetrization:

$$f(x) = \sum_{S\subset[r]} \widehat{f}(x) \prod_{i\in S} x_i$$

has corresponding symmetrization given by

$$\widetilde{f}(r,x) = \sum_{S\subset[n]} \widehat{f}(S) \prod_{i\in S} r_i x_i = \sum_{S\subset[n]} \widehat{f}(S) x_S r_S.$$

Since x_i is a symmetric random variable, it follows that $x_i \sim r_i x_i$ so in fact $f \sim \widetilde{f}$.

Remark 2.19. More generally, if f has Fourier expansion consisting of only symmetric basis functions then $\widetilde{f} \sim f$ and so all the qth norms of f and $\widetilde{f}$ coincide.

Another way to think about $\widetilde{f}$ is that fixing a $y \in \Omega^n$, we have that $\widetilde{f}_{x\to y}$ is the Boolean function whose Fourier coefficient is $f^{=S}(x)$. A consequence of Parseval's identity that the second moments of f and $\widetilde{f}$ always coincide:

Claim 2.20. *Let $f \in L^2(\Omega^n, \pi^{\otimes n})$ and let $\widetilde{f}$ be its symmetrization. Then $\left\|\widetilde{f}\right\|_2 = \|f\|_2$.*

However, in general we would not expect $\widetilde{f}$ and f to agree on the qth moments, but it turns out that we can sandwich f between symmetrized applications of the noise operator T_ρ.

Theorem 2.21. *[Bou80] Let $f \in L^2(\Omega^n, \pi^{\otimes n})$ be any function over a product space and $q > 1$. Then*

$$\left\|\widetilde{T_{c_q} f}\right\|_q \leq \|f\|_q \leq \left\|\widetilde{T_2 f}\right\|_q$$

for some constant $0 \leq c_q \leq 1$ dependent only on q.

We only prove the upper bound in Theorem 2.21 which is all we need in the sequel, and refer the interested reader to [Bou80] for the proof of the lower bound. In order to gain some intuition for this statement, we define the "coordinate-by-coordinate" noise operator as follows, and show how this is related to the symmetrization operation.

Definition 2.22. For any $i \in [n]$ and $\rho \in \mathbb{R}$, we define

$$T_\rho^i(f) := \sum_{S \not\ni i} f^{=S} + \sum_{S \ni i} \rho f^{=S}.$$

This coordinate-wise noise operator gives us a nice reformulation of symmetrization. Indeed, for $r = (r_1, \ldots, r_n) \in \{\pm 1\}^n$, we have

$$T_r f(x) = T_{r_1}^1 \ldots T_{r_n}^n f(x) = \sum_{S \subset [n]} r_S f^{=S}(x) = \widetilde{f}(r, x).$$

This means that we should think of symmetrization as smoothing but with the vector $r \in \{\pm 1\}^n$. As a first step to proving the upper bound of Theorem 2.4, we first rewrite it as

$$\left\|T_{1/2} f(x)\right\|_q \leq \left\|\mathop{\mathbb{E}}_r T_r f(x)\right\|_q$$

which suggests that we should compare $T_{1/2}$ with T_r for $r \in \{\pm 1\}$. By thinking of the "smoothing" properties of the noise operator T_ρ, it is probably unsurprising that the following claim is true.

Claim 2.23. *For any $f \in L^2(\Omega, \pi)$, we have for r uniformly random over ± 1,*

$$\left\|T_{1/2} f\right\|_q \leq \|T_r f\|_q .$$

Proof. This hypercontractivity statement is really about random variables. Consider the decomposition $f(x) = f^{\emptyset}(x) + f^{=1}(x)$ so that $T_\alpha f(x) = f^{\emptyset}(x) + \alpha f^{=1}(x) =: Z_1 + \alpha Z_2$. Let Z_2' be an independent copy of Z_2. Then because $\mathbb{E}[Z_2] = \mathbb{E}[Z_2'] = 0$, we can write

$$\begin{aligned}
\left\|Z_1 + \tfrac{1}{2} Z_2\right\|_q &= \mathbb{E}[\left|Z_1 + \tfrac{1}{2} Z_2 - \tfrac{1}{2}\,\mathbb{E}[Z_2']\right|^q] \\
&= \mathbb{E}\left[\left|\mathbb{E}\left[Z_1 + \tfrac{1}{2} Z_2 - \tfrac{1}{2} Z_2'\right]\right|^q\right] \\
&\leq \mathbb{E}\left[(Z_1 + \tfrac{1}{2}(Z_2 - Z_2'))^q\right] \\
&\leq \left\|Z_1 + \tfrac{1}{2}(Z_2 - Z_2')\right\|_q
\end{aligned}$$

where the first inequality follows from Jensen's inequality. The upshot of doing this is that $Z_2 - Z_2'$ is now a symmetric random variable, and in particular we know that for a symmetric random variable x, we have $x \sim rx$ for $r \in \{-1, +1\}$. Consequently, we can write

$$\begin{aligned}
\left\|Z_1 + \tfrac{1}{2}Z_2\right\|_q &\le \left\|Z_1 + \tfrac{1}{2}(Z_2 - Z_2')\right\|_q \\
&\le \left\|Z_1 + \tfrac{r}{2}(Z_2 - Z_2')\right\|_q \\
&= \left\|(\tfrac{1}{2}Z_1 + \tfrac{r}{2}Z_2) + (\tfrac{1}{2}Z_1 - \tfrac{r}{2}Z_2')\right\|_q \\
&\le \left\|\tfrac{1}{2}Z_1 + \tfrac{r}{2}Z_2\right\|_q + \left\|\tfrac{1}{2}Z_1 - \tfrac{r}{2}Z_2'\right\|_q \\
&= \left\|\tfrac{1}{2}Z_1 + \tfrac{r}{2}Z_2\right\|_q + \left\|\tfrac{1}{2}Z_1 + \tfrac{r}{2}Z_2'\right\| \\
&= \|T_r f\|_q .
\end{aligned}$$

as desired, where the third inequality follows from the triangle inequality and the fourth equality follows from $r \sim -r$. □

With this result in hand, we can now prove the upper bound in Theorem 2.21 by induction.

Proof of Theorem 2.21. By induction it suffices to prove that

$$\left\|T^1_{1/2} f\right\|_q \le \left\|T^1_{r_i} f\right\|_q .$$

Let $x = (x_1, x')$. We fix x' and consider the restriction of the last $n-1$ coordinates to x'. The key property here is that the coordinate-wise noise operator commutes with restriction to that coordinate so we can just apply Claim 2.23. More precisely, we can write

$$\begin{aligned}
\left\|T^i_{1/2} f(x)\right\|_q &= \left\|\left\|(T^i_{1/2} f)\big|_{x'}(x_1)\right\|_{q,x_1}\right\|_{q,x'} \\
&= \left\|\left\|T_{1/2} f\big|_{x'}(x_1)\right\|_{q,x_1}\right\|_{q,x'} \\
&\le \left\|\left\|T_{r_1} f\big|_{x'}(x_1)\right\|_{q,x_1}\right\|_{q,x'} \\
&= \left\|T^i_{r_i} f(x)\right\|_q
\end{aligned}$$

as desired. □

With this preparation in place, we can prove the hypercontractivity inequality.

Proof of Theorem 2.16. Let $g = (T_2 f)^{\le r}$. By Jensen's inequality, we have

$$\mathop{\mathbb{E}}_x\left[\left(f^{\le r}\right)^4\right] \le \mathop{\mathbb{E}}_x\left[\mathop{\mathbb{E}}_r\left[\widetilde{g}\big|_x(r)^4\right]\right].$$

Now, by applying the usual hypercontractive inequality (Lemma 2.10) over the Boolean hypercube to g, and using the fact that $g^{=S}(x)$ are the Fourier coefficients of the Boolean function $\widetilde{g}(x,\cdot)$, it follows that

$$\begin{aligned}
\mathop{\mathbb{E}}_x\left[\mathop{\mathbb{E}}_r\left[\widetilde{g}\big|_x(r)^4\right]\right] &\le 2^{O(r)}\mathop{\mathbb{E}}_x\left[\mathop{\mathbb{E}}_r[\widetilde{g}(x,r)^2]^2\right] \\
&= 2^{O(r)}\mathop{\mathbb{E}}_x\left[\left(\sum_{|S|\le i} g^{=S}(x)^2\right)^2\right].
\end{aligned}$$

That is, we have

$$\begin{aligned}
\mathop{\mathbb{E}}_x\left[\left(f^{\le r}\right)^4\right] &\le 2^{O(r)}\mathop{\mathbb{E}}_x\left[\left(\sum_{|S|\le r} 2^{2|S|} f^{=S}(x)\right)^2\right] \\
&\le 2^{O(r)}\mathop{\mathbb{E}}_x\left[\left(\sum_{|S|\le r} f^{=S}(x)^2\right)^2\right].
\end{aligned}$$

Next, we split up this term and apply the Cauchy-Schwarz inequality:

$$\begin{aligned}
\mathop{\mathbb{E}}_x\left[\left(\sum_{|S|\le r} f^{=S}(x)^2\right)^2\right] &= \mathop{\mathbb{E}}_x\left[\sum_{|I|\le r}\sum_{\substack{S\supset I\\|S|\le r}} f^{=S}(x)^2\left(\sum_{T:|T|\le r, S\cap T=I} f^{=T}(x)^2\right)\right] \\
&\le \sum_{|I|\le r}\mathop{\mathbb{E}}_{x_I}\left[\left(\sum_{S\supset I:|S|\le r}\mathop{\mathbb{E}}_{x_{S\setminus I}}[f^{=S}(x_S)^2]\right)\left(\sum_{T\supset I:|T|\le i}\mathop{\mathbb{E}}_{x_{T\setminus I}}[f^{=T}(x_T)^2]\right)\right] \\
&= \left(\sum_{|I|\le r}\sum_{S\supset I:|S\ \le r}\mathop{\mathbb{E}}_{x_S}\left[f^{=S}(x_S)^2\right]\right)\max_{\substack{|I|\le r\\ y_I\in\{\pm1\}^I}}\left(\sum_{T\supset I:|T|\le r}\mathop{\mathbb{E}}_{x_{T\setminus I}}[f^{=T}(y_I,x_{T\setminus I})^2]\right) \\
&\le 2^r\left\|f^{\le r}\right\|_2^2\max_{|I|\le r, y_I\in\{\pm1\}^I}\left(\sum_{T\supset I:|T|\le r}\mathop{\mathbb{E}}_{x_{T\setminus I}}\left[f^{=T}(y_I,x_{T\setminus I})^2\right]\right)
\end{aligned}$$

where for the last inequality we note that each $f^{=S}(\cdot)$ term in the first summation appears exactly $2^{|S|}\le 2^r$ times. For the other term, we can do a principle of inclusion-exclusion type counting to relate it to restrictions:

$$\begin{aligned}
\sum_{T\supset I:|T|\le r}\mathop{\mathbb{E}}_{x_{T\setminus I}}[f^{=T}(y_I,x_{T\setminus I})^2] &\le \sum_{T\supset I}\mathop{\mathbb{E}}_{x_{T\setminus I}}[f^{=T}(y_I,x_{T\setminus I})^2] \\
&= \sum_{T\supset I}\mathop{\mathbb{E}}_{x_{T\setminus I}}\left[\left(\sum_{J\subset I}(-1)^{|I|-|J|}(f_{J\to y_J})^{=T\setminus I}(x_{T\setminus I})\right)^2\right] \\
&\le 2^r\sum_{T\supset I}\sum_{J\subset I}\mathop{\mathbb{E}}_{x_{I\setminus J}}\left[\left(f_{J\to y_J}\right)^{=T\setminus I}(x_{T\setminus I})^2\right] \\
&= 2^r\max_{J\subset I}\left\|f_{J\to y_J}\right\|_2^2
\end{aligned}$$

where we applied the Cauchy-Schwarz inequality to obtain the final inequality. Putting everything together gives us the desired conclusion. □

Finally, we deduce Theorem 2.4 from Theorem 2.16. This part of the argument is quite similar to the argument in Theorem 2.9, where we used the Hölder inequality trick in Lemma 2.12. We first package this part of the argument into the following useful corollary.

Corollary 2.24. *Let* $f \in L^2(\{\pm 1\}^n, \mu_p^{\otimes n})$ *be a* monotonic function *and* $r \leq n$. *Suppose* $\mu_p(f_{J\to 1}) \leq \mu_p(f) + \delta$ *for all* $|J| \leq r$ *and* $\mu_p(f) < \delta$. *Then*

$$\mathbb{E}[(f^{\leq r})^2] \leq C^r \delta^{1/3} \mu_p(f)$$

for some absolute constant $C > 0$.

Proof. Let $g_i := f_{i\to 1} - f_{i\to 0}$, then for $S' \subset [n]$ with $|S'| \leq r-1$, we have

$$\begin{aligned}\mu_p((g_i)_{S'\to 1}) &= \mu_p\left(f_{S'\cup\{i\}\to 1} - f_{s\to 1,\{i\}\to 0}\right)\\ &\leq \mu_p(f) + \delta - \mu_p(f_{i\to 0})\\ &\leq \mu_p(f_{i\to 1}) + \delta - \mu_p(f_{i\to 0})\\ &= \mu_p(g_i) + \delta,\end{aligned}$$

where we used in both inequality the monotonicity of f. Now, by Theorem 2.16, it follows that for some $\widetilde{C} > 1$, we have

$$\left\|g^{\leq r}\right\|_4 \leq \widetilde{C}^r \left\|g^{\leq r}\right\|_2^{\frac{1}{2}} \left(\|g\|_2^2 + \delta\right)^{1/2}. \tag{2}$$

Combining (2) with Lemma 2.12 (which we can apply in this p-biased setting because of Remark 2.13, it follows that

$$\begin{aligned}\left\|g^{\leq r}\right\|_2^2 &\leq \left\|g^{\leq r}\right\|_4 \|g\|_2^{3/2}\\ &\leq \widetilde{C}^r \left\|g^{\leq r}\right\|_2^{1/2} \|g\|_2^{3/2} \left(\|g\|_2^2 + \delta\right)^{1/2}\end{aligned}$$

which rearranges to $\left\|g^{\leq r}\right\|_2 \leq C^r \|g\|_2 \delta^{1/3}$. □

Proof of Theorem 2.4. Let β be such that $C^{2K} \exp(-\beta K)^{1/3} = \frac{1}{2}$ for the value of the constant C from Corollary 2.24. Suppose for the sake of contradiction that $\mu_p(f_{J\to 1}) \leq \mu_p(f) + \exp(-\beta K)$ for all sets J of size $2K$. By Corollary 2.24, it follows that

$$\mathbb{E}[(f^{\leq 2K})^2] \leq C^{2K} \exp(-\beta K)^{1/3} \mu_p(f) = \frac{\mu_p(f)}{2}$$

which implies that $\mathbb{E}[(f^{\geq 2K})^2] = \|f\|_2^2 - \mathbb{E}[(f^{\leq 2K})^2] \geq \mu_p(f)/2$. But by definition of the p-biased influence (see (1)), it follows that $p(1-p)I[f] \geq 2K\,\mathbb{E}[(f^{\leq 2K})^2]$ and so $p(1-p)I[f] > K\mu_p(f)$ which gives the desired contradiction. □

2.1.3 Boosting Theorem 2.4 in the graph setting

Theorem 2.25 (Another variant of Friedgut's sharp threshold theorem)**.** *Let $\mathcal{A}$ be a monotone graph property with a coarse threshold. Then for all positive α and C, there exists $\delta, \varepsilon, p_0, K > 0$ such that if $0 < p < p_0$ for some $\alpha \le \mathbb{P}_{Z\sim G(n,p)}[Z \in \mathcal{A}] \le 1 - \alpha$ then there is a graph M on K vertices such that $\mathbb{P}_{Z\sim G(n,p)}[M \in Z] > \varepsilon$ and*

$$\mathbb{P}_{Z\sim G(n,p)}[Z \cup M \in \mathcal{A}] > \mathbb{P}_{Z\sim G(n,p)}[Z \in \mathcal{A}] + \delta.$$

Proof. We begin by encoding the graph property $\mathcal{A}$ as a Boolean function $f\colon \{\pm 1\}^{\binom{n}{2}} \to \{0,1\}$ the natural way: we encode the edge set of the graph the obvious way and then set $f = \mathbf{1}_{\mathcal{A}}$ to be the indicator function of the property. The main observation is that for a $S \subset \binom{[n]}{2}$ of size K, we have $|\widehat{f}(S)| \le (4p(1-p))^{|S|/2}$ and summing S over its orbit Θ (of homomorphic images) it follows that since $\mathbb{P}_{Z\sim G(n,p)}[S \in Z] = |\Theta(S)| p^{|S|}(1-p)^{\binom{n}{2}-|S|}$, we have

$$\sum_{S\in\Theta} \widehat{f}(S)^2 \le O(\mathbb{P}_{Z\sim G(n,p)}[S \in Z])$$

In other words, if $\mathbb{P}_{Z\sim G(n,p)}[S \in Z]$ is small, the orbit of the graph contributes very little to the 2-norm of f. We can therefore proceed as before: first, we use the Russo-Margulis Theorem (Theorem 1.6) to obtain $K > 0$ such that $pI[f] \le K\mu_p(f)(1-\mu_p(f))$. Let β be such that $C^{2K}\exp(-\beta K)^{1/3} = \frac{1}{2}$ for the value of the constant C from Corollary 2.24. Let

Suppose for the sake of contradiction that $\mu_p(f_{J\to 1}) \le \mu_p(f) + \exp(-\beta K)$ for all sets J of size $2K$ such that $\mathbb{P}_{Z\sim G(n,p)}[S \in Z] \ge \exp(-\beta K)$. Note that for all $I \subset \binom{[n]}{2}$ such that $|I| \le 2k$ we have $\|f_{I\to 1}\|_2^2 \le \exp(-\beta K)$ and so the same proof of Corollary 2.24 continues to work. It therefore follows that

$$\mathop{\mathbb{E}}_{\mu_p}[(f^{\le 2K})^2] \le C^{2K}\exp(-\beta K)^{1/3}\mu_p(f) = \frac{\mu_p(f)}{2}$$

which implies that $\mathbb{E}[(f^{\ge 2K})^2] = \|f\|_2^2 - \mathbb{E}[(f^{\le 2K})^2] \ge \mu_p(f)/2$.

But by definition of the p-biased influence, it follows that $p(1-p)I[f] \ge 2K\,\mathbb{E}[(f^{\le 2K})^2]$ and so $p(1-p)I[f] > K\mu_p(f)$ which gives the desired contradiction. □

It is immediate to deduce Theorem 2.1 from Theorem 2.25.

3 Baby example of sharp threshold: 2-colorability of hypergraphs

In this section, we demonstrate an application of Theorem 2.1 to prove the sharp threshold of 2-(vertex)-coloring of k-uniform hypergraphs for $k > 2$. This is an important lemma in results about the threshold for k-sat problems [AM02], the latter being itself a fundamental question in probability. The reason we study hypergraphs rather than graphs is that 2-colorability of graphs has a coarse threshold, since 2-colorability corresponds to being bipartite and is witnessed by the local property of absence of odd cycles.

Theorem 3.1. *Let $k > 2$ be fixed and let $H(n,p)$ be the random k-uniform hypergraph on n vertices. Then the threshold for $H(n,p)$ being non-2-colorable is sharp.*

Suppose for the sake of contradiction that 2-colorability of k-uniform hypergraph is coarse. Then the statement of the threshold theorem (Theorem 2.1) basically provides the existence of a constant size "Magical" graph M such that its existence boosts the probability of being non-2-colorability by a lot. On first thought this may seem plausible; for instance the presence of the graph formed by all possible k-edges among $(2k-1)$ nodes (i.e. the complete k-uniform hypergraph $\mathcal{H}_k$ on $(2k-1)$ nodes) would result in non-2-colorability.

The point is that the threshold probability is $p \asymp n^{-(k-1)}$ and at this p it is very unlikely for $H(n,p)$ to contain such $\mathcal{H}_k$. Indeed, the expected number of such $\mathcal{H}_k$ in $H(n,p)$ is $\asymp n^{(2k-1)-(k-1)\binom{2k-1}{k-1}} = o(1)$. And therein lies what I find remarkable about Theorem 2.1: it is able to capture the fact that these kinds of obstructions are in some sense "bogus" and only see the real obstructions in the form of the magical graphs M.

Now, to prove Theorem 3.1, we begin by formalizing the above heuristic regarding obstructions.

Claim 3.2. *If $\mathbb{P}[M \subset H(n,p)]$ is bounded away from 0, then M is 2-colorable.*

Proof. By definition, for every subhypergraph $M' \subset M$, the expected number of copies of M' in $H(n,p)$ is $\Omega(1)$. Since the threshold probability is $\asymp n^{-(k-1)}$, we have that

$$\mathbb{E}[\# \text{ of } M' \text{ in } H(n,p)] = n^{|V(M')|-(k-1)|E(M')|}.$$

That is, every set of r edges in M meets at least $(k-1)r$ vertices. By Hall's marriange theorem, it follows that we can find a mapping g from $E(H)$ to r-tuples of distinct vertices such that for all $v \in V(H)$ there exists a unique edge e with $v \in g(e)$. Then the 2-coloring of M can be formed by assigning each edge a unique pair of vertices and then assigning them different colors. ◻

Next, we apply Theorem 2.1 to the (hyper)graph property $\mathcal{A}$ of being non-2-colorable. We restrict to the positive fraction of instances $Z \sim H(n,p)$ for which the two properties in the theorem holds. Then we know that for an ε fraction of possible $\varphi(M)$, adding $Z \cup \varphi(M)$ is not 2-colorable.

However, Claim 3.2 shows that $M = \{v_1, \ldots, v_t\}$ is itself 2-colorable; fix such a coloring $\psi \colon [t] \to [2]$. For some t-tuple of vertices $u_1, \ldots, u_t$, we say that $\{u_1, \ldots, u_t\}$ is an *obstruction* to Z if the coloring $\widetilde{\psi}$ specified by $\widetilde{\psi}(u_i) = \psi(i)$ cannot be extended to give a valid coloring of of Z; that is, if there does not exist some coloring $\varphi \colon V(Z) \to \{0,1\}$ such that the restriction to $\{u_1, \ldots, u_t\}$ agrees with $\widetilde{\psi}$ i.e. $\varphi|_{\{u_1,\ldots,u_t\}} = \widetilde{\psi}$. Using this language, we can rephrase the observation in the previous paragraph as effectively saying that an ε fraction of t-tuples $\{u_1, \ldots, u_t\} \in \binom{V}{t}$ are obstructions to Z.

We use a supersaturated version of the hypergraph Kővari-Sós-Turán theorem to extract a subset of vertices which form many t-tuples of obstructions to Z.

Theorem 3.3 ([ES83]). *For every positive integers k and t and $0 < \gamma \leq 1$, there exists $0 < \gamma'$ such that for sufficiently large n, if H is a t-regular hypergraph on vertex set $[n]$, then there exists $\gamma' n^{kt}$ copies of the complete t-partite graph with k vertices in each part. (i.e. γ' fraction of sampling t possible k-tuples $(v_1^1, \ldots, v_1^k), \ldots, (v_t^1, \ldots, v_t^k)$ have the property that for any $f \colon [t] \to [k]$ we have $(v_1^{f(1)}, \ldots, v_t^{f(t)}) \in E(H)$)*

By applying Theorem 3.3, to our $\varepsilon\binom{n}{t}$ t-tuple obstructions to Z, we can obtain complete t-partite graphs G where the edges in G correspond to t-tuples of vertices that form obstructions to Z.

Claim 3.4. *If we add t hyperedges to Z, such that the corresponding k-tuples form a t-partite graph of obstructions, then the resulting hypergraph is non-2- colorable.*

This would give the desired contradiction because combining Claim 3.4 with Theorem 3.3 implies that for any random t hyperedges, with probability ε' the resulting hypergraph is not 2-colorable. Consequently, adding $\omega(n)$ hyperedges to Z would almost surely make it not 2-colorable, which would then contradict the first condition of Theorem 2.1. We finish by proving Claim 3.4.

Proof of Claim 3.4. Let $e_1, \dots e_t$ be the t hyperedges furnished by the claim, and suppose for the sake of contradiction that there is a proper 2-coloring ψ of $Z \cup \{e_1, \dots, e_t\}$. In other words, for each $i \in [t]$ there is some $v_i \in e_i$ such that $\psi(e_i) = \sigma(i)$. But $(v_1, \dots, v_t)$ was assumed to be bad but it also agrees in color with σ, and so we should not have been able to extend it to get ψ in the first place. □

To summarize, the proof proceeded in two steps:

- First, we extract a Magical graph M alá Theorem 2.1 and then show that $M \notin \mathcal{A}$ because $\mathbb{P}[M \in G(n, p)]$ is bounded away from zero.
- Next, we use this property of M to find gadgets in the graph Z upon whose inclusion would force Z to be in $\mathcal{A}$, and then we show that these gadgets are well-dstributed so that $G(n, \varepsilon p)$ almost surely would not be able to avoid them. This would then imply that almost surely we have $Z \cup G(n, \varepsilon p) \in \mathcal{A}$, which would then contradict the first property in Theorem 2.1. This step is typically the involved step.

For further illustrations of using Theorem 2.1 to prove that various properties have sharp thresholds, we refer the reader to the excellent survey [Fri05] which covers this example of hypergraph 2-colorability and more.

4 (Brief) Historical interlude

In the previous section, we discussed the sharp threshold of some vertex colorings. In the remainder of this article, we will only be concerned with *edge colorings* and in particular Ramsey properties of the following flavor:

Definition 4.1. Given graph G and H and an integer $r > 2$, we write $G \longrightarrow (H)_r$ if every r-coloring of the edges of G contains a monochromatic copy of H.

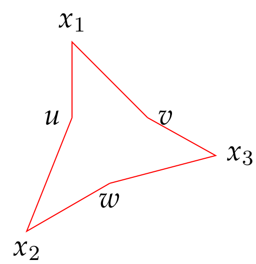
The goal of this section is to briefly discuss the results in [FRRT06] which show that $G(n, p) \longrightarrow (K_3)_2$ has a sharp threshold. There is also a long history of sharp thresholds of Ramsey properties. For space reasons we will omit this and direct interested readers to consult the introductions of [FRRT06, FKSS22]. First, we state the key result in [FRRT06].

Theorem 4.2 ([FRRT06]). *There exist positive constants c_0 and c_1 and a function $c(n)$ satisfying $c_0 \leq c(n) \leq c_1$ such that, for every positive ε,*

$$\lim_{n\to\infty} \mathbb{P}[G(n, p) \longrightarrow (K_3)_2] = \begin{cases} 1 & \textit{if } p \geq (1+\varepsilon)c(n) \cdot n^{-\frac{1}{2}}, \\ 0 & \textit{if } p \leq (1-\varepsilon)c(n) \cdot n^{-\frac{1}{2}}. \end{cases}$$

Let $\mathcal{A}$ denote the property of containing a monochromatic K_3 in any two coloring of the edges. As sketched in the previous section, we need to first show that the M given by Theorem 2.1 does not lie in $\mathcal{A}$; i.e. there is a two coloring of M avoiding triangles. We do this in detail in the following section, so we omit this deduction for now.

We will put aside M for now, and discuss what kind of gadgets we would like to find so that upon adding it to an instance $Z \in G(n, p)$ we would land in $\mathcal{A}$. The most natural idea is to find a gadget of the following form:



where $ux_1, vx_1, vx_2, wx_3, wx_2, ux_2 \in Z$. This would then force uvw to be a monochromatic blue triangle. The issue is that we would not expect our graph Z to be dense enough to contain many gadgets of the above form, and so in particular it is highly likely that $G(n, \varepsilon p)$ would just avoid all of these gadgets entirely.

At a high level, this is where the existence of M comes into play. The fact that $\mathbb{P}[Z \cup \varphi(M)|Z] \geq \varepsilon$ implies that there are many sets of vertices in G such that planting a homomorphic copy of M on them "destroys" all triangle-free colorings (we call these sets of vertices *bad*), so the set of triangle-free colorings of Z are very constrained. In some sense, we can use these *bad* sets to encode triangle-free colorings, and in [FRRT06] a Szemérédi regularity style argument is utilized to reveal some underlying structure in these *bad* sets. We will expand upon the idea of triangle-free colorings of Z being constrained and show how to use the *hypergraph containers method* to cluster colorings together in the following section.

5 Sharp threshold for $G(n, p) \to (K_3)_3$, following [FKSS22]

In this section, we provide an exposition of [FKSS22] in the specific instance of the three color Ramsey problem for triangles, which in some sense is the "smallest" example that improves upon Section 4. The most general statement that can be proven with the methods developed in [FKSS22] is the following:

Definition 5.1. A graph H is *collapsible* if for every edge e of H and every endpoint a of e, there is an edge f of H and a graph homomorphism (i.e. mapping of vertices that preserves edge relations) from $H \setminus f$ to $H \setminus e$ that maps both endpoints of f to a.

Definition 5.2. The *2-density* of a graph H is defined to be

$$m_2(H) := \max \left\{ \max_{\emptyset \neq F \subset H} \frac{|E(F)| - 1}{|V(F)| - 2}, \frac{1}{2} \right\}.$$

Definition 5.3. A graph H is *strictly 2-balanced* if $m_2(F) < m_2(H)$ for every subgraph $F \subsetneq H$.

We refer the reader to the introduction of [FKSS22] for a discussion for why these definitions are in some sense natural.

Theorem 5.4. *[FKSS22, Theorem 1.2] Suppose that H is a strictly 2-balanced, collapsible graph that is not a forest and $r > 2$ is an integer. There exist positive constants c_0 and c_1 and a function $c(n)$ satisfying $c_0 \le c(n) \le c_1$ such that, for every positive ε,*

$$\lim_{n\to\infty} \mathbb{P}[G(n,p) \to (H)_r] = \begin{cases} 1 & \text{if } p \ge (1+\varepsilon)c(n)\cdot n^{-1/m_2(H)}, \\ 0 & \text{if } p \le (1-\varepsilon)c(n)\cdot n^{-1/m_2(H)}. \end{cases}$$

At the end of this section, we outline how one might conceivably generalize what is presented in this section to the general case. We leave the interested reader to consult [FKSS22] for the full proof. Henceforth we restrict to the setting of $H = K_3$ and $r = 3$.

The proof goes by way of contradiction using Friedgut's sharp threshold theorem, as illustrated in Section 3. For the reader's convenience, we recall the variant of Friedgut's sharp threshold theorem that we will be using.

Theorem 5.5 (Variant of Friedgut's sharp threshold Theorem)**.** *Let $\mathcal{A}$ be a graph property with a coarse threshold. Then there exists $p = p(n)$, $\varepsilon > 0$ and a fixed graph M with $\mathbb{P}[M \in G(n,p)] > \varepsilon$ such that with probability at least ε the following hold for $Z \sim G(n,p)$:*

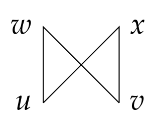
$$\mathbb{P}[Z \cup G(n,\varepsilon p) \in \mathcal{A} | Z] \le \frac{1}{2}$$

(here the randomness is over $G(n,\varepsilon p)$) and

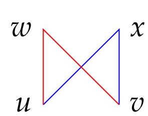
$$\mathbb{P}_\varphi[Z \cup \varphi(M) \in \mathcal{A}|Z] \ge \varepsilon,$$

(here the randomness is over φ) where $\varphi(M)$ is a uniformly random copy of M in K_n.

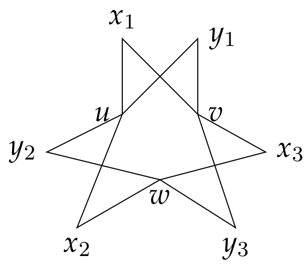
Let $Z \sim G(n,p)$ be a typical instance of $G(n,p)$. If we want to prove the existence of a monochromatic triangle in this 3-color Ramsey problem, we are motivated to consider the existence of the following gadgets that serve as constraints on our edge colorings. Following [FKSS22], we call the configuration of edges



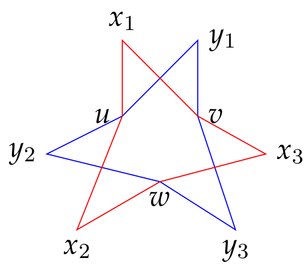
a *star* supported on uv and we call



a *rainbow star* supported on uv. We also say that uv is *forced* to a color, which in this example would be green. The most natural way to show that a monochromatic triangle is "forced" is to try to find configurations of the following form obtained by "gluing together" three stars. We call the following configuration of edges a *constellation* supported on uvw.



And we call



a *rainbow constellation* supported on uvw. In this example, if uvw were also edges in Z then we would obtain a contradiction as uvw would be forced to a green triangle. The main conclusion from this discussion is the following observation:

> **Observation.** Let $Z \sim G(n,p)$ and let C be a given proper 3-coloring of Z. Let uvw be a triangle forced to some color by C. Then C cannot be extended to give a proper 3-coloring of $Z \cup \{u,v,w\}$.

First attempt. Suppose for the sake of contradiction that $G(n,p) \to (K_3)_3$ has a coarse threshold. Let $Z \sim G(n,p)$. Then by Theorem 5.5, it follows that there exists a Magical graph M with property of "boosting the probability of $\to (K_3)_3$". In the following outline, we use the symbol (💡) to denote key observations or ideas.

Roughly speaking, we proceed in four steps:

1. We show that for every proper coloring of Z, there will be many edges (precisely, $\Omega(n^2)$ edges) forced to some color.
2. We show that since $\Omega(n^2)$ edges of Z are forced to some color, then actually $\Omega(n^3)$ triangles of Z are also forced to some color. This step is far from obvious, and we will discuss the subtleties in this step in due course.
3. We show that it is very unlikely (with exponentially small probability) for $G(n, \varepsilon p)$ to avoid all these forced triangles from a given proper coloring of Z.
4. We combine the probability estimate in the previous step with a union bound over all proper colorings of Z.

Now, we elaborate on each of the above steps in more detail.

1. Suppose for the sake of contradiction that $o(n^2)$ edges are forced to some color. Recall that by Theorem 5.5, the Magical graph M we obtain has $O(1)$ edges. This implies

$$\mathbb{P}_{\varphi}[\varphi(M) \text{ contains a forced edge}] \leq |E(M)| \cdot o(1) = o(1).$$

Furthermore, we claim that

$$\mathbb{P}_{\varphi}[Z \cap \varphi(M) \neq \emptyset \text{ or } \exists \text{ triangle } T \subset Z \cup \varphi(M) \text{ such that } |E(T \cap \varphi(M))| = 2] = o(1). \tag{3}$$

To prove this, it suffices to note that since M has $O(1)$ edges, we have

$$\mathbb{P}_{\varphi}[Z \cap \varphi(M) \neq \emptyset] \leq |E(M)| \cdot p = o(1).$$

Furthermore, if there is some triangle T such that $|E(T \cap \varphi(M))| = 2$, then note that the final edge of T is specified by the two edges lying in the copy of M, that is:

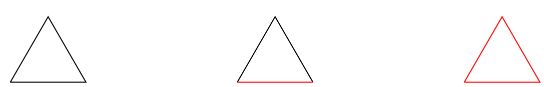
$$\mathbb{P}_{\varphi}[\exists \text{ triangle } T \subset Z \cup \varphi(M) \text{ such that } |E(T \cap \varphi(M))| = 2] = p = o(1).$$

Consequently, since (b) in Theorem 5.5 gives a positive proportion of φ providing a copy $\varphi(M)$ of M satisfying the "boosting" conditions, this implies that we can find a $\varphi(M)$ such that:

(i) $Z \cup \varphi(M) \longrightarrow (K_3)_3$,

(ii) none of $e \in \varphi(M)$ is forced; that is, there exists a list of (at least) two possible colors for each edge of M,

(iii) The only triangles in $Z \cup \varphi(B)$ are of the following form:



where the color indicates in where the edge lies: Z (black) and $\varphi(B)$ (red).

Here comes the key claim (💡):

Claim 5.6. *If M is a graph such that $\mathbb{P}[M \subset G(n,p)] = \Omega(1)$, then M is 2-choosable in a way that avoids monochromatic K_3.*

We can think of this as the analogue of Claim 3.2 in Section 3. To prove Claim 5.6, we first massage the condition $\mathbb{P}[M \subset G(n,p)] = \Omega(1)$ into a more usable condition using the following well known fact.

Definition 5.7. Define the *edge-vertex ratio* of a graph H by $\rho(H) := \frac{|E(H)|}{|V(H)|}$. Define the *maximum edge-vertex ratio* of a subgraph of H to be $m(H) := \max_{H' \subset H} \rho(H')$.

Proposition 5.8 ([Bol81])**.** *For any graph H, $p = n^{-1/m(H)}$ is the threshold for $G(n,p)$ containing H as a subgraph.*

Proof of Claim 5.6. First, since M is a graph such that $\mathbb{P}[M \subset G(n,p)] = \Omega(1)$, it follows that $m(M) \leq 2$ which in turn implies that either the minimum degree of M is at most 3 or M is 4-regular. We endeavour to reduce to the case of $M = K_5$.

To that end, we first reduce to the case of M being 4-regular. To that end we will show that we may assume for every $v \in M$ we have $|E(N(v))| \geq 5$. Suppose otherwise, so that $\exists v \in V(M)$ such that $|E(N(v))| \leq 4$. Then that means we can find an orientation of the edges of $G[N(v)]$ in which every vertex has out-degree at most one. We claim that we can extend every monochromatic-K_3-free 3-coloring of $M \setminus \{v\}$ to M. Indeed, for every $u \in N(v)$ assign uv a color that differs from the out-edge from u. There are no monochromatic triangles in $M \setminus \{v\}$ by assumption, and this coloring avoids any monochromatic involving v as this would involve an edge lying in $N(v)$. Consequently, this allows us to successively "peel off" such vertices v until we are left with a 4-regular graph.

Now, we show that if M is 4-colorable then $M = K_5$. Suppose otherwise. If $|E(N(v))| > 5$ for some $v \in V(M)$ then $M = K_5$ since M is a connected, 4-regular graph. That means we may assume that $|E(N(v))| = 5$ for all $v \in V(M)$. Fix some vertex v and write $N(v) = \{u_1, u_2, u_3, u_4\}$ and suppose WLOG that $u_1u_3 \notin E(N(v))$. Then we must have $N(u_1) = \{v, w, u_2, u_4\}$ for some $w \in V(M)$. However, w is not adjacent to v, u_2, u_4 since these vertices all have their four neighbors already specified. However, this would then imply that $|E(N(u_1))| \leq 3$ which is a contradiction as desired.

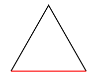
This last part is comparably the least interesting: it remains to show that K_5 is 2-choosable in a way that avoids a monochromatic K_3. This part is a case-check; I could not do it in a more succinct way than [FKSS22] so I will just reproduce verbatim their rather clean argument here:

> If some colour, say red, contains a 5-cycle, then we may colour this 5-cycle red and the complementary 5-cycle not red. If some colour class, say red, contains an edge, say e, not in a triangle, then we may colour $K_5 \setminus e$ without monochromatic triangles (this is possible as K_5 is minimally non-2-choosable) and colour e red.
>
> If none of the above is true, then each colour induces one of the following graphs: $K_3, K_4, K_4^-, K_5 \setminus K_3$ or two triangles sharing a vertex. If some colour, say red, induces $K5 \setminus K_3$, then we colour $K_{2,3}$ with red, the remaining edge of $K_5 \setminus K_3$ with not red and the edges of the K_3 in the complement with two different colours other than red. If one of the colours, say red, induces K_4 or K_4^-, then colour a C_4 with red and its diagonal with a colour other than red. Each of the remaining, uncoloured four edges can close at most one monochromatic triangle, as red is not available anywhere outside of the K_4 we have already coloured; thus we may colour them one-by-one. This leaves the case where every colour class is either K_3 or two triangles sharing a vertex. But this is impossible, since 3 does not divide $2|E(M)| = 20$.

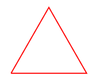
□

Combining Claim 5.6 and (ii) gives a proper coloring of $\varphi(M)$ using the available colors. In particular, by (iii) it follows that this coloring of $Z \cup \varphi(M)$ that extends the initial monochromatic K_3-free 3-coloring of Z that we started with does not introduce any monochromatic K_3-free triangles:

by definition, there are no monochromatic triangles of the form



and Claim 5.6 ensures that there are no monochromatic triangles of the form



as desired. However, this implies that $Z \cup \varphi(B) \not\to (K_3)_3$ which contradicts the second property of Theorem 5.5. Therefore, there must be $\Omega(n^2)$ edges forced to some color.

2. One way to derive a contradiction is to use the fact that every proper coloring of Z forcing many edges to the same color to show that we must therefore force a monochromatic triangle in Z, which would then give a contradiction as desired. However, such a guess is too naïve. Indeed, even though we forced $\Omega(n^2)$ edges to some color, it is conceivable however that they form a triangle-free graph.

   Remark 5.9. If we were instead trying to prove a sharp threshold for the Ramsey property of $G(n,p) \to (H)_r$ where H is a bipartite graph, then because of the Kővari-Sós-Turán theorem we would be able to circumvent this additional subtlety. This is also why we chose to work with K_3, it illustrates some of the main difficulties of the problem.

   This is where the gadgets that we precisely defined come into play. The key observation is that to force a triangle, we might as well find a rainbow constellation and that we can basically think of forcing edges as finding rainbow stars. A rainbow constellation is a 3-fold blow-up of a rainbow star which implies:

   Claim 5.10. *If a partial (i.e. we may choose to not color some edges) $\{R, B\}$-coloring of K_n contains $\Omega(n^4)$ many rainbow stars, then it must also contain $\Omega(n^9)$ many rainbow constellations.*

   We prove this in Subsection 5.2 (see Lemma 5.26). In Subsection 5.2, we also show how to "bootstrap" (💡):

   $\{\text{Claim 5.10}\}+\{\text{containers}\}+\{\text{second-moment}\} \longrightarrow \{\text{every partial proper 3-coloring of } Z \text{ forces } \Omega(n^3) \triangle\}$.

   We defer a further discussion of the the above techniques of "containers" and "second-moment [method]" to Subsection 5.2.

   Lemma 5.11. *Let $\varepsilon > 0$. Suppose $Z \sim G(n,p)$ for some $p = \Theta(n^{-1/2})$. Then with probability $> 1 - \varepsilon$, for every partial coloring of Z with three colors, if $\Omega(n^2)$ edges of Z are forced then $\Omega(n^3)$ triangles of Z are forced.*

   Remark 5.12. If the fact that we consider the stronger condition of partial colorings right now is confusing, it will become clear soon; we state Lemma 5.11 in this strengthened form with foresight towards an application of the hypergraph containers method that becomes apparent soon.

3. Now we use the first property in Theorem 5.5. Since with positive probability, $Z \cup G(n, \varepsilon p) \not\to (K_3)_3$ we should bound the probability that $G(n, \varepsilon p)$ avoids the many forced triangles from the previous step. To that end, we can apply Janson's inequality which we recall here.

   Theorem 5.13 (Janson's inequality [Jan90])**.** *Suppose Ω is a finite set and let $B_1, \ldots, B_k$ be a sequence of (not necessarily distinct) subsets of Ω, and let $R \sim \Omega_p$ for some $p \in [0, 1]$. For each $i \in [k]$, let X_i be the indicator of the event A_i that $B_i \subset R$ and let $X := \sum_i A_i$. Then, for any $0 \le t \le \mathbb{E}[X]$,*

$$\mathbb{P}[X \le \mathbb{E}[X] - t] \le \exp\left(-\frac{t^2}{2\sum_{i\sim j} \mathbb{P}[A_i \cap A_j]}\right).$$

   Lemma 5.14. *Let $p = \Theta(n^{-1/2})$ and let $\mathcal{S}$ be a set of $\Omega(n^3)$ triangles in K_n. Then*

$$\mathbb{P}[G(n, \varepsilon p) \cap \mathcal{S} = \emptyset] \le \exp\left(-O(\varepsilon n^2 p)\right).$$

   Proof. Let $G \sim G(n, \varepsilon p)$ and let $X = |G \cap \mathcal{S}|$ so that $\mathbb{E}[X] = \Omega((\varepsilon p)^3 n^3)$. Let A_i be the indicator that the ith triangle of $\mathcal{S}$ lies in G. Then we can compute

$$\sum_{i\sim j} \mathbb{P}[A_i \cap A_j] = \sum_{j=1}^{3} \sum_{\substack{T_1, T_2 \in \mathcal{S} \\ |T_1 \cap T_2| = j}} \mathbb{P}_{G\sim G(n,\varepsilon p)}[T_1 \cup T_2 \subset G] = O\left((\varepsilon p)^4 n^2\right).$$

   Substituting into Theorem 5.13 gives the desired conclusion. □

4. The goal is to find a contradiction to the first property in Theorem 5.5. To that end, we can combine Lemma 5.14 with a union bound over all colorings:

$$\frac{1}{2} \overset{?}{\le} \mathbb{P}[Z \cup G(n, \varepsilon p) \not\to (K_3)_3] \le \#\text{ 3-colorings of } Z \cdot \exp\left(-O(\varepsilon n^2 p)\right),$$

   and $\#$ 3-colorings of $Z \le 3^{|E(Z)|}$ which we typically expect to be on the order of $3^{n^2 p/2}$. However, plugging this back we do not get a contradiction to the inequality marked with a '?'.

Second attempt. What went wrong? The issue is really that we were too brutal with our union bound over all possible colorings in the final step. Instead we can get quite a lot of savings by restricting to partial colorings (i.e. colorings of $Z \sim G(n, p)$ that only assign colors to some subset of edges):

Theorem 5.15. *Suppose $G(n, p) \to (K_3)_3$ has a coarse threshold. Then there exists a constant $\varepsilon > 0$ and a subsequence $p = \Theta(n^{-1/2})$ such that for $Z \sim G(n, p)$ with probability at least ε:*

- *The hypergraph $Z \cup G(n, \varepsilon p) \not\to (K_3)_3$ with probability at least $\frac{1}{2}$.*
- *There exists a family $\mathcal{C}$ of partial colorings of Z with three colors such that:*
  - *(i) $|\mathcal{C}| = \exp(o(n^2 p))$,*
  - *(ii) every proper coloring of Z extends some $C \in \mathcal{C}$,*

(iii) every $C \in \mathcal{C}$ forces $\Omega(n^2)$ edges to some color.

This is incidentally what we were referring to in Remark 5.12. By combining the fact that Lemma 5.11 continues to work for partial colorings with the family $\mathcal{C}$ of partial colorings in Theorem 5.15 as well as the application of Janson's inequality in Lemma 5.14, we obtain the following improved union bound:

$$\frac{1}{2} \leq \mathbb{P}[Z \cup G(n, \varepsilon p) \not\to (K_3)_3] \leq |\mathcal{C}| \cdot \exp\left(-O(n^2 p)\right) = o(1)$$

which furnishes the desired contradiction as desired.

Roadmap. The above proof outline is basically complete barring two important steps: the proof of the partial coloring container Theorem 5.15 as well as the proof of the fact that we can force many triangles (Lemma 5.11) in Step 2. In the following, we establish both of these. We will first prove Theorem 5.15 in Subsection 5.1 because we think it is an illustrative example of the philosophy of using the hypergraph container method to "beat the union bound". It is apriori not clear how this philosophy is applicable to prove Lemma 5.11 and we aim to motivate the application of hypergraph containers in Subsection 5.2.

5.1 Containers for Coloring

Recall that the key point of this step is to show that we do not need to take a union bound over all colorings. To do so we show we show that colorings which "force" $\Omega(n^2)$ many edges are actually "well-clustered". We recall the specific statement that we aim to prove.

Theorem 5.16. *Suppose $G(n, p) \to (K_3)_3$ has a coarse threshold. Then there exists a constant $\varepsilon > 0$ and a subsequence $p = \Theta(n^{-1/2})$ such that for $Z \sim G(n, p)$ with probability at least ε:*

- *The hypergraph $Z \cup G(n, \varepsilon p) \not\to (K_3)_3$ with probability at least $\frac{1}{2}$.*
- *There exists a family $\mathcal{C}$ of partial colorings of Z with three colors such that:*
  - *(i) $|\mathcal{C}| = \exp(o(n^2 p))$,*
  - *(ii) every proper coloring of Z extends some $C \in \mathcal{C}$,*
  - *(iii) every $C \in \mathcal{C}$ forces $\Omega(n^2)$ edges to some color.*

To that end, we will appeal to the hypergraph containers method developed by [ST15, BMS15] which at a high level shows that independent sets in 'natural' hypergraphs are 'clustered'.

Definition 5.17. For a hypergraph $\mathcal{H}$, we defined $\Delta_t(\mathcal{H})$ to be the maximum degree of a t-element set of vertices. Precisely, we define

$$\Delta_t(\mathcal{H}) := \max\{\deg_{\mathcal{H}}(T) : T \subset V(\mathcal{H}) \text{ and } |T| = t\}$$

Theorem 5.18 (Hypergraph containers)**.** *[FKSS22, Theorem 4.7] For every positive integer k and all positive reals ε and K, there exist an integer t and a positive real δ such the following holds. Suppose that a nonempty k-uniform (multi)hypergraph $\mathcal{G}$ with vertex set V and a positive real τ satisfy $\Delta_\ell(\mathcal{G}) \leq K\tau^{\ell-1}\frac{e(\mathcal{G})}{v(\mathcal{G})}$ for every $\ell \in [k]$. Then there exists a function $f : \mathcal{P}(V)^t \to \mathcal{P}(V)$ with the following properties:*

(i) For every $I \subset V$ satisfying $e(\mathcal{G}(I)) \leq \delta\tau^k e(\mathcal{G})$, there are $S_1, \ldots, S_T \subset I$ with at most $\tau v(\mathcal{G})$ elements each such that $I \subset f(S_1, S_2, \ldots, S_t)$.

(ii) For every $S_1, \ldots, S_t \subset V$, the set $f(S_1, \ldots, S_t)$ induces fewer than $\varepsilon e(\mathcal{G})$ edges in $\mathcal{G}$.

For a proof of this formulation of the container theorem, we refer the interested reader to [FKSS22, Appendix B]. We also refer the interested reader to the wonderful survey [BMS18] for an account of the many applications of hypergraph containers.

Roadmap. Let $Z \sim G(n,p)$. For the rest of this section we first begin by constructing a hypergraph $\mathcal{T}$ (where $V(\mathcal{T}) = E(Z) \times [3]$) such that proper colorings of Z correspond to independent sets in $\mathcal{T}$. Next, we need to show that $\mathcal{T}$ is a "natural" hypergraph to invoke Theorem 5.18. Specifically, we check that $\mathcal{T}$ satisfies the condition on degree sequences in Theorem 5.18.

Let $Z \sim G(n,p)$. In order to construct $\mathcal{T}$, we show how to encode proper colorings via the "boosting" subgraph M. More precisely, by Theorem 5.5 with probability at least ε, we have $\mathbb{P}[Z \cup G(n, \varepsilon p)$ is 3-colorable$] \geq \frac{1}{2}$ while for at least ε proportion of $M \in \mathcal{Z}$, we have that $Z \cup M$ is not 3-colorable. We henceforth restrict to these ε fraction of instances Z and suppose that the fixed graph M have K edges.

Now, the key idea is to track colorings of Z such that taken in conjunction with some proper coloring M does not create any monochromatic triangles in $M \cup Z$. (One might object that $M \cap Z$ could contain some edges and this creates complications; this is however an unlikely event per a similar calculation as that in (3).) Precisely, let $\mathcal{M}$ be the collection of all the automorphic copies of M in K_n and let

$$\mathcal{M}_Z := \{M \in \mathcal{M} : Z \cup M \text{ not 3-colorable}\}.$$

Definition 5.19. Let $Z \sim G(n,p)$ and let $B \subset K_n$ be such that $B \cap Z = \emptyset$. A coloring $\varphi \colon Z \to [3]$ is *consistent* with a coloring ψ of B if $\varphi \cup \psi$ does not create any monochromatic triangles.

Definition 5.20. Let $Z \sim G(n,p)$. A coloring $\varphi \colon Z \to [3]$ is *enlargeable* if there exists $M \in \mathcal{M}_Z$ such that $M \cap Z = \emptyset$ and a coloring $\psi \colon M \to [3]$ such that φ is consistent with ψ.

Since enlargeable colorings of Z must necessarily not be proper colorings, it suffices to construct a hypergraph $\mathcal{T}$ with $V(\mathcal{T}) = E(Z) \times [3]$ such that hyperedges of $\mathcal{T}$ encode these enlargeable colorings. In other words, we construct a hypergraph $\mathcal{T}$ such that

$$E(\mathcal{T}) \leftrightarrow \{\text{enlargeable colorings of } Z\} \leftrightarrow \{\text{improper colorings of } Z\}.$$

First, we show that a positive fraction of Z has the property that all $M \in \mathcal{M}_Z$ is well-behaved. Because of our desire to track stars, it is natural to consider the following quantity:

$$I(M, Z) := \{S \setminus M : S \in \mathcal{S} \text{ is such that } S \cap M, S \cap Z \neq \emptyset\}$$

where $\mathcal{S}$ denotes the triangles in K_n.

Claim 5.21. *Fix $\varepsilon > 0$. For $Z \sim G(n,p)$ (for the parameters as defined earlier), the expected number of $M \in \mathcal{M}_Z$ that do not satisfy the following properties is at most $0.1\varepsilon^2|\mathcal{M}_Z|$:*

(1) The edge sets of Z and M are disjoint,

(2) Every nonempty set in $I(M,Z)$ has 2 elements,

(3) The sets in $I(M,Z)$ are pairwise disjoint,

(4) The family $I(M,Z)$ contains at most L sets.

where L is chosen such that $K^L/L! \leq 0.1\varepsilon^2$.

We defer the proof of Claim 5.21 to the end of the section. None of these conditions should be surprising; a triangle is specified by two edges, and since Z is sparse we would expect there to be very few triangles that contain two edges of M and one edge of Z. Similarly, since Z is sparse and M has $O(1)$ size, we would expect there to be very few "hitting set" triangles that intersect Z and some $M \in \mathcal{M}_Z$. For the purpose of constructing our containers for the coloring via the hypergraph $\mathcal{T}$, the last property is relevant so that we can bound the size of the edges in $\mathcal{T}$. The earlier properties of $I(M,Z)$ are useful for checking that the codegree conditions of $\mathcal{T}$ satisfies those in Theorem 5.18.

Let $\mathcal{M}'_Z \subset \mathcal{M}_Z$ be those M that satisfy the four properties listed in Claim 5.21. We say that $Z \in G(n,p)$ is *nice* if $|\mathcal{M}_Z| \geq \varepsilon/2 \cdot |\mathcal{M}|$. Then a consequence of Claim 5.21 is that since with probability ε we have that $|\mathcal{M}_Z| \geq \varepsilon|\mathcal{M}|$, it follows that

$$\mathbb{P}[Z \text{ is nice}] \geq \varepsilon - \mathbb{P}[|\mathcal{M}_Z| - |\mathcal{M}'_Z| \leq \varepsilon/2|\mathcal{M}|] \geq \varepsilon - \frac{\mathbb{E}[|\mathcal{M}_Z| - |\mathcal{M}'_Z|]}{\varepsilon/2 \cdot |\mathcal{M}|} \geq \varepsilon - \frac{4\lambda}{\varepsilon} \geq \varepsilon/2.$$

We henceforth restrict our attention to this positive density of nice Z.

Definition 5.22. A given coloring C of Z is *enlargeable* if and only if there exists $M \in \mathcal{M}_Z$, a proper 3-coloring ψ of M and a 3-coloring $\varphi_{C,\psi} \colon \bigcup I(M,Z) \to [3]$ of all the elements in $I(M,Z)$ that is consistent with ψ and furthermore $\varphi_{C,\psi} \subset C$.

This motivates constructing the hypergraph $\mathcal{T}$ with:

- $V(\mathcal{T}) = E(Z) \times [3]$, and
- $E[\mathcal{T}]$ contains edges $\{(e, \varphi_{C,\psi}(e)) : e \in \bigcup I(M,Z)\}$ for all choices of $M \in \mathcal{M}$ and proper colorings ψ of M for which $\varphi_{C,\psi}$ is well-defined.

Apriori it is possible that $\mathcal{T}$ has some very high degree sets (i.e. $T \in \binom{V}{t}$ such that $\deg_{\mathcal{T}}(T)$ is large) that makes $\mathcal{T}$ not satisfy the codegree conditions of Theorem 5.18. To that end, we will first prune out these high degree vertices to a subhypergraph $\mathcal{T}' \subset \mathcal{T}$ that contains about as many edges as $\mathcal{T}$ but with more well-behaved codegree conditions.

Lemma 5.23. *Let $\mathcal{T}$ be the hypergraph we defined earlier corresponding to some $Z \sim G(n,p)$. Then with probability at least ε, there exists a subhypergraph $\mathcal{T}' \subset \mathcal{T}$ that has at least $|E(\mathcal{T})| - \varepsilon/8 \cdot |\mathcal{M}|$ edges that satisfies:*

- $\Delta_1(\mathcal{T}') = O\left(\varepsilon^{-1}\frac{|\mathcal{M}|}{n^{3/2}}\right)$, *and*

- $\Delta_2(\mathcal{T}') = o\left(\frac{|\mathcal{M}|}{n^{3/2}}\right)$.

We prove Lemma 5.23 at the end of the section. Next, we combine these pieces with Theorem 5.18 to deduce Theorem 5.16.

Proof of Theorem 5.16. We restrict to $\mathcal{T}'$ that satisfies Lemma 5.23 which occurs with probability at least ε, and we may also assume that $|E(Z)| \lesssim n^2p$. This implies the following bound on the number of edges in $\mathcal{T}'$

$$|E(\mathcal{T}')| \leq 3 \cdot O(n^2p) \cdot \Delta_1(\mathcal{T}') = O(\varepsilon^{-1}|\mathcal{M}|).$$

Note that every edge of $\mathcal{T}'$ has cardinality at most $2\max_{M\in\mathcal{M}'_Z} |I(M,Z)| \leq 2L$ by Claim 5.21, and so we can write $\mathcal{T}' = \bigsqcup_{i=1}^{2L} \mathcal{T}_i$ where $\mathcal{T}_i$ is the subhypergraph of $\mathcal{T}'$ that comprises edges of $\mathcal{T}'$ of cardinalities i. The whole point of this step is to split $\mathcal{T}'$ into i-regular subhypergraphs to apply the hypergraph container lemma for each subhypergraph, since Theorem 5.18 only applies to regular hypergraphs.

Define $U := \{u \in [2L] : |E(\mathcal{T}_u)| \gtrsim \varepsilon/L \cdot |\mathcal{M}|\}$ and fix an arbitrary $u \in U$ so that by Lemma 5.23, we have $\Delta_1(\mathcal{T}_u) \leq \Delta_1(\mathcal{T}') = O(L/\varepsilon^2|E(\mathcal{T}_u)|/|V(\mathcal{T}_u)|)$ and $\Delta_2(\mathcal{T}_u) \leq \Delta_2(\mathcal{T}') \leq o(L/\varepsilon \cdot |E(\mathcal{T}_u)|/|V(\mathcal{T}_u)|)$. Consequently, by applying Theorem 5.18, we obtain $t(\varepsilon, L)$ and a collection $\mathcal{C}_u$ of at most

$$\left(\sum_{i=0}^{o(|E(Z)|)} \binom{O(|E(Z)|)}{i}\right)^t \leq \exp(o(|E(Z)|))$$

subsets of $E(Z) \times [3]$ such that:

- Every proper coloring $\psi\colon Z \to [3]$ of Z is contained (where we view the coloring naturally as a subset of $E(Z) \times [3]$) in a member of $\mathcal{C}_u$.
- Every member of $\mathcal{C}_u$ induces fewer that $O(\varepsilon/L|\mathcal{M}|)$ edges in $\mathcal{T}_u$.

To finish up, we aggregate the information from all the $\mathcal{T}_i$: let our final collection of partial colorings Ψ be defined as all partial colorings $\psi\big|_C$ where C is a set of the form $C = \bigcap_{u\in U} C_u$ (i.e. we restrict to consider colorings ψ of Z in the $\mathcal{C}_u$ that assign the same color to edges in C).

It remains to check that Ψ has all the claimed properties in the statement of the theorem. To that end, we note that:

- For each $u \in U$, every proper coloring ψ of Z is contained in some $C_u \in \mathcal{C}_u$, and so the proper coloring ψ extends ψ_C.
- We claim that every partial coloring $\psi_C \in \Psi$ forces at least $\Omega(n^2)$ edges to some color. Indeed, since $|E(\mathcal{T}_u[C])| \leq |E(\mathcal{T}_u[C_u])| \leq O(\varepsilon/L\cdot|\mathcal{M}|)$ if $u \in U$ and we also have $|E(\mathcal{T}_u[C])| \leq |E(\mathcal{T}_u)| \leq O(\varepsilon/L\cdot|\mathcal{M}|)$ if $u \notin U$. It follows that

$$|E(\mathcal{T}'[C]) = \sum_{i=1}^{2L} |E(\mathcal{T}_u[C])| \leq 2L \cdot O(\varepsilon/L \cdot |\mathcal{M}|) = O(\varepsilon|\mathcal{M}|),$$

  which in turns implies that

$$|E(\mathcal{T}[C])| \leq |E(\mathcal{T}'[C])| + |E(\mathcal{T})| - |E(\mathcal{T}')| = O(\varepsilon|\mathcal{M}|). \tag{4}$$

However, to finish up we note that for any $C \subset E(Z) \times [3]$, if we denote the number of edges forced to a color by C as f_C, then

$$|E(\mathcal{T}[C]) \geq |\mathcal{M}'_Z| - f_C \cdot K\frac{|\mathcal{M}|}{n^2}$$

because every edge lies in at most $K\frac{|\mathcal{M}|}{n^2}$ copies of M, and for every $M \in \mathcal{M}'_Z$ that does not contain any edges whose colors are forced by C, we can find an edge in $\mathcal{T}[C]$ corresponding to M: by definition for any $e \in I(M, Z)$ we have that there are at least two possible choices of color for $e \cap M$. Since M is 2-choosable it follows that we can construct a proper coloring ψ of M that is consistent with Z, and this ψ would then give rise to an edge $\{(e, \varphi_{C,\psi}(e)) : e \in I(M, Z)\}$ in $\mathcal{T}$. That is, our construction of $\mathcal{T}$ ensures that for some coloring C the edges of $\mathcal{T}[C]$ witnesses the proper colorings of $M \in \mathcal{M}'_Z$ that are consistent with C.

Combining this with (4) implies that $f_C \geq \Omega(\varepsilon/K \cdot n^2)$ since $|\mathcal{M}'_Z| \geq (\varepsilon/2) \cdot |\mathcal{B}|$.

□

Deferred proofs.

Proof of Claim 5.21. Since for any fixed $M \in \mathcal{M}_Z$, we have that $\mathbb{E}[Z \cap M] = p|M| = o(1)$, it follows from Markov's inequality that the fraction of $\mathcal{M}_Z$ that does not satisfy (1) is $o(1)$. Fix $M \in \mathcal{M}$. Suppose a nonempty set in $I(M, Z)$ has only one element. Then it follows that there is a triangle that contains two elements of M which fixes the edge that lies in Z, and therefore this occurs with probability p. The maximum number of triangles that correspond to a singleton element in $I(M, Z)$ is $\binom{|M|}{2}$. Consequently, on expectation the fraction of $M \in \mathcal{M}_Z$ that does not satisfy (2) is at most $p\binom{K}{2} = o(1)$.

Remove all the elements from $\mathcal{M}_Z$ that do not satisfy (1) and (2) to get $\mathcal{M}'_Z$. The expected size of $\mathcal{M}'_Z$ satisfies $\mathbb{E}\,|\mathcal{M}'_Z| = (1 - o(1))|\mathcal{M}_Z|$. Having ensured that (2) holds, it follows that for any fixed $M \in \mathcal{M}'_Z$ the expected number of pairs of $S_1, S_2 \in I(M, Z)$ satisfying $S_1 \cap S_2 \neq \emptyset$ is at most $p^3 n = o(1)$. Consequently, on expectation the fraction of $M \in \mathcal{M}_Z$ that does not satisfy (3) is also $o(1)$.

Finally, let $\mathcal{M}''_Z \subset \mathcal{M}_Z$ consist of all the elements that satisfy (1), (2) and (3). We have $\mathbb{E}\,|\mathcal{M}''_Z| = (1 - o(1))|\mathcal{M}_Z|$. For $M \in \mathcal{M}''_Z$ and $Z \sim G(n, p)$ let $i(M, Z)$ be the largest integer ℓ such that there exists triangles $T_1, \ldots, T_\ell$:

- $T_1, \ldots, T_\ell \in M \cup Z$,
- $|T_i \cap M| = 1$ for all $1 \leq i \leq \ell$,
- $T_1 \setminus M, \ldots, T_\ell \setminus M$ are pairwise disjoint.

It suffices to prove that $\mathbb{P}[i(B, Z) > L] \leq 0.1\varepsilon^2$. To that end, it suffices to note that

$$\mathbb{P}[i(B, Z) > L] \leq \frac{\mathbb{E}[i(B, Z)]}{L!} = \frac{(|M| \cdot n)^L p^{2L}}{L!} \leq 0.1\varepsilon^2$$

by the given assumption on L, as desired. □

Proof of Lemma 5.23. We begin by noting that

$$\mathop{\mathbb{E}}_{Z\sim G(n,p)}\left[\sum_{v\in V(\mathcal{T})}\deg_{\mathcal{T}}(v)^2\right]\le O\left(\frac{|\mathcal{M}|^2}{n^{3/2}}\right) \tag{5}$$

where we note that colorings add constant factors so we can just consider the underlying graph, and so for any $e\in E(K_n)$ and $i\in[3]$, we have $\deg_{\mathcal{T}}(e,i)\le O(\#\{\text{triangle } T\ni e : |(T\setminus\{e\})\cap Z|=1\}\cdot\frac{|\mathcal{M}|\cdot K}{\binom{n}{2}})$, and $\mathbb{E}_e\,\#\{\text{triangle } T\ni e : |(T\setminus\{e\})\cap Z|=1\}^2\le O(p^3n^2)=O(n^{1/2})$.

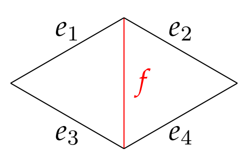
Furthermore, we claim that

$$\mathop{\mathbb{E}}_{Z\sim G(n,p)}\left[\sum_{T\in\binom{V(\mathcal{T})}{2}}\deg_{\mathcal{T}}(T)^2\right]\le o\left(\frac{|\mathcal{M}|^2}{n^{3/2}}\right). \tag{6}$$

To prove this, we define for distinct edges $e_1,e_2\in E(K_n)$ the quantity $\gamma_2(e_1,e_2)$ that records the number of triples of (T_1,T_2,M) where T_1,T_2 are triangles and $M\in\mathcal{M}$ such that:

- $e_1\in T_1$, $e_2\in T_2$,
- $T_1\cap M=T_2\cap M$ and $|T_1\cap M|=1$,
- $(T_1\cup T_2)\setminus(M\cup\{e_1,e_2\})\subset Z$.

Pictorially, $\gamma_2(e_1,e_2)$ basically counts objects (technically it encodes such objects in terms of the two triangles and also M) that look like:



where $e_1,e_2,e_3,e_4\in Z$ and $f\in M$. It is not difficult to check that

$$\mathbb{E}\left[\sum_{T\in\binom{V(\mathcal{T})}{2}}\deg_{\mathcal{T}}(T)^2\right]\le O\left(\mathbb{E}\left[\sum_{e_1,e_2\in Z}\gamma_2(e_1,e_2)^2\right]\right)=O\left(\frac{|\mathcal{M}|^2K^2}{\binom{n}{2}^2}\cdot p^6n^4\right)=o\left(\frac{|\mathcal{M}|^2}{n^{3/2}}\right).$$

These suffice to establish the desired conclusion because by iteratively removing $\varepsilon/16\cdot|\mathcal{M}|$ hyperedges of $\mathcal{T}$ that contain some vertex of $V(\mathcal{T})$ with largest degree and then iteratively removing $\varepsilon/16\cdot|\mathcal{M}|$ hyperedges of $\mathcal{T}$ that contain some $\binom{V(\mathcal{T})}{2}$ with largest degree, we get $\mathcal{T}'$ with at least $|E(\mathcal{T})|-\varepsilon/8\cdot|\mathcal{M}|$ hyperedges. By assumption of maximality, it is not difficult to check that: $\Delta_1(\mathcal{T}')\le\frac{16}{\varepsilon}\sum_{v\in V(\mathcal{T})}\deg_{\mathcal{T}}(v)^2$ and $\Delta_2(\mathcal{T}')\le\frac{16}{\varepsilon}\sum_{T\in\binom{V(\mathcal{T})}{2}}\deg_{\mathcal{T}}(T)^2$ which implies the desired conclusion by combining (5), (6) and Markov's inequality. □

5.2 Stars $\implies$ Constellations

Lemma 5.24. *Let $\varepsilon>0$. Suppose $Z\sim G(n,p)$ for some $p=\Theta(n^{-1/2})$. For every $\beta_1>0$ there exists $\beta_2>0$ with the following property: with probability $>1-\varepsilon$ and a partial coloring of Z with three colors, if β_1n^2 edges of Z are* *forced then $\beta_2 n^3$ triangles of Z are forced.*

Given the intimate relationships between forcing edges/triangles and rainbow stars/constellations, we will instead prove the following.

Lemma 5.25. *[Equivalent formulation of Lemma 5.25] Let $\varepsilon > 0$. Suppose $Z \sim G(n,p)$ for some $p = \Theta(n^{-1/2})$. For every $\beta_1 > 0$ there exists $\beta_2 > 0$ with the following property: with probability $> 1-\varepsilon$, for every partial coloring ψ of Z with three colors, if ψ has $\beta_1 n^2$ rainbow stars then C also has $\beta_2 n^3$ rainbow constellations.*

To deduce Lemma 5.25 from Lemma 5.25 we need to show that in typical instances of $Z \sim G(n,p)$ we have that if there are $\Omega(n^3)$ rainbow constellations then they are typically supported by distinct triangles. This follows by "ℓ^2 control" via the Cauchy-Schwarz inequality.

Deducing Lemma 5.25 from Lemma 5.25. For a triangle T and an instance $Z \sim G(n,p)$, we write $C(T,Z)$ for the number of constellations in Z supported on T. Then because a constellation θ supported on T has the property that $v \in \theta \setminus T$ has degree 2, it follows that

$$\mathop{\mathbb{E}}_{Z\sim G(n,p)}\left[\sum_{\text{triangle } T} C(T,Z)^2\right] \le O(n^3).$$

An application of Markov's inequality implies that with high probability (by enlarging constants necessarily) we have that $\sum_{\text{triangle } T} C(T,Z)^2 \le O(n^3)$. Furthermore, by Lemma 5.25, with high probability if a partial coloring C of Z with three colors has $\Omega(n^2)$ forced edges of Z (that is, it has $\Omega(n^2)$ rainbow stars) it follows that C has $\Omega(n^3)$ rainbow constellations and let $\mathcal{T}$ be the collection of triangles that supports at least one of these rainbow constellations. By the Cauchy-Schwarz inequality, it follows that

$$\Omega(n^6) = \left(\sum_{T\in\mathcal{T}} C(T,Z)\right)^2 \le |\mathcal{T}|\cdot\left(\sum_{T\in\mathcal{T}} C(T,Z)^2\right) \le |\mathcal{T}|\cdot\left(\sum_{\text{triangles } T} C(T,Z)^2\right) = |\mathcal{T}|\cdot O(n^3),$$

which rearranges to give the desired bound on $|\mathcal{T}|$. □

We first claim that this desired relationship between stars and constellations holds for $E(K_n)\times[3]$ (identified with 3-colored K_n in the obvious way). We abuse notation in what follows, and we talk about rainbow stars in $E(K_n)\times[3]$ we mean the natural identification in the first component with the edges, and the corresponding color of each edge in the second component.

Lemma 5.26. *Any 3-coloring of K_n that contains $\Omega(n^4)$ rainbow stars also contains $\Omega(n^9)$ rainbow constellations.*

This lemma basically states that if we have in some sense close to the maximal possible density of rainbow stars then we should also have close to the maximal possible density of rainbow constellations. We can thereby think of Lemma 5.25 as a suitably sparsified version of Lemma 5.26: indeed the "expected maximal number of rainbow stars" is $\sim p^4n^4 = \Theta(n^2)$ and the "expected maximal number of rainbow constellations" is $\sim p^6n^9 = n^3$.

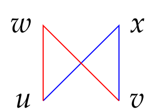
The key observation for Lemma 5.25 is that a rainbow constellation is a 3-fold blow-up of a rainbow star. To gain some intuition for the counts, we introduce the following notation for the homomorphism density of a 3-colored graph F in a host 3-colored graph G:

$$\hom(F,G) := \frac{\#\{\text{homomorphisms from } F \text{ to } G\}}{|V(G)|^{|V(F)|}}$$

where a (colored) homormorphism $\varphi\colon V(\widetilde{F}) \to V(\widetilde{G})$ where $\widetilde{F}$ and $\widetilde{G}$ are the projection to the first coordinate (i.e. underlying graph) of 3-colored $F \subset E(K_n) \times [3]$ to 3-colored $G \subset E(K_n) \times [3]$ such that if $uv \in E(\widetilde{F})$ then $\varphi(u)\varphi(v) \in E(\widetilde{G})$ and furthermore the color of uv and color of $\varphi(u)\varphi(v)$ are the same. Another interpretation of $\hom(F,G)$ is that it is the probability that a random map $\varphi\colon V(\widetilde{F}) \to V(\widetilde{G})$ is a valid colored homomorphism. Since a rainbow constellation is a 3-fold blow-up of a rainbow-constellation, it is reasonable to expect that if for some $G \subset E(K_n) \times [3]$ the probability p_1 that a random map $\varphi_1\colon V(\text{star}) \to V(\widetilde{G})$ is a valid colored homomorphism satisfies $p_1 = \Omega(1)$ then we would expect that the probability p_2 that a random map $\varphi_2\colon V(\text{constellation}) \to V(\widetilde{G})$ satsifies $p_2 \geq p_1^3 = \Omega(1)$. Now unwinding the definitions shows that $p_1 = \Omega(1) \Leftrightarrow \exists\, \Omega(n^3)$ rainbow stars in G and $p_2 = \Omega(1) \Leftrightarrow \exists\, \Omega(n^9)$ rainbow constellations in G.

Consequently, in order to prove Lemma 5.25 the main idea would be somehow to transfer this relationship between stars and constellations from K_n to the sub-sampled/percolated $G(n,p)$. To that end, we build the following hypergraphs of stars and constellations; both hypergraphs live on the vertex set $E(K_n) \times [3]$:

- The edges of 4-uniform $\mathcal{R}_s$ encodes all the rainbow stars, so that $|E(\mathcal{R}_s)| = \Omega(n^4)$. More concretely, suppose red $\leftrightarrow$ 1, blue $\leftrightarrow$ 2, green $\leftrightarrow$ 3 then

  corresponds to the edge $\{(uw,1),(vw,1),(ux,2),(xv,2)\}$ in $\mathcal{R}_s$.

- The edges of the 12-uniform $\mathcal{R}_c$ encodes all the rainbow constellations so that $|E(\mathcal{R}_c)| = \Omega(n^9)$.

In this language, an equivalent formulation of Lemma 5.26 is the following:

Lemma 5.27. *For every $C \subset E(K_n) \times [3]$ and every $\beta_1 > 0$, there exists β_2 such that we have that if $|E(R_1[C])| \geq \beta_1 n^4$ then $|E(R_2[C])| \geq \beta_2 n^9$.*

Recall that the goal is to show that with high probability, every coloring of $Z \sim G(n,p)$ with small number of rainbow constellations also have a small number of rainbow stars. In the proof outline, we mentioned that one way to do so is to combine hypergraph containers, the second moment method and Lemma 5.27. It is perhaps now more clear why hypergraph containers will come in useful: it allows us to have some savings on the union bounding over all colorings of Z.

Now, we employ the hypergraph containers method on the hypergraph $\mathcal{R}_c$ to "cluster" colorings that have few rainbow constellations. Because a constellation θ supported on triangle T has the property that $v \in \theta \setminus T$ has degree 2, the codegree conditions of Theorem 5.18 are satisfied with $\tau = o(p)$. Let γ_2 be a sufficiently small constant to be determined. By Theorem 5.18, there exists constant $t(\gamma_2)$, $\gamma_1(\gamma_2)$ and $f\colon \mathcal{P}(E(K_n) \times [3])^t \to E(K_n) \times [3]$ such that:

(†) For every partial 3-colouring ψ of K_n with fewer than $\Omega(n^3)$ rainbow constellations, there are $S_1, \dots, S_t \subset E(K_n) \times [3]$ consisting of at most $o(n^{3/2})$ edges each such that $\psi \subset f(S_1, \dots, S_t)$.

(‡) For every $S_1, \dots, S_t \subset E(K_n) \times [3]$, the set $f(S_1, \dots, S_t)$ induces $\lesssim \gamma_2 n^9$ edges in $\mathcal{R}_c$, and so by Lemma 5.27 $f(S_1, \dots, S_t)$ also induces $\lesssim \gamma_1 n^4$ edges in $\mathcal{R}_s$.

Let $Z \sim G(n, p)$ with $p = \Theta(n^{-1/2})$. Basically, to prove Lemma 5.25, we suppose for the sake of contradiction that we have a partial coloring ψ of Z with $\beta_1 n$ rainbow stars but less than $\beta_2 n^3$ rainbow constellations. Then we apply the hypergraph container as above to obtain $S_1, \dots, S_t$ such that $\psi \subset f(S_1, \dots, S_t)$. In particular, we can arrange for at most $\exp(o(n^{3/2}))$ containers for each of these partial colorings ψ. Since each of these contains induces $\lesssim \gamma_1 n^3$ edges in $\mathcal{R}_s$, we would guess that by taking γ_2 sufficiently small so that γ_1 is sufficiently small and in particular so that $\gamma_1 < \beta_1$, then with high probability since $f(S_1, \dots, S_t)$ has $\lesssim \gamma_1 n^4$ rainbow triangles, then $\psi \subset f(S_1, \dots, S_t)$ would have $\lesssim \beta_1 p^4 n^4 = \beta_1 n^2$ rainbow stars and that would give a contradiction.

However, generally speaking, bounding upper tails is hard, and in this case apriori we may not be able to guarantee that the upper tails decay sufficiently fast (we need exponential decay). One of the nice ideas in [FKSS22] is to use the second moment method to demonstrate that the number of stars in Z is typically concentrated. That is, we would expect $\mathcal{R}_s[E(Z) \times [3]]$ to be concentrated and therefore not too large. In particular, if there are many *rainbow* stars in $\psi = (E(Z) \times [3]) \cap f(S_1, \dots, S_t)$, then there are very few rainbow stars in $\mathcal{R}_s[E(Z) \times [3]] \setminus f(S_1, \dots, S_t)$.

Let $\pi(\cdot)$ be the projection to the first coordinate; that is, π extracts the underlying (uncolored) graph structure. Since $f(S_1, \dots, S_t)$ induces $\lesssim \gamma_1 n^4$ edges in $\mathcal{R}_s$, it follows that there are many rainbow stars in $\mathcal{R}_s[\pi(f(S_1, \dots, S_t)) \times [3]] \setminus f(S_1, \dots, S_t)$. In other words, we are reduced to the problem of bounding the probability that $\mathcal{R}_s[E(Z) \times [3]] \setminus f(S_1, \dots, S_t)$ avoids the many rainbow stars in $\mathcal{R}_s[\pi(f(S_1, \dots, S_t)) \times [3]] \setminus \mathcal{R}_s[E(Z) \times [3]] \setminus f(S_1, \dots, S_t)$. This lower tail event is something that we can get an exponential tail bound for by appealing to Janson's inequality (Theorem 5.13). We summarize this proof idea (💡) with the following schematic:

$$\begin{Bmatrix}\text{upper tail bound on}\\ \#\text{ rainbow stars in }\psi\end{Bmatrix} \xrightarrow[\#\text{(uncolored) stars in } Z]{\text{second moment}} \begin{Bmatrix}\text{lower tail bound on }\#\text{ rainbow}\\ \text{stars in } \mathcal{R}_s[E(Z)\times[3]]\setminus\psi\end{Bmatrix}.$$

We formalize the above proof sketch in what follows.

Proof of Lemma 5.25. Let $\mathcal{E}$ be the event that there is a color ψ that fails the conditions of the Lemma; that is, suppose ψ has $\beta_1 n^2$ rainbow stars but less than $\beta_2 n^3$ rainbow constellations. Let γ_2 be a sufficiently small constant to be determined. By Theorem 5.18, there exists constant $t(\gamma_2)$, $\gamma_1(\gamma_2)$, $S_1, \dots, S_t \subset E(K_n) \times [3]$ and $f : \mathcal{P}(E(K_n) \times [3])^t \to E(K_n) \times [3]$ such that:

(†) For every partial 3-colouring ψ of K_n with fewer than $\Omega(n^3)$ rainbow constellations, there are $S_1, \dots, S_t \subset E(K_n) \times [3]$ consisting of at most $o(n^{3/2})$ edges each such that $\psi \subset f(S_1, \dots, S_t)$.

(‡) For every $S_1, \dots, S_t \subset E(K_n) \times [3]$, the set $f(S_1, \dots, S_t)$ induces $\lesssim \gamma_2 n^9$ edges in $\mathcal{R}_c$, and so by Lemma 5.27 $f(S_1, \dots, S_t)$ also induces $\lesssim \gamma_1 n^4$ edges in $\mathcal{R}_s$.

Let $\beta > 0$ be a constant to be determined.

Let S_Z be the the number of stars supported by edges of Z. Then by Chebyshev's inequality, it follows that

$$\mathbb{P}_{Z\sim G(n,p)}[S_Z > (1+\beta)\,\mathbb{E}[S_Z]] \leq \frac{\mathrm{Var}[S_Z]}{\beta^2\,\mathbb{E}[S_Z]^2} = \frac{O(p^7n^6 + p^6n^5)}{\beta^2 p^4 n^4} = o(1). \tag{7}$$

That is, with high probability we have that $S_Z \leq (1+\beta)\,\mathbb{E}[S_Z] = (1+\beta)p^4\binom{n}{2}$. That is, there are at most $6(1+\beta)p^4\binom{n}{4}$ edges in $\mathcal{R}_s[E(Z)\times[3]]$. However, there are also at least $\beta_1 n^2$ edges in $\mathcal{R}_s[\psi]$. Since $\psi = (E(Z)\times[3]) \cap f(S_1,\ldots,S_t)$ it follows that by choosing β sufficiently small we can ensure that

$$|E(\mathcal{R}_s[E(Z)\times[3]] \setminus \mathcal{R}_s[f(S_1,\ldots,S_t)])| \leq 3p^4\binom{n}{4}.$$

However, since by assumptions $|E(\mathcal{R}_s[f(S_1,\ldots,S_t)])| \leq \gamma_1 n^4$ by (‡), it follows that

$$\mathbb{E}\,|E(\mathcal{R}_s[E(Z)\times[3]] \setminus \mathcal{R}_s[f(S_1,\ldots,S_t)])| \geq 6(1-\gamma_1)p^4\binom{n}{4}.$$

By choosing γ_2 sufficiently small we can ensure that $\gamma_1 \leq \frac{1}{3}$. To finish up, we apply Janson's inequality to the random variable $\alpha(S_1,\ldots,S_t) := |E(\mathcal{R}_s[E(Z)\times[3]] \setminus \mathcal{R}_s[f(S_1,\ldots,S_t)])|$. It is not difficult to check that the denominator in Janson's inequality (Theorem 5.13) is $\Theta(p^7n^6) = n^{5/2}$ (this is essentially the same calculation that went in (7), which gives

$$\mathbb{P}_{Z\sim G(n,p)}\left[\alpha \leq 3p^4\binom{n}{4}\right] = \mathbb{P}_{Z\sim G(n,p)}\left[\alpha \leq \mathbb{E}[\alpha] - p^4\binom{n}{4}\right] \leq \exp\left(-O\left(\frac{n^2}{n^{5/2}}\right)\right) = \exp\left(-O(n^{3/2})\right).$$

To finish up, let π denote projection onto the first coordinate (i.e. the process of extracting the underyling subgraph). We union bound over all the containers to get that

$$\mathbb{P}[\mathcal{E}] = \mathbb{P}[S_Z > (1+\beta)\,\mathbb{E}[S_Z]] + \sum_{(S_1,\ldots,S_t)} \mathbb{P}\left[\pi(S_1\cup\ldots\cup S_t) \subset Z \cap \alpha(S_1,\ldots,S_t) \leq 3p^4\binom{n}{4}\right].$$

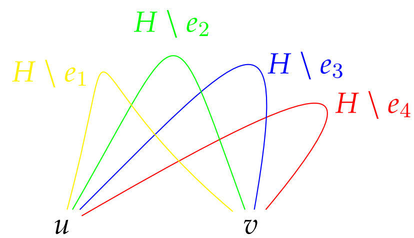
Finally, note that $\pi(S_1\cup\ldots\cup S_t)\subset Z$ is an increasing event, while $\alpha(S_1,\ldots,S_t) \leq 3p^4\binom{n}{4}$ is a decreasing event. Consequently, we can use Harris' inequality to write:

$$\begin{aligned}
\mathbb{P}[\mathcal{E}] &\leq \mathbb{P}[S_Z > (1+\beta)\,\mathbb{E}[S_Z]] + \sum_{(S_1,\ldots,S_t)} \mathbb{P}\left[\pi(S_1\cup\ldots\cup S_t)\subset Z\right]\mathbb{P}\left[\alpha(S_1,\ldots,S_t) \leq 3p^4\binom{n}{4}\right] \\
&\leq o(1) + \exp\left(-O(n^{3/2})\right)\sum_{m=0}^{o(n^{3/2})}\binom{n^2}{m}\cdot(3\cdot 2^t)^m \\
&\leq o(1) + \exp\left(-O(n^{3/2})\right)\exp\left(o(n^{3/2})\right) = o(1),
\end{aligned}$$

as desired. □

5.3 Generalizing the argument to prove Theorem 5.4

In the general setting of H and $r \geq 3$, the most natural generalizations of *star* would be the following gadget, in which we "glue" together $(r-1)$ copies of H with an edge missing:



It is also natural to define the analogous notion of constellations, where we take a "blow-up" of the star relative to H; that is, we glue one copy of the star to each edge of H. We can basically repeat the same proof as before, but now each of Subsection 5.1 and Subsection 5.2 becomes considerably more involved; in these sections we often appealed to the facts such as " a constellation θ supported on triangle T has the property that $v \in \theta \setminus T$ has degree 2" to bound certain quantities, which is no longer true for these general stars and constellations. Overall, it is much more difficult to count these generalized gadgets and their intersection patterns, which accounts for the technical proofs in [FKSS22].

There are additional technicalities posed for r large and H such that $|E(H)|$ is large, in which the way we built $\mathcal{T}$ in Subsection 5.1 gives containers that are not efficient enough. We refer the interested reader to [FKSS22, Section 5] for the details on how to construct $\mathcal{T}$ and its associated containers more efficiently.

References

[AM02] D. Achlioptas and C. Moore. The asymptotic order of the random k-sat threshold. In *The 43rd Annual IEEE Symposium on Foundations of Computer Science, 2002. Proceedings.*, pages 779–788, 2002.

[Bal13] Deepak Bal. On Sharp Thresholds of Monotone Properties: Bourgain's Proof Revisited. *arXiv preprint arXiv:1302.1162*, 2013.

[Bec75] William Beckner. Inequalities in Fourier analysis. *Ann. of Math. (2)*, 102(1):159–182, 1975.

[BKM23] Amey Bhangale, Subhash Khot, and Dor Minzer. Effective bounds for restricted 3-arithmetic progressions in $\mathbb{F}_p^n$. *arXiv preprint arXiv:2308.06600*, 2023.

[BMS15] József Balogh, Robert Morris, and Wojciech Samotij. Independent sets in hypergraphs. *J. Amer. Math. Soc.*, 28(3):669–709, 2015.

[BMS18] József Balogh, Robert Morris, and Wojciech Samotij. The method of hypergraph containers. In *Proceedings of the International Congress of Mathematicians—Rio de Janeiro 2018. Vol. IV. Invited lectures*, pages 3059–3092. World Sci. Publ., Hackensack, NJ, 2018.

[Bol81] Béla Bollobás. Threshold functions for small subgraphs. *Math. Proc. Cambridge Philos. Soc.*, 90(2):197–206, 1981.

[Bon70] Aline Bonami. Étude des coefficients de Fourier des fonctions de $L^p(G)$. *Ann. Inst. Fourier (Grenoble)*, 20:335–402, 1970.

[Bou80] J. Bourgain. Walsh subspaces of L^p-product spaces. In *Seminar on Functional Analysis, 1979–1980 (French)*, pages Exp. No. 4A, 9. École Polytech., Palaiseau, 1980.

[BT87] B. Bollobás and A. Thomason. Threshold functions. *Combinatorica*, 7(1):35–38, 1987.

[EKL24] Shai Evra, Guy Kindler, and Noam Lifshitz. Polynomial Bogolyubov for special linear groups via tensor rank. *arXiv preprint arXiv:2404.00641*, 2024.

[EKLM24] David Ellis, Guy Kindler, Noam Lifshitz, and Dor Minzer. Product mixing in compact lie groups. *arXiv preprint arXiv:2401.15456*, 2024.

[ES83] Paul Erdős and Miklós Simonovits. Supersaturated graphs and hypergraphs. *Combinatorica*, 3(2):181–192, 1983.

[FKSS22] Ehud Friedgut, Eden Kuperwasser, Wojciech Samotij, and Mathias Schacht. Sharp thresholds for Ramsey properties. *arXiv preprint arXiv:2207.13982*, 2022.

[Fri98] Ehud Friedgut. Boolean functions with low average sensitivity depend on few coordinates. *Combinatorica*, 18(1):27–35, 1998.

[Fri99] Ehud Friedgut. Sharp thresholds of graph properties, and the k-sat problem. *J. Amer. Math. Soc.*, 12(4):1017–1054, 1999. With an appendix by Jean Bourgain.

[Fri05] Ehud Friedgut. Hunting for sharp thresholds. *Random Structures Algorithms*, 26(1-2):37–51, 2005.

[FRRT06] Ehud Friedgut, Vojtech Rödl, Andrzej Ruciński, and Prasad Tetali. A sharp threshold for random graphs with a monochromatic triangle in every edge coloring. *Mem. Amer. Math. Soc.*, 179(845):vi+66, 2006.

[Gro75] Leonard Gross. Logarithmic Sobolev inequalities. *Amer. J. Math.*, 97(4):1061–1083, 1975.

[Jan90] Svante Janson. Poisson approximation for large deviations. *Random Structures Algorithms*, 1(2):221–229, 1990.

[KLLM21] Peter Keevash, Noam Lifshitz, Eoin Long, and Dor Minzer. Global hypercontractivity and its applications. *arXiv preprint arXiv:2103.04604*, 2021.

[KLLM24] Peter Keevash, Noam Lifshitz, Eoin Long, and Dor Minzer. Hypercontractivity for global functions and sharp thresholds. *J. Amer. Math. Soc.*, 37(1):245–279, 2024.

[KLS23] Nathan Keller, Noam Lifshitz, and Ohad Sheinfeld. Improved covering results for conjugacy classes of symmetric groups via hypercontractivity. *arXiv preprint arXiv:2310.18107*, 2023.

[KM22] Tali Kaufman and Dor Minzer. Improved optimal testing results from global hypercontractivity. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science—FOCS 2022*, pages 98–109. IEEE Computer Soc., Los Alamitos, CA, [2022] ©2022.

[Lac22] Michael T Lacey. Threshold: Explaining Bourgain. 2022.

[Mar74] G. A. Margulis. Probabilistic characteristics of graphs with large connectivity. *Problemy Peredači Informacii*, 10(2):101–108, 1974.

[Min21] Dor Minzer. Lecture notes for Topics in Combinatorics: Analysis of Boolean Functions, February 2021.

[O'D14] Ryan O'Donnell. *Analysis of Boolean functions*. Cambridge University Press, New York, 2014.

[Rus82] Lucio Russo. An approximate zero-one law. *Z. Wahrsch. Verw. Gebiete*, 61(1):129–139, 1982.

[ST15] David Saxton and Andrew Thomason. Hypergraph containers. *Invent. Math.*, 201(3):925–992, 2015.

[Zak23] Dmitrii Zakharov. Spherical sets avoiding orthogonal bases. *arXiv preprint arXiv:2310.06821*, 2023.

[Zha21] Yu Zhao. *Generalizations and applications of hypercontractivity and small-set expansion*. PhD thesis, Columbia University, 2021.