Extremal Problems in Graph Theory and a Hypergraph Packing Problem

Attila Dankovics

A thesis submitted for the degree of
Doctor of Philosophy

Department of Mathematics
The London School of Economics
and Political Science

22nd December 2021

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work, with the exceptions outlined below.

I declare that this thesis consists of 67742 words.

Statement of co-authored work

I confirm that Chapter 2 and Section 1.1 are based on [14], which is my own work.

I confirm that Chapter 3 and Section 1.2 are joint work with Fabio Botler, Jan Corsten, Nóra Frankl, Hiệp Hàn, Andrea Jiménez and Jozef Skokan.

I confirm that Chapter 4 and Section 1.3 are joint work with Peter Allen and Julia Böttcher.

Abstract

This thesis improves on the best result of an open problem, showing that Hamiltonian graphs with low independence number are pancyclic. It also describes graphs with the most triangle-free 5- and 6-colourings and generalises a packing theorem for degenerate graphs from graphs to hypergraphs.

A graph on n vertices is called pancyclic if it contains a cycle of every length $3 \leq \ell \leq n$. Given a Hamiltonian graph G with independence number at most k, we are looking for the minimum number of vertices $f(k)$ that guarantees that G is pancyclic. The problem of finding $f(k)$ was raised by Erdős in 1972, who showed that $f(k) \leq 4k^4$, and conjectured that $f(k) = \Theta(k^2)$. Improving on a result of Lee and Sudakov, we show that $f(k) = O(k^{11/5})$.

Let $k \geq 3$ and $r \geq 2$ be natural numbers. For a graph G, let $F(G, k, r)$ denote the number of colourings of the edges of G with colours $1, \dots, r$ such that, for every colour $c \in \{1, ..., r\}$, the edges of colour c contain no complete graph on k vertices K_k. Let $F(n, k, r)$ denote the maximum of $F(G, k, r)$ over all graphs G on n vertices. The problem of determining $F(n, k, r)$ was first proposed by Erdős and Rothschild in 1974, and has so far been solved only for $r = 2, 3$, and a small number of other cases. In this thesis we consider the question for the cases $k = 3$ and $r = 5$ or $r = 6$. We approximately determine the value $F(n, 3, 5)$ and $F(n, 3, 6)$ for large values of n. We also prove a stability result for both cases. This is joint work with F. Botler, J. Corsten, N. Frankl, H. Hàn, A. Jiménez and J. Skokan.

Given $D \geq 1$, whenever n is sufficiently large, if we are given any family of D-degenerate graphs of individual orders at most n, with maximum degree $c\frac{n}{\log(n)}$, and total number of edges at most $(1-\varepsilon)\binom{n}{2}$, then the family packs into the complete graph K_n, as proved by Allen, Böttcher, Hladký, and Piguet. If we add the condition that a linear fraction of the degenerate graphs have linearly many leaves, we can weaken the condition on the total number of edges to at most $\binom{n}{2}$ and still obtain a packing of the family into K_n, as proved by Allen, Böttcher, Clemens and Taraz. In this thesis we generalise both results to hypergraphs of any given uniformity. This is joint work with P. Allen and J. Böttcher.

Acknowledgements

First and foremost, I would like to thank my supervisors, Prof. Jozef Skokan, Prof. Julia Böttcher and Prof. Peter Allen for the recommendation of problems, their valuable advice and their continuous support during my PhD studies. I would like to thank them for motivating me during difficult times and for encouraging me to participate in workshops and conferences around the world.

I would like to thank LSE for providing me with a scholarship that made my PhD studies possible. I would like to thank Kate Barker, Enfale Farooq, Rebecca Lumb, Sarah Massey and Ed Perrin for their administrative support and for creating a professional environment I enjoyed working in.

I would like to thank all my fellow PhD students at LSE for being great to work with and all the fun we had together. I would especially like to thank my close friends from the department, Jan Corsten, Keat Eng Hng and Stanislav Kucera.

I would like to thank my other collaborators, Fabio Botler, Nóra Frankl, Hiệp Hàn and Andrea Jiménez. They have been a pleasure to work with, and I learned from them a lot.

I would like to thank Rita Alexiev and T Venczel for proofreading some of my work.

Finally, I would like to thank my parents and brother for supporting me throughout my life.

Contents

1 Introduction

Extremal problems in graph theory refer to problems of the following form. Given a family of graphs $\mathcal{G}$ and a graph parameter T, for which graph $G \in \mathcal{G}$ is $T(G)$ maximal?

We can also ask what the value of this maximal $T(G)$ is. More broadly, any statement of the form "For all $G \in \mathcal{G}$ we have $T(G) \leq c$" or "There exists $G \in \mathcal{G}$ such that $T(G) \geq c$" are statements belonging to the field of extremal graph theory.

One of the most well known classical extremal graph theory results is due to Mantel.

Theorem 1.0.1 (Mantel [42]). *Graphs with n vertices that contain no triangles have no more than $\left\lfloor \frac{n^2}{4} \right\rfloor$ edges.*

In this case, the family of graphs is "graphs with n vertices that contain no triangles" and the parameter is "number of edges". Mantel also proved that this result is sharp. Indeed the balanced complete bipartite graph on n vertices fulfills all conditions and has exactly $\left\lfloor \frac{n^2}{4} \right\rfloor$ edges.

In Chapter 2 and Chapter 3, we consider questions of this format. A graph is *Hamiltonian* if it contains a cycle going through all vertices. A graph on n vertices is called *pancyclic* if it contains a cycle of every length $3 \leq \ell \leq n$. It is widely studied what cycle lengths must show up in graphs. One graph property of interest is the sum of the reciprocals of cycle lengths in the graph, introduced by Erdős and Hajnal [20].

In Chapter 2, the family of graphs is "Hamiltonian but not pancyclic graphs on n vertices" and the parameter is independence number. This problem was also introduced by Erdős, and we improve on the best known result. In Chapter 3, the family is "graphs on n vertices" while the parameter is "number of triangle free colourings with 5 (or 6) colours".

In Section 1.1 and Section 1.2, we introduce these problems in more detail, provide some background on the history and relevance of the problems and state our main results on the topics.

Graph packing problems can be formulated the following way. Given a set of graphs $G_1, \ldots, G_k$ and a host graph H, can we embed the vertices of each of $G_1, \ldots, G_k$ into H such that edges are assigned to edges and each edge of H is used at most once. Many classical problems can be phrased as graph packing problems. For example an equivalent definition of Hamiltonicity is the following. A graph G on n vertices is Hamiltonian if the complement of G and the cycle C_n pack into the complete graph K_n. Especially useful are results that not only state that such an assignment is possible, but also provide an efficient algorithm.

A packing is *perfect* if each edge of the host graph is used exactly once. A classical packing problem is whether the complete graph can be perfectly packed with triangles. The first solution is by Kirkman [38], who proved that if the obvious divisibility requirements are met, this is always possible. As the question got popularised by Steiner [54], these packings are known as *Steiner triples*.

Hypergraphs are a generalisation of graphs. While in graphs each edge connects exactly two vertices, in hypergraphs an edge can contain any number of vertices. Packing questions can be more generally asked for hypergraphs. In Section 1.3, we pose such a problem, and in Chapter 4, we provide a simple algorithm with a complicated analysis that generates the required packing.

1.1 Low Independence Number and Hamiltonicity Implies Pancyclicity

A *Hamilton cycle* of a graph is a cycle that passes through all its vertices. A graph is *Hamiltonian* if it contains a Hamilton cycle as a subgraph. It is difficult to decide whether a graph contains a Hamilton cycle, therefore it is valuable to establish useful sufficient conditions for Hamiltonicity. The most well known sufficient condition is by Dirac [15].

Theorem 1.1.1 (Dirac [15])**.** *Let G be a graph on n vertices. If each vertex of G has at least $\frac{n}{2}$ neighbours, then G is Hamiltonian.*

A graph is *pancyclic* if it contains a cycle of every length $3 \leq \ell \leq n$, where n denotes the number of vertices. By definition pancyclicity implies Hamiltonicity.

Although the converse is not true, it is often the case that conditions that imply Hamiltonicity turn out to also imply pancyclicity. A famous meta-conjecture of Bondy [10] states that almost all non-trivial sufficient conditions of Hamiltonicity also imply pancyclicity with the possible exception of a few graphs. An example of this is the following theorem by Bondy [9].

Theorem 1.1.2 (Bondy [9])**.** *Let G be a graph on n vertices. If each vertex of G has at least $\frac{n}{2}$ neighbours, then G is either pancyclic or the complete balanced bipartite graph $K_{\frac{n}{2},\frac{n}{2}}$.*

The independence number of a graph G is the size of the largest stable set, denoted by $\alpha(G)$. A famous result of Chvátal and Erdős [12] states that if $\kappa(G) \geq \alpha(G)$, where $\kappa(G)$ is the vertex connectivity of G, then G is Hamiltonian. Keevash and Sudakov [35] showed that the similar but stronger condition $\kappa(G) \geq 600\alpha(G)$, is sufficient to conclude pancyclicity.

In this thesis we study a connection between Hamiltonicity, pancyclicity and independence number. Assuming G is a Hamiltonian graph with independence number at most k, we are looking for the minimum number of vertices $f(k)$ that guarantees that G is pancyclic. The problem of finding $f(k)$ was raised by Erdős, who showed that $f(k) \leq 4k^4$ and conjectured that a stronger statement holds.

Conjecture 1.1.3 (Erdős [19])**.** *There are constants c_1 and c_2 such that for all k we have $c_1k^2 \leq f(k) \leq c_2k^2$.*

The following construction due to Erdős provides a lower bound for this conjecture. Let $Q_1, \ldots, Q_k$ be cliques of size $k-2$ and add an edge between successive cliques (including between the last and the first) such that these edges are independent. It is easy to check that this graph is Hamiltonian, has independence number k and $k(k-2)$ vertices, but does not contain a cycle of length $k-1$.

The result of Erdős was later improved by Keevash and Sudakov [35], who showed that $f(k) \leq 150k^3$ holds and by Lee and Sudakov [40], who proved that $f(k) = O(k^{7/3})$ holds.

Here we improve their results.

Theorem 1.1.4. *There exists $c > 0$, such that if G is a Hamiltonian graph with $n \geq ck^{11/5}$ vertices and independence number at most k, then G is pancyclic. In other words $f(k) \leq c'k^{11/5}$.*

1.2 Maximum Number of Triangle-free Edge Colourings

A fundamental theorem of graph theory by Turán [55] asserts that among the graphs on n vertices that do not contain a complete graph on k vertices K_k, the complete balanced $(k-1)$-partite graph, also known as the Turán graph $T_{k-1}(n)$, has the largest number of edges $t_{k-1}(n)$. Clearly, no matter what subset of edges from a Turán graph we take, the resulting graph is also K_k-free. A natural question is: How many K_k-free graphs are there? In 1976 Erdős, Kleitman and Rothschild proved that the number of K_k-free graphs is asymptotically the same as the number of subgraphs of $T_{k-1}(n)$ [22].

Similarly, no matter how we colour the edges of the Turán graph, the edges of the same colour form a graph with no K_k in it. We call such a colouring K_k-free. Hence $T_{k-1}(n)$ has $r^{t_{k-1}(n)}$ K_k-free colourings with r colours. A natural question is whether we can find a graph with more K_k-free colourings, and if yes, what's the largest possible number of such colourings.

Let $k \geq 3$ and $r \geq 2$ be natural numbers. By a *colouring* of a graph $G = (V, E)$ with r colours here we mean *edge-colouring*, that is a function $f : E \to \{1, \dots, r\}$. In this context we refer to the numbers $1, \dots, r$ as colours and by the colour class of c we mean $f^{-1}(c)$. A colouring is K_k-free, if no colour class contains a copy of K_k. For a graph G, let $F(G, k, r)$ denote the number of K_k-free colourings of G with r colours. Let $F(n, k, r)$ denote the maximum of $F(G, k, r)$ over all graphs G on n vertices. Let $T_{k-1}(n)$ denote the Turán graph on n vertices with clique size $k-1$. Let $t_{k-1}(n)$ denote the number of edges of $T_{k-1}(n)$. Then the above lower bound obtained from the Turán graph can be restated as

$$F(n, k, r) \geq r^{t_{k-1}(n)}. \tag{1.2.1}$$

The problem of determining $F(n, k, r)$ was first proposed by Erdős and Rothschild in 1974 [17, 18]. They conjectured that in the case $k = 3$ and $r = 2$ the lower bound (1.2.1) is sharp for large enough n, and furthermore that $T_2(n)$ is the unique extremal graph. Their conjecture was proved by Yuster [56] who also proved an approximate version of the statement for general k and $r = 2$.

Improving on Yuster's results, Alon, Balogh, Keevash and Sudakov fully resolved the cases where $r = 2$ and $r = 3$ for large values of n.

Theorem 1.2.1 (Alon, Balogh, Keevash, Sudakov [4]). *For $k \geq 3$ and $n \geq n_0(k)$, we have $F(n,k,2) = 2^{t_{k-1}(n)}$ and $F(n,k,3) = 3^{t_{k-1}(n)}$. Moreover, the corresponding unique extremal graph is $T_{k-1}(n)$.*

In their paper [4] the authors also noted that the case $r > 3$ is more challenging as the behavior of $F(n,k,r)$ changes. Indeed, they proved that if $r > 3$ then $F(n,k,r)$ is exponentially larger than $r^{t_{k-1}(n)}$. To prove this, they provided the following construction: in T_{k-1} take two maximal independent sets, and replace them with complete bipartite graphs. They also determined the approximate values of $F(n,3,4)$ and $F(n,4,4)$. Subsequently, Pikhurko and Yilma [47] improved on their result, showing that $F(n,3,4) = F(T_4(n),3,4)$ and $F(n,4,4) = F(T_9(n),4,4)$ and the corresponding extremal graphs are unique.

More recently, Pikhurko, Staden and Yilma [46] proved that for every n,k,r there is a complete multipartite graph G such that $F(n,k,r) = F(G,k,r)$. This graph is not necessarily unique and not necessarily balanced. They also devised an optimisation problem with $R_r(k)$ variables, where $R_r(k)$ is the r-colour Ramsey-number for K_k. Very recently, Pikhurko and Staden [45] proved a stability result, stating all asymptotically optimal graphs are close to one of the solutions of their optimization problem. Their proofs use Szemerédi's regularity lemma and a symmetrisation method.

In this thesis we construct many examples of approximately extremal graphs for $k = 3$, $r = 5$ and every $n \in \mathbb{N}$ which are not Turán-graphs. However, all of them are complete multipartite and for each n there is a Turán-graph that is approximately extremal as well. Whether there is always at least one extremal Turán-graph and whether all extremal graphs are complete multipartite is still an open question.

Our approach and methods are different from that of Pikhurko, Staden and Yilma, but our results can be interpreted in a way to draw parallels between the results. The focus of Section 3.4 is to prove Theorem 3.4.3, which can be interpreted as an optimization problem with only 2^r variables ($r \in \{5,6\}$) and extra constraints. Our optimisation problem is different to the one proposed by Pikhurko et al. and contains less variables. Then in Section 3.6, we solve this optimization problem thus proving a stability result.

In the current thesis we consider the case $k = 3$ and $r \in \{5,6\}$. Let $\varphi_r(G) = F(G,3,r)$ and $\varphi_r(n) = F(n,3,r)$. We find the approximate value of $\varphi_5(n)$ and $\varphi_6(n)$. We also prove stability results for both cases.

We use the following simplified definition of *edit distance* to state our stability results.

Definition 1.2.2 (Edit distance)**.** Given two graphs $G_1 = (V, E_1)$ and $G_2 = (V, E_2)$ on the same vertex set, we define their edit distance as $d(G_1, G_2) = |E_1 \triangle E_2|$, where $\triangle$ denotes the symmetrical difference of sets.

This definition is equivalent to only allowing edge deletion and edge insertion of the usual edit operations, but it is sufficient for us.

Assumption 1.2.3. *For simplicity we will assume that $|V(G)|$ is divisible by 24 in the future. Thus when we define balanced graphs, we don't have to consider parts of different size. It will be easy to see that the error term introduced by this is negligible.*

We first state the technical stability result for $r = 6$ as that is the easier of the two.

Definition 1.2.4. Let V be a set of vertices. We define $G_6(V)$ as the set of balanced complete 8-partite graphs that has vertex set V.

Theorem 1.2.5 (Stability for $r = 6$)**.** *For every $0 < \delta < 10^{-30}$ there is an n_0 and $0 < \delta'$ such that for all graphs $G = (V, E)$ on $n > n_0$ vertices the following holds. Suppose $\varphi_6(G) \geq 3^{n^2/4}4^{3n^2/16-\delta n^2}$. Then there is a $G' \in G_6(V)$ such that $d(G, G') < \delta' n^2$.*

Calculating the number of K_3-free colourings for these graphs gives the following corollary.

Corollary 1.2.6. *For every n we have $\varphi_6(n) \leq 3^{n^2/4}4^{3n^2/16+o(n^2)}$.*

In our unpublished paper with Botler, Corsten, Frankl, Hàn, Jiménez and Skokan, we prove that the complete balanced 8-partite graph is the unique extremal graph for $r = 6$. This is not part of the present thesis.

The case of $r = 5$ is more complicated, as the structure of approximately optimal graphs is more varied.

Definition 1.2.7. Let V be a set of vertices with $|V| = n$. We define $G_5(V)$ as the set of complete 8-partite graphs that has vertex set V and has the following part sizes: $(n/4, n/4, a, a, b, b, n/4 - a - b, n/4 - a - b)$ for some $0 \leq a, b$ and $a + b \leq n/4$; or $(a, a, n/4 - a, n/4 - a, b, b, n/4 - b, n/4 - b)$ for some $0 \leq a, b \leq n/4$.

Note: In edge cases this might be a complete 4-partite or 6-partite graph instead of 8-partite, these graphs are part of $G_5(V)$ as well.

Theorem 1.2.8 (Stability for $r = 5$). *For every $0 < \delta < 10^{-30}$ there is an n_0 and $0 < \delta'$ such that for all graphs $G = (V, E)$ on $n > n_0$ vertices the following holds. Suppose $\varphi_6(G) \geq 6^{n^2/4-\delta n^2}$. Then there is a $G' \in G_5(V)$ such that $d(G, G') < \delta' n^2$.*

Calculating the number of K_3-free colourings for these graphs gives the following corollary.

Corollary 1.2.9. *For every n we have $\varphi_5(n) \leq 6^{n^2/4+o(n^2)}$.*

The fact that the aforementioned graphs all have asymptotically equal number of K_3-free colourings makes the case $r = 5$ particularly difficult and interesting. We are currently working on finding the exact solution for this case.

1.3 Packing Degenerate Hypergraphs

A *packing* of a family $\mathcal{G} = \{G_1, \ldots, G_k\}$ of hypergraphs into a hypergraph H is a colouring of the edges of H with the colours $0, 1, \ldots, k$ such that the edges of colour i form an isomorphic copy of G_i for each $1 \leq i \leq k$. The packing is *perfect* if no edges have colour 0. We will often say an edge is *covered* in a packing if it has colour at least 1, and *uncovered* if it has colour zero.

Packing problems have been studied for several decades. Classical theorems and conjectures of extremal graph theory can often be written as packing problems. For example, Turán's theorem can be read as the statement that if the n-vertex G does not have too many edges, then G and K_r pack into K_n. However packings in this context are usually very far from being perfect packings, with a large fraction of $E(H)$ uncovered. By contrast, in Chapter 4 we are interested in *almost perfect packings* and *perfect packings*, that is, packings in which $o\big(e(H)\big)$ edges are uncovered. One of the first problems asking for perfect packings in graphs is the problem of *Steiner-systems* and it is over a century old. Plücker [48] in 1835 found perfect packings of $\frac{1}{3}\binom{n}{r}$ copies of K_3 into K_n for various values of n, and more generally, Kirkman [38] in 1847 solved the problem for all values of n. Unaware of Kirkman's results, in 1853 Steiner re-asked the question [54] and popularised it. More generally, one can ask the following question.

Question 1.3.1. *Given $2 \leq k \leq r$, for which values of n does the complete k-uniform hypergraph $K_n^{(k)}$ have a perfect packing with copies of $K_r^{(k)}$?*

A packing of this form is called a *combinatorial design*. There are some simple divisibility conditions on n which are necessary for an affirmative answer. Recently and spectacularly, Keevash [32] proved that for sufficiently large n these conditions are also sufficient. He used a method called *randomised algebraic construction*, setting aside a structure with algebraic properties allowing him to absorb whatever is leftover after an almost perfect packing. Keevash reached an almost perfect packing by using *Rödl nibble*.

Rödl nibble was first introduced by Rödl in 1985 [50]. Given the packing problem in Question 1.3.1, the idea is to choose a small number of copies of $K_r^{(k)}$ in the host hypergraph, then throw out intersecting ones. Next, update our host graph and repeat until we have an almost perfect packing. This method is more difficult to use in problems where the graphs that we want to pack are larger, as they will almost always intersect if we don't take extra precautions. This is one of the reasons we do not use Rödl nibble in our present work.

Keevash's result was reproved independently, using a more combinatorial method, by Glock, Kühn, Lo and Osthus [26], who were also able to extend the result to pack with arbitrary fixed hypergraphs [26]. Their method, called *iterative absorption*, was also based on absorption, but the structure they set aside had basic combinatorial properties instead of algebraic. They applied this method repeatedly to make the leftover part more and more structured. After the leftover is structured enough, they can use an absorber, set aside at the start, that can absorb everything that's left. The toroidal n-queens problem asks how many ways n queens can be placed on a nxn chessboard, where the board is considered on the surface of the torus, such that no pair of queens attack each other. In 2021, Bowtell and Keevash used a random greedy algorithm and a complex absorption method, utilising ideas of the iterative absorption and randomised algebraic construction, to assymptotically resolve the problem for $n \equiv 1, 5 \pmod 6$.

In general, a proof using the *absorption method* can be described as follows. First, we set aside a structure in the host graph, called the absorber. Next, we almost perfectly pack into the rest of the graph. Finally, we prove that whatever is leftover, together with our absorber it can be perfectly packed into. This last step usually relies on properties of the absorber established at their construction, as well as properties of the leftover maintained through the almost perfect packing process. Generally speaking, the stronger an absorber we can create, the less careful we have to be in the almost perfect packing part. This method was used, described and popularised under the name absorption by Rödl, Ruciński and Szemerédi in 2006 [51]. We note that a similar method was already used by Krivelevich in 1997

[39].

Intuitively, perfect packing results are hard precisely because every edge must be used. If the hypergraphs $\mathcal{G}$ were embedded in order to H, on coming to the last hypergraph of $\mathcal{G}$ we would need to find that a hole is left in H of precisely the right shape to accommodate it; this clearly requires some foresight in the packing. If some edges remain uncovered at the end, this difficulty decreases.

In our work we consider D-degenerate hypergraphs. First, let us define the term. An ordering of $V(G)$ is D-degenerate if for each vertex v, there are at most D edges of G whose final vertex is v. We say G is D-degenerate if $V(G)$ has a D-degenerate ordering. In particular, we define *trees* as connected 1-degenerate hypergraphs. We also need to define the maximum degree $\Delta(G) := \max_{v\in V(G)} \big|\{e \in E(G) : v \in e\}\big|$. Of course, every hypergraph of bounded maximum degree has automatically bounded degeneracy. Note that an n-vertex D-degenerate hypergraph has less than Dn edges, and so trivially has maximum degree at most Dn.

Our results can be interpreted as hypergraph analogues of the two widest known tree packing conjectures with a maximum degree $\frac{cn}{\log(n)}$ condition.

Conjecture 1.3.2 (Ringel's conjecture)**.** *Given n and a tree on* $n+1$ *vertices,* $2n+1$ *copies of this tree pack into the complete graph* K_{2n+1}.

Ringel stated this conjecture in 1968 [49]. Simple calculations show that the packing in this conjecture is a perfect packing. Early results towards proving the conjecture only packed simple trees, for example stars or paths. The first general result was by Böttcher, Hladký, Piguet and Taraz in 2016 [11]. They proved an almost perfect packing version of Ringel's conjecture for bounded degree graphs. Joos, Kim, Kühn and Osthus [31] proved the perfect packing result, assuming bounded degree as well. Most recently, Montgomery, Pokrovskiy and Sudakov [43] and later Keevash and Staden [33] proved Ringel's conjecture for all sufficiently large n.

Note, that both the almost perfect packing results of Böttcher, Hladký, Piguet and Taraz [11] and the perfect packing results of Joos, Kim, Kühn and Osthus [31] allow for much more general families of trees than Ringel's conjecture.

Conjecture 1.3.3 (Gyárfás' conjecture)**.** *Given n and a family of trees* $T_1, \dots, T_n$ *with* $|V(T_i)| = i$*, the family packs into the complete graph* K_n.

Gyárfás stated this conjecture, also known as the *tree packing conjecture*, in 1978 [28]. Note, that once again the conjecture requires a perfect packing. Similarly to

Ringel's conjecture, the initial results were about very specific tree-types. In 2013, Balogh and Palmer [5] proved that the largest $n^{\frac{1}{4}}$ trees can be packed if none of them are stars. This was an important step, as in general spanning or almost spanning trees are much harder to pack than smaller trees. The first general almost perfect packing result is again by Böttcher, Hladký, Piguet and Taraz [11] for bounded degree graphs. Joos, Kim, Kühn and Osthus [31] in their already mentioned paper also proved Gyárfás' conjecture for all large values of n and bounded degree trees.

Moving one step forward from trees, in 1967, attending a conference in Oberwolfach, Ringel posed the following problem.

Problem 1.3.4 (Oberwolfach problem)**.** *Given an odd number n and a two-regular graph F on n vertices, for what n and F can we perfectly pack copies of F into K_n?*

In 2021, Glock, Joos, Kim, Kühn and Osthus showed that this is always possible, if n is sufficiently large, no matter what F is [27]. They use an absorption method, utilising tools from several already mentioned papers, including Rödl nibble and Keevash's proof of the existence of designs. Later, Keevash and Staden solved a generalised version of the Oberwolfach problem [34]. They proved that any quasirandom dense large graph in which all degrees are equal and even can be decomposed into any given collection of two-factors.

After this short detour to graph packing results, let us return to hypergraph packings. In 2021, Ehard and Joos proved that a family of uniform bounded degree hypergraphs packs into any quasirandom host graph almost perfectly. With this result they address questions of Kim, Kühn, Osthus and Tyomkyn [37], as well as Keevash [36]. This result is similar to our Theorem 4.7.1, which proves the same type of statement for bounded degeneracy and a maximum degree of $\frac{cn}{\log(n)}$, therefore it can be applied to a wider family of graphs. Their results however apply to the partite setting and sparser graphs.

Our results

In the entirety of Chapter 4 we consider hypergraphs that are r-*uniform*, that is each edge of a hypergraph G is an r-element subset of $V(G)$. We will refer to $(r-1)$-element vertex sets as *semi-edges*.

Our first main result says that we can approximately pack the complete graph $K_n^{(r)}$ with hypergraphs of bounded degeneracy and not too large maximum degree.

Theorem 1.3.5. *For each $r \geq 2$, each $\gamma > 0$ and each $D \in \mathbb{N}$ there exist $c > 0$ and a number n_0 such that the following holds for each integer $n > n_0$. Suppose that $(G_t)_{t\in[t^*]}$ is a family of D-degenerate r-uniform hypergraphs, each of which has at most n vertices and maximum degree at most $\frac{cn}{\log n}$. Suppose further that the total number of edges of $(G_t)_{t\in[t^*]}$ is at most $\binom{n}{r} - \gamma n^r$. Then $(G_t)_{t\in[t^*]}$ packs into the complete graph $K_n^{(r)}$.*

Our second main result is a bit more complicated. If we insist that the graphs G_t each have linearly many vertices of degree 1, and in addition these graphs are not too close to spanning, then we can upgrade 'covering almost all the edges' to a perfect packing.

Theorem 1.3.6. *For each $r \geq 2$, every D and $\mu > 0$ there are n_0 and $c > 0$ such that for every $n \geq n_0$, the following holds. Suppose that $(G_t)_{t\in[t^*]}$ is a family of D-degenerate r-uniform hypergraphs, each of which has at most $n - \lfloor \mu n \rfloor$ vertices, at least $\lfloor \mu n \rfloor$ leaves and maximum degree at most $\frac{cn}{\log n}$. Suppose further that the total number of edges of $(G_t)_{t\in[t^*]}$ is exactly $\binom{n}{r}$. Then $(G_t)_{t\in[t^*]}$ perfectly packs into the complete graph $K_n^{(r)}$.*

Note that with this theorem we also prove reasonable hypergraph analogues of Ringel's conjecture and Gyárfás' conjecture for typical hypergraphs. Trees for us are 1-degenerate hypergraphs. In Ringel's conjecture all graphs are non-spanning, which this theorem can handle. In case of Gyárfás' conjecture, some graphs are close to spanning, but the technical version of this theorem which is stated as Theorem 4.1.4 can handle that. Our two extra restrictions are the requirement of μn leaves as well as the maximum degree $\frac{cn}{\log n}$. Both of these are true for *typical* 1-degenerate hypergraphs. Here by typical we mean that if we randomly generate the back-edges in degeneracy order to get a tree, this tree will have these properties with high probability.

Most of the work of Chapter 4 is to analyse a natural randomised algorithm which packs almost-spanning hypergraphs. The graph version of this algorithm was previously analysed by Allen, Böttcher, Hladký and Piguet [3] and further by Allen, Böttcher, Clemens and Taraz [2]. They proved Theorem 1.3.5 and Theorem 1.3.6 for $r = 2$ respectively. Some of the analysis carries over to hypergraphs, but there are points where a new idea is needed which we will highlight.

2

Low Independence Number and Hamiltonicity Implies Pancyclicity

2.1 Introduction

A *Hamilton cycle* of a graph is a cycle that passes through all its vertices. A graph is *Hamiltonian* if it contains a Hamilton cycle as a subgraph. It is difficult to decide whether a graph contains a Hamilton cycle, therefore it is valuable to establish useful sufficient conditions for Hamiltonicity. A graph is *pancyclic* if it contains a cycle of every length $3 \leq \ell \leq n$, where n denotes the number of vertices. The independence number of a graph G is the size of the largest stable set, denoted by $\alpha(G)$.

In this chapter we study a connection between Hamiltonicity, pancyclicity and independence number. Assuming G is a Hamiltonian graph with independence number at most k, we are looking for the minimum number of vertices $f(k)$ that guarantees that G is pancyclic.

In this section we introduce an equivalent theorem to Theorem 1.1.4.

Definition 2.1.1. For $\beta > 0$, let $SC(\beta)$ denote the following statement, called *short cycle statement*. There exists $c > 0$ such that given a Hamiltonian graph G with $n \geq ck^{\beta}$ vertices and independence number at most k, and any subset of vertices W of size at most $20k^2$, we can find a cycle of length $n - 1$ containing all the vertices from W.

Theorem 2.1.2 (Lee, Sudakov [40])**.** *For all $\beta \geq 2$, assuming $SC(\beta)$ the following statement holds. There exists $c' > 0$, such that if G is a Hamiltonian graph with $n \geq c'k^{\beta}$ vertices and independence number at most k, then G is pancyclic. In other words $f(k) \leq c'k^{\beta}$.*

The above theorem is implicitly proved in [40]. To see this one follows their *Proof of Theorem 1.1*. This gives the stronger conclusion if their Theorem 2.1 is replaced by $SC(\beta)$.

Theorem 2.1.3. *$SC(11/5)$ holds.*

Using Theorem 2.1.2 and Theorem 2.1.3, Theorem 1.1.4 immediately follows.

For a proof of Theorem 2.1.3 we substantially extend the methods of [40]. The improvement comes from Lemma 2.3.9 and the inductive approach to proving Lemma 2.3.5, which is made possible by Lemma 2.3.11. Proving these new lemmas constitutes most of Section 2.3. Before that, we will state some definitions and prove a basic structural proposition in Section 2.2.

The goal of Section 2.2 is to prove Theorem 2.1.3.

2.2 Definitions, earlier results

The core idea of the proof of Theorem 2.1.3, as well as those of previous results, is the following. We break down the graph into parts along the Hamilton cycle, we call these parts *arcs*. Next we show that if we have certain edge configurations between these arcs, then we can also find a cycle of length $n-1$. Finally, we prove that a graph contains either a sparse subgraph that contradicts the independence number constraint, or a certain type of expander that implies the existence of one of the aforementioned edge configurations. In this section we state the basic definitions and prove Proposition 2.2.9, which shows the existence of the *arc-system* we will use in Section 2.3.

We make no attempt to find the optimal value of c in Theorem 1.1.4. For this reason we can ignore small rounding errors and thus will omit all floor and ceiling signs. We fix a large constant c, how large we actually need will come from later calculations so we do not specify at this point. We can also assume that $ck^{11/5} \leq 4k^4$ by the result of Erdős [19], thus by setting c large we can assume that k is also large.

Assumption 2.2.1. *From this point we assume for a contradiction to Theorem 2.1.3 the following:*

- *G is a Hamiltonian graph with $n \geq ck^{11/5}$ vertices,*
- *G has independence number at most k,*
- *W is a subset of $V(G)$ with at most $20k^2$ vertices,*

- *G has no cycle of length $n-1$ containing W,*
- *H is a Hamilton cycle in G.*

We call a vertex of G *problematic* if it has degree at most $2k$ or is an element of W. There are at most $2k^2$ vertices with degree at most $2k$ (by the greedy algorithm for finding independent sets), so there are at most $22k^2$ problematic vertices. Let $\widetilde{W}$ denote the set of problematic vertices. The motivation for calling these vertices problematic comes from Proposition 2.2.2.

We call a cycle C a *contradicting cycle* if it has length $n-k \leq |C| \leq n-1$ and contains all problematic vertices.

The following proposition shows that in the graphs we are considering no contradicting cycle exists, thus justifying their name.

Proposition 2.2.2 ([40] Proposition 3.1)**.** *If G satisfies Assumption 2.2.1, then there is no contradicting cycle in G.*

We say two vertices of G are *consecutive* if they are neighbours in H. A set of vertices is *continuous* if they form a path in H. For a subset $A \subseteq V(G)$ the *continuous closure* of the set is the $\overline{A}$ continuous set of minimum size that contains it (in general this might not always be unique, but we will use it only in cases when it is).

Next we define arc-systems, which are the objects that we will primarily use in the rest of the section.

Definition 2.2.3. A family of subsets of $V(G)$ (where the graph G has a fixed Hamilton cycle H) denoted by $\mathcal{A}$ is called an *arc-system* and its elements *arcs* if the following properties hold:

(i) For all $A \neq B \in \mathcal{A}$, we have $\overline{A} \cap \overline{B} = \emptyset$, that is, the continuous closures of arcs are pairwise disjoint,

(ii) For all $A \in \mathcal{A}$, we have $\overline{A} \cap \widetilde{W} = \emptyset$, that is, the continuous closure of each arc has no problematic vertex in it,

(iii) For all $A \in \mathcal{A}$, no two vertices in A are consecutive.

Remark 2.2.4. *If A is an arc and $|\overline{A}| \leq k+2$, then A is an independent set. Indeed, if there was an edge $\{u, v\}$ where $u, v \in A$, then using the longer path between u and v in H and the edge $\{u, v\}$ we would get a contradicting cycle (see Figure 2.1a).*

Definition 2.2.5. We call an arc system $\mathcal{A}$ *independent* if it has the property that for all A in $\mathcal{A}$ we have $|\overline{A}| \leq k/6$.

We use $k/6$ in the definition instead of $k + 2$ for technical reasons, as in later proofs we will want to remove vertices from up to six arcs to find a contradicting cycle.

We say the *size* of the arc-system is $|\mathcal{A}|$ and the *length* of the arc system is $\min_{A\in\mathcal{A}} |A|$.

Proposition 2.2.6. *Given c_1 and c_2, there exists c such that if we assume Assumption 2.2.1, then there is an independent arc-system in the graph G of size c_1k^2 and length $c_2k^{1/5}$.*

Proof. We start with the empty arc-system. Removing the problematic vertices from H, we obtain a set $\mathcal{P}$ of at most $22k^2$ paths. From this set we will construct an arc-system $\mathcal{A}$ with the desired properties.

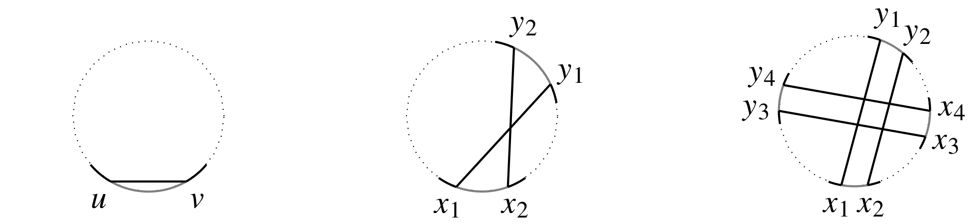
While there is a path $\{v_1, v_2, \dots, v_m\} = P \in \mathcal{P}$ such that $m \geq 2c_2k^{1/5}$ we do the following. Remove P from $\mathcal{P}$. Add $\{v_{2c_2k^{1/5}+1}, v_{2c_2k^{1/5}+2}, \dots, v_m\}$ to $\mathcal{P}$. Add $\{v_1, v_3, \dots, v_{2c_2k^{1/5}-1}\}$ to $\mathcal{A}$. In words, we remove the first $2c_2k^{1/5}$ vertices of P and form an arc from every second vertex in it, and add that arc to $\mathcal{A}$. Since we choose such sets, by definition $\mathcal{A}$ fulfills properties (ii) and (iii) of arc-systems. Also since k is large enough we have $2c_2k^{1/5} \leq k/6$, so $\mathcal{A}$ is independent by definition. The change to $\mathcal{P}$ ensures that property (i) will also hold for $\mathcal{A}$.

At the end of this process we have at most $(22k^2)(2c_2k^{1/5})$ leftover vertices (from paths shorter than $2c_2k^{1/5}$), we removed $22k^2$ problematic vertices at the start, and the half of the other vertices were used to form arcs in $\mathcal{A}$. So we have at least $\frac{ck^{11/5}-(22k^2)(2c_2k^{1/5}+1)}{2c_2k^{1/5}}$ arcs in $\mathcal{A}$, which is more than c_1k^2 if c is large enough. □

Definition 2.2.7. We denote by M_2 the matching of size two, or equivalently two independent edges. We say a graph is *M_2-free* if it doesn't have two independent edges (or equivalently in bipartite graphs, there is a vertex that is incident to all edges).

Definition 2.2.8. We say an arc-system is *simple* if it is independent and for each pair of arcs, the subgraph of G induced by them is M_2-free.

Proposition 2.2.9. *Given c_1 and c_2, there exists c such that if we assume Assumption 2.2.1, then in the graph G there is a simple arc-system of size c_1k^2 and length $c_2k^{1/5}$.*



(a) An arc with an edge inside

(b) An intersecting M_2 between arcs

(c) Two non-intersecting M_2

Figure 2.1: Contradicting cycles implied by edge configurations

Proof. We fix an independent arc-system $\mathcal{A}$ provided by Proposition 2.2.6. We draw the vertices of the graph G on a circle in the plane in the order of the cycle H and connect neighbouring vertices with line segments. If there are two independent edges $\{x_1, y_1\}$, $\{x_2, y_2\}$ between arcs A_1 and A_2, then these can intersect on this drawing or not.

If they intersect, then we immediately find a contradicting cycle the following way. Starting from H, remove the edges of the shorter path between x_1, x_2 and, similarly, remove the edges of the shorter path between y_1, y_2 and instead add the edges $\{x_1, y_1\}$ and $\{x_2, y_2\}$ (see Figure 2.1b). This gives a cycle with at least $n - 4c_2k^{1/5} > n - k$ vertices. By Proposition 2.2.2, this is a contradiction to Assumption 2.2.1.

If we find two pairs of arcs each with non-intersecting M_2 such that these two M_2 ($\{x_1, y_1\}$, $\{x_2, y_2\}$ between A_1 and A_2 and $\{x_3, y_3\}$, $\{x_4, y_4\}$ between A_3 and A_4) intersect each other on the drawing, then again we can find a contradicting cycle the following way. From H remove the edges of the shorter path between x_1, x_2; y_1, y_2; x_3, x_4; y_3, y_4 and instead add the edges $\{x_1, y_1\}$, $\{x_2, y_2\}$, $\{x_3, y_3\}$, $\{x_4, y_4\}$ (see Figure 2.1c). This gives a cycle with at least $n - 8c_2k^{1/5} > n - k$ vertices. By Proposition 2.2.2, this is a contradiction to Assumption 2.2.1.

This implies that if we look at the graph on the arcs as vertices where a pair of arcs form an edge if the subgraph of G induced by them contains an M_2, then this is a planar graph. That implies 5-colourability. By taking the majority colour, we get a simple arc-system with size $\frac{c_1}{5}k^2$ and length $c_2k^{1/5}$. □

In some cases we want to consider an auxiliary graph, in which the arcs are the vertices.

Definition 2.2.10. The *arc-graph* of an arc-system $\mathcal{A}$ is the graph $G_{\mathcal{A}}$ where the vertices are the arcs in the system and there is an edge between two arcs in the arc-graph if and only if there is an edge between the two arcs in G. We call the edges of the arc-graph *arc-edges*.

2.3 Key lemmas, proof of the main theorem

The goal of this section is to prove Lemma 2.3.5, which states that arc-systems with certain size and length contain large independent sets. This will imply Theorem 1.1.4. To prove it, we will use induction and two structural lemmas. Lemma 2.3.9 states that an arc system contains either a large independent set or a subsystem of arcs that are expanding. Lemma 2.3.11 states that an expanding arc system must have a certain edge-configuration that we will call a *semi-triangle*. Finally we show that this implies the existence of a contradicting cycle, which is impossible.

Given an arc-system $\mathcal{A}$, let $G[\mathcal{A}]$ denote the subgraph induced by all vertices in the arcs of $\mathcal{A}$.

Lemma 2.3.1. *Given a simple arc-system $\mathcal{A}$ with length a, size b, and m arc-edges in the corresponding arc-graph, there is an independent set in $G[\mathcal{A}]$ of size $ab-m$.*

Proof. Since the arc-system is simple, the edges of $G[\mathcal{A}]$ corresponding to a single arc-edge e can be covered by a single vertex of $G[\mathcal{A}]$ (that is, there is a vertex v in $G[\mathcal{A}]$ such that all the edges corresponding to e are incident to v). Removing these vertices we get an independent vertex set in $G[\mathcal{A}]$ of size $ab - m$. □

Now we define the function that we want to work with.

Definition 2.3.2. Let $g(a, b)$ denote the largest number such that given a simple arc-system $\mathcal{A}$ of length a and size b there is always an independent set of size $g(a, b)$ in $G[\mathcal{A}]$.

Setup 2.3.3. *For all $i \in \mathbb{N}$ we define the following constants that we will use in the following section; $a_i = 10 \cdot 3^i$, $b_i = 1000^i \cdot 4^{i^2}$.*

The lemma that we want to prove in this section is the following.

Definition 2.3.4. For every $p \in \mathbb{N}$ let $AI(p)$ denote the following statement, called the *arc-independence statement*. For every $x \geq 1$ with $a_p x \leq k/6$ we have

$$g\left(a_p x, b_p x^{p(p-1)/2}\right) \geq x^p + 1\,.$$

Lemma 2.3.5. *For every $p \in \mathbb{N}$, $AI(p)$ holds.*

This means that given $\Theta\left(x^{p(p-1)/2}\right)$ arcs with each having at least $\Theta(x)$ vertices, we can find an independent set of size $\Theta(x^p)$.

First, we prove Theorem 1.1.4 using Lemma 2.3.5.

Proof of Theorem 1.1.4. Using Lemma 2.3.5 for $p = 5$ and $x = k^{1/5}$, together with Proposition 2.2.9, we get an independent set of size $k + 1$, and therefore a contradiction. This proves Theorem 2.1.3 and thus Theorem 1.1.4. □

To prove Lemma 2.3.5, we will use induction. To prepare for the induction step, first we prove Lemmas 2.3.9 and 2.3.11.

Definition 2.3.6. Given an arc-system $\mathcal{A}$ of the graph G and a subset of an arc $X \subseteq A \in \mathcal{A}$, we say that the *arc-neighbourhood* of X is

$$N_{\mathcal{A}}(X) = \{B \in \mathcal{A} : (\exists b \in B)(\exists x \in X)\{b, x\} \in E(G)\}.$$

We sometimes write $N_{\mathcal{A}}(v)$ meaning $N_{\mathcal{A}}(\{v\})$ for simplicity. We denote by $d_{\mathcal{A}}(X)$ the size of the arc-neighbourhood, that is $d_{\mathcal{A}}(X) = |N_{\mathcal{A}}(X)|$.

Definition 2.3.7. Using the constants from Setup 2.3.3 and given $p > 1$ integer, we say an arc A is *good* in the arc-system $\mathcal{A}$ if to at least half of the vertices $v \in A$ we can assign a set of

$$4b_{p-1}x^{(p-1)(p-2)/2}$$

arcs in $N_{\mathcal{A}}(v)$, each arc of $\mathcal{A}$ being assigned to at most one $v \in \mathcal{A}$. If an arc is not good, we call it *bad*.

Also, we say that a subset X of an arc $A \in \mathcal{A}$ is *expanding* in $\mathcal{A}$ if $(\forall Y \subseteq X) d_{\mathcal{A}}(Y) \geq |Y| 4b_{p-1}x^{(p-1)(p-2)/2}$.

The definition of good and expanding depends on p, but it will always be clear from the context which p is meant.

We will use the following simple proposition in the proof of Lemma 2.3.9.

Proposition 2.3.8. *Given an arc-system $\mathcal{A}$ of a graph G, the function $d_{\mathcal{A}}$ is submodular.*

Proof. We trivially have

$$|N_{\mathcal{A}}(A)| + |N_{\mathcal{A}}(B)| = |N_{\mathcal{A}}(A) \cup N_{\mathcal{A}}(B)| + |N_{\mathcal{A}}(A) \cap N_{\mathcal{A}}(B)|$$

and $|N_{\mathcal{A}}(A \cap B)| \leq |N_{\mathcal{A}}(A) \cap N_{\mathcal{A}}(B)|$ as the former is a subset of the latter. Furthermore, $N_{\mathcal{A}}(A) \cup N_{\mathcal{A}}(B) = N_{\mathcal{A}}(A \cup B)$. Putting these together we get

$$d_{\mathcal{A}}(A) + d_{\mathcal{A}}(B) \geq d_{\mathcal{A}}(A \cup B) + d_{\mathcal{A}}(A \cap B)\,.$$

□

Lemma 2.3.9. *Let the constants a_i and b_i be as defined in Setup 2.3.3. For each integer $p > 1$, given a simple arc-system $\mathcal{A}$ of size $b_p x^{p(p-1)/2}$ and length $a_p x$, either there is an independent set of size $x^p + 1$ in $G[\mathcal{A}]$ or there is a non-empty $\mathcal{A}' \subseteq \mathcal{A}$ such that for all $A \in \mathcal{A}'$, A is good in $\mathcal{A}'$.*

Proof. We observe that an arc A is good if and only if it has an expanding subset of size at least $|A|/2$, by Hall's theorem.

We define a process, by the end of which we either have the required independent set or the good subset. Let $\mathcal{A}_0 = \mathcal{A}$ be the given arc system of size $b_p x^{p(p-1)/2}$ and length $a_p x$. Let $\mathcal{B}_0$ be the empty system. In step t we will define arc-systems $\mathcal{A}_t$ and $\mathcal{B}_t$ as follows.

If each $A \in \mathcal{A}_{t-1}$ is good in $\mathcal{A}_{t-1}$, then we define $\mathcal{A}' = \mathcal{A}_{t-1}$, $\mathcal{B} = \mathcal{B}_{t-1}$ and the process terminates.

Otherwise we take a bad arc A from $\mathcal{A}_{t-1}$ and a maximal expanding set X in it. Note that $|X| < |A|/2 = a_p x/2$ in this case. We say that $Y \subseteq X$ is *tight* if $d_{\mathcal{A}_{t-1}}(Y) < (|Y| + 1)4b_{p-1}x^{(p-1)(p-2)/2}$. Let B denote $A \setminus X$. Now for every $v \in B$ there is a tight set T_v such that $d_{\mathcal{A}_{t-1}}(T_v \cup \{v\}) < (|T_v| + 1)4b_{p-1}x^{(p-1)(p-2)/2}$ (by the maximality of X). Let T denote $\cup_{v \in B} T_v$.

First we claim that if a set T_i is the union of i tight sets, then

$$d_{\mathcal{A}_{t-1}}(T_i) \leq (|T_i| + i)4b_{p-1}x^{(p-1)(p-2)/2} \tag{2.3.1}$$

holds, which we will prove by induction. We can assume $T_i = T_{i-1} \cup T'$ where T_{i-1} is the union of $i - 1$ tight sets and T' is tight. Then

$$\begin{aligned} d_{\mathcal{A}_{t-1}}(T_i) &\leq d_{\mathcal{A}_{t-1}}(T_{i-1}) + d_{\mathcal{A}_{t-1}}(T') - d_{\mathcal{A}_{t-1}}(T_{i-1} \cap T') \\ &\leq \big((|T_{i-1}| + i - 1) + (|T'| + 1) - |T_{i-1} \cap T'|\big)4b_{p-1}x^{(p-1)(p-2)/2}\,, \end{aligned} \tag{2.3.2}$$

where the first inequality is an application of Proposition 2.3.8. The second inequality follows by induction on T_{i-1}, tightness of T' and expansion of $T' \cap T_{i-1}$.

This proves inequality (2.3.1).

Next, we claim that

$$d_{\mathcal{A}_{t-1}}(B) \leq 2a_p x 4 b_{p-1} x^{(p-1)(p-2)/2} . \tag{2.3.3}$$

First we see that for each $v \in B$ we have

$$|N_{\mathcal{A}_{t-1}}(v) \setminus N_{\mathcal{A}_{t-1}}(T_v)| \leq 4 b_{p-1} x^{(p-1)(p-2)/2} ,$$

since T_v is expanding, as it is a subset of X, and $T_v \cup \{v\}$ is not expanding. This implies

$$|N_{\mathcal{A}_{t-1}}(B) \setminus N_{\mathcal{A}_{t-1}}(T)| \leq a_p x 4 b_{p-1} x^{(p-1)(p-2)/2} . \tag{2.3.4}$$

By its definition, T is the union of tight sets, therefore it is also the union of at most $|T|$ tight sets. As T is a subset of X, $|T| \leq a_p x/2$. Thus, using inequality (2.3.1) on T gives us

$$d_{\mathcal{A}_{t-1}}(T) \leq a_p x 4 b_{p-1} x^{(p-1)(p-2)/2} . \tag{2.3.5}$$

(2.3.4) and (2.3.5) together imply (2.3.3). So in this step we define $\mathcal{B}_t = \mathcal{B}_{t-1} \cup \{B\}$ and $\mathcal{A}_t = \mathcal{A}_{t-1} \setminus \{A\}$. Note that the total size of $\mathcal{A}_t$ and $\mathcal{B}_t$ is $b_p x^{p(p-1)/2}$.

If by the end of this process we have a non-empty good arc-system $\mathcal{A}'$, then we have found what we are looking for.

If $\mathcal{A}'$ is empty, then we have an arc-system $\mathcal{B}$ with length at least $a_p x/2$ and size $b_p x^{p(p-1)/2}$. We count the arc-edges of $G_{\mathcal{B}}$ in the following way. We assign each arc-edge to the arc that was added to $\mathcal{B}$ in the earlier step. By the Equation (2.3.3) property of B proven in the process, each arc will be assigned at most $2a_p x 4 b_{p-1} x^{(p-1)(p-2)/2}$ arc-edges this way. Thus $G_{\mathcal{B}}$ has at most

$$b_p x^{p(p-1)/2} 2 a_p x 4 b_{p-1} x^{(p-1)(p-2)/2}$$

arc-edges. This means $G_{\mathcal{B}}$ has an edge density of at most

$$\frac{b_p x^{p(p-1)/2} 2 a_p x 4 b_{p-1} x^{(p-1)(p-2)/2}}{\binom{b_p x^{p(p-1)/2}}{2}} \leq \frac{d}{b_p x^{p-2}}$$

where d is a constant not depending on x or b_p.

We take a subset C of $\mathcal{B}$ of size x^{p-1} with the minimal amount of arc-edges. G_C

(a) Type 1 (b) Type 2

Figure 2.2: Semi-triangles

will have at most the same edge density as $G_{\mathcal{B}}$. Therefore, G_C has at most

$$\binom{x^{p-1}}{2}\frac{d}{b_p x^{p-2}} \leq \frac{dx^p}{b_p}$$

edges. Thus, by Lemma 2.3.1, $G[C]$ has an independent set of size $\frac{a_p}{2}x^p - \frac{dx^p}{b_p} \geq x^p + 1$ as $a_p \geq 6$ and $b_p \geq d$. $\square$

Definition 2.3.10. Let us fix a direction on the Hamilton cycle H so we can order the vertices of an arc $u < v$ if u is before v in the given direction. We say that three arcs (A, B, C) form a *semi-triangle* if they are in the given order and there exists $a_1 < a_2 \in A, b_1 < b_2 \in B, c_1 < c_2 \in C$ such that one of the following conditions holds (see Figure 2.2):

- Type 1: $\{a_1, c_1\}, \{a_2, b_1\}, \{b_2, c_2\} \in E(G)$ and A and B are not consecutive arcs,
- Type 2: $\{a_1, b_1\}, \{a_2, c_1\}, \{b_2, c_2\} \in E(G)$.

Note that a Type 2 semi-triangle gives us a contradicting cycle (see Figure 2.3a), therefore it can not exist. Using the good arc-system given by Lemma 2.3.9, we will show the existence of certain Type 1 semi-triangles. Later, using those semi-triangles, we find a contradicting cycle.

Given an arc A we get the *main part* of it by taking the second half of it in the given order. That is, $M(A) \subseteq A$, $|M(A)| = |A|/2$ and for all $v \in M(A), u \in A \setminus M(A)$ we have $v > u$. We define *leftover part* as $L(A) = A \setminus M(A)$. For an arc-system $\mathcal{A}$, $M(\mathcal{A})$ and $L(\mathcal{A})$ are the sets of the main and leftover parts, respectively, of each arc in the system.

Lemma 2.3.11. *Let the constants a_i and b_i be as defined in Setup 2.3.3. For each $p > 1$, if $AI(p-1)$ holds, then the following is true.*

Fix a simple arc-system $\mathcal{A}$ of length $a_p x$ and an arc $A \in \mathcal{A}$. Assume that there is an assignment of arcs to vertices with the following properties:

- *The range of the assignment is A', a subset of A of size at least $|A|/6$,*
- *To each $v \in A'$ at least $b_{p-1}x^{(p-1)(p-2)/2}+1$ arcs are assigned from $N_{M(\mathcal{A})}(v)$,*
- *None of the arcs are assigned to more than one vertex.*

Then either there are arcs $B, C \in \mathcal{A}$ such that (A, B, C) is a semi-triangle of Type 1, or there is an independent set in $G[\mathcal{A}]$ of size $x^p + 1$.

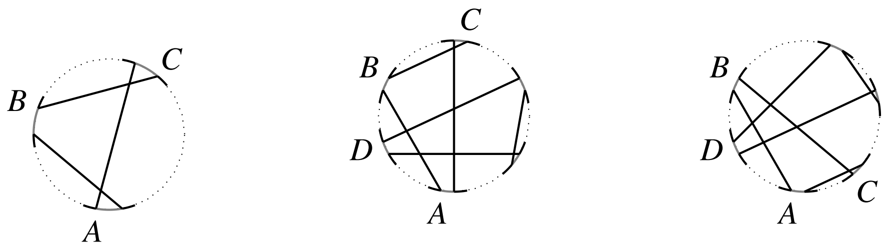
Proof. If any of the assigned neighbouring arcs is consecutively after A, we unassign it. Let us fix $v \in A'$. We look at the corresponding leftover arcs of the assigned neighbours of v, which form an arc-system $\mathcal{A}_v$ of size $b_{p-1}x^{(p-1)(p-2)/2}$ and length $a_p x/2$. Using $AI(p-1)$ and by $a_p \geq 2a_{p-1}$ there is an independent set J_v in $G[\mathcal{A}_v]$ of size $x^{p-1}+1$. Taking x of these sets, which we can do because $a_p x/6 > x$, we get either an independent set of size more than $x^p + 1$ or an edge $\{b, c\}$ between J_w and J_u. Let B and C be the arcs containing b and c, respectively. Without loss of generality, we can assume that A, B, C are in this order on the Hamilton cycle. Then (A, B, C) is a semi-triangle, because of the edge proving that $M(B)$ is a neighbour of w, the edge proving that $M(C)$ is a neighbour of u and the edge $\{b, c\}$ going between $L(B)$ and $L(C)$. Since Type 2 semi-triangles cannot exist, this must be a Type 1 semi-triangle. □

With this, we are ready to prove the main lemma of the section.

Proof of Lemma 2.3.5. We will use induction to prove the lemma. As the base case, we observe that $g(x+1, 1) = x+1$ by Remark 2.2.4. Also, $g(2x, 2x) \geq x^2+1$, since an arc-system of size $2x$ can have at most $\frac{2x(2x-1)}{2}$ arc-edges, therefore by Lemma 2.3.1 it has an independent set of size $4x^2 - \frac{2x(2x+1)}{2} \geq x^2 + 1$.

For the induction step we can assume $p \geq 3$ and that the lemma is true for all lower values of p.

Let $\mathcal{A}'$ be a simple arc system of size $b_p x^{p(p-1)/2}$ and length $a_p x$. Applying Lemma 2.3.9 on $M(\mathcal{A}')$, we obtain a good subsystem $\mathcal{M}'$, or we find an independent set of size $x^p + 1$ and we are done. For each arc in $M \in \mathcal{M}'$, there is an arc $A \in \mathcal{A}'$ such that $M(A) = M$. Let $\mathcal{A}$ denote the set of such A's. Thus, $M(\mathcal{A})$ is good by definition. Using Lemma 2.3.11, we get either a Type 1 semi-triangle or the desired independent set. We define the length of a semi-triangle (A, B, C) as the number of arcs between A and B (note that this is at least 1 by definition). Let (A, B, C)



(a) Type 2 semi-triangle (b) Type 1 semi-triangles, $\mathcal{A}_C$ case (c) Type 1 semi-triangles, $\mathcal{A}_B$ case

Figure 2.3: Contradicting cycles implied by semi-triangles

be a semi-triangle of Type 1 in $\mathcal{A}$ with the shortest length. Let D denote the arc consecutively after A and $\mathcal{A}_A$ those arcs between A, B; $\mathcal{A}_B$ those arcs between B, C; and $\mathcal{A}_C$ those arcs between C, A.

Recall that by the definition of good, there is a subset of vertices $D' \subseteq D$, with $|D'| \geq |D|$, such that we can assign $4b_{p-1}x^{(p-1)(p-2)/2}$ neighbouring arcs to each. For each vertex in D', using the pigeonhole principle, at least $b_{p-1}x^{(p-1)(p-2)/2}+1$ of the assigned arcs are in either $\mathcal{A}_A, \mathcal{A}_B$ or $\mathcal{A}_C$. Using the pigeonhole principle again, at least $|D|/6$ vertices of D are assigned at least $b_{p-1}x^{(p-1)(p-2)/2}+1$ neighbouring arcs from either $\mathcal{A}_A, \mathcal{A}_B$ or $\mathcal{A}_C$. If this is $\mathcal{A}_A$, then by applying Lemma 2.3.11 to A and $\mathcal{A}_A$ we find either a shorter Type 1 semi-triangle contradicting the minimality of (A, B, C), or the desired independent set. If this is $\mathcal{A}_B$ or $\mathcal{A}_C$, then we get a new Type 1 semi-triangle.

Considering this new Type 1 semi-triangle together with (A, B, C), we get one of the arrangements on Figures 2.3b or 2.3c. We call the vertices incident to the edges used in the semi-triangles *relevant vertices*. Note that the order of these six arcs as well as that of the relevant vertices is determined by the definition of Type 1 semi-triangles and the definition of the sets $\mathcal{A}_B$ and $\mathcal{A}_C$. We construct a contradicting cycle the following way: starting from the fixed Hamilton cycle, we remove the edges between the pairs of relevant vertices in the six arcs and add the six edges used in the two semi-triangles. Note that we removed at most $6 \cdot k/6$ edges as $\mathcal{A}$ is an independent arc-system. This graph is clearly 2-regular, and connectivity can be checked using the Figures as the order of vertices used in them is established.

As finding a contradicting cycle is impossible, we must have found the independent set we are looking for at one of these steps. $\square$

3

Maximum Number of Triangle-free Edge Colourings

3.1 Introduction

Let $k \geq 3$ and $r \geq 2$ be natural numbers. By a *colouring* of a graph $G = (V, E)$ with r colours here we mean *edge-colouring*, that is a function $f : E \to \{1, \ldots, r\}$. In this context we refer to the numbers $1, \ldots, r$ as colours and by the colour class of c we mean $f^{-1}(c)$. A colouring is K_k-free, if no colour class contains a copy of K_k. For a graph G, let $F(G, k, r)$ denote the number of K_k-free colourings of G with r colours. Let $F(n, k, r)$ denote the maximum of $F(G, k, r)$ over all graphs G on n vertices.

In this chapter we consider the case $k = 3$ and $r \in \{5, 6\}$. Let $\varphi_r(G) = F(G, 3, r)$ and $\varphi_r(n) = F(n, 3, r)$. We find the approximate value of $\varphi_5(n)$ and $\varphi_6(n)$. We also prove stability results for both cases.

In this section we state stronger, technical versions of Theorem 1.2.8 and Theorem 1.2.5. The reason for this is that these stronger technical theorems will be useful when we want to find the exact extremal graph for the Erdős-Rothschild problem. While the search for the exact result is beyond the scope of this thesis, we state and prove the following technical results.

We first state the technical stability result for $r = 6$ as that is the easier of the two. For two bit vectors $\boldsymbol{\varepsilon}^{(1)}, \boldsymbol{\varepsilon}^{(2)} \in \{0, 1\}^n$, with $\boldsymbol{\varepsilon}^{(1)} = (\varepsilon_1^{(1)}, \ldots, \varepsilon_n^{(1)})$, we define their *distance* as their Hamming-distance, which is the number of indices where they differ.

To state our theorems, we need the following notation.

Definition 3.1.1. For a graph G, a positive integer n, a bit vector $\boldsymbol{\varepsilon} \in \{0, 1\}^n$ and vertex partitions $V_i^0 \cup V_i^1 = V(G)$ for all $i \in \{1, \ldots, n\}$, we define

$$V(\boldsymbol{\varepsilon}) = \bigcap_{j \in [n]} V_j^{\varepsilon_j} .$$

Theorem 3.1.2 (Technical stability for $r = 6$)**.** *For every $0 < \delta < 10^{-30}$ there is an n_0 such that for all graphs G on $n > n_0$ vertices the following holds. Suppose $\varphi_6(G) \geq 3^{n^2/4} 4^{3n^2/16 - \delta n^2}$. Then there are balanced bipartitions $V_i^0 \cup V_i^1 = V(G)$ for each $i \in [6]$ and $\mathcal{E} = \{\boldsymbol{\varepsilon}^{(1)}, \ldots, \boldsymbol{\varepsilon}^{(8)}\}$ with $\boldsymbol{\varepsilon}^{(j)} \in \{0,1\}^6$, furthermore, there is a partition of $\mathcal{E}$ into $\mathcal{E}_0$ and $\mathcal{E}_1$ such that*

(i) We have
$$\left| \bigcup_{\boldsymbol{\varepsilon} \in \mathcal{E}} V(\boldsymbol{\varepsilon}) \right| \geq \left(1 - 10^5 \sqrt{\delta}\right) n .$$

(ii) We have
$$\left| \bigcup_{\boldsymbol{\varepsilon} \in \mathcal{E}_0} V(\boldsymbol{\varepsilon}) \right| = \left| \bigcup_{\boldsymbol{\varepsilon} \in \mathcal{E}_1} V(\boldsymbol{\varepsilon}) \right| \pm 10^8 \sqrt{\delta} n .$$

(iii) For any pair $(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}') \in \mathcal{E}^2$ we have
$$e\Big(V(\boldsymbol{\varepsilon}), V(\boldsymbol{\varepsilon}')\Big) \geq \Big|V(\boldsymbol{\varepsilon})\Big| \cdot \Big|V(\boldsymbol{\varepsilon}')\Big| - 10^7 \delta n^2 .$$

(iv) For any $\boldsymbol{\varepsilon} \in \mathcal{E}$ we have
$$e\Big(V(\boldsymbol{\varepsilon})\Big) \leq 10^7 \delta n^2 .$$

(v) Each pair $(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}') \in \mathcal{E}_0 \times \mathcal{E}_1$ has distance three and each pair $(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}') \in {\mathcal{E}_0}^2 \cup {\mathcal{E}_1}^2$ has distance four.

(vi) Moreover, for any pair $(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}') \in \mathcal{E}^2$ we have
$$\Big|V(\boldsymbol{\varepsilon})\Big| = \Big|V(\boldsymbol{\varepsilon}')\Big| \pm 10^4 \sqrt[4]{\delta} n .$$

Next we state the technical stability result for the more complex $r = 5$ case.

Theorem 3.1.3 (Technical stability for $r = 5$)**.** *For every $0 < \delta < 10^{-30}$ there is an n_0 such that for all graphs G on $n > n_0$ vertices the following holds. Suppose $\varphi_5(G) \geq 6^{n^2/4 - \delta n^2/4}$. Then there are balanced bipartitions $V_i^0 \cup V_i^1 = V(G)$ for each $i \in [5]$, an integer $t \in \{4, 6, 8\}$ and $\mathcal{E} = \{\boldsymbol{\varepsilon}^{(1)}, \ldots, \boldsymbol{\varepsilon}^{(t)}\}$ with $\boldsymbol{\varepsilon}^{(j)} \in \{0,1\}^5$, furthermore, there is an $s \in [5]$ and a partition of $\mathcal{E}$ into $\mathcal{E}_0$ and $\mathcal{E}_1$ such that*

(i) *We have*

$$\left|\bigcup_{\boldsymbol{\varepsilon}\in\mathcal{E}} V(\boldsymbol{\varepsilon})\right| \geq \left(1 - 10^5\sqrt[4]{\delta}\right) n\,.$$

(ii) *We have*

$$\left|\bigcup_{\boldsymbol{\varepsilon}\in\mathcal{E}_0} V(\boldsymbol{\varepsilon})\right| = \left|\bigcup_{\boldsymbol{\varepsilon}\in\mathcal{E}_1} V(\boldsymbol{\varepsilon})\right| \pm 10^8\sqrt{\delta}n\,.$$

(iii) *For any pair* $(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}') \in \mathcal{E}^2$ *we have*

$$e\Big(V(\boldsymbol{\varepsilon}), V(\boldsymbol{\varepsilon}')\Big) \geq \Big|V(\boldsymbol{\varepsilon})\Big| \cdot \Big|V(\boldsymbol{\varepsilon}')\Big| - 10^7\delta n^2\,.$$

(iv) *For any* $\boldsymbol{\varepsilon} \in \mathcal{E}$ *we have*

$$e\Big(V(\boldsymbol{\varepsilon})\Big) \leq 10^7\delta n^2\,.$$

(v) *Each pair* $(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}') \in \mathcal{E}_0 \times \mathcal{E}_1$ *has distance three and each pair* $(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}') \in \mathcal{E}_0^2 \cup \mathcal{E}_1^2$*has distance two or four.*

(vi) *Moreover, pairs of distance four form a perfect matching in* $\mathcal{E}$ *and for each of these matched pairs* $(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}')$ *we have*

$$\Big|V(\boldsymbol{\varepsilon})\Big| = \Big|V(\boldsymbol{\varepsilon}')\Big| \pm 10^4\sqrt[4]{\delta}n\,.$$

Following the implied partitions, it is easy to prove the main theorems using these technical results. The goal of the following sections is to prove these two theorems.

In Section 3.2 we introduce basic notation that we'll use throughout the chapter and cite a container theorem by Balogh et al.[6] which will be a key tool in proving our results. In Section 3.4 we prove a technical theorem, Theorem 3.4.3, which provides an upper bound on the number of colourings and some structural properties of any graph and corresponding containers which comes close to this upper bound. This result is based on the structure of edges appearing in exactly 3 containers, which is the main novelty in this chapter. In Section 3.5 we prove matching lower bounds by providing a family of constructions and colourings on them. In Section 3.6 we build on Theorem 3.4.3 to prove Theorem 3.1.2 and Theorem 3.1.3.

3.2 Preliminaries

We denote by $\varphi_r(G)$ the number of triangle-free r-colourings of $E(G)$, and by $\varphi_r(n) := \max\{\varphi_r(G) : G$ is a graph with n vertices$\}$. We denote by $tr(G)$ the number of triangles in G.

Suppose we are given a graph G and subgraphs $C_1, \ldots, C_r \subseteq G$ with $E(G) = E(C_1) \cup \ldots \cup E(C_r)$. We denote by $\Phi(C_1, \ldots, C_r)$ the family of all colourings χ in which $\chi^{-1}(i) \subseteq E(C_i)$ for all $i \in [r]$. If each C_i is triangle-free, then every such colouring is triangle-free, hence we have $\varphi_r(G) \geq |\Phi(C_1, \ldots, C_r)|$ in this case. Furthermore, the number of such colourings is easy to count. For $i \in [r]$, let $M_i = M_i(C_1, \ldots, C_r) \subseteq G$ denote the graph on $V(G)$ whose edges are those of G contained in exactly i of the subgraphs $C_1, \ldots, C_r$. We denote by $m_i = m_i(C_1, \ldots, C_r)$ the number of edges in M_i. We then have

$$|\Phi(C_1, \ldots, C_r)| = \prod_{i \in [r]} i^{m_i} \qquad \text{and} \qquad \sum_{i \in [r]} i \cdot m_i = \sum_{i \in [r]} |C_i|, \tag{3.2.1}$$

We will make use of the container theorem below proved by Mousset, Nenadov and Steger [44] using the hypergraph container method of Balogh, Morris and Samotij [7] and Saxton and Thomason [53]. We use an equivalent formulation of their result as stated by Balogh et al. in [6]. This approach was introduced by Hàn and Jiménez [29] using ideas of Clemens, Das and Tran [13].

Theorem 3.2.1 ([6, Theorem 3.2])**.** *There exists constant C such that for every graph G on $n > C$ vertices there exists a collection $\mathcal{C} = \mathcal{C}(G)$ of subgraphs of G such that the following holds:*

(a) every triangle-free subgraph $G' \subseteq G$ is a subgraph of some $C \in \mathcal{C}$,

(b) $tr(C) \leq n^{25/9}$ for every $C \in \mathcal{C}$,

(c) $|\mathcal{C}| \leq \exp\left(n^{16/9}\right)$. □

Theorem 3.2 from [6] is stated for $G = K_n$ only, however, by taking the intersection of each container C with G we obviously obtain the family $\mathcal{C} = \mathcal{C}(G)$ from above.

For a graph G on n vertices, let us denote by $C(G)$ the family of all graphs $C \subseteq G$ on n with $tr(C) \leq n^{25/9}$. An immediate corollary of Theorem 3.2.1 provides the following.

Corollary 3.2.2. *There is a constant n_0, such that for $n > n_0$ and all graphs G on n vertices and every $r \in \mathbb{N}$, we have*

$$\varphi_r(G) \leq \exp\left(r \cdot n^{16/9}\right) \cdot \max_{C_1,\ldots,C_r \in \mathcal{C}(G)} |\Phi(C_1, \ldots, C_r)| .$$

3.3 An approximate upper bound for many colours

In this section, we showcase the strength of the container method on the Erdős-Rothschild problem. The following theorem and proof is by Hàn and Jiménez. Note that all previous theorems of this type made use of Szemerédi's regularity lemma and were much more technical. In Section 3.4 we will strengthen this theorem for 5 and 6 colours.

Theorem 3.3.1. *We have $\varphi_r(n) \leq \left(\frac{r}{2}\right)^{(1+o(1))\binom{n}{2}}$.*

Theorem 3.3.1 basically follows from the following claim and Corollary 3.2.2.

Claim 3.3.2. *Let $\varepsilon > 0$, let $r \geq 6$ be an integer and let $C_1, \ldots, C_r$ be graphs on $[n]$ with $e(C_i) \leq (1/2 + \delta)\binom{n}{2}$. Then we have*

$$|\Phi(C_1, \ldots, C_r)| \leq \left(\left(\frac{1}{2} + \delta\right) r\right)^{\binom{n}{2}} .$$

Proof. Let $m := e(G)$ and let $i_0 = \sum_{i \in [r]} e(C_i)/m \leq r(\frac{1}{2} + \delta)\binom{n}{2}/m$. Note that, by convexity, we have

$$\begin{aligned} |\Phi(C_1, \ldots, C_r)| &\leq \prod_{i \in [r]} i^{m_i} \leq i_0^{m_1 + \ldots + m_r} \\ &\leq \left(r\left(\frac{1}{2} + \delta\right)\frac{n^2}{2m}\right)^m . \end{aligned}$$

Since $r \geq 6 > 2e$, it follows from elementary calculus that $(r(\frac{1}{2} + \delta)\frac{n^2}{2m})^m$ is monotone increasing in m. The claim now follows since $m \leq \binom{n}{2}$. □

To finish the proof of Theorem 3.3.1, we would like to bound the number of edges based on the number of triangles. The problem of finding the best such bound is known as the Erdős-Rademacher problem, first asked by Erdős [21] in 1955. The problem was resolved Liu, Pikhurko and Staden [41] in 2017. Their answer is quite complex, therefore we will make use of the following simpler theorem of Bollobás [8] instead.

Theorem 3.3.3 (Bollobás, [8])**.** *A graph with n vertices and m edges has at least $\frac{n}{9}(4m - n^2)$ triangles.*

Proof of Theorem 3.3.1. Let G be an n vertex graph and let $C(G)$ be the collection provided by Theorem 3.2.1. Let $C_1, \ldots, C_r \in C(G)$ with

$$|\Phi(C_1, \ldots, C_r)| = \max_{C_1,\ldots,C_r \in C(G)} |\Phi(C_1, \ldots, C_r)| \, .$$

Using Theorem 3.3.3, we deduce that $e(C_i) \leq n^2/4 + 9/4 \cdot n^{16/9} = (1/2 + o(1))\binom{n}{2}$ for every $i \in [r]$. Hence, using Corollary 3.2.2 and Claim 3.3.2, we deduce

$$\begin{aligned} \varphi_r(G) &\leq \exp\left(r \cdot n^{16/9}\right) \cdot |\Phi(C_1, \ldots, C_r)| \\ &\leq \left(\frac{r}{2}\right)^{(1+o(1))\binom{n}{2}} . \end{aligned}$$

□

3.4 An approximate upper bound for five and six colours

In this section, we improve the upper bound from Theorem 3.3.1 for five and six colours. These improved upper bounds are asymptotically best possible in n.

Theorem 3.4.1. *For every graph G on n vertices, the number of 5-colourings of $E(G)$ without monochromatic triangles is at most*

$$6^{n^2/4+o(n^2)}.$$

Theorem 3.4.2. *For every graph G on n vertices, the number of 6-colourings of $E(G)$ without monochromatic triangles is at most*

$$3^{n^2/4}4^{3n^2/16+o(n^2)}.$$

In fact we shall prove a more general result, which will be useful for the stability results we will discuss later.

Theorem 3.4.3. *There exists $c > 0$ such that for every $10^{-20} > \delta > 0$ there is an n_0 such that for all $n > n_0$ and all graph G on n vertices the following holds. Suppose* *(i)* $C_1, \ldots, C_5 \subseteq G$, *are such that* $|\Phi(C_1, \ldots, C_5)| \geq 6^{n^2/4 - \delta n^2}$ *or*

(ii) $C_1, \ldots, C_6 \subseteq G$, *are such that* $|\Phi(C_1, \ldots, C_6)| \geq 3^{n^2/4} 4^{3n^2/16 - \delta n^2}$.

and in either case $tr(C_i) \leq n^{25/9}$ *for all* C_i.

Then all the C_i*'s and* $M_3(C_1, \ldots, C_r)$ *are* $c\delta n^2$*-close to being balanced complete bipartite. Furthermore,* $m_1, m_5, m_6 \leq c\delta n^2$*. In case (i) we have* $m_2 + 2m_4 = n^2/4 \pm c\delta n^2$*. In case (ii) we have* $m_2 + 2m_4 = 3n^2/8 \pm c\delta n^2$*. In particular, we have* $|\Phi(C_1, \ldots, C_5)| \leq 6^{n^2/4 + c\delta n^2}$ *and* $|\Phi(C_1, \ldots, C_6)| \leq 3^{n^2/4} 4^{3n^2/16 + c\delta n^2}$.

Remark 3.4.4. *Our proof provides the constant* $c = 10^6$ *and we make no effort to find the lowest possible* c*. We don't believe the dependency on* δ *can be improved, apart from the constant* c*. If you start with our extremal examples and remove* δn^2 *arbitrary edges, the conditions hold but you cannot hope for anything better.*

Theorem 3.4.1 and Theorem 3.4.2 easily follow from Theorem 3.4.3 and Corollary 3.2.2.

We will use the following theorem of Füredi [25].

Theorem 3.4.5 (Füredi, [25])**.** *Every triangle-free graph with at least* $\frac{n^2}{4} - t$ *edges has a bipartite subgraph with at least* $\frac{n^2}{4} - 2t$ *edges.*

3.4.1 Upper bound proof for five colours

We will now prove Theorem 3.4.3 for $r = 5$. By the triangle removal lemma [52], there is some $\delta' > 0$, so that every graph with at most $\delta' n^3$ triangles can be made triangle-free by removing $\delta n^2/5$ edges. Hence, for every $i \in [5]$ and every large enough n, there are triangle-free subgraphs $\tilde{C}_i \subseteq C_i$ with $e(\tilde{C}_i) \geq e(C_i) - \delta n^2/5$. It is easy to see now that

$$\left|\Phi(\tilde{C}_1, \ldots, \tilde{C}_5)\right| \geq 5^{-\delta n^2} \left|\Phi(C_1, \ldots, C_5)\right| \geq 6^{n^2/4 - 2\delta n^2}. \qquad (3.4.1)$$

Define $\alpha_1, \ldots, \alpha_5 \in \mathbb{R}$ so that $e(\tilde{C}_i) = (1/2 - \alpha_i)\binom{n}{2}$ and note that $-1/(2n-2) \leq \alpha_i \leq 1/2$ for every $i \in [5]$. Furthermore, let $\alpha = (\alpha_1 + \ldots + \alpha_5)/5$ and define $\mu_3 \in [0, 1]$ by $m_3 = \mu_3 \binom{n}{2}$. The theorem follows from the following three claims.

Claim 3.4.6. *We have*

$$\mu_3 \leq \frac{3}{4}\left(\frac{75}{2} + \frac{75}{8}\alpha\right)\alpha + \frac{1}{2} + O\left(\frac{1}{n}\right).$$

Claim 3.4.7. *We have*

$$\alpha \leq \frac{1}{5\log(4)} \cdot \left((4\log(3) - 3\log(4))\mu_3 + \frac{3\log(4)}{2} - 2\log(3) + 16\log(6) \cdot \delta \right).$$

Claim 3.4.8. *We have* $\alpha \leq 0.015$.

Before proving the claims, we show how they imply the theorem.

Proof of Theorem 3.4.3. Putting together Claim 3.4.6 and Claim 3.4.7 we get

$$\alpha \leq \frac{1}{5\log(4)} \cdot \left((4\log(3) - 3\log(4)) \frac{3}{4} \left(\frac{75}{2} + \frac{75}{8}\alpha \right) \alpha + 16\log(6) \cdot \delta \right) + O\left(\frac{1}{n}\right).$$

Since n is large enough we can assume that $O\left(\frac{1}{n}\right) \leq \delta$ and thus we can rearrange the quadratic inequality to

$$0 \leq \alpha^2 - b\alpha + c\delta,$$

where $b = -4 + \frac{5\log(4)}{4\log(3)-3\log(4)} \frac{32}{225} \approx 0.18485$ and $c = 1 + \frac{16\log(6)}{4\log(3)-3\log(4)} \frac{32}{225} \approx 18.3083$.

By solving for α, we get that either

$$\alpha \leq \frac{b - \sqrt{b^2 - 4c\delta}}{2} \qquad \text{or} \qquad \alpha \geq \frac{b + \sqrt{b^2 - 4c\delta}}{2} > 0.1.$$

By Claim 3.4.8, the second case is impossible. A simple calculation shows that $b - \frac{3c\delta}{b} < \sqrt{b^2 - 4c\delta}$ and thus we have $\alpha \leq \frac{3c\delta}{2b} < 150\delta$, which implies that $\alpha_i < 750\delta$ for every $i \in [5]$. Therefore, C_i is $10000\delta n^2$ close to being balanced complete bipartite (using Theorem 3.4.5, we get a large bipartite subgraph, which we can balance by moving a few vertices and make it complete by adding the missing edges). Using Claim 3.4.6 we also get that $\mu_3 < 1/2 + 100000\delta$ and therefore $M_3(C_1, \ldots, C_5)$ is 1000000δ close to being balanced complete bipartite.

By (3.4.1) and (3.2.1), we have

$$6^{n^2/4 - 2\delta n^2} \leq \left| \Phi(\tilde{C}_1, \ldots, \tilde{C}_5) \right| \leq \prod_{i=1}^{5} i^{m_i}$$

and $\sum_{i \in [5]} i \cdot m_i = \sum_{i \in [5]} e(\tilde{C}_i)$. Using our bounds on μ_3 we know that $m_3 \leq n^2/4 + 400\delta$. Simple linear optimisation on m_i shows that the only solutions to these constraints are in the form $m_1, m_5, m_6 \leq 10000\delta n^2$ and $m_2 + 2m_4 = \frac{n^2}{4} \pm 10000\delta n^2$. □

It remains to prove Claims 3.4.6, 3.4.7 and 3.4.8.

Proof of Claim 3.4.6. By applying Theorem 3.4.5 we get bipartite graphs $B_i \subseteq \tilde{C}_i$ such that $e(B_i) = (1/2 - 2\alpha_i)\binom{n}{2} - O(n)$ for every $i \in [5]$. We will denote by $P_1(B_i)$ and $P_2(B_i)$ the smaller and larger part of vertices respectively in the bipartition (breaking ties arbitrarily). Let $G_3 = M_3(\tilde{C}_1, \tilde{C}_2, \tilde{C}_3, \tilde{C}_4, \tilde{C}_5)$ and define $x_i \in [0, 1/2]$ such that $v(P_1(B_i)) = \left(\frac{1}{2} - x_i\right) n$.

We say an edge e is *missing*, if there is an $i \in [5]$ such that $e \notin \tilde{C}_i$ and e shares a vertex with both parts of B_i. We say an edge e is *extra* if there is an $i \in [5]$ such that $e \in \tilde{C}_i$ and e is contained in one part of B_i.

For a triangle T we define $e_i(T) = |E(T) \cap E(\tilde{C}_i)|$ and $e_s(T) = \sum_{i=1}^{5} e_i(T)$. Note that, if a triangle T has no missing or extra edges, then $e_i(T) \in \{0, 2\}$. Furthermore, if a triangle T is contained in G_3, then $e_s(T) = 9$. This implies that each triangle in G_3 has a missing or an extra edge. We assign to each triangle of G_3 a missing or an extra edge such that, if it has both, then we choose a missing one.

The number of missing edges in $\tilde{C}_i$ is at most $2(\alpha_i - x_i^2)\binom{n}{2} + O(n)$ and each of them is assigned to at most n triangles. The number of extra edges in $\tilde{C}_i$ is at most $\alpha_i\binom{n}{2} + O(n)$ and each of them is assigned to at most $\left(\frac{1}{2} + x_1\right) n$ triangles as the third vertex must be on the same part of the bipartition (otherwise it would have a missing edge as well, as $\tilde{C}_i$ is triangle-free).

Let $tr(H)$ denote the number of triangles in a graph H. From the above discussion, we get

$$\begin{aligned}
tr(G_3) &\leq \sum_{i=1}^{5} \left(2(\alpha_i - x_i^2)\binom{n}{2}n + \alpha_i\binom{n}{2}\left(\frac{1}{2} - x_i\right)n\right) + O(n^2) \\
&= \binom{n}{3}\sum_{i=1}^{5}\left(\frac{15}{2}\alpha_i - 6\left(x_i - \frac{\alpha_i}{4}\right)^2 + \frac{3\alpha_i^2}{8}\right) + O(n^2) \\
&\leq \binom{n}{3}\sum_{i=1}^{5}\left(\frac{15}{2}\alpha_i + \frac{3\alpha_i^2}{8}\right) + O(n^2).
\end{aligned}$$

Since $\alpha_i \leq 5\alpha$ and $\sum_{i=1}^{5} \alpha_i = 5\alpha$, we have

$$tr(G_3) \leq \left(\frac{75}{2} + \frac{75}{8}\alpha\right)\alpha\binom{n}{3} + O(n^2)\,. \tag{3.4.2}$$

We now apply Theorem 3.3.3 in the following equivalent form,

$$tr(G_3) \geq \frac{4}{3}\left(\mu_3 - \frac{1}{2}\right)\binom{n}{3}. \tag{3.4.3}$$

Putting (3.4.2) and (3.4.3) together and organising the inequality, we get

$$\mu_3 \leq \frac{3}{4}\left(\frac{75}{2} + \frac{75}{8}\alpha\right)\alpha + \frac{1}{2} + O\left(\frac{1}{n}\right).$$

□

Proof of Claim 3.4.7. By (3.4.1) and (3.2.1), we have

$$6^{n^2/4-2\delta n^2} \leq \left|\Phi(\tilde{C}_1,\ldots,\tilde{C}_5)\right| \leq \prod_{i=1}^{5} i^{m_i}$$

and $\sum_{i\in[5]} i \cdot m_i = \sum_{i\in[5]} e(\tilde{C}_i)$. A simple optimisation shows that $\prod_{i=1}^5 i^{m_i}$ is maximised under this constraint for fixed m_3 when $m_1 = m_5 = 0$ and $2m_2 + 4m_4$ is as large as possible. We conclude that

$$6^{(1/4-2\delta)n^2} \leq \prod_{i\in[5]} i^{m_i} \leq 3^{m_3} \cdot 4^{(\sum_{i=1}^5 e(\tilde{C}_i)-3m_3)/4} \tag{3.4.4}$$

By taking logarithms we obtain

$$\begin{aligned}\log(6)\cdot\left(\frac{1}{2}-4\delta\right)\cdot\frac{n^2}{2} &\leq \log(3)\cdot m_3 + \log(4)\cdot\frac{\sum_{i=1}^5 e(\tilde{C}_i) - 3m_3}{4}\\ &\leq \left(\log(3)\cdot\mu_3 + \log(4)\cdot\frac{\frac{5}{2}-5\alpha-3\mu_3}{4}\right)\cdot\frac{n^2}{2}.\end{aligned}$$

By cancelling $\frac{n^2}{2}$ and rearranging we obtain

$$\frac{5}{4}\cdot\log(4)\cdot\alpha \leq \left(\log(3) + \frac{3}{4}\cdot\log(4)\right)\mu_3 + \frac{5\log(4)}{8} - \frac{\log(6)}{2} + 4\delta\log(6).$$

The claim now follows, since $\frac{5}{2}\cdot\log(4) - 2\log(6) = \frac{3}{2}\log(4) - 2\log(3)$. □

Proof of Claim 3.4.8. Note that (3.4.4) is maximised under the condition

$$\sum_{i\in[5]} m_i = \sum_{i\in[5]} e(\tilde{C}_i)$$

if m_3 is as large as possible, i.e. when

$$3m_3 = \sum_{i\in[5]} e(\tilde{C}_i)$$

or in other words when $\mu_3 = 5(1/2-\alpha)/3$. This implies

$$6^{(1/4-2\delta)n^2} \leq 3^{5(1/2-\alpha)/3\cdot\binom{n}{2}} \leq 3^{5(1/2-\alpha)n^2/6}.$$

Taking logarithms and solving for α, we get

$$\alpha \leq \frac{1}{2} - \frac{3\log(6)}{10\log(3)} + \frac{12\log(6)}{5\log(3)}\cdot\delta \leq 0.01073 + 3.915\delta \leq 0.015.$$

□

3.4.2 Upper bound proof for six colours

This proof is very similar to the $r=5$ case. We include it in this thesis, as these calculations are essential to see that the proof actually works.

Similarly as in the 5-colour case, for large enough n, there is a triangle-free subgraph $\tilde{C}_i \subseteq C_i$ for every $i \in [6]$ with

$$\left|\Phi(\tilde{C}_1,\ldots,\tilde{C}_6)\right| \geq 3^{n^2/4}4^{3n^2/16-2\delta n^2}. \tag{3.4.5}$$

Define $\alpha_1,\ldots,\alpha_6 \in \mathbb{R}$ so that $e(\tilde{C}_i) = (1/2-\alpha_i)\binom{n}{2}$ and note that $-1/(2n-2) \leq \alpha_i \leq 1/2$ for every $i \in [6]$. Furthermore, let $\alpha = (\alpha_1+\ldots+\alpha_6)/6$ and define $\mu_3 \in [0,1]$ by $m_3 = \mu_3\binom{n}{2}$. The theorem follows from the following three claims.

Claim 3.4.9. *We have*

$$\mu_3 \leq \frac{3}{4}\left(45 + \frac{27}{2}\alpha\right)\alpha + \frac{1}{2} + O\left(\frac{1}{n}\right).$$

Claim 3.4.10. *We have*

$$\alpha \leq \left(\frac{2\log(3)}{3\log(4)} - \frac{1}{2}\right)\mu_3 + \frac{1}{4} - \frac{\log(3)}{3\log(4)} + \frac{8}{3}\delta.$$

Claim 3.4.11. *We have* $\alpha \leq 0.016$.

Before proving the claims, we show how they imply the theorem.

Proof of Theorem 3.4.3. Putting together Claim 3.4.9 and Claim 3.4.10 and rearranging the quadratic inequality, we get

$$0 \le \alpha^2 - b\alpha + c\delta$$

where $b = \frac{286\log(2)-180\log(3)}{54\log(3)-81\log(2)} \approx 0.15404$ and $c = 12$ (for large enough n).

By solving for α, we get that either

$$\alpha \le \frac{b - \sqrt{b^2 - 4c\delta}}{2} \qquad \text{or} \qquad \alpha \ge \frac{b + \sqrt{b^2 - 4c\delta}}{2} > 0.1.$$

By Claim 3.4.11, the second case is impossible. A simple calculation shows that $b - \frac{3c\delta}{b} < \sqrt{b^2 - 4c\delta}$ and thus we have $\alpha \le \frac{3c\delta}{2b} < 125\delta$, which implies that $\alpha_i < 750\delta$ for every $i \in [6]$. The rest of the proof is completely analogous to the proof in the 5-colour case. □

It remains to prove Claims 3.4.9, 3.4.10 and 3.4.11.

Proof of Claim 3.4.9. Analogously as in Claim 3.4.6, we derive

$$tr(G_3) \le \left(45 + \frac{27}{2}\alpha\right)\alpha\binom{n}{3} + O(n^2). \tag{3.4.6}$$

The claim again follows from using (3.4.6) and Theorem 3.3.3 in the form (3.4.3). □

Proof of Claim 3.4.10. By (3.4.5) and (3.2.1), we have

$$3^{n^2/4}4^{3n^2/16-2\delta n^2} \le \left|\Phi(\tilde{C}_1, \dots, \tilde{C}_6)\right| \le \prod_{i=1}^{6} i^{m_i}$$

and $\sum_{i\in[6]} i \cdot m_i = \sum_{i\in[6]} e(\tilde{C}_i)$. A simple optimisation shows that $\prod_{i=1}^{6} i^{m_i}$ is maximised under this constraint for fixed m_3 when $m_1 = m_5 = m_6 = 0$ and $2m_2 + 4m_4$ is as large as possible. We conclude that

$$3^{n^2/4}4^{3n^2/16-2\delta n^2} \le \prod_{i\in[6]} i^{m_i} \le 3^{m_3} \cdot 4^{(\sum_{i=1}^{6} e(\tilde{C}_i) - 3m_3)/4} \tag{3.4.7}$$

By taking logarithms we obtain

$$\left(\frac{\log(3)}{4} + \log(4) \cdot \left(\frac{3}{16} - 2\delta\right)\right) n^2 \leq \log(3) \cdot m_3 + \log(4) \cdot \frac{\sum_{i=1}^{6} e(\tilde{C}_i) - 3m_3}{4}$$
$$\leq \left(\log(3) \cdot \frac{\mu_3}{2} + \log(4) \cdot \left(\frac{3}{8} - \frac{3\alpha}{4} - \frac{3\mu_3}{8}\right)\right) n^2 .$$

By cancelling n^2 and rearranging we obtain

$$\frac{3}{4} \cdot \log(4) \cdot \alpha \leq \log(3) \left(\frac{\mu_3}{2} - \frac{1}{4}\right) + \log(4) \left(\frac{3}{16} - \frac{3}{8} \cdot \mu_3 + 2\delta\right),$$

which implies the claim. □

Proof of Claim 3.4.11. Note that (3.4.7) is maximised under the condition

$$\sum_{i\in[6]} m_i = \sum_{i\in[6]} e(\tilde{C}_i)$$

if m_3 is as large as possible, i.e. when

$$3m_3 = \sum_{i\in[6]} e(\tilde{C}_i)$$

or in other words when $\mu_3 = 2(1/2 - \alpha)$. This implies

$$3^{n^2/4} 4^{3n^2/16 - 2\delta n^2} \leq 3^{2(1/2-\alpha)\cdot\binom{n}{2}} \leq 3^{(1/2-\alpha)n^2}.$$

Taking logarithms and solving for α, we get

$$\alpha \leq \frac{1}{4} - \frac{3\log(4)}{16\log(3)} + \frac{2\log(4)}{\log(3)} \cdot \delta \leq 0.01341 + 2.524\delta \leq 0.016.$$

□

3.5 Extremal configurations

In this section we construct families of graphs that satisfy all the conditions of our stability theorems, Theorem 3.1.2 and Theorem 3.1.3, and give a lower bound on the triangle-free colourings these graphs have. These lower bounds are asymptotically close to our upper bounds, proving that our upper bounds are asymptotically sharp with an $e^{o(n^2)}$ error factor.

Recall that we assume that the number of vertices n is divisible by 24.

We define the following matrices

$$A_1 = \begin{pmatrix} 0&0&0&0&0\\ 0&1&1&1&1\\ 1&1&1&0&0\\ 1&0&0&1&1\\ 1&0&1&0&1\\ 1&1&0&1&0\\ 1&0&1&1&0\\ 1&1&0&0&1 \end{pmatrix} A_2 = \begin{pmatrix} 0&0&0&0&0\\ 0&1&1&1&1\\ 1&1&1&0&0\\ 1&0&0&1&1\\ 1&0&1&0&1\\ 1&1&0&1&0\\ 0&1&0&0&1\\ 0&0&1&1&0 \end{pmatrix} A_3 = \begin{pmatrix} 0&0&0&0&0&0\\ 1&1&1&1&0&0\\ 1&1&0&0&1&1\\ 0&0&1&1&1&1\\ 1&0&1&0&1&0\\ 0&1&0&1&1&0\\ 0&1&1&0&0&1\\ 1&0&0&1&0&1 \end{pmatrix}$$

and for a matrix $A \in \{A_1, A_2, A_3\}$ let $\mathcal{E}_A = \{\boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_8\}$ be the set of its row vectors. In these matrices columns represent colours while rows represent disjoint sets of vertices. The details of how exactly follows.

3.5.1 Extremal configurations for five colours

For $A = A_1$ and $a, b, n \in \mathbb{N}$ with $a+b \le n/4$. Let c denote $n/4 - a - b$ and consider a partition of $[n]$ into sets $V(\boldsymbol{\varepsilon}_1) \sqcup \cdots \sqcup V(\boldsymbol{\varepsilon}_8)$ such that

(i) $|V(\boldsymbol{\varepsilon}_1)| = |V(\boldsymbol{\varepsilon}_2)| = n/4$ and

(ii) $|V(\boldsymbol{\varepsilon}_3)| = |V(\boldsymbol{\varepsilon}_4)| = a$ and

(iii) $|V(\boldsymbol{\varepsilon}_5)| = |V(\boldsymbol{\varepsilon}_6)| = b$ and

(iv) $|V(\boldsymbol{\varepsilon}_7)| = |V(\boldsymbol{\varepsilon}_8)| = c$.

Let $G_{A_1}\big(V(\boldsymbol{\varepsilon}_1), \dots, V(\boldsymbol{\varepsilon}_8)\big)$ be the complete 8-partite graph with the partition classes $V(\boldsymbol{\varepsilon}_1), \dots, V(\boldsymbol{\varepsilon}_8)$ and let $\mathcal{B}^{(5)}_{A_1}(n)$ denote the set of all graphs obtained this way. Note that in edge cases these graphs might be 4-partite or 6-partite instead of 8-partite. These graphs are also included in $\mathcal{B}^{(5)}_{A_1}(n)$.

Proposition 3.5.1 (Lower bound for $r = 5$, part 1)**.** *For each $G \in \mathcal{B}^{(5)}_{A_1}(n)$ we have $\varphi_5(G) \ge 6^{n^2/4}$.*

Proof. Let $G = G_{A_1}\big(V(\boldsymbol{\varepsilon}_1), \dots, V(\boldsymbol{\varepsilon}_t)\big)$. Let us define $\pi(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}')$ as the set of colours c for which $\boldsymbol{\varepsilon}_c \ne \boldsymbol{\varepsilon}'_c$. We colour the edges of G according to the pattern $\pi = \pi_{A_1}$, i.e., for any $(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}') \in \mathcal{E}_A^2$ and $uv \in V(\boldsymbol{\varepsilon}) \times V(\boldsymbol{\varepsilon}')$ we colour uv with any of the colours in $\pi(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}')$. For example, an edge between vertices u, v where $u \in \boldsymbol{\varepsilon}_3$ and $v \in \boldsymbol{\varepsilon}_5$ can have colours 2 or 5, as their rows differ at these coordinates.

Each of these colourings is triangle-free as $\pi_A^{-1}(c)$, for any $c \in [5]$, is bipartite. Moreover, the number of such colourings is

$$\prod_{\{\boldsymbol{\varepsilon},\boldsymbol{\varepsilon}'\}\in\binom{\mathcal{E}_A}{2}} |\pi(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}')|^{|V(\boldsymbol{\varepsilon})|\cdot|V(\boldsymbol{\varepsilon}')|} = 3^{\frac{n^2}{4}} \cdot 4^{\frac{n^2}{16}+a^2+b^2+c^2} \cdot 2^{4ab+4ac+4bc}$$

which is easily verified to be $6^{n^2/4}$. □

Note that this lower bound is not sharp, for example we can get further colourings by permuting the colours. This also applies to Proposition 3.5.2 and Proposition 3.5.3.

For $A = A_2$ and $a, b, n \in \mathbb{N}$ with $a, b \leq n/4$. Let c denote $n/4 - a$, d denote $n/4 - b$ and consider a partition of $[n]$ into non-empty sets $V(\boldsymbol{\varepsilon}_1) \sqcup \cdots \sqcup V(\boldsymbol{\varepsilon}_8)$ such that

(i) $|V(\boldsymbol{\varepsilon}_1)| = |V(\boldsymbol{\varepsilon}_2)| = a$ and

(ii) $|V(\boldsymbol{\varepsilon}_3)| = |V(\boldsymbol{\varepsilon}_4)| = b$ and

(iii) $|V(\boldsymbol{\varepsilon}_5)| = |V(\boldsymbol{\varepsilon}_6)| = d$ and

(iv) $|V(\boldsymbol{\varepsilon}_7)| = |V(\boldsymbol{\varepsilon}_8)| = c$.

Let $G_{A_2}(V(\boldsymbol{\varepsilon}_1), \ldots, V(\boldsymbol{\varepsilon}_8))$ be the complete 8-partite graph with the partition classes $V(\boldsymbol{\varepsilon}_1), \ldots, V(\boldsymbol{\varepsilon}_8)$ and let $\mathcal{B}_{A_2}^{(5)}(n)$ denote the set of all graphs obtained this way. Note that in edge cases these graphs might be 4-partite or 6-partite instead of 8-partite. These graphs are also included in $\mathcal{B}_{A_2}^{(5)}(n)$.

Proposition 3.5.2 (Lower bound for $r = 5$, part 2)**.** *For each $G \in \mathcal{B}_{A_2}^{(5)}(n)$ we have $\varphi_5(G) \geq 6^{n^2/4}$.*

The proof is analogous to the proof of Proposition 3.5.1 and therefore omitted.

3.5.2 Extremal configurations for six colours

For $A = A_3$ and $n \in \mathbb{N}$ consider a partition of $[n]$ into non-empty sets $V(\boldsymbol{\varepsilon}_1) \sqcup \cdots \sqcup V(\boldsymbol{\varepsilon}_8)$ such that $|V(\boldsymbol{\varepsilon}_i)| = n/8$ for all $i \in \{1, \ldots, 8\}$. Let $G_{A_3}(V(\boldsymbol{\varepsilon}_1), \ldots, V(\boldsymbol{\varepsilon}_8))$ be the complete 8-partite graph with the partition classes $V(\boldsymbol{\varepsilon}_1), \ldots, V(\boldsymbol{\varepsilon}_8)$ and let $\mathcal{B}_{A_3}^{(6)}(n)$ denote the set of all graphs obtained this way.

Proposition 3.5.3 (Lower bound for $r = 6$)**.** *For each* $G \in \mathcal{B}_{A_3}^{(6)}(n)$ *we have* $\varphi_6(G) \geq 3^{n^2/4}4^{3n^2/16}$.

The proof is analogous to the proof of Proposition 3.5.1 and therefore omitted.

3.6 Stability

In this section we prove Theorem 3.1.3 and Theorem 3.1.2 with the help of Theorem 3.4.3.

We can assume that $V(G) = [n]$, we will do so for convenience. A vertex partition $V^0 \cup V^1 = V(C)$ of the graph C is called a *t-bipartition* if $|E(C) \triangle K(V^0, V^1)| \leq t$, and it is called *balanced* if $|V^0| = |V^1| \pm 1$. Furthermore, for graphs $C_1, \ldots, C_r$ with bipartitions $V_i^0 \cup V_i^1 = [n]$ and $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_r) \in \{0,1\}^r$ we define the partition class

$$V(\boldsymbol{\varepsilon}) = \bigcap_{i \in [r]} V_i^{\varepsilon_i},$$

and set $v(\boldsymbol{\varepsilon}) = |V(\boldsymbol{\varepsilon})|$. Let $\mathrm{dist}(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}')$ denote the Hamming distance of the vectors $\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}' \in \{0,1\}^k$. Moreover, we call a pair $\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}'$ an *m_i-pair (or an m_i-edge)* if their Hamming distance is i. Let us define the following constants, which will be used in both proofs.

$$\gamma = 2 \cdot 10^6 \delta\,, \alpha = 10\sqrt{\gamma}\,, \beta = 10\sqrt{\alpha}\,.$$

3.6.1 Stability proof for five colours

Proof of Theorem 3.1.3. By Theorem 3.2.1 and Corollary 3.2.2 taking $C_1, \ldots, C_5 \in \mathcal{C}(G)$ such that $|\Phi(C_1, \ldots, C_5)|$ is maximal, we have $tr(C_i) < n^{25/9}$ and $|\Phi(C_1, \ldots, C_5)| \geq 6^{n^2/4 - 2\delta n^2}$.

By Theorem 3.4.3 there are balanced γn^2-bipartitions $V_i^0 \cup V_i^1 = [n]$ of C_i, $i \in [5]$, and $X \cup Y = [n]$ of M_3. We consider the partition $V(G) = \bigcup_{\boldsymbol{\varepsilon} \in \{0,1\}^5} V(\boldsymbol{\varepsilon})$ and let

$$\mathcal{E}_X = \{\boldsymbol{\varepsilon} \in \{0,1\}^5 : |V(\boldsymbol{\varepsilon}) \cap X| > \alpha n\}.$$

Let $\mathcal{E}_Y$ be defined analogously. Trivially $V(\mathcal{E}_X) = \bigcup_{\boldsymbol{\varepsilon} \in \mathcal{E}_X} V(\boldsymbol{\varepsilon})$ has size

$$|V(\mathcal{E}_X)| \geq |X| - 2^5 \alpha n \qquad \text{and} \qquad |V(\mathcal{E}_Y)| \geq |Y| - 2^5 \alpha n. \tag{3.6.1}$$

This establishes properties (i) and (ii) for $\mathcal{E}_X$ and $\mathcal{E}_Y$.

Note that $\mathcal{E}_0$ will be chosen to be a subset of $\mathcal{E}_X$ and similarly $\mathcal{E}_1$ will be chosen to be a subset of $\mathcal{E}_Y$. For now we work on establishing the needed properties for $\mathcal{E}_X$ and $\mathcal{E}_Y$. For many of the properties this implies them trivially for subsets, for others we will discuss the changes when we define $\mathcal{E}_0$ and $\mathcal{E}_1$ at the end of the proof.

It will be more convenient to work only with edges of G whose profiles are consistent with the the partition classes containing its end vertices. More precisely, we call $uv \in G$ a *good edge (w.r.t. $\mathcal{E}_X \cup \mathcal{E}_Y$)* if there are $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_5), \boldsymbol{\varepsilon}' = (\varepsilon'_1, \dots, \varepsilon'_5) \in \mathcal{E}_X \cup \mathcal{E}_Y$ such that $u \in V(\boldsymbol{\varepsilon})$ and $v \in V(\boldsymbol{\varepsilon}')$ and

$$\{i \colon uv \in C_i\} = \{i \colon \varepsilon_i \neq \varepsilon'_i\} .$$

In other words, for each $i \in [5]$ with $\varepsilon_i \neq \varepsilon'_i$ the edge uv is contained in the bipartite graph $C_i\big[V_i^{\varepsilon_i}, V_i^{\varepsilon'_i}\big]$ while $uv \notin C_i$ if $\varepsilon_i = \varepsilon'_i$. Note that in this case uv is an m_d-edge where $d = \mathrm{dist}(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}')$. and a good m_d-edge is a witness for the existence of a pair $\{\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}'\} \in \binom{\mathcal{E}_X \cup \mathcal{E}_Y}{2}$ with $d = \mathrm{dist}(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}')$.

As each $V_i^0 \cup V_i^1 = [n]$ is an γn^2-bipartition, there are at most $5\gamma n^2$ pairs of vertices from $V(\boldsymbol{\varepsilon}) \times V(\boldsymbol{\varepsilon}')$, over all $(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}') \in (\mathcal{E}_X \cup \mathcal{E}_Y)^2$, which do not form good edges. Since $X \cup Y$ is an γn^2-partition of M_3 and $m_0, m_1, m_5 \leq 2\gamma n^2$ from Theorem 3.4.3, we derive from the above and $v(\boldsymbol{\varepsilon})v(\boldsymbol{\varepsilon}') > \alpha^2 n^2 > 10\gamma n^2$ that

$$\begin{gathered}\text{each pair } (\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}') \in \mathcal{E}_X \times \mathcal{E}_Y \text{ has distance three and} \\ \text{each pair } (\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}') \in \mathcal{E}_X^2 \cup \mathcal{E}_Y^2 \text{ has distance two or four.}\end{gathered} \tag{3.6.2}$$

This establishes that properties (iii), (iv) and (v) hold.

Let B_i be the bipartite graph induced by the good edges of G on the bipartition $V_i^0 \cup V_i^1$ and let $H = \bigcup_{i \in [5]} B_i$. Moreover, let $N_i = M_i(B_1, \dots, B_5)$, let $n_i = e(N_i)$ and call an edge an n_i-edge if it is contained in N_i. Since all but at most $5\gamma n^2 + 2^6 \alpha n^2$ edges of G are good we have

$$\sum_{i \in [5]} m_i - n_i = e(G) - e(H) \leq 2^7 \alpha n^2 \tag{3.6.3}$$

and, for each $i \in [5]$,

$$|m_i - n_i| \leq 2^7 \alpha n^2. \tag{3.6.4}$$

Furthermore, as each B_i is triangle-free and the restriction of each colouring $\varphi \in \Phi(C_1, \dots, C_5)$ of G to H is a colouring in $\Phi_H(B_1, \dots, B_5)$, we have

$$\prod_{i \in [5]} i^{n_i} = |\Phi(B_1, \dots, B_5)| \geq |\Phi(C_1, \dots, C_5)| 5^{-2^7 \alpha n^2} \geq 6^{\frac{n^2}{4} - 2^7 \alpha n^2}. \tag{3.6.5}$$

We now set out to establish property (vi) and that $|\mathcal{E}| \leq 8$. From Theorem 3.4.3 and (3.6.3) we have

$$n_2 + 2n_4 \geq m_2 + 2m_4 - 2^8 \alpha n^2 \geq \frac{n^2}{4} - 2^9 \alpha n^2. \tag{3.6.6}$$

Due to (3.6.2) we have $n_2(X,Y) + 2n_4(X,Y) = 0$ and as $n_i = n_i(X) + n_i(Y) + n_i(X,Y)$ for each $i \in [5]$, we conclude therefore that either

$$n_2(X) + 2n_4(X) \geq \frac{n^2}{8} - 2^9 \alpha n^2 \qquad \text{or} \qquad n_2(Y) + 2n_4(Y) \geq \frac{n^2}{8} - 2^9 \alpha n^2. \tag{3.6.7}$$

Hence, as $X \cup Y$ is a equipartition, either X or Y will satisfy the presumption of the following claim.

Claim 3.6.1. *If $n_2(X) + 2n_4(X) > \frac{2}{5}|X|^2 + 2^6 \beta n^2$ then there are $\boldsymbol{\varepsilon}_0, \boldsymbol{\varepsilon}_1 \in \mathcal{E}_X$ with Hamming distance four and $|V(\boldsymbol{\varepsilon}_0)|, |V(\boldsymbol{\varepsilon}_1)| > \beta n$. The analogous statement holds for Y.*

Proof. Let K be the largest clique in $N_2[X]$ and consider $\mathcal{K} = \{\boldsymbol{\varepsilon} \colon V(K) \cap V(\boldsymbol{\varepsilon}) \neq \emptyset\} \subseteq \mathcal{E}_X$. By possibly swapping the upper indices of V_i^1 and V_i^0 for some $i \in [5]$ we may assume that $(0,0,0,0,0) \in \mathcal{K}$. Then, by definition of N_2 each of the remaining element in $\mathcal{K}$ has exactly two 1-entries and two of them differ at exactly two entries. This implies that any two elements from $\mathcal{K} \setminus \{(0,0,0,0,0)\}$ coincide in (exactly) one 1-entry, hence, the family corresponds to a 2-uniform intersecting family on a ground set of size five. By the Erdős-Ko-Rado Theorem [23] we know that the set has size at most four. Hence, $|\mathcal{K}| \leq 5$ and Turán's Theorem [55] implies that $N_2[X]$ contains at most $\frac{4}{5}\frac{|X|^2}{2}$ edges. From the presumption on X we then obtain $n_4(X) > 2^5 \beta n^2$. Therefore there must be an n_4-edges in X with ends vertices in partition classes of size larger than βn. This edge is a witness for the desired pair in $\mathcal{E}_X$ with Hamming distance four. □

By possibly renaming the partition classes X and Y suppose that X satisfies (3.6.7). Hence, Claim 3.6.1 provides a pair $\boldsymbol{\varepsilon}_0, \boldsymbol{\varepsilon}_1 \in \mathcal{E}_X$ with Hamming distance four and let $s \in [5]$ be the (unique) coordinate where $\boldsymbol{\varepsilon}_0, \boldsymbol{\varepsilon}_1$ agree. By possibly reordering the C_i's and swapping the upper indices of V_i^1 and V_i^0 for some $i \in [5]$ we may also assume that $\boldsymbol{\varepsilon}_0 = (0,0,0,0,0)$ and $\boldsymbol{\varepsilon}_1 = (1,1,1,1,0)$.

From (3.6.2) we know that the Hamming distances of any $\boldsymbol{\varepsilon} \in \mathcal{E}_Y$ to $\boldsymbol{\varepsilon}_0$ and to $\boldsymbol{\varepsilon}_1$ are both three. Therefore each $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_5) \in \mathcal{E}_Y$ must satisfy $\varepsilon_5 = 1$ since otherwise $(\varepsilon_1, \dots, \varepsilon_4)$ must contain three ones and three zeroes. Hence, $V(\boldsymbol{\varepsilon}) \subseteq Y \cap V_5^1$ for each $\boldsymbol{\varepsilon} \in \mathcal{E}_Y$. Recall from (3.6.1) that $V(\mathcal{E}_Y) \subseteq Y \cap V_5^1$ covers at least $|Y| - 2^5\alpha n$ vertices of $Y \cap V_5^1$, hence all but at most $2^5\alpha n + 1$ vertices of V_5^1 as $|V_5^1| = |Y| \pm 1$. Thus the sizes of those $V(\boldsymbol{\varepsilon})$ with $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_5) \in \mathcal{E}_X$, $\varepsilon_5 = 1$, all together is at most $|V_5^1 \cap X| = |V_5^1 \setminus Y| \le 2^5\alpha n + 1$. With this in mind we set

$$\mathcal{E}'_Y = \mathcal{E}_Y \qquad \text{and} \qquad \mathcal{E}'_X = \{(\varepsilon_1, \dots, \varepsilon_5) \in \mathcal{E}_X : \varepsilon_5 = 0\} \tag{3.6.8}$$

and from (3.6.1) we have $|V(\mathcal{E}'_X)| \ge |V(\mathcal{E}_X)| - 2^5\alpha n - 1 > |X| - 2^7\alpha n$.

Claim 3.6.2. *For each $\tilde{\boldsymbol{\varepsilon}} \in \mathcal{E}'_X$ there is at most one element at distance four to $\tilde{\boldsymbol{\varepsilon}}$ in $\mathcal{E}'_X$. Furthermore, if $v(\tilde{\boldsymbol{\varepsilon}}) > \beta n$ then there is an $\hat{\boldsymbol{\varepsilon}} \in \mathcal{E}'_X$ at distance four to $\tilde{\boldsymbol{\varepsilon}}$ which, moreover, satisfies $v(\tilde{\boldsymbol{\varepsilon}}) = v(\hat{\boldsymbol{\varepsilon}}) \pm \beta n/2$.*

Proof. The first part is clear as $\varepsilon_5 = 0$ for all $(\varepsilon_1, \dots, \varepsilon_5) \in \mathcal{E}'_X$. For a contradiction suppose that there exists $\tilde{\boldsymbol{\varepsilon}} = (\tilde{\varepsilon}_1, \dots, \tilde{\varepsilon}_5) \in \mathcal{E}'_X$ with $v(\tilde{\boldsymbol{\varepsilon}}) > \beta n$ which has Hamming distance two to all $\boldsymbol{\varepsilon} \in \mathcal{E}'_X$, in particular to $\boldsymbol{\varepsilon}_0 = (0,0,0,0,0) \in \mathcal{E}'_X$ and $\boldsymbol{\varepsilon}_1 = (1,1,1,1,0) \in \mathcal{E}'_X$.

Let $A = V(\boldsymbol{\varepsilon}_0)$, $B = V(\boldsymbol{\varepsilon}_1)$ and let $L \cup R = V(\tilde{\boldsymbol{\varepsilon}})$ be an equitable partition of $V(\tilde{\boldsymbol{\varepsilon}})$. We modify H by deleting all vertices not contained in $V(\mathcal{E}'_X \cup \mathcal{E}'_Y)$ and further, deleting the edges between A and L and those between B and R, and adding all possible edges edges between L and R. Let H' denote the so obtained graph. Furthermore, define a partition of $V(\mathcal{E}'_X \cup \mathcal{E}'_Y)$ by

- $U(\boldsymbol{\varepsilon}_0) = A \cup L$ and $U(\boldsymbol{\varepsilon}_1) = B \cup R$,
- $U(\tilde{\boldsymbol{\varepsilon}}) = \emptyset$ and $U(\boldsymbol{\varepsilon}) = V(\boldsymbol{\varepsilon})$ for each $\boldsymbol{\varepsilon} \in (\mathcal{E}'_X \cup \mathcal{E}'_Y) \setminus \{\boldsymbol{\varepsilon}_0, \boldsymbol{\varepsilon}_1, \tilde{\boldsymbol{\varepsilon}}\}$.

Let $U_i^0 \cup U_i^1 = [n]$, $i \in [5]$, be the corresponding bipartition defined by

$$U_i^j = \bigcup_{\substack{\boldsymbol{\varepsilon} \in \mathcal{E}'_X \cup \mathcal{E}'_Y \\ \varepsilon_i = j}} U(\boldsymbol{\varepsilon}), \quad \text{so that} \quad U(\boldsymbol{\varepsilon}) = \bigcap_{i \in [5]} U_i^{\varepsilon_i} \quad \text{for all } \boldsymbol{\varepsilon} \in \mathcal{E}'_X \cup \mathcal{E}'_Y.$$

For each $i \in [5]$ let B'_i be the bipartite graph $H'_i\big[U_i^0 \cup U_i^1\big]$. Further, let

$N'_i = M_i(B'_1, \dots, B'_5)$ and $n'_i = e(N'_i)$. Note that the edges deleted from H either has an end vertex in a partition class $V(\boldsymbol{\varepsilon})$ with $\boldsymbol{\varepsilon} \in \mathcal{E}_X \setminus \mathcal{E}'_X$ or are edges in N_2. On the other hand, the n_2-edges between L and B, those between R and A as well as the new edges between L and R are now edges of N'_4. All remaining edges of H do not change their type, hence

- $n'_3 \geq n_3 - 2^6\alpha n^2$,
- $n'_2 \geq n_2 - n_2(L \cup R, A \cup B) - 2^6\alpha n^2$, and
- $n'_4 \geq n_4 + n_2(L, B) + n_2(R, A) + |L||R| - 2^6\alpha n^2$.

Up to renaming L and R suppose that $n_2(L, B) + n_2(R, A) \geq n_2(L, A) + n_2(R, B)$. Then

$$\begin{aligned} 2n_2(L, B) + 2n_2(R, A) &\geq n_2(L, B) + n_2(R, A) + n_2(L, A) + n_2(R, B) \\ &= n_2(L \cup R, A \cup B)\,, \end{aligned}$$

and hence $n'_2 + 2n'_4 \geq n_2 + 2n_4 + 2|L||R| - 2^8\alpha n^2$. As all the B'_i are triangle-free we obtain from $|L||R| \geq \beta^2 n^2/5$ and (3.6.5) that

$$\begin{aligned} |\Phi(B'_1, \dots, B'_5)| = 2^{n'_2} 3^{n'_3} 4^{n'_4} &\geq 3^{n_3 - 2^6\alpha n^2} 2^{n_2 + 2n_4 + 2|L||R| - 2^8\alpha n^2} \\ &> 6^{\frac{n^2}{4} - 2^9\alpha n^2} 4^{|L||R|} > 6^{\frac{n^2}{4} + \gamma n^2}\,. \end{aligned}$$

This, however, yields a contradiction to the upper bound given by Theorem 3.4.3 and hence, $\tilde{\boldsymbol{\varepsilon}}$ must be at distance four to some $\hat{\boldsymbol{\varepsilon}} \in \mathcal{E}'_X$.

To show that $v(\tilde{\boldsymbol{\varepsilon}}) = v(\hat{\boldsymbol{\varepsilon}}) \pm \beta n/2$ we argue in a similar manner. If, say, $v(\tilde{\boldsymbol{\varepsilon}}) > v(\hat{\boldsymbol{\varepsilon}}) + \beta n$ then we make the two sets equitable by moving vertices from $V(\tilde{\boldsymbol{\varepsilon}})$ to $V(\hat{\boldsymbol{\varepsilon}})$ and changing the edges accordingly. This increases the number of n_4-edges by at least $\beta^2 n^2/20$ which would again yield a contradiction to the upper bound given by Theorem 3.4.3. □

In view of the lemma we choose $\mathcal{E}_0$ to consist of those $\tilde{\boldsymbol{\varepsilon}} \in \mathcal{E}'_X$ with $v(\tilde{\boldsymbol{\varepsilon}}) > \beta n$ together with the unique corresponding $\hat{\boldsymbol{\varepsilon}} \in \mathcal{E}'_X$ at distance four. Hence, $V(\mathcal{E}_0)$ covers all but at most $2^5\beta n$ vertices of X. Further, note that the projections of $\boldsymbol{\varepsilon} \in \mathcal{E}_0$ and of $\boldsymbol{\varepsilon}' \in \mathcal{E}'_Y$ to the first four coordinates yields vectors with distance two, while the projections of $\boldsymbol{\varepsilon}$ and $\boldsymbol{\varepsilon}'$ with $\{\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}'\} \in \binom{\mathcal{E}_0}{2} \cup \binom{\mathcal{E}'_Y}{2}$ yields vectors with distance two or four. Hence, $|\mathcal{E}_0 \cup \mathcal{E}'_Y| \leq 8$ and as $\mathcal{E}'_Y \neq \emptyset$ we have $|\mathcal{E}_0| \in \{2, 4, 6\}$. Note that if $\mathcal{E}_0$ has size two then the largest value of $n_2(X) + 2n_4(X)$ is achieved by a balanced complete bipartite graph. In this case $n_2(X) = 0$ and $n_4(X) \leq \frac{|X|^2}{4}$. If

$\mathcal{E}_0$ has size four then the optimum is achieved by any four partite graph consisting of two pairs of equitable parts, say the first pair of sizes a and the second with sizes b and $b \pm 1$ with $2a + 2b = |X| \pm 1$. In this case, $n_2(X) \leq 4ab + 2a$ and $n_4(X) \leq a^2 + b^2 + b$, hence, $n_2(X) + 2n_4(X) \leq 2(a+b)^2 + 2(a+b)$. For $\mathcal{E}_X$ of size six the optimum is achieved by any six partite graph consisting of three pairs of equitable parts, say the first of sizes a, the second of sizes b and the third of sizes c and $c \pm 1$ with $2a+2b+2c = |X| \pm 1$. In this case, $n_2(X) \leq 4ab+4ac+4bc+2a+2b$ and $n_4(X) \leq a^2+2b^2+2c^2+c$, hence, $n_2(X)+2n_4(X) \leq 2(a+b+c)^2+2(a+b+c)$.

Hence, in any of these cases, we have $n_2(X)+2n_4(X) \leq \frac{|X|^2}{2}+5|X|$ and by (3.6.6) we derive that $n_2(Y) + 2n_4(Y) > \frac{2}{5}|Y|^2 + 2^6\beta^2 n^2$. By Claim 3.6.1 there is a pair of elements at distance four in $\mathcal{E}'_Y$ and the same argument as in Claim 3.6.2 shows that for each $\tilde{\boldsymbol{\varepsilon}} \in \mathcal{E}'_Y$ with $v(\tilde{\boldsymbol{\varepsilon}}) > \beta n$ there is an $\hat{\boldsymbol{\varepsilon}} \in \mathcal{E}'_Y$ at distance four to $\tilde{\boldsymbol{\varepsilon}}$ and $v(\tilde{\boldsymbol{\varepsilon}}) = v(\hat{\boldsymbol{\varepsilon}}) \pm \beta n/2$. We choose $\mathcal{E}_1$ to consist of those $\boldsymbol{\varepsilon} \in \mathcal{E}'_Y$ with $v(\boldsymbol{\varepsilon}) > \beta n$ together with the unique corresponding $\hat{\boldsymbol{\varepsilon}} \in \mathcal{E}'_Y$ at distance four. This establishes property (vi).

This yields the desired families $\mathcal{E}_0$ and $\mathcal{E}_1$. Since we only lost vertex groups of size at most βn and at most 2^5 of them, properties (i) and (ii) still hold. □

3.6.2 Stability proof for six colours

By Theorem 3.2.1 and Corollary 3.2.2 taking $C_1, \ldots, C_5 \in C(G)$ such that $|\Phi(C_1, \ldots, C_5)|$ is maximal, we have $tr(C_i) < n^{25/9}$ and $|\Phi(C_1, \ldots, C_5)| \geq 6^{n^2/4 - 2\delta n^2}$.

Proof of Theorem 3.1.2. By Theorem 3.2.1 and Corollary 3.2.2 taking $C_1, \ldots, C_5 \in C(G)$ such that $|\Phi(C_1, \ldots, C_6)|$ is maximal, we have $tr(C_i) < n^{25/9}$ and $|\Phi(C_1, \ldots, C_6)| \geq 3^{n^2} 4^{3n^2/16 - 2\delta n^2}$.

By Theorem 3.4.3 there are balanced γn^2-bipartitions $V_i^0 \cup V_i^1 = [n]$ of C_i, $i \in [6]$, and we consider the partition $V(G) = \bigcup_{\boldsymbol{\varepsilon} \in \{0,1\}^6} V(\boldsymbol{\varepsilon})$. Consider

$$\mathcal{E}_0 = \mathcal{E}_X = \{\boldsymbol{\varepsilon} \in \{0,1\}^6 \colon |V(\boldsymbol{\varepsilon}) \cap X| > \alpha n\}$$

and let $\mathcal{E}_1 = \mathcal{E}_Y$ be defined analogously. Trivially $V(\mathcal{E}_X) = \bigcup_{\boldsymbol{\varepsilon} \in \mathcal{E}_X} V(\boldsymbol{\varepsilon})$ has size

$$|V(\mathcal{E}_X)| \geq |X| - 2^6 \alpha n \qquad \text{and} \qquad |V(\mathcal{E}_Y)| \geq |Y| - 2^6 \alpha n. \tag{3.6.9}$$

This establishes properties (i) and (ii).

It will be more convenient to work only with edges of G whose profiles are consistent with the the partition classes containing its end vertices. More precisely, we call $uv \in G$ a *good edge (w.r.t. $\mathcal{E}_X \cup \mathcal{E}_Y$)* if there are $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_6), \boldsymbol{\varepsilon}' = (\varepsilon_1', \ldots, \varepsilon_6') \in \mathcal{E}_X \cup \mathcal{E}_Y$ such that $u \in V(\boldsymbol{\varepsilon})$ and $v \in V(\boldsymbol{\varepsilon}')$ and $\{i \colon uv \in C_i\} = \{i \colon \varepsilon_i \neq \varepsilon_i'\}$. In other words, for each $i \in [6]$ with $\varepsilon_i \neq \varepsilon_i'$ the edge uv is contained in the bipartite graph $C_i\big[V_i^{\varepsilon_i}, V_i^{\varepsilon_i'}\big]$ while $uv \notin C_i$ if $\varepsilon_i = \varepsilon_i'$. Note that in this case uv is an m_d-edge where $d = \mathrm{dist}(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}')$ and a good m_d-edge is a witness for the existence of a pair $\{\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}'\} \in \binom{\mathcal{E}_X \cup \mathcal{E}_Y}{2}$ with $d = \mathrm{dist}(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}')$.

As each $V_i^0 \cup V_i^1 = [n]$ is an γn^2-bipartition, there are at most $6\gamma n^2$ pairs of vertices from $V(\boldsymbol{\varepsilon}) \times V(\boldsymbol{\varepsilon}')$, over all $\{\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}'\} \in \binom{\mathcal{E}_X \cup \mathcal{E}_Y}{2}$, which do not form good edges. Since $X \cup Y$ is an γn^2-bipartition of M_3 and $m_0, m_1, m_5, m_6 \leq \gamma n^2$, by Theorem 3.4.3, we derive from the above and $v(\boldsymbol{\varepsilon}) v(\boldsymbol{\varepsilon}') > \alpha^2 n^2 > \gamma n^2 + 3\gamma n^2 + 6\gamma n^2 > 10\gamma n^2$ that

$$\begin{gathered} \text{each pair } (\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}') \in \mathcal{E}_X \times \mathcal{E}_Y \text{ has distance three and} \\ \text{each pair } (\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}') \in \mathcal{E}_X^2 \cup \mathcal{E}_Y^2 \text{ has distance two or four.} \end{gathered} \tag{3.6.10}$$

This establishes that properties (iii) and (iv) hold.

Let B_i be the bipartite graph induced by the good edges of G on the bipartition

$V_i^0 \cup V_i^1$ and let $H = \bigcup_{i\in[6]} B_i$. Moreover, let $N_i = M_i(B_1, \ldots, B_6)$, let $n_i = e(N_i)$ and call an edge an n_i-edge if it is contained in N_i. Since all but at most $6\gamma n^2 + 2^6\alpha n^2$ edges of G are good we have

$$\sum_{i\in[5]} m_i - n_i = e(G) - e(H) \le 6\gamma n^2 + 2^6\alpha n^2 \le 2^7\alpha n^2 \tag{3.6.11}$$

and for each $i \in [6]$

$$|m_i - n_i| \le 2^7\alpha n^2. \tag{3.6.12}$$

Furthermore, as each B_i is triangle-free and the restriction of each colouring $\varphi \in \Phi(C_1, \ldots, C_5)$ of G to H is a colouring in $\Phi_H(B_1, \ldots, B_5)$, we have

$$\prod_{i\in[5]} i^{n_i} = |\Phi(B_1, \ldots, B_5)| \ge |\Phi(C_1, \ldots, C_5)| 6^{-2^7\alpha n^2} \ge 3^{n^2/4} 4^{3n^2/16} 6^{-2^8\alpha n^2}. \tag{3.6.13}$$

Due to (3.6.10) and (3.6.11), we have $n_3(X,Y) = n_3(H) \ge e(M_3) - 2^7\alpha n^2 \ge n^2/4 - 2^8\alpha n^2$.

Our next goal is to prove $|\mathcal{E}| \le 8$, the rest of the properties will easily follow.

Claim 3.6.3. $|\mathcal{E}_X|, |\mathcal{E}_Y| \ge 4$

Proof. We first note that for each $i \in [6]$

$$|N_3(X,Y) \cap E(B_i)| \ge \frac{n^2}{8} - 2^9\alpha n^2 \ge \frac{n^2}{8} - 2^9\alpha n^2.$$

Also, we have that

$$\sum_{i\in[6]} |N_3(X,Y) \cap E(B_i)| = |\{(e,i) : e \in N_3(X,Y) \cap E(B_i)\}| \le 3n_3(X,Y) \le 3(n^2/4 + \gamma n^2)\,.$$

Hence, for each $i \in [6]$ we have that $|N_3(X,Y) \cap E(B_i)| \le \frac{n^2}{8} + \frac{1}{6}\gamma n^2 \le \frac{n^2}{8} + 2^9\alpha n^2$. Altogether, for each $i \in [6]$,

$$|N_3(X,Y) \cap E(B_i)| = \frac{n^2}{8} \pm 2^9\alpha n^2\,.$$

Therefore, for each $i \in [6]$, the following holds

$$|X \cap V_i^0| = |X \cap V_i^1| = |Y \cap V_i^0| = |Y \cap V_i^1| = \frac{n}{4} \pm 2^5 \alpha n\,, \tag{3.6.14}$$

and hence, $|\mathcal{E}_X|, |\mathcal{E}_Y| \geq 2$.

From now on in this proof, for a subset $U \subseteq V(G)$, we use the notation $\tilde{U} = X \cap U$. Suppose that $2 \leq |\mathcal{E}_X| \leq 3$. Then, for some ordering $\boldsymbol{\varepsilon}_1, \dots \boldsymbol{\varepsilon}_{2^6}$ of the elements of $\{0,1\}^6$, we have that $|\tilde{V}(\boldsymbol{\varepsilon}_i)| \leq \alpha n$ for each $i \in \{4, \dots, 2^6\}$ and therefore

$$\left| \bigcup_{i \geq 4} \tilde{V}(\boldsymbol{\varepsilon}_i) \right| \leq 2^6 \alpha n. \tag{3.6.15}$$

Since $2^6 \alpha n < \frac{n}{4} \pm 2^5 \alpha n$, due to (3.6.14) and (3.6.15), for each $i \in [6]$, there are distinct $k_i, t_i \in \{1,2,3\}$ such that

$$\tilde{V}(\boldsymbol{\varepsilon}_{k_i}) \subseteq \tilde{V}_i^1 \;\text{ and }\; \tilde{V}(\boldsymbol{\varepsilon}_{t_i}) \subseteq \tilde{V}_i^0$$

Hence, we can assume that for a fixed $i \in [6]$, by possibly interchanging V_i^1,V_i^0, that there distinct $k, \ell, t \in \{1,2,3\}$ such that

$$(\tilde{V}(\boldsymbol{\varepsilon}_k) \cup \tilde{V}(\boldsymbol{\varepsilon}_\ell)) \subseteq \tilde{V}_i^1 \;\text{ and }\; \tilde{V}(\boldsymbol{\varepsilon}_t) \;\subseteq \tilde{V}_i^0. \tag{3.6.16}$$

Moreover, since $\boldsymbol{\varepsilon}_k \neq \boldsymbol{\varepsilon}_l$ there is some $j \in [6]$ with $j \neq i$ such that

$$\tilde{V}(\boldsymbol{\varepsilon}_k) \subseteq \cap V_j^1 \;\text{ and }\; (\tilde{V}(\boldsymbol{\varepsilon}_t) \cup \tilde{V}(\boldsymbol{\varepsilon}_\ell)) \subseteq \tilde{V}_j^0, \tag{3.6.17}$$

by possibly interchanging V_j^1,V_j^0, and $\boldsymbol{\varepsilon}_k$, $\boldsymbol{\varepsilon}_\ell$.

Due to (3.6.15), (3.6.16) and (3.6.17), we have that $|\tilde{V}_i^0| \leq |\tilde{V}(\boldsymbol{\varepsilon}_t)| + 2^6 \alpha n$ and $|\tilde{V}_j^1| \leq |\tilde{V}(\boldsymbol{\varepsilon}_k)| + 2^6 \alpha n$. Hence, using (3.6.14), we obtain that

$$|\tilde{V}(\boldsymbol{\varepsilon}_t)| \geq n/4 - 2^7 \alpha n \quad \text{and} \quad |\tilde{V}(\boldsymbol{\varepsilon}_k)| \geq n/4 - 2^7 \alpha n. \tag{3.6.18}$$

Furthermore, by (3.6.14) and (3.6.16), we have $|\tilde{V}(\boldsymbol{\varepsilon}_k)| + |\tilde{V}(\boldsymbol{\varepsilon}_l)| \leq |\tilde{V}_i^0| \leq n/4 + 2^5 \alpha n$, which combined with (3.6.18) yields

$$|\tilde{V}(\boldsymbol{\varepsilon}_l)| \leq 2^8 \alpha n.$$

Now, due to (3.6.14), (3.6.15), the previous inequality and the fact that $2^9 \alpha n <$

$\frac{n}{4} \pm 2^5\alpha n$, we have that for each $i \in [6]$, $k_i, t_i \in \{k, t\}$, and $\boldsymbol{\varepsilon}_k$, $\boldsymbol{\varepsilon}_t$ are at distance 6. Due to (3.6.18), $\boldsymbol{\varepsilon}_k, \boldsymbol{\varepsilon}_t \in \mathcal{E}_X$, and we have obtained a contradiction to 3.6.10.

□

From Theorem 3.4.3 and (3.6.11) we have

$$n_2 + 2n_4 \geq m_2 + 2m_4 - 2^7\alpha n^2 \geq \frac{3n^2}{8} - 2^8\alpha n^2. \tag{3.6.19}$$

It follows from (3.6.10) that we have $n_2(X,Y) + 2n_4(X,Y) = 0$. Since $n_i = n_i(X) + n_i(Y) + n_i(X,Y)$ for each $i \in [6]$, we conclude therefore that either

$$n_2(X) + 2n_4(X) \geq \frac{3n^2}{16} - 2^7\alpha n^2 \qquad \text{or} \qquad n_2(Y) + 2n_4(Y) \geq \frac{3n^2}{16} - 2^7\alpha n^2. \tag{3.6.20}$$

Hence, as $X \cup Y$ is a equipartition, either X or Y will satisfy the presumption of the following claim.

Claim 3.6.4. *If $n_2(X) + 2n_4(X) > \frac{5}{12}|X|^2 + 2^6\alpha n^2$ then there are $\boldsymbol{\varepsilon}_0, \boldsymbol{\varepsilon}_1 \in \mathcal{E}_X$ with Hamming distance four and $|V(\boldsymbol{\varepsilon}_0)|, |V(\boldsymbol{\varepsilon}_1)| > \alpha n$. The analogous statement holds for Y.*

Proof. Let K be the largest clique in $N_2[X]$ and consider $\mathcal{K} = \{\boldsymbol{\varepsilon} \colon V(K) \cap V(\boldsymbol{\varepsilon}) \neq \emptyset\} \subseteq \mathcal{E}_X$. By possibly swapping the upper indices of V_i^1 and V_i^0 for some $i \in [6]$ we may assume that $(0,0,0,0,0,0) \in \mathcal{K}$. Then, by definition of N_2 each of the remaining element in $\mathcal{K}$ has exactly two 1-entries and two of them differ at exactly two entries. This implies that any two elements from $\mathcal{K} \setminus \{(0,0,0,0,0,0)\}$ coincide in (exactly) one 1-entry, hence, the family corresponds to a 2-uniform intersecting family on a ground set of size six. By the Erdős-Ko-Rado Theorem [23] we know that the set has size at most five. Hence, $|\mathcal{K}| \leq 6$ and Turán's Theorem [55] implies that $N_2[X]$ contains at most $\frac{5}{12}|X|^2$ edges. From the presumption on X we then obtain $n_4(X) > 2^5\alpha n^2$. Therefore there must be an n_4-edge in X with end vertices in partition classes of size larger than αn. This edge is a witness for the desired pair in $\mathcal{E}_X$ with Hamming distance four. □

By possibly renaming the partition classes X and Y suppose that X satisfies (3.6.20). Hence, the claim provides a pair $\boldsymbol{\varepsilon}_0, \boldsymbol{\varepsilon}_1 \in \mathcal{E}_X$ with Hamming distance four. By possibly reordering the C_i's and swapping the upper indices of V_i^1 and V_i^0 for some $i \in [6]$ we assume that $\boldsymbol{\varepsilon}_0 = (0,0,0,0,0,0)$ and $\boldsymbol{\varepsilon}_1 = (1,1,1,1,0,0)$.

We now define set $\mathcal{E}'_X$. If the last coordinate of $\boldsymbol{\varepsilon} \in \mathcal{E}_X$ is 0, then $\boldsymbol{\varepsilon} \in \mathcal{E}'_X$ and if the last coordinate of $\boldsymbol{\varepsilon} \in \mathcal{E}_X$ is 1, then $\overline{\boldsymbol{\varepsilon}} \in \mathcal{E}'_X$, where $\boldsymbol{\varepsilon} + \overline{\boldsymbol{\varepsilon}} = (1,1,1,1,1,1)$. The set $\mathcal{E}'_Y$ is defined analogously. Since no pair $\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}'$ of elements in $\mathcal{E}_X \cup \mathcal{E}_Y$ is at distance 6, we have that $|\mathcal{E}'_X| = |\mathcal{E}_X|$, $|\mathcal{E}'_Y| = |\mathcal{E}_Y|$ and

$$|\mathcal{E}'_X \cup \mathcal{E}'_Y| = |\mathcal{E}_X \cup \mathcal{E}_Y| \,. \tag{3.6.21}$$

Moreover, due to (3.6.10), we have that

$$\begin{gathered}\text{each pair } (\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}') \in \mathcal{E}'_X \times \mathcal{E}'_Y \text{ has distance three and} \\ \text{each pair } (\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}') \in (\mathcal{E}'_X)^2 \cup (\mathcal{E}'_Y)^2 \text{ has distance two or four.}\end{gathered} \tag{3.6.22}$$

Moreover, each $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_6) \in \mathcal{E}'_X \cup \mathcal{E}'_Y$ satisfies $\varepsilon_6 = 0$. From (3.6.22) we know that the Hamming distances of any $\boldsymbol{\varepsilon} \in \mathcal{E}'_Y$ to $\boldsymbol{\varepsilon}_0$ and to $\boldsymbol{\varepsilon}_1$ are both three. Therefore each $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_6) \in \mathcal{E}'_Y$ must satisfy $\varepsilon_5 = 1$ since otherwise $(\varepsilon_1, \ldots, \varepsilon_4)$ must contain three ones and three zeroes. Also, the following claim holds.

Claim 3.6.5. *Each* $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_6) \in \mathcal{E}'_X$ *satisfies* $\varepsilon_5 = 0$.

Proof. For the sake of contradiction suppose that there is $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_6) \in \mathcal{E}'_X$ with $\varepsilon_5 = 1$. Due to (3.6.22), $\boldsymbol{\varepsilon}$ is at distance 2 or 4 from $\boldsymbol{\varepsilon}_0$ and $\boldsymbol{\varepsilon}_1$ and hence, there is $i \in \{1,2,3,4\}$ such that $\varepsilon_i \neq \varepsilon_j \in \{1,2,3,4\} - \{i\}$. Again due to (3.6.22), the distance between $\boldsymbol{\varepsilon}$ and $\boldsymbol{\varepsilon}' = (\varepsilon'_1, \ldots, \varepsilon'_6) \in \mathcal{E}'_X$ is three, and since $\varepsilon_5 = \varepsilon'_5$ and $\varepsilon_6 = \varepsilon'_6$, we must have that $(\varepsilon_1, \ldots, \varepsilon_4)$ and $(\varepsilon'_1, \ldots, \varepsilon'_4)$ are at distance 3. Moreover, $(\varepsilon'_1, \ldots, \varepsilon'_4)$ is at distance 2 with $(0,0,0,0)$ and $(1,1,1,1)$, which implies that $(\varepsilon'_1, \ldots, \varepsilon'_4)$ has exactly two 1-entries. It implies that $\varepsilon_i = \varepsilon'_i$ and there are at most $\binom{3}{2} = 3$ elements in $\mathcal{E}'_Y$, a contradiction to Claim 3.6.3. ◻

Claim 3.6.5 implies that $|\mathcal{E}'_X \cup \mathcal{E}'_Y| \leq 8$. Moreover, $|\mathcal{E}_X \cup \mathcal{E}_Y| \leq 8$ by (3.6.21) and due to Claim 3.6.3, we have $|\mathcal{E}| = |\mathcal{E}_X \cup \mathcal{E}_Y| = 8$.

We now need to show that each pair $(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}') \in \mathcal{E}_X^2 \cup \mathcal{E}_Y^2$ has distance four. But this is clear from the fact that it can be realized. This establishes property (v).

Finally we show that for each $(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}') \in (\mathcal{E}_X \cup \mathcal{E}_y)^2$ we have $v(\boldsymbol{\varepsilon}) = v(\boldsymbol{\varepsilon}') \pm \beta n$. This proof is analogous to the proof of (vi) in Theorem 3.1.3. Since X, Y is an equipartition we only need to consider pairs $(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}') \in \mathcal{E}_X^2 \cup \mathcal{E}_Y^2$. If, say, $v(\boldsymbol{\varepsilon}) > v(\boldsymbol{\varepsilon}') + \beta n$ then we make the two sets equitable by moving vertices from $V(\boldsymbol{\varepsilon})$ to $V(\boldsymbol{\varepsilon}')$ and changing the edges accordingly. This increases the number of n_4-edges by at least $\beta^2 n^2/5$ which would yield a contradiction to the upper bound given by Theorem 3.4.3. This establishes property (vi).

□

4

Packing Degenerate Hypergraphs

4.1 Introduction

Recall, in the entirety of Chapter 4 we will consider hypergraphs that are r-*uniform*, that is each edge of a hypergraph G is an r element subset of $V(G)$. We will refer to $(r-1)$-element vertex sets as *semi-edges*.

A *packing* of a family $\mathcal{G} = \{G_1, \ldots, G_k\}$ of hypergraphs into a hypergraph H is a colouring of the edges of H with the colours $0, 1, \ldots, k$ such that the edges of colour i form an isomorphic copy of G_i for each $1 \le i \le k$. The packing is *perfect* if no edges have colour 0. We will often say an edge is *covered* in a packing if it has colour at least 1, and *uncovered* if it has colour zero.

In our work we consider D-degenerate hypergraphs. First, let us define the term. An ordering of $V(G)$ is D-degenerate if for each vertex v, there are at most D edges of G whose final vertex is v. We say G is D-degenerate if $V(G)$ has a D-degenerate ordering. In particular, we define *trees* as connected 1-degenerate hypergraphs.

Our results

In this section we restate Theorem 1.3.5 and Theorem 1.3.6 in stronger, more technical forms. We will use these forms for the remainder of the chapter.

To state our more technical results, we need the following concept of quasirandomness.

Definition 4.1.1 (density, quasirandom)**.** The *density* of a hypergraph H is the number $e(H)/\binom{v(H)}{r}$. Suppose that H is a hypergraph with n vertices and with density p. We say that H is (α, L)-*quasirandom* if for every set $S \subseteq \binom{V(H)}{r-1}$ of size at most L, the number of vertices v which form an edge with every member of S is $|\mathsf{N}_H(S)| = (1 \pm \alpha)p^{|S|}n$.

Our first main technical result then says that we can approximately pack any sufficiently quasirandom H with hypergraphs of bounded degeneracy and not too large maximum degree. Note that in particular $K_n^{(r)}$ is sufficiently quasirandom.

Theorem 4.1.2. *For each $r \geq 2$, each $\gamma > 0$ and each $D \in \mathbb{N}$ there exist $\xi, c > 0$ and a number n_0 such that the following holds for each integer $n > n_0$. Suppose that H is any n-vertex hypergraph which is $(\xi, 4(r+1)rD+3)$-quasirandom. Suppose that $(G_t)_{t\in[t^*]}$ is a family of D-degenerate r-uniform hypergraphs, each of which has at most n vertices and maximum degree at most $\frac{cn}{\log n}$. Suppose further that the total number of edges of $(G_t)_{t\in[t^*]}$ is at most $e(H) - \gamma n^r$. Then $(G_t)_{t\in[t^*]}$ packs into H, and furthermore the hypergraph of leftover edges is $(\gamma, (r+1)rD+3)$-quasirandom.*

Our second main technical result is a bit more complicated. If we insist that a small fraction of the hypergraphs G_t each have linearly many vertices of degree 1, and in addition those graphs are not too close to spanning, then we can upgrade 'covering almost all the edges' to a perfect packing. Specifically, we can perfectly pack a collection of hypergraphs defined as follows.

Definition 4.1.3 ((μ, n)-sequence)**.** We say that a sequence $(G_i)_{i\in[m]}$ of hypergraphs is a *D-degenerate (μ, n)-hypergraph sequence with maximum degree* Δ if

(G 1) G_i is D-degenerate and for each $v \in V(G_i)$ we have $\deg(v) \leq \Delta$ for each $i \in [m]$,

(G 2) $v(G_i) \leq n$ for each $1 \leq i \leq m - \lfloor \mu n^{r-1} \rfloor$, and

(G 3) $v(G_i) \leq n - \mu n$ and G_i has at least μn leaves for each i with $m - \lfloor \mu n^{r-1} \rfloor < i \leq m$.

We also call the G_i with $m - \lfloor \mu n^{r-1} \rfloor < i \leq m$ the *special* hypergraphs of the sequence.

Theorem 4.1.4. *For each $r \geq 3$, every D and $\mu > 0$ there are n_0 and $\xi, c > 0$ such that for every $n \geq n_0$ and every m, the following holds. Suppose that H is any n-vertex hypergraph which is $(\xi, (r+1)rD+3)$-quasirandom with density at least μ. Every D-degenerate (μ, n)-hypergraph sequence $(G_i)_{i\in[m]}$ with maximum degree $\frac{cn}{\log n}$ such that $\sum_{i\in m} e(G_i) \leq e(H)$ packs into H.*

For future use, we will state a technical version of Theorem 4.1.2 in which we not only pack into a quasirandom host hypergraph and obtain at the end a quasirandom hypergraph of leftover edges, but also obtain a collection of additional quasirandomness conditions on the packing as a whole. Although we do consider this theorem to be a main result of this paper, since it is rather more complicated to state, we defer it to Section 4.7, where it is Theorem 4.7.1. We note that it is this technical version of Theorem 4.1.2 which we need to prove Theorem 4.1.4.

We now briefly discuss the optimality of our conditions in Theorems 4.1.2 and 4.1.4. First, we observe that for packing into quasirandom hypergraphs, the maximum degree bound is optimal up to the value of c, even if we want only to embed one single spanning hypergraph. To see this, consider H a typical n-vertex hypergraph generated by selecting r-sets as edges independently with probability $\frac{1}{2}$, and G an n-vertex hypergraph which is the disjoint union of $\frac{1}{10}\log n$ stars (that is, hypergraphs consisting of an $(r-1)$-set centre which forms an edge with each one of the remaining vertices, and no further edges) of equal size. This G is 1-degenerate and has maximum degree less than $20r\frac{n}{\log n}$, but a standard calculation (essentially the same as for the 2-uniform graph case) shows that G is unlikely to be contained in H, even though H is very quasirandom.

Our proofs do allow D to grow with n; we believe (but did not check carefully) the dependency is roughly as $\log\log\log n$, but this is presumably not optimal. On the other hand, D cannot be as big as $10\log n$, since a typical random hypergraph is unlikely to contain any given hypergraph with $9n\log n$ edges.

We cannot allow all hypergraphs to be spanning in Theorem 4.1.4. For example, if H is obtained by choosing r-sets as edges independently with probability either $\frac{1}{2}$ or $\frac{1}{2}+\varepsilon$, depending on whether the r-set contains one of the first $n/2$ vertices or not, then H will likely be (ξ, Q)-quasirandom for any given $\xi > 0$ and natural number Q, provided $\varepsilon > 0$ is sufficiently small. However, some $\Theta(n)$ vertices will be in $\Theta(n^{r-1})$ more edges than others, so that we certainly cannot pack any collection of regular spanning hypergraphs. In fact, if for example each G_s is a matching, then we need that $\Theta(n^{r-1})$ of the matchings are $\Theta(n)$ vertices away from being spanning to correct the imbalance in degrees. This shows that the restriction on the size of the hypergraphs in $(G3)$ is optimal up to the choice of constants. However, we do not believe that it is necessary to have many hypergraphs with many leaves. We should note that one cannot simply omit this condition, because for example for $r = 2$ no collection of cycles can perfectly pack K_{2n}, due to a parity obstruction: cycles use an even number of edges at each vertex, but K_{2n} has odd degree vertices. However for the case $D = 1$ (i.e. forests) we believe one can omit the condition entirely (as the leaves should allow for parity correction). This is demonstrated to be true in [1] for graphs.

Proof outline

To pack a collection of **g**uest hypergraphs $\mathcal{G}$ into a **h**ost hypergraph H, in which each guest hypergraph has order at most $(1-\delta)n$, we do the following. We take hypergraphs in $\mathcal{G}$ in succession. For each G, we embed vertex by vertex into $K_n^{(r)}$ in a degeneracy order, at each time embedding to a vertex of $K_n^{(r)}$ chosen uniformly at random subject to the constraints that we do not re-use a vertex previously used in embedding G, or an edge used in embedding a previous hypergraph. This procedure succeeds with high probability, and after each stage of embedding a hypergraph, the unused edges in $K_n^{(r)}$ are quasirandom (in a sense we will later make precise). The analysis of this process is what gives the quasirandomness statements of Theorem 4.7.1.

To allow for spanning hypergraphs, we modify this slightly. We adjust the degeneracy order so that (for some small $\delta > 0$) the last δn vertices are strongly independent (that is, no pair of them is contained in an edge of G) and all have the same degree. This can be done while at worst multiplying the degeneracy of the order by r. Then for each hypergraph we follow the above procedure to embed the first $(1-\delta)n$ vertices, and finally complete the embedding arbitrarily using a matching argument. We will see that this last step is with high probability always possible. The only point to be careful about is that we have to split $E(K_n^{(r)})$ into a very dense *bulk* main part, whose edges we use only for the embedding of the first $(1-\delta)n$ vertices, and a sparse *reservoir* which we use only for the completion; we do this randomly. This is the hypergraph analogue of the main result of [3].

To obtain a perfect packing, we perform a different modification. We use the above procedure to pack all the non-special guest graphs of our (μ, n)-hypergraph sequence, and observe that the resulting leftover edges form a quasirandom graph H. We then need to pack the special graphs. We first remove from each special graph a strongly independent set of vertices of degree 1: the *omitted leaves*. We pack the resulting collection of graphs into H, using the randomised algorithm described above. What now remains is, for each guest graph, to pack in addition the omitted leaves. By definition, each omitted leaf is in exactly one edge of the guest graph, together with its *parent* semi-edge, and all the vertices of the parent semi-edge have been packed. This reduces our remaining packing to a matching problem: each semi-edge u of $K_n^{(r)}$ has some *dangling leaves*, namely the collection of omitted leaves (over all special guest graphs) whose parent semi-edge is embedded to u. We need to match the dangling leaves at u to the vertices which form leftover edges of $K_n^{(r)}$ with u, such that we do not use any edge of H twice and such that we do not match dangling leaves of any graph G_i with vertices which were used for embedding the rest of G_i, or for another dangling leaf of G_i.

It is relatively easy to ensure that we do not use leftover edges multiple times. We *orient* the edges of H by marking one vertex in each edge, and we insist on matching dangling leaves at $u \in S(H)$ to an edge containing u where none of the vertices in u are marked. This reduces the problem to ensuring that we pick distinct edges of H to embed the dangling leaves at u. Note, that matching a dangling leave to an edge means embedding the dangling leaf (which is a vertex in one of the graphs we want to pack) to the marked vertex of the edge we matched to. Therefore, we need to ensure that nothing is embedded to that vertex in the host graph, from the relevant guest graph. In order for this to work, we of course need that the number of dangling leaves at u is equal to the number of edges containing u whose marked vertex is not in u. We can obtain this by performing the marking uniformly at random and then correcting the resulting small errors.

We then go through all the $\binom{n}{r-1}$ semi-edges of H in turn, and at each semi-edge u embed all its dangling leaves. We do this as follows. We draw a bipartite *leaf matching graph* whose parts are the dangling leaves at u and the marked vertices of edges containing u, putting an edge from a dangling leaf of a guest graph G_i to a marked vertex if the marked vertex has not previously been used in the embedding of G_i. We choose uniformly at random a perfect matching in the leaf matching graph, and embed the dangling leaves accordingly. The difficulty with this — and the reason we need to pick a matching uniformly at random — is that when we embed dangling leaves at one semi-edge, we need to remove a few edges from the leaf matching graphs of other semi-edges. We will argue that the leaf matching graphs begin with a strong quasirandom property and are only changed slightly through the entire process, so that they are still quasirandom at the end; this is enough to guarantee the desired matchings always exist and hence complete a perfect packing. This second part is the hypergraph analogue of the main result of [2].

Organisation of the chapter

In the following Section 4.2, we introduce notation, the probabilistic tools we need, and state and prove a result on random matchings in bipartite graphs.

In Section 4.3, we reduce Theorem 4.1.2 to a technical statement, Theorem 4.3.1 which is convenient for our proof, formally state the algorithm sketched above which proves it (i.e. which produces a packing), state the main lemmas which we need to prove it, and finally give the proof of Theorem 4.3.1 from these lemmas.

In Section 4.4, we analyse the embedding of the first $(1-\delta)n$ vertices of a single guest hypergraph into (what remains of) the bulk of the host hypergraph, and argue that our randomised embedding is likely to succeed and maintain various desirable properties. We build on this in Section 4.5 to analyse the sequential embedding of the first $(1-\delta)n$ vertices of all guest hypergraphs into the bulk. Finally in Section 4.6 we argue that it is always possible to complete the embedding in the reservoir, which finishes the proof of Theorem 4.1.2.

In Section 4.7 we state and prove Theorem 4.7.1, the above mentioned technical version of Theorem 4.1.2 which also gives various quasirandomness properties of the packing. Most of the work here is done in the preceding sections, so this section mainly consists of using established probabilistic estimates to prove the desired quasirandomness properties are likely to hold.

In Section 4.8 we prove Theorem 4.1.4. Again, most of the work is done by Theorems 4.1.2 and 4.7.1, and what remains to do is formalise the above sketched packing algorithm and prove that it works with high probability.

4.2 Preliminaries

4.2.1 Notation

We will almost always work with r-uniform hypergraphs; when we write 'hypergraph' without specifying the uniformity, it should be understood that the hypergraph considered is r-uniform.

We refer to vertex sets of size $r-1$ as *semi-edges*. Let $S(G)$ denote the set of semi-edges of a hypergraph G, that is $S(G) = \binom{V(G)}{r-1}$, the set of unordered $r-1$-subsets of the vertex set. Given a vertex set $A \subseteq V(G)$, we write $\mathsf{N}_G(A)$ for the *neighbourhood* of A, that is, the collection of subsets B of $V(G) \setminus A$ such that $A \cup B \in E(G)$. We will often write $\mathsf{N}_G(v)$ for the set $\mathsf{N}_G(\{v\})$ of semi-edges. Note that the neighbourhood of A in a hypergraph is a $(r-|A|)$-uniform hypergraph on

$V(G) \setminus A$. However, when A is a semi-edge, we will usually think of $\mathsf{N}_G(A)$ as a set of vertices in $V(G) \setminus A$ (even though formally it is a set of singleton sets of vertices). Finally, if S is any subset of $\binom{V(G)}{k}$, we write $\mathsf{N}_G(S)$ for the intersection $\bigcap_{A \in S} \mathsf{N}_G(A)$.

We call a set of vertices $A \subseteq V(G)$ *strongly independent* if for any pair of different vertices $u, v \in A$ we have $\mathsf{N}_G(\{u, v\}) = \emptyset$. For any $k \leq r-1$ we call a hypergraph G a *k-star with centre A* if $|A| = k$ and for each $e \in E(G)$ we have $A \subseteq e$. The *leaves* of the star G are the vertices $\{v \in V(G) : \{v\} \cup A \in E(G)\}$.

For any given natural number ℓ, we write $[\ell] := \{1, 2, \ldots, \ell\}$. The definition of degenerate graphs naturally suggests to label the vertices of a hypergraph by integers. Suppose that the vertex set of a hypergraph G is $V(G) = [\ell]$ and $i \in V(G)$. We write $\mathsf{N}^-(i) = \mathsf{N}(i) \cap \binom{[i-1]}{r-1}$ and $\deg^-(i) = |\mathsf{N}^-(i)|$ for the *left-neighbourhood* and the *left-degree* of i. We make use of the natural order on $[\ell]$ also in other ways, like referring to sets of the form $[\ell_1] \subseteq V(G)$ and $\{\ell_2, \ell_2 + 1, \ldots, \ell\} \subseteq V(G)$ as *initial vertices* and *final vertices*, respectively. We write $R(v) := \{u : \exists y \in \mathsf{N}_G^-(v) \text{ with } u = \max(y)\}$. That is, $R(v)$ denotes the set of vertices which are the second-to-last vertices in hyperedges of G whose last vertex is v.

The hypergraphs to be packed are denoted G_s and referred to as **g**uest hypergraphs. By contrast, during our packing procedure, we shall work with **h**ost hypergraphs H_s which are obtained from the original $K_n^{(r)}$ by removing what was used previously.

An *orientation* of a hypergraph $H = (V, E)$ is an oriented hypergraph $\vec{H}$ on V which contains, for each undirected edge $e \in E$, exactly one *directed edge* (v, e) where v is a vertex in e. We say the edge e is *directed towards* v. The *outdegree* $\deg^+_{\vec{H}}(s)$ of a semi-edge s in an oriented hypergraph $\vec{H}$ is the number of vertices v in $V(\vec{H})$ such that $\big(v, s \cup \{v\}\big)$ is an edge of $\vec{H}$; the set of these vertices v is the *outneighbourhood* $N^+_{\vec{H}}(s)$ of s. The *indegree* $\deg^-_{\vec{H}}(v)$ of a vertex v in an oriented hypergraph $\vec{H}$ is the number of s in $S(H)$ such that $\big(v, s \cup \{v\}\big)$ is an edge of $\vec{H}$; the set of these hyperedges s is the *inneighbourhood* $N^-_{\vec{H}}(v)$ of v.

We write $0 < a \ll b$ to mean that given $b > 0$, any sufficiently small a will make our calculations work. In other words, there is a monotone increasing function $f\colon \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ with $f(b) \leq b$ such that any choice $0 < a \leq f(b)$ will work. We define longer strings similarly; thus $0 \ll a \ll b \ll c$ means that $0 \ll b \ll c$ and in addition $0 \ll a \ll b$, where the two monotone functions need not be the same. We can always find constants satisfying such a sequence by choosing from the right-hand side: in the example above we would choose first c, then a sufficiently small b, and given that a sufficiently small a.

4.2.2 Probabilistic tools

In this subsection we give the notation and tools we need to analyse our randomised processes. These are taken directly from [2, Section 2.2] and we refer the reader there for exposition and proofs.

Let Ω be a finite probability space. A *filtration* $\mathcal{F}_0, \mathcal{F}_1, \ldots, \mathcal{F}_n$ is a sequence of partitions of Ω such that $\mathcal{F}_i$ refines $\mathcal{F}_{i-1}$ for all $i \in [n]$. In our application, the partition $\mathcal{F}_i$ is given by all possible histories of the run of one of our algorithms up to time i. (For more explanation see [3].) We say that a function $f : \Omega \to \mathbb{R}$ is *$\mathcal{F}_i$-measurable* if f is constant on each part of $\mathcal{F}_i$. Further, for any random variable $Y : \Omega \to \mathbb{R}$ the *conditional expectation* $\mathbb{E}(Y|\mathcal{F}_i) : \Omega \to \mathbb{R}$ and the *conditional variance* $\mathrm{Var}(Y|\mathcal{F}_i) : \Omega \to \mathbb{R}$ of Y with respect to $\mathcal{F}_i$ are defined by

$$\begin{aligned} \mathbb{E}(Y|\mathcal{F}_i)(x) &= \mathbb{E}(Y|X), \\ \mathrm{Var}(Y|\mathcal{F}_i)(x) &= \mathrm{Var}(Y|X), \end{aligned} \qquad \text{where } X \in \mathcal{F}_i \text{ is such that } x \in X\,.$$

Suppose that we have an algorithm which proceeds in m rounds using a new source of randomness Ω_i in each round i. Then the probability space underlying the run of the algorithm is $\prod_{i=1}^m \Omega_i$. By *history up to time t* we mean a set of the form $\{\omega_1\} \times \cdots \times \{\omega_t\} \times \Omega_{t+1} \times \cdots \Omega_m$, where $\omega_i \in \Omega_i$. We shall use the symbol $\mathscr{H}_t$ to denote any particular history of such a form. By a *history ensemble up to time t* we mean any union of histories up to time t; we shall use the symbol $\mathscr{L}$ to denote any one such. Observe that there are natural filtrations associated to such a probability space: given times $t_1 < t_2 < \ldots$ we let $\mathcal{F}_{t_i}$ denote the partition of Ω into the histories up to time t_i.

Our first tool is the well known Chernoff bound about the sum of independent Bernoulli variables.

Theorem 4.2.1 (Chernoff bounds, [30, Theorem 2.10]). *Suppose X is a random variable which is the sum of a collection of independent Bernoulli random variables. Then we have for $\delta \in (0, 3/2)$*

$$\mathbb{P}\big[X > (1+\delta)\mathbb{E}X\big] < e^{-\delta^2 \mathbb{E}X/3} \quad \textit{and} \quad \mathbb{P}\big[X < (1-\delta)\mathbb{E}X\big] < e^{-\delta^2 \mathbb{E}X/3}\,.$$

When analysing randomised algorithms, usually one has to deal with a sum of random variables which are not independent, but rather are *sequentially dependent*. When we use this term, what we mean is that there is some filtration $\mathcal{F}_0$, $\mathcal{F}_1, \ldots,$ $\mathcal{F}_n$ associated to the randomised algorithm such that the random variable X_i is $\mathcal{F}_i$-measurable for each i. The filtration in question will generally be a filtration defined by histories up to a sequence of increasing times in the algorithm; when it is clear from the context we will simply say that the random variables are sequentially dependent without specifying the filtration explicitly. In this chapter, we need the following concentration inequality for sequentially dependent random variables, which is a corollary of Freedman's inequality [24] deduced in [2].

Corollary 4.2.2 (Corollary 6, [2])**.** *Let Ω be a finite probability space, and $(\mathcal{F}_0, \mathcal{F}_1, \ldots, \mathcal{F}_n)$ be a filtration. Suppose that we have $R > 0$, and for each $1 \leq i \leq n$ we have an $\mathcal{F}_i$-measurable non-negative random variable Y_i, nonnegative real numbers $\tilde{\mu}, \tilde{\nu}$ and an event $\mathcal{E}$.*

(a) Suppose that either $\mathcal{E}$ does not occur or we have $\sum_{i=1}^{n} \mathbb{E}\big[Y_i \big| \mathcal{F}_{i-1}\big] \leq \tilde{\mu}$, and $0 \leq Y_i \leq R$ for each $1 \leq i \leq n$. Then

$$\mathbb{P}\left[\mathcal{E} \text{ and } \sum_{i=1}^{n} Y_i > 2\tilde{\mu}\right] \leq \exp\left(-\frac{\tilde{\mu}}{4R}\right).$$

(b) Suppose that either $\mathcal{E}$ does not occur or we have $\sum_{i=1}^{n} \mathbb{E}\big[Y_i \big| \mathcal{F}_{i-1}\big] = \tilde{\mu} \pm \tilde{\nu}$, and $0 \leq Y_i \leq R$ for each $1 \leq i \leq n$. Then for each $\tilde{\varrho} > 0$ we have

$$\mathbb{P}\left[\mathcal{E} \text{ and } \sum_{i=1}^{n} Y_i \neq \tilde{\mu} \pm (\tilde{\nu} + \tilde{\varrho})\right] \leq 2\exp\left(-\frac{\tilde{\varrho}^2}{2R(\tilde{\mu} + \tilde{\nu} + \tilde{\varrho})}\right).$$

In particular, if $\tilde{\nu} = \tilde{\varrho} = \tilde{\mu}\tilde{\eta} > 0$ and $\tilde{\eta} \leq \frac{1}{2}$, then

$$\mathbb{P}\left[\mathcal{E} \text{ and } \sum_{i=1}^{n} Y_i \neq \tilde{\mu}(1 \pm 2\tilde{\eta})\right] \leq 2\exp\left(-\frac{\tilde{\mu}\tilde{\eta}^2}{4R}\right).$$

Finally, let us note that we shall be using many statements of the form

$$\text{with probability at least } p\text{, provided event } \mathcal{A} \text{ we get event } \mathcal{B}. \tag{4.2.1}$$

We emphasize that such statements are not statements about conditional probabilities. That is, the meaning of (4.2.1) is $\mathbb{P}[\mathcal{A} \setminus \mathcal{B}] \leq 1 - p$. A prototypical example is *with probability at least* $1 - o(1)$*, if a given randomized algorithm does not fail,* *then it produces an output with certain desired properties.*

4.2.3 Degenerate hypergraphs

We need to bound $\sum_{x\in V(G)} \deg(x)^2$ for degenerate hypergraphs G.

Lemma 4.2.3. *Let G be an n-vertex hypergraph with degeneracy D and maximum degree Δ. Then we have*

$$\sum_{x\in V(G)} \deg(x)^2 \leq rDn\Delta\,.$$

Proof. Suppose that for each $1 \leq i \leq \Delta$ there are w_i vertices in G of degree i. We have $\sum_{i=1}^{\Delta} iw_i = \sum_{x\in V(G)} \deg(x) = re(G) \leq rDn$. By convexity of the function $z \mapsto z^2$, among all choices of non-negative reals w_i such that $\sum_{i=1}^{\Delta} iw_i \leq 2Dn$, the one maximising $\sum_{i=1}^{\Delta} i^2 w_i$ is $w_1 = \cdots = w_{\Delta-1} = 0$ and $w_\Delta = rDn/\Delta$. The maximum attained is $rDn\Delta$, proving the lemma. □

We also need to show that degenerate hypergraphs contain large strongly independent sets all of whose vertices have the same degree.

Lemma 4.2.4. *Let G be a D-degenerate n-vertex hypergraph. Then there exists an integer $0 \leq d \leq rD$ and a set $I \subseteq V(G)$ with $|I| \geq (rD+1)^{-4}n$ which is independent, and all of whose vertices have the same degree d in G.*

Proof. We first claim that at least $(rD+1)^{-1}n$ vertices of G have degree at most rD. Indeed, if this were false then there would be more than $rDn/(rD+1)$ vertices of G all of whose degrees are at least $rD+1$, so that we obtain $e(G) > Dn$, which contradicts D-degeneracy of G. Let $0 \leq d \leq rD$ be chosen to maximise the number of vertices in G of degree d, and let S be the set of vertices in G with degree d. We thus have $|S| \geq (rD+1)^{-2}n$. Now let I be a maximal independent subset of S. Each vertex of I has at most $d \leq rD$ incident edges, so that the number of vertices of G which are either in I or in an edge containing a vertex of I is at most $(r-1)rD|I| + |I| \leq (rD+1)^2|I|$. By maximality this set of vertices covers S, hence $|I| \geq (rD+1)^{-2}|S| \geq (rD+1)^{-4}n$, as desired. □

The following auxiliary lemma takes an arbitrary family of graphs we want to pack and produces a family with not too many members and the same bound on maximum degree and degeneracy.

Lemma 4.2.5 (compression lemma)**.** *Let $\mathcal{G} = (G_i)_{i\in[s]}$ be a family of D-degenerate hypergraphs with maximum degree at most Δ, with $\sum_{i=1}^{s} e(G_i) \le \binom{n}{r}$ and $v(G_i) \le n$ for all $i \in [s]$. Then there is a family of hypergraphs $\mathcal{G}' = (\check{G}_i)_{i\in[s^*]}$ with $s^* \le \frac{2n^{r-1}}{(r-1)!}$ such that $\sum_{i=1}^{s^*} e(\check{G}_i) \le \binom{n}{r}$, such that for each $i \in [s^*]$ we have $v(\check{G}_i) \le n$, $\Delta(\check{G}_i) \le \Delta$, and $\check{G}_i$ is D-degenerate, and such that $\mathcal{G}$ perfectly packs into $\mathcal{G}'$.*

Proof. We successively modify the family $\mathcal{G}$ as follows. If there are two hypergraphs $G, G' \in \mathcal{G}$ with $v(G), v(G') \le n/2$, we replace G and G' with the disjoint union $G \cup G'$. We repeat this until no further such pairs exist, giving $\mathcal{G}'$.

Observe that the maximum degree and the degeneracy of the hypergraphs in $\mathcal{G}$ is the same as in $\mathcal{G}'$. Furthermore $\mathcal{G}'$ is a packing of $\mathcal{G}$. Finally, there is at most one hypergraph in $\mathcal{G}'$ with less than $n/2$ vertices. Hence all but at most one hypergraph has at least $n/(2r)$ edges. We conclude that the total number s^* of hypergraphs in $\mathcal{G}'$ satisfies $(s^* - 1)n/(2r) \le (1-\gamma)\binom{n}{r}$, and hence $s^* \le 2n^{r-1}/(r-1)!$. Finally, we let the hypergraphs $(G'_s)_{s=1}^{s^*}$ be obtained from the hypergraphs $\mathcal{G}'$ by adding if necessary isolated vertices to each in order to obtain n-vertex hypergraphs. □

4.2.4 Random matchings in bipartite graphs

There are two places in this chapter where we reduce 'completing' embeddings to a perfect matching problem in bipartite graphs. In both cases, we can show that the bipartite graph in question is suitably quasirandom, and this not only implies that a perfect matching exists, it also implies that a uniform random perfect matching is about equally likely to use any given edge (which, in both cases, we need). The following statement, which is rather similar to [2, Lemma 20], formalises this. The proof is very similar to that given in [2], but we provide it for completeness.

Lemma 4.2.6 (matching lemma)**.** *Assume $0 \ll \frac{1}{m} \ll \varepsilon, \varrho \ll \mu \ll 1$. Let $F = F[U,W]$ be a bipartite graph with $|U| = |W| = (1\pm\varepsilon)m$ such that*

(M 1) $\deg_F(x) = (1\pm\varepsilon)\mu m$ *for all $x \in U \cup W$, and*

(M 2) $\deg_F(u,u') = (1\pm\varepsilon)\mu^2 m$ *for all but at most $\frac{m^2}{\log m}$ pairs $\{u,u'\} \in \binom{U}{2}$,*

and let $F' = F'[U,W]$ be a spanning subgraph of $F[U,W]$ such that

(M 3) $\deg_F(x) - \deg_{F'}(x) < \varrho m$ *for all $x \in U \cup W$.*

Then F' has a perfect matching and for a perfect matching σ chosen uniformly at random among all perfect matchings in F' and for all $uw \in E(F')$ we have

$$\mathbb{P}[\sigma(u) = w] \le \frac{2}{\mu m}\,.$$

This lemma is a straightforward consequence of a lemma (Lemma 4.2.9) on random matchings in super-regular pairs by Felix Joos (see [37]) and the degree-codegree characterisation of super-regular pairs (Lemma 4.2.8) provided by Duke, Lefmann, and Rödl in [16]. The proof of this lemma is the only place in this chapter where we use the concept of a regular pair.

Definition 4.2.7 (density, (ε, d)-regular, (ε, d)-super-regular)**.** Let G be a graph and $U, W \subseteq V(G)$ be disjoint vertex sets. The *density* of the bipartite graph $G[U, W]$ is

$$d_G(U, W) = \frac{e(G[U, W])}{|U||W|}\,.$$

We say that $G[U, W]$ is (ε, d)*-regular* if for all $U' \subseteq U$ and $W' \subseteq W$ with $|U'| \geq \varepsilon|U|$ and $|W'| \geq \varepsilon|W|$ we have

$$d_G(U', W') = d \pm \varepsilon\,.$$

The graph $G[U, W]$ is (ε, d)*-super-regular* if it is (ε, d)-regular and for all $u \in U$ and for all $w \in W$ we have

$$\deg_{G[U,W]}(u) = (d \pm \varepsilon)|W|, \qquad \text{and} \qquad \deg_{G[U,W]}(w) = (d \pm \varepsilon)|U|\,.$$

It is well-known that regular pairs are forced by a degree-codegree condition; we use the following formulation due to Duke, Lefmann, and Rödl in [16].

Lemma 4.2.8 (degree-codegree condition [16])**.** *Assume* $0 < \varepsilon' < 2^{-200}$ *and let* $G[U, W]$ *be a bipartite graph with parts* U *and* W *of size* $|U| = |W| = n$ *and density* $d = d_{G[U,W]}(U, W)$*. If*

(i) $\deg_{G[U,W]}(u) > (d - \varepsilon')|W|$ *for all* $u \in U$*, and*

(ii) $\deg_{G[U,W]}(u, u') < (d + \varepsilon')^2|W|$ *for all but at most* $2\varepsilon'|U|^2$ *pairs* $\{u, u'\} \in \binom{U}{2}$*,*

then $G[U, W]$ *is* $\big((\varepsilon')^{\frac{1}{6}}, d\big)$*-regular.* □

If we choose a perfect matching uniformly at random in a super-regular pair then each edge is roughly equally likely to appear in the matching, as was shown by Joos (see [37]).

Lemma 4.2.9 (perfect matchings in super-regular pairs [37, Theorem 4.3])**.** *Assume* $0 \ll \frac{1}{m'} \ll \varepsilon' \ll d \ll 1$*. Let* $G[U, W]$ *be an* (ε'', d)*-super-regular graph with*

$|U| = |W| = m'$. *Then* $G[U,W]$ *contains a perfect matching. Moreover, for a perfect matching* σ *chosen uniformly at random among all perfect matchings in* $G[U,W]$ *and for all* $uw \in E(G)$ *we have*

$$\mathbb{P}[\sigma(u) = w] = (1 \pm (\varepsilon'')^{\frac{1}{20}})\frac{1}{dm'} \,. \qquad \square$$

The proof of the matching lemma simply combines these two lemmas.

Proof of Lemma 4.2.6. By $(M1)$ and $(M3)$, for all $x \in U \cup W$ we have

$$\deg_{F'}(x) = \big(\mu \pm (\varrho + \varepsilon)\big)m \tag{4.2.2}$$

By $(M2)$ and $(M3)$, for all but at most $\frac{m^2}{\log m}$ pairs $\{u, u'\} \in \binom{U}{2}$ we have

$$\deg_{F'}(u, u') = \big(\mu^2 \pm (2\varrho + \varepsilon)\big)m \,. \tag{4.2.3}$$

We want to apply Lemma 4.2.8 with $d = d_{F'}(U, W) = \mu \pm (\varrho + \varepsilon)$ and $\varepsilon' = 5(\varrho + \varepsilon)\mu^{-1}$ to conclude that $F'[U, W]$ is super-regular, and now check the conditions of this lemma. By (4.2.2), for $u \in U$ we have

$$\deg_{F'}(u) = \big(d \pm 2(\varrho+\varepsilon)\big)m = \big(d \pm 2(\varrho+\varepsilon)\big)\frac{|W|}{1 \pm \varepsilon} = \big(d \pm 4(\varrho+\varepsilon)\big)|W| > (d-\varepsilon')|W| \,,$$

and similarly for $w \in W$ we have $\deg_{F'}(w) = \big(d \pm 4(\varrho + \varepsilon)\big)|U| > (d - \varepsilon')|U|$. By (4.2.3), for all but at most $\frac{m^2}{\log m}$ pairs $\{u, u'\} \in \binom{U}{2}$ we have

$$\begin{aligned}\deg_{F'}(u, u') &\le \big(d^2 + (2\varrho + \varepsilon)\big)m \le \big(d^2 + (2\varrho + \varepsilon)\big)\frac{|W|}{1 - \varepsilon} \le \big(d^2 + 4(\varrho + \varepsilon)\big)|W| \\ &< \big(d^2 + 2\varepsilon' d + (\varepsilon')^2\big)|W| = (d + \varepsilon')^2|W| \,,\end{aligned}$$

where the last inequality uses $d = \mu \pm (\varrho + \varepsilon)$ and $\varepsilon' = 5(\varrho + \varepsilon)\mu^{-1}$. We conclude that, if $\frac{m^2}{\log m} \le 2\varepsilon'|U|^2$, which holds for $\log m > 10/\varepsilon'$, then F' is $\big((\varepsilon')^{\frac{1}{6}}, d\big)$-regular by Lemma 4.2.8. Since $\deg_{F'}(x) = \big(d \pm 2(\varrho + \varepsilon)\big)|U|$ for all $x \in U \cup W$, it follows that F' is $\big((\varepsilon')^{\frac{1}{6}}, d\big)$-super-regular.

Hence we can apply Lemma 4.2.9 to F' with

$$m' = |U| = (1 \pm p)m \,, \quad \varepsilon'' = (\varepsilon')^{\frac{1}{6}} \,, \quad \text{and} \quad d = \mu \pm (\varrho + \varepsilon) \,,$$

and conclude that F' has a perfect matching and that for a perfect matching σ chosen uniformly at random among all perfect matchings of F' and for all $uw \in E(F')$ we have

$$\mathbb{P}[\sigma(u) = w] = \left(1 \pm (\varepsilon')^{\frac{1}{120}}\right)\frac{1}{d(1 \pm \varepsilon)m} \leq \frac{2}{\mu m},$$

as required. □

4.3 Almost perfect hypergraph packing

In this section we give a randomised algorithm which almost perfectly packs almost spanning hypergraphs, and a modification which allows for almost perfect packing of spanning hypergraphs. We state the main lemmas which show that the randomised algorithm is likely to succeed, and assuming them prove Theorem 4.1.2.

4.3.1 Reducing the first main theorem

We deduce Theorem 4.1.2 from the following technical result.

Theorem 4.3.1 (Approximate packing technical result)**.** *For each $r \geq 2$, $\gamma > 0$ and each $D \in \mathbb{N}$ there exist numbers $n_0 \in \mathbb{N}$ and $c, \xi > 0$ such that the following holds for each $n > n_0$. Suppose that $\widehat{H}$ is an $(\xi, (r+1)D+3)$-quasirandom hypergraph with n vertices and density $p > 0$. Suppose that $s^* \leq 2n^{r-1}/(r-1)!$. Suppose that for each $s \in [s^*]$ the hypergraph G_s is a hypergraph on vertex set $[n]$, with maximum degree at most $\frac{cn}{\log n}$, such that $\deg^-(x) \leq D$ for each $x \in V(G_s)$ and such that the last $(D+1)^{-4}n$ vertices of $[n]$ form a strongly independent set in G_s, and all have the same degree d_s in G_s.*

Suppose further that the total number of edges of $(G_s)_{s\in[s^*]}$ is at most $(p-3\gamma)\binom{n}{r}$. Then $(G_s)_{s\in[s^*]}$ packs into $\widehat{H}$. In addition, the hypergraph of leftover edges is $(\gamma, (r+1)D+3)$-quasirandom.

We briefly explain how to deduce Theorem 4.1.2 from this result.

Proof of Theorem 4.1.2. Given $(G_t)_{t\in t^*}$ to pack, we create a sequence $(G'_s)_{s\in[s^*]}$ by applying Lemma 4.2.5 to the graphs $(G_t)_{t\in[t^*]}$. By Lemma 4.2.5, we have the required $s^* \leq 2n^{r-1}/(r-1)!$.

Now, for each G'_s with $s \leq s^*$ we choose an order on $V(G'_s)$ as follows. First, we pick an order witnessing D-degeneracy of G'_s. Next, we pick an integer $0 \leq d_s \leq 2D$ and a strongly independent I_s set of $(rD+1)^{-4}n$ vertices each of which has degree d_s in G'_s and change the order by moving these vertices to the end. Such an integer d_s and strongly independent set exist by Lemma 4.2.4. Observe that moving vertices to the end of the order cannot increase the left-degree of vertices which are not moved, which therefore have at most D left-neighbours in the new order. The moved vertices have degree at most rD, all of which are left-neighbours. The result is an ordering of $V(G'_s)$ with degeneracy at most rD, as required for Theorem 4.3.1 with input $\gamma/3$ and rD. Then Theorem 4.3.1 returns the desired packing. □

Proof of Theorem 4.3.1

For the proof of Theorem 4.3.1, we need some algorithms and definitions. We give these now along with a sketch of the proof. At this level, the proof is very similar to that in [3]. We will do a bit of extra analysis (as compared to the proof in [3]) in order to obtain additional properties of the packing (in order to later prove Theorem 4.7.1), and the analysis itself will need some modifications to deal with hypergraphs.

We prove Theorem 4.3.1 by analysing a randomised algorithm, which we call *PackingProcess*, that packs the guest hypergraphs G_s into $\widehat{H}$. We prove that this algorithm succeeds with high probability. In this algorithm we assume that the last δn vertices of each hypergraph G_s form an strongly independent set, where $\delta < (D+1)^{-4}$ is to be chosen later.

PackingProcess begins by splitting the edges of the input hypergraph $\widehat{H}$ into a *bulk* H_0 and a *reservoir* H_0^* by independently selecting edges into the latter with probability chosen such that $e(H_0^*) \approx \gamma\binom{n}{r}$. As a result, the hypergraphs H_0 and H_0^* are with high probability quasirandom.

Now *PackingProcess* proceeds in s^* stages. In each stage s, it runs a randomised embedding algorithm, called *RandomEmbedding* and explained below, to embed the first $n - \delta n$ vertices of G_s into the bulk H_{s-1}. Then in the *completion phase* the last δn vertices of G_s are embedded into the reservoir H_{s-1}^*. Since there are exactly δn vertices of G_s left to embed and exactly δn vertices of $V(\widehat{H})$ unused so far in this stage, we want to find a bijection between these. Since all neighbours of each yet unembedded vertex are already embedded, this completion amounts to choosing a system of distinct representatives. The completion phase relies on choosing a random matching in a super-regular bipartite graph. Now H_s and H_s^* are defined simply by removing the edges used in this embedding.

Both *RandomEmbedding* and the completion phase may *fail* at any stage s;

this means that it is not possible to embed a certain part of G_s. In that case *PackingProcess* fails, too. If *PackingProcess* does not fail then it always produces a valid packing of (G_s) into H. So, we need to show that *PackingProcess* (see Algorithm 1) succeeds with positive probability.

Algorithm 1: *PackingProcess*

Input : hypergraphs $G_1, \dots, G_{s^*}$, with G_s on vertex set $[n]$ such that the last δn vertices of G_s form a strongly independent set; a hypergraph $\widehat{H}$ on n vertices
choose H_0^* by picking edges of $\widehat{H}$ independently with probability $\gamma\binom{n}{r}/e(\widehat{H})$;
let $H_0 = \widehat{H} - H_0^*$;
for $s = 1$ **to** s^* **do**
  run *RandomEmbedding*(G_s,H_{s-1}) to get an embedding φ_s of $G_s[[n-\delta n]]$ into H_{s-1};
  let H_s be the hypergraph obtained from H_{s-1} by removing the edges of $\varphi_s(G_s[[n-\delta n]])$;
  choose uniformly at random an extension φ_s^* of φ_s embedding all of $V(G_s)$, and embedding $E(G_s) \setminus E(G_s[[n-\delta n]])$ into $E(H_{s-1}^*)$;
  let H_s^* be the hypergraph obtained from H_{s-1}^* by removing the edges $\varphi_s^*(E(G_s) \setminus E(G_s[[n-\delta n]]))$;
end

For describing our randomised embedding algorithm *RandomEmbedding* we need the following definitions. We shall use the symbol $\hookrightarrow$ to denote embeddings produced by *RandomEmbedding*. We write $G \hookrightarrow H$ to indicate that the hypergraph G is to be embedded into H. Also, if $t \in V(G)$, $v \in V(H)$ and $A \subseteq V(H)$ then $t \hookrightarrow v$ means that t is embedded on v, and $t \hookrightarrow A$ means that t is embedded on a vertex of A. If $\vec{t} \in V(G)^k$, $\vec{v} \in V(H)^k$ ordered vertex sets then $\vec{t} \hookrightarrow \vec{v}$ means that $\vec{t}$ is embedded on $\vec{v}$ in the given order. If ψ is an embedding $V(G) \to V(H)$ and $A \subseteq V(G)$ is a vertex set then we write $\psi(A) = \{\psi(v) | v \in A\}$ for the image of A. Similarly if $\mathcal{A} \subseteq \mathcal{P}(V(G))$ is a system of vertex sets then we write $\psi(\mathcal{A}) = \{\psi(A) | A \in \mathcal{A}\}$.

Definition 4.3.2 (partial embedding, candidate set)**.** Let G be a hypergraph with vertex set $[v(G)]$, and H be a hypergraph with $v(H) \geq v(G)$. Further, assume $\psi_j \colon [j] \to V(H)$ is a *partial embedding* of G into H for $j \in [v(H)]$, that is, ψ_j is a hypergraph embedding of $G[[j]]$ into H. Finally, let $t \in [v(H)]$ be such that

$\mathsf{N}^-(t) \subseteq P([j])$. Then the *candidate set of* t (with respect to ψ_j) is

$$C^j_{G\hookrightarrow H}(t) = \mathsf{N}_H\Big(\psi_j\big(\mathsf{N}^-_G(t)\big)\Big)\,.$$

RandomEmbedding (see Algorithm 2) randomly embeds a guest hypergraph G into a host hypergraph H. The algorithm is simple: we iteratively embed the first $(1-\delta)n$ vertices of G randomly to one of the vertices of their candidate set which was not used for embedding another vertex already.

Algorithm 2: *RandomEmbedding*

Input : hypergraphs G and H, with $V(G) = [v(G)]$ and $v(H) = n$
$\psi_0 := \emptyset$;
$t^* := \min(v(G), (1-\delta)n)$;
for $t = 1$ **to** t^* **do**
  if $C^{t-1}_{G\hookrightarrow H}(t) \setminus \mathrm{im}(\psi_{t-1}) = \emptyset$ **then** halt with failure;
  choose $v \in C^{t-1}_{G\hookrightarrow H}(t) \setminus \mathrm{im}(\psi_{t-1})$ uniformly at random;
  $\psi_t := \psi_{t-1} \cup \{t \hookrightarrow v\}$;
end
return ψ_{t^*}

To show that *PackingProcess* does not fail at any stage, we shall show that the host hypergraph H_s constructed in *PackingProcess* in embedding stage s is quasirandom in the sense of Definition 4.1.1. In fact, in order to analyse the completion phase of *PackingProcess* we need quasirandomness of the pair (H_s, H^*_0), where H^*_0 is the initial reservoir. We now define this *coquasirandomness* of a pair of hypergraphs. Recall that quasirandomness of one hypergraph means that common neighbourhoods of semi-edges are always about the size one would expect in a random hypergraph of a similar density. Coquasirandomness of two hypergraphs means that the intersection of a common neighbourhood in the first hypergraph and another in the second hypergraph has about the size one would expect in two independent random hypergraphs of the respective densities.

Definition 4.3.3 (coquasirandom)**.** For $\alpha > 0$ and $L \in \mathbb{N}$, we say that a pair of hypergraphs (F, F^*), both on the same vertex set V of order n and semi-edge set $S(V)$ and with densities p and p^*, respectively, is (α, L)*-coquasirandom* if for every set $S \subseteq S(V)$ of at most L semi-edges and every subset $R \subseteq S$ we have

$$|\mathsf{N}_F(R) \cap \mathsf{N}_{F^*}(S \setminus R)| = (1 \pm \alpha)p^{|R|}(p^*)^{|S\setminus R|}n\,.$$

The reader should make the following important observation: while we are currently thinking of the reservoir H_0^* as being a hypergraph of small but positive density, we will also want to make use of the same analysis when H_0^* is a zero-edge hypergraph in order to consider the embedding of the non-spanning graphs. Coquasirandomness of (F, F^*) makes sense if F^* has no edges (and so $p^* = 0$): it reduces to quasirandomness of F. Similarly, in the following setting, and consequently in most of the following lemmas, we assume that F^* has density either bounded away from zero or equal to zero. If the final δn vertices of each G_s are isolated vertices, and H_0^* has density zero, then the embedding loop of *PackingProcess* effectively just runs *RandomEmbedding* repeatedly to embed the first $n - \delta n$ vertices of each G_s into H_{s-1} and then removes the used edges to form H_s. We will see this algorithm explicitly (as *PackingProcess2*) in Section 4.7, and will use the lemmas below and in the following two sections for its analysis.

With this we can state the setting of our main lemmas and fix various constants which we will use throughout this chapter.

Setting 4.3.4. *Let $D, n, r \in \mathbb{N}$ and $\gamma > 0$ be given. Without loss of generality, we may suppose γ is sufficiently small to play the rôle of μ in Lemma 4.2.6. We define*

$$\begin{aligned}
&Q = D(r+1) + 3\,, \quad s_{\max} = \frac{2n^{r-1}}{(r-1)!}\,, \quad \eta \le \frac{\gamma^Q}{200Q}\,,\\
&\delta = \frac{\gamma^{10Q}\eta}{10^6 Q^4}\,, \quad C = 40Q\exp\left(1000Q\delta^{-2}\gamma^{-2Q-10}\right),\\
&\alpha_x = \frac{\delta^2}{10^8 C^3 Q}\exp\left(\frac{10^8 C D^3 Q^3 r^3 r!\delta^{-2}\gamma^{-2Q}(x - s_{\max})}{n^{r-1}}\right) \qquad \textit{for each } x \in \mathbb{R},\\
&\varepsilon = \frac{\alpha_0\delta^8\gamma^{10Q}}{1000CQr}\,, \quad c = 10^{-8}C^{-2}r^{-2}2^{-2Q}Q^{-4}\varepsilon^8 \quad \textit{and} \quad \xi = \alpha_0/100\,,
\end{aligned} \tag{4.3.1}$$

and we require that $10D\eta$ to be small enough to play the rôle of ε in Lemma 4.2.6 for input $\mu = \gamma^D$.

Let $G_1, G_2, \ldots, G_{s^*}$ (for some $s^* \le s_{\max}$) be hypergraphs on $[n]$, such that for each s and $x \in V(G_s)$ we have $\deg^-_{G_s}(x) \le D$, such that $\Delta(G_s) \le cn/\log n$, and such that the final δn vertices of G_s all have degree d_s and form an strongly independent set.

Let H_0 and H_0^* be two edge-disjoint hypergraphs on the same vertex set of order n such that (H_0, H_0^*) is $(\frac{1}{4}\alpha_0, Q)$-coquasirandom, and $\sum_{s\in[s^*]} e(G_s) \le e(H_0) - \gamma n^r$. Suppose that H_0^* has either zero edges or at least $(\gamma - \alpha_0)\binom{n}{r}$ edges.

Note that in (4.3.1) we give numbers α_x which we call 'constant' even though n appears in their definition. Observe that α_x is strictly increasing in x. We will be interested only in values $0 \leq x \leq s_{\max}$ (though it is technically convenient to have the definition for all $x \in \mathbb{R}$), and it is easy to check that neither α_0 nor $\alpha_{s_{\max}}$ depends on n.

Remark 4.3.5. *Note that for the proof to work it is sufficient to choose the constants in the following manner.*

$$D^{-1}, \gamma, r^{-1} \gg \eta \gg \delta \gg C^{-1} \gg \alpha_0 \gg \varepsilon \gg c \tag{4.3.2}$$

We give exact formulas so our calculations can be properly checked, but most of the time one should simply think of the constants as being of very different orders of magnitude.

The next lemma justifies splitting our host graph into bulk and reservoir, maintaining quasirandomness.

Lemma 4.3.6. *For each $r \geq 2$, $D \in \mathbb{N}$ and each $\gamma > 0$, and for each n sufficiently large, let us suppose that the constants Q, α_0 and ξ are as in Setting 4.3.4.*

Suppose that $\widehat{H}$ is a (ξ, Q)-quasirandom hypergraph of order n and density $p \geq 3\gamma$. Let H_0^* be a q-random subgraph of $\widehat{H}$, where $q = \gamma/p$. Let H_0 be the complement of H_0^* in $\widehat{H}$. Then with probability at least $1 - n^{-5C}$, we have that $e(H_0^*) = (1 \pm \alpha_0)\gamma\binom{n}{r}$ and the pair (H_0, H_0^*) is $\left(\frac{1}{4}\alpha_0, Q\right)$-coquasirandom.

The next lemma states that coquasirandomness of (H_s, H_0^*) is preserved when we embed into the bulk.

Lemma 4.3.7. *For each $r \geq 2$, $D \in \mathbb{N}$ and each $\gamma > 0$, and for each n sufficiently large, the following holds with probability at least $1 - n^{-4C}$. Suppose that the constants and $G_1, G_2, \ldots, G_{s^*}$ and the hypergraph $H_0 \cup H_0^* = H$ and the constant Q are as in Setting 4.3.4. When* PackingProcess *is run, for each $s \in [s^*]$ either* PackingProcess *fails before completing stage s, or the pair (H_s, H_0^*) is (α_s, Q)-coquasirandom.*

The next lemma estimates the probability that a single execution of *RandomEmbedding* succeeds.

Lemma 4.3.8. *For each $r \geq 2$, $D \in \mathbb{N}$, each $\gamma > 0$, and any sufficiently large n, let $Q, \delta, \eta, \alpha_0, \alpha_{s_{\max}}, \varepsilon$ and c be as in Setting 4.3.4. Given any $\alpha_0 \leq \alpha \leq \alpha_{s_{\max}}$, let G*

be a hypergraph on vertex set $[n]$ *with maximum degree at most* $cn/\log n$ *such that* $\deg^-(x) \leq D$ *for each* $x \in V(G)$*, and let* H *be any* (α, Q)*-quasirandom* n*-vertex hypergraph with at least* $\gamma\binom{n}{r}$ *edges. When* RandomEmbedding *is run then it fails with probability at most* n^{-5C}*.*

For the following lemma we need to define the following property which states that the random choice of the extension has desirable random properties.

Our final two main lemmas concern the completion phase of *PackingProcess*. The first states that the completion phase is likely to delete very few edges at any vertex of H_0^*.

Lemma 4.3.9. *For each* $r \geq 2$*,* $D \in \mathbb{N}$ *and* $\gamma > 0$*, let* n *be sufficiently large. Suppose that the constants and* $G_1, G_2, \ldots, G_{s^*}$ *and* H *are as in Setting 4.3.4, and suppose* H_0^* *has at least* $(\gamma - \alpha_0)\binom{n}{r}$ *edges. When* PackingProcess *is run, with probability at least* $1 - n^{-3C}$ *one of the following three events occurs.*

- PackingProcess *fails.*
- *There is some* $s \in [s^*]$ *such that* (H_s, H_0^*) *is not* (α_s, Q)*-coquasirandom.*
- *For each* $s \in [s^*]$ *and* $y \in S(H_s^*)$ *we have* $\deg_{H_0^*}(y) - \deg_{H_s^*}(y) \leq 10r!\gamma^{-D}D\delta n$*, and* (H_s, H_s^*) *is* (η, Q)*-coquasirandom.*

We will show in the proof of Theorem 4.3.1 that the first two events are unlikely, so that the likely event is the last one.

Our last lemma states that with high probability, at any stage s, provided (H_{s-1}, H_{s-1}^*) is sufficiently coquasirandom, running *RandomEmbedding* to partially embed G_s into H_{s-1} is likely to give a partial embedding which can be completed to an embedding of G_s using H_s^*.

Lemma 4.3.10. *For each* $r \geq 2$*,* $D \in \mathbb{N}$ *and each* $\gamma > 0$*, and for each* n *sufficiently large, let the constants be as in Setting 4.3.4. Suppose that* G *is a hypergraph on* $[n]$*, such that we have* $\deg^-(x) \leq D$ *for each* $x \in V(G)$*, we have* $\Delta(G) \leq cn/\log n$*, and such that the final* δn *vertices of* G *form a strongly independent set, and all have degree* d*. Suppose* (H, H^*) *are a pair of* (η, Q)*-coquasirandom hypergraphs on* n *vertices, and* H *is* (α_{s^*}, Q)*-quasirandom, with* $e(H) = p\binom{n}{r}$ *and* $e(H^*) = (1 \pm \eta)\gamma\binom{n}{r}$*, where* $p \geq \gamma$*. When* RandomEmbedding *is run to embed* $G[[n-\delta n]]$ *into* H*, with probability at least* $1 - n^{-2C}$ *it returns a partial embedding* φ *which can be extended to an embedding* φ^* *of* G *into* $H \cup H^*$*, with all the edges using a vertex in* $\{n - \delta n + 1, \ldots, n\}$ *mapped to* H^**.*

Let us briefly explain why we cannot simply perform the whole embedding in the quasirandom $\widehat{H}$, but have to split it into a bulk and a reservoir. In order to analyse *RandomEmbedding*, we require that the bulk is very quasirandom, but *RandomEmbedding* is very well-behaved and preserves this good quasirandomness. In contrast, we are not able to show that the completion stage, where we choose a system of distinct representatives for the remaining vertices, is so well-behaved. If we used the bulk for this embedding the errors would rapidly become unacceptably large. However, to show that choosing such a system of distinct representatives is possible, we do not need much quasirandomness. Thus the reservoir H_s^* does rapidly lose its quasirandomness (compared to H_s), but it is sufficient for the completion.

We now argue that our main lemmas imply Theorem 4.3.1.

Proof of Theorem 4.3.1. We can assume that $p > 3\gamma$ as the statement is vacuous otherwise.

Suppose that we run *PackingProcess* on the input hypergraphs $G_1, \ldots, G_{s^*}$. For the course of the analysis of this run, we shall first ignore possible failures during the completion phase. That is, if any failure during the completion phase occurs, we ignore it and continue embedding using *RandomEmbedding* into the bulk. Clearly, this does not change behaviour of future rounds of *RandomEmbedding* or the evolution of the bulk.

It is clear that *PackingProcess* does not fail (in the *RandomEmbedding* stage) unless at least one of the following exceptional events occurs:

(i) (H_0, H_0^*) is not $(\frac{1}{4}\alpha_0, Q)$-coquasirandom.

(ii) *RandomEmbedding* proceeded through stages $s = 1, \ldots, s_0$ (for some $s_0 \in [s^* - 1]$) without failure, the pairs (H_s, H_0^*) are (α_s, Q)-coquasirandom for $s < s_0$, and (H_{s_0}, H_0^*) is not an (α_r, Q)-coquasirandom pair.

(iii) *RandomEmbedding* proceeded through stages $s = 1, \ldots, s_0$ (for some $s_0 \in \{0, \ldots, s^* - 1\}$) without failure, the hypergraphs H_s are (α_s, Q)-quasirandom for $s \leq s_0$. Then, in stage $s_0 + 1$, *RandomEmbedding* fails.

Lemma 4.3.6 gives an upper bound on the probability of the event in (i). Lemma 4.3.7 gives an upper bound on the probability of all the events in (ii). For each fixed $s_0 \in \{0, \ldots, s^* - 1\}$, the event in (iii) can be bounded using Lemma 4.3.8. Thus, the probability that *PackingProcess* fails in the *RandomEmbedding* part is at most $n^{-5C} + n^{-4C} + n^{-5C}$.

Let us now analyse the completion phases of *PackingProcess*. If *PackingProcess* fails in one of the completion phases then one of the following events occurs:

(iv) One of the events described under (i)-(iii).

(v) None of (i)-(iii) occurs. *RandomEmbedding* and the completion phase proceed successfully through the first s_0 stages (for some $r \in \{1, \dots, s^* - 1\}$. For $s \in [s_0]$ all the pairs (H_s, H_0^*) are (α_s, Q)-coquasirandom. However, there is a stage $s \in [r]$ where (H_s, H_s^*) is not (η, Q)-coquasirandom.

(vi) None of (i)-(iii) occurs. *RandomEmbedding* and the completion phase proceeds successfully through the first s_0 stages (for some $s_0 \in \{0, \dots, s^* - 1\}$, and throughout all the pairs (H_s, H_0^*) and (H_s, H_s^*) are (α_s, Q)-coquasirandom and (η, Q)-coquasirandom, respectively. In stage $s_0 + 1$, *RandomEmbedding* successfully embeds but the completion phase fails.

Lemma 4.3.9 bounds the probability of the event in (v) by n^{-3C}. Finally, Lemma 4.3.10 bounds the probability of events in (vi) for each given s_0 by n^{-2C}. Thus, the total probability of failure due to (v) or (vi) is at most $n^{-2C} + n^{-3C}$.

We conclude that *PackingProcess* packs the hypergraphs $G_1, \dots, G_{s^*}$ into $\widehat{H}$, and none of the above bad events occur, with probability at least $1 - n^{-C}$. This in particular gives the desired packing. In addition, we see that $(H_{s^*}, H_{s^*}^*)$ is (η, Q)-coquasirandom. Let H' be the hypergraph of leftover edges, i.e. $E(H') = E(H_{s^*}) \cup E(H_{s^*}^*)$. Given a set S of semi-edges of size at most Q, the set $\mathsf{N}_{H'}(S)$ is partitioned into sets

$$\mathsf{N}_{H_{s^*}}(R) \cap \mathsf{N}_{H_{s^*}^*}(S \setminus R)$$

as R runs over all subsets of S. Since these sets have sizes controlled by $(H_{s^*}, H_{s^*}^*)$ being (η, Q)-coquasirandom, we see that H' is $(2^Q\eta, Q)$-quasirandom, as desired. □

It thus remains to prove all the main lemmas from this section. Lemmas 4.3.6 and 4.3.7 are proven in Section 4.5. We actually prove a stronger statement than Lemma 4.3.8 in Section 4.4. This stronger form, Lemma 4.4.3 is also needed for proving Lemma 4.3.7. Lemmas 4.3.9 and 4.3.10 are proven in Section 4.6.

4.4 Staying on a diet

In this section we consider the running of *RandomEmbedding* to embed one degenerate hypergraph G_s into a quasirandom hypergraph H_{s-1}. Since we will only consider one stage s, to avoid a profusion of subscripts we write G in place of G_s and H in place of H_{s-1}.

The basic strategy here is broadly similar to that in [3]. Our main aim is to show that during the embedding of G into H, if X is the set of vertices which have been used in the embedding at any given time t, then X looks like a random set of vertices in that it intersects any given common neighbourhood in about as many vertices as a uniform random set of the same size would do.

Definition 4.4.1 (diet condition, codiet condition)**.** Let H be a hypergraph on n vertices and $p\binom{n}{r}$ edges, and let $X \subseteq V(H)$ be any vertex set. We say that the pair (H, X) satisfies the (β, L)*-diet condition* if for every set $S \subseteq S(H)$ of at most L semi-edges we have $|\mathsf{N}_H(S) \setminus X| = (1 \pm \beta)p^{|S|}(n - |X|)$.

Let H, H^* be two hypergraphs on vertex set V of order n vertices with semi-edge set $S(V)$ and $p\binom{n}{r}$ and $p^*\binom{n}{r}$ edges, respectively, and let $X \subseteq V$ be any vertex set. We say that the triple (H, H^*, X) satisfies the (β, L)*-codiet condition* if for every set $S \subseteq S(V)$ of at most L semi-edges and for every subset $R \subseteq S$ we have

$$\Big|\big(\mathsf{N}_H(R) \cap \mathsf{N}_{H^*}(S \setminus R)\big) \setminus X\Big| = (1 \pm \beta)p^{|R|}(p^*)^{|S\setminus R|}(n - |X|)\ .$$

Observe that the (β, L)-diet condition holding for $(H, \emptyset)$ is simply the statement that H is (β, L)-quasirandom, and similarly for the codiet condition. To see why it is enough to show the diet condition holds for (H, X) where X is the set of vertices used up to some given time t in the embedding, consider the embedding of vertex $t + 1$ of G. The only way *RandomEmbedding* can fail is if there is no vertex in the candidate set which is not contained in X. But the candidate set is precisely a common neighbourhood of some at most D semi-edges in H, namely the semi-edges to which were embedded the left-neighbourhood of $t + 1$. So the diet condition tells us how many vertices in the candidate set are not covered by X, and in particular that that number is not zero.

In order to argue that we maintain the diet condition, we introduce the cover condition. Roughly speaking, this states that for any given v in H and any short interval of vertices (of length εn) of G, about the 'right fraction' of vertices x in the interval have v in their final candidate set when *RandomEmbedding* is run. To make precise what we mean by 'the right fraction' some care is needed. How likely it is that v is in the final candidate set of x depends on the number of semi-edges in the left-neighbourhood of x. Therefore we will partition $V(G)$ according to this number of previous neighbours. Thus we define

$$X_{i,d} := \{x \in V(G) \colon i \le x < i + \varepsilon n, |\mathsf{N}^-(x)| = d\} .$$

When G is given with a D-degenerate ordering it is enough to consider $d \in \{0, 1, \dots, D\}$. So if H is quasirandom and has $p\binom{n}{r}$ edges, then for an arbitrary $v \in V(H)$, we would expect that about a p^d-fraction of vertices x in each $X_{i,d}$ have v in their final candidate sets (at time $x - 1$).

Definition 4.4.2 (cover condition)**.** Suppose that G and H are two hypergraphs such that H has order n, the vertex set of G is $[n]$, and G has density p. Suppose that numbers $\beta, \varepsilon > 0$ and $i \in [n - \varepsilon n]$ are given. We say that a partial embedding ψ of G into H, which embeds $\bigcup \mathsf{N}^-(x)$ for each $i \le x < i + \varepsilon n$, satisfies the (ε, β, i)*-cover condition* if for each $v \in V(H) \setminus \operatorname{im} \psi$, and for each $d \in \mathbb{N}$, we have

$$\left|\left\{x \in X_{i,d} : v \in \mathsf{N}_H\big(\psi(\mathsf{N}^-(x))\big)\right\}\right| = (1 \pm \beta) p^d |X_{i,d}| \pm \varepsilon^2 n .$$

Note that a corresponding condition for $d = 0$ is trivial, even with zero error parameters.

The main idea of the analysis is to show that if the diet condition holds up to some given time t, then it is unlikely that the cover condition fails at or before time t, and similarly if the diet and cover conditions hold up to time t then the diet condition is unlikely to fail at time $t + 1$ (and so *RandomEmbedding* does not fail before this time either). We wrap this up in the following lemma.

Lemma 4.4.3 (Diet-and-cover lemma)**.** *For each $D \in \mathbb{N}$, each $\gamma > 0$, and any sufficiently large n, let $s_{\max}, Q, \delta, \eta, \alpha_0, \alpha_{s_{\max}}, \varepsilon$ and c, C be as in Setting 4.3.4. Let $\alpha \in [\alpha_0, \alpha_{s_{\max}}]$ be arbitrary. Let G be a hypergraph on vertex set $[n]$ with maximum degree at most $cn/\log n$ such that $\deg^-(x) \le D$ for each $x \in V(G)$, and let H be any (α, Q)-quasirandom n-vertex hypergraph with at least $\gamma\binom{n}{r}$ edges. Suppose in addition that H^* is a hypergraph on $V(H)$ with $\hat{p}\binom{n}{r}$ edges such that (H, H^*) is (η, Q)-coquasirandom and either $\hat{p} \ge (1 - \eta)\gamma$ or $\hat{p} = 0$. Finally, fix any set Z of vertices of H with $|Z| \ge c^{-1} \log n$. Then with probability at least $1 - n^{-5C}$ all of the following good events hold.*

(a) When RandomEmbedding *is run it does not fail and generates a sequence $(\psi_i)_{i \in [n - \delta n]}$ of partial embeddings of G into H.*

(b) For each $t \in [n - \delta n]$ the pair $(H, \operatorname{im} \psi_t)$ satisfies the $(C\alpha, Q)$-diet condition.

(c) For each $1 \le t \le n+1-\varepsilon n$, the partial embedding $\psi_{t+\varepsilon n-2}$ of G into H satisfies the $(\varepsilon, C\alpha, t)$-cover condition.

(d) For each $t \in [n-\delta n]$, the triple $(H, H^*, \operatorname{im}\psi_t)$ satisfies the $(2\eta, Q)$-codiet condition.

(e) For each $t \in [n-\delta n]$, we have $|Z \setminus \operatorname{im}\psi_t| = (1 \pm C\alpha)\frac{n-t}{n}|Z|$.

Observe that conclusion (a) of this lemma is the conclusion of Lemma 4.3.8, so that proving Lemma 4.4.3 also proves Lemma 4.3.8.

The main difficulty is to establish that the cover and diet conditions hold. We will see that the other two conditions are easy byproducts. The reason for the difficulty is that the error terms in the cover and diet conditions for small times t feed back into the calculations which will establish the cover and diet conditions for larger times t, so that the errors grow. To bound their growth, we define a new sequence of error terms, which we need only in the proof of Lemma 4.4.3. There are three important points to note about the following sequence β_t. It is an increasing sequence, we have $\beta_0 = 2\alpha$, and β_n/β_0 is bounded by a constant which does not depend on α (though it does depend on D, γ and δ). This last observation will turn out to be crucial for the analysis of *PackingProcess*.

Definition 4.4.4. Given Q and $\alpha, \delta, \gamma > 0$, we define

$$\beta_t := 2\alpha \exp\left(\tfrac{1000Q\delta^{-2}\gamma^{-2Q-10}t}{n}\right) . \tag{4.4.1}$$

We will mainly take t integer in the range $[0, n]$, but it is convenient to allow t to be any real number. In particular, for each $t \ge 0$, we have

$$\begin{aligned} &\tfrac{1}{n}\int_{i=0}^{t} 1000Q\delta^{-2}\gamma^{-2Q-10}\beta_i \,\mathrm{d}i \\ \le &2\alpha \int_{i=-\infty}^{t} \frac{1000Q\delta^{-2}\gamma^{-2Q-10}}{n} \exp\left(\tfrac{1000Q\delta^{-2}\gamma^{-2Q-10}i}{n}\right) \mathrm{d}i = \beta_t \,. \end{aligned} \tag{4.4.2}$$

One should read (4.4.2) as saying that when we want to estimate a parameter of the process *RandomEmbedding* at some time t, even if we make in each step i an error which is a rather large multiple of the current error β_i, the total error is still bounded by β_t.

We split the proof of Lemma 4.4.3 into two parts. The cover lemma (Lemma 4.4.5) states that if the (β_t, Q)-diet condition holds for $(H, \operatorname{im}\psi_i)$ for each $i \in [t-1]$, then it is very unlikely that the $(\varepsilon, 20Q\beta_t, t)$-cover condition fails for $\psi_{t+\varepsilon n-2}$. Note that the time $t+\varepsilon n-2$ is the first time at which the $(\varepsilon, 20Q\beta_t, t)$-cover condition is guaranteed to be determined, since at this time all left-neighbours of all vertices $t, t+1, \ldots, t+\varepsilon n-1$ have certainly been embedded.

Lemma 4.4.5 (Cover lemma)**.** *For each D, each $\gamma > 0$ and sufficiently large n, let Q, $s_{\max}$, α_0, $\alpha_{s_{\max}}$, ε, δ and c be as in Setting 4.3.4. Suppose that $\alpha_0 \le \alpha \le \alpha_{s_{\max}}$ and G is a hypergraph on vertex set $[n]$, with $\deg^-(x) \le D$ for each $x \in [n]$, with maximum degree at most $cn/\log n$, and suppose that H is an n-vertex hypergraph of density at least γ. Let β_t for $0 \le t \le n$ be defined as in (4.4.1) and assume that $\beta_n \le \frac{1}{10}$. Let t with $1 \le t \le n - \delta n - \varepsilon n + 1$ be fixed.*

When RandomEmbedding *is run to embed G into H, with probability at most n^{-6C} the following holds. For each $0 \le i \le t-1$ the (β_t, Q)-diet condition holds for $(H, \operatorname{im}\psi_i)$, but the $(\varepsilon, 20Q\beta_t, t)$-cover condition does not hold for $\psi_{t+\varepsilon n-2}$.*

Note that for any $0 \le t \le n-\delta n-\varepsilon n$, if *RandomEmbedding* runs up to time t and the (β_t, Q)-diet condition holds for $(H, \operatorname{im}\psi_t)$, then by choice of ε the $(2\beta_t, Q)$-diet condition holds deterministically for $(H, \operatorname{im}\psi_j)$ for each $t+1 \le j \le t+\varepsilon n$. In particular *RandomEmbedding* cannot fail before time $t+\varepsilon n$.

The diet lemma (Lemma 4.4.6) states that when the (β_i, Q)-diet condition holds for $(H, \operatorname{im}\psi_i)$ for each $i \in [t-1]$, and the $(\varepsilon, 20Q\beta_i, i)$-cover condition holds for $\psi_{i+\varepsilon n-2}$ for each $i \in [t+1-\varepsilon n]$, then it is unlikely that the (β_t, Q)-diet condition fails for $(H, \operatorname{im}\psi_t)$. We also obtain the desired codiet condition.

Lemma 4.4.6 (Diet lemma)**.** *For each $r \ge 2$, $D \in \mathbb{N}$, each $\gamma > 0$, and any sufficiently large n, let Q, α_0, α_{2n}, ε, δ and η be as in Setting 4.3.4. For any $t \le (1-\delta)n$, and $\alpha_0 \le \alpha \le \alpha_{2n}$ the following holds. Suppose that G is a hypergraph on $[n]$ such that $\deg^-(x) \le D$ for each $x \in [n]$, and H is an (α, Q)-quasirandom hypergraph with n vertices with $p\binom{n}{r}$ edges, with $p \ge \gamma$. Suppose furthermore that H^* is a hypergraph on $V(H)$ and $\hat{p}\binom{n}{r}$ edges with either $\hat{p} \ge (1-\eta)\gamma$ or $\hat{p} = 0$, such that (H, H^*) satisfies the (η, Q)-coquasirandomness condition. Let β_t for $0 \le t \le n$ be defined as in (4.4.1) and assume that $\beta_n \le \frac{1}{10}$. Finally let Z be any subset of $V(H)$ of size at least $c^{-1}\log n$.*

When RandomEmbedding *is run to embed G into H, with probability at most n^{-6C} the following event occurs.*

- *For each $1 \le j \le t-1$ the (β_j, Q)-diet condition holds for $(H, \operatorname{im}\psi_j)$, and*
- *for each $1 \le j \le t+1-\varepsilon n$ the $(\varepsilon, 20Q\beta_j, j)$-cover condition holds for $\psi_{j+\varepsilon n-2}$, and*

- *the* (β_t, Q)*-diet condition does not hold for* $(H, \operatorname{im}\psi_t)$*, or the* $(2\eta, Q)$*-codiet condition does not hold for* $(H, H^*, \operatorname{im}\psi_t)$*, or* $|Z \setminus \operatorname{im}\psi_t| \neq (1 \pm \beta_t)\frac{n-t}{n}|Z|$.

Since the hypergraphs G and H are fixed in the proof of Lemma 4.4.3, in this section we drop the subscript in the notation $C^j_{G\hookrightarrow H}(x)$ and write simply $C^j(x)$. We now show that Lemmas 4.4.5 and 4.4.6, whose proofs are deferred to later in this section, imply Lemma 4.4.3.

Proof of Lemma 4.4.3. Given Q and $\gamma, \delta > 0$, we have C as in Setting 4.3.4. Given $\alpha > 0$, we define β_t for each $0 \leq t \leq n$ as in (4.4.1). We claim that with high probability, when we run *RandomEmbedding*, the following event $\mathcal{A}$ occurs: the algorithm *RandomEmbedding* does not fail, for each $1 \leq t \leq n - \delta n$ the pair $(H, \operatorname{im}\psi_t)$ satisfies the (β_t, Q)-diet condition and the triple $(H, H^*, \operatorname{im}\psi_t)$ satisfies the $(2\eta, Q)$-codiet condition, and for each $\varepsilon n - 1 \leq t \leq n - \delta n$ the $(\varepsilon, 20Q\beta_{t-\varepsilon n+2}, t - \varepsilon n + 2)$-cover condition holds for ψ_t.

Indeed, if the event $\mathcal{A}$ does not occur then there is a first time witnessing its failure. Let us calculate what is the probability that this first time is t. We can thus assume $\mathcal{A}$ does not fail before time t. This in particular means that the (β_j, Q)-diet condition holds for $(H, \operatorname{im}\psi_j)$ for each $1 \leq j < t$, and the $(\varepsilon, 20Q\beta_{j-\varepsilon n+2}, j - \varepsilon n + 2)$-cover condition holds for ψ_j for each $\varepsilon n - 1 \leq j < t$.

Firstly, we show that *RandomEmbedding* cannot fail at time t, then we use Lemma 4.4.5 to show that with high probability the $(\varepsilon, 20Q\beta_{t-\varepsilon n+2}, t - \varepsilon n + 2)$-cover condition holds for ψ_t. Because the (β_{t-1}, Q)-diet condition holds for (H, ψ_{t-1}), picking $S = \psi_{t-1}(\mathsf{N}^-(t))$, we have $\left|C^{t-1}(t) \setminus \operatorname{im}\psi_{t-1}\right| = \left|\mathsf{N}_H(S) \setminus \operatorname{im}\psi_{t-1}\right| > 0$. It follows that *RandomEmbedding* cannot fail at time t. Now by Lemma 4.4.5, the probability of the $(\varepsilon, 20Q\beta_{t-\varepsilon n+2}, t - \varepsilon n + 2)$-cover condition failing is at most n^{-6C}.

Secondly, we use Lemma 4.4.6 to show that with high probability neither diet condition fails at time t. More precisely, by Lemma 4.4.6, the probability that the (β_t, Q)-diet condition fails for $(H, \operatorname{im}\psi_t)$, or the $(2\eta, Q)$-codiet condition fails for $(H, H^*, \operatorname{im}\psi_t)$, is at most n^{-6C}.

We conclude that the probability that a given t is the first time that we witness event $\mathcal{A}$ failing is at most $2n^{-6C}$. Taking a union bound over the at most n choices of t, we see that with probability at least $1 - n^{-5C}$ the good event from the statement of Lemma 4.4.3 holds, i.e., that *RandomEmbedding* does not fail, and by the choice of C and by (4.4.1), for each $1 \leq t \leq (1 - \delta)n$ the pair $(H, \operatorname{im}\psi_t)$ satisfies the $(C\alpha, Q)$-diet condition and the triple $(H, H^*, \operatorname{im}\psi_t)$ satisfies the $(2\eta, Q)$-codiet condition, and for each $1 \leq t \leq n+1-\varepsilon n$ the embedding $\psi_{t+\varepsilon n-2}$ satisfies the $(\varepsilon, C\alpha, t)$-cover condition, as desired. □

For the next proof we need the following notation. Given $1 \leq \ell \leq d \leq D$ we call a sequence $(a_1, \ldots, a_\ell)$ whose entries are in $[d]$ an (ℓ, d)-*pattern* if we have $\sum_{i=1}^{\ell} a_i = d$.

Next for a vertex $x \in X_{t,d}$ we define a pattern. Let E_x denote the following hyperedge set: $\{e \in E(G) : e \subseteq [x], x \in e\}$, that is the hyperedges that contain x as last vertex. For each vertex $y \in [x-1]$ let $f(y)$ denote the number of edges $e \in E_x$ such that $y \in e$ and $y = \max(e \setminus \{x\})$, i.e. y is the second largest vertex in e. We need to remove the zeroes from this sequence in order to obtain the desired pattern. To that end, let ℓ denote the number of $y \in [x-1]$ such that $f(y)$ is non-zero, and let $a_1, \ldots, a_\ell$ be the nonzero values of $f(y)$ as y runs from 1 to $x-1$ in order. By construction, this is a $(\ell, |E_x|)$-pattern which we call the *pattern associated to* x.

If $\mathbf{a}$ is any (ℓ, d)-pattern with $1 \leq \ell \leq d \leq D$, we let $X_{t,\mathbf{a}}$ be defined as the set of vertices in $X_{t,d}$ whose associated pattern is $\mathbf{a}$. We note that for each d there are 2^{d-1} possible patterns.

We now prove the cover lemma.

Proof of Lemma 4.4.5. Let $e(G) = p\binom{n}{r} \geq \gamma\binom{n}{r}$. Let $\mathcal{D}$ be the event that the (β_t, Q)-diet condition holds for each $(H, \operatorname{im}\psi_i)$ with $1 \leq i \leq t-1$. We fix a vertex $v \in V(H)$. Observe that if $v \in \operatorname{im}\psi_{t+\varepsilon n-2}$ then there is nothing to prove, so when it is necessary we will assume $v \notin \operatorname{im}\psi_{t+\varepsilon n-2}$. We also fix $1 \leq \ell \leq d \leq D$ and an (ℓ, d)-pattern $\mathbf{a}$. Let s_i denote $\sum_{j=1}^{i} a_j$ and $s_0 = 0$. Define $\mathcal{B}_{v,\mathbf{a}}$ as the event that $\mathcal{D}$ holds, and that v and $\mathbf{a}$ witness the failure of the following condition,

$$\left|\left\{x \in X_{t,\mathbf{a}} : v \in \mathsf{N}_H\big(\psi(\mathsf{N}^-(x))\big)\right\}\right| = (1 \pm 20Q\beta_t)p^d|X_{t,\mathbf{a}}| \pm \varepsilon^2 n/2^d .$$

Our aim is to show that

$$\mathbb{P}\left[\mathcal{B}_{v,\mathbf{a}}\right] \leq n^{-7C} D^{-1} 2^{-D} . \tag{4.4.3}$$

A union bound over the choices of v, d and $\mathbf{a}$ then gives the lemma.

Our strategy for proving (4.4.3) is as follows. Ideally, we would like to assert that for each $x \in X_{t,\mathbf{a}}$ the probability of $v \in C^{x-1}(x)$ is roughly p^d and apply Corollary 4.2.2 to bound the probability of the bad event $\mathcal{B}_{v,\mathbf{a}}$. Note that at time

$i = 0$, we have $v \in C^i(x)$, and as i increases, the set $C^i(x)$ changes exactly at times when a vertex $y \in R(x)$ (i.e. a vertex which is second-to-last in a hyperedge with last vertex x) is embedded. This ideal strategy is not possible because the indicator variables $\mathbb{1}_{\{v \in C^{x-1}(x)\}}$ may not be sequentially dependent as x ranges over $X_{t,\mathbf{a}}$: the sets $R(x)$ may interleave each other. For a vertex $y \in R(x)$, we write $F_x(y)$ for the set of semi-edges $e \in \mathsf{N}^-_G(x)$ such that $y = \max(e)$ (i.e. $F_x(y)$ is the set of edges which witness $y \in R(x)$). We write $W_x(y)$ for the event that for each semi-edge $e \in F_x(y)$ we have $\{v\} \cup \psi(e) \in E(H)$. That is, the edges that have y as their second to last vertex do not stop x from having v in their final candidate set. We refine our previous strategy as follows. Let $y_1, \ldots, y_l$ be the vertices of $R(x)$ in increasing order, and we define

$$\widetilde{W}_x(k) = \big(W_x(y_1) \cap W_x(y_2) \cap \ldots \cap W_x(y_k)\big) \quad \text{so} \quad \big\{v \in C^{x-1}(x)\big\} = \bigcap_{k=1}^{\ell} \widetilde{W}_x(k)\,. \tag{4.4.4}$$

The event $\widetilde{W}_x(\ell)$, of course, equals the entire intersection (4.4.4). However, this more complicated way of expressing (4.4.4) gets us into the setting of sequential dependence as in Corollary 4.2.2.

More formally, given $1 \le k \le \ell$ and $y \in V(G)$, we define random variables $Y_{k,1}, \ldots, Y_{k,t+\varepsilon n-2}$ as follows. Let $Y_{k,y}$ be the number of vertices $x \in X_{t,\mathbf{a}}$ such that y is the kth vertex of $R(x)$ in increasing order and $\widetilde{W}_x(k)$ holds. In other words, $Y_{k,y}$ counts the number of $x \in X_{t,\mathbf{a}}$ where y is the kth element of $R(x)$ and which immediately after embedding y could still be embedded to v (i.e. all the semi-edges of $\mathsf{N}^-(x)$ which are contained in $[y]$ have been embedded to semi-edges that make a hyperedge with v). Observe that $Y_{k,1}, \ldots, Y_{k,t+\varepsilon n-2}$ are by definition sequentially dependent.

We write $s_k = \sum_{i=1}^k a_i$ for the number of semi-edges of $\mathsf{N}^-(x)$ which are fully embedded once we embed the kth element of $R(x)$. These quantities are by definition of a pattern the same for each $x \in X_{t,\mathbf{a}}$. For each $0 \le k \le \ell$, we let $\mathcal{Y}_k$ be the event that $(1 \pm 10\beta_t)^{s_k} p^{s_k} |X_{t,\mathbf{a}}| \pm s_k \varepsilon^2 n/(d2^d)$ vertices $x \in X_{t,\mathbf{a}}$ have the property that $\widetilde{W}_x(k)$ holds. Observe that the event $\mathcal{Y}_k$ is precisely the statement that

$$\sum_{y=1}^{t+\varepsilon n-2} Y_{k,y} = (1 \pm 10\beta_t)^{s_k} p^{s_k} |X_{t,\mathbf{a}}| \pm s_k \varepsilon^2 n/(d2^d)\,. \tag{4.4.5}$$

Our bad event then satisfies

$$\mathcal{B}_{v,\mathbf{a}} \subseteq \left(\mathcal{D} \text{ and not } \mathcal{Y}_d\right),$$

because we have $s_k \leq d$ and $(1 \pm 10\beta_t)^d = 1 \pm 20Q\beta_t$. In order to bound the probability of $\mathcal{B}_{v,\mathbf{a}}$ we cover $\mathcal{B}_{v,\mathbf{a}}$ with ℓ events, each of whose probabilities we can bound with Corollary 4.2.2. For this purpose we define the event

$$\mathcal{E}_k = \mathcal{Y}_{k-1} \text{ and } \mathcal{D}$$

for each $1 \leq k \leq \ell$. Note that $\mathcal{E}_1 = \mathcal{D}$ since $\mathcal{Y}_0$ holds trivially with probability one. We thus have

$$\mathcal{B}_{v,\mathbf{a}} \subseteq \left(\mathcal{D} \text{ and not } \mathcal{Y}_d\right) \subseteq \bigcup_{1\leq k\leq \ell} \left(\mathcal{E}_k \text{ and not } \mathcal{Y}_k\right).$$

Our aim then is to show that for each $1 \leq k \leq \ell$ we have

$$\mathbb{P}[\mathcal{E}_k \text{ and not } \mathcal{Y}_k] \leq n^{-7C}\ell^{-1}D^{-1}2^{-D}. \tag{4.4.6}$$

Note that this and a union bound over the ℓ choices of k gives (4.4.3).

To establish (4.4.6) we would like to apply Corollary 4.2.2. Hence we need to argue that either $\mathcal{E}_k$ fails, or we can estimate $\sum_{y=1}^{t+\varepsilon n-2} \mathbb{E}[Y_{k,y}|\mathscr{H}_{y-1}]$, where $\mathscr{H}_{y-1}$ is the history of embedding decisions taken in *RandomEmbedding* up to and including the embedding of vertex $y-1$. To this end, for $y \in [t+\varepsilon n-2]$ let $Z_{k,y}$ be the number of vertices $x \in X_{t,\mathbf{a}}$ such that y is the kth vertex of $R(x)$ and $\widetilde{W}_x(k-1)$ holds. Also let $Y_{k,y,x}$ and $Z_{k,y,x}$ be indicator variables that x is counted in $Y_{k,y}$ and $Z_{k,y}$ respectively. Then the quantity $Z_{k,y,x}$ is determined by $\mathscr{H}_{y-1}$ and

$$\mathbb{E}[Y_{k,y,x}|\mathscr{H}_{y-1}] = Z_{k,y,x} \cdot \mathbb{P}\big[W_x(y)|\mathscr{H}_{y-1}\big]. \tag{4.4.7}$$

Observe further that

$$\sum_{y=1}^{t+\varepsilon n-2} Z_{k,y} = \sum_{y=1}^{t+\varepsilon n-2} Y_{k-1,y}, \tag{4.4.8}$$

because both sums count the number of vertices $x \in X_{t,a}$ such that the first $k-1$ vertices of $R(x)$ are embedded to H so that $\widetilde{W}_x(k-1)$ holds, in the first sum grouped by the kth vertex of $R(x)$ and in the second sum by the $(k-1)$st vertex.

Assume now that $x, y \in V(G)$ are fixed and that $\mathscr{H}_{y-1}$ does not witness that $\mathcal{E}_k$

fails, and let us bound $\mathbb{P}\big[W_x(y)|\mathscr{H}_{y-1}\big]$. Since $\mathscr{H}_{y-1}$ does not witness that $\mathcal{E}_k$ fails and $\mathcal{D} \subseteq \mathcal{E}_k$, the (β_t, Q)-diet condition holds for $(H, \operatorname{im}\psi_{y-\varepsilon n})$, where we have to subtract εn in the index of $\psi_{y-\varepsilon n}$ because y could be as large as $t+\varepsilon n-2$ (and we only know that the diet condition holds up to time $t-1$). This implies that for each set S of semi-edges in H with $|S| \le Q$ we have

$$\begin{aligned}\big|\mathsf{N}_H(S) \setminus \operatorname{im}\psi_{y-1}\big| &= (1 \pm \beta_t)p^{|S|}(n-y+\varepsilon n) \pm \varepsilon n \\ &= (1 \pm \beta_t)p^{|S|}(n-y+1) \pm 2\varepsilon n = (1 \pm 2\beta_t)p^{|S|}(n-y+1)\,,\end{aligned}$$

where the last inequality follows from $\gamma \le p$ and $\varepsilon \le \alpha\gamma^Q \le \frac{1}{2}\beta_t\gamma^Q$. We conclude that the $(2\beta_t, Q)$-diet condition holds for $(H, \operatorname{im}\psi_{y-1})$.

Given ψ_{y-1}, we define a set $N \subseteq C^{y-1}(y) \setminus \operatorname{im}\psi_{y-1}$ of vertices of G with the following property: if we embed y to any $u \in N$, then we can still embed x to v. That is, we put $u \in C^{y-1}(y) \setminus \operatorname{im}\psi_{y-1}$ in N if and only if all the semi-edges $\mathsf{N}^-(x) \cap [y]$ are embedded by $\psi_{y-1} \cup \{y \hookrightarrow u\}$ to semi-edges of H which form edges together with v. Note that this definition forces $N \cap \operatorname{im}\psi_{y-1} = \emptyset$.

We can write N differently: $N = \mathsf{N}_H(S) \setminus \operatorname{im}\psi_{y-1}$ where S is the collection of semi-edges $\psi_{y-1}\big(\mathsf{N}^-(y)\big)$ together with semi-edges of the form $\{v\} \cup \psi_{y-1}(f)$ where $y \notin f$ and $f \cup \{y\}$ is a semi-edge of $F_x(y)$. There are $\deg^-(y)$ semi-edges of the first type, and a_k of the second type. Using this observation, we can estimate $|N|$ using the diet condition. Observe that these semi-edges are all distinct: by definition the semi-edges $\psi_{y-1}\big(\mathsf{N}^-(y)\big)$ contain only vertices in $\operatorname{im}\psi_{y-1}$, while v is not in $\operatorname{im}\psi_{y-1}$. Furthermore they are genuinely semi-edges, i.e. they have $r-1$ vertices: for the first type this is obvious since ψ_{y-1} is injective, while for the second type we need to observe that any such f has $r-2$ vertices by definition and $\psi_{y-1}(f)$ does not contain v.

Since $\deg^-(y) \le D$ we have

$$\begin{aligned}\big|C^{y-1}(y) \setminus \operatorname{im}\psi_{y-1}\big| &= (1 \pm 2\beta_t)p^{\deg^-(y)}(n-y+1) \quad \text{and} \\ |N| &= (1 \pm 2\beta_t)p^{a_k+\deg^-(y)}(n-y+1)\,.\end{aligned}$$

Therefore we have

$$\mathbb{P}\big[W_x(y)|\mathscr{H}_{y-1}\big] = \frac{|N|}{\big|C^{y-1}(y) \setminus \operatorname{im}\psi_{y-1}\big|} = (1 \pm 10\beta_t)p^{a_k}\,.$$

We conclude from (4.4.7) that

$$\sum_{y=1}^{t+\varepsilon n-2} \mathbb{E}(Y_{k,y}|\mathscr{H}_{y-1}) = (1 \pm 10\beta_t)p \sum_{y=1}^{t+\varepsilon n-2} Z_{k,y}\,, \tag{4.4.9}$$

unless $\mathcal{E}_k$ fails. Further, unless $\mathcal{E}_k$ fails, we have

$$\sum_{y=1}^{t+\varepsilon n-2} Z_{k,y} \overset{(4.4.8)}{=} \sum_{y=1}^{t+\varepsilon n-2} Y_{k-1,y} \overset{(4.4.5)}{=} (1 \pm 10\beta_t)^{s_{k-1}} p^{s_{k-1}} |X_{t,\mathbf{a}}| \pm \tfrac{s_{k-1}\varepsilon^2 n}{d2^d}\,.$$

Plugging this into (4.4.9), and noting $s_{k-1}+1 \le s_k$, we get that $\mathcal{E}_k$ fails or we have

$$\sum_{y=1}^{t+\varepsilon n-2} \mathbb{E}[Y_{k,y}|\mathscr{H}_{y-1}] = (1 \pm 10\beta_t)^{s_k} p^{s_k} |X_{t,\mathbf{a}}| \pm \tfrac{(s_{k-1}+0.5)\varepsilon^2 n}{d2^d}\,.$$

Since $0 \le Y_{k,y} \le \deg(y)$ for each y, we can thus apply Corollary 4.2.2 with the event $\mathcal{E} = \mathcal{E}_k$, with $R = \Delta(G)$, with $\tilde{\mu} \pm \tilde{\nu} = (1 \pm 10\beta_t)^{s_k} p^{s_k} |X_{t,\mathbf{a}}| \pm \big(s_{k-1} + 0.5\big)\varepsilon^2 n d^{-1} 2^{-d}$, and with $\tilde{\varrho} = \frac{1}{2}\varepsilon^2 n d^{-1} 2^{-d}$ to conclude that

$$\begin{aligned}\mathbb{P}\left[\mathcal{E}_k \text{ and not } \mathcal{Y}_k\right] &= \mathbb{P}\left[\mathcal{E}_k \text{ and } \sum_{y=1}^{t+\varepsilon n-2} Y_{k,y} \neq \mu \pm (\nu + \varrho)\right] \\ &\le 2\exp\left(\frac{-\tilde{\varrho}^2}{2R(\tilde{\mu}+\tilde{\nu}+\tilde{\varrho})}\right)\,.\end{aligned}$$

Substituting $\Delta(G) \le cn/\log n$, and because $c \le 10^{-4}C^{-1}r^{-1}2^{-2D}D^{-4}\varepsilon^4/(100+r)$ and $d \le D$, we obtain (4.4.6) as desired. □

We deduce the diet lemma from the following simplified version.

Lemma 4.4.7. *For each $r \ge 2$, $D \in \mathbb{N}$, each $\gamma > 0$, and any sufficiently large n, let Q, α_0, α_{2n}, ε, δ and η be as in Setting 4.3.4. For any $t \le (1-\delta)n$, and $\alpha_0 \le \alpha \le \alpha_{2n}$ the following holds. Suppose that G is a hypergraph on $[n]$ such that $\deg^-(x) \le D$ for each $x \in [n]$, and H is an (α, Q)-quasirandom hypergraph with n vertices with $p\binom{n}{r}$ edges, with $p \ge \gamma$. Suppose furthermore that H^* is a hypergraph on $V(H)$ and $\hat{p}\binom{n}{r}$ edges with either $\hat{p} \ge (1-\eta)\gamma$ or $\hat{p} = 0$, such that (H, H^*) satisfies the (η, Q)-coquasirandomness condition. Let β_t for $0 \le t \le n$ be defined as in (4.4.1) and assume that $\beta_n \le \frac{1}{10}$. Finally let Z' be any subset of $V(H)$ of size at least $c^{-1}\log n$.*

When RandomEmbedding *is run to embed G into H, with probability at most n^{-7C} the following event occurs.*

- *For each $1 \le j \le t-1$ the (β_j, Q)-diet condition holds for $(H, \operatorname{im}\psi_j)$, and*
- *for each $1 \le j \le t+1-\varepsilon n$ the $(\varepsilon, 20Q\beta_j, j)$-cover condition holds for ψ_j, and*
- *we have $|Z' \setminus \operatorname{im}\psi_t| \neq \big(1 \pm \frac{1}{4}\beta_t\big)\frac{n-t}{n}|Z'|$.*

Before we prove this lemma, let us briefly observe that it implies Lemma 4.4.6.

Proof of Lemma 4.4.6. Observe that the difference between Lemmas 4.4.6 and 4.4.7 is that the probability bound in Lemma 4.4.7 is stronger and the error term on the size of $|Z \cap \operatorname{im}\psi_t|$ is smaller, and that there are a few more ways in which we can enter the unlikely event of Lemma 4.4.6, namely there can be a failure of the diet or codiet conditions at time t.

Suppose that $R \subseteq S \subseteq S(H)$ are sets of semi-edges, with $|S| \le Q$. Then R witnesses a failure of the (β_t, Q)-diet condition for $(H, \operatorname{im}\psi_t)$ if and only if we have

$$\big|\mathsf{N}_H(R) \setminus \operatorname{im}\psi_t\big| \neq (1 \pm \beta_t)p^{|R|}(n-t)\,.$$

Observe that by (α, Q)-quasirandomness of H we have $\big|\mathsf{N}_H(R)\big| = (1 \pm \alpha)p^{|R|}n$. Letting $Z' = \mathsf{N}_H(R)$, if we have $|Z' \cap \operatorname{im}\psi_t| = \big(1 \pm \frac{1}{4}\delta\beta_t\big)\frac{t}{n}|Z|$, then we have

$$\begin{aligned}\big|\mathsf{N}_H(R) \setminus \operatorname{im}\psi_t\big| &= (1 \pm \alpha)p^{|R|}n\Big(1 - \big(1 \pm \tfrac{1}{4}\delta\beta_t\big)\tfrac{t}{n}\Big) = (1 \pm \alpha)p^{|R|}\big(n - t \pm \tfrac{1}{4}\delta\beta_t n\big)\\ &= (1 \pm \alpha)p^{|R|}\big(1 \pm \tfrac{1}{4}\beta_t\big)(n-t) = (1 \pm \beta_t)p^{|R|}(n-t)\end{aligned}$$

where the first equality on the second line uses $n - t \ge \delta n$ and the second that $\alpha = \frac{1}{2}\beta_0 \le \frac{1}{2}\beta_t$ and that β_t is sufficiently small. In particular, what this calculation establishes is that if R witnesses a failure of the (β_t, Q)-diet condition for $(H, \operatorname{im}\psi_t)$, then the corresponding $Z' = \mathsf{N}_H(R)$ witnesses the low-probability event of Lemma 4.4.7 occurring. Note that since $|Z'| \ge (1-\alpha)p^{|R|}n \ge \frac{1}{2}\gamma^Q n$, the condition on $|Z'|$ of Lemma 4.4.7 is indeed satisfied.

If $\hat{p} = 0$, then the $(2\eta, Q)$-codiet condition for $(H, \hat{H}, \operatorname{im}\psi_t)$ is implied by the (β_t, Q)-diet condition for $(H, \operatorname{im}\psi_t)$ since $\beta_t < \eta$. If $\hat{p} \ge (1-\eta)\gamma$, then a similar calculation shows that if R and S witness a failure of the $(2\eta, Q)$-codiet condition for $(H, H^*, \operatorname{im}\psi_t)$ then the corresponding $Z' = \mathsf{N}_H(R) \cap \mathsf{N}_{H^*}(S \setminus R)$ witnesses the low-probability event of Lemma 4.4.7 occurring. This calculation holds with rather more room to spare since η is much larger than α, and we omit the details. Taking a union bound over the at most $n^{2rQ}+1$ choices of R and S, and of Z in Lemma 4.4.6, we observe that the probability that any one of the corresponding Z' for Lemma 4.4.7 witnesses the low-probability event occurring is at most $2n^{2rQ}n^{-7C} < n^{-6C}$. This is the required upper bound on the probability of the unlikely event of Lemma 4.4.6. □

We now prove Lemma 4.4.7.

Proof of Lemma 4.4.7. Observe that if ψ_{t-1} satisfies the (β_{t-1}, Q)-diet condition, *RandomEmbedding* cannot fail at time t, so ψ_t exists. To begin with, we show that in any short interval of time, not too many vertices can be embedded to Z'. This analysis is not particularly accurate; we need it for the more accurate analysis that follows.

Claim 4.4.8. *For every $0 \le j \le t-1$, if the (β_j, Q)-diet condition holds for $(H, \operatorname{im}\psi_j)$, then with probability at least $1-n^{-10C}$ we have*

$$\big|Z' \cap (\operatorname{im}\psi_{\min(j+\varepsilon n,t)} \setminus \operatorname{im}\psi_j)\big| \le 4\varepsilon\gamma^{-D}\delta^{-1}|Z'|\,.$$

Proof. At each time $j+1 \le i \le \min(j+\varepsilon n, t)$, we embed the vertex i to the set $C^{i-1}(i) \setminus \operatorname{im}\psi_{i-1}$. This set is a common neighbourhood of some at most D semi-edges, from which we remove $\operatorname{im}\psi_j$ and a further at most εn vertices. Since $(H, \operatorname{im}\psi_j)$ satisfies the diet condition, we conclude that $\big|C^{i-1}(i) \setminus \operatorname{im}\psi_{i-1}\big| \ge \frac{3}{4}\gamma^D\delta n - \varepsilon n \ge \frac{1}{2}\gamma^D\delta n$. The probability of embedding i to Z' is thus at most $2\gamma^{-D}\delta^{-1}|Z'|n^{-1}$. By Corollary 4.2.2, the probability that more than $4\gamma^{-D}\delta^{-1}\varepsilon|Z'|$ vertices i with $j+1 \le i \le \min(j+\varepsilon n, t)$ are embedded to Z', is at most $\exp\big(-\frac{1}{2}\gamma^{-D}\delta^{-1}\varepsilon|Z'|\big) \le n^{-9C}$ since $|Z'| \ge c^{-1}\log n$ and by choice of c. ▫

We now state a claim that if the diet condition holds up to time $t-\varepsilon n$, then for any given large set $T \subseteq Z'$, with high probability either the cover condition fails at some time before $t-\varepsilon n$, or ψ_t embeds about the expected fraction of each interval of εn vertices to T.

Claim 4.4.9. *For every $1 \le j \le t-\varepsilon n+1$, and for every $T \subseteq V(H) \setminus \operatorname{im}\psi_j$ with $|T| \ge \frac{1}{2}\gamma^D\delta|Z'|$, if the (β_j, Q)-diet condition holds for $(H, \operatorname{im}\psi_j)$, then with probability at least $1-n^{-9C}$, one of the following occurs.*

(a) ψ_t does not have the $(\varepsilon, 20Q\beta_j, j)$-cover condition, or

(b) $\big|\{x : j \le x < j+\varepsilon n, \psi_{t-1}(x) \in T\}\big| = (1 \pm 40Q\beta_j)\frac{|T|\varepsilon n}{n-j}$.

We defer the proof of this claim to later, and move on to state a second claim, which we will deduce from Claim 4.4.9. Let $\ell = \lfloor \frac{t}{\varepsilon n} \rfloor$. We claim it is likely that either we witness a failure of the diet or cover conditions before time t, or the set $Z' \setminus \operatorname{im} \psi_{\ell\varepsilon n}$ has about the expected size.

Claim 4.4.10. *With probability at least* $1 - n^{-8C}$*, one of the following holds.*

(a) The (β_j, Q)*-diet condition fails for* $(H, \operatorname{im} \psi_j)$ *for some* $1 \le j \le t-1$*, or*

(b) the $(\varepsilon, 20Q\beta_j, j)$*-cover condition fails for* ψ_{t-1} *for some* $1 \le j \le t+1-\varepsilon n$*, or*

(c) we have

$$\left|Z' \setminus \operatorname{im} \psi_{\ell\varepsilon n}\right| = |Z'| \prod_{k=0}^{\ell-1} \left(1 - \left(1 \pm 40Q\beta_{k\varepsilon n}\right) \tfrac{\varepsilon n}{n - k\varepsilon n}\right). \tag{4.4.10}$$

Before proving these claims, we show that they imply the lemma. Suppose that the likely event of Claim 4.4.10 holds, and, if $\ell\varepsilon n < t$, that the likely event of Claim 4.4.8 with $j = \ell\varepsilon n$ holds. Taking logs, we have

$$\begin{aligned}
&\log\left|Z' \setminus \operatorname{im} \psi_{\ell\varepsilon n}\right| \\
&= \log\left|Z'\right| + \sum_{k=0}^{\ell-1} \log\left(1 - (1 \pm 40Q\beta_{k\varepsilon n})\tfrac{\varepsilon n}{n-k\varepsilon n}\right) \\
&= \log\left|Z'\right| + \sum_{k=0}^{\ell-1} \left(\log \tfrac{n-(k+1)\varepsilon n}{n-k\varepsilon n} + \log\left(1 \pm \tfrac{40Q\beta_{k\varepsilon n}\varepsilon n}{n-(k+1)\varepsilon n}\right)\right) \\
&= \log\left|Z'\right| + \log\left(1 - \ell\varepsilon\right) \pm 2\sum_{k=0}^{\ell-1} \tfrac{40Q\beta_{k\varepsilon n}\varepsilon}{1-(k+1)\varepsilon}\,,
\end{aligned}$$

where the final equality holds since $1 - (k+1)\varepsilon \ge \delta$, and hence by choice of ε the quantity $\frac{40Q\beta_{k\varepsilon n}\varepsilon}{1-(k+1)\varepsilon}$ is close to 0. By the likely event of Claim 4.4.8 we assumed, if $\ell\varepsilon n < t$ then the number of vertices embedded to $Z' \setminus \operatorname{im} \psi_{\ell\varepsilon n}$ by ψ_t is at most $4\varepsilon\delta^{-1}\gamma^{-Q}|Z' \setminus \psi_{\ell\varepsilon n}|$. If $\ell\varepsilon n = t$ then the same estimate holds trivially. We conclude $|Z' \setminus \psi_{\ell\varepsilon n}| = |Z' \setminus \psi_t| \pm 4\varepsilon\delta^{-1}\gamma^{-Q}|Z'|$, and so

$$\left|Z' \setminus \operatorname{im} \psi_t\right| = |Z'| \cdot \frac{n - t \pm \varepsilon n}{n} \cdot \exp\left(\pm 80Q\delta^{-1}\varepsilon \sum_{k=0}^{\ell-1} \beta_{k\varepsilon n}\right) \pm 4\varepsilon\delta^{-1}\gamma^{-Q}|Z'|\,. \tag{4.4.11}$$

Now since β_x is increasing in x, we can estimate

$$80Q\delta^{-1}\varepsilon\sum_{k=0}^{\ell-1}\beta_{k\varepsilon n}\leq\tfrac{1}{n}80Q\delta^{-1}\sum_{x=0}^{\ell n-1}\beta_x\leq\tfrac{1}{n}\int_{x=-\infty}^{\ell n}80Q\delta^{-1}\beta_x\,\mathrm{d}x\overset{(4.4.2)}{\leq}\tfrac{1}{16}\beta_{\ell n}\leq\tfrac{1}{16}\beta_t\,.$$

Since $\frac{1}{16}\beta_t$ is small, we have $\exp\big(\pm\frac{1}{16}\beta_t\big)=1\pm\frac{1}{8}\beta_t$. Plugging this into (4.4.11) we get

$$\big|Z'\setminus\operatorname{im}\psi_t\big|=|Z'|\cdot\Big(\frac{n-t\pm\varepsilon n}{n}\cdot\big(1\pm\tfrac{1}{8}\beta_t\big)\pm 4\varepsilon\delta^{-1}\gamma^{-Q}\Big)=\big(1\pm\tfrac{1}{4}\beta_t\big)|Z'|\tfrac{n-t}{n}\,,$$

where the final equality uses the fact that ε is tiny compared to β_t, δ^2 and γ^Q. This concludes the proof of the lemma, modulo the proofs of Claim 4.4.9 and Claim 4.4.10, which we now provide.

Proof of Claim 4.4.9. Let j and T be as in the statement, and suppose that the likely event of Claim 4.4.8 holds for Z' and j. Fix $0\leq d\leq D$. We want to show how to make use of the $(\varepsilon,20Q\beta_j,j)$-cover condition for ψ_j (which we have when Part (a) fails) to deduce that the assertion of Part (b) holds with high probability. That is, we consider the number of vertices in $X_{j,d}$ embedded to T. In order to apply Corollary 4.2.2, we want to estimate the sum over $x\in X_{j,d}$ of the probability that x is embedded to T, conditioning on ψ_{x-1}, that is, we need to estimate the number

$$\frac{\big|T\cap C^{x-1}(x)\setminus\operatorname{im}\psi_{x-1}\big|}{\big|C^{x-1}(x)\setminus\operatorname{im}\psi_{x-1}\big|}\,. \tag{4.4.12}$$

By the diet condition, we have $\big|C^{x-1}(x)\setminus\operatorname{im}\psi_j\big|=(1\pm\beta_j)p^d(n-j)$. Since $j<t\leq(1-\delta)n$, since $x\leq j+\varepsilon n$, since $p\geq\gamma$, and by choice of ε, we have

$$\big|C^{x-1}(x)\setminus\operatorname{im}\psi_{x-1}\big|=(1\pm 2\beta_j)p^d(n-j)\,, \tag{4.4.13}$$

thus providing a bound on the denumerator in (4.4.12). (Note that this bound on the denumerator does not depend on the choice of $x\in X_{j,d}$.) Now x is embedded uniformly at random into $C^{x-1}(x)\setminus\operatorname{im}\psi_{x-1}$, so it remains to determine the sum of the numerators in (4.4.12). We rewrite the sum as

$$\sum_{x\in X_{j,d}}\big|T\cap C^{x-1}(x)\setminus\operatorname{im}\psi_{x-1}\big|=\sum_{v\in T}\big|\{x\in X_{j,d}:v\in C^{x-1}(x)\setminus\operatorname{im}\psi_{x-1}\}\big|\,.$$

We split this sum into two cases. For $v\notin\operatorname{im}\psi_{j+\varepsilon n}$, by definition $v\in C^{x-1}(x)\setminus$

$\operatorname{im}\psi_{x-1}$ holds if and only if $v \in C^{x-1}(x)$, and by the $(\varepsilon, 20Q\beta_j, j)$-cover condition we have

$$\left|\{x \in X_{j,d} : v \in C^{x-1}(x) \setminus \operatorname{im}\psi_{x-1}\}\right| = (1+20\beta_j)p^d|X_{j,d}| \pm \varepsilon^2 n\,.$$

Observe that this quantity is bounded between 0 and εn. For $x \in \operatorname{im}\psi_{j+\varepsilon n}$, we use the trivial bound $0 \le \left|\{x \in X_{j,d} : v \in C^{x-1}(x) \setminus \operatorname{im}\psi_{x-1}\}\right| \le \varepsilon n$, which we write as

$$\left|\{x \in X_{j,d} : v \in C^{x-1}(x) \setminus \operatorname{im}\psi_{x-1}\}\right| = (1+20\beta_j)p^d|X_{j,d}| \pm \varepsilon^2 n \pm \varepsilon n\,.$$

Since there are by the good event of Claim 4.4.8 at most $4\varepsilon\gamma^{-D}\delta^{-1}|Z'|$ vertices $x \in T$ such that $x \in \operatorname{im}\psi_{j+\varepsilon n}$, we obtain the estimate

$$\begin{aligned}\sum_{x\in X_{j,d}} \left|T \cap C^{x-1}(x) \setminus \operatorname{im}\psi_{x-1}\right| &= \sum_{v\in T} \left|\{x \in X_{j,d} : v \in C^{x-1}(x) \setminus \operatorname{im}\psi_{x-1}\}\right| \\ &= |T|\Big((1+20\beta_j)p^d|X_{j,d}| \pm \varepsilon^2 n\Big) \pm 4\varepsilon\gamma^{-D}\delta^{-1}|Z'| \cdot \varepsilon n \\ &= |T|\Big((1+20\beta_j)p^d|X_{j,d}| \pm 10\varepsilon^2\gamma^{-2D}\delta^{-2} n\Big)\,, \end{aligned} \tag{4.4.14}$$

where the second line uses the bounds we calculated above and the third our assumption $|T| \ge \frac{1}{2}\gamma^D\delta|Z'|$.

We can thus apply Corollary 4.2.2, setting $\mathcal{E}$ to be the event that the $(\varepsilon, 20D\beta_j, j)$-cover condition holds for ψ_j. Combining (4.4.13) and (4.4.14), the expected number of vertices of $X_{j,d}$ embedded to T is

$$\begin{aligned}&\frac{(1\pm 20Q\beta_j)p^d|T||X_{j,d}| \pm 10\varepsilon^2\gamma^{-2D}\delta^{-2}n|T|}{(1\pm 2\beta_j)p^d(n-j)} = \\ &= (1\pm 30Q\beta_j)\frac{|T||X_{j,d}|}{n-j} \pm 20\varepsilon^2\gamma^{-3D}\delta^{-3}|T|\,,\end{aligned}$$

where we use $n-j \ge \delta n$ and $p \ge \gamma$. By Corollary 4.2.2, with $R = 1$, the probability that the $(\varepsilon, 20Q\beta_j, j)$-cover condition holds for ψ_j and the outcome differs from this by more than $\varepsilon^2|T|$ is at most $2\exp(-\varepsilon^4|T|) \le n^{-10C}$, where the inequality holds by our assumptions on $|T|$ and $|Z'|$ and choice of c. Taking the union bound over the $D+1$ choices of d and the unlikely events of Claim 4.4.8, we conclude that with probability at most n^{-9C} the $(\varepsilon, 20Q\beta_j, j)$-cover condition holds for ψ_j

and the number of vertices x with $j \leq x < j + \varepsilon n$ embedded to T is not equal to

$$(1 \pm 30Q\beta_j)\frac{|T|\varepsilon n}{n-j} \pm 40(D+1)\varepsilon^2\gamma^{-3D}\delta^{-3}|T| = (1 \pm 40Q\beta_j)\frac{|T|\varepsilon n}{n-j},$$

where the final equality uses our choice of ε tiny compared to D^{-1}, β_j, γ^{3D} and δ^4. This is what we wanted to show. ▫

Proof of Claim 4.4.10. We set $T_k = Z' \setminus \operatorname{im}\psi_{k\varepsilon n}$. Suppose that $|T_k| \geq \frac{1}{2}\gamma^D\delta|Z'|$, and note that this assumption holds for $k = 0$ trivially. On this assumption, we can apply Claim 4.4.9 with $T = T_k$ and obtain that with probability at least $1 - n^{-9C}$ either a failure of the diet or the cover condition is witnessed before time k, or we have

$$|T_{k+1}| = |T_k|\Big(1 - \big(1 \pm 40Q\beta_{k\varepsilon n}\big)\frac{\varepsilon n}{n - k\varepsilon n}\Big).$$

Inductively, taking the union bound over $0 \leq k \leq \ell - 1$, with probability at least $1 - n^{1-9C} > 1 - n^{-8C}$ one of the following occurs. Either we witness a failure of the diet or cover condition before time $\ell\varepsilon n$, or for each $1 \leq k \leq \ell$ we have $|T_{k-1}| \geq \frac{1}{2}\gamma^D\delta|Z'|$ and hence

$$\big|T_k\big| = |Z'|\prod_{i=0}^{k-1}\Big(1 - (1 \pm 40Q\beta_{i\varepsilon n})\tfrac{\varepsilon n}{n - i\varepsilon n}\Big) \geq \tfrac{1}{2}\gamma^D\delta|Z'|.$$ ▫

□

4.5 Maintaining quasirandomness

In this section we provide the proofs of Lemma 4.3.6 and Lemma 4.3.7.

4.5.1 Initial coquasirandomness

We begin with the easy proof of Lemma 4.3.6, which states that splitting the edges of a quasirandom hypergraph randomly gives a coquasirandom pair with high probability.

Proof of Lemma 4.3.6. Using Theorem 4.2.1 we see that the densities p_0 and p_0^* of H_0 and H_0^* satisfy

$$p_0 = (1 \pm \tfrac{\alpha_0}{1000Q})(p - \gamma) \quad \text{and} \quad p_0^* = (1 \pm \tfrac{\alpha_0}{1000Q})\gamma \tag{4.5.1}$$

with probability at least $1-n^{-6C}$, giving the first part of Lemma 4.3.6.

Now, let $R \subseteq S \subseteq S(\widehat{H})$ be two sets of size at most Q. By quasirandomness of $\widehat{H}$ we have $|\mathsf{N}_{\widehat{H}}(S)| = (1\pm\xi)p^{|S|}n$. Observe that each vertex of $\mathsf{N}_{\widehat{H}}(S)$ appears with probability $q^{|R|}(1-q)^{|S\setminus R|}$ in $\mathsf{N}_{H_0^*}(R)\cap\mathsf{N}_{H_0}(S\setminus R)$. Hence,

$$\mathbb{E}\left[\left|\mathsf{N}_{H_0^*}(R)\cap\mathsf{N}_{H_0}(S\setminus R)\right|\right] = q^{|R|}(1-q)^{|S\setminus R|}(1\pm\xi)p^{|S|}n\,.$$

Observe also that for distinct vertices in $\mathsf{N}_{\widehat{H}}(S)$ the events whether these appear in $\mathsf{N}_{H_0^*}(R)\cap\mathsf{N}_{H_0}(S\setminus R)$ are independent. Using again Theorem 4.2.1, with probability at least $1-n^{-Q-10}$ we have that

$$\left|\mathsf{N}_{H_0^*}(R)\cap\mathsf{N}_{H_0}(S\setminus R)\right| = q^{|R|}(1-q)^{|S\setminus R|}(1\pm 2\xi)p^{|S|}n\,. \tag{4.5.2}$$

Taking the union bound we conclude that (4.5.2) holds for all $S\subseteq S(\widehat{H})$ with $|S|\le Q$ and $R\subseteq S$ with probability at least $1-n^{-6C}$.

Now, assume that (4.5.1) holds. Then the right-hand side of (4.5.2) can be rewritten as

$$\begin{aligned}(1\pm 2\xi)\gamma^{|R|}(p-\gamma)^{|S\setminus R|}n &= (1\pm 2\xi)\left(\frac{p_0^*}{1\pm\frac{\xi_0}{1000Q}}\right)^{|R|}\left(\frac{p_0}{1\pm\frac{\alpha_0}{1000Q}}\right)^{|S\setminus R|}n\\ &= (1\pm 2\xi)(1\pm\tfrac{\alpha_0}{100})(p_0^*)^{|R|}p_0^{|S\setminus R|}\\ &= \left(1\pm\tfrac{1}{10}\alpha_0\right)(p_0^*)^{|R|}p_0^{|S\setminus R|}\,.\end{aligned}$$

We conclude that (H_0^*,H_0) is $\left(\frac{1}{10}\alpha_0,Q\right)$-coquasirandom with probability at least $1-n^{-5C}$. □

4.5.2 Maintaining coquasirandomness

In this subsection we prove Lemma 4.3.7. We need to show that, provided coquasirandomness is maintained up to stage $s-1$ and *RandomEmbedding* does not fail, it is likely that coquasirandomness holds after stage s, when G_s is embedded into H_{s-1} and we obtain H_s. Let us briefly sketch the idea (for convenience focusing only on quasirandomness of H_s). We fix a set $R\subseteq S(\widehat{H})$ with $|R|\le Q$, and consider the running of *PackingProcess* up to stage s. We want to show that it is very unlikely that R witnesses the failure of H_s to be quasirandom, since then the union bound over choices of R tells us that it is likely that H_s is quasirandom. In other words, we want to know that $\left|N_{H_s}(R)\right|$ is very likely close to the expected size. We write

$$\left|N_{H_s}(R)\right| = \left|N_{H_0}(R)\right| - Y_1 - \cdots - Y_s \,,$$

where $Y_i = \left|N_{H_i-1}(R)\right| - \left|N_{H_i}(R)\right|$ is the change at step i, and apply Corollary 4.2.2 to show that the sum $Y_1 + \cdots + Y_s$ is very likely to be close to its expectation. So proving Lemma 4.3.7 boils down to estimating accurately $\mathbb{E}(Y_i|H_{i-1})$. We now sketch how this goes.

Observe that Y_i is equal to the number of 1-stars in H_{i-1} whose leaves are R with centre $v \in \mathsf{N}_{H_{i-1}}(R)$ (that is, the hypergraph obtained from R by adding to each semi-edge of R the vertex v to make a hyperedge of H_{i-1}), at least one of whose hyperedges is used in embedding G_i to H_{i-1}. By linearity of expectation, $\mathbb{E}(Y_i|H_{i-1})$ is equal to the sum, over 1-stars in H_{i-1} whose leaves are R, of the probability that at least one edge in the star is used in embedding G_i. We will see that this probability is about the same for any given star S, and the problem is to calculate it. To do this we need to consider the running of *RandomEmbedding*.

First, we find accurate bounds on the probability any one of a small set U of vertices being used in a short time interval in *RandomEmbedding*. That is, we condition on a give ψ_{t-1} and suppose $U \cap \operatorname{im} \psi_{t-1} = \emptyset$. We then let *RandomEmbedding* embed the following vertices $t, t+1, \ldots, t+\varepsilon n - 1$. If we embedded each such vertex to an unused vertex uniformly at random, it would have a roughly $\frac{|U|}{n-t}$ chance of being embedded to U, and we would estimate a probability $\frac{\varepsilon n|U|}{n-t}$ of using any vertex of U during all the εn embeddings. Note that we do not consider in this heuristic either the fact that the number of used vertices is decreasing or that we might embed to two different vertices of U. These two facts do affect the probability, but (because ε is tiny) the effect is negligible. What is important is that of course we do not embed a vertex x uniformly to the unused vertices, but rather to the unused vertices in its candidate set, which intersects U in an unknown amount. We use the diet condition to control the number of vertices in the candidate set of x that are unused, and the cover condition to control (on average) the size of the intersection with U. A complication is that at time $t-1$ the $(\varepsilon, C\alpha, t)$-cover condition which we want to use has not been decided: to get around this, we suppose it is unlikely to fail.

Lemma 4.5.1. *For each $r \geq 2$, $D \in \mathbb{N}$ and $\gamma > 0$, let $Q, \delta, \alpha_0, \alpha_{s_{\max}}, C, \varepsilon$ be as in Setting 4.3.4. The following holds for any $\alpha_0 \leq \alpha \leq \alpha_{s_{\max}}$ and all sufficiently large n. Suppose that G is a hypergraph on $[n]$ such that $\deg^-(x) \leq D$ for each $x \in V(G)$,* *and H is an (α, Q)-quasirandom hypergraph with n vertices and $p\binom{n}{r}$ edges, with $p \geq \gamma$. Let $1 \leq k \leq 2C$ be fixed and $U \subseteq V(H)$ with $|U| = k$ be arbitrary. When* RandomEmbedding *is run to embed G into H, for any $1 \leq t \leq n + 1 - (\delta + \varepsilon)n$, if the history $\mathscr{H}_{t-1}$ up to and including embedding $t - 1$ is such that $|U \cap \operatorname{im} \psi_{t-1}| = 0$, the $(C\alpha, Q)$-diet condition holds for $(H, \operatorname{im} \psi_{t-1})$, and the probability, conditioned on $\mathscr{H}_{t-1}$, of the $(\varepsilon, C\alpha, t)$-cover condition failing is at most n^{-3}, we have*

$$\mathbb{P}\Big(\big|U \cap \operatorname{im} \psi_{t+\varepsilon n-1}\big| \geq 1 \Big| \mathscr{H}_{t-1}\Big) = (1 \pm 10C\alpha)\frac{k\varepsilon n}{n-t} .$$

Furthermore, for any $0 \leq q < \varepsilon n$, we have

$$\mathbb{P}\Big(\big|U \cap \operatorname{im} \psi_{t+q-1}\big| \geq 1 \Big| \mathscr{H}_{t-1}\Big) \leq 2\varepsilon\gamma^{D}\delta k .$$

The main technical difficulty in this proof is to rigorously allow for embedding to multiple vertices of U. To deal with this, we define the following *Modified RandomEmbedding*, which generates a sequence of embeddings with an identical distribution to *RandomEmbedding*, but which in addition generates a sequence of *reported* vertices. The modification we make is simple: at each time $1 \leq t' \leq n - \delta n$, *RandomEmbedding* chooses a vertex of $C^{t'-1}_{G\hookrightarrow H}(t') \setminus \operatorname{im} \psi_{t'-1}$. In *Modified RandomEmbedding*, we instead choose a vertex w of $C^{t'-1}_{G\hookrightarrow H}(t') \setminus (\operatorname{im} \psi_{t'-1} \setminus U)$, and report this vertex. If the reported vertex w is not in $\operatorname{im} \psi_{t'-1}$, we set $\psi_{t'} = \psi_{t'-1} \cup \{t' \hookrightarrow w\}$, as in *RandomEmbedding*. If the reported vertex is in $\operatorname{im} \psi_{t'-1}$ (which happens only if $w \in U$) we choose w' uniformly at random in $C^{t'-1}_{G\hookrightarrow H}(t') \setminus \operatorname{im} \psi_{t'-1}$, and set $\psi_{t'} = \psi_{t'-1} \cup \{t' \hookrightarrow w'\}$. It is not too hard to estimate the expected number of vertices of U reported, and easy to show that the contribution due to multiple reports of vertices of U is tiny. The probability of *RandomEmbedding* using v is the same as the probability that *Modified RandomEmbedding* reports v at least once, which we can thus calculate.

Proof of Lemma 4.5.1. Instead of *RandomEmbedding*, we consider *Modified RandomEmbedding* as defined above, which creates the same embedding distribution. For each i, let $r(i)$ be the vertex reported by *Modified RandomEmbedding* at time i.

Note that since the $(C\alpha, Q)$-diet condition holds for $(H, \operatorname{im} \psi_{t-1})$, for each $t \leq$

$x < t + \varepsilon n$, setting $S = \psi_{x-1}(\mathsf{N}^-(x))$, we have

$$\begin{aligned}\left|C^{x-1}(x) \setminus \operatorname{im}\psi_{x-1}\right| \pm k &= \left|\mathsf{N}_H(S) \setminus \operatorname{im}\psi_{t-1}\right| \pm \varepsilon n \pm k \\ &= (1 \pm C\alpha)p^{|\mathsf{N}^-(x)|}(n-t) \pm \varepsilon n \pm k \qquad (4.5.3) \\ &= (1 \pm 2C\alpha)p^{|\mathsf{N}^-(x)|}(n-t)\,.\end{aligned}$$

To begin with, we obtain an easy (but not very accurate) upper bound on the probability of a vertex $v \in U$ being used by a vertex x with $t \leq x < t + \varepsilon n$. By (4.5.3), the probability that any such x is embedded to v, conditioned on ψ_{x-1}, is at most $2\gamma^D \delta n^{-1}$, and so by the union bound the probability that some vertex x with $t \leq x < t + \varepsilon n$ is embedded to v is at most $2\varepsilon\gamma^D\delta$. Summing over $v \in U$ we obtain the 'furthermore' statement of the lemma.

We shall use the following two auxiliary claims.

Define E as the random variable counting the times when a vertex in U is reported by *Modified RandomEmbedding* in the interval $t \leq x < t + \varepsilon n$,

$$E = \left| \left\{ x \in [t, t+\varepsilon n) \ : \ r(x-1) \in U \right\} \right| \,.$$

The probability that *RandomEmbedding* uses vertices of U in the interval $t \leq x < t + \varepsilon n$, conditioning on $\mathscr{H}_{t-1}$, is equal to the probability that *Modified RandomEmbedding* reports a vertex of U at least once in that interval, which probability is by definition at least

$$\mathbb{E}\Big[E \mid \mathscr{H}_{t-1}\Big] - \sum_{\ell=2}^{\varepsilon n} \mathbb{P}\Big(\big|\{x \in [t, t+\varepsilon n) \ : \ r(x-1) = v\}\big| \geq \ell \mid \mathscr{H}_{t-1}\Big)\,.$$

Our first claim estimates $\mathbb{E}\big[E \mid \mathscr{H}_{t-1}\big]$.

Claim 4.5.2. *We have that*

$$\mathbb{E}\Big[E \mid \mathscr{H}_{t-1}\Big] = k\Big((1 \pm 4C\alpha)\frac{\varepsilon n}{n-t} \pm 4(D+1)\varepsilon^2\gamma^{-2D}\delta^{-2}\Big)\,.$$

Our second claim is that the sum in the expression above is small.

Claim 4.5.3. *We have that*

$$\sum_{\ell=2}^{\varepsilon n} \mathbb{P}\Big(\big|\{x \in [t, t+\varepsilon n) \ : \ r(x-1) = v\}\big| \geq \ell \mid \mathscr{H}_{t-1}\Big) \leq 8k^2\varepsilon^2\gamma^{-2D}\delta^{-2}$$
$$\leq 8kC\varepsilon^2\gamma^{-2D}\delta^{-2} \,.$$

By choice of ε, we have $64(D+1)C\varepsilon^2\gamma^{-2D}\delta^{-2} < C\alpha\varepsilon\delta^{-1}$, so putting the two claims together we have a proof of the Lemma.

Proof of Claim 4.5.2. Let us fix a vertex $v \in U$. By linearity of expectation, it is enough to estimate the random variable E' counting the number of times when v is reported by *Modified RandomEmbedding* in the interval $t \leq x < t + \varepsilon n$.

By linearity of expectation, we have

$$\begin{aligned} \mathbb{E}\Big[E' \mid \mathscr{H}_{t-1}\Big] &= \sum_{x=t}^{t+\varepsilon n-1} \mathbb{P}\big(v \text{ is reported at time } x \big| \mathscr{H}_{t-1}\big) \\ &= \sum_{x=t}^{t+\varepsilon n-1} \mathbb{E}\left[\frac{\mathbb{1}\{v \in C^{x-1}(x)\}}{|C^{x-1}(x) \setminus (\operatorname{im}\psi_{x-1} \setminus \{v\})|}\bigg|\mathscr{H}_{t-1}\right] \\ &= \sum_{x=t}^{t+\varepsilon n-1} \mathbb{E}\left[\frac{\mathbb{1}\{v \in C^{x-1}(x)\}}{|C^{x-1}(x) \setminus \operatorname{im}\psi_{x-1}| \pm k}\bigg|\mathscr{H}_{t-1}\right] . \end{aligned} \tag{4.5.4}$$

Using (4.5.3), we get

$$\mathbb{E}\Big[E' \mid \mathscr{H}_{t-1}\Big] = \sum_{x=t}^{t+\varepsilon n-1} \frac{\mathbb{P}\big(v \in C^{x-1}(x) \big| \mathscr{H}_{t-1}\big)}{(1 \pm 2C\alpha)p^{|\mathsf{N}^-(x)|}(n-t)} \,.$$

Splitting this sum up according to $|\mathsf{N}^-(x)|$, and again using linearity of expectation, we have

$$\mathbb{E}\Big[E' \mid \mathscr{H}_{t-1}\Big] = \sum_{d=0}^{D} \frac{\mathbb{E}\big(|\{x \in X_{t,d} : v \in C^{x-1}(x)\}| \big| \mathscr{H}_{t-1}\big)}{(1 \pm 2C\alpha)p^d(n-t)} \,.$$

We now need to estimate $\mathbb{E}\big(|\{x \in X_{t,d} : v \in C^{x-1}(x)\}| \big| \mathscr{H}_{t-1}\big)$. We do this by separating two cases. In the case that the $(\varepsilon, C\alpha, t)$-cover condition holds and $v \notin \operatorname{im}\psi_{t+\varepsilon n-1}$, the cover condition gives

$$|\{x \in X_{t,d} : v \in C^{x-1}(x)\}| = (1 \pm C\alpha)p^d|X_{i,d}| \pm \varepsilon^2 n \,.$$

If either the $(\varepsilon, C\alpha, t)$-cover condition fails, or $v \in \operatorname{im}\psi_{t+\varepsilon n-1}$, we use the trivial bound that the given set size is in $[0, \varepsilon n]$ since $|X_{t,d}| \leq \varepsilon n$, and write

$$|\{x \in X_{t,d} : v \in C^{x-1}(x)\}| = (1 \pm C\alpha)p^d|X_{i,d}| \pm \varepsilon^2 n \pm \varepsilon n\,.$$

The second case holds with probability at most $n^{-3} + 2\varepsilon\gamma^D\delta$, by our assumption on the likelihood of the cover condition failing and by our easy upper bound on the probability of embedding some x to v. Putting these together, we get

$$\begin{aligned}\mathbb{E}\big(|\{x \in X_{t,d} : v \in C^{x-1}(x)\}|\big|\mathscr{H}_{t-1}\big) &= (1 \pm C\alpha)p^d|X_{i,d}| \pm \varepsilon^2 n \pm \big(n^{-3} + 2\varepsilon\gamma^D\delta\big)\varepsilon n \\ &= (1 \pm C\alpha)p^d|X_{i,d}| \pm 3\gamma^D\delta\varepsilon^2 n\,.\end{aligned}$$

Substituting this in, we have

$$\begin{aligned}\mathbb{E}\Big[E' \mid \mathscr{H}_{t-1}\Big] &= \sum_{d=0}^{D} \frac{(1 \pm C\alpha)p^d|X_{t,d}| \pm 3\gamma^D\delta\varepsilon^2 n}{(1 \pm 2C\alpha)p^d(n-t)} \\ &= (1 \pm 4C\alpha)\tfrac{\varepsilon n}{n-t} \pm 4(D+1)\varepsilon^2\gamma^{-2D}\delta^{-2}\,,\end{aligned}$$

where the last equality uses $p \geq \gamma$ and $n - t \geq \delta n$.

By linearity of expectation, we have

$$\mathbb{E}\Big[E \mid \mathscr{H}_{t-1}\Big] = k\mathbb{E}\Big[E' \mid \mathscr{H}_{t-1}\Big]\,,$$

which proves the claim. □

Proof of Claim 4.5.3. Since the $(C\alpha, Q)$-diet condition holds for $(H, \operatorname{im}\psi_{t-1})$, since $p \geq \gamma$, and since $n - t \geq \delta n$, for each $x \in [t, t + \varepsilon n)$, when we embed x we report a uniform random vertex from a set of size at least $\frac{1}{2}\gamma^D\delta n$. The probability of reporting a vertex in U when we embed x is thus at most $2k\gamma^{-D}\delta^{-1}n^{-1}$, conditioning on $\mathscr{H}_{t-1}$ and any embedding of the vertices $[t, x)$. Since the conditional probabilities multiply, the probability that at each of a given ℓ-set of vertices in $[t, t + \varepsilon n)$ we report v is at most $2^\ell k^\ell \gamma^{-\ell D}\delta^{-\ell}n^{-\ell}$. Taking the union bound over choices of ℓ-sets, we have

$$\sum_{\ell=2}^{\varepsilon n}\mathbb{P}\Big(\big|\{x\in[t,t+\varepsilon n)\ :\ r(x-1)=v\}\big|\geq\ell\mid\mathscr{H}_{t-1}\Big)\leq$$
$$\leq\sum_{\ell=2}^{\varepsilon n}\binom{\varepsilon n}{\ell}2^{\ell}k^{\ell}\gamma^{-\ell D}\delta^{-\ell}n^{-\ell}\leq\sum_{\ell=2}^{\varepsilon n}\big(2k\varepsilon\gamma^{-D}\delta^{-1}\big)^{\ell}\leq\tfrac{4k^2\varepsilon^2\gamma^{-2D}\delta^{-2}}{1-2\varepsilon\gamma^{-D}\delta^{-1}}$$
$$\leq 8k^2\varepsilon^2\gamma^{-2D}\delta^{-2}\,,$$

where we use the bound $\binom{\varepsilon n}{\ell}\leq(\varepsilon n)^{\ell}$ and sum the resulting geometric series. ▫

□

We now extend this to deal with intervals of any length, and replace the single history in which the cover condition is not too likely to fail with a sufficiently large collection of histories.

Lemma 4.5.4. *Given $D\in\mathbb{N}$ and $\gamma>0$, let $Q,\delta,\alpha_0,\alpha_{s_{\max}},C,\varepsilon$ be as in Setting 4.3.4. Then the following holds for any $\alpha_0\leq\alpha\leq\alpha_{s_{\max}}$ and all sufficiently large n. Suppose that G is a hypergraph on $[n]$ such that $\deg^-(x)\leq D$ for each $x\in V(G)$, and H is an (α,Q)-quasirandom hypergraph with n vertices and $p\binom{n}{r}$ edges, with $p\geq\gamma$. Let $0\leq t_0<t_1\leq n-\delta n$. Let $\mathscr{L}$ be a history ensemble of* RandomEmbedding *up to time t_0, and suppose that $\mathbb{P}(\mathscr{L})\geq n^{-4C}$. Then the following hold for any fixed $1\leq k\leq 2C$ and set of vertices $U\subseteq V(H)$ with $|U|=k$. If $|U\cap\operatorname{im}\psi_{t_0}|=0$ then we have*

$$\mathbb{P}(|U\cap\operatorname{im}\psi_{t_1}|=0|\mathscr{L})=(1\pm 100Ck\alpha\delta^{-1})\big(\tfrac{n-1-t_1}{n-t_0}\big)^k\,.$$

We should note that the numerator $n-1-t_1$ might look strange, since ψ_{t_1} has $n-t_1$ unused vertices. The extra -1 is absorbed in the error term, and will make for neater cancellation in a future lemma.

Proof. We divide the interval $(t_0,t_1]$ into $\ell:=\lceil(t_1-t_0)/\varepsilon n\rceil$ intervals, all but the last of length εn. Let $\mathscr{L}_0:=\mathscr{L}$. Let, for each $1\leq i<\ell$, the set $\mathscr{L}_i$ be the embedding histories up to time $t_0+i\varepsilon n$ of *RandomEmbedding* which extend histories in $\mathscr{L}_{i-1}$ and are such that $|U\cap\psi_{t_0+i\varepsilon n}|=0$. Let $\mathscr{L}_\ell$ be the embedding histories up to time t_1 extending those in $\mathscr{L}_{\ell-1}$ such that $|U\cap\psi_{t_1}|=0$. Thus we have

$$\mathbb{P}(|U\cap\operatorname{im}\psi_{t_1}|=0|\mathscr{L})=\mathbb{P}(\mathscr{L}_\ell)/\mathbb{P}(\mathscr{L}_0)\,.$$

Finally, for each $1 \le i \le \ell$, let the set $\mathscr{L}'_{i-1}$ consist of all histories in $\mathscr{L}_{i-1}$ such that the $(C\alpha, Q)$-diet condition holds for $(H, \operatorname{im}\psi_{t_0+(i-1)\varepsilon n})$ and the probability that the $(\varepsilon, C\alpha, t_0+1+(i-1)\varepsilon n)$-cover condition fails, conditioned on $\psi_{t_0+(i-1)\varepsilon n}$, is at most n^{-3}. In other words, $\mathscr{L}'_i$ is the subset of $\mathscr{L}_i$ consisting of typical histories, satisfying the conditions of Lemma 4.5.1.

We now determine $\mathbb{P}(\mathscr{L}_\ell)$ in terms of $\mathbb{P}(\mathscr{L}_0)$, and in particular we show inductively that $\mathbb{P}(\mathscr{L}_i) > n^{-9C/2}$ for each i. Observe that for any time t, the probability (not conditioned on any embedding) that either the $(C\alpha, Q)$-diet condition fails for $(H, \operatorname{im}\psi_j)$ for some $j \le t$ or that the $(\varepsilon, C\alpha, t+1)$-cover condition has probability greater than n^{-3} of failing, is at most n^{-5C} by Lemma 4.4.3. In other words, for each i we have $\mathbb{P}(\mathscr{L}_i \setminus \mathscr{L}'_i) \le n^{-5C}$. Thus by Lemma 4.5.1 we have

$$\begin{aligned}
\mathbb{P}(\mathscr{L}_i) &= \big(1 - k(1 \pm 10C\alpha)\tfrac{\varepsilon n}{n-t_0-(i-1)\varepsilon n}\big)\mathbb{P}(\mathscr{L}'_{i-1}) \pm n^{-5C} \\
&= \big(1 - k(1 \pm 10C\alpha)\tfrac{\varepsilon n}{n-t_0-(i-1)\varepsilon n}\big)\big(\mathbb{P}(\mathscr{L}_{i-1}) \pm n^{-5C}\big) \pm n^{-5C} \\
&= \big(1 - k(1 \pm 20C\alpha)\tfrac{\varepsilon n}{n-t_0-(i-1)\varepsilon n}\big)\mathbb{P}(\mathscr{L}_{i-1}),
\end{aligned}$$

where the final equality uses the lower bound $\mathbb{P}(\mathscr{L}_{i-1}) \ge n^{-9C/2}$. Similarly, using the 'furthermore' statement of Lemma 4.5.1, we have

$$\mathbb{P}(\mathscr{L}_\ell) = \big(1 \pm 4\varepsilon\gamma^D\delta k\big)\mathbb{P}(\mathscr{L}_{\ell-1}).$$

Putting these observations together, we can compute $\mathbb{P}(\mathscr{L}_\ell)$:

$$\mathbb{P}(\mathscr{L}_\ell) = \big(1 \pm 4\varepsilon\gamma^D\delta k\big)\mathbb{P}(\mathscr{L}_0)\prod_{i=1}^{\ell-1}\Big(1 - k(1 \pm 20C\alpha)\tfrac{\varepsilon n}{n-t_0-(i-1)\varepsilon n}\Big).$$

Observe that the approximation $\log(1+x) = x \pm x^2$ is valid for all sufficiently small x. In particular, since $n - t_0 - (i-1)\varepsilon n \ge n - t_1 \ge \delta n$ and by choice of ε, for each i we have

$$\log\Big(1 - k(1 \pm 20C\alpha)\tfrac{\varepsilon n}{n-t_0-(i-1)\varepsilon n}\Big) = -k(1 \pm 30C\alpha)\tfrac{\varepsilon n}{n-t_0-(i-1)\varepsilon n}.$$

Thus we obtain

$$\begin{aligned}\log \mathbb{P}(\mathscr{L}_\ell) &= \log \mathbb{P}(\mathscr{L}_0) \pm 8\varepsilon\gamma^D\delta k - \sum_{i=1}^{\ell-1} k(1 \pm 30C\alpha)\tfrac{\varepsilon n}{n-t_0-(i-1)\varepsilon n} \\ &= \log \mathbb{P}(\mathscr{L}_0) \pm 8\varepsilon\gamma^D\delta k - k(1 \pm 40C\alpha)\int_{x=0}^{(\ell-1)\varepsilon n} \tfrac{1}{n-t_0-x}\,\mathrm{d}x \\ &= \log \mathbb{P}(\mathscr{L}_0) \pm 8\varepsilon\gamma^D\delta k - k(1 \pm 50C\alpha)\big(\log(n-t_0) - \log(n-1-t_1)\big) \\ &= \log \mathbb{P}(\mathscr{L}_0) + k\log\tfrac{n-1-t_1}{n-t_0} \pm 8\varepsilon\gamma^D\delta k \pm 50kC\alpha\log\delta^{-1}\,, \end{aligned} \tag{4.5.5}$$

where we use $t_1 \leq n - \delta n$, and we justify that the integral and sum are close by observing that for each i in the summation, if $(i-1)\varepsilon n \leq x \leq i\varepsilon n$ then we have

$$\tfrac{1}{n-t_0-i\varepsilon n} \leq \tfrac{1}{n-t_0-x} \leq \tfrac{1}{n-t_0-(i-1)\varepsilon n} \leq (1+\alpha)\tfrac{1}{n-t_0-i\varepsilon n}\,,$$

where the final inequality uses $n - t_0 - i\varepsilon n \leq n - t_1 \leq \delta n$ and the choice of ε. By choice of ε, and since $\delta^{-1} > \log\delta^{-1}$, this gives the lemma. Furthermore, (4.5.5), and the fact $t_1 \leq n-\delta n$, imply that $\mathbb{P}(\mathscr{L}_\ell) \geq n^{-9C/2}$. Since the $\mathscr{L}_i$ form a decreasing sequence of events the same bound holds for each $\mathscr{L}_i$. □

Next, suppose that F is a small labelled induced subhypergraph of G whose vertices are in $[n-\delta n]$, and suppose that we have a not necessarily induced copy of F in H, with $\varrho : V(F) \to V(H)$ the embedding witnessing this. Suppose in addition Z is a small set of vertices of H disjoint from $\varrho(V(F))$. We now estimate the probability that *RandomEmbedding* embeds F to F^*, preserving the labels (i.e. we produce an embedding extending ϱ) and does not use any vertex of Z. For the moment, the reader should think of F as being a single vertex.

The idea is the following. Suppose $F = \{u\}$ is a single vertex. We embed u to $\varrho(u)$ if and only if, as we run through $V(G)$, for each semi-edge $S \in \mathsf{N}^-_G(u)$, we embed S to a semi-edge which forms an edge of H together with $\varrho(u)$, all the while not using the vertices $\varrho(u)$ or Z. We then embed u to $\varrho(u)$, and complete the embedding without using Z. The point of phrasing it like this is that we can estimate accurately (using the diet condition) the probability of embedding $S \in \mathsf{N}^-_G(u)$ to a semi-edge which forms an edge of H with $\varrho(u)$: wherever we embed the first $r-2$ vertices of S, what matters is where we embed the final vertex to complete the embedding of S. So we can split up the embedding of G into a collection of intervals (in which we embed vertices which are neither u nor in the set $R(u)$ of maximum vertices of S) together with the embeddings of the vertices of $R(u)$ and finally of u. We can use Lemma 4.5.4 to estimate the chance that $\varrho(u)$ or Z is used in an interval. Finally, if we know that $\mathsf{N}^-_G(u)$ is embedded to a collection of semi-edges which all form edges with $\varrho(u)$ and $\varrho(u)$ is not used when we come to embed u, then we can also estimate accurately the probability of embedding u to $\varrho(u)$.

For more general F, there are two points to change in the above sketch. First, there are several different vertices in $V(F)$ and consequently we need to split the embedding up into more intervals. Second, it is no longer necessarily the case that for each $S \in \mathsf{N}^-_G(u)$ the embedding of the last vertex is what matters. Rather, what matters is the embedding of the last vertex of $S \setminus V(F)$. Note that if $S \setminus V(F) = \emptyset$, then $S \cup \{u\}$ is an edge of F.

In what follows, we also slightly extend the rôle of Z: we allow for considering only part of the run of *RandomEmbedding* up to some point $n - \tilde{\delta}n$, where $\delta \le \tilde{\delta} \le 1 - \delta$.

Lemma 4.5.5. *For each $r \ge 2$, $D \in \mathbb{N}$, and $\gamma > 0$, let constants $Q, \delta, \varepsilon, C, \alpha_0, \alpha_{s_{\max}}$ be as in Setting 4.3.4. Then the following holds for any $\alpha_0 \le \alpha \le \alpha_{s_{\max}}$ and all sufficiently large n. Suppose that G is a hypergraph on $[n]$ such that* $\deg^-(x) \le D$ *for each $x \in V(G)$, and H is an (α, Q)-quasirandom hypergraph with n vertices and $p\binom{n}{r}$ edges, with $p \ge \gamma$. Let $\delta \le \tilde{\delta} \le 1-\delta$. Let F be an induced subhypergraph of G with $0 \le |V(F)| \le C$, and suppose $V(F) \subseteq [n - \tilde{\delta}n]$. Let $\varrho : V(F) \to V(H)$ be an embedding of F. Finally let $Z \subseteq V(H) \setminus \varrho(V(F))$ be of size at most C. When* RandomEmbedding *is run to embed G into H, the probability that $\psi_{n-\tilde{\delta}n}$ extends ϱ and* $\operatorname{im}\psi_{n-\tilde{\delta}n} \cap Z = \emptyset$ *is*

$$\left(1 \pm 200C(|V(F)| + 2|Z|)\alpha\delta^{-1}\right)^{(2+2D)|V(F)|+1} p^{-|E(F)|} n^{-|V(F)|} \tilde{\delta}^{|Z|} .$$

Observe that the number of labelled copies of F in H is easily seen (by quasirandomness of H) to be roughly $p^{|E(F)|}n^{|V(F)|}$, so this lemma states that we embed F in G to a copy chosen (roughly) uniformly at random and this is roughly independent of the set of vertices left unused.

Proof. If $|V(F)| = 0$, the lemma statement follows directly from Lemma 4.5.4 with $t_0 = 0$ and $t_1 = n - \tilde{\delta}n$, so we now assume $|V(F)| \ge 1$. Let $V(F) = \{x_1, \dots, x_\ell\}$ in increasing order.

We define *relevant vertices* in the following way: for each edge $e \in E(G) \setminus E(F)$ whose final vertex is in $V(F)$, we let the last vertex in $e \setminus V(F)$ be a relevant vertex.

Let $z_1, \ldots, z_k$ be the increasing sequence of relevant vertices. Note that $k \leq D\ell$ as each vertex of F has at most D left neighbours. Let S_i denote the set of hyperedges $e \in E(G)$ such that $\max(e)$ is in $V(F)$, and $z_i = \max\big(e \setminus V(F)\big)$. We say that a partial embedding ψ_{z_i} of $[1, \ldots, z_i]$ is *good for* S_i if for each hyperedge $e \in S_i$, the set $\psi_{z_i}\big(e \setminus V(F)\big) \cup \varrho\big(e \cap V(F)\big)$ is a hyperedge of H.

Let $j_0 = 0$, and for each $1 \leq i \leq \ell - 1$, let j_i be the index such that $z_{j_i} < x_i < z_{j_{i+1}}$. Let $j_\ell = k$. Observe that the interval $\big[1, n - \tilde{\delta}n\big]$ is split by removing $\{x_1, \ldots, x_\ell, z_1, \ldots, z_k\}$ into $k + \ell + 1$ intervals $I_1, \ldots, I_{k+\ell+1}$ in increasing order. For convenience, when two vertices in $\{x_1, \ldots, x_\ell, z_1, \ldots, z_k\}$ are consecutive in $V(G)$, we say there is an interval between them, but it contains zero vertices (in order to justify having exactly $k + \ell$ intervals). Thus the interval I_1 is either $[1, z_1 - 1]$ or $[1, x_1 - 1]$, depending on whether $z_1 < x_1$ or not, and so on, and the final interval $I_{k+\ell+1}$ is the interval $[\max(V(F)) + 1, n - \tilde{\delta}n]$ (or the empty interval if $\max(V(F)) = n - \tilde{\delta}n$).

We consider the run of *RandomEmbedding* split up in this way, with the various intervals and the embeddings of the individual vertices $\{x_1, \ldots, x_\ell, z_1, \ldots, z_k\}$. We obtain an increasing sequence $t_0, \ldots, t_{2k+2\ell+1}$ of *critical times*, where $t_0 = 0$, $t_1 = \max(I_1)$ is the last vertex of I_1 (or 0 if $I_1 = \emptyset$), t_2 is whichever is smaller from x_1 and z_1, and so on. In general, t_{2i-1} will be the last vertex of I_i (if this interval is empty, we set $t_{2i-1} = t_{2i-2}$) and t_{2i} is the ith vertex of $\{x_1, \ldots, x_\ell, z_1, \ldots, z_k\}$. Correspondingly, we define a nested collection of events $\mathscr{L}_0, \ldots, \mathscr{L}_{2k+2\ell+1}$. The first, $\mathscr{L}_0$, is the sure event, and the last is the event that $\psi_{n-\tilde{\delta}n}$ extends ϱ and has image disjoint from Z whose probability we want to estimate. In general, we let $\mathscr{L}_i$ be the event that after embedding t_i it is not yet impossible to have the desired event. More formally, $\mathscr{L}_{2i-1}$ is the event that $\mathscr{L}_{2i-2}$ occurs and furthermore no vertex of I_i was embedded to $\varrho\big(V(F)\big) \cup Z$, and $\mathscr{L}_{2i}$ is the event that $\mathscr{L}_{2i-1}$ occurs and furthermore, if $t_{2i-1} = x_j$ then x_j is embedded to $\varrho(x_j)$, while if $t_{2i-1} = z_j$ then ψ_{z_j} is good for S_j and z_j is not embedded to any vertex of $\varrho\big(V(F)\big) \cup Z$.

Since these events are nested, we have

$$\mathbb{P}\big[\mathscr{L}_{2k+2\ell+1}\big] = \mathbb{P}[\mathscr{L}_0] \prod_{i=1}^{2k+2\ell+1} \mathbb{P}\big[\mathscr{L}_i \big| \mathscr{L}_{i-1}\big]\,, \tag{4.5.6}$$

and we will see that we can estimate each of the conditional probabilities in this product accurately. These estimates will depend on an assumption that $\mathbb{P}[\mathscr{L}_i] \geq n^{-2C}$, which we will justify by establishing this bound for the smallest event $\mathscr{L}_{2k+2\ell+1}$ at the end of the proof. We first state as a claim the bounds on the various $\mathbb{P}\big[\mathscr{L}_i\big|\mathscr{L}_{i-1}\big]$ which we need: there are three types, depending on whether t_i is the end of an interval, or some x_j, or some z_j. Though the formulae are somewhat complicated, we will see that most of the main terms end up cancelling. We need one more definition: for each i, we let $\ell_i := \big|V(F) \setminus [1, t_i]\big|$ be the number of vertices of F still to embed once t_i is embedded.

Claim 4.5.6. *If t_i is the end of the interval I_j, then we have*

$$\mathbb{P}\big[\mathscr{L}_i\big|\mathscr{L}_{i-1}\big] = \big(1 \pm 200C(\ell_i + |Z|)\alpha\delta^{-1}\big)\Big(\tfrac{n-1-t_i}{n-t_{i-1}}\Big)^{\ell_i+|Z|} . \tag{4.5.7}$$

If $t_i = z_j$, then we have

$$\mathbb{P}\big[\mathscr{L}_i\big|\mathscr{L}_{i-1}\big] = \big(1 \pm 100C\ell_i\alpha\big)p^{|S_j|} . \tag{4.5.8}$$

If $t_i = x_j$, then we have

$$\mathbb{P}\big[\mathscr{L}_i\big|\mathscr{L}_{i-1}\big] = \big(1 \pm 100C\alpha\big)\frac{1}{p^{\deg^-_G(x_j)}(n-t_i)} . \tag{4.5.9}$$

Before proving this claim, we use it to prove Lemma 4.5.5. We first point out which terms cancel in the product (4.5.6). If t_i is the end of the interval I_j, and so t_{i+2} is the end of the interval I_{j+1}, observe that $t_{i+1} = t_i + 1$ and so the numerator of the fraction in (4.5.7) for i is the same as the denominator in (4.5.7) for $i+2$. Thus these terms either cancel exactly (if $\ell_i = \ell_{i+2}$), or with an extra factor $n - 1 - t_i = n - t_{i+1}$ if $\ell_i = \ell_{i+2} + 1$. The latter case occurs exactly when $t_{i+1} = x_j$, in which case the factor $n - t_{i+1}$ cancels with the same factor in the denominator of (4.5.9) for $i + 1$. Putting these observations together, when we multiply together all the terms from the various cases of (4.5.7) and of (4.5.9) we have a collection of error terms, the term $n^{-\ell-|Z|}$ from the term $\mathbb{P}\big[\mathscr{L}_1\big|\mathscr{L}_0\big]$, a term $\prod_{j=1}^{\ell} p^{-\deg^-_G(x_j)}$ from the various cases of (4.5.9), and finally a term $\big(\tilde{\delta}n - 1\big)^{|Z|}$ from the final case $t_{2k+2\ell+1}$. The various terms from (4.5.8) simply give us a collection of error terms and the term $\prod_{j=1}^{k} m_j$. Putting all this together, from (4.5.6) we have

$$\begin{aligned}\mathbb{P}\big[\mathscr{L}_{2k+2\ell+1}\big] &= 1\cdot\left(\prod_{i=1}^{2k+2\ell+1}\big(1\pm 200C(\ell_i+|Z|)\alpha\delta^{-1}\big)\right)n^{-\ell-|Z|}\\ &\quad\cdot\prod_{j=1}^{\ell}p^{-\deg^-_G(x_j)}\prod_{j=1}^{k}p^{|S_j|}\big(\tilde{\delta}n-1\big)^{|Z|}\\ &= 1\cdot\left(\prod_{i=1}^{2k+2\ell+1}\big(1\pm 200C(\ell_i+2|Z|)\alpha\delta^{-1}\big)\right)n^{-\ell}\\ &\quad\cdot\prod_{j=1}^{\ell}p^{-\deg^-_G(x_j)}\prod_{j=1}^{k}p^{|S_j|}\tilde{\delta}^{|Z|}\,.\end{aligned}$$

Observe now that if e is an edge whose last vertex is in $V(F)$, then it is counted exactly once in $\sum_{j=1}^{\ell}\deg^-_G(x_j)$, namely for $x_j=\max(e)$. If $e\setminus V(F)$ is non-empty, it is also counted exactly once in $\sum_{j=1}^{k}|S_j|$, namely for $z_j=\max\big(e\setminus V(F)\big)$. If $e\setminus V(F)=\emptyset$ then e is not counted in $\sum_{j=1}^{k}|S_j|$, and if e is any edge whose last vertex is not in $V(F)$, then it is counted in neither sum. Thus we have $\sum_{j=1}^{\ell}\deg^-_G(x_j)-\sum_{j=1}^{k}|S_j|=e(F)$, and we get

$$\mathbb{P}\big[\mathscr{L}_{2k+2\ell+1}\big]=1\cdot\left(\prod_{i=1}^{2k+2\ell+1}\big(1\pm 200C(\ell+2|Z|)\alpha\delta^{-1}\big)\right)n^{-\ell}p^{-e(F)}\tilde{\delta}^{|Z|}\,,$$

which is as claimed in the lemma statement. What remains is to prove the Claim.

Proof of Claim 4.5.6. We begin with (4.5.7), which is given immediately by Lemma 4.5.4 with $U=\varrho\big(V(F)\setminus[1,t_{i-1}]\big)$, by our assumption $\mathbb{P}[\mathscr{L}_{i-1}]\geq n^{-2C}$.

To prove the other two statements of this Claim, we will first consider only histories in which the $(C\alpha,Q)$-diet condition holds for $(H,\operatorname{im}\psi_t)$ for each time t, and the probability of the $(\varepsilon,C\alpha,t)$-cover condition failing for any $\psi_{t+\varepsilon n-2}$ is less than n^{-4}. Observe that by Lemma 4.4.3, the total probability of histories not satisfying these conditions is at most n^{-4C}, which we will see can be absorbed in the error term.

We now prove (4.5.8). By the $(C\alpha,Q)$-diet condition, we embed z_j uniformly at random to a set of size $(1\pm C\alpha)p^{\deg^-_G(z_j)}(n-t_i-1)$. We need to estimate the number of these vertices which are good for S_j and not in $\varrho\big(V(F)\big)$. Observe that v is in $C^{z_j-1}(z_j)$ and good for S_j if and only if it is in the common neighbourhood of the semi-edge $\mathsf{N}^-_G(z_j)$, and also of all the semi-edges in the set

$$W := \Big\{\psi z_j - 1\big(e \setminus (V(F) \cup \{z_j\})\big) \cup \varrho\big(e \cap V(F)\big) \,:\, e \in S_j\Big\}\,.$$

Observe that, since we are in $\mathscr{L}_{i-1}$, the elements of W are sets each of $r-1$ vertices. This collection of semi-edges has size $\deg^-_G(z_j) + |S_j|$: that is, no two semi-edges listed are equal. This is automatic for $\mathsf{N}^-_G(z_j)$ and for W, because both of these sets are obtained from sets of edges of G by removing z_j, which is in all of them. So we just need to check that no edge in $\mathsf{N}^-_G(z_j)$ is also in $\{S \setminus \{z_j\} : S \in S_j\}$. Notice that every edge e of S_j has by definition its largest vertex in $V(F)$, and since z_j is not in $V(F)$ in particular the largest vertex of e is not embedded by ψ_{z_j-1}, whereas all vertices of each edge of $\mathsf{N}^-_G(z_j)$ are embedded by ψ_{z_j-1}. Thus e is not in $\mathsf{N}^-_G(z_j)$ as required.

Using the $(C\alpha, Q)$-diet condition with the set $\mathsf{N}^-_G(z_j) \cup W$, we see that the number of vertices of $C^{z_j-1}(z_j) \setminus \operatorname{im}\psi_{z_j-1}$ which are good for S_J is $(1 \pm C\alpha)p^{\deg^-_G(z_j)+|S_j|}(n-t_i-1)$. Taking account finally of the histories in which the diet or cover condition fails, we see

$$\mathbb{P}\big[\mathscr{L}_i \big| \mathscr{L}_{i-1}\big] = \frac{(1 \pm C\alpha)p^{\deg^-_G(z_j)+|S_j|}(n-t_i-1)}{(1 \pm C\alpha)p^{\deg^-_G(z_j)}(n-t_i+1)} \pm n^{-4C}\mathbb{P}[\mathscr{L}_{i-1}]^{-1}\,,$$

which by the assumption $\mathbb{P}[\mathscr{L}_{i-1}] \geq n^{-2C}$ gives (4.5.8).

Finally, we prove (4.5.9). When we embed x_j, provided $\mathscr{L}_{i-1}$ holds, we have $\varrho(x_j) \notin \operatorname{im}\psi_{x_j-1}$ and $\varrho(x_j) \in C^{x_j-1}(x_j)$. Thus we simply need to estimate the size of $|C^{x_j-1}(x_j) \setminus \operatorname{im}\psi_{t_i-1}|$ (to which we embed x_j uniformly), which we can do using the diet condition: it has size $(1 \pm C\alpha)p^{\deg^-_G(x_j)}(n-t_i+1)$. Taking account of the possibility of the cover or diet conditions failing, we get

$$\mathbb{P}\big[\mathscr{L}_i \big| \mathscr{L}_{i-1}\big] = \frac{1}{(1 \pm C\alpha)p^{\deg^-_G(x_j)}(n-t_i+1)} \pm n^{-4C}\mathbb{P}[\mathscr{L}_{i-1}]^{-1}\,,$$

which by the assumption $\mathbb{P}[\mathscr{L}_{i-1}] \geq n^{-2C}$ gives (4.5.9). ▫

□

The immediate case of this lemma which we need is the case that F is an edge of G contained in $[n - \delta n]$ and $Z = \emptyset$. We use this to estimate the probability that, again for fixed G and H, at least one edge in a given 1-star in H is used by *RandomEmbedding*.

Lemma 4.5.7. *For each $r \geq 2$, $D \in \mathbb{N}$, and $\gamma > 0$, let the constants $Q, \delta, \varepsilon, \alpha_0, \alpha_{s_{\max}}, C$ be as in Setting 4.3.4. Then the following holds for any $\alpha_0 \leq \alpha \leq \alpha_{s_{\max}}$ and all sufficiently large n. Suppose that G is a hypergraph on $[n]$ such that $\deg^-(x) \leq D$ for each $x \in V(G)$, with at least $n/4$ edges and maximum degree $\Delta(G) \leq n/\log n$, and H is an (α, Q)-quasirandom hypergraph with n vertices and $p\binom{n}{r}$ edges, where $p \geq \gamma$. Let $u_1, \ldots, u_k$ be semi-edges and v a vertex of H for some $k \leq Q$, and suppose $u_i v$ is an edge of H for each i. When* RandomEmbedding *is run to embed G into H, the probability that there is at least one $u_i v$ to which some edge of G is embedded is*

$$\left(1 \pm 1000Cr\alpha\delta^{-1}\right)^{2(D+1)r+1} p^{-1} n^{-r} \cdot r! k e(G') ,$$

where G' is the subhypergraph of G induced by $[n - \delta n]$.

Proof. Given $u_1, \ldots, u_k, v$ and G and H, let S be the event that there is at least one $u_i v$ to which some edge of G is embedded.

Fix any given $u_i v$. By Lemma 4.5.5, the probability that any one edge of G' is embedded to $u_i v$ in any given order is $(1 \pm 200Cr\alpha\delta^{-1})^{2(D+1)r+1} p^{-1} n^{-r}$. Since at most one edge in at most one order is so embedded, these events are disjoint, and summing over all edges of G' and all of the $r!$ possible orders, we see that the probability that some edge of G' is embedded to $u_i v$ in some order is

$$\left(1 \pm 200Cr\alpha\delta^{-1}\right)^{2(D+1)r+1} p^{-1} n^{-r} \cdot r! e(G') .$$

By linearity of expectation, the expected number of edges $u_i v$ embedded to by *RandomEmbedding* is, by Lemma 4.5.5 and linearity of expectation,

$$E := \left(1 \pm 200Cr\alpha\delta^{-1}\right)^{2(D+1)r+1} p^{-1} n^{-r} \cdot r! E(G') \cdot k ,$$

and by inclusion-exclusion, we have

$$E - \sum_{1 \leq i < i' \leq k} \mathbb{P}\big[u_i v \text{ and } u_{i'} v \text{ are embedded to by } \textit{RandomEmbedding}\big] \leq \mathbb{P}[S] \leq E .$$

We thus simply have to show that the above sum, which has $\binom{k}{2} \leq \binom{Q}{2}$ terms, is small. We will show that the probability of *RandomEmbedding* embedding to any two fixed edges $uv, u'v$ is small. This probability is equal to the sum over triples $x, x' \in S(G), y \in V(G)$ such that $xy, x'y \in E(G)$ of the probability that

$x \hookrightarrow u$, $x' \hookrightarrow u'$ and $y \hookrightarrow v$. For any given $y \in V(G)$ there are at most $\deg_G(y)^2$ choices of (x, x'). If $r = 2$ then by Lemma 4.2.3, there are at most $2Dn\Delta(G)$ such triples. If $r > 2$ we use the trivial upper bound $(re(G))^2$. It is now enough to make the estimate for one such triple. Assuming the $(C\alpha, Q)$-diet condition holds throughout *RandomEmbedding*, we embed each vertex of x, x' and y uniformly at random into a set of size at least $\frac{1}{2}p^D\delta n \geq \frac{1}{2}\gamma^D\delta n$, so the probability of the event $x \hookrightarrow u, x' \hookrightarrow u', y \hookrightarrow v$ is at most $8\gamma^{-(2r-1)D}\delta^{-(2r-1)}n^{-(2r-1)}$. Finally, the probability of the $(C\alpha, Q)$-diet condition failing for some $(H, \operatorname{im}\psi_i)$ is by Lemma 4.4.3 at most n^{-5C}. Putting this together, we have

$$\mathbb{P}[S] = \left(1 \pm 200Cr\alpha\delta^{-1}\right)^{2(D+1)r+1} p^{-1}n^{-r} \cdot r!ke(G') \\ \pm \binom{Q}{2} \sum_{y \in V(G)} \deg_G(y)^2 \cdot 8\gamma^{-(2r-1)D}\delta^{-(2r-1)}n^{-(2r-1)} \pm n^{-5C}\,. \quad (4.5.10)$$

Because $e(G') \geq n/(2r)$ the first term in the above is $\Theta(n^{-(r-1)})$, while since $\Delta(G) \leq n/\log n$ the other two terms are of asymptotically smaller order. Since n is sufficiently large, this gives the desired result. □

Finally, we use Lemma 4.5.7 to control the way in which subsets of common neighbourhoods shrink during the embedding process. This in particular is what we need to prove Lemma 4.3.7, though we will use the more general statement again to prove Theorem 4.7.1.

Lemma 4.5.8. *For each $r \geq 2$, $D \in \mathbb{N}$ and $\gamma > 0$, let the constants $Q, \delta, \varepsilon, \alpha_0, \alpha_{2n}, C$ be as in Setting 4.3.4. Then the following holds for any $0 \leq s \leq s' \leq s^*$. Suppose that when* PackingProcess *is run up to and including stage s. Let $k \leq Q$, and let $u_1, \ldots, u_k$ be semi-edges of H_s. Fix any $X \subseteq \mathsf{N}_{H_s}(u_1, \ldots, u_k)$ satisfying $|X| \geq \varepsilon n$. Then with probability at least $1 - n^{-5C}$, when* PackingProcess *is run further up to and including stage s', if it does not fail and if H_i is (α_i, Q)-quasirandom for each $s \leq i \leq s'$, we have*

$$\left|X \cap \mathsf{N}_{H_{s'}}(u_1, \ldots, u_k)\right| = \left(1 \pm \tfrac{1}{2}\alpha_{s'}\right)\left(\frac{p_{s'}}{p_s}\right)^k |X|\,.$$

Proof. We claim that, for a given $u_1, \ldots, u_k$, s and X, the probability of s' being the first time after s at which the given equation fails is at most $1 - n^{-6C}$, and by taking a union bound over the at most $s_{\max}$ choices of s' we conclude the lemma statement.

For each $s \le i \le s'$, let $X_i := X \cap \mathsf{N}_{H_i}(u_1, \ldots, u_k)$, and for each $s+1 \le i \le s'$ let $Y_i = \big|X_i \setminus X_{i-1}$. Thus $X_s = X$, and we want to know how much smaller $X_{s'}$ is than X_s. This quantity is simply $Y_{s+1} + \cdots + Y_{s'}$. We aim to apply Corollary 4.2.2 with $\mathcal{E}$ being the event that after each stage i with $s \le i \le s'$, the hypergraph H_i is (α_i, Q)-quasirandom, and we have

$$|X_i| = (1 \pm \tfrac{1}{2}\alpha_i)\Big(\frac{p_i}{p_s}\Big)^k |X|\,.$$

By definition, the failure event of the above claim is contained in $\mathcal{E}$, and so it suffices to estimate the probability that $\mathcal{E}$ occurs and the given equation does not hold. We therefore assume in the estimates which follow that we are in $\mathcal{E}$. The probability space in which we work is the set of all possible histories of *RandomEmbedding*, and the sequence of partitions required by Corollary 4.2.2 is given by the histories up to increasing times $s \le i \le s'$ of *RandomEmbedding*.

For each i, let G'_i be the subhypergraph of G_i induced by $[n - \delta n]$, and let p_i be such that $p_i\binom{n}{r} = e(H_i) = e(H_0) - \sum_{j=1}^{i} e(G'_j)$. Then by Lemma 4.5.7 and linearity of expectation, we have

$$\begin{aligned}\mathbb{E}[Y_i|H_{i-1}] =&(1 \pm \tfrac{1}{2}\alpha_{i-1})\Big(\frac{p_{i-1}}{p_s}\Big)^k |X|\big(1 \pm 1000Cr\alpha_{i-1}\delta^{-1}\big)^{(2D+2)r+1}\\ &\cdot p_{i-1}^{-1} n^{-r} r! k e(G'_i)\\ =&\big(1 \pm 2000Cr\alpha_{i-1}\delta^{-1}\big)^{(2D+2)r+1}|X| p_{i-1}^{k-1} p_s^{-k} n^{-r} r! k e(G'_i)\,.\end{aligned}$$

We now need to estimate the sum $\sum_{i=s+1}^{s'} \mathbb{E}(Y_i|H_{i-1})$, on the assumption that each H_{i-1} is (α_{i-1}, Q)-quasirandom. We first estimate the sum of the main terms, using the facts that $e(G'_i) = (p_{i-1} - p_i)\binom{n}{r}$, that $i \le s_{\max}$, that for any $x \in [0,1]$, any $0 < h < 1$ and any integer $a \ge 1$ we have $(x+h)^a - x^a = ahx^{a-1} \pm 2^a h^2$, and that $p_{i-1} - p_i \le 4Dn^{1-r}$:

$$\begin{aligned}\sum_{i=s+1}^{s'} |X| p_{i-1}^{k-1} p_s^{-k} n^{-r} r! k e(G'_i) &= \sum_{i=s+1}^{s'} k p_{i-1}^{k-1}(p_{i-1} - p_i) p_s^{-k}|X|(1 \pm rn^{-1})^r\\ &= |X|(1 \pm 2r^2 n^{-1}) p_s^{-k} \sum_{i=s+1}^{s'} \Big((p_{i-1}^k - p_i^k) \pm 16D^2 2^k n^{2-2r}\Big)\\ &= |X| p_s^{-k}(p_s^k - p_{s'}^k) \pm 64D^2 r^2 2^k n^{-1}|X| p_s^{-k}\,.\end{aligned}$$

Next, we estimate the sum of the error terms:

$$\begin{aligned}\sum_{i=s+1}^{s'} 10^6 CD^2 Qr^2 r!\delta^{-1}\alpha_{i-1}p_{i-1}^{k-1}p_s^{-k}e(G_i')n^{-r}|X| \leq \\ &\leq \int_{-\infty}^{s'} 10^7 CD^3 Qr^2 r!\delta^{-1}\gamma^{-2Q}n^{1-r}|X|\alpha_x \,\mathrm{d}x \\ &\leq \tfrac{1}{8}\alpha_{s'}\gamma^Q|X|\,,\end{aligned}$$

where we use $e(G_i') \leq Dn$ and the final inequality is by (4.3.1). Putting these two estimates together, we have

$$\sum_{i=s+1}^{s'} \mathbb{E}(Y_i|H_{i-1}) = \big(p_s^k - p_{s'}^k\big)p_s^{-k}|X| \pm \tfrac{1}{4}\alpha_{s'}\gamma^Q|X| := \tilde{\mu} \pm \tilde{\nu}\,.$$

The range of each Y_i is at most $k\Delta(G_i) \leq \frac{cQn}{\log n}$. We apply Corollary 4.2.2 with $\tilde{\varrho} = \varepsilon\gamma^Q|X|$ and $\mathcal{E}$ as above. We obtain that the probability that

$$\sum_{i=s+1}^{s'} Y_i \neq \big(p_s^k - p_{s'}^k\big)p_s^{-k}|X| \pm (\alpha_{s'}\gamma^Q|X|/4 + \varepsilon\gamma^Q|X|) = \big(p_s^k - p_{s'}^k\big)p_s^{-k}|X| \pm \tfrac{1}{2}\alpha_{s'}\gamma^Q|X|$$

is at most

$$2\exp\left(\frac{-\varepsilon^2\gamma^Q|X|^2}{2\frac{cQn}{\log n}(\tilde{\mu}+\tilde{\nu}+\tilde{\varrho})}\right) < n^{-6C}\,,$$

where the last inequality is by choice of $|X|$ and c.

Observe that if the above likely event occurs, we have

$$|X_{s'}| = |X| - \big(p_s^k - p_{s'}^k\big)p_s^{-k}|X| \pm \tfrac{1}{2}\alpha_{s'}\gamma^Q|X| = \left(\frac{p_{s'}}{p_s}\right)^k|X| \pm \tfrac{1}{2}\alpha_{s'}\gamma^Q|X|\,,$$

which implies the desired likely equation of the lemma. Taking the union bound over the at most $s_{\max}$ choices of s' we obtain the lemma statement. □

We are now in a position to prove Lemma 4.3.7.

Proof of Lemma 4.3.7. We define $\hat{p}$ by $e(H_0^*) = \hat{p}\binom{n}{r}$. By assumption we have either $\hat{p} = (1 \pm \eta)\gamma$ or $\hat{p} = 0$.

Our aim is to show that with high probability, for a given s, either *PackingProcess* fails before completing stage s or the pair (H_s, H_0^*) is (α_s, Q)-coquasirandom. Let S be a set of at most Q vertices in $V(H_0^*)$, and let $R \subseteq S$. Suppose that these two sets witness a first failure at some time s' of $(\alpha_{s'}, Q)$-coquasirandomness. Recall that for $(H_{s'}, H_0^*)$ to be $(\alpha_{s'}, Q)$-coquasirandom means that $\mathsf{N}_{H_{s'}}(R) \cap \mathsf{N}_{H_0^*}(S \setminus R)$ has about the size one would expect if both hypergraphs were random. Observe that if $\hat{p} = 0$ and $S \setminus R \neq \emptyset$ then $\mathsf{N}_{H_0^*}(S \setminus R)$ is empty, and consequently there is nothing to check. We therefore assume that either $\hat{p} = (1 \pm \eta)\gamma$ or $S \setminus R = \emptyset$ in what follows (and in the latter case we interpret the expression $\hat{p}^{|S \setminus R|}$ as 1).

We apply Lemma 4.5.8 with $s = 0$, with s' as above, with $u_1, \ldots, u_k$ the semi-edges of R, and with $X = \mathsf{N}_{H_0}(R) \cap \mathsf{N}_{H_0^*}(S \setminus R)$. Since (H_0, H_0^*) is $(\frac{1}{4}\alpha_0, Q)$-coquasirandom, this set X satisfies $|X| \geq \varepsilon n$ as required. We see that with probability at least $1 - n^{-5C}$ we have

$$\begin{aligned}\left|\mathsf{N}_{H_{s'}}(R) \cap \mathsf{N}_{H_0^*}(S \setminus R)\right| &= \left(1 \pm \tfrac{1}{2}\alpha_{s'}\right)\left(\tfrac{p_{s'}}{p_0}\right)^{|R|}|X| \\ &= \left(1 \pm \tfrac{1}{2}\alpha_{s'}\right)\left(1 \pm \tfrac{1}{4}\alpha_0\right)\left(\tfrac{p_{s'}}{p_0}\right)^{|R|} p_0^{|R|}\hat{p}_0^{|S \setminus R|}n \\ &= \left(1 \pm \alpha_{s'}\right)p_{s'}^{|R|}\hat{p}_0^{|S \setminus R|}n\,,\end{aligned}$$

which precisely says that we do *not* have a witness to a failure of $(\alpha_{s'}, Q)$-coquasirandomness of $(H_{s'}, H_0^*)$.

Taking the union bound over all choices of $R \subseteq S$ and S of size at most Q, and applying Lemma 4.4.3, we see that the following event has probability at most n^{2rQ-5C}. The pair (H_i, H_0^*) is (α_i, Q)-coquasirandom for each $0 \leq i \leq s-1$, but either *RandomEmbedding* fails to embed G_s or (H_s, H_0^*) is not (α_s, Q)-coquasirandom. Taking now the union bound over all choices of $1 \leq s \leq s_{\max}$, and recalling that (H_0, H_0^*) is by assumption $\left(\frac{1}{4}\alpha_0, Q\right)$-coquasirandom, we conclude that the probability that for any $1 \leq s \leq s_{\max}$, *RandomEmbedding* fails to embed G_s or the pair (H_s, H_0^*) fails to be (α_s, Q)-coquasirandom is at most n^{-4C}. This completes the proof. □

4.6 Completing spanning embeddings

Recall that we complete the embedding of each hypergraph G_s by embedding the final δn vertices using only edges of H_{s-1}^*. From Setting 4.3.4, these vertices form a strongly independent set and each of them has degree d_s. Lemma 4.3.9 states that it is very likely, provided *PackingProcess* does not fail and provided (H_s, H_0^*) is coquasirandom for each s, that only a few edges of H_0^* are used at any given semi-edge to form H_s^*, and hence (H_s, H_s^*) is also coquasirandom. Complementing this, Lemma 4.6.1 states that this coquasirandomness guarantees that completing the embedding is possible. We prove these two lemmas in this section, beginning with Lemma 4.3.10.

Recall that Lemma 4.3.10 states that it is likely that the partial embedding φ_s of each G_s provided by *RandomEmbedding* can be extended to an embedding φ_s^* of G_s, with the completion edges used for the extension lying in H^*. Since the neighbours of each of the last δn vertices of G_s are embedded by φ_s, the set of candidate vertices

$$C_s^*(x) := \big\{v \in V(H_{s-1}^*) \setminus \operatorname{im}\varphi_s : \varphi_s(y) \in \mathsf{N}_{H_{s-1}^*}(v) \text{ for each } y \in \mathsf{N}_{G_s}(x)\big\}$$

for each x of these last δn vertices in $V(H_{s-1}^*) \setminus \operatorname{im}\varphi_s$ are already fixed, and the desired φ_s^* exists if and only if there is a system of distinct representatives for the $C_s^*(x)$ as x ranges over the last δn vertices of G_s. Recall that Lemma 4.4.3 states in particular that $(H^*, \operatorname{im}\varphi_s)$ is likely to satisfy the $(2\eta, Q)$-codiet condition, which implies both that $C_s^*(x)$ is of size roughly $p^{d_s}\delta n$ for each of these last x, and also that the collection of sets is well-distributed (in a sense we will make precise later). We would like to apply Lemma 4.2.6, for which we need to prove that our bipartite graph is super-regular. The codiet condition provides the degree and codegree condition on one half of the graph, and the following lemma provides the other half.

Lemma 4.6.1. *For each $r \geq 2$, $D \in \mathbb{N}$ and $\gamma > 0$, let the constants be as in Setting 4.3.4. Suppose that G is a hypergraph on vertex set $[n]$, with $\deg^-(x) \leq D$ for each $x \in V(G)$, with maximum degree at most $cn/\log n$ and whose last δn vertices all have degree d, where $0 \leq d \leq D$, and form a strongly independent set. Suppose that H is an (α_{s^*}, Q)-quasirandom n-vertex hypergraph with $p\binom{n}{r}$ edges, where $p \geq \gamma$, and that H^* is a hypergraph on $V(H)$ with $(1 \pm \eta)\gamma\binom{n}{r}$ edges such that (H, H^*) forms an (η, Q)-coquasirandom pair. When* RandomEmbedding *is run to embed the first $n - \delta n$ vertices of G into H, with probability at least $1 - n^{-4C}$ we have that for all $v \in V(H^*) \setminus \operatorname{im}\psi_{n-\delta n}$*

$$\Big|\big\{x \in V(G) \,:\, n - \delta n < x \leq n,\ \psi_{n-\delta n}(\mathsf{N}^-(x)) \subseteq \mathsf{N}_{H^*}(v)\big\}\Big| = (1 \pm 10D\eta)\gamma^d\delta n\,.$$

The proof of this lemma is similar to the proof of Lemma 4.4.5. We use (ℓ, d)-patterns as described before the proof of Lemma 4.4.5.

Proof. Fix $v \in V(H^*)$, fix $1 \leq \ell \leq d$ and an (ℓ, d)-pattern $\mathbf{a}$. Let s_k denote $\sum_{i=1}^k a_i$. Let I be the last δn vertices of G, which by assumption form a strongly independent set. Let $X_\mathbf{a}$ denote the set of vertices from the last δn vertices of G that have associated pattern $\mathbf{a}$. For $x \in X_\mathbf{a}$ let $R(x)$ denote the set of vertices t, for which there is an edge $e \in G$ such that x is the last vertex in it and t is second to last. Let for each such t the set $E_{t,x}$ be the set of edges that caused t to be in $R(x)$. Given $t \in R(x)$ and an embedding ψ_{t-1} of $\{1, \ldots, t-1\}$, we say the *correct set* of t is

$$A_{x,t} = \left(\bigcap_{e \in E_{x,t}} \mathsf{N}_{H^*}\big(\psi_{t-1}(e \setminus \{t,x\}) \cup \{v\}\big) \right) \cap \left(V(H) \setminus \operatorname{im} \psi_{t-1}\right).$$

In other words, if at time $t-1$ the embedded part of $\mathsf{N}^-(x)$ has been embedded to $\mathsf{N}_{H^*}(v)$, then $A_{x,t}$ is the set of vertices to which we can embed t and maintain this property.

Given x among the last δn vertices of G and $1 \leq k \leq \ell$, suppose that t is the kth vertex of $R(x)$. We let $\mathsf{N}_k^-(x)$ denote the set of semi-edges in $\mathsf{N}(x)$ which are contained in $[t]$. Let $\mathcal{Y}_k$ denote the event that $\mathsf{N}_k^-(x)$ is embedded to $\mathsf{N}_{H^*}(v)$ for about as many x as one would expect. More formally, $\mathcal{Y}_k$ is the event that

$$\left|\left\{x \in I : \psi_{n-\delta n}\big(\mathsf{N}_k^-(x)\big) \subseteq \mathsf{N}_{H^*}(v)\right\}\right| = (1 \pm 10k\eta)\gamma^{s_k}\delta n. \tag{4.6.1}$$

Observe that what we want is to show that $\mathcal{Y}_\ell$ holds with sufficiently high probability for each choice of $v \notin \operatorname{im} \psi_{n-\delta n}$, since $s_\ell = d$ and $\ell \leq D$.

Let $\mathcal{B}$ be the event that the $(2\eta, Q)$-codiet condition fails at some time $t \leq n - \delta n$. We define the Bernoulli random variable $Z_{k,x,t}$ to be equal to 1 if t is the kth vertex of $R(x)$ and we have $\psi_{n-\delta n}\big(\mathsf{N}_{k-1}^-(x)\big) \subseteq \mathsf{N}_{H^*}(v)$, and zero otherwise. Note that $Z_{k,x,t}$ depends only on ψ_{t-1} and not the later part of the embedding. Let $Y_{k,x,t} := Z_{k,x,t} \cdot \mathbb{1}_{\psi_{n-\delta n}(t) \in A_{x,t}}$. That is, $Z_{k,x,t}$ tells us before embedding t if we have a chance for x to contribute to the set in (4.6.1), and $Y_{k,x,t}$ tells us if, having embedded t, it does contribute or not.

We want to show that if $\mathcal{Y}_{k-1}$ occurs, then $\mathcal{Y}_k$ is very likely to occur. We will then show this implies the lemma. Observe that $\mathcal{Y}_k$ is the event that

$$\sum_{t=1}^{n-\delta n} \sum_{x=n-\delta n+1}^{n} Y_{k,x,t} = (1 \pm 10k\eta)\gamma^{s_k}\delta n.$$

Furthermore, $\mathcal{Y}_{k-1}$ implies that $\sum_{t=1}^{n-\delta n} \sum_{x=n-\delta n+1}^{n} Z_{k,x,t} = (1 \pm 10(k-1)\eta)\gamma^{s_{k-1}}\delta n$.

We would like to calculate $\sum_{t=1}^{n-\delta n}\sum_{x=n-\delta n+1}^{n}\mathbb{E}(Y_{k,x,t}|\mathscr{H}_{t-1})$, where $\mathscr{H}_{t-1}$ denotes the embedding history of *RandomEmbedding* up to and including embedding $t-1$. Given a time t, if t is the kth vertex of $R(x)$, then at time $t-1$ the first $k-1$ vertices of $R(x)$ have already been embedded, so $Z_{k,x,t}$ is determined. Thus we have $\mathbb{E}(Y_{k,x,t}|\mathscr{H}_{t-1}) = \mathbb{P}\Big(\psi_t(t)\in A_{x,t}\big|\mathscr{H}_{t-1}\Big)\cdot Z_{k,x,t}$. Suppose that at time $t-1$ we have not seen a witness that $\mathcal{B}$ fails. If $v\in \operatorname{im}\psi_{t-1}$, then $v\in\operatorname{im}\psi_{n-\delta n}$ and there is nothing to prove. So we may suppose $v\notin\operatorname{im}\psi_{t-1}$. Using the $(2\eta,Q)$-codiet condition once with $S=\mathsf{N}^-(t)\cup\{\psi_{t-1}(e\setminus\{t,x\})\cup\{v\}:e\in E_{x,t}\}$ (these two sets are disjoint since $v\notin\operatorname{im}\psi_{t-1}$) and $R=\mathsf{N}^-(t)\subseteq S$ and once with $S=R=\mathsf{N}^-(t)$, we obtain

$$\mathbb{P}\Big(\psi_t(t)\in A_{x,t}\Big|\mathscr{H}_{t-1}\Big)=\frac{(1\pm2\eta)(1\pm\eta)\gamma^{a_k}p^{|\mathsf{N}^-(t)|}(n-t+1)}{(1\pm2\eta)p^{|\mathsf{N}^-(t)|}(n-t+1)}=(1\pm6\eta)\gamma^{a_k}\,.$$

Therefore, if $\overline{\mathcal{B}}$ and $\mathcal{Y}_{k-1}$ hold, we have

$$\sum_{t=1}^{n-\delta n}\sum_{x=n-\delta n+1}^{n}\mathbb{E}(Y_{k,x,t}|\mathscr{H}_{t-1})=(1\pm10(k-1)\eta)(1\pm6\eta)\gamma^{s_k}\delta n\,.$$

Applying Corollary 4.2.2 with $\tilde{\varrho}=\eta\gamma^k\delta n$, we deduce that the probability that $\mathcal{Y}_k$ fails is exponentially small. Indeed the probability that $\overline{\mathcal{B}}$ holds but $\sum_{t=1}^{n-\delta n}\sum_{x=n-\delta n+1}^{n}Y_{k,x,t}\neq(1\pm10k\eta)\gamma^k\delta n$ is at most $2\exp\big(-\frac{\eta^2\gamma^{2k}\delta^2n^2\log n}{2Dcn^2}\big)\le n^{-5C}$, where we use that $\sum_{x=n-\delta n+1}^{n}Y_{k,x,t}\le\deg(t)\le\frac{cn}{\log n}$.

As $\mathcal{Y}_0$ holds trivially with probability one, by a union bound over the choices of ℓ, $\mathbf{a}$, k and v we obtain that the probability that $\overline{\mathcal{B}}$ holds but there is some $1\le k\le d$ for which $\mathcal{Y}_k$ fails is at most n^{10-5C}. Finally, Lemma 4.4.3 states that $\mathcal{B}$ holds with probability at most n^{-5C}, giving the lemma statement by the union bound. □

Given a hypergraph G on $[n]$ whose last δn vertices are strongly independent, if *RandomEmbedding* is run to give an embedding φ of the first $n-\delta n$ vertices of G into the n-vertex hypergraph H, and H^* is a hypergraph on $V(G)$, we define the *completion graph* $C(G,H^*,\varphi)$ to be the bipartite graph with parts $X=V(G)\setminus[n-\delta n]$ and $Y=V(H)\setminus\operatorname{im}\varphi$, and an edge xy for $x\in X$ and $y\in Y$ whenever we have $\varphi\big(\mathsf{N}_G(x)\big)\subseteq NBH_{H^*}(y)$. Recall that each stage of *PackingProcess* first uses *RandomEmbedding* to embed the first $n-\delta n$ vertices of G_s into H_s, then picks a uniform random extension to complete the embedding of G_s using edges of H_s^*. In other words, what is picked is a uniform random perfect matching in

$C(G_s, \varphi, H_s^*)$. The following lemma guarantees that $C(G_s, \varphi, H_s^*)$ satisfies the conditions of Lemma 4.2.6 which tell us how the uniform random perfect matching behaves.

Lemma 4.6.2. *For each $r \geq 2$, $D \in \mathbb{N}$ and $\gamma > 0$, let the constants be as in Setting 4.3.4. Suppose that G is a hypergraph on vertex set $[n]$, with $\deg^-(x) \leq D$ for each $x \in V(G)$, with maximum degree at most $cn/\log n$ and whose last δn vertices all have degree d, where $0 \leq d \leq D$, and form a strongly independent set. Suppose that H is an (α_{s^*}, Q)-quasirandom n-vertex hypergraph with $p\binom{n}{r}$ edges, where $p \geq \gamma$, and that H^* is a hypergraph on $V(H)$ with $(1 \pm \eta)\gamma\binom{n}{r}$ edges such that (H, H^*) forms an (η, Q)-coquasirandom pair. When* RandomEmbedding *is run to give an embedding φ of the first $n - \delta n$ vertices of G into H, with probability at least $1 - 2n^{-4C}$ it succeeds and the completion graph $C(G, \varphi, H^*)$ satisfies the following. For every $x \in X$, we have $\deg(x) = (1 \pm 2\eta)\gamma^d\delta n$, and for all but at most $\frac{Dcn}{\log n}$ choices of x' we have in addition $\deg(x, x') = (1 \pm 2\eta)\gamma^{2d}\delta n$. For every $y \in Y$, we have $\deg(y) = (1 \pm 10D\eta)\gamma^d\delta n$.*

Proof. When *RandomEmbedding* is run to produce a partial embedding φ of G into H, by Lemma 4.4.3 with probability at least $1-n^{-5C}$ the algorithm succeeds and the triple $(H, H^*, \operatorname{im}\varphi)$ satisfies the $(2\eta, Q)$-codiet condition. By Lemma 4.6.1, with probability at least $1-n^{-4C}$ in addition we have, for every vertex v of $V(H^*) \setminus \operatorname{im}\varphi$,

$$\left|\left\{x \in V(G) \;:\; n-\delta n < x \leq n,\; \varphi(\mathsf{N}^-(x)) \subseteq \mathsf{N}_{H^*}(v)\right\}\right| = (1\pm 10D\eta)\gamma^d\delta n \;. \quad (4.6.2)$$

Suppose that both good events occur, which happens with probability at least $1 - 2n^{-4C}$. Let I denote the last δn vertices of G. Since the triple $(H, H^*, \operatorname{im}\varphi)$ satisfies the $(2\eta, Q)$-codiet condition we get, for every vertex x of I,

$$\left|\left\{v \in V(H^*) \setminus \operatorname{im}\varphi \;:\; \varphi(\mathsf{N}^-(x)) \subseteq \mathsf{N}_{H^*}(v)\right\}\right| = (1 \pm 2\eta)\gamma^d\delta n \;, \qquad (4.6.3)$$

and for every pair of vertices x, x' of I such that $\mathsf{N}_G(x) \cap \mathsf{N}_G(x') = \emptyset$ we get

$$\left|\left\{v \in V(H^*) \setminus \operatorname{im}\varphi \;:\; \varphi(\mathsf{N}^-(x)) \cup \varphi(\mathsf{N}^-(x')) \subseteq \mathsf{N}_{H^*}(v)\right\}\right| = (1 \pm 2\eta)\gamma^{2d}\delta n \;. \qquad (4.6.4)$$

For any given $x \in I$, the set $\mathsf{N}_G(x)$ contains at most D semi-edges. Since no semi-edge of G is in more than $\frac{cn}{\log n}$ edges of G, we see that there are at most $\frac{Dcn}{\log n}$ vertices $x' \in I$ such that $\mathsf{N}_G(x) \cap \mathsf{N}_G(x') \neq \emptyset$.

By definition of $C(G,\varphi,H^*)$, (4.6.2) gives the statement about degrees in Y, while (4.6.3) and (4.6.4) give the degree and codegree statements for X. □

The completion lemma, Lemma 4.3.10, now follows easily.

Proof of Lemma 4.3.10. The assumptions of Lemma 4.3.10 match those of Lemma 4.6.2, so with probability at least $1-2n^{-4C}$ *RandomEmbedding* succeeds in producing an embedding φ of the first $n-\delta n$ vertices of G into H, and the completion graph $C(G,\varphi,H^*)$ satisfies the degree and codegree conditions of that lemma.

We apply Lemma 4.2.6 with $F=F'=C(G,\varphi,H^*)$, with $m=\delta n$, with $\mu=\gamma^d$, with $\varepsilon=10D\eta$, and any valid $\varrho>0$. Since $\frac{m^2}{\log m}\geq\frac{\delta^2n^2}{\log n}>\frac{Dcn^2}{\log n}$, the degree and codegree conditions from Lemma 4.6.2 give us $(M1)$ and $(M1)$, while $(M3)$ is trivially true. Thus $C(G,\varphi,H^*)$ contains a perfect matching, which corresponds to the existence of an extension of φ to the desired φ^*. □

Finally, we prove Lemma 4.3.9. The idea is as follows. We fix a semi-edge y, give an upper bound for the expected number of edges used at y in each stage, and apply Corollary 4.2.2 to show that the actual outcome is with high probability not much larger than this upper bound. In order to obtain this upper bound on expectation, we will need to use again the observation just made that with probability at least $1-2n^{-4C}$ the output of *RandomEmbedding* gives us a completion graph $C(G,\varphi,H^*)$ which satisfies the conditions of Lemma 4.2.6.

For each $z\in S(G_s)$, we define the *completion degree* of z, written $\deg^*(z)$, to be the degree of z in the hypergraph G_s^c whose vertex set is $V(G_s)$ and whose edge set is all edges of G_s which have exactly one vertex in the final δn vertices and the other $r-1$ in the first $n-\delta n$ vertices. Then the number of edges of H_0^* at y used in stage s is $\deg_{G_s}(z)$ where z is the semi-edge of G_s embedded to y. Note that since every edge of G_s^c has exactly one vertex among the final δn vertices, and each of these final δn vertices is in d_s edges, we have

$$\sum_{y\in S(G_s^c)}\deg^*(y)=r\delta n d_s\ . \tag{4.6.5}$$

Let S_0 denote the semi-edges in G_s^c that have no vertices in the last δn, S_1 those that have one vertex in the last δn and S_2 those that have at least two.

Proof of Lemma 4.3.9. Fix $y\in S(H_0^*)$. For each $s\in[s^*]$, let Y_s be the number of edges of H_0^* at y used in stage s. For $z\in S_2$ we have $\deg^*(z)=0$ since the last δn

vertices of G_c^s are strongly independent, so we have

$$Y_s = \sum_{z\in S(G_s^c)} \deg^*(z)\mathbb{1}_{z\hookrightarrow y} = \sum_{z\in S_0} \deg^*(z)\mathbb{1}_{z\hookrightarrow y} + \sum_{z\in S_1} \deg^*(z)\mathbb{1}_{z\hookrightarrow y} + \sum_{z\in S_2} 0\cdot \mathbb{1}_{z\hookrightarrow y} \tag{4.6.6}$$

We define $\mathcal{E}_s$ to be the event that *PackingProcess* completes stages up to and including $s-1$, that (H_{s-1}, H_0^*) is (α_{s-1}, Q)-coquasirandom, and that for each $z \in S(H_0^*)$ we have $\deg_{H_0^*}(z) - \deg_{H_{s-1}^*}(z) \le 10r!\gamma^{-D}D\delta n$. Observe if $\mathcal{E}_s$ holds, that if $R \subseteq S$ are any two collections of semi-edges of H_{s-1}^* with $|S| \le Q$, since (H_{s-1}, H_0^*) is (α_{s-1}, Q)-coquasirandom, we have

$$\left|\mathsf{N}_{H_{s-1}}(R) \cap \mathsf{N}_{H_0^*}(S\setminus R)\right| = (1\pm\alpha_{s-1})p_{s-1}^{|R|}\gamma^{|S\setminus R|}n\,.$$

We therefore have

$$\left|\mathsf{N}_{H_{s-1}}(R) \cap \mathsf{N}_{H_{s-1}^*}(S\setminus R)\right| = (1\pm\eta)p_{s-1}^{|R|}\gamma^{|S\setminus R|}n\,,$$

since the neighbourhood in question is smaller in size by at most $10r!\gamma^{-D}D^2\delta n$; by choice of δ this is tiny compared to $\eta\gamma^Q n$. Thus (H_{s-1}, H_{s-1}^*) is also (η, Q)-coquasirandom when $\mathcal{E}_s$ holds. Thus $\mathcal{E}_s$ implies the required coquasirandomness.

Suppose that $\mathscr{H}_{s-1}$ is an arbitrary history of *PackingProcess* up to and including stage $s-1$ which is in $\mathcal{E}_s$. We begin by estimating $\mathbb{E}(Y_s|\mathscr{H}_{s-1})$.

To estimate the desired expectation, we use Lemma 4.5.5. If $z \in S_0$, then we need to consider the $(r-1)!$ possible ways that z can be mapped to y, each of which mappings actually occurs with probability at most $2n^{r-1}$. If $z \in S_1$, then let z' denote the $r-2$ vertices of Z which are in the first δn vertices and let $z'' = z \setminus z'$. There are $(r-1)!$ different ways to map z' to y; fix one such and let y'' be the vertex of y not mapped to. By Lemma 4.5.5, the probability that z' is embedded to y and y'' is not in the image of φ when we run *RandomEmbedding*is at most $2n^{r-2}\delta$. In order for z to then be mapped to y, we need that the final vertex z'' of z is embedded by the random matching to the final vertex y'' of y. If $y''z''$ is not an edge of $C(G_s, \varphi_s, H_{s-1}^*)$, then this probability is zero. Otherwise, it is the probability that $y''z''$ is used in a uniform random perfect matching of $C(G_s, \varphi_s, H_{s-1}^*)$. Since $\mathscr{H}_{s-1}$ is in $\mathcal{E}_s$, as in the proof of Lemma 4.3.10, with probability at least $1-2n^{-4C}$ *RandomEmbedding* produces φ_s such that the conditions of Lemma 4.2.6 are met for $C(G_s, \varphi_s, H_{s-1}^*)$, and in this case the probability that the chosen perfect matching uses $x''y''$ is at most $\frac{2}{\gamma^{d_s}\delta n}$, where d_s is the degree of the last δn vertices of G_s.

Putting this together, we obtain

$$\mathbb{P}[z \hookrightarrow y|\mathscr{H}_{s-1}] \le 2(r-1)!n^{-(r-1)} \quad \text{if } z \in S_0, \text{ and} \tag{4.6.7}$$

$$\begin{aligned}\mathbb{P}[z \hookrightarrow y|\mathscr{H}_{s-1}] &\le 4(r-1)!\gamma^{-d_s}n^{-(r-1)} + 2n^{-4C} \\ &\le 5(r-1)!\gamma^{-d_s}n^{-(r-1)} \quad \text{if } z \in S_1\,.\end{aligned} \tag{4.6.8}$$

For (4.6.8), the term n^{-4C} covers the possibility that *RandomEmbedding* produces φ_s which does not give a completion graph with the required degree and codegree conditions.

We now put (4.6.5), (4.6.7) and (4.6.8) into (4.6.6). It follows that when $\mathscr{H}_{s-1}$ is in $\mathcal{E}_s$ we have

$$\mathbb{E}[Y_s|\mathscr{H}_{s-1}] \le 5(r-1)!\gamma^{-d_s}n^{-(r-1)} \cdot r\delta n d_s \le 5r!D\gamma^{-D}\delta n^{-(r-2)}\,.$$

Given any $1 \le s' \le s^*$, we estimate the probability that $\mathcal{E}_s$ holds for each $1 \le s \le s'$ but we have $\deg_{H_0^*}(y) - \deg_{H_{s'}^*}(y) > 10r!\gamma^{-D}D\delta n$.

Since $0 \le Y_s \le \Delta(G_s)$ holds for each s, and since $s' \le s_{\max}$, we can apply Corollary 4.2.2, with $\mathcal{E} = \bigcap_{s=1}^{s'} \mathcal{E}_s$, to give

$$\mathbb{P}\Big[\mathcal{E} \text{ and } \sum_{i=1}^{s'} Y_s > 10r!D\gamma^{-D}n^{-(r-2)} \cdot s_{\max}\Big] \le \exp\big(-\tfrac{20r!D\gamma^{-D}n^{-(r-2)}s_{\max}}{\Delta(G_s)}\big) < n^{-4C}\,,$$

where the final inequality is since $\Delta(G_s) \le cn/\log n$ and by choice of c. Taking the union bound over all choices of s' and y, we see that the probability that neither of the first two events of the lemma statement occurs, and yet $\mathcal{E}_s$ fails for any s, is at most n^{-3C} as required. □

4.7 Quasirandom packing

The main theorem of this section is the following, which states that if we want to pack almost-spanning hypergraphs, we can obtain a packing with several additional quasirandomness properties. To state these conveniently, we need a few definitions.

We first define weight functions on our guest graphs G_s. Given a nonnegative integer q and an r-uniform hypergraph G_s whose vertex set $V(G_s)$ is $[v(G_s)]$ (i.e. it is the first $v(G_s)$ natural numbers), we say ω_s is a *q-weight function on G_s* when ω_s is a function whose domain is $\binom{V(G_s)}{q}$, the q-vertex subsets of $V(G_s)$, and whose codomain is $\mathbb{R}_{\geq 0}$. The *support* of ω_s is the set $\big\{\mathbf{f} \in \binom{V(G_s)}{q} \,:\, \omega_s(\mathbf{f}) > 0\big\}$, and we say that ω_s is *free* if all q-sets in its support contain no edges of G_s. We will say that the support of ω_s is contained in a set $X \subseteq V(G_s)$ if each q-set in the support is in X. We write

$$\Delta(\omega_s) := \max_{x \in V(G_s)} \sum_{\substack{\mathbf{f} \in \binom{V(G_s)}{q} \\ x \in \mathbf{f}}} \omega_s(\mathbf{f}) \quad \text{and} \quad \mathrm{sum}(\omega_s) := \sum_{\mathbf{f} \in \binom{V(G_s)}{q}} \omega_s(\mathbf{f})\,.$$

Given an embedding $\varphi_s : V(G_s) \to V(H)$, where H is the r-uniform host hypergraph, and an ordered q-set $\mathbf{e}$ of distinct vertices of H, we define $\omega_s(\mathbf{e})$ to be the weight φ_s assigns to $\mathbf{e}$, with the natural order on G_s corresponding to the order of $\mathbf{e}$. More formally, if $\varphi_s^{-1}(e_1) < \varphi_s^{-1}(e_2) < \cdots < \varphi_s^{-1}(e_q)$, we set

$$\omega_s(\mathbf{e}) = \omega_s\big(\{\varphi_s^{-1}(e_1), \ldots, \varphi_s^{-1}(e_q)\}\big)\,.$$

If some e_i is not in the image of φ_s, or if one of the above inequalities does not hold, we set $\omega_s(\mathbf{e}) = 0$.

Given an r-uniform hypergraph H, a vertex $v \in V(H)$, and a collection $X_1, \ldots, X_{r-1}$ of subsets of $V(H)$, we write $\mathsf{N}_H(v; X_1, \ldots, X_{r-1})$ for the collection of ordered $(r-1)$-sets whose ith element is in X_i for each $i \in [r-1]$ and which are (as unordered sets) contained in $\mathsf{N}_H(v)$.

Theorem 4.7.1. *For each $r \geq 2$, each $\gamma > 0$ and each $D, K \in \mathbb{N}$ there exist numbers $n_0 \in \mathbb{N}$ and $\delta, c, \xi > 0$ with $\delta \leq \gamma$ such that the following holds for each $n > n_0$, with $Q = (r+1)D + 3$. Suppose that $\widehat{H}$ is an (ξ, Q)-quasirandom r-uniform hypergraph with n vertices and $\hat{p}\binom{n}{r}$ edges, where $\hat{p} > 0$. Suppose that $s^* \leq 2n^{r-1}/(r-1)!$ and that for each $s \in [s^*]$ the r-uniform hypergraph G_s is on vertex set $[n - \delta n]$, with maximum degree at most $\frac{cn}{\log n}$, such that $\deg^-(x) \leq D$ for each $x \in V(G_s)$. Let p satisfy $p\binom{n}{r} = \hat{p}\binom{n}{r} - \sum_{s=1}^{s^*} e(G_s)$, and suppose $p \geq \gamma$. Given any $0 \leq q \leq Q$, let $\omega_s : \binom{V(G_s)}{q} \to \big[0, \frac{cn}{\log n}\big]$ be a q-weight function on G_s for each $s \in [s^*]$. Let $X, X_1, \ldots, X_q \subseteq V(\widehat{H})$ satisfy $|X|, |X_1|, \ldots, |X_q| \geq \gamma n$, and suppose that for each $1 \leq i \leq q$ the pair $\big(\widehat{H}, V(\widehat{H}) \setminus X_i\big)$ satisfies the (ξ, Q)-diet condition.*

There is a randomised algorithm (whose running is unaffected by the choice of $q, \omega_s, X, X_1, \ldots, X_q$) which with probability at least $1 - n^{-K}$ returns maps $\varphi_s : G_s \to \widehat{H}$ for each $s \in [s^*]$ which form a packing of $(G_s)_{s \in [s^*]}$ into $\widehat{H}$, such that the hypergraph H of leftover edges is (γ, Q)-quasirandom, and in addition the *following hold for every $\delta \le \tilde{\delta} \le 1-\delta$, every $U \subseteq V(\widehat{H})$ with $|U| \le Q$, and every collection $T_1, \dots, T_\ell$ of distinct semi-edges in $V(\widehat{H})$ with $0 \le \ell \le Q$.*

(QUASI 1) *For every $S \subseteq [s^*]$ with $|S| \le Q$, if $\left|X \cap \mathsf{N}_{\widehat{H}}(T_1, \dots, T_\ell)\right| \ge \frac{1}{2}\gamma^Q n$, we have*

$$\left|X \cap \mathsf{N}_H(T_1,\dots,T_\ell) \setminus \bigcup_{s\in S} \varphi_s([n-\tilde{\delta}n])\right| = \\ = (1\pm\gamma)\tilde{\delta}^{|S|}\left(\tfrac{p}{\hat{p}}\right)^\ell \left|X \cap \mathsf{N}_{\widehat{H}}(T_1,\dots,T_\ell)\right| .$$

(QUASI 2) *Suppose that for each $s \in [s^*]$ the weight function ω_s is free. Then for every ordered q-set $\mathbf{e}$ in $V(\widehat{H}) \setminus U$, if $\sum_{s=1}^{s^*} \mathrm{sum}(\omega_s) \ge \gamma n^{q+1}$, we have*

$$\sum_{s=1}^{s^*} \omega_s(\mathbf{e}) \mathbb{1}_{U \cap \varphi_s([n-\tilde{\delta}n]) = \emptyset} = (1\pm\gamma)\tilde{\delta}^{|U|} n^{-q} \sum_{s=1}^{s^*} \mathrm{sum}(\omega_s) .$$

(QUASI 3) *Suppose that $q = r-1 \ge 2$. For every $s \in [s^*]$ and $v \in V(\widehat{H})$, if $\mathrm{sum}(\omega_s) \ge \gamma n$, the support of ω_s is contained in $[n-\tilde{\delta}n]$, and $\Delta(\omega_s) \le \frac{cn}{\log n}$, then we have either $v \in \varphi_s([n-\tilde{\delta}n])$ or*

$$\sum_{\mathbf{e} \in \mathsf{N}_H(v; X_1, \dots, X_{r-1})} \omega_s(\mathbf{e}) = (1\pm\gamma)\frac{p}{\hat{p}} \cdot \left|\mathsf{N}_{\widehat{H}}(v; X_1,\dots,X_{r-1})\right| \\ \cdot n^{1-r} \mathrm{sum}(\omega_s) .$$

Observe that (QUASI 2) for $q = 0$ does have content. There is only one ordered 0-set of vertices of G_s; if for example we choose to give it weight 1 in each G_s, then (QUASI 2) counts the number of s such that U is disjoint from $\varphi_s([n-\tilde{\delta}n])$.

By taking the union bound, we can ask for a packing in which our properties hold for polynomially many different choices of $q, \omega_s, X, X_1, \dots, X_q$ simultaneously.

The proof of Theorem 4.7.1 mainly consists of using the tools developed in Sections 4.4 and 4.5. As mentioned in the sketch in Section 4.3.1, we use the cases of the various lemmas in which H_0^* has no edges, and $H_0 = \widehat{H}$. That is, the randomised algorithm of Theorem 4.7.1 is the following *PackingProcess2* and the various lemmas of the preceding three sections apply to it.

Proof of Theorem 4.7.1. To begin with, we prove that *PackingProcess2* is likely to succeed. This is much the same as the corresponding proof for Theorem 4.3.1,

Algorithm 3: *PackingProcess2*

Input : hypergraphs $G_1, \dots, G_{s^*}$, with G_s on vertex set $[n - \delta n]$; a hypergraph $\widehat{H}$ on n vertices

let $H_0 = \widehat{H}$;

for $s = 1$ **to** s^* **do**

  run *RandomEmbedding*(G_s,H_{s-1}) to get an embedding φ_s of G_s into H_{s-1};

  let H_s be the hypergraph obtained from H_{s-1} by removing the edges of $\varphi_s(G_s)$;

end

so we omit the details. In order that *PackingProcess2* does not succeed, there must exist some smallest $1 \le s \le s^*$ such that *PackingProcess2* runs up to and including stage $s-1$, and $H_{s'}$ is $(\alpha_{s'}, Q)$-quasirandom for each $1 \le s' \le s-1$, but either *RandomEmbedding* fails to embed G_s, or H_s is not (α_s, Q)-quasirandom. By Lemma 4.3.8, the probability of *RandomEmbedding* failing to embed G_s is at most n^{-5C}, and by Lemma 4.3.7, if G_s is successfully embedded, the probability that H_s is not (α_s, Q)-quasirandom is at most n^{-4C}. Taking the union bound over the choices of s, we see that the probability of either failure at any stage is at most n^{-3C}. In what follows, we refer to the likely event as PackingProcess2 *behaves* and we will generally assume it occurs.

We next prove (*QUASI* 1), with the error bound $\frac{1}{2}\gamma$ in place of γ. Note that the claim that H is (γ, Q)-quasirandom is a special case of this statement obtained by taking $X = V(\widehat{H})$ and $S = \emptyset$, since for any given collection of distinct $(r-1)$-sets $T_1, \dots, T_\ell$, where $\ell \le Q$, we have $\left|\mathsf{N}_{\widehat{H}}(T_1, \dots, T_\ell)\right| = (1 \pm \xi)\hat{p}^\ell n > \frac{1}{2}\gamma^Q n$.

Given a set X, we fix $S = \{s_1, \dots\}$ in increasing order, ℓ and $T_1, \dots, T_\ell$, and $\tilde{\delta}$, as in the statement; we assume that $\left|X \cap \mathsf{N}_{\widehat{H}}(T_1, \dots, T_\ell)\right| \ge \frac{1}{2}\gamma^Q n$ otherwise there is nothing to prove. The idea is the following. We split up the running of *PackingProcess2* into several parts: the interval before embedding G_{s_1}, embedding G_{s_1}, the interval between G_{s_1} and G_{s_2}, and so on. We use Lemma 4.5.8 to determine how $X \cap \mathsf{N}_{H_j}(T_1, \dots, T_\ell)$ shrinks as j increases through the first interval, how $X \cap \mathsf{N}_{H_j}(T_1, \dots, T_\ell) \setminus \varphi_{s_1}([n - \tilde{\delta} n])$ shrinks as j increases through the second interval, and so on. We use Lemma 4.4.3 to determine how much of each set is covered by the next $\operatorname{im} \varphi_{s_i}([n - \tilde{\delta} n])$, and observe that embedding any one graph can only change $\mathsf{N}_{H_j}(T_1, \dots, T_\ell)$ by a tiny amount.

Claim 4.7.2. *For each* $-1 \le i \le |S|$ *we have with probability at least* $1-3(i+1)n^{-5C}$

$$\Big|X \cap \mathsf{N}_{H_{s_{i+1}-1}}(T_1,\dots,T_\ell) \setminus \bigcup_{j=1}^{i} \varphi_{s_j}([n-\tilde{\delta}n])\Big| = (1 \pm C\alpha_{s_{i+1}-1})^{3i+3}\tilde{\delta}^{\max(0,i)}\left(\frac{p_{(s_{i+1}-1)}}{\hat{p}}\right)^{\ell} \cdot \big|X \cap \mathsf{N}_{\widehat{H}}(T_1,\dots,T_\ell)\big|\,,$$

where we set $s_{|S|+1|} := s^*+1$ *and* $s_0 = 1$.

Proof. The statement for $i=-1$ is a triviality. Suppose now that $i \ge 0$ and the statement holds for $i-1$. Let

$$Z = X \cap \mathsf{N}_{H_{s_i-1}}(T_1,\dots,T_\ell) \setminus \bigcup_{j=1}^{i-1} \varphi_{s_j}([n-\tilde{\delta}n])\,,$$
$$Z' = X \cap \mathsf{N}_{H_{s_i-1}}(T_1,\dots,T_\ell) \setminus \bigcup_{j=1}^{i} \varphi_{s_j}([n-\tilde{\delta}n]) \quad \text{and}$$
$$Z'' = X \cap \mathsf{N}_{H_{s_i}}(T_1,\dots,T_\ell) \setminus \bigcup_{j=1}^{i} \varphi_{s_j}([n-\tilde{\delta}n])\,.$$

If $i=0$, then we have $Z=Z'$. If $i>0$, by Lemma 4.4.3(e) with the given Z, with probability at least $1-n^{-5C}$, we have $|Z'| = \big|Z \setminus \varphi_{s_i}([n-\tilde{\delta}n])\big| = (1\pm C\alpha_{s_i})\tilde{\delta}|Z|$. In either case, we conclude that with probability at least $1-n^{-5C}$ we have

$$|Z'| = (1 \pm C\alpha_{s_i})^{3i+1}\tilde{\delta}^{\max(0,i)}\left(\frac{p_{s_i-1}}{\hat{p}}\right)^{\ell} \cdot \big|X \cap \mathsf{N}_{\widehat{H}}(T_1,\dots,T_\ell)\big|\,.$$

Observe that since G_{s_i} has no vertex of degree more than $\frac{cn}{\log n}$, the number of vertices in $\mathsf{N}_{H_{s_i-1}}(T_1,\dots,T_\ell)$ which are not in $\mathsf{N}_{H_{s_i}}(T_1,\dots,T_\ell)$ is at most $Q\frac{cn}{\log n}$, and since $e(G_s) \le Dn$ we have $p_{s_i-1} - p_{s_i} < n^{-0.5}$. Putting these together we have with probability at least $1-n^{-5C}$

$$|Z''| = (1 \pm C\alpha_{s_i})^{3i+2}\tilde{\delta}^{\max(0,i)}\left(\frac{p_{s_i}}{\hat{p}}\right)^{\ell} \cdot \big|X \cap \mathsf{N}_{\widehat{H}}(T_1,\dots,T_\ell)\big|\,.$$

Finally, we consider the embeddings of the graphs G_s with $s_i < s \le s_{i+1}-1$. By Lemma 4.5.8, with probability at least $1-n^{-5C}$, if for each s with $s_i \le s \le s_{i+1}-1$ the graph H_s is (α_s, Q)-coquasirandom, we have

$$\big|Z'' \cap \mathsf{N}_{H_{s_{i+1}-1}}(T_1,\dots,T_\ell)\big| = \left(1 \pm \tfrac{1}{2}\alpha_{s_{i+1}-1}\right)\left(\frac{p_{s_{i+1}-1}}{p_{s_i}}\right)^{\ell}|Z''|\,,$$

and this implies the desired equation for i. Summing the probability bounds we obtain the claimed probability of this equation holding. ▫

Taking the $i = |S|$ case of Claim 4.7.2, and a union bound over choices of S, ℓ, $T_1 \dots, T_\ell$ and $\tilde{\delta}$, by choice of $\alpha_{s_{\max}}$ we see that (*QUASI* 1), with error bound $\frac{1}{2}\gamma$, holds with probability at least $1 - Qn^{Qr+Qr^2+1} \cdot 3n^{-5C} \geq 1 - n^{-4C}$ provided *PackingProcess2* behaves. Taking account of the possibility of misbehaviour, we see that (*QUASI* 1) holds with probability at least $1 - n^{-2C}$. Note that since $n - \tilde{\delta}n$ is required to be in $[n]$, the number of choices of $\tilde{\delta}$ over which we take a union bound is at most n.

We next prove (*QUASI* 2). Given $q \leq Q$ and free q-weight functions ω_s on G_s for each $s \in [s^*]$, fix $U \subseteq V(\widehat{H})$, $\tilde{\delta}$, and an ordered q-set $\mathbf{e}$ in $V(\widehat{H}) \setminus U$. Let for each $1 \leq s \leq s^*$ the random variable $Y_s := \omega_s(\mathbf{e})$ by running *PackingProcess2* to obtain embeddings φ_s of each G_s. Let $\mathcal{E}$ be the event that *PackingProcess2* succeeds and that H_s is (α_s, Q)-quasirandom for each $s \in [s^*]$. Given F in the support of ω_s, since ω_s is free, F is a q-set which contains no edges of G_s. By Lemma 4.5.5, provided that H_{s-1} is (α_{s-1}, Q)-quasirandom, the probability that the ith vertex of F (in the natural order on $[n - \delta n]$) is embedded to the ith vertex of $\mathbf{e}$, and in addition $\operatorname{im} \varphi_s([n - \tilde{\delta}n]) \cap U = \emptyset$, is

$$\left(1 \pm 600CQ\alpha_{s-1}\delta^{-1}\right)^{(2+2D)Q+1} p_{s-1}^{0} n^{-q} \tilde{\delta}^{|U|} = \left(1 \pm \tfrac{1}{4}\gamma\right) n^{-q} \tilde{\delta}^{|U|} .$$

By linearity of expectation, if H_{s-1} is (α_{s-1}, Q)-quasirandom we have

$$\mathbb{E}[Y_s | H_{s-1}] = \operatorname{sum}(\omega_s) \cdot \left(1 \pm \tfrac{1}{4}\gamma\right) n^{-q} \tilde{\delta}^{|U|} .$$

Since $0 \leq Y_s \leq \frac{cn}{\log n}$, by Corollary 4.2.2 we have

$$\sum_{s=1}^{s^*} Y_s = \sum_{s=1}^{s^*} \operatorname{sum}(\omega_s) \cdot \left(1 \pm \tfrac{1}{2}\gamma\right) n^{-q} \tilde{\delta}^{|U|}$$

with probability at least $1 - n^{-5C}$. Here we use the fact $\sum_{s=1}^{s^*} \operatorname{sum}(\omega_s) \geq \gamma n^{q+1}$ and choice of c. Taking the union bound over choices of U, $\tilde{\delta}$ and $\mathbf{e}$, and taking account of the possibility of *PackingProcess2* not behaving, we see that (*QUASI* 2) holds with probability at least $1 - n^{-2C}$ as desired.

We now complete the proof of Theorem 4.7.1 by proving (*QUASI* 3). Our calculations in the preceding parts go through for $r = 2$, but for this part we need to assume $r \geq 3$. Fix s and $v \in V(\widehat{H})$, and sets $X_1, \ldots, X_{r-1}$ in $V(\widehat{H})$ each of size at least γn. Suppose that ω_s is a $(r-1)$-weight function on G_s whose support is contained in $[n-\tilde{\delta}n]$ such that $\mathrm{sum}(\omega_s) \geq \gamma n$. Recall that we need to prove that it is likely that either $v \in \varphi_s([n-\tilde{\delta}n])$, or we have control of $\sum_{\mathbf{e}\in \mathsf{N}_H(v;X_1,\ldots,X_{r-1})} \omega_s(\mathbf{e})$. We will assume for now that *PackingProcess2* behaves, and take account of its small failure probability later. Since there is nothing to prove if $v \in \varphi_s([n-\tilde{\delta}n])$, we will assume this event does not occur.

The proof of (*QUASI* 3) is fairly involved, so we now sketch the idea. First, we argue that when *RandomEmbedding* embeds the first $n-\tilde{\delta}n$ vertices of G_s into H_{s-1}, we can estimate $\sum_{\mathbf{e}\in \mathsf{N}_{H_{s-1}}(v;X_1,\ldots,X_{r-1})} \omega_s(\mathbf{e})$ fairly accurately. We then argue that the embedding of what remains of G_s does not change the result much; that is, we obtain an estimate for $\sum_{\mathbf{e}\in \mathsf{N}_{H_s}(v;X_1,\ldots,X_{r-1})} \omega_s(\mathbf{e})$. Finally, we argue that as the remaining hypergraphs $G_{s+1}, \ldots, G_{s^*}$ are embedded, we can maintain an estimate of $\sum_{\mathbf{e}\in \mathsf{N}_{H_{s'}}(v;X_1,\ldots,X_{r-1})} \omega_s(\mathbf{e})$. It is easy to estimate (using Lemma 4.5.5) the conditional expected change Y_i when any one G_i is embedded, and what we would like to do is to show that $\sum_{i=s+1}^{s^*} Y_i$ is close to the sum of the conditional expected changes. However the range of these Y_i is too large for Corollary 4.2.2 to give useful results. What we do to get around this is to define *capped* random variables; that is, we define a random variable Y_i' which is equal to the minimum of the actual change when G_i is embedded and a fixed number, the *cap*. The cap is small enough that we can apply Corollary 4.2.2 to show concentration of $\sum_{i=s+1}^{s^*} Y_i'$, and we then argue that it is very likely that $Y_i = Y_i'$ for each i and that the conditional expectations of the two random variables are very close. This final part is where we use our assumption $r \geq 3$ and where we need to do most work.

The following claim establishes the first step in the above argument. In this claim we will write φ_s although we formally do not assume *RandomEmbedding* succeeds in embedding all of G_s but only the first $n-\tilde{\delta}n$ vertices and should really therefore write instead the partial embedding of these first vertices.

Claim 4.7.3. *When* PackingProcess2 *is run up to and including the embedding of the first $n-\tilde{\delta}n$ vertices of G_s, with probability at least $1-3n^{-3C}$, we have either $v \in \varphi_s([n-\tilde{\delta}n])$ or*

$$\sum_{\mathbf{e}\in \mathsf{N}_{H_{s-1}}(v;X_1,\ldots,X_{r-1})} w_s(\mathbf{e}) = \big(1 \pm 5rC\alpha_s\big)p_{s-1} \cdot |X_1| \ldots |X_{r-1}| n^{1-r} \mathrm{sum}(\omega_s) \,.$$

Proof. We prove this claim by arguing that it is likely that the total weight of semi-edges in G_s whose first element is mapped to X_1 is roughly $\frac{|X_1|}{n}\mathrm{sum}(\omega_s)$, and inductively prove that the total weight of semi-edges in G_s whose first i elements are mapped to $X_1, \ldots, X_i$ respectively is likely to be roughly $\frac{|X_1|\ldots|X_i|}{n^i}\mathrm{sum}(\omega_s)$, for each $2 \le i \le r-2$. We finally use this to prove the claim statement. In each of these steps, the critical point is that we can keep control of the probability that any given vertex x of G_s is embedded to any given set X_i: we know how large is the set of vertices to which we will embed x from the diet condition, and we will prove a version of the diet condition for each set X_i. We begin by stating and proving this last condition.

We first claim that with probability at least $1 - 2n^{-3C}$, when *PackingProcess2* is run up to and including the embedding of G_{s-1}, it succeeds, H_{s-1} is (α_{s-1}, Q)-quasirandom, and for any set R of distinct semi-edges of H_{s-1} with $|R| \le Q$, and any $1 \le i \le q$ we have $\left|\mathsf{N}_{H_{s-1}}(R) \cap X_i\right| = (1 \pm \alpha_{s-1})p_{s-1}^{|R|}|X_i|$. The first two parts of this are part of the event that *PackingProcess2* behaves, whose failure probability is at most n^{-3C}, so we need to establish the final part. We use Lemma 4.5.8, once for each X_i and for each choice of R, to establish that (for each fixed such choice) we have with probability at least $1 - n^{-5C}$ that either *PackingProcess2* does not behave, or

$$\left|\mathsf{N}_{H_{s-1}}(R) \cap X_i\right| = \left(1 \pm \tfrac{1}{2}\alpha_{s-1}\right)\left(\tfrac{p_{s-1}}{\hat{p}}\right)^{|R|} \cdot \left|\mathsf{N}_{H_{s-1}}(R) \cap X_i\right| = (1 \pm \alpha_{s-1})p_{s-1}^{|R|}|X_i|\,,$$

where the second equality is since $(\widehat{H}, V(\widehat{H}) \setminus X_i)$ satisfies the (ξ, Q)-diet condition. Taking the union bound over choices of i and R we obtain the desired claim.

We now build on this to prove the desired version of the diet condition. That is, we prove that assuming that the likely event of the first claim occurs, when *RandomEmbedding* is run to embed the first $n - \tilde{\delta}n$ vertices of G_s into H_{s-1}, with probability at least $1 - n^{-4C}$ it succeeds, for each $1 \le t \le n - \tilde{\delta}n$ the pair $\big(H_{s-1}, \varphi_s([t])\big)$ satisfies the $(C\alpha_s, Q)$-diet condition, and for any set R of distinct semi-edges of H_{s-1} with $|R| \le Q$, and any $1 \le i \le q$ we have

$$\left|\mathsf{N}_{H_{s-1}}(R) \cap X_i \setminus \varphi_s([t])\right| = (1 \pm 2C\alpha_s)p_{s-1}^{|R|}\tfrac{n-t}{n}|X_i|\,.$$

This follows directly from the first claim and Lemma 4.4.3, with $Z = \mathsf{N}_{H_{s-1}}(R) \cap X_i$, taking the union bound over t and choices of R and i.

We now prove by induction on i the following. Suppose that the likely event of the first claim above holds. For each $0 \le i \le r-2$, with probability at least $1-n^{-4C}-in^{-5C}$, when *RandomEmbedding* is run to embed the first $n-\tilde{\delta}n$ vertices of G_s into H_{s-1}, we have

$$\sum_{(x_1,\dots,x_{r-1})} \omega_s(\{x_1,\dots,x_{r-1}\}) = (1 \pm 5iC\alpha_s)\mathrm{sum}(\omega_s)\prod_{j=1}^{i}\tfrac{|X_j|}{n},$$

where the sum runs over vectors with $x_1 < x_2 < \dots < x_{r-1}$ members of $[n-\tilde{\delta}n]$, such that $\varphi_s(x_j) \in X_j$ for each $1 \le j \le i$.

The induction statement is trivial for $i=0$. Suppose now $i \ge 1$, and the statement holds for $i-1$. For each $1 \le t \le n-\tilde{\delta}n$, let

$$Y_t = \sum_{(x_1,\dots,x_{r-1})} \omega_s(\{x_1,\dots,x_{r-1}\})\mathbb{1}_{\varphi_s(t)\in X_i},$$

where the sum runs over vectors with $x_1 < x_2 < \dots < x_{r-1}$ members of $[n-\tilde{\delta}n]$, such that $\varphi_s(x_j) \in X_j$ for each $1 \le j \le i$ and $x_i = t$. Observe that $0 \le |Y_t| \le \Delta(\omega_s) \le \frac{cn}{\log n}$ by definition. Letting $\mathscr{H}_{t-1}$ denote the embedding history of *RandomEmbedding* up to and including the embedding of $t-1$, we have

$$\mathbb{P}[\varphi_s(t) \in X_i | \mathscr{H}_{t-1}] = \frac{(1 \pm 2C\alpha_s)p_{s-1}^{\deg^-(t)}\frac{n+1-t}{n}|X_i|}{(1 \pm C\alpha_s)p_{s-1}^{\deg^-(t)}(n+1-t)} = (1 \pm 4C\alpha_s)\tfrac{|X_i|}{n},$$

where we use the $(C\alpha_s, Q)$-diet condition for $\big(H_{s-1}, \varphi_s([t-1])\big)$ to estimate the denominator and the second claim for the numerator. Assuming the likely event of the $i-1$ induction statement, we conclude

$$\sum_{t=1}^{n-\tilde{\delta}n} \mathbb{E}[Y_t|\mathscr{H}_{t-1}] = (1 \pm 4C\alpha_s)\tfrac{|X_i|}{n} \cdot (1 \pm 5(i-1)C\alpha_s)\mathrm{sum}(\omega_s)\prod_{j=1}^{i}\tfrac{|X_j|}{n},$$

and by Corollary 4.2.2, with $\mathcal{E}$ the event that the diet condition, and the variant diet conditions of the second claim, and the equation of the $i-1$ statement, all hold, we see that the probability that $\mathcal{E}$ holds and the equality of the i statement fails is at most n^{-5C}. This gives the desired induction statement for i.

Finally, we prove the main Claim. The argument is very similar to the above induction, except that we want to consider $(x_1,\dots,x_{r-1})$ such that not only $\varphi_s(x_{r-1}) \in X_{r-1}$, but also $\varphi_s(\{x_1,\dots,x_{r-1}\}) \in \mathsf{N}_{H_{s-1}}(v)$. Here we need to assume $v \notin \varphi_s([n-\tilde{\delta}n])$. We define similarly as above

$$Y'_t = \sum_{(x_1,\dots,x_{r-1})} \omega_s(\{x_1,\dots,x_{r-1}\})\mathbb{1}_{\varphi_s(t)\in X_i}\,,$$

where the sum runs over vectors with $x_1 < x_2 < \cdots < x_{r-1}$ members of $[n-\tilde{\delta}n]$, such that $\varphi_s(x_j) \in X_j$ for each $1 \le j \le r-1$, and $\varphi_s(\{x_1,\dots,x_{r-1}\}) \in \mathsf{N}_{H_{s-1}}(v)$, and $x_{r-1} = t$. For any t, and any $x_1 < x_2 < \cdots < x_{r-2} < t$, we have

$$\begin{aligned}&\mathbb{P}[\varphi_s(t) \in X_{r-1} \text{ and } \varphi_s(\{x_1,\dots,x_{r-2},t\}) \in \mathsf{N}_{H_{s-1}}(v)] = \\ &\quad = \frac{(1\pm 2C\alpha_s)p_{s-1}^{\deg^-(t)+1}\frac{n+1-t}{n}|X_{r-1}|}{(1\pm C\alpha_s)p_{s-1}^{\deg^-(t)}(n+1-t)} = (1\pm 4C\alpha_s)p_{s-1}\frac{|X_{r-1}|}{n}\,,\end{aligned}$$

where we obtain the extra factor of p_{s-1} in the numerator by looking at the common neighbourhood of $\varphi_s\big(\mathsf{N}^-_{G_s}(t)\big)$ and the extra semi-edge $\{\varphi_s(x_1),\dots,\varphi_s(x_{r-2}),v\}$. Note that this extra semi-edge is not in the set $\varphi_s\big(\mathsf{N}^-_{G_s}(t)\big)$ since it contains v and v is not in $\varphi_s([n-\tilde{\delta}n])$. An essentially identical application of Corollary 4.2.2 as above now gives us the claim. □

Next, we show that the embedding of the rest of G_s does not have much effect.

Claim 4.7.4. *When* PackingProcess2 *is run up to and including the embedding of* G_s*, with probability at least* $1-4n^{-3C}$*, we have either* $v \in \varphi_s([n-\tilde{\delta}n])$ *or*

$$\sum_{\mathbf{e}\in \mathsf{N}_{H_s}(v;X_1,\dots,X_{r-1})} w_s(\mathbf{e}) = \big(1\pm 10rC\alpha_s\big)p_s\cdot|X_1|\cdots|X_{r-1}|n^{1-r}\mathrm{sum}(\omega_s)\,.$$

Proof. Let $\mathcal{E}$ be the intersection of the likely event of Claim 4.7.3 and the event that *RandomEmbedding*, run to embed G_s into H_{s-1}, succeeds and for each $1 \le t \le v(G_s)$ the pair $\big(H_{s-1},\varphi_s([t])\big)$ satisfies the $(C\alpha_s,Q)$-diet condition. By Lemma 4.4.3 and Claim 4.7.3, $\mathcal{E}$ occurs with probability at least $1-4n^{-3C}$.

Suppose that when *RandomEmbedding* embeds G_s into H_{s-1}, at some time t we have $\varphi_s(t) = v$. If $t \le n-\tilde{\delta}n$, then the claim statement is true trivially, so we may assume $t > n-\tilde{\delta}n$. We now consider the embeddings of vertices from t onwards. Each vertex that we embed completes the embedding of at most D edges of G_s, and therefore is responsible for at most D semi-edges of $\mathsf{N}_{H_{s-1}}(v)$ not being in $\mathsf{N}_{H_s}(v)$. These D semi-edges have weight at most $\frac{Dcn}{\log n}$.

Since t has degree at most $\frac{cn}{\log n}$ in G_s, there are at most $\frac{cn}{\log n}$ vertices of G_s

which are in an edge with t, and only when we embed one of these vertices can we remove a semi-edge from $\mathsf{N}_{H_{s-1}}(v)$. Provided that we are in the event $\mathcal{E}$, at each time t' when we embed such a vertex, by the diet condition we do so to a set of size at least $\frac{1}{2}\gamma^D\delta n$. The expected weight of the semi-edges we remove from $\mathsf{N}_{H_{s-1}}(v)$ by one such vertex embedding, conditioning on the previous embedding history, is thus at most $2\gamma^{-D}\delta^{-1}\mathrm{sum}(\omega_s)n^{-1}$. We apply Corollary 4.2.2, with $\tilde{\mu} = \tilde{\nu} = 2\gamma^{-D}\delta^{-1}\mathrm{sum}(\omega_s)n^{-1} \cdot \frac{cn}{\log n}$ and $\tilde{\varrho} = \varepsilon\mathrm{sum}(\omega_s)$, to bound the probability that $\mathcal{E}$ occurs and the total weight removed, i.e.

$$\sum_{\mathbf{e}\in\mathsf{N}_{H_{s-1}}(v)\setminus\mathsf{N}_{H_s}(v)} \omega_s(\mathbf{e})$$

exceeds $2\varepsilon\mathrm{sum}(\omega_s)$. We obtain the upper bound

$$2\exp\Big(-\frac{\tilde{\varrho}^2}{2\frac{Dcn}{\log n}(2\tilde{\mu}+\tilde{\varrho})}\Big) < n^{-5C}\,,$$

where the inequality uses the choice of c. This bound, together with the choice of ε, and the facts $p_{s-1} - p_s \le Dn/\binom{n}{r}$, $|X_1|\cdots|X_{r-1}|n^{1-r} \ge \gamma^{r-1}$, and $p_s \ge \gamma$, establish the claim. □

We now need to deal with the embedding of $G_{s+1}, \ldots, G_{s^*}$. What we want to do is establish the following claim.

Claim 4.7.5. *With probability at least* $1 - n^{-K-2r}$*, for each* $s \le s' \le s^*$*, either* $v \in \varphi_s([n-\tilde{\delta}n])$ *or we have*

$$\sum_{\mathbf{e}\in\mathsf{N}_{H_{s'}}(v;X_1,\ldots,X_{r-1})} w_s(\mathbf{e}) = \big(1 \pm 30rC\alpha_{s'}\big)p_{s'} \cdot |X_1|\ldots|X_{r-1}|n^{1-r}\mathrm{sum}(\omega_s)\,.$$

Proof. Observe that to establish this claim, by the union bound it is enough to show that, for a given s' with $s < s' \le s^*$, with probability at most n^{-K-5r} we have that s' is minimal such that (4.7.1) fails, since Claim 4.7.4 already establishes this bound for $s' = s$. Furthermore, defining a weight function ω' on $\binom{V(\widehat{H})}{r-1}$ by

$$\omega'(y) = \sum_{\substack{\mathbf{e}\in\mathsf{N}_{H_s}(v;X_1,\ldots,X_{r-1})\\ \mathbf{e}\text{ is an ordering of }y}} \omega_s(\mathbf{e})$$

we see that the equation of Claim 4.7.5 is equivalent to

$$\sum_{y \in \mathsf{N}_{H_{s'}}(v)} \omega'(y) = \big(1 \pm 30rC\alpha_{s'}\big)p_{s'} \cdot |X_1| \cdots |X_{r-1}| n^{1-r} \mathrm{sum}(\omega_s)\,, \tag{4.7.1}$$

and we will aim to prove this version, in which we simply work with weights on semi-edges of $\widehat{H}$.

Let, for each $s < i \le s^*$,

$$Y_i = \sum_{y \in \mathsf{N}_{H_{i-1}}(v) \setminus \mathsf{N}_{H_i}(v)} \omega'(y)\,.$$

To begin with, we observe that, provided (4.7.1) holds for $i-1$ and H_{i-1} is (α_{i-1}, Q)-quasirandom, we can calculate $\mathbb{E}[Y_i | H_{i-1}, \omega']$.

Claim 4.7.6. *Suppose that for some $s < i \le s^*$, (4.7.1) holds for $i-1$ and H_{i-1} is (α_{i-1}, Q)-quasirandom. Then we have*

$$\mathbb{E}[Y_i | H_{i-1}, \omega'] = \big(1 \pm 10^4 Cr\alpha_{i-1}Q\delta^{-1}\big) r! |X_1| \cdots |X_{r-1}| n^{1-2r} \mathrm{sum}(\omega_s) e(G_i)\,.$$

Proof. By definition of Y_i we have

$$Y_i = \sum_{y \in \mathsf{N}_{H_{i-1}}(v)} \omega'(y) \cdot \mathbb{1}_{yv \text{ is used when embedding } G_i}$$

and therefore

$$\mathbb{E}[Y_i | H_{i-1}, \omega'] = \sum_{y \in \mathsf{N}_{H_{i-1}}(v)} \omega'(y) \cdot \mathbb{P}[yv \text{ is used when embedding } G_i \mid H_{i-1}, \varphi_s]\,.$$

Since H_{i-1} is (α_{i-1}, Q)-quasirandom, Lemma 4.5.5 yields

$$\mathbb{E}[Y_i | H_{i-1}, \omega'] = \sum_{y \in \mathsf{N}_{H_{i-1}}(v)} \omega'(y) \cdot (1 \pm 200Cr\alpha_{i-1}\delta^{-1})^{2+2D} \frac{r! e(G_i)}{p_{i-1} n^r}\,,$$

and since (4.7.1) holds for $i-1$ we get

$$\begin{aligned}\mathbb{E}[Y_i|H_{i-1},\omega'] &= \big(1\pm 30rC\alpha_{i-1}\big)p_{i-1}|X_1|\cdots|X_{r-1}|n^{1-r}\mathrm{sum}(\omega_s)\\ &\quad\cdot(1\pm 200Cr\alpha_{i-1}\delta^{-1})^{2+2D}\frac{r!e(G_i)}{p_{i-1}n^r}\\ &= \big(1\pm 10^4Cr\alpha_{i-1}Q\delta^{-1}\big)r!|X_1|\cdots|X_{r-1}|n^{1-2r}\mathrm{sum}(\omega_s)e(G_i)\,. \quad\square\end{aligned}$$

What we would like to do now is to use Corollary 4.2.2 to estimate $\sum_{i=s+1}^{s^*}Y_i$. However the range of the Y_i might be too large for this. To that end, we define, for each $s+1\le i\le s^*$, the random variables

$$Y_i' := \min\Big(Y_i, 2K'c\frac{\mathrm{sum}(\omega_s)}{\log n}\Big) \quad \text{with} \quad K' = 10^{10}K\gamma^{-1}r^2CD^3\delta^{-1}\,.$$

The 'capped' random variable Y_i' trivially does not have an excessively large range, and so we can use Corollary 4.2.2 to show that $\sum_{i=s+1}^{s^*}Y_i'$ is very likely to be close to the corresponding sum of conditional expectations. For $r\ge 3$, we will now prove that for each i, with very high probability we have $Y_i'=Y_i$, so that in particular (by the union bound) the statement holds simultaneously for all $s<i\le s^*$ and $\mathbb{E}[Y_i'|H_{i-1},\omega']$ is very close to $\mathbb{E}[Y_i|H_{i-1},\omega']$. This is what we need to establish (*QUASI* 3). We should note that our proof does *not* go through for $r=2$, and indeed for $r=2$ one can find examples where it is likely that $Y_i'\ne Y_i$ for some i. For a specific class of weight functions ω_s, and with $X_1=V(\widehat{H})$, in [2] a different proof of (*QUASI* 3) is given; the approach there should generalise to prove the entire $r=2$ case of (*QUASI* 3), but we have not checked the details.

Let us define the event $\mathsf{CapE}(i)$ as $Y_i > 2K'c\frac{\mathrm{sum}(\omega_s)}{\log n}$. This means that the set $\mathsf{N}_{H_{i-1}}(u)\setminus\mathsf{N}_{H_i}(u)$ of semi-edges has a large total weight. We want to separate cases when this comes from few edges and when this comes from many edges. Therefore we define the set

$$W=\Big\{y\in\mathsf{N}_{H_{i-1}}(v)\,\Big|\,\omega'(y)>\frac{\mathrm{sum}(\omega_s)}{(\log n)^{10}}\Big\}$$

as the *heavy* semi-edges and

$$L=\Big\{y\in\mathsf{N}_{H_{i-1}}(v)\,\Big|\,\frac{\mathrm{sum}(\omega_s)}{(\log n)^{10}}\ge\omega'(y)>0\Big\}$$

as the *light* semi-edges. We define $\mathsf{CapE1}(i)$ as $\sum_{y\in W\setminus\mathsf{N}_{H_i}(v)}\omega'(y) > K'c\frac{\mathrm{sum}(\omega_s)}{\log n}$

meaning that the heavy edges hit the cap and $\mathsf{CapE2}(i)$ as $\sum_{y\in L\setminus \mathsf{N}_{H_i}(v)} \omega'(y) > K'c\frac{\mathrm{sum}(\omega_s)}{\log n}$ meaning that the light edges hit the cap. Clearly $\mathsf{CapE}(i)$ implies at least one of $\mathsf{CapE1}(i)$ and $\mathsf{CapE2}(i)$ holds. By definition of $\mathrm{sum}(\omega_s)$, we have $|W| \le (\log n)^{10}$.

In order to prove $\mathsf{CapE}(i)$ is unlikely, it is enough to prove that each of $\mathsf{CapE1}(i)$ and $\mathsf{CapE2}(i)$ is unlikely. We begin with $\mathsf{CapE1}(i)$. The idea here is that in order for $\mathsf{CapE}(i)$ to occur, a reasonably large matching of semi-edges of W must form edges with v which are used in embedding G_i; we prove that any one such large matching is unlikely and take the union bound over all such matchings. The next claim justifies that a large matching must be chosen.

Claim 4.7.7. *Let S be a set of semi-edges of $\widehat{H}$ such that $\sum_{y\in S} \omega'(y) \ge a(r-1)\Delta(\omega_s)$. There is a subset $S' \subseteq S$ such that $|S'| = a$ and S' is a matching.*

Proof. We greedily pick semi-edges y from S, add them to S' and remove any semi-edge from S that shares at least one vertex with y. Since a semi-edge has $r-1$ vertices, in each step we remove semi-edges with total weight at most $(r-1)\Delta(\omega_s)$, and therefore the number of steps before removing all semi-edges from S is at least a. If we stop after a steps, we have $|S'| = a$, and by construction S' is a matching. □

We now show that $\mathsf{CapE1}(i)$ is unlikely.

Claim 4.7.8. *Suppose that H_{i-1} is (α_{i-1}, Q)-quasirandom and $r > 2$. Then we have*

$$\mathbb{P}\big[\mathsf{CapE1}(i)\,\big|\,H_{i-1}, \omega'\big] \le n^{-K-10r}\,.$$

Proof. Let $\psi_1, \ldots, \psi_{n-\delta n}$ the embedding sequence generated by *RandomEmbedding* when called to embed G_i into H_{i-1}. Let $\mathcal{E}$ be the event that for each $1 \le j \le n-\delta n$ the $(C\alpha, Q)$-diet condition holds for $\big(H_{i-1}, \mathrm{im}(\psi_j)\big)$. By Lemma 4.4.3 we have

$$\mathbb{P}\big[\mathcal{E}\,\big|\,H_{i-1}, \omega'\big] \ge 1 - n^{-5C}\,.$$

For the following calculations we assume $\mathcal{E}$ holds.

By Claim 4.7.7, since $K'c\frac{\mathrm{sum}(\omega_s)}{\log n} > (K+20r)(r-1)\Delta(\omega_s)$, if $\mathsf{CapE1}(i)$ holds then there is a matching of $a := K{+}20r$ semi-edges $y_1 = (y_{1,1}, \ldots, y_{1,r-1}), \ldots, y_a = (y_{a,1}, \ldots, y_{a,r-1}) \in W$ such that $y_j v$ is an edge of H_{i-1} which is used in the embedding of G_i for each $1 \le j \le a$. That is, there exists $x \in V(G_i)$ and a matching of semi-edges $z_1 = (z_{1,1}, \ldots, z_{1,r-1}), \ldots, z_a = (z_{a,1}, \ldots, z_{a,r-1})$ in

$\mathsf{N}_{G_i}(x)$, such that φ_i maps x to v, and $z_{j,\ell}$ to $y_{j,\ell}$ for each $j \in [a]$ and $\ell \in [r-1]$. We fix such a choice, and will later take the union bound over all possible such choices. Let E denote the event that $\varphi_i(z_j, \ell) = y_{j,\ell}$ for each $j \in [a]$ and $\ell \in [r-1]$.

We order the pairs $(z_{j,k}, y_{j,k})$ by increasing $z_{j,k}$ and for convenience rename them to pairs $((x_1, y'_1), \dots, (x_{a(r-1)}, y'_{a(r-1)}))$. Let, for each $\ell \in [a(r-1)]$, the event E_ℓ be that $\varphi_i(x_\ell) = y'_\ell$. Let $\tilde{E}_\ell$ denote $\cap_{h=1}^{\ell} E_h$. Since $\mathcal{E}$ holds, we have

$$\mathbb{P}\big[E_\ell \,\big|\, H_{i-1}, \omega', \tilde{E}_{\ell-1}\big] \le 2\gamma^{-D}\delta^{-1}n^{-1}\,.$$

This implies

$$\begin{aligned}\mathbb{P}[E|H_{i-1}, \omega'] &= 1 - \mathbb{P}[\mathcal{E}|H_{i-1,\omega'}] + \prod_{\ell=1}^{a(r-1)} \mathbb{P}\big[E_\ell \,\big|\, H_{i-1}, \omega', \tilde{E}_{\ell-1}\big] \\ &\le n^{-5C} + \big(2\gamma^{-D}\delta^{-1}n^{-1}\big)^{a(r-1)} \le \big(\tfrac{1}{4}\gamma^D\delta n\big)^{-a(r-1)}\,.\end{aligned}$$

Taking the union bound over choices of x, $y_{j,\ell}$ and $z_{j,\ell}$, we see

$$\begin{aligned}\mathbb{P}\big[\mathsf{CapE1}(i) \,\big|\, H_{i-1}, \omega'\big] &\le n\binom{|W|}{a}\binom{\mathsf{N}_{G_i}(x)}{a}((r-1)!)^a \cdot \big(\tfrac{1}{4}\gamma^D\delta n\big)^{-a(r-1)} \\ &\le n(\log n)^{10a} \cdot \big(\tfrac{cn}{\log n}\big)^a \big(4^{r-1}r!\gamma^{-D(r-1)}\delta^{1-r}\big)^a \cdot n^{-a(r-1)} \\ &\le n^{2-a(r-2)} \le n^{-K-10r}\,,\end{aligned}$$

as desired. □

To deal with $\mathsf{CapE2}(i)$, we also consider the vertex-by-vertex embedding of G_i. The critical point is that each vertex embedding embeds at most D edges of G_i, and therefore is responsible for removing at most D light semi-edges from $\mathsf{N}_{H_i}(v)$. Since these semi-edges are light, the maximum weight removed is at most $D \cdot \mathrm{sum}(\omega_s)(\log n)^{-10}$, and the extra log factors here (as compared to heavy semi-edges) give us better concentration.

Claim 4.7.9. *Suppose that H_{i-1} is (α_{i-1}, Q)-quasirandom and $r > 2$. Then we have*

$$\mathbb{P}\big[\mathsf{CapE2}(i) \,\big|\, H_{i-1}, \omega'\big] \le n^{-K-10r}\,.$$

Proof. Again, let $\psi_1, \dots, \psi_{n-\delta n}$ the embedding sequence generated by *RandomEmbedding* when called to embed G_i into H_{i-1}. Let $\mathcal{E}$ be the event that for each $1 \leq j \leq n - \delta n$ the $(C\alpha, Q)$-diet condition holds for $\big(H_{i-1}, \mathrm{im}(\psi_j)\big)$. By Lemma 4.4.3 we have

$$\mathbb{P}\big[\mathcal{E}\,\big|\,H_{i-1}, \omega'\big] \geq 1 - n^{-5C}\,,$$

and for the following calculations we assume $\mathcal{E}$ holds.

Let us fix $x \in V(G_i)$ and assume $\varphi_i(x) = v$; we will later take a union bound over the choices of x.

Let $x_1 < \cdots < x_\ell$ be the vertices of G_i which are maximum vertices of some semi-edge in $\mathsf{N}_{G_i}(x)$ and which are greater than x. Note that there can be some at most D semi-edges of $\mathsf{N}_{G_i}(x)$ all of whose vertices are smaller than x. Trivially $\ell \leq \Delta$. For each x_j let S_j be the set of semi-edges in $\mathsf{N}_{G_i}(x)$ whose last element is x_j, and note $|S_j| \leq D$ since G_i is D-degenerate. We define

$$X_j := \sum_{y \in \varphi_i(S_j) \cap L} \omega'(y)\,.$$

We have $0 \leq X_j \leq D \cdot \mathrm{sum}(\omega_s)(\log n)^{-10}$ for each j. Taking account of the at most D light semi-edges which come entirely before x, we see that in order for $\mathsf{CapE2}(i)$ to occur, we need

$$D \cdot \mathrm{sum}(\omega_s)(\log n)^{-10} + \sum_{j=1}^{\ell} X_j > K'c\tfrac{\mathrm{sum}(\omega_s)}{\log n}\,,$$

and so $\sum_{j=1}^{\ell} X_j > \frac{1}{2}K'c \cdot \mathrm{sum}(\omega_s)(\log n)^{-1}$. Our aim is now to show that this is unlikely.

When we embed x_j, by the diet condition we embed it uniformly at random to a set of size at least $\frac{1}{2}\gamma^D \delta n$ vertices; at most, all of $\mathrm{sum}(\omega_s)$ could be distributed over these vertices (making semi-edges with the remaining embedded parts of S_j) and so we have

$$\mathbb{E}[X_j | H_{i-1}, \omega', \psi_{x_j - 1}] \leq \mathrm{sum}(\omega_s) \cdot 2\gamma^{-D}\delta^{-1}n^{-1}\,.$$

Summing over j, within $\mathcal{E}$ we get

$$\sum_{j=1}^{\ell} \mathbb{E}[X_j | H_{i-1}, \omega', \psi_{x_j - 1}] \leq \frac{cn}{\log n} \cdot \mathrm{sum}(\omega_s) \cdot 2\gamma^{-D}\delta^{-1}n^{-1}\mathrm{sum}(\omega_s) \leq \tfrac{1}{4}K'c\tfrac{\mathrm{sum}(\omega_s)}{\log n}$$

by choice of K'.

Applying Corollary 4.2.2(*a*) with the given $\mathcal{E}$, with $\tilde{\mu} = \frac{1}{4}K'c\frac{\mathrm{sum}(\omega_s)}{\log n}$, and with $R = D \cdot \mathrm{sum}(\omega_s)(\log n)^{-10}$ we then obtain that conditioning on H_{i-1} and ω_s we have

$$\begin{aligned}\mathbb{P}\left[\mathcal{E} \text{ and } \sum_{j=1}^{\ell} X_j \geq \frac{1}{4}K'c\frac{\mathrm{sum}(\omega_s)}{\log n}\right] &\leq 2\exp\left(-\frac{\frac{1}{4}K'c\frac{\mathrm{sum}(\omega_s)}{\log n}}{4D \cdot \mathrm{sum}(\omega_s)(\log n)^{-10}}\right) \\ &\leq n^{-K-20r}\,,\end{aligned}$$

and by the union bound over choices of x we have the claim. □

By Claims 4.7.8 and 4.7.9, we see that for each $s < i \leq s^*$, provided that *PackingProcess2* behaves, we have $\mathbb{P}[\mathsf{CapE}(i)|H_{i-1}, \omega'] \leq 2n^{-K-10r}$. Thus, by the union bound, with probability at least $1 - n^{-3C} - 2n^{-K-9r}$ we have $Y_i' = Y_i$ for each $s < i \leq s^*$. In addition, since $Y_i \leq \mathrm{sum}(\omega_s)$ trivially, provided *PackingProcess2* behaves we have

$$\begin{aligned}\mathbb{E}[Y_i'|H_{i-1}, \omega'] &> \mathbb{E}[Y_i|H_{i-1}, \omega'] + (n^{-3C} - 2n^{-K-9r}) \cdot \mathrm{sum}(\omega_s) \\ &> (1 + n^{-4})\mathbb{E}[Y_i|H_{i-1}, \omega']\,.\end{aligned}$$

By Corollary 4.2.2, and since $\mathrm{sum}(\omega_s) \geq \gamma n$, with probability at least $1 - n^{-K-6r}$, conditioning on ω' and H_s, if *PackingProcess2* behaves and the good event of Claim 4.7.4 occurs, we have

$$\sum_{i=s+1}^{s'} Y_i' = (1 \pm \varepsilon) \sum_{i=s+1}^{s'} \mathbb{E}[Y_i'|H_{i-1}, \omega']\,,$$

and we claim that this equation is inconsistent with s' being the first failure time of (4.7.1). Summing up the failure probabilities, this then establishes the Claim.

Supposing that (4.7.1) holds up to and including $s' - 1$ and the likely events above occur, we have

$$
\begin{aligned}
&\sum_{y \in \mathsf{N}_{H_{s'}}(v)} \omega'(y) \\
&= \sum_{y \in \mathsf{N}_{H_s}(v)} \omega'(y) - \sum_{i=s+1}^{s'} Y_i \\
&= \sum_{y \in \mathsf{N}_{H_s}(v)} \omega'(y) - (1 \pm \varepsilon) \sum_{i=s+1}^{s'} \mathbb{E}[Y_i' | H_{i-1}, \omega'] \\
&= \sum_{y \in \mathsf{N}_{H_s}(v)} \omega'(y) - (1 \pm \varepsilon)(1 \pm n^{-4}) \sum_{i=s+1}^{s'} \mathbb{E}[Y_i | H_{i-1}, \omega'] \\
&= \sum_{y \in \mathsf{N}_{H_s}(v)} \omega'(y) \\
&\quad - (1 \pm 2\varepsilon) \sum_{i=s+1}^{s'} \left(1 \pm 10^4 C r \alpha_{i-1} Q \delta^{-1}\right) r! |X_1| \cdots |X_{r-1}| n^{1-2r} \mathrm{sum}(\omega_s) e(G_i) \,,
\end{aligned}
$$

and the main term of this last line is, by the good event of Claim 4.7.4,

$$
\begin{aligned}
&p_s \cdot |X_1| \cdots |X_{r-1}| n^{1-r} \mathrm{sum}(\omega_s) - r! |X_1| \cdots |X_{r-1}| n^{1-2r} \mathrm{sum}(\omega_s) \sum_{i=s+1}^{s'} e(G_i) \\
&= (1 \pm \varepsilon) p_{s'} |X_1| \cdots |X_{r-1}| n^{1-r} \cdot \mathrm{sum}(\omega_s) \,,
\end{aligned}
$$

which is as desired. What remains is to estimate the error term

$$
\begin{aligned}
&10 r C \alpha_s p_s |X_1| \cdots |X_{r-1}| n^{1-r} \mathrm{sum}(\omega_s) \\
&\qquad + \sum_{i=s+1}^{s'} 10^5 C r \alpha_{i-1} Q \delta^{-1} r! |X_1| \cdots |X_{r-1}| n^{1-2r} \mathrm{sum}(\omega_s) e(G_i) \\
&\quad \le |X_1| \cdots |X_r| n^{1-r} \mathrm{sum}(\omega_s) \cdot \left(10 r C \alpha_s p_s + \sum_{i=s+1}^{s'} 10^5 C D r \alpha_{i-1} Q \delta^{-1} r! n^{1-r}\right) \\
&\quad \le |X_1| \cdots |X_r| n^{1-r} \mathrm{sum}(\omega_s) \cdot \left(10 r C \alpha_s p_s + \int_{x=-\infty}^{s'} 10^5 C D r \alpha_x Q \delta^{-1} r! n^{1-r} \mathrm{d}x\right) \\
&\quad \le |X_1| \cdots |X_r| n^{1-r} \mathrm{sum}(\omega_s) \left(10 r C \alpha_s p_s + \gamma \alpha_{s'}\right) \\
&\quad \le 20 r C \alpha_s p_s |X_1| \cdots |X_r| n^{1-r} \mathrm{sum}(\omega_s) \,,
\end{aligned}
$$

where the penultimate line is by choice of the α_x.

Putting these together, the above likely events imply

$$\sum_{y\in \mathsf{N}_{H_{s'}}(v)} \omega'(y) =(1\pm\varepsilon)p_{s'}|X_1|\cdots|X_{r-1}|n^{1-r} \cdot \mathrm{sum}(\omega_s) \pm 20rC\alpha_s p_s|X_1|\cdots|X_r|n^{1-r}\mathrm{sum}(\omega_s)\,,$$

which is precisely the statement that (4.7.1) holds for s'. This completes the proof of Claim 4.7.5. ◻

Finally, Claim 4.7.5, for $s' = s^*$, together with the union bound over v and $s\in[s^*]$, establishes *(QUASI 3)*. ☐

4.8 Perfect packings

In this section we prove Theorem 4.1.4. We deduce Theorem 4.1.4 from Theorem 4.1.2 and the following technical result.

Theorem 4.8.1. *For every $D\ge 1$ and $r\ge 3$, and every sufficiently small $\mu>0$ there are n_0 and $\xi, c>0$ such that for every $n\ge n_0$ the following holds. Suppose that $\widehat{H}$ is a r-uniform $(\xi, D(r+1)+3)$-quasirandom hypergraph with n vertices. Suppose that $s^*=\lfloor \mu n^{r-1}\rfloor$, and that the hypergraphs $(G_s)_{s\in[s^*]}$ all have maximum degree $\frac{cn}{\log n}$, have exactly $n-\lfloor\mu n\rfloor$ vertices, and have at least μn vertices of degree 1. Suppose further that $\sum_{s\in[s^*]} e(G_s)=e(\widehat{H})$. Then $(G_s)_{s\in[s^*]}$ packs into $\widehat{H}$.*

Before sketching the proof of Theorem 4.8.1, we show that it, together with Theorem 4.1.2, implies Theorem 4.1.4.

Proof of Theorem 4.1.4. Without loss of generality, we may assume μ is sufficiently small for Theorem 4.8.1 for input D and r. Given D, r, μ, let γ' be small enough to play the rôle of ξ in Theorem 4.8.1 for input D, μ. Suppose that n_0 is also large enough, and $\xi, c>0$ are small enough, for this application of Theorem 4.8.1 and also for Theorem 4.1.2 for input r, D and $\gamma=\min(\gamma', \frac{1}{2}\mu^2)$.

Let n and the hypergraphs $\widehat{H}$ and $(G_i)_{i\in[m]}$ be given. Without loss of generality, we may assume $\sum_{i\in[m]} e(G_i)=e(\widehat{H})$, since otherwise we can add new hypergraphs consisting of single edges to our family until this condition is satisfied. We can further assume that the $\lfloor\mu n^{r-1}\rfloor$ special hypergraphs come last in the sequence.

We apply Theorem 4.1.2, with input as above, to pack the hypergraphs G_i with $1\le i\le s^*-\lfloor\mu n^{r-1}\rfloor$ into $\widehat{H}$, and let $\widetilde{H}$ be the hypergraph of leftover edges. By Theorem 4.1.2 such a packing exists and $\widetilde{H}$ is $(\gamma, 4D(r+1)r+3)$-quasirandom.

We now create a sequence $G'_1, \dots, G'_{s'}$ of hypergraphs with $s' := \lfloor \mu n^{r-1} \rfloor$ by taking the special hypergraphs $G_{s^*-\lfloor \mu n^{r-1}\rfloor+1}, \dots, G_{s^*}$ and adding to each isolated vertices until they have $n - \lfloor \mu n \rfloor$ vertices. By Theorem 4.8.1 we can pack $(G'_i)_{i\in[s']}$ into $\widetilde{H}$. Ignoring the extra isolated vertices, this gives us a packing of $G_{s^*-\lfloor \mu n^{r-1}\rfloor+1}, \dots, G_{s^*}$ into $\widetilde{H}$ and therefore a packing of $G_1, \dots, G_{s^*}$ into $\widehat{H}$ as desired. □

We now sketch the proof of Theorem 4.8.1. For this, we need to define various constants, which we will take as fixed throughout this section. We do not provide explicit dependencies for all constants (mainly because Lemma 4.2.9, which we used to prove Lemma 4.2.6, does not provide explicit dependencies), but explain now how to choose them.

Definition 4.8.2 (Choice of constants)**.** The constants D, r and μ are provided as input to Theorem 4.8.1, with the assumption that μ is sufficiently small for Lemma 4.2.6. We let ϱ be returned by Lemma 4.2.6 for input μ. We choose $\nu \le 10^{-3} r^{-2} \varrho\mu$ sufficiently small to play the rôle of ε in Lemma 4.2.6.

We choose $\gamma \le \frac{1}{100}(r!)^{-2r}\mu^{3r}\nu^{2r}2^{-3r}r^{-3r}$ sufficiently small for Lemma 4.8.5 below, and given γ we let ξ and c be sufficiently small for Lemma 4.8.5 with input D, r, μ, ν and γ, and in addition we let c be sufficiently small for various explicit calculations in this section. Finally, we suppose n_0 is large enough for Lemma 4.8.5 below for the given inputs, and also large enough such that $\frac{1}{8r!r}\mu\nu n_0$ is large enough to play the rôle of m in Lemma 4.2.6 with inputs as above, and sufficiently large for various explicit calculations in this section.

Given $n \ge n_0$, the density p is determined by

$$p = \lfloor \mu n^{r-1} \rfloor \lfloor \nu n \rfloor \binom{n}{r}^{-1} = \left(1 \pm \tfrac{1}{2}\right)\frac{\mu\nu}{r!} \,. \tag{4.8.1}$$

In what follows, we will assume $s^* := \lfloor \mu n^{r-1} \rfloor$.

We start by creating an almost perfect packing, which omits ℓ leaves (i.e. vertices of degree 1) in each hypergraph.

Definition 4.8.3 (corresponding subgraph sequence)**.** Given a sequence $(G_s)_{s\in[s^*]}$ of hypergraphs, we say that $(G'_s)_{s\in[s^*]}$ is a *corresponding subgraph sequence omitting ℓ leaves* if for each $s \in [s^*]$ we have $G'_s = G_s - V_s + I_s$ for a strongly independent set V_s of leaves in G_s with $|V_s| = \ell$, and a set I_s of new and isolated vertices with $|I_s| = \ell$.

The addition of the isolated set I_s is purely for technical reasons: it guarantees that the G'_s have exactly $n - \mu n$ vertices, which allows for easier formulae than the $n - \mu n - \ell$ we would otherwise have. The restriction that the set V_s is strongly independent means that there is a bijection between the set V_s and the set of edges of G_s which are not in G'_s; we do not pick two leaves which lie in the same edge of G_s.

We first find a packing $(\varphi'_s)_{s\in[s^*]}$ of the corresponding subgraph sequence $(G'_s)_{s\in[s^*]}$ with the additional properties guaranteed by Lemma 4.8.5 below. We will apply Lemma 4.8.5 (which is a corollary of Theorem 4.7.1) with the weight function on G'_s defined as follows.

Definition 4.8.4 (weights)**.** Let $(G_s)_{s\in[s^*]}$ be a sequence of n-vertex hypergraphs each with a strongly independent set V_s of ℓ leaves, and $(G'_s)_{s\in[s^*]}$ be a corresponding subgraph sequence, H be an n-vertex hypergraph, and $\varphi'_s\colon V(G'_s) \to V(H)$ be an injection for each $s \in [s^*]$. For $s^* - \lfloor \mu n^{r-1} \rfloor < s \leq s^*$ we define for each $x \in S(G_s)$ the weight

$$w_s(x) = \big|\{y \in \mathrm{N}_{G_s}(x)\colon y \text{ is a leaf of } G_s \text{ in } V_s\}\big|\,,$$

and for each $y \in S(H)$ the weight

$$w_s(y) = w_s\big({\varphi'}_s^{-1}(v)\big)\,.$$

Further, for each $y \in S(H)$ we define

$$w(v) = \sum_{s\in[s^*]} w_s(y)\,.$$

Since each set V_s of omitted leaves is a strongly independent set in G_s, the weight of any semi-edge containing an omitted leaf is 0. Thus the entire weight of G_s, which is ℓ, the number of omitted leaves, is supported on the semi-edges on the vertices of G'_s. The packing of the G'_s is then guaranteed by the following lemma.

Lemma 4.8.5 (almost perfect packing lemma)**.** *Given $D \geq 1$, $r \geq 2$ and $\mu > 0$, provided we have*

$$0 < c \ll \xi \ll \gamma \ll \nu, \mu, \frac{1}{D}, \frac{1}{r}$$

the following holds. Let $\widehat{H}$ be a $(\xi, (r+1)D+3)$-quasirandom hypergraph with n vertices. Let $(G_s)_{s\in[s^*]}$ be a D-degenerate hypergraph sequence with maximum *degree $\frac{cn}{\log n}$ in which each graph contains at least μn leaves and $\sum_{s\in[s^*]} e(G_s) = e(\widehat{H})$, and let $(G'_s)_{s\in[s^*]}$ be a corresponding subgraph sequence omitting $\lfloor \nu n\rfloor$ leaves. Then there exists a packing $(\varphi'_s)_{s\in[s^*]}$ of $(G'_s)_{s\in[s^*]}$ into $\widehat{H}$ with leftover H which is $(\gamma^3, 2)$-quasirandom such that for $p = \lfloor \mu n^{r-1}\rfloor\lfloor \nu n\rfloor\binom{n}{r}^{-1}$ and for all $y \in S(H)$ and $s, s' \in [s^*]$ we have*

- *(P 1)* $w(y) = (1 \pm \gamma^3)\frac{pn}{r}$,
- *(P 2)* $\left|\mathsf{N}_H(y) \setminus \operatorname{im}\varphi'_s\right| = (1 \pm \gamma^3)\mu pn$,
- *(P 3)* $\left|\mathsf{N}_H(y) \setminus (\operatorname{im}\varphi'_s \cup \operatorname{im}\varphi'_{s'})\right| = (1 \pm \gamma^3)\mu^2 pn$ *if* $s \neq s'$,

for all $u \in V(H)$ and $y \in S(H)$ with $u \notin y$ we have,

- *(P 4)* $\sum_s w_s(y)\mathbb{1}_{u\notin \operatorname{im}\varphi'_s} = (1 \pm \gamma^3)\mu\frac{pn}{r}$,

and for all $u \in V(H)$ we have

- *(P 5) If $u \notin \operatorname{im}\varphi'_s$ then $\sum_{y\colon yu\in E(H)} w_s(y) < \frac{10r!p^2n}{\mu}$.*

This lemma is deduced from Theorem 4.7.1 in Section 4.8.3. Given this packing of the G'_s, what remains is to extend each φ'_s by packing the remaining leaves of G_s. At each semi-edge $y \in S(H)$, we have a collection of $w(y)$ leaves to pack coming from the various hypergraphs G_s; we say these are the *leaves dangling at y*. We next choose an orientation $\vec{H}$ of H such that $N^+_{\vec{H}}(y) = w(y)$ for each semi-edge y in H. The idea here is that we will then embed the leaves dangling at y to the out-neighbours of y. The following lemma, which we prove in Section 4.8.2, tells us that a random orientation can be slightly modified to obtain the desired orientation: this will allow us to show that $\vec{H}$ inherits the quasirandomness properties of H guaranteed by Lemma 4.8.5.

Lemma 4.8.6 (orientation lemma)**.** *Given $p > 0$, $r \geq 2$ and $0 < \gamma < 2^{-2r-3}p^{2r}r^{-3r}$, let H be an n-vertex r-uniform $(\gamma^3, 2)$-quasirandom hypergraph of density p with semi-edge weights $w\colon V(H) \to \mathbb{N}_0$ such that $w(y) = (1 \pm \gamma^3)\frac{pn}{r}$ for all $y \in S(H)$ and such that we have $\sum_{y\in V} w(y) = e(H)$. If $\vec{H}_0$ is a uniform random orientation of H, then with probability tending to 1 as $n \to \infty$ there is an orientation $\vec{H}$ of H such that for all $y \in S(H)$ we have*

- *(O 1)* $\deg^+_{\vec{H}}(y) = w(y)$, *and*
- *(O 2)* $\left|\{u \in V(\vec{H})\colon yu \in E(H) \text{ and } yu \text{ is oriented differently in } \vec{H} \text{ and } \vec{H}_0\}\right| \leq \gamma^2 n$.

To help with packing the leaves dangling at each semi-edge y, we define the following auxiliary bipartite graphs.

Definition 4.8.7 (leaf matching graphs)**.** Given $y \in S(\vec{H})$, we define the *leaves at* y to be the set

$$L_y := \Big\{x \colon \exists s \text{ such that } x \in V(G_s) \setminus V(G'_s) \text{ and } x{\varphi'}_s^{-1}(y) \in E(G_s)\Big\}$$

Let the *leaf matching graph* F_y be the bipartite graph with parts L_y and $N^+_{\vec{H}}(y)$, and edges xu with $x \in L_y$ and $u \in N^+_{\vec{H}}(y)$ whenever $u \notin \operatorname{im} \varphi'_s$ for the s such that $x \in V(G'_s)$.

Observe that a perfect matching in F_y defines an assignment of the leaves at (all preimages of) y to $N^+_{\vec{H}}(y)$ which extends the packing of $(G'_s)_{s\in[s^*]}$. In what follows, we will be able to choose $\vec{H}$ such that for each y, the graph F_y has equal parts of size roughly $\frac{1}{r}pn$ and every vertex has degree roughly $\frac{1}{r}\mu pn$. We will further see that each F_y satisfies a codegree condition which by Lemma 4.2.6 implies that F_y has a perfect matching.

If we simply chose a perfect matching in each F_y to embed all the leaves $\bigcup_y L_y$, then we would almost have a perfect packing—each edge of $\widehat{H}$ would be used exactly once—but it could be the case that multiple leaves of some G_s (not in the same L_y) are embedded to a single $u \in V(\widehat{H})$. To avoid this, we find perfect matchings in each F_y one at a time and update the leaf matching graphs by removing edges which are no longer useable. We choose these perfect matchings uniformly at random at each step, and we will argue (using Lemma 4.2.6) that our updated leaf matching graphs lose only very few (at most $\varrho pn/r$) edges at any given vertex, and thus continue to satisfy the conditions of Lemma 4.2.6.

Making this precise, assume $V(\vec{H}) = \{1, \dots, n\}$, and set $F_y^{(0)} := F_y$ for each $y \in S(\vec{H})$. We use the following algorithm *MatchLeaves*.

Assuming *MatchLeaves* succeeds, we then, for each $s \in [s^*]$ and each $x \in V(G_s)$, set

$$\varphi_s(x) = \begin{cases} \varphi'_s(x) & \text{if } x \in \operatorname{dom}(\varphi'_s) \\ \sigma_y(x) & \text{if } x \in L_y\,. \end{cases} \tag{4.8.2}$$

This is a perfect packing of $(G_s)_{s\in[s^*]}$ into $\widehat{H}$ by construction. Thus to prove Theorem 4.8.1, what we need to do is prove Lemmas 4.8.5 and 4.8.6 above, and argue that *MatchLeaves* succeeds with positive probability. We do the last of these steps first, in the following subsection.

Algorithm 4: *MatchLeaves*

Input : a hypergraph sequence $(G_s)_{s\in[s^*]}$, a corresponding subgraph sequence $(G'_s)_{s\in[s^*]}$ omitting ℓ leaves, and associated leaf matching graphs $(F_y^{(0)})_{y\in S([n])}$

Output : matchings $(\sigma_y)_{y\in S([n])}$ of the omitted leaves to feasible image vertices as given by the leaf matching graphs

pick an arbitrary order $S([n]) = \{y_1, \dots, y_{\binom{n}{r-1}}\}$;
for $i = 1$ **to** $\binom{n}{r-1}$ **do**
  let $y = y_i$;
  let σ_y be a uniform random perfect matching in $F_y^{(i-1)}$;
  for $j = i + 1$ **to** $\binom{n}{r-1}$ **do**
    let $B_{y_j} := \{xu \in E(F_{y_j}^{(i-1)}) \colon \exists s \text{ such that } x \in V(G'_s) \text{ and } \sigma_y^{-1}(u) \in V(G_s)\}$;
    let $F_{y_j}^{(i)} := F_{y_j}^{(i-1)} - B_{y_j}$;
  end
end
return $(\sigma_y)_{y\in S([n])}$;

4.8.1 The proof of Theorem 4.8.1

We are now in a position to prove (assuming Lemmas 4.8.5 and 4.8.6) Theorem 4.8.1. As mentioned above, what this boils down to is, given the leftover H from Lemma 4.8.5, showing that we can choose $\vec{H}$ to inherit the required quasirandomness properties (which imply that the leaf matching graphs F_y satisfy the degree-codegree condition of Lemma 4.2.6) and arguing that during the running of *MatchLeaves*, it is unlikely that more than $\varrho pn/r$ edges are lost at any vertex of any F_y, which (by Lemma 4.2.6) implies that *MatchLeaves* is likely to complete and therefore the desired packing exists. For the choice of $\vec{H}$, we use Chernoff's inequality and union bounds together with Lemma 4.8.6. For the analysis of *MatchLeaves*, we use Corollary 4.2.2 together with union bounds.

Proof of Theorem 4.8.1. Given r, D and $\mu > 0$ as in the statement of Theorem 4.8.1, we choose constants as described at the beginning of this subsection. We set $Q = D(r+1) + 3$.

Suppose that $n \geq n_0$. Let $\widehat{H}$ be a (ξ, Q)-quasirandom hypergraph with n vertices, and let $(G_s)_{s\in[s^*]}$ be a sequence of D-degenerate hypergraphs, each on $n - \lfloor \mu n \rfloor$ vertices, with maximum degree at most $\frac{cn}{\log n}$, and containing at least μn leaves. Suppose further that $\sum_{s\in[s^*]} e(G_s) = e(\widehat{H})$.

For each $s \in [s^*]$, we choose a strongly independent set V_s of $\lfloor \nu n \rfloor$ leaves of G_s, and let G'_s be the corresponding hypergraph with vertices $(V(G_s) \setminus V_s) \cup I_s$, where I_s is a set of $|V_s|$ isolated vertices. We define weights $w_s(y)$ as in Definition 4.8.4.

Let for each s the map φ'_s be the embedding of G'_s into $\widehat{H}$ provided by Lemma 4.8.5, with H the $(\gamma^3, 2)$-quasirandom hypergraph of leftover edges, and suppose that the properties (*P* 1)–(*P* 5) hold, with the weight function w as in Definition 4.8.4. By construction, H has $\lfloor \mu n^{r-1} \rfloor \lfloor \nu n \rfloor = p\binom{n}{r}$ edges, so has density p.

The following claim establishes the existence of the desired orientation of H.

Claim 4.8.8. *There exists an orientation $\vec{H}$ of H such that for each semi-edge $y \in S(H)$ we have $w(y) = \deg^+_{\vec{H}}(y)$, and in addition for each $s, s' \in [s^*]$ we have*

(P'2) $\left|N^+_{\vec{H}}(y) \setminus \operatorname{im} \varphi'_s\right| = (1 \pm \gamma)\frac{\mu p n}{r}$, *and*

(P'3) $\left|N^+_{\vec{H}}(y) \setminus (\operatorname{im} \varphi'_s \cup \operatorname{im} \varphi'_{s'})\right| = (1 \pm \gamma)\frac{\mu^2 p n}{r}$ *if $s \neq s'$.*

Proof. H is $(\gamma^3, 2)$-quasirandom and of density p, and by (*P* 1) we have $w(y) = (1 \pm \gamma^3)\frac{pn}{r}$ for all $y \in S(H)$. This, together with our choice of γ, verifies that H satisfies the conditions of the orientation lemma (Lemma 4.8.6).

Let $\vec{H}_0$ be a uniform random orientation of H. Given $y \in S(H)$ and $s \in [s^*]$, by (*P* 2) and Theorem 4.2.1, with probability at least $1 - \exp\big(-\frac{\gamma^6 \mu p n}{12}\big)$ we have

$$\left|N^+_{\vec{H}_0}(y) \setminus \operatorname{im} \varphi'_s\right| = (1 \pm 3\gamma^3)\frac{\mu p n}{r} .$$

Similarly, given $y \in S(H)$ and $s, s' \in [s^*]$ with $s \neq s'$, by (*P* 3) and Theorem 4.2.1, with probability at least $1 - \exp\big(-\frac{\gamma^6 \mu^2 p n}{12}\big)$ we have

$$\left|N^+_{\vec{H}_0}(y) \setminus (\operatorname{im} \varphi'_s \cup \operatorname{im} \varphi'_{s'})\right| = (1 \pm 3\gamma^3)\frac{\mu^2 p n}{r} .$$

Taking the union bound, and by Lemma 4.8.6, with probability at least $1 - 2n^{3r+6}\exp\big(-\frac{\gamma^6 \mu^2 p n}{12}\big) - o(1)$ each of the above good events holds for each $y \in S(H)$ and each $s, s' \in [s^*]$ with $s \neq s'$, and in addition there is an orientation $\vec{H}$ of H satisfying conclusions (*O* 1) and (*O* 2) of Lemma 4.8.6.

For sufficiently large n we have $1 - 2n^{3r+6}\exp\big(-\frac{\gamma^6 \mu^2 p n}{12}\big) - o(1) > 0$, so we fix $\vec{H}_0$ and $\vec{H}$ satisfying all these properties. By (*O* 1) the orientation $\vec{H}$ satisfies $\deg^+_{\vec{H}}(y) = w(y)$ for each $y \in S(H)$, as desired. Given $y \in S(H)$ and $s, s' \in [s^*]$

with $s \neq s'$, by $(O2)$ we have

$$\left|N^+_{\vec{H}}(y) \setminus \operatorname{im}\varphi'_s\right| = \left|N^+_{\vec{H}_0}(y) \setminus \operatorname{im}\varphi'_s\right| \pm \gamma^2 n = (1 \pm 3\gamma^3)\tfrac{\mu pn}{r} \pm \gamma^2 n = (1 \pm \gamma)\tfrac{\mu pn}{r}\,,$$

where the final inequality is by choice of γ. This verifies $(P'2)$. Similarly, we have

$$\begin{aligned}\left|N^+_{\vec{H}}(y) \setminus (\operatorname{im}\varphi'_s \cup \operatorname{im}\varphi'_{s'})\right| &= \left|N^+_{\vec{H}_0}(y) \setminus (\operatorname{im}\varphi'_s \cup \operatorname{im}\varphi'_{s'})\right| \pm \gamma^2 n \\ &= (1 \pm 3\gamma^3)\tfrac{\mu^2 pn}{r} \pm \gamma^2 n = (1 \pm \gamma)\tfrac{\mu^2 pn}{r}\,,\end{aligned}$$

giving $(P'3)$. □

From now on, we assume $\vec{H}$ satisfies the properties of the above claim, which we use instead of $(P2)$ and $(P3)$. What remains is to analyse *MatchLeaves*, which means showing that the leaf matching graphs $F_y^{(t-1)}$ satisfy the conditions of Lemma 4.2.6 at every step t. We first show that the leaf matching graphs $F_y^{(0)}$ satisfy the conditions $(M1)$ and $(M2)$ required to play the rôle of F in Lemma 4.2.6, with $m := pn/r$. The main work is then to argue that sufficiently few edges are deleted from $F_y^{(0)}$ to form $F_y^{(t-1)}$ that it can play the rôle of F', i.e. that $(M3)$ is satisfied. For this last part, we actually need only to show specifically that the condition holds for $F_{y_t}^{(t-1)}$, where y_t is the tth semi-edge in our arbitrary order on $S(H)$.

Property $(M1)$: Given any $y \in S(H)$ and any $x \in V\big(F_y^{(0)}\big)$, we separate two cases. If $x \in L_y$ is in the hypergraph G_s, then by $(P'2)$ we have $\deg_{F_y^{(0)}}(x) = \left|\mathrm{N}^+_{\vec{H}}(y) \setminus \operatorname{im}\varphi'_s\right| = (1 \pm \gamma)\frac{\mu pn}{r}$. If $x \in N^+_{\vec{H}}(y)$, then by $(P4)$ we have $\deg_{F_y^{(0)}}(x) = \sum_s w_s(y)\mathbb{1}_{x \notin \operatorname{im}\varphi'_s} = (1 \pm \gamma^3)\frac{\mu pn}{r}$. In either case, since $p > \gamma$ this verifies $(M1)$.

Property $(M2)$: Given any $y \in S(H)$ and any $u, u' \in L_y$, if $u \in V(G_s)$ and $u' \in V(G_{s'})$, where $s \neq s'$, then by $(P'3)$ we have $\deg_{F_y^{(0)}}(u, u') = \left|\mathrm{N}^+_{\vec{H}}(y) \setminus (\operatorname{im}\varphi'_s \cup \operatorname{im}\varphi'_{s'})\right| = (1 \pm \gamma)\frac{\mu^2 pn}{r}$. Again since $\gamma < p$ this is as required by $(M2)$, and we only need to show that the number of $u, u' \in L_y$ which are both in G_s for some $s \in [s^*]$ is at most $\frac{p^2n^2}{4\log(pn/2)}$. But any given G_s has at most $w_s(y) \leq \frac{cn}{\log n}$ vertices in L_y, so that for a given u there are at most $\frac{cn}{\log n}$ choices of u' with $u, u' \in V(G_s)$ for some $s \in [s^*]$. Since $|L_y| \leq n$ we conclude that there are at most $\frac{cn^2}{\log n} < \frac{p^2n^2}{4\log(pn/2)}$ pairs $u, u' \in L_y$ such that $u, u' \in V(G_s)$ for some $s \in [s^*]$. This completes the verification of $(M2)$.

Property $(M3)$: This property does not hold deterministically, but we shall show that it holds for all t with high probability. For this purpose we define the following events. For each $t \in \binom{n}{r-1}$, let $\mathcal{E}_t$ be the event that for each $x \in V(F_{y_t}^{(0)})$ we have

$$\deg_{F_{y_t}^{(0)}}(x) - \deg_{F_{y_t}^{(t-1)}}(x) \le \tfrac{1}{2r}\varrho pn\,, \tag{4.8.3}$$

that is, $\mathcal{E}_t$ implies the event that $(M3)$ holds for $F = F_t^{(0)}$ and $F' = F_t^{(t-1)}$. We shall prove the following claim below, but first show how it implies the theorem.

Claim 4.8.9. *With probability at least* $1 - n^{-1}$*, for every* $t \in \left[\binom{n}{r-1}\right]$ *the event* $\mathcal{E}_t$ *holds.*

If $\mathcal{E}_t$ holds then all conditions of Lemma 4.2.6 are satisfied for $F = F_t^{(0)}$, $F' = F_t^{(t-1)}$. In particular, the perfect matching in $F_t^{(t-1)}$ required by *MatchLeaves* exists, and so *MatchLeaves* does not fail at time t. Thus Claim 4.8.9 states that with high probability, *MatchLeaves* does not fail at any time, and thus the maps φ_s defined in (4.8.2) form (as explained there) the required perfect packing, proving Theorem 4.8.1.

To prove Claim 4.8.9, let for each $t \in \binom{n}{r-1}$ the history $\mathscr{H}_{t-1}$ consist of the collection of matchings $\sigma_1, \ldots, \sigma_{t-1}$ obtained by running *MatchLeaves* up to and including step $t-1$. Note that giving $\mathscr{H}_{t-1}$ determines whether $\mathcal{E}_t$ holds or not.

Claim 4.8.10. *For each* $t \in \left[\binom{n}{r-1}\right]$*, each* $u \in N^+_{\vec{H}}(y_t)$ *and each* $s \in [s^*]$ *the following holds. Either* $\mathcal{E}_t$ *does not occur or a random perfect matching* σ_t *in* $F_t^{(t-1)}$ *satisfies*

$$\mathbb{P}\left[\sigma_t^{-1}(u) \in V(G_s) | \mathscr{H}_{t-1}\right] \le \frac{2rw_s(y_t)}{\mu pn}\,.$$

Proof. If $\mathcal{E}_t$ occurs, then all properties $(M1)$–$(M3)$ from Lemma 4.2.6 are satisfied with $F = F_t^{(0)}$ and $F' = F_t^{(t-1)}$, and thus, a random matching σ_t in F' satisfies for any given edge $xu \in E\big(F_t^{(t-1)}\big)$

$$\mathbb{P}\left[xu \in \sigma_t \middle| \mathscr{H}_{t-1}\right] \le \frac{2}{\mu m} = \frac{2r}{\mu pn}\,.$$

Taking the union bound over the $w_s(y_t)$ choices of $x \in L_{y_t}$ in G_s, the claim follows. □

We now verify Claim 4.8.9. We shall first argue that the claimed probability bound follows from a probability bound, given in (4.8.4), which is of the right form to use Corollary 4.2.2. Indeed, let $\mathcal{A}_t$ be the event that $\mathcal{E}_i$ holds for each $1 \le i < t$ but $\mathcal{E}_t$ does not hold. Observe that if for each t the event $\mathcal{A}_t$ does not hold, then

$\mathcal{E}_t$ holds for each $t \in \left[\binom{n}{r-1}\right]$. In particular, by the union bound over $t \in \left[\binom{n}{r-1}\right]$ it suffices to show that for each fixed $t \in \left[\binom{n}{r-1}\right]$ we have $\mathbb{P}[\mathcal{A}_t] \le n^{-r-1}$. Further, by another union bound over the at most $v(F^{(0)}_{y_t}) = 2w(y_t) \le 2n$ different $x \in V(F^{(0)}_{y_t})$ and since $\mathcal{A}_t \subseteq \bigcap_{1\le i\le t-1} \mathcal{E}_i$ it is enough to show that for a fixed t and $x \in V(F^{(0)}_{y_t})$

$$\mathbb{P}\Big[\bigcap_{1\le i\le t-1} \mathcal{E}_i \quad\text{and}\quad \deg_{F^{(0)}_{y_t}}(x) - \deg_{F^{(t-1)}_{y_t}}(x) > \tfrac{1}{2r}\varrho pn\Big] \le n^{-r-3}\,. \tag{4.8.4}$$

The remainder of this proof is devoted to establishing this bound. We will use Corollary 4.2.2 for this purpose, with the good event $\bigcap_{1\le i\le t-1} \mathcal{E}_i$. To that end, define for each $1 \le i \le t-1$ the random variable

$$Y_i := \deg_{F^{(i-1)}_{y_t}}(x) - \deg_{F^{(i)}_{y_t}}(x)$$

and observe that

$$\deg_{F^{(0)}_{y_t}}(x) - \deg_{F^{(t-1)}_{y_t}}(x) = \sum_{i=1}^{t-1} Y_i\,.$$

To apply Corollary 4.2.2 we need to find the range of each Y_i and the expectation of each Y_i, conditioned on the history $\mathscr{H}_{i-1}$. This is encapsulated in Claim 4.8.11.

Claim 4.8.11. *For each $1 \le i \le t-1$, we have $0 \le Y_i \le \frac{cn}{\log n}$. Furthermore, either some $\mathcal{E}_i$ with $1 \le i \le t-1$ does not occur, or we have $\sum_{i=1}^{t-1} \mathbb{E}[Y_i|\mathscr{H}_{i-1}] \le 40p^2r!r\mu^{-2}n$.*

Proof. We first show $0 \le Y_i \le \frac{cn}{\log n}$. There are two cases to consider. First, if $x \in L_{y_t}$, then x is in G_s for some $s \in [s^*]$. An edge xu of $F^{(i-1)}_{y_r}$ is removed to form $F^{(i)}_{y_r}$ only if u is assigned a leaf of G_s in σ_i. Since there are at most $w_s(y_i) \le \frac{cn}{\log n}$ such leaves, we have $Y_i \le \frac{cn}{\log n}$ in this case. Second, if $x \in N^+_{\vec{H}}(y_t)$, and x is assigned a leaf of G_s in σ_i, then we remove all edges of $F^{(i-1)}_t$ from x to leaves of G_s to form $F^{(i)}_t$. Since σ_i is a matching, this happens for at most one $s \in [s^*]$. There are at most $w_s(y_t) \le \frac{cn}{\log n}$ such leaves of G_s, so also in this case we have $Y_i \le \frac{cn}{\log n}$.

We now bound above the sum of conditional expectations. Again, there are two cases to consider. First, if $x \in L_{y_t}$, then let s be such that $x \in V(G_s)$. Suppose that $\mathscr{H}_{i-1}$ is a history up to and including σ_{i-1} such that $\mathcal{E}_i$ holds. Recall that $\sigma_i^{-1}(u) \in V(G_s)$ means that some leaf of G_s is matched to u in σ_i, and that we write $\{uy_i, u\} \in E(\vec{H})$ to mean that the edge uy_i, directed towards u, is in $\vec{H}$. By linearity of expectation, we have

$$\begin{aligned}\mathbb{E}\big[Y_i\big|\mathscr{H}_{i-1}\big] &= \sum_{u\in \mathsf{N}_{F_{y_t}^{(i-1)}}(x)\,:\,\{uy_i,u\}\in E(\vec{H})} \mathbb{P}\big[\sigma_i^{-1}(u)\in V(G_s)\big|\mathscr{H}_{i-1}\big]\\ &\le \sum_{u\in \mathsf{N}_{F_{y_t}^{(i-1)}}(x)\,:\,\{uy_i,u\}\in E(\vec{H})} w_s(y_i)\tfrac{2r}{\mu pn} \le 2p^2n\cdot w_s(y_i)\cdot\tfrac{2r}{\mu pn}\,,\end{aligned}$$

where the first inequality is by Claim 4.8.10 and the second holds since if $u\in \mathsf{N}_{F_{y_t}^{(i-1)}}(x)$ and $\{uy_i,u\}\in E(\vec{H})$ then in particular $u\in \mathsf{N}_H(y_i,y_t)$, and since $\big|\mathsf{N}_H(y_i,y_t)\big|\le 2p^2n$ by $(\gamma^3,2)$-quasirandomness of H. Summing over i, either some $\mathcal{E}_i$ with $i\in[t-1]$ does not hold, or we have

$$\sum_{i=1}^{t-1}\mathbb{E}\big[Y_i\big|\mathscr{H}_{i-1}\big] \le \sum_{i=1}^{t-1} 2p^2n\cdot w_s(y_i)\cdot\tfrac{2r}{\mu pn} = \tfrac{4rp}{\mu}\sum_{i=1}^{t-1} w_s(y_i) \le \tfrac{4rp}{\mu}\sum_{i=1}^{\binom{n}{r-1}} w_s(y_i)\,.$$

Since $\sum_{i=1}^{\binom{n}{r-1}} w_s(i) = \lfloor \nu n\rfloor$ counts the number of leaves removed from G_s to form G'_s, we obtain that either some $\mathcal{E}_i$ with $i\in[t-1]$ does not hold, or

$$\sum_{i=1}^{t-1}\mathbb{E}\big[Y_i\big|\mathscr{H}_{i-1}\big] \le \tfrac{40p^2}{\mu^2}n\,,$$

as desired.

Finally, we consider the case $x\in N^+_{\vec{H}}(y_t)$. Suppose that a leaf of G_s is assigned to x by σ_i, and that x is adjacent to $w_s(y_t)$ leaves of G_s in $F_{y_t}^{(i-1)}$. Then the edges to these leaves are exactly the edges at x removed from $F_{y_t}^{(i-1)}$ to form $F_{y_t}^{(i)}$. Suppose that $\mathscr{H}_{i-1}$ is a history up to and including σ_{i-1} such that $\mathcal{E}_i$ holds. Recall that a leaf of G_s can only be assigned to x by σ_i if $\{xy_i,x\}\in E(\vec{H})$ and $x\notin \operatorname{im}\varphi'_s$. By linearity of expectation, we have

$$\begin{aligned}\mathbb{E}\big[Y_i\big|\mathscr{H}_{i-1}\big] &= \sum_{s\in[s^*]} \mathbb{1}_{x\in N^+_{\vec{H}}(y_i)}\mathbb{P}\big[\sigma_i^{-1}(x)\in V(G_s)\big|\mathscr{H}_{i-1}\big]\cdot w_s(y_t)\\ &\le \sum_{s\in[s^*]} \mathbb{1}_{x\in N^+_{\vec{H}}(y_i)}\mathbb{1}_{x\notin\operatorname{im}\varphi'_s} w_s(y_t)w_s(y_i)\tfrac{2r}{\mu pn}\,,\end{aligned}$$

where the second line follows by Claim 4.8.10. Summing over i, either some $\mathcal{E}_i$

with $i \in [r-1]$ does not hold, or we have

$$\begin{aligned}
\sum_{i=1}^{t-1} \mathbb{E}\big[Y_i \big| \mathscr{H}_{i-1}\big] &\le \sum_{i=1}^{t-1} \sum_{s\in[s^*]} \mathbb{1}_{x\in N^+_{\vec{H}}(y_i)} \mathbb{1}_{x\notin \operatorname{im}\varphi'_s} w_s(y_t) w_s(y_i) \tfrac{2r}{\mu p n} \\
&\le \sum_{s\in[s^*]} \sum_{i=1}^{\binom{n}{r-1}} \mathbb{1}_{x\in N^+_{\vec{H}}(y_i)} \mathbb{1}_{x\notin \operatorname{im}\varphi'_s} w_s(y_t) w_s(y_i) \tfrac{2r}{\mu p n} \\
&= \sum_{s\in[s^*]} \sum_{i:\{xy_i,x\}\in E(\vec{H})} \mathbb{1}_{x\notin \operatorname{im}\varphi'_s} w_s(y_t) w_s(y_i) \tfrac{2r}{\mu p n} \\
&\le \sum_{s\in[s^*]} \tfrac{2r w_s(y_t)}{\mu p n} \mathbb{1}_{x\notin \operatorname{im}\varphi'_s} \sum_{i:xy_i\in E(H)} w_s(y_i) \\
&\le \sum_{s\in[s^*]} \tfrac{2r w_s(y_t)}{\mu p n} \cdot \tfrac{10 r! p^2 n}{\mu} = \tfrac{20 p r! r}{\mu^2} \sum_{s\in[s^*]} w_s(y_t)\,,
\end{aligned}$$

where the last inequality is by (P 5). By definition of $w(y_t)$, by (P 1) and by choice of γ we have $\sum_{s\in[s^*]} w_s(y_t) = w(y_t) \le \frac{2}{r}pn$, so we conclude that either some $\mathcal{E}_i$ with $i \in [t-1]$ does not hold, or we have

$$\sum_{i=1}^{t-1} \mathbb{E}\big[Y_i \big| \mathscr{H}_{i-1}\big] \le \frac{20 p r! r}{\mu^2} \cdot \frac{2}{r} pn\,,$$

as desired. ▫

Recall that $p \le 2\mu\nu/r!$ by (4.8.1). Thus $40 p r! r \mu^{-2} \le 160\nu\mu^{-1} \le \frac{1}{4r}\varrho$, where the final inequality is by choice of ν. It follows that $40 p^2 r! r \mu^{-2} \le \frac{1}{4r}\varrho p$. Thus using Claim 4.8.11 and Corollary 4.2.2, with $R = \frac{cn}{\log n}$, with $\tilde{\mu} = \frac{1}{4r}\varrho p n$, and with the event $\mathcal{E} = \bigcap_{i=1}^{r-1} \mathcal{E}_i$, we get

$$\mathbb{P}\Big[\bigcap_{1\le i\le r-1} \mathcal{E}_i \quad \text{and} \quad \sum_{i=1}^{r-1} Y_i > \tfrac{1}{2r}\varrho p n\Big] \le \exp\big(-\tfrac{\tilde{\mu}}{4R}\big) \le n^{-r-3}\,,$$

where the final inequality is by choice of c. This establishes (4.8.4), so completes the proof. □

4.8.2 Proof of the orientation lemma

In this subsection we prove Lemma 4.8.6. For this we need the following two definitions. The *directed neighbourhood* of a semi-edge $y \in S(\vec{H})$ to vertex $v \in y$

in a directed hypergraph $\vec{H}$ is

$$\mathsf{N}^{v}_{\vec{H}}(y) = \left\{u \colon \{yu, v\} \in E(\vec{H})\right\},$$

i.e. all vertices that form an edge with y, which is directed towards v. Also, if $y_1, \ldots, y_k \in V(H)$ is a sequence of vertices, and $1 \le i \le j \le k$ are integers, then we write $y_{i..j}$ for the set $\{y_i, \ldots, y_j\}$.

The idea of the proof is the following. We first argue that a uniform random orientation $\vec{H}_0$ of H is likely to have $\mathsf{N}_{\vec{H}_0}(y)$ very close to $w(y)$ for every semi-edge y, and in addition it is likely to inherit the quasirandomness of H. Assuming both these likely events hold, we then iteratively modify $\vec{H}_0$ as follows. If there is no $y \in S(H)$ whose out-degree is currently larger than $w(y)$, we are done. If there is such a y, there exists also $y' \in S(H)$ whose out-degree is smaller than $w(y')$. We identify (using the quasirandomness) a short chain of edges whose orientations we can change in order that overall the out-degree of y is decreased by one, that of y' is increased by one, and all other semi-edges remain unchanged. We repeat this until we obtain the desired $\vec{H}$. It remains to argue that we do not alter the orientation of too many edges at any given semi-edge: to that end, we choose the chain of edges at each step randomly and argue that it is unlikely any given semi-edge appears too often in the chosen chains.

Proof of Lemma 4.8.6. By the given quasirandomness of H we know that $\deg_H(y) = (1 \pm \gamma^3)pn$ and $|\mathsf{N}_H(y) \cap \mathsf{N}_H(z)| = (1 \pm \gamma^3)p^2 n$ for every $y \neq z \in S(H)$. Using Theorem 4.2.1, we see that a.a.s. for every $y \neq z \in S(H)$ and $v \in z$ we have

$$\deg^{+}_{\vec{H}_0}(y) = (1 \pm 2\gamma^3)\frac{pn}{r} \text{ and } \left|\mathsf{N}^{+}_{\vec{H}_0}(y) \cap \mathsf{N}^{v}_{\vec{H}_0}(z)\right| = (1 \pm 2\gamma^3)\frac{p^2 n}{r^2}.$$

From now on, we fix an arbitrary orientation $\vec{H}_0$ satisfying these two properties, and we will show that for any such $\vec{H}_0$ there is an orientation $\vec{H}$ of H satisfying $(O\,1)$ and $(O\,2)$.

Starting with $\vec{H}_0$, we successively switch the orientations of sets of $2r-2$ edges, thus producing a sequence of oriented hypergraphs $(\vec{H}_i)_{0 \le i \le t}$, with $\vec{H}_t$ being the desired $\vec{H}$. For any such oriented hypergraph $\vec{H}_i$ and every semi-edge $y \in S(H)$ we define the potential $\varphi_i(y) := \deg^{+}_{\vec{H}_i}(y) - w(y)$ and

$$\varphi(\vec{H}_i) := \sum_{y \in S(H)} |\varphi_i(y)|.$$

Observe that $\varphi(\vec{H}_i)$ is always nonnegative and even, and it is equal to zero exactly when (O 1) is satisfied. To begin with, we have $|\varphi_0(y)| \le 3\gamma^3 pn$ for every semi-edge y, so that $\varphi(\vec{H}_0) \le 3\gamma^3 pn\binom{n}{r-1}$.

The algorithm *PathSwitch*, with input two semi-edges y and z, finds a way of changing the orientations of some $2r-2$ edges such that the out-degree of y is increased by 1 and that of z is decreased by 1, with all other semi-edges keeping their out-degree the same.

Algorithm 5: *PathSwitch*

```
Input : A directed hypergraph H⃗′ and two disjoint semi-edges y, z ∈ S(H⃗′).
let y_1, ..., y_{r−1} be a uniform random order of the vertices in y ;
let z_1, ..., z_{r−1} be a uniform random order of the vertices in z ;
let H⃗′_0 := H⃗′ ;
for i = 1 to r − 1 do
    let a := y_{i..i+r−2} ;
    let b := z_{i..i+r−2} ;
    let X := N^{y_i}_{H⃗′_{i−1}}(a) ∩ N^+_{H⃗′_{i−1}}(b) \ (y_{1..i+r−2} ∪ z_{1..i+r−2}) ;
    halt with failure if X = ∅ ;
    let v be a uniform random vertex from X ;
    let y_{i+r−1} := v ;
    let z_{i+r−1} := v ;
    let H⃗′_i be H⃗′_{i−1} − {av, y_i} − {bv, v} + {av, v} + {bv, z_i} ;
end
return H⃗′_{r−1} ;
```

Observe that *PathSwitch* either halts with failure, or performs as described: that is, in the returned hypergraph the out-degree of y is increased by one, that of z is decreased by one, and all other semi-edges have the same out-degree as before, with the orientation of in total $2r-2$ edges changed. To see this, observe first that the definition of X ensures that all the edges whose orientations we choose are distinct, so that we indeed change the orientation of $2r-2$ edges. Consider now the set of vertices $y_{1..2r-2}$. In step i, we change the orientation of the edge $y_{i..i+r-1}$ from being directed towards y_i, to being directed towards y_{i+r-1}. Thus this change increases the out-degree of $y_{i..i+r-2}$, and decreases that of $y_{i+1..i+r-1}$, by one, and affects no other semi-edge. Overall, the result of these changes among the vertices $y_{1..2r-2}$ is to increase the out-degree of y in $\vec{H}'_{r-1}$, and decrease that of $y_{r..2r-2}$, by one. Similarly, the changes among the vertices $z_{1..2r-2}$ have the effect of decreasing the out-degree of z in $\vec{H}'_{r-1}$, and increasing that of $z_{r..2r-2}$, by one.

But $y_{r..2r-2} = z_{r..2r-2}$ by construction, so that this semi-edge overall has unchanged out-degree in $\vec{H}'_{r-1}$ as required.

Next, using the algorithm *PathSwitch* as a subroutine, the algorithm *OrientationSwitch* describes how orientations are switched. In every iteration of this algorithm, we change the potential of two semi-edges $y, z \in S(H)$ with $\varphi_i(y) < 0$ and $\varphi_i(z) > 0$ by applying *PathSwitch*. In case that

$$\left|\{us \in E(H) \colon us \text{ is oriented differently in } \vec{H}_i \text{ and } \vec{H}_0\}\right|$$

gets too large in some round i and for some semi-edge s, we let the algorithm halt with failure. However, we will see in the following that this happens with probability tending to 0.

Algorithm 6: *OrientationSwitch*

```
let t := φ(H⃗_0)/2 ;
for i = 0 to t − 1 do
    if ∃ s ∈ S(H⃗_0) with
      |{us ∈ E(H): us is oriented differently in H⃗_i and H⃗_0}| > γ²n
        then halt with failure;
    choose semi-edges y, z ∈ S(H⃗_i) with φ_i(y) < 0 and φ_i(z) > 0 ;
    let H⃗_{i+1} be the result of PathSwitch on H⃗_i, y, z ;
end
return H⃗ := H⃗_t ;
```

Observe that by our assumption on γ, provided n is sufficiently large *PathSwitch* does not halt with failure when called by *OrientationSwitch*, because the relevant directed (common) neighbourhoods can only be of size less than $4r$ if the failure condition of *OrientationSwitch* is satisfied for some semi-edge. Furthermore, by construction if *OrientationSwitch* does not halt with failure, then at each step i we have $\varphi(\vec{H}_{i+1}) = \varphi(\vec{H}_i) - 2$, so the returned $\vec{H}$ has $\varphi(\vec{H}) = 0$ and thus satisfies (*O* 1). Finally, by definition, not failing implies that $\vec{H}$ also satisfies (*O* 2). Thus what remains to prove is that *OrientationSwitch* succeeds with positive probability.

Note that when *PathSwitch* is run, it changes the orientation of an edge contained in a semi-edge x only if x is contained in either $y_{1..2r-2}$ or $z_{1..2r-2}$ (as constructed by *PathSwitch*), and the number of edges containing x whose orientation changes is at most four (for the semi-edge $y_{r..2r-2} = z_{r..2r-2}$). Thus in order to prove that *OrientationSwitch* succeeds with positive probability, it is enough to upper bound the number of times any given $x \in S(\vec{H}_0)$ is a subset of $y_{1..2r-2}$ or $z_{1..2r-2}$. This is given by the following claim.

Claim 4.8.12. *Given $x \in S(\vec{H}_0)$, with probability at least $1 - 2r\exp(-\sqrt{n})$, the number of times x is a subset of $y_{1..2r-2}$ or $z_{1..2r-2}$ is at most $2^{2r+3}p^{-2r}r^{3r}\gamma^3 n$.*

Proof. We fix $x \in S(\vec{H}_0)$. To begin with, we upper bound the probability that x is a subset of $y_{1..2r-2}$ for call i to *PathSwitch* with semi-edges y and z, working in the graph $\vec{H}_{i-1}$. This probability depends on $|x \cap y|$, but not on z or on the graph $\vec{H}_{i-1}$. If *OrientationSwitch* fails, the probability is zero, so we will from now assume that *OrientationSwitch* has not failed.

When we run *PathSwitch*, we choose $y_r, \dots, y_{2r-2}$ in order, at each step (we claim) choosing from a set of size at least $\frac{1}{2}p^2 n/r^2$. To see that the claimed set size is valid, observe that the set X of *PathSwitch* is obtained from a set of the form $\mathsf{N}^+_{\vec{H}_0}(y') \cap \mathsf{N}^v_{\vec{H}_0}(z')$ by removing vertices in edges at y' or z' whose orientation is changed in $\vec{H}_{i-1}$ from that in $\vec{H}_0$, and a further at most $3r-3$ vertices previously used in *PathSwitch*. Since *OrientationSwitch* has not failed, the number of vertices removed due to orientation changes is at most $2\gamma^2 n$, so we obtain the claimed set size by choice of γ.

For each $r \le i \le 2r-2$, when we choose y_i the probability of choosing a vertex of x is by the above claim at most $(r-1) \cdot \frac{2r^2}{p^2 n}$, and the probability that we end up choosing all of x is thus at most

$$\left((r-1) \cdot \frac{2r^2}{p^2 n}\right)^{(r-1)-|x \cap y|} \le \left(\frac{2r^3}{p^2 n}\right)^{(r-1)-|x \cap y|} .$$

We now use this bound to estimate how often x is a subset of $y_{1..2r-2}$ over all calls to *PathSwitch* by *OrientationSwitch*. For each $0 \le j \le r-1$, the number of choices of a semi-edge y such that $|x \cap y| = j$ is at most $\binom{r-1}{j} \cdot n^{(r-1)-j} \le 2^r n^{r-1-j}$. The number of times that any given y is input to *PathSwitch* by *OrientationSwitch* is at most $|\varphi_0(y)| \le 3\gamma^3 pn$. Thus the total number of calls to *PathSwitch* with some semi-edge y which intersects x in j vertices is at most $3\gamma^3 pn \cdot 2^r n^{r-1-j} = 3\gamma^3 2^r pn^{r-j}$. We see that the number of times that x is a subset of $y_{1..2r-2}$ from calls to *PathSwitch* with input y where $|x \cap y| = j$, is stochastically dominated by a binomial random variable with $3\gamma^3 2^r pn^{r-j}$ tries and success probability $\left(\frac{2r^3}{p^2 n}\right)^{(r-1)-j}$. By the Chernoff bound, Theorem 4.2.1, the probability that the number of successes exceeds

$$2\cdot 3\gamma^3 2^r pn^{r-j}\cdot\left(\tfrac{2r^3}{p^2 n}\right)^{(r-1)-j} = 3\cdot 2^{2r-j}\gamma^3 p^{-2r+3+2j} r^{3r-3-3j} n \leq 2^{2r+2} p^{-2r} r^{3r-1}\gamma^3 n\,.$$

is at most $\exp(-\sqrt{n})$. Taking the union bound over choices of j, we see that the total number of times that we have $x\subseteq y_{1..2r-2}$ is with probability at least $1-r\exp(-\sqrt{n})$ bounded above by $2^{2r+2}p^{-2r}r^{3r}\gamma^3 n$

We obtain the same estimate for the number of times that $x\subseteq z_{1..2r-2}$ by the same argument replacing y with z throughout, and the claim follows. ▫

We can now complete the proof that *OrientationSwitch* succeeds with positive probability. Indeed, if it fails there is some $x\in S(\vec{H}_0)$ which witnesses the failure. Since by choice of γ we have $2^{2r+3}p^{-2r}r^{3r}\gamma^3 n\leq\gamma^2 n$, by Claim 4.8.12 the probability that a given x is such a witness is at most $2r\exp(-\sqrt{n})$. Taking the union bound, the probability of a witness existing is at most $2r\binom{n}{r-1}\exp(-\sqrt{n})$, which tends to zero as $n\to\infty$, as desired. □

4.8.3 Proof of the almost perfect packing lemma

In this section we prove Lemma 4.8.5, which is an easy corollary of Theorem 4.7.1.

Proof of Lemma 4.8.5. Given D, r, μ, ν, set γ, we choose $\gamma>0$ such that $\gamma^3\leq\frac{1}{2}\mu\nu$. We let $\delta,c,\xi>0$ be returned by Theorem 4.7.1 with input r, $\frac{1}{2}\gamma^3$, D, $K=1$, and suppose n is sufficiently large for this theorem. Recall Theorem 4.7.1 provides that $\delta\leq\gamma$, and we without loss of generality also suppose $\xi\leq\frac{1}{100}\gamma^3$.

Given $\widehat{H}$ and $(G'_s)_{s\in[s^*]}$ as in Lemma 4.8.5, let $\hat{p}$ be the density of $\widehat{H}$, i.e. $\widehat{H}$ has $\hat{p}\binom{n}{r}$ edges. We have $p\binom{n}{r}=\hat{p}\binom{n}{r}-\sum_{s\in[s^*]}e(G'_s)=\lfloor\mu n^{r-1}\rfloor\lfloor\nu n\rfloor$ as in Definition 4.8.2.

For each $s\in[s^*]$ we create an $n-\delta n$-vertex hypergraph G''_s by adding $n-\delta n-v(G'_s)$ isolated vertices to G'_s, which we put at the end of the degeneracy order. We apply Theorem 4.7.1 to pack the hypergraphs $(G''_s)_{s\in[s^*]}$ into $\widehat{H}$, with input as above, with $X=X_1=\cdots=X_{r-1}=V(\widehat{H})$, and with ω_s the $(r-1)$-weight function obtained from w_s of Lemma 4.8.5 for each $s\in[s^*]$ as follows: we give each ordered $(r-1)$-set the weight w_s of the underlying unordered set. Note that the required diet conditions reduce to (ξ,Q)-quasirandomness of $\widehat{H}$. Suppose that the likely outcome of Theorem 4.7.1 occurs, with φ''_s being the packing of G''_s for each $s\in[s^*]$. Then the hypergraph H of leftover edges is $(\frac{1}{2}\gamma^3,Q)$-quasirandom. In addition, setting $\tilde{\delta}=\mu$, we have (*QUASI* 1)–(*QUASI* 3).

Observe that for each $s \in [s^*]$ we have $\mathrm{sum}(\omega_s) = (r-1)!\lfloor \nu n \rfloor$, because each omitted leaf is counted exactly $(r-1)!$ times, once for each ordering of its parent semi-edge. We set φ'_s to be the restriction of φ''_s to the first $n - \mu n$ vertices of G''_s, so that the φ'_s form a packing of the G'_s into $\widehat{H}$, also with leftover edges H. Note that $\mathrm{im}\,\varphi'_s = \varphi''_s([n-\mu n])$ for each s.

We now verify (P1). Given any semi-edge $y \in S(\widehat{H})$, we let $\mathbf{e}$ be an arbitrary ordering of y. By (QUASI2) with $U = \emptyset$, we have

$$\begin{aligned}\sum_{s=1}^{s^*} \omega_s(\mathbf{e}) &= (1 \pm \tfrac{1}{2}\gamma^3) n^{-(r-1)} \sum_{s=1}^{s^*} \mathrm{sum}(\omega_s) \\ &= (1 \pm \tfrac{1}{2}\gamma^3) n^{-(r-1)} \lfloor \mu n^{r-1} \rfloor (r-1)! \lfloor \nu n \rfloor \\ &= (1 \pm \tfrac{1}{2}\gamma^3) n^{-(r-1)} (r-1)! p \binom{n}{r} \\ &= (1 \pm \gamma^3) \tfrac{pn}{r} ,\end{aligned}$$

where the final equality uses the approximation $\binom{n}{r} = (1 \pm n^{-0.5}) \frac{n^r}{r!}$, valid for all sufficiently large n. This verifies (P1) since $w_s(y) = \omega_s(\mathbf{e})$.

For (P2) and (P3), which we need to verify for each semi-edge y and each $s \neq s' \in [s^*]$, we use (QUASI1), with $\ell = 1$ and $T_1 = y$, with respectively $S = \{s\}$ and $S = \{s, s'\}$. By (ξ, Q)-quasirandomness of $\widehat{H}$, we have $\left|\mathrm{N}_{\widehat{H}}(y)\right| = (1 \pm \xi)\hat{p}n$, and so

$$\left|\mathrm{N}_H(y) \setminus \mathrm{im}\,\varphi'_s\right| = (1 \pm \tfrac{1}{2}\gamma^3)\mu\tfrac{p}{\hat{p}} \cdot (1 \pm \xi)\hat{p}n = (1 \pm \gamma^3)\mu p n ,$$

as required for (P2). Essentially the same calculation gives (P3) and we omit the details.

For (P4), with a given $u \in V(\widehat{H})$ and semi-edge $y \in S(\widehat{H})$ that does not contain u, we apply (QUASI2), this time with $U = \{u\}$, and with an arbitrary ordering $\mathbf{e}$ of y. We obtain

$$\sum_{s=1}^{s^*} \omega_s(\mathbf{e}) \mathbb{1}_{U \cap \mathrm{im}\,\varphi'_s = \emptyset} = (1 \pm \tfrac{1}{2}\gamma^3)\mu n^{-(r-1)} \sum_{s=1}^{s^*} \mathrm{sum}(\omega_s) = (1 \pm \gamma^3)\mu \tfrac{pn}{r}$$

by the same calculation as for (P1), which as with (P1) gives (P4).

Finally we verify (*P* 5). Given $v \in V(\widehat{H})$, note that

$$\left|\mathsf{N}_{\widehat{H}}(v; X_1, \ldots, X_{r-1})\right| = (r-1)!\left|\mathsf{N}_{\widehat{H}}(v)\right| = (1 \pm \xi)(r-1)!\hat{p}n \cdot \binom{n-1}{r-2}\tfrac{1}{r-1},$$

where the final inequality is by (ξ, Q)-quasirandomness of $\widehat{H}$ and the observation that each semi-edge in $\mathsf{N}_{\widehat{H}}(v)$ contains $\binom{r-1}{r-2} = r-1$ sets of size $r-2$ which make a semi-edge with v. From (*QUASI* 3), we have that either $v \in \operatorname{im}\psi'_s$ or

$$\begin{aligned}\sum_{\mathbf{e}\in\mathsf{N}_H(v;X_1,\ldots,X_{r-1})} \omega_s(\mathbf{e}) &= (1 \pm \tfrac{1}{2}\gamma^3)\tfrac{p}{\hat{p}} \cdot (1 \pm \xi)(r-1)!\hat{p}n\binom{n-1}{r-2}\tfrac{1}{r-1} \cdot n^{1-r}\operatorname{sum}(\omega_s) \\ &= (1 \pm \gamma^3)p(r-1)!\tfrac{n^{r-1}}{(r-1)!}n^{1-r}(r-1)!\lfloor \nu n \rfloor \\ &= (1 \pm \gamma^3)p(r-1)! \cdot (1 \pm \tfrac{3}{4})\tfrac{pr!n}{\mu} \le 2(r-1)!\frac{r!p^2n}{\mu},\end{aligned}$$

where the final line uses (4.8.1). Observe that each semi-edge contributing to (*P* 5) contributes $(r-1)!$ times in the above sum (once for each ordering) and so this gives (*P* 5). □

Bibliography

[1] P. Allen, J. Böttcher, D. Clemens, J. Hladký, D. Piguet and A. Taraz, *The tree packing conjecture for trees of almost linear maximum degree*, arXiv:2106.11720.

[2] P. Allen, J. Böttcher, D. Clemens and A. Taraz, *Perfectly packing graphs with bounded degeneracy and many leaves*, arXiv:1906.11558.

[3] P. Allen, J. Böttcher, J. Hladký and D. Piguet, *Packing degenerate graphs*, Adv. Math. 354 (2019), 106739.

[4] N. Alon, J. Balogh, P. Keevash and B. Sudakov, *The number of edge colorings with no monochromatic cliques*, J. London Math. Soc. (2) 70.2 (2004), 273–288.

[5] J. Balogh and C. Palmer, *On the Tree Packing Conjecture*, SIAM J. Discrete Math. 27.4 (2013), 1995–2006.

[6] J. Balogh, N. Bushaw, M. Collares, H. Liu, R. Morris and M. Sharifzadeh, *The typical structure of graphs with no large cliques*, Combinatorica 37.4 (2017), 617–632.

[7] J. Balogh, R. Morris and W. Samotij, *Independent sets in hypergraphs*, J. Amer. Math. Soc. 28.3 (2015), 669–709.

[8] B. Bollobás, *Extremal graph theory*, vol. 11, London Mathematical Society Monographs, Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], London-New York, 1978, xx+488.

[9] J. A. Bondy, *Pancyclic graphs*, Proceedings of the Second Louisiana Conference on Combinatorics, Graph Theory and Computing (Louisiana State Univ., Baton Rouge, La., 1971), 1971, 167–172.

[10] J. A. Bondy, *Pancyclic graphs: recent results*, Infinite and finite sets (Colloq., Keszthely, 1973; dedicated to P. Erdős on his 60th birthday), Vol. I, 1975, Colloq. Math. Soc. János Bolyai, Vol. 10, 181–187.

[11] J. Böttcher, J. Hladký, D. Piguet and A. Taraz, *An approximate version of the tree packing conjecture*, Israel J. Math. 211.1 (2016), 391–446.

[12] V. Chvátal and P. Erdős, *A note on Hamiltonian circuits*, Discrete Math. 2 (1972), 111–113.

[13] D. Clemens, S. Das and T. Tran, *Colourings without monochromatic disjoint pairs*, European J. Combin. 70 (2018), 99–124.

[14] A. Dankovics, *Low independence number and Hamiltonicity implies pancyclicity*, J. Graph Theory 95.2 (2020), 181–191.

[15] G. A. Dirac, *Some theorems on abstract graphs*, Proc. London Math. Soc. (3) 2 (1952), 69–81.

[16] R. A. Duke, H. Lefmann and V. Rödl, *A fast approximation algorithm for computing the frequencies of subgraphs in a given graph*, SIAM J. Comput. 24.3 (1995), 598–620.

[17] P. Erdős, *Some new applications of probability methods to combinatorial analysis and graph theory*, Proceedings of the Fifth Southeastern Conference on Combinatorics, Graph Theory and Computing (1974), 39–51. Congressus Numerantium, No. X.

[18] P. Erdős, *Some of my favourite problems in various branches of combinatorics*, Matematiche (Catania) 47.2 (1992), Combinatorics 92 (Catania, 1992), 231–240 (1993).

[19] P. Erdős, *Some problems in graph theory* (1974), 187–190. Lecture Notes in Math., Vol. 411.

[20] P. Erdős, *Some recent progress on extremal problems in graph theory*, Proceedings of the Sixth Southeastern Conference on Combinatorics, Graph Theory and Computing (Florida Atlantic Univ., Boca Raton, Fla., 1975), 1975, 3–14. Congressus Numerantium, No. XIV.

[21] P. Erdős, *Some theorems on graphs*, Riveon Lematematika 9 (1955), 13–17.

[22] P. Erdős, D. J. Kleitman and B. L. Rothschild, *Asymptotic enumeration of K_n-free graphs*, Colloquio Internazionale sulle Teorie Combinatorie (Rome, 1973), Tomo II, 1976, 19–27. Atti dei Convegni Lincei, No. 17.

[23] P. Erdős, C. Ko and R. Rado, *Intersection theorems for systems of finite sets*, Quart. J. Math. Oxford Ser. (2) 12 (1961), 313–320.

[24] D. A. Freedman, *On tail probabilities for martingales*, Ann. Probability 3 (1975), 100–118.

[25] Z. Füredi, *A proof of the stability of extremal graphs, Simonovits' stability from Szemerédi's regularity*, J. Combin. Theory Ser. B 115 (2015), 66–71.

[26] S. Glock, D. Kühn, A. Lo and D. Osthus, *The existence of designs via iterative absorption: hypergraph F-designs for arbitrary F*, arXiv:1611.06827.

[27] S. Glock, F. Joos, J. Kim, D. Kühn and D. Osthus, *Resolution of the Oberwolfach problem*, J. Eur. Math. Soc. (JEMS) 23.8 (2021), 2511–2547.

[28] A. Gyárfás and J. Lehel, *Packing trees of different order into K_n*, Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976), vol. 18, Colloq. Math. Soc. János Bolyai, North-Holland, Amsterdam, 1978, 463–469.

[29] H. Hàn and A. Jiménez, *Improved Bound on the Maximum Number of Clique-Free Colorings with Two and Three Colors*, SIAM J. Discrete Math. 32.2 (2018), 1364–1368.

[30] S. Janson, T. Łuczak and A. Ruciński, *Random graphs*, Wiley-Interscience, 2000.

[31] F. Joos, J. Kim, D. Kühn and D. Osthus, *Optimal packings of bounded degree trees*, J. Eur. Math. Soc. (JEMS) 21.12 (2019), 3573–3647.

[32] P. Keevash, *The existence of designs*, arXiv:1401.3665.

[33] P. Keevash and K. Staden, *Ringel's tree packing conjecture in quasirandom graphs*, arXiv:2004.09947.

[34] P. Keevash and K. Staden, *The generalised Oberwolfach problem*, arXiv:2004.09937.

[35] P. Keevash and B. Sudakov, *Pancyclicity of Hamiltonian and highly connected graphs*, J. Combin. Theory Ser. B 100.5 (2010), 456–467.

[36] P. Keevash, *Hypergraph matchings and designs*, Proceedings of the International Congress of Mathematicians—Rio de Janeiro 2018. Vol. IV. Invited lectures, World Sci. Publ., Hackensack, NJ, 2018, 3113–3135.

[37] J. Kim, D. Kühn, D. Osthus and M. Tyomkyn, *A blow-up lemma for approximate decompositions*, Trans. Amer. Math. Soc. 371.7 (2019), 4655–4742.

[38] T. P. Kirkman, *On a problem in combinations*, Cambridge and Dublin Math. J. 2 (1847), 191–204.

[39] M. Krivelevich, *Triangle factors in random graphs*, Combin. Probab. Comput. 6.3 (1997), 337–347.

[40] C. Lee and B. Sudakov, *Hamiltonicity, independence number, and pancyclicity*, European J. Combin. 33.4 (2012), 449–457.

[41] H. Liu, O. Pikhurko and K. Staden, *The exact minimum number of triangles in graphs with given order and size*, Forum Math. Pi 8 (2020), e8, 144.

[42] W. Mantel, *Problem 28*, Wiskundige Opgaven 10.60-61 (1907), 320.

[43] R. Montgomery, A. Pokrovskiy and B. Sudakov, *A proof of Ringel's conjecture*, Geom. Funct. Anal. 31.3 (2021), 663–720.

[44] F. Mousset, R. Nenadov and A. Steger, *On the number of graphs without large cliques*, SIAM J. Discrete Math. 28.4 (2014), 1980–1986.

[45] O. Pikhurko and K. Staden, *Stability for the Erdős-Rothschild problem*, arXiv:2105.09991.

[46] O. Pikhurko, K. Staden and Z. B. Yilma, *The Erdős-Rothschild problem on edge-colourings with forbidden monochromatic cliques*, Math. Proc. Cambridge Philos. Soc. 163.2 (2017), 341–356.

[47] O. Pikhurko and Z. B. Yilma, *The maximum number of K_3-free and K_4-free edge 4-colorings*, J. Lond. Math. Soc. (2) 85.3 (2012), 593–615.

[48] J. Plücker, *System der analytischen Geometrie, auf neue Betrachtungsweisen gegründet, und insbesondere eine ausführliche Theorie der Kurven dritter Ordnung enthaltend*, Duncker und Humboldt, Berlin, 1835.

[49] G. Ringel, *Problem 25*, Theory of Graphs and its Applications (Proc. Int. Symp. Smolenice 1963), Prague: Czech. Acad. Sci., 1963.

[50] V. Rödl, *On a packing and covering problem*, European J. Combin. 6.1 (1985), 69–78.

[51] V. Rödl, A. Ruciński and E. Szemerédi, *Perfect matchings in uniform hypergraphs with large minimum degree*, European J. Combin. 27.8 (2006), 1333–1349.

[52] I. Z. Ruzsa and E. Szemerédi, *Triple systems with no six points carrying three triangles*, Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976), Vol. II, vol. 18, Colloq. Math. Soc. János Bolyai, North-Holland, Amsterdam-New York, 1978, 939–945.

[53] D. Saxton and A. Thomason, *Hypergraph containers*, Invent. Math. 201.3 (2015), 925–992.

[54] J. Steiner, *Combinatorische Aufgaben*, J. Reine Angew. Math. 45 (1853), 181–182.

[55] P. Turán, *Eine Extremalaufgabe aus der Graphentheorie*, Mat. Fiz. Lapok 48 (1941), 436–452.

[56] R. Yuster, *The number of edge colorings with no monochromatic triangle*, J. Graph Theory 21.4 (1996), 441–452.