

AN EXPONENTIAL UPPER BOUND FOR INDUCED RAMSEY NUMBERS

LUCAS ARAGÃO, MARCELO CAMPOS, GABRIEL DAHIA,
RAFAEL FILIPE, AND JOÃO PEDRO MARCIANO

ABSTRACT. The induced Ramsey number $R_{\text{ind}}(H; r)$ of a graph H is the minimum number N such that there exists a graph with N vertices for which all r -colourings of its edges contain a monochromatic induced copy of H . Our main result is the existence of a constant $C > 0$ such that, for every graph H on k vertices, these numbers satisfy

$$R_{\text{ind}}(H; r) \leq r^{Crk}.$$

When $r = 2$, this resolves a conjecture of Erdős from 1975. For $r > 2$, it answers a question of Conlon, Fox and Sudakov in a strong form.

1. INTRODUCTION

The Ramsey number of a graph H , denoted by $R(H)$, is the minimum number N for which every red/blue colouring of the edges of K_N , the complete graph on N vertices, contains a monochromatic copy of H . Ramsey [28] showed that $R(H)$ is finite for every graph H , with Erdős and Szekeres [18] providing the first explicit upper bound in 1935. In an influential paper from 1947, Erdős [15] proved an exponential lower bound for the case $H = K_k$, which established

$$2^{k/2} \leq R(K_k) \leq 4^k. \tag{1}$$

Since then, the lower bound of Erdős has only been improved by a constant factor [34]. On the other hand, the upper bound in (1) has seen several improvements over the years [35, 9, 30], culminating in an exponential improvement by Campos, Griffiths, Morris, and Sahasrabudhe [8]. A second (much shorter) proof of this result was given in [2], which also extended it to the setting of r -colourings, though only for fixed r and sufficiently large k .

Another major open problem in the area is the natural extension of $R(H)$ to the setting of induced subgraphs. We will write $G \xrightarrow{\text{ind}} H$ to denote the following property: for any red/blue colouring of the edges of G , there exists an induced monochromatic copy of H (that is, a copy of H which is induced in G , and all of its edges have the same colour). We then define

$$R_{\text{ind}}(H) = \min \{v(G) : G \xrightarrow{\text{ind}} H\}.$$

In particular, observe that we have $R_{\text{ind}}(K_k) = R(K_k)$ for every $k \in \mathbb{N}$, since every copy of K_k in a graph G is also an induced subgraph of G . For general graphs H , on the other hand, Erdős [17] remarked that even “the existence of [the induced Ramsey number] is not at all obvious.”

Deuber [14], Erdős, Hajnal, and Pósa [19] and Rödl [29] independently established in the 1970s that $R_{\text{ind}}(H)$ is finite for every graph H . While none of these works provide an explicit dependency

During this work, Marcelo Campos was supported by Serrapilheira (grant R-2412-51283), Rafael Filipe was supported by CNPq, and João Pedro Marciano was supported by a FAPERJ Bolsa Nota 10.

on k , the number of vertices of H , Erdős [17] later remarked that the best bound that one can deduce from the proofs in [14, 19, 29] is of the form

$$R_{\text{ind}}(H) \leq 2^{2^{k^{1+o(1)}}}.$$

Nevertheless, Erdős [16, 17] conjectured, first implicitly in 1975 and then explicitly in 1984, that the function $R_{\text{ind}}(H)$ should grow at most exponentially as a function of $v(H) = k$ for every H . Note that, if true, this would be best possible, since we have $R_{\text{ind}}(K_k) = R(K_k) \geq 2^{k/2}$ by (1).

Using the techniques of Rödl [29], one can prove Erdős' conjecture for bipartite H . The problem is much harder for non-bipartite H , however, and the next significant advance was not obtained until almost 25 years later, by Kohayakawa, Prömel, and Rödl [23]. By taking G (the host graph) to be a random graph built using projective planes, they showed, among other results, that

$$R_{\text{ind}}(H) \leq k^{O(k \log k)} \quad (2)$$

for every graph H with k vertices. Fox and Sudakov [20] later provided an explicit, pseudorandom graph attaining the bound in (2). To achieve this, they pioneered a versatile approach to Ramsey-type theorems in the setting where one fixed H is forbidden as an induced subgraph, which became influential for other notorious conjectures, like the Erdős–Hajnal problem (see e.g. [6, 26, 27]).

A few years later, Conlon, Fox, and Sudakov [11] removed a factor of $\log k$ from the exponent in (2) and showed, using an explicit graph, that

$$R_{\text{ind}}(H) \leq k^{O(k)}. \quad (3)$$

In order to prove (3), the authors of [11] developed a general method for proving Ramsey-type theorems using pseudorandom properties of the host graph G . For example, their method also allowed them to improve a celebrated result of Graham, Rödl, and Ruciński [22] on the Ramsey numbers of bounded degree graphs.

The main result of this paper confirms Erdős' conjecture for all graphs H .

Theorem 1.1. *There exists a constant $C > 0$ such that*

$$R_{\text{ind}}(H) \leq 2^{Ck}$$

for every graph H with k vertices.

We will also consider the induced Ramsey number for r -colourings. Define $R_{\text{ind}}(H; r)$ to be the r -colour induced Ramsey number of a graph H ; that is, the minimum number of vertices of a graph G such that every r -colouring $c: E(G) \rightarrow [r]$ of the edges of G contains an induced monochromatic copy of H . The techniques used in [11] and [23] do not work in this more general setting, and provide no bounds when $r \geq 3$, but Fox and Sudakov [21] introduced a different approach in 2009, which can be used to show that¹

$$R_{\text{ind}}(H; r) \leq r^{O(rk^2)}. \quad (4)$$

The large gap between (4) and the known bounds for the $r = 2$ case motivated Conlon, Fox, and Sudakov [12, Problem 3.5] to ask if one could show that, for fixed $r \in \mathbb{N}$,

$$R_{\text{ind}}(H; r) \leq 2^{k^{1+o(1)}}.$$

¹The authors of [21] only state the weaker bound $R_{\text{ind}}(H; r) \leq r^{O(rk^3)}$, because their focus in that paper was on fixed k and large r , but their proof can easily be modified to give the bound (4).

We resolve this problem in a very strong form, obtaining a bound that generalises Theorem 1.1 for every $r \geq 2$. In fact, our proof works directly in this more general setting.

Theorem 1.2. *There exists a constant $C > 0$ such that*

$$R_{\text{ind}}(H; r) \leq r^{Crk} \quad (5)$$

for every $r \geq 2$ and every graph H with k vertices.

We remark that, up to the value of the constant C , the bound (5) matches the classical upper bound of Erdős and Szekeres [18] on the r -colour Ramsey numbers $R(K_k; r) = R_{\text{ind}}(K_k; r)$, which was only recently improved by a small exponential factor in [2]. The best known lower bound is of the form $R(K_k; r) \geq 2^{\Omega(rk)}$, see [1, 10, 36, 32]. Our method will moreover imply the stronger statement that there exists a graph G with $N = r^{Crk}$ vertices such that every r -colouring of the edges of G contains an induced monochromatic copy of *every* graph H on k vertices. In fact, we will show that almost every graph G with N vertices has this property.

1.1. An overview of our approach. Unlike the earlier approaches in [23, 11, 20], where the authors developed ingenious deterministic algorithms to embed H in a pseudorandom host graph G , we will instead use a relatively simple vertex-by-vertex embedding strategy in a truly random graph G . Specifically, we consider the Erdős–Rényi random graph $\mathbb{G}(N, 1/2)$, where each edge of K_N is included independently at random with probability $1/2$; equivalently, we can choose a (labelled) graph G on N vertices uniformly at random.

Our strategy will crucially exploit the fact that the edges of $\mathbb{G}(N, 1/2)$ are chosen randomly, rather than relying on pseudorandom properties that are also satisfied by $G \sim \mathbb{G}(N, 1/2)$. We will show that, with (extremely) high probability, every r -colouring of the edges of $G \sim \mathbb{G}(N, 1/2)$ contains an induced monochromatic copy of an arbitrary k -vertex graph H .

One way to approach the problem is via an Erdős–Szekeres-type induction. In other words, we might generalise to the setting in which we want to find an induced copy of H_i in colour i , apply the induction hypothesis inside a subset $U \subset V(G)$ of size δN to find an induced copy of H_i^- (that is, H_i minus a vertex) in colour i for some $i \in [r]$, and then attempt to use the randomness between U and the rest of the vertices to extend this copy of H_i^- to a copy of H_i .

There are several major obstacles to utilizing such a strategy. First, and most obviously, there may be no edges of colour i between U and $V(G) \setminus U$, in which case we have no chance of extending H_i^- to a copy of H_i in colour i . We can easily deal with this issue, however, by instead using the induction hypothesis to find an induced copy of $H_i^- \subset G[U]$ in colour i for *every* $i \in [r]$, and then considering the colour that is used most often between U and $V(G) \setminus U$.

A second (and more challenging) issue is that the colouring of the edges inside the set U is allowed to depend on the (random) edges between U and $V(G) \setminus U$. We will deal with this problem by taking a union bound over all possible colourings of the edges of $G[U]$. However, this does not come for free: it requires us to prove an extremely strong bound on the probability of failure in each step of the induction.

In order to prove such a strong bound, we must strengthen our induction hypothesis. Indeed, if we only have one copy of H_i^- , the probability that no vertex of $V(G) \setminus U$ extends this fixed copy of H_i^- to an induced copy of H_i in G is essentially $(1 - 2^{-k})^N$, which is much too large to beat the roughly $r^{\delta^2 N^2}$ choices in our union bound. In order to reduce this failure probability, we will need

to instead find many copies of H_i^- in $G[U]$ that are moreover “well-distributed” in a certain precise sense, which we define in Section 2 (see Definition 2.2). We say that a hypergraph is (p, R) -Janson if its edges are well-distributed in this sense, since this definition resembles (and was inspired by) the condition in Janson’s inequality.

There is still, however, one further (and even more critical) obstruction to this approach: we must also handle all colourings of the edges between U and $V(G) \setminus U$, and in this case we cannot do so simply by taking a union bound, since there are too many possible colourings. In order to deal with this more serious obstacle, we will use the method of hypergraph containers.

1.2. Hypergraph containers. The method of hypergraph containers, which was introduced in 2015 by Balogh, Morris, and Samotij [4] and Saxton and Thomason [33], is a flexible technique for controlling the probability that a random set avoids some forbidden substructure (see the surveys [5, 31]). Roughly speaking, the basic container lemma implies that the sets that avoid these substructures are “clustered”, in the sense that they can be covered by a relatively small number of sets that contain only few copies of the forbidden substructures.

The container method has been used by several different sets of authors to prove Ramsey theoretic properties of random graphs. For example, it was used by Nenadov and Steger [25] and Mousset, Nenadov, and Samotij [24] to find monochromatic copies of fixed subgraphs in r -colourings of sparse random graphs, and by Conlon, Dellamonica, La Fleur, Rödl, and Schacht [13] and Balogh and Samotij [3] to bound the induced Ramsey numbers of graphs and hypergraphs.

In particular, the authors of [13] significantly improved the best-known upper bound for the induced Ramsey number of an arbitrary s -uniform hypergraph with k vertices when $s \geq 3$. Their result was then improved when $s = 2$ (that is, for graphs) in [3], where the authors introduced a new container lemma with a dramatically improved dependency on the size of the forbidden structure, and applied it to give a bound of the form

$$R_{\text{ind}}(H; r) \leq r^{O(rk^2)}. \quad (6)$$

However, further improvements to the dependency of the container lemma would not advance the bound beyond (6), and we must therefore take a different approach.

Our first significant departure from these earlier applications of the container method is in how we apply this method. In [3, 13, 24, 25], the authors embed the whole of H in a single step, encoding the edge sets of induced copies of H using a hypergraph whose vertex set consists of r copies of $E(K_n)$. In contrast, we will have not just one, but roughly $r^{|U|^2}$ different hypergraphs, one for each of the possible colouring of $G[U]$. Our hypergraphs will encode the vertex sets (not the edge sets) of monochromatic induced copies of H_i^- in $G[U]$ and our aim will be to find a (p, R) -Janson collection of (also monochromatic and induced) copies of H_i in an arbitrary r -colouring of the edges between U and $V(G) \setminus U$.

Although our final goal is to find a (p, R) -Janson collection of copies of H_i , applying the method of containers to find a single copy is already instructive. This will require introducing a correspondence between independent sets in our hypergraph and pairs of neighbourhoods $N_G(v)$ and $N_{G_i}(v)$ (where G_i is the subgraph of i -coloured edges) that do not extend any copy of H_i^- . In this simplified setting, applying a standard hypergraph container lemma is sufficient to conclude that the probability that a vertex does not extend any copies of H_i^- is exponentially small.

Changing from extending a single copy to finding a (p, R) -Janson collection of copies of H_i introduces significant new difficulties, which further set our argument apart from those in [3, 13, 24, 25]. The general strategy will be to have a second induction on the Janson parameter R : we will show that adding a vertex to U increases R with very high probability, which results in a “richer” family of copies of H_i . In this way, after adding δN vertices, R will be as large as we need.

However, even the first step of this incrementing argument requires containers for sets where the hypergraph is not (p, R) -Janson, rather than the classical container lemma from [4, 33] or the “efficient” container lemma proved in [3], see Theorem 4.1. For this purpose, our main tool will be a strengthening of the container method, proved recently by Campos and Samotij [7]. We will combine this result with an “efficient” container lemma in a novel and surprising way, resulting in a flexible method to prove new, more powerful container theorems.

1.3. Containers for sets that are not well-distributed. One of the main challenges of carrying out the above outline is that our induction hypothesis is a *global property* (being (p, R) -Janson) of the hypergraph that encodes monochromatic induced copies of H_i in colour i , whereas previous container lemmas are only able to handle *local properties* of the forbidden substructures.

Our most important technical contribution is a novel method that allows one to prove container theorems that can handle such global properties that depend on local behaviour. The first step of this method is to represent the global property as edges, of potentially linear size, in a hypergraph. We then decompose the independent sets of this hypergraph into independent sets of a small collection of *container hypergraphs*. Together, the container hypergraphs encode all of the local obstructions to this global property, which effectively separates the local obstructions from the global obstructions. This separation leads to a more favourable setting: to deal with local obstructions, we have a vast array of tools at our disposal, including traditional container theorems.

We illustrate this method by proving a general container theorem for sets that are not (p, R) -Janson, Theorem 4.1, which we moreover expect to have further applications. The proof of this result starts with an application of the aforementioned result of Campos and Samotij [7] (see Theorem 4.3), yielding a decomposition of the sets that are not (p, R) -Janson into a collection of independent sets in container hypergraphs. Roughly speaking, Theorem 4.3 states that each container hypergraph that does not have a “large local part” has the following property: a random independent set I with $q|V|$ vertices looks very similar to a binomial random subset V_q of the vertices of the hypergraph. More precisely, all sets L that are not edges of the container hypergraph satisfy

$$\mathbb{P}(L \subset I) \approx \mathbb{P}(L \subset V_q). \quad (7)$$

From (7), we will be able to deduce that the random independent set I is (p, R) -Janson with positive probability, contradicting the fact that independent sets are not (p, R) -Janson. It then follows that all container hypergraphs have a large local part, and we can apply another container theorem (Theorem 3.4, also proved in [7]) to the local part of each container hypergraph to finish the proof.

The next section provides key definitions for the rest of the paper and an important reduction. In particular, we define two events, one encoding the induction hypothesis and the other encoding the failure to contain a (p, R) -Janson collection of copies of H_i . We then state our key probabilistic lemma, which says that the probability of these two events happening simultaneously is very small, and prove that Theorem 1.2 follows from it by induction. In the end of Section 2, we discuss the structure of the remainder of the paper.

2. REDUCTION TO A KEY LEMMA

This section is devoted to the reduction of Theorem 1.2 to a key probabilistic lemma. To do that, we will first give two important definitions, Definitions 2.1 and 2.2. The first will allow us to reason about induced copies of a graph H as edges in a hypergraph, while the other is the key notion of how “well-distributed” the edges of a hypergraph are. This notion will underpin our container theorem, so immediately afterwards, we note a few simple properties of this latter notion that will be useful later.

Using these two definitions, we will further define an event to represent our inductive assumption, Definition 2.6, and another to stand for the failure of completing an induction step, Definition 2.7. We follow this with the statement of the key probabilistic lemma, Lemma 2.8, which bounds the probability of these events happening simultaneously, and a proof that Theorem 1.2 can be deduced from Lemma 2.8. This section concludes with a summary for the rest of the paper, where we provide the tools that we will use to prove the key lemma in Section 6.

2.1. Preliminaries. The first definition that we need is that of the hypergraph which encodes copies of some graph $F \subset G'$ that is also induced in G . We will use Definition 2.1 with G' being the subgraph defined by the edges in some colour $i \in [r]$, and G being the underlying (random) graph. In what follows and in the rest of the paper, we will identify a hypergraph with its edge set.

Definition 2.1. Given graphs F, G' and G such that $G' \subset G$, define $\mathfrak{J}_{F,G',G}$ to be the $v(F)$ -uniform hypergraph with vertex set $V(G)$, and

$$\mathfrak{J}_{F,G',G} = \{L \subset V(G) : F \cong G'[L] = G[L]\}.$$

That is, each hyperedge of $\mathfrak{J}_{F,G',G}$ corresponds to a copy of $F \subset G'$ that is induced in G .

As we will apply the container method to a hypergraph that is closely related to $\mathfrak{J}_{F,G',G}$, we emphasise that it has $V(G)$ for its vertex set. This is in contrast with using $E(G)$ for the vertices of the auxiliary hypergraph, the more common definition in applications of the container method.

We now formalise what it means for a hypergraph to be “well-distributed.” A crucial aspect of this definition is that we let ourselves assign weights to the edges of the hypergraph using measures. We say that a measure ν is *supported* on a hypergraph $\mathcal{G}$ if ν is non-zero only on the edges of $\mathcal{G}$. A hypergraph $\mathcal{G}$ is (p, R) -Janson when there exists a measure ν with certain properties that is supported on $\mathcal{G}$.

We call this property (p, R) -Janson because, under some extra assumptions, one can apply Janson’s inequality and conclude that a p -random subset of the hypergraph’s vertices is an independent set with probability at most $\exp(-R)$. Even though this is the original motivation for the definition, we will not require these extra assumptions or use Janson’s inequality.

Definition 2.2. Let $\mathcal{G}$ be a hypergraph and $p > 0$. For measures $\nu : \mathcal{G} \rightarrow \mathbb{R}_{\geq 0}$, we define

$$e(\nu) = \sum_{E \in \mathcal{G}} \nu(E) \quad \text{and} \quad \Lambda_p(\nu) = \sum_{\substack{L \subset V(\mathcal{G}) \\ |L| \geq 2}} d_\nu(L)^2 p^{-|L|}, \quad (8)$$

where

$$d_\nu(L) = \sum_{L \subset E \in \mathcal{G}} \nu(E)$$

is the degree of the set L in the measure ν . For every $R > 0$, we say that $\mathcal{G}$ is (p, R) -Janson if there exists a measure $\nu : \mathcal{G} \rightarrow \mathbb{R}_{\geq 0}$ such that

$$\Lambda_p(\nu) < \frac{e(\nu)^2}{R}. \quad (9)$$

If $R = 0$, then every hypergraph is (p, R) -Janson.

We mention three very simple, but useful observations about this property, for future reference. The first is so simple that we omit its proof.

Observation 2.3. *If $\mathcal{G}$ is a (p, R) -Janson hypergraph for $p > 0$ and $R \geq 0$, then for all $p' \geq p$ and $R' \leq R$, it is also (p', R') -Janson.*

The next observation is that we can normalise any measure ν satisfying (9) to choose the value it takes in $e(\nu)$.

Observation 2.4. *Let $p, R > 0$, and let $\mathcal{G}$ be a hypergraph that is (p, R) -Janson. For every $y > 0$, there exists $\nu : \mathcal{G} \rightarrow \mathbb{R}_{\geq 0}$ such that*

$$e(\nu) = y \quad \text{and} \quad \Lambda_p(\nu) < \frac{y^2}{R}.$$

Proof. As $\mathcal{G}$ is (p, R) -Janson, there exists $\tilde{\nu} : \mathcal{G} \rightarrow \mathbb{R}_{\geq 0}$ such that $\Lambda_p(\tilde{\nu}) < e(\tilde{\nu})^2/R$, which implies that $e(\tilde{\nu}) > 0$ by $\Lambda_p(\cdot) \geq 0$. One can now verify that taking $\nu = y\tilde{\nu}/e(\tilde{\nu})$ completes the proof. $\square$

Finally, we observe being (p, R) -Janson is monotone with respect to taking subhypergraphs.

Observation 2.5. *Let $R \geq 0$ and $p > 0$, and let $\mathcal{G}$ and $\mathcal{G}'$ be hypergraphs satisfying $\mathcal{G}' \subset \mathcal{G}$. If $\mathcal{G}'$ is (p, R) -Janson, then $\mathcal{G}$ is also (p, R) -Janson.*

Proof. The observation follows immediately from the fact that any ν supported on $\mathcal{G}'$ is also supported on $\mathcal{G}$, by assigning measure zero to every $E \in \mathcal{G} \setminus \mathcal{G}'$. $\square$

2.2. The key lemma. As outlined in Subsection 1.1, our proof will be by induction on k , and we will need a very strong bound on probability of each inductive step failing, in order to allow for a union bound over colourings of $G[U]$ for some set U of size δN . We will now define the event representing our induction hypothesis, Definition 2.6, and the event that represents the failure of completing an induction step, Definition 2.7. Once these are defined, we will state the key lemma, Lemma 2.8, and prove that our main theorem is a straightforward consequence of it.

To avoid repetition, we will let, in this section and also throughout the paper²

$$C = 300, \quad \delta = \delta(r) = \frac{1}{r^{50}} \quad \text{and} \quad p = p(r, k) = \frac{1}{2^{25} k^2 r^4}. \quad (10)$$

Whenever the specific value of these constants is important, we will remind the reader of them.

As previously mentioned, we have a very strong induction hypothesis. To state it, we introduce some notation. Given $c : E(G) \rightarrow [r]$ and a colour $i \in [r]$, we will denote by $G_i^{(c)}$ the subgraph consisting of the edges of G that are coloured i by c . We will omit the dependency of $G_i^{(c)}$ on c when the colouring is evident from context, writing simply G_i .

²Except in our container theorems, where we will let p vary in an interval that includes this value.

We will assume that for every sufficiently large subset $W \subset V(G)$, and every collection of graphs $F_1, \dots, F_r$ such that

$$v(F_1), \dots, v(F_r) \leq k \quad \text{and} \quad \sum_{i=1}^r v(F_i) = rk - 1,$$

the following holds: in any r -colouring of $G[W]$, there is a colour $i \in [r]$ for which the copies of $F_i \subset G_i[W]$ are well-distributed. To be precise, and using the definitions that we have just introduced, we will have that in this colour $i \in [r]$, the hypergraph $\mathfrak{J}_{F_i, G_i, G}[W]$ is $(p, p|W|)$ -Janson.

Definition 2.6. Given $r \in \mathbb{N}$ and $s_1, \dots, s_r \in \mathbb{N}$, let $\mathbf{s} = (s_i)_{i \in [r]}$ and let $\mathcal{E}(\mathbf{s})$ be the family of graphs G with the following property. For all graphs $F_1, \dots, F_r$ satisfying

$$\sum_{i=1}^r v(F_i) = \sum_{i=1}^r s_i - 1 \quad \text{and} \quad v(F_i) \leq s_i \quad \text{for each } i \in [r],$$

for every $W \subset V(G)$ with

$$|W| \geq \frac{\delta}{8r} v(G),$$

and every colouring $c : E(G[W]) \rightarrow [r]$, there is $i \in [r]$ such that $\mathfrak{J}_{F_i, G_i, G}[W]$ is $(p, p|W|)$ -Janson.

We now define the event that corresponds to the failure of the induction step, $\mathcal{B}(\mathbf{H})$.

Definition 2.7. Given a collection of graphs $H_1, \dots, H_r$, let $\mathbf{H} = (H_i)_{i \in [r]}$ and let $\mathcal{B}(\mathbf{H})$ be the family of graphs G with the following property. There exists a colouring $c : E(G) \rightarrow [r]$ such that, for every $i \in [r]$, the hypergraph $\mathfrak{J}_{H_i, G_i, G}$ is not $(p, p v(G))$ -Janson.

To simplify the notation, whenever all the $(H_i)_{i \in [r]}$ are equal to a single graph H , we denote this by $\mathcal{B}(H; r)$. Generalising the notation used in the introduction, we write

$$G \xrightarrow[r]{\text{ind}} H$$

to denote that, for any r -colouring of the edges of G , there exists an (induced) monochromatic copy of H . Observe that if $\mathbb{P}(G \in \mathcal{B}(H; r)) < 1$, then there exists a graph G such that $G \xrightarrow[r]{\text{ind}} H$.

We can now state the result from which we will deduce Theorem 1.2.

Lemma 2.8. *For all $k, r \in \mathbb{N}$ with $r \geq 2$ and for all graphs $H_1, \dots, H_r$ with at most k vertices, if $N \in \mathbb{N}$ satisfies*

$$N \geq r^{C(k+t)} \tag{11}$$

for $t = \sum_{i=1}^r v(H_i)$, then

$$\mathbb{P}(G \in \mathcal{B}(\mathbf{H}) \cap \mathcal{E}(\mathbf{s})) \leq 2^{-\delta^2 N^2},$$

where $G \sim \mathbb{G}(N, 1/2)$, $\mathbf{H} = (H_i)_{i \in [r]}$ and $\mathbf{s} = (v(H_i))_{i \in [r]}$.

To prove that $\mathbb{P}(G \in \mathcal{B}(H; r)) < 1$, we will find, for each $G \in \mathcal{B}(H; r)$, a large set $W \subset V(G)$ and graphs $H_1, \dots, H_r$ such that $G[W] \in \mathcal{B}(\mathbf{H}) \cap \mathcal{E}(\mathbf{s})$. Then, we will apply Lemma 2.8 to each $G[W]$ and take a union bound over choices of W and $H_1, \dots, H_r$ to complete the proof.

Proof that Lemma 2.8 implies Theorem 1.2. Fix

$$N = r^{5Crk}, \quad (12)$$

a choice which makes the constant in Theorem 1.2 equal to $5C$. Our goal is to show that there exists G on N vertices such that $G \xrightarrow[r]{\text{ind}} H$. As previously observed, it will suffice to prove that

$$\mathbb{P}(G \in \mathcal{B}(H; r)) < 1, \quad (13)$$

for $G \sim \mathbb{G}(N, 1/2)$. The next claim will let us bound this probability with a union bound.

Claim 2.9. *For every $G \in \mathcal{B}(H; r)$, there are $W \subset V(G)$ and $H_1, \dots, H_r \subset K_k$ such that*

$$|W| \geq \left(\frac{\delta}{8r}\right)^{rk-t} N \quad (14)$$

for $t = \sum_{i=1}^r v(H_i)$, and

$$G[W] \in \mathcal{E}(\mathbf{s}) \cap \mathcal{B}(\mathbf{H}) \quad (15)$$

where $\mathbf{H} = (H_i)_{i \in [r]}$ and $\mathbf{s} = (v(H_i))_{i \in [r]}$.

Proof. Choosing $H_1, \dots, H_r = H$ and $W = V(G)$ trivially satisfies (14) and $G[W] \in \mathcal{B}(\mathbf{H})$ by $G \in \mathcal{B}(H; r)$, which is a choice of $\mathbf{H}$ and W that satisfies the assumptions in the claim. The existence of one such choice means that we can take $\mathbf{H} = (H_i)_{i \in [r]}$ and $W \subset V(G)$ to minimise $t = \sum_{i=1}^r v(H_i)$ for $H_1, \dots, H_r \subset K_k$ among the choices that satisfy $G[W] \in \mathcal{B}(\mathbf{H})$ and (14). We claim that this $\mathbf{H}$ and W also satisfies $G[W] \in \mathcal{E}(\mathbf{s})$, and therefore (15).

Indeed, if $G[W] \notin \mathcal{E}(\mathbf{s})$, then, by definition, there are graphs $H'_1, \dots, H'_r$ with

$$\sum_{i=1}^r v(H'_i) = t - 1, \quad (16)$$

a set $W' \subset W$ with

$$|W'| \geq \frac{\delta}{8r} |W|,$$

and a colouring $c : G[W'] \rightarrow [r]$ such that $\mathfrak{J}_{H'_i, G_i, G}[W']$ is not $(p, p|W'|)$ -Janson for all $i \in [r]$. But then W' and $\mathbf{H}' = (H'_i)_{i \in [r]}$ also satisfy $G[W'] \in \mathcal{B}(\mathbf{H}')$ and

$$|W'| \geq \frac{\delta}{8r} |W| \geq \left(\frac{\delta}{8r}\right)^{rk-(t-1)} N,$$

since W satisfies (14). Therefore, W' would also satisfy (14), and the existence of W' and $\mathbf{H}'$ would, by (16), contradict the minimality of t in the original choice of W and $\mathbf{H}$, so $G[W] \in \mathcal{E}(\mathbf{s})$. ■

Applying Claim 2.9, we can take a union bound over the choices of $\mathbf{H}$ and $W \subset V(G)$ in that statement, obtaining as a result

$$\mathbb{P}(G \in \mathcal{B}(H; r)) \leq \sum_{\mathbf{H}} \sum_{\substack{W \subset V(G) \\ |W| \geq \delta^{2rk} N}} \mathbb{P}(G[W] \in \mathcal{B}(\mathbf{H}) \cap \mathcal{E}(\mathbf{s}_{\mathbf{H}})), \quad (17)$$

where $\mathbf{s}_{\mathbf{H}} = (v(H_i))_{i \in [r]}$, and we have bounded $(\delta/8r)^{rk-t} N \geq \delta^{2rk} N$ for all $t \geq 0$ by $r \geq 2$ and our choice of $\delta = r^{-50} \leq 1/(8r)$.

Now, with the goal of bounding each term in (17) by applying Lemma 2.8, fix $W \subset V(G)$ with $|W| \geq \delta^{2rk} N$ and $\mathbf{H} = (H_i)_{i \in [r]}$, a collection of graphs, each with at most k vertices. Observe that letting $G' = G[W]$, we have $G' \sim \mathbb{G}(N', 1/2)$ where $N' = |W|$, and also that $\mathbf{H}$ trivially satisfies

$$t = \sum_{i=1}^r v(H_i) \leq rk. \quad (18)$$

The only assumption in Lemma 2.8 that remains to be checked is that N' satisfies (11), which it does because

$$N' \geq \delta^{2rk} N \geq r^{-100rk} r^{5Crk} \geq r^{4Crk} \geq r^{C(k+t)}$$

by our choice of $N = r^{5Crk}$ in (12), the fixed values of $\delta = r^{-50}$ and $C = 300$ in (10), and (18).

Applying Lemma 2.8 to $G' = G[W]$ and $\mathbf{H}$, we conclude that

$$\mathbb{P}(G[W] \in \mathcal{B}(\mathbf{H}) \cap \mathcal{E}(\mathbf{s}_{\mathbf{H}})) \leq 2^{-\delta^2 |W|^2} \leq 2^{-\delta^{5rk} N^2}, \quad (19)$$

for all $|W| \geq \delta^{2rk} N$. As W and $\mathbf{H}$ were arbitrary, (19) holds for all terms in (17).

Replacing (19) in (17), bounding the choices for $\mathbf{H}$ by $k^r 2^{r \binom{k}{2}}$ and for $W \subset V(G)$ by 2^N , we have

$$\mathbb{P}(G \in \mathcal{B}(H; r)) \leq k^r 2^{r \binom{k}{2}} 2^N 2^{-\delta^{5rk} N^2} \leq 2^{-\delta^{5rk} N^2/2} < 1 \quad (20)$$

where the final inequalities hold since

$$\frac{\delta^{5rk} N^2}{2} > N + rk + r \binom{k}{2} > 0$$

by our choice of $N = r^{5Crk}$ and $\delta = r^{-50}$. We have established (13) and completed the proof. $\square$

Remark 2.10. Observe that the bound in (20) is sufficiently small for us to take another union bound, this one over all graphs H with k vertices, for which there are at most $2^{\binom{k}{2}}$ choices. The conclusion is that, with high probability, if $G \sim \mathbb{G}(N, 1/2)$, then $G \xrightarrow[r]{\text{ind}} H$ for all such H .

2.3. Structure of the paper. The bulk of the paper is concerned with proving Lemma 2.8. However, before we proceed to the proof of Lemma 2.8, we will need some tools.

First, we will have two expository sections that contain many of the ideas that we use later, but in simpler settings. In Section 3, this simpler setting corresponds to showing that the probability that a vertex extends no copy of H_i^- in a fixed colour $i \in [r]$ is exponentially small. This section shows how to map neighbourhoods that fail to extend a copy of H_i^- to independent sets in a certain hypergraph $\mathcal{H}$, which will illustrate how the method of hypergraph containers can be applied to this problem. It is also where we explain how our induction hypothesis, the event $\mathcal{E}(\mathbf{s})$, yields a supersaturation result for our application(s) of containers.

In Section 4, we present the proof of a container theorem for “sets that are not (p, R) -Janson” (see Theorem 4.1, and also Theorem 4.6, which implies it). While this result is not sufficiently strong to prove Lemma 2.8, its proof already presents the general argument that we will later adapt to prove Theorem 5.4, the container theorem that we do use in the proof of Lemma 2.8. We see that argument as a general method to reduce a problem about sets avoiding a global property to a simpler problem, where the sets only need to avoid a local property. In this particular instance, we deal with this local property by using a standard hypergraph container lemma (Theorem 3.4).

Section 5 contains the proof that if the hypergraph $\mathfrak{J}_{H_i, G_i, G}[U]$ is (p, R) -Janson, then adding a new vertex to U produces a $(p, R+1)$ -Janson hypergraph with probability $1 - \exp(-\gamma|U|)$ for some

constant $\gamma = \gamma(r) > 0$ that only depends on r . The proof of this statement mirrors the proof in Section 3, but replacing the hypergraph container lemma there by the (yet unproven) Theorem 5.4, the container theorem tailored to our application. We postpone the proof of Theorem 5.4 to the last section since it is the most technical part of the paper.

In Section 6, we prove Lemma 2.8 in two stages. The first and easier stage uses a double counting argument to show that if $\mathfrak{J}_{H_i, G_i, G}[S]$ is roughly $(p, \delta pN/r)$ -Janson for all subsets S with $|S| \geq \delta^{2/3}N$, then $\mathfrak{J}_{H_i, G_i, G}$ is (p, pN) -Janson. The second stage proves Lemma 2.8 by induction on the Janson parameter of the family $\mathfrak{J}_{H_i, G_i, G}[U]$. To prove the inductive step, we use Lemma 5.1 to show that the probability one cannot increase this Janson parameter by adding a vertex to U is very small. This is also the part where we take a union bound over colourings of $G[U]$, which ends up being somewhat technical because we must be careful with the randomness outside $G[U]$.

The last section, Section 7, contains the proof of Theorem 5.4. As we mentioned before, this proof follows the blueprint of Section 4, but there are extra details. One is that we want to combine the family of copies of H_i containing v with the copies that are already in U , which does not contain v . The second is that the hypergraph $\mathcal{H}$ that we use to deal with the extensions, defined in Section 3, is not the same as $\mathfrak{J}_{H_i, G_i, G}$, the hypergraph appearing in the event whose probability we must bound. Related to this is the fact that the hypergraph defined by $\mathcal{H}$ corresponds to copies of H_i^- , and we are interested in copies of H_i . These complications make the proof more technical, but it turns out the approach we developed in Section 4 is sufficiently flexible to deal with them.

3. WARM-UP TO THE EXTENSION LEMMA

In this section, we give an essentially complete proof of a weak “extension lemma”, Lemma 3.1. Very roughly, this result bounds the probability that a fixed vertex of G fails to complete any copy of a graph F . We will describe the setting for this result in Subsection 3.1, connecting it with the application of its stronger variant in the proof of Lemma 2.8.

Of central importance in the proof of Lemma 3.1 is a hypergraph container theorem, here stated as Theorem 3.4, which motivates many of the definitions that follow. The most significant of these is that of the auxiliary hypergraph $\mathcal{H}$, defined in Subsection 3.2; the other definitions will, above all, assist in reasoning about it.

As we will later see, each edge of this auxiliary hypergraph represents one way to extend, using a fixed vertex v , an (induced) copy of F^- to a copy of F . Observation 3.3 then fulfils a crucial requirement to apply the container method, establishing that independent sets in $\mathcal{H}$ correspond to pairs of graphs (G', G) in which the neighbourhood of v fails to extend every F^- in a fixed set U .

3.1. Extension lemmas. The setting of Lemma 3.1 is that $\tilde{G}'$ and $\tilde{G}$ are m -vertex graphs such that $\tilde{G}' \subset \tilde{G}$ and $\mathfrak{J}_{F^-, \tilde{G}', \tilde{G}}[W]$ is $(p, 2pm)$ -Janson for every $W \subset V(\tilde{G})$ with $|W| \geq m/16$. Let $U = V(\tilde{G})$ and let G be the random graph $\mathbb{G}(m+1, 1/2)$ conditioned on $G[U] = \tilde{G}$. The lemma bounds the probability that G contains a subgraph G' with the following properties:

- (a) $G'[U] = \tilde{G}'$,
- (b) the vertex $v \in V(G) \setminus U$ does not have small degree in G' , and
- (c) v does not extend any copy of $F^- \subset \tilde{G}'$ to an induced F in G using the edges in G' .

In our application, U will be a set of vertices for which the colouring $c : E(G) \rightarrow [r]$ is fixed, v will be a vertex not in U such that $d_{G_i}(v, U) \geq |U|/4r$ for some colour $i \in [r]$, and we will set

$\tilde{G} = G[U]$ and $\tilde{G}' = G_i[U]$. In this setting, we could use this lemma (when $r = 2$) to conclude that it is extremely likely that v extends some i -coloured copy of $F^- := H_i^-$ to a copy of $F = H_i$.

Lemma 3.1. *Let $m, s, k \in \mathbb{N}$ with $s < k \leq m$, $p = 2^{-20}k^{-2}$ and F be a graph on $s + 1$ vertices. Further let $\tilde{G}'$ and $\tilde{G}$ be graphs on m vertices satisfying $\tilde{G}' \subset \tilde{G}$.*

If $\mathfrak{I}_{F^-, \tilde{G}', \tilde{G}}[W]$ is $(p, 2pm)$ -Janson for every $W \subset U = V(\tilde{G})$ with $|W| \geq m/16$, then

$$\mathbb{P} \left(\exists G' \subset G : \begin{array}{l} G'[U] = \tilde{G}', \quad d_{G'}(v) \geq m/8 \\ \text{and } \mathfrak{I}_{F, G', G} = \emptyset \end{array} \mid G[U] = \tilde{G} \right) \leq 2^{-m/32}, \quad (21)$$

where $G \sim \mathbb{G}(m + 1, 1/2)$ and $V(G) = U \cup \{v\}$.

We refer to Lemma 3.1 as a weak extension lemma because it involves the event $\{\mathfrak{I}_{F, G', G} = \emptyset\}$, whereas our main extension result, Lemma 5.1, involves instead $\{\mathfrak{I}_{F, G', G} \text{ is not } (p, R)\text{-Janson}\}$. The proof of Lemma 3.1 presents most of the ideas that we will require in Section 5, while avoiding some technicalities from dealing with this more complicated property of $\mathfrak{I}_{F, G', G}$.

3.2. The auxiliary hypergraph. We want to bound the probability in (21) using the method of hypergraph containers. To accomplish that, the first step is to represent the graphs $G' \subset G$ such that $G'[U] = \tilde{G}'$ and $\mathfrak{I}_{F, G', G} = \emptyset$ as independent sets in an auxiliary hypergraph $\mathcal{H}$. This auxiliary hypergraph will be a function of $\tilde{G}'$ and $\tilde{G}$ only, which crucially means that $\mathcal{H}$ does not depend on the edges of G (or $G' \subset G$) between v and U .

Define $w \in V(F) \setminus V(F^-)$ and note that the only randomness in the event inside the probability in (21) comes from the edges of G between U and v , since we condition on the event $\{G[U] = \tilde{G}\}$. Our goal is to have each edge of $\mathcal{H}$ correspond to one way of extending a copy of F^- in $\tilde{G}'$, which is also induced in $\tilde{G}$, to a (still induced) copy of F , by adding the vertex v in the role of w . To be precise, for each edge $L \in \mathfrak{I}_{F^-, \tilde{G}', \tilde{G}}$, fix a bijection $\phi_L : L \rightarrow V(F^-)$ such that $F^- \cong_{\phi_L} \tilde{G}'[L] = \tilde{G}[L]$. Now define $\mathcal{H}$ to be the hypergraph with vertex set $U \times \{0, 1\}$ and edge set

$$\mathcal{H} = \left\{ E_L : L \in \mathfrak{I}_{F^-, \tilde{G}', \tilde{G}} \right\}, \quad (22)$$

where

$$E_L = \left\{ \left(u, \mathbb{1}[\phi_L(u) \in N_F(w)] \right) : u \in L \right\},$$

for every $L \in \mathfrak{I}_{F^-, \tilde{G}', \tilde{G}}$. In words, the vertex set of $\mathcal{H}$ is two copies of U , corresponding to the neighbours and non-neighbours of w in F , and its edge set has the following property: if $u \in N_{G'}(v)$ for every $(u, 1) \in E_L$ and $u \notin N_G(v)$ for every $(u, 0) \in E_L$, then $L \cup \{v\} \in \mathfrak{I}_{F, G', G}$.

To formalise that independent sets in $\mathcal{H}$ correspond to pairs of graphs (G', G) of our interest, denote the non-neighbours of $v \notin U$ by

$$N_G(v)^c = U \setminus N_G(v) \quad (23)$$

and, for each $A \subset V(\mathcal{H})$ and $i \in \{0, 1\}$, let

$$A^{(i)} = \{u \in U : (u, i) \in A\}.$$

We repeat the following property, previously stated without these definitions, for future reference.

Observation 3.2. Let G' and G be graphs satisfying $G' \subset G$, $G'[U] = \tilde{G}'$ and $G[U] = \tilde{G}$. If $L \in \mathfrak{I}_{F-, \tilde{G}', \tilde{G}}$ and $E = E_L \in \mathcal{H}$, then we can extend L to a copy of $F \subset G'$ induced in G , i.e. $L \cup \{v\} \in \mathfrak{I}_{F, G', G}$, whenever

$$E^{(0)} \subset N_G(v)^c \quad \text{and} \quad E^{(1)} \subset N_{G'}(v).$$

The most useful consequence of this fact, which allows applying the method of hypergraph containers, is more conveniently stated with some additional notation. Let

$$\Gamma_G = \left\{ G' \subset G : \begin{array}{l} G'[U] = \tilde{G}', \ d_{G'}(v) \geq m/8 \\ \text{and } \mathfrak{I}_{F, G', G} = \emptyset \end{array} \right\} \quad (24)$$

be the collection of G' whose existence is the event in the probability of (21). To index vertex subsets of $\mathcal{H}$ by graphs G' and G , further let

$$\iota(G', G) = \{(u, 0) : u \in N_G(v)^c\} \cup \{(u, 1) : u \in N_{G'}(v)\}, \quad (25)$$

and note that if $I = \iota(G', G)$, then

$$I^{(0)} = N_G(v)^c \quad \text{and} \quad I^{(1)} = N_{G'}(v). \quad (26)$$

To justify one of the premises in Observation 3.3, recall that we have conditioned the distribution of $G \sim \mathbb{G}(m+1, 1/2)$ on $\{G[U] = \tilde{G}\}$ in (21).

Observation 3.3. If $G[U] = \tilde{G}$ and $G' \in \Gamma_G$, then $\iota(G', G)$ is an independent set in $\mathcal{H}$.

Proof. Let $E = E_L \in \mathcal{H}$, and suppose by contradiction that $E \subset I = \iota(G', G)$. It then follows from (26) that $E^{(1)} \subset N_{G'}(v)$ and $E^{(0)} \subset N_G(v)^c$, so Observation 3.2 implies that $L \cup \{v\} \in \mathfrak{I}_{F, G', G}$. Therefore, the hypergraph $\mathfrak{I}_{F, G', G}$ is not empty, and we could not have $G' \in \Gamma_G$ by definition, (24). This contradiction means that $E \not\subset I$, which, as E was an arbitrary edge of $\mathcal{H}$, means that $I = \iota(G', G)$ is an independent in $\mathcal{H}$. $\square$

3.3. Containers. Having connected independent sets in $\mathcal{H}$ to pairs of graphs (G', G) such that $G' \in \Gamma_G$, we can proceed to the next step of the proof, which is applying a container theorem to $\mathcal{H}$. In this section, we use a simpler version of the container theorem than the one required in Section 5, since this avoids technical complications that arise in the proof of Lemma 5.1.

Even though a previous result of Balogh and Samotij [3, Theorem 2.1] would most likely suffice to prove Lemma 3.1, we prefer to adapt (in Appendix B) the mildly stronger theorem of Campos and Samotij [7]. This result provides a small family of containers for the independent sets of $\mathcal{H}$, with each container inducing a subhypergraph that is not (p, R) -Janson.

Theorem 3.4 (Campos and Samotij [7, modified Theorem A]). *Let $\mathcal{G}$ be an s -uniform hypergraph with n vertices. For all $0 < \zeta \leq 1$ and $0 < p \leq \zeta/(8s^2)$, there is a family $\mathcal{S} \subset 2^{V(\mathcal{G})}$ and functions*

$$\phi : \mathcal{I}(\mathcal{G}) \rightarrow \mathcal{S} \quad \text{and} \quad \psi : \mathcal{S} \rightarrow 2^{V(\mathcal{G})} \quad (27)$$

such that:

- (i) For each $I \in \mathcal{I}(\mathcal{G})$, we have $\phi(I) \subset I \subset \psi(\phi(I))$.
- (ii) Each $S \in \mathcal{S}$ has at most $8s^2pn/\zeta$ elements.
- (iii) For every $S \in \mathcal{S}$, letting $X = \psi(S)$, $\mathcal{G}[X]$ is not $(p, \zeta p|X|)$ -Janson.

We will apply Theorem 3.4 with $\mathcal{G} = \mathcal{H}$ and define $\mathcal{X} = \{\psi(S) : S \in \mathcal{S}\}$ to be our container family. Combining item (i) in the statement of that theorem with Observation 3.3 guarantees that whenever G' and G satisfy $G[U] = \tilde{G}$ and $G' \in \Gamma_G$, there is some container $X \in \mathcal{X}$ for which $\iota(G', G) \subset X$. Our goal is now to show that $|X^{(0)}| \leq (1 - \gamma)|U|$ for some constant $\gamma > 0$ using item (iii) in Theorem 3.4 and a supersaturation result. This will be sufficient to obtain the probability bound in Lemma 3.1, because $N_G(v)^c \subset X_0$ and this non-neighbourhood is a binomial random set, due to $G \sim \mathbb{G}(m + 1, 1/2)$.

To prove the required supersaturation result, we will use our assumption that $\mathfrak{J}_{F^-, \tilde{G}', \tilde{G}}$ is $(p, 2pm)$ -Janson for every $W \subset U$ with $|W| \geq m/16$. However, to connect this assumption with the fact that $\mathcal{H}[X]$ is not $(p, p|X|)$ -Janson for a container $X \in \mathcal{X}$, we must relate the Janson properties of the hypergraphs $\mathcal{H}$ and $\mathfrak{J}_{F^-, \tilde{G}', \tilde{G}}$.

Towards that goal, let $\pi : U \times \{0, 1\} \rightarrow U$ be the projection of each pair (u, i) onto its first coordinate u . Observe that if $E = E_L \in \mathcal{H}$ for some $L \in \mathfrak{J}_{F^-, \tilde{G}', \tilde{G}}$, then $\pi(E) = L$. Moreover, for every $L \in \mathfrak{J}_{F^-, \tilde{G}', \tilde{G}}$, there is $E \in \mathcal{H}$ such that $\pi(E) = L$. We conclude that $\pi(\mathcal{H}) = \mathfrak{J}_{F^-, \tilde{G}', \tilde{G}}$, where we extend the application of π from a single vertex to hypergraphs $\mathcal{G}$ by

$$V(\pi(\mathcal{G})) = \pi(V(\mathcal{G})) \quad \text{and} \quad E(\pi(\mathcal{G})) = \{\pi(E) : E \in \mathcal{G}\}.$$

Observation 3.5. *If $\pi : U \times \{0, 1\} \rightarrow U$ is the projection onto the first coordinate, then*

$$\pi(\mathcal{H}) = \mathfrak{J}_{F^-, \tilde{G}', \tilde{G}}.$$

The next lemma will allow us to prove that if $\mathfrak{J}_{F^-, \tilde{G}', \tilde{G}}$ is (p, R) -Janson, then so is $\mathcal{H}$. We state it in greater generality than we need, but it is a trivial exercise to observe that π satisfies this more general condition. Lemma 3.6 is in fact a corollary of Lemma 7.2, and we will therefore defer its proof to when that other result is proven (see Section 7 and Appendix C).

Lemma 3.6. *Let $R \geq 0$ and $p > 0$. Further let $\mathcal{G}$ be a hypergraph and $\pi : V(\mathcal{G}) \rightarrow U$ be a function satisfying*

$$|\pi(E)| = |E| \quad \text{for every } E \in \mathcal{G}. \quad (28)$$

If $\mathcal{G}$ is not (p, R) -Janson, then $\pi(\mathcal{G})$ is also not (p, R) -Janson.

We now record the trivial observation that π satisfies (28) when $\mathcal{G} = \mathcal{H}$ for future reference.

Observation 3.7. $|\pi(E)| = |E|$ for every $E \in \mathcal{H}$.

By item (iii) in Theorem 3.4 and Lemma 3.6, it follows that $\pi(\mathcal{H}[X])$ is not $(p, p|X|)$ -Janson for each container X . However, our assumption in Lemma 3.1 is that $\pi(\mathcal{H})[W]$ is $(p, 2pm)$ -Janson for all large sets $W \subset U$, and this does not immediately imply anything about the size of X . For example, we might try to apply this assumption with $W = \pi(X)$, and hope that

$$\pi(\mathcal{H})[W] \subset \pi(\mathcal{H}[X]), \quad (29)$$

but unfortunately (29) does not always hold: if $X = U \times \{0\}$, then we have $\pi(\mathcal{H})[\pi(X)] = \pi(\mathcal{H})$ and $\mathcal{H}[X] = \emptyset$. In order to deal with this issue, we will instead apply our assumption to the set $W = X^{(0)} \cap X^{(1)}$. By the following lemma, this choice *does* satisfy (29).

Lemma 3.8. *For all $Y \subset U \times \{0, 1\}$, we have*

$$\pi(\mathcal{H})[W] \subset \pi(\mathcal{H}[Y]), \quad (30)$$

where $W = Y^{(0)} \cap Y^{(1)}$.

Proof. Take $L \in \pi(\mathcal{H})[W]$ with the goal of proving that $L \in \pi(\mathcal{H}[Y])$. By definition of W ,

$$L \subset Y^{(0)} \cap Y^{(1)}. \quad (31)$$

Moreover, since $L \in \pi(\mathcal{H})$, there is $E \in \mathcal{H}$ such that $\pi(E) = L$. Therefore,

$$E \subset \pi^{-1}(L) = \{(u, a) : u \in L, a \in \{0, 1\}\},$$

but we also have, by (31), that

$$\pi^{-1}(L) \subset (Y^{(0)} \times \{0\}) \cup (Y^{(1)} \times \{1\}) = Y,$$

and so $E \subset Y$. We conclude that $E \in \mathcal{H}[Y]$ and also $L = \pi(E) \in \pi(\mathcal{H}[Y])$, as required. $\square$

Now, taking $W = X^{(0)} \cap X^{(1)}$ for some container $X \in \mathcal{X}$ and applying Lemma 3.8, we have that $\pi(\mathcal{H})[W] \subset \pi(\mathcal{H}[X])$. Using item (iii) of Theorem 3.4, Observation 3.5, Lemma 3.6 and Observation 2.5, we will conclude that $\mathfrak{J}_{F^-, \tilde{G}', \tilde{G}}[W]$ is not $(p, 2pm)$ -Janson (see Claim 3.9). Our assumption in Lemma 3.1 then implies that $|W| < m/16$, so every $X \in \mathcal{X}$ satisfies

$$|X^{(0)} \cap X^{(1)}| < \frac{m}{16},$$

by definition of W , and, since $X^{(0)} \cup X^{(1)} \subset U$, it follows that

$$|X^{(0)}| + |X^{(1)}| = |X^{(0)} \cup X^{(1)}| + |X^{(0)} \cap X^{(1)}| < \left(1 + \frac{1}{16}\right)m. \quad (32)$$

To establish that $X^{(0)}$ is not large when $I = \iota(G', G) \subset X$, recall that it follows from $G' \in \Gamma_G$ that

$$|X^{(1)}| \geq d_{G'}(v) \geq \frac{m}{8} \quad (33)$$

where the first inequality holds because $N_{G'}(v) \subset X^{(1)}$ by (26). Combining (32) and (33) yields

$$|X^{(0)}| \leq \left(1 - \frac{1}{16}\right)m,$$

which, together with

$$N_G(v)^c \subset X^{(0)} \quad (34)$$

will allow us to use the randomness in the distribution of $G \sim \mathbb{G}(m+1, 1/2)$ to bound the probability of (34) by an exponentially small term. The resulting upper bound is sufficiently strong to overcome the union bound over all $X \in \mathcal{X}$, so we can now formalise the proof of Lemma 3.1.

Proof of Lemma 3.1. Applying Theorem 3.4 to $\mathcal{H}$, defined in (22), with $\zeta = 1$ we obtain a family $\mathcal{S}$ and functions ϕ, ψ satisfying items (i)–(iii) in Theorem 3.4, so let

$$\mathcal{X} = \{\psi(S) : S \in \mathcal{S}\}.$$

With the goal of taking a union bound over $X \in \mathcal{X}$, fix G on $m+1$ vertices satisfying $G[U] = \tilde{G}$, and observe that $\iota(G', G) \in \mathcal{I}(\mathcal{H})$ for all $G' \in \Gamma_G$ by Observation 3.3. Now, it follows from item (i) of Theorem 3.4 that there exists $X \in \mathcal{X}$ such that $\iota(G', G) \subset X$, and therefore

$$N_G(v)^c = I^{(0)} \subset X^{(0)} \quad \text{and} \quad N_{G'}(v) = I^{(1)} \subset X^{(1)}, \quad (35)$$

where $I = \iota(G', G)$. Using (35), we further refine $\mathcal{X}$ to

$$\mathcal{X}' = \{X \in \mathcal{X} : |X^{(1)}| \geq m/8\}$$

and preserve the property that for every $G' \in \Gamma_G$, there is $X \in \mathcal{X}'$ such that $\iota(G', G) \subset X$, because

$$|X^{(1)}| \geq d_{G'}(v) \geq \frac{m}{8} \quad (36)$$

follows from $G' \in \Gamma_G$, defined in (24). That is, if $X \in \mathcal{X}$ contains some $\iota(G', G)$, then it is also in $\mathcal{X}'$ by (36), and crucially, for every $G' \in \Gamma_G$, there exists $X \in \mathcal{X}'$ such that

$$N_G(v)^c \subset X^{(0)}. \quad (37)$$

Taking a union bound over $\mathcal{X}'$ then yields, for the probability in (21),

$$\mathbb{P}(\exists G' \subset G : G' \in \Gamma_G \mid G[U] = \tilde{G}) \leq \sum_{X \in \mathcal{X}'} \mathbb{P}(N_G(v)^c \subset X^{(0)}), \quad (38)$$

where we used (37) to bound the probability of the event $\{\iota(G', G) \subset X\}$ by that of the event $\{N_G(v)^c \subset X^{(0)}\}$. We shift our focus to obtaining an upper bound for the probability of this event that holds for every $X \in \mathcal{X}'$, so we fix $X \in \mathcal{X}'$ and set $W = X^{(0)} \cap X^{(1)}$. The following claim, together with the assumption that $\mathfrak{J}_{F^-, \tilde{G}', \tilde{G}}[W']$ is $(p, 2pm)$ -Janson for every $W' \subset U$ with $|W'| \geq m/16$, will allow us to bound the size of $X^{(0)}$ from above.

Claim 3.9. $\mathfrak{J}_{F^-, \tilde{G}', \tilde{G}}[W]$ is not $(p, 2pm)$ -Janson.

Proof. By Observation 3.5 and Lemma 3.8, we have

$$\mathfrak{J}_{F^-, \tilde{G}', \tilde{G}}[W] = \pi(\mathcal{H})[W] \subset \pi(\mathcal{H}[X]).$$

Therefore, if we establish that the hypergraph $\pi(\mathcal{H}[X])$ is not $(p, 2pm)$ -Janson, then we are done, because being $(p, 2pm)$ -Janson is increasing with respect to inclusion, by Observation 2.5.

To show this, observe first that $\mathcal{H}[X]$ is not $(p, p|X|)$ -Janson by item (iii) in Theorem 3.4 and our choice of $\zeta = 1$. Since $|X| \leq 2m = v(\mathcal{H})$, and recalling Observation 2.3, it follows that the hypergraph $\mathcal{H}[X]$ is also not $(p, 2pm)$ -Janson. Thus, by Lemma 3.6, we deduce that $\pi(\mathcal{H}[X])$ is not $(p, 2pm)$ -Janson, and, by our previous reasoning, that neither is $\mathfrak{J}_{F^-, \tilde{G}', \tilde{G}}[W]$, as claimed. ■

Since we know, by assumption, that $\mathfrak{J}_{F^-, \tilde{G}', \tilde{G}}[W']$ is $(p, 2pm)$ -Janson for every $W' \subset U$ with $|W'| \geq m/16$, Claim 3.9 implies that $|W| < m/16$ and therefore

$$|X^{(0)}| + |X^{(1)}| = |X^{(0)} \cup X^{(1)}| + |X^{(0)} \cap X^{(1)}| < m + \frac{m}{16}.$$

As $X \in \mathcal{X}'$, we have $|X^{(1)}| \geq m/8$, and therefore

$$|X^{(0)}| < \left(1 + \frac{1}{16}\right)m - |X^{(1)}| \leq m + \frac{m}{16} - \frac{m}{8} = \left(1 - \frac{1}{16}\right)m,$$

that is,

$$|U \setminus X^{(0)}| \geq \frac{m}{16}. \quad (39)$$

We conclude from (39) that

$$\mathbb{P}(N_G(v)^c \cap (U \setminus X^{(0)}) = \emptyset) = 2^{-|U \setminus X^{(0)}|} \leq 2^{-m/16}, \quad (40)$$

where we used that $G \sim \mathbb{G}(m+1, 1/2)$ and $v \notin U$. Replacing (40) in (38), we obtain

$$\sum_{X \in \mathcal{X}'} \mathbb{P}(N_G(v)^c \subset X^{(0)}) \leq |\mathcal{X}'| 2^{-m/16}. \quad (41)$$

To bound the size of $\mathcal{X}'$ by $|\mathcal{X}|$, we enumerate the latter using that each container X is a function of $S \in \mathcal{S}$. As each $S \in \mathcal{S}$ satisfies

$$|S| \leq 16ps^2m \leq 2^{-16}m$$

by item (ii) in Theorem 3.4 and our choice of $\zeta = 1$, where the last inequality holds by our choice of $p = 2^{-20}k^{-2}$ and the assumption that $v(F) = s \leq k$, we have

$$|\mathcal{X}'| \leq |\mathcal{X}| \leq \sum_{t=0}^{2^{-16}m} \binom{2m}{t} \leq 2^{m/32}, \quad (42)$$

where, recall, $\mathcal{H}$ has $2m$ vertices. Combining (42) and (41), we complete the proof:

$$\mathbb{P}(\exists G' \subset G : G' \in \Gamma_G \mid G[U] = \tilde{G}) \leq 2^{m/32} 2^{-m/16} \leq 2^{-m/32}$$

as required. $\square$

4. A GENERAL CONTAINER THEOREM FOR NON-JANSON SETS

In this section, we prove a preliminary version of our main technical result.

Theorem 4.1. *Let $s, n \in \mathbb{N}$ with $s \leq n$, and let $q, p, R, \eta > 0$ satisfy*

$$q \leq \frac{1}{16}, \quad p \leq \frac{q}{2^{10}s^2}, \quad R \geq 2^{-6}pn \quad \text{and} \quad \eta = 2^{-2s-2}.$$

For every s -uniform hypergraph $\mathcal{H}$ with n vertices, there exists a family $\mathcal{X} \subset 2^{V(\mathcal{H})}$ with

$$|\mathcal{X}| \leq \left(\frac{2}{q}\right)^{8qn} \quad (43)$$

such that the following hold.

- (i) *If $L \subset V(\mathcal{H})$ and $\mathcal{H}[L]$ is not $(p/q, \eta R)$ -Janson, then $L \subset X$ for some $X \in \mathcal{X}$.*
- (ii) *For each $X \in \mathcal{X}$, the hypergraph $\mathcal{H}[X]$ is not (p, R) -Janson.*

We now discuss the proof of Theorem 4.1. The first step is defining the auxiliary hypergraph

$$\mathcal{J} = \{L \subset V : \mathcal{H}[L] \text{ is } (p/q, \eta R)\text{-Janson}\},$$

where $V = V(\mathcal{H})$. It follows immediately from the definition that sets $L \subset V$ for which $\mathcal{H}[L]$ is not $(p/q, \eta R)$ -Janson are not edges of $\mathcal{J}$, but, more importantly, Observation 2.5 implies that each such L is an independent set in $\mathcal{J}$.

We can now see that item (i) in Theorem 4.1 could be equivalently phrased as “for all $I \in \mathcal{I}(\mathcal{J})$, there exists $X \in \mathcal{X}$ such that $I \subset X$ ”. This formulation suggests applying a container theorem to this hypergraph, but the edges of $\mathcal{J}$ could have size comparable to $n = |V|$. To avoid that issue, we first reduce to an alternate setting involving independent sets in s -uniform hypergraphs.

To state the theorem that we use in that reduction, we need two definitions. The first is standard: we say that a hypergraph $\mathcal{C}$ covers another hypergraph $\mathcal{G}$ when $\mathcal{G} \subset \langle \mathcal{C} \rangle$, where

$$\langle \mathcal{C} \rangle = \{L \subset V(\mathcal{C}) : \exists A \in \mathcal{C}, A \subset L\}$$

is the up-set of $\mathcal{C}$. The second is that of the non-strict link of a hypergraph, which is deceptively similar to the ordinary notion of hypergraph link.

Definition 4.2. For a hypergraph $\mathcal{G}$ and a set Y , let

$$\partial_Y \mathcal{G} = \{E \setminus Y : E \in \mathcal{G}\}$$

denote the non-strict link of $\mathcal{G}$ with respect to Y .

Interestingly, this reduction to s -uniform hypergraphs is proven by applying a container theorem (Theorem 4.3, below) that has no dependency on the uniformity of the hypergraph. It is not immediate that Theorem E in [7] implies Theorem 4.3, so we give a short (and, in fact, self-contained) proof of the result we will use in Appendix A. In the statement and in the rest of the paper, when $0 \leq q \leq 1$, we write V_q to denote a q -random subset of V .

Theorem 4.3 (Campos and Samotij [7, modified Theorem E]). *Let $\mathcal{G}$ be a hypergraph with vertex set V . For all $\alpha, q \in \mathbb{R}$ satisfying $0 < q \leq \alpha < 1$, there exists a family $\mathcal{T} \subset 2^V$ and a function $\varphi : \mathcal{I}(\mathcal{G}) \rightarrow \mathcal{T}$ such that:*

- (a) *For each $I \in \mathcal{I}(\mathcal{G})$, we have $\varphi(I) \subset I$.*
- (b) *Each $T \in \mathcal{T}$ has at most qn/α elements, where $n = |V|$.*
- (c) *For every $T \in \mathcal{T}$, there exists a hypergraph $\mathcal{C}_T$ with vertex set V that covers $\mathcal{G}$ and satisfies*

$$\mathbb{P}(L \subset V_q \mid V_q \in \mathcal{I}(\partial_T \mathcal{G})) > (1 - \alpha)^{|L|} q^{|L|} \quad (44)$$

for all $L \notin \mathcal{C}_T$; moreover, for all $I \in \mathcal{I}(\mathcal{G})$ such that $T = \varphi(I)$, we have $I \in \mathcal{I}(\mathcal{C}_T)$.

We will apply Theorem 4.3 to $\mathcal{J}$, obtaining as a result a family of sets $T \in \mathcal{T}$, each with a corresponding hypergraph $\mathcal{C}_T$. Now, for each $I \in \mathcal{I}(\mathcal{J})$, there is some $T \in \mathcal{T}$ such that $I \in \mathcal{I}(\mathcal{C}_T)$ by item (c) in Theorem 4.3. Therefore,

$$\mathcal{I}(\mathcal{J}) \subset \bigcup_{T \in \mathcal{T}} \mathcal{I}(\mathcal{C}_T). \quad (45)$$

However, we also want the hypergraphs in the right-hand side of (45) to be s -uniform. The first step is replacing each $\mathcal{C}_T$ by its up-set.

Observation 4.4. *Let $\mathcal{G}$ be a hypergraph. If $I \in \mathcal{I}(\mathcal{G})$, then $I \in \mathcal{I}(\langle \mathcal{G} \rangle)$.*

Proof. Assume that $I \notin \mathcal{I}(\langle \mathcal{G} \rangle)$. Hence, there is $E \in \langle \mathcal{G} \rangle$ such that $E \subset I$, and by definition of the up-set, there is also $E' \in \mathcal{G}$ such that $E' \subset E \subset I$, so $I \notin \mathcal{I}(\mathcal{G})$. $\square$

The trivial observation that completes the reduction to s -uniform hypergraphs says that Observation 4.4 also holds if we replace $\mathcal{C}_T$ by a subhypergraph. In our case, the implication is that if $\mathcal{C}'_T \subset \langle \mathcal{C}_T \rangle$ for all $T \in \mathcal{T}$, then

$$\mathcal{I}(\mathcal{J}) \subset \bigcup_{T \in \mathcal{T}} \mathcal{I}(\mathcal{C}'_T).$$

The subhypergraph $\mathcal{C}'_T$ that we will take is simply the set of edges with size s , i.e.

$$\mathcal{C}'_T = \langle \mathcal{C}_T \rangle_{=s}$$

where we define

$$\mathcal{G}_{=s} = \{E \in \mathcal{G} : |E| = s\}$$

for any hypergraph $\mathcal{G}$ and $s \in \mathbb{N}$.

Having reduced the problem to a collection of hypergraphs with edges of size s , we can apply another container theorem to each $\mathcal{C}'_T$. Recall that we have previously used the next statement in Section 3, and that its easy deduction from a result in the literature [7, Theorem A] can be found in Appendix B.

Theorem 3.4 (Campos and Samotij [7, modified Theorem A]). *Let $\mathcal{G}$ be an s -uniform hypergraph with n vertices. For all $0 < \zeta \leq 1$ and $0 < p \leq \zeta/(8s^2)$, there is a family $\mathcal{S} \subset 2^{V(\mathcal{G})}$ and functions*

$$\phi : \mathcal{I}(\mathcal{G}) \rightarrow \mathcal{S} \quad \text{and} \quad \psi : \mathcal{S} \rightarrow 2^{V(\mathcal{G})} \quad (27)$$

such that:

- (i) For each $I \in \mathcal{I}(\mathcal{G})$, we have $\phi(I) \subset I \subset \psi(\phi(I))$.
- (ii) Each $S \in \mathcal{S}$ has at most $8s^2pn/\zeta$ elements.
- (iii) For every $S \in \mathcal{S}$, letting $X = \psi(S)$, $\mathcal{G}[X]$ is not $(p, \zeta p|X|)$ -Janson.

Consider the following recap of our overview so far. First, we take an $I \in \mathcal{I}(\mathcal{J})$, and from Theorem 4.3 and Observation 4.4 we obtain $\varphi(I) = T \subset I$ and $\mathcal{C}'_T$ such that $I \in \mathcal{I}(\mathcal{C}'_T)$. Then, applying Theorem 3.4 with $\mathcal{G} = \mathcal{C}'_T$ yields a family $\mathcal{S}_T$ from which we retrieve a container $X \supset I$. This container has the property that $\mathcal{C}'_T[X]$ is not $(p, 2^{-8}p|X|)$ -Janson, and we want to take X to be the container for this fixed I in Theorem 4.1. It is easy to deduce the claimed bound (43) on the number of such containers from the bounds on $|\mathcal{T}|$ and $|\mathcal{S}_T|$ given by the corresponding container theorems. However, it is not yet clear how to deduce that $\mathcal{H}[X]$ is not (p, R) -Janson when we only know the Janson properties of $\mathcal{C}'_T[X]$.

To show that $\mathcal{H}[X]$ is not (p, R) -Janson, we fix an arbitrary $\nu : \mathcal{H}[X] \rightarrow \mathbb{R}_{\geq 0}$ with the objective of establishing that

$$\Lambda_p(\nu) \geq \frac{e(\nu)^2}{R} \quad (46)$$

which, recall, is the definition of what it means to not be (p, R) -Janson. We will consider two cases, depending on the relation between ν and ν' , the restriction of ν to $\mathcal{C}'_T[X]$, where $T = \varphi(I)$. The first (easier) case is when $e(\nu') \geq e(\nu)/2$. Here, we can easily show that (46) follows from $\mathcal{C}'_T[X]$ not being $(p, 2^{-8}p|X|)$ -Janson (see Claim 4.7).

In the other, more delicate case, we will have that $\nu'' = \nu - \nu'$ satisfies

$$e(\nu'') \geq \frac{e(\nu)}{2}. \quad (47)$$

Our goal is now to obtain bounds relating $\Lambda_p(\nu)$ and $e(\nu'')^2$ – it will be easy to see that combining them with (47) will reach (46). Concretely, we will show that

$$2^{2s}\Lambda_p(\nu) > \mathbb{E}[\Lambda_{p/q}(\nu''_q) \mid V_q \in \mathcal{I}(\partial_T \mathcal{J})] \geq \frac{\mathbb{E}[e(\nu''_q)^2 \mid V_q \in \mathcal{I}(\partial_T \mathcal{J})]}{\eta R} \geq \frac{e(\nu'')^2}{\eta R} \quad (48)$$

where $\nu''_q : \mathcal{H}[X] \rightarrow \mathbb{R}_{\geq 0}$ is defined by

$$\nu''_q(E) = \frac{\mathbb{1}[E \subset V_q]}{\mathbb{P}(E \subset V_q \mid V_q \in \mathcal{I}(\partial_T \mathcal{J}))} \nu''(E).$$

The simple proof of Claim 4.8 will establish, by inspecting the definitions, that

$$\mathbb{E}[e(\nu''_q) \mid V_q \in \mathcal{I}(\partial_T \mathcal{J})] = e(\nu''),$$

which implies the rightmost inequality of (48) by convexity. To prove the second inequality, we will in fact show that (deterministically) whenever $V_q \in \mathcal{I}(\partial_T \mathcal{J})$, we have

$$\Lambda_{p/q}(\nu_q'') \geq \frac{e(\nu_q'')^2}{\eta R}. \quad (49)$$

Proving (49) will require a trivial observation about independent sets in the non-strict link of hypergraphs with respect to a set T .

Observation 4.5. *Let $\mathcal{G}$ be a hypergraph and $T \subset V(\mathcal{G})$. If $I \in \mathcal{I}(\partial_T \mathcal{G})$, then $I \in \mathcal{I}(\mathcal{G})$.*

Proof. For each $E \in \mathcal{G}$, observe that $E \setminus T \subset E$ and $E \setminus T \in \partial_T \mathcal{G}$ by definition, so $\mathcal{G} \subset \langle \partial_T \mathcal{G} \rangle$. The statement now follows from Observation 4.4. $\square$

With Observation 4.5, we will see that the definition of $\mathcal{J}$ and the monotonicity of being $(p/q, \eta R)$ -Janson will imply (49), see Claim 4.9. The remaining inequality, Claim 4.10, establishes that

$$2^{2s} \Lambda_p(\nu) > \mathbb{E}[\Lambda_{p/q}(\nu_q'') \mid V_q \in \mathcal{I}(\partial_T \mathcal{J})]. \quad (50)$$

To prove it, we will crucially rely on the fact that, for all $E \in \mathcal{H}[X] \setminus \mathcal{C}'_T$,

$$\mathbb{P}(E \subset V_q \mid V_q \in \mathcal{I}(\partial_T \mathcal{J})) > \left(\frac{q}{2}\right)^{|E|}$$

since $\mathcal{H}[X] \setminus \mathcal{C}'_T$ and $\mathcal{C}_T$ are disjoint, and $\mathcal{C}_T$ satisfies item (c) of Theorem 4.3 with $\alpha = 1/2$.

The preceding overview and proof strategy in fact proves Theorem 4.6, which strengthens the original statement and adds the characterization of sets L for which $\mathcal{H}$ is not $(p/q, \eta R)$ -Janson as independent sets in the auxiliary hypergraph $\mathcal{J}$. To obtain this stronger statement, we will redefine $\mathcal{J}$ in (52) to consider instead sets L for which $\mathcal{H}[L]$ is $(p/(q-p), \eta R)$ -Janson. We will then deduce Theorem 4.1 from Theorem 4.6 by applying the latter with q being $q+p$.

Theorem 4.6. *Let $s, n \in \mathbb{N}$ with $s \leq n$, and let $p, q, \alpha, R, \eta > 0$ satisfy*

$$p \leq \frac{1}{2^{11}s^2}, \quad 2p \leq q \leq \alpha < 1, \quad R \geq 2^{-6}pn \quad \text{and} \quad \eta \leq \frac{(1-\alpha)^{2s}}{4}. \quad (51)$$

Furthermore, for every s -uniform hypergraph $\mathcal{H}$ with n vertices, let

$$\mathcal{J} = \{L \subset V(\mathcal{H}) : \mathcal{H}[L] \text{ is } (p/(q-p), \eta R)\text{-Janson}\}. \quad (52)$$

There exists a family $\mathcal{Y} \subset 2^{V(\mathcal{H})} \times 2^{V(\mathcal{H})}$ and functions

$$g : \mathcal{I}(\mathcal{J}) \rightarrow \mathcal{Y} \quad \text{and} \quad f : \mathcal{Y} \rightarrow 2^{V(\mathcal{H})}$$

such that:

- (1) *For every $I \in \mathcal{I}(\mathcal{J})$, if $g(I) = (S, T)$, then $S \cup T \subset I \subset f(S, T)$.*
- (2) *Each $(S, T) \in \mathcal{Y}$ satisfies*

$$|S| \leq 2^{11}ps^2n \quad \text{and} \quad |T| \leq qn/\alpha. \quad (53)$$

- (3) *For all $Y \in \mathcal{Y}$, the hypergraph $\mathcal{H}[X]$ is not (p, R) -Janson, where $X = f(Y)$.*

Before proving Theorem 4.6, let us quickly observe that it implies Theorem 4.1.

Proof that Theorem 4.6 implies Theorem 4.1. Apply Theorem 4.6 to $\mathcal{H}$ with $\alpha = 1/2$ and q replaced by $q + p$. Note that

$$p \leq \frac{q}{2^{10}s^2} \leq \frac{1}{2^{11}s^2} \quad \text{and} \quad 2p \leq q + p \leq \alpha,$$

since $q \leq 1/16$. As a result, we obtain functions g, f and a family $\mathcal{Y}$, so let

$$\mathcal{X} = \{f(S, T) : (S, T) \in \mathcal{Y}\}.$$

Observe that

$$|\mathcal{X}| \leq \sum_{m=0}^{2^{11}ps^2n} \binom{n}{m} \sum_{t=0}^{2(q+p)n} \binom{n}{t} \leq \left(\frac{2}{q}\right)^{8qn} \quad (54)$$

by enumerating every possible S and T , and combining the bound on their sizes, (53), with

$$2^{11}ps^2n \leq 2(q+p)n \leq 4qn \leq \frac{n}{4}. \quad (55)$$

This choice of $\mathcal{X}$ therefore satisfies the bound on the size of the family in Theorem 4.1.

Take an arbitrary $L \subset V(\mathcal{H})$ such that $\mathcal{H}[L]$ is not $(p/q, \eta R)$ -Janson, and note that $L \notin E(\mathcal{J})$, by our choice of q as $q + p$, and hence $L \in \mathcal{I}(\mathcal{J})$, by Observation 2.5. Applying item (1) of Theorem 4.6 implies that there is $(S, T) \in \mathcal{Y}$ such that $L \subset f(S, T)$, which establishes item (i) of Theorem 4.1. Item (ii) of Theorem 4.1 follows from our choice of $\mathcal{X}$ and item (3) of Theorem 4.6. $\square$

We now proceed to prove Theorem 4.6.

Proof of Theorem 4.6. Apply Theorem 4.3 with $\mathcal{G} = \mathcal{J}$ and parameters q and α to obtain $\mathcal{T}$ and φ . Now, for each $T \in \mathcal{T}$, there is $\mathcal{C}_T$ satisfying item (c) in Theorem 4.3, so we let $\mathcal{C}'_T = \langle \mathcal{C}_T \rangle_s$ be the edges of $\langle \mathcal{C}_T \rangle$ with size s . As $\mathcal{C}'_T$ is s -uniform and $p \leq 1/(2^{11}s^2)$ by (51), we can apply Theorem 3.4 with $\mathcal{G} = \mathcal{C}'_T$ and $\zeta = 2^{-8}$, obtaining as a result $\mathcal{S}_T$, ψ_T and ϕ_T .

Fix an $I \in \mathcal{I}(\mathcal{J})$ and note that if $T = \varphi(I)$, then it follows from the “moreover” part in item (c) of Theorem 4.3 that $I \in \mathcal{I}(\mathcal{C}_T)$. Combining this with Observation 4.4 and $\mathcal{C}'_T \subset \langle \mathcal{C}_T \rangle$, we conclude that $I \in \mathcal{I}(\mathcal{C}'_T)$. As we have applied Theorem 3.4 with $\mathcal{G} = \mathcal{C}'_T$ for every $T \in \mathcal{T}$, and since $I \in \mathcal{I}(\mathcal{C}'_T)$, we obtain, by item (i), sets $S = \phi_T(I) \in \mathcal{S}_T$ and $X = \psi_T(S)$, where $T = \varphi(I)$, such that

$$S \cup T \subset I \subset X. \quad (56)$$

We then define $g(I) = (S, T)$ and $f(S, T) = X$ for $T = \varphi(I)$ and $S = \phi_T(I)$ and set

$$\mathcal{Y} = \{g(I) : I \in \mathcal{I}(\mathcal{J})\},$$

which, by (56) and the fact that I was arbitrary, is a definition that satisfies item (1) of Theorem 4.6. Moreover, each $T \in \mathcal{T}$ and $S \in \mathcal{S}_T$ satisfy

$$|T| \leq qn/\alpha \quad \text{and} \quad |S| \leq 2^{11}s^2pn \quad (57)$$

by item (b) in Theorem 4.3 and item (ii) in Theorem 3.4 with our choice of $\zeta = 2^{-8}$, which proves item (2). It remains only to show that item (3) holds.

Take an arbitrary $(S, T) \in \mathcal{Y}$ with $X = f(S, T)$ with the goal of showing that $\mathcal{H}[X]$ is not (p, R) -Janson. To do so, it suffices to show that, for any $\nu : \mathcal{H}[X] \rightarrow \mathbb{R}_{\geq 0}$, we have

$$\Lambda_p(\nu) \geq \frac{e(\nu)^2}{R}, \quad (58)$$

so we fix a measure $\nu : \mathcal{H}[X] \rightarrow \mathbb{R}_{\geq 0}$ and want to establish (58).

Recall that we have defined $\mathcal{C}'_T = \langle \mathcal{C}_T \rangle_{=s}$, where $\mathcal{C}_T$ is the hypergraph given by item (c) in Theorem 4.3 for $T \in \mathcal{T}$. Let ν' be the restriction of ν to $\mathcal{H}[X] \cap \mathcal{C}'_T[X]$, i.e. for each $E \in \mathcal{H}[X]$, let

$$\nu'(E) = \begin{cases} \nu(E) & \text{if } E \in \mathcal{C}'_T[X], \\ 0 & \text{otherwise.} \end{cases}$$

Claim 4.7. *If the measure ν' satisfies*

$$e(\nu') \geq \frac{e(\nu)}{2}, \quad (59)$$

then (58) holds.

Proof. Assume that (59) holds. We have

$$\Lambda_p(\nu) \geq \Lambda_p(\nu') \geq \frac{2^8 e(\nu')^2}{p|X|} \geq \frac{4e(\nu')^2}{R} \geq \frac{e(\nu)^2}{R},$$

first because $\nu' \leq \nu$ and $\Lambda_p(\cdot)$ is monotone increasing, second since $\mathcal{C}'_T[X]$ is not $(p, 2^{-8}p|X|)$ -Janson by item (iii) of Theorem 3.4 and our choice of $\zeta = 2^{-8}$, then because $R \geq 2^{-6}pn$ by (51), and the last step is due to (59). $\blacksquare$

We now define the measure $\nu'' = \nu - \nu'$, which corresponds to the restriction of ν to the hypergraph $\mathcal{H}' := \mathcal{H}[X] \setminus \mathcal{C}'_T = \mathcal{H}[X] \setminus \mathcal{C}'_T[X]$. By Claim 4.7, we may assume that

$$e(\nu'') = e(\nu) - e(\nu') > \frac{e(\nu)}{2}, \quad (60)$$

otherwise we are done.

With the goal of defining a random measure ν''_q on the hypergraph induced by the random set $X_q = V_q \cap X$, where V_q is a q -random subset of V , we introduce some notation. First, let

$$P_q(E) = \mathbb{P}(E \subset V_q \mid V_q \in \mathcal{I}(\partial_T \mathcal{J})) \quad (61)$$

and observe that

$$P_q(E) > (1 - \alpha)^{|E|} q^{|E|} \quad (62)$$

for all $E \in \mathcal{H}'$, by (44), since the hypergraphs $\mathcal{H}'$ and $\mathcal{C}_T$ are disjoint. Indeed, $\mathcal{H}$ is s -uniform, which means that so are $\mathcal{H}'$ and $\mathcal{H}' \cap \mathcal{C}_T$. The only edges of $\mathcal{C}_T$ that could be edges of $\mathcal{H}'$ thus have size exactly equal to s . But every s -sized edge of $\mathcal{C}_T$ is also an edge of $\mathcal{C}'_T = \langle \mathcal{C}_T \rangle_{=s}$, and is therefore not in $\mathcal{H}' = \mathcal{H}[X] \setminus \mathcal{C}'_T$.

Now, let $\nu''_q : \mathcal{H}' \rightarrow \mathbb{R}_{\geq 0}$ be defined by

$$\nu''_q(E) = \frac{\mathbb{1}[E \subset V_q]}{\mathbb{P}(E \subset V_q \mid V_q \in \mathcal{I}(\partial_T \mathcal{J}))} \nu''(E).$$

We will first show that

$$\mathbb{E}[e(\nu''_q) \mid V_q \in \mathcal{I}(\partial_T \mathcal{J})] = \sum_{E \in \mathcal{H}'} \nu''(E) = e(\nu''),$$

which will allow us to relate $e(\nu''_q)^2$ to $e(\nu'')^2$.

Claim 4.8.

$$\mathbb{E}[e(\nu''_q)^2 \mid V_q \in \mathcal{I}(\partial_T \mathcal{J})] \geq e(\nu'')^2.$$

Proof. The definitions of $e(\nu_q'')$ and ν_q'' ,

$$e(\nu_q'') = \sum_{E \in \mathcal{H}'} \nu_q''(E) = \sum_{E \in \mathcal{H}'} \frac{\nu''(E) \mathbb{1}[E \subset V_q]}{P_q(E)},$$

imply that

$$\mathbb{E}[e(\nu_q'') \mid V_q \in \mathcal{I}(\partial_T \mathcal{J})] = \sum_{E \in \mathcal{H}'} \nu''(E) = e(\nu''), \quad (63)$$

since, for all $E \in \mathcal{H}'$, we have that

$$\mathbb{E}[\mathbb{1}[E \subset V_q] \mid V_q \in \mathcal{I}(\partial_T \mathcal{J})] = P_q(E)$$

from (61), the definition of $P_q(E)$. The claim now follows from (63) by Jensen's inequality. $\blacksquare$

The next step is relating $\Lambda_{p/(q-p)}(\nu_q'')$ and $e(\nu_q'')$ when $V_q \in \mathcal{I}(\partial_T \mathcal{J})$.

Claim 4.9. *If $V_q \in \mathcal{I}(\partial_T \mathcal{J})$, then*

$$\Lambda_{p/(q-p)}(\nu_q'') \geq \frac{e(\nu_q'')^2}{\eta R}.$$

Proof. It follows from $V_q \in \mathcal{I}(\partial_T \mathcal{J})$ and Observation 4.5 that $V_q \in \mathcal{I}(\mathcal{J})$, and hence $\mathcal{H}[V_q]$ is not $(p/(q-p), \eta R)$ -Janson, by the definition of $\mathcal{J}$, (52). In particular, it follows that

$$\Lambda_{p/(q-p)}(\nu_q'') \geq \frac{e(\nu_q'')^2}{\eta R}$$

as ν_q'' is also a measure supported on $\mathcal{H}[V_q]$ by $\mathcal{H}' \subset \mathcal{H}$. $\blacksquare$

Our final inequality bounds $\Lambda_{p/(q-p)}(\nu_q'')$ in expectation by $\Lambda_p(\nu)$, up to an exponential factor.

Claim 4.10.

$$\mathbb{E}[\Lambda_{p/(q-p)}(\nu_q'') \mid V_q \in \mathcal{I}(\partial_T \mathcal{J})] < (1 - \alpha)^{-2s} \Lambda_p(\nu).$$

Proof. Recall that $d_{\nu_q''}(L)$ is defined as

$$d_{\nu_q''}(L) = \sum_{L \subset E \in \mathcal{H}'} \nu_q''(E) = \sum_{L \subset E \in \mathcal{H}'} \frac{\nu''(E) \mathbb{1}[E \subset V_q]}{P_q(E)},$$

where the last equality is using the definition of ν_q'' , and hence we can write, for every $L \subset V$,

$$d_{\nu_q''}(L)^2 = \sum_{L \subset E_1 \in \mathcal{H}'} \frac{\nu''(E_1)}{P_q(E_1)} \sum_{L \subset E_2 \in \mathcal{H}'} \frac{\nu''(E_2)}{P_q(E_2)} \mathbb{1}[E_1 \cup E_2 \subset V_q].$$

Observe that the event $V_q \in \mathcal{I}(\partial_T \mathcal{J})$ is decreasing and also that, for every $E_1, E_2 \in \mathcal{H}'$, the event $E_1 \cup E_2 \subset V_q$ is increasing. We can therefore use Harris' inequality to bound, for $E_1, E_2 \in \mathcal{H}'$,

$$\mathbb{P}(E_1 \cup E_2 \subset V_q \mid V_q \in \mathcal{I}(\partial_T \mathcal{J})) \leq \mathbb{P}(E_1 \cup E_2 \subset V_q) = q^{|E_1 \cup E_2|} = q^{2s - |E_1 \cap E_2|} \quad (64)$$

because $\mathcal{H}'$ is s -uniform. Taking the conditional expectation and applying (64), we obtain

$$\mathbb{E}[d_{\nu_q''}(L)^2 \mid V_q \in \mathcal{I}(\partial_T \mathcal{J})] < (1 - \alpha)^{-2s} \sum_{L \subset E_1 \in \mathcal{H}'} \nu(E_1) \sum_{L \subset E_2 \in \mathcal{H}'} \nu(E_2) q^{-|E_1 \cap E_2|} \quad (65)$$

where we used $P_q(E) > (1 - \alpha)^s q^s$, by (62) and since $\mathcal{H}'$ is s -uniform, and the fact that $\nu'' \leq \nu$.

By the definition (8) of $\Lambda_{p/(q-p)}(\nu''_q)$,

$$\mathbb{E}[\Lambda_{p/(q-p)}(\nu''_q) \mid V_q \in \mathcal{I}(\partial_T \mathcal{J})] = \sum_{\substack{L \subset V \\ |L| \geq 2}} \mathbb{E}[d_{\nu''_q}(L)^2 \mid V_q \in \mathcal{I}(\partial_T \mathcal{J})] \left(\frac{q-p}{p}\right)^{|L|}, \quad (66)$$

and then applying (65) to each $d_{\nu''_q}(L)$ term in (66) yields

$$\mathbb{E}[\Lambda_{p/(q-p)}(\nu''_q) \mid V_q \in \mathcal{I}(\partial_T \mathcal{J})] < (1-\alpha)^{-2s} \sum_{\substack{L \subset V \\ |L| \geq 2}} \sum_{L \subset E_1 \in \mathcal{H}'} \sum_{L \subset E_2 \in \mathcal{H}'} \frac{\nu(E_1)\nu(E_2)}{q^{|E_1 \cap E_2|}} \left(\frac{q-p}{p}\right)^{|L|}$$

or, equivalently,

$$\mathbb{E}[\Lambda_{p/(q-p)}(\nu''_q) \mid V_q \in \mathcal{I}(\partial_T \mathcal{J})] < (1-\alpha)^{-2s} \sum_{\substack{E_1, E_2 \in \mathcal{H}' \\ |E_1 \cap E_2| \geq 2}} \frac{\nu(E_1)\nu(E_2)}{q^{|E_1 \cap E_2|}} \sum_{\ell=2}^s \left(\frac{q-p}{p}\right)^\ell \binom{|E_1 \cap E_2|}{\ell} \quad (67)$$

by first choosing $E_1, E_2 \in \mathcal{H}'$ and then $L \subset E_1 \cap E_2$, grouping terms according to $\ell = |L|$. Bounding the innermost sum in (67) for fixed $E_1, E_2 \in \mathcal{H}'$ with $|E_1 \cap E_2| \geq 2$ then yields

$$\sum_{\ell=2}^s \left(\frac{q-p}{p}\right)^\ell \binom{|E_1 \cap E_2|}{\ell} \leq \left(\frac{q}{p}\right)^{|E_1 \cap E_2|},$$

which replaced in (67) and simplified, results in

$$\mathbb{E}[\Lambda_{p/(q-p)}(\nu''_q) \mid V_q \in \mathcal{I}(\partial_T \mathcal{J})] < (1-\alpha)^{-2s} \sum_{\substack{E_1, E_2 \in \mathcal{H}' \\ |E_1 \cap E_2| \geq 2}} \frac{\nu(E_1)\nu(E_2)}{p^{|E_1 \cap E_2|}}. \quad (68)$$

To complete the proof, note that

$$\sum_{\substack{E_1, E_2 \in \mathcal{H}' \\ |E_1 \cap E_2| \geq 2}} \frac{\nu(E_1)\nu(E_2)}{p^{|E_1 \cap E_2|}} \leq \sum_{\substack{L \subset V \\ |L| \geq 2}} \sum_{L \subset E_1 \in \mathcal{H}} \sum_{L \subset E_2 \in \mathcal{H}} \frac{\nu(E_1)\nu(E_2)}{p^{|L|}} = \sum_{\substack{L \subset V \\ |L| \geq 2}} d_\nu(L)^2 p^{-|L|},$$

where the last term is equal to $\Lambda_p(\nu)$ by definition, so we obtain, replacing it back in (68), the inequality that we wanted. $\blacksquare$

Observe that Claim 4.9 implies that

$$\mathbb{E}[\Lambda_{p/(q-p)}(\nu''_q) \mid V_q \in \mathcal{I}(\partial_T \mathcal{J})] \geq \frac{\mathbb{E}[(\nu''_q)^2 \mid V_q \in \mathcal{I}(\partial_T \mathcal{J})]}{\eta R}. \quad (69)$$

Combining Claim 4.10 and Claim 4.8 with (69) yields

$$(1-\alpha)^{-2s} \Lambda_p(\nu) \geq \frac{e(\nu'')^2}{\eta R}$$

and thus, since we are in the case where (60) holds, it follows that

$$\Lambda_p(\nu) \geq \frac{e(\nu)^2}{R} \quad (70)$$

by our choice of η satisfying $4\eta \leq (1-\alpha)^{2s}$. As (70) was exactly our goal, (58), and ν was arbitrary, we conclude that $\mathcal{H}[X]$ is not (p, R) -Janson. Moreover, our choice of $(S, T) \in \mathcal{Y}$ was also arbitrary, so we have established that item (3) holds, and the proof is complete. $\square$

5. EXTENDING COLLECTIONS OF COPIES

In this section, we use a novel container theorem to prove the core statement that we need in the proof of Lemma 2.8. We defer the proof of this container theorem to Section 7, since that is the most technical part of the entire argument.

The setting is very similar to Lemma 3.1, but now we will be able to extend many copies of F^- to F by adding a single vertex v to $U = V(\tilde{G}) = V(\tilde{G}')$. Moreover, we will be able to show that the set of copies created by adding v to U will be well-distributed in relation to the copies of F fully contained in U . To obtain this stronger conclusion, we assume that, besides $\mathfrak{I}_{F^-, \tilde{G}', \tilde{G}}[W]$ being (p, R) -Janson for every $W \subset U = V(\tilde{G})$ with $|W| \geq m/(8r)$, we also have that $\mathfrak{I}_{F, G', G}[U]$ is (p, R') -Janson. Under these circumstances, the lemma states that, when G is distributed as $\mathbb{G}(m+1, 2)$ conditioned on $\{G[U] = \tilde{G}\}$, the following holds with extremely high probability: for every choice of $G' \subset G$ such that $N_{G'}(v)$ is not too small, the hypergraph $\mathfrak{I}_{F, G', G}$ is $(p, R' + 1)$ -Janson.

Lemma 5.1. *Let $m, k, r, s \in \mathbb{N}$ with $r \geq 2$, $s < k \leq m$, and let*

$$p = \frac{1}{2^{25} k^2 r^4}, \quad m \geq r^{Ck}, \quad R = 2^{-5} r^{-1} p m \quad \text{and} \quad 0 \leq R' \leq \frac{R}{16}. \quad (71)$$

Further let F , $\tilde{G}'$ and $\tilde{G}$ be graphs such that $v(F) = s + 1 < m$, $\tilde{G}' \subset \tilde{G}$ and $v(\tilde{G}) = m$.

If $\mathfrak{I}_{F, \tilde{G}', \tilde{G}}$ is (p, R') -Janson and $\mathfrak{I}_{F^-, \tilde{G}', \tilde{G}}[W]$ is (p, R) -Janson for every $W \subset U = V(\tilde{G})$ with $|W| \geq m/(8r)$, then

$$\mathbb{P} \left(\exists G' \subset G : \begin{array}{l} G'[U] = \tilde{G}', \quad d_{G'}(v) \geq m/(4r) \text{ and} \\ \mathfrak{I}_{F, G', G} \text{ is not } (p, R' + 1)\text{-Janson} \end{array} \mid G[U] = \tilde{G} \right) \leq 2^{-m/(32r)}, \quad (72)$$

where $G \sim \mathbb{G}(m+1, 1/2)$ and $V(G) = U \cup \{v\}$.

The first change that we need to make to the proof in Section 3 is to replace Γ_G , defined in (24), by Ψ_G , which is just the collection of $G' \subset G$ satisfying the event in (72):

$$\Psi_G = \left\{ G' \subset G : \begin{array}{l} G'[U] = \tilde{G}', \quad d_{G'}(v) \geq m/(4r) \text{ and} \\ \mathfrak{I}_{F, G', G} \text{ is not } (p, R' + 1)\text{-Janson} \end{array} \right\}. \quad (73)$$

That is, we now require that $\mathfrak{I}_{F, G', G}$ is not $(p, R' + 1)$ -Janson, instead of requiring to it be empty as in Lemma 3.1. The argument follows very closely the one in Section 3, so we briefly summarise the ideas in the proof of Lemma 3.1, referring to some of the definitions in that section as we progress.

Recall that we defined

$$\mathcal{H} = \left\{ E_L \subset U \times \{0, 1\} : L \in \mathfrak{I}_{F^-, \tilde{G}', \tilde{G}} \right\},$$

where

$$E_L = \left\{ \left(u, \mathbb{1}[\phi_L(u) \in N_F(w)] \right) : u \in L \right\},$$

with $\phi_L : L \rightarrow V(F^-)$ being a fixed bijection and w being the unique vertex in $V(F) \setminus V(F^-)$. We then defined

$$\iota(G', G) = \left\{ (u, 0) : u \in N_G(v)^c \right\} \cup \left\{ (u, 1) : u \in N_{G'}(v) \right\}$$

in (25), proved Observation 3.3, which states that if $G[U] = \tilde{G}$ and $G' \in \Gamma_G$, then $\iota(G', G) \in \mathcal{I}(\mathcal{H})$, and applied a container theorem to bound the probability of that event. However, this is not immediately possible here, since Observation 3.3 is not true if we replace $G' \in \Gamma_G$ by $G' \in \Psi_G$, the analogous collection for this section: requiring $\mathfrak{I}_{F, G', G}$ to not be $(p, R' + 1)$ -Janson instead of

$\mathfrak{J}_{F,G',G} = \emptyset$ means that we are not interested in independent sets in $\mathcal{H}$, but in vertex subsets I such that $\mathcal{H}[I]$ is not $(p, R' + 1)$ -Janson.

We remedy that situation by relying on a container theorem for such sets, like the one that we proved in Section 4. However, Theorem 4.1 is not adequate for several reasons, which we discuss while introducing some notation and new definitions. We then prove the analogue of Observation 3.3 for this section, Observation 5.3, and state the container theorem that we end up using, Theorem 5.4.

It will be helpful to partition the copies of $F \subset G'[U \cup \{v\}]$ in two natural classes. The first one consists of the copies of F that use v , which correspond to copies of $F^- \subset \tilde{G}'[U]$ that are extended with the addition of $v \notin U$. Every other copy of $F \subset G'$, i.e. those that do not use v , belong in the second class, and are contained in $\tilde{G}' = G'[U]$. The next definition will relate the hypergraph of copies of F^- that can be extended with v and the hypergraph of the resulting copies of F .

Definition 5.2. For a hypergraph $\mathcal{G}$ and a vertex v not in $V(\mathcal{G})$, let

$$\bar{\partial}_v \mathcal{G} = \{E \cup \{v\} : E \in \mathcal{G}\}$$

denote the edge-wise inclusion of v in $\mathcal{G}$.

Like in Section 3, let $\pi : V(\mathcal{H}) \rightarrow U$ be the projection onto the first coordinate and define $\pi_v = \bar{\partial}_v \circ \pi$. Recalling Observation 3.5, that is,

$$\pi(\mathcal{H}) = \mathfrak{J}_{F^-, \tilde{G}', \tilde{G}},$$

we now relate $\mathfrak{J}_{F,G',G}$ to both $\mathfrak{J}_{F, \tilde{G}', \tilde{G}}$ and $\mathcal{H}$ when $G[U] = \tilde{G}$ and $G'[U] = \tilde{G}'$. We will use this fact to conclude that, if $\mathfrak{J}_{F,G',G}$ is not $(p, R' + 1)$ -Janson, then neither is $\pi_v(\mathcal{H}[I]) \cup \mathfrak{J}_{F, \tilde{G}', \tilde{G}}$ when $I = \iota(G', G)$.

Observation 5.3. For all graphs G' and G such that $G[U] = \tilde{G}$, $G' \subset G$ and $G'[U] = \tilde{G}'$, if $I = \iota(G', G)$, then

$$\pi_v(\mathcal{H}[I]) \cup \mathfrak{J}_{F, \tilde{G}', \tilde{G}} \subset \mathfrak{J}_{F, G', G}. \quad (74)$$

The inclusion in Observation 5.3 is in fact an equality, but we will not use that fact, and therefore avoid giving its (trivial) proof for the sake of brevity. As we will see, Observation 5.3 follows easily from expanding the definitions, especially after we recall (26), that is, if $I = \iota(G', G)$, then

$$I^{(0)} = N_G(v)^c \quad \text{and} \quad I^{(1)} = N_{G'}(v). \quad (75)$$

Proof of Observation 5.3. It follows immediately from

$$G'[U] = \tilde{G}' \quad \text{and} \quad G[U] = \tilde{G} \quad (76)$$

that $\mathfrak{J}_{F, \tilde{G}', \tilde{G}} \subset \mathfrak{J}_{F, G', G}$, so it only remains to show that

$$\pi_v(\mathcal{H}[I]) \subset \mathfrak{J}_{F, G', G}.$$

Let $E \in \mathcal{H}[I]$ be of the form $E = E_L$ for $L \in \mathfrak{J}_{F^-, \tilde{G}', \tilde{G}}$ and recall that $\pi(E) = L$, so our goal is to show that

$$L \cup \{v\} \in \mathfrak{J}_{F, G', G}, \quad (77)$$

where $\pi_v(E) = L \cup \{v\}$ by the definition of π_v . As $L \in \mathfrak{J}_{F^-, \tilde{G}', \tilde{G}}$, (77) follows from Observation 3.2 using (75), (76) and the fact that $E \subset I$. $\square$

In analogy to the proof in Section 3, by Observation 5.3 and the fact that the Janson property is increasing, Observation 2.5, it suffices to have a family of containers $\mathcal{X}$ with the following property. For all $\iota(G', G) = I \subset V(\mathcal{H})$ such that $\pi_v(\mathcal{H}[I]) \cup \mathfrak{J}_{F, G', G}$ is not $(p, R' + 1)$ -Janson, there is $X \in \mathcal{X}$ with $I \subset X$. This is the statement of Theorem 5.4, the container theorem that we need to prove Lemma 5.1. We state that theorem below, but defer its proof, an implementation of the methods discussed in Section 4 tailored to this specific setting, to Section 7.

Continuing the comparison with Lemma 3.1, ideally each $X \in \mathcal{X}$ would be such that $\pi(\mathcal{H}[X])$ is not (p, R) -Janson. We are unable to prove such a statement, because the inclusion of v by π_v adds a constraint in the one-degrees of the vertices in $\mathcal{H}$. To deal with this extra constraint, we (roughly) delete a small proportion of vertices to reduce the maximum degree.

Implementing this modification to our method yields something slightly weaker that nonetheless suffices: we show that if $X \in \mathcal{X}$ is sufficiently large, then there is Y covering almost all of X such that $\pi(\mathcal{H}[Y])$ is not (p, R) -Janson. Our final note before the statement is that, despite applying Theorem 5.4 with the function π being a projection, as we previously defined it in this section, we state the theorem in a slightly more general setting.

Theorem 5.4. *Let $n, r, s \in \mathbb{N}$ with $n \geq s$ and $r \geq 2$, and let $q, p, R, R', \eta \in \mathbb{R}$ satisfy*

$$0 < q < \frac{1}{8}, \quad 0 < p \leq \frac{q}{2^{11}rs^2}, \quad R = 2^{-6}r^{-1}pn, \quad 0 \leq R' \leq \frac{R}{16} \quad \text{and} \quad \eta = p^4 \left(\frac{q}{2} \right)^{4s}. \quad (78)$$

Further let $\mathcal{F}$ be a $(s+1)$ -uniform hypergraph with vertex set U that is (p, R') -Janson, let $\mathcal{H}$ be an s -uniform hypergraph with vertex set V , where $|V| = n$, and let $\pi : V \rightarrow U$ satisfy

$$|\pi(L)| \geq \frac{|L|}{2} \quad \text{for every } L \subset V \quad \text{and} \quad |\pi(E)| = |E| \quad \text{for every } E \in \mathcal{H}. \quad (79)$$

Finally, let v be a vertex not in U . There exists a family $\mathcal{X} \subset 2^V$ with

$$|\mathcal{X}| \leq \left(\frac{2}{q} \right)^{2qn} \quad (80)$$

such that the following hold.

- (1) *If $I \subset V$ and $\pi_v(\mathcal{H}[I]) \cup \mathcal{F}$ is not $(p, R' + \eta R)$ -Janson, then $I \subset X$ for some $X \in \mathcal{X}$.*
- (2) *For each $X \in \mathcal{X}$ with $|X| \geq n/(8r)$, there exists $Y \subset X$ with*

$$|Y| \geq |X| - 2^{-8}r^{-1}n \quad (81)$$

such that $\pi(\mathcal{H}[Y])$ is not (p, R) -Janson.

We are now ready to prove Lemma 5.1.

Proof of Lemma 5.1. The first step is applying Theorem 5.4 with $q = 2^{-15}r^{-2}$ and $\mathcal{F} = \mathfrak{J}_{F, \tilde{G}', \tilde{G}}$, so we must check that these choices satisfy the assumptions of the theorem.

We assumed that $\mathcal{F} = \mathfrak{J}_{F, \tilde{G}', \tilde{G}}$ is (p, R') -Janson and $(s+1)$ -uniform, and $\mathcal{H}$ being s -uniform follows from its definition and the fact that $v(F^-) = v(F) - 1$. To apply Theorem 5.4 with this choice of $\mathcal{H}$, we implicitly set

$$V = U \times \{0, 1\}, \quad n = 2m \quad \text{and} \quad R = 2^{-5}r^{-1}pm = 2^{-6}r^{-1}pn$$

and therefore the value of R coincides in both statements, resulting in the condition $0 \leq R' \leq R/16$ also being satisfied by the identical assumption in Lemma 5.1. Furthermore, our choices for the parameters $0 < q = 2^{-15}r^{-2} < 1/8$ and $p > 0$ satisfy

$$p = \frac{1}{2^{25}k^2r^4} < \frac{q}{2^{10}s^2r^2}$$

because $s < k$, and we have checked that all the conditions in (78) hold.

We now check that π satisfies (79). The first assumption follows trivially from $V = U \times \{0, 1\}$ and $\pi : V \rightarrow U$ being a projection into the first coordinate, while the other requirement is Observation 3.7. This concludes the checking of the assumptions and requirements in Theorem 5.4.

Applying Theorem 5.4, we obtain a family $\mathcal{X}$ satisfying (80) and items (1) and (2) in its statement. The next claim, a simple combination of Observation 5.3, the definition of Ψ_G and the choice of q , shows that item (1) holds for $I = \iota(G', G)$ when $G' \in \Psi_G$, where, recall,

$$\iota(G', G) = \{(u, 0) : u \in N_G(v)^c\} \cup \{(u, 1) : u \in N_{G'}(v)\}.$$

Claim 5.5. *Let G' and G be graphs and let $I = \iota(G', G)$. If $G[U] = \tilde{G}$ and $G' \in \Psi_G$, then the hypergraph $\pi_v(\mathcal{H}[I]) \cup \mathcal{F}$ is not $(p, R' + \eta R)$ -Janson.*

Proof. Recall that the definition of Ψ_G , (73), implies that $G'[U] = \tilde{G}'$, so we can apply Observation 5.3 to conclude that

$$\pi_v(\mathcal{H}[I]) \cup \mathcal{F} \subset \mathfrak{J}_{F, G', G}, \quad (82)$$

where we replaced $\mathcal{F} = \mathfrak{J}_{F, \tilde{G}', \tilde{G}}$ in (74).

It also follows from $G' \in \Psi_G$ that $\mathfrak{J}_{F, G', G}$ is not $(p, R' + 1)$ -Janson, so we can combine (82) with the fact that being Janson is increasing, Observation 2.5, to deduce that $\pi_v(\mathcal{H}[I]) \cup \mathcal{F}$ is also not $(p, R' + 1)$ -Janson. But now, as we chose

$$\eta = p^4 \left(\frac{q}{2}\right)^{4s}, \quad R = 2^{-5}r^{-1}pm, \quad q = 2^{-15}r^{-2} \quad \text{and} \quad p = 2^{-25}k^{-2}r^{-4},$$

one can verify that

$$\eta R = 2^{-4s-5}r^{-1}p^5q^{4s}m \geq 1$$

by $m \geq r^{Ck}$, $s \leq k$ and $C = 300$. We therefore conclude that $\pi_v(\mathcal{H}[I]) \cup \mathcal{F}$ is not $(p, R' + \eta R)$ -Janson, because being (p, R) -Janson is decreasing in R by Observation 2.3. $\blacksquare$

By item (1) in Theorem 5.4 and Claim 5.5, for all graphs G' and G such that $G[U] = \tilde{G}$ and $G' \in \Psi_G$, there exists $X \in \mathcal{X}$ such that $\iota(G', G) \subset X$. Let

$$\mathcal{X}' = \{X \in \mathcal{X} : |X^{(1)}| \geq m/(4r)\},$$

and we will show that there is also $X \in \mathcal{X}'$ such that $\iota(G', G) \subset X$.

Claim 5.6. *If $G[U] = \tilde{G}$ and $G' \in \Psi_G$, then there exists $X \in \mathcal{X}'$ such that $\iota(G', G) \subset X$.*

Proof. Fix G with $G[U] = \tilde{G}$ and $G' \in \Psi_G$, and let $X \in \mathcal{X}$ be such that $I = \iota(G', G) \subset X$. By the definition of ι , we have

$$N_{G'}(v) = I^{(1)} \subset X^{(1)} \quad \text{and} \quad N_G(v)^c = I^{(0)} \subset X^{(0)},$$

and therefore

$$|X^{(1)}| \geq d_{G'}(v) \geq \frac{m}{4r} \quad (83)$$

where the last inequality is due to $G' \in \Psi_G$. We conclude that $X \in \mathcal{X}'$. $\blacksquare$

Taking a union bound over choices of $\mathcal{X}'$, we can bound the probability in (72) from above by

$$\mathbb{P}(\exists G' \subset G : G' \in \Psi_G \mid G[U] = \tilde{G}) \leq \sum_{X \in \mathcal{X}'} \mathbb{P}(\mathcal{N}_G(v)^c \subset X^{(0)}) \quad (84)$$

using Claim 5.6 and replacing the event $\{\iota(G', G) \subset X\}$ by $\{\mathcal{N}_G(v)^c \subset X^{(0)}\}$, which it implies by (75). We now want an upper bound for the probability of this event for each $X \in \mathcal{X}'$.

Claim 5.7. *For every $X \in \mathcal{X}'$, we have*

$$|U \setminus X^{(0)}| \geq \frac{m}{16r}. \quad (85)$$

Proof. By item (2) in Theorem 5.4, there is $Y \subset X$ with

$$|Y| \geq |X| - \frac{n}{2^{8r}} = |X^{(0)}| + |X^{(1)}| - \frac{n}{2^{8r}} \geq |X^{(0)}| + \left(\frac{1}{4r} - \frac{1}{2^{7r}}\right)m \quad (86)$$

such that the hypergraph $\pi(\mathcal{H}[Y])$ is not (p, R) -Janson, where we used (83) and $n = 2m$.

Taking $W = Y^{(0)} \cap Y^{(1)}$, observe that $\mathfrak{J}_{F^-, \tilde{G}', \tilde{G}}[W]$ is not (p, R) -Janson. To check that, recall that the hypergraph $\pi(\mathcal{H}[Y])$ is not (p, R) -Janson by item (2) in Theorem 5.4. It then follows from

$$\mathfrak{J}_{F^-, \tilde{G}', \tilde{G}}[W] = \pi(\mathcal{H})[W] \subset \pi(\mathcal{H}[Y])$$

by Observation 3.5 and Lemma 3.8 and the fact that being (p, R) -Janson is increasing, Observation 2.5, that $\mathfrak{J}_{F^-, \tilde{G}', \tilde{G}}[W]$ cannot be (p, R) -Janson.

Since we have assumed in the statement of Lemma 5.1 that $\mathfrak{J}_{F^-, \tilde{G}', \tilde{G}}[W]$ is (p, R) -Janson whenever $W \subset U$ satisfies $|W| \geq m/(8r)$, the fact that choosing $W = Y^{(0)} \cap Y^{(1)}$ results in a subhypergraph that is not (p, R) -Janson implies that

$$|Y^{(0)} \cap Y^{(1)}| < \frac{m}{8r}. \quad (87)$$

Manipulating (87), we obtain

$$|Y| = |Y^{(0)} \cup Y^{(1)}| + |Y^{(0)} \cap Y^{(1)}| < \left(1 + \frac{1}{8r}\right)|U|,$$

which combined with (86) yields

$$|X^{(0)}| < \left(1 - \frac{1}{16r}\right)m,$$

and hence (85), as desired. ■

Applying Claim 5.7 to each term in (84), we obtain

$$\mathbb{P}(\mathcal{N}_G(v)^c \subset X^{(0)}) = \mathbb{P}(\mathcal{N}_G(v)^c \cap (U \setminus X^{(0)}) = \emptyset) = 2^{-|U \setminus X^{(0)}|} \leq 2^{-m/(16r)} \quad (88)$$

for each $X \in \mathcal{X}'$, using that $G \sim \mathbb{G}(m+1, 1/2)$. Replacing (88) back in (84) yields

$$\sum_{X \in \mathcal{X}'} \mathbb{P}(\mathcal{N}_G(v)^c \subset X^{(0)}) \leq |\mathcal{X}'| 2^{-m/(16r)}.$$

Now, we can bound the size of $\mathcal{X}'$ using (80) and $2qn \leq m/(2^{13}r^2)$, where the latter holds by $n = 2m$ and our choice of $q = 2^{-15}r^{-2}$, to obtain

$$\mathbb{P}(\exists G' \subset G : G' \in \Psi_G \mid G[U] = \tilde{G}) \leq \left(\frac{2}{q}\right)^{2qn} 2^{-m/(16r)} \leq 2^{m/(2^9r) - m/(16r)} \leq 2^{-m/(32r)}$$

since $r \geq 2$. □

6. PROOF OF LEMMA 2.8

The purpose of this section is to give a proof of Lemma 2.8, restated below. We begin with an intuitive and informal overview of the proof, with the purpose of motivating the intermediate results in the section, and then introduce the details and technicalities in the subsections that follow.

Lemma 2.8. *For all $k, r \in \mathbb{N}$ with $r \geq 2$ and for all graphs $H_1, \dots, H_r$ with at most k vertices, if $N \in \mathbb{N}$ satisfies*

$$N \geq r^{C(k+t)} \quad (11)$$

for $t = \sum_{i=1}^r v(H_i)$, then

$$\mathbb{P}(G \in \mathcal{B}(\mathbf{H}) \cap \mathcal{E}(\mathbf{s})) \leq 2^{-\delta^2 N^2},$$

where $G \sim \mathbb{G}(N, 1/2)$, $\mathbf{H} = (H_i)_{i \in [r]}$ and $\mathbf{s} = (v(H_i))_{i \in [r]}$.

Our informal overview of the proof of Lemma 2.8 starts with a statement of our setting and strategy. We assume that $G \in \mathcal{B}(\mathbf{H}) \cap \mathcal{E}(\mathbf{s})$, i.e. G admits a “bad” colouring $c : E(G) \rightarrow [r]$ in which the copies of $H_i \subset G_i$ that are induced in G are not (p, pN) -Janson, even though G satisfies the inductive assumption, represented here by the event $\mathcal{E}(\mathbf{s})$. Our goal is to show that such G are extremely rare when $G \sim \mathbb{G}(N, 1/2)$, which we accomplish by applying Lemma 5.1 with $\tilde{G} = G[U]$ and $\tilde{G}' = G_\ell^{(c)}[U]$ for a certain vertex subset U and a specific colour $\ell \in [r]$. Proving the existence of this set U is not difficult, and is the main purpose of the intermediate results in this section.

To reach a point where we can apply Lemma 5.1, that is, to show that this set U exists, we combine Lemma 6.5 and Lemma 6.6. The proof of the former lemma uses the induction hypothesis to conclude that $\mathfrak{J}_{H_i^-, G_i^{(c)}, G}[W]$ is $(p, p|W|)$ -Janson for all “bad” colourings c and all large $W \subset U$.

In the proof Lemma 6.6, we (roughly) construct a set $U \subset V(G)$ vertex-by-vertex, starting from an arbitrary vertex subset of size δN . Adding a vertex v to U increments the Janson parameter of $H_i \subset G_i[U]$ for some colour i , in the sense that if $\mathfrak{J}_{H_i, G_i, G}[U]$ was (p, R_i) -Janson for some $R_i \geq 0$, then $\mathfrak{J}_{H_i, G_i, G}[U \cup \{v\}]$ is $(p, R_i + 1)$ -Janson. The set U is complete when there are no more vertices whose addition to U would increase the Janson parameter of H_i for some $i \in [r]$, and we show that the final size of U is at most $2\delta N$.

These are the two preliminaries that we require before the proof of Lemma 2.8, which we briefly and informally discuss now. With U given by Lemma 6.5 and Lemma 6.6, we will apply Lemma 5.1 to the graphs $\tilde{G} = G[U]$ and $\tilde{G}' = G_\ell^{(c)}[U]$ for a certain colour ℓ and many vertices $v \notin U$ when the colouring c is bad, relying on the independence of these events for each v to obtain the required bound on their joint probability. These vertices $v \notin U$ are chosen first to ensure that their degree to U is not small, so one colour $i_v \in [r]$ also has sufficiently many neighbours in U . We then select ℓ as the majority colour among the i_v , and further restrict to those v for which $i_v = \ell$.

To formally implement this outline, we first address a technicality: we are not able to directly prove that the final collection of copies of H_ℓ is (p, pN) -Janson, only $(p, 2^{-9}r^{-1}\delta pN)$ -Janson³. As we will see, this is not a problem, because we show in Lemma 6.2, using double counting, that a hypergraph $\mathcal{G}$ that is not (p, pN) -Janson contains a subset S of $\delta^{2/3}N$ vertices such that $\mathcal{G}[S]$ is not $(p, 2^{-9}r^{-1}\delta pN)$ -Janson, and we will be able to work entirely inside this set S . The next subsection

³This is because we can only apply Lemma 5.1 for $R' \leq R/16$, where the collection of copies of H_ℓ^- contained in W is (p, R) -Janson for every $W \subset U$ with $|W| \geq |U|/(8r)$, and the event $\mathcal{E}(\mathbf{s})$ only tells us that this hypergraph is $(p, p|W|)$ -Janson.

proves Lemma 6.2, and the rest of the section implements the above outline to show that U exists, and finally to prove Lemma 2.8 in Subsection 6.3.

6.1. Changing to another bad event. First, recall the definition of the bad event $\mathcal{B}(\mathbf{H})$.

Definition 2.7. Given a collection of graphs $H_1, \dots, H_r$, let $\mathbf{H} = (H_i)_{i \in [r]}$ and let $\mathcal{B}(\mathbf{H})$ be the family of graphs G with the following property. There exists a colouring $c : E(G) \rightarrow [r]$ such that, for every $i \in [r]$, the hypergraph $\mathcal{I}_{H_i, G_i, G}$ is not $(p, pv(G))$ -Janson.

As previously mentioned, the first stage in the proof is to show that we can replace $\mathcal{B}(\mathbf{H})$ by an alternative event $\mathcal{B}'(\mathbf{H})$. The main advantage of this change is that it reduces the Janson parameter by a factor of order $r^{-1}\delta$, at the cost of assuming that it only holds for a single subset S of size $|S| \geq \delta^{2/3}N$.

Definition 6.1. Given a collection of graphs $H_1, \dots, H_r$, let $\mathbf{H} = (H_i)_{i \in [r]}$ and let $\mathcal{B}'(\mathbf{H})$ be the family of graphs G with the following property. There exists $S \subset V(G)$ with $|S| \geq \delta^{2/3}v(G)$ and a colouring $c : E(G[S]) \rightarrow [r]$ such that $\mathcal{I}_{H_i, G_i, G}[S]$ is not $(p, 2^{-9}r^{-1}\delta pv(G))$ -Janson for all $i \in [r]$.

It is crucial that the original bad event $\mathcal{B}(\mathbf{H})$ is contained in the variant $\mathcal{B}'(\mathbf{H})$, which we now prove with a simple double counting argument.

Lemma 6.2. For every $k \in \mathbb{N}$ and graphs $H_1, \dots, H_r$,

$$\mathcal{B}(\mathbf{H}) \subset \mathcal{B}'(\mathbf{H}),$$

where $\mathbf{H} = (H_i)_{i \in [r]}$.

As we need to work with measures in the Janson property, the following observation, albeit a trivial consequence of the definitions, will be useful when proving Lemma 6.2.

Observation 6.3. For all $s \in \mathbb{N}$ and hypergraphs $\mathcal{G}$, if $\vartheta_1, \dots, \vartheta_s : \mathcal{G} \rightarrow \mathbb{R}_{\geq 0}$ satisfy

$$\vartheta = \sum_{i=1}^s \vartheta_i,$$

then,

$$e(\vartheta) = \sum_{i=1}^s e(\vartheta_i) \quad \text{and} \quad d_{\vartheta}(L) = \sum_{i=1}^s d_{\vartheta_i}(L)$$

for all $L \subset V(\mathcal{G})$.

We can now prove Lemma 6.2.

Proof of Lemma 6.2. Assume that $G \notin \mathcal{B}'(\mathbf{H})$ and let $V = V(G)$ and $N = |V|$. Let $c : E(G) \rightarrow [r]$ be an arbitrary r -colouring of the edges of G , and observe that, by Definition 6.1, for every $S \subset V$ with $|S| \geq \delta^{2/3}N$ there exists $j \in [r]$ such that $\mathcal{I}_{H_j, G_j, G}[S]$ is $(p, 2^{-9}r^{-1}\delta pN)$ -Janson. Thus, if for each $j \in [r]$, we define

$$\mathbf{V}_j = \left\{ S \subset V : |S| = \delta^{2/3}N \text{ and } \mathcal{I}_{H_j, G_j, G}[S] \text{ is } (p, 2^{-9}r^{-1}\delta pN)\text{-Janson} \right\}$$

then

$$\bigcup_{j=1}^r \mathbf{V}_j = \binom{V}{\delta^{2/3}N}$$

and hence there exists $i \in [r]$ such that

$$|\mathbf{V}_i| \geq \frac{1}{r} \binom{N}{\delta^{2/3}N}, \quad (89)$$

so fix such an i . For each $S \in \mathbf{V}_i$, as $\mathfrak{J}_{H_i, G_i, G}[S]$ is $(p, 2^{-9}r^{-1}\delta pN)$ -Janson by assumption, there exists $\nu_S : \mathfrak{J}_{H_i, G_i, G}[S] \rightarrow \mathbb{R}_{\geq 0}$ satisfying

$$e(\nu_S) = 1 \quad \text{and} \quad \Lambda_p(\nu_S) < \frac{2^9 r}{\delta p N} \quad (90)$$

by Observation 2.4. Now, define $\nu : \mathfrak{J}_{H_i, G_i, G} \rightarrow \mathbb{R}_{\geq 0}$ by

$$\nu = \sum_{S \in \mathbf{V}_i} \nu_S$$

and note that if

$$\Lambda_p(\nu) < \frac{e(\nu)^2}{pN}, \quad (91)$$

then $\mathfrak{J}_{H_i, G_i, G}$ is (p, pN) -Janson and therefore $G \notin \mathcal{B}(\mathbf{H})$ by definition, since c is an arbitrary r -colouring of the edges of G .

Recall that $V = V(\mathfrak{J}_{H_i, G_i, G})$ and also that, by definition,

$$\Lambda_p(\nu) = \sum_{\substack{L \subset V \\ |L| \geq 2}} d_\nu(L)^2 p^{-|L|} = \sum_{\substack{L \subset V \\ |L| \geq 2}} \left(\sum_{S \in \mathbf{V}_i} d_{\nu_S}(L) \right)^2 p^{-|L|}$$

where the last equality is due to Observation 6.3. Further observe that since ν_S is supported only on $\mathfrak{J}_{H_i, G_i, G}[S]$, then we can only have $d_{\nu_S}(L) > 0$ if $L \subset S$. Hence, denoting the subfamily

$$\mathbf{T}_L = \{S \in \mathbf{V}_i : L \subset S\}$$

for every $L \subset V$, we have

$$\Lambda_p(\nu) = \sum_{\substack{L \subset V \\ |L| \geq 2}} \left(\sum_{S \in \mathbf{T}_L} d_{\nu_S}(L) \right)^2 p^{-|L|} \leq \sum_{\substack{L \subset V \\ |L| \geq 2}} |\mathbf{T}_L| \sum_{S \in \mathbf{T}_L} d_{\nu_S}(L)^2 p^{-|L|}, \quad (92)$$

where the last step holds by the Cauchy–Schwarz inequality. Now, note that, as every L in the sums of (92) satisfies $|L| \geq 2$, we can bound $|\mathbf{T}_L|$ for these L by

$$|\mathbf{T}_L| \leq \binom{N-2}{\delta^{2/3}N-2} \leq \delta^{4/3} \binom{N}{\delta^{2/3}N} \leq r \delta^{4/3} |\mathbf{V}_i|,$$

where we used (89) in the last inequality. Substituting this into (92) and using (90) together with the definition of $\Lambda_p(\nu)$ yields

$$\Lambda_p(\nu) \leq r \delta^{4/3} |\mathbf{V}_i| \sum_{S \in \mathbf{V}_i} \Lambda_p(\nu_S) < r \delta^{4/3} |\mathbf{V}_i|^2 \frac{2^9 r}{\delta p N} = 2^9 r^2 \delta^{1/3} \frac{|\mathbf{V}_i|^2}{pN}. \quad (93)$$

Note that, since $e(\nu_S) = 1$ for all $S \in \mathbf{V}_i$ by (90), it follows from Observation 6.3 that

$$e(\nu) = |\mathbf{V}_i|,$$

which, replaced in (93), results in our goal, (91),

$$\Lambda_p(\nu) < 2^9 r^2 \delta^{1/3} \frac{e(\nu)^2}{pN} < \frac{e(\nu)^2}{pN}$$

where the last inequality is due to our choice of $\delta = r^{-50}$. Therefore, the hypergraph $\mathfrak{J}_{H_i, G_i, G}$ is (p, pN) -Janson, which implies that $G \notin \mathcal{B}(\mathbf{H})$ since the colouring $c : E(G) \rightarrow [r]$ was arbitrary. $\square$

6.2. Finding a set U to apply Lemma 5.1. Using Lemma 6.2, we can bound

$$\mathbb{P}(G \in \mathcal{B}(\mathbf{H}) \cap \mathcal{E}(\mathbf{s})) \leq \mathbb{P}(G \in \mathcal{B}'(\mathbf{H}) \cap \mathcal{E}(\mathbf{s})) \quad (94)$$

where $G \sim \mathbb{G}(N, 1/2)$, which leads us to the second stage in the proof of Lemma 2.8. In it, we will apply Lemma 5.1 to bound $\mathbb{P}(G \in \mathcal{B}'(\mathbf{H}) \cap \mathcal{E}(\mathbf{s}))$, but the setup of this application requires some work. First we show, with a deterministic argument, that for any graph $G \in \mathcal{B}'(\mathbf{H})$ with a “bad” colouring c , we can find a set U that satisfy the requirements of Lemma 5.1.

Before describing the concrete properties of U , we establish a correspondence between $G \in \mathcal{B}'(\mathbf{H})$, the sets S that appear in Definition 6.1, and these bad colourings c . To do that, it will be helpful to define the common setting for the rest of this section. Fix then $k \in \mathbb{N}$ and $s_1, \dots, s_r \in \mathbb{N}$ such that $s_i \leq k$ for each $i \in [r]$. Further fix graphs $H_1, \dots, H_r$ such that $v(H_i) \leq s_i$ for all $i \in [r]$, let $\mathbf{s} = (s_i)_{i \in [r]}$ and $\mathbf{H} = (H_i)_{i \in [r]}$, and fix $N \in \mathbb{N}$ satisfying (11).

Definition 6.4. For a graph G , define the collection $\mathbf{S}(G)$ by

$$\mathbf{S}(G) = \left\{ (S, c) : \begin{array}{l} S \subset V(G) \text{ with } |S| \geq \delta^{2/3}N \text{ and } c : E(G[S]) \rightarrow [r] \\ \text{such that } \forall i \in [r], \mathfrak{J}_{H_i, G_i, G}[S] \text{ is not } (p, 2^{-9}r^{-1}\delta pN)\text{-Janson} \end{array} \right\}.$$

Recall from Definition 6.1 that if $G \in \mathcal{B}'(\mathbf{H})$, then there exist $S \subset V(G)$ and $c : E(G[S]) \rightarrow [r]$ such that $(S, c) \in \mathbf{S}(G)$. We will next show that if $(S, c) \in \mathbf{S}(G)$, then for every large subset $W \subset S$, the hypergraph $\mathfrak{J}_{H_i^-, G_i, G}[W]$ is $(p, p|W|)$ -Janson. To prove that, we will require the event $\mathcal{E}(\mathbf{s})$, whose definition we recall for the reader’s convenience.

Definition 2.6. Given $r \in \mathbb{N}$ and $s_1, \dots, s_r \in \mathbb{N}$, let $\mathbf{s} = (s_i)_{i \in [r]}$ and let $\mathcal{E}(\mathbf{s})$ be the family of graphs G with the following property. For all graphs $F_1, \dots, F_r$ satisfying

$$\sum_{i=1}^r v(F_i) = \sum_{i=1}^r s_i - 1 \quad \text{and} \quad v(F_i) \leq s_i \quad \text{for each } i \in [r],$$

for every $W \subset V(G)$ with

$$|W| \geq \frac{\delta}{8r} v(G),$$

and every colouring $c : E(G[W]) \rightarrow [r]$, there is $i \in [r]$ such that $\mathfrak{J}_{F_i, G_i, G}[W]$ is $(p, p|W|)$ -Janson.

The following lemma is the only place in the proof of Lemma 2.8 where we will use the event $\mathcal{E}(\mathbf{s})$. Since later we will choose the set U to be a subset of S , the lemma immediately implies that $\mathfrak{J}_{H_i^-, G_i, G}[W]$ is $(p, p|W|)$ -Janson for every large subset $W \subset U$.

Lemma 6.5. *For all $G \in \mathcal{E}(\mathbf{s})$ and $(S, c) \in \mathbf{S}(G)$, the hypergraph $\mathfrak{J}_{H_i^-, G_i, G}[W]$ is $(p, p|W|)$ -Janson for every $i \in [r]$ and every $W \subset S$ with $|W| \geq \delta N / (8r)$.*

Proof. Fix an arbitrary $i \in [r]$, let $F_j = H_j$ for all $j \in [r] \setminus \{i\}$, and take $F_i = H_i^-$, a choice that satisfies

$$\sum_{j=1}^r v(F_j) = \sum_{j=1}^r s_j - 1. \quad (95)$$

By $G \in \mathcal{E}(\mathbf{s})$ and (95), for every $W \subset S \subset V(G)$ with $|W| \geq \delta N/(8r)$, there exists a colour $\ell \in [r]$ such that $\mathfrak{J}_{F_\ell, G_\ell, G}[W]$ is $(p, p|W|)$ -Janson. However, as

$$p|W| \geq 2^{-9}r^{-1}\delta pN,$$

we conclude that $\mathfrak{J}_{F_\ell, G_\ell, G}[W]$ is $(p, 2^{-9}r^{-1}\delta pN)$ -Janson. Since the Janson property is increasing by Observation 2.5, it follows that $\mathfrak{J}_{F_\ell, G_\ell, G}[S]$ is also $(p, 2^{-9}r^{-1}\delta pN)$ -Janson.

Now, recall that $\mathfrak{J}_{H_j, G_j, G}[S]$ is not $(p, 2^{-9}r^{-1}\delta pN)$ -Janson for all $j \in [r]$ since $(S, c) \in \mathbf{S}(G)$. Combining our choice of $F_j = H_j$ for all $j \neq i$ with the fact that $\mathfrak{J}_{F_\ell, G_\ell, G}[S]$ is $(p, 2^{-9}r^{-1}\delta pN)$ -Janson, the only remaining possibility is that $\ell = i$. It follows that, for every $W \subset S$ satisfying $|W| \geq \delta N/(8r)$, the hypergraph $\mathfrak{J}_{H_i, G_i, G}[W]$ is $(p, p|W|)$ -Janson. Since i was arbitrary, this completes the proof of the lemma. $\square$

The remaining properties that U will have are all guaranteed simultaneously by the way we construct it. For each $i \in [r]$, the hypergraph $\mathfrak{J}_{H_i, G_i, G}[U]$ will be (p, R_i) -Janson for some $R_i \geq 0$, and $\mathfrak{J}_{H_i, G_i, G}[U \cup \{v\}]$ will not be $(p, R_i + 1)$ -Janson for any $v \notin U$. The proof of Lemma 6.6 shows that we can find such a U with a routine induction.

Lemma 6.6. *Let G be a graph. If $(S, c) \in \mathbf{S}(G)$, then there exist $U \subset S$ and $R_1, \dots, R_r \in \mathbb{Z}_{\geq 0}$ such that*

$$|U| = \delta N + \sum_{i=1}^r R_i \quad \text{and} \quad R_1, \dots, R_r \leq \frac{p|U|}{2^{9r}} \quad (96)$$

which further satisfy, for all $i \in [r]$,

- (a) $\mathfrak{J}_{H_i, G_i, G}[U]$ is (p, R_i) -Janson, and
- (b) $\mathfrak{J}_{H_i, G_i, G}[U \cup \{v\}]$ is not $(p, R_i + 1)$ -Janson for all $v \in S \setminus U$.

Proof. Observe first that every set $U \subset S$ of size δN satisfies item (a) with $R_i = 0$ for all $i \in [r]$, since $\mathfrak{J}_{H_i, G_i, G}[U]$ is $(p, 0)$ -Janson by definition. Now take U of size $\delta N + \sum_{i=1}^r R_i$ maximising $\sum_{i=1}^r R_i$ among the choices that satisfy item (a). Note that $\mathfrak{J}_{H_i, G_i, G}[U \cup \{v\}]$ is (p, R_i) -Janson for every $v \in S$, since $\mathfrak{J}_{H_i, G_i, G}[U]$ is (p, R_i) -Janson, and by Observation 2.5, so item (b) holds by the maximality of $\sum_{i=1}^r R_i$. It therefore remains to prove the inequality in (96).

Suppose for a contradiction that $R_i > p|U|/(2^{9r})$ for some $i \in [r]$ and observe that

$$R_i > \frac{p|U|}{2^{9r}} \geq \frac{\delta pN}{2^{9r}},$$

so $\mathfrak{J}_{H_i, G_i, G}[U]$ is $(p, 2^{-9}r^{-1}\delta pN)$ -Janson by Observation 2.3. Since $U \subset S$, Observation 2.5 implies that $\mathfrak{J}_{H_i, G_i, G}[S]$ is also $(p, 2^{-9}r^{-1}\delta pN)$ -Janson, contradicting the fact that $(S, c) \in \mathbf{S}(G)$. $\square$

6.3. The proof. Having established that there is a U satisfying most of the requirements of Lemma 5.1, we will combine the following trivial consequence of the Chernoff bound with a pigeonhole argument to find many vertices v whose degree to U in some colour $\ell \in [r]$ is not small.

Observation 6.7. *Let $G \sim \mathbb{G}(N, 1/2)$, let $U \subset S \subset V(G)$ and set*

$$S' = \{v \in S \setminus U : d_G(v, U) > |U|/4\}.$$

If $2^{10} \leq |U| \leq |S|/4$, then

$$\mathbb{P}(|S'| \leq |S|/4) \leq \exp(-2^{-6}|U||S|).$$

Proof. For a vertex $v \in S \setminus U$, the Chernoff bound implies that

$$\mathbb{P}(v \notin S') = \mathbb{P}(d_G(v, U) \leq |U|/4) \leq \exp(-|U|/16),$$

and observe that these events are independent for distinct vertices of $S \setminus U$. If $|S'| \leq |S|/4$, then at least $|S|/2$ vertices of $S \setminus U$ fail to be in S' . Taking a union bound then yields

$$\mathbb{P}(|S'| \leq |S|/4) \leq 2^{|S|} \exp\left(-\frac{|S|}{2} \frac{|U|}{16}\right) \leq \exp(-2^{-6}|U||S|),$$

where last inequality uses that $|U| \geq 2^{10}$. $\square$

With all of the above, the proof of Lemma 2.8 proceeds by fixing the sets S, U and the colouring of $G[U]$, all of which we will take union bounds over. We will use Observation 6.7 to find many vertices whose degree to U in colour ℓ is not small, and we will then be able to check that the graphs $\tilde{G}' = G_\ell^{(c)}[U]$ and $\tilde{G} = G[U]$ satisfy all the conditions required by Lemma 5.1. To complete the proof, it will then suffice to observe that for each v the events bounded by Lemma 5.1 are independent, and hence the probabilities multiply.

Proof of Lemma 2.8. Recall first that we fixed $p = (2^{25}k^2r^4)^{-1}$ in (10). We claim that if $v(H_i) = 1$ for some $i \in [r]$, then trivially we have $\mathcal{B}(\mathbf{H}) = \emptyset$. Indeed, every non-empty 1-uniform hypergraph is (p, R) -Janson for all $p > 0$ and R , since $\Lambda_p(\nu) = 0$ regardless of the measure ν . We may therefore assume that $v(H_i) \geq 2$ for all $i \in [r]$.

By Lemma 6.2, we have $\mathcal{B}(\mathbf{H}) \subset \mathcal{B}'(\mathbf{H})$, and therefore

$$\mathbb{P}(G \in \mathcal{B}(\mathbf{H}) \cap \mathcal{E}(\mathbf{s})) \leq \mathbb{P}(G \in \mathcal{B}'(\mathbf{H}) \cap \mathcal{E}(\mathbf{s})),$$

where $G \sim \mathbb{G}(N, 1/2)$, here and in every probability statement in this proof. It will also be convenient to assume that every such graph shares the same vertex set $V(G) = [N]$.

Let $\mathbf{U}$ denote the collection of tuples $(S, U, (\tilde{G}_i)_{i \in [r]})$ with the following properties. The sets $U \subset S \subset [N]$ satisfy

$$|S| \geq \delta^{2/3}N \quad \text{and} \quad |U| = \delta N + \sum_{i=1}^r R_i, \quad \text{with} \quad 0 \leq R_1, \dots, R_r \leq \frac{p|U|}{2^{9r}}. \quad (97)$$

Moreover, letting

$$\tilde{G} = \bigcup_{i \in [r]} \tilde{G}_i, \quad (98)$$

each tuple $(S, U, (\tilde{G}_i)_{i \in [r]}) \in \mathbf{U}$ satisfies, for all $i \in [r]$, that $V(\tilde{G}_i) = U$,

(1) $\mathfrak{J}_{H_i^-, \tilde{G}_i, \tilde{G}}[W]$ is $(p, p|W|)$ -Janson for every $W \subset U$ with $|W| \geq |U|/(8r)$, and

(2) $\mathfrak{J}_{H_i, \tilde{G}_i, \tilde{G}}[U]$ is (p, R_i) -Janson.

This collection is important because of the following claim. Before stating it, define $f(G) = (S, c)$ to map each $G \in \mathcal{B}'(\mathbf{H}) \cap \mathcal{E}(\mathbf{s})$ to a fixed choice of $(S, c) \in \mathbf{S}(G)$.

Claim 6.8. *For each $G \in \mathcal{B}'(\mathbf{H}) \cap \mathcal{E}(\mathbf{s})$ and $(S, c) = f(G)$, there is $\sigma = (S, U, (\tilde{G}_i)_{i \in [r]}) \in \mathbf{U}$ for which $\tilde{G}_i = G_i^{(c)}[U]$ and the hypergraph $\mathfrak{J}_{H_i, \tilde{G}_i, \tilde{G}}[U \cup \{v\}]$ is not $(p, R_i + 1)$ -Janson for all $i \in [r]$ and all $v \in S \setminus U$.*

Note that the latter property of U and every $v \in S \setminus U$ in Claim 6.8 cannot be defined only in terms of $(\tilde{G}_i)_{i \in [r]}$, because it also depends on the colouring $c : E(G[S]) \rightarrow [r]$ in $(S, c) \in f(G)$.

Proof of Claim 6.8. Fix $G \in \mathcal{B}'(\mathbf{H}) \cap \mathcal{E}(\mathbf{s})$ and $(S, c) = f(G)$. Let $U \subset S$ be given by Lemma 6.6, which implies that it satisfies (97). Further set $\tilde{G}_i = G_i^{(c)}[U]$ for every $i \in [r]$, and recall that $\tilde{G} = \bigcup_{i \in [r]} \tilde{G}_i$ by (98). It follows from item (a) in Lemma 6.6 that this choice satisfies item (2) in the definition of $\mathbf{U}$ and also that $\mathfrak{J}_{H_i, G_i, G}[U \cup \{v\}]$ is not $(p, R_i + 1)$ -Janson for all $i \in [r]$ and all $v \in S \setminus U$, where the values of each R_i are given by Lemma 6.6.

To prove that this choice also satisfies item (1) in the definition of $\mathbf{U}$, observe first that Lemma 6.5 implies that $\mathfrak{J}_{H_i^-, G_i, G}[W]$ is $(p, p|W|)$ -Janson for every $i \in [r]$ and every $W \subset S$ with $|W| \geq \delta N/(8r)$. As $U \subset S$, this conclusion also holds whenever $W \subset U$ and $|W| \geq |U|/(8r)$, since $|U| \geq \delta N$ by (97). The final observation is that

$$\mathfrak{J}_{H_i^-, \tilde{G}_i, \tilde{G}}[W] = \mathfrak{J}_{H_i^-, G_i, G}[W]$$

by our choice of $\tilde{G}_i = G_i^{(c)}[U]$, so the previous reasoning indeed establishes item (1). $\blacksquare$

Now, for $\sigma = (S, U, (\tilde{G}_i)_{i \in [r]})$, let $\mathcal{A}(\sigma)$ be the collection of pairs of graphs G and colourings $c : E(G) \rightarrow [r]$ such that

- (a) $G_i[U] = G_i^{(c)}[U] = \tilde{G}_i$ for all $i \in [r]$, and
- (b) the hypergraph $\mathfrak{J}_{H_i, G_i, G}[U \cup \{v\}]$ is not $(p, R_i + 1)$ -Janson for all $i \in [r]$ and all $v \in S \setminus U$.

By Claim 6.8, we know that if $G \in \mathcal{B}'(\mathbf{H}) \cap \mathcal{E}(\mathbf{s})$, then there exists $\sigma \in \mathbf{U}$ and a colouring $c : E(G) \rightarrow [r]$ such that $(G, c) \in \mathcal{A}(\sigma)$. Taking a union bound over choices of $\sigma \in \mathbf{U}$, but, crucially, *not* over the choices of $c : E(G) \rightarrow [r]$, then yields

$$\mathbb{P}(G \in \mathcal{B}'(\mathbf{H}) \cap \mathcal{E}(\mathbf{s})) \leq \sum_{\sigma \in \mathbf{U}} \mathbb{P}(\exists c \in [r]^{E(G)} : (G, c) \in \mathcal{A}(\sigma)). \quad (99)$$

Most of the remainder of the proof will be dedicated to proving the following claim.

Claim 6.9. *For every $\sigma \in \mathbf{U}$,*

$$\mathbb{P}(\exists c \in [r]^{E(G)} : (G, c) \in \mathcal{A}(\sigma)) \leq 2^{-8r\delta^2 N^2}. \quad (100)$$

To prove Claim 6.9, we will modify $\mathcal{A}(\sigma)$ until the event(s) whose probability we need to bound become(s) $\{\exists G' \subset G : (G', G) \in \mathcal{M}_{v, \ell}\}$, where

$$\mathcal{M}_{v, \ell} = \left\{ (G', G) : \begin{array}{l} G'[U] = \tilde{G}_\ell, \ d_{G'}(v, U) \geq |U|/(4r) \text{ and} \\ \mathfrak{J}_{H_\ell, G', G}[U \cup \{v\}] \text{ is not } (p, R_\ell + 1)\text{-Janson} \end{array} \right\}. \quad (101)$$

We will then observe that not only the definition of $\mathcal{M}_{v, \ell}$ in (101) corresponds to the event whose probability Lemma 5.1 bounds if we take $F = H_\ell$, $R' = R_\ell$ and $\tilde{G}' = \tilde{G}_\ell$, but also that the current setting satisfies the assumptions to apply that lemma.

To make the proof easier to follow, we will first establish several intermediate claims towards Claim 6.9. Because of that, we will fix $\sigma = (S, U, (\tilde{G}_i)_{i \in [r]}) \in \mathbf{U}$ and abbreviate $\mathcal{A}(\sigma) = \mathcal{A}$ until the proof of Claim 6.9.

The first step towards proving (100) is to replace the set S with a large subset $S' \subset S$ by discarding vertices with small degree into U . By Observation 6.7, S contains such a subset with very high probability. More precisely, let

$$S' = S'(G) = \{v \in S \setminus U : d_G(v, U) > |U|/4\} \quad (102)$$

for every graph G , and define

$$\mathcal{A}' = \{(G, c) \in \mathcal{A} : |S'(G)| \geq |S|/4\}.$$

Claim 6.10.

$$\mathbb{P}(\exists c \in [r]^{E(G)} : (G, c) \in \mathcal{A}) \leq \mathbb{P}(\exists c \in [r]^{E(G)} : (G, c) \in \mathcal{A}') + \exp(-2^{-6}\delta^{5/3}N^2). \quad (103)$$

Proof. Recalling (97) in the definition of $\mathbf{U}$ and that $p \leq 1$, we have

$$|U| \leq \delta N + \frac{p|U|}{2^9} \leq \delta N + \frac{|U|}{2},$$

which, since we also have $|U| \geq \delta N$, implies that

$$\delta N \leq |U| \leq 2\delta N. \quad (104)$$

As $2\delta N \leq |S|/4$, we can apply Observation 6.7 to S and U , obtaining as a result

$$\mathbb{P}(|S'| \leq |S|/4) \leq \exp(-2^{-6}|U||S|) \leq \exp(-2^{-6}\delta^{5/3}N^2), \quad (105)$$

where the last inequality follows from $|S| \geq \delta^{2/3}N$ by (97) and $|U| \geq \delta N$ by (104). The claim now follows from splitting of $\mathcal{A}$ into $\mathcal{A}'$ and $\mathcal{A} \setminus \mathcal{A}'$ and bounding the latter using (105). $\blacksquare$

We now focus our attention on bounding the first term in the right-hand side of (103). Recalling that our goal is to apply Lemma 5.1 and that this lemma requires only a single subgraph, instead of a colouring, we will restrict our attention to the subgraph G_ℓ whose colour $\ell \in [r]$ is the majority colour of the neighbourhood of most vertices in S' .

Claim 6.11.

$$\mathbb{P}(\exists c \in [r]^{E(G)} : (G, c) \in \mathcal{A}') \leq \sum_{\ell \in [r]} \sum_{\substack{A \subset S \\ |A| \geq \frac{|S|}{4r}}} \mathbb{P}(\exists G' \subset G : (G', G) \in \bigcap_{v \in A} \mathcal{M}_{v, \ell} \text{ and } G[U] = \tilde{G}).$$

Proof. Let G be a graph, and suppose that there exists a colouring $c : E(G) \rightarrow [r]$ such that $(G, c) \in \mathcal{A}'$. We claim that $G[U] = \tilde{G}$, and that there exists a subgraph $G' \subset G$, a colour $\ell \in [r]$ and a subset $A \subset S$ with $|A| \geq |S|/(4r)$, such that $(G', G) \in \mathcal{M}_{v, \ell}$ for every $v \in A$. Claim 6.11 will then follow by taking a union bound over the choices of ℓ and A .

To show this, observe first that $G[U] = \tilde{G}$ holds because $(G, c) \in \mathcal{A}' \subset \mathcal{A}$, where $\tilde{G} = \bigcup_{i \in [r]} \tilde{G}_i$ was defined in (98). Next, note that for each $c : E(G) \rightarrow [r]$ and $v \in S'(G)$, the colouring c partitions the edges connecting v and U into r sets. As each $v \in S'(G)$ satisfies

$$d_G(v, U) > \frac{|U|}{4}$$

by the definition of $S'(G)$, (102), there exists a colour $j(v) = j \in [r]$ such that

$$d_{G_j}(v, U) > \frac{|U|}{4r}. \quad (106)$$

By the pigeonhole principle and (106), then, there exists a colour $\ell \in [r]$ such that, letting

$$A = A(G, c) = \{v \in S'(G) : j(v) = \ell\},$$

we have, as a consequence of $(G, c) \in \mathcal{A}'$, that

$$|A| \geq \frac{|S'(G)|}{r} \geq \frac{|S|}{4r}.$$

Letting $G' = G_\ell$ be the graph of edges spanned by colour ℓ completes the proof of the claim because $A \subset S \setminus U$ and $(G, c) \in \mathcal{A}$ satisfy item (b) in the definition of the event $\mathcal{A}$, and this is the last property in the definition of $\mathcal{M}_{v,\ell}$ that we had yet to show holds for this choice of G' . ■

The final claim that we need before the proof of Claim 6.9 is the observation that the events $\mathcal{M}_{v,\ell}$ are independent for each $v \notin U$ when conditioned on $\{G[U] = \tilde{G}\}$. This is a simple consequence of the conditioning causing $\mathcal{M}_{v,\ell}$ to depend only on the subgraph $G[v, U]$ corresponding to the edges between v and U . Although Claim 6.11 has $\{G[U] = \tilde{G}\}$ as an intersecting event, we can instead condition on it because its probability is non-zero.

Claim 6.12. *For any set $A \subset U$, we have*

$$\mathbb{P}\left(\exists G' \subset G : (G', G) \in \bigcap_{v \in A} \mathcal{M}_{v,\ell} \mid G[U] = \tilde{G}\right) = \prod_{v \in A} \mathbb{P}\left(\exists G' \subset G : (G', G) \in \mathcal{M}_{v,\ell} \mid G[U] = \tilde{G}\right).$$

Proof. First recall that $\mathcal{M}_{v,\ell}$ is defined in (101) as

$$\mathcal{M}_{v,\ell} = \left\{ (G', G) : \begin{array}{l} G'[U] = \tilde{G}_\ell, \ d_{G'}(v, U) \geq |U|/(4r) \text{ and} \\ \mathfrak{I}_{H_\ell, G', G}[U \cup \{v\}] \text{ is not } (p, R_\ell + 1)\text{-Janson} \end{array} \right\}.$$

Now, observe that after conditioning on $G[U]$, the existence of a subgraph G' satisfying the three properties in the definition of $\mathcal{M}_{v,\ell}$ depends only on the edges of $G[v, U]$. Since these edges are chosen independently for each vertex $v \in A$, the claim follows. ■

We can now combine the previous claims to prove Claim 6.9.

Proof of Claim 6.9. Fix $\sigma = (S, U, (\tilde{G}_i)_{i \in [r]}) \in \mathbf{U}$ and let $\mathcal{A} = \mathcal{A}(\sigma)$. By Claim 6.10, we have

$$\mathbb{P}(\exists c \in [r]^{E(G)} : (G, c) \in \mathcal{A}) \leq \mathbb{P}(\exists c \in [r]^{E(G)} : (G, c) \in \mathcal{A}') + \exp(-2^{-6}\delta^{5/3}N^2). \quad (107)$$

Claim 6.11 then implies that the first term in this right-hand side is at most

$$\mathbb{P}(\exists c \in [r]^{E(G)} : (G, c) \in \mathcal{A}') \leq \sum_{\ell \in [r]} \sum_{\substack{A \subset S \\ |A| \geq \frac{|S|}{4r}}} \mathbb{P}\left(\exists G' \subset G : (G', G) \in \bigcap_{v \in A} \mathcal{M}_{v,\ell} \mid G[U] = \tilde{G}\right), \quad (108)$$

where we moved the non-zero probability event $\{G[U] = \tilde{G}\}$ from the intersection to the conditioning, so each probability term inside the two sums of (108) satisfies

$$\mathbb{P}\left(\exists G' \subset G : (G', G) \in \bigcap_{v \in A} \mathcal{M}_{v,\ell} \mid G[U] = \tilde{G}\right) = \prod_{v \in A} \mathbb{P}\left(\exists G' : (G', G) \in \mathcal{M}_{v,\ell} \mid G[U] = \tilde{G}\right) \quad (109)$$

as a consequence of Claim 6.12.

We now want to apply Lemma 5.1 to obtain an upper bound for the probabilities on the right-hand side of (109), recalling that (101), the definition of $\mathcal{M}_{v,\ell}$, corresponds to the event whose probability we bound in (72) if we take $F = H_\ell$, $R' = R_\ell$ and $\tilde{G}' = \tilde{G}_\ell$. To check that the choice of parameters for this application is admissible, first observe that $s = v(H_\ell) - 1$ by assumption, that

$$m = |U| \geq \delta N \geq \delta r^{C(k+t)} \geq r^{Ck},$$

because $t \geq 1$, $C = 300$ and $\delta = r^{-50}$, and also that $v \notin U$ by $v \in A \subset S \setminus U$. It follows from $(S, U, (\tilde{G}_i)_{i \in [r]}) \in \mathbf{U}$ and the definition of $\mathbf{U}$ that not only

$$R_\ell \leq \frac{p|U|}{2^{9r}} = \frac{R}{16}$$

by (97) and our choice of $R = 2^{-5}r^{-1}p|U|$ in (71), but also that $\tilde{G}_\ell \subset \tilde{G}$ by (98). Finally, again by the properties of the tuples in $\mathbf{U}$, we have that

- (1) $\mathfrak{J}_{H_\ell, \tilde{G}_\ell, \tilde{G}}[U]$ is (p, R_ℓ) -Janson, and
- (2) $\mathfrak{J}_{H_\ell^-, \tilde{G}_\ell, \tilde{G}}[W]$ is (p, R) -Janson for every $W \subset U$ with $|W| \geq |U|/(8r)$, because

$$p|W| \geq \frac{p|U|}{8r} \geq 2^{-5}r^{-1}p|U| = R$$

and the Janson property is increasing (Observation 2.3),

so we can apply Lemma 5.1.

Applying Lemma 5.1, we then obtain, for fixed $\ell \in [r]$ and $v \in S \setminus U$, that

$$\mathbb{P}\left(\exists G' \subset G : (G', G) \in \mathcal{M}_{v, \ell} \mid G[U] = \tilde{G}\right) \leq 2^{-|U|/(2^5 r)},$$

which, replaced in (109), yields

$$\mathbb{P}\left(\exists G' \subset G : (G', G) \in \bigcap_{v \in A} \mathcal{M}_{v, \ell} \mid G[U] = \tilde{G}\right) \leq 2^{-|U||A|/(2^5 r)} \leq 2^{-\delta^{5/3}N^2/(2^7 r^2)} \quad (110)$$

by $|A| \geq |S|/(4r) \geq \delta^{2/3}N/(4r)$, where $A \subset S \setminus U$, and $|U| \geq \delta N$. Substituting (110) in (108) and bounding the number of choices for ℓ and A respectively by r and 2^N then yields

$$\mathbb{P}(\exists c \in [r]^{E(G)} : (G, c) \in \mathcal{A}') \leq r 2^N 2^{-\delta^{5/3}N^2/(2^7 r^2)}. \quad (111)$$

To deduce the claim, we combine (107) and (111) to obtain

$$\mathbb{P}(\exists c \in [r]^{E(G)} : (G, c) \in \mathcal{A}) \leq r 2^N 2^{-\delta^{5/3}N^2/(2^7 r^2)} + \exp(-2^{-6}\delta^{5/3}N^2) \leq 2^{-8r\delta^2N^2}$$

by $r \geq 2$ and our assumptions that $N \geq r^{C(k+t)}$ and $\delta = r^{-50}$. ■

We now apply Claim 6.9 in every term of (99) to obtain

$$\mathbb{P}(G \in \mathcal{B}'(\mathbf{H}) \cap \mathcal{E}(\mathbf{s})) \leq \sum_{\sigma \in \mathbf{U}} 2^{-8r\delta^2N^2} \quad (112)$$

To bound the right-hand side of (112), we count the number of tuples $\sigma = (S, U, (\tilde{G}_i)_{i \in [r]})$ in $\mathbf{U}$. There are at most 2^{2N} choices for both $S \subset V$ and $U \subset V$, and at most $2^{|U|^2}$ choices for each $\tilde{G}_i$. It follows from $|U| \leq 2\delta N$ in (104) that there are at most

$$2^{r|U|^2} \leq 2^{4r\delta^2N^2}$$

tuples $(\tilde{G}_i)_{i \in [r]}$, which replaced back in (112) yields

$$\mathbb{P}(G \in \mathcal{B}'(\mathbf{H}) \cap \mathcal{E}(\mathbf{s})) \leq 2^{2N+4r\delta^2N^2-8r\delta^2N^2} \leq 2^{-\delta^2N^2}$$

because we assumed that $N \geq r^{C(k+t)}$ and $\delta = r^{-50}$. □

7. CONTAINERS FOR NON-JANSON SETS

In this section, we prove our main technical result, and with it complete the proof of Theorem 1.2. The statement of Theorem 5.4 has three components that differ from Theorem 4.1: another hypergraph $\mathcal{F}$, which is (p, R') -Janson by assumption, a function π and a vertex v not in the set $U = V(\mathcal{F})$. Recall that when applying this theorem, we will take $\mathcal{F}$ to correspond to copies of F completely contained in U , and π to be the projection from $U \times \{0, 1\}$ onto U . To explain the role of v in the statement of Theorem 5.4, recall Definition 5.2,

$$\bar{\partial}_v \mathcal{G} = \{E \cup \{v\} : E \in \mathcal{G}\},$$

the edge-wise inclusion of v in a hypergraph $\mathcal{G}$.

Rather than obtaining containers for sets $L \subset V$ such that $\mathcal{H}[L]$ is not $(p/q, \eta R)$ -Janson, Theorem 5.4 provides containers for sets $L \subset V$ such that $\pi_v(\mathcal{H}[L]) \cup \mathcal{F}$ is not $(p, R' + \eta R)$ -Janson, where $\pi_v = \bar{\partial}_v \circ \pi$. Another difference between this theorem and Theorem 4.1 is in the properties of the containers $X \in \mathcal{X}$. Previously, we concluded that $\mathcal{H}[X]$ was not (p, R) -Janson, but we were not able to establish the same thing here due to vertices with high-degree. Instead, what we show is that whenever $X \subset V$ has linear size, we have a set $Y \subset X$ containing almost all elements of X such that $\pi(\mathcal{H}[Y])$ is not (p, R) -Janson.

Theorem 5.4. *Let $n, r, s \in \mathbb{N}$ with $n \geq s$ and $r \geq 2$, and let $q, p, R, R', \eta \in \mathbb{R}$ satisfy*

$$0 < q < \frac{1}{8}, \quad 0 < p \leq \frac{q}{2^{11}rs^2}, \quad R = 2^{-6}r^{-1}pn, \quad 0 \leq R' \leq \frac{R}{16} \quad \text{and} \quad \eta = p^4 \left(\frac{q}{2}\right)^{4s}. \quad (78)$$

Further let $\mathcal{F}$ be a $(s+1)$ -uniform hypergraph with vertex set U that is (p, R') -Janson, let $\mathcal{H}$ be an s -uniform hypergraph with vertex set V , where $|V| = n$, and let $\pi : V \rightarrow U$ satisfy

$$|\pi(L)| \geq \frac{|L|}{2} \quad \text{for every } L \subset V \quad \text{and} \quad |\pi(E)| = |E| \quad \text{for every } E \in \mathcal{H}. \quad (79)$$

Finally, let v be a vertex not in U . There exists a family $\mathcal{X} \subset 2^V$ with

$$|\mathcal{X}| \leq \left(\frac{2}{q}\right)^{2qn} \quad (80)$$

such that the following hold.

- (1) *If $I \subset V$ and $\pi_v(\mathcal{H}[I]) \cup \mathcal{F}$ is not $(p, R' + \eta R)$ -Janson, then $I \subset X$ for some $X \in \mathcal{X}$.*
- (2) *For each $X \in \mathcal{X}$ with $|X| \geq n/(8r)$, there exists $Y \subset X$ with*

$$|Y| \geq |X| - 2^{-8}r^{-1}n \quad (81)$$

such that $\pi(\mathcal{H}[Y])$ is not (p, R) -Janson.

We now highlight differences between the proof of Theorem 5.4 and the one in Section 4. The first one is already in the auxiliary hypergraph $\mathcal{J}'$, which is defined here by

$$\mathcal{J}' = \left\{ L \subset V : \pi_v(\mathcal{H}[L]) \cup \mathcal{F} \text{ is } (p, R' + \eta R)\text{-Janson} \right\}.$$

Note that, as the Janson property is increasing by Observation 2.5, each $L \subset V$ for which the hypergraph $\pi_v(\mathcal{H}[L]) \cup \mathcal{F}$ is not $(p, R' + \eta R)$ -Janson is also an independent set in $\mathcal{J}'$.

The start of the proof of Theorem 5.4 is analogous to the proof of Theorem 4.6: we apply Theorem 4.3 with $\mathcal{G} = \mathcal{J}'$, define $\mathcal{C}'_T = \langle \mathcal{C}_T \rangle_s$ for each $T \in \mathcal{T}$, and apply Theorem 3.4 with $\mathcal{G} = \mathcal{C}'_T$. This application of Theorem 3.4 also provides our candidate containers X , and the bulk of

the argument is proving that they fulfil the conclusions of the theorem, in particular that $\pi(\mathcal{H}[Y])$ is not (p, R) -Janson for some large $Y \subset X$.

Towards determining that $\mathcal{H}[X]$ was not (p, R) -Janson for fixed $X \in \mathcal{X}$, in the previous proof we took an arbitrary measure $\nu : \mathcal{H}[X] \rightarrow \mathbb{R}_{\geq 0}$ and showed that

$$\Lambda_p(\nu) \geq \frac{e(\nu)^2}{R}, \quad (113)$$

by considering two cases, depending on whether $e(\nu')$ was sufficiently large, where ν' was the restriction of ν to $\mathcal{C}'_T[X]$. Our goal in item (2) of Theorem 5.4 is to establish something like (113) for measures μ supported on $\pi(\mathcal{H}[Y])$ for some large $Y \subset X$. In order to do so, we will need to introduce some additional machinery, which will allow us to relate Janson properties of $\mathcal{H}$ and those of $\pi(\mathcal{H})$.

To reason about the Janson properties under the effect of π , we define the pullback of a measure ϑ with respect to π . The resulting measure, denoted by $\vartheta \hat{\circ} \pi$, distributes the mass of $E \in \pi(\mathcal{G})$ equally among edges that are entirely contained in its pre-image.

Definition 7.1. Let $\mathcal{G}$ be a hypergraph, U be a set, and let $\pi : V(\mathcal{G}) \rightarrow U$. If $\vartheta : \pi(\mathcal{G}) \rightarrow \mathbb{R}_{\geq 0}$ is a measure, then $\vartheta \hat{\circ} \pi : \mathcal{G} \rightarrow \mathbb{R}_{\geq 0}$ is defined by

$$\vartheta \hat{\circ} \pi(E) = \frac{\vartheta(\pi(E))}{|\{E_0 \in \mathcal{G} : \pi(E_0) = \pi(E)\}|} \quad (114)$$

for all $E \in \mathcal{G}$.

The crucial property of pullback measures is that, using them, we can show that for any hypergraph $\mathcal{G}$, if $\pi(\mathcal{G})$ is (p, R) -Janson, then so is $\mathcal{G}$.

Lemma 7.2. Let $R > 0$ and $p > 0$. Further let $\mathcal{G}$ be a hypergraph, U be a set, $\pi : V(\mathcal{G}) \rightarrow U$ be a function satisfying

$$|\pi(E)| = |E| \quad \text{for every } E \in \mathcal{G},$$

and $\vartheta : \pi(\mathcal{G}) \rightarrow \mathbb{R}_{\geq 0}$ be a measure. If

$$\Lambda_p(\vartheta) < \frac{e(\vartheta)^2}{R},$$

then

$$\Lambda_p(\vartheta \hat{\circ} \pi) < \frac{e(\vartheta \hat{\circ} \pi)^2}{R}.$$

In particular, if $\pi(\mathcal{G})$ is (p, R) -Janson, then so is $\mathcal{G}$.

As the proof of Lemma 7.2 is just checking that the definitions fit nicely together, we will postpone it, and the proofs of intermediate results that we require, to Appendix C. Nevertheless, we reference two of those results, Observation 7.3 and Lemma 7.4, because they will also be useful in the proof of Theorem 5.4.

Observation 7.3. Let $\mathcal{G}$ be a hypergraph and let $\pi : V(\mathcal{G}) \rightarrow U$ for some set U . Further let $\vartheta : \pi(\mathcal{G}) \rightarrow \mathbb{R}_{\geq 0}$ be a measure. If $E' \in \pi(\mathcal{G})$, then

$$\sum_{\substack{E \in \mathcal{G} \\ \pi(E) = E'}} \vartheta \hat{\circ} \pi(E) = \vartheta(E').$$

Observation 7.3 almost immediately implies the second useful result about pullback measures that we use in the proof of Theorem 5.4.

Lemma 7.4. *Let $\mathcal{G}$ be a hypergraph. For all $\pi : V(\mathcal{G}) \rightarrow U$ and $\vartheta : \pi(\mathcal{G}) \rightarrow \mathbb{R}_{\geq 0}$, we have*

$$e(\vartheta \hat{\circ} \pi) = e(\vartheta).$$

With the definition of the pullback measure and its crucial property, we can state the main inequalities in the proof that $\pi(\mathcal{H}[Y])$ is not (p, R) -Janson. We will start with a simplified view of the proof, and add details as we proceed. For technical reasons, it will be useful to assume instead the converse of our goal, i.e. we will fix one $X \in \mathcal{X}$ such that, for all large Y , the hypergraph $\pi(\mathcal{H}[Y])$ is (p, R) -Janson, and reach a contradiction. In particular, $\pi(\mathcal{H}[X])$ is (p, R) -Janson, so there is a measure $\mu : \pi(\mathcal{H}[X]) \rightarrow \mathbb{R}_{\geq 0}$ such that

$$\Lambda_p(\mu) < \frac{e(\mu)^2}{R} \tag{115}$$

and the pullback $\nu = \mu \hat{\circ} \pi$ satisfies

$$\Lambda_p(\nu) < \frac{e(\nu)^2}{R}$$

by Lemma 7.2.

Recall that a critical step in the proof of the simpler version of our container theorem was the inequality

$$\mathbb{E}[\Lambda_{p/q}(\nu_q'') \mid V_q \in \mathcal{I}(\partial_T \mathcal{J})] \geq \frac{\mathbb{E}[e(\nu_q'')^2 \mid V_q \in \mathcal{I}(\partial_T \mathcal{J})]}{\eta R}, \tag{116}$$

established in Claim 4.9 via the characterization of independent sets in $\mathcal{J}$. The analogous statement here depends on the definition of $\mathcal{J}'$, which implies that when $I \subset V(\mathcal{J}')$ is independent, then $\pi_v(\mathcal{H}[I]) \cup \mathcal{F}$ is not $(p, R' + \eta R)$ -Janson. To avoid too many technical details at once, let us assume that $R' > 0$, and that we have $\pi(\cdot)$ instead of $\pi_v(\cdot)$. These simplifications mean that $\pi(\mathcal{H}[I]) \cup \mathcal{F}$ is not $(p, R' + \eta R)$ -Janson when $I \in \mathcal{I}(\mathcal{J}')$.

In this simpler setup, we now use the assumption that $\mathcal{F}$ is (p, R') -Janson and $R' > 0$ to choose $\rho : \mathcal{F} \rightarrow \mathbb{R}_{\geq 0}$ with

$$\Lambda_p(\rho) < \frac{e(\rho)^2}{R'}. \tag{117}$$

A case analysis, the same as in Claim 4.7 in the proof of Theorem 4.6, will allow us to focus on $\mathcal{H}' = \mathcal{H}[X] \setminus \mathcal{C}'_T$, so we define $\mu_q : \pi(\mathcal{H}[X]) \rightarrow \mathbb{R}_{\geq 0}$ by

$$\mu_q(E) = \frac{\gamma \cdot \mathbf{1}[E \in \pi(\mathcal{H}'[V_q])]}{P'_q(E)} \mu(E),$$

where we will choose $\gamma = \sqrt{8\eta}$ and

$$P'_q(E) = \mathbb{P}(E \in \pi(\mathcal{H}'[V_q]) \mid V_q \in \mathcal{I}(\partial_T \mathcal{J}')).$$

The definition of $\mathcal{H}'$ will allow us to assume that

$$P'_q(E) > \left(\frac{q}{2}\right)^{|E|},$$

cf. (62), so μ_q is well-defined.

The inequality corresponding to (116) in this simplified overview of the proof will then be

$$\mathbb{E}[\Lambda_p(\rho + \mu_q) \mid V_q \in \mathcal{I}(\partial_T \mathcal{J}')] \geq \frac{\mathbb{E}[e(\rho + \mu_q)^2 \mid V_q \in \mathcal{I}(\partial_T \mathcal{J}')] }{R' + \eta R}, \quad (118)$$

which, note, has Λ_p instead of $\Lambda_{p/q}$ due to some still undiscussed numerics related to the fact that, in this case, the value of η is much smaller here than in Theorem 4.1. In the other direction, we would ideally like to show that

$$\mathbb{E}[\Lambda_p(\rho + \mu_q) \mid V_q \in \mathcal{I}(\partial_T \mathcal{J}')] \leq \Lambda_p(\rho) + 2\gamma\sqrt{\Lambda_p(\rho)\Lambda_p(\mu)} + \gamma^2\left(\frac{q}{2}\right)^{-2s}\Lambda_p(\mu), \quad (119)$$

and

$$\mathbb{E}[e(\rho + \mu_q)^2 \mid V_q \in \mathcal{I}(\partial_T \mathcal{J}')] \geq (e(\rho) + \gamma e(\mu))^2, \quad (120)$$

both of which, when combined with (115), (117) and our choice of γ , yield

$$\mathbb{E}[\Lambda_p(\rho + \mu_q) \mid V_q \in \mathcal{I}(\partial_T \mathcal{J}')] < \frac{\mathbb{E}[e(\rho + \mu_q)^2 \mid V_q \in \mathcal{I}(\partial_T \mathcal{J}')] }{R' + \eta R}$$

a direct contradiction of (118). Although we can establish inequalities resembling (119) and (120) with methods similar to those in Section 4, we have not justified some of our assumptions.

The first detail that we overlooked in the preceding overview is the assumption that $R' > 0$. In the case $R' = 0$, we can simply take $\rho = 0$. To handle both cases together, we use Observation 2.4 when $R' > 0$ and assume instead that

$$e(\rho) = \sqrt{R'} \quad \text{and} \quad \Lambda_p(\rho) < 1,$$

which is sufficient for our purposes.

The other omission in our discussion so far is that we simplified the definition of $\mathcal{J}'$. The correct definition means that, to satisfy something like (118), we require a measure supported on $\pi_v(\mathcal{H}[I]) \cup \mathcal{F}$, instead of $\pi(\mathcal{H}[I]) \cup \mathcal{F}$, where $I \in \mathcal{I}(\mathcal{J}')$. To address this change, we extend $\mu : \pi(\mathcal{H}[X]) \rightarrow \mathbb{R}_{\geq 0}$ to a measure $\bar{\mu} : \pi_v(\mathcal{H}[X]) \rightarrow \mathbb{R}_{\geq 0}$ by setting

$$\bar{\mu}(E \cup \{v\}) = \mu(E)$$

for all $E \in \pi(\mathcal{H})$, recalling that $v \notin U \supset \pi(V)$.

Adjusting for this seemingly innocuous change requires some technicalities. To prove the appropriate version of (118), with $\bar{\mu}_q$ replacing μ_q , we now need an upper bound for $\sum_{u \in \pi(X)} d_\mu(u)^2$. This is how we use Lemma 7.5 below, and is the main reason why we prove item (2) in Theorem 5.4 for a large $Y \subset X$ instead of all of $X \in \mathcal{X}$; it is also why it is simpler to prove that same item by contradiction. When we account for this term in the final version of (118), it dominates the term for $\Lambda_p(\nu)$, so a simpler bound for the latter suffices, and we do not need to consider $\Lambda_{p/q}$.

The proof of Lemma 7.5 is straightforward: we simply restrict the measure to avoid vertices with high degree. If we were only dealing with hypergraphs then implementing this idea would be a triviality, but working with measures involves a few tedious calculations, so we postpone the proof of this lemma to Appendix C.

Lemma 7.5. *Let $s \in \mathbb{N}$, $R, p, \beta > 0$, and $\mathcal{G}$ be an s -uniform hypergraph. If for every $W \subset V(\mathcal{G})$ with $|W| \geq (1 - \beta)v(\mathcal{G})$, we have that $\mathcal{G}[W]$ is (p, R) -Janson, then there exists $\mu : \mathcal{G} \rightarrow \mathbb{R}_{\geq 0}$ with*

$$e(\mu) = \sqrt{R}, \quad \Lambda_p(\mu) < \frac{e(\mu)^2}{R} \quad \text{and} \quad \sum_{v \in V(\mathcal{G})} d_\mu(v)^2 \leq \frac{2s^2 e(\mu)^2}{\beta v(\mathcal{G})}. \quad (121)$$

We are now ready to prove Theorem 5.4; for ease of reference, let us restate Theorem 4.3.

Theorem 4.3 (Campos and Samotij [7, modified Theorem E]). *Let $\mathcal{G}$ be a hypergraph with vertex set V . For all $\alpha, q \in \mathbb{R}$ satisfying $0 < q \leq \alpha < 1$, there exists a family $\mathcal{T} \subset 2^V$ and a function $\varphi : \mathcal{I}(\mathcal{G}) \rightarrow \mathcal{T}$ such that:*

- (a) *For each $I \in \mathcal{I}(\mathcal{G})$, we have $\varphi(I) \subset I$.*
- (b) *Each $T \in \mathcal{T}$ has at most qn/α elements, where $n = |V|$.*
- (c) *For every $T \in \mathcal{T}$, there exists a hypergraph $\mathcal{C}_T$ with vertex set V that covers $\mathcal{G}$ and satisfies*

$$\mathbb{P}(L \subset V_q \mid V_q \in \mathcal{I}(\partial_T \mathcal{G})) > (1 - \alpha)^{|L|q^{|L|}} \quad (44)$$

for all $L \notin \mathcal{C}_T$; moreover, for all $I \in \mathcal{I}(\mathcal{G})$ such that $T = \varphi(I)$, we have $I \in \mathcal{I}(\mathcal{C}_T)$.

Proof of Theorem 5.4. Apply Theorem 4.3 with $\mathcal{G} = \mathcal{J}'$, where

$$\mathcal{J}' = \left\{ L \subset V : \pi_v(\mathcal{H}[L]) \cup \mathcal{F} \text{ is } (p, R' + \eta R)\text{-Janson} \right\},$$

and parameters q and $\alpha = 1/2$ to obtain $\mathcal{T}$ and φ . Now, for each $T \in \mathcal{T}$, there is $\mathcal{C}_T$ satisfying item (c) in Theorem 4.3, so we let $\mathcal{C}'_T = \langle \mathcal{C}_T \rangle_s$ be the edges of $\langle \mathcal{C}_T \rangle$ of size s . Since $\mathcal{C}'_T$ is s -uniform and $p \leq 1/(2^{11}rs^2)$ by assumption, we can apply Theorem 3.4 with $\mathcal{G} = \mathcal{C}'_T$ and $\zeta = 2^{-8}r^{-1}$. As in the proof of Theorem 4.6 we obtain φ, ψ_T, ϕ_T that satisfy

- (i) For each $I \in \mathcal{I}(\mathcal{J}')$, we have $\phi_T(I) \subset I \subset \psi_T(\phi_T(I))$, for $T = \varphi(I)$.
- (ii) Each $S \in \mathcal{S}_T$ has at most $2^{11}rps^2n$ elements.
- (iii) For every $S \in \mathcal{S}_T$, letting $X = \psi_T(S)$, $\mathcal{C}'_T[X]$ is not $(p, 2^{-8}r^{-1}p|X|)$ -Janson.

Setting $f(I) = X = \psi_T(S)$ for $T = \varphi(I)$ and $S = \phi_T(I)$, we define

$$\mathcal{X} = \{f(I) : I \in \mathcal{I}(\mathcal{J}')\},$$

which, by item (i) and the fact that I was arbitrary, is a definition that satisfies item (1) in Theorem 5.4.

As in the proofs of Theorem 4.1 and Theorem 4.6 (cf. (54), (55) and (57)), to show that the size of $\mathcal{X}$ is suitable, we count $X \in \mathcal{X}$ by choosing $T \in \mathcal{T}$ and then $S \in \mathcal{S}_T$. We therefore obtain

$$|\mathcal{X}| \leq \sum_{T \in \mathcal{T}} |\mathcal{S}_T| \leq \sum_{m=0}^{2^{11}ps^2n} \binom{n}{m} \sum_{t=0}^{2qn} \binom{n}{t} \leq \left(\frac{2}{q}\right)^{2qn} \quad (122)$$

combining item (ii) above with item (b) in Theorem 4.3 and

$$2^{11}rps^2n \leq 2qn \leq \frac{n}{4}.$$

The bound in (122) matches (80) in the statement, so it only remains to show that item (2) holds.

Now, assume by contradiction that there exists $I \in \mathcal{I}(\mathcal{J}')$ such that

$$\begin{aligned} f(I) = X \in \mathcal{X} \text{ satisfies } |X| \geq n/(8r), \text{ and} \\ \pi(\mathcal{H}[Y]) \text{ is } (p, R)\text{-Janson for every } Y \subset X \text{ with } |Y| \geq |X| - 2^{-8}r^{-1}n. \end{aligned} \quad (*)$$

This is the converse of item (2) in Theorem 5.4, so contradicting it will complete the proof.

We want to apply Lemma 7.5 with $\mathcal{G} = \pi(\mathcal{H}[X])$ and $\beta = 2^{-9}r^{-1}$. To do that, we first verify that this hypergraph satisfies the assumptions in the lemma.

Claim 7.6. *For all*

$$W \subset \pi(X) \quad \text{with} \quad |W| \geq (1 - 2^{-9}r^{-1})|\pi(X)|, \quad (123)$$

the hypergraph $\pi(\mathcal{H}[X])[W]$ is (p, R) -Janson.

Proof. Fix W satisfying (123) and observe that letting $Y = \pi^{-1}(W) \cap X$, we have

$$\pi(\mathcal{H}[Y]) \subset \pi(\mathcal{H}[Y])[W] \subset \pi(\mathcal{H}[X])[W]$$

where the first containment holds because $\pi(E) \subset W$, for all $E \subset Y$, by our choice of $Y \subset \pi^{-1}(W)$ and the second holds since $Y \subset X$.

As $\pi(X \setminus Y) = \pi(X) \setminus W$ and we assumed that $|L| \leq 2|\pi(L)|$ for every $L \subset V$, we have that

$$|Y| = |X| - |X \setminus Y| \geq |X| - 2|\pi(X \setminus Y)| = |X| - 2|\pi(X) \setminus W| \geq |X| - 2^{-8}r^{-1}n. \quad (124)$$

It follows from (124) and (*) that $\pi(\mathcal{H}[Y])$ is (p, R) -Janson, but this is an increasing property by Observation 2.5, so $\pi(\mathcal{H}[X])[W]$ is also (p, R) -Janson. $\blacksquare$

Applying Lemma 7.5 with $\mathcal{G} = \pi(\mathcal{H}[X])$ and $\beta = 2^{-9}r^{-1}$, we obtain $\mu : \pi(\mathcal{H}[X]) \rightarrow \mathbb{R}_{\geq 0}$ satisfying

$$e(\mu) = \sqrt{R}, \quad \Lambda_p(\mu) < 1 \quad (125)$$

and

$$\sum_{u \in \pi(X)} d_\mu(u)^2 \leq \frac{2s^2 e(\mu)^2}{\beta |\pi(X)|} \leq \frac{2^{10} r s^2 e(\mu)^2}{|\pi(X)|} \leq \frac{2^{14} r^2 s^2 e(\mu)^2}{n}, \quad (126)$$

where the last inequality follows from the assumptions that $|L|/2 \leq |\pi(L)|$ for every $L \subset V$ and $|X| \geq n/(8r)$.

Let $\nu : \mathcal{H}[X] \rightarrow \mathbb{R}_{\geq 0}$ be the pullback measure of μ with respect to π , i.e. $\nu = \mu \hat{\circ} \pi$. As π satisfies

$$|\pi(E)| = |E| \quad \text{for all } E \in \mathcal{H}[X] \subset \mathcal{H}$$

by (79), we can apply Lemma 7.2 with $\mathcal{G} = \mathcal{H}[X]$ and $\vartheta = \mu$ to conclude that

$$\Lambda_p(\nu) < \frac{e(\nu)^2}{R}. \quad (127)$$

Now take $T = \varphi(I)$, and recall that $\mathcal{C}'_T = \langle \mathcal{C}_T \rangle_{=s}$, where $\mathcal{C}_T$ is the cover given by item (c) in Theorem 4.3. Let ν' be the restriction of the measure ν to $\mathcal{H}[X] \cap \mathcal{C}'_T[X]$, that is, for each $E \in \mathcal{H}[X]$, let

$$\nu'(E) = \begin{cases} \nu(E) & \text{if } E \in \mathcal{C}'_T[X], \\ 0 & \text{otherwise.} \end{cases}$$

Claim 7.7. *The measure ν' satisfies*

$$e(\nu') < \frac{e(\nu)}{2}.$$

Proof. Indeed, we have

$$\frac{e(\nu)^2}{R} > \Lambda_p(\nu) \geq \Lambda_p(\nu') \geq \frac{2^8 r e(\nu')^2}{p|X|} \geq \frac{4e(\nu')^2}{R}$$

first by (127), second because $\nu \geq \nu'$ and $\Lambda_p(\cdot)$ is monotone increasing, then since the hypergraph $\mathcal{C}'_T[X]$ is not $(p, 2^{-8}r^{-1}p|X|)$ -Janson by item (iii) of Theorem 3.4 and our choice of $\zeta = 2^{-8}r^{-1}$, and finally because $R = 2^{-6}r^{-1}pn$ by assumption. $\blacksquare$

We now define the measure $\nu'' = \nu - \nu'$, which corresponds to the restriction of ν to the hypergraph $\mathcal{H}' := \mathcal{H}[X] \setminus \mathcal{C}'_T = \mathcal{H}[X] \setminus \mathcal{C}'_T[X]$. By Claim 7.7, we have

$$e(\nu'') = e(\nu) - e(\nu') > \frac{e(\nu)}{2}. \quad (128)$$

Also define μ'' to be the restriction of μ to $\pi(\mathcal{H}')$. Applying Observation 7.3 with $\vartheta = \mu$ and $\nu = \mu \hat{\circ} \pi$, then using (128) and Lemma 7.4, yields

$$e(\mu'') = \sum_{E \in \pi(\mathcal{H}')} \mu(E) = \sum_{E \in \pi(\mathcal{H}')} \sum_{\substack{E_0 \in \mathcal{H} \\ \pi(E_0) = E}} \nu(E_0) \geq \sum_{E \in \mathcal{H}'} \nu''(E) = e(\nu'') > \frac{e(\nu)}{2} = \frac{e(\mu)}{2}. \quad (129)$$

With the goal of defining a random measure μ''_q , let, for all $E \in \pi(\mathcal{H}')$,

$$P'_q(E) = \mathbb{P}(E \in \pi(\mathcal{H}'[V_q]) \mid V_q \in \mathcal{I}(\partial_T \mathcal{J}')). \quad (130)$$

Claim 7.8. *For all $E \in \pi(\mathcal{H}')$,*

$$P'_q(E) \geq \mathbb{P}(E_0 \subset V_q \mid V_q \in \mathcal{I}(\partial_T \mathcal{J}')) > \left(\frac{q}{2}\right)^s \quad (131)$$

where $E_0 \in \mathcal{H}'$ is fixed and satisfies $\pi(E_0) = E$.

Proof. The first inequality in (131) follows from the fact that if $E_0 \subset V_q$ for $E_0 \in \mathcal{H}'$, then we have $\pi(E_0) \in \pi(\mathcal{H}'[V_q])$. Also notice that, by the same argument used to establish (62), the hypergraphs $\mathcal{H}'$ and $\mathcal{C}_T$ are disjoint. Therefore, we can use (44) in item (c) of Theorem 4.3 to obtain

$$\mathbb{P}(E_0 \subset V_q \mid V_q \in \mathcal{I}(\partial_T \mathcal{J}')) > \left(\frac{q}{2}\right)^s. \quad \blacksquare$$

Now, let $\gamma = \sqrt{8\eta} > 0$ (a choice made with foresight) and $\mu''_q : \pi(\mathcal{H}') \rightarrow \mathbb{R}_{\geq 0}$ be defined by

$$\mu''_q(E) = \mu''(E) \frac{\gamma \cdot \mathbf{1}[E \in \pi(\mathcal{H}'[V_q])]}{P'_q(E)}.$$

Also define the extension $\bar{\mu}''_q$ supported on $\pi_v(\mathcal{H}'[X_q])$, where $X_q = V_q \cap X$, by

$$\bar{\mu}''_q(E \cup \{v\}) = \mu''_q(E) \quad \text{for all } E \in \pi(\mathcal{H}'[X_q]).$$

Our goal now is to show that analysing $\bar{\mu}''_q$ for our choice of γ contradicts $I \in \mathcal{I}(\mathcal{J}')$.

To reason about properties of $\mathcal{J}'$, we define a measure ρ supported on $\mathcal{F}$. If $R' > 0$, then we can apply Observation 2.4 with $\mathcal{G} = \mathcal{F}$ and $y = \sqrt{R'}$, since $\mathcal{F}$ is (p, R') -Janson, to obtain $\rho : \mathcal{F} \rightarrow \mathbb{R}_{\geq 0}$ satisfying

$$\Lambda_p(\rho) < \frac{e(\rho)^2}{R'} \quad \text{and} \quad e(\rho) = \sqrt{R'}. \quad (132)$$

If, on the other hand, $R' = 0$, then we take the measure

$$\rho = 0. \quad (133)$$

Regardless of the value of R' , or the choice of ρ as either (132) or (133), we have

$$\Lambda_p(\rho) < 1 \quad \text{and} \quad e(\rho) = \sqrt{R'}, \quad (134)$$

which, together on being supported on $\mathcal{F}$, are the only properties that we will use of ρ .

Our goal is to obtain inequalities bounding $\Lambda_p(\rho + \bar{\mu}''_q)$ from $e(\rho + \bar{\mu}''_q)^2$, so we start easily, relating the expected value of $e(\bar{\mu}''_q)$ to $e(\mu'')$.

Claim 7.9.

$$\mathbb{E}[e(\bar{\mu}_q'') \mid V_q \in \mathcal{I}(\partial_T \mathcal{J})] = \gamma e(\mu'').$$

Proof. The definitions of $e(\bar{\mu}_q'')$, $\bar{\mu}_q''$ and μ_q'' ,

$$e(\bar{\mu}_q'') = \sum_{E \cup \{v\} \in \pi_v(\mathcal{H}')} \bar{\mu}_q''(E \cup \{v\}) = \sum_{E \in \pi(\mathcal{H}')} \mu''(E) \frac{\gamma \cdot \mathbb{1}[E \in \pi(\mathcal{H}'[V_q])]}{P'_q(E)},$$

imply that

$$\mathbb{E}[e(\bar{\mu}_q'') \mid V_q \in \mathcal{I}(\partial_T \mathcal{J}')] = \sum_{E \in \pi(\mathcal{H}')} \gamma \mu''(E) = \gamma e(\mu''),$$

since, for all $E \in \pi(\mathcal{H}')$, we have that

$$\mathbb{E}[\mathbb{1}[E \in \pi(\mathcal{H}'[V_q])] \mid V_q \in \mathcal{I}(\partial_T \mathcal{J}')] = P'_q(E) \quad (135)$$

from the definition of $P'_q(E)$, (130). ■

Combining Claim 7.9 with (129), we have

$$\mathbb{E}[e(\bar{\mu}_q'') \mid V_q \in \mathcal{I}(\partial_T \mathcal{J}')] > \gamma \frac{e(\mu)}{2}. \quad (136)$$

Recall that the definition of $d_\rho(L)$ for any set $L \subset U$ is

$$d_\rho(L) = \sum_{L \subset E \in \mathcal{F}} \rho(E).$$

As ρ is supported over $\mathcal{F}$, we have $\rho(E) = 0$ when $E \not\subset U = V(\mathcal{F})$. We can use this observation to obtain the expansion

$$\Lambda_p(\rho + \bar{\mu}_q'') = \sum_{\substack{L \subset U \\ |L| \geq 2}} d_\rho(L)^2 p^{-|L|} + 2 \sum_{\substack{L \subset U \\ |L| \geq 2}} d_\rho(L) d_{\bar{\mu}_q''}(L) p^{-|L|} + \sum_{\substack{L \subset U \cup \{v\} \\ |L| \geq 2}} d_{\bar{\mu}_q''}(L)^2 p^{-|L|} \quad (137)$$

where the first two terms range over $L \subset U$ instead of $L \subset U \cup \{v\}$ because if $v \in L$, then $d_\rho(L) = 0$ as $v \notin V(\mathcal{F})$ by assumption. The first sum in (137) is now exactly $\Lambda_p(\rho)$, which we can immediately bound with (134), but we need some simple claims before we analyse the other two sums. We start with an observation that relates the degrees $d_{\bar{\mu}_q''}$ to the degrees $d_{\mu_q''}$.

Claim 7.10. *For all $L \subset U \cup \{v\}$, we have*

$$d_{\bar{\mu}_q''}(L) = d_{\mu_q''}(L \setminus \{v\}).$$

Proof. Let $E_v = E \cup \{v\}$ for each $E \in \pi(\mathcal{H}')$. The definitions of $d_{\bar{\mu}_q''}$ and $\bar{\mu}_q''$ imply that, for every $L \subset U \cup \{v\}$,

$$d_{\bar{\mu}_q''}(L) = \sum_{L \subset E_v \in \pi_v(\mathcal{H}')} \bar{\mu}_q''(E_v) = \sum_{L \subset E_v \in \pi_v(\mathcal{H}')} \mu_q''(E_v \setminus \{v\}) = \sum_{L \setminus \{v\} \subset E \in \pi(\mathcal{H}')} \mu_q''(E) = d_{\mu_q''}(L \setminus \{v\}),$$

where the last equality is the definition of $d_{\mu_q''}(L \setminus \{v\})$. ■

We can now use Claim 7.10 to relate $\Lambda_p(\bar{\mu}_q'')$ and $\Lambda_p(\mu_q'')$.

Claim 7.11.

$$\Lambda_p(\bar{\mu}_q'') = \left(1 + \frac{1}{p}\right) \Lambda_p(\mu_q'') + \frac{1}{p^2} \sum_{u \in U} d_{\mu_q''}(u)^2. \quad (138)$$

Proof. First, recall the definition of $\Lambda_p(\bar{\mu}_q'')$,

$$\Lambda_p(\bar{\mu}_q'') = \sum_{\substack{L \subset U \cup \{v\} \\ |L| \geq 2}} d_{\bar{\mu}_q''}(L)^2 p^{-|L|}.$$

Splitting that sum according to whether $v \in L$ or not, we obtain

$$\Lambda_p(\bar{\mu}_q'') = \sum_{\substack{L \subset U \\ |L| \geq 2}} d_{\mu_q''}(L)^2 p^{-|L|} + \sum_{u \in U} d_{\mu_q''}(u)^2 p^{-2} + \sum_{\substack{L \subset U \\ |L| \geq 2}} d_{\bar{\mu}_q''}(L \cup \{v\})^2 p^{-|L|-1}, \quad (139)$$

where the second term corresponds to $L = \{u, v\}$ for $u \in U$, that is, the case $v \in L$ and $|L| = 2$. Applying Claim 7.10 in the third sum of (139) and using the definition of $\Lambda_p(\mu_q'')$ yields (138). ■

Now, we prove a deterministic upper bound for $\Lambda_p(\mu_q'')$ in terms of $\Lambda_p(\mu'')$. We do not optimise this bound, like the analogous one in Section 4 (cf. Claim 4.10), since that will not be necessary in this proof.

Claim 7.12.

$$\Lambda_p(\mu_q'') \leq \gamma^2 \left(\frac{q}{2} \right)^{-2s} \Lambda_p(\mu'').$$

Proof. Recall that

$$P'_q(E) > \left(\frac{q}{2} \right)^s$$

for all $E \in \pi(\mathcal{H}')$, by Claim 7.8, so by definition of $d_{\mu_q''}$ and μ_q'' , we deterministically have that

$$d_{\mu_q''}(L)^2 = \left(\sum_{L \subset E \in \pi(\mathcal{H}')} \mu''(E) \frac{\gamma \cdot \mathbf{1}[E \in \pi(\mathcal{H}'[V_q])]}{P'_q(E)} \right)^2 \leq \gamma^2 \left(\frac{q}{2} \right)^{-2s} d_{\mu''}(L)^2, \quad (140)$$

holds for all $L \subset U$, by ignoring the indicators. The inequality in the claim now follows from the definition of $\Lambda_p(\cdot)$:

$$\Lambda_p(\mu_q'') = \sum_{\substack{L \subset U \\ |L| \geq 2}} d_{\mu_q''}(L)^2 p^{-|L|} \leq \gamma^2 \left(\frac{q}{2} \right)^{-2s} \sum_{\substack{L \subset U \\ |L| \geq 2}} d_{\mu''}(L)^2 p^{-|L|} = \gamma^2 \left(\frac{q}{2} \right)^{-2s} \Lambda_p(\mu''). \quad \blacksquare$$

We will now combine the bound given by Lemma 7.5 for the sum of the square of the μ -degrees, (126), with Claims 7.11 and 7.12 and another simple calculation to complete the proof of a deterministic inequality relating $\Lambda_p(\bar{\mu}_q'')$ and $e(\mu)^2$.

Claim 7.13.

$$\Lambda_p(\bar{\mu}_q'') < \frac{\gamma^2}{2\sqrt{\eta}}.$$

Proof. Repeating what we did in (140), we have

$$\sum_{u \in U} d_{\mu_q''}(u)^2 \leq \gamma^2 \left(\frac{q}{2} \right)^{-2s} \sum_{u \in U} d_{\mu''}(u)^2 \leq \gamma^2 \left(\frac{q}{2} \right)^{-2s} \sum_{u \in U} d_{\mu}(u)^2 \quad (141)$$

where the last step is using that $\mu'' \leq \mu$. Now, recall that the support of μ is $\pi(\mathcal{H}[X])$, so if an edge $E \subset U$ is not fully contained in $\pi(X)$, then $\mu(E) = 0$. As so, $d_{\mu}(u) = 0$ if $u \notin \pi(X)$ and thus

$$\sum_{u \in U} d_{\mu}(u)^2 = \sum_{u \in \pi(X)} d_{\mu}(u)^2.$$

We conclude that (141) is at most

$$\sum_{u \in U} d_{\mu_q''}(u)^2 \leq \gamma^2 \left(\frac{q}{2}\right)^{-2s} \sum_{u \in U} d_\mu(u)^2 = \gamma^2 \left(\frac{q}{2}\right)^{-2s} \sum_{u \in \pi(X)} d_\mu(u)^2 \leq \gamma^2 \left(\frac{q}{2}\right)^{-2s} \frac{2^{14} r^2 s^2 e(\mu)^2}{n} \quad (142)$$

where the last step is

$$\sum_{u \in \pi(X)} d_\mu(u)^2 \leq \frac{2^{14} r^2 s^2 e(\mu)^2}{n},$$

the inequality in (126).

Combining (142) with Claims 7.11 and 7.12 thus yields

$$\Lambda_p(\bar{\mu}_q'') \leq \gamma^2 \left(\frac{2}{q}\right)^{2s} \left(\left(1 + \frac{1}{p}\right) \Lambda_p(\mu'') + \frac{2^{14} r^2 s^2 e(\mu)^2}{p^2 n} \right). \quad (143)$$

Also recall that we chose

$$p \leq \frac{q}{2^{11} r s^2}, \quad q < \frac{1}{8}, \quad R = 2^{-6} r^{-1} p n, \quad \text{and} \quad \eta = p^4 \left(\frac{q}{2}\right)^{4s},$$

in (78), and that

$$\Lambda_p(\mu'') \leq \Lambda_p(\mu) < \frac{e(\mu)^2}{R} \quad \text{and} \quad e(\mu) = \sqrt{R}$$

by $\mu'' \leq \mu$ and (125), so (143) is at most

$$\Lambda_p(\bar{\mu}_q'') < \gamma^2 \left(\frac{2}{q}\right)^{2s} \left(1 + \frac{1}{p} + \frac{2^8 r s^2}{p} \right) \leq \gamma^2 \left(\frac{2}{q}\right)^{2s} \frac{2^{10} r s^2}{p} \leq \frac{\gamma^2}{2\sqrt{\eta}}$$

as desired. ■

Replacing the bound of Claim 7.13 in (137) and then taking the conditional expectation with respect to $\{V_q \in \mathcal{I}(\partial_T \mathcal{J}')\}$ yields

$$\mathbb{E}[\Lambda_p(\rho + \bar{\mu}_q'') \mid V_q \in \mathcal{I}(\partial_T \mathcal{J}')] < 1 + 2 \sum_{\substack{L \subset U \\ |L| \geq 2}} d_\rho(L) \frac{\mathbb{E}[d_{\bar{\mu}_q''}(L) \mid V_q \in \mathcal{I}(\partial_T \mathcal{J}')] }{p^{|L|}} + \frac{\gamma^2}{2\sqrt{\eta}}. \quad (144)$$

In particular, this inequality motivates our final claim.

Claim 7.14. *If $L \subset U$, then*

$$\mathbb{E}[d_{\bar{\mu}_q''}(L) \mid V_q \in \mathcal{I}(\partial_T \mathcal{J}')] = \gamma d_{\mu''}(L).$$

Proof. Note that $v \notin L$ if $L \subset U$. It then follows from Claim 7.10 and the definition of $d_{\mu_q''}(L)$ that

$$d_{\bar{\mu}_q''}(L) = d_{\mu_q''}(L) = \sum_{L \subset E \in \pi(\mathcal{H}')} \mu''(E) \frac{\gamma \cdot \mathbb{1}[E \in \pi(\mathcal{H}'[V_q])]}{P'_q(E)},$$

so taking the conditional expectation yields, for all $L \subset U$,

$$\mathbb{E}[d_{\bar{\mu}_q''}(L) \mid V_q \in \mathcal{I}(\partial_T \mathcal{J}')] = \sum_{L \subset E \in \pi(\mathcal{H}')} \gamma \mu''(E) = \gamma d_{\mu''}(L)$$

since

$$\mathbb{E}[\mathbb{1}[E \in \pi(\mathcal{H}'[V_q])] \mid V_q \in \mathcal{I}(\partial_T \mathcal{J}')] = P'_q(E),$$

(135), holds for all $E \in \pi(\mathcal{H}')$. ■

Substituting the bound in Claim 7.14 in the middle sum of (144) and applying the Cauchy–Schwarz inequality, we obtain

$$\gamma \sum_{\substack{L \subset U \\ |L| \geq 2}} \frac{d_\rho(L) d_{\mu''}(L)}{p^{|L|}} \leq \gamma \left(\sum_{\substack{L \subset U \\ |L| \geq 2}} \frac{d_\rho(L)^2}{p^{|L|}} \right)^{1/2} \left(\sum_{\substack{L \subset U \\ |L| \geq 2}} \frac{d_{\mu''}(L)^2}{p^{|L|}} \right)^{1/2} = \gamma \sqrt{\Lambda_p(\rho)} \sqrt{\Lambda_p(\mu'')} \leq \gamma,$$

which, replaced back in (144), yields

$$\mathbb{E}[\Lambda_p(\rho + \bar{\mu}_q'') \mid V_q \in \mathcal{I}(\partial_T \mathcal{J}')] < 1 + 2\gamma + \frac{\gamma^2}{2\sqrt{\eta}}. \quad (145)$$

Now, to bound the expected edge mass of the measure $\rho + \bar{\mu}_q''$, observe that

$$\mathbb{E}[e(\rho + \bar{\mu}_q'') \mid V_q \in \mathcal{I}(\partial_T \mathcal{J}')] = e(\rho) + \mathbb{E}[e(\bar{\mu}_q'') \mid V_q \in \mathcal{I}(\partial_T \mathcal{J}')] \geq \sqrt{R'} + \frac{\gamma\sqrt{R}}{2},$$

where the last inequality is due to (134) and (136). Jensen's inequality thus implies that

$$\mathbb{E}[e(\rho + \bar{\mu}_q'')^2 \mid V_q \in \mathcal{I}(\partial_T \mathcal{J}')] \geq \left(\sqrt{R'} + \frac{\gamma\sqrt{R}}{2} \right)^2 > (1 + 4\gamma) \left(R' + \frac{\gamma^2 R}{8} \right) \quad (146)$$

because $R \geq 16R'$ by assumption, and we can choose $\gamma < 1/4$. We fix $\gamma = \sqrt{8\eta}$, which is less than $1/4$ because

$$\eta = p^4 \left(\frac{q}{2} \right)^{4s} < \frac{1}{2^7}.$$

This choice, replaced in (145), implies that

$$\mathbb{E}[\Lambda_p(\rho + \bar{\mu}_q'') \mid V_q \in \mathcal{I}(\partial_T \mathcal{J}')] < 1 + 4\gamma = \frac{(1 + 4\gamma)(R' + \gamma^2 R/8)}{R' + \eta R} \leq \frac{\mathbb{E}[e(\rho + \bar{\mu}_q'')^2 \mid V_q \in \mathcal{I}(\partial_T \mathcal{J}')] }{R' + \eta R}$$

and the last step is (146).

To reach a contradiction, observe that if $V_q \in \mathcal{I}(\partial_T \mathcal{J}')$ then $V_q \in \mathcal{I}(\mathcal{J}')$ follows from Observation 4.5, and hence $\pi_v(\mathcal{H}'[V_q]) \cup \mathcal{F}$ is not $(p, R' + \eta R)$ -Janson, by the definition of $\mathcal{J}'$. In particular, it follows that

$$\Lambda_p(\rho + \bar{\mu}_q'') \geq \frac{e(\rho + \bar{\mu}_q'')^2}{R' + \eta R}$$

since this measure is supported on $\pi_v(\mathcal{H}'[V_q]) \cup \mathcal{F}$, so we cannot have

$$\mathbb{E}[\Lambda_p(\rho + \bar{\mu}_q'') \mid V_q \in \mathcal{I}(\partial_T \mathcal{J}')] < \frac{\mathbb{E}[e(\rho + \bar{\mu}_q'')^2 \mid V_q \in \mathcal{I}(\partial_T \mathcal{J}')] }{R' + \eta R}$$

like we previously established, and we have a contradiction. We conclude that item (2) in Theorem 5.4 holds, and the proof is complete. $\square$

ACKNOWLEDGEMENTS

We would like to greatly thank Rob Morris for carefully reading the paper, and for the many improvements and corrections he suggested. We are also grateful to Wojciech Samotij and Julian Sahasrabudhe for helpful comments and discussions on the presentation and proof.

REFERENCES

- [1] H. L. Abbott. Lower bounds for some Ramsey numbers. *Discrete Math.*, 2(4):289–293, 1972.
- [2] P. Balister, B. Bollobás, M. Campos, S. Griffiths, E. Hurley, R. Morris, J. Sahasrabudhe, and M. Tiba. Upper bounds for multicolour Ramsey numbers. *arXiv:2410.17197*, 2024.
- [3] J. Balogh and W. Samotij. An efficient container lemma. *Discrete Anal.* Article 17, 56 pp., 2020.
- [4] J. Balogh, R. Morris, and W. Samotij. Independent sets in hypergraphs. *J. Amer. Math. Soc.*, 28(3):669–709, 2015.
- [5] J. Balogh, R. Morris, and W. Samotij. The method of hypergraph containers. In *Proceedings of the International Congress of Mathematicians—Rio de Janeiro 2018*, volume IV, pages 3059–3092. World Sci. Publ., Hackensack, NJ, 2018.
- [6] M. Bucić, T. Nguyen, A. Scott, and P. Seymour. Induced subgraph density. I. A loglog step towards Erdős–Hajnal. *Int. Math. Res. Not. IMRN*, (12):9991–10004, 2024.
- [7] M. Campos and W. Samotij. Towards an optimal hypergraph container lemma. *arXiv:2408.06617*, 2024.
- [8] M. Campos, S. Griffiths, R. Morris, and J. Sahasrabudhe. An exponential improvement for diagonal Ramsey. *Ann. Math.*, to appear.
- [9] D. Conlon. A new upper bound for diagonal Ramsey numbers. *Ann. Math.*, 170(2):941–960, 2009.
- [10] D. Conlon and A. Ferber. Lower bounds for multicolor Ramsey numbers. *Adv. Math.*, 378, 2021. Paper No. 107528, 5 pp.
- [11] D. Conlon, J. Fox, and B. Sudakov. On two problems in graph Ramsey theory. *Combinatorica*, 32(5):513–535, 2012.
- [12] D. Conlon, J. Fox, and B. Sudakov. Recent developments in graph Ramsey theory. In *Surveys in combinatorics 2015*, volume 424 of *London Math. Soc. Lecture Note Ser.*, pages 49–118. Cambridge Univ. Press, Cambridge, 2015.
- [13] D. Conlon, D. Dellamonica, S. La Fleur, V. Rödl, and M. Schacht. A note on induced Ramsey numbers. In *A Journey Through Discrete Mathematics*, pages 357–366. Springer, Cham, 2017.
- [14] W. Deuber. A generalization of Ramsey’s theorem. In *Infinite and Finite Sets*, volume 10 of *Colloq. Math. Soc. János Bolyai*, pages 323–332. North-Holland, Amsterdam-London, 1975.
- [15] P. Erdős. Some remarks on the theory of graphs. *Bull. Amer. Math. Soc.*, 53:292–294, 1947.
- [16] P. Erdős. Problems and results on finite and infinite graphs. In *Recent advances in graph theory (Proc. Second Czechoslovak Sympos., Prague, 1974)*, pages 183–192. (loose errata). Academia, Prague, 1975.
- [17] P. Erdős. On some problems in graph theory, combinatorial analysis and combinatorial number theory. In *Graph theory and combinatorics (Cambridge, 1983)*, pages 1–17. Academic Press, London, 1984.
- [18] P. Erdős and G. Szekeres. A combinatorial problem in geometry. *Compositio Math.*, 2:463–470, 1935.
- [19] P. Erdős, A. Hajnal, and L. Pósa. Strong embeddings of graphs into colored graphs. In *Infinite and finite sets*, volume 10 of *Colloq. Math. Soc. János Bolyai*, pages 585–595. North-Holland, Amsterdam-London, 1975.
- [20] J. Fox and B. Sudakov. Induced Ramsey-type theorems. *Adv. Math.*, 219(6):1771–1800, 2008.
- [21] J. Fox and B. Sudakov. Density theorems for bipartite graphs and related Ramsey-type results. *Combinatorica*, 29(2):153–196, 2009.
- [22] R. L. Graham, V. Rödl, and A. Ruciński. On bipartite graphs with linear Ramsey numbers. *Combinatorica*, 21(2):199–209, 2001.
- [23] Y. Kohayakawa, H. J. Prömel, and V. Rödl. Induced Ramsey numbers. *Combinatorica*, 18(3):373–404, 1998.
- [24] F. Mousset, R. Nenadov, and W. Samotij. Towards the Kohayakawa-Kreuter conjecture on asymmetric Ramsey properties. *Combin. Probab. Comput.*, 29(6):943–955, 2020.
- [25] R. Nenadov and A. Steger. A short proof of the random Ramsey theorem. *Combin. Probab. Comput.*, 25(1):130–144, 2016.
- [26] T. Nguyen, A. Scott, and P. Seymour. Induced subgraphs density. IV. New graphs with the Erdős–Hajnal property. *arXiv:2307.06455*, 2023.
- [27] T. Nguyen, A. Scott, and P. Seymour. Induced subgraph density. VII. The five-vertex path. *arXiv:2312.15333*, 2023.
- [28] F. P. Ramsey. On a problem of Formal Logic. *Proc. London Math. Soc.*, 30(4):264–286, 1929.

- [29] V. Rödl. The dimension of a graph and generalized Ramsey theorems. Master's thesis, Charles University, 1973.
- [30] A. Sah. Diagonal Ramsey via effective quasirandomness. *Duke Math. J.*, 172(3):545–567, 2023.
- [31] W. Samotij. Counting independent sets in graphs. *European J. Combin.*, 48:5–18, 2015.
- [32] W. Sawin. An improved lower bound for multicolor Ramsey numbers and a problem of Erdős. *J. Combin. Theory Ser. A*, 188, 2022. Paper No. 105579, 11 pp.
- [33] D. Saxton and A. Thomason. Hypergraph containers. *Invent. Math.*, 201(3):925–992, 2015.
- [34] J. Spencer. Ramsey's theorem—a new lower bound. *J. Combin. Theory Ser. A*, 18:108–115, 1975.
- [35] A. Thomason. An upper bound for some Ramsey numbers. *J. Graph Theory*, 12(4):509–517, 1988.
- [36] Y. Wigderson. An improved lower bound on multicolor Ramsey numbers. *Proc. Amer. Math. Soc.*, 149(6): 2371–2374, 2021.

APPENDIX A. PROOF OF THEOREM 4.3

This section is dedicated to the proof of Theorem 4.3, which we restate for the reader's convenience.

Theorem 4.3 (Campos and Samotij [7, modified Theorem E]). *Let $\mathcal{G}$ be a hypergraph with vertex set V . For all $\alpha, q \in \mathbb{R}$ satisfying $0 < q \leq \alpha < 1$, there exists a family $\mathcal{T} \subset 2^V$ and a function $\varphi : \mathcal{I}(\mathcal{G}) \rightarrow \mathcal{T}$ such that:*

- (a) *For each $I \in \mathcal{I}(\mathcal{G})$, we have $\varphi(I) \subset I$.*
- (b) *Each $T \in \mathcal{T}$ has at most qn/α elements, where $n = |V|$.*
- (c) *For every $T \in \mathcal{T}$, there exists a hypergraph $\mathcal{C}_T$ with vertex set V that covers $\mathcal{G}$ and satisfies*

$$\mathbb{P}(L \subset V_q \mid V_q \in \mathcal{I}(\partial_T \mathcal{G})) > (1 - \alpha)^{|L|q^{|L|}} \quad (44)$$

for all $L \notin \mathcal{C}_T$; moreover, for all $I \in \mathcal{I}(\mathcal{G})$ such that $T = \varphi(I)$, we have $I \in \mathcal{I}(\mathcal{C}_T)$.

The following proof is essentially a subset of the corresponding argument in [7], but the exact statement of Theorem 4.3 admits a simpler, self-contained proof that in particular does not assume the result of Campos and Samotij.

Proof. Given $I \in \mathcal{I}(\mathcal{G})$, let $T \subset I$ be maximal with respect to

$$\mathbb{P}(T \subset V_q \mid V_q \in \mathcal{I}(\mathcal{G})) \leq (1 - \alpha)^{|T|q^{|T|}}. \quad (147)$$

Set $\varphi(I) = T$ and $\mathcal{T} = \{\varphi(I) : I \in \mathcal{I}(\mathcal{G})\}$. We claim that these choices satisfy the requirements of Theorem 4.3. Notice that item (a) trivially holds, since $\varphi(I)$ is defined as a subset of I .

Now, take an arbitrary $T \in \mathcal{T}$. We have, on one hand,

$$(1 - q)^{n - |T|q^{|T|}} = \mathbb{P}(V_q = T) \leq \mathbb{P}(T \subset V_q \wedge V_q \in \mathcal{I}(\mathcal{G})), \quad (148)$$

where the inequality is using that $T \in \mathcal{I}(\mathcal{G})$, and, on the other hand,

$$\mathbb{P}(T \subset V_q \wedge V_q \in \mathcal{I}(\mathcal{G})) \leq \mathbb{P}(T \subset V_q \mid V_q \in \mathcal{I}(\mathcal{G})) \leq (1 - \alpha)^{|T|q^{|T|}} \quad (149)$$

due to (147). Combining (148) and (149), we obtain

$$(1 - q)^{n - |T|q^{|T|}} \leq (1 - \alpha)^{|T|q^{|T|}},$$

which, as $x \mapsto (1 - x)^{1/x}$ is decreasing on $(0, 1)$, implies $|T| \leq qn/\alpha$. But we took T arbitrarily, so the above establishes item (b).

To show that item (c) holds, take any $T \in \mathcal{T}$ and set

$$\mathcal{C}_T = \left\{ L \subset V(\mathcal{G}) : \mathbb{P}(L \subset V_q \mid V_q \in \mathcal{I}(\partial_T \mathcal{G})) \leq (1 - \alpha)^{|L|} q^{|L|} \right\}. \quad (150)$$

Observe that not only $\mathcal{C}_T$ is a cover of $\mathcal{G}$, but also $\mathcal{G} \subset \mathcal{C}_T$: for all $E \in \mathcal{G}$, we have

$$\mathbb{P}(E \subset V_q \mid V_q \in \mathcal{I}(\partial_T \mathcal{G})) = 0,$$

because $V_q \in \mathcal{I}(\partial_T \mathcal{G})$ and $E \subset V_q$ together imply $(E \setminus T) \in \mathcal{I}(\partial_T \mathcal{G})$, which directly contradicts the definition of $\partial_T \mathcal{G}$. The definition in (150) also immediately implies that, for all $L \notin \mathcal{C}_T$,

$$\mathbb{P}(L \subset V_q \mid V_q \in \mathcal{I}(\partial_T \mathcal{G})) > (1 - \alpha)^{|L|} q^{|L|}$$

holds, and that is exactly (44) in item (c). It remains only to establish the “moreover” part of item (c).

Our goal is to show that for all $I \in \mathcal{I}(\mathcal{G})$, if $\varphi(I) = T$, then $I \in \mathcal{I}(\mathcal{C}_T)$, so we take $L \in \mathcal{C}_T$ and must determine that $L \not\subset I$. Note that $L \not\subset T$ follows from $L \in \mathcal{C}_T$, as otherwise

$$\mathbb{P}(L \subset V_q \mid V_q \in \mathcal{I}(\partial_T \mathcal{G})) = \mathbb{P}(L \subset V_q) = q^{|L|} > (1 - \alpha)^{|L|} q^{|L|},$$

because $\{V_q \in \mathcal{I}(\partial_T \mathcal{G})\}$ and $\{L \subset V_q\}$ are independent when $L \subset T$.

We claim that

$$\mathbb{P}(L \cup T \subset V_q \mid V_q \in \mathcal{I}(\mathcal{G})) = \mathbb{P}(T \subset V_q \mid V_q \in \mathcal{I}(\mathcal{G})) \mathbb{P}(L \setminus T \subset V_q \mid V_q \in \mathcal{I}(\partial_T \mathcal{G})). \quad (151)$$

First, observe that we can reveal $T_q = T \cap V_q$ and then $V'_q = V_q \setminus T$, obtaining as a result

$$\mathbb{P}(L \cup T \subset V_q \mid V_q \in \mathcal{I}(\mathcal{G})) = \mathbb{P}(T \subset V_q \mid V_q \in \mathcal{I}(\mathcal{G})) \mathbb{P}(L \setminus T \subset V'_q \mid T \subset V_q \wedge V_q \in \mathcal{I}(\mathcal{G})).$$

But $\{T \subset V_q\} = \{T_q = T\}$ and $V_q \in \mathcal{I}(\mathcal{G})$ are together equivalent to $V'_q \in \mathcal{I}(\partial_T \mathcal{G})$, so we have

$$\mathbb{P}(L \setminus T \subset V_q \setminus T \mid T \subset V_q \wedge V_q \in \mathcal{I}(\mathcal{G})) = \mathbb{P}(L \setminus T \subset V'_q \mid V'_q \in \mathcal{I}(\partial_T \mathcal{G})),$$

and therefore

$$\mathbb{P}(L \cup T \subset V_q \mid V_q \in \mathcal{I}(\mathcal{G})) = \mathbb{P}(T \subset V_q \mid V_q \in \mathcal{I}(\mathcal{G})) \mathbb{P}(L \setminus T \subset V'_q \mid V'_q \in \mathcal{I}(\partial_T \mathcal{G})). \quad (152)$$

Finally, $L \setminus T$ and $V_q \setminus V'_q \subset T$ are disjoint and T_q is independent of $\{V'_q \in \mathcal{I}(\partial_T \mathcal{G})\}$, so

$$\mathbb{P}(L \setminus T \subset V'_q \mid V'_q \in \mathcal{I}(\partial_T \mathcal{G})) = \mathbb{P}(L \setminus T \subset V_q \mid V_q \in \mathcal{I}(\partial_T \mathcal{G})) \quad (153)$$

and substituting (153) in (152) yields (151).

Now, partitioning L according to its intersection with T , we have

$$\mathbb{P}(L \subset V_q \mid V_q \in \mathcal{I}(\partial_T \mathcal{G})) = \mathbb{P}(L \cap T \subset V_q) \mathbb{P}(L \setminus T \subset V_q \mid V_q \in \mathcal{I}(\partial_T \mathcal{G}))$$

because $\{L \cap T \subset V_q\}$ and $\{L \setminus T \subset V_q\}$ are independent, and so are $\{L \cap T \subset V_q\}$ and $\{V_q \in \mathcal{I}(\partial_T \mathcal{G})\}$. Therefore,

$$\mathbb{P}(L \setminus T \subset V_q \mid V_q \in \mathcal{I}(\partial_T \mathcal{G})) = \mathbb{P}(L \subset V_q \mid V_q \in \mathcal{I}(\partial_T \mathcal{G})) q^{-|L \cap T|}. \quad (154)$$

Combining (154) with (151), we obtain

$$\mathbb{P}(L \cup T \subset V_q \mid V_q \in \mathcal{I}(\mathcal{G})) = \mathbb{P}(L \subset V_q \mid V_q \in \mathcal{I}(\partial_T \mathcal{G})) \mathbb{P}(T \subset V_q \mid V_q \in \mathcal{I}(\mathcal{G})) q^{-|L \cap T|},$$

and now we can use that our choice of T satisfies (147), and L satisfies (150) to obtain

$$\mathbb{P}(L \cup T \subset V_q \mid V_q \in \mathcal{I}(\mathcal{G})) \leq (1 - \alpha)^{|T| + |L|} q^{|T| + |L| - |L \cap T|} \leq (1 - \alpha)^{|L \cup T|} q^{|L \cup T|} \quad (155)$$

because $0 < \alpha < 1$.

Taking $T' = L \cup T$, it is obvious that $T \subset T'$. We have picked $T \subset I$ to be maximal satisfying (147), so it follows from (155) that if $T' \subset I$, then we would have picked it instead of T . Therefore, $T' \not\subset I$ and thus $L \not\subset I$. As we took $L \in \mathcal{C}_T$ arbitrarily, we conclude that $I \in \mathcal{I}(\mathcal{C}_T)$ and so item (c) holds, completing the proof. $\square$

APPENDIX B. PROOF OF THEOREM 3.4

We recall the statement of Theorem 3.4 for the reader's convenience.

Theorem 3.4 (Campos and Samotij [7, modified Theorem A]). *Let $\mathcal{G}$ be an s -uniform hypergraph with n vertices. For all $0 < \zeta \leq 1$ and $0 < p \leq \zeta/(8s^2)$, there is a family $\mathcal{S} \subset 2^{V(\mathcal{G})}$ and functions*

$$\phi : \mathcal{I}(\mathcal{G}) \rightarrow \mathcal{S} \quad \text{and} \quad \psi : \mathcal{S} \rightarrow 2^{V(\mathcal{G})} \quad (27)$$

such that:

- (i) *For each $I \in \mathcal{I}(\mathcal{G})$, we have $\phi(I) \subset I \subset \psi(\phi(I))$.*
- (ii) *Each $S \in \mathcal{S}$ has at most $8s^2pn/\zeta$ elements.*
- (iii) *For every $S \in \mathcal{S}$, letting $X = \psi(S)$, $\mathcal{G}[X]$ is not $(p, \zeta p|X|)$ -Janson.*

To prove it, we need one of the container theorems of Campos and Samotij [7, Theorem A], which we state below as Theorem B.1. Recall from [7] that when $\mathcal{C}$ is a hypergraph and $0 < p < 1$, we denote by

$$w_p(\mathcal{C}) = \sum_{E \in \mathcal{C}} p^{|E|}$$

what is called the p -weight of $\mathcal{C}$.

Theorem B.1 (Campos and Samotij [7, Theorem A]). *Let $\mathcal{G}$ be an s -uniform hypergraph with n vertices. For every $0 < p' \leq 1/(8s^2)$, there exists a family $\mathcal{S} \subset 2^{V(\mathcal{G})}$ and functions*

$$\phi : \mathcal{I}(\mathcal{G}) \rightarrow \mathcal{S} \quad \text{and} \quad \psi : \mathcal{S} \rightarrow 2^{V(\mathcal{G})} \quad (156)$$

such that:

- (A) *For each $I \in \mathcal{I}(\mathcal{G})$, we have $\phi(I) \subset I \subset \psi(\phi(I))$.*
- (B) *Each $S \in \mathcal{S}$ has at most $8s^2p'n$ elements.*
- (C) *For every $S \in \mathcal{S}$, letting $X = \psi(S)$, there exists a hypergraph $\mathcal{C}$ on X with*

$$w_{p'}(\mathcal{C}) \leq p'|X| \quad (157)$$

that covers $\mathcal{G}[X]$ and satisfies $|E| \geq 2$ for all $E \in \mathcal{C}$.

Proof of Theorem 3.4. We apply Theorem B.1 to $\mathcal{G}$ with parameter $p' = p/\zeta \leq 1/(8s^2)$ and obtain ϕ , ψ and $\mathcal{S}$. Items (i) and (ii) in Theorem 3.4 are direct consequences of items (A) and (B) in Theorem B.1, so it remains only to show that item (C) implies item (iii). Fix $S \in \mathcal{S}$ and the corresponding $X = \psi(S)$, and let $\mathcal{C}$ be the cover of $\mathcal{G}[X]$ given by item (C) in Theorem B.1, and notice that by (157) we have

$$w_{p'}(\mathcal{C}) \leq p'|X| = p|X|/\zeta. \quad (158)$$

We can use (158) and combine the assumption $\zeta \leq 1$ with the fact that every edge in $\mathcal{C}$ has size at least 2 to bound the p -weight of $\mathcal{C}$ from its p' -weight:

$$w_p(\mathcal{C}) = \sum_{E \in \mathcal{C}} p^{|E|} = \sum_{E \in \mathcal{C}} (p/\zeta)^{|E|} \zeta^{|E|} \leq \zeta^2 \sum_{E \in \mathcal{C}} (p/\zeta)^{|E|} = \zeta^2 w_{p'}(\mathcal{C}) \leq \zeta p|X|. \quad (159)$$

To show that $\mathcal{G}[X]$ is not $(p, \zeta p|X|)$ -Janson, take any measure $\nu : \mathcal{G}[X] \rightarrow \mathbb{R}_{\geq 0}$, so our goal is to establish that

$$\Lambda_p(\nu) \geq \frac{e(\nu)^2}{\zeta p|X|}.$$

Observe first that

$$\Lambda_p(\nu) = \sum_{\substack{L \subset V(\mathcal{H}), \\ |L| \geq 2}} d_\nu(L)^2 p^{-|L|} \geq \sum_{L \in \mathcal{C}} d_\nu(L)^2 p^{-|L|}, \quad (160)$$

since every edge in $\mathcal{C}$ has size at least 2 by item (C) in Theorem B.1 and all the terms in the sum are non-negative. Massaging (160), we can apply the Cauchy–Schwartz inequality to obtain

$$\Lambda_p(\nu) \geq \frac{1}{w_p(\mathcal{C})} \left(\sum_{L \in \mathcal{C}} d_\nu(L)^2 p^{-|L|} \right) \left(\sum_{L \in \mathcal{C}} p^{|L|} \right) \geq \frac{1}{w_p(\mathcal{C})} \left(\sum_{L \in \mathcal{C}} d_\nu(L) \right)^2. \quad (161)$$

Now, note that, as $\mathcal{C}$ is a cover of $\mathcal{G}[X]$, we have

$$\sum_{L \in \mathcal{C}} d_\nu(L) = \sum_{L \in \mathcal{C}} \sum_{L \subset E} \nu(E) \geq \sum_{E \in \mathcal{G}[X]} \nu(E) = e(\nu). \quad (162)$$

Combining (159) and (162) with (161), we obtain

$$\Lambda_p(\nu) \geq \frac{e(\nu)^2}{\zeta p|X|},$$

which completes the proof because ν was arbitrary. $\square$

APPENDIX C. PROPERTIES OF MEASURES

C.1. The pullback measure. The main goal of this subsection is to prove Lemma 7.2, which we restate for convenience. Observe that Lemma 3.6 is a direct corollary of this statement.

Lemma 7.2. *Let $R > 0$ and $p > 0$. Further let $\mathcal{G}$ be a hypergraph, U be a set, $\pi : V(\mathcal{G}) \rightarrow U$ be a function satisfying*

$$|\pi(E)| = |E| \quad \text{for every } E \in \mathcal{G},$$

and $\vartheta : \pi(\mathcal{G}) \rightarrow \mathbb{R}_{\geq 0}$ be a measure. If

$$\Lambda_p(\vartheta) < \frac{e(\vartheta)^2}{R},$$

then

$$\Lambda_p(\vartheta \hat{\circ} \pi) < \frac{e(\vartheta \hat{\circ} \pi)^2}{R}.$$

In particular, if $\pi(\mathcal{G})$ is (p, R) -Janson, then so is $\mathcal{G}$.

The missing proof of Lemma 7.2 is a trivial combination of Lemma 7.4, which says that $e(\vartheta \hat{\circ} \pi) = e(\vartheta)$, and Lemma C.2, which establishes $\Lambda_p(\vartheta \hat{\circ} \pi) \leq \Lambda_p(\vartheta)$. Towards these two lemmas, we recall the statement of Observation 7.3 about pullback measures.

Observation 7.3. *Let $\mathcal{G}$ be a hypergraph and let $\pi : V(\mathcal{G}) \rightarrow U$ for some set U . Further let $\vartheta : \pi(\mathcal{G}) \rightarrow \mathbb{R}_{\geq 0}$ be a measure. If $E' \in \pi(\mathcal{G})$, then*

$$\sum_{\substack{E \in \mathcal{G} \\ \pi(E) = E'}} \vartheta \hat{\circ} \pi(E) = \vartheta(E').$$

Proof. Fix $E' \in \pi(\mathcal{G})$. We have by (114) that

$$\sum_{\substack{E \in \mathcal{G} \\ \pi(E) = E'}} \vartheta \hat{\circ} \pi(E) = \sum_{\substack{E \in \mathcal{G} \\ \pi(E) = E'}} \frac{\vartheta(E')}{|\{E \in \mathcal{G} : \pi(E) = E'\}|} = \vartheta(E')$$

□

We can now prove Lemma 7.4, an easy consequence of the definition of the pullback measure.

Lemma 7.4. *Let $\mathcal{G}$ be a hypergraph. For all $\pi : V(\mathcal{G}) \rightarrow U$ and $\vartheta : \pi(\mathcal{G}) \rightarrow \mathbb{R}_{\geq 0}$, we have*

$$e(\vartheta \hat{\circ} \pi) = e(\vartheta).$$

Proof. Expanding the definition of $e(\vartheta \hat{\circ} \pi)$ and using Observation 7.3, we obtain

$$e(\vartheta) = \sum_{E' \in \pi(\mathcal{G})} \vartheta(E') = \sum_{E' \in \pi(\mathcal{G})} \sum_{\substack{E \in \mathcal{G} \\ \pi(E) = E'}} \vartheta \hat{\circ} \pi(E).$$

But each $E \in \mathcal{G}$ appears on the right-hand side exactly once, only when $E' \in \pi(\mathcal{G})$ satisfies $\pi(E) = E'$. Therefore,

$$\sum_{E' \in \pi(\mathcal{G})} \sum_{\substack{E \in \mathcal{G} \\ \pi(E) = E'}} \vartheta \hat{\circ} \pi(E) = \sum_{E \in \mathcal{G}} \vartheta \hat{\circ} \pi(E) = e(\vartheta \hat{\circ} \pi)$$

as we wanted to show. □

The next lemma requires an assumption about π , motivating its appearance in the statements of Lemma 7.2 and Theorem 5.4. The proof is easy, and follows from expanding the definitions and a simple counting argument. It also requires defining the uniformity-preserving pre-image of $L' \subset \pi(V(\mathcal{G}))$, the set

$$\overleftarrow{\pi}(L') = \{L \subset V(\mathcal{G}) : \pi(L) = L' \text{ and } |L| = |L'|\}.$$

Lemma C.1. *Let $\mathcal{G}$ be a hypergraph, let $\pi : V(\mathcal{G}) \rightarrow U$ satisfy*

$$|\pi(E)| = |E| \quad \text{for every } E \in \mathcal{G},$$

and let $\vartheta : \pi(\mathcal{G}) \rightarrow \mathbb{R}_{\geq 0}$ be a measure. For all $L' \subset \pi(V(\mathcal{G}))$, if $\lambda = \vartheta \hat{\circ} \pi$, then

$$d_{\vartheta}(L') = \sum_{L \in \overleftarrow{\pi}(L')} d_{\lambda}(L). \quad (163)$$

Proof. Fix $L' \subset \pi(V(\mathcal{G}))$. The definition of $d_{\vartheta}(L')$, combined with Observation 7.3, yields

$$d_{\vartheta}(L') = \sum_{L' \subset E' \in \pi(\mathcal{G})} \vartheta(E') = \sum_{L' \subset E' \in \pi(\mathcal{G})} \sum_{\substack{E \in \mathcal{G} \\ \pi(E) = E'}} \lambda(E) = \sum_{\substack{E \in \mathcal{G} \\ L' \subset \pi(E)}} \lambda(E), \quad (164)$$

where the last step holds because each $E \in \mathcal{G}$ with $L' \subset \pi(E)$ appears exactly once in the second-to-last sum. On the other hand, the right-hand side of (163) is, by definition, equal to

$$\sum_{L \in \overleftarrow{\pi}(L')} d_{\lambda}(L) = \sum_{L \in \overleftarrow{\pi}(L')} \sum_{L \subset E \in \mathcal{G}} \lambda(E). \quad (165)$$

We also know that $|L| = |L'|$ for all $L \in \overset{\leftarrow}{\pi}(L')$. It then follows from $|\pi(E)| = |E|$ for all $E \in \mathcal{G}$ that $\pi|_E$ is a bijection, so there is a unique $L \subset E$ satisfying $\pi(L) = L'$ and $|\pi(L)| = |L|$. The conclusion is that

$$\sum_{L \in \overset{\leftarrow}{\pi}(L')} \sum_{L \subset E \in \mathcal{G}} \lambda(E) = \sum_{\substack{E \in \mathcal{G} \\ L' \subset \pi(E)}} \lambda(E),$$

which, together with (164), (165) and the fact that L' was arbitrary, completes the proof. $\square$

The proof of Lemma 7.2 will be complete once we establish Lemma C.2, which also admits a simple proof from Lemma C.1.

Lemma C.2. *Let $R > 0$ and $p > 0$. Further let $\mathcal{G}$ be a hypergraph, $\pi : V(\mathcal{G}) \rightarrow U$ be a function satisfying*

$$|\pi(E)| = |E| \quad \text{for every } E \in \mathcal{G}.$$

For all $\vartheta : \pi(\mathcal{G}) \rightarrow \mathbb{R}_{\geq 0}$, we have

$$\Lambda_p(\vartheta \hat{\circ} \pi) \leq \Lambda_p(\vartheta).$$

Proof. Let $V = V(\mathcal{G})$, $\lambda = \vartheta \hat{\circ} \pi$ and recall (8), the definition of Λ_p ,

$$\Lambda_p(\vartheta) = \sum_{\substack{L' \subset \pi(V) \\ |L'| \geq 2}} d_\vartheta(L')^2 p^{-|L'|}.$$

As $|\pi(E)| = |E|$ for every $E \in \mathcal{G}$, we can use Lemma C.1 to conclude that

$$\Lambda_p(\vartheta) = \sum_{\substack{L' \subset \pi(V) \\ |L'| \geq 2}} \left(\sum_{L \in \overset{\leftarrow}{\pi}(L')} d_\lambda(L) \right)^2 p^{-|L'|} \geq \sum_{\substack{L' \subset \pi(V) \\ |L'| \geq 2}} \sum_{L \in \overset{\leftarrow}{\pi}(L')} d_\lambda(L)^2 p^{-|L|} \quad (166)$$

where the last inequality uses that $d_\lambda(L)$ is always non-negative and that $|L| = |L'|$ for $L \in \overset{\leftarrow}{\pi}(L')$.

Now, whenever $d_\lambda(L) > 0$ for $L \subset V$, we know that there is $L' \subset \pi(V)$ such that $L \in \overset{\leftarrow}{\pi}(L')$. We conclude that the rightmost part of (166) is at least

$$\Lambda_p(\vartheta) \geq \sum_{\substack{L' \subset \pi(V) \\ |L'| \geq 2}} \sum_{L \in \overset{\leftarrow}{\pi}(L')} d_\lambda(L)^2 p^{-|L|} \geq \sum_{\substack{L \subset V \\ |L| \geq 2}} d_\lambda(L)^2 p^{-|L|} = \Lambda_p(\lambda)$$

where the last step is the definition, and the proof is complete. $\square$

C.2. General properties. Now, we prove Lemma 7.5, which, despite its length, is as simple as restricting the original measure to another that zeroes the mass of edges containing high degree vertices.

Lemma 7.5. *Let $s \in \mathbb{N}$, $R, p, \beta > 0$, and $\mathcal{G}$ be an s -uniform hypergraph. If for every $W \subset V(\mathcal{G})$ with $|W| \geq (1 - \beta)v(\mathcal{G})$, we have that $\mathcal{G}[W]$ is (p, R) -Janson, then there exists $\mu : \mathcal{G} \rightarrow \mathbb{R}_{\geq 0}$ with*

$$e(\mu) = \sqrt{R}, \quad \Lambda_p(\mu) < \frac{e(\mu)^2}{R} \quad \text{and} \quad \sum_{v \in V(\mathcal{G})} d_\mu(v)^2 \leq \frac{2s^2 e(\mu)^2}{\beta v(\mathcal{G})}. \quad (121)$$

Proof. Observe that taking $W = V(\mathcal{G})$, we conclude that $\mathcal{G} = \mathcal{G}[W]$ is (p, R) -Janson by assumption. Therefore, we can apply Observation 2.4 to obtain $\mu : \mathcal{G} \rightarrow \mathbb{R}_{\geq 0}$ such that

$$e(\mu) = \sqrt{R} \quad \text{and} \quad \Lambda_p(\mu) < \frac{e(\mu)^2}{R}. \quad (167)$$

Take such a μ minimizing $\sum_{v \in V(\mathcal{G})} d_\mu(v)^2$, and assume by contradiction that

$$\sum_{v \in V(\mathcal{G})} d_\mu(v)^2 > \frac{2s^2 e(\mu)^2}{\beta v(\mathcal{G})}. \quad (168)$$

Now, take

$$W = \{v \in V(\mathcal{G}) : d_\mu(v) \leq s e(\mu) / (\beta v(\mathcal{G}))\} \quad (169)$$

observe that it satisfies $|W| \geq (1 - \beta)v(\mathcal{G})$ since $\mathcal{G}$ is s -uniform, and therefore $\mathcal{G}[W]$ is (p, R) -Janson by assumption. We conclude that there is a measure $\mu' : \mathcal{G}[W] \rightarrow \mathbb{R}_{\geq 0}$ which satisfies

$$\Lambda_p(\mu') < \frac{e(\mu')^2}{R} \quad \text{and} \quad e(\mu') = \sqrt{R}, \quad (170)$$

again by Observation 2.4 applied with $y = \sqrt{R} > 0$. Moreover, we claim that setting

$$\mu'' = (1 - \tau)\mu + \tau\mu',$$

for a suitable $0 < \tau \leq 1$, also yields a measure satisfying

$$\Lambda_p(\mu'') < \frac{e(\mu'')^2}{R}, \quad e(\mu'') = \sqrt{R} \quad (171)$$

and

$$\sum_{v \in V(\mathcal{G})} d_{\mu''}(v)^2 < \sum_{v \in V(\mathcal{G})} d_\mu(v)^2. \quad (172)$$

If μ'' satisfies both (171) and (172), it contradicts the minimality of our original choice of μ .

We start the proof of our claims by showing that (171) holds. The equality $e(\mu'') = \sqrt{R}$ follows by linearity of $e(\cdot)$ and the fact that both μ and μ' have edge measure equal to $\sqrt{R}$. Expand $\Lambda_p(\mu'')$ as

$$\Lambda_p(\mu'') = (1 - \tau)^2 \Lambda_p(\mu) + 2\tau(1 - \tau) \sum_{\substack{L \subset V(\mathcal{G}) \\ |L| \geq 2}} d_\mu(L) d_{\mu'}(L) p^{-|L|} + \tau^2 \Lambda_p(\mu'). \quad (173)$$

Applying the Cauchy–Schwarz inequality to the second term in (173), we obtain

$$\sum_{\substack{L \subset V(\mathcal{G}) \\ |L| \geq 2}} \frac{d_\mu(L) d_{\mu'}(L)}{p^{|L|}} \leq \left(\sum_{\substack{L \subset V(\mathcal{G}) \\ |L| \geq 2}} \frac{d_\mu(L)^2}{p^{|L|}} \right)^{1/2} \left(\sum_{\substack{L \subset V(\mathcal{G}) \\ |L| \geq 2}} \frac{d_{\mu'}(L)^2}{p^{|L|}} \right)^{1/2} = \sqrt{\Lambda_p(\mu) \Lambda_p(\mu')}. \quad (174)$$

Replacing (174) back in (173) and simplifying yields

$$\Lambda_p(\mu'') \leq (1 - \tau)^2 \Lambda_p(\mu) + 2\tau(1 - \tau) \sqrt{\Lambda_p(\mu) \Lambda_p(\mu')} + \tau^2 \Lambda_p(\mu') = ((1 - \tau) \Lambda_p(\mu)^{1/2} + \tau \Lambda_p(\mu')^{1/2})^2,$$

which, by (167) and (170), establishes (171):

$$\Lambda_p(\mu'') < \frac{1}{R} ((1 - \tau)e(\mu) + \tau e(\mu'))^2 = \frac{e(\mu'')^2}{R}.$$

Towards establishing (172), we expand

$$\sum_{v \in V(\mathcal{G})} d_{\mu''}(v)^2 = (1 - \tau)^2 \sum_{v \in V(\mathcal{G})} d_{\mu}(v)^2 + 2\tau(1 - \tau) \sum_{v \in V(\mathcal{G})} d_{\mu}(v)d_{\mu'}(v) + \tau^2 \sum_{v \in V(\mathcal{G})} d_{\mu'}(v)^2, \quad (175)$$

an expression whose terms we will bound separately. Recall that $d_{\mu'}(v) = 0$ for $v \notin W$, so

$$\sum_{v \in V(\mathcal{G})} d_{\mu}(v)d_{\mu'}(v) = \sum_{v \in W} d_{\mu}(v)d_{\mu'}(v)$$

but now, (169) implies that

$$\sum_{v \in W} d_{\mu}(v)d_{\mu'}(v) \leq \frac{s e(\mu)}{\beta v(\mathcal{G})} \sum_{v \in W} d_{\mu'}(v) = \frac{s^2 e(\mu)^2}{\beta v(\mathcal{G})} \quad (176)$$

since $\mathcal{G}$ is s -uniform and $e(\mu') = e(\mu)$. Using (168) in (176), we obtain, for the second term,

$$\sum_{v \in W} d_{\mu}(v)d_{\mu'}(v) < \frac{1}{2} \sum_{v \in V(\mathcal{G})} d_{\mu}(v)^2. \quad (177)$$

The s -uniformity of $\mathcal{G}$ and $e(\mu) = e(\mu')$ also imply an easy bound on the third term in (175):

$$\sum_{v \in V(\mathcal{G})} d_{\mu'}(v)^2 \leq \left(\sum_{v \in V(\mathcal{G})} d_{\mu'}(v) \right)^2 = (e(\mu)s)^2 \leq \frac{\beta v(\mathcal{G})}{2} \sum_{v \in V(\mathcal{G})} d_{\mu}(v)^2 \leq \beta v(\mathcal{G}) \sum_{v \in V(\mathcal{G})} d_{\mu}(v)^2, \quad (178)$$

where the first inequality holds because d_{μ}' is always non-negative, and the second is (168).

We can now replace (177) and (178) in (175) and use that τ is positive to obtain, after simplification, that

$$\sum_{v \in V(\mathcal{G})} d_{\mu''}(v)^2 < (1 - \tau + \beta v(\mathcal{G})\tau^2) \sum_{v \in V(\mathcal{G})} d_{\mu}(v)^2.$$

Choosing τ to satisfy

$$0 < \tau < \min \left\{ 1, \frac{1}{\beta v(\mathcal{G})} \right\}$$

results in

$$\sum_{v \in V(\mathcal{G})} d_{\mu''}(v)^2 < \sum_{v \in V(\mathcal{G})} d_{\mu}(v)^2$$

which contradicts the fact that μ minimises $\sum_{v \in V(\mathcal{G})} d_{\mu}(v)^2$ and completes the proof. $\square$

UERJ, R. SÃO FRANCISCO XAVIER, 524 - MARACANÃ, RIO DE JANEIRO, BRASIL
Email address: lucas.aragao@uerj.br

IMPA, ESTRADA DONA CASTORINA 110, JARDIM BOTÂNICO, RIO DE JANEIRO, 22460-320, BRASIL
Email address: {marcelo.campos, gabriel.dahia, rafael.santos, joao.marciano}@impa.br