Sharp Thresholds for Ramsey Properties

Henry Fleischmann

April 2024

Contents

1 Introduction

A classic question in extremal graph theory is: how large can a graph be before it necessarily has property A? For example, when property A is "contains a copy of the complete graph on k vertices, K_k," this is famously answered by Turán's Theorem.

Theorem 1.1 (Turán's Theorem, [Tur41])**.** *Every n-node graph that does not contain K_{r+1} as a subgraph has at most $\left(1-\frac{1}{r}\right)\cdot\frac{n^2}{2}$ edges.*

The complete r-partite graph with (almost) equal-sized parts shows that this bound is tight. A second important example is that of Ramsey properties: what is the smallest k such that any 2-coloring of the edges of K_k contains a monochromatic triangle?

These sorts of extremal questions remain well-motivated even in a non-deterministic setting. Let $G(n,p)$ be the Erdős-Rényi random graph on n nodes, where each edge is added independently with probability p [ER60]. One way to formulate such questions in this random setting is to ask the following: how large does p need to be (as a function of n) so that $G(n,p)$ has property A with high probability? In trying to rigorously analyze questions of this type, threshold phenomena naturally emerge. We define a sequence of probabilities $\hat{p}$ to be a *threshold* for a property A if, for any $p \ll \hat{p}$, the probability that $G(n,p) \in A$ tends to 0 and, for $p \gg \hat{p}$, the probability that $G(n,p) \in A$ tends to 1. A priori, it is not even immediately obvious which properties have thresholds. Certainly some properties do not: for example, the property defined by, for n even, containing an edge, and, for n odd, having at least $n^2/4$ total edges. But, imposing a little structure on the properties, an important theorem of Bollobás and Thomason [BT87] implies that all nontrivial monotone graph properties have a threshold. Here a property A is *monotone* if $H \in A$ and $H \subseteq G$, then $G \in A$, and it is *nontrivial* if it is neither the empty set nor all graphs. The properties "contains a copy of K_k" and "any 2-coloring of the edges contains a monochromatic copy of H," for any fixed graph H, are both nontrivial and monotone and hence have thresholds.

However, Bollobás and Thomason's theorem is purely existential. As such, considerable efforts were made to actually identify the threshold functions of important monotone properties in subsequent work. For example, in the sequence of papers [RR93, RR94, RR95], Rödl and Ruciński completely determined the threshold for being r-Ramsey for any fixed graph H. (That is, the property "any r-coloring of the edges contains a monochromatic copy of H.")

While these lines of work resolved some of the questions about the existence and locations of thresholds, it is also interesting to consider their more granular properties. To motivate this, we revisit the original paper of Erdős and Rényi introducing $G(n,p)$ [ER60]. In that paper, they showed that $p = n/\log n$ is a threshold for $G(n,p)$ being connected. More specifically, they showed that, for all $1 > \varepsilon > 0$, if $p = (1+\varepsilon)n/\log n$ then $G(n,p)$ is asymptotically almost surely connected and, if $p = (1-\varepsilon)n/\log n$, $G(n,p)$ is asymptotically almost surely *not* connected. That is, the threshold for connectivity is *sharp*; perturbing the threshold probability by a constant factor toggles the probability of obtaining the property from 0 to 1. This is in contrast with properties with a *coarse* threshold, such as containing a triangle.

Although it is relatively straightforward to prove sharpness of the thresholds for certain properties—connectivity, for example—for many years proving sharp thresholds for Ramsey properties seemed entirely out of reach. That is, until Friedgut gave an immensely powerful characterization of all properties with a coarse threshold [FB+99]. Soon after, in a breakthrough paper, Friedgut et al. managed to combine Friedgut's criterion with a sparse generalization of the Szemerédi regularity lemma to prove that being 2-Ramsey for a triangle has a sharp threshold [FRRT03].

More than 10 years later, sharp thresholds were proved for a variety of other Ramsey properties (see for example [SS18]), leveraging the strength of the hypergraph container lemma [BMS18, ST15]. Recently, Friedgut et al. [FKSS22] established a unifying framework for proving sharp thresholds for a multitude of these properties using the hypergraph container lemma, including the original result of [FRRT03].

The goal of this exposition is to give a largely self-contained proof that the property of being 2-Ramsey for a triangle has a sharp threshold (Theorem 6.1). Along the way, we introduce and apply a plethora of broadly applicable powerful tools. This includes:

- Friedgut's criterion for symmetric monotone properties with a coarse threshold (Theorem 3.2), along with its more practical dichotomy corollary (Theorem 3.16).

- Fourier analytic methods (introduced in Definition 3.1 and Proposition 3.3).
- The hypergraph container lemma (Theorem 4.3).
- Janson's inequality (Theorem 2.1).

The proof of Friedgut's criterion makes heavy use of Fourier analysis and relies crucially on the idea of so-called "modest" graphs, notions developed in Sections 3.1 and 3.2, respectively. The core of the proof of Friedgut's criterion, which we sketch in Section 3.3, involves a series of approximations of the characteristic function of a monotone symmetric property with a coarse threshold, with the final approximation being the characteristic function of an approximating property. Section 3.4 focuses on reframing the result in preparation for applying it in the proof of Theorem 6.1.

Section 4 is entirely devoted to the hypergraph container lemma (Theorem 4.3). Besides introducing the container lemma and breaking down all of its moving parts, we give a short proof of the upper bound on thresholds for the property of being Ramsey for a triangle as a warmup application.

In Section 5, we analyze the hypergraph of triangles in K_n in considerable detail, developing some of its key properties used in the proof of Theorem 6.1 in the subsequent section.

Finally, in Section 6, we combine all of these tools to give a proof that the property of being 2-Ramsey for a triangle has a sharp threshold. The starting point is the dichotomy corollary of Friedgut's criterion, Theorem 3.16, and we spend Section 6.1 introducing notation to properly leverage this result. We also derive several simple, useful consequences. In Section 6.2 we carefully construct a hypergraph and apply the container lemma to said hypergraph to facilitate a more efficient union bound over colorings of our random graph. The purpose of Section 6.3 is then to effectively take advantage of this union bound, again making use of the container lemma. We complete the proof in Section 6.4, leveraging the power of Janson's inequality.

2 Preliminaries

In this section we formally establish some of the notation, definitions, and tools we will use throughout.

Asymptotic Notation. Let $f, g : \mathbb{N} \to \mathbb{R}_{>0}$. If there exists some constant $c > 0$ and $n_0 \in \mathbb{N}$ large enough such that, for all $n \geq n_0$, $f(n) \leq c \cdot g(n)$, then we write $f = O(g)$ and $g = \Omega(f)$. If $f = O(g)$ and $f = \Omega(g)$, then $f = \Theta(g)$. We will often write $f = O(1)$ and $f = \Omega(1)$, where 1 here means the constant function 1. If

$$\lim_{n\to\infty} f(n)/g(n) = 0,$$

then $f = o(g)$ and $g = \omega(f)$.

Graphs. Let G be a (hyper)graph. We use $V(G)$ to denote its vertex set, $E(G)$ to denote its edge set, $v(G)$ to denote the number of vertices in G, and $e(G)$ or $|G|$ to denote the number of edges in G.

Random graphs. Let n be an integer and $p \in [0, 1]$. All of our random graphs will be generated via the Erdős-Rényi random graph model $G(n, p)$. Namely, each edge of the complete graph on n vertices is added independently with probability p. Technically $G(n, p)$ is a product probability space, but we will often abuse notation to consider it as a graph for convenience. When we do treat $G(n, p)$ as a probability space, $A \subseteq G(n, p)$ denotes a subset of the edge subgraphs of K_n, μ_p is the product measure on the space, and $\mu_p(A)$ is the measure of A. Alternatively, $\mu_p(A)$ is the probability that $G(n, p)$ is exactly one of the edge subgraphs in A.

Graph properties. A *(graph) property* is a subset of

$$\bigcup_{n=2}^{\infty} \mathcal{P}(E(K_n)).$$

A property A can then also be viewed as a sequence of sets (or function of n) $A_2, A_3, \ldots$ where $A_n = A \cap \mathcal{P}(E(K_n))$, the graphs G with property A and $v(G) = n$. As stated in the introduction, a property is *monotone* if $H \in A$ and $H \subseteq G$ implies $G \in A$. A property is *nontrivial* if it is neither the empty set nor all graphs.

A sequence of probabilities $p = p(n)$ is a *threshold* for a property A if, for $p' = o(p)$,

$$\lim_{n\to\infty} \Pr[G(n, p') \in A] = 0$$

and, for $p'' = \omega(p)$,

$$\lim_{n\to\infty} \Pr[G(n, p'') \in A] = 1.$$

By a celebrated theorem of Bollobás and Thomason [BT87], all nontrivial, monotone properties have a threshold. Such properties can be further categorized by their probabilistic behavior near the threshold probability. We say a property A has a *sharp* threshold (or is *sharp*) if, for p a threshold for A, for all $1 > \varepsilon > 0$,

$$\Pr[G(n, (1 - \varepsilon)\hat{p}) \in A] = o(1)$$

and

$$\Pr[G(n, (1 + \varepsilon)\hat{p}) \in A] = 1 - o(1).$$

Otherwise, we say the property A has a *coarse* threshold (or is *coarse*).

Most natural properties are also *symmetric*. That is, if $G \in A$ and $v(G) = n$, then, for any automorphism π of the complete graph K_n, $\pi(G) \in A$. Stated differently, the property is independent of the choice of labellings of the vertices in how we generate graphs via the $G(n, p)$ model.

Janson's Inequality. Janson's inequality gives a strong upper bound on the tail probabilities of random variables counting the number of sets from a given collection are contained in a binomially generated random subset (e.g., $G(n, p)$). Let $(A_1, \ldots A_k)$ be a sequence of (not necessarily distinct) events. Let $X = \sum_{i=1}^{k} 1_{A_i}$. Then,

$$\begin{aligned}\mathrm{Var}(X) &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2 \\ &\leq \sum_{A_i \sim A_j} \mathbb{E}[1_{A_i} 1_{A_j}] \\ &= \sum_{A_i \sim A_j} \Pr[A_i \cap A_j],\end{aligned}$$

where $A_i \sim A_j$ if the events A_i and A_j are dependent (and the sum over $A_i \sim A_j$ varies over $i, j \in [k]$, including when $i = j$). The upper-bound on $\mathrm{Var}(X)$

$$\mathrm{Var}'(A_1, \ldots A_k) := \sum_{A_i \sim A_j} \Pr[A_i \cap A_j]$$

is the *pseudo-variance* of $(A_1, \ldots A_k)$. We can now state Janson's inequality.

Theorem 2.1 (Janson's inequality [Jan90])**.** *Let Ω be a finite set, and let $B_1, \ldots B_k$ be a sequence of (not necessarily distinct) subsets of Ω. Let $p \in [0, 1]$ and let $R \subset \Omega$ be generated by keeping each element of Ω independently with probability p. For each $i \in [k]$, let A_i be the event that $B_i \subseteq R$. Let $X = \sum_{i=1}^{k} 1_{A_i}$. Then, for all $0 \leq t \leq \mathbb{E}[X]$,*

$$\Pr[X \leq \mathbb{E}[X] - t] \leq \exp\left(\frac{-t^2}{2\,\mathrm{Var}'(A_1, \ldots, A_k)}\right).$$

3 Friedgut's Criterion for Properties with a Coarse Threshold

We begin by motivating and contextualizing Friedgut's criterion. One rather simple graph property is that of containing a triangle. It is not hard to show that the property is coarse; intuitively this is a result of the limited dependence between different triangles in K_n. Similarly "local" properties, e.g., properties defined by containing a member of a finite list of bounded subgraphs, have a similarly bounded dependency between occurrences of those bounded subgraphs in K_n. It turns out that properties defined in such a way always have coarse thresholds. We can prove this by a coupling argument.

Lemma 3.1 (Lemma 4.6 of [FKSS22])**.** *Let $K > 0$ be a constant. Let $H_1, H_2, \dots H_\ell$ be graphs with $|H_i| \leq K$ for all $i \in [\ell]$, and let A be the (monotone symmetric) graph property defined by $G \in A$ if and only if $H_i \subseteq G$ for some $i \in [\ell]$. Then, A has a coarse threshold.*

Proof. Let $\mu_p(A) = \Pr[G(n,p) \in A]$, as usual. Now, fix $c \in (0,1)$ and consider $\mu_{cp}(A)$. By viewing the random graph $G(n,cp)$ as a random subgraph of a random graph drawn from $G(n,p)$ (with each edge kept independently with probability c), it is clear that $\mu_{cp}(A) \geq c^K \cdot \mu_p(A)$. This is because for each random graph $G(n,p) \in A$, the coupled random subgraph $G(n,cp)$ still contains an H_i subgraph with probability at least c^K. Hence, letting p be a threshold probability for A (e.g., such that $\mu_p(A) = 1/2$), the result follows. □

With the sufficient condition for coarse thresholds given by Lemma 3.1 in mind, consider the following:

Conjecture 3.1 (Ludicrously optimistic conjecture)**.** Let A be a monotone symmetric graph property with a coarse threshold. Then, there is a finite list of bounded subgraphs $H_1, H_2, \dots, H_\ell$ such that $G \in A$ if and only if $H_i \subset G$ for some $i \in [k]$.

That is, all monotone symmetric graph properties with coarse thresholds can be characterized locally by checking containment of bounded size subgraphs. Unforunately, Conjecture 3.1 is false.

Example 3.1 (Combining properties with coarse and sharp thresholds)**.** A simple idea for building a counterexample to Conjecture 3.1 is to take the intersection of two monotone symmetric graph properties A_1 and A_2, with A_1 having a sharp threshold and A_2 having a coarse threshold. If the the threshold for A_1 is lower than the threshold for A_2, then the monotone symmetric property determined by $A_1 \cap A_2$ will have a coarse threshold at the threshold of A_2. This is because, in order to have property A_2, we must be at the threshold for A_2, and, around that threshold, we will have property A_1 with probability $1 - o(1)$. Crucially, if A_1 is a property not determined by containing bounded size subgraphs, then $A_1 \cap A_2$ will be a property with a coarse threshold not determined by containing bounded size subgraphs.

Concretely, let A_1 be the property of being connected and A_2 tbe property of containing a copy of K_4. As mentioned in the introduction, the former famously has a sharp threshold at $p = \frac{\log n}{n}$ [ER60]. The latter has a coarse threshold by Lemma 3.1, and, since there are $\Theta(n^4)$ copies of K_4 in K_n and each is added with probability p^6, by Markov's inequality, the threshold probability for containing a K_4 must be $\Omega(n^{-2/3})$. Hence, the property of being connected and containing a K_4 has a coarse threshold but is not determined by containing a subgraph from a list of bounded size subgraphs. This disproves Conjecture 3.1.

Surprisingly however, Conjecture 3.1 is not far from the truth. Let A be a monotone symmetric property. Suppose that A has a coarse threshold. Let $p = p(n)$ be a threshold probability for A so that $\mu_p(A)$ is bounded above and below by a constant. To understand A, it would be enough to understand another monotone symmetric graph family B such that $\mu_p(A \Delta B)$ is arbitrarily small, where $A \Delta B$ is the symmetric difference of the sets A and B. Then, up to rare events, checking membership in B suffices for checking membership in A. Excitingly, by adopting this more flexible view, our dream from Conjecture 3.1 is now a reality:

Theorem 3.2 (Friedgut's Coarse Threshold Criterion [FB$^+$99])**.** *Fix $\varepsilon > 0$. There exists a function $k(\varepsilon, c)$ such that, for any n and any monotone symmetric family of graphs A on n vertices with a threshold probability p such that $p \cdot \mu'_p(A) \leq c$, there exists another monotone symmetric family of graphs B such that:*

1. *The maximum size of a minimal subgraph in B is bounded by $k(\varepsilon, c)$, and*
2. *$\mu_p(A \Delta B) \leq \varepsilon$.*

In the statement of Theorem 3.2, $\mu'_p(A)$ is the derivative of the function $\mu_p(A)$ with respect to the parameter p. A minimal subgraph in B is one not containing any other members of B as a strict edge subgraph. Before we strive to understand why Theorem 3.2 is true, we consider why it achieves our dream. First of all, why does this say anything at all about graph properties with coarse thresholds?

Remark 3.1. Let $A = A(n)$ be a monotone graph family with a coarse threshold. Then, there is some threshold probability $\hat{p}$ and some constants $1 > c, \rho > 0$ such that $\rho < \mu_{c\hat{p}}(A) < \mu_{\hat{p}}(A) < 1 - \rho$.

Now, consider the function $\mu_p(A)$ on the interval $[c\hat{p}, \hat{p}]$. The average slope on the interval is at most $(1 - 2\rho)/((1-c)\hat{p})$. Hence, there exists some threshold probability p in the interval $[c\hat{p}, \hat{p}]$ such that $p \cdot \mu'_p(A) \leq$

$(1-2\rho)/(1-c)$. That is, by the coarseness assumption on A, there is some threshold probability p satisfying that $p \cdot \mu_p'(A)$ is bounded by a constant. We will implicitly consider such a p throughout our discussion of Friedgut's criterion. See Figure 1.

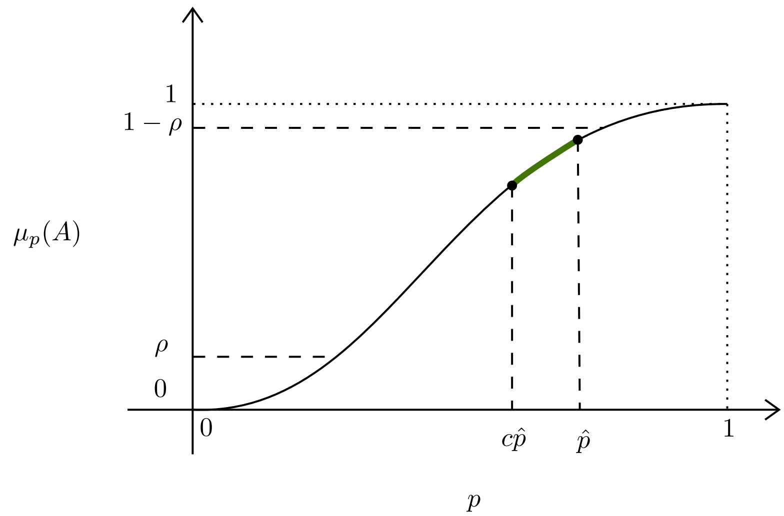

Figure 1: Plot of p versus $\mu_p(A)$ from Remark 3.1.

Now, for a fixed property A, ε and c are constants, so that all minimal subgraphs in B are a subset of all graphs on at most $k(\varepsilon, c)$ edges. Up to isomorphism, this is a constant number of graphs. Monotonicity then means then that there is some finite sublist $H_1, H_2, \ldots, H_\ell$ of these graphs on at most $k(\varepsilon, c)$ edges such that $G \in B$ if and only if $H_i \subset G$ for some $i \in [\ell]$. That is, B is one of the simple monotone graph families described in Conjecture 3.1!

3.1 Boolean Fourier analysis

The proof of Friedgut's criterion makes heavy use of Boolean Fourier analysis. In this section, we introduce some basic definitions and consequences of those definitions. Consider the vector space of functions from $\{0,1\}^{\binom{n}{2}} \to \mathbb{R}$, viewing these functions as functions from all subgraphs of K_n to $\mathbb{R}$. The most important function we will consider from this vector space is the *characteristic function* of a set. Namely, for $A \subset \{0,1\}^{\binom{n}{2}}$, its characteristic function 1_A satisfies:

$$1_A(H) = \begin{cases} 1 & \text{if } H \in A, \\ 0 & \text{if } H \notin A. \end{cases}$$

A simple basis for this vector space is the standard one composed of the functions that are 1 on a single element of $\{0,1\}^{\binom{n}{2}} \to \mathbb{R}$ and 0 on the rest; the existence of this basis shows that this space has dimension $2^{\binom{n}{2}}$.

However, it turns out that there is a much more convenient basis to consider in the context of properties of $G(n,p)$ and the corresponding probability measure μ_p. This is the *Fourier basis*, attributed in this setting to Talagrand [Tal94]. For convenience, we define $q := 1-p$.

Definition 3.1 (Fourier basis)**.** Let $U_\emptyset$ be the all-ones function. Then, for $e \in E(K_n)$, define

$$U_e(H) = \begin{cases} -\sqrt{q/p} & \text{if } e \in H; \\ \sqrt{p/q} & \text{if } e \notin H. \end{cases}$$

For all other edge subgraphs $R \subseteq K_n$, define

$$U_R(H) = \prod_{e \in R} U_e(H).$$

The collection of functions U_R, for $R \subseteq K_n$, form the *Fourier basis* of $\{0,1\}^{\binom{n}{2}} \to \mathbb{R}$ with respect to the natural inner product on the measure space $G(n,p)$. That is, the inner product $\langle \cdot, \cdot \rangle$ defined by, for $f, g : G(n,p) \to \mathbb{R}$,

$$\langle f, g \rangle = \sum_{H \subseteq K_n} f(H) g(H) \mu_p(H).$$

Proposition 3.3. *The family of functions U_R, for $R \subseteq K_n$, form an orthonormal basis of the space of functions from $G(n,p) \to \mathbb{R}$.*

Proof. First, note that for all $e \in K_n$, U_e satsifies $\mathbb{E}_{H \sim G(n,p)}[U_e(H)] = 0$ and $\mathbb{E}_{H \sim G(n,p)}[U_e(H)^2] = 1$. Now, let $R, S \subseteq K_n$. Then,

$$\begin{aligned} \langle U_R, U_S \rangle &= \mathbb{E}_{H \sim G(n,p)}[U_R(H) \cdot U_S(H)] \\ &= \mathbb{E}_{H \sim G(n,p)} \left[\prod_{e \in R} U_e(H) \cdot \prod_{e \in S} U_e(H) \right]. \end{aligned}$$

Then, by independence of the edges in $G(n,p)$, we have $\langle U_R, U_S \rangle = 1$ if $R = S$ and 0 otherwise. □

Now, for $f : G(n,p) \to \mathbb{R}$, its *Fourier coefficients* are just the coefficients of its Fourier basis vectors when expressed in terms of the Fourier basis:

$$\hat{f}(H) := \langle f, U_H \rangle,$$

for $H \subseteq K_n$. Then,

$$f = \sum_{H \subset K_n} \hat{f}(H) U_H.$$

We can define L^r-norms with respect to the measure μ_p. That is,

$$\|f\|_r^r := \mathbb{E}_{H \sim G(n,p)}[|f(H)|^r] = \sum_{H \subset K_n} |f(H)|^r \mu_p(H).$$

Then, the following is a simple consequence of Proposition 3.3.

Proposition 3.4 (Parseval's Identity)**.** *Let $f : G(n,p) \to \mathbb{R}$. Then,*

$$\|f\|_2^2 = \mathbb{E}_{H \sim G(n,p)}[f(H)^2] = \sum_{H \subseteq K_n} \hat{f}^2(H).$$

The following remark is specifically relevant to our discussion of Theorem 3.2.

Remark 3.2 (Bound on the Fourier coefficients of a characteristic function)**.** Let A be the restriction of a symmetric monotone graph property onto edge subgraphs of K_n. Note in particular that

$$\mu_p(A) = \|1_A\|_2^2 = \sum_{S \subseteq K_n} \widehat{1_A}(S)^2.$$

Hence, for each edge subgraph $S \subseteq K_n$, $\widehat{1_A}(S)^2 \leq \mu_p(A) \leq 1$, so $|\widehat{1_A}(S)| \leq 1$. Define $\mathcal{O}(S)$ to be the orbit of S under automorphisms of K_n. Then, by symmetry of 1_A, using that $\widehat{1_A}(S) = \widehat{1_A}(S')$ for $S' \in \mathcal{O}(S)$, we also have

$$|\widehat{1_A}(S)| \leq 1/\sqrt{|\mathcal{O}(S)|}.$$

We will make repeated use of the strategy of treating the orbit of $S \subseteq K_n$ all at once and applying the symmetry assumption. For this reason, we introduce one more piece of notation.

Definition 3.2. Let $V_S := \sum_{H \in \mathcal{O}(S)} U_H$, that is, the sum over all Fourier basis vectors corresponding to subgraphs of K_n isomorphic to S.

To reiterate, for a monotone symmetric property A, by symmetry, $\langle 1_A, U_S \rangle = \langle 1_A, U_H \rangle$ for all $H \in \mathcal{O}(S)$. As such, we can decompose the Fourier expansion of 1_A in terms of the V_S's instead of U_H's.

3.2 Modest graphs

At the heart of Friedgut's proof of Theorem 3.2 is the notion of *modest* graphs. To properly define them, we first need a few definitions.

Definition 3.3 (Balanced graphs)**.** Let G be a graph. Then, G is *balanced* if, for all nonempty subgraphs $H \subseteq G$, the average degree of H is at most the average degree of G. If this inequality is strict for all nontrivial subgraphs, we say that G is *strictly balanced*.

Example 3.2 (Balanced and unbalanced graphs)**.** All regular graphs are balanced. This is because the maximum average degree of a subgraph of a k-regular graph H is k, and H itself has average degree k. In contrast, every graph H with a node of degree strictly less than the average degree is unbalanced, since removing that node increases the average degree.

Closely related to the notion of balanced graphs is the expected number of occurrences of the graph in $G(n,p)$.

Definition 3.4 (Expected number of occurrences of a subgraph)**.** Let $G \subseteq K_n$. Define

$$\mathbb{E}[G] := \mathbb{E}[|\{G' \in \mathcal{O}(G) \,:\, G' \subseteq G(n,p)\}|].$$

We are now ready to define modest graphs.

Definition 3.5 (Modest graphs)**.** Given constants $c_1, c_2, L > 0$, call a graph G (c_1, c_2, L)-*modest* if $c_1 \leq \mathbb{E}[G] \leq c_2$, $|G| \leq L$, and G is balanced. When the choice of the constants is apparent from context, we simply call G a *modest* graph.

The graph property B we will build approximating the property A will have small unions of modest graphs as its minimal graphs. Since L is a constant, given a choice of c_1, c_2, and L, there are a constant number of total modest graphs up to isomorphism. We can enumerate them: $S_1, \ldots, S_\ell$.

Remark 3.3. In general, given a graph H, $\mathbb{E}[H] = |\mathcal{O}(H)|p^{|H|}$. But, assuming $p = o(1)$, if $|H|$ is a constant, then $\mathbb{E}[H] = \Theta(n^{v(H)}p^{|H|})$. Suppose $p(n) = n^{-1/\delta}$ and we wish to have $\mathbb{E}[H] = \Theta(1)$. Then, we must have $v(H) = |H|/\delta$, i.e., H must be average degree 2δ for n large enough. Hence, asymptotically, our constraints on modest graphs determine their average degree too.

Moreover, if $\mathbb{E}[H] = O(\log^k n)$ for some k, any graph G with $|G|$ bounded by a constant having $\mathbb{E}[G] = \Theta(1)$ implies that $\mathbb{E}[H] = O(1)$. This is because G being of constant size and $\mathbb{E}[G] = \Theta(1)$ means $p = \Theta(n^{-\frac{v(G)}{|G|}})$.

Why are modest graphs the right object to consider? Consider the Fourier expansion of the characteristic function $f := 1_A$:

$$f = \sum_H \hat{f}(H) U_H.$$

We will construct the monotone property B approximating A by devising a series of approximations of f, with the last approximation being the characteristic function of B. The first step in approximating f will be to drop the terms in its Fourier expansion which do not correspond to modest graphs, applying Lemma 3.5.

Lemma 3.5 (Lemma 4.2 of [FB+99])**.** *Let A be a monotone symmetric family of graphs of n vertices, and let $f = 1_A$ be its characteristic function. Suppose that $p \cdot \mu_p'(A) \leq c$, for some constant c. Then, for all $\varepsilon > 0$, there exists c_1, c_2, L and sufficiently large n such that*

$$\sum_{H \text{ not } (c_1,c_2,L)\text{-modest}} \hat{f}^2(H) \leq \varepsilon.$$

Namely, by Parseval's identity, this means the total contribution of non-modest graphs to $\|f\|_1 = \|f\|_2^2$ can be made arbitrarily small. Several important ingredients go into proving Lemma 3.5; these are the basis for each property of modest graphs.

Bounding the size of modest graphs. The fact that we can bound the size of modest graphs is a consequence of a more general lemma:

Lemma 3.6 (Lemma 2.3 of [FB+99], [Tal94])**.** $q \cdot p \cdot \mu_p'(A) = \sum_H \hat{f}^2(H)|H|$.

As an immediate corollary, we have:

Corollary 3.7. $\sum_{H:|H|\geq L} \hat{f}^2(H) \leq \frac{q \cdot p \cdot \mu_p'(A)}{L}$.

Crucially, since A is a monotone graph property with a coarse threshold, we will have $p \cdot q \cdot \mu_p'(A)$ bounded by a constant. Then, we can choose an appropriately large constant bound L on the sizes of our graph to ensure that the remaining graphs have arbitrarily low contribution to $\|f\|_2^2$.

Lower bounding the expectation of modest graphs. In order to lower bound the expectation of modest graphs, we simply need to show that the total contribution of graphs with sufficiently low expectation is negligible. Lemma 3.8 shows the total contribution of the orbit of a graph to the Fourier expansion of f can be bounded by a multiple of the expected number of occurrences of that graph.

Lemma 3.8. *Let G be a graph with $|G| \leq L$. Then, for sufficiently large n and $p = o(1)$,*

$$\sum_{S \in \mathcal{O}(G)} \hat{f}^2(S) \leq 2^{2L} \cdot \mathbb{E}[G].$$

Proof. Let $S \in \mathcal{O}(G)$. We first bound $\hat{f}^2(S) = \langle f, U_S \rangle^2$. We have

$$\begin{aligned} |\hat{f}(S)| &\leq \sum_{H \subseteq S} \left(\frac{q}{p}\right)^{|H|-|S|/2} p^{|H|} q^{|S|-|H|} \\ &\leq p^{|S|/2} q^{|S|/2} 2^{|S|} \leq 2^L p^{|G|/2}, \end{aligned}$$

So, $\hat{f}(S)^2 \leq 2^{2L} \cdot p^{|G|}$. Now,

$$\sum_{S \in \mathcal{O}(G)} \hat{f}^2(S) \leq |\mathcal{O}(G)| \cdot 2^{2L} \cdot p^{|G|} = 2^{2L} \cdot \mathbb{E}[G].$$

□

Since from the previous section we can bound the size of the graphs we consider by a constant, there are only a constant number of possible non-isomorphic graphs. Hence, we can set a constant lower bound threshold on the expected number of occurrences of a graph so that ignoring the contribution of all graphs of bounded size with lower expected number of occurrences has a negligible effect on our approximation of f.

The balanced condition of modest graphs. The balanced condition of modest graphs is a consequence of the constant upper bound of the expected number of occurrences of modest graphs. This comes from Lemma 3.9. For ease of notation, we establish the convention that the expectation of the empty graph is 1.

Lemma 3.9 (Lemma 4.4 of [FB+99])**.** *Let H be a graph with $|H| \leq L$ for some constant L. Then, for some some $c > 0$, for every strict edge subgraph $H' \subsetneq H$,*

$$\sum_{S \in \mathcal{O}(H)} \hat{f}^2(S) \leq c \cdot \max_{R \subsetneq H'} \{\mathbb{E}[H']/\mathbb{E}[R]\}.$$

Now, suppose that $\mathbb{E}[H] \leq c_2$ for some constant c_2. Suppose additionally that H is not balanced. Namely, then there exists a minimal subgraph $H' \subsetneq H$ with the average degree of H' strictly greater than the average degree of H and strictly greater than all strict subgraphs of H' (this second condition means H' is strictly balanced). But then $\mathbb{E}[H'] = o(1)$ and $\mathbb{E}[H'] = o(\mathbb{E}[H''])$ for all $H'' \subsetneq H'$. So, by Lemma 3.9, the total contribution of $\mathcal{O}(H)$ to the Fourier expansion of f is trivial.

Remark 3.4. We will not prove Lemma 3.9, but the intuition is as follows. The order of the expected number of copies of a fixed graph H appearing in $G(n,p)$ depends only on its average degree. Nonetheless, higher average degree graphs are rarer, so, intuitively, the distribution of unbalanced graphs is skewed by the rarity of their denser subgraphs. Indeed, if H appears in $G(n,p)$, all of its lower expectation subgraphs do as well. This skewed distribution is realized in the smaller contribution of the Fourier coefficients of the orbit of bounded unbalanced graphs to $\|f\|_2^2$.

Upper bounding the expectation of modest graphs. The last step in proving Lemma 3.5 is then to upper bound the expected number of occurrences of graphs whose orbit has nontrivial contribution to the Fourier expansion of f.

Lemma 3.10. *Suppose that $H \subseteq K_n$ such that*

$$\sum_{R \in \mathcal{O}(H)} \hat{f}(R)^2 = \Omega(1),$$

and $|H| \leq L$, for some constant L. Then, $\mathbb{E}[H] = O(1)$.

Again the proof is rather technical, so we only provide a sketch of the argument.

Proof sketch. Let $\tilde{H}$ be the subgraph of H with $\mathbb{E}[\tilde{H}]$ minimal and suppose

$$\sum_{R \in \mathcal{O}(H)} \hat{f}(R)^2 \geq C_1.$$

By a fourth moment estimate it is possible to bound $\mathbb{E}[\tilde{H}]$ above by a constant depending on C_1.

There are then three cases. If $\tilde{H}$ is a single edge, since it has bounded expectation, $p \sim C_2/n^2$ for some constant C_2. All bounded H have $\mathbb{E}[H] = O(1)$ for this choice of p. When $\tilde{H} = H$, we are done by the fourth moment estimate.

The third and final case is when $\tilde{H} \subsetneq H$, but $\tilde{H}$ is not a single edge. By Lemma 3.9, we know that $\tilde{H}$ must be balanced. We can further assume $\tilde{H}$ is strictly balanced by shifting our attention to a minimal subgraph of $\tilde{H}$ of the same average degree (this only changes the expectation by a constant). Hence, we must have $\mathbb{E}[\tilde{H}] = \Omega(1)$ by Lemma 3.8.

The key insight is that, by fixing a copy of $\tilde{H}$ and considering the probability space with those edges fixed, it is possible to iterate this argument on the graph $H \setminus \tilde{H}$. Namely, the minimal expectation subgraph of $H \setminus \tilde{H}$ in this new probability space will have constant bounded expected number of occurrences. Repeating only at most a constant number of times (since $|H| \leq L$), we will get that H is a disjoint union of graphs with a constant bounded expected number of occurrences. This then implies $\mathbb{E}[H] = O(1)$, as desired. □

3.3 Approximating the characteristic function

To prove Theorem 3.2, we prove a sequence of approximations of f and ultimately round f to another function g, the characteristic function of a monotone symmetric property B approximating A. These approximations will take advantage of the notion of modest graphs. While we will state all of the main lemmas required, we will blackbox some of their proofs.

Recall that we may enumerate all modest graphs $S_1, S_2, \ldots, S_\ell$, for some appropriately chosen constant parameters c_1, c_2, and L. For any graph S, define C_S to be the collection of graphs G such that the union of all modest subgraphs contained in G is S. See Figure 3.3 for an example. Note that the C_S partition the subgraphs of K_n. Also note that, if $G \in C_S$, then $S \subseteq G$. We will see that this partition has several other convenient properties after applying a few initial approximations to f (see Lemma 3.11).

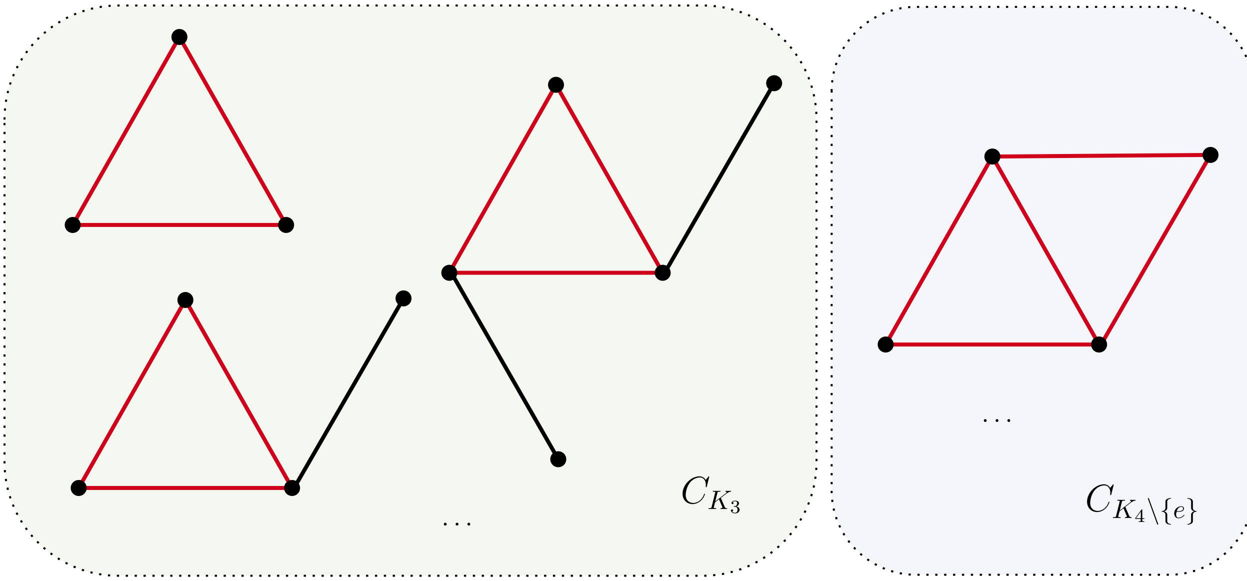

Figure 2: Examples of C_{K_3} and $C_{K_4\setminus\{e\}}$ for modest graphs being 2-regular. Edges of modest subgraphs are colored in red.

Restricting to Fourier coefficients corresponding to modest graphs. Our first appproximation of f, g_1, is the restriction of the Fourier expansion of f to the terms corresponding to modest graphs. Namely,

$$g_1 = \sum_{S \text{ modest}} \hat{f}(S) V_S,$$

with modesty defined for an implicit, appropriate choice of constants c_1, c_2, and L and V_S as defined in Definition 3.2. In particular, by Lemma 3.5, for any $\varepsilon > 0$, there is a choice of c_1, c_2, and L such that

$$\|f - g_1\|_2^2 = \sum_{S \text{ immodest}} \hat{f}^2(S) \leq \varepsilon.$$

Restricting to graphs containing a bounded number of modest graphs. A crucial property of a modest graph S is that $\mathbb{E}[S] \leq c_2$, where c_2 is a constant. By Markov's inequality, having more than $\ell c_2/\varepsilon$ copies of any particular modest graph S_i happens with probability at most ε/ℓ. By the union bound, the probability that more than $\ell c_2/\varepsilon$ copies of any of the ℓ modest graphs appear is then at most ε.

Now, consider S such that S is the union of modest graphs and S contains more than $\ell c_2/\varepsilon$ copies of some modest graph S_i. The appearance of any graph in C_S is a certificate for the rare event of more than $\ell c_2/\varepsilon$ copies of S_i occurring. Hence, the total probability measure of all such C_S must be at most ε. Let the union of all such C_S be $\mathcal{C}_1$. Define our second approximation of f, g_2, such that

$$g_2(S) = \begin{cases} g_1(S) & \text{if } S \notin \mathcal{C}_1; \\ 1 & \text{otherwise.} \end{cases}$$

Namely,

$$\|f - g_2\|_2^2 \leq \sum_{S \in \mathcal{C}_1} \mu_p(S) \cdot 1 + \|f - g_1\|_2^2 \leq 2\varepsilon.$$

Removing graph classes of low measure. Consider the graphs remaining after removing the graphs in $\mathcal{C}_1$ from the previous step. Note that these are covered by a finite number of parts M from the partition induced by the C_S's since the size of the union of the modest graphs contained in them is bounded from the previous step. Hence, we can define g_3 to be g_2 except we set g_3 to be 1 on graphs $H \in C_S$ such that $\mu_p(C_S) \leq \varepsilon/M$. Denote the union of C_S such that $\mu_p(C_S) \leq \varepsilon/M$ as $\mathcal{C}_2$. Namely, then

$$\|g_3 - f\|_2^2 \leq \sum_{S \in \mathcal{C}_2} \mu_p(S) \cdot 1 + \|f - g_2\|^2 \leq 3\varepsilon.$$

Let $\mathcal{C}$ be the union of the remaining C_S not contained in $\mathcal{C}_1$ or $\mathcal{C}_2$. At this point, we observe several other useful properties of partitioning the subgraphs of K_n by the union of the modest graphs contained in them.

Lemma 3.11. *Let $C_S \subseteq \mathcal{C}$ such that $C_S \neq \emptyset$ (i.e., S can be formed as a union of modest graphs). Then,*

1. *S has the same average degree as the modest graphs.*
2. *S is balanced.*
3. *There exist constants c_1' and c_2' such that, for all such S, $c_1' \leq \mathbb{E}[S] \leq c_2'$.*

Proof. Since S is an edgewise union of modest graphs, it has average degree at least the average degree of the modest graphs, δ. If S has average degree strictly greater than δ, then $\mathbb{E}[S] = o(1)$ since the size of S is bounded by a constant (see Remark 3.3). But, then $\mu_p(C_S) = o(1)$ by Markov's inequality, so we remove C_S when removing $\mathcal{C}_2$.

Then, since S has average degree δ, S must be a vertex disjoint union of modest graphs. Since modest graphs are balanced and any subgraph of S is then a vertex disjoint union of subgraphs of modest graphs, the average degree of any subgraph of S is at most δ (and S is balanced).

The final property is an immediate consequence of S having average degree δ and the number of possible S being bounded a constant (M) as a result of $C_S \subseteq \mathcal{C}$ and $C_S \neq \emptyset$. ☐

Uniformizing the approximation. The next modification uniformizes the approximation of f in preparation for rounding it to a characteristic function of a symmetric graph property. For $C_S \subseteq \mathcal{C}$ with $C_S \neq \emptyset$, for all $H \in C_S$, set

$$g_4(H) = \mathbb{E}[g_3 | C_S].$$

For $H \notin \mathcal{C}$, set $g_4(H) = g_3(H) = 1$. It turns out that this uniformization does not substantially weaken the approximation of f:

Lemma 3.12 (Lemma 4.14 of [FB+99]). *For all $\tau > 0$, for $C_S \subseteq \mathcal{C}$ with $C_S \neq \emptyset$,*

$$\lim_{n \to \infty} \mu_p(\{H : |g_3(H) - \mathbb{E}[g_3 | C_S]| > \tau, H \in C_S\}) \to 0.$$

Because there are only M total sets $C_S \subseteq \mathcal{C}$, for any $\tau, \rho > 0$, there exists n large enough so that

$$\mu_p(\{H : |g_3(H) - \mathbb{E}[g_3 | C_S]| > \tau, H \in C_S\}) \leq \rho$$

for all $C_S \subseteq \mathcal{C}$. Additionally, observe that, for $H \in \mathcal{C}$, $\mu_p(H) \cdot g_3(H)^2$ is uniformly bounded since $\|g_3\|_2$ is bounded by the triangle inequality (using that $\|f\|_2 \leq 1$). Hence, $\mu_p(H) \cdot |g_3(H)|$ is also uniformly bounded. Thus, since for each $C_S \subseteq \mathcal{C}$, $\mu_p(C_S) \geq \varepsilon/M$, $|\mathbb{E}[g_3 | C_S]|$ is bounded by a constant, so $|g_4|$ is bounded by some constant C. Then,

$$\|f - g_4\|_2^2 \leq (C+1)^2 \cdot \rho + 3\varepsilon + (3\varepsilon + \tau)^2 \leq 4\varepsilon,$$

for ρ and τ sufficiently small.

Making the approximation Boolean. The next approximation of f, g_5, will be a Boolean function constant on each C_S. For S such that $C_S \subseteq \mathcal{C}$ nonempty, define g_5 to be 1 on C_S or 0 on C_S, whichever approximates g_4 on C_S better. For $H \notin \mathcal{C}$, set $g_5(H) = g_4(H) = 1$.

Lemma 3.13. $\|g_5 - f\|_2^2 \leq 2 \|g_4 - f\|_2^2$.

Proof. Let h be the best approximation of f which is constant on each C_S (with respect to $\|\cdot\|_2$). We will compute the value of h directly. If h takes constant value x_* on C_S, then x_* must minimize the equation

$$(1-x)^2\mu_p(\{H : f(H)=1\}\cap C_S) + x^2\mu_p(\{H : f(H)=0\}\cap C_S),$$

so,

$$x_* = \frac{2\mu_p(\{H : f(H)=1\}\cap C_S)}{2\mu_p(C_S)} = \mu_p(f(H)=1 : H \in C_S).$$

Note also that if $\mu_p(f(H)=1 : H\in C_S) > 1/2$ then g_5 must be 1 on C_S and if $\mu_p(f(H)=1 : C_S) < 1/2$ it must be 0 (and otherwise g_5 can be assigned a Boolean value on C_S arbitrarily). The contribution of C_S to $\|f-h\|_2^2$ is

$$\mu_p(C_S)(\mu_p(f(H)=1 : C_S)(1-x_*)^2 + (1-\mu_p(f(H)=1 : C_S))x_*^2) = \mu_p(C_S)\cdot x_*(1-x_*).$$

In contrast, the contribution of C_S to $\|f-g_5\|_2^2$ is

$$\mu_p(C_S)\min(\mu_p(f(H)=1 : C_S), 1-\mu_p(f(H)=1 : C_S)) = \mu_p(C_S)\cdot\min(x_*, 1-x_*).$$

which is at most 2 times the contribution of C_S to $\|f-h\|_2^2$. Hence,

$$\|f-g_5\|_2^2 \le 2\|f-h\|_2^2 \le 2\|f-g_4\|_2^2,$$

as desired, using that g_4 is constant on each C_S. □

We then get that g_5 is a good approximation of f as a consequence of g_4 being a good approximation of f. Namely, $\|g_5-f\|_2^2 \le 8\varepsilon$.

Making the approximation monotone. Our approximation g_5 is now a characteristic function, but may not yet be monotone. This is the final piece of the puzzle in proving Theorem 3.2. Note that g_5 is symmetric since g_5 is constant on each C_S and, if $H \in C_S$, then $\mathcal{O}(H)\subseteq C_S$. Our next approximation of f will be defined to be both Boolean and monotone. Consider the graphs S such that

$$\Pr[f(T)=1 : T\in C_S] \notin (\tau, 1-\tau),$$

for an appropriately chosen τ. These graphs are convenient for the purposes of rounding; we can round each such C_S non-empty to its majority value while incurring a low penalty in our approximation (we will see this formally momentarily). For this reason, call such S *decisive.*

The analysis of Lemma 3.13 shows that the union of C_S over the set of indecisive graphs S has small total measure. Call this set $\mathcal{C}_3$. Namely, each indecisive graph S contributes at least $\tau^2\mu_p(C_S)$ to $\|f-h\|_2^2$. So, the total contribution will be at least $\tau^2\mu_p(\mathcal{C}_3)$. This is bounded by $2\|f-g_4\|_2^2 \le 8\varepsilon$, so $\mu_p(\mathcal{C}_3)\le 8\varepsilon/\tau^2$.

Since $\mu_p(\mathcal{C}_1\cup\mathcal{C}_2\cup\mathcal{C}_3) \le 2\varepsilon + 8\varepsilon/\tau^2$, we can make this arbitrarily small by choosing an appropriately small ε. Then, we can modify f on these sets without sacrificing much in terms of our approximation. an arbitrary modification would not ensure monotonicity.

As such, we define g_6 so that g_6 is 1 on C_S if and only if S has some subgraph R with $C_R \not\subseteq \mathcal{C}_1\cup\mathcal{C}_2\cup\mathcal{C}_3$. This ensures monotonicity. Additionally, it turns out that g_6 and g_5 are actually equal on $C_S \not\subseteq \mathcal{C}_1\cup\mathcal{C}_2\cup\mathcal{C}_3$ for an appropriate choice of τ, making it simple to analyze the approximation.

Lemma 3.14 (Consistency of decisive graphs, Lemma 4.8 of [FB+99])**.** *If $R\subseteq S$ and $C_R, C_S \not\subseteq \mathcal{C}_1\cup\mathcal{C}_2\cup\mathcal{C}_3$, then*

$$\mathbb{E}[f|C_R] \ge 1-\tau \implies \mathbb{E}[f|C_S] \ge 1-2\tau.$$

Now, let $\tau < 1/4$ and suppose that $C_R \not\subseteq \mathcal{C}_1\cup\mathcal{C}_2\cup\mathcal{C}_3$. Lemma 3.14 implies that, if g_5 is 1 on C_R, then, if $R\subseteq S$ and S is decisive, g_5 is 1 on C_S as well. So, we have

$$\|g_6-f\|_2^2 \le \mu_p(\mathcal{C}_1\cup\mathcal{C}_2\cup\mathcal{C}_3) + \|f-g_5\|_2^2 \le 8\varepsilon + 2\varepsilon + 128\varepsilon = 138\varepsilon.$$

Letting $\varepsilon' = \varepsilon/138$, we have that g_6 is a characteristic function of a symmetric, monotone graph family B with $\mu_p(A\Delta B)\le\varepsilon'$. Crucially, the maximum size of a minimal subgraph in B is bounded by a constant. Indeed, for $C_S \not\subseteq \mathcal{C}_1\cup\mathcal{C}_2\cup\mathcal{C}_3$, for each $H\in C_S$, $S\subseteq H$. Moreover, $g(H)=1$ if and only if $g(S)=1$ and $|S|$ is bounded by a constant by definition of $\mathcal{C}_1$ and the fact that modest graphs have constant bounded size. These S such that $C_S \not\subseteq \mathcal{C}_1\cup\mathcal{C}_2\cup\mathcal{C}_3$ and $g(S)=1$ are exactly the minimal graphs in B. This completes the proof of Theorem 3.2.

3.4 A dichotomy corollary of Friedgut's Coarse Threshold Criterion

In this section we derive a useful corollary of Theorem 3.2. This alternative form will play an integral role in the proof that the property of being 2-Ramsey for a triangle has a sharp threshold.

Our first goal is to show that there are bounded size subgraphs of K_n which "boost" the probability of having the monotone property indicated by f. In contrast to Theorem 3.2, these subgraphs are on a specific set of edges, not some isomorphism class of subgraphs. This will come from a simple averaging argument. First, observe that we could have defined g_6 in the proof of Theorem 3.2 so that g_6 is 1 on C_S if and only if there is a subgraph $R \subseteq S$ with $C_R \nsubseteq \mathcal{C}_1 \cup \mathcal{C}_2 \cup \mathcal{C}_3$ and $\mathbb{E}[f|C_R] \geq 1 - \tau/2$. This is because the additional contribution to $\|g_6 - f\|_2^2$ can be made negligible by making τ sufficiently small. The consequence of doing this is that, then, for all C_S for which $g_6(C_S) = 1$, we have $\mathbb{E}[f|C_S] \geq 1 - \tau$ (using Lemma 3.14).

Fix some such S. Each element of C_S has exactly one copy of S as a subgraph (or else the union of its modest subgraphs will be larger than S). We can then partition C_S into subsets $C_{s_1}, C_{s_2}, \ldots$ where $s_1, s_2, \ldots$ enumerates all specific copies of S in K_n. By symmetry, $\mathbb{E}[f|C_s] = \mathbb{E}[f|C_{s'}]$ for all s, s' specific copies of S in K_n. Additionally, the probability that $H \in C_s$ given that $H \in C_S$ is equal for all specific copies of S, s. Hence, for any specific copy s of S in K_n, we have

$$\mathbb{E}[f|C_s] = \mathbb{E}[f|C_S].$$

Fix some specific copy of S in K_n, s. Let B_s be the space of graphs having s as a subgraph. Let $C_s := C_S \cap B_s$, as above. Now, since $\mathbb{E}[f|B_s \setminus C_s] \geq \mathbb{E}[f|C_s]$ since A is a monotone property, we have

$$\mathbb{E}[f|B_s] \geq \mathbb{E}[f|C_s] = \mathbb{E}[f|C_S].$$

Namely, $\mathbb{E}[f|B_s] \geq 1 - \tau$. This motivates the following definition.

Definition 3.6 (Booster)**.** Let A be a monotone graph property. Then, for $p \in [0,1]$, $\delta > 0$, $B \subset K_n$ is a (p, δ)-*booster* if

$$\Pr[G(n,p) \cup B \in A] \geq \Pr[G(n,p) \in A] + \delta.$$

The graph s in the above discussion is a $(p, \rho - \tau)$-booster for our monotone property A: its inclusion in the random graph "boosts" the probability the random graph will have property A (here ρ is the absolute constant depending only on A from Remark 3.1). Also, nothing about the above argument depended on the choice of copy s, so all isomorphic copies of s are also $(p, \rho - \tau)$-boosters.

Also, observe that some copy of S is contained in $G(n,p)$ with at least constant probability. This is because S is a disjoint union of modest graphs, and, hence, $\Theta(1)$ copies arise in expectation, $|S| \leq K$ for some constant K, and, crucially, S is balanced. Disjointness comes from the removing low measure graph classes step. As such, the pseudovariance of the number of copies of S arising in $G(n,p)$ will be $O(1)$. (Consider squaring the sum of indicators for each fixed copy of S. The result is a sum of indicators of unions of pairs of fixed copies of S with each union appearing a constant bounded number of times. The union of the copies of S have average degree at least that of S, see Lemma 3.11, and there are a constant number of unions up to isomorphism.) As such, by Janson's inequality, we get that the probability that no copies of S occur in $G(n,p)$ is bounded by a constant away from 0.

We can now state a useful corollary of Friedgut's result, following from the above discussion.

Theorem 3.15. *Suppose that A is a nontrivial monotone property with a coarse threshold. Let $\rho > 0$ and let $p = p(n)$ such that for n sufficiently large, $\rho \leq \mu_p(A) \leq 1 - \rho$. Then, there there exists δ, η, and K such that there is a family $\mathcal{B}$ of (p, δ)-boosters, each of size at most K, satisfying*

$$\Pr[B \subseteq G(n,p) \text{ for some } B \in \mathcal{B}] > \eta.$$

When we apply Theorem 3.15 to prove sharp thresholds using the hypergraph container method it will be convenient to consider not only boosters, but those boosters that, when added to our random graph, cause it to have the desired monotone graph property A. This inspires the following definition.

Definition 3.7 (Active Booster)**.** Let A be a monotone graph property. A graph $B \in \mathcal{B}$ is an *active booster* for $H \subseteq K_n$ if $H \cup B \in A$.

We can now state an extremely useful consequence of Theorem 3.15.

Theorem 3.16 (Proposition 4.4 of [FKSS22]). *Let A be a nontrivial, monotone graph property with a coarse threshold. Let $\hat{p}$ be such that $\mu_{\hat{p}}(A) = 1/2$. Then there exists positive constants ε, δ, K and sequence of probabilities $p = p(n)$ with $p = \Theta(\hat{p})$ such that:*

1. *For all families $\mathcal{F}$ of graphs with*

$$\Pr[H \in G(n,p) \text{ for some } H \in \mathcal{F}] < \varepsilon,$$

there is a (p,δ)-booster of size at most K not in $\mathcal{F}$.

2. *For any family $\mathcal{B}$ of (p,δ)-boosters, sampling $S \sim G(n,p)$, the following hold with probability at least ε:*
   (a) $\Pr[S \cup G(n, \varepsilon p) \in A] \leq 1/2$*;*
   (b) at least $\varepsilon|\mathcal{B}|$ elements of $\mathcal{B}$ are active boosters for S.

The primary utility of the first implication in Theorem 3.16 is to facilitate pruning low probability problematic candidate boosters. In the case of symmetric graph properties (the only kind we will consider), we can then turn the (p,δ)-booster B_0 guaranteed by the first condition into a large family of non-problematic boosters of size at most K by just considering all copies of B_0 contained in K_n. Symmetricity implies that all are (p,δ)-boosters. We can also disregard problematic low probability instances of S sampled from $G(n,p)$ by leveraging the fact that the latter dichotomy conditions hold with probability at least ε.

Suppose we had a family $\mathcal{B}$ of (p,δ)-boosters, perhaps resulting from a procedure as described above. The dichotomy in the second implication tells us that, apart from a low probability case, given our random graph S sampled from $G(n,p)$, unioning a random subgraph $G(n,\varepsilon p)$ with S only yields a graph with property A with bounded probability. Meanwhile, a constant proportion of boosters in $\mathcal{B}$ are active, meaning that their union with S yields a graph with property A. This is especially remarkable since the graphs in $\mathcal{B}$ are only of constant size.

Proof of Theorem 3.16. Since A is a nontrivial, monotone graph property with a coarse threshold, there exists constants $\rho > 0$, $c_1 < c_2$ with

$$\rho \leq \mu_{c_1\hat{p}}(A) \leq \mu_{c_2\hat{p}}(A) \leq 1 - \rho$$

on an infinite subsequence. We will need to apply a more careful version of the argument in Remark 3.1. Define $C := 4c_2/(c_2 - c_1)$. The following technical claim usefully allows us to say more than just $\mu'_p(A) \leq C/p$.

Claim 3.17 (Claim 4.5 of [FKSS22]). *For all $\gamma > 0$, there exists $p \in (c_1 \cdot \hat{p}, c_2 \cdot \hat{p})$ such that*

1. $\mu'_p(A) \leq C/p$ *and*
2. $\mu_{(1+\gamma)p}(A) - \mu_p(A) \leq C \cdot \gamma$.

Let p be as in Claim 3.17 with $\gamma = \varepsilon$. Now, let $\mathcal{B}$ be the family of (p,δ)-boosters from applying Theorem 3.15. Let B be a booster sampled uniformly at random from $\mathcal{B}$. For $S \subseteq K_n$, define

$$g(S) := \Pr_B[S \cup B \in A]$$

and

$$h(S) := \Pr[S \cup G(n, \varepsilon p) \in A].$$

Note that, to show the second part of Theorem 3.16, all we need to do is show

$$\Pr_{H \sim G(n,p)}[g(H) \geq \varepsilon \text{ and } h(H) \leq 1/2] \geq \varepsilon.$$

We can do this using Markov's inequality. Indeed,

$$\begin{aligned}\Pr[g(G(n,p)) < \varepsilon] &= \Pr[1 - g(G(n,p)) > 1 - \varepsilon] \\ &< \frac{1 - \mathbb{E}[g(G(n,p))]}{1-\varepsilon} \\ &\leq \frac{1 - \mu_p(A) - \delta}{1 - \varepsilon}.\end{aligned}$$

The first inequality is Markov's inequality, and the second uses the definition of a (p, δ)-booster. Next, observe that if $S \subseteq K_n$ such that $S \in A$, then $h(S) = 1$, by monotocity of A. Also observe that for any $\gamma > 0$,

$$\Pr[G(n,(1+\gamma)p) \in A] \geq \Pr[G(n,p) \cup G(n,\gamma p) \in A],$$

where the graphs sampled from $G(n,p)$ and $G(n,\gamma p)$ are independent on the right hand side. This is because each edge is added independently with probability $p+(1-p)\gamma p \leq (1+\gamma)p$ in the latter case. Then, combining these observations with Markov's inequality and Claim 3.17, we have

$$\begin{aligned}\Pr[h(G(n,p)) > 1/2] &= \mu_p(A) + \Pr_{H \sim G(n,p)}[h(H) - 1_A(H) > 1/2] \\ &< \mu_p(A) + 2(\mathbb{E}[h(G(n,p))] - \mathbb{E}[1_A(G(n,p))]) \\ &= 2\mathbb{E}[h(G(n,p))] - \mu_p(A) \\ &\leq 2\mu_{(1+\varepsilon)p}(A) - \mu_p(A) \\ &\leq 2C \cdot \varepsilon + \mu_p(A).\end{aligned}$$

Finally, combining our two bounds from Markov's inequality, we get

$$\begin{aligned}\Pr[g(G(n,p)) \geq \varepsilon \text{ and } h(G(n,p)) \leq 1/2] &\geq 1 - 2C\cdot\varepsilon - \mu_p(A) - \frac{1-\mu_p(A)-\delta}{1-\varepsilon} \\ &\geq \delta/2 \geq \varepsilon,\end{aligned}$$

for sufficiently small ε. Since ε need only depend on C and δ, this yields the desired result. $\square$

4 Introducing the Hypergraph Container Lemma

In this section we introduce the hypergraph container lemma, Theorem 4.3, a powerful tool which can be used, among other applications, to show that the property of being Ramsey for a triangle has a sharp threshold. The lemma, introduced in [BMS15, ST15], provides a unified framework for covering all independent sets in a hypergraph with a lower order size collection of subsets. That is, it provides a small collection of "containers" such that every independent set in the hypergraph belongs to some container. Moreover, each container is itself almost independent; its induced subhypergraph includes few hyperedges. We make the descriptors "lower order," "small," and "few" explicit in the statement of the lemma.

As an example of the power of the container lemma, we give a simple proof of a (tight) upper bound on thresholds for the property of being Ramsey for a triangle. We first introduce some notation.

Definition 4.1. We say that G is *r-Ramsey* for a graph H if any r-coloring of the edges of G contains a monochromatic copy of H. We denote this $G \to (H)_r$.

In order to apply the container lemma, we need to connect the property of being r-Ramsey for a bounded graph to the non-r-colorability of some hypergraph. We will consider the $|H|$-uniform family of hypergraphs $\mathcal{H}$ where $V(\mathcal{H}) = E(K_n)$. The hyperedges correspond exactly to edge subgraph copies of H contained in K_n. We next make the notion of hypergraph coloring precise.

Definition 4.2 (Hypergraph coloring)**.** Let $\mathcal{H}$ be a hypergraph. A coloring of $V(\mathcal{H})$ with r colors is a function $\chi : V(\mathcal{H}) \to [r]$. If no edge in $E(\mathcal{H})$ is monochromatic, that is, for all $e \in E(\mathcal{H})$, $|\chi(e)| \geq 2$, then χ is a *proper r-coloring* of $\mathcal{H}$. If $\mathcal{H}$ admits a proper r-coloring, we say $\mathcal{H}$ is (properly) r-colorable.

This is akin to the definition of proper vertex colorings in the graph setting: a proper r-coloring of a graph is exactly an r-coloring of the vertices such that every edge is not monochromatic (i.e., adjacent vertices are colored different colors). Observe that, for $S \subseteq E(K_n)$, $\mathcal{H}[S]$ is r-colorable if and only if $S \not\to (H)_r$. Namely, an r-coloring of $V(\mathcal{H}[S])$ corresponds exactly to an r-edge coloring of S and each color class in the r-coloring of $V(\mathcal{H}[S])$ being an independent set corresponds exactly to S not containing any monochromatic copies of H.

The usual subset S that we consider will be a random graph $G(n, p)$. We can now state the our main result for this section.

Theorem 4.1 (Upper bound on r-Ramsey for a triangle threshold)**.** *For every fixed $r \geq 2$, there exists constants C and n_0 such that, if $p \geq Cn^{-1/2}$, then $G(n, p) \to (H)_r$ asymptotically almost surely.*

A slightly weaker version of Theorem 4.1 was originally proved by Frankl and Rödl [FR86] for $r = 2$, and the full statement of Theorem 4.1 was first shown in [RR94]. Our treatment here is based on [BMS18].

The high-level idea of the proof of Theorem 4.1 is as follows. First, we apply Ramsey's theorem to say that there is a constant such that any r-edge coloring of a complete graph of that size contains a monochromatic copy of K_3. We use this to prove a supersaturation result about the prevalence of monochromatic copies of K_3. Namely, we show that there are $\Omega(n^3)$ monochromatic triangles in any r-coloring of K_n.

Next, we apply the hypergraph container lemma. Roughly speaking, the hypergraph container lemma tells us that all of the independent sets of the 3-uniform hypergraph $\mathcal{H}$ with edges corresponding to triangles are contained in a small collection of subhypergraphs (called *containers*). Each container has the additional property of having few hyperedges. This near-independence property of the containers is crucial since it allows us to exploit our supersaturation of monochromatic triangles to show that $G(n, p) \not\to (K_3)_r$. Namely, it allows us to show that $G(n, p)$ being contained in the union of r containers is a relatively rare event. Finally, using the fact that the total number of containers is small, we apply a union bound and deduce the result.

First we prove our supersaturation lemma.

Lemma 4.2 (Supersaturation of monochromatic triangles)**.** *There exists some constant $\alpha > 0$ such that any r-coloring of K_n contains at least αn^3 monochromatic triangles.*

Proof. Fix an arbitrary r-coloring $\chi : E(K_n) \to [r]$. By Ramsey's theorem, there exists a constant R such that any r-coloring of the edges of K_R contains a monochromatic triangle. Now, observe that there are $\Theta(n^R)$ copies of K_R contained in K_n. By Ramsey's theorem, χ induces a monochromatic triangle in each of them. In total, up to double counting, this yields $\Theta(n^R)$ monochromatic triangles. Moreover, each monochromatic triangle is contained in at most $\Theta(n^{R-3})$ copies of K_R. Hence, there are $\Omega(n^3)$ monochromatic triangles in K_n, as desired. □

4.1 The Hypergraph Container Lemma

In order to formally state and apply the hypergraph container lemma, we need a new definition. For a hypergraph $\mathcal{H}$ and some integer ℓ, let

$$\Delta_\ell(\mathcal{H}) = \max_{\substack{A \subset V(\mathcal{H}),\\ |A| = \ell}} |\{e : e \in E(\mathcal{H}), A \subseteq e\}|.$$

For $\ell = 1$, this is the maximum degree of the hypergraph. In general, $\Delta_\ell(\mathcal{H})$ is the maximum number of edges containing a specific subset of vertices of size ℓ. We are now ready to state the theorem.

We will use a formulation of the hypergraph container lemma from [FKSS22]; this formulation is slightly stronger than the original formulations in [BMS15, ST15].

Theorem 4.3 (Theorem 4.7 of [FKSS22])**.** *For all $\varepsilon, K > 0$ and $k \in \mathbb{N}$, there exists $t \in \mathbb{N}$ and $\delta > 0$ such that the following holds. Let $\mathcal{H}$ be a k-uniform (multi)hypergraph and $\tau > 0$ such that*

$$\Delta_\ell(\mathcal{H}) \leq K\tau^{\ell-1} \cdot \frac{e(\mathcal{H})}{v(\mathcal{H})}$$

for each $\ell \in [k]$. Then, there exists a collection of subsets $\mathcal{C}$ of $V(\mathcal{H})$ and a function $f : \mathcal{P}(V(\mathcal{H}))^t \to \mathcal{C}$ such that:

1. *For every $I \subseteq V(\mathcal{H})$ with $e(\mathcal{H}[I]) \leq \delta\tau^k e(\mathcal{H})$, there are $S_1, \dots, S_t \subseteq I$, each of size at most $\tau v(\mathcal{H})$, such that $I \subseteq f(S_1, \dots S_t)$.*

2. *For each $S_1, \dots, S_t \subseteq V(\mathcal{H})$, $f(S_1, \dots S_t)$ induces at most $\varepsilon e(\mathcal{H})$ hyperedges.*

Before continuing, let us attempt to break down the various moving parts in Theorem 4.3. In applications, the parameters ε, K, and k will often be fixed constants. Then, the resulting t and δ from Theorem 4.3 will also be constants.

In order to apply Theorem 4.3, it is necessary to bound $\Delta_\ell(\mathcal{H})$ for all ℓ. The fraction $\frac{e(\mathcal{H})}{v(\mathcal{H})}$ is on the order of the average degree of $\mathcal{H}$ (for constant uniformity k), so the bound on $\Delta_\ell(\mathcal{H})$ is on the order of the average degree of $\mathcal{H}$ scaled by $\tau^{\ell-1}$. The smaller the τ these bounds hold for, the smaller the collection of containers will be.

The function f, the *fingerprint* function, ensures that each independent set is contained in some container: $\mathcal{C}$ is the collection of containers. Indeed, the first implication of Theorem 4.3, implies that if $I \subseteq V(\mathcal{H})$ is independent, then $I \subseteq f(S_1, \dots S_t) \in \mathcal{C}$. (In this stronger formulation of the container lemma we also have that each set inducing a small number of edges is contained in some container, but we will not need this for our example application.)

The fact that there are a small number of containers in $\mathcal{C}$ is a consequence of how an independent set $I \subseteq V(\mathcal{H})$ is mapped to a container. Indeed, its container membership is determined by $f(S_1, \dots S_t)$ for $S_1, \dots, S_t \subseteq I$ such that each S_i has at most $\tau v(\mathcal{H})$ elements. As such, the number of containers covering all $I \subseteq V(\mathcal{H})$ with $e(\mathcal{H}[I]) \leq \delta\tau^k e(\mathcal{H})$ can be naively bounded by

$$\left(\sum_{i=0}^{\tau v(\mathcal{H})} \binom{v(\mathcal{H})}{i}\right)^t \leq \left(\frac{ev(\mathcal{H})}{\tau v(\mathcal{H})}\right)^{t\tau v(\mathcal{H})}.$$

If $\tau = o(1)$ and ε, K, and k are constants (so that t and δ are constants), then this gives us

$$|\mathcal{C}| = \exp(o(v(\mathcal{H}))),$$

since $f(S_1, \dots, S_t)$ with $|S_i| > \tau v(\mathcal{H})$ for some $i \in [t]$ can be arbitrarily set to $f(S_1', \dots, S_t')$ for some $S_1', \dots, S_t'$ with $|S_i'| \leq \tau v(\mathcal{H})$ for all $i \in [t]$. This improves upon the trivial $\exp(O(v(\mathcal{H})))$ bound on the size of the collection of containers.

The parameter ε controls the sparsity of our containers. This comes from the second implication of Theorem 4.3, since each distinct $f(S_1, \dots S_t)$ can be assumed to be one of the containers via the discussion above. Oftentimes it will be useful to set this sparsity parameter ε to conflict with a constant coming from a supersaturation result. This is exactly what we will do in our application of the container lemma in this section.

For our application in this section, we will be applying Theorem 4.3 to the 3-uniform hypergraph $\mathcal{H}$ with $V(\mathcal{H}) = E(K_n)$ and the hyperedges corresponding to triangles in K_n. So, $\Delta_1(\mathcal{H}) = n - 2$, $\Delta_2(\mathcal{H}) = \Delta_3(\mathcal{H}) = 1$, $v(\mathcal{H}) = \binom{n}{2}$, and $e(\mathcal{H}) = \binom{n}{3}$. Setting $k = 3$ and $K = 6$ and only considering $\tau = n^{-1/2}$ then yields the following.

Corollary 4.4. *Let $\mathcal{H}$ be the 3-uniform hypergraph with $V(\mathcal{H}) = E(K_n)$ and the hyperedges corresponding to triangles in K_n. Then, for all constants $\varepsilon > 0$, there exist constant $t \in \mathbb{N}$, $\mathcal{C} \subset \mathcal{P}(V(\mathcal{H}))$, and a function $f : (\mathcal{P}(V(\mathcal{H}))^t \to \mathcal{C}$ such that:*

1. *For every independent set $I \subset V(\mathcal{H})$, there exist $S_1, \dots, S_t \subseteq I$, each with at most $n^{3/2}$ elements, such that $I \subseteq f(S_1, \dots, S_t)$.*

2. *For each $C \in \mathcal{C}$, $e(\mathcal{H}[C]) \leq \varepsilon \cdot n^3$.*

4.2 An upper bound on thresholds for being Ramsey for a triangle

We are now ready to combine our supersaturation result (Lemma 4.2) and the hypergraph container lemma (Corollary 4.4) to prove Theorem 4.1.

Proof of Theorem 4.1. First fix r, the number of colors we will be considering. Now, let α be as in Lemma 4.2 with $r+1$ colors and $\varepsilon = \alpha/(r+1)$ in Corollary 4.4 (yielding $f, \mathcal{C}$, and t, with $\mathcal{H}$ the hypergraph).

Let $p = \Omega(n^{-1/2})$ and let $R \subset K_n$ be a random graph generated by $G(n,p)$. Suppose that R is not r-Ramsey for a triangle. That is, there is a coloring of the edges of R with r colors admitting no monochromatic triangles. Hence, $R \subset C_1 \cup C_2 \cdots \cup C_r$ for some $C_1, C_2, \dots, C_r \in \mathcal{C}$. We can extend the r-edge coloring of R to an $(r+1)$-edge coloring of K_n by coloring $K_n \setminus R$ all the $(r+1)^{\text{th}}$ color. Then, by Lemma 4.2, $K_n \setminus R$ must contain at least $\frac{\alpha n^3}{(r+1)}$ triangles. This is because each C_i contains at most εn^3 monochromatic triangles, and K_n contains at least αn^3 monochromatic triangles overall. Hence, since each edge is in at most n triangles, $K_n \setminus R$ must contain at least $\frac{\alpha n^2}{(r+1)} = \varepsilon n^2$ edges. This is the rare event we were looking for!

Now, let $\mathcal{S}$ be a collection of t-tuples of sets of size at most $n^{3/2}$ such that, for each $\mathbf{S} = (S_1, \dots, S_t) \in \mathcal{S}$, $f(\mathbf{S}) \in \mathcal{C}$ is unique. For $(\mathbf{S}_1, \dots, \mathbf{S}_r) \in \mathcal{S}^r$, let $A_{(\mathbf{S}_1, \dots \mathbf{S}_r)}$ be the event that

$$\bigcup_{i=1}^{r} \bigcup_{S \in \mathbf{S}_i} S \subseteq G(n,p)$$

and

$$G(n,p) \setminus \left(\bigcup_{i=1}^{r} \bigcup_{S \in \mathbf{S}_i}^{t} S \right) \subseteq \bigcup_{i=1}^{r} f(\mathbf{S}_i) \setminus \left(\bigcup_{i=1}^{r} \bigcup_{S \in \mathbf{S}_i} S \right).$$

That is, $A_{(\mathbf{S}_1, \dots \mathbf{S}_r)}$ is the event that the $\mathbf{S}_i$'s in $(\mathbf{S}_1, \dots \mathbf{S}_r)$ can determine the containers for an r-edge coloring of $G(n,p)$ without monochromatic triangles. Note that the two subevents making up $A_{(\mathbf{S}_1, \dots \mathbf{S}_r)}$ are actually independent. Then, by definition of $A_{(\mathbf{S}_1, \dots \mathbf{S}_r)}$ and our independence observation,

$$\begin{aligned}
\Pr[G(n,p) \not\to (K_3)_r] &\leq \Pr\left[A_{(\mathbf{S}_1, \dots \mathbf{S}_r)} \text{ for } (\mathbf{S}_1, \dots, \mathbf{S}_r) \in \mathcal{S}^r\right] \\
&\leq \sum_{s \leq rtn^{3/2}} \binom{n^2/2}{s} p^s \cdot (1-p)^{\varepsilon n^2} \\
&\leq \sum_{s \leq rtn^{3/2}} \left(\frac{en^2 p}{2s}\right)^s \exp(-\varepsilon p n^2),
\end{aligned}$$

In the second step we use our rare event to bound the probability of the second part of $A_{(\mathbf{S}_1, \dots \mathbf{S}_r)}$. In the last step we use $(1-x) \leq e^{-x}$ for all $x \in \mathbb{R}$ and $\binom{n}{r} \leq (en/r)^r$. So, for $p \geq Cn^{-1/2}$ for sufficiently large C, we can then bound this sum by $\exp(-\varepsilon p n^2/2)$, yielding the desired result. ☐

5 Properties of the Hypergraph of Triangles in K_n

Before moving onto the proof that being 2-Ramsey for a triangle has a sharp threshold, we take a closer look at the hypergraph considered in the last section. We observe five useful properties, (A1)–(A5).

Concretely, let $\mathcal{H}$ be the 3-uniform hypergraph with $V(\mathcal{H}) = E(K_n)$ and hyperedges corresponding to all triangles in K_n, as seen in Section 4. Let $p_{\mathcal{H}} = n^{-1/2}$. Denote the random sub-hypergraph of $\mathcal{H}$ obtained by keeping each vertex independently with probability p by $\mathcal{H}_p$ (this is equivalently $\mathcal{H}[G(n,p)]$). The first property we will make use of is the fact that $\mathcal{H}$ is vertex transitive.

(A1) *Symmetry.* The group of automorphisms $\mathrm{Aut}(\mathcal{H})$ acts transitively on $V(\mathcal{H})$.

This is immediate: the permutation swapping vertices u and v and fixing the remaining vertices is an automorphism, for example.

The next property is that we can apply the hypergraph container lemma (Theorem 4.3) to $\mathcal{H}$.

(A2) *Non-clusteredness.* For all $\ell \in [3]$, $\Delta_\ell(\mathcal{H}) = O(p_{\mathcal{H}}^{\ell-1} \cdot \frac{e(\mathcal{H})}{v(\mathcal{H})})$.

We already checked this property in Section 4. A crucial property, and the only one whose proof will not be self-contained in this exposition, is that $p_{\mathcal{H}}$ is a threshold probability for the property of being 2-Ramsey for a triangle. We proved that $p_{\mathcal{H}}$ is an upper bound of the threshold probability in Section 4 as a sample application of the hypergraph container lemma. For the proof of the lower bound, we refer the reader to [FR86, LRV92].

(A3) *Weak threshold.* $p_{\mathcal{H}} = n^{-1/2}$ is a threshold probability for $G(n,p) \to (K_3)_2$.

The next property is an important necessary property for proving that $G(n,p) \to (K_3)_2$.

(A4) *Colorability of typical bounded-size subsets.* For all $K > 0$, for all $S \subset \mathcal{H}_{p_{\mathcal{H}}}$ with $|S| \leq K$, $\mathcal{H}[S]$ is 2-colorable, asymptotically almost surely.

That is, typical bounded-size subgraphs of $G(n, p_H)$ cannot be 2-Ramsey for a triangle. Note that Lemma 3.1 implies that, if there exists $K > 0$ such that the probability of containing a subgraph $H \subseteq G(n, p_{\mathcal{H}})$ with $|H| \leq K$ with $H \to (K_3)_2$ is bounded below by some constant, then the property of $G(n,p) \to (K_3)_2$ in fact has a coarse threshold. This uses property (A3).

Proposition 5.1. *Property* (A4) *holds.*

We begin with a helpful claim bounding the degree of typical bounded size subgraphs of $G(n, p_{\mathcal{H}})$.

Claim 5.2. *For all constants $K > 0$, asymptotically almost surely (a.a.s.), for all subgraphs $H \subseteq G(n, p_{\mathcal{H}})$ with $|H| \leq K$, $e(H)/v(H) \leq 4$.*

Proof. Let $k \leq K$, and let $S \subseteq [n]$ with $|S| = k$. Observe that

$$\Pr[|G(n, p_{\mathcal{H}})[S]| > 2k] = O(n^{-k-1/2}),$$

since $p_{\mathcal{H}} = n^{-1/2}$. Then, union bounding over all $\binom{n}{k}$ subsets of size k in $[n]$, we get that

$$\Pr[\exists S \subseteq [n], |S| = k \text{ s.t. } |G(n, p_{\mathcal{H}})[S]| > 2k] = O(n^{-1/2}).$$

Finally, union bounding over all $1 \leq k \leq K$, we get that

$$\Pr[\exists S \subseteq [n], |S| \leq K \text{ s.t. } |G(n, p_{\mathcal{H}})[S]| > 2|S|] = O(n^{-1/2}),$$

yielding the desired result. □

Proof of Proposition 5.1. Fix $K > 0$. Let $G_{p_{\mathcal{H}}} \sim G(n, p_{\mathcal{H}})$. Let $S \subseteq G_{p_{\mathcal{H}}}$ with $|S| \leq K$. Denote $G_{p_{\mathcal{H}}}[S]$ by H and suppose for the sake of contradiction that $H \to (K_3)_2$, and there is some $v \in V(H)$ such that $H \setminus v \not\to (K_3)_2$. If the former condition holds but the latter does not, we can restrict to a smaller subgraph S to make the latter hold. A.a.s., by Claim 5.2, the average degree of H is at most 4. Then, either H has minimum degree less than 4 or H is 4-regular.

By assumption, any extension of a 2-edge coloring of $H \setminus v$ without monochromatic triangles to H introduces a monochromatic triangle including v. Consider $H[N(v)]$ (where $N(v)$ denotes the neighborhood of v). Suppose that the edges of $H[N(v)]$ can be oriented so that each vertex has outdegree ≤ 1. Then, we can extend any 2-edge coloring of $H \setminus v$ without monochromatic triangles to H as follows. For $u \in N(v)$, color edge (v, u) the same color as the outgoing edge from u in the orientation of the edges of $H[N(v)]$ if such an outgoing edge exists. Otherwise color edge (v, u) arbitrarily. Now, any triangle (v, u_1, u_2) with $u_1, u_2 \in N(v)$ includes the outgoing edge from either u_1 or u_2 (as it includes the edge (u_1, u_2)) and hence, since it also includes both edges (v, u_1) and (v, u_2), cannot be monochromatic. See Figure 3.

Now, if $e(H[N(v)]) \leq 4$, then we can orient the edges of $H[N(v)]$ so that each vertex has at most one outgoing edge. For each node of degree 1, orient its incident edge to be outgoing from it. Then, all the unoriented edges form a cycle, and directing the edges in one direction about the cycle yields the desired orientation. Thus, we can assume that for each $v \in V(H)$, $e(H[N(v)]) \geq 5$. Concretely, if there is $v \in V(H)$ with $e(H[N(v)]) \leq 4$, since $H \to (K_3)_2$, we must have $H \setminus v \to (K_3)_2$ by the above.

This implies that H is 4-regular. In fact, we also get that, for each $v \in V(H)$, $e(H[N(v)]) = 6$. Indeed, suppose there were $v \in V(H)$ with $e(H[N(v)]) = 5$ with neighbors u_1, u_2, u_3, and u_4 and $(u_3, u_4) \notin H$. Then, let w be the one unaccounted for neighbor of u_3. It cannot be adjacent to any of v, u_1, or u_2 since all four of their neighbors are accounted for. These are the other three neighbors of u_3, and, hence, $e(H[N(u_3)]) = 3$, a contradiction. See Figure 5.

Finally, H being 4-regular with $e(H[N(v)]) = 6$ for all $v \in V(H)$ implies that $H = K_5$. But K_5 admits a 2-coloring without any monochromatic triangles: e.g., decompose K_5 as the union of two 5-cycles and color one 5-cycle red and the other blue. This yields the desired contradiction. □

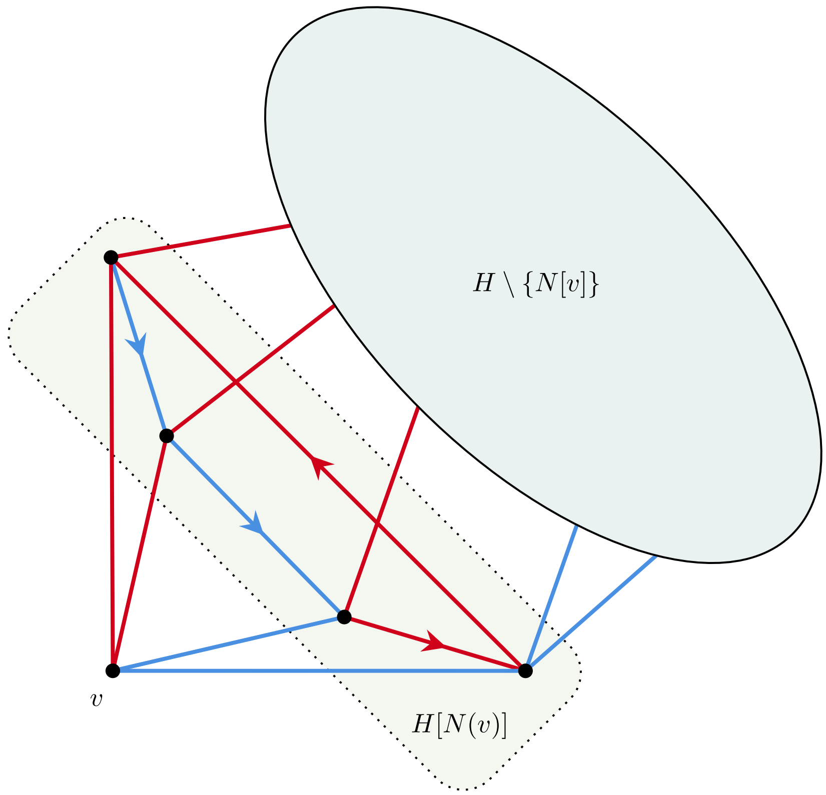

Figure 3: Transferring edge-orientations of $H[N(v)]$ to extensions of edge-colorings of $H \setminus v$.

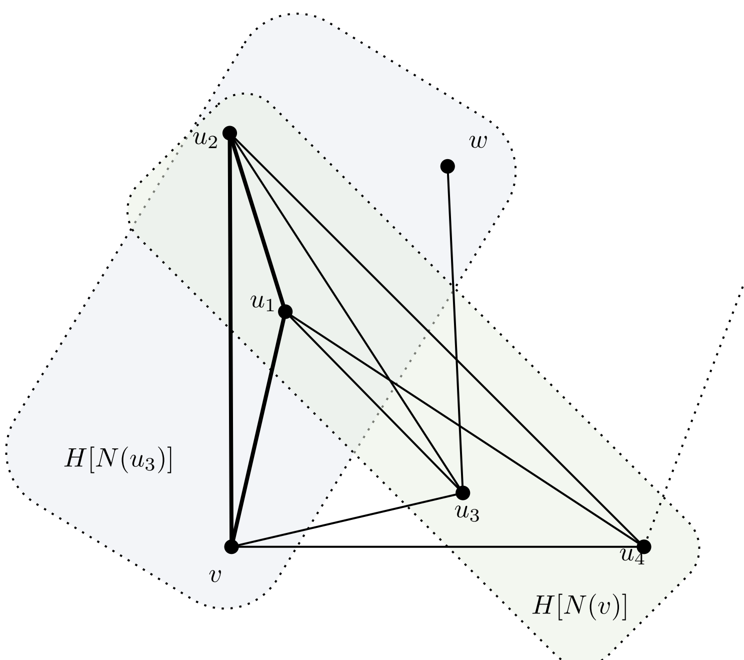

Figure 4: Diagram showing that, for each $v \in V(H)$, $e(H[N(v)]) = 6$.

The fifth and final property we require is a supersaturation result for an application of the hypergraph container lemma on a hypergraph related to $\mathcal{H}$. At a high-level, it will allow us to argue that relevant partial proper colorings of $\mathcal{H}$ have a large number of edge violations in any extension of the coloring to the full hypergraph.

To make this concrete, we begin with a definition.

Definition 5.1 (Forced triangle)**.** Let φ be a partial coloring of $\mathcal{H}$. Then, the 4-tuple of hyperedges (T, T_1, T_2, T_3) forms a *forced triangle* centered at T if there are two vertices in each of T_1, T_2, and T_3 such that φ colors them all the same color and T is composed of the other three vertices in T_1, T_2, and T_3.

Definition 5.1 is called a *forced triangle* because if the partial coloring is a partial proper coloring, then any extension of the coloring to include the union of the vertices in the forced triangle includes a violating hyperedge (which amounts to a monochromatic triangle when interpreted as a graph). See Figure 5.

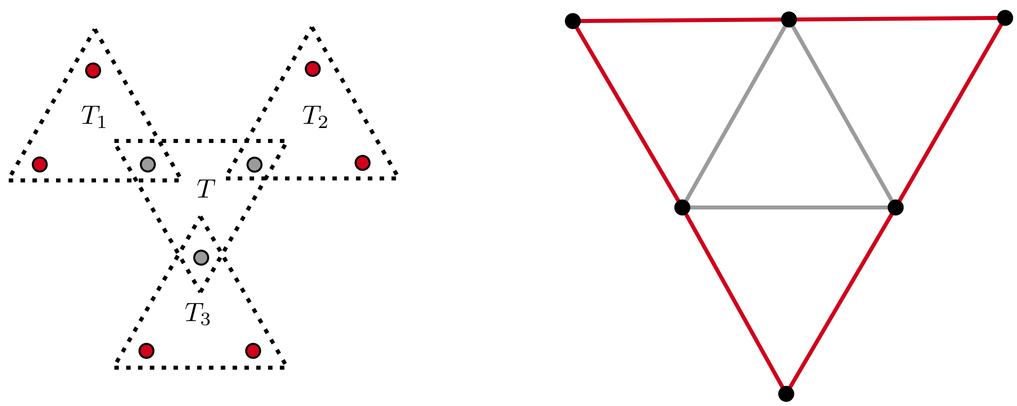

Figure 5: Red forced triangle (T, T_1, T_2, T_3) and its corresponding graphical representation.

(A5) *Partial coloring supersaturation.* Any partial coloring of $\mathcal{H}$ coloring at least two out of three vertices of $\Omega(n^3)$ hyperedges the same color induces $\Omega(n^6)$ forced triangles.

Proposition 5.3. *Property* (A5) *holds.*

Proof. Let φ be a partial coloring of $\mathcal{H}$ coloring at least two out of three vertices of $\Omega(n^3)$ hyperedges the same color. Assume without loss of generality that φ colors two out of three vertices of $\Omega(n^3)$ hyperedges red, and restrict φ to only consider its red colorings (treat all vertices colored blue as uncolored). It will be more convenient to view φ as a partial coloring of the edges of K_n. As usual, the vertices in $\mathcal{H}$ correspond to edges in K_n, and the hyperedges in $E(\mathcal{H})$ correspond to triangles in K_n.

Now, since each $e \in E(K_n)$ belongs to at most $n-2$ triangles, there exists $\Omega(n^2)$ total $e \in E(K_n)$ such that e is the third edge in $\Omega(n)$ triangles with the other two edges colored red. Let $S \subset E(K_n)$ be the set of all such edges.

In particular, fix $(u, v) \in S$. Let $X_{(u,v)}$ be the set of non-u, non-v vertices in each of the $\Omega(n)$ triangles including (u, v) with other two edges colored red. Every triangle of vertices $X_{(u,v)}$ is formed by three edges forming the third edges of a triple of triangles with other two edges colored red. Namely, every triangle of vertices in $X_{(u,v)}$ corresponds to a forced triangle (and the three triangles with two edges colored red each have two vertices in $X_{(u,v)}$ and the vertex between the red edges either u or v). See Figure 5. Since all the forced triangles formed this way are distinct when varying (u, v) over S, this yields $\Omega(n^6)$ distinct forced triangles. □

Remark 5.1. In [FKSS22], they prove that the conjunction of more general versions of properties (A1)–(A5) is sufficient for showing sharp thresholds for a much broader class of Ramsey properties on hypergraphs. In their more general setting, establishing (A4) and (A5) requires considerably more work.

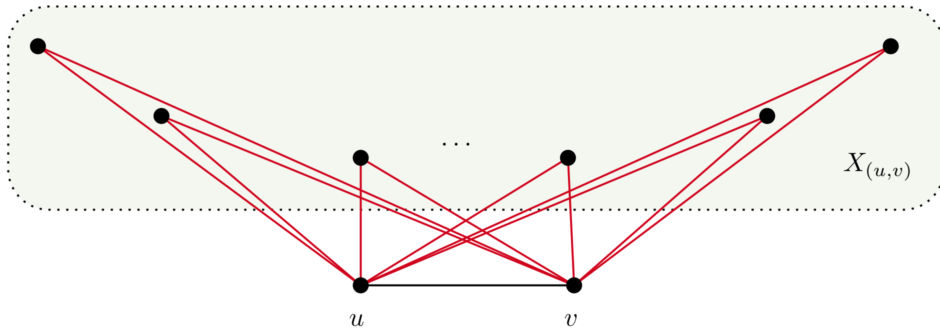

Figure 6: Structure of $X_{(u,v)}$ for some $(u, v) \in S$

6 The Threshold for Being Ramsey for a Triangle is Sharp

The focus of the rest of this exposition is the unifying method of proving sharpness of thresholds for Ramsey properties from [FKSS22]. We will focus our attention on proving that the property of being 2-Ramsey for a triangle has a sharp threshold, using properties (A1)–(A5) from Section 5. More rigorously, we have the following.

Theorem 6.1 (The Threshold for Being Ramsey for a Triangle is Sharp)**.** *Let A be the property $G \to (K_3)_2$. Then, there exists a sequence $p = p(n)$ with $p = \Theta(n^{-1/2})$ such that, for all $\varepsilon > 0$,*

$$\Pr[G(n, p(1 - \varepsilon)) \in A] = o(1)$$

and

$$\Pr[G(n, p(1 + \varepsilon)) \in A] = 1 - o(1).$$

Throughout the proof, we will make heavy use of the hypergraph container lemma, Theorem 4.3, introduced in Section 4. As such, it will be convenient to adopt the hypergraph view of these coloring properties as in Section 4 and 5. That is, let $\mathcal{H}$ be the 3-uniform hypergraph with $V(\mathcal{H}) = E(K_n)$ and hyperedges corresponding to all triangles in K_n, and let $p_{\mathcal{H}} = n^{-1/2}$, as in Section 5. Let $S \subseteq E(K_n)$. A proper 2-coloring of $\mathcal{H}[S]$ is precisely a 2-coloring of the edges S in $E(K_n)$ without any monochromatic triangles.

Our goal will be to derive a contradiction with the dichotomy version of Friedgut's criterion, Theorem 3.16. In particular, we start by assuming for the sake of contradiction that the property of $G \to (K_3)_2$ has a coarse threshold. Then, Theorem 3.16 applies. Whenever we use p for the remainder of this section, we refer to p from the infinite sequence admitted by Theorem 3.16. For simplicity, we can assume that $p \leq K p_{\mathcal{H}}$ for the same K bounding the booster size in the first implication of the theorem. Using the first implication of Theorem 3.16, we can enforce that the booster we find is suitably nice (by including low probability, problematic boosters in the family $\mathcal{F}$). Then, given an appropriate choice of $\mathcal{F}$, the first implication will yield a (p, δ)-booster B_0. Our family of boosters $\mathcal{B}$ for the second implication will then be the family of isomorphic copies of B_0 in K_n (and, by the symmetry of the property $G \to (K_3)_2$, these are all (p, δ)-boosters themselves). It will be helpful to view the family of sets $\mathcal{B}$ as a $|B_0|$-uniform hypergraph (also on vertex set $E(K_n) = V(\mathcal{H})$).

Since the second implication holds with probability at least ε, we can also disregard low probability "bad" random graphs $S \sim G(n, p)$ to focus instead on suitably nice S. Our starting point will be the family $\mathcal{S}$ of $S \subseteq E(K_n)$ for which the second implication holds, and we will gradually refine $\mathcal{S}$.

What properties of B_0 and S are useful to us? Our hope is to eventually show that the dichotomy offered by the second implication of Theorem 3.16 cannot hold. We choose properties of B_0 and S so that having at least a constant fraction of $\mathcal{B}$ being active boosters, i.e., $S \cup B \in A$ for $B \in \mathcal{B}$, implies $\Pr[S \cup G(n, \varepsilon p) \in A] = 1 - o(1)$, contradicting (a) of the second implication of Theorem 3.16. Roughly, this will come from showing that, for any proper 2-coloring ψ of $\mathcal{H}[S]$, any extension of this coloring to all of $V(\mathcal{H})$ induces $\Omega(e(\mathcal{H}))$ many monochromatic edges. Then we can leverage Janson's inequality, Theorem 2.1, to show that, for any choice of coloring ψ, $S \cup G(n, \varepsilon p)$ very likely includes one of these edges. Finally, applying a careful union bound—coming from considering partial proper colorings and the hypergraph container lemma—we will be able to conclude the result.

6.1 Applying the Friedgut Coarse Threshold Criterion

A few observations are in order. We begin by interpreting the second implication of Theorem 3.16 when applied to the property of being 2-Ramsey for a triangle.

Observation 6.2. *We may assume that both $\mathcal{H}[S]$ and $\mathcal{H}[B_0]$ are 2-colorable.*

When $S \in \mathcal{S}$, $\mathcal{H}[S]$ is 2-colorable since, by (a) of the dichotomy,

$$\Pr[S \cup G(n, \varepsilon p) \not\to (K_3)_2] \geq 1/2 > 0.$$

We can ensure that $\mathcal{H}[B_0]$ is 2-colorable by choosing the constant bound in (A4) larger than the K from Theorem 3.16 (and adding the family of $o(1)$ total probability problematic graphs to $\mathcal{F}$). This is actually the only property we will enforce for B_0.

We now introduce several definitions. For each $S \in \mathcal{S}$, for at least an ε fraction of $B \in \mathcal{B}$, $S \cup B \to (K_3)_2$. We refer to this subset of $\mathcal{B}$ as $\mathcal{B}_S$. This is precisely the subset of activated boosters for S. That is,

$$\mathcal{B}_S := \{B \in \mathcal{B} \,:\, \mathcal{H}[S \cup B] \text{ is not 2-colorable}\}.$$

Since both $\mathcal{H}[S]$ and $\mathcal{H}[B]$ are each individually 2-colorable by Observation 6.2, the only limiting factor for 2-colorability of $\mathcal{H}[S \cup B]$ is the the set of edges in $\mathcal{H}[S \cup B]$ intersecting both S and B. Call each such edge an *interacting* edge. Indeed, given any pair of proper 2-colorings ψ and φ for S and $B \in \mathcal{B}_S$, respectively, there will be a monochromatic interacting edge (and all monochromatic edges in the coloring $\psi \cup \varphi$ will be interacting edges). More generally, if ψ is instead a partial proper 2-coloring of $\mathcal{H}[S]$ (i.e., not all elements of S are assigned colors), and φ is still a proper 2-coloring of $\mathcal{H}[B]$, then $\psi \cup \varphi$ is a proper partial 2-coloring of $\mathcal{H}[S \cup B]$ if and only if no interacting edges are monochromatic. It will be essential to consider partial colorings later.

This motivates the notion of the *interface* between S and $B \in \mathcal{B}$. Namely, we define

$$I(S, B) := \{A \setminus B \,:\, A \in E(\mathcal{H}) \text{ interacting between } S \text{ and } B\}.$$

Note that we restrict our definition of the interface to the subsets of the interacting edges in S. The idea is that $I(S, B)$ contains all the necessary information to determine whether a partial proper 2-coloring ψ of S and proper 2-coloring φ of B are *consistent*. That is, the colorings agree on any vertices colored by both, and no new monochromatic edges are induced by combining the colorings. It will be crucial to carefully control the structure of $I(S, B)$.

6.1.1 Controlling the interfaces

We can control the structure of $I(S, B)$ for $B \in \mathcal{B}_S$ by restricting $\mathcal{S}$. Indeed, let $\mathcal{B}'_S$ be $\mathcal{B}_S$ after only retaining active boosters with all of the following additional properties:

(B1) B and S are disjoint.

(B2) Each set in $I(S, B)$ has cardinality exactly 2.

(B3) The sets in $I(S, B)$ are disjoint.

(B4) The collection $I(S, B)$ contains at most L sets, for L some constant depending only on ε and K.

The constant L is defined implicitly so that $\frac{K^L}{L!} \leq \varepsilon^2/8$. Define our first restriction of $\mathcal{S}$ as follows:

$$\mathcal{S}' := \{S \in \mathcal{S} \,:\, |\mathcal{B}'_S| \geq \varepsilon e(\mathcal{B})/2\}.$$

The following lemma shows that, for a typical S, few active boosters fail one of (B1)–(B4).

Lemma 6.3. *If $S \sim G(n, p)$, then the expected number of $B \in \mathcal{B}$ failing any of* (B1)–(B4) *is at most $\varepsilon^2 e(\mathcal{B})/4$.*

Now, by Lemma 6.3 and Markov's inequality, for $S \sim G(n,p)$,

$$\Pr[|\mathcal{B}_S \setminus \mathcal{B}'_S| \geq \varepsilon e(\mathcal{B})/2] \leq \frac{2\varepsilon^2 e(\mathcal{B})}{4\varepsilon e(\mathcal{B})} = \varepsilon/2.$$

Hence, since for $S \in \mathcal{S}$, $|\mathcal{B}_S| \geq \varepsilon e(\mathcal{B})$ and, by definition of $\mathcal{S}$, $\Pr[S \in \mathcal{S}'] \geq \varepsilon/2$. (Recall that $\mathcal{S}$ was defined as $S \sim G(n,p)$ satisfying the second implication of Theorem 3.16 with total probability at least ε.)

Proof of Lemma 6.3. Let X_1, X_2, X_3, and X_4 be the number of $B \in \mathcal{B}$ failing conditions (B1), (B2), (B3), and (B4), respectively. First,

$$\begin{aligned}\mathbb{E}[X_1] &= \sum_{B\in\mathcal{B}} \Pr[B \cap S \neq \emptyset] \\ &\leq \sum_{B\in\mathcal{B}} \mathbb{E}[|B \cap S|] \\ &= \sum_{v\in V(\mathcal{B})} p \cdot \deg_{\mathcal{B}}(v) \\ &= \Theta(e(\mathcal{B})) \cdot \Theta(n^{-1/2}) = o(e(\mathcal{B})),\end{aligned}$$

using that $\mathcal{B}$ has constant uniformity. Next, assuming that (B1) holds, the only bad case for (B2) is when the sets in $I(S,B)$ have size 1. So,

$$\mathbb{E}[X_2 - X_1] \leq \sum_{B\in\mathcal{B}} \binom{|B|}{2} \cdot p \leq e(\mathcal{B}) \cdot \binom{K}{2} \cdot \Theta(n^{-1/2}) = o(e(\mathcal{B})).$$

That is, for each booster $B \in \mathcal{B}$, for it to induce a set in $I(S,B)$ of size 1, there must be an interacting hyperedge between S and B intersecting B on 2 elements and intersecting S on 1 element. But, for any two elements of B, there is a single hyperedge in $\mathcal{H}$ containing them both, and the third vertex is in S with probability p.

Now, assuming that (B1) and (B2) hold, the only bad case for (B3) is when there exist hyperedges A, A' interacting between S and B with $|A \cap S| = |A' \cap S| = 2$ and $|A \cap A' \cap S| = 1$. For a given $B \in \mathcal{B}$, we can choose a vertex in B in $|B|$ ways, a hyperedge of $\mathcal{H}$ involving that vertex in $\Delta_1(\mathcal{H})$ ways, the vertex for the interacting edges to overlap in S on in 2 ways, and the other vertex in B of the other interacting edge in $|B|-1$ ways. So,

$$\mathbb{E}[X_3 - X_2 - X_1] \leq \sum_{B\in\mathcal{B}} |B| \cdot \Delta_1(\mathcal{H}) \cdot 2 \cdot (|B|-1) \cdot p^3 = O(n^{3/2}) = o(e(\mathcal{B})).$$

Finally, assuming that (B1), (B2), and (B3) hold, we bound $|I(S,B)|$. In order for $|I(S,B)| \geq \ell$ for some $\ell \in \mathbb{N}$, there must be $A_1, \ldots, A_\ell \in E(\mathcal{H})$ interacting between S and B. There are at most $\frac{(|B|\cdot\Delta_1(\mathcal{H}))^\ell}{\ell!}$ such sequences up to permutation, and, by our assumed properties, the union of each such sequence intersects S on exactly 2ℓ elements. So,

$$\Pr[|I(S,B)| \geq \ell] \leq \frac{p^{2\ell}(|B| \cdot \Delta_1(\mathcal{H}))^\ell}{\ell!} \leq \frac{K^\ell}{\ell!}.$$

Now, letting $\ell = L$ as in (B4), we have that $\Pr[|I(S,B)| \geq L] \leq \varepsilon^2/8$. Then,

$$\mathbb{E}[X_4 - X_3 - X_2 - X_1] \leq \sum_{B\in\mathcal{B}} \Pr[|I(S,B)| \geq L] \leq \varepsilon^2 e(\mathcal{B})/8.$$

Combining all of our bounds, we get

$$\mathbb{E}[X_1 + X_2 + X_3 + X_4] \leq \varepsilon^2 e(\mathcal{B})/8 + o(e(\mathcal{B})) \leq \varepsilon^2 e(\mathcal{B})/4,$$

as desired. ∎

6.2 Partitioning the colorings of S

Now we focus our attention on $S \in \mathcal{S}'$. Consider a proper 2-coloring of $\mathcal{H}[S]$, ψ. Since for all $B \in \mathcal{B}'_S$ we know $\mathcal{H}[S \cup B]$ is not properly 2-colorable, for any choice of a proper 2-coloring of $\mathcal{H}[B]$, φ, $\varphi \cup \psi$ admits a monochromatic edge in $\mathcal{H}[S \cup B]$.

But, none of this uses any of the properties (B1)–(B4)! For $v \in V(\mathcal{H})$, say that v is *forced* by the partial coloring ψ of $\mathcal{H}$ if v belongs to a hyperedge in $\mathcal{H}$ with both other vertices colored the same color by ψ. Let $F(\psi) \subseteq V(\mathcal{H})$ be the set of all vertices forced by ψ. One major consequence of properties (B1)–(B4) is that we can show results of the following kind.

Proposition 6.4. *For any $S \in \mathcal{S}'$ and any proper 2-coloring ψ of $\mathcal{H}[S]$, $|F(\psi)| = \Omega(v(\mathcal{H}))$.*

Proof. Let $S \in \mathcal{S}'$ and ψ be a proper 2-coloring of $\mathcal{H}[S]$. For each $B \in \mathcal{B}'_S$, $|F(\psi) \cap B| \geq 1$. Indeed, if this were not the case, then, by (B1), (B2), and (B3), for each $A \setminus B \in I(S, B)$, ψ colors the two vertices in $A \setminus B$ two different colors. So, namely, in any extension of ψ to a coloring also coloring B, the activated edges will not be monochromatic. But then, since $\mathcal{H}[B]$ is properly 2-colorable, for any proper 2-coloring φ of $\mathcal{H}[B]$, $\varphi \cup \psi$ is a proper 2-coloring of $\mathcal{H}[S \cup B]$, contradicting the fact that B is an activated booster.

Then, since each $v \in \mathcal{H}$ belongs to at most $e(\mathcal{B})/v(\mathcal{H})$ boosters (by symmetry of $\mathcal{B}$), we get

$$|F(\psi)| \geq |\mathcal{B}'_S| v(\mathcal{H})/e(\mathcal{B}) \geq \varepsilon v(\mathcal{H})/2 = \Omega(v(\mathcal{H})),$$

as desired. □

In Section 6.3, we will apply the hypergraph container lemma and the partial coloring supersaturation property (A5) to show that, in fact, any extension of a proper 2-coloring ψ of $\mathcal{H}[S]$ for $S \in \mathcal{S}'$ to a 2-coloring of all of $\mathcal{H}$ contains $\Omega(e(\mathcal{H}))$ monochromatic edges (Theorem 6.12). This is almost enough to conclude Theorem 6.1, except that we will need to apply a union bound over too many different 2-colorings of S. To preempt this issue, we will apply the hypergraph container lemma to partition the set of proper 2-colorings of S into a lower order size collection, setting up the application of the container lemma carefully so as to still be able to ultimately conclude something akin to Proposition 6.4.

6.2.1 The color forcing hypergraph

Given $S \in \mathcal{S}'$, our goal is to define a hypergraph $\mathcal{T}$ such that, for any partial 2-coloring of S, if it consistent with a proper 2-coloring of some $B \in \mathcal{B}'_S$, then there is some hyperedge in $E(\mathcal{T})$ demonstrating this fact. In this way, we will be able to relate independent sets of $\mathcal{T}$ to proper 2-colorings of S and make use of the hypergraph container lemma.

Note that a partial coloring ψ of S can be viewed as a subset of $S \times [2]$: for each $s \in S$ colored by ψ, include the pair $(s, \psi(S))$. For each $s \in S$ not colored by ψ, include both the pairs $(s, 1)$ and $(s, 2)$. We will use this set-based view of colorings in defining $\mathcal{T}$.

Definition 6.1 (The color forcing hypergraph)**.** We let $\mathcal{T}$ be the (multi)hypergraph with vertex set $S \times [2]$ with (multi)set of edges defined as follows. For each active booster $B \in \mathcal{B}'_S$ and proper 2-coloring φ of B, we add to $\mathcal{T}$ all hyperedges corresponding to (not necessarily proper) 2-colorings of $\cup_{A \setminus B \in I(S,B)} A \setminus B \subseteq S$ consistent with φ.

Let $B \in \mathcal{B}'_S$. As a result of our efforts in defining $\mathcal{B}'_S$, we can say a substantial amount about what 2-colorings of $\cup_{A \setminus B \in I(S,B)} A \setminus B$ consistent with φ look like. Let $T = A \setminus B \in I(S, B)$ and define

$$N_{\mathcal{H}}(T, B) := \{b \in B \,:\, T \cup \{b\} \in E(\mathcal{H})\}.$$

Namely, these are the vertices of B that are part of interacting edges between S and B, intersecting S on T.

Now, we know that $| \cup_{A \setminus B \in I(S,B)} A \setminus B| \leq 2L$, by (B4). Additionally, using (B1), (B2), and (B3), for each $\{t_1, t_2\} = T \in I(S, B)$, either t_1 and t_2 are colored different colors or, if they are colored the same color, then each element of $N_{\mathcal{H}}(T, B)$ is colored the opposite color under φ.

Formalizing the discussion above, the following is a simple but important consequence of Definition 6.1.

Lemma 6.5. *Let ψ be a proper 2-coloring of $\mathcal{H}[S]$. Then, when viewed as a subset of $S \times [2]$, ψ corresponds to an independent set in $\mathcal{T}$.*

Proof. Suppose that $\mathcal{T}[\psi]$ contains some edge T. Then, by Definition 6.1, there exists some $B \in \mathcal{B}'_S$ and proper 2-coloring φ of $\mathcal{H}[B]$ such that T encodes a proper 2-coloring ψ_φ of $\cup_{A\setminus B\in I(S,B)} A\setminus B$ consistent with φ. But, $T \subset \mathcal{T}[\psi]$ implies that ψ_φ and ψ in fact agree on $\cup_{A\setminus B\in I(S,B)} A\setminus B$. This implies that $\psi \cup \varphi$ is a proper 2-coloring of $\mathcal{H}[S \cup B]$, contradicting the fact that B is an activated booster. $\square$

By design, we also have a nice analogue of Proposition 6.4 in the setting of our color forcing hypergraph $\mathcal{T}$. Namely, subhypergraphs of $\mathcal{T}$ corresponding to partial colorings which force few vertices must contain many edges. This is extremely convenient for applications of the hypergraph container lemma (where we will be able to easily restrict the number of edges inside our containers).

Lemma 6.6. *If ψ is a partial 2-coloring of S, then*

$$e(\mathcal{T}[\psi]) \geq e(\mathcal{B}'_S) - |F(\psi)| \cdot \Delta_1(\mathcal{B}).$$

Proof. Let $B \in \mathcal{B}'_S$ such that $B \cap F(\psi) = \emptyset$. It suffices to show that B induces some edge in $\mathcal{T}[\psi]$.

For each $T \in I(S,B)$, ψ does not color both nodes in T the same color (or else the third vertex of any activated edge A between S and B including T would be forced). Hence, there is a coloring of $\cup_{T\in I(S,B)} T$ which extends ψ and, for each $T \in I(S,B)$ colors the two vertices in T different colors (this crucially uses properties (B1), and (B2), (B3)). Then, that coloring is in fact consistent with every proper 2-coloring φ of B (and, in particular, some coloring φ) and thus corresponds to an edge in $\mathcal{T}[\psi]$. $\square$

Remark 6.1. For the count in Lemma 6.6 to be accurate, it is crucial that $\mathcal{T}$ is defined as a multihypergraph. This is precisely why we need the hypergraph container lemma phrased more generally in terms of multihypergraphs (Theorem 4.3).

6.2.2 Applying the container lemma to the color forcing hypergraph

Our next goal will be to show that all proper 2-colorings of S are extensions of a relatively small family of partial 2-colorings of S. To do this, we want to apply the hypergraph container lemma to $\mathcal{T}$. There are two major obstacles to doing this. First, we would need to show that $\mathcal{T}$ satisfies the degree bound hypotheses in the container lemma (Theorem 4.3). Rather than showing this for $\mathcal{T}$, it will be enough for us to apply the container lemma to a subhypergraph $\mathcal{T}' \subseteq \mathcal{T}$. In any case, to get the necessary degree bounds, we will need to further restrict $\mathcal{S}'$. It will suffice to show that there is a relatively dense subhypergraph $\mathcal{T}' \subseteq \mathcal{T}$ satisfying

$$\Delta_1(\mathcal{T}') = O\left(\frac{e(\mathcal{T}')}{|S|}\right)$$

and

$$\Delta_2(\mathcal{T}') = o\left(\frac{e(\mathcal{T}')}{|S|}\right).$$

We will use the bound on the $\Delta_2(\mathcal{T}')$ to derive bounds on $\Delta_\ell(\mathcal{T}')$ for $\ell \in [2L]$ such that $\ell \geq 3$. (Observe that we are exploiting (B4) to have an upper bound on the size of edges in $\mathcal{T}'$.)

Note that $O(e(\mathcal{T}')/|S|)$ is of the order of the average degree of $\mathcal{T}'$. Hence, our goal is merely to remove any dramatic disparities between the maximum and average degrees. It turns out that this is easily achievable by repeatedly removing edges from sets of nodes of maximum degree. We can bound the effect of this greedy algorithm using the sum of the squares of degrees.

Proposition 6.7. *Let $\mathcal{G}$ be a (multi)hypergraph. For all $t, m \in \mathbb{N}$, there is $\mathcal{G}' \subseteq \mathcal{G}$ with $e(\mathcal{G}') \geq e(\mathcal{G}) - m$ and*

$$\sum_{T\in\binom{V(\mathcal{G})}{t}} \deg_{\mathcal{G}}(T)^2 \geq m \cdot \Delta_t(\mathcal{G}').$$

Proof. Assume $m \leq e(\mathcal{G})$ or else the result is trivial. Fix $t \in \mathbb{N}$. We obtain $\mathcal{G}'$ by removing edges from $T \in \binom{V(\mathcal{G})}{t}$ with maximum degree. Think of

$$\sum_{T\in\binom{V(\mathcal{G})}{t}} \deg_{\mathcal{G}}(T)^2$$

as each $T \in \binom{V(\mathcal{G})}{t}$ contributing its degree, $\deg_{\mathcal{G}}(T)$, to each of the $\deg_{\mathcal{G}}(T)$ hyperedges it belongs to. Then, whenever we remove an edge, it removes a contribution of at least $\Delta_t(\mathcal{G}')$ from this sum (since otherwise we should have removed a different edge). This yields the desired result. □

The following technical lemma allows us to take advantage of Proposition 6.7. In the lemma, p is drawn from the same sequence of probabilities we have been using.

Lemma 6.8 (Lemmas 5.7 and 5.8 of [FKSS22]). *If $S \sim G(n,p)$, then:*

$$\mathbb{E}\left[\sum_{v \in V(\mathcal{T})} \deg_{\mathcal{T}}(v)^2\right] \leq \frac{C \cdot e(\mathcal{B})^2}{p \cdot v(\mathcal{H})}, \tag{1}$$

and

$$\mathbb{E}\left[\sum_{T \in \binom{V(\mathcal{T})}{2}} \deg_{\mathcal{T}}(T)^2\right] \leq \frac{\sigma \cdot e(\mathcal{B})^2}{p \cdot v(\mathcal{H})} \tag{2}$$

where $\sigma = o(1)$ and C a constant depending on K and L.

We suppress the proof of Lemma 6.8. Now, by Markov's inequality, these expectation bounds will give us the desired degree bounds on $\mathcal{T}$. Indeed, let $\tau = \sigma^{1/(2L+1)}$, where σ is the parameter in (2). We restrict $\mathcal{S}'$ to $\mathcal{S}''$ by only retaining the $S \in \mathcal{S}'$ such that both the following conditions hold:

$$\sum_{v \in V(\mathcal{T})} \deg_{\mathcal{T}}(v)^2 \leq \frac{16\varepsilon^{-1} C e(\mathcal{B})^2}{|S|},$$

and

$$\sum_{T \in \binom{V(\mathcal{T})}{2}} \deg_{\mathcal{T}}(T)^2 \leq \frac{\tau^{2L} e(\mathcal{B})^2}{|S|}.$$

By the Chernoff bound and Markov's inequality,

$$\Pr(S \in \mathcal{S}'') \geq \Pr[S \in \mathcal{S}'] - \Pr[|S| \geq 2pv(\mathcal{H})] - \varepsilon/8 - 2\tau \geq \varepsilon/4,$$

for large enough n. With our final restriction to $\mathcal{S}$ made, we are finally ready to apply the container lemma.

Lemma 6.9. *Let $S \in \mathcal{S}''$. Then, there exists a family $\mathcal{C}$ of partial colorings of S such that:*

1. *All proper 2-colorings ψ of S are extensions of some coloring in $\mathcal{C}$.*
2. *The family $\mathcal{C}$ satisfies $|\mathcal{C}| = \exp(o(|S|))$.*
3. *Every partial coloring ψ in $\mathcal{C}$ forces $\Omega(v(\mathcal{H}))$ vertices.*

Note that the trivial bound on the size of such a family $\mathcal{C}$ is $2^{|S|}$, so we are exploiting the interdependency of the proper 2-colorings in some nontrivial way.

Proof of Lemma 6.9. Fix $S \in \mathcal{S}''$. By applying Proposition 6.7 twice to $\mathcal{T}$ with $m = \varepsilon \cdot e(\mathcal{B})/16$ and $t = 1$ and then $t = 2$, we get $\mathcal{T}' \subseteq \mathcal{T}$ with $e(\mathcal{T}') \geq e(\mathcal{T}) - \varepsilon \cdot e(\mathcal{B})/8$ such that

$$\Delta_1(\mathcal{T}') \leq \frac{2^8 \varepsilon^{-2} C \cdot e(\mathcal{B})}{|S|}$$

and

$$\Delta_2(\mathcal{T}') \leq \frac{16\varepsilon^{-1} \cdot \tau^{2L} \cdot e(\mathcal{B})}{|S|}.$$

Note that then

$$e(\mathcal{T}') \leq 2|S| \cdot \Delta_1(\mathcal{T}') \leq 2^9 \varepsilon^{-2} C \cdot e(\mathcal{B}),$$

since there are $2|S|$ vertices in $\mathcal{T}'$.

An issue emerges. The container lemma as phrased in Theorem 4.3 only applies to uniform (multi)hypergraphs, but $\mathcal{T}'$ has hyperedges of varying even sizes between 2 and $2L$. Fortunately, there is an easy fix. Partition $\mathcal{T}'$ into L distinct u-uniform subhypergraphs, $\mathcal{T}_u$ for $u \in [2L]$ even. This covers all of the hyperedges in $\mathcal{T}'$ by (B2), (B3), and (B4). Each independent set in $\mathcal{T}$ is also an independent set in $\mathcal{T}'$ and $\mathcal{T}_u$ for each u. When we apply the hypergraph container lemma to one of the hypergraphs $\mathcal{T}_u$, the resulting collection of containers corresponds exactly to a set of partial 2-colorings of S such that every proper 2-coloring of S is an extension of one of the partial 2-colorings.

While we could apply the container lemma to any single $\mathcal{T}_u$ to get the first two implications of Lemma 6.9, we need to tread carefully to simultaneously achieve the third implication. We want to make use of Lemma 6.6, so we need to control the number of edges our containers induce in $\mathcal{T}'$. We will achieve this by intersecting the containers resultant from applying the container lemma to each distinct $\mathcal{T}_u$.

We now make this concrete. We apply the container lemma to each of subhypergraphs with sufficiently many edges. Namely, define

$$U = \{u \in [2L] \,:\, \varepsilon \cdot e(\mathcal{B})/(16L) < e(\mathcal{T}_u)\}.$$

Observe that $e(\mathcal{T}) \geq \varepsilon e(\mathcal{B})/4$ since each $S \in \mathcal{S}''$ has at least that many activated boosters. Hence we also have $e(\mathcal{T}') \geq \varepsilon e(\mathcal{B})/8$. Now, fix $u \in U$. We check that it satisfies the hypotheses of the container lemma. First,

$$\Delta_1(\mathcal{T}_u) \leq \Delta_1(\mathcal{T}') \leq \frac{2^8\varepsilon^{-2}C \cdot e(\mathcal{B})}{|S|} \leq \frac{2^{13}\varepsilon^{-3}CL \cdot e(\mathcal{T}_u)}{v(\mathcal{T}_u)}.$$

Also, for $\ell \in [2, 2L]$,

$$\Delta_\ell(\mathcal{T}_u) \leq \Delta_2(\mathcal{T}_u) \leq \Delta_2(\mathcal{T}') \leq \frac{2^5\varepsilon^{-1}\tau^{\ell-1}e(\mathcal{B})}{v(\mathcal{T}_u)} \leq \frac{2^9\varepsilon^{-2}L\tau^{\ell-1}e(\mathcal{T}_u)}{v(\mathcal{T}_u)}.$$

So, choosing appropriate constant parameters $\varepsilon_{\mathcal{T}}, K_{\mathcal{T}}$ in Theorem 4.3 (depending only on ε, L, C), we get a collection $\mathcal{C}_u$ of partial colorings (the image of $f : \mathcal{P}(V(\mathcal{T}))^t \to \mathcal{P}(V(\mathcal{T}))$ applied to all t-tuples of at most $\tau v(\mathcal{T})$-element subsets) such that:

1. For each proper 2-coloring ψ of S, there exists $C \in \mathcal{C}_u$ so that $\psi \subseteq C$. (That is, ψ is a extension of the partial coloring of S corresponding to C.)

2. For each $C \in \mathcal{C}_u$, $e(\mathcal{T}_u[C]) \leq \varepsilon e(\mathcal{B})/(16L)$.

3. $|\mathcal{C}_u| \leq \left(\frac{ev(\mathcal{T}_u)}{\tau v(\mathcal{T}_u)}\right)^{\tau v(\mathcal{T}_u)} = \exp(o(|S|))$.

Finally, let $\mathcal{C}$ be the collection of $|U|$-wise intersections of containers, one from each $\mathcal{C}_u$, $u \in U$. (U is nonempty by construction by the pigeonhole principle.) Since $|U| \leq 2L$ and L is constant, $|\mathcal{C}| = \exp(o(|S|))$. Each proper 2-coloring ψ of S is contained in some container in $\mathcal{C}_u$ for each $u \in U$, and, hence, is contained in some container in $\mathcal{C}$.

The last thing to check is that $\mathcal{C}$ achieves the third implication of the lemma. By the container lemma and definition of U, each $\psi \in \mathcal{C}$ induces at most $2L \cdot \varepsilon e(\mathcal{B})/16L = e(\mathcal{B})/8$ edges in $\mathcal{T}'$ and then at most $\varepsilon e(\mathcal{B})/8 + \varepsilon e(\mathcal{B})/8 = \varepsilon e(\mathcal{B})/4$ edges in $\mathcal{T}$. Then, by Lemma 6.6, since $|\mathcal{B}'_S| \geq \varepsilon e(\mathcal{B})/2$,

$$|F(\psi)| \geq \frac{\varepsilon e(\mathcal{B})}{4\Delta_1(\mathcal{B})} = \Omega(v(\mathcal{H})),$$

by symmetry of $\mathcal{B}$. □

6.3 Turning forced vertices into monochromatic edges

The goal of this section is to apply our partial coloring supersaturation property (A5) to show that, for any $\psi \in \mathcal{C}$ resultant from Lemma 6.9, not only does ψ force $\Omega(v(\mathcal{H}))$ vertices, but ψ also forces $\Omega(e(\mathcal{H}))$

monochromatic hyperedges in any extension to a 2-coloring of all of $\mathcal{H}$. As a reminder, we refer to the structures which induce monochromatic hyperedges as *forced triangles*, see Definition 5.1.

To do this, we require an intermediate theorem and a technical lemma. The theorem uses the container lemma to transfer (A5) to the sparse setting (and to be able to apply the container lemma, the proof relies crucially on (A2)).

Theorem 6.10. *Let $S \sim G(n,p)$. With probability $1-o(1)$, for every partial 2-coloring ψ of S such that ψ colors two out of three vertices of $\Omega(v(\mathcal{H}))$ hyperedges the same color, ψ induces $\Omega(e(\mathcal{H}))$ forced triangles.*

The next step is to say the $\Omega(e(\mathcal{H}))$ forced triangles actually amount to $\Omega(e(\mathcal{H}))$ monochromatic edges in any extension of the partial 2-coloring to a 2-coloring of $\mathcal{H}$. For $T \in E(\mathcal{H})$ and $S \subseteq V(\mathcal{H})$ let $f(T,S)$ be the number of 4-tuples of hyperedges (T, T_1, T_2, T_3) such that $T_1, T_2, T_3 \in E(\mathcal{H}[S])$ and T intersects each of T_1, T_2, and T_3 on one vertex and, and $|T \cup T_1 \cup T_2 \cup T_3| = 9$. Namely, each such 4-tuple is a minimally intersecting configuration of hyperedges in S which could, when colored, yield a forced triangle centered at T. For an example of this, see Figure 5. (In this example, the red vertices are those that could be colored to yield a forced triangle.) We can upper bound the expectation of the sum of the squares of $f(T,S)$ for any given $S \sim G(n,p)$.

Lemma 6.11. *Let $S \sim G(n,p)$. Then,*

$$\mathbb{E}\left[\sum_{T\in E(\mathcal{H})} f(T,S)^2\right] = O(e(\mathcal{H})).$$

Using Theorem 6.10, Lemma 6.11, and the Cauchy-Schwarz inequality we can then prove the main result of this section.

Theorem 6.12 (Forced vertices to forced monochromatic edges)**.** *Let $\varepsilon > 0$ be a constant, and let $S \sim G(n,p)$. Then, with probability at least $1-\varepsilon$, for every partial 2-coloring ψ of S, if $|F(\psi)| = \Omega(v(\mathcal{H}))$, then ψ induces $\Omega(e(\mathcal{H}))$ distinct centers of forced triangles.*

Proof. Let $S \sim G(n,p)$. By Lemma 6.11 and Markov's inequality, there exists a constant C such that

$$\sum_{T\in E(\mathcal{H})} f(T,S)^2 \le Ce(\mathcal{H})$$

with probability at least $1-\varepsilon/2$. Moreover, the implication of Theorem 6.10 holds with probability $1-\varepsilon/2$ (for large enough n). So, with probability at least $1-\varepsilon$, S has both of these properties.

Now, let ψ be a partial 2-coloring of such an S with $|F(\psi)| = \Omega(v(\mathcal{H}))$. Then, since the implication of Theorem 6.10 holds, ψ induces $\Omega(e(\mathcal{H}))$ forced triangles. Let $\mathcal{H}_\psi \subseteq \mathcal{H}$ composed only of edges T with $f(T,S) \ge 1$. Observe that $e(\mathcal{H}_\psi)$ is the number of distinct centers of forced triangles induced by ψ. Now, by definition of $\mathcal{H}_\psi$,

$$\sum_{T\in\mathcal{H}_\psi} f(T,S) = \Omega(e(\mathcal{H})).$$

But, by the Cauchy-Schwarz inequality,

$$\begin{aligned}\left(\sum_{T\in\mathcal{H}_\psi} f(T,S)\right)^2 &\le e(\mathcal{H}_\psi)\cdot\sum_{T\in\mathcal{H}_\psi} f(T,S)^2\\ &\le e(\mathcal{H}_\psi)\cdot C\cdot e(\mathcal{H}),\end{aligned}$$

with the last inequality using our assumption on S. Combining the above implies $e(\mathcal{H}_\psi) = \Omega(e(\mathcal{H}))$, as desired.

□

Now we prove Lemma 6.11. The proof follows from a simple degree of freedom argument after expanding the defintion of the expectation.

Proof of Lemma 6.11. Fix $T \in \mathcal{H}$. Define $\mathcal{F}_T$ to be the collection of subsets of $V(\mathcal{H})$ corresponding to the union of supports of non-T vertices of possible forced triangles centered at T. That is, $\mathcal{F}_T$ is the collection of $T_1 \cup T_2 \cup T_3 \setminus T$ for $T_1, T_2, T_3 \in E(\mathcal{H})$ such that (T, T_1, T_2, T_3) contributes to $f(T, S)$ (e.g., the red vertices in Figure 5).

Then, expanding the definition of $f(T, S)$ as a sum of indicator random variables, we have

$$\mathbb{E}[f(T,S)^2] = \sum_{F, F' \in \mathcal{F}_T} p^{\mathrm{supp}(F) \cup \mathrm{supp}(F')}.$$

Each $F, F' \in \mathcal{F}_T$ can overlap on 0, 1, 2, or all 3 of the outer hyperedges composing them. In each case, they overlap on 0, 2, 4, and or all 6 vertices in their supports, respectively. There are $O(n^6)$, $O(n^5)$, $O(n^4)$, and $O(n^3)$ instances of each type of pair respectively, yielding a total contribution of

$$O(n^6 \cdot p^{12} + n^5 \cdot p^{10} + n^4 \cdot p^8 + n^3 \cdot p^6) = O(1),$$

since $p = \Theta(p_{\mathcal{H}}) = \Theta(n^{-1/2})$. This yields the desired result by linearity of expectation. □

The proof of Theorem 6.10 is rather technical, so we only include a sketch of the proof here.

Sketch of the proof of Theorem 6.10. The first step is to define a 6-uniform hypergraph $\mathcal{R}$ with vertex set $V(\mathcal{H}) \times [2]$ and hyperedges encoding all of the sets in $\mathcal{F}_T \times [2]$ for each T from the proof of Lemma 6.11. Namely, the hyperedges of $\mathcal{R}$ correspond to the unions of vertices involved in all possible forced triangles in $\mathcal{H}$ (excluding the vertices in the central triangle), along with the (shared) color of the vertices in the forced triangle. It is simple to show via degree of freedom counting arguments that $e(\mathcal{R}) = \Theta(n^6)$, and, for $1 \le t \le 6$,

$$\Delta_t(\mathcal{R}) = O(n^{5-t}) = O\left(p_{\mathcal{H}}^{t-1} \cdot \frac{e(\mathcal{R})}{v(\mathcal{R})}\right).$$

This means we can apply the hypergraph container lemma (Theorem 4.3) to $\mathcal{R}$ with $\tau = p_H$ to get a relatively small family $\mathcal{C}$ of containers (each corresponding to partial colorings) containing all partial colorings of $V(\mathcal{H})$ inducing few forced triangles. Crucially, (A5) implies that for each partial 2-coloring ψ of $V(\mathcal{H})$ with $\psi \subseteq C \in \mathcal{C}$ we have a strong bound on the number of hyperedges with two vertices colored the same color.

The next step is to suppose that $S \sim G(n, p)$ violates the statement for the theorem. That is, let S be such that there is a partial 2-coloring ψ of S coloring at least two out of three vertices the same color in at least $c_1 v(\mathcal{H})$ hyperedges, but simultaneously inducing at most $c_2 e(\mathcal{H})$ forced triangles (for constants c_1, c_2 suitably defined with respect to the constants resultant from the container lemma). Choosing c_2 such that $\psi \subset C$ for some $C \in \mathcal{C}$, one can argue that the random sample is "biased" toward C. It is possible to show that such a bias toward a given container C is sufficiently unlikely (using concentration inequalities such as Janson's inequality, see Theorem 2.1) so that, even after union bounding over all containers, the probability of such a bias toward any container is $o(1)$. □

6.4 Tying it all together

As a reminder, our goal is to derive a contradiction with the dichotomy version of Friedgut's criterion, Theorem 3.16. By Theorem 6.12, choosing the ε for the theorem to be $\varepsilon/8$, with probability at least $1 - \varepsilon/8$, for $S \sim G(n, p)$, for every partial 2-coloring $\psi \in \mathcal{C}$, ψ induces $\Omega(e(\mathcal{H}))$ distinct centers of forced triangles. Then, since $\Pr[S \in \mathcal{S}''] \ge \varepsilon/4$, by Lemma 6.9, with probability at least $\varepsilon/8$, there exists a family $\mathcal{C}$ of partial colorings of S such that:

1. All proper 2-colorings ψ of S are extensions of some coloring in $\mathcal{C}$.
2. The family $\mathcal{C}$ satisfies $|\mathcal{C}| = \exp(o(|S|))$.
3. Every partial coloring $\psi \in \mathcal{C}$ induces $\Omega(e(\mathcal{H}))$ distinct centers of forced triangles.

By the Chernoff bound, for any constant $c > 1$, adding the assumption that

$$pv(\mathcal{H})/c \le |S| \le cpv(\mathcal{H})$$

only decreases the probability of this event to at least $\varepsilon/8 - o(1)$.

Now, suppose that we could show that, for a given $\psi \in \mathcal{C}$, $G(n, \varepsilon p) \cup S$ does not include one of the distinct centers of a forced triangle induced by ψ with probability at most P_J. Then, union bounding over all $\psi \in \mathcal{C}$, we would get that $G(n, \varepsilon p) \cup S$ does not include one of distinct centers for some ψ with probability at most $P_J \cdot \exp(o(|S|))$. Since all proper 2-colorings of $G(n, \varepsilon p) \cup S$ are extensions of proper 2-colorings of S, and these are in turn extensions of some partial 2-coloring $\psi \in \mathcal{C}$, this would imply that $G(n, \varepsilon p) \cup S$ is properly 2-colorable with probability at most $P_J \exp(o(|S|))$. That is, $S \cup G(n, \varepsilon p)$ is 2-Ramsey for a triangle with probability at least $1 - P_J \exp(o(|S|))$. But, since $S \in \mathcal{S}'' \subseteq \mathcal{S}$, $S \cup G(n, \varepsilon p)$ should be 2-Ramsey for a triangle with probability at most 1/2. Hence, if we show $P_J = \exp(-\Omega(|S|)) = \exp(-\Omega(p \cdot v(\mathcal{H})))$, we will get the desired contradiction, proving Theorem 6.1.

The last ingredient of the proof of Theorem 6.1 is then showing that

$$P_J = \exp(-\Omega(p \cdot v(\mathcal{H}))) = \exp(-\Omega(n^{3/2})).$$

The appropriate tool for the job turns out to be Janson's inequality. Using Janson's inequality, we prove the following.

Lemma 6.13. *Let $\mathcal{E} \subset E(\mathcal{H})$ with $|\mathcal{E}| = \Omega(e(\mathcal{H}))$. Then,*

$$\Pr[E \not\subseteq G(n, \varepsilon p) \text{ for all } E \in \mathcal{E}] \le \exp(-\Theta(n^{3/2})).$$

Proof. Enumerate the edges in $\mathcal{E}$ and let A_i be the event that $E_i \subseteq G(n, \varepsilon p)$. Now,

$$\mathrm{Var}'(A_1, \ldots, A_{|\mathcal{E}|}) \le e(\mathcal{H})p^3 + v(\mathcal{H})n^2 p^5 = \Theta(n^{3/2}).$$

The first term is the contribution from edges and themselves, and the second is from pairs of hyperedges intersecting on a single vertex. (The latter case corresponds to two triangles overlapping on an edge.) Letting $X = \sum_{i \in [|\mathcal{E}|]} 1_{A_i}$, $\mathbb{E}[X] = \Theta(p^3 e(\mathcal{H})) = \Theta(n^{3/2})$. Then, by Janson's inequality,

$$\Pr[X = 0] \le \exp(-\Theta(n^{3/2})),$$

as desired. □

The fact that $P_J = \exp(-\Omega(n^{3/2}))$ then follows, completing the proof of Theorem 6.1.

7 Closing Remarks

Over the course of this exposition, we showed that the property of being 2-Ramsey for a triangle has a sharp threshold (Theorem 6.1). A number of powerful and widely applicable tools arose along the way. This includes:

- Friedgut's criterion for symmetric monotone properties with a coarse threshold (Theorem 3.2) and its more practical dichotomy corollary (Theorem 3.16).
- Boolean Fourier analytic methods (introduced in Definition 3.1 and Proposition 3.3).
- The hypergraph container lemma, at least as stated in [FKSS22] (Theorem 4.3).
- Jansen's concentration inequality (Theorem 2.1).

Although our discussion was focused on random graphs for simplicity, most of it immediately generalizes to random sets (with each element added independently at random from some universe set with probability p). Indeed, some of the earliest work studying probability thresholds was in this setting ([Sch16, VdW27]).

Our proof of Theorem 6.1 is based on the more general argument given in [FKSS22]. In that work, they show that r-colorability of any hypergraph satisfying suitably generalized versions of (A1)–(A5) has a sharp threshold. This implies sharp thresholds for a wide range of properties. For example, an argument akin to Section 6 can be used to prove that the properties of being r-Ramsey for K_k, for any fixed k, and r-Ramsey for H bounded and nearly-bipartite[1] have sharp thresholds.

While much is now known about monotone graph properties with a sharp threshold, there is considerable room for future work. One avenue is to consider thresholds of properties of more structured random objects. For example, in the $G(n,p)$ model, all of the random graphs are subgraphs of K_n. Another perfectly valid model is to start with some other family of "host" graphs G on n vertices and keep each edge independently with probability p. Perhaps the most basic property to consider is connectivity, and problems of this kind have already been studied: for hypercubes, there is a sharp threshold at $p = 1/2$ [Bur77]; for Cartesian products of complete graphs, so called "Hamming graphs," a sharp threshold is known [FvdHH16]; and for $\Omega(\sqrt{n})$-regular, $\Omega(\sqrt{n})$-edge connected graphs, a sharp threshold is also known [GLL21]. Although the graph families considered in [GLL21] are more general than in the other works, they are still significantly constrained. One interesting future direction would be to try and characterize the families of host graphs for which connectivity has a sharp threshold. The answer is not *all* graphs as illustrated by the following example.

Example 7.1 (Host graphs without a sharp threshold for connectivity)**.** Consider the family of graphs composed of two disjoint copies of $K_{n/2}$ with a cut edge between them. In order for a subgraph of this family to be connected, it must contain the cut edge and must be connected on either side of the cut edge. The property of containing the cut edge evidently has a coarse threshold at $p = \Theta(1)$. Moreover, connectivity on either side of the cut edge is ensured for $p = \omega(\frac{\log n}{n})$.

The example still holds if $f(n) = o(n/\log n)$ edges are added between the disjoint copies of $K_{n/2}$. In this case, in order to have one of those edges with constant probability, we need $p = \Omega(1/f(n)) = \omega(\frac{\log n}{n})$. The threshold will also still be coarse around $p(n) = 1/f(n)$.

Notably, none of the previous work on these problems made use of the hypergraph container method. One promising aspect of the container method in this context is that the subgraphs certifying connectivity can potentially be small, unlike in the general case, facilitating the creation of bounded uniformity encoding hypergraphs as in Section 6.2.

Imposing an ambient metric space is another way to add structure to the $G(n,p)$ model. Random geometric graphs are an example of this. One model is to build a graph by selecting n points from the unit disk, independently and uniformly at random, connecting them by an edge if their Euclidean distance is at most r. The threshold parameter is then r rather than the p from before. A sharp threshold for connectivity is known in this setting [GK99], and thresholds for other properties have been studied, for example planarity [BKMS20]. However, for properties other than connectivity, sharpness of the threshold probabilities is poorly understood, even though sometimes the sharpness of the respective properties in the $G(n,p)$ model is known (e.g., in [ŁPW94] they prove that there is a sharp threshold for planarity at $p = 1/n$). Once again, perhaps the hypergraph container method provides a way forward.

References

[BKMS20] Ahmad Biniaz, Evangelos Kranakis, Anil Maheshwari, and Michiel Smid. Plane and planarity thresholds for random geometric graphs. *Discrete Mathematics, Algorithms and Applications*, 12(01):2050005, 2020.

[BMS15] József Balogh, Robert Morris, and Wojciech Samotij. Independent sets in hypergraphs. *Journal of the American Mathematical Society*, 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*, pages 3059–3092. World Scientific, 2018.

[1] A graph G is nearly bipartite if there exists $e \in E(G)$ such that $G \setminus e$ is bipartite.

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

[Bur77] Ju D Burtin. The probability of connectedness of a random subgraph of an n-dimensional cube. *Problemy Peredaci Informacii*, 13(2):90–95, 1977.

[ER60] Paul Erdős and Alfréd Rényi. On the evolution of random graphs. *A Magyar Tudományos Akadémia*, 5(1):17–60, 1960.

[FB$^+$99] Ehud Friedgut, Jean Bourgain, et al. Sharp thresholds of graph properties, and the k-sat problem. *Journal of the American Mathematical Society*, 12(4):1017–1054, 1999.

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

[FR86] Peter Frankl and Vojtech Rödl. Large triangle-free subgraphs in graphs without K_4. *Graphs and Combinatorics*, 2(1):135–144, 1986.

[FRRT03] Ehud Friedgut, Vojtech Rödl, Andrzej Rucinski, and Prasad Tetali. A sharp threshold for random graphs with a monochromatic triangle in every edge coloring. In *Memoirs of the American Mathematical Society*, 2003.

[FvdHH16] Lorenzo Federico, Remco van der Hofstad, and Tim Hulshof. Connectivity threshold for random subgraphs of the hamming graph. 2016.

[GK99] Piyush Gupta and Panganamala R Kumar. Critical power for asymptotic connectivity in wireless networks, 1999.

[GLL21] Anupam Gupta, Euiwoong Lee, and Jason Li. The connectivity threshold for dense graphs. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 89–105. SIAM, 2021.

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

[ŁPW94] Tomasz Łuczak, Boris Pittel, and John C Wierman. The structure of a random graph at the point of the phase transition. *Transactions of the American Mathematical Society*, 341(2):721–748, 1994.

[LRV92] Tomasz Luczak, Andrzej Ruciński, and Bernd Voigt. Ramsey properties of random graphs. *Journal of Combinatorial Theory, Series B*, 56(1):55–68, 1992.

[RR93] Vojtech Rödl and Andrzej Rucinski. Lower bounds on probability thresholds for Ramsey properties. *Combinatorics, Paul Erdos is eighty*, 1(317-346):8–40, 1993.

[RR94] Vojtech Rödl and Andrzej Ruciński. Random graphs with monochromatic triangles in every edge coloring. *Random Structures & Algorithms*, 5(2):253–270, 1994.

[RR95] Vojtěch Rödl and Andrzej Ruciński. Threshold functions for Ramsey properties. *Journal of the American Mathematical Society*, 8(4):917–942, 1995.

[Sch16] I Schur. Über die kongruenz $x^m + y^m \equiv z^m \pmod{p}$. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 25:114–116, 1916.

[SS18] Mathias Schacht and Fabian Schulenburg. Sharp thresholds for Ramsey properties of strictly balanced nearly bipartite graphs. *Random Structures & Algorithms*, 52(1):3–40, 2018.

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

[Tal94] Michel Talagrand. On Russo's approximate zero-one law. *The Annals of Probability*, pages 1576–1587, 1994.

[Tur41] Paul Turán. On an extremal problem in graph theory. *Matematikai és Fizikai Lapok*, 48:436–452, 1941.

[VdW27] Bartel Leendert Van der Waerden. Beweis einer baudetschen vermutung. *Nieuw Arch. Wiskunde*, 15:212–216, 1927.